

Teleimersja

Spis treści

Wstęp	2
I. Wprowadzenie: Grafika i wizualizacja.....	7
II. Teleimersja: Wirtualna obecność.....	16
III. Tele – imersja: Telekonferencje jutra	31
IV. SOFT (oprogramowanie dla teleimersji); BAMBOO(platforma i język)	39
V. Zastosowanie teleimersji.....	45
1. Biuro Przyszłości i Teleimersja	45
2. Biuro przyszłości (system projekcji ściennej)	70
3. Projekt “Being There”.....	88
4. Lampy Cieniowe.....	115
5. Automatycznie rekonfigurujące się rzutniki szeroko obrazowe.....	150
6. Teleimersja w medycynie	159
Zakończenie	212
Dodatki.....	230

Anna Szatanik

Wstęp

Teleimersja

Teleimersja to:

- nowe medium, synteza wideokonferencji i rzeczywistości wirtualnej, umożliwiające ludziom oddalonym o tysiące kilometrów kontakt tak naturalny, jakby byli w tym samym pokoju.
- sztandarowa aplikacja konsorcjum Internet2. Wymaga olbrzymich mocy obliczeniowych i superszybkich sieci, których rozwój będzie miał pozytywny wpływ na wiele dziedzin.
- w wielu przypadkach zastąpi podróże służbowe.

Teleimersja pozwala użytkownikom w oddalonych geograficznie od siebie miejscach do współpracy w tym samym czasie w symulowanym środowisku, tak jakby byli w tym samym rzeczywistym pomieszczeniu. Jest to najnowsze połączenie technologii medialnych:

- trójwymiarowe skanowanie obrazu
- technologie projekcyjne i prezentacyjne
- technologie śledzące ruch
- technologie audio
- robotyka

Znaczące wymagania dla systemu teleimersyjnego, jak szerokopasmowość, małe opóźnienie, sprawiają, że jest to jedno z najnowszych i wymagających zastosowań.

Koncepcje i nowe wyzwania

W teleimersyjnym środowisku komputery rozpoznają obecność i ruchy osób, fizycznych i wirtualnych przedmiotów, śledzą ruch osób i przedmiotów i pokazują je w realnych oddalonych geograficznie immersyjnych środowiskach na ekranach stereoimersyjnych. To wymaga pobierania próbek i resyntezy fizycznego środowiska zarówno twarzy użytkowników i ciała, co jest nowym wyzwaniem,

które wpłynę na pojawiające się technologie jak wydobywanie głębi obrazu i wyprostowanie zniekształceń.

Środowiska teleimersyjne przyspieszą i ułatwią nie tylko interakcję między użytkownikami, ale również między użytkownikami a modelami sterowanymi komputerowo i symulacjami. To będzie wymagało poszerzenia granic obrazu komputerowego, śledzenia ruchu, odpowiedniego monitora i technologii dostarczających obraz. Umożliwi użytkownikom uzyskanie interesującego doświadczenia i podstawy do włączenia w większym stopniu tej technologii do całego systemu. Ten nowy model współpracy między człowiekiem a komputerem dzieli się na kategorię zastosowania sieci o najwyższym stopniu zaawansowania i to jest najnowsze wyzwanie dla Internet2. Kolejnym celem tej interakcji jest przyspieszenie rozwoju lepszych narzędzi do współpracy między obszarami badań a edukacyjnym oraz do promowania postępów w badaniach z wykorzystaniem wirtualnej rzeczywistości, co oznacza, że ten plan będzie wymagał wielu rozwiązań następnej generacji Internetu.

Potencjalne sposoby zastosowania?

Teleimersyjne systemy mają potencjał do znacznego zmieniania edukacyjnych, naukowych i przemysłowych modeli. Pokażą pełną moc w systemach, gdzie posiadanie zrekonstruowanych prawdziwych przedmiotów w trój wymiarze połączone z trójwymiarowymi wirtualnymi przedmiotami jest kluczowe dla prawidłowego wykonania zadania. Możliwy może być również przypadek, gdzie niektóre zadania nie byłyby możliwe do wykonania bez posiadania kombinacji takich informacji (informacje rzeczywiste + informacje wirtualne).

Jest kilka sposobów zastosowania, które skorzystają z użycia teleimersji. Mechanizm projektowania przy pomocy komputera i różne medyczne zastosowania, są to tylko dwoma, które bardzo skorzystają.

Na przykład, grupa projektantów będzie w stanie współpracować interaktywnie w oddalonych od siebie miejscach. Będą mogli manipulować wirtualnym modelem zaczynając od konceptu, przedyskutować każdy krok projektu, wykonywać wymagane oceny i symulacje a nawet kończyć cykl wyprodukowaniem danej części na frezarce.

W przypadku zastosowania w medycynie – jak tele-radiologii lub do nagłych diagnoz, dostępność takich technologii w miejscach , gdzie jest to fizycznie niemożliwe mogłoby potencjalnie ratować życie. Statki morskie i platformy wiertnicze są dobrymi przykładami takich miejsc.

Telekabina

Telekabina reprezentuje immersyjny interfejs następnej generacji. Może być postrzegana jako połączenie wszystkich możliwych interfejsów. Biuro ukazane jest jako jeden prostokąt w większej wirtualnej biurowej przestrzeni. Płótna, na których pokazywany jest obraz są stereo-immersyjnymi powierzchniami biurka jak i ścian. Gdy podłączymy się do sieci, ściany wydają się przezroczystymi przejściami do innych telekabin. Naszym celem jest sprawienie by każdy fizyczny, potencjalnie nieregularny lub okrągły przedmiot mógł przejąć rolę płótna. Książki użytkownika, chusteczka a nawet ciało mogło by służyć takiemu celowi.



Obraz pokazuje wyobrażenie artysty o telekabinie lub biurze przyszłości.

Taki system reprezentuje unifikację wirtualnej rzeczywistości i wideokonferencji i dostarcza możliwości do pełnej integracji wirtualnej rzeczywistości do życia codziennego biura. Fizyczne i wirtualne środowiska zjednoczone w otrzymywaniu danych i na obrazie. Ta kombinacja oferuje nowy model służący ludzkiej komunikacji i współpracy.

Osiągnięcia

W ostatnich latach praca nad projektem rozwijała się w dwóch kierunkach:

1. Teleimersyjną technologię mediów – odtwarzanie trójwymiarowego obrazu, technologie rzutowania obrazu i wyświetlania, technologie śledzenia ruchu.
2. Strumień oprogramowania – systemy oprogramowania do wsparcia zastosowania teleimersji.

Dzięki takiemu ukierunkowaniu dokonał się najważniejszy postęp :

- 2/3 kamerowy system stereo do odtwarzania w trójwymiarze
- prototypy „Biuro przyszłości” ,”Biuro teraźniejsze”, Biuro trójwymiarowe”
- prototyp telekabiny używanej do dwu- i trój- stronnej współpracy
- pierwsze tanie kamery wbudowane w ekran do wyświetlania
- precyzyjnie składane kamery z obrazem panoramicznym
- inteligentne rzutniki do wychwytywania modeli dynamicznych za pomocą techniki wykorzystującej niedostrzegalne strukturalne światło oraz pokazujące obraz dużej rozdzielczości na zaprojektowanych do tego powierzchniach – płótnach
- SOFT: oprogramowanie, które wspiera przesyłane graficznie informacje w pomniejszych wirtualnych środowiskach przez powiększenie go na tle siatki danych.
- BAMBOO: platforma i język, niezależny mechanizm umożliwiający dynamiczną rekonfigurację zastosowań. Jest to zarówno unowocześnienie jak i środowisko przekazu.

Na ostatniej konferencji zademonstrowany był system, który zjednoczył przyswajanie danych do trójwymiaru (prawdziwych danych), syntetycznych przedmiotów(dane wirtualne) i interakcję między użytkownikiem i obiektami w 3D przy użyciu wirtualnego laserowego wskaźnika. Uczestnicy sesji immersyjnej byli nie tylko w stanie widzieć się nawzajem, ale mogli również współpracować i brać czynny udział w procesie projektowania.

Rola „Zaawansowanych Sieci i Usług”

Projekt teleimersji jest wspierany przez „Advanced Network & Services”. Firma działa jako katalizator w łączeniu współpracy między uznanymi ekspertami w dziedzinie rzeczywistości wirtualnej, pod przewodnictwem Jaron’a Lanier’a. Jej

celem jest postęp w nauce przez przyśpieszenie użycia komputerów, zastosowań, i technologii z nimi związanych.

Współpraca nad projektem.

Obecnie nad projektem współpracują cztery grupy:

1. Advanced Network & Services, Armonk, New York, kierownik naukowy: Jaron Lanier
2. Brown University, Providence, Rhode Island, Czołowy badacz: Dr. Andries van Dam (1997/obecnie).
3. University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, Czołowy badacz: Dr. Henry Fuchs (1997/obecnie)
4. University of Pennsylvania, Philadelphia, Pennsylvania, Czołowy badacz: Dr. Ruzen Bajcsy (1997/obecnie).

W przeszłości kilka innych grup współpracowało nad tym projektem:

1. Naval Postgraduate School, Monterey, California, Czołowy badacz: Dr. Michael Zyda (1997/99)
2. Carnegie Mellon University, Pittsburgh, Pennsylvania, Czołowy badacz: Dr. Randy Pausch (1997/98)
3. Columbia University, New York, Czołowy badacz: Dr. Steven K. Feiner (1997/98)
4. University of Illinois at Chicago, Chicago, Illinois, Czołowy badacz: Dr. Thomas A. DeFanti (1997/98)
5. University of Southern California, Los Angeles, California, Czołowy badacz Dr. Ulrich Neumann (1997/98)

Współpraca rozwija się również z Internet2 Working groups, zwłaszcza z grupą IPPM

I. Wprowadzenie: Grafika i wizualizacja.

Wstęp

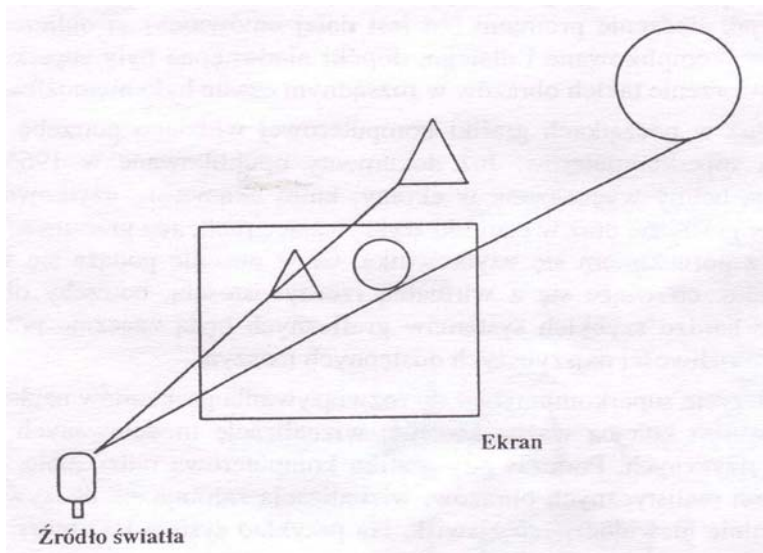
Zastosowanie grafiki komputerowej obejmuje wiele dziedzin tworzenia i odtwarzania obrazów. Zastosowanie grafiki komputerowej to nie tylko nowoczesne, interfejsowe biblioteki użytkownika (np. Apple Macintosh, X Windows), modelowanie trójwymiarowe, ale również animacja komputerowa, interaktywna edycja obrazów wideo, wizualizacja wyników np. tomografii komputerowej i rezonansu magnetycznego, wirtualna rzeczywistość, i również teleimersja. Wraz ze wzrostem wydajności komputerów wzrosła możliwość wykorzystania obrazów tworzonych przez komputerowe systemy graficzne jak i samych technik ich tworzenia.

Większość z systemów graficznych wymaga dużych zasobów obliczeniowych. Techniki pozwalające na tworzenie realistycznych obrazów, takie jak śledzenie promieni są obliczeniowo bardzo skomplikowane, dzięki wysokim możliwościom obliczeniowym dzisiejszych komputerów jest to możliwe. Już dokumenty opublikowane w 1965 roku opisują hełmy wyposażone w ekrany, które ukazywały użytkownikowi obrazy graficzne oraz czujniki wykrywające ruch, aby przesuwając obraz wraz z poruszaniem się użytkownika - wiąże się to z wirtualną rzeczywistością. Użycie superkomputerów do rozwiązywania problemów naukowych wprowadza kolejną ważną kwestię: wizualizację modelowych systemów fizycznych. Podczas gdy grafika komputerowa radzi sobie z tworzeniem realistycznych obrazów, wizualizacja zajmuje się ukazywaniem normalnie niewidocznych zjawisk. Na przykład system taki może utworzyć obraz cząsteczki DNA.

Techniki graficzne:
Śledzenie promieni

Większość technik przetwarzających i przedstawiających obrazy jest używana do przedstawiania obiektów trójwymiarowych na dwuwymiarowej płaszczyźnie. Te techniki, aby powstał rzut dwuwymiarowy, muszą uwzględniać określenie oświetlenia, a także wizualną informację o powierzchni.

Przestrzeń obiektu składa się z abstrakcyjnych obiektów w przestrzeni trójwymiarowej, które są rzutowane na dwuwymiarową płaszczyznę.



Rzutowanie na powierzchnię dwuwymiarową

Modele oświetlenia

Istnieją trzy podstawowe metody dołączania oświetlenia do obrazu graficznego:

1. Model empiryczny. Modele empiryczne starają się, aby powstały obraz wydawał się identyczny ze swoim odpowiednikiem w rzeczywistości. Na przykład modele te wykorzystują techniki interpolacyjne, aby utworzyć obraz płynnego cieniowania obiektów.
2. Model przejściowy. Modele przejściowe używają bardziej złożonej geometrii, aby powstał dokładniejszy obraz. Na przykład uwzględniona jest prawdziwa geometria obiektów graficznych, a w odniesieniu do tekstury i powierzchni uwzględnia się wyższy stopień szczegółowości.
3. Model analityczny. Modele analityczne używają opisów zjawisk fizycznych w celu uzyskania jak najbardziej realistycznej reprodukcji obrazu graficznego. Na przykład w modelu analitycznym są zawarte prawa energii świetlnej i sposób, w jaki światło reaguje z powierzchniami o różnych współczynnikach odbicia i załamania.

Wszystkie algorytmy tworzenia realistycznych obrazów muszą brać pod uwagę, że podczas modelowania prawdziwych obiektów różne ich części są konstruowane z różnych materiałów, które mogą mieć różne stopnie przezroczystości i odbicia światła. Wynika z tego, że w tworzonym obrazie widoczne mogą być tylko fragmenty kątów i przednich części obiektów. Nosi to nazwę „problemu zasłoniętego obrazu”. Z tego powodu wiele algorytmów tworzenia obrazów bada względny kolor i intensywność każdego punktu porównując jego atrybuty

oświetlania. Wszystkie wymienione algorytmy zakładają realizowanie iteracji dla przestrzeni trójwymiarowej. Opisany będzie tu tylko jeden algorytm, wykorzystywany do śledzenia promieni.

Piksel lub element obrazu(ang. pixel –picture element) jest najmniejszym punktem wyświetlanym na ekranie. Elementem objętości jest woksel (ang. Volume element).

Śledzenie promieni

Podstawowa koncepcja realizacji algorytmu wykorzystywanego do śledzenia promieni zakłada uwzględnienie wszystkich możliwych źródeł światła i sposobu, w jaki wpływają one na poszczególne obiekty znajdujące się w przestrzeni obiektowej. Promienie są prowadzone z każdego źródła światła przez każdy piksel znajdujący się w przestrzeni dwuwymiarowej, a następnie do przestrzeni obiektowej. Ten obiekt, do którego promienie dotrą jako do pierwszego napotkanego na swojej drodze, jest obiektem, który pojawia się na powierzchni widocznej. Gdy promień pochodzący ze źródła światła dotrze do powierzchni, następuje modelowanie kilku zjawisk fizycznych po pierwsze, każdy obiekt, który znajduje się za punktem przecięcia promienia z powierzchnią, pozostaje w cieniu dla danego źródła światła co z kolei wskazuje na to, że żadne ze źródeł światła w cieniu nie będzie oświetlało tego punktu przecięcia. Po drugie, jeżeli materiał tworzący powierzchnię powoduje odbicie lub refrakcję światła, z powierzchni rozchodzą się promienie wtórne, których zadaniem jest modelowanie tych zjawisk. Takie modele wtórne mogą następnie także oświetlać inne punkty znajdujące się w przestrzeni obiektowej.

Złożoność obliczeń przebiegu promieni

Do określenia przebiegu promieni niezbędne są obliczenia, oparte na przecięciu promieni w określonej wiązce światła z obiektem umieszczonym w przestrzeni obiektowej. Dowolny punkt X położony na promieniu może być opisany za pomocą równania:

$$X = P + \lambda D, \lambda \geq 0$$

Gdzie P jest początkowym punktem promienia, λ jest odległością wzdłuż promienia, a D jest kierunkiem (wyrażonym za pomocą współrzędnych x, y, z).

Najprostszym obiektem, dla którego można przygotować model, jest sfera. Może ona być opisana za pomocą określenia jej środka i promienia. Dowolny punkt X na powierzchni sfery może być opisany jako:

$$|X - V| = r$$

gdzie V jest środkiem sfery, a r – jej promieniem. Podstawiając pierwsze równanie do drugiego i upraszczając je, otrzymujemy równanie kwadratowe:

$$\lambda^2 + (2(P-V) \times D)\lambda + (P-V)^2 - r^2 = 0$$

Rozwiązaniem tego równania są punkty promienia, które przecinają powierzchnię sfery. Dalsze przekształcania tego równania wskazują, że znalezienie punktów przecięcia wymaga wykonania 17-stu działań dodawania lub odejmowania, 17-stu działań mnożenia i jedno obliczenie pierwiastka kwadratowego, czyli łącznie 35 działań zmiennoprzecinkowych.

Typowa przestrzeń obiektowa może składać się z tysięcy sfer, wielu źródeł światła i można założyć, że powierzchnie sfer zarówno odbijają, jak i rozpraszają światło (co zwiększa liczbę promieni, których przebieg należy opisać). Zakładając powierzchnię złożoną z 1000x1000 pikseli, 1000 sfer i przeciętnie 5 promieni przypadających na jeden piksel, liczba działań potrzebnych do określenia przebiegu promieni wynosi 1000x1000 pikseli x 5 promieni na piksel x 1000 sfer na promień x 35 działań na sfer, czyli 175 bilionów działań zmiennoprzecinkowych – jest to poważne zadanie obliczeniowe. Znaczne wymagania co do mocy obliczeniowej nie są charakterystyczne tylko dla przypadku śledzenia przebiegu promieni świetlnych; występują one zawsze, jeżeli zostanie zastosowany algorytm syntezy obrazu. Jednym ze sposobów zapewnienia takich sposobów obliczeniowych jest rozdzielenie różnych operacji na zestaw procesorów równoległych.

Równoległe przetwarzanie graficzne

Najwcześniejszym wykorzystaniem procesorów równoległych w przetwarzaniu graficznym było ich użycie do potokowego przetwarzania różnych faz operacji graficznych. Systemy opracowane przez Computer Generated Imagery (CGI) w latach 70-tych umożliwiły podzielenie procesu obliczeniowego na cztery fazy i połączenie ich w jeden potok:

1. Menadżer sceny. Reprezentacje obiektów, przechowywane w pamięci masowej, są przekazywane do systemu przez menadżera sceny, po uprzednim posortowaniu obiektów w celu określenia ich widocznych powierzchni.
2. Procesor geometryczny. Posortowane obiekty są przesyłane z menadżera sceny do procesora geometrycznego. Procesor ten realizuje obliczenia i przetwarzanie niezbędne dla rzutowania obiektu na płaszczyznę obrazową,

ustalenia, które fragmenty obiektów są aktualnie widoczne w płaszczyźnie obrazu, oraz ustalenia gradientów poszczególnych barw.

3. Procesor wideo. Procesor wideo eliminuje zasłonięte części obiektów, a także określa wartość intensywności oświetlenia.
4. Wyświetlacz. Wyświetlacz jest zwykle graficznym monitorem barwnym o wysokiej rozdzielczości.

Przetwarzanie potokowe tych działań umożliwia wprowadzenie równoległości do przetwarzania graficznego, lecz zapotrzebowanie systemów graficznych na przetwarzanie równoległe wymaga o wiele wyższego poziomu równoległości. Najbardziej rozbudowany przypadek polega na przypisaniu procesora do każdego piksela danej sceny. Koszty opracowania i wykorzystania takiego systemu byłyby jednak wysokie i dlatego tworzy się takie systemy, w których procesor jest przypisywane do większej części(regionu) sceny, przedstawiającej zbiór pikseli.

Wyświetlacz 8x8

Typowym rozwiązaniem umożliwiającym przypisanie regionów sceny do zbioru procesorów jest zbudowanie maszyny z 64-ma procesorami ułożonymi w tablicy o wymiarach 8x8. Obraz ekranu może być wówczas podzielony cyklicznie(każdy co ósmy piksel w rzędzie jest przypisany do tego samego procesora i każdy co ósmy piksel w kolumnie jest także przypisany do tego samego procesora). Dzięki temu uzyskuje się podział obliczeń realizowanych dla każdego obiektu między zbiór procesorów. Ponieważ niektóre obiekty wymagają bardziej skomplikowanych obliczeń niż inne, takie cykliczne przypisanie pikseli do procesorów sprawia, że obciążenie z tytułu wykonywania obliczeń staje się bardziej równomierne.

Sprzęt specjalizowany: Relity Engine opracowany przez SGI

Relity Engine jest specjalizowanym systemem graficznym. Jego celem jest przedstawianie obiektów, które są oświetlone, gładkie, cieniowane i posiadają teksturę. System ten składa się z wielopłytkowego przyspieszacza, zainstalowanego w stacji roboczej MIPS, jest maszyną wykorzystywana do przetwarzania potokowego, która może przetwarzać dane teksturalne w taki sposób, jak gdyby były to obiekty rzeczywiste, a nie powtarzające się fragmenty wzorców.

Odwzorowanie tekstury polega na wykorzystaniu skanowanych wcześniej, rzeczywistych „tekstur”, które są następnie odwzorowywane na powierzchnie

tworzonego obrazu. Kartowanie tekstury ma na celu zwiększenie realistyczności przedstawianej powierzchni. Przykładem odwzorowania tekstury może być przedstawienie tekstury trawy w programie symulacji lotu, które służy realistycznemu przedstawieniu powierzchni terenu. Innymi przykładami są liście na drzewie, ściana z cegieł przedstawiona jako element budynku tworzonego w pamięci wirtualnej, lub nawet zwykła etykieta adresowa, która jest dodawana do obrazu przesyłki pocztowej.

Reality Engine został utworzony jako zbiór następujących elementów:

- kolejek FI FO,
- procesora poleceń,
- procesorów geometrii,
- generatorów fragmentów,
- procesorów obrazów.

Kolejka **FI FO** (ang. **First-In – First-Out**, czyli pierwszy na wejściu-pierwszy na wyjściu) pozwala na takie uporządkowanie danych, że dane opuszczają kolejkę w takim samym porządku, jak wchodziły do kolejki.

Elementy te są połączone w taki sposób, że dane wejściowe wpływają przez kolejkę FI FO do procesora poleceń. Procesor ten rozdziela dane wejściowe na dowolną liczbę procesorów geometrii. Procesor geometrii wykorzystuje procesor Intel i860 i używa go do ustalenia oświetlenia obiektów, określenia ich współrzędnych, rzutowania poligonów na trójkąty składowe. Dane wychodzące z procesora geometrii są formowane jako dane wejściowe dla generatorów fragmentów. Dane przepływają przez generatory fragmentów w postaci potoku. Generator fragmentów pobiera fragmenty (tzn. trójkąty z procesora geometrii) i wykonuje rastrowanie, a także realizuje wszelkie obliczenia potrzebne w celu określenia tekstury. Fragmenty pochodzące z generatora fragmentów są następnie rozdzielane na poszczególne procesory obrazów. Procesor obrazów finalizuje obliczenia dotyczące obrazu i przekazuje obraz do bufora. Bufor ten jest podzielony na obszary, z których każdy jest sterowany przez jeden procesor obrazu.

Reality Engine może być konfigurowany z 5, 10 lub 20 generatorami fragmentów i z 80, 160 lub 320 generatorami obrazów. Kombinacja taka jest masowym, równoległym systemem graficznym, który jest w stanie zrealizować odwzorowanie

tekstury dla ponad miliona trójkątów w ciągu jednej sekundy, co z kolei pozwala na tworzenie obrazów i ich interaktywne przetwarzanie w czasie rzeczywistym.

Wizualizacja

Wykorzystanie systemów graficznych jest ściśle powiązane z wieloma aplikacjami. Możliwość modelowania różnych układów fizycznych za pomocą komputera staje się niemal bezużyteczna, jeśli jej wyniki nie mogą być „widziane”. Systemy wizualizacji wykorzystywane dla celów naukowych są włączane do wysoko wydajnych aplikacji, dzięki czemu użytkownik może obserwować symulowane modele. Różnica między przetwarzaniem graficznym i wizualizacją polega na tym, że przetwarzanie graficzne wykorzystuje się w celu wyświetlania określonych obrazów, natomiast wizualizacja wykorzystuje techniki graficzne w celu wyświetlania informacji.

Wizualizacja polega na tworzeniu widzialnych obrazów przedmiotów i obiektów, których nie jesteśmy w stanie zobaczyć; może ona także polegać na wyświetlaniu, w postaci wizualnej – informacji ilościowej i jakościowej.

Wykorzystując obrazy wizualne można przekazywać informacje jakościowe w sposób bardziej zrozumiały niż zestawy liczb. Celem wizualizacji jest umożliwienie percepcji danych, a bardzo wiele informacji, które są przetwarzane i tworzone przez liczne aplikacje wykorzystujące komputery, może być zrozumiane przez użytkowników dopiero wtedy, gdy zostaną przetworzone do postaci wizualnej.

Przykłady wizualizacji naukowej

Niektórymi przykładami wizualizacji naukowej są:

- przepływ powietrza wokół samochodu lub samolotu,
- przepływ oleju przez substancje porowate,
- atomowa struktura molekuł,
- molekularna struktura materiałów,
- globalne prądy oceaniczne,

- symulowane wypadki samochodowe,
- modele klimatyczne i pogodowe.

Specjalizowany sprzęt do wizualizacji: Gigawall

Projekt wizualizacji, realizowany przez Army High Performance Computing Research Center wykorzystuje zestaw superkomputerów i posługuje się nim dla potrzeb wielkiego systemu wizualizacji naukowej, który nazywa się Gigawall. Gigawall jest systemem zdolnym do obsługi w ciągu jednej sekundy gigabajtów pikseli ułożonych w ściany wizualne o wymiarach 6 x 10 stóp (około 1.8 x 3.0 metra), z zachowaniem odświeżania ekranu 60 razy na sekundę.

Gigawall umożliwia interaktywną wizualizację przez przesuwanie i zmianę skali na powierzchni w rzeczywistości wirtualnej. Użytkownik, jeżeli dysponuje specjalnymi okularami do obserwacji stereoskopowych, może obserwować przetworzenia w trzech wymiarach, patrząc z różnych punktów obserwacji, nawet patrząc ze środka bryły!

Gigawall wykorzystuje wiele zestawów Reality Engine, MasPar MP-2s, TMC Connection Machines lub Ncube-2s, które są połączone razem za pomocą HiPPI lub sieci ATM i podłączone do szybkiego zbioru macierzy dyskowych umożliwiającym prowadzenie operacji wejścia/wyjścia.

Rzeczywistość wirtualna

Rzeczywistość wirtualna jest technologią, która została opracowana po to, aby użytkownik został otoczony trójwymiarowym środowiskiem wygenerowanym przez komputer, w którym można doświadczać kontaktów ze „światem wirtualnym”. Taki świat wirtualny może być wzmacniany interaktywnie za pomocą specjalnych urządzeń, które pozwalają na tworzenie bodźców wizualnych, dotykowych i dźwiękowych.

Technologia rzeczywistości wirtualnej (Virtual Reality –VR) jest bardzo przydatna w wielu aplikacjach.

- Wizualizacja naukowa. Rzeczywistość wirtualna wspomaga działanie naukowców związane z analizą jakościową i interpretacją olbrzymich ilości danych.
- Edukacja. Dobrze znanym systemem rzeczywistości wirtualnej wykorzystywanym w edukacji jest symulator lotów stosowany w czasie szkolenia pilotów.
- Działania w zagrożonym środowisku. Podłączenie systemu rzeczywistości wirtualnej do zdalnego systemu robotów pozwala na realizowanie działań w takim środowisku, które jest niebezpieczne dla ludzi, jak na przykład na

obszarach składowisk materiałów nuklearnych, które muszą zostać oczyszczone.

- Chirurgia wirtualna. System VR podłączony do mechanizmu robota może umożliwić chirurgom zdalne przeprowadzenie operacji.
- Gry. Przyjemność gry może być znacznie wzmocniona, gdy gracz może prowadzić grę w sposób bardziej interaktywny. Rzeczywistość wirtualna pozwala graczowi na zaistnienie w świecie gry.

Równoległość w rzeczywistości wirtualnej

Istnieją trzy poziomy możliwości wykorzystania równoległości w systemach rzeczywistości wirtualnej:

- Środowisko. Równoległe można symulować różne środowiska.
- Obiekty. Można równoległe przetwarzać różne obiekty.
- Atrybuty. Każdy obiekt może składać się z kilku połączonych ze sobą, mniejszych elementów. Istnieje możliwość wykorzystania równoległości do obliczania atrybutów każdego obiektu.

CAVE

CAVE jest projektem z dziedziny rzeczywistości wirtualnej, który został opracowany w National Center for Supercomputer Applications (NCSA). CAVE jest skrótem od nazwy „CAVE Automatic Virtual Environment –CAVE Automatyczne Środowisko Wirtualne”. CAVE jest trójwymiarowym obszarem w rzeczywistości wirtualnej, który był zaprojektowany dla potrzeb wizualizacji naukowej. Był on projektowany jako obszar o rozmiarach 10x10x10, z trzema ścianami dla projekcji tylnej i z podłogą dla projekcji dolnej. Ściany i podłoga składają się z wysoko rozdzielczych obrazów barwnych (o rozdzielczości 1280 na 1024, piksele obrazowe 24-bitowe), tworzonych przez wysoko wydajne stacje robocze skonstruowane przez Silicon Graphics. CAVE powstał w wyniku zapotrzebowania na przekształcenie przetworzeń graficznych realizowanych za pomocą monitora i systemów graficznych w taki system, który byłby systemem bardziej interaktywnym i umożliwiał realizację wizualizacji, także z wewnątrz przestrzeni.

CAVE jest zaprojektowany w taki sposób, że użytkownik może bezpośrednio działać w przestrzeni trójwymiarowej. CAVE wykorzystuje bryłę sześcianu, która

służy do wywołania uczucia otoczenia użytkownika sferą. Projektanci CAVE wykorzystali technologię rzutowania obiektów trójwymiarowych na powierzchnię dwuwymiarową w celu „połączenia” krawędzi ścian, dzięki czemu powstało wrażenie sfery. Użytkownik używa specjalnych okularów, które wraz z wyświetlonym obrazem trójwymiarowym, pozwalają użytkownikowi oglądać trójwymiarowe objekty. Dzięki synchronizacji ekranu zmniejsza się jego migotanie. Ponieważ trzy ściany obrazu są wyświetlane w projekcji tylnej, nie występują na nich cienie, lecz cienie powstałe w podłodze, która jest rzutowana w projekcji dolnej. CAVE został wyposażony w specjalne czujniki elektromagnetyczne, umieszczone właśnie w podłodze, które śledzą użytkownika i umożliwiają tworzenie cieni.

CAVE jest prawdziwym połączeniem zastosowań graficznych, bardzo szybkich sieci i wysoko wydajnego przetwarzania danych. Wielkość mocy obliczeniowej, potrzebnej do ciągłego aktualizowania w czasie rzeczywistym skomplikowanych modeli, o wiele przekracza dostępną obecnie moc obliczeniową. Na przykład aby usunąć efekt migotania obrazu, musi on być odświeżany 60 razy na sekundę. Z kolei aby 60 razy w ciągu jednej sekundy przekazać obraz jednej ściany w CAVE potrzeba możliwości przesyłania w sieci prawie 2 gigabajtów danych. I to musi być wykonane dla czterech ścian!

CAVE jest projektem badawczym w toku. Niektóre z badanych problemów obejmują rzutowanie na wszystkie sześć ścian sześcianu, obniżanie kosztów przedsięwzięcia i zwiększenie solidności, w celu umożliwienia tworzenia modeli, które mogłyby funkcjonować w mniej „pilnowanych i kontrolowanych” środowiskach (jak np. muzeach czy szkołach). CAVE został po raz pierwszy przedstawiony w czasie konferencji SIGGRAPH w 1992 roku, był także prezentowany w tym samym roku w czasie konferencji Supercomputing.

II. Teleimersja: Wirtualna obecność.

Podczas wideokonferencji użytkownicy porozumiewają się tylko za pomocą słów, a istotny jest również przekaz niewerbalny, który ginie podczas takiego przekazu. W czasie wideokonferencji nie daje się na przykład nawiązać kontaktu wzrokowego, gdyż kamera i ekran nie mogą być w tym samym miejscu. Taka rozmowa staje się sztuczna, bo nie można patrzeć sobie w oczy, co na ogół sprzyja

atmosferze obopólnego zaufania. Co więcej, uczestnicy konferencji nie są w stanie uzmysłwić sobie wzajemnego usytuowania, a więc nie mają prostego sposobu zwrócenia na siebie uwagi, wyrażenia aprobaty czy też dezaprobaty. Teleimersja – to nowy sposób kontaktów międzyludzkich możliwy dzięki technikom cyfrowym daje złudzenie, że człowiek przebywa w tym samym pomieszczeniu, co osoby, które w rzeczywistości znajdują się tysiące kilometrów od nas. Łączy ona techniki wizualizacji oraz interakcji z otoczeniem – charakterystyczne dla rzeczywistości wirtualnej – z nowymi osiągnięciami wizualizacyjnymi, które wychodzą poza tradycyjne możliwości kamer wideo. Zamiast obserwować uczestników i ich otoczenie z jednego tylko punktu, stanowiska teleimersji przedstawiają człowieka jako “ruchomą rzeźbę”. Nie wyróżnia się żadnego punktu obserwacji, co powoduje, że wszyscy uczestnicy znajdują się pozornie w tej samej przestrzeni i postrzegają ją tak samo, tak jak naprawdę wygląda. Teleimersja została pomyślana nie tylko jako medium pozwalające udoskonalić wideokonferencje, lecz także jako idealne narzędzie do rozwijania nowatorskich technik telekomunikacyjnych; szczególnie dla Internet2 – konsorcjum zajmującego się zaawansowanymi technikami teleinformatycznymi w USA. Jeśli sieć może obsłużyć teleimersję, to zapewne podoła każdemu innemu zadaniu, gdyż oprócz tradycyjnych wymagań, jak szybki i niezawodny przepływ danych, warunkiem sukcesu tej aplikacji są bardzo małe opóźnienia w przekazywaniu informacji (ponadto opóźnienia nie mogą się zmieniać w trakcie transmisji).

Wirtualna rzeczywistość i sieci

Teleimersja wykorzystuje zarówno techniki wirtualnej rzeczywistości i szybkie sieci, jak i widzenia komputerowego oraz projektowania interfejsów użytkownika. Zajmowanie się nią wymaga więc pewnego doświadczenia we wszystkich tych dziedzinach.



Telewspółpracownicy oddaleni od siebie o setki kilometrów dyskutują o modelu anatomicznym wygenerowanym przez komputer jak o rzeczywistym przedmiocie. Helm użytkownika pomaga komputerowi w lokalizacji położenia głowy i jej orientacji, co jest niezbędne do wytworzenia właściwego obrazu na ekranie. W przyszłości urządzenia te nie będą konieczne.

W 1965 roku Ivan Sutherland, powszechnie uznawany za ojca grafiki komputerowej, zaproponował coś, co nazwał “obrazowaniem doskonałym”. Dzięki niemu użytkownik miałby wrażenie przebywania w rzeczywistym otoczeniu, które jednak byłoby całkowicie tworzone przez komputer. Sutherland nazwał to “światem wirtualnym”, terminem używanym w filozofii estetyki, szczególnie w pismach Suzanne K. Langer. W 1968 roku Sutherland po raz pierwszy stworzył wirtualny świat za pomocą urządzenia zwanego wyświetlaczem nagłownym. Był to rodzaj hełmu z dwoma ekranami naprzeciwko oczu, co dawało wrażenie przebywania w trójwymiarowej przestrzeni. Kiedy poruszało się głową, komputer błyskawicznie dokonywał obliczeń i wytwarzał obraz, jaki powinniśmy widzieć w tej pozycji, aby zachować złudzenie przebywania w nieruchomym otoczeniu. W latach osiemdziesiątych, Jaron Lanier – pionier wirtualnej rzeczywistości, znalazł się dzięki Scientific American u steru pierwszej firmy, która sprzedawała uniwersalne narzędzia do tworzenia wirtualnych światów i przebywania w nich. Firmę nazwał VPL(Visual Programing Language lub Virtual Programing

Language). W połowie lat osiemdziesiątych VPL zaczęła sprzedawać sprzęt do tworzenia rzeczywistości wirtualnej i była szeroko znana ze swoich rękawic.



Trzy osoby w trzech różnych miastach mogą dzięki temu pomieszczeniu przebywać w tej samej wirtualnej przestrzeni.

W połowie lat osiemdziesiątych w firmie VPL przeprowadzono pierwsze eksperymenty z czymś, co zdecydowano nazwać wirtualną rzeczywistością. Pojęcie to łączy ideę wirtualnych światów z techniką sieci: umieszcza się uczestników wyposażonych w wyświetlacze hełmowe w tej samej wirtualnej przestrzeni. W 1989 roku firma VPL wypuściła na rynek przyrząd RB2 (skrót od Reality Built for Two – rzeczywistość dla dwojga), który pozwalał dwóm osobom przebywać w tym samym wirtualnym świecie. Ciekawym aspektem wirtualnej rzeczywistości jest to, że uczestnicy widzą wymodelowane przez komputer sylwetki innych osób – tzw. awatary (w hinduizmie zejście, wcielenie bóstwa). Mimo że ówczesne możliwości komputerów pozwalały na modelowanie awatar w bardzo ograniczony sposób – były to nadzwyczaj uproszczone figurki, jak z kreskówek. W bardzo przybliżony sposób oddawały rysy twarzy, ale modelowały wiernie ruchy ludzi, których wyobrażeniami były w wirtualnym świecie, dając w ten sposób poczucie obecności, emocji i centrum zainteresowania. Na początku w wirtualnych światach poruszały się osoby, które były fizycznie blisko siebie, ale zrobiono także parę doświadczeń odsuwając je na znaczną odległość. Zaaranżowali sesje w wirtualnej rzeczywistości z uczestnikami z Japonii i Kalifornii, a także z Niemiec i Kalifornii. W tych próbach

nie stawiano sieci zbyt dużych wymagań, gdyż należało przekazywać tylko informacje o ruchach uczestników, a nie szczegóły ich postaci, tak jak w teleimersji. Badania nad sieciami komputerowymi rozpoczęły się w tym samym okresie. Pierwsza sieć, Arpanet, powstała pod koniec lat sześćdziesiątych. Zainspirowała ona powstanie innych podobnych sieci, aż wreszcie w latach osiemdziesiątych połączyły się one w jedną Sieć – Internet. W miarę rozwoju Internetu tworzono “sieci szkieletowe”. Są one sieciami w sieci, po których dane przesyłane są z wyjątkowo dużą przepustowością i prędkością. Pozwala to szybko przekazywać informacje na duże odległości. Wśród sieci szkieletowych warto wyróżnić te przeznaczone do zastosowań naukowych – NSFnet (NSF – National Science Foundation) z późnych lat osiemdziesiątych i vBNS z połowy lat dziewięćdziesiątych. Oba te przedsięwzięcia odegrały dużą rolę w inspirowaniu nowatorskich zastosowań Internetu, takich jak WWW. W 1998 roku rozpoczęto budowę nowej supersieci, zwanej Abilene, która obsługuje uniwersyteckie konsorcjum Internet2. Dziś Abilene łączy ponad 170 uczelni w USA. Gdyby jedynym celem Internet2 było udostępnienie szybkiej (tj. szerokopasmowej) sieci, to wystarczyłaby Abilene. Jednak Internet2 miał dalsze cele, m.in. stworzenie nowych protokołów przesyłania danych, zapewniających nie tylko dużą przepustowość, lecz także krótkie czasy oczekiwania (tzn. opóźnienia spowodowane obróbką danych w czasie ich przesyłania). Internet2 miał pewien kłopot: żadne z ówczesnych zastosowań nie było dla sieci wystarczającym wyzwaniem. Rozwój informatyki zazwyczaj opierał się na założeniu, że zawsze znajdą się zastosowania dla nowych, wydajniejszych urządzeń, nawet jeśli nie wiemy dokładnie, jakie to będą zastosowania. W przypadku nowych szerokopasmowych sieci wiara ta nie wystarczała. Należało jakoś przetestować nową sieć. Allan H. Weis, jeden z twórców NSFnet, przewodził fundacji Advanced Network and Services, która zarządzała biurami projektów Internet2. Użył on terminu “teleimersja” do opisu idealnego, najważniejszego zastosowania szybkiej sieci. Zaproponował utworzenie organizacji National Tele-immersion

Initiative i objęcie nad nią nadzoru naukowego. Była to kontynuacja prac nad wirtualnymi światami. Mimo że wiele składników niezbędnych do działania teleimersji, na przykład obrazowanie, wymagało jeszcze opracowania, najważniejszym problemem było stworzenie odpowiedniego sposobu

przedstawiania ludzi i ich otoczenia. Być może nie od razu jest jasne, czym teleimersja różni się od wideokonferencji.

Naczelną zasadą teleimersji jest to, że każdy z uczestników musi mieć własny punkt widzenia na modelowany świat, właściwie nie jeden, lecz dwa (dla każdego oka, aby dać złudzenie perspektywy). Co więcej, uczestnicy wspólnej sesji powinni mieć możliwość poruszania się, tak więc punkty widzenia ciągle się zmieniają. Scena powinna być rejestrowana w sposób nie wyróżniający żadnego punktu (inaczej niż przez kamerę, która filmuje scenę z określonej pozycji). Każde miejsce, wraz z ludźmi i przedmiotami tam się znajdującymi, musi być obserwowane ze wszystkich stron naraz i przedstawiane jako "ruchoma, trójwymiarowa rzeźba". Każde ze zdalnych stanowisk otrzymuje pełną informację niezbędną do wytworzenia obrazów dla obu oczu użytkownika. Zbieranie informacji musi być tak szybkie, aby działało to w czasie rzeczywistym, z opóźnieniem co najwyżej ułamek sekundy. Modeluje się wtedy awatarę uczestnika z taką częstotliwością, że powstaje złudzenie ciągłego ruchu. Zaczynamy mieć złudzenie ruchu, kiedy obrazy pojawiają się w tempie około 12.5 klatek na sekundę, ale złudzenie staje się pełne dopiero przy 25 klatkach na sekundę, a lepiej stosować jeszcze większe częstotliwości. Kontury poruszających się uczestników oraz przedmiotów można rejestrować na wiele różnych sposobów. Już w 1993 roku Henry Fuchs z University of North Carolina w Chapel Hill zaproponował metodę zwaną morzem kamer. Porównuje się w niej obrazy z wielu kamer. W typowych scenach z udziałem ludzi jest dużo szczegółów widocznych z więcej niż jednej kamery, jak na przykład fałda na swetrze. Porównując komputerowo obrazy tych samych szczegółów z różnych kamer, można utworzyć trójwymiarowy model danej sceny. Technikę tę testowano, lecz nie w czasie rzeczywistym. Warto wspomnieć o pracach Takeo Kanadego z 1995 roku uwieńczonych pokazem "Virtualized Reality"- Zwirtualizowana rzeczywistość w Carnegie Mellon University. Zamontowano wtedy na sferycznej konstrukcji 51 kamer, zwróconych do środka. Obrazów nie przetwarzano w czasie rzeczywistym, więc urządzenie nie nadawało się do teleimersji. Używano zapisu na taśmach wideo, które potem analizowano za pomocą komputerów. Ruzena Bajcsy, kierująca GRASP (General Robotics, Automation, Sensing and Perception) Laboratory w University of Pennsylvania, zainteresowała się koncepcją morza kamer w czasie rzeczywistym. Od 1994 roku, wraz z kolegami

z uniwersytetów w Chapel Hill i Carnegie Mellon, pracowała nad grupami dwu-trzech kamer działających w czasie rzeczywistym, nadających się do zastosowań w rzeczywistości wirtualnej. Bajcsy i jej współpracownik Kostas Daniilidis podjęli się zbudowania pierwszego na świecie morza kamer pracujących w czasie rzeczywistym. Co więcej, system miał być modularny i możliwy do rozbudowy, tak aby dało go się użyć w różnych miejscach i sytuacjach. System powstawał przy współdziałaniu naukowców z Chapel Hill, którzy byli odpowiedzialni za przetwarzanie danych o „ruchomych rzeźbach” w celu wytworzenia realistycznie wyglądających scen dla każdego użytkownika systemu. Jednak morze kamer nie rozwiązuje jeszcze problemu. Przypuśćmy, że wszystkie te kamery skierowane są na czystą, gładką białą ścianę. Ponieważ nie ma żadnych szczegółów, kamery nie dostarczają żadnej informacji pomocnej w odtworzeniu sceny. Człowiek może spoglądać na biały mur i nie mieć wątpliwości, co widzi. Ludzie nie martwią się, że ta biała plama to nie ściana, lecz nieskończenie głęboka, biała przepaść. Zwykle nie polegamy wyłącznie na informacji wizualnej – mamy w głowach model pokoju, co ogranicza zbyt nieprawdopodobne interpretacje odbieranych obrazów. Niestety, dla współczesnych kamer cyfrowych czoło człowieka lub podkoszulek są równie trudne do zinterpretowania jak biała ściana, a dzisiejsze oprogramowanie nie radzi sobie z tym problemem. Naukowcy z Chapel Hill wymyślili nową metodę, która okazała się obiecująca: ISL (Imperceptible Structural Light – niedostrzegalne światło strukturalne). Zwykle świetlówki migają 50 razy na sekundę, a więc tak często, że oko tego nie dostrzega. Podobnie ISL wydaje się ciągłym, białym światłem, ale w rzeczywistości są to szybkozmienne przebiegi rastrujące, widoczne tylko dla wyspecjalizowanych, zsynchronizowanych kamer. Standardowe wzory rzutowane

przez ISL wyświetlane są na gładkich powierzchniach, co pozwala zespołowi kamer



odtworzyć geometrię sceny.

Ten obraz powstał przez porównanie obrazów z dwu kamer systemu teleimersji. Kolory pokazują wyniki pierwszych oszacowań odległości różnych punktów postaci od miejsca, z którego prowadzono obserwację.

Po raz pierwszy zademonstrowano teleimersję 9 maja 2000 roku, łącząc wirtualnie trzy miejsca. W pokoju do teleimersji w Chapel Hill znajdowało się kilkanaście osobistości. Użytkownicy siedzą przy symulowanym biurku przyszłości. Przez okulary z polaryzującymi szklami widać dwie ściany pokoju przekształcone w okna, za którymi można było ujrzeć dwa inne pokoje, w których siedzieli inni ludzie. (Okulary gwarantowały, że każde oko otrzymywało trochę inny obraz, co daje efekt stereoskopii). To wszystko nie przypominało wirtualnych światów. Związek postrzeganych osób i ich otoczenia z rzeczywistością był dobrze wyczuwalny. Nie było to jednak perfekcyjne przedstawienie. Widać było “szum” – wyglądało to, jakby w innych pokojach sypał śnieg. Częstotliwość przekazywanych obrazów była niska (2–3 klatki na sekundę), a opóźnienia sięgały sekundy. W dodatku tylko w Chapel Hill był dostęp do w pełni teleimersyjnej instalacji. Ten wirtualny świat nie był uproszonym modelem, ale wiernym odbiciem rzeczywistości opartym na prawdziwych parametrach. W późniejszym pokazie (w październiku 2000 roku) śnieżenie już prawie nie występowało, a jakość obrazu i szybkość transmisji były znacznie lepsze. Jednak największą innowację wprowadzili naukowcy z Brown University pracujący pod kierunkiem Andriesa van Dama. Pojawili się oni na sesji teleimersji z wirtualnymi przedmiotami, które istniały tylko w ich komputerze. Było to interesujące połączenie technik wirtualnych światów i uczucia przebywania

w tym samym miejscu z inną osobą. Na biurku pomiędzy użytkownikami umieszczono wirtualny model gabinetu (o boku około pół metra), używano symulowanych wskaźników laserowych, aby wspólnie modyfikować wystrój wnętrza tego pomieszczenia. Tą próbę przeprowadzali Robert C. Zeleznik z Brown University i Jaron Lanier.

NA CZYM POLEGA TELEIMERSJA



Ten bardzo uproszczony schemat pokazuje, jak w przyszłości system teleimersji pozwoli dwóm osobom współpracować w projektowaniu silnika.

1. Niedostrzegalne światło strukturalne

Jest ono odbierane przez ludzkie oko jak zwykle białe światło, ale w rzeczywistości to krótko wyświetlane różne desenie, umożliwiające komputerowi lokalizację obszarów pozbawionych szczegółów, na przykład czoła.

2. Morze kamer

Ukryte kamery wideo obserwują scenę z różnych punktów, co umożliwia wytworzenie jej trójwymiarowego modelu. Kamery można umieścić za otworami w ekranie, tak jak na ilustracji, albo na suficie, jednak wtedy ekran musi selektywnie odbijać światło.

3. Wirtualne lustro

Użytkownik może sprawdzić, jak go postrzegają inni, na urządzeniu zwanym wirtualnym lustrem. W tym hipotetycznym przykładzie mężczyzna zdecydował, że będzie ukazywał się innym bardziej elegancko ubrany niż w rzeczywistości. Oprogramowanie pozwalające na takie metamorfozy jeszcze nie istnieje, ale pierwsze próby są w toku.

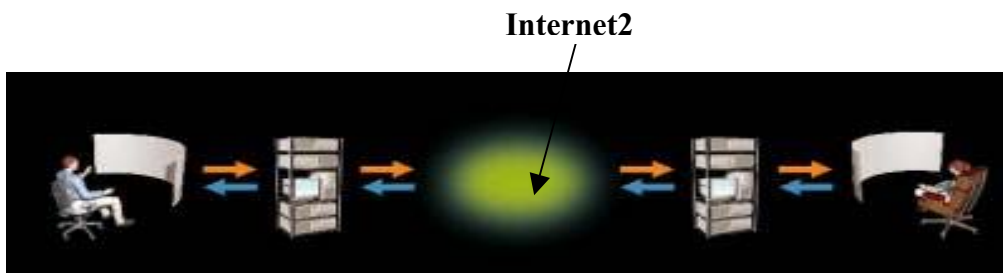
4. Wirtualne przedmioty

Symulowane obiekty pojawiają się w wirtualnej przestrzeni pomiędzy uczestnikami sesji teleimersyjnej. Można nimi manipulować, tak jakby były rzeczywistymi przedmiotami. Badania w ramach prac objętych National Tele-immersion Initiative dotyczą także technik integracji wirtualnych przedmiotów wytworzonych przez uczestników sesji używających różnego, niezgodnego ze sobą oprogramowania.

5. Ekran

Dzisiejsze prototypy wyświetlają dwa nakładające się obrazy o różnej polaryzacji światła i wymagają noszenia polaryzujących okularów, tak aby każde oko widziało inny obraz. W przyszłości technika ta zostanie zastąpiona przez wyświetlacze autostereoskopowe, które będą wysyłać do każdego oka odpowiedni obraz, eliminując konieczność zakładania specjalnych okularów.

Przeływ informacji



Teleimersja wymaga dużych mocy obliczeniowych i szerokopasmowej sieci.

Od nadawcy...

Procesory pracujące równolegle przetwarzają obrazy z kamer wideo w trójwymiarowy model sceny.

... do odbiorcy

Dane z modelu wytworzonego po stronie nadawczej są podstawą wygenerowania obrazów, jakie powinno widzieć każde oko uczestnika. Proces jest powtarzany wiele razy na sekundę, aby uwzględnić ruchy głowy.

Tworzenie trójwymiarowego obrazu

1. Zestaw kamer wideo obserwuje uczestników i ich otoczenie z różnych punktów. Każda kamera rejestruje wiele obrazów na sekundę.



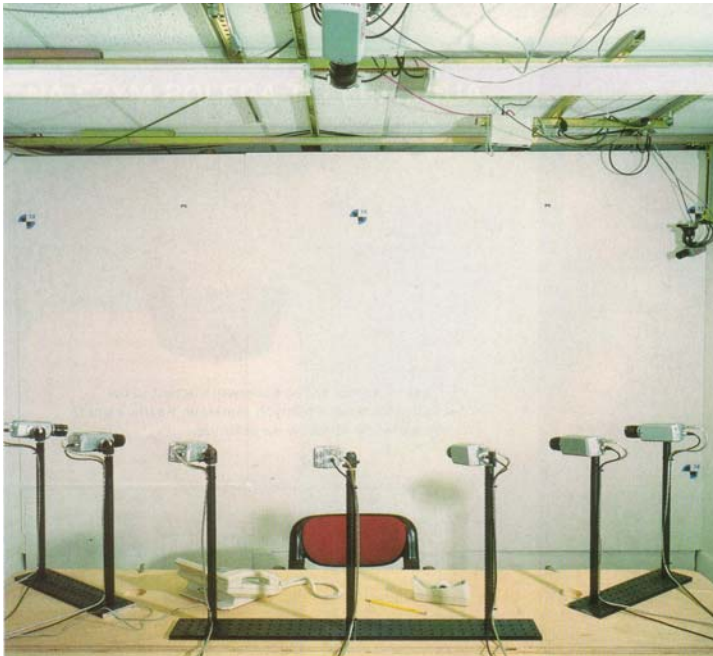
2. Obrazy są grupowane w pokrywające się częściowo podzbiory trzech obrazów.



3. Dla każdej trójki komputer oblicza “mapę niezgodności” odzwierciedlającą różnice między obrazami z kamer. Różnice te są przeliczane na odległości od płaszczyzny odniesienia. Na tej podstawie tworzy się “płaskorzeźbę” obserwowanej sceny.



4. Powyższe dane są podstawą do stworzenia chwilowego modelu trójwymiarowego całej sceny, który już nie zależy od punktu obserwacji. W trakcie tych obliczeń można zredukować szum i przekłamania.



Siedem kamer filmuje uczestnika sesji teleimersyjnej w Chapel Hill.

Poza systemem podglądania sceny głównymi elementami teleimersji są komputery, sieć, wyświetlacze i inne urządzenia służące do komunikowania się z użytkownikiem. Każdy z tych składników został udoskonalony i wymaga dalszych dopracowań. Teleimersja zachłannie używa mocy obliczeniowej komputerów. Gdy tylko było to możliwe, stosowano komponenty, które są wykorzystywane w sprzęcie powszechnego użytku, aby przybliżyć w ten sposób dzień, kiedy teleimersja wyjdzie poza laboratorium. Dziś używa się dziesiątek takich procesorów w każdej stacji teleimersji. Są one w pecetach stojących rzędami na półce albo w płytach z układami elektronicznymi wmontowanymi w stojaki wielkości chłodziarki. Większość mocy obliczeniowej poświęca się zbieraniu danych. Morze kamer to

zespół nakładających się na siebie trójek wideokamer. Obecnie najczęściej używa się siedmiu kamer do filmowania jednej osoby siedzącej za biurkiem – są one tak skonfigurowane, że działają jak zestaw pięciu trójek. Można powiedzieć, że osiem dwugigahercowych procesorów

Pentium ze wspólną pamięcią powinno obsłużyć trójkę kamer w czasie rzeczywistym. Tego typu zestawy prawdopodobnie ukażą się na rynku przed końcem roku. Mimo spadku cen komputerów, tak jak w ostatnich dziesięcioleciach, trzeba będzie jeszcze poczekać, zanim teleimersja stanie się na tyle tania, by wejść do powszechnego użytku. Ocenia się, że koszt zestawu ośmioprocessorowego wyniesie 30–50 tys. dolarów, a niezbędnych będzie kilka w każdym stanowisku teleimersyjnym (jeden zestaw na każdą trójkę kamer). Potrzebne będą też procesory do pozostałych obliczeń. Nie wiadomo jeszcze, ile kamer wymagać będzie konkretne zastosowanie, ale można oceniać, że siedem to absolutne minimum do prowadzenia nieformalnej rozmowy, a 60 do obrazowania najbardziej skomplikowanych scen, takich jak zdalny pokaz operacji chirurgicznej, konsultacje lub szkolenia. Przetwarzanie danych z kamer wideo to nie wszystko. Potrzeba dalszych procesorów do obliczenia, jak ma wyglądać scena, gdy na przykład użytkownik rusza głową. Na początku używano dużego, wyspecjalizowanego komputera do zastosowań graficznych, ale niedawno udało się przejść na zwykłe procesory z tanimi kartami graficznymi – używa się ich po jednym do generowania obrazu dla każdego oka. Potrzebne są także dodatkowe procesory do innych zadań, takich jak integracja danych z trzech kamer, generowanie niedostrzegalnego światła strukturalnego, obliczenia dotyczące położenia głowy użytkownika, przetwarzanie danych interfejsu użytkownika i symulowanie obrazów wirtualnych przedmiotów. Istotą teleimersji jest minimalizacja czasu oczekiwania (opóźnień), to będą konieczne duże moce obliczeniowe, aby móc stosować algorytmy predykcyjne. Sygnały rozchodzą się we włóknie optycznym z prędkością około 2/3 prędkości światła w próżni. Tak więc sygnał potrzebuje 25–50 ms, by przebiec światłowodem przez całe Stany Zjednoczone. Do tego dochodzą inne opóźnienia spowodowane na przykład przełączaniem sygnałów przez rutery. Niektóre ważne aspekty wirtualnego świata wymagają, by opóźnienia nie przekraczały właśnie 30–50 ms. Większe powodują zmęczenie i dezorientację, utratę poczucia przestrzeni, a w najgorszym

przypadku – nudności. Nawet posługując się najszybszymi komputerami, używa się algorytmów predykcyjnych, aby skompensować efekt opóźnień spowodowanych przesyłaniem sygnału na duże odległości. Z tych właśnie powodów obecnie testuje się system tylko pomiędzy stacjami znajdującymi się na wschodnim wybrzeżu USA.

Jednym z obiecujących kierunków badań przez następnych parę lat jest przesyłanie danych z teleimersji do dużych superkomputerów i uzyskanie w czasie rzeczywistym dostępu do dużych mocy obliczeniowych. W takim przypadku superkomputer musi być tak potężny, aby skompensować dodatkowe opóźnienia spowodowane przesyłaniem danych do i z miejsca, w którym się znajduje. Najważniejszym czynnikiem jest przepustowość sieci. Wymagania zależą od złożoności sceny i od rodzaju jej zastosowania – im bardziej złożona scena, tym większa przepustowość. Można założyć, że większość sceny, a w szczególności tło, nie zmienia się i nie trzeba przysyłać informacji o niej w każdym obrazie. Dane opisujące człowieka siedzącego za biurkiem, bez szczegółów otoczenia, przesyłane z niewielką prędkością dwu klatek na sekundę, wymagają przepustowości sieci rzędu 20 Mb/s, ale czasem trzeba nawet 80 Mb/s. Jednak z czasem wymagania te zmniejszą się, gdyż wprowadzone zostaną lepsze techniki kompresji danych. Każda stacja teleimersji uczestnicząca we wspólnej sesji musi otrzymywać dane od wszystkich innych, tak więc wymagania należy odpowiednio zwielfokrotnić. “Ostatni kilometr” połączenia sieciowego, doprowadzający dane do ośrodków komputerowych na wydziałach informatyki w USA, to dziś na ogół linia OC3 o przepustowości 155 Mb/s – co akurat wystarcza do ustanowienia trójstronnej sesji przy niewielkiej częstotliwości obrazów. Jednak linia OC3 ma stokrotnie większą przepustowość od linii, które dziś uważamy za szybkie i kosztuje odpowiednio drożej. W nadchodzących latach może uda się stworzyć taką wersję teleimersji, która nie będzie wymagać okularów polaryzacyjnych ani żadnych innych urządzeń. Ken Perlin z New York University zbudował prototyp takiego wyświetlacza autostereoskopowego. Obecnie teleimersja jest 100 razy droższa od innych form telekomunikacji, a poza tym wymaga udoskonaleń. Wszyscy mają nadzieję, że w ciągu najbliższych pięciu lat na tyle zostanie dopracowana i stanie się, że będzie jej można używać w ograniczonym zakresie, a do powszechnego użytku wejdzie za lat 10. Gdy teleimersja już się upowszechni, zapewne znajdzie się dla niej wiele ważnych zastosowań. Zespoły inżynierów będą współpracować na duże odległości

przy komputerowym projektowaniu nowych urządzeń, którymi da się manipulować, tak jak prawdziwymi modelami we wspólnej pracowni. Archeolodzy z całego świata będą mogli zdalnie uczestniczyć w najważniejszych etapach prac wykopaliskowych.

Wybitni eksperci budowlani lub przemysłowi będą odwiedzać miejsca zagrożeń lub katastrof, nie tracąc czasu na dojazd. Możliwe, że teleimersja zacznie być postrzegana jako konkurencja dla podróży lotniczych – inaczej niż wideokonferencja. Mimo że w bliskiej przyszłości mało kto stwierdzi, że teleimersja jest równie dobra jak fizyczna obecność, to może ona zastąpić spotkania biznesowe, konsultacje zawodowe, doksztalcanie, pokazy na targach itp. Za jakieś 10 lat sesja teleimersyjna w dużym stopniu wyprze podróże służbowe. I to nie tylko dlatego, że będzie tańsza i lepszej jakości, lecz także dlatego, że coraz większa liczba lotów zmniejszy ich bezpieczeństwo, zacznie brakować terenów i zwiększy się zanieczyszczenie środowiska. Prawdopodobnie teleimersja zmieni także stosunki międzyludzkie. Na przykład ci, którzy zastanawiali się, z czego będą żyć artyści w czasach, gdy coraz trudniej egzekwować prawa autorskie (wskutek coraz powszechniejszego kopiowania ich utworów przez Internet), często proponowali, że wyjściem są płatne osobiste prezentacje, gdyż bezpośredni kontakt z twórcą ma większą wartość niż zapis na wideo. Dzięki teleimersji prezentacje takie staną się łatwe, tanie i przysporzą dochodów artystom. Warto pamiętać, że przed XX wiekiem wszystkie sztuki były “interaktywne”. Muzycy mieli bezpośredni kontakt z audytorium, podobnie aktorzy, a nawet poeci recytujący swe dzieła. Teleimersyjne wersje muzyki, poezji i dramatu, w których ważne są bezpośredni kontakt i intymność, mogą stać się odpowiedzią na kryzys w egzekwowaniu praw autorskich. Bez wątplenia teleimersja ma także wady. Jej pierwsi użytkownicy niekiedy zastanawiali się, czy nie odsłania ona zbyt dużo; telefony i wideokonferencje pozwalają uczestnikom lepiej kontrolować, co pokazują innym – na przykład na rozłączenie się lub wyjście z pola widzenia kamery. „Jestem przekonany, że w miarę nabierania doświadczenia odkryjemy właściwe interfejsy użytkownika (takie jak np. wirtualne lustro) oraz zasady zachowania, które rozwiążą problemy tego typu. Ta nowa technologia z pewnością będzie wywierać duży wpływ na społeczeństwo”- Jaron Lanier. Teleimersja to narzędzie ułatwiające

kontakty międzyludzkie, technologie komunikacji międzyludzkiej dają nam możliwość odczuwania empatii, tak więc są czymś, co wspiera moralne zachowania.

Jaron Lanier, informatyk, często nazywany jest ojcem rzeczywistości wirtualnej. Poza tym to specjalista od programowania wizualnego, technik symulacji oraz zastosowań sieci szerokopasmowych. Pełni funkcję dyrektora ds. naukowych w Advanced Network and Services, fundacji z siedzibą w Armonk (w Stanie Nowy Jork), która daje fundusze i lokale na rozwój konsorcjum Internet2.

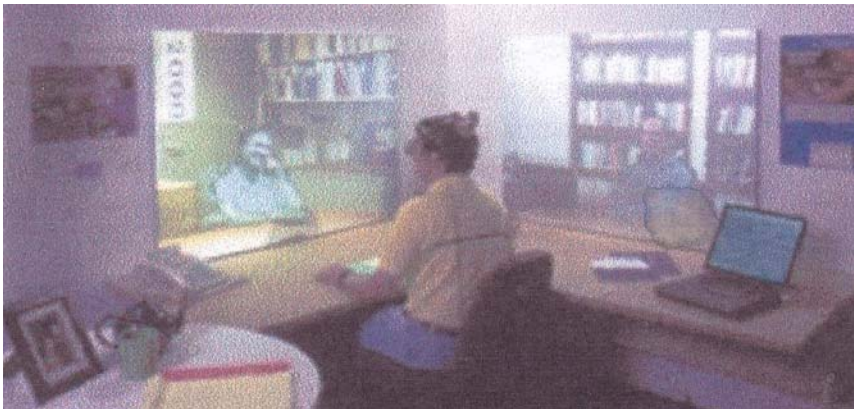
III. Tele – imersja: Telekonferencje jutra

Transmisja rzeczywistych środowisk w rzeczywistym czasie może stać się najbardziej wizjonerskim zastosowaniem Internetu 2.

Kiedy osiągi sieciowe na poziomie Internetu 2 staną się rzeczą zwykłą kiedyś w przyszłości, jakie aplikacje najlepiej wykorzystają ten milowy krok w szerokości pasma? Obecnie wykorzystuje się już takie zaawansowane zastosowania jak laboratoria wirtualne, biblioteki cyfrowe czy studiowanie na odległość. Jaron Lanier, który był jednym z przewodników w pracach nad rzeczywistością wirtualną w latach 80-tych, prowadzi teraz próby zatwierdzenia sieci jutra przy użyciu rodzącej się technologii znanej jako tele-imersja: przesyłania na dużą odległość trójwymiarowych scen rzeczywistych rozmiarów, które są poprawnie sampilowane i wyświetlane w czasie rzeczywistym przy użyciu zaawansowanej grafiki komputerowej i technik wizyjnych. Takie kopiowanie obrazów codziennej rzeczywistości w dużych ilościach powinno doprowadzić do bardziej naturalistycznych środowisk telekonferencyjnych (mniej podróży w interesach), większej wierności odtwarzania w przekazywaniu wiadomości i wydarzeń rozrywkowych (wysoka jakość będzie tak łatwa do osiągnięcia jak dzisiaj niska rozdzielczość), a nawet do tele-obecności w odległych miejscach podobnej do teleportacji ze Star Trecka.

Po trzech latach starań Lanier i jego współpracownicy przedstawiają pierwszą wersję demonstracyjną dla udowodnienia swojej koncepcji – wirtualne spotkanie trzech osób, które sprawia, że dzisiejsze wideokonferencje wyglądają jak filmy z taśm ośmiomilimetrowych sprzed lat. Na Uniwersytecie Północnej Karoliny w

Chapel Hill, zaproszeni goście mogą zobaczyć wyświetlane na dwóch ścianach w rzeczywistym czasie i rozmiarach obrazy naukowców siedzących przy swoich biurkach na Uniwersytecie Pensylwanii w Filadelfii i w Armonk w Nowym Jorku, w laboratorium sponsora projektu Advanced Network & Services, gdzie Lanier jest głównym naukowcem tego przedsięwzięcia. Uczestnik konferencji z Północnej Karoliny znajduje się w czymś, co zostało nazwane “tele-kabiną”, ma założone okulary polaryzacyjne i srebrne urządzenie śledzące na głowie i może się poruszać dookoła i widzieć wygenerowany przez komputer trójwymiarowy stereoskopowy obraz pozostałych dwóch uczestników telekonferencji, w którym zawartość wizualna bloku przestrzeni otaczającej górną część ciała uczestnika i elementów jego miejsca pracy jest zasadniczo odtwarzana przez grafikę komputerową. Dzięki temu otrzymujemy bardziej pełnowymiarowe i ściśle przedstawienie rzeczywistych środowisk, niż jest to możliwe przy użyciu istniejącej technologii wideo. Mimo że wersja demonstracyjna jest daleka od doskonałości, przesyłanie danych odbywa się w dwie strony zamiast w trzy i widoczne są zakłócenia, pokazuje ona to, co Lanier określa jako “maksymalną zbieżność rzeczywistego świata i grafiki komputerowej”.



Naukowiec z Uniwersytetu Północnej Karoliny spotyka się wirtualnie z kolegami z Pensylwanii i Nowego Jorku w rzeczywistych rozmiarach i w trójwymiarze. Wyświetlanie tele-imersji w rzeczywistym czasie osiąga się dzięki zaawansowanym technikom wizji komputerowej, graficznym i sieciowym.

Narodowe Przedsięwzięcie Tele-imersji (NNTI), pod taką nazwą znany jest ten trwający wciąż projekt, został zaproponowany przez Allana H. Weisa, założyciela Advanced Network & Services, jednej z firm, które stworzyły pierwotny szkielet Internetu. Dzięki zyskom ze sprzedaży firmy firmie America Online, Advanced

Network & Services stało się instytucją badawczą, fundującą prace nad przodującymi zastosowaniami technologii sieci komputerowych. Potem firma zatrudniła Laniera i dano mu zespół współpracowników w Armonk oraz budżet dla zapewnienia dotacji dla naukowców. Głównymi badaczami są Henry Fuchs z UNC (Un. Płn. Karoliny), Andries van Dam z Uniwersytetu Brown i Ruzena Bajcsy z Uniwersytetu Pensylwanii. W projekt zaangażowani byli również naukowcy z Podyplomowej Szkoły Morskiej, Uniwersytetu Carnegie Mellon, Uniwersytetu Columbia, Uniwersytetu Illinois w Chicago oraz Uniwersytetu Południowej Karoliny.

Pokonywanie Zagadnień Technicznych

“Jest dość dużo trudności”, mówi Lanier, “Pierwszą jest jak wyczuć odległe miejsce w czasie rzeczywistym wystarczająco szybko i z taką jakością, żeby można było ponownie je wyświetlić i żeby to dobrze wyglądało. Pojawia się zestaw problemów z wizją, problemów graficznych i sieciowych i to wszystkie ze sobą połączone. Poza tym pojawia się problem jak stworzyć konfigurację widoku fizycznego, która będzie współdziałać z iluzją rzeczywistości”. Zaczynając od zdobycia sceny całkowitego trójwymiarowego wyobrażenia niezależnego od jakiegokolwiek perspektywy, Lanier wraz ze swoim zespołem optowali na rzecz technik wizyjnych wykorzystując “morze” kamer wideo. Dla jak najmniejszych strat na jakości na rzecz wydajności zastosowali zachodzące na siebie trójki kamer wideo z większymi marginesami informacji o przedstawianym obrazie, aby zapewnić pełniejsze pokrycie powierzchni wizualnych niż przy dwójkach kamer. W zaawansowanym zastosowaniu telekonferencyjnym ustawia się siedem kamer wideo w łuku 120-stopniowym przed każdym uczestnikiem, gdzie kamery są sampłowane w nachodzących na siebie triadach - optymalne ustawienie, jeśli weźmiemy pod uwagę ograniczenia sieciowe. Celem Laniera są obrazy o wyższej rozdzielczości z zestawem 60 kamer, który może być używany w tele-szkoleniach medycznych.



Przy użyciu zestawu 7 kamer naukowcy nagrali wizualną i wymiarową “testową” zawartość sceny, na którą składał się manekin człowieka i szachownica jako wzór kalibracyjny.

Używając standardowych elementów, tak, aby inni użytkownicy mogli zastosować tele-immersję, zespół NTII wybrał laboratoryjne cyfrowe kamery wideo

Sony z szybkimi złączami standardu IEEE 1394, które są też sprzedawane przez Apple Computer pod nazwą FireWire. “Chcemy, aby w końcowym efekcie do tele-immersji zastosować można było standardowe konfiguracje”, wyjaśnia Lanier. “Jednym z następnych zadań będzie napisanie książki kucharskiej, w której udokumentujemy to, co robimy, aby inni mogli zrobić to samo i na tym budować.”

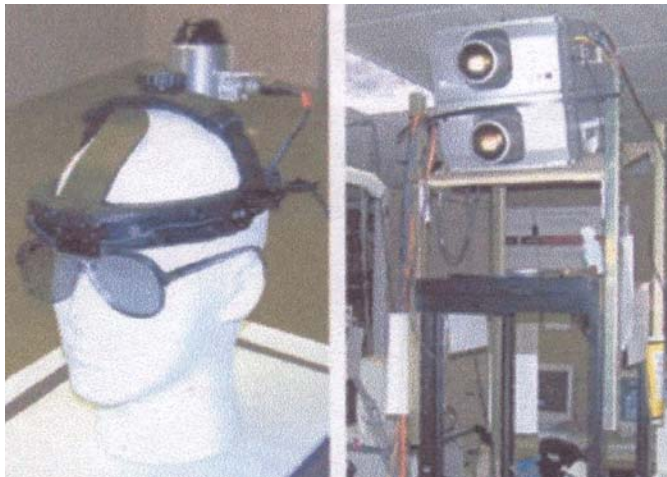
Inne Przeszkody

Jednym z problemów graficznych, który trzeba przezwyciężyć są niejasności powierzchni, z którymi mózg ludzki radzi sobie bez problemu, a z którymi komputery mają duże trudności. W normalnym oświetleniu pomieszczenia, takim jak sufitowe żarówki fluorescencyjne czy biurkowe lampy jarzeniowe, na pustej ścianie nie są wyświetlane żadne tekstury powierzchni, co sprawia, że programy do rozpoznawania wzorów gubią się w pracy.

Aby właściwie rejestrować puste, gładkie ściany, a także obiekty błyszczące, zespół prowadzi badania nad techniką stworzoną na UNC nazywaną “niedostrzegalne światło strukturalne”. Przy użyciu tej technologii oraz istniejącego oświetlenia pomieszczenia, scena jest oświetlona przez coś, co wydaje się być zwykłymi reflektorami, które jednak zawierają w swoim oświetleniu strukturalnie

monochromatyczne wzory geometryczne. Wzory te drgają, aby były niedostrzegalne dla ludzi, ale zsynchronizowana kamera wideo jest w stanie odbierać te wzory kalibracji wizualnej, tak, aby wyeliminować dwuznaczności koloru, kształtu i współczynnik odbicia obiektów (takich jak ekrany, puste ściany, przejścia, a nawet czoło człowieka).

Wspomniany powyżej cel uzyskania sceny i przekazu w czasie rzeczywistym podkreśla wyzwania graficzne i sieciowe. Na przykład, szybkości samplowania są różne w zależności od zawłości i ruchu w scenie, gdy algorytmy rozpoznawania wzorów tworzą model zawartości wizualnej przechwycony przez zestaw kamer wideo; dokładność danego algorytmu musi być oceniana w stosunku do opóźnienia w czasie, jakie może spowodować. Aby zmniejszyć opóźnienia obliczeniowe, dokonuje się pewnej optymalizacji, takiej jak dzielenie sceny na odcinki, tak aby można się było bardziej skupić na dokładnym przechwyceniu wyrazów twarzy. Rozpoznawanie cech to podstawowa technologia Eyematic, firmy z Los Angeles, w której Lanier jest również głównym naukowcem.



Wyświetlacz tele-imersyjny składa się z urządzenia śledzącego ruchy obserwatora zakładanego na głowę oraz z okularów polaryzacyjnych, które dają poczucie głębi z wyświetlanego widoku z oka prawego i lewego.

Trzymając się strategii użycia standardowych elementów i urządzeń, zespół używa również standardowych komputerów do przetwarzania informacji wizualnych. “Bierzemy systemy od najczęstszych “podejrzanych” – Dell i IBM”, mówi Lanier. “Zapelniamy nimi półki i każemy im obliczać rozwiązania. Dobrą zasadą jest mieć jeden szybki zestaw procesorów na kamerę” Najbardziej pracochłonną pracą w tym projekcie było skracanie czasu obliczania danych przez procesor. “Strasznie

męczące jest dochodzenie do tego, dlaczego dane sterowniki są wolniejsze niż według nas powinny być albo dlaczego pamięć nie oczyszcza się tak szybko jak powinna.”

Inne zagadnienia przetwarzania wizji w tele-imersji to poszczególne ograniczenia sieci, takie jak szerokość pasma, opóźnienia i protokoły. W różnych sytuacjach trzeba określić jaki rodzaj informacji najlepiej jest przesyłać przez sieć. Mogłyby to być surowe dane z poszczególnych kamer cyfrowych, skompresowane dla szybszego przesyłania lub połączone sygnały z zestawu kamer, które są przetwarzane dla podstawowych obliczeń parametrów sceny i wysyłane jako dane trójwymiarowe. Wybór będzie zależeć od czynników konfiguracji, w tym od ilości kamer. “Chcemy dać ludziom formułę na to, jak idealnie skonfigurować system dla danej sieci”, wyjaśnia Lanier. “To również będzie zawarte w naszej książce kucharskiej.”

Kiedy informacja wizualna jest otrzymana w odległym miejscu, jest ponownie wyświetlana jako wygenerowani komputerowo ludzie i otoczenie przez komputer przeznaczony specjalnie do tej operacji. W przypadku obecnej wersji demonstracyjnej, użyto systemu SGI Onyx 2 Reality Monster, wykorzystywanego do najbardziej wymagających symulacji rzeczywistości wirtualnej, ale Lanier chce, aby wkrótce wyświetlanie mogło być dokonywane również przez tańsze systemy. “Sprawdzaliśmy pewne rozwiązania z Pentium z kartami do gier, więc tanie wyświetlanie nie jest tak odległe”, dodaje. Podczas tegorocznej konferencji Siggraph, był pod wrażeniem zestawu zwykłych Sony Play Station z dodatkową ilością pamięci, która mogła podołać temu zadaniu. A jednak nadal pozostają problemy w znalezieniu najlepszych technik do przedstawiania różnych rodzajów obiektów; na przykład, włosy są lepiej wyświetlane jako chmura punktów, a skóra jako wielokąty z teksturami.

Tele-Rzeczywistość

Według Laniera wrażenie i użyteczność tele-imersji są dość różne od wideokonferencji. “Jeśli właściwie wyświetla się ludzi, ma się wrażenie, że są oni prawdziwi. Jest to ich “komputerowa wersja” – są trochę szklani, nie tacy wypełnieni jak w rzeczywistości, ale i tak ma się dość silne wrażenie ich obecności, można nawiązać kontakt wzrokowy, przekazać swój nastrój i odczytać ich nastrój, a

to wszystko dlatego, że są to obrazy trójwymiarowe, mające rzeczywiste wymiary, a nie małe, płaskie obrazy wideo.”

Aby zwiększyć wrażenie rzeczywistości, stereoskopia jest konieczna do odczucia głębi. Obecnie używany w tele-imersji system do osiągnięcia tego wrażenia składa się z pary projektorów przednich na każdą wyświetlaną scenę i z okularów ze spolaryzowanymi soczewkami używanymi przez uczestników w celu oddzielenia na nowo stworzonego widoku prawego i lewego oka. Lanier jest świadomy, że jest to bardzo nienaturalna technologia dla użytkownika, ponieważ zmienia typowy wygląd wielu uczestników. “Omawialiśmy eliminację okularów w obliczeniach, ale nigdy nie poszliśmy za tym”, mówi “ponieważ ostatecznym rozwiązaniem będą autostereoskopowe ekrany, które nie będą wymagały zakładania okularów.” Jednym z takich systemów, którego prototyp zostały stworzony w Media Research Lab Uniwersytetu Nowojorskiego przez Kena Perlina, zostanie włączony w wersji demonstracyjnej następnej generacji NTII.

Właściwe ustawienie punktu widzenia uczestnika telekonferencji dla odpowiedniego ponownego wyświetlania oddalonych scen wymaga teraz użycia kolejnego dziwnego urządzenia: stworzonego na UNC urządzenia śledzącego opartego na urządzeniu śledzącym HiBall 3rd Tech, które “siedzi” na głowie użytkownika jak srebrna solniczka. Ustanawia orientację fizyczną patrzącego przez wyczuwanie pozycji w stosunku do zamieszczonych na suficie podczerwonych diod emitujących światło. Lanier oczekuje, że obniżenie kosztów wynikające z Prawa Moora będzie dotyczyło również procesorów wideo, co w rezultacie doprowadzi do tego, że śledzenie pozycji będzie dokonywane przez czujniki rozmieszczone jak tapeta w pokoju, dzięki czemu nie trzeba będzie zakładać niewygodnego urządzenia na głowę.



W najnowszej demonstracji tele-imersji naukowcy poruszają obiekt wirtualny, przecierając szlak dla wyświetlania w czasie rzeczywistym prototypów CAD/CAM i innych współpracujących aplikacji roboczych.

Obecnie wprowadzana jest do demonstracji wizja Laniera dotycząca wspólnych miejsc pracy w tele-imersji. W pierwszej konfiguracji uczestnicy zajmują oddzielne miejsca. W następnej wersji lokalni uczestnicy i odległe sceny będą na siebie zachodzić dla naturalnego przekazywania formuł na tych samych tablicach wirtualnych czy przy przemieszczaniu obiektów wirtualnych przez sieć. Na wykopaliskach archeologicznych, na przykład, miejscowi uczestnicy mogliby przekazać wirtualne wyobrażenie znaleziska, aby oddalony ekspert mógł je bliżej zbadać. Dyskutuje się również nad obiektami wirtualnymi o szczególnych cechach – jak “lampa prywatności” skonstruowana przez naukowca z Kolumbii, Stevena Feinera. Lampa ta to wirtualne źródło oświetlenia, które mogłoby wyznaczyć pewien obszar, którego zawartość byłaby zakryta przed jednym lub większą ilością uczestników, którzy nie mieliby odpowiedniego poziomu zwolnienia.

Mimo że wszystkie systemy opracowywane w NTII mają być skalowane, na dzień dzisiejszy ilości przestrzeni wizualnej, jakie są przesyłane w demonstracji – masy otaczające trzech siedzących uczestników – bliskie są ograniczeniom szerokości pasma, nawet w Ineternece 2. Spowodowane jest to wąskim „gardłem”, znanym jako problem ostatniej mili: Połączenia w miasteczkach studenckich i instytucjach badawczych innych niż centra super-komputerów (oraz firm i mieszkań prywatnych) nadal mają okablowanie na dość dużo mniejszym paśmie w stosunku do pełnego potencjału Internetu 2 i przedsięwzięć internetowych następnej generacji.

Kolejnym ograniczeniem możliwości rzeczywistego czasu tele-imersji jest fizyka samego światła. Fotony przemieszczające się przez światłowody poruszają się dużo poniżej podstawowej prędkości światła, powodując odczuwalne opóźnienia na dużych odległościach. W dużej różnorodności zastosowań oglądający są w stanie przyjąć opóźnienie rzędu 30 - 50 milisekund. W niektórych przypadkach, krytyczne elementy, takie jak ruch głowy i rąk mogą być wspomagane przez algorytmy prognozujące, które w efekcie przyspieszają przechwytywanie sceny poza dany moment w czasie, przewidując następną pozycję w geście ręki czy ruchu głowy, ale one również mają ograniczenia. Może udałoby się wyeliminować opóźnienia używając super-komputerów. Jednak na dziś maksymalną odległością dla uzyskania

wydajnej tele-imersji jest mniej-więcej szerokość Stanów Zjednoczonych. Lanier uważa, że techniki tele-imersji, które dzisiaj prezentuje, są fundamentami do budowy biura jutra, w którym kilku użytkowników z całego kraju będzie mogło współpracować tak, jakby byli w tym samym pomieszczeniu. Gdyby obrazy były skalowane, można byłoby przysyłać większe sceny, takie jak konferencje prasowe, przedstawienia baletowe czy wydarzenia sportowe. Gdyby zestawy kamer zamienić ze stojących w przenośne, oglądający mogliby dokonać tele-obecności w miejscach bardzo odległych lub niebezpiecznych. Jaron Lanier chce nie tylko zastosowania tele-imersji w Internecie następnej generacji – oczekuje, że fundamentalnie zmieni ona nasze patrzenie na świat rzeczywisty i wirtualny.

IV. SOFT (oprogramowanie dla teleimersji); BAMBOO(platforma i język)

SOFT – strukturalne oprogramowanie dla teleimersji.

Wirtualne środowiska (trójwymiarowe zespoły oprogramowania) borykają się z odległym i powracającym problemem. Problem pojawia się na początku, tak jakby każde badanie i każda firma tworzyła dynamiczną wirtualną rzeczywistość po raz pierwszy. Zostało opublikowanych i przetestowanych wiele zagadnień o urządzeniach dla użytkowników rzeczywistości wirtualnej i o technikach wizualizacji danych, ale rzadko można je zaobserwować. Pomysły w różnych laboratoriach rzadko są łączone i testowane. Na przykład czy ktoś używa idei od Sketcha, aby utworzyć startowe warunki dla symulacji zachowania się konstrukcji drewnianych pod dużym działaniem środowiska zewnętrznego(trzęsienie ziemi). Jest kilka takich symulatorów do tego typu problemu ale nie są one używane tak często jak powinny z powodu problemów z wprowadzaniem danych. W obecnym obszarze badań, nie ma wystarczającej ilości czasu na obmyślanie nowych narzędzi i na korzystanie z sugestii użytkowników. Kolejny przykład: były przeprowadzane kliniczne próby, aby zademonstrować zalety chirurgicznych symulacji, ale nie ma jeszcze chirurgicznego przypadku użycia takiej symulacji przez kogoś innego niż wynalazców. Te narzędzia są takie niezwykle i niepodzielne, że mogą tylko przetrwać w warunkach laboratoryjnych. W społecznościach naukowych, badacze mogą korzystać z wzajemnych prac, ponieważ zostały ustalone pewne stałe określające środowisko zastosowania. Na przykład społeczności UNIX można było używać narzędzi modułarnych jak GREP, bez bycia jedynymi z wynalazców. Praca wszystkich jest połączona poprzez polecenia tekstowe i pliki. Gdyby społeczność wirtualnej rzeczywistości mogła dzielić się swoimi doświadczeniami, wtedy zaczęto by pracować jako społeczność a nie jako pojedyncze laboratoria. W taki sposób dokonuje się rozwoju w sposób wykładnikowy a nie liniowy. W oprogramowaniu

wirtualnej rzeczywistości istnieją pewne platformy pośredniczące, ale nie zostały bliżej zbadane. Można stworzyć standard dla statycznej siatki geometrycznej i mieć nadzieję, że będzie to tym dla VR co text dla UNIX. Istnieje VRML 1 i nie działa to w taki sposób, a statyczna geometria nie jest największym problemem. Można znaleźć standardowy sposób opisywania zachowania i zmusić każdego do przestrzegania go. Problem w tym, że albo kończy się na wąskim zakresie zachowań (VRML2), lub wymaga aby każdy wierzył w jednokierunkowy rozwój środowiska. Jedną szczególnie trudną rzeczą w środowiskach wirtualnych jest to aby eksperymentować ze zmiennymi pomysłami na temat ciągłej kontroli, rzeczywistej interakcji i innymi rzeczami dziejącymi się natychmiast, trzeba zmieniać najbardziej zewnętrzne i wewnętrzne pętle. Wszystkie poprzednie grupy badawcze były w stanie związywać najbardziej zewnętrzne pętle i dzielić się wewnętrznymi modułami. Jedynie SOFT jest w stanie zjednoczyć siły naukowców i społeczności wirtualnej rzeczywistości.

Jaron Lanier

Motywacja

Wprowadzenie w życie standardowej telekabin

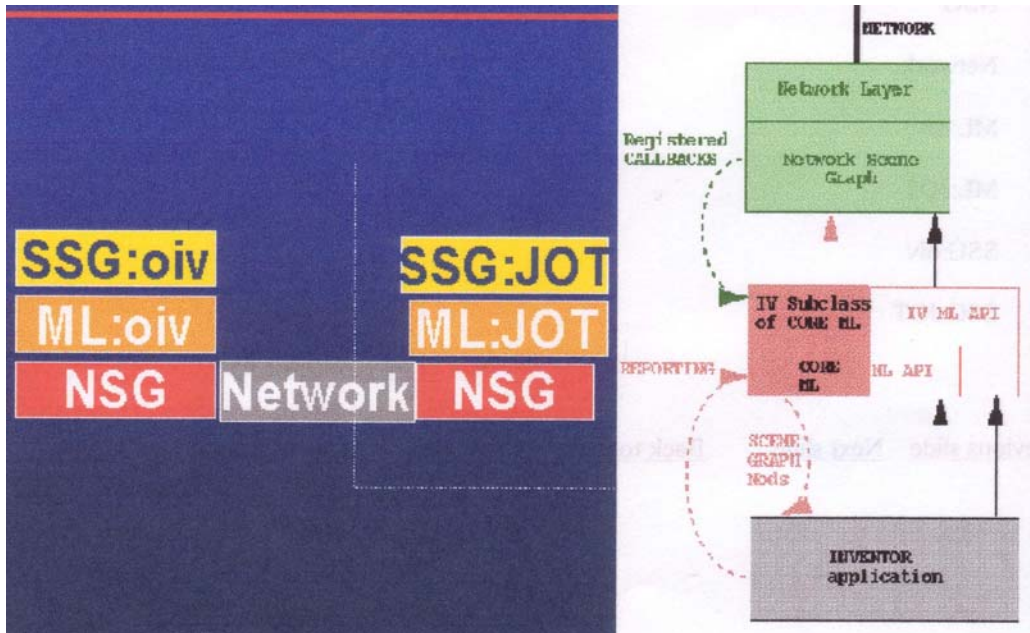
- połączenie prawdziwych i syntetycznych światów
- sieć – umożliwia istnienie i przyszłe trójwymiarowe zastosowania

Rozwój rozmieszczonych platform wirtualnej rzeczywistości

- możliwość łatwego współdziałania

łatwość w dzieleniu się zastosowaniami

Architektura SOFT



Określenie struktury wspólnych danych: węzeł

- Tag- co znaczy „węzeł”
- Field list – SG dane definiujące „węzeł”
- Zawartość kodu – dodatkowe dane do transmisji

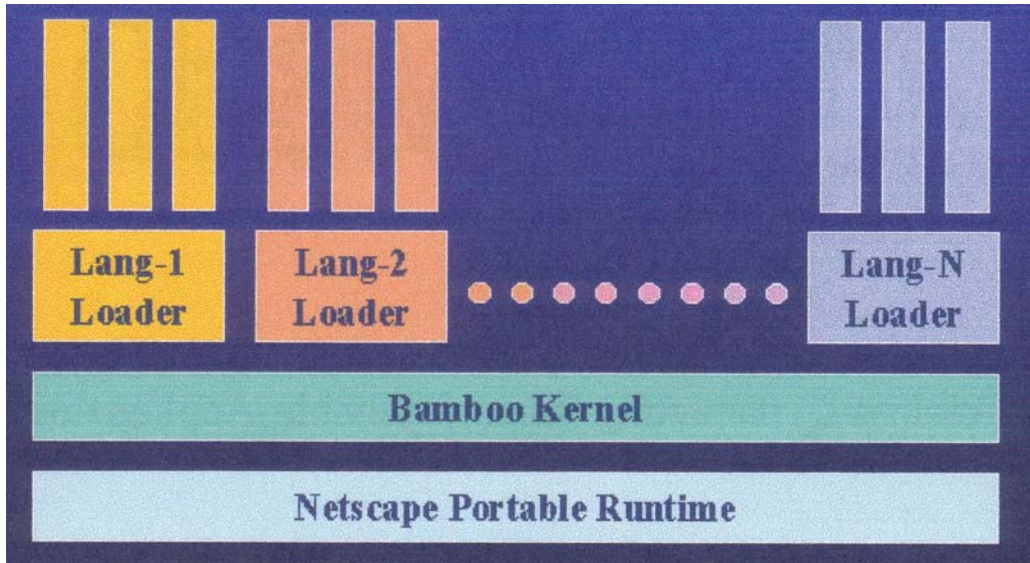
Wprowadzenie podziału struktury danych:

- Protokół transmisji
- Zawartość danych
- Zarządzanie sesją

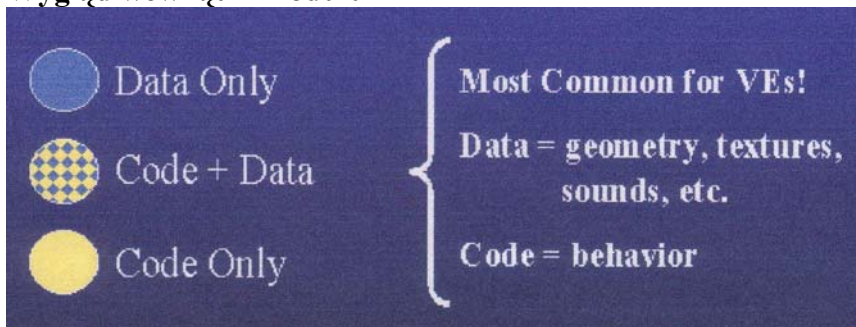


BAMBOO - platforma i język niezależny mechanizm umożliwiający dynamiczną rekonfigurację aplikacji

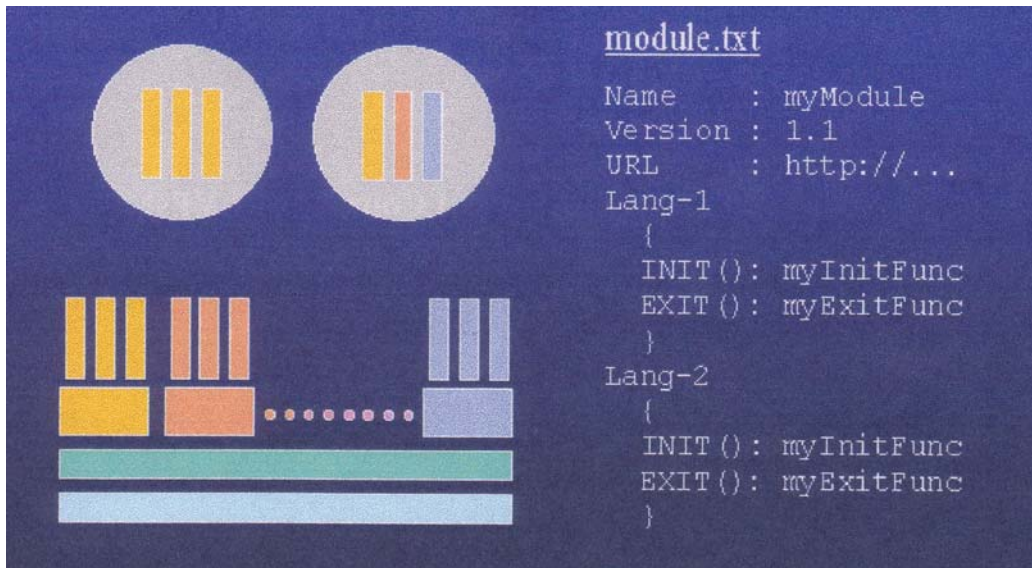
Działanie



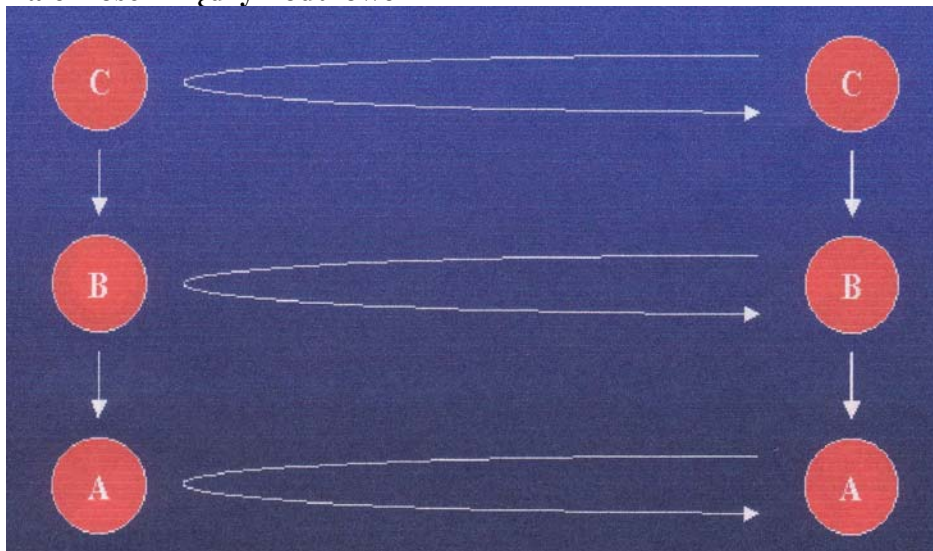
Wygląd wewnątrz modułu



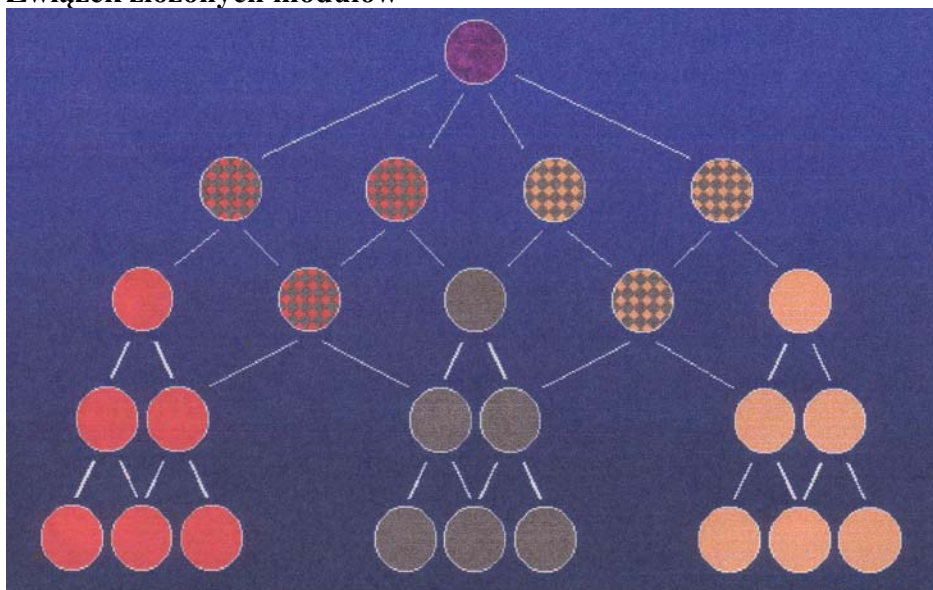
Sporządzanie mapy modułów na wejściu



Zależności międzymodułowe



Związek złożonych modułów



LSVE

Dynamiczne protokoły

- umożliwia nowym uczestnikom dołączyć do istniejącego środowiska

Obszar udziału zarządzania (AOIM)

- zarządzanie transmisją umożliwiające wysoką zdolność dzielenia

Trwałość/stałość

- możliwość stanu zapamiętywania w sieci

Bamboo –w skrócie

Bamboo

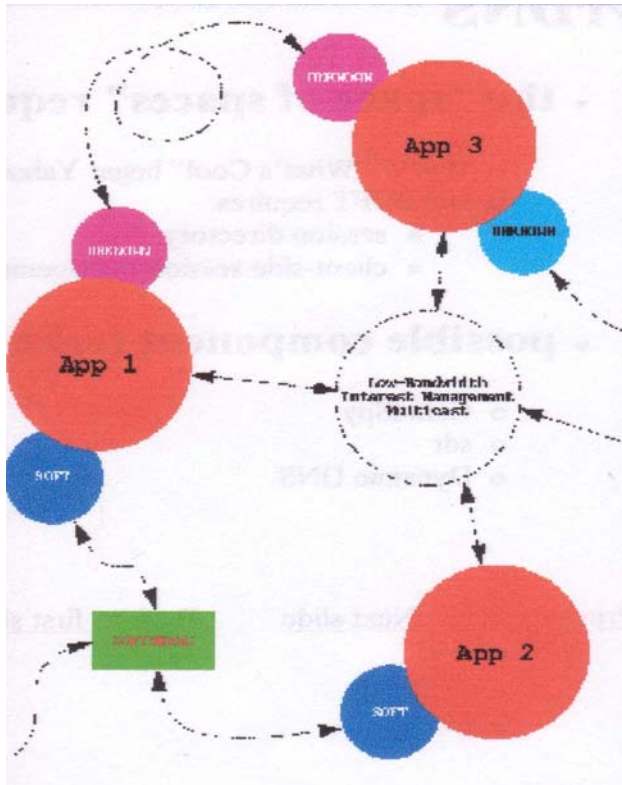
- otwarte źródło

NPSNET V

- narzędzie do określonych VE modułów Bamboo

ActiVE

- dynamicznie-rozciągliwy VE zbudowany z NPSNET V



SOFT utrzymuje dokładnie/blisko związane rozmiary podzielonego środowiska
... i ułatwia dostosowanie zapisu aplikacji graficznych przez wykres podziału sceny
... i SG – niezależny atut nadawania pierwszeństwa

Miejsce w przestrzeni:

- SOFT wymaga:
 - katalogu sesji
 - zarządzania sesją po stronie klienta

Możliwe składowe technologie:

- **GameSpy**
- **Sdr**
- **Dynamiczne DNS**

Wybór protokołu Behavior

- **Opis ułożenia plików protokołu w XML.**
 - **równe połowy stosu Protokołu Transmisji Wirtualnej Rzeczywistości**
- **Możliwość dynamicznego załadownia nowego opisu protokołu XML w czasie wykonywania programu, pozwalając na:**
 - **szybszą produkcję protokołów**
 - **protokoły adaptacyjne, które reagują na przeciążenie sieci**
 - **ładowanie protokołu w czasie wykonywania programu**

V. Zastosowanie teleimersji.

1. Biuro Przyszłości i Teleimersja

Uniwersytet w Północnej Karolinie w Chapel Hill jest jedynym z Centrów Naukowo Technologicznych zajmujących się grafiką dla NSF(National Science Foundation), innymi partnerami są Brown University, Uniwersytet w Kaliforni, Cornell University i Uniwersytet w Utah. Katedra Informatyki w Chapel Hill przeprowadza badania istotne dla STC(Graphics and Visualization Science and Technology Center) w obszarze telekolaboracji - skalowanie/ skanowanie ultradźwiękowe i laparoskopowe. W przyszłości STC sponsorował badania nad monitorami noszonymi na głowie przed oczami nad telekolaboracją dla projektów mechanicznych i precyzyjnej architektury. Uniwersytet w Północnej Karolinie w Chapel Hill jest również centrum multimedialnej sieci łączącej 5 STC miejsc. To zapewnia symultaniczne połączenie video konferencje i przesyłanie danych między miejscami.

Biuro przyszłości

Pewnego dnia biura będą wyposażone w system komputerowego podglądu i grafiki, które będą pracowały w zespole aby dać idealny obraz i zdolności immersyjne.

Teleimersja

Praca nad tym odbywa się przy współpracy z Advanced Network & Services, Uniwersytetów Browna i Pensylwanii. Duża część pracy była rezultatem wspólnego zaangażowania w Narodową Incjatywę o Teleimersji.

Obecnie naukowcy są wspierani przez Narodową Fundację Naukową(NSF), Centrum Naukowo Techniczne Grafiki i Wizualizacji Komputerowej(STC). Datki w postaci sprzętu i podstawowego finansowania otrzymywane są z Intel Corporation i Laboratorium Mitsubishi.

Biuro przyszłości : Ujednolicone podejście do modelowania w oparciu o obraz i przestrzenne immersyjne wyświetlanie.

Wprowadzone zostały pomysły, technologie i początkowe rezultaty dla stworzenia Biura Przyszłości w oparciu o ujednolicone zastosowanie obrazu i grafiki komputerowej w systemie, który łączy i korzysta z pojęć the CaveTM, systemów wyświetlenia i modelowania w oparciu o obraz. Podstawą jest użycie technik rzeczywistego odtwarzania obrazu przez komputer i dynamicznego wydobywania głębi widocznej w biurze, powierzchni (ściany, meble, przedmioty, ludzie), następnie rzutuje się obraz na te przedmioty (używa się ich jak ekranów) i dostraja się ostrość lub interpretuje zmiany na powierzchni.

W pierwszym przypadku można użyć przedmiotów codziennego użytku o nieregularnych kształtach jako powierzchni projekcyjnych i rzutować grafiki o dużej rozdzielczości na te powierzchnie. W drugim przypadku można przysyłać dynamiczne/ ruchome modele za pomocą sieci do odległego miejsca. W końcu można zinterpretować ruch (zmiany dynamiczne) na ekranie w celu śledzenia ruchu, interakcji lub użycia powiększenia rzeczywistego obrazu. Aby zrozumieć symultaniczne uchwycenie obrazu i jego pokazywanie wyobrażamy sobie biuro przyszłości, gdzie światła sufitowe są zastąpione kontrolowanymi komputerem kamerami i inteligentnymi rzutnikami, które używane są do wychwytywania poruszających się modeli za pomocą niezauważalnych strumieni światła i do projekcji obrazu o dużej rozdzielczości na zaprojektowane do tego ekrany. Poprzez wykonywanie obu czynności jednocześnie można automatycznie dostroić i

dynamicznie modyfikować zmiany geometryczne, gęstościowe, rozdzielcze wynikające z nieregularnych lub zmieniających się powierzchni lub nakładających się na siebie obrazów projekcyjnych. Nowym podejściem do modelowania ruchu w oparciu o obraz jest użycie optymalnie rozszczepionego światła, które może wychwycić (zgodnie z szybkością pracy komputera) głębię i odbite światło. Systemy urządzeń nie są jeszcze niezauważalne, ale można zademonstrować przyszłe zmiany w laboratorium. Nowym zamierzeniem jest wyregulowanie obrazu na nieregularnych powierzchniach, zostanie zastosowane systemu dwu pasmowego strukturywania projekcji, aby gromadzić obrazy, tak aby przy obrocie głowy obracał się również obraz. Zaprezentowane będą tu początkowe urządzenia do ogólnego obserwowania, biurowy wystrój wnętrza, i początkowe pokazy naśladowania i obrazowania ruchu.



Rys 1 Rysunek pomocniczy dla rozróżnienia pomiędzy wyświetlanym obrazem i realnymi obiektami na rysunku.



Rys 2 Poglądowy rysunek biura przyszłości (realne obiekty na rysunku i wyświetlany obraz). Przez zastąpienie normalnych świateł biurowych projektorami, dzięki jednemu projektorowi można uzyskać dokładną kontrolę nad całym światłem w biurze. Z pomocą zsynchronizowanych kamer, geometrii

i odbicia informacji może być ona wychwytywana dla wszystkich z widzialnych powierzchni w biurze, tak aby jedna kamera mogła zarówno wyświetlać obrazy na powierzchni, nakładać je na siebie i interpretować zmiany na powierzchni. zaplanowano do pomocy w rozróżnieniu pomiędzy wyświetlanym obrazem i realnymi obiektami na rysunku.

Celem jest pragnienie wybudowania bardziej użytecznego systemu do teleobecności i telewspółpracy między oddalonymi osobami. Użyto „morza kamer” i niewidzialnych wiązek światła do wydobywania trójwymiarowego obrazu i odtworzenie go na odległość. Wszystkie te pomysły z definicji są futurystyczne (nie istnieje jeszcze całkowicie działający system). Obecnie zademonstrowane zostały rezultaty i nowe metody i pomysły możliwe do zrealizowania w przyszłości. W punkcie 2. zaprezentowane będą podstawowe komponenty systemu bez szczegółowego opisu urządzeń i podejście do utworzenia biura przyszłości. W punkcie 4. omówione będzie obecne oprzyrządowanie, w części 5. pracę jaką należy jeszcze wykonać i tematy przyszłych badań.

Interfejsy do telekolaboracji

Podczas gdy systemy telewspółpracy używające dwuwymiarowych „mówiących głów” i dzielonych białych tablic znacznie się rozwinęły na przestrzeni lat, nowy system komunikacji często ogranicza tę część współpracy, która odbywała by się między rozmówcami, gdyby rozmawiali twarzą w twarz w tym samym pokoju. Zidentyfikowano kilka zadań dla których system kamer przemysłowych dostarcza tylko ograniczonego wsparcia. Poza ograniczonymi rozwiązaniami, częścią problemu jest to, że użytkownicy są zmuszeni utrzymywać dwa oddzielne EGOCENTRA (pojęcie gdzie i z kim przebywają) – jedno na miejscu, a drugie u ich współpracowników. Uczestniczący muszą przemyśleć i dopasować zachowanie tak aby pasowało do transmisji teleimersyjnej. Jedną z alternatyw jest urządzić wspólne wirtualne środowisko, w którym użytkownik zakłada HMD-monitor noszony na głowie, znika ze świata rzeczywistego, i wchodzi do środowiska wirtualnego – znajduje się wewnątrz pokazu, tam na przykład może obserwować wirtualne obiekty wraz z dwuwymiarowymi awatarami swoich współpracowników. Najbardziej oczywiste problemy przy tego typu eksperymentach z interfejsami to problemy ergonomiczne – na przykład rozmiar, ciężar, ruchliwość, ograniczona rozdzielczość i ograniczone pole widzenia. Ponadto, dla immersyjnych HMD-sów wyłączenie człowieka z jego wygodnego rzeczywistego otoczenia może ograniczyć ich

zdolność do interakcji z innymi ludźmi i prawdziwymi przedmiotami w środowisku. Atrakcyjniejszą alternatywą jest ściągnięcie ekranu z głowy użytkownika, i zamiast tego użycie SID przestrzennego ekranu immersyjnego. Ekran ten otacza widza panoramą obrazów. Te obszary projekcyjne są zazwyczaj w rozmiarach całych pomieszczeń stąd mogą obsługiwać złożoną liczbę widzów, i są zazwyczaj używane w połączeniu z urządzeniami projekcyjnymi, które rzutują obraz z pozycji przedniej lub tylnej. Najprawdopodobniej najbardziej znanymi przykładami urządzeń SID do ogólnego zastosowania są CAVETM Cave Automated Virtual Enviroment, oraz powiązane systemy z „kafelkowymi” ekranami, Power Wall i Infinity WallTM. Jest kilka przykładów udanej telekolaboracji, gdzie użytkownicy widzą i współpracują ze swoimi oddalonymi współpracownikami przy pomocy systemu CAVETM lub jemu podobnych. Takie gdzie jest zarówno miejsce na ekrany jak i na urządzenia projekcyjne. Naukowcy chcą, aby ten system działał na wzór telefonu –SID wbudowane w życie biura. Taki wysiłek nie zwróciłby się w zyskach jeśli używany byłby tylko do telekolaboracji, jeśli jednak byłoby to wysokiej jakości można by tego używać do codziennej dwuwymiarowej pracy przy komputerze, video, i obrazach immersyjnych trójwymiarowych. Jednak konstrukcja takich SID nie tylko wymaga bardzo ostrożnej budowy i zmontowania, ale jeszcze pewne dane charakterystyczne zmieniają się w zależności od czasu i czynników środowiska jak temperatura lub wibracje. Takie zmienne w czasie czynniki obejmują wewnętrzne i zewnętrzne parametry projektora, równowagę intensywności i koloru, wyrównywanie brzegów i zespalenie obrazu. Te problemy są najczęściej rozwiązywane przez okresowe skalowanie projektorów, jednak staje się to niezwykle trudne i mniej niezawodne jeśli wzrasta liczba projektorów. Urządzenia symulujące lot borykają się z tego typu problemami od pewnego czasu, i chociaż rozwinęły one systemy cyfrowego skalowania, systemy te zazwyczaj są wysoce wyspecjalizowane, co zwiększa koszty tworzenia i komplikuje generalnie cały projekt. System „samoskalownia” o szerszym zastosowaniu byłby lepszy, bo każdy mógłby modyfikować obszar projekcyjny i konfigurację projektorów w zależności od potrzeb. Gdyby można było modyfikować powierzchnię projekcyjną można by spontanicznie włączać dodatkowe obrazy i system włączył je do projekcji. Jeśli można by mieć swobodę przy mocowaniu projektora, można by dodawać projektory tak aby rzucane przez nie obrazy się nakładały co pozwoli na zwiększenie

rozdzielczości ekranu, intensywności obrazu, lub na zwiększenie generalnie pola projekcji.

Infrastruktura i zastosowanie telekolaboracji

Istnieje ogromnie dużo pracy w obszarze infrastruktury i zastosowania telekolaboracji, nie wspominając już o pracy potrzebnej we Współpracy Opartej o Zastosowanie Komputera. Wizja Biura przyszłości, to taka w której praca ta mogłaby zostać wykorzystana w nowy i ekscytujący sposób. Biuro będzie wyglądać jako szczególnie przykuwający uwagę interfejs służący do codziennych zastosowań.

Spośród istniejących dokonań w dziedzinie telekolaboracji, jedynym który próbuje dostarczyć podobnego do biura interfejsu jest TelePort. System TelePort wykorzystuje ekrany o rozmiarach ścian, każdy z nich pokazuje sztuczną scenę, która jest łączona z obrazem przesyłanym z oddalonego rzeczywistego miejsca naszych współmówców. Gdy uczestnicy telepołączenia poruszają się, ich pozycje są śledzone tak, że ich obraz jest odtwarzany z prawidłowej perspektywy. TelePort jest wbudowany w pomieszczenie, jest ostrożnie zaprojektowany tak by pasował do obrazu odtwarzanego pokoju. Celem tego jest sprawienie aby odtwarzany pokój wydawał się przedłużeniem pomieszczenia prawdziwego. Używają oni dokładnie skonstruowanych geometrycznych modeli dla stworzenia biurowego środowiska, oraz ludzki kształt otrzymany przez oddzielenie oddalonych uczestników od ich oryginalnego środowiska. Zamiast wybudowania specjalnego systemu telekolaboracji, który przypomina biuro, naukowcy chcą stworzyć funkcję do tworzenia podziału środowiska różnych biur.

Każdy milimetr przy Każdej milisekundzie

Jak dalece oddaleni współpracownicy i ich środowiska powinny wydawać się odległe? Choć odpowiedni do pewnych zadań, obraz ludzi dwuwymiarowy jak na video nie skutecznie wywołuje sens bycia z inną osobą potrzebny do skutecznej międzyludzkiej komunikacji. Lepiej widzieć i współdziałać z naszymi współpracownikami w trójwymiarze, w warunkach maksymalnie zbliżonych do tych jakie istnieją w fizycznym pomieszczeniu: gestykulowaniu, wskazywanie, chodzenie, machanie, używanie wszystkich delikatnych niuansów werbalnej i niewerbalnej komunikacji. Możliwością uatrakcyjnienia wizualnego byłoby użycie systemu modelowania lub odtwarzania obrazów w wysokiej jakości w trójwymiarze. Potrzebna będzie pojemność pracująca w rozmiarach pomieszczenia,

ponieważ potrzebna jest mobilność, ale też dlatego że chcemy widzieć więcej niż jednego uczestnika i chcemy widzieć ich w ich naturalnych środowiskach (biurach). System miałby być podobny do tego, w którym miejscowe i oddalone biuro wydają się fizycznie połączone wzdłuż pewnego połączenia jak na przykład wyznaczonej ściany, która działa wtedy jako SID. Jednak w przeciwieństwie do systemu, który nakłada dwuwymiarowy przekaz video odległych uczestników na obraz utworzonego wirtualnego biura, można zobaczyć trójwymiarową rekonstrukcję oddalonego biura i to co zawiera (łącznie z ludźmi) . Stąd będziemy mieli zdolność do uchwycenia i odtworzenia na odległość dynamicznego obrazu całego biura.

Gdyby istniał dostęp do dynamicznego modelu całego biura, łącznie z powierzchniami SID, automatycznie można by było korygować zmiany we właściwościach geometrycznych SID zależnych od zmian czasu. Ponadto jeśli kilka kamer mogłoby wychwytywać powierzchnię projekcyjną z różnych kątów, powinniśmy być w stanie zaobserwować zależny od zmian intensywności i koloru obraz na wyznaczonych powierzchniach projekcyjnych. Innymi słowy, w czasie pobierania modelu biura w oparciu o obraz można by dokonać autoskalowania wszystkich wyznaczonych powierzchni projekcyjnych. SID mógłby być niemal wszędzie albo być niemal gdziekolwiek w biurze. Nie ważne by było nawet jeśli powierzchnie projekcyjne były by nieregularne, lub gdyby geometria obrazu zmieniała się względem czasu, model biura utworzony w oparciu o obraz wskazywałby te zmiany. I jeśli można by było poświęcić pewną część zasięgu dynamiki w projektowanych obrazach, można by nawet użyć informacji o powierzchniowym odbiciu aby wytłumaczyć drobne zmiany w intensywności. Kluczowymi zaletami tego zespolonego podejścia jest to że autoskalowanie i projekcja są dokonywane przez to samo urządzenie.

Jeśli ktoś ma dostęp do dynamicznego modelu biura, łącznie z jego pracownikami, można by wydobyć lepszej jakości dane, przypisać znaczenie tym ‘lepszym’ danym, i potem w realnym czasie interpretować i reagować na ruch lub zderzenia obiektów, w celu wyśledzenia ruchu, interakcji, lub powiększenia obrazu AR. Z takimi możliwościami można by zastosować typ interakcji obecnej w Luminous Room (świecącym pomieszczeniu), gdzie kamery i rzutniki funkcjonują jako I/O Bulbs. W ten sposób można by śledzić ruch ludzkich dłoni tak by mogły one sięgnąć po i poruszać dryfującym w powietrzu trójwymiarowym modelem, lub

można by wykryć zderzenia między realnymi i wirtualnymi przedmiotami. Tak aby przedmioty wirtualne mogły być umieszczone na biurku.

Rysunek 1. przedstawia poglądowy szkic biura przyszłości, powielony i w użyciu w trzech różnych miejscach. Na suficie osadzone są kamery i projektory, użycie jasno zabarwionego materiału ściany projekcyjnej i biurka, i połączone użycie SID dla symultanicznej projekcji w oparciu o obraz i model geometryczny.

Aby osiągnąć wymienione wyżej cechy poboru obrazu, skalowania, i projekcji w nieustannie zmieniającym się wyglądzie biura z zarówno realnymi jak i transmitowanymi przez projektor pracownikami, wyzwaniem jest możliwość kontroli światła w biurze „na każdym milimetrze i w każdej milisekundzie”.

2 KLUCZOWE KOMPONENTY

Pomysł na biuro przyszłości łączy kilka kluczowych obszarów informatyki, komponentów które mogą być wyliczone oddzielnie z opisu ich właściwego zastosowania.

2.1 Tworzenie modelu w oparciu o dynamiczny obraz

Jednym z głównych komponentów systemu jest moduł, który uchwyci w sposób ciągły i w realnym czasie, na bazie wizerunku biura, model środowiska biurowego łącznie ze wszystkimi z wyznaczonych powierzchni, na których obraz jest wyświetlany. Ważne jest określenie/wydobycie głębi obrazu. Wiele technik używanych tradycyjnie związanych z „czasem realnym” jest trudne z powodu wysokich wymagań komputerowych i szerokopasmowych. W CMU, został wykorzystany wyspecjalizowany sprzęt głębi realnego czasu wraz z architekturą stereo systemu. Może pomieścić dane wejściowe z sześciu kamer i tworzyć trzydzieści klatek na sekundę, 256 x 240 ustawiać mapę głębi z intensywnością obrazu. Ma również zdolność do tworzenia szacunkowych obliczeń niewiadomych dla każdego pixela. Jedną z zalet tej techniki jest natychmiastowa próba i uchwycenie naturalnej głębi stereo. Dla porównania, używając laserowego skanera nie można ukończyć zeskanowanego obrazu w pojedynczej ramie ze względu na wypaczone kształty jak obiekty z przesuwających się obrazów. Każda technika, która zależy od obliczeń zrobiona z różnych klatek - próbek pobieranych w różnym czasie, wyłączając metodę strukturalnego światła, będzie mieć z tym problem.

Został też wynaleziony inny system wykorzystujący głębię realnego czasu na Uniwersytecie Columbia. Zademonstrowali oni zdolność do tworzenia 512x480 oszacowań głębi przy 30 Hz z dokładnością do 0.3 %. Ich technika opiera się na precyzyjnym fizycznym modelu wszystkich optycznych elementów systemu zależnych o bodźców i obliczeń: funkcje optycznego transferu, badanie i próbkowanie obrazu i operatory stopnia skupienia. Projektują one obraz z wysoką częstotliwością na ekran poprzez tą samą optyczną ścieżkę. Zaletą tego systemu nad głębią stereo, jest to że nie trzeba martwić się o problem odpowiedników. Jedynym zmartwieniem jest zniekształcająca faktura o wysokiej częstotliwości, która może być projektowana na ekran. To mogłoby być bardzo niepożądane jeśli użytkownik chce być w obszarze pobierania obrazu.

2.2 Odtwarzanie

Ta wizja biura przyszłości wymaga zdolności do generowania obrazów, które w czasie projekcji na obszar projekcyjny wydają się poprawne w stosunku do obrotu głowy obserwatora. Jest to prawda również dla systemów takich jak CAVE™, jednak w tym przypadku sytuacja jest trochę inna gdyż CAVE™ jest przeznaczony raczej do gładkich powierzchni. Przyszłe zdolności w generowaniu obrazu pozwolą na wykonanie projekcji o zwiększonych rozmiarach na przypadkowych powierzchniach.

Interesująca jest technika do zastosowania komputerowej grafiki w oprawie scenicznej teatru, z jej pomocą modeluje się wygląd dekoracji (tła) z perspektywy widowni. Bez korekty te dekoracje wydawałyby się rozmazane. W omawianym systemie pojawia się ten sam problem, ale przy zastosowaniu wielu projektorów. Należy zrobić to tak, aby projektowany obraz wyglądał poprawnie z perspektywy użytkowników. Technika ta została poszerzona do prezentowania rzutników optycznych i pokazywania poszerzonej metody radiacyjnej do symulowania właściwości oświetlenia.

2.3 Przestrzennie Immersyjne Wyświetlacze

Najbardziej znanym wyświetlaczem jest CAVE™ w dziedzinie grafiki. CAVE™ istnieje w wielu formach, z reguły jest skonfigurowany jako system prawo, lewo, lub tylny ścienny. W niektórych instalacjach używane są lustra ponad CAVE™, które rzuca obraz na podłogę. W czasie gdy CAVE™ dostarcza obrazu stereo

otaczającego użytkownika środowiska, obecne instalacje są ograniczone do jednego projektora na każdą ścianę. CAVE™ nie ma do czynienia z mieszaniem intensywności i nie ma metody łapania geometrycznego obrazu środowiska, co jest słuszne gdyż to nie było w kwestii tego systemu.

Wojskowy/lotniczy przemysł symulatorów jest pełen licznych przykładów przestrzennie immersyjnych ekranów. Te systemy typowo używają CRT projektorów, które wymagają częstego skalowania. Również one często ograniczają się do dopasowywania łączy ekrany zamiast uważania całego systemu jako czegoś co musi zostać połączone bez widocznych ‘szwów’. Kolejną techniką symulacji lotów jest umieszczenie wyświetlacza o dużej rozdzielczości w centrum widoku użytkownika i projektowanie obrazu o niskiej rozdzielczości na jego otaczający ekran. Jest to efektywne, ale też może dojść do ukazania się złącza w miejscu spotkania się obszarów o wysokiej i niskiej rozdzielczości. Złącze to jest problemem, gdyż często przeszkadza w osiągnięciu celu uczucia że jest się gdzieś indziej - użytkownicy widzą, że jest to tylko niedoskonały ekran projekcyjny.

Ekran w kształcie kopuły są kolejnym przykładem. Takie układy są często ograniczane do projektorów o wysokiej rozdzielczości i rzadko używają mechanizmu do wylapywania głębi lub informacji o powierzchni projekcyjnej danej sceny. Istnieje metoda, która koryguje zniekształcenia kopuły poprzez modelowanie jej za pomocą pięcio stopniowej siatki wielokątnej i GUI do manipulowania współrzędnymi tej siatki, ale to nie dzieje się w czasie realnym ani automatycznie, wymagana jest bezpośrednia interwencja użytkownika. Ta metoda przeznaczona jest do użytku, tylko wtedy gdy układ jest przesunięty lub wypada ze stałego ustawienia, to nie jest metoda przeznaczona do uaktualnienia projekcji w rzeczywistym świecie takich jak zmiana kształtu obszaru projekcyjnego czy occlusion properties.

Końcowym ważnym punktem dla wszystkich systemów jest to, że bazują one na specjalnych geometrycznych ustawieniach i mają stałe nie zmieniające się środowisko z jednolitą idealną powierzchnią projekcyjną. Również, miały one kontrolę nad każdą kwestią odnośnie systemu wyświetlania, od oświetlenia do precyzyjnego skalowania projektorów. Rozwiązaniem problemu projekcji na przypadkowe powierzchnie w rzeczywistym czasie, są procedury automatycznego skalowania.

3 METODY

W tym rozdziale opisany jest sens używania kamer i projektorów, które mogą być kierowane albo w trybie wyłapywania obrazu lub jego odtwarzania. Jeśli chodzi o tryb wyłapywania obrazu, projektory i kamery mogą być użyte razem aby uzyskać głębię. Jeśli chodzi o tryb odtwarzania modelu budowanego w oparciu o obraz mogą być używane do projekcji geometrycznej i fotometrycznej skorygowanych obrazów na ekrany projekcyjne (o potencjalnie nieregularnych kształtach).

Modelowanie w oparciu o obraz dynamiczny

Wydobycie Głębi

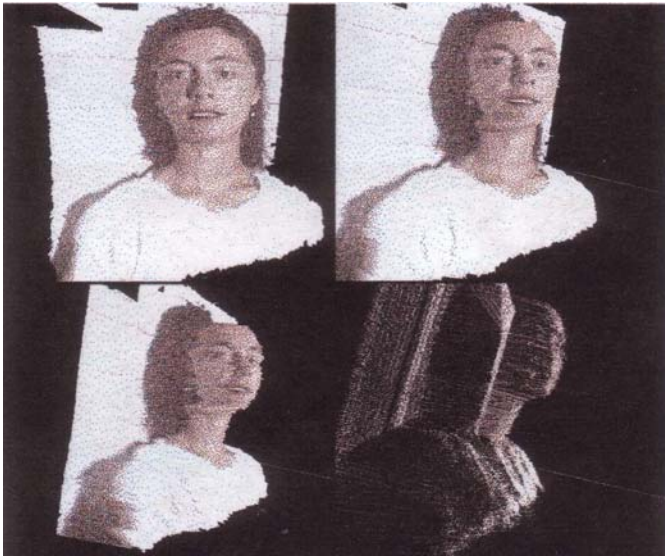
Metoda wydobywania głębi obrazu biura powinna funkcjonować w obszernych roboczych obszarach z powierzchniami o fakturze o wysokiej częstotliwości jak i powierzchniach które nie mają faktury. Do stworzenia modelu powierzchni projekcyjnej, potrzebujemy dużej dokładności aby projektowanie obrazu było poprawne. Aby dokonać modelu dynamicznie zmieniającego się obrazu potrzebujemy większej szybkości uaktualniania obrazu aby pokazać ruch o potencjalnie zmniejszonej rozdzielczości lub precyzji. Ta technika nie powinna przeszkadzać - tak aby ludzie mogli normalnie pracować w środowisku jej zastosowania. To uniemożliwia użycie laserów lub innych metod inwazyjnych.



Rys. 2. Obecne instalacje biura przyszłości a) kamery i cyfrowe projektory światła są zamontowane blisko sufitu b) ściany składają się z białych styropianowych tablic zamontowanych pionowo przy stanowiskach. Jak pokazane jest na rys. b), są one używane jako powierzchnie wyświetlające dla przestrzennie – immersyjnych wyświetlaczy.

System ten obecnie używa jednej kamery wideo i jednego projektora w parze, mimo że większa ilość kamer mogłaby funkcjonować w połączeniu z jednym projektorem. Problem powiązania/odpowiedniości jest rozwiązany przez projekcje dwójkowo zakodowanych poziomych pasków. Kamera wychwytuje x ilość następujących po sobie obrazów, tworzy podwójny obraz i generuje x -bitowy kod dla poziomego paska dla każdego piksela korespondującego z tym kodem dla tego poziomego paska, który został pokazany w tym pikselu. To pozwala na rozróżnienie między 2 do potęgi n -tej poziomymi paskami. Tabela triangulacyjna utworzona w oparciu o dane skalowania pozwala na tryliniową interpolację, aby obliczyć przekrój promieni pseudo-kamery i projektowanych poziomych powierzchni. Współrzędne trójwymiarowego obrazu punktów powierzchni obrazowanych przez każdy piksel są później używane wraz z informacją o kolorze do skompletowania modelu opartego o obraz. Poziome paski mogą być rzucane powtórnie aby obliczyć głębokość dla obrazu dynamicznego. Ta sama metoda może zostać użyta do skanowania ścian, ludzi i poruszających się przedmiotów o różnych stopniach skomplikowania dla prognozowania obrazów pikselowych. Wybór kamery i projektora aby skanować obraz trójwymiarowy jest wymagana przez fakt że powierzchnie projekcyjne mogą nie posiadać odpowiedniej faktury aby zapewnić wystarczającą ilość wskazówek o odpowiednikach. Użycie tego samego projektora do skanowania i odtwarzania zapewnia zespolenie dwóch zadań i jedynym dodatkowym komponentem jest kamera. Zwiększenie szybkości kosztem dokładności zaprowadziło do dwóch typów par kamera-projektor. Relatywnie statyczne powierzchnie projekcyjne jak ściany i meble są wzorowane dokładniej i w wolniejszym trybie przez zewnętrzne układy kamera-projektor niż ludzie i poruszające się przedmioty, które są skanowane przez wewnętrznie patrzący układ kamera-projektor. Trudną częścią użycia dwóch oddzielnych przyrządów dla wydobycia głębi jest skalowanie. Używamy najpierw do odnalezienia parametrów kamery przy użyciu systemu szachownicy na płaskiej powierzchni. Potem ta metoda używana jest do skalowania projektora w odniesieniu do tej samej płaskiej powierzchni. Połączenie tych dwóch daje powiązanie między kamerą i projektorem. Aby odnaleźć związek między układami projektor-kamera, najpierw określona jest transformacja między dwiema kamerami poprzez oglądanie wspólnego wzorca szachownicy na płaskiej powierzchni. Potem, przy użyciu opisanej wyżej metody,

te dwie pary/ układy są skalowane w odniesieniu do pojemności roboczej. Procedura jest łatwiejsza gdy stożki ścięte obu kamer pokrywają się znacznie.



Rys. 3. Przykładowe rezultaty (nieruchome sylwetki uaktualniane na żywo) wydobywania głębi używając binarnie zakodowanego strukturalnego światła.

Wykrycie zmian w geometrii obrazu przez obraz kamery nie jest ‘żwawe’ gdy powierzchnia projekcyjna nie posiada odpowiedniej faktury. Jednak można dokonywać projekcji na powierzchnie dowolne. Dowolna faktura sama w sobie nie będzie niewidzialna dla ludzkiego oka. Wykryte zmiany w geometrii obrazu na przestrzeni czasu mogłyby być rezultatem albo właściwych zmian w powierzchni albo przesunięcia w skalowaniu systemu.

Użycie kamer pozwala na samo skalowanie systemu okresowo aby zrekompensować za błędy spowodowane czynnikami środowiska jak temperatura, albo wibracje.

KOLOR I ODBICIE

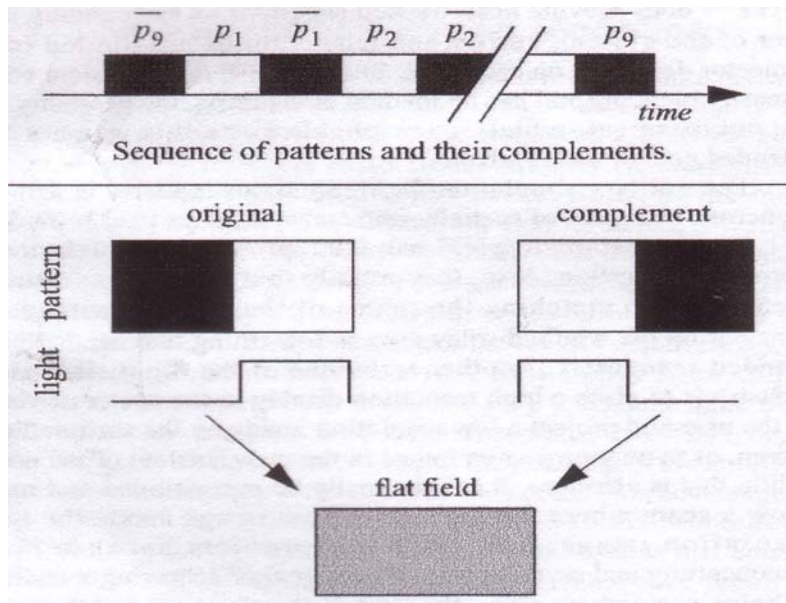
Projektor używany jest w dwóch trybach, wydobywania obrazu i odtwarzania. Aby otrzymać informacje o kolorze powierzchni projektor służy jako źródło jasnego światła razem z synchronizowaną kamerą. Jednak tryby te mogą być pozostawiane poprzez wstawianie białych ramek pomiędzy ramki obrazu projektowanego. W dwójkowo kodowanej pulsowo modulacji (PCM) projektorów , tylko kilka bitów jest używanych do projekcji białych ramek gdy inne bity mogą zostać użyte do projekcji ramek obrazu przy zmniejszonej rozdzielczości koloru.

Obecnie obraz ten jest wyświetlany przy pomocy czarnego obrazu ukazującego się po białym wzorze i obserwowaniu powstałego ciemnego i jasnego obrazu z jednej perspektywy aby ocenić funkcje odbicia pikselowego. Funkcja ta jest podstawowo używana do progowania obrazu rzucanego przez dwójkowo kodowane strukturalne światło przy założeniu że reakcja kamery jest linearna względem intensywności. Krzywe reakcji kamery mogą być obliczone przy pomocy wyświetlania obrazu przy pomocy innych poziomów nasilenia. By ukończyć ten model utworzony w oparciu o obraz, mogą być pobierane próbki powierzchni tej sceny ze złożonych punktów widoku aby oszacować funkcję obustronnego rozprzewadzenia odbicia (BRDF).

Kamera używana jest do pikselowego wydobywania głębi jak i do wydobywania koloru. Ponieważ obie procedury dzielą tę samą oś optyczną, nie ma mowy o dryfowaniu obrazu. Podobnie, ten sam projektor jest używany do projekcji wzorów światła strukturalnego, dla wydobywania głębi, i do odtwarzania na powierzchniach o wydobywanej głębi.

3.2 Niewykrywalne Światło Strukturalne

W odniesieniu do światła opisanego w sekcji 3.1, celem jest sprawienie aby wydawało się ono przeciętnemu widzowi jak nic więcej niż żarzące się białe światło, a nie szereg przeblyskujących podwójnych wzorów. Metodą do robienia tego jest użycie światła niewykrywalnego. Kombinacja podziału czasu i techniki zniesienia światła do ukrycia wzorów to gwałtowna projekcja serii obrazów w białym świetle. Rysunek 3 ukazuje cykl obrazów projektowanych w czasie. Światło to wykorzystuje n-tą ilość wzorów aby rozłożyć 2 do n-tej poziomych pasków. Okres powtarzania wzorów byłby na przykład 1/60 sekundy przy 60Hz szybkości wydobywania głębi. Obraz 4. pokazuje jak dany wzór p_i i uzupełnienie p_i są zespolone przez system wizualny w taki sposób że cykl ten wydaje się projekcją płaskich obrazów lub białego światła. To samo podejście może zostać zastosowane do projekcji tego światła jak również grafiki wideo, ułatwiając autoskalowanie do wyznaczonego obszaru projekcji. Użyty został cyfrowy projektor świetlny, który używa PCM do projektowania wzoru obrazu i jego uzupełnienia. Zsynchronizowana kamera może określić/zmierzyć to światło przez integrowanie światła w czasie projekcji obrazu.



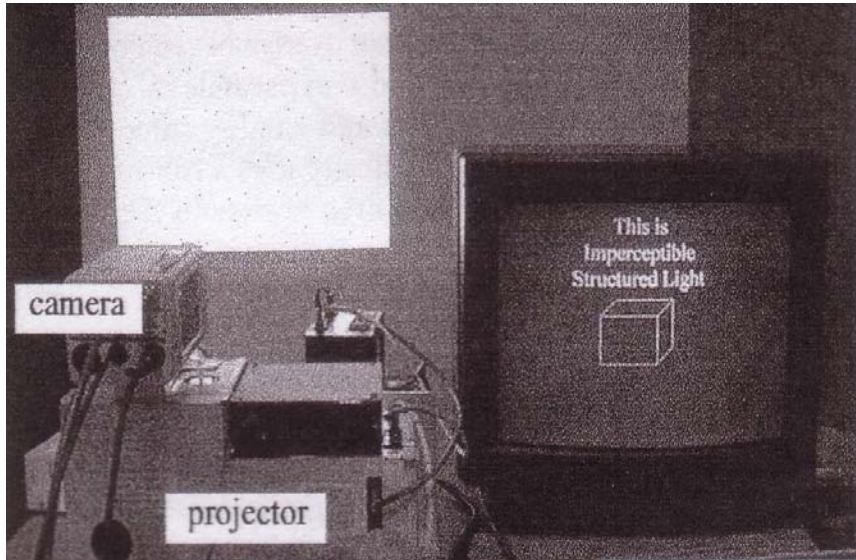
Rys. 4. Wzór i uzupełnienie są wizualnie połączone, rezultatem jest pojawienie się płaskiego pola lub „białego” światła.

Ograniczony dostęp do cyfrowego projektora obecnie ogranicza zdolność do zastosowania całkowicie niwychwytywalnego światła, możemy zademonstrować po kolei pobieranie obrazu w czasie rzeczywistym i ten typ światła. Rysunek 6 pokazuje użycie podobnego podejścia do włączenia światła strukturalnego do nieruchomego obrazu w przeciwieństwie do białego światła. Wspólna praca naukowców i wynalazców takich projektorów wprowadzi zupełnie nowe bogactwo możliwości do odtwarzania.

3.3 Geometryczna Rejestracja

Informacje o głębi obrazu uzyskane z różnych kamer muszą zostać pospinaane razem aby skompletować geometrię obrazu projekcji. Jest to wymagane do generowania poprawnego obrazu z pojedynczego projektora dla wielu złożonych powierzchni i dla łączenia intensywności jeśli obrazy z projektorów się nakładają. Obrazy są pobierane z odrębnych kamer i istnieją w różnych układach współrzędnych. Więc aby zespolić wydobywany obraz korespondujące ze sobą punkty z nakładających się obrazów są wykorzystywane. Punkty te są generowane przy użyciu światła strukturalnego niewidocznego dla szeregów i kolumn pikseli. Kod dwójkowy danego punktu w sposób unikalny identyfikuje korespondujące piksele projektora. Pary pikseli w dwóch kamerach dzielą kod dwójkowy i są użyte do obliczenia transformacji między zbiorami danych o głębi. W innym przypadku, geometryczna

odpowiedniość/korespondencja między pikselami dwóch różnych projektorów jest ustalana przez obserwację obszaru rzucanego przez jeden projektor a pokrytego obrazem innego projektora. Zakładamy że każda para obrazów jest pokryta przynajmniej częściowo (w przypadku jednej trzeciej to pokrycie jest całkowite).



Rys. 5. Niezauważalne strukturalne światło jest demonstrowane w laboratorium. Cyfrowy projektor z lewej strony wyświetla tekst pokazany na

monitorze i jego uzupełnienie, jednakże tekst może być tylko widziany z synchronizowaną kamerą, taką jaka jest umieszczona na projektorze powyżej.

Dołączone zdjęcie oscyloskopu pokazuje impulsy, które odpowiadają/korespondują w krótkim czasie, kiedy wzór (tekst w tym przypadku) jest wyświetlany.



Rys. 6. Niezauważalne strukturalne światło jest „osadzone” w zdjęciu. (a) Początkowe zdjęcie (Tokaj, Węgry). (b) Binarne zdjęcie, które jest niezauważalne. (c) Dwa zdjęcia połączone i sporządzone mapy odpowiednio podzielonych sekwencji bitów. (d) ostateczny rezultat, pokazujący zdjęcie początkowe wyświetlone na ścianie, podczas gdy „osadzony”: na nim niezauważalny obraz jest wychwycony i wyświetlony na monitorze (na dole po lewej stronie).

3.4 Odtwarzanie

Celem jest generowanie obrazów które wydają się poprawne dla obserwatora w czasie projekcji na obszar projekcyjny (o potencjalnie nieregularnych kształtach) . Ponieważ obserwator może poruszać się po biurze , obecnie używa się magnetycznego śledzenia ruchu głowy aby określić obecną pozycję obserwatora. Dostarczane dane do algorytmu są modelem powierzchni, wewnętrznymi i zewnętrznymi parametrami projektora, pozycji obserwatora, i pożądanego obrazu (takiego obrazu jaki chcemy aby widz zobaczył). Ten pożądaný obraz będzie typowo wynikiem konwencjonalnego odtwarzania obrazu trójwymiarowego.

Algorytm może pracować dla każdego typu powierzchni (na przykład NURB), pod warunkiem że model odtwarzania obrazu jest dokładny. Odtwarzanie może się odbyć przy użyciu wielu różnych metod (na przykład kalkowanie promieniem) ;obecne urządzenie wykorzystuje fakturę obszaru projekcyjnego OpenGL aby osiągnąć odpowiednią szybkość pracy sprzętu. Ta technika projekcji jest rozwinięciem poprawnie perspektywicznego mapowania faktury która może zostać użyta do dowolnej projekcji dwu wymiarowych obrazów na geometrię w czasie rzeczywistym.

Opisujemy dwa pasmowe podejście do odtwarzania obrazu trójwymiarowego na potencjalnie nieregularne powierzchnie. W pierwszym podejściu obliczymy ‘pożądaný obraz’ poprzez odtwarzanie trójwymiarowych scen z perspektywy widza. Ten pożądaný obraz przechowywany jest jako mapa fakturowa. W drugim podejściu faktura jest skutecznie rzutowana, z perspektywy obserwatora, na wielokątny model obszaru projekcyjnego. Obszar ten (z wyświetloną nań mapą faktury), jest potem odtwarzany z perspektywy projektora. Gdy użytkownik się porusza pożądaný obraz się zmienia i też jest projektowany zgodnie z nową pozycją użytkownika.

Wiele projektorów może być użytych do zwiększenia obszaru projekcji. Aby zapewnić całkowite pokrycie obszaru obrazem, każda część obszaru musi być widoczna dla przynajmniej jednego projektora. Ustawienie projektora w drugim typie podejścia pozostaje niezmiennie. Nie ma powodu aby nakładanie się obrazu projektorów przeszkadzało w powstawaniu efektownego obrazu z punktu widzenia użytkownika. Rysunek 7.

Aby sprecyzować kierunek oglądania projektowanej faktury za pomocą jednego oka, potrzebna jest jedynie pozycja użytkownika a nie orientacja. Pole widzenia, które zawiera wszystkie wielokąty sztucznego przedmiotu jest wystarczające dla stożka projekcji faktury. Stożek ten może zostać 'przystrzyżony' jeśli wykracza poza obszar stożka modelu obszaru projekcyjnego. Stożek ten jest nastawiony zgodnie z widokiem użytkownika na model wielokąta powierzchni projekcyjnej. Parametry stożka obserwatora w czasie pierwszego pasma i faktura projekcji w czasie drugiego są identyczne.

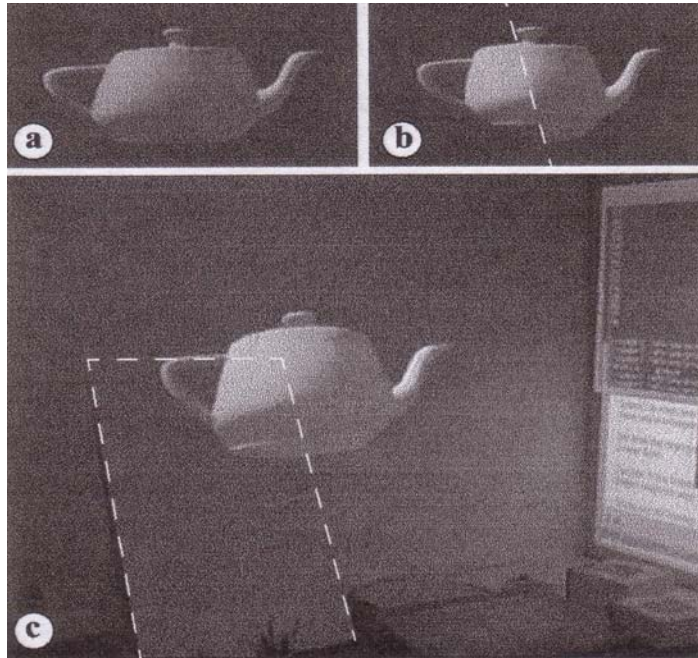
Zakładamy że projektory nie mają właściwości zniekształceń promienistych, stąd projektory można nastawić za pomocą punktowej kamery.

Dla optyczności cyfrowego projektora z 'mikrolustrem' (DMD), co jest ważnym założeniem. Jednak jeśli projektor wykazuje zniekształcenia promieniste musimy wstępnie zniekształcić odtwarzany obraz zanim jest on przesyłany do ramy buforowej projektora. To wstępne zniekształcenie może być wykonane przy pomocy nie linearnego trójwymiarowego wypaczenia geometrii powierzchni projekcyjnej lub przy użyciu dwuwymiarowego wypaczenia z mapowaniem faktury.

WYZWANIA

Prędkość

Dwu pasmowa metoda odtwarzania składa się z normalnego trójwymiarowego odtwarzania w pierwszym paśmie, po czym następuje drugie pasmo które mapuje pożądaną fakturę na powierzchnię projekcyjną. Dodatkowy koszt algorytmu pochodzi od przeniesienia ramy buforowej z pierwszego pasma do pamięci faktury odtwarzania modelu razem z zastosowanym mapowaniem w ciągu drugiego pasma. Stąd jest to podstawą aby uprościć geometrię powierzchni projekcyjnej.



Rys. 7. Interpretacja wielopowierzchniowa. (a) dzbanek do herbaty wyświetlony na pojedynczą płaską powierzchnię. (b) zniekształcona projekcja w wyniku wprowadzenia drugiej płaskiej powierzchni. (c) Ostateczna poprawiona projekcja uzyskana przez wyciągnięcie/usunięcie wyświetlonej geometrii powierzchni i następnie zastosowanie dwupasmowej interpretacji programu.

Krawędzie drugiej płaskiej powierzchni – wykazują pochylenie w stronę ściany na biurko (c) – jest zakreślona linią przerywaną.

Tworzenie równoległości

Jeśli więcej niż jeden projektor jest używany, każdy projektor może mieć oddzielny silnik odtwarzania. Odtwarzanie może odbywać się równoległe, i każdy silnik projekcji musi tylko załadować te części modelu obszaru projekcyjnego, które są widoczne z pozycji odpowiedniego projektora. Istnieją pewne wstępne dowody na to że ta metoda jest szybsza względem konwencjonalnego odtwarzania przy użyciu większej liczby projektorów, złożonymi wyświetlaczami, lub złożonymi modelami graficznymi. Powodem tego jest to że pierwsze pasmo algorytmu (konwencjonalne odtwarzanie) musi być wykonane tylko raz, a potem drugie pasmo może zostać wykonane dla każdego projektora lub nawet dla każdego wielokąta na powierzchni projekcyjnej.

Jeśli pierwsze pasmo jest również tworzone równolegle, wszystkie korespondujące linie graficzne powinny dotrzeć do projektowanego syntetycznego modelu w tym samym czasie. To mogło by być problemem jeśli model będzie się dynamicznie zmieniał i gdy każda graficzna linia musi odczytywać obraz modelu przy każdej zmianie obrazu. Kolejne problemy to jeśli model powierzchni jest dynamiczny jak i gdy silnik projekcji są na różnych urządzeniach.

Skrytość/Uśpienie

Jest obecna w tym systemie wewnętrzna uśpioność na dodatek do tradycyjnego śledzenia ruchu i uśpienie odtwarzania spowodowane dwupasmową metodą rysowania modeli. Dla dużych modeli, czas odtwarzania może być inny dla różnych projektorów więc istnieje pojęcie wewnętrznego opóźnienia projektora w czasie odtwarzania. Jeśli wszystkie projektory są kierowane z jednego urządzenia, wtedy ustawienia portów obrazu w obrębie pojedynczego programu odtwarzania dla każdego projektora i synchroniczne uaktualnianie ram buforowych wyeliminuje ten problem.

3.5 Generowanie funkcji łączenia/mieszania obrazu

Aby opisać funkcje mieszania używa się tradycyjnej frazy używanej w dziedzinie grafiki komputerowej 'wartości alfa'. Kiedy buduje się jakiś rodzaj łączonego multi projektora SID, należy zmierzyć się z dwoma problemami związanymi z przejściem między projektowanymi obrazami: można albo tak zaprojektować system aby projektowane obrazy się nie nakładały i by mogły być tak dopasowywane aby się ledwo dotykać i nie tworzyć widocznych spójń, lub można pozwolić projektowanym obrazom się nakładać i zastosować pewne środki łączenia. Ten drugi sposób typowo używa funkcji linearnej rampy lub krzywej cosinus aby wygładzić przejście między obrazami projektowanymi z różnych projektorów. Projektanci CAVE™ ćwiczą pierwszą opcję poprzez ograniczanie systemu do dobrze zdefiniowanego, relatywnie prostego ustawienia ekranu przy braku nakładania się obrazów z projektorów. Chcielibyśmy być zdolni rzutować obraz na dowolne i potencjalnie nieregularne powierzchnie projekcyjne w biurze, oznacza to jednak, że nie można zastosować pierwszego podejścia gdyż nie zakłada się żadnej kontroli nad powierzchnią. Ponadto lepsze będzie bardziej elastyczne ustawienie za pomocą którego zbiór projektorów może zostać użyty do projekcji na ta samą powierzchnię aby osiągnąć większą rozdzielczość.

Wprowadzono funkcję specjalną/dodatkową poprzez przypisywanie wartości alfa między 0 i 1 dla każdego piksela w każdym projektorze, i jak to zostało opisane w sekcji 3.3 zapewniono, że każdy punkt wyświetlany odpowiadający pojedynczemu pikselowi kamery ma sumę alfa równą jeden. Stąd projektory mają tę samą reakcję natężenia. Są dwa przypadki: punkt wynikający z projekcji z tylko jednego projektora i punkt wynikający z projekcji wielu projektorów. W pierwszym przypadku rozwiązanie jest trywialne. Jednak w drugim przypadku, w przypadku gdzie występuje pokrywanie się projektowanych obrazów, tworzymy z wartości alfa funkcję odległości do początku/końca obszaru nakładającego się, z ograniczeniem że wartości alfa punktów z różnych projektorów, korespondujących z tymi samymi punktami w przestrzeni, muszą razem wynosić jeden. Aby przypisać różne dodatki dla pikseli projektorów, tak właściwie tworzony jest obraz alfa dla każdego projektora. Obraz ten zawiera $(1 - \text{pożądana wartość alfa})$ dla każdego piksela. Ten obraz alfa jest odtwarzany jako ostatni. W OpenGL instalacji jest to osiągnięte za pomocą przezroczystej faktury. Kamera w systemie zamkniętej pętli zapewnia fotometrycznie poprawne nawet przy różnej jasności projektorów i nawet gdy powierzchnia projekcyjna nie ma jednolitych wartości odbicia. Mimo że projektory cyfrowe reagują z linearną intensywnością, używają korekcji de-gamma. Użycie obrazu alfa kompensuje korekcję de-gamma.

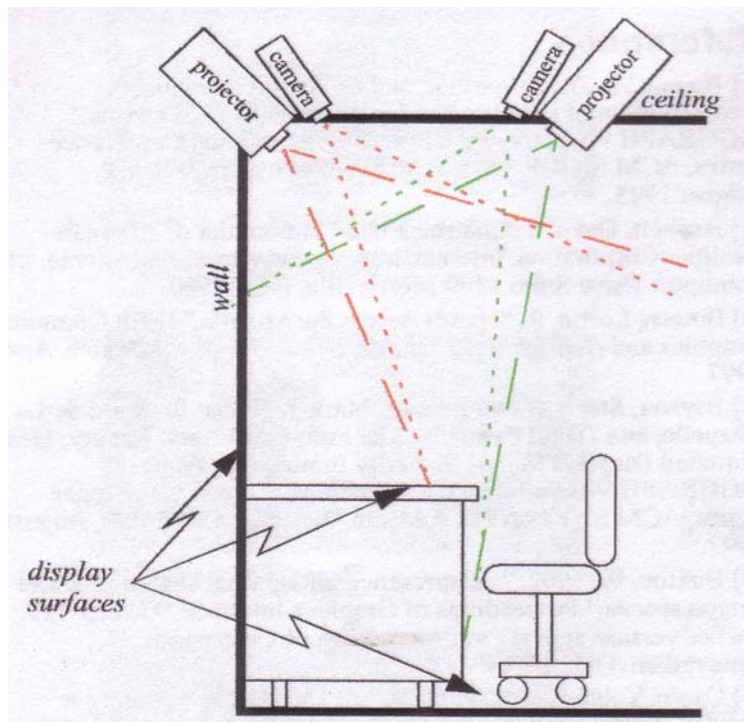
3.6 Uproszczenie Danych o Głębi

Modelowanie w oparciu o dynamiczny obraz całego biura zakończy się zbiorami ogromnej ilości danych, pod warunkiem że dane te były by dla każdego piksela dla zbioru kamer, pojawiających się z szybkością prawidłową dla obrazu video. Jednak większość tych danych powinna być w znacznej mierze ze sobą powiązana zarówno czasowo jak i przestrzennie. Tak naprawdę mało prawdopodobne jest aby wygląd większej części biura zmienił się diametralnie w ciągu krótkiego okresu czasu, jest to zwłaszcza prawda dla biurowych ścian i większości wyznaczonych obszarów projekcyjnych. Sensownym jest próbowanie uproszczenia danych tak aby system nie musiał mieć do czynienia z taką horrendalną ilością danych. Na przykład, Radim Sara i Ruzena Bajcsy na Uniwersytecie Pensylwanii utworzyli zbiór danych głębi biura, które ma około pół miliona wierzchołków (kątown). Metoda upraszczania musi być jednak ostrożna by nie upraszczać w obszarach gwałtownych zmian lub dużej krzywizny gdzie informacje mogą ulec zgubieniu.

Niestety, dynamiczna natura i obecność hałasu w opisywanym systemie, odrzuca użycie dobrze ustalonego algorytmu uproszczenia. Metoda która obecnie jest używana nie jest dobrze zdefiniowanym matematycznym podejściem, ale raczej metodą opartą o własne doświadczenie ludzi zajmujących się nią, która wyprodukowała jakościowo zadowalające rezultaty oparte na wspólnych wartościach zbiorów danych. Najpierw zastosowano metodę krzywizny używając metody tangensowej, a potem użyto metody odległości Euklidesa w stosunku do pozostałych punktów. Wybrano ten szczególny cykl kroków, ponieważ metoda krzywizny jest z reguły znacznie skuteczniejsza w eliminowaniu punktów niż ta druga. Taka metodyka powoduje eliminację szybkości o 80-90% bez widocznej utraty informacji. Jest tak ponieważ większość przedmiotów w środowisku biurowym jest miejscowo bardziej płaska.

3.7 Śledzenie ruchu

Opisywane biuro przyszłości z pewnością mogło by z powodzeniem zostać użyte tylko w dwuwymiarowym trybie, ale jest bardziej sensownym rozważenie jego dodatkowego użycia jako trójwymiarowej wizualizacji środowiska. Potrzebna jest zdolność do śledzenia ruchu głowy oglądającego, aby odtwarzać obraz zgodny z perspektywą oglądającego. Interesujące jest, że dla oglądania za pomocą jednego oka nie potrzebna jest znajomość orientacji oka, ponieważ odtwarzany obraz wyjątkowo określany przez pozycję oka widza. Dla oglądania stereoskopowego potrzeba jedynie być w stanie albo śledzić orientację oka i głowy użytkownika, albo obu oczu. System ten wykorzystuje projekcje syntetycznych obiektów na prawdziwe powierzchnie, dla których wydobyty model powierzchni jest przyjęty za poprawny w przestrzeni. Jakikolwiek błąd w odczytywaniu urządzenia śledzącego ruch w przestrzeni spowoduje widocznie niepoprawny zapis między projektowanymi obrazami powierzchni projekcyjnej. Ta sytuacja jest podobna do systemu powiększania rzeczywistości, gdzie tradycyjnie zapis oparty o wizję jest użyty do korekcji odczytów urządzenia śledzącego ruch.



Rys. 8. Cyfrowe projektory i kamery pracują razem w celu wychwycenia głębi, koloru, i powierzchni „odbijającej” informacje o obiektach i uczestnikach w danej scenie. Podzespoły projektora są również używane w celu wyświetlania; wychwycona głębia wyświetlonej powierzchni, kolor, i „odbite” informacje mogą być użyte do autokalibracji wyświetlanych powierzchni, tak żeby wyświetlane obrazy były geometrycznie i fotometrycznie prawidłowe z punktu widzenia obserwatora, jak również zachodzące na siebie projekcje były właściwie nałożone na siebie - „wtopione”.

4 OBECNE ZASTOSOWANIE

Ponieważ z definicji, zrealizowanie planu konstrukcji biura przyszłości jest wciąż odległe, istotny jest stały postęp i pewne obiecujące wyniki. Zastosowano roboczy system używając faktur projekcyjnych. Dokonano również odrębnych demonstracji 1-systemu wydobycia głębi funkcjonującego przy 3Hz, 2- niewidzialne światło strukturalne, 3-początkowe eksperymenty w łączeniu intensywności.

Rozmiar biura jest 10x10 stóp, w którym umieszczone jest pięć cyfrowych projektorów świetlnych o rozdzielczości 800x600 i dwie wideo kamery o rozdzielczości 640x480 (Rysunek 4). Patrzące na świat zewnętrzny pięć projektorów jest sterowanych symultanicznie z SGI Nieskończoności Rzeczywistości. Cztery spośród tych projektorów rzucają obraz na relatywnie pionowe powierzchnie

a jeden rzuca obraz na dół na biurko i na podłogę. Projekторы świetlne z dwójkowo kodowanym PCM rzucają 24 bitowy kolor w szybkością uaktualniania 60Hz. Projekторы umożliwiają nie dokładną względem osi projekcję z małym offsetem(równoważnością) bez znacznych problemów z ostrością. Wychwytywacze obrazu kamery to SGIO2 i obecnie używany przełącznik wideo, aby przełączać między poszczególnymi kamerami wideo w czasie wychwytywania obrazu ze środowiska. W bardzo bliskiej przyszłości odtwarzanie będzie odbywało się za pomocą innych portów widoku pojedynczego okna i może zostać użyte do synchronicznego uaktualniania wszystkich ram buforowych projektorów. System ten również obejmuje magnetyczne urządzenie śledzące Ascension (wniebowstąpienie) dla śledzenia pozycji użytkownika.

Ściany biura są zrobione z względnie niedrogich płyt styropianowych. Odrębny układ projektor-kamera wykorzystany jest do tworzenia obrazu dynamicznego opartego o mały roboczy model o objętości 3x3x2 stóp, i może być używany do tworzenia modelu ludzkiej twarzy. System tworzy mapę głębi o rozmiarach 640x240 o szybkości 3 uaktualnień na sekundę, na którym pokazywany jest przekaz video. Używając bezpośredniego cyfrowego interfejsu do cyfrowych projektorów z PC, możemy projektować wzory które nie są wychwytywane przez ludzkie oko, ale widoczne z zsynchronizowanej kamery. Zdolność do zmiany kodowania dwójkowego PCM światła dla projektorów pozwala na użycie różnych bitów dla różnych celów; w szczególności możliwości do „zespalania” PROM dla projektorów co pozwala na zademonstrowanie niewychwytalnego światła (Rysunek 5). Z PROM, system dwójkowy może być wyświetlany 24 razy na 60Hz klatkę-czyli co 694 mikrosekundę. Stąd system i jego uzupełnienie mogą być pokazywane sekwencyjnie co około 1.4 milisekundę. Zsynchronizowana kamera z czasem naświetlania krótszym niż 700 mikrosekund została użyta na obrazie 7. Używając PROM można teoretycznie projektować i pobierać $60 \times 24 = 1440$ wzorów dwójkowych na sekundę. Obecne wychwytywacze ram/klatek, jednak mogą przetwarzać tylko 60 obrazów na sekundę. Lepszy cyfrowy interfejs od DLP pozwoli na odtwarzanie obrazów stereo przy 60 Hz.

Od uczestników oczekuje się siedzenia na krzesłach przez większość czasu, obecne ustawienie pozwala użytkownikom przeciętnej wysokości (poniżej 6 stóp =2 metry) aby wstawali i poruszali się bez blokowania projekcji na ściany pod warunkiem że robią to przynajmniej cztery stopy (ok. 130 cm) od ściany.

Ustawienie Biura Przyszłości pozwala na stabilność jeśli chodzi o większą ilość par kamer i projektorów aby albo zwiększyć rozdzielczość wydobywanych powierzchni lub aby zwiększyć rozdzielczość odtwarzanego obrazu na powierzchni projekcyjnej. System ten inny niż sprzęt komputerowy kosztuje około \$35,000. Przewiduje się minimalne koszty utrzymanie projektorów, kamer, powierzchni projekcyjnych, ponieważ w tym systemie stosuje się metody samo skalowania.

5.PRZYSZŁA PRACA

Integracja pobierania sceny z odtwarzaniem w taki sposób, że pobieranie odbywa się w sposób niewidzialny, a przynajmniej nie będzie przeszkadzało w pracy. To pociągnie za sobą kombinację kontroli światła i kamer (możliwie z szerokim polem widzenia i dużą rozdzielczością). Razem z partnerami w GRASP Laboratorium na Uniwersytecie Pensylwanii bada się ciąg opcji między surową kontrolą nad wszystkimi światłami a między małą lub całkowitym brakiem kontroli nad światłami ale przy użyciu zbioru kamer i techniki opartej o pasywną/bierna korelację. Zadaniem jest otrzymać system o złożonej korelacji do eksperymentów z pobieraniem światła strukturalnego i środowiskiem odtwarzania.

Jako część pobierania obrazu, można wykryć zmiany w powierzchni projekcyjnej i dopasować odpowiedni system odtwarzania. Obecnie można tego dokonać przy nie interaktywnej szybkości. Ostatecznie należy również zbadać metody jak wykrywać zmiany powierzchni dla celów takich jak : śledzenie ruchu i informacje o gestach.

Należy również poprawić generowanie obrazu z lepszym połączeniem poprzez badanie metod odtwarzania obrazu, aby skonstruować obraz docelowy ze złożonych obrazów odniesienia. Jest możliwość że system odtwarzania posługujący się algorytmami multi-projektorów i multi-powierzchni projekcyjnych, okaże się skuteczny. Na dodatek, należy dokonać korekty dla niespójności odbicia powierzchni, i wykorzystać te informacje podczas czasu emisji aby dostroić odtwarzany obraz.

Przyszłe plany to wykorzystać ten system w telekolaboracji wykorzystującej wszechstronny projekt mechaniczny i produkcję z partnerami w NSF Centrum Nauki i Technologii Grafiki i Wizualizacji. Wykorzystać również CAVE™ i mu podobne systemy do połączenia kilku laboratoriów poprzez sieć szybkiego przepływu informacji z nowym środowiskiem odtwarzania immersyjnego.

6 Podsumowanie

Pokazane zostały początkowe rezultaty nowego pół-immersyjnego odtwarzania w podobnym do biura środowisku, takim które łączy pobieranie obrazu i odtwarzanie. Rozwinięto techniki pobierania geometrii nieregularnych powierzchni, służące potem do modyfikowania przebiegu odtwarzania, aby umożliwić projekcję na nieregularne powierzchnie tak by projektowany obraz wydawał się poprawny dla obserwującego, którego lokalizacja jest znajoma dla systemu. Opisana metoda 'wstrzykiwania' światła strukturalnego na obraz które jest niewidzialne dla widzów ale wychwytywane przez zsynchronizowane kamery. Techniki te mogą zostać wykorzystane w innych środowiskach pokazowych, które składają się ze złożonej liczby projektorów lub które wykorzystują odtwarzanie złożonych geometrii.

W podsumowaniu, główny trąd w dziedzinie informatyki od kilkudziesięciu lat to przejście od jednego do wielu, od bycia ograniczonym przez dostępność środków do angażowania środków bez względu na ich lokalizację. Jednym z obszarów pozostających poza wpływem tego typu rozwoju było odtwarzanie komputerowe i obszar w którym wyniki mogą być zaprezentowane. Opisany system nadrabia to poprzez umożliwianie jakiegokolwiek przedmiotowi lub grupie przedmiotów na bycie obszarami projekcyjnymi. Od teraz nie trzeba tłoczyć dużej ilości informacji na małym ekranie komputerowym, gdyż można mieć tyle przestrzeni do pokazania informacji ile jest wokół nas w określonym środowisku biurowym. Wszystko może być powierzchnia projekcyjną – ściana lub stół, i też gdziekolwiek – w biurze, lub sali konferencyjnej. Oczywiście system ten musi stawić czoła wielu wyzwaniom, ale mogą one zostać pokonane poprzez wzrost mocy sprzętu do grafiki i generalnie sprzętu komputerowego.

2. Biuro przyszłości (system projekcji ściiennej)

Przez ostatnie dwadzieścia lat, obecność technologii komputerowej wzrosła w przemyśle, biznesie i handlu w Ameryce i nie tylko. W dzisiejszym środowisku biurowym, poczta elektroniczna zajęła miejsce notatki biurowej zapisywanej na papierze (rozsyłanej między pracownikami danej firmy). Pracownicy mają komputery zamiast maszyn do pisania. Faxy i pager'y są powszechne. Ten marsz ku przyszłości jest silny i stabilny i nie widać żadnych oznak jego spowolnienia.

„Biuro Przyszłości” (OOTF) – grupa badawczo-naukowa na Uniwersytecie Północnej Karoliny w Chapel Hill poświęca swoje wysiłki na eksplorowanie tego co następne dwadzieścia lat pokaże na świecie w dziedzinie biur i sal konferencyjnych. Biuro przyszłości, jak widzą to członkowie OOTF, nie będzie składało się tylko z prostego komputera z prostokątnym ekranem. Pracownicy nie będą zmuszeni do pracowania na urządzeniach przy pomocy tych samych klawiatur i myszek, których używamy dzisiaj. Przyszłość użyje biura samego w sobie jako powierzchni projekcyjnej. Zamiast dzisiejszych wirtualnych projektorów, samo biurko będzie stanowiło duży komputer a ściany biura będą funkcjonowały jako dodatkowa powierzchnia operacyjna. Członkowie OOTF mówią: „ piksele będą wszędzie”. Biura zostaną „powiększone” za pomocą sensorów, które będą ciągle donosiły o stanie biura do komputera. Wtedy komputer będzie mógł zinterpretować te dane i ustosunkować się do nich (odpowiednio zareagować). Te same sensory mogą też być użyte do ocenienia czy odpowiednie kroki zostały podjęte. Ten zamknięty układ pozwala na kilka niezwykłych możliwości. Na przykład, kamery mogłyby obserwować pracownika co pozwoliłoby komputerowi na odpowiednie interpretowanie jego gestów wykonywanych rękoma. Mikrofony mogłyby nagrywać dźwięki obecne w biurze i mogłyby być użyte przez pracownika do wydawania poleceń głosem (poleceń dla komputera). Urządzenia śledzące ruch mogłyby zostać użyte by tworzyć trójwymiarowe grafiki z zachowaniem odpowiedniej perspektywy. Możliwości są nieograniczone. Ostatnim aspektem biura przyszłości jest tele-współpraca. Dzisiejsza globalna ekonomia oznacza, że pracownicy często znajdują się w odległych miejscach od swoich współpracowników, pracując nad tym samym projektem. Tele-współpraca pozwoliłaby użytkownikom znajdującym się w biurach na całym świecie na spotkanie się w wirtualnej rzeczywistości i pracowanie przy wymianie informacji.

Tele-współpraca używałaby tych samych sensorów i systemów projekcji jak reszta biura.

Obecny status biura przyszłości

Obecnie grupa OOTF koncentruje się na dwóch działaniach badawczych. Pierwsza część projektu dotyczy samej tele-kolaboracji (jako element NTII Narodowej Inicjatywy Teleimersji. Naukowcy OOTF zastosowali prototypowy system, który umożliwi oddalonym od siebie współpracownikom na wirtualne spotkanie. System ten używa morza kamer do wydobycia odpowiedniego koloru i głębi z obrazu biura,

potem tworzy cyfrowy opis tej sceny, transportuje go przez sieć i pokazuje na ekranie odległego komputera. Prototyp systemu został z sukcesem użyty do połączenia trzech oddalonych miejsc w jedną wirtualną konferencję. Drugi rodzaj działań koncentruje się na części projekcyjnej tego systemu. Aby zrealizować wspomniany koncept „ piksele wszędzie”, grupa OOTF zastosowała prototypowy system projekcji, który używa morza projektorów i czujników dostarczającym obraz o bardzo dużej rozdzielczości. Co jednak ważniejsze, system projekcji istniejących ścian sali konferencyjnej jako powierzchni projekcji, ma opcję automatycznego skalowania i rekonfiguracji (ponownego ustawienia). Grupa OOFT nazwała ten system „ The DisplayWall”.

System Projekcji Ściennej „The DisplayWall”

System ten jest pierwszą próbą stworzenia modelu projekcji. Nie jest to pierwszy system, który rozszerza obszar projekcji poza tradycyjne ramy CRT(cathode ray tube = ekrany katodowe) monitorów. Ten prototyp został zainstalowany w sali konferencyjnej Uniwersytetu Północnej Karoliny w Chapel Hill. System ten używa ośmiu projektorów i cyfrowych kamer o wysokiej rozdzielczości do odbierania bodźców.

System DisplayWall składa się z trzech rodzajów oprogramowania. Pierwszy moduł odpowiedzialny jest za fotometryczne skalowanie. Niestety obecne projektory wykazują duży stopień braku spójności obrazu na ekranie projekcyjnym. Na przykład, jasność (mierzona w jednostkach Lumen) w środku rzucanego przez projektor obrazu może być dwukrotnie większa od tej w rogach tego obrazu. Dodatkowo wartości Lumenów różnych projektorów też się różnią i to znacznie. Obok tych różnic w jasności obrazu , każdy projektor wykazuje brak spójności kolorów. Projektory wykorzystują trzy dyskretne kanały kolorów (czerwony, zielony, niebieski) do symulowania palety kolorów. Każdy z tych kanałów wykazuje zmiany na przestrzeni obszaru projekcyjnego. Różnice istnieją również między poszczególnymi projektorami. Celem fotometrycznego skalowania jest mierzenie i opisywanie tych różnych zmian. Użytkownicy mogą korzystać z systemu począwszy od surfowania w sieci do naukowych pokazów.

Prototyp

Grupa OOFT zainstalowała prototyp systemu DisplayWall w sali konferencyjnej. Zbudowano ten system ze standardowych, dostępnych w sklepie komponentów. Prototyp ten wykorzystuje osiem pierścieni projektorów (PTU). Każdy pierścień zawiera projektor z ciekłokrystalicznym ekranem Proxima DP6850. Każdy projektor może pokazywać obraz aż do rozdzielczości 1024x 768 przy 1500 ANSI Lumenów. Projektory Proxima pozwalają na sterowany komputerem zoom, skupienie, równowagę kolorów, jasność i kontrast. Te funkcje są dostępne przez port komunikacyjny RS-232.

Z przodu każdego projektora znajduje się PTU(element kierujący bezpośrednio odbiorem). PTUsy są sterowane komputerowo przez port komunikacyjny RS-232. PTU pozwala na dwojaką swobodę przez pozwalanie komputerowi na kontrolę kątów systemu obracającego/ kierującego (osi pionowej i poziomej). Do każdego PTU przymocowano lustro. Przez kontrolę systemu prowadzącego/obracającego i jego parametrów ustawienia, system może obrócić promień odbicia lustra w każdym kierunku. Lustro jest użyte do wycelowania obrazu projektora na ekran projekcyjny. Trzecie lustro potrzebne jest do skorygowania podczas procesu wyrównywania ostrości/jakości obrazu.

Drugim ważnym składnikiem jest moduł skalowania geometrycznego. System ten zapewnia dużą swobodę odnośnie umieszczenia i orientacji projektorów. Jeżeli cyfrowa kamera „widzi” dany obraz, projektory mogą być umieszczone gdziekolwiek w pokoju i mogą być różnie ustawione. Ta dowolność narzuca skalowanie/ kalibrowanie każdego projektora, gdyż system musi znać dokładną pozycję obrazu wyświetlanego. Przez geometryczne skalowanie, możemy umieścić wszystkie z projektorów we wspólnej skoordynowanej, zespolonej ramie. Jest to ważny krok do stworzenia spójnego obrazu.

Trzecim ważnym komponentem jest moduł odtwarzania. Odpowiedzialny jest on za użycie danych zgromadzonych w częściach geometrycznego i fotometrycznego skalowania w dostarczaniu jednolitego przekazu projektowanego obrazu. Musi to połączyć działanie ośmiu projektorów w system do odtwarzania jednolitego „bez widocznych szwów” obrazu. Te trzy elementy składają się na „serce” systemu DisplayWall. Razem umożliwią użytkownikowi pracę przy pomocy interaktywnego komputerowego środowiska z obrazem o wysokiej rozdzielczości.

Panel Sterowania

Panel sterowania działa jako centrum nerwowe systemu DisplayWall. Oprogramowanie panelu mieści się w PC (komputerze osobistym) i kontroluje funkcjonowanie projektorów i PTU. Pozwala również użytkownikom na zapisywanie obecnego ustawienia/konfiguracji dla późniejszego użytku. Na dodatek, pozwala użytkownikom na uruchomienie różnych programów skalowania.

Panel sterowania pozwala użytkownikom na kontrolowanie kilku ustawień/funkcji projektora :zoom, ostrość, jasność, kontrast, równowaga czerwonego i zielonego koloru, moc. Dzięki niemu obraz i wprowadzone dane mogą zostać wybierane poprzez standardowe przyciski i cyfry jak na odbiorniku

radiowym. Te same standardowe narzędzia pozwalają użytkownikowi na zmiany kąta ustawienia projektora. Obraz 3 pokazuje zdjęcie panelu kontrolnego.

Panel kontrolny umożliwia również zastosowanie dwóch geometrycznych programów skalowania. Pierwszy program to procedura kalkulacji matrycy. Drugi to zastosowanie skalowania światłem strukturalnym. Te dwa programy zostaną omówione szerzej w punkcie poświęconym Skalowaniu Geometrycznemu.

Fotometryczne skalowanie

Skalowanie fotometryczne jest integralną częścią systemu DisplayWall. Zmiany zarówno w kolorze i jasności pomiędzy poszczególnymi projektorami powodują powstawanie nieoczekiwanych tworów w powstałym obrazie z połączonych projektorów. Nawet brak spójności w obrębie obrazu tego samego projektora daje niepożądane efekty. Aby skorygować te zmiany, kilka obliczeń i pomiarów musi zostać dokonanych dla każdego projektora.

Pierwszym krokiem w fotometrycznym skalowaniu jest dopasowanie poziomów czerni i bieli dla każdego projektora. W warunkach idealnych projektor przybrałby wartości dla jasności od 0 do 255 i przemienił ten numer w całkowite natężenie/intensywność. Zero oznaczałoby czerń a 255 biel. Niestety projektory nie są zdolne do pokazywania prawdziwej czerni z powodu światła przesiąkającego zaciemniony panel ciekłokrystaliczny. Jasność obecna przy pokazywaniu poziomu natężenia 255 jest nazywana poziomem białym.

Pierwszym krokiem w dopasowywaniu poziomu czerni jest ustawienie jasności każdego projektora na najniższy poziom ustawienia. Potem używamy spektralnego radiometru z podziałką/skalą zdolną do dokładnego pomiaru blasku/światła aby dokonać pomiaru poziomu czerni do każdego projektora. Najjaśniejszy poziom czerni jest używany jako linia startowa, i dla wszystkich innych projektorów, jasność jest ustawiona tak by pasować do ustawionego poziomu czerni. Kiedy już wszystkie poziomy czerni są dopasowane, możemy przejść do poziomu bieli. Ustawienie jasności przy każdym projektorze może zostać użyte do zmiany powiększenia. Pierwszym krokiem jest ustawienie funkcji jasności na każdym projektorze maksymalnie do góry. Radiometr spektralny jest użyty wtedy ponownie do obliczenia blasku/światła przy natężonym poziomie o wartości 255. Najciemniejsza wartość jest użyta jako poziom startowy dla poziomu białego.

Ustawienia pozostałych projektorów zostają zrównane w wartości poprzez zrównywanie wartości poziomu bieli z poziomem startowym.

Kiedy już zarówno poziom bieli jak i czerni zostaną dopasowane, następnym krokiem jest dopasowanie wszystkich ośmiu krzywych gamma do pojedynczej kanonicznej funkcji gamma. Dla ułatwienia, definiujemy tą funkcja kanoniczną gamma jako funkcję liniową od czerni do bieli. Każdy projektor ma własną, unikalną krzywą gamma, która opisuje w jaki sposób projektor mapuje wartości zawarte w przekazie video. Ta krzywa gamma jest zazwyczaj wyrażona jako funkcja wartości całkowitej/integralnej od 0 do 255. Rysunek 4 pokazuje kształt tej krzywej. Jednym z rezultatów krzywej gamma jest to że właściwy blask przy wartości A jest nie koniecznie równy połowie wartości blasku przy intensywności równej $2xA$. Brak linearności w krzywej gamma powoduje poważne problemy w tym systemie, stąd tworzenie linearności jest integralną częścią procesu skalowania.

Przy użyciu radiometrów spektralnych, pobierane są pomiary wartości intensywności dla wszystkich 256 wartości dyskretnej w sygnale video. To pozwala na przyporządkowywanie określonego natężenia i natężenia właściwego emitowanego przez projektor. Później obliczamy tabelę odwrotności LUT która pozwoli na szybkie utworzenie linearnej krzywej gamma. Używamy LUT do mapowania wyidealizowanych wartości linearnych z wartościami realnymi używanymi przez projektor. Na skutek tego natężenie równa $2xA$ ma wartość blasku dwukrotnie większą od natężenia A . Każdy projekt ma unikalny LUT który użyty zostaje w procesie korekty krzywych gamma. Ten Lut nazywany jest „LUT korekcji gamma”.

Dodatkowy złożony poziom powstaje z nakładających się na siebie obszarów Pola Projektji. Zarówno poziom bieli i czerni są jaśniejsze dla nakładających się na siebie pikseli niż dla nie nakładających się. Są cztery rodzaje pikseli w ekranie projekcji. Nazywamy je klasami pikseli. Pierwsza klasa zawiera piksele, które nie nachodzą na inne. Druga klasa zawiera piksele, które nachodzą na jakiś inny jeden piksel. Trzecia zawiera piksele, które pokrywają jeszcze dwa inne, a trzecia klasa zawiera te piksele, które nachodzą na trzy inne. Z powodu geometrycznych ograniczeń są jedynie te cztery rodzaje pikseli.

Każda klasa musi być traktowana oddzielnie by poprawnie schować nachodzące na siebie fragmenty. Każdy LUT korekcji gamma jest powielany cztery

razy, po razie dla każdej klasy. Te nowe LUTy są dopasowywane do wzrostu poziomu bieli i czerni dla tych klas.

Geometryczne skalowanie

Następnym krokiem w skalowaniu systemu jest skalowanie geometryczne. Ta część systemu używa cyfrowej kamery do podglądu obszaru projekcji każdego projektora, i do obliczania ramy współrzędnych dla ujednoliconego systemu. Jednak zanim jakiegokolwiek skalowanie geometryczne może się odbyć najpierw kamera musi zostać odpowiednio nastawiona (zeskalowana).

Idealna soczewka nie zniekształcałaby światła gdy przechodzi przez jej soczewkę. Niestety żadne soczewki nie są idealne. Jeśli soczewka jest używana do dokładnych pomiarów, te zniekształcenia muszą zostać skorygowane. Procedura typowego skalowania kamery przybliża zniekształcenia soczewki poprzez założenie, że uproszczony model wyłapie większość tych zniekształceń. W założeniu tego modelu są również dwa typy zniekształceń. Pierwszy to zniekształcenia radialne, a drugi to zniekształcenia tangensowe.

Radialne zniekształcenie odnosi się do wykrzywiania światła jako funkcji odległości od optycznego środka soczewki. Ten typ zniekształcenia jest też często nazywany ‘beczkowym zniekształceniem’ lub ‘poduszkowym zniekształceniem’. Ten typ zniekształcenia charakteryzuje się trzema typami parametrów. Zniekształcenia tangensowe odnoszą się do położenia optycznego centrum w stosunku do centrum projektowanego obrazu.

Aby wyskalować cyfrową kamerę napisano formularz, który używa systemu kalibracji wynalezionego w Image Laboratory Intelu. Ten system oblicza pięć parametrów zniekształceń przez badanie licznych zdjęć. Przy około 15 zdjęciach obrazów próbnych o znanych rozmiarach, tych pięć parametrów może zostać dokładnie określonych. Tych pięć parametrów dokładnie opisuje wszystkie istotne wewnętrzne parametry kamery

Gdy już te wewnętrzne parametry są znane, system musi określić zewnętrzne parametry. Te parametry określają położenia i orientacje kamery odnośnie powierzchni projekcyjnej. Ponieważ ta powierzchnia w tym prototypie jest ograniczona do położenia w płaszczyźnie, prosta macierz 3×3 która może zostać użyta do pomieszczenia zewnętrznych parametrów. Więcej szczegółów na rysunku

5, ta maczyca jest określona w skali i potrzeba rozwiązać do niej osiem niewiadomych .

Aby znaleźć te niewiadome, musimy mieć 16 wartości do dodania do systemu linearnego który określa macierz. Szesnaście wartości odnosi się do ośmiu współrzędnych dwuwymiarowych. Cztery spośród współrzędnych pochodzą ze znanych punktów na płaszczyźnie, która zawiera obszar projekcyjny. Te cztery punkty zostały obliczone przed rozpoczęciem procesu kalibracji (skalowania), i są zaznaczone na powierzchni projekcyjnej za pomocą czterech współrzędnych w obrębie pola widzenia aparatu cyfrowego. Aby odnaleźć pozostałe cztery dwuwymiarowe współrzędne, zdjęcie tych zaznaczonych współrzędnych jest robione za pomocą cyfrowego aparatu. Obraz ten jest później odkształcany w taki sposób by usunąć radialne i tangensowe zniekształcenia. Współrzędne pikseli dla każdego z czterech zaznaczonych współrzędnych są zapisywane i wykorzystywane do skompletowania tych 16 wartości potrzebnych do rozwiązania systemu linearnego. Tych osiem współrzędnych określa macierz, która jest potem używana do transformacji współrzędnych pikseli na współrzędne obszaru projekcji. Współrzędne tego obszaru służą jako rama współrzędnych dla DisplayWall.

Kiedy nastąpi całkowite wyskalowanie kamery, system zaczyna algorytm skalowania światłem strukturalnym. Algorytm ten określa mapowanie dla każdego piksela projektora względem ramy współrzędnych. Algorytm ten również odpowiada za promieniste/radialne i tangensowe/odbite zniekształcenia w soczewce (obiektywie) projektora.

Algorytm światła strukturalnego funkcjonuje przez iluminowanie siatki danych na, każdy projektor. W tym systemie użyto siatki 10 na 10 „plamek Gaussa”. One są całkowicie oświetlone w środku i przechodzą w czerń wraz ze wzrostem promienia. Ten system plamek umożliwia algorytmowi na określenie położenia środka plamki z pikselową dokładnością. To jest ważne ponieważ cyfrowa kamera ma ograniczoną rozdzielczość, i jakkolwiek błąd w znalezieniu środka plamki spowoduje oczywiste zerwanie ciągłości geometrycznego obrazu.

Każdy projektor jest skalowany oddzielnie. Najpierw system zaświeca wszystkich 100 elementów na sieci i cyfrowa kamera wychwytuje obraz. Następnie, system koryguje zniekształcenia. Potem system stosuje kilka standardowych technik komputerowego obrazu aby określić środek każdej plamki na obrazie pikselowych

współrzędnych. Dokonuje się tego z dokładnością co do piksela poprzez obliczanie zważonego centrum w oparciu o przestrzenne rozłożenie natężenia tych plamek. Dla każdej plamki środkowe umieszczenie jest przenoszone na wspólny system współrzędnych poprzez zastosowanie macierzy

Do tej pory system zlokalizował 100 znaków, i obliczył położenie każdej plamki we wspólnym systemie współrzędnych. Jednak aby dokończyć proces przyporządkowywania, system musi znać współrzędne pikseli systemu dla każdej plamki. Aby tego dokonać system używa dwójkowo zakodowanego algorytmu światła strukturalnego aby zidentyfikować która plamka jest którą. Każda plamka obdarzona jest unikalnym numerem identyfikacyjnym od 1 do 100, te numery mogą być przedstawione za pomocą siedmiu bitów. System projekcji najpierw wyświetla te dane, których najmniej ważnym bitem jest 1. Cechy których najmniej ważny bit to 0 są pozostawione na czarno. Obraz zostaje pobrany kamerą cyfrową, poprawiony ze zniekształceń, i przeanalizowany aby określić które cechy zostały wyświetlone. Ten sam proces powtarzany jest jeszcze sześć razy, po jednym dla każdego bitu potrzebnego do określenia numerów identyfikacji.

Po zanalizowaniu wszystkich siedmiu podwójnie zakodowanych obrazów, system zna numery identyfikacji każdej plamki. Ten numer pozwala systemowi na określenie współrzędnych pikseli projektora dla danej plamki i dopasować te współrzędne do odpowiednich współrzędnych na powierzchni projekcyjnej. Ten proces powtarzany jest dla każdego projektora wchodzącego w skład systemu. Po ukończeniu tego procesu system ma 100 próbek w przyporządkowywaniu pikseli projektora do współrzędnych ekranu projekcyjnego.

Aby dokonać obliczeń całego przyporządkowywania, system może interpolować między pomierzonymi próbkami. Jednak system ten nie wykonuje jeszcze tej interpolacji. Zamiast tego system ten tworzy sieciowy opis powierzchni projekcji.

Ponieważ system ten zna położenie każdej próbki w przestrzeni pikselowej projektora, siatka może zostać obliczona dla każdego projektora przez znalezienie triangulacji Delaunay próbnych punktów. Dokonano triangulacji danych, ponieważ sprzęt do grafiki komputerowej dostarcza bardzo szybkiej linearnej interpolacji dla trójkątów. Podczas gdy linearna interpolacja nie jest najbardziej dokładną metodą interpolowania, odkryto poprzez eksperymentowanie że przy takiej gęstości próbkowania 100 próbek na jeden projektor, linearna interpolacja jest wystarczająco

dokładna że tylko kilka ‘tworów niechcianych’ pojawia się w ostatecznym projektowanym obrazie. Siatki te wraz ze wspólnym systemem współrzędnych, są później przesyłane do algorytmu przetwarzającego siatkę.

Następnym etapem w skalowaniu następuje odwrócenie surowych danych określonych w porcji światła strukturalnego na format który może zostać użyty do odtwarzania w realnym czasie w tym systemie. Dane określone przy użyciu plamek Gaussa zawierają wszystkie informacje potrzebne do projekcji, ale wiele zadań które mogłyby spowolnić odtwarzanie obrazu, mogą zostać ukończone przed tym jak moduł odtwarzania zacznie pracować. Pierwszym krokiem jest znalezienie prostokąta o maksymalnej wielkości. W tym prototypie, plik sieciowy zawiera osiem nakładających się na siebie siatek, po jednej na każdy projektor. Aby określić powierzchnię projekcji dla DisplayWall należy znaleźć największy możliwy obszar prostokątny, który zawierałby się w połączeniu wszystkich ośmiu siatek. Zamiast rozwiązywania tego problemu poprzez obliczanie największego pola prostokątnego, wykonujemy operację powiększania prostokąta, który przywróci lokalnie powiększony prostokąt obszarowo jest zbliżony do całkowitego obszaru powierzchni projekcyjnej. Tworzone to jest przez interaktywne zwiększanie wysokości i szerokości prostokąta, którego początek jest w samym środku zespolonych ze sobą siatek. To się powtarza aż do momentu kiedy prostokąt nie może zostać powiększony bez wyjścia poza ramy obszaru zespolonych siatek. Następnym krokiem jest określenie położenia faktury połączeń ‘szwów’. Ponieważ system ten jest zbudowany na bazie Open GL 6, trzeba znaleźć sposoby na ominięcie jej ograniczeń. Jednym z takich ograniczeń jest maksymalny rozmiar textury/faktury. Niestety, planowane jest wybudowanie ekranu projekcyjnego o dużej rozdzielczości i maksymalny rozmiar faktury nie jest wystarczająco duży. Aby usunąć tą przeszkodę, należy podzielić ekran o wysokiej rozdzielczości na pojedyncze kawałki/płytki fakturowe/strukturalne. Aby uzyskać wartości interaktywnych ram w procesie odtwarzania, dzieli się siatkę na złączach fakturowych. Tak dzieje się zawsze gdy takie dwie płytki fakturowe sąsiadują ze sobą. Siatka ta jest dzielona poprzez wstawianie dodatkowych wierzchołków do obszaru siatki przy każdym brzegu trójkąta który przecina spojenie faktury. Nowy wierzchołek jest obliczany przez linearne interpolowanie między istniejącymi wierzchołkami. Kiedy już nowe wierzchołki zostaną wstawione, siatka jest

naprawiana poprzez wykonywanie nowej triangulacji Delaunay na wierzchołkach siatki.

Trzeci etap to przydzielenie każdego piksela do jednej z klas nakładania się na siebie pikseli. Klasa taka musi zostać określona aby poprawnie ukryć nakładające się obszary i do połączenia ośmiu projektorów w jeden spójny system. Działania Boolean pomiędzy opisem poszczególnych siatek każdego projektora są użyte do wykonania tych kalkulacji, i określana jest mapa która umieszcza każdy piksel każdego projektora w jednej z czterech klas pikseli.

Ostatecznym krokiem modułu skalowania jest określenie maski alfa dla każdego piksela. Taka maska używana jest do osłabiania jasności piksel za pikselem. Dla każdego obszaru na ekranie projekcyjnym, całkowita wartość alfa musi równać się jeden. Wartości alfa dla pikseli z pierwszej klasy, gdzie nie ma nachodzenia na siebie, są sprowadzane do jeden. Dla innych pikseli alfa jest rozmieszczane po wszystkich nakładających się pikselach. Wartość alfa dla nakładających się pikseli jest obliczana przez sprawdzanie klasy danego piksela, odległości piksela od brzegu projektowanego obrazu, i wartości gęstości pikseli charakterystycznej dla danego projektora. Gęstość pikseli jest używana, ponieważ projektory o większej gęstości pikseli są zdolne do odtwarzania dokładniejszego obrazu i powinno się im dać pierwszeństwo użycia.

Po obliczeniu alfa maski modułu geometrycznego skalowania jest zakończone. Wszystkie dane zgromadzone w czasie procesu geometrycznej kalibracji są zapisywane w pliku danych do użycia dla urządzenia odtwarzającego. W czasie odtwarzania, informacje są ładowane do pamięci i używane do przyspieszenia procesu odtwarzania.

Odtwarzanie

Ta część systemu jest początkowo-końcową częścią tego systemu przywoływanego przez użytkowników systemu. Moduł ten jest oparty o VNC, urządzenia które pozwalają użytkownikom na sterowanie odległego serwera X i przesyłanie danych przez sieć do lokalnych urządzeń. Użyto kodu źródłowego VNC jako punktu startowego dla tego urządzenia odtwarzającego i jak jest to konieczne dodano specyficzny kod Displaywall.

Kiedy następuje uruchomienie, pierwszym zadaniem jakie wykonuje moduł, jest załadowanie zgromadzonych w czasie skalowania systemu danych. Dane Lut ze

skalowania fotometrycznego jest załadowywane do tabeli wyszukiwania dostępnych w wielu urządzeniach SGI. LUT dostarczony przez SGI mają głębię 12 bitów, ale informacje z tego systemu LUT mają głębie tylko o wartości 8. Używa się dwóch dodatkowych bitów jako flag wskazujących klasę pikseli dla każdego piksela. Używając dwóch bitów do określania klas pozwala na symultaniczne załadowywanie czterech LUTów dla każdego projektora. Ładujemy te cztery wklejając dane, używając tych dwóch mniej ważnych bitów do wskazania LUTu w oparciu o klasę pikseli.

Kiedy już Lut są załadowane, system jest gotowy do rozpoczęcia pracy. Obraz o wysokiej rozdzielczości jest odtwarzany do buforu poza ekranem. To umożliwia użycie standardowych serwera X i operatora okna. Serwer X otrzymuje wszystkie polecenia z ekranu i wyniki są odtwarzane na typowym prostokątnym ramie. Jedyna różnica w tym że ten ramowy bufor ma bardzo dużą rozdzielczość i nie jest przypisany do żadnego przyrządu odtwarzającego. Zamiast tego jest to prawdziwy bufor umieszczony w głównej pamięci.

Następny krok to podzielenie ogromnej faktury na mniejsze kawałki, które pasują do maksymalnych ograniczeń określanych przez Open GL. W obecnym prototypie te kawałki mają rozmiary 1024x1024. Te kawałki są później przesyłane do procedur wyspecjalizowanego odtwarzania DisplayWall.

W tej części modułu odtwarzania, następuje odtwarzanie osobno z każdego projektora. Najpierw są wydobywane fragmenty faktury z wirtualnych buforów. Następnie odtwarzana jest siatka trójkątów, jeden fragment po drugim. Mapowanie faktury (trójkątów) dokonuje linearnego zniekształcenia/wykrzywienia wirtualnej ramy buforu.

Kolejny krok to zastosowanie maski alfa obliczonej w części przetwarzania siatki. Każdy projektor ma unikalną maskę alfa. Maską ta zawiera stałą od 0 do 1 dla każdego piksela w ramce buforowej. Te stałe są mnożone przez liczbę kolorów. To pozwala na gładkie przejście między obrazami projektorów w nachodzących na siebie obszarach. Kiedy nastąpi zastosowanie maski alfa, system potrzebuje oznaczyć piksele ich właściwą klasą. Aby tego dokonać, system najpierw wypełnia bufor flagami dla każdej klasy. Flagami są te najmniej ważne bity/kawałki. Oznacza to, że każdemu pikselowi w tym buforze przypisane jest 0, 1,2, lub 3, w zależności od klasy danego piksela. Następnym krokiem jest wykonanie ukierunkowanej w stronę bitów AND pomiędzy rama buforu a `MAX_UNSIGNED_INT-3`. W wyniku

tego następuje ustawienie dwóch najmniej ważnych bitów koloru dla każdego piksela na 00.

Następny krok to wykonanie związane z bitami OR między buforem ramowym i buforem związanym z klasą pikseli. Wynik, zmagazynowany w buforze ramowym, zawiera flagę informującą o klasie pikseli w dwóch najmniej ważnych bitach i informacje o kolorze w pozostałych. Ten ostateczny bufor ramowy jest wymieniany do aktywnego buforu ramowego i wysyłany do projektorów. W czasie gdy przewód odtwarzający SGI przesyła informacje do projektorów, angażuje on też sprzęt LUT do współpracy z buforem ramowym. W wyniku tego każdy piksel wskazuje na LUT, dopasowując bity mniejsze (mniejszego rzędu) do klasy pikseli i bity duże (węższego rzędu) do odpowiedniego koloru. Każdy piksel jest przesyłany do projektora po przejściu korekty gamma i dopasowania kontrastu dla nachodzących na siebie regionów. Powstaje wtedy całkowicie zintegrowany obraz o wysokiej rozdzielczości. Zdjęcia od 6 do 8 pokazują działanie tego systemu.

DisplayWall (ścienny ekran projekcyjny)

Prototyp DisplayWall jest ekranem o wysokiej rozdzielczości zdolnym do odtwarzania obrazu w rzeczywistym czasie. Użytkownicy mogą serfować w sieci, oglądając zdjęcia w wysokiej rozdzielczości, usprawniać oprogramowanie, i dokonywać prezentacji slajdów na ekranie o rozdzielczości większej blisko osiem razy od standardowych monitorów. Niestety, te cechy nie są unikalne i ekrany o podobnych właściwościach istnieją w kilku innych laboratoriach.

Jednak, obecny prototyp wykazuje kilka funkcji nieobecnych w dotychczasowych modelach. Możliwość rekonfigurowania. Użytkownicy mogą tworzyć nowe konfiguracje używając panelu sterowania, który umożliwia całkowitą kontrolę nad projektorami i PTU. Użytkownik może też szybko przełączać pomiędzy skomplikowanymi konfiguracjami za pomocą kliknięcia myszki. Na przykład użytkownik może użyć dużego formatu ekranu dla grupowych prezentacji, lub pomniejszyć obraz dla osobistego użytku.

Kolejną unikalną właściwością systemu jest to, że proces skalowania odbywa się całkowicie w oprogramowaniu. Inne systemy wymagają żmudnych mechanicznych ustawień projektorów aby osiągnąć poprawny zapis obrazu. To może zabrać godziny a nawet dni. Ten system wykonuje skalowanie geometryczne w całkowicie zautomatyzowanym procesie przy użyciu strukturalnego światła i

cyfrowej kamery. Ponowne skalowanie DisplayWall może zostać dokonane po mniej niż 10 minutach.

Ostanią cechą systemu jest umiejętność radzenia sobie z różnorodnym natężeniem pikseli na ekranie. Pozostałe systemy wymagają aby piksele były rozmieszczone równomiernie po powierzchni ekranu. Ten system w pełni akceptuje wstawki o wysokiej rozdzielczości, które pozwalają użytkownikom na strategiczne rozmieszczenie większej liczby pikseli w miejscach potrzebnych. Na przykład, w prezentacji naukowej, naukowcy będą mogli zgromadzić więcej pikseli w danym obszarze. Co pozwoli im na obejrzenie większej liczby szczegółów tego wybranego obszaru, bez utraty obrazu całości z ekranu projekcyjnego (który będzie pokazywany przy niskiej rozdzielczości). Wcześniej naukowcy musieli wybierać między przybliżaniem obszarów z utratą całości, a powiększaniem i utratą obrazu o wysokiej rozdzielczości.

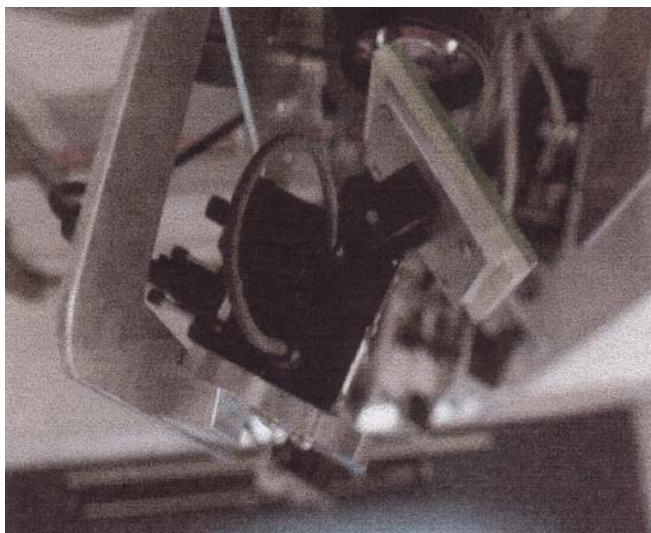
DisplayWall Przyszłości

Dużo pracy musi jeszcze zostać wykonane nad tym systemem, zanim będzie mógł się stać dojrzałą i integralną częścią Biura Przyszłości. Największą przeszkodą będzie odejście od płaskiej powierzchni projekcyjnej i używanie jakiegokolwiek powierzchni. Kolejne zadanie to poprawienie algorytmu tak aby poziom pokazu mógł wzrosnąć do punktu w którym możliwe będzie zastosowanie trójwymiarowych projekcji graficznych jak na przykład w czasie teleimersji. Przyszłym planem jest osiągnąć system od wyspecjalizowanych urządzeń SGI w kierunku sieci komputerów, a by obniżyć koszty, i sprawić aby było to rozsądne rozwiązanie dla innych laboratoriów badawczych.

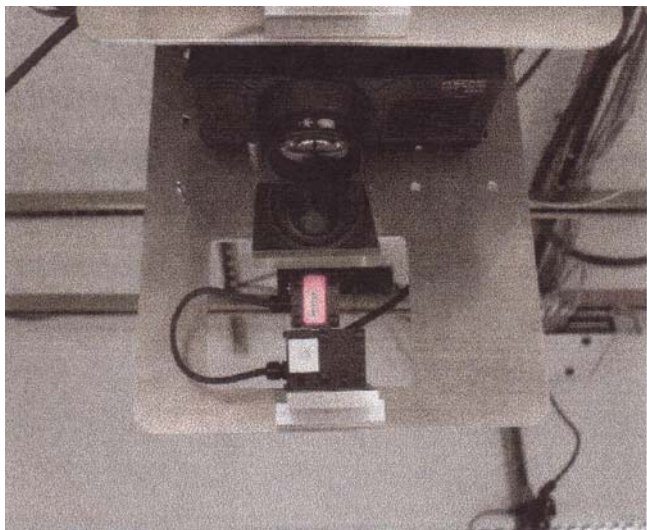
Zastosowanie tego prototypu nauczyło OOTF wiele o ekranach o wysokiej rozdzielczości i problemach z jakimi trzeba się zmierzyć przy tworzeniu Biura przyszłości. Po drodze rozwiązano kilka kluczowych kwestii i znaleziono sposoby na usprawnienie funkcjonalności systemu. Odniesiono kilka sukcesów w pracy nad tym projektem, ale wciąż jeszcze długa droga zanim zostanie zbudowany system przyszłości. Atrakcyjniejszą alternatywą jest ściągnięcie ekranu z głowy użytkownika, i zamiast tego użycie SID przestrzennego ekranu imercyjnego. Ekran ten otacza widza panoramą obrazów. Te obszary projekcyjne są zazwyczaj w rozmiarach całych pomieszczeń stąd mogą obsługiwać złożoną liczbę widzów, i są

zazwyczaj używane w połączeniu z urządzeniami projekcyjnymi które rzutują obraz z pozycji przedniej lub tylnej. Najprawdopodobniej najbardziej znanymi przykładami urządzeń SID do ogólnego zastosowania są CAVE™ Cave Automated Virtual Environment, oraz powiązane systemy z ‘kafelkowymi’ ekranami Power Wall i Infinity Wall™. Jest kilka przykładów udanej telekolaboracji, gdzie użytkownicy widzą i współpracują ze swoimi oddalonymi współpracownikami przy pomocy systemu CAVE™ lub mu podobnych. Takie ogromne systemy wymagają znacznej przestrzeni podłoża, na przykład w laboratorium gdzie jest zarówno miejsce na ekrany jak i na urządzenia projekcyjne. Ten system powinien działać na wzór telefonu – SID wbudowane w życie biura. Podczas gdy taki wysiłek nie zwróciłby się w zyskach jeśli użyty jest tylko do telekolaboracji, jeśli jednak byłoby to wysokiej jakości można by tego używać do codziennej dwuwymiarowej pracy przy komputerze, video, i obrazach immersyjnych trójwymiarowych. Jednak konstrukcja takich SID nie tylko wymaga bardzo ostrożnej budowy i zmontowania, ale jeszcze pewne dane charakterystyczne zmieniają się w zależności od czasu i czynników środowiska jak temperatura lub wibracje. Takie zmienne w czasie czynniki obejmują wewnętrzne i zewnętrzne parametry projektora, równowagę intensywności i koloru, wyrównywanie brzegów, i zespalanie obrazu. Te problemy są najczęściej rozwiązywane przez okresowe skalowanie projektorów, jednak to podejście staje się niezwykle trudne i mniej niezawodne jeśli wzrasta liczba projektorów. Urządzenia symulujące lot borykają się z tego typu problemami od pewnego czasu, i choć rozwinęły one systemy cyfrowego skalowania, systemy te zazwyczaj są wysoce wyspecjalizowane, co zwiększa koszty tworzenia i komplikuje generalnie cały projekt. System samo skalowania o szerszym zastosowaniu byłby lepszy bo każdy mógłby modyfikować obszar projekcyjny i konfigurację projektorów w zależności od potrzeb. Gdyby można by było modyfikować powierzchnie projekcyjną można by spontanicznie włączać dodatkowe obrazy i system włączy je do projekcji. Jeśli można by mieć swobodę przy mocowaniu projektora, można by dodawać projektory tak aby rzucany przez nie obrazy nakładały się, co pozwoli na zwiększenie rozdzielczości ekranu, intensywności obrazu, lub na zwiększenie generalnie pola projekcji.

Rysunek 1

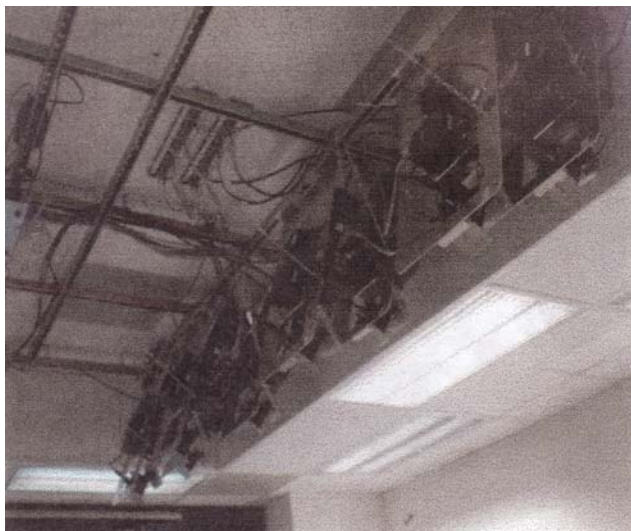


Boczny widok projektora/ platforma PTU



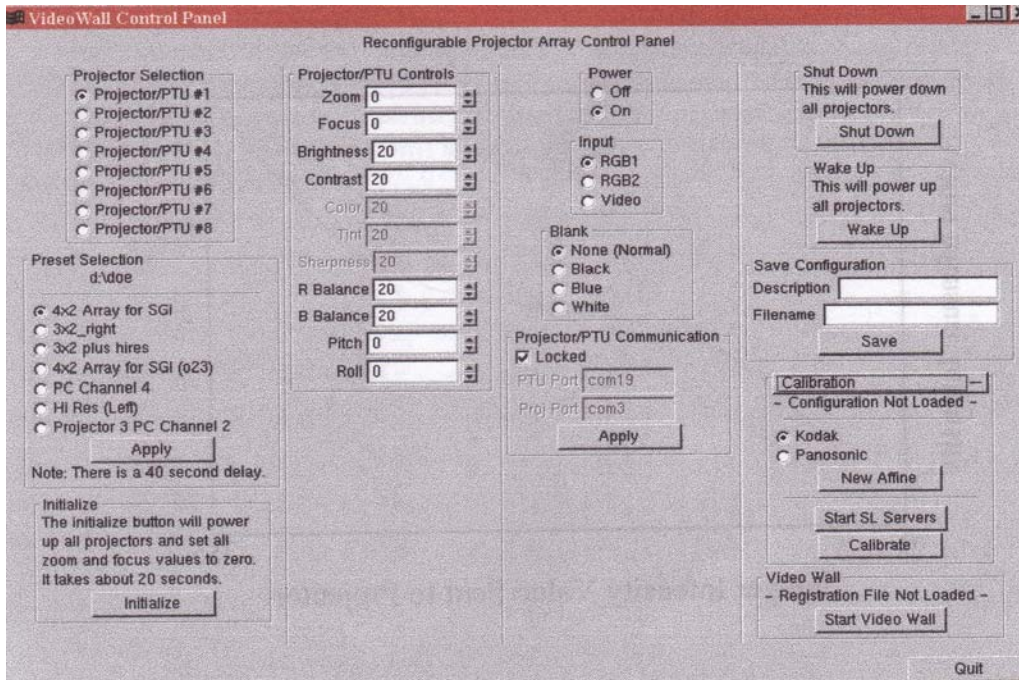
Widok projektora/ platformy PTU od frontu

Rysunek 2



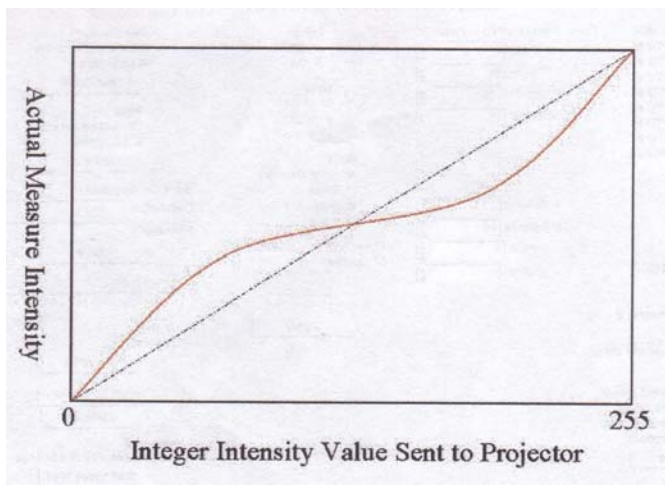
Osiem projektorów/ platform PTU zawieszonych pod sufitem

Rysunek 3



Ujęcie z ekranu monitora panelu sterowania Systemu Projektacji Ściennej

Rysunek 4



Wykres powyżej pokazuje typową krzywą gamma. Krzywa gamma przedstawia mapowanie pomiędzy wartością intensywności wysłaną do projektora a aktualną intensywnością światła emitowanego przez ten projektor. Aktualna intensywność jest mierzona przez użycie skalibrowanego czujnika, zwanego radiometrem spektralnym(miernik promieniowania).

Rysunek 5

$$\begin{bmatrix} x_1 & y_1 & 1 & 0 & 0 & 0 & -x'_1 \cdot x_1 & -x'_1 \cdot y_1 \\ 0 & 0 & 0 & x_1 & y_1 & 1 & -y'_1 \cdot x_1 & -y'_1 \cdot y_1 \\ x_2 & y_2 & 1 & 0 & 0 & 0 & -x'_2 \cdot x_2 & -x'_2 \cdot y_2 \\ 0 & 0 & 0 & x_2 & y_2 & 1 & -y'_2 \cdot x_2 & -y'_2 \cdot y_2 \\ x_3 & y_3 & 1 & 0 & 0 & 0 & -x'_3 \cdot x_3 & -x'_3 \cdot y_3 \\ 0 & 0 & 0 & x_3 & y_3 & 1 & -y'_3 \cdot x_3 & -y'_3 \cdot y_3 \\ x_4 & y_4 & 1 & 0 & 0 & 0 & -x'_4 \cdot x_4 & -x'_4 \cdot y_4 \\ 0 & 0 & 0 & x_4 & y_4 & 1 & -y'_4 \cdot x_4 & -y'_4 \cdot y_4 \end{bmatrix} \cdot \begin{bmatrix} a_{1,1} \\ a_{1,2} \\ a_{1,3} \\ a_{2,1} \\ a_{2,2} \\ a_{2,3} \\ a_{3,1} \\ a_{3,2} \end{bmatrix} = \begin{bmatrix} x'_1 \\ y'_1 \\ x'_2 \\ y'_2 \\ x'_3 \\ y'_3 \\ x'_4 \\ y'_4 \end{bmatrix}$$

$$C = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & 1 \end{bmatrix}$$

Następujący system linearny definiuje kolineację macierzy między dwie grupy czterech dwuwymiarowych punktów. Rozwiązanie dla systemu linearnego stanowi macierz 3x3, C, ta sporządza mapę współrzędnych (x_i, y_i) do współrzędnych (x'_i, y'_i) . C jest określona dodanej wielkości, dlatego tylko osiem wartości musi być obliczonych. Cztery współrzędne w poszczególnych współrzędnych systemu musi być znanych, żeby rozwiązać wartość ósmą – niewiadomą $a_{i,j}$.

Rysunek 6



System projekcji ściennej pokazuje Deklarację Niepodległości. Trzy strony pokazane tu są wzięte z Biblioteki Kongresu. Każda strona jest za duża, żeby pasowała do standardowego monitora, ale system projekcji ściennej pokazuje trzy strony jednocześnie z wolnymi pikselami.

Rysunek 7



Dwóch użytkowników bada zeskanowany o wysokiej rozdzielczości banknot dwudziesto dolarowy.

Rysunek 8



Użytkownicy porównują mikrodruk na prawdziwym banknocie dwudziestodolarowym z wirtualnym. Mikrodruk na banknocie wyświetlonym na ścianie dużo łatwiej odczytać niż na prawdziwym banknocie.

3. Projekt “Being There”

Projekt Being There jest częścią badań nad Biurem Przyszłości.

Wprowadzenie

Naukowcy pracują nad nową metodą opartą o projekt do wizualizacji utworzonych prawdziwych lub wyobrażonych miejsc. Zasadnicza różnica między tym podejściem a poprzednimi to że w tym przypadku użyto naturalnej wielkości

ekranów projekcyjnych, które zbliżone są rozmiarami do projektowanego nań obrazu. Na skutek tego powstaje środowisko wirtualne które jest zarówno wizualnie jak i przestrzennie realne, i zapewnia użytkownikowi silnego poczucia immersji.

Celem jest odtworzenie realnych ale oddalonych miejsc, z jednoczesnym sprawieniem by użytkownik miał realne poczucie obecności w miejscach rzeczywistych. Należy umożliwić ludziom na doświadczanie odtwarzania sławnych miejsc jak Monticello, dom prezydenta Thomasa Jeffersona. Obecnie prace wykonywane są w laboratorium badawczym, ale przyszłe plany to wykorzystać przestrzeń muzeum ze zmianami wystaw prezentowanych z odległych miejsc.

Raskar zbadał użycie projektorów do wyświetlania prawdziwych przedmiotów poprzez projektowanie obrazów modeli komputerowych, które znacznie przypominają geometrie tych przedmiotów. Poszerzony został zamysł wirtualnego środowiska o rozmiarach świata realnego, przy użyciu powierzchni projekcyjnych umożliwiających projekcję obrazu bez zmniejszania.

Zalety i Ograniczenia

Użycie obszarów projekcyjnych, które rozmiarowo przypominają rozmiar prezentowanych obrazów w rzeczywistości ma wiele zalet nad metodami tradycyjnymi, tak jak HMD VR i CAVR™. Podobnie jak CAVE™ ta metoda oferuje obraz o szerokim polu widzenia, które wypełnia również obszar peryferyjny zakresu widzenia użytkownika. Obraz peryferyjny jest konieczny do utrzymania orientacji przestrzeni podczas nawigowania w wirtualnym środowisku i zapewnia lepsze poczucie immersji.

Fizyczne ustawienie powierzchni projekcyjnych pozwala użytkownikowi na poruszanie się w rzeczywistości wirtualnej. Chodzenie rzeczywiste daje użytkownikowi silniejsze poczucie obecności niż takie wirtualne latanie i poruszanie się (pozostając tak naprawdę w miejscu), ale dzieje się tak kosztem wykorzystania znacznie większej przestrzeni.

Ogólnie, odmienne stopnie podobieństwa geometrii projektowanej sceny powodują powstanie spektrum powierzchni projekcyjnych w zakresie od pojedynczych płaskich ekranów do powierzchni projekcyjnych dokładnie odpowiadających geometrii odtwarzanej sceny. Kiedy obszary projekcyjne są znacznie odmienne od geometrii obrazu (jak w przypadku CAVE™) wrażliwość na błędy w odbiorze

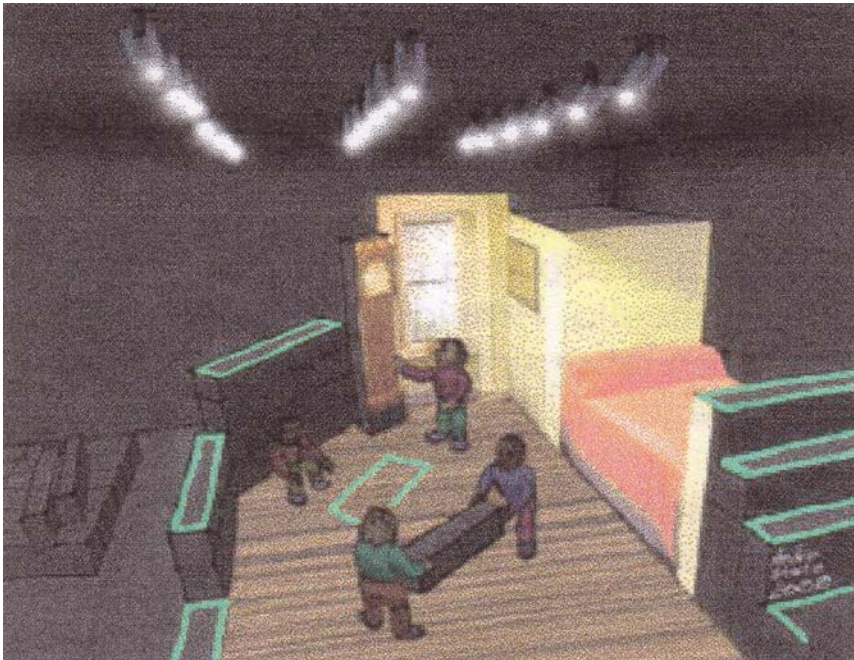
(spowodowana przez uśpienie systemu, błędy w skalowaniu projektora i urządzenia śledzącego ruch...) jest duża. Jeśli jednak powierzchnie te obszarowo odpowiadają geometrii obrazu, wrażliwość na takie błędy jest zminimalizowana.

W najlepszym przypadku, metoda ta w sposób skuteczny oferuje auto-stereo obraz, który może obsługiwać większą liczbę osób których ruch nie jest śledzony przy jednoczesnym badaniu tej samej sceny.

Jednak budowanie takiej samej, z uwzględnieniem szczegółów, repliki obrazu jest z reguły nie możliwe. Dlatego użyto uproszczonego zestawu powierzchni projekcyjnych, które pasują do geometrii pokazywanego obrazu. Na przykład, początkowe struktury budowy wnętrza i obiektów architektonicznych średniej budowy (ścian, kolumn, szafek, stołów...) może być z reguły przybliżone przy pomocy prostych kształtów (sześciokątów, walców...). Ceną za to jest ponowne wprowadzenie pewnego uśpienia w reakcji obrazu, i częściowa utrata auto stereo-obrazu. Ruch oczu użytkownika musi być śledzony.

Kolejnym znacznym ograniczeniem tej metody jest fakt że wymaga ona ogromnej fizycznej przestrzeni aby wiernie odtworzyć duże sceny. Ta metoda nie jest również odpowiednia do obrazów dynamicznych, które posiadają duże poruszające się przedmioty (na przykład samochody). W końcu, mimo że naukowcy posiadają pomysł na umieszczenie projektorów w takich miejscach aby uniemożliwić zasłanianie obrazu przez użytkownika, uważamy za naturalne chęć podchodzenia i dotyknięcia powierzchni projekcyjnych. W planowanym ustawieniu muzeum, może

być koniecznym odgródzenie obszaru w pobliży ekranów.



Wymagania odnośnie dużej przestrzeni fizycznej

Obecnie prac przeprowadzane są w laboratorium badawczym, ale planuje się użyć wnętrza o rozmiarach magazynu, w którym będzie można odtworzyć większość scen studyjnych.

Status

Dokonano pewnych wstępnych eksperymentów, w których geometria sceny jest przybliżana poprzez bloki styropianowe, i metoda dwupasmowego odtwarzania użyta jest do generowania poprawnych obrazów w odniesieniu do pola widzenia użytkownika. Rezultaty są zachęcające.

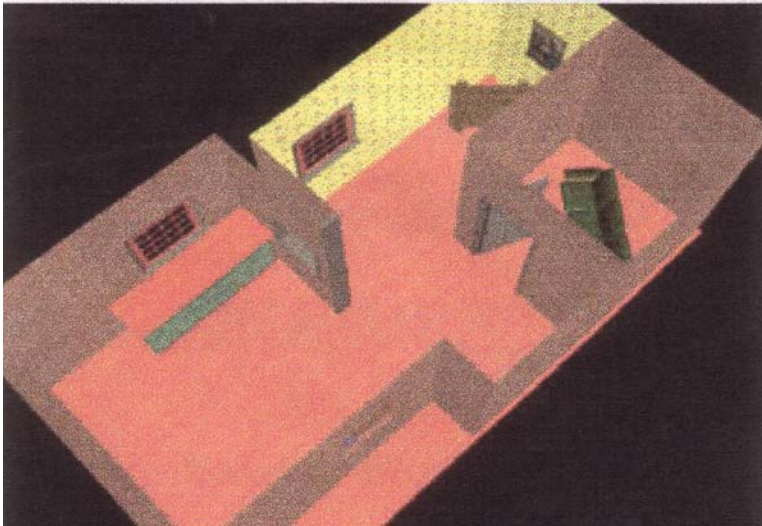
Obecnie tylko zademonstrowano pomysły za pomocą scen wymyślonych, ale zebrano i przygotowuje się modele tworzone w oparciu o obraz o bardzo wysokiej jakości z biblioteki Monticello. Uzyskano te modele przy użyciu laserowego skanera 3RD TECH podczas kilku dniowej wycieczki do Monticello.

Wyzwania

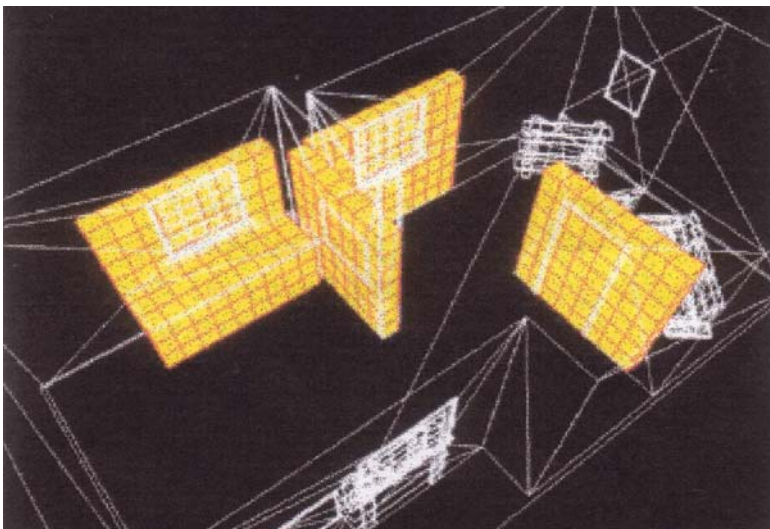
Należy utworzyć model i zbudować dobry ekran projekcyjny co jest wyzwaniem nie spotykanym w innych metodach opierających się o prace projektorów. Kolejnymi problemami tej metody są systemy odtwarzania posługujące się złożoną liczbą projektorów. Problemy te obejmują cienie, wewnętrzne odbicie, nakładanie się projektowanych obrazów, dalekie od ideału właściwości obszarów projekcyjnych,

ustawienie i skalowanie projektorów, różnoraki podgląd, i zarządzanie odtwarzaniem. Ponadto, aby odtworzyć prawdziwe miejsca, potrzebne są efektywne metody do pobierania danych, przetwarzania ich, przechowywania i tworzenia użytecznych do interaktywnego odtwarzania.

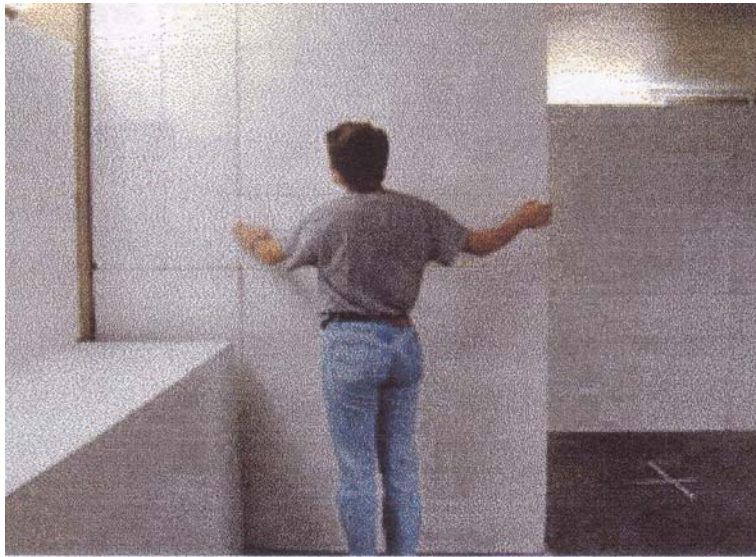
Wstępne rezultaty



Najpierw potrzebny jest model sceny. Tutaj pokazany jest syntetyczny model utworzony przy użyciu oprogramowania do tworzenia modeli. Dla modeli realnych miejsc możemy użyć skanera zakresowego w celu uzyskania informacji o geometrii i kamery do uchwycenia kolorów.

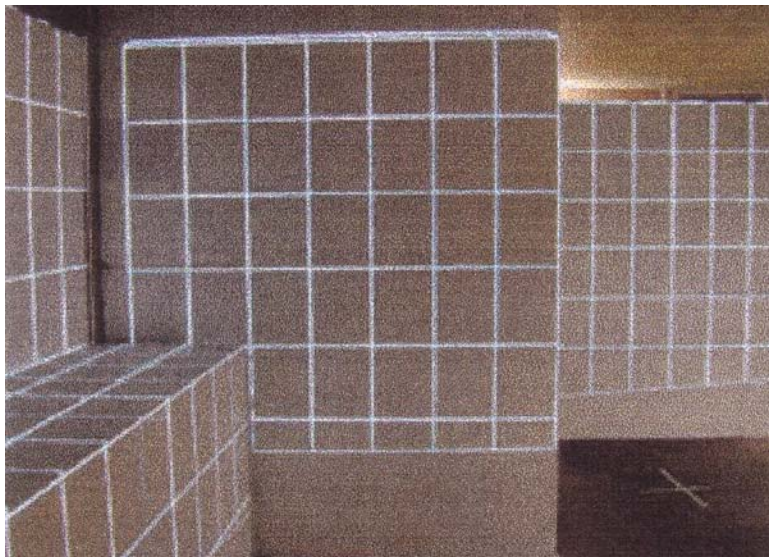


Potem tworzy się model powierzchni projekcyjnej przy użyciu jednego z oprogramowań do ułożenia prymitywnych bloków blisko powierzchni modelu, danego obrazu, który chcemy odtwarzać. Rozmiar bloków jest taki sam jak realnych bloków styropianowych których używa się w tym eksperymencie.

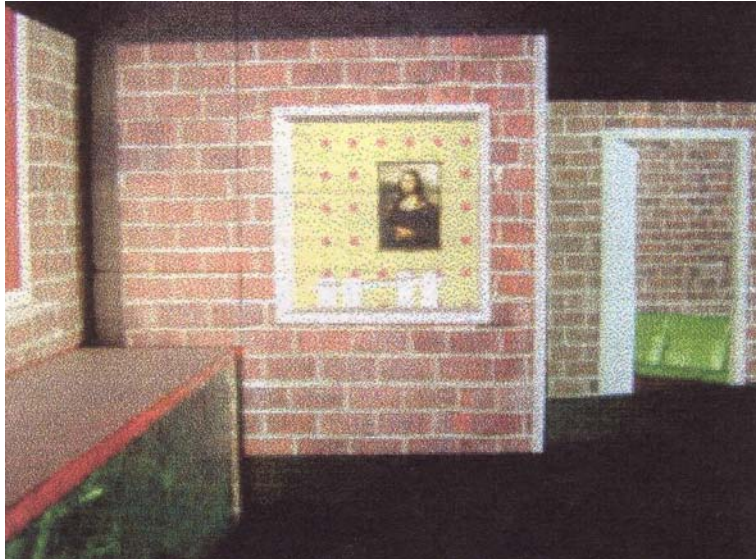


Powierzchnie projekcyjne są konstruowane w laboratorium, zgodnie z modelem tej powierzchni, przy użyciu przemysłowych bloków styropianowych. Bloki te zostały ufundowane przez Redi-Form, Inc.

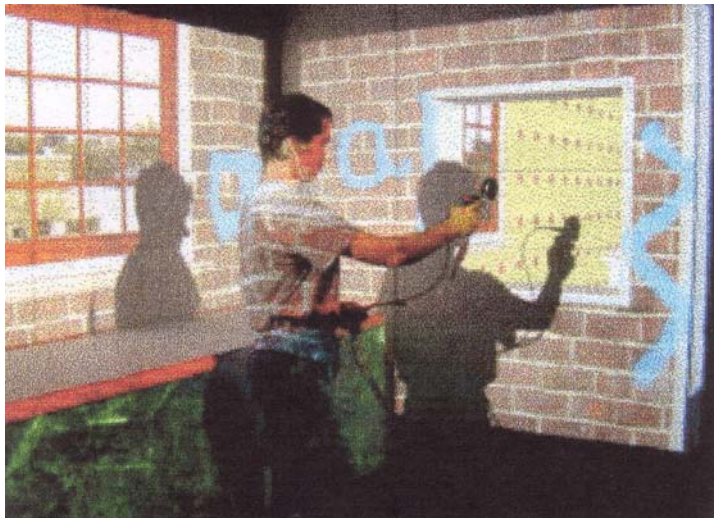
Użyto sześciu projektorów i dwóch 3RDTECH urządzeń do śledzenia ruchu w tym eksperymencie. Tych sześć projektorów jest najpierw skalowanych zgodnie z ramą odniesienia wspólnego urządzenia śledzącego ruch.



Gdy projektory są już skalibrowane, ramka modelu obszaru projekcyjnego jest rzucana na bloki styropianowe. Używa się tej metody do dostrojenia i oszacowania układu bloków styropianowych.



Ponieważ obrazy na powierzchni projekcyjnej powinny być zależne od pola widzenia, trzeba śledzić ruch użytkownika i jego położenie. Metoda dwupasmowa, używająca mapowania faktury, używana jest do generowania poprawnego obrazu dla użytkownika.



Użytkownik może również współpracować z wirtualnym środowiskiem. Tutaj użyto drugiego HiBall urządzenia do śledzenia ruchu do śledzenia pozycji wirtualnego sprayu. Użytkownik na zdjęciu maluje spray'em na sztukaterii okiennej, która nie jest fizycznie dostępna.

Ten przykład pokazuje użyteczność zdolności poruszania się wokół wirtualnych obiektów, co daje użytkownikowi naturalność w przestrzennym pracowaniu ze środowiskiem wirtualnym.

Dioramy o rzeczywistym kształcie powstałe w oparciu o obraz.

Zaprezentowany będzie w tym rozdziale pewien pomysł i wstępne wyniki nowej metody wykorzystującej funkcję projektora do odtwarzania prawdziwych i

wyimaginowanych miejsc. Celem jest osiągnięcie takich odtworzeń, które będą zarówno wizualnie jak i przestrzennie realistyczne, zapewniając małej liczbie relatywnie nieobciążonych użytkowników silne poczucie immersji jak razem poruszają się po wirtualnym miejscu.

Zamiast używania osadzonych na głowie albo uniwersalnych systemów odtwarzania wykorzystujących projektor, skorzystano z doświadczeń pracy nad rzeczywistością wizualnie powiększoną i lampami cieniowymi. Przy użyciu prostych białych bloków skonstruowano statyczny model, który przypomina kształt, rozmiar i ułożenie przestrzenne danego miejsca. Potem projektuje się na bloki obraz dynamiczny przekształcając martwy model w wizualnie wierną reprodukcję danego miejsca. Pewne zalety tej metody to obraz szerokiego pola widzenia, możliwość poruszania się po danym miejscu, zmniejszona wrażliwość na błędy w śledzeniu ruchu, zmniejszona wrażliwość na uśpienie systemu, auto-stereoskopowy obraz.

Wprowadzenie

Wielu z nas odnajduje wspomnienia o miejscach, które kiedyś odwiedziliśmy, być może jako dorośli, być może jako dzieci. Kombinacja obrazów, dźwięków, zapachów, i doświadczenie przestrzeni często pozostawiają długotrwałe piętno w pamięci. Muzea i parki tematyczne często próbują wywołać podobne uczucia poprzez fizyczne odtwarzanie miejsc, które są odległe lub nawet uległy zatraceniu. Na przykład, Centrum Walta Disneya w Orlando na Florydzie (USA) zawiera obszar „Gablota Świata”, gdzie małe kawałki miast z całego świata są odtwarzane z fasadami jak w filmach i pracownikami, którzy są właściwymi etnicznymi mieszkańcami tych miast. Po drugiej stronie tego realno-wirtualnego spektrum są czysto wirtualne przykłady jak Londyński Podróżnik, który posługuje się wirtualną rzeczywistością aby pozwolić turystom na zwiedzanie Londynu przed właściwym doń przyjazdem.



Rys. 1 Użyto bloków styropianowych do skonstruowania geometrycznie uproszczonej repliki sceny. Bloki te użyte jako powierzchnia, na którą wyświetlane są obrazy zależne od tego jaka scena jest wyświetlana. W uproszczonej replice pominięta została większość z geometrycznych złożoności sceny, takich jak ramy okna, parapety.

Naukowcy szukają czegoś pomiędzy całkowicie realnymi i wirtualnymi przykładami, taki przestrzennie realistyczny i wizualnie wirtualny hybryd. Można będzie cieszyć się bogactwem, elastycznością i dynamiczną naturą grafiki komputerowej, przy jednoczesnym zachowaniu pewnych myślowych i fizycznych czynności powiązanych z chodzeniem i badaniem realnego miejsca. Umożliwi to użytkownikom realistyczne poczucie obecności zarówno w sztucznym jak i realnym świecie. Obecnie pracuje się nad tym w laboratorium, naukowcy oczekują dnia kiedy muzea wyznaczą obszar do dioram sławnych miejsc jak na przykład Monticello, USA, dom prezydenta Thomasa Jeffersona w Charlstonville w Virgini.

Dioramy o rzeczywistym kształcie powstałe w oparciu o obraz

Osadzone na głowie odtwarzacze (HMD) są powszechnie używane do immersyjnej wizualizacji wirtualnego środowiska. Z funkcją śledzeni ruchu głowy, HMD może dostarczyć stereo obraz wirtualnego świata w rzeczywistym świecie i zależny od widza. Niestety, widok przez HMDsy jest zazwyczaj bardzo wąski, i może obniżać zdolność do skutecznego nawigowania w środowisku, i zmniejszać poczucie obecności w środowisku wirtualnym. Ponadto, HMDsy typowo zasłaniają większość czy nawet cały obraz peryferyjny realnego świata i bliskich ludzi, co sprawia, że użytkownik niechętnie się porusza.

System wizualizacji oparty o odtwarzanie projektorów jak CAVE™ otacza użytkownika ogromnymi ekranami, które są przymocowane w środowisku (ale nie na użytkowniku). Ekran są oświetlane obrazem stereo o poprawnym obrazie stereo z projektorów świetlnych umieszczonych za ekranami. Metoda ta poprawia pole widzenia, ogranicza ona obszar (który typowo ma wymiary 3x3x3 metry sześcienne), po którym użytkownik może swobodnie się poruszać w czasie projekcji przy eksplorowaniu środowiska wirtualnego. Umieszczenie ekranów daleko by zamknąć wystarczającą ilość przestrzeni do swobodnego poruszania się nie koniecznie jest dobrym rozwiązaniem z wielu powodów. Na przykład, kiedy oglądający jest w środku CAVE patrząc na średniej wielkości przedmiot bardzo blisko niego, oddalone ekrany muszą być bardzo wysokie aby umożliwić projekcję całego obrazu danego przedmiotu. W tym rozdziale zaprezentowana będzie nowa metoda oparta o użycie projektora używaną do rekonstrukcji realnych i sztucznych miejsc. Ta metoda została zainspirowana przez wystawę Michaela Neimark 'Displacement' w 1984 r. Użył on obrotowej kamery wideo aby sfilmować nieruchomy pokój gościnny, pełen mebli i nieruchomych ludzi. Pokój i meble zostały później pomalowane na biało, a uchwycony obraz został rzucony z powrotem na ściany przy użyciu obrotowego projektora który został dokładnie zgrany z kamerą. Jest to interaktywny model, który sięga poza i generalizuje zamysł Neimarka. Używając prostych białych bloków budowlanych konstruuje się statyczny model, który przypomina kształtem, rozmiarem, i układ przestrzenny oddalonego (sztucznego) miejsca. Potem projektuje się dynamiczny (ruchomy) obraz na te bloki, przekształcając 'martwy' fizyczny model w wierną wzrokowo reprodukcję tego miejsca. Poza zwiększaniem przestrzennego realizmu wirtualnego środowiska, ta metoda ma wiele innych zalet nad zastosowaniem tradycyjnego osadzonego na głowie wyświetlacza typu CAVE™.

Przeprowadzono pewne wstępne eksperymenty przy użyciu sztucznie stworzonej scenerii, i rezultaty są zachęcające. Podczas gdy następuje postęp w pracach nad osiągnięciem celu – do odtwarzania realnych miejsc jeszcze nie istnieje całkowity system. Zaprezentowana mieszanina zademonstrowanych wyników nowej metody i możliwe do zrealizowania zamysły w przyszłości. Trzeba odróżniać między tymi dwoma na wypadek gdyby „czytelnik” skłonny był wierzyć, że zastosowano już coś czego być może jeszcze nawet nie skonstruowano.

W następnej części zaprezentowany jest podstawowy zamysł stojący za tą metodą i rozważone są jego zalety i ograniczenia. Następna kwestia to zastosowania systemu opartego o tą nową metodę, i wyzwania.

Nowe podejście /Nowa metoda.

Oczywiście jednym ze sposobów na realistyczne odtworzenie miejsca jest wybudowanie dokładnej kopii, łącznie ze wszystkimi detalami, kolorami i właściwościami odbicia/fotometrycznymi. Za wyjątkiem stałych wyznaczonych miejsc ('Gablota Świata' Disneya). Jest to właściwie nie praktyczne.

Zamiast tego ta metoda była ściśle oparta o obserwację, że gdy oświetlony jest fizyczny przedmiot białym światłem, jego powierzchnia odbija fale świetlne o określonej długości które odbieramy jako kolor. Z kolei pewne fizyczne atrybuty mogą zostać efektywnie włączone do źródła światła w celu osiągnięcia efektu zmysłowo odpowiadającego przy użyciu kolorowego światła i białego obiektu. Z lampami cieniującymi(cieniowymi) tylko geometryczna struktura obrazu musi zostać powielona, a cyfrowe projektory świetlne wypełniają detale powierzchni i kolor. Struktury geometryczne realnych miejsc są zazwyczaj skomplikowane i tworzenie dokładnych geometrycznych replik może okazać się karkołomnym/zniechęcającym zadaniem. . Zamiast tego ta metoda wykorzystuje geometrycznie uproszczoną kopię danej sceny jako powierzchni projekcyjnej. Ta uproszczona powierzchnia jest łatwiejsza do zaplanowania i zbudowania. Poza wypełnianiem szczegółów kolorem tej powierzchni, używa się projektorów do odtwarzania obrazu w zależności od widza, gdzie należy nadrobić różnice geometryczne między obrazem wirtualnym a uproszczonym fizycznym modelem (powierzchnia projekcyjną).

Zastosowanie lamp cieniujących(cieniowych) jest odpowiednie do tego celu. Na przykład, podstawowe struktury w budowaniu wnętrz i średniej wielkości obiektów architektonicznych (ścian, kolumn, szafek, kredensów, stołów...)może zazwyczaj zostać przybliżone przy pomocy prostych komponentów (sześciątów, walców..). Główne cechy architektoniczne, które odpowiadają uproszczonemu fizycznemu modelowi zatrzymują trójwymiarowy obraz autostereoskopowy, podczas gdy każdy inny szczegół musi być pokazany poprzez rzutowanie obrazu zależnego od pola widzenia.

Ogólnie, różne stopnie przybliżenia geometrii sceny tworzą spektrum powierzchni projekcyjnych od pojedynczych płaskich ekranów jak CAVE™ do obszarów projekcyjnych, które dokładnie odpowiadają geometrii danej sceny. Dla osiągnięcia celu, należy uprościć obszar projekcyjny jak to tylko możliwe przy jednoczesnym zachowaniu pustej przestrzeni gdzie użytkownik może chodzić, schylać się... jednak jest wiele zalet w posiadaniu powierzchni projekcyjnych tak podobnych do geometrii sceny jak to tylko możliwe.

Zalety

Mimo że początkową motywacją to dostarczenie większego realizmu przestrzeni dla użytkownika, ta metoda ma wiele innych zalet, które są również kluczowe dla osiągnięcia celu odtwarzania realnych miejsc. Niektóre z tych zalet są wymienione poniżej.

Obraz o szerokim FOV i peryferyjny

Ludzki wzrok ma efektywne pole widzenia FOV, które obejmuje około 200 stopni horyzontalnie na 150 stopni pionowo. Wiele spośród dostępnych w handlu HMDsów ma relatywnie wąskie pole widzenia, wahające się między 30 a 70 stopni po przekątnej. Wąskie FOV zostało pokazane (w prawdziwych środowiskach) w celu pokazania ludzkiej słabości w nawigacji, manipulacji, świadomości przestrzennej, i aby zakłócić naszą koordynację oczu i głowy, oraz percepcję kształtu i przestrzeni. Obraz peryferyjny jest potrzebny do podtrzymywania samo-orientacji podczas ruchu, i są dowody sugerujące że wąskie pole widzenia umniejsza zdolność do nawigowania w środowisku efektywnie, i zmniejsza poczucie obecności w przestrzeni wirtualnej.

Jedną z wielu motywacji systemu CAVE i tej metody jest dostarczenie obrazu o bardzo szerokim FOV. Uzyskuje się to poprzez użycie ogromnych ekranów projekcyjnych, które są przymocowane w środowisku w pewnej odległości od użytkownika. Takie postępowanie ma również inne zalety nad HMD – użytkownik już nie potrzebuje nosić na sobie i swojej głowie ciężkiego sprzętu, a szybkie obroty głowy już nie wytwarzają poważnego efektu pływania.

Prawdziwe Chodzenie

Z tą metodą, fizyczna aranżacja powierzchni projekcyjnych pozwala użytkownikowi na prawdziwe chodzenie w środowisku wirtualnym. Prawdziwe chodzenie daje silniejsze poczucie obecności niż chodzenie w miejscu i wirtualne latanie, ale dzieje się tak kosztem wykorzystania dużej fizycznej przestrzeni.

Z systemami takimi jak CAVE, środowisko wirtualne, które użytkownik bada może być większe niż zamknięta przez ekrany powierzchnia. W takim przypadku jednak, użytkownicy mogą jedynie nawigować przy użyciu mniej naturalnych metod jak chodzenie w miejscu lub wirtualne latanie.

Zmniejszona Wrażliwość na Błędy w śledzeniu ruchu

Podobnie jak ustawienia projektorów generalnego użytku jak CAVE™, gdzie powierzchnia projekcyjna jest umocowana w środowisku raczej niż do użytkownika (jak w przypadku HMD), ta metoda jest relatywnie mało wrażliwa na błędy w oszacowaniu orientacji głowy. Jednak w przeciwieństwie do systemów używających projektorów generalnego użytku, ta metoda jest również relatywnie mało wrażliwa na błędy w oszacowywaniu pozycji głowy.

Diagram na rysunku 2 ilustruje geometrię ustawienia projektorów generalnego użytku z lotu ptaka. To zostanie pokazane do zilustrowania błędów w powierzchni horyzontalnej. Błąd pionowy zostałby wydobyty w podobny sposób. Punkt V reprezentuje pewien punkt na graficznym modelu to znaczy punkt na powierzchni pionowej. Jeśli urządzenie śledzące donosi, że punkt widzenia jest w E_1 , to następny odtwarzany obraz punktu V pojawi się na fizycznej powierzchni projekcyjnej w punkcie D_1 . jednak jeśli punkt oka jest umieszczony w punkcie E_2 , wtedy obraz punktu V powinien być odtworzony na powierzchni projekcyjnej w punkcie D_2 . W wyniku tego, użytkownik doświadczy błędu obrazu kąтового θE .

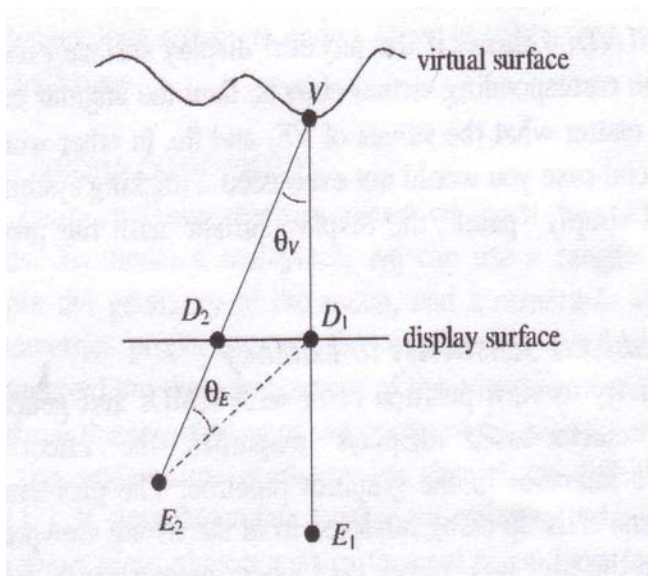
Należy zauważyć że kształt powierzchni projekcyjnej jest nieistotny; ponieważ punkt D_2 nie jest odtwarzany, nie liczy się gdzie powinien się pojawić (gdziekolwiek na linii VE_2), użytkownik wciąż będzie doświadczał tego samego błędu, θE . Podobnie absolutna pozycja punktu oka nie jest czynnikiem krytycznym, co liczy się to względna pozycja oka odnośnie odtwarzanego punktu D_1 . Tak na prawdę błąd kątowy θE zależy tylko od trzech parametrów: VE_2 , odległość od punktu wirtualnego obrazu do punktu oka użytkownika; VD_1 odległość od wirtualnego obrazu do fizycznej powierzchni projekcyjnej; i θ_v , kąt pomiędzy

obliczonym i rzeczywistym położeniem oka. Zważywszy na te trzy parametry i pewne trygonometryczne obliczenia, błąd kątowy może zostać obliczony jako:

$$\theta_E = \arcsin \left(\frac{\sin(\theta_V) |VD_1|}{\sqrt{|VD_1|^2 + |VE_2|^2 + 2|VD_1||VE_2| \cos(\theta_V)}} \right)$$

Te dwa wykresy powierzchni na rysunku 3 ilustrują błąd kąтового widzenia przez przestrzeń przybliżoną rozmiarami i kształtem do CAVE™, dla odległości między powierzchniami projekcyjnymi VD_1 od punktu wirtualnego w odległym rogu powierzchni (0,0) 10 i 50 cm. Aby pokryć tą powierzchnię θ_V waha się pomiędzy 0 a 90 stopni, a VE_2 waha się między 0 i 2 metry.

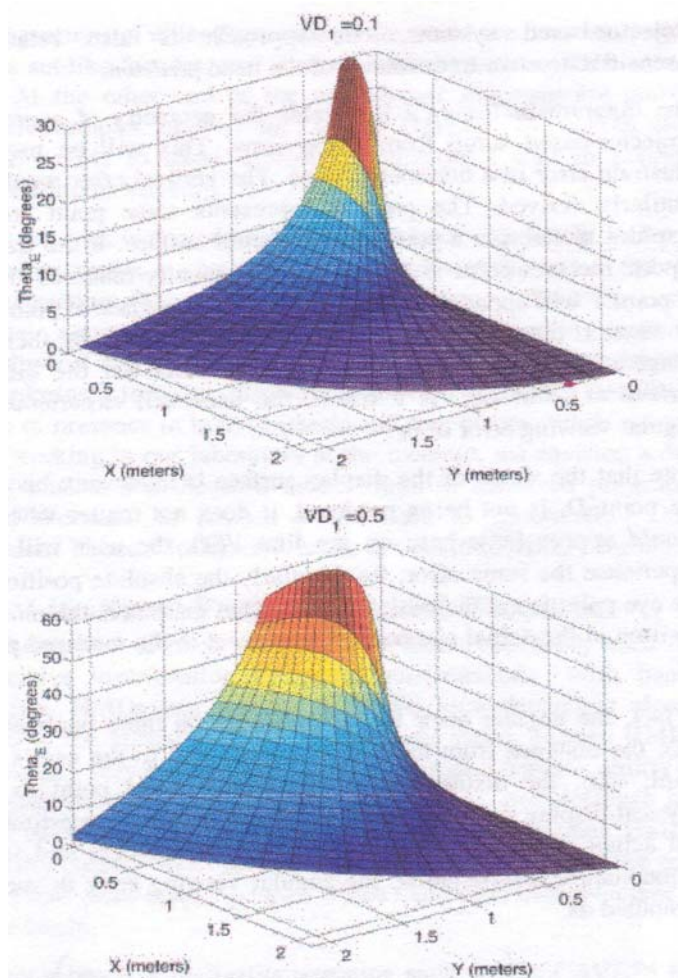
Z powyższego równania i rzutów błąd kątowy (a) wzrasta razem ze wzrostem kąta θ_V między oszacowanymi i właściwymi punktami oka (b) a zmniejsza się gdy odległość VE_2 od punktu wirtualnego obrazu do właściwego punktu oglądania wzrasta, i (c) wzrasta gdy odległość VD_1 między realną i wirtualną powierzchnią wzrasta.



Rys. 2

Pierwszy i drugi przypadek są na równi znane większości ludziom, którzy mają doświadczenie z konwencjonalną grafiką komputerową i perspektywiczną projekcją. Jednak ostatni przypadek jest rzadziej rozpoznawany lub omawiany w odniesieniu do konwencjonalnej grafiki komputerowej, przeważnie z powodu tego, że powierzchnia projekcyjna i wirtualne przedmioty są z reguły bardzo odmienne. Ważną rzeczą do zauważenia jest to że nawet gdy powierzchnia projekcyjna nie jest

dokładnie umieszczona i nie jest dokładnie w kształcie jak by tego wymagała prezentowana nań porcja obrazu, jeśli jest stosunkowo blisko, i jeśli widz nie podchodzi za blisko do powierzchni projekcyjnej to błąd kątowy może być stosunkowo mały.



Rys. 3

Jeżeli $VD_1 = 0$, to znaczy, jeżeli powierzchnia projekcyjna dokładnie odpowiada odpowiednim przedmiotom wirtualnym, wtedy błąd kątowy θ_e równa się 0 bez względu na VE_2 i θ_v . Innymi słowy, w tym specjalnym przypadku, nie potrzebne by nawet było specjalne urządzenie do śledzenia ruchu – po prostu ‘malowałyby’ się powierzchnię projekcyjną odpowiednimi kolorami.

Zmniejszona Wrażliwość na Uśpienie

Czułość na błąd pozycji oglądania w wyświetlaczach typu HMD i opierających się na projektorach ogólnego użycia, powiększa efekt nieuniknionych przestojów/uśpienia w przepływie grafiki. Problem polega na tym że scena kończy się będąc odtwarzaną ze złej perspektywy widzenia: z perspektywy miejsca w którym widz się znajdował niż gdzie właściwie jest. Takie błędy oglądania

spowodowane uśpieniem mogą przełamać poczucie obecności immersji, i mogą nawet wywołać chorobę lokomocyjną.

W przeciwieństwie, ponieważ ta metoda cieszy się ograniczoną wrażliwością na błąd śledzenia ruchu, w związku z tym też na uśpienia systemowe. Ludzie w typowy sposób przekładają głowy stosunkowo powoli, i nawet szybkie obracanie głowy powoduje tylko stosunkowo małe zmiany w pozycji oka.

Wynikiem ograniczonej wrażliwości na uśpienie może zostać łatwo zaobserwowane w tym prototypowym systemie, gdzie uśpienia systemowe są całkiem zauważalne dla obrazów świata zewnętrznego przez okno (gdzie VD_1 jest całkiem duże), ale są znacznie mniej widoczne na parapetach i sztukaterii (gdzie VD_1 jest stosunkowo małe).

Naturalne Auto-Stereoskopowe Widzenie

Kiedy geometria powierzchni projekcyjnej odpowiada dokładnie powierzchni wirtualnej środowiska, obrazy sceny projektowane na powierzchnie projekcyjne są obrazami stereoskopowymi trójwymiarowymi, bez wykorzystania żadnej przestrzennej stereo projekcji. Użytkownik ma możliwość oglądania właściwej powierzchni trójwymiarowej. Kiedy układ powierzchni projekcyjnych jest uproszczonym modelem danej sceny, tylko części tej sceny, które odpowiadają powierzchni projekcyjnej zatrzymują auto-stereoskopowy obraz, a pozostałe części muszą zostać powielone przy użyciu specjalnych trójwymiarowych stereo projekcji. W wstępnym eksperymencie, istnieją już duże ściany w obrazie, które odpowiadają geometrycznie uproszczonym powierzchniom projekcyjnym. Nawet bez użycia stereo projekcji, wirtualny obraz wciąż wygląda trójwymiarowo, ponieważ większość powierzchni jest bardzo przybliżona do właściwej, pożądanej geometrii.

Wysoka Przestrzenna i Geometryczna Wierność

W tradycyjnym odtwarzaniu grafiki, przestrzenne i graficzne właściwości przedmiotu są przedstawione w obrazie. Rozdzielczości tych obrazów mogą ograniczyć dokładność z którą właściwości geometryczne są prezentowane. Jednak użycie tej metody, te części modelowej sceny które odpowiadają geometrii powierzchni projekcyjnej mogą zostać przedstawione z bardzo dużą przestrzenną i geometryczną wiernością– bo przecież, poza kolorem, użytkownik ogląda prawdziwy fizyczny przedmiot.

Ostrość oczna

Ostrość oka dostarcza sygnałów o głębi w obrazie otaczającego nas świata. Przy typowej komputerowej grafice, HMD VR i CAVE, ostrość oka nie została włączona do wizualizacji, i wszystko w wirtualnych środowiskach jest ostre (w sensie obrazu). Z tą metodą i zastosowaniem lamp cieniujących, użytkownik teraz może wybiórczo skupiać swoje oko na dowolnej części obrazu, umożliwiając użytkownikowi bardzo naturalny sposób wizualizacji środowiska wirtualnego.

Powiększona Rzeczywistość

Ponieważ fizyczny układ powierzchni projekcyjnych jest przestrzennie podobny do geometrii sceny, jest stosunkowo łatwo dodać prawdziwe przedmioty do środowiska wirtualnego. Na przykład, jeśli układ powierzchni projekcyjnych przypomina wirtualne biurko, wtedy możemy postawić wazonik na tym biurku poprzez ustawienie prawdziwego wazonu na układzie powierzchni projekcyjnych, które przypominają biurko. Dodatkowe projektory mogą później zostać użyte do prawidłowego oświetlenia tych dodanych prawdziwych przedmiotów w celu symulowania efektu wirtualnego oświetlenia w wirtualnym środowisku.

Haptyka

Jeżeli układ powierzchni projekcyjnych jest wystarczająco blisko do powierzchni sceny, można go użyć do umożliwienia widzowi doświadczenia tego, co Fred Brooks nazwał, statyczną haptyką. W doświadczalnym ustawieniu Brook użył niskiej jakości, tanich i łatwych do skonfigurowania struktur aby skonstruować bardzo podobny fizyczny model kuchni. Użytkownik nosi na sobie HMD i porusza się wewnątrz tego fizycznego modelu, będąc w stanie dotknąć go i również oglądać szczegółowy obraz kuchni. Efekt realnego czucia powierzchni jest bardzo silnym czynnikiem w osiągnięciu poczucia obecności. Zgodnie z doświadczeniem wizualne poczucie użytkownika jest silniejsze od drobnych błędów w lokalizowaniu i powodowanych brakiem drobnych szczegółów modelu fizycznego.

Wady i Ograniczenia

Prawdopodobnie największym ograniczeniem tego zamysłu jest specyficzna natura zastosowania systemu. Metoda ta nie jest przeznaczona do ogólnej wizualizacji, gdzie użycie CAVE™ i podobnych wyświetlaczy oraz HMD miałyby więcej sensu.

Jak wskazuje tytuł, zamysł ten jest przede wszystkim pożyteczny przy tworzeniu dioram naturalnej wielkości miejsc realnych i wyimaginowanych.

Aspekt Fizyczny

Po podjęciu decyzji o użyciu metody odtwarzania danych scen, następnymi ważnymi problemami są aspekty fizyczne związane z potrzebą posiadania modelu powierzchni projekcyjnej, i przestrzeni potrzebnej do umieszczenia projektorów, kamer, i samego fizycznego modelu. Wymagania dotyczące fizycznego modelu jest bardziej uciążliwe teraz niż będzie w przyszłości. Jak to zostało opisane wcześniej istnieją pewne narzędzia, które są pomocne, i na pewno zostaną użyte pewne mocniejsze i skomputeryzowane narzędzia. Odnośnie wymagań przestrzeni fizycznej, obecnie pracuje się w laboratorium badawczym, więc nie ma możliwości przeprowadzać symulacji miejsc większych niż laboratorium. Można użyć pomieszczenia o kształcie magazynu, w którym będzie można odtworzyć większość scen studyjnych. Należy też zauważyć, że nie zawsze trzeba przydzielać przestrzeń fizyczną dla wszystkiego co znajduje się na obrazie, należy zawsze jednak przydzielić/wyznaczyć taki obszar dla tych części sceny, w których będzie się poruszał użytkownik. Na przykład na Rysunku 1, nie ma przydzielonej przestrzeni fizycznej dla scen ze świata zewnętrznego widzianych przez okno.

Zawartość Sceny

Wszystkie zalety są do użycia jeśli fizyczne powierzchnie projekcyjne utrzymane są tak by były podobne do powierzchni na nich projektowanych. Oznacza to że obrazy dynamiczne z dużymi przedmiotami, które poruszają się sporo lub odkształcają znacznie, nie są dobrymi kandydatami do tej metody. Jednak, małe przedmioty które poruszają się blisko dużych powierzchni (na przykład owad chodzący po ścianie) nie powinny utrudniać pracy tego systemu. Ponadto, co istotne dla osiągnięcia celu, większość muzeów wystawia statyczne eksponaty, których nie wolno dotykać zwiedzającym.

Czynniki projekcyjne

Generalnie, ta metoda nie zawsze pozwala na tylną projekcję. Na przykład, na rysunku 1, obie strony ściany środkowej muszą być oświetlone i nie ma tam właściwie miejsca między nimi aby umieścić projektory. Mimo że metoda tylnej projekcji ma większe wymagania co do przestrzeni, ma też tą zaletę że użytkownicy nie zasłaniają światła od projektorów. Zamiast tego będzie użyte 'morze

projektorów' ustawionych na suficie przestrzeni rekonstrukcyjnej. Kwestia rozmiaru projektorów i kosztów nie będą duże w przyszłości. Rozmiar i koszt projektorów wciąż gwałtownie maleją. Dziś można zakupić projektor, który waży mniej niż 5 funtów za mniej niż \$5,000. Nowe odblaskowe mikro wyświetlacze skonstruowane w oparciu o standardową technologię CMOS dają nadzieję na mniejsze, 'inteligentniejsze', i tańsze cyfrowe projektory.

Zastosowanie

Zastosowano bardzo wczesny prototyp do zademonstrowania zamysłu nowej metody, ale jeszcze nie istnieje kompletny system.

Przegląd

System utworzony w oparciu o tę nową metodę będzie składał się z następujących głównych składników:

- **Tworzenie modelu sceny.** Model sceny może być albo wyimaginowany bądź realny.
- **Modelowanie powierzchni projekcyjnej.** Model powierzchni projekcyjnej musi zostać obliczony tak by pasował do powierzchni wyświetlanej na nim sceny.
- **Fizyczna konstrukcja powierzchni projekcyjnej.** Fizyczne ustawienie musi być skonstruowane zgodnie z obliczonym modelem powierzchni sceny.
- **Umieszczenie i skalowanie projektora.** Projektory są umieszczane w odpowiednich miejscach względem modelu fizycznego, i są potem skalowane aż do momentu odnalezienia geometrycznego dopasowania do obrazu fizycznego modelu.
- **Śledzenie ruchu.** Głowa każdego użytkownika musi być śledzona w celu zgromadzenia zgodnych z perspektywą obrazów.
- **Odtwarzanie.** Ta funkcja odpowiedzialna jest za generowanie poprawnego obrazu na powierzchniach projekcyjnych (według perspektywy użytkownika). Zajmuje się ona również obrazami nachodzącymi na siebie a pochodzącymi od różnych projektorów, i tworzy obrazy które łączą się bez widocznych szwów na nakładających się obszarach.

Każdy z tych składników omawiany jest szerzej w poniższych podrozdziałach, w których zaprezentowane są główne założenia, wyzwania, rozwiązania i inne możliwe sposoby zmierzenia się z tymi wyzwaniami.

Modelowanie Sceny

Główne wyzwanie w tym obszarze powstaje gdy chcemy odtworzyć prawdziwe miejsca. Aby utworzyć model takiego miejsca, można użyć skanera zasięgowego aby pobrać próbki geometrii danej sceny, i kamery do pobrania próbki właściwości fotometrycznych danej sceny. Niektóre z napotkanych kwestii to zagospodarowanie tą ogromną ilością pobranych informacji, zapisywanie zdjęć z informacjami o zasięgu, i zapis zasięgu danych z różnych punktów widzenia. Jeśli zależne od punktu oglądania odbicie powierzchniowe ma być symulowane, wtedy o wiele więcej zdjęć prawdziwej sceny musi być zrobionych, a metoda oświetlonego pola może być użyta w celu odtworzenia obrazów w zależności od punktu patrzenia na obszar projekcyjny.

Obecnie najbardziej interaktywną metodą odtwarzania dla VR jest wciąż tradycyjna metoda z-buforowa. Stąd, odtwarzanie obrazu na powierzchni z danych zasięgu do generowania modelu wielokątowego jest konieczne. Uproszczenie siatki jest stosowane do zmniejszenia ilości wielokątów poprzez eliminację zbędnych informacji o geometrii. Podczas odtwarzania, zdjęcia scen realnych mogą być fakturowane na uproszczonej siatce.

Modelowanie i Konstrukcja Powierzchni Projekcyjnej

Potrzeba modelowania i budowania nie prostych fizycznych powierzchni projekcyjnych jest wyzwaniem nie spotykanym w innych tego typu metodach. Jednym z głównych celów jest to że fizyczne powierzchnie projekcyjne powinny być łatwe do ustawienia. To sprawiło że rozważa się zastosowanie podobnych do klocków Lego™ w tej metodzie.

Z wybranym zbiorem kształtów, trzeba znaleźć metodę do automatycznego obliczania układu tych bloków, aby najlepiej trafić w kształt powierzchni projektowanego obrazu. To jest bardzo trudny problem, i nie istnieje opis tego typu problemów w literaturze fachowej.

W tym prototypie używa się białych bloków styropianowych o wymiarach 0.25x1.2x0.3 metrów sześciennych, do ułożenia fizycznych powierzchni projekcyjnych. Aby utworzyć model powierzchni projekcyjnej w komputerze,

zastosowano program, który pozwoli na ręczne projektowanie układu bloków w celu dopasowania powierzchni do powierzchni modelu projektowanego obrazu. Model ten jest potem używany jako projekt do generalnej konstrukcji ustawienia fizycznego.

W przyszłości będą użyte projektory i modele analityczne powierzchni projekcyjnych do pomagania użytkownikom w nawigacji w obszarze fizycznego modelu. Biorąc pod uwagę pustą przestrzeń z zeskalowanymi projektorami, zautomatyzowany program może odtworzyć zarys bloku na podłodze, czekając na pracownika aby potwierdził położenie bloku. Podobne kroki są stosowane dla pozostałych bloków, łącznie z tymi ułożonymi na stosie.

Umieszczenie i Skalowanie Projektorów

Umieszczenie projektora

Po utworzeniu powierzchni projekcyjnej, trzeba zdecydować gdzie umieścić projektory. Możemy też chcieć znać minimalną liczbę projektorów potrzebnych do pokrycia całego obszaru projekcyjnego. Problem ten jest podobny do problemu umieszczenia kamery w modelowaniu w oparciu o obraz, ale jest utrudniony przez fakt że, poza właściwościami, także należy obliczyć orientacje projektorów.

Metoda do obliczenia pozycji i orientacji układu projektorów może wymagać wzięcia pod uwagę następujących elementów:

- fizyczne i środowiskowe ograniczenia przy zamontowywaniu projektora,
- maksymalizowanie powierzchni obejmującej projekcją dla każdego projektora,
- minimalizowanie powierzchni nakładania się obrazów,
- utrzymywanie pewnej rozdzielczości obrazu na powierzchni projekcyjnej,
- minimalizowanie możliwego zasłaniania obrazu przez użytkownika.
- minimalizowanie wewnętrznego odbicia pomiędzy powierzchniami projekcyjnymi
- minimalizowanie liczby projektorów.

Aby zminimalizować wewnętrzne odbicie światła, najlepiej unikać projekcji światła na duże powierzchnie pod bardzo ukośnym kątem.

Jeszcze nie dokonano zautomatyzowania procesu umieszczania projektora. To jest wciąż otwarty do badań rozdział. Obecnie umieszczone są projektory (używając ruchomych kratek) w miejscach które spełniają powyższe wymagania, potem z reguły wypróbować się kilka opcji przed ustawieniem ostatecznej opcji.

Skalowanie projektora

Kiedy projektory są umieszczone i nakierowane odpowiednio względem powierzchni projekcyjnej, dostraja się ich odległość skupienia aby uzyskać najlepszą ostrość na powierzchni projekcyjnej. Następnym krokiem jest skalowanie projektorów aby odnaleźć ich geometryczny stosunek względem modelu projektowanej sceny i modelu powierzchni projekcyjnej.

Aby wyskalować projektor, należy znaleźć układy par odpowiedników. Każda taka para składa się ze współrzędnych trójwymiarowego punktu w przestrzeni, a odpowiadające im punkty (ich współrzędne) na płaszczyźnie obrazu projektora. Przy wystarczającej liczbie takich par odpowiedników (przynajmniej 6, z których cztery trójwymiarowe punkty nie mogą mieć wspólnej płaszczyzny). Możemy rozwiązać parametry projekcji przy użyciu metod linearnych najmniejszego kwadratu, lub przy użyciu nieliniarnych metod optymalizacji. Aby odnaleźć parę odpowiedników, rzucamy znany dwuwymiarowy punkt z projektora, który pojawia się jako promień. Potem przemieszcza się małą powierzchnię projekcyjną aby przeciąć ten promień i użyć urządzenia śledzącego do zmierzenia trójwymiarowej pozycji przecięcia. Te trójwymiarowe punkty są potem przekształcane z systemu współrzędnych urządzenia śledzącego ruch na system współrzędnych modelu sceny, przy użyciu znanej transformacji. Zestaw par odpowiedników jest używany do rozwiązania parametrów projekcji projektora względem systemu współrzędnych sceny/obrazu.

Powyższa ręczna metoda odnajdywania par odpowiedników może być zmusna i podatna na błędy. W przyszłości, automatyczne skalowanie projektorów przy użyciu takich kamer może zostać zastosowane, możliwie jako część zautomatyzowanego algorytmu umieszczenia.

Śledzenie ruchu

Ponieważ użytkownikowi wolno swobodnie chodzić w obrębie potencjalnie dużej powierzchni, potrzebne jest urządzenie śledzące o obszernym zasięgu do śledzenia pozycji głowy. Jeśli kończyny użytkownika mogą również 'pracować ze

środowiskiem' wtedy też ich ruch musi być śledzony. Wymagana przestrzenna rozdzielczość i dokładność zależą w dużej mierze od największego VD_1 danego modelu (rysunek 2). Im większe odległości między pożądanymi modelami i fizycznymi powierzchniami projekcji, tym ważniejsza jest wierność przestrzenna każdego śledzonego obrazu.

Użyte są dwa 3rdTech (trzeciej generacji) HiBall urządzenia śledzące w tym eksperymencie. Te urządzenia mogą pokryć swym zasięgiem szeroki obszar i posiadać wystarczająco wysoką precyzję i niską uśpioność. Jedno urządzenie śledzące jest używane do śledzenia głowy użytkownika a drugie do śledzenia pozycji i orientacji wirtualnej puszki do sprayu. Każda pozycja zwrócona przez te urządzenia śledzące jest przekształcana w system współrzędnych modelu sceny.

Odtwarzanie

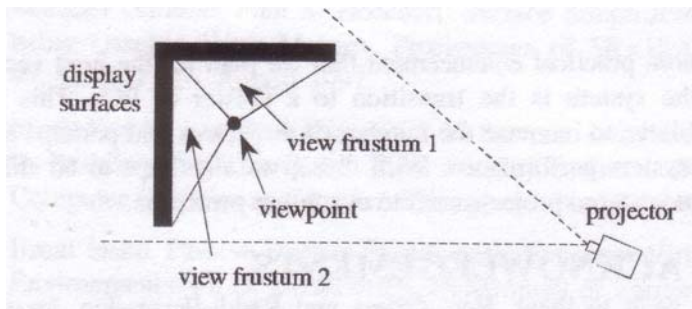
Generowanie Poprawnych pod względem Perspektywy Obrazów

Dla każdego projektora, używa się dwupasmowej metody odtwarzania aby zgromadzić poprawny obraz dla użytkownika. Podczas pierwszego pasma, model sceny jest normalnie odtwarzany z pozycji oka użytkownika. Powstający obraz jest wczytywany z powrotem do ramy buforowej, aby był użyty w następnym paśmie. W drugim paśmie model powierzchni projekcyjnej jest odtwarzany z pozycji projektora, z powierzchniami projekcyjnymi obrzucanymi fakturą obrazu utworzonego podczas pierwszego pasma. Mapa fakturowa jest mapowana na powierzchnie projekcyjne przy użyciu mapowania projekcyjnej faktury, wyświetlanej z pozycji oka użytkownika. Obraz odczytywany z powrotem z ramy buforowej może okazać się 'wąskim gardłem' dla odtwarzania, gdyż transfer obrazu o wysokiej rozdzielczości może być bardzo wymagający dla pamięciowej szerokości taśmy dla wielu odtwarzających systemów.

Porcjowanie Powierzchni Projekcyjnej

Aby zapewnić panoramiczny obraz użytkownikowi, trzeba wyświetlać obrazy na wszystkich powierzchniach projekcyjnych otaczających użytkownika. Sytuacja z rysunku 4: dwie ściany są oświetlane projektorem a użytkownik jest pomiędzy tymi dwiema ścianami. Podczas pierwszego pasma odtwarzania, trzeba odtworzyć obraz danej sceny z punktu widzenia widza. Jednak, nie można ustawić stożka projekcji aby generować całkowity obraz z obu ścian. Aby zebrać obraz całkowity obu ścian,

można ustawić dwa różne stożki, każdy z nich będzie patrzył na ścianę i odtwarzał dwa obrazy. Podczas drugiego pasma odtwarzania, powierzchnie projekcyjne, które są zbliżone budową do ścian są mapowane fakturą obrazem korespondującym z tym utworzonym w czasie pierwszego pasma. To wymaga aby powierzchnie projekcyjne były podzielone na grupy tak aby każda grupa mogła być mapowana fakturą innego obrazu. W tym prototypie, podzielono powierzchnie projekcyjne ręcznie.



Rys. 4

Bezszwowe Łączenie Obrazów Projektowanych

Regiony obszaru projekcyjnego, które są oświetlane licznymi projektorami wydają się jaśniejsze, sprawiając że nakładające się obszary wyraźnie odróżniają się od reszty dla oka widza. Aby sprawić by te połączenia nakładających się na siebie obrazów były niewidoczne, można użyć techniki alfa do łączenia aby zmniejszyć intensywność projekcji każdego projektora w obrębie nakładania się na siebie projektowanych obrazów. Jednak takie obszary projekcyjne mogą mieć części wklęsłe, które mogą spowodować że nakładające się obszary, z punktu widzenia projektora, mogą być widoczne jako nie ciągłe. Tradycyjne metody nie mogą poradzić sobie z tym problemem zbyt skutecznie, jednak jest lepsza metoda.

Kolejnym problemem łączenia obrazu projektowanego z różnych projektorów jest brak równości kolorów pomiędzy sąsiadującymi projektorami. Majunder używa tabeli wyszukiwania kolorów aby skorygować te różnice między projektorami.

Nie zastosowano żadnej korekty dla nakładających się obszarów i różnic koloru w tym prototypie systemu.

Inne Kwestie i Wyzwania

Wewnętrzne Odbicie

Ze wstępnych eksperymentów zaobserwowano, że wewnętrzne odbicie światła z projektorów może być poważnym problemem. Odbicia te obniżają jakość projektowanych obrazów poprzez zmniejszanie ich nasycenia kontrastu i kolorów. Jeszcze nie istnieje dobre rozwiązanie tego problemu, ale poniższe wskazówki będą pomocne:

- używać mniej odbłaskowych (bardziej rozpraszający) materiałów do powierzchni projekcyjnych;
- unikać nadmiernie ukośnych kątów projekcji na duże powierzchnie.

W przyszłości użyte będą metody zautomatyzowane, które będą brały pod uwagę tą drugą wskazówkę.

Cienie

Głównym problemem użycia metody przedniej projekcji jest to, że projektory mogą być blokowane przez użytkownika. Może to wywoływać przeszkadzające cienie. Można umieścić projektory wyżej i bliżej powierzchni projekcyjnych, ale to nie jest dobre rozwiązanie, gdyż projektory będą wtedy ustawione na bardzo ostrych kątach w stosunku do powierzchni projekcyjnej. Możliwym rozwiązaniem jest posiadanie co najmniej jednego ‘wspierającego/pomocniczego’ projektora dla każdego projektora właściwego, wtedy projektor i jego wsparcie będą oświetlały tą samą powierzchnię, ale z dwóch różnych pozycji ale nie w tym samym czasie. Kamery wideo i urządzenia śledzące ruch są potrzebne do określenia które projektory zostały zasłonięte przez użytkownika i wtedy te wspierające projektory zostaną uruchomione w celu wypełnienia obrazem zacienionych miejsc na powierzchni projekcyjnej.

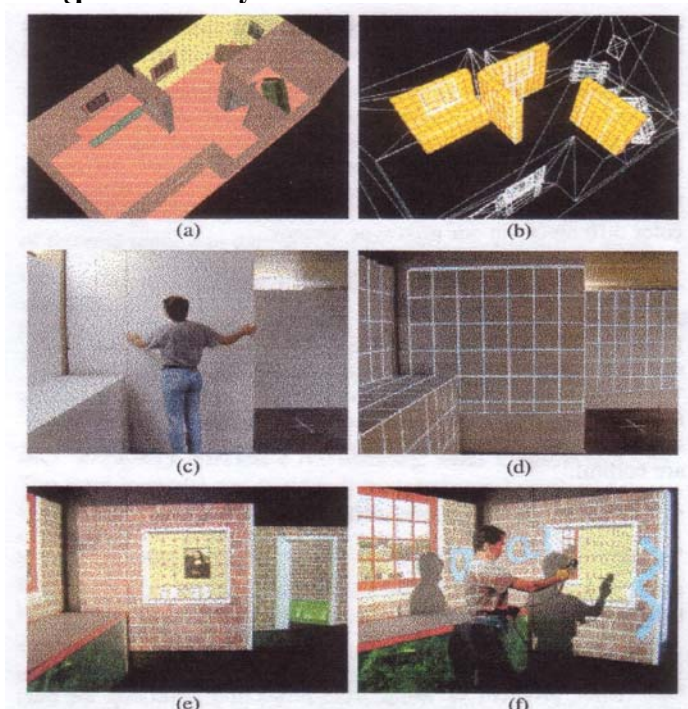
Złożona liczba widzów

Jeśli powierzchnie projekcyjne pasują kształtem geometrii projektowanego na nich obrazu, wtedy większa liczba oglądających powinna być w stanie eksplorować wirtualne środowisko w tym samym czasie, i każdy będzie obserwował zawsze obraz poprawny, bez potrzeby śledzenia ruchu użytkowników. Jednak to ekstremalne wymaganie realistycznego modelu jest nierealistyczne i nierozsądne dla większości scen jakie sobie można wyobrazić. Oznacza to że wciąż problemem pozostaje kwestia potrzeby podtrzymywania projekcji wielu obrazów symultanicznie w zależności od ruchów oczu i pozycji użytkowników. Z tym samym problemem zmagają się również systemy CAVE™

i podobne. Niestety nie ma jeszcze żadnego praktycznego rozwiązania gotowego do użycia i podtrzymywania więcej niż kilku (dwóch może trzech) symultanicznych widoków zależnych o ruchu głowy użytkownika.

Problem ten obejmuje trudności w śledzeniu ruchu, odtwarzaniu, i wyświetlaniu. Największym problemem powiązaniem z większą ilością widzów jest samo wyświetlanie. Podobnie jak konwencjonalne systemy używające projektorów generalnego użytku, dotyczą zarówno praktyczne ograniczenia w obrębie technologii projektorów i zachowanie fizyczne światła. Tradycyjne metody projekcji obrazów symultanicznie na te same powierzchnie obejmuje podział czasu i podział fazy. Podział czasu jest ograniczony przez szybkość, z którą można zmieniać projektowane przez rzutniki obrazy, i przez zmniejszającą się ilość światła odpowiadającą krótszym przerwom projekcji jak dodaje się więcej obrazów. Podział fazy (spolaryzowanego obrazu) jest ograniczona przez podstawowe sposoby jakimi można dokonać polaryzacji światła, sposobami jakimi projektory wewnętrznie przetwarzają światło (czasami niezależnie je polaryzując), i skutkami polaryzacji na jakość ostatecznego projektowanego obrazu. Zdolność dostarczania niezależnych kanałów przepływu obrazu do oczu każdego widza jest, w co jednak można wątpić, zaleta jaką osadzone na głowie wyświetlacze mają nad wyświetlaczami wykorzystującymi projektory.

Wstępne rezultaty



Rys. 5 Prototyp (a) Syntetyczny model sceny, (b) Model powierzchni wyświetlającej, (c) Wstępna fizyczna aranżacja modelu powierzchni wyświetlającej, (d) Weryfikacja fizycznego modelu przez użycie ramy sitowej wyświetlanej przez projektory, (e) Ostateczny rezultat z zachowującym dobre proporcje obrazie wygenerowanym w czasie rzeczywistym, (f) Użytkownik wirtualnie maluje ściany spray'em na wirtualnych ścianach.

Zastosowano prototyp systemu do zademonstrowania pomysłów. W tym prototypie dokonuje się symulacji tylko części wymyślnego modelu pokoju (rysunek 5), z powodu ograniczonej przestrzeni w laboratorium badawczym – które ma około 8.5x4.5x2.5 metrów sześciennych. Stworzono powierzchnie projekcyjne używając konstrukcyjnych bloków styropianowych (rys. 5c). Zostało użytych sześć projektorów, każdy z których miał rozdzielczość obrazu 1024x768 pikseli. Użyto dwóch HiBall urządzeń śledzących ruch - jeden do śledzenia głowy użytkownika a drugi do śledzenia ruchu wirtualnej puszkii sprayu. Użyto 32-procesorowego, 8-kanalowego SGI Onyx2 Infinite Reality2 ('Reality Monster') do pobierania wszystkich obrazów dla projektorów w czasie rzeczywistym. Z powodu ograniczonej ilości projektorów i stosunkowo powolnego dwu-pasmowego odtwarzania, nie załączono funkcji stereo-projeckcji w tym systemie.

Uzyskane rezultaty z pracy tego prototypu były interesujące, zwłaszcza przy patrzeniu przez wirtualne przejście na ścianie łączącej pomieszczenie realne z wirtualnym, idąc z jednej strony ściany do drugiej. Także ciekawym jest zobaczyć jak zewnętrzny obraz porusza się przez okno, mijając okna, w czasie gdy same ramy i otaczające okno ściany pozostają w miejscu. Na rysunku 5f, pokazany jest użytkownik wykonujący wirtualne malowanie spray'em na ścianach i framudze okna. Celem jest zademonstrowanie użyteczności bycia zdolnym do fizycznego poruszania wirtualnych przedmiotów – co daje użytkownikowi naturalną swobodę w pracowaniu z wirtualnym środowiskiem.

Naukowcy są w trakcie budowania modelu, w skład którego weszły by ściany na prawo od kamery z rysunku 5 (e) aby zapewnić większy panoramiczny (immersyjny) efekt.

Wnioski

Naukowcy są usatysfakcjonowani rezultatami pierwszego prototypu. Zdolność do chodzenia w utworzonym środowisku jest unikalna. Krokiem następnym będzie wypróbowanie modelu obrazu pobranego w Monticello. To będzie o wiele większe wyzwanie, ponieważ meble w pomieszczeniach zawierają o wiele więcej zakrzywionych powierzchni. I stąd nie są dobre do projektowania na powierzchni zbudowanej ze styropianowych bloków. Przejście do bardziej złożonego modelu będzie wymagało badań nad bardziej odpowiednimi kształtami i rozmiaracmi bloków, jak i zautomatyzowanego dopasowywania modelu fizycznego (realnego) do modelu wirtualnego.

Naukowcy naprowadzają projektory, można dzięki temu uzyskać bardzo precyzyjną kontrolę nad oświetleniem w pokoju. Ponadto śledzony jest ruch oczu użytkownika, można uwydatnić to doświadczenie poprzez usuwanie spowodowanych przez użytkownika cieni przy użyciu projektorów wspierających, i oświetlając ciało użytkowników przy jednoczesnym nie oślepieniu ich.

Więszym udogodnieniem jakie planują zastosować naukowcy w następnej wersji systemu jest przejście do zespołu komputerów zespolonych ze sobą. To pozwoli na zwiększenie liczby projektorów i być może również skali działania systemu. Pozwoli to na zastosowanie stereo projekcji w przyszłym prototypie.

4. Lampy Cieniowe

Ożywianie Rzeczywistych Obiektów Przy Użyciu Oświetlenia Opartego na Obrazach

Przedstawiony jest tutaj nowy wymiar trójwymiarowej grafiki komputerowej z wykorzystaniem projektorów do ożywiania obiektów fizycznych w świecie rzeczywistym. Chodzi o to, aby zastąpić obiekt fizyczny, wraz z jego kolorem, teksturą i właściwościami materiału, neutralnym obiektem i wyświetlanym obrazem, odtwarzając oryginalny (lub zmieniony) wygląd bezpośrednio na tym obiekcie fizycznym. Ze względu na to, że metoda ta ma skutecznie “podnosić” właściwości wizualne obiektu do projektora, nazywa się te projektory lampami cieniowymi. W tym rozdziale omówione jest główne zagadnienie pełnego i ciągłego oświetlania nieszablonowych obiektów fizycznych przy użyciu systemów wieloprojektorowych i zaprezentowany zestaw nowych technik, dzięki którym proces oświetlania jest praktyczny i przystępny. Demonstracja zastosowania tych

technik w wielu różnych sytuacjach i przedstawienie pierwszych rezultatów odtwarzania przestrzeni wirtualnych o rzeczywistych wymiarach.

Wstęp

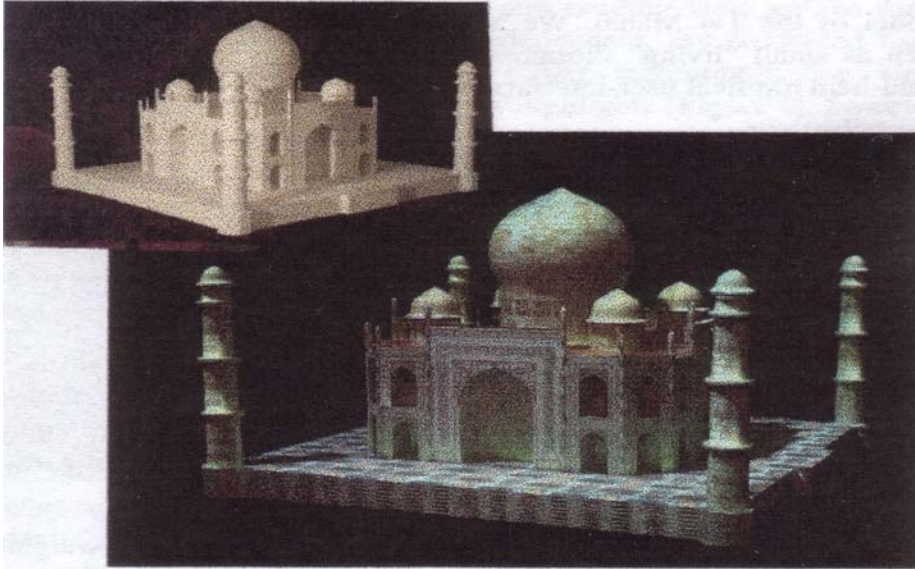
Grafika na świecie

Tradycyjnie techniki grafiki komputerowej próbowały zawsze “uchwycić” świat realny w komputerze, a potem odtworzyć go wizualnie. Późniejsze działania doprowadziły w końcowym efekcie do sytuacji odwrotnej – do “wkładania” grafiki komputerowej do świata realnego. Działania takie miały miejsce przede wszystkim w tworzeniu efektów specjalnych do filmów i w *powiększanej rzeczywistości* w rzeczywistym czasie. W ostatnim czasie pojawił się nowy trend, aby wykorzystywać projektory świetlne do wyświetlania obrazów bezpośrednio w rzeczywistym otoczeniu. Przykładami tego mogą być *Luminous Room* (Świetlny Pokój) i *Office of the Future* (Biuro Przyszłości). To, do czego dążą naukowcy z Uniwersytetu Północnej Karoliny, to bardziej kompletne rozwinięcie tych pomysłów – użycie trójwymiarowej grafiki komputerowej i animacji bezpośrednio w rzeczywistym świecie, który nas otacza.

Stymulacja i przekazywanie pomysłów

Mimo dużych postępów w grafice komputerowej, architekci i osoby zajmujące się zagospodarowaniem przestrzennym miasta nadal uciekają się do budowania modeli fizycznych jeśli zależy im na zadowoleniu klientów czy wyborców. Architekci uważają, co potwierdzają również książki o tej tematyce, że mimo że prawdą jest to, że projektanci nie mogą już obejść się bez narzędzi CAD, to “komputer nie jest w stanie zastąpić styczności z prawdziwym materiałem, fizycznych kształtów i narastania relacji przestrzennych.” Nawet w dzisiejszej erze animacji komputerowej, animatorzy najpierw rzeźbią fizyczny model obiektu, a dopiero potem tworzą modele komputerowe. Tak było w przypadku Geriego w “Geri's Game” (filmie stworzonym przez Pixar Animation Studios). Jednym z powodów, dla którego takie sentymenty i praktyki nadal są obecne jest to, że człowiek ma jakby “intuicyjne połączenie” z modelem fizycznym. Nie trzeba przesuwac żadnych ikonek ani suwaków, nie trzeba patrzeć przez żadne wyświetlacze, ani ich na siebie zakładać. Można za to chodzić wokół obiektów, przybliżać się i oddalać, aby uzyskać właściwą wielkość obrazu, można się przyglądać i skupiać wzrok na interesujących komponentach, a wszystko to w bardzo wysokiej jakości wizualnej, przestrzennej i

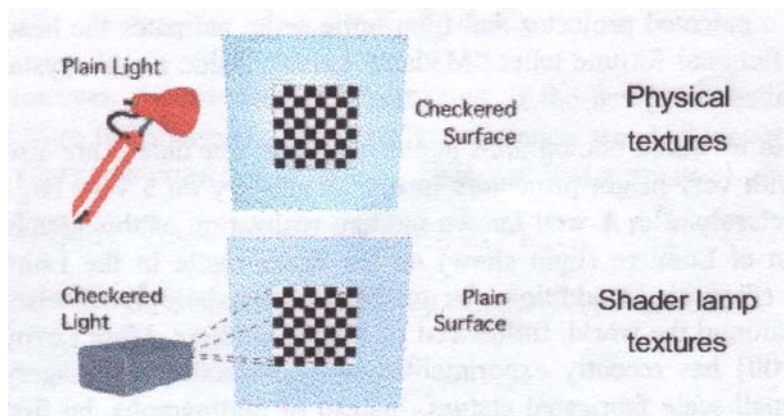
czasowej. Ambitny celem jest uzyskanie zalet naturalnego interfejsu fizycznego przy użyciu lamp cieniowych. Główne zalety, jakie trzeba będzie uzyskać, to autostereoskopowy sposób widzenia obiektów połączony z bogactwem grafiki komputerowej.



Rys1. Na górze fizyczny model Tadź Mahal, na dole ten sam model oświetlony lampami cieniowymi.

Oświetlanie oparte na obrazach

Kiedy oświetlamy obiekt rzeczywisty białym światłem, jego powierzchnia odbija poszczególne długości fal światła. Dzięki temu, że nasz sposób postrzegania właściwości powierzchni zależy tylko od widma światła, jakie dochodzi do naszych oczu, jesteśmy w stanie przesunąć i przemieszczać obiekty na ścieżce optycznej, pod warunkiem, że widmo światła, jakie dotrze do naszych oczu, pozostanie wystarczająco podobne. Do źródła światła można dołączać wiele atrybutów fizycznych i wyświetlać je na neutralnym obiekcie, aby uzyskać obiekt, który będzie postrzegany jako taki sam jak fizyczny. Można w ten sposób wyświetlić nawet nierealistyczny wygląd przedmiotu. Koncepcja ta zilustrowana jest na rys. 2. Można korzystać z cyfrowych projektorów świetlnych i grafiki komputerowej, aby stworzyć lampy cieniowe, dzięki którym można skutecznie odtwarzać lub łączyć różne właściwości powierzchni, tak statycznie, dynamicznie, jak i interaktywnie. Mimo że teoria wskazuje na to, że wyniki takie osiąga się tylko z niewielką liczbą powierzchni i właściwości, to z doświadczenia wynika, że obrazy są dosyć realistyczne i efektowne w szerokiej gamie zastosowań.



Rys. 2 Koncepcja lamp cieniowych.

Korzystanie z fizycznych modeli jako podstawy do wyświetlania obrazu jest rzeczą dość nietypową w grafice komputerowej, jednak inaczej rzecz ma się w przypadku architektów, artystów i animatorów komputerowych. Poza tym pojawiają się cały czas różne metody zautomatyzowanej produkcji trójwymiarowej, np. wytwarzanie obiektów świetlnych, stereolitografia czy łączone osadzanie. Można się również spodziewać, że w niedalekiej przyszłości obrazy trójwymiarowe będziemy drukować i przysyłać faxem.

W poprzednich rozdziałach przedstawione zostały wstępne pojęcia i wyniki pracy w warunkach warsztatowych. Po dodatkowych badaniach i testach nad tym pomysłem przedstawiono tę metodę bardziej dokładnie i omówiono metody praktyczne. Przedstawiono w tej pracy wyniki użycia wielokrotnych lamp cieniowych do ożywiania obiektów fizycznych – od prostych do bardziej złożonych, poczynając od gładkiego wazonu, a kończąc na stosunkowo bardzo skomplikowanym modelu Tadź Mahal. Przedstawiono również kilka zastosowań metody, takich jak małe “żywe” dioramy (ułożenia obrazów, które tworzą efekt trójwymiarowy), modele w ludzkiej skali i trzymane w ręce obiekty fizyczne z interfejsem użytkownika.

Dokonania

Prezentacja lamp cieniowych jako nowy sposób wizualizacji trójwymiarowej grafiki komputerowej. W metodzie tej traktuje się oświetlenie jako trójwymiarową projekcję perspektywiczną z lampy i jako takie może ono zostać wygenerowane przy użyciu tradycyjnej trójwymiarowej grafiki komputerowej. Przedstawia się techniki, które mogą zastąpić nie tylko tekstury, czyli komponent rozproszony, ale można nimi także zastąpić praktycznie każdy obraz BRDF (funkcja

dwukierunkowego rozpraszania odbicia).

Przedstawione są nowe algorytmy, dzięki którym proces oświetlania staje się praktyczny. Najpierw znajduje się proste równanie dostosowania świecenia, które nadzoruje proces wyświetlania, a potem przedstawiane są metody poprawy odpowiadającej intensywności świecenia.

Nowy algorytm określania ciężkości pikseli i wyliczania intensywności wtapiania podczas przechodzenia w różne rejony wpływu projektora przy braku ciągłości głębi.

Poprzednie prace

Teatr i rozrywka. “Displacements” Naimarka, śpiewające głowy w “Haunted Mansion” Disneya i pokazy Son et Lumiere na zabytkach architektury to przykłady idei najbliższych pojęciu jawnego dzielenia, a następnie łączenia geometrii fizycznej i właściwości wizualnych rzeczywistej sceny.

Naimark korzystał z obrotowej kamery filmowej do sfilmowania salonu, w którym było pełno mebli i ludzi. Następnie pokój i meble pomalowano na kolor biały (neutralny), a przechwycony wcześniej obraz został wyświetlony na ścianach przy użyciu obrotowego projektora, który był ustawiony dokładnie tak jak wcześniej kamera. Takie krytyczne ustawienie urządzeń nagrywających i odtwarzających jest popularne w większości obecnych sposobów prezentacji, które wykorzystują nagrane wcześniej obrazy lub sekwencje obrazów. Niewielkim, ale za to imponującym przykładem tej idei jest wyświetlenie wcześniej nagranych filmów do ożywienia czterech neutralnych popiersi śpiewających mężczyzn w “Haunted Mansion” Walta Disneya. Innym przykładem jest użycie opatentowanego zestawu projektor – światłowodów do zaminowania głowy fikcyjnej wróżki “Madame Leota” wewnątrz prawdziwej kryształowej kuli.

Na bardzo dużą skalę w zastosowaniach architektonicznych używa się również slajdów zmodyfikowanych zdjęć powiększonych ze szczegółami do wyświetlania obrazów za pomocą bardzo jasnych projektorów. Bardzo znanym nowoczesnym wykorzystaniem tej idei jest Son et Lumiere (pokaz świetlny) na zamku Blois w Dolinie Loire we Francji. Marc Levoy, zainspirowany przez Son et Lumiere, przeprowadził ostatnio eksperyment wyświetlania obrazu na pomniejszonych modelach pomników. Zamiast użyć zdjęć, najpierw wyświetla się obraz zapisanego modelu trójwymiarowego, podobnego do powyższej techniki, a

następnie ręcznie ustawia projektor, aby był geometrycznie wyjustowany w stosunku do wyświetlanego obrazu. Hypermask99 jest wyjątkiem jeśli chodzi o justowanie automatyczne, składa się z wyświetlania animowanej twarzy na poruszającą się maskę do opowiadania historii. Wszystkie te systemy tworzą wspaniałe wizualizacje. Jednakże niewygodny proces dopasowywania może zająć do kilku godzin, nawet dla pojedynczego projektora. Nowa technika omija ten problem poprzez stworzenie trójwymiarowego porozumienia geometrycznego przy użyciu znanych komputerowych technik wizyjnych opisanych w następnych podpunktach, a potem przechodzi poza zwykłe wyświetlanie obrazu do odtwarzania wartości współczynnika odbicia.

Dotykowe interfejsy świetlne

Projekt Pokoju Świetlnego traktuje ustawioną parę kamera-projektor jako lampę wejście/wyjście do zwracania i wyświetlania obrazu na płaskie powierzchnie w rzeczywistym otoczeniu fizycznym pomieszczenia czy wydzielonego miejsca pracy. Obecna praca jest oddzielną od tego projektu, ale uzupełnia go. Główna różnica polega na tym, że pierwsza praca skupia się głównie na interakcji z informacjami poprzez interfejsy oświetlane i dotykowe. Przykładami takiego nastawienia mogą być takie zastosowania jak “Światło Oświetlające” i “URP” (plany zagospodarowania przestrzeni miejskich). Ten ostatni przykład jest najbardziej podobny do aktualnych prac, zwłaszcza pod względem interaktywnej symulacji cieni budynków zmieniających się pod wpływem światła słonecznego. Metoda ta najpierw rozpoznaje dwuwymiarowe obiekty fizyczne (“fizyczne ikony” budynków) leżące na płaszczyźnie, następnie śledzi ich dwuwymiarową pozycję i orientację na płaszczyźnie, a na koniec oświetla scenę z góry, aby odtworzyć właściwe cienie spowodowane zmieniającą się pozycją słońca. Naukowcy z Uniwersytetu Pn. Karoliny zainteresowani są przede wszystkim użyciem obiektów fizycznych jako prawdziwych trójwymiarowych urządzeń do wyświetlania w bardziej ogólnej grafice komputerowej, wizualizacji i zastosowaniach estetycznych (artystycznych).

Tworzenie modeli i wyświetlanie architektury na podstawie zdjęć

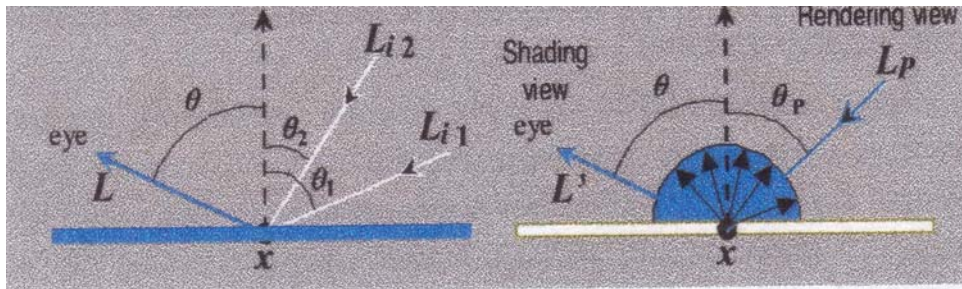
W projekcie o nazwie “Fasada” niewielki zestaw fotografii został użyty do stworzenia modeli i wyświetlania zabytków architektonicznych. Jest to dobry przykład metody łączonej wykorzystującej geometrię i obrazy do odtwarzania fizycznych struktur stworzonych rękami ludzkimi. Główne wyzwania do pokonania w tej metodzie związane są z zagadnieniami okluzji, próbkowania oraz mieszania,

które pojawiają się podczas ponownego wyświetlania obrazów na modele geometryczne. Autorzy tego projektu stają naprzeciw tym wyzwaniom wyposażeni w obrazy komputerowe i modele analityczne, następnie, używając lamp cieniowych naukowcy muszą rozwiązywać te problemy przy użyciu rzeczywistych (wyświetlanych) obrazów i modeli fizycznych. Wygodnym byłoby użycie narzędzi Fasady do zbudowania łączonej geometrii i modelu obrazu miasteczka uniwersyteckiego, by potem użyć technik lamp cieniowych do zanimowania skalowanego modelu fizycznego do skutecznego stworzenia “żywej dioramy” tego miasteczka.

Aby zrealizować zastosowanie tej metody, trzeba, pomiędzy innymi rzeczami, mieć metodę zaginania obrazu, tak, aby “pasował” do modelu fizycznego, aby ten był postrzegany przez obserwatorów jako poprawny. Pewne ograniczone dwuwymiarowe efekty zostały osiągnięte w tworzeniu modelu pojawiania się kurtyn teatralnych, tak aby wyglądały na poprawne z perspektywy widowni. Projekt Biura Przyszłości przedstawia techniki wyświetlania obrazów na powierzchni niepłaskie. Aktualnie naukowcy korzystają z technik, które opierają się na wyżej opisanych aby oświetlać potencjalnie niewypukłe lub niepołączone zestawy obiektów i przedstawiają nowe techniki zajmujące się zagadnieniami wyrównywania, okluzji, próbkowania i mieszania.

Proces wyświetlania

Przedstawiono tutaj ideę zamiany terminologii w relacjach pomiędzy oświetlaniem i odbiciem w odtwarzaniu odpowiedniej jasności na danej powierzchni. Jak widać na rys. 3, promieniowanie w danym kierunku w punkcie (x), który dał BRDF w świecie rzeczywistym (po lewej), można naśladować zmieniając BRDF i oświetlając ten punkt z odpowiednio dobranego źródła światła, np. z pikselu projektora (po prawej). Poniżej określone są równania na dostosowanie jasności dla ustalenia koniecznej intensywności pikselu projektora, dla danej pozycji i orientacji obserwatora i sceny wirtualnej. Jeśli chodzi o bardziej systematyczny schemat wyświetlania, opisany jest pojęcie dzielenia widoku wyświetlania – tradycyjnego wirtualnego widoku kamery od widoku cieniowego – pozycji obserwatora dla obliczeń oświetlenia.



Rys. 3 (po lewej) Jasność w punkcie w kierunku (θ, ϕ) . (po prawej) Jasność, która jest wynikiem oświetlenia z lampy projektorowej. Dzięki przestawieniu parametrów na ścieżce optycznej, można zrównać oba odczyty.

Po pierwsze należy rozpatrzyć równanie wyświetlania, które są zasadniczo przybliżeniem optyki geometrycznej. Jasność w widocznym punkcie powierzchni (x) w kierunku (θ, ϕ) , która dotrze do obserwatora fizycznego wygląda sceny to:

równanie nr (1)

$$L(x, \theta, \phi) = g(x, \theta, \phi)(L_e(x, \theta, \phi) + h(x, \theta, \phi))$$

gdzie

równanie nr (2)

$$h(x, \theta, \phi) = \int_i F_r(x, \theta, \phi, \theta_i, \phi_i) L_i(x, \theta_i, \phi_i) \cos(\theta_i) d\omega_i$$

a $g(x, \theta, \phi)$ to termin geometryczny (widoczność i odległość). $L_e(x, \theta, \phi)$ to emitowana jasność w danym punkcie (różne od zera tylko dla źródeł światła); $F_r(x, \theta, \phi, \theta_i, \phi_i)$ to BRDF punktu. Całka w $h(x, \theta, \phi)$ odpowiada za całe odbicie przypadkowej jasności $L_i(x, \theta_i, \phi_i)$ od kątów przestrzennych $d\omega_i$. Jasność ma wymiary energii na jednostkę czasu, obszaru i kąta przestrzennego.

Jeśli traktujemy lampę projektorową jako emiter punktowy, to jasność wynikająca z bezpośredniego oświetlenia przez projektor w tym samym punkcie powierzchni w odległości $d(x)$, ale ze współczynnikiem rozproszonego odbicia $k_u(x)$ jest dana przez:

równanie nr (3)

$$L'(x, \theta, \phi) = g(x, \theta, \phi) k_u(x) I_p(x, \theta_p, \phi_p) \cos(\theta_p) / d(x)^2$$

gdzie $I_p(x, \theta_p, \phi_p)$ = intensywność jasności projektora w kierunku (θ_p, ϕ_p) i jest

związana z nieciągłą wartością pikselu przez filtrowanie i odwzorowanie tonu.

Odtwarzamy jasność $L'(x, \theta, \phi)$ równą $L(x, \theta, \phi)$ dla danego położenia obserwatora przez rozwiązanie równania (3) dla I_p :

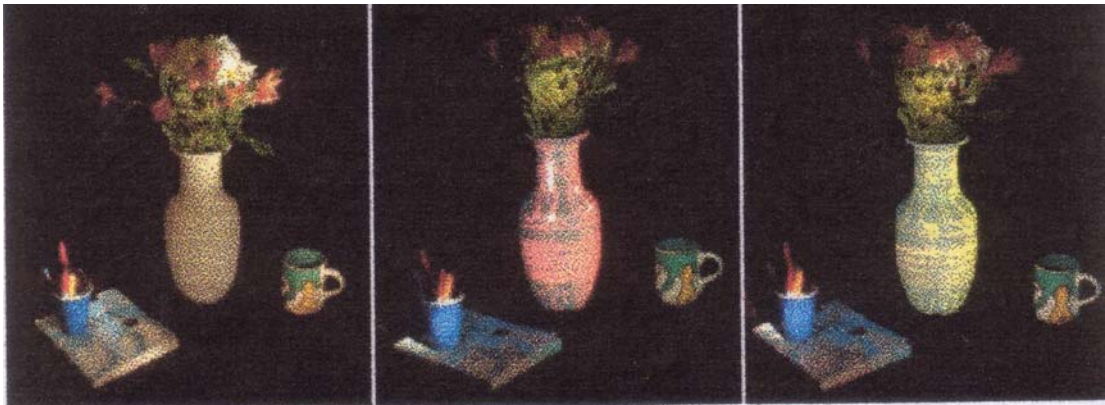
$$I_p(x, \theta_p, \phi_p) = \frac{L(x, \theta, \phi) d(x)^2}{k_u(x) \cos(\theta_p)} \quad \text{for } k_u(x) > 0.$$

Dlatego, tak długo jak współczynnik rozproszonego odbicia $k_u(x)$ jest różny od zera dla wszystkich długości fal przedstawionych w $L(x, \theta, \phi)$, to można skutecznie przedstawić właściwość powierzchni odpowiednią jasnością pikseli. Jednak w praktyce zakres wartości, jakie można przedstawić jest ograniczony przez jasność, zasięg dynamiczny oraz rozdzielczość pikseli projektora.

Proces wyświetlania zawiera w sobie dwa punkty widzenia: obserwatora i projektora. Prostą metodą byłoby najpierw wyświetlić obraz taki, jaki jest widziany przez obserwatora, który jest przedstawiony przez $L(x, \theta, \phi)$, a następnie użyć tradycyjnych technik wyświetlania opartych o obrazy, aby dopasować obraz i wygenerować obraz z projektora z poprawioną intensywnością świecenia, przedstawiony przez $I_p(x, \theta_p, \phi_p)$. Można również zaimplementować zależne od obserwatora cieniowanie w warunkach oświetlenia statycznego dla zmieniającej się pozycji obserwatora. Można ominąć dopasowywanie obrazu w przypadku gdy przedmiot wyświetlania jest taki sam jak obiekt wirtualny. Przy wyświetlaniu jednorzebiegowym traktujemy punkt widzenia przemieszczającego się obserwatora jako widok cieniowy. Wtedy proces syntezy obrazu wymaga wyświetlania sceny z punktu widzenia projektora, przy użyciu macierzy projekcji perspektywicznej, która odpowiada parametrom wewnętrznym i zewnętrznym projektora, a następnie dostosowana jest jasność. Oddzielenie tych dwóch punktów widzenia daje ciekawy temat do dyskusji. Na przykład, obliczenia widoczności i cieniowania niezależnego od punktu widzenia dla projektora statycznego mogą być dokonywane tylko raz, nawet jeśli punkt widzenia obserwatora jest zmienny.

Aby doprowadzić do interaktywnej implementacji w czasie rzeczywistym, korzysta się z konwencjonalnego wyświetlania trójwymiarowego API, które jest jedynie przybliżone do równania ogólnego wyświetlania. Obliczenie BRDF jest podzielone na: kierunkowe zależne od punktu widzenia oraz niezależne od punktu widzenia rozproszone i otaczające. Obliczeń cieniowania niezależnego od punktu

widzenia można dokonać przyjmując, że widok wyświetlania i cieniowania jest taki sam. (Cienie wirtualne, także niezależne od punktu widzenia, wyliczane są przy użyciu tradycyjnej dwuprzebiegowej techniki bufora cienia.) Jednak dla cieniowania niezależnego od punktu widzenia, takiego jak kierunkowe rozjaśnianie wybranych miejsc (Rys. 4) nie ma istniejącej podpory do oddzielenia tych dwóch punktów widzenia. W dodatku opisane są wymagane modyfikacje.



Rys. 4 (po lewej) Podstawą fizyczną do prezentacji jest wazon o białym rozmytym kolorze. (w środku i po prawej) Efekty zależne od punktu widzenia, takie jak kierunkowe rozjaśnianie wybranych miejsc, można wygenerować śledząc pozycję obserwatora i wyświetlając obrazy na wazonie.

Rozpraszanie drugorzędne

Lampy cieniowe mają pewne ograniczenia w dziedzinie właściwości powierzchni, jakie można odtworzyć. Poza tym, ponieważ używa się powierzchni neutralnych, drugorzędne rozpraszanie jest nie do uniknięcia i potencjalnie może wpłynąć na jakość wyników. Kiedy obiekt - podstawa prezentacji jest całkowicie rozmyty, to rozpraszanie drugorzędne może czasem wyjść na korzyść. Relacje geometryczne, nazywane też czynnikami kształtu, pomiędzy częściami obiektów fizycznych naturalnie są takie same jak relacje pomiędzy częściami obiektu wirtualnego. Rozważmy rozwiązanie RADIOSITY (metody wyświetlania opartej na dokładnej analizie odbicia światła od dużych powierzchni) dla działki i w scenie wirtualnej, gdzie m to ilość źródeł światła, a n to ilość działek:

równanie (5) $B_{i\text{-zamierzony}}$

$$B_{i\text{-intended}} = k_{d_i} \sum_j B_j F_{i,j} = k_{d_i} \left(\sum_m B_m F_{i,m} + \sum_n B_n F_{i,n} \right)$$

W równaniu tym k_d to odbicie rozproszone, B_j to jasność działki j , a F_{ij} to czynnik formy pomiędzy działkami. Korzystając z lamp cieniowych do odtwarzania efektu bezpośredniego oświetlenia (po dostosowaniu jasności) możemy wygenerować efekt m źródeł światła:

równanie (6)

$$B_{i-direct} = k_{d_i} \sum_m B_m F_{i,m}$$

Jednakże z powodu rozproszenia drugorzędowego, jeśli powierzchnie neutralne mają odbicie rozproszone k_u to odbierana jasność będzie zawierała drugorzędne rozproszenie z powodu n działek, co daje równanie (7)

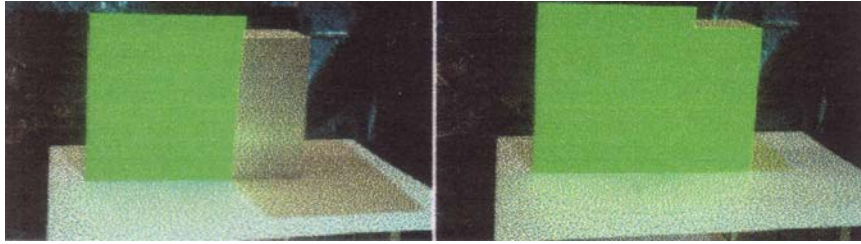
$$B_{i-actual} = B_{i-direct} + B_{i-secondary} = k_{d_i} \sum_m B_m F_{i,m} + k_u \sum_n B_n F_{i,n}$$

Różnica pomiędzy żadaną a odbieraną jasnością wynosi:

$$(k_{d_i} - k_u) \sum_n B_n F_{i,n}$$

równanie (8)

Z tego powodu, w przypadkach, w których k_d i k_u są podobne, otrzymujemy “przybliżoną jasność za darmo” – projekcja nawet prostego oświetlenia powoduje wiarygodne “wylewanie” kolorów na sąsiadujące części obiektów fizycznych. Z powyższego równania wynika, że drugorzędny udział powierzchni neutralnych z pewnością nie jest dokładny, nawet jeśli pierwszy odskok odtworzymy dokładnie. Różnica jest jeszcze większa kiedy obiekt wirtualny ma nielampercyjne wartości odbijające. Obecnie badane są metody *odwrotnego globalnego oświetlenia*, tak aby wyświetlony obraz mógł dać dokładniejszy efekt *globalnego oświetlenia*. Rys 5. przedstawia zielony i biały papier z rozlaniem z naturalnego białego i wyświetlonego zielonego oświetlenia. W tym szczególnym przypadku drugorzędne rozproszenie poza brzegi poziomej białej powierzchni poniżej oświetlanego przedmiotu jest podobne dla obu ścieżek.



Rys. 5 (po lewej) Zielony papier oświetlony białym światłem. (po prawej) Biała rozmyta powierzchnia oświetlona zielonym światłem.

Oświetlanie wszystkich widocznych powierzchni

Jeśli mamy dany obiekt fizyczny, to jaki będzie dobry zestaw punktów widzenia dla lamp, tak aby każda widoczna powierzchnia była oświetlana przynajmniej przez jedną lampę? Problemem tym zajmował się [Stuerzlinger99] i znalazł, przy użyciu algorytmu hierarchicznej widzialności, zestaw punktów widzenia kamer, dzięki któremu na każdą widoczną część powierzchni obraz nałożony jest przynajmniej raz. Problem określenia optymalnej liczby punktów widzenia jest trudny ze względu na brak algorytmu do ich obliczania i podobny jest do problemu galerii sztuki znanego w dziedzinie geometrii obliczeniowej.

Metody

Oświetlanie obiektów fizycznych oparte na obrazach było badane przez wielu naukowców. Dwa główne wyzwania, które sprawiały, że wysiłki skupiały się do tej pory na drogich i, na dużą skalę wdrożeniach: (a) Pierwsze wyzwanie to problem rejestracji geometrycznej, który się objawia jako dopasowanie projekcji prostego obrazu dwuwymiarowego do obiektu. Projekcja urządzenia perspektywicznego ma aż 11 stopni swobody (5 wewnętrznych i 6 zewnętrznych), dlatego też każda próba ręcznego wyjustowania obrazu z obiektem może być niesamowicie pracochłonna i męcząca. (b) Drugi problem, który do tej pory nie był badany, to całkowite oświetlenie nieszablonowych obiektów fizycznych w obecności cieni spowodowanych okluzją. Z pojawieniem się cyfrowych projektorów i wyświetlania trójwymiarowej grafiki w czasie rzeczywistym, możliwe stało się wykorzystanie nowego sposobu oświetlania opartego na obrazach. Rozwiązuje się ten problem tworząc trójwymiarowe porozumienie geometryczne pomiędzy częściami zestawu.

Dostosowywanie do potrzeb i wyrównywanie

Jednym z ważnych zadań tworzenia efektywnych wizualizacji jest stworzenie połączenia pomiędzy obiektami fizycznymi i graficznymi funkcjami pierwotnymi, które uwydatnią te obiekty podczas wyświetlania. Na przykład, jak określić, jaka

tekstura powinna zostać użyta dla ścian modelu budynku albo jakie kolory będą lepsze dla obiektu fizycznego? Potrzebny jest obiekt fizyczny oraz jego geometryczny odpowiednik trójwymiarowy i rzeczywisty lub żądana właściwość powierzchni. Pojawiło się wiele rozwiązań sprzętowych i oprogramowania do skanowania i drukowania obiektów trójwymiarowych oraz przechwytywania i tworzenia modeli graficznych z teksturami i o dużej jakości szczegółów. Można interaktywnie dostosowywać obraz do własnych potrzeb “malując” bezpośrednio na obiektach fizycznych. Rezultaty interakcji użytkownika mogą zostać wyświetlone na obiektach, a także przechowywane w komputerze. Idealną sytuacją byłoby użycie interfejsu użytkownika z większymi możliwościami, aby tworzyć i edytować pierwotne funkcje graficzne różnych kształtów, kolorów i tekstur. Aby właściwie ustawić projektor, najpierw trzeba ustawić projektor w przybliżonej pozycji, a następnie dostosować do jego geometrycznej relacji z obiektem. Relacja ta jest wyliczana przez odnajdowanie wewnętrznych parametrów projektora i sztywnego odwzorowania pomiędzy tymi dwoma systemami współrzędnych. Jest to klasyczny problem wizji komputerowej. Przyjmujemy zbiór cech charakterystycznych o znanych pozycjach na obiekcie fizycznym i odnajdujemy odpowiadające piksele projektora, które je oświetlają. To pozwala wyliczyć macierz projekcji perspektywicznej 3×4 w skali, która jest rozkładana, aby znaleźć wewnętrzne i zewnętrzne parametry projektora. W procesie wyświetlania wykorzystane są te same parametry wewnętrzne i zewnętrzne, tak aby wyświetlane obrazy były właściwie ustawione w stosunku do obiektów fizycznych.

Korekta intensywności świecenia

Intensywność wyświetlanego obiektu jest modyfikowana przy użyciu równania (4), aby uwzględnić współczynnik odbicia powierzchni neutralnej, orientację w danym punkcie oraz odległość w stosunku do projektora. Ponieważ normalne powierzchni używane do wyliczania poprawki $1/\cos(\theta_p)$ dostępne są tylko dla wierzchołków wielokątnych modeli graficznych, używa się przewodu wyświetlania do przybliżonej interpolacji. Oświetlamy białą rozmytą wersję modelu graficznego (lub model odpowiadający właściwemu $k_u(x)$ modelu fizycznego) białym wirtualnym światłem umieszczonym w miejscu lampy projektorowej i wyświetlamy je z czarną mgłą dla uzyskania kwadratowego osłabienia odległości. Intensywność świecenia

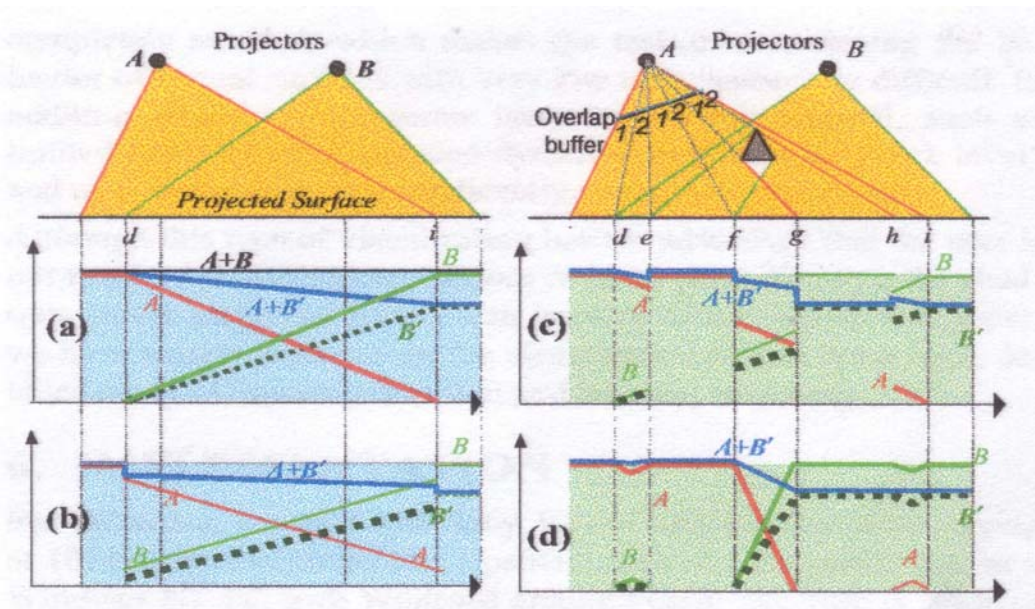
wynikająca z tego działania jest gładka na zakrzywionych powierzchniach dzięki interpolacji cieniowania i jest przeciwnie proporcjonalna do czynnika $(d(x)^2 k_u(x) \cos(\theta_p))$. Aby bardziej efektywnie wykorzystać ograniczony zakres dynamiczny projektorów, nie oświetlana jest powierzchnia, gdzie $\theta_p > 60$ (ponieważ zakres $1/\cos(\theta_p)$ wynosi od 2 do nieskończoności). Dzięki temu unika się niskiej prędkości samplowania wyświetlanych pikseli na skośnych powierzchniach oraz minimalizuje się sztuczności niewłaściwego justowania spowodowane jakimikolwiek błędami kalibracji geometrycznej. Podczas obliczeń znajdowania miejsc nakładania się na siebie marginesów (co opisane jest poniżej), bardzo skośne powierzchnie są uznawane za nieoświetlone przez dany projektor. Rys. 7 jest przykładem takiego działania.

Okluzje i zakładki (nakładające się na siebie marginesy)

Jeżeli chcemy uzyskać pełne oświetlenie sceny, bardzo oczywistym wyborem są dodatkowe projektory. Wybór ten prowadzi jednak do zaistnienia trudniejszego problemu, a mianowicie łączenia obrazów wieloprojektorowych tak, aby nie było widać łączeń. Można przyjąć bardzo proste rozwiązanie, które będzie wymagać ustawienia projektorów w ten sposób, aby dany fragment powierzchni oświetlany był tylko przez jeden projektor. Jednakże są dwa główne zagadnienia, które trzeba rozpatrzyć w przypadku nakładania się na siebie obrazów z projektorów typu CRT, LCD czy DLP, które zmuszają do użycia technik stapiania (lub techniki płynnego przechodzenia) intensywności świecenia. Pierwsze zagadnienie to brak równości kolorów pomiędzy sąsiadującymi projektorami, wynikający z procesu produkcyjnego projektorów i przesuwania się koloru pod wpływem temperatury w trakcie ich użytkowania. Drugie to chęć zminimalizowania wrażliwości na małe błędy w ocenie parametrów czy różnic mechanicznych kalibracji geometrycznej.

Technika stapiania jest często używana, gdy chce się wygenerować fotomozaiki panoramiczne bez widocznych łączeń poprzez łączenie kilku widoków z pojedynczego miejsca. Podobne techniki wykorzystywane są w wieloprojektorowych wyświetlaczach o szerokim polu widzenia oraz w dwuwymiarowych układach projekcji płaskich. W takich przypadkach rejon zakładki to zazwyczaj (dobrze zdefiniowany) przylegający rejon na powierzchni wyświetlania, a także w buforze klatek każdego projektora. W algorytmie użytym w

projektach [Szelinski97, Raskar99], intensywność świecenia pikselu jest ważona jako proporcjonalna do odległości euklidesowej od najbliższego pikselu granicznego (udział zerowy) (wyświetlanego) obrazu. Ciężkości te są mnożone do intensywności świecenia w końcowym wyświetlanym obrazie i znajdują się w przedziale [0,1]. Piksele w pobliżu granicy źródła obrazu mają niewielki udział, więc przejście do następnego źródła obrazu jest płynne. To prowadzi do często spotykanego spadku intensywności, co widać na rysunku 6(a). W idealnych warunkach i po przyjęciu równowartości koloru, ciężar wpływu obu projektorów $A+B$ dodaje się i wynosi 1. Nawet jeśli reakcja kolorystyczna projektora B jest inna niż projektora A (jest osłabiona, co pokazano jako B'), wynikające z tego $A+B'$ (kolor niebieski) płynnie przechodzi w rejonie zakładki.



Rys. 6 Pomiary intensywności przy użyciu metod stapiania. Wykresy pokazują udział projektorów A, B i B' i wynikającą akumulację $A+B$ oraz $A+B'$ wzdłuż oświetlonej powierzchni. Technika, pokazana w punkcie (d), tworzy płynne przejścia ciężaru.

Na rysunku: Projectors – projektory / Overlap buffer – bufor zakładki / Projected surface – wyświetlana powierzchnia

Ta strategia przydzielania ciężkości działa bardzo dobrze, kiedy docelowy obraz oświetla gładką ciągłą powierzchnią na zakładce i naokoło niej. W tym przypadku model fizyczny jest zazwyczaj złożony z niewypukłych obiektów lub ze zbioru niepołączonych obiektów, co powoduje cienie, podzielone rejony zakładek oraz, co najważniejsze, rejony zakładek zawierające nieciągłości głębi, co pokazane jest na

rys. 6(c) z prostym okluderem. Jeżeli istnieją nierówne reakcje kolorystyczne, wynikający z nich rozdział ciężkości $A+B'$ ma rażące nagłe zmiany, jak w punktach f i g. Sytuacja ta jest analogiczna do wyświetlania opartego o obrazy (IBR), gdzie dopasowanie pojedynczego obrazu o uwydatnionej głębi do obrazu docelowego prowadzi do realizacji dysokluzyjnych. Kiedy obrazy wieloźródłowe są dopasowywane do docelowego obrazu, to kolor przypisany danemu pikselowi musi zostać wyprowadzony (albo od pojedynczego obrazu, w którym na siebie nachodzą lub) jako wyważona kombinacja zgodnych pikseli z obrazów źródłowych. Stapianie, które w rzeczywistości rozmywa wynik, jest zazwyczaj konieczne do pokonania (pomniejszych) różnic koloru w zgodnych pikselach w obrazach wchodzących i do ukrycia efektu zjawy (spowodowanego małymi błędami justowania). Jednym z niewielu dostępnych rozwiązań jest rozwiązanie zaproponowane przez [Debevec98], w którym reguluje się intensywność świecenia przy użyciu ciężkości proporcjonalnych do kątów pomiędzy widokiem docelowym a widokami źródłowymi. Jak napisano we wspomnianej pracy: “nie gwarantuje to, że ciężkości będą płynnie przechodzić przez powierzchnie sceny. Rezultatem tego może być pojawienie się na obrazie widocznych łączeń w miejscach, gdzie na sąsiadujące wielokąty wyświetlane są bardzo różne kombinacje obrazów.” Wykresy na rys. 6(b) przedstawiają próbkę schematu ważenia opartego na podobnej idei i związane z nim problemy. Poniżej przedstawione jest ogólne rozwiązanie wykorzystujące nowy algorytm stapiania, który pasuje do IBR jak również do lamp cieniowych. Algorytm ten oparty jest na następujących wskazówkach:

1. Suma ciężkości intensywności zgodnych pikseli projektorów wynosi jeden, tak aby intensywności były znormalizowane.
2. Ciężkości pikseli projektora wzdłuż powierzchni fizycznej zmieniają się płynnie w miejscach zakładek i naokoło nich, aby różnice w kolorach pomiędzy projektorami nie stwarzały widocznych nieciągłości w wyświetlanym obrazie oraz
3. rozprowadzanie ciężkości intensywności dla danego projektora w jego buforze klatek jest płynne, aby małe błędy kalibracyjne lub różnice mechaniczne nie powodowały ostrych brzegów.

W praktyce łatwiej jest uzyskać (lub podtrzymać) dokładną kalibrację geometryczną niż zapewnić jednolitość koloru pomiędzy zestawem projektorów przez dany okres czasu. To sprawia, że warunek (2) jest ważniejszy niż warunek (3). Jednakże nie

zawsze można spełnić warunki (2) lub (3) (na przykład, jeśli okluder przesunie się bliżej płaszczyzny, co spowoduje, że $f=g$ na rys. 6) i powyższe punkty traktuje się jak wskazówki, a nie jak zasady.

Te trzy wskazówki sugerują rozwiązanie problemu stapiania bez ograniczeń ciężkości przy nieciągłości głębi i w miejscach granic cieni. Tradycyjne metody stapiania wykorzystują odległość do najbliższego piksela granicznego do znajdowania ciężkości. W tej metodzie najpierw odnajduje się piksele odpowiadające rejonom oświetlonym przez jeden projektor i przypisuje się im ciężkość intensywności wynoszącą 1. Następnie - dla każdego pozostałego piksela - znajduje się najkrótszą euklidesową odległość do piksela o wartości 1, ignorując ścieżki, które przecinają miejsca nieciągłości głębi. Ta przypisana ciężkość jest odwrotnie proporcjonalna do odpowiadającej odległości. Rys 6(d) pokazuje wynik nowego algorytmu stapiania dla dwóch projektorów. Nawet przy różnych reakcjach kolorystycznych (A+B') algorytm ten generuje płynne przejścia na płaszczyźnie powierzchni w obecności cieni i urywanych zakładek. Można używać tego algorytmu dla trzech lub więcej projektorów bez konieczności wprowadzania modyfikacji.

W praktycznej realizacji wykorzystuje się dwa bufory - bufor zakładki i bufor głębi. Bufor głębi jest aktualizowany przy wyświetlaniu modelu graficznego. Bufor zakładki zawiera wartości liczb całkowitych, które mają wskazywać ilość nachodzących na siebie projektorów dla każdego piksela. Rejony zachodzące na siebie (np. podwójnie lub więcej razy) są wyliczane przy użyciu tradycyjnego algorytmu bufora cienia. Algorytm ten jest następujący:

W każdym projektorze,

Obliczyć granice zakładek pomiędzy rejonami w liczbie 1 i >1

Obliczyć nieciągłości głębi używając progu

Dla każdego pikselu, zaktualizować najkrótszą odległość do rejonu, w którym ilość zakładek = 1

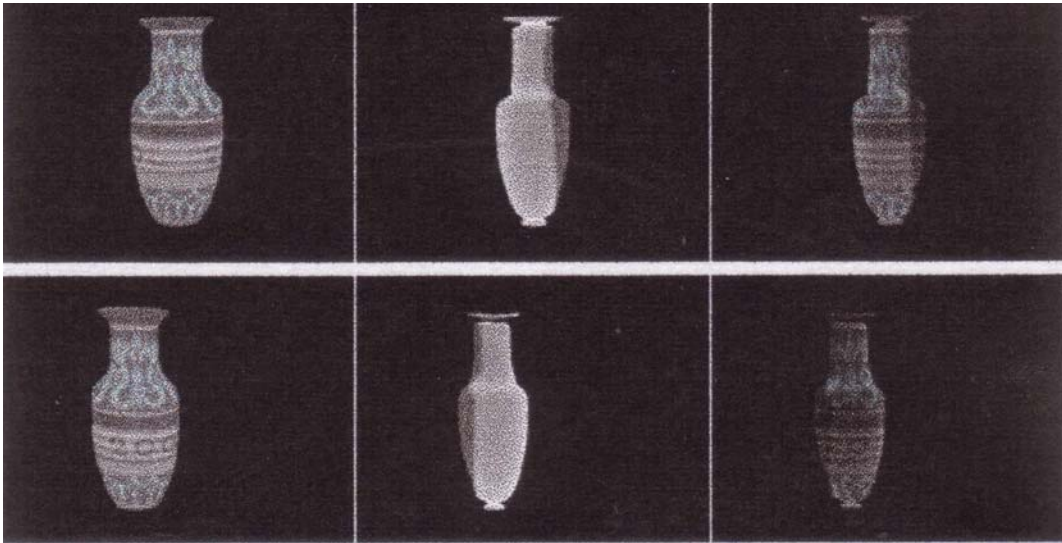
Dla każdego pikselu, dla którego ilość zakładek > 1 w każdym projektorze

Odnaleźć wszystkie zgodne piksele w innych projektorach

Przypisać ciężkości odwrotnie proporcjonalne do najkrótszej odległości

Dla niektórych pikseli w rejonie zakładki, takim jak rejony $[h,i]$ nie można znaleźć najbliższego piksela o ilości zakładek równej 1, więc najkrótsza odległość jest ustawiana na dużą wartość. Takie działanie pierwszorzędnie redukuje ciężkości w

odizolowanych rejonach, a także zmniejsza niepotrzebne strefy przejść. Rys. 7 przedstawia zestaw obrazów do oświetlania wazonu, które mają poprawione ciężkości i intensywność świecenia.



Rys. 7 Oświetlanie wazonu (po lewej) Wyświetlane obrazy (w środku). Obrazy ciężkości intensywności, zawierające eliminację zbędnych części i poprawki orientacji powierzchniowej i zakładek. (po prawej) Obrazy o poprawionej intensywności.

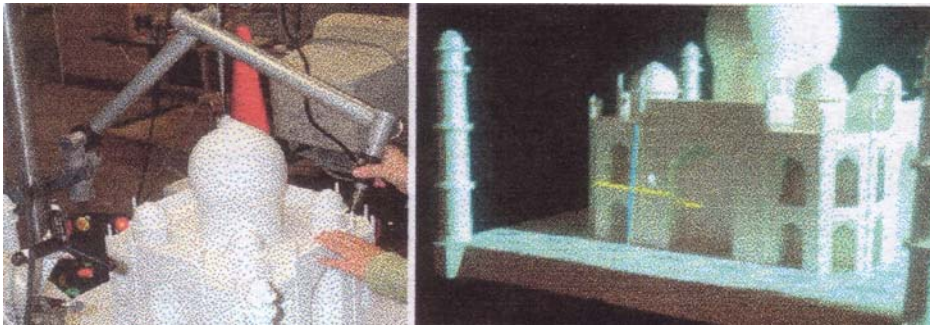
Ograniczenia

Głównym ograniczeniem dla lamp cieniowych jest zależność od właściwości neutralnej powierzchni fizycznej i kontrolowanego (ciemnego) oświetlenia otaczającego. Nie można całkowicie uniknąć problemu spowodowanego drugorzędym rozproszeniem, co sprawia że odtwarzanie zachowania powierzchni wirtualnych przy niskim współczynniku odbicia jest bardzo trudne. Poza tym, ograniczenia tradycyjnych projektorów, takie jak ograniczona głębokość pola, zredukowany zakres dynamiczny spowodowany “czarnym poziomem” i niejednorodnością, mogą w sposób znaczący wpłynąć na jakość wizualną.

Mimo że ten typ wizualizacji posiada tę zaletę, że użytkownik nie musi zakładać okularów ani hełmów wirtualnych, to cienie na wyświetlanych powierzchniach mogą bardzo przeszkadzać. W opisie skupiono się głównie na aspekcie wizualizacji, ale konieczne jest dokładniejsze studium interakcji człowieka i związanych z nią zagadnień.

Implementacja (realizacja projektu)

W zestawie użyto projektorów SONY VPLK6000U wyświetlających w rozdzielczości 1024x768. Programy do wyświetlania OpenGL działają na Windows NT PC z kartą graficzną Wildcard. Wazon został wykonany z gliny i ma wymiary, w przybliżeniu, 12x12x35 cm. Model Tadź Mahal jest drewniany i pomalowany sprayem na biało. Jego wymiary to w przybliżeniu 70x70x35 cm. Oba modele zostały zeskanowane, (w każdym przypadku zajęło to około 30 min.) przy pomocy trójwymiarowego skanera dotykowego z sondą, którego odczyty mają dokładność do 0,5 mm. Model wazonu składa się z 7 000 trójkątów, a model Tadź Mahal z około 21 000 trójkątów i 15 map tekstur. Do stworzenia efektu kierunkowego rozjaśniania użyto optycznego systemu śledzącego Dyna Sight Origin Instruments do śledzenia położenia obserwatora.



Rys. 8 (po lewej) Trójwymiarowy skaner dotykowy z sondą służący do stworzenia trójwymiarowego modelu obiektu rzeczywistego. (po prawej) Projektory są kalibrowane w stosunku do modelu przez odnajdowanie, które piksele (centrum krzyżowania się) oświetlają znane trójwymiarowe punkty podstawowe.

Każdy projektor jest kalibrowany przez odnajdowanie pikseli, które oświetlają zbiór 20 znanych trójwymiarowych punktów podstawowych na modelu fizycznym. Dokonuje się kalibracji przez przesuwanie wyświetlanej siatki nitek w przestrzeni obrazu projektora, tak aby jej centrum zbiegało się ze znanymi punktami podstawowymi. Matryca projekcji perspektywicznej 3x4 i jej rozłożenie na parametry wewnętrzne i zewnętrzne projektora wyliczane są przy użyciu Matlab. Proces wyświetlania wykorzystuje te parametry, tak aby wyświetlane obrazy były wyjustowane w stosunku do modelu. Skalibrowanie każdego projektora zajmuje mniej niż 5 minut. Zazwyczaj błąd ponownego wyświetlania jest mniejszy niż 2

piksele i obrazy z dwóch projektorów wyglądają na modelu na dopasowane geometrycznie. Ciężkości intensywności świecenia dla pikseli projektorów są wyliczane podczas przetwarzania wstępnego i zajmuje mniej więcej 10 sekund na każdy projektor. Podczas wyświetlania, intensywność jest modyfikowana przy użyciu alfa-mieszania, które jest dostępne przy sprzęcie graficznym.

Zastosowania

W swojej najprostszej formie lampy cieniowe mogą być używane, by dynamicznie zmieniać kolory codziennych obiektów lub dodawać im tymczasowe znaki. Na przykład, inżynierzy mogą zaznaczyć interesujący ich obszar, na przykład miejsca odwiertów bez zmieniania wyglądu obiektu fizycznego. Jak widać na rys. 1, można wyświetlać cienie wirtualne na skanowanych modelach. Osoby zajmujące się zagospodarowaniem przestrzeni miejskich mogą przesuwając budynki i pokazywać efekty globalne w trójwymiarze na blacie stołu zamiast na monitorze komputera. W pokazach scenicznych można zmieniać nie tylko położenie kurtyn, ale także symulować pory roku czy starzenie się obiektów na scenie. Zamiast przypadkowego wyświetlania laserowych obrazów wektorowych, naukowcy chcą tworzyć kształty przy użyciu wyświetlaczy laserowych na ogromnych budynkach przez kalibrowanie urządzeń laserowych w stosunku do trójwymiarowych modeli tych budynków. Można także symulować ruch, przez wyświetlanie zmieniającej się tekstury na stacjonarne, symetryczne obiekty obrotowe. Można również wygenerować interesujące nie-fotorealistyczne efekty.

Śledzony obserwator

Przy użyciu prostego śledzenia ruchów obserwatora zademonstrowano jak gliniany wazon może wyglądać jak metalowy lub plastikowy. Łatwo jest wyświetlać zależne od położenia obserwatora efekty, takie jak odbicia. Koncepcję tę można rozwinąć i wykorzystać do większych projektów. Rzeźbiarze często używają glinianych modeli dużych pomników nim je stworzą. Mogłoby być dla nich pomocne, gdyby zobaczyli jak formy geometryczne, które stworzyli będą wyglądać w różnych materiałach i w różnych warunkach w otoczeniu innych obiektów. Przez wyświetlanie punktów lub linii prowadnic (np. ramy sitowej) z modeli komputerowych rzeźbiarze mogą sprawdzić poprawność geometryczną modeli glinianych. Oświetlenie oparte na obrazach może być bardzo efektywnie wykorzystywane w studiach filmowych, gdzie wykonuje się pracochłonne zadanie budowania modeli miniaturowych, a potem nakłada się na nie zaktualizowane

szczególności. Aby wstawić syntetyczne postaci do przeglądowego widoku zbioru modeli miniaturowych, można wyświetlić sylwetkę poruszającej się postaci wirtualnej, która według kamery śledzącej ruch wygląda na perspektywicznie poprawną. Będzie to przewodnikiem w ustawianiu w czasie końcowego przetwarzania ponieważ nie będą wymagane wewnętrzne parametry kamery.

Śledzone obiekty

Można oświetlać obiekty, tak aby tekstury powierzchni wyglądały na przyklejone do obiektów nawet gdy te są w ruchu. W takim przypadku, możemy wyświetlać zaktualizowane kierunkowe rozjaśnianie nawet dla obserwatora statycznego. Na przykład, w oknach salonów wystawowych lub w salonach pokazowych można pokazać poruszający się model produktu ze zmieniającymi się kolorami lub innymi cechami. W systemie eksperymentalnym śledzony "pędzel" został użyty do namalowania na śledzonym poruszającym się prostopadłościanie trzymanym przez użytkownika. (rys. 9). Obecność modelu fizycznego pozwala na naturalne dotykowe sprzężenie zwrotne. Dwa główne problemy tej techniki to potrzeba zakładania urządzenia śledzącego i złe justowanie dynamiczne spowodowane opóźnieniem urządzenia śledzącego.



Rys.9 Śledzony "pędzel" maluje na śledzonym prostopadłościanie.

Proporcjonalne zwiększanie rozmiarów.

Rozpoczęto badania nad rozwinięciem tej metody w kierunku stworzenia modeli wirtualnych środowisk ludzkich rozmiarów, po których będzie można się poruszać. Zamiast budować dokładne szczegółowe repliki obiektów do wyświetlania, użyto wersji uproszczonych. Na przykład, podstawowe struktury wnętrza budynków i obiektów architektonicznych o średnich rozmiarach (ściany, kolumny, szafki, stoły itp.) można przedstawić w przybliżeniu przy użyciu prostych elementów (pudełek, tub itp.) W tym przypadku wykorzystano klocki ze styropianu. Główne cechy architektoniczne, które są dopasowywane do uproszczonych modeli fizycznych,

utrzymują trójwymiarową autostereofonię, ale inne szczegóły muszą być przedstawiane przy użyciu obrazów zależnych od widoku. Eksperyment symulacji wnętrza budynku dowodzi, że takie ustawienie może dać silniejsze uczucie przebywania w wyświetlanym środowisku w porównaniu z CAVETM, ponieważ użytkownik naprawdę może przemieszczać się w środowisku wirtualnym. Jednakże z powodu obecności ogromnych powierzchni wklęsłych (np. kątów pomieszczenia), problem ponownego odbijania światła staje się bardziej poważny. Ponadto, ponieważ prawie wszystkie powierzchnie wokół użytkownika muszą być oświetlane, łatwiej jest użytkownikowi zamykać (zasłaniać) niektóre projektory. Dlatego tak bardzo ważne jest ustawienie projektorów i ta sprawa (oprócz innych zagadnień) pozostaje najpoważniejszym wyzwaniem.

Pomysły

System modelowania w glinie oparty na lampach cieniowych może być użyty jako trójwymiarowa wersja “łączenia kropek”, aby pomóc modelarzowi ocenić jego projekt. Na przykład, dwa zsynchronizowane projektory mogą sukcesywnie wyświetlać obrazy różnych części modelu trójwymiarowego na czerwono i zielono. Właściwe ustawienie gliny weryfikowane jest przez żółte oświetlenie. Po utworzeniu kształtu ta sama lampa cieniowa może zostać użyta do przeprowadzenia w procesie malowania modelu lub w nakładaniu rzeczywistego materiału o odpowiadających właściwościach odbicia.

Istnieje możliwość zrealizowania interaktywnego systemu dotykowego z sondą skanującą z zamkniętą pętlą weryfikacji rekonstruowania powierzchni (mozaikowanie) poprzez nieprzerwane wyświetlanie uwypuklonych obrazów częściowej rekonstrukcji obiektu, który jest skanowany. Dzięki temu osoba dokonująca skanowania będzie wiedziała, jaka jest wymagana gęstość punktów, będzie знаła rejony, które nie zostały zsampleowane oraz obecne odchylenie obiektu wirtualnego w stosunku do odpowiadającego mu obiektu fizycznego.

Użyteczny dwuręczny zestaw trójwymiarowego modelowania i trójwymiarowego malowania wymagałby śledzenia punktu widzenia obserwatora, urządzeń wejściowych oraz obiektu o surowych kształtach (np. kuli). Użytkownik może dosłownie stworzyć i dodawać właściwości do obiektu wirtualnego, który jest właściwie ustawiony w stosunku do kuli.

Wnioski

Opisana została w tym rozdziale nowa technika wizualizacji trójwymiarowej grafiki

komputerowej, która wykorzystuje projektory świetlne i obiekty fizyczne do generowania obrazów o wysokiej jakości szczegółów bezpośrednio w świetle użytkownika. Mimo że metoda ma ograniczenia w porównaniu z tradycyjnym wyświetlaniem grafiki na ekranach komputerów, daje jednak nowy sposób interakcji z obrazem syntetycznym. Przedstawiono nowe techniki praktycznego zastosowania oświetlania obiektów nieszablonywych opartego o obrazy. Proces wyświetlania wymaga przede wszystkim określenia punktu widzenia użytkownika, kształtu obiektu graficznego, właściwości odbijania i oświetlenia. W tradycyjnej grafice komputerowej czy powiększanej rzeczywistości z hełmami wirtualnymi, wszystkie obrazy, które widzimy, mają zredukowaną rozdzielczość czasową (szybkość klatkowania) lub przestrzenną (piksele). W przypadku lamp cieniowych, punkt widzenia i kształt utrzymywany jest w najlepszej rozdzielczości, a tylko dodawany kolor ma ograniczoną rozdzielczość. Ta metoda wizualizacji może być bardzo efektywna w wielu zastosowaniach, na przykład w szkoleniach, projektach architektonicznych, sztuce i rozrywce.

Dodatki

Jak wynika z opisu w sekcji 3, podczas gdy widok wyświetlania zdefiniowany przez parametry projektora pozostaje niezmienny, widok cieniowy jest zależny od pozycji obserwatora. Przedstawiona jest niewielka modyfikacja tradycyjnego ustawienia widoku, aby rozdzielić dwa widoki przy użyciu przykładowego OpenGL API.

Przestrzennie Powiększana Rzeczywistość:

Ożywianie modeli fizycznych dzięki wyświetlanym obrazom

Mimo iż na rynku dostępne są systemy grafiki komputerowej wysokiej jakości, architekci i projektanci nadal często budują pomniejszone fizyczne modele budynków czy projektów. Modele fizyczne dają wyobrażenie obiektu, które ma wysoką rozdzielczość, może być oglądane w trój-wymiarze ze wszystkich stron przez wiele osób na raz, bez systemów wyświetlających czy okularów trójwymiarowych i można je fizycznie zmieniać (na przykład, można zdjąć dach modelu budynku, aby pokazać projekt wnętrza). Jednakże takie modele fizyczne są statyczne w swojej strukturze i cechach – w gruncie rzeczy są bez życia. Z drugiej strony, wysokiej jakości systemy graficzne są niesamowicie elastyczne, dzięki

czemu obserwatorzy mogą zobaczyć alternatywne struktury, elewacje, tekstury, częściowe przekroje perspektywiczne, a nawet efekty dynamiczne, takie jak zmiana oświetlenia, poruszające się samochody, ludzie itp.

Zaprezentowana będzie kombinacja tych metod, która bazuje na poprzednich osiągnięciach i demonstrowanych technikach Przestrzennie Powiększanej Rzeczywistości, do których używa się projektorów. Chodzi przede wszystkim o to, aby skierować wielokrotne projektory świetlne umieszczone na suficie do wewnątrz, aby oświetlić i powiększyć umieszczone na blacie fizyczne modele budynków czy produktów. To podejście gwarantuje bardzo efektowne łączone wizualizacje, które mają w sobie zalety tak tradycyjnych modeli fizycznych jak i nowoczesnej grafiki komputerowej, efektywnie “ożywiają” zwykłe modele fizyczne.

Omówione będą również dwa nowe wyzwania, jakie mogą się pojawić, kiedy wielokrotne kierowane do wewnątrz projektory będą używane do wizualnego powiększania modeli fizycznych. Pierwszym wyzwaniem jest określenie statycznej relacji pomiędzy połączonymi systemami kierowanych do wewnątrz projektorów a fizycznymi modelami na blacie. Drugim wyzwaniem jest wygenerowanie obrazu przy użyciu nakładania obrazów na siebie i potencjalnie nieprzerwanych projekcji padających na powierzchnie modeli fizycznych.

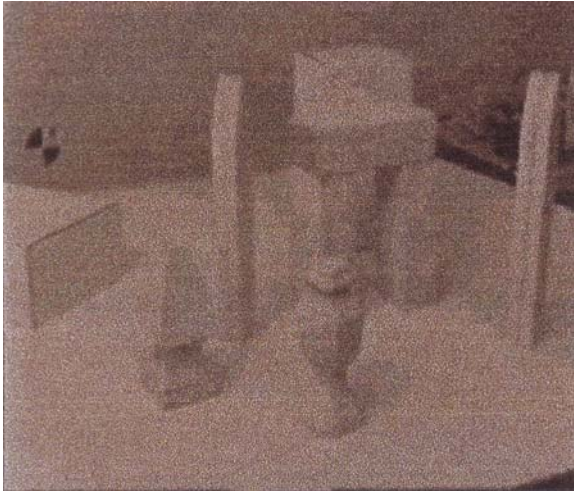
Wstęp



Rys. 1 Dwa różne spojrzenia na modele fizyczne powiększone przez wyświetlane obrazy. (Modele fizyczne, które są podstawą do tej prezentacji, pokazane są na rys.2)

Wcześniej przedstawione było ogólnie pojęcie Przestrzennie Powiększanej Rzeczywistości (SAR), w której obiekty fizyczne są powiększane przez obrazy, które są bezpośrednio zintegrowane ze środowiskiem użytkownika, a nie tylko z jego polem widzenia. Na przykład, obrazy mogą być wyświetlane na rzeczywistych

obiektach przy użyciu projektorów świetlnych lub bezpośrednio osadzone w środowisku dzięki płaskim wyświetlaczom. W tym konkretnym przypadku wielokrotne projektory umieszczone na suficie są kierowane do wewnątrz, aby oświetlały i mogły powiększać ustawione na blacie fizyczne pomniejszone modele budynków i innych obiektów.



Rys. 2 Modele fizyczne, które były podstawą do prezentacji z rys. 1. Te obiekty fizyczne wykonane są z drewna, cegły i kartonu.

Takie ustawienie zapewnia imponujące łączone pokazy, które mają zalety dotychczas zarezerwowane tylko dla modeli fizycznych lub graficznych. Podobnie jak tradycyjne modele fizyczne, powiększone modele fizyczne mogą być oglądane w trójwymiarze, z jakiegokolwiek miejsca wokół stołu, przez wiele osób na raz, bez okularów trójwymiarowych. Oprócz tego, tak jak rzecz ma się z typowym przedstawieniem graficznym, można z łatwością pokazywać różne właściwości powierzchni zmieniając oświetlenie, obiekty poruszające się oraz podawać różne pomocne informacje dwuwymiarowe. Dodatkowo, przy użyciu okularów trójwymiarowych można dokonywać wirtualnych modyfikacji modelu fizycznego, dodając lub usuwając elementy, pokazując strukturę wewnętrzną itp. W obu przypadkach wielokrotne projektory skupione do wewnątrz mogą być wykorzystywane, kiedy chce się uzyskać wysoką rozdzielczość i wysokie nasycenie obrazu. Kiedy używa się więcej niż jednego projektora, aby stworzyć obraz wirtualny, trzeba rozwiązać dwa główne problemy. Po pierwsze, trzeba skalibrować środowisko do prezentacji i uzyskać rejestrację statyczną pomiędzy projektorami, obiektami na polu działania i systemem śledzenia (tracking). Po drugie, trzeba rozwiązać problem generowania obrazów bez widocznych łączeń poprzez uzyskanie

rejestracji geometrycznej pomiędzy zachodzącymi na siebie projekcjami (z kilku projektorów). Należy użyć statycznych kamer wideo i standardowych technik aktywnej wizji komputerowej aby wygenerować konieczne trójwymiarowe przedstawienie parametrów projektora i powierzchni rzeczywistych obiektów (a także system współrzędnych urządzenia śledzącego, jeśli chce się otrzymać trójwymiarowy obraz wirtualny). Problem uzyskania obrazu bez widocznych łączeń przy użyciu wielokrotnych projektorów był badany dla prostych konfiguracji przez [InfoMural] [Raskar98d] [Raskar99a] [Trimensions] [Panoram].

Zastosowania

Opisana w tej pracy fizyczno-graficzna metoda przedstawiania oparta na modelach ma pewne ograniczenia w porównaniu z czysto fizycznymi lub graficznymi metodami. Jednakże daje ciekawą nową metodę realizowania efektywnych wyobrażeń obiektów wirtualnych, o wysokiej rozdzielczości i właściwości powierzchni współistniejących z rzeczywistym światem. Dwoma przykładami zastosowań są powiększane prezentacje modeli architektonicznych jednego lub więcej budynków oraz powiększane prezentacje części i procedur dla pracowników taśmy montażowej lub dla techników napraw.

W pierwszym przypadku architekt mógłby przedstawić klientom imponującą formę skanowanego modelu proponowanego budynku lub kompleksu, który można obejrzeć ze wszystkich stron. Najprostszym zastosowaniem tej metody, przyjmując że modele mają rozproszony biały kolor, jest użycie jej do “nakładania” różnych kolorów i tekstur na powierzchnie modeli fizycznych. (Tekstury mogłyby także pokazywać pewien poziom trójwymiarowych zaburzeń powierzchni przy użyciu, na przykład bump-mappingu, czyli mapowania nierówności powierzchni). Dodatkowo można byłoby pokazać klientom jak wyglądałby obiekt w różnych warunkach oświetleniowych, w tym nocnych z włączonymi światłami, w świetle dnia w czasie różnych pozycji słońca, a także w różnych porach roku. Można byłoby również pokazać klientom wewnętrzną strukturę budynku włączając w to rury, kable elektryczne itp.

W drugim przykładzie pracownik taśmy montażowej mógłby zostać poprowadzony przez proces produkcyjny poprzez powiększone przestrzennie informacje. Hełmy do rzeczywistości wirtualnej AR (AR- powiększana rzeczywistość) były używane do tego celu w firmie Boeing. Używając opisywanych tu technik można uzyskać te

same efekty bez potrzeby korzystania z hełmów do rzeczywistości wirtualnej, za to przy użyciu projektorów skierowanych do wewnątrz rzucających obrazy lub tekst instrukcji na białą powierzchnię.

Łączona wizualizacja modeli

W czysto wirtualnych środowiskach (VE - virtual environments) przedstawia się graficzne modele rzeczywistych obiektów, zazwyczaj wraz z wirtualnymi obiektami generowanymi komputerowo. Z drugiej strony, podstawowe znaczenie Powiększanej Rzeczywistości (AR od augmented reality) to powiększanie obiektów fizycznych przy pomocy generowanych komputerowo obiektów wirtualnych. W przypadku środowisk wirtualnych (opartych na hełmach do rzeczywistości wirtualnej), w których ma się wrażenie przebywania w wyświetlanej rzeczywistości, użytkownik widzi fizyczne i wirtualne obiekty w takiej samej ograniczonej przestrzennej i czasowej rozdzielczości i wierności odtwarzania. Jedną z zalet opartej na projektorach Powiększanej Przestrzennej Rzeczywistości, tak jak i optycznie przeglądanej powiększanej rzeczywistości (korzystającej z hełmów do rzeczywistości wirtualnej), jest to, że zachowana jest wierność przestrzenna i czasowa obiektów fizycznych i tylko dodatkowe dane wyświetlane są w ograniczonej rozdzielczości. Dla porównania, w Powiększanej Rzeczywistości z przeglądaniem wideo przy użyciu hełmów do rzeczywistości wirtualnej obrazy obiektów wirtualnych są przedstawiane i nakładane na siebie przy użyciu obrazów wideo obiektów fizycznych, użytkownik widzi obiekty fizyczne i wirtualne w takiej samej ograniczonej rozdzielczości i wierności odtwarzania.

W tym przypadku przedstawione jest wykorzystanie Przestrzennej Powiększanej Rzeczywistości w konkretnym przypadku, w którym obiekt fizyczny powiększany przez wyświetlany obraz jest obiektem zainteresowania – jest fizycznym modelem, który odpowiada strukturze modelu graficznego. Najprostszym przykładem na to jest wizualizacja (np. budynku), która wykorzystuje i fizyczny i graficzny model tego budynku. Model fizyczny ma właściwą strukturę lub kształt, ale nie ma kolorów ani tekstur. Model graficzny podstawowo zawiera strukturę (identyczną do struktury modelu fizycznego), kolory, tekstury i jakiegokolwiek inne cechy powierzchni. Dodatkowo model graficzny może zawierać czysto wirtualne komponenty, które nie mają odpowiednika fizycznego. W efekcie użytkownik ogląda połączony model graficzny i fizyczny i korzysta z zalet obu rodzajów modeli.

Ustawienia projektorów

Kiedyś wielokrotne projektory wyświetlające i zachodzące na siebie obrazy używane były głównie do tworzenia obrazów panoramicznych. W takich prezentacjach użytkownik zazwyczaj stoi przed wyświetlanymi obrazami lub wewnątrz ogromnego pola widzenia środowiska prezentacji. Taki rodzaj prezentacji nazywa się systemem projekcyjnym “patrzenia ze środka na zewnątrz”. W większości przypadków ustawia się projektory tak, aby sąsiadujące projekcje lekko nachodziły na siebie. Obszar wyświetlanych powierzchni równocześnie oświetlany przez dwa lub trzy projektory zazwyczaj jest obszarem (dobrze zdefiniowanym) pojedynczym i przyległym. Poza tym, współrzędne zgodnego piksela projektora zmieniają się jednostajnie. Jest to podobne do jednostajnego uporządkowania zgodnych pikseli w parach kamer stereo.

Wyobraźmy sobie stół otoczony i oświetlony przez zbiór umieszczonych na suficie projektorów, przy którym użytkownicy mogą obejrzeć, a może nawet zmieniać obraz z któregośkolwiek miejsca naokoło stołu. Taki rodzaj projekcji nazywa się “patrzeniem z zewnątrz do wewnątrz”. Można sobie wyobrazić wykorzystywanie projektorów do wyświetlania na prostych powierzchniach takich jak kostka czy kula i tworzenie systemu wizualizacji typu “kryształowa kula”. Przy innym ustawieniu patrzy się na wklęsłą półkulistą misę oświetlaną by pokazać dwuwymiarowe lub trójwymiarowe obrazy zmieniające się pod wpływem pozycji obserwatora, które można obejść dookoła.

Istotną sprawą jest systemem wizualizacji, w którym można zmieniać właściwości dwuwymiarowe takie jak kolor czy tekstura, a może również właściwości trójwymiarowe znanych trójwymiarowych obiektów fizycznych, które same tworzą powierzchnię wyświetlania. Wcześniej pokazano jak wyświetlać perspektywicznie poprawne obrazy na gładkich, ale nie płaskich powierzchniach. W tym przypadku, dzięki obecności wklęsłych obiektów lub zbioru niepołączonych obiektów, obszary styku pomiędzy dwoma lub więcej projektorami niekoniecznie do siebie przylegają, a zgodne piksele nie zachowują jednostajnego uporządkowania. Jest to główna różnica, która stawia nowe wyzwania do uzyskania obrazów obiektów wirtualnych, które na miejscach styku nie będą miały widocznych łączeń.

Użyteczność

Z jednej strony, jeśli możliwe jest uzyskanie szczegółowego modelu fizycznego obiektu, oczywiście jest, że model ten będzie miał większą rozdzielczość, będzie

bardziej czuły, lepszy dla oczu – ogólnie lepszy niż jakikolwiek obraz, jaki może zaoferować nam rzeczywistość wirtualna w przypadku modelu statycznego. Z drugiej strony, wyraźnie widać że czysta rzeczywistość wirtualna ma tę zaletę, że można pokazać użytkownikowi “wszystko”, co statyczne czy poruszające się, bez potrzeby użycia modelu fizycznego. Metoda łączonej Przestrzennie Powiększanej Rzeczywistości ma do zaoferowania kilka zalet każdej z tych sytuacji -jeśli możliwe jest zdobycie modelu fizycznego. Taka kombinacja ma znaczący potencjał. Nawet proste demonstracje statyczne są bardzo imponujące, jasne, wyraźne i łatwe do oglądania. Ogólnie, zakładając, że chce się powiększyć obiekt fizyczny dwu lub trójwymiarowymi informacjami graficznymi, ma się kilka alternatyw. Na przykład, można użyć hełmów do wirtualnej rzeczywistości – wideo lub optycznych. Jedną z głównych zalet Przestrzennie Powiększanej Rzeczywistości uzyskanej przez projektory świetlne jest to, że użytkownik nie musi zakładać hełmu. (W [Bryson97] i [SARpaper]) opisane są różne zalety przestrzennie wyświetlaczy, dających wrażenie przebywania w wyświetlanej rzeczywistości, w stosunku do hełmów - dla wirtualnej rzeczywistości i powiększanej rzeczywistości.) W Powiększanej Rzeczywistości oglądanej przez wideo czy przy czystej rzeczywistości wirtualnej fizyczne i wirtualne obiekty są pokazywane przy ograniczonej rozdzielczości pikseli i prędkości klatek, to znaczy przy ograniczonych parametrach wyświetlania projektora, takich jak rozdzielczość bufora klatek, pole widzenia czy odległość od oświetlanego obiektu. Rozdzielczość przestrzenna i czasowa statycznej sceny jest niezależna od miejsca użytkownika i jego ruchów. Dzięki temu przy użyciu złożonego zestawu projektorów można będzie przedstawiać obraz, tekst czy nawet drobne elementy o dużo wyższej rozdzielczości.

Jeśli trzeba zmienić tylko właściwości powierzchniowe rzeczywistych obiektów, to kalibracja, dostosowywanie do indywidualnych potrzeb i wyświetlanie są o wiele łatwiejsze. W takim przypadku wyświetlanie jest niezależne od oglądającego, nie jest konieczna projekcja stereo i wiele osób może w tym samym czasie widzieć powiększenie. Nawet jeśli obiekty wirtualne nie są właściwościami powierzchni, ale znajdują się blisko rzeczywistych powierzchni, na których są wyświetlane, łatwo jest przystosować oczy do oglądania ich. Większość ustawień do Przestrzennie Powiększanej Rzeczywistości zawiera w sobie opisane powyżej zalety. Niestety, taki łączony sposób wizualizacji ma swoje wady. Metoda ta nie może być określona jako lepsza od czysto fizycznych lub czysto graficznych modeli, ale jest lepsza od

każdego w pewnych aspektach i w pewnych okolicznościach, ale w innych aspektach i okolicznościach jest gorsza. Na przykład, trzeba uzyskać (na przykład używając tych metod) graficzny odpowiednik modelu fizycznego. Poza tym, jeśli powierzchnie modelu fizycznego nie są czysto-białe, może się zdarzyć, że nie będzie można zamknąć pewnych części fizycznych. Jest to jedna z zalet Powiększonej Rzeczywistości oglądanej przez wideo. W podobnej sytuacji, w przypadku, gdy chce się wyświetlić wirtualne obiekty trójwymiarowe, które są stosunkowo daleko przed lub za powierzchnią modelu fizycznego, mogą być niewielkie sprzeczności pomiędzy modelami fizycznymi i wirtualnymi. Nie oceniano jeszcze tych sprzeczności, ale jeśli model jest wystarczająco przezroczysty fotometrycznie, tzn. wystarczająco biały i jeśli jest się w stanie kontrolować otaczające światło, to nie powinno być problemów.

Metody

Stworzono prostą manualną metodę modyfikacji właściwości powierzchni wielu ustawianych na blacie modeli fizycznych. Metoda ta polega na manualnym dostosowywaniu wyświetlania współrzędnych tekstur obrazu, aby je dopasować wizualnie do obiektów fizycznych. Mimo że metoda nie jest skomplikowana, znaczna większość osób, które widziały prezentacje wykonywane tą metodą stwierdza, że rezultaty są naprawdę imponujące.

Stworzono prostą automatyczną metodę modyfikacji właściwości powierzchni modelu fizycznego i dodawania trójwymiarowych modeli wirtualnych. Nie przedstawiano jeszcze całej metody, tylko jej elementy. Kompletna metoda powiększania modeli fizycznych wymaga określenia na początku, jakie są relacje między różnymi komponentami w środowisku i jakie są ich parametry. Komponenty te to kamery wideo, projektory świetlne, modele fizyczne i system śledzenia zmian pozycji obserwatora (head-tracking). Te elementy omówione są w części "kalibracja", poniżej. Następnie, użytkownik prawdopodobnie będzie musiał interaktywnie połączyć części modelu graficznego z odpowiadającymi im częściami modelu fizycznego, chyba że będzie chciał zmienić części modelu fizycznego. Działanie takie nazywa się "dostosowywaniem do potrzeb". Ostatecznie, w czasie dokonywania prezentacji używa się zaawansowanych technik wyświetlania, aby powiększyć model fizyczny wykorzystując perspektywicznie poprawne obiekty wirtualne dla użytkowników zaopatrzonych w urządzenie śledzące.

Kalibracja

Należy wykorzystać wielokrotne kamery wideo montowane na suficie i skierowane do wewnątrz do zapisania informacji geometrycznych o modelu fizycznym. Same kamery mogą zostać skalibrowane przy użyciu prostego wzoru kalibracji, takiego jak kostka z dokładnie naklejoną szachownicą na każdym widocznym boku. Po wyliczeniu wewnętrznych i zewnętrznych parametrów kamer można usunąć wzór kalibracyjny. Wyświetlając aktywne strukturalne światło z projektorów, skalibrowane pary kamer wideo mogą obliczyć głębię sceny. Funkcje pierwotne strukturalnego światła mogą być zwartym zbiorem binarnie kodowanych kropek wyświetlanych przez każdy projektor. Łącząc ze sobą wartości głębii wyliczone przez każdą z pary kamer można stworzyć trójwymiarowy obraz powierzchni całego modelu fizycznego. Ponieważ użyte zostaną wielokrotne projektory, konieczne jest stworzenie jednoznacznego i nieprzerwanego geometrycznego obrazu modelu fizycznego, tak aby można było wyświetlać zachodzące na siebie obrazy bez widocznych łączeń. Uzyskany model fizyczny może być przechowany jako model wielokątny – *model grafiki*. Podczas uzyskiwania głębii można też określić odpowiedniość dwuwymiarowych współrzędnych pikseli danego projektora w stosunku do trójwymiarowego miejsca oświetlonego przez te piksele. Jeśli znane są dwuwymiarowe piksele dla sześciu lub więcej punktów powierzchni trójwymiarowej, to można skalibrować projektor i znaleźć jego parametry wyświetlania. Jeśli podczas prezentacji zostanie użyty system śledzenia ruchów obserwatora, to można obliczyć przejście pomiędzy systemem współrzędnych urządzenia śledzącego a systemem współrzędnych dźwięku poprzez użycie odczytów czujnika urządzenia śledzącego w różnych miejscach i przyporządkowanie pozycji czujnika po obliczeniu ich przy pomocy triangulacji ze skalibrowanymi parami kamer.

Kiedy więcej niż jeden projektor oświetla część modelu fizycznego, należy się upewnić czy wyświetlane obrazy są geometrycznie wyjustowane. Jest to proces analogiczny do tworzenia mozaiki ze zdjęć zrobionych aparatem poprzez doczepianie tych zdjęć do siebie. Należy wyliczyć odpowiedniości pomiędzy pikselami wielokrotnych projektorów. Każda kamera obserwuje, który piksel różnych kamer oświetla ten sam punkt powierzchni na modelu fizycznym. Z tych obserwacji może zostać niebezpośrednio obliczony zestaw odpowiadających sobie pikseli projektorów.

Dostosowywanie do indywidualnych potrzeb

Jednym z ważnych zadań w uzyskiwaniu efektywnej powiększanej rzeczywistości jest stworzenie połączenia pomiędzy obiektami fizycznymi a funkcjami pierwotnymi grafiki, które powiększą te obiekty w czasie wyświetlania. Przykładami pierwotnych funkcji grafiki są linie, tekst, wielokąty z nałożonymi teksturami a nawet kompletne obiekty trójwymiarowe (wirtualne).

Na przykład, która tekstura powinna zostać użyta dla zewnętrznej powierzchni modelu budynku? Jakie kolory będą lepsze dla modelu fizycznego? Ogromnie ważny jest interfejs użytkownika w tworzeniu pierwotnych funkcji grafiki z różnymi kształtami, kolorami i teksturami. Podobny interfejs użytkownika wymagany jest przy ustawianiu i justowaniu pierwotnej funkcji grafiki tak, aby była poprawnie wyświetlana na żądanej części modelu fizycznego.

Wyświetlanie

Jeśli chce się zmienić tylko powierzchniowe wartości modelu fizycznego, takie jak kolor czy tekstura, nie trzeba całkowicie wyliczać trójwymiarowych modeli graficznych odpowiadających modelowi fizycznemu czy parametrów wyświetlania dla projektorów świetlnych. Na przykład, jeżeli użytkownik chce zmienić kolor na jednej ścianie budynku modelu architektonicznego, to wystarczy, że znajdzie zestaw pikseli z jednego lub więcej projektorów oświetlających tę ścianę budynku. Piksele te mogą zostać określone interaktywnie bez wyraźnych odpowiedników trójwymiarowych. Piksele mogą być kolorowane lub można na nie nałożyć wcześniej ukształtowane tekstury, aby zmienić wygląd ściany (ścian) budynku. Kiedy zmieniane są tylko wartości powierzchniowe (dla powierzchni rozproszonych), to wyświetlanie jest uznawane za niezależne od oglądającego i nie jest potrzebny system śledzenia ruchów obserwatora.

Kiedy na powierzchni obiektów fizycznych trzeba wyświetlić obiekty wirtualne korzystając z systemu śledzenia ruchów obserwatora, to można użyć dwuprzebiegowej metody wyświetlania opisanej w [Raskar98a]. Przy użyciu tej metody obiekty wirtualne objawiają się jako perspektywicznie poprawne nawet jeśli odpowiadających im obiektów fizycznych nie ma na planie. Podczas pierwszego przebiegu żądany obraz obiektu wirtualnego dla użytkownika jest obliczany i przechowywany jako mapa tekstur. Podczas drugiego przebiegu tekstura jest efektywnie wyświetlana z punktu widzenia użytkownika na wielokątny model

graficzny modelu fizycznego. Następnie ten wielokątny model graficzny, wraz z nałożoną na niego żadaną teksturą obrazu, wyświetlany jest z punktu widzenia projektora. Uzyskuje się to w czasie rzeczywistym, przy użyciu tekstur projekcyjnych. Zazwyczaj konieczny jest trzeci przebieg wyświetlania, aby upewnić się czy zachodzące na siebie wyświetlane z wielokrotnych projektorów obrazy są geometrycznie wyjustowane. Kiedy obrazy z wielokrotnych projektorów zachodzą na siebie, oświetlenie modelu fizycznego w tym miejscu może być o wiele większe niż oświetlenie w miejscach oświetlanych tylko przez jeden projektor. Dlatego więc, poza geometrycznym justowaniem pomiędzy wyświetlanymi obrazami, konieczna jest normalizacja intensywności. Problem generowania obrazów bez widocznych łączeń przy użyciu wielokrotnych projektorów był już wcześniej rozpatrywany w przypadku wyświetlaczy z dużym polem widzenia [Panoram][Trimensions][Raskar99a] czy w przypadku płaskich projekcji $m \times n$ [Czernuszenko97]. W takich przypadkach obszar, na którym obrazy zachodzą na siebie, jest zazwyczaj obszarem przylegającym na powierzchni wyświetlania, jak również w buforze klatek każdego projektora. Intensywność świecenia pikseli projektora jest regulowana w oparciu o techniki wtapiania – feathering (znane również pod nazwami intensity roll-off lub soft-edge), tak, aby zachodzące na siebie obrazy zmieszały się tworząc jednorodny obraz bez widocznych łączeń. W przypadku wielokrotnych projektorów skierowanych do wewnątrz, jeśli mamy pojedynczy wypukły obiekt rzeczywisty oświetlany przez wyświetlany obraz kwadratowy, to obszar zachodzenia na siebie brzegów dla jakichkolwiek dwóch projektorów także jest sąsiadujący. Jednakże zazwyczaj model fizyczny jest wykonany z niewypukłych obiektów lub zbiór niepołączonych obiektów, co sprawia, że obszar zachodzenia na siebie brzegów jest dzielony w buforze klatek każdego projektora. W pracy [Raskar99a] opisano i zademonstrowano technikę rozmywania obrazu w celu osiągnięcia wyjustowania geometrycznego i normalizacji intensywności świecenia dla stworzenia obrazów o niewidocznych łączeniach wyświetlanych z kilku projektorów na raz. Metoda rozmywania obrazu może być wykorzystywana nawet jeśli pojedynczy sąsiadujący obraz łączony nie jest kwadratowy lub powierzchnia oświetlana nie jest płaska. Jednakże gdy obraz łączony nie jest sąsiadujący, to trzeba określić w buforze klatek projektora zestaw pikseli, które oświetlają powierzchnię oświetlaną również przez co najmniej jeden dodatkowy projektor. Używając prostego algorytmu wzrastającego dla danego

obszaru w buforze klatek można określić te różne wyspy zachodzących na siebie obszarów. Technika rozmywania obrazu opisana w [Raskar99a] może zostać użyta dla każdej z takich wysp.

Kwestie rejestracyjne

W powiększanej rzeczywistości, podtrzymywanie wyobrażenia, że wirtualne i rzeczywiste obiekty współistnieją ze sobą, wymaga właściwego wyjustowania i rejestracji obiektów wirtualnych w stosunku do obiektów rzeczywistych. Tradycyjne metody powiększania rzeczywistości wykorzystują skupiony na obiekcie system współrzędnych do wyświetlania syntetycznych obiektów, a metody Przestrzenie Powiększanej Rzeczywistości wykorzystują system ustalonych współrzędnych świata do wyświetlania takich obiektów. Jednakże w obu przypadkach pojawiają się błędy statyczne i dynamiczne powodowane przez różne czynniki, takie jak opóźnienia systemu, zaburzenia optyczne i błąd obliczeniowy urządzenia śledzącego; są one trudne do poprawienia przez dostępne techniki. Wymagania trackingowe (warunków śledzenia) dla rejestracji w Przestrzenie Powiększanej Rzeczywistości są podobne do wymagań systemów SID-VR, ponieważ rzeczywiste i wirtualne obiekty leżą w tym samym systemie ustalonych współrzędnych świata. Dlatego też błędy kalibracji statycznej znacząco wpływają na rejestrację. Błędy te zawierają niepoprawne oceny przejść pomiędzy narzędziami wyświetlającymi, urządzeniem śledzącym i systemem współrzędnych świata. W Przestrzenie Powiększanej Rzeczywistości z hełmami wirtualnymi, błędy w ocenianych parametrach objawiają się tym, że obiekty wirtualne “pływają” w stosunku do rzeczywistych. W Przestrzenie Powiększanej Rzeczywistości takie błędy prowadzą do całkiem innych typów sztucznych wytworów.. Na przykład, kiedy dodatkowy obraz tylko modyfikuje właściwości powierzchni, to wyświetlany obraz jest niezależny od oglądającego i pozostaje zarejestrowany z statycznymi rzeczywistymi obiektami. Jeśli trójwymiarowe obiekty są wyświetlane na części modelu fizycznego, któremu mają odpowiadać, to tak jak opisano w [Cave][Raskar98c], błędy dynamiczne wywołują tylko ścinanie obiektów wirtualnych, zamiast “pływania”. W przypadku wyświetlania pływających obiektów trójwymiarowych, niewłaściwa rejestracja dynamiczna jest podobna do Przestrzenie Powiększanej Rzeczywistości z hełmami wirtualnymi. Efekt taki może pojawi się również, jeśli w interakcji z obiektami wirtualnymi przesunie się je w stosunku do obiektów rzeczywistych.

Wnioski

W tym rozdziale przedstawione zostało powiększanie modeli fizycznych poprzez otaczanie ich projektorami świetlnymi i wyświetlanie niezakłóconych na łączeniach obrazów na powierzchniach tych obiektów. Model ten wydaje się być szczególnie wydajny, gdy właściwości powierzchniowe danych obiektów rzeczywistych, takie jak kolor czy tekstura mają zostać zmodyfikowane. Wielu użytkowników na raz może stać wokół wyświetlanego obrazu i oglądać zmodyfikowaną powierzchnię bez projekcji stereoskopowej, okularów czy hełmów. Opisano też jak można użyć takiego ustawienia do powiększania fizycznych modeli przez wyświetlanie perspektywicznie poprawnych trójwymiarowych obiektów wirtualnych. Dzięki metodzie wizualizacji łączonej można powiększać modele o białych rozsianych powierzchniach mieszając obrazy z wielokrotnych projektorów. Jednakże obecnie technika ta wydaje się mieć pewne ograniczenia w zakresie wizualizacji i nie jest dobra do skomplikowanych interakcji z modelami wirtualnymi. Trzeba też pamiętać o problemie utożsamiania jeśli modele fizyczne o ostrych brzegach oświetlane są przez obrazy o niskiej rozdzielczości. Innym problemem mogą też być cienie.

Pomysły te na pewno będą ulepszone, jak i związane z nimi algorytmy, należy stworzyć lepsze programy dostosowywania do indywidualnych potrzeb i doprowadzić do wykorzystywania tej metody do wielu innych zastosowań. Na obecnym etapie, nakładano tylko różne kolory i tekstury na powierzchnie modeli fizycznych. Przyszłe plany to: stworzyć pełen system ze śledzeniem ruchów obserwatora do wyświetlania trójwymiarowych obiektów wirtualnych.

W końcowym efekcie metoda ta zapewni imponujące wizualizacje łączone, które zawierają w sobie zalety tak tradycyjnych modeli fizycznych jak i nowoczesnej grafiki komputerowej, ożywiając zwykłe modele fizyczne.

Przebieg eksperymentu

Zademonstrowano prosty model fizyczny oświetlany przez dwa projektory wideo i powiększany przez nakładanie na niego różnych kolorów i tekstur. Scena stworzona jest z białych obiektów drewnianych, pudełek kartonowych i z cegły. Tekstury i kolory są nakładane interaktywnie przy użyciu programu Adobe Photoshop.

W przypadku elewacji wystarczy wyznaczyć cztery punkty, aby uzyskać konieczne zagięcia tekstur. Kolory mogą być zmieniane interaktywnie (np. malowane sprayem).. Kiedy włączy się światła w pomieszczeniu, można zobaczyć jak prosty jest model fizyczny.. Takie ustawienie z systemem śledzenia ruchów obserwatora,

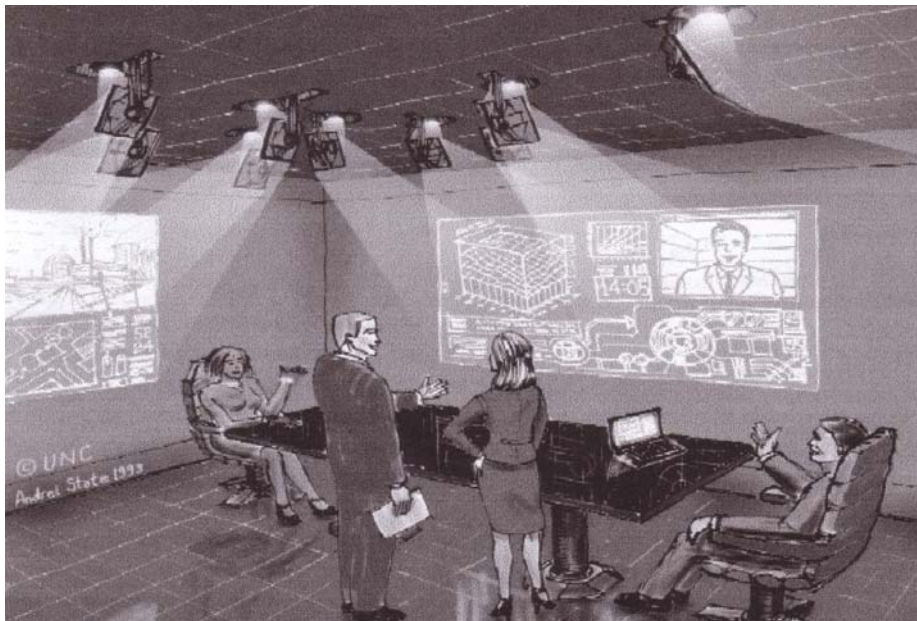
wielokrotnymi projektorami i modelami architektonicznymi będą bardziej przyjemne dla oglądającego niż obrazy czysto wirtualne lub wymagające użycia hełmów.

Wyświetlacze wieloprojektorowe

korzystające z rejestracji opartej o kamery

Opisane są tu nowe techniki wyświetlania obrazów panoramicznych przez system wieloprojektorowy, ze szczególnym naciskiem na utrzymanie rejestracji graficznej w wielu różnych konfiguracjach dotyczących obserwatora, projektorów świetlnych i powierzchni wyświetlania. Minimalizowane są błędy rejestracji geometrycznej poprzez użycie kamer, dzięki którym otrzymywane są powierzchnie wyświetlania i geometrie projektorów świetlnych. Korzystając z otrzymanej geometrii, używa się technik łączonych, dzięki którym otrzymuje się obrazy bez widocznych łączeń w wielu różnych konfiguracjach geometrycznych, z uwzględnieniem różnic w działaniu i tolerancji błędów. Takie techniki zajmują się problemami rejestracji geometrycznej, które większość systemów pomija przy konstrukcji (np. płaskie ściany i dokładnie wyjustowane projektory) lub przez ograniczenia (np. stojący w miejscu obserwator). Użytkownik, który nie jest ekspertem w tej dziedzinie, jest w stanie w ciągu 15 minut ustawić i skalibrować konfigurację prototypu tego systemu złożonego z dwóch kamer, dwóch projektorów i systemu śledzenia ruchu obserwatora.

5. Automatycznie rekonfigurujące się rzutniki szeroko obrazowe



Celem tego projektu jest rozwinięcie dynamicznego rekonfigurującego multi-projektora i całego systemu rzutowania. Będą to złożone z osadzonych na suficie projektory – każdy komputerowo sterowaną osłoną, panelem i ostrością (zoom) oraz z kamerami o pełnym skalowniu i obrocie o 360°. Taki system steruje:

1. automatyczną rekonfiguracją różnych obrazów, rozdzielczością i jasnością/kontrastem
2. wyrównywaniem ostrości obrazu bez pogorszenia jakości przy różnorodnym skonfigurowaniu

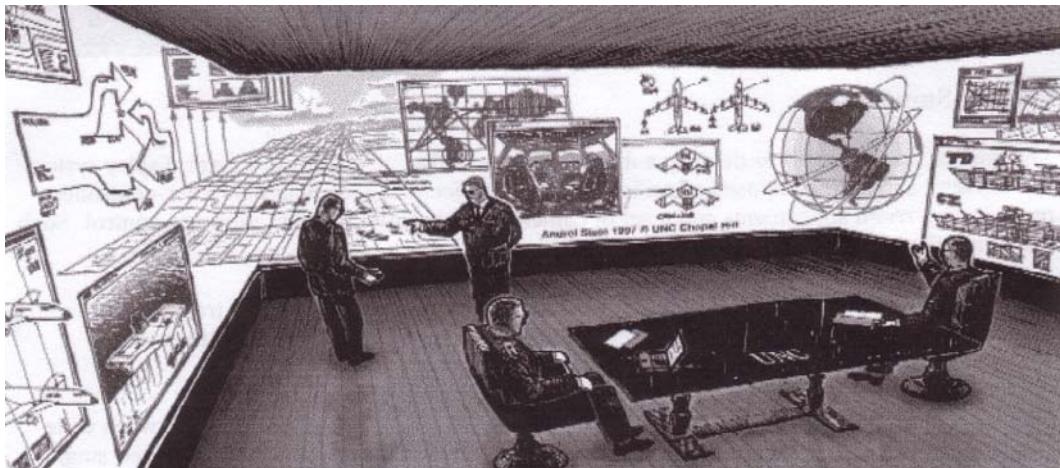
Proste złożenie z elementów kupionych w sklepie. Taki system pozwoliłby użytkownikom przejść łatwo od szeroko obrazowych, wielo ściennych prezentacji (prezentacji slajdów), do jednościennych prezentacji o większej rozdzielczości obrazu i jaśniejszym kontraście trójwymiarowych naukowych pokazów.

Projekty poboczne to:

Prototyp projektu z czterema i ośmioma rzutnikami, optyka graficzna komputerów, poprawa koloru i skalowanie. Praca grupy jest po części sponsorowana przez Ministerstwo Energii i Centrum NSF. Grupa jest częścią Grupy Biura Przyszłości.

Prototyp ośmio projektowy

Po stworzeniu prototypu cztero projektorowego rozpoczęto pracę nad nowym systemem z ośmioma projektorami . Rozwinięto system, który łączy projektory w jeden logiczny „projektor” i zapewnia taki rodzaj wykonania zadań, który wystarcza by użyć systemu jako interaktywny detoskop. Następny system będzie wykorzystywał środowisko Linux i usprawniał wyniki do poziomu utrzymania w czasie rzeczywistym grafiki w 3D.

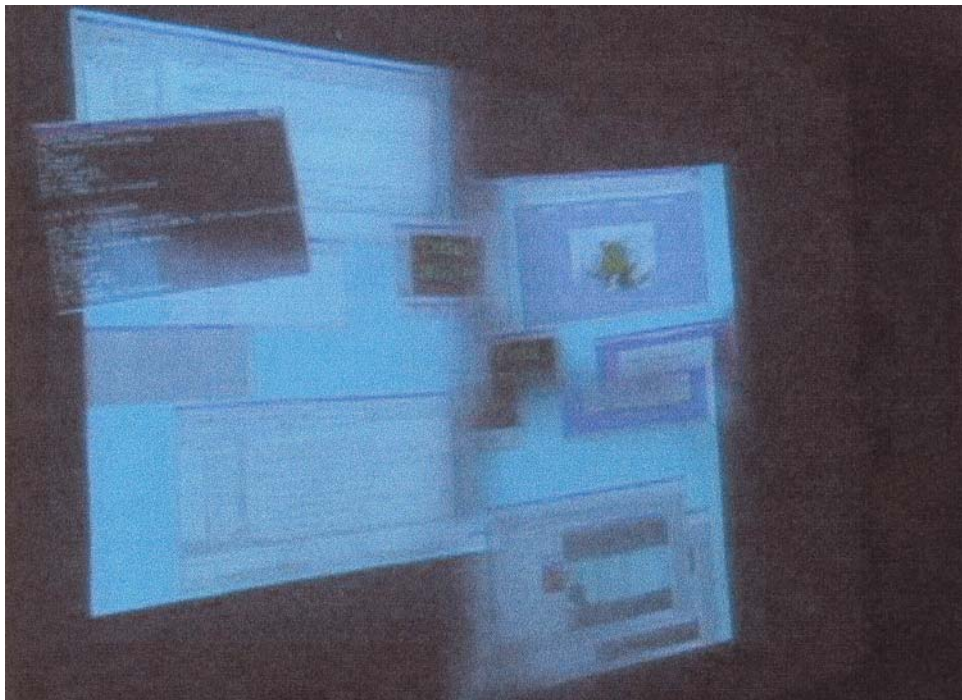
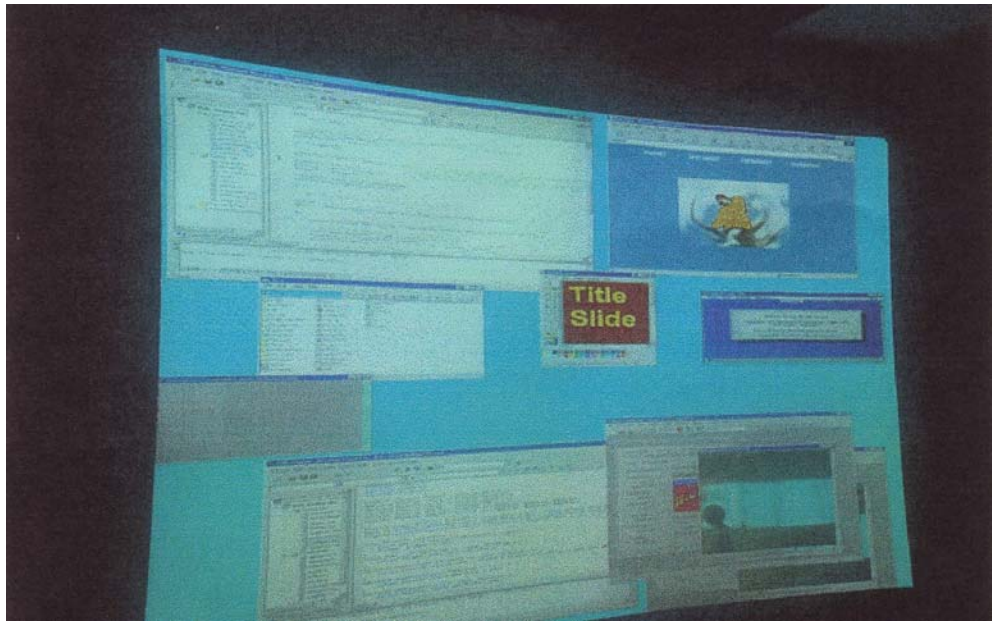


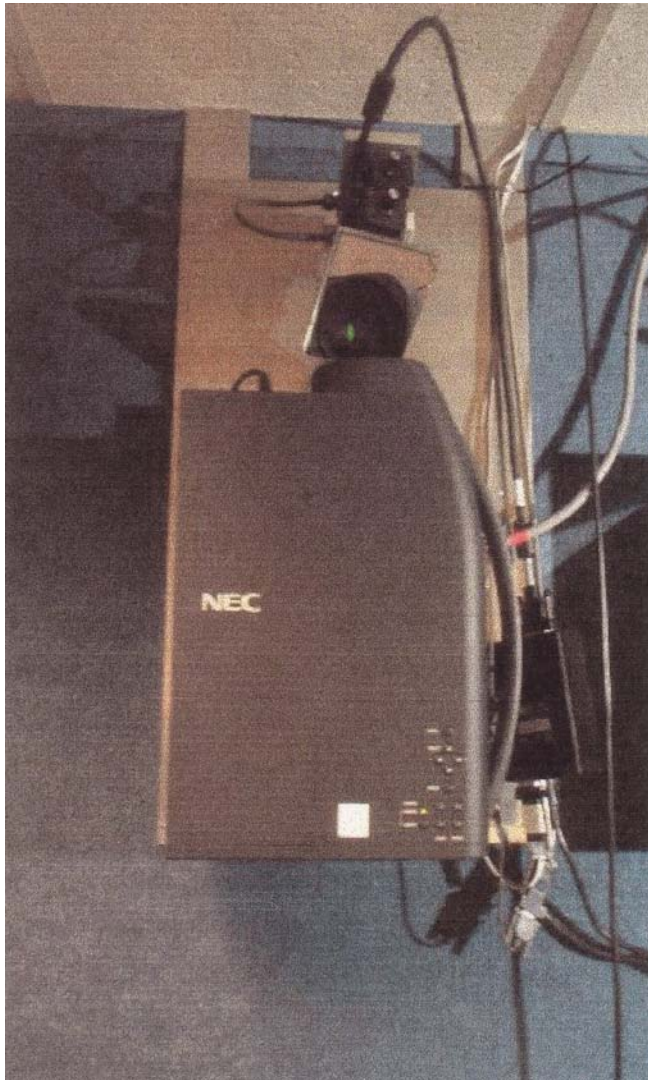
Prototyp PTU cztero projektorowy

Po uzyskaniu rozsądnego projektu za pomocą symulatora, zbudowano cztero rzutnikowy prototyp. Zamontowano rekonfigurowalny rzutnik o szerokim obrazie ze zrównanymi i geometrycznie połączonymi obrazami. Szybkość tego systemu nie

była jeszcze wystarczająca do interaktywnego zastosowania, jednak był to udany prototyp i pierwszy pracujący system.

- wyrównane liniowo i połączone dwa na dwa zestawione obrazy z projektora
- to samo ustawienie, ale PTU się przemieszczają. Obrazy poruszały się na ścianie, to zdjęcie podkreśla układ czterech projektorów.





Projektor/PTU System symulacyjny

Omawiany projektor i system symulacji jest zaprojektowany aby ułatwić projekt procesu dla systemu projekcji szeroko obrazowej. Z powodu unikalnej natury problemu często projektuje się nowe konfiguracje projektów, które dostarczają potrzebną funkcjonalność i łatwości przystosowania. Pełne fizyczne prototypy są zarówno drogie jak i czasochłonne w budowaniu. W celu przyspieszenia procesu projektowania i zapewnieniu szybkiego powstawania modelu pomysłów – zaprojektowano system symulacyjny, który pozwala na szybkie tworzenie nowych konfiguracji prototypów.

Wstawki interaktywne o wysokiej rozdzielczości

Celem tego projektu było stworzenie projekcyjnego zdalnego regulowania ostrości obrazu o wysokiej rozdzielczości i jednoczesnym utrzymaniu użytkownika w przekonaniu o rozumieniu widzianego obrazu (kontekstu). Czynnikiem ograniczającym przy tradycyjnych systemach projekcyjnych jest fakt, że gęstość pixeli jest równomiernie rozprowadzana na sieci. Ten układ

pracuje sprawnie zazwyczaj, ale nie zawsze. Na przykład zeskanowany obraz o wysokiej rozdzielczości 35mm zdjęcie. Cyfrowy obraz mógłby mieć rozdzielczość większą niż 5.000x 5.000 pikseli. Aby zobaczyć taki obraz użytkownik miałby dwie możliwości. Po pierwsze, użytkownik mógłby oglądać wersję tego zdjęcia o niższej rozdzielczości. To spowodowałoby zanik wielu szczegółów ze źródłowego obrazu. Co w przypadku zdjęć naukowych, mogą być szczegóły, które użytkownik powinien zobaczyć. Druga opcja to pokazywanie małego wycinka obrazu źródłowego, przez co można utracić ogólny widok. Użytkownik mógłby stracić orientację. Sposobem na rozwiązanie tego problemu jest stworzenie rzutnika o zmiennej gęstości pixeli poprzez użycie dwóch projektorów w zespole. Jeden projektor był używany aby zapewnić nisko-rozdzielczy ogólny obraz całość, a drugi projektor jest do pokazywania małego obszaru, ale o wysokiej rozdzielczości. Użytkownik powinien zaobserwować zarówno detale jak i obraz ogólny.

Projekt i ustawienie projektu (fizyczne ustawienie)

Projekt ten wykorzystywał dwa projektory. Jednemu z nich przypisano określoną pozycję i oś obrotu w pokoju. Projektor ten użyty jest do ogólnego obrotu o niskiej rozdzielczości. Drugi projektor był osadzony obok tarczy PTU. Urządzenie to miało lustro przymocowane do ruchomej osi. Lustro było użyte do kontroli pozycji i orientacji drugiego projektora. Był on użyty do wstawek o wysokiej rozdzielczości. Aby zobaczyć największą różnicę między fragmentami obrazu wysokiej i niskiej rozdzielczości, potrzebna była największa różnica w rozmiarze obrazów rzutowanych na ścianę. Najlepszym rozwiązaniem byłyby soczewki o wąskim polu widzenia(FOV). Niestety takie soczewki były i drogie i trudne do zdobycia. Umieszczono projektor o mniejszej rozdzielczości bliżej do ekranu projekcyjnego. Na dodatek przestawiono ustawienie zoom'u na rzutniku o wysokiej rozdzielczości do największego skupienia (FOV), a zoom projektora o niskiej rozdzielczości przestawiono odwrotnie – ustawiając soczewkę na najszersze skupienie(FOV). Widocznym rezultatem była zmiana gęstości pixeli. W idealnym systemie wąskie soczewki dałyby lepszy rezultat. Użyto również standardowej NTSC video kamery do skalowania i wychwytywania zmian ruchu laserowego wskaźnika (z tym były problemy). Użyto Windows NT z dwu kanałową karta video jako platformę rozprawdzającą. Ta dwu kanałowa karta pozwoliła użyć 2048 x 768 ramy buforowej. Podzielono ją na dwie połówki

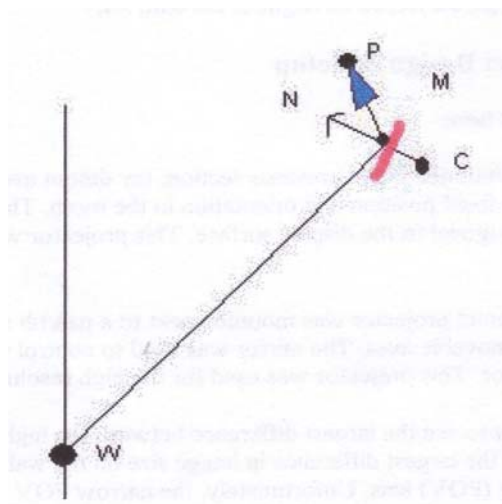
każda po 1024x768, zostały wysłane do projektora. Użyto dodatkowo systemu Matrox Meteor II video do wychwytywania obrazu video z kamery NTSC.

Interakcja

Planowano operowanie na dwóch interakcyjnych trybach. Pierwszy pozwolił użytkownikowi użyć zwykłej myszki do zaznaczenia obszaru zainteresowania. Druga metoda polegała na użyciu laserowego wskaźnika, który mógłby dostarczyć bardziej naturalnych środków wyboru danych obszarów. Jednak nie udało się zastosować laserowego wskaźnika. Ponadto zaopatrzonego użytkownika w wybór systemu kontroli łączenie z trybem włączania i wyłączenia obu projektorów. Ta cecha była przeznaczona bardziej do celów demonstracyjnych, jednak może być pożyteczna, gdy użytkownik chce wyłączyć np. projektor o niskiej rozdzielczości. Inna z funkcji kontroli pozwala użytkownikowi na powiększenie obszaru który go interesuje. Przy normalnych trybie – wstawki o wysokiej rozdzielczości są geometrycznie w tej samej skali co przy niskiej rozdzielczości. Jednak w tym trybie zoom –użytkownik może podwoić wielkość wstawki o wysokiej rozdzielczości, co pozwala na pokazanie nawet więcej szczegółów o wysokiej rozdzielczości. Łamie to jednak ciągłość geometryczną między obrazami o wysokiej i niskiej rozdzielczości.

Wyzwania techniczne

Napotkano na wiele technicznych wyzwań zanim można było zainstalować pracujący system. Pierwszy problem to zdecydowanie się na użycie właściwych : obrotownika i rumpla dla PTU. NA początku zarysowano system równań dla kątów przy każdej osi PTU.



P -to położenie centrum projektora a niebieskie wypełnienie przedstawia pojemność projekcji. P jest przymocowane a centrum pojemności projekcji jest przymocowane pod kątem. C – jest centrum osi rotacji PTU. C jest również przymocowane w miejscu, N –jest normalnym wektorem lustra (na różowo) i zależy od obrotu PTU względem C. M-jest położeniem przekroju centrum promienia objętości projekcji i lustra. W –jest punktem na ścianie. P i C są przymocowane. Znając pozycję (x, y, z) dla W chcemy znaleźć kąt obrotu względem C, który da nam takie N, że będzie to dwusieczna kąta PMW. Problem ten można podsumować, że należy znaleźć funkcję W (która jest położona x, y, z – określona), która obraca prawidłowy obrotnik i prowadnica $F(x, y, z)$ (obrotnik, rumpel). Najpierw rozwiązywano ręcznie, potem przy wykorzystaniu pakietu MALTAB. Po kłopotach z osiągnięciem odpowiednich wartości użyto skalowania światłem strukturalnym jako proces przygotowujący. Użyto procesu ze strukturalnym światłem, który łączy i koordynuje systemy działania kamer, projektor o niskiej i wysokiej rozdzielczości. Wprowadzono odpowiedniki na siatkę danych o wymiarach 30 na 30. Kamera o wyższej rozdzielczości albo soczewka o wąskiej ogniskowej zapewniłaby gęstszą siatkę danych. Kolejną trudnością jest uzyskanie realnej geometrycznej ciągłości między obrazami projektora. Są techniki, które rozwiązują ten problem dla ustawienia statecznych projektów, ale one używają procesu, który nie będzie funkcjonował w rzeczywistym czasie. Jeden z projektów w systemie jest ruchomy, nowe rozwiązanie musiało zostać zastosowane. Poza problemami powstającymi z poruszania centrum projekcji (CoP) jest więcej kwestii, które powstają z powodu ustawienia PTU. Ważniejszym od tego problemem jest tak zwany efekt „toczenia”. Jak lustro porusza się z jednej strony na drugą, obraz wydaje się zwiijać/toczyć poprzez ekran. To jest bezpośrednio związane z kątem obrotu lustra. Aby pokazać dokładnie połączone obrazy, system musi zrównoważyć ten efekt.

Zamontowanie

Zmontowanie zostało dokonane w Open GL. Najpierw wprowadzono kod światła i obliczono funkcje między projektorami i kamerami. Potem wbudowano tabele poglądu LUT zarówno dla ekranu w zamocowanym projektorze (dla pracy za pomocą myszki) i monitora kamery (dla pracy za pomocą laserowego wskaźnika). To pozwoliło zadać systemowi pytanie: „Myszka jest w punkcie X,

y, z. Jakie powinny być parametry obrotownika i osi obrotu aby zrównać w linii te dwa projektory?”

Z tymi wbudowanymi tabelami, utworzono pętlę powtarzającą się sekwencję w Open GL, która przeciągała by obrazy o niskiej rozdzielczości „wyprostowywała”, poprawiała ostrość na ścianie, a potem za pomocą alfa maski zaczerniała tę część, która ma być wypełniona wstawką o wysokiej rozdzielczości. Potem przyciągnęłaby obraz o wysokiej rozdzielczości, skorygowała ostrość na ekranie ściennym i użyła maski alfa do połączenia dwóch obrazów (o niskiej i wysokiej rozdzielczości). Przy użyciu myszki, trzeba wyświetlić współrzędne pikselowe i znaleźć odpowiedni trój wymiarowy punkt, który wybrała myszka.

Demonstracja

Pierwszej żywej demonstracji dokonano 12 maja 2000 roku.

Trudności i kompromisy



Pierwszym kompromisem było użycie metody światła strukturalnego do skalowania. W rezultacie powstał lepszy system. Jakikolwiek matematyczne derywacje polegały na dokładnych pomiarach i umiejscowieniu projektorów, luster i PTU. Proces skalowania brał pod uwagę dokładnie te wszystkie pozycje. Wielkim kompromisem było zrezygnowanie z laserowego wskaźnika. Przy jego użyciu można było wybierać dane punkty wyświetlanego obrazu. Miał być to laser z osłoną filtrową ograniczającą z przodu kamery (niestety był zbyt drogi). Następnie próbowano wyizolować czerwony strumień kolorowej kamery video, z nadzieją, że czerwony laser będzie tak jasny w świetle czerwonym, że będzie go można odróżnić od projektowanego obrazu. Jednak tak się nie stało. Włączono uchwyconą ramkę z czerwonego kanału światła kamery. Intensywny czerwony ukazuje się jako białe plamy a czerwony o niskiej intensywności ukazuje się jako czarny. Biała kropka po lewej (n obrazku) jest to laserowy wskaźnik, a reszta białych obszarów pochodzi od projektora. Zainstalowano

95% kodu potrzebnego do wykonania laserowego wychwytywania ruchu, I można by było ukończyć pracę, gdyby znalazł się odpowiedni filtr – ten element zostanie dodany do systemu w przyszłości. Nieprzewidzianym problemem było użycie funkcji `glUnproject` – nie dało to bardzo dokładnych rezultatów. Jak przewidziano (x,y,z) współrzędne, które zostały wrócone z funkcji były na promieniu określonym przez centrum projekcji i odpowiedni piksel. Jednak wartość głębi była przemieszczona o znaczna ilość. Ten błąd był rezultatem cyfrowego braku precyzji buforowej głębokości lub inwersji maczy. Celem było znalezienie punktu na ekranie projekcyjnym na który użytkownik wskazuje wskaźnikiem. Wiadome było, gdzie znajduje się centrum lokacji projektora i , że punkt (x,y,z) wrócił z `gl Unproject` i z Cop projektora, punkt ten określiły promień, który zawierał punkt szukany. Potem znaleziono przekrój tego promienia. Punkt przekroju był punktem, którego punkt wybiera się myszką. Były pewne fizyczne ograniczenia, które powstrzymały sukces projektu. Obok filtra, ten projekt mógł użyć wąsko ogniskowych soczewek dla projektora o wysokiej rozdzielczości. To pozwoliło by na większe zróżnicowanie rozdzielczości pikseli. Ten typ soczewek pozwoliłby też na umieszczenie tych dwóch projektorów znacznie bliżej siebie w pomieszczeniu, minimalizując zniekształcenia.

Kolejnym problemem była kamera NSTC. Po pierwsze kamera ta jest ograniczona do 640x480 rozdzielczości. Na dodatek, soczewki miały całkiem szerokie FOV. To sprawiało, że znaczne części obrazu uległy utracie. Metoda skalownia światłem strukturalnym wracała siatką danych o wymiarach 30x30. To było spowodowane dwoma czynnikami: rozdzielczość i FOV. Kamera o większej rozdzielczości i odpowiednie soczewki pozwoliły by na siatkę o odpowiednio większej gęstości obrazu. Rezultatem tego był gładszy ruch PTU, i mniejsza intensywność „przeskoków” obrazu w czasie przemieszczania widoku.

Interaktywne wstawki o wysokiej rozdzielczości / skalowanie

Algorytm rzeczywisty mógłby zostać użyty znacznie łatwiej, gdyby kąty obrotu mogły być odczytywane z tabelki pogładowej LUT. To pozwoliłoby na wybudowanie takiej tabelki nie martwiąc się o szybkość. Aby to wybudować zastosowano algorytm światła strukturalnego do przestrzeni poruszającego się projektora. Mapowanie odbywa się poprzez pośrednią przestrzeń: przestrzeń pikselową kamery. Ten pośredni krok jest tym czego potrzeba do zastosowania

laserowego wskaźnika. Po zrealizowaniu programu o strukturalnym świetle posiadano plik tekstowy, który określa mapowanie między tymi trzema powierzchniami: x i y nieruchomego projektora, u i v kamery i jednostki obrotnicy. Algorytm światła strukturalnego funkcjonuje przez budowanie LUT dla siatki danych w ustawionych projektorach.

Dla każdego obrazu:

Iluminowany obraz w (x_i, y_i)

Lokalizowanie obrazu w kamerze (u_i, v_i)

Iluminowanie centrum projekcji obrazu projektora ruchomego.

W pobliskim sąsiedztwie, osie obrotnic są ułożone liniowo. Miejscowo system ten zachowuje się liniowo. Za każdym powtórzeniem próbuje przesunąć obrotnicę coraz bliżej pożądanego kąta. Ten proces pozwala na określenie czy dana odległość jest wystarczająco bliska.

6. Teleimersja w medycynie

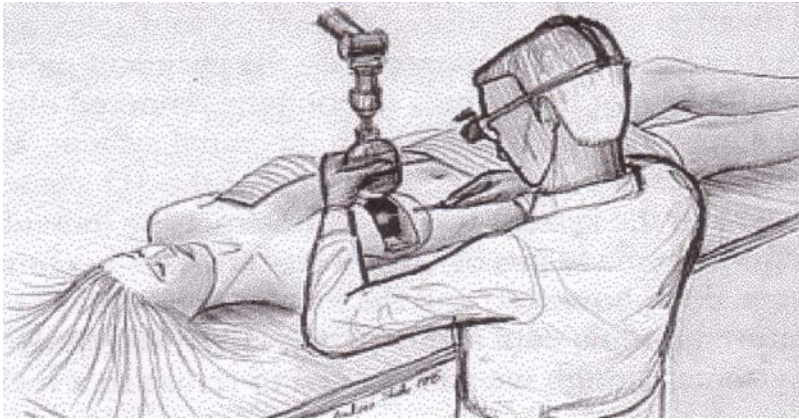


Zespół badawczy pracuje nad wynalezieniem i zastosowaniem systemu, który pozwoli lekarzowi bezpośrednio zajrzeć do ciała pacjenta, używając „rzeczywistości rozszerzonej” –AR. AR łączy grafikę komputerową z obrazami z realnego świata. Ten projekt używa obrazowania ultradźwiękiem (echografia), obrazu laparoskopowego, ekranu w postaci okularów (HMD) i komputera o wysokiej jakości grafiki do tworzenia żywych obrazów, które łączą sterowane komputerem obrazy z żywym przekazem video z ciała pacjenta. System AR wykorzystujący obrazowanie ultradźwiękowe lub laparoskopowe mógłby stać się silnym narzędziem

używanych do asystowania i kierowania poczynaniami lekarza podczas różnych typów medycznych procedur.

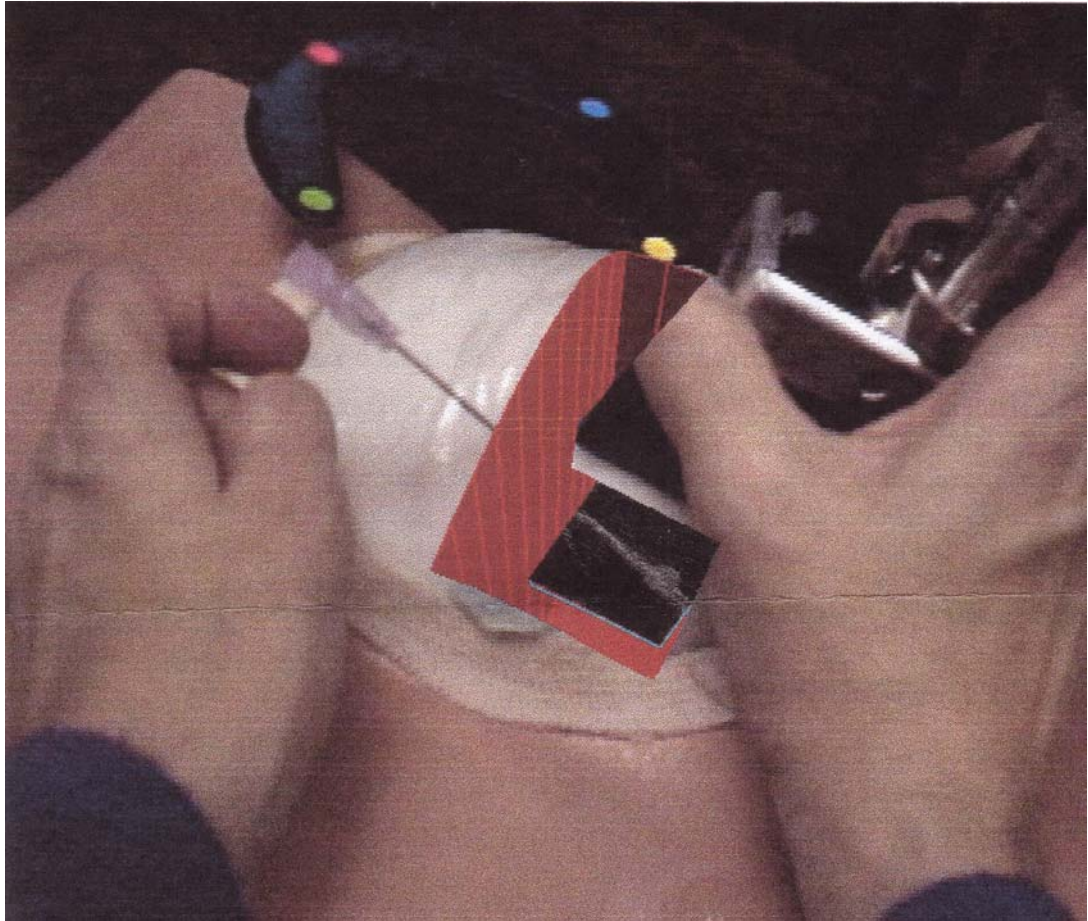
Ekran w postaci okularów(HMD)

Hełmy HMD zaopatrzone w odpowiednie właściwości do powiększania rzeczywistości nie są dostępne na rynku komercyjnym. Istniejące hełmy HMD modyfikuje się do potrzeb projektu.



Biopsja piersi przy pomocy ultradźwięków

Biopsja piersi przy pomocy wprowadzonych ultradźwięków byłaby używana w celach diagnostycznych, zwłaszcza zamiast interwencji chirurgicznej. Kierowanie ultradźwiękiem jest również stosowane przy precyzyjnej lokalizacji zmian chorobowych przed biopsją, jak również przy przecyszczaniu cyst. Kierowanie ultradźwiękiem w tego typu ingerencjach jest trudne do nauczenia i wykonania. Potrzeba dobrej koordynacji i umiejętności trójwymiarowej wizualizacji by prowadzić igłę biopsyjną do docelowego obszaru tkanki za pomocą monitorowania ultradźwiękiem. Technologia komputerowa rozszerzonej rzeczywistości może znacznie ułatwić zarówno nauczenie się jak i wykonywanie kierowanych ultradźwiękiem ingerencji w ciele pacjenta. Obecnie trwają prace nad systemem, który pomoże lekarzowi w wykonywaniu biopsji kierowanej ultradźwiękowo. Wyniki pierwszych badań na fantomach i jednym człowieku były obiecujące.



Wczesne systemy rozszerzonej rzeczywistości

Początkowe systemy wizualizacji ultradźwięków, zawierały zdolność pokazywania małej ilości pojedynczych ultradźwiękowych ujęć płodu nałożonych na zdjęcie podbrzusza ciężarnej pacjentki. System ten używał tradycyjnych technik by połączyć zakłócenia w ultradźwiękowym odbiorze z cyfrowym odtwarzaniem video z kamery zamontowanej na głowie. W początkowym systemie ostrość była słaba z powodu słabego systemu śledzenia ruchów, obraz był niewyraźny i nie było trójwymiarowego kształtu płodu. Pierwszym udogodnieniem systemu było zrekonstruowanie pojedynczych ujęć ultradźwiękowych na dźwięk. Ten algorytm wykorzystuje UNC – maszynę generującą wysoko zaawansowaną grafikę Pixel-Planes 5. Unowocześnienia były powolne jednak obraz był wciąż mało przekonujący. Stąd zaczęto przechodzić do systemu zawierającego nagromadzenie danych on-line i ich porządkowanie off-line. Te powstałe obrazy ustanowiły poziom jakości dla późniejszych systemów. W końcu gdy wspomagające technologie polepszyły się zaczęto wykorzystywać je w tym systemie. Powiększenie(rozszerzenie) zostało polepszone przez nową kamerę i metody

ultradźwiękowego skalowania i przez ulepszenie w używanych instrumentach do śledzenia ruchu.

System średniozaawansowany

Grupa badawcza zajęła się prototypem systemu powiększenia rzeczywistych obiektów w oparciu o SGI Onyx Infinite Reality (IR- nieskończona rzeczywistość) – stanowisko pracy zaopatrzone w system o wysokiej jakości grafice – wyposażone w sprzęt video Sirius Video-jednostkę wychytującą obrazy w czasie rzeczywistym. Ten system wykorzystywał zdolność do strukturyzowania obrazu przy dużej szybkości dostępnej w IR. Sirius wykorzystywał zarówno kamerę HMD jak i kamerę ultradźwiękową. Obraz z kamery był wyświetlany w tle; obraz kamery ultradźwiękowej jest przenoszony do pamięci strukturalnej i pokazywany na wielokątach emitowanych przez sondę ultradźwiękową w środku syntetycznego nacięcia w ciele skanowanego pacjenta. Obraz jaki ten system pokazał użytkownikowi przypomina oferowany przez wcześniejsze rozwiązania w oparciu o on-line i rekonstrukcję dźwięku. Ten system mógł podtrzymać szybkość ujęć 10 Hz dla zarówno zmieniania obrazu odczytu, ultradźwiękowego zbierania obrazu i również dostarcza ultradźwiękowy obraz o wysokiej rozdzielczości i dostraja obraz do szesnastu milionów pixeli. System ten użył technikę zaprojektowaną do poprawiania obrazu pochodzącego z magnetycznego systemu monitorowania noszonego na głowie poprzez śledzenie punktów odniesienia na obrazie video. Inne techniki – jak przewidywanie interpolowanie uprzednich odczytów, uporządkowanie danych w komputerze okazały się pożyteczne w dalszej redukcji rejestrowanych błędów. W przeciwieństwie do początkowego systemu, ujęcia ultradźwiękowe nie są rekonstruowane dźwiękowo, ale są pokazywane bezpośrednio jako półprzezroczyste wieloboki z ultradźwiękowymi obrazami video nanoszonych na nie. Wielkość takich bezpośrednio naniesionych ultradźwiękowych ujęć może dać coś na kształt zbioru danych dźwiękowych. Śledzenie ultradźwiękowym próbnikiem zostało wykonane z dużą precyzją, za pomocą urządzenia do mechanicznego śledzenia, które dostarczyło prawidłowej rejestracji między poszczególnymi ujęciami ultradźwiękowymi. Ten system został użyty do zademonstrowania możliwości użycia powiększonego obrazu by podkreślić wizualizację przy operacji laparoskopowej.

Obecny system

Ostatnio przeniesiono prototyp systemu rozszerzonej rzeczywistości z SGI Onyx Infinite Reality do departamentu SGI Reality Monster. Wykorzystano cyfrowe dane wejściowe-zdolności Reality Monster - przez równoczesne wychwycenie obrazów z kamer HMD i obrazów ultradźwiękowych i wprowadzono do pamięci strukturalnej. Ten nowy system używa tylko pojedynczego optoelektrycznego systemu śledzenia z większą i całkowitą precyzją niż w systemie wcześniejszym. Nowe zastosowanie tego systemu wymaga postępu w systemach śledzenia, obróbce obrazu, monitorów. Współpracując z Uniwersytetem w Utah zaprojektowano i zbudowano wyświetlacz video zamontowany na głowie i eksperymentowano z alternatywami hełmów HMD. Obecnie osiągalne systemy śledzenia ograniczają zdolność do dokładnego rejestrowania prawdziwych i syntetycznych obrazów, należy prowadzić badania nad poprawieniem, udoskonaleniem systemu śledzenia. Zaczynając rozwijać system do chirurgii laparoskopowej odkryto potrzebę rozwoju instrumentów i algorytmów dla szybszego rozwoju trójwymiarowej geometrii w ciele przez laparoskop poprawiając system śledzenia.

Poprzednie podprojekty dla rozszerzonej rzeczywistości

Podczas budowania systemu, pracowano szczegółowo nad technologią rozszerzonej rzeczywistości, szczególnie nad rejestracją prawdziwych i syntetycznych światów, przestrzennie i czasowo.

Wizualizacja wielowymiarowych danych naukowych: Zastosowanie idei ze sztuki oraz psychologii percepcyjnej

Jest to wielodyscyplinarny projekt badawczy w celu odkrycia nowych interakcyjnych narzędzi wizualizacyjnych oraz zrozumienie wielowymiarowych danych naukowych oraz zjawisk fizycznych jakie mierzą. Narzędzia zostaną rozwinięte i ocenione w bliskiej współpracy z naukowcami w trzech dziedzinach: neurobiolodzy studiujący rozwój nerwowy oraz choroby poprzez biologiczne odwzorowanie, naukowcy płynności obliczeniowej badający przepływ krwi przez arterie oraz geografowie wykorzystujący zdalne wyczuwanie w celu monitoringu środowiskowego oraz zarządzania złożami. Wyodrębnione są wspólne wzorce z zagadnień problemowych w tych wielodyscyplinarnych dziedzinach w celu rozwinięcia interakcyjnych metafor oraz technik wizualizacyjnych, które można uogólnić i które znajdują szerokie zastosowanie.

Projekt ten rozwija nowe metodologie oceny wizualizacji, dziedzinę, do której

dopiero zaczęto się zwracać. Porównuje skuteczność zastosowań wizualizacyjnych w kilku interaktywnych oraz statycznych środowiskach obliczeniowych i środowiskach wyświetlania włączając w to 4-ścienny Cave, wirtualne środowisko 40'x 40' z hełmami wirtualnymi, stanowiska pracy z trójwymiarowym systemem śledzenia ruchów obserwatora, graficzne stacje robocze, papier oraz trójwymiarowe wyjście szybkiego prototypowania. Wirtualne środowiska stwarzające wrażenie przebywania w nich będą poddane badaniu ponieważ wartość tych nietradycyjnych środowisk pracy nie została jeszcze ustalona i stanowią również okazję do zbadania fundamentalnie różniących się interakcyjnych metafor. Zostaną dokonane porównania dla obu przypadków: interaktywnych i statycznych, przy pomocy odpowiedniej technologii określonej dla każdego zastosowania.

Ten projekt łączy doświadczenie zaczerpnięte ze sztuki z percepcyjną psychologią i z nich czerpie pomysły. Przez kilka wieków artyści rozwinęli tradycję technik aby stworzyć wizualne reprezentacje dla konkretnych celów komunikacyjnych. Historia sztuki dostarcza języka aby zrozumieć tę wiedzę. Zaczerpnięto inspirację z malarstwa, rzeźby, szkicu oraz grafiki projektowej i użyję się tych technik do problemów naukowych.

Poza inspiracją, psychologia percepcyjna dostarcza również drugiego zestawu wiedzy, która dotyczy problemów naukowej wizualizacji. Ocena skuteczności metod wizualizacyjnych jest trudna ponieważ nie tylko trudno jest określić i skodyfikować cele ale również testy, które mogą je ocenić w znaczący sposób, są trudne do zaprojektowania i przeprowadzenia. Oceny te są pokrewne z ocenami funkcjonowania ludzkiego systemu percepcji. Psychologowie percepcyjni od wielu dekad przeprowadzają eksperymenty na zrozumienie percepcji i pomogą oni rozwinąć metodologię oraz ekspertyzę w celu oceny metod wizualizacyjnych w bliskiej współpracy z biologami, badaczami przepływu płynów, geografami, artystami oraz naukowcami komputerowymi. Podczas gdy poszczególne składniki tego projektu są same w sobie znaczące to jednak aspekty współpracy są najważniejsze. Ważne idee zaczerpnięte ze sztuki oraz dziedziny percepcji zasugerują niezwykle innowacyjne idee wizualizacyjne. Zastosowanie nowych technik wizualizacyjnych oraz współpraca naukowców w innych dziedzinach stworzy unikalną możliwość zatwierdzenia technik oraz upewnienia się, że stanowią one odpowiedź na potrzeby naukowe. Ponieważ techniki będą rozwijane z zastosowaniem w wielu dziedzinach, prawdopodobne jest, iż znajdą dalsze

zastosowanie w tych i innych dyscyplinach naukowych. Zebrany zespół zapewnia ekspertyzę w każdej dziedzinie i już przedstawił protokół przebiegu wspólnej pracy. Szerszy wpływ proponowanych badań leży nie tylko w dziedzinie technologii informacyjnej, gdzie nowe metody pomogą naukowcom w wielu dyscyplinach aby bardziej oddziaływali na dane i bardziej je rozumieli oraz zdobyli wgląd w zjawiska fizyczne jakie reprezentują, jak i w poszczególnych dziedzinach jakie będą badane. Studium przepływu krwi mogłoby prowadzić do lepszego zrozumienia oraz leczenia zwyrodnień naczyń sercowych. Zrozumienie wczesnego stadium rozwoju nerwowego mogłoby doprowadzić do powstania nowych terapii w przypadku wad wrodzonych, zaburzeń genetycznych i innych chorób. Postęp w dziedzinie zdalnego odczuwania mógłby dostarczyć bardziej skutecznego monitoringu zasobów oraz pozwolić na ogólną poprawę w globalnej jakości życia.

Elektroniczne Książki wieku Teleimmersji: Nowy paradygmat w nauczaniu procedur chirurgicznych

Teleimmersja stanie się nowym środkiem przekazu dla osób oddalonych od siebie aby pracowały i dzieliły się doświadczeniami w immersyjnym trójwymiarowym środowisku wirtualnym, tak jakby przebywały w tej samej fizycznej przestrzeni. Immersyjne książki elektroniczne, które w efekcie łączą „maszynę czasu” z trójwymiarowym hiper medium, dostarczą jeszcze jednego ważnego wymiaru który umożliwia rejestrowanie doświadczeń w których widz, znajdujący się w trójwymiarowej rekonstrukcji, może dosłownie przechodzić przez poszczególne sceny oraz poruszać się wstecz i naprzód w czasie. Podczas gdy istnieje wiele dziedzin zastosowania takich nowych technologii (np. projektowe i wirtualne tworzenie prototypów, utrzymywanie i naprawa, paleontologiczne i archeologiczne rekonstrukcje), w centrum zainteresowania znajdzie się ważne ze względów społecznych oraz stanowiące wyzwanie dla technologii, nauczanie chirurgicznego postępowania w przypadku skomplikowanych, potencjalnie śmiertelnych, obrażeń ciała.

Dzisiaj, szybkość wprowadzania innowacji chirurgicznych gwałtownie wzrosła lecz mechanizmy szkolenia cierpią z powodu nieelastycznej koordynacji czasowej, wydłużonych w czasie zobowiązań oraz ograniczonej treści. Tradycyjne instrukcje utrwalone na taśmie video przez długi czas były dostępne dla chirurgów w celu nauczania nowych procedur ale to podejście jest tylko marginalnie skuteczne ze

względu na utratę punktu widzenia, iż ta metoda ogranicza się do narracji, brakuje jej głębokości percepcji, interaktywności jak i informacji; innymi słowy, doświadczenie zdobywane poprzez oglądanie video nie odpowiada w stopniu dostatecznym przebywaniu w pomieszczeniu i obserwacji procedur. W tym projekcie jego inicjator rozwinie nowy wymiar nauczania chirurgicznych procedur, który pozwala na oglądanie oraz badanie (w czasie i przestrzeni) poprzednich chirurgicznych działań tak jakby widział tam był. Dodatkową korzyścią jest instrukcja z ust oryginalnego chirurga lub innego instruktora, jak również zintegrowane trójwymiarowe ilustracje, adnotacje oraz bieżące dane medyczne. Praktykanci powinni mieć możliwość swobodnego poruszania się w powiększonej do rozmiarów naturalnych, wiernej i trójwymiarowej graficznej rekonstrukcji oryginalnych, dziejących się w różnym czasie wydarzeń, zatrzymując, cofając lub idąc naprzód w czasie, aby zaspokoić swoją ciekawość lub pozbyć się wątpliwości.

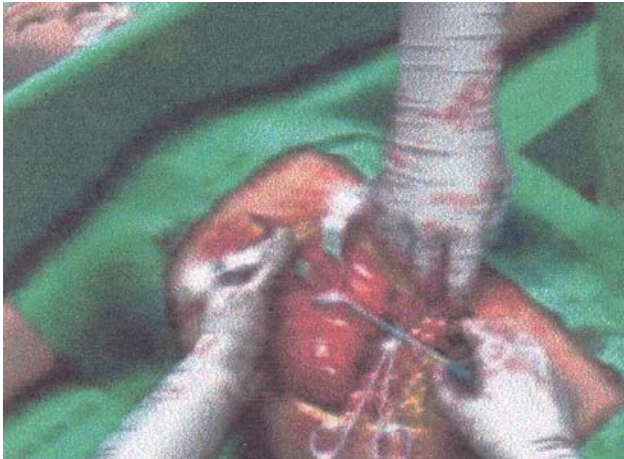
Aby stworzyć taką rzeczywistość, inicjator projektu oraz jego zespół gromadzą ekspertów z różnych dziedzin przy współpracy z którymi będzie mógł wykorzystać ich poprzednią pracę nad teleimersją, zarejestrowaniem trójwymiarowych scen dziejących się w różnym czasie, metaforami interakcyjnymi, technikami „kinematograficznymi” oraz narzędziami pozwalającymi dostosowywać program do własnych potrzeb.

Projekt Prof. Andy van Dam

Pięć najnowocześniejszych projektów informatycznych z Uniwersytetu Brown znalazło się pomiędzy zwycięzcami bardzo konkurencyjnej Narodowej Fundacji Naukowej przyznającej nagrody za badania technologiczne. Te niesamowite projekty, które – jeśli zostaną w pełni zrealizowane – pozwolą studentom medycyny brać udział w operacjach chirurgicznych, które mogły mieć miejsce miesiące temu, a osobom całkowicie sparaliżowanym poruszać ramieniem robota jeśli tylko o tym pomyślą – przyniosły naukowcom 5 milionów dolarów.

Można obserwować wykwalifikowanego chirurga przeprowadzającego skomplikowaną operację, stojąc tuż obok niego. Aby się lepiej przyjrzeć, zaglądamy chirurgowi przez prawe ramię. Potem przechodzimy na drugą stronę stołu, skąd oglądamy operację z innej perspektywy. Mimo że chirurg ma przed sobą trudne zadanie, cierpliwie wyjaśnia każdy problem, jaki napotyka, każdą technikę, jakiej

używa. Jeśli chcemy jeszcze raz obejrzeć jakąś metodę, chirurg chętnie się zgadza – bez ryzyka dla pacjenta.



Rys.1 W projekcie ITR prowadzonym przez Brown i UNC technologia “Teleimersji” będzie dostosowana, aby szkolić chirurgów. Studenci będą mogli naturalnie poruszać się w obrębie graficznej rekonstrukcji operacji, która cechuje się naturalnymi rozmiarami, wysoką rozdzielczością i trójwymiarowością.

Jest to świat rzeczywistości wirtualnej. Operacja ta mogła być przeprowadzona wiele miesięcy temu, w odległości tysięcy mil, ale jeśli zrealizowane zostanie marzenie profesora informatyki Andries van Dama i jego współpracowników, to studenci medycyny i inni będą mogli być świadkami przeprowadzonej wcześniej operacji, tak jakby działa się w danym momencie. Narodowa Fundacja Naukowa nagrodziła niedawno van Dama i jego zespół 1 750 000 dolarów, aby nadal realizowali swoją wizję. Proponowany projekt łączy dwie długoterminowe pasje van

Dama – interaktywne książki elektroniczne (z czułą zawartością, która zmienia się dynamicznie) i trójwymiarową grafikę immersyjną. Projekt, zatytułowany “Książki Elektroniczne dla Ery Teleimersji”, skupia ekspertów z kilku różnych dziedzin, między innymi medycyny, i bazuje na poprzednich pracach informatyków z zespołu. Zespół wraz z van Dajem prowadzi Gregory Welch, informatyk z Uniwersytetu Północnej Karoliny w Chapel Hill. Od lat studenci medycyny oglądali procedury chirurgiczne na wideo, co według van Dama “chirurdzy, w najlepszym wypadku, uważają za marginalnie skuteczne”. Film wideo, mówi van Dam, ma określony punkt widzenia i brakuje mu trójwymiarowości. “Podsumowując to krótko, mówi, oglądanie filmu wideo nie jest nawet w połowie zbliżone do oglądania tej procedury na żywo.” Wizja van Dama przypomina trójwymiarowy wehikuł czasu pozwalający chirurgom i studentom chodzić – wirtualnie – naokoło pełnowymiarowej rekonstrukcji właściwego wydarzenia w wysokiej rozdzielczości i poruszać się w czasie – do przodu i w tył. Istniejące narzędzia do “przechwytywania scen” w różnym czasie nie są dokończzone, mówi van Dam. Narzędzia autorskie i techniki kinematyczne dla tego nowego środka są właściwie nieznanne.

Van Dam i jego zespół skupiają się na chirurgii urazowej, choć technologia, którą planują rozwijać ma wiele innych możliwych zastosowań. Uczenie chirurgii urazowej, mówią, jest szczególnie trudne, ponieważ niektóre sytuacje zdarzają się tak rzadko, że wielu chirurgów spotyka się z nimi tylko kilka razy w całej swojej karierze. Najbardziej skutecznym sposobem nauczania się właściwej reakcji na poszczególne urazy jest obserwowanie operacji osobiście, a potem branie aktywnego udziału w całej procedurze, ale takie możliwości rzadko się zdarzają. Immersyjne “książki elektroniczne” sprawiłyby, że wiele osób mogłoby w takich działaniach brać udział. Van Dam przyznaje, że żadna z tych rzeczy nie nadejdzie szybko. Kiedy wygaśnie dotacja z NSF w 2004 roku “będziemy w Kitty Hawk, mając nadzieję wzbić się na pół minuty na wysokość 20 stóp”. Jednak przewiduje, że do 2010 roku ich lot znacznie się wydłuży. Lot ten, jak napisał van Dam w swoim uzasadnieniu dla NSF, może zawierać następujący scenariusz: Dr Smith, chirurg i wiodący ekspert w tej dziedzinie, wykorzystuje innowacyjną technikę do zreperowania serca młodej dziewczyny. Podczas operacji pani doktor otoczona jest przez morze kamer, które rejestrują model sali operacyjnej i wszystkich osób w niej przebywających. Zapisywany jest każdy ruch i każdy szczegół. Po operacji, będąc już w swoim biurze, dr Smith zakłada swoje trójwymiarowe okulary, co czyni

również kilkoro jej kolegów z całego kraju. Wszyscy są teraz “wewnątrz” filmu wideo tej procedury i mogą swobodnie omawiać jej aspekty nie narażając przy tym zdrowia pacjenta. Następnie, specjalista medialny tworzy zaopatrzoną w adnotacje trójwymiarową wersję “książki elektronicznej”, która natychmiast będzie osiągalna – przez Internet – dla studentów i chirurgów na całym świecie. Teraz i oni mogą wejść do środka sali operacyjnej dr Smith, usłyszeć jak omawia operację, przyglądać się jej metodom z jakiegokolwiek punktu widzenia, a nawet przesuwać się w czasie, żeby jeszcze raz obejrzyć wcześniejszy krok. Van Dam chce, aby w końcowym efekcie użytkownicy mogli czuć i używać narzędzi chirurgicznych, ale symulacja zmysłu dotyku jest, jak mówi, szczególnym wyzwaniem.

Projekt van Dama jest częścią współpracy między wieloma instytucjami, której fundatorem jest NSF. Uniwersytet Północnej Karoliny prowadzi obecnie prace nad uzupełniającą częścią tego projektu, która znana jest jako “uzyskiwanie i rekonstrukcja sceny w czasie rzeczywistym”.

To, o czym mowa to pomysł telefonu XXI wieku, telefonu, dzięki któremu ludzie czują, że są ze sobą w jednym pomieszczeniu. “Chcemy przenieść pomysł teleobecności trzy poziomy realizmu wyżej”- mówi van Dam.

Tworzenie modeli i wizualizacja odległości pomiędzy kośćmi w stawach

Wprowadzenie

Przedstawiona w tym rozdziale będzie metoda wizualizacji odległości pomiędzy kośćmi w stawach. Wizualizacja odległości pomiędzy kośćmi w sposób przystępny charakteryzuje struktury trójwymiarowe i relacje przestrzenne w złożonych stawach. Ta metoda wykorzystuje zarówno ukryte jak i parametryczne sposoby przedstawiania kości. Te dwa rodzaje obrazów mają dodatkowe atuty w różnego rodzaju obliczeniach. Obraz ukryty - pole odległości - jest to pole skalarne przedstawione na zwykłej siatce. Określa ono daną odległość od punktu na siatce do najbliższego punktu na kości. Obliczamy pola odległości z obrazów parametrycznych, które są gładkimi wielopłaszczyznowymi powierzchniowymi modelami kości, które można obracać lokalnie. Modele te są składane ze zdjęć tomograficznych kości .

Obliczamy odległości pomiędzy kośćmi w prosty i wydajny sposób przy użyciu pól odległości. Tworzymy wizualne obrazy obliczonych odległości pomiędzy kośćmi używając map kolorystycznych i konturowania izometrycznego.

Animacje tych przestawień były szczególnie pomocne w próbach zrozumienia zmian w relacjach pomiędzy kośćmi. Użyto tej techniki oddzielnie w trzech połączonych strukturach stawów dłoni – stawu promieniowo-nadgarstkowego, kości nadgarstkowych i stawu nadgarstkowo-śródręcznego.

Mimo że badania struktury i kinematyki stawów były dokonywane głównie *in vitro* przy użyciu takich technik jak dzielenie tkanki miękkiej, elektromagnetyczne śledzenie ruchu i tomografia zwłok, to ostatnio wprowadzono badania wykorzystujące techniki trójwymiarowe. Nie wykorzystywano wcześniej pól odległości do przedstawiania kości, choć pola odległości były już wykorzystywane do innych celów w robotyce i w grafice komputerowej. Te dwie techniki wizualizacji danych skalarnych, których użyto – mapy kolorystyczne i konturowanie izometryczne – są dobrze znanymi naukowymi metodami wizualizacyjnymi.

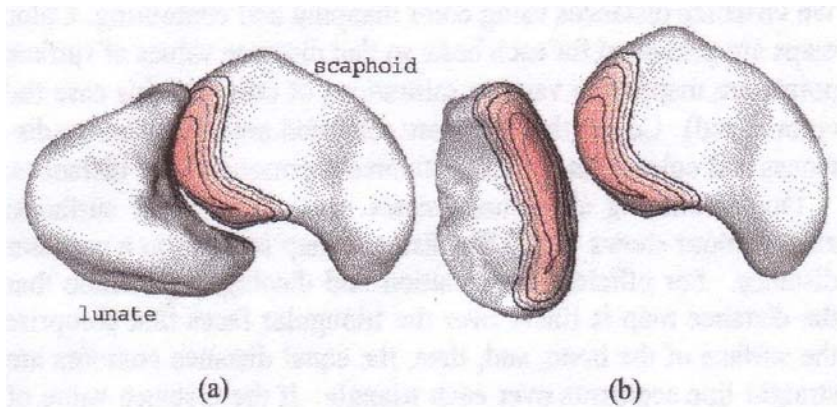
Materialy i metody

Pobieramy trójwymiarowe obrazy struktur i kinematyki stawów używając technologii tomograficznej. Kontury powierzchni kości są ręcznie składane z dwuwymiarowych kawałków, które składają się na tomograficzne obrazy stawów.

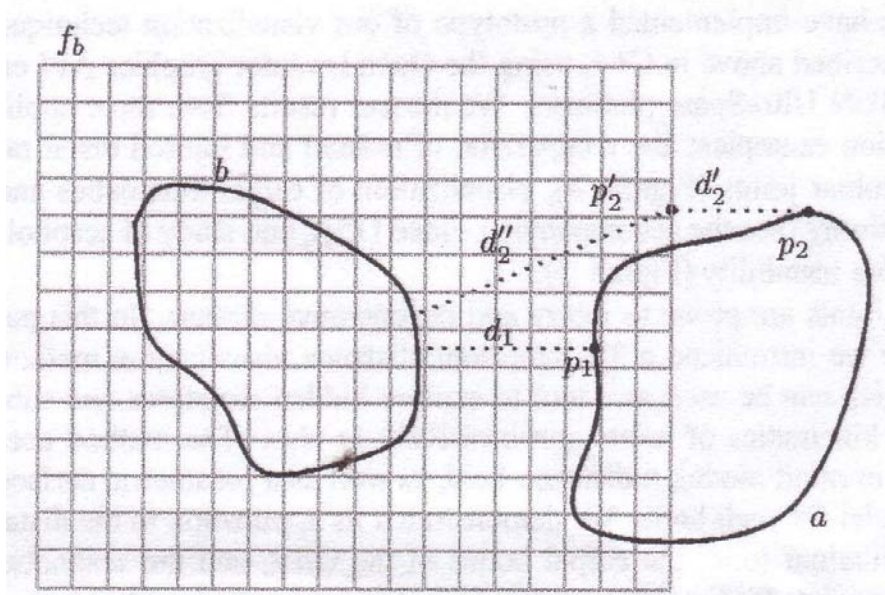
Powierzchnia kości rekonstruowana jest przez dopasowywanie wielopłaszczyznowej powierzchni do tych konturów, co daje gładką, sparametryzowaną w każdym punkcie, C^2 ciągłą powierzchnię kości. Składająca się z wielu zachodzących na siebie warstw struktura wielopłaszczyznowego obrazu powierzchni, która jest oparta na geometrii różnicowej, ma kilka zalet, między innymi elastyczność w dostosowywaniu kształtów bez kosztownych ograniczeń ruchu oraz płynne przejścia i jednorodność elementów.

Obliczamy pola odległości dla każdej kości używając zrekonstruowanych wielopłaszczyznowych modeli kości. Pole odległości to pole skalarne, które określa daną minimalną odległość od danego punktu do powierzchni kości. Pole skalarne jest przechowywane nad prostopadłościannym otaczającym kość, jako zestaw przetworzonych na postać cyfrową danych. Znak używany jest po to, by odróżnić wnętrze kości od jej powierzchni. Zachowywanie pól odległości jako obrazów ukrytych ma ważne zalety dla takich działań geometrycznych jak obliczenia odległości, wykrywanie kolizji, generowanie poziomu powierzchni, wykonywanie

działań logicznych oraz testy wewnątrz/zewnątrz. Przetwarzamy pola odległości w zwykłych siatkach trójwymiarowych, by przyspieszyć przeglądanie.



Rys. 1 Normalny staw nadgarstkowo-śródręczny. Na kości nałożona jest mapa kolorów i kontury. Natężenie czerwieni (ciemniejszy obszar na czarno-biało), na powierzchniach kości odnosi się do odległości do najbliższego punktu na przeciwległej kości. Im obszar jest bardziej czerwony, tym mniejsza odległość. Maksymalna zobrazowana odległość to 5 mm. Linie zarysu rozrysowane są w przerwach 0,5 mm. (a) Kości w ich właściwym wyglądzie anatomicznym. (b) Kości przesunięte, by wyraźniej pokazać powierzchnie stawowe.



Rys. 2 Dwuwymiarowa ilustracja otrzymywania odległości od punktów p_1 i p_2 do kości b . f_b to odległość prostopadłościanu dla kości b . Znalezione są najmniejsze wartości odległości do kości b na skrzyżowaniach siatek. Korzystamy z trójkątnej interpolacji przestrzennej dla wartości wewnątrz siatki. Ponieważ p jest wewnątrz prostopadłościanu, to odległość od p_1 do b równa się $f_b(p_1) = d_1$. Dla p_2 najpierw odnajdujemy odległość do najbliższego punktu p'_2 na

prostopadłościennie odległości, a potem odległość pomiędzy p_2 i b jest obliczana w przybliżeniu na podstawie wzoru:

$$d'_2 + f_b(p'_2) = d''_2 + d''_2.$$

Odnajdowanie odległości pomiędzy kośćmi.

Po wygenerowaniu pól odległości obliczamy odległość od dowolnego punktu p do powierzchni b , jak niżej. Każda powierzchnia kości otoczona jest przez prostopadłościennie odległości, nad którym przetwarzane jest na postać cyfrową pole odległości. Punkt p może się znajdować w lub poza prostopadłościennie odległości f_b . Punkty poza prostopadłościennie odległości są daleko od kości, więc nie muszą być dokładne, ale muszą być gładkie. Upewniamy się czy punkty w obszarze, który nas interesuje (tzn. powierzchnie stawowe) znajdują się w obrębie prostopadłościennego odległości. Rys. 2 ilustruje tę procedurę. Rozważamy dwa przypadki, aby znaleźć odległość:

p jest w f_b : odnajdujemy f_b dla p ,

p jest **poza** f_b : najpierw odnajdujemy odległość do najbliższego punktu p' na granicy f_b , a następnie dodajemy wartość odległości, którą otrzymujemy po odnalezieniu f_b dla p' . Wykorzystujemy trójkątną interpolację przestrzenną opartą na sekwencji krzywych Beziera do interpolowania przykładowych wartości pomiędzy miejscami siatki. Trzeba zauważyć, że przy użyciu zdefiniowanej procedury znajdowania odległości punkt-kość, możemy znaleźć odległość od każdego wierzchołka na modelu powierzchniowym do każdej innej sąsiadującej kości. Te wartości odległości są na bieżąco przeliczane dla każdej klatki w animacji opartej na kinematyce stawów.

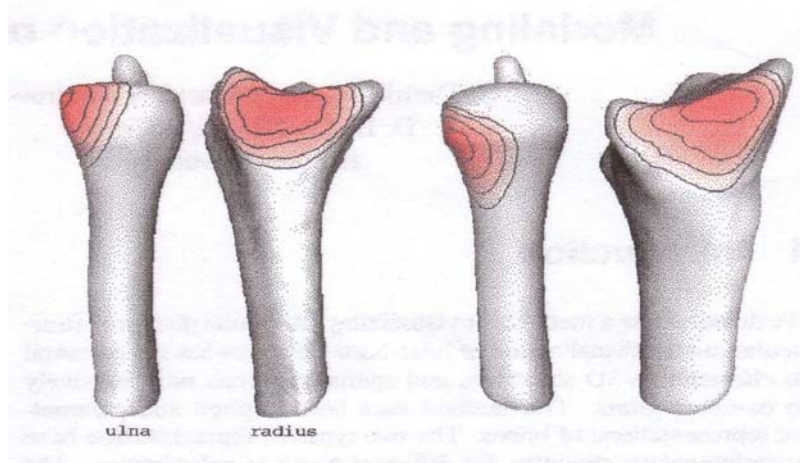
Wizualizacja

Obrazowane są odległości przy użyciu map kolorów i konturowania. Mapy kolorów generowane są dla każdej kości tak, aby wartości odległości punktów powierzchniowych odpowiadały różnemu natężeniu kolorów (w tym przypadku jest to kolor czerwony). Kolory ciemniejsze oznaczają mniejsze odległości, a kolory jaśniejsze – większe odległości.

Algorytm do konturowania tworzy linie izometryczne na powierzchniach kości; każdy kontur pokazuje, gdzie mapa odległości jest równa stałej odległości. Aby obliczenia i wyświetlanie przebiegały sprawnie, przyjmujemy, że mapa odległości jest liniowa ponad trójkątnymi płaszczyznami, które składają się na powierzchnię kości, a więc równe kontury odległości są prostymi segmentami linii

nad każdym trójkątem. Jeśli wartość odległości dla konturu d jest w zasięgu wartości odległości wierzchołków, to nad tym trójkątem generowany jest segment linii konturowej.

Konturowanie jest bardzo przydatne do grupowania odległości i w tym sensie uzupełnia technikę map kolorów. Odległości poza wartością progową nie są ani kolorowane ani konturowane. Są przedstawiane jako białe powierzchnie. Rys. 1 i 3 przedstawiają te dwie techniki.



Rys. 3: (a) Odwrócony do zewnątrz normalny staw promieniowo-nadgarstkowy (prawe przedramię). (b) Odwrócony do zewnątrz staw promieniowo-nadgarstkowy z urazem (lewe przedramię tego samego pacjenta). Lewe przedramię jest pokazane w odbiciu lustrzanym dla łatwiejszego porównania. Wyraźne przesunięcie pozycji w obszarze stawowym. Natężenie czerwieni na powierzchniach kości przedstawia odległość do najbliższego punktu na przeciwległej kości. Bardziej czerwone obszary są sobie bliższe. Kości są przesunięte, aby wyraźniej pokazać powierzchnie stawów. Maksymalna zobrazowana odległość wynosi 5 mm. Linie konturowe narysowane są w odstępach 1 mm.

Rezultaty i podsumowanie.

Zaimplementowano prototyp opisaną powyżej technikę wizualizacyjną z użyciem narzędzia Open Inventor w C++ z API na platformie SUN UltraSparc. Zaprezentowano rezultaty przykładów trzech zastosowań: porównania normalnego stawu i stawu po urazie (Rys. 3), wizualizacji kinematyki i anatomii nadgarstka i badania niestabilności stawu nadgarstkowo-śródręcznego (Rys. 1)

Stawy są podatne na urazy i choroby zwyrodnieniowe. W tym rozdziale przedstawiona została technika trójwymiarowej wizualizacji odległości pomiędzy

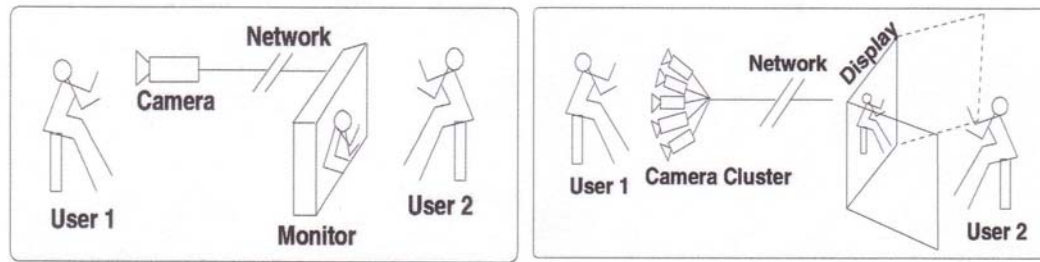
kości, która może być użyta jako narzędzie do badania ukrytych struktur i subtelnej kinematyki stawów bezinwazyjnie *in vivo*. Metoda ta wykorzystuje model ukryty, pole odległości jak również parametryczny model powierzchniowy dla każdej kości. Zademonstrowano użycie tej metody w stawach promieniowo-nadgarstkowych, kości nadgarstkowych i stawach nadgarstkowo-śródręcznych. Pierwsze rezultaty wskazują na to, że ta metoda może być bardzo użyteczna w badaniach normalnej anatomii i kinematyki skomplikowanych stawów, takich jak nadgarstek. Prototyp tej wizualizacji będzie wystawiony na nadchodzącym Spotkaniu Amerykańskiego Zrzeszenia Chirurgii Dłoni. Ta technika może mieć również zastosowanie w badaniach chorób nadgarstka, takich jak artretyzm reumatyczny, wewnątrz nadgarstkowe naderwanie/osłabienie więzadeł oraz syndrom kanału nadgarstkowego.

Obecna praca prezentuje nowy algorytm do prowadzenia telekonferencji immersyjnych, który rozwiązuje problem dopasowywania i łączenia wielu obrazów w jeden dla stworzenia jednego widoku panoramicznego bez widocznych łączeń. W paradygmacie telekonferencji immersyjnych jedna *klatka* telekonferencji to panorama, która jest konstruowana przez urządzenie wyczuwające złożone obrazy. Klatki te są wyświetlane w oddalonym miejscu na powierzchni wyświetlania, która otacza użytkownika tworząc immersyjne wrażenie obecności i uczestnictwa w telekonferencji. Ten algorytm wydajnie tworzy klatki panoramiczne do sesji telekonferencyjnej, które są geometrycznie wyrównane i których intensywność jest ujednolicona. W tym rozdziale zademonstrowany jest prototyp, który przechwytuje obrazy z czujnika złożonych obrazów, wyrównuje je, aby tworzyły spójną klatkę panoramiczną i wyświetla te panoramiczne klatki na powierzchni wyświetlania z prędkością 30 kl./s

1. Wstęp

Czynnikiem ograniczającym, najczęściej kojarzonym z obecną technologią telekonferencyjną jest szerokość pasma sieci i jego wpływ na rozmiar obrazu i częstotliwość klatkowania. Inną, rzadziej wspominaną rzeczą ograniczającą użyteczność systemów telekonferencyjnych, jest pole widzenia przedstawiane przez obraz. Pole widzenia typowej kamery jest małe i ma niską rozdzielczość i jeśli używa się go jako podstawy telekonferencji, nie tworzy on przekonującego wrażenia obecności i immersji. Nawet kiedy całkowicie rozwiąże się problem szerokości pasma sieci, pozostaje relatywnie niska rozdzielczość przestrzenna i wąskie pole widzenia. Opisany w tym rozdziale algorytm został zaprojektowany do systemów telekonferencyjnych, które starają się być *immersyjne* i w których

obrazy mają być złożone z kawałków tworzących wspólnie panoramę lub rozległy obraz (zob. Rys. 1).



Obecna telekonferencja

Imersyjna telekonferencja

Rysunek 1. Różnica pomiędzy obecną i immersyjną telekonferencją

Mimo że telekonferencje immersyjne również obciążają sieć szerszą szerokość pasma, to jednak posuwają się o krok naprzód od systemów wąskiego pola widzenia i otwierają większe możliwości dla sesji, które wywołują wrażenie obecności użytkowników.

Decyzja o przechwytywaniu, przesyłaniu i wyświetlaniu rozległych obrazów spowodowana jest ceną w kilku dziedzinach. Pojawia się też problem tworzenia obrazów o szerokim polu widzenia bez widocznych łączeń, bez wprowadzania wizualnych artefaktów (sztuczności) nie do przyjęcia i takiej ilości obliczeń, z powodu których technika ta stałaby się niepraktyczna dla konferencji w czasie rzeczywistym. Rozwiązania jedno-czujnikowe, o wysokiej rozdzielczości i szerokim polu widzenia są bardzo drogie i nie ogólnie dostępne, co sprawia, że większość naukowców uważa, że obrazy immersyjne powinny być tworzone z urządzeń złożonego odwzorowania. Takie urządzenia są budowane z zestawu kamer, ustawiane z lustrami lub bez, aby ich zbiorowe pole widzenia generowało żądaną panoramę. Urządzenia te nie tworzą obrazów bez widocznych łączeń ponieważ ich ustawienie, nawet gdy jest dokładnie skalibrowane, jest niedokładne i prowadzi do widocznych artefaktów na granicach obrazowych, algorytm do tworzenia rozległych obrazów bez widocznych łączeń z zestawu obrazów podstawowych, przy dwóch konkretnych założeniach: że czujniki są ustawione tylko w przybliżeniu i że algorytm musi działać w czasie rzeczywistym. Algorytm wykorzystuje proste założeń geometrycznych, które są używane, gdy kamery są połączone dla uzyskania przybliżonego wspólnego wirtualnego środka projekcji.

Przedstawiony algorytm ustawienia geometrycznego tworzy odwzorowanie tekstur, które koduje krzywiznę uzyskaną w czasie inicjalizacji. Krzywizna stworzona jest w taki sposób, aby łączyć wszystkie pojedyncze obrazy z urządzenia

złożonego odwzorowania w jeden panoramiczny obraz. Odwzorowanie tekstur, które wylicza się tylko raz, jest dostarczane onlinowej części wyświetlającej systemu i jest w niej używane jako odwzorowanie do wyświetlania w czasie rzeczywistym obrazu z kamer na docelową powierzchnię wyświetlania. Algorytm fotometryczny, który tu zaproponowano, również ma dwie części składowe. W pierwszej fazie wykorzystuje technikę wielorozdzielczościowej krzywej składanej, aby wygenerować *odwzorowanie ciężkości w oparciu piksele*. W drugiej fazie koduje się to odwzorowanie do użytku online, aby uzyskać wyrównanie fotometryczne na granicach kamer. Stosując offlinowy algorytm dla uzyskania krzywizny i odwzorowania ciężkości opartego na pikselach, a następnie wykorzystując wyniki w onlinowym algorytmie wyświetlania, można przechwytywać, wyrównywać i wyświetlać obrazy panoramiczne o wysokiej rozdzielczości z prędkością 30 klatek/s.

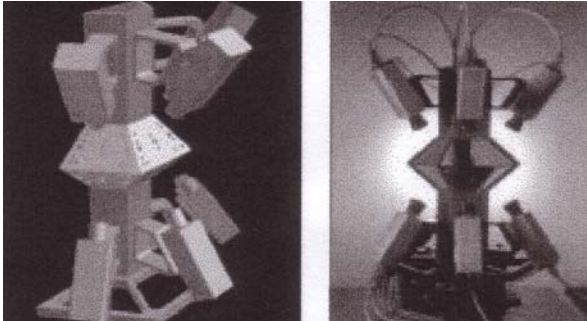
Przechwytywanie Obrazów Panoramicznych

Postępy w fotografii cyfrowej sprawiły, że tworzenie scen panoramicznych jest dziś zjawiskiem powszechnym. Stworzono również kilka czujników, które tworzą obrazy panoramiczne na pojedynczej płaszczyźnie obrazu przez użycie ciągłych powierzchni zwierciadlanych. W takich przypadkach nie jest wymagane wyrównywanie obrazu ani jego dopasowywanie, ponieważ ciągły obraz panoramiczny jest tworzony na jednym czujniku. Ta metoda wymaga łączenia wielu obrazów, ale rozdzielczość obrazu końcowego jest znacznie wyższa.

Metoda ta sprawia, że urządzenie złożonego odwzorowania może być stworzone z jakiegokolwiek zestawu kamer, który przypomina system ze wspólnym środkiem projekcji. Nie wymagana jest dokładność, ponieważ algorytm ma rekompensować złe ustawienie i aberracje w ustawieniu. Konkretny zestaw kamer użyty był jako urządzenie wejścia - jest zbudowany z wielu kamer, aby stworzyć symulację kamery wirtualnej z poziomym polem widzenia 360° i pionowym polem widzenia 90° . Ze względu na to, że obraz z jakiegokolwiek zestawu będzie się wydawał poprawny tylko wtedy, gdy kamery fizyczne będą miały ten sam środek projekcji, należało w przybliżeniu ustawić środki optyczne kamer. Wiele kamer nie może znajdować się w tym samym punkcie fizycznym, dlatego używa się luster, aby odwzorować optyczny środek danej kamery w innej lokacji fizycznej. Na każdy 60-stopniowy 'plaster' poziomy użyte są dwie kamery: jedna kamera na niższą część

zestawu, a druga na jego wyższą część. Cały układ mechaniczny ma ok. 400 mm długości, 450 mm wysokości i 450 mm szerokości.

Rys. 2 przedstawia geometrię zestawu (w CAD):



Rysunek 2. Jedna strona grupy kamer

Mimo że zestaw kamer, z którego korzystano, został skonstruowany z dużą dokładnością przez współpracowników na Uniwersytecie Utah, wymaga użycia algorytmu, który będzie zaprezentowany. Szczególnie dwa aspekty tego zestawu kamer sprawiają, że dopasowanie geometryczne i wyrównanie intensywności są trudne: ustawienie lustek powierzchni przedniej oraz niezgodności pomiędzy poszczególnymi kamerami. Zestaw został zaprojektowany, by dostosować lustra powierzchni przedniej, które przylegają do walcowanego metalu. Ustawienie lustek jest niedokładne, co prowadzi do niewłaściwego dopasowania obrazu. Błędy obrazu zdominowane są jednak przez różnice pomiędzy kamerami. Wykorzystano 12 identycznych kamer (JVC model TK-1270U), jedna na każde lustro, ale różnice pomiędzy kamerami są znaczne. Istniejące aberracje optyczne i mechaniczne sprawiają, że nie można wygenerować obrazu panoramicznego z zestawu kamer poprzez bezpośrednie dopasowywanie i rozmywanie. Zamiast tego, dla uzyskania poprawnego ustawienia użyto algorytmu dopasowania geometrycznego opartego na cechach.

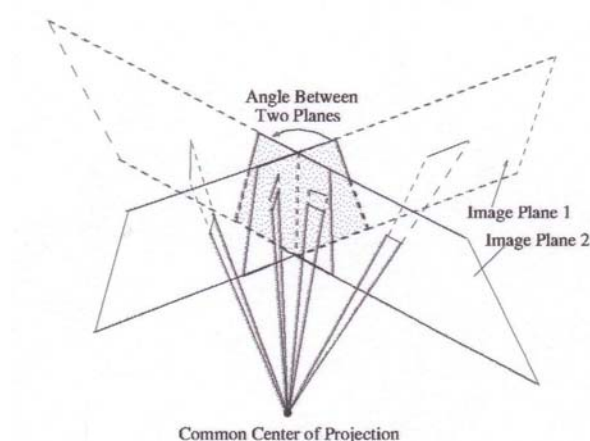
Obrazy z Zestawu Kamer. Projekt zestawu kamer zapewnia mały margines (region zachodzenia na siebie fragmentów obrazu) pomiędzy sąsiadującymi kamerami ze względu na nie-zeroowy rozmiar apertury kamer. W idealnym modelu kamery z obiektywem otworkowym nie byłoby marginesu.

Typowy obraz z kamery w zestawie widoczny jest w lewym górnym rogu tablicy nr 1. Ważna jest część zaznaczony na zielono. Każda kamera widzi trapezowe lustro, od którego odbija się przestrzeń świata. Kamera widzi też w swoim obrazie część siebie. Trapezowy rejon, w każdym obrazie, jest

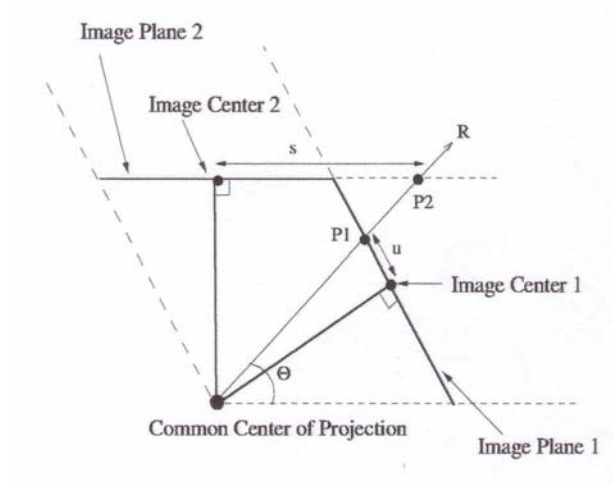
geometrycznie i fotometrycznie dopasowany do sąsiednich, tworzy panoramiczny obraz świata ze wspólnego środka projekcji. Problem dopasowania geometrycznego powodowany jest przez złe dopasowanie mechaniczne oraz przez różnice pomiędzy wewnętrznymi parametrami kamer. Odmienność fotometryczna spowodowana jest odmiennością balansu koloru oraz uzyskaną przez kamery intensywnością. Duża odmienność w intensywności obrazu z różnych kamer sprawia, że skorzystano z techniki opartej na cechach, której działanie w dopasowaniu geometrycznym jest solidniejsze niż w metodach opartych na korelacji. Obrazy uzyskane z kamer wykorzystywane są jako dane wejściowe do wyznaczenia współrzędnych rogów rejonu trapezowego w każdym obrazie. Te punkty narożne przyjmują poprawkę na zniekształcenia obiektywów. Ponieważ brzegi tego trapezowego rejonu w świecie rzeczywistym są liniami prostymi, czworokąt, jaki by się wytworzył z niezniekształconych punktów narożnych, dawałby rejon trapezowy.

Analiza Obrazów z Zestawu Kamer.

Sześć płaszczyzn obrazu zestawu kamer na dodatniej półpłaszczyźnie Z, gdzie oś Z odpowiada pionowej osi zestawu kamer na rys. 1. Zakładamy, że środek projekcji jest w miejscu pierwotnym. Pionowe płaszczyzny obrazu tworzą z płaszczyznami sąsiednimi kąt dwuścienny, wynoszący w przybliżeniu 127° , przy czym środek optyczny jest jednakowo odległy od środka projekcji. W takiej konfiguracji każda płaszczyzna obrazu będzie skierowana do wspólnego środka projekcji, a na brzegach będzie połączona z sąsiadującymi płaszczyznami tworząc trapez. Rys. 3 przedstawia dwie płaszczyzny wirtualnych obrazów w takiej konfiguracji. Płaszczyzny te są, dla jasności, rozciągane poza zasięg obrazu. W granicach obrazu, projekcja jednej płaszczyzny na drugiej jest zaznaczona przez przyciemnione rejonny.

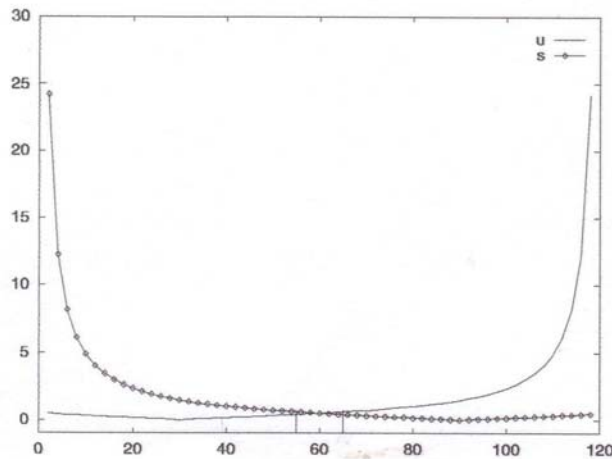


Rysunek 3. Zachodzące na siebie obszary w 3 D.



Rysunek 4. Analiza zachodzących na siebie obszarów

W wielkim kole płaszczyzny $Z=0$, która jest podstawą płaszczyzn obrazu z rys. 3, dolne brzegi sześciu płaszczyzn obrazu łączą się z sąsiednimi pod kątem 120° i każdy leży naprzeciwko środka projekcji, pod kątem 60° .



Rysunek 5. Związek pomiędzy u i s.

Dla uproszczenia rozważane są tylko dwie z sześciu powyższych płaszczyzn. Interesuje nas relacja pomiędzy systemami współrzędnych (u,v) i (s,t) tych dwóch płaszczyzn. Dwuwymiarowy odpowiednik tej konfiguracji, przedstawiony na rys. 4, tworzy się z brzegów podstawy tych dwóch płaszczyzn na płaszczyźnie $Z=0$. W dwuwymiarze interesuje nas relacja pomiędzy u i s . Łatwo rozszerzyć to na trójwymiar, włączając również v i t . Promień, R , ze środka projekcji przecina płaszczyznę (punkty) obrazu, np. $P1$ i $P2$. Na rys. 5 odległość tych wyświetlanych

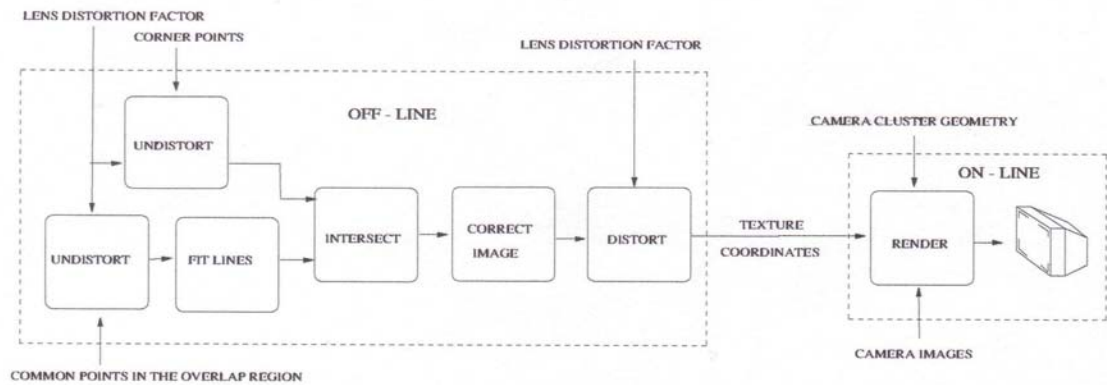
punktów P1 i P2 od środka optycznego odpowiadających im płaszczyzn obrazu jest nakreślana naprzeciwko kąta 0, który promień optyczny tworzy z linią poziomą. Odległości te są bezpośrednio powiązane ze współrzędnymi płaszczyzn obrazu u i s. W rejonie otaczającym skrzyżowanie tych dwóch linii obrazu (55 do 65 stopni), zmiana u i s jest niewielka, a dla małej zmiany u, zmiana s jest w przybliżeniu taka sama, co wskazuje na liniową relację pomiędzy tymi dwoma wartościami w rejonie zachodzenia na siebie marginesów. Oczywiście kiedy kąt dwuścienny płaszczyzn obrazu jest bliski 180° , to liniowość obszaru marginesów jest rozszerzona. Ta lokalnie liniowa relacja pomiędzy współzrędnymi dwóch płaszczyzn obrazu jest ważną wartością geometryczną, którą wykorzystano w tej metodzie do dopasowywania obrazów do panoramy. Ponieważ rejon marginesu jest wąski, a kąt pomiędzy płaszczyznami duży, założenie liniowe jest dokładne.

Następny rozdział omawia geometrię, która wymaga tego algorytm od urządzenia złożonego odwzorowania. Rozdział 2 przedstawia szczegóły algorytmu do zapewniania dopasowania geometrycznego, włączając w to odnajdowanie odpowiedniości i nadrabianie zniekształceń obiektywu. Rozdział 3 przedstawia przykłady i wyniki aplikacji, która była używana do wyświetlania obrazów panoramicznych z zestawu w czterościennym CAVE w czasie rzeczywistym. W czwartym rozdziale przedstawiony jest algorytm do uzyskiwania fotometrycznego rozmycia w granicach kamery, a rozdział 5 zawiera podsumowanie i omawia zalety metody.

2. Algorytm Dopasowania Geometrycznego: Zarys.

Poniżej przedstawione są założenia, które przyjęto do uzyskiwania właściwego dopasowania geometrycznego obrazów przechwyconych z zestawu kamer:

- Środki projekcji wszystkich kamer w zestawie są w przybliżeniu takie same.
- Rejon marginesu łączenia sąsiadujących obrazów jest wąski, 2 – 10% wymiaru obrazu. Np., dla obrazu 500x500, margines wynosi 10-50 pikseli.
- Pole widzenia (długości ogniskowych) kamer mogą być nieznane i nie są wymagane przez algorytm.
- Czynniki zniekształceń obiektywów radialnych kamer są używane, jeśli są znane.
- Przyjmuje się, że balans koloru pomiędzy obrazami jest różny.
- Dana jest geometria zestawu kamer.



Rysunek 6. Kolejność kroków

Rys. 6 przedstawia schemat procesu dopasowania geometrycznego. Algorytm przechodzi przez następującą sekwencję kroków. W zależności od systemu obiektywów, obrazy z kamer mogą wykazywać znaczne zniekształcenia obiektywów radialnych. Gdy pojawiają się zniekształcenia, jak w przypadku tego systemu eksperymentalnego, ocena zniekształceń dla każdej kamery może być dokonana przy użyciu standardowych procedur. Cztery punkty narożne rejonu trapezowego obrazów wszystkich kamer wyciąga się automatycznie z obrazów trapezowych luster widzianych z kamer, przy użyciu prostego przetwarzania migawek obrazu z kamer (wykrywanie zlokalizowanego brzegu). Te punkty narożne to rogi rejonu zielonego, widocznego w lewym górnym rogu tablicy nr 1. Następną czynnością jest sprofilowanie odpowiadających punktów pomiędzy dwiema sąsiednimi kamerami w wąskim rejonie marginesu. W następnym rozdziale opisana jest metoda jednolitego sprofilowania tych wspólnych odpowiadających punktów. Ze względu na to, że rejon marginesu jest wąski, a sprofilowanie odpowiadających punktów jest jednolite, najmniejszego kwadratu dopasowania tych sprofilowanych punktów na każdej płaszczyźnie obrazu niezależnie, obrazowałyby się nawzajem na dwóch płaszczyznach obrazu. Linie te nazywa się *liniami łączenia*. Parametryzacja tych linii będzie różna aż do poziomu czynnika skali, w zależności od pola widzenia (długości ogniskowych) kamer. Rozwiązano ten problem bez dokładnej znajomości pola widzenia wykorzystując odwzorowanie tekstur. W geometrii tego zestawu kamer każda kamera ma trzech sąsiadów i tym samym trzy linie łączenia. Te trzy linie określają czworokątny rejon wewnątrz pierwotnego rejonu trapezowego określonego przez punkty narożne. Określa to rejon wewnątrz trapezu bez marginesów. Następnie tworzy się z tego czworokątnego rejonu odwzorowanie tekstur na komputerowym modelu geometrii zestawu kamer. Kiedy

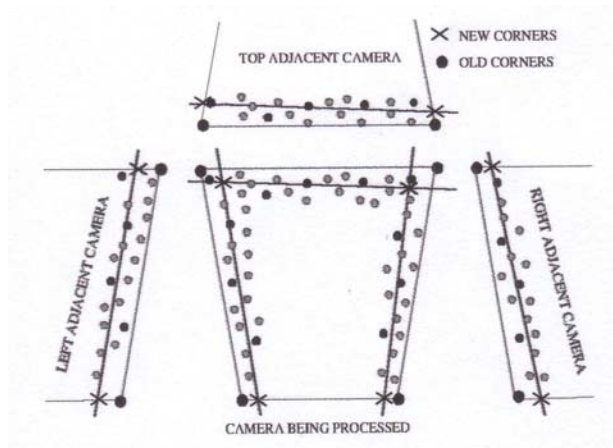
to samo stanie się ze wszystkimi obrazami, widok panoramiczny może zostać wyświetlony ze środka geometrii zestawu kamer.

Uzyskiwanie Odpowiadających Punktów.

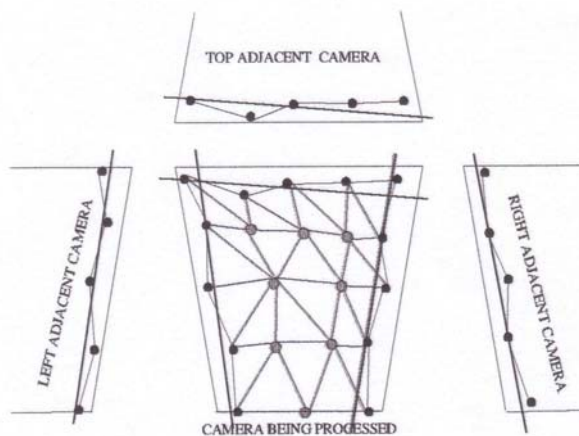
Należy odnaleźć odpowiadające punkty na sąsiednich kamerach przy użyciu projektora, który emituje wzory na ścianie wyczuwane przez sąsiednie kamery w zestawie. Aby uzyskać odwzorowanie pomiędzy współrzędnymi projektora (x,y) a współrzędnymi kamery (u_i, v_i) , przetwarzany jest obraz kamery i . Kiedy dwie kamery, np., i,j , widzą światło wyświetlane z tego samego pikselu projektora (x,y) odpowiednio w punktach (u_i, v_i) i (u_j, v_j) , wtedy $((u_i, v_i), (u_j, v_j))$ są to odpowiadające piksele dla tej pary kamer. Ponieważ przyjęto wspólny środek projekcji, nie trzeba znać odległości zestawu od ściany. Poza tym, ze względu na to samo założenie, rejon marginesu jest stały, bez względu na głębię obiektu widzianego przez kamery. Jeśli założenie o wspólnym środku projekcji jest nieważne, odległość do obiektu zaczyna być ważna w łączeniu sąsiednich obrazów tego obiektu. Te kamery nie mają dokładnie tego samego wirtualnego środka projekcji ze względu na różnice powstałe podczas ich produkcji. Jednakże odległość od środka projekcji każdej kamery do wspólnego wirtualnego środka jest bardzo mała. Dzięki temu procedura samplowania odpowiadających punktów jest bardzo prosta. Zestaw można obracać, dzięki czemu różne pary sąsiednich kamer mogą widzieć ścianę, przy czym nie jest ważna odległość od niej; procedurę tę powtarza się dla różnych rejonów marginesów. Projektor używany jest tylko dla wygody i automatyzacji.

Odnajdowanie Linii Łąceń.

Odpowiadające punkty zebrane z pary kamer (i, j) przedstawiane są jako $((u_i, v_i), (u_j, v_j))$. Najmniejszy kwadrat linii, np. L_i , pasuje do (u_i, v_i) , a inny najmniejszy kwadrat linii, np. L_j , pasuje do (u_j, v_j) . Obrazem L_i na kamerze j jest L_j , ze względu na liniową relację pomiędzy $u_i(v_i)$ i $U_j(v_j)$ w rejonie marginesu. Stała proporcjonalności w tej relacji liniowej przypisywana jest różnicom w polach widzenia. Jeśli pola widzenia kamer są takie same, wtedy przejście z $u_i(v_i)$ do $u_j(v_j)$ można wykonać przez przesunięcie równoległe i obrót.



Rysunek 7. Linie łączące punkty w narożnikach



Rysunek 8. Korekta Obrazu Oparta na Odpowiedności.

Rys. 7 przedstawia działanie tej części algorytmu dla jednej kamery. Punkty narożne uzyskane z trapezu obrazu kamery są przedstawione jako duże okrągłe kropki. Wspólne odpowiadające punkty pomiędzy sąsiednimi kamerami to małe szare kropki. Odpowiadające punkty to małe czarne kropki. Grube linie przedstawiają trapez utworzony przez punkty narożne, a linie przerywane to linie najmniejszego kwadratu dopasowania. Interesują nas punkty skrzyżowań, które zaznaczone są krzyżykami. Te krzyżyki to nowe punkty narożne, a czworokąt przez nie utworzony, to dostosowany rejon, który nas interesuje. Wnętrze wielokąta to maksymalny rejon obrazu, na którym nie ma marginesów.

Kiedy przetwarzanie dla wszystkich kamer jest zakończone, tekstury tych czworokątnych obrazów są odwzorowane na powierzchnie trapezowe (wielokąty) tworząc geometrię, która przypomina geometrię zestawu kamer. W idealnym układzie należy odwzorować teksturę czworokątnego rejonu z rys. 7, na trapezowej geometrii kamer. Widok tej grupy obrazów ze środka geometrii zestawu kamer powinien mieć niewidoczne łączenia pomiędzy obrazami. Ze względu na różne

niedokładności i błędy obliczeń, obrazy generowane według powyższej metody będą miały marginesy widoczne w obu sąsiadujących kamerach lub połączone regiony, które nie będą widoczne w żadnej z tych kamer. Dlatego trzeba wnosić dokładniejsze poprawki opisane poniżej.

Korekta Obrazu Oparta na Odpowiedniości.

Metoda ta tworzy krzywiznę obrazu opartą na pikselach na dwóch sąsiednich obrazach, dzięki czemu kilka odpowiadających punktów nachodzi na siebie na granicy wspólnego obrazu tworząc dopasowanie geometryczne bez widocznych łączy.

Linie łączące pomiędzy kamerami i i j to $L_{ij}(\alpha)$ na płaszczyźnie obrazu kamery i , oraz odpowiadająca jej linia $L_{ji}(\beta)$ na płaszczyźnie obrazu kamery j . Odpowiadające punkty narożne tych dwóch płaszczyzn obrazów otrzymywane są z $(L_{ij}(0), L_{ji}(0))$ i z $(L_{ij}(1), L_{ji}(1))$. Ponieważ punkty końcowe to punkty odpowiadające, a relacja pomiędzy współrzędnymi jest liniowa, obraz $L_{ij}(\alpha)$, $0 \leq \alpha \leq 1$, na kamerze i to $L_{ji}(\alpha)$, $0 \leq \alpha \leq 1$ na kamerze j . Parametryzacja ta eliminuje potrzebę znajomości czynnika skali, wnoszonego przez różniące się pola widzenia różnych kamer.

Podczas pierwszego kroku poprawiania obrazu wybieramy n punktów ze zbioru odpowiadających punktów $((u_i, v_i), (u_j, v_j))$, które są bardzo blisko linii łączenia L_{ij} i L_{ji} . Zakładamy, że punkty te to $((u_{ik}, v_{ik}), (u_{jk}, v_{jk}))$, $1 \leq k \leq n$. Linie łączące te wybrane punkty (u_{ik}, v_{ik}) $((u_{jk}, v_{jk}))$ tworzą kawałkami liniową aproksymację (przybliżenie) $M_{ij}(M_{ji})$ dla $L_{ij}(L_{ji})$. Zakładając, że M_{ij} i M_{ji} to nowe granice łączenia, to następnym krokiem jest parametryzacja M_{ij} i M_{ji} . Weźmy najbliższy punkt w L_{ij} do (u_{ik}, v_{ik}) będzie $L_{ij}(\alpha_k)$. Przyporządkujemy $M_{ij}(\alpha_k)$ będzie (u_{ik}, v_{ik}) i $M_{ji}(\alpha_k)$ będzie (u_{jk}, v_{jk}) . Jeżeli nie ma przyporządkowanych punktów do $M_{ij}(0)(M_{ji}(0))$ i $M_{ij}(1)(M_{ji}(1))$ wtedy przyporządkujemy $M_{ij}(0)=L_{ij}(0)(M_{ji}(0)) = L_{ji}(0)$ i $M_{ij}(1)=L_{ij}(1)(M_{ji}(1)=L_{ji}(1))$. Wszystkie inne wartości $M_{ij}(M_{ji})$ są interpolowane pomiędzy przyjętymi wartościami. Ta kawałkami liniowa aproksymacja uważana jest za nową granicę łączącą, jak na rys. 8. Zapewnia to dopasowanie przynajmniej w punktach pośrednich, bo są one w rzeczywistości punktami odpowiadającymi. Dlatego M_{ij} i M_{ji} dają współrzędne tekstury dla linii granicznej pomiędzy kamerami i i j w procesie wyświetlania.

Na rys. 8 i 7, punkty, które są wybrane jako bliskie linii granicznej, są zaznaczone małymi ciemnymi kropkami na obu sąsiednich płaszczyznach. Odległość tych punktów od linii granicznej jest przesadzona na rysunku dla lepszego zobrazowania.

Aby rozszerzyć powyższy proces w kierunku poprawiania zniekształceń obiektywów także wewnątrz obrazu, wybiera się wiele punktów przykładowych w granicach trapezu. Tym przykładowym punktom przypisuje się współrzędne tekstury poprzez interpolowanie pomiędzy współrzędnymi tekstury niezniekształconych punktów granicznych. Na rys. 8 są to szare punkty.

Opis. Jak opisano powyżej, ten algorytm wylicza region łączenia dla każdej kamery i przypisuje ten wielokątowy region łączenia jako odwzorowanie tekstury obrazowi pochodzącemu z każdej kamery zestawu. Odwzorowanie tekstury jest następnie dzielone przez odpowiedniości innych punktów na poprawne nieliniowe zjawiska jak np. zniekształcenia obiektywów. Całkowity wpływ stosowania odwzorowania tekstury w każdej kamerze zestawu to nieliniowe tworzenie krzywizn. Ogólna „kawałkowa” krzywizna w formie odwzorowania tekstury skonstruowanego przez algorytm, to bardzo dobry sposób nadrabiania tych wszystkich zjawisk nieliniowych, które powodują zniekształcenia, widoczne marginesy i przede wszystkim sprawiają, że obraz nie wygląda na jednolity.

Na tablicy nr 1 pokazane są wyniki każdego kroku dopasowania górnego obrazu do sąsiednich. W prawym górnym rogu pokazany widok z kamery przed dokonaniem dopasowania geometrycznego. Środkowe zdjęcie z lewej strony przedstawia ten sam obraz, po odnalezieniu linii łączącej. Należy zauważyć, że poprawia się dopasowanie wzdłuż granicy z lewą sąsiednią kamerą. Środkowe zdjęcie z prawej strony i dolne z lewej przedstawiają ten sam obraz po odnalezieniu odpowiednio dolnej i prawej linii łączącej. Ostatnie zdjęcie, w prawym dolnym rogu przedstawia obraz po korekcie obrazu opartej na odpowiedności.

3. Wyświetlanie i wyniki. Wyświetlanie obrazu dla uzyskania dopasowania geometrycznego wideo w czasie rzeczywistym jest ważną częścią tego algorytmu. Niezniekształcony obraz jest doklejany do linii łączących i poprawiany metodami korekty obrazu. Poprawiony obraz postrzegany jest jako odwzorowanie tekstury i jest odwzorowywany na trójkątną trapezową płaską powierzchnię przedstawiając konkretną geometrię zestawu kamer (w tym przypadku jest to widok lustrzany zestawu). Wszystkie obrazy są odwzorowywane na geometrię płaską takich samych

rozmiarów, która jest określana przez fizyczną geometrię zestawu. Dzięki temu eliminuje się czynnik skali w relacji pomiędzy współrzędnymi dwóch sąsiednich płaszczyzn obrazu spowodowany różnicami w polu widzenia. Ze względu na to, że faktyczny obraz po dokonaniu poprawek nie jest dokładnym trapezem, jaki widać na rys.8, odwzorowanie tekstury na powierzchnię trapezową spowodowałoby zniekształcenia. Jednak w praktyce takie odchylenie tekstury od trapezu jest nie większe niż kilkupikselowe i efekt zniekształcenia jest nieznaczny.

Trzeba podkreślić, że wszystkie powyższe wyliczone współrzędne tekstury na obrazie są otrzymywane po usunięciu zniekształceń obiektywów. Jednak współrzędne tekstury muszą w końcowym efekcie być przedstawione na ramie współrzędnych, która rekompensuje zniekształcenia każdego poszczególnego obiektywu. Ponieważ odwzorowujemy tekstury ruchomego strumienia wideo z kamer, modeluje się współrzędne przed zastosowaniem zniekształcenia obiektywów przy użyciu odwrotnego przekształcenia zniekształcenia obiektywów. Następnie współrzędne te są przypisywane właściwym współrzędnym geometrycznym samego modelu zestawu kamer. Szczególnie dla uproszczenia przyjmuje się kwadratowy obraz $I'(U', V')$, który jest surowym obrazem z kamery uzyskiwanym na żywo. Obraz ten jest poprawiany pod względem zniekształceń obiektywów do $I(U, V)$. Poprawiony obraz jest odwzorowany na płaszczyznę $P(U, V)$, na której punkt $P(u, v)$ ma przypisaną wartość koloru $I(u, v)$. W celu uzyskania odwzorowania tekstury na żywo, przypisuje się wartość koloru $I'(u', v')$ do $P(u, v)$, który jest zniekształconą współrzędną (u, v) . Proces ten poprawia również zniekształcenia obiektywów, ponieważ lokacja fizyczna $P(u, v)$ przyciąga piksel $I'(u', v')$ do właściwej lokacji, która jest równa niezniekształceniu obiektywów.

Wyniki tego algorytmu przedstawione są na tablicach 2 i 3 na końcu rozdziału. Tablice kolorów przedstawiają migawki wyświetlania w czasie rzeczywistym ruchomych obrazów z zestawu 10 kamer. Po wygenerowaniu pliku tekstury można było przesunąć zestaw kamer w różne miejsca i uzyskać geometrycznie dopasowany obraz panoramiczny środowiska naokoło zestawu. Na tablicy nr 2 przedstawiony jest panoramiczny widok trzech takich środowisk. Na tablicy 3 – dwóch środowisk. Dwa z tych środowisk przedstawiają ludzi i są migawkami jednej klatki panoramicznego obrazu wideo. Fotometryczne różnice kolorów z różnych kamer powodują, że widoczne jest łączenie poszczególnych części panoramy. Obecnie naukowcy pracują nad tym problemem. Ponieważ istnieją

pewne błędy w ustawieniu środka projekcji różnych kamer w jednym punkcie, widać w różnych miejscach niewielkie odchylenia od dopasowania. Proces zbierania odpowiadających punktów dla zestawu 10 kamer trwa ok. 20 s. Offlinowy proces generowania plików tekstury, przeprowadzany po uzyskaniu odpowiadających pikseli, trwa w przybliżeniu 5 ms na kamerę. Proces generowania plików tekstury dla całego zestawu trwa ok. 60-70 ms. Wyświetlanie i synchronizacja była wykonywane na SGI 02 z procesorem R10000.

Złożono pierwszy prototyp immersyjnego środowiska telekonferencyjnego we współpracy z naukowcami z Uniwersytetu Illinois w Chicago. Uzyskano obrazy z zestawu kamer z UNC i przesyłano je przez Internet do Laboratorium Wizualizacji Elektronicznej Uniwersytetu w Chicago (UIC). Wstępnie wyliczone odwzorowanie tekstury jest przesyłane z UNC do UIC. Otrzymują oni obrazy przez sieć i wykorzystując wstępnie wyliczone odwzorowanie tekstury wyświetlają dopasowaną geometrycznie panoramę wideo na systemie CAVE.

4. Brak Łąceń Fotometrycznych.

Granice kamer są widoczne mimo poprawnego dopasowania geometrycznego z powodu różnic fotometrycznych pomiędzy dwiema kamerami. To znaczy, że jest duża różnica pomiędzy kamerami we wrażliwości na kolor, co sprawia, że obrazy nie pasują do siebie. Dla rozwiązania tego problemu zastosowano metodę podobną do rozwiązania problemu dopasowania geometrycznego. Tworzy się algorytm, który ma część offlinową i onlinową. Algorytm offline będzie użyty tylko raz – do zdefiniowania odwzorowania ciężkości w oparciu o piksele. Następnie dokonuje się rozmywania w trybie online, przy użyciu graficznego przetwarzania potokowego. Dla stworzenia odwzorowania ciężkości używa się klasycznego algorytmu do rozmywania dwóch obrazów. Termin *krzywa składana obrazu* odnosi się do technik cyfrowych używanych do zapewniania płynnego przejścia w sąsiedztwie granicy pomiędzy dwoma obrazami. Dobra krzywa składana obrazu zapewnia płynne przejście, a oprócz tego zachowuje większość pierwotnych informacji o obrazie. Używając techniki składania średniej ważonej, najpierw dzieli się obraz na środkowo-przepustowy zbiór części. Inna krzywa używana jest dla każdego środkowo-przepustowego komponentu po to, by rozmyć przejścia pomiędzy granicami kolorów. Na koniec elementy środkowo-przepustowe są podsumowywane dla wygenerowaniażądanego rozmytego na obrazu.

Definicje.

Podrozdział ten przedstawia przegląd definicji i różnych operacji na obrazie, które są użyte dla rozmycia granic pomiędzy częściami obrazu. Załóżmy, że G_0 to obraz pierwotny. Można uzyskać sekwencję dwuprzepustowo filtrowanych obrazów G_0, G_1, \dots, G_N poprzez zapętlenie małej funkcji korygującej w stosunku do obrazu. Jeśli wyobrazimy sobie, że obrazy te leżą jeden na drugim, otrzymamy stożkową strukturę danych znaną jako piramida. Jeśli G_0 ma rozmiar $(2^N + 1)$ przez $(2^N + 1)$, to będzie $N+1$ poziomów piramidy. Ta operacja na obrazie na poziomie $l - 1$ dla wygenerowania obrazu na poziomie l nazywa się redukcją (reduce).

Więc dla $0 < l < N$
$$G_l = \text{REDUCE}(G_{l-1})$$

Jądro Gaussa jest używane jako filtr dwuprzepustowy. Definiuje się je przez ciężkości $w(m,n)$, gdzie $-2 \leq m,n \leq 2$, stąd piramida, jaką otrzymano nazywa się piramidą Gaussa. Wyrażenie matematyczne funkcji redukcji jest następujące:

$$G_l(i, j) = \sum_{m=-2}^2 \sum_{n=-2}^2 w(m, n) G_{l-1}(2i + m, 2j + n)$$

Gdzie $0 \leq i, j < q$ i G_l jest $q \times q$ obrazu.

Każdy piksel (i,j) w G_l reprezentuje obciążoną średnią pod-układu pikseli G_{l-1} o wymiarach 5×5 , ze środkiem w punkcie (i,j) . Każdy z nich po kolei reprezentuje średnią pod-układu pikseli w G_{l-2} . W ten sposób możemy dojść do ciężkości w G_0 , aby znaleźć odpowiadającą funkcję korygującą, W_l , która zapętlona bezpośrednio z G_0 wygeneruje obraz w G_l . To daje nam pojedyncze przekształcenie matematyczne r_l dla wygenerowania G_l z G_0 .

$$G_l = r_l(G_0), 0 < l < N$$

Piramida Gaussa jest zbiorem dwuprzepustowo filtrowanych obrazów. Teraz definiujemy operację rozszerzania (EXPAND) na G_l , którą można podsumować jako operację przesamplowania. Zakładamy, że $G_{l,k}$ to obraz otrzymany przez rozszerzenie G_l k razy. Więc

$$G_{l,k} = \text{EXPAND}(G_{l, k-1})$$

Wyrażenie matematyczne dla operacji rozszerzenia jest następujące:

$$G_{l,k}(i, j) = 4 \sum_{m=-2}^2 \sum_{n=-2}^2 G_{l, k-1}\left(\frac{2i + m}{2}, \frac{2j + n}{2}\right)$$

Gdzie $0 \leq i, j < q$ i $G_{l, k-1}$ jest $q \times q$ obrazu. To oznacza, że $G_{l, 1}$ jest takiej samej wielkości jak G_{l-1} , a $G_{l,l}$ takiej samej wielkości jak obraz pierwotny. Zakładając, że

e_l to przemiana dla uzyskania $G_{l,l}$. $G_{l,l} = e_l(G_l)$

Teraz określamy sekwencję dwuprzepustowych obrazów L_0, L_1, \dots, L_N , $0 < l < N$,

$$L_l = G_l - G_{l+1, l}$$

a $L_N = G_N$. Nazywa się to piramidą Laplace'a. Tak samo jak piramidę Gaussa można było uzyskać bezpośrednio zapętłając W_l z G_0 , podobnie każde L_l można uzyskać bezpośrednio zapętłając funkcję korygującą ($W_l - w_{l+1}$) z G_0 . Tak więc można uzyskać pojedyncze przekształcenie matematyczne h_l , aby z G_0 wygenerować L_l .

$$L_l = h_l(G_0), 0 < l < N$$

Ponownie $L_{l,l} = e_l(L_l)$

Kroki zastosowane dla uzyskania piramidy Laplace'a można odwrócić, aby dokładnie przywrócić pierwotny obraz G_0 . Najpierw rozszerza się $L_N = G_N$, a potem dodaje do L_{N-1} , aby wygenerować G_{N-1} . Następnie G_{N-1} jest rozszerzane i dodawane do L_{N-2} , aby wygenerować G_{N-2} . Kontynuując operację można wygenerować G_0 . Można to zapisać jako

$$G_0 = \sum_{l=0}^N L_{l,l}$$
$$\Rightarrow G_0 = \sum_{l=0}^N e_l(L_l)$$

Zarys Algorytmu.

Algorytm oparty na powyższych operacjach, do rozmywania granic obrazów pomiędzy kamerami jest offlinowym algorytmem, który nie jest w stanie uzyskać niewidocznych łączeń w czasie rzeczywistym. Trzeba go zmodyfikować, aby wygenerować odwzorowanie ciężkości w oparciu o piksele i zastosować je w czasie rzeczywistym.

Panoramę G tworzy się z obrazów I_0, I_1, \dots, I_9 pochodzących z różnych kamer, przy życiu algorytmu dopasowania geometrycznego. Granice pomiędzy tymi obrazami są widoczne z powodu różnic fotometrycznych. Następująca procedura opisuje metodę generowania panoramicznego obrazu GB z niewidocznymi łączeniami z geometrycznie dopasowanego obrazu panoramicznego G . Najpierw generujemy po kolei piramidy Laplace'a P_0, P_1, \dots, P_9 . Każdy poziom l piramidy P_i , $0 \leq i \leq 9$ jest generowany przez

$$P_{i,l} = h_l(I_i)$$

Piramidę Laplace'a BP dla obrazu panoramicznego G tworzy się przez składanie P0, P1, ..., P9, według opisu. Zakładamy, że funkcja składania dla poziomu l to f_l .

$$BP_l = f_l(P1_l, P2_l, \dots, P9_l)$$

gdzie BP_l to poziom l piramidy Laplace'a BP. Aby wygenerować ostateczny rozmyty obraz G_B ,

$$G_B = \sum_l BP_{l,l}$$

$$\Rightarrow G_B = \sum_l e_l(BP_l)$$

Można więc całe przekształcenie zapisać jako,

$$G_B = \sum_l e_l(f_l(P1_l, P2_l, \dots, P9_l))$$

$$\Rightarrow G_B = \sum_l e_l(f_l(h_l(I1), h_l(I2), \dots, h_l(I9)))$$

Jeżeli znamy dokładne definicje rodziny funkcji e_l , f_l i h_l , otrzymujemy zamknięte rozwiązanie dla G_B poprzez rozszerzenie powyższego równania. Dla każdego $G_B(x,y)$ można znaleźć zbiór pikseli $S_{x,y}$ w G i odpowiadające im ciężkości $w(u,v)$ dla każdego pikselu $(u,v) \in S_{x,y}$, tak aby

$$G_B(x, y) = \sum_{S_{x,y}} w(u, v) \cdot G(u, v)$$

Mimo że wydaje się, że da to funkcję dla każdego $G_B(x,y)$, to w praktyce tylko piksele w sąsiedztwie granic będą miały więcej niż jeden piksel w $S_{x,y}$. Stąd, kiedy znalezione są $S_{x,y}$ i odpowiadające im ciężkości, należy rozsądnie wykorzystać graficzne przetwarzanie potokowe do wyliczenia $G_B(x,y)$.

Obecnie stosuje się tę wersję rozmycia fotometrycznego w celu uzyskania rozmycia w czasie rzeczywistym. Wynik wygląda tak jak zdjęcie 3 na tablicy nr 3, które jest uzyskiwane przez zastosowanie offlinowej wersji algorytmu. Jak pokazują tablice kolorów, rozmywanie wygładza łączenia obrazów dwóch kamer i nie są one widoczne. Jeśli połączone to zostanie z właściwym dopasowaniem geometrycznym, efekt jest naprawdę imponujący. Sprawdzone obecnie techniki konieczne do użycia graficznego przetwarzania potokowego, tak aby można było uzyskać działanie systemu w czasie rzeczywistym.

5. Podsumowanie i Wnioski.

W tym rozdziale przedstawiono algorytm do generowania panoramicznych obrazów w czasie rzeczywistym z danych otrzymanych z zestawu kamer ze wspólnym środkiem projekcji. Jednym nowatorskim aspektem jest użycie odwzorowania tekstur, które koduje krzywiznę obrazu oraz korekty zniekształceń obiektywów dla uzyskania dopasowania geometrycznego. Algorytm oddziela dopasowanie geometryczne, dokonywane jako krok wstępny w trybie offline, od onlinowego procesu wyświetlania. To oddzielenie pozwala na efektywne użycie konwencjonalnego graficznego wyświetlania potokowego dla uzyskania wyświetlania w czasie rzeczywistym. Kroki te są instrumentalne w uzyskiwaniu obrazów panoramicznych o wysokiej rozdzielczości w czasie rzeczywistym, które może być użyte do zapisywania ruchomych obrazów z zestawu kamer, które mają zbliżony wspólny środek projekcji.

Drugą cechą charakterystyczną tego algorytmu jest *lokalność* aplikacji. Dane z każdej kamery przetwarzane są oddzielnie, a algorytm zachowuje globalne dopasowanie poprzez sekwencję lokalnych dopasowań. Dzięki temu łatwo jest dopasować poszczególne kamery w zestawie, ponieważ algorytm musi być zastosowany tylko do kamery, która ma być dopasowana.

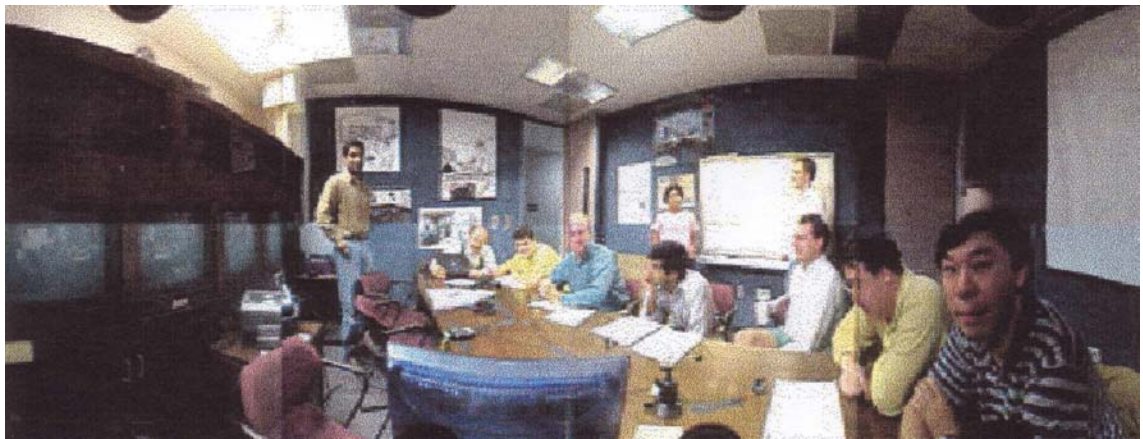
W końcu, w obliczu dużych różnic kolorów pomiędzy obrazami z różnych kamer, algorytm wykorzystuje metodę światła strukturalnego i koncentruje się na dopasowaniu cech zamiast na metodach korelacji, na które różnice w kolorze mogą mieć znaczny wpływ. Przedstawiliśmy również sposób uzyskania rozmycia fotometrycznego w czasie rzeczywistym na granicach obrazów w geometrycznie dopasowanej panoramie wideo. Do tego celu algorytm wykorzystuje konwencjonalne graficzne przetwarzania potokowe.



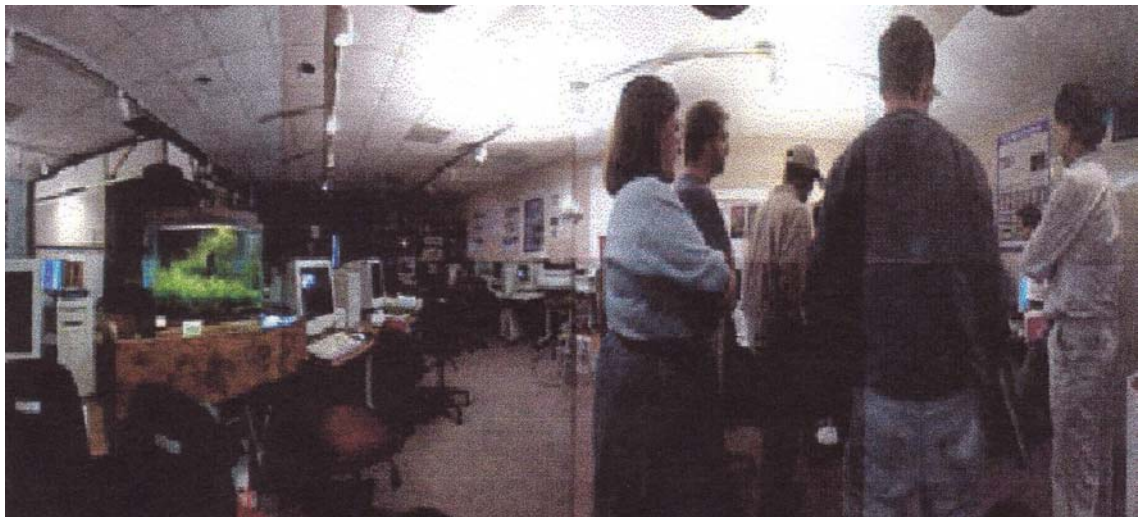


Tablica 1. Górny lewy rysunek: Obrazy z typowej grupy kamer. Następne pięć obrazów pokazuje kroki algorytmu „zszywania” – łączenia obrazów - z górnej kamery. Środkowy lewy: Obraz po spięciu(złączeniu) z lewym . Środkowy prawy: Obraz po złączeniu z dolnym. Dolny lewy: Obraz po złączeniu prawym. Dolny prawy: Obraz po bazującej na odpowiedniości korekcie obrazu.





Tablica 2. Na tych rysunkach zostało pokazanych kilka panoram wygenerowanych z grup kamer. Są to zdjęcia pojedynczych klatek w panoramicznym obrazie, kolejno wygenerowanych w czasie rzeczywistym.





Tablica 3. Górny i środkowy rysunek: O dwie więcej panoramy wygenerowane przez grupę kamer. Środkowy rysunek: Środkowa panorama po zastosowaniu algorytmu łączenia fotometrycznego w poprzek granic kamery.

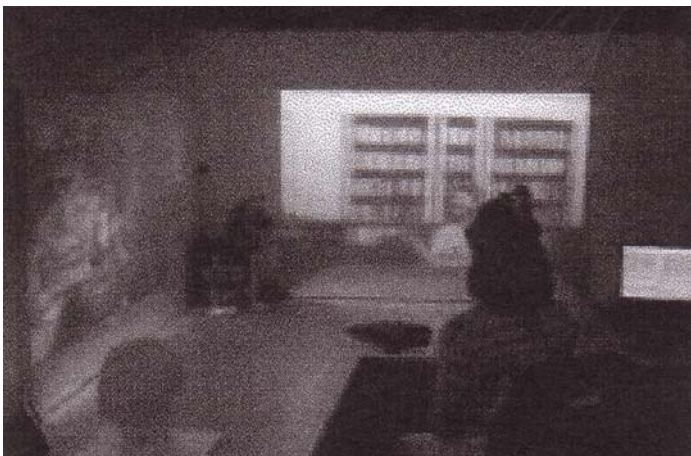
VII. Uzyskiwanie Sceny Niezależnej od Punktu Patrzenia w Tele-Obecności

Tele-imersja to nowe medium, które pozwala użytkownikom z różnych, oddalonych od siebie miejsc korzystać z tej samej przestrzeni wirtualnej. Użytkownik jest zanurzony w wyświetlanym trójwymiarowym świecie, który jest przesyłany z odległego miejsca. Dla uzyskania takiego trójwymiarowego obrazu stosuje się dwu- i trójobiektywowe techniki obrazu stereoskopowego. Wyzwaniem jest obliczenie danych o zwartym zakresie stereo przy wysokiej częstotliwości klatkowania, ponieważ uczestnicy nie mogą się z łatwością porozumiewać, jeśli cykl przetwarzania czy opóźnienia sieciowe są długie. Poza tym, nowy widok otrzymanego świata trójwymiarowego musi być tak dokładny jak tylko jest to możliwe. W tym rozdziale omówione będą oba zagadnienia - prędkości i

dokładności – i przedstawiona metoda łączenia ruchu i obrazu stereo w celu zwiększenia szybkości i odporności.

Wstęp

Moc dzisiejszych procesorów ogólnego zastosowania i graficznych oraz duża szerokość pasma ostatnich generacji internetowych zapewniają konieczną infrastrukturę dla systemów Mieszanej Rzeczywistości, które mogą wzmacniać wrażenia użytkownika i stwarzać wrażenie tele-obecności. Tele-imersja pozwala użytkownikom z fizycznie odległych miejsc współpracować na wspólnej przestrzeni, która łączy rzeczywistość lokalną z oddaloną. Koncepcja tele-imersji łączy wrażenia wizualne, słuchowe i dotykowe. Do chwili obecnej zajmowano się tylko częścią wizualną i we współpracy z UNC (z Henrym Fuchsem i jego współpracownikami) oraz z Advanced Network & Services (Jaronem Lanierem) poczyniono znaczące kroki w kierunku realizacji tele-imersji wizualnej.



Rys. 1 Użytkownik w Chapel Hill, uzbrojony w okulary polaryzujące i optyczne urządzenie śledzące porozumiewa się z dwoma odległymi użytkownikami z Filadelfii (po lewej) i z Armonk (po prawej). Wyświetlane stereoskopowo odległe sceny trójwymiarowe są składane z wchodzących strumieni teksturowanych danych trójwymiarowych opisujących użytkowników i ze statycznego tła uzyskanego offline.

Rysunek 1 najlepiej ilustruje obecne osiągnięcia. Zdjęcie to zostało zrobione podczas pełnej demonstracji na UNC. Użytkownik ma założone bierne okulary polaryzujące i optyczne urządzenie śledzące, które określa pozycję głowy. Na dwóch ścianach wyświetlane są dwie rzeczywistości – z Filadelfii i z Armonk, które są wyświetlane stereoskopowo ze spolaryzowanych par projektorów. Statyczne części tych dwóch scen są niezależnymi od punktu widzenia trójwymiarowymi

obrazami uzyskanymi off-line. Trójwymiarowe obrazy osób na pierwszym planie uzyskiwane są w czasie rzeczywistym i przesyłane przez sieć. Projekcje na ścianach są dynamicznie zmieniane, odpowiednio do punktu widzenia użytkownika, i aktualizowane przez rekonstrukcje rzeczywistego świata w rzeczywistym czasie dla zwiększenia wrażenia, że uczestnicy siedzą przy tym samym stole konferencyjnym. Ta pierwsza na świecie prezentacja potwierdza słuszność przewidywań Raymonda Kurzweila co do roku 2019:

Można zrobić właściwie cokolwiek i z kimkolwiek, bez względu na bliskość fizyczną... Rozmowy telefoniczne odbywają się na zasadzie przesyłania obrazów trójwymiarowych wysokiej jakości, wyświetlanych bezpośrednio w oku i przez soczewki słuchowe. Pojawiły się też trójwymiarowe obrazy holograficzne. W obu przypadkach użytkownikom wydaje się, że są blisko tej drugiej osoby... W taki sposób można się nabrać myśląc, że jakaś osoba jest fizycznie obecna, gdy w rzeczywistości jest wyświetlana przez łącze elektroniczne. Większość "spotkań" nie wymaga fizycznej obecności w miejscu spotkania.

Długa jeszcze droga nim będzie można rzeczywiście "oszukać" zmysły użytkownika. Mimo to jednak można już zacząć testować to nowe medium w dziedzinie tele-współpracy. Uczestnicy tych eksperymentów stwierdzili, że środowisko tele-imersyjne jest lepsze od tradycyjnego sposobu – wideo-konferencji.

Istnieją dwa alternatywne podejścia do odległej imersji, których nie badano. Pierwsze z nich to przeprowadzanie dużych wideo-konferencji: okrężna projekcja dwuwymiarowych obrazów panoramicznych. Sposób ten wymaga jedynie poprawnego dopasowania kilku widoków, jednak pozbawiony jest wrażenia głębi i nie pozwala na jakąkolwiek trójwymiarową interakcję z obiektami realnymi i wirtualnymi. Druga technologia jest bardziej zbliżona do opisywanej i wykorzystuje trójwymiarowe graficzne opisy odległych uczestników. Technika taka mogłaby w przyszłości zostać połączona z metodą opisywaną, gdyby można było uzyskiwać modele w oparciu o obecne punkty surowej głębi. Jest to kolejny sposób patrzenia na dwie skrajności – niezależność od modelu kontra zależność od modelu - w trójwymiarowych opisach scen czy w kontrowersji dół-góra kontra góra-dół. Przy założeniu, że mamy do czynienia z osobami, w przyszłości mogą być stosowane lub uzyskiwane modele osób o wysokiej jakości. Obecnie jednak tele-współpraca oparta na wcieleniach jest na poziomie odwzorowań wyglądających jak kreskówki.

Porównując tele-imersję do klasycznej powiększanej rzeczywistości dowiadujemy się, że śledzenie ruchów głowy użytkownika i częstotliwość odświeżania obrazu w czasie rzeczywistym stwarzają raczej małe problemy. Różnica natomiast jest w fakcie, że, po pierwsze, wykorzystywane wyświetlacze to wyświetlacze przestrzennie powiększane, a nie helmy do rzeczywistości wirtualnej, a po drugie, że elementy mieszane to nie wcześniej przygotowane doskonale obiekty wirtualne, ale otrzymywane na bieżąco dane rzeczywistego zasięgu. Dodatkowo, dane te przesyłane są przez sieć przed wyświetleniem. Wrażenie obecności bardzo zależy od reakcji w rzeczywistym czasie i od poprawnej oceny głębi w stosunku do pozycji obserwatora. W tym projekcie omówione będzie trójwymiarowe uzyskiwanie dynamicznych części sceny w czasie rzeczywistym, którymi na rys. 1 są osoby na pierwszym planie. Postanowiono przyjąć zasadę uzyskiwania sceny niezależnego od punktu widzenia. Po uzyskaniu migawki sceny z odległego miejsca przesyła się ją jako przedstawioną z uwzględnieniem systemu współrzędnych świata. Wyświetlanie z nowego punktu widzenia wymaga tylko pierwotnych transformacji standardowo wpisanych w każdy procesor graficzny. Oprócz uzyskania rzeczywistego czasu, widok powinien być wolny od błędów, tak aby użytkownik nie odbierał niewłaściwej głębi przez swoje polaryzujące okulary stereoskopowe. Podstawowe pytanie, jakie się pojawia, to jak uzyskać jak najlepsze odwzorowanie w czasie rzeczywistym. Należy podkreślić, że te kryteria są ostrzejsze niż, na przykład, w przypadku nawigacji. Nawigacyjna stereoskopia celuje w odwzorowanie oparte na powłoce wypukłej podczas gdy użytkownik musi mieć możliwość widzenia takich szczegółów jak profil twarzy rekonstruowany z widoków frontowych.

Inne Prace dotyczące Uzyskiwania Obrazu

Istniejące prace (roczne bibliografie opracowywane przez Azriela Rosenfelda) na temat wszystkich aspektów stereoskopii. Użycie stereoskopii w wyświetlaniu opartym na obrazach jest bardzo szczegółowo omówione w niedawno opublikowanej pracy Narayanana i Kanade. Mimo że w wielu pracach używane są terminy “rzeczywistość wirtualizowana” i “powiększana rzeczywistość”, w tej pracy omawiany jest problem reakcyjnej teleobecności podczas gdy metody wyświetlania opartego na obrazach starają się wymienić statyczny model graficzny na rzeczywisty w trybie off-line.

Metody stereoskopowe można skwalifikować ze względu na dopasowanie lub na schemat rekonstrukcji. Jeśli chodzi o dopasowanie, rozróżniamy rekonstrukcje oparte na rzadkich cechach i rekonstrukcje zagęszczonej głębi. Niektóre metody omawiają prawdopodobną naturę dopasowania ze szczególnym naciskiem na problem okluzji. Metody oparte na polu bazują na korelacji i podkreślają reakcje w czasie rzeczywistym. Jest również metoda podkreślająca wirtualizowaną rzeczywistość. System ten przechwytyje ruch osoby przez zestaw 51 kamer. Zestawy kamer otaczających są także bardzo odpowiednie w technikach opartych na trójwymiarowych pikselach, takich jak wycinanie przestrzeni. Przetwarzanie odbywa się w trybie off-line i nie jest powiedziane jak można byłoby w ten sposób wykorzystywać je w teleobecności, poza off-linową rekonstrukcją struktur statycznych.

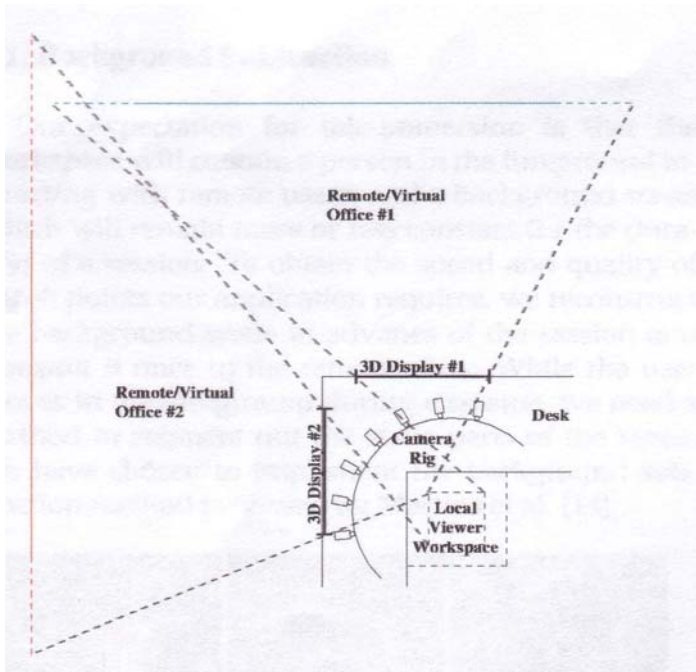
Jeśli chodzi o rekonstrukcję, ostatnie metody można byłoby sklasyfikować jako silne lub słabe (lub jako samo-kalibrowane). Metody samo-kalibracji dają rekonstrukcję metryczną z różnych punktów widzenia z dokładnością, jaka jest wykorzystywana jedynie do ograniczonych zastosowań powiększanej rzeczywistości, jak na przykład manipulacji wideo, w której jakość głębi nie ma znaczenia. Słabo skalibrowane metody dają wyświetlanie w czasie rzeczywistym i są odpowiednie do powiększania scen jedynie z obiektami syntetycznymi.

Ostatnio, dzięki pracom nad geometrią rzutową, stał się popularny paradygmat generacji widoku. Zamiast kalibrować kamery w stosunku do systemu współrzędnych świata, wyliczane są podstawowe matryce par obrazów. Następnie obrazy te są poprawiane i budowane są mapy różnic gęstości. Nowy widok można wyliczyć przez stworzenie nowych matryc w odniesieniu do każdego danego widoku. Naukowcy uważają, że poza kalibracją niewiele się zyskuje wykorzystując matryce podstawowe zamiast projekcyjnych. W obu podejściach istnieje znaczne opóźnienie w tworzeniu map różnic głębi. Aby zminimalizować problem odpowiedniości, stworzono metody zależne od punktu widzenia, które są kalibrowane, ale polegają na sylwetkach obiektów.

Budowa i ogólny opis systemu

Kabina tele-imersyjna zaprojektowana została po to, by uzyskać trójwymiarowy model lokalnego użytkownika i środowiska do wyświetlania i interakcji w odległych lokacjach oraz by zapewnić lokalnemu użytkownikowi wrażenie immersji przez śledzenie ruchów głowy i obrazy stereoskopowe wyświetlane na dużych ekranach.

Rys. 2 przedstawia typowe ustawienie. Użytkownik ma do dyspozycji powierzchnię 1 m swojej przestrzeni do pracy przy biurku, po której może się swobodnie poruszać. Użytkownicy z odległych lokacji wyświetlani są przez pary projektorów na ekranach o wymiarach 90x120 cm. Użytkownik ma założone lekkie okulary polaryzujące i urządzenie śledzące ruch głowy do kierowania funkcją wyświetlania stereoskopowego.



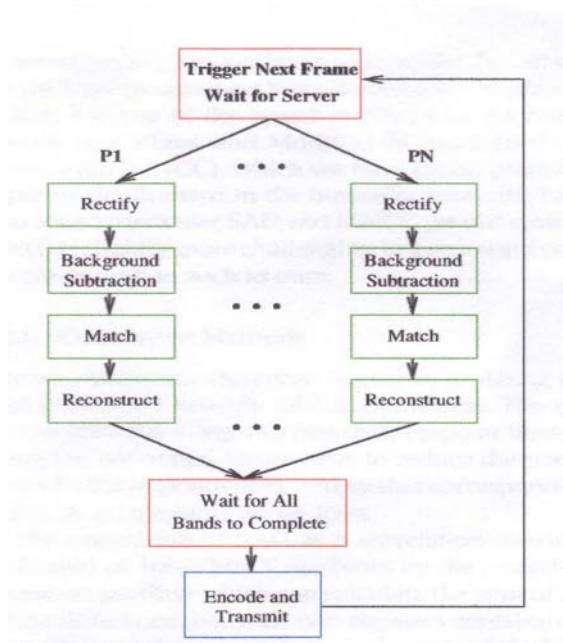
Rysunek 2. Układ kamer w tele-kabinie



Rysunek 3. Układ kamer, widok użytkownika.

Siedem kamer z interfejsem firewire ustawionych jest w łuku przy 15 stopniach oddzielenia, tak aby 'otoczyć' użytkownika i uniknąć przerwania obrazu, które mogłoby zostać spowodowane marginesem, na którym rekonstrukcja się zatrzymuje. W obecnym ustawieniu żadne z miejsc biorących udział w projekcie nie ma pełnej wersji tele-kabiny. Zamiast tego, jak przedstawia rys. 2, miejsce wyświetlania oraz miejsce uzyskiwania obrazu składa się z tylko jednego zestawu kamer, jak pokazuje rys. 3. Kamery te wykorzystywane są do obliczania dwu- lub

trójbiektowych stereoskopowych map głębi dla zachodzących na siebie par lub trójek. Na przykład, łączona trójbiektowa rekonstrukcja zilustrowana na rys. 10 została wyliczona z pięciu trójek $\{C_0, C_1, C_2\}$, $\{C_1, C_2, C_3\}$, $\{C_2, C_3, C_4\}$, $\{C_3, C_4, C_5\}$ i $\{C_4, C_5, C_6\}$. Przeszkodą techniczną w łączeniu obrazów z różnych kamer jest to, że każda rekonstrukcja dokonywana jest na innym komputerze. Ponieważ cyfrowy kabel firewire nie jest 'podzielny', trzeba przesyłać obrazy do sąsiadujących maszyn przez Ethernet/TCP/IP, co dodaje około 130 milisekund do ogólnego czasu obliczeniowego systemu. Dane szybkości reakcji i jakości głębi są niezmiernie ważne w aplikacjach imersyjnych. Aby poprawić częstotliwość klatkowania tego systemu, zastosowano kilka technik dla zredukowania ciężaru obliczeniowego, zwłaszcza w wymagającym zadaniu dopasowywania korelacji koniecznym do wygenerowania map gęstości głębi. Oczywiście, najprostszą metodą dla konstruktora byłoby kupić więcej szybszych komputerów. Ten system zbudowano na serwerach Pentium III 550 MHz (jeden na każdy zrekonstruowany widok) i odpowiednio wyrównano ten kod.



Rysunek 4. System równoległy.

Ogólna struktura równoległego systemu przedstawiona jest na rys. 4. Jeden z serwerów działa jako serwer wejściowy do wprowadzania danych przez interfejs firewire. Kiedy wszystkie rekonstrukcje są gotowe na następną klatkę, serwer wejściowy uwalnia wszystkie kamery jednocześnie. Każdy komputer otrzymuje obraz z jednej lub dwóch kamer oraz przesyła i otrzymuje obrazy potrzebne jemu i jego sąsiadowi. W obrębie każdej czwórki komputerów, obrazy są dzielone na

cztery równe pasma i każdy procesor opracowuje poszczególne pasmo. Wiązka dla każdego procesora przetwarza, oddziela tło, dopasowuje i rekonstruuje punkty w swoim paśmie obrazu. Kiedy wszystkie procesory kończą przeliczanie, mapy tekstur i głębi są przesyłane przez TCP/IP do odległego wyświetlacza. Dane te są dekodowane jako 3-(320x240) nieoznakowane płyty CHAR obrazów (RGB) tekstury, plus jeden krótki, nieoznakowany pląt obrazu, w którym wartości 1/z zostały przeskalowane na nieoznakowane krótkie, a piksele tła i niedopasowane piksele pierwszego planu są flagowane. Ogółem jest to 3 megabity na widok na klatkę.

Oddzielanie tła

Miejsce pracy w tele-imersji będzie się składać z osoby na pierwszym planie, która będzie współdziałać z oddalonymi użytkownikami oraz z tła, które będzie pozostawało mniej więcej stałe przez czas trwania sesji. Aby uzyskać szybkość i jakość punktów głębi, jakich wymaga taka aplikacja, rekonstruuje się scenę tła przed sesją, a potem raz przesyła się ją do oddalonych miejsc. Ponieważ użytkownik porusza się w czasie sesji, potrzebna jest metoda oddzielania statycznych części sceny. Wykorzystano metodę oddzielania tła zaproponowaną przez Martinsa i in.



Rysunek 5.

Sekwencja N (2 lub więcej) obrazów tła B_i uzyskiwana jest przed każdą sesją. Z tego zestawu oblicza się średni pikselowy obraz tła $B = 1/N \sum_i B_i$, a następnie obliczona jest różnica pikselowa pomiędzy B i B_i , $D = 1/N \sum_i (B - B_i)$.

Podczas sesji tele-imersji każdy główny obraz I jest oddzielany od statycznego średniego tła $I_D = B - I$, obraz binarny tworzony jest przez porównanie $I_B = I_D > T \times D$, gdzie T to próg konfigurowalny (zazwyczaj używamy $T = 7$). Te sprogowane obrazy różnic mają dość duże zakłócenia. Nad I_B dokonywana jest seria żłobień i rozszerzeń, aby wyostrzyć tło. Operacje morfologiczne wprowadzane są przez rozdzielne sploty z maskami jedynek rozmiaru W . W żłobieniu, jeśli piksel po

splocie ma wartość mniejszą niż $W \times W$, to ustawiany jest jako zero; w rozszerzaniu, jeśli piksel ma wartość większą niż zero, ustawiany jest na jedynkę. Sekwencja wygląda następująco: żłobić $W = 3$, rozszerzyć $W = 7$, żłobić $W = 17$, rozszerzyć $W = 9$. Typowe wyniki przedstawia rys. 5.

Dopasowywanie Dwu i Trójbiektywowe

Starając się zachować szybkość i jakość map zagęszczonej głębi stereoskopowej, zbadano dużą liczbę technik odpowiedniości korelacji. W szczególności skupiono się na sumie całkowitych różnic (SAD), z powodu prędkości, jaka wynika z poszczególnych operacji sprzętu oraz ze Zmodyfikowanej Znormalizowanej Korelacji Krzyżowej (MNCC), która, jak odkryto, tworzy bardzo dobre mapy głębi w przypadku dwubiektywowym. Testowano również trójbiektywowe SAD i MNCC z opisywanym systemem, co jest trudniejsze w konfiguracji kamer okrężnych.

Metody Korelacji

Algorytm rekonstrukcji zaczyna się od odebrania obrazów z dwóch lub trzech silnie skalibrowanych kamer. System przetwarza te obrazy, tak aby linie epipolarne leżały wzdłuż poziomych rzędów obrazów dla zredukowania przestrzeni szukania odpowiedniości i tak aby odpowiadające punkty leżały na tych samych liniach obrazów.

Wyliczanie SAD jako odległości korelacji jest wspomagane na maszynach Intel/MMX przez operację Assemblera psadbw, która może obliczyć sumę całkowitych różnic pomiędzy dwoma rejestrami zawierającymi po 8 CHAR wartości, w pojedynczym cyklu. Ogólnie wyliczanie SAD wygląda następująco:

$$\text{corr}_{SAD}(I_L, I_R) = \sum_W |I_L - I_R|$$

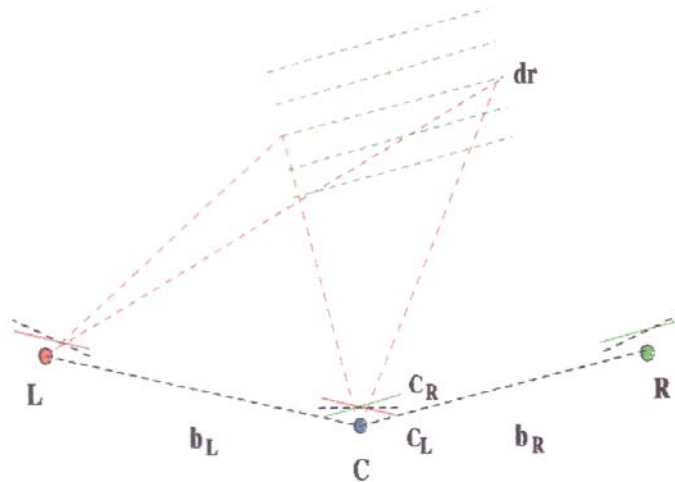
dla okna W w przetworzonych obrazach I_L i I_R . Różnica d określa względną pozycję okna w obrazie lewym i prawym. Lepszą metryką odpowiedniości jest zmodyfikowana znormalizowana korelacja krzyżowa (MNCC).

$$\text{corr}_{MNCC}(I_L, I_R) = \frac{2 \text{cov}(I_L, I_R)}{\sigma^2(I_L) + \sigma^2(I_R)}$$

gdzie I_L i I_R to lewy i prawy przetworzony obraz na wybranych oknach korelacji.

Dla każdego pikselu (u, v) w lewym obrazie, powyższe metryki tworzą profil korelacji $c(u, v, d)$, gdzie różnica d rozciąga się na akceptowalne wartości całkowite. Wybrane pary są maksimumami (MNCC) lub minimumami (SAD) w tym profilu.

Nierównoległe Konfiguracje Trójobiektywowe



Rysunek 6. Trójka trójobiektywnych kamer.

Trójobiektywowe zwężenie epipolarne jest dobrze znaną techniką weryfikacji odpowiedniości i polepszania jakości danych zasięgu stereoskopowego. Oparta jest na fakcie, że dla hipotetycznego dopasowania (pendant) $[u, v, d]$ w parze obrazów jest jednoznaczna lokacja, którą można przewidzieć w obrazie kamery trzeciej, co do której możemy oczekiwać, że znajdziemy w niej dane tego samego punktu świata [5]. Hipoteza jest właściwa, jeśli linie epipolarne dla pierwotnego punktu $[u, v]$ i domniemane dopasowanie (pendant) $[u-d, v]$ krzyżuje się w obrazie trzeciej kamery. Najczęściej używanym schematem do wykorzystania tego zwężenia jest ustawienie trójki kamer pod właściwym kątem, pozwalając na dopasowanie wzdłuż rzędów i kolumn obrazu wzorcowego. Konfiguracja telekabiny, zilustrowana na rys. 2, nie pozwala na ustawienie czy przetworzenie trójek płyt obrazów z kamer, tak by były współpłaszczyznowe, więc jest drożej wykorzystywać zwężenie trójobiektywowe.

Idąc za obserwacją Okutomi i Kanade'a zoptymalizowano sumę wartości korelacji w odniesieniu do faktycznej wartości głębi zamiast do różnicy. Traktuje się trójkę kamer $\{L, C, R\}$ jako dwie niezależne pary stereoskopowe $\{L, C_L\}$ i $\{C_R, R\}$. Rys. 6 ilustruje założenie ogólne: obojętnej wielkości różnica dla pary wzorcowej $\{C_R, R\}$ reprezentuje powierzchnię o stałej głębi (w odniesieniu do tej pary) w świecie, jednak dla lewej pary $\{L, C_L\}$ powierzchnia ta zawiera pewien przedział

odległości, czyli różnic. Istnieją dwa podejścia do wykorzystywania zwięzienia trójobiektywowego w określonej konfiguracji kamer. Pierwsza metoda, którą zastosowano, wstępnie obliczała obrazy korelacji dla rozpiętości różnic w lewej parze kamer, a następnie wyliczona korelacja dla każdego testowanego $[u_R, v_R, d_R]$ była dodawana do liczby wstępnie wyliczonej dla odpowiadającego $[u_L, v_L, d_L]$. Wynikiem tego są duże tablice funkcji korelacji dla lewej pary obrazów.

Drugą metodą była próba uniknięcia dużych tablic funkcji poprzez oddzielne odnajdowanie najlepszych N ekstremów na powierzchniach korelacji dla obu par obrazów. Te wydzielone hipotezy były następnie sprawdzane krzyżowo dla określenia czy wspólny punkt głębi powodował wzrost wyników jakiegokolwiek pary. Zachowywana była ważna para hipotez z najlepszym wynikiem. Metoda ta wymagała mniej przestrzeni dla tablic funkcji, miała jednak znaczne własne wymagania w stosunku do systemu, ponieważ trzeba było zachowywać wydzielone hipotezy.

Kiedy robiono przegląd opisywanego systemu w celu wyrównania go i polepszenia jego prędkości, odkryto, że przez używanie segmentacji pierwszego planu do rozpatrzenia jest tylko połowa jednej trzeciej pikseli obrazu wzorcowego C_R . Dzięki temu obliczenie całego profilu korelacji dla każdego pikselu oddzielnie staje się wykonalne. Aby obliczyć sumę wyników korelacji wstępnie oblicza się tablicę funkcji danej lokacji w C_L odpowiadającej obecnemu pikselowi w C_R (w oparciu o relację przetwarzania prawo-lewo). Oblicza się również bliskość liniową różnicy

$$d_L = M(u_{CR}, v_{CR}) \times d_R + b(u_{CR}, v_{CR})$$

w $[u_{CL}, v_{CL}]$, która wznosi się z tego samego punktu głębi, co $[u_{CR}, v_{CR}, d_R]$.

Podczas obliczania wyniku korelacji $\text{corr}_R(u_{CR}, v_{CR}, d_R)$ patrzymy na odpowiadające mu $[u_{CL}, v_{CL}]$ i obliczamy d_L , a potem wyliczamy wynik korelacji $\text{corr}_L(u_{CL}, v_{CL}, d_L)$. Wybieramy różnicę d_R która wynosi:

$$\text{corr}_T = \text{corr}_L(u_{CL}, v_{CL}, d_L) + \text{corr}_R(u_{CR}, v_{CR}, d_R)$$

Metodę tę można podsumować w następujący sposób:

Pikselowa Stereoscopia Trójobiektywowa

Pixelwise Trinocular Stereo

Step 1: Precompute lookup table for C_L locations corresponding to C_R locations, and approximation lookup tables M and b

Step 2: Acquire image triple $\langle L, C, R \rangle$

Step 3: Rectify $\langle L, C_L \rangle$ and $\langle C_R, R \rangle$ independently.

Step 4: Calculate foreground mask for C_R

Step 5: for every foreground pixel

Step I: $corr_{best} = INVALID,$
 $d_{best} = INVALID$

Step II: for every disparity $d_R \in D_r$

Step i: compute $corr_R(u_{C_R}, v_{C_R}, d_R)$

Step ii: lookup $[u_{C_L}, v_{C_L}]$

Step iii: compute $\hat{d}_L = M(u_{C_R}, v_{C_R}) \times$
 $d_R + b(u_{C_R}, v_{C_R})$

Step iv: compute $corr_L(u_{C_L}, v_{C_L}, \hat{d}_L)$

Step v: $corr_T = corr_L + corr_R$

Step vi: if $corr_T$ better than $corr_{best}$

$corr_{best} = corr_T$

$d_{best} = d_R$

Step 6: Goto 2

Osiągi i wyniki

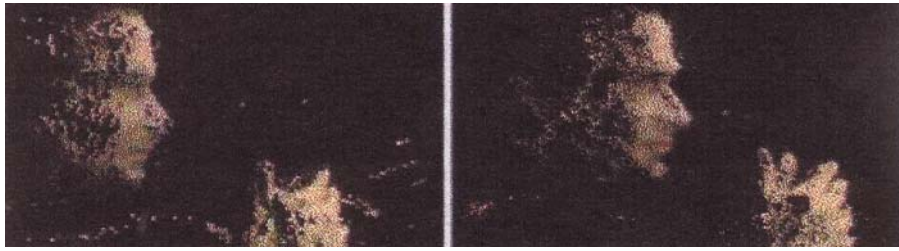
Metody wykorzystujące SAD były szybsze niż zastosowania oparte na wdrożeniu MNCC. Wszystkie aplikacje działały na serwerze zestawu PIII 550 MHz w czasie 1 sekundy lub mniej, włączając w to uzyskiwanie obrazu, transfer i przesyłanie rekonstrukcji do wyświetlacza. Czas trwania procesów w różnych systemach przedstawiony jest w tabeli 1.

Step	SAD	MNCC	Tri-SAD	Tri-MNCC
Rectify	49	50	49	48
Background	18	18	18	18
Matching	182	261	390	791
Reconstruct	6	6	7	6
Total	446 ms	520 ms	662 ms	1067 ms
fps	2.2	1.9	1.5	0.9

Tabela 1. Synchronizacja dla wdrażania online metod korelacji. Klatki na sekundę (fps – frames per second) zawierają wartość 160 ms wychwyconego czasu i ponad 6 ms transmisji (przekazu) sieciowej.



Rysunek 7. Trójobiektywowa trójka.



a.

b.

Rysunek 8. Odtworzenie rekonstrukcji, widok profilu. (a) dwuobiektywowa MNCC; (b) trójobiektywowa MNCC.



Rysunek 9. Widoki z siedmiu kamer.



Rysunek 10. Pięć trójobiektywnych rekonstrukcji połączonych i odtworzonych, obraz zmienia się kolejno.

Kolejną rzeczą, interesującą naukowców w tele-imersji, jest jakość i zagęszczenie punktów głębi. Mimo że czasy obliczeniowe były dłuższe, wysoka jakość trójobiektywnych map głębi sprawia, że są pożądaną alternatywą dla szybszych, ale bardziej zakłócanych obrazów zasięgu SAD. Rys. 7 ilustruje trójkę trójobiektywową, a rys. 8 wynikające z niej wyświetlane mapy głębi dla dwuobiektywowego MNCC (prawa para) rys. a i trójobiektywowego MNCC – rys. b. Polepszenie mapy głębi przez użycie zwężenia trójobiektywowego jest ewidentne w redukcji zakłóceń i udoskonalaniu szczegółów profilu.

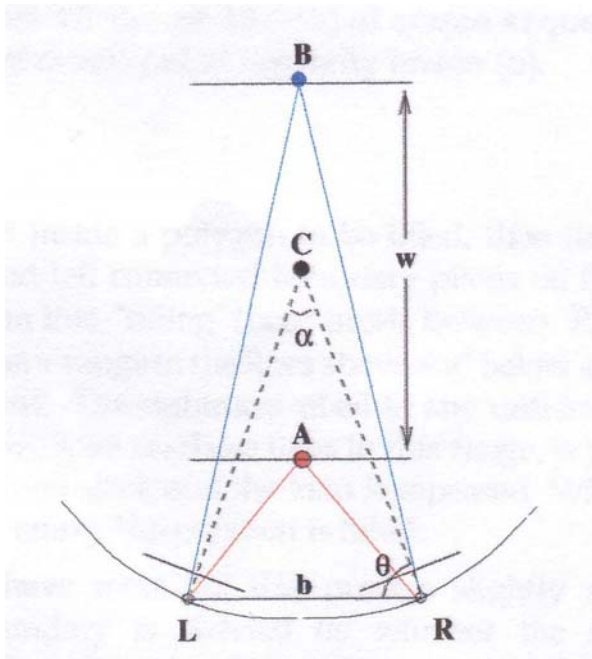
Dodatkowym wyzwaniem w zastosowaniu takiego zestawu siedmiu kamer jest połączenie wielu rekonstrukcji w pojedynczy obraz przy wyświetlaniu. Obecnie

należy polegać na poprawności kalibracji wspólnej klatki wzorcowej dla wszystkich kamer.

Rys. 9 pokazuje pełen zestaw obrazów z kamer dla pojedynczej klatki w obecnym zestawie kamer w tele-kabinie. Z tego zestawu obrazów oblicza się 5 zrekonstruowanych widoków z nachodzących na siebie trójek. Rys. 10 przedstawia obrót profilu pełnego zestawu 164 000 punktów głębi obliczonych przy użyciu trójobiektywowego MNCC dla klatki z rys. 9.

Powiększenia Oparte na Ruchu

Najbardziej kosztowną operacją w rekonstrukcji stereoskopowej jest samo dopasowywanie korelacji, ogólnie proporcjonalnej do $N \times M \times D$ dla obrazów o rozmiarach $N \times M$, gdzie D to testowane wartości różnic. Używając oddzielenia tła w takiej aplikacji zredukowano liczbę pikseli rozpatrywanych w procesie szukania do połowy, a nawet jednej trzeciej całego $N \times M$. Aby jeszcze bardziej zredukować koszt, należy zredukować D , czyli ilość różnic rozpatrywanych dla każdego pozostałego pikselu.



Rysunek 11. Zbliżony układ par stereo dla pracy wielkości w

Na rys.11. widać diagram pary nachylonych kamer. Jeśli chcielibyśmy uzyskać i podtrzymać niezakłócone wrażenie immersji, nie można zbyt ograniczać ruchów użytkownika w obrębie jego stereoskopowego miejsca pracy. Dla głębi miejsca pracy $w = 1$ m różnica może wynosić od $d = -61$ pikseli w punkcie A, 75 cm od kamer do $d = 87$ pikseli w punkcie B, 175 cm od kamer. Widać wyraźnie, że

rozpiętość różnic, wynosząca 148, nie zezwala na wyczerpujące wyszukiwanie odpowiedniości w aplikacjach w czasie rzeczywistym.

Dalsze obserwacje rekonstrukcji stereoskopowej w trybie online pokazują, że przy wysokich częstotliwościach klatkowania wystąpi znaczące podobieństwo pomiędzy następującymi po sobie obrazami. Można wykorzystać ten czasowy związek, aby następnie zoptymalizować obliczenia online. Propozycją jest prosta segmentacja obrazu, oparta na znajdowaniu regionów obrazu różnic, który zawiera tylko wąski przedział wartości różnicy. Używając obliczenia przepływu optycznego na region można ocenić lokację danego regionu w kolejnych klatkach i zawęzić wyszukiwanie różnicy D_i .

Uzupełniający charakter przepływu optycznego oraz kalkulacje stereoskopowe są dobrze znane. Odpowiedniość stereoskopowa zakłócana jest przez kombinatorykę przeszukiwania przedziału różnic oraz przez okluzję i niejednorodność powierzchni. Obliczanie struktury z ruchu jest zazwyczaj drugorzędne i wrażliwe na zakłócenia, a także nie jest w stanie ustanowić czynnika skali. Wykorzystanie czasowego związku pomiarów głębi i przepływu może przyjąć dwie formy: może być wykorzystane do poprawienia jakości lub poprawności wyliczonych wartości głębi i trójwymiarowego ruchu lub, jak w tym przypadku, jako środek do optymalizacji obliczeń w celu osiągnięcia działania w czasie rzeczywistym. Niewielkie jest prawdopodobieństwo, że metody, które wyliczają poprawne modele trójwymiarowe przy użyciu wielokrotnych technik, np. programowania liniowego, będą pomocne w takich aplikacjach w czasie rzeczywistym, jak powyższa. Inne proponowane metody dla robotów autonomicznych ograniczają lub w innym przypadku zależą od relatywnego ruchu, którego w przypadku swobodnie poruszającego się człowieka nie da się kontrolować.

Metodę łączenia podziału różnic i przepływu optycznego można podsumować w następujących krokach:

Krok 1: Dokonać ładowania początkowego przez wyliczenie pełnej mapy różnic dla pierwszej pary stereoskopowej w sekwencji.

Krok 2: Użyć opcji zalewania, aby podzielić mapę różnic na prostokątne okna zawierające wąski zakres różnic.

Krok 3: Obliczyć przepływ optyczny na okno dla prawych i lewych sekwencji wygładzonych i przetworzonych obrazów dla mających w danym momencie miejsce klatek.

Krok 4: Dopasować pozycję okna różnicy i zakresy różnicy w odniesieniu do szacowanego przepływu.

Krok 5: Przeszukać okna pod kątem odpowiedniości wykorzystując przypisany zakres różnic, wybierając "najlepsze" wartości korelacji ze wszystkich okien i różnice przypisane lokacji każdego pikselu.

Krok 6: Powrócić do Kroku 2.

Segmentacja zalewania (flood-fill)

Zawężanie zmian w różnicach na okno zasadniczo dzieli całą powierzchnię na kawałki, których głębia jest prawie równomierna. Na przykład obraz zakrzywionej powierzchni będzie podzielony na wiele sąsiadujących ze sobą okien, podobnie będzie z płaską powierzchnią oddalającą się od kamer pod wysokim kątem. Zasadniczo, okna te są małymi prawie czołowymi płaskimi fragmentami powierzchni.

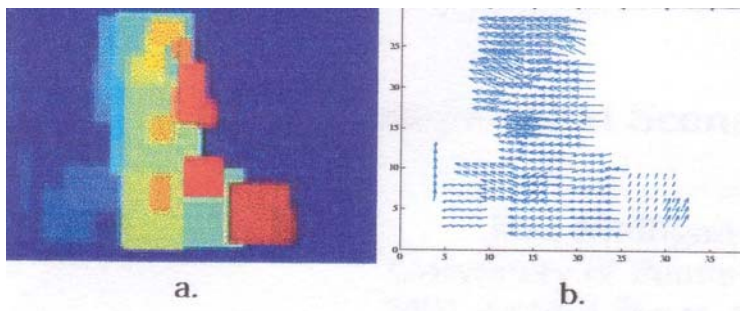
Używa się progu na maksymalną całkowitą rozbieżność różnicy jako regionów definiujących zawężenie i pozwala, żeby regiony na siebie zachodziły. Zachowuje się tylko prostokątne okna obrazu, żeby nie otrzymać wypukłej powłoki lub bardziej skomplikowanej struktury, ponieważ ogólnie szybciej jest zastosować jakieś operacje na większym prostokątnym oknie niż na bardziej skomplikowanej strukturze. Regiony wybiera się przy użyciu zalewania lub wypełniania ziarnistego prostego algorytmu wypełniania wielokątów z grafiki komputerowej. Zastosowano wersję linii skanowania, która wybiera lokację pikselu wewnątrz wielokąta, która ma zostać wypełniona, a następnie odnajduje połączone lewe i prawe piksele graniczne na danej linii skanowania, 'wypełniając' piksele pomiędzy nimi. Następnie bada się piksele w tym samym zasięgu x na liniach powyżej i poniżej. Piksel wysunięty najbardziej na prawo w jakimkolwiek niewypełnionym niegranicznym polu na tych liniach w tym zasięgu zostaje wypchnięty na stos i pętla się powtarza. Kiedy kupka jest pusta, wielokąt zostaje wypełniony.

Nieznacznie zmodyfikowano ten proces, aby granica była wyznaczana przez to czy wartość różnicy obecnego pikselu mieści się w progu (± 5) pierwszego pikselu zaziarnionego. Zaczyna się od maski ważnych lokacji różnic w obrazie

różnic. W tej pracy wypełnianie oznacza zaznaczanie lokacji na masce, które zostały włączone do jakiegoś regionu różnicy i uaktualnianie lewej górnej i prawej dolnej współrzędnej pikselu obecnej ramki ograniczającej okno. Kiedy nie ma już pikseli przylegających do obecnego regionu, które są w granicach różnicy pierwotnego ziarna, następny niewypełniony piksel maski jest używany do zaziarniania nowego okna. Kiedy wszystkie lokacje pikseli na masce są ustalone, segmentacja jest zakończona.



Rysunek 12. Klatka 12(lewa) sekwencji stereo (a), i obliczona różnica obrazu(b).



Rysunek 13. Wydobycie okien podobnych różnic (a) i obliczona odkształcenia na okno (b).

Na koniec małe regiony uznaje się za zakłócenia i kasuje się je. Zbliżone lub nachodzące okna są łączone jeśli średnie różnice regionów są podobne. Rys. 12 i 13 przedstawiają obraz, jego mapę różnic oraz prostokątne regiony uzyskane poprzez zalanie.

Przeptyw na Okno.

Obliczenia przepływu optycznego przybliżają pole ruchu obiektów poruszających się w stosunku do kamer, w oparciu o znane równanie stałości jasności obrazu:

$I_x v_x + I_y v_y + I_t = 0$, gdzie I to jasność obrazu, a I_x , I_y oraz I_t to częściowe pochodne I w odniesieniu do x , y , t , a $v = [v_x, v_y]$ to szybkość obrazu. Użyto standardowego

kwadratowego algorytmu przyznawania najmniejszych wartości do wyliczania wartości v w oparciu o zmniejszanie

$$e = \sum_{W_i} (I_x v_x + I_y v_y + I_t)^2$$

dla pikseli w obecnym oknie W_i . Dla każdego okna różnicy przyjmuje się, że pole ruchu jest stałe w całym regionie W_i i oblicza się pojedynczą wartość dla pikselu centralnego. Rys 13 b ilustruje wartości przepływu optycznego obliczanego na okno.

Gdy są dane regiony obrazu, należy dostosować ich lokalizację do przyjętego przepływu dla obrazu prawego i lewego. Ogólnie trzeba zmusić to okno do rozszerzenia zamiast je przesuwać, jeśli przewidujemy, że lewa współrzędna okna zostanie przesunięta do góry lub w lewo przez przepływ prawy lub lewy. Jeśli przewidujemy, że prawa współrzędna przesunie się w dół lub w prawo, odpowiednio się ją powiększa.

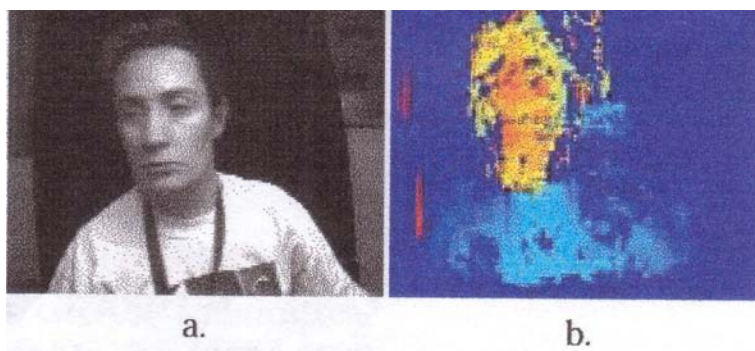
Ze względu na to, że okna się przesunęły z powodu obiektów poruszających się w głębi, trzeba także dostosować zakres różnic $D(t)=[d_{min}, d_{max}]$ dla każdego okna wykorzystując ocenę prędkości przepływu:

$$D(t + dt) = [\min(d_{min} + v_{xl}dt - v_{xr}dt, d_{min}), \max(d_{max} + v_{xl}dt - v_{xr}dt, d_{max})].$$

Odpowiedniość Okien.

Odpowiedniość w oparciu o okna zachowuje się tak samo, jak opisane to zostało powyżej. W przypadku takich okien różnic każde okno może mieć dowolny rozmiar, ale będzie mieć relatywnie mniej różnic do sprawdzenia. Ze względu na to, że takie obrazy są przetwarzane, aby można było wyrównać linie epipolarne w stosunku do linii skanowania, okna będą miały te same współrzędne y w obrazach prawych i lewych. Gdy dany jest zakres różnic, można wyciągnąć żądane okno z prawego obrazu przy

$x_r = x_l - d$. Dopasowywanie korelacji i przypisywanie ważnych par mapom różnic przebiega według zwykłej procedury.



Rysunek 14.

Rys. 14 przedstawia wyniki przeniesienia okna różnicy z klatki 12 do 18 w sekwencji obrazu stereoskopowego. Podmiotem jest odwzorowanie i przesuwanie obrazu od prawej do lewej. Szczegółowo zbadano wymianę pomiędzy jakością i poziomem skomplikowania w tej metodzie. Poziom skomplikowania odpowiedniości stereoskopowej w proponowanym systemie okien jest mniej więcej połową poziomu w pełnych obrazach, w zależności od ilości klatek w czasie używanych do oceniania przepływu optycznego. Pokazano eksperymentalnie, że takie rekonstrukcje oparte na oknach wypadają korzystnie w porównaniu z rekonstrukcjami generowanymi przez korelację na pełnym obrazie, nawet po kilku klatkach przesunięcia przez oceniany przepływ optyczny. Zaobserwowane średnie odchylenia w obliczonych różnicach wynosiły mniej niż 1 piksel, a maksymalne standardowe odchylenie wynosiło 4.4 piksele.

Podsumowanie

Przedstawiono część uzyskiwania sceny pierwszego testu tele-imersji wizualnej. Wykorzystano metodę niezależną od punktu widzenia i stworzono trójwymiarowy opis w odniesieniu do systemu współrzędnych świata. Uzyskano okrężny widok przy przepustowości 2 klatki/s. W przyszłości można będzie osiągnąć 10 klatek/s bez poświęcania jakości głębi. Aby wyeliminować wartości wyizolowane, w najbliższej przyszłości należy zająć się problemami rozjaśniania i okluzji. Ostatecznym celem jest wykonywanie systematycznych eksperymentów i studiowanie wrażenia obecności oraz niuansów komunikacji podczas współpracy.

“View-independent Scene Acquisition for Tele-Presence”

Zakończenie

Teleimersję definiuje się jako „poczucie częściowej obecności” z oddalonymi osobami i ich środowiskami, które czują się tak, jakby były w jednym

pomieszczeniu. Jeden z pierwszych gości w systemie teleimersji zauważył: „ To jest tak, jakby ktoś wziął piłę łańcuchową i wyciął dziurę w ścianie i widać następny pokój”. Ten rodzaj teleimersji różni się znacznie od konwencjonalnej wideo konferencji, w teleimersji obraz użytkownika odległego środowiska zmienia się dynamicznie jak on porusza głową.

To również różni się od teleimersji oglądanej jako sieciowych współpracujących środowisk wirtualnej rzeczywistości, która używa awatarów do przedstawienia współpracujących ze sobą osób i nawet nie próbują się „wtopić” w prawdziwe i syntetyczne obiekty. Takie podejście jest nastawieniem się na zgłębienie abstrakcyjnych danych; nasze wyobrażenie (wizja), zamiast, realistycznie rozszerzonej naszej własnej fizycznej przestrzeni, stanowi to wyzwanie dla tworzenia próbek środowiska, transmisji, rekonstrukcji, prezentacji, i wzajemnego oddziaływania(interakcji) użytkowników. Inne podejście skupia się na realistycznym przedstawieniu uczestników w telekonferencji nie stosuje się wiadomości na temat rozległego lokalnego środowiska niezbędnego do utrzymania jednolitego (bez szwów) połączenia prawdziwego i syntetycznego środowiska.

Zastosowania dla tej formy teleimersji będą zawierały immersyjne książki elektroniczne, które w praktyce połączą „wehikuł czasu „ z trójwymiarowym hipermedium, powiększając o dodatkowy wymiar, będzie można nagrywać doświadczenia, w których użytkownik jest „zatopiony” w trójwymiarowej rekonstrukcji, może on dosłownie przechadzać się w danej scenie(miejscu) lub poruszać się do tyłu, na przód. Chociaż jest wiele możliwości zastosowań w dziedzinach takich jak nowatorskie technologie (np. projektowanie i wirtualne tworzenie pierwowzorów, konserwacja i naprawa, rekonstrukcje paleontologiczne i archeologiczne), nastawienie będzie na przydatność społeczną i wyzwanie technologiczne wiodących zastosowań, uczenie się trudnych chirurgicznych umiejętności, potencjalnie śmiertelnych obrazów.

Jak holograficzne środowiska będą pracować

W filmie „Star Trek”, można było zobaczyć załogę statku kosmicznego „Enterprise”, która mieszkała poza miejscem własnej wyobraźni w pokoju nazywanym „holodeck” (pokład holograficzny). Ten pokład był olbrzymim

holograficznym pokojem projekcji, który pozwalał załodze oddziaływać na siebie wzajemnie z projekcjami tak, jakby byli w wielkiej grze wideo.

Dziś naukowcy rozwijają nowe technologie komunikacji, które pozwolą użytkownikom na wzajemne oddziaływanie (interakcję) i komunikację wewnątrz symulowanego środowiska, nawet jeśli są tysiące kilometrów oddaleni od siebie.



Teleimersja pozwoli nam manipulować holograficznymi obrazami, tak, jakby były one prawdziwymi obiektami.

Większość z podstawowych komponentów dla tej sieci jest w miejscu pozwalającym na rozwój teleimersji. Teleimersja jest odpowiedzią społeczności naukowej na platformę holograficzną z filmu. Poprzez połączenie kamer i technologii telefonicznej, wideokonferencje umożliwiają wymianę w czasie rzeczywistym większą liczbę informacji niż kiedykolwiek było to możliwe bez obecności fizycznej poszczególnych osób w jednym pokoju (centrali). Teleimersja wyniosła wideokonferencję na wyższy poziom. Stworzyła „centralę”, czyli symulowane środowisko, które pozwala wszystkim spotkać się razem w jednym wirtualnym pokoju i nikt nie będzie musiał opuszczać swojego fizycznego położenia, żeby to zrobić.

Poza wideo konferencją

W biurze przyszłości nie trzeba używać telefonu, żeby się do kogoś dodzwonić. Trzeba tylko rozkazać komputerowi, żeby zadzwonił do poszczególnych osób: „Komputerze, zadzwoń do biura Joe w Londynie”. Wtedy jedna ze ścian biura

zabłyśnie i Joe będzie siedział naprzeciwko nas, tak jakby był tam w rzeczywistości. Z drugiej strony, Joe doświadczy tego samego immersyjnego połączenia. Taka jest idea teleimersji: przenieść dwie lub więcej osób odległych od siebie, razem do jednego symulowanego otoczenia biura. Podróże służbowe będą praktycznie niepotrzebne. Ten typ komunikacji zabierze nas poza ograniczenia wideo konferencji. Do telekonferencji z inną osobą służy Webcom. Taki rodzaj wideo konferencji przez internet nie jest doskonałą (perfekcyjną formą komunikacji). Obraz jest zbliżony do czasu rzeczywistego, ale występują tam opóźnienia, które powodują nagłe ruchy obrazu wideo. Jeśli ktoś wyjdzie z pola widzenia jednej kamery nie można tej osoby dłużej widzieć. Podczas sesji teleimersyjnej ludzie nie będą mogli wyjść z pola widzenia kamery. Właściwie, mogą się rozglądać po swoim biurze przez patrzenie na wyświetlany ekran pod różnym kątem. To wygląda tak, jakbyśmy mieli okno w ścianie i wyglądali przez nie. Środowiska holograficzne będą generowane przez komputery z szybkością obliczeniową tysiąc razy szybszą od dzisiejszych PC.

Kroki służące do komunikacji holograficznego środowiska:

1. Komputer rozpoznaje obecność i ruchy ludzi i przedmiotów.
2. Komputer śledzi te obrazy.
3. Komputer wyświetla obrazy na stereoimmersyjnych powierzchniach.

Pracuje nad tym kilka grup: NTII(National Teleimersion Initiative) wprowadzają nas w wirtualne środowiska.

W maju 200 roku, naukowcy z Uniwersytetu Pn. Karoliny, Uniwersytetu w Pensylwanii i Zaawansowane Sieci i Usługi przekroczyli kamień milowy w rozwoju tego rodzaju technologii. Urzytkownik w biurze na UNC w Chapel Hill, NC, był zdolny zobaczyć trójwymiarowe jak żywe obrazy kolegów, oddalonych o setki mil, jeden w Filadelfii i drugi z nich w Nowym Jorku.

Teleimersja tylko brzmi podobnie do wirtualnej rzeczywistości, są pewne kluczowe różnice między tymi dwoma technologiami. Wirtualna rzeczywistość pozwala użytkownikowi zmieniać sceny, które widzimy. Możliwości aplikacji w immersyjnym holograficznym środowisku są nieskończenie wielkie. Gry wideo nie będą wymagały joystick'ów, będzie można stać się uczestnikiem gry, walczyć z potworami i zdobywać punkty przy lądowaniu. Zamiast podróżować tysiące mil, aby zobaczyć krewnych będziemy zdolni uścisnąć się z innymi ludźmi używającymi

tej samej technologii. Lekarze i żołnierze będą w stanie używać teleimersji do szkoleń w symulowanym środowisku

Budowanie holograficznego środowiska

Wczesne prototypy wyświetlaczy teleimersyjnych wymagają od użytkownika noszenia specjalnych okularów i urządzenia na głowie, które śledzi punkty obrazu, na które użytkownicy patrzą. Z drugiej strony, ludzie pojawiający się jako trójwymiarowe obrazy są śledzeni układem siedmiu zwyczajnych wideo kamer, podczas gdy dwie inne kamery wideo wychwytyują rzeczywisty model projektowanego(wyświetlanego) światła w poszczególnych pokojach w celu obliczenia odległości. Pozwala to na właściwą głębię do przetworzenia na wyświetlacz(ekran). W ten sposób oglądający poruszając głową w prawo, może zobaczyć "korespondencyjnie" obrazy, które byłyby widziane jeśli byłby on aktualnie w pokoju z osobą na ekranie. Obrazy na ekranie są rozdzielane i polaryzowane do stworzenia różnych obrazów dla każdego poszczególnego oka. Następnie okulary łączą te obrazy tak, aby mózg rozpoznał tylko jeden obraz trójwymiarowy. Ten proces jest podobny do starego sposobu działania okularów, służących do oglądania trójwymiarowych filmów. Wczesne eksperymenty jak ten na UNC w maju były podobne do normalnej wideo konferencji. Wyświetlacze sceny były odświeżane tylko trzy razy na sekundę, wytwarza to „poszarpany” obraz. Jeżeli ta szybkość mogłaby być udoskonalona do dziesięciu klatek na sekundę, stworzono by wyświetlany obraz bez „szwów”, to będzie jak patrzenie przez okno na inną osobę. Naukowcy rozwijają nowe technologie w celu utrzymania tego typu komunikacji, wliczając w to:

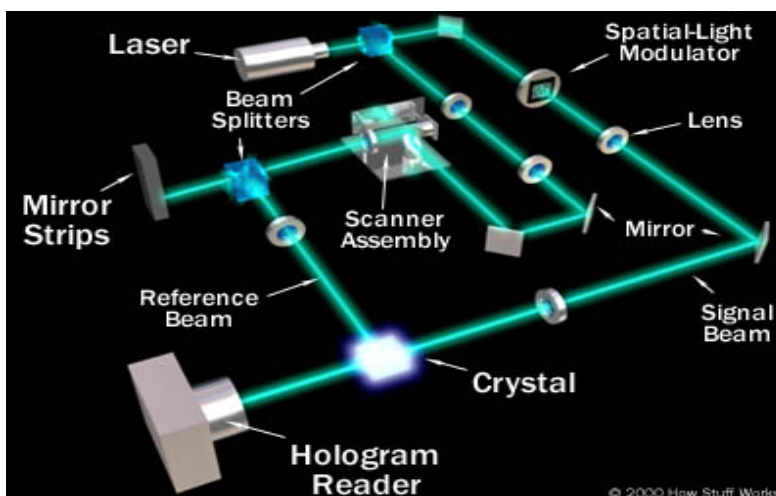
- Internet2 – zastąpi obecną infrastrukturę Internetu. Ta nowa sieć będzie miała wyższą częstotliwość (szerokość pasma) i szybkość tysiąc razy większą od dzisiejszego Internetu. Wysoka częstotliwość i szybkość sieci jest potrzebna do przekazywania ogromnej liczby danych, którą będzie produkować teleimersja.
- Technologie wyświetlania obrazu – stereoimersyjny wyświetlacz będzie miał przedstawiać czysty obraz transmitowanych scen.
- Czujniki haptyczne – będą pozwalały ludziom dotykać obrazy wyświetlane tak jakby były realne.

- Biurowe superkomputery będą wykorzystywać tryliony obliczeń potrzebnych do działania środowiska holograficznego. Inną możliwością do utrzymania tych środowisk będzie sieć komputerów, która podzieli możliwości obliczeniowe.

Teleimersja zatreze różnicę pomiędzy rzeczywistymi i wygenerowanymi przez komputer obrazami. To będzie fundamentalna technologia telekomunikacyjna, prawie całkowicie wyeliminuje jazdę do pracy w godzinach szczytu. Zamiast tradycyjnych sposobów komunikacji, ludzie będą mogli uczestniczyć w scenie spotkań przez wyświetlenie się w pokoju konferencyjnym firmy. Jeśli praca wymagać będzie podróży służbowych, będzie można wciąż być w domu na obiad przez teleimersję, wyświetlenie się w rodzinnej kuchni. Ponieważ ta technologia jest wciąż we wczesnym stadium rozwoju, jej możliwości są rzeczywiście nieograniczone.

Jak będzie działała pamięć holograficzna

Urządzenia, które używają światła do zapamiętywania i czytania danych były głównym sposobem gromadzenia danych przez blisko dwa dziesięciolecia. Dyski kompaktowe zrewolucjonizowały gromadzenie danych we wczesnych latach 80-tych, pozwalają zapamiętywać multi-megabajty danych na dysku. Dyski kompaktowe posiadają średnicę jedynie 12 centymetrów i grubość około 1,2 milimetra. W 1997 roku udoskonalono wersję dysku kompaktowego, nazwano ją „digital versatile disc” (DVD) – uniwersalny dysk cyfrowy. W sprzedaży były takie DVD, które pozwalały gromadzić (zmieścić) pełnej długości filmy na jednym dysku.



W urządzeniu pamięci holograficznej, wiązka lasera dzieli się na dwie wiązki, te dwie wiązki oddziałują w kryształach – w medium, gdzie zapamiętywane są przetworzone hologramy danych.

Płyty CD i DVD są podstawową metodą gromadzenia takich danych jak: muzyka, oprogramowanie, video itp. Płyta CD może zmieścić 783 megabajtów danych, co jest równoważne jednej godzinie i piętnastu minutom muzyki, ale Sony planował wykonać 1,3 gigabajta (GB), dużej pojemności CD. Dwustronne, dwuwarstwowe DVD mogą pomieścić 15,9 GB danych, to jest około ośmiogodzinny film. Te konwencjonalne środki do przechowywania danych odpowiadają dzisiejszej potrzebie gromadzenia danych, ale technologie przechowywania muszą rozwijać się, aby dotrzymać kroku wzrastającym wymaganiom konsumenta. Wszystkie sposoby przechowywania CD, DVD i magnetyczne zapamiętują bity informacji na powierzchni danego środka nagrywającego. W następstwie wzrastających możliwości gromadzenia danych naukowcy pracują obecnie nad metodą optyczną, nazywają ją pamięcią holograficzną, która będzie działać „poniżej” powierzchni i używać pojemności danego środka nagrywającego, służącego do gromadzenia danych. Zamiast tylko używać powierzchni, trójwymiarowe sposoby zapamiętywania danych będą zdolne do zapamiętania większej ilości informacji w mniejszej przestrzeni i oferują szybszy czas przesyłu (transferu) danych.

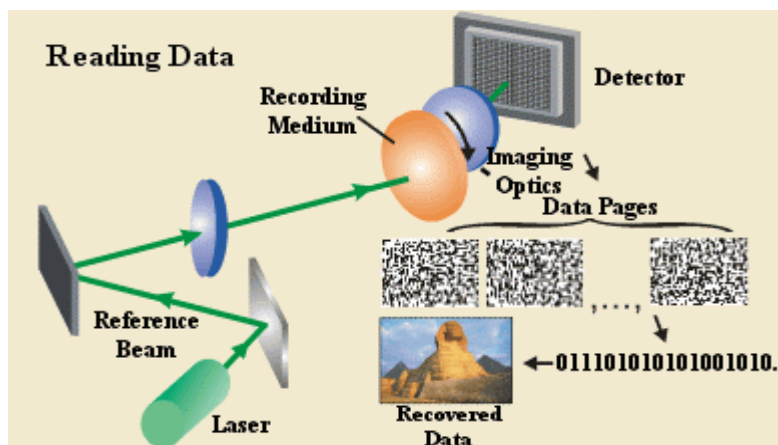
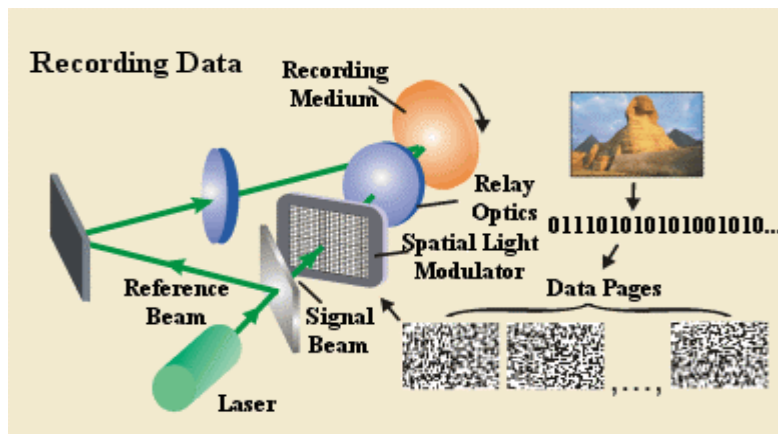
Podstawy

Pamięć holograficzna oferuje możliwość zapamiętania 1 terabajta (TB) danych w kryształach cukru rozmiaru sześciangu. Terabajt danych równa się tysiącowi gigabajtów, 1 milionowi megabajtów lub 1 trylionowi bajtów. Dane z CD większych niż 1 tysiąc mogą odpowiadać holograficznym systemom pamięci. Większość z komputerowych twardych dysków może pomieścić tylko 10-40 GB danych, a to tylko mała część tego, co mógłby pomieścić system pamięci holograficznej. Naukowiec z Polaroida Pieter J. Van Heerden pierwszy zaproponował pomysł holograficznego (trójwymiarowego) sposobu zapamiętywania danych we wczesnych latach 60-tych. Dziesięć lat później naukowcy z laboratoriów RCA zademonstrowali technologię przez wgrzywanie 500 hologramów do kryształu niobianu litu z domieszką żelaza i 550 hologramów w obrazach o wysokiej rozdzielczości we wrażliwym na światło polimerze. Brak tanich części i postępu w magnetycznych i półprzewodnikowych pamięciach zatrzymał w miejscu rozwój holograficznego gromadzenia danych. Przez ostatnie 10 lat, DARPA (Defense

Advanced Research Project Agency – Agencja Obrony Zaawansowanych Projektów Badań) i wysoko zaawansowanych technologicznie gigantów jak IBM i Lucent's Bell Labs zapoczątkowali na nowo rozwój holograficznej pamięci. Prototypy rozszerzone przez Lucent i IBM różniły się nieznacznie, ale większość z (HDSS – Systemy holograficznego gromadzenia danych) bazują na tej samej koncepcji. Podstawowe komponenty potrzebne do HDSS:

- Niebiesko-zielony argonowy laser
- Urządzenie służące do rozszczepienia promieni lasera
- Lustro do skierowania (ukierunkowania promieni lasera)
- Panel LCD (przestrzenny modulator światła)
- Okulary do skupienia(wyostrzenia) promieni lasera
- Kryształ niobianu litu lub fotopolimer
- Kamera CCD (Charge-coupled device – element ze sprzężeniem ładunkowym)

Kiedy niebiesko-zielony argonowy laser jest zapalony, urządzenie rozpraszające tworzy dwa promienie lasera. Jeden, nazywany jest przedmiotem lub promieniem znakomitym(nadzwyczajnym), który przechodzi prosto, odbija się od lustro i dalej przebiega przez przestrzenny modulator światła(SLM). SLM jest wyświetlaczem ciekłokrystalicznym (LCD –Liquid Crystal Display), pokazuje on strony surowych(nie obrobionych) danych binarnych jako jasne i ciemne miejsca(skrzynki). Informacje ze stron kodu binarnego są noszone przez promień nadzwyczajny dookoła wrażliwego na światło kryształu niobianu litu. Niektóre systemy używają fotopolimera w miejsce kryształu. Drugi promień, nazywany promieniem skierowanym, wystrzelony jest na zewnątrz urządzenia do rozpraszania promieni i odbiera on oddzielną ścieżkę do kryształu. Promień lasera jest rozszczepiany na dwie wiązki. Kiedy dwa promienie spotkają się, obraz interferencyjny, który jest stworzony zapamiętuje dane przenoszone przez promień nadzwyczajny w określony obszar w kryształce – dane są zapamiętywane jako hologram.



Te dwa diagramy pokazują, jak informacje są zapamiętywane i odzyskiwane przez system holograficznego zapamiętywania danych.

Zaletą holograficznego systemu zapamiętywania jest kompletna strona danych, które mogą być szybko odbierane w określonym czasie. Następnie po odebraniu i rekonstrukcji holograficznych stron zapamiętanych danych w kryształach, skierowany promień jest zaświecany w kryształach pod dokładnie tym samym kątem, pod którym wychodził do zapamiętania tej strony danych. Każda ze stron danych jest zapamiętywana w różnych innych obszarach kryształu, bazując na kącie pod którym promień skierowany przebywa swoją drogę. Podczas rekonstrukcji, promień będzie ugiął się (ulegał dyfrakcji) przechodząc przez kryształ, pozwala to odtworzyć oryginalną stronę danych, która była zapamiętana na początku. Ta zrekonstruowana strona jest wtedy wyświetlana na kamerę CCD, która interpretuje i przesyła cyfrową informację do komputera. Kluczowym komponentem holograficznego systemu zapamiętywania danych jest kąt pod którym drugi (skierowany promień) jest zapalany w kryształach w celu odebrania danych. On musi dopasować dokładnie początkowy kąt skierowanego promienia. Różnica tylko jednej tysięcznej milimetra jest błędem i odebranie strony danych zakończy się niepowodzeniem.

Biurowe holograficzne przechowywanie danych

Po więcej niż trzydziestu latach badań i postępu, biurowy holograficzny system przechowywania (HDSS) jest na wyciągnięcie ręki. Są jeszcze wciąż problemy z dostrojeniem, które muszą być rozwiązane przed tym jak urządzenie do „bardziej ścisłego” przechowywania będzie do nabycia na rynku, ale naukowcy z IBM sugerują, że będą mieli gotowe urządzenie HDSS nie wcześniej niż w 2003 roku. Te wczesne holograficzne urządzenia do przechowywania danych będą miały pojemność 125 GB i szybkość transferu danych około 40 GB na sekundę – wystarczająco szybko, żeby przesłać cały film DVD w 30 sekund. Dlaczego więc rozwój HDSS trwa tak długo?, i co jeszcze zostało do zrobienia? Kiedy pomysł HDSS zaproponowano pierwszy raz, komponenty do skonstruowania takiego urządzenia były dużo większe i bardziej kosztowne. Na przykład laser do takiego systemu w latach 60-tych miałby 6 stóp długości. Obecnie, wraz z rozwojem elektroniki dla konsumenta(odbiorecy) laser jest podobny do tego, jakiego używa się w odtwarzaczach płyt CD, taki sam mógłby być używany w HDSS. Wyświetlacze LCD nie były znane do 1968 roku, i pierwszy był bardzo drogi. Obecnie, wyświetlacz LCD jest dużo tańszy i bardziej złożony(skomplikowany), od tego, który był rozpowszechniony 30 lat temu. Dodatkowo czujniki CCD nie były osiągalne w ostatnim dziesięcioleciu. Prawie całe urządzenie HDSS może być zrobione z elementów kupionych w sklepie, co oznacza, że mogłoby być ono produkowane masowo. Mimo tego, że części do HDSS są bardziej dostępne dzisiaj, niż było to możliwe w latach 60-tych, istnieją pewne problemy techniczne, które muszą być rozwiązane. Na przykład, jeśli zbyt dużo stron będzie przechowywanych w jednym kryształ, wytrzymałość poszczególnych hologramów jest zmniejszona. Jeżeli jest za dużo hologramów przechowywanych w kryształ laser skierowany zazwyczaj odbiera hologram, który nie jest zapalany pod odpowiednim kątem. Hologram wyłapuje za dużo tła z innych hologramów przechowywanych dookoła niego. To jest właśnie wyzwanie do ustawienia w szeregu (w linii prostej) wszystkich komponentów w mało kosztującym systemie. Naukowcy są pewni, że ta technologia będzie się rozwijała w następnych dwóch lub trzech latach, aby sprostać wyzwaniom. Z takimi technologiami na rynku będziemy mogli nabyć wkrótce pierwszy holograficzny odtwarzacz pamięci. Ten podobny do DVD dysk będzie miał pojemność 27 razy większą niż 4.7 GB DVD osiągalne obecnie, i urządzenie

odtworzone będzie posiadało szybkość odtwarzania danych 25 razy większą od dzisiejszych najszybszych odtwarzaczy DVD.

Jak rozszerzona rzeczywistość będzie działać

Gry wideo zabawiały nas przez blisko 30 lat, już we wczesnych latach 70-tych. Grafika komputerowa stała się dużo bardziej wyszukana od tego czasu, i wkrótce grafika gier komputerowych okazywała się ostatecznie zbyt rzeczywista. Przez następne 10 lat naukowcy planowali „wyciągnąć” grafikę z ekranów telewizji lub wyświetlaczy komputerowych i włączono je do środowisk światów rzeczywistych. Ta nowa technologia, nazywana rozszerzoną rzeczywistością, będzie dodatkowo zacierało linię pomiędzy tym co rzeczywiste i tym co jest wygenerowane przez komputer, poprzez wzmocnienie tego, co widzimy, słuch, węch, dotyk.



Wyświetlona na monitorze rozszerzona rzeczywistość będzie pokrywać się z grafiką wygenerowaną przez komputer w realnym świecie.

Porównując rzeczywistość wirtualną, która tworzy immersyjne wygenerowane przez komputer środowiska, i świat rzeczywisty, rozszerzona rzeczywistość jest bliższa rzeczywistości. Rozszerzona rzeczywistość dodaje grafikę, dźwięk, dotyk, i zapach do naturalnego świata, tak jak on egzystuje. Możemy spodziewać się, że gry wideo doprowadzą do rozwoju rozszerzonej rzeczywistości, ponieważ ta technologia będzie posiadać niezliczone zastosowania. Wszyscy począwszy od turystów do szwadronu wojska odniosą korzyści ze zdolności do umiejscowienia wygenerowanych przez komputer grafik w swojej dziedzinie(zakresie

zainteresowań). Rozszerzona rzeczywistość zmieni prawdziwie sposób patrzenia na świat. Obraz samych siebie chodzących albo jadących po ulicy. Z wyświetlaczami rozszerzonej rzeczywistości, które wyglądają jak zwykła para okularów, informacyjne grafiki będą pojawiać się w polu widzenia użytkownika, również dźwięk będzie się zgadzał z tym co zobaczy użytkownik. Te wzbogacenia będą stale odświeżane, odzwierciedlając ruchy głowy użytkownika.

Rozszerzając nasz świat

Podstawową zasadą działania rozszerzonej rzeczywistości są nakładające się na siebie grafiki, dźwięk, inne zmysły ponadto środowisko świata realnego w czasie rzeczywistym. Telewizja wyświetlała takie grafiki przez dziesięciolecia, ale wyświetlała grafiki statyczne, które nie dostosowują się do ruchu kamery. Rozszerzona rzeczywistość jest daleko bardziej zaawansowaną technologią niż to, co można zobaczyć w audycjach telewizyjnych, chociaż wczesna wizja rozszerzonej rzeczywistości miała swój początek w telewizyjnych wyścigach, w meczach football'u, takich jak Racef/x. Stworzona przez SporTVision. Te systemy wyświetlały grafiki tylko z jednego punktu widzenia. Następna generacja systemów rozszerzonej rzeczywistości będzie wyświetlać grafikę z perspektywy każdego z użytkowników biorących udział w sesji. Rozszerzona rzeczywistość jest wciąż we wczesnym stadium badań i rozwoju na niektórych uniwersytetach i w wysoko zaawansowanych technologicznie firmach. Prawdopodobnie, pod koniec tego dziesięciolecia zobaczymy pierwszy system rozszerzonej rzeczywistości na rynku masowym (jeden z naukowców nazwał ten system „Walkman 21 wieku”). To co rzeczywistość wirtualna próbuje robić, to nie tylko nakładane grafiki na środowisko rzeczywiste w rzeczywistym czasie, ale również zmiana tych grafik do przystosowania do ruchów głowy, oczu użytkownika, tak żeby zawsze grafiki pasowały do danej perspektywy. Trzy komponenty potrzebne do tego, żeby system rozszerzonej rzeczywistości działał to:

- wyświetlacze montowane na głowie
- system śledzenia
- mobilna siła obliczeń komputerów



Photo courtesy Columbia University Computer Graphics and User Interfaces Lab

Wczesny prototyp przenośnego zestawu dla systemu rozszerzonej rzeczywistości.

Metą dla ludzi rozwijających rzeczywistość rozszerzoną będzie połączenie tych trzech komponentów w jedną jednostkę, umiejscowioną w noszonym na pasku urządzeniu, które bezprzewodowo przekazuje informacje na wyświetlacz, który jest podobny do zwykłej pary okularów.

KOMPONENTY

Wyświetlacze montowane na głowie.

Jak tylko monitory pozwoliły nam zobaczyć tekst i grafikę wygenerowany przez komputery, HMDS (head-mounted displays) będą zdolne pokazać grafikę i tekst stworzony przez systemy rozszerzonej rzeczywistości. Jak dotąd, szczególnie HMD'sy nie były tworzone z myślą o rozszerzonej rzeczywistości. Większość z wyświetlaczy przypomina niektóre typy okularów narciarskich, które były stworzone oryginalnie dla rzeczywistości wirtualnej. Są dwa podstawowe typy hełmów (HMDS):

- video
- optyczne

Wyświetlacz video zasłania otaczające użytkownika środowisko, używa małe kamery video przymocowane na zewnątrz okularów do wychwytywania obrazów.

W środku wyświetlacza obraz video jest odtwarzany w czasie rzeczywistym i grafika jest nakładana na video. Jedynym problemem z używaniem kamer video są opóźnienia, oznacza to, że są pewne opóźnienia w dostosowaniu obrazu dochodzącego do obserwatora, kiedy porusza głową.



Wyświetlacze dla rozszerzonej rzeczywistości są wciąż nieporęczne; ale inwestorzy wierzą, że są w stanie stworzyć wyświetlacz, który przypomina z wyglądu parę okularów.

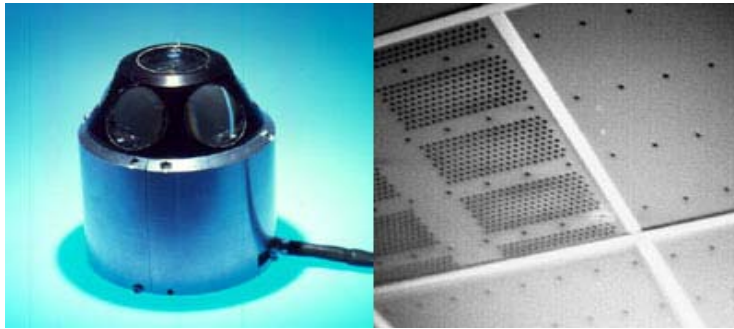
Większość z firm, które stworzyły optyczne wyświetlacze (HMDS) zajęły się handlem. Sony stworzyła półprzezroczysty wyświetlacz, którego używają naukowcy, nazywa się Glasstron. Blair MacIntyre, kierownik laboratorium rozszerzonej rzeczywistości w Georgie Tech, wierzy, że takie wyświetlacze (Microvision's Virtual Retinal Displays) dobrze się zapowiadają dla systemu AR. To urządzenie obecnie używa światła do malowania obrazów na siatkówce przez szybkie przemieszczanie się źródła światła (wzdłuż i w szerz na siatkówce). Problem z wyświetlaczami Mikrowizyjnymi jest taki, że obecnie kosztują one około \$10.000. Wyświetlacze skanujące siatkówkę są obiecujące, ponieważ mogą być małej wielkości. Będą wyglądać jak normalna para okularów, która posiada źródło światła na stronie wyświetlanych obrazów na siatkówkę.

Śledzenie i orientacja.

Największym wyzwaniem, któremu naukowcy zajmujący się AR muszą sprostać jest potrzeba dokładnej znajomości, w jakim położeniu znajduje się użytkownik w odniesieniu do jego/jej otoczenia. Pojawia się tu problem dodatkowy, a mianowicie śledzenie ruchów oczu i głowy użytkownika. System śledzenia rozpoznaje ruchy i

wyświetla grafikę odnoszącą się do środowiska świata realnego, użytkownik widzi każdą daną chwilę. Obecnie, zarówno wyświetlacze video i optyczne mają charakterystyczne opóźnienia w pokrywającym się materiale, jest to spowodowane obecnie osiągalną technologią śledzenia. Dla rozszerzonej rzeczywistości osiągnięto jej pełny potencjał, który musi być użyteczny zarówno na wyjściu i wejściu. Obecnie, najlepsze technologie śledzenia osiągalne są dla szeroko otwartych obszarów GPS(Global Positioning System). Jakkolwiek GPS osiąga dokładność około 10 do 30 metrów, co nie jest złe w gruncie rzeczy, ale nie jest wystarczająco dobre dla rozszerzonej rzeczywistości, która potrzebuje dokładności mierzonej w milimetrach i mniejszej. System AR byłby bezwartościowy jeśli grafika byłaby wyświetlana około 10 do 30 metrów dalej od tej na którą patrzyłby w danej chwili użytkownik. Istnieją parę sposobów na zwiększenie dokładności śledzenia. Na przykład, wojsko używa wielokrotnych sygnałów gps. To jest „różnicowy” GPS, który wymaga używania obszaru, który jest właśnie obserwowany. Wtedy system używałby odbiornika GPS z anteną, taka lokalizacja jest bardzo dobrze znana i precyzyjnie śledzona w obrębie danego obszaru. To będzie pozwalało użytkownikowi poznać dokładnie jak niedokładny jest odbiornik GPS, i będzie można dostosować go odpowiednio do systemu rozszerzonej rzeczywistości. „Różnicowy” GPS pozwala na lepszą dokładność. Bardziej dokładny system dopiero się rozwija, nazywany jest „Kinematyczny GPS czasu rzeczywistego”, może on osiągać dokładność na poziomie centymetrów. Śledzenie jest łatwiejsze i na mniejszej przestrzeni nie na dużej. Naukowcy z UNC- Chapel Hill rozwijają bardzo precyzyjny system, który pracuje w obrębie 500 stóp kwadratowych. System śledzenia nazywany jest „HiBall” jest optoelektronicznym systemem śledzenia złożonym z dwóch części:

- czujniki optyczne – zamontowane na 6-u użytkownikach
- diody emitujące światło podczerwone(LED) osadzone w specjalnych panelach pod sufitem.



System śledzenia Hiball używa czujników optycznych i diod LED osadzonych w specjalnych panelach pod sufitem, śledzące ruch, o dużym zasięgu.

System ten wykorzystuje znaną lokalizację LED, znaną geometrię czujników optoelektrycznych zamontowanych na głowie użytkownika i specjalnego algorytmu do obliczenia i zawiadomienia o pozycji użytkownika i kierunku(orientacji). Ten system analizuje liczbowo ruch (pozycję) z dokładnością do mniej niż dwóch milimetrów, i pozycję kątową mniej niż 3 stopnie. System ten aktualizuje dane więcej niż 1500 Hz, i opóźnienia utrzymują się około 1 milisekundy.

Mobilna siła obliczeń komputerów

Dla systemu AR nadającego się do noszenia nie ma wciąż wystarczającej mocy obliczeniowej komputerów, żeby stworzyć grafiki 3-D stereo. Naukowcy używają jak dotąd cokolwiek mogą dostać, laptopy, PC. Laptopy dopiero zaczynają być wyposażane w jednostki do przetwarzania grafiki (GPU). Toshiba właśnie dodała Nvidia GPU do swoich notebook'ów, który jest zdolny przetworzyć więcej niż 17 milionów trójkątów na sekundę i 286 milionów pikseli na sekundę, co może umożliwić CPU - intensywność programów, takich jak gry 3D. Wciąż notebook'i pozostają daleko w tyle – Nvidia rozwija zwyczajny 300 MHz procesor do grafiki 3D dla dobrze zapowiadającej się konsoli do gier Microsoftu „Xbox”, która produkuje 150 milionów wielokątów na sekundę – i wieloboki te są bardziej skomplikowane od trójkątów. Przenośne, możliwe do zastosowania w praktyce systemy 3D nie będą osiągalne do końca 2005 roku.

Zastosowanie Rozszerzonej Rzeczywistości

Przewyciężono wyzwania, rzeczywistość wirtualna będzie prawdopodobnie przenikać do wszystkich dziedzin naszego życia. AR posiada potencjał do użycia w każdej gałęzi przemysłu, takich jak:

- Utrzymanie i Budowa – prawdopodobnie będzie to jedno z pierwszych użycie rozszerzonej rzeczywistości. Zakreślacze (markery) będą przymocowane do

poszczególnych przedmiotów, nad którymi pracuje osoba, i system AR może rysować grafikę zaczynając od góry. Rozszerzona rzeczywistość ma prostszą formę, od kiedy system powinien znać tylko gdzie znajduje się użytkownik w odniesieniu do przedmiotu, na który on/ona patrzy się. Nie jest konieczne śledzenie dokładnej fizycznej lokalizacji użytkownika.

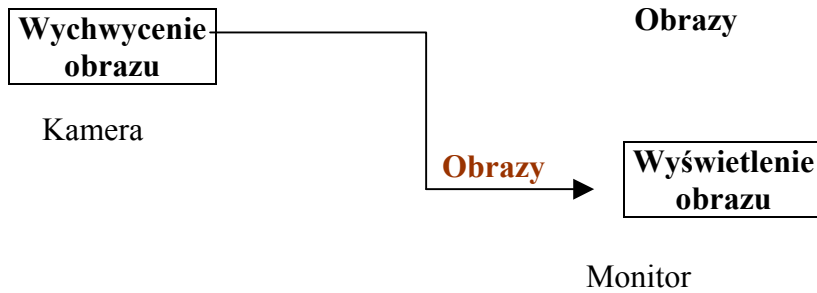
- Wojsko – wojsko obmyślało użycie rzeczywistości rozszerzonej przez dekady. Właściwie, Biuro Badań Morskich sponsorowało niektóre badania nad rzeczywistością rozszerzoną Agencja Obrony Zaawansowanych Projektów Badań(DARPA) sfinansowała projekt projektorów HMD w celu rozwoju wyświetlaczy, które mogą być łączone z przenośnym systemem informacji. Idea w tym przypadku jest taka, że system AR może mieć do dyspozycji wojsko z istotnymi informacjami na temat ich otoczenia(środowiska), takimi jak pokazywanie, gdzie są wejścia na przeciwległych końcach budynku, w pewnym stopniu jest to widzenie tak, jak przez promienie rentgenowskie. Wyświetlacze AR mogą również zwracać uwagę na ruchy wojska, i dawać żołnierzom zdolność do przemieszczania się gdzieś, gdzie wróg nie może ich zobaczyć.
- Natychmiastowa Informacja – turyści i studenci będą mogli używać tego systemu, żeby nauczyć się więcej na temat ważnych wydarzeń historycznych. Wyobraźmy sobie, że użytkownik spaceruje po polu walki podczas np. Wojny Domowej i ogląda odtworzone wydarzenia historyczne na zamontowanym na głowie wyświetlaczu rzeczywistości rozszerzonej. Użytkownik będzie „wtopiony” w wydarzenie, i obraz będzie panoramiczny.
- Gry – gra może być wyświetlana w realnym świecie dookoła użytkownika, i będzie on mógł, dosłownie, jako jedna z postaci z gry. Naukowiec z Australii stworzył prototyp gry, która łączy w sobie Quake’a, popularną grę video, z rozszerzoną rzeczywistością. Umieścił model miasteczka uniwersyteckiego do oprogramowania gry. Teraz, kiedy używa tego systemu, gra otacza go, tak jakby spacerował po tym miasteczku. Istnieją setki potencjalnych aplikacji(zastosowań) dla tej technologii, gry i rozrywka są jednymi z najbardziej oczywistych. Każdy system, który daje ludziom natychmiastową informację, nie wymaga badań nad jego częściami, jest zobowiązany do bycia wartościowym dla każdego, niemalże w każdej

dziedzinie. Systemy AR będą natychmiast rozpoznawać na co patrzy użytkownik, odbierać i wyświetlać dane związane z tym widokiem.

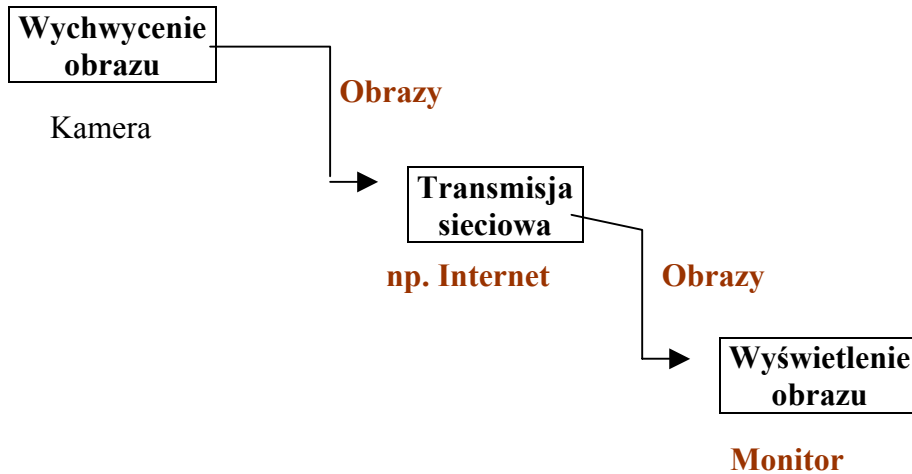
Dodatki

Porównanie transmisji teleimersji i videokonferencji

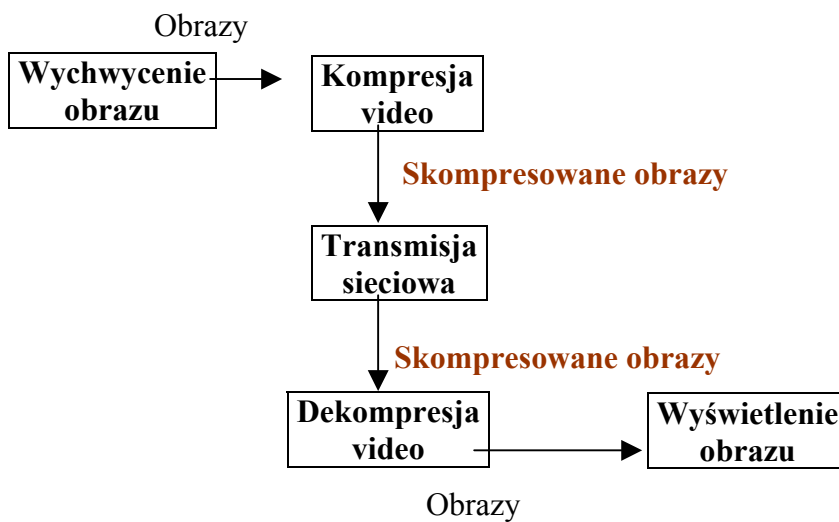
Wideokonferencja 2D



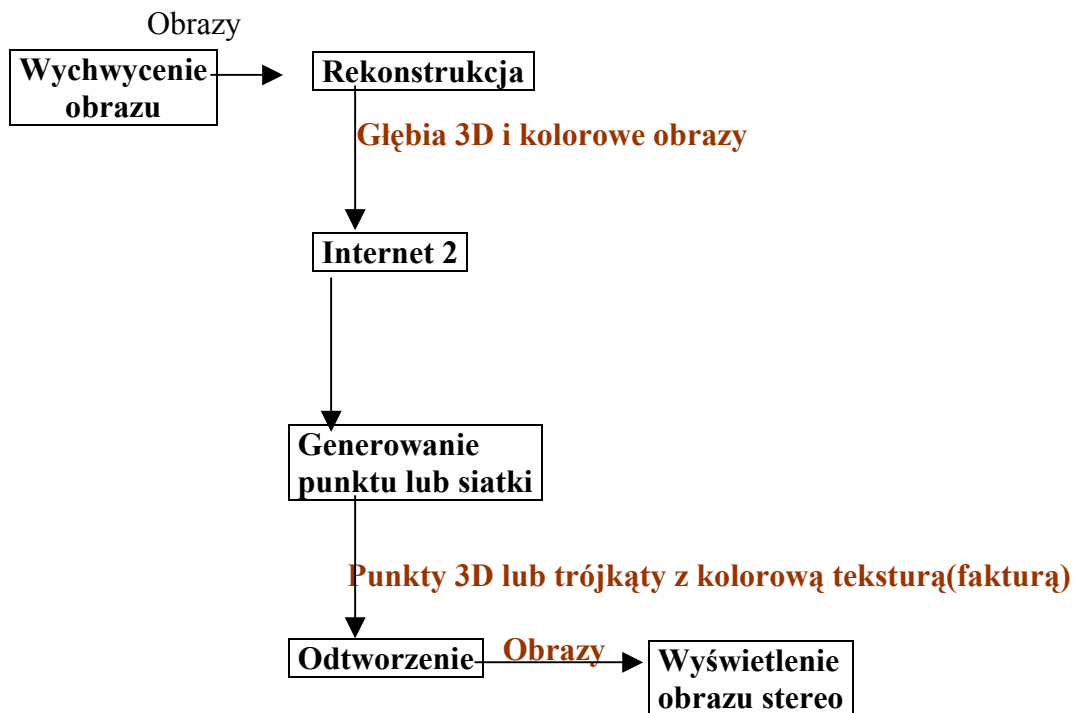
Jedno miejsce dla transmisji



Nowy typ danych



Teleimersja 3D



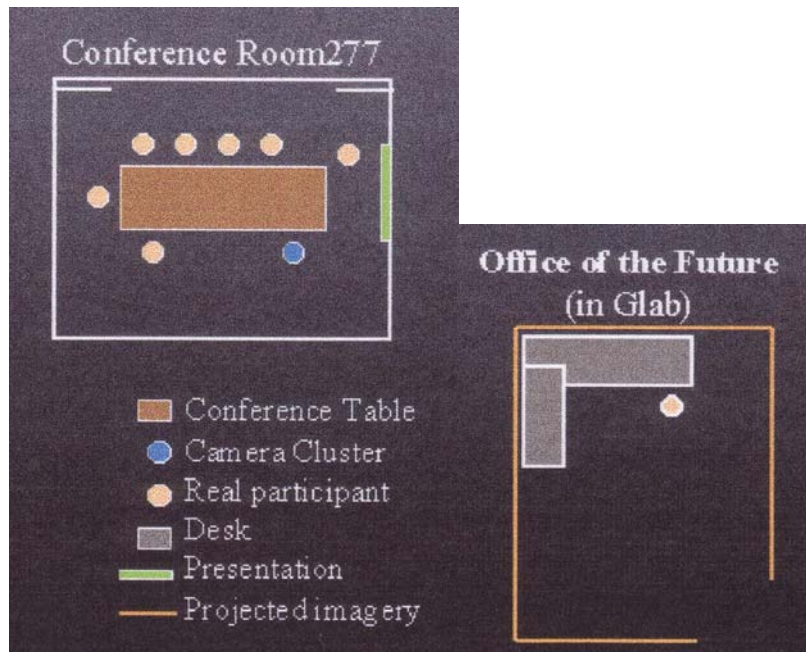
Szybkość transmisji danych w teleimersji:

$$76800 \cdot 40 \cdot 5 \cdot 1,0 \cdot 5 = 77 \text{Mbps na uczestnika}$$

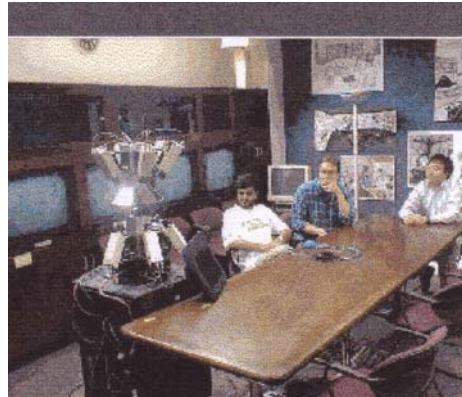
Cele:

- 320x240 rozdzielczość
- 5 klatek na sekundę
- 5 zrekonstruowanych obrazów
- 2 uczestników zlokalizowanych jednocześnie

Schemat pokoju konferencyjnego i Biura przyszłości



Wyniki sesji teleimersyjnej w 1998 roku



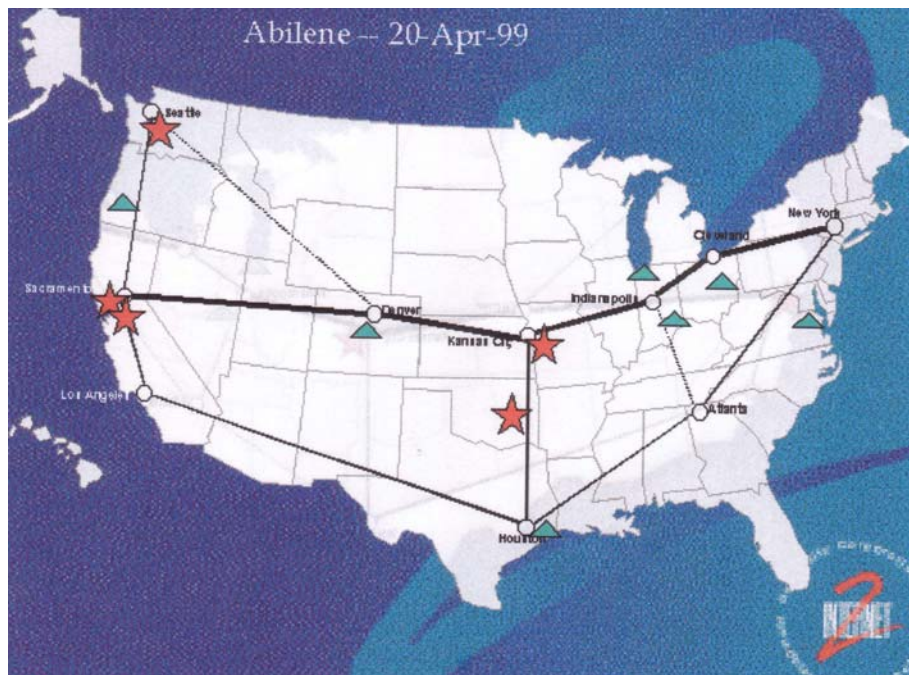
Pomysł dr. Gary Bishop : Biuro Niedalekiej Przyszłości



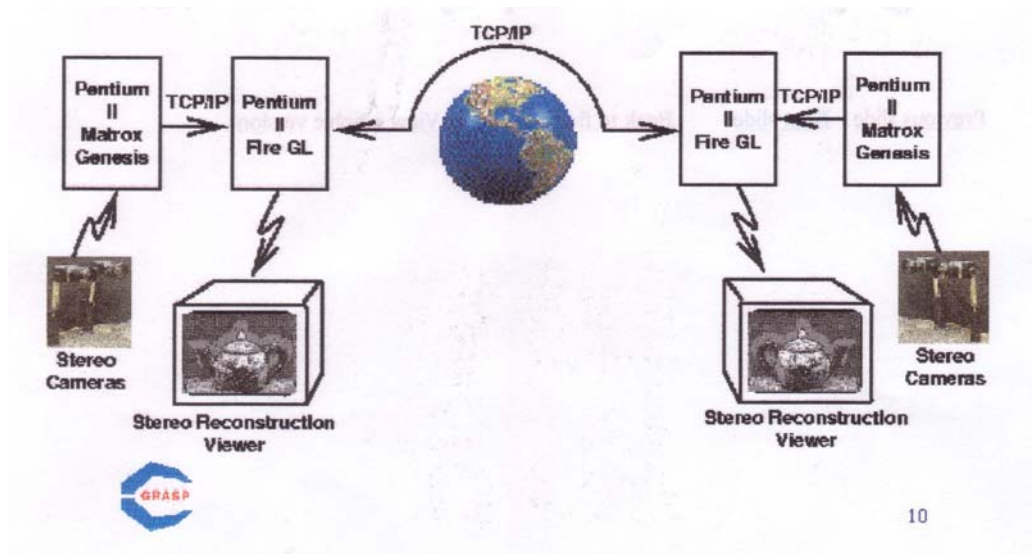
Biuro dr. Ruzena Bajcsy



Infrastruktura 20 kwiecień 1999r.



Teleimersja dla zwykłego obywatela



Teleimersja dotyczy:

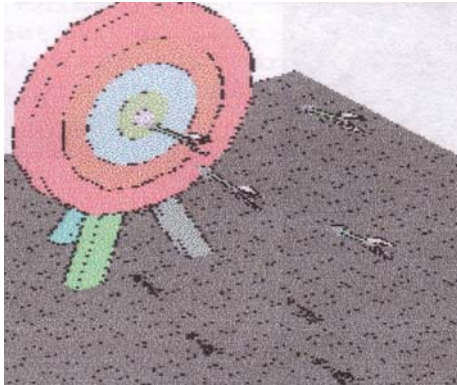
- poczucia prawdziwego świata
- wzajemnego oddziaływania z prawdziwym światem
- wizualizacji realnych i syntetycznych danych
- transmisji (sieci)

Teleimersja wymaga dużej siły obliczeniowej komputerów i szerokopasmowej sieci.

Cel – poprawa sposobu, w jaki ludzie używają komputerów

- perspektywy społeczne: rozwój „społeczeństwa zamiast odosobnionych laboratoriów”
- perspektywy funkcjonalne: współpraca przy projektach odległych
- perspektywy systemowe

Sketch



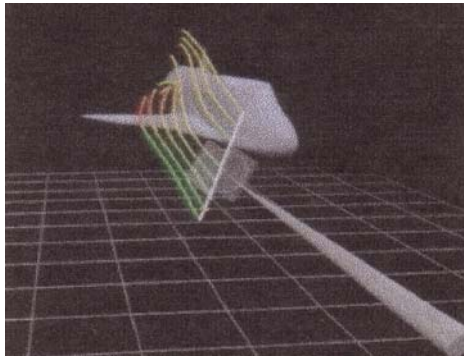
Szkicowanie jest często wspaniałym sposobem przekazu idei ukrywając fakt, iż idea ta znajduje się na etapach kształtowania. Ma wiele korzyści. Prędkość jest wysoka ponieważ używa się zbliżonych obrazów wizualnych za pomocą prostych narzędzi (ołówki i papier). Nie ma potrzeby na precyzję lub wyspecjalizowaną wiedzę. W zasadzie precyzja może spowodować zamieszanie u widza, który może dojść do wniosku, iż idea jest bardziej kompletna i stała niż to wygląda w rzeczywistości. Stosuje się również niskiego poziomu korektę i rewizję. Większość trójwymiarowych modelujących systemów komputerowych sprawdza się w generowaniu dowolnych szkiców dokładnych modeli trójwymiarowych i wspierają wysokiego poziomu edytowanie oraz rewizję. Zastosowanie SKETCH ma na celu połączenie ze sobą korzyści jakie niosą obie metody ręcznej prezentacji oraz dokładnego modelowania komputerowego aby stworzyć warunki dla szybkiej konceptualizacji oraz edytowania zbliżonych obrazów trójwymiarowych. Aby to osiągnąć, SKETCH wykorzystuje proste, niefotorealistyczne odtwarzanie oraz czysto ręczny interfejs który opiera się na uproszczonym rysowaniu linii prymitywów i pozwala na to aby wszystkie operacje przebiegały w środowisku trójwymiarowym. Dzięki temu, iż możemy oglądać obrazy w Internecie, wykorzystywać programy graficzne oraz dokonywać obliczeń, komputery stały się częścią naszego życia. Jednak jest coś trudnego w komputerach; idea, iż jest to 'maszyna' jest wciąż żywa. Wiele razy wydaje się, że dobrze byłoby zbudować komputer gdzie można by było robić notatki lub pisać naprędce kilka słów, tak jak na zwykłej kartce papieru. W tym rozdziale będzie opisane graficzne oprogramowanie komputerowe jakiego będziemy używać w przyszłości w dziedzinie grafiki komputerowej. Jest to nowy program nazywany Sketch. W dzisiejszych czasach jest jeszcze wiele osób, które wciąż uważają, iż używanie komputerów jest trudne. Nie jest jeszcze łatwo z nich korzystać w taki sposób w jaki

chcielibyśmy aby one funkcjonowały. Chodzi o to jak naszkicować jakiś rysunek na kartce papieru tak aby był on dwuwymiarowy. W taki sam sposób rysuje się kilka linii i nagle rysunek staje się trójwymiarowy. Najpierw szkicujemy rysunek na kartce papieru w tradycyjny sposób. Projektant nadaje rysunkowi trochę objętości. Najpierw wykonujemy szkic ręcznie a potem na komputerze. Jednak później trzeba nadal dorysować boczny i górny obraz.

Najpierw artysta, przy pomocy ołówka, bardzo nieprecyzyjnie szkicuje stół. Następnie, na komputerze rysuje stół bardziej precyzyjnie, z punktu widzenia szkicu. Potem, musiał naszkicować jeszcze jeden rysunek aby pokazać boki oraz górę stołu. Wtedy potrzebowałby powrócić jeszcze raz do komputera i przerysować boczny obraz tak aby był bardziej precyzyjny. Jeżeli użyjemy nowego oprogramowania wygląda to zupełnie inaczej. Obraz staje się trójwymiarowy w procesie szkicowania kiedy tylko rysujemy trzy linie (nogi od stołu są rysowane pod stołem), rysuje się tylko tę część która jest widoczna. Ale komputer wypełnia część pod stołem z drugiej strony. To jest jak szkicowanie na kartce papieru, ale rysunek automatycznie staje się trójwymiarowy! (obracanie przedmiotów). Możemy je widzieć pod różnym kątem. To niesamowite, że możemy obserwować przedmioty pod różnym kątem. Czy istnieją jakieś zasady szkicowania, na przykład okna?

Oto kilka przykładów zasad. Sześcian, może być opisany jako szerokość, wysokość i głębokość. I w taki sposób się go rysuje. Ale istnieją inne zasady. W zasadzie rysuje się przedmioty tak jakbyśmy rysowali je odręcznie. Nie trzeba rysować dokładnie, to komputer pracuje nad precyzją rysunku, sprawi, że będą pasować dokładnie. Upłynie jeszcze trochę czasu zanim będzie można z tego korzystać. Niesamowitą rzeczą jest to, iż rysujemy coś zupełnie przypadkowo i automatycznie rysunek staje się trójwymiarowy. Będzie to popularnym narzędziem w dziedzinie projektowania. Jest to najnowsza wersja która ukazała w 1996 r. Komputer, w którym ten system działa jest wciąż kosztowny, ale wkrótce będzie bardziej dostępny. Oprogramowanie będzie wciąż ulepszane a ceny komputerów tak szybko spadają, iż, będzie można korzystać z tego oprogramowania na PC. Dzięki temu oprogramowaniu można stać się pomysłowym projektantem!

Trójwymiarowa Biblioteka Ikonek



Założone przez NASA, NSF Centrum Graficzne i Wizualizacyjne na Uniwersytecie Brown rozwinęło trójwymiarowe interfejsy użytkownika do graficznych oraz immersyjnych zastosowań wizualizacji naukowej. Najnowszy projekt interfejsu użytkownika przy współpracy z NASA to rozwinięcie samodzielnej biblioteki służącej tworzeniu i wzajemnemu oddziaływaniu trójwymiarowych ikonek. Rdzeniem biblioteki jest zestaw klocków służących tworzeniu zwykłych trójwymiarowych ikonek oraz funkcji pozwalających na bezpośrednią manipulację ikonkami. Biblioteka będzie połączona z systemem Wirtualnego Tunelu Aerodynamicznego NASA jak również jest udostępniona publicznemu użytkownikowi.

Poza podstawowymi ikonkami klocków, biblioteka wspiera interakcyjne cienie oraz system kontroli ręcznej nawigacji (np. nadawanie ruchu postępowego i wstecznego, wirtualny obrót manipulatora kulowego oraz przesuwanie płaszczyzny filmu) za pomocą jednego przycisku myszki. Te systemy kontroli kamery pochodzą z systemów używanych w SKETCH.

W szczególności rozwinięto interakcyjne techniki umieszczania sond w środowisku trójwymiarowym, kontrolowanie zarówno przestrzennych jak i nie przestrzennych parametrów kilku technik wizualizacyjnych, używanie przyrządów wchodzących o sześciostopniowej swobodzie, nawigacja trójwymiarowym środowiskiem, zarządzanie komentarzami głosowymi, jak również rozwój dostępu do typowych przyrządów wchodzących tworzących szybkie prototypy.