



PORTFOLIO: System projekcji stereoskopowej.

Autorzy: Michał Turek

Centrum Inteligentnych Systemów Informatycznych Akademia Górniczo-Hutnicza im. Stanisława Staszica al. Mickiewicza 30, 30-059 Kraków
budynek C-2 pokój 426 tel.: 12 617 44 53 www.isi.agh.edu.pl isi@agh.edu.pl



Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Społecznego

Inteligentne budynki - wykorzystanie symulacji sterowanych danymi do analizy i krótkoterminowej predykcji zachowań pieszych.

Jakub Porzycki, Jarosław Wąs, Robert Lubaś, Marcin Mycek

AGH University of Science and Technology
Institute of Applied Computer Science
al. Mickiewicza 30, 30-059 Kraków, Poland
porzycki,jarek,rlubas, mycek@agh.edu.pl

1 Wprowadzenie

Nieustanny rozwój technologii spowodował, że prawie wszystkie nowoczesne budynki, najczęściej biurowe i usługowe, są w pełni zautomatyzowane i centralnie zarządzane. Oświetlenie, dostęp do pomieszczeń i drzwi awaryjnych, systemy audio-wizyjne, windy, ogrzewanie - wszystkie te aspekty zarządzania infrastrukturą mogą być połączone razem. Daje to sposobność do zwiększenia na niepotykaną na wcześniejszą skalę komfortu i bezpieczeństwa użytkownika takich budynków.

Wyzwanie stanowią wydajne mechanizmy kontroli takim układem. Większość podukładów w takim systemie ma duże opóźnienie np. na efekty włączenia ogrzewania trzeba czekać od paru minut do paru godzin, zależnie od budynku. Z tego powodu ważne jest analizowanie historycznych danych oraz przewidywanie sytuacji na podstawie dostępnych informacji. Powszechnie stosowana jest pierwsza metoda, jednak bazowanie jedynie na historii nie pozwala na dużą responsywność systemu oraz powoduje jego złe funkcjonowanie w przypadku sytuacji nietypowej. Drugie podejście jest znacząco trudniejsze, ale pozwala na reagowanie na sytuację w precyzyjnie dopasowany sposób. Brakuje jednak na obecną chwilę systemów, które realizują tą funkcjonalność.

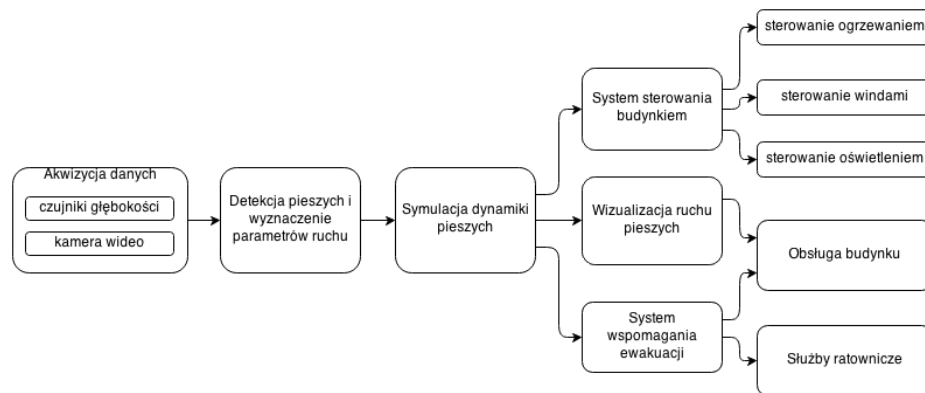
2 Inteligentne budynki czyli o wykorzystaniu symulacji sterowanych danymi

2.1 Symulacje sterowane danymi (SSD)

Symulacje sterowane danymi (*SSD*) są koncepcją łączącą klasyczne symulacje komputerowe z analizą danych rzeczywistych. W kontekście dynamiki pieszych łączą one modele symulacji dynamiki tłumów z metodami wyznaczania parametrów ruchu pieszych.

Wyniki symulacji zawsze w znaczący sposób, zależą od stanu początkowego. Dla tego tak istotne jest używanie rzeczywistych informacji jako danych wejściowych. Im dłużej trwa symulacja tym warunki początkowe mają większe znaczenie. Im większa niepewność wyznaczenia początkowych położenia, parametrów,

celów do osiągnięcia przez pieszych etc., tym większa niepewność otrzymanych wyników w symulacji. *SSD* przeciwdziała tej sytuacji wykorzystując nie tylko jako stan początkowy estymację rzeczywistej sytuacji, ale również na bieżąco uwzględniając nowe pomiary w celu poprawy wyników. W ten sposób jeśli choć częściowe dane mogą być dostarczane do symulacji, przez dowolnie długi okres czasu będzie ona aktualna i możliwie dokładna w swoich przewidywaniach dostosowując się do zmieniających się warunków, zapewniając najdłuższą możliwą i aktualną prognozę sytuacji w systemie.



Rysunek 1. Przykładowy schemat przepływu danych w prezentowanym systemie. Informacje z czujników przekazywane są do modułu wykrywającego pieszych, informacje o pozycji i parametrach ruchu pieszych używane są jako dane wejściowe do symulacji, które potem są przekazywane do: systemu zarządzania budynkiem oraz do obsługi budynku za pomocą modułu odpowiadającego za wizualizację. Po wykryciu zagrożenia uruchamiany jest system wspomaganie ewakuacji, który dostępny jest dla obsługi budynku oraz służb ratowniczych.

2.2 Wzrost funkcjonalności obiektu przy wykorzystaniu SSD

Jak wspomniano we wstępie zagadnienie centralnego i zaplanowanego sterowania dużymi obiektami jest trudne i często nierealizowalne ze względu na potencjalnie zmieniające się warunki. Nawet dla najprostszych kompleksów jak np. biuro nieplanowane zebranie znacząco zaburza typowy cykl dzienny i zmienia rozkład osób w obiekcie.

Jednym z aspektów, który często wprowadza dyskomfort u użytkowników obiektów i/lub generuje duże koszty jest oświetlenie. Duża część wykorzystywanych systemów oświetleniowych charakteryzuje się długim okresem zanim lampy uzyskają pełną świetność (w szczególności np. wysokoprężne lampy sodowe). Takie urządzenia podczas pracy zużywają mało energii, ale mają długi czas załączania i ich częste włączanie oraz wyłączenie znacząco redukuje czas życia

zwiększając koszty utrzymania instalacji. Drugi typ oświetlenia charakteryzuje się większą responsywnością, ale również zwiększonym zużyciem energii elektrycznej. W obu przypadkach wyłączenie oświetlenia, gdy jest ono niepotrzebne przez uzasadniony ekonomicznie czas, pozwala zaoszczędzić na kosztach utrzymania obiektu. Problemem jest wykrywanie potencjalnych początków i końców takich okresów. Sytuacje gdy nie ma osób w danym rejonie łatwo jest wykryć. Jednak biorąc pod uwagę potencjalną bezwładność systemu, idea inteligentnych budynków wspomaganych *SSD* pozwoliłaby przewidywać końce takich okresów z odpowiednim wyprzedzeniem i odpowiednio na nie reagować.

Takie rozwiązanie pozwala też znacząco zwiększyć komfort użytkownika obiektu. Stosowane obecnie rozwiązania (np. fotokomórki) realizują funkcjonalność dynamicznego sterowania oświetleniem, ale często kosztem dużego dyskomfortu u użytkowników. Wyłączenia się oświetlenia podczas bezruchu, nie wykrycie osoby przez czujnik, szybka i nieprzyjemna zmiana jasności spowodowana nagłym załączeniem się oświetlenia - wszystkie te niepożądane zjawiska spowodowane są faktem, że obecnie stosowane systemy reagują jedynie na sytuację obecną. Ich zdolność przewidywania jest ściśle określona i opiera się raczej na odpowiednim ich rozmieszczeniu (np. fotokomórka włączająca światło w korytarzu umieszczona znacząco przed korytarzem).

W praktyce występuje również wiele innych zagadnień generujących duże koszty albo dyskomfort użytkownika. Ogrzewanie pomieszczeń albo chłodzenie za pomocą klimatyzacji w sposób znaczący zależy od ilości osób w pomieszczeniu i również posiada dużą bezwładność czasową. Zagadnienia sterowania i potencjalnego wykorzystania *SSD* są dokładnie analogiczne jak w przypadku sterowania oświetleniem. Znajomość stanu układu z wyprzedzeniem czasowym pozwala lepiej zarządzać takim systemem. Taką informację można też wykorzystać do mniej oczywistych zastosowań, jak na przykład ustawianie wind na piętrach gdzie spodziewamy się większej ilości osób chcących zmienić piętro lub odpowiedniego przydziału osób z obsługi hotelu czy lotniska do rejonów, gdzie w niedalekiej przyszłości będą potrzebne.

Zebrane dane mogą też zostać wykorzystane do globalnej analizy. Ponieważ *SSD* pozwala odtworzyć stan całego obiektu na podstawie częściowych pomiarów, umożliwia to przeanalizowanie, gdzie np. powinny zostać umieszczone ławki, automaty z napojami etc.

2.3 Poprawa bezpieczeństwa obiektu dzięki SSD

Drugim obszarem zastosowania koncepcji inteligentnych budynków wspomaganych *SSD* jest poprawa bezpieczeństwa użytkowanych obiektów. Drogi ewakuacyjne wyznaczane dla obiektów są zazwyczaj ściśle określone i dla przewidzianego w scenariuszach ewakuacyjnych rozmieszczenia ludzi sprawdzają się znakomicie.

Należy jednak podkreślić dwa fakty. Po pierwsze w przypadku gdy rozkład ludzi w obrębie obiektu jest znacznie inny niż zakładany - np.: w skutek zebrania wszystkich pracowników firmy albo w czasie lunchu - mogą istnieć lepsze drogi

ewakuacyjne zmniejszające całkowity czas ewakuacji z obiektu. Ponadto zaplanowane drogi mogą zawodzić jeśli nastąpi nieprzewidziane zdarzenie np.: wizyta dużej ilości osób w obiekcie przekraczającego jego maksymalną pojemność lub awaria uniemożliwiająca wykorzystanie części drogi ewakuacyjnej (np.: zawalenie się schodów czy pożar). Oba zastrzeżenia wynikają z jednej cechy - drogi ewakuacyjne są statyczne i zależne od określonych, niezmiennych założeń.

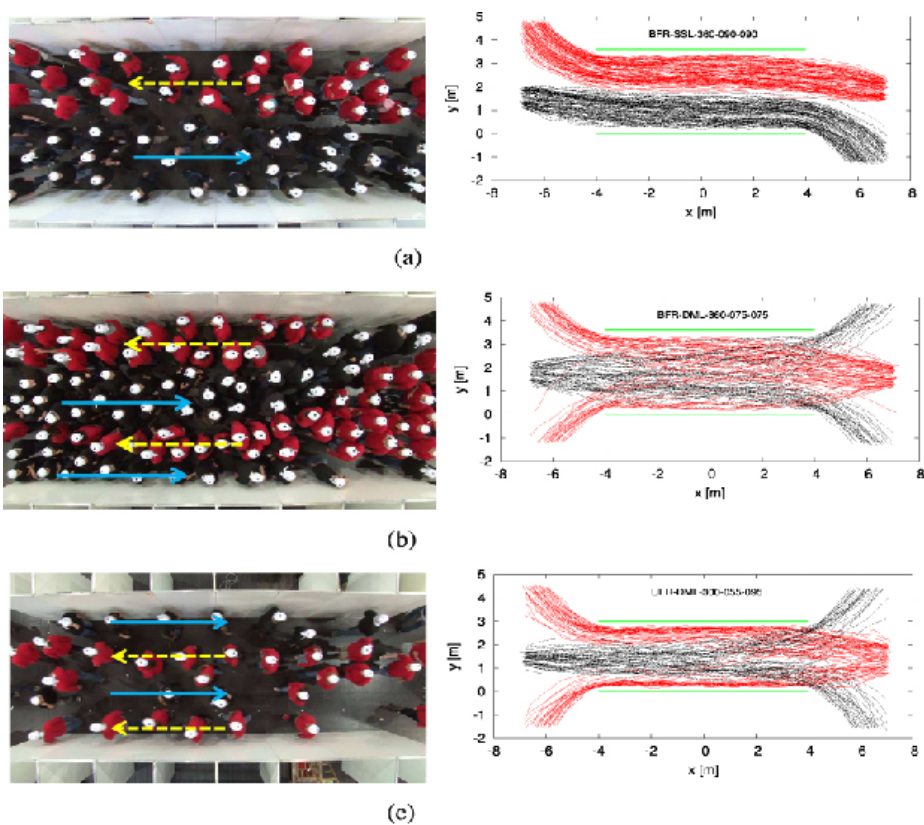
Wykorzystanie SSD pozwala reagować na nieprzewidziane sytuacje - w dynamiczny sposób dobierając drogi ewakuacyjne, reagując na zaburzenia w rozmieszczeniu osób albo nieprzewidziane sytuacje. Symulacja na bieżąco korygowana ciągłym dopływem danych pozwala też posiadać informację o pełnym stanie układu. Co więcej wykorzystanie symulacji zachowanie tłumów, pozwala wykryć część ewentualnych zagrożeń zanim one faktycznie nastąpią. Ponadto umożliwia to reakcje w przypadku wykrycia, w trakcie ewakuacji np. zatoru, oraz stanowi nieocenione wsparcie dla policji i straży pożarnej jako źródło informacji.

3 Ruch pieszych

Tłum pieszych może być realistycznie symulowany za pomocą różnych modeli dynamiki i zachowań tłumu. Do jak najwierniejszego odwzorowania rzeczywistości niezmiernie ważne jest przeprowadzenie ilościowej i jakościowej weryfikacji modelu. Walidacja ilościowa polega na konfrontacji danych liczbowych uzyskanych w wyniku symulacji z danymi empirycznymi. Podczas analizy jakościowej model sprawdzany jest pod kątem występowania zjawisk samoorganizacji. Do takich kolektywnych zjawisk zaliczamy:

- Zatlóczenie (*ang. jamming*): W sytuacji dużej gęstości w tłumie pojawiają się różnorodne zjawiska, np. duża liczba ludzi opuszcza pomieszczenie w tej samej chwili. Zjawisko tamowania jest typowe dla wąskiego gardła, gdzie przepływ ludzi jest ograniczony przez drzwi lub zwiężenia. Inne rodzaje zatlóczenia pojawiają się w przypadku ruchu pieszych w przeciwnych kierunkach - dwie grupy pieszych wzajemnie się blokują.
- Formowanie się strug (*ang. lane formation*): Gdy dwie grupy poruszają się w przeciwnych kierunkach, pojawia się efekt spontanicznego łamania symetrii. W ruchu pieszych dochodzi do samoorganizacji polegającej na formowaniu się strug w ruchu wśród ludzi poruszających się w jednym kierunku. Zjawisko to jest widoczne na rys.: 2. W ten sposób silne oddziaływanie pomiędzy pieszymi jest zredukowane i możliwe jest poruszanie się z większą prędkością.
- Oscylacje przepływu (*ang. flow oscillations*): W ruchu pieszych w obie strony w wąskim gardle można zaobserwować zmiany oscylacyjne kierunku ruchu. Kiedy jeden pieszy jest w stanie przejść przez wąskie gardło, ruch w tym samym kierunku (podążaniem za poprzednikiem) dla pozostałych pieszych staje się łatwiejszy. Kierunek przepływu zmienia się, kiedy w wyniku chwilowo mniejszego ruchu, przez wąskie gardło przejdzie pieszy poruszający się w innym kierunku.

- Panika (*ang. panic*): W sytuacji paniki, może pojawić się wiele nieintuicyjnych w interpretacji zjawisk. W fenomenie szybki jest wolny (*ang. faster-is-slower effect*)[12] wyższa pożądana prędkość pieszych prowadzi do wolniejszego ruchu w tłumie. W efekcie zamrożenie-przez-ogrzewania (*ang. freezing-by-heating effect*) [13] zwiększenie wahań może prowadzić do bardziej uporządkowanego stanu.



Rysunek 2. Naturalne zjawisko formowania się strug w tłumie przemieszczającym się w dwóch przeciwnych kierunkach.

Analizy zachowań tłumu z nagrań wideo wskazują, że oprócz ruchu chaotycznego w dynamice pieszych można doszukać się regularnych zachowań. Rezultaty badań pokazują, że regularne zachowanie odnosi się tylko do pieszych zorientowanych na konkretny cel. Nieregularne wzorce w ruchu występują na przykład u turystów i dzieci, które uczą się obierania optymalnej strategii ruchu.[14],[15].

Podsumowując zachowania tłumu można wyróżnić cztery główne obserwacje:

1. Piesi zazwyczaj obierają najkrótszą drogę do następnego celu, która ma kształt wielokąta. Jeżeli występuje alternatywna droga tej samej długości, pieszy preferuje ścieżkę po której może poruszać w linii prostej jak najdłużej i zmieniać kierunek jak najpóźniej. Pieszy woli unikać wydłużania drogi lub poruszania się w kierunku przeciwnym do kierunku ruchu nawet gdy bezpośrednia ścieżka jest zatłoczona.
2. Pieszy preferuje ruch ze swoją prędkością pożądaną, jest dla niego komfortowa i zachowuje ją do momentu w którym nie będzie zmuszony przyspieszyć w celu osiągnięcia celu na czas. Badania pokazują, że prędkość pożądana w tłumie ma rozkład normalny.
3. Piesi zachowują pewien dystans do innych pieszych i przeszkód. Dystans ten zmniejsza się wraz ze wzrostem pośpiechu i gęstości w tłumie. Osoby nieporuszające się (oczekiwanie na pociąg lub posiłek w jadalni) są równomiernie rozłożone na dostępnej powierzchni. Osoby, znając siebie, mogą tworzyć grupy, które są jednostki, które zachowują się podobnie do pojedynczych przechodniów. Wielkość grup mają rozkład Poissona.
4. Piesi zazwyczaj nie odzwierciedlają swojej strategii zachowań w każdej sytuacji na nowo, ale działają mniej lub bardziej automatycznie (jak doświadczony kierowca).

4 Metody monitorowania zachowań pieszych

4.1 Kamery

Obecnie najpopularniejszym urządzeniem używanym do akwizycji danych o ruchu pieszych jest kamera wideo. W większości są to kamery należące do różnorodnych systemów monitoringu. Do niedawna jedyną powszechnie stosowaną metodą analizy takich nagrań była bezpośrednia obserwacja przez wyznaczoną do tego zadania osobę.

Od kilku lat można zaobserwować narastający trend polegający na coraz częstszym stosowaniu różnorodnych algorytmów automatycznej analizy obrazów. Okazuje się jednak, że utworzenie algorytmów dokładnie wyznaczających parametry ruchu pieszych w dowolnie wybranych warunkach nie jest zadaniem łatwym. Istnieje kilka znanych metod wyznaczania pozycji oraz parametrów ruchu pieszych przy użyciu obrazu z kamery. Spośród istniejących rozwiązań, można wyróżnić dwa, najbardziej obiecujące, typy algorytmów, tj.:

- grupowanie spójnych wektorów ruchu,
- detekcja głów przy użyciu transformaty Hougha.

Grupowanie spójnych wektorów ruchu Istnieje duża grupa metod wykrywania pieszych oparta na grupowaniu spójnych wektorów ruchu [24, 25, 26, 27]. Ich wspólnym punktem jest obserwacja, że charakterystyczne cechy pochodzące od jednej osoby będą miały podobne właściwości (np. kierunek ruchu) oraz podobną lokację w przestrzeni. Ogólny schemat działania tych algorytmów można przedstawić następująco:

1. wykryj charakterystyczne punkty na obrazie,
2. oblicz przemieszczenie znalezionych punktów względem poprzedniej ramki,
3. połącz w grupy punkty o podobnym przemieszczeniu,
4. ograniczenie zawartości grup tak aby należały do nich tylko punkty znajdujące się blisko w przestrzeni
5. wyznaczenie pozycji pieszego na podstawie pozycji punktów z danej grupy.



Rysunek 3. Przykład działania algorytmu wykrywania pieszych opartego o grupowanie spójnych wektorów ruchu [26].

Metody oparte o grupowanie spójnych wektorów ruchu wykazują dużą skuteczność wykrywania pieszych - do 94% [26]. Jednak wydajność tych algorytmów znacznie spada w przypadku niskich częstotliwości rejestracji obrazu.

Detekcja głów przy użyciu transformaty Hougha Inną metodą automatycznego wykrywania pieszych przy użyciu kamer wideo, jest wykorzystanie kamer patrzących prostopadłe w dół. W takim kadrze najlepiej widocznym elementem pieszego jest głowa. Założenie to pozwala wykorzystać właściwości transformaty Hougha - mapuje ona dany obraz do przestrzeni parametrów opisujących każde możliwe wystąpienie poszukiwanego kształtu (w tym wypadku okręgu).

Jako, że okrąg może być opisany równaniem:

$$(x - x_c)^2 + (v - y_c)^2 = r^2$$

to przestrzeń parametrów jest zdefiniowana przez trzy parametry: x_c , y_c , r . Detekcja okręgów przy użyciu tej metody polega na tym, że każdy piksel na wykrytych krawędziach kształtu „głosuje” na każdy możliwy zestaw parametrów x_c , y_c , r który mógł się przyczynić do powstania okręgu zawierającego ten piksel. W ten sposób zestawy parametrów odpowiadające za faktycznie istniejące okręgi zdobywają dużo „głosów” i jest wykrywany jako maksimum w przestrzeni parametrów.



Rysunek 4. Przykład detekcji okręgów (głów pieszych) przy użyciu transformaty Hougha. Zarówno rysunek A jak i B pokazują rezultaty detekcji dokonywane na tej samej ramce obrazu. W przypadku rys. A poziom progów jest zbyt wysoki - niektóre głowy są pominięte. Na rys. B poziom progów jest zbyt niski - wszystkie głowy są wykryte, ale pojawiają się również fałszywe detekcje.

Kluczowym elementem opisywanego algorytmu jest właściwe dobranie parametrów detekcji krawędzi oraz warunków jakie muszą spełnić punkty w przestrzeni parametrów aby zostały uznane za odpowiedniki faktycznie istniejących okręgów. W zależności od przyjętych wartości może dojść do pominięcia niektórych głów okręgów (błąd *false positive*) lub fałszywej detekcji rejonów nie będących okręgami (błąd *false negative*), patrz rys.: 4. Żadna z tych sytuacji nie jest pożądana, ale lepszym rozwiązaniem jest zgoda na wystąpienie fałszywych detekcji niż pominięcie rzeczywiście pojawiających się głów. W dalszym etapie fałszywe detekcje są odfiltrowywane przy użyciu wybranych klasyfikatorów (np. sieci neuronowych).

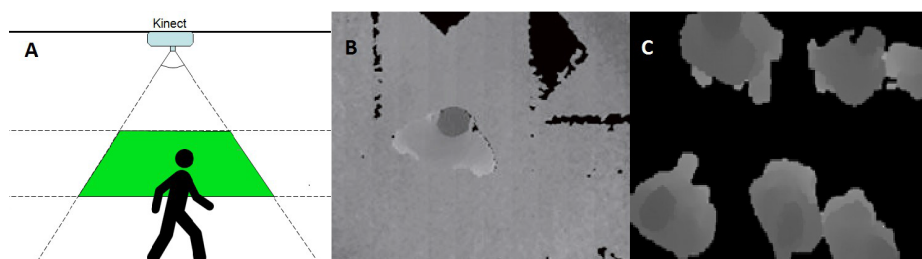
Monitoring wizyjny - podsumowanie Dużym plusem wykorzystania kamer wideo w automatycznym zbieraniu danych o parametrach ruchu pieszych jest możliwość wykorzystania już istniejących instalacji sprzętowych. Ponieważ systemy monitoringu wizyjnego są najbardziej rozpowszechnione, są również najlepiej przetestowane. W przeciwieństwie do innych nowo pojawiających się technologii istnieje pewność poprawnego funkcjonowania kamer wideo w różnych warunkach atmosferycznych i oświetleniowych.

Należy jednak zwrócić uwagę, że algorytmy automatycznej analizy obrazu, zazwyczaj są projektowane dla specyficznych sytuacji - danej geometrii przestrzeni, ułożenia kamery, natężenia ruchu. W szczególności jakość działania znacznej większości algorytmów znacznie spada w przypadku dużej gęstości tłumu. Obecnie żadne z rozwiązań nie stało się ogólnie akceptowalnym standardem. Dla wszystkich można wyróżnić sytuacje w których radzą sobie dobrze jak i takie w których występują częste błędy. Ponadto analiza danych obrazowych wymaga zazwyczaj dużej mocy obliczeniowej, co w przypadku pokrycia kamerami dużego budynku może mieć kluczowe znaczenie.

4.2 Czujniki głębokości

Pomimo iż, skanery 3D i czujniki głębokości były znane już od wielu lat, to dopiero pojawienie się na rynku urządzeń takich jak Asus Xtion i Microsoft Kinect było impulsem do szerokich badań nad wykorzystaniem czujników głębokości do wyznaczania parametrów ruchu pieszych [21, 20, 22, 23].

Sugerowanym przez autorów rozwiązaniem, jest wykorzystanie czujników przymocowanych do sufitu, patrzących pionowo w dół. Przy takim ustawieniu czujników czubek głowy pieszych odpowiada lokalnemu minimum w mapie głębokości. Ogólny schemat wykrywania pieszych przy użyciu map głębokości przedstawiony jest na rys.: 5.



Rysunek 5. Ogólny schemat algorytmów wykrywania pieszych z wykorzystaniem czujników głębokości. Czujnik zamocowany u sufitu, skierowany pionowo w dół, w jego polu widzenia można wyróżnić zakres wysokości, w której spodziewane jest pojawienie się czubka głowy przechodzących pod spodem osób (A). Przykładowa mapa głębokości (B), wyseparowane sylwetki pieszych (C).

Wykorzystanie czujników głębokości pozwala uzyskać bardzo wysoką wykrywalność pieszych - do 99% nawet w gęstym tłumie oraz dobrą estymację położenia pieszego w kolejnych klatkach. Z kolei dużym minusem urządzeń takich jak Kinect czy Xtion, korzystających do wyznaczania map głębokości z podczerwonego światła strukturalnego, jest ich wrażliwość na światło słoneczne. Powoduje to powstawanie wielu „dziur” w mapach głębokości (przykładem mogą być czarne pola na rys.:5 B). Ograniczają one użycie tego rodzaju czujników do zamkniętych budynków. Ponadto wymagają one dość rozbudowanego okablowania. Stworzenie dedykowanego urządzenia do wykrywania ruchu pieszych, mogłoby wyeliminować wspomniane minusy przy zachowaniu dużej dokładności pomiaru.

4.3 Znaczniki i czujniki noszone przez pieszego

Alternatywnym podejściem do akwizycji informacji na temat ruchu pieszego jest wykorzystanie znaczników i aktywnych czujników noszonych przez pieszego, które albo w pośredni sposób dostarczają informacji o ruchu pieszych albo w znaczny sposób upraszczają jego lokalizację.

Znaczniki takie jak w szczególności RFID (*Radio-frequency identification* - identyfikacja radiowa) od dawna stosowane są w rozwiązaniach przemysłowych i korporacyjnych. Rozwiązanie takie składa się z dwóch elementów: układu czytnika oraz układu znacznika. Znacznik wykorzystując fale radiowe emituje co pewien interwał czasu komunikat zawierający w zależności stosowanego standardu kodowania różne informacje. Układ czytnika natomiast ma za zadanie wykrywać i identyfikować znaczniki. W przypadku najpopularniejszych, pasywnych znaczników układ czytnika stanowi również źródło zasilania dla znacznika, który nie posiada własnego źródła energii a na potrzeby emisji gromadzi energię z otaczającego go pola elektromagnetycznego. Najczęściej stosowane znaczniki operują w odległości do 50 cm, ale część rozwiązań umożliwia pracę w zasięgu do paru metrów.

Stosuje się również tzw. aktywne czujniki czyli urządzenia wykrywające ruch noszone przez pieszych. Do niedawna jedynymi tego typu urządzeniami były pedometry (*krokomierze*), które wykrywają ruch poprzez wykrywanie pojedynczych kroków. Wraz ze wzrostem popularności smartfonów instytucja osobnego pedometru odeszła w zapomnienie oraz pojawiła się możliwość połączenia zbieranych danych z innymi sensorami obecnymi w telefonie komórkowym - magnetometrem i odbiornikiem GPS.

4.4 Sugerowane rozwiązanie

Do użytku w prezentowanym systemie, proponujemy rozwiązanie modułowe, które może działać z dowolnie wybranym źródłem danych wejściowych pod warunkiem utrzymania zdefiniowanego interfejsu. Możliwe będzie sterowanie symulacjami przy użyciu najwygodniejszych w użyciu sensorów i wybranej ich ilości.

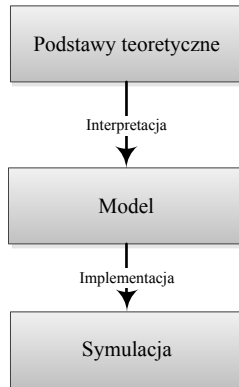
Najbardziej optymalnym rozwiązaniem wydaje się jednak być wykorzystanie czujników głębokości w postaci dedykowanego urządzenia integrującego w sobie sam czujnik głębokości, mikrokomputer dokonujący analizy danych i moduł umożliwiający połączenie z centralnym serwerem (np. moduł WiFi).

Tokeny jako urządzenia o najwyższej potwierdzonej dokładności, mogą zostać wykorzystane do testów i walidacji innych systemów pozyskiwania wiedzy o dynamice pieszych.

5 Modelowanie ruchu pieszych

Modelowanie ruchu pieszych rozpoczyna się od sformułowania niezbędnych podstaw teoretycznych. Prawidłową interpretację założeń teoretycznych prowadzi do stworzenia *modelu*. W momencie zrozumienia teorii przystępuje się do redukcji stopnia swobody poprzez wyspecyfikowanie kilku alternatyw, w taki sposób, by nie utracić spójności modelu. Podejście to jest przedstawione w schematyczny sposób na rysunku 6

Przykładem redukcji może być wybór odpowiedniej reprezentacji przestrzeni lub pozbawienie kontaktu werbalnego w populacji. Efektem implementacji modelu jest program komputerowy symulujący przyjęte założenia teoretyczne z



Rysunek 6. Interpretacja teorii prowadzi do sformułowania modelu, który następnie zostaje zaimplementowany i w efekcie powstaje symulacja.

uwzględnieniem wyżej wspomnianej redukcji. W przypadku dynamiki ruchu, teoria dopuszcza pewne założenia dotyczące zachowań pieszych (*reakcja na otoczenie, orientacja, metoda wyboru drogi, aspekty psychologiczne*). Istnienie takich założeń nie determinuje jednego sposobu interpretacji, czy reprezentacji w modelu.

Tłum pieszych może być postrzegany jako wiele rozproszonych cząstek w systemie. Z technicznego punktu widzenia ciągłe modele ruchu pieszych są podobne do dynamiki molekularnej (*ang. MD – molecular dynamics*). Symulacje *MD* opierają się na numerycznym rozwiązaniu równania Newtona zachodzących interakcji między cząsteczkami. Model *social force* [18] wykazuje właśnie to podobieństwo.

Modele można scharakteryzować według następujących kategorii [19]:

- skali – rozmiar problemu w odniesieniu do czasu i przestrzeni,
- rozdzielczości – poziom detali niezbędnych do reprezentacji przestrzeni,
- wierności odwzorowania – zestaw parametrów mający wpływ na model (wiek, wysokość, waga i sprawność pieszych).

Podczas budowania modelu niezbędne jest podjęcie kilku decyzji, niektórych bezpośrednio lub częściej pośrednio. Poniżej w tabeli 1 przedstawiono niektóre z nich. Głównym wyborem dla tworzenia modelu ruchu pieszych jest kryterium ciągły *vs* dyskretny oraz stochastyczny *vs* deterministyczny.

Dopóki istnieje czynnik wpływu na ludzkie zachowanie, który nie może być do końca poznany, wynik ewakuacji oraz ruch tłumu w ogólnym przypadku nie może zostać przewidziany – odnosi się to do *modeli stochastycznych*. Wybór pomiędzy modelem ciągłym a dyskretnym nie może zostać determinowany tylko aspektem teoretycznym. Biorąc pod uwagę skalowalność modelu i wykorzystanie go w realnych przykładach (ewakuacje z budynków o różnych rozmiarach) różnice w kryterium ciągły *vs* dyskretny stają się ważne. Innymi słowy, wybór zależy od podjętych decyzji implementacyjnych.

Dokładny	\longleftrightarrow	Ogólny
Dyskretny	\longleftrightarrow	Ciągły
Numeryczny	\longleftrightarrow	Analityczny
Stochastyczny	\longleftrightarrow	Deterministyczny
Ilościowy	\longleftrightarrow	Jakościowy
Makroskopowy	\longleftrightarrow	Mikroskopowy

Tablica 1. Kryteria tworzenia modelu.

Odnosząc się do kategorii rozdzielnosci modelu ruchu pieszych, można wymienić dwa główne pojęcia: *modele makroskopowe* oraz *mikroskopowe*. Modele makroskopowe nie reprezentują ogólnej teorii dynamiki tłumu i są zarezerwowane do specyficznych zastosowań. Modele mikroskopowe można określać za pomocą następujących cech:

- szczegółowe przedstawienie przestrzeni,
- osobna reprezentacja pieszych,
- jednolity algorytm ruchu,
- rozważanie osobistych cech i umiejętności pieszych.

5.1 Reprezentacja rzeczywistości

Najczęściej wykorzystuje się dwa podejścia w modelowaniu ruchu pieszych. Podejście makroskopowe opiera się na równaniach przepływu. Do klasyki należy hydrodynamiczny model Paulsa [16]. Brak możliwości symulacji konkretnych osób jest największym ograniczeniem modelu – dlatego, iż większość modeli opiera się na mikroskopowym podejściu (każdy pieszy traktowany jest indywidualnie).

Rozważając reprezentację przestrzeni w modelach mikroskopowych, istnieją dwa podejścia: *dyskretne* oraz *ciągłe*. Wybór jednego z nich jest ściśle związany z implementacją. Mocnym argumentem przemawiającym za ciągłą reprezentacją jest jej stosunkowa prostota i skalowalność takiego modelu. Użycie dyskretnego czasu w modelu zazwyczaj prowadzi do dyskretnej reprezentacji przestrzeni, a to z kolei narzuca wykorzystanie regularnej siatki. Szczególnym rodzajem dyskretnych modeli są modele oparte o teorię automatów komórkowych (*ang. Cellular Automata models*).

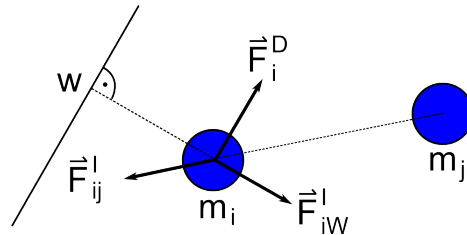
5.2 Model dyskretny

W modelach opartych o automaty komórkowe powierzchnia jest podzielona na kwadratowe komórki (*ang. cell*), które mogą być zajmowane maksymalnie przez jednego pieszego. Ruch pieszych jest determinowany głównie poprzez aktualną konfigurację sąsiedztwa. Dzięki temu model jest niezwykle wydajny, a implementacja i odpowiednia optymalizacja daje dobre rezultaty na komputerach osobistych. Do wad modeli opartych o CA można zaliczyć trudność w modelowaniu ruchu pieszych w różnych kierunkach (*ang. counter flow*). Taki problem

nie występuje w modelach symulujących ruch samochodów, ponieważ ruch na jednym pasie ruchu występuje tylko w jednym kierunku (*ang. uni-directional*) oraz kierowcy mają tendencję do reagowania tylko na samochody znajdujące się w pobliżu.[17]

5.3 Model ciągły

Jak już zostało wspomniane w sekcji 5 ciągle modele dynamiki pieszych można interpretować jako układ poruszających się cząstek. Równania ruchu opierają się na klasycznej mechanice Newtona. Obecnie najbardziej popularny model ciągły to model sił społecznych – *social forces*. Piesi są reprezentowani jako koła o różnym promieniu w dwuwymiarowej przestrzeni. Model *social forces* jest oparty na ciągłej przestrzeni oraz czasie.



Rysunek 7. Model sił społecznych - *social force*.

Zastosowanie modelu *social forces* do symulacji dynamiki pieszych w dużych obiektach takich jak stadion piłkarski posiada dwie główne wady:

1. W odniesieniu do interakcji między pieszymi liczba operacji na jednostkę czasu rośnie kwadratowo w zależności od liczby pieszych.
2. Nie jest jasne jak określić prędkość i siły odpychające dla dowolnego układu.

6 Podsumowanie

Obecnie większość systemów monitoringu jest wykorzystywana do weryfikacji *post factum* przebiegu pewnych wydarzeń. W ostatnich latach prowadzone są badania nad wykorzystaniem systemów monitoringu do automatycznej i natychmiastowej reakcji na różnego rodzaju zdarzenia - np. wystąpienie silnych zakłóceń typowego ruchu, bójka, pożar . . . itp. System, którego koncepcja przedstawiona jest w tym artykule idzie krok dalej, pozwala na przewidywanie sytuacji jakie mogą nastąpić w przyszłości, a także wykorzystanie tej wiedzy do poprawy komfortu i bezpieczeństwa użytkowników danego budynku.

Wykorzystanie dopracowanych i dobrze zwalidowanych modeli dynamiki pieszych w połączeniu z nowymi możliwościami wyznaczania parametrów ruchu

pieszych pozwala na stworzenie sterowanych danymi symulacji zachowań użytkowników danego budynku. Pozwalają one nie tylko na detekcja aktualnego rozmieszczenia osób w budynku, ale również predykcje ich pozycji w przyszłości.

Prezentowane rozwiązanie pozwala na wykrycie części potencjalnych zagrożeń dla bezpieczeństwa, zanim one jeszcze nastąpią. Zastosowanie proponowanego systemu może również przyspieszyć ewakuację budynku w sytuacjach zagrożenia, np. poprzez sterowanie drogami ewakuacyjnymi. Równoległe przeprowadzenie symulacji dla różnych możliwych sytuacji, oraz przetestowanie efektów różnych działań, może pomóc w podjęciu najlepszych decyzji, zarówno obsłudze budynku jak i służbom ratowniczym.

Również, wiele do zaoferowania prezentowany system ma w kontekście zwiększenia komfortu użytkownika budynku, przy minimalizacji kosztów użytkowania. Może on zostać wykorzystany do automatycznego sterowania takimi elementami jak oświetlenie, ogrzewanie czy położenie wind. Wreszcie znajomość wzorców zachowań pieszych w budynku w połączeniu z predykcją ich położenia może zostać z powodzeniem wykorzystana w działaniach marketingowych czy akcjach promocyjnych.

Bibliografia

- [1] Burstedde C.K., Klauck K., Schadschneider A., Zittartz J.: Simulation of Pedestrian Dynamics using a 2-dimensional Cellular Automaton, *Phys. Rev. A* **295** (2001) 507–525.
- [2] Dijkstra J., Jessurun A.J., Timmermans H.: A Multi-Agent Cellular Automata System for Visualising Simulated Pedestrian Activity, *Proceedings of ACRI*, (2000) 29–36.
- [3] Dudek–Dyduch E., Was J.: Knowledge Representation of Pedestrian Dynamics in Crowd. Formalism of Cellular Automata. *Proceedings of ICAISC, Lecture Notes in Artificial Intelligence* **4029** (2006) 1101–1110.
- [4] Hall E.T.: *The Hidden Dimension*. Garden City, New York (1966)
- [5] Helbing D., Molnar P.: A Social Force Model for Pedestrian Dynamic, *Phys. Rev. E* **51**, 4284–4286 (1995)
- [6] Lubaś R., Miller J., Mycek M., Porzycki J., Wąs J.: Three different approaches in pedestrian dynamics modeling, *New results n dependability and computer systems, Advances in Intelligent Systems and Computing Vol. 224*, 285-294, (2013)
- [7] Narimatsu K., Shiraishi T., Morishita S.: Acquisiting of Local Neighbour Rules in the Simulation of Pedestrian Flow by Cellular Automata, *Proceedings of 6th ACRI, LNCS* **3305**, Amsterdam (2004) 211–219
- [8] Pauls J.: *Movement of people*, DiNenno (1995)
- [9] Was J., Gudowski B., Matuszyk P.J.: Social Distances Model of Pedestrian Dynamics, *International Conference on Cellular Automata For Research & Industry, Perpignan, France, Lecture Notes in Computer Science Vol. 4173* (2006)
- [10] Wąs J.: Cellular Automata Model of Pedestrian Dynamics for Normal and Evacuation Conditions, *Proceedings of Intelligent Systems Design and Applications, Wroclaw, IEEE CS Washington Brussels Tokyo* (2005) 154-159
- [11] Wąs J., Lubaś R., Izak P.: Projektowanie ewakuacji z obiektów użyteczności publicznej z wykorzystaniem symulacji komputerowych, *Ochrona przeciwpożarowa w budownictwie : seminarium naukowo-techniczne : Zakopane, 4–6 października 2012 r., Izba Rzeczoznawców Stowarzyszenia Inżynierów i Techników Pożarnictwa, Stowarzyszenie Inżynierów i Techników Pożarnictwa. Oddział Małopolska, Komenda Wojewódzka Państwowej Straży Pożarnej w Krakowie. — Kraków : Wydawnictwo SITP. Oddział Małopolska, cop. 2012*
- [12] Helbing, D., Farkas, I., Vicsek, T.: Simulating dynamical features of escape panic. *Nature* 407 (2000) 487–490
- [13] Helbing, D., Farkas, I., Vicsek, T.: Freezing by heating in a driven mesoscopic system. *Phys. Rev. Lett.* 84 (2000) 1240–1243
- [14] Video films by Thomas Arns, Wannenstr. 22, 70199 Stuttgart, Germany.
- [15] Helbing, D. *Behavioral Science* 36, 298–310 (1991).

- [16] J. Pauls, Movement of People, The SFPE Handbook of Fire Protection Engineering, 3-263, 1995
- [17] K. Nagel and M. Schreckenberg, A cellular automaton model for freeway traffic, Journal de Physique I, 1992, 2, 2221–2229, 12
- [18] D. Helbing, P. Molnar, Social Force Model for Pedestrian Dynamics, PHYSICAL REVIEW E, 1995, 51, 4282
- [19] Nagel, Kai and Wolf, Dietrich E. and Wagner, Peter and Simon, Patrice, Two-lane traffic rules for cellular automata: A systematic approach, Phys. Rev. E, 1998, 58, 1425–1437, Aug, American Physical Society
- [20] X. Zhang, Water Filling: Unsupervised People Counting via Vertical Kinect Sensor. Advanced Video and Signal-Based Surveillance (AVSS), 2012 IEEE Ninth International Conference on. IEEE, 2012.
- [21] S. Seer, N. Brändle, C Ratti. Kinects and Human Kinetics: A New Approach for Studying Crowd Behavior. arXiv preprint arXiv:1210.2838 (2012).
- [22] L. Steyaert, Crowd analysis using TOF-technology: Blob detection. Perso-nendetectie. 2011-2012.
- [23] R. Lubaś, M. Mycek, J. Porzycki, J. Waś, Dynamic data driven simulation of pedestrian movement with automatic validation, Traffic and granular flow '13: 25–27 September 2013, Jülich, Germany.
- [24] M. Komorkiewicz, Real-Time Detection of Movement in Prohibited Direction for Video Surveillance System. Image Processing & Communications 17.4 (2012): 251-264.
- [25] V. Rabaud, S. Belongie, Counting crowded moving objects. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 1. IEEE, 2006.
- [26] G. Brostow, R. Cipolla. Unsupervised bayesian detection of independent motion in crowds. Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on. Vol. 1. IEEE, 2006.
- [27] A. M. Cheriyyadat, B. Bhaduri, R.J. Radke. Detecting multiple moving objects in crowded environments with coherent motion regions. Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on. IEEE, 2008.

Kraków 30 maja 2014

SPIN - raport z prac przeprowadzonych w celu wytworzenia polaryzacyjnego systemu projekcji stereoskopowej.

Michał Turek

1. Wstęp

Projekt ma na celu opracowanie technologii projekcji stereoskopowej stosowanej w lokalnych (kompaktowych) systemach prezentacji multimedialnych. Celem jest umożliwienie uruchomienia takiego systemu jako narzędzia promującego dowolny obiekt fizyczny w niemalże każdej organizacji (firma, urząd, instytucja kulturalna etc). Dotychczasowe systemy projekcji stereoskopowej są jedynie domeną kin trójwymiarowych lub atrybutem wyszukanych i zamkniętych dla modyfikacji produktów multimedialnych sprzedawanych seryjnie (filmy 3D, gry 3D). Realizacja projektu sprawi, iż systemy takie staną się ogólnodostępne. Planowane jest opracowanie oraz upowszechnienie gotowego zestawu rozwiązań technicznych umożliwiających stworzenie systemu prezentacji stereoskopowej opartego o projekторы multimedialne, specjalnie stworzone oprogramowanie renderujące, system filtrowania polaryzacyjnego, urządzenia HID i ekran ścienny (tzw. anty-depolaryzacyjny). Cały system będzie konfigurowany tak, aby możliwa była stereoskopowa projekcja własnoręcznie przygotowanej prezentacji trójwymiarowej. System ten ma być ponadto łatwy w budowie i aplikowalny w jak największym przedziale zastosowań, głównie promocyjnych i demonstracyjnych (na przykład jako prezentacja projektu architektonicznego, wizualizacja siedziby firmy, zabytku, wnętrza urzędu, eksponatu muzealnego itp.). Jednocześnie ma być dostępny dla większej liczby korzystających z niego odbiorców (dopuszcza jednocześnie dużą liczbę obserwatorów wspomnianej projekcji stereoskopowej).

Całość prac przeprowadzonych w ramach projektu podzielono na kilka etapów:

- Stworzenie fizycznej instalacji umożliwiającej projekcję - opartej o ekran anty-depolaryzacyjny i parę projektorów multimedialnych.
- Opracowanie technik renderowania sygnału stereoskopowego, czyli podstawowych założeń umożliwiających różnicowanie klatek animacji dla prawego i lewego oka odbiorcy, opracowanie rachunku umożliwiającego generowanie takiego sygnału wizyjnego.
- Stworzenie oprogramowania, które wykorzystując możliwości współczesnych akcelerowanych sprzętowo kart graficznych wygeneruje wysokowydajną animację trójwymiarową bazującą na dwóch niezależnych i jednocześnie prezentowanych oknach graficznych (dla prawego i lewego oka).

- Opracowanie ergonomicznego interfejsu dla użytkownika instalacji, umożliwiającego sterowanie projekcją stereoskopową.
- Zaprojektowanie narzędzia umożliwiającego rejestrowanie modeli 3D, wraz z prowadzeniem procesu ich optymalizacji.
- Rozważania na temat technik szybkiego pozyskiwania modeli 3D do prezentacji stereoskopowej na bazie pierwotnych rzeczywiście.

Część prac była zrealizowana wcześniej (co umożliwiło wstępne rozpoznanie możliwości skonstruowania przedmiotowej instalacji) i znacznie ułatwiło opracowanie ścieżki rozwiązania.

2. Opracowanie fizycznej instalacji opartej o ekran anty-depolaryzacyjny i parę projektorów multimedialnych

Konieczna do wyświetlania infrastruktura oparta jest o specjalnie przygotowany ekran (tzw. anty-depolaryzacyjny) oraz parę projektorów. Określenie "anty-depolaryzacyjny" oznacza, iż ekran posiada właściwość polegającą na utrzymywaniu polaryzacji światła od niego odbitego. Zwykły materiał typowo wykorzystywany jako ekran (biała powierzchnia) powoduje, że spolaryzowane światło odbijające się od takiego ekranu traci polaryzację. Wspomnianą właściwość utrzymują tzw. ekrany srebrne, do których wytworzenia stosuje się lakier metalizowany. Jednak nie każdy tego typu lakier spełnia wymagania. W czasie wstępnych eksperymentów okazało się, że istotna jest na przykład granulacja proszku aluminiowego, stosowanego jako podstawowy składnik lakieru metalizowanego.

Wcześniejsze prace (przygotowawcze) dotyczyły ustalenia:

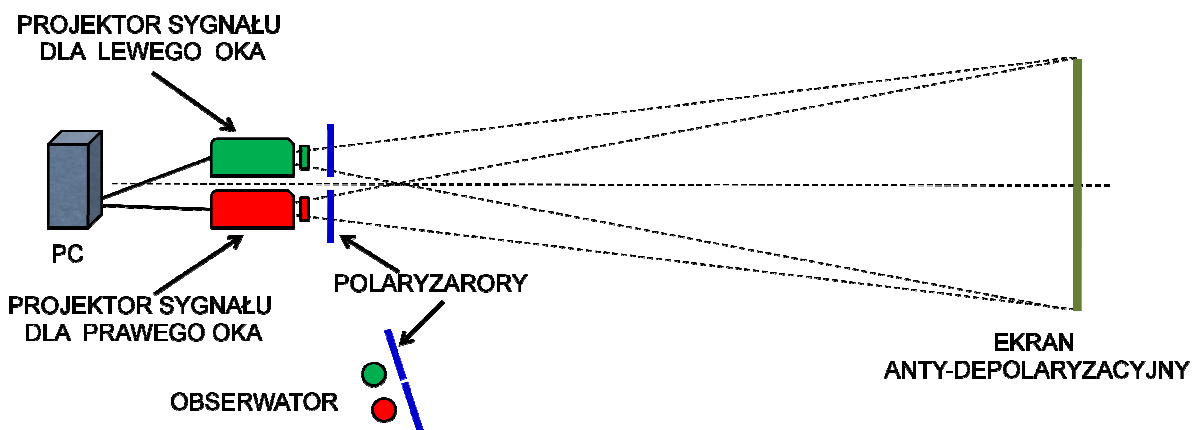
- optymalnego składu pigmentu minimalizującego negatywne skutki depolaryzacji światła odbitego (granulacji proszku aluminiowego w lakierze bezbarwnym poliuretanowym),
- właściwego podkładu - nie wpływającego na barwę światła odbitego,
- domieszek utrzymujących właściwą (zgodną z oryginałem) temperaturę światła białego odbitego od ekranu,
- domieszek umożliwiających ocieplenie barwy odbitego światła - konieczne do zrekompensowania ubytku siły składowych ciepłych po odbiciu od ekranu srebrnego.

Odpowiednią granulację proszku aluminiowego dobrano na podstawie testów serii kilkunastu różnych próbek takich lakierów, gdzie oświetlano polakierowaną powierzchnię światłem spolaryzowanym (pochodzącym z projektora), tłumiono światło odbite za pomocą polaryzatora umieszczonego poprzecznie do polaryzatora wcześniejszego i finalnie mierzono siłę światła światłomierzem fotograficznym. Do testów stosowano światło białe (zawierające wszystkie postrzegane przez człowieka

składowe). Najlepsze wyniki (największe tłumienie polaryzatorów) osiągnięto dla granulacji aluminium 40 um. Lakier z takiej próbki wykorzystano do budowy ekranu w instalacji doświadczalnej. Zastosowanie granulacji z jeszcze większym ziarnem nie pogarszało wyniku związanego z tłumieniem, jednak wpływało niekorzystnie na utrzymanie wierności odwzorowania barw (światło odbite rozpraszało się w przypadkowych kierunkach ulegając rozszczepieniu).

Po dobraniu granulacji konieczne okazało się także użycie właściwego podkładu pod warstwę lakieru metalizowanego. Lakier taki oświetlany silną wiązką światła (materiałem kryjącym jest jedynie proszek aluminiowy) nie izolował od tego światła powierzchni podkładu. Wykorzystano ten fakt - stosując jako podkład lakier w kolorze żółtym. Spowodowało to ocieplenie barwy odbitego światła, znacznie poprawiając temperaturę bieli zniekształconą wcześniej przez aluminium (kolor srebrny). Do lakieru metalizowanego także dodano niewielką ilość pigmentu w kolorze żółtym.

W instalacji zastosowano standardowe projektory DLP firmy BenQ (model MP 511) [5]. Konieczne jest tu zastosowanie technologii DLP [9], gdyż projektory tego typu jako jedyne emitują światło nie polaryzowane (w projektorach typu LCD lub reflective-LCD prowadzona jest polaryzacja światła wewnątrz urządzenia).



Rys 1. Schemat stworzonej instalacji doświadczalnej

Rozdzielenie sygnału wymagało wytworzenia i dobrania par polaryzatorów, zlokalizowanych w okularach oraz na drodze pomiędzy źródłem światła a ekranem (Rys. 1) [7]. Polaryzatory występują w wariacie umożliwiającym wymuszenie polaryzacji liniowej lub spiralnej. Przeprowadzono testy tłumienia światła dla obydwu wariantów, a ponieważ wyniki były porównywalne - wybrano polaryzatory liniowe (jako łatwiej dostępne) [10]. Dodatkowo warstwy polaryzujące światło (w polaryzatorach) można nanosić na szkło lub przezroczach polimerowych. Pierwszy wariant jest bardziej trwały, drugi łatwiejszy w obróbce i łatwiej adaptowalny do wszelkiego typu opravek. Tu właśnie trudności z osadzeniem polaryzatorów szklanych z okularach spowodowały, że wybrano polaryzatory foliowe (przezroczce polimerowe o grubości 0,8 mm z naniesioną warstwą polaryzującą światło). Takie same polaryzatory zamontowano przez projektorami. Konieczne było odsunięcie polaryzatorów od

obiektów projektorów (dystans około 10 cm), gdyż z powodu wysokiej temperatury wytwarzanej przez żarówkę - odkształcały się w krótkim czasie. W czasie osadzania polaryzatorów rozważono także możliwość wbudowania ich do wnętrza projektora, lecz ten sam czynnik - wysoka temperatura panująca w urządzeniu - wykluczyło to rozwiązanie.

Ponieważ polaryzatory znajdują się pomiędzy projektorem a ekranem i zależnie od wyboru oka mają na celu polaryzowanie (poziome lub pionowe) światła - konieczne jest podanie na nie światła jeszcze nie spolaryzowanego. Takie światło emitują tylko projektory typu DLP i jedynie takie nadają się do pracy w instalacji. Przeprowadzono także eksperymenty z projektorami typu reflective LCD [8]. W tego typu projektorach układ optyczny rozszczepia światło z żarówki na trzy niezależne kanały i przepuszcza je przez trzy niezależne matryce LCD. Każda z matryc odpowiada za jedną ze składowych (RGB) pikseli obrazu. Każda polaryzuje światło - zależnie od kanału poziomo lub pionowo. Polaryzacja światła z poszczególnych kanałów nie jest jednak zunifikowana (np. dla składowej zielonej jest pionowa, dla pozostałych dwóch - pozioma). Wymuszono to na etapie konstruowania projektora odpowiednim ułożeniem polaryzatorów przed i za matrycą LCD. W ramach eksperymentów dokonano próby unifikowania polaryzacji składowych RGB - starając się wytworzyć dwa projektory posiadające natywnie zunifikowaną polaryzację pionową (w jednym z nich) i poziomą (w drugim). Nie było jednak wykonalne (Toshiba, Hitachi), gdyż niemożliwy okazał się fizyczny demontaż polaryzatorów (odklejenie ich od matryc) w celu ich obrócenia - tą ścieżkę rozwiązania zatem zarzucono.

Finalnie stworzono klatkę dla dwóch projektorów DLP, umożliwiającą ich stabilne posadowienie - tak, aby pod wpływem wstrząsu nie przemieszczały się względem siebie. Ustawiono je nad sobą tak, aby emitowały obrazy dokładnie w to samo miejsce ekranu anty-depolaryzacyjnego. Następnie osadzono przez nimi dwa polaryzatory: z polaryzacją poziomą dla prawego oka, oraz z pionową dla lewego. Linie sygnałowe projektorów podłączono do dwóch wyjść karty graficznej komputera PC. Zastosowano identyczne konfiguracje temperatury barw, identyczne okablowanie oraz identyczne rozdzielczości prezentacji - aby maksymalnie zunifikować obrazy dla prawego i lewego oka. Uboczną cechą (i zaletą) rozwiązania jest to, że projektory (identyczne) mogą być sterowane równocześnie przy pomocy jednego pilota, a ich konfiguracja jest ciągle identyczna.

3. Opracowanie technik renderowania sygnału stereoskopowego

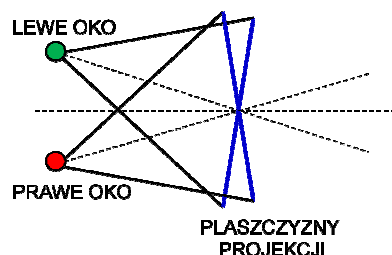
Jak wiadomo, oprogramowanie które przygotowuje sygnał wizyjny dla systemu projekcji stereoskopowej opartego o polaryzację światła musi przygotować dwa okna graficzne (dwa obrazy) jednocześnie - z przeznaczeniem dla prawego i lewego oka. Treść obydwu obrazów będzie różnicowana tak, aby symulować rozłokowanie geometryczne dwojga

oczu (dwóch kamer) z którymi dane obrazy będą powiązane. Przy różnicowaniu treści obrazów dla prawego i lewego oka wykorzystano rachunek tzw. macierzy projekcji (umożliwia on sprawne korygowanie przestrzeni projekcji). Zdefiniowano dzięki temu globalną transformatę, która będzie aplikowana do odpowiedniej sytuacji korygując ułożenie obiektów w przestrzeni projekcji. Macierz projekcji (będąca faktycznie tablicą liczb zmiennoprzecinkowych o wymiarach 4×4) zawiera zakodowaną transformatę, będącą złożeniem:

- transformaty zapewniającej wybrane rzutowanie projekcji (*perspective*, *ortho*, *frustum* itp.)
- transpozycji w osi X pozycji kamery w prawo bądź w lewo (zależnie od wyboru oka) w lokalnym układzie współrzędnych. Symuluje to rozplanowanie poziome oczu (gdyż oczy - a w konsekwencji kamery je emulujące nie znajdują się w tym samym położeniu)
- kadrowania kamery, które jest skonfigurowane tak, aby kamera spoglądała na punkt odległy od niej zgodnie z osią Z (głębokości) w lokalnym układzie współrzędnych
- korekt mających na celu poziome przesunięcie płaszczyzny projekcji stereoskopowej w przypadku stosowania projekcji typu *frustum*

Wspomniane złożenie transformat jest osiągnięte poprzez zwykłe mnożenie macierzy przez siebie. Należy tu zaznaczyć, że kolejność mnożenia będzie jednocześnie kolejnością wykonywanych operacji transformacji geometrycznych. Dotyczyć ona będzie wszystkich siatek wielokątów będących elementarnymi komponentami obiektów trójwymiarowych prezentowanych w przyszłej przestrzeni projekcji sceny stereoskopowej. Wektor pozycji dowolnego wierzchołka w takiej siatce mnożony przez wspomnianą macierz przybierze nową postać - opisującą nową pozycję wierzchołka po zaaplikowaniu transformaty (wykonaniu transformacji macierzą).

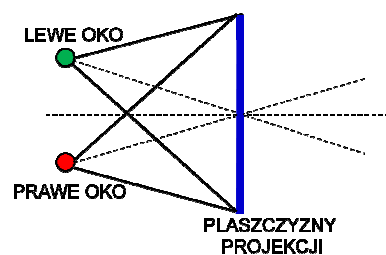
Do celów doświadczalnych stworzono dwa niezależne systemy rzutowania przestrzeni projekcji; perspektywiczny (*perspective projection*) i oparty na ostrosłupie (*frustum projection*). Pierwszy wariant posiada wady geometrii, jednak często wprowadzany jest w grach 3D - wprowadzane przez niego zniekształcenia dają efekty podobne do powodowanych przez niedoskonałości układów optycznych kamer. Pozwala to na osiągnięcie większego stopnia naturalności sceny sugerując, że jej źródłem są kamery (rys. 2).



Rys 2. Geometria płaszczyzn projekcji przy zastosowaniu *perspective projection*

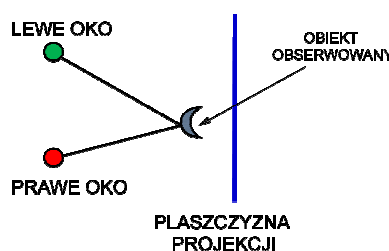
Dostępne w bibliotekach 3D funkcje (metody) umożliwiające generowanie macierzy transformacji dla *perspective projection* pozwalają na określenie "ogniskowej kamery" emulowanej w takiej transformacji. Przeprowadzono eksperymenty z użyciem różnych ogniskowych dobierając najbardziej naturalną i skutkującą wyprodukowaniem najbardziej atrakcyjnej graficznie prezentacji.

W przypadku generowania macierzy transformacji dla trybu *frustum projection* możliwości bibliotek 3D są większe. Przede wszystkim możliwe jest wprowadzenie asymetrycznego odwzorowania przestrzeni 3D dla każdego z oczu. Jest to istotne z uwagi na fakt, iż obserwator sceny postrzegając głębię przestrzeni wytworzoną na skutek stosowania projekcji stereoskopowej faktycznie nadal patrzy na zwykłą płaszczyznę (ekran). Przy budowie macierzy projekcji w wariantcie *frustum* konieczne jest dodatkowe korygowanie symetrii obrazu prawego i lewego - aby użytkownik mógł spoglądać na dwa równoległe obrazy, a nie obrazy przesunięte kątowno tzw. „*Parallel axis asymmetric projection*” [1]. Jest to najbardziej prawidłowa postać odwzorowania i powinna być stosowana w dalszych doświadczeniach (Rys 3) - jako wariant matematycznie idealny lepiej nadaje się do zastosowań stereoskopowych.

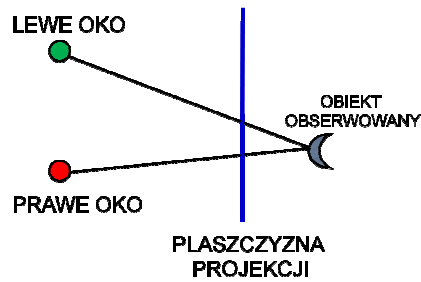


Rys 3. Geometria płaszczyzn projekcji przy zastosowaniu *frustum projection* z przesunięciem (*Parallel axis asymmetric projection*)

Kolejnym problemem związanym z uzyskaniem maksymalnego poziomu poprawności odwzorowania, jest brak bieżącej wiedzy o odległości punktu, na który aktualnie patrzy obserwator sceny. Nie jest zatem możliwe wygenerowanie zawsze właściwej treści macierzy transformacji. Gdy użytkownik patrzy bliżej, niż założono przy konfigurowaniu systemu projekcji stereoskopowej, mamy do czynienia z tzw. dodatnią paralaksą przy postrzeganiu stereoskopowym (Rys. 4). Gdy jest odwrotnie (obserwator patrzy dalej niż założono) - powstaje paralaksa ujemna (Rys. 5).

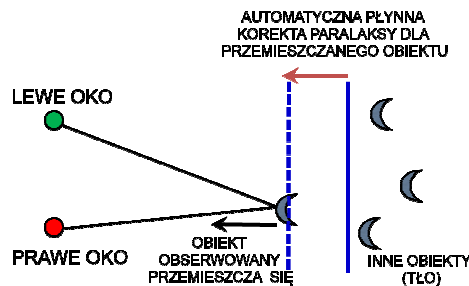


Rys. 4: Dodatnia paralaksa w projekcji stereoskopowej: Użytkownik spogląda na obiekt znajdujący się bliżej niż założona płaszczyzna projekcji.



Rys. 5: Ujemna paralaksa w projekcji stereoskopowej: Użytkownik spogląda na obiekt znajdujący się dalej niż założona płaszczyzna projekcji.

Nie jest możliwe pozyskanie wiedzy o odległości obiektu będącego w centrum zainteresowania. Teoretycznie - można spowodować przeniesienie uwagi na taki punkt wymuszając ją ruchem obiektów w scenie stereoskopowej lub innym podobnym środkiem. Ciekawym wydaje się rozwiązanie umożliwiające animowane płynnej zmiany wartości macierzy transformacji dla projekcji stereoskopowej - w celu przyciągnięcia uwagi. Obrazuje to rysunek 6.



Rys. 6: Animowane (płynne) przemieszczanie płaszczyzny projekcji stereoskopowej w celu ogniskowania uwagi na konkretnym elemencie przestrzeni projekcji.

Wprowadzenie systemu animowania przemieszczania płaszczyzny projekcji uwzględnia naturalne możliwości człowieka w dziedzinie postrzegania głębi stereoskopowej. Ograniczenia postrzegania związane z tempem przemieszczania uwagi pomiędzy pierwszym planem, a tłem znajdującym się daleko. Obecnie trwają prace nad dalszą optymalizacją tego procesu.

4. Optymalizacja wydajnościowa procesów renderowania dwóch klatek jednocześnie - na potrzeby projekcji stereoskopowej.

Głównym założeniem było stworzenie wysokowydajnego systemu wizualizacji, który w pełni wykorzystywałby dostępne narzędzia sprzętowe (akcelerator 3D). Typowa biblioteka graficzna służąca do renderowania scen trójwymiarowych w czasie rzeczywistym korzysta z pojedynczego potoku wielokątów oraz pojedynczego lub podwójnego bufora rasteryzacji (w opcji zakładającej wykorzystanie podwójnego buforowania). W projektowanym rozwiązaniu wszystkie te struktury muszą być

zaprojektowane tak, aby było możliwe wydajne renderowanie dwóch klatek animacji - eksponowanych jednocześnie.

Dużych problemów przysporzyło prowadzenie dwóch niezależnych i współbieżnych procesów rasteryzacji na bazie banku tekstur załadowanych wcześniej do pamięci akceleratora 3D. Okazało się, że tekstury są odwzorowywane w sposób wymagający wyłączości w zakresie dostępu procesu rasteryzującego do pamięci kadry graficznej. Liczne kolizje w dostępie do tekstur, skutkujące drastycznym spadkiem wydajności oraz "niedorysowaniami" skłoniły do stworzenia dwóch niezależnych banków tekstur - z przeznaczeniem dla prawego i lewego oka. Podobnie (tu już przewencyjnie) postąpiono w przypadku potoków wielokątów - także stworzono dwa niezależne.

Wszelkie siatki wielokątów, przypisane do materiałów powierzchni i zawarte w komponentach graficznych sceny trójwymiarowej (modelach 3D) są optymalizowane pod kątem wydajności renderowania. Optymalizacja polega na grupowaniu ich zgodnie z przydzielonym materiałem powierzchni i tworzeniem jednego zbioru wielokątów dla wszystkich kształtów przypisanych do tego samego materiału (tekstury) - niezależnie od ich pochodzenia (modelu 3D, do którego były przypisane). Macierz transformacji dla projekcji stereoskopowej (omówiona w rozdziale poprzednim) może być przekazana do Vertex shadera - programu, który pracując bezpośrednio na GPU akceleratora graficznego zastosuje ją jako transformatę projekcji stereoskopowej. Vertex shader zaimplementowano dla języka NVidia CG i można go wykorzystywać na sprzęcie od tego producenta oraz urządzeniach kompatybilnych.

5. Zaprojektowanie narzędzia umożliwiającego rejestrowanie modeli 3D, wraz z prowadzeniem procesu ich optymalizacji

Kolejnym problemem w całości rozwiązany w ramach prac było dostarczenie mechanizmu, który w prosty dla jego użytkownika sposób pozwoli na definiowanie treści sceny 3D. Tu przyjęto założenie, iż system będzie służył głównie do prezentowania jednostkowych obiektów trójwymiarowych reprezentowanych przez bryły sztywne lub zbiory takich brył (obiekty z ruchomymi elementami). Dopuszczono także wprowadzenie prostego systemu animowania, bazującego na eksponowaniu samopowtarzalnej sekwencji brył sztywnych składającej się na taką animację. jako nośnik plikowy wybrano format 3ds - bardzo popularny, umożliwiający przeniesienie informacji o kształcie bryły, cechach jej powierzchni (materiałach) a także o teksturach (wraz z koordynatami tekstur przypisanymi do wierzchołków trójkątów w zakodowanych w pliku siatkach 3D). Format ma postać binarną, co skraca czas ładowania prezentacji.

W systemie projekcji zaimplementowano mechanizmy auto-zoom i auto-transpose dla kamery, automatycznie (w rybie płynnej animacji) pozycjonujące rejestrowany przez użytkownika obiekt w kadrze kamery.

Ponieważ plik zawierający model 3D (także w formacie 3DS) nie określa żadnego przelicznika skali, jaką należy zastosować wobec obiektu (finalny rozmiar obiektu zadany jest jedynie współrzędnymi wierzchołków składających się na siatki 3D opisujące bryłę) - konieczne było stworzenie transformacji unifikującej skalę przestrzeni projekcji wykorzystaną prezentacji.

6. Stworzenie interfejsu umożliwiającego interaktywne wprowadzanie wizualnych efektów specjalnych do prezentacji stereoskopowej

Jak wspomniano, do projektu zaangażowano zaawansowany silnik 3D napisany nad OpenGL i stworzony od podstaw. Jego funkcjonalność pozwala na pełną kontrolę przestrzeni projekcji, także emulowanie licznych efektów specjalnych związanych z oświetleniem sceny, elementami mgły, rozmywania, przezroczystości itp. Wszystkie te efekty wprowadzono w systemie prezentacji stereoskopowej.

Wygenerowane gotowe sceny stereoskopowe są interaktywne. Umożliwiono w nich przemieszczanie kamery, modyfikację kadru, wprowadzono system kontroli położenia kamery z uwzględnieniem konfigurowanych każdorazowo stref dla kamery zabronionych. W obecnej wersji jako urządzenie HID wykorzystano standardową mysz PC, naśladując ergonomię interfejsu stosowaną w większości gier 3D: kamera przemieszcza się po powierzchni wirtualnego terenu, nad którym znajduje się prezentowany obiekt. System re-pozycjonowania kamery umożliwia "rozglądanie się" przemieszczając punkt, na który patrzy kamera po okręgu wokół niej oraz dodatkowo po osi Y (pionowej).

Dodatkowo - wyprowadzono z systemu interfejs umożliwiający modyfikowanie cech prezentowanych w scenie obiektów wyróżnionych - poprzez wykorzystanie różnorodnych zdarzeń generowanych przez urządzenia HID stacji PC. Interfejs ten umożliwia również sterowanie dodatkowymi funkcjami korygującymi globalne parametry prezentacji: nanoszenie mgły, korekty oświetlenia czy parametrów projekcji.

Dodano także kolejny interfejs - pozwalający na wyprowadzanie (wyświetlanie) wartości parametrów projekcji stereoskopowej i interaktywne modyfikowanie ich podczas pracy systemu.

7. Badania w kierunku uzyskania optymalnych metod pozyskiwania modeli 3D

Duży problem stanowi pozyskanie modeli trójwymiarowych będących wizualnymi odpowiednikami rzeczywistych obiektów będących docelowo przedmiotami prezentacji. Wysokowydajne obrazowanie obiektu rzeczywistego bryłą trójwymiarową zdatną do

zintegrowania z projektowanym systemem prezentacji stereoskopowej wymaga dodatkowo odpowiedniego jej formatowania (reprezentowania). Nie ulega wątpliwości, że najbardziej popularnym, elastycznym, wydajnym i precyzyjnym systemem reprezentowania kształtów 3D są siatki wielokątów. Należy zatem rozważyć możliwe drogi pozyskiwania takich siatek. Gdy pierwowzór (obiekt rzeczywisty) posiada stosunkowo prostą geometrię, możliwe jest jego wymodelowanie w ogólnie dostępnych narzędziach do obróbki kształtów 3D. Dostatecznie wierny oryginałowi odpowiednik można następnie umieścić w stworzonym systemie. Jednak w przypadkach mniej dogodnych (a takie z pewnością będą stanowiły większość) koniecznej jest obrazowanie przy użyciu urządzeń skanujących. Obiekty o niewielkich rozmiarach i możliwe do przemieszczania najwygodniej jest skanować przy użyciu skanerów rotacyjnych[4]. Idea takiego skanowania polega na oświetlaniu obiektu wiązką laserową poziomo pod różnymi kątami (obiekt jest obracany wokół swojej osi Y) i konsekwentnym odczytywaniu odległości urządzenia skanującego od powierzchni obiektu. Po wykonaniu pełnego "okręgu" wokół skanowanego obiektu uzyskujemy jego obwiednię złożoną z pozycji wierzchołków na jego powierzchni i na zadanej płaszczyźnie poziomej. Następnie przemieszczamy urządzenie skanujące nieznacznie po osi Y i powtarzamy proces - uzyskując druga obwiednię. Proces powtarzamy wielokrotnie uzyskując siatkę 3D.

Inna popularną dziś techniką skanowania 3D jest analiza różnicowa obrazu 2D pochodzącego z dwóch kamer rozstawionych kątowno (ich ułożenie jest znane) i obserwujących obiekt pod różnymi kątami [6]. Wynik (w tym przypadku) - to jedynie fragment powierzchni bryły wyeksponowany w stronę kamer. Konieczne jest zatem wielokrotne skanowanie obiektu - w celu uzyskania kompletnej siatki 3D obrazującej ten obiekt.

Kolejne rozwiązanie zakłada emisję specjalnie spreparowanej maski graficznej, oświetlając jej treścią skanowany obiekt (przy użyciu projektora multimedialnego). Jednocześnie obiekt jest obserwowany przez kamerę [3] a na podstawie zniekształconego obrazu maski (po odbiciu od skanowanego obiektu) pozyskiwany jest kształt 3D.

Niezależnie od wybranej techniki - wyniki skanowania (dzięki wprowadzonym w systemie procesom automatycznego skalowania i pozycjonowania kamer oraz zastosowaniu popularnego formatu opisu kształtów trójwymiarowych) będą kompatybilne bez jakiegokolwiek obróbki. Ta naturalnie jest możliwa - jednak dotyczyć będzie głównie modyfikowania cech powierzchni obiektu (kolory, tekstura) - w celu zwiększenia atrakcyjności całej prezentacji.

W ramach eksperymentu przeprowadzono pełny proces przygotowania i integracji modelu trójwymiarowego w stworzonym systemie. Wykorzystano tu skaner rotacyjny firmy Roland [4] i kilka przedmiotów pełniących rolę doświadczalnych pierwowzorów fizycznych. Wyniki były zadowalające

Naturalnie - poza angażowaniem różnorodnych technik skanowania trójwymiarowego - możliwe jest użycie wszelkich innych dostępnych metod pozyskiwania modeli trójwymiarowych . Ma to szczególne znaczenie w sytuacji, gdy przedmiotem prezentacji będzie model już istniejący - a wytworzony na drodze interaktywnego wizualnego modelowania 3D (np. projekt architektoniczny nowego budynku, model/wirtualny prototyp produktu itp.). Adaptacja w tym przypadku może się sprowadzać jedynie do automatycznej redukcji liczby wielokątów w przypadku gdy ich ilość mogłaby negatywnie wpłynąć na wydajność renderowania animacji stereoskopowej w czasie rzeczywistym

Materiały źródłowe

[1] Steinicke, "Realistic Perspective Projections for Virtual Objects and Environments", ACM Transactions on Graphics, Vol. 30, No. 5, Article 112, Publication date: October 2011.

[2] Rotary Scanners, <http://www.shapegrabber.com/sol-products-3d-auto-inspection-Ai310.shtml>

[3] 3D WebCam Scanners, <http://www.david-3d.com/>

[4] Roland 3D laser Scanners, <http://www.rolanddga.com/products/scanners/>

[5] BenQ Projectors, <http://benq.com>

[6] Mobile 3D Scanners <http://www.creaform3d.com/en/ndt-solutions/handyscan-3d-laser-scanners>

[7] Polarizers, <http://en.wikipedia.org/wiki/Polarizer>

[8] Reflective LCD projectors, <http://www.pcmag.com/article2/0,2817,1157687,00.asp>

[9] DLP projection, http://en.wikipedia.org/wiki/Digital_Light_Processing

[10] 3DExpert, http://en.wikipedia.org/wiki/Digital_Light_Processing