



Prof. dr hab. Ewa Grabska  
Zakład Projektowania i Grafiki Komputerowej  
Wydział Fizyki, Astronomii i Informatyki Stosowanej  
Uniwersytet Jagielloński

Kraków, 10. 08. 2017

## RECENZJA ROZPRAWY DOKTORSKIEJ

- Tytuł rozprawy:** Automatyzacja procesu przenoszenia mimiki aktora na docelową siatkę twarzy
- Autor rozprawy:** mgr inż. Damian Pęszor
- Promotor rozprawy:** prof. dr hab. inż. Konrad Wojciechowski

### I. Informacje ogólne

Oceniana rozprawa doktorska mgr inż. Damiana Pęszora zatytułowana „Automatyzacja procesu przenoszenia mimiki aktora na docelową siatkę twarzy” stanowi opracowanie o objętości 129 stron, zawierające 64 rysunki, jedną tabelę, spis literatury obejmujący 50 pozycji oraz elektroniczną wersję rozprawy w formacie PDF na płycie CD.

### II. O przedmiocie rozprawy

Rozprawa pana mgr inż. Damiana Pęszora dotyczy automatyzacji procesu animacji modelu ludzkiej twarzy w wysokiej rozdzielczości z użyciem technik akwizycji ruchu.

Najbardziej znaną metodą sterowania ruchem jest *animacja artystyczna*, w której animator odpowiada za kształtowanie ruchu. Podstawą tej metody jest interpolacja. Wykreowanie realistycznego ruchu obiektu przez animatora należy do bardzo pracochłonnych zadań wymagających dużej koncentracji uwagi. Obecnie coraz częściej stosowana jest *animacja sterowana za pomocą danych*. Łatwość zarejestrowania (przechwycenia) ruchu obiektu fizycznego i przeniesienia go na model cyfrowy obiektu (ang. motion capture) jest szybszą metodą dającą zadowalające rezultaty w wielu zastosowaniach. Etapami reje-

stracji ruchu są jego pomiary, zamiana na postać cyfrową i zapamiętanie ruchu obiektu. Do pomiarów ruchu wykorzystuje się czujniki elektromagnetyczne, znaczniki optyczne, czy szkielety cyfrowe. Innymi słowy ruch rzeczywisty w tej metodzie jest zamieniany na dane numeryczne, a następnie odwzorowywany na model. Inna metoda sterowania ruchem polega na animacji proceduralnej wykorzystującej model obliczeniowy, np. symulacje fizyczne czy modelowanie i symulacje zachowań.

W rozprawie Autor zajmuje się animacją i odtwarzaniem mimiki twarzy. Używana jest zaawansowana technika zwana przechwytywaniem występu aktora (ang. performance capture), która poza animacją przechwytuje również mimikę aktora.

W wielu wypadkach stosowane do tej pory techniki animacji twarzy nie są w pełni zadawalające zarówno jeżeli idzie o jakość animacji, jak i czasochłonność i poniesione koszty.

### **III. O postawionym w rozprawie celach**

Cel jaki postawił sobie w rozprawie doktorant to zaproponowanie i zaimplementowanie systemu komputerowego, którego rolą byłoby automatyczne wspomaganie animacji twarzy we wszystkich etapach, od wstępnego przetwarzania modelu twarzy do wynikowej animacji twarzy. Zastosowane zostały techniki, które polegają na umieszczeniu na skórze twarzy aktora znaczników odbijających światło podczerwone oraz analizie barwy skóry nagrywanej w formie wideo. Istotnym elementem w poszerzeniu zakresu zastosowań zaproponowanego systemu było wzięcie pod uwagę antropometrycznych cech twarzy pozwalających na generalizację proponowanych rozwiązań.

Cel postawiony przez Autora należy do podstawowych z dziedziny animacji.

### **IV. Przegląd zawartości rozprawy**

**Rozdział 1** rozważa rolę wirtualnych postaci z reakcjami naśladowczymi zachowania ludzkie w zastosowaniach takich jak symulacja i gry komputerowe oraz robotyka. Okazuje się, że w dążeniu do wizualnego realizmu największe trudności w animacji postaci 3D związane są z animacją twarzy i odtworzeniem naturalnej mimiki. Powszechnie stosuje się więc animację sterowaną przez dane odnoszącą się do samej twarzy (ang. performance capture). Polega ona na umieszczeniu na skórze twarzy aktora znaczników, które odbijają światło podczerwone lub analizie barwy skóry, która jest nagrywana w formie wideo. Na zakończenie rozdziału Autor formułuje cele, które pragnie zrealizować w rozprawie.

**Rozdział 2** przedstawia techniki modelowania obiektów w przestrzeni trójwymiarowej. Opisane są metody modelowania za pomocą funkcji sklejana, podziału powierzchni oraz siatki wielokątów. Autor charakteryzuje poszczególne metody i uzasadnia wybór reprezentacji modelu twarzy za pomocą siatki wielokątów. Ten typ reprezentacji jest szeroko stosowany w przemyśle rozryw-

kowym. Twarz reprezentowaną w ten sposób można łatwo poddawać różnym modyfikacjom. Odrębnym problemem jest pozyskiwanie realistycznego modelu twarzy w stanie statycznym. Obecnie taki model uzyskuje się poprzez akwizycję rzeczywistych danych (np. skanowanie z użyciem skanera 3D). Wymagania, które ma spełniać poprawna animacja są następujące. Po pierwsze, każda użyta w animacji reprezentacja twarzy, tzw. element statyczny animacji powinien wyglądać realistycznie. Po drugie, należy uzyskać realistyczną reprezentację ruchu. Po trzecie, wymagana jest realistyczna reprezentacja cech osobniczych. W dalszej części rozdziału Autor analizuje rolę animacji artystycznej oraz animacji sterowanej przez dane w uzyskaniu poprawnej animacji twarzy. Na zakończenie naszkicowane są etapy procesu automatycznej animacji twarzy, które stanowią tematykę dalszej części rozprawy.

**Rozdział 3** ma charakter przeglądu, ale również dostarcza informacji na temat metody użycia siatki wielokątów do reprezentacji twarzy. Autor porównuje pionierską metodę opisaną animacji za pomocą interpolacji pomiędzy dwoma siatkami reprezentującymi klatki kluczowe, a automatycznym odczytywaniem położenia znaczników odbijających światło podczerwone. Metoda pionierska bazuje na pozyskaniu informacji o przemieszczeniu się wybranych wierzchołków z klatek kluczowych, natomiast w metodzie ze znacznikami każdy ze znaczników steruje większą liczbą wierzchołków siatki w swoim otoczeniu. Rozdział kończy rozważania na temat komputerowych metod animacji wspomaganých przez animatora umożliwiających lokalne zmiany mimiki oraz modyfikację tekstur.

**Rozdział 4** przedstawia zaproponowaną w rozprawie koncepcję animacji twarzy, która polega na połączeniu dwóch technik animacji, z których jedna jest oparta o znaczniki, natomiast druga o dane wideo. Celem połączenia technik jest wyeliminowanie wad poszczególnych technik oraz wzmocnienie ich zalet.

**Rozdział 5** prezentuje rozwiązania najważniejszych zagadnień związanych z proponowaną metodą animacji twarzy. Dla techniki animacji z użyciem znaczników podstawowym zagadnieniem jest wybór metody otrzymywania modelu twarzy i wydobywanie informacji o zmianie położenia znaczników w trakcie animacji modelu. Autor rozprawy zdecydował się na wybór modeli twarzy uzyskanych za pomocą trójwymiarowego skanera wykorzystującego jako metodę wizyjną światło strukturalne. Siatka trójkątów reprezentująca twarz uzyskana w procesie skanowania jest zmodyfikowana do potrzeb animacji za pomocą algorytmu przeszukiwania wszerz działającym na grafie reprezentującym siatkę. W ten sposób usuwa się z modelu części, które nie są poddawane animacji. Zarejestrowany model może zawierać dziury w siatce wynikające z rozproszenia światła. Do wypełniania dziur zastosowano dwa istniejące algorytmy. Pierwszy algorytm zaawansowanej metody frontальной generacji siatki wypełnia dziurę obszarem płaskim złożonym z trójkątów, a drugi algorytm dostosowuje ten obszar do kształtu modelu. Następnym krokiem po przygotowaniu modelu twarzy do animacji jest detekcja na podstawie cech antropometrycznych punktów charakterystycznych twarzy wykorzystywanych do lokalizacji znaczników, niezbędnych do sterowania ekspresją mimiczną. Autor szczegó-

łowo opisuje algorytmy wykrywania punktów charakterystycznych i algorytmy dopasowania animowanego modelu do modelu referencyjnego z wierzchołkami odpowiadającymi punktom charakterystycznym i położeniu znaczników na twarzy aktora. Animowany model twarzy, zostaje podzielony na segmenty, których wierzchołki reprezentują znaczniki. Przynależność wierzchołków do segmentów nie jest jednoznaczna. Autor opisuje dwa algorytmy rozwiązujące ten problem.

Druga technika animacji twarzy oparta o dane wideo używana jest w sytuacji, gdy animowany model ma reprezentować aktora. Zgodność modelu z klatką filmu zarejestrowaną na początku stanowi podstawę zdefiniowania odpowiedniości pomiędzy wierzchołkami modelu a pikselami z nagrania. W każdej kolejnej klatce filmu zachowanie tej odpowiedniości wymaga odpowiedniej modyfikacji funkcji przekształcenia wierzchołków.

W połączeniu technik animacji istotną rolę odgrywa potraktowanie techniki opartej o dane wideo jako źródła danych dla metody animacji na podstawie znaczników. Autor dokonuje interesującej analizy danych dla aplikacji korzystającej z animacji wysokiej jakości. Okazuje się, że w takim przypadku, aby otrzymać realizowalną animację komputerową należy wybrać najbardziej reprezentatywne próbki danych, które odpowiadają znacznikom używanym w animacji z trajektorią ruchu znaczników.

**Rozdział 6** prezentuje przykładowe wyniki uzyskane za pomocą opracowanej przez Autora metody animacji twarzy.

**Rozdział 7** zawiera podsumowanie oraz pokazuje wybrane kierunki kontynuacji badań.

## **V. Ocena wyników rozprawy**

Trzeba stwierdzić, że cel rozprawy został osiągnięty. Autor zaproponował i wstępnie przetestował automatyczną metodę animacji twarzy na podstawie danych ruchu. Wykorzystał dwa systemy akwizycji ruchu: system, w którym używane są znaczniki odbijające światło podczerwone oraz system wykorzystujący kamery światła widzialnego. Połączył dwie techniki animacji oparte o te systemy i w ten sposób wyeliminował najistotniejsze wady poszczególnych technik oraz wzmocnił ich zalety.

Mocną stroną rozprawy jest niewątpliwie podjęcie bardzo ambitnej tematyki dotyczącej automatycznej metody animacji twarzy. Autor najwyraźniej zna poruszany temat i jego literaturę oraz ma własne pomysły, które umie należyście opisać i zrealizować. Zaproponowana metoda animacji twarzy została zweryfikowana praktycznie na przykładach modeli nieoświetlonych, jak i modeli dodatkowo oświetlonych i ten aspekt rozprawy jest odpowiednio udokumentowany.

Przedstawione rozwiązanie ma zarówno wartość praktyczną, jak i poznawczą. Rozprawa jest merytorycznie poprawna i spójna logicznie.

Rozprawa ma również i słabsze punkty.

1. Wysoko oceniam umiejętność Autora opisywania trudnych zagadnień w sposób zrozumiały, natomiast mam zastrzeżenia do komentowania wyrażeń matematycznych. Przykładem może być komentarz do wzoru (5.4). W dodatku A dotyczącym wyznaczenia gradientu w metodzie wizyjnej brakuje objaśnień wielu oznaczeń.
2. We wstępie Autor stwierdził, że celem pracy było stworzenie algorytmu, który umożliwi całkowitą automatyzację procesu animacji twarzy. W pracy podano wyniki uzyskane na podstawie stworzonego algorytmu, natomiast brakuje podstawowych danych dotyczących jego implementacji.

## 5. Konkluzja

W sumie nie mam jednak poważniejszych zastrzeżeń do rozprawy i oceniam ją pozytywnie. Oceniana rozprawa doktorska spełnia wymagania, jakie ustawa o stopniach i o tytule naukowym przewiduje dla rozpraw doktorskich i na tej podstawie wnoszę o dopuszczenia jej Autora – mgr inż. Mariana Pęszora – do publicznej obrony.

