

RECENZJA

rozprawy doktorskiej mgr. inż. Michała Chwesiuka pt.: „Pomiar charakterystyk czułości na kontrast w peryferyjnych obszarach widzenia oraz ich wykorzystanie do przyspieszania syntezy obrazów”

Promotor rozprawy: dr hab. inż. Radosław Mantiuk, prof. ZUT

1. Aktualność i znaczenie tematyki rozprawy

W rozprawie podjęto problem zbadania charakterystyk układu wzrokowego człowieka, które odpowiednio wykorzystane, mogą zapewnić wzrost efektywności metod syntezy i renderingu obrazów graficznych. Obecnie istnieje duże zapotrzebowanie na syntetyczne grafiki przypominające zdjęcia wykonane aparatem fotograficznym, które wraz z wizualizacjami 3D znalazły zastosowanie w medycynie, w projektowaniu, w różnego rodzaju symulacjach, a także w grach komputerowych. Jednakże, do przygotowania fotorealistycznych wizualizacji wykorzystuje się algorytmy o wysokiej złożoności, wymagające użycia komputerów o dużej mocy obliczeniowej. Wymagania takich algorytmów są szczególnie trudne do spełnienia w przypadku systemów, gdzie kluczowym czynnikiem jest czas renderowania obrazu. Dla takich obszarów poszukuje się rozwiązań, które będą w stanie sprostać ograniczeniom czasowym, przy jednoczesnym zachowaniu zadawalającej jakości obrazu. W ostatnich latach dużym zainteresowaniem cieszą się metody bazujące na analizie ruchu oka i charakterystyce przetwarzania informacji pozyskiwanych za pomocą układu wzrokowego. Dzięki dynamicznemu rozwojowi algorytmów przetwarzania obrazów oraz technologii umożliwiających śledzenie ruchu oka, w szczególności w zakresie wideo-okulografii, możliwym staje się pozyskiwanie aktualnych pozycji skupienia wzroku z dużą precyzją, w nieuciążliwy dla obserwatora sposób.

Prace we wspomnianych obszarach są ciągle rozwijane, więc podjęcie przez Autora rozprawy tematu, który rozszerza możliwości efektywnej syntezy obrazów z wykorzystaniem metod okulografii uważam za aktualne i celowe.

2. Zawartość rozprawy

Rozprawa ma charakter eksperymentalny z elementami analitycznymi. Stanowi ona spójną tematycznie całość. Składa się z pięciu rozdziałów wraz z wprowadzeniem i podsumowaniem, a każdy rozdział, poza ostatnim, kończy się własnym podsumowaniem. We wprowadzeniu Autor rozprawy krótko scharakteryzował problem oraz przedstawił motywacje uzasadniające podjęcie analizowanego tematu. Następnie zdefiniował tezę pracy:

Uwzględnienie progów widzialności informacji dla peryferyjnych obszarów widzenia umożliwia zmniejszenie częstotliwości próbkowania, a co za tym idzie przyspieszenie syntezy obrazów.

Natomiast, jako główny cel swojej pracy Autor postawił:

Pomiar charakterystyk układu wzrokowego człowieka w kontekście dostrzegalnych progów kontrastów bodźców wzrokowych, uwzględniając kierunkowość widzenia oraz kolor, a następnie opracowanie analitycznego modelu wskazującego prawdopodobieństwo spostrzeżenia bodźca o zadanym kontraście.

W mojej opinii cel pracy powinien być lekko przeformułowany. Część: „uwzględniając kierunkowość widzenia oraz kolor”, należałoby zmienić na „uwzględniając ich kolor i kierunkowość widzenia”. W pierwszej wersji nie bardzo wiadomo, o jaki kolor chodzi.

Cel główny pracy został uzupełniony o 5 celów szczegółowych, wśród których znalazły się:

- (1) *Opracowanie środowiska eksperymentalnego do pomiaru progu czułości na kontrast dla widzenia peryferyjnego. Opracowanie oraz implementacja eye trackera o częstotliwości zapewniającej precyzyjną rejestrację ruchów sakadycznych.*
- (2) *Opracowanie i wykonanie eksperymentu mierzącego czułość na kontrast dla bodźców chromatycznych i achromatycznych w peryferyjnych obszarach widzenia.*
- (3) *Analiza wyników i opracowanie modeli czułości na kontrast dla układu wzrokowego człowieka w zależności od częstotliwości sygnału, typu bodźca i kąta widzenia.*
- (4) *Opracowanie i wykonanie eksperymentu mierzącego opóźnienie percepcji obiektów spowodowane zmianą z widzenia peryferyjnego na centralne.*
- (5) *Aplikacja opracowanych modeli w algorytmach syntezy obrazów uwzględniającym (pisownia oryginalna) widzenie kierunkowe. Wykorzystanie modeli w wyświetlaczach stereoskopowych z akomodacją.*

Rozdział pierwszy stanowi kompendium wiedzy na temat anatomii i fizjologii układu wzrokowego. Przedstawiono w nim budowę oka oraz proces przetwarzania sceny, pozyskanej za pomocą wzroku. Omówiono cechy widzenia centralnego i peryferyjnego w kontekście czułości na natężenie światła, percepcji barw oraz postrzegania ruchu. W rozdziale tym scharakteryzowano także składowe ruchy oka wynikające ze sposobu działania układu wzrokowego. Zdefiniowano również pojęcie kontrastu i funkcji czułości na kontrast. Wiedza ta jest kluczowa dla udowodnienia postawionej tezy i osiągnięcia założonego celu.

W rozdziale drugim przedstawiono zestaw eksperymentów ukierunkowanych na wyznaczenie charakterystyki czułości na kontrast. Dokonany w tej części rozprawy przegląd literaturowy wskazuje, że w tym obszarze zrealizowano już szereg podobnych badań, również z użyciem technologii śledzenia ruchu oka. Badania zaprezentowane w rozprawie doktorskiej stanowią ich rozszerzenie polegające na zastosowaniu nowych bodźców o nowych kierunkach kolorów, łączących składniki chromatyczne i achromatyczne. Eksperymenty, które wymuszają utrzymanie wzroku we wskazanym miejscu są niekomfortowe dla obserwatora, ze względu na nienaturalny sposób pracy układu wzrokowego. Pojawienie się bodźca poza obszarem widzenia centralnego powoduje naturalny odruch spojrzenia w jego kierunku. Odruch ten może spowodować niewłaściwą realizację eksperymentu, skutkującą dokonaniem pomiarów odpowiadających zakresowi widzenia centralnego

a nie, zgodnie z założeniami, widzenia peryferyjnego. Z tego względu w zaproponowanych w rozprawie testach, zmieniono wykorzystywany dotychczas sposób prezentacji bodźca w obszarze peryferyjnym z migającego, ograniczonego w czasie, na ciągły, a kontrolę kierunku spojrzenia zapewniono poprzez użycie okulografu (ang. *eye tracker*). W celu uniknięcia niepoprawnych pomiarów, w momencie wykrycia niepożądanego ruchu oczu, stosowano mechanizm wygaszania bodźca. Zapewniono w ten sposób warunki dla realizacji dokładniejszych pomiarów czułości na kontrast w obu obszarach widzenia. Dodatkowo, dla zwiększenia efektywności eksperymentów użyto metody psychometrycznej, która pozwoliła na adaptacyjny, wynikający z przebiegu eksperymentu, sposób doboru kontrastu wyświetlanego bodźca.

W dalszej kolejności zaproponowano model czułości na kontrast dla obliczenia prawdopodobieństwa wykrycia bodźca wzrokowego w scenie. Opracowanie tego modelu inspirowane było wcześniejszymi badaniami opublikowanymi w literaturze, w stosunku do których model ten został uproszczony. Stanowi on prostą zależność liniową czułości na kontrast od wartości kontrastu oraz odległości bodźca od centralnego obszaru widzenia. Pominięto w nim wcześniej stosowaną dodatkowo luminancję tła.

Osobną część rozdziału drugiego stanowi analiza dynamiki percepcji przy zmianie kierunku widzenia. Badania te ukierunkowane zostały na wykrycie zależności czasowych pomiędzy szybkością zmiany kierunku widzenia z peryferyjnego na centralny a wymaganą szybkością wygenerowania sceny w systemie renderingu czasu rzeczywistego. Dla oceny tej zależności zbudowany został, w oparciu o technologię śledzenia ruchu oka, eksperyment, w którym reprezentacja geometryczna bodźca dostosowywana była do aktualnego punktu spojrzenia obserwatora. W badaniach tych użyto jednak innego niż poprzednio urządzenia rejestrującego ruchy oczu, którego typ i częstotliwość próbkowania były porównywalne z okulografem stosowanym wcześniej. Zmiana ta nie została uzasadniona w rozprawie. Efektem przeprowadzonego eksperymentu są rekomendacje dotyczące częstotliwości odświeżania obrazu i próbkowania sygnału ruchu oka, które mają szansę zredukować możliwość dostrzeżenia operacji przerysowania sceny do obrazu o bardziej szczegółowej geometrii. W ostatnim podrozdziale tej części rozprawy zaprezentowano opracowany, w zespole badawczym Autora, system pomiaru ruchu oka, który może spełnić te wymagania. Ze względu na fakt, że autorami okulografu, jak i algorytmu wykrywania położenia źrenicy są inni członkowie zespołu, trudno ocenić udział Autora rozprawy w rozwoju tego urządzenia.

Z punktu widzenia planowanych do osiągnięcia celów, eksperymenty zostały starannie zaprojektowane i przeprowadzone. Przyjęta metoda badań była weryfikowana przed przystąpieniem do właściwych testów, z uwzględnieniem wyników uzyskiwanych w innych, powiązanych tematycznie, pracach. Zwrócono również uwagę na specyfikę działania układu wzrokowego człowieka, dostosowując do niej częstość i odstępy prowadzonych eksperymentów.

W trzecim rozdziale rozprawy zaprezentowano możliwości i sposoby użycia uzyskanych wyników w konkretnych zastosowaniach grafiki komputerowej. Jednym z nich była synteza obrazów realizowana w środowisku wirtualnej rzeczywistości (ang. *virtual reality*, VR). Zaproponowany model czułości na kontrast, na podstawie przeprowadzonego studium literaturowego, wykorzystany został do opracowania maski zapewniającej niższą gęstość próbkowania sceny w obszarze widzenia peryferyjnego. Dodatkowo, uproszczono proces rekonstrukcji obszarów nieobjętych próbkowaniem do prostej interpolacji barycentrycznej z użyciem kolorów sąsiednich pikseli. W mojej ocenie opis tej metody potraktowano zbyt lakonicznie, szczególnie, że w artykule [W6] jej zastosowanie wskazywane jest jako osiągnięcie badawcze. Dalsze zwiększenie efektywności procesu syntezy uzyskano poprzez

opracowanie takiej implementacji algorytmu wyznaczającego kolory pikseli należących do maski, która zakłada wykorzystanie zasobów karty graficznej. W efekcie wprowadzonych uproszczeń oraz na podstawie przeprowadzonych eksperymentów, potwierdzono skrócenie czasu renderingu sceny przy jednoczesnym utrzymaniu zadowalającej jej jakości.

W drugiej części tego rozdziału podjęto, niezbyt udaną, próbę zaprezentowania hybrydowej metody dekompozycji obrazu opracowanej dla środowiska wyposażonego w wyświetlacz stereoskopowy z akomodacją. W oparciu o ten opis oraz artykuł oznaczony jako [W7] można wnioskować, że model ten posłużył do opracowania reguły decydującej o wyborze metody dekompozycji obrazu w obszarze centralnego i peryferyjnego widzenia. Bazując na wykonanym studium literaturowym, na potrzeby tych badań wybrano dwie metody: liniowego mieszania (*ang. linear blending LB*) dla widzenia peryferyjnego oraz syntezy punktów świetlnych (*ang. Light-Field Synthesis LFS*) dla widzenia centralnego. Skuteczność nowej metody zbadano w eksperymentach pozwalających na porównanie wyników uzyskanych w trakcie renderingu zrealizowanego każdą z trzech metod niezależnie – LB, LFS oraz hybrydową. Środowisko dla badań stanowił zbudowany, w zespole badawczym Autora rozprawy, wielopłaszczyznowy wyświetlacz punktów świetlnych. Ocenie poddano zarówno jakość, jak i efektywność całego procesu z wykorzystaniem różnych scen. Wyniki eksperymentów wskazują, że zaproponowana metoda hybrydowa uzyskała lepsze parametry czasowe renderingu przy prawie niezauważalnej różnicy jakości rekonstruowanej sceny. Ponownie, dodatkowe zwiększenie efektywności uzyskano poprzez opracowanie implementacji metody LFS, dostosowanej do realizacji w środowisku procesorów graficznych. Ze względu na fakt, że wspomniany artykuł posiada dziesięciu autorów i brak jest jednoznacznego wskazania ich udziału w jego powstaniu, trudno jest mi określić stopień zaangażowanie Autora w ten zakres prac.

Ostatnią część rozprawy stanowi podsumowanie przeprowadzonych badań i wnioski końcowe bazujące na uzyskanych wynikach. W rozdziale tym, Autor uzasadnił potwierdzenie postawionej tezy oraz naszkicował kierunki prac, które dalej zamierza rozwijać.

Pozostałe części rozprawy stanowią dobrze opracowane spisy rysunków, tabel, własnych publikacji Autora i publikacji cytowanych w rozprawie. Wśród tych ostatnich znalazło się 110 pozycji pokrywających zakres tematyczny rozprawy. Nie zauważono natomiast w pracy streszczenia w języku angielskim.

3. Ogólna ocena rozprawy

Autor rozprawy wykazał się gruntowaną wiedzą z zakresu funkcjonowania układu wzrokowego, metod pozyskiwania i analizy sygnału ruchu oka, metod syntezy obrazów graficznych oraz renderingu czasu rzeczywistego.

Za najważniejsze oryginalne elementy tej rozprawy uważam:

- zaproponowaną strategię badań czułości na kontrast w obszarze peryferyjnego widzenia, z wykorzystaniem bodźców o różnej charakterystyce i użyciem okulografu dla stworzenia bardziej naturalnych dla oka warunków eksperymentów,
- wyznaczenie charakterystyki oraz opracowanie modelu czułości na kontrast, który wskazuje prawdopodobieństwo wykrycia bodźca wzrokowego z uwzględnieniem jego cech i kierunkowości widzenia,

- ocenę dynamiki percepcji w przypadku zmiany kierunku spojrzenia pod kątem wymaganej szybkości renderingu obrazu, z jakości dedykowanej obszarowi widzenia peryferyjnego na niezbędną dla widzenia centralnego,
- opracowanie rozwiązań umożliwiających wykorzystanie zasobów karty graficznej.

Interesująca jest również część pracy potwierdzająca realizowalność przyjętej koncepcji:

- implementacja opracowanego modelu w celu zwiększenia efektywności syntezy obrazu w środowisku wirtualnej rzeczywistości,
- zastosowanie opracowanego modelu w hybrydowym algorytmie dekompozycji obrazu do zastosowań w środowisku badawczym wykorzystującym wielopłaszczyznowy wyświetlacz punktów świetlnych.

Uzyskanie powyższych wyników pozwala na potwierdzenie osiągnięcia głównego celu pracy i większości celów szczegółowych. Potwierdzają one również słuszność postawionej tezy. Na podstawie zaprezentowanej rozprawy nie można jednak potwierdzić spełnienia, takich szczegółowych celów jak:

- *opracowanie oraz implementacja eye trackera o częstotliwości zapewniającej precyzyjną rejestrację ruchów sakadycznych* – chociaż prezentacja takiego okulografu znalazła się w przedstawionej rozprawie, to jednak jego opis oraz referencyjne publikacje, nie wskazują Autora rozprawy jako autora tego rozwiązania,
- *opracowanie modeli czułości na kontrast dla układu wzrokowego człowieka w zależności od częstotliwości sygnału, typu bodźca i kąta widzenia* – w rozprawie przedstawiono tylko jeden model, bazujący na kontraście bodźca i jego odległości od punktu centralnego widzenia.

Ponadto należy podkreślić, że rezultaty badań przedstawione w niniejszej rozprawie doktorskiej były wcześniej publikowane w siedmiu pracach ze współudziałem Autora rozprawy, w tym w dwóch publikacjach indeksowanych w bazie Web of Science, z czego jedna posiada Impact Factor równy 3.780. Wyniki te zostały więc w jakimś stopniu zweryfikowane w środowisku naukowym. Wszystkie te publikacje mają jednak kilku autorów i brak mi danych, aby szczegółowo ocenić wkład w te prace Autora rozprawy. Natomiast, należy zauważyć, że w trzech publikacjach nazwisko Autora rozprawy znalazło się na pierwszym miejscu na liście autorów. Zbiór publikacji nie jest imponujący, jednak, moim zdaniem, zadowolający na tym etapie rozwoju naukowego.

Praca doktorska ma również słabe strony i niedociągnięcia, które przedstawiłam w dwóch grupach – uwagi ogólne, obejmujące pytania i wątpliwości, oraz uwagi szczegółowe.

Uwagi ogólne

1. Dla prezentacji osiągnięć zastosowano w rozprawie formę bezosobową lub osobową sformułowaną w liczbie mnogiej. Taka prezentacja utrudnia ocenę wkładu Autora rozprawy w prowadzone badania:
 - *„Aby sprawdzić, czy nasz eksperyment został przygotowany poprawnie, wykonaliśmy pomiary czułości kontrastu dla bodźca achromatycznego....”*,

- „Bezpośrednie porównanie wskazuje zaniżoną czułość na kontrast uzyskaną w naszych pomiarach, co wyjaśnialiśmy różną wielkością bodźców wzrokowych w obu eksperymentach..”,
 - „natomiast fioletowa linia wskazuje progi kontrastu uzyskane w naszych badaniach...”.
2. W rozprawie Autor kilkakrotnie używa sformułowania „autorski algorytm”, „autorska procedura” lub „autorskiej płytce”. Określenia tego użyto, na przykład, w odniesieniu do algorytmu wykrywania żrenicy, opublikowanego w pracy [108] spisu literaturowego rozprawy, wśród autorów którego brakuje Autora rozprawy. Wymaga to więc wyjaśnienia czy w pozostałych przypadkach określenie to stosowano w stosunku do własnej pracy, czy też autorstwo dotyczy któregoś z członków zespołu badawczego.
 3. Wyniki badań potwierdzające skuteczność opracowanej metody uzyskane zostały dla stosunkowo małej liczby uczestników, jeśli weźmie się pod uwagę różnorodność oceny kolorów i zdolności ich postrzegania. Choć zaangażowani uczestnicy zapewne są użytkownikami systemów renderingu czasu rzeczywistego i ze względu na wiek oraz doświadczenie stanowią wiarygodną próbkę badawczą, to jednak większa liczba uczestników dawałaby silniejsze potwierdzenie poprawności dobranych progów czułości na kontrast.
 4. Rozprawa napisana została dość niestarannie. Poprawność języka pozostawia wiele do życzenia – szczególnie w trzecim rozdziale pracy. Czytelnik ma wrażenie, że rozprawa powstała jako bezpośrednio, czasem nietrafione, tłumaczenie artykułów z języka angielskiego. Powoduje to również niepotrzebne powtarzanie treści, które można było zaprezentować zbiorczo w jednym miejscu. Dobrym przykładem jest tu powtórzenie w rozprawie równania liniowego konwersji kolorów z przestrzeni XYZ do przestrzeni LMS lub powtarzanie opisu znaczenia kalibracji dla każdego użycia okulografu. Zastosowanie numeracji wzorów mogłoby rozwiązać pierwszy z wspomnianych problemów. Ponadto w pracy odnaleźć można wiele błędów edytorskich oraz wyrażen żargonowych.

Uwagi szczegółowe

- Wzór kontrastu Michelsona, prezentowany na początku strony 16, zawiera błąd w mianowniku – powinno być $L_{max} + L_{min}$
- W tekście rozdziału 2 można znaleźć zdanie, w którym występują dwa błędy. Pierwsza rzecz, która zwraca uwagę jest niepoprawnie zapisana nazwa autora publikacji oznaczonej numerem [101], Vatson zamiast Vanston. Druga uwaga dotyczy całego zdania – wydaje się ono niefortunnym tłumaczeniem zdania z artykułu oznaczonego w bibliografii rozprawy jako [W3].

„Współcześnie Vatson i Crognale [101] w swoich badaniach sugerują, że **wzrost rozmiaru bodźca chromatycznego w obszarze peryferyjnym nie zmienia się** gdy bodziec odbierany jest przez czopki typu L i M, w przeciwieństwie do czopków typu S, dla których postrzegalny **próg kontrastu zwiększa się.**”

- część zdania „Yxy zostały zmierzone spektrometrem, aby zostać skonwertowane do LMS”, ze strony 27 wydaje się być automatycznym tłumaczeniem translatora.

- „Pomiary czułości na kontrast dla achromatycznego bodźca wzorkowego wykazują, że szczytową zdolność rozpoznawania wzorców HVS osiąga dla bodźca wzorkowego o częstotliwości od 2 do 4 cpd.” Czy ten wniosek da się odczytać z rysunku 2.9 lub 2.11 ?
- Na rysunku 2.14 przyjęto błędne oznaczenie dla jednej serii danych – jest C1 powinno być raczej A1.
- Na stronie 45 pojawia się stwierdzenie *czułości bodźca na kontrast* – czy mamy tu do czynienia z pewnym uproszczeniem wypowiedzi ?
- Również na stronie 45 – „Standardowo za tę wartość przyjmuje się $\theta = 3.5$, jednak w proponowanym modelu przyjęto za $\theta = 2$ ”, z tekstu rozprawy nie wynika, czy wartość tę przyjęto za artykułem [19] czy przeprowadzono własne badania w tym kierunku ?
- Na stronie 68 zaproponowano niefortunny opis rysunku 3.6. Pierwszą sceną nazwano tą, która znajduje się po prawej stronie rysunku, drugą natomiast tą, którą umieszczono po lewej stronie. Naturalny sposób analizy rysunku sugerowałby odwrotną kolejność oznaczenia.
- Na stronie 71, na wykresie 3.8, po prawej jego stronie, nie zamieszczono informacji na temat różnic statystycznych wyników uzyskanych dla masek wskazujących 54% i 27% pikseli. Czy ich brak ma jakieś uzasadnienie ? W opisie wykresu pojawił się komentarz, „Prawdopodobieństwa bliskie 50% zwykle powodują brak istotności statystycznej. Dla większych prawdopodobieństw przerywana linia zastępowana jest przez niebieską linię”. Jak w tym kontekście należy traktować prawdopodobieństwo równe 77% ?
- Na stronie 75 pojawiło się sformułowanie „liniowej metody najmniejszych kwadratów”, czy chodziło o *regresję liniową metodą najmniejszych kwadratów* ? Ponadto dla określenia tej regresji użyto akronimu LFS tymczasem w artykule [W7], w którym opublikowano badania akronim LFS używany jest do określenia syntezy obrazów metodą punktów świetlnych *Light Field Synthesis*.
- Przykłady sformułowań żargonowych:
 - na stronie 67 pojawia się określenie – *temporalnego anty-aliasingu*,
 - na stronie 68 – *w ramach testów wydajnościowych ray tracera porównano ...*
 - na wielu stronach - *Wielopłaszczyznowy Wyświetlacz Light Fieldów, generowanie obrazów light field’ów ...*,
 - na stronie 75 – *w offline’owych przykładach*.

5. Wniosek końcowy

Podsumowując stwierdzam, że rozprawa doktorska mgr. inż. Michała Chwesiuka stanowi oryginalne rozwiązanie problemu naukowego i wykazuje ogólną wiedzę Autora z zakresu informatyki. Jest ona również świadectwem umiejętności Autora w zakresie prowadzenia badań naukowych. W mojej ocenie, przedstawiona do recenzji rozprawa, mimo pewnych niedociągnięć, spełnia wymagania stawiane rozprawom doktorskim przez obowiązujące przepisy. Wnoszę o dopuszczenie Pana mgr. inż. Michała Chwesiuka do dalszych etapów przewodu doktorskiego.

Katarzyna Kowalska

