

Bogdan Smółka
Instytut Automatyki
Politechnika Śląska

Gliwice, 06.09.2019

Recenzja rozprawy doktorskiej

Tytuł rozprawy:

Development of video motion magnification techniques for biomedical applications

Autor rozprawy:

mgr inż. Mateusz Popek

Promotorzy rozprawy:

dr hab. inż. D. Robert Iskander
dr inż. Monika E. Danielewska

Przedmiotem rozprawy są techniki wizualizacji i pomiaru subtelnych zmian sceny obrazu odwzorowywanej przez strumień wideo. Tematyka rozprawy jest bardzo aktualna i stanowi przedmiot intensywnych badań naukowych prowadzonych od kilku lat w wiodących ośrodkach naukowych.

Potrzeba tworzenia algorytmów uwidaczniających drobne zmiany w treści obrazu wynika w dużej mierze z własności ludzkiego układu wzrokowego, cechującego się bardzo wysoką przestrzenną zdolnością rozdzielczą lecz niską wrażliwością na bardzo szybkie lub powolne zmiany analizowanej sceny obrazowej. Zainteresowanie technikami wzmocnienia zmian czasowych w obrazach, stymulowane jest także potencjalnymi możliwościami aplikacyjnymi do pomiaru zjawisk, w których obiekt zmienia się w sposób nieobserwowalny przez układ wzrokowy człowieka.

Można tu wskazać między innymi na monitoring rytmu serca lub częstości oddychania na podstawie analizy wideo, detekcję zmian barwy skóry lub jej temperatury związanej z rytmem serca lub przeżywanymi emocjami, a także drgania budynków czy konstrukcji np. pod wpływem wiatru lub innych czynników. Paleta potencjalnych zastosowań jest bardzo szeroka, a różnorodne przykłady zostały zaprezentowane w bardzo atrakcyjnej formie w serwisie YouTube. Prezentacja pod tytułem *Eulerian Video Magnification*, przygotowana na podstawie pracy [3], została



odwiedzona ponad pół miliona razy, co z pewnością znacznie zwiększyło zainteresowanie tematyką wzmocnienia małych zmian w poszczególnych ramkach sekwencji wideo. Ponadto techniki wizyjne charakteryzują się bardzo niską inwazyjnością i stosunkowo niskimi kosztami, co wpływa bardzo istotnie na ich atrakcyjność aplikacyjną.

Algorytmy wzmocnienia zmian obrazowych bazują głównie na koncepcji dekompozycji ramek strumienia wideo za pomocą metod opartych na piramidach obrazów, a następnie na ich analizie czasowej. Dotychczas publikowane prace koncentrowały się raczej na jakościowym uwydatnianiu drobnych zmian w treści obrazu. Ich autorzy wskazywali na różnorodne możliwości wzmocnienia mikroprzemieszczeń, czy słabych zmian jasności lub barwy obiektów, nie koncentrowali się jednak na pomiarze parametrów wykrywanych zmian, takich jak na przykład ich częstotliwości czy rozkłady spektralne.

W rozprawie zaproponowano rodzinę algorytmów, pozwalających na wzmocnienie subtelnych zmian w strumieniu ramek wideo za pomocą radialnie modyfikowanej, dwuwymiarowej transformacji Hilberta. Zaproponowana metoda dokonuje dekompozycji ramek wideo za pomocą transformacji piramidowej, a następnie analizuje zmiany fazy w reprezentacji częstotliwościowej. Zmiany fazy poddawane są amplifikacji, a obraz wyjściowy budowany jest z poszczególnych składowych piramidy.

Oryginalnym wkładem Autora rozprawy jest zaproponowanie trzech metod uwydatniania małych zmian w obrazie wykorzystujących techniki analizy częstotliwościowej. Ponadto autor dokonał walidacji skuteczności opracowanych algorytmów na sekwencjach wideo drgającego głośnika akustycznego. Podjęto także próbę zastosowania utworzonych metod do pomiaru pulsacji rogówki oka. Tak więc rozprawa ma charakter zarówno teoretyczny, jak i eksperymentalny.

Recenzowana praca została zredagowana w języku angielskim, składa się z 5 rozdziałów i liczy 55 stron. Została ona opatrzona streszczeniem w języku angielskim i polskim, spisem publikacji autora rozprawy, spisem treści, wykazem ilustracji i tabel, a także spisem używanych w pracy skrótów. Ponadto rozprawa zawiera liczący 58 pozycji spis literatury, który bardzo dobrze odzwierciedla bieżący stan wiedzy na temat uwydatniania subtelnych zmian w sekwencjach wideo.

Rozprawa zredagowana została niezwykle lakonicznie. Zawiera ona wprawdzie w miarę dokładny opis opracowanych algorytmów oraz przeprowadzonych badań, jednakże jest on bardzo ogólny. Autor nie podaje wielu szczegółów, które są niezbędne do weryfikacji osiągniętych rezultatów przez innych badaczy. Recenzent nie znalazł także informacji o udostępnieniu kodów opracowanych metod, co pozwoliłoby na odszukanie niepodanych w rozprawie informacji.

Rozprawa została starannie przygotowana od strony redakcyjnej. Autor zdecydował się na wykorzystanie formatu edycyjnego pozwalającego na umieszczanie ilustracji czy komentarzy na marginesach manuskryptu, co ułatwia jego czytanie, a zwłaszcza nanoszenie uwag. Niestety potencjał wybranego szablonu nie został w pełni wykorzystany, albowiem marginesy na większości stron są puste, a w najważniejszym rozdziale 4 nie ma ich wcale.

R. Smolka

Rozdział pierwszy, liczący niecałe 2 strony, stanowi bardzo zwięzłe wprowadzenie do rozprawy. Autor skrótowo przedstawił koncepcję metod uwydatniania zmian w obrazie, wskazał na pierwsze prace, które zapoczątkowały rozwój metod analizy drobnych zmian w obrazie, przedstawił strukturę swojej dysertacji oraz jej główny, oryginalny wkład w rozwój metod analizy sekwencji wideo. Rozdział ten zawiera też tezę, która stwierdza, że udoskonalone techniki uwydatniania zmian czasowych w sekwencji obrazów mogą być stosowane do pomiaru sygnałów biomedycznych.

W rozdziale 2, Autor rozprawy dokonał przeglądu istniejącej literatury poświęconej metodom amplifikacji zmian czasowych w sekwencjach wideo. Przedstawił także koncepcję struktury reprezentacji piramidowej ramek sekwencji wideo. W rozdziale tym wprowadzono podstawowe definicje związane z transformacją Hilberta, Rieszsa oraz wykorzystanych w pracy ich uogólnień. Przedstawiono także opis sygnału monogenicznego, a także scharakteryzowano różne techniki pomiarowe pozwalające na analizę przebiegów czasowych zjawisk, w których obserwuje się mikroprzemieszczenia.

W rozdziale 3 przedstawiono stanowisko badawcze, które posłużyło do przeprowadzenia badań. Opisano poszczególne elementy aparatury: głośnik, wzmacniacz, generator sygnałów, oscyloskop, przetwornik ultradźwiękowy i 3 kamery. Następnie szczegółowo zbadano parametry dystorsji użytych kamer. Najważniejsza część tego rozdziału poświęcona została opisowi opracowanych przez Autora rozprawy algorytmów opartych na bankach filtrów Gabora, filtrach Rieszsa oraz sterowalnych, radialnych filtrach Hilberta. Opis jest bardzo zdawkowy, jednakże strukturę nowych algorytmów dobrze prezentują zamieszczone diagramy. W dalszej części skupiono się na ekstrakcji sygnału czasowego, uwydatnionych przez opracowane metody drobnych zmian przestrzennych poddawanego analizie obiektu oraz wprowadzono zmodyfikowaną miarę podobieństwa rozkładów zmiennych losowych. Rozdział kończy się przykładem zastosowania opracowanych algorytmów do analizy zjawiska pulsacji rogówki.

Rozdział czwarty podsumowuje rezultaty przeprowadzonych eksperymentów. Zbadano charakterystyki częstotliwościowe użytych do eksperymentów kamer, a następnie wyznaczono korelację pomiędzy wynikami uzyskanymi za pomocą opracowanej metodologii, a odczytami ultradźwiękowego urządzenia pomiarowego, traktowanego jako układ odniesienia. Obiektem pomiarowym była membrana głośnika akustycznego, wzbudzanego sygnałami sinusoidalnymi o różnej częstotliwości. Liczne wykresy oraz wyniki analizy statystycznej potwierdziły skuteczność zaproponowanego podejścia w przypadku drgań o jednej, dwóch i trzech wybranych częstotliwościach. Rozdział kończy się krótką analizą najbardziej interesującego przypadku drgań o szerokim spektrum częstotliwości. Porównano pulsację rogówki zmierzoną aparatem ultrasonograficznym oraz opracowaną metodą wizyjną. Przedstawiono wykresy czasowe odkształcenia rogówki, poddano analizie rozkłady mocy widmowej oraz zilustrowano otrzymane wyniki za pomocą analizy czasowej rozkładu ZAM.

R. Smolka

Jak wspomniano wcześniej, rozprawa jest bardzo zwięzła, co niestety bardzo utrudnia zrozumienie i ewentualną weryfikację uzyskanych wyników. Na podkreślenie zasługuje jednak fakt publikacji wyników przedstawionych w rozprawie w dwóch artykułach zamieszczonych w IEEE Transactions on Instrumentation and Measurement o wysokim współczynniku wpływu, (IF=3.067). Tym niemniej należałoby oczekiwać, że rozprawa zawierać będzie więcej szczegółów niż publikacje. Tak niestety nie jest. Przykładowo brak jest w rozprawie informacji zawartych w [56] na temat segmentów użytych do binaryzacji obrazu, zdawkowo potraktowano rolę korekcji gamma, której poświęcono znacznie więcej uwagi w publikacjach [57, 58]. W rozprawie nie porównano opracowanej metody z tzw. podejściem eulerowskim [3], pomimo, że można się z nim zapoznać w pracy [57] autora.

Autor rozprawy wykazał się bardzo dobrą znajomością zagadnień związanych z analizą sygnałów i przetwarzaniem obrazów. Opracowane algorytmy cechują się dużym stopniem zaawansowania i stanowią cenny wkład w rozwój wizyjnych metod pomiarowych. Teza rozprawy została w dużym stopniu uprawdopodobniona, aczkolwiek pełne potwierdzenie przydatności zaproponowanej metodologii wymaga jeszcze sporego wysiłku badawczego.

Czytając pracę nasunęło mi się kilka uwag krytycznych, które nie wpływają na moją ogólnie pozytywną ocenę rozprawy. Prosiłbym o dyskusję najważniejszych kwestii podczas obrony rozprawy.

1. W rozprawie nie podano jakiego rodzaju filtru użyto do budowy piramidy Gaussa. Kształt filtru jest istotny, ponieważ od jego postaci zależą w dużej mierze charakterystyki częstotliwościowe obrazów w piramidzie. W równaniu (2.2) brakuje współczynnika $w(m,n)$, (ten sam błąd można znaleźć w słynnej pracy [13]). Brakuje także informacji na temat postępowania w przypadku gdy $(2i+m)$ lub $(2j+m)$ jest liczbą nieparzystą. We wzorze (2.3) indeks przy G powinien być l , zamiast i .
2. Definicja transformacji Hilberta jest niespójna. Równanie (2.5) powinno zostać usunięte, a pozostawione tylko równanie (2.7). W przeciwnym razie należałoby zmienić definicję sygnału analitycznego i wprowadzić i do równania definicyjnego (2.4). W równaniu (2.12) dla porządku powinno się wprowadzić znak ujemny.
3. Nie doszukałem się w rozprawie znaczenia parametru λ_0 użytego w rysunku 2.3.
4. Razi brak informacji o roli parametru P , który występuje pod wzorem (2.15).
5. W rozprawie brak jest informacji na temat konstrukcji filtrów Gabora. Nie wiadomo co ma przedstawiać rysunek 3.14.
6. W rozprawie nie uzasadniono w sposób przekonujący konieczności modyfikowania definicji względnej entropii Kullbacka-Leiblera. Istnieje przecież wiele miar, w których bierze się pod uwagę wartości bezwzględne odchyłek pomiędzy rozkładami. Poza tym, własności tak określonej miary są nieznane, a jej przydatność do oceny rozbieżności pomiędzy dwoma rozkładami nie jest oczywista.
7. Na stronie 41, autor twierdzi, że dokładność pomiaru częstotliwości wynosi 0.0 Hz. Nie wiadomo czy to błąd pisarski, czy niezręczne określenie.

A. Sankowski

8. Niejasne jest w jaki sposób przeprowadzono filtrację medianową na sekwencji wideo. Czy wyznaczano medianę z pikseli na poszczególnych pozycjach ramek (jak sugeruje rysunek z pracy [56]) czy też wyznaczano medianę z poszczególnych wierszy tych ramek (jak można domyślić się z rysunku zamieszczonego w rozprawie).
9. W rozprawie nie poświęcono uwagi wpływowi korekcji gamma na wyniki filtracji sekwencji wideo, mimo że kwestia ta była dość szczegółowo omówiona w artykułach, których Doktorant jest pierwszym autorem.
10. Kluczowy fragment rozprawy (4.5) jest niedopracowany. Podano tylko 2 wykresy i diagram ZAM. Korelacja pomiędzy pomiarami uzyskanymi za pomocą aparatu ultrasonograficznego i przy użyciu zaproponowanej metody nie została oceniona w sposób obiektywny.

Mimo tych niedociągnięć, należy stwierdzić, że Autor rozprawy wykazał się świetną znajomością zagadnień przetwarzania obrazów i wizji komputerowej, biegłością w zastosowaniu metod częstotliwościowych, a także bardzo dobrą orientacją w literaturze poświęconej zagadnieniom filtracji sygnałów wideo..

Przedstawione w rozprawie doktorskiej wyniki prowadzonych prac badawczych stanowią oryginalny wkład magistra inżyniera Mateusza Popka w rozwój Inżynierii Biomedycznej. Autor rozprawy wykazał się także bardzo dobrym opanowaniem warsztatu badawczego, a w rezultacie również znaczną dojrzałością naukową. Sformułowane w rozprawie problemy naukowe zostały rozwiązane w sposób właściwy, a jej teza została uprawdopodobniona.

Uwzględniając wysoki poziom naukowy pracy, a w szczególności stworzenie oryginalnego systemu wzmacniania drobnych zmian w strukturze obrazu i przeprowadzenie jego weryfikacji eksperymentalnej, recenzowaną dysertację oceniam pozytywnie i stwierdzam, że spełnia ona warunki stawiane ustawowo rozprawom doktorskim, zgodnie z obowiązującą ustawą o stopniach i tytule naukowym. Wnoszę zatem o jej przyjęcie oraz o dopuszczenie do publicznej obrony.



Bogdan Smolka