

Prof. dr hab. inż. Bożena Kostek, prof. nadzw. PG
Katedra Systemów Multimedialnych
Wydział Elektroniki, Telekomunikacji i Informatyki
Politechnika Gdańska

Gdańsk, dnia 15 maja 2007 r.

**Recenzja rozprawy doktorskiej
mgr Krzysztofa Tyburka
pt. „Klasyfikacja instrumentów strunowych w multimedialnych
bazach danych ze uwzględnieniem artykulacji pizzicato”**

Recenzowana rozprawa doktorska mgr inż. Krzysztofa Tyburka ma charakter eksperymentalny. Liczy 140 stron i dodatkowo składa się z 30-stronicowego Dodatku, który powstał na prośbę Recenzentów jako uzupełnienie rozprawy doktorskiej. Na rozprawę składają się Wstęp, Spis oznaczeń, dziewięć rozdziałów, Podsumowanie oraz Literatura. Z kolei Dodatek zawiera uzupełnienia do wybranych rozdziałów rozprawy.

**Problem naukowy rozprawy - ocena tematu i zakresu opiniowanej
rozprawy**

Tematyka rozprawy dotyczy obserwowanego obecnie gwałtownego rozwoju ilościowego multimedialnych zasobów danych dostępnych cyfrowo (w tym również danych muzycznych), w którym ważnym elementem praktycznym tego rozwoju jest problem wyszukiwania określonych informacji przechowywanych w tych multimedialnych zasobach. W ciągu ostatnich kilkunastu lat opracowano wiele metod i algorytmów klasyfikacji dźwięków instrumentów muzycznych, jednak problem ten jest na tyle złożony i nie do końca rozwiązany, że pozostaje nadal aktualny. W oparciu o te badania powstał standard MPEG 7, który zawiera m.in. zestawienie użytecznych deskryptorów, które mogą być wykorzystane w procesie parametryzacji danych fonicznych, w tym dźwięków instrumentów muzycznych, ale nie zawsze deskryptory niskiego poziomu zawarte w standardzie MPEG 7 są wystarczające w celu zapewnienia skutecznej klasyfikacji.

Doktorant postawił sobie za cel stworzenie mechanizmów pozwalających na rozpoznanie dźwięków instrumentów muzycznych, a w szczególności instrumentów strunowych z artykulacją pizzicato. Jest to pewien wycinek badań z zakresu rozpoznawania dźwięków instrumentów muzycznych, które są aktualnie prowadzone w licznych ośrodkach w kraju i zagranicą, a więc stanowi przyczynek do bardzo aktualnej tematyki.

Rozprawę rozpoczyna Wstęp, w którym Doktorant przedstawił genezę tematu, pokrótce przywołał podstawowe deskryptory ujęte w standardzie MPEG 7, cel rozprawy, a także sformułował dwie tezy w brzmieniu:

1. Istnieje taki wektor cech, który pozwoli na skuteczne rozpoznanie przebiegów dźwiękowych instrumentów strunowych z artykulacją pizzicato.

2. Nowo zaproponowany wektor cech zawiera deskryptory wynikające z analizy przestrzeni widmowej badanych próbek dźwiękowych.

Sądzę, że pierwsza teza powinna zostać uściślona, gdyż Doktorant rozpoznaje w sposób automatyczny klasę instrumentu muzycznego w oparciu o sparametryzowane dźwięki z artykulacją pizzicato, ale nie zajmuje się rozpoznawaniem przebiegów dźwiękowych, jak to zostało sformułowane w tezie. W przypadku drugiej tezy, sformułowanie to ma charakter informacyjny, a powinno być o charakterze koniecznym do uwodnienia.

Rozdziały 1-3 rozdziały mają charakter wprowadzający. Rozdziały te przywołują zagadnienia podstawowe z zakresu akustyki fal, sygnału cyfrowego, problemów związanych z kwantyzacją sygnału czy jego analizą. W moim odczuciu przedstawione zagadnienia mają dość ogólne powiązanie z rozprawą i dla zachowania bardziej spójnego charakteru rozprawy można by z nich zrezygnować, pozostawiając jedynie rozdział dotyczący instrumentów muzycznych ze szczególnym uwzględnieniem uzupełnienia z Dodatku oraz powołując się na stosowną literaturę. Taki też podręcznikowy charakter, nie związany bezpośrednio z ocenianą rozprawą, ma piąty rozdział pracy dotyczący baz danych i systemów zarządzania bazą danych. Ponadto, w rozdziale 5. zawarto przykłady nie związane z danymi multimedialnymi, a w szczególności z danymi muzycznymi. Dopiero ostatnie półtorej strony tego 26-stronicowego rozdziału odnosi się w sposób ogólny do danych multimedialnych i dodatkowo uzupełnienie zawarte w Dodatku przybliży pokrótce zagadnienia baz muzycznych. Sądzę, że gdyby Doktorant zastosował konwencję umieszczenia kilku zdań wprowadzających na początku każdego rozdziału, wtedy ta nadmiarowość byłaby możliwa do uniknięcia.

Zagadnienia związane z parametryzacją dźwięków instrumentów muzycznych zostały przedstawione w rozdziale czwartym. Rozdział ten ma charakter bardzo zwięzły i dobrze się stało, że w uzupełnieniu do rozprawy (patrz rozdział 4 Dodatku) pojawił się szerszy przegląd źródeł bibliograficznych dotyczących parametryzacji dźwięków instrumentów muzycznych, zwłaszcza, że zagadnienia te są bardzo dobrze reprezentowane w literaturze tematu. Niedosyt budzi jednak brak odwołań do prac bezpośrednio zajmujących się aspektami artykulacyjnymi instrumentów muzycznych. Cennym uzupełnieniem tego rozdziału jest też ogólne przedstawienie deskryptorów zapisanych w standardzie MPEG 7 (patrz rozdział 6 Dodatku), gdyż prowadzenie eksperymentów w szeroko pojętej dziedzinie wyszukiwania informacji muzycznej (ang. *Music Information Retrieval*) wymaga odniesienia się do obowiązującego standardu w tym obszarze.

Wybrane algorytmy klasyfikujące przedstawione zostały w rozdziale 6. W treści tego rozdziału znalazły się również przykłady wyników klasyfikacji dźwięków instrumentów muzycznych, przez co rozdział ten ma charakter niespójny, gdyż nie wiadomo czy jest to już część eksperymentalna rozprawy, czy tylko przedstawienie wyników działania poszczególnych algorytmów w zależności od zastosowanych parametrów procesu klasyfikacyjnego. W szczególności dotyczy to podrozdziału 6.5 pt. „Przygotowanie danych testowych”, Doktorant podaje w nim wnioski natury ogólnej dotyczące klasyfikacji na podstawie przeprowadzonego eksperymentu (tak przynajmniej można się domyślać), ale nie wiadomo, dla przykładu, jaka była zawartość wektora cech wykorzystywanego w tym eksperymencie.

