

R e c e n z j a

rozprawy doktorskiej mgr. inż. Konrada Kluwaka
pt. „**Modele i algorytmy detekcji znaczników w problemach analizy ruchu
obiektów**”

1. Obszar problemowy rozprawy

Tematyka rozprawy dotyczy wybranych problemów analizy danych pochodzących z sensorów kinematyki, kinetyki i elektromiografii dynamicznej (EMG), związanych z analizą ruchu człowieka. Ze względu na charakter proponowanych rozwiązań i typ analizowanych danych rozprawę można zaliczyć do obszarów badawczych informatyki, w szczególności obejmujących metodyki „sztucznej inteligencji” i „rozpoznawania obrazów” służące do „wizualnej analizy dynamicznych scen” i „interpretacji danych o ruchu człowieka”.

Możliwość automatycznej analizy/oceny ruchu obiektu, analizy wykorzystującej reprezentacje wielomodalne (parametry antropometryczne, kinetyczne, kinematyczne itp.) wiąże się z koniecznością rozwiązania szeregu problemów cząstkowych obejmujących m.in: akwizycję danych wielomodalnych, reprezentację cech ruchu (tagowanie), redukcję wymiarowości danych, identyfikację postawy człowieka, trajektorii ruchu czy też interakcji z innymi osobami. Tematyka analizy ruchu człowieka jest znana od wielu lat i wciąż bardzo aktualna. Powszechnie stosowane rozwiązania chociażby z obszaru metod sztucznych inteligencji z powodzeniem wykorzystywane są w procesach rozpoznawania obrazów, jednak ich zastosowanie w sytuacjach wymagających semantycznej oceny interakcji pomiędzy obiektami (ludźmi) ogranicza się do stosunkowo niewielu przypadków i wciąż stanowi duże wyzwanie.

W ogólności problemy analizy ruchu obiektów wymagają znacznych nakładów obliczeniowych. Wiąże się to koniecznością przetwarzania ogromnych ilości danych, których rozmiar wyklucza możliwości czasowe aktualnie dostępnych metod i narzędzi. Powoduje to, że czasowo efektywne metody analizy ruchu charakteryzuje niewielka dokładność a każda

WPLYNĘŁO

29-09-2021

ANJ-IT/160/2021

1/6

próba zwiększenia ich dokładności okupiona jest dużym wzrostem czasu obliczeń. Należy zaznaczyć, że stopień „trudności” tego typu problemów rośnie jeszcze bardziej, gdy w procesie analizy uwzględnia się złożone pozy i akcje człowieka (np. rozważane w rozprawie uderzenia mieczem, czy chód osoby po endoprotezoplastyce) jak również interakcje zachodzące między dwoma osobami (np. zmiana z pozycji leżącej na pozycję siedzącą). Warto zauważyć, że możliwość analizy ruchu, jest warunkowana sposobem identyfikacji (tagowaniem) i właściwym wyborem charakterystycznych cech pozyskiwanych parametrów wielomodalnych. Odpowiednie tagowanie pozwala na redukcję rozmiaru danych i przyspiesza wyznaczanie pożądaných informacji. Przedstawiona możliwość zmniejszenia złożoności obliczeniowej analizy ruchu obiektów stanowi główną motywację opracowania modelu i budowy czasowo efektywnych algorytmów detekcji znaczników (tagów) wielomodalnych parametrów ruchu leżących u podstaw prowadzonych przez Doktoranta badań.

Biorąc pod uwagę, że prowadzone prace ograniczają się do analizy ruchu człowieka opiniowana rozprawa koncentruje się na problemie, którego istotę można zrekonstruować następująco: Dany jest zbiór parametrów wielomodalnych obejmujących informacje pochodzące z sensorów kinematyki, kinetyki i elektromiografii dynamicznej ruchu człowieka. Poszukiwana jest odpowiedź na pytanie: Czy czynności wykonane przez człowieka są zgodne z wcześniej zadanymi referencyjnymi wzorcami zachowań (np. wzorcowymi uderzeniami miecza, poprawnym transportem pacjenta, itp.)? Dla tak zdefiniowanego problemu prowadzone są poszukiwania modelu oraz algorytmów detekcji znaczników (tagów postawy, tagów trajektorii oraz tagów interakcji) zapewniających efektywne rozpoznawanie zadanych aktywności człowieka.

Uważam, że opiniowana rozprawa podejmuje ważny i aktualny problem. Elementy nowości przejawiają się w propozycji autorskiej metody analizy sekwencji danych wielomodalnych. Podjęcie przedstawionej problematyki jest uzasadnione zarówno ze względów poznawczych, jak i możliwości wielu praktycznych zastosowań związanych m.in. z systemami wsparcia treningów/zawodów sportowych, rehabilitacji pooperacyjnej, opieki nad osobami niepełnosprawnymi.

2. Kompozycja i treść rozprawy

Opiniowana rozprawa liczy 162 strony i składa się z wprowadzenia, 7 rozdziałów podzielonych na trzy części, 2 dodatków (19 stron), bibliografii, spisu rysunków, spisu tabel, streszczenia w j. polskim oraz streszczenia w j. angielskim. Załączona lista źródeł bibliograficznych (192 pozycje) obejmuje ważniejsze pozycje literaturowe z zakresu

przedmiotu pracy. W bibliografii występuje 24 współautorskich publikacji Doktoranta. Dwie z nich zostały opublikowane w czasopismach: Sensors (IF = 3.576) oraz Acta Polytechnica Hungarica (IF=1.806) indeksowanych w bazie Scopus oraz Journal Citation Reports.

W rozprawie wyodrębnić można trzy zasadnicze części. W części pierwszej (obejmującej rozdziały 1-4) sformułowano tezy, cel główny rozprawy oraz przedstawiono aktualny stan wiedzy z zakresu analizy ruchu obiektów, akwizycji wielomodalnych danych ruchu, detekcji danych wielomodalnych. W szczególności przedstawiono przestrzenno-czasowy model ludzkiego ciała oraz wprowadzono autorski model hierarchicznej reprezentacji informacji wielomodalnej bazującej na metodzie tagowania Extracted Tag (EXTag).

Tezy rozprawy zostały sformułowane poprawnie, choć pozostawiają pewien niedosyt odczuwalny w zakresie ich uzasadnienia. Poza krótką dyskusją zawartą we „Wprowadzeniu” brakuje szerszego wyjaśnienia, z jakiego powodu Doktorant zdecydował się na takie, a nie inne rozwiązanie. Warto zaznaczyć, że do postawionych tez w dalszej części pracy Doktorant się nie odwołuje. W tym kontekście istotny niedostatek stanowi brak podsumowań kolejnych rozdziałów, które mogłyby referować uwagi autorskie obejmujące krytyczną analizę oraz przedstawienie argumentów stojących za przyjętym sposobem prowadzonych badań.

W części drugiej (rozdziały 5-7) zawierającej główne (poznawcze) wyniki rozprawy przedstawiono rezultaty prowadzonych badań ilościowych. Celem eksperymentów była praktyczna weryfikacja opracowanego modelu tagowania Extracted Tag (EXTag).

Pierwszy obszar eksperymentów (rozdział 5) dotyczył procesu tagowania pozy człowieka w trakcie uprawiania treningu (identyfikacja uderzeń długim mieczem). Na szczególną uwagę zasługuje dedykowana dla tego typu analiz metoda tagowania. W metodzie tej przyjęto, że detekcja tagu pozy człowieka jest możliwa dzięki wyznaczeniu „odległości” od tzw. pozy reprezentatywnej.

Kolejna część eksperymentów (rozdział 6) dotyczyła procesu tagowania określonych akcji człowieka. Doktorant przedstawił zastosowanie rekurencyjnych sieci neuronowych (RNN) typu Long Short-Term Memory (LSTM) do tagowania (akcji, trajektorii) jednostki chorobowej na podstawie ścieżek chodu pacjentów pooperacyjnych (osoby po endoprotezoplastyce oraz udarze). W prezentowanej metodzie przyjęto założenie, że proces tagowania jest synchronizowany uderzeniem prawej stopy o podłoże.

Trzecia część eksperymentów (rozdział 7) dotyczyła z kolei procesu tagowania interakcji między dwoma osobami. W szczególności Doktorant podjął się próby oceny (wzg. ergonomii pracy) czynności wykonywanych przez personel medyczny. Ergonomię pracy badano na przykładzie podnoszenia pacjentów z pozycji leżącej na łóżku do pozycji siedzącej. Jednym



z rezultatów prowadzonych badań jest autorska baza otagowanych danych, interakcji pacjent-pielęgniarz (rejestrowane są dwie osoby: podnosząca i podnoszona). Badania prowadzone były przy współpracy z Centrum Badawczo-Rozwojowym Polsko-Japońskiej Akademii Technik Komputerowych w Bytomiu.

Wśród niedostatków tej części rozprawy w pierwszej kolejności należy zaliczyć brak modeli matematycznych opracowanych modeli i metod tagowania oraz analizy złożoności obliczeniowej proponowanych algorytmów. Z tego względu trudno jest ocenić możliwość wykorzystania proponowanego podejścia w bardziej złożonych (niż te analizowane w rozprawie) przypadkach – trudno oszacować na przykład: licznosc grupy osób, dla których interakcja może być skutecznie identyfikowana? Szkoda również, że Doktorant nie przeprowadził porównania proponowanego rozwiązania z opisywanymi w literaturze podejściami. Tego typu badania pozwoliłyby określić obszar zastosowań opracowanej metody (niekoniecznie ograniczający się do aktywności sportowej i wsparcia medycznego).

Pomimo wskazanych niedostatków należy podkreślić, że przyjęty sposób narracji tej części rozprawy jest na ogół poprawny i świadczy o dobrym przygotowaniu oraz wysokich kompetencjach Doktoranta.

3. Oryginalne osiągnięcia

Do najważniejszych osiągnięć rozprawy, wyróżniających je spośród dostępnych w literaturze przedmiotu, można zaliczyć:

1. Autorski model hierarchicznej reprezentacji informacji wielomodalnej bazującej metodzie tagowania Extracted Tag (EXTag).
2. Metodę tagowania pozy człowieka bazującą na wyznaczeniu „odległości” od tzw. pozy reprezentatywnej.
3. Metodę redukcji wymiaru danych wielomodalnych, opisujących cechy pozy człowieka bazującą na analizie głównych składowych - PCA.
4. Metodę tagowania jednostki chorobowej na podstawie ścieżek chodu (metoda bazująca na rekurencyjnych sieci neuronowych typu Long Short-Term Memory (LSTM)).
5. Metodę tagowania interakcji międzyludzkich zachodzących pomiędzy dwoma osobami.

Uwzględniając wymienione osiągnięcia naukowe uważam, że Doktorant zrealizował cel rozprawy. Uzyskane rezultaty potwierdzają kompetencje Doktoranta z zakresu technik programowania, metod przetwarzania obrazów, metod sztucznej inteligencji, jak i planowania/realizacji eksperymentów komputerowych. Dowodzą również, że Doktorant potrafi podejmować i samodzielnie realizować zaplanowane cele badawcze.

4. Uwagi i komentarze dotyczące rozprawy

Lektura rozprawy skłania do kilku uwag, tak ogólniejszej, jak i bardziej szczegółowej natury:

1. W rozdziale 5 przeprowadzona została redukcja wymiarowości danych wielomodalnych, której efektem było ograniczenie liczby atrybutów analizowanych danych z 117 do 5. Interesującym w tym kontekście wydaje się być pytanie, czy możliwe jest określenie minimalnego zbioru atrybutów, które dla każdego z analizowanych zagadnień: tagowanie pozy, tagowanie trajektorii, tagowanie interakcji, gwarantuje właściwą (z zadaną dokładnością) identyfikację ruchu obiektów?
2. W rozdziałach 5-7 przedstawiono możliwość wykorzystania opracowanych rozwiązań w rzeczywistych zastosowaniach. W tym kontekście interesująca wydaje się ocena możliwości implementacji opracowanych algorytmów w systemie doradczym. Czy z perspektywy przeprowadzonych badań można odpowiedzieć na przykładowe pytanie: Jak mógłby wyglądać proces wspierania decyzji w takim systemie?
3. Interesująca wydaje się również dyskusja na temat porównania proponowanych rozwiązań do rezultatów innych prac (np. doktorskich) i przedstawianych tam metod np. *Śledzenie ruchu postaci ludzkiej w systemie wielokamerowym*, Tomasz Krzeszowski, Promotor: prof. dr hab. inż. Bogdan Kwolek, Gliwice 2013.

Uwagi szczególne:

1. Brak stosownej korekty redakcyjnej dotyczy sposobów prezentacji rysunków. Większość z nich nie posiada legendy i opisu stosowanych oznaczeń. Część z nich stanowi zrzuty ekranów, które są całkowicie nieczytelne (np. rys. 2.6; 3.4, 3.8, 3.9, 3.12, itp.) - przyjęcie takiego sposobu prezentacji wyników bardzo utrudnia lekturę pracy.
2. Wielokrotnie tekst jest pisany w różnych trybach np.: str. 32 (3-osoba) „Autor współpracował”, str. 34 (1-osoba) „uczestniczę”, itp.
3. W rozprawie występują nieliczne błędy językowe i edycyjne, np. str.3 brak kropki po słowie „pomiarowych”, str. 39 „zdarzeniem może być stopą”, „dotknięcie stopa”, itp.

Pomijając powyżej wymienione uwagi należy zaznaczyć, że rozprawa jest zredagowana poprawnie, posiada właściwą strukturę i proporcje. Błędy stylistyczne występują sporadycznie.

5. Konkluzja

Reasumując stwierdzam, że w recenzowanej rozprawie doktorskiej mgr inż. Konrada Kluwaka rozwiązany został oryginalny problem badawczy, polegający na opracowaniu modelu

oraz algorytmów detekcji znaczników (tagów postawy, tagów trajektorii oraz tagów interakcji) zapewniających efektywne rozpoznawanie zadanych aktywności człowieka. Doktorant wykazał się znajomością podstawowej literatury przedmiotu rozprawy, umiejętnościami budowy i komputerowej implementacji metod przetwarzania danych i sztucznej inteligencji, a także umiejętnościami prowadzenia eksperymentów obliczeniowych.

Uważam, że opiniowana rozprawa spełnia warunki stawiane przez obowiązującą ustawę o stopniach i tytule naukowym w określeniu do rozpraw doktorskich (Ustawę z dnia 3 lipca 2018 r. Przepisy wprowadzające ustawę – Prawo o szkolnictwie wyższym i nauce) w dyscyplinie **informatyka techniczna i telekomunikacja** i wnioskuję o jej dopuszczenie do publicznej obrony.

A handwritten signature in blue ink, appearing to read "G. regoer Zuzia".