

Politechnika Śląska
Wydział Automatyki, Elektroniki i Informatyki

Rozprawa doktorska

mgr inż. Magdalena Pawlyta

Klasyfikacja zachowań postaci ludzkiej z
wykorzystaniem uczenia maszynowego na
podstawie trajektorii punktów
charakterystycznych w reprezentacji 3D i 2D

Promotor: prof. dr hab. inż. Konrad Wojciechowski,
Promotor pomocniczy: dr inż. Przemysław Skurowski,

STRESZCZENIE

Niniejsza rozprawa jest próbą wykazania, w jakim stopniu zdolność do rozpoznawania zachowań/akcji w systemach wizyjnych zależy od reprezentacji śledzonych punktów, ich liczby oraz położenia i orientacji kamery w przypadku reprezentacji 2D.

W tym celu utworzono bazę danych ruchu 3D, zawierającą 9 unikatowych akcji prostych. Do jej stworzenia wykorzystano technologię Motion Capture, co umożliwiło uzyskanie referencyjnych, niezakłóconych danych. W celu uniknięcia zafałszowania wyników zaproponowanych modeli opracowano specjalny algorytm podziału danych na dowolną liczbę podzbiorów w celu przeprowadzenia walidacji krzyżowej.

Na podstawie danych 3D utworzona została następnie baza danych ruchu 2D. W tym celu opracowano model wirtualnej kamery, z modyfikowalnymi parametrami zarówno wewnętrznymi jak i zewnętrznymi. Kamera ta ustawiana była na wirtualnej scenie w różnych lokalizacjach, a następnie wykonano szereg rzutowań perspektywicznych. Wyniki tych rzutowań stanowiły źródło danych 2D.

Przeprowadzono szereg eksperymentów obejmujących klasyfikację wybranych akcji zarówno w przestrzeni 3D i 2D. Ponieważ celem rozprawy nie było opracowanie idealnego modelu klasyfikującego zachowania postaci ludzkiej, a zbadanie aspektów wpływających, na jakość klasyfikacji, do eksperymentów wykorzystano dwie, dość generyczne architektury głębokich sieci neuronowych – sieć CNN i sieć LSTM.

W pracy wykazano wpływ zmiany reprezentacji danych z 3D na 2D, na jakość klasyfikacji. Wykazano także, że redukcja, wymiarowości przestrzeni wejściowej, do pewnego stopnia poprawia, jakość klasyfikacji. Stopień redukcji zależy od wykorzystywanego modelu, oraz przestrzeni danych – inny w przestrzeni 3D i inny w przestrzeni 2D.

Dodatkowo przebadano wpływ rzutowania perspektywicznego na ogólną dokładność klasyfikacji, oraz na rozpoznawalność poszczególnych zachowań/akcji. W tym celu wprowadzono dwie dodatkowe miary, mające związek ze specyfiką utworzonej bazy danych, oraz opracowano metody wizualizacji wyników w formie specjalnych kopuł. Uzyskane wyniki zostały szczegółowo przeanalizowane i przedyskutowane oddzielnie dla każdej z akcji. Pozwoliło to na wykazanie wpływu, na jakość klasyfikacji, przy czym uzyskane wyniki w niektórych przypadkach nie są intuicyjne.

Słowa kluczowe: klasyfikacja zachowań postaci ludzkiej, uczenie maszynowe, głębokie sieci neuronowe