

Abstrakt

Niniejsza praca dotyczy propozycji kompleksowych rozwiązań do rozpoznawania ludzkich zachowań z niekontrolowanych, realistycznych danych wideo, takich jak filmy, Internet oraz nagrania z monitoringu. Zaproponowaliśmy rozwiązanie, umożliwiające jednoczesne rozpoznawanie w strumieniu wideo wielu indywidualnych akcji wielu osób w ruchu i ich jednoczesną lokalizację w czasie i w przestrzeni.

W celu rozwiązania praktycznego problemu utrzymania ciągłej lokalizacji w przestrzeni na podstawie obrazu z ruchomych kamer monitoringu, opracowaliśmy metodę ciągłej kalibracji, nawigacji i estymacji współrzędnych świata danego obiektu uchwyconego przez kamerę.

Zbudowaliśmy bibliotekę programistyczną do obsługi kamer monitoringu w zakresie ciągłego znajdowania położenia we współrzędnych świata punktu odpowiadającego pikselowi w bieżącym obrazie pochodzącym z kamery monitoringu, przy założeniu, że punkt ten w rzeczywistości leży na płaszczyźnie podłoża. Następnie zbadaliśmy potencjał tej biblioteki do korekcji zniekształceń wideo i ciągłego śledzenia rozpoznanych w strumieniu wideo poruszających się obiektów.

Opracowaliśmy bezszkieletową reprezentację dla grup lokalnych trajektorii wyznaczonych funkcji ruchu oraz ruchomych obiektów. Bezszkieletowa reprezentacja pozwala uprościć opis sceny, wyodrębnić określone cechy obiektu za pomocą parametrów trajektorii i wydzielić z wykrytych obszarów ruchomych obrazu ruchy poszczególnych osób w sekwencji wideo.

Zaproponowaliśmy również bardziej uogólnioną formę reprezentacji bezszkieletowej jako zbiór sekwencji symboli z powszechnie używanego alfabetu, z których każda sekwencja symboli ma indywidualny czas rozpoczęcia i zakończenia. Zademonstrowaliśmy możliwość rozpoznawania akcji przy użyciu naszego podejścia.

Podjęte wyzwanie związane z rozpoznawaniem akcji w strumieniu wideo w czasie zbliżonym do rzeczywistego zostało rozwiązane przy użyciu zoptymalizowanych lokalnych reprezentacji czasoprzestrzennych dla reprezentacji ruchu oraz podejścia Bag-of-Visual-Words do rozpoznawania akcji.

Przetestowaliśmy algorytmy na danych ze świata rzeczywistego z publicznych zbiorów wideo (Hollywood2, UCF101, HMDB51) oraz bezpośrednio z miejskich sieci kamer monitoringu. W oparciu o zaproponowane metody zrealizowaliśmy inteligentny system monitoringu SAVA klasy przemysłowej i zademonstrowaliśmy jego działanie w czasie zbliżonym do rzeczywistego w rzeczywistym systemie monitoringu. Tym samym potwierdziliśmy, że nasze podejścia osiągnęły możliwości zastosowania zgodne ze stanem techniki oraz dostępnym potencjałem obliczeniowym pokazując skuteczność i wydajność sugerowanych metod.