

OKARMA Krzysztof¹
MAZUREK Przemysław¹

Zmodyfikowana hybrydowa metoda nieliniowej estymacji tła dla wizyjnych systemów śledzenia ruchu pojazdów

Inteligentne Systemy Transportowe,
Wizyjna detekcja ruchu,
Estymacja tła

Streszczenie

W artykule przedstawiono efektywną metodę estymacji tła przydatną w technikach śledzenia ruchu pojazdów na drodze w Intelligentnych Systemach Transportowych. Ze względu na wpływ szumu oraz zakłóceń powodowanych przez inne ruchome obiekty na wyniki estymacji tła najlepsze wyniki zapewnia stosowanie technik nieliniowych z dodatkowym pomijaniem lub uśrednianiem sąsiednich klatek sekwencji wideo. W celu weryfikacji algorytmów estymacji tła możliwe jest wykorzystanie wskaźników podobieństwa obrazu lub porównawczych wskaźników jakości obrazu wymagających jednak znajomości referencyjnego obrazu tła. Dzięki ich zastosowaniu możliwe jest również adaptacyjne przełączanie pomiędzy algorytmami estymacji tła w zależności od wykrytych lokalnych zmian w obrazie powodowanych przez poruszające się obiekty, a także pominięcie klatek lub ich fragmentów znacząco różniących się z tego powodu od pozostałych. W artykule zaproponowano i zweryfikowano modyfikację uprzednio proponowanego hybrydowego podejścia do estymacji tła [5], dzięki czemu uzyskano szybszą zbieżność estymacji oraz zwiększenie stabilności uzyskiwanych wyników.

A MODIFIED HYBRID METHOD OF NONLINEAR BACKGROUND ESTIMATION FOR VISION BASED VEHICLE TRACKING SYSTEMS

Abstract

In the paper an effective method of background estimation is presented, useful for vehicles tracking on the road in Intelligent Transportation Systems. Due to the influence of noise and contaminations caused by some other moving objects on the results of the background estimation the best results can be achieved using the nonlinear algorithms with additional skipping or averaging the neighbouring video frames. In order to verify the background estimation algorithms some image similarity metrics and full-reference image quality assessment methods, which require the reference background image, can be applied. Due to their usage an adaptive switching between the background estimation algorithms can also be applied, depending on the detected local changes of consecutive images caused by moving objects as well as skipping some frames or their fragments which differ significantly from the others. In the paper a proposition and verification of the modified version of previously proposed hybrid approach to background estimation [5] is presented, which allows the faster convergence of background estimation and increase of the stability of obtained results.

1. WSTĘP

Współczesne Inteligentne Systemy Transportowe (ITS) w coraz większej mierze wykorzystują techniki wizyjne oparte na analizie obrazów i sekwencji wideo pozyskanych z kamer montowanych w bezpośrednim sąsiedztwie drogi. Pomimo tego, iż informacje uzyskiwane w sposób automatyczny z obrazów często są obciążone błędami uniemożliwiającymi ich pełne wykorzystanie np. jako materiał dowodowy, również z powodów formalnych (ekstrakcja informacji wymaga zazwyczaj przetwarzania pozyskanych obrazów), stanowią one atrakcyjną alternatywę dla tradycyjnych sposobów pomiaru np. pętli indukcyjnych wykorzystywanych do pomiaru natężenia ruchu dla potrzeb ITS. Głównymi powodami dynamicznego rozwoju tego rodzaju technik jest stosunkowo niski koszt sprzętu o dobrych parametrach, jak również znaczny postęp technik analizy obrazów, jaki ma miejsce przez co najmniej kilkanaście ostatnich lat.

Jednym z zasadniczych elementów wizyjnych systemów automatycznego nadzoru ruchu jest możliwość śledzenia pojazdów. Możliwe są tutaj do wykorzystania różne podejścia, spośród których szczególnie atrakcyjne wydają się być algorytmy śledzenia przed detekcją (Track-Before-Detect) umożliwiające zwiększenie zasięgu systemu oraz odporności na zakłócenia spowodowane np. zmiennymi warunkami atmosferycznymi. Jednym z najistotniejszych elementów warunkujących poprawność i wysoką dokładność śledzenia jest przetwarzanie wstępne obrazów związane z estymacją i usuwaniem tła, stanowiące wciąż aktywny obszar zainteresowania badaczy [3,7], dzięki czemu znacząco zmniejsza się ilość danych przekazywanych do algorytmu śledzenia powodując wzrost jego wydajności.

¹ Wyższa Szkoła Techniczno-Ekonomiczna w Szczecinie, Wydział Transportu Samochodowego;
71-244 Szczecin; ul. Klonowica 14. Tel: +48 91 424-08-75, Fax: +48 91 424-08-76 E-mail: okarma@wste.szczecin.pl, mazurek@wste.szczecin.pl

2. ALGORYTMY ESTYMACJI TŁA W SEKWENCJACH WIDEO

Najprostszym podejściem do zagadnienia estymacji tła w sekwencji wideo jest dokonanie detekcji pikseli obrazu, które różnią się w obrębie określonej liczby sąsiednich klatek. W podstawowej wersji mamy do czynienia z algorytmem różnicowym, dla którego piksele zgodne z ich odpowiednikami z poprzedniej klatki są traktowane jako tło, a pozostałe interpretowane jako reprezentujące ruchomy obiekt. Oczywiście tego rodzaju podejście nie jest w żadnej mierze odporne na wpływ szumów i zakłóceń, zmiennych warunków oświetleniowych, czy nawet wpływu różnego rodzaju warunków atmosferycznych powodujących zmiany jasności niewielkich elementów na obrazie (np. liście drzew poruszające się na wietrze, opady śniegu itp.) [4]. W celu ograniczenia ich wpływu można stosować dodatkowe progowanie np. oparte na analizie wielkości obiektów sklasyfikowanych jako ruchome. Metoda taka jest co prawda szybka i łatwa w implementacji, jednak ma poważne ograniczenia związane z wymaganiem stałych warunków oświetleniowych, a także z prędkością pojazdów oraz wydajnością kamery (liczba klatek na sekundę możliwych do rejestracji, transmisji i przetworzenia).

Inną liniową metodą estymacji tła jest uśrednianie określonej liczby klatek sekwencji wideo, jednak jest to metoda powolna o dużym zapotrzebowaniu na pamięć, dlatego też zazwyczaj jest stosowana w zmodyfikowanej wersji jako średnia ruchoma (Moving Average) obejmująca określoną liczbę poprzednich klatek lub też w postaci wygładzania wykładniczego. W niektórych zastosowaniach wykorzystywana jest także średnia ważona, w której minimalizowany jest wpływ wcześniejszych klatek na rzecz ostatnio zarejestrowanych. Dodatkową stosowaną modyfikacją liniowych metod estymacji tła jest oznaczanie pikseli sklasyfikowanych jako reprezentujące obiekty ruchome w celu eliminacji ich wpływu na wyznaczone wartości średnie jasności pikseli reprezentujących tło.

Jedną z istotniejszych niedogodności liniowych algorytmów estymacji tła jest konieczność doboru właściwego progów, zazwyczaj bazującego na pojedynczej wartości jasności, co niekoniecznie jest właściwym rozwiązaniem w wypadku złożonego obrazu tła. W takiej sytuacji może być zastosowane rozwiązanie oparte na uśrednianiu gaussowskim, jednak nie jest ono również pozbawione wad związanych z inicjalizacją oraz aktualizacją danych, a także koniecznością podziału modeli na reprezentujących tło oraz pierwszy plan [6].

2.1 Inicjalizacja liniowych algorytmów estymacji tła

Niezależnie od zastosowanego algorytmu estymacji tła, wyniki jego działania, w szczególności szybkość zbieżności są zależne w znacznej mierze od inicjalizacji algorytmu. W najprostszym wypadku możliwy jest start od wartości zerowych, co odpowiada całkowicie czarnemu obrazowi lub maksymalnych (obraz biały). Jednym z klasycznych algorytmów estymacji tła jest wygładzanie wykładnicze zgodne z zależnością

$$B_t^{EXP}(u, v) = \alpha \cdot B_{t-1}^{EXP}(u, v) + (1 - \alpha) \cdot I_t(u, v) \quad (1)$$

gdzie: α - parametr wygładzania (wartość zbliżona do 1),
 $I(u, v)$ - jasność piksela o współrzędnych (u, v) ,
 t - numer kroku.

Analizując zbieżność algorytmu wygładzania wykładniczego można łatwo zauważyć, iż korzystniejszym rozwiązaniem jest inicjalizacja wartością średnią zakresu jasności, co umożliwi przyspieszenie zbieżności algorytmu dla tej samej wartości współczynnika wygładzania α . Dzięki zastosowaniu jego wartości bliskiej jedności uzyskuje się redukcję wpływu szumu na uzyskiwane wyniki. Pomimo stosunkowo wysokiej czułości na szum impulsowy algorytm wygładzania wykładniczego zapewnia stabilne wyniki estymacji tła, jednak jego zasadniczą wadą jest stosunkowo powolna zbieżność w początkowej fazie działania.

Klasyczne liniowe podejście do estymacji tła oparte jest na założeniu, iż jasność wynikowego piksela jest liniową kombinacją jasności tego piksela z poprzednich klatek sekwencji wideo. Zakładając, iż obiekty (pojazdy) są ruchome, poruszając się względem nieruchomego tła (drogi i jej otoczenia), dobrą estymację tła można uzyskać po stosunkowo dużej liczbie kroków odpowiadających dużej liczbie zarejestrowanych klatek sekwencji wideo, tak aby wpływ jasności obiektów ruchomych widocznych na poszczególnych klatkach stał się nieistotny względem pikseli reprezentujących tło. Przykładowo jasny obiekt widoczny przez 50 ze 100 klatek w określonym fragmencie obrazu z ciemnym tłem spowodowałby znaczące rozjaśnienie wartości średnich, natomiast w wypadku jego obecności przez 50 klatek z 1000 wpływ ten znacznie maleje. Z tego też względu wszystkie algorytmy liniowe bazujące na uśrednianiu wartości wymagają długiego czasu działania niezbędnego do ustabilizowania wartości poszczególnych pikseli estymaty tła.

2.2 Nieliniowa estymacja tła

W celu uzyskania szybszej zbieżności możliwe jest zastosowanie algorytmów nieliniowych, spośród których większość oparta jest na filtracji medianowej. Biorąc pod uwagę właściwości filtru medianowego związane z odpornością na gwałtowne lokalne zmiany jasności (zazwyczaj szum impulsowy), można uzyskać dobre efekty także w algorytmach estymacji tła ze względu na fakt, iż tego rodzaju zmiany jasności pikseli są również charakterystyczne dla ruchu obiektów na statycznym tle [1].

Podstawowa wersja algorytmu medianowego zakłada sortowanie jasności dla określonego piksela w ciągu określonej liczby kolejnych klatek sekwencji wideo i wybór elementu środkowego (mediany) jako wyniku. Ze względu na fakt, iż na drodze mogą znajdować się zarówno pojazdy od niej jaśniejsze, jak również ciemniejsze (np. czarne), w typowych sytuacjach jasność wynikowego piksela jest zbliżona do jasności piksela reprezentującego tło.

W celu zredukowania wpływu szumu oraz przyspieszenia obliczeń możliwe jest wykorzystanie algorytmu medianowego na podstawie jedynie wybranych klatek sekwencji wideo z określonym odstępem czasowym pomiędzy nimi (temporal downsampling). W takim wypadku znaczącemu zmniejszeniu ulega również wpływ poruszających się obiektów na końcowy wynik estymacji tła.

2.3 Hybrydowa estymacja tła

W jednej z poprzednich prac [5] zaproponowano wykorzystanie podejścia hybrydowego łączącego zalety szybkiej zbieżności nieliniowego algorytmu medianowego oraz stabilność działania wygładzania wykładniczego. Zaproponowany algorytm oparty został na inicjalizacji z wykorzystaniem filtra medianowego z wykorzystaniem wybranych klatek, której wynik stanowi estymatę startową dla algorytmu wygładzania wykładniczego działającego w drugiej fazie.

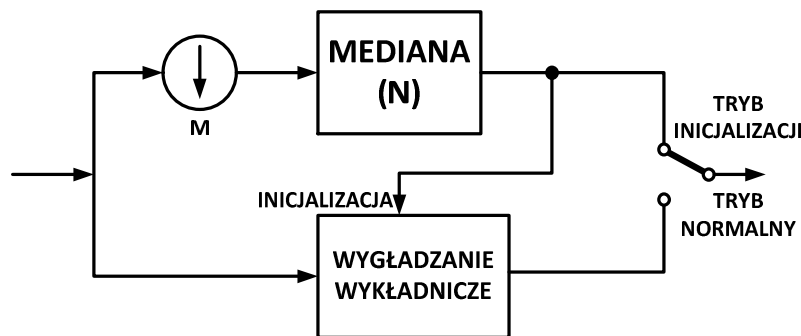
Zasada działania omawianego algorytmu może być przedstawiona za pomocą następującego wzoru;

$$B_i^H(u, v) = \begin{cases} \text{MEDIAN}\{I_i(u, v), I_{i-M}(u, v), \dots, I_{i-i \cdot M}(u, v), \dots, I_{i-(N-1)M}(u, v)\} & : t < T \\ \alpha \cdot B_{i-1}^H(u, v) + (1 - \alpha) \cdot I_i(u, v) & : t \geq T \end{cases} \quad (2)$$

gdzie: T - moment przełączenia,

M - liczba opuszczanych klatek w algorytmie medianowym.

Hybrydowy algorytm estymacji tła wykorzystuje wartość mediany obliczanej dla jasności pikseli z N klatek wyznaczonych z zachowaniem określonego odstępem (M) do inicjalizacji algorytmu wygładzania wykładniczego, co przedstawiono na rysunku 1.



Rys.1. Zasada działania hybrydowego algorytmu estymacji tła

3. PROPONOWANY ZMODYFIKOWANY ALGORYTM ESTYMACJI TŁA

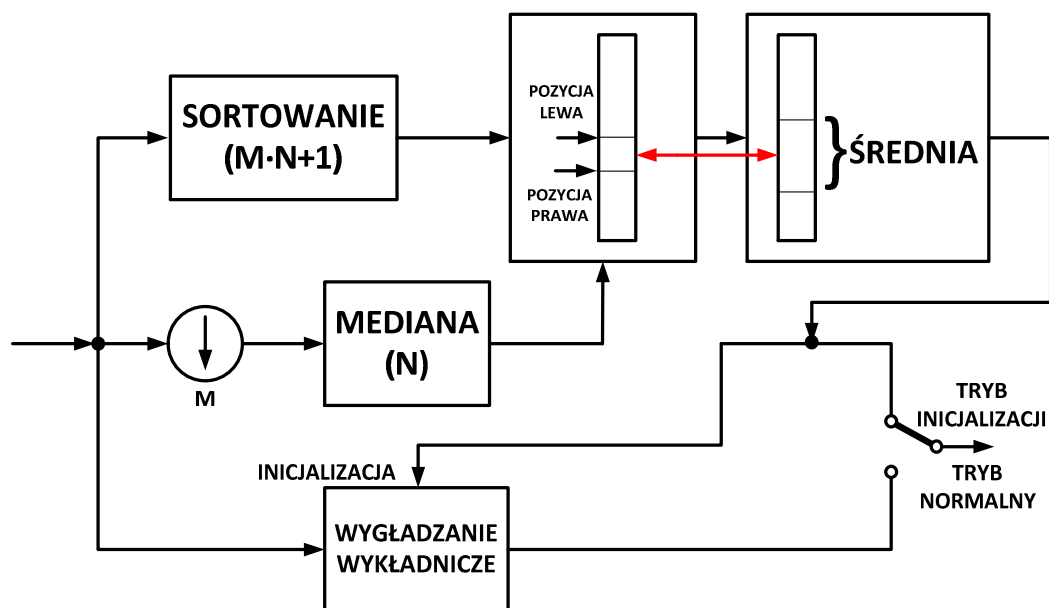
Ze względu na fakt, iż głównym celem zaproponowanego uprzednio algorytmu hybrydowego jest zapewnienie szybszej zbieżności do obrazu tła w porównaniu do klasycznej techniki uśredniania, jak również algorytmu wygładzania wykładniczego, jego modyfikacje powinny dotyczyć przede wszystkim fazy inicjalizacji. W celu weryfikacji uzyskiwanych wyników przyjęto, iż referencyjnym obrazem tła jest obraz uzyskiwany w wyniku uśrednienia 10000 klatek testowej sekwencji wideo. Biorąc pod uwagę zmiany globalnego oświetlenia sceny w czasie rejestracji sekwencji wideo, nie jest możliwe uzyskanie identycznego rezultatu w krótszym czasie (na podstawie fragmentu sekwencji wideo), jednak możliwe jest osiągnięcie obrazu o znacznym stopniu podobieństwa z nieznacznymi różnicami jasności pikseli.

Proponowana modyfikacja dotycząca sposobu wyznaczania estymaty tła do inicjalizacji algorytmu wygładzania wykładniczego polega na eliminacji wpływu obiektów ruchomych wpływających na wartość uzyskiwanej mediany wraz z uśrednieniem wartości jej bliskich. W pierwszym kroku następuje wyznaczenie wartości mediany dla każdego pikseli na podstawie N klatek z pominięciem określonej liczby $M-1$ klatek pomiędzy próbkami, w ten sam sposób jak w metodzie przedstawionej poprzednio. Następnie wykonywane jest sortowanie wartości jasności danego pikseli ze wszystkich klatek w tym samym fragmencie sekwencji wideo (bez pomijania klatek), po czym wyznaczane są skrajne pozycje uprzednio wyznaczonej mediany w uzyskanym ciągu. Wynikiem końcowym dla danego pikseli jest wartość średnia elementów ciągu w oknie o ustalonej szerokości K elementów, gdzie środek okna stanowi pozycja wyznaczona jako średnia arytmetyczna z wyznaczonych pozycji skrajnych. Zasadę tę ilustruje rysunek 2, w którym zaznaczone zostało środkowe położenie wartości wyznaczonej estymaty mediany w uzyskanym posortowanym ciągu wartości jasności pikseli w kolejnych $M \cdot N + 1$ klatkach sekwencji wideo. Proponowany algorytm może być opisany jako:

$$B_i^H(u, v) = \begin{cases} \text{MEAN}\{I_{pos-(K-1)/2}(u, v), \dots, I_{pos}(u, v), \dots, I_{pos+(K-1)/2}(u, v)\} & : t < T \\ \alpha \cdot B_{i-1}^H(u, v) + (1 - \alpha) \cdot I_i(u, v) & : t \geq T \end{cases} \quad (3)$$

gdzie: K - szerokość okna uśredniania,

pos - średnia pozycja wartości równej wyznaczonej estymacie mediany w ciągu.



Rys.2. Zasada działania proponowanego zmodyfikowanego hybrydowego algorytmu estymacji tła

4. WERYFIKACJA ZBIEŻNOŚCI ALGORYTMÓW Z UŻYCIEM METOD PORÓWNAWCZEJ OCENY JAKOŚCI OBRAZÓW

Jedną z najczęściej stosowanych technik weryfikacji działania algorytmów przetwarzania obrazów jest porównanie uzyskiwanych wyników z obrazem referencyjnym, przy założeniu, iż jest on dostępny. Podejście takie może być stosowane w szczególności przy ocenie działania algorytmów stratnej kompresji obrazów i sekwencji wideo, czy też filtracji obrazów. Ze względu na fakt, iż zazwyczaj celem algorytmu przetwarzania obrazu jest poprawa jego jakości (z wyłączeniem technik analizy obrazu prowadzących do ekstrakcji informacji), typowym podejściem jest wykorzystanie wskaźników jakości obrazów.

Klasyczne metody oceny jakości obrazów [2], stosowane od wielu lat oparte są na błędzie średniokwadratowym (Mean Squared Error - MSE). Jedną z częściej stosowanych jest szczytowy stosunek sygnału do szumu (Peak Signal to Noise Ratio - PSNR) wyrażany w decybelach definiowany jako:

$$PSNR = 10 \cdot \log_{10} \sum_{u,v} \left(\frac{k^2}{|B(u,v) - Q(u,v)|^2} \right) \quad (4)$$

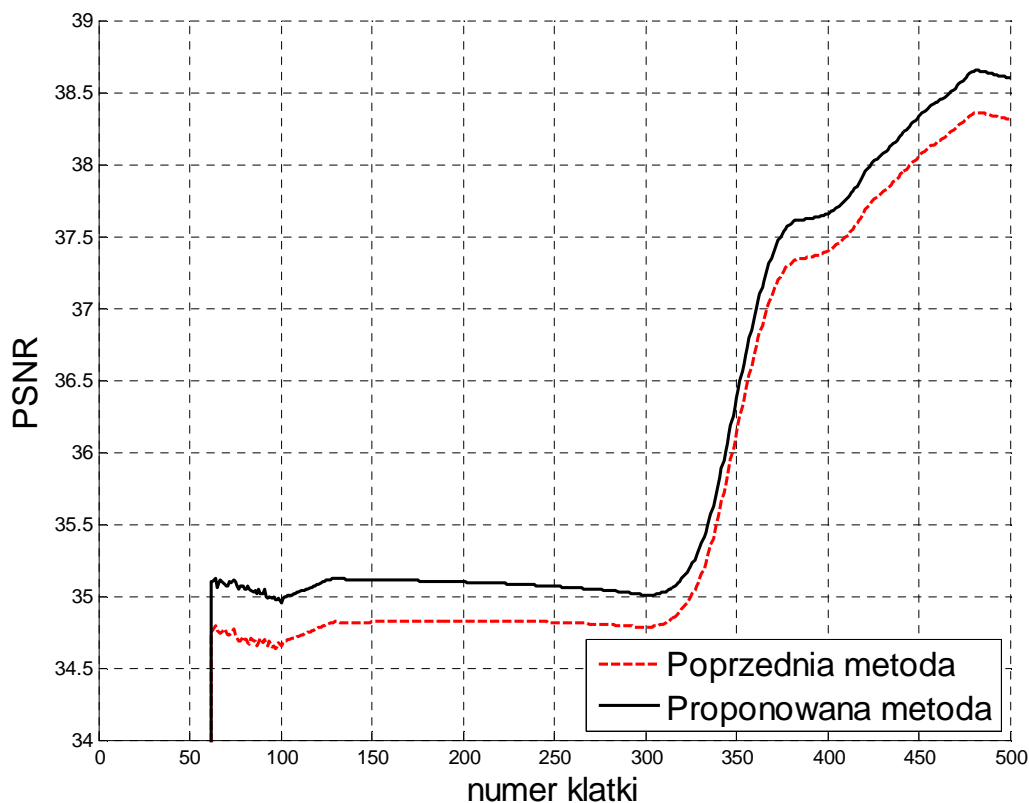
gdzie: $B(u,v)$ - wartość jasności piksela estymaty tła o współrzędnych (u,v) ,
 $Q(u,v)$ - wartość jasności piksela referencyjnego obrazu tła o współrzędnych (u,v) ,
 k - liczba poziomów szarości (zwykle 256).

Ze względu na niską korelację tego rodzaju wskaźników z ocenami subiektywnymi, w ostatnich latach zaproponowane zostało nowe podejście do zagadnienia oceny jakości obrazów cyfrowych, bazujące na ocenie podobieństwa obrazów. Najpopularniejszym tego rodzaju wskaźnikiem jest podobieństwo strukturalne (Structural Similarity – SSIM) definiowane jako średnia wartość wskaźników lokalnych wyznaczanych w obrębie przesuwającego okna Gaussa o rozmiarze 11×11 pikseli zgodnie z następującą zależnością [8]:

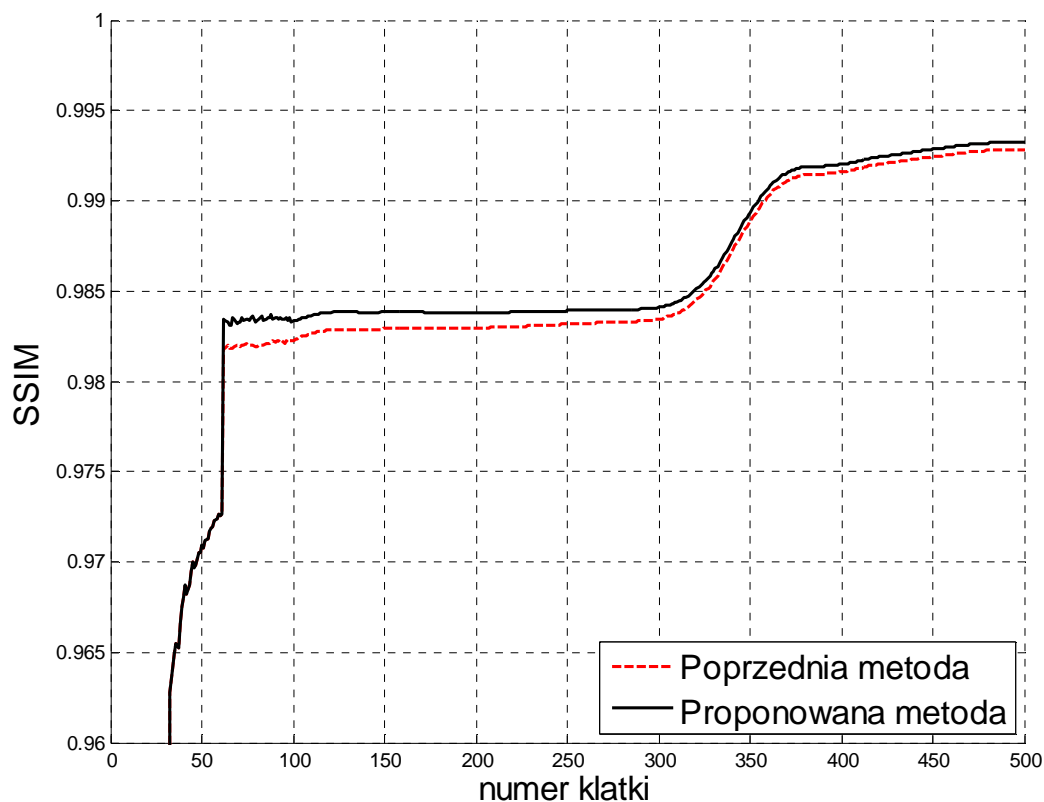
$$SSIM = \frac{(2 \cdot \bar{x}\bar{y} + C_1) \cdot (2 \cdot \sigma_{xy} + C_2)}{(\bar{x}^2 + \bar{y}^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (5)$$

gdzie: C_1, C_2 - stałe współczynniki zapobiegające dzieleniu przez zero,
 \bar{x}, \bar{y} - wartości średnie obrazu referencyjnego (x) i ocenianego (y) w obrębie lokalnego okna,
 σ^2 - wariancja jasności dla fragmentu obrazu wewnątrz lokalnego okna,
 σ_{xy} - kowariancja jasności wewnątrz lokalnego okna.

Traktując obraz tła uzyskany w wyniku uśredniania 10000 próbek jako referencyjny, wykonano badania eksperymentalne z wykorzystaniem zarejestrowanych testowych sekwencji wideo, wyznaczając wskaźniki podobieństwa uzyskanej estymaty tła z tym obrazem w postaci wartości PSNR oraz SSIM. Wyniki uzyskane dla metody hybrydowej oraz jej zmodyfikowanej wersji zaproponowanej w artykule zostały przedstawione na rysunkach 3 oraz 4.



Rys.3. Wartości współczynnika PSNR dla estymacji tła uzyskanej poprzednią metodą hybrydową oraz proponowaną metodą zmodyfikowaną



Rys.4. Wartości współczynnika SSIM dla estymacji tła uzyskanej poprzednią metodą hybrydową oraz proponowaną metodą zmodyfikowaną

Wyniki oceny podobieństwa referencyjnego obrazu tła oraz uzyskanych estymat zostały wyznaczone dla tej samej liczby próbek służących do estymacji mediany ($N=61$) z odstępem czterech klatek (co odpowiada wartości $M=5$), przy czym dla proponowanej metody przyjęto $K=25$ wartości poddanych uśrednianiu. Przykładowe klatki sekwencji testowej, referencyjny obraz tła oraz zaznaczone piksele sklasyfikowane jako tła zostały przedstawione na rysunku 5.



Rys.5. Przykładowy obraz wejściowy (klatka sekwencji testowej), referencyjny obraz tła oraz obraz z zaznaczonymi pikselami sklasyfikowanymi jako tło

5. WNIOSKI

Uzyskane rezultaty świadczą o poprawie zbieżności metody, co pozwala na poprawę dokładności działania algorytmów śledzenia pojazdów, w których wykorzystywana jest estymacja tła w celu jego usunięcia z obrazu wejściowego dla algorytmu śledzenia ruchu.

Zaproponowane podejście łączy zalety zarówno algorytmu medianowego, jak również algorytmów liniowych, czyli średniej ruchomej oraz wygładzania wykładniczego. Efekty zastosowania obu rodzaju algorytmów w sposób niezależny zostały przedstawione w postaci graficznej we wcześniejszym artykule [5], przy czym warto zwrócić uwagę na wrażliwość obu metod na pojawiające się ruchome obiekty, co zostało w znacznej mierze zredukowane w proponowanym algorytmie.

6. BIBLIOGRAFIA

- [1] Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting Moving Objects, Ghosts and Shadows in Video Streams. IEEE Transactions on Pattern Analysis and Machine Intelligence vol. 25, no. 10, str. 1337–1342, 2003.
- [2] Eskicioglu, A.: Quality Measurement for Monochrome Compressed Images in the Past 25 Years. Proceedings of the International Conference on Acoustics Speech & Signal Processing, str. 1907–1910, Istanbul, Turkey, 2000.
- [3] Maddalena, L., Petrosino, A.: A Self-organizing Approach to Background Subtraction for Visual Surveillance Applications. IEEE Transactions on Image Processing, vol. 17, no. 7, str. 1168–1177, 2008.
- [4] Okarma, K., Mazurek, P.: Background Estimation Algorithm for Optical Car Tracking Applications. Machinebuilding and Electrical Engineering no. 7–8, str. 7–10, 2006.
- [5] Okarma, K., Mazurek, P.: Nonlinear background estimation methods for video vehicle tracking systems. Archives of Transport Systems Telematics vol. 4 issue 4, str. 42–48, 2011.
- [6] Piccardi, M.: Background subtraction techniques: a review. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, The Hague, Netherlands, str. 3099–3104, October 2004.
- [7] Reddy, V., Sanderson, C., Lovell, B.C.: A Low-Complexity Algorithm for Static Background Estimation from Cluttered Image Sequences in Surveillance Contexts. EURASIP Journal on Image and Video Processing, Article ID 164956, 14 str., 2011.
- [8] Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image Quality Assessment: From Error Measurement to Structural Similarity. IEEE Transactions on Image Processing, vol. 13, no. 4, str. 600–612, 2004.

7. PODZIĘKOWANIA

Artykuł powstał dzięki wsparciu w ramach grantu Ministerstwa Nauki i Szkolnictwa Wyższego nr N509 399136 „Estymacja trajektorii ruchu pojazdów z wykorzystaniem analizy bayesowskiej oraz algorytmów cyfrowego przetwarzania obrazów“.