

Zbigniew CZAPLA, Wiesław PAMUŁA

LOKALIZACJA OBIEKTÓW NA OBRAZACH CYFROWYCH Z WIDEOREJSTRATORA RUCHU DROGOWEGO

Streszczenie. Artykuł przedstawia zagadnienia analizy scen ruchu drogowego. Jako metodę przetwarzania wstępnego w detekcji cech punktowych obiektów wykorzystano konwersję obrazu do dwuwarstwowego modelu danych. Do dwuwarstwowego modelu danych obrazu zastosowano detektory cech punktowych FAST i SUSAN. Do lokalizacji obiektów nadaje się – i jest prosta obliczeniowo – detekcja cech punktowych obiektów (przy zastosowaniu dwuwarstwowego modelu danych).

OBJECT LOCATION IN DIGITAL IMAGES FROM A ROAD TRAFFIC VIDEORECORDER

Summary. The paper presents issues of road traffic scene analysis. Image conversion into the two-layer data model is used as a preliminary processing method in detection of object feature points. The feature points detectors FAST and SUSAN are applied to two-layer data model of an image. Detection of object feature points using the two-layer data model is suitable for the object location and it is computationally simple.

1. WSTĘP

We współczesnych systemach transportowych do sterowania wykorzystywane są parametry ruchu drogowego. Stosuje się wiele różnych metod ich wyznaczania, m.in. można wykorzystać analizę obrazów cyfrowych otrzymywanych z wideorejstratora ruchu drogowego [1].

Analiza obrazu obejmuje różne techniki: segmentację obrazu, indeksację obiektów oraz ich klasyfikację. Danymi wejściowymi w tego rodzaju analizie jest obraz, natomiast danymi wyjściowymi jest zestaw atrybutów opisujących zawartość obrazu. Segmentacja obrazów jest techniką, w której tworzone są obiekty przez łączenie obszarów obrazu zgodnie z wybranym kryterium jednorodności. Podstawowym kryterium jednorodności jest ustalony zakres wartości pikseli. Najbardziej znane metody segmentacji to: segmentacja przez podział obrazu,

przez rozrost obrazu, wykorzystująca wykrywanie krawędzi oraz oparta na statystyce [2], [3], [4].

Przed analizą obrazu często przeprowadzane jest jego przetwarzanie, które jest przeprowadzane w celu poprawy jakości obrazu, (np. przez redukcję zakłóceń, zwiększenie kontrastu) lub w celu uwypuklenia cech obrazu przydatnych do jego analizy, (np. przez wyodrębnianie krawędzi i binaryzację). Danymi wejściowymi przy przetwarzaniu obrazu jest obraz źródłowy, natomiast danymi wyjściowymi jest przetworzony obraz wynikowy [2], [3], [4].

Obrazy cyfrowe przechowywane są w plikach graficznych, których rozmiary są zazwyczaj duże. Rozmiar plików graficznych można zmniejszyć przeprowadzając kompresję, która może być bezstratna lub stratna. Kompresja bezstratna pozwala na odtworzenie niezmiennego obrazu źródłowego, stratna jest bardziej efektywna, natomiast jednak powoduje utratę części informacji obrazowej. Znanych jest wiele metod kompresji zarówno stratnych, jak i bezstratnych [5], [6].

W metodach przetwarzania i analizy obrazów można stosować wstępne przetwarzanie obrazów, które ma na celu zmianę formatu danych w postać bardziej odpowiednią dla wykorzystywanych algorytmów.

Dwuwarstwowy model danych obrazów cyfrowych jest reprezentacją obrazu, która powstaje w wyniku kodowania predykcyjnego dostarczającego informacji o charakterze zmian w otoczeniu poszczególnych pikseli i może być wykorzystane do lokalizacji obiektów na obrazie.

W artykule przeprowadzono porównanie wyników detekcji cech punktowych obiektów z użyciem krawędzi oraz z wykorzystaniem dwuwarstwowego modelu danych. Wybrano też do oceny detektory narożników SUSAN i FAST, które najczęściej są stosowane w śledzeniu obiektów w sekwencji obrazów [11].

2. DWUWARSTWOWY MODEL DANYCH OBRAZÓW CYFROWYCH

Dwuwarstwowy model danych obrazu cyfrowego tworzy nową reprezentację obrazu, składającą się z dwóch warstw. Jedna z warstw zawiera wartości bazowe pikseli, druga składa się z wartości różnicowych, określających różnicę między wartością piksela a odpowiadającą mu wartością bazową [7].

Wiele obrazów zawiera obszary pikseli o zbliżonych wartościach. Takie obszary mogą zostać opisane przez ciągi wartości pikseli, zawierające wartość bazową obszaru oraz sekwencję wartości różnicowych, które odnoszą się do poszczególnych pikseli opisywanego obszaru. Reprezentacja całego obrazu tworzona jest przez zbiór ciągów pikseli wszystkich obszarów obrazu. Zbiór wartości różnicowych wszystkich pikseli tworzy warstwę różnicową dwuwarstwowego modelu danych, natomiast zbiór wartości bazowych – warstwę bazową modelu. Każda z warstw jest kodowana oddzielnie. Założenie, że wartość różnicowa jest liczbą całkowitą z zakresu od 0 do 15 pozwala na kodowanie w jednym bajcie dwóch wartości różnicowych. Wartości bazowe obszarów są kodowane bez zmian – jedna wartość w jednym bajcie.

2.1. Konwersja obrazu do modelu dwuwarstwowego

Konwersja obrazów cyfrowych do dwuwarstwowego modelu danych przeprowadzana jest na obrazach o rozdzielczości poziomów jasności 8 bitów na piksel. Taka rozdzielczość jest wystarczająca do większości zastosowań przemysłowych.

Obraz źródłowy, poddawany konwersji do modelu dwuwarstwowego, opisywany jest przez dwuwymiarową macierz X , na podstawie której, w wyniku skanowania, tworzony jest wektor obrazu Y . Wektory konwersji D i B przeznaczone są do przechowywania wartości różnicowych pikseli (wektor D) i wartości bazowych obszarów (wektor B). Obraz źródłowy dzielony jest na kwadratowe bloki o rozmiarze 2×2 piksele, a następnie bloki te są kolejno konwertowane do wektora obrazu Y .

Po przekonwertowaniu wszystkich bloków obrazu do wektora Y ustalana jest pierwsza wartość bazowa i jest ona zapisywana do wektora konwersji B :

$$b_0 = y_0. \quad (1)$$

Pierwszą wartością zapisaną do wektora konwersji D jest znacznik wartości bazowej c :

$$d_0 = c. \quad (2)$$

Odczytywane są elementy wektora obrazu Y i uwzględniane są dwa przypadki. W pierwszym z nich wartość różnicowa mieści się na 4 bitach i wtedy jest ona zapisywana do wektora konwersji D :

$$d_i = \Delta y_i. \quad (3)$$

W drugim przypadku wartość różnicowa nie mieści się na 4 bitach i wtedy ustalana jest nowa wartość bazowa; do wektora konwersji D zapisywany jest znacznik wartości bazowej c , a do wektora konwersji B zapisywana jest nowa wartość bazowa. Indeksy elementów wektorów określają bieżącą zapisywaną pozycję, z uwzględnieniem przesunięcia między indeksami wektorów B i C :

$$\begin{aligned} d_i &= c, \\ b_{i-\Delta i} &= y_i. \end{aligned} \quad (4)$$

Strumień wyjściowy S składa się z dwóch strumieni składowych S_1 oraz S_2 :

$$S = S_1 + S_2. \quad (5)$$

Strumień składowy S_1 powstaje przez połączenie sąsiednich wartości różnicowych wektora konwersji D w jeden bajt:

$$s_i = d_i \cdot 16 + d_{i+1}, \quad (6)$$

a strumień składowy S_2 tworzony jest bezpośrednio przez wektor konwersji B :

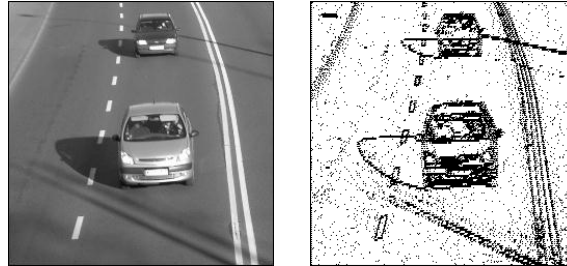
$$S_2 = B. \quad (7)$$

2.2. Konwersja powrotna obrazu do modelu jednowarstwowego

Konwersja powrotna obrazu do modelu jednowarstwowego odbywa się w odwrotnej kolejności niż konwersja obrazu do modelu dwuwarstwowego i obejmuje: podzielenie strumienia S na strumienie składowe S_1 i S_2 , wydzielenie wektorów konwersji B i D , odtworzenie wektora obrazu Y oraz odtworzenie macierzy obrazu X .

2.3. Wyniki konwersji do modelu dwuwarstwowego

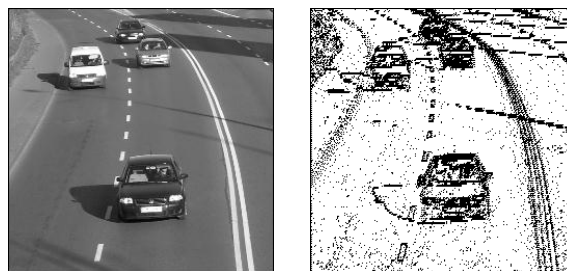
Konwersji do modelu dwuwarstwowego poddano cztery obrazy testowe o różnym poziomie szczegółowości. Wszystkie obrazy testowe miały rozdzielczość poziomów jasności 8 bitów na piksel i rozdzielczość przestrzenną 256 x 256 pikseli. Wyniki konwersji do modelu dwuwarstwowego przedstawiają odpowiednio rysunki 1 do 4.



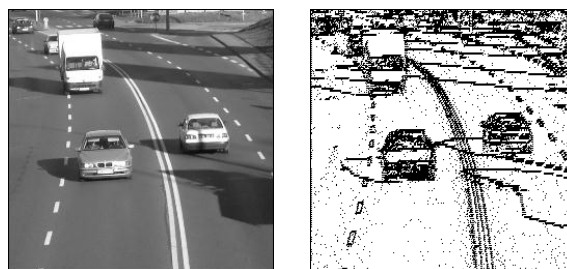
Rys. 1. Wynik konwersji obrazu 1
Fig. 1. Result of conversion of image 1



Rys. 2. Wynik konwersji obrazu 2
Fig. 2. Result of conversion of image 2



Rys. 3. Wynik konwersji obrazu 3
Fig. 3. Result of conversion of image 3



Rys. 4. Wynik konwersji obrazu 4
Fig. 4. Result of conversion of image 4

Po lewej stronie rysunków znajduje się obraz źródłowy, po prawej – obraz przedstawiający rozmieszczenie wartości bazowych. We wszystkich przypadkach uzyskano zmniejszenie rozmiaru reprezentacji obrazu o ponad 20%. Rozkład punktów bazowych odpowiada zawartości obrazu i jest przydatny do lokalizacji obiektów znajdujących się na obrazach.

3. DETEKTORY CECH PUNKTOWYCH OBIEKTÓW

Cechy punktowe obiektów pozwalają zmniejszyć rozmiar ich opisu na obrazie i mogą ułatwić przeprowadzenie operacji odnajdywania korespondujących elementów obiektów na kolejnych obrazach w sekwencji. Wymagają uważnego doboru progów dyskryminacji dla uzyskania istotnych do śledzenia elementów obiektów. Ważnymi własnościami punktów charakterystycznych są:

- mała czułość na zmiany orientacji, przesunięcia obiektów,
- odporność na szum, wahania wartości pikseli obiektów w czasie – wywołane zmianą oświetlenia pola obserwacji,
- mała złożoność obliczeniowa,
- łatwość przeprowadzania operacji dopasowania w algorytmach śledzenia.

Cechy punktowe wyznacza się, analizując zmiany wartości pikseli lub wyliczając statystyki wartości w kołowym obszarze wokół danego piksela. Przekroczenie zadanego progu detekcji wskazuje na wystąpienie cechy punktowej. W zastosowaniach ważne są detektory „narożników” Harrisa, SUSAN, FAST [8], [9], [10].

Detektor Harrisa wylicza cechy rozkładu gradientów w sąsiedztwie danego piksela. Na podstawie współczynników macierzy autokorelacji wartości pikseli M_i wynosi:

$$M_i = w_\sigma * \begin{bmatrix} A & C \\ C & B \end{bmatrix} \quad (8)$$

$$A = (w_{x\sigma} * P_i)^2 \quad B = (w_{y\sigma} * P_i)^2 \quad C = (w_{x\sigma} * P_i)(w_{y\sigma} * P_i),$$

gdzie: w_σ , $w_{x\sigma}$, $w_{y\sigma}$, – wektory wag odpowiednio: dla wyliczania gausowskiego filtru wygładzającego, gradientu w kierunku poziomym, gradientu w kierunku pionowym.

Określona zostaje miara obecności „narożników” R_H :

$$R_H = \det(M_i) - k \cdot \text{trace}(M_i)^2 \quad (9)$$

$$R_H = (M_{i11}M_{i22} - M_{i12}M_{i21}) - k \cdot (M_{i11} + M_{i22})^2.$$

Wyliczanie cech punktowych Harrisa wymaga wyliczania kierunkowych gradientów zmian jasności pikseli i stąd wykluczona jest możliwość wykorzystania reprezentacji z użyciem zaproponowanego dwuwarstwowego modelu danych.

SUSAN pozwala sklasyfikować zawartość kołowego otoczenia piksela. Używa się aproksymacji koła o średnicy do kilkunastu pikseli (najczęściej 3, 7, 11 – rys. 5). Wartości pikseli w sąsiedztwie $P_i(j)$ porównuje się z wartością piksela centralnego $P_i(x_0, y_0)$, zliczając piksele różniące się o wartość *progu*. Różnica liczby pikseli sąsiedztwa L i różniących się E_i określa R_S – miarę obecności narożników. Wartość progu określa zdolność do eliminacji szumów i czułość współczynnika. Dla $L=K/2$ współczynnik wskazuje na występowanie

„narożników”, dla $L=3K/4$ – krawędzi, gdzie: K – liczba pikseli w sąsiedztwie – w kole, otaczającym piksel centralny.

$$e(j) = \begin{cases} 1 & \text{gdy } |P_i(j) - P_i(x_0, y_0)| \leq \text{próg} \\ 0 & \text{gdy } |P_i(j) - P_i(x_0, y_0)| > \text{próg} \end{cases}$$

$$E_i = \sum_{j=1}^K e(j) \quad (10)$$

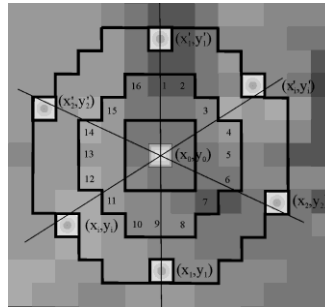
$$R_s = \begin{cases} L - E_i & \text{gdy } E_i < L \\ 0 & \end{cases}$$

Najmniej złożony obliczeniowo algorytm jest podstawą do wyliczania miary cech detektora FAST (ang. *Features from Accelerated Segment Test*); cechy wyznaczane są na podstawie szybkiego sprawdzenia obwodu. Wyznaczona wielkość określa długość nieprzerwanego wycinka obwodu okręgu, wyrażoną liczbą pikseli, którego piksele różnią się od wartości piksela – środka okręgu – o ustaloną wartość. Gdy liczba przekroczy zadany próg, uznaje się środek za punkt charakterystyczny.

$$l(j) = \begin{cases} 1 & \text{gdy } |P_i(j) - P_i(x_0, y_0)| > \text{próg} \\ 0 & \text{gdy } |P_i(j) - P_i(x_0, y_0)| \leq \text{próg} \end{cases} \quad (11)$$

$$R_F = \sum_{j \in \text{obwód}} l(j)$$

Rozpatrywany jest zwykle okrąg o średnicy 7 pikseli (rys. 5), który ma 16 pikseli na obwodzie. Piksele, dla których R_F wynosi 9, najczęściej są narożnikami (jest to też najmniejsza wartość jednoznacznie wskazująca na to, że okrąg nie zawiera krawędzi).



Rys. 5. Kołowe otoczenia piksela do wyznaczania SUSAN i FAST
Fig. 5. Pixel neighbourhoods for evaluating SUSAN and FAST

3.1. Implementacja detektorów cech punktowych

Korzystając ze strumienia wartości bazowych S_2 (7) reprezentacji obrazu, można zaproponować uproszczone wersje detektorów. Wartości bazowe w niewielkim otoczeniu (a takie są otoczenia detekcji) określają piksele istotne dla obecności obiektu, można zatem na ich podstawie określić cechy tego obiektu.

Analizując liczbę wartości bazowych w otoczeniu lub na obwodzie, można uzyskać równoważny detektor cech punktowych.

Detektor SUSAN: zliczane są piksele z wartościami bazowymi w kole piksela centralnego.

$$e_B(j) = \begin{cases} 1 & \text{gdy } P_i(j) \rightarrow \text{wartość bazowa} \\ 0 & \end{cases}$$

$$E_i = \sum_{j=1}^K e_B(j) \quad (12)$$

$$R_S = \begin{cases} L - E_i & \text{gdy } E_i < L \\ 0 & \end{cases}$$

Detektor FAST: sprawdzany jest ciąg wartości bazowych na obwodzie otoczenia piksela centralnego.

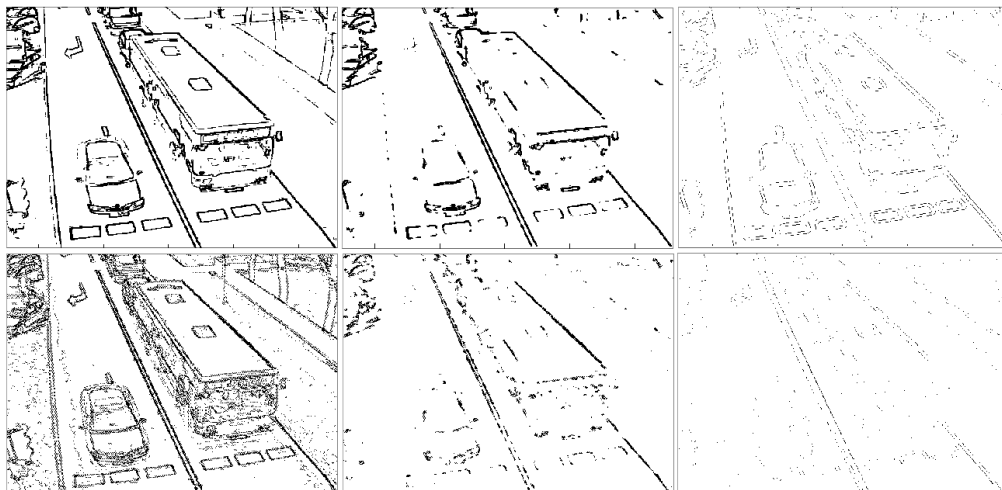
$$l_B(j) = \begin{cases} 1 & \text{gdy } P_i(j) \rightarrow \text{wartość bazowa} \\ 0 & \end{cases} \quad (13)$$

$$R_F = \sum_{j \in \text{obwód}} l_B(j)$$

Implementacja nie wymaga znajomości wartości bazowych, a jedynie ich pozycji względem analizowanych pikseli centralnych. Operacje składają się z sumowania stanów obecności, a więc z operacji realizowanych z użyciem najprostszycy środków.

3.2. Wyniki porównania detekcji cech punktowych

Uproszczone detektory zaimplementowano za pomocą pakietu matematycznego Matlab i przeprowadzono badanie detekcji cech punktowych za pomocą bazy danych obrazów ze stanowisk rejestracji ruchu drogowego. Jako materiału porównawczego użyto wyników z badań detektorów cech punktowych, wykorzystujących krawędzie obiektów dla detekcji. Krawędzie obiektów były wyznaczone z użyciem gradientu morfologicznego, wyliczanego w kwadratowym otoczeniu 3 x 3 piksele. Przykładowe wyniki ilustruje rys. 6.



Rys. 6. Pierwszy wiersz: krawędzie obiektów, cechy punktowe: SUSAN, FAST, drugi wiersz: wartości bazowe, cechy punktowe: SUSAN, FAST

Fig. 6. 1st row: object edges, feature points: SUSAN, FAST, 2nd row: base values, feature points: SUSAN, FAST

Detekcja z wykorzystaniem wartości bazowych jest mniej czuła; generowana jest mniejsza liczba cech punktowych, których lokalizacja odpowiada lokalizacji cech uzyskanych

dla krawędzi obiektów. Nadaje się ona do śledzenia obiektów w sekwencji obrazów, a wykorzystywane algorytmy cechują się niewielką złożonością obliczeniową.

4. PODSUMOWANIE

Lokalizacja obiektów jest ważnym elementem analizy scen ruchu drogowego. Konwersja obrazu cyfrowego do reprezentacji dwuwarstwowej może być stosowana jako metoda przetwarzania wstępnego w algorytmie lokalizacji obiektów. Detekcja cech punktowych z użyciem wartości bazowych dwuwarstwowej reprezentacji obrazu spełnia wymagania dla śledzenia obiektów w sekwencji obrazów. Atrakcyjną cechą zaproponowanych detektorów jest ich bardzo niska złożoność obliczeniowa, która ułatwia sprzętową realizację z użyciem programowalnych układów logicznych.

Bibliografia

1. Datka S., Sucharzewski W., Tracz M.: Inżynieria ruchu. Wydawnictwa Komunikacji i Łączności, Warszawa 1999.
2. Gonzales R.C., Woods R.E.: Digital Image Processing. Pearson Prentice Hall, New Jersey 2008.
3. Pratt W.K.: Digital Image Processing. John Wiley & Sons, Inc., Hoboken, New Jersey 2007.
4. Tadeusiewicz R., Korohoda P.: Komputerowa analiza i przetwarzanie obrazów. Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1997.
5. Salomon D.: Data compression. The Complete Reference. Springer-Verlag, New York Berlin, Heidelberg 2000.
6. Sayood K.: Kompresja danych – wprowadzenie. Wydawnictwo RM, Warszawa 2002.
7. Czapla Z.: The Two-Layer Data Model of Vehicles Digital Images, [in:] Piecha J., Węgrzyn T. (eds.): Transaction on Transport Systems Telematics and Safety. Silesian University of Technology Academic Press, Gliwice 2009, p. 82-89.
8. Harris C., Stephens M.: A combined corner and edge detector. Proc. of the 4th ALVEY Conference, 1988, p. 147–151.
9. Smith S.M., Brady J.M.: SUSAN – a new approach to low level image processing. Int. J. Comput. Vis., Vol. 23, No.1, 1997, p. 45–78.
10. Rosten E., Drummond T.: Fusing points and lines for high performance tracking, [in:] 10th IEEE International Conference on Computer Vision. Vol. 2, Beijing, China, Springer 2005, p. 1508-1515.
11. Lowe D.G.: Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, Vol. 60. No. 2, 2004, p. 91-110.

Recenzent: Prof. dr hab. inż. Romuald Szopa