

Robert ŚWITA, Zbigniew SUSZYŃSKIPOLITECHNIKA KOSZALIŃSKA, KATEDRA SYSTEMÓW MULTIMEDIALNYCH I SZTUCZNEJ INTELIGENCJI
ul. Śniadeckich 2, 75-453 Koszalin**Segmentacja sekwencji obrazów metodą korelacyjną****Mgr inż. Robert ŚWITA**

Absolwent Wydziału Elektroniki, Telekomunikacji i Informatyki Politechniki Gdańskiej. Od 2010 roku doktorant na Wydziale Elektroniki i Informatyki Politechniki Koszalińskiej w Katedrze Systemów Multimedialnych i Sztucznej Inteligencji. Zainteresowania naukowe: algorytmy tworzenia grafiki 3D, przetwarzanie dźwięku i obrazu, techniki multimedialne.



e-mail: robert.swita@wp.pl

Dr hab. inż. Zbigniew SUSZYŃSKI

Kierownik Katedry Systemów Multimedialnych i Sztucznej Inteligencji na WEiI Politechniki Koszalińskiej. Studia i obrona pracy doktorskiej – Wydział Techniki Elektronowej na Politechnice Kijowskiej. Kolokwium habilitacyjne w Instytucie Technologii Elektronowych PAN w Warszawie. Zainteresowania: termografia aktywna, tomografia termiczna, przetwarzanie sygnałów i obrazów oraz zastosowania algorytmów genetycznych i sztucznych sieci neuronowych.



e-mail: zbigniew.suszynski@tu.koszalin.pl

Streszczenie

Artykuł przedstawia nową metodę segmentacji sekwencji obrazów termicznych wyodrębniając obszary o różnych właściwościach cieplnych. Metoda oparta jest na korelacji położenia i kształtu segmentów w poszczególnych kadrach sekwencji. Segmentacja pozwala zmniejszyć liczbę analizowanych obszarów do kilku tysięcy razy, co stwarza realne możliwości praktycznego wykorzystania tomografii termicznej. Opisana metoda jest porównana z algorytmami klasteryzacji K-Means i FCM. Zaletą algorytmu korelacyjnego jest automatyczne wyznaczanie liczby segmentów wyjściowych.

Słowa kluczowe: segmentacja, obrazy termiczne, korelacja, K-means, FCM.

Segmentation of the image sequence using the correlation method**Abstract**

This paper presents a new method for segmentation of thermal image sequences. Its aim is to divide the sequence into segments with different thermal properties. The described algorithm is based on measurements of the position and shape correlation of the segments in successive frames of the sequence. It is composed of several stages. The first stage consists of segmenting consecutive frames of the sequence (Fig. 2). The second step is analysis of the similarity of each segment in each frame with respect to all other segments of all frames and synthesis of the intermediate segments (Fig. 4). The intermediate segments form the segmented output image using the depth buffer technique to resolve multiple pixel-to-segment assignments (Fig. 6). This method is a basis for the thermal analysis of solids, which results in discovering depth profiles of thermal properties for each area. The segmentation reduces the number of the analyzed areas down to a few thousand times, which creates real opportunities for practical application of thermal tomography. The new algorithm has been compared with the K-means algorithm [2], and FCM [6], which minimizes the sum of pixel value deviations from the centers of the segments they are assigned to, for all frames of the sequence (Tab. 1). The advantage of the correlation method is automatic determination of the number of output segments in the image and maintaining the constant segmentation error when increasing the number of the processed frames.

Keywords: segmentation, thermal images, correlation, K-means, FCM.

1. Wstęp

Odpowiedź temperaturowa badanego obiektu na pobudzenie energetyczne jest zależna od cech pobudzenia i parametrów cieplnych materiału. Modyfikując widmo pobudzenia i wyznaczając odpowiedź badanego obiektu w dziedzinie częstotliwości możemy przeprowadzać analizę częstotliwościową obrazów termicznych w celu wyznaczenia objętościowych parametrów cieplnych. Analiza częstotliwościowa pozwala znacznie rozszerzyć złożoność pobudzeń i obiektów w stosunku do analizy czasowej, która może być trudna ze względu na parametryczny charakter modeli termicznych. Analiza widmowa odpowiedzi temperaturowej, przy

znany widmie pobudzenia pozwala wyznaczyć profil głębokościowy właściwości cieplnych badanych struktur [1]. Bardzo istotną cechą składowych częstotliwościowych widma strumienia cieplnego wzbudzonego w obiekcie jest zależność głębokości penetracji od częstotliwości. Ta cecha wykorzystywana jest przy analizie profilu cieplnego ciał stałych. Wynika z niej zależność pola temperaturowego obiektu od głębokości dyfuzji składowych zmiennych strumienia cieplnego. Przy zmieniającym się profilu głębokościowym parametrów cieplnych powoduje to zmiany rejestrowanej temperatury, dlatego obrazy termiczne zarejestrowane w kolejnych sekwencjach czasu lub dla kolejnych składowych widma zmieniają się w sposób zależny od cech obiektu. W szczególności umożliwia to detekcję i identyfikację defektów badanej struktury, co wymaga rozwiązania problemu odwrotnego z zastosowaniem metod optymalizacyjnych. Poza prostymi jednowymiarowymi przypadkami nie można tego problemu rozwiązać metodami analitycznymi, dlatego stosuje się metody numeryczne, algorytmy genetyczne i sztuczne sieci neuronowe [2].

Obrazy termiczne, podobnie jak inne obrazy diagnostyczne, zawierają od kilkudziesięciu tysięcy do kilku milionów pikseli. Analiza sygnału temperaturowego w każdym pikselu sekwencji, zawierającej od kilkudziesięciu do kilkuset kadrów, jest bardzo kosztowna obliczeniowo. Konieczne staje się zatem znalezienie i zgrupowanie pikseli o zbliżonych charakterystykach czasowych lub częstotliwościowych, których „uśrednione” charakterystyki będą mogły posłużyć jako parametry wejściowe do algorytmu wyznaczającego właściwości cieplne obszarów ograniczonych kolejnymi segmentami. W artykule przedstawiono metodę wyznaczania takich obszarów na podstawie kilku charakterystycznych kadrów wybranych z sekwencji. Kadry zostały poddane segmentacji pikselowej, a następnie rozpoznane segmenty, na podstawie utworzonej tablicy korelacji pomiędzy wszystkimi segmentami wszystkich kadrów, posłużyły do utworzenia obrazu wynikowego, zawierającego obszary o podobnych charakterystykach a więc i głębokościowych profilach właściwości cieplnych.

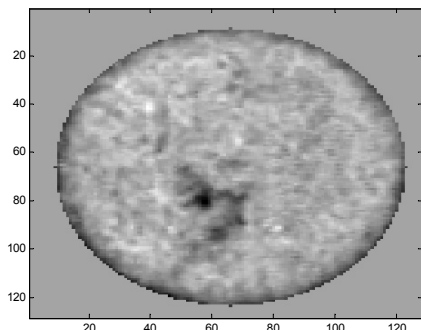
Ponieważ celem segmentacji jest wyznaczenie segmentów złożonych z pikseli o podobnych wartościach w poszczególnych kadrach, zastosowano segmentację pikselową, tworzącą w ogólności segmenty niespójne, składające się z nie połączonych ze sobą fragmentów. Zadanie segmentacji pikselowej polega na znalezieniu takich środków segmentów, dla których suma odchyżeń pikseli od przypisanych im środków jest jak najmniejsza. Problem takiej segmentacji jest problemem NP zupełnym nawet w przypadku podziału na dwa segmenty [3].

2. Metoda korelacyjna segmentacji sekwencji obrazów

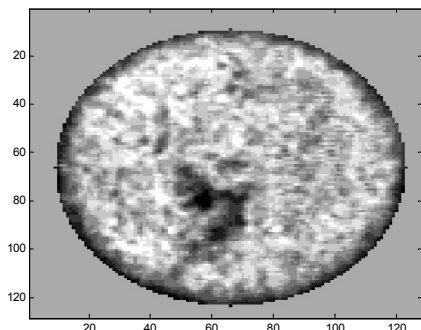
Opisywana metoda identyfikuje obszary uwzględniając charakter zmian kontrastu w dziedzinie czasu lub częstotliwości. Pozwala to na wydzielenie i rozróżnienie niejednorodności cieplnych

badanych struktur z punktu widzenia ich właściwości termicznych. Metoda składa się z kilku etapów.

W pierwszym, dokonuje się klasycznej segmentacji każdego kadru sekwencji (rys. 1 i rys. 2). Ze względu na dynamiczne zmiany kontrastów poszczególnych obszarów (dotyczące nawet znaku kontrastu), kolejne obrazy segmentacyjne ulegają istotnym zmianom.



Rys. 1. Przykładowy kadr źródłowy sekwencji (odwrócona paleta szarości)
Fig. 1. Exemplary source frame of the sequence (inverted gray palette)



Rys. 2. Kadr po segmentacji pikselowej
Fig. 2. Frame after point segmentation

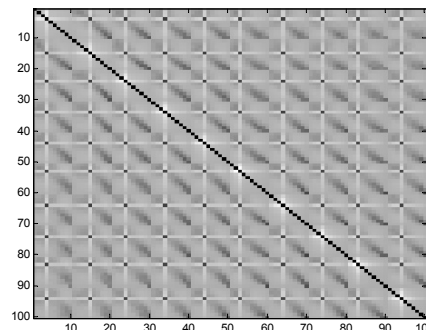
Drugi etap segmentacji korelacyjnej polega na syntezie tych obrazów. Rozpoczyna się ona analizą podobieństwa każdego segmentu, w każdym kadrze, w stosunku do wszystkich innych segmentów wszystkich kadrów (łącznie z kadrem aktualnego segmentu). Następnie tworzona jest kwadratowa i symetryczna macierz korelacji (rys. 3) o wymiarach $(K \cdot \hat{S}) \times (K \cdot \hat{S})$ (K - liczba kadrów, \hat{S} - liczba segmentów w kadrze) ze współczynnikami korelacji Pearsona dla poszczególnych par segmentów. Współczynnik korelacji Pearsona segmentów X i Y jest kowariancją odchyłeń wartości ich pikseli od środków segmentów, znormalizowaną przez produkt ich odchyłeń standardowych:

$$\text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y} = \frac{\sum_{i=1}^P (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^P (x_i - \bar{x})^2 \sum_{i=1}^P (y_i - \bar{y})^2}} \quad (1)$$

Współczynniki w macierzy korelacji określają stopień podobieństwa pomiędzy segmentami X i Y .

Macierz korelacji jest symetryczna, $\text{corr}(x, y) = \text{corr}(y, x)$, dlatego wyliczenia wystarczy przeprowadzić tylko dla jej połowy. Nierówność Cauchy'ego-Schwarza zapewnia, że współczynnik korelacji jest, co do wartości bezwzględnej, mniejszy od jedności. Z nierówności wynika również, że jeśli $x < y$ (lub $y < x$), korelacja jest na pewno większa od zera. Duże wartości dodatnie współczynnika oznaczają duże podobieństwo kształtu segmentów. Jeśli segmenty nie mają części wspólnej, korelacja jest zawsze ujemna i wskazuje na podobieństwo badanego segmentu do dopełnienia (negacji) drugiego segmentu. Współczynnik korelacji Pearsona

jest w przypadku segmentów binarnych taki sam dla ich dopełnień. Dla korelacji badanego segmentu z dopełnieniem drugiego, przyjmuje wartość ze znakiem przeciwnym. Jeśli segmenty nie są podobne, korelacja jest bliska zera. Twierdzenie odwrotne nie jest jednak prawdziwe. Zerowy współczynnik korelacji nie musi oznaczać, że segmenty nie są do siebie podobne, bowiem efekt korelacji dodatniej i ujemnej może się zniósć.

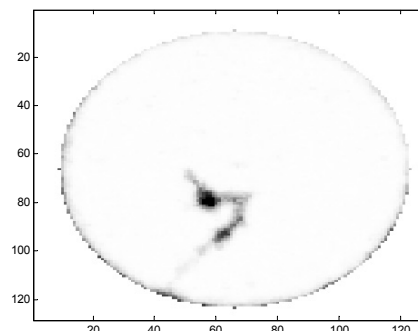


Rys. 3. Macierz korelacji dla liczby kadrów $K = 10$ i segmentów $\hat{S} = 10$
Fig. 3. Correlation matrix for frame count $K = 10$ and segment count $\hat{S} = 10$

Dla każdego i -tego segmentu tworzony jest i -ty obraz pośredni (rys. 4) będący sumą wszystkich segmentów ważonych (piksele w obrębie każdego segmentu, przed ważeniem, mają wartość jeden). Waga $w_{i,j}$ każdego j -tego segmentu jest współczynnikiem korelacji tego segmentu z segmentem i -tym, odniesiony do liczby pikseli segmentu j .

$$\text{seg}_i = \sum_{j=1}^{K \cdot \hat{S}} w_{ij} \text{seg}_j = \sum_{j=1}^{K \cdot \hat{S}} \frac{\text{corr}(\text{seg}_j, \text{seg}_i)}{\|\text{seg}_j\|} \text{seg}_j \quad (2)$$

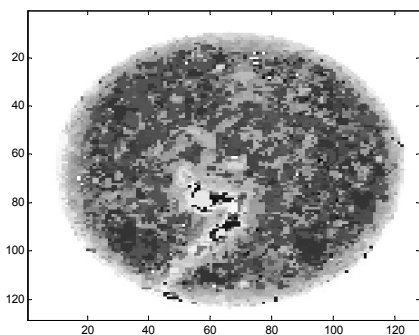
W ten sposób uprzywilejowane zostały małe segmenty, które w przeciwnym razie mogłyby zostać wyeliminowane przez większe podczas syntezy końcowego obrazu korelacyjnego. Większe segmenty korelują z innymi częściej niż segmenty mniejsze, dlatego waga została skompensowana ze względu na wielkość segmentów. Warto zwrócić uwagę, że obrazy pośrednie, podobnie jak segmenty z kadrów źródłowych, nie muszą być obszarami spójnymi, gdyż powstały w wyniku segmentacji pikselowej.



Rys. 4. Przykładowy segment pośredni dla przypadku $K = \hat{S} = 10$
Fig. 4. Exemplary intermediate segment for $K = \hat{S} = 10$

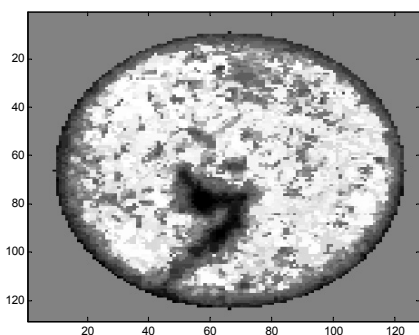
Segmenty pośrednie służą do wyznaczenia końcowego obrazu segmentacyjnego. Powstaje on w wyniku dodawania poszczególnych segmentów. Ponieważ nie są to rozłączne względem siebie zbiory pikseli, należało rozwiązać problem jednoznacznej przynależności tych pikseli, które mogą znajdować się w dwóch lub większej liczbie segmentów pośrednich. Wykorzystana została tu technika bufora głębokości obrazu i w takich przypadkach o przynależności piksela decydowała jego wartość w poszczególnych segmentach – piksel przydzielano do segmentu, w którym miał on największą wartość.

Maksymalna liczba segmentów w tak powstałym obrazie wynikowym, wynosi $K \cdot S$. Niektóre segmenty zostały jednak wyeliminowane przez działanie bufora głębokości, dlatego należało zmienić indeksację obszarów w obrazie wynikowym (rys. 5).



Rys. 5. Obraz wyjściowy po zmianie indeksacji
Fig. 5. Reindexed output image

Segmenty w obrazie wyjściowym można również indeksować zgodnie z ich wartością średnią w obrazach źródłowych. Dzięki temu obraz końcowy jest podobny do obrazów źródłowych, a analiza poprawności segmentacji jest łatwiejsza (rys. 6).



Rys. 6. Obraz wynikowy z indeksami segmentów proporcjonalnymi do ich wartości średniej
Fig. 6. Output image with segment indexes proportional to segment mean values

3. Metoda K-Means i FCM

Segmentacja korelacyjna została porównana z algorytmem klasteryzacji K-Means (MacQuinn, 1967 [4]), dla którego parametrem wejściowym jest sekwencja obrazów i macierz $K \times S$ środków segmentów (K - liczba kadrów, S - liczba segmentów wyjściowych). Wartości środków segmentów rozłożono początkowo równomiernie dla poszczególnych kadrów. Aby wyeliminować możliwość przydziału pikseli o zbliżonych wartościach do dwóch różnych klastrów, po przydzieleniu pikseli do najbliższych segmentów, środki klastrów są przeliczane do wartości średniej ze wszystkich pikseli należących do danego klastra i ponownie przeprowadzana jest procedura przydziału. Iteracja przerywana jest wówczas, gdy piksele przestaną zmieniać swoje przyporządkowanie do określonych klastrów. Metoda K-Means dąży w ten sposób, do ustalenia normalnego rozkładu pikseli w poszczególnych segmentach na histogramie wartości pikseli.

Algorytm K-Means jest zawsze zbieżny, ale o szybkości zbieżności i dokładności lokalnego rozwiązania decyduje wybór początkowy środków klastrów (Pena et al., 1999 [5]). Standardowo przyjmuje się wartości losowe przy jednorodnym rozkładzie prawdopodobieństwa wszystkich wartości w zbiorze (inicjalizacja Forgy'ego [6]). D.Arthur i S.Vassilvitskii w [7] przedstawiają algorytm K-Means++, w którym środki początkowe wybierane są ze zbioru próbek danych wejściowych przy założeniu prawdopodobieństwa próbki proporcjonalnym do odległości od już wyznaczonych centroidów. Obie metody zakładają jednak wykorzysta-

nie statystyczne algorytmu, który dostarcza za każdym razem nieco innego rozwiązania, i oczywiście, inny jest również wówczas błąd końcowej aproksymacji charakterystyk. Z tego względu zastosowany został wektor początkowych środków klastrów wyznaczonych równomiernie w zakresie zmienności wartości pikseli w kadrze.

Algorytm segmentacji FCM (*Fuzzy C-Means*) [8] korzysta z logiki rozmytej, zakładając przynależność pikseli do danego segmentu zgodnie z określoną funkcją przynależności μ (*soft membership*). Po każdej iteracji, środki klastrów dla każdego kadru modyfikowane są przez wyznaczenie środka ciężkości segmentów, zgodnie z równaniem (3):

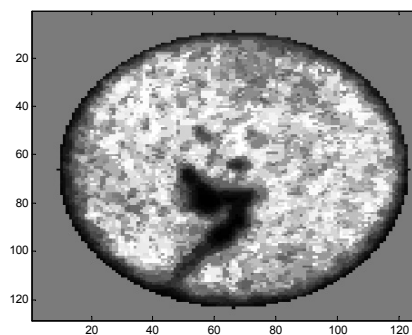
$$\bar{x}_j = \frac{\sum_{i=1}^P \mu(x_i, \bar{x}_j) x_i}{\sum_{i=1}^P \mu(x_i, \bar{x}_j)} \quad (3)$$

Funkcja przynależności jest określana dla każdej pary i -tego pikseli i j -tego segmentu jako:

$$\mu(x_i, \bar{x}_j) = \frac{(x_i - \bar{x}_j)^{-2/(r-1)}}{\sum_{j=1}^S (x_i - \bar{x}_j)^{-2/(r-1)}} \quad (4)$$

gdzie: r - stopień „rozmycia” przynależności pikseli, $r \geq 1$, S - liczba segmentów

Algorytm K-Means jest zatem szczególnym przypadkiem metody FCM, dla którego funkcja przynależności przyjmuje wartości binarne, równe jedności dla pikseli należących wyłącznie do określonego segmentu (*hard membership*).



Rys. 7. Obraz wynikowy uzyskany po segmentacji FCM
Fig. 7. Output image obtained with the FCM method

4. Średni błąd bezwzględny segmentacji

Najważniejszym parametrem jakości segmentacji pikselowej jest średni błąd bezwzględny segmentacji, który został określony jako suma odchyień wartości wszystkich P pikseli od przyporządkowanych im środków segmentów dla wszystkich K kadrów:

$$\varepsilon = \frac{1}{KP} \sum_{k=1}^K \sum_{i=1}^P \min_{j \in \{1, S\}} |x_{ki} - \bar{x}_{kj}| \quad (5)$$

Segmentacja pikselowa, nie uwzględniająca wartości pikseli sąsiednich w stosunku do przetwarzanego pikseli, będzie starała się minimalizować funkcję opisującą błąd bezwzględny z równania (5). Jest to również funkcja kosztu dla algorytmu K-Means, w którym dane zostały sformatowane do dwuwymiarowej tablicy wartości kolejnych pikseli w funkcji numeru kadru, które możemy nazwać charakterystykami pikseli. Proszę zauważyć, że w funkcji celu, którą algorytm K-Means minimalizuje, odchylenia wartości pikseli od środków przypisanego mu segmentu dla wszystkich kadrów są sumowane.

Błąd bezwzględny można również przedstawić jako sumę błędów bezwzględnych przypadających na poszczególne segmenty i zobrazować na charakterystyce błędu bezwzględnego w funkcji numeru segmentu:

$$\varepsilon = \sum_{j=1}^S \frac{1}{KP} \sum_{k=1}^K \sum_{i=1}^{P_{seg(j)}} |x_{kji} - \bar{x}_{kj}| = \sum_{j=1}^S \varepsilon_{seg(j)} \quad (6)$$

5. Wyniki i wnioski

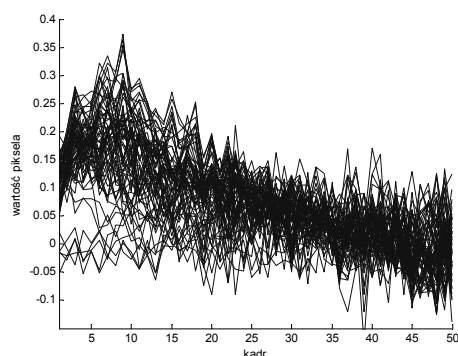
W tab.1 zostały porównane średnie błędy bezwzględne segmentacji sekwencji termicznych obrazów fazowych dla metody korelacyjnej, K-Means i FCM. Przyjęta została liczba segmentów wyjściowych S uzyskana automatycznie z metody korelacyjnej. Podane zostały również czasy obliczeń. Liczba iteracji dla metody FCM wynosiła 20, a parametr rozmycia $r = 1.05$. Dla metody K-Means nie ograniczono liczby iteracji – wykonywana była aż do osiągnięcia zbieżności.

Tab. 1. Średnie błędy bezwzględne segmentacji metodą korelacji obrazów segmentów, K-Means i FCM

Tab. 1. Absolute mean segmentation errors for correlation, K-Means and FCM algorithms

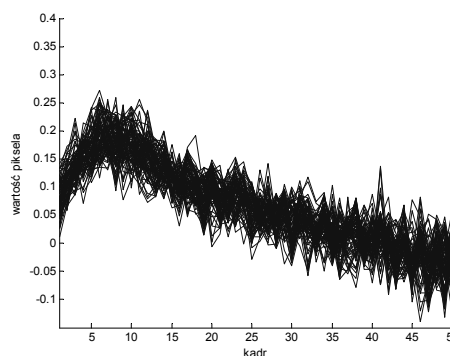
Liczba kadrów x segm. $K \times S$	Liczba segm. wyj. S	Metoda korelacyjna		Metoda K-Means		Metoda FCM	
		ε [°]	t [s]	ε [°]	t [s]	ε [°]	t [s]
10x10	88	0,7464	11,3	0,6315	9,5	0,6061	23,6
25x10	176	0,8470	50,4	0,7325	24,3	0,7129	79
50x10	241	0,9856	180	0,8713	55,7	0,8565	169
10x20	165	0,7184	38,7	0,5829	11,6	0,5625	44
25x20	330	0,8373	190	0,6862	39	0,6667	146
50x20	501	0,9906	700	0,7986	110	0,7800	340

Segmentacja grupuje piksele o podobnych charakterystykach ich wartości w funkcji numeru kadru. Na rys. 8 i rys. 9 przedstawione zostały charakterystyki wszystkich pikseli wybranego segmentu (defektu widocznego na kadrach) po segmentacji metodą korelacji i K-Means dla przypadku $K=50$ i $S = 241$.



Rys. 8. Charakterystyki pikseli wybranego segmentu w segmentacji korelacyjnej
Fig. 8. Characteristics of chosen segment pixels after segmentation with the correlation method

Segmentacja obrazów termicznych w bardzo dużym stopniu redukuje liczbę koniecznych do przetwarzania danych bez utraty znaczących informacji o właściwościach badanego obiektu. Zaprezentowana metoda w porównaniu z algorytmem K-Means daje w wyniku większe, ale porównywalne błędy segmentacji. Czas wykonywania algorytmów K-Means i FCM jest również krótszy (użyte zostały wysoko zoptymalizowane algorytmy MATLAB'a).



Rys. 9. Charakterystyki pikseli wybranego segmentu po K-Means
Fig. 9. Characteristics of chosen segment pixels after K-Means

K-Means posiada jednak wady, do których zaliczyć należy brak automatycznego wyznaczania liczby segmentów, zależność końcowej segmentacji od macierzy $K \times S$ początkowych środków segmentów oraz słabą zbieżność algorytmu. Metoda korelacyjna jest heurystyką, która wykorzystuje ważoną funkcję przynależności segmentów do segmentu pośredniego, natomiast algorytm klasteryzacji zakłada, że piksel może należeć wyłącznie do jednego segmentu. Jeśli funkcja przynależności nie jest binarna, algorytm K-Means modyfikuje się do takich algorytmów jak FCM lub KHM (*K-Harmonic Means*), które są szybciej zbieżne i prowadzą do rozwiązań o mniejszym błędzie segmentacji. Metody te jednak znacznie ustępują efektywności algorytmowi K-Means.

Dla pojedynczych obrazów łatwo wskazać poprawną liczbę segmentów wyjściowych. Dużo trudniej oszacować tę liczbę dla sekwencji obrazów. Zaletą metody korelacyjnej jest właśnie automatyczne wyznaczanie liczby segmentów w obrazie wyjściowym i utrzymywanie stałego błędu segmentacji przy zwiększaniu liczby kadrów wejściowych przetwarzanej sekwencji.

Praca współfinansowana przez Ministerstwo Nauki i Szkolnictwa Wyższego, grant NN515247037.

6. Literatura

- [1] Suszyński Z.: Termofalowe metody badania materiałów i przyrządów elektronicznych, Monografia habilitacyjna, Wydawnictwo Uczelniane Politechniki Koszalińskiej, 2001.
- [2] Arsoba R., Suszyński Z.: Application of photoacoustic method and evolutionary algorithm for determination of thermal properties of layered structure, *Journal de Physique IV - Proceedings*, Vol. 117, pp. 1-6, October 2004.
- [3] Drineas P., Frieze A., Kannan R., Vempala S., Vinay V.: Clustering large graphs via the singular value decomposition, *Mach. Learn.*, 56(1-3):9-33, 2004.
- [4] MacQueen J.B.: Some methods for classification and analysis of multivariate observations, In *proceedings of the fifth Berkeley Symposium on Mathematical Statistics and Probability*, L. M. L. Cam and J. Neyman, eds., vol. 1, pp. 281-297, UC Press 1967.
- [5] Pena J. M., Lozano J. A., Larranaga P.: An empirical comparison of four initialization methods for the K-Means algorithm, *Pattern Recognition Letters* 20 (10), pp. 1027-1040, 1999.
- [6] Anderberg M. R.: *Cluster analysis for applications*, Academic Press, New York, 1973.
- [7] Arthur D., Vassilvitskii S.: *K-Means++: the advantages of careful seeding*, Technical Report 2006-13, Stanford InfoLab, 2006.
- [8] Hamerly G., Elkan Ch.: Alternatives to the k-means algorithm that find better clusterings, In *proceedings of the ACM conference on information and knowledge management (CIKM)*, pp. 600-607, 2002.

otrzymano / received: 20.11.2012

przyjęto do druku / accepted: 03.06.2013

artykuł recenzowany / revised paper