

Paweł Gardziński  
Krzysztof Kowalak  
Łukasz Kamiński  
Sławomir Maćkowiak  
Katedra Telekomunikacji Multimedialnej i Mikroelektroniki  
ul. Polanka 3, 61-131 Poznań,  
tel. + (48) 6653900, fax: + (48) 6653899,  
email: {pgardzinski, kkowalak, lkaminski, smack}@multimedia.edu.pl  
web: www.multimedia.edu.pl



Kraków, 27-29 czerwca 2016

## SEPARACJA SYLWETEK OSÓB PRZY POMOCY DYSKRETNEGO RÓWNANIA POISSONA ORAZ ROZSZERZONEGO ALGORYTMU SEGMENTACJI WODODZIAŁOWEJ

### HUMAN SILHOUETTE SEGMENTATION USING DISCRETE POISSON EQUATION AND EXTENDED WATERSHED ALGORITHM

**Streszczenie:** W niniejszym artykule przedstawiono autorską metodę separacji sylwetek osób dla zastosowań w dozorcze wizyjnym. Zaproponowane rozwiązanie wykorzystuje dyskretne równanie Poissona oraz kombinację zmodyfikowanego algorytmu segmentacji wododziałowej z algorytmem rozrostu regionu. Badania zostały przeprowadzone na powszechnie dostępnej bazie testowej PETS 2006. Otrzymane wyniki potwierdzają skuteczność przedstawionej metody.

**Abstract:** In this paper a novel approach on human silhouette segmentation for surveillance systems was proposed. The described solution uses discrete Poisson equation and a combination of extended watershed algorithm with Region Growing algorithm. Experiments were performed on a commonly known database PETS 2006 and the results show that the proposed solution achieves high precision and accuracy.

**Słowa kluczowe:** algorytm wododziałowy, dozór wizyjny, dyskretne równanie Poissona, segmentacja sylwetek

**Keywords:** discrete Poisson equation, silhouette segmentation, surveillance systems, watershed algorithm

## 1. WSTĘP

Dozór wizyjny jest ciągle rozwijającą się branżą, zaś zapotrzebowanie na inteligentne systemy dozоровe stale rośnie. Zagadnienia takie jak wykrywanie wtargnięcia na teren chroniony, czy pomiar natężenia ruchu nie są już problemem, a producenci kamer oferują powyższe rozwiązania już zaimplementowane w ich urządzeniach. Wciąż jednak rozwijane są metody rozpoznawania lub detekcji bardziej skomplikowanych zdarzeń. Przykładem najpopularniejszych prac dotyczących rozpoznawania zdarzeń są rozwiązania oparte o analizę sylwetek obiektów [1] lub konturów [5]. Innego rodzaju rozwiązania wykorzystują ukryty model Markova [2]

czy też stosując deskrytory CSSD (ang. Curvature scale space descriptor) lub ARTD (ang. Angular radial transform descriptor) objęte standardem MPEG-7 [8,7,3,11]. Wszystkie te rozwiązania mają jednak słaby punkt, którym jest problem z przysłaniającymi się postaciami. Jeżeli przykładowo dwie obserwowane osoby przemieszczą się tak, że ich wyznaczone sylwetki (lub kontury) się połączą, to skuteczność detekcji zdarzeń na takich sylwetkach spada do zera lub konieczne jest zastosowanie metod predykcji. Powyższy problem jest częściowo rozwiązywany poprzez wykorzystanie rozwiązań opartych na histogramach zorientowanych gradientów (ang. Histogram of Oriented Gradients) [2, 12]. Jednak te rozwiązania charakteryzują się znacznie niższą skutecznością oraz wymagają wcześniejszego uczenia. Ponadto, dla osiągnięcia wyższej skuteczności, konieczne są zbiory uczące liczone w dziesiątkach tysięcy obrazów, co wciąż nie zapewnia pokrycia wszystkich przypadków.

W związku z powyższym, niniejszy artykuł zawiera propozycję metody pozwalającej na rozwiązanie problemu przysłaniających się postaci poprzez zastosowanie separacji obiektów na podstawie dyskretnego równania Poissona wyliczonego dla sylwetek połączonych. Rozwiązanie to, jako dane wejściowe, przyjmuje obrazy binarne sylwetek, dla których wyznaczane są deskryptory obliczone z dyskretnego równania Poissona [10]. Następnie, wykorzystując autorski, zmodyfikowany algorytm wododziałowy, dokonywana jest estymacja liczby osobnych obiektów oraz wyznaczane są granice między nimi.

## 2. OPIS ALGORYTMU

Rozwiązanie przedstawione w niniejszym artykule składa się z dwóch głównych części. Część pierwsza polega na wyznaczaniu deskryptora sylwetki przy pomocy dyskretnego równania Poissona (dla uproszczenia notacji nazwijmy wynik deskryptorem Poissona, w skrócie DP). Następnie w części drugiej dokonywana jest

segmentacja wododziałowa, podczas której rozpoznawane są maksima lokalne oznaczające potencjalne obiekty oraz granice między nimi.

## 2.1. Deskryptor sylwetki oraz dyskretne równanie Poissona

Rozważmy sylwetkę  $S$  rozpiętą na siatce punktów odległych od siebie o  $d$ , otoczoną konturem zamkniętym  $\partial S$ . Następnie w każdym punkcie  $(x, y) \in S$  umieszczamy zbiór cząsteczek i wprawiamy je w ruch losowy (tj. w każdym kolejnym kroku cząsteczka z równym prawdopodobieństwem może przemieścić się do jednego z czterech sąsiadów w rozważanej siatce). Cząsteczki poruszają się ruchem losowym, póki nie trafią na kontur  $\partial S$ . Mierząc średnią odległość  $F(x, y)$  w ruchu losowym danego punktu od konturu stosujemy przybliżoną formę równania Poissona, która wyraża się wzorem:

$$F(x, y) = d + \frac{1}{4}(F(x + d, y) + F(x - d, y) + F(x, y + d) + F(x, y - d)) \quad (1)$$

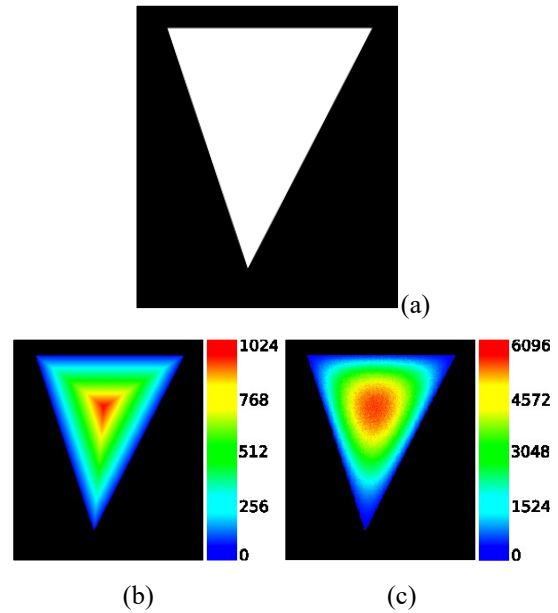
Powyższe równanie może być wyznaczone rekurencyjnie w następujący sposób: na granicy  $S$  tj. dla  $(x, y) \in \partial S$  wartość  $F(x, y) = 0$ , zaś dla pozostałych punktów  $(x, y)$  należących do  $S$  wartość  $F(x, y)$  jest równa średniej arytmetycznej bezpośrednich sąsiadów plus stała (wyrażająca odległość do bezpośredniego sąsiada). Dla uproszczenia obliczeń przyjmujemy  $d = 1$ , co odpowiada jednostkowej odległości między sąsiednimi punktami obrazu.

Transformata odległościowa jest deskryptorem, który bardzo przypomina DP. Jednak różnica polega na tym, że każdy punkt  $(x, y) \in S$  zawiera jedynie informację na temat najbliższego punktu konturu  $\partial S$ . Kolejne wartości deskryptora wyznaczonego z dyskretnego równania Poissona reprezentują coraz bardziej wygładzony kontur. W zastosowaniu w dozorze wizyjnym, kończyny obserwowanych osób są reprezentowane relatywnie niskimi wartościami w porównaniu do tułowia, co pozwala na efektywne wyróżnienie kilku postaci w złączonym konturze. Niniejsze rozwiązanie zostało wybrane właśnie ze względu na tę cechę. Porównanie deskryptora DP z wynikiem transformacji odległościowej został przedstawiony na rys. 1.

Tak przygotowany deskryptor DP, jest wykorzystywany do wyszukania maksimów lokalnych, służących za ziarna dla algorytmu rozrostu regionu, oraz ustalenia granic między złączonymi obiektami.

## 2.2. Algorytm segmentacji wododziałowej

Drugą częścią prezentowanego rozwiązania jest wyznaczenie liczby oddzielnych obiektów oraz ich separacja z połączonej sylwetki (obrazu wejściowego). Etap ten jest kombinacją dwóch ogólnie znanych metod seg-



Rys. 1. Porównanie wyniku transformacji odległościowej (b) z deskrytorem wyznaczonym z dyskretnego równania Poissona (c) dla obrazu wejściowego (a).

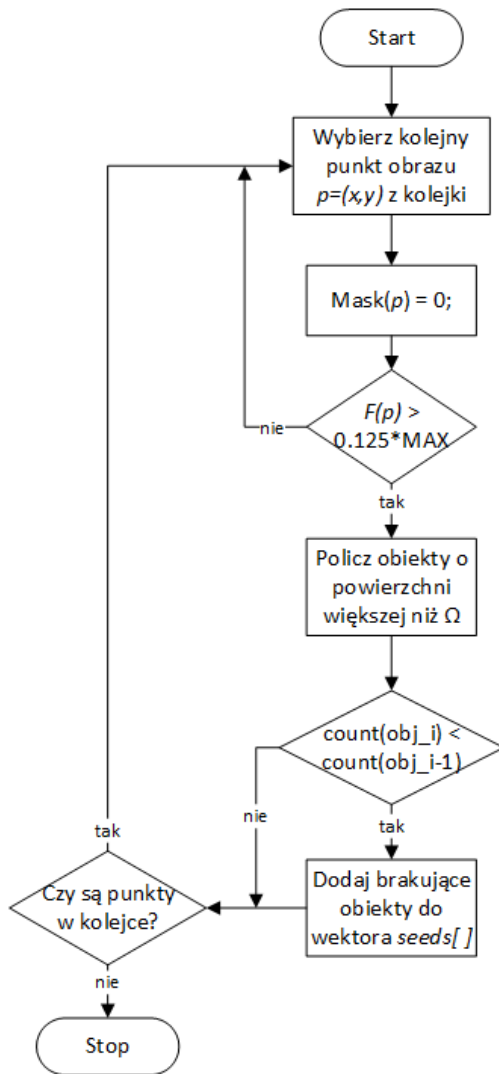
mentacji wraz z autorskimi modyfikacjami służącymi do wyszukania ziaren oraz granic.

Pierwszym krokiem w tej części algorytmu jest posortowanie punktów  $p = (x, y)$  obrazu wejściowego (czyli deskryptora DP wyznaczonego w punkcie 2.1) w kolejności rosnącej względem wartości  $F(x, y)$  lub inaczej  $F(p)$ . Po czym można przystąpić do zmodyfikowanego algorytmu segmentacji wododziałowej, którego schemat został przedstawiony na rys. 2.

Oznaczenia na diagramie są następujące:

- $Mask(p)$  maska wykorzystywana w trakcie działania algorytmu segmentacji wododziałowej,
- $MAX$  jest największą wartością deskryptora DP jaka wystąpiła (przykładowo wartość 6096 na rys. 1),
- $\Omega$  jest progiem służącym do odfiltrowania obszarów o zbyt małej powierzchni i zależy od rozmiarów obrazu wejściowego,
- $count(obj_i), count(obj_{i-1})$  są odpowiednio liczbami obiektów w danym kroku przetwarzania oraz kroku poprzednim,
- $seeds[.]$  jest wektorem wyjściowym zawierającym pozycje obiektów.

W wyniku działania powyższego algorytmu uzyskujemy zbiór punktów maksimów lokalnych deskryptora DP. Tak wyznaczony zbiór punktów służy jako zbiór ziaren (ang. *seed*) do ostatniego kroku postępowania drugiej części zaproponowanego rozwiązania.

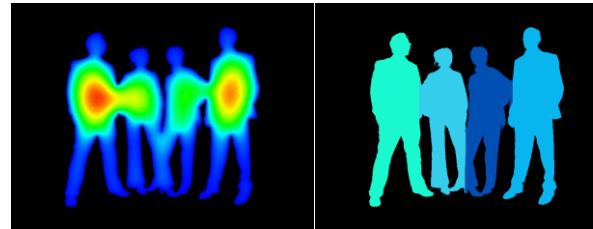


Rys. 2. Schemat blokowy algorytmu znajdowania ziarna poprzez segmentację wododziałową.

Drugim krokiem tej części algorytmu jest zastosowanie algorytmu rozrostu regionu (ang. *Region Growing*) w celu wyznaczenia oddzielnych obiektów z sylwetki połączonej. Dokładny opis tego algorytmu można znaleźć w [4, 6]. Obszarem rozrostu jest sylwetka połączona, zaś jako ziarna przyjmowany jest zbiór punktów *seeds*. Przykładowy wynik końcowy działania algorytmu został przedstawiony na rys.3.

### 3. WYNIKI BADAŃ EKSPERYMENTALNYCH

Skuteczność działania zaprezentowanego algorytmu separacji sylwetek osób została sprawdzona w dwóch etapach. W pierwszym etapie wykonane zostały testy oraz optymalizacje parametrów algorytmu. Testy te były przeprowadzone na zróżnicowanym zbiorze obrazów testowych, m.in. *people* (przedstawiony na rys.3). Druga część badań polegała na sprawdzeniu skuteczności na zbiorze obrazów testowych pochodzących z bazy PETS 2006 [9]. Skuteczność separacji sylwetek była porównana względem ręcznie oznaczonych obiektów, czyli tak zwanego ground truth.



(b)

(c)

Rys. 3. Przykładowy rezultat działania zaproponowanego rozwiązania separacji sylwetek. Obraz wejściowy (a), wyznaczony deskryptor DP (b) oraz wynik rozdzielania sylwetki (c).

Dzięki zastosowaniu ground truth możliwa była ewaluacja skuteczności działania algorytmu przy pomocy dwóch popularnych metryk *precision* oraz *recall*, które są zdefiniowane następująco:

$$precision = \frac{TP}{TP+FP}, \quad (2)$$

$$recall = \frac{TP}{TP+FN}, \quad (3)$$

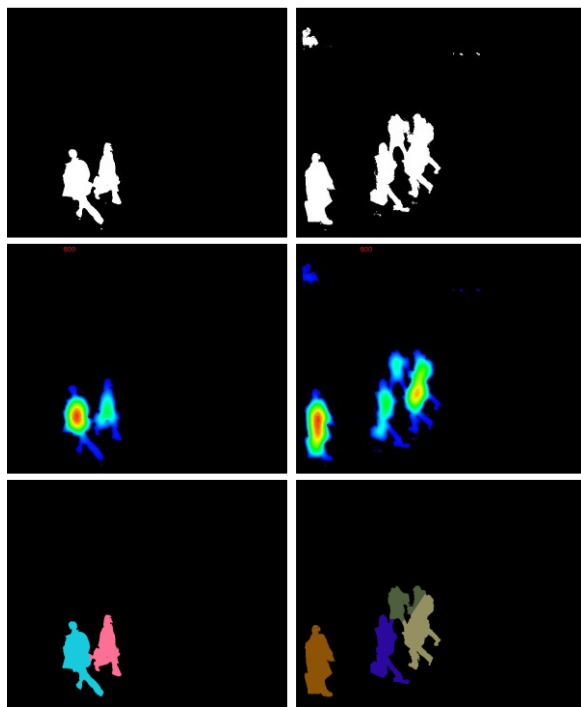
gdzie: *TP* oznaczana zbiór prawdziwe pozytywny (ang. True Positive), *FP* jest zbiorem fałszywie pozytywnym (ang. False Positive), natomiast *FN* jest zbiorem fałszywie negatywnym (ang. False Negative). W wyniku przeprowadzonych badań zaprezentowane rozwiązanie osiągnęło skuteczność *precision* = 95% oraz *recall* = 90%. Przykładowy wynik działania algorytmu podczas testów na zbiorze obrazów PETS 2006 został przedstawiony na rys. 4.

Skuteczność przedstawionej metody jest bardzo wysoka, zaś wykorzystanie deskryptora wyznaczonego z dyskretnego równania Poissona pozostawia wciąż wiele możliwości dodatkowego zastosowania oraz dalszego rozwoju.

### 4. PODSUMOWANIE

Przedstawiona w niniejszym artykule metoda separacji sylwetek osób jest autorskim rozwiązaniem problemu przysłaniających się postaci występujących w przetwarzanych sekwencjach dozoru wizyjnego. Jest to powszechne zjawisko, które często obniża skuteczność działania algorytmów opartych na analizie sylwetek lub konturów obserwowanych obiektów.

Przetwarzanie w przedstawionej metodzie zostało podzielone na dwa etapy. Etap pierwszy polega na wyznaczeniu deskryptora sylwetki z dyskretnego równania Poissona. W etapie drugim wykonywane są jednocześnie dwa postępowania algorytmiczne: zmodyfikowana seg-



Rys. 4. Przykładowe wyniki podziału dla wybranych obrazów ze zbioru PETS 2006.

mentacja wododziałowa oraz rozrost regionu. Według wiedzy autorów jest to pierwsze wykorzystanie deskryptora DP w celu separacji sylwetek. Zaprezentowany deskryptor może być z powodzeniem wykorzystany do innych zastosowań w systemach dozoru wizyjnego, jak to zostało zaprezentowane w [1,10].

Przetawiona metoda ma stanowić podstawę do dalszego przetwarzania m.in. w celu analizy zdarzeń oraz prewencji w dziedzinie dozoru wizyjnego. Zaś wyniki badań eksperymentalnych pozwalają sądzić, że zastosowanie zaproponowanego algorytmu pozwoli znacząco zwiększyć skuteczność działania takich systemów.

#### PODZIĘKOWANIA

Praca finansowana ze środków przyznanych przez Ministerstwo Nauki i Szkolnictwa Wyższego na działalność statutową w ramach zadania pn. „Teoria i algorytmy wielowymiarowego przetwarzania sygnałów”.

#### LITERATURA

- [1] L. Gorelick, M. Blank, E. Shechtman, M. Irani, and R. Basri, “Actions as space-time shapes,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 12, pp. 2247–2253, Dec. 2007
- [2] Y. Huang, H. Yang, and P. Huang, “Action recognition using hog feature in different resolution video sequences,” in *2012 International Conference on Computer Distributed Control and Intelligent Environmental Monitoring (CDCIEM)*, March 2012, pp. 85–88.

- [3] K. Mokhtarian and M. Bober, “Curvature Scale Space Representation: Theory, Applications and MPEG-7 Standardisation”, Kluwer Academic (now Springer), 2003.
- [4] W. K. Pratt, “*Digital Image Processing 4th Edition*”, John Wiley & Sons, Inc., Los Altos, California, 2007
- [5] S. Maćkowiak, P. Gardziński, Ł. Kamiński, K. Kowalak. „Human Activity Recognition in Multiview Video”, *11th IEEE International Conference on Advanced Video and Signal-Based Surveillance*, South Korea, 2014, pp. 148-153.
- [6] M. Petrou and P. Bosdogianni, *Image Processing the Fundamentals*, Wiley, UK, 2004.
- [7] B.S. Manjunath, P. Salembier, and T. Sikora, “Introduction to MPEG-7: Multimedia Content Description Interface” eds. Wiley, 2002
- [8] M. Bober, “MPEG-7 Visual Shape Descriptors,” *IEEE Trans. Circuits and Systems for Video Technology*, vol. 1, no. 6, 2001.
- [9] “PETS 2006 Benchmark Dataset” <http://www.cvg.reading.ac.uk/PETS2006/data.html>
- [10] L. Gorelick, M. Galun, E. Sharon, A. Brandt, and R. Basri, “Shape Representation and Classification Using the Poisson Equation,” *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, Dec. 2006.
- [11] “The MPEG-7 XM software repository,” [http://www.lis.ei.tum.de/research/bv/topics/mmdb/e\\_mpeg7.html](http://www.lis.ei.tum.de/research/bv/topics/mmdb/e_mpeg7.html), 2005.
- [12] H.Qian, J.Zhou, Y.Yuan, „Visual-based fall detection using histogram of oriented gradients of Poisson distance image” in *Chinese Automation Congress (CAC)*, 2015, pp. 657-662