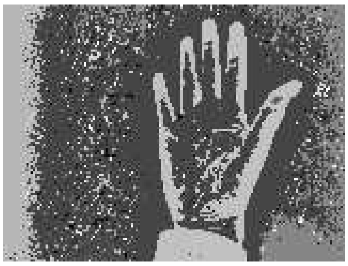
(e) the sign distribution of the relative potential in frame 5

(f) the sign distribution of the relative potential in frame 6

Fig. 7 The sign distribution of the relative potential values for the “clenching hand” video sequence

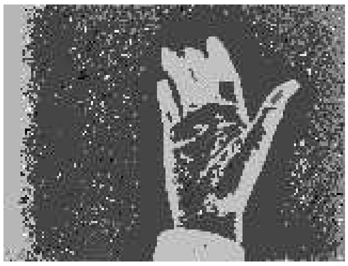


(a) the segmentation result on the plane of frame 1

(b) the segmentation result on the plane of frame 2

(c) the segmentation result on the plane of frame 3

(d) the segmentation result on the plane of frame 4



(e) the segmentation result on the plane of frame 5

(f) the segmentation result on the plane of frame 6

Fig. 8 The segmentation results for the “clenching hand” video sequence

Fig. 9 The sequence of clenching hand area segmented from the video

Fig. 10 is a sequence of the TV broadcaster, which has the size 176×144 for each frame. In the experiment, the sign distribution of the relative potential in the 3D space is recorded. The result is shown in a 2D form in Fig. 11, where the sign distribution in each frame is shown separately. According to Fig. 11, the main area of the broadcaster can be segmented as several connected regions including the face. Moreover, the tracking of the face is also possible based on the segmentation result because the result includes connected 3D areas through the whole sequence.

The segmentation result is shown in Fig. 12, where the 3D result is shown in a 2D form and different areas are represented by different gray-scale values. Especially, for the application of automatic expression recognition, in Fig. 13 the segmentation of the face on each frame are put together to show the tracking of the face in the sequence. The segmentations of the lip in each frame are also put together in Fig. 14. Based on the segmentation results, further recognition can be carried out.


(a) frame 1


(b) frame 2


(c) frame 3


(d) frame 4


(e) frame 5


(f) frame 6

Fig. 10 The video sequence of a TV presenter


(a) the sign distribution of the relative potential in frame 1


(b) the sign distribution of the relative potential in frame 2


(c) the sign distribution of the relative potential in frame 3


(d) the sign distribution of the relative potential in frame 4


(e) the sign distribution of the relative potential in frame 5


(f) the sign distribution of the relative potential in frame 6

Fig. 11 The sign distribution of the relative potential values for the “TV presenter” video sequence


(a) the segmentation result on the plane of frame 1


(b) the segmentation result on the plane of frame 2


(c) the segmentation result on the plane of frame 3


(d) the segmentation result on the plane of frame 4


(e) the segmentation result on the plane of frame 5


(f) the segmentation result on the plane of frame 6

Fig. 12 The segmentation results for the “TV presenter” video sequence

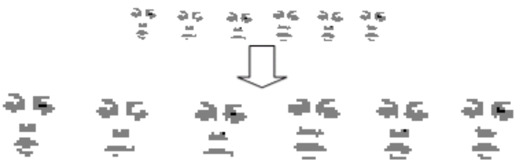


Fig. 13 The sequence of the face (including eyebrows, eyes, nose and lip) segmented from the video of the TV presenter

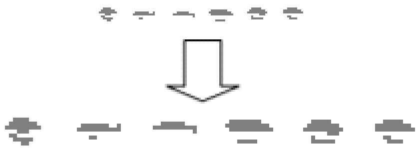


Fig. 14 The sequence of the moving lip segmented from the video of the TV presenter

Fig. 15 is the sequence of a person raising his arms, which has the size 160×120 for each frame. In the experiment, the sign distribution of the relative potential in the 3D space is recorded. The result is shown in a 2D form in Fig. 16, where the sign distribution in each frame is shown separately. According to Fig. 16, the main area of the person can be segmented as a connected region with the same sign of the relative potential value. Moreover, the tracking of the person’s action is also possible based on the segmentation result because the result includes connected 3D area of the person through the whole sequence. The segmentation result is shown in Fig. 17, where the 3D result is shown in a 2D form and different areas are represented by different gray-scale values. For the application of automatic posture and behavior recognition, in Fig. 18 the segmentation of the person on each frame are put together to show the tracking of his action in the sequence. Based on the segmentation results, further recognition can be carried out.


(a) frame 1

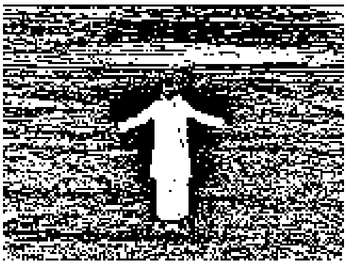

(b) frame 2

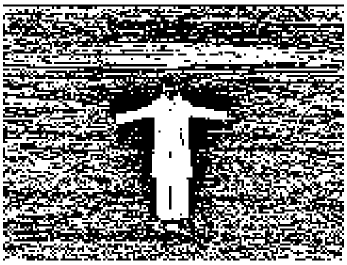

(c) frame 3

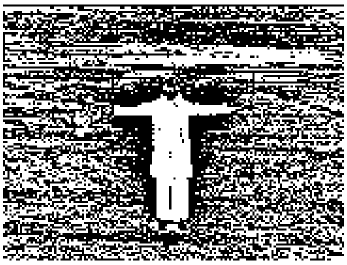

(d) frame 4

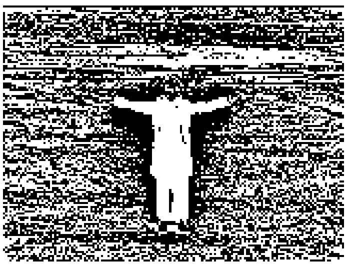

e) frame 5

Fig. 15 The video sequence of a person raising the arms


(a) the sign distribution of the relative potential in frame 1


(b) the sign distribution of the relative potential in frame 2


(c) the sign distribution of the relative potential in frame 3


(d) the sign distribution of the relative potential in frame 4

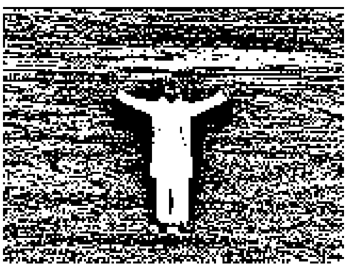
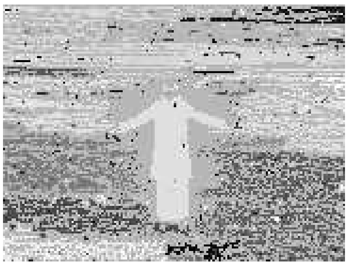
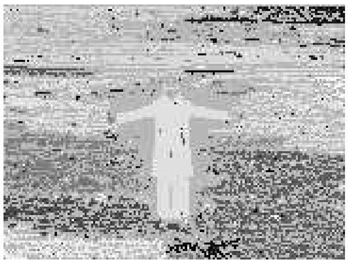
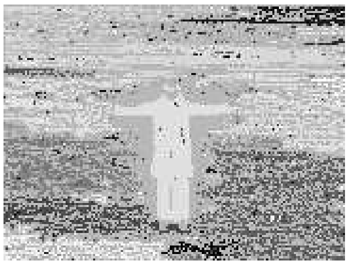
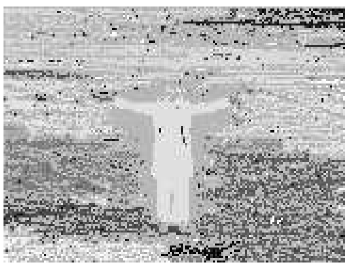

(e) the sign distribution of the relative potential in frame 5

Fig. 16 The sign distribution of the relative potential values for the “raising arms” video sequence


(a) the segmentation result on the plane of frame 1


(b) the segmentation result on the plane of frame 2


(c) the segmentation result on the plane of frame 3


(d) the segmentation result on the plane of frame 4

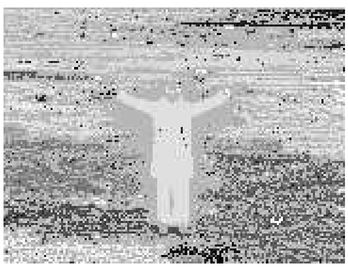

(e) the segmentation result on the plane of frame 5

Fig. 17 The segmentation results for the “raising arms” video sequence

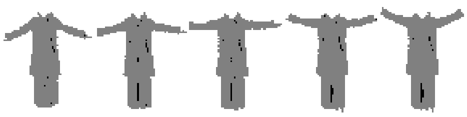


Fig. 18 The sequence of the “raising arms” action segmented from the video

Fig. 19 is a traffic sequence, which has the size 160×120 for each frame. In the experiment, the sign distribution of the relative potential in the 3D space is recorded. The result is shown in a 2D form in Fig. 20, where the sign distribution in each frame is shown separately. According to Fig. 20, the main area of the car at the front can be segmented as a connected region with the same sign of the relative potential. The tracking of the front car is possible based on the segmentation result because the result includes connected 3D area of the car through the whole sequence. The segmentation result is shown in Fig. 21, where the 3D result is shown in a 2D form and different areas are represented by different gray-scale values. For the application of automatic car tracking, in Fig. 22 and Fig. 23 the segmentations of the front car in each frame are put together to show the vertical and horizontal movement respectively. In Fig. 22, the vertical positions of the car in each frame are kept unchanged to show the vertical translation through the sequence. In Fig. 23, the horizontal positions of the car in each frame are kept unchanged to show the horizontal translation through the sequence. Then the movement of the car can be estimated for further analysis.

(a) frame 1

(b) frame 2

(c) frame 3

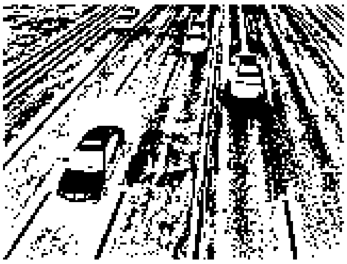
(d) frame 4

Fig. 19 A traffic video sequence

(a) the sign distribution of the relative potential in frame 1

(b) the sign distribution of the relative potential in frame 2

(c) the sign distribution of the relative potential in frame 3



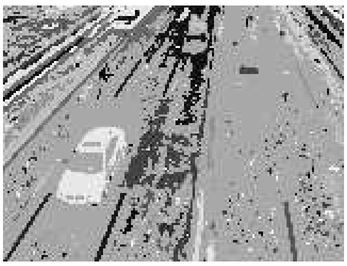
(d) the sign distribution of the relative potential in frame 4

Fig. 20 The sign distribution of the relative potential values for the traffic video sequence

(a) the segmentation result on the plane of frame 1

(b) the segmentation result on the plane of frame 2

(c) the segmentation result on the plane of frame 3



(d) the segmentation result on the plane of frame 4

Fig. 21 The segmentation results for the traffic video sequence

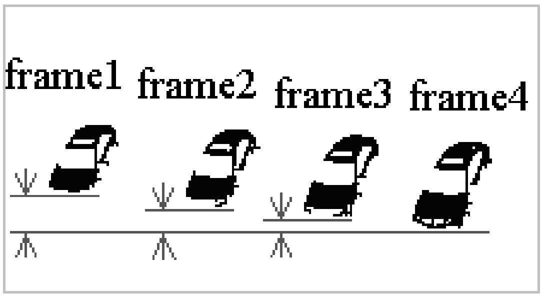


Fig. 22 The segmented car and its movement on the vertical direction in the image sequence

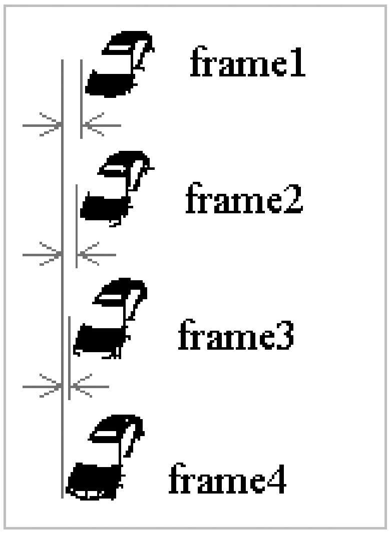


Fig. 23 The segmented car and its movement on the horizontal direction in the image sequence

5 Conclusion

In this paper, a novel method of image sequence analysis is presented inspired by physical electro-static field. The spatial property of the 3D relative potential field is studied, based on which the segmentation of image sequences in the 3D signal space is proposed. The experimental results indicate that the 3D relative potential field of image sequences can serve as a natural representation of 3D region border for object segmentation. Moreover, the segmented 3D bodies from the sequence provide a convenient way for object tracking and analysis. The effectiveness of the relative potential method is based on some unique characteristics of its mathematical form, which serves as a suitable model for the representation of the local-global relevance between image points. The structure information of the sequence can be revealed by the relative potential transform. The sign distribution of the relative potential values can serve as the feature of 3D region border, based on which image sequence segmentation can be performed. Further work will study detailed properties of the 3D relative potential field for image sequences, and its application on other tasks of image sequence processing will also be investigated.

References:

[1] Antonios Oikonomopoulos, Ioannis Patras, Maja Pantic, Spatiotemporal localization and categorization of human actions in unsegmented image sequences, *IEEE Transactions on Image Processing*, Vol. 20, No. 4, 2011, pp. 1126-1140.

[2] Kanglin Chen, Dirk A. Lorenz, Image sequence interpolation based on optical flow, segmentation, and optimal control, *IEEE Transactions on Image Processing*, Vol. 21, No. 3, 2012, pp. 1020-1030.

[3] Iulian Udroiu, Ioan Tache, Nicoleta Angelescu, Ion Caciula, Methods of measure and analyse of video quality of the image, *WSEAS Transactions on Signal Processing*, Vol. 5, No. 8, 2009, pp. 283-292.

[4] Radu Dobrescu, Matei Dobrescu, Dan Popescu, Parallel image and video processing on distributed computer systems, *WSEAS Transactions on Signal Processing*, Vol. 6, No. 3, 2010, pp. 123-132.

[5] D. J. Hurley, M. S. Nixon and J. N. Carter, Force field feature extraction for ear biometrics, *Computer Vision and Image Understanding*, Vol. 98, No. 3, 2005, pp. 491-512.

[6] X. D. Zhuang and N. E. Mastorakis, The Curling Vector Field Transform of Gray-Scale Images: A Magneto-Static Inspired Approach, *WSEAS Transactions on Computers*, Vol. 7, Issue 3, 2008, pp. 147-153.

[7] G. Abdel-Hamid and Y. H. Yang, Multiscale Skeletonization: An electrostatic field-based approach, *Proc. IEEE Int. Conference on Image Processing*, Vol. 1, 1994, pp. 949-953.

[8] Luo, B., Cross, A. D. and Hancock, E. R., Corner Detection Via Topographic Analysis of Vector Potential, *Pattern Recognition Letters*, Vol. 20, No. 6, 1999, pp. 635-650.

[9] Andrew D. J. Cross and Edwin R. Hancock, Scale-space vector field for feature analysis, *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1997, pp. 738-743.

[10] K. Wu and M. D. Levine, 3D part segmentation: A new physics-based approach, *IEEE International symposium on Computer Vision*, 1995, pp. 311-316.

[11] N. Ahuja and J. H. Chuang, Shape Representation Using a Generalized Potential Field Model, *IEEE Transactions PAMI*, Vol. 19, No. 2, 1997, pp. 169-176.

[12] T. Grogorishin, G. Abdel-Hamid and Y.H. Yang, Skeletonization: An Electrostatic Field-Based Approach, *Pattern Analysis and Application*, Vol. 1, No. 3, 1996, pp. 163-177.

[13] Xiao-Dong Zhuang, Nikos E. Mastorakis, A magneto-statics inspired transform for structure representation and analysis of digital images, *WSEAS Transactions on Computers*, Vol. 8, No. 5, 2009, pp. 874-883.

[14] X. D. Zhuang, N. E. Mastorakis, A novel field-source reverse transform for image structure representation and analysis, *WSEAS Transactions on Computers*, Vol. 8, No. 2, 2009, pp. 376-385.

[15] P. Hammond, *Electromagnetism for Engineers: An Introductory Course*, Oxford University Press, USA, forth edition, 1997.

[16] I. S. Grant and W. R. Phillips, *Electromagnetism*, John Wiley & Sons, second edition, 1990.

[17] Terence W. Barrett, *Topological foundations of electromagnetism*, World Scientific series in contemporary chemical physics, Vol. 26, World Scientific, 2008.

[18] Minoru Fujimoto, *Physics of classical electromagnetism*, Springer, 2007.

[19] Gustavo Carneiro, Allan D. Jepson, Flexible Spatial Configuration of Local Image Features, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, 2007, pp. 2089-2104.

[20] C. R. Shyu, C. E. Brodley, A. C. Kak, A. Kosaka, A. Aisen, L. Broderick, Local versus global features for content-based image retrieval, *IEEE Workshop on Content-Based Access of Image and Video Libraries*, 1998, pp. 30-34.

[21] Y. Shelepin, A. Harauzov, V. Chihman, S. Pronin, V. Fokin, N. Foreman, Incomplete image perception: Local features and global description, *International Journal of Psychophysiology*, Vol. 69, Issue 3, 2008, pp. 164.

[22] Aude Oliva, Antonio Torralba, Building the gist of a scene: the role of global image features in recognition, *Progress in brain research*, Vol. 155, 2006, pp. 23-36.

[23] Yuntao Qian, Rongchun Zhao, Image segmentation based on combination of the global and local information, *International*

Conference on Image Processing, Vol. 1, 1997, pp. 204-207.

[24] Dimitri A. Lisin, Marwan A. Mattar, Matthew B. Blaschko, Erik G. Learned-Miller, Mark C. Benfield, Combining Local and Global Image Features for Object Class Recognition, *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 03, 2005, pp. 47.

[25] Takahiro Toyoda, Osamu Hasegawa, Random Field Model for Integration of Local Information and Global Information, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, 2008, pp. 1483-1489.

[26] J.A. Montoya-Zegarra, J. Beeck, N. Leite, R. Torres, A. Falcao, Combining Global with Local Texture Information for Image Retrieval Applications, *10th IEEE International Symposium on Multimedia*, 2008, pp. 148-153.

[27] M. Aly, P. Welinder, M. Munich, P. Perona, Automatic discovery of image families: Global vs. local features, *16th IEEE International Conference on Image Processing*, 2009, pp. 777-780.