

Image Sequence Analysis Based on the 3D Relative Potential Inspired by Physical Electro-Static Field

XIAODONG ZHUANG^{1,2} and N. E. MASTORAKIS^{1,3,4}

1. WSEAS Research Department, Agiou Ioannou Theologou 17-23, 15773, Zografou, Athens, GREECE (xzhuang@worldses.org)
2. Automation Engineering College, Qingdao University, Qingdao, CHINA
3. Department of Computer Science, Military Institutions of University Education, Hellenic Naval Academy, Terma Hatzikyriakou, 18539, Piraeus, GREECE
mastor@wseas.org <http://www.wseas.org/mastorakis>
4. Technical University of Sofia, BULGARIA

Abstract: - A novel three-dimensional relative potential field is proposed for image sequence analysis, which is inspired by the physical electro-static potential. The spatial property of the 3D relative potential is studied, based on which the 3D volume segmentation can be implemented in image sequences. The experimental results of test sequences and real-world video sequences prove that the proposed method provides an effective and convenient way of object segmentation and tracking in video sequences, which can effectively support further analysis and recognition.

Key-Words: - image sequence processing, three-dimensional relative potential, electro-statics, video segmentation, object tracking

1 Introduction

Image sequence processing is one of the significant research topics in digital image processing, which has significant value in theory and practice [1-4]. An image sequence can be regarded as a 3D signal $g(x,y,t)$. The 3D signal of image sequence can have much more information than a single 2D image $f(x,y)$. Many real-world applications are based on image sequence processing such as security monitoring, traffic surveillance, medical image sequence analysis, etc. For such practical tasks, the trend is advanced intelligent analysis which aims at accurate object recognition, human gesture recognition and even behavior recognition. The practical tasks continuously give new requirements of processing techniques, which has attracted lots of research interest and efforts.

Recently, physics inspired methodology has attracted more and more research interest in image processing, which exhibits the ability of effective feature extraction and analysis [5-14]. Electro-static field plays an important role in such methods [5,11,12,14]. The fundamental idea underlying the methods inspired by physical fields is the transform from one form of the field to another (i.e. from field source to its potential) so that the feature of interest can be revealed [5,14]. Some unique characteristics of the physical fields are exploited in such methods, which leads to impressively effective processing

results. Preliminary results indicate the promising wide application of such methods in practical tasks.

Currently, most physics-inspired methods concentrate on the processing of single 2D images. In the author's previous research, the authors have studied a series of novel methods inspired by electro-magnetics for single 2D image analysis [6,13,14]. In this paper, the physical field inspired methodology is extended to image sequence analysis based on the authors' previous work. In fact, the relative potential and most vector field methods in the previous work have some similarity in the definition of the virtual field, including the relative potential field, diffusing vector field, curling vector field and compressing vector field [6,13,14]. All of them define the virtual field by introducing the measurement of pixel difference $g(i,j)-g(x,y)$ into their definitions. Such virtual fields belong to the "relative field", which are defined with the relative value of one pixel compared to another (i.e. $g(i,j)-g(x,y)$) [6,13,14]. In this paper, the 3D relative potential field is studied as a typical example of extending the "relative field" for 2D images to 3D image sequence analysis. Because the similarity in definition, other "relative field" methods in the authors' previous research may be extended to their 3D forms in a similar way.

The method described in this paper is named "three-dimensional relative potential", which

exploits the relationship between the source and the potential in the physical electro-static field. The spatial property of the 3D relative potential is investigated by both theoretic analysis and experiments, which proves that the positive-negative sign distribution of the 3D relative potential value is a natural and convenient representation for 3D volume separation and segmentation. The segmented 3D volumes in the 3D relative potential field provide an effective object tracking method in image sequences, which can be used in further intelligent recognition.

2 The 3D relative potential field of image sequences

In the physical electro-static field, the potential is determined by the source (i.e. the charge distribution) [15-18]. Therefore, the potential field can reflect some characteristics of the source. This relationship between the field and its source can be exploited in image transform, in which the image is regarded as the source (i.e. the pixels are regarded as discrete charges) and the generated virtual field may reveal important features of the images. The attraction of physical field inspired methods is the possibility of a natural representation of image structure or components without artificially set parameters (such as thresholds in image segmentation). In this paper, a general form of virtual potential field for 3D image sequence is proposed, which is inspired by the physical electro-static field.

2.1 The electro-static potential and its spatial property

The formula of the physical electro-static potential generated by a charge q is [15-18]:

$$V = \frac{1}{4\pi\epsilon} \cdot \frac{q}{r} \quad (1)$$

where V is the electro-static potential at a space point. q is the charge quantity. r is the distance between the charge and the space point. ϵ is a physical constant.

For a charge distribution ρ in the space, the potential generated by ρ on point (x,y) is [15-18]:

$$V = \frac{1}{4\pi\epsilon} \int_V \frac{\rho \cdot d\tau}{r} \quad (2)$$

where V is the electro-static potential at a space point. The integral in Equation (2) is for the 3D area where the charge distribution ρ exists.

Many image processing techniques involves local operations in the image, i.e. local image features are extracted and analyzed [19-21]. Local image features usually have the form of a function $f(x,y,z)$ defined in the three-dimensional space of the image sequence. On the other hand, the analysis of the image also requires consideration of the neighbouring area of each image point in order to get practically useful results. The local and global analyses are both needed in many image processing tasks [22-27]. Generally speaking, neighbouring points have stronger relevance than remote points, i.e. the closer the distance, the stronger the relevance. In many image processing tasks, it is necessary to consider the balance between the strong local relevance of close neighbouring points and a wide range of weaker relevance of remote points. Equation (2) indicates that the potential of a charge q on a space point (i.e. the impact of q on that point) is in direct proportion to the reciprocal of the distance r . The mathematical form of the distance reciprocal in Equation (2) inspires the representation of the local-global relevance between image points. For a point p in the space, the near charge distribution in the small local neighboring area has greater impact on p 's potential than remote charge distribution. On the other hand, no matter how far the distance is, remote charge distribution still has relatively weak impact on p 's potential. Moreover, the accumulation of the weak impacts of wide-range remote charge distribution can not be neglected. The above characteristic of the distance reciprocal form in Equation (2) is quite suitable for the requirement of image analysis that both local and global relevance between image points should be considered.

2.2 A general form of virtual potential field for images

The electro-static potential has a suitable mathematical form to model the local-global relevance of image points. Here a general form of virtual image potential field is proposed with the electro-static analogy. For image analysis, not only the distance between two image points but also the relationship between their gray-scale should be considered. Therefore, a general continuous form of virtual potential field for image sequences is proposed as:

$$V_c^m(x, y, z) = A \cdot \int_a^b \int_b^c \int_c^m \frac{f(g(a,b,c), g(x,y,z))}{r_{(a,b,c) \rightarrow (x,y,z)}^m} da \cdot db \cdot dc \quad (3)$$

where $V_c^m(x,y,z)$ is the continuous image potential value at point (x,y,z) . (x,y,z) and (a,b,c) are coordinates in the 3D space. Here the 3D space corresponds to the image sequence, which includes the single image as a special 2D case. A is a predefined constant value. g is the gray-scale value of image points in the sequence (here gray-scale image sequences are considered). f is a function defined according to specific image processing tasks, which represents the relationship between the gray-scale values of point (x,y,z) and (a,b,c) . r is the distance between (x,y,z) and (a,b,c) . m is a constant that affect the reciprocal's decreasing rate with the increasing distance r . The multiple integral in Equation (3) is on the three-dimensional image space. For a specific processing task, the function f , the constants A and m should be pre-defined according to the specific processing purpose.

For digital image sequences, the discrete form of the virtual potential field is presented as the discrete form of Equation (3):

$$V_d^m(x,y,z) = A \cdot \sum_{k=0}^{D-1} \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} \frac{f(g(i,j,k), g(x,y,z))}{r_{(i,j,k) \rightarrow (x,y,z)}^m} \quad (4)$$

$(k \neq z \text{ or } j \neq y \text{ or } i \neq x)$

where $V_d^m(x,y,z)$ is the discrete image potential on point (x,y,z) . A is a predefined constant value. H and W are the height and width of the digital image respectively. D is the depth of the image sequence (i. e. the frame number), which may represent the temporal position in a video sequence, or the spatial position in a sequence of section scanning. g is the gray-scale value of image points. f is a function defined according to specific image processing tasks, which represents the relationship between the gray-scale values of point (x,y,z) and (i,j,k) . r is the distance between (x,y,z) and (i,j,k) . m is a constant that affect the reciprocal's decreasing rate with the increasing distance r . Equation (4) imitates the form of the physical electro-static potential. It is defined in a flexible form, which extends the physical formula of (2) by introducing the function f and constant m adjustable for different processing tasks.

2.3 The definition of 3D relative potential field

For some important image processing tasks such as edge detection and segmentation, the difference between pixels (i.e. their gray-scale difference) is the factor of major consideration. In this paper, the 3D relative potential is proposed for gray-scale image sequences based on the general form of discrete image potential, where the function $f(g(i,j,k),$

$g(x,y,z))$ in Equation (4) is specialized as the difference between the gray-scale values of the two points (x,y,z) and (i,j,k) in the sequence:

$$V_R^m(x,y,z) = A \cdot \sum_{k=0}^{D-1} \sum_{j=0}^{H-1} \sum_{i=0}^{W-1} \frac{g(i,j,k) - g(x,y,z)}{r_{(i,j,k) \rightarrow (x,y,z)}^m} \quad (5)$$

$(k \neq z \text{ or } j \neq y \text{ or } i \neq x)$

where $V_R^m(x,y,z)$ is the discrete relative potential at the point (x,y,z) . A is a predefined constant value. D , H and W are the depth, height and width of the image sequence respectively. g is the gray-scale value of the points in the sequence. r is the distance between (x,y,z) and (i,j,k) . m is a constant that affect the reciprocal's decreasing rate with the increasing distance r .

Compared with the mathematic form of the electro-static potential, the proposed relative potential has two major differences. One is the replacement of the charge with the gray-scale difference, which can make the relative potential represents the difference of one point between others in the sequence. This is the reason why the virtual potential is called "relative". The other is the m -th power of the distance r . Thus the adjustment of the value m can change the decreasing rate of the relevance between image points with the increasing distance r according to the requirement of a specific task.

3 The spatial property of the 3D relative potential field

In Equation (5), the relevance between two image points with distance r is represented quantitatively by the reciprocal of r^m . The value of relative potential is virtually the weighted sum of the gray-scale difference between the image point (x,y,z) and all other points, and the weight is the factor of relevance, i.e. the reciprocal of r^m .

According to the definition of the image relative potential in Equation (5), the relative potential value of a point p is mainly affected by its local neighboring area in the 3D space. The local neighboring area consists of two classes of points. One class is those in the same region of p (i.e. with similar gray-scale of p), the other is those in the different region. To investigate the properties of the relative potential field, experiments are carried out for a series of simple test image sequences. For these simple test image sequences, the gray-scale difference in the same region is zero. Thus the relative potential of p is mainly affected by the gray-scale difference between p 's region and its adjacent

area. Suppose A and B are two adjacent areas shown in Fig. 1. p_a and p_b are two border points at different border sides. p_a is in area A and p_b is in area B . g_a and g_b are the gray-scale of area A and B respectively. According to the above discussion, the sign of p_a 's relative potential is determined by $g_b - g_a$, while the sign of p_b 's relative potential is determined by $g_a - g_b$. Thus the signs of p_a and p_b are opposite. Therefore, in the 3D relative potential field, the sign of the relative potential value will reverse across the border between adjacent 3D areas. In another word, the sign of the relative potential values in two different adjacent 3D bodies (or 3D volumes) is different. This spatial feature of the relative potential field can be exploited in image sequence analysis.

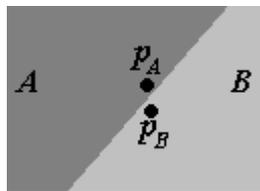
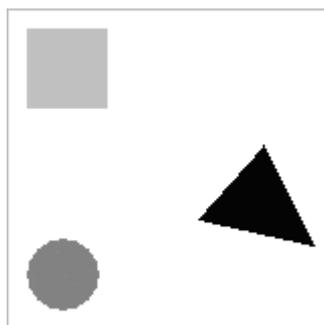


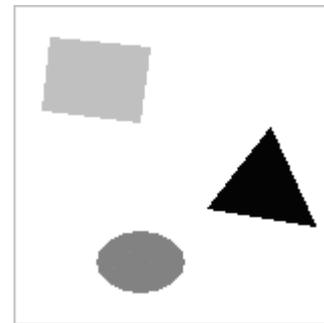
Fig. 1 p_a and p_b on different sides of the area border

In order to study the spatial property of the 3D relative potential field, experiments are carried out on a group of test image sequences. The simple test sequences are of the size 160×160 for each frame. Each sequence has 6 frames. The 3D relative potential field is calculated for each sequence. When computing the relative potential values, the constant m in Equation (5) is pre-defined as $m=3$. Fig. 2 and Fig. 3 show the experimental results for a typical test sequence, which combines the shifting, rotating and deforming simultaneously. In the experiment, the sign distribution of the 3D relative potential field is recorded for each frame.

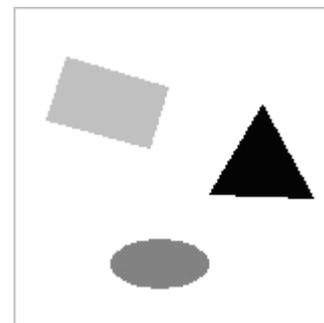
In Fig. 2, there are three simple shapes (square, triangle and circle). They move by shifting, rotating and deforming. Such test is more complex in the motion of objects, which aims to investigate the relative potential distribution for sequences with objects of combinational motion.



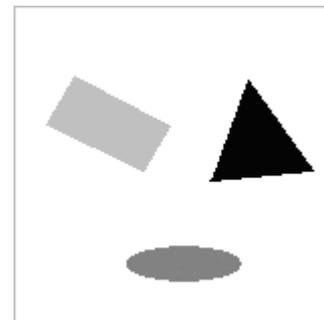
(a) frame 1



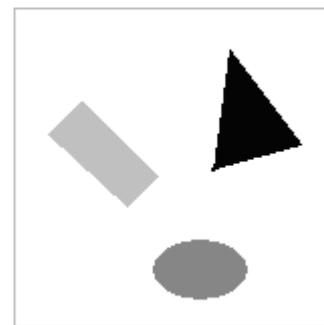
(b) frame 2



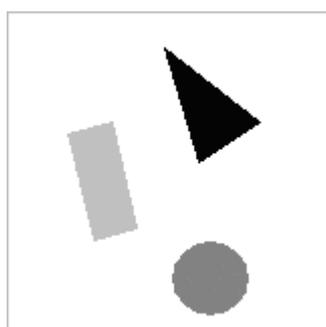
(c) frame 3



(d) frame 4



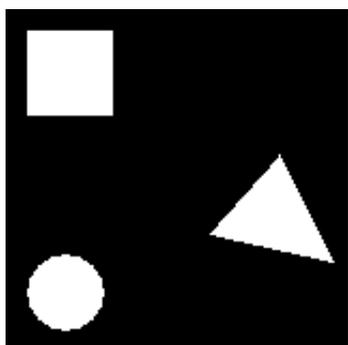
(e) frame 5



(f) frame 6

Fig. 2 The frames of the “combinational motion” sequence

In Fig. 3, the sign distribution of the relative potential values are shown for the Sequence of “combinational motion”, where white points represent the positive sign and the black points represents the negative sign. It is indicated that the sign of the relative potential value will reverse across the border of two adjacent 3D bodies.



(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. 3 The sign distribution of the 3D relative potential value on each frame of the “combinational motion” sequence

The above experimental results indicate the importance and possible application of the sign distribution in the 3D relative potential field. Since the sign of the relative potential reverses across the border of adjacent object areas, it may serve as the feature of 3D border. This spatial feature can be exploited in the segmentation of 3D volumes for image sequences, which is proved by the following experiments in the next section.

4 3D segmentation of image sequence in the relative potential field

Based on the above analysis, 3D segmentation is implemented for the simple test sequences, which separates each connected 3D volumes with the same sign of relative potential from others. Since the sign of relative potential are opposite in the two different adjacent spatial areas, this can provide the basis of object segmentation and tracking in the image sequence. Based on the experimental results, a method of 3D body division in the relative potential field is proposed as following:

Step1: Calculate the 3D relative potential field;

Step2: Obtain the sign distribution of the relative potential field;

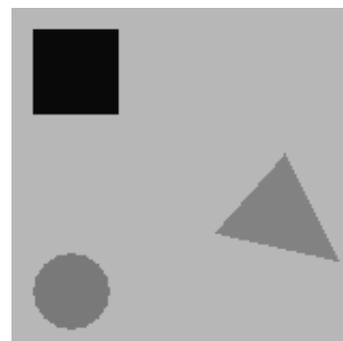
Step3: Group the adjacent points with the same sign of relative potential into connected 3D bodies. In the grouping process, the adjacent points of 6-connection in the 3D space for a point p at (x,y,z) is investigated. The points of 6-connection for (x,y,z) include: $(x+1,y,z)$, $(x-1,y,z)$, $(x,y+1,z)$, $(x,y-1,z)$, $(x,y,z+1)$ and $(x,y,z-1)$. If any of the six adjacent points has the same sign of relative potential as p , it is grouped into the 3D region which p belongs to. The obtained connected 3D regions are the result of segmentation for the image sequence.

The obtained set of connected regions is the result of 3D volume division for the gray-scale image sequence. Therefore, the 3D signal space of the sequence is segmented to several 3D bodies. Each 3D body is a connected 3D area in which the points have the same sign of relative potential value.

4.1 The 3D segmentation for the testing image sequence

The segmentation result for a typical test image sequence is shown in Fig. 4. Fig. 4 shows the 3D segmentation result for the sequence of "combinational motion" according to the sign distribution in Fig. 3, where different 3D bodies are represented by different gray-scale values for

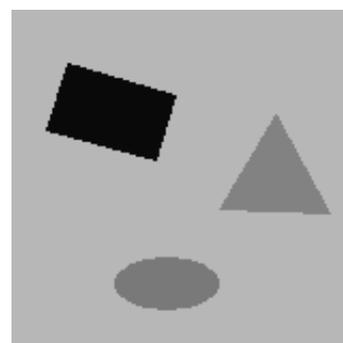
visualization. The segmented 3D bodies are shown as a series of sections in 2D form. In another word, each 2D section of the bodies on the corresponding frame plane is shown separately, and in each section the areas of different 3D bodies are of different gray-scale values.



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



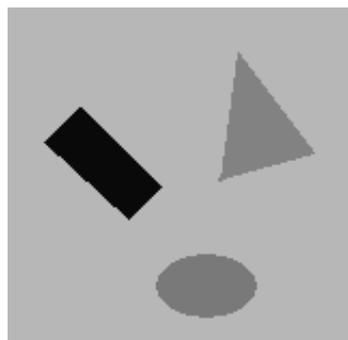
(b) the segmentation result on the plane of frame 2



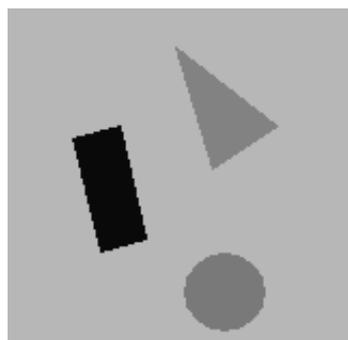
(c) the segmentation result on the plane of frame 3



(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. 4 The sections of the segmented 3D bodies on each frame of the “combinational motion” sequence

In order to show the segmentation results in a clearer way, the sections shown in Fig. 4 are overlaid together in Fig. 5, which can clearly show the movement of each object through the sequence. In Fig. 5, only the region borders in each section are shown, and the arrows show the moving direction of each object.

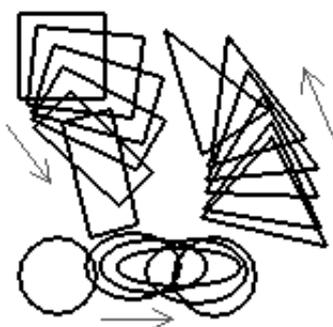


Fig. 5 The overlaid result of the 2D sections in Fig. 4

In the above experimental results, the regions of the simple objects are effectively separated from others in each single frame. Moreover, because the segmentation is implemented in the 3D space, the results are 3D bodies (or 3D volumes) in the space. Each 2D region segmented in a single frame is just a section of a corresponding 3D body. In another

word, each segmented 3D body is formed by a corresponding object region which runs through the frame sequence (maybe with the movements of shifting, rotating or deforming). Therefore, the 3D segmentation result can provide a basis of object tracking for further analysis and recognition.

4.2 The 3D segmentation for the real-world image sequences

To investigate the practical application of the above 3D segmentation method, experiments have also been carried out for real world image sequences. Real world image sequences are more complex than the test sequences. It is much more difficult to handle the real world images because of noise, shadow, sheltering between moving objects, etc. In this section, experiments are carried out on lots of real world sequences for several hot topics of research and application. These topics include: hand tracking and segmentation (for gesture recognition), face tracking and segmentation (for expression recognition), the tracking and segmentation of the speaker’s lip (for automatic lip reading by computers), moving person tracking and segmentation (for posture recognition and human behavior identification), and vehicle tracking and segmentation (for vehicle type recognition). Such experiments aim at the proving of the proposed method’s practical effectiveness, and also possible improvements of the method in practical use. Some of the sequences used in the experiments are captured by the authors with a digital camera. The others are obtained from public image databases on the Internet.

Fig. 6 is the sequence of a clenching hand, 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. 7, where the sign distribution in each frame is shown separately. According to Fig. 7, the main area of the clenching hand can be segmented as a connected region of the same sign of relative potential value. Moreover, the tracking of the hand is also possible based on the segmentation result because in the result the hand area is a connected 3D area through the whole sequence. The segmentation result is shown in Fig. 8, where the 3D result is shown in a 2D form and different areas are represented by different gray-scale values. In Fig. 9, the segmentation on each frame are put together to show the tracking of the hand’s clenching process. Based on the segmentation result, further hand gesture recognition can be carried out. The behavior

identification of the hand can also be studied based on the segmentation.



(a) frame 1



(b) frame 2



(c) frame 3



(d) frame 4



(e) frame 5



(f) frame 6

Fig. 6 The video sequence of a clenching hand



(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. 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



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



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

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

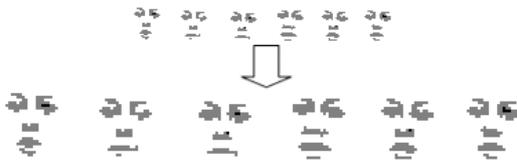


Fig. 13 The sequence of the face (including eyebrows, eyes, nose and lip) segmented from the video of the TV presenter

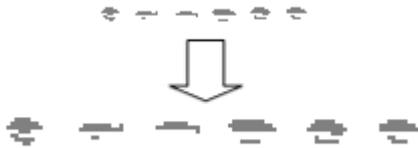


Fig. 14 The sequence of the moving lip segmented from the video of the TV presenter

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

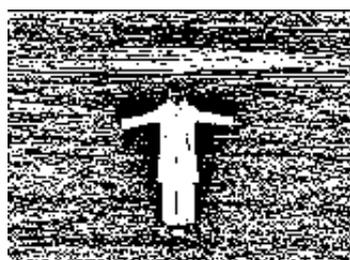


e) frame 5

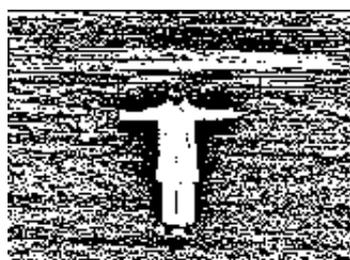
Fig. 15 The video sequence of a person raising the arms



(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

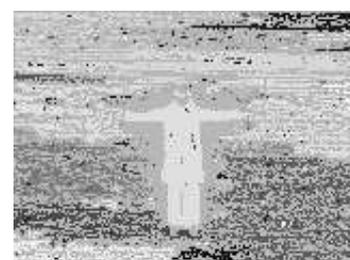
Fig. 16 The sign distribution of the relative potential values for the “raising arms” video sequence



(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

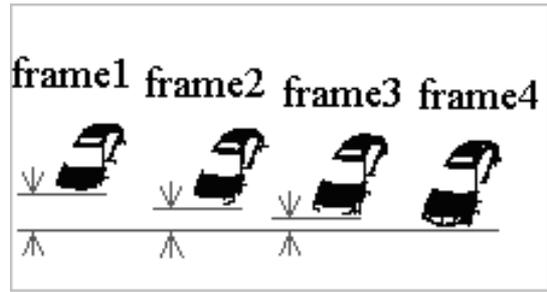


Fig. 22 The segmented car and its movement on the vertical direction in the image sequence

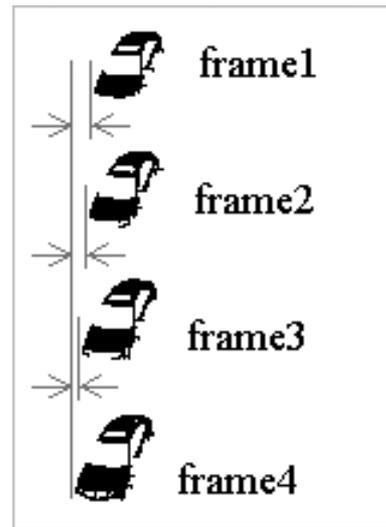


Fig. 23 The segmented car and its movement on the horizontal direction in the image sequence

5 Conclusion

In this paper, a novel method of image sequence analysis is presented inspired by physical electrostatic 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 Udroi, 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.