Applied Mathematics & Information Sciences
An International Journal

http://dx.doi.org/10.12785/amis/070417

Perfect Snapping: An Accurate and Efficient Interactive Image segmentation Algorithm

Qingsong Zhu[1,2,3,4], *Guanzheng Liu*[6], *Zhanyong Mei*[1], *Qi Li*[5], *Yaoqin Xie*[1,2,3,4,*] *and Lei Wang*[1,2,3,4,*]

[1]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, 518055, Shenzhen, China
[2]School of Medicine, Stanford University, Stanford, California, USA
[3]Key Lab for Health Informatics, Chinese Academy of Sciences, 518055, Shenzhen, China
[4]Key Lab for Low-Cost Healthcare, Chinese Academy of Sciences, 518055, Shenzhen, China
[5]School of Software Engineering, University of Science and Technology of China, 230026, Hefei, China
[6]School of Engineering, Sun Yat-Sen University, 510000, Guangzhou, China

Received: 30 Dec. 2012, Revised: 16 Feb. 2013, Accepted: 3 Mar. 2013
Published online: 1 Jul. 2013

Abstract: Interactive image segmentation is a process that extracts a foreground object from an image based on limited user input. In this paper, we propose a novel interactive image segmentation algorithm named Perfect Snapping which is inspired by the presented method named Lazy Snapping technique. In the algorithm, the mean shift algorithm with a boundary confidence prior is introduced to efficiently pre-segment the original image into homogeneous regions (super-pixels) with precise boundary. Secondly, Gaussian Mixture Model (GMM) clustering algorithm is used to describe and to model the super-pixels. Finally, a Monte Carlo based Expectation Maximization (EM) algorithm is used to perform parametric learning of mixture model for priori knowledge. Experimental results indicate that the proposed algorithm can achieve higher segmentation quality with higher efficiency.

Keywords: Interactive Image Matting, Mean Shift Algorithm, Lazy Snapping

1. Introduction

Interactive image segmentation [1,2] has been an important technique in image processing and video editing, which refers to the problem of softly extracting the foreground objects from a single input image. With the rapid development of digital image processing techniques, image segmentation has become possible to segment the foreground objects on an individual pixel level. And a variety of image segmentation algorithms have been proposed and used in post image and video editing. Purpose of image segmentation is to specify which parts of the image belong to the foreground and which parts belong to background. Usually, a user imposes certain hard constraints for segmentation by indicating certain pixels (seeds) that absolutely have to be part of the foreground and certain pixels that have to be part of the background. Intuitively, these hard constraints provide clues on what the user intends to segment.

1.1. Image Matting Algorithms

The early matting algorithms are based on known background. The blue screen matting [3] was used for live action matting, whose principle is to photograph the subject against a constant-colored background (typically blue and green). Recently, in the field of image matting study, many natural image matting approaches have been proposed. In natural image matting processing, moderate user interactions are essential. In the Knockout [4] method, the algorithm starts from the known foreground and background of the trimap and extrapolates the known foreground and background colors into the unknown region to estimate the alpha matte In Ruzon and Tomasi's approach [5], a statistical method is proposed to analyze the color samples of the foreground and background and the estimation of In [6], a new image matting algorithm based on principal components analysis (PCA) is introduced to analyze the foreground and background samples. This method utilizes the projection method to estimate and have also a considerable computation cost.

* Corresponding author e-mail: yq.xie@siat.ac.cn, wang.lei@siat.ac.cn

In [7], a Bayesian framework based image matting approach is proposed. In this method, color samples of foreground and background are clustered and modeled as mixture Gaussian distribution. An estimation named maximum a posterior (MAP) is introduced to calculate foreground, background and simultaneously for each pair of the foreground and background in a Bayesian framework. The final estimate of is chosen from the pair of the foreground and background that provides the maximum likelihood. Although the algorithm achieved a better result of matting, the algorithm is quite complex and has a slower processing speed compared with Knockout method. In [8], the Poisson Matting approach for natural images matting of complex scenes is presented. The basic idea behind the algorithm is to formulate the problem of natural image matting as one of solving Poisson equations with the matte gradient field. Unlike previous methods that optimize a pixel's alpha matte in a statistical manner, the Poisson Matting method operates directly on the gradient of the matte, which reduces greatly the error caused by mis-classification of color samples in a complex scene. But the shortcomings of the algorithm include two aspects. Firstly, when the foreground and background colors are very similar, the matting equation becomes ill-conditioned. Secondly, when the matte gradient estimated in global Poisson Matting largely biases the true values, more user interaction is required.

In [9], a novel approach called Flash Matting is proposed. This algorithm can robustly recover the matte from flash/no-flash images, even for scenes in which the foreground and the background are similar or the background is complex. In [10], an image matting approach based on Belief Propagation is presented. The approach can achieve better matting results only with less trimap restriction by utilizing the discrete set of value to formulate the matting problem as energy minimum problem. But the algorithm is quite time consumed subject to the iterative processing. Another scribble-based method is proposed in [11]. The algorithm has a better interaction performance and can achieve a better result only by a few scribbles restriction. But it has a higher computation cost and a lack of color statistic characteristic. A known trimap is essential for the above image matting algorithms. In [11,12,13,14], some interactive Graph Cuts [12] based approaches are introduced. Grab Cut method [13] makes user draw a rectangle around the periphery of the foreground object, and then extracts the foreground objects accurately by image segmentation and feathered process. Grow Cut approach [14] is a geodesic distance based image matting algorithm which utilizes the surface of "hard constraint pixels" that user calibrated to grow outside to complete image matting, and it is difficult to operate the texture region using the method. Lazy Snapping [15] is another well-known image matting algorithm. User only draws a few lines in different places. The region that some lines appear is regarded as foreground region and the region that others lines appear is regarded as background region, and user can separate the object from the background by these lines. However, the imposed lines drawn by user must satisfy sufficiently to represent the colors species in the foreground and background region. Otherwise user must constantly add new lines until get satisfactory results. Moreover, the over-segmentation problem of the method remains to be solved.

1.2. Interactive Image Segmentation Algorithms

Graph Cut theory have been a classical algorithm to find the optimal MAP estimation of various image segmentation problems defined over an Markov Random Field Model. Although the maxflow/mincut approach was introduced early into foreground objects extraction of image, their advantage was exploited based on the work of Boykov et al. [16,17] and their characterization of functions that can be optimized by using Graph Cuts [18]. Graph Cuts Algorithm is widely used in a large quantity of image segmentation problems like multi-camera scent reconstruction [19,20] and image/video segmentation [21,22]. Recently, there are a number of interactive image segmentation algorithms with different human interaction ways. Some representative interactive segmentation methods such as Intelligent Scissor [23], Image Snapping [24], which needs to densely put points or draws curves around the foreground objects boundary, and thus reguires an amount of user interaction especially for the object with some complex shapes. In order to make it easy for user to segment the foreground objects that user is interested in, a more general image segmentation framework based on Gra-ph Cut [25,26] is proposed, which uses a Graph to represent the image and utilize the max-flow/min-cut algorithm to extract accurately the foreground objects. Grab Cut algorithm [27] further simplies the user interaction way only by drawing a rectangle bounding box to cover user interested object, Grab Cut can extract quickly a relatively accurate object. In [28], the Graph Cut method is combined with the random walker approach for a better segmentation quality with a considerable time/space complexity. In [29], a novel geodesic distance based image interactive segmentation approach is presented, in which all unmarked pixels are classified into "Background" and "foreground" by using the calculated geodesic distance to the user-provided scribbles, and thus only less scribbles are needed to achieve desirable segmentation results.

In this paper, a novel image segmentation algorithm named Perfect Snapping [30] is proposed. The algorithm can be divided into the following steps:

a. Use mean shift algorithm with a boundary confidence prior to efficiently pre-segment the original image into homogeneous regions (super-pixels);

b. Perform mainly description and modeling for the super-pixels by Gaussian Mixture Model clustering algorithm;

c. Complete the parametric learning of mixture model for priori knowledge.

Extensive experimental results have been implemented and compared with classical algorithm to show its advantage.

2. Perfect Snapping Algorithm

2.1. Graph Cut with Pre-segmentation

The main idea of the Graph Cut model is to construct an energy function and use weighted graph mapping and network flow theory to convert the global labeling problem to the maximum-flow/minimum-cut problem of the corresponding weighted graph. To indicate the classification of each pixel, we can construct pixel-based Markov Random Field energy function as:

$$E(X) = \sum_{i \in v} E_1(x_i) + \lambda \sum_{i,j \in \varepsilon} E_2(x_i, x_j) \tag{1}$$

Where v is the set of all nodes and is the set of all arcs connecting adjacent nodes. $E_1(x_i)$ is the likelihood energy which measures the energy consumption that a node is defined as foreground or background, and $E_2(x_i, x_j)$ is the prior energy that denotes the cost when the labels of adjacent nodes i and j are x_i and x_j respectively. In order to simplify Graph Cut model, we usually use some pre-segmentation algorithms to segment the original image into some small regions which are regarded as the nodes of the weighted graph to construct the Graph Cut model. Compared with traditional Graph Cut method viewing the pixel as node, the approach greatly simplify the topological structure of weighted graph and reduced the computation cost. In this paper we introduce a mean shift based pre-segment algorithm with boundary prior in place of the watershed method appeared in Lazy Snapping [15]. Mean shift algorithm is an efficient tool used for feature space analysis. To make segmentation results similar in color and continuous in space $l^*u^*v^* \sim x^*y^*$, we perform a mean shift filtering on an original image in 5-D feature space. Assume that the probability density function of 5-D feature space is $f(x)$:

$$\nabla f(x) \propto \sum_{i=1}^{n} (x - x_i) \nabla k\left(\left\| h^{-1}(x - x_i) \right\|^2\right) \tag{2}$$

Where $x_i \in W_{h,z}$, $W_{h,z}$ represents 5-D super-spheroid with center at points x_i and has a $h = \{h_s, h_c\}$ bandwidth. h_s and h_c represent the bandwidth of space and color, respectively. Let the function $g(x) = -k'(x)$ and the corresponding new kernel $G(x) = \lambda g \cdot \|x\|^2$, then the density of new kernel is described by:

$$\nabla f'(x) \propto \sum_{i=1}^{n} (x - x_i) g\left(\left\| h^{-1}(x - x_i) \right\|^2\right) \tag{3}$$

To improve the filtering speed, the pixels are only relegated to the corresponding model attractive regions. After filtering, the model attractive regions are executed recursion and combination according to regions adjacency graph algorithm, color bandwidth and the size of region. To obtain accurate pre-segment results, the mean shift algorithm is extended to incorporate a boundary confidence prior. Suppose that the gradient of a continuous surface $w(x,y)$ at (x,y) is the vector pointing toward the direction of largest increase on the surface as:

$$\nabla \hat{w}(x,y) = \frac{\partial w}{\partial x}\mathbf{i} + \frac{\partial w}{\partial y}\mathbf{j} \tag{4}$$

Any Cartesian x-y coordinate system can be chosen since it is easy to verify that the gradient magnitude and an edge orientation as:

$$\hat{\omega} = \|\nabla \hat{\omega}(x,y)\| = \left[\left(\frac{\partial w}{\partial x}\right)^2 + \left(\frac{\partial w}{\partial y}\right)^2\right]^{\frac{1}{2}} \tag{5}$$

$$\hat{\theta} = \arctan\left(\frac{\partial w}{\partial y} \Big/ \frac{\partial w}{\partial x}\right) \tag{6}$$

After finishing gradient estimation, every pixel in the image is associated with an edge magnitude $\hat{\omega}$ and an edge orientation $\hat{\theta}$. Let $\hat{\omega}_{[1]} < \cdots < \hat{\omega}_{[k]} < \hat{\omega}_{[k+1]} < \cdots < \hat{\omega}_{[N]}$ be the ordered set of distinct magnitudes values. Therefore, for any pixel, its edge magnitude $\hat{\omega}_{[k]}$ is replaced with the probability:

$$\delta_k = \mathrm{prob}[\hat{\omega} \leq \hat{\omega}_{[k]}] \tag{7}$$

Note that δ_k is the percentile of the cumulative gradient magnitude distribution. While the weight of each pixel i is described as:

$$\psi_i = 1 - [\alpha_i \varepsilon_i + (1 - \alpha_i)\zeta_i] \tag{8}$$

Where ε_i and ζ_i represent the estimated gradient magnitude and the confidence in the presence of an edge pattern, respectively. The nearer the pixels to an edge, the smaller its weight is. The above process can pre-segment the original image into many small regions whose edges are described well and whose color is consistent. In this paper, we define this region as super-pixel and use it to construct Graph Cut model. Compared with traditional single-pixel based model, our method can simplify the number of nodes and weighted edges of weighted graph topological structure and reduce the computation cost and memory consumption.

2.2. Character Description and Clustering

To extract the feature information of super-pixel, usual methods are to compute the feature average of all sample points of the region, which leads to the lack of spatial color correlation between pixels. Thus, in this paper we introduce a Gaussian Mixture Model clustering algorithm to describe the super-pixel. Denote a super-pixel i by $s_i = \{\mu_i, \Sigma_i\}$, where μ_i and Σ_i represent mean and variance of color feature of region i respectively. In Equation (1), to compute $E_1(x_i)$, the user can define the background seeds and the background seeds, the super-pixels of unknown regions can be clustered by Gaussian Mixture Model. The mean colors of the foreground and background clusters are denoted as $\{G_1^F, G_2^F, \cdots, G_M^F\}$ and $\{G_1^B, G_2^B, \cdots, G_N^B\}$ respectively, where M and N are the clusters of foreground and background respectively. Then, for each super-pixel, minimum distance from its color cluster G_i to foreground clusters can be expressed as:

$$D_i^F = \min\nolimits_{n\in[1,M]}^{i} \mathrm{dis}(G_i, G_n^F) \tag{9}$$

$$D_i^B = \min\nolimits_{n\in[1,N]}^{i} \mathrm{dis}(G_i, G_n^B) \tag{10}$$

Therefore $E_1(x_i)$ is defined as follows:

$$\begin{cases} E_1(x_i=1)=\infty & E_1(x_i=0)=0 \quad \forall i\in B \\ E_1(x_i=1)=0 & E_1(x_i=0)=\infty \quad \forall i\in F \end{cases} \tag{11}$$

$$E_1(x_i=1) = D_i^F (D_i^F + D_i^B)^{-1} \quad \forall i \in U \tag{12}$$

$$E_1(x_i=0) = D_i^B (D_i^F + D_i^B)^{-1} \quad \forall i \in U \tag{13}$$

Here, U is the uncertain (not labeled) super-pixel set. The third equation guarantees the super-pixels to have the label with similar colors to foreground or background. We define $E_2(x_i, x_j)$ as a function of the color gradient between two super-pixels i and j:

$$E_2(x_i,x_j) = |x_i - x_j| \cdot \exp\{-\phi \,\mathrm{dis}^2(G_m, G_n)\} \tag{14}$$

$$\phi = \left(|Z|^{-1} \sum_{m,n\in Z} \mathrm{dis}^2(G_m, G_n) \right)^{-\varepsilon} \tag{15}$$

$$\mathrm{dis}(G_m, G_n) = \frac{\sqrt{2}}{2}\sqrt{KLD(G_m\|G_n) + 2KLD(G_n\|G_m)} \tag{16}$$

$$KLD(G_m\|G_n) = \int G_m(x) \log\frac{G_m(x)}{G_n(x)}\mathrm{d}x \tag{17}$$

Where $KLD(\cdot)$ is abbreviation to Kullback Leibler Divergence, which is used to measure quantitatively the distance between Gaussian features. To perform parametric learning of mixture model for interactive priori knowledge, EM algorithm is usually better selection. EM algorithm is suitable for maximum likelihood based Graph Cut segmentation model. To overcome the problem of slow convergence speed of traditional EM algorithm, a Monte Carlo based EM (MCEM) acceleration algorithm is introduced. The main idea is to combine MCEM algorithm and Newton-Raphson algorithm and use Monte Carlo simulation to realize E-step of EM algorithm, which can not only preserve the advantage of EM algorithm but also effectively improve the convergence of EM algorithm. Finally, the full description of MCEM algorithm can be given. Firstly (E-step), use $p(\theta|\theta^{(i)}, Y)$ as the posterior distribution density function of θ with adding the data Z, let $Q(\theta|\theta^{(i)}, Y)$ be E-step integral, given sampling spots $\{z_1, Z_2, \cdots, z_m\}$ from $p(Z|\theta^{(i)}, Y)$ Computing:

$$\hat{Q}(\theta|\theta^{(i)}, Y) = \frac{1}{m}\sum_{j=1}^{m} \log p(\theta|z_j, Y) \tag{18}$$

Secondly (M-step), maximizing the function $\hat{Q}(\theta|\theta^{(i)}, Y)$ to work out $\theta_{EM}^{(i)}$ and satisfy (let $\Theta = \log p(\theta|Y)$):

$$\hat{Q}(\theta_{EM}^{(i)}|\theta^{(i)}, Y) = \max_{\theta} \hat{Q}(\theta|\theta^{(i)}, Y) \tag{19}$$

$$\theta^{(i+1)} = \theta^{(i)} + \left(-\left.\frac{\partial^2\Theta}{\partial\theta^2}\right|_{\theta^{(i)}} \right) \times \left[\int \frac{\partial\Theta}{\partial\theta} p(z|y, \theta^{(i)})\mathrm{d}z \Big|_{\theta^{(i)}} \right] (\theta_{EM}^{(i)} - \theta^{(i)}) \tag{20}$$

Thus arose an iteration: $\theta^{(i)} \to \theta^{(i+1)}$, and then perform the iteration operation for the above E-step and M-step until $\|\theta^{(i+1)} - \theta^{(i)}\|$ or $\|Q(\theta^{(i+1)}|\theta^i, Y) - Q(\theta^i|\theta^i, Y)\|$ approaches infinitesimal. By the Geweke Law of Large Numbers, we have:

$$\hat{Q}(\theta|\theta^{(i)}, Y) \to Q(\theta|\theta^{(i)}, Y) \tag{21}$$

3. Experimental Result

To demonstrate the performance of our proposed approach, we first test it on some public images. We also compare our algorithm to Graph Cut [12], Grab Cut [13], Lazy Snapping [15]. Our algorithm starts with following initial parameters: $\lambda = 50, \alpha = 0.2, \varepsilon = 0.95, M = N = 5$. The system is running on a P4-2GHz desktop with 1GB RAM. Figure 1 shows the pre-segmentation comparison of watershed appeared in Lazy Snapping [15] and mean shift used in our algorithm. The left column are the original test images, the 2nd and 3rd columns displayed the segmentation results by watershed and mean shift with a boundary confidence prior.

In comparison, the over-segmentation phenomenon of watershed is very serious, which necessarily leads to higher time complexity of sequent Graph Cut model. To compare, our method which uses mean shift incorporating a boundary confidence prior can effectively control the over-segmentation phenomenon and the number of regions pre-segmented is less than 1% of the watershed method.

The algorithm is also compared with *Graph Cut*, *Grab Cut* and *Lazy Snapping* and the results are as shown

Figure 1: Pre-segmentation comparison of watershed and mean shift with a boundary confidence prior algorithm.

in Figure 2. In Figure 2, the top row is the original test images with a quick object marking step: the red lines are drawn to indicate the foreground and the blue lines to indicate the background. The 2rd, 3th, 4th and 5th rows displayed the segmentation results by Graph Cut, Grab Cut, Lazy Snapping and Our method respectively. In comparison, the proposed method outperforms in complex scenes (the extraction of a pair of thin and long tentacles in "Butterfly", color similarity of foreground and background in "Fish", background complexity in "Starfish" and "Boy") and also gives better segmentation results compared with Graph Cut, Grab Cut and Lazy Snapping.

In Figure 3, we select another test image to further compare four different methods. In the test image, a curl of hair of a little girl is a particularly challenging object with the characteristic of thin/long. Segmentation of the special thin/long object such as human hair and animal antenna is currently difficult to do better. From the results, we can see that Graph Cut, Grab Cut and Lazy Snapping either detect a few background pixels as foreground or miss some important information of curl hair or hair braids. By the proposed algorithm, more clean and complete hair marked with red circle can be extracted successfully. The segmentation results show distinct improvement in comparison with previous methods.

In addition, comparison of average execution time of the four methods is provided in Figure 3. By comparison, we can see that the average time required of our method extracting foreground objects less than Graph Cut, Grab Cut and Lazy Snapping. For Lazy Snapping, in order to obtain a general satisfied result it need execute many interactions, and for any interaction operation it needs rerun the whole Graph Cut model and thus decreases its efficiency.

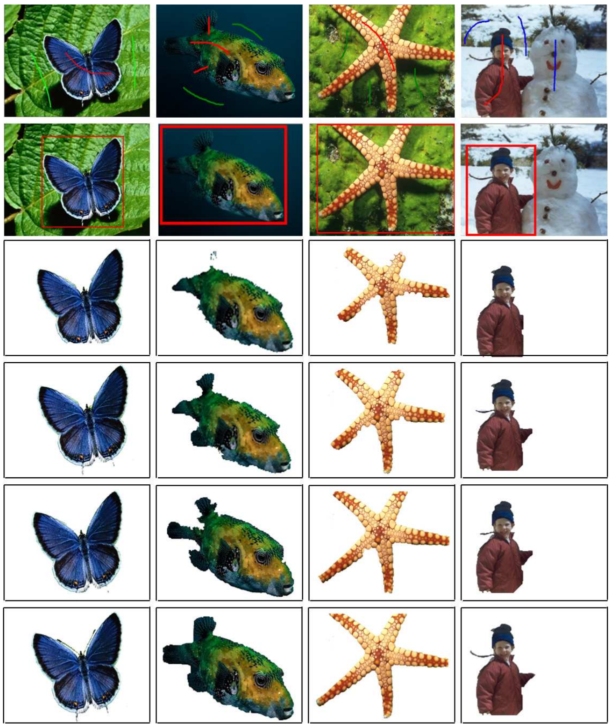


Figure 2: Some comparative results by the four methods.

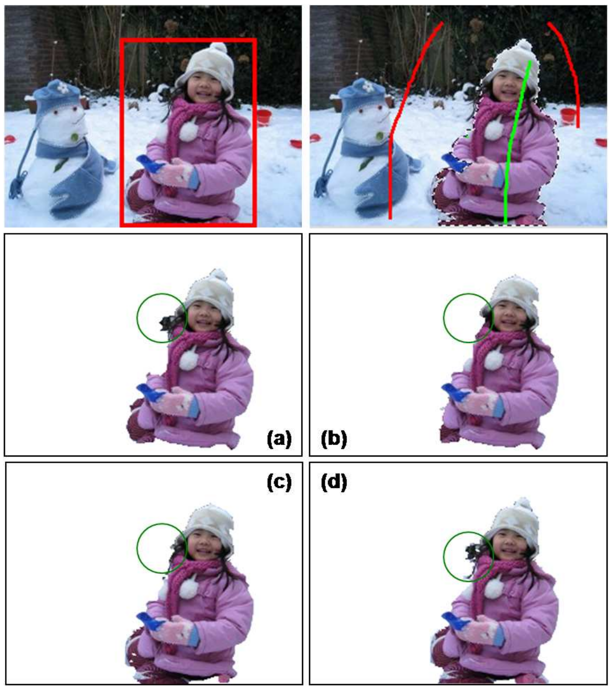


Figure 3: Segmentation Results of different methods on a testing image with complex scene including a curl of hair of a little girl. (a)Graph Cut, (b)Grab Cut, (c)Lazy Snapping, (d)Our proposed method.

4. Conclusion

In this paper, we propose a more effective interactive image matting method compared with Graph Cut, Grab

Cut and Lazy Snapping. Firstly, our method uses mean shift algorithm with a boundary confidence prior to efficiently pre-segment the original image into homogeneous regions (super-pixels) with precise boundary. Secondly, we introduce Gaussian Mixture Model clustering algorithm to describe and model the super-pixels. Finally, a Monte Carlo based EM acceleration algorithm is presented to perform parametric learning of mixture model for priori knowledge. The experimental results show that our algorithm can outperform in both matting quality and efficiency.

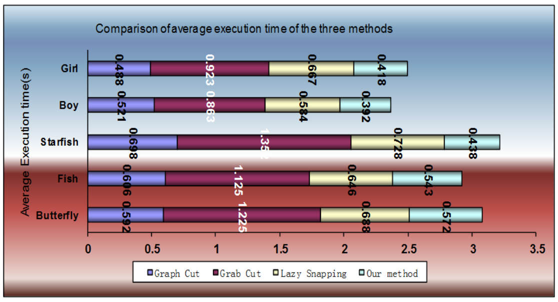


Figure 4: Comparison of average execution time of Graph Cut, Grab Cut, Lazy Snapping and the proposed method.

Acknowledgement

This study has been financed partially by the Projects of National Natural Science Foundation of China (Grant No. 50635030, 60932001, 61072031, 61002040), the National Basic Research (973) Program of China (Sub-grant 6 of Grant No. 2010CB732606) and the Knowledge Innovation Program of the Chinese Academy of Sciences, and was also supported by the grants of Introduced Innovative R&D Team of Guangdong Province: Robot and Intelligent Information Technology, the grants of Introduced Innovative R&D Team of Guangdong Province: Image-Guided therapy technology.

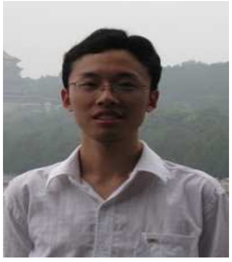
References

[1] Q. Zhu, Z. Zhang, Y. Xie, Applied Mathematics & Information Sciences **6**, 363-369 (2012).
[2] J. Wang and M. Cohen, Foundations and Trends in Computer Graphics and Vision **3**, (2007).
[3] A.R. Smith and J.F. Blinn, In proceedings of SIGGRAPH 259-268 (1996).
[4] A. Berman, P. Vlahos, A. Dadourian, U. S patent 134-345(2000).
[5] M. Ruzon and C. Tomasi, In proceedings of CVPR 18-25 (2000).
[6] P. Hillman, J. Hannah, D. Renshaw, In proceedings of CVPR 1063-1068 (2001).
[7] Y.Y. Chuang, B. Curless, D. Salesin, In proceedings of CVPR 264-271 (2001).
[8] J. Sun, J. Jia, C.K. Tang, H.Y. Shum, In proceedings of SIGGRAPH 315-321 (2004).
[9] J. Sun, Y. Li, S.B. Kang, H.Y. Shum, In proceedings of SIGGRAPH 259-268 (2006).
[10] J. Wang and M.F. Cohen, In proceedings of ICCV 936-943 (2005).
[11] A. Levin, D. Lischinskid, Y. Weiss, In proceedings of CVPR 61-68 (2006).
[12] Y. Boykov and M.P. Jolly, In proceedings of ICCV 105-112 (2001).
[13] C. Rother, A. Blake, V. Kolmogorov, In proceedings of SIGGRAPH 309-314 (2004).
[14] V. Vezhnevets and V. Konouchine, In proceedings of Graphicon 150-156 (2005).
[15] Y. Li, J. Sun, C.K. Tang, H.Y. Shum, In proceedings of SIGGRAPH 303-308 (2004).
[16] Y. Boykov, O. Veksler, R. Zabih, In proceedings of CVPR, 648C655, (1998).
[17] Y. Boykov, O. Veksler, R. Zabih, IEEE Trans. Pattern Anal. Mach. Intell **23**, 1222-1239 (2001).
[18] V. Kolmogorov and R. Zabih, IEEE Trans. Pattern Anal. Mach. Intell **26**, 147-159 (2004).
[19] H. Ishikawa and D. Geiger, In proceedings of ECCV 232-248 (1998).
[20] V. Kolmogorov and R. Zabih, In proceedings of ICCV 508-515 (2001).
[21] Y. Li, J. Sun, H.Y. Shum, ACM Trans. Graph **24**, (2005).
[22] C. Rother, V. Kolmogorov, A. Blake, ACM Trans. Graph **23**, (2004).
[23] F. Alizadeh and A. Goldberg, Technical Report STAN-CS-92-1410, Stanford University (1992).
[24] R.J. Anderson and J.C. Setubal, In proceedings of SPAA 168-177 (1992).
[25] Y. Boykov, O. Veksler, R. Zabih, IEEE Transactions on Pattern Analysis and Machine Intelligence **20**, 1222-1239 (2001).
[26] Y. Boykov and M.P. Jolly, In proceedings of ICCV **1**, 105-112 (2001).
[27] C. Rother, V. Kolmogorov, A. Blake, ACM Transactions on Graphics **23**, 309-314 (2004).
[28] A.K. Sinop and L. Grady, In proceedings of ICCV 1-8 (2007).
[29] X. Bai and G. Sapiro, In proceedings of CVPR 1-8 (2007).
[30] Q. Zhu, L. Shao, Q. Li, Y. Xie, In proceedings of MMM 125-133 (2013).

Qingsong Zhu received the BS and MS degrees in computer science from University of Science and Technology of China (USTC), Hefei, China. He is currently an assistant professor in Research Centre for Medical Robotics and Minimally Invasive Surgical Devices, Institute of Biomedical and Health Engineering, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. His current research interests focus on computer vision, statistical pattern recognition, machine learning and robotics. He is a member of the IEEE.

Guanzheng Liu received the M.S. degree in control theory and control engineering from Zhejiang University, Hangzhou, China. He received Ph.D degree at Chinese Academy of Sciences. His research is biomedical signal processing. He is now Biomedical Engineering Program, Sun Yat-sen University, Guangzhou, China, as a professor Assistant professor in Biomedical Engineering.

Zhanyong Mei received the M.S. degree from Chengdu University of Technology, Chengdu, China, in 2008. She is currently a PhD student at the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. Her research interest is biomedical signal processing and gait analysis.



Qi Li received the BS degree in 2010 from University of Science and Technology of China (USTC), Hefei, China. He is currently a postgraduate student in School of software engineering of USTC. He is also working as a visiting student in Shenzhen Institutes of Advanced Technology, Chine- se Academy of Sciences. His research interests include digital image processing.

Yaoqin Xie Professor, supervisor of the doctoral students and Overseas High-Caliber Personnel in Shenzhen, China. He received his B.Eng, M.Eng, and Ph.D. from Tsinghua University of China in 1995, 1998, and 2002, respectively, and won the excellent doctor degree dissertation of Tsinghua University. He joined Stanford University as a postdoctoral fellow from 2006 to 2008. Dr. Xie's research areas have been focused on image-guided surgery. He has published more than 90 papers. Dr. Xie was working in Peking University of China from 2002, and joined Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, on 2010. He is a committee member of Chinese Society of Molecular Imaging, committee member of Chinese Society of Medical Imaging Physics, editorial board member of Chinese Medical Imaging Technology, member of American Association of Physicist in Medicine (AAPM), Chinese Society of Biomedical Engineering. He is also the referee of IEEE Transactions on Information Technology in BioMedicine.

Lei Wang received his B.Eng in Information and Control Engineering and Ph.D in Biomedical Engineering, in 1995 and 2000, respectively, from Xi'an Jiaotong University, China. He was with University of Glasgow and Imperial College London during 2000-2008. He is now with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, as a full professor. Dr. Wang's research interests focus on Body Sensor Network. He has published more than 200 scientific papers, authored four book chapters and filed 60 patents.