A Method of Hand Contour Tracking based on GVF, Corner and Optical flow

Ke Du[1, a], Ying Shi[1,b] , Jie Chen[2,c], MingJun Sun[1], Jie Chen[1], ShuHai Quan[1]

[1] School of Automation, Wuhan University of Technology, Hubei 430070, China

[2]WuHan GUIDE INFRARED CO,WuHan,Hubei,430205,China

[a]dukework@163.com, [b]a_laly@163.com, [c]chenjie85930431@163.com

Keywords: GVF Snake, SUSAN Corner, Optical flow

Abstract. Since bare human hand is highly articulated and deformable, its tracking is more challenging than rigid object tracking. In this paper, an automatic approach is proposed to track hand contour on the basis of gradient vector flow (GVF Snake), corner detection and sparse optical flow. The skin-color model and motion information help to reduce the impact of local minimal under complex background. To lower the sensitivity of GVF Snake to initial contour, a curvature-difference-based corner detection is proposed to effectively detect enough corners on hand contour and sparse optical flow is applied to ensure a robust tracking of these corners even with violent gesture change. The experiments have shown a promising hand contour tracking using real image sequence under the usual lighting condition in lab.

Introduction

In the field of HCI and virtual environment, hand gesture is regarded as one of the most natural and unconstrained communication manner. Since bare human hand is highly articulated and deformable, its tracking is more challenging than rigid object tracking [1]. Usually, rectangular or ellipse is used to locate human hand region in traditional hand tracking[2,3], since the contour or fingertips of complicated hand gesture could not be marked, it is difficult to understand the hand gesture. For a better understanding, the hand silhouette or contour is always applied to hand gesture recognition.

Kass [4] firstly proposed Snakes, where the object boundary is described as a parameter curve and is determined by using an energy minimization; however, the boundary concavities cannot be captured by Snakes. Xu [5] proposed a GVF Snake by using a new external force to enlarge the capture range and enable converge to boundary concavities.

As far as the image with complicated background is concerned, GVF Snake still could not give a good object contour for its sensitivity to the local minimal and initial curve. Many researches focused on these issues [6]. For a single hand situation, the modified extra-energies-based algorithms and prediction-based algorithms appeared [7].

The former aims to decrease local minimal via adding the extra energies to traditional snake model [8-10]. Extra energies do help to reduce the effect of local minimal, but sometimes result in the disability of curve convergence to precise object boundary. And due to more computation cost for extra energies, these algorithms are time-consuming.

The latter is proposed to lower the sensitivity to initial contour via a dynamic prediction on the object contour [11-13]. Yang [14] proposed a GVF Snake combined with skin color region tracking. The skin color model could rapidly capture precise hand shapes only for smooth hand deformation from frame to frame but the predictions always fail for violent changes of hand shape [15].

In this paper, a method is proposed on the base of hand feature and the traditional energies, and the impact of local minimal on hand region segmentation is reduced by using Gaussian skin model and the corners on hand contour. And we lower the sensitivity to initial contour by fitting a curve of these corners as the initial contour in current frame for its similarity to the real hand contour. In addition, we track violent changes on hand shape via tracking the selected corners which better describe the motion feature in real-time tracking [16].

Cheng[17] proposed a GVF Snake method based on the Canny operator and thinning method to get a consecutive boundary. For a better segmentation performance, SUSAN corner detection could be used for a precise detection on the points of hand fingertips and their roots. However, these points are not enough for fitting a good hand contour. Cheng[18] and Malagi [19] proposed a modified GVF snake method on the base of SUSAN corner detection. Both the corners on hand contour with high contour curvature and SUSAN corners were used to the hand contour fitting. In this paper, the corners on hand contour will be screened before the tracking according to the sign change on forward and backward curvature difference for a high efficiency.

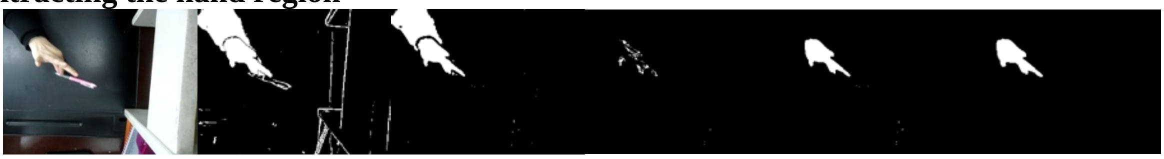
To robustly track hand contour, the optical flow was introduced to track the corners on hand contour between frames [20, 21]. For few corners are concerned, the sparse optical flow is applied within the motion sub region to reduce the computation [22]. And these corners are screened by the curvature-difference-based corner detection in successive frame. And then GVF Snake adopts the fitting curve of these corners as initial contour to detect the hand contour.

The paper is organized as follows: Firstly, hand motion region is located by using skin color model within the motion sub region segmentation using background subtraction, and then its contour is captured by GVF Snake model. Secondly, the corners on this contour are detected by using curvature-difference-based corner detection, and tracked by using sparse optical flow. At last, the corners are interpolated in next frame to form a closed curve as the initial curve of GVF Snake model, and then the hand contour is captured. By repeating these steps, we can obtain a robust and fast tracking on the hand contours in video frame.

Initializing the hand contour

Color image segmentation or background subtraction often fails to extract the hand region due to the complicated background and varying illumination. Usually, Gaussian mixture model is applied to deal with these problems. With the consideration on the color feature of human skin, Gaussian mixture skin model is combined with background subtraction to detect the hand region [32, 33]. Our work addresses this problem by using both Gaussian mixture skin model and a flood fill method to mark a complete hand region. Then the contour of maximum region with cross-sectional area greater than a given threshold is supposed to be the initial one, and let it converge to the accurate hand contour via using GVF Snake.

Extracting the hand region



(a) (b) (c) (d) (e) (f)

Fig. 1. Hand region detection in the 500th frame: (a)original color image; (b)difference image; (c)difference image after removing noise; (d)skin detection image; (e)flood filled image; (f) hand region image

To meet speed requirement, the first frame of input video is taken as the background, and motion area is obtained by subtracting the background from next frames (see Fig. 1b), and it can be seen that there is a lot of noise and shadows. The morphological erosion operator is adopted to reduce noise, as shown in Fig. 1c. To eliminate the interruptions of shadows and non-skin moving objects, Gaussian mixture skin model defined by color statistic is applied [23,24].

$$P(x \mid skin)=\sum_{i=1}^{N} w_i \frac{1}{(2\pi)^{\frac{3}{2}}\left|\Sigma_i\right|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu_i)^{\mathrm{T}}\Sigma_i^{-1}(x-\mu_i)} \qquad \textbf{(2-1)}$$

where N is the number of mixture components defined as 16, x is an RGB color vector, and the contribution of i^{th} Gaussian is determined by a scalar weight ω_i, mean vector μ_i, and diagonal covariance matrix Σ_i. Two separate mixture models for skin and non-skin classes are trained, and a particular RGB value is labeled as skin color if

$$\frac{\mathrm{P}(x \mid skin)}{\mathrm{P}(x \mid \neg skin)} > \theta \qquad \textbf{(2-2)}$$

where $0 \le \theta \le 1$. In order to obtain a good skin detection for most images, let θ=0.8. This skin color model sometimes fails for the existence of holes in hand region and noise in small region when the skin region is too bright or dark (see Fig. 1d for example). Thus, a flood fill method is proposed to stuff the missing skin area (see Fig. 1e). In order to reduce the effect of varying luminance, the $\mathrm{Y'C_b'C_r'}$ color space is chosen for its skin-tone independence on luminance. The flood fill algorithm is as follows:

Convert RGB image $I_{RGB}(x,y)$ to $\mathrm{YC_bC_r}$ image $I_{YC_bC_r}(x,y)$ and nonlinear transform $\mathrm{YC_bC_r}$ image $I_{YC_bC_r}(x,y)$ to $\mathrm{Y'C_b'C_r'}$ image $I_{Y'C_b'C_r'}(x,y)$. For points (x,y) of nonzero value in $I_{RGB}(x,y)$ and points detected by Gaussian mixture skin model in $I_{seed}(i)$, $i = 1, 2, 3 \cdots n$, if

$$\begin{cases} I(C_b')(x,y)+T_1 \ge I_{seed}(C_b')(i), & \text{and} \\ I(C_b')(x,y)\text{-}T_2 \le I_{seed}(\mathrm{C}_b')(i), & \text{and} \\ I(C_r')(x,y)+T_1 \ge I_{seed}(\mathrm{C}_r')(i), & \text{and} \\ I(C_r')(x,y)\text{-}T_2 \le I_{seed}(\mathrm{C}_r')(i), & \end{cases} \qquad \textbf{(2-3)}$$

then mark the points $I_{RGB}(x,y)$ as skin point. T1, T2 are the minimum and maximum thresholds, respectively. $I(C_b')$, $I(C_r')$ are values of chromatic component in the $\mathrm{Y'C_b'C_r'}$ space, respectively.

After the skin area detection, the cross-sectional area sizes are counted for all connected domain. If the size of the greatest is greater than a given threshold, the corresponding region is extracted, otherwise there is no hand image exiting (see Fig.1f).

Capturing the initial hand contour

After the preceding process, the relatively complete hand region is obtained, and now we need to capture the hand contour for the convenience of hand gesture tracking. The GVF Snake is a kind of curve $X(s) = \left[x(s) \quad y(s)\right]$, $s \in [0,1]$ which moves through the spatial domain of an image to minimize the energy functional:

$$E=\int_0^1 \frac{1}{2}[\alpha\left|X'(s)\right|^2 + \beta\left|X''(s)\right|^2] + E_{ext}(X(s))ds \qquad \textbf{(2-4)}$$

$$E_{int}=\int_0^1 \frac{1}{2}[\alpha\left|X'(s)\right|^2 + \beta\left|X''(s)\right|^2] \qquad \textbf{(2-5)}$$

where α and β are weighting parameters which control the snake's tension and rigidity respectively. The first term is the internal force, which controls the curve changes, while the second term E_{ext} is the external force, which pulls the curve to desired features. Different E_{ext} can be constructed in different models. The minimized E of a snake must satisfy the Euler equation

$$\alpha X''(s) - \beta X''''(s) - \nabla E_{ext} = 0 \qquad \textbf{(2-6)}$$

A new static external force field V takes the place of the traditional external force, $V(x,y)=[u(x,y),v(\mathrm{x,y})]$, and is deduced by minimizing the energy functional:

$$E=\iint \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + \left|\nabla f\right|^2 \left|\mathbf{v} - \nabla f\right|^2 dxdy \qquad \textbf{(2-7)}$$

where f is the edge map of gray level image, and ∇f is the gradient of edge map, the parameter μ is a regularization parameter. Numerical solution of the GVF field is as follows:

$$u_t(x,y,t) = \mu\nabla^2 u(x,y,t) - (u(x,y,t) - f_x(x,y))(f_x^2(x,y) + f_y^2(x,y)) \qquad \textbf{(2-8a)}$$

$$v_t(x,y,t) = \mu\nabla^2 v(x,y,t) - (v(x,y,t) - f_y(x,y))(f_x^2(x,y) + f_y^2(x,y)) \qquad \textbf{(2-8b)}$$

In order to lower the sensitivity to initial contour and reduce calculation cost, we use the hand region image (see Fig. 2b) as the initial contour of GVF Snake, and then only the GVF field in an exterior rectangular region(see Fig. 2d), which contains the hand image, needs to be calculated, as shown in Fig. 2d and 2e.

(a) (b) (c) (d) (e)

Fig. 2. Hand contour initialization：(a)Hand contour initialization; (b)Hand region image; (b) hand region in an exterior rectangular region; (d)GVF external forces in u direction; (e)GVF external forces in v direction

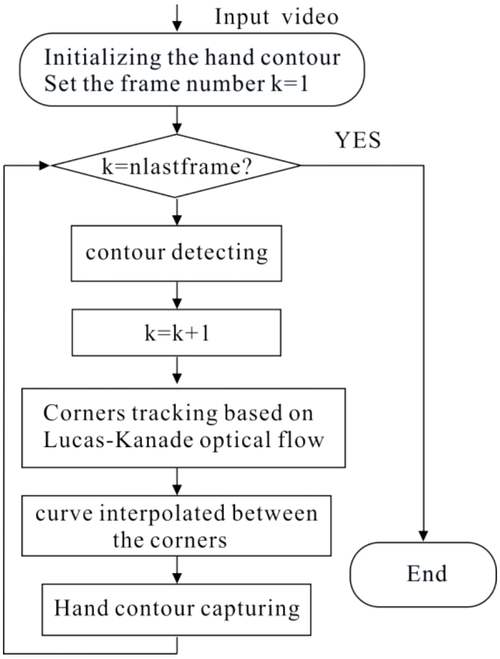


Fig.3. Framework of hand contour tracking

TRACKING THE HAND CONTOUR

For the subsequent hand gesture recognition, the tracking of hand contour is necessary. On the base of the initial hand contour captured above, we proposed an algorithm to robustly track hand contour, which can also be applied to track the contour of other objects. Hand contour tracking in this paper includes four steps, as follows (the framework shows in Fig.3):

Step 1: Detect corners on the hand contour in previous frame.

Step 2: Track corners on the base of Lucas-Kanade (LK) optical flow in current frame.

Step 3: Initialize GVF Snake with the curve fitted by the corners.

Step 4: Capture hand contour. And then, go to the step 1 for processing next frame.

In this part, the detail of above-mentioned method is focused.

A. Corner detection

Corner are good features in real-time tracking because they are close to the original object and can reliably match to the object points in the next frame. We need more points with high curvature to insure the accurate and rapid capture of the GVF Snake on those sharp areas. In this paper, curvature-difference-based corner detection is proposed to obtain enough corners. Calculate curvature k at every point of the contour obtained by GVF Snake, and let $k(i)$ be the curvature of the i^{st} point on the contour sequence:

$$k(i)=\frac{\dot{x}(i)\ddot{y}(i)-\ddot{x}(i)\dot{y}(i)}{(\dot{x}(i)^2+\dot{y}(i)^2)^{1.5}} \quad \textbf{(3-1)}$$

where $\dot{x}(i)$, $\ddot{x}(i)$ denote the first and second derivatives of $x(i)$, respectively; $\dot{y}(i)$, $\ddot{y}(i)$ denote the first and second derivatives of $y(i)$, respectively.

The following is the detail of curvature-difference-based corner detection.

Step 1: After hand region detection, apply Canny edge detection to obtain a binary edge-map.

Step 2: Check corners at the edge using SUSAN approach and save to Sequence list S in order, the total point number is denoted as N.

Step 3: Insert the Nth point to the head of the list S and the first point to the end of the list S.

Step 4: Calculate $k(i)$ at the every point in the list S, where 1<i<N+1and initialize i to 2.

Step 5: If i=N+1, end the program, else i++.

Step 6: If $(k(i-1)-k(i)\times(k(i)-k(i+1))<0$, go to step 7, else jump to step 5.

Step 7: Save the point $S(i)$ as the corner points to Sequence list P, and loop execute step 5.

B. Optical flow based corner tracking

Optical flow is a good method for tracking the points between frames, especially in the case of violent changes. But this method calculates all pixels of the whole image. Lucas and Kanade[31] proposed a sparse optical flow method which is realized by processing the pixels within a small neighborhood when tracking a point. In this paper, the LK method is used to fast track the corners on hand contour. LK optical flow is based on the three hypotheses, i.e. brightness constancy, temporal persistence and spatial coherence. The details of the method can be found in the work [31].

C. GVF Snake on the base of corners and optical flow

As we can see, there are many convex areas and concave areas on the hand contour. Therefore, GVF Snake might take much iteration to converge to those areas because of the disturbance from internal energy of snake and the initial position of contour. Usually, it is not helpful to regard the hand contour in previous frame as the initial one in current frame. To address this problem, we proposed a contour initialization method which interpolates the corners tracked in current frame to form an initial contour. The details are as follows:

Step 1: Detect $P(t-1)=\{P_{1,t-1},P_{2,t-1},\cdots,P_{i,t-1},\cdots P_{L,t-1}\}$ as the corners on hand contour at time t-1, $0<i<L$ and L is the total number of the corners.

Step 2: Track $P(t-1)$ by LK method to obtain $P(t)=\{P_{1,t},P_{2,t},\cdots,P_{i,t},\cdots P_{L,t}\}$, $0<i<L$.

Step 3: Delete abnormal points, denote point $P_{i,t}$ with the coordinate $(x_{i,t},y_{i,t})$. If $d=\left|x_{i,t}-x_{i-1,t}\right|+\left|y_{i,t}-y_{i-1,t}\right|+\left|x_{i+1,t}-x_{i,t}\right|+\left|y_{i+1,t}-y_{i,t}\right|>50$, Delete $P_{i,t}$ from $P(t)$.

Step 4: Linearly fit a closed curve C(t) by the corner points to insure the distance of the adjacent two points less than two-pixel wide.

Step 5: Take C(t) as the initial curve of GVF Snake at the time t, get the hand contour. Repeat step 1 to 5, hand contour is tracked.

EXPERIMENTS

In this section, we present the result of hand contour tracking. Two video sequences with the size of 320×240 under the usual lighting condition in lab are recorded to test the efficiency of our hand contour tracking algorithm. In two experiments, we take the first frame as the background for quick motion detection. The two experiments are performed with the parameter value, α=0.1 and β=0.1, which determine the tension and rigidity of the snake respectively.

The first experiment is performed on the sequence of video1, which contains a fast moving hand with nearly no hand shape change. Figs. 4a, 4b and 4c show the result of the traditional contour tracking method, which takes the previous contour as the initial contour in the current frame. In contrast, Figs. 4d, 4e and 4f show the result of the tracking method of GVF Snake on the base of corner and sparse optical flow we proposed. In the experiment, the red curves present hand contours and the blue points are the traced corners from the previous frame.

(a) (b) (c) (d) (e) (f)

Fig.4. Tracking on the video1:(a) (b) (c) the result of the traditional contour tracking method;(d)(e)(f) the result of our approach

In the video1, the hand does not appear until 48th frame, and it completely moves into the camera view in the 70th frame. As shown in Figs. 4a and 4d, the algorithm proposed in section 2 automatically outlines hand contour as the initial one for tracking. At the 75th frame, Fig. 4b shows the fingertip of index finger with a complex convex-concave shape cannot be captured by the traditional contour tracking, while it can be tracked precisely in Fig. 4e. After that, the hand moves fast downward. Fig. 4c shows the index finger information is lost because of the local minimal of GVF Snake. Fig. 4f illustrates an accurate result of tracking hand contours for tracking corners from the previous frame.

The second experiment is performed on the sequence of video 2, which contains the skin-like color object and a violent hand change. In Fig. 5, the result hand contour is marked by red curve; the blue points are the traced corners from the previous frame; the yellow points are the corners detected using curvature-difference-based corner detection.

(a)409th frame (b)426th frame (c)449th frame (d)484th frame (e)504th frame (f)544th frame

Fig. 5. Hand contour tracking on the video2

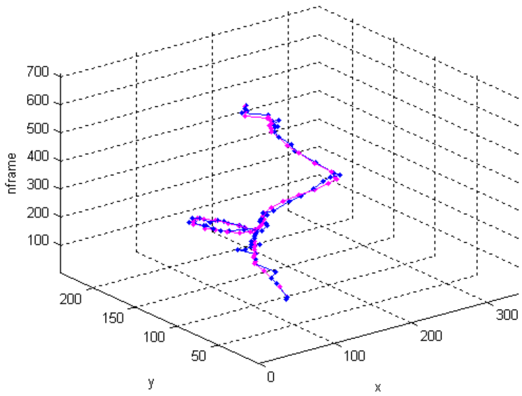


Fig. 6. The 3D figure of mass center trace of video2

The initial contour is firstly detected in the 409th frame based on the cross-sectional area size, as shown in Fig. 5a. Along with hand moving into the video, the whole hand contour can be precisely captured with the accurate contour of complex convex-concave shape in Fig. 5b. When the hand gesture changes smoothly, Fig. 5c shows we still can track them well. And then the hand begins to move towards the pen in skin-like color, some local minimal come forth. Because the initial contour in the current frame is a closed curve fitted by the tracked corners, the hand contour can be separated from the pen, as shown in Fig. 6d. Subsequently, violent gesture changes occur in the 504th and 544th frames, by using our method, the hand contours can be robustly tracked as Fig. 5e and 5f illustrate.

Fig. 6 shows two 3D traces of mass center for video2 from frame 1 to frame 710, where x, y and nframe represent x coordinate, y coordinate and the number of the frame, respectively. One trace is the real mass center trace of video1, manually labeled in magenta. The other one is the mass center trace calculated by the tracking method in this paper, presented in blue. As can be seen, the contour of hand can be robust tracked over 710 frames, and the results of our tracking method are in good agreement with the real mass center trace.

CONCLUSION

In this paper, we have proposed a novel automatic hand contour tracking approach based on GVF Snake without the complex extra energies for high efficiency. The method consists of skin-color GMM model based the flood fill algorithm, curvature-difference-based corner detection and sparse optical flow. Our approach mainly addresses two problems, one is about the sensitivity of GVF Snake to local minimal, and the other one is the sensitivity to initial curve. To reduce the impact of local minimal, we segment hand region by using motion information and skin-color GMM model. For its sensitivity to initial contour, we combine GVF Snake with optical flow. The experiments show a robust tracking of violent change hand gesture and the sensitivity to initial contour is lowered. Both experimental results indicate that this algorithm yields accurate and robust hand contour tracking and the system is fully automatic and bootstraps itself. It works well by the real-time tracking of the hand contour.

References

[1] RT. Collins, “Mean-shift blob tracking through scale space,” Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on; 2003. pp. II-234-240 vol.232.

[2] S. Bilal, R. Akmeliawati, MJE. Salami, AA. Shafie, EM. Bouhabba, “A hybrid method using haar-like and skin-color algorithm for hand posture detection, recognition and tracking,” Mechatronics and Automation (ICMA), 2010 International Conference on; 2010. pp. 934-939.

[3] P. Zhigeng, L. Yang, Z. Mingmin, S. Chao, T. Xing, et al, “A real-time multi-cue hand tracking algorithm based on computer vision,” Virtual Reality Conference (VR), 2010 IEEE; 2010. pp. 219-222.

[4] M. Kass, A. Witkin, D. Terzopoulos, “Snakes: active contour models,” Proceedings - First International Conference on Computer Vision. London, Engl: IEEE; 1987. pp. 259-268.

[5] X. Chenyang, JL. Prince, “Snakes, shapes, and gradient vector flow,” Image Processing, IEEE Transactions on 1998,7:359-369.

[6] Z. Zhou, J. Wu, W. Yang, et al, “Face contour tracking based on mean shift and GVF Snake,” Information Technology Journal, 2013, 12(10): 1884.

[7] F. Zhang, X. Zhang, K. Cao, et al, “Contour extraction of gait recognition based on improved GVF Snake model,” Computers & Electrical Engineering, 2012, 38(4): 882-890.

[8] Q. Li, Y. Luo, D. Xiao, “Video object contour tracking using improved dual-front active contour,” Computational Intelligence. Springer Berlin Heidelberg, 2006: 855-865.

[9] K. Takaya, “Tracking a video object with the active contour (snake) predicted by the optical flow,” 2008 Canadian Conference on Electrical and Computer Engineering, Vols 1-4; 2008. pp. 356-359.

[10] R . Wang, Y. Wang, J. Zhou, et al, “Active contours with neighborhood-extending and noise-smoothing generalized gradient vector flow external force,” Information Technology and Applications (ITA), 2013 International Conference on. IEEE, 2013.

[11] M. Niethammer, A. Tannenbaum, “Dynamic geodesic snakes for visual tracking,” Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on. IEEE, 2004, 1: I-660-I-667 Vol. 1.

[12] Y. Rathi, N. Vaswani, A. Tannenbaum, et al, “Particle filtering for geometric active contours with application to tracking moving and deforming objects,” Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. IEEE, 2005, 2: 2-9.

[13] Z. Zhou, J. Wu, D. Wu, “Human Body Contour Tracking Based on GVF-snake and DESO*,” Journal of Computational Information Systems, 2013, 9(18): 7525-7532.

[14] Y. Yang, Z. Wang, M. Zhang, et al, “Skin Region Tracking using Hybrid Color Model and Gradient Vector Flow,” Machine Vision and Human-Machine Interface (MVHI), 2010 International Conference on. IEEE, 2010: 243-246.

[15] S. Zhao, X. Song, W. Tan, et al, “A novel approach to hand gesture contour detection based on GVF Snake model and skin color elliptical model,” Computer Application and System Modeling (ICCASM), 2010 International Conference on. IEEE, 2010, 5: V5-381-V5-384.

[16] J. Shi, C. Tomasi, “Good features to track,” Computer Vision and Pattern Recognition, 1994. Proceedings CVPR'94., 1994 IEEE Computer Society Conference on. IEEE, 1994: 593-600.

[17] J. Cheng, R. Xue, W. Lu, et al, “Segmentation of medical images with Canny operator and GVF snake model,” Intelligent Control and Automation, 2008. WCICA 2008. 7th World Congress on. IEEE, 2008: 1777-1780.

[18] J. Cheng, C. Liu, “Image segmentation with GVF Snake and corner detection,” Computer Science and Software Engineering, 2008 International Conference on. IEEE, 2008, 1: 1017-1020.

[19] VP . Malagi, GG. SivaSankari, “Object recognition based on GVF and SUSAN in wireless sensor network,” Signal and Image Processing (ICSIP), 2010 International Conference on. IEEE

[20] E. Jayabalan, A. Krishnan, “Object Detection and Tracking in Videos Using Snake and Optical Flow Approach,” Computer Networks and Information Technologies. Springer Berlin Heidelberg, 2011: 299-301.

[21] G. Qiaoyun, X. Xuemei, A. Li, Q. Mo, “Computation of real-time optical flow based on corner features,” Intelligence Information Processing and Trusted Computing (IPTC), 2010 International Symposium on; 2010. pp. 345-348.

[22] BD. Lucas, T. Kanade, “Iterative image registration technique with an application to stereo vision,” Proceedings of the 7th International Joint Conference on Artificial Intelligence. Vancouver, BC, Can; 1981. pp. 674-679.

[23] HL. Ribeiro, A. Gonzaga, “Hand Image Segmentation in Video Sequence by GMM: a comparative analysis,” Computer Graphics and Image Processing, 2006. SIBGRAPI'06. 19th Brazilian Symposium on. IEEE, 2006: 357-364.

[24] MJ. Jones, JM. Rehg, “Statistical color models with application to skin detection.” International Journal of Computer Vision 2002,46:81-96.

[25] SM. Smith, JM. Brady, “SUSAN—A new approach to low level image processing,” International journal of computer vision, 1997, 23(1): 45-78.