

Real-Time Object Tracking using Gradient Vector Flow

Abstract. In this paper an object tracking system with utilizing optical flow technique, and Gradient Vector Flow (GVF) active contours is presented. Optical flow technique is less sensitive to background structure and does not need to build a model for the background of image so it would need less time to process the image. GVF active snakes have good precision for image segmentation. However, due to the high computational cost, they are not usually applicable. Since precision is one of the important factors in the image segmentation, several methods have been developed to overcome the computational speed. In this paper, we, first, recognize the moving object. Then, the object frame with some pixels surrounding to it, was created. Then, this new frame is sent to the GVF filed calculation procedure. Contour initialization is obtained based on the selected pixels. This approach increases the calculation speed, and therefore makes it possible to use the contour for the tracking. The system was built, and tested with a microcomputer. The results show a speed of 10 to 12 frames per second which is considerably suitable for object tracking approaches.

Streszczenie. W artykule przedstawiono system śledzenia obiektu z wykorzystaniem techniki Optic Flow oraz Gradiend Vector Flow. Wykrywanie ruchomego obiektu stanowi pierwszy etap działania, następnie ramka zawierająca obiekt przesyłana jest do algorytmu GVF, gdzie określany jest zarys obiektu. Dzięki temu podejściu możliwe jest wykorzystanie, wymagającego obliczeniowo GVF w śledzeniu obiektów. Przedstawiono wyniki eksperymentalne. (Wykorzystanie metody GVF w śledzeniu obiektów w czasie rzeczywistym)

Keywords: Object tracking, Active contour, GVF, Optical flow.

Słowa kluczowe: Śledzenie obiektów, aktywny kontur, GVF, Optical Flow.

Introduction

Object tracking is one of the fundamental tasks in the image processing techniques. In the past two decades many researches has been conducted in this field and many methods and algorithms have been introduced. These methods can be summarized into three major groups.

First method is point tracking. In this method objects are represented as points in consecutive frames and these points are related to each other based on previous object situation which can be in form of position and movement.

This method needs an external mechanism for object recognition in each frame. For a single object tracking, Kalman and particle filters [1], and for multi object tracking Joint Probabilistic Data Association Filter (JPDAF) and Multi Hypothesis Tracking (MHT) are some of methods that are widely used [2]. The second method named kernel tracking. Kernel mostly considers the shape and object appearances. For example, kernel can be rectangular or elliptical shape with a defined histogram. Objects are tracked in the frames with estimation of their kernel trajectory. These trajectories are defined in form of translation, rotation and scaling. The last group is silhouette tracking which tracking is accomplished by estimating the object boundaries in each frame. Silhouette tracking uses the encoded information inside the object boundaries. This information can be in form of appearance density and shape models which are mostly edge maps. If the shape model in known Silhouette of object can be tracked using the image matching or contour evolution. This method is used when articulated objects such as human body, hand, face are to be tracked.

In the contour tracking method, first an initial contour is selected, and then this contour is evolved in each frame until it reaches to the position of object in current frame. In this method, the object motion must be in a manner that the object boundary in the current frame overlaps with the previous and following frames. There are two methods that are commonly used to evolve the contour:

- using state space
- minimizing the energy function

In this paper, the third method, i.e. the contour tracking was considered. In the following, a brief background is presented.

Tissainayagam et al. at [3] tried to present a tracking system utilizing the MHT, segmentation with contour and point tracking. In the first part of this work, the image edges

are extracted using the MHT method, and then the contours are evaluated from these edges. The extracted contours are then mixed together for recognizing the objects. In the second part of the work, tracking is obtained from the first frame using MHT and some desirable key feature points extracted from edges.

For validation the proposed method in each frame, the extracted key feature points are compared to the segmented image. If an acceptable numbers of points are on the boundary or inside the object, then the tracking is considered to be valid.

Cheng et. al. at [4] proposed the contour tracking for non-rigid objects with the complex backgrounds images. They used Kalman filter with some changes which has resulted in a new filter called "structural Kalman filter". This new filter takes into account the relation between the different parts of the moving object. This is done to overcome the measurement uncertainties when the object is overlapped by some the object in the scene. In this method, the estimation of the next object position is also achieved.

Koschan et. al. at [5] used the active shape models for the non-rigid object tracking and changed the model so that it can be used for color images.

Ning et. al. at [6] introduce a tracking system which uses a model based on kinematics of human motion and uses image matching and Hierarchical searching methods was presented. They also used a divide and conquer method to reduce the search space which utilizes a skeleton like model of human body. Then, an algorithm is used to update the joint angels of the model. For error calculations, a pose estimation function using boundary and region information is presented. As an initial condition, the estimated first frame is compared with 6 pre-defined poses and the one which lies with the least error, was selected as the initial condition.

Kim et. al. at [7] tried to solve the active contour tracking short comings such as problems with intensity changes, object edge deformation, existence of other edges (complex back ground), and the edge noises in image with a system based on active shape models. Since active shape model is used, counters relations in different frames has been improved in comparison with the traditional active contour model.

Razi and Azadi [8] at presented a novel algorithm for moving object detection which is based on GVF active contours. The method proposed in their work improves the

GVF contour so the contour reaches the image edges in some concavities that the original GVF could not reach. This paper implements this method as a base for tracking system.

In the next section, the optical flow method is presented.

Optical Flow

The main objective in optical flow method is to estimate velocity of pixels using temporal intensity changes in the sequence of images [9,10].

One of the common assumptions for starting the work is that the intensity just moves from in successive frames:

$$(1) \quad I(\vec{x}, t) = I(\vec{x} + \vec{u}, t + 1)$$

where $I(\vec{x}, t)$ is image intensity as a function of position vector $\vec{x} = (x, y)^T$ and time $\vec{u} = (u_1, u_2)$. is the 2D velocity vector.

It should be noted that the above assumption is rarely satisfied and the real assumption is that the radiance of object surface is constant in successive frames (for example there are no mirror surfaces in the scene, the light source is a point which is far away so movements of the source does not have any effects on scene. Also the object is not rotating and there are no another reflections in scene). Although these conditions seem to be unreal, the equation works fine in practice. A

Now let's assume that the intensity equation can be estimated with Taylor series as:

$$(2) \quad I(\vec{x} + \vec{u}, t + 1) = I(\vec{x}, t) + \vec{u} \cdot \nabla I(\vec{x}, t) + I_t(\vec{x}, t) + H.O.T$$

where $\nabla I \equiv (I_x, I_y)$ and I_t are spatial and temporal derivatives of image intensity I and H.O.T. stands for the higher order terms of Taylor series.

If we neglect the higher order terms and substitute equation 2 into equation 1, we obtain:

$$(3) \quad \nabla I(\vec{x}, t) + I_t(\vec{x}, t) = 0$$

The above equation relates the velocity to the time-space derivatives and is commonly known as gradient constraint.

The resulted velocity vector \vec{u} is known as image velocity or the optical flow. If only two frames are accessible or temporal derivative of image intensity cannot be calculated, a new gradient constraint can be achieved by substituting I_t with:

$$(4) \quad \delta I_t(\vec{x}, t) \equiv I_t(\vec{x}, t + 1) - I_t(\vec{x}, t)$$

It is clear that equation 1 is not sufficient for estimating the velocities since two parameters of the speed are unknown. Gradient constraint just limits the velocity vector to a family of lines normal to ∇I with normal distance of $I_t / \|\nabla I\|$ from the base. Since only normal velocity can be

calculated from equation (3), and there is no way to calculate tangent velocity, the complete velocity vector cannot be calculated (aperture problem). One way to overcome this problem is to assume that the velocity is the same in the neighbor pixels but there is a possibility that no velocity satisfy these constraints. Therefore, we would like to minimize the least square means of errors [11]:

$$(5) \quad E(\vec{u}) = \sum_x g(\vec{x}) [\vec{u} \cdot \nabla I(\vec{x}, t) + I_t(\vec{x}, t)]^2$$

where $g(\vec{x})$ is a weighting factor which is commonly assumed to have a Gaussian distribution. The velocity

vector which minimizes the equation is assumed to be estimated velocity of image.

To calculate the minimum of the equation (5), the derivative of the equation is set to be zero. Thus we have:

$$(6-1) \quad \frac{\partial E}{\partial u_1} = \sum_x g(\vec{x}) [\vec{u}_1 \cdot I_x^2 + u_2 I_x I_y + u_2 I_x I_t + I_x I_t] = 0$$

$$(6-2) \quad \frac{\partial E}{\partial u_2} = \sum_x g(\vec{x}) [\vec{u}_2 \cdot I_y^2 + u_1 I_x I_y + I_x I_t] = 0$$

Above equations can be written in matrix form:

$$(7) \quad M \vec{u} = b$$

where the M and b are matrices as:

$$(8) \quad M = \begin{pmatrix} \sum_x g I_x^2 & \sum_x g I_x I_y \\ \sum_x g I_x I_y & \sum_x g I_y^2 \end{pmatrix}, \vec{b} = - \begin{pmatrix} \sum_x g I_x I_t \\ \sum_x g I_x I_t \end{pmatrix}$$

M has a rank of 2 and estimated velocity is calculated from the equation below:

$$(9) \quad \hat{u} = M^{-1} b$$

Using the resulted estimation, we can set a constraint to find the moving object:

$$(10) \quad |u_1| + |u_2| < \tau$$

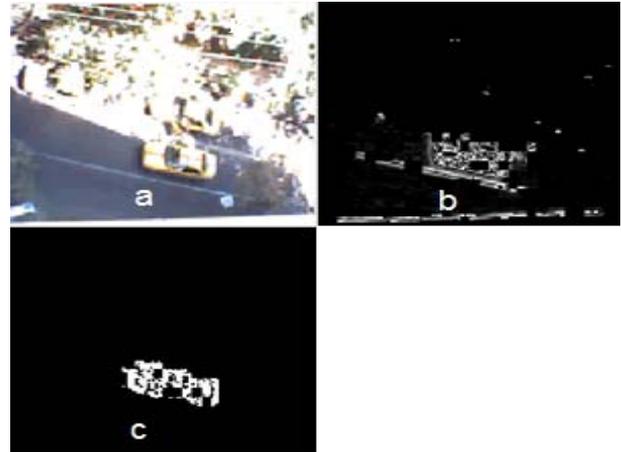


Fig.1. a) original image, b) extracted edges which edges from previous frame is subtracted from it, the moving object after threshold.

Pixels which satisfy the above constraint are considered to be the moving object and are subtracted from the edges of previous frame. Then the image is converted to a binary image using a threshold value.

Gradient Vector Flow (GVF) Active Contour

Kass et. al. [12] introduced a new term called "Active Snake". An active snake (contour) is a curve like:

$$(11) \quad X(s) = [x(s), y(s)], s \in [0..1]$$

Which tries to minimize the following energy function?

$$(12) \quad E = \int_0^1 \left[\frac{1}{2} (\alpha |x'(s)|^2 + \beta |x''(s)|^2 + E_{ext}(x(s))) \right] ds$$

where α and β are curvature and hardness factors respectively. x' , and x'' are the first and second derivatives for $X(s)$.

External energy term (E_{ext}) is defined so that in the desired positions such as object boundaries, has the minimum value. E_{ext} is commonly selected like this:

$$(13) \quad E_{ext}^{(1)}(x, y) = -|\nabla I(x, y)|^2$$

$$(14) \quad E_{ext}^{(2)}(x, y) = -|\nabla(G_{\sigma}(x, y) * I(x, y))|^2$$

To extend the capture range of snake α must have a high value, though this makes the edge of image blurry.

The snake which minimizes the energy function must satisfy the following Euler equation:

$$(15) \quad \alpha x''(s) - \beta x'''(s) - \nabla E_{ext} = 0$$

The above equation can be considered as an equation for equilibrium between two internal and external forces:

$$(16) \quad F_{int} + F_{ext} = 0$$

For finding an answer to the above equation, both $X(s)$ and parameter "s" are assumed to be time variable. Then, the temporal derivative of $X(s, t)$ is set equal to the left side of equation (15). When the curve becomes stable the equation (15) is satisfied.

$$(17) \quad x_t(s, t) = \alpha x''(s, t) - \beta x'''(s, t) - \nabla E_{ext}$$

The method that Prince and Xu [23, 24] presented, introduces a new external force which is called "gradient vector flow" and is defined as:

$$(18) \quad F_{ext}^{(g)} = v(x, y)$$

Substituting $-\nabla E_{Ext}$ with this term leads to:

$$(19) \quad x_t(s, t) = \alpha x''(s, t) - \beta x'''(s, t) - v$$

The curve that can satisfy the above equation is called a GVF snake. The gradient vector flow field is defined as:

$$(20) \quad V(x, y) = (u(x, y), v(x, y))$$

Which minimizes the following function:

$$(21) \quad \varepsilon = \iint [\mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |v - \nabla f|^2] dx dy$$

It can be proven that the equation (20) can be solved with the two following Euler equations:

$$(22) \quad \begin{aligned} \mu \nabla^2 u - (u - f_x)(f_x^2 + f_y^2) &= 0 \\ \mu \nabla^2 v - (v - f_y)(f_x^2 + f_y^2) &= 0 \end{aligned}$$

where μ is a parameter which controls the dominance of the first and second terms of integral. These equations need a lot of calculations and processing resources which slows down the object tracking techniques. To solve this problem, different methods have been proposed [13-16]. One important problem that has a great effect on the system speed and contours convergence is the initial contour selection. In this paper, for initial contour selection, we selected some points over or near the edges of the object. Since the edges of moving object are extracted in the optical flow section, there is no need to calculate them. For reducing the time needed for calculation of GVF field and increasing the image processing speed, first a window containing the moving object and some pixels around it, is selected. Then the GVF field is calculated in this new window. In the following section, this results obtained based on our method is presented.

Experimental Result

Experiments have been conducted in both indoor and outdoor environments. In both environments the system has an acceptable operation.

Methods and algorithms mentioned in the text are written in Code Gear RAD studio environment with DELPHI programming language. In all of the experiments, the threshold for converting the image to binary is set in a range from 50 to 120 (50 to 60 for low intensity, 70 to 90 for normal intensity and 120 for high intensity situations). Factor μ is set to be 0.2 in all experiments.

The developed software is tested on a laptop system with a core 2 due 2.55 GHz CPU and 4GB of RAM.

In the following section a brief discussion is presented.

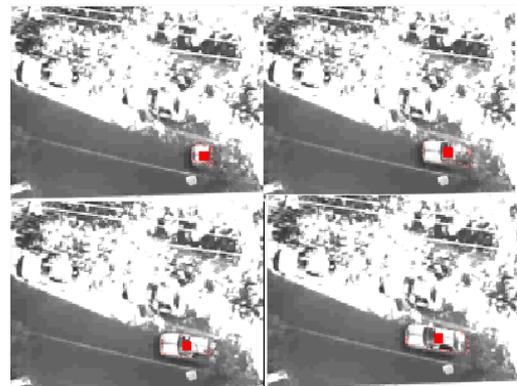


Fig.2. Tracking when the object is covered by another static object in the scene



Fig.3. tracking of a passing car in street

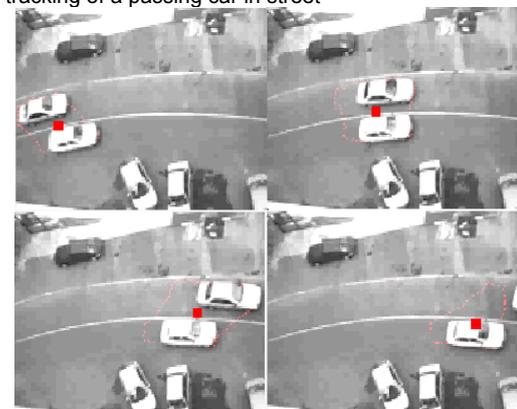


Fig.4. A case with two moving object. Due to the assumption in this paper the system fails to operate correctly

Discussion

In our research work, we combine GVF tracking method with optical flow. Then, we demonstrated that the tracking has good speed. In fact, the GVF method provides good resolutions, while optical flow provides good speed for the tracking. Fig. 2 to 8 shows the tracking pictures which are for different tacking situations. These situations were considered to examine our algorithm. For instance, as shown in the Fig. 2, the track is for a situation where part of moving object was covered by another object. Also, in Fig. 3 tracking are done for a passing car in street. In our work, we assumed that there is only one moving object. Therefore, if another moving object appears in the scene, the tracker fails to track the objects. Fig. 4 shows such a case where two moving cars are present in the scene. Fig. 5 to 8 shows the tracker performance in the indoor

environment. Situations vary from low to high intensity which is from simple to complex backgrounds. Therefore, in all of the situations, the system is capable of tracking.

Conclusion

In this paper, a system is presented which utilizes the optical flow and GVF active contours methods for object tracking purposes. Active contours allow the user to have more details about object in comparison with other methods such as kernels or many other models. Without code optimization, the system processes 10 to 12 images with size of 160x120 pixels which is good enough for many of object tracking.



Fig.5. Tracking indoor with normal intensity and simple background

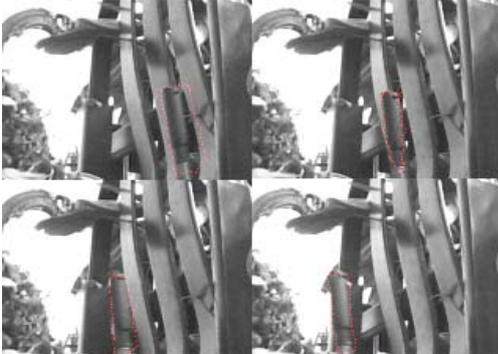


Fig.6. Tracking indoor with relatively high intensity and complex background

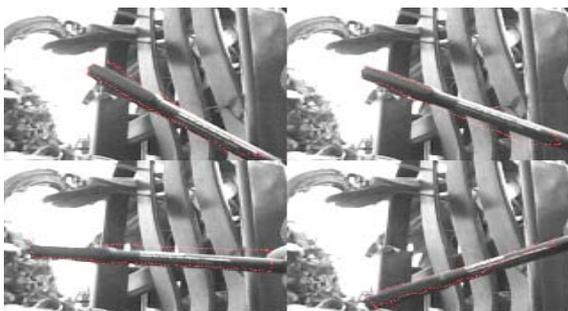


Fig.7. Tracking indoor: A Situation where the orientation of object changes

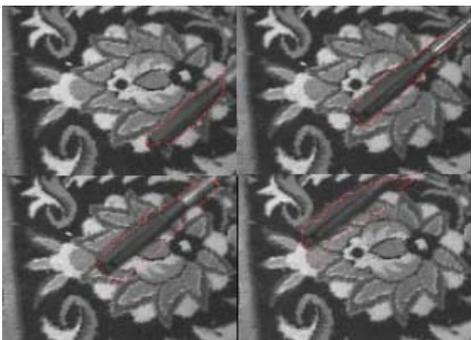


Fig.8. Tracking with relatively low intensity and complex background



Fig.9. The implemented hardware for tracking

REFERENCES

- [1] H.TANIZAKI, Non-Gaussian state-space modeling of non-stationary time series. *J. Amer. Statist. Assoc.* 82, 1032–1063, 1987
- [2] Yilmaz, A., Javed, O., and Shah, M., Object tracking: A survey. *ACM Comput. Surv.* 38, 4, 2006, Article 13, 45 pages
- [3] P. Tissainayagam, and D. Suter, Object tracking in image sequences using point features, *Pattern Recognition* 38, 2005, 105-113
- [4] D. S. Jang, S. W. Jang, and H. I. Choi, "2D human body tracking with Structural Kalman Filter", *Pattern Recognition* 35, 2002, 2041 – 2049
- [5] A. Koschan, S. Kang, , J. Paik, B. Abidi, and M. Abidi, Color active shape models for tracking non-rigid objects, *Pattern Recognition Letters* 24, 2003, 1751–1765
- [6] H. Ning, T. Tan, L. Wang, and W. Hu, "Kinematics-based tracking of human walking in monocular video sequences", *Image and Vision Computing* 22, 2004, 429–441
- [7] W. Kim, and J. J. Lee, "Object tracking based on the modular active shape model", *Mechatronics* 15, 2005, 371–402
- [8] I. Razi, and S. Azadi, "Presenting a novel algorithm for moving object detection", 6th ICEENG Conference, Egypt, 27-29 May, 2008
- [9] D. J. Fleet, and Y. Weiss, "Optical Flow Estimation", *Mathematical Models in Computer Vision: The Handbook*, Springer, 2005, 239-258
- [10] J. L.Barron, and N. A. Thacker, "Tutorial: Computing 2D and 3D Optical Flow", 2004, Tina Memo No. 2004-012
- [11] LUCAS, B. D. and KANADE, T., An iterative image registration technique with an application to stereo vision. *International Joint Conference on Artificial Intelligence*, 1981.
- [12] M. Kass, A. Witkin, , D. Terzopoulos, "Snakes: active contour models", *Int. J. Comput. Vision* 1, 1988, 321–332
- [13] J. Cheng and S. W. Foo, Dynamic Directional Gradient Vector Flow for Snakes, *IEEE Transactions on Image Processing*, vol. 15, no. 6, June 2006, pp. 1563-1571.
- [14] H. Yu, M. S. Pattichis and M. B. Goens, Robust Segmentation of Freehand Ultrasound Image Slices Using Gradient Vector Flow Fast Geometric Active Contours, *IEEE Southwest Symposium on Image Analysis and Interpretation*, 2006, pp.115 – 119
- [15] C. Kiser, C. Musial and P. Sen , Accelerating active contour algorithms with the Gradient Diffusion Field, 19th International Conference on Pattern Recognition, 2008 (ICPR 2008), 1 - 4
- [16] N. Paragios, O. Mellina-Gottardoand, V. Ramesh, "Gradient Vector Flow Fast Geodesic Active Contours", 8th IEEE International Conference on Computer Vision (ICCV 2001) Proceedings., 2001, vol.1, 67 – 73

Authors: Dr. Sassan Azadi, E-mail: sazadi@semnan.ac.ir
 Esmaeil Pourjame , E-mail : pourjame@sun.semnan.ac.ir
 Alireza Nikkerdar , Email : anikkerdar@qums.ac.ir
 Mostafa Sharifi : E-mail : mostafa_sharifi@sun.semnan.ac.ir