

Improved C-V Level Set Algorithm and its Application in Video Segmentation

Jinsheng XIAO, Benshun YI, Xiaoxiao QIU

School of Electronic Information, Wuhan University, Wuhan, China

Email: js_xiao@tom.com

Received March 23, 2009; revised May 15, 2009; accepted June 21, 2009

ABSTRACT

Image segmentation method based on level set model has wide potential application for its excellent segmentation result. However its complex computing restricts its application in video segmentation. In order to improve the speed of image segmentation, this paper presents a new level set initialization method based on Chan-Vese level set model. After a simple iterative, we can separate out the outline of objects. Experiments show that the method is simple and efficient, with good separation effects. The improved Chan-Vese method can be applied in video segmentation.

Keywords: Image Segmentation, Level Set, C-V Model, Video Segmentation

1. Introduction

Image segmentation is intended to separate objects from the image, and the corresponding border is gained at the same time. In recent years, Researchers in the theory and technology of image segmentation has achieved fruitful research results, and active contours extraction is one of the important research results. The research of active contour can be divided into two groups: parameterized active contour based on Snake model which was proposed by Kass [1]; geometric active contour based on Level set methods which were first introduced by Osher and Sethian [2] for capturing moving fronts. Level set methods overcome the weaknesses of other algorithms. Its segmentation results are not sensitive to the initial position and the topology adaptability is strong. Level set method is a powerful tool of curve evolution, which can effectively deal with cusp and has a strong ability to separate complex structure of objects.

Chan-Vese level set model (C-V model) proposed by Chan and Vese [3] was integrated with the ideal of level set and Mumford-Shah model [4]. Being different from the traditional model based on the deformation parameters and geometry active contour model, this model does not rely on the gradient of image when extracting the boundary of objects, so the images with gradient edge

meaningless and ambiguous verge can get a good segmentation. However, it has a weakness like general level set model, amount of computation. At this stage, works on C-V model mainly concentrate on revising its model, such as J. Li [5], etc, through improved C-V model, upgraded the capture of outline from local to the whole image; Y. Y. Gong [6], etc, by amending the C-V model, multi-objects can be extracted based on single level set, and so on. In this paper, we propose a improved C-V model, which can greatly improve the efficiency of segmentation, and can be applied to real-time video image segmentation.

2. Description of C-V Model

C-V model, which is integrated with the thinking of level set and Mumford-Shah model, does not take advantage of gradient information, but minimizes the energy function to evolve curve [3]. Assume that image $I(x, y)$ is formed by two regions: objects (C_o) and background (C_b), which is separated by the evolving curve C in Ω . The constants c_o, c_b , depending on C , are the averages of image I inside C (C_o) and outside C (C_b) respectively. Chan and Vese introduced the energy function $E(C, c_o, c_b)$, defined by

$$E(C, c_o, c_b) = \mu L(C) + \nu S(C) + \lambda_o \int_{C_o} |I - c_o|^2 dxdy + \lambda_b \int_{C_b} |I - c_b|^2 dxdy \quad (1)$$

where $L(C)$ is the length of the curve C , and $S(C)$ is the area of C_o , $\mu, \nu \geq 0$, $\lambda_o, \lambda_b > 0$ are fixed parameters. Therefore the energy function is minimized if the curve is on the boundary of the object. Optimization (1), we can get the ultimate location of segmentation line C , as well as the unknown c_o, c_b .

$$\{C^o, c_o^o, c_b^o\} = \inf_{C, c_o, c_b} E(C, c_o, c_b) \quad (2)$$

Using the Heaviside function $H(z)$, and the one-dimensional Dirac measure $\delta(z)$, and defined, respectively, by

$$H(z) = \begin{cases} 1, z \geq 0 \\ 0, z < 0 \end{cases}, \quad \delta(z) = \frac{d}{dz} H(z) \quad (3)$$

Partial differential equations, gotten by Chan and Vese using Euler-Lagrange method, are as follows: (4)–(7)

$$c_o(\phi) = \frac{\int_{\Omega} I(x, y) H(\phi(x, y)) dxdy}{\int_{\Omega} H(\phi(x, y)) dxdy} \quad (4)$$

$$c_b(\phi) = \frac{\int_{\Omega} I(x, y) (1 - H(\phi(x, y))) dxdy}{\int_{\Omega} (1 - H(\phi(x, y))) dxdy} \quad (5)$$

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left[\mu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) - \nu - \lambda_o (I - c_o)^2 + \lambda_b (I - c_b)^2 \right] \quad (6)$$

$$\phi(0, x, y) = \phi_0(x, y), \quad \text{in } \Omega \quad (7)$$

$$\begin{cases} H_\varepsilon(z) = \frac{1}{2} \left[1 + \frac{2}{\pi} \arctan \left(\frac{z}{\varepsilon} \right) \right] \\ \delta_\varepsilon(z) = \frac{1}{\pi} \cdot \frac{\varepsilon}{\varepsilon^2 + z^2} \end{cases} \quad (8)$$

In the numerical calculations, regularizing Function (8) is used to replace $H(z), \delta(z)$ respectively. So that the gradient flow Equation (6) roles in all of the level set, and we can automatically monitor the empty goal with the internal region, and make the overall energy function to the minimum.

Let's disperse the equation in ϕ , use a finite differences implicit scheme. Recall first the usual notations: let h be the space step, Δt be the time step, and $(x_i, y_j) = (ih, jh)$ be the grid points, where $1 \leq i, j \leq M$. Let $\phi_{i,j}^n = \phi(n\Delta t, x_i, y_j)$ be an approximation of $\phi(t, x, y)$. Knowing ϕ^n , we can get and $c_o(\phi^n), c_b(\phi^n)$ using (4) and (5). The finite differences are

$$\Delta_{\pm}^x \phi_{i,j} = \pm(\phi_{i\pm 1,j} - \phi_{i,j}), \quad \Delta_{\pm}^y \phi_{i,j} = \pm(\phi_{i,j\pm 1} - \phi_{i,j})$$

Chan and Vese compute ϕ^{n+1} through (11).

$$L = \frac{\mu}{h} \Delta_x^x \left(\frac{\Delta_+^x \phi_{i,j}^n}{\sqrt{(\Delta_+^x \phi_{i,j}^n)^2 + (\Delta_+^y \phi_{i,j}^n)^2}} \right) + \frac{\mu}{h} \Delta_y^y \left(\frac{\Delta_+^y \phi_{i,j}^n}{\sqrt{(\Delta_+^x \phi_{i,j}^n)^2 + (\Delta_+^y \phi_{i,j}^n)^2}} \right) \quad (9)$$

$$R = -\nu - \lambda_o (u_{0,i,j} - c_o(\phi_{i,j}^n))^2 + \lambda_b (u_{0,i,j} - c_b(\phi_{i,j}^n))^2 \quad (10)$$

$$\frac{\phi_{i,j}^{n+1} - \phi_{i,j}^n}{\Delta t} = \delta_h(\phi_{i,j}^n) [L + R] \quad (11)$$

From the Equation (6), can be seen, the definition of partial differential equations involving image function $I(x, y)$ is domain-wide map data, and the definition of other two unknown c_o, c_b is also image definition of the region, with the overall characteristics. Hence, updating level set function is in the entire defined region, the computation is large [3].

3. Improved C-V Model

From the above analysis, we know that C-V method has a grate calculation. If we can effectively reduce the amount of computation, C-V method can be more widely applied. In traditional level set methods, it is necessary to initialize the level set function Φ as a signed distance

function Φ_0 . Curve C divides the plane into internal and external regions. $\Phi(x, y) = \pm d, d$ is the distance from point (x, y) to curve C. Generally, the distance of internal and external points are negative and positive respectively and signed distance function needs to be re-initialized. C-V method generally defines the symbol distance function (SDF) as a cone, with particularly complex calculation.

Lie [8] demonstrates that the presence of signed distance function is not inevitable. Lie imposed a binary level set model. We will introduce this idea into C-V method. Define radius of the closed curve C as infinite, C represents a straight line in plane Ω , which will be divided into upper regional Ω_u and lower regional Ω_d . Initialization function ϕ_0 is defined as:

$$\phi_0(x, y) = \begin{cases} -\rho_u, & (x, y) \in \Omega_u \\ \rho_d, & (x, y) \in \Omega_d \end{cases} \quad \rho_u, \rho_d > 0 \quad (12)$$

So first level set evolution, the curve of the internal and external simplified curve of the upper and lower regions, the calculation of c_o, c_b are very simple, The calculation of difference operator is very simple too, as only points on the boundary of upper and lower region are non-zero constant, and the remaining places are zero value; The initialization level set function is fixed constants, as well as $\delta(z)$. In a linear mesh C at the point, follow the Reference [7] approach, iterative Formula (9) can be transformed as follows:

$$\frac{\phi_{i,j}^{n+1} - \phi_{i,j}^n}{\Delta t} = \begin{cases} \delta_h(\phi_{i,j}^n)R, & j < j_0 - 1, j > j_0 \\ \delta_h(\phi_{i,j}^n) \left(-\frac{u}{h} \text{sign}(\rho_u - \rho_d) + R \right), & j = j_0 \\ \delta_h(\phi_{i,j}^n) \left(\frac{u}{h} \text{sign}(\rho_u - \rho_d) + R \right), & j = j_0 - 1 \end{cases} \quad (13)$$

Obviously, compared with the traditional C-V method, the calculation of our method is much smaller in the first level set iterative.

In order to ensure the level set method not departure from SDF, the time step must be very small, usually 0.1 s, which increasing the evolution of time. Since the existence of SDF is not inevitable, we appropriately increase the time step. Using larger time step can speed up the evolution, but may cause error in the boundary location if the time step is chosen too large. There is a tradeoff between choosing larger time step and accuracy in boundary location. Usually, we use $\Delta t \leq 10.0$ for the most images.

We list out some of the experimental results in the following paper. In our numerical experiments, we generally choose the parameters as follows: $\lambda_o = \lambda_b = 1$, $\nu = 0$, $h = 1$ (the space step), $\Delta t = 2$ (the time step), $\varepsilon = 1$.

In the first experiment, we chose the test image cavern. Jpg (Figure 1(a)), whose size is 200×200 . The region of

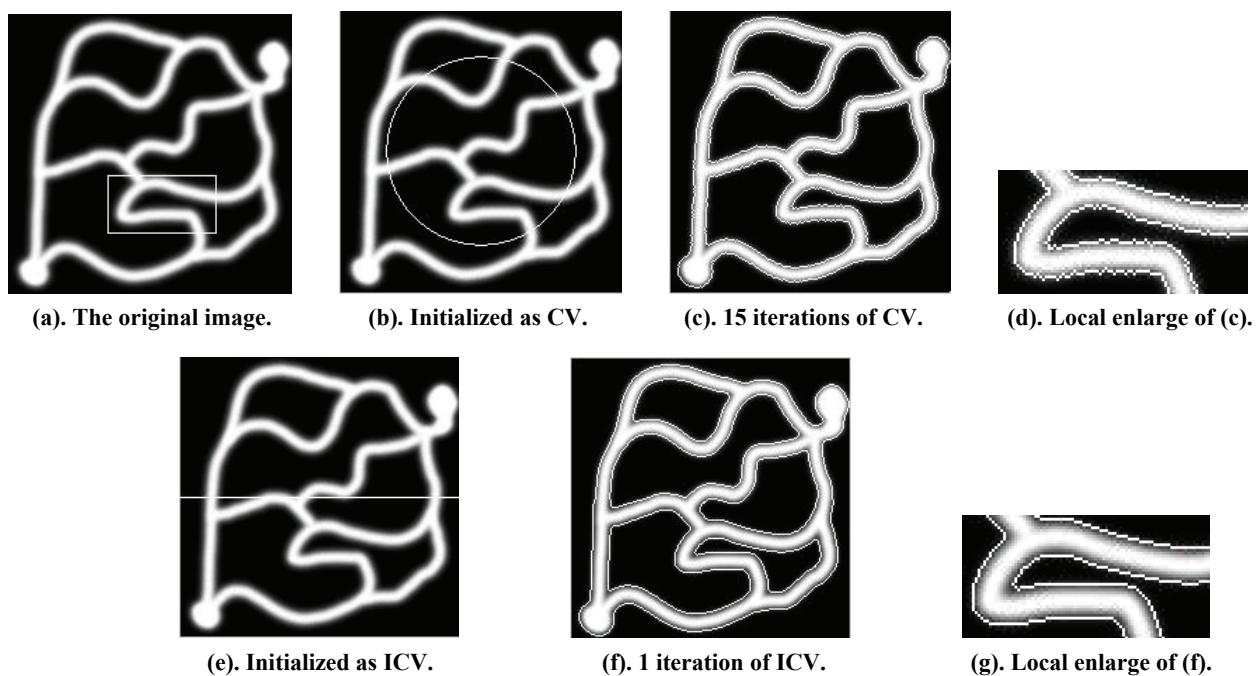


Figure 1. The binary image.

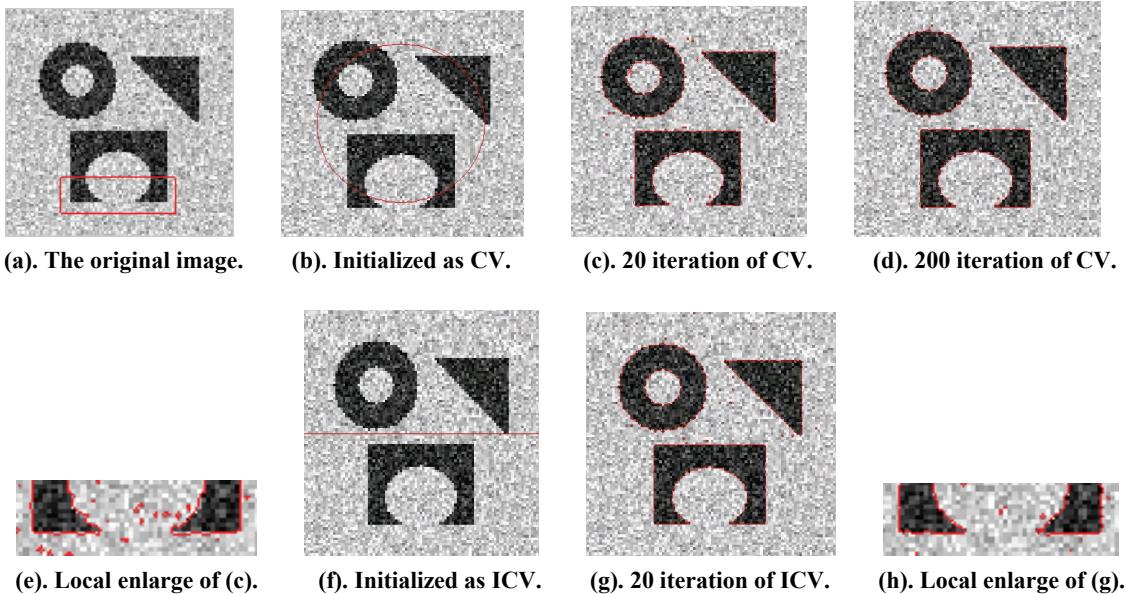


Figure 2. The noisy image.

interest is specified by the white box. It needs 15 iterative to achieve the ideal state of division when the level set function is initialized as $\phi_0(x, y) = 67 - \sqrt{(x-100)^2 + (y-100)^2}$, and time-consuming is 1.578 s. Figures 1(b), 1(c) show the traditional method of initialization(CV method) and the final result of division. With our improved method (ICV method), $\rho_u = -1, \rho_d = 1$, we can get the desired effect of split after one iterative and the time-consuming is 0.016s (Figures 1(e), 1(f)). To compare the segmentation results of the two methods in more detail, we show a zoomed version of the results in Figure (d) and Figure (g) for a region delineated by the white box in Figure (a). It's clear to see that we can get a smoother curve split with our approach.

In Figure 2, we show how our arithmetic and traditional C-V methods work on a noisy synthetic image. The region of interest is specified by the white box (Figure 2(a)). Both of the two methods can automatically detect the objects. If we use traditional C-V model initialization method, $\phi_0(x, y) = 33 - \sqrt{(x-50)^2 + (y-50)^2}$, the outline of objectives are separated after 20 times of iterative, but the right-angle region isn't well separated(Figures 2(c), 2(e)). After 200 iterative, the situation is improved. However, it spends 20 times iteration to achieve a perfect result with our method. The rectangular region is also divided perfect (Figures 2(g), 2(h)).

The test results of cameraman.tif are showed in Figure 3. Size of the picture is 256×256 . In Figure 3(a), the level set function was initialized as signed distance function: $\phi_0(x, y) = -\sqrt{(x-128)^2 + (y-128)^2} + 85$. Figure 3(b) and Figure 3(c) are results of 10 iterations and 400 iterations

respectively. In Figure 5(d) level set function was initialized as a linear curve. We can get the photographer's outline after only one time of iteration. But there is a lot of noise divisions on the lake, noises would gradually reduce after multiple iterations. Figure 3(e) is the result of 10 iterations. After 150 iterations, we get a stable state. We can see that the noise has been significantly reduced, but not completely filtered out.

4. Video Image Segmentation Based on ICV

C-V model is the key to resolve the two issues separate, and categories of backgrounds and objectives in the video images are unknown. C-V model can not be directly applied to the whole map. So we get the movement area by analyzing the characteristics of H.264 codec first of all, and then the improved C-V model is applied to the movement region which was detected. This article focuses on the moving target in the same scene, the video image processing and experimental data are based on a single static camera, so the background image is static. According the principles of H.264 video codec, motion vector of macro block as a background is usually zero value, and only motion vector of macro block in the movement region is non-zero, so we can get the moving region through the distribution of motion vector.

When inter coding being chosen, the coding frame need to use the frame before (reference frame) for movement searching, and then motion vector of each block is get. We illustrate in Figure 4 the above remarks. In Figure 4(a), the large box represents 16×16 macro block, and the small box represents 16×8 or 8×16 block, black line is motion vector. By the vector distribu-

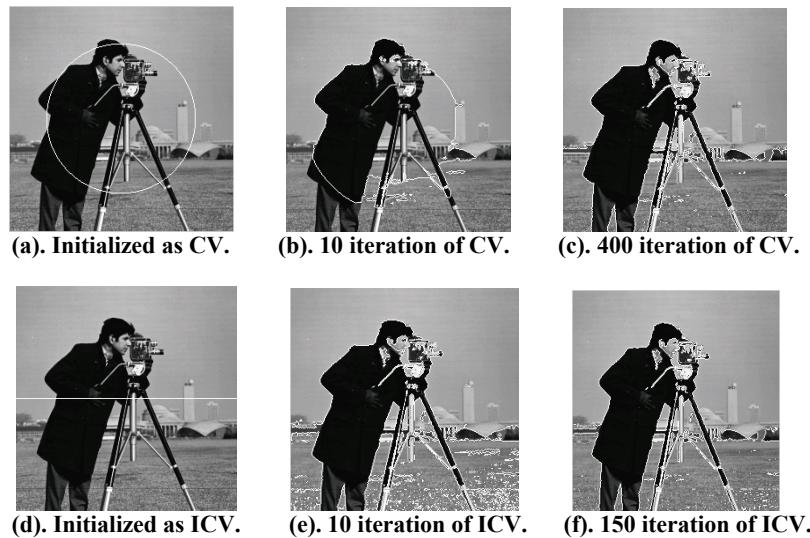


Figure 3. The complex background image.



Figure 4. Moving target separated.

tion map, we can frame the regional campaign, as shown in Figure 4(b).

Figure 4(c) is a real-time video segmentation results map, target detection of the system and tracking algorithm and real-time are tested respectively. We choose the parameters as follows: $v = 0$, $\lambda_o = \lambda_b = 1$, $\mu = 0.01 \times 255 \times 255$, $\varepsilon = 1$, $\Delta t = 2$, $h = 1$, $\rho_u = -1$, $\rho_d = 1$. The regional campaign is extracted, and the moving target is tracked very well. In Figure 4(c), however, we can see that the outline of some of the goal in frame is not very ideal, sometimes the background is also included. It is because the delimitation of regional campaign is mainly based on motion estimation which processed by video coding. If light is changed, or other factors, the reference block selecting is inaccurate in motion estimation, resulting in the coding block is mistakenly believed as campaign block, so the error division is produced. The brightness weight of hands is very close to the brightness weight of the wall. As target tracking algorithm uses the brightness information of pixel, when the goal and background have similar brightness information, the algorithm will get the wrong track.

H.264 codec set the frame rate of 25 fps. The size of the regional movement we get by the motion vector information of decoder is an important factor of computing time, which affect the moving target detection and tracking module. Larger the regional campaign is, more time the target detection and tracking module need. We have real-time measured the running time of the module, which is in 15–19 ms range. When the ICV model is evolved in the whole frame, the time-consuming is not more than 20 ms, fully meet the real-time requirements.

5. Conclusions

In this paper, we have improved the image segmentation efficiency based on the C-V model from initialization and algorithms simplifying. The improved C-V model guarantees the division in effect while greatly improves the efficiency of the division. We apply it to the real-time H.264 video codec system. The experimental results show that, for the image of simple background, it can partition the outline of objects with dramatic speed and

high efficiency segmentation, using the methods proposed in this paper. However, for the complex background image, the outline of objects can be accurately divided, but there will be some regional background mistakenly separated. How to filter this noise fast is the next step that we need to improve. This article improves the C-V method to separate video images, which can be used to real-time detect and track the moving targets.

6. References

- [1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models [J]," *International Journal of Computer Vision*, Vol. 1, No. 4, pp. 321–331, 1987.
- [2] S. Osher and J. A. Sethian, "Fronts propagating with curvature dependent speed: Algorithms based on hamilton-jacobi formulations [J]," *Journal of Computational Physics*, Vol. 79, pp. 12–49. 1988.
- [3] F. T. Chan and L. Vese, "Active contours without edges [J]," *IEEE Transaction Image Processing*, Vol. 10, No. 2, pp. 266–277, 2001.
- [4] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems [J]," *Communication of Pure Applied Mathematics*, Vol. 42, No. 5, pp. 577–685, 1989.
- [5] J. Li, X. Yang, and P. F. Shi, "A fast level set approach to image segmentation based on mumford-shah model [J]," *Chinese Journal of Computers*, Vol. 25, No. 11, pp. 1175 –1183. 2002.
- [6] Y. Y. Gong, X. N. Luo, H. Huang, G. J. Liao, and Y. Zhang, "Multi-objects extracted based on single level set [J]," *Chinese Journal of Computers*, Vol. 30, No. 1, pp. 120–128, 2007.
- [7] J. S. Xiao, H. Feng, and B. S. Yi, "Finite difference method for semilinear parabolic differential inclusions [J]," *Journal of Wuhan University, Natural Sciences Edition*, Vol. 52, No. 3, pp. 262–266, 2006.
- [8] J. Lie, M. Lysaker, and X. C. Tai, "A binary level set model and some applications to mumford-shah image segmentation," *IEEE Transactions on Image Processing*, Vol. 15, No. 5, pp. 1171–1181, 2006.