韓國컴퓨터情報學會 論文誌
第14卷 第7號, 2009. 7.

2009-14-7-2-2

# Using Mean Shift Algorithm and Self-adaptive Canny Algorithm for Improvement of Edge Detection

Shin Seong-Yoon*, Pyo Seong Bae**

# 경계선 검출의 향상을 위한 Mean Shift 알고리즘과 자기 적응적 Canny 알고리즘의 활용

신 성 윤*, 표 성 배**

## 요 약

Edge detection is very significant in low level image processing. However, majority edge detection methods are not only effective enough cause of the noise points' influence, even not flexible enough to different input images. In order to sort these problems, in this paper an algorithm is presented that has an extra noise reduction stage at first, and then automatically selects the both thresholds depending on gradient amplitude histogram and intra class minimum variance. Using this algorithm, can fade out almost all of the sensitive noise points, and calculate the propose thresholds for different images without setting up the practical parameters artificially, and then choose edge pixels by fuzzy algorithm. In finally, get the better result than the former Canny algorithm.

## Abstract

전경계선 검출은 저수준 영상 처리에서 매우 중요하다. 하지만, 대부분의 경계선 검출 방법들은 노이즈 포인트들의 영향으로 효과적이지 못하며 서로 다른 입력 영상에서도 유연하지 못하다. 이 문제를 해결하기 위하여 본 논문에서는 먼저 외부 노이즈 제거 단계를 제시하였고, 다음으로 기울기 폭 히스토그램과 내부 클래스 최소 변이에 따른 양쪽 임계치의 자동 선택을 제시하였다. 이 알고리즘을 사용하여 민감한 노이즈 포인트들의 대부분을 줄일 수 있었고 실제 파라미터를 인위적으로 세팅하지 않고 서로 다른 영상을 위한 목적 임계치를 계산하며, 퍼지 알고리즘에 의하여 경계선 픽셀들을 선택하였다. 결론적으로 이전의 Canny 알고리즘보다 훨씬 더 좋은 결과를 얻을 수 있었다.

▸ Keyword : 경계선 검출(edge detection), Canny 알고리즘(Canny algorithm), 기울기 폭 히스토그램(Gradient Amplitude Histogram), 내부 클래스 최소 변이(Intra Class Minimum Variance)

---

• 제1저자 : 신성윤
• 투고일 : 2009. 06. 22, 심사일 : 2009. 06. 25, 게재확정일 : 2009. 07. 18.
* 군산대학교 컴퓨터정보공학과 교수 ** 인덕대학 컴퓨터 소프트웨어과 교수

# Ⅰ. Introduction

In a digital image, the boundary of objects usually contains much information; although just include some few quantities of pixels out of the whole image. Therefore edge detection is an important and challenged task in image processing, image analysis and computer vision field. Hitherto the Canny algorithm already become one of the most suitable and common used method among the others, so select Canny algorithm for edge detection is still main stream. However, there are two drawbacks plague it all the time, the first is it can't avoid the sensitive noise points effect, the second is it's parameters of both thresholds should be set artificially. Hypothetical situation, take the same image, stay in different circumstance, the both thresholds have to be set different from each other by person, let alone the different image. The result of being stubborn in the original Canny algorithm manifest the user have to set parameters every time. So that there are many masters research building thresholds automatically method thorough the world [1][2][3], and bring the alleged self adaptive methods who have also need to set the scale coefficient of threshold. That can't shrug off the human factor as well.

There is an approach try to get rid of human factor influence and select the most suitable thresholds automatically hinge on the gradient amplitude histogram and intra class variance, and it's supposed to be the propose method to improve original Canny algorithm. By besides, it can fade out the sensitive noise points use Mean Shift algorithm for extra smoothing in advance.

# Ⅱ. Related Work

Edge detection is a very important feature-extraction method. A huge number of edge detectors has been developed from different perspectives(e.g., [4][5]). Although many edge-detection evaluation methods have been developed in the past years, this is still a challenging and unsolved problem. The major challenge comes from the difficulty in choosing an appropriate performance measure of the edge-detection results. In [6], we show the superiority of the ratio-contour algorithm over other existing algorithms for detecting salient closed boundary in a set of detected edges.

Depth edges directly represent shape features that are valuable information in computer vision [7]. Unfortunately, few research results have been reported that provide only depth discontinuities without computing 3D information at every pixel in the input image of a scene.

Recently, the use of structured light was reported to compute 3D coordinates at every pixel in the input image [8][9]. However, the fact that this approach needs a number of structured light images makes it hard to be applicable in realtime. One notable technique was reported recently for non-photorealistic rendering [10]. They capture a sequence of images in which different light sources illuminate the scene from various positions. Then they use shadows in each image to assemble a depth edge map. This technique was applied to finger spelling recognition [11].

Till now, how to set the low and high thresholds has no solid criteria, so decide them depend on experience can't avoid from influent of individual. But the Otsu's method assumes that the image to be thresholded contains two classes of pixels, the foreground and background, then calculates the optimum threshold separating those tow classes so that theitowombined spread is minimal. This method inspire the masters explore its scale and for more than one threshold selale andin many fields. According to Otsu's description, minimizing the intra class variance $\sigma_w^2(t)=w_1(t)\sigma_1^2(t)+w_2(t)\sigma_2^2(t)$, is the same as maximizing inter class variance $\sigma_b^2(t)=\sigma^2-\sigma_w^2(t)=w_1(t)w_{2(t)}[\mu_1(t)-\mu_2(t)]^2$, so from both of these aspects can find out the propose thresholds, and this portion is the kernel in recently years research.

## III. Improvement of Mean Shift and Self-adaptive Canny Algorithm

### 3.1 Improvement of Mean Shift Smoothing

The Mean Shift algorithm is a method finds the cluster center by constantly iteration and its basic idea is moving a shift window on the gradient direction of the feature space. Take brief explanation of clustering problem. Given a set of data points $\{x_i\}$ in a d-dimensional Euclidean space $R^d$, assign a label $l_i$ to each point $x_i$ base don proximity to high density regions in the space, see the following diagram(Fig 1)

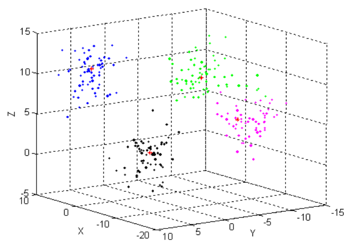

그림 1. 클러스터링 다이어그램
Fig. 1. Clustering Diagram

The multivariate kernel density estimate is defined as: , $f(x)=\frac{1}{nh^d}\sum_{i=1} k(\frac{x-x_i}{h})$ using the Epanechnikov kernel:

$$K_E(x)=\begin{cases}(2c_d)^{-1}(d+2)(1-x^Tx) & x^Tx<1 \\ 0 & otherwise\end{cases}$$

,and we obtain the Mean Shift vector as:

$$M_h(x)=\frac{h^2}{d+2}\frac{\nabla f(x)}{f(x)}=\frac{1}{n_x}\sum_{x_i\in S_h(x)} x_i - x$$

, and the Mean Shift iteration, derived from the Mean Shift vector :

$$M(x): y_{k+1}=\frac{1}{n_x}\sum_{x_i\in S_h(y_k)} x_i$$

, where $n_k$ is the number of points in a sphere of radius $h$, and $h$ is the window radius[4]. The following diagrams(Fig 2,3) show the multivariate kernel density estimate:

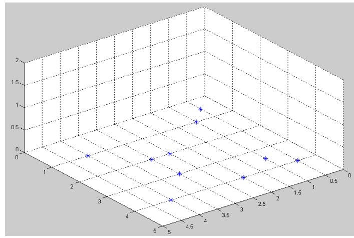

그림 2. 예제 데이터 셋 {Xi}
Fig. 2. Example Data Set {Xi}

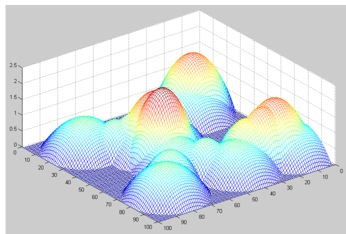

그림 3. 다중변이 커널 강도
Fig. 3.Multivariate Kernel Density

$$f(x)=\sum_{i=1}^{n} K(x-x_i)$$

$$K_E(x)=\begin{cases}(1-x^Tx & x^Tx<1 \\ 0 & otherwise\end{cases}$$

The Mean Shift procedure smooths the image while preserving and sharpening its discontinuities. It should be recalled that smoothing has a positive effect to reduce sensitive noise points and ensure robust edge detection; while a negative effect on information loss. Clearly, a fundamental tradeoff between loss of information and noise reduction is crucial.

This procedure make true in computer as the following steps:

Initialize data set: $I(i,j)\rightarrow(i,j,I(i,j)^*C)$

For each $j=1,\ldots, n$

Initialize $k=1$ and $y_k=x_j$.

Repeat: Compute $y_{k+1}$using the Mean Shift iteration; $k\leftarrow k+1$

Until convergence $(y_{k+1}-y_k\langle\varepsilon)$.

Assign $I_{smoothed}(x_j(1),x_j(2))=y_k(3)$.

### 3.2 Self-adaptive Canny Technique

The original Canny algorithm adopt the finite differential of adjacent area's first partial derivate to calculate the data matrix $I(x,y)$'s gradient amplitude and orientation. The partial derivate of $x$ and $y$'s 2 matrixes, which are $P_x[i,j]$ and $P_y[i,j]$:

$$P_x[i,j]=(i[i,j+1]-I[i,j]+I[i+1,j+1]-I[i+1,j])/2$$
$$P_y[i,j]=(i[i,j]-I[i+1,j]+I[i,j+1]-I[i+1,j+1])/2$$

Pixel's gradient amplitude and orientation use convert formula rectangular to polar to figure out: the gradient amplitude is $M[i,j]=\sqrt{(P_x[i,j])^2+(P_y[i,j])^2}$, and the gradient orientation is $\Theta[i,j]=\arctan(P_y[i,j]/P_x[i,j])$. Then in order to extract the single pixel thick edge, we have to make out thinned amplitude image. However, near the position of maxima $M[i,j]$ will appear ridge; maintaining the maxima regional change point of amplitude can determine the edge position accurately, and this procedure is the alleged non-maxima suppression. In this section, the Canny algorithm use the magnitude of 3X3, 8 directions adjacent area interpolate all pixels of amplitude matrix $M[i,j]$ along the gradient direction. On every point, compare the adjacent area center pixel $m[i,j]$ with the two interpolation result along the gradient direction. If adjacent area center point' amplitude $M[i,j]$ is not large than both interpolation results, assign the edge corresponding $m[i,j]$ to zero. This procedure details the ridge into one pixel magnitude, and maintains the gradient amplitude at the same time.

After the above steps, we have to determine the double thresholds and connect the edge. Undergone the non-maxima suppression and set both high threshold $H_{TH}$ and low threshold $H_{TL}$ for the sub-image $N[i,j]$ of gradient histogram, assign zero to the pixel whose gray vale is less than the threshold, then get two edge images $T_h[i,j]$ and $T_l[i,j]$. The $T_h[i,j]$ who through high threshold include very few fake edges, while $T_l[i,j]$ is opposition, consist of detail information but many more fake edges. Therefore combine both of them can get plausible edge [6].

According to these descriptions, the both thresholds determination can't refrain from the effect of personal factor, so need some method can absolutely exclude the personal factor and automatically determinate the thresholds hinge on different images. Inspirited by the Otsu's method [7], adopt the improved approach which based on gradient amplitude histogram and intra class variance to make out the both thresholds and then undertake non-maxima suppression to select the proper points, accomplish edge detection finally.

The gradient amplitude through non-maxima suppression is compose of L level, modulus maxima separate into three classes: $C_0$, $C_1$, $C_2$, $C_0$ class is the pixel that didn't belong to edge, $C_2$ class is the pixel belong to edge, $C_1$ class include the rest limbo pixels. Assume $N_i$ is the amount of pixels whose modulusis $i$, $P_i$ is the ratio of this kind of pixel to the whole pixels: $P_i=\frac{n_i}{N}, p_i\geq 0, \sum_{i=0}^{L-1} P_i=1$, let $C_0$ include the pixels whose modulusis $[0, 1, \ldots, k]$, $C_1$ include the pixels whose modulusis $[k+1, k+2, \ldots, m]$, $C_2$ include the pixels whose modulusis $[m+1, m+2, \ldots, l-1]$: $\mu_T=\mu(l-1)=\sum_{i=0}^{l-1} ip_i$,

$$\begin{cases} \omega_0(k)=\sum_{i=0}^{k} P_i, \omega_1(k,m)=\sum_{k+1}^{m} p_i, \\ \omega_2(m)=\sum_{m-1}^{l-1} p_i, \mu_0(k)=\frac{\sum_{i=0}^{k} ip_i}{\omega_0}, \\ \mu_1(k,m)=\frac{\sum_{i=k+1}^{m} ip_i}{\omega_1}, \mu_2(m)=\frac{\sum_{i=m+1}^{l-1} ip_i}{\omega_2}, \end{cases}$$

$$\begin{cases} \sigma_0^2 = \dfrac{\sum_{i=0}^{k}(i-\mu_0)^2 p_i}{\omega_0}, \\ \sigma_1^2 = \dfrac{\sum_{i=k+1}^{m}(i-\mu_1)^2 p_i}{\omega_1}, \\ \sigma_2^2 = \dfrac{\sum_{i=m+1}^{l-1}(i-\mu_2)^2 p_i}{\omega_2} \end{cases}$$

Propose the evaluate function "based on gradient amplitude histogram and minimized intra class variance determine the two thresholds": $J(k,u) = Arg\min(\sigma_w^2) = Arg\min(w_0\sigma_0^2 + w_1\sigma_1^2 + w_2\sigma_2^2)$.

Minimized intra class variance reflects the difference between the different classes should be minima. Through this procedure convert itself into first order statistical data, get the result has the advantages that easy to figure out and code out. There is the inference:

$$J(k,m) = \int_0^k (i-\mu_0(k))^2 p_i di + \int_{k+1}^m (i-\mu_1(k,m))^2 p_i di + \int_{m+1}^{l-1} (i-\mu_2(m))^2 p_i di,$$

$$\frac{\partial J(k,m)}{\partial k} = (k-\mu_0(k))^2 p_k - 2\mu_0^{(1)}(k)\int_0^k (g-\mu_0(k))p_i di - (k-\mu_1(k,m))^2 p_k - 2\mu_1^{(1)}(k,m)\int_m^{k+1}(g-\mu_1(k,m))p_i di.$$

From mathematical statistics we know that: if $\int_0^k (g-\mu_0(k))p_i di = 0$ ,than

$$\frac{\partial j(k,m)}{\partial k} = [k-\mu_0(k)]^2 p_k - [k-\mu_1(k,m)]^2 p_k.$$

Let $\dfrac{\partial J(k,m)}{\partial k} = 0$ , and simplify the equation than get $2k-\mu_0(k)-\mu_1(k,m)=0$ . In a similar way: $2m-\mu_1(k,m)-\mu_2(m)=0$ .

From the above proof process, get the results that $m$, $k$ are the most suitable extreme points for split. After that use fuzzy control method filter the pixel between the both thresholds.

When adopting both thresholds to filter the maximum modulus undergone the non-maxima suppression, fade out the point whose modulus is less than the low threshold as the non boundary point, while maintain the point that is large than the high threshold as the boundary point. But the pixel between the both thresholds can't decide. Because there is one regional peak in any boundary and its vertical direction, hence decide whether the maximum modulus between both thresholds is the boundary or not. According to some point maximum modulus' direction, determine the direction of boundary, and select its line adjacent area at boundary vertical direction, there are two pixels on this point's every side. If its maximum modulus is the most among its adjacent area, it is boundary point, otherwise is not. Using fuzzy control algorithm extract the pixel between the thresholds, set up membership function depending on maximum modulus $Mf_{2^j}(x,y)$,and undertake edge extraction and connection:

$$\mu(Mf_{2^j}(x,y)) = \begin{cases} 1, Mf_{2^j}(x,y) \ge m, \\ 0, Mf_{2^j}(x,y) \le k, \\ \{1+[q_{\max} - Mf_{2^j}(x,y)]^2\}^{-1}, k < Mf_{2^j}(x,y) < m \end{cases}$$

The $q_{max}$ is the largest one among the 5 pixels which perpendicular to edge direction. And the membership show that, if the maximum modulus of decision point is not the largest among the adjacent area, this point must not belong to the edge; if it is, it belongs to the edge definitely.

## Ⅳ. Experiment and Result

Take experiment to compare the original Canny method with the improved method combined Mean Shift smoothing and self adaptive thresholds selection, the result is obvious different:

그림 4. 원 영상
Fig. 4. The original image

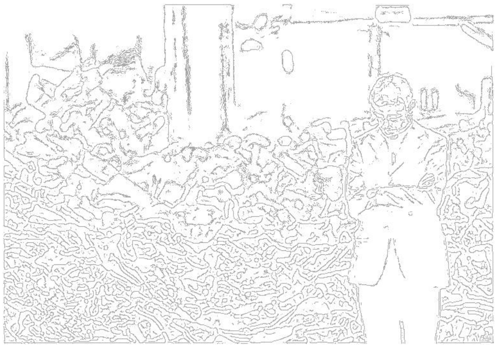

그림 5. σ=0.4, Lratio=0.4, Hratio=0.8일 때 원 Canny 알고리즘
Fig. 5. The original Canny algorithm, where σ=0.4, Lratio=0.4, Hratio=0.8;

그림 6. 자기 적응적 Canny 알고리즘의 결과
Fig. 6. The result of using self adaptive Canny algorithm

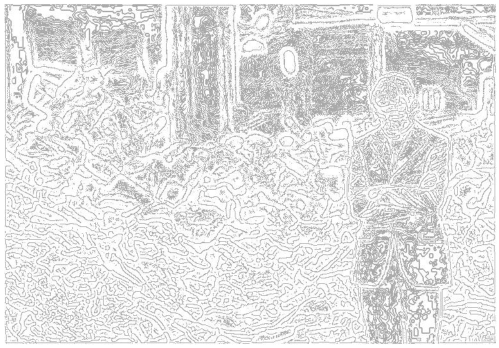

그림 7. hs=32, hr=16일 때 Mean Shift 스무딩에 자기 적응적 Canny 알고리즘을 합한 방법의 결과
Fig. 7. The result of Mean Shift smoothing plus self adaptive Canny algorithm, where hs=32, hr=16

The self adaptive Canny algorithm should detect the propose boundary of objects automatically from technique aspect, and from experience result, it really depict the contour we want, even some inner details, such as the outline of suit's lapel. On the other hand, the surplus of precise texture can't reframe from involving noise points and useless background some or less, so make the present pixels a bit chaos. Then use Mean Shift algorithm smooth the raw image ahead of time, theoretically speaking, this procedure should eliminate the sensitive noise points and simplify the following work, from experience result, it indeed do it. But tackle its effect this time is the natural drawback of Mean Shift that will blur the outline of objects, therefore make the consequence of the series ragged lines. So it is clear that the above combination of Mean Shift and self adaptive Canny can detect the more practical edge, but import extra chaos some or less at the same time. The kernel of how to make good use of this method's advantage will hinge on the coefficients of Mean Shift and the only one of Canny against the original one's three. The following table shows the merits and demerits between the traditional and improved methods:

표 1. 제안 방법의 장점과 단점
Table. 1 Merits and Demerits of Proposed Method

| | Time Cost | Contour Resolution | Details Obtain |
|---|---|---|---|
| Canny | 0.24㎲ | Low | Medium |
| Auto Canny | 3.56㎲ | High | More |
| Mean Shift +Auto Canny | 604.60㎲ | Medium | Less |

(The Mean Shift didn't accelerate.)

## IV. Conclusion

This paper presented a method can detect more practical edge base on the idea of self adaptive Canny and Mean Shift algorithm. The Mean Shift eliminates the sensitive noise points at first, and then improves the original Canny algorithm by intra class variance minimum theory, select the both thresholds automatically, in the end can get better result than the original approach. However, both of the methods' natural drawbacks lower the efficiency each other, and how to set the necessary parameters of both seem as very important.

## Reference

〔1〕 R.S.Anand, Pradeep Kumar, "Flaw detection in radiographic weld images using morphological approach", NDT&E International, 2006.
〔2〕 Kiranyaz S, Ferreira M, Gabbouj M, "Automatic object extraction over multi scale edge field for multimedia retrieval", IEEE Trans. Image Processing, 2006.
〔3〕 Hongshan Yu, Yaonan Wang, "An improved Canny edge detection algorithm", Engineering and Application of Computer, 2004.
〔4〕 K.W. Bowyer, C. Kranenburg, and S. Dougherty, "dge detector evaluation using empirical ROC curves,"Computer Vision and Image Under standing, Vol. 84, No. 1, pp. 77 - 03, 2001.
〔5〕 S. Konishi, A. L. Yuille, J. M. Coughlan, and S. C. Zhu, "tatistical edge detection: learning and evaluating edge cues," IEEE Transactions on Pattern Analysis andMachine Intelligence, vol. 25, no. 1, pp. 57 - 4, 2003.
〔6〕 S.Wang, T. Kubota, J. M. Siskind, and J.Wang, "Salient closed boundary extraction with ratio contour,"IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 27, no. 4, pp. 546 - 61, 2005.
〔7〕 I. Weiss and M. Ray, "Model-based recognition of 3D object from single vision," IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 116-128, 2001.
〔8〕 Sukhan Lee, Jongmoo Choi, Daesik Kim, Byungchan Jung, Jaekeun Na, and Hoonmo Kim, "An Active 3D Robot Camera for Home Environment," Proceedings of 4th IEEE Sensors Conference, 2004.
〔9〕 D. Scharstein and R. Szeliski, "High-Accuracy Stereo Depth Maps Using Structured Light," IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Vol. 1, pp. 195-202, 2003.
〔10〕 R. Raskar, K. H. Tan, R. Feris, J. Yu, and M. Turk, "Non-photorealistic Camera: Depth Edge Detection and Stylized Rendering Using Multi-Flash Imaging," Proceedings of ACM SIGGRAPH Conference, Vol. 23, pp. 679-688, 2004.
〔11〕 R. Feris, M. Turk, R. Raskar, K. Tan, and G. Ohashi, "Exploiting Depth Discontinuities for Visionbased Fingerspelling Recognition," IEEE Workshop on Real-Time Vision for Human-Computer Interaction, 2004.

## 저 자 소 개

**신 성 윤**
2003년 2월 군산대학교
컴퓨터과학과 이학박사
2006년~현재 :군산대학교
컴퓨터정보공학과
교수
관심분야: 비디오 인덱싱, 비디오
요약, 멀티미디어,
색채공학

**표 성 배**
2004년 2월 : 숭실대학교
컴퓨터학부 공학박사
1992년~현재: 인덕대학 컴퓨터
소프트웨어과 교수
관심분야: 멀티미디어, 비디오
색인화, 비디오 요약