KNOWLEDGE-BASED BOUNDARY EXTRACTION OF MULTI-CLASSES OBJECTS

Hae Chul Choi, Ho Chul Shin, Jin Sung Lee, Ju Hyun Cho and Seong Dae Kim

Division of Electrical Engineering, Department of Electrical Engineering & Computer Science,
Korea Advanced Institute of Science & Technology, Korea
E-mail: chc@sdvision.kaist.ac.kr

ABSTRACT

We propose a knowledge-based algorithm for extracting an object boundary from low-quality image like the forward looking infrared image. With the multi-classes training data set, the global shape is modeled by multispace KL(MKL)[1] and curvature model. And the objective function for fitting the deformable boundary template represented by the shape model to true boundary in an input image is formulated by *Bayes rule.* Simulation results show that our method has more accurateness in case of multi-classes training set and performs better in the sense of computation cost than point distribution model(PDM)[2]. It works well in distortion under the noise, pose variation and some kinds of occlusions.

1. INTRODUCTION

Recently, various applications using image data are developed caused by powerful computational ability of computer, increase of data storage capacity, expansion of communication bandwidth and demands on multi-media service. Boundary that is connective edge contour describing object border is one of the most important information of the image data. Using the boundary, we can segment image into object region and background region, and establish the identity of an object. In the computer vision and recognition area, it is useful features describing an object, and it can reduce recognition error raised by background. Also it is necessary component for object-based coding in communication area. Because of internal edges of an object, background edges and occlusion by other objects, it is difficult to extract the boundary of interested object from image. Therefore many approaches have been proposed up to the present. We can divide these algorithms to two methodologies, knowledge-free and knowledge-based. In this paper, knowledge means geometric or statistical information about shapes of specific objects and it can be obtained from training set.

Knowledge-free algorithms based on clustering, morphological filtering and a watershed method segment homogeneous regions by means of local features such as gray level, texture and color. These local features may be useful, however these methods are hard to find complete boundary like closed loop. Kass[3] introduced deformable contours to model complex shape. When the energy function is designed appropriately, their work performs well even if the boundary is deformed a little. Knowledge-free algorithm employing local features can be applicable independently to object types, but it is sensitive to neighborhood edges, noise and occlusion because of having no processing to global shape.

The knowledge-based methods can generally find more accurate boundary of an object than knowledge-free methods, but need more computational cost. Wang and Staib[2] used PDM derived by the principal component analysis and employed the statistical shape distribution acquired from training sets. It is relatively robust to noise and initial parameters. However, because it is interested in only edge information, the local template may fluctuate near to the corner of an object, and that results in many iteration numbers to converge. The conventional knowledge-based methods have not managed multi-classes training set. That is, they have been applied when a training set is made from one class object or some objects having very similar shape each other. If multi-classes training set obtained from different objects is given and we don't know which one among the objects exist in an input image, the methods may not work well. The reason is that the shape statistics of multi-classes training set having large variation causes the basis vectors to have low energy compaction and the shape model represented by the basis vectors includes high reconstruction error. Also the determinant of initial condition is embarrassed with the ambiguity of object type in an input image.

In this paper, our approach uses global shape information obtained from training set to extract accurate boundary. To handle multi-classes objects, all training images are grouped by shape of boundary instead of identity. And then we derive the basis vector set from each group clustered to have minimum reconstruction error. The objective function for fitting the deformable boundary template represented by the shape model to true boundary of an object is formulated by *Bayes rule.* This function implies the probability that a shape similar to shapes of a group exists in an input image. So it can help to select one of basis vector sets obtained from training clusters

2. SHAPE MODELING

Because of the boundary deformation raised by noise and pose variation of object, we can not extract boundary by means of simply template matching between input image and each boundary image of training set. Besides, as the number of the training image increases, it needs heavier computational cost. So general knowledge-based methods represent the shape of boundary with feature vector S and then the feature is modeled by specific basis vectors as follows.

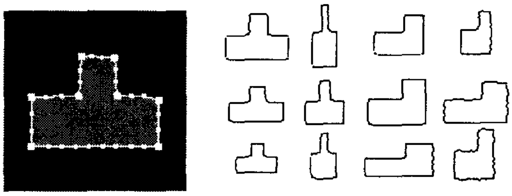
$$\mathbf{S} = \mathbf{a}_{initial} + \mathbf{B}_{basis}\mathbf{w}_{weight} \tag{1}$$

We represents a boundary of each shape of training sets as labeled points like figure 1(a). To find labeled points from a boundary, the critical points having high curvature are extracted from the boundary and then equally spaced points between the critical points are interpolated along the boundary. Each of M aligned training shapes is described by a position vector $\mathbf{L}^i=[x^i(1), y^i(1), x^i(2), y^i(2),\dots, x^i(N), y^i(N)]^T$ $(i=1,\dots,M)$, where N is the number of total labeled point in a training image.

Now, we have to determine what basis vector set is useful to boundary extraction. In image recognition area, it is designed for discrimination the different characteristic of an object from other objects. However, in boundary extraction algorithm where we find boundary in input image similar to one of objects in training set, it is used in reconstruction the original boundary with boundary of training set. So we model the shapes of training set using KL transform which has minimum reconstruction error. According to KL transform, any shape $\mathbf{L}$ in the training set can be approximated by using the mean shape $\bar{\mathbf{L}}$ and the first t eigenvectors $\mathbf{Q}=(\mathbf{q}_1|\mathbf{q}_2|\dots|\mathbf{q}_t)$ of covariance matrix about the mean.

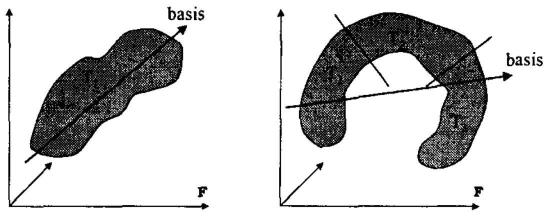
$$\mathbf{L} = \bar{\mathbf{L}} + \mathbf{Qb} \tag{2}$$

where $\mathbf{b}=(b_1,b_2,\dots,b_t)^T$ is a weighting vector which indicates how much variation is exhibited with respect to each of the eigenvectors. Equation (2) allows us to generate plausible shapes that are not part of the training set.



(a) (b)

Figure 1. (a) Boundary representation by labeled points. (b) Multi-classes training set including 2 object types



(a) (b)

Figure 2. Feature projection to basis vector

If shape features distribute linearly like figure 2(a), they are projected well to basis vector set. But multi-classes training set may follow nonlinear distribution like figure 2(b). That causes high projection error and high reconstruction error. In pattern recognition area, MKL proposed by Cappelli[1] tries to solve this problem. It clusters all training images into some groups and then finds separated KL basis for each cluster. To minimize projection error fitted to KL, it performs unsupervised clustering.

Using MKL, we expand the shape model to multi-classes training set. If all training boundary images are clustered into some groups, any boundary feature, $\mathbf{L}$, in training images can be approximated by mean and separated KL basis of each cluster.

$$\mathbf{L} \approx \bar{\mathbf{L}}_j + \mathbf{Q}_j\mathbf{b} \tag{3}$$

where j indicates a group of training images, $\bar{\mathbf{L}}_j$ and $\mathbf{Q}_j$ is mean and KL basis obtained from training boundaries belong to j-th cluster. Using this modeling, we define reconstruction error with respect to a $\mathbf{L}$ as follow.

$$Error(\mathbf{L}, j) = \left|\mathbf{L} - \left(\bar{\mathbf{L}}_j + \mathbf{Q}_j\mathbf{b}\right)\right| \tag{4}$$

From unsupervised clustering where each boundary of training images is projected to each separated KL basis set of training groups and it is moved into cluster having minimum error, basis vector sets of the proposed shape model is obtained like figure 3. Then a position vector $\mathbf{L}_j$ of boundary belong to j-th training group is modeled as follows.

$$\mathbf{L}_j = \bar{\mathbf{L}}'_j + \mathbf{Q}'_j\mathbf{b} \tag{5}$$

where $\bar{\mathbf{L}}'_j$ and $\mathbf{Q}'_j$ is obtained from j-th training image group after unsupervised clustering

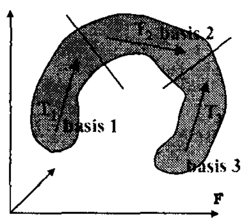


Figure 3. The proposed basis vector of shape model

To solve the fluctuation problem of PDM caused by using only edge information, we incorporate a statistical curvature model about curvature distribution into the above shape model. A curvature vector is denoted by $\mathbf{R}^i=[\, r^i(1), r^i(2),\ldots, r^i(N)]^T$ (i=1, 2,..., M), where $r^i(n)$ is curvature value of the n -th labeled point in the i-th training image. The mean curvature $\bar{r}(n)$ and the standard deviation $\sigma_r(n)$ are easily obtained from the alignment process of training images. Given these statistics of curvature, a curvature image is modeled as noise-corrupted version.

$$R = R_{\bar{\mathbf{L}}'_j+\mathbf{Q}'_j\mathbf{b}'} + n = R_{\mathbf{L}_j} + n \qquad (6)$$

where a curvature image R that represents the curvature value of each boundary pixel is approximated with a curvature image of the boundary $\mathbf{L}_j$ constructed by equation (5) and additive and independent noise. The noise is characterized by a standard deviation of curvatures in the training set. More detail shape model including curvature information as well as edge, can prevent the fluctuation near corner region.

3. OBJECTIVE FUNCTION

In order to apply the proposed shape model to the boundary extraction problem, we derive an objective function that is similarity measure between true boundary in input image and the boundary template generated by the proposed shape model. Deforming the local boundary template in the direction of maximizing the result of this function, we can fit the template to true boundary. Let $\mathbf{p}_j = (s,\theta,T_x,T_y,b_{j,1},b_{j,2},\ldots,b_{j,t})$ and $(\bar{x}_{j,n},\bar{y}_{j,n})$ denote a deformation parameter and mean position of j-th training cluster set respectively, by which the n-th mean position is deformed to $(x(\mathbf{p}_j,n), y(\mathbf{p}_j,n))$ as follows [2].

$$x(\mathbf{p}_j,n) = s\cos\theta\left(\bar{x}_{j,n} + \sum_{k=1}^{t} Q_{j,k}b_{j,k}\right) - s\sin\theta\left(\bar{y}_n + \sum_{k=1}^{t} Q_{j,k'}b_{j,k}\right) + T_x$$

$$y(\mathbf{p}_j,n) = s\sin\theta\left(\bar{x}_{j,n} + \sum_{k=1}^{t} Q_{j,k}b_{j,k}\right) + s\cos\theta\left(\bar{y}_n + \sum_{k=1}^{t} Q_{j,k'}b_{j,k}\right) + T_y \qquad (7)$$

where s, θ, and (T_x, T_y) are scale, rotation and translation parameters respectively, and $Q_{j,k}(Q_{j,k'})$ is $2n$-th($2n$+1-th) row and k-th(k'-th) column element of Q_j.

First, based on the position of boundary pixels, the degree of fitness between deformed boundary, $\mathbf{Lp}_j$, and true boundary can be represented as follows for a given deformation parameter $\mathbf{p}_j$ [2].

$$M_{\text{edge}}(\mathbf{p}) = \sum_{i=1}^{t+4}\left(-\frac{(p_{j,i}-m_{j,i})^2}{2\sigma_{j,i}^2}\right) + \frac{1}{\sigma^2}\sum_{n=1}^{N} E(x(\mathbf{p}_i,n), y(\mathbf{p}_i,n)) \qquad (8)$$

where m_i and σ_i is the mean and standard deviation of the i-th element in $\mathbf{p}_j$ respectively, and E is an edge image of input image.

Secondly, in order to take the input curvature image R and high curvature point distribution of training sets into consideration, the proposed method gets the curvature template $R\mathbf{p}_j$ for a particular value of $\mathbf{p}_j$, and then finds the maximum of $\Pr(R\mathbf{p}_j\,|R)$ over $\mathbf{p}_j$. This can be expressed as $\Pr(R|R\mathbf{p}_j)\cdot \Pr(R\mathbf{p}_j)/\Pr(R)$ by *Bayes Rule*. Since $\Pr(R|R\mathbf{p}_j)$ is equivalent to $\Pr(R=R\mathbf{p}_j+n)$ or $\Pr(n=R-R\mathbf{p}_j)$ in the statistical curvature model, the Bayesian objective function of curvature term is simplified as follows.

$$M_{\text{curvature}}(R\mathbf{p}_j) = \ln \Pr(R\mathbf{p}_j) + \sum_{k=1}^{N}\frac{1}{\sigma_r^2(k)} R(x(\mathbf{p}_j,k), y(\mathbf{p}_j,k))\bar{r}(k) \qquad (9)$$

where the probability density $\Pr(n)$ is assumed to be independent for each point and follow Gaussian with zero mean and standard deviation σ_n, and $R(x,y)$ is curvature value of (x,y) pixel. For each value of $R\mathbf{p}_j$, $\bar{r}(i)$ is set along the boundary represented by labeled points and the standard deviation of noise is substituted by the standard deviation of curvature obtained from training set.

Assuming the events about edge and curvature are independent each other, $\Pr(\mathbf{Lp}_j,R\mathbf{p}_j|\text{input image})$ is equivalent to $\Pr(\mathbf{Lp}_j|E)\cdot \Pr(R\mathbf{p}_j|R)$. From the above discussions, the combined Bayesian objective function over $\mathbf{p}_j$ is derived as follows.

$$M_{com}(\mathbf{p}_j) = 2\sum_{i=1}^{t+4}\left(-\frac{(p_{j,i}-m_{j,i})^2}{2\sigma_{j,i}^2}\right) + \sum_{n=1}^{N}\left(\frac{1}{\sigma_n^2}E\big(x(\mathbf{p}_j,n), y(\mathbf{p}_j,n)\big) + \frac{1}{\sigma_{j,r}^2(n)}R(x(\mathbf{p}_j,n), y(\mathbf{p}_j,n))\bar{r}_j(n)\right) \qquad (10)$$

By this analytic formulation, we can compute efficiently the optimum of the objective function $M_{\text{com}}(\mathbf{p}_j)$ that implies the probability that a shape similar to shapes of i-th training group exists in an input image, too.

4. EXPERIMENTAL RESULTS

For implementation of our algorithm, we obtained an edge image and a curvature image using Canny edge detector [4] and Curvature Scale Space [5] method from an input image. Also, we used the steepest decent methods to solve the maximum *a posteriori* problem of $M_{com}(\mathbf{p}_j)$. In experience, we calculated errors in each iteration time. The error means the distance between a boundary point on deformed template and the closest point on its true boundary.

To prove the performance of our shape model in multi-classed training images, we got 24 training boundary images like figure 1(b). Figure 4 shows that the proposed method using separated basis vector sets worked better than PDM derived from all training images in sense of accurateness and convergence speed, because the our shape model has lower reconstruction error. Where the basis vector set was selected by objective function among basis vector sets grouped by unsupervised clustering.

Figure 5 is the simulation result of our shape model including curvature and PDM interesting only edge in one class training images. The proposed method converged rapidly because the detection of high curvature points help local template to avoid the fluctuation near the corner.

In the figure 6, the sensitivity to noise was measured by adding different amounts of zero mean Gaussian noise to the synthetic images. It shows that our method is robust to noise and particular occlusion by other object.

When an infrared image is given as input, the experimental results of the proposed method and snake [3] are figure 7. As shown in the figure, it is very hard to extract boundary from the low-quality image like an infrared image without global information obtained from training set.

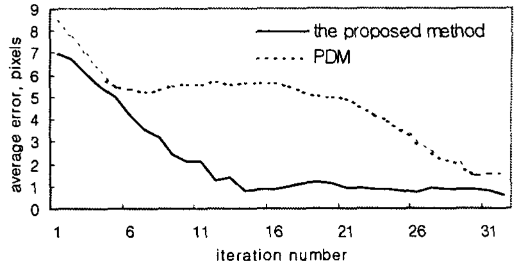


Figure 4. Boundary extraction

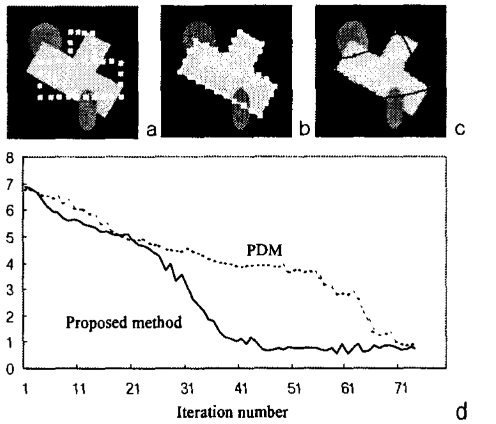


Figure 5. Iteration number for convergence, (a) Initial labeled points, (b) Final result by the proposed method (c) Final result of snake algorithm (e) Average error measure with respect to iteration numbers

5. CONCLUSION

We proposed a new boundary extraction scheme that can be applied to multi-classes training set.

The proposed method models shapes of different objects with multiple basis sets based on MKL, and curvature statistics is incorporated into the shape model. The objective function that is similarity measure between true boundary in input image and the boundary template generated by the shape model is formulated by *Bayes rule*.

By experimental results, we have shown faster convergence of the objective function and more accurate extraction than PDM. Also it was proved to be robust to noise, occlusion, and background edges.

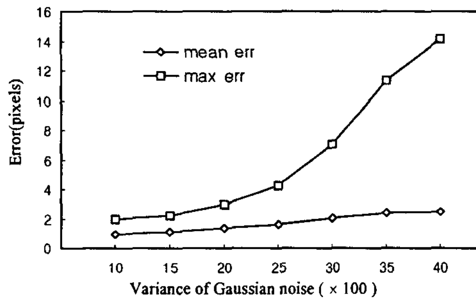


Figure. 6. Noise robustness experiment

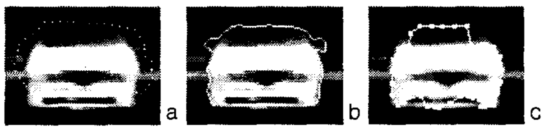


Figure 7. The results of infrared images containing vehicle, (a) input image and initial contour, (b) proposed method. (c) snake algorithm

6. REFERENCES

[1] R. Cappelli, D. Maio, and D. Maltoni, "Multispace KL for Pattern Representation and Classification," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 23, No. 9, pp. 977~996, September 2001

[2] Y. Wang, L.H. Staib, "Boundary Finding with Prior Shape and Smoothness Methods," *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 22, No. 7, pp. 738~743, July 2000

[3] M. Kass, A. Witkin, D. Terzopoulos, "Snakes : Active Contour Model", *International Journal of Computer Vision*, pp. 321 ~ 331, 1988.

[4] J. F. Canny, "a Computational Approach to Edge-detection," *IEEE Trans. Pattern Anyaysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679-698, 1986.

[5] F. Mokhtarian and R. Suomela, "Curvature Scale Space for Image Point Feature Detection," *in Proceedings of 7th International Conference on Image Processing*, Vol. 1, pp. 206-210, 1999.

[6] H. C. Choi and and S. D. Kim, "Boundary Extraction Using Statistical Shape Descriptor," *IEE Electronics Letters*, Vol. 38, No. 22, pp. 1330-1332, October 2002.