Object Detection Based on Two Level Fast Matching

Yumin Dou, Mao Ye*, Pei Xu, Lishen Pei and Zhenghua Liu

*School of Computer Science and Engineering Center for Robotics,
University of Electronic Science and Technology of China
Chengdu 611731, P.R. China
cvlabuestc@163.com*

Abstract

Shape template matching is an important approach in object detection and recognition. In this paper, we propose a fast and novel object detection method, which represents edge map contours with salient points and retrieves the target object by using a backtracking method with two stages from coarse matching to fine matching. Our method has two main contributions. One is we propose the way to represent contour structure with salient points. The other is that the object detection method proposed can directly operate on the real images, which improves its practicability. According to the experimental results, the proposed approach improves the speed than previous works.

Keywords: Salient Point, Corner Point, Distance Image, Backtracking

1. Introduction

Object detection is still a difficult task, which has attracted increasing research attention, and the fast and accurate detection is the essential factor. Many target class detection approaches have been proposed, which always have the following procedure. Firstly, the feature filter is obtained by a certain method, then a sliding window is conducted on the query image and a similar measure is given between the feature filter and window patch. The approach of extracting the feature filter includes the training-based or the template matching based.

For training-based methods, the feature filter is learned from many samples. For example, texture or gradient features which are extracted from the images are used to learn the general tolerance terms [1, 7, 13]. Although the features based on texture or gradient are useful to real applications, the texture-based method can't have a better performance when the target is texture-less or no textured. Addressing this problem, the researchers [26, 27] presented a learning filter based on local contour features which was used to detect object. Ferrari *et al*. used a contour segment network to represent line segments to retrieve image in [10] and [11].

However a classifier which is used to detect object in traditional learning methods [18, 29] usually need more hundreds of samples. In real world application, sometimes it is an impossible and boring work to collect lots of samples, such as in airport security checking, and image indexing on the internet.

To overcome the above limitation, researchers proposed the non-training method, template matching method, which can be categorized into two classes, *i.e*. the feature template matching and the shape template matching. For the forehand, in [38], the authors proposed to extract features like larks, hog from example image as template. The latter, a hand-drawn contour is used as the matching shape template, which is useful and manual to real applications. Contour is important information to represent image structures in machine vision and object recognition. Compared to other image cues *e.g*. texture, feature point etc; the outline contour is invariant to lighting conditions and variations in object color and texture.

Researchers proposed template matching approaches previously in [34-37]; unfortunately, they will suffer severe degradation of performance or even failure, when they are placed in the cluttered scenes. In [32, 33], the authors consider only images and their gradients to seek the target instances and also can't robust to the strong cluttered background.

In recent years, there are seldom works on this direction. Hinterstoisseri *et al*. proposed a robust encoding of gradient information of edge map [31], which delivers fast matching time. However, it cannot show a more desired result in cluttered scenes. In [19], Liu *et al*. demonstrated a fast directional chamfer matching (FDCM) algorithm, which improved the matching accuracy with chamfer distance from multiple directions and reduced its computational cost significantly at the same time. However, when confronting with the cluttered background, the computational costs increase drastically. It is because their approach depends on the method of extracting line segments at the first stage.

Motivated by the above observation, we propose an object detection approach which includes two stages, the coarse stage and the fine stage. At the coarse stage, a set of detection hypotheses are obtained by using a conventional $2D$ distance image of middle points to match the template. At the fine stage, we use a shape descriptor which describe the relative spatial arrangement and compare the similarity between structures of detection hypothesis with the shape template (in Figure1).

Our main contributions can be summarized as follows:

Firstly, we propose the way to represent contours structure with salient points, which utilize the corner extraction technique and ellipse support region model. Secondly, according to the Euclidean distance image index, we can track back the line segment in accord with the salient point and conduct a fine matching processing. Experiments have shown that the detection method proposed can directly operate on the real image, which improves its practicability and is robust to strong clutter scenes.
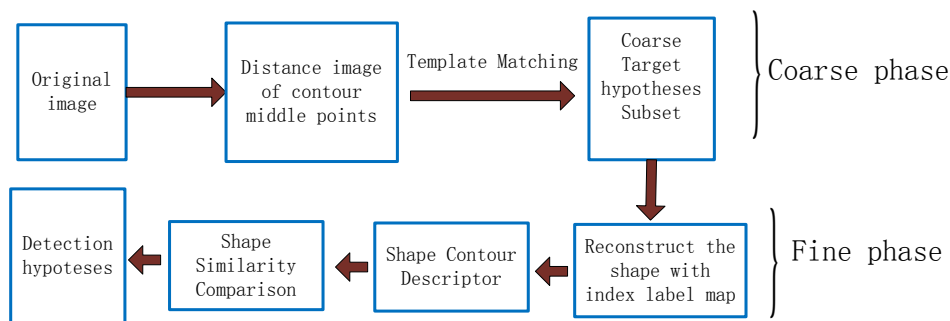
Figure 1. Top Row is the Coarse Phase; Low Row is the Fine Phase

In the following, the method of extracting salient points is first introduced in Section 2. The details of shape matching approach with index are described in Section 3. The experiment results are presented in Section 4. Finally, conclusions are given in Section 5.

2. Course Matching Phase

Accurate detection target in images is not an easy work. In this section, our task is to find the target instances by using the 2D object function in images. After the preliminary filtering, we can obtain a target instance hypothesis subset. The processing procedure includes three stages. Firstly, we extract the contours in images, and seek the key points which can represent the contours structure effectively. Secondly, the query contour images are approximated by small line segments which connected with the key points. Then, the middle points in line segments are transformed into Euclidean distance images

in Figure7. Lastly, a traditional sliding window approach is used to compare the similarity with shape template.

2.1. Salient Point Extraction

In this paper, the salient points include corners and ellipse vertexes [20]. An ellipse support region model is used to represent image edges based on corner point extraction. The corner points are extracted from image edges by using a fast corner detector based on the chord-to-point distance accumulation(CPDA) technique [3,12] because CPDA technique does not use any derivatives and selects the chord-length L that is less sensitive to local variation and noise on the curve and can conduct multi-scale detection. There are three phases for extracting the corner points. Firstly, an edge extraction and selection process are performed, that is to say, planar curves are extracted by using the Canny edge detector from the grey image. Secondly, Gaussian convolution operations with different variances are conducted on the curves to smooth it. Thirdly, accumulation curvature estimation is the critical factor, which is used from a moving chord along a curve to each point on the curve.

The improvement of CPDA is that obtained a set of probable candidate points before CPDA curvature estimation. Choosing the critical point is the basic step. Awrangjeb *et al*. find the difference of the distance between a point on an original curve to the corresponding point on the smoothed curve, when an original curve is smoothed by Gaussian [3]. As can be seen at left most in Figure 2, when the point is nearer the position of corner, the distance becomes higher and vice versa. When the corner points were obtained, the next step is to find the ellipse vertexes in the following subsection.
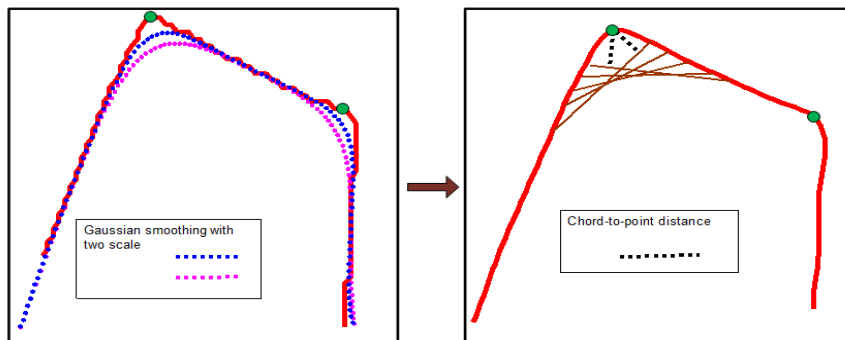

Figure 2. Flow Chart of CPDA

2.2. Ellipse Support Region Model

Due to the disturbance of noise, only single point whose curvature is the maximum is often not the vertex of curve fragment. To find the vertex of curve fragment robustly, we utilize the curvature support region (Figure4), actually the area magnitude of the curve fragment curvature represents the changes of mean curvature. Support region is the area encircled with curvature curve and the axis.

In different situations, according to the rate of major axis and minor axis, we set a constant value, through which we can decide whether to select the curve vertex (Figure3). As a result, we can make sure whether the mean curvature larger than the given threshold. If it is true, there is a true vertex, otherwise the curve fragment is regarded as a straight line fragment, and that is to say, there is not vertex.

In this section, we consider the global characteristic of the curve fragment, and then the contours can be approximate better by connecting these points. We utilize the curve fragment between the adjacent corners as support region [14-15] (in Figure4). When the sum of point curvature in support region is lower than a fixed threshold, there will

produce an ellipse vertex between these two corner points along the corresponding edge. From [14-15], the curvature of an elliptic curve in edge map can be obtained as follow,

$$K(x) = \frac{ba^4}{((bx)^2 - (ax)^2 + a^4)^{3/2}} \tag{1}$$

When set $K_{max} = K(0) = \dfrac{b}{a^2}$, $K_{min} = K(a) = \dfrac{a}{b^2}$, the support region of curvature is as follow,

$$\int K(x)dx = \frac{ba^4}{((bx)^2 - (ax)^2 + a^4)^{1/2}} \tag{2}$$

And we have

$$\overline{K} = \int_{-a}^{a} k(x)dx / 2a ,$$

$$T = R \times \overline{K} = R \times K_{max} \cdot a / b \tag{3}$$

Which can reflect the varies of curvature of the edge curve robustly in accordance, because we consider the global characteristic of the edge curve from the formula (2).

The rate between major axis and minor axis reflects the degree of bending. From (3), $\overline{K}$ is the mean curvature which is obtained by computing curvature of the whole points on the curve fragment, R is a constant value. We find the relation of the rate with the mean curvature of the curve fragment, and transform the single point curvature computing into the mean curvature computing of curve fragment, which enhance the robustness of processing procedure [14-15]. We can set a threshold T, comparing to the value of the mean curvature. When it is large than 0.167 (Figure 3), the curve fragment is approximating to a straight line fragment, we think that there are not vertex. Otherwise, there are.

After these salient points (ellipse vertexes and corner points) are found out in Figure 5, the shape contours can be approximated by connecting these points into line segments. Then we seek the middle points on these line segments and compute the distance image.
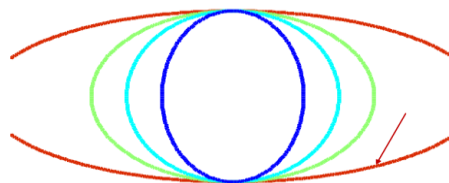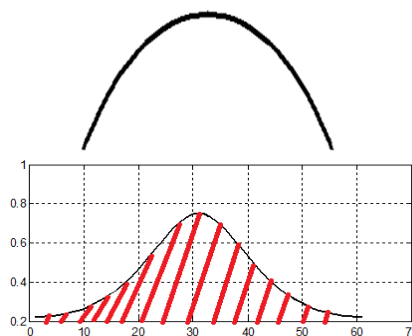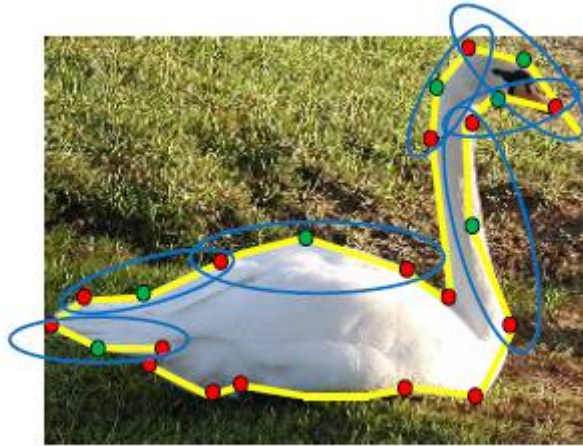

Figure 3. Ellipse Curvature

Figure 4. Support Region (Red). Top: Curve Vertex Candidate; Low: Curve Curvature

Figure 5. Finding Ellipse Model Vertex (Green points) Red Points are Corner Points

2.3. Distance Transform

A sequential algorithm for computing the exact Euclidean distance [22,25-26] is a popular and robust distance measurement method to align two maps, which has linear time complexity and can be used for computing the DT for a wide class of distance functions, *e.g.* L_p and chamfer metrics. Let $W = \{w_i\}$ and $Z = \{z_j\}$ be the sets of edge map points. The distance between these two sets is

$$dist(W, Z) = (\sum_{i=1}^{k} |w_i - z_j|^2)^{1/2} \tag{4}$$

Though the chamfer DT is substantially faster to compute, we choose the method of computing Euclidean distance because it is reversible, that is to say, given a pixel position in distance image, we can obtain the index label position which can be retained in computing the Euclidean distance. In some sense it is a one-to-many mapping. When compute the distance image, we can retain the index of middle points in the label map. One index which is nearest to the pixel position in contour image may map many pixels in distance image in Figure 6. Sometimes the exact Euclidean distance is required for some applications, because only the EDT can produce an accurate metric and allows exact reconstruction of the object. When the distance metric is the L2 metric (Euclidean distance), then the procedure can be implemented simply using only integer arithmetic. In Figure 7 right most, a distance image is generated quickly which is only with one fifth ms [22], and at the same time we can obtain an index label map corresponding to the each pixel of the distance map.

We conduct on Euclidean distance map by using the traditional template matching method and get many detection hypotheses at the first phase. When there are many bounding box to meet the shape, we performed non-maxima suppression by retaining only the lowest cost hypotheses among the ones which had significant overlap.
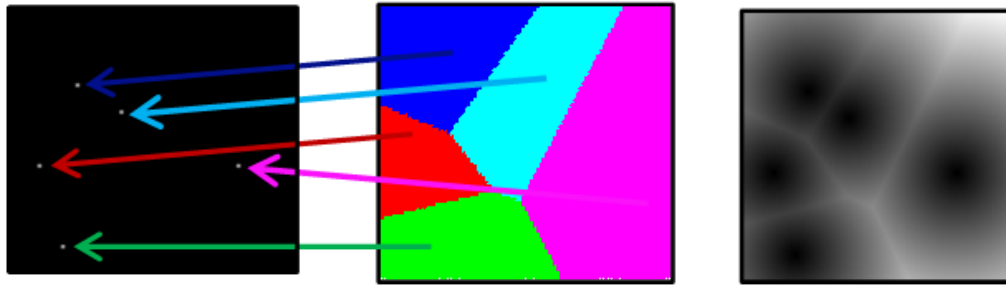
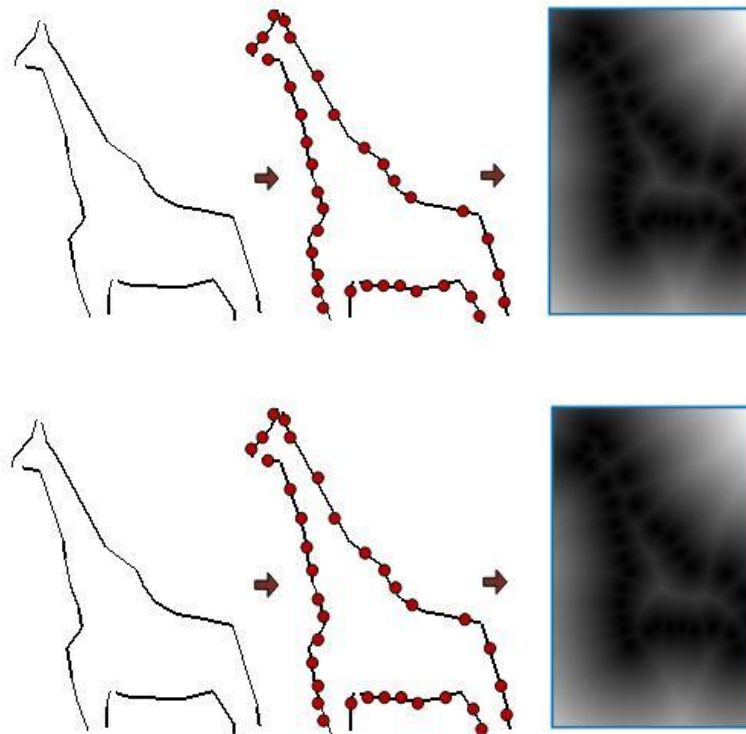**Figure 6. Left: Original Image; Middle: Label Map; Right: Distance Map**

Figure 7. Shape Line Segments to Red Points, Then Red Points to Distance Image. Left are Shape Line Segments. Middle: Red Points are Middle Points of the Line Fragments. Right: Euclidean Distance Image

3. Fine Matching Phase

Fine matching will conduct on the target hypotheses subsets which have been obtained at the coarse matching phase. Each target hypothesis is a matching window. When the matching window set was obtained, at template matching position in the window, according to the index label map, we track back the middle points which generate the distance at the position and the salient fragments in the original image (Figure 6). These fragments can be used to reconstruct a shape which will be utilized to fine matching at the second phase. We use shape contour descriptor composed of these points to match the shape template at the fine stage by measuring the similarity with the template. Consider to the structure relations of the salient points, there are two components at the fine phase, shape contour descriptor and shape similarity comparison [6, 8].

3.1. Shape Contour Descriptor

In this section, the first task is to represent the shape with salient line segments which are obtained by looking up the index of middle points in the label map. At this phase, we obtained a subset of the detected windows which are filtered by the first coarse phase, and each is composed of a set of middle points obtained from line segments. Then from these middle points we will track back to the original line segments. The start or end point of each line segment is salient points. As a result, various contours which are composed of the line segments are used to compute the similarity with the shape template at the second phase. Then in the fine phase we can see that a target object detection task is formulated as a matching problem.

Because the salient points are prominent to imitate the structure of the object outline contour than uniform sampled points and, it is not necessary that several sampled points on the same straight line, we will treat the salient points instead of the sampled points. As can be seen, M. Donoser *et al.* used the sampled points with uniform distribution in [8] hence it is different from [8] in this article.

Our shape descriptor is also based on angles which describe the relative spatial arrangement of the salient points. We calculate an angle $\angle P_j O_k P_i$ between two chords $\overline{O_k P_j}$, $\overline{O_k P_i}$ from a reference point O_k. We denote $\angle P_j O_k P_i$ as β and there are a fixed number of points between P_j and P_i. Firstly, O_k is fixed, only P_j and P_i point moves clockwise with β varies constantly changing. Secondly, O_k move clockwise one position, P_j repeat the first step. Finally, we obtain a $(n-1) \times (n-1)$ matrix (see Figure8), as follow

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1,N-1} \\ \vdots & \ddots & \vdots \\ a_{N-1,1} & \cdots & a_{N,N-1} \end{pmatrix}$$

As can be seen that the two matrix maps are similar as shown in Figure 9. It is intuitive that our matrix implies the relations between each salient point as reference point on the shape outline contours with the other points. As a root point, the reference point is circled by all the other points.

A reasonable similarity measure is important except the consensus matching point sequences. In this paper, we apply each point (from 1 to n) on the shape as a start center, and connect the point with other $n-1$ points, we can obtain a serial angle differ from [8], these angles form into a matrix (see Figure8).

3.2. Shape Similarity Comparison

In order to find the optimal matching shape, we use the Pareto framework which is an optimal approach of multiple targets based on part and dissimilarity metric. Due to the cause of occlusion and deformation, the global shape descriptor cannot describe the similarity of target shape perfectly. So the partial matching would be a good selection.

In this paper, we compute the similarity of two dimensional non-rigid shapes with Pareto framework [6, 8]. The key is to choose an appropriate metric scheme which can find the optimal combination between part and dissimilarity metric. Given a two dimensional object shape, we denote X' as its part. The area of the object shape is denoted by the measure S_x. We interpreted the partiality as the portion of the area of the

selected parts. According to the partiality, we use the Gromov-Hausdorff distance [8] as the dissimilarity. Therefore, we can see that the Pareto distance $dist_s(X,Y)$ describes the tradeoff between the partiality cut from the objects and the dissimilarity.
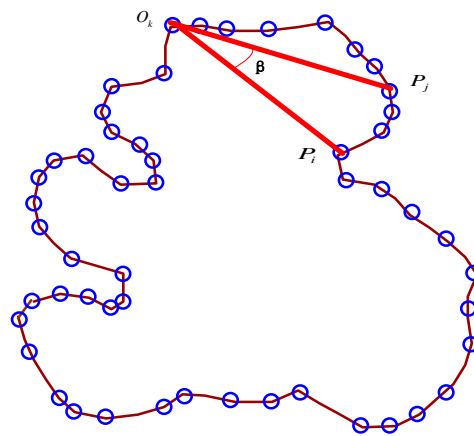
Figure 8. Shape Descriptor is based on Salient Points

When the two matrix map obtained in Figure 9, we will find the best matching similarity between template and target edge map. We use the Pareto framework to have the multiple optimal processes [6, 8]. Two quantities are defined. One is partiality $\lambda(X,Y)$, which describes the length of the parts. The other is dissimilarity $\varepsilon(X,Y)$, which measures the dissimilarity between the two parts, where X and Y are two contour parts of the shape, when the part is smaller with the value higher. Our task is to find the minimum descriptor difference for all partialities, and decide whether it is our target object [6, 8]. Finally, transfer a vector to a scalar value, in order to find the optimum value. We use a Salukwadze distance [6] based on the Pareto frontier by

$$\Phi(X^*,Y^*) = (\lambda(X^*,Y^*), \varepsilon(X^*,Y^*)) \tag{5}$$

$$dist_s(X,Y) = \inf_{(X^*,Y^*)} |\Phi(X^*,Y^*)|_1 \tag{6}$$

Where $|.|_1$ is the L1-norm of the vector.

4. Experiments and Discussions

To evaluate the approach proposed, we conducted two sets of experiments on challenging real scene datasets. The detection algorithm proposed was executed 100 times on a 2.66 GHz PC with 2 G byte memory. Note that, in all our experiments we emphasize the speed and the improved accuracy of our approach compared to FDCM.

ETHZ dataset consists of five different classes: applelogos, bottles, giraffes, mugs, and swans, which contains total of 255 images. Each image contains one or more objects from each class respectively. In this dataset, the objects have various variations in viewpoint, appearance, size, and non-rigid deformation. In our work, we firstly put the template processing into a middle point set. Then the same process to the detection image. We transform the image into $2D$ Euclidean distance image, and use the traditional sliding window approach to obtain a subset of matching windows, which reduce one dimension cycle. In the second stage, we conduct the partial similarity match algorithm on the subset. Due to the first stage filter process, the subset become very small, we only

need a trivial computational time. From Figure 10, we can see an exact detection in the image. Each image shows detected targets for one or more detections.
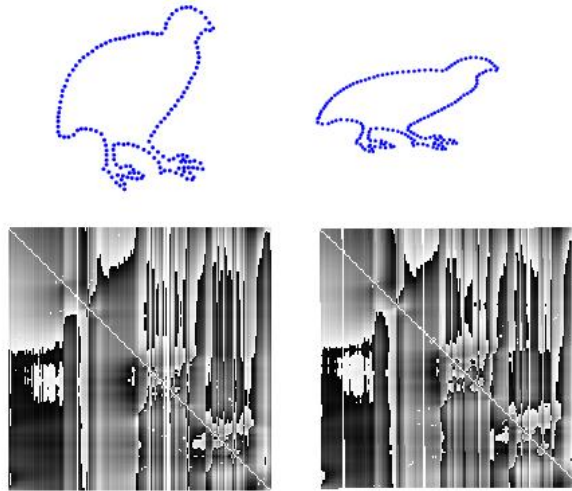
Figure 9. Shape Similarity Comparison

To quantitative evaluation, we use detection rate (DR) vs. false positive per image (in Figure11). We compare our method with the popular contour based object detection methods [5, 9-10]. All these methods use the traditional training method or template matching method, *e.g*. FDCM.

Among the compared methods, we have achieved best detection rate for category Giraffes, which is the most difficult class in this dataset. Our method achieves a comparable result to [5, 9, 10] for category Swans, Bottles, Apple logos and Mugs (in Figure 11).

Table 1 shows the comparison of detection time of each class in ETHZ dataset. Top two rows show the detection time of FDCM and proposed respectively and, the third row is the improved detection rate of each class and total five classes. We can see the detection time have been improved differently in the third row, the minimum is 5.78%, the maximum is 34.90% and the total average rate of five classes is 16.46%.

Table 1. Detection Time (Seconds) Comparison of the Average of each Class and Total Average of Five Classes

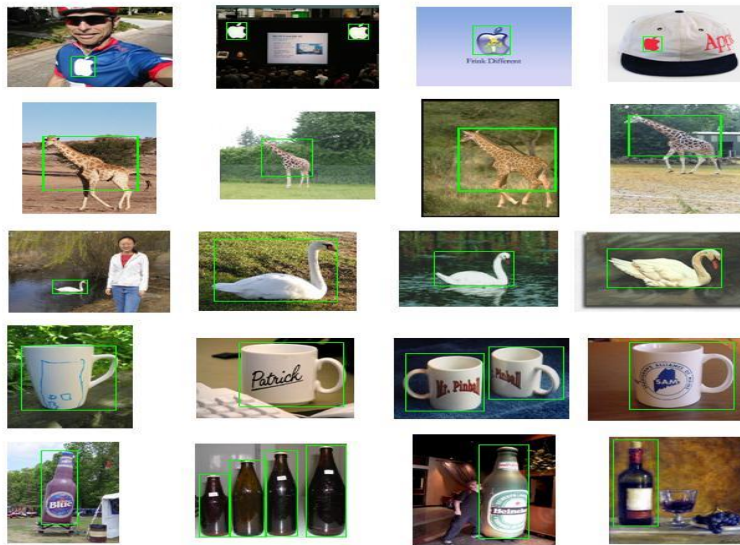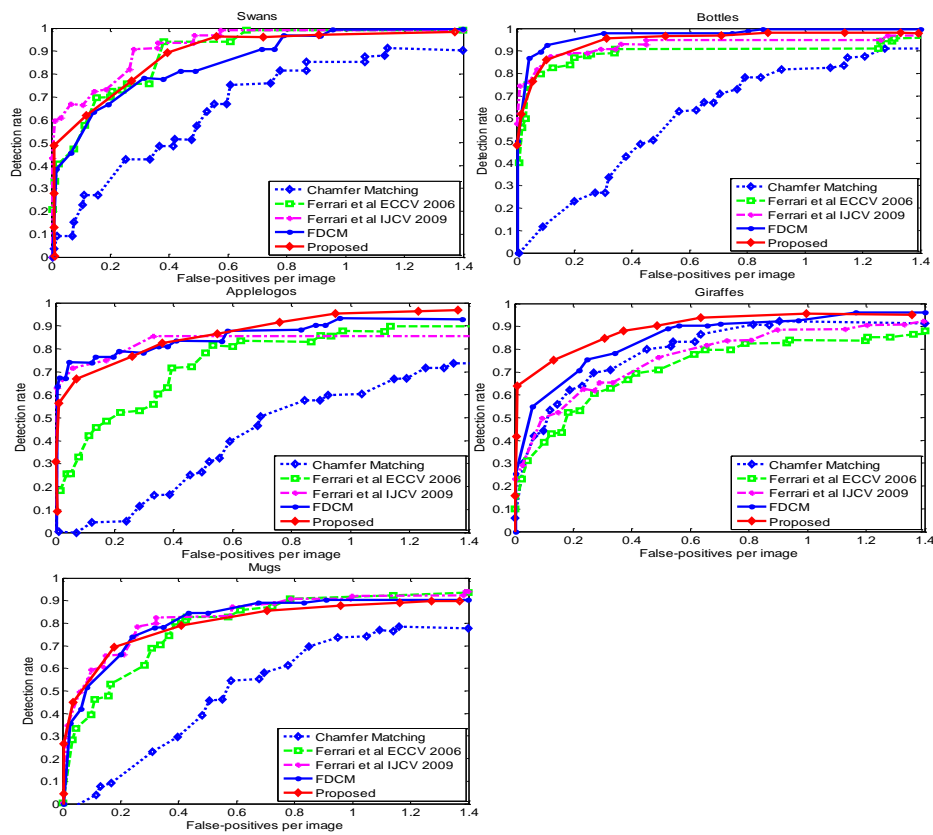|	Bottles	Applelogos	Giraffes	Mugs	Swans	Average (total)
FDCM	5.6095	14.4040	8.0431	6.9475	6.8595	8.3727
Proposed	5.2851	11.3265	7.2432	6.2308	4.4652	6.8573
Improved rated	5.78	21.37	9.95	10.32	34.90	16.46 (%)

Figure 10. Target Object Detection

Figure 11. FPPI of the Comparison Results on ETHZ

INRIA horse dataset, which consists 170 images with one or more horses viewed from the side. Horses appear in images at several scales, various perspectives and against occlusions and cluttered backgrounds. Horse's action changes are more than the ETHZ database classes. Since Liu *et al*., do not provide the detection result information on this dataset in [19], we will make a comparison with other state of art methods, M2HT+IKSVM and KAS in [21]. Note that to seek the salient points, we conduct on this experiment with the real images directly rather than edge images of the dataset. However,

it is an impossible task to extracting line segments by using RANSAC fitting algorithm on the real image, because it is more sensitive to the surrounding noise when the number of trivial line segments is increasing. The extracting salient point method is more stable in the whole processing phase either on real image or on edge map, because it avoids looking around for the fitting point as RANSAC algorithm. As shown in Figure 12, there are the detection results and Figure 13 shows that our approach is comparable with M2HT+IKSVM and significantly better than KAS [21], even we need not to have a training process.

Figure 12. Object Detection of Salient Point Representation Horse

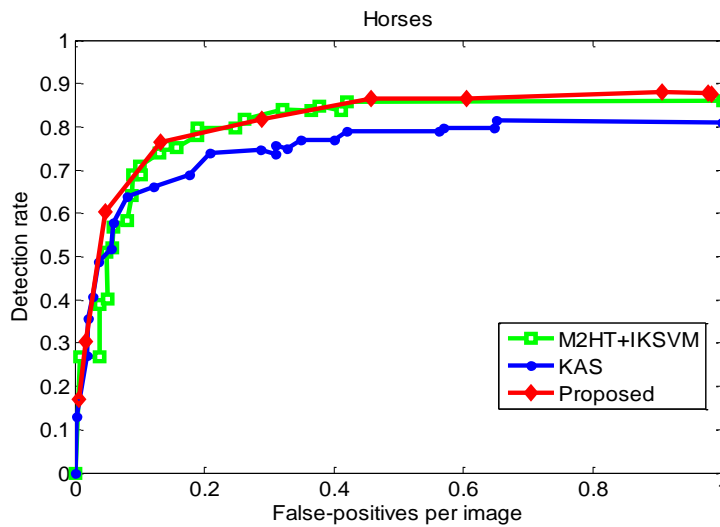**Figure 13. ROC Curve and Several Localization Results on the INRIA Horses Dataset**

5. Conclusion

In this paper, we have proposed an improved and efficient approach to detect target object based on the FDCM [19], which is an improved two round solution based on distance matching map, and then used an fast multiple object optimal method to partial shape matching. The experimental results show our approach is very efficient and

practical. In our work, we creatively introduce a detection method of two level stages. Firstly we demonstrate a matching process of salient point Euclidean distance image; on the other hand we obtain the detection instances on shape structure image by using Pareto optimal method. In conclusions, our method makes target detection more usable in many computer vision applications. In the future work, we will apply the approach to object outline contour detection and target boundary location.

Acknowledgments

References

[1]   S. An, P. Peursum, W. Liu and S. Venkatesh, "Efficient Algorithms for Sub-window Search in Object Detection and Localization", IEEE Conf. Computer Vision and Pattern Recognition, Miami, FL,USA, June 20-25, **(2009)**.

[2]   M. Awrangjeb and G. Lu. "An improved curvature scale-space corner detector and a robust corner matching technique for transformed image identification", IEEE Trans. Image Process, vol. 12, no. 17, **(2008)**.

[3]   M. Awrangjeb, G. Lu, C. S. Fraser and M. Ravanbakhsh, "A Fast Corner Detector Based on the Chord-to-Point Distance Accumulation Technique", 2009 Digital Image Computing: Techniques and Applications, Melbourne, VIC, December **(2009)**, pp. 519–525.

[4]   M. Awrangjeb and G. Lu, "Robust image corner detection based on the chord-to-point distance accumulation technique", IEEE Trans. Multimedia, vol. 6, no. 10, **(2008)**.

[5]   H. Barrow, J. Tenenbaum, R. Bolles and H. Wolf, "Parametric correspondence and chamfer matching: Two new techniques for image matching", In 5th Int. Joint Conf. on Artif. Intel, San Francisco, CA, USA, January **(1977)**.

[6]   A. Bronstein, M. Bronstein, A. Bruckstein and R. Kimmel, "Partial similarity of objects or how to compare a centaur to a horse", International Journal of Computer Vision, vol. 2, no. 84, **(2008)**.

[7]   G. Csurka, C. R. Dance, L. Fan, J. Willamowski and C. Bray, "Visual categorization with bags of key-points", Proc. European Conf. Computer Vision, Prague, May 11-14, **(2004)**.

[8]   M. Donoser, H. Riemenschneider and H. Bischof, "Efficient partial shape matching of outer contours", ACCV, Xi'an, China, September 23-27, **(2009)**.

[9]   V. Ferrari, L. Fevrier, F. Jurie and C. Schmid, "Groups of Adjacent Contour Segments for Object Detection", PAMI, vol. 1, no. 30, **(2008)**.

[10] V. Ferrari, F. Jurie and C. Schmid, "From images to shape models for object detection", Intl. Journal of Comp. Vis., vol. 3, no. 87, **(2009)**.

[11] V. Ferrari, T. Tuytelaars and L. Van Gool, "Object Detection with Contour Segment Networks", ECCV, Graz,Austria, May 7-13, **(2006)**.

[12] J. H. Han and T. Poston, "Chord-to-point distance accumulation and planar curvature: a new approach to discrete curvature", Pattern Recognition Letters, vol. 10, no. 22, **(2001)**.

[13] R. Huang, T. Li, M. Ye and Y. Dou, "Unconstrained Face Verification by Optimally Organizing Multiple Classifiers", International Journal of Control, Automation and Systems, vol. 4, no. 12, **(2014)**.

[14] X. C. He and N. H. C. Yung, "Curvature Scale Space Corner Detector with Adaptive Threshold and Dynamic Region of Support", Proceedings of the 17th International Conference on Pattern Recognition, Cambridge, August 23-26, **(2004)**.

[15] X. C. He and N. H. C. Yung, "Corner detector based on global and local curvature properties", Optical Engineering, vol. 5, no. 47, **(2008)**.

[16] F. Jurie and C. Schmid, "Scale-invariant Shape Features for Recognition of Object Categories", CVPR, Washington, DC, USA, June 27-July 2, **(2004)**.

[17] M. Leordeanu, M. Hebert and R. Sukthankar, "Beyond local appearance: Category recognition from pairwise interactions of simple features", CVPR, Minnesota, USA, June 18-23, **(2007)**.

[18] T. Li, M. Ye and J. Ding, "Discriminative Hough context model for object detection", The Visual Computer, vol. 1, no. 30, **(2014)**.

[19] M.-Y. Liu, O. Tuzel, A. Veeraraghavan and R. Chellappa, "Fast Directional Chamfer Matching", Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Francisco, CA, June 13-18, **(2010)**.

[20] E. Loupias, N. Sebe, S. Bres and J.-M. Jolion, "Wavelet-based Salient Points for Image Retrieval", 2000 Proceedings International Conference on Image Processing, Vancouver, BC, Canada, September 10-13, **(2000)**.

[21] S. Maji and J. Malik, "Object Detection using a Max-Margin Hough Transform", CVPR, Miami, Florida, USA, June 20-25, **(2009)**.

[22] M. Calvin, R. Qi, and V. Raghavan, "A Linear Time Algorithm for Computing Exact Euclidean c Transforms of Binary Images in Arbitrary Dimensions", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 2, no. 25, **(2003)**.

[23] F. Mokhtarian and A. K. Mackworth, "Scale-based description and recognition of planar curves and two-dimensional shapes", IEEE Trans. Pattern Anal. Mach. Intell., vol. 1, no. 8, **(1986)**.

[24] F. Mokhtarian and R. Suomela, "Robust image corner detection through curvature scale space", IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 20, **(1998)**.

[25] P. David, "Distance Transforms: Properties and Machine Vision Applications", Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing, vol. 1, no. 54, **(1992)**.

[26] R. Azriel and J. Pfaltz, "Sequential operations in digital picture processing", Journal of the Association for Computing Machinery, vol. 4, no. 13, **(1966)**.

[27] J. Shotton, A. Blake and R. Cipolla. "Contour-Based Learning for Object Detection", ICCV, Beijing, China, October 17-21, **(2005)**.

[28] J. Shotton, A. Blake and R. Cipolla, "Multi-scale categorical object recognition using contour fragments", IEEE Trans. on Pattern Anal and Mac. Intel., vol. 7, no. 30, **(2008)**.

[29] H. Wang, M. Ye and S. Yang, "Shadow Compensation and Illumination Normalization of Face Image", Machine Vision and Applications, vol. 6, no. 24, **(2013)**.

[30] X. Zhang, M. Lei, D. Yang, Y. Wang and L. Ma, "Multi-scale curvature product for robust image corner detection in curvature scale space", Pattern Recognition Letters, vol. 5, no. 28 **(2007)**.

[31] S. Hinterstoisser, C. Cagniart, S. Ilic, P. Sturm, N. Navab, P. Fua and V. Lepetit, "Gradient response maps for real-time detection of texture-less objects", PAMI, vol. 5, no. 34, **(2012)**.

[32] S. Hinterstoisser, V. Lepetit, S. Ilic, P. Fua and N. Navab, "Dominant Orientation Templates for Real-Time Detection of Texture-Less Objects", IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, CA, June 13-18, **(2010)**.

[33] M. Muja, R. Rusu, G. Bradski and D. Lowe, "Rein - a Fast, Robust, Scalable Recognition Infrastructure", in International Conference on Robotics and Automation, Shanghai, China, May 9-13, **(2011)**.

[34] G. Borgefors, "Hierarchical Chamfer Matching: a Parametric Edge Matching Algorithm", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 6, no. 10, **(1988)**.

[35] D. Huttenlocher, G. Klanderman and W. Rucklidge, "Comparing Images Using the Hausdorff Distance", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 9, no. 15, **(1993)**.

[36] D. Gavrila and V. Philomin, "Real-Time Object Detection for "smart" Vehicles", in IEEE International Conference on Computer Vision, Kerkyra, September 20-27, **(1999)**.

[37] C. F. Olson and D. P. Huttenlocher, "Automatic Target Recognition by Matching Oriented Edge Pixels", IEEE Transactions on Image Processing, vol. 1, no. 6, **(1997)**.

[38] P. Xu, M. Ye, X. Li, L. Pei and P. Jiao, "Object Detection Using Voting Spaces Trained by Few Samples", Optical Engineering, vol. 9, no. 52, **(2013)**.

Authors

Yumin Dou, received his master's degree in Computer Science from Chengdu University of Technology in 2010. He is currently a Ph.D. candidate at the School of Computer Science and Engineering of University of Electronic Science and Technology University of China. His research interests include pattern recognition and computer vision.

Mao Ye, received his Ph.D. degree in Mathematics from Chinese University of Hong Kong, in 2002. He is currently a professor and Director of CVLab at University of Electronic Science and Technology of China. His current research interests include machine learning and computer vision. In these areas, he has published over 70 papers in leading international journals or conference proceedings.

Pei Xu, received his BS degree in computer science and technology from SiChuan University of Science and Engineering, ZiGong, China, in 2008 and his MS degree in condensed matter physics from University of Electronic Science and Technology of China, Chengdu, China, in 2011. He is currently a PhD student in University of Electronic Science and Technology of China, Chengdu, China. His current research interests include machine learning and computer vision

Lishen Pei, received the B.S. degree in computer science and technology from Anyang Teachers College, Anyang, China, in 2010. She has been taking the successive master-doctor program since September 2010. She is currently a Ph.D. candidate in University of Electronic Science and Technology of China, Chengdu, China. Her current research interests include action detection and action recognition in computer vision.

Zhenghua Liu, received his BS degree in computer science from Anhui University of Technology, Maanshan, China,in 2012.He is currently a postgraduate student in University of Electronic Science and Technology of China,Chengdu,China. His current research interests is image retrieval.