

Object Detection by Matching Salient Line Segments

Yumin Dou, Pei Xu, Mao Ye*, Renjie Huang and Tao Li

*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 method to represent edge maps by using salient line segments, which is used to detect objects based on shape template matching. Firstly, image edges are computed and, on these image edge fragments, corner points are extracted. Then, a parabola model is proposed to represent image edges. Secondly, based on salient points, a directional chamfer matching framework is used to compute the similarity between template image and corresponding locations in target image. Our method has two main contributions. One is we apply the way to represent contours with salient points to the framework, which is used to detect the target rapidly and accurately. The other is that our method can directly operate on real images, which improves its practicability. A series of experiments are performed on two benchmark data sets. Comparing with previous work, our method is much more time-saving and memory-saving.

Keywords: *Salient point, Chord-to-point, Chamfer distance image*

1. Introduction

Outline contour (silhouette) can efficiently represent image structures with large spatial extents. Compared to other image cues, outline contour is invariant to lighting conditions and variations in object color and texture. Researchers like to utilize extensive prior information in the form of feature such as Sift, Hog. Learning algorithms attempt to get the classified information from these training samples. Traditional methods [1-2] which use a classifier to detect object need hundreds or thousands of samples. In real world application, usually, it is impossible to obtain lots of samples, such as airport security check, and image indexing on the internet. If there is only one sample for the interesting object, in order to detect object in the target images, using contours of the target as the matching template is a better strategy.

There are some works based on shape features which are suited to represent objects in cluttered scene. In [5-6], Ferrari *et.al* use a contour segment network to represent line segments in image retrieving task. They present a collection of scale-invariant local shape features to form chains of contour segments. But the training data need to be aligned impliedly and the detector should get highest response at the annotated area, which may not hold up. In [9], a new class of distinguished regions based on convex local contours in the image is proposed. This work combines cost functions of two tangential edges near a circle to represent object. As one weakness, their work is unlikely to be robust to viewpoint changes. In [10], the authors use cliques of fully-interconnected parts that are a pair wise geometric relationship between them to represent object. However, the learned models are sparse collections of features, so the edge information is not fully exploited in that work. In [14], the authors present a local contour-based feature to detect object. A boosting

procedure is used to select features and estimate features parameters. But these kind of features are harder to reuse within other recognition frameworks. Moreover, their sensitivity to scale limits their applicability. Even these approaches mentioned above have impressive results in cluttered background, but they show lower performances for time critical applications due to high computational complexity.

In many real world applications, there are always not enough samples to train. So method based on shape template matching is useful and meaningful to real applications. In recent years, Liu *et.al.* demonstrate a fast directional chamfer matching(DCM) algorithm in [11], which improves the accuracy of chamfer matching while significantly reducing its computational cost. But at the first phase of DCM, RANSAC algorithm is used to fit line segments, which generates lots of line segments and is sensitive to cluttered background. The number of line segments significantly influences the detection speed. Seeking a method of reducing line segments and applying it to object detection would be a critical task.

To overcome the above problem, we propose an efficient and novel approach called salient line segment representation(SLSR) and apply the shape representation to DCM [11] which is a template matching framework. First, a parabola model is proposed 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[2-7]. Two adjacent corner points whose curvature is lower than a fixed threshold are used to produce a parabola vertex between these two corner points along the corresponding edge. Then we define salient points to describe the shape of object which contain parabola points and corner points. Based on the salient points, DCM method is used to compute the similarity of template and sliding windows.

The remainder of this paper is organized as follows. In Section 2, we first introduce the salient point extraction process, and then show the flow chart of our method. In Section 3, we detail the parabola model. The subsequent two sections present the details of our method. Then in Section 6 a chamfer distance transform matching method is presented. In Section 7, the experimental results are presented to evaluate our method. We conclude this paper in Section 8.

2. Preliminary

The process of our method is illustrated in Figure 1. The input image includes two types, one is shape template image, the other is query image. Firstly, from a(left column), we extract corner points by using CPDA method[2-7]. Then, a parabola model is proposed to represent image edges. This parabola model is constructed by two adjacent corner points whose curvature is lower than a fixed threshold G , where G is a parameter to discriminate straight lines and parabolas (see Figure 3).

In the parabola model, a parabola point between the corresponding adjacent corner points is introduced and two lines between parabola point and two corner points are approximately represented an edge. We regard parabola points and corner points as salient points. Secondly, in right column, query image is operated as input image by 'a' (left column). Then we can obtain the edge map which consists of line segments connected with salient points. Based on the edge map, we compute the chamfer distance image with DCM framework [11]. Thirdly, we operate shape template image with 'a' (left column) and obtain a shape structure. Slide windows move on the distance image and each window matches with the template. Finally, we compute the similarity of the both and find the target hypothesis.

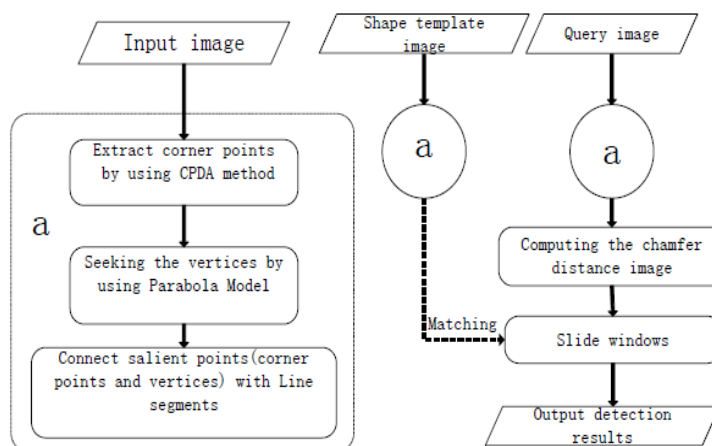


Figure 1. Flow Chart of our Propose.

3. Object Representation

To represent edge simply and efficiently, we propose to use a parabola model to approximate edges whose curvature is lower enough. The aim of parabola model is to approximate low curvature edges which sometime seem like arcs. Previous works[5,6,14] use the dense sampling points on the edge to represent a shape of object. However, we adopt a novel and efficient way to represent shapes.

First, we extract the outlines of the image. Canny operator method is used to compute the edge map and a set of edge fragments is generated by traditional edge-linking approach. The i th edge fragment of this image is defined as $E^i = \{X^i, Y^i\}$, where $X^i = \{x_1^i, x_2^i, \dots, x_N^i\}$, $Y^i = \{y_1^i, y_2^i, \dots, y_N^i\}$, and N is the number of pixels on the i th edge fragment. x_j^i and y_j^i are the coordinates of the j th pixel on the i th edge fragment. Then corner points are detected on each edge fragment, where we use CPDA algorithm to find corner points[2,7], because CPDA algorithm's robust property outperform the curvature scale space corner detector[8]. On each edge fragment, the corner points are $C_i = \{c_1^i, c_2^i, \dots, c_M^i\}$ where M is the number of the corner points on the i th edge fragment and $c_j^i = \{xc_j^i, yc_j^i\}$ are the coordinates of the j th corner point on the i th edge fragment.

To simplify the expression of symbol, we only just consider the corner points of the i th edge fragment. For any two adjacent corner points on the i th edge fragment is written as c^j and c^{j+1} , we define the coordinates of all pixels, between c^j and c^{j+1} , are x^j and y^j , where $x^j = \{x_1^j, x_2^j, \dots, x_N^j\}$, $y^j = \{y_1^j, y_2^j, \dots, y_N^j\}$ and N is the number of pixels on the edge fragment between c^j and c^{j+1} . In fact, $x_1^j = xc_j$, $x_N^j = xc_{j+1}$, $y_N^j = yc_{j+1}$. And one can find that $x^j \subseteq X$ and $y^j \subseteq Y$.

If the curvature of adjacent corner points c^j and c^{j+1} is lower than a threshold G , then c^j and c^{j+1} are written as $c^{k,1}$ and $c^{k,2}$ ($k=1,2,\dots,p_0$), where p_0 is the number of parabola models, and these two corner points are corresponding to a parabola model $p^k = (c^{k,1}, o^k, c^{k,2})$ which means the k th parabola model on the edge fragment (see

Figure 2). In variables $c^{k,1}$ and $c^{k,2}$, the superscripts $k,1$ and $k,2$ mean two terminal points of the k th parabola model. $o^k = (xo^k, yo^k)$ is the parabola point on the corresponding edge fragment between $c^{k,1}$ and $c^{k,2}$. So how to compute the location of the parabola vertex o^k ?

4. Seeking Parabola Vertices

When the mean curvature range from $c^{k,1}$ to $c^{k,2}$ ($k = 1, 2, \dots, p_0$) is lower than the fixed parameter G , the shape between $c^{k,1}$ and $c^{k,2}$ is more like a parabola but not a straight line (see Figure 3). So we let the parabola point o^k is the point whose curvature is the maximum on the edge fragment between $c^{k,1}$ and $c^{k,2}$. Our aim is to localize the $o^k = (xo^k, yo^k)$ on the parabola model. Considering a parabola is given by

$$(x - z_x)^2 + C_0 = 2py, \quad (1)$$

where z_x is a parameter representing a translation value satisfied $z_x \in (xc_j, xc_{j+1})$ on x coordinate and C_0, p are another two parameters, with $x \in (xc_j, xc_{j+1})$.

In this case, we know two terminal points on the parabola $(xc_j, yc_j), (xc_{j+1}, yc_{j+1})$ already. The first and second order derivatives of y are $y' = \frac{1}{p}(x - z_x)$ and $y'' = \frac{1}{p}$, respectively. The curvature at each pixel on the parabola can be calculated as

$$K(x) = \left| \frac{y''}{(1 + (y')^2)^{\frac{3}{2}}} \right| = \frac{\frac{1}{p}}{(1 + (\frac{1}{p}(x - z_x))^2)^{\frac{3}{2}}}.$$

When $x = z_x$, we have $K_{\max}(z_x) = K(z_x) = \frac{1}{p}$. It is easy to find that

$$p = \frac{1}{K(z_x)}. \quad (2)$$

Here, we introduce a variable:

$$K_{mean} = \frac{1}{N} \sum_{z_x=x_1^j}^{x_N^j} K(z_x).$$

Considering the global property of the edge fragment[8], K_{mean} represents its average curvature between corner points x_1^j and x_N^j . Assume each $z_x \in (x_1^j, x_N^j)$ corresponds to a parabola. We use the way formula (3) in the range of $[z_x - \frac{p}{2}, z_x + \frac{p}{2}]$ to express the mean

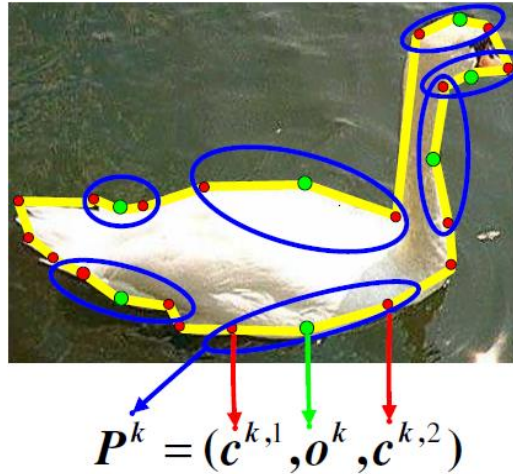


Figure 2. The Parabola Model. The Red Points are the Corner Points. The Green Points are the Parabola Points. And a Parabola Model is Contained by a Blue Ellipse

curvature on each parabola:

$$\bar{K}(z_x) = \frac{1}{p} \int_{z_x - \frac{p}{2}}^{z_x + \frac{p}{2}} \frac{\frac{1}{p}}{\left(1 + \left(\frac{1}{p}(x - z_x)\right)^2\right)^{\frac{3}{2}}} dx. \quad (3)$$

Let $t = (x - z_x) / p$, then

$$\bar{K}(z_x) = \frac{1}{p} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{(1+t)^{\frac{3}{2}}} dt = \frac{\sin(\arctan \frac{1}{p}(x - z_x)) \Big|_{z_x - \frac{p}{2}}^{z_x + \frac{p}{2}}}{p} = \frac{T}{p}. \quad (4)$$

Where T is a constant value.

So the coordinate of o^k is computed by

$$xo^k = \arg \min_{z,x} \left| \bar{K}(z_x) - K_{mean} \right| = \arg \min_{z,x} \left| \frac{T}{p} - K_{mean} \right| \quad (5)$$

$$\text{s.t. } \bar{K}(z_x) \geq G \text{ and } \left| \bar{K}(z_x) - K_{mean} \right| < \delta,$$

Where δ is the neighborhood parameter of K_{mean} . We just consider $z_x \in (x_1^j, x_N^j)$ which satisfy $\bar{K}(z_x) \geq G$ and $\left| \bar{K}(z_x) - K_{mean} \right| < \delta$. $\bar{K}(z_x) \geq G$ means, at each z_x , the curvature is large enough to ensure the point (z_x, z_y) is not on a straight line. If $\bar{K}(z_x) < G$, the curvature of a parabola is too small to view as a parabola (see Figure 3). And $\left| \bar{K}(z_x) - K_{mean} \right| < \delta$ means the curvature of the estimated parabola point o^k can reflect and approximate enough the global curvature. When xo^k is obtained, we can find

that yo^k from the corresponding z_x , where yo^k is in $[yc_j, yc_{j+1}]$. Formula (5) combines the global curvature of all parabolas with mean curvature of each parabola, where we can choose parabola points between two corner points.

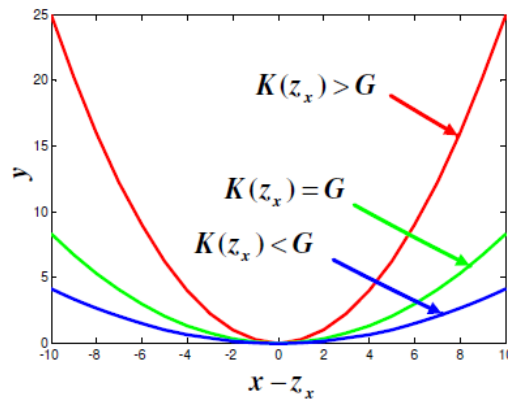


Figure 3. Parabolas with Different Curvatures.



Figure 4. Salient Points. The Green Points are the Parabola Points. The Red Points are the Corner Points.

5. Salient Point

Our parabola model is evolution of the curvature scale space corner detector[8], but very different. Our parabola model is used to represent edge fragments between corner points, but the corner detector is used to search corner points on edge fragments[8]. This parabola model is some kind of interpolation between corner points to fit edges.

After we find out all parabola models on object template, there are two kinds of points on edge fragments(see Figure 4): corner points and parabola points. Pairs of two adjacent corner points without parabola between them approximate straight lines of edges. Triples of a parabola point and two corner points approximate edge curves. Here, we call these points as salient points. In previous works[12-16], there have also been the definition of salient points. In these two works, salient points are more like interest points in salient regions. In this paper, we define the salient points are the collection of corner and parabola points which can be used to approximate and fit edges. Different from previous works in [12-16] our salient points are all on edge fragments to represent the contours of object(See Figure 5). In Figure 6, we present an example of linear representation, where five classes of ETHZ

are respectively modeled by line segments with SLSR approach. As shown, we can see it has give a good approximation of the object shape.

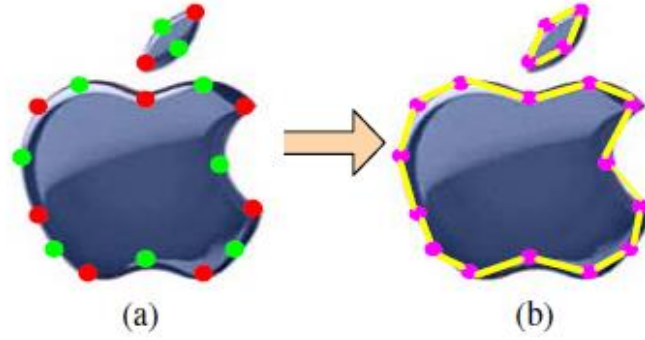


Figure 5. (a)The Red Points are the Parabola Points. The Green Points are the Corner Points. (b)The Pink Points are Salient Points.

6. Distance Measurement

Chamfer distance[11] is a popular and robust distance measurement method to align two maps. Let $W = \{w_i\}$ and $Z = \{z_j\}$ be the sets of edge map points. The distance between these two sets is

$$dist(W, Z) = \frac{1}{n} \sum_{w_i \in W} \min_{z_j} |w_i - z_j|, \quad (6)$$

where n is the number of elements of W . For each pixel w_i , we define a distance transform image

$$D_Z(W_i) = \min_{z_j \in Z} |w_i - z_j|, \quad (7)$$

$D_Z(W_i)$ specifies the distance from pixel w_i to the nearest edge pixel in Z . Chamfer matching based on Chamfer distance can tolerate small rotations, misalignments, occlusions and deformations. Formula (6) just contains distance measurement of pixel location.

In fact, each point on edge fragment has a direction term $\psi(Z)$ which represents the orientation. So we use DCM[11] method to combine location distance measurement and orientation distance measurement. Then, DCM score is written as

$$D_Z(w_i, \psi(w_i)) = \min_{z_j \in Z} |w_i - z_j| + \lambda |\psi(w_i) - \psi(z_j)|, \quad (8)$$

where λ is a weighting parameter between location measurement and orientation measurement. $\psi(\cdot)$ is quantized into Q (default $Q = 60$) levels from $[0, \pi)$. $|\psi(w_i) - \psi(z_j)|$ means the distance between the two quantization levels of $\psi(w_i)$ and $\psi(z_j)$. Let $\Psi = \{\psi_1, \psi_2, \dots, \psi_Q\}$ be the Q quantized orientation levels. So formula(10) can be rewritten as:

$$D_Z(w_i, \psi(w_i)) = \min_{\psi_i \in \Psi} (|w_i - z_j|_{\psi_i} + \lambda |\psi(w_i) - \psi_i|). \quad (9)$$

where $\left|w_i - z_j\right|_{\psi_k}$ means we just consider edge points whose orientation are ψ_k . In formula (7), DCM score is computed on the pixel view. In formula (8), DCM score is computed on the orientation view. In fact, these two formulas are equivalent. That's because a edge map that just contains edge pixels is corresponding to Q maps according to the orientation.

Next we adopt the distance transform integral[11] to represent integral distance measurement. Integral images are used for fast calculation of region and linear sums[11] as an intermediate of image representations. Let $P = \{p_1, p_2, \dots, p_s\}$ be the set of pair of adjacent salient points, where S is the number of pairs and $p_i = (s_i, e_i)$ is a pair of adjacent salient points on a same edge whose start point and end point are s_i and e_i . Edge points on P is defined as $EP_W = \{ep_{[s_i, e_i]}\}_{i=1,2,\dots,S}$. The distance transform integral is

$$ID(w, \psi_k) = \sum_{w_i \in ep_{[w_0, w]}} D_Z(w_i, \psi_i), \quad (10)$$

where the w_0 is the intersection of an image boundary with the line passing though w and having the direction ψ .

So the directional chamfer matching score with distance transform integral form is written as:

$$dist(W, Z) = \frac{1}{n} \sum_{ep_{[s_k, e_k]} \in EP_W} [ID(e_i, \psi_{ep_{[s_i, e_i]}}) - ID(s_i, \psi_{ep_{[s_i, e_i]}})]. \quad (11)$$

We use the formula(11) to match template and windows which are from target image by sliding window. In off-line processing, the locations of salient points of template image are obtained which are saved in the set of pairs of salient points P . For each sliding window is the same size of template, w in formula(11) is the pixel set whose elements are from P . $dist(W, Z)$ is the similarity of between template image and each sliding window. If $dist(W, Z) < \tau$, the corresponding window contains the object of template image where τ is a threshold parameter. In [11], Liu et.al. propose to using trivial line segments in each sliding window to compute $dist(W, Z)$. Comparing their work, our method is much more time-saving and memory-saving. We just use parabola model and corner points to obtain salient line segments and compute similarity $dist(W, Z)$ just on the corresponding locations of template.

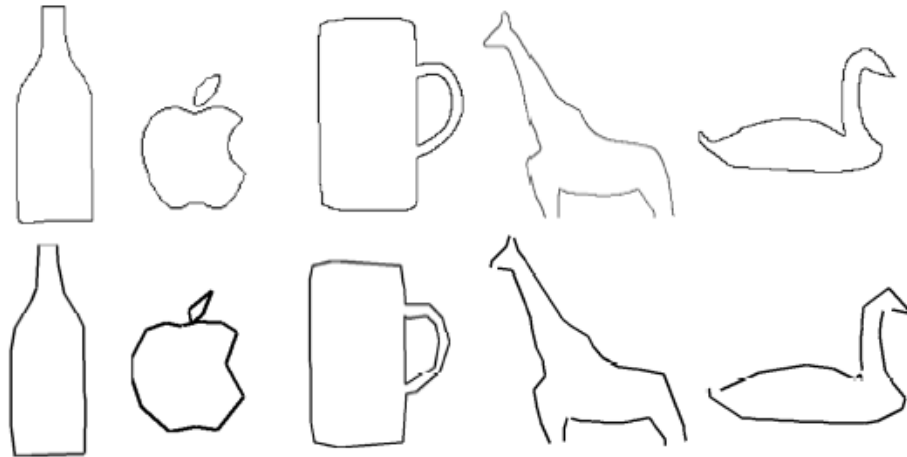


Figure 6. Top Row: Edge Image of ETHZ Model. Bottom Row: Linear Representation of the Edge Image with SLSR

7. Experiments

In this section, we conduct on two challenging image sets with our SLSR approach. One is ETHZ shape class dataset and the other is INRIA horses dataset. All experiments were conducted on a 2.2 GHz PC with 2 GB of RAM. We first briefly discussed for DCM[11]'s work step. a) RANSAC method is used to fit the small line segment, which can represent template image and the query image. b) Line segment in the query image is Quantified n (default 60) directions, and the chamfer distance image is generated in each direction. c) Indexed by the number of line segments of the target template, detection algorithm seek the object in the distance image of each direction by means of sliding window .

And then to our method, in the first phase we use SLSR to depict the query image instead of RANSAC, due to RANSAC method is slow and sensitive to cluttered background. The other steps are identical with the above. Note that, in the first experiments we emphasize the speed and the improved accuracy of our method comparing to DCM. In the second dataset, we emphasize the conduction on real image too.

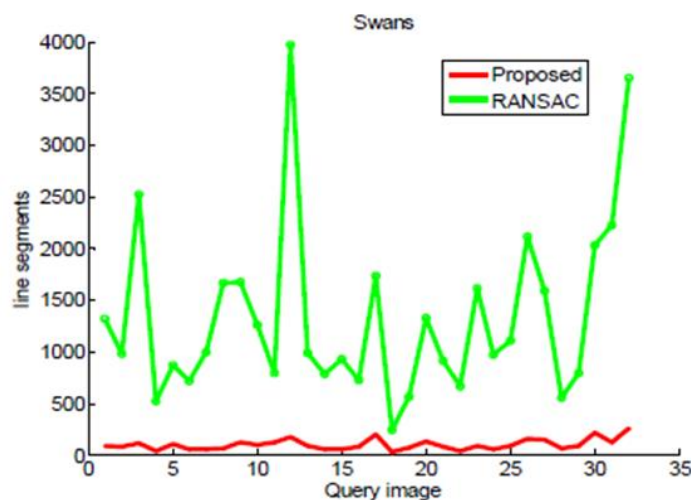


Figure 7. The Number of Line Segment Representing the Detected Image in ETHZ (swans)

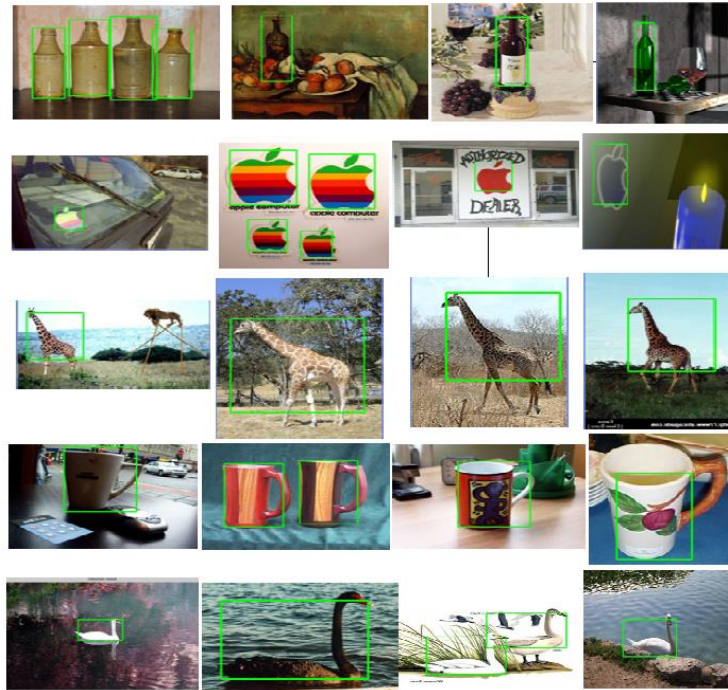


Figure 8. Salient Point Representation Image Detection

7.1. On Edge Map

Experiments are conducted with ETHZ shape class dataset, which includes 5 classes (bottles, applelogos, mugs, giraffes, swans) and contains 255 images all together. In each image there is one or more objects, and they have large variations in appearance, scale and viewpoint.

By using SLSR method, we get the line segment representation of the query edge image. Then we align the template with the object using the direction chamfer matching method in the detection stage. As shown from Figure 8, we can see the detection results with SLSR. Due to a good approximation of the object shape, and the number of line segments of largely reduced(see Figure 7), detection accuracy and speed has been improved. We present results that are very competitive with DCM' RANSAC scheme and show robustness to clutter, and inner-class variations.

There are diagrams of false positive per image vs. detection rate, We compared our approach with oriented chamfer matching[15] and recent studies by Ferrari *et.al* [4,5] and by Liu *et.al*. [11] From DCM, we replace RANSAC method with SLSR, which could reduce the number of lines largely and also remove the disturbance of the trivial line segments. Comparison of 5 classes, the four classes' performance including swans, giraffes, bottles and applelogos, each of the detection positive rate is better than DCM, while mugs is comparable to DCM in Figure 9.

RANSAC method which fits numerous small line segments in DCM[11] is used to represent image contours, then they generate chamfer distance map by processing these line segments in multiple directions. So the idea we reduce the number of representation line segments of the image arises spontaneously. Extract edge contour curves, and seek salient points on the contour curves. As shown in Figure 7, the number of line segments which are connected between salient points, is greatly reducing. The number of line segments can be obtained by using RANSAC averages about 1365 in each image, and in average is $1365/60 = 22$ in each direction, i.e., when generate the distance transform image, operating 22 times in each direction. But SLSR in this paper, the number is 125, and the average number of line segments is $125/60 = 2$. SLSR is the average 2 at each direction, so that is $(22 - 2) * 60 = 1200$. less than 20 line segments each cycle in average is calculated, each image is

calculated less than 1200 line segments in average. when come to scale(default 15) in [11], $1200 * 15 = 18000$. That is to say, SLSR can save 18000 cycle each image in average. Meanwhile, the number of segments of representation image greatly reduced, the division of the direction can be appropriately reduced relative range from 60 to 40, can reduce the number of detection cycles. Because the large number of trivial line segments generated by RANSAC, it leads to a long detection time. But the detection speed by using SLSR is 16.06% higher per image in average than RANSAC under the similar or better precision(see Table 1).

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)
DCM	5.6095	14.4040	8.0431	6.9475	6.8595	8.3727
SLSR	5.2635	10.4009	7.6220	6.8897	4.9623	7.0277

RANSAC, which is the first step of DCM, would conduct on the data set of edge image, rather than on real images sets, because it could be an almost impossible task. So it need an edge extraction approach in the pretreatment process such as the best Berkeley's algorithm to generate edge image. As well as we known, this algorithm is quite time-consuming, and to prepare extracted Edge Image is a very tedious work. SLSR can save this processing step, and directly extract line segments in the nature image.

We can see the detection rate diagram(Figure 9), SLSR method is superior to DCM method. It is because the line segment with SLSR is longer than with DCM in average, the image generated from it is clear, less noise. Distance image is sensitive to noise, so splitting it in n(default 60) directions can well remove the impact of a complex environment. Extracting salient point greatly reduces the number of line segments, Which can make the image simple, and object detection can be more stable.

7.2. On Real Images

The second experiment we conduct on the challenging INRIA horses 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. Since Liu et.al. do not provide the information on this dataset in[11], we will make a comparison with the state of art methods, M2HT+IKSVM and Kas in [13]. Note that this experiment is done without edge images, but directly on the real images to seek salient points. Horse's action changes is more than the ETHZ database class. RASANC fitting algorithm on the real image segment is an impossible task, because it is more sensitive to the surrounding noise. The salient point extracting method is more stable in the whole processing phase ether on real image or on edge map, which is to avoid looking around for the fitting point. As shown in Figure 10, there are the detection results and in Figure 11. we can see our approach is comparable with M2HT+IKSVM and significantly better than KAS[13], even we need not to have a training process.

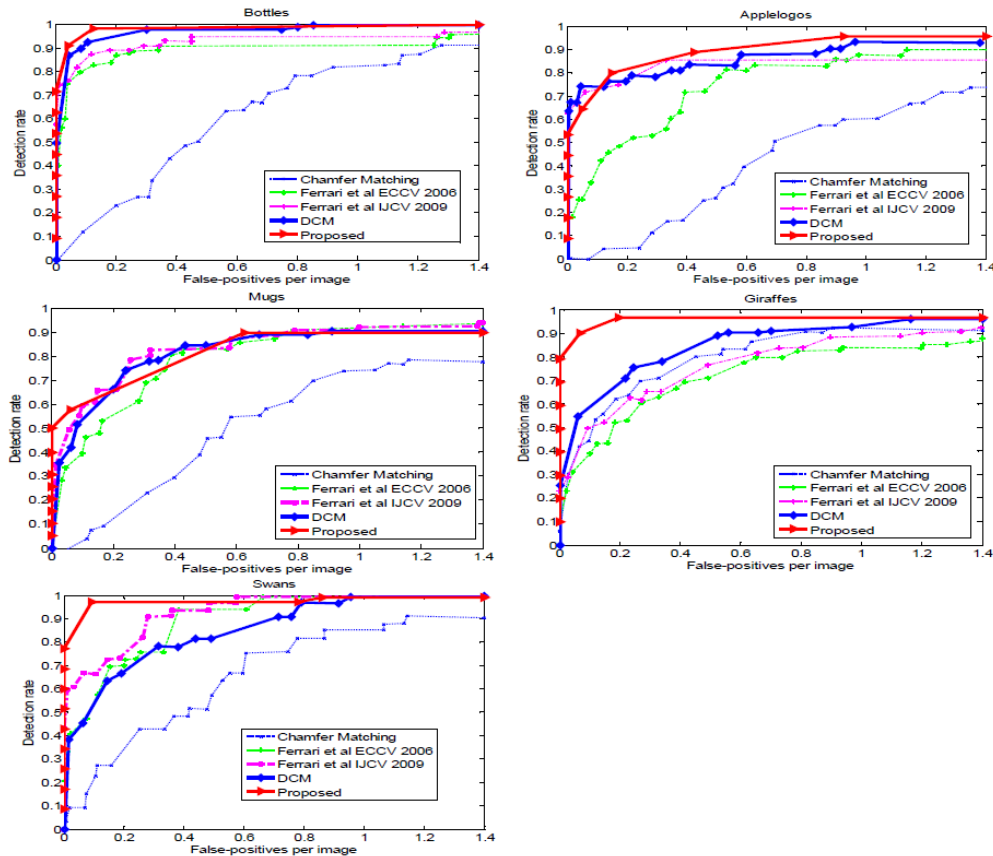


Figure 9. ROC Curve and Several Localization Results on the ETHZ Shape Dataset

8. Conclusion

We propose an efficient and novel approach for representing image contours with salient points. We present a parabola model for seeking the salient points while significantly reducing the number of salient line segments. According to the methods, the detection accuracy and speed have been improved largely. As shown all the above, we can see the pros and cons of detection affect on the contour extraction in proportion. Finding appropriate salient points is a very significant task, this paper aims to use the least number of line segment representation nearest edge image, and to improve speed in the case of maintaining detection accuracy. Finally, in this paper it would be desirable to detect the partial occlusion target fast in the future.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (61375038).

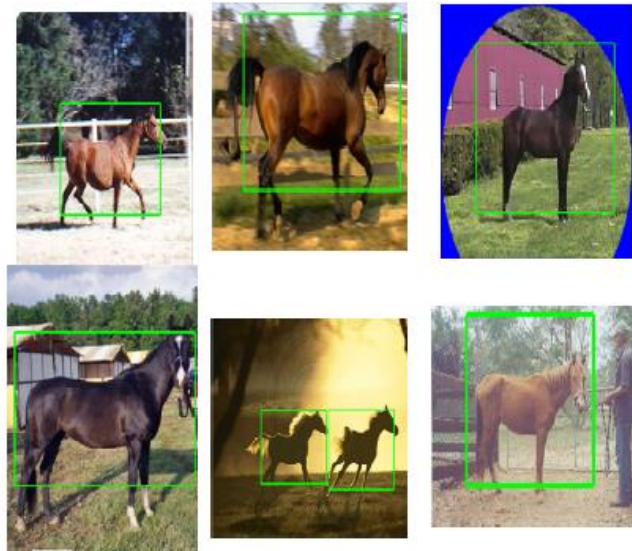


Figure 10. Salient Point Representation Horse

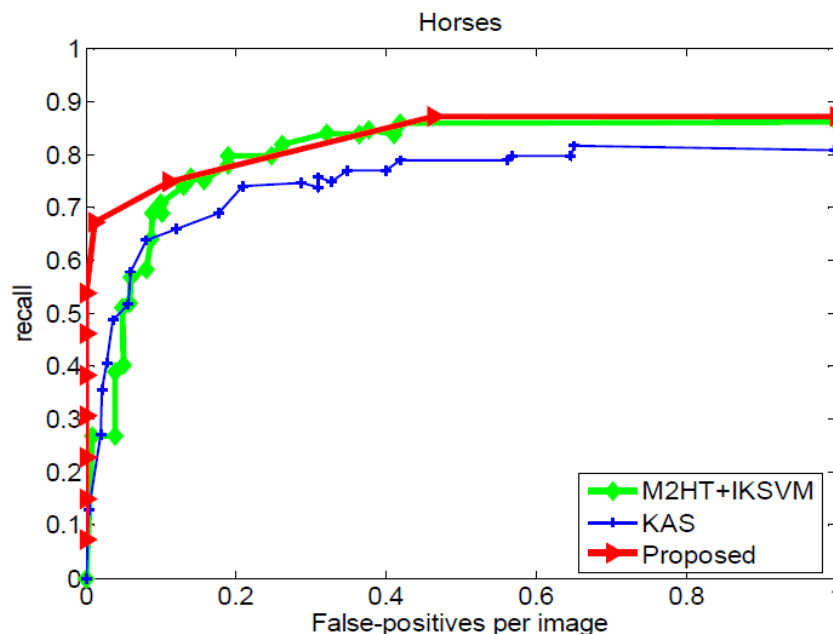


Figure 11. ROC Curve and Several Localization Results on the INRIA Horses Dataset

References

- [1] S. An, P. Peursum, W. Liu and S. Venkatesh, "Efficient Algorithms for Subwindow Search in Object Detection and Localization", IEEE Conf. Computer Vision and Pattern Recognition, Miami, FL, USA, June 20-25 (2009).
- [2] M. Awrangjeb, G. Lu, C. S. Fraser and M. Ravanbakhsh, "A Fast Corner Detector Based on the Chord-to-Point Distance Accumulation Technique", IEEE Conf. Digital Image Computing: Techniques and Applications, Melbourne, VIC, Australia, December 1-3 (2009).
- [3] G. Csurka, C. R. Dance, L. Fan, J. Willamowski and C. Bray, "Visual categorization with bags of keypoints:", Proc. European Conf. Computer Vision, Prague, May 11-14 (2004), pp.1-22.
- [4] V. Ferrari, L. Fevrier, F. Jurie and C. Schmid, "Groups of Adjacent Contour Segments for Object Detection", PAMI, vol. 30, no. 1, (2008), pp. 36-51.
- [5] V. Ferrari, F. Jurie and C. Schmid, "From images to shape models for object detection", Intl Journal of Comp. Vis, (2009).

- [6] V. Ferrari, T. Tuytelaars and L. Van Gool, "Object Detection with Contour Segment Networks", ECCV Graz, Austria, May 7-13 (2006).
- [7] J. H. Han and T. Poston, "Chord-to-point distance accumulation and planar curvature: a new approach to discrete curvature", Pattern Recognition Letters, vol. 22, (2001), pp. 1133-1144.
- [8] 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, vol. 2, Aug. 23-26 (2004), pp. 791-794.
- [9] F. Jurie and C. Schmid, "Scale-invariant Shape Features for Recognition of Object Categories", CVPR, Washington, DC, USA, June 27-July 2 (2004).
- [10] 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).
- [11] M.-Yu 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).
- [12] E. Loupas, 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, Sept 10-13 (2000).
- [13] S. Maji and J. Malik, "Object Detection using a Max-Margin Hough Transform", CVPR, Miami, Florida, USA, June 20-25 (2009), pp. 1038 - 1045 .
- [14] J. Shotton, A. Blake, and R. Cipolla, "Contour-Based Learning for Object Detection", ICCV, Beijing China, Oct. 17-21 (2005).
- [15] J. Shotton, A. Blake and R. Cipolla, "Multi-scale categorical object recognition using contour fragments", IEEE Trans. on Pattern Anal. and Mac. Intel., vol. 30, no. 7, (2008), pp. 1270-1281.
- [16] Q. Tian, N. Sebe and M. S. Lew, "Image Retrieval Using Wavelet-based Points", Journal of Electronic Imaging, vol.10, no. 4, (2001), pp.835-849.

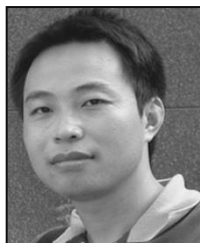
Authors



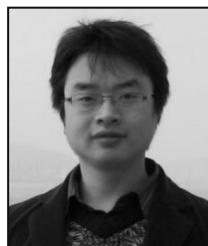
Yumin Dou received his master's degree in Computer Science from Chengdu University of Technology in 2010. He is currently a Ph.D. candidate at the School of Computer Science and Engineering of University of Electronic Science and Technology University of China. His research interests include pattern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



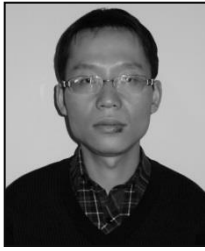
Pei Xu received his BS degree in computer science and technology from SiChuan University of Science and Engineering, ZiGong, China, in 2008 and his MS degree in condensed matter physics from University of Electronic Science and Technology of China, Chengdu, China, in 2011. He is currently a PhD student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



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.



Renjie Huang received his master's degree in Computer Science from Southwest University in China in 2006. He is currently a Ph.D. candidate at the School of Computer Engineering and Technology of University of Electronic Science and Technology of China. He is also an instructor at the School of Computer and Information Science of Southwest University in China. His research interests include machine learning and computer vision.



Tao Li received his M.E. degree from Central South University, Changsha, China, in 2006. He is now a Ph.D. student at the School of Computer Engineering and Technology of University of Electronic Science and Technology, Chengdu, China. His current research interests are machine vision, visual surveillance and object detection.

