

Shape Representation Robust to the Sketching Order Using Distance Map and Direction Histogram

Kiwon Yun, Junyeong Yang, and Hyeran Byun

Dept. of Computer Science, Yonsei University, Seoul, Korea, 120-749
{wallace,gundid,hrbyun}@cs.yonsei.ac.kr

Abstract. We propose a powerful shape representation to recognize sketches drawn on a pen-based input device. The proposed method is robust to the sketching order by using the combination of distance map and direction histogram. A distance map created after normalizing a free-hand sketch represents a spatial feature of shape regardless of the writing order. Moreover, a distance map which acts a spatial feature is more robust to shape variation than chamfer distance. Direction histogram is also able to extract a directional feature unrelated to the drawing order by using the alignment of the spatial location between two neighboring points of the stroke. The combination of these two features represents rich information to recognize an input sketch. The experiment result demonstrates the superiority of the proposed method more than previous works. It shows 96% recognition performance for the experimental database, which consists of 28 freehand sketches and 10 on-line handwritten digits.

Keywords: Shape representation, Sketch recognition, Sketching order, On-line handwriting recognition.

1 Introduction

Recently, as pen-based input devices became popular in the education area by the advance of tablet PC, it is increasing the needs to recognize not only handwritten characters, but also sketches with various shapes. The spatial information is used as an important factor in recognizing handwritten sketches because it illustrates the shape of the input object. However, contrary to printed characters, it has a large shape variation in terms of each individual's in handwritten characters [1]. Therefore, the variation of input must be strongly considered when describing the features of a sketch in the handwriting recognition; especially for writing independent sketch. Fortunately, the additional information of the writing, such as the number of strokes or the end points for each stroke, can be used to recognize on-line handwriting. In particular, the directional information of the writing for each stroke is another significant factor, in addition to the spatial information of writing. However, when the system tries to represent the directional feature of a sketch, it must consider that users can draw each sketch regardless of some

decisive factors for directional feature, such as the order of input or the number of strokes.

A lot of study groups have over the past decades researched how to illustrate the feature information of freehand writing [2]. Also, their research has a widespread application in various environments [3], [4], [5]. Chan and Yeung [5] proposed a feature representation method to recognize handwriting by using fundamental properties of symbol consisting of a primitive set. This method has a weakness, which makes it difficult to recognize complex symbols or general sketches. Matsakis [6] represented shapes consisting of multi-strokes using aligned points that are invariant to the order of writing. However, this method does not consider the different number of strokes that can be drawn by different writers. It is also impossible to apply this method to symbols having loop like 0, 8 and ∞ . Belongie [7] proposed shape context method using distance and direction between sampling points as feature information. It is less accurate in recognizing shapes that have large variations, but it can represent sketches as rich features simply by their distance and direction. Laviola [8] proposed a symbol representation method combined with fourteen different features such as the number of strokes, aspect ratio, intersection points, angle histogram and so on. It demonstrates that a combination feature consisting of several different kinds of slight features has enough accuracy to recognize a symbol.

In this paper, we propose a shape representation method with a high discriminating power using a combination of spatial and directional feature to recognize input sketch on Tablet PC. First of all, we use distance map, given the distance to the nearest edge at every pixel in the shape to represent spatial information. By the way, chamfer matching using distance map cannot handle large shape variation like handwriting because it just uses the edge of strokes as a measure [9]. In order to solve this problem, we use a distance map as a robust concerning large variations of on-line handwriting. This is due to the fact that it has more detail and rich spatial information of strokes. Moreover, we add the direction of strokes into feature data of input to increase recognition accuracy. In this work, we propose the representation method of directional features to be invariant to the order and the number of strokes using direction histogram.

After this section, we show how the proposed shape representation method can be robust in regards to writing order and variation by using distance map and direction histogram in section 2. In section 3 we show how to make features using the proposed method and to recognize the new input using them. Section 4 shows the results of our recognition experiments on a database including handwritten digits and freehand sketches. Finally, we present conclusions in section 5.

2 Shape Representation

In this chapter, we propose a shape representation method to be robust in its sketching order. The proposed method consists of two steps. One is to normalize freehand sketch and another is to extract features from normalized sketch data. Fig. 1 shows the conceptual diagram of the proposed method. For the the normalization step, control points are extracted using curvature properties of

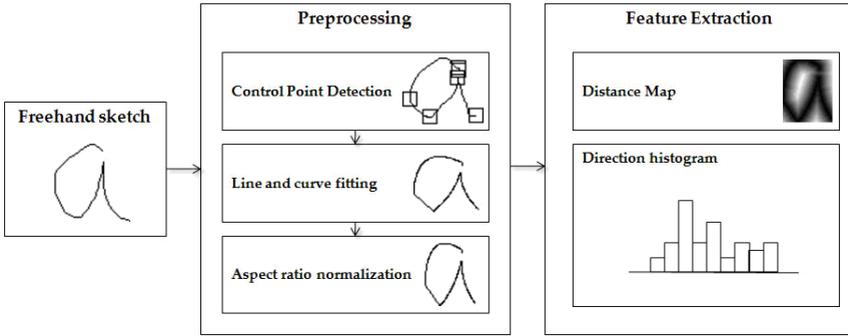


Fig. 1. System overview. After a couple of preprocessing from freehand sketch, the feature of sketch can be represented with distance map and direction histogram.

input sketch. Then, noises are removed by line and curve fitting using detected control points. Finally, the result of fitting process is interpolated into regular size, $m \times n$, to normalize aspect ratio. The second step is to extract spatial and directional features from normalized sketch data. In order to represent spatial feature, we use distance map computed by distance transform. This method helps the system reduce errors which can occur within the same category because features made solely by edges can be sparse. Next, in order to represent directional feature, we have to consider a variation of strokes drawn by different people for the same shape. This is because the order of drawing for a shape can not be the same for different writers and its effects on directional features for the shape. In order to solve this problem, we proposed directional feature which is invariant toward the order of drawing by the process of alignment between neighboring sampling points.

2.1 Sketch Normalization

A stroke is defined as a set of point, $\mathbf{x} = (x, y)^T$, on 2D plane. The set consists of p points between pen-down point and pen-up point. A stroke, \mathbf{s}_i , is defined as $\mathbf{s}_i = \{\mathbf{x}_1, \dots, \mathbf{x}_p\}$. A sketch or online handwritten symbol is consisted of a set $\{\mathbf{s}_1, \dots, \mathbf{s}_s\}$ having above one stroke. We define the set of the control point, \mathbf{P} , as the part of input points in which dramatically change the direction, such as the corner, a high curvature, and junction. By the definition, \mathbf{P} satisfies $\mathbf{p}_i \in \{\mathbf{s}_1, \dots, \mathbf{s}_s\}, \forall \mathbf{p}_i \in \mathbf{P}$. About a point \mathbf{x}_i , we define the nearest control point on a spatial domain as $\hat{\mathbf{p}}$ and define two vectors \mathbf{d}_{xp} and \mathbf{d}_{xx} where $\mathbf{d}_{xp} = \mathbf{x}_i - \hat{\mathbf{p}}$ and $\mathbf{d}_{xx} = \mathbf{x}_i - \mathbf{x}_{i+1}$. Now, the dot product, $\mathbf{D}_p(\mathbf{x}_i)$, of two vectors is expressed as $\mathbf{D}_p(\mathbf{x}_i) = (\mathbf{d}_{xp}^T \mathbf{d}_{xx}) / (\|\mathbf{d}_{xp}\| \|\mathbf{d}_{xx}\|)$. If $\mathbf{D}_p(\mathbf{x}_i)$ is more than the threshold, it means that the curvature of \mathbf{x}_i is small, i.e., the point is not a control point. We apply this rule to all input points and the point that satisfies the rule is added at a set of control points. This algorithm is summarized in Table 1. However, inadequate points remain around the corner of the stroke. We simply eliminate

Table 1. The proposed feature point detection algorithm based on the dot product

<p>A stroke $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_k]$, $\mathbf{x}_i \in \mathfrak{R}^2$ is given. Define \mathbf{P} as the set of control points and ϕ as the threshold. Initially, \mathbf{P} is the empty set.</p> <p>Initialization: $\mathbf{P} \leftarrow \mathbf{P} \cup \mathbf{x}_0$</p> <p>Iteration: For $i = 1$ to k Compute the nearest control point of \mathbf{x}_i. Compute the dot product $\mathbf{D}_p(\mathbf{x}_i)$. If $\mathbf{D}_p(\mathbf{x}_i) < \phi$ $\mathbf{P} \leftarrow \mathbf{P} \cup \mathbf{x}_i$</p>
--

control points which are close together because they have a tendency of being together.

When we extracted control points from an stroke by using the proposed feature point detection algorithm, we classify each control point pair as a line or a curve. To this work, we used $\mathbf{L}_2/\mathbf{L}_1$ proposed in [10] where $\mathbf{L}_1(i, j) = \|\mathbf{x}_i - \mathbf{x}_j\|$ and $\mathbf{L}_2(i, j) = \sum_{l=i}^j \|\mathbf{x}_l - \mathbf{x}_{l+1}\|$, given the control point pair consisting of two neighbors, \mathbf{P}_i and \mathbf{P}_j . By \mathbf{L}_1 and \mathbf{L}_2 , each control point pair is classified as a curve, if $\mathbf{L}_2/\mathbf{L}_1$ is above the threshold, ϕ_c , otherwise it is classified as a line. The control point pair, which is classified as a line is beautified from the line, which pass through the start and end point of the pair. In the case of a curve, we expressed the curve as a quadratic Bezier curve by using the proposed algorithm in [10].

Lastly, the sketch which is fitted by line and curve is interpolated into regular size to reduce the effect of different aspect ratio. A sketch is interpolated with the scale ratio which adjusts the width and height of sketch into m and n . In this paper, we define the interpolated sketch as $\mathbf{I}(x, y) \in \mathfrak{R}^{m \times n}$.

2.2 Spatial Information Using Distance Map

We propose a spatial feature representation method using a distance map to be robust to shape variation as well as sketching order in handwritten sketches. The proposed spatial feature is trivially robust to sketching order because a distance map represents a captured image of sketch. In chamfer matching, first proposed by Barrow [11], distance transform creates a distance map, which represents how far each pixel is from the edges. Lower value is assigned to pixels near edges and higher value is assigned to pixels far from edges. By chamfer matching, the edge of new input is superimposed on the distance maps of each template. Then, the input is recognized as the category of the template with the lowest sum of the edge hits. Here, the sum of the edge hits is called chamfer distance. However, original chamfer matching using chamfer distance lacks robustness concerning the environment with a large shape variation like on-line handwriting [9]. On the

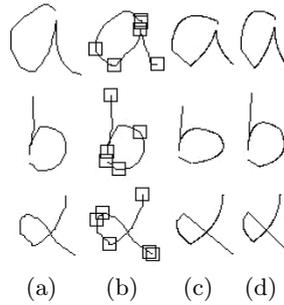


Fig. 2. Examples of the normalization result. (a) Freehand sketch. (b) Control point detection. (c) Line and curve fitting. (d) Aspect ratio normalization.

other hand, the case of using a distance map for spatial information of sketch gives us a richer spatial information because each pixel on distance map has more detail for each spatial feature. Given normalized sketch $\mathbf{I}(x, y)$ in regular size, $m \times n$, a distance map of the sketch is defined as $\mathbf{C}(x, y) \in \mathbb{R}^{m \times n}$ by the proposed algorithm in [12]. Fig. 3 shows examples of a distance map.



Fig. 3. Examples of a distance map. Input sketches are on top, and distance maps created by input sketches are on bottom. Black area is more closed to the edges, and white area is farther from edges in distance maps.

As we proposed, feature information using distance map can solve the another problem of original chamfer matching which is that chamfer distance considers edge pixels just as meaningful information. In this case, this problem is shown in Fig. 4. In contrast, our proposed approach not only solves this problem, but it also creates a robust feature toward shape variation. This is because all pixels on distance map can be meaningful to represent the shape.

2.3 Direction Histogram

We propose invariant directional feature using a histogram toward sketching order or the number of strokes. The directional information of input strokes is one of the significant features, in addition to the spatial information of writing. Thus, this factor is often used to recognize shape in several research area [9], [5], [8] [6]. When feature of shape is represented using directional information, one significant factor has to be considered. The same shape can be drawn in a different order by different persons. As a result, feature can describe the same sketch in a different way. For example, if two people try to draw input '1' on screen, one person can draw it from top to bottom, while another person can

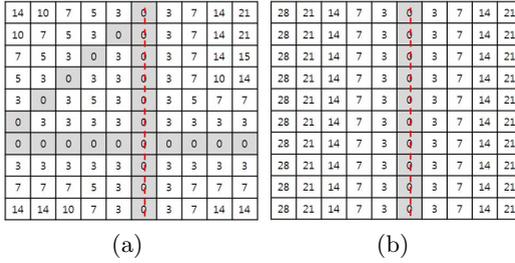


Fig. 4. The problem of original chamfer matching. The number of each pixel is distance to the nearest edge and the dotted line is the edge of new input. (a) is the comparison between a template for digit 4 and a new input for digit 1, and (b) is the comparison between a template for digit 1 and a new input for digit 1. The chamfer distance is 0 in both cases, so this input can be recognized as both 4 and 1.

draw it from bottom to top for the same shape. In order to solve this problem, we propose a direction invariant method by using the process of alignment between neighbors of sampling points

Given normalized sketch $\mathbf{I}(x, y)$, the set of re-sampling points is define as $\{\mathbf{X}_i \dots \mathbf{X}_k\}$, where $\mathbf{X}_i = (x_i, y_i)^T$ and the number of points in the set of normalized strokes is defined as k . Then direction of input points is given as :

$$\mathbf{D}(i) = \tan^{-1}\left(\frac{y_{j+1} - y_j}{x_{i+1} - x_i}\right) \tag{1}$$

In order to make direction invariant feature, we modified the proposed algorithm in [6]. Two neighbor points are sequentially selected from first two points in the set of re-sampling points. A primitive stroke is made by these two selected points. Then, each primitive stroke can be one of three types : horizontal, vertical, diagonal. Given two points, \mathbf{X}_i and \mathbf{X}_j , a primitive stroke is horizontal if $x_i = x_j$, and vertical if $y_i = y_i$, and diagonal if $x_i \neq x_j$ and $y_i \neq y_j$. A primitive stroke is aligned by the regular order. In other words, if $x_i < x_j$ in horizontal strokes, or if $y_i < y_j$ in vertical and diagonal strokes, then let them leave. In other cases, two points of the primitive stroke swap their positions. As a result, the direction of a primitive stroke made by two points is always same. By this directional alignment, $\mathbf{D}(i)$ is $0 \leq \mathbf{D} \leq \pi$ as Eq. 1, while if the directional alignment process is not performed, $\mathbf{D}(i)$ is $-\pi \leq \mathbf{D} \leq \pi$.

By the way, if these direction are used as feature without any process, then feature information can be sparse. For this reason, direction angle quantize as n level to create direction histogram. In conclusion, this direction histogram represents invariant directional information toward the order of strokes and the number of strokes.

3 Sketch Recognition

In this paper, the spatial feature using distance map and directional feature using direction histogram are used to represent the features of shape. In order to

use these two pieces of information as a feature vector, we first normalize them within an uniform range. First of all, the spatial feature is normalized in the following way. We first compute the maximum value of $\mathbf{C}(x, y)$, \mathbf{C}_{max} , for each distance map. Then, all the value of $\mathbf{C}(x, y)$ divided by \mathbf{C}_{max} are normalized within $0 \sim 1$ value. On the other hand, the directional feature using direction histogram is also normalized within $0 \sim 1$ divided by the number of points in an entire stroke. Finally, these two feature vector can be used as the feature vector to represent sketch because they are normalized within an uniform range. In this paper, we define the size of distance map, m and n , as $m=50$ and $n=40$. These values are from the result of experiments to decide which value is the best way to illustrate a shape. Also, the quantization level for a direction histogram is defined as the most efficient value, 10, by the experiment. Eventually, the final feature vector of shape is combined by the spatial vector and the directional vector. The spatial vector is combined by each row of $\mathbf{C}(x, y)$ in distance map and the directional vector is quantized as 10 direction. In other words, the final feature vector of shape is defined as $\mathbf{F} = (\mathbf{C}^T \mathbf{D}^T)^T \in \mathfrak{R}^{(m \times n + 10) \times 1}$. After all, templates can be made using this feature vector for each sketch. At this time, the feature vector is computed by the mean of total training data. In other words, given m training data and feature vector \mathbf{F}_i for a sketch, the feature vector of a template, \mathbf{T} , is defined as $\mathbf{T} = \sum_{i=1}^m \mathbf{F}_i / m$. In conclusion, the new input sketch is recognized as the category of the template with the lowest Euclidian distance value by comparisons between the feature vector of input and the feature vector of templates.

4 Experiment

For an experiment, we collected a sketch database while writers draw given data-set with random sketching order. First, we gave 28 predefined shapes to 12 person and let them draw the shapes by using the pen of Tablet PC. Each person repeatedly drew each shape 3 times. Therefore, we collected a total of 1008 ($28 \times 12 \times 3$) shapes. We utilized 336 sketches among the 1008 sketches for a training database, while others were utilized for an experimental database. Second, for an experiment to recognize the online handwritten digits, we asked 40 persons draw $0 \sim 9$ digits by using the pen of Tablet PC and they were also drawn 3 times for each digit. Therefore, we collected a total of 1200 ($40 \times 10 \times 3$) shapes for digits. We utilized 900 sketches among the 1200 sketches for a training database, while others were utilized for an experimental database. Fig. 5 shows examples of two kinds of experimental databases.

We built templates according to each subject by using two training databases and evaluated the accuracy of sketch recognition about two experimental databases. The first experiment compares the accuracy between the proposed method using distance map and original chamfer matching. It is shown that the accuracy of original chamfer matching is 83.5% over 636 digits. On the other hand, it is shown that the accuracy of the case using the distance map is 94.2%. This result demonstrates how spatial feature representation using a distance map is a better method than original chamfer matching. The second experiment is

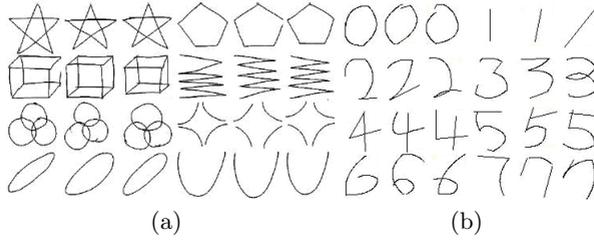


Fig. 5. Examples of sketch in our database. (a) figure sketches (b) numeric symbols.

the comparison experiment concerning the combination of two different features. We experiment with three different cases. One case uses only a directional feature, another uses only a spatial feature, and the other uses both together. It is shown that the accuracy of the case only using directional feature is 52.5%, while another case only using distance map as spatial feature is 87.7%. However, in the case of using both features together, the accuracy is highly increased to 93%. It demonstrates that it is a powerful feature for a shape representation which combines a spatial feature using a distance map and a directional feature using direction histogram. Fig. 6 shows the experimental result of the two experiments with numeric symbols in our database.

Table 2 shows that differences of the accuracy by giving different weight to spatial and directional features in the proposed shape representation method. We use both sketch database and online handwritten digit database in this experiment. It is shown that our proposed method with both spatial feature and directional feature has a high accuracy rate. In particular, the case which is given 0.4 weight to distance map and 0.6 weight to directional feature has the highest accuracy, which is the average rate 96%, on our database. It proves that the shape representation method using distance map and direction histogram is robust in regards to shape variation and invariant toward the direction of strokes.

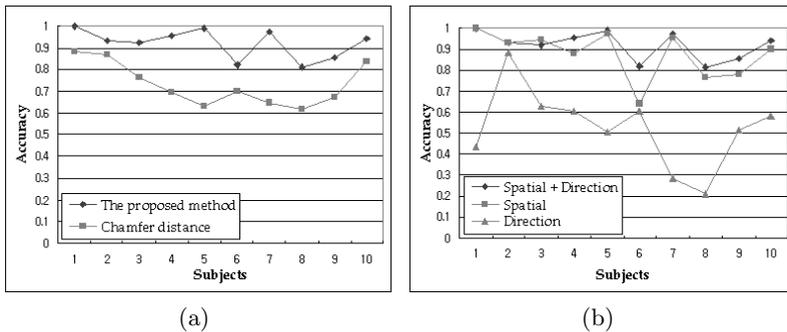


Fig. 6. The experimental result using numeric symbols. (a) The difference of accuracy between the proposed method and original chamfer matching. (b) The result of combination of spatial feature and directional feature.

Table 2. The result of experiment about all sketches in our database by different weight. In this table, w_1 is the weight to spatial feature, and w_2 is the weight to directional feature. Also, all values are the average of recognition accuracy in each database. Bold text is the highest accuracy rate for different weight.

Weight	Freehand Sketches	Numeric Symbols	Both Sketches
$w_1 : 1.0, w_2 : 0.0$	0.93	0.85	0.89
$w_1 : 0.8, w_2 : 0.2$	0.97	0.87	0.92
$w_1 : 0.6, w_2 : 0.4$	0.98	0.90	0.94
$w_1 : 0.5, w_2 : 0.5$	0.98	0.92	0.95
$w_1 : 0.4, w_2 : 0.6$	0.98	0.93	0.96
$w_1 : 0.2, w_2 : 0.8$	0.97	0.88	0.93
$w_1 : 0.0, w_2 : 1.0$	0.64	0.52	0.58

5 Conclusion and Future Work

In this paper, in order to recognize a sketch, we proposed the shape representation method being the combination of spatial and directional features on Tablet PC environment. It is obvious that the sketching order did not have to be considered in the spatial feature using distance map. Also, the experimental result demonstrated that the proposed template based on distance map is superior to the original chamfer matching. Furthermore, a directional feature is a good supplement feature to improve discriminant power. The proposed directional feature is more powerful on real application because it is unrelated to the drawing order. The experiment result showed 96% recognition performance about the experimental database which consist of 28 freehand sketches and 10 online handwritten digits.

In future work, we will study for making a classifier to recognize handwriting sketches more accurately by using this proposed feature representation method. Moreover, we will add more feature information to improve accuracy by giving weight to a discrimination area against other sketches. At this time, we will add alphanumeric symbols and complex mathematical expression symbols to experimental database with previous database. Through this research, we will make an unitive module to recognize various sketches, enabling us to go beyond characters, such as mathematical expression, geometrical figures and graph.

Acknowledgments. This research was supported by MIC, Korea under ITRC IITA-2008-(C1090-0801-0046).

References

1. Connell, S.D., Jain, A.K.: Template-based online character recognition. *Pattern Recognition* 34(1), 1–14 (2001)
2. Plamondon, R., Smithies, S.: On-line and off-line handwriting recognition: A comprehensive survey. *IEEE Trans. on Pattern Analysis and Recognition* 22(1), 63–84 (2000)

3. LaViola, J.J., Zeleznik, R.C.: MathPad2: a system for the creation and exploration of mathematical sketches. *ACM Trans. on Graphics* 23(3), 432–440 (2004)
4. Smithies, S., Novins, K., Arvo, J.: A handwriting-based equation editor. In: *Proc. of Graphics Interface*, Kingston, pp. 84–91 (1999)
5. Chan, K.F., Yeung, D.Y.: PenCalc: A Novel Application of On-Line Mathematical Expression Recognition Technology. In: *Proc. of the Sixth International Conf. on Document Analysis and Recognition*, Seattle, p. 774 (2001)
6. Matsakis, N.E.: Recognition of handwritten mathematical expression. Master's thesis, Massachusetts Institute of Technology (1999)
7. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. *IEEE Trans. Pattern Analysis and Machine Intell.* 24(4), 509–522 (2002)
8. LaViola, J., Zeleznik, R.: A Practical Approach for Writer-Dependent Symbol Recognition Using a Writer-Independent Symbol Recognizer. *IEEE Trans. on Pattern Analysis and Machine Intell.*, 1917–1926 (2007)
9. Thayananthan, A., Stenger, B., Torr, P., Cipolla, R.: Shape context and chamfer matching in cluttered scenes. In: *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Madison, vol. 1, pp. 127–133 (2003)
10. Sezgin, T.M., Stahovich, T., Davis, R.: Sketch Based Interfaces: Early Processing for Sketch Understanding. In: *Proc. 2001 Workshop Perceptive User Interfaces*. ACM Press, Orlando (2001)
11. Barrow, H.G., Tenenbaum, J.M., Bolles, R.C., Wolf, H.C.: Parametric correspondence and chamfer matching: Two new techniques for image matching. In: *Proc. 5th Int. Joint Conf. Artificial Intell.*, Cambridge, Massachusetts, pp. 659–663 (1977)
12. Borgefors, G.: Hierarchical chamfer matching: A parametric edge matching algorithm. *IEEE Trans. Pattern Analysis and Machine Intell.* 10(6), 849–865 (1988)