Pattern Matching for Industrial Object Recognition Using Geometry Based Vector Mapping Descriptor

D.S. Pae[1]*, Y.T. You[2], T. K. Kang[3] and M.T. Lim[1]†

[1]School of Electrical Engineering, Korea University of Seoul, Republic of Korea.
[2]Department of Mechanics, Korea University of Seoul, Republic of Korea.
[3]Department of Human Intelligence and Robot Engineering University, Cheonan, Republic of Korea.

*Presenting author: paeds915@korea.ac.kr
†Corresponding author: mlim@korea.ac.kr

Abstract

This paper introduces pattern matching algorithm from industrial camera based on geometric features. In industry production line, object recognition using vision is a challenging problem. In the object recognition, feature is an important element that represents object's state. Although its large amount of information contains location, rotation and scale difference, geometric feature is hard to get because it is sensitive to noise. To overcome the weakness, we propose two types of geometric features. Then geometric features are detected in order to construct descriptor. The geometry based Vector Mapping Descriptor (VMD) for pattern matching is proposed to effectively match salient feature points between different images under geometric transformation regardless of missing or additional feature points. VMD represents the correlation of features that includes Euclidean distance and angle. The group of one to one corresponding feature points on different images results in the completed object matching. The proposed algorithm is invariant to translation, rotation and scaling difference. To demonstrate the performance of the proposed algorithm, we conducted an experiment with both reference image data and real-time industrial camera. The result provides accurate and robust feature matching.

Keywords: Pattern Matching, Industrial Camera, Geometric Feature, Vector Mapping Descriptor

Introduction

Recently, numerous studies on pattern matching have been conducted in the industrial area. Some researches focus on complex image scene like crowded place and many buildings in city. These algorithms are based on obtaining certain point (feature) from image pyramid. Therefore the features are robust for each image pyramid. Lowe [1] suggested Scale Invariant Feature Transform (SIFT) algorithms. Ker et al. [2] improved SIFT algorithm by implementing the Principal Components Analysis (PCA) to the normalized gradient patch. Mikolajczyk et al. [3] proposed an extension of the SIFT descriptor and named Gradient Location and Orientation Histogram (GLOH). Other previous studies performed the object recognition through the shape based retrieval approach. Ling et al. [4] suggested the inner distance shape context (IDSC) which is extension of shape context (SC) [5]. SC is defined as comparing similarity of shape contexts for corresponding points. The inner distance is defined as the shortest path within edge of object shape and results in invariant shape articulation.
Matching can be performed by image segmentation method. Zhang and Ji [6] proposed image segmentation using a Bayesian network to object detection. Ferrari [7] also proposed contour segments for object detection. Dynamic Programming was used for distorted and occluded object retrieval by Petrakis [8].
However the previous methods are disadvantage to simple image like industrial component recognition where its background is mono-color and simple object. Firstly the current algorithms are too heavy. In industrial, the speed is one of the most important issues.

However the previous works did not satisfy the demand. The background is mono-color and the object is simple which indicates that there is less salient feature information. This causes the problem when the objects are occluded. They cannot build the relation between occluded objects due to the lack of number of features. Occlusion can cause building the insufficient relation. Therefore it leads to miss object matching.

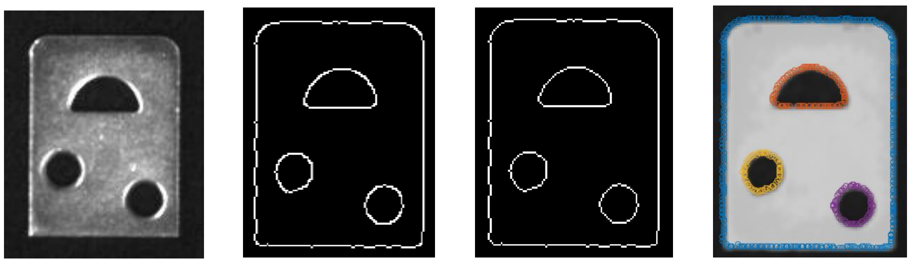
In this paper, we address the challenge of improving the efficiency and reliability of object matching in image processing. Pre-processing for object detection is applied on input images at a lower level of abstraction and its purpose is to reduce undesired outliers and enhance the useful image data which is important for further processing. The feature extraction process is an important procedure of object matching system. The feature contains unique, relevant information relating to the model object and target object. The geometry based Vector Mapping Descriptor (VMD) for pattern matching is proposed for object matching. In the proposed method, VMD for each feature is developed in order to obtain corresponding feature points between the model image and target image. Object matching is performed by the constructed descriptor by finding corresponding feature points that cope with both distortion and occlusion and it is invariant to geometric transformations.

The rest of this paper is organized as follows. Section 2 introduces the preprocessing by edge-enhancement using image sub-pixeling. Section 3 proposes two types of geometric features. Section 4 represents the VMD algorithm between model and target. In section 5, our experimental result is conducted with three different cases on both image data set and real-time camera. Section 6 discusses the result on previous experimental section and concludes proposed algorithm with the future works of what is left and how to improve our system.

2. Contour Detection in Sub-Pixel Range

The geometric feature is obtained from the contour of object. Thus, we need to detect contour to get geometric information. Contour detection process has three steps. First step is edge extraction. In edge extraction, we use canny edge algorithm. Second step is edge thinning. Result of edge extraction has shape of thick line. It is hard to get geometric information from thick line. Edge thinning method makes thick line to thin line. The third step is edge linking. Edge linking method links the edge and neighbors. Linked edges are contours of the object.

General method, explained earlier, has low accuracy and hard to extract geometric feature from contour. So we add two processes for improving performance. We apply sub-pixel approximation to increase accuracy. Afterward, we take a process called edge enhancement. Edge enhancement is required to get smoother edge for later stage's geometric feature extraction.



(a) Input image (b) Edge extraction (c) Thinning (d) Edge linking

Figure 1. Contour detection progress

2.1 Sub-pixel

Edge detection algorithms such as Sobel Operator, Robert's cross operator, Prewitt's operator calculate the edge's position information as pixel units. However pixel units have limitations in detecting accurate features such as circles and lines. The sequence of pixels becomes straight instead of becoming smooth. Application of sub pixel [9] units in image processing algorithms is suggested to improve the accuracy of position information.

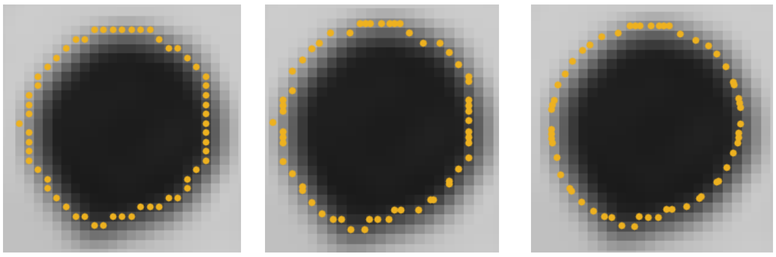
2.2 Edge Enhancement

The detected sub pixel unit edge is more suitable than pixel unit edge but it can be processed further to smoother edge. Therefore the edge is improved for geometric feature extraction. We proposed the method which revises the location of the edge point by using the relation of the neighbor points. In this case, we use two fitting methods that are line fitting and circle fitting [10]. Each fitting equation is shown as follows in (1) and (2).

$$E_l = \frac{1}{N}\sum_{i=1}^{N}\frac{(ax_i + by_i - 1)^2}{a^2 + b^2}. \tag{1}$$

$$E_c = \frac{1}{N}\sum_{i=1}^{N}\left(\sqrt{(x_i - c_x)^2 + (y_i - c_y)^2} - r\right)^2. \tag{2}$$

where N is number of edge points, edge point pixel position (x_i, y_i), secondary line equation parameters a and b, circle center position (c_x, c_y) and circle radius r. Calculate fitting errors and choose smaller one. If revise distance is longer than one pixel, the edge point is fixed.



(a) result of edge linking (b) sub-pixeling (c) edge enhancement

Figure 2. Results on sub-pixeling and edge enhancement

3. Geometric Feature Extraction

Geometric features are extracted from contour of the object. One object has contour set $\mathbf{C} = \{C_1, \cdots, C_k\}$. Contour element C_i has edge points set $\mathbf{E}_{c_i} = \{C_i(x_1, y_1), \cdots, C_i(x_n, y_n)\}$. To present the object, we need the features that are accurate and robust at environment. In this section we define the geometric features that are used for creating Vector Mapping Descriptor. We define two types of geometric features: (a) circle and its center, and (b) line segments and their intersection points. This section is divided into two parts as two main geometric features are employed. We begin with detecting the circle information (a). Once this step is completed, remaining contours are examined to determine whether it is a line or not. The defined line segments are extended imaginarily to produce the intersections among the lines used for salient feature points.

3.1 Circle Feature Detection

The least square circle fitting method is used to represent the circle along the contour. According to the following equation (2) we calculate parameter to minimize the circle fitting error. If circle fitting error is smaller than user parameter, we define the center of circle as geometric feature point. If not, we consider the contour non-circular regions.

3.2 Line Feature Detection

Once inspection of circles from the object is completed, the remaining contour regions are indicated as lines. The least square line fitting algorithm (1) is used to detect line segments. Firstly one point $C_i(x_p, y_p)$ is selected from the remainder contour. The selected point is the starting point of the line segments if the point and its sequence of the neighbor points $\{C_i(x_p, y_p), C_i(x_{p+1}, y_{p+1}), \cdots, C_i(x_{p+q}, y_{p+q})\}$ construct the line, L_k by using equation (1). The line needs to be formed by satisfying equation (1) within condition of initial number (q) of neighbor point. Once the starting point $C_i\left(x_p, y_p\right)$ is established, the initial line, L_k is formed and its seed starts to grow, adding neighbor points, until the equation (1) is not satisfied. This procedure is repeated until all of the remaining contours are inspected. The noise causes the separation of one long line segments into several line segments. The line is compared with neighbor lines by equation (1) so that it can be combined or one left to be separated. Therefore the line segments are completely detected. However the line fitting still contains error. We need to minimize the errors by shortening the length of the line until it has least amount of errors. The shortening process is done by cutting both side of line segment points until it has minimum error. The parameter can be determined to minimum range of cutting of line segment points.

Now we have line segments with least error. However, the shortening process causes the line segments to have different length which means they have different starting point and ending point. These different results in line segments cause the instability in vertex and these features are unable to be used as salient feature. Thereby we apply the intersection of lines as salient feature. This creates another concept of vertex. The defined extended line segments cause the intersections. The intersection points are robust regardless of differences in length of line segments thus can be used as salient feature.

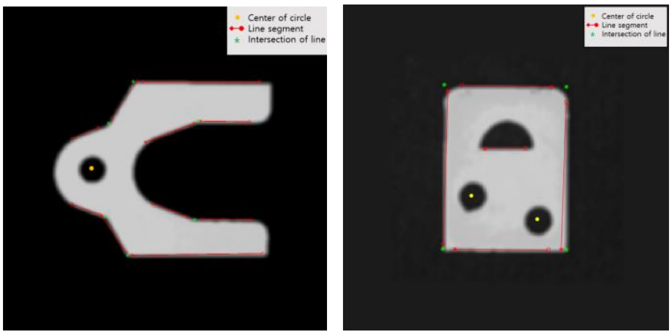


Figure 3. Results on geometric feature extraction

4. Constructing Vector Map Descriptor and Matching Algorithm

In this section we describe the three steps of our proposed geometric matching method: constructing the descriptor, one to one feature matching and finding desired point with completed object matching. Firstly we need the descriptor that is unique identity for an object. Then we compare each descriptors from one feature to the other features. Lastly, the position of objects in target is calculated according to the user's setting point from object on model.

4.1 Constructing Vector Mapping Descriptor

Descriptor is defined as containing unique information of the object. This descriptor is used as recognizing the objects in different scene. The target objects could be single or multi. Our descriptors are created from geometrical feature F which defined at the Chapter 3. The relationship between the defined features is represented as vectors which are separated to form of distance descriptors and angle descriptors. Figure 4 represents the feature locations and feature vectors contain distances and angles.

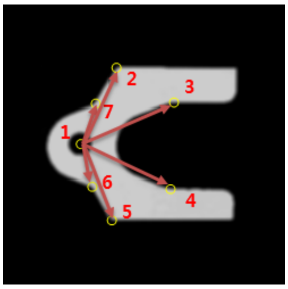


Figure 4. Vector mapping descriptor

Vector between two feature points can represent as follows:

$$\vec{v}_{ij} = F_i - F_j. \quad (3)$$

where $i \in (1,\cdots,n)$ and $j \in (1,\cdots,n)$, and n represents the number of features. Its distance descriptor D and angle descriptor Θ have n x n matrix size. The element of distance descriptor d_{ij} and angle descriptor θ_{ij} are denoted as follows:

$$d_{ij} = \left\|\vec{v}_{ij}\right\|, \ \theta_{ij} = \angle\vec{v}_{ij}. \quad (4)$$

4.2 One to One Feature Matching Using VMD

Object matching indicates that one feature in object can only have one corresponding feature in another scene. Thus, the corresponding feature set should only contain unique correspondences between features. Object matching indicates that one feature in an object can only have one corresponding feature in another object. Thus, the corresponding feature set should only contain unique correspondences between features. To achieve one to one feature matching, we use the row matrices of object descriptor that are feature descriptor distance $\bar{D}_i$ and feature descriptor angle $\bar{\Theta}_i$. In this paper, we use subscript M to state model variable, e.g. I_M, subscript T to state target variable, I_T. The rotation and scaling relationship for arbitrary features F_M from model to feature F_T in target are denoted respectively, by

$$\sigma = d_{M_{ij}} / d_{T_{ij}}, \quad \Delta\theta = \theta_{M_{ij}} - \theta_{T_{ij}}. \quad (5)$$

Once the arbitrary scale factor and rotation angle are calculated, assume that the target image I_T is enlarged by σ and rotated by $\Delta\theta$.

The feature descriptor distance and the feature descriptor angle are similar to the enlarged and rotated feature descriptor distance and feature descriptor angle.

Vector Mapping Descriptor algorithm

for i from 1 to N
  for j from 1 to N
    $\sigma = d_{M_{ij}} / d_{T_{ij}}, \ \Delta\theta = \theta_{M_{ij}} - \theta_{T_{ij}}$
    $E_{matching} = \sum_{k=1}^{N} \sqrt{\left(d_{M_{ik}} - \sigma d_{T_{ik}}\right)^2 + \left(\theta_{M_{ik}} - \theta_{T_{ik}} - \Delta\theta\right)^2}$
    save feature pair in array
  end
  Find feature pair (F_P) having minimum matching error
  if F_P < Threshold
    Calculate matching position, angle and scale.
  end
end

4.3 Completed Object Matching and Desired Detection Point

The geometric relations between the set of corresponding feature points in two different images are derived from the rotational angle and distance. Once all sets of corresponding feature points are obtained, it is compared with number of feature points in model image to the number of feature points in target image. The features in the model image are standard. The desired point (x_T, y_T) at target image can be calculated from each corresponding feature points; the x direction and y direction vectors of model, $\vec{v}_{xi}$ and $\vec{v}_{yi}$, rotation angle, $\Delta\theta$ and feature points from target $\left(F_{T_x}, F_{T_y}\right)$. The desired point from the original image can be detected in transformed image respectively by,

$$\begin{bmatrix} x_T \\ y_T \end{bmatrix} = \begin{bmatrix} \cos(\Delta\theta) & -\sin(\Delta\theta) \\ \sin(\Delta\theta) & \cos(\Delta\theta) \end{bmatrix} \begin{bmatrix} \vec{v}_{xi} \\ \vec{v}_{yi} \end{bmatrix} + \begin{bmatrix} F_{T_x} \\ F_{T_y} \end{bmatrix}. \tag{6}$$

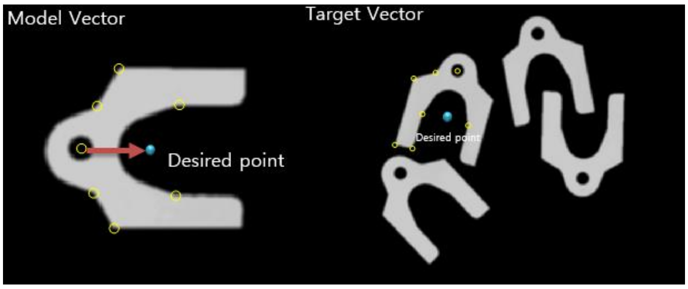


Figure 5. Desired detection point in target object

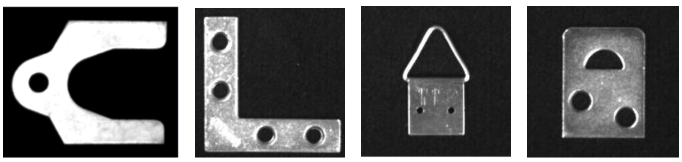
5. Experimental Result

5.1 Experimental Environment

The experimental environment is set as industrial inspection system; mono-color background is used and simple object is recognized. Mono-color background leads the clear separation between the background and foreground. The clear separation also helps the clear edge detection in preprocessing on section 2. The clear edge has close relationship on feature accuracy on section 3 and this results from the desired feature, center of circle and intersection of lines, detection as mentioned on section 3. To reduce the time consuming we limited the intersection of line feature to be formed on the edge and its neighbor only. Finally the Vector Mapping Descriptor algorithm is performed by using the relation between the geometric features.

5.2 Experimental Matching Result

We performed the experiment with four different model images and tested with 20 different target images. The average errors for each models are calculated to evaluate the proposed algorithm that is provided in Table 1. The total process times are shown in Table 2.


(a) Model 1 (b) Model 2 (c) Model 3 (d) Model 4
Figure 6. Model image

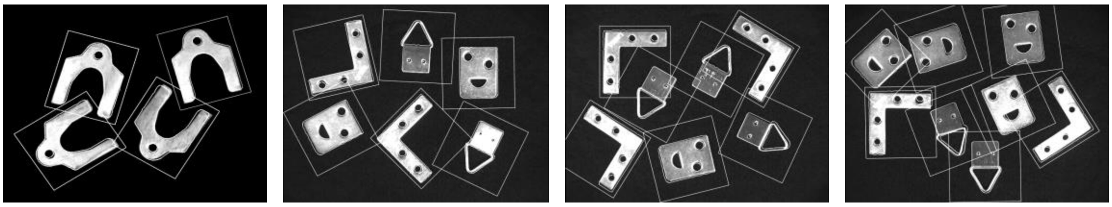


Figure 7. Results on target image

Table 1. Average error in geometric accuracy

Model	*X-axis (pixel)*	*Y-axis (pixel)*	*Scale (rate)*	*Angular (deg)*
Model 1	0.4838	1.0302	0.0132	1.0135
Model 2	0.5625	0.6150	0.0240	5.4371
Model 3	1.1881	1.4383	0.0094	0.9606
Model 4	0.8402	0.7481	0.0072	0.4835

Table 2. Processing Time

Processing step	*Average Computation Time (s)*
Filter	0.1062
Edge detection	0.0464
Edge linking	0.4586
Feature extraction	0.0842
Matching	0.0536
Total	0.749

The result of Table 1 shows efficiency of our method. In X-axis and Y-axis location accuracy, all average errors are lower than 1.5 pixel. Also, total process time is shorter than 1 second as shown in table 2.

6. Conclusion

In this paper, we start from demonstrating how to obtain fine images from coarse images. It results in obtaining robust geometric features. The strong geometric features are introduced for creating VMD algorithm which is proposed to solve the object recognition problem. Based on the results, it is shown that the proposed matching method is invariant to arbitrary geometric transformation: translations, rotations, scale changes, extra or missing feature points and the occlusion. The speed of matching algorithm is also fast enough for industrial recognition system. Moreover the VMD algorithm, proposed algorithm can be performed with any case if the feature of object in model and target are similar.

In future, we have a plan to improve the performance of VMD algorithm. We will find the optimal parameters for recognition system. Our research will focus in 3D data sets and their occlusion as well. The distance and angle descriptor extend to x, y and z axis can be one of the solution. Representing 3D image to 2D image by projection is also considered as solution. Image matching and registration are the foundation for many computer vision and its application such as navigation and security surveillance by recognizing the desired objects.

References

[1] Lowe, David G. "Distinctive image features from scale-invariant keypoints." International Journal of Computer Vision, 2004, Vol. 60, No.2, pp. 91-110.

[2] Ke, Yan, and Rahul Sukthankar. "PCA-SIFT: A more distinctive representation for local image descriptors." Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, 2004, Vol. 2, pp. 506-513.

[3] Mikolajczyk, Krystian, and Cordelia Schmid. "A performance evaluation of local descriptors." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, Vol. 27, No. 10, pp. 1615-1630.

[4] Ling, Haibin, and David W. Jacobs. "Shape classification using the inner-distance." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, Vol. 29, No. 2, pp. 286-299.

[5] Belongie, Serge, Jitendra Malik, and Jan Puzicha. "Shape matching and object recognition using shape contexts." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, Vol. 24, No. 4, pp. 509-522.

[6] Zhang, Lei, and Qiang Ji. "Integration of multiple contextual information for image segmentation using a bayesian network." Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, 2008, pp. 1-6.

[7] Petrakis, Euripides G. M., Aristeidis Diplaros, and Evangelos Milios. "Matching and retrieval of distorted and occluded shapes using dynamic programming." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, Vol. 24, No .11, pp. 1501-1516.

[8] Ferrari, Vittorio, et al. "Groups of adjacent contour segments for object detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, Vol. 30, No. 1, pp. 36-51.

[9] Canny, John. "A computational approach to edge detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, Vol. 6, pp. 679-698.

[10] Gander, Walter, Gene H. Golub, and Rolf Strebel. "Least-squares fitting of circles and ellipses." BIT Numerical Mathematics, 1994, Vol. 34, No. 4, pp. 558-578.