MVA '98 IAPR Workshop on Machine Vision Applications, Nov. 17-19, 1998, Makuhari, Chiba, Japan

6—4 A Practical Method to Recognize a Flat Object Using a Combination of Intensity Profiles

Yasukazu Okamoto * Eiji Tanaka
Kansai Research Laboratories
Toshiba Corporation

Abstract

We present a method to recognize a flat object using a combination of intensity profiles.

We are developing an automatic system that classifies packages according to their destinations, and also computes the delivery charges.

Our method represents an object as a set of intensity profiles on the line segments that connect two points. Because an intensity profile on a line segment normalized on the line length is invariant to affine transformation, similar line segments in a reference image can be recognized by only comparing the intensity profiles. The corresponding corner point of the object is detected by voting the corner points of the recognized line segment.

Experimental results show that the system can recognize a package tag and ID code with an objective accuracy.

1 Introduction

The recent rise in the number of packages being transported has created a need for a package sorting system. We are developing an automatic system that classifies packages according to their destinations, and also computes the delivery charges. In this paper, we propose a system which employs a new practical method for recognizing an affine deformed flat object such as a package tag.

The size and weight of a package, the destination code and the package ID code specified on the package tag are necessary to classify the package and to determine the delivery charge. The proposed system not only measures the package size, but also identifies the ID code and the destination code by observing images of a package carried on a conveyer.

In order to construct such a system, two key problems have to be solved. The first problem is that of acquiring a well-focused high-resolution image. Since conventional TV camera has insufficient resolution to read the ID code and the destination code characters, we equipped the system with a line scan camera to obtain an image with sufficient resolution. However, a line scan camera needs a fine focus control because of the high resolution. For this reason, we measure the package size with stereo images and adaptively control the focus according to the package size. Hence, the system is able to capture a high-resolution image of a package in focus with the line scan camera and recognize the ID and the destination code.

The second problem is that the resolution of the line scan image differs depending on the scanning direction. Whereas the resolution aligned with the scan line is defined by the size of sensor cell, the resolution vertical to the scan line is defined by the velocity of the object motion. Therefore, the line-scanned image of a tag on the package is deformed by an affine transformation.

Some methods previously developed are available for recognizing an affine deformed flat object. Murase et al. [1] represented an object as a pattern feature, and developed a system that recognizes a rotating object. By generating a quantities of pattern features that are gradually deformed from the original pattern feature while changing deformation parameters, they derived a relation between deformation parameters and the principal components of the pattern features by principal component analysis. Although, this method is robust to noise because it uses pattern features, the problem is that the relation becomes ambiguous when the number of deformation parameters increases.

The Geometric Hashing technique [2] represents an object as a set of geometric features, such as points or lines. A geometric relation of plural features is encoded to a value that is invariant to projection, and the object is recognized with the invariant value. However, this method also has problems in that the invariance against noise is unstable and the number of point combinations is huge. Since an affine invariance is computed from four points, The system must test the great number of ${}_nC_4$ point combinations.

In this paper, we propose a practical method for recognizing a package tag using a combination of intensity profiles. Our method represents an object as a set of intensity profiles normalized on the line length. An intensity profile is an intensity pattern

*Address: 6–26 Motoyama minami-cho, 8-chome, Higashi-nada-ku, KOBE 658-0015 JAPAN. E-mail:yaskaz@krl.toshiba.co.jp

feature on the line segment that connect two feature points.

The proposed method has three advantages.

1. It is robust to noise because it uses profile pattern features.

2. It can recognize a deformed object with any affine transformation because the profile pattern normalized on the line length is invariant to affine transform.

3. It is fast because the system has only to examine $_nC_2$ combinations.

2 System

Figure 1 shows an outline of the system. This system consists of two PCs and three cameras. First, the system takes a sideview and an overview picture of a package moving on a conveyer. The corner of the package is detected and the system measures the width, height and depth of the package by stereo matching. Next, the system adaptively controls the focal length of the line scan camera according to the measured package height. Subsequently, the system captures a high-resolution image of a package with the line scan camera, and recognizes the ID characters from the line scan image.

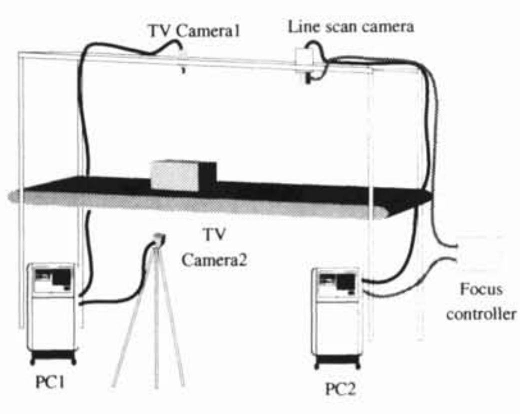

Figure 1: System Overview

3 Algorithm

The system selects a certain image as a reference image, and recognizes an object by comparing intensity profiles in an input image with the profiles in the reference image.

The corner points in the reference image are preliminarily detected, and a set of intensity profiles on the line segments connecting the detected corner points is obtained. The system makes a reference table to recognize the object with the set of intensity profiles.

Because an intensity profile on a line segment normalized on the line length is invariant to affine transformation, similar line segments in a reference image can be recognized by only comparing the intensity profiles. The corresponding corner point of the object is detected by voting the corner points of the recognized line segment.

In the following, the algorithm is described in detail.

3.1 Intensity Profile

First, the system selects a region of the object from the reference image. The recognition object in this system is a package tag, and some character strings are written in the address field of the package tag. The region in which character strings are written is inappropriate for object recognition, because the texture of the region differs depending on each sample. Therefore, a region that contains a logotype printed on the package tag is selected to represent the object.

Next, detecting corner points by Harris corner detector [3], the system selects the corner points in the object region, and generates an intensity profile on the line segment that connects selected corner points.

Suppose corner point P_0, and P_1 in the reference image are projected on point P'_0, and P'_1 in an input image with the following affine transformation.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} e \\ f \end{pmatrix}. \quad (1)$$

A point P on the line segment P_0P_1 lying in $P = sP_0 + (1-s)P_1$ is projected on a point $P' = sP'_0 + (1-s)P'_1$ with the above transformation. Therefore, an intensity profile on the line segment normalized on the line length is invariant to the affine transformation, because a positional relationship on a line segment is uniquely maintained.

3.2 Profile Similarity

The profile in an input image is compared with the profiles in the reference image, and similar profiles are selected.

Suppose $\mathcal{S} = \{\mathbf{S}_1, \ldots, \mathbf{S}_n\}$ is a set of n normalized intensity profiles in the reference image, and $\mathbf{S}_i = (s_i^1, s_i^2, \ldots, s_i^l)^T$ is an individual intensity on the profile. l is the length of the profile.

The Euclidean distance of profile $\mathbf{S}_i$ and $\mathbf{S}_j$ is defined as follows.

$$d(\mathbf{S}_i, \mathbf{S}_j) = \sum_{m=1}^{l} (s_i^m - s_j^m)^2. \quad (2)$$

Selection of similar profiles is equivalent to selection of the profiles such that the Euclidean distance does not exceed the certain threshold value d_t. This selection of a similar intensity profile costs $O(nl)$ computational complexity.

However, this computational complexity can be reducted using Karuhunen-Loève expansion[4]. Let $\mathbf{V}$ be the covariance matrix of the vectors in $\mathcal{S}$. We denote the eigenvalues of $\mathbf{V}$ as λ_i, and the corresponding eigenvectors as $\mathbf{e}_i$. The eigenvalues λ_i are ordered in the descending order. $\mathbf{S}_i$ in $\mathcal{S}$ is represented sufficiently with only first $K, (K < n)$ eigenvectors. Therefore, $\mathbf{S}_i$ is approximated sufficiently by the following $\widetilde{\mathbf{S}}_i$.

$$\widetilde{\mathbf{S}}_i = \sum_{j=1}^{K} \alpha_{ij} e_j. \tag{3}$$

α is the projection of the profile onto the eigenspace.

$$\alpha_{ij} = \mathbf{S}_i^T e_j. \tag{4}$$

Because the eigenvectors are orthogonal to each other, and the norm equals to 1, the distance of profiles is represented by the following equation.

$$d(\widetilde{\mathbf{S}}_i, \widetilde{\mathbf{S}}_j) = \sum_{m=1}^{K} (\alpha_{im} - \alpha_{jm})^2 \tag{5}$$

Using Karuhunen-Loève expansion, computational complexity is eliminated to $O(K(n+l))$.

Computational complexity can be reduced further by using the following equation.

$$\sum_{m=1}^{t} (\alpha_{im} - \alpha_{jm})^2 \leq \sum_{m=1}^{K} (\alpha_{im} - \alpha_{jm})^2 \quad (t \leq K) \tag{6}$$

Equation 6 shows that the distance must exceed the threshold value d_t when the distance computed with the first t eigenvalues exceeds the threshold value. Therefore, the system need not further computation when the distance computed with the first t eigenvalues exceeds the threshold value.

3.3 Reference Table

Figure 2 shows a reference table made from the intensity profiles in the reference image. The system recognizes an object in an input image, using the reference table.

Each rectangle in the left of the figure shows a list of α value in equation 4 computed with the intensity profiles in the reference image. The list is ordered in increasing order. The rectangles show the α value corresponding to the first K eigenvalues.

Each rectangle in the right of the figure shows a data table corresponding to a line segment in the reference image. The table contains the first K α values computed with the intensity profile on the line segment, and index numbers of the corner points which are the two endpoints of the line segment. The segment data table is linked from the α value list, and can be retrieved from the α value.

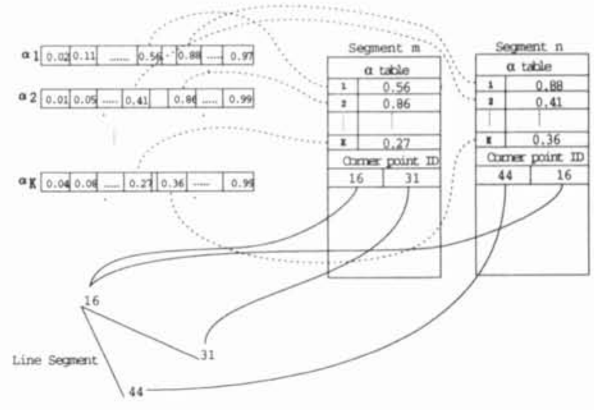

Figure 2: Reference table

3.4 Recognition by voting

In the following, the recognition flow using the above reference table is shown.

First, detecting corner points from an input image, the system chooses an arbitrary corner point P_0, and a point P_i in the neighborhood of P_0.

Next, the system retrieves segment data from the reference table whose intensity profiles are similar to the normalized intensity profile on line segment P_0P_i. The projection α'_j value is computed by equation 4 from the normalized intensity profile on the line segment P_0P_i. The system search the α_j value list to find the values similar to α'_j. The system iterates the above process increasing j, and retrieve the segment data $\mathbf{S}_a$ that is similar to the profile on segment $P_0P_i(\sum_{j=1}^{K}(\alpha_j - \alpha_{ij})^2 < d_t)$.

The corner point indices in the retrieved segment data show the end points of the similar line segment in the reference image. If the corner point indexes show P_a and P_b in the reference image, it implies that the point P_0 is likely to correspond to the point P_a.

Due to the ambiguity of the profile pattern, plural segments may be obtained by the retrieval. However, correct correspondence to the point P_0 is recovered by the voting method as shown in Figure 3. The system retrieves segments similar to the segment P_0P_i while replacing different candidates of P_i. Subsequently, the system votes to the points corresponding to P_0 represented by the retrieved segments.

Although the correspondence can be ambiguous, the highest vote P_a in the result of voting for the points corresponding to P_0 indicates the correct correspondence to P_0.

The package tag is recognized after verifying the correspondence of line segments that include the point P_a. The position of a character string is computed from the recognized package tag position, and the ID characters are recognized from the clipped image on the computed position.

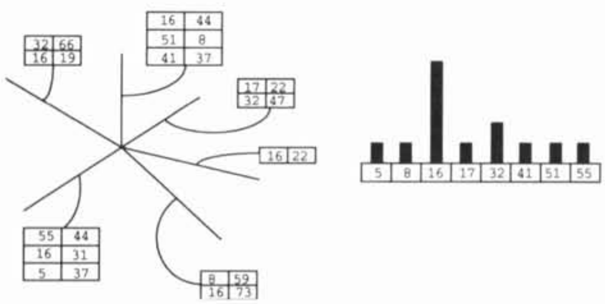

Figure 3: Voting

4 Experimental Result

We have carried out some experiments to evaluate the proposed system. 52 images of three packages are used in the experiment with a variety of package directions and the tag directions. Each package is 60cm, 20cm or 10cm in height. Figure 4 shows a recognition result obtained by the proposed system. The larger rectangle overlaid on the original image shows the recognized tag position and direction. The smaller rectangle shows the position of tag ID code derived from the tag position. Figure 5 shows another recognition result. The tag ID code image that is clipped with the recognition result is shown below. The recognized tag ID code is shown in the right part of the image. The proposed system recognized 48 tags in 52 images, and read all the ID code on the recognized tags properly. The experimental results show the efficiency of the proposed system.

5 Conclusion

We have presented an automatic system that classifies packages and a method for recognizing package tag used in the system.

Our method represents an object as a set of intensity profiles on the line segments that connects two points. Because an intensity profile on a line segment normalized on the line length is invariant to affine transformation, similar line segments in a reference image can be recognized by only comparing the intensity profiles. The corresponding corner point of the object is detected by voting the corner points of the recognized line segment.

Experimental results show that the system can recognize a package tag and ID code with an objective accuracy.

References

[1] H. Murase and S. Nayer, "Visual Learning and Recognition of 3-D Objects from Appearance", International Journal of Computer Vision, vol4-1, pp.5-24, 1995.

[2] H.J. Wolfson and Y. Lamdan, "Transformation Invariant Indexing", in Geometric Invariance in Computer Vision, Edited by J.L. Mundy and A. Zisserman, pp.335-353, MIT Press, 1992.

[3] C. Harris, "Geometry from Visual Motion", in Active Vision, Edited by A. Blake and A. Yuille, pp.263-284, MIT Press, 1993.

[4] E. Oja, "Subspace Methods of Pattern Recognition", Research Studies Press, 1983.

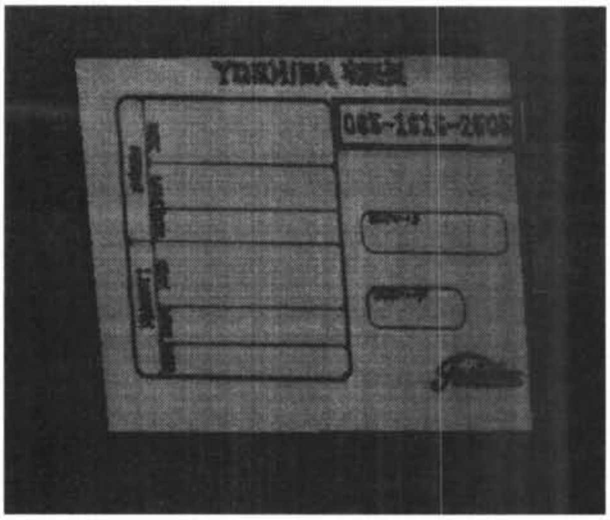

Figure 4: Recognized package tag

Figure 5: Recognitition result of ID code