A SYSTEM FOR RECOGNIZING FORMS AND CHARACTERS ON CYLINDRICAL OBJECTS

Tapio Seppänen, Pasi Hyttinen, Matti Pietikäinen, Hannu Hahtala *), and Jukka Rajala *)

University of Oulu, FIN-90570 Oulu, FINLAND
*) Sypal Inc., FIN-78250 Varkaus, FINLAND

ABSTRACT

This paper presents an experimental system capable of recognizing forms and printed characters positioned on cylindrical objects in an industrial application. Forms are classified with structural pattern recognition methods while characters are recognized with a hybrid classifier utilizing both statistical and structural information. Training of form and character classifiers is performed using a teach-by-showing principle.

INTRODUCTION

The goals of machine vision based document analysis include recognizing document in a digital image and extracting information contained in the document. The implementations usually apply structural pattern recognition methods to recognize the documents. A short review of structure recognition methods is given in [1]. Papers [2] and [3] concentrate on forms processing which is also a topic of our application. The ultimate target in our application is text recognition. There is extensive literature on optical character recognition. Good reviews of OCR are given in [4] and [5], while [6] presents many successful applications.

This paper presents a system capable of recognizing forms and printed characters in an industrial application. The application involves cylinder like packages which are transferred around a factory via a conveyor belt. A form to be analyzed by machine vision is attached to the cylinder (see Fig. 1). The cover material on the cylinder includes unknown texture and visual objects. The character set of the application contains only numbers but the character classifier was designed to accept any character set.

The application software runs on a 80486 based PC under OS/2. The execution time is currently about one second which includes time to recognize the form and interpret the text fields.

FORM RECOGNITION

A form includes text fields with enclosing rectangular boxes. Text contains letters and numbers to be located and recognized. The form also has a barcode on it, which will serve as a landmark for form recognition.

The recognition task is performed by structural pattern recognition. The basic idea is to define a form specific set of test points which is projected on the input image. A local neighborhood is investigated near each test point in order to determine the degree of match to model data.

A complication is the cover of the cylinder which may have almost anything printed on it. Text, lines, barcodes, and figures are typical texture on the cover. Light reflections are also possible if the cover is of plastic material.

The form might not be in an upright position but may be rotated slightly, which introduces distortions to the image. The form may also have been attached upside down to the cylinder.

The enclosing box of the form is rectangular in shape. The perspective distortion introduced by cylinders with a small radius of curvature, however, transforms straight lines to curves and makes recognition difficult.

A distinct landmark on the form is a barcode in a fixed location. There are barcodes also in the surrounding cover material, but due to the fixed structure of the form the neighborhood of each detected barcode can be checked to reject false alarms. The basic problem is then to reliably detect all barcodes in the image. We developed a nonlinear filter which produces strong responses in locations containing a barcode [8]. See Fig. 2 for an example result.

The thresholded response peaks of the filter are mapped to an adaptively thresholded original image. The neighborhood of each peak is analyzed to locate the four corner points of the barcode candidate. If four corners are found, the candidate is accepted for further analysis. The rotation and the scale of the form are calculated from the known geometry of the barcode.

Wrong barcode candidates are rejected in a verification phase in which the surroundings of each candidate are checked for other landmarks that belong to the form. The lines of the enclosing and inner boxes are used for this purpose. They must be located in correct positions and directions. Some of the test points are defined on empty areas of the form. A set of test points is selected from the box lines by the user in the training phase of the system.

The number of successful matches of the test points must exceed a threshold value for acceptance of the barcode candidate. As the correct barcode has been identified, the form has also been recognized and its pose is known. Fig. 3 demonstrates results obtained by the algorithm steps described above.

Training of the form recognizer includes construction of a form description file. The file format is designed to contain the location of the barcode, enclosing boxes and test points. One of the form corners is selected as the coordinates origin.

CHARACTER RECOGNITION

Recognition of characters within required text boxes is performed by a 2-level hybrid classifier. The primary classifier performs an initial classification by statistical pattern recognition and the secondary classifier is used if necessary to make the final classification by template matching. The secondary classifier is needed because the primary classifier confuses some characters.

The primary classifier is a kNN classifier (k=1) which utilizes as features the descriptors generated by the complex Fourier transform of the outer contour of characters. The classifier ant the features were selected according to our earlier comparison of various methods [7]. Descriptors are modified to achieve invariance to scale, rotation, starting point and translation. The classifier is trained with a teach-by-showing technique. Some characters resemble each other to the extent that they are confused during classification. The confusions are detected automatically in the training phase. Mutually exclusive groups are then defined each of which contains characters which are confused with each other. The primary classifier forwards a character to the secondary classifier if it decides it belongs to one of the confusion groups.

A complication exists due to the cylindrical background of forms. The features used by the primary classifier are not fully invariant to the 3-D perspective distortions. The classifier is therefore trained by taking samples from several locations on the cylinder surface. The position of the character on the cylinder must be used to select the correct model database during the classification phase.

During the training of the secondary classifier, only one set of images is taken in which characters are viewed directly from above. The other sets are generated automatically by the training subsystem through simulating perspective projections. Confusion groups are determined for each location of training. In practice, confusion groups in different locations resemble each other because the same characters are confused everywhere on the cylinder surface. The generated confusion groups are, therefore, finally combined in order to create as few groups as possible.

The secondary classifier performs template matching by fitting selected skeleton points of character models to adaptively thresholded character images. Only those character models are used in matching which belong to the confusion group in question. The best match determines the final result. The result is accepted only if the matching score is above a predefined threshold value. The models are generated automatically in the training stage with a teach-by-showing method. The models are tested by classifying several examples with the secondary classifier. Confusion groups are collected from misclassified characters. If some of the groups are not empty, the operator selects a small number of additional test points to enhance discrimination power.

During classification, the secondary classifier is also influenced by the character position on the cylinder. The skeleton points defining the character skeleton are transformed in 3-space to make as good a fit as possible. Orthogonal projection works sufficiently well. The rotation and translation of the form and the size of the cylinder are already known at this stage, which is necessary for the transformation to take place.

Two text strings of a form must be recognized in this application. The result is verified by first recognizing a check sum field printed on the form and then comparing this against a check sum computed from recognized characters. Use of the check sum field increases reliability because the system knows when the recognition fails.

CONCLUSION

A machine vision system was described which is capable of recognizing forms and printed characters on cylindrical objects. The forms are classified with structural pattern recognition methods while characters are recognized with a hybrid classifier utilizing both statistical and structural information. A practical approach was developed for recognizing characters under highly varying 3-D perspective distortions.

The forms are recognized by using a barcode as a landmark and a set of test points. The barcode detection is performed robustly with a nonlinear filter even on a cluttered background.

The primary character classifier categorizes characters with a nearest-neighbor classifier using invariant Fourier-descriptors. Those characters which the primary classifier decides belong to confusion groups are forwarded to the secondary classifier. The secondary classifier performs final categorization using a template matching algorithm. The confusion groups are prior information which is produced during the training phase.

REFERENCES

[1] Watanabe T, Luo Q & Sugie N, "Structure recognition methods for various types of documents", Machine Vision and Applications, No. 6, 1993, pp. 163-176.

[2] Casey RG & Ferguson DR, "Intelligent forms processing", IBM Systems Journal, Vol. 29, No. 3, 1990, pp. 435-450.

[3] Taylor SL, Fritzson R & Pastor JA, "Extraction of data from preprinted forms", Machine Vision and Applications, No. 5, 1992, pp. 211-222.

[4] Mori S, Suen CY & Yamamoto K, "Historical review of OCR research and development", Proc. IEEE, Vol. 80, No. 7, July 1992, pp. 1029-1057.

[5] Pavlidis T, "Recognition of printed text under realistic conditions", Pattern Recognition Letters, No. 14, April 1993, pp. 317-326.

[6] Proc. IEEE, Vol. 80, No. 7, July 1992.

[7] Kauppinen H, Seppänen T & Pietikäinen M, "Shape classification with autoregressive and Fourier-based methods: a comparative study", 8th Scandinavian Conference on Image Analysis, 1993, Tromsø, Norway, pp. 1171-1177.

[8] Seppänen T. & Hyttinen P, "A gradient distribution based method for barcode detection", TUT Symposium on signal processing, 1994, Tampere, Finland, pp. 109-112.

Figure 1. An example image of an application situation.

Figure 2. Steps in barcode detection.

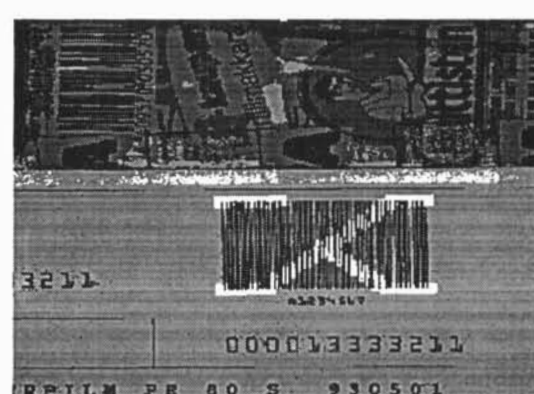

Figure 3. Final result of barcode locationing.