The trends in Figs. 7 and 8 are consistent for both data sets. The five-best-choices filtering, followed by the minimum distance measure, produces the best results, both at the extrema of the curves and in passing closest to the ideal (0 percent, 0 percent) point. The minimum distance measure alone is next best, and then the Euclidean distance. Fig. 8 shows the limitation of the five-best-choices assumption; the error rate cannot be reduced below nine percent because nine percent of the five best choices are in error and are, consequently, being accepted by the system without question. This assumption is much less problematic for the 96×96 data, where about 0.1 percent of the five best choices are in error.

Alterations to the recognition system and changes in the basic CQA assumptions might yield improved results. The six airplanes are fundamentally very similar objects. As noise is added to the images, and the resolution is reduced, the consistent differences between the different airplanes disappear. Techniques such as multiframe averaging or multiple views would increase the classification success [10] and, as a consequence, improve the results of CQA. Further refinements of the CQA technique are possible. For example, currently only magnitudes of distances in feature space are used to generate a scalar for CQA normalization; more powerful methods might be possible if vector feature space measures are used.

V. Conclusion

This correspondence has presented a method for making an object identification system, originally tested by recognizing a small set of objects, more general. This generality comes from the ability to reject objects that are not known to the recognition system. The resulting classification system still has the same computational complexity as before the addition of CQA. The strengths of the method include an emphasis on nonruntime computation and no dependence on data description techniques.

It is worthwhile to consider briefly the system time complexity. For an image of size $n \times n$, we calculate feature vectors that are complete moment sets (CMS) of order m. An order m CMS contains $O(m^2)$ elements, and requires $O(m^4)$ operations for normalization [15]. If we have p airplanes in our model database we will put v views of each into our library, then the overall library generation complexity is $O(pv(m^2n^2 + m^4 + m^2))$. The lone m^2 in the time complexity is from an overall library normalization performed at the end of library generation.

Once the library is generated, the CQA measures can be calculated. Calculation of CQA 1 for the entire library is $O(pvm^2)$. Calculation of the CQA 2 normalization measures is $O(p^2v^2m^2)$. All this represents precomputation prior to classification runtime. For the experiments presented here $n = 128$, $m = 3$, $p = 6$, and $v = 500$. In contrast, runtime complexity for classification is $O(n^2m^2 + m^4 + m^2 + pvm^2)$, with additional $O(1)$ time required for CQA 1 and 2. The runtime time complexity could be reduced by using parallel hardware applied along any or all of n, m, p, and v [16].

After unknown object rejection, a second concept presented in this correspondence is that of raising classification dependability, by rejecting less reliable classifications of objects that are in the data base of known objects. Thus if an action can be taken only if certainty of identification is 95 percent or better, this method provides a means for achieving this level of dependability for a subset of the total possible set of unknown feature vectors. This second analysis stage shares the computational characteristics of the first stage. It is also not dependent on data representations.

Overall, the CQA techniques presented allow an identification system, based on a limited object set, to function in a much broader context. Aside from the generality, alternative applications are also possible: in an industrial environment, such an enhanced system would allow fast bin picking, where the vision system would only need to recognize the parts that were required, and not all the types of parts in the bin. This would mean a system speedup, by reducing the size of the search space (along the same lines as [8]). CQA could also be used to recognize the deviation of defective parts from the ideal in a quality evaluation system.

References

[1] K. S. Fu, "Robot vision for machine part recognition," *Robot. Robot Sensing Syst.*, vol. 442, pp. 2–14, Aug. 25, 1983.

[2] P. J. Besl and R. C. Jain, "Three-dimensional object recognition," *Comput. Surveys*, vol. 17, pp. 75–145, 1985.

[3] R. C. Bolles and R. A. Cain, "Recognizing and locating partially visible objects: The local feature-focus method," *Int. J. Robot. Research.*, vol 1, Fall 1982.

[4] W. E. L. Grimson and T. Lozano-Perez, "Recognition and localization of overlapping parts from sparse data in two and three dimensions," in *Proc. Int. Conf. Robotics and Automation*, pp. 61–66, 1985.

[5] J. L. Turney, T. N. Mudge, and R. A. Volz, "Experiments in occluded parts recognition," in *Proc. Intelligent Robots: 3rd Int. Conf. Robotic Vision and Sensory Controls*, vol. 449, 1983 pp. 719–725.

[6] A. Kalvin, E. Schonberg, J. T. Schwartz, and M. Sharir, "Two-dimensional, model-based, boundary matching using footprints," in *Int. J. Robotics Res.*, vol. 5, pp. 38–55, Winter 1986.

[7] G. J. Ettinger, "Large hierarchical object recognition using libraries of parameterized model subparts," in *Proc. IEEE Computer Society Conf. Computer Vision and Pattern Recognition (CVPR)*, 1988, pp. 32–41.

[8] C. Goad, "Special purpose automatic programming for 3D model based vision," in *Proc. ARPA Image Understanding Workshop*, 1983, pp. 94–104.

[9] A. P. Reeves, R. J. Prokop, S. E. Andrews, and F. P. Kuhl, "Three dimensional shape analysis using moments and Fourier descriptors," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 10, pp. 937–943, 1988.

[10] A. P. Reeves, R. J. Prokop, and R. W. Taylor, "Shape analysis of three-dimensional objects using range information," in *Proc. 1985 IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, June 1985, pp. 452–457.

[11] D. H. Ballard and C. M. Brown, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 1982.

[12] T. P. Wallace and P. Wintz, "An efficient, three-dimensional aircraft recognition algorithm using normalized Fourier descriptors," *Comput. Graphics Image Processing*, vol. 3, pp. 99–126, 1980.

[13] S. A. Dudani, K. J. Breeding, and R. B. McGhee, "Aircraft identification by moment invariants," *IEEE Trans. Comput.*, vol. C-26, pp. 39–46, Jan. 1977.

[14] *100X 3D Scanner: User's Manual and Application Programming Guide*. Technical Arts Corporation, 1986.

[15] A. P. Reeves and A. Rostampour, "Shape analysis of segmented objects using moments," in *Proc. 1981 Pattern Recognition and Image Processing Conf.*, Dallas, TX, 1981, pp. 171–174.

[16] O. R. Mitchell, A. P. Reeves, and T. A. Grogan, "Algorithms and architectures for global shape analysis in time-varying imagery," *SPIE Proc., Robotics and Industrial Inspection*, vol. 360, pp. 190–197, Aug. 1982.

A Gray-Level Threshold Selection Method Based on Maximum Entropy Principle

ANDREW K. C. WONG, MEMBER, IEEE, AND P. K. SAHOO

Abstract—**The gray-level threshold selection method for image segmentation presented here is based on the maximum entropy principle. The optimal threshold value is determined by maximizing the *a posteriori* entropy subject to certain inequality constraints which are derived by means of special measures characterizing uniformity and the shape of the**

Manuscript received November 3, 1987; revised December 22, 1988. This work was supported in part by the Natural Science and Engineering Research Council of Canada.

A. K. C. Wong is with the Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada N2L 3G1.

P. K. Sahoo is with the Department of Mathematics, University of Louisville, Louisville, KY 40292.

IEEE Log Number 8927237.

0018-9472/89/0700-0866$01.00 ©1989 IEEE

Reproduced with permission of copyright owner. Further reproduction prohibited.

regions in the image. For this purpose, the authors use both the gray-level distribution and the spatial information of an image. The effectiveness of our method is demonstrated by its performance on some real-world images. An extension of this method to chromatic images has been provided.

I. Introduction

Image thresholding is essentially a pixel classification problem. Its basic objective is to classify the pixels of a given image into two classes: those pertaining to an object and those pertaining to the background. While one includes pixels with gray values that are below or equal to a certain threshold, the other includes those with gray values above the threshold. Thresholding is a popular tool for image segmentation. It is widely used in halftone reproduction [12], [14], [27], automatic target recognition [3], design of visual navigation system for autonomous land vehicles [25], [29], [30], industrial applications of computer vision [9], and biomedical image analysis [4], [6], [19], [23]. Over the years many threshold selection techniques have been proposed. For a survey of threshold selection techniques, readers may refer to [26] and [31].

In general, threshold selection techniques can be broadly divided into two groups, namely, global and local thresholding. A global technique may be point-dependent or region-dependent. The thresholding method is point-dependent if the threshold value is determined solely from the pixel gray tone as represented by gray-level histogram and is independent of the gray tone of the neighborhood of a pixel. On the other hand, a method is called region-dependent if the threshold value is determined from the local property within a neighborhood of a pixel. A global thresholding technique is one that thresholds the entire image with a single threshold value, whereas a local thresholding technique is one that partitions a given image into subimages and determines a threshold for each of these subimages.

Though point-dependent global thresholding methods are comparatively simple, they use only a small part of the independent information provided by the image. Therefore, these techniques are not always useful. For example, in halftone reproduction, global point-dependent methods rarely produce good pictures.

In this correspondence we propose a global threshold selection method which uses both gray-level distribution and spatial information. The work is organized as follows. In Section II, we briefly review some recent methods on thresholding. These methods have not been considered in the recent survey article [26] or in [31]. Section III describes the proposed thresholding method and its theoretical justifications. Section IV reports the effectiveness of our thresholding method when applied to some real-world images. In Section V, we show how this method is extended for segmenting chromatic images. Finally, in Section VI, we provide some concluding remarks regarding the method proposed.

II. Review of Recent Works

function $h(x, y)$ at gray level t is a binary image function $h_t(x, y)$. The explicit representation of $h_t(x, y)$ is expressed as

$$h_t(x, y) = \begin{cases} a_1, & \text{if } h(x, y) \leq t \\ a_2, & \text{if } h(x, y) > t. \end{cases}$$

The constant a_1 is usually 0 and a_2 is $l-1$.

Let $f_0, f_1, \cdots, f_{l-1}$ be the observed gray-level frequencies of an image. The probability of the gray level i in an image can be estimated as

$$p_i = \frac{f_i}{N_0}, \qquad N_0 = \sum_{i=0}^{l-1} f_i, \qquad i = 0, 1, \cdots, l-1$$

where N_0 is the total number of pixels in the image and l is the total number of gray levels. Dunn *et al.* [5] proposed a uniform error threshold selection method which equalizes the probability of misclassification between the object and the background pixels in an image. The optimal threshold value in uniform error method is obtained by finding a threshold $t \in Z_l^+$ such that

$$b(t)(b(t)+1) - (a^2(t) + c(t)) = 0$$

where

$a(t)$ = Pr {a pixel has gray value > t}

$b(t)$ = Pr {two adjacent pixels have gray values > t}

$c(t)$ = Pr {four neighboring pixels have gray values > t}.

The probabilities $a(t)$, $b(t)$, and $c(t)$ are estimated by examining 2×2 neighborhoods in the image for fixed t. The uniform error method is a global region-dependent thresholding method. Aviad and Lozinskii [1] proposed a thresholding method based on minimizing the ambiguity which results from the different contexts in which gray levels are classified as object or background. Pavlidis and Wolberg [22] have given an algorithm based on a model of the image distortion. Forte and Sahoo [7] introduced a measure for the loss of information due to reduction on the range of a real random variable and proposed a thresholding method based on the loss of information criterion. Their method belongs to the class of point-dependent global techniques. Another information theoretic method that recently appeared in the literature is due to Beattie [2]. Other recent publications related to automatic threshold selection include [10], [13], and [17].

III. Maximum Entropy Thresholding

The success of the maximum entropy method for solving problems in image reconstruction [32], [33] and other areas [13] has encouraged us to apply this method to threshold selection. As