Proceedings of Student/Faculty Research Day, CSIS, Pace University, May 7th, 2004

Interactive Flag Identification Using a Fuzzy-Neural Technique

Eduardo Hart, Sung-Hyuk Cha, Charles Tappert

Abstract

We propose an interactive system for identifying flags in photos taken from natural scenes. The system is interactive in two respects. First, because segmentation can be a difficult problem, users are asked to crop the whole flag area from a photo. Second, the user makes the final decision by selecting one of the top choices obtained from the machine classification system. The proposed system utilizes a color-based image retrieval technique. For experimental purposes a large number of flag images are synthetically generated from a small number of original ones in order to increase the reference image database. Feature extraction is based on a four-layer Fuzzy-Neural algorithm. A nearest neighbor classifier produces a sorted list of candidate choices. Recognition accuracy of these choices varies from 85% to 90% depending on whether the correct flag is among the first 10 or 20 top choices, respectively, from a set of 186 flags.

Keywords: *flag recognition ,image retrieval*

1. Introduction

Classifying an unknown flag in a scene is challenging due to the diversity of the data and to the complexity of the identification process. Flag images in scenes are often blurred, shaded and, because of wind or movement, furled (curled and possibly folded). These distortions greatly increase the difficulty of identifying flags embedded in real photos. We present a flag recognition system using the content-based, image-retrieval technique. Given a database of known flag images and an unknown flag image as input, the system retrieves all similar flag images in the database. Other applications of content-based, image-retrieval techniques have received attention (see the extensive survey [1]), and color components have been widely used to retrieve images [1-6]. We adopted this color-based image retrieval technique and, by comparing the color components of two flag images, we retrieve all similar flags.

Our proposed system requires preprocessing to locate and to segment a flag from the surrounding background in a photo. To perform this preprocessing, we propose an interactive recognition system similar to that of an earlier flower recognition system [7]. Users select the region of interest from a photo by selecting many points of the flag's edge and then the system retrieves all the similar flags. Similar approaches can be found in sign recognition systems where the studies have focused on identifying commercial, informative, and traffic signs. In two key studies [8, 9] the signs to be identified had two common characteristics: regular shapes (polygonal) and no motion (signs move together with the background). Because the signs were difficult to segment automatically, the best results in these studies required user intervention. Once the sign is located, those systems executed automatic adaptations, like rotation, slant, and tilt correction, that facilitated accurate identification of letters or symbols. Those adaptations, however, cannot easily be applied to flags because flags are usually non-static (waving, folded, etc.). Interestingly, not being static is an advantage for flag segmentation because the motion could be used to assist an automatic segmentation process. Being furled (not flat and two-dimensional), however, not only increases the difficulty of applying automatic shape correction to flags but it also increases the difficulty of accurately identifying symbols or letters as defining features because portions of the flags may be obscured.

The proposed system consists of three main stages: segmentation, feature extraction, and classification. Although automatic flag segmentation is possible due to the non-static property, for this preliminary study we have the user locate the flag by selecting many points of the perimeter containing the flag. Extracting features for the flag identification system is based on this concept. The target features are the colors and their corresponding percentage contribution to the selected portion of the flag. Identifying colors in real images is very difficult due to the shadows and bright areas. To solve this problem we developed a four-layer fuzzy-neural approach. The first two are based on fuzzy algorithms that help segment the image in true color areas and doubtful areas. In the third layer a small neural network is trained with the information of the areas found as true colors. This neural network is used to decide the true color of the doubtful pixels. The classification process uses the nearest neighbor method, and most of the references are synthetically generated from a small number of original flag images. For this study the references are based on 186 different flag classes and 36 samples per class. The testing set is composed for 33 different classes and one sample per class. We tested the system by comparing the results of feature extraction based on crisp logic and on the four-layer Fuzzy-Neural approach. The crisp algorithm is based on a set of if-else statements to process the HSV in order to decide the corresponding true color.

The subsequent sections are organized as follows. Section 2 discusses the flag image database of flags from 186 different countries, section 3 the feature extraction process, and section 4 the flag retrieval system with recognition accuracy results. Finally, section 5 draws conclusions of this work.

2. Flag Image Database Construction

The flag image database construction is the most important step of the flag recognition system. There are three kinds of flag images: plain, synthetic, and natural-scene. First, plain flag images are clean graphic images as shown in the upper left portion of Figure 1. They are used for the truth displayed to the users, and there are 186 of them representing the various countries.[1] Second, synthetic flag images are those synthetically generated from the plain flags as shown in Figure 1; they are used as references for matching. Finally, natural-scene flag images are obtained from a digital camera, and the flags in these images are often occluded or waving; they are used for testing.

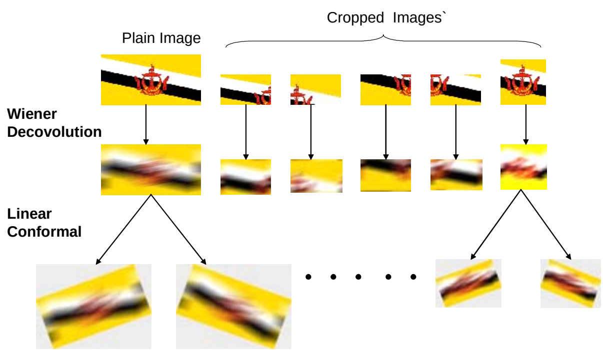

Figure 1. Data generation for Brunei.

To generate the synthetic flag images from the basic plain ones, we first cropped each plain image in five sections, four of them representing ¼ of the image, and the fifth one a central area of 30% of the total image. Six more samples were generated by applying *Wiener Deconvolution* to the 6 first images, the original and the five generated ones. We used Wiener Decovolution to simulate the effect of camera movement or out of focus. From this set of 12 samples we created 24 more samples by using the *Linear Conformal Transform*. The rotation angle for this function was set to 20 and -20 degrees. The idea here is to cover the possibility of cropping rotated flags. Figure 1 depicts an example for Brunei. In all, there were 36 synthetic flags for each plain flag.

3. Color Feature Extraction

A previous study described the main characteristics of flags: colors, symbols, object counting, or color distribution [11]. It is apparent that using only one type of feature is not sufficient to successfully classify flags. For example, in Figure 2 (column 1) color alone differentiates between the flags. In column 2, both flags have the same three colors and what differentiates them is that the top one has symbols. Column 3 is more complicated because both flags have symbols and the same colors, but the number of objects is different with the flag at the top having 6 and the one at the bottom 8. The last example (column 4) cannot be solved by using the first three approaches – the flags contain

[1] The plain flag images come from the World Flag Database, http://www.flags.net/.

the same three colors, do not have symbols, and have the same number of objects. Color distribution, however, makes the discrimination because one flag has a horizontal and the other a vertical arrangement.

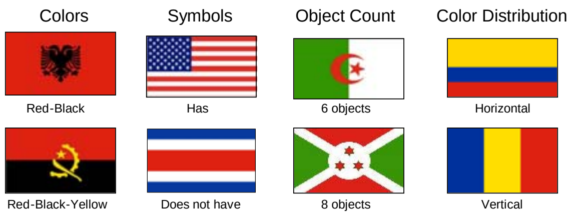

Figure 2. Feature extraction alternatives

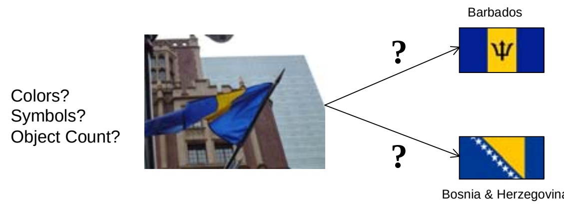

Figure 3. Limitations of feature extraction in real photos.

Because there are many factors present when a photograph is taken, a real picture doesn't look as perfect as in the images above. Movement (the flag is waving), light, focus, distance, and position of the camera make it difficult to recognize features. Most of the actual photos look something like the flag in Figure 3 where it is not possible to determine the number of objects, symbols, colors (besides red and yellow) or color distribution. For this reason the feature vector for this preliminary design of the system considers only colors and percentages of color participation. All colors are clustered into nine colors using HSV color values. Thus, flag images are represented as a vector F_i where i = 1…9 and each element represents the percentage of the color. Due to shadows or bright areas real images of flags are composed not only by their true colors but also by fictitious additional colors. When analyzing color values using an algorithm based on crisp logic, shaded or bright pixels that not correspond to black or white areas are mistakenly identified as those. Identifying flags based on color recognition requires having an adequate algorithm capable of recognizing true colors in these doubtful areas. An attempt to solve this dilemma is by a four-layered approach based on two fuzzy logic algorithms and a neural network (see figure 4). The fuzzy algorithms aim to segment the image in true color areas and shaded or bright areas. Areas found as shaded or bright are considered as doubtful areas. After segmenting the image, a small neural network is trained with the information of the areas found as true colors. This neural network is used to decide the true color of the doubtful pixels.

In the first layer analyzing the HSV values of every pixel by a fuzzy logic process does the segmentation. This fuzzy algorithm uses the color components to decide whether they correspond to black, white, true color or by default doubtful color. If the output is black or white then the HSV values and its color identification (black or white) are stored in a vector. Let V1 be a vector that stores the H_i , S_i , V_i, Id_i of every $pixel_i$ found as true color where Id_j (j = 1..9) is an identification number of nine basic colors; black, white, red, yellow, orange, green, cyan, blue, and magenta. If the output is true color then, using only the H value as input in a second layer, a fuzzy logic algorithm decides its color from a set of 7 basic colors (excluding black and white). In this case also the pixel's HSV values and color identification are store in V1. If the output is doubtful color (by default) then the pixel's coordinates (i, j) are stored in a vector V2. Using the information gathered in V1, a neural network is trained having the H_i , S_i , V_i as input and their corresponding Id_i as output. Once the neural network gets trained it is used to define the true color of the doubtful pixels. Here the input of the neural network are the H_i , S_i , V_i of the (i, j) position stored in vector V2 and the output is the corresponding true color. The following figure depicts the four-layer approach.

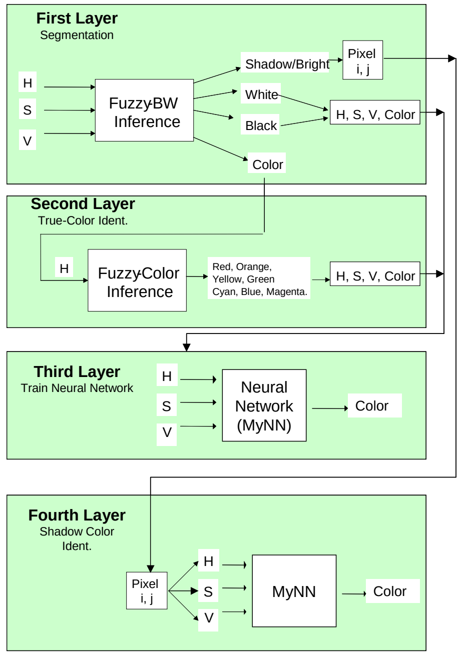

Figure 4. Four-layer Fuzzy-Neural Feature extraction

The next image shows some of the results after replacing the colors of original photos with the true colors found using the four-layer approach or using and algorithm based on crisp logic.

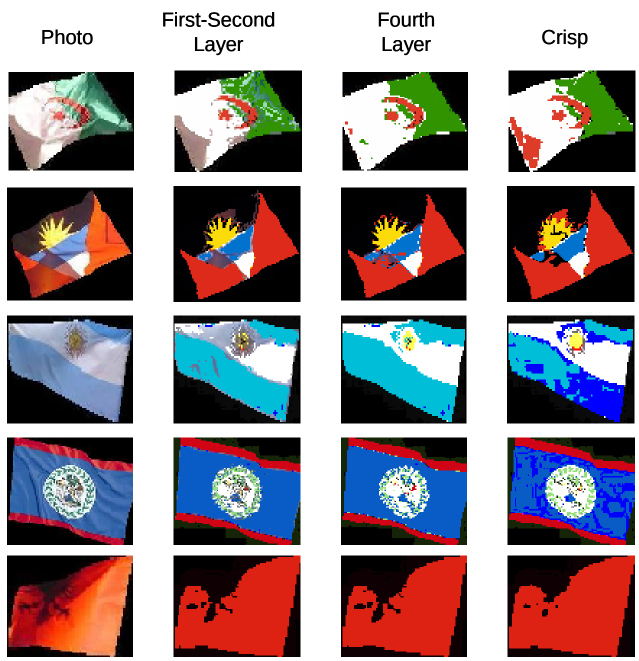

Figure 5. Feature extraction based on Four-layer Fuzzy-Neural and Crisp logic.

4. Flag Image Retrieval and Results

The classification process uses an image retrieval technique involving a similarity measure. As we indicated above, our reference is generated from a set of 186 different classes and 36 samples per class. We tested the system on synthetically generated flag images and on real photos of flags. Let F_i be a plain flag image where the subscript, *i*, indicates *i*th country flag. Let $S_{i,j}$ be synthetic flag image of *i*th country and *j*th synthesized flag. Finally, let Q be an unknown flag photo image taken from a digital camera. Then the inputs to our system are $S_{i,j}$ and Q. The output is a sorted list of F_i's according to $d(S_{i,j}, Q) = \sqrt{\sum_{k=1}^{9}(S_{i,j,k} - Q_k)^2}$

The output of a typical pattern recognition system is $\arg\min_{S_{ij}} \sqrt{\sum_{k=1}^{9}(S_{i,j,k} - Q_k)^2}$, i.e., the top match. Instead, we retrieve all flag images in order of similarity. Figure 6 shows the recognition performance graph where *x* axis is the number of retrieved flags, often called 'hits' and *y* axis is the percentage of correctly retrieved flags within the top *k* hits.

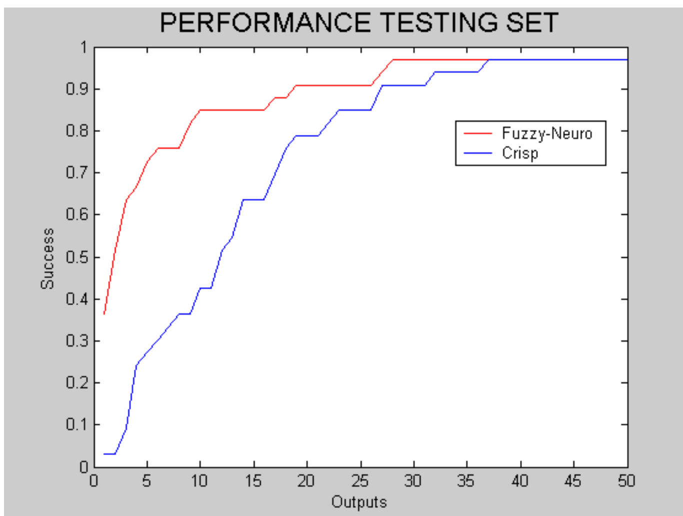

Figure 6. Performance of the Testing Set.

To test the system interactively we developed a graphical user interface in which the user is allowed to crop the image and the system displays the first three matches. Users can also request to see the next three matches, etc. Figure 7 depicts 10 original photos, the cropped area, and the 10 first outputs after the retrieval.

5. Conclusion

We presented an interactive flag recognition system that identifies flags embedded in photos of natural scenes. Since obtaining a large volume of flag images is time-consuming and difficult, we generated a large number of synthetic flag images from plain flag images. Applying a color-based image retrieval technique based on a four-layer Fuzzy-Neural feature extraction and Nearest Neighbor classifier, we reported good flag identification results. The proposed system is an interactive system because of two reasons. First, users are asked to select the region of interest by cropping the perimeter of the flag area. Second, the system does not automatically identify the flag to its respective country but lists the countries based on the color similarity.

References

[1] Arnold W.M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, Ramesh Jain, “Content-Based Image Retrieval at the End of the Early Years”, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol 22, no. 12, December 2000.

[2] Sung-Hyuk Cha, “A Fast Hue-Based Colour Image Indexing Algorithm”, Machine Graphics & Vision, vol 11, no 2/3, pp285-295, 2002

[3] Sung-Hyuk Cha, Srikanth Munirathnam, “Comparing Colour Images using Angular Histogram Measures”, in Proceedings of 5th JCIS, vol II, CVPRIP, p139-142, 2000

[4] G. Pass, R. Zabih, J. Miller, “Comparing Images using Color Coherence”, in ACM International Multimedia Conference, pp65-73, 1998

[5] H. S. Sawhney, J. L. Hafner, “Efficient Color Histogram Indexing”, in International Conference on Image Processing, vol1, pp66-70, 1994

[6] M. J. Swain, D. H. Ballard, “Color Indexing”, International Journal of Computer Vision, vol7, pp11-32, Nov. 1991

[7] G. Nagy, J. Zou, Interactive Visual Pattern Recognition, *Proc. Int. Conf. Pattern Recognition XIV*, vol. 2, pp. 478-481, Quebec City, 2002.

[8] Yang Jie, Gao Jiang, “An Automatic Sign Recognition and Translation System", ACM International Conference Proceeding of the 2001 workshop on Perceptive user interfaces, pp 1-8

[9] Lee Jung-Hum, Jo Kang-Hyum, “Traffic Sign Recognition by Division of Characters and Symbols Regions”, Proceedings of the 7th Korea-Russia International Symposium, KORUS 2003, *pp*324 – 328

[10] Foschi G. Patricia, “Feature Extraction for Image Mining”, *Romberg* , International Workshop on Multimedia Information Systems, p 1-7
.
[11] FireFly, “Guide to Flags of the World”,Published by Firefly Books Ltd. 2003. ISBN 1-55297-813-3

[12] N. Ito, Y. Shimazu, T. Yokoyama, Y. Matushita, “Fuzzy Logic Based Non-Parametric Color Image Segmentation with Optional Block Processing”, pp 119-126, Association of Computing Machinery, 1995 ACM 0-89791-737-5.

[13] Y. Choi, R. Krishnapuran, “Image Enhancement based on Fuzzy Logic”, pp 167-170, Proceedings of the 1195 International Conference on Image Processing, 0-8186-7310-9/95, 1995 IEEE.

[14] Y. Chen, H. Hwang, B. Chen, “Color Image Analysis Using Fuzzy Set Theory”, pp 242- 245, Proceedings of the 1995 International Conference on Image Processing, 0-8186-7310-9/95, 1995 IEEE.

[15] F. Chung, B. Fung, “Fuzzy Color Quantization and its Application to Scene Change Detection”, pp 157-162, MIR’03, November 7, 2003, Berkeley, California, USA.

Figure 7. Samples of first 10 outputs after cropping and classifying real photos.