Reconocimiento en color mediante un solo canal basado en el preprocesado del histograma

Single-channel color pattern recognition by histogram-based preprocessing

M.Corbalán(1), M.S.Millán(2), M.J.Yzuel(3)

1. Dpt. Ingeniería Electrónica, EUPM, Universidad Politécnica de Cataluña, montserrat.corbalan@upc.es, Av.Bases de Manresa 61-73, E-08240 Manresa, España
2. Dpt. Óptica y Optometría, EUOOT, Universidad Politécnica de Cataluña, millan@oo.upc.es, E-08222 Terrassa, Spain
3. Dpt. Física, Facultat de Ciencies, Universidad Autónoma de Barcelona, Maria.Yzuel@uab.es, E-08193 Bellaterra, Spain

RESUMEN:

Se propone la mejora del proceso de reconocimiento en color mediante un preprocesado basado en el análisis del histograma del objeto a reconocer en los tres canales rojo, verde y azul. Como resultado del preprocesado se obtiene una imagen binaria y se realiza una única correlación. El método proporciona mejor discriminación entre objetos con una forma similar pero con diferente distribución del color. Además, el método mejora el proceso de reconocimiento, ya que se reduce el número de canales a procesar y aumenta la capacidad de discriminación.

Palabras clave: Color, procesado de imagen digital, correlación óptica, filtro sólo de fase.

ABSTRACT:

Preprocessing based on the histogram analysis of the target in the three red, green and blue channels is proposed to improve color pattern recognition. As a result of the preprocessing, single binary images are synthesized and monochannel correlation can be carried out. This approach provides better discrimination between objects with a similar shape but a different color distribution. The method improves color recognition processing by reducing the number of channels and increasing discrimination capability.

Key words: Color, digital image processing, optical correlator, phase-only filters

REFERENCES.

1. Introduction

The introduction of color increases the information content in pattern recognition and therefore discrimination can be considerably improved. A color camera usually captures three color channels: R (red), G (green) and B (blue). RGB multichannel decomposition has been associated with matched filtering and optical correlation techniques in black and white. Badiqué et al.\(^2\) converted a three-dimensional (R, G, B) vector into a two-dimensional vector by projecting the vector onto a color plane. By obtaining an optimal generalized color plane, better discrimination is achieved. Pei and Hsu\(^3\) have proposed a complex logarithm representation of the color object information, based on the structure of the human retina to obtain a recognition system that is invariant to both rotations and scale changes of the color objects. Several transformations of the RGB channels to another set of channels have been proposed.\(^4\) In the case of Ref. 5, recognition of polychromatic objects was made possible with a subsystem of only two channels (T and D) based on models of human vision. Pérez et al.\(^8\) proposed the selection of \(n\) quasi-monochromatic channels based on the study of spectral reflectance of the color of the objects. Swain and Ballard\(^9\) developed a color-indexing algorithm to identify an object by comparing its colors with the colors of each object in a database. The total areas covered by each color are computed and compared by histogramming the images and intersecting the histograms. Swain's method identifies objects entirely on the basis of color, but it ignores information about their shape. In Ref. 10, 3D images are created in which the third coordinates

---

[72x759]

carries the color information. 3D correlation is applied and color pattern recognition is studied. Filtering operations on the three dimensional spectrum of color images are studied in Ref. 11. In Ref. 12 the three RGB channels of a color image are encoded in a single monochromatic image that encodes, in complex values, the color content of the image.

In this paper we propose a method based on preprocessing both the target and the scene. In a deterministic recognition process the target is generally known. First, we generate a scene with the target on a black background. Preprocessing is based on the histogram analysis of the target in the three channels (RGB). A single binary channel is obtained from the information of the three channels for both the target and the scene. We thus intend to improve the discrimination between similarly shaped objects, but with different colors. We will see that the method can be particularly useful in analyzing complex scenes consisting of multicolored objects on non-uniform color backgrounds.

2. Preprocessing procedure

We consider multicolor objects. The histogram of the target in the three R, G, B channels are carried out. For this operation, the target is in a scene with black background. In each R, G, B channel we remove the gray level equal to zero (background) from the histogram. \( V_i^g \) is the number of pixels of the gray level \( g \) in the \( i \) channel. The total number of pixels or area of the target in the \( i \) channel is \( V_i^T \), where
\[
V_i^T = \sum_{g=0}^{255} V_i^g , \tag{1}
\]
Preprocessing is carried out channel by channel. In each channel \( i \), the values \( V_i^g \) are placed in decreasing order and we generate a numerical sequence \( \{ V_i^g \}_h \). Thus, the first number of the sequence corresponds to the mode of the histogram, i.e., the gray level that corresponds to \( \{ V_i^g \}_h \) is the mode. From this first value \( \{ V_i^g \}_1 \) we add the following values \( \{ V_i^g \}_2, \{ V_i^g \}_3 \ldots \) of the numerical sequence until we exceed a given percentage of the total target area in channel \( i \). We introduce a threshold \( U \) for the target area. We select the lowest number of gray levels \( N_i \) so that the condition:
\[
\sum_{j=1}^{N_i} \{ V_i^g \}_j \geq U \cdot V_i^T \tag{2}
\]
is satisfied in each channel \( i \). Let us represent the set of gray levels selected in each channel by \( \{ g \}_{N_i} \).

These \( N_i \) gray levels correspond to the area in each channel R, G, B, defined by the selective threshold \( U \). The outline of this operation for the target \( t(x,y) \) is shown in Figure 1 where \( t(x,y) \) is the target component of the \( i \) channel.

Binarization is then applied pixel by pixel in each channel and a new binary component \( F_i \) is obtained:
\[
F_i(x,y) = \begin{cases} 
1 & \text{if } \{ g(x,y) \} = \{ g \}_{N_i} , \\
0 & \text{otherwise} 
\end{cases} \tag{3}
\]
where \( \{ g(x,y) \} \) is the gray level of pixel \( (x,y) \) in the \( i \)th image component.

A single binary image \( \hat{F} \) is obtained by applying the OR operation to the resulting \( F_i \)
\[
\hat{F} = \{ F_a \} \text{OR} \{ F_b \} \text{ OR} \{ F_c \} \tag{4}
\]
The operation OR was selected to obtain a large area of the target. The outline of the operation of Eqs. (3) and (4) are shown in Figure 2a. From the new binary target \( \hat{T}(x,y) \) we generate a matched filter for a recognition process based on correlation.

The scene may contain several multicolor objects and a color background. Once the scene has been captured in the RGB components, Eqs. (2) and (3) are used taking into account the same selected gray levels \( \{ g \}_{N_i} \) that were extracted from the target (Fig. 1) and is obtained a binary scene by the preprocessing given in Eqs. (3) and (4). Figure 2b shows this procedure. Therefore, the recognition process is carried out on a single channel that works with binary scenes. This represents an advantage both for experimental application of the procedure and for data compression. As the scene is binary it can be displayed in an optical real-time correlator that uses a ferroelectric spatial light modulator$^{13,14}$.

The advantage is that these devices are binary and can reach speeds of high transmission.

![Fig. 1.- Outline for obtaining the set of gray levels \( \{ g \}_{N_i} \) in each \( i \) channel from the \( i \) target (Eq. 2), \( h_i \) is the histogram of the target in the \( i \) channel.](image-url)

Fig. 2.- Preprocessing applied to the target (a) and to the scene (b) by application of Eqs. (3) and (4).

3. Experimental results

To show the potential advantages of our approach, we choose a scene with very similar objects on a non-overlapping color background (Fig. 3a). The color scene contains two birds with the same shape but different color distribution on a background non-uniform and this makes conventional recognition more difficult. We call this type of scenes, scenes with continuous histograms. The target is Object 1.

Fig. 3.- Color scenes: a) continuous histogram, b) discrete histogram. In each case the target is object 1.

Figures 4a, 4b, and 4c show the histograms of the target on a black background in the three R, G and B channels. In the upper right corners the target components are also show. One can see that the histograms that are obtained are continuous. Using each histogram we build up the numerical sequence \( V_i^g \) and we obtain the set \( \{ g \}_{N_i} \) of gray levels by applying Eq. 2. In our experiment we will apply \( U=0.25 \) in each channel. The selected gray level \( \{ g \}_{N_i} \) for the red channel is 11, for the green channel it is 14, and for the blue channel it is 4. After applying Eqs. (3) and (4) to the target we obtain a single binary target shown in Fig. 4d. A phase only filter is matched to this binary target.

Fig. 4.- Histograms and images of the target in: a) R, b) G and c) B channels. d) Preprocessed target.
Using the same set $\{g\}_N$ obtained from the target, Eqs. (3) and (4) are applied to the scene (Fig. 3a). A single binary image is thus obtained and it is shown in Fig. 5a. Using the scene of Fig. 5a and the target of Fig. 4d, the correlation plane obtained is shown in Fig. 5b. We obtain recognition of Object 1 and discrimination of Object 2. The correlation peaks are very narrow and the correlation plane shows high signal to noise ratio. On the other hand, if we had used three POFs matched to the respective RGB target components and we had performed RGB multichannel correlation with the scene of Fig. 6, we would have obtained the correlation planes shown in Fig. 7. It can be seen that the correlation peaks obtained for Object 2 in all the three channels reach high values and make the recognition difficult due to the resemblance of the two objects in the three R, G and B channels (Fig. 6). Recognition is clearly improved when preprocessing is applied (Fig. 5b).

In order to test the proposed preprocessing with other type of objects and scenes we have applied to objects with well defined areas of uniform color. We want to recognize a color object and to discriminate against another object whose color areas have the same shape but different color although belonging to the same color range (Fig. 3b). This election is considered as a difficult problem in multichannel RGB recognition process based on correlation. Moreover, the scene contains the objects are on a non-overlapping color background. This fact turns out more difficult to recognize the objects in the scene. If we had used three POFs matched to the RGB target components and we had performed RGB multichannel correlation with the scene of Fig. 3b, the maximum values of correlation peak that are obtained for Obj. 2 if we normalized with respect to the maximum value of autocorrelation are 0.78, 0.64, 1.07 for channels R, G, and B, respectively.

The histograms of the target (object 1 in Figure 3b) on a black background are processed and are shown in Figures 8a, 8b, and 8c. It can be seen that the histograms for this type of scenes are discrete. Using each histogram we obtain the set $\{g\}_N$ of gray levels by applying Eq. 2. Again we use a $U$ equal to 0.25 in each channel. The selected gray levels $\{g\}_N$ for the red channel are 111 and 184, for the green channel they are 100 and 136, and for the blue channel they are 50 and 70. After applying Eqs. (3) and (4) to the target we obtain a single binary target shown in Fig. 8d. A phase only filter is matched to this binary target. In the same way, by using the same set of gray levels obtained from the target, Eqs. (3) and (4) are applied to the scene and a single binary image is then obtained (Fig. 9a). Using the scene of Fig. 9a and the target of Fig. 8d, the correlation plane obtained is shown in Fig. 9b. We obtain recognition of Object 1 and discrimination of Object 2. Therefore, recognition is clearly improved.
when preprocessing is applied. This preprocessing method was applied to different types of scenes and backgrounds and an improvement in recognition was obtained in all cases.

Conclusions

We propose a method for improving color pattern recognition. A preprocessing based on the histogram analysis of the target on black background in the three R, G, and B channels is proposed. It is carried out channel by channel and a single channel that moreover is binary is obtained.

To show the improvement that the method produces in the recognition process two tests color scenes are used. The two scenes have objects with a similar shape but a different color distribution on color backgrounds. They have been selected so that they have either continuous or discrete histograms to show that the method improves very clearly the recognition in both cases. We have shown that the method can be particularly useful in analyzing complex scenes that with a conventional recognition process based on correlation is not possible.

Besides, we may conclude that the proposed method is more compact because we reduce the number of channels from three to one, and also that it is more discriminative and easier to apply experimentally because the images that are obtained are binary.

Acknowledgements

The authors thank the Ministerio de Ciencia y Tecnologia MCYT (projects DPI2000-0991 and BFM2000-0036-C02-01), and to the Generalitat de Catalunya (Xarxa Temàtica de Processament Optic d’Imagenes 2003XT 00081) for financial support.