Color image segmentation by supervised pixel classification in a color texture feature space. Application to soccer image segmentation.

Nicolas Vandenbroucke, Ludovic Macaire and Jack-Gérard Postaire
Laboratoire d'Automatique I^{3}D - Université des Sciences et Technologies de Lille
Cité Scientifique - Bâtiment P2 - 59655 Villeneuve d'Ascq - FRANCE
nv@i3d.univ-lille1.fr

Abstract

In this paper, we describe a new approach to color image segmentation which is considered as a supervised pixel classification problem. The pixel classification algorithm analyses the color texture features, that is to say the texture features which are computed by tacking into account the color components of the neighbor pixels. We determine the most discriminating color texture features among a multidimensional set of color texture features by means of an iterative feature selection procedure associated to an information criterion. We successfully apply our new approach to soccer image segmentation.

1. Introduction

Several authors have proposed soccer game image analysis systems designed to obtain accurate information about the soccer game and about player behavior [1, 4]. In the framework of soccer player tracking, we propose an original color image segmentation method which consists of the classification of *player pixels*, that is to say, the pixels which represent the players. We suppose that the team of each player is identified by the color of its soccer suit. Consequently, the player pixels which represent the players with the same soccer suit constitute one class of player pixels. We can consider five different classes namely, the players of the two teams, the two goalkeepers and the referees.

1.1. The previous approach

In our last two papers, each player pixel was classified according to the mean of each color component of its neighbor player pixels [5, 6]. Although the color components of a pixel are usually the trichromatic components R (Red), G (Green) and B (Blue), many other color components exist, and we considered those which come from well known color representation systems [3]. In this former approach, we determined the three most discriminating color components among this set, which constitute a *hybrid color space*. The color pixels were classified in this hybrid color space. Because the soccer player's suits are textured, however, the classification of the player pixels, which was based only on the mean of each color component, was imperfect.

1.2. The new approach

In this paper, we propose to take into account *color texture features* in order to improve the classification. For each player pixel, we compute well known texture features such as the mean, the variance or the skewness as a function of the color components of its neighbors. We use a set of N_t texture features which can be processed as a function of a set of N_c color components. We consider that each player pixel is represented by a point whose coordinates are the $N_c \times N_t$ color texture features. Such a classification algorithm requires a very large computation time since the dimension of the color texture feature space is $N_c \times N_t$. The computational burden is reduced by projecting the color texture feature space onto a lower dimensional space thanks to the selection of the most discriminating color texture features. More over, in our new approach, we propose a new information criterion which takes into account the correlation between the color texture features.

1.3. Scheme

In the second section of this paper, we present a supervised learning scheme which consists of determining the most discriminating color texture features thanks to an interactive selection of *player windows*. In the third section, we describe the pixel classification algorithm which assigns each player pixel to a specific class according to the values of the color texture features of its neighborhood. In the last section, we apply our previous and our new approaches to soccer images and we compare the results. We show that

0-7695-0750-6/00 $10.00 © 2000 IEEE

the new approach improves the reliability of the classification of the player pixels.

2. The supervised learning scheme

In order to determine the color texture feature space, we require samples which are representative of the classes of pixels. At the beginning of a supervised learning scheme, we interactively select several *player windows* from a set of *presegmented learning images*. In these images, the ground is withdrawn thanks to an adapted Ohlander's algorithm in order to extract the player pixels [2]. The player windows contain player pixels which represent players in different situations (running, pushing the ball, dribbling...) and in different positions (facing the camera, back to the camera...). The images of figure 1 represent a set of selected player windows.

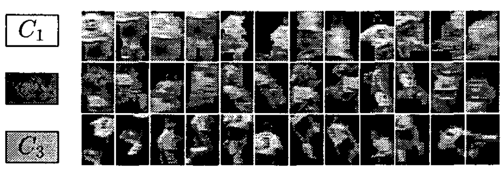

Figure 1. Set of selected player windows

The size of these player windows depends on the mean size of the players in the image.

Let us denote by C_j a class of player pixels ($j = 1, ..., N$ where N is the number of classes). For each class C_j, we interactively select the same number N_ω, of player windows $\omega_{i,j}$, where i indicates the serial number of the player window ($i = 1, ..., N_\omega$). The player windows are used to determine the color texture feature space which yields the best separation between the classes.

2.1. Color texture features extraction

A player pixel neighborhood can be characterized by color texture feature values which are computed by tacking into account the color components of the neighbor player pixels.

Here, we use a non exhaustive list of texture features. The *mean* of the pixel values in a neighborhood, the *median* and the *mode* evaluate the central value of this neighborhood. The variability of the pixel values around a central value is estimated by the *variance* or its square root, the *standard deviation*. The *skewness* estimates the degree of asymmetry of the pixel values around a central value. The variance and the skewness can be evaluated around the mean, the median or the mode. Let N_t, be the number of available texture features which are computed with one color component.

The pixels of a color image are usually digitized with the (R, G, B) color representation system. Nevertheless, the R, G and B color components are not always adapted to a specific problem of color image segmentation. In digital color imaging, many other color representation systems exist [3]. For our study, we use the more classical color representation systems which are derived from the (R, G, B) system, such as: the primaries system CIE (X, Y, Z), the normalized systems (R_n, G_n, B_n) and CIE (X_n, Y_n, Z_n), the perceptually uniform systems CIE (L^*, a^*, b^*) and CIE (L^*, u^*, v^*), the Faugeras's system $(A, C1, C2)$, the Ballard's system (wb, rg, by), the television systems (Y', I', Q') and (Y', U', V'), different perceptual systems (L, C, H) and the Ohta's system $(I1, I2, I3)$. There is a total of N_c available color components which constitute the different above mentioned color representation systems.

By taking into account the $N_f = N_c \times N_t$ available color texture features, we define an N_f-dimensional color texture feature space. The large dimension of this space inevitably generates redundancy. Furthermore, the classification algorithm is very costly in computation time. So, our goal is to look for the best subset of color texture features for discriminating the different classes of pixels.

2.2. Color texture feature space selection

In order to reduce the dimension of the color texture feature space, we use an iterative feature selection procedure. At each step k of this procedure, we consider several candidate color texture feature spaces for which we compute their discriminating power thanks to an information criterion J. At the beginning of this procedure ($k = 1$), we consider the N_f mono-dimensional candidate spaces defined by each of the available color texture features. The candidate space which maximizes J is the best one for discriminating the N classes. We select this space as the first one and we associate it, in the second step of the procedure ($k = 2$), to each of the $(N_f - 1)$ remaining candidate color texture features in order to constitute $(N_f - 1)$ bi-dimensional candidate spaces. We consider that the bi-dimensional space which maximizes J is the best plane for discriminating the classes... The procedure is iterated until stabilization of the value of J. Let k_0, the rank of the iteration process which corresponds to the beginning of the stabilization of J. k_0 is the dimension D of the color texture feature space. This classical multiple discriminant analysis method does not yield the optimal solution but a satisfying one which is less computation time consuming.

The evaluation of the discriminating power supposes that the more the classes are well separated and compact in the candidate color texture feature space, the higher the discriminating power of the selected features is. That leads us to choose measures of classes separability and compactness as measures of the discriminating power.

At each step k of the procedure and for each of the $(N_f - k + 1)$ k-dimensional candidate color texture feature spaces, we define, for each player window $\omega_{i,j}$ associated to the class C_j, a *feature vector* $X_{i,j} = [x^1_{i,j}, ..., x^k_{i,j}]^{\mathrm{T}}$ where $x^k_{i,j}$ is the k^{th} color texture feature.

The measure of compactness of each class C_j is defined by the intra-class dispersion matrix Σ_C:

$$\Sigma_C = \frac{1}{N_\omega \times N} \times \sum_{j=1}^{N} \sum_{i=1}^{N_\omega} (X_{i,j} - M_j)(X_{i,j} - M_j)^{\mathrm{T}}$$

where $M_j = [m^1_j, ..., m^k_j]^{\mathrm{T}}$ is the *mean vector* of the k color texture features of the class C_j.

The measure of the class separability is defined by the inter-class dispersion matrix Σ_S:

$$\Sigma_S = \frac{1}{N} \times \sum_{j=1}^{N} (M_j - M)(M_j - M)^{\mathrm{T}}$$

where $M = [m^1, ..., m^k]^{\mathrm{T}}$ is the mean vector of the k color texture features for all the classes.

The most discriminating set of color texture features maximizes the information criterion:

$$J = \mathrm{trace}\Big((\Sigma_C + \Sigma_S)^{-1}\Sigma_S\Big).$$

In order to only select features which are not correlated, we measure, at each step $k \geq 2$ of the procedure, the correlation between the candidate color texture feature and each of the $k - 1$ other color texture features constituting the considered space. An *observation vector*, denoted $X^f = [x^f_{1,1}, ..., x^f_{i,j}, ..., x^f_{N_\omega,N}]$ is associated to each color texture feature. The correlation $\mathrm{cor}(X^f, X^{f'})$ between two features is defined by:

$$\mathrm{cor}(X^f, X^{f'}) = \frac{\mathrm{cov}(X^f, X^{f'})}{\sigma^f \times \sigma^{f'}}$$

where $\mathrm{cov}(X^f, X^{f'})$ is the covariance between the two features:

$$\mathrm{cov}(X^f, X^{f'}) = \sum_{j=1}^{N} \sum_{i=1}^{N_\omega} \frac{(x^f_{i,j} - m^f) \times (x^{f'}_{i,j} - m^{f'})}{N \times N_\omega}$$

and σ^f is the standard deviation of the candidate color texture feature for all the classes and m^f is the mean of the candidate color texture feature for all the classes.

The correlation value ranges between 0 and 1. The closer is the correlation to 1, the more correlated the two features are. If one of the computed correlation is higher than the value 0.75, the candidate color texture feature is rejected.

Thanks to this iterative procedure, we select the set of the most discriminating color texture features among the N_f available ones.

3. Player pixel classification

In order to classify a player pixel P, we compute its color texture feature vector $X_P = [x^1_P, ..., x^D_P]^{\mathrm{T}}$ in the above determined D-dimensional color texture feature space. For that, we consider the set of the neighbor player pixels falling into a neighborhood of P (see figure 2). This neighborhood is defined by a *neighborhood window* which is centered on the player pixel P and whose the size is equal to the player window size.

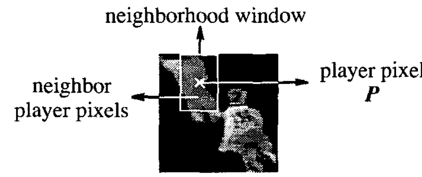

Figure 2. Player pixel neighborhood

For each class C_j, we evaluate the euclidean distance D^P_j between X_P and the mean vector $M_j = [m^1_j, ..., m^D_j]^{\mathrm{T}}$ of the class C_j in the D-dimensional color texture feature space. A minimum distance decision rule is used to assign P to the class C_j for which D^P_j is minimum.

4. Experimental results

In order to illustrate our method, we choose a test soccer image (see figure 3) which contains three different classes of player pixels. In this image, the ground is withdrawn. For the supervised learning scheme, we also select a set of different player windows (see figure 1).

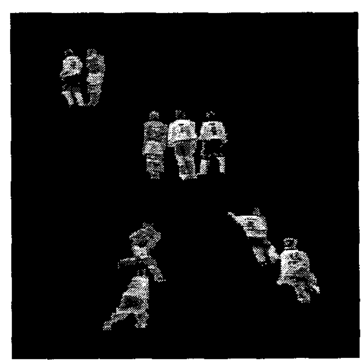

Figure 3. Test image

We apply our previous approach to the image of figure 3 (see figure 4). In this approach, we only use the mean feature and the information criterion which is presented in section 2.2. By reading the confusion matrix (see table 1),

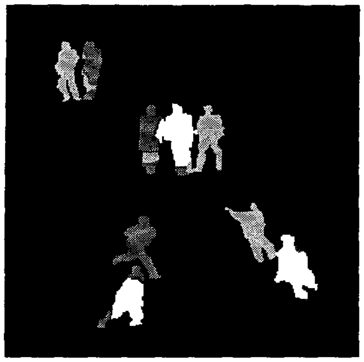

Figure 4. Previous approach results

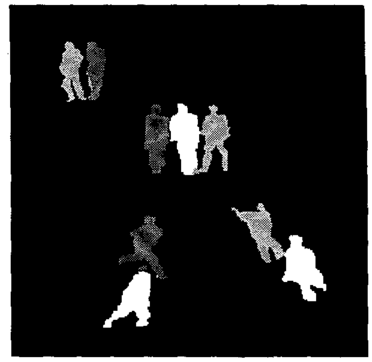

Figure 5. New approach results

we conclude that many player pixels which belong to the classes C_1 and C_2 are miss classified and that the error rate is higher than 8 %.

	C_1	C_2	C_3
C_1	87.94 %	10.52 %	1.54 %
C_2	3.93 %	87.41 %	8.66 %
C_3	0.83 %	0.06 %	99.11 %

Table 1. Previous approach confusion matrix

We apply our new approach in which we use several texture features. The determined color texture feature space is a 4-dimensional one and is constituted by the mean value of the a^* color component of the system (L^*, a^*, b^*), the mean value of the chroma color component of the system (Y', U', V'), the standard deviation around the median associated to a saturation color component and the skewness around the mean associated to the L^* color component of the system (L^*, a^*, b^*). The image of figure 5 contains the player pixels which are classified in this space by our classification algorithm. The player pixels are correctly assigned with an error rate lower than 2 %. By reading the confusion matrix (see table 2), we conclude that the player pixels which belong to all the classes are well classified.

	C_1	C_2	C_3
C_1	97.51 %	1.40 %	1.09 %
C_2	0.12 %	98.67 %	1.21 %
C_3	0.12 %	0.00 %	99.88 %

Table 2. New approach confusion matrix

4.1. Conclusion

The new color image segmentation approach proposed here shows the contribution of color texture features to supervised pixel classification. The determined color texture feature space is specific to an application and depends on the chosen texture features and the criterion used. We are presently studying different information criteria and other color texture features in order to determine the ones which yield the best classification results. Finally, we want to extend our approach to the partitioning of color images in homogeneous regions.

References

[1] K. Matsui, M. Iwase, M. Agata, T. Tanaka, and N. Ohnishi. Soccer image sequence computed by a virtual camera. In *IEEE International Conference on Computer Vision and Pattern Recognition*, pages 860–865, Santa Barbara, 1998.

[2] R. Ohlander, K. Price, and D. R. Reddy. Picture segmentation using a recursive region splitting method. *Computer Graphics and Image Processing*, 8:313–333, 1978.

[3] G. Sharma and H. J. Trussell. Digital color imaging. *IEEE Transactions on Image Processing*, 6(7):901–932, 1997.

[4] S. Sudo and S. Ozawa. Scene analysis of soccer game. In *IAPR International Conference on Quality Control by Artificial Vision*, pages 119–123, Trois-Rivières, 1999.

[5] N. Vandenbroucke, L. Macaire, and J.-G. Postaire. Color pixels classification in an hybrid color space. In *IEEE International Conference on Image Processing*, volume 1, pages 176–180, Chicago, 1998.

[6] N. Vandenbroucke, L. Macaire, C. Vieren, and J.-G. Postaire. Contribution of a color classification to soccer players tracking with snakes. In *IEEE International Conference on System, Man, and Cybernetics*, volume 4, pages 3660–3665, Orlando, 1997.