Multiresolution Color Image Analysis

Hubert Konik, Alain Tremeau and Bernard Laget
Equipe Ingénierie de la Vision, Laboratoire Traitement du Signal et Instrumentation
Site GIAT Industries, Saint-Etienne Cedex , France

Abstract

Because of the color space size, a color image contains too much information. Consequently, before analyzing it, its information has to be reduced without a loss of relevant data. In order to reach that, the multiresolution process is well adapted. Indeed, it enables a spatiocolor approach considering color and spatial distribution at the same time. More precisely, the construction of a color pyramid is presented. Finally, its utilisation during a segmentation step is discussed.

Introduction

The final aim of our research is an automatic vision system for visual aspects quantification of any surfaces. Generally speaking, color information is not enough took into account in computer vision. Yet, without considering this information, many problems are inevitably unsolvable because, among the criteria that enable one to judge the homogeneity of an area, color information is particularly the main parameter in visual evaluation process[1]. Let consider for example textile surfaces, that can be decomposed both as a color image and a luminance image. This second one has to be then divided into an image of structure and an image of texture[2]. Let now present an appropriate multiresolution tool to analyze color images.

Interest of a Spatiocolor Approach

Color image analysis is still limited because of a very important amount of data[3], considering that the RGB color space is three dimensional. To minimize this information, we can either work on the color space either on the image itself. Thus, the most general case deals with images coded with 16 $\overline{\mathrm{M}}$ of RGB colors! In order to reduce this information, different kinds of process are used, as the quantization process for example. It first selects some representative colors from the gamut of the image and then assigns each pixel to one of them.

Nevertheless, most of these processes take only into account the information contained in the color space, without considering the spatial distribution of color data in the image. Consequently, the chosen representative colors can be in poor agreement with the real image. Other methods have been tested, among them the multiresolution approach seems to be one of the most relevant[4,5]. In fact, it simulates the human vision system sharpening each studied area until it is homogeneous enough. Our approach consists in using a pyramidal process to implement the multiresolution approach. By definition, a pyramidal tool contains the same image at different resolution levels decreasing from one to another. It allows us to compute, thanks to the spatiocolor information, a new set of representative colors that is more relevant as regards to the original image. Actually, the main interest of this pyramidal tool is to combine both spatial sampling and frequency sampling. Let now present the construction of a color pyramid.

Color Pyramid Construction

The use of multiresolution or pyramidal techniques in computer vision has been studied by many authors since Tanimoto and Pavlidis[6]. Such a tool has been initially built using only gray values. In fact, a pyramid is a hierarchy of fine to coarse resolution versions of an image, where the resolution decreases usually twofold between consecutive levels. Let be $2^n \times 2^n$ the original image , the levels are then of sizes :

$$2^{n-1}\times 2^{n-1},\ 2^{n-2}\times 2^{n-2},\ \dots,\ 2\times 2,\ 1\times 1. \tag{1}$$

As a result, an entire pyramid contains 4^{n+1}-1 elements. Generally, the values of the current level are computed by convolving the gray values at the previous level with a K×K kernel and by sampling them at half the current spatial frequency.

Thus, the value of each element (x,y) at level h is computed as follows :

$$f_h(x,y) = \sum_{i=0}^{K-1}\sum_{j=0}^{K-1} w(i,j).f_{h-1}(2x+i-z, 2y+j-z), \tag{2}$$

where z is $\left\lfloor \frac{K-1}{2} \right\rfloor$.

By definition, the K^2 pixels (2×+i-z,2y+j-z) at level h-1 are the sons of (x,y) at level h. When (x,y) is used to compute an element at level h+1, each of them is one of his father.

Different forms of the generating kernel w(i,j) have been studied by Burt[7]. The gaussian one tends to preserve the shape of the objects and the contrast of the image[1]. It is defined as follows :

$$[w(i,j)] = \begin{Bmatrix} 0.0169 & 0.0481 & 0.0481 & 0.0169 \\ 0.0481 & 0.1369 & 0.1369 & 0.0481 \\ 0.0481 & 0.1369 & 0.1369 & 0.0481 \\ 0.0169 & 0.0481 & 0.0481 & 0.0169 \end{Bmatrix} \tag{3}$$

Such a kernel defines the overlapping gaussian pyramid, presented in Figure 1.

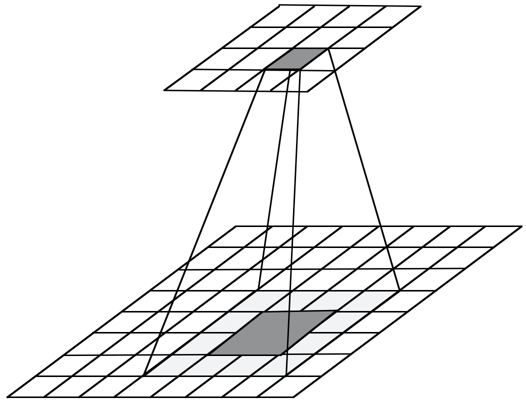

Figure 1. Construction of the levels in an overlapping pyramid

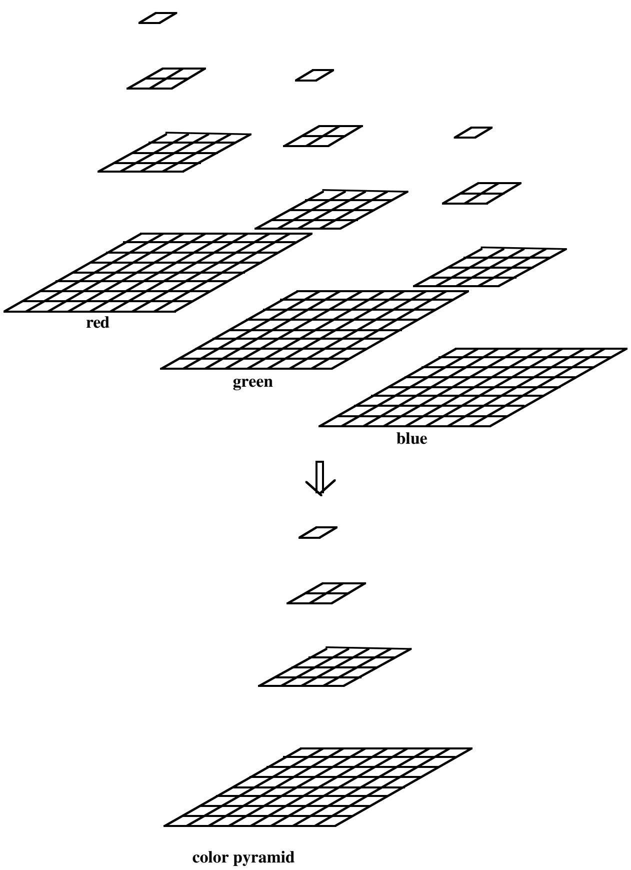

Figure 2. The RGB color gaussian pyramid

Considering the three dimensional color space, a solution is to define one gaussian pyramid for each component, that can be the RGB or the $L^*a^*b^*$ ones. The color pyramid is then obtained by combining the three pyramids. As an example, the RGB color gaussian pyramid is illustrated in fig. 2. The choice of RGB components is justified in noting that an RGB element falls every time into the RGB space. Moreover, the RGB space is discrete and can be coded with integer values contrary to the La^*b^* one.

Such a tool simulates the human vision in its attention focusing, through an individual and a contextual analysis of the regions[8]. Moreover, this structure is used by defining links between pixels at adjacent levels. In fact, pixels are classified by linking them at successive levels according to a similarity criterion based on color values. Let present the obtained pyramid on the test image “mandrill”. Each color component RGB is coded on 8 bits. Figure 3, 4, 5 and 6 shows respectively the levels 0, 1, 2 and 3. All levels are presented with the same resolution as the original one. This example shows that color information is well propagated through the pyramid. In fact, the coarser the resolution is, the more homogeneous the relevant areas appear. The filtering smoothes the texture in accordance with human visual analysis, especially in the coat. Its associated color is then still characteristic at level 3.

Figure 3. Test image “mandrill”

Figure 4. Level 1

Figure 5. Level 2

Figure 6. Level 3

Let now present some possible uses of color pyramid for the problem of color image segmentation.

Using the Color Pyramid for Image Segmentation

Generally speaking, image segmentation is the process that partitions an image into some meaningful regions. In our case, the obtained regions must be homogeneous in some color sense, using the following criterion. Segmentation techniques have to take into account local informations that are important in the human visual process. On this subject, the color pyramidal tool is attractive because the lower resolutions provide a global view of the image, while the higher resolutions provide the details.

First of all, let $f_h^l(x,y)$ represents the color value of the pixel (x,y) on the l-axis at the level h of the pyramid.

$$f_h(x,y) = (f_h^1(x,y),\ f_h^2(x,y),\ f_h^3(x,y)) \qquad (4)$$

defines now the color vector associated to this pixel. Let precise that we can use the RGB or the $L^*a^*b^*$ color spaces in our process.

Let consider (x,y) as the current color element at level h+1. The problem is to decide if it is closer enough to its sons in order to be a good root for the region.

We can then compute the standard deviation between (x,y) at level h and its sons :

$$\sigma_h^l(x,y) = \sqrt{\frac{1}{K^2}\sum_{i=0}^{K-1}\sum_{j=0}^{K-1}\left(f_h^l - f_{h-1}^l(2x+i-z, 2y+j-z) \right)^2} \qquad (l = 1,2,3) \qquad (5)$$

Finally, in order to define a color homogeneity criterion, we compute the following dispersion :

$$d_h(x,y) = \frac{1}{3}\cdot\sum_{l=1}^{3}\sigma_h^l(x,y) \qquad (6)$$

The more homogeneous a set of colors is, the more sensitive to the color contrast this distance is. On the contrary, the more inhomogeneous a set of color is, the less sensitive to the color contrast this distance is. Thus, in order to stress the inhomogeneity of an area, we use the standard deviation criterion previously defined.

Specifically, we allow nodes to refuse to link to any of their parents if the father's value is some number m times the color dispersion[9] and we consider the area is homogeneous.

To illustrate the interest of this measure of spatiocolor homogeneity of the current area, let consider its representation in between 0 and 255 at levels 1, 2 and 3 (see Figures 7, 8 and 9). When the area is inhomogeneous, it appears greater. As the iterations progress, the tendency to merge neighbouring spatiocolor areas exhibits the robustness of this construction.

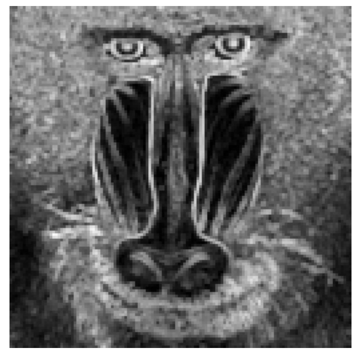

Figure 7. Color dispersion image at level 1

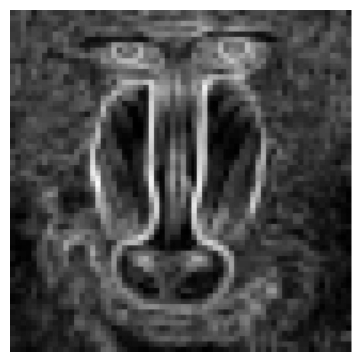

Figure 8. Color dispersion image at level 2

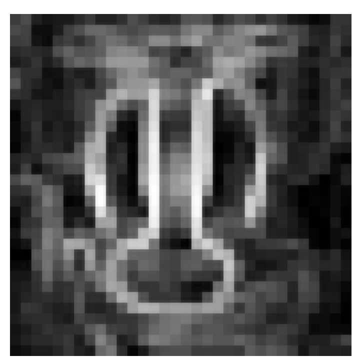

Figure 9. Color dispersion image at level 3

Other criteria can be used, such as an "interest measure"[10] based on a local comparison of the contents of cells on successive levels of the pyramid or on links quality taking into account euclidean distance between elements[11].

Finally, the color image segmentation process is realized as follows. Each homogeneous area is represented by a root at an optimal level in the color pyramid, from which a top-down process can be done until the full resolution image, keeping at each level the closest sons. Then, a post refine process is applied to each inhomogeneous area. Nevertheless, in this case, the aggregation is spatially limited to the areas that are not yet segmented, according to color constraints less restrictive.

Conclusion

The gaussian pyramid improves color treatments in computer vision. In fact, different resolutions of an image compress its relevant information. Some possible uses of this tool for segmentation are discussed. Actually, we work on color fractional pyramid[12] to increase the number of levels. In fact, the number of cells in the next higher level is 1/4 of that in the lower one and some applications have shown that this growth rate may be too fast. Moreover, the conventional pyramidal structure is sometimes too rigid and limited with its inherent limitations (especially elongated regions[2]). We then have introduced a new approach for luminance images : using multiple localized pyramids in an image, called "local-base pyramids". It simulates the human vision in its attention focusing, through an individual and a contextual analysis of the regions. We then work actually on the way to use this new concept in color image analysis.

References

1. R. Gershon, "Aspects of perception and computation in color vision," *Computer Vision Graphics and Image Processing*, vol. **32**, pp. 244-277, 1985.
2. H. Konik, "Contribution de l'approche pyramidale à la segmentation des images texturées", *PhD thesis,* Université de Saint-Etienne, 1994.
3. A. Tremeau, "Contribution des modèles de la perception visuelle à l'analyse d'image couleur", *PhD thesis,* Université de Saint-Etienne, 1993.
4. M. D. Levine, "Region analysis using a pyramid data structure", *In Structured Computer Vision*, Tanimoto S. & Klinger A. Eds, New-York, pp. 57-100, 1980.
5. J. Liu et al., "Multiresolution color image segmentation", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. **16** (7), pp. 689-700, 1994.
6. S. Tanimoto et al., "A hierarchical data structure for picture processing", *Computer Graphics and Image Processing*, vol. **4**, pp. 104-119, 1975.
7. P. J. Burt, "Fast filter transforms for image processing", *Computer Graphics and Image Processing*, vol. **16**, pp. 20-51, 1981.
8. A. B. Watson, "Ideal shrinking and expansion of discrete sequences", *NASA Technical Memorandum* 88202, pp. 1-18, January 1986.
9. H. J. Antonisse, "Image segmentation in pyramids", *Computer Graphics & Image Processing*, vol. **19**, pp. 367-383, 1982.
10. A. D. Gross et al., "Multiresolution object detection and delineation", *Computer Vision Graphics and Image Processing*, vol. **39**, pp. 102-115, 1987.
11. T. H. Hong et al., "Compact region extracting using weighted pixel linking in a pyramid", *IEEE Trans. on Pattern Anal. Machine Intell.*, vol. **PAMI-6** (2), pp. 229-237, 1984.
12. P. J. Burt, "Fast filter transforms for image processing", *Computer Graphics and Image Processing*, vol. **16**, pp. 20-51, 1981.
13. H. Konik et al., "Segmentation d'images par utilisation de pyramides à bases locales", *Traitement du Signal,* vol. **10 (4),** pp. 283-295, 1993.