AUTOMATIC COLOR PALETTE

J. Delon[†], *A. Desolneux*[††], *J.L. Lisani*[†††] *A.B. Petro*[†††]

[†]CMLA, ENS Cachan, France
email: julie.delon@cmla.ens-cachan.fr
[††]MAP5, Universit Paris 5, France
email: desolneux@math-info.univ-paris5.fr
[†††]Univ. Illes Balears, Spain
email: joseluis.lisani@uib.es, anabelen.petro@uib.es

ABSTRACT

Color palettes are an important tool for color image analysis, since they are the initial point of different techniques such as quantization or indexing. This paper presents a new method for the automatic construction of a color palette, which adjusts dynamically its number of colors according to the visual content of the image. The method is based on appropriately segmenting the HSI color space, which is achieved by individually partitioning the histograms associated to each color component. As a result we obtain a hierarchical color palette, which represents the color image with a reduced number of colors.

1. INTRODUCTION

The human visual system (HVS) is a complex and precise entity, it is able both to distinguish millions of colors and to describe an image by naming a few colors. This last characteristic derives from the HVS ability of grouping colors with similar tonality and of assigning a unique name to each group. Humans perform this process automatically but, computationally, it is a very difficult task.

The simplest way for computationally describing a color image is by means of its color palette which contains the more representative colors in the image just like the palette of a painter.

The color palette has different applications in many color image fields. It is an important tool in color quantization techniques, whose ultimate goal is to reduce the number of colors of an image with minimum distortion ([1]). Another application is color indexing, which is used both to compress color information ([2]) and to effectively retrieve information from image databases ([3]).

In view of these applications, we deduce the importance of obtaining a correct color palette, with a small number of colors but enough to produce an acceptable image representation. Moreover the color palette construction must be an automatic and fast process.

We present a new approach for the construction of the color palette that takes into account these previous remarks. This approach is based on the HSI color space properties and the use of a novel and parameter-free technique for histograms segmentation.

The paper is structured as follows. The basic ideas of the method are exposed in the next section; in Section 3, a new algorithm for histogram segmentation is presented; Section 4 is devoted to display and comment some results on color palette construction; the conclusions are presented in the final section.

2. COLOR PALETTE CONSTRUCTION

As pointed out in the previous section our goal is, given any color image, to automatically generate a palette that contains the "more representative" colors in the image.

The notion of "representative color" must be clarified. We start by choosing a color system representation, for that we have several options: RGB, HSI, $L^*a^*b^*$, etc. We decide for HSI , since it separates pure color information (hue and saturation) from brightness (intensity). Moreover, the notions of hue, saturation and intensity are intuitive and have a physical meaning ([4]).

Next, we want to partition the color space of the image (that is, the set of all its colors) in a minimum number of regions. For each region in this segmentation one color is then selected to form the palette.

Several methods can be devised to segment the HSI color space. One option is to perform a 3D segmentation but we take a different approach. From our study of color images we have concluded that not all the color components

This work has been partially financed by the Centre National d'Etudes Spatiales (CNES), the Office of Naval Research under grant N00014-97-1-0839, the Direction Générale des Armaments (DGA), the Ministerio de Ciencia y Tecnologia under grant TIC2002-02172.

(H, S, and I) have the same importance when comparing two colors. For example, a slight variation in the hue component may produce a big change in the perception of the color, while a variation of intensity may be slightly perceived. Therefore, it makes no sense to mix together the 3 components information in a 3D segmentation.

We make the following ordering of the color components according to their relevance to the definition of colors: hue, saturation, and intensity. Therefore, we propose the following 3-steps segmentation method for partitioning the HSI image space:

1. *Consider the histogram of hue values for the colors of the image. Partition this histogram using the technique described in Section 3.*
2. *For each one of the regions obtained in the previous step do the following:*
   (a) *Consider the saturation of the colors in the region.*
   (b) *Construct a histogram of saturations and segment it as in step 1.*
3. *For each one of the regions in the previous step, repeat steps 2a and 2b but now considering the intensity of the colors instead of their saturation.*

Observe that from the first step of the method we obtain a number of colors which is successively increased when adding saturation and intensity information. As a result of the algorithm we end up with a minimum set of color regions. A color can be chosen as a representative for each region by simply computing the mean value of the HSI components of the colors in the region. This set of minimum colors forms the final color palette.

In the next section, the techniques used to segment each color component histogram are described in detail.

3. HISTOGRAM SEGMENTATION

In 2003, Agnès Desolneux, Lionel Moisan and Jean-Michel Morel ([5]) defined a new parameter-free method for the detection of meaningful events in data. An event is called ε-meaningful if its expectation under the a contrario random uniform assumption is less than ε. Let us state what this definiton yields in the case of the histogram modes.

3.1. Uniform hypothesis. Meaningful intervals and gaps of a histogram.

In all the section, we will consider a discrete histogram r, that is N points distributed on L values $\{1, ..., L\}$. For each discrete interval $[a, b]$ of $\{1, ..., L\}$, $r(a, b)$ will represent the proportion of points in the interval. For each interval $[a, b]$ of $\{1, ..., L\}$, we note $p(a, b) = \frac{b-a+1}{L}$ the relative length of the interval. The value $p(a, b)$ is also, under the uniform assumption, the probability for a point to be in $[a, b]$. Thus, the probability that $[a, b]$ contains at least a proportion $r(a, b)$ of points among the N is given by the binomial tail $\mathcal{B}(N, Nr(a, b), p(a, b))$, where $\mathcal{B}(n, k, p) = \sum_{j=k}^{n} \binom{n}{j} p^j (1-p)^{n-j}$. The number of false alarms of $[a, b]$ is:

$$NFA([a, b]) = \frac{L(L+1)}{2} \mathcal{B}(N, Nr(a, b), p(a, b)).$$

Thus, an interval $[a, b]$ is said ε-meaningful if it contains "more points" than the expected average, in the sense that $NFA([a, b]) \leq \varepsilon$. In the same way, an interval $[a, b]$ is said to be an ε-meaningful gap if it contains "less points" than the expected average.

Now, these binomial expressions are not always easy to compute, especially when N is large. In practice, we adopt the large deviation estimate to define meaningful intervals and gaps. First, we define the relative entropy of an interval $[a, b]$ (with respect to the prior uniform distribution p) by

$$\begin{aligned} H([a, b]) \quad = \quad & r(a, b) \log \frac{r(a,b)}{p(a,b)} + \\ & +(1 - r(a, b)) \log \frac{1-r(a,b)}{1-p(a,b)}. \end{aligned}$$

$H([a, b])$ is the Kullback-Kleiber distance between two Bernouilli distributions of parameters $r(a, b)$ and $p(a, b)$ ([6]), that is $H([a, b]) = \mathrm{KL}(r(a, b)||p(a, b))$.

Definition 1 *An interval $[a, b]$ is said to be an ε-**meaningful interval** (resp. ε-**meaningful gap**) if $r(a, b) \geq p(a, b)$ (resp. $r(a, b) \leq p(a, b)$) and if its relative entropy $H([a, b])$ is*

$$H([a, b]) > \frac{1}{N} \log \frac{L(L+1)}{2\varepsilon}$$

3.2. Monotone hypothesis

How can we detect meaningful intervals and gaps if we know that the observed objects follow a non-uniform distribution (e.g. decreasing or increasing)? We want now to define the meaningfulness of an interval with respect to the decreasing hypothesis (the definitions and results for the increasing hypothesis can be deduced by symmetry). We will call $\mathcal{D}(L)$ the space of all decreasing densities on $\{1, 2, ..., L\}$ and $\mathcal{P}(L)$ be the space of normalized probability distributions on $\{1, 2, ..., L\}$.

If $r \in \mathcal{P}(L)$ is the normalized histogram of our observations, we need to estimate the density $p \in \mathcal{D}(L)$ in regards to which the empirical distribution r has the "less meaningful" gaps and intervals, which is summed up by the optimization problem

$$\tilde{r} = argmin_{p \in \mathcal{D}(L)} \min_{[a,b] \in \{1,2,...,L\}} KL(r(a, b)||p(a, b)).$$

The meaningfulness of intervals and gaps can then be defined relatively to this distribution $\tilde{r}$. Note that the uniform distribution is a particular case of decreasing density, which means that this formulation strengthens the previous theory: if there is no meaningful interval or gap in regards to the uniform hypothesis, there will be no meaningful interval or gap in regards to the decreasing hypothesis.

However, this optimization problem is not easy to solve. We choose to slightly simplify it by approximating $\tilde{r}$ by the Grenander estimator $\bar{r}$ of r ([7]), which is defined as the nonparametric maximum likelihood estimator restricted to decreasing densities on the line.

Definition 2 *The histogram $\bar{r}$ is the unique histogram which achieves the minimal Kullback-Leibler distance from r to $\mathcal{D}(L)$, i.e.*

$$KL(r||\overline{r}) = \min_{p \in \mathcal{D}(L)} KL(r||p).$$

It has been proven ([8]) that $\bar{r}$ can easily be derived from r by an algorithm called "Pool Adjacent Violators" that leads to a unique decreasing step function $\bar{r}$.

Now, the definitions of meaningful interval and gaps are analogous to the ones introduced in the uniform case, the uniform prior being just replaced by the global decreasing estimate $\overline{r}$ of the observed normalized histogram r.

Definition 3 *Let r be a normalized histogram. We say that an interval $[a, b]$ is ε-meaningful for the decreasing hypothesis (resp. an ε-meaningful gap for the decreasing hypothesis) if $r(a, b) \geq \overline{r}(a, b)$ (resp. $r(a, b) \leq \overline{r}(a, b)$) and $H_{\overline{r}}([a, b]) = \mathrm{KL}(r(a, b)||\overline{r}(a, b)) > \frac{1}{N} \log \frac{L(L+1)}{2\varepsilon}$.*

We are now able to define precisely what we called "to be almost decreasing on a segment".

Definition 4 *We say that a histogram **follows the decreasing (resp. increasing) hypothesis** on an interval if it contains no meaningful gap for the decreasing (resp. increasing) hypothesis on the interval.*

3.3. Acceptable Segmentations

The aim of our segmentation is to split the histogram in separated "modes". We will call "mode" an interval on which the histogram follows the increasing hypothesis on a first part and the decreasing one on the second part.

Definition 5 *We say that a histogram r **follows the unimodal hypothesis** on the interval $[a, b]$ if it exists $c \in [a, b]$ such that r follows the increasing hypothesis on $[a, c]$ and r follows the decreasing hypothesis on $[c, b]$.*

Such a segmentation exists. Indeed, the segmentation defined by all the minima of the histogram as separators follows obviously the unimodal hypothesis on each segment. But if there are small fluctuations it is clear that it is not a reasonable segmentation (see the left part of Figure 1). Fortunately, a segmentation following the unimodal hypothesis on each segment is generally not unique. We present a procedure that finds a segmentation much more reasonable than the segmentation defined by all the minima. We want to build a minimal (in terms of numbers of separators) segmentation, which leads us to introduce the notion of "acceptable segmentation".

Definition 6 *Let r be a histogram on $\{1, ..., L\}$. We will say that a segmentation s of r is **acceptable** if it verifies the following properties:*

- *r follows the unimodal hypothesis on each interval $[s_i, s_{i+1}]$.*
- *there is no interval $[s_i, s_j]$ with $j > i + 1$, on which r follows the unimodal hypothesis.*

It is clear in the discrete case that such a segmentation exists: we can start with the limit segmentation containing all the minima of r and gather the consecutive intervals together until both properties are verified. It is the principle used in the next algorithm:

Fine to Coarse (FTC) Segmentation Algorithm:

1. *Define the finest segmentation (i.e. the list of all the minima) $S = \{s_1, \ldots, s_n\}$ of the histogram.*
2. *Repeat:*

   Choose i randomly in $[2, length(S)-1]$. If the modes on both sides of s_i can be gathered in a single interval $[s_{i-1}, s_{i+1}]$ following the unimodal hypothesis, group them. Update S.

   Stop when no more unions of successive intervals follows the unimodal hypothesis.
3. *Repeat step 2 with the unions of j intervals, j going from 3 to $length(S)$.*

A result of this algorithm is shown on Figure 1. The left part of the figure shows the initialization of the algorithm (all the minima), and the final result is on the right.

4. RESULTS

We display three examples of color palettes obtained with the proposed method. For each one of the examples we display 3 images. The first one is the original image; the second one is the image obtained by replacing the original colors with the colors in the palette; and the third one shows

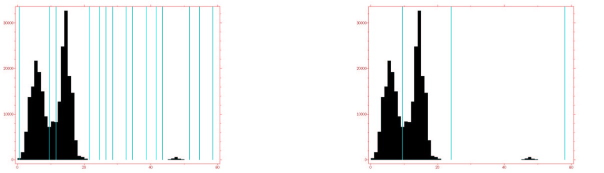

Fig. 1. Left: all the minima of the histogram. Right: remaining minima after the fine to coarse algorithm.

the hierarchical palette constructed with the algorithm described in Section 2. Each row in this image displays the colors obtained at each step of the algorithm. Remark that the number of colors increases when saturation and intensity information is added to the results obtained for the hue component.

5. CONCLUSIONS

We present an automatic and adapted method for the automatic construction of a color palette for any given image. The palette represents the main colors in the image and it is constructed using a hierarchical algorithm that attaches more importance to the hue and saturation components than to the intensity component in the HSI color representation.

Other options, such as split and merge techniques, are being currently studied in order to improve the obtained color palette.

Acknowledgement: The authors gratefully thank F. Mettini for her kind authorization to use the "ladybug" image.

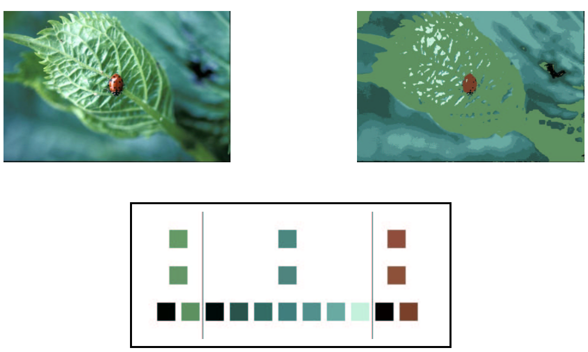

Fig. 2. *Top left:* Original image "ladybug". *Top right:* Resulting image with 11 colors after hue, saturation and intensity steps. *Bottom:* Color palette. The hue values for this image are represented in the histogram of Figure 1. Since the histogram is divided in three regions, the first row in the color palette contains just three colors (corresponding to leaf, background and ladybug, respectively). As information on saturation and intensity is added, the number of colors increases.

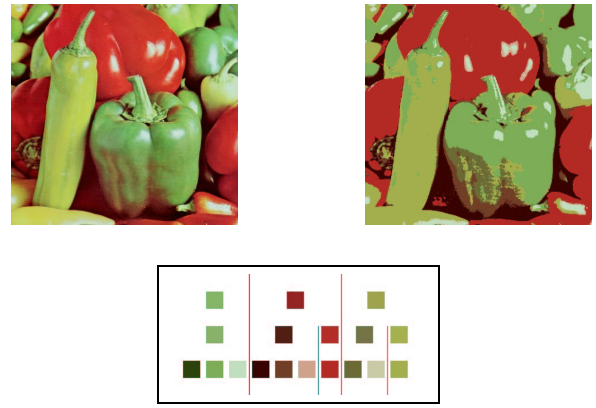

Fig. 3. *Top left:* Original image "pepper". *Top right:* Resulting image with 10 colors after hue, saturation and intensity steps. *Bottom:* Color palette.

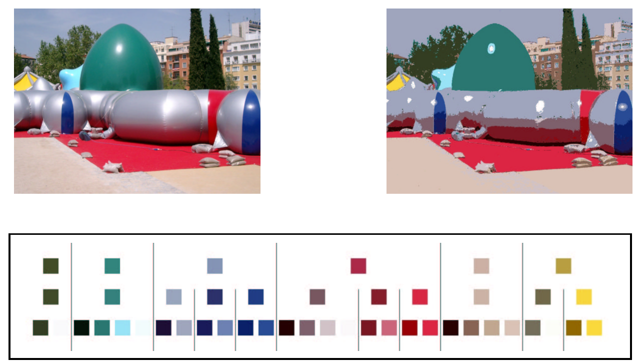

Fig. 4. *Top left:* Original image "madrid". *Top right:* Resulting image with 28 colors after hue, saturation and intensity steps. *Bottom:* Color palette.

6. REFERENCES

[1] I. Andreadis A. Atsalakis, N. Papamarkos, "On estimation of the number of image principal colors and color reduction through self-organized neural networks," *Int. Journal of Imaging Systems and Technology*, vol. 12, no. 3, pp. 117–127, 2002.

[2] K.N. Plataniotis and A.N. Venetsanopoulos, *Color Image Processing and Applications*, Springer, 2000.

[3] B. M. Mehtre M. S. Kankanhalli and J. Kang Wu, "Cluster-based color matching for image retrieval," *Pattern Recognition*, vol. 29, no. 4, pp. 701–708, 1986.

[4] G. Wyszecki and W.S.Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, Wiley, 1984.

[5] A. Desolneux, L. Moisan, and J.M. Morel, "A grouping principle and four applications," *IEEE Trans. on PAMI*, vol. 25, no. 4, pp. 508–513, 2003.

[6] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley Series in Telecommunications. John Wiley & Sons Inc., New York, 1991, Wiley-Interscience.

[7] U. Grenander, "On the theory of mortality measurement, part II," *Skand. Akt.*, vol. 39, pp. 125–153, 1956.

[8] M. Ayer, H.D. Brunk, G.M. Ewing, W.T. Reid, and E. Silverman, "An empirical distribution function for sampling with incomplete information," *The Annals of Math. Statistics*, vol. 26, no. 4, pp. 641–647, 1955.