Measurement and Display of Color Image Differences Based on Visual Attention*

Alain Trémeau, Eric Dinet, and Eric Favier

Université Jean Monnet, Equipe Ingénierie de la Vision, Lab. T.S.I. UMR 5516, 3 rue Javelin Pagnon, BP 505, 42007 Saint-Etienne Cedex 1, France

Current subjective methods for evaluation and comparison of image processing techniques are generally not sufficiently relevant and accurate, essentially because they are not correlated with human observation. We propose a method to measure and display color image differences derived from the basic characteristics of human visual perception. This method provides images of color differences. These images can be used to define a global relative measure to predict image quality.

Journal of Imaging Science and Technology 40: 522–534 (1996)

Introduction

Measuring performance of image processing techniques is essential to evaluate results given by an image process, to improve the quality of an image process, or to compare different image processes. For example, measuring performance can provide a feedback path by which a system can modify its strategy during processing.[1]

In earlier work, comparison was realized between one image produced by a human and another one produced by the algorithm to be tested. This approach requires a reference image corresponding to human observation, which is usually not available in automatic processing, rendering the proposed measures useless for most processing techniques.[1] For example in segmentation the reference image used to evaluate the final result is that of human performance in segmenting the image. Performance parameters are then measured separately for each region.

Rather than having an a priori knowledge about regions in the initial image, we propose in this report to define performance parameters according to the original studied image and to the visual observation independently of the reference image. Moreover we suggest developing measures that are independent of image processes. Measures proposed in this study have been developed to be applied indifferently to segmentation, quantization, compression, or other image processing techniques. Images compared here illustrate the evaluation of two segmentation processes.

Current subjective methods for evaluation and comparison of image processing techniques are inadequate. In the same way, current objective quantitative measures are not effective because proposed criteria do not accurately reflect the amount of disagreement between an original and a test image.[2] Most of these criteria are linked to contrast, sharpness, spatial frequency, or intensity parameters computed pixel by pixel or region by region.[1,3–5] Then error measures are computed according to distortions produced between image parameters and weighted according to their importance.[2,6] Nevertheless because these error measures are not evaluated in terms of their correlation with a human observer, they are inadequate to compare images accurately. Indeed, image processing techniques require accurate measures of subjective differences to predict image quality.

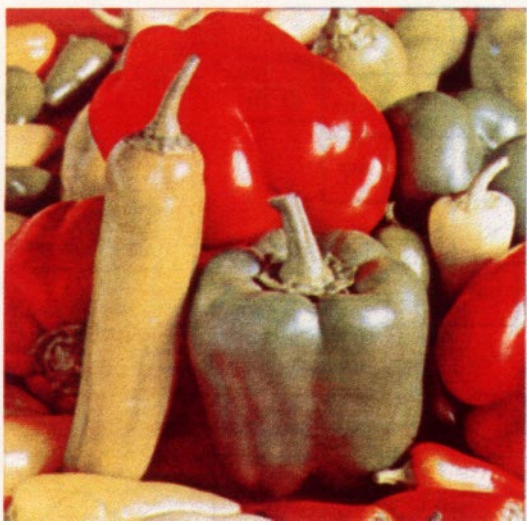
Although measures to compare images have been improved during the last few years, we believe they do not make enough reference to visual observation. Even if experience and theory show that it is extremely difficult to define an objective measure that takes into account most significant phenomena linked to visual observation, it has been shown that we can, nevertheless, define a heuristic measure from subjective criteria to evaluate image differences.[7,8] Such a heuristic measure is based on subjective criteria that have been selected both because they are linked to psychometric observations and because they relate to independent but complementary features.

Three criteria are considered in this study: the brightness difference, the dispersion difference, and the emergence difference. These criteria correspond to physical measures linked to psychometric observations. Among these measures, contrast is one of the image quality features most commonly used.[1,4,5] In this report, contrast is computed through the evaluation of brightness differences. The spatial frequency features that are also commonly used to measure spatial activity, such as noise, texture, or uniformity, are calculated here to define dispersion differences and emergence differences; local and surround excitation and inhibition effects also are taken into account.[5,9] All these measures are successively presented in the following sections. These criteria are defined locally according to a neighborhood function to emphasize the weight of certain pixels like the principle focus of attention.[10] Moreover, using local rather than global measures leads to accurate evaluation of image quality, because the measure is fitted to critical areas rather than over the whole image.[9]

These three criteria are combined to define a local image correlation measure. By definition, the criteria involved in this combination are independent, because they are derived from a priori independent psychometric attributes and because they relate to different physical attributes that are really independent. It could be interesting to analyze quantitatively the interdependence of these criteria by statistical evaluations, but it would require a lot of testing and observation, whereas the results themselves seem to satisfy the independence hypothesis, as evidenced by Figs. 6, 8, and 9. These figures successively present the three criteria ranged from 0 to 1 (i.e., from

Original manuscript received January 30, 1996. Revised September 12, 1996.

* Presented in part at the 3rd IS&T/SID Color Imaging Conference, November 7–10, 1995, Scottsdale, Arizona.



(a)

(b)

(c)

Figure 1. (a) Image Peppers; (b) result achieved with Segmentation Process 1; (c) result achieved with Segmentation Process 2.

most important differences to no difference) and displayed through an 8-bit digital image.

We will show that to reflect human observation accurately, measures need to be defined according to the degree of homogeneity of the image, such as textured or edge areas. The measure of edge differences is dissociated from other measures because it is strongly correlated to them, especially to the dispersion difference measure. Consequently, this measure cannot be multiplexed to the others but requires specific processing. We will present at the end of this report a masking process from which we can adjust our local image correlation measure according to edge differences. From a perceptual point of view, edge elements seem to play a specific part in the observation of images.[1,5,8] In most cases, edge elements prevail over surrounding region characteristics because they focus the attention more. That is the reason for a masking process.

The local correlation measure is computed for each pixel. It can therefore be displayed initially as an image of color differences. We show that results emphasize distortions observed between the original image and the test image. Then we propose a global measure of "average misfit" between images from the spatial distribution of distortions on the image of local correlations. Other developments will be necessary to define a more accurate global measure. They will be presented in a subsequent report because they are outside the scope of this study.

Fields of Study

Image *Peppers* shown in Fig. 1(a) and two segmentation processes, described as Process 1 and Process 2, have been chosen to illustrate parameters and criteria that have been computed in this study to underline and to quantify the most noticeable image differences. The segmented images are shown in Figs. 1(b) and 1(c).

In the following notation, I is the reference image and J is the image to study. For example, I can be the original image and J an image resulting from a segmentation process. From these two images we compute three new images: the local brightness difference image B, the local dispersion difference image C, and the local emergence difference image E, which will be used to compute the comparison image D. These images are computed according to the formulae introduced in the following sections. For example, Figs. 6(a), 8(a), 9(a), and 11(a) show, respectively, images B, C, E, and D computed from the result achieved with Process 1 for image *Peppers* of Fig. 1(a). In the same way Figs. 6(b), 8(b), 9(b), and l l(b)

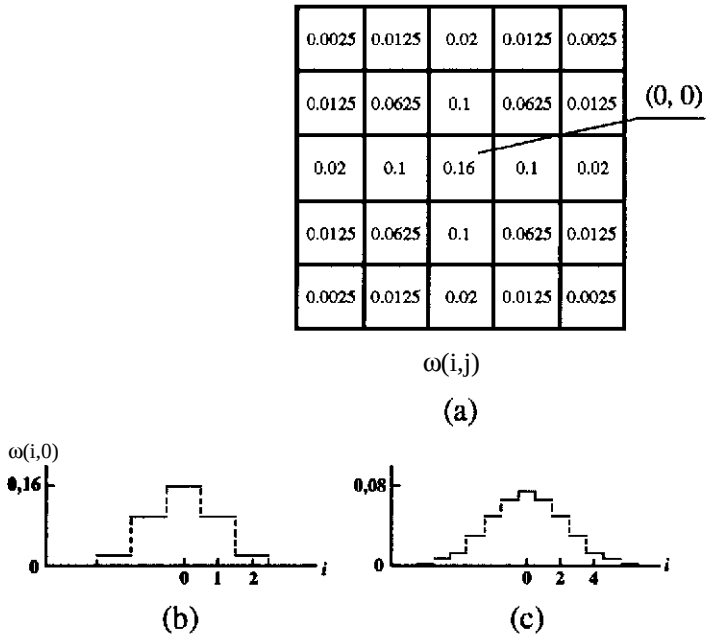


Figure 2. Weighting function ω used in this study:[12] (a) ω defined at (i, j) with $-2 \le i, j \le 2$; (b) ω viewed according to i axis; (c) ω rescaled and expanded twice; (d) the weighting function ω converges rapidly to a characteristic Gaussian-like form.

correspond to the results provided by Process 2 from the image *Peppers*.

By definition, segmented images present homogeneous regions that can be displayed by only one representative color per region. Likewise quantized images present homogeneous sets of colors that can be displayed by only one representative color per set. Consequently, whatever the process applied to an image, it is important to analyze whether representative computed colors are really relevant for the image under study. In that way we propose first to compute an image of differences and second to define measurements to analyze these differences from the image of comparison. Rather than computing an absolute measure of these differences, we propose to compute a relative measure. This allows us either to improve the quality of an image process or to compare different image processes.

Notations and Formulas

Definition of a Weighting Function. Let us introduce several notations and formulas used in this report to define our criteria of comparison.

Let P be the image plane and

$$V(x,y) = \{(x + i,\ y + j) \in P \mid -2 \le i, j \le 2\} \tag{1}$$

be the set of pixels that belong to the neighborhood of pixel (x, y). This neighborhood must not be too large, because the larger its size the fewer the peripheral pixels clustered in the neighborhood linked to the central pixel.[11] Moreover, the smaller this neighborhood is, the more image processing techniques emphasize local information, such as edges and details, without increasing their size.

Let ω be a weighting function that is normalized, symmetric, and unimodal, as described in Fig. 2. This function gives to each pixel a weight that is proportional to its distance from the central point [see Figs. 2(a) and 2(b)]. It is represented as a discrete function defined only for integer values on i and j axes. The weighting function ω can be rescaled and adjusted for fractional values, as shown in Figs. 2(c) and 2(d). With weights initially chosen [see Fig. 2(a)], its shape converges rapidly to a characteristic form that closely resembles the Gaussian probability density function.[12]

Even if this function does not correspond to the mathematical model that has been defined to simulate the focus of attention of visual acuity by Burt,[13] it can be considered as a good descriptor for a 5×5 kernel for which the goal is to emphasize central vision relative to the periphery. Thus, when we use this kernel in a convolution we underline the contribution of the central pixel and, in the same way, take into account the influence of the surrounding pixels that belong to the neighborhood under study.

Definition of an Average Function and a Variance Function. Let

$$\mu_\omega^K(x,y) = \sum_i \sum_j \omega(i,j) \cdot f^K(x + i,\ y + j) \tag{2}$$

be, according to weighting function ω, the weighted average value of the $f^K(x + i, y + j)$ values of the neighborhood $V(x, y)$ centered at pixel (x, y). The value $f^K(x + i, y + j)$ represents

- the gray level of pixel $(x + i, y + j)$ when $K = 0$,
- the trichromatic components of pixel $(x + i, y + j)$ when $K = 1, 2, 3$. For example, f^1, f^2, and f^3 will, respectively, represent the R, G, and B values of pixel $(x + i, y + j)$.

Let

$$\begin{aligned}\left(\sigma_\omega^K(x,y)\right)^2 &= \sum_i \sum_j \omega(i,j) \cdot \left[f^K(x + i,\ y + j) - \mu_\omega^K(x,y)\right]^2 \\ &= \sum_i \sum_j \omega(i,j) \cdot \left[f^K(x + i,\ y + j)\right]^2 - \left[\mu_\omega^K(x,y)\right]^2\end{aligned} \tag{3}$$

be, according to weighting function ω, the variance around the weighted average value of the $f^K(x + i, y + j)$ values of the neighborhood $V(x, y)$ centered at pixel (x, y).

The two functions μ_ω^K and σ_ω^K are then locally defined by convolving the neighborhood of each pixel with the weighting function introduced above (see Fig. 3).

Definition of Covariance Functions. Let

$$\begin{aligned}\mathrm{cov}_\omega^K(x,y) &= \sum_i \sum_j \omega(i,j) \cdot \left[f_I^K(x + i,\ y + j) - \mu_{I_\omega}^K(x,y)\right] \\ &\quad \bullet \left[f_J^K(x + i,\ y + j) - \mu_{J_\omega}^K(x,y)\right] \\ &= \sum_i \sum_j \omega(i,j) \cdot f_I^K(x + i,\ y + j) \cdot f_J^K(x + i, y + j) \\ &\quad - \mu_{I_\omega}^K(x,y) \cdot \mu_{J_\omega}^K(x,y)\end{aligned} \tag{4}$$

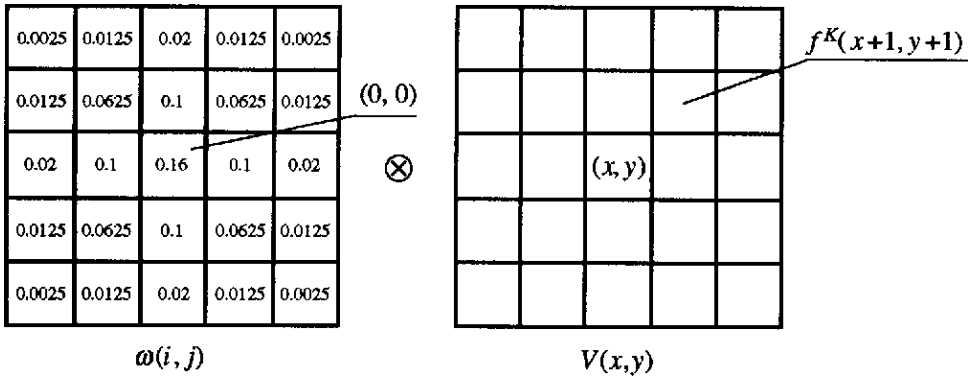


Figure 3. Convolving the neighborhood $V(x, y)$ by weighting function ω to define an average value for pixel (x, y).

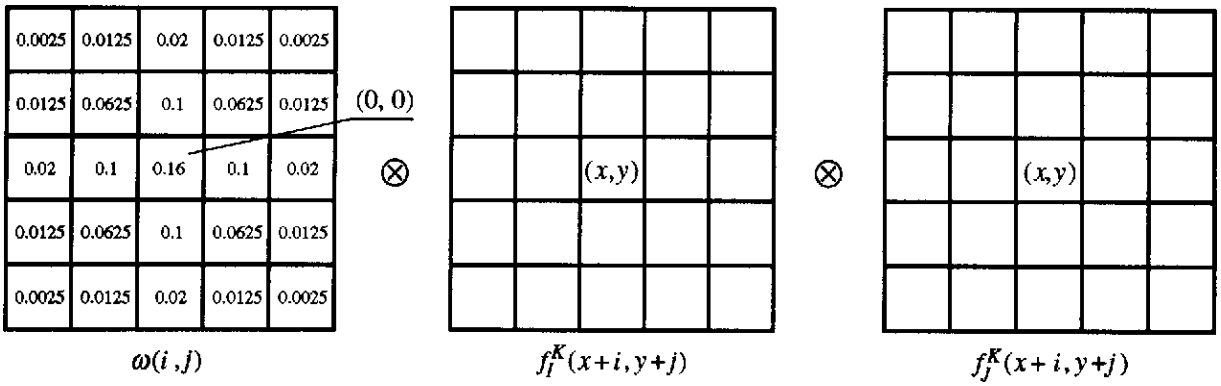


Figure 4. Cross-convolving of the two color distributions f_I^K and f_J^K with kernel $\omega(i, j)$ to define a covariance value for pixel (x, y).

be the covariance of the $f^K(x+i, y+j)$ values of the neighborhood $V(x, y)$ centered at pixel (x, y) and computed between two images I and J according to kernel ω. The values $f_I^K(x,y)$ and f_J^K (x, y) represent the respective $f^K(x, y)$ values for Image I and Image J. Likewise, $\mu_{I_\omega}^K$ (x, y) and $\mu_{J_\omega}^K$ (x, y) represent the weighted average values of the respective $f^K(x, y)$ values for Image I and Image J.

The function $\mathrm{cov}_\omega^K(x,y)$ is locally defined by cross-convolving the two color distributions of images I and J with the weighting function ω under study (see Fig. 4).

In contrast to functions μ_ω^K and σ_ω^K, which characterize only the local color distribution of an image, the function cov_ω^K enables us to analyze simultaneously two local color distributions and differences that can appear between these two distributions.

Whatever the descriptors used to characterize color distributions, we can indifferently use scalar or vectorial notations. We can use any Euclidean metric as norms L_1, L_2, or L_∞. Nevertheless, considering there is no order relation between colors, it does not make any sense to define an average color value for a set of colors or to compare two averaged color values when the colors involved do not have the same hue. Consequently, rather than using a 3-dimensional color descriptor to characterize color distributions and their differences we will use a combination of the three 1-dimensional descriptors computed from the three color components f^1, f^2, and f^3.

Then, by using norm L_1, the color variance value and the color covariance value will simply be defined according to the following formulae:

$$\sigma_\omega(x,y) = \sum_{K=1}^{3} \sigma_\omega^K(x,y), \tag{5}$$

$$\mathrm{cov}_\omega(x,y) = \sum_{K=1}^{3} \mathrm{cov}_\omega^K(x,y). \tag{6}$$

Another way to measure differences of distribution between two images consists in analyzing each distribution according to the other one.

Let

$$\left[e_{IJ}^K(x,y)\right]^2 = \sum_i \sum_j \omega(i,j) \cdot \left[f_I^K(x+i, y+j) - f_J^K(x,y)\right]^2 \tag{7}$$

be a measure of dispersion of the central value $f_J^K(x,y)$ of Image J relative to the f_I^K values of Image I on the neighborhood $V(x, y)$ centered at point (x, y) (see Fig. 5).

In the same way we can define the measure $e_{JI}^K(x,y)$, which is symmetrical with the first one, or we can define two measures $e_{II}^K(x,y)$and $e_{JJ}^K(x,y)$, which compute the variance of the neighborhood values around the central values, respectively, for Image I and Image J.

Remark. All the previous examples have been computed by convolving the neighborhood under study with kernel ω to emphasize the weight of the central pixel.

To analyze only the color distribution of a neighborhood without making reference to the central pixel, we can use measures such as those previously introduced but changing the weight of each kernel element by the same constant value. For example, the average and variance values of neigborhood $V(x, y)$ can be defined for an image I by:

$$\mu_I^K(x,y) = \frac{1}{25} \sum_{i=-2}^{2} \sum_{j=-2}^{2} f_I^K(x+i, y+j), \tag{8}$$

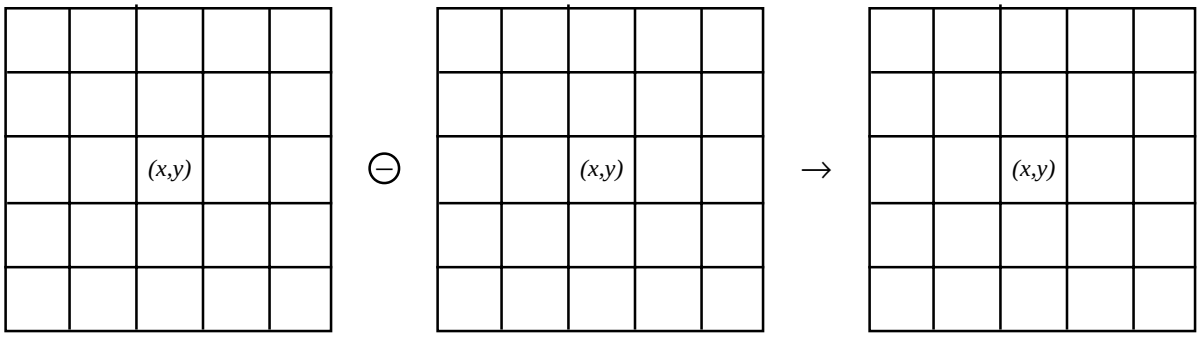


Figure 5. Measure of dispersion between the $f_J^K(x,y)$ and the $f_I^K(x+i, y+j)$ values.

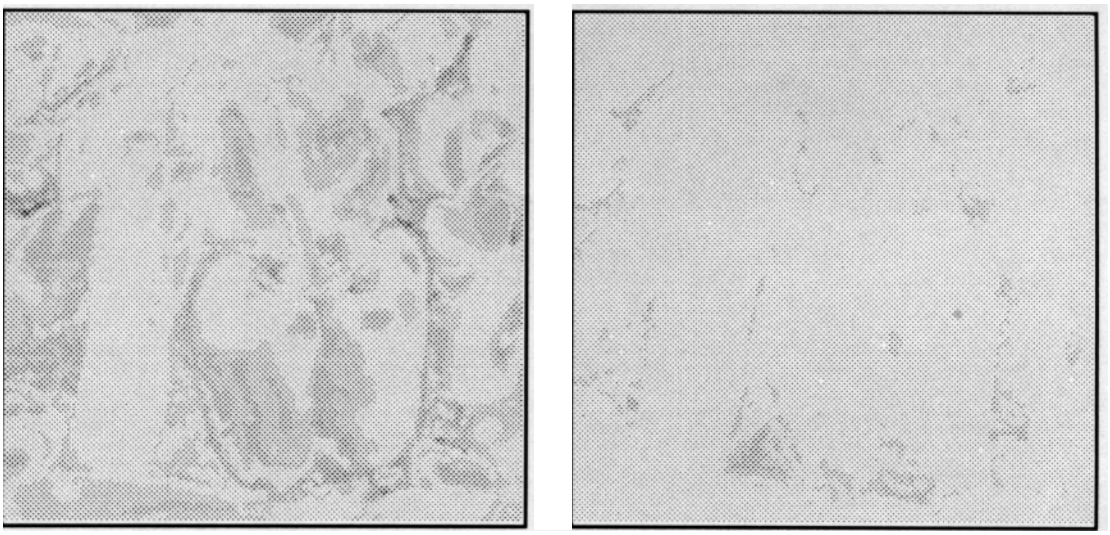


Figure 6. (a) Brightness differences between Figs. 1(a) and 1(b), (b) brightness differences between Figs. 1(a) and 1(c).

$$\left[\sigma_I^K(x,y)\right]^2 = \frac{1}{25}\sum_{i=-2}^{2}\sum_{j=-2}^{2}\left[f_I^K(x+i,y+j) - \mu_I^K(x,y)\right]^2. \quad (9)$$

Local Brightness Differences

Three criteria have been used to define a color correlation measure. The first one involves the local brightness differences defined as follows:[10]

$$B^K(x,y) = 1 - \frac{\left|\log\mu_{I_\omega}^K(x,y) - \log\mu_{J_\omega}^K(x,y)\right|}{\log L_{\max}^K - \log L_{\min}^K} \text{ with } K = 0, \quad (10)$$

and where

$$L_{\max}^K = \max\left\{f_I^K(x,y), f_J^K(x,y) | (x,y) \in P\right\},$$
$$L_{\min}^K = \min\left\{f_I^K(x,y), f_J^K(x,y) | (x,y) \in P\right\}.$$

The values $L_{\max}^K$ and $L_{\min}^K$, respectively, represent for K = 0 the highest and lowest brightness values of the two images under study.

Thus $B^K(x,y) \in [0,1]$ with $B^K(x,y) = 1$ when brightness difference is minimal and $B^K(x,y) = 0$ when brightness difference is maximal.

The B^K values are computed according to the logarithm of the f^K values to take into account the perception of brightness, which is not linearly proportional to luminance or color differences but approximately linearly proportional to the logarithms of these quantities.[14]

Let B be the image corresponding to the B^K values when $K = 0$. This image can be rescaled in the range [0, 255] and can be viewed as a gray-level image to show pixels for which brightness differences are the highest (black pixels) or the lowest (white pixels). For example, Fig. 6 shows the brightness differences between the images in Fig. 1(a) and 1(b) and between the images in Fig 1(a) and 1(c).

Values given by Eq. 10 involve only *brightness differences* without taking into account chromatic differences. It is possible to generalize Eq. 10 to chromatic differences by considering the B^K values for K = 1, 2 or 3, i.e., color component by color component. Image B′ of *local chromatic differences* will be computed according to the following formula:

$$B'(x,y) = \alpha_1 \cdot B^1(x,y) + \alpha_2 \cdot B^2(x,y) + \alpha_3 \cdot B^3(x,y) \text{ such that } B'(x,y) \in [0,1]. \quad (11)$$

Image B'' of *local color differences* will be computed with the following formula:

$$[B''(x,y)]^{1/2} = \beta_1 \cdot [B^0(x,y)]^2 + \beta_2 \cdot [B'(x,y)]^2 \text{ such that } B''(x,y) \in [0,1]. \quad (12)$$

Most of the time color differences observed between two images are essentially due to brightness differences rather than to chromatic differences. Color differences

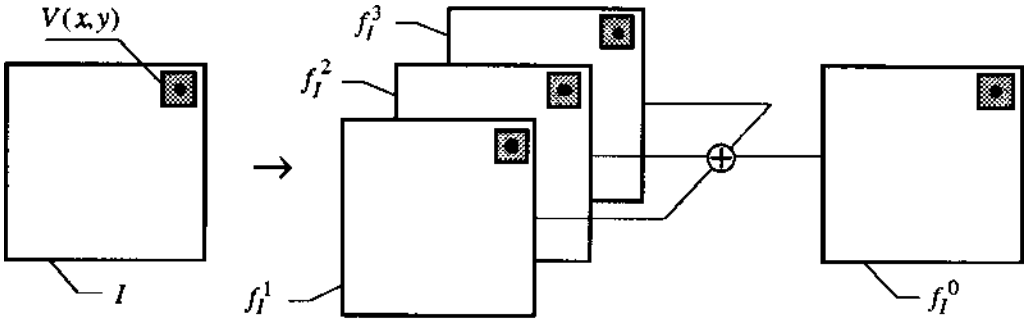


Figure 7. For each color image there are three corresponding color component images from which we can compute a gray-level image characteristic of the brightness of the color image.

(a) (b)

Figure 8. (a) Dispersion differences between Figs. l(a) and l(b), (b) dispersion differences between Figs. l(a) and l(c).

are primarily perceived in terms of brightness differences, because our visual sensitivity to brightness differences is more accurate than our visual sensitivity to chromatic differences. Moreover, by stating that in color image analysis, the brightness value f^0 is computed from the three color component values f^1, f^2, and f^3 (see Fig, 7), we can consider that the calculation of brightness differences also leads to computation of chromatic differences. Consequently, in our study we have taken into account only local brightness differences.

It is important to note that function μ_ω^K given by Eq. 2 can be applied to the brightness values f^0 and to the three chromatic component values f^1, f^2, and f^3 but cannot be applied to the trichromatic color value (f^1, f^2, f^3), because no color relation order can be defined. Indeed it does not make any sense to define an average color value for a set of colors and to compare two average color values unless the colors involved have roughly the same hue.

Local Dispersion Differences

The second criterion used to define the color correlation measure is linked to the local dispersion difference, which is expressed as follows:[10]

$$\mathrm{cor}(x,y) = \frac{1}{3}\sum_{K=1}^{3} \mathrm{cor}^K(x,y) \tag{13}$$

with

$$\mathrm{cor}^K(x,y) = \frac{\mathrm{cov}_\omega^K(x,y)}{\sigma_{I_\omega}^K(x,y)\cdot\sigma_{J_\omega}^K(x,y)}, \tag{14}$$

where $\sigma_{I_\omega}^K(x,y)$ and $\sigma_{J_\omega}^K(x,y)$ represent the variance values of the respective f^K values for Image I and Image J.

Then $\mathrm{cor}(x,y) \in [-1,1]$ with $\mathrm{cor}(x,y) = 1$ when the dispersion difference is minimal (i.e., correlation maximal), $\mathrm{cor}(x,y) = 0$ when the dispersion difference is maximal (i.e., correlation minimal), and $\mathrm{cor}(x,y) = -1$ when the dispersion difference is minimal inverse (i.e., correlation maximal inverse).

To avoid problems of uncertainty about a null variance we will consider the following definition:

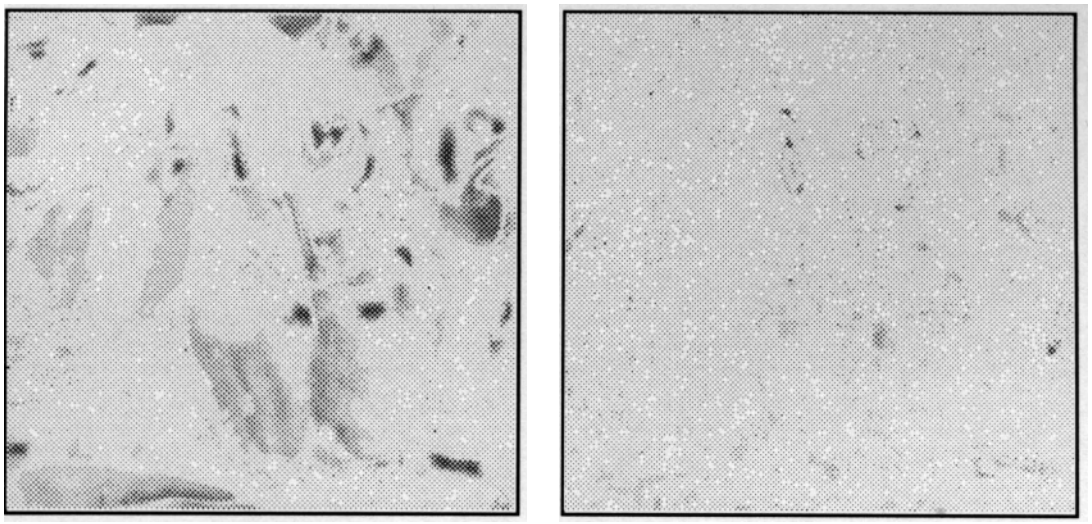
$$\begin{aligned}
\mathrm{cor}^K(x,y) &= \frac{\mathrm{cov}_\omega^K(x,y)}{\sigma_{I_\omega}^K(x,y)\cdot\sigma_{J_\omega}^K(x,y)} && \text{if } \sigma_{I_\omega}^K(x,y) > \varepsilon \text{ and } \sigma_{J_\omega}^K(x,y) > \varepsilon,\\
&= 1 && \text{if } \sigma_{I_\omega}^K(x,y) \le \varepsilon \text{ and } \sigma_{J_\omega}^K(x,y) \le \varepsilon,\\
&= 0 && \text{otherwise.}
\end{aligned}$$

Let C be the image corresponding to the $|\mathrm{cor}(x,y)|$ values. This image can be scaled between 0 and 255 and can be viewed as a gray-level image to show pixels where dispersion differences are the greatest or the lowest. We assume that the sign of the correlation value does not matter in the case under study.

Consequently black pixels correspond to the greatest dispersion differences and white pixels correspond to the lowest dispersion differences. For example, Figs. 8(a) and 8(b) show the dispersion differences between Image (a) and Images (b) and (c) of Fig. 1.

Local Emergence Differences

The last criterion used to define the color correlation measure is linked to the local emergence differences and is stated as follows:



(a) (b)

Figure 9. (a) Emergence differences between Figures 1(a) and 1(b); (b) emergence differences between Figs. 1(a) and 1(c).

$$E^K(x,y) = 1 - \frac{\left\| \left[e^K_{II_\omega}(x,y) - e^K_{IJ_\omega}(x,y) \right] \cdot \left[e^K_{JJ_\omega}(x,y) - e^K_{JI_\omega}(x,y) \right] \right\|}{\left(e^K_{\max} - e^K_{\min} \right)^2} \quad (15)$$

with $K = 0$, where $e^K_{II_\omega}(x,y)$ and $e^K_{JJ_\omega}(x,y)$ represent, respectively, for Image I and for Image J, the *local emergence* of the central pixel according to its neighborhood. In the same way, $e^K_{IJ_\omega}(x,y)$ and $e^K_{JI\omega}(x,y)$ represent, respectively, for Images I and J and for Images J and I, the *local cross-emergence* of a pixel of an image according to the corresponding neighborhood on the other image.

Thus $\left(e^K_{II_\omega}(x,y) - e^K_{IJ_\omega}(x,y) \right)$ represents the *difference of local emergence* of a pixel (x,y) of Image I with regard to Image I and to Image J when pixel (x,y) is compared with its neighborhood.

In the same way, the other cross-difference represents the difference of local emergence of a pixel (x,y) of Image J with regard to Image I and to Image J when pixel (x,y) is compared with its neighborhood. Whatever the change that intervenes in an image, it could be analyzed not only in the context of its own neighborhood, but also in the context of the original neighborhood.

In Eq. 15

$$e^K_{\max} = \max \left\{ \left| e^K_{NN_\omega}(x,y) - e^K_{NM_\omega}(x,y) \right| (x,y) \in P;\ N = I,J \ \text{ and } M = I,J \right\}$$

represents the highest emergence value measured into the two images under study, and

$$e^K_{\min} = \min \left\{ \left| e^K_{NN_\omega}(x,y) - e^K_{NM_\omega}(x,y) \right| (x,y) \in P;\ N = I,J \ \text{ and } M = I,J \right\}$$

represents the lowest emergence value measured into the two images under study.

Then $E^K(x,y) \in [0,1]$ with $E^K(x,y) = 1$ when emergence difference is minimal and $E^K(x,y) = 0$ when emergence difference is maximal.

Let E be the image corresponding to the E^K values when $K = 0$. This image can be scaled between 0 and 255 and can be viewed as a gray-level image to show pixels where emergence differences are the highest (black pixels) or the lowest (white pixels). For example, Figs. 9(a) and 9(b) show the emergence differences between Image (a) and Images (b) and (c) of Fig. 1.

Values given by Eq. 15 involve only brightness emergence differences without taking into account chromatic emergence differences. As for brightness differences, it will be possible to generalize Eq. 15 to chromatic emergence differences by considering E^K values for $K = 1$, 2, or 3, i.e., color component by color component. Image E' of *local chromatic emergence differences* will be computed according to the following formula:

$$E'(x,y) = \alpha_1 \cdot E^1(x,y) + \alpha_2 \cdot E^2(x,y) + \alpha_3 \cdot E^3(x,y) \text{ such that } E'(x,y) \in [0,1]. \quad (16)$$

Image E'' of *local color emergence differences* will be computed with the following formula:

$$[E''(x,y)]^{1/2} = \beta_1 \cdot [E^0(x,y)]^2 + \beta_2 \cdot [E'(x,y)]^2 \text{ such that } E''(x,y) \in [0,1]. \quad (17)$$

We refer to the justifications previously given to understand these two definitions.

Measure of Local Image Correlations

As mentioned in the Introduction, each of the three criteria previously presented is scaled into the same range of values from 0 to 1 and is defined independently of the others. Therefore the measure of local image correlations can be based on an orthogonal representation. In other words, the three criteria are combined on the base of norm L_2 as shown in Fig. 10.

Thus

$$D(x,y) = \frac{\sqrt{\left[B^0(x,y) \right]^2 + \left[\text{cor}(x,y) \right]^2 + \left[E^0(x,y) \right]^2}}{\sqrt{3}}. \quad (18)$$

Consequently $D(x,y) \in [0,1]$ with $D(x,y) = 1$ when *local image correlations* are maximal (i.e., minimal brightness, dispersion, and emergence differences) and $D(x,y) = 0$ when *local image correlations* are minimal.

Let D be the image corresponding to the D values. This image can be scaled between 0 and 255 and can be viewed as a gray-level image to show pixels where local image correlations are the highest or are the lowest. For example,

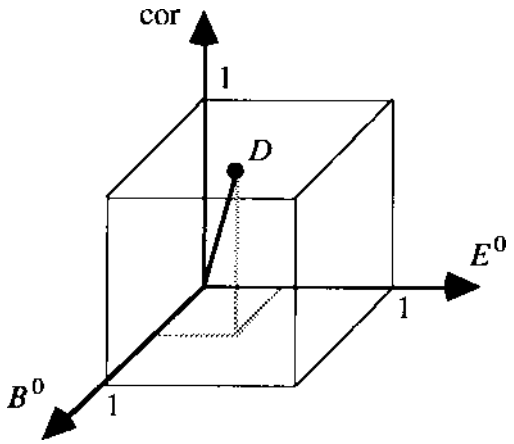


Figure 10. Combination of the three criteria B^0, cor, and E^0 to define a measure of local image correlations.

Figs. 11(a) and 11(b) show image correlations measured, respectively, between Image (a) and Images (b) and (c) of Fig. 1.

A false color look-up table has been used in the display process to emphasize different types of local image correlations. As shown by Fig. 11(c), red color corresponds to pixels where local image correlation is the highest and purple color corresponds to pixels where local image correlation is the lowest.

We observe in Figs. 11(a) and 1l(b) that contours appear well correlated, whatever the process under study and whatever the difference of magnitude of the contours between the original image and the processed images. As we can see in Figs. 8(a) and 8(b), this contour correlation is essentially provided by the computation of dispersion differences.

Note that the images used in this study to illustrate our measure of color image differences result from a segmentation process. Moreover, these images have been naturally displayed according to each average color value computed for each region. Consequently, the more homogeneous a region is in the original image, the less color differences are in the region between the original image and the segmented one. On the other hand, the more the spatiocolor distribution of a region is textured, the greater these differences become. In other words, color differences resulting from a segmentation process are essentially linked to dispersion differences, but other color differences can nevertheless superimpose over the dispersion differences, such as emergence differences linked to the loss of details in segmented images, or luminance differences linked to the fading of brightness areas.

This is the reason that the results obtained using all three measures look quite similar to those using only the dispersion criterion, as shown by Fig. 11(a) and Fig. 8(a). For other image processes, such as image compression, for example, the observations would be different. Actually color differences are just as much dependent on the images considered as on the processes under study, which justifies the use of the three criteria previously defined.

Adjustment of the Measure to Specifilc Image Analysis Problems

A More Important Weight for Brightness Differences. In Eq. 18, the same weight has been given to each component. Nevertheless it is possible to amplify the weight of one component by rescaling and expanding its distribution in the range [0,1] toward the central value or toward the boundaries.

For example, in Eq. 10 the computation of local brightness differences has been normalized according to the highest and the lowest brightness differences measured in the whole images under study.

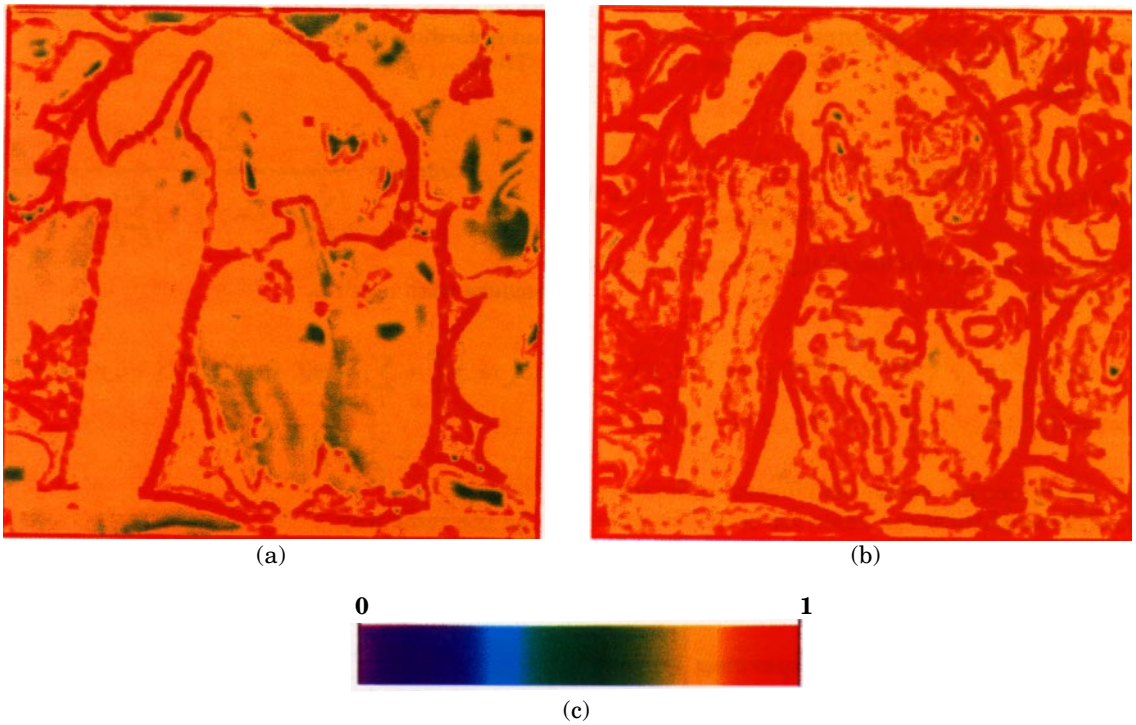
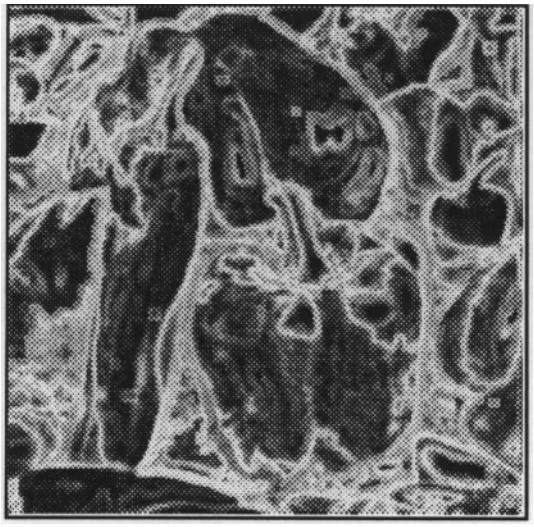
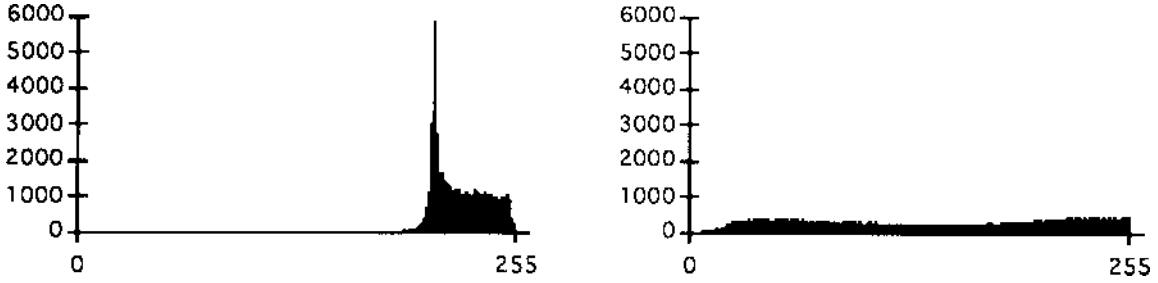


Figure 11. (a) Image of *local image correlations* between Figs. 1(a) and 1(b); (b) image of *local image correlations* between Figs. 1(a) and 1(c); (c) false color look-up table used to display images of *local image correlations*: red color corresponds to pixels where local image correlation is the highest and purple color corresponds to pixels where local image correlation is the lowest.



(a)



(b) (c)

Figure 12. (a) Brightness differences between Fig. l(a) and Fig. l(b) computed with Eq. 19; (b) histogram of the image of Fig. 6(a); (c) histogram of Image 12(a).

We can introduce another equation that normalizes the calculation of local brightness differences according to the highest and the lowest local brightness differences computed not on the whole images, but only on the neighborhood under study.

Thus

$$B_V^K(x,y) = 1 - \frac{\left|\log \mu_{I_\omega}^K(x,y) - \log \mu_{J_\omega}^K(x,y)\right|}{\log L_{\max,V}^K - \log L_{\min,V}^K} \tag{19}$$

represents another measure of local brightness differences when $K = 0$, where

$$L_{\max,V}^K = \max\left\{f_I^K(x',y'), f_J^K(x',y') \middle| (x',y') \in V(x,y)\right\},$$
$$L_{\min,V}^K = \min\left\{f_I^K(x',y'), f_J^K(x',y') \middle| (x',y') \in V(x,y)\right\}.$$

Thus the B_V^K values are better distributed in the interval [0, 1] and the brightness contrasts are better discriminated [see Figs. 12(a) and 12(b), in contrast with Figs. 6(a) and 12(c)]. It is equivalent to giving a greater weight to the component of local brightness differences than to the other components.

A More Important Weight for Dispersion Differences. In the same way, we could provide a more important weight to local dispersion differences by using the following formula:

$$\mathrm{cor}(x,y) = \frac{\mathrm{cov}_\omega(x,y)}{\sigma_{I_\omega}(x,y) \cdot \sigma_{J_\omega}(x,y)}, \tag{20}$$

which measures local dispersion differences according to the variance of the color distribution of Images I and J and according to norm L_2.

Thus

$$\mathrm{cor}(x,y) = \frac{\sum_{K=1}^{3} \mathrm{cov}_\omega^K(x,y)}{\sqrt{\sum_{K=1}^{3}\left[\sigma_{I_\omega}^K(x,y)\right]^2} \cdot \sqrt{\sum_{K=1}^{3}\left[\sigma_{J_\omega}^K(x,y)\right]^2}}, \tag{21}$$

because with norm L_2

$$\left[\sigma_{I_\omega}(x,y)\right]^2 = \sum_i \sum_j \omega(i,j) \cdot \sum_{K=1}^{3}\left[f_I^K(x+i,y+j) - \mu_{I_\omega}^K(x,y)\right]^2 = \sum_{K=1}^{3}\left(\sigma_{I_\omega}^K\right)^2. \tag{22}$$

In the same way

$$\mathrm{cov}_\omega(x,y) = \sum_i \sum_j \omega(i,j) \cdot \sum_{K=1}^{3}\left[f_I^K(x+i,y+j) - \mu_{I_\omega}^K(x,y)\right] \bullet \left[f_J^K(x+i,y+j) - \mu_{J_\omega}^K(x,y)\right] = \sum_{K=1}^{3} \mathrm{cov}_\omega^K(x,y). \tag{23}$$

We must keep in mind that even if norm L_2 gives more relevant results in terms of difference measurements, it is

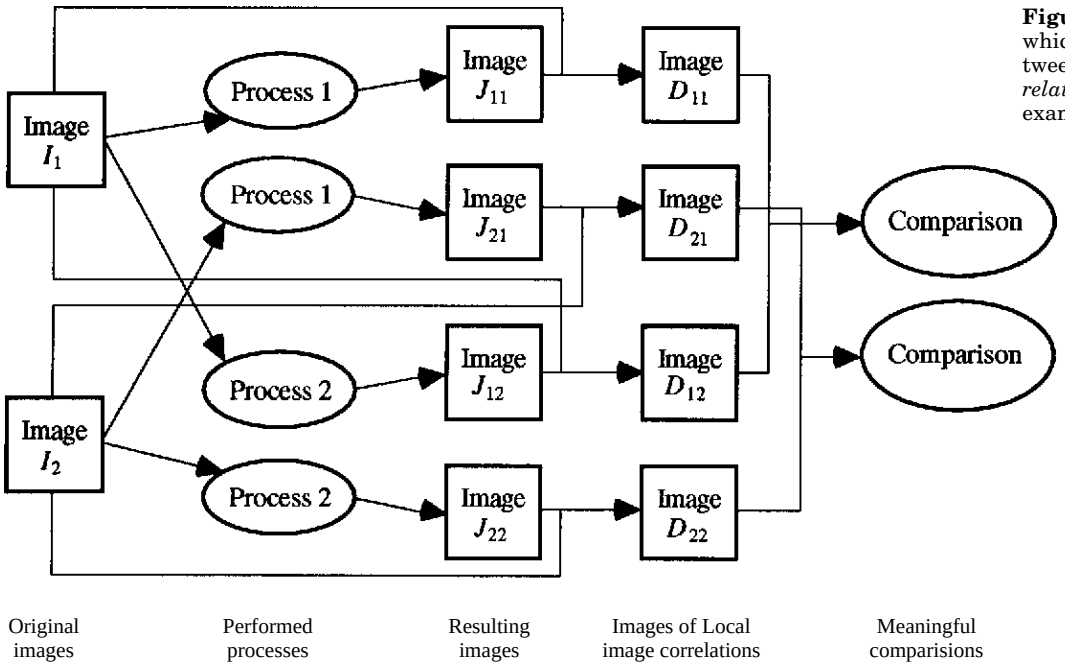


Figure 13. Cases of study for which relative comparison between images of *local image correlations* make sense (Refer, for example, to Figs. 1 and 11).

less adapted to characterize color dispersion. Indeed, considering that no order relation exists between colors, it does not make any sense to define an average color value for a set of colors when colors involved do not have the same hue.

According to the L metrics used before, some distribution features can be more robust and more relevant. It is then essential to choose the best metric among L_1, L_2, or L_∞ metrics to improve the quality of descriptors.[15]

In this section we have seen that the measure under study should amplify the weight of one criterion over the others. This is required by the fact that the cost of errors is not identical for each image analysis problem. For this reason the proposed measure has properties that can be adjustable to a specific image analysis problem, as advised by Yasnoff, Miu, and Bacus.[2]

Discussion

The representation of local image correlations is locally defined (i.e., pixel by pixel) from two entry images, one of which is used as a reference image. This representation emphasizes brightness, dispersion, and emergence differences that appear between the entry images according to the local distribution of colors. This representation can then be used only for a relative comparison between images provided the reference image is the same in all cases of study (see Fig. 13).

For example, images provided by Process 1 and Process 2 are compared with the same original image. The segmented image shown in Fig. 1(b) is compared with the image *Peppers* shown in Fig. 1(a), as is the segmented image of Fig. 1(c). Results of these comparisons are represented, respectively, in Figs. 11(a) and 11(b). As these two images of comparison refer to the same original image, they can be compared with each other. For example, it will be interesting to know for which segmentation process the most noticeable differences appear and if pixels highlighted by differences are clustered in the image plane.

To compare one image of *local image correlations* with another, we can use global descriptors such as those we present in the following section.

Definition of Descriptors for a Global Comparison. Generally to compare two image distributions we refer to their histograms. For example, to compare Figs. 11(a) and 11(b), we can refer to histograms given in Figs. 14(a) and 14(b). Nevertheless, as we can see in these two figures, histograms are not relevant enough to define an absolute descriptor of similarity between images with sufficient accuracy.

Indeed, in our cases of study, we are faced with relatively similar histograms. It is then absolutely necessary to use robust and accurate features to characterize small differences that appear between histograms.

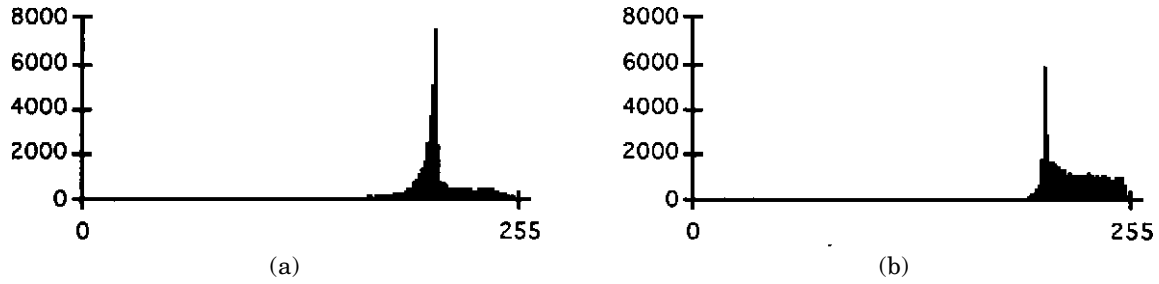


Figure 14. (a) Histogram of the image of *local image correlations* shown in Fig. 11(a); (b) histogram of the image of *local image correlations* shown in Fig. 11(b).

Among descriptors we can use, we suggest a measure based on both numbers of pixels having high-local-image correlations and of pixels having low-local-image correlations. This measure must estimate the number of poorly correlated pixels relative to the number of correlated pixels. It can be defined on the basis of the two following criteria.

Let

$$H_D = \left\{ (x,y) \in P \middle| D(x,y) \geq (1 - r_h) \right\} \tag{24}$$

be the set of pixels for which *local image correlation* is higher than $(1 - r_h)$, and let

$$L_D = \left\{ (x,y) \in P \middle| D(x,y) < r_l \right\} \tag{25}$$

be the set of pixels for which *local image correlation* is lower than r_l.

Then

$$R_{H_D} = \frac{\text{card}\, H_D}{\text{card}\, D - \text{card}\, H_D} \quad \text{and} \quad R_{L_D} = \frac{\text{card}\, L_D}{\text{card}\, D - \text{card}\, L_D} \tag{26}$$

represent, respectively, the proportion of pixels of high correlation and the proportion of pixels of low correlation in Image D under study. These two complementary proportions can be jointly used to derive a global descriptor that quantizes color image differences.

For example, $r_h = r_l = 10\%$, $R_{H_D} = 9.8\%$ and $R_{L_D} = 0\%$ for the histogram corresponding to Fig. 11(a). In the same way, $R_{H_D} = 46.3\%$ and $R_{L_D} = 0\%$ for the histogram of Fig. 11(b). According to these criteria, Process 2 seems to give better results than does Process 1. Consequently, from a signal processing point of view, Process 2 introduces less local distortion in the image than does Process 1. These criteria can be improved by filtering the image with processes like neighborhood averaging to smooth noise and to decrease details in the image, such as isolated pixels or small size areas.

Nevertheless, the criteria R_{H_D} and R_{L_D} alone do not accurately reflect the amount of disagreement between the original image and the one studied.[6] Moreover, they do not agree with human visual observation.

The disadvantage of histogram-based methods is the loss of spatial information in computing the histogram.[1] That leads us to introduce other distribution features more relevant to the spatial distribution of data in an image, as recommended by Yasnoff, Miu, and Bacus.[2] Moreover, it has been shown that methods based on histograms are not robust to aspects such as slight shift, for example, unless we consider cumulative histograms with distribution features as central moments.[15]

Until now we have taken into account only local differences that appear between one original image and one test image, without according any importance to the spatial distribution of colors into these images. Human observation is more sensitive to distortions observed in large homogeneous areas than to distortions appearing in textured areas or along contours.[16] Even if the three criteria previously defined have been evaluated as proportional to the local degree of homogeneity (from the variance of neighborhood pixels), they do not accurately reflect perceptual disturbances. This is due to the fact that, first, they do not take into account region features such as size, spatial density, or surface area, and, second, they do not distinguish textured areas and contour areas.

From a local point of view it does not make any sense to take into account region size, for example. That can be done only from a global analysis of the spatial distribution of image elements. To distinguish textured areas from homogeneous areas, we can use a local homogeneous criterion, such as those defined by Balasubramanian and Allebach[16] or by Trémeau, Calonnier, and Laget.[17] Next we can compute the degree of importance of these areas and weight measures previously defined according to the selected criterion.

Actually, when we analyze the image of local correlations in a global way, we are more interested in detection of poorly correlated elements than well-correlated elements and in analysis of spatial distribution of poorly correlated elements relative to the spatial distribution of well-correlated elements.

Consequently, we first propose to threshold the image of local image correlations in three classes according to two values r_h and r_l (see Eqs. 24 and 25), and we next propose to analyze the image provided by this segmentation in terms of spatial density, surface area, and surface size. After this step we compute one or several descriptors of image comparison.

A More or Less Important Weight for Edge Differences. It is also important to distinguish textured areas from contour elements. To achieve this distinction we first compute color edges from each image under study. Then we compute color edge differences according to the following formula (see Fig. 15):

$$C(x,y) = 1 - \frac{\left| \text{Edge}_I(x,y) - \text{Edge}_J(x,y) \right|}{\text{Edge}_I(x,y) + \text{Edge}_J(x,y)}, \tag{27}$$

where $\text{Edge}_I(x,y)$ corresponds to the color edge measure defined by Cumani and computed on Image I.[18]

Next we can define a mask operator M, which discriminates local image correlations linked to homogeneous or textured areas from local image correlations linked to edge elements. The mask operator M has been empirically defined from a thresholding process of color edge elements as follows:

$$\begin{aligned} &M(x,y) = 1 \ \text{ if } \ \text{Edge}_I(x,y) \ \text{ or } \ \text{Edge}_J(x,y) \geq r_e \cdot \text{Edge}_{\max}, \\ &M(x,y) = 0 \ \text{ otherwise,} \end{aligned} \tag{28}$$

with

$$\text{Edge}_{\max} = \max\left\{ \text{Edge}_I(x,y), \text{Edge}_J(x,y) \middle| (x,y) \in P \right\}$$

and where $r_e = 0.02$ in reference to Cumani's scheme.[18] It has been applied to the local image correlation measure D to introduce a new measure D_C stated as follows:

$$\begin{aligned} D_C(x,y) &= D(x,y) \quad \text{if} \quad M(x,y) = 0, \\ D_C(x,y) &= C(x,y) \quad \text{otherwise.} \end{aligned} \tag{29}$$

Figure 16 shows results provided by this measure D_C, which takes into account color edge elements.

In the context of image analysis it seems more important to underline the presence of a contour than to analyze its magnitude. Consequently, we suggest using Measure D for the general case of study. Indeed we can justify this by the fact that in visual observation, edge attributes focus the attention more than contrast attributes linked to the magnitudes of these edges.

Inversely, in the context of image comparison, it is not only the presence of contours between two images that is important, but also the differences of color edge magnitudes.

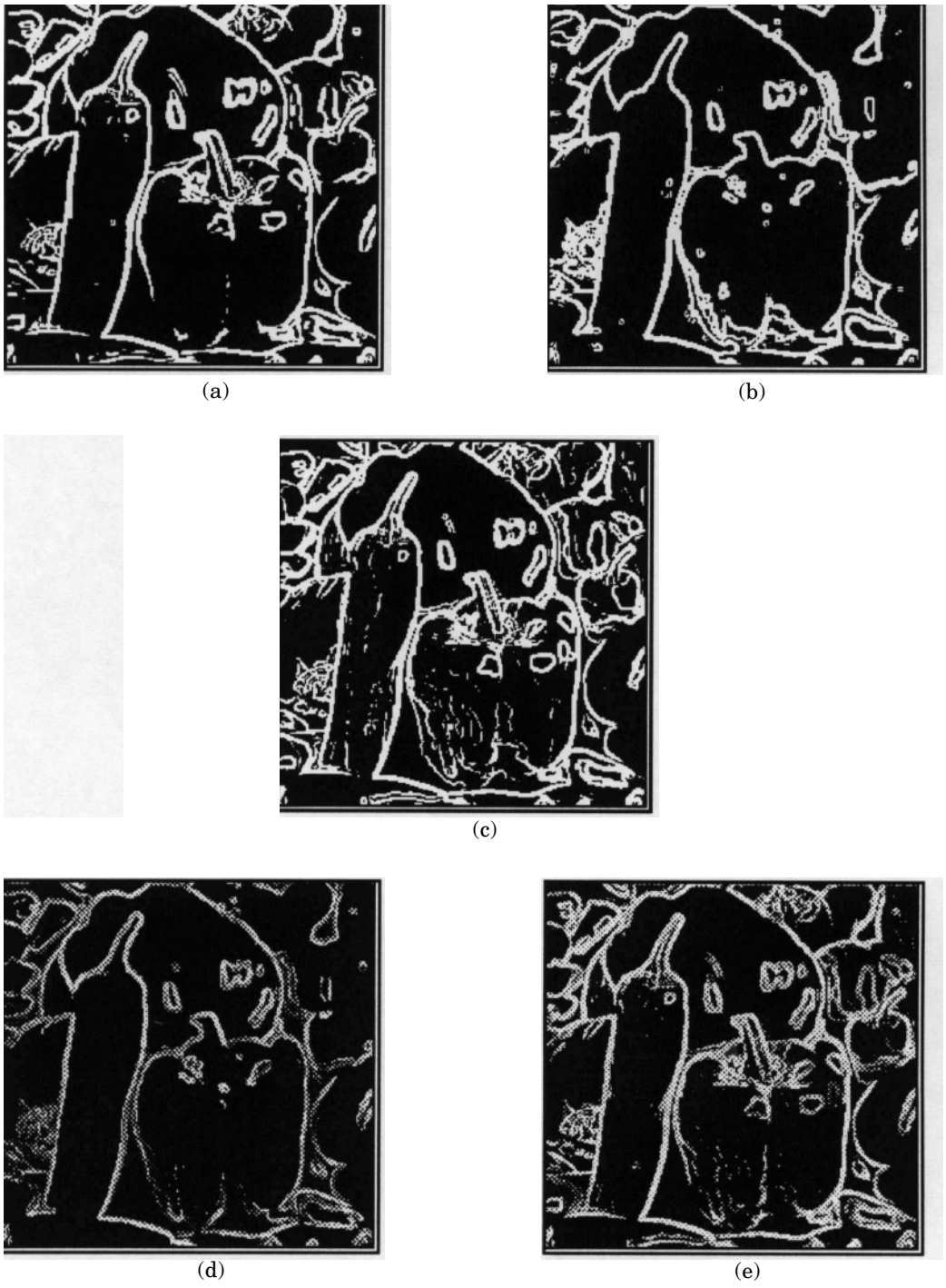


Figure 15. (a) Edge elements detected in Fig. 1(a); (b) edge elements detected in Fig. 1(b); (c) edge elements detected in Fig. 1(c); (d) color edge differences between Fig. 15(a) and Fig. 15(b); (e) color edge differences between Fig. 15(a) and Fig. 15(c). In Figs. 15(d) and 15(e), white pixels correspond to color edge differences equal to zero.

As an illustration, compare Figs. 11(a) and 16(a) [also 11(b) and 16(b)] with Figs. 1(a) and 1(b) [and 1(a) and 1(c)]. From a general point of view it seems that contours are relatively well preserved by the segmentation process, i.e., on these elements the local correlation measure is maximal, as we can see in Fig. 11. Inversely a more precise analysis shows that contours have been more perceptibly modified than the first measure points out, i.e., on these elements the local correlation measure can be relatively small, as we see in Fig. 16. This last measure also shows that if color edge differences are no more important than other color differences, color edges may not appear on the image of local image correlations.

Conclusion

The measure of color image differences proposed in this report is based on visual observation. This measure enables us to evaluate results of a process or to compare different processes in terms of image quality. The measure presents two characteristics: it can be displayed as an image of local image disturbances, and it can be computed

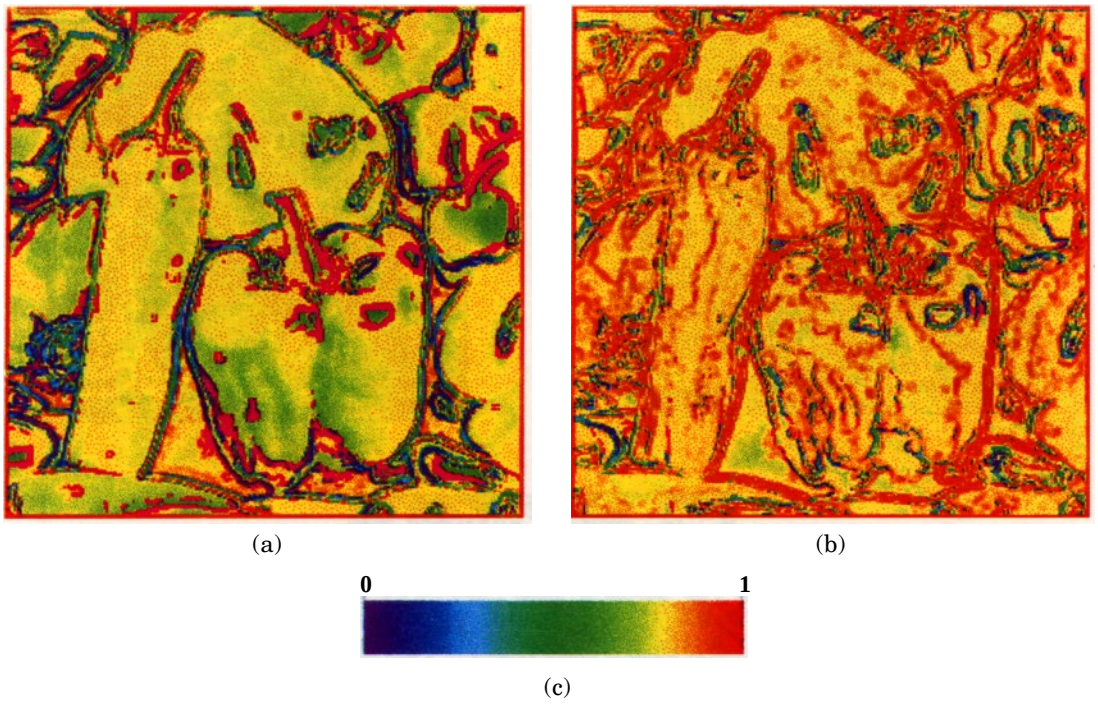


Figure 16. (a) Image of local image correlations between Figs. 1(a) and 1(b); (b) image of local image correlations between Figs. 1(a) and 1(c). These images have been computed according to measure D_C defined by Eq. 29.

as a quantitative global attribute of image differences. Thus, better than the mere measure, it allows us to analyze and to understand why a process gives good or bad results on an area of the image under study. Indeed, measuring performance of a process is a key factor in providing a feedback path by which a system can modify its strategy during processing to improve its result.

This measure does not require any a priori knowledge about the reference image nor about the test process. It results from a good trade-off between various attributes involved in image difference observation to reflect accurately the amount of disturbance that appears between one reference image and one test image. Even if this measure has been developed in terms of independent process features, it can be adjusted component by component to perform better according to processes or images under study.

In this report we have essentially dealt with local color image differences without giving too much detail about the global measure that can be computed from these differences. It will be interesting to extend the field of study to this point of view. Such an investigation is included in our current research area. We also plan to test the accuracy of this global measure with different kinds of reference images and different processes. In that way it will be interesting to extend this field of study to different color spaces to test their performance accurately in terms of image quality. In earlier works this testing has been done without taking into account parameters linked to visual observation.

References

1. M. D. Levine and A. M. Nazif, Dynamic measurement of computer generated image segmentations, *IEEE Trans. Patt. Anal. Mach. Intell.* **7**:155 (1985).
2. W. A. Yasnoff, J. K. Miu, and J. W. Bacus, Error measures for scene segmentation, *Patt. Recog.* **9**: 217 (1977).
3. L. M. Delves, R. Wilkinson, C. J. Oliver, and R. G. White, Comparing the performance of SAR image segmentation algorithms, *Int. J. Remote Sens.* **13**: 2121 (1992).
4. P. G. J. Barten, Evaluation of subjective image quality with the square-root integral method, *J. Opt. Soc. Amer. A* **7**: 2024 (1990).
5. A. Inoue and J. Tajima, Adaptive quality improvement method for color images, *Proc. SPIE* **2179**:429 (1994).
6. J. Liu and Y. H. Yang, Multiresolution color image segmentation, *IEEE Trans. Patt. Anal. Mach. Intell.* **16**: 689 (1994).
7. A. Trémeau, V. Pugnet, E. Dinet, and B. Laget, A local color correlation measure for color image comparison, *Proceeding of IS&T/SID 3rd Color Imaging Conference*: 119 (1995).
8. R. E. Jacobson, An evaluation of image quality metrics, *Proc. of the Royal Photographic Society Symposium, London*: **7** (1994).
9. F. X. Lukas and Z. L. Budrikis, Picture quality prediction based on a visual model, *IEEE Trans. Comm.* **COM-30**: 1679 (1982).
10. E. Dinet, Contribution de modéles de vision humaine à la définition de méthodes multirésolutions en vision artificielle, Thèse de l'université Jean Monnét, Saint-Etienne, 1994.
11. A. Rosenfeld, Connectivity in digital pictures, *J. Asso. Comput. Mach.* **17**: 146 (1970).
12. P. J. Burt, Fast filter transforms for image processing, *Comput. Graph. Image Process.* **16**: 20 (1981).
13. P. J. Burt, Attention mechanisms for vision in a dynamic world, *Proc. 9th ICPR*: 977 (1988).
14. G. H. Graham et al., *Vision and visual perception*, New York, Wiley & Sons,1965.
15. M. Stricker and M. Orengo, Similarity of color images, *Proc. SPIE* **2420**: 381 (1995).
16. R. Balasubramanian and J. Allebach, A new approach to palette selection for color images, *J. Imaging Technol.* **17**: 284 (1991).
17. A. Trémeau, M. Calonnier, and B. Laget, Color quantization error in terms of perceived image quality, *Proc. IEEE Conf. on Acoustics, Speech and Signal Processing* **5-I**: 93 (1994).
18. A. Cumani, Edge detection in multispectral image, *CVGIP: Graph. Models Image Process.* **53**: 40 (1991).