

# Pair correlation integral for fractal characterization of three-dimensional histograms from color images

Julien CHAUVEAU, David ROUSSEAU, and François CHAPEAU-BLONDEAU

Laboratoire d'Ingénierie des Systèmes Automatisés (LISA),  
Université d'Angers, 62 avenue Notre Dame du Lac, 49000 Angers, France.

**Abstract.** The pair correlation integral is used to assess the intrinsic dimensionality of the three-dimensional histogram of RGB color images. For application in the bounded colorimetric cube, this correlation measure is first calibrated on color histograms of reference constructed with integer dimensionality. The measure is then applied to natural color images. The results show that their color histogram tends to display a self-similar structure with noninteger fractal dimension. Such a fractal organization in the colorimetric space can have relevance for image segmentation or classification, or other areas of color image processing.

**Key words:** Color image, Color histogram, Fractal, Pair correlation integral, Feature extraction and analysis.

## 1 Introduction

In image processing, histograms of the pixel values are useful to many purposes [1]. For instance, they can serve for the characterization and correction of image acquisition, or for segmentation operations, or for indexing in image data bases. For gray-level or monocomponent images, the histogram is a simple one-dimensional data structure, straightforward to visualize and to handle. However, multicomponent images, including multispectral images, are becoming more and more pervasive. The histogram of a multicomponent image, accordingly, is a multidimensional data structure, and multidimensional data can exhibit complex organization [2]. For multicomponent images, the approach which is very often followed in current image processing, is to consider separately the monodimensional marginal histogram of each component. This leads to simple processing, but which may lose important information contained in the dependence between components. The full multidimensional histogram contains richer information, but it is more complicated to handle, and little is known concerning the structure of the multidimensional histogram of multicomponent images, starting with RGB color images. In the present paper, we are going to show that the histograms of natural RGB color images tend to exhibit nontrivial multiscale organization, with self-similarity across scales, and a fractal character. This seeks to contribute in the direction of better knowledge and more effective exploitation of multidimensional histograms of images.

Other fractal properties have already been reported for images, but these essentially concerned the spatial organization of images. Especially, for natural images, inherent multiscale structures and details existing in natural scenes, tend to induce fractal properties detectable in the spatial organization of the images. This is for instance conveyed by self-similar power-law evolutions observed for the spatial-frequency spectra and the associated spatial correlation functions of many natural images [3–7]. The fractal properties we are going to report here are distinct. They are not observed in the spatial organization of the images, but in their colorimetric organization, and these are new observations. The colors that are present in an image, and the frequency of their occurrences, tend also to distribute over the colorimetric space in a fractal, self-similar, way.

## 2 Pair correlation measures

We consider an RGB color image with  $N$  pixels. Each one of the three ( $R, G, B$ ) components can vary among  $Q$  possible values, from 0 to  $Q - 1$ . In the colorimetric RGB space, each pixel, indexed by  $n$ , maps into a triplet defining a point  $\mathbf{X}_n = (R, G, B)$ , for  $n = 1$  to  $N$ . The  $N$  points  $\mathbf{X}_n$  which distribute among the  $Q^3$  colorimetric cells form the three-dimensional histogram we will be considering for the image. For a characterization of the histogram organization in RGB space, we evaluate the number of pairs of points which are separated by a distance  $\leq r$ . It is defined as

$$C_2(r) = \sum_{n=1}^N \sum_{n'>n}^N \Gamma(r - \|\mathbf{X}_n - \mathbf{X}_{n'}\|), \quad (1)$$

with the Heaviside function  $\Gamma(u) = 1$  if  $u \geq 0$  and  $\Gamma(u) = 0$  otherwise. We can consider  $C_1(r) = C_2(r)/N$  which is the average number of pairs per pixel of the image. This  $C_1(r)$  reflects in the RGB space, for a typical pixel, the average number of neighbors within a distance  $\leq r$  in the three-dimensional histogram. The largest possible pixel separation in the colorimetric cube  $[0, Q - 1]^3$  is  $r_{\max} = \sqrt{3}(Q - 1)$  and accordingly, at  $r = r_{\max}$  the value of  $C_2(r)$  saturates at  $N(N - 1)/2$  which is the total number of pairs. Normalization by this total number of pairs yields  $C_0(r) = C_2(r)/(N(N - 1)/2)$ . This  $C_0(r)$  is a non-decreasing function which starts close to zero at  $r = 0$  and saturates at 1 at large  $r$  approaching  $r_{\max}$ . For points in space, such a function  $C_0(r)$  is known as the correlation integral [8, 9]. It is a measure which is used to characterize the organization of sets of points in space. For points uniformly scattered in the three-dimensional space, the pair correlation integral  $C_0(r)$ , or any of its non-normalized versions  $C_1(r)$  or  $C_2(r)$ , would vary as  $r^3$ . For more complex sets of points, a remarkable feature is to register a power-law variation as  $r^D$  with a possibly noninteger exponent  $D$  differing from 3. Such a behavior characterizes a nontrivial self-similar organization of the points. Self-similarity arises from invariance of the power-law under a change of scale: if the scale in  $r$  changes as  $r \rightarrow ar$ , a corresponding change of scale in  $C_0$  as  $C_0 \rightarrow \alpha^D C_0$  makes the whole

structure look similar. The organization of the points in space, as characterized by the correlations  $C_0(r)$  versus  $r$ , has no characteristic scale and is self-similar across scales. This confers to it a fractal character [10, 11].

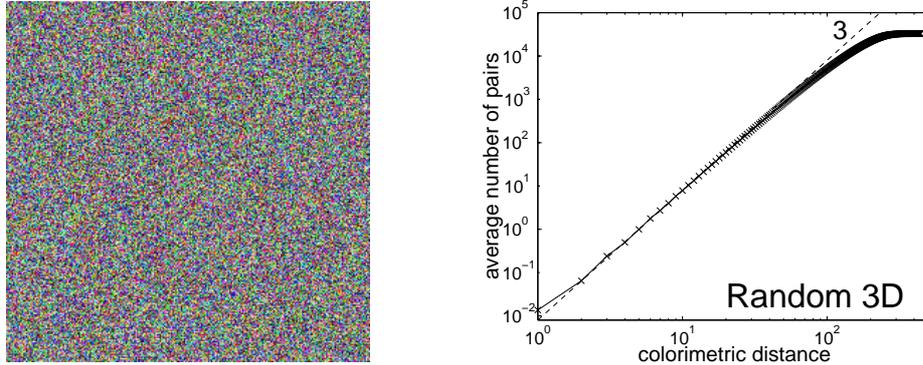
Pair correlation measures similar to  $C_0(r)$  or  $C_1(r)$  or  $C_2(r)$ , have been used to characterize fractal structures in various processes, for instance in chaotic attractors, or in porous media, or in diffusion-limited aggregates, or in percolation clusters, or in the distribution of stars and galaxies [8, 10, 11]. Here, we shall use the measure  $C_1(r)$  to exhibit a fractal organization in three-dimensional histograms of color images.

### 3 Calibration of the measure for images

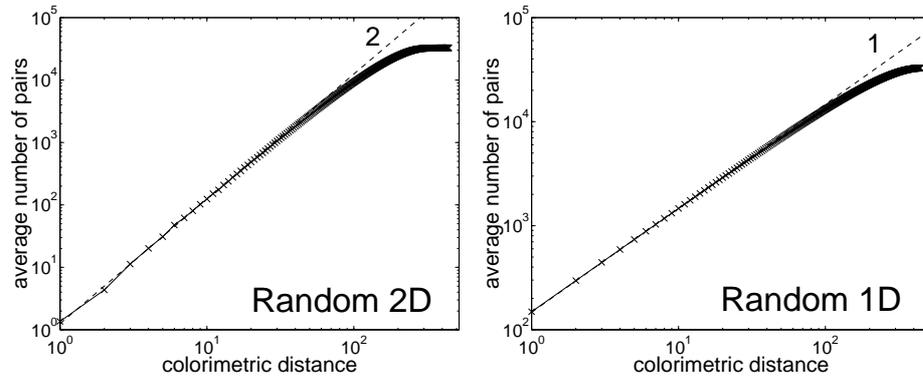
We want to use the average number of pairs  $C_1(r) = C_2(r)/N$  from Eq. (1), in order to characterize the distribution of points of the histogram in the colorimetric cube  $[0, Q - 1]^3$ . First, for calibration purpose, it is appropriate to test the evolution of the correlation measure  $C_1(r)$  on reference images with known statistical properties. This is especially useful because, when applied to images, there are no simple a priori models to theoretically predict the evolutions of  $C_1(r)$ . Even in the simple case, evoked above, of points uniformly scattered in the three-dimensional space, the simple prediction  $C_1(r) \propto r^3$  is valid only in a homogeneous unbounded space. But since the colorimetric cube  $[0, Q - 1]^3$  has well defined boundaries, when it is uniformly filled with points one may expect a deviation from  $C_1(r) \propto r^3$ , due to boundary effects. Accordingly, one may expect the behavior  $C_1(r) \propto r^3$  to be valid at small  $r$  and to gradually disrupt as  $r$  approaches the linear size  $Q - 1$  of the colorimetric cube.

For calibration of the correlation measure  $C_1(r) = C_2(r)/N$  from Eq. (1), we start with the simple random image with size  $N = 256 \times 256$  pixels, for which the color components  $R$ ,  $G$  and  $B$  at each pixel are randomly picked in  $[0, Q - 1 = 255]$  with uniform probability. A resulting realization of such a random image is shown in Fig. 1. The corresponding pair correlation measure  $C_1(r)$  has been evaluated and is also presented in Fig. 1.

The log-log plot of Fig. 1 shows an evolution as  $C_1(r) \propto r^3$  which is well verified, except at large and small  $r$ . At large distance  $r$  approaching  $r_{\max} = \sqrt{3}(Q - 1)$ , there is a saturation of the count  $C_1(r)$  which tends to grow more slowly than  $r^3$  due to the boundary effect. At small  $r$  close to  $r = 1$ , the count  $C_1(r)$  which operates on the discrete grid of the colors in the histogram, deviates from the model  $C_1(r) \propto r^3$  which assumes a continuous distribution of points. At small  $r$ , the average number of neighbors on the discrete grid differs slightly from the volume of the sphere of radius  $r$ , and produces fluctuations of the count around a straight line behavior in log-log coordinates. These fluctuations become rapidly negligible as  $r$  increases above 1, as visible in Fig. 1. Also at small  $r$ , the statistical estimation is less accurate, due to the low average density of neighbors at small  $r$ , with the random uniform model of Fig. 1. There is a total of  $N = 2^{16}$  pixels in the image, for  $2^{24}$  colorimetric cells in the histogram; this leads to a



**Fig. 1.** Left: RGB color image with size  $256 \times 256$  pixels and three random  $R$ ,  $G$  and  $B$  components uniform in  $[0, Q - 1 = 255]$ . Right: Average number of pairs  $C_1(r)$  in its three-dimensional histogram, as a function of the colorimetric distance  $r$ ; the dashed line has slope 3.



**Fig. 2.** Average number of pairs  $C_1(r)$  in the color histogram, as a function of the colorimetric distance  $r$ , when the histogram lies in a plane (left), or on a line (right), in the colorimetric cube  $[0, Q - 1 = 255]^3$ . The dashed lines have slope 2 and 1.

(low) uniform density of  $2^{-8}$  point per colorimetric cell of side 1, yielding the order of magnitude of  $10^{-2}$  for the average number of neighboring points at  $r = 1$  in Fig. 1. Yet, within these limits at large and small  $r$ , the evolution as  $C_1(r) \propto r^3$  is well identified in the results of Fig. 1, and provides a consistent characterization of the uniform color histogram as a three-dimensional manifold.

We also tested the correlation measure  $C_1(r)$  on another random test image whose color histogram is deliberately chosen as a two-dimensional manifold. This is achieved with a random image with size  $N = 256 \times 256$  pixels, for which the components  $R$  and  $G$  at each pixel are randomly picked in  $[0, Q - 1 = 255]$  with uniform probability; then the remaining component is constructed as  $B = (R + G)/2$ . This ensures that, in the colorimetric cube  $[0, Q - 1]^3$ , the histogram

of the random image lies in a plane: the principal diagonal plane with a vertex at  $(0, 0, 0)$  and a vertex at  $(Q - 1, Q - 1, Q - 1)$  in the colorimetric cube  $[0, Q - 1]^3$ . The corresponding pair correlation measure  $C_1(r)$  is presented in Fig. 2(left).

The log-log plot of Fig. 2(left) shows an evolution of the pair correlation measure as  $C_1(r) \propto r^2$ , within the limits at large and small  $r$ . This again provides a consistent characterization of the uniform color histogram as a two-dimensional manifold here.

Next, we have tested  $C_1(r)$  on a histogram constructed as a one-dimensional manifold. For the histogram,  $N = 256 \times 256$  points are distributed at random with uniform density on the principal diagonal line joining vertices  $(0, 0, 0)$  and  $(Q - 1, Q - 1, Q - 1)$  of the colorimetric cube  $[0, Q - 1 = 255]^3$ . This corresponds to images with random uniform components  $R$ ,  $G$  and  $B$  constrained by  $R = G = B$ . The log-log plot of Fig. 2(right) shows an evolution of the pair correlation measure as  $C_1(r) \propto r^1$ , providing, as before, a consistent characterization of the uniform color histogram as a one-dimensional manifold here.

Together, the results of Figs. 1–2 demonstrate the ability of the correlation measure  $C_1(r)$  to provide a consistent identification of the dimension of the manifold formed by the color histogram in the colorimetric cube. These results, obtained with random test images, will serve to us as reference in order to interpret the behavior of  $C_1(r)$  measured on natural color images.

## 4 Characterization of natural color images

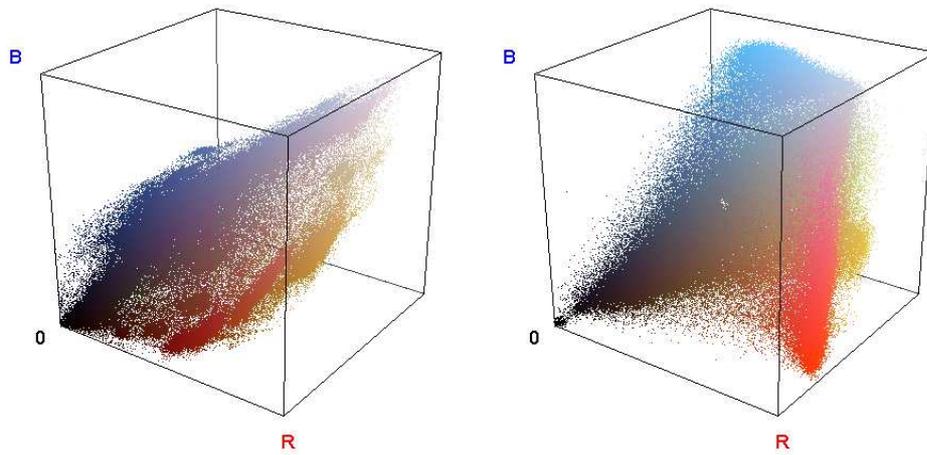
We have considered various common RGB color images, with size  $N = 256 \times 256$  pixels and  $Q = 256$  levels, and represented in Fig. 3. For illustration, two three-dimensional color histograms are depicted in Fig. 4.

For each of the color images of Fig. 3, we have computed the average number of pairs  $C_1(r) = C_2(r)/N$  from Eq. (1), in order to characterize the distribution of points of the histogram in the colorimetric cube  $[0, Q - 1]^3$ . The resulting evolutions of  $C_1(r)$  are reported in Fig. 5.

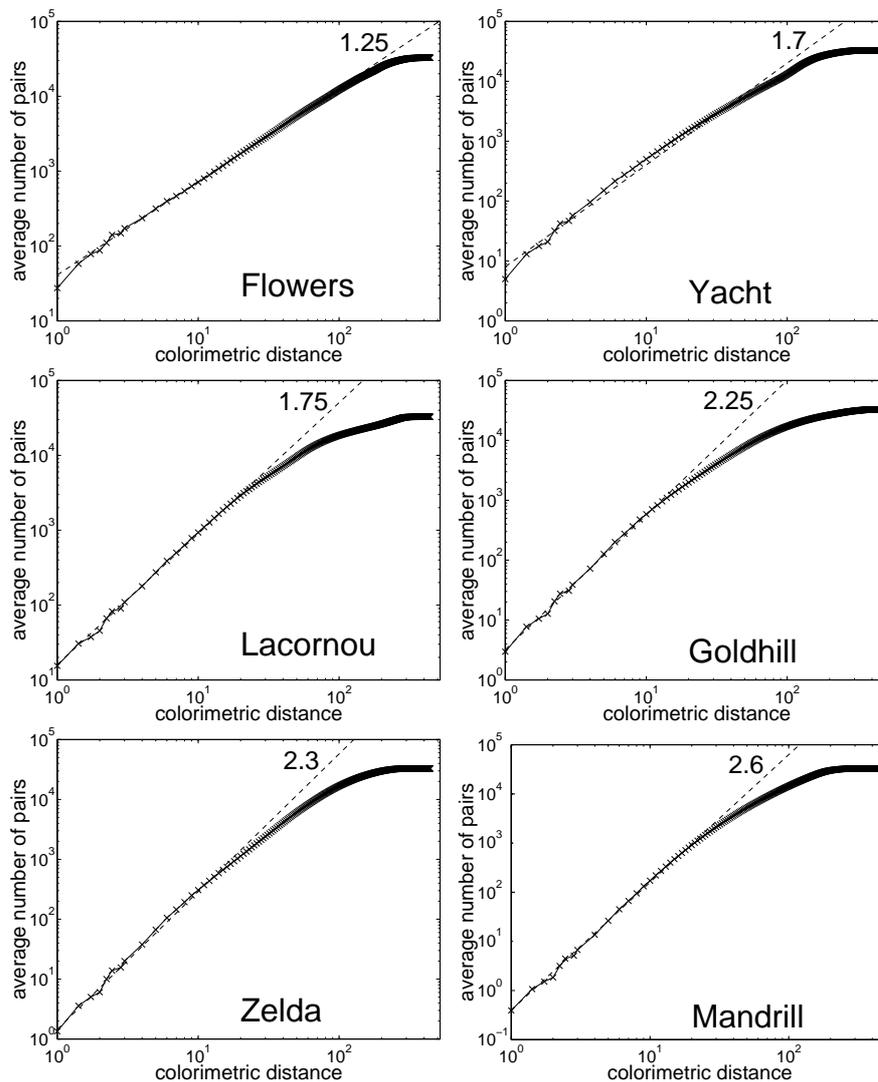
For each log-log plot in Fig. 5, a dashed line with slope  $D$  is indicated which provides a fit to the data. This corresponds for the pair correlation measure  $C_1(r)$  to a power-law model as  $C_1(r) \propto r^D$ . The linear fits in Fig. 5 best apply at intermediate range of  $r$  where the practical implementation of the correlation integral is known to be best suited for dimensionality estimation [9]. In Fig. 5, noninteger values are observed for the power-law exponent  $D$ . Such power-law behavior of the pair correlation integral, with noninteger exponent  $D$ , identifies a nontrivial self-similar organization characterizing a fractal structure [9–11]. For the three-dimensional color histograms, this reveals an organization with no characteristic scale. The points in the histogram tend to form clusters with no characteristic size, but clusters with many sizes occurring in a self-similar way, and assessed by a noninteger fractal dimension  $D$ . In association, voids with no points may be expected to exist at all scales in the fractal histogram.



**Fig. 3.** Six RGB color images with size  $256 \times 256$  pixels, and  $Q = 256$  levels.



**Fig. 4.** Color histogram in the RGB colorimetric cube  $[0, Q - 1 = 255]^3$  for image “Yacht” (left) and image “Mandrill” (right) from Fig. 3.



**Fig. 5.** Average number of pairs  $C_1(r)$  in the color histogram, as a function of the colorimetric distance  $r$ , for the six RGB color images of Fig. 3. The slope is indicated of the dashed line fitted to the data.

## 5 Discussion

In this study, we have used the pair correlation integral estimated via Eq. (1), in order to assess the intrinsic dimensionality of the three-dimensional histogram of RGB color images. For application in the bounded colorimetric cube, the

correlation measure has been first calibrated on color histograms of reference constructed with integer dimensionality. The pair correlation integral has then been evaluated on natural color images. Our observation is that for natural images with sufficient colorimetric variability, the three-dimensional histogram tends to display a self-similar structure with a noninteger fractal dimension which is specific to a given image. The registration of possible fractal structures in the color distribution of natural images is a new feature. The preliminary results presented here open a way in this direction. Additional observations and analyses are needed, in order to confirm the colorimetric fractal properties on more extensive surveys of natural images, and in order to appreciate the conditions of existence and the possible origins of them. Other approaches or estimation methods can also be used for alternative characterizations of fractal colorimetric properties in images. We currently have complementary studies under way in this direction.

We again emphasize that the present fractal properties observed in the color distribution of natural images are different from, but possibly related to, other fractal properties previously reported for natural images in their spatial organization [3–7]. Briefly stated, previous results [3–7] dealt with fractal distribution of the pixels in space, while the present results deal with fractal distribution of the pixels in the colorimetric cube.

The registration of possible fractal structures in the colorimetric organization of natural images may have several significant consequences for color image processing. Estimating the intrinsic dimensionality of data sets, as done here for color images, is helpful to determine the possibility of new coordinate systems allowing more parsimonious representation. Also, a self-similar fractal organization with no characteristic scales, may have bearing for various processing performed with the color histograms: for instance, when implementing segmentation methods defining classes by looking for peaks in the color histogram, or when implementing subquantization of the color components or color reduction operations. In addition, the fractal dimension  $D$  may serve as a useful index for various image processing tasks, such as image characterization, classification or indexing. All these issues related to a possible fractal organization of the colors, may form interesting areas to explore for color image processing [12] and enlarge the applications of fractal geometry to imaging [13].

**Acknowledgement:** Julien CHAUVEAU acknowledges support from *La Communauté d'Agglomération du Choletais*, France.

## References

1. Russ, J.C.: *The Image Processing Handbook*. CRC Press, Boca Raton (1995).
2. Landgrebe, D.: Hyperspectral image data analysis. *IEEE Signal Processing Magazine* **19**(1) (Jan. 2002) 17–28.
3. Burton, G.J., Moorhead, I.R.: Color and spatial structure in natural scenes. *Applied Optics* **26** (1987) 157–170.

4. Ruderman, D.L., Bialek, W.: Statistics of natural images: Scaling in the woods. *Physical Review Letters* **73** (1994) 814–817.
5. Ruderman, D.L.: Origins of scaling in natural images. *Vision Research* **37** (1997) 3385–3398.
6. Turiel, A., Parga, N., Ruderman, D.L., Cronin, T.W.: Multiscaling and information content of natural color images. *Physical Review E* **62** (2000) 1138–1148.
7. Hsiao, W.H., Millane, R.P.: Effects of occlusion, edges, and scaling on the power spectra of natural images. *Journal of the Optical Society of America A* **22** (2005) 1789–1797.
8. Grassberger, P., Procaccia, I.: Characterization of strange attractors. *Physical Review Letters* **50** (1983) 346–349.
9. Camastra, F.: Data dimensionality estimation methods: A survey. *Pattern Recognition* **36** (2003) 2945–2954.
10. Mandelbrot, B.B.: *The Fractal Geometry of Nature*. Freeman, San Francisco (1983).
11. Gouyet, J.F.: *Physics and Fractal Structures*. Springer, Berlin (1996).
12. Sharma Ed., G.: *Digital Color Imaging Handbook*. CRC Press, Boca Raton (2003).
13. Turner, M.J., Andrews, P.R., Blackledge, J.M.: *Fractal Geometry in Digital Imaging*. Academic Press, New York (1998).