



**Repositorio Institucional de la Universidad Autónoma de Madrid**

<https://repositorio.uam.es>

Esta es la **versión de autor** de la comunicación de congreso publicada en:  
This is an **author produced version** of a paper published in:

CompSysTech '03: Proceedings of the 4th international conference on Computer systems and technologies: e-Learning. New York: ACM, 2003. 242-249

**DOI:** <http://dx.doi.org/10.1145/973620.973661>

**Copyright:** © 2003 ACM

El acceso a la versión del editor puede requerir la suscripción del recurso  
Access to the published version may require subscription

## Statistics of Natural Images Using Hash Fractal Image Compression

Kostadin Koroutchev and Jose R. Dorronsoro

**Abstract:** *Natural images form very small subset of all images. In spite of this fact, the direct computation of their block densities is not possible. On the other hand, the existence of various successful image compression methods, in particularly, the fractal compression, indicates that the compression somehow is able to capture and use at least part of the natural image statistics. In this work we show how hash based fractal image compression can be used to derive quite precise the entropies of  $4 \times 4$  patches of the natural images. We state that the probability density in first order factorize to the probability densities of the contrast, the brightness and the index of the codebook blocks.*

**Key words:** *Fractal image compression, hash fractal code algorithm, natural image.*

### INTRODUCTION

The success of the loose compression methods for natural images shows that essentially their statistics are far from uniform. The compression ratios are of order of one hundred without significant visible lost of the quality of image. That shows that, at least the information relevant to the human visual system in the natural images can be represented as very sparse subset of all possible images.

In fact, natural image information is not distributed uniformly over the image: there are parts that are most relevant to human visual system, while other parts are far less relevant.

Given the natural assumption that the purpose of the visual system is to extract those parts with maximal information, a very important question is how to measure this information. Notice that this question is not only relevant for basic human visual processing but also for a number of everyday information processing systems such as for instance efficient image compression.

The natural approach to that information measurement is to determine first the particular statistics that natural images should possess and a large effort has been undertaken in that direction [5,2].

However, it is quite easy to see that direct natural image statistics will be quite difficult to come by. For instance, it can be seen that using entropy lossless compression, one can achieve in the ideal case a compression rate of about 2.8 bits per pixel [7]. Thus, about 45 bits are needed to compress a  $4 \times 4$  pixel square, which implies that one needs many times  $2^{45}$   $4 \times 4$  blocks to compute reasonable block frequency values, something that is evidently not possible. This has led to the study of several simplified image models, whose statistics are easier to compute and hence to use. Examples of this are locally averaged gray levels, luminance values or gradients, as for all of them, significant histograms can be computed over individual images. Although giving useful information, it is clear that gray levels and luminance or even gradient-based edge power statistics represent a first level, coarse grain approach to natural image statistics. However, other sources of, at least implicit, natural image statistics are concrete image processing techniques. A particular relevant example is Fractal Image Compression (FIC).

Moreover, efficient fractal compression essentially depends on the successful capture of some kind of image statistics. Let us briefly review it.

In Fractal Image Compression (FIC) [1,6] an  $N \times N$  gray level image  $I$  to be compressed is partitioned using a quadtree structure into a set  $R$  of blocks of  $K \times K$  pixels called ranges and for any one  $R$  of them, a same size image block  $D$ , taken from a certain codebook set,  $D$  is chosen so that  $R$  is similar to  $D$  in some sense. Typical  $K$  values are 16, 8 or 4, while  $N$  usually is 512. To obtain the  $D$  set, baseline FIC first decimates  $I$  by

averaging every  $2 \times 2$  block and then somehow selects a set of possibly overlapping  $K \times K$  blocks from the decimated image. This basic domain set is finally enlarged with their blocks' isometric images to arrive at the final domain set  $D$ . Now, in baseline FIC, to compress a given range  $R_0$ , a domain  $D_0$  is obtained such that

$$D_0 = \arg \min_D \min_{s,o} \|R - (sD + o1)\|^2 \quad (1)$$

The component  $sD + o1$  above is the gray level transformation of  $D_0$ , with  $s$  the contrast factor and the offset  $o$  the luminance shifting.  $R_0$  is then compressed essentially by the triplet  $D_0, s_0, o_0$ , with  $s_0$  and  $o_0$  the minimizing contrast and luminance parameters. The decoding process recovers an approximation to the range  $R_0$  by simply iterating the above linear transformation upon the matching domain  $D_0$ . As FIC proceeds by matching each range with the closest domain from the same overall image, it is intuitively clear that efficient fractal compression relies on some kind of implicit statistical analysis of the image to be compressed. This fact has been made explicit in hash based FIC, a novel image compression procedure introduced by the authors [3]. Hash based FIC, quickly reviewed in the next section, first organizes domains in a certain hash table and then performs for a given range  $R$  a simplified full range-domain block comparison only with the domains in a single "hash bin", whose index depends on  $R$ . As shown in [3], hash based FIC achieves quality and compression rates comparable with those given by state of the art FIC methods with a much better time performance particularly in the low image quality range. Given the basic statistic nature of hash tables, it is clear that a key ingredient for this performance must somehow lie in an efficient capture of block image statistics. In particular, hash FIC suggests a procedure to derive mean and variance normalized block statistics. We shall show in the next section how to compute these block statistics, after which new image entropy, that we shall call Normalized Block (NB) entropy, will be presented. Moreover, we shall also introduce a NB image representation that visually incorporates an image's normalized block densities. In the third section we shall illustrate its computation on the familiar Lena image and draw the conclusions.

### Normalized block image statistics

The starting point of hash FIC is the fact that the pixel approximation  $r_{ij} \approx d_{ij} + o$  derived from (1) also gives range-domain mean  $\langle r \rangle \approx s \langle d \rangle + o$  and variance  $\sigma(r) \approx |s| \sigma(d)$  approximations, that in turn, translates the approximation parameter estimation in (1) to the following normalized block comparison

$$\tilde{r}_{ij} = \frac{r_{ij} - \langle r \rangle}{\sigma(r)} \approx \frac{d_{ij} - \langle d \rangle}{\sigma(d)} = d_{ij} \quad (2)$$

This suggests to perform for a given range  $R$  the full block comparisons in (1) only for those domains "close" to  $R$  according to (2). Hash FIC uses for this [3] the block hash function  $h(B)$  defined by

$$h(B) = \sum_{h=2}^H \left( \left\lfloor \frac{b_{i(h)j(h)} - \langle b \rangle}{\lambda \sigma(b)h} + B \right\rfloor \% C \right) C^{h-1} = \sum_{h=1}^H b_h C^{h-1}, \quad (3)$$

where  $B$  is a centering parameter,  $C$  is taken to cover the range of the fractions in (3) and  $\lambda$  controls the spread of the argument of the floor function so that it approximately has a uniform behavior. Typical values for these parameters in 8 bit gray level images are  $H=5$  (corresponding to the choice of the 4 block corners plus its center),  $B = 8$ ,  $C=16$  and  $\lambda = 0.4$  for  $4 \times 4$  domains (the ones we shall use in our illustration). They are chosen in such a way that (3) gives a base  $C$  expansion of  $h(B)$  for which the expansion coefficients are approximately uniformly distributed.

Notice that, by the above discussion, for fractal image compression (and even vector quantization codebook compression) to be successful, it must somehow capture the statistics of normalized blocks  $\tilde{B}$ . Moreover, it suggests a simple procedure to estimate normalized block frequencies  $p(\tilde{B})$  over a relatively small set of natural images. More precisely, assume we have fixed a certain set  $\mathbf{S}$  of natural images. Then we can take the blocks of a given image  $I$  as a source of domain blocks for the compression of the images in  $\mathbf{S}$  and proceed to the FIC compression of all the images in  $\mathbf{S}$  using these domains. In this setting the "natural" density of a normalized domain block  $\tilde{B}$  can be measured by the number of times its hash bin is used to compress ranges from the set  $\mathbf{S}$ , divided by the total number of hash bins. Flat blocks, that is, those with variance less than 2, are excluded from the above, as they mainly are amplifications of the high frequency noise that may be present in  $I$ .

There are two ways of exploiting the densities  $p(\tilde{B})$ . As a first one, observe that, since a full block  $B$  is obviously equivalent to its  $(\tilde{B}, \bar{B}, \sigma(B))$  representation,  $p(\tilde{B})$  could also be seen as one of the components of what could be the full block density  $p(B) = p(\tilde{B}, \bar{B}, \sigma(B))$  and could thus be taken as a first step to obtain the "true" natural image block densities  $p(B)$ . Further results in that direction will be discussed elsewhere. On the other hand,  $p(\tilde{B})$  can be used to define a new "normalized block" (NB) image entropy  $H_{NB}(I)$  over  $I$  as

$$H_{NB} = E_{NB} \left[ \log \frac{1}{p(\tilde{B})} \right] = - \sum p(\tilde{B}) \log p(\tilde{B}).$$

Furthermore, there is a simple procedure to visualize the  $p(\tilde{B})$  values over  $I$ . To do so, assume that each block is indexed as  $B_{ij}$  according to the position in  $I$  of its upper left corner. Then, once they have been adequately normalized, the densities  $p(\tilde{B}_{ij})$  can be projected back to  $I$  as "block" luminance  $\tilde{s}_{ij}$ , resulting in what we may call again the normalized block (NB) representation  $I_{NB}$  of  $I$ . In the next section we shall use both  $E_{NB}(I)$  and  $I_{NB}$  to compare this block normalized image information with other information measures over  $I$ .

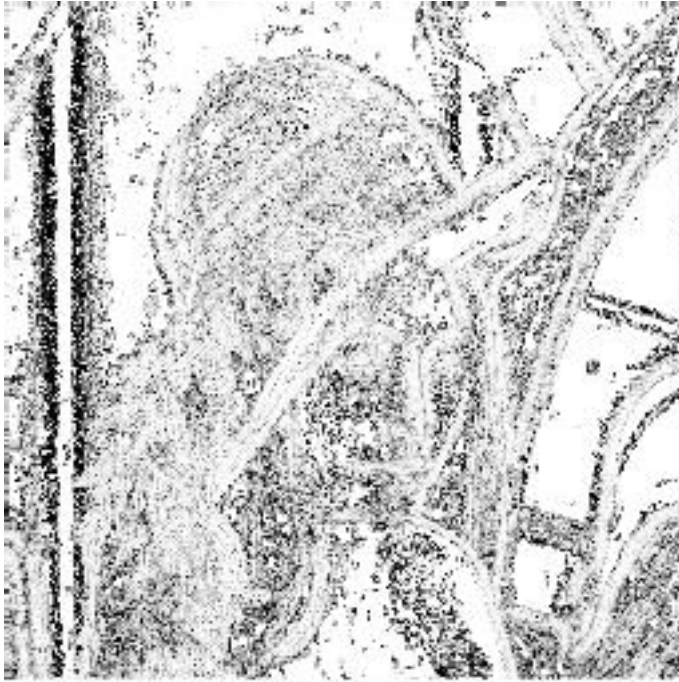


Fig. 1 The information due to the normalized block.



Fig.2 The information due to the variation of the blocks. This is the main structural information component.

We shall illustrate the just defined block entropy and image representation over the well-known Lena image, for which we shall compute its NB statistics. To do so, we shall use all  $4 \times 4$  blocks from a  $256 \times 256$  decimation of Lena as a domain source, using for each block its upper left corner coordinate as its index in its NB representation. As the natural image set  $\mathbf{S}$  to be used to derive the block statistics  $p(\tilde{B})$  we shall work with one set of 2600  $1024 \times 1024$  central cuts of pictures taken from the image server <http://hlab.phys.rug.nl/archive.html>. These images were first used in [2]. All of them have been compressed using hash FIC against Lena's domains. As mentioned before, flat

blocks were dropped, bringing the initial domain number of about 64.000 blocks down to 49.329 non-flat blocks.

In order to get the NB representation of the Lena image, the interval count from 0 to 2000 was linearly mapped to the 8 bit gray level interval from 0 to 255, with a cutoff to 255 for all counts above 2000.

Figure 1 shows the NB representation of Lena, while figure 2 gives a gray level variance representation of each block. White blocks correspond to 0 variance, while black blocks correspond to the maximal block variance of 110. Recall that a long standing hypothesis in image processing [4] is that maximal gradient or variance blocks (that is, the image edges) carry most of an image's information.

## Results

We estimate the following quantities: The total entropy of the images per block,  $H(x, y, s, \sigma, \bar{B})$ , the entropy of the normalized blocks  $H(x, y, s)$  bits and the entropy of the contrast and brightness  $H(\sigma, \bar{B})$ . From these entropies we derive the mutual information  $I(x, y, s \parallel \sigma, \bar{B})$  that is 2.2503 bits. Also we can measure in similar way the mutual information  $I(\sigma \parallel \bar{B})$  that is 0.2140 bits. The mutual information  $I(x, y, s \parallel \sigma, \bar{B})$  is only 9.6% of the total entropy, and the mutual information  $I(\sigma \parallel \bar{B})$  is only 2% of the entropy  $H(\sigma \parallel \bar{B})$ . The results are shown in the first column of Table 1. These quantities show that in first order of approximation, the probability distribution of 4 x 4 blocks can be represented as a product of 3 terms: the probability of the normalized block, the probability of the standard deviation of the block and the probability of the gray level of the block. That is:

$$p(x, y, s, \sigma, b) \approx p(x, y, s) p(\sigma) p(b)$$

Quantity	Value[bits]	Value[bits]
$N$	91533069	45996977
$\log_2 N$	26.45	25.45
$H(x, y, s, \sigma, \bar{B})$	23.45	23.14
$H(x, y, s)$	15.36	15.36
$H(\sigma, \bar{B})$	10.34	10.33
$I(x, y, s \parallel \sigma, \bar{B})$	2.25	2.56
$I(\sigma, \bar{B})$	0.21	0.22
$I / H_{total}$	9.6%	11.1%
$I(\sigma \parallel \bar{B}) / H(\sigma, \bar{B})$	2.07%	2.10%

Table 1. See the text.

Further insight in the NB representation and entropy can be gained from figures 1 and 2. First observe that NB representation is not uniform and has a non trivial structure. For instance, the most informative blocks seem to be those corresponding to vertical

boundaries. This can be due to the fact that the Lena image has practically no horizontal edges while the natural images compressed with Lena's domains do have them, and have to be compressed using rotations of Lena's vertical edge domains. Notice also that very complex areas do not seem to carry a lot of information. For instance, the "tails" of the hat have a weaker representation than, say, Lena's face. In other words, simple "shape" areas carry more NB information. Furthermore, recall that any block  $B$  can be directly reconstructed from its normalization  $\bar{B}$  and its variance  $\sigma(B)^2$  and mean  $\bar{B}$ . Of these three components, the mean  $\bar{B}$  carries rather little information and it is thus intuitively clear that the NB and variance representations somehow decompose the standard luminance representation, something that is also visually in agreement with figures of Lena and 1. Moreover the NB and variance representations are quite independent, something shown by the very small (4.6 %) correlation.



Fig.3 Dependency between the normalized blocks and the variance of the image. It is clear that the main dependence is concentrated near the edges.

#### Robustness:

The statistics performed over 2600 images give some  $91000000 = 2^{26.44}$  blocks. Therefore, the entropies would be reliable only if the entropies measured are significantly stable and bellow  $2^{26}$ . The stability of the entropies can be shown, using the half of the images and performing the same statistics on these images. If the total entropies are not more or less the same than the estimation of the information is not stable and the size of the sample must be increased. The results over the half of the sample shows that the entropy measures are indeed stable and that the mutual information  $H(x,y,s||\sigma,b)$  is not an artifact of the lower number of the samples.

The stable estimation of the mutual information shows that actually there is some weak statistical dependence between the "profile" of the block, represented by  $x,y,s$  and its contrast and brightness, that deserve future investigation. A preliminary idea of the type of the dependence can be given by plotting the quantity

$< p(x, y, s, \sigma, \bar{B}) / (p(x, y, s) p(\sigma, \bar{B})) >_{s, \bar{B}}$ . The result is shown in figure 3. The dependence is concentrated among the blocks with high contrast. There are some indications that one bit of that information can be spurious, because the range of values that can be used in high contrast block is limited.

## **CONCLUSIONS AND FUTURE WORK**

In this work we have shown how hash based fractal image compression can be used to derive density function estimates for mean-variance normalized image blocks. We have applied these densities to compute a novel normalized block image entropy measure and to derive an associated image representation, which we have compared with other similar image information measures. NB entropy and representation seem to be one of the two components in a decomposition of an image's gray level representation, the other being edge information.

Future work will proceed along two different lines. On the one hand, the properties of the NB entropy and representation shall be further studied. The exact dependence of the interrelation between the normalized block and the contrast and brightness level can be investigated.

Some results valid about blocks of size 3 x 3 does not hold on blocks 4 x 4, for example the blocks 4 x 4 do not form simple boundaries figures. It is promising to make statistics on blocks of size 6x6 in order to find what are the exact properties of the blocks in order to carry a lot of information.

## **Acknowledgments**

This work was partially supported by Spain's CICYT, TIC 01-572 and CAM 02-18. We thank to Antonio Turiel for fruitful discussions.

## **REFERENCES**

- [1] Y. Fisher, *Fractal Image Compression--Theory and Application*. New York: Springer--Verlag 1994.
- [2] J.H. van Hateren and A. van der Schaaf, "Independent component filters of natural images compared with simple cells in primary visual cortex." *Proc.R.Soc.Lond. B*, 265:359-366, 1998.
- [3] K. Koroutchev and J. Dorronsoro, "Hash-like Fractal Image Compression with Linear Execution Time", *Iberian Conference on Pattern Recognition and Image Analysis, IbPRIA 2003*.
- [4] D. Marr, *Vision*, W.H. Freeman and Co., NY 1982
- [5] D.L. Ruderman, "The statistics of natural images", *Network: Computation in Neural Systems*, vol. 5, 517-548, 1994.
- [6] D. Saupe, R. Hamzaoui, and H. Hartenstein, "Fractal image compression: an introductory overview," in: *Fractal Models for Image Synthesis, Encoding, and Analysis*, D. Saupe and J. Hart (eds.), SIGGRAPH '96, New Orleans, 1996.



[7] M. Weinberger, G. Seroussi, G. Sapiro, ``LOCO-I: A Low Complexity, Context-Based, Lossless Image Compression Algorithm," Proc. of the IEEE Data Compression Conference, Snowbird, Utah, March-April 1996.

#### **ABOUT THE AUTHORS**

Universidad Autonoma de Madrid, {kostadin.koruchev,,jose.dorransoro}@ii.uam.es.