

THE 2D-DCT COEFFICIENT STATISTICAL BEHAVIOUR: A COMPARATIVE ANALYSIS ON DIFFERENT TYPES OF IMAGE SEQUENCES

M. Mitrea^{a,b,*}, F. Prêteux^b, A. Vlad^{a,c}, C. Fetita^b

^aFaculty of Electronics and Telecommunications, "Politehnica" University, Bucharest–Romania

^bARTEMIS Project Unit, GET / INT, Evry–France

^cThe Research Institute for Artificial Intelligence, Romanian Academy

This paper presents a method designed to analyse the distribution of the 2D-DCT (bi-Dimensional Discrete Cosine Transform) coefficient hierarchy computed over three types of image sequences: colour video, grey-level X-ray CT (computerised tomography) images and computer simulated noise images. This method is a statistical approach which combines in a new way the following four tests: (1) the χ^2 (Chi-square) test on concordance between experimental data and a theoretical probability density function, (2) the ρ (Ro) test on correlation, (3) the Fisher F test on equality between two variances, and (4) the Student T test on equality between two means. Such an approach was compulsory so as to mathematically overcome the dependency existing among successive images in the considered sequences. The results obtained on natural sequences (either video or medical images) are compared to those corresponding to computer simulated sequences, interesting differences being pointed out and discussed. The overall results may play a central role in a large variety of image/video processing applications: compression, segmentation, retrieval, protection (*e.g.* cryptography and watermarking). For instance, we successfully applied them to the design of a new robust watermarking method for colour video.

(Received March 4, 2004; accepted March 10, 2004)

Keywords: 2D-DCT coefficient hierarchy, Statistical analysis, Colour video, X-ray CT medical images, Computer simulated noise images

1. Introduction

Finding a probability law which is able to accurately model the probability density functions -p.d.f.- characterising the values taken by the 2D-DCT (bi-Dimensional Discrete Cosine Transform) coefficients has always been a challenging research topic [1]-[5].

The transform was computed either on the whole image [1], or on blocks (*e.g.* 8×8 , 16×16 , or 32×32 pixels) [2]-[5]. All these papers considered the coefficients which correspond to certain spatial frequencies and some well known p.d.f.s: Gaussian, Laplacian, Gamma, generalised Gaussian, and Rayleigh p.d.f.s. According to the targeted application and to the spatial frequency, one of these laws was found out as being the most suitable in the coefficient modelling. The support for such a selection consisted of theoretical (*i.e.* the Central Limit Theorem) and quantitative results (the Chi-square and Kolmogorov-Smirnov tests on concordance were applied) which were validated by some experiments (*e.g.* entropic coding applications). However, the dependency existing among the pixels in an image was each time neglected; hence, there is no mathematical support for the obtained results. Moreover, the experimental data consisted each time of a few test images.

A different approach is presented in this paper. First of all, we do not consider individual images but sequences of images. Then, we are no longer interested in the p.d.f. corresponding to a

* Corresponding author: mihai.mitrea@int-evry.fr

spatial frequency but in those characterising the values taken by a certain rank in the upper part of the 2D-DCT coefficient hierarchy. This hierarchy is built-up for each image in the sequence by applying the 2D-DCT to the whole image and by sorting the corresponding coefficients into a decreasing order, see Fig. 1. We also offer a solution to rigorously handle the dependency existing: (1) among different pixels in the same image and (2) among successive images in the same sequence. The experimental data were large, consisting of three types of image sequences:

1. We considered 100 video sequences, each of them with $L=17500$ frames and with different aspect ratios. They were inhomogeneous, containing: indoor and outdoor scenes, high and low motion activity, unstable and arbitrary lighting conditions, and complex backgrounds.
2. The medical data came out from a research study on the human airways [6]. They were acquired in clinical routine at the Central Radiology Service of Pitié Salpêtrière Hospital in Paris, France with three different X-ray scanners (courtesy of Professor P. A. Grenier). The data set consisted of grey-level axial thorax CT (computerised tomography) images acquired from 55 patients suffering from various lung diseases, 200-400 images per patient.
3. For the comparison sake, we also used three types of grey-level computer simulated image sequences: (a) with uniform distribution, un-correlated; (b) Gamma distributed with 6 degrees of freedom, uncorrelated and (c) Gamma distributed with 6 degrees of freedom, correlated [7].

We emphasise that these three types of image sequences correspond to three different types of 2D random processes: *a priori*, we may expect a total lack of stationarity for the first case, a certain degree in the second case while the third case is stationary by the very way in which it was generated. Hence, the method we developed and the results we obtained can be adapted and generalised to another types of image sequences: seismic, industrial control, holography, *etc.*

Our interest in the p.d.f. characterising the upper part of the 2D-DCT coefficients came out from a study we devoted to video watermarking, [8]; hence, in the sequel we shall consider the coefficients having the r rank in the following interval: $r \in [1;192]$. That is, we consider the largest $R=192$ coefficients, except for the direct component, see Fig. 1.

The structure of this paper is the following: Section 2 presents the statistical approach we developed, Section 3 is devoted to experimental results while Section 4 stands for conclusions and perspectives.

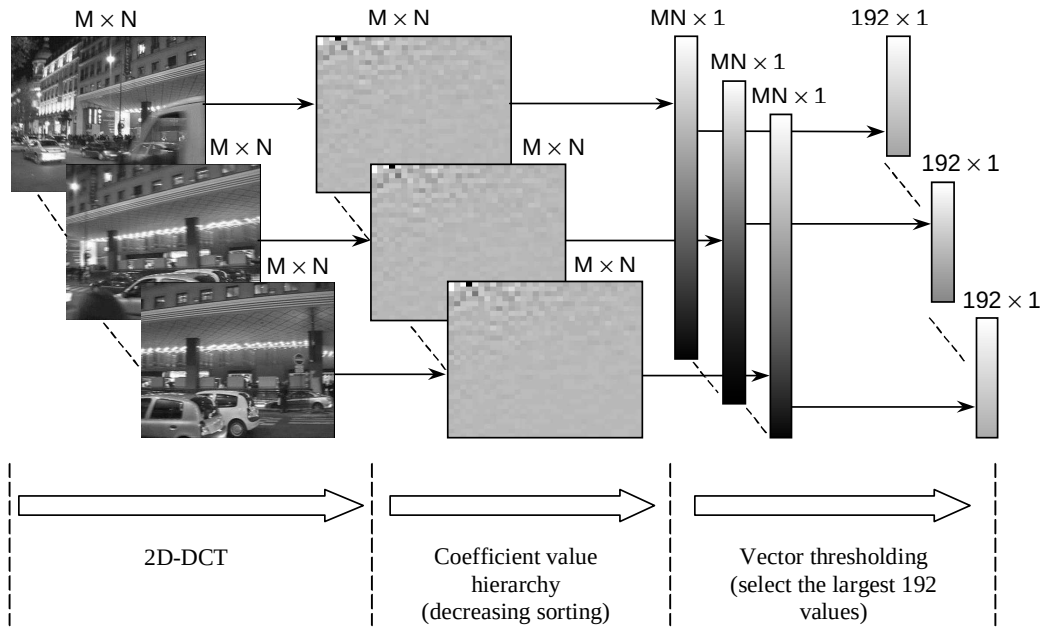


Fig. 1 Obtaining the experimental data. For each and every image in the sequence, the 2D-DCT is computed, then the coefficients are sorted into a decreasing order and the largest $R=192$ are to be statistically investigated.

2. Method presentation

We shall further present the method we developed in order to accurately determine the p.d.f. characterising the 2D-DCT coefficient hierarchy of an image sequence.

Be there an L length sequence (that is, a sequence of L images) and be $r \in [1; R]$, $R = 192$, an arbitrarily chosen rank in the 2D-DCT coefficient hierarchy. Our procedure follows as such:

Step 1: Pre-processing

For each image in the sequence:

- 1-1. compute the 2D-DCT on the whole image;
- 1-2. sort the coefficients in a decreasing order;
- 1-3. record the coefficient corresponding to the r rank.

As a consequence, L values are obtained, one for each image. Be these values denoted by $\{x_1, x_2, \dots, x_L\}$.

Step 2: Overcoming the dependency among successive images in the sequence

Partition the L values into D classes, by means of a fixed period sampling. When D is large enough (e.g. $D = 375$ for video frames or $D = 300$ for medical images), a partition into D classes with independent elements is obtained: $\{x_1, x_{D+1}, x_{2D+1}, \dots\}$ is the first class, $\{x_2, x_{D+2}, x_{2D+2}, \dots\}$ is the second class, ..., and $\{x_D, x_{2D}, x_{3D}, \dots\}$ is the last class.

Step 3: The statistical analysis

For each class in the partition:

- 3-1. verify whether the elements in the considered class are Gaussian distributed or not; this is to be carried out by means of a χ^2 statistical test [9]; when running the test, the mean value and the variance are to be estimated on the respective elements;
- 3-2. verify whether the considered D value is large enough so as to afford the independence among the elements in the same class; this is to be carried out by means of a ρ statistical test [9]; note that a ρ test is meaningful only when the χ^2 test is passed (the ρ test should be applied only on Gaussian data);
- 3-3. verify the homogeneity of the elements, from the variance point of view; this is to be carried out by means of an F statistical test on equality between two variances [10]; when running the F test, the elements are randomly split up into two sets; note that an F test is meaningful only when the χ^2 and ρ tests are passed (the F test should be applied only on independent Gaussian data);
- 3-4. verify the homogeneity of the elements, from the mean value point of view; this is to be carried out by means of a T statistical test on equality between two means [10]; here again the elements of the class are randomly split up into two sets; note that a T test is meaningful only when the χ^2 , ρ , and F tests are passed (as there is no *a priori* information about the theoretical variances, the T test should be applied only on independent Gaussian data sets which have the same variance);

We emphasise that each type of statistical tests we considered is to be applied D times, once for each partition class.

Step 4: Fusion the results obtained on each class (in order to provide information characterising the whole sequence)

- 4-1. for each type of test, compute the relative number of tests which are not passed (i.e. compute the ratio of the number of tests which are not passed to the D value).
- 4-2. verify the homogeneity among the D classes in the partition, from the variance point of view; this is to be carried out by means of an F test; $D-1$ such tests are to be run, each time on consecutive classes (for example first class vs. second class, second class vs. third class, and so on);
- 4-3. verify the homogeneity among the classes from the mean values point of view; i.e., resume Step 4-2, this time by considering a T test.

Looking back on the method we propose, several remarks should be made:

1. Step 2 was compulsory in order to overcome the dependency existing among the successive images in the sequence. Such a dependency is very strong when considering natural video and medical images. On the other hand, for the computer simulated random images we considered, this Step should be skipped over and the statistical tests directly applied.
2. Step 3 verifies whether each of the D classes in the partition confirms a Gaussian p.d.f. Moreover, as the mean values and the variances were independently estimated on each class, we have no information whether the D Gaussian p.d.f.s would be identical or not. To answer this question, we included the Steps 4-2 and 4-3 in our algorithm.
3. The homogeneity investigation carried out in Steps 3-3, 3-4, 4-2 and 4-3 is compulsory when the 2D-DCT is computed on natural video and medical images. Note that from the theoretical point of view, each image in such a sequence may be *a priori* considered as a sample of a non-stationary 2D random process. On the other hand, when considering computer simulated images, they can be well approximated as samples of an ergodic 2D random process and hence, these four Steps lose their reason.
4. In order to investigate the coefficients with the ranks in the $[1; R]$ interval, the method should be applied R times.

3. Experimental results

We first apply our method to colour video. We considered a sequence of $L = 17500$ frames (about 10 minutes). Just for illustration, Fig. 2.a displays an arbitrarily chosen frame. The frames were represented in the HSV (hue-saturation-value) system while the 2D-DCT was applied on the V component.

We sampled the sequence with a period of $D = 375$ frames (*i.e.* 15 seconds) and we ran the four types of statistical tests at an $\alpha = 0.05$ significance level. Fig. 2.b synthesises the results obtained when running the χ^2 tests, see Steps 3-1 and 4-1. The rank in the hierarchy is represented on abscissa while the ordinate corresponds to the relative number of tests which were not passed (*i.e.*, the ratio of the number of tests which were not passed to the D value). Fig. 2.b brings into evidence two intervals on the rank axis for which more than 80% of the tests are passed, namely: $6 \leq r \leq 15$ and $120 \leq r \leq 192$. In other words, the values corresponding to these two intervals do not refute the Gaussian behaviour. Fig. 2.c was built up in order to *a posteriori* validate the $D = 375$ numerical value: for practically all the ranks, more than 80% of the ρ tests -see Steps 3-2 and 4-1- were passed. This means that the $D = 375$ is a large enough value as to practically break the dependency among frames. (Note that, for comparison sake, the same axes were considered in Figs. 2.b and 2.c). As concerning the tests involved in the Steps 3-3, 3-4, 4-2 and 4-3, they were each time passed by more than 80% of the classes.

We obtained the same type of results for all the 100 video sequences we considered in our experiments.

As a second level in our experiments, we considered a sequence of medical images, see Fig. 3. The sequence has about $L = 16000$ images and we used a $D = 300$ sampling period. Fig. 3.a illustrates a torax axial CT image, where both lungs are visible. The results of Steps 3-1 and 4-1 are depicted in Fig. 3.b and those corresponding to Steps 3-2 and 4-1 in Fig. 3-c. Note that Fig. 3 shows the same kind of plots as Fig. 2. When comparing Fig. 3.b to Fig. 2.b, a larger number of ranks for which the coefficients are Gaussian distributed may be noticed. However, the two intervals determined for natural video are also confirmed by the medical data set we considered. The homogeneity investigation was again successful: more than 80% of the tests required by the Steps 3-3, 3-4, 4-2 and 4-3 were passed.

Finally, we considered computer simulated noise images. Fig. 4 illustrates the results corresponding to $L = 17500$ images obeying a 6 degrees of freedom Gamma p.d.f., with prescribed auto-correlation function [7]. This time, all the coefficients may be considered as having a Gaussian behaviour, Fig. 4.b. We also verified whether the coefficients are independent or not, by running the

ρ tests, Fig. 4.c. Note that although the 2D-DCT coefficients were computed on independent images, such an experiment was compulsory: it is known that a linear filter applied to an uncorrelated random process may induce correlation. As the computation of the 2D-DCT means, in fact, a linear filtering, the coefficients may be correlated. However, when inspecting Fig. 4.c, it can be noticed that such a correlation can be practically neglected.

We obtained the same results for the other two types of computer simulated image sequences we considered: uncorrelated uniform distributed and uncorrelated Gamma distributed.

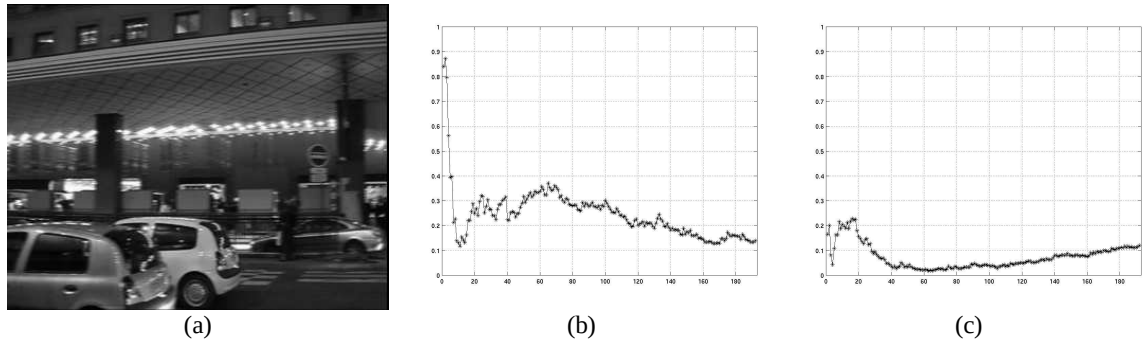


Fig. 2. Experimental results obtained on colour video: (a) sample image; b),c) the relative number of Chi-square and Ro tests which were not passed vs. the rank in the hierarchy.

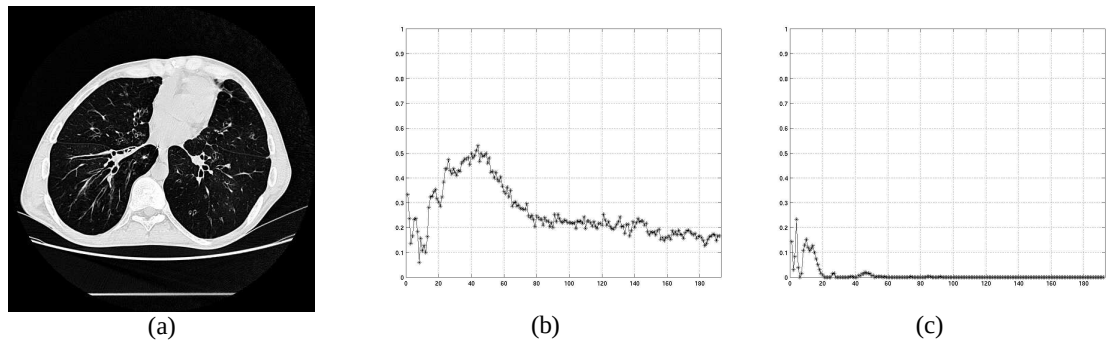


Fig. 3. Experimental results obtained on medical data set: (a) sample image; b),c) the relative number of Chi-square and Ro tests which were not passed vs. the rank in the hierarchy.

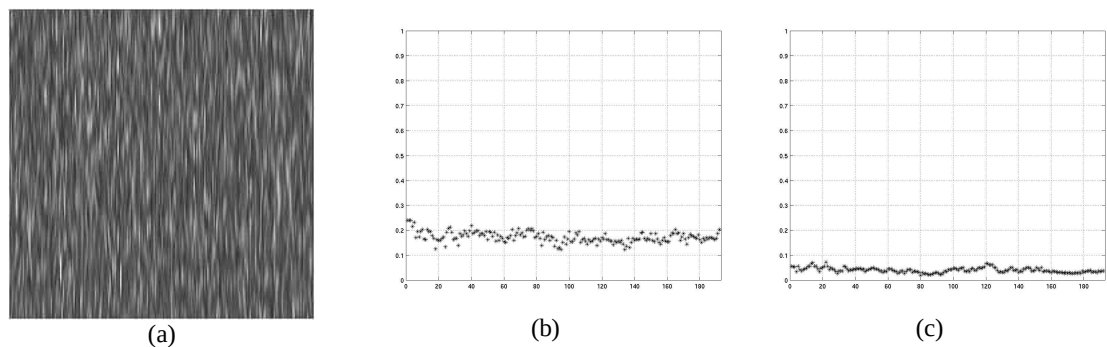


Fig. 4. Experimental results obtained on computer simulated noise images: (a) sample image; b),c) the relative number of Chi-square and Ro tests which were not passed vs. the rank in the hierarchy.

4. Discussion and final remarks

To conclude with, this paper presents a new statistical approach which is able to investigate the p.d.f. corresponding to a certain rank in the 2D-DCT coefficients hierarchy while considering multiple data sets sampled from the same natural image sequence.

When applying it to natural colour video, our method brings into evidence two intervals on the rank axis, for which more than 80% of the χ^2 tests are passed, at an $\alpha = 0.05$ significance level. Consequently, we thus validated the hypothesis according to which the values taken by the r rank, where $6 \leq r \leq 15$ or $120 \leq r \leq 192$, are Gaussian distributed. Concerning the middle ranks, $15 < r < 120$, the Gaussian assumption is not likely to hold. The first 4 ranks in the hierarchy should not be considered as having a Gaussian p.d.f.. When applying the method to medical images, the number of ranks for which the coefficients may be considered as having a Gaussian p.d.f. is larger and includes the two intervals reported for video. Here again, some intervals refute the Gaussian behaviour. As concerning the computer simulated noise images, all the three types under investigation pointed to a very nice Gaussian behaviour, for each considered rank. This may be a consequence of the fact that each image in a computer simulated sequence may be considered as a sample of an ergodic 2D random process and hence, the Central Limit Theorem is likely to hold.

We successfully used the theoretical results we obtained in this paper when designing a robust watermarking schema for colour video. In [8], we proposed an oblivious robust watermarking method for colour still image databases, which enables a 64 bit message (a logo, serial number, *etc.*) to be recovered. The mark is generated by means of an SS-CDMA (Spread Spectrum – Code Division Multiple Access) technique [8] and embedded into the largest 1024 2D-DCT coefficients computed out of the V components corresponding to a sequence of 32 images. In order to achieve a very good quality for the marked images – that is, in order to deal with a transparent watermarking method – we had to set the power of the mark to $1/16$. Despite this very small value, the method featured an excellent robustness in terms of JPEG compression, print & scan, small rotations and Stirmark attack [11]. When trying to extend the method in [8] from still images toward video, several problems arose [12]. First of all, the problem of transparency: some artefacts which are imperceptible when inspecting the individual frames are outlined in video. Hence, we were not allowed to embed our mark into the highest ranks of the 2D-DCT hierarchy. Repeated experiments pointed out that values $r > 50$ are able to provide transparency. We chose the length of the sequence $L = 1024$, the power of the mark $\sigma^2 = 1/512$, and we considered two cases: (1) the mark is embedded into the coefficients $50 \leq r \leq 113$ and (2) the mark is embedded into the coefficients $128 \leq r \leq 191$. We obtained better results concerning the robustness in the latter case. Although this may be surprising, it can be easily explained when considering the fact that the CDMA-based procedure assumes the 2D-DCT coefficients as being Gaussian distributed when recovering the mark. Hence, when considering a CDMA technique in the 2D-DCT space, the mark should be embedded into the coefficients belonging to the two intervals on the rank axis we pointed out as having Gaussian distributions.

Note that the interest of the results obtained in this paper is not restricted to the watermarking field. In fact, such results may stand for a first step in an accurate investigation of the natural image sequences with information theory tools, [13]. It can be said that there are always some ranks in the 2D-DCT hierarchy which can be accurately modelled by stationary Gaussian information sources (one information source for each rank). As far as we know, this is the first time when stationary sources modelling the natural video sequences are brought into evidence. We emphasise that these sources were connected to the hierarchy of the coefficients. Note that several other frequency-rank dependencies have been already reported in the literature, for various fields: linguistics, physics, *etc.* [14], [15], [16].

Further on, for natural images, we tried to illustrate the role of each type of coefficient in the image quality, see Fig. 5. The 2D-DCT was computed on the V component of the original image in Fig. 5.a and the coefficients were sorted. Fig. 5.b was obtained by considering only the coefficients in the $1 \leq r \leq 5$ and $15 < r < 120$ intervals, *i.e.* the non-Gaussian coefficients, and by further computing the 2D-IDCT (the InverseDCT). Fig. 5.c was obtained by considering only the coefficients in the $6 \leq r \leq 15$ and $120 \leq r \leq 192$ intervals, *i.e.* the Gaussian coefficients. In both situations we kept the H

and S components unchanged. When comparing Fig. 5.b to Fig. 5.c, the same (poor) visual quality may be noticed, although the energy corresponding to the coefficients in Fig. 5.b is about 10 times larger than the energy corresponding to the coefficients in Fig. 5.c. This may intuitively point to a larger redundancy in the non-Gaussian coefficients than in the Gaussian ones: as they bring the same visual information, there is no need for a larger energy. This idea is strengthened by the known result – *e.g.* [17] – according to which the Gaussian distribution has the most uncertainty (*i.e.* the smallest redundancy) among all distributions of a given second moment. We resumed this experiment for 100 individual still images and we obtained the same result.

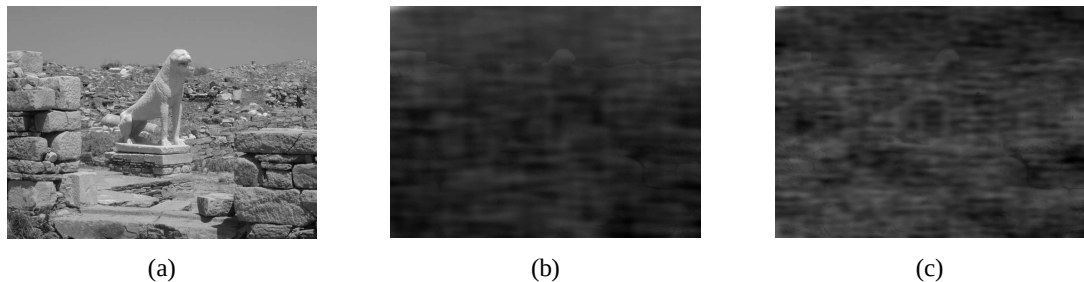


Fig. 5. The difference between Gaussian and non-Gaussian coefficients. The original image in Fig. 5.a was reconstructed either out of non-Gaussian (Fig. 5.b) or Gaussian (Fig. 5.c) coefficients. Although the non-Gaussian coefficients have a larger energy than the Gaussian ones, they lead to images with the same (poor) quality.

To conclude with, these accurate mathematical results can be useful in a very large spectrum of applications: entropic coding, image analysis and understanding, image prediction, security, *etc.* In our future work, we shall investigate how to extend the developed approach to other types of image transforms, as wavelets. Such a statistical approach will be applied to the challenging issue of video sequences indexing. The principle would be to describe the visual content as a mixture of density probability laws and to derive an associated statistical descriptor. Defining a similarity measure directly on the probability space will allow to evaluate the performances of such a new approach in terms of video content-based retrieval.

References

- [1] W. Pratt, Digital Image Processing, John Willey, 1978.
- [2] R. Reininger, J. Gibson, "Distributions of the Two-Dimensional DCT Coefficients for Images", IEEE Trans. on Communications Vol. COM-31, No. 6, 835-839 (1983).
- [3] F. Müller, "Distribution Shape of Two-Dimensional DCT Coefficients of Natural Images", Electronics Letters **29**(22), 1935-1936 (1993).
- [4] R. Joshi, T. Fischer, "Comparison of Generalized Gaussian and Laplacian Modeling in DCT Image Coding", IEEE Signal Processing Letters **2**(5), 81-82 (1995).
- [5] E. Lam, J. A. Goodman, "A Mathematical Analysis of the DCT Coefficient Distributions for Images" in IEEE Trans. on Image Processing **9**(10), 1661-1666 (2000).
- [6] P. A. Grenier, C. Beigelman-Aubry, C. Fetita, F. Prêteux, M. W. Brauner, S. Lenoir, "New frontiers in CT imaging of airway disease", Eur Radiol, **12**, 1022-1044 (2002).
- [7] A. Vlad, M. Ferecatu, P. Oprişan, "Statistical noise simulation for image processing purposes", Opt. Eng. **35**, 1590-1599 (1996).
- [8] M. Mitrea, F. Prêteux, A. Vlad, "Robust Watermarking Method for Colour Still Image Databases", The 2002 Tyrrhenian Intl. Workshop on Digital Communications – Advanced Method for Multimedia Signal Processing, Capri Italy, pp. 311-318.
- [9] R. E. Walpole, R. H. Myres, Probability and Statistics for Engineers and Scientists, 4th ed., MacMillan Publishing, 1989.
- [10] B. R. Frieden, Probability, Statistical Optics and Data Testing, Springer-Verlag, N.Y., 1983.

- [11] F. Petitcolas, R. Anderson, M. Kuhn, "Attacks on copyright marking systems" in Proc of the Second workshop on information hiding, David Aucsmith Ed., in Lecture Notes in Computer Science Vol. 1525 Portland, USA, 1998.
- [12] M. Mitrea, Françoise Prêteux, Adriana Vlad, "Spread Spectrum Colour Video Watermarking in the DCT Domain", presented at ROMOPTO 2003, Sept. 2003, Constanta-Romania, to be published in Proc. SPIE.
- [13] Al. Spataru, "Fondements de la théorie de la transmission de l'information", Presses Polytechniques Romandes, Laussane, 1987, ISBN 2-88074-133-0.
- [14] I. Kanter, D. A. Kessler, "Markov Processes: Linguistics and Zipf's Law", Physical Review Letters, **74**(22), 4559-4562 (1995).
- [15] A. Vlad, A. Mitrea, M. Mitrea, "Two frequency-rank laws for letters in printed Romanian", Procesamiento del Lenguaje Natural, Revista N° 24, Septiembre de 2000, pp. 153-160.
- [16] A. Vlad, A. Mitrea, M. Mitrea, "A Corpus – based Analysis of how Accurately Printed Romanian Obeys Some Universal Laws", Chap. 15 in "A Rainbow of Corpora: Corpus Linguistics and the Languages of the World", Wilson, Andrew/Rayson, Paul/McEnery Tony Editors, Lincom-Europa Publishing House, Munich, 2002, pp. 153-165.
- [17] A. Cohen, A. Lapidoth, "The Gaussian Watermarking Game", in IEEE Trans. on Information Theory, **48**(6), 1639-1667 (2002).