Texture Indexes and Gray Level Size Zone Matrix Application to Cell Nuclei Classification

Guillaume THIBAULT [1], Bernard FERTIL [1], Claire NAVARRO [2], Sandrine PEREIRA [2], Pierre CAU [2], Nicolas LEVY [2], Jean SEQUEIRA [1] and Jean-Luc MARI [1]

1) LSIS Laboratory, Aix-Marseille II University, France

2) INSERM UMR 910, Medical Genetic and Functional Genomic, Medical School, France

Abstract: *In this paper, we present a study on the characterization and the classification of textures. This study is performed using a set of values obtained by the computation of indexes. To obtain these indexes, we extract a set of data with two techniques: the computation of matrices which are statistical representations of the texture and the computation of "measures". These matrices and measures are subsequently used as parameters of a function bringing real or discrete values which give information about texture features. A model of texture characterization is built based on this numerical information, for example to classify textures. An application is proposed to classify cells nuclei in order to diagnose patients affected by the Progeria disease.*

***Keywords*:** texture indexes, gray level size zone matrix (GLSZM), cell nuclei classification.

1. INTRODUCTION

Pattern recognition is a major part of artificial intelligence that aims to automate the identification of typical situations. It is a major objective for many applications: handwritten character recognition (optical character recognition, automatic reading of postal letters and bank checks, etc.), video surveillance (facial recognition), medical imaging (ultrasound, CAT scan, Magnetic Resonance Imaging), etc. At the heart of the pattern recognition issue, there is a first and unavoidable step: shape characterization. Indeed, in order to recognize an object or a person, it is necessary to describe it by defining its characteristics (morphological, geometrical, textural etc.) and then to find and identify these characteristics on the digital source under investigation. For this reason, it is often helpful to distinguish two classes of characteristics: the shape using global analysis algorithms or outline analysis algorithms and the texture.

The aim of this paper is to create a model to classify culture skin fibroblast nuclei in patients affected by Progeria disease (otherwise known as the Hutchinson-Gilford syndrome). This rare disease (which affects about one hundred patients in the world) is cause by a mutation in a gene encoding lamins A and C, two proteins localized at the nuclear periphery (lamina) and within nucleoplasm [5]. Progeria patients exhibited an accelerate aging. The presence of mutated lamin A protein resulted in abnormal nuclear shape and texture (not homogeneous), evidenced by the immunodetection of lamin A/C (primary antibody directed against lamin A/C, secondary antibody coupled to Fluoresceine Iso Thio Cyanate) (see fig. 1). Digitized pictures of immunostained nuclei were sampled using a conventional epifluorescent microscope (Leica DMR) coupled to a Princeton-Roper camera. The two interest of this study are i/ to design an automatic classification of abnormal nuclei to be compared with the classification made by an expert microscopist through the analysis of more nuclei than possible by the expert; ii/ to follow up using this automated procedure the eventual nuclear changes induced by therapeutic drugs.

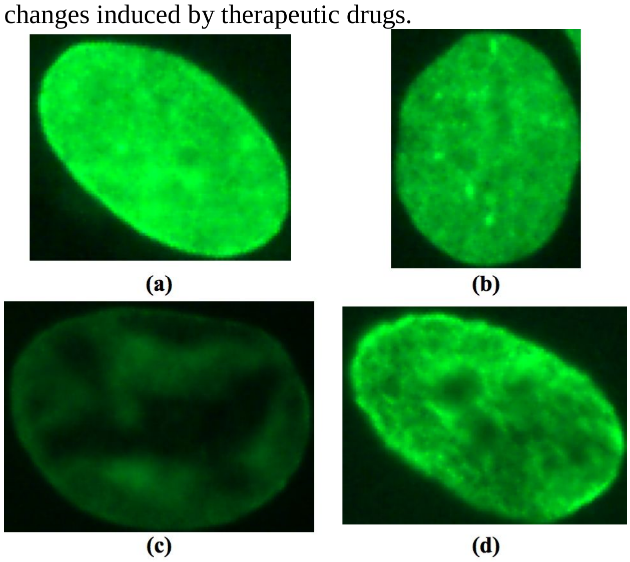

Fig.1 – Examples of nuclei highlighted with FITC : above two homogeneously textured nuclei (a and b) and below two nuclei with an inhomogeneous texture.

The first work on characterization and classification shape's of a nucleus [10, 11] made it possible to obtain a classification success rate of more than 95% thanks to the creation and use of dedicated shape indexes. The texture characterization method, on the other hand, was less satisfactory. This original approach, based on indexes initially obtained for shape characterization, was not able to achieve a success rate of more than 85%. This rate is inferior to that of expert's repeatability rate (which corresponds to the percentage of nuclei classified in the same way by an expert on two successive analysis, and is between 86 and 89%). The result of this first attempt, being inferior to the repeatability rate, we wished to improve this result and obtain one near to that given by shape classification.

OUR CONTRIBUTION

With this in mind, we present classification and validation methods used, followed by the characterization techniques used in order to improve the results of texture classification. Cooccurence matrices and Haralick indexes will be presented first. The Run Length Matrix is used and modified in order to create a novel texture homogeneity characterization method. Our contribution introduces a new method based on the construction and analysis of statistical matrices that represent the texture. All these techniques have been studied and validated by the model developed in order to solve the texture characterization issue.

2. CLASSIFICATION

The aim of classification is to attribute a class to each object being studied. As mentioned earlier, the objective of this sub-task is to determine whether a cell's nucleus has a normal (homogeneous) or an abnormal (inhomogeneous) texture.

The classification methods are said to be supervised, as they require a reference expert analysis. In this study we benefit from of biologists and geneticists knowledge who have specified classes (*healthy* and *pathological*) and subclasses (*normal* and *irregular* shapes, *homogeneous* and *non-homogeneous* textures, etc.).

A classification model is usually built using a learning method, with the help of data divided into known classes. Though applied to a specific problem, the model must be capable of being generalized (in so far as data is concerned). With this objective, the data is separated into two groups: a learning sample and a validation sample. The classifier must have the same performance rate through learning and validation. It is necessary to construct a characteristic vector for each data prior the classification phase. The vector must be relevant to the problem in order to allow accurate classification and prediction. The major risk when providing too many characteristics to the classifier is *rote learning*. The greater the vector's dimension, the greater the flexibility of the model and the better the classification, but the greater the likelihood that the model's performance will be poor for a data set not used during the validation. Each model must then be systematically validated and the best classification with the validation sample obtained. Due to the few elements in the sample at our disposal, validation was done according to the *Leave One Out* protocol [8].

The method of classification chosen for the model is *logistic regression* [7]. It is a linear model particularly well adapted to classification problems with two classes:

$P(Y/\vec{x}) = \dfrac{e^{f(\vec{x})}}{1+e^{f(\vec{x})}}$ with $\vec{x} = (x_1, \dots x_n)$ the characteristic vector of the initial data, $f(\vec{x}) = \sum_{i=1}^{n} \alpha_i x_i$ and $P(Y/\vec{x})$ the conditional probability P of the variable $\vec{x}$ to belong to the class Y. To estimate the coefficients α_i of the model, the maximum likelihood method is often used, which maximizes the probability of obtaining values observed on the learning sample. It consists in finding the parameters that optimize the likelihood function $\pounds(\alpha, Y) = P^Y [1-P]^{1-Y}$. Logistic regression is preferred over discriminate analysis [3] because of its greater reliability, versatility, the few restrictions that it imposes on the variables and the clarity of its results.

Nevertheless, other methods will also be used in order to compare results. These methods are more complex, non-linear, and are obtained through different conception techniques: neural networks [9], k-nearest neighbors [13], and random forests [14].

3. GRAY LEVEL COOCCURRENCE MATRIX

The cooccurence matrix technique is one of the oldest and most efficient methods of statistical texture representation. This method defines texture according to gray level special distribution and characterizes texture by means of second order statistics. In order to accomplish this, it is interested in the relationships that exist between the gray scale of pixels of the texture for a given displacement vector *d*. The resulting matrix is of size $N \times N$, where N is the number of gray levels in the texture. For a given displacement vector $\vec{d} = (d_x, d_y)$, an element *(i,j)* of the matrix is defined by the number of pixels in the texture that have a gray level *j* at a distance *d* from a pixel of gray level *i*. This can be written as follows:

$$M_d(i,j) = card\left\{ \begin{array}{l} ((r,s),(r+d_x, s+d_y)) / \\ I(r,s) = i, I(r+d_x, s+d_y) = j \end{array} \right\}$$

Figure 2 shows an example of cooccurence matrix calculation.

Image

1	2	3	4
1	3	4	4
3	2	2	2
4	1	4	1

=>

Gray Level (i)	Cooccurrences (j) 1	2	3	4
1	0	1	1	3
2	1	4	2	0
3	1	2	0	2
4	3	0	2	2

Fig.2 – Example of the calculation of a cooccurrence matrix for a 4x4 image with 4 gray levels with for a displacement vector (0,1).

In this study we are not only interested in the neighbors but rather the neighborhood in general. Thus, for a given spacing distance E, four matrices are calculated, one for each of the four displacement vectors $(0,E)$, (E,E), $(E,0)$ and $(-E,E)$, these are then averaged to combine all the extracted information. From the resulting reduced matrix, 15 second-order texture indexes (Haralick features [6]) are extracted allowing the characterization of the texture.

Next, a systematic study, aimed at finding the best subset of indexes, was undertaken. The best result was obtained for a subset of 8 indexes, on images reduced to 32 gray levels with a distance of one pixel. This obtains a classification success rate of 90% by logistic regression. Figure 3 illustrates the distribution of probabilities as given by the model.

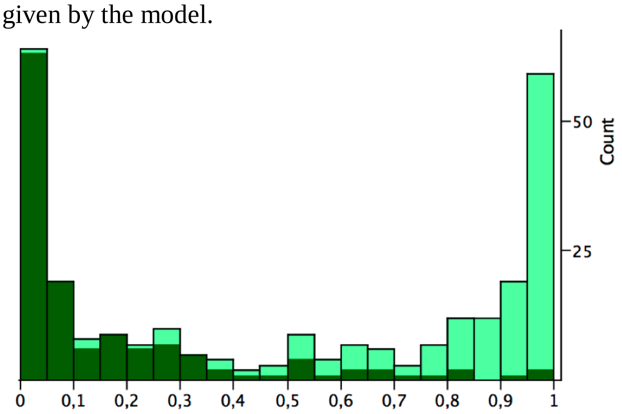

Fig.3 – Distribution of probabilities as given by the model. In light green (resp. dark green), elements with homogeneous (resp. inhomogeneous) texture. The nearer the probability to 1, the more homogeneous the nuclei's texture.

A high concentration rate at both extremities can be seen, which shows that the elements are well separated. But 40 ambiguous cases exists (for which the given probability is near the decision value of 0.5, between 0.3 and 0.7). Moreover, 8 important errors appear: dark green (respectively light green) elements in the right hand (respectively left hand) column. These errors correspond to elements for which the model gives a probability upper than 0.8 (respectively lower than 0.2) but which have an inhomogeneous (respectively homogeneous) texture. The presence of these non-negligible errors and ambiguities has lead us to improve the model.

4. GRAY LEVEL RUN LENGTH MATRIX

The Gray Level Run Length Matrix is a statistical texture characterization method [2,4,6]. This method consists in counting the number of pixel segments having the same intensity in a given direction, then representing the results in a matrix. A direction (0°, 45°, 90° or 135°) and a number of gray levels are decided on beforehand. The value contained in the matrix's *(l,n)* square is equal to the number of segments of length *l* and gray level *n*. This implies that the matrix's number of columns is dynamic, as it is determined by the length of the longest segment. By design, this calculation is symmetrical and consequentially, it is unnecessary to consider the four complementary directions (180°, 225°, 270° or 315°, in this example 8 possible directions between a given pixel and its neighbors are taken into account). Figure 4 shows an example of the calculation of a Run Length Matrix:

Image

1	2	3	4
1	3	4	4
3	2	2	2
4	1	4	1

=>

Gray Level (i)	Run Length (j)			
	1	2	3	4
1	4	0	0	0
2	1	0	1	0
3	3	0	0	0
4	3	1	0	0

Fig.4 – Example of the calculation of a run length matrix for a 4x4 image in 0° direction and for 4 gray levels.

Once the matrix obtained, 11 indexes are calculated [12] to determine the vector that characterizes the texture. To establish our model, the matrix for a given gray level and for four directions was calculated. Then, for each index, the average value of the four directions was taken. A systematic study found that the best model was obtained for a set of 7 indexes and 32 gray levels. The classification success rate was 84.81% by logistic regression, which is inferior to the rate obtained with the cooccurence matrix and the Haralick features (90%).

5. NEW METHOD: GRAY LEVEL SIZE ZONE MATRIX

A homogeneous texture is composed of large areas of the same intensity, and not of small groups of pixels or segments in any given direction. To take this fact into consideration, it was necessary to take into account, in a matrix, the size of each area with pixels of the same intensity level. This matrix was calculated according to the *Run Length Matrix* principle: the value of the matrix's *(s,n)* square is equal to the number of areas of size *s* and of gray level *n*. Figure 5 shows an example of the calculation of such a matrix, baptized *Size Zone Matrix*.

Image

1	2	3	4
1	3	4	4
3	2	2	2
4	1	4	1

=>

Gray Level (i)	Size Zone (j)			
	1	2	3	4
1	2	1	0	0
2	1	0	1	0
3	0	0	1	0
4	2	0	1	0

Fig.5 – Example of the calculation of a Size Zone Matrix for a 4x4 image with 4 gray levels.

The resulting matrix has a fixed number of lines equal to the number of gray levels and a dynamic number of columns, determined by the size of the largest area. The more homogeneous the texture, the *wider* and flatter the matrix will be. This matrix has the advantage of not requiring calculations in several directions, which are replaced by tagging different areas. However, specifying the number of gray levels is still necessary, but this renders the calculations robust to noise. The 11 same indexes as for the Run Length Matrix with 32 gray levels are calculated. The classification success rate for these 11 indexes is 91.11% by logistic regression, which is the best result obtained by all the techniques used so far. This improvement can be clearly seen by comparing figures 3 and 6. A better distribution of the elements at both extremities and a diminution of 16 the ambiguous cases (then 29 ambiguous cases) is visible. However, six important errors remain.

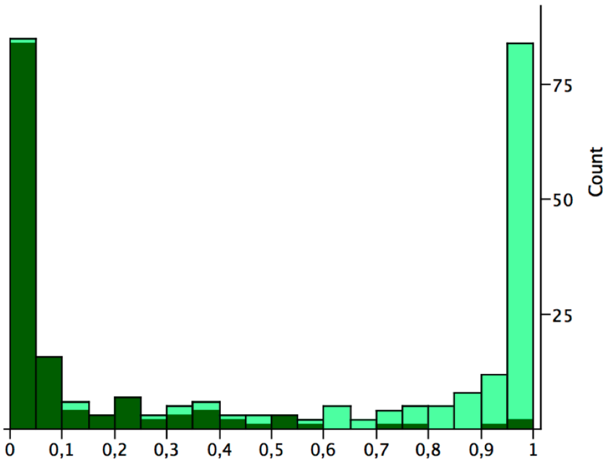

Fig.6 – Distribution of probabilities as given by the model. In light green (resp. dark green), the elements with a homogeneous (resp. inhomogeneous) texture. The nearer the probability to 1, the more homogeneous the nuclei's texture.

However, when examining the data, the indexes and the false positive results, it became clear that a specific texture case was not being correctly characterized and was the cause of the remaining errors: nuclei with large homogeneous areas, but with high variations in the intensity between the areas, making them inhomogeneous textured nuclei.

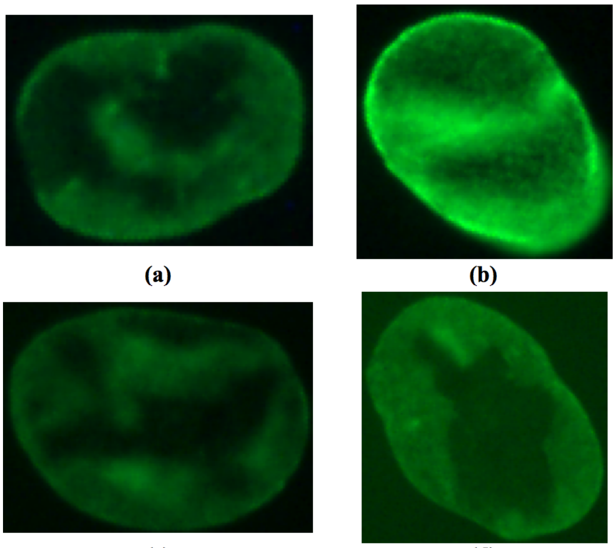

Fig.7 – Examples of false positive results: nuclei with large homogeneous areas, but with high intensity variations between these areas.

To characterize these nuclei types, two new indexes, which are weighted variances of gray level or area size, are needed:

$$Var_N = \sqrt{\frac{1}{N \times S}\sum_{n=1}^{N}\sum_{s=1}^{S}\left(n \times M(n,s) - \mu_N\right)^2}$$

with $\mu_N = \frac{1}{N \times S}\sum_{n=1}^{N}\sum_{s=1}^{S} n \times M(n,s)$

$$Var_S = \sqrt{\frac{1}{N \times S}\sum_{n=1}^{N}\sum_{s=1}^{S}\left(s \times M(n,s) - \mu_S\right)^2}$$

with $\mu_S = \frac{1}{N \times S}\sum_{n=1}^{N}\sum_{s=1}^{S} s \times M(n,s)$

with N and S the dimensions of the matrix and $M(n,s)$ the matrix's element of coordinates (n,s). The more the texture consists of large areas with high intensity variations between them, the higher the value of the Var_N index. In the case of a more homogeneous texture, the value of this index will be low. The same is true for the Var_S index, concerning area size.

In this way two new texture indexes are added to the 11 previous ones, making a total of 13 indexes. An extensive study was once again undertaken, using four different classification methods: logistic regression, *k*-nearest neighbors, random forests and neural networks. For nearest neighbors, we tested *k* equal to the number of indexes plus 1 to 30 and k equal 1 to 30. Best result is obtain with k equal to 1. Neural networks is a multi-layer perceptron, with a hidden layer. We tested various numbers of nodes in hidden layer: number of node of input layer divide by 2 to 6. Best result is obtain for 2. This study proved that the best subset is composed of 12 indexes (only the LRHGLE index is not used) on images reduced to 32 gray levels and classified by logistic regression. The different method's performances can be seen and compared in figure 8: better distribution of the elements at both extremities and number of ambiguous cases reduces to 16. Thanks to the use of the two new indexes in this new model composed of the 12 most pertinent indexes, the classification success rate of 92.59% has been reached. For the best configuration of indexes, performances between regression logistic and multi-layers perceptron are comparable.

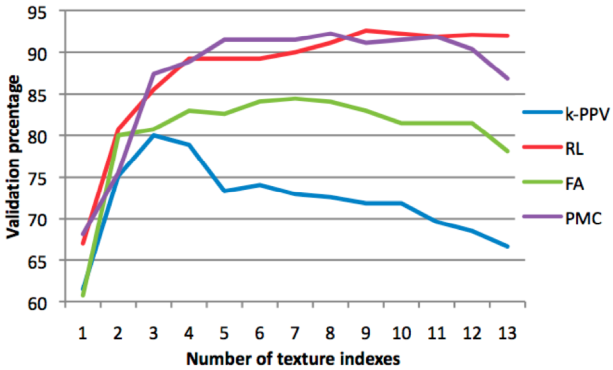

Fig.8 – Performance of the different classification methods with respect to the number of indexes used: k nearest neighbor (kNN), logistic regression (LR), random forests (RF) and multi-layers perceptron (MLP).

Our model's validity is illustrated in figure 9, which shows the distribution of probabilities as given by the model. The high concentration rate at both extremities of the histogram and the near absence of ambiguous cases (having a probability near the decision value of 0.5) show the efficiency of the classification and the pertinence of the choice of indexes.

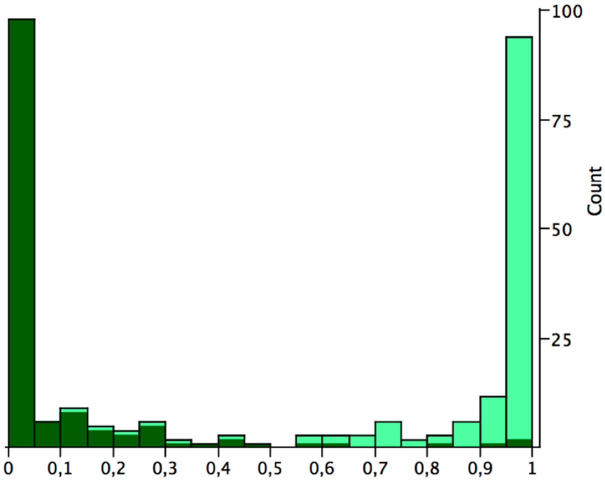

Fig.9 – Distribution of probabilities as given by the model. In light green (resp. dark green), the elements with a homogeneous (resp. inhomogeneous) texture. The nearer the probability to 1, the more homogeneous the nuclei's texture.

6. CONCLUSION AND FUTURE WORK

In this article, the problem of texture classification, applied to cell nuclei classification, has been covered. The main goal was to achieve pertinent texture homogeneity characterization. To do so, two existing texture characterization methods have been presented: the cooccurence matrix with Haralick features and the Run Length Matrix. Applied to this particular problem these methods did not obtain a satisfactory success rate (90%). For this reason, a novel method of texture homogeneity characterization has been presented. It consists in tagging, then counting the size of areas of the same intensity level. This allows a matrix, representative of the texture's homogeneity, to be found. In order to improve the pertinence, two new texture characterization indexes have also been determined. When combined and applied to nuclei texture classification, this method and these indexes obtain a success rate of 92.6%.

The initial aim of our work was to classify nuclei into health and unhealthy groups. Expert diagnosis showed that it was necessary to examine both the shape and the texture of the nuclei in order to achieve this objective. Our shape classification model (based on shape indexes) presented in [10] achieved a classification success rate of 95.4% when applied to nuclei's shape, but was only 86.9% successful with respect to the initial problem. The model presented in this article obtains a 92.6% success rate when applied to nuclei's texture, and when combined with the shape classification model, the overall success rate with respect to the initial problem reaches 87.7%. Even though this near one percent improvement may seem slight, it is in fact significant. As there is a large intersection between abnormally shaped nuclei and inhomogeneous textured nuclei, most of the abnormally textured ones are already classified by shape. The best possible improvement could only have been of 1%, so a 0.8% improvement is highly pertinent with respect to the probabilities.

By studying the intersections between the different classes of nuclei, the following conclusion was reached: only 94% of nuclei can be classified by their shape and / or texture alone. As a consequence, in our future work, complementary models will have to be established in order to characterize the unfrequented diagnostic criteria and thus further improve the initial problem's classification success rate. These unfrequented diagnostic elements are related to the presence of holes and focis. Holes are areas of the nuclei in which no lamin A/C are present, which leads to an absence of a reaction to the marker (cf. figure 10 a and b). A foci is a small, near circular area of high intensity (cf. figure 10 c and d).

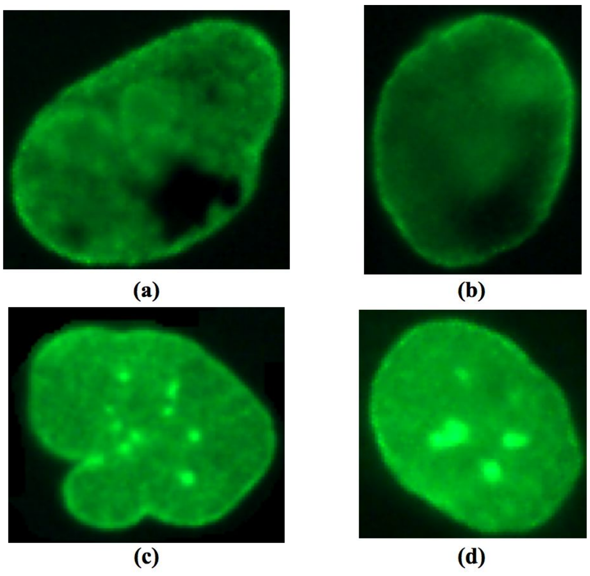

Fig.10 – Examples of nuclei with holes (a and b) or focis (c and d).

To detect and characterize these elements, we wish to determine an original approach, based on texture representation by its *volume below the surface* (cf. figure 11). This representation will allow the extraction of troughs and peaks, amongst which the 3D representation of holes and focis can be found. Once extracted, these volumes could then be characterized by a 3D extension of 2D shape indexes.

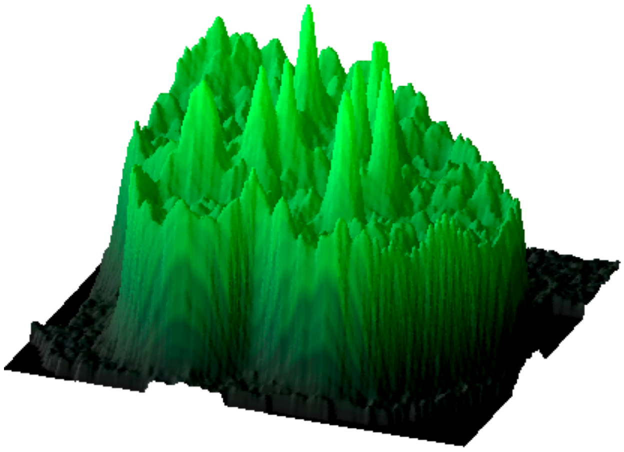

Fig.11 – Representation of the texture of the nucleus shown in fig. 10*c* by its volume from elevation map.

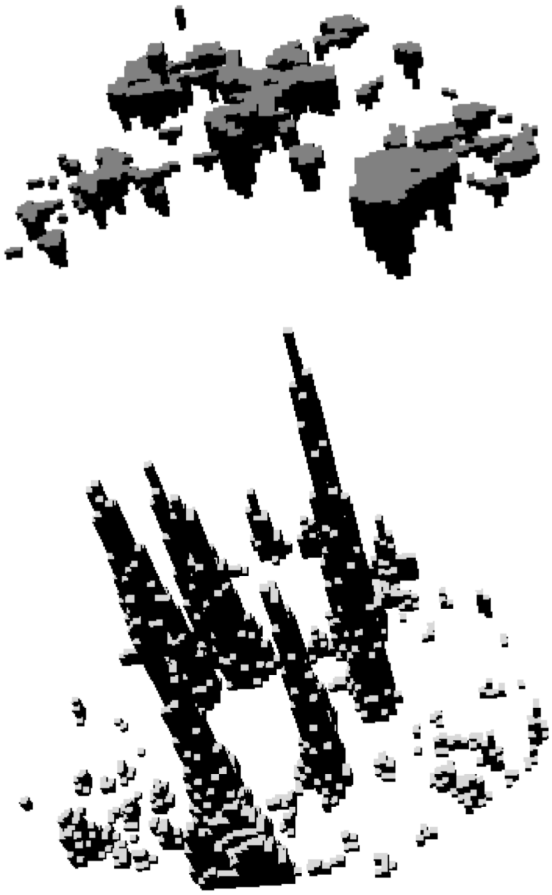

Fig.12 – Result of the extraction of troughs (above) and peaks (below) on the volume illustrated in fig. 11.

7. REFERENCES

[1] Chen Y. Q., Dixon M. S., Thomas D. W., "Statistical Geometrical Features For Texture Classification", Pattern Recognition, vol. 28, n° 4, p 537-552, 1995.

[2] Chu A., Sehgal C.M., Greenleaf J. F., "Use of gray value distribution of run length for texture analysis". *Pattern Recognition Letters*, vol. 11, n° 6, p. 415-419, 1990.

[3] Fisher R. A., “The use of multiple measurements in taxonomic problems”, *Annals of Eugenics* vol. 7, p. 179-188, 1936.

[4] Galloway M. M., “Texture analysis using grey level run lengths”, *Computer Graphics Image Processing*, vol. 4, p. 172-179, July, 1975.

[5] De Sandre-Giovannoli A., Bernard R., Cau P., Navarro C., Amiel J., Boccaccio I., Lyonnet S., Stewart C. L., Munnich A., Merrer M. L., Levy N., “Lamin A Truncation in Hutchinson-Gilford Progeria”, *Science,* vol. 300, n° 5628, p. 2055, June, 2003.

[6] Haralick R.M., Shanmugam K., Dinstein I., “Textural features for image classification”, IEEE *Transactions on Systems, Man and Cybernetics*, vol. 3, p. 610-621, 1973.

[7] Hosmer D., Lemeshow S., “*Applied Logistic Regression”*, John Wiley & Sons, Toronto, 1989.

[8] Martens H., Dardenne P., “Validation and verification of regression in small data sets”, Chemometrics *and intelligent laboratory systems*, vol. 44, n° 1-2, p. 99-121, 1998.

[9] McCulloch W. S., Pitts W., “A logical calculus of the ideas immanent in nervous activity”, *Bulletin of Mathematical Biophysics*, vol. 5, p. 115-133, 1943.

[10] Thibault G., Devic C., Horn J.-F., Fertil B., Sequeira J.,Mari J.-L., “Classification of cell nuclei using shape and texture indexes”, *International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision (WSCG),* Plzen, Czech Repulic, p. 25-28, February, 2008.

[11] Wolberg W., StreetW., Heisey D., Mangasarian O., “Computer-Derived Nuclear Features Distinguish Malignant From Benign Breast Cytology”, *Human Pathology*, vol. 26, Elsevier, New York, NY, United States, p. 792-796, July, 1995.

[12] Xu D., Kurani A., Furst J., Raicu D., “Run-Length Encoding For Volumetric Texture”, *International Conference on Visualization, Imaging and Image Processing (VIIP)*, p. 452-458, 2004.

[13] Shakhnarovich, G., Darrel, T., Ydik, P., “Nearest Neighbor Methods in Learning and Vision: Theory and Practice”, *The MIT Press*, 2006.

[14] Breiman, L., “Random forests”, *Machine Learning*, vol. 45, p. 5-32, 2001.