Plastic Surface Similarity Measurement Based on Textural and Fractal Features

DAN POPESCU*, RADU DOBRESCU

Politehnica University of Bucharest, Department of Control and Computers, 313 Splaiul Independentei, Bucharest 6, Romania

The paper presents two methods for similarity measurement of plastic textured images, and also for identification and localization of defective region in textured images. In order to have less sensitive statistic features regarding to rotated image, we introduce the notion of average co-occurrence matrix. With the purpose of algorithm validation, the images are divided into four equivalent regions. For the proper region similarity measurement, a decision theoretic method is used. In the fractal approach, we consider a new fractal dimension derived from box-counting algorithm, named effective fractal dimension, with an increasing efficiency for texture classification. Two experimental studies, one for statistical features and one for fractal type features, in a plastic simulated wood case, validate the algorithms. The algorithms are implemented in Visual C++ and Matlab. They allow the simultaneously display of both the investigated region, and the Euclidian distance between this and a reference image. The results confirm the fact that the distances between the regions without defect are relatively small, and the distance between a region with defect and a region without defect is relatively large. Also, the results show that features extracted from average co-occurrence matrix and the effective fractal dimension have a good discriminating power.

Keywords: *texture similarity, statistic features, fractal dimension, plastic simulated wood, defect detection*

It is very hard to define rigorously the texture into an image. The texture can be considered like a structure which is composed of many similar elements (patterns) named textons or texels, in some regular or continual relationship.

Texture analysis has been studied using various approaches, like statistical type (characteristics associable with grey level histogram, grey level image difference histogram, grey level co-occurrence matrices and the features extracted from them, autocorrelation based features, power spectrum, edge density per unit of area), fractal type (box counting fractal dimension), and structural type. The statistical approach uses features to characterize the stochastic properties of the grey level distribution in the image. The most powerful statistical method to textured image analysis is based on features extracted from the Grey-Level Co-occurrence Matrix (GLCM), [1]: contrast, energy, entropy, homogeneity, and variance. GLCM is a second order statistical measure of local image variation and it gives the joint probability of occurrence of grey levels of two pixels separated spatially by a fixed vector distance $d = (\Delta x, \Delta y)$.

The fractal based texture classification is another approach that correlates texture coarseness and fractal dimension. Thus, a method to relieve the texture of a surface is to calculate and combine different forms of fractal dimension [2-4].

Fractals have high power in low frequencies, which enables them to model processes with long periodicities. Many plastic surfaces have a statistical quality of roughness and self-similarity at different scales. Fractals are very useful and became popular in modeling these properties in image processing [5].

For the box counting basic algorithm, the image must be of binary type [6, 7]. The method consists in dividing the image, successively, in equivalent squares with normalized length $r = \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, ...$, and computing every time the number $N(r)$ of squares covered by the object image. The dividing process is limited by the image resolution. The fractal dimension can be obtained plotting $logN(r)$ for different values of $log(1/r)$, and calculating the slope of the resulting curve. A linear regression (1) is performed using the logarithmic coordinates, $x = log(1/r)$, $y = logN(r)$. The regression slope, a, is used to determine the box counting fractal dimension FD (2).

$$y = a\,x + b \tag{1}$$

The notation significances in equation (2) are the following: $x_i = \log(1/r_i)$, $y_i = \log(N(r_i))$, n – the number of partitions, $i = 1,2,3,\ldots,n$ – the function points in the graphical representation.

$$a = FD = \frac{n\sum_{i=1}^{n} x_i y_i - \left(\sum_{i=1}^{n} x_i\right)\left(\sum_{i=1}^{n} y_i\right)}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \tag{2}$$

There are two important kind of problems that texture analysis attempts to solve in plastic production investigation: a) are the plastic textures similar in all the regions? b) Is there a surface defect and where it is?

Texture classification involves deciding what texture class an observed image belongs to. Thus, one needs to have an a priori knowledge of the classes to be recognized [8- 12]. The major focus of this paper is the classification process, based on statistical features (especially derived from medium co-occurrence matrix), and fractal type features. Another objective is the defect detection and localization, based of texture analysis of regions which are obtained by image partition.

* *email: dan_popescu_2002@yahoo.com*

Experimental part

For the purpose of algorithm validation, two experimental studies have been conducted. The first study is focused on region similarity measurement of textured images, namely a plastic simulated wood type (fig. 1 and fig. 4) based on texture feature extracted from average co-occurrence matrix. Also, this study deals with defect identification and localization (fig. 3). With this end in view, the whole image, is partitioned in four equivalent regions like in figura 1. Different textured regions are compared based on minimum distance between measured features which are derived from medium co-occurrence matrices (contrast, energy, entropy, homogeneity, and variance). The second study presents the use of the fractal dimension in the texture similarity measurement of the plastic surfaces.

It is well known that the second order characteristics, which derived from the co-occurrence matrices offer good results relatively to texture classification. In order to increase the efficiency of these features we introduced the notion of *average co-occurrence matrix* (1).

The elements of a co-occurrence matrix C_d depend upon

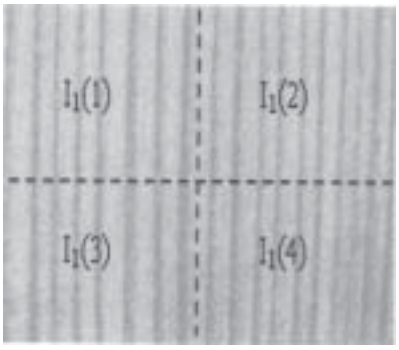

Fig.1. Four regions image partition of image I_1

displacement $d = (\Delta x, \Delta y)$:

$$C_d(i,j) = Card\{((x,y),(t,v)) / I(x,y) = i, \; I(t,v) = j, \; (x,y), (t,v) \in N \times N_I \; (t,v) = (x+\Delta x, y+\Delta y)\} \quad (3)$$

We consider increasing $(2d+1)$x$(2d+1)$ symmetric neighborhoods, d = 1, 2, 3,...,10. Thus, for 3 . 3 neighborhood, d = 1; for 5 . 5 neighborhood, d = 2; for 7 . 7 neighborhood, d = 3, and so on. Inside each neighborhood there are 8 principal directions: 1, 2, 3, 4, 5, 6, 7, 8 (fig. 2) and we evaluated the co-occurrence matrices $C_{d,k}$ corresponding to the vector distances determined by the central point and the neighborhood edge point in the k direction (k = 1,2,...,8).

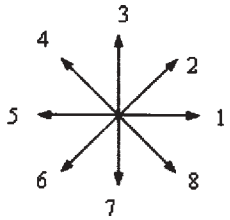

Fig.2. The principal directions for the average co-occurrence matrix calculus

For each neighborhood type, we define an average co-occurrence matrix C_{dav} (1):

$$C_{dav} = 1/8(C_{d,1} + C_{d,2} + C_{d,3} + C_{d,4} + C_{d,5} + C_{d,6} + C_{d,7} + C_{d,8}), \; d = 1,2,...,10 \quad (4)$$

In order to quantify the spatial dependence of gray level values, from average co-occurrence matrices C_d, d = 1, 2,..., 10, we calculate various textural features like Contrast – Con_d (5), Energy – Ene_d (6), Entropy – Ent_d (7), Homogeneity – Omo_d (8) and Variance – Var_d (9).

$$Con_d = \sum_{i=1}^{L}\sum_{j=1}^{L}(i-j)^2 C_d(i,j) \quad (5)$$

$$Ene_d = \sum_{i=1}^{L}\sum_{j=1}^{L} C_d(i,j)^2 \quad (6)$$

$$Ent_d = -\sum_{i=1}^{L}\sum_{j=1}^{L} C_d(i,j)\log(C_d(i,j)) \quad (7)$$

$$Omo_d = \sum_{i=1}^{L}\sum_{j=1}^{L}\frac{C_d(i,j)}{1+|i-j|} \quad (8)$$

$$Var_d = \frac{1}{L}\sum_{i=1}^{L}\sum_{j=1}^{L}[C_d(i,j) - \overline{C_d(i,j)}]^2 \quad (9)$$

In the preceding notations, *LxL* is the dimension of co-occurrence matrices.

For the purpose of the evaluation of the discriminating power of the selected features we calculate the Euclidian distances between two regions: a reference one, for example $I_1(1)$ and a measured one, for example $I_1(2)$. The Euclidian distance $D\{I_1, I_2\}$ between two images I_1 and I_2, which are characterized by the feature vectors $[C_1,E_1,Et_1,O_1,V_1]^T$ and $[C_2,E_2,Et_2,O_2,V_2]^T$ is expressed by the following relation

$$D(I_1,I_2) = \sqrt{(C_1-C_2)^2+(E_1-E_2)^2+(Et_1-Et_2)^2+(O_1-O_2)^2+(V_1-V_2)^2} \quad (10)$$

where:

$C = Con, E = Ene, Et = Ent, O = Omo, V = Var.$

Because the ranges of the initial characteristics can differ too much, for efficient Euclidian distance calculation, the characteristics are necessary to be normalized. So, we considered that all the feature values which correspond to the reference region are equal to 1.

For the second experimental study, the texture edges were extracted and the fractal dimension calculate. Because the images are grey level type, for the binary thresholds, an interval was chosen, based on the request that a definite texture exist in the contour image. The texture edges are extracted from the binary image, by a local logical operator [13]. The fractal dimension spectrum of contour images were represented and the mean fractal dimension *MDF* compute d. The slope of the log-log curve is evaluated by the linear regression method. The existence, inside an image, of textured regions with different fractal dimension requires alternative methods for *FD* estimation [13 - 15]. Because the analysed regions are textured, their contour images (edges) are full of edge pixels, and the first points in the log-log representations give a partial *FD* equal with 2. Therefore, we proposed another fractal dimension named *effective fractal dimension* (*EFD*), [16]. The $EFD = a_E$ is calculated by the omission of the first points in the log-log representation (the points of the form (x_i, x_i^2), i = 1,2,...,k) .

$$a_E = \frac{n\sum_{i=k+1}^{n} x_i y_i - \left(\sum_{i=k+1}^{n} x_i\right)\left(\sum_{i=k+1}^{n} y_i\right)}{n\sum_{i=k+1}^{n} x_i^2 - \left(\sum_{i=k+1}^{n} x_i\right)^2} \quad (11)$$

The threshold assessment, used for edge extraction, constitutes a problem of the fractal dimension evaluation in the grey level image case. Thus, we proposed some methods for threshold establishment [17, 18]:

a. The values for which the contour conserve the texture of image;

b. The value for which there are the most of points to contour;

c. The values for which the contour pixel set is a nonempty one. In this case, one must calculate *MFD*.

The algorithm, which was implemented in MATLAB, consists of the following steps:

1) Reading and converting of the color image in 32 grey levels;

2) Converting of the 32 grey levels image to a binary level image using a corresponding threshold *T*;

3) Extraction of the image contour using 3x3 neighborhoods;

4) Computing of the fractal dimensions *FD*, and *EFD*, from the contour image, applying the box-counting algorithm.

Results and discussion

For algorithm testing and program validation we used two textured images I_1 and I_2, each partitioned in four regions $I_i(1)$, $I_i(2)$, $I_i(3)$, $I_i(4)$, $i = 1,2$ (fig. 1 and fig. 3), and a region from another image (fig. 4), similar with I_1, namely $I_3(1)$. All the regions are 256 . 256 images with 32 grey levels. The similar regions are the following pairs: [$I_1(1)$, $I_1(2)$], [$I_1(1)$, $I_1(3)$], and [$I_1(1)$, $I_1(4)$]. We considered two sets of feature values. The first set derives from the regions of the images I_1 and I_3. With this set we tested the similarity between plastic simulated wood type surfaces. Both images are similar with the exception of the texture finesse (The image I_3 is more fine than the image I_1).

The second set of texture values derives from image I_2, which represents a deal board with defect (two knots). With this set we tested the defect detection and region identification based on textures extracted from average co-occurrence matrices. Textural features like Con_d, Ene_d, Ent_d, Omo_d, and Var_d are calculated. We consider that the textures of the regions $I_1(1)$ and $I_2(1)$ are the reference ones. Therefore, all the feature values for $I_1(1)$ and $I_2(1)$ are equal with 1. For example, in the case $d = 10$, the normalized results are presented in table 1 (I_1, and I_3), and table 2 (I_2).

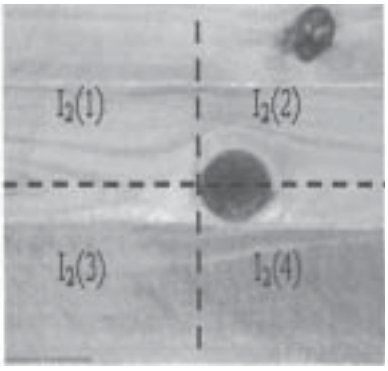

Fig.3. Four regions image partition of image I_2

Fig.4. Region 1 from image I_3

Table1

NORMALIZED VALUES OF FEATURES CORRESPONDING TO REGIONS INSIDE IMAGES I_1 AND I_3

Image	Con_d	Ene_d	Ent_d	Omo_d	Var_d
$I_1(1)$	1	1	1	1	1
$I_1(2)$	1.02	1.19	0.87	1.02	0.88
$I_1(3)$	0.99	0.96	0.96	1	1.01
$I_1(4)$	1.01	1.05	0.90	1.03	0.97
$I_3(1)$	1.04	1.43	0.52	1.22	0.82

Table 2

NORMALIZED VALUES OF FEATURES CORRESPONDING TO REGIONS INSIDE IMAGE I_2

Image	Con_d	Ene_d	Ent_d	Omo_d	Var_d
$I_2(1)$	1	1	1	1	1
$I_2(2)$	0.91	0.67	5	0.90	5.67
$I_2(3)$	0.96	0.98	1.44	0.99	1.58
$I_2(4)$	0.89	0.42	3.22	0.86	3.28

From table 1 it is calculated the Euclidian distances (9) between the reference image $I_1(1)$, with normalized values 1, and the test regions $I_1(2)$, $I_1(3)$, $I_1(4)$, and $I_3(1)$. The results are presented in table 3 and one can observe that the distances between two different regions, like $D\{I_1(1),I_3(1)\}$, are greater than distances between two similar regions, like $D\{I_1(1),I_1(2)\}$, $D\{I_1(1),I_1(3)\}$ or $D\{I_1(1),I_1(4)\}$. In order to appreciate the efficiency of the presented algorithm, we analyzed the most unfavorable case, namely the maximum distance between the reference image and the regions coming from the same image. Thus, $D\{I_1(1),I_3(1)\}$ is bigger than $\max\{D\{I_1(1),I_1(k)\}\}$.

Table 3

EUCLIDIAN DISTANCES BETWEEN

Distances	Values	Distances	Values
$D\{I_1(1),I_1(2)\}$	0.26	$D\{I_2(1),I_2(2)\}$	7.61
$D\{I_1(1),I_1(3)\}$	0.06	$D\{I_2(1),I_2(3)\}$	0.73
$D\{I_1(1),I_1(4)\}$	0.12	$D\{I_2(1),I_2(4)\}$	4.64
$D\{I_1(1),I_3(1)\}$	0.71		

Also, from the table of distances, it can be noticed that the same idea of region comparison based on statistical features extracted from average co-occurrence matrix, can be used to defect detection and localization. So, there are defects in regions $I_2(2)$, and $I_2(4)$, inside of image I_2. The defect is bigger in region $I_2(2)$ than in region $I_2(4)$.

The second case study implies the calculus of two estimations of box-counting based fractal dimensions. From this point of view we considered the algorithm which uses the the most contour pixels. The first estimation is FD and the second estimation is EFD. It can be seen that EFD is smaller than FD.

The graphical results are presented in figure 5, for $I_1(1)$, and figure 6, for $I_3(1)$.

The full division box-counting algorithm (FD algorithm) considers the linear regresion slope from all the points The vector [v] represents the division vector, and the vector [w] represents the resulting box-counting vector (the numbers of squares which contain edge pixels). The vectors [x] and [y] represent the component logarithms of [v] and [w]:

$$x = \log v, y = \log w \qquad (12)$$

Thus, FD is the fractal dimension, which is evaluated from (*x*,*y*) representation, namely log-log representation.

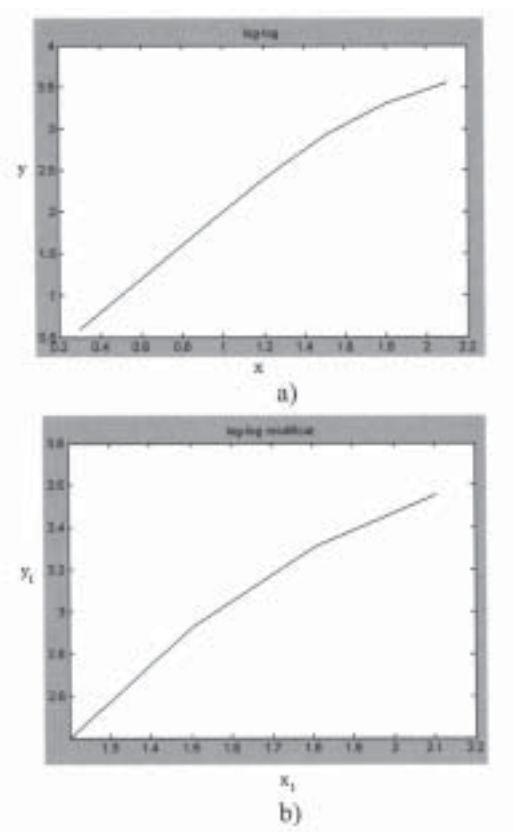

Fig. 5. The curve log-log for the box-counting fractal dimension evaluation for region $I_1(1)$: a) For *FD* evaluation, b) For *EFD* evaluation(modified log-log)

The effective division box counting algoritm (ED algoritm) considers (*x*,*y*) points, begining after a start point, corresponding to first value of *x* wherefore y ≠ 2x. The new representation is (*xe*,*ye*), namely modified log-log:

$$xe = \log v, \quad ye = \log w. \qquad (13)$$

Thus, EFD is the effective fractal dimension, which is evaluated from (*xe*,*ye*) representation.

For $I_1(1)$ and $I_3(1)$, the results are the following:

a) Image $I_1(1)$, threshold T = 140.
[*x*] = [0.301 0.602 0.903 1.204 1.505 1.806 2.107]
[*y*] = [0.602 1.204 1.806 2.403 2.932 3.312 3.555]
[*v*] = [2 4 8 16 32 64 128]
[*w*] = [4 16 64 253 855 2052 3587]
FD_1 = 1.685
Start point (1.204, 2.403)
[*xe*] = [1.204 1.505 1.806 2.107]
[*ye*] = [2.403 2.932 3.312 3.555]
EFD_1 = 1.274
b) Image $I_3(1)$, threshold T = 140
[*x*] = [0.301 0.602 0.903 1.204 1.505 1.806 2.107]
[*y*] = [0.602 1.204 1.806 2.377 2.923 3.369 3.649]
[*v*] = [2 4 8 16 32 64 128]
[*w*] = [4 16 64 238 837 2339 4459]
FD_3 = 1.731
Start point (1.204, 2.377)
[*xe*] = [1.204 1.505 1.806 2.107]
[*ye*] = [2.377 2.923 3.369 3.649]
EFD_3 = 1.417

We can observe that the discriminating efficiency is better for *EFD* than *FD*:

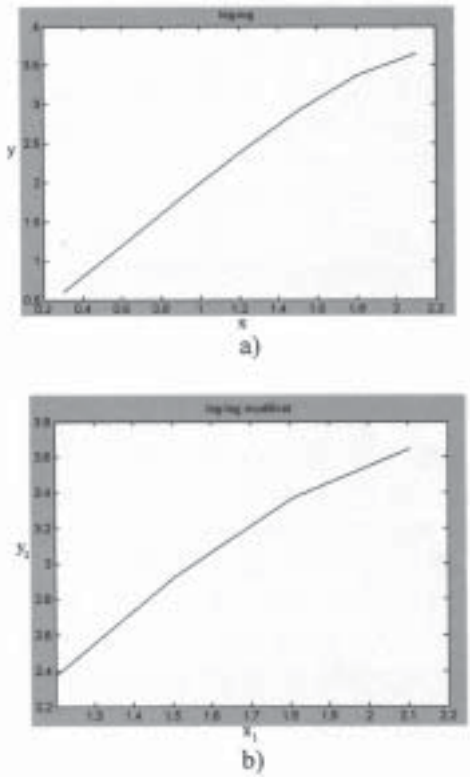

Fig. 6. The curve log-log for the box-counting fractal dimension evaluation for region $I_1(1)$: a) For *FD* evaluation, b) For *EFD* evaluation (modified log-log)

$$\frac{FD_3 - FD_1}{FD_3} = 2.65\%, \qquad \frac{EFD_3 - EFD_1}{EFD_3} = 10\%$$

Conclusions

Because it is considered an average co-occurrence matrix, the presented algorithm is relatively insensible to image translation and rotation. The results confirm that the statistic second order features, extracted from average co-occurrence matrices, offer a good discriminating power, both in texture identification process and in defect detection and identification. The main application of the algorithm consists in texture identification and classification, and defect detection in the regions of textured images (like plastic surfaces). The additional features, like effective fractal box-counting dimension, increase the power of discrimination for texture identification and classification. The efficiency of the defect detection and localization depends upon the range of image partition and the texture element dimension.

References

1. HARALICK, R.M. et al, IEEE Trans. on Systems, Man. And Cybernetics, vol.SMC-3, nr.6, nov.1973, p. 610
2. HARALICK, R.M., SHAPIRO, L.G., Computer and Robot Vision, Add.-Wesley, Pub. Co., 1992
3. TAMURA, H., MORI, S., YAMAWAKI, T., IEEE Trans. On Systems, Man and Cybernetics. 6(4), 1976, p. 460
4. NIBLACK, W. et al. - The QBIC Project: Querying Images by Content Using Color, Texture and Shape, Proc. of the Conference Storage and Retrieval for Image and Video Databases, SPIE vol.1908, 1993, p.173
5. MANDELBROT, B. B., Fractal Geometry of Nature, Freeman, New York, 1982
6. OLTEANU M., PAUN V.-P., TANASE M., Rev. Chim. (Bucure°ti), **56**, nr.1, 2005, p. 97

7. PAUN V.-P., OLTEANU M., MARIN A., Rev. Chim. (Bucureºti), **58**, nr.3, 2007, p.355
8. LIU, F., PICARD, R.W., IEEE Transactions on Pattern Analysis and Machine Intelligence 18(7), 1996, p. 722
9. MANJUNATH, B.S., MA, W.Y. - Texture features for browsing and retrieval of large image data, IEEE Transactions on Pattern Analysis and Machine Intelligence, (Special Issue on Digital Libraries), Vol. 18 (8), 1996, p. 837
10. KAPLAN, L.M. et al. - Fast texture database retrieval using extended fractal features, in Storage and Retrieval for Image and Video Databases VI
(SETHI, I K, JAIN, R C, eds.), Proc SPIE 3312, 1998, p. 162
11. SMITH, J., Integrated Spatial and Feature Image System: Retrieval, Analysis and Compression, Ph. D. Thesis, Columbia University, 1997
12. PAUN V.-P., Mat. Plast., **40**, nr. 1, 2003, p. 23
13. DENG,Y., A Region Representation for Image and Video Retrieval, Ph. D. thesis, University of California, Santa Barbara, 1999
14. MA, W.Y., Netra: A Toolbox for Navigating Large Image Databases, Ph. D. thesis, University of California, Santa Barbara, 1997
15. IORDACHE D., PUSCA S., TOMA G., PÃUN V.-P., STERIAN A., MORARESCU C., Lect Notes Comput SC, **3980**, 2006, p.804
16. OLTEANU, M., PAUN V.-P., TÃNASE, M.., Rev. Chim. (Bucureºti), **56**, nr.7, 2005, p.781
17. POPESCU, D., DOBRESCU, R., AVRAM, V., MOCANU, ST., WSEAS Trans. on Systems, Issue 8, Vol.5, 2006, p. 1932
18. POPESCU, D., DOBRESCU, R., NICOLAE, M., AVRAM, V., Algorithm Based On Medium Co-Occurrence Matrix For Image Region Classification, Proceedings of the 7th WSEAS International Conference on Signal Processing, Computational Geometry and Artificial Vision (ISCGAV'07), Athens, Greece, August 24-26, 2007, p. 119

Manuscript received: 10.03.2008