# A NOTE ON THE CORRELATION AND HIGUCHI DIMENSIONS FOR IMAGE ANALYSIS

Andreea UDREA[1], Mircea OLTEANU[2]

*In this paper we present a method for the estimation of the correlation dimension of the attractor and the Higuchi dimension of the spatial series that is extracted from a measurement (image) of a skin lesion. We are interested in melanoma (deadliest skin cancer) early detection. The grey scale image and each image color channel are separately investigated. The correlation dimension calculated for the red image channel and the grey scale image and, also, the Higuchi dimension for the grey scale image represent relevant melanoma descriptors.*

**Keywords**: attractor correlation dimension, Higuchi dimension, image analysis

## 1. Introduction

Nonlinear time series analysis (NTSA) is a branch of chaos theory and its methods have been developed in the past 25 years, being motivated by the existence of deterministic chaos in many systems from biology, medicine, chemistry, physics and electronics [1]. An objective of NTSA is to characterize the chaotic behaviour of a system based on a measurement (a time series) by using the invariants of the system: attractor's correlation dimension, entropy, Lyapunov exponents, Higuchi dimension.

A discrete time series is a set of time-ordered observations generally spaced at equal time intervals:

$$Y = \{Y_i \mid i = 1, N\}. \tag{1}$$

Considering the observations made by far in biology and medicine, one can expect complicated functional behaviours (e.g., the cardiac rhythm is chaotic), respectively, specific self-similarities in the macroscopic organization of organ tissues (e.g., liver, skin) [2], [3]. This fractal organization is the result of the dynamical evolution over time and can be assimilated to attractors. In this case, any image (measurement) captures information on the system that can be investigated by using NTSA methods. The only problem is that NTSA refers to time series and that an image is a static 2D signal. There are precedents for image

[1] Assoc.Prof., Dept.of Automatic Control and Systems Engineering, University POLITEHNICA of Bucharest, Romania, e-mail: andreea.udrea@acse.pub.ro
[2] Prof., Dept. of Mathematics, University POLITEHNICA of Bucharest, Romania, Romania, e-mail: mirolteanu@yahoo.co.uk

transformation into 1D series [4]-[7] in order to be investigated with NTSA methods.

A very important problem in dermatology is melanoma (deadliest type of skin cancer) early detection. This is a difficult task even for the dermatologists and the development of automatic tool for diagnosis was a major research area in the last 10 years. In this context, we have investigated potential methods to produce trustworthy descriptors for skin lesions when using a smartphone camera to acquire images [8], [9]. In this paper we present the results obtained by using the correlation dimension of the attractor and the Higuchi dimension (chapter 2) in terms of sensitivity and specificity in melanoma diagnosis. For our study we analyse the lesions images (grey scale and on each color channel) and obtain a set of 8 descriptors. Their performances are presented in chapter 3.

## 2. Mathematical method

### 2.1. The correlation dimension

Considering a time series, one can reconstruct in an appropriate embedding space the attractor from the time series generated by the system, and estimate its correlation dimension.

Taken's Embedding Theorem states that, if the fractal dimension ($d_f$) of the attractor is known, the attractor can be reconstructed from an 1D time series in a higher dimensional space and the embedding space dimension ($d_E$) is $2d_f+1$ [10].

In our case, we do not know $d_f$, so we estimate it using the false nearest neighbours method as described in [7]. The delay $\tau$ for the embedding is determined by using the autocorrelation function decay method [1].

Given a series $Y$ of length $N$, its embedding in $d_E$ dimension using the delay $\tau$ is:

$$(y_i, y_{i+\tau}, \ldots, y_{i+(d_E-1)\tau}) \longrightarrow x_i \tag{2}$$

satisfying $N' + (d_E - 1)\tau \le N$.

The correlation integral is defined by the following expression:

$$C(\varepsilon) = \lim_{N\to\infty} \frac{1}{N^2} \sum_{i,j=1}^{N} H(\varepsilon - |x_i - x_j|), \tag{3}$$

where $H(x)$ is the Heaviside function and $\varepsilon$ the imposed distance between points.

The correlation dimension of the attractor is calculated using the formula:

$$d_C = \lim_{\varepsilon \to 0} \frac{\ln C(\varepsilon)}{\ln \varepsilon}. \tag{4}$$

**2.2. The Higuchi fractal dimension**

In order to estimate the Higuchi fractal dimension associated to a time series *Y*, one has to consider a finite set of observations *Y(j), j=1:N*, an interval *k*, and generate *k* new series in the following fashion :

$$Y_m^k : Y(m), Y(m+k), \ldots, Y(m+\left[\frac{N-m}{k}\right]),$$

where: *m=1 : k* provides the initial element *Y(m)* for each new series and *k* is the delay between elements. For each series $X_m^k$ the length $L_m(k)$ is evaluated using the formula:

$$L_m(k) = \frac{N-1}{\left[\frac{N-m}{k}\right]k^2}\left(\sum_{i=1}^{\left[\frac{N-m}{k}\right]} \left[X(m+ik) - X(m+(i-1)k)\right]\right), \tag{5}$$

where: $\frac{N-1}{\left[\frac{N-m}{k}\right]k^2}$ is called the normalization factor.

The total medium length *L(k)* is calculated for all series generated with the same delay *k* and different *m* values:

$$L(k) = \sum_{m=1}^{k} L_m(k). \tag{6}$$

The procedure is repeated for each *k=1*: *kmax*. If the series has fractal properties over *k*, then:

$$L(k) \sim k^{-d_H}, \tag{7}$$

where $d_H$ is called the Higuchi dimension.

**2.3. Image analysis**

In order to obtain high quality images for analysis, we are using an app specifically designed [8] to acquire high-resolution, well focused, shadows free images of skin lesions. The lesions are also guaranteed to be completely imaged and centred. Each lesion has 3 images taken. For our study, we use a database of

140 lesions (3 images each): 115 benign and 25 melanomas in early stages (with associated histopathological results).

From each image we extract the region of interest - the lesion and we analyse the information from the corresponding grey scale image representation and also from each color channel: red, green and blue; we have 4 matrixes denoted by $I_{gr}$, $I_R$, $I_G$ and $I_B$ associated to each image of a lesion.

Considering, for example, a grey scale image $A$ of $W$-width and $H$-height:

$$A = (a_{ij})_{\substack{i=1,W \\ j=1,H}}, \quad \forall a_{ij} \in \{0,\ 1,\ \ldots,\ 255\}; \tag{8}$$

one can transform it from its 2D representation (8) into an 1D representation by using the *vec* operator (9):

$$vec : A(R) \to \mathbb{R}^{W \cdot H},\ vec(A) = col(a_{i1},\ a_{i2}, \ldots,\ a_{iH}),\ i = \overline{1,W}. \tag{9}$$

Moreover, we know where the lesion lies in the image, so we are going to apply the above procedure only inside the previously identified lesion contour:

$$Y = \{y_i \mid i = 1, S\}, \tag{10}$$

where $S$ is the area (in pixels) delimited by the lesion's contour.

After normalization, we obtain the following "spatial series": $Y_{gr}$, $Y_R$, $Y_G$ and $Y_B$, that correspond to $I_{gr}$, $I_R$, $I_G$ and $I_B$ matrixes and estimate $d_c$ and $d_H$ for each spatial series.

In terms of computational duration, the $d_H$ estimation method runs in less than 1 second while the $d_C$ estimation methods is time consuming (more than a minute in most of the cases).

## 3. Results and conclusions

### 3.1. Results

The analysis set (140 lesions) was divided into a training set (12 melanomas and 57 benign lesions) and a test set (13 melanomas and 58 benign nevi). For each lesion we have 3 associated images – so 3 results for each descriptor. The final value of a lesion's descriptor is the mean of the 3 values. The requirements for a relevant descriptor are made in terms of sensitivity (*Ss*) and specificity (*Sp*) of the classification:

$$Ss = \frac{TP}{TP + FN} \cdot 100\ ;\ \ Sp = \frac{TN}{TN + FP} \cdot 100, \tag{11}$$

where: TP (true positive) stands for the number of correctly identified melanomas, FN (false negative) is the number of melanomas that are not correctly identified, TN (true negative) is the number of benign lesions that are correctly identified and FP (false positive) is the number of benign lesions that are not correctly identified as normal.

The mean dimension value (calculated on each color channel or on the intensity matrix) is considered to be a relevant descriptor for melanoma detection if it classifies the lesions in the training set with specificity ($Sp_{tr}$) higher than 80%. Considering this condition, the mean $d_c$ and $d_H$ values ($V_d$) that best separates the lesions in benign and cancerous are determined and the associated sensitivities ($Ss_{tr}$) and specificities ($Sp_{tr}$) are calculated on the training set. Next, the specificity ($Sp_{test}$) and sensitivity ($Se_{test}$) are calculated on the test set and the obtained results can be found in Table I.

*Table 1*

**Sensitivities and specificities values for the training and test set – $d_c$ and $d_H$**

| | $Ss_{tr}$ $d_c$ [%] | $Sp_{tr}$ $d_c$ [%] | $V_d$ $d_c$ | $Se_{test}$ $d_c$ [%] | $Sp_{test}$ $d_c$ [%] | $Ss_{tr}$ $d_H$ [%] | $Sp_{tr}$ $d_H$ [%] | $V_d$ $d_H$ | $Se_{test}$ $d_H$ [%] | $Sp_{test}$ $d_H$ [%] |
|---|---|---|---|---|---|---|---|---|---|---|
| Red color channel | 76.92 | 80. 7 | 1.53 | 83 | 81.03 | 58.3 | 80. 7 | 1.39 | 46.15 | 74.14 |
| Green color channel | 50 | 80.7 | 1.28 | 53.85 | 75.86 | 41.67 | 80.7 | 1.4 | 46.15 | 72.41 |
| Blue color channel | 58 | 80. 7 | 1.32 | 61.53 | 77.59 | 50 | 80. 7 | 1.4 | 53.85 | 74.14 |
| Grey scale image | 75 | 85.96 | 1.39 | 76.92 | 82.75 | 66.67 | 80. 7 | 1.38 | 61.53 | 79.31 |

These results show that the methods proposed in this article lead to good classification performances when considering the correlation dimension estimated for $Y_R$ and $Y_{gr}$ and the Higuchi dimension estimated for $Y_{gr}$.

### 3.2. Conclusions

NTSA methods prove efficient in the context of image analysis, providing reliable texture descriptors.

By taking into consideration the correlation dimension calculated for the red image channel and the grey scale image and, also, the Higuchi dimension for the grey scale image along with other texture [9], shape [8] and color descriptors of a skin lesion, we expect to create an automatic diagnosis algorithm that equals the dermatologist's accuracy.

**Acknowledgement**

The work of Andreea Udrea was supported by Project SOP HRD - PERFORM /159/1.5/S/138963.

## R E F E R E N C E S

[1] *Holger Kantz, T. Schreiber*, "Nonlinear Time Series Analysis", University Press, Cambridge, 2004.

[2] *Catalin Vasilescu, A. Herlea, B. Ivanov, R. Dobrescu, F. Talos*, "A survey on differences between intestinal and diffuse type of gastric carcinoma", in Interdisciplinary Applications of Fractal and Chaos Theory, Editura Academiei Romane, Bucuresti, 2004.

[3] *Mircea Olteanu, R Dragomir and M. Tanase*, "The detection of modified structure in human tissues by computing the local fractal dimension", in Interdisciplinary applications of fractal and chaos theory, Editura Academiei Romane, 2004, pp. 255-264.

[4] *Torsten Mattfeldt*, "Spatial pattern analysis using chaos theory: A nonlinear deterministic approach to the histological texture of tumours", in Fractals in Biology and Medicine, Mathematics and Biosciences in Interaction 1998, pp. 50-72.

[5] *Wlodzimierz Klonowski, R. Stepien and P. Stepien*, "Simple fractal method of assessment of histological images for application in medical diagnostics", in Nonlinear Biomedical Physics,vol. 4, no. 7, 2010.

[6] *Helmut Ahammer*, "Higuchi Dimension of Digital Images", in Plos ONE, 2011.

[7] *Andreea Udrea, M. Olteanu*, "Image analysis based on the study of the attractor of a time series", in Buletinul Stiintific UPB, seria A, nr.4, 2010, pp. 11-20.

[8] *Andreea Udrea, C. Lupu*, "Real-time acquisition of quality verified non-standardized color images for skin lesions risk assessment – a preliminary study", in 18th International Conference on System Theory, Control and Computing, Sinaia, 2014, pp. 199-204.

[9] *Tanja Maier, D. Kulichova, K. Schotten, R. Astrid, T. Ruzicka, C. Berking, A. Udrea*, Accuracy of a smartphone application using fractal image analysis of pigmented moles compared to clinical diagnosis and histological result, in European Academy of Dermatology and Venereology, vol. 29, iss. 4, 2015, pp. 663-667.

[10] *Floris Takens*, "Detecting strange attractors in turbulence", in Dynamical Systems and Turbulence, Lectures Notes in Mathematics, vol. 898, Springer-Verlag, 1981, pp. 366-381.