

# COMPUTER PROGRAM FOR IMAGE TEXTURE ANALYSIS IN PhD STUDENTS LABORATORY

Piotr Szczypiński<sup>1</sup>, Marcin Kociołek<sup>1</sup>,  
Andrzej Materka<sup>1</sup>, Michał Strzelecki<sup>1</sup>

<sup>1</sup> Institute of Electronics, Technical University of Lodz, 18 Stefanowskiego Str., 90-924 Lodz, Poland, email: [pms][kociolek][materka][mstrzel]@ck-sg.p.lodz.pl

**Abstract:** The following paper reports results on building software for image texture analysis. The main goal was to create an efficient tool for quantitative analysis of magnetic resonance images. The software was written to compute a variety of texture features (parameters) based on image histogram, co-occurrence and run-length matrices, image gradient, auto-regressive model and wavelet transform. Calculated features may be presented in a form of text list or in images (feature maps).

**Keywords:** image texture, texture features, feature maps, magnetic resonance imaging

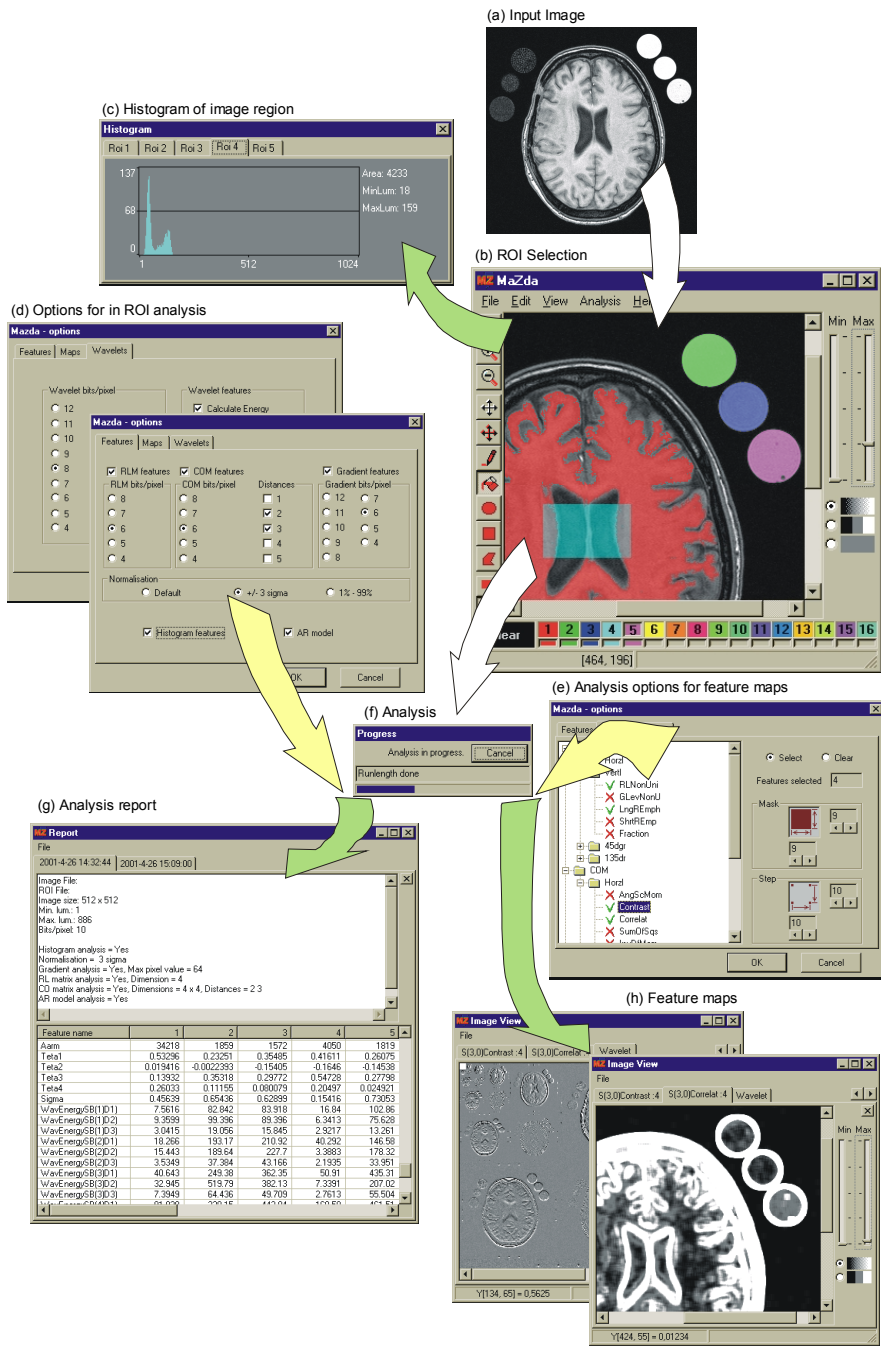
## 1. INTRODUCTION

At the beginning of 1998 the COST B11 project started. The main objective of the project is to “coordinate and focus recent developments in quantitative magnetic resonance imaging, in particular texture analysis, to maximize the amount of clinical diagnostic information extracted from this exciting technique” [4]. There was a need for development of specialized software tool for calculation of texture features in images that would be accepted and applied by all research groups and PhD students cooperating within the COST B11. Developing such software, called MaZda (acronym for co-occurrence matrix – Macierz Zdarzeń in Polish), is a contribution of Polish group. The program code has been written in C++ and compiled for a graphical user interface of Windows 9x/NT system.

## 2. TWO PATHS OF ANALYSIS

The first step of analysis is to load image file (Fig. 1.a). MaZda allows to open Windows bitmap files, unformatted binary files (RAW) and other file formats specific to magnetic resonance image scanners. The object of analysis is gray-scale image with 4 up to 12 bits per pixel.

There are two main paths of analysis. The first computes feature values within defined regions of interest (ROI) – the output data is a text table. The second computes features within a rectangular window sliding in steps over the image area – the output data is a collection of images (feature maps) in which a feature value is represented by a pixel luminance.



**Fig.1.** Block diagram of steps of analysis in MaZda software

## 2.1. Analysis Within ROI

Concerning the former path, the analysis starts by defining a ROI (Fig. 1.b). The user may select from the image up to 16 regions of interest using built-in drawing tools (lines, rectangles, circles and polygons shapes). Regions are marked with various, unique colors. They may be moved, copied, redefine, cleared or saved to a disk file for further use if necessary. There is no constraint imposed on the final shape of a region, it may be consistent or not, different regions may also overlap.

MaZda can compute more than 300 features. Not all of them are of user interest for every image. Moreover it takes time to compute them. The choice which group of features to calculate can be made by setting up the analysis options (Fig. 1.d).

Once the image is loaded, ROI selected and analysis options chosen, one can start the analysis process (Fig. 1.f). When the process stops the report form is shown with a new tab-page created that keeps the analysis result (Fig. 1.g). Every tab-page has a title giving the exact time at which the analysis was performed. Each page contains two panels, first with data header text and second with features table. Data header contains information on the type of image that have been analyzed and on the analysis options. Feature table includes all calculated texture features. Features calculated for individual regions of interest are grouped in columns; table rows carry values of particular features.

Analysis report can be saved into a disk file in text format or binary format.

## 2.2. Feature Maps

Feature maps calculation is a kind of image nonlinear filtration. Feature value, coded by a gray-level of single pixel in the output image, is calculated within the area of rectangular mask. The user may define vertical and horizontal dimensions of the mask (Fig. 1.e). During the analysis process, mask slides over the image surface by a given vertical and horizontal step in order to fill equally the output image with calculated features. A user can also define these step values. It should be taken into consideration that smaller steps increase the resulting image resolution but also increase the time of computation.

There are three groups of features, based on co-occurrence matrix, run-length matrix and gradient transform, that may create feature maps. The choice of features to compute can be made by setting up the analysis options in the feature maps tree control (Fig. 1.e).

Feature maps are presented in the feature maps window (Fig. 1.h). Every image is placed on its own tab-page. Every tab-page has a title stating what type of feature value was used for image creation. User can read a value at specified position of the image by pointing it with a mouse cursor. The feature value is shown on a status bar.

Additionally, a wavelet transform may be calculated. The image of the transform is presented in a separate tab-page of feature maps window.

Feature maps can be saved into a disk file in specifically developed binary format for storing floating point images.

## 3. COMPUTATION ALGORITHMS

The MaZda software was written to compute a variety of texture features based on image histogram, co-occurrence, run-length and gradient matrices [1,2], auto-regressive

model [5] and wavelet transform [7]. In this paragraph computation algorithms used in MaZda are briefly described.

### 3.1. Histogram-based features

Within defined ROI a histogram distribution of the image is computed. Assuming that  $h(i)$  is a normalized histogram value for a gray-level  $i$  and  $N$  denotes number of gray levels in the image – nine features are calculated: *Mean*, *Variance*, *Skewness*, *Kurtosis*, *Perc.1%*, *Perc.10%*, *Perc.50%*, *Perc.90%* and *Perc.99%*.

$$\text{Mean} = \mu = \sum_{i=1}^N ih(i) \quad (1)$$

$$\text{Variance} = \sigma^2 = \sum_{i=1}^N (i - \mu)^2 h(i) \quad (2)$$

$$\text{Skewnes} = \sigma^{-3} \sum_{i=1}^N (i - \mu)^3 h(i) \quad (3)$$

$$\text{Kurtosis} = \sigma^{-4} \sum_{i=1}^N [(i - \mu)^4 h(i)] - 3 \quad (4)$$

$$\text{Perc.K\%} = [k] \quad \left\{ k : \left( \sum_{i=1}^k h(i) \right) \left( \sum_{i=1}^N h(i) \right)^{-1} = K\% \right\} \quad (5)$$

### 3.2. Gradient-based features

For every pixel within defined ROI or within the sliding mask an absolute value of image gradient is calculated

$$G(x, y) = \sqrt{(i_{x+1,y} - i_{x-1,y})^2 + (i_{x,y+1} - i_{x,y-1})^2} \quad (6)$$

where  $i_{x,y}$  is pixel luminance at  $(x, y)$  position. Assuming that  $M$  is a total number of pixels within ROI, five gradient-based features are calculated: *GrMean*, *GrVariance*, *GrSkewness*, *GrKurtosis* and *Grads>0* – ratio of nonzero absolute gradient values to all gradient values calculated.

$$\text{GrMean} = M^{-1} \sum_{x,y \in \text{ROI}} G(x, y) \quad (7)$$

$$\text{GrVariance} = M^{-1} \sum_{x,y \in \text{ROI}} (G(x, y) - \text{GrMean})^2 \quad (8)$$

$$\text{GrSkewness} = M^{-1} \text{GrVariance}^{-\frac{3}{2}} \sum_{x,y \in \text{ROI}} (G(x, y) - \text{GrMean})^3 \quad (9)$$

$$\text{GrKurtosis} = M^{-1} \text{GrVariance}^{-2} \sum_{x,y \in \text{ROI}} (G(x, y) - \text{GrMean})^4 - 3 \quad (10)$$

### 3.3. Run length matrix based features

Let  $p(i,j)$  be the number of times there is a run of length  $j$  having gray level  $i$ . Let  $N_g$  be the number of gray levels and  $N_r$  be the number of runs. There are 5 run length matrix-based features computed for 4 directions of run (horizontal, vertical, 45° and 135° run).

$$ShrtREmph = C^{-1} \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^{-2} p(i, j) \quad (11)$$

$$LngREmph = C^{-1} \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 p(i, j) \quad (12)$$

$$GLEvNonUni = C^{-1} \sum_{i=1}^{N_g} \left( \sum_{j=1}^{N_r} p(i, j) \right)^2 \quad (13)$$

$$RLNonUni = C^{-1} \sum_{j=1}^{N_r} \left( \sum_{i=1}^{N_g} p(i, j) \right)^2 \quad (14)$$

$$Fraction = C \left( \sum_{j=1}^{N_r} \sum_{i=1}^{N_g} jp(i, j) \right)^{-1} \quad (15)$$

The coefficient  $C$  is defined as

$$C = \sum_{j=1}^{N_r} \sum_{i=1}^{N_g} p(i, j) \quad (16)$$

### 3.4. Co-occurrence matrix based features

Co-occurrence matrix, the second-order histogram, is a square matrix with  $N_g \times N_g$  dimensions;  $N_g$  is a number of image gray levels. The co-occurrence matrix element  $h_{k,l}$  is defined as

$$h_{k,l} = \sum_{\substack{x,y \in ROI \\ x+p,y+q \in ROI}} \left\{ \begin{array}{l} 1 \text{ if } (i_{x,y} = k \wedge i_{x+p,y+q} = l) \\ 0 \text{ elsewhere} \end{array} \right\} + \sum_{\substack{x,y \in ROI \\ x+p,y+q \in ROI}} \left\{ \begin{array}{l} 1 \text{ if } (i_{x,y} = l \wedge i_{x+p,y+q} = k) \\ 0 \text{ elsewhere} \end{array} \right\}$$

There are 11 features computed for a chosen parameters  $p$  and  $q$ . Definitions of these features can be found in [1].

### 3.5. Autoregressive model parameters

The autoregressive model assumes that pixel intensity is weighted sum of neighboring pixel intensities. In MaZda software the model given by:

$$i_{x,y} = \theta_1 (i_{x-1,y} - \mu) + \theta_2 (i_{x-1,y-1} - \mu) + \theta_3 (i_{x,y-1} - \mu) + \theta_4 (i_{x-1,y} - \mu) + \mu + e_s \quad (17)$$

where  $e_s$  denotes an independent and identically distributed noise and  $\mu$  is a mean of pixel intensities for a given ROI. Parameters  $\theta_1$ ,  $\theta_2$ ,  $\theta_3$  and  $\theta_4$  are computed within the ROI for minimal noise variance (*Sigma*).

### 3.6. Wavelet-based features

The discrete wavelet transform (DWT) for images is a linear operation that transforms  $2^n \times 2^n$  image matrix ( $n$  is positive integer value) into a matrix of the same dimensions. The transposed input image is multiplied by transform matrix; the result is transposed and multiplied once again by the same transform matrix. Half of transform matrix rows may be considered as smoothing filter (L) coefficients, while the other as sharpening filter (H) coefficients. Hence, the obtained matrix consists of four square quarters (marked as LL,

HL, LH and HH and called sub-bands). The DWT is computed repeatedly for LL sub-band until  $2 \times 2$  pixel square remains.

At the present, only a simple Harr wavelet transform is considered in MaZda software and only one feature (*Energy*) is calculated. The *Energy* feature is computed for each sub-band. Thus, the total number of feature values based on the wavelet transform may differ according to input image dimensions.

$$Energy = M^{-1} \sum_{x,y \in ROI} i_{x,y}^2 \quad (18)$$

where the parameter  $i_{x,y}$  is a resulting matrix element. Summation is performed for every pixel  $(x,y)$  located in the defined ROI and  $M$  is a total number of these pixels.

## 4. CONCLUSIONS

Techniques of texture features computation described here are quite well known. Despite this, there was a need for implementation of these methods in a single software tool. The MaZda software has been tested for magnetic resonance image analysis in several research-centers in Europe (Poland, France, Great Britain, Czech Republic, Denmark, Spain and Belgium) [4,6]. It was applied for kidney and liver image analysis, analysis of tumor texture; for texture classification (ROI analysis) and image segmentation (feature maps).

The MaZda software is accessible for downloading on the Internet page at:

<http://www.eletel.p.lodz.pl/cost/download.html>

## ACKNOWLEDGEMENTS

This work was performed within the framework of COST B11 European project.

## REFERENCES

- [1] Haralick R., Shanmugam K., Dinstein I., *Textural Features for Image Classification*, IEEE Transactions on Systems, Man and Cybernetics, **3**, 6, 1973, 610-621.
- [2] Haralick R., *Statistical and Structural Approaches to Texture*, Proceedings IEEE, **67**, 5, 1979, 786-804.
- [3] R. Lerski, K. Straughan, L. Shad, D. Boyce, S. Bluml and I. Zuna, MR Image Texture Analysis – An Approach to Tissue Characterization, *Magnetic Resonance Imaging*, **11**, 1993, 873-887.
- [4] *COST B11 – Quantitation of Magnetic Resonance Image Texture*, <http://www.uib.no/costb11/>
- [5] Materka A., Strzelecki M., *Texture Analysis Methods – A Review*, COST B11 report, Brussels, June 1998.
- [6] Lundervold A., *Feature maps*, COST B11 report, Heidelberg, October 2000 (<http://www.uib.no/costb11/>)
- [7] J. R. Williams, K. Amaratunga: Introduction to wavelets engineering, *International journal for numerical methods in engineering*, vol. 37 1994 , pp.2365-2388