Intelligent Medical and Diagnostic Systems

SYNTHESIS OF STATIC MEDICAL IMAGES WITH USING MACHINE LEARNING METHODS

Łukasz Piątek, Grzegorz Owsiany

***Abstract**: A part of collection of new algorithms for synthesis of melanocytic skin lesion images is briefly described. Presented approach in the developed algorithms based on a semantic conversion of textual description of melanocytic skin lesions into hybrid (vector-raster type) digital respective images of such lesions. It is assumed, that the developed methodology could be successfully used in education of dermatology students.*

***Keywords**: TDS, melanocytic skin lesion, image synthesis.*

***ACM Classification Keywords**: I.3.3 Computing Methodologies – Computer Graphics – Picture/Image Generation*

Introduction

In the past few years an increasing interest in images of melanocytic skin lesions can be clearly noticed. These images can be treated as a visual support in diagnosing of *malignant melanoma*, currently one of the most dangerous type of tumors [Topwik, 1998]. But on the other hand the lack of professional computer databases, containing images of such lesions, could be clearly noticed. This situation (at least in Poland) appears from owing the specific interpretation of the personal data protection act. Namely, the current interpretation of this law imposes the necessity for obtain patient permission not only for making in a hospital (or in clinic) a photo of a real lesion, but also approval for publishing or handing such image to any scientific research institution, i.e. specialized in image processing.

These above mentioned reasons inspired us to start the research over the development of effective algorithms for synthesis of static medical images, but generally to synthesize static images of melanocytic skin lesion. It seems, that the collection of developed algorithms – implemented in the specialized computer program (called *ImageSYNTHESIZER* [Piątek, 2010]) - allow to create large, multi-category informational database, which could be successfully used in teaching of medicine students.

Structure of the source database

In our initial research [Hippe and Piątek, 2005] we use informational (textual) database, already discussed in [Hippe, 1999], contains information about almost **550** real cases of anonymous patients lesions, confirmed by histopatological tests. At present, prepared database contains (***i***) **53** textual description of melanocytic lesions in combination with (***ii***) respective digital photographs of such lesions (Fig.1) obtained from [Triller et al., 2008]. To increase the dataset each image (Fig.2a) has been rotated by **90**, **180** and **270** degrees (Fig.2b-d) – such procedure provides a set of **212** real images of lesions.

a)

<Asymmetry>	<Border>	<Color>					<Diversity of structure>					<TDS>	<Diagnose>
symmetric lesion	0.0	absent	present	absent	absent	absent	absent	present	absent	present	absent	absent	1.5	Blue nevus

b)

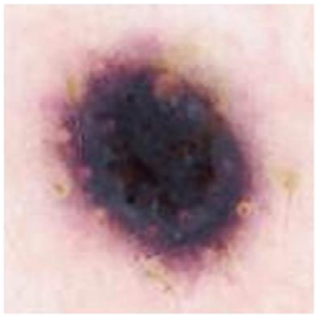

*Figure 1. An example **textual data vector** (**a**) and **digital photograph** (**b**) of real lesion (Dermal nevus type) [Triller et al., 2008], stored in informational database of melanocytic skin lesions [Hippe, Grzymała-Busse oraz Piątek, 2009a]*

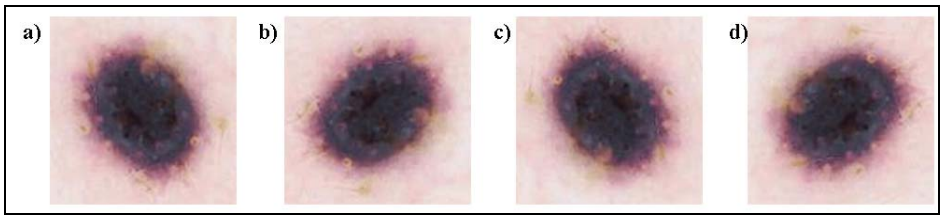

*Figure 2. Digital image (i.e. photograph) of real lesion (Dermal nevus type) (**a**) [Triller et al., 2008] and **3** additional images (**b**, **c** and **d**), obtained by applying the operation of four-fold symmetry axis, perpendicular to the plane of the image*

Each textual description of the lesion (Fig.1a) is saved in the form of 15-th component data vector, which values transmit information about presence or lack of specific symptoms of a give lesion. These symptoms – in machine learning called *descriptive attributes* – are (*i*) *asymmetry*, (*ii*) *border*, (*iii*) *color* and (*iv*) *diversity of structure*. In fact, classification of every real case formally relies on the application of the ABCD *rule* [Stolz et al., 2006], in which A (*Asymmetry*) shows a result of evaluation of lesion shape's asymmetry, B (*Border*) – estimates the character of the rim of lesion, C (*Color*) – stands for number of colors (one or more, from six allowed) and D (*Diversity of structure*) – identifies number of structures (one or more, from five allowed). Elements of ABCD *rule* enumerate four main symptoms of the mentioned lesion, and these element are used to compute value of the TDS (*Total Dermatoscopy Score*) parameter [Braun-Falco et al., 1990], according to the Eq. (1) presented below:

$$TDS = 1.3 * Asymmetry + 0.1 * Border + 0.5 * \Sigma\, Color + 0.5 * \Sigma\, Diversity\ of\ structures \quad (1)$$

For example, for a case described by a vector with values presented below:

- ***A**symmetry* equal to **symmetric change**,
- ***B**order* equal to **0**,
- ***C**olor* equal to **four** selected colors observed in a lesion, and
- ***D**iversity of structure* equal to **four** selected structures presented in a lesion

the **TDS** parameter value equals to 4.0 (Eq. (2)):

$$\mathbf{TDS} = \mathbf{1.3} * 0 + \mathbf{0.1} * 0 + \mathbf{0.5} * (0+0+1+1+1+1) + \mathbf{0.5} * (1+1+1+1+0) \quad (2)$$

Due to the **TDS** value (Table 1) the analyzed lesion can be assigned to one of four allowable categories of melanocytic skin lesion, namely: *Benign nevus, Blue nevus, Suspicious nevus* or *Malignant melanoma*.

***Table 1**. Classification of melanocytic lesions in dependence of **TDS** value*

TDS value	Lesion classification
TDS < **4.76** and **lack** of blue color	*Bening nevus*
TDS < **4.76** and blue color is **presented**	*Blue nevus*
4.76 <= **TDS** < **5.45**	*Suspicious nevus*
TDS >= **5.45**	*Malignant melanoma*

Methodology of the research

In our initial research [Hippe and Piątek, 2005] the algorithms of transformation of a single case from the textual informational database [Hippe, 1999] allowed to obtain only one synthetic lesion image. At present, developed algorithms are improved to generate the exhaustive number of synthesized images [Kulikowski, 2005], corresponding to symptoms displayed by a given lesion. The set of new developed algorithms define the hybrid (*vector-raster* type) approach to synthesis of medical images. Precisely, developed methodology determine a connection of (***i***) *vector* type procedures applied in graphics, combined with (***ii***) *raster* graphics operations, and (***iii***) elements of *machine learning* methods. A sequence of application of the developed algorithms is divided into separate modules – i.e. (i) *learning* module and (ii) module of specify *synthesis process* (Fig.3). In this paper the attention is focused only on part related to the algorithms of mapping lesions' asymmetry (processes no. 2 and no. 5 at the Fig.3) .

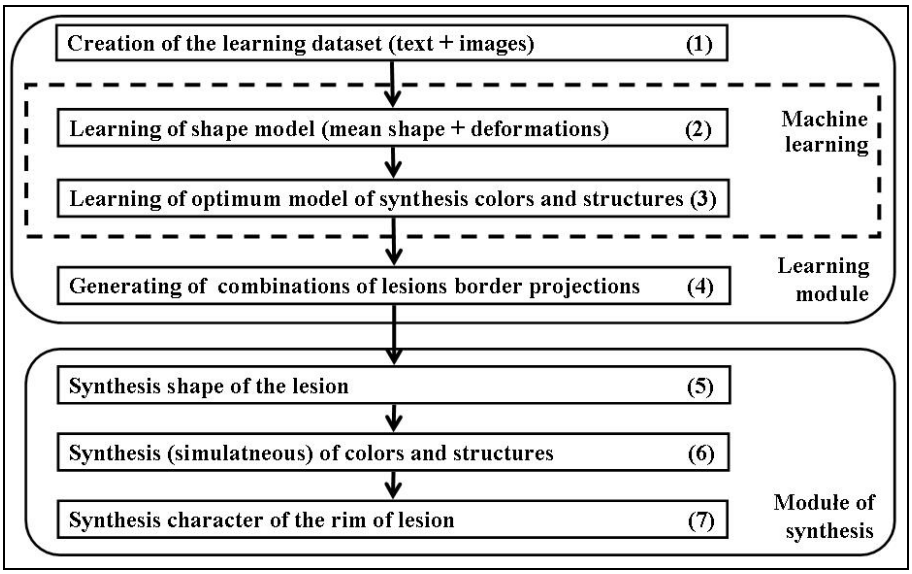

Figure 3. A sequence of application of the developed algorithms

It should be stressed, that in our research we concern on synthesis of melanocytic skin lesion images from only two most dangerous groups of such lesions, i.e. *Nevus* and *Melanoma* [Stolz et al., 2006]. Precisely, Nevus group contains five types of lesions, i.e. *Junctional nevus, Junctional and dermal nevus, Atypical/dysplastic nevus, Dermal nevus* and *Palmo-plantar nevi*. *Melanoma* group include two types of lesions – *Superficial Melanoma* and *Nodular Melanoma*.

Synthesis of the lesions' asymmetry

Mapping asymmetry of a lesion (i.e. shape of a lesion) in synthetic image is divided into separates module as follows at Fig. 4, right into (***i***) *learning* module and (***ii***) module of specify *synthesis process*.

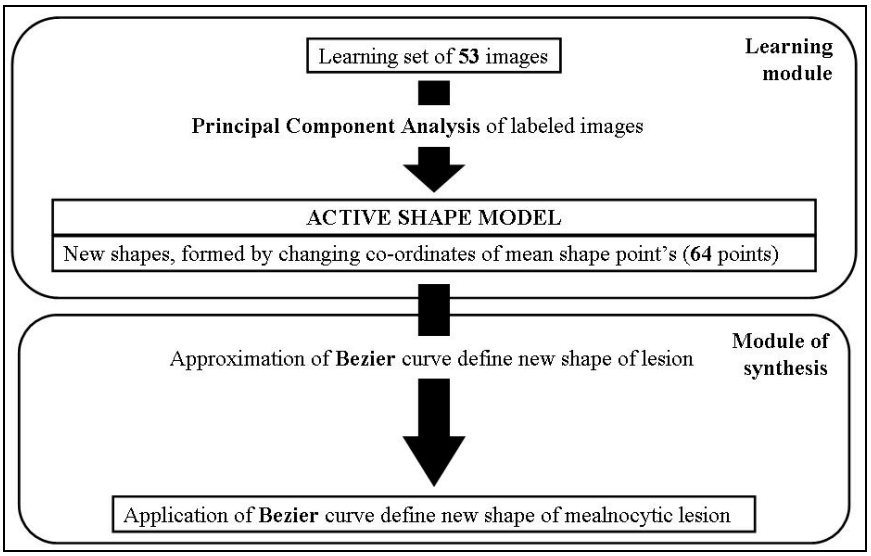

Figure 4. A sequence of selected operations in process of synthesis of an image of lesion shape's asymmetry

At the first module the literature algorithm called ***ASM*** (***A****ctive* ***S****hape* ***M****odel* [Cootes et al., 1994]) is applied. ***ASM*** could be define as a structural information about the *mean shape* of object placed onto digital images, joined with information about the permitted *deviation* from the mean shape, observed in the learning set of images. ***ASM*** models are obtain by statistical analysis of *point distribution model* ***PDM***, based on the set of points labeled manually onto the set of all learning images, with the required conditions. Such control points (so called landmarks) of each training image had to represent a required correspondence (Fig.5).

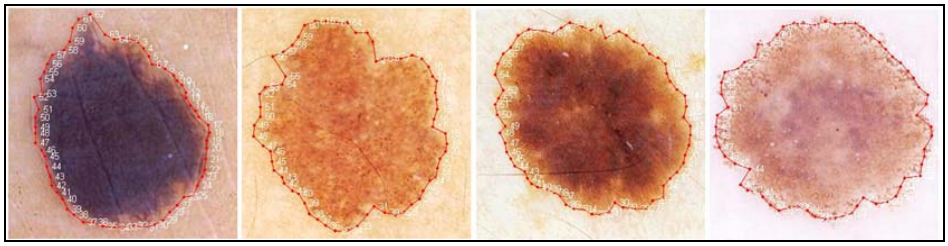

Figure 5. Selected images from the training set (for symmetric lesions) – each with marked 64 landmark points

In other words, every shape ***X*** from the training set is represent as an n-point polygon in images coordinates $\mathbf{X} = (x_1,y_1,\ldots,x_n,y_n)^T$. Every point with coordinates equal (x_n,y_n) for n equal form 1 to 64is defined in the place of intersections of 32-fold symmetry axis with an edge of the lesions' shape. Finally, each new shape can be obtain according to the Eq. (3):

$$X = \Gamma + P_t * b \tag{3}$$

where:

Γ - it is a mean shape of all images from the training set,

$\mathbf{P_t = [u_1,u_2,\ldots,u_t]}$ - includes t first eigenvectors of the model covariance matrix, and

$\mathbf{b} = [b_1, b_2, \ldots, b_t]^T$ – it is so called shape model vector, contains shape model parameters for each of the selected eigenvectors.

Components of the shape model vector $\mathbf{b}_i$ can be fixed according Eq. (4) within the following range:

$$-s * \sqrt{\lambda_i} \geq b_i \leq s * \sqrt{\lambda_i} \tag{4}$$

where:

λ_i – it is a selected eigenvalue correspond to eigenvector $\mathbf{u}_i$, and

s – it is a constant value, in our research experimentally equal to **1**.

Active Shape Model – i.e. mean shape and permitted deformation of this shape, observed in the set of training images - for symmetric lesions are presented at Fig.6-9.

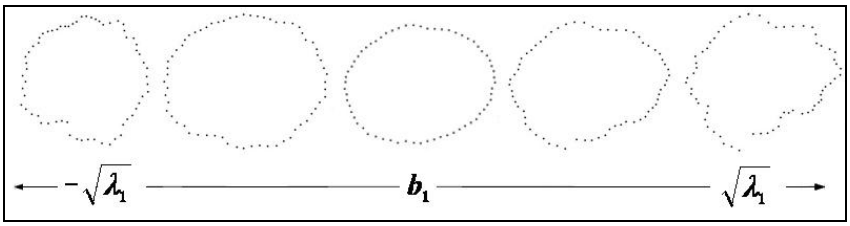

Figure 6. Changing of the mean shape for symmetric lesion, by changing value for the first eigenvector u_1 (1st eigenvector contains 47.9% information about the symmetric lesion shape)

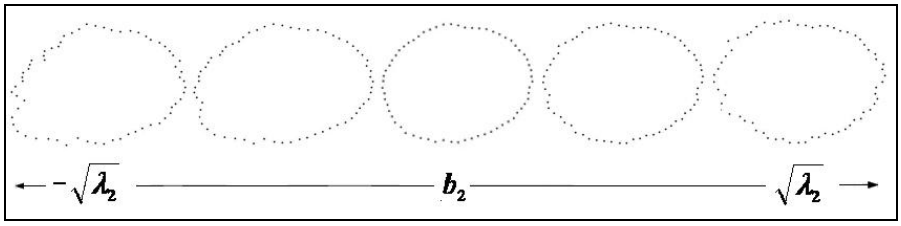

Figure 7. Changing of the mean shape for symmetric lesion, by changing value for the second eigenvector u2 (2nd eigenvector contains 16.8% information about the symmetric lesion shape)

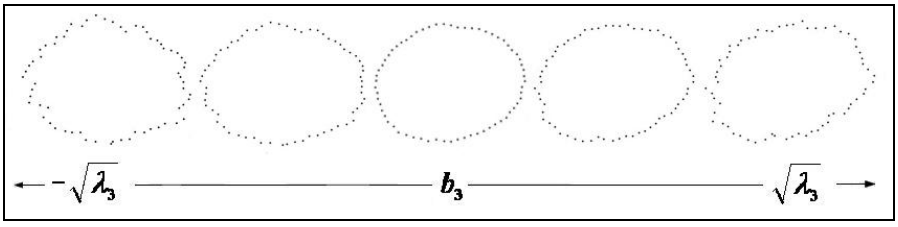

Figure 8. Changing of the mean shape for symmetric lesion, by changing value for the third eigenvector u3 (3rd eigenvector contains 10.8% information about the symmetric lesion shape)

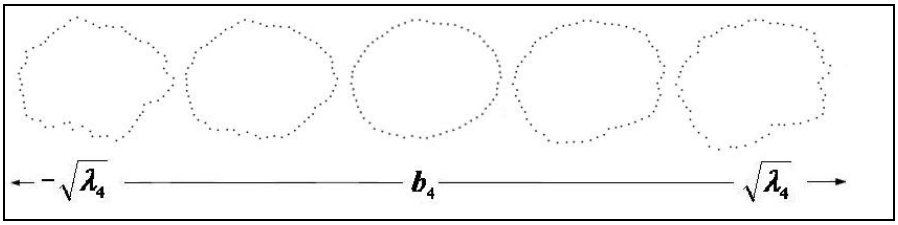

Figure. 9. Changing of the mean shape for symmetric lesion, by changing value for the fourth eigenvector u4 (4th eigenvector contains 8.8% information about the symmetric lesion shape)

Based on the connections of **64** control points of shape **X** (calculate according to the Eq. (3)) the control polygon of the curve is defined. Finally – in the process of synthesis of lesions shape – by using ***the Casteljau*** algorithm [Matusiewicz, 2008] each segment line between neighboring landmark points of the control polygon is splitted with a fixed ratio **t / (t-1)**. This process is performed until arriving at the single point of a curve, corresponding to the parameter **t**. Curve generated with this method is called as a ***Bezier*** curve. According to the developed methodology of synthesis lesion asymmetry we can achieve reliable shapes of synthetic images, similar to shapes of the real melanocytic skin lesions. Example shapes for various types of lesions' asymmetry – i.e. *symmetric lesion, one-axial asymmetry lesion* and *two-axial asymmetry lesion* – is presented below at Fig.10.

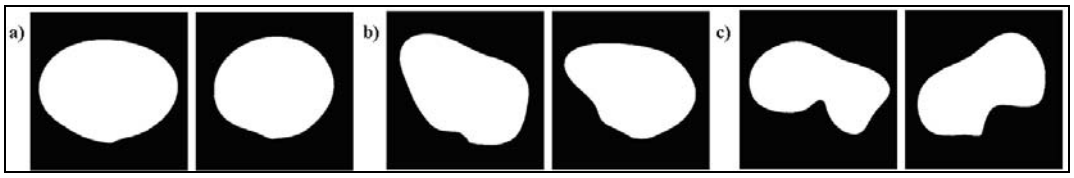

Figure 10. Examples of shapes defined with Bezier curve for symmetric lesion (a), one-axial asymmetry lesion (b) and two-axial asymmetry lesion (c)

Time complexity

Time complexity [Papadimitriou, 2002] of algorithm of mapping lesion's asymmetry for (***i***) learning module (see Eq. (5)) and for (***ii***) module of synthesis process (see Eq. (6)) is presented below:.

$$O(M^2 (M + N)) \quad (5)$$

$$O(N (M + N)) \quad (6)$$

where:

M – is a number of shapes from the set of learning images, equal 22 for *symmetric lesions*, 13 for *one-axial asymmetry lesion* and 18 for *two-axial asymmetry lesion*, whereas

N – is a number of control points, labeled onto every image (shape) from the training set. Value of N is constant and equal to 64.

It could be assumed, that those values of time complexity allow to generate synthetic images of lesion in statu nascendi, it means at the time when physician can need such images.

Conclusion

In research described here, we succeeded in obtaining random synthesis of the most important symptom of melanocytic skin lesions, namely ***A****symmetry*. Combination of algorithms described here and also discussed in [Hippe et al., 2009] follow a new approach to hybridization of synthesis of static medical images (*vector-raster* type), but in general images of melanocytic skin lesions. The set of developed algorithms – already implemented

in specialized computer program (called ***Image*SYNTHESIZER** [Piątek, 2010]) – could be used to create large, multi-category informational database, which can be useful in teaching of medicine students, but also in a practice of less experienced dermatologists.

In our future research we consider two main directions. The first problem had to be solved is connected with synthesis of asymmetry. Precisely, apart from synthesis shape's asymmetry, we consider develop additional algorithms for synthesis asymmetry of location selected colors and structures, presented in a given lesion. The second improvement, based on adding the 3rd dimension of synthesize images, should allow to synthesize more reliable images of various type of melanocytic skin lesion (i.e. nipple lesions).

Bibliography

[Cootes et al., 1994] T.F.Cootes, A.Hill, C.J.Taylor, J.Haslam. *The Use of Active Shape Models for Locating Structures in Medical Images*. In: Image and Vision Computing 12(1994,6)355-366.

[Hippe, 1999] Z.S.Hippe. *Computer Database 'NEVI' on Endangerment by Melanoma*. In: TASK Quarterly 3(1999,4)483-488.

[Hippe et al., 2009] Z.S.Hippe, J.W.Grzymała-Busse, Ł.Piatek. *From the Research on Synthesis of Static Images of Melanocytic Skin Lesions*. In: E.Kącki, M.Rudnicki, J.Stempczyńska (eds.) Advances in Intelligence and Soft Computing 65 (*Computers in Medical Activity*). Springer-Verlag, Berlin-Heidelberg (2009), pp. 223-229.

[Hippe and Piątek, 2005] Z.S.Hippe, Ł.Piatek. *From research on the database of simulated medical images*. In: Proc. of the Conference *Databases for Science and Technology*. Gdańsk (Poland), 2005, pp. 29-38.

[Kulikowski, 2005] J.Kulikowski. *The foundations of the structural description of distracted databases of expert knowledge*. In: Proc. of the Conference *Databases for Science and Technology*. Gdańsk (Poland), 2005, pp. 25-27.

[Matusiewicz, 2008] Z.Matusiewicz. *Matematyka dla grafików komputerowych* (in Polish). In: Publishing house of the University of Information Technology and *Management*. Rzeszów (Poland), 2008.

[Papadimitriou, 2002] C.H. Papadimitriou. *Złożoność obliczeniowa* (in Polish). WNT, Warszawa (Poland), 2002.

[Piątek, 2010] Ł.Piątek. *Synthesis of static medical images, by an example of melanocytic skin lesions*. PhD Dissertation. Gdansk University of Technology, Gdańsk, 2010.

[Stolz et al., 2006] W.Stolz, O.Braun-Falco, P.Bilek, M.Landthaler, W.H.C.Burgdorf, A.B.Cognetta. *Altas dermatoskopii*. In: D.Czelej, M.Michalska-Jakubus, M.Ziarkiewicz, Czelej Sp. z o.o., Lublin (Poland), 2006.

[Topwik, 1998] E.Topwik. *WHO Melanoma Programme. Rozpoznanie i kliniczne leczenie czerniaka* (in Polish). In: Publishing house of Polish Foundation of European Oncology School, Warszawa (Poland), 1998.

[Triller et al., 2008] R.Triller, G. Aitken, T.Luc. Atlas of Dermatoscopy of Pigmented Skin Tumors. http://www.pless.fr/derma-toscopie/ (2008).

Authors' Information

Łukasz Piątek *– adjunct in (i) Department of Expert Systems and Artificial Intelligence and (ii) Institute of Bioinformatics, University of Information Technology and Management; Sucharskiego 2 Str., 35-225 Rzeszow, Poland ; e-mail: lpiatek@wsiz.rzeszow.pl*

Major Fields of Scientific Research: Bioinformatics, Computer Vision and Graphics, Artificial Intelligence

Grzegorz Owsiany *– assistant lecturer in Department of Mathematics and Informatics, University of Information Technology and Management; Sucharskiego 2 Str., 35-225 Rzeszów, Poland ; e-mail: gowsiany@wsiz.rzeszow.pl*

Major Fields of Scientific Research: Algorithmic