Semantic simulation of the strangeness in an artificial immune system

1. Randomized dynamic generation of selected features of melanocytic skin lesions

Zdzisław S. Hippe[1)], Jerzy W. Grzymała-Busse[1,2)], Łukasz Piątek[3)]

[1)] Department of Expert Systems and Artificial Intelligence, University of Information Technology and Management, 35-225 Rzeszów, Poland
[2)] Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045, USA
[3)] Department of Distributed Systems, University of Information Technology and Management, 35-225 Rzeszów, Poland

E-mail: {ZHippe, LPiatek}@wenus.wsiz.rzeszow.pl, {Jerzy@ku.edu}

Keywords: dynamic generation, melanocytic lesions, random image building

Abstract

In this paper, the methodology of generating images of melanocytic skin lesions is briefly outlined. The developed methodology proceeds essentially in two steps. In the first one, semantic description of skin lesions of anonymous patients is carefully analyzed to catch important features (symptoms) and to mine their logical values. Then, data gained in this step, were used to control a specific simulation process, in which the simulated lesion's image is randomly put together from *a priori* pre-defined fragments (textures). In this way, a single textual vector representing a distinct lesion, can produce a collection of several images of a given category. The quality of simulated images, verified by an independent expert (R.K.), was found to be quite satisfactory.

1. Introduction

It is roughly two years since we begun in our group research devoted to generation of simulated images of melanocytic skin lesions and their classification [1]. Real lesions are treated as elements of the strangeness within the regular healthy skin. However the immune system itself, owing to its inherent properties, undertakes stepwise action in order to minimize their influence, or even to reject them. Effectively functioning immune mechanisms may deactivate foreign cells, however, it is generally very difficult for the immune system to prevent transformation of a *suspicious* lesion into melanoma *malignant* [2]. For that reason, the ability to diagnose correctly melanocytic skin lesions plays the fundamental role in conducting of the effective treatment.

2. Methodology of the research

The main goal of our research was to develop a computer program system suitable for the reliable classification of melanocytic skin lesions. Due to personal data protection act, both making and publishing of real photographs of melanocytic lesions requires patients' approval, which obstructs in many cases creating informational databases, having instructional character, for less experienced medical doctors. It was assumed, that by the application of the developed in our research generator of images can intent on the constraint of the use of real digital pictures in favour of simulated images, representing symptoms of melanocytic lesions with required precision. To accomplish our intention, we run the research on the semantic conversion of textual informational database into a base containing respective images of lesions. The source (textual) databased, described in [3], concerns real cases of the investigated disease; it contains vector description of four categories (classes) of melanocytic lesions (namely: *Benign nevus, Blue nevus, Suspicious nevus* and *Melanoma malignant*) for 548 anonymous patients. Each case is described by a vector with 15 elements; their values transmit information about presence or lack of specific symptoms of a given lesion. These symptoms (in machine learning language called *descriptive attributes*) were: the type of *asymmetry*, character of the *border* of a lesion, combination of *colours* and *structures* observed, the value of the **TDS**-parameter (**T**otal **D**ermatoscopy **S**core) and category to which each case has been classificated. The attribute <***Asymmetry***> defines the symmetry of a lesion along two axes crossing at a slant of right angle [4]. Logical values of this attribute can be: *symmetric change* (numeric value in the base = **0**), *one-axial asymmetry* (numeric value = **1**) and *two-axial asymmetry* (numeric value = **2**). Definition of the character of a rim of lesion relies on its splitting into eight equal parts by four axes

crossed in a point, and assigning 0 or 1, if the border between a lesion and the skin is diffuse or sharp, respectively. This causes, that value of the attribute <***Border***> oscillates between **0** and **8**. At the same time <***Colour***> can have six allowed values: *black*, *blue*, *dark-brown*, *light-brown*, *red* and *white*. The attribute <***Structure***> can have 5 logical values: *branched streaks*, *pigment dots*, *pigment globules*, *pigment network* and *structureless areas*. All attributes related to colour and structure of a lesion, have assigned in the base **0** (absence) or **1** (presence). The value of the 14-th element, i.e. **TDS** parameter is computed according to the ABCD rule [5]:

TDS = 1,3 * <Asymmetry> + 0,1 * <Border> + 0,5 * Σ <Colour> + 0,5 * Σ <Structure>

The 15-th element of the anonymous patient vector contains the information about the class of the disease, fixed by an expert (an experienced medical doctor, thoroughly trained in dermatology). In this paper, simulation algorithms for mapping of <***Asymmetry***> and <***Border***> of skin lesion are briefly dealt with, whereas the simulation of <***Colour***> and <***Structure***> was already described in [1].

3. Random selection of chosen features of lesions

Methodology of generation of images elaborated by us consists of the composition of pre-defined fragments of images of melanocytic lesions; these fragments are called here textures. Such procedure seems to be quite effective in relation to two characteristic attributes of images i.e. <colour of a lesion> and <diversity of structure>. The simulation of <**asymmetry**> of lesions and character of their rim (<**border**>) requires the special approach, based on random selection of allowed logical values for these attributes and combining them in an exhaustive way into a **set** of simulated images. In this way, each anonymous patient vector from the source (textual) database creates a collection of simulated images. These images, according to Kulikowski [6], should be treated as synonyms in the field of images.

4. Simulation of lesion's asymmetry

The process of mapping the type of asymmetry consists in generating the lesion's image by a combination of four basic structures, created from *a priori* prepared parts (of the size of 200*200 pixels) (see Fig. 1):

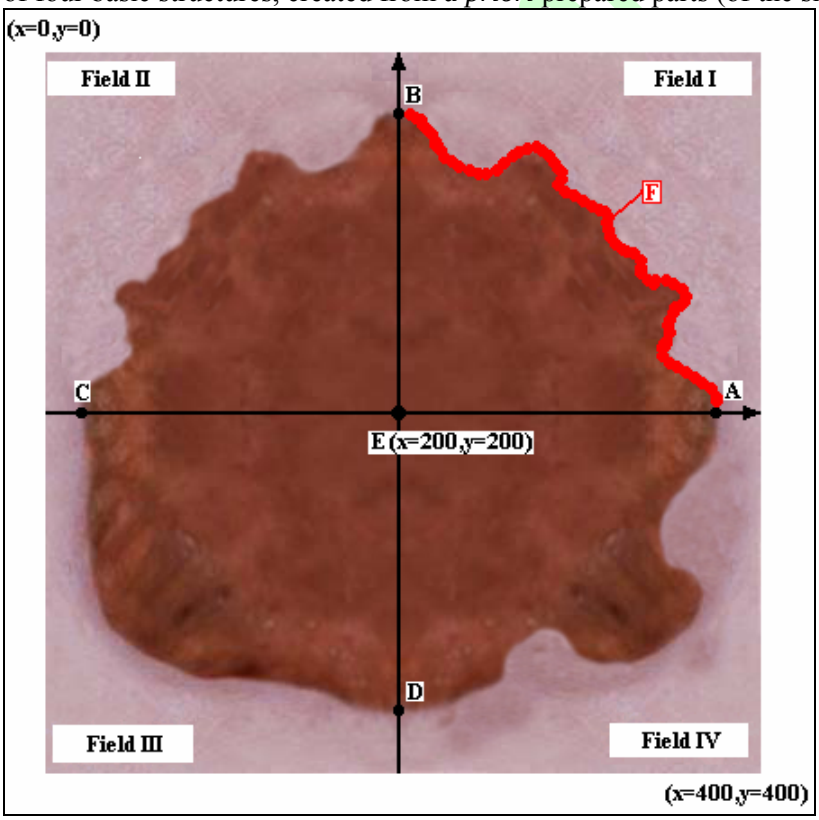


Fig. 1. Partition of an image and arrangement of its parts

All attributes of a texture that stands for a constructing part (a quarter) of an image, are different due to the diversity of lesion asymmetry. It should be mentioned, however, that the four components of a simulated image differ each other only by a shape of the curve **F**. Usually, this shape is mirrored, while comparing,

for example part I and part II (see Fig. 1). Due to various shapes of the curve **F** there are 6 available types of construction elements, responsible for the simulation of asymmetry (see Fig. 2).

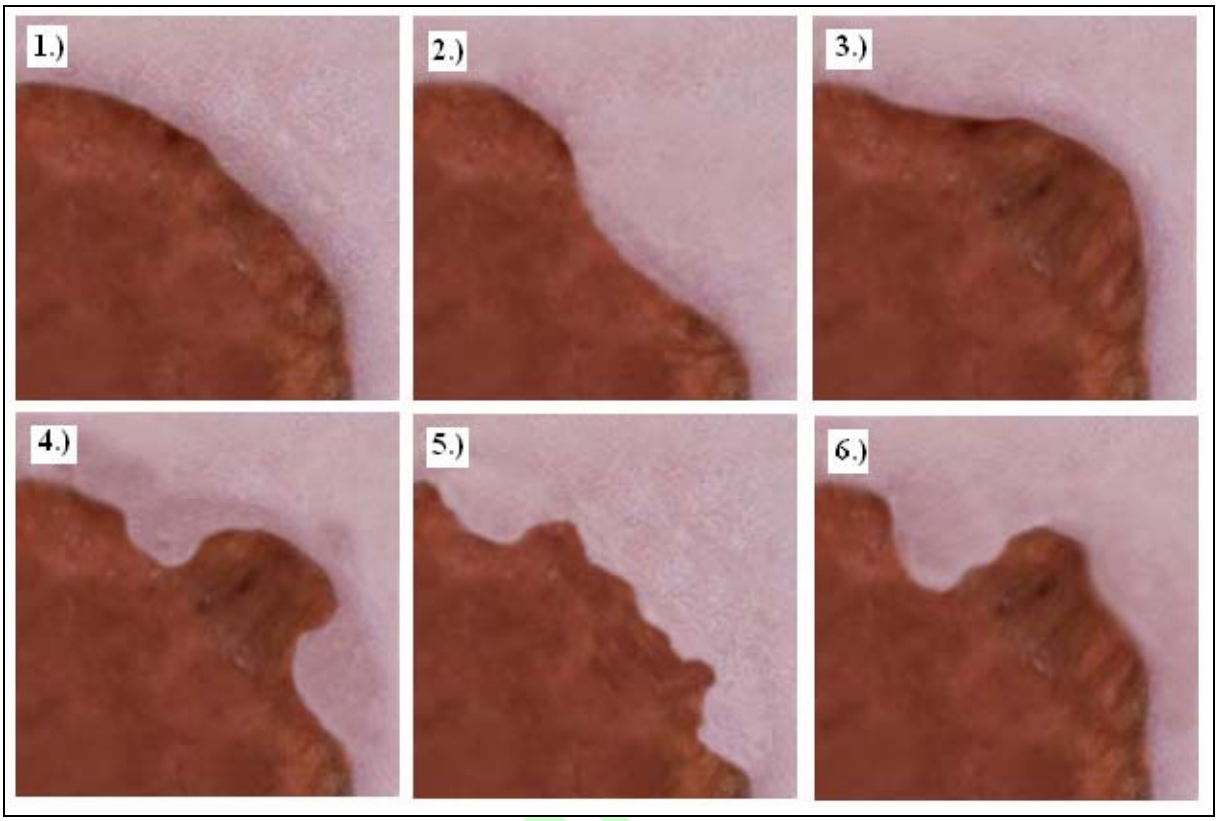


Fig. 2. Pre-defined fragments of images (quarters), with a different form of the curve F

Further, there are three possible ways of mapping of the attribute <***Asymmetry***>, depending on its real value. In the case of <*symmetric change*> there is only one way of mapping a fragment randomly selected out of two possible fragments (various shapes of the curve **F**), which is next placed in each of four fields of the main square. Then, for <*a lesion asymmetric along one axis*>, after selection of two different parts, one of them is repeated onto the first and the fourth field, while the second part is put onto the two remaining fields labelled **II** and **III**. Finally, for <*a lesion asymmetric along to perpendicular axes*>, three different construction parts (textures) are chosen, and one of them (randomly selected) is repeated onto the fields labeled **I** and **II**, whereas two remaining fragments are randomly put in the fields **III** and **IV**. Some consecutive steps of the simulation of <*asymmetry*> are shown in Fig. 3, 4 and 5.

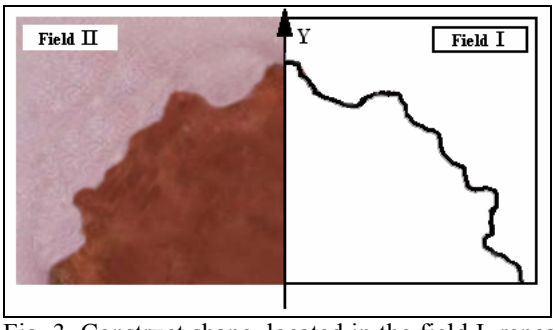


Fig. 3. Construct shape, located in the field I, repeated onto the field II (an operation of two-fold symmetry axis was used)

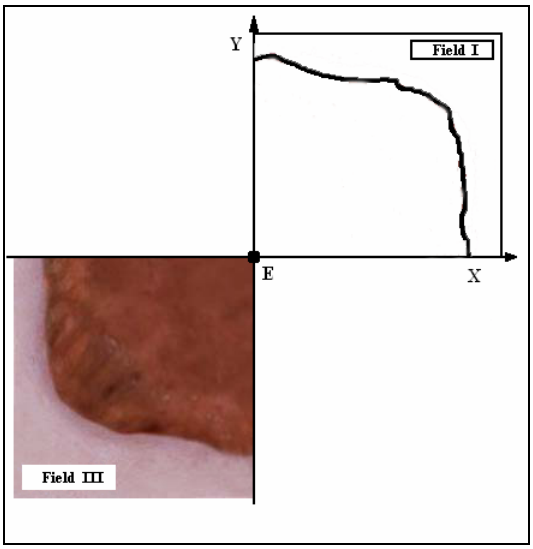


Fig. 4 Construct shape, located in the field I, repeated onto the field III (an operation of symmetry center was applied)

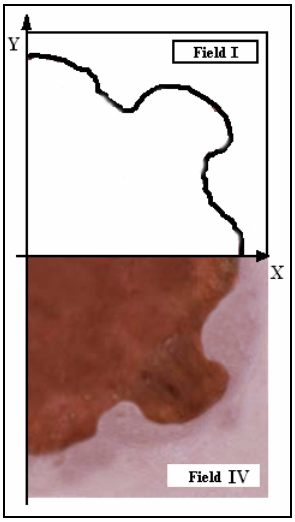
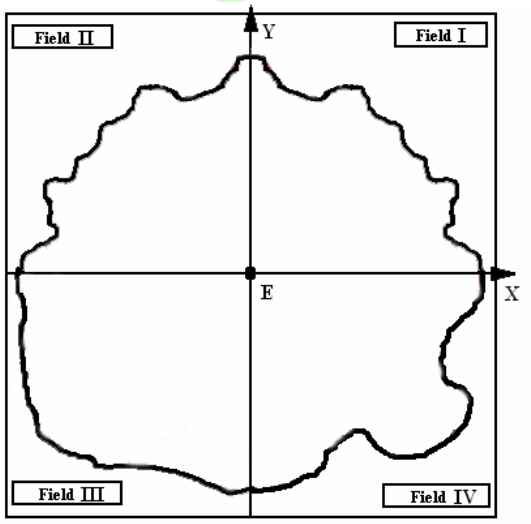


Fig. 5 The operation of two-fold symmetry axis executed along the axis X

After placing all four construct fragments in particular fields of the main square, we receive an image characteristic for the type of asymmetry, described by a given textual data vector in the source database (Fig. 6).



Picture 6. Way of placement of particular textures onto the simulated image

5. Simulation of lesion's border

Diagnosing of the symptom <*border*> by medical doctors relies on splitting of a lesion into eight regular parts, and then on counting how many of them displayed sharp transition towards the skin (count=1), and how many displayed diffuse transition (count=0). In this way, the numerical value of this symptom is in the range <*border*> $\in$ (0,8). However, the distribution of possible "isomers" of transforms rises up to 218 possible combinations of sharp/diffuse transitions. The exhaustive collection of all possible combinations of border transitions are shown in Fig. 7.

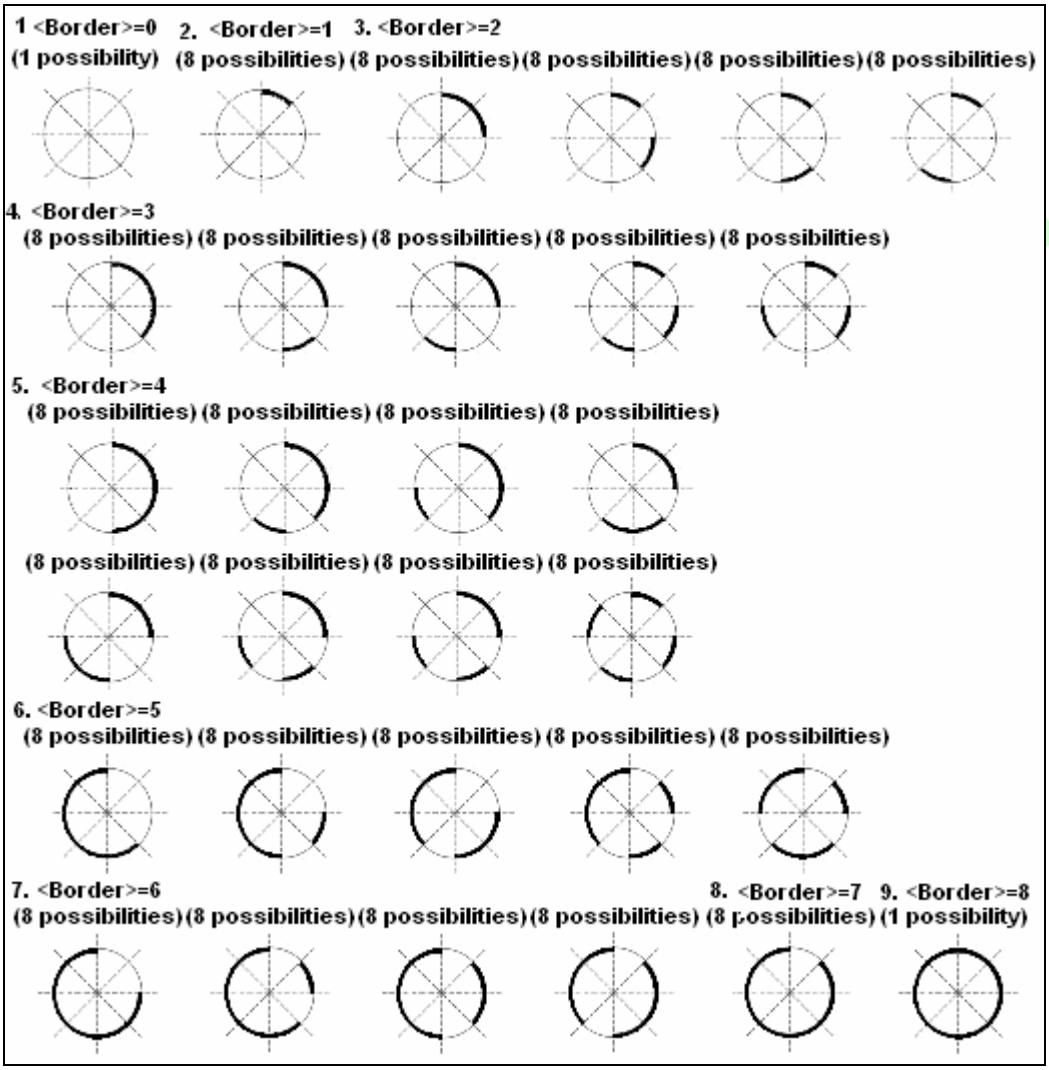


Fig. 7. Exhaustive collection of sharp/diffuse combinations of fragments of melanocytic lesions. Sharp transition from a lesion towards the skin is displayed here by means of a thick line, whereas thin line represents "fuzzy" transition. Each combination of transitions (except for *border*=0 and *border*=8) is multiplied by 8, because a set of 8 new transition can be generated applying the operation of eight-fold symmetry axis, perpendicular to the plane of the figure. Transitions for the *border*=1 and *border*=8 can be treated as distinct representations, applying the approach of Schoenflies points groups [7].

All of them are then applied in the superposition with previously simulated asymmetry of melanocytic lesions.

6. Program implementation

Recent implementation of the developed generator of simulated images of melanocytic lesions is based on the language **PHP** (**P**rogramming **H**ypertext **P**re-processor), combined with the use of graphic library **GD** [8]. Pre-defined textures, necessary for the reasonable simulation (202 in number) of lesions, were defined in **PNG** (**P**ortable **N**etwork **G**raphics) format, each texture contains various number of combinations of descriptive features. Programming code dynamically generates website, allowing the user (e.g. trainee medical doctor) to choose one of three options:

- **study** – (both simulated image of a lesion and the respective data-vector are shown),
- **quiz** – (only generated image is shown, and trainee medical doctor, using his/her knowledge and experience, determines symptoms of the analyzed lesion and classifies it), or
- **pictures** – (here, real (non-simulated) digital photograph(s) of melanocytic lesions, accompanied with an expert diagnosis, are displayed).

All these options were carefully evaluated; simulated images were thoroughly tested in a demanding estimation procedure, executed by an expert (dermatologist).

7. Summary and conclusion

In research described here, we succeeded in obtaining random simulation of the two important symptoms of real melanocytic skin lesions, namely: <***Asymmetry***> and <***Border***>. The developed algorithm enables to generate the exhaustive number of simulated images, corresponding to symptoms contained in a given lesion, originally described by a textual vector from the source database. It was found, that simulated images considerably broaden the informational source database, and can be successfully used in the process of training less experienced medical doctors.
The next problem to be solved in our future research, seems to be an attempt to make a combination of already used way of simulation of colours and structural diversity of lesions, with the approach based on the randomized superposition of those symptoms into the simulated images.

8. Acknowledgements

We would like to express our gratitude to Mr. Robert Kijowski, for his helpful comments and expert assistance in the research.

9. References

[1] Z.S. Hippe, Ł. Piątek:
From research on the database of simulated medical images
Proc. The Conferene Databases for Science and Technology, Gdańsk 25-27.09.2005, pp. 225-230 (in Polish).

[2] J. Gołąb, M. Jakóbisiak, W. Lasek:
Immunology
Polish Scientific Publishers PWN, Warsaw 2004 (in Polish).

[3] Z.S. Hippe:
Computer Database 'NEVI' on Endangerment by Melanoma
TASK Quarterly 1999(3,No4)483-488.

[4] S. Bajcar, L. Grzegorczyk:
The Atlas of Diagnostics of Melanocitic Lesions
Jagiellonian University Editorial Office, Cracow 2002, pp. 14, 28 (in Polish).

[5] O. Braun-Falco, W. Stolz, P. Bilek, T. Merkle, M. Landthaler:
Das dermatoskop. Eine Vereinfachung der Auflichtmikroskopie von pigmentierten Hautveranderungen.
Hautarzt 1990(40)131.

[6] J.L. Kulikowski:
The foundations of the structural description of distracted databes of expert knowledge (in Polish)
Proc. The Conference Databases for Science and Technology, Gdańsk 25-27.09.2005, pp. 29-38 (in Polish).

[7] R.A. Alberty, R.J. Silbey:
Physical Chemistry
J. Wiley & Sons, Inc., New York, 1992, pp. 426-430.

[8] http://www.boutell.com/gd/.