

# Nonlinear model of qualitative classification on base of wave-packet decomposition in space of Volterra kernels.

Akhmetshin A.M., Shutin D.V.

Dnepropetrovsk State University

Email: akhm@mail.dsu.dp.ua, faria@email.dp.ua

## ABSTRACT

In the given paper the new approach to the usage of nonlinear models for qualitative classification has been presented. The method is oriented on classification of texture images and stochastic signals. The wave packet decompositions of the initial sequence were interpreted as input and output signals of virtual nonlinear system. The peculiarities of nonlinear transfer characteristics were than considered for performing a qualitative classification. This technique was applied to classification of gastritis on the base of IR spectrograms of the mucous membrane, x-ray CT and x-ray images of the bone trabecular structure, affected with osteoporosis. It was shown, that the proposed method requires much smaller resolution and it is also more sensitive to the changes in the initial sequence.

**Keywords:** Nonlinear identification, Volterra kernels, wave-packet decomposition.

## 1. INTRODUCTION

Stochastic processes (signals) find their application in almost every natural system. This fact very often makes them an object of research. In terms of information theory, signal acquired from the system can be interpreted as a message, transmitted to the receiver. The receiver's main role is to "read" the message correctly, in other words, to identify some distinctive characteristics of the message (signal) and classify it. The correct classification of the information received plays one of the most important roles in further processing. In case of medical data processing we would most often speak of correct diagnostics. Therefore, choosing the proper method of signal classification would increase the probity of the results received and quality of the analysis as a whole.

Current methods of data classification are based on linear and nonlinear models. Linear models are most often simple enough. They can be easily simulated and analyzed, but these models are quite rough and unsuitable for the complex processes. On the other hand, the usage of nonlinear models lacks simplicity, but it does provide an exactness and accurateness of the results obtained.

In the given paper we presented a method of

qualitative classification of the stochastic signals and images, which is based on nonlinear models and wave-packet decomposition.

## 2. METHOD

Let's consider three grained texture images (Fig.1.). To increase the sensitivity of the processing procedure, it was proposed to analyze the texture image by the cut. The cut should preserve the spectral information of the image however it will lose information of any objects that are on the texture.

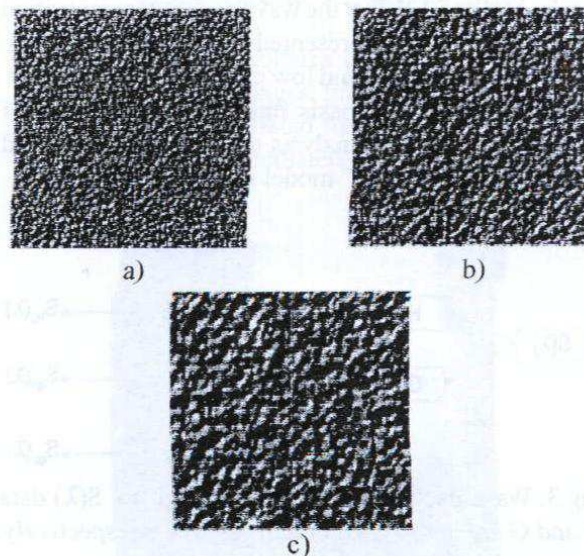


Fig.1. Three grained textures.

The taken cuts  $S(\lambda)$  of the presented images are shown on Fig.2. In case of stochastic signals, it would be convenient to use AR models[1] for describing them, i.e.

$$S(\lambda) = \sum_{i=1}^M P_i S(\lambda_{k-i}) + a(\lambda_{k-i}) ; k=1, \dots, N. \quad (1)$$

In terms of the model (1), observable data  $S(\lambda)$  can be interpreted as an output of a linear system excited with a stationary white noise  $a(\lambda_k)$ . This approach is very efficient when the processes, being simulated, are stationary. However in many cases these signals are non-stationary. It means that the input of our system must be nonstationary as well. In this case we should use another approach to specify the input of the system so that the



requirements mentioned above were satisfied. To achieve this, the wave packet decomposition[2] was used. It was

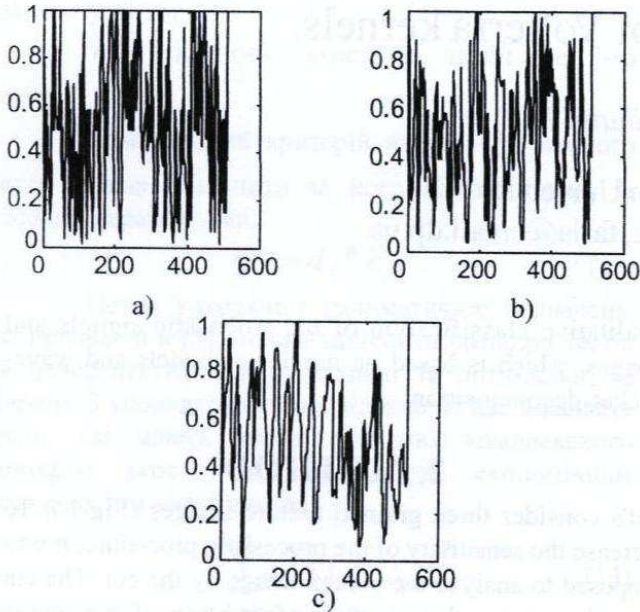


Fig.2. Slices of the grained textures, presented on the Fig.1.

shown by Mallat S.[2], that the wave packet decomposition (till the 2 level) could be presented as it is shown on Fig.3.  $H$  and  $G$  are special high and low pass filters respectively. They play a role of the basis functions in wave-packet decomposition. Further analysis of  $S(\lambda)$  was conducted in terms of "input-output" model shown on Fig.4.

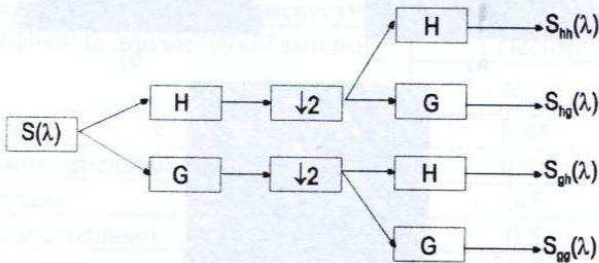


Fig.3. Wave packet decomposition of given  $S(\lambda)$  data;  $H$  and  $G$  are lowpass and highpass filters respectively;  $\downarrow 2$  – dyadic downsampling.

It was assumed, that  $S_{hh}(\lambda)$  and  $S_{gh}(\lambda)$  are the input and output of the model respectively. Practicability of this approach has following grounds:

1.  $S_{hh}(\lambda)$  and  $S_{gh}(\lambda)$  are sharing the same "genetic root" (i.e.  $S(\lambda)$ ) and have concerted Fourier spectra. It ensues from the idea of wave packet decomposition.
2.  $S_{hh}(\lambda)$  and  $S_{gh}(\lambda)$  are non-stationary alternating-sign sequences, which envelope curves are correlated. This fact is the main reason for using model, presented in Fig.4.

The main idea of the analysis is the connection between  $S_{hh}(\lambda)$  and  $S_{gh}(\lambda)$  through functional second order Volterra series. As long as the second order Volterra kernel  $F(\tau_1, \tau_2)$  is an image it should reflect the changes in the signal, making them more obvious for a visual observation. This fact specified the choice of the second order Volterra kernel as a major informative characteristic

for the qualitative classification. The kernels of the system with arbitrary input and output sequences can be calculated in the frequency domain by non-parametric method of Powers[3].

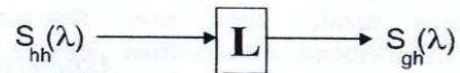


Fig.4. Hypothetical model to form  $S_{gh}(\lambda)$ ;  $L$  - nonlinear operator, characterizing the system excited with  $S_{hh}(\lambda)$ .

Volterra kernels, calculated for the grained textures(Fig.1) are shown on Fig.5.

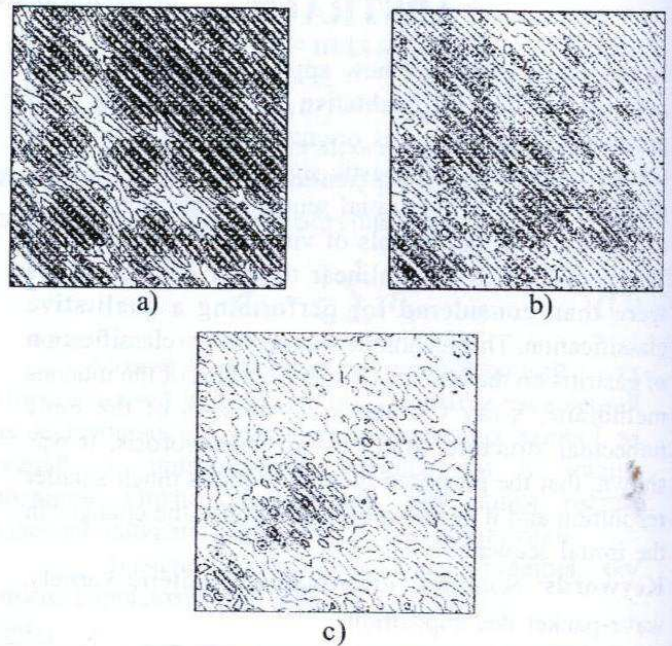


Fig.5. Volterra kernels, calculated for textures, shown on the Fig.1.

### 3. APPLICATION OF THE PROPOSED METHOD

The proposed method has been applied to qualitative characterization of different stochastic processes, in particular to characterization X-Ray CT and X-Ray images of a coxa, affected with osteoporosis and IR spectrograms of gastric mucos membrane. X-Ray CT images of coxa are presented on Fig.6. The main advantage of using X-Ray CT images is that they give the full information of the bone topology, revealing the local changes of trabeculae. This fact makes the diagnostics of osteoporosis on the base of the tomographic images very effective, especially, on the beginning stages of the disease development. X-Ray CT images, shown on Fig.6. present coxae of a woman at the age of 45. The right coxa (b) is affected more deeply with osteoporosis than the left one (a). This fact has been proved by earlier clinical analysis. White lines show the places where the cuts have been taken from the image for the analysis. It should be noted, that only the oscillating component was used for the analysis because it is responsible for the trabecular bone structure. Therefore, only this component should be analyzed for osteoporosis diagnostics. The Volterra kernels, calculated by the



proposed method from these component are shown on Fig.7. Considering calculated kernels, it can be seen that the peculiarities of the kernel topology reflect the

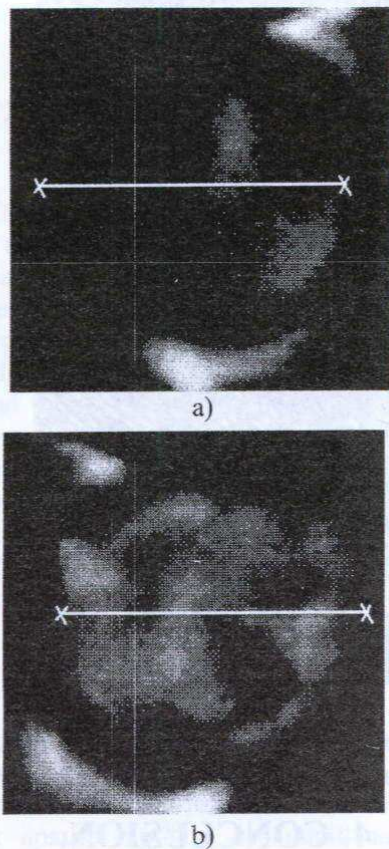


Fig.6. X-Ray CT images of the human coxae, affected with osteoporosis.  
a) normal coxa; b) affected with the disease.

changes in bone trabecular structure and therefore they can be used for osteoporosis diagnostics. The proposed method was also applied to the diagnostics of osteoporosis on the base of X-Ray images. One of the classical methods for osteoporosis classification on the base of X-Ray images is an X-Ray densitometry techniques, which are based on dual energy x-ray absorptiometry[4]. This approach gives good results, but the resulting images could not be processed any further, because the brightness of the pixels is connected to the density of the bone and any nonlinear transformation would destroy this brightness/density relationship. Therefore, the method, which could be applied to the X-Ray images acquired with a usual x-ray camera, would be of a great use. On Fig.8. there are two X-Ray images of the left coxa of a person suffering from a quickly developing osteoporosis. The treatment began in 1997 but the drug therapy failed to stop the disease development and in two years the coxa was replaced with prosthesis. It was assumed that a cut of the X-Ray image (the projection of trabecular structure on a plain surface) should contain averaged sizes of trabeculae. Therefore, local changes in trabeculae bone structure should be reflected in local characteristics of oscillating component of a slice. From looking at calculated images of second order Volterra

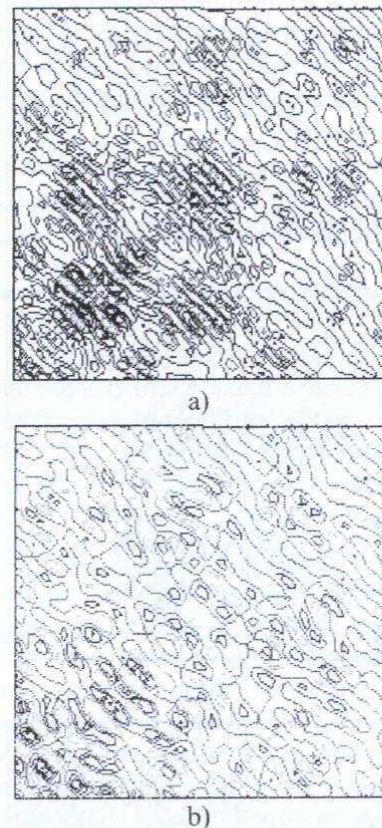


Fig.7. Second order Volterra kernels, calculated for the X-Ray CT images, presented on Fig.5.  
a) kernel calculated for the normal coxa;  
b) kernel calculated for the ill coxa.

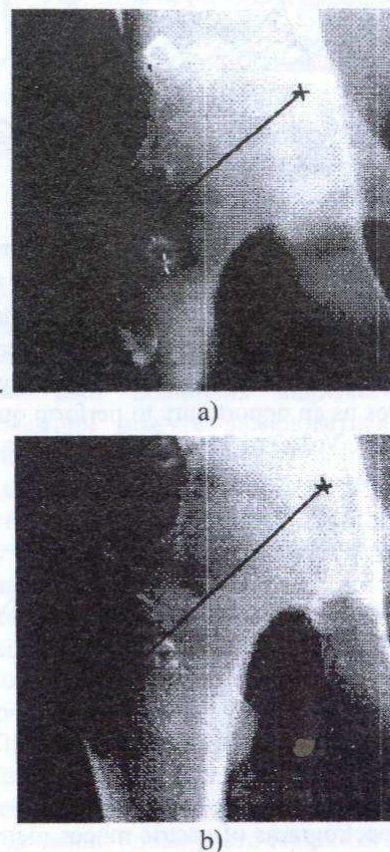


Fig.8. X-Ray images of the coxa, affected with osteoporosis. a) Year 1998; b) Year 1999.



kernels (Fig.9.) it can be noticed that the topology of the image changes as the disease develops. To our mind, the peculiarities of the topology of the kernels cohere with the trabecular structure – the less uniform kernel is, the thicker trabeculae are.

The proposed method was also tested on gastritis classification on the base of IR spectroscopy data (Fig.10.) of the gastric mucos membrane. In comparison

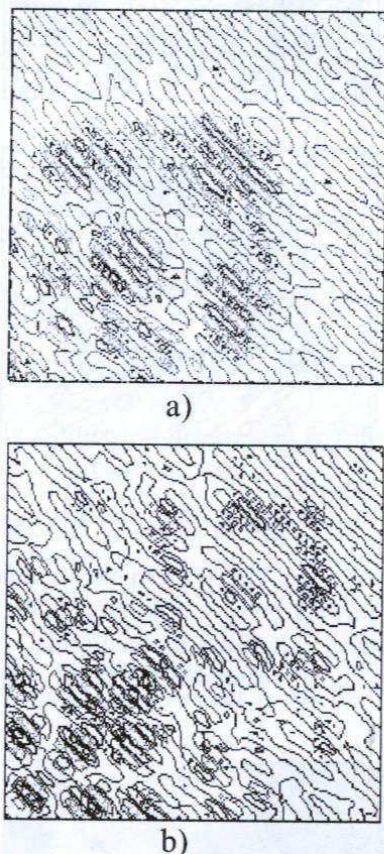


Fig.9. Kernels, calculated for the X-Ray images. presented on Fig.7.

a) Year 1998; b) Year 1999

with other methods of testing for gastroenterological diseases, IR spectroscopy is the most harmless and easy to be conducted. The use of the proposed method gives very good results, especially, when the calculated kernels were further processed (Fig.11.) [5]. Contrasting images on Fig.11 gives us an opportunity to perform qualitative distinctions of Volterra kernels for “normal” and “affected” cases.

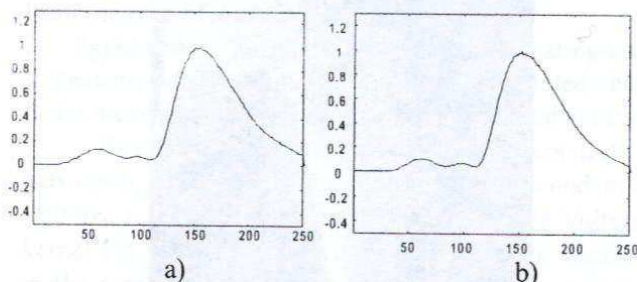


Fig.10. IR spectrograms of gastric mucos membrane; a) normal; b) gastritis;

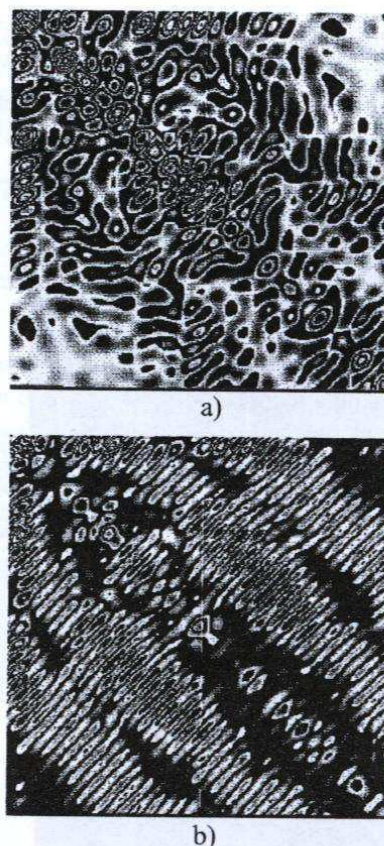


Fig.10. Kernels, calculated for the IR spectrograms presented on Fig.9. a) normal case. b) ill case.

#### 4. CONCLUSION.

It has been shown that the proposed method proved to be sensitive enough to be successfully applied to the analysis of trabecular bone structure on the base of both X-Ray CT and X-Ray images. The results of IR spectroscopy data analysis have also proved that this method can be effectively used for gastritis diagnostics. Nonlinear approach and wave-packet decomposition, implemented in the method, make it effective in analysis of complex stochastic processes, especially, when the physico-mathematical model of the processes being studied is very complex.

#### REFERENCES.

- [1]. Box G.E.P., Jenkins G.M., *Time series forecasting and control*, San Francisco, 1970.
- [2]. Mallat S., “A theory for multidimensional signal decomposition: the wavelet representation”, *IEEE Trans. On Pattern Analysis*, v.11, pp.674-693, 1989.
- [3]. Powers, E.J., “Applications of digital polyspectral analysis to nonlinear systems modeling and nonlinear wave phenomena,” *Proc. Workshop on Higher-Order Spectral Analysis*, pp. 73-77, Vail, Colorado, June 1989.
- [4]. Lafferty F.W. “Correlation of dual-energy X-ray absorptiometry, quantitative computed tomography, and single photon absorptiometry with spinal and non-spinal fractures”. *Osteoporosis Int* 6:407-415, 1996.
- [5]. Akhmetshin A., “Pseudocoherent phase-contrast imaging method with varying focus depth: an effective tool of low-contrast radiological image analysis”, *SPIE Proc., Medical Image Processing*, v.3338, pp.904-915, 1998.



# СТРУКТУРНИЙ МЕТОД ОПИСУ І ВИДІЛЕННЯ ВІДРІЗКІВ ПРЯМИХ ЛІНІЙ У КОНТУРІ.

Володимир Калмиков, Дмитро Калмиков, Тетяна Романенко  
Інститут проблем математичних машин і систем НАН України

02187 Київ 187 просп. Академіка Глушкова 42, тел. 266 1369, E-mail: [kvg@immssp.kiev.ua](mailto:kvg@immssp.kiev.ua)

The structural method of the description and allocation of pieces of direct lines in a contour is offered which is based on a hypothesis about allocation of pieces of any direct lines in primary visual cerebral cortex. The offered hypothesis demonstrates one of probable mechanisms of display and recognition of pieces of direct any lengths and directions by the neurophysiology means, and can be used as a working hypothesis at experimental researches of visual perception. The received way of the description of image contours is not economical, but the algorithm, received on its basis, allows to restore the contour information as a sequence of pieces of straight lines, which was lost while the image was discretized.

## 1. ВСТУП.

Запропонований структурний метод опису і виділення відрізків прямих ліній у контурі бінарного об'єкту має в основі гіпотезу про виділення відрізків довільних прямих ліній у первинній зоровій корі мозку людини, що розроблена на основі наступних відомих даних про механізм зорового сприйняття [1].

1. Зображення у полі зору ока проектується на сітку, що утворена зоровими рецепторами, яку можна розглядати як якийсь двомірний дискретний простір;

2. Кожний нейрон поля зору ока (гангліозна клітина) збирає інформацію від деякої множини зорових рецепторів, що її називатимемо рецептивним полем. Середня відстань  $\tau$  між гангліозними клітинами в центральній частині поля зору ока визначає дискретність поля зору. Кожній гангліозній клітині відповідає нейрон первинної зорової (стриарної) кори мозку. Будемо називати такі нейрони первинними нейронами.

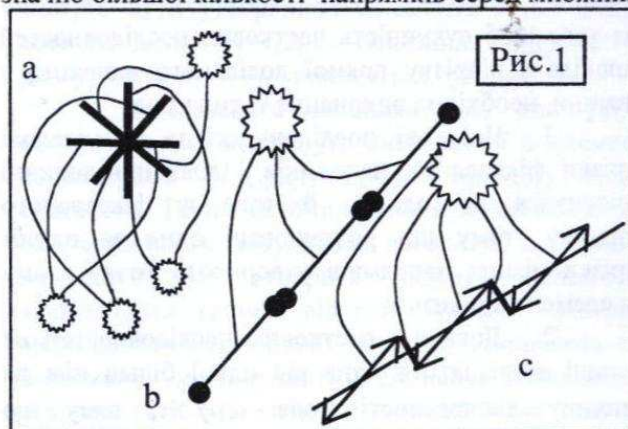
3. У стріарній корі виявлені нейрони, що збуджуються при пред'явленні у визначеному місці поля зору відрізка визначеної довжини і визначеного напрямку – рис. 1а (зірочками позначені нейрони, дугами – зв'язки їх зі стимулами, що ті нейрони збуджують). Розмір відрізків, на які реагують такі нейрони, складає менше  $2^\circ$  поля зору. Будемо називати такі нейрони лінійними. Кількість відмінних один від одного напрямків відрізків, що перетинають одну точку, обмежена і залежить, очевидно, від кількості первинних нейронів, що утворюють рецептивні поля лінійних нейронів, конфігурації цих рецептивних полів і не перевищує 2 - 3 десятків. Природно припустити, що кожній

гангліозній клітині в центральній частині поля зору відповідає лінійний нейрон.

4. У стріарній корі виявлені також кінцеві нейрони, що збуджуються якщо в даній точці поля зору має місце кінець відрізка прямої. Природно також припустити, що кожній гангліозній клітині в центральній частині поля зору відповідає кінцевий нейрон.

## 2. ГІПОТЕЗА.

Отже, по сигналах лінійних нейронів можна судити про те, чи є в поле зору ока відрізки прямих деяких фіксованих напрямків, наприклад, як на рис. 1а. Чим довше відрізок прямої, що розглядається, тим більше лінійних нейронів даного напрямку, відповідних відрізкам, що взаємно продовжуються, видають сигнали – 1б. У той же час загальновідомо, що зір людини виділяє відрізки прямих дуже великої, значно більшої кількості напрямків серед множини



довільних ліній, ніж тієї, що виявлена експериментально. Дійсно, із звичайних повсякденних спостережень з очевидністю випливає, що відмінності між напрямками двох довільних відрізків достатньо великої довжини можуть вимірюватися кутовими хвилинами і, навіть, секундами. Для цього потрібно значно більша дозвольна спроможність, ніж можуть забезпечити самі по собі сигнали лінійних нейронів. У такому випадку повинен існувати спеціальний зоровий механізм, що реалізований у вигляді зв'язків між нейронами, який забезпечує виділення і розпізнавання відрізків прямих ліній довільної довжини і довільного напрямку при розумних витратах кількості нейронів.

Природно припустити, що сигнали лінійних нейронів про відрізки фіксованих напрямків використовуються для побудови образу відрізка прямої довільної довжини і довільного напрямку.



Якщо напрямок деякого відрізка прямої довільної довжини в полі зору не збігається цілком із жодним із тих фіксованих напрямків, що виділяються за допомогою лінійних нейронів, тобто цей напрямок є проміжним між двома сусідніми фіксованими напрямками, то такий відрізок не може бути відображений сигналами лінійних нейронів, що відповідають продовжувачим один одного відрізкам обмеженої довжини та одного фіксованого напрямку. У той же час окремі лінійні нейрони, що мають напрямок найближчий до напрямку відрізка, що розглядається у полі зору, видають відповідні сигнали. Проте відрізки, що відповідають цим нейронам, уже не завжди утворюють зв'язану послідовність, тобто кінець попереднього відрізка не завжди збігається з початком наступного – мал. 1с. У залежності від того, наскільки напрямок відрізка у полі зору не збігається з фіксованими напрямками, послідовність продовжувачих один одного відрізків фіксованої довжини переривається, "спрацьовують" кінцеві нейрони і починається нова послідовність продовжувачих один одного відрізків фіксованої довжини. Ця нова послідовність зміщена щодо попередньої в напрямку, перпендикулярному фіксованому напрямку відрізків. Такі послідовності, що складаються з одного або деякої кількості продовжувачих один одного відрізків фіксованої довжини і напрямку, будемо називати *частковими*. Для того щоб сукупність часткових послідовностей відповідала відрізку прямої довільного напрямку і довжини, необхідно виконання таких умов:

1. Часткові послідовності та їх складові відрізки фіксованих напрямків і довжини повинні відноситися до одного й того ж фіксованого напрямку, тому що розташовані один за одним відрізки різних напрямків утворюють точку зламу або елемент кривизни;

2. Довжини часткових послідовностей не повинні відрізнятися одна від одної більш ніж на величину дискретності поля зору  $\tau$ , тому що розташовані один за одним відрізки різної довжини утворюють точку зламу або елемент кривизни;

3. Відстані між кінцями і початками розташованих одна за одною часткових послідовностей не повинні перевищувати величини дискретності поля зору  $\tau$  тому що в іншому випадку між ними може бути поміщений відрізок, що відрізняється по довжині або напрямку.

4. Напрямки зміщення наступних часткових послідовностей відносно попередніх не змінюються.

Умови 1,3,4 можуть бути легко виконані шляхом утворення рецептивних полів нейронів більш високих порядків із лінійних і кінцевих нейронів у якості рецепторів. Умова 2 може бути здійснена шляхом порівняння електричних зарядів, отриманих від нейронів сусідніх часткових послідовностей, оскільки величина передаваного заряду залежить від довжини відрізка у полі зору.

Запропонована гіпотеза демонструє один із можливих механізмів відображення і розпізнавання відрізків прямих довільних довжин і напрямків засобами, характерними для нейрофізіології і може бути використана в якості робочої гіпотези при експериментальних дослідженнях зорового сприйняття.

### 3. ВІДНОВЛЕННЯ ВІДРІЗКІВ ПРЯМИХ У ОПИСІ КОНТУРУ.

Далі розглянемо ту ж задачу виділення відрізків прямих у процесі структурного аналізу контурів об'єктів бінарного зображення з використанням основних положень розглянутої гіпотези.

Зображення задане в двомірному просторі на прямокутній ґратці  $M \times N$ , як деяка функція  $V(m, n) = (0, 1)$ ,  $m = 1, 2, \dots, M$ ,  $n = 1, 2, \dots, N$ . Зображення включає об'єкти - сукупності компактно розташованих пікселів, таких що  $v(m, n) = 1$ , у той час як  $V(m, n) = 0$  для фонових пікселів. Об'єкти звичайно описуються контурами - замкнутими ламаними лініями, що відокремлюють піксели, що належать об'єкту, від пікселів, що належать фону.

Як результат дискретизації вихідного зображення відрізки прямих і кривих довільного напрямку, що спочатку утворювали контур, перетворилися у відрізки вертикальних і горизонтальних ліній - *граней*, що чергуються. Спільні точки сусідніх граней будемо називати *точками зламу*. Представлення контурів у вигляді граней незручно з ряду причин, головною з них є втрата представлення контуру як послідовності відрізків прямих (або кривих) ліній довільного напрямку. Звичайно для відновлення відрізків прямої по окремих точках використовують методи, алгоритми сгладжування, що дозволяють одержати приблизне (а не точне) рішення навіть при відсутності поміх.

У роботі [2] був запропонований опис контурів бінарних об'єктів послідовностями L-елементів. Розглянемо виділення відрізків прямих у послідовностях L-елементів. Складові елементарні відрізки - L-елементи утворюються парами, що складаються з розташованих один за одним горизонтального і вертикального (або навпаки, вертикального і горизонтального) відрізків. Перший відрізок може мати довільну довжину:  $l = 1, 2, \dots, L$ , де  $L$  - максимально можлива довжина відрізка в даній реалізації, у той час як довжина другого відрізка в парі дорівнює  $l$  (у деяких випадках вона дорівнює 0), що цілком очевидно впливає з визначення алгоритму дискретизації. L-елементи відповідають відрізкам обмеженої довжини у викладеній вище гіпотезі про виділення відрізків прямих довільного напрямку.

Параметрами L-елемента є:

-  $l$  - довжина L-елемента, у якості якої приймається довжина першого відрізка; як уже



згадувалося,  $l$  приймає значення з ряду натуральних чисел;

-  $g$  - напрямок L-елемента, у якості якого приймається напрямок першого відрізка (для певності приймемо  $g = 0$  - "нагору",  $g = 1$  - "управо",  $g = 2$  - "униз",  $g = 3$  - "уліво");

-  $q = 0, 1$  - напрямок другого відрізка (0 - поворот управо щодо напрямку першого відрізка, 1 - поворот уліво щодо напрямку першого відрізка).

Розглянемо послідовності L-елементів, що утворюють відрізки прямих, для випадку  $\Delta y > (x(0, \text{ вважаючи, що для } \Delta x > y(0, \text{ результат буде симетричним. Такі послідовності звичайно утворюються в результаті роботи алгоритму дискретизації відрізків по методу Брезенхама.$

Як і для випадку зорового сприйняття, L-елементи, що складають відрізок прямої, співпадають за напрямками першого і другого відрізків, відрізняються по довжині не більш, ніж на величину дискретності простору - один піксел.

Насамперед відзначимо, що  $n = \Delta x$  - кількість L-елементів, що утворюють відрізок. Ціла частина від ділення  $\Delta y / \Delta x - l = (\text{ЦЧ}) \Delta y / \Delta x$  - це довжина деякої частини L-елементів, що складають відрізок. Довжина інших L-елементів складає  $l + 1$ . Кількість таких L-елементів -  $n^*$  дорівнює залишку від ділення  $\Delta y / \Delta x$ , тобто  $n^* = (O)y / (x, \text{ у той час як кількість L-елементів довжини } l \text{ складає } n - n^*. \text{ L-елементи довжин } l, l + 1, \text{ що утворюють відрізок прямої, рівномірно розподілені в межах цього відрізка. Якщо при відображенні відрізка прямої L-елементами } n^* = n - n^*, \text{ то L-елементи довжини } l \text{ і } l + 1 \text{ чергуються. Якщо } n^* > n - n^*, \text{ то L-елементи довжини } l \text{ рівномірно розподіляються між L-елементами довжини } l + 1, \text{ і навпаки, у протилежному випадку. Послідовність з однакових L-елементів довжини } l, \text{ що закінчується L-елементом довжини } l + 1 \text{ або } l - 1, \text{ будемо називати комплексним L-елементом або K-елементом. Послідовність з однакових L-елементів довжини } l, \text{ що починається L-елементом довжини } l + 1 \text{ або } l - 1, \text{ а далі продовжується L-елементами довжини } l \text{ також є K-елементом. Такий K-елемент будемо називати зворотним. Кількість K-елементів, що утворюють відрізок прямої визначається як } \min\{(n - n^*), n^*\}. \text{ Параметрами K-елементу, таким чином, є:$

- параметри першого вхідного в нього L-елемента (для зворотних K-елементів - другого);

-  $k$  - кількість L-елементів, що утворюють K-елемент;  $k = (\text{ЦЧ})n / j$  будемо називати "довжиною" K-елементу; зворотні K-елементи будемо позначати знаком "-" при величині  $k$ .

-  $r = -1, 0, +1$ ; значення  $r = -1$  або  $r = +1$  визначають відповідно зменшення або збільшення довжини останнього в порівнянні з попереднім L-елементом на 1. Якщо  $r = 0$ , то всі L-елементи даного K-елементу однакові.

Через несумісність величин  $n$  і  $j$  деякі K-елементи містять  $j + 1$  L-елементів. Їхня кількість визначається як  $j^* = (O)n / j$ . K-елементи, що містять

різні кількості L-елементів, рівномірно розподіляються в межах відрізка.

Послідовність із  $k_1$  однакових K-елементів довжини  $k$ , що починається або закінчується K-елементом довжини  $k + 1$  або  $k - 1$ , будемо називати K1-елементом. Кількість K1-елементів, що утворюють відрізок прямої  $j_1 = \min\{(j - j^*), j^*\}$ . Параметрами K-елементу, таким чином, є:

- параметри першого K-елементу, що входить в нього;

-  $k_1$  - кількість K-елементів, що утворюють K1-елемент;  $k_1 = (\text{ЦЧ})j / j_1$  будемо називати "довжиною" K1-елемента; зворотні K-елементи будемо позначати знаком "-" при величині  $k_1$ ;

-  $r_1 = -1, 0, +1$ ; значення  $r_1 = -1$  або  $r_1 = +1$  визначають відповідно зменшення або збільшення довжини першого або останнього в послідовності K-елемента в порівнянні з іншими K-елементами даного K1-елемента на 1. Якщо  $r_1 = 0$ , то всі K-елементи даного K1-елемента однакові.

При несумісності величин  $j$  і  $j_1$  деякі K1-елементи містять  $j_1 + 1$  K-елементів. Їхня кількість визначається як  $j_1^* = (O)j / j_1$ . K1-елементи, що містять різні кількості K-елементів рівномірно розподіляються в межах відрізка. У цьому випадку можуть бути утворені K2-елементи в кількості  $j_2 = \min\{(j_1 - j_1^*), j_1^*\}$ . При несумісності величин  $j_2$  і  $j_1$  можуть бути утворені K3-елементи і так далі, аж доти для деякого  $i = 1, 2, \dots$  - показника сумірності (або рангу) значення  $j_i$  і  $j_{i-1}$  виявляться сумірні.

Із сказаного випливає, що для усякого відрізка прямої може бути визначений K-елемент із параметрами  $\{l, g, q, k, r, \dots, k_b, r_b, i\}$ , причому єдиним способом. Такий засіб опису не є ощадливим, проте дає можливість побудувати алгоритм, що дозволяє визначити, чи утворить задана послідовність чергуючихся граней відрізок прямої. Алгоритм визначається перевіркою умов, чи утворюють грані L-елемент, чи утворюють L-елементи K-елемент, чи утворюють K-елементи K1-елемент і так далі доти чергова пара граней перестане разом задовольняти умовам приналежності до відрізка прямої, вже утвореному раніше переглянутими гранями. Тоді відрізок прямої вважається розпізнаним, а наступна пара граней, певне, уже не є продовженням відрізка прямої і є початком іншого відрізка.

Стосовно до зорового сприйняття сказане вище може позначати наступне. Якщо часткові послідовності, що складають довільний відрізок прямої відрізняються одна від одної на величину  $\tau$  по довжині, то вони повинні бути розподілені рівномірно уздовж відрізка прямої, подібно тому як K-елементи неоднакової довжини рівномірно розподіляються в межах відрізка прямої. Якби в процесі зорового сприйняття не існувало механізму такого рівномірного розподілу, то відрізки прямих, складені з часткових послідовностей довжини тільки довжини  $l$  або тільки довжини  $l + \tau$  не розрізнялися б один від одного, чого насправді не відбувається. У такому випадку запропонована гіпотеза повинна



бути доповнена ще одною умовою перевірки рівномірного розподілу часткових послідовностей різноманітної довжини в межах відрізка прямої.

Подібно L-елементам часткові послідовності мають утворювати структури, подібні K-елементам, що можуть бути реалізовані у вигляді нейронів час-

ткових послідовностей старших порядків, рецептивні поля яких утворені нейронами часткових послідовностей молодших порядків.

Алгоритм виділення послідовності K-елементів у послідовності L-елементів поданий у вигляді таблиці, де  $t = 0, 1, 2, \dots$  - номер L-елементу у послідовності.

$k$	$l_t, l_{t-1}$	$l_t = l_{t-1}$	$l_t = l_{t-1} + 1$	$l_t = l_{t-1} - 1$	$ l_t - l_{t-1}  > 1$
$k=1$	$k_i++; t++;$	$k_i++; t++;$	$k_i++; t++; k_i = (-1) * k_i;$	$k_i++; t++; k_i = (-1) * k_i;$	$k_i=1; i++;$
$k>1$	$k_i++; t++;$	$k_i++; i++; t++; r_i = +1; k_i = 1;$	$k_i++; i++; t++; r_i = -1; k_i = 1;$		
$k<-1$	$k_i--; t++;$	$k_i = 1; i++;$			

Розглянемо приклади роботи алгоритму. На рис. 2,3 наведені дві різні частини контуру у вигляді відрізків прямих ліній та послідовностей L-елементів, не пошкоджені завадами. Послідовність L-елементів на мал.2 -  $de$  утворена при дискретизації одного відрізка  $DE$  розмірами  $|\Delta x|=62; |\Delta y|=45$ . Послідовність L-елементів на мал.3 -  $abc$  - утворена при дискретизації двох зв'язаних відрізків  $AB$  та  $BC$ , що незначно відрізняються нахилом, загальними розмірами  $|\Delta x|=53; |\Delta y|=36$ . Розміри

утворює один K2-елемент, у той час як послідовність  $abc$  утворює два K2-елементи, розділяючи послідовність L-елементів саме у точці  $b$ .

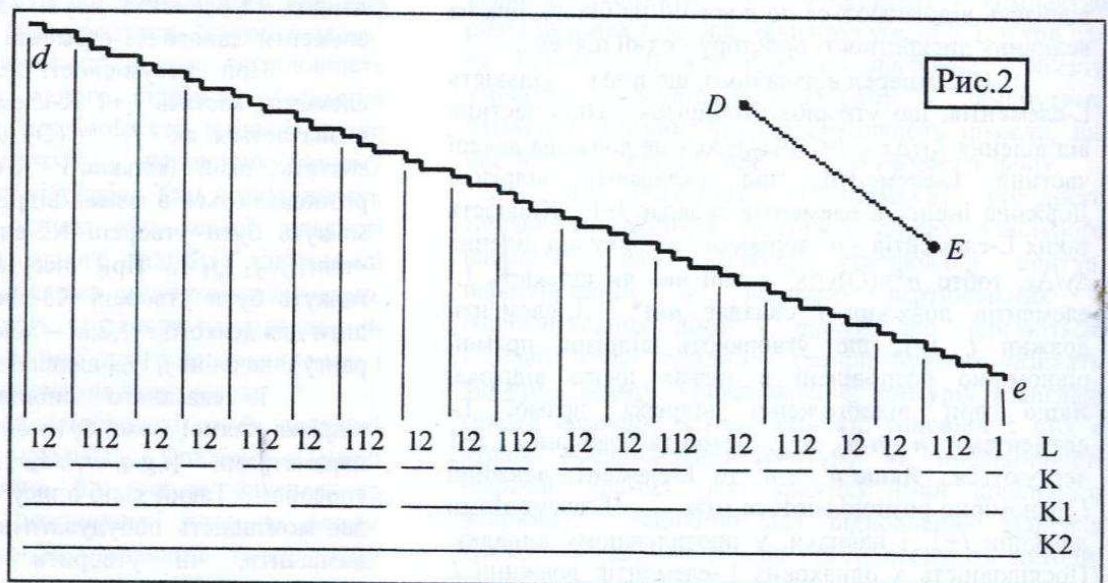


Рис.2

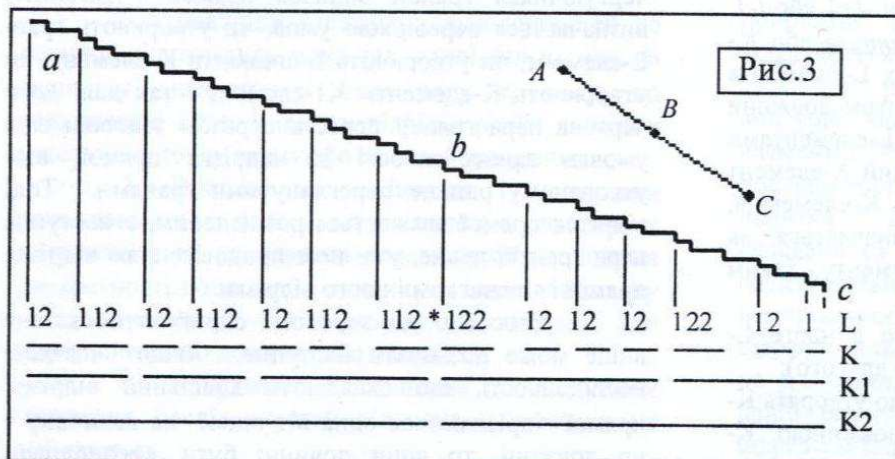


Рис.3

#### 4.ВИСНОВКИ.

Запропоновано метод та алгоритм структурного аналізу відрізків прямих, що дозволяє відновити опис контуру у дискретному просторі, як послідовність відрізків прямих. Проведено зіставлення з механізмами зорового сприйняття, показано їх подібність відносно операцій над елементами зображення, що дозволило сформулювати гіпотезу, придатну до експериментальної перевірки.

#### ЛІТЕРАТУРА.

1. Хьюбел Д. "Глаз, мозг, зрение." - Москва: Мир, 1990 - 239 с.
2. Калмыков В.Г. Вишневский В.В. Анализ контуров объектов в бинарных изображениях. // Математические машины и системы - 1997 - №2 - с.68- 71.



# IMPROVEMENT OF LOW – CONTRAST MEDICAL IMAGE ANALYSIS SENSITIVENESS AND EXACTNESS USING MARKOV RANDOM FIELD MULTIDIMENSIONAL SEGMENTATION METHOD

A.M. Akhmetshin, A.E. Fedorenko

Dnepropetrovsk State University

## Abstract

In this work algorithm of medical images visual properties improvement for image analysis simplification was realized. The exactness and sensitiveness increasing of different image objects detection is achieved via gradient correction and segmentation. Pixels classification proceeds using Markov random field (MRF) conception. In order to increase the input segmentation data the multiresolution Wavelet decomposition method was proposed to take in consideration the information about more coarse region forms and borders. Using a number of "approximation" images generated by Wavelet transform together with the initial image allows us to make the resulting clusterisation more exact.

## 1. INTRODUCTION

In the area of medical images processing the resulting inaccuracy is usually conditioned by both the absence of sufficient input data amount and weak photos contrast. Accurate definition of weak contrast regions is more convenient to be accomplished with any image segmentation (clusterisation) methods. There are two different segmentation ways: "unsupervised" and "supervised". The obvious weakness of supervised algorithms such as K – means method or fuzzy C – means method [1],[2],[4] often lies in the absence of knowledge about true number of image objects, which must be detected. Due to this problem we used unsupervised clusterisation method. Moreover, during the segmentation process local spatial image characteristics of MRF model were taken into account what allows us to consider the input image not only as a simple set of different brightness pixels but also as a complex textured object.

Lets consider some gray – scaled image  $Y=\{y_s\}$ , where  $s$  is a pixel characterized by its coordinates:  $s=(i,j)$ ,  $i \in \{1..N\}$ ,  $j \in \{1..M\}$ ,  $y_s$  is a gray level of  $s$ . The segmented version of  $Y$  is  $X=\{x_s\}$ ;  $x_s=k$  means that pixel  $s$  can belong to the segmentation class  $k$ .

The most known pixels classification method for the segmentation process is the maximization of a posteriori

probability (MAP), which is usually defined by Bayesian relationship:

$$P(X | Y) = P(Y | X) \cdot P(X) \quad (1)$$

$P(X) \sim p(x_s=k)$  is a priori probability of pixel  $s$  belonging to segmentation class  $k$ ,  $P(Y|X) \sim p(y_s=y | x_s=k)$  is a conditional distribution of pixels intensity within certain class  $k$ . Both a priori and a posteriori probabilities are defined with respect to MRF model.

## 2. GAUSSIAN MARKOV RANDOM FIELD MODEL

Let an image  $Y$  be modeled on a finite lattice  $L$  by Gaussian Markov random field (GMRF) where  $Y=\{y_{ij}; 0 \leq i \leq M, 0 \leq j \leq N\}$  and  $L=\{(i,j); 0 \leq i \leq M, 0 \leq j \leq N\}$ . A neighborhood system  $\eta_s$  is a set of all the pixels which are the nearest neighbors of pixel  $s$ .  $\eta_s$  must satisfy the following requires:

- $s$  is not an element of  $\eta_s$ ;
- if  $k \in \eta_s$ , then  $s$  should be an element of  $\eta_k$ .

The joint probability distribution of pixel brightness values is:

$$P(y) = \frac{1}{(2 \cdot \pi)^{M \cdot N / 2} (\det \Sigma)^{1/2}} \cdot e^{-\frac{1}{2} [y]^T [\Sigma]^{-1} [y]} \quad (2)$$

where  $y=[y_{11}, y_{12}, \dots, y_{MN}]^T$ , and  $\Sigma$  is a covariance matrix of  $y$ .

In addition given image must satisfy the Markovian property:

$$P(y_{ij} | y_{pq}, (p,q) \neq (i,j)) = P(y_{ij} | y_{kl}, (k,l) \in \eta_{ij}) \quad (3)$$

The nearest pixels spatial interaction is described statistically by Gibbs distribution.

## 3. GIBBS DISTRIBUTION

Lets define a clique as a set of one or two neighboring pixels. Given image  $Y$  can be modeled as MRF with respect to  $\eta$  when and only when its joint distribution is defined according to Hammersley – Clifford theorem:

$$P(X) = \frac{1}{Z} \cdot \exp\{-U(y)\}, \quad (4)$$



where  $Z$  is a normalization constant,  $U(y)$  is an energy function value, which defines the generalized interaction of certain pixel with it's neighbors:

$$U(y_s) = \sum_{c \in C} V_c(y_s) \quad (5)$$

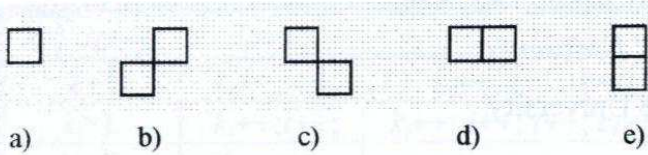


Fig. 1. Different cliques used in work and pixels mutual spatial location within each kind of clique. a) Single clique; b)– e) Double cliques

$V_c$  is a potential function of type  $c$  which defines a degree of clique pixels interaction in a way of fuzzy pixels similarity evaluation. Membership function of the fuzzy relationship "  $y_s$  and  $y_l$  are similar":

$$\Phi(y_s, y_l) = \frac{|y_s - y_l|}{N} \quad (6)$$

Two - pixel clique potential  $V_c$  is defined by value  $\Phi(s,l)$ , thus it is an increasing function in a closed interval  $[0,\beta]$

$$V_c(y_s) = \Phi(y_s, y_l | l \in \eta_s) \quad (7)$$

It corresponds to simplified fuzzy inference rule: "if pixels  $s$  and  $l$  are similar, then  $V_c$  is  $V$ ". Based on the given principle the a priori probability as well as the conditional distribution defined in (1) are calculated.

A priori probability:

$$p(x_s) = \exp\{-U(x_s \sim y_t, t \in \eta_s)\} \quad (8)$$

where

$$U(x_s \sim y_t, t \in \eta_s) = \sum_{l=1}^7 \frac{|y_l - y_t|}{N}, \quad \forall l \in \eta_s \quad (9)$$

Conditional on certain class probability:

$$p(y_s | x_s) = \exp\{-U(y_s | x_s \sim y_t, t \in \eta_s)\} \quad (10)$$

where

$$U(y_s | x_s \sim y_t, t \in \eta_s) = \frac{|y_s - y_t|}{N} \quad (11)$$

## 4. METHOD

Proposed method is based on neighboring pixels gradient analysis. Image clusterisation algorithm consists of several steps:

1. For extraction of homogeneous regions, the initial image is exposed to local covariance transform using following equation:

$$\tilde{a}_{ij} = a_{ij} \cdot (1 - c_{ij}) \quad (12)$$

where

$$c_{ij} = \frac{(q_{ij} - \mu_{ij}) \cdot (q_{ij} - \mu_{ij})^T}{\text{Norm}} \quad (13)$$

$q_{ij}$  is a row of the nearest eight neighboring to  $(i,j)$  pixels brightness values,  $\mu_{ij}$  is mean of  $q_{ij}$ . This operation picks out all image homogenous regions (fig. 2 (b)).

2. Homogeneous regions contrast increasing is achieved by gradient correction and histogram equalization method. Unfortunately, clear histogram equalization method application not always allows us to receive desired exactness of image details. Therefore, before using the histogram equalization algorithm the image gradient correction is accomplished by:

$$b_{ij} = \tilde{a}_{ij} - \Delta a_k \quad (14)$$

where  $\Delta a_k$  is a brightness difference on direction  $k$ ,  $k=1..8$ :

$$\Delta a_k = a_{nm} - \tilde{a}_{ij}, \quad (m,n) \in \eta_{ij} \quad (15)$$

It can be seen that there are eight different values of brightness gradient for each pixel. Therefore it's necessary to define the best gradient meaning according to some criterion. In this work choosing of gradient values is accomplished with famous feature detection algorithm according to entropy minimization principle (fig.2 (c)).

3. Achieved on stage 2 image is then used as initial for segmentation. Afterwards it is corrected with Bayesian segmentation method, based on MRF theory. Under classification of each pixel the following rule is used:

$$x_s = \arg \max_t P_t(X | Y) \quad (16)$$

where  $P_t(X|Y)$  is calculated for each of eight directions  $t$  from pixel  $s$ . During clusterisation process together with the input image two "low-pass" images found by "herringbone" algorithm of two-dimensional discrete Wavelet transformation are processed to evaluate common region forms. In this case we have deal with multidimensional version of unsupervised clusterisation based on MRF model (fig.2 (d)).

## 5. RESULTS

Given algorithm was tested on series of digital mammograms and on several tomographic images. Testing results show that the application of the method gives quite good results in the increasing of exactness of different medical images analysis in the sense of more accurate detail extraction and contrast improvement of the researched image object forms and borders (fig. 3). Due to clusterisation described method has a real possibility to be used as a method of suspected regions extraction before their direct recognition.



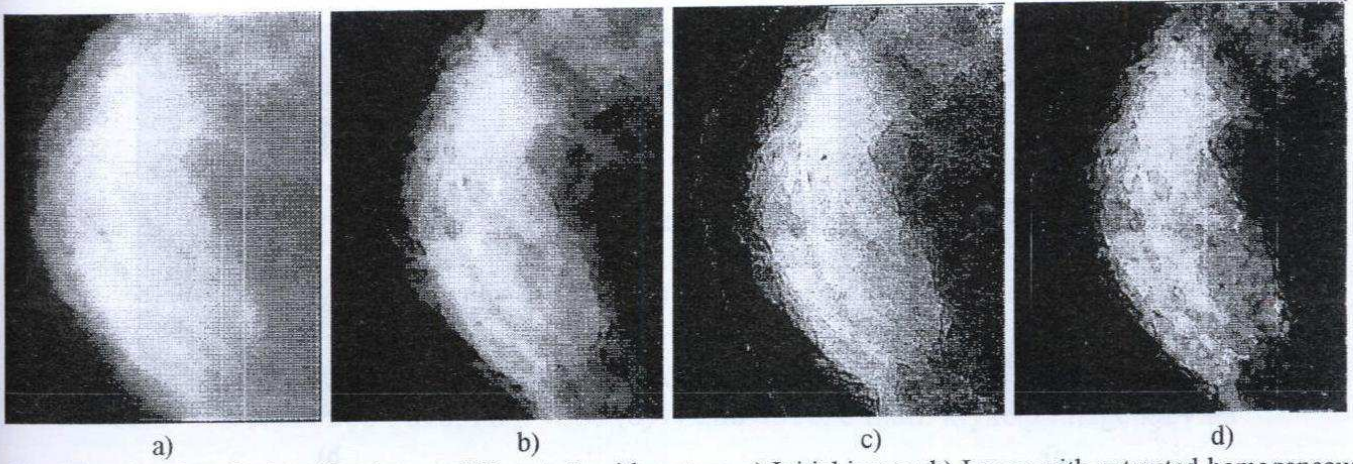


Fig. 2 Results of method application on different algorithm steps. a) Initial image; b) Image with extracted homogeneous regions; c) Image (b) after gradient correction; d) Multidimensional MAP – segmentation result with application of MRF model and Wavelet decomposition.

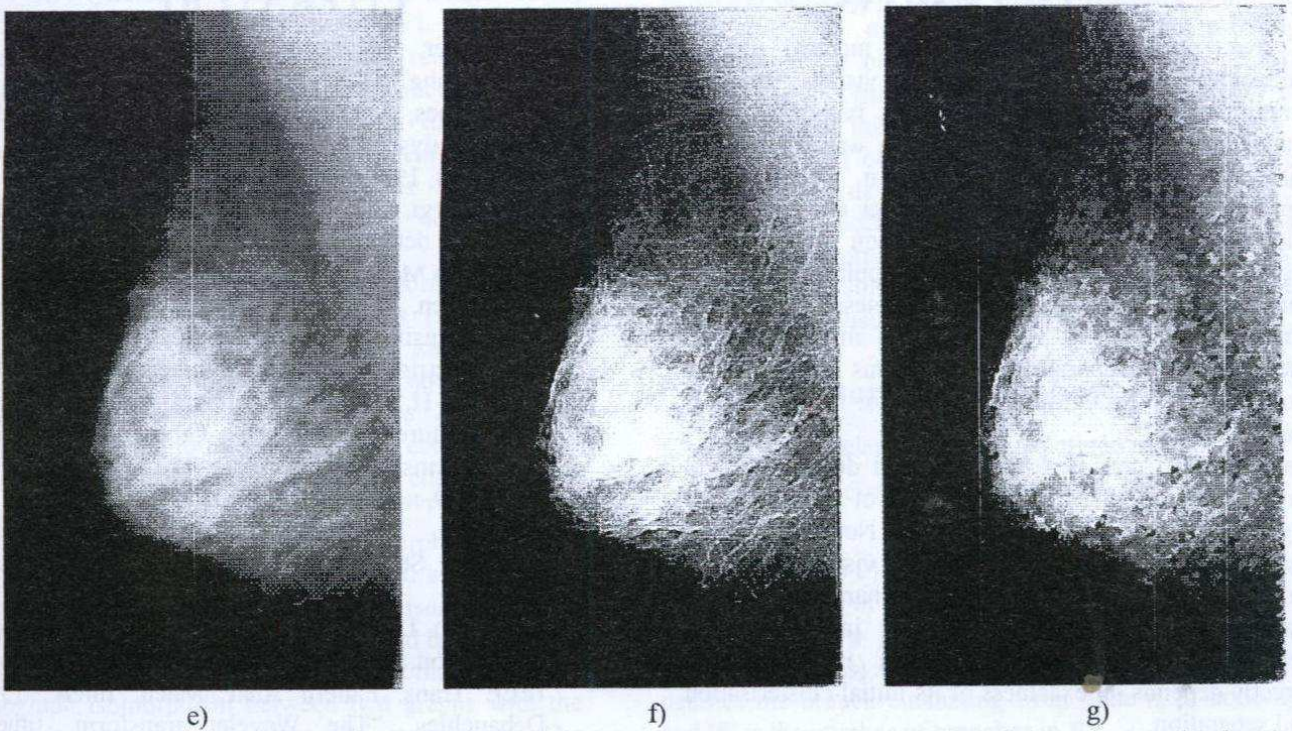
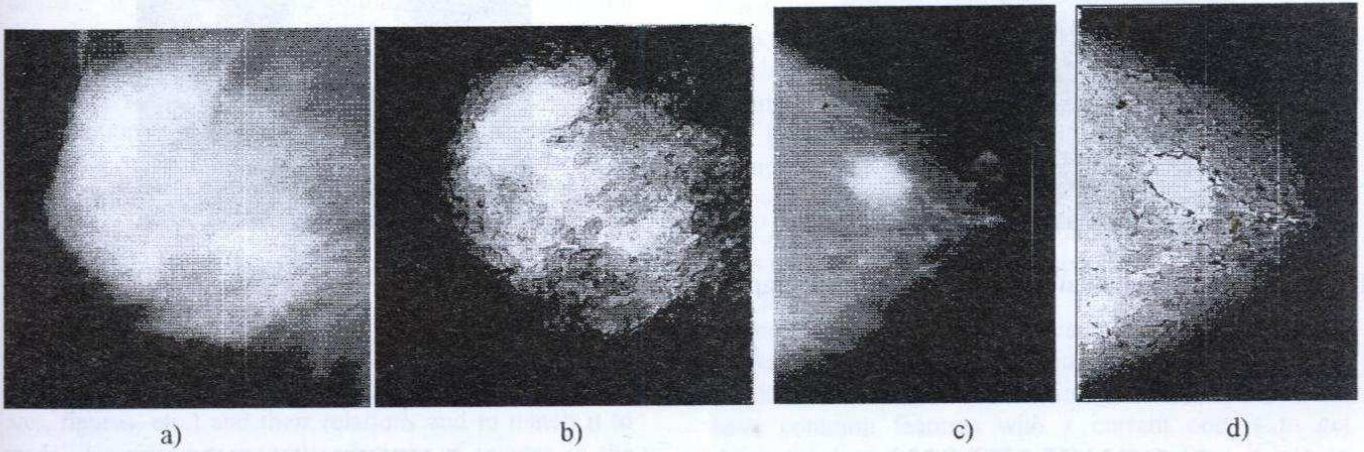


Рис. 3 MAP – segmentation method tested on digital mammograms. a), c), e) – Initial images, b), d), f), g) – Segmentation result.



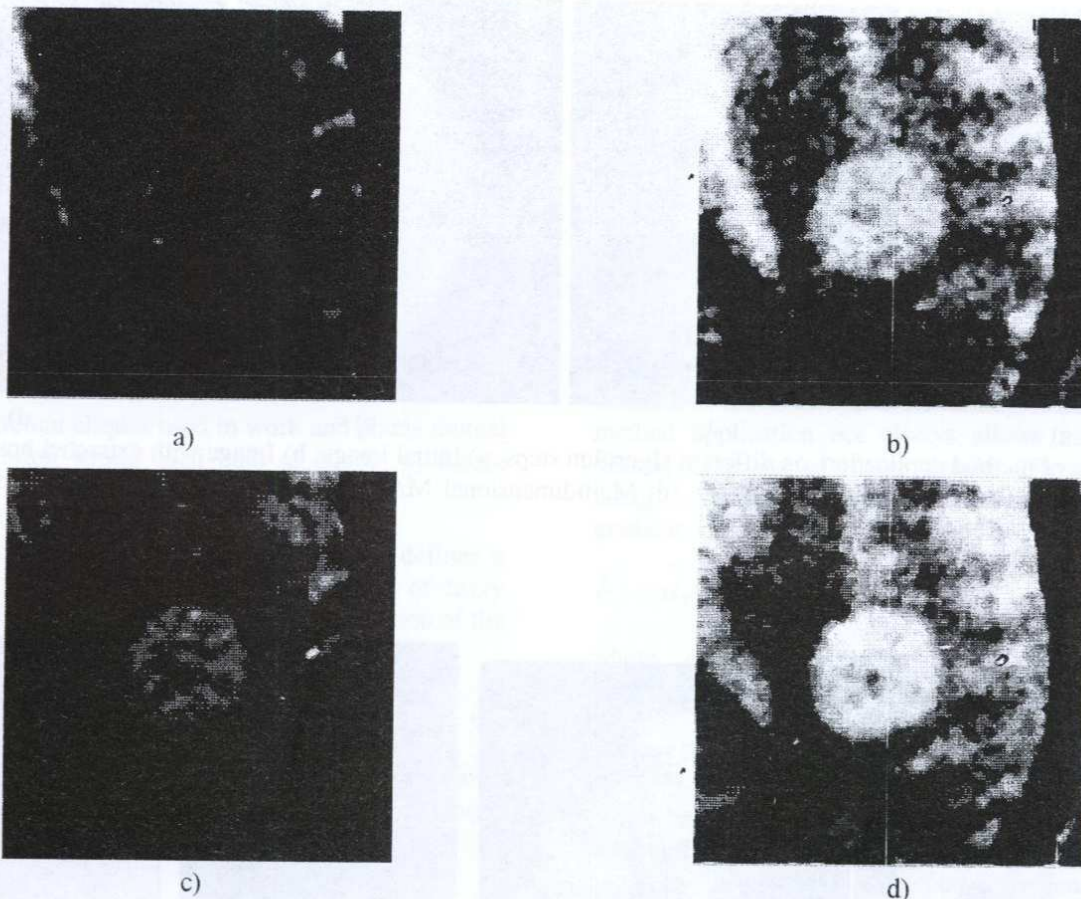


Fig.4. MAP – segmentation method tested on tomographic images. a), c) – Initial images, b), d) - Segmentation result.

## 6. CONCLUSIONS

In given paper we have developed medical images processing method for image visual analysis simplification. Proposed technique is based on local spatial image features analysis, which was carried out with the help of Markov random field model. Application of multiresolution Wavelet decomposition allows us to make additional correction of segmented regions. Results of algorithm application show considerable improvement of exactness and visual image characteristics. Thus proposed method can be used in tasks of separating of suspicious tumor regions from normal background tissues for their further classification.

As it can be seen, performed method doesn't contain any classifier for the final recognition of ill and normal regions within certain medical image. Nevertheless, the described problem of image visual analysis simplification in a way of its visual characteristics and features improvement plays a very important role, because the final classification of image objects directly depends on exactness of its initial clusterisation and separation.

## LITERATURE

- [1] R. Kohler, "A segmentation system based on thresholding", *CGIP*, vol. 15, pp 319 – 338, 1981
- [2] R.C. Dubes, A.K. Jain, "Random fields models in image analysis", *J. Applied Stat.*, vol. 16, no 2, pp. 131 – 164, 1989
- [3] M. Callergi, L.P. Clarke, "Markov Random Field for tumor detection in digital mammography", *IEEE Trans. On Medical Imaging*, vol. 14, no 3, 1995
- [4] C.H. Chen, G. G. Lee, "Multiresolution wavelet analysis using gaussian markov random fields", *Dep. Electrical Computer Engineering, UMD*, 1998
- [5] H. Derin, H. Elliott, "Modeling and segmentation of noisy textured images using Gibbs random fields", *IEEE Trans. Pattern Anal. Machine Intell.*, vol PAMI – 9, no 1, pp 39 – 55, 1987
- [6] J. Besag, "On the statistical analysis of dirty pictures", *Stat. Soc.*, vol. 48, no 3, pp 259 – 302, 1986
- [7] S. Geman, D. Geman, "Stochastic relaxation, Gibbs distribution, and Bayesian restoration of images", *IEEE Trans. Pattern Anal. Mach. Intell.*, 1984I. Debauchies, "The Wavelet transform, time – frequency localization and signal analysis", *IEEE Trans. Inform. Theory*, no 36, pp 961 – 1005, 1981



# Attributed relational graphs based description and retrieval of shoeprints in the image database system SHARS

Ch. Feist<sup>1</sup>, V. Kiiko<sup>2</sup>, R. Knöfel<sup>1</sup>, G. Stanke<sup>1</sup>

<sup>1</sup>Gesellschaft zur Förderung angewandter Informatik e. V. (GfAI),  
Albert-Einstein-Str. 16, D-12489 Berlin

Tel.: +49 (0)30 6392-1610, Fax: +49 (0)30 6392-1661, e-mail: stanke@gfai.de

<sup>2</sup>Int. Research and Training Centre of Information Technologies and Systems of National Academy of Sciences of Ukraine, Prospect Academica Glushkova 40, 252022 Kiev  
Tel.: +380 (0)44 266-3018, Fax: +380 (0)44 266-1570, e-mail: kiyko@image.kiev.ua

## Abstract

In this paper attributed relational graphs based approach implemented in the image database system SHARS for archival and retrieval of shoeprints from a database is described. The initial description of the shoeprints is a set of drawn by user figures, which then is transformed to an attributed relational graph (ARG) describing shoeprints in terms of some primitives and their relations. To retrieve shoeprints from the database this ARG is matched with all ARGs in the database using proposed exact matching algorithms.

## 1. Introduction

So called "geometric" approach for retrieval of images from the database is to get a structural description of the image as a set of extracted geometric features (lines, arcs, figures, etc.) and their relations and to match it to model descriptions or to descriptions of images in the database. The automatic detection of such features is a subject of many researches but still remains on in general unsolved problem. One of the recent works in this field [1] represents an algorithm for search or adjusting figures consisting of line segments and arcs in grey scale images. This algorithm is based on a computationally efficient procedure for defining similarity between two images.

One of the powerful tools for describing structured objects are attributed relational graphs (ARGs). The nodes of ARG represent primitives or subpatterns of structured objects, and branches between nodes represent mutual relations between these primitives [2]. To match two graphs we need to find a one-to-one correspondence between their vertices and between their edges such that an incidence relationship is preserved [3]. This matching of two graphs is called graph isomorphism. If the matching is fulfilled between a graph and a subgraph of another graph, then it is called subgraph isomorphism or graph monomorphism. Terms of exact and best matching of ARGs are also important. Under exact matching there is subgraph isomorphism between two graphs with the same attributes on correspondent vertices and edges. The best matching means an optimal solution of the subgraph isomorphism problem when the distance between two matched graphs is minimal.

A number of algorithms are proposed for matching attributed relational graphs [2,4-10]. It is known that the graph matching problem belongs to the class of NP complete problems and computational complexity of any conventional search algorithm may grow exponentially when increasing the size of matched graphs. Some proposed algorithms have polynomial complexity for special types of graphs but still have exponential complexity at the worst case.

In [10] the algorithm for detection subgraph optimal isomorphism between two ARGs is proposed which is known to have an  $O(N^4)$  average case complexity and an  $O(N^5)$  in the worst-case complexity. In [11] the algorithm for detection subgraph isomorphism between two planar graphs in time  $O(c^{w \log w} n)$  is described where  $w$  and  $n$  are numbers of nodes in input and reference graphs and  $c$  is some constant value.

Widely used way to retrieve images from database that have common features with a current one is to get structural description of this image, transform it into an ARG, then to match it with ARGs representing structures of the other images in the database. In this paper some algorithms for description and retrieval of shoeprints along this way are considered.

The definition of the attributed relational graph and also steps of the description of the shoeprints in SHARS are shown in Section 2. Section 3 introduces some subgraph isomorphism based retrieval algorithms. Experimental results are presented in Section 4, and finally, the conclusions are given in Section 5.

## 2. Description of shoeprints

An attributed relational graph is defined as [5]

$$G=(N, B, A, E, G_N, G_B),$$

where

$N=\{n_1, n_2, \dots, n_{|N|}\}$  is a finite set of nodes;  $|N|$  is the number of nodes in  $N$ ;

$B=\{b_1, b_2, \dots, b_{|B|}\}$  is a set of ordered node pairs (or directed branches) i.e.,  $b=(n_i, n_j)$  for some  $1 \leq i, j \leq |N|$  denotes the branch emanating from node  $n_i$  to node  $n_j$ , and  $|B|$  is the number of branches in  $B$ ;

$A$  is an alphabet of node attributes;

$E$  is an alphabet of branch attributes;

$G_N : N \rightarrow 2^{|A|}$  is a function that defines the node attributes;



$G_B: B \rightarrow 2^{|B|}$  is a function that defines the branch attributes.

Let us consider now separate steps of description of the shoeprints and representing them by ARGs in image database system SHARS. The initial description of the shoeprint is a set of geometrical figures drawn by user in the image of this shoeprint. User can also define one or more characteristic subsets of this set, which must be on retrieved shoeprints.

Each of currently used figures (line, half of the circle, circle, polyline, polygon, square, triangle, rectangle, parallelogram, rhombus, rings, wave, tooth and position) contains so called control points and to draw or modify any figure in the image only these points must be defined.

On the next step every figure is represented by ARG with some attributes on its branches and nodes. Every node or branch of this graph has one context mark or attribute (centre of the circle, centre of half of the circle, end of line segment, line segment) and none or some of the following metric attributes: coordinates, radius, length and direction of the line segment, cross angle and relation of lengths of two neighbour line segments. The number of used metric attributes depends on some options of the program and also of the possibility for the user to define the position and orientation of the shoeprint in the image as well as the resolution of this image.

As an example let us consider attributed relational graph  $G_p$  of the figure „polyline“ that consists of two line segments. This graph consists of three nodes and two branches. The nodes of the graph  $G_p$  have the context mark „end of line segment“ and the branches have the context mark „line segment“. The nodes of the graph have metric attributes „coordinates“ if: 1) the resolution of the image is known, 2) the user has estimated the position and direction of the shoeprint in the image by drawing the figure „position“ and 3) the option „Coordinates“ is switched on. The middle node of the graph in any case has also metric attributes „lengths relation of neighbour line segments“ and „angle between neighbour line segments“. The branches of the graph have the metric attribute „length“, if the resolution of the image is known, and the metric attribute „direction“, if the user has defined the orientation of the shoeprint in the image and the option „Angle“ is switched on.

It follows that

$$G_p = (N, B, A, E, G_N, G_B),$$

where

$$N = \{n_1, n_2, n_3\};$$

$$B = \{b_1, b_2\};$$

$A = \{ \text{end of the line segment, coordinates, lengths relation of neighbour line segments, angle between neighbour line segments} \}.$

$$E = \{ \text{length, direction} \}.$$

Attributed relational graphs that define characteristic groups of figures in the shoeprint are similar to ARGs of the figure „polyline“ but instead of the node attribute „end of line segment“ they have the node attribute „label of figure“.

### 3. Retrieval algorithms

Retrieval algorithms are based on matching (subgraph isomorphism) of the input ARG with ARGs of shoeprints in the database and also on the distance measure between these graphs. For calculating the distance between two ARGs the concept of error-correcting transformation [5,7,10] is adopted. The cost of matching two ARGs (or distance between these graphs) is defined as the cost of the sequence of transformations with minimal total cost that must be performed on one of the two ARGs in order to produce another ARG. These transformation operations are: *node insertion, node deletion, branch deletion, branch insertion, node label substitution and branch label substitution*. The costs corresponding to each operation are:  $w_{ni}$ ,  $w_{nb}$ ,  $w_{bi}$ ,  $w_{bd}$ ,  $w_{ns}$  and  $w_{bs}$  respectively. It is reasonable in our case to assume that  $w_{nd} = w_{bd} = 0$ ,  $w_{ni} > 0$ ,  $w_{bi} > 0$ ,  $w_{ns} = w_{bs} = 1$  and the distance  $dist(x,y)$  between two numeric attributes  $x$  and  $y$  is equal to  $v = \text{abs}(x-y)$  if  $v$  exceeds some threshold value and to 0 otherwise. It can be defined that we have the exact matching if the distance between two graphs is equal to 0 and the best matching otherwise. In both cases the part of one graph must be defined that has the minimal distance with another graph but for exact matching this distance is equal to 0 and less computational time is usually required.

ARGs, corresponding to currently use in SHARS figures, are either chains or cycles. That is why the way of checking exact matching between two ARGs during moving lesser graph along greater one can be used. The computational complexity of this algorithm is  $O(M*N)$ .

There are some criteria for matching of input and reference descriptions of shoeprints implemented in the system SHARS. By default two descriptions are matched if at least one figure in input description is exact matched with one or more figures in the reference shoeprint. The user may also define the number of pairs of figures or group of figures in the input shoeprint that must be matched with ARGs of the reference shoeprints.

To match the group of figures we need to find subgraph isomorphism between two graphs. One of them is the complete (fully connected) graph such that nodes of this graph represent different figures in the shoeprint from the database and each of these nodes is connected to all other nodes of the graph. The other graph  $G$  represents the chain of figures in the input shoeprint. Each node of this graph has a label of the corresponding figure in the group and number of nodes in  $G$  is equal to number of figures in the group. The following algorithm checks some necessary conditions of subgraph isomorphism between these two graphs. During the first part of the algorithm for each node of the graph  $G$  the list of exact matched with this node figures from the reference description is constructed. During the next step all possible exclusings of those figures from these lists that do not satisfy to local relations between the neighbour figures in the group are carried out. It means that each figure  $F$  must be excluded from the list if there are no in neighbour lists two other figures that satisfy with  $F$  to local relations of neighbour relations of figures



in the group. The process of excluding figures is stopped when additionally none of the figures can be excluded. Therefore the input and reference graphs are considered as matched if the final lists of these figures are not empty in spite of the fact that there is exactly subgraph isomorphism between these graphs only when each of lists has only one label or figure and all these labels are different.

For matching ARGs that represent figures of other types than currently used in SHARS one of the algorithms [2,4-10] can be implemented. For this aim new algorithm for exact matching ARGs was also developed which is described below.

The main problem using the error-correction distance measure is the computational complexity of the matching that may grow up exponentially when the graph size increases. To reduce the complexity original graphs can be decomposed into smaller subgraphs (basic attributed relational graphs or BARGs) and the matching of the original graphs to fulfil through the matching of these decomposed subgraphs [10]. The BARG represents a one-level tree, which consists of a node (root), and every node (descendant) connected to that node by one emanating or incoming branch. The cost of matching two BARGs  $G_k$  and  $G_p$  can be defined as

$$\text{Distance}(G_k, G_p) = w_{ns} * \text{dist}(r_k, r_p) + \text{dist}(b's, e's),$$

where

$\text{dist}(r_k, r_p)$  is the distance between root node labels and attributes of two BARGs that is computed depending on their data types,

$k$  and  $p$  are the number of branches connected to the root node BARGs being matched, and

$\text{dist}(b's, e's)$  is calculated as the minimum of a weighted bipartite graph constructed from branches of BARGs as their nodes and its structure is as shown in Fig. 1.

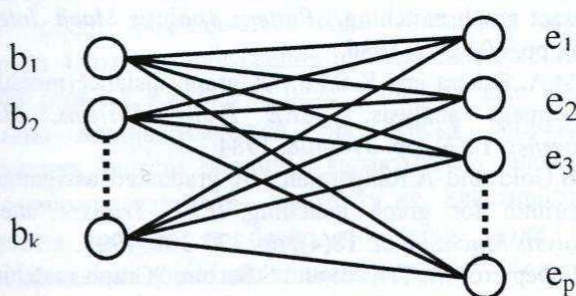


Fig. 1. Weighted bipartite graph corresponding to branches distance.

The minimum of a weighted bipartite graph is defined as the complete matching with the minimum sum of the matching edges weights. There are some algorithms for such matching with an  $O(N^2)$  average case complexity and an  $O(N^3)$  worst-case complexity [12].

Exact matching between branches of two BARGs means that  $\text{dist}(b's, e's) = 0$ , i.e., there is one-to-one correspondence between branches of these BARGs and correspondent branches have the same attributes. To

define such exact matching it is necessary to exclude from bipartite graph in Fig. 1 those branches that have non-zero weights and then to use one of the algorithms f.e., [13] to find assignment for remain branches.

Let  $G_i$  and  $G_r$  be input and a reference ARGs. The first step of the algorithm for checking exact matching between these two graphs consists in their decomposition into BARGs. Then the distances between decomposed BARGs are defined, it means that a distance matrix  $D$  between BARGs of reference and input graphs is constructed. The  $D_{ij}$  element of this matrix represents the distance between the  $i$ th BARG in the input graph and the  $j$ th BARG in the reference graph, and  $D_{ij}=0$  if there is exact matching between these two BARGs and  $D_{ij}>0$  otherwise. On the base of distance matrix for every node of lesser graph (having less nodes than another one) a list of BARGs of greater graph that have exact matching with this node and not lesser number of connected branches is made out.

During the next step of the algorithm we scan constructed lists of BARGs and exclude every BARG  $B$  in the current list if the list of BARGs in the neighbour node is either empty or contains only BARGs that are not neighbours to  $B$  in the reference graph. During this step can be also helpful some additional tests, f.e. : a) if some list  $L$  contains only one BARG this BARG can be excluded from all other lists or b) if the number of different BARGs in all lists less than the number of nodes in input graph the input and reference graphs can not be matched.

This process is stopped when additionally none of the BARGs can be excluded. Therefore the input and reference graphs are not matched if the final lists of BARGs are empty and are exact matched if each of these lists contains only one BARG and all these final BARGs are different. In the case when final lists contain more than one BARG additional testing is needed.

Let us estimate computational complexity of the algorithm. Suppose that the input and reference graphs have  $M$  and  $N$  nodes relatively. To decompose both  $G_i$  and  $G_r$  into BARGs we need quadratic time for each graph, i.e., this step has a time complexity of  $O(M^2 + N^2)$ . The computational complexity of exact matching between two BARGs is  $O(M*N)$ , assuming that any node is connected to all other nodes in the graph, and the computational complexity of calculating the distances between all BARGs of both input and reference graphs is  $O(M^2 * N^2)$ . The computational complexity of the second (BARGs excluding) step of the algorithm is  $O(M^2 * N^3)$ ,  $M < N$  in the worst case. In summary, the computational complexity of the proposed algorithm for checking the exact matching between two ARGs is also  $O(M^2 * N^3)$  in the worst case.

## 4. Experiments

During experiments 150 shoeprints were encoded and saved in the database. After that the testing group of the 25 different shoeprints was formed and shoeprints of this group were then encoded and retrieved by some users.



Each shoeprint from this group has one or more shoeprints of the same type in the database, which differ from each other by size, position and many other features, for example some of these shoeprints are only parts of other ones (Fig. 2).

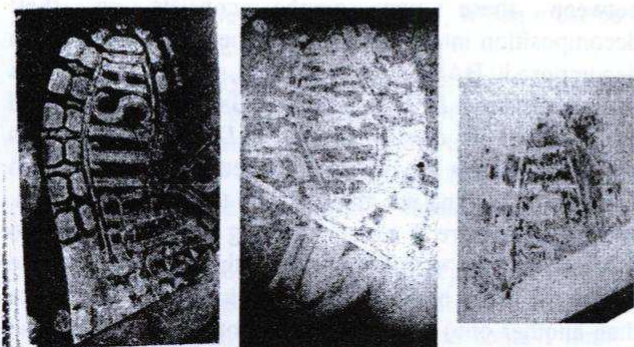


Fig. 2. An example of the same type shoeprints in the database.

Many shoeprints in the database are not supplied by metric information because of the absence of data about resolution of the image or the position of the shoeprint in this image. The number of retrieved shoeprints by one figure or by one pair of figures for such shoeprints is not so small. As follows from experimental data the most effective possibility for the retrieval of most shoeprints from a database consists in defining of a group of figures in the input shoeprint. This possibility provides a sharp reduction of number of retrieved shoeprints and preserves the majority of the same type retrieved shoeprints from the database.

It is important that retrieval of the shoeprints from database is fulfilled using not the names and parameters of figures but using attributed graphs that represent these figures and also on the base of subgraph isomorphism of these graphs. This fact reduces dependence of the retrieval results on subjective actions of the user during describing or coding input shoeprints because in many cases the same object in the shoeprint can be described by figures that have different names and/or different numbers of its components but attributed graphs that represent these figures can be matched. Introduced options and some possibilities for the definition of an inquiry to a database (retrieval by figure, by pairs or by groups of figures) can be also useful in this direction.

## 5. Conclusions

In this paper ARG based approach and algorithms for the retrieval of shoeprints from a database are proposed. The proposed algorithms are based on the exact matching of ARGs that describe input and reference shoeprints. The distance is calculated based on the concept of error-correcting transformations with taking in account some peculiarities of the shoeprints description and retrieval problem. From experimental results the following remarks can be concluded.

The developed tools can be used for the description and retrieval of shoeprints from a database. The proposed algorithms for detection subgraph isomorphism between input and reference ARGs and also the introduced different possibilities for the definition of an inquiry to a database reduce dependence of the retrieval results on subjective actions of the user during describing or coding input shoeprints. On the next step the algorithms for optimal subgraph isomorphism will be tested and added to the system for the further reduction of this dependency.

The implemented ARGs based tools and algorithms are an addition but not a substitution for the other method implemented in SHARS that is based on coding shoeprints by rather global geometrical features, that describe shoeprint in its different parts.

## References

1. I.Aksak, Ch. Feist and others, "Detection of the Objects with the Given Shape on the Grey-Valued Pictures," CAIP'97 "Computer Analysis of Images and Patterns", 1997, Kiel, p.551-558
2. A. Sanfelio and K.S. Fu, "A distance between attributed relational graphs for pattern recognition," *IEEE Trans. Systems, Man Cybernet.* **13(3)**, pp. 353-362, 1983.
3. Narsingh Deo, "Graph Theory with Applications to Engineering and Computer Science," Prentice-Hall, Englewood Cliffs, New Jersey, 1995.
4. A.K.C. Wong, M. You and S.C. Chan, "An algorithm for graph optimal monomorphism," *IEEE Trans. Systems, Man Cybernet.* **20(3)**, pp. 628-636, 1990.
5. W.Tsai and K.S.Fu, "Subgraph error-correcting isomorphism for syntactic pattern recognition," *IEEE Trans. Systems, Man Cybernet.* **13(1)**, pp. 48-62, 1983.
6. M.A. Eshera and K.S.Fu, "An image understanding system using attributed symbolic representation and inexact graph matching," *Pattern Analysis Mach. Intell.* **8(5)**, pp. 604-617, 1986.
7. M.A. Eshera and K.S.Fu, "A graph distance measure for image analysis," *IEEE Trans. Systems, Man Cybernet.* **14(3)**, pp. 398-408, 1984.
8. S.Gold and A.Rangarajan, "A graduated assignment algorithm for graph matching," *IEEE Trans. Pattern Analysis Mach. Intell.* **18(4)**, pp. 377-388, 1996.
9. F.Depiero, M. Trivedi and S.Serbin, "Graph matching using a direct classification of node attendance," *Pattern Recognition*, **29(6)**, pp. 1031-1048, 1996.
10. Yasser El-Sonbaty and M.A.Ismail, "A new algorithm for subgraph optimal isomorphism," *Pattern Recognition*, **31(2)**, pp. 205-218, 1998.
11. D.Eppstein, "Subgraph isomorphism in planar graphs and related problems," Department of information and computer science, University of California, Irvine, CA 92717, Tech. Report 94-25, p. 27, 1994.
12. S.Yamada and T.Hasai, "An efficient algorithm for the linear assignment problem," *Elect. Comm. Japan*, Part 3, **73(12)**, pp. 28-36, 1990.
13. O.Оре. Теория графов. Москва, 352 с, 1968.



# ВИКОРИСТАННЯ "ВЛАСНИХ ОБЛИЧ" ДЛЯ РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ

Левко О. В., Нюнькін К. М.

Інститут проблем штучного інтелекту

83050, м.Донецьк, проспект Богдана Хмельницького, 84

Тел.: (0622)926082, факс: (0622)926082

E-mail: kmn@iai.donetsk.ua

## ABSTRACT

In this paper usage of a method of image decomposition on eigenfaces for human face recognition is considered. The experiments permitting are conducted to define limitations of this approach.

## ВСТУП

Комп'ютерне розпізнавання людських облич привертає в останній час все більше уваги. З одного боку це пов'язано з численними задачами, в яких може застосовуватися ця техніка, з іншого - навалює зростаючі потужності сучасних комп'ютерів роблять можливим вирішення таких задач.

Всю безліч підходів до розпізнавання облич можна поділити на дві великі групи: методи, що використовують окремі деталі зображення, й алгоритми, що розглядають все зображення цілком. В роботі розглядається один з підходів другий групи - розклад зображення по "власним обличчам" або розклад Karhunen-Loeve [1.2], що вже досить давно використовується для стиску зображень. Поняття "власні обличчя" вводяться по аналогії з власними векторами і мають ту же властивість: їх можна вибрати як базис, по якому може бути розкладений довільний елемент простору облич.

## 1. ПРОСТІР ОБЛИЧ

Зображення в комп'ютері представляється матрицею окремих пікселів, яка має  $V$  рядків по  $N$  пікселів у кожному. У випадку сірого напівтонового зображення значення пікселів визначаються яскравістю відповідної точки зображення. Таке зображення можна уявити у вигляді вектору. Для цього необхідно з'єднати послідовно всі рядки зображення, отримавши один довгий рядок. Значення кожного пікселя зображення буде

відповідати одній компоненті вектору. Отриманий вектор належить простору усіх можливих зображень, що мають розмір  $N \times V$  пікселів. Цей вектор має вимірність  $N \times V$ . Ясно, що працювати з векторами такої великої вимірності дуже складно.

При вирішенні задачі розпізнавання людських облич нас цікавлять не будь-які зображення, а тільки зображення облич. Всі обличчя схожі друг на друга: всі вони мають ніс, очі, губи й інші характерні деталі, розташовані на певних місцях. Тому зображення облич в просторі зображень компактно згруповані і утворюють порівняно невеликий кластер. Отже, повний простір зображень є надмірним для опису облич і можна побудувати спеціальний простір, що буде описувати обличчя оптимальним чином, маючи при цьому значно меншу вимірність, ніж вихідний простір зображень. Природно цей простір назвати простором облич [3].

Нехай  $I=N \times V$  - кількість пікселів у зображенні та число компонентів вектору простору зображень;  $K$  - кількість зображень облич в навчальній вибірці;  $X$  - матриця вимірності  $K \times I$ , що визначається навчальною вибіркою.  $k$ -ий стовбчик матриці відповідає  $k$ -му зображенню обличчя навчальної множини.

Надмірність простору зображень є наслідком того, що кожний піксель обличчя корелює з іншими пікселями і коваріаційна матриця для множини облич не є діагональною.

Для оптимального подання облич необхідно знайти такий простір, в якому коваріаційна матриця для навчальної множини облич буде діагональною. Ця умова буде виконуватися, якщо за базисні вектори простору взяти власні вектори. Базисні вектори простору облич називають основними компонентами. Введемо матрицю  $Y$ , що містить основні компоненти,  $k$ -ий стовбчик матриці - подання  $k$ -го зображення обличчя в просторі облич. Її коваріаційна матриця  $\Sigma_Y = Y \cdot Y^T$  буде діагональною.

Для отримання основних компонентів використовуються лінійні перетворення. Нехай  $P$  -



матриця перетворення, тоді:  $Y=P^T*X$ ,  $X=P*Y$ . В [3] запропоновано за матрицю  $P$  взяти власні вектори матриці  $\Sigma_X$ :  $\Sigma_Y*P=P*\Lambda$ , де  $\Lambda$  - діагональна матриця, що містить власні значення матриці  $\Sigma_X$ . Тоді  $\Sigma_X=P^T*P*\Lambda=\Lambda$ , тобто,  $\Sigma_Y$  - діагональна матриця, що містить власні вектори  $\Sigma_X$ , чого і вимагалось досягнути. Для знаходження власних векторів використовується метод обертань [4].

Величина власного значення визначає важливість або вагу відповідного йому власного вектору для подання простору. Власні значення можна відсортувати за убаванням, та відкинути вектори, що відповідають найменшим власним значенням, не вносячи цим істотної помилки в подальше розпізнавання [3]. При цьому вимірність простору ще зменшиться в декілька раз.

Кожний із знайдених власних векторів визначає точку в просторі зображень; тобто являє собою деяке зображення. Більше того, отримані зображення схожі на обличчя. Тому їх аналогічно до векторів називають "власними обличчями" (ВО), підкреслюючи цим їхній зв'язок зі зображеннями облич. Приклад зображень перших ВО, упорядкованих за убаванням відповідних їм власних значень, наведений на малюнку 1.



Мал. 1. Приклад перших ВО

Зазначимо, що перше ВО є середнім всіх облич, поданих в навчальній множині; друге відповідає напрямку найбільших змін; третє - напрямку наступних по величині змін; і т. д. Зображення будь-якого обличчя може бути уявлене в вигляді лінійної комбінації ВО. На малюнку 2 показаний приклад розкладу одного з облич (тільки перші чотири члена розкладу).



Мал. 2. Приклад розкладу зображення обличчя по базису ВО

Після побудови базису ВО, з'являється можливість подання довільного зображення в знайденому базисі. Причому коефіцієнт розкладу при першому власному векторі може служити детектором наявності людського обличчя на зображенні. Експериментально встановлено, що

коефіцієнт розкладу при першому векторі для зображень, що містять обличчя, складає в середньому 0.96; а для інших зображень - 0.83. Проведені дослідження показали можливість успішного застосування цього методу. Відносна кількість зображень, що були невірно ідентифіковані (обличчя / не обличчя), не перевищувала 2%.

Коефіцієнти розкладу являють собою вектор, вимірність якого дорівнює кількості ВО і звичайно складає декілька десятків. Маючи коефіцієнти розкладу зображення довільного обличчя і знаючи набір ВО, по якому цей розклад виконувався, можна відновити вхідне зображення. На малюнку 3 наведені вхідне і відновлені зображення одного і того ж обличчя при різній кількості коефіцієнтів розкладу. Точність відновлення тим більша, чим більше ВО використовується при розкладі. Тим не менше вже при 50 членах розкладу відновлене зображення мало відрізняється від оригіналу. Таким чином, замість зображень облич можна розглядати вектори в  $n$ -мерном просторі.



Мал. 3. Відновлення зображення обличчя людини по різній кількості ВО: а) вхідне зображення; б) - е) відновлене по 100, 50, 25 і 10 ВО відповідно

Для відновлення зображень на мал.3 ВО обчислювалися по всім 212 зображенням (приблизно по 20 зображень облич кожної з 11 людей) і їх було отримано також 212. Після цього отримані ВО ранжувалися в порядку убавання відповідних їм власних значень, і при відновленні використовувалося тільки 100, 50, 25 або 10 перших з них.

Слід підкреслити, що для успішного застосування цього методу зображення облич повинні бути справді "схожі" друг на друга, тобто обличчя повинні мати однаковий масштаб, ракурс і розташуватися в одному і тому ж місці на зображенні. Бажано, щоб обличчя займало все зображення. Це дозволить виключити вплив фону.

## 2. РОЗПІЗНАВАННЯ ЛЮДСЬКИХ ОБЛИЧ

Подання зображень людських облич як векторів в просторі ВО дає можливість вирішувати задачу розпізнавання облич. Для цього спочатку за декількома десятками зображень довільних облич



формується ВО. Після цього декілька зображень кожного з "відомих" людей розкладаються по цим ВО і вектори коефіцієнтів розкладу зберігаються в базі даних як еталони для цих людей. Зазначимо, що зображення, що розкладаються, могли і не використовуватися при формуванні ВО. База даних еталонних векторів в будь-який момент може бути доповнена новими обличчями. Набір ВО при цьому залишається незмінним. Кожне нове введені зображення обличчя людини також розкладається по тим же самим ВО, що й еталонні зображення.

Еталонні вектори, що отримані розкладом по ВО зображень обличчя кожної людини, утворюють класи. Тепер можна ставити і вирішувати задачу приналежності вектору, отриманого розкладом по ВО зображення довільного обличчя, одному з класів.

## 2.1. БАЗИ ДАНИХ

В експериментах по розпізнаванню обличчя використовувались зображення з декількох баз даних. Оскільки розміри обличчя на зображеннях різних баз даних істотно розрізняються, експерименти з зображеннями кожної бази даних проводилися незалежно від інших. У базах були представлені обличчя людей різної статі та расової приналежності, з вусами, бородами, окулярами та без них. Варіювався також вираз обличчя. Нижче наведені більш детальні параметри кожної з баз даних.

*Манчестерська база даних:* 214 зображень розміром 64x64 (20 зображень обличчя кожного з 9 чоловіків і 3 жінок). Мають місце слабкі повороти голови, трохи розрізняється масштаб, на окремих кадрах наявні окуляри.

*База даних дослідної лабораторії Оліветті:* 400 зображень розміром 92x112 (по 10 зображень обличчя кожного з 35 чоловіків і 5 жінок). На зображеннях однієї й тієї ж людини розрізняються вираз обличчя та ракурс. Масштаби практично однакові.

*База даних Йельського університету:* 165 зображень розміром 320x243 (по 11 зображень обличчя кожного з 14 чоловіків та 1 жінки). На зображеннях однієї й тієї ж людини розрізняються вираз обличчя, умови освітлення і наявність окулярів. Масштаб та ракурс практично однакові.

*База даних університету Берну:* 150 зображень розміром 512x342 (по 10 зображень обличчя кожного з 15 чоловіків). На зображеннях однієї й тієї ж людини розрізняються вираз обличчя і ракурс. Масштаб практично однаковий.

*База даних MIT:* 450 зображень розміром 128x120 (по 27 зображень обличчя кожного з 16 чоловіків). На зображеннях однієї й тієї ж людини

розрізняються вираз обличчя, ракурс, масштаб та умови освітлення. Тільки серед зображень цієї бази є бічні нахили голови.

## 2.2. РЕЗУЛЬТАТИ ЕКСПЕРИМЕНТІВ

Перед виконанням розпізнавання зображення всіх баз даних були заздалегідь оброблені: фон був обрізаний і залишено тільки обличчя. Для порівняння манчестерська база даних, на зображеннях якої фон слабкий, була оброблена також і без попереднього виділення обличчя.

Експерименти проводилися слідуєчим чином. З зображень бази даних формувалася навчальна вибірка, розмір якої варіювався від всієї бази до одного зображення на людину. По зображенням навчальної вибірки обчислювалися ВО. Після цього зображення кожного з наведених в базі людей ділилися на дві рівні частини. По першій половині обчислювалися еталонні вектори коефіцієнтів розкладу, а зображення другої розглядалися як невідомі і підлягали розпізнаванню. Після цього половини мінялися місцями. Еталонні вектори коефіцієнтів розкладу, стосовні до однієї людини, осереднювалися для отримання середнього вектору.

Кожне нове зображення обличчя невідомої людини розкладалося по тим же самим ВО, що й еталонні зображення. Вважалася, що обличчя на введеному зображенні належить тій людині, чий середній еталонний вектор виявлявся найближчим.

Частки невірно розпізнаних обличчя при різних розмірах навчальної вибірки і кількості ВО наведені в таблиці 1.

Результати експериментів показують, що даний метод розпізнавання людських обличчя надто чутливий до фону на зображенні. В першій серії експериментів зображення манчестерської бази даних були оброблені разом з фоном (перший рядок в табл.1). Частка помилок при розпізнаванні в цьому випадку була дуже висока і досягала половини. Це, звичайно, менше, ніж імовірність помилок при випадковому виборі, але значно більше, ніж можна допустити в реально працюючій системі. Усунення фону дозволило в 4-5 раз знизити частку помилок (третій рядок в табл.1).

Ще один шлях зниження помилок - використання більш досконалої процедури розпізнавання. Так пошук найближчого з всіх еталонних векторів (а не середнього) дозволяє зменшити імовірність помилки в півтора рази (другий рядок в табл. 1). Це означає, що кластери, утворені еталонними векторами окремих людей, не мають сферичної форми і погано рознесені в просторі.



Найкращі результати були отримані для зображень бази даних Йельського університету, незважаючи на те, що зображення обличчя однієї й тієї ж людини розрізнялися виразом, умовами освітлення та наявністю окулярів. Отже ці умови не такі важливі як масштаб обличчя і ракурс голови, які на цих зображеннях практичні однакові.

Більш за все помилок було зареєстровано при розпізнаванні обличчя на зображеннях бази даних MIT, серед яких були наявні бічні нахили голови та істотна зміна розміру обличчя. Справді, у випадку

бічного нахилу голови зображення обличчя сильно викривляється і мало схоже на зображення, де голова не нахилена. Зображення, на яких обличчя мають різні розміри також значно розрізняються.

При обробці зображень з баз даних дослідної лабораторії Оливетти й університету Берну, що містять зображення обличчя при поворотах і вертикальних нахилах, отриман невеликий відсоток помилок, отже такі зміни ракурса припустимі для систем розпізнавання, що основані на поданні людських обличчя через власні обличчя.

Табл.1 Результати експериментів

N	База Даних	Кількість зображень в навчальній вибірці	Кількість ВО	Розмір зображення	Еталони/Невідомі		
					Всі/всі	1/2	2/1
1	Manchester*	212	64	64x64	4.7%	48.1%	50.0%
		66	64		12.3%	49.0%	40.7%
		33	32		13.2%	51.0%	46.3%
		11	11		20.8%	56.7%	50.0%
2	Manchester**	212	64	64x64	0%	35.6%	35.2%
3	Manchester	212	32	32x48	3.3%	15.4%	10.2%
		33	32		3.3%	17.3%	9.3%
		11	11		9.4%	26.9%	13.9%
4	ORL	400	32	92x112	4.0%	21.0%	16.0%
		40			5.5%	18.0%	20.0%
5	Yale	135	32	112x128	2.22%	5.0%	12.0%
6	Bern	150	32	96x112	6.7%	16.0%	13.3%
7	MIT	430	32	64x64	29.1%	34.4%	44.7%

\* Манчестерська база даних оброблена без виділення облич.

\*\* При розпізнаванні використовувався найближчий еталонний вектор.

## ВИСНОВКИ

Таким чином, розклад зображень людських обличчя на ВО може бути використаний при побудові системи розпізнавання людей за їхніми обличчями.

Методи, що основані на ВО дозволяють розпізнавати людей за їхніми обличчями незалежно від вираза обличчя та умов освітлення.

Цей метод має високу чутливість до навколишнього фону, розміру обличчя на зображенні, а також до бічних нахилів голови. При створенні реально працюючої системи, що розпізнає, необхідно передбачити автоматичне визначення положення, розміру і нахилу обличчя на зображенні з тим, щоб компенсувати викривлення, які вносяться, і усувати з розгляду навколишній фон.

Слід також приділити особливу увагу застосуванню ефективних засобів вирішення задачі розпізнавання образів, що здатні істотно підвищити якість розпізнавання.

## ЛІТЕРАТУРА

1. Turk M. A., Pentland A. P. Face Recognition Using Eigenfaces.//Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1991.-p. 586-591.
2. Pentland A., Moghaddam B., Starner T., Turk M. View-based and Modular Eigenfaces for Face Recognition.//Proc. IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, 1994.
3. Romdhani S. Face Recognition Using Principal Component Analysis./The Msc Thesis, 1996.-57p.
4. Березин И.С., Жидков Н.П. Методы вычислений. - М., Наука, 1965.



# CLASSIFICATION OF TEXTURE IMAGES WITH THE USAGE OF ARTIFICIAL NEURAL NETS AND HIGHER-ORDER STATISTICS

A.M. Akhmetshin, D.A. Mikhalev  
Dnepropetrovsk state university

## Abstract

The object of the research is photos of metal-graphic microstructures.

The aim of the work is the choice of algorithmic and elaboration of software for effective classification of texture images with the help of neural nets, and also classification the process of exactness rise researches of questions and diminution of general processing time of texture images.

The method of the research is comparative analysis of two classifiers, the first one is built on Hamming neural net, the second one is built on Kohonen neural net.

Results of the researches were employ for classification of the photos of metal-graphic microstructures, which were obtain through digital photoscanner.

## 1. Introduction

The concept of the texture is more often used in different areas of a science and engineering, such as hylology and research of mining and metallurgical processes, the control of the quality of industrial production, image processing of the metal-structures. It follows from the fact, that the surfaces of a great number of real objects in these fields are textures and under condition of their homogeneity, they can be described with the help of a comparatively little number of parameters. By virtue of this, fact, the solution to this problems of the texture classification is actual, in particular, in geology and geodesy, in a snap analysis of micro- and of macrostructures in metallurgy.

The general scheme of texture classification is shown in the fig. 1.

As well as in the system of pattern recognition, two main tasks can be determined here: the first one is feature extraction and the second one is classification of textures to the evaluated features.

Two different approach to extraction of texture features is considered in the work. The first one is connected with higher-order statistics and is based on calculation of third-order cumulant. The second one is based on the statistics of the second order and consists in calculation 10 texture features from spatial gray-level dependence (SGLD) matrices.

Structurally the classifier is represented by two principally different schemes of artificial neural net. The

first classifier, which is used during calculation of third-order cumulant, consists in application of Hamming neural net, the second one, which is used during calculation of statistics of second order consists in application Kohonen's self-organizing maps.

There are 5 test images of microstructures, each of which represents the separate class. The size of each image for a simplicity of a illustration is selected 128×128. All these images are shown on fig. 2.

One of the main problems during the classification of textures consists in uncertainty of texture images orientation. In this connection each texture, presented on the fig. 2, in addition can be rotated on 90°, 180° and 270°. Thus, 20 images from 5 classes are presented.

It is required to interpret the input image correctly and refer it to one of 5 classes.

## 2. Feature extraction

### 2.1. Higher-order statistics

For the selection of informative features of textures the apparatus of higher-order statistics can be used, and as features the estimation of 3-rd order cumulants can be used [6].

Cumulant is one of the numerical characteristics of random variables, it is closed to concept of the moment of the high order. In practice, for the finite sampling of data  $\{x(n)\}_{n=0}^{N-1}$  the cumulantes are calculated according the following formulas:  
second-order cumulant:

$$\hat{C}_{xy}(k) = \frac{1}{N_3} \sum_{n=N_1}^{N_2} x^*(n)y(n+k), \quad (2.1)$$

third-order cumulant:

$$\hat{C}_{xyz}(k,l) = \frac{1}{N_3} \sum_{n=N_1}^{N_2} x^*(n)y(n+k)z(n+l), \quad (2.2)$$

where  $N_1, N_2 \in [0, N]$ .  $N_3$  are usually set equal to  $N$  and calculated their estimations are asymptotically unbiased.

The algorithm for application of the higher-order statistics apparatus (algorithm 1) for obtaining the matrix of informative features of textural images consists of the following steps [8]:

1. Texture image  $I(x, y)$  is transformed in one-dimensional array  $y(k)$  by joining of successive rows in a sequence.



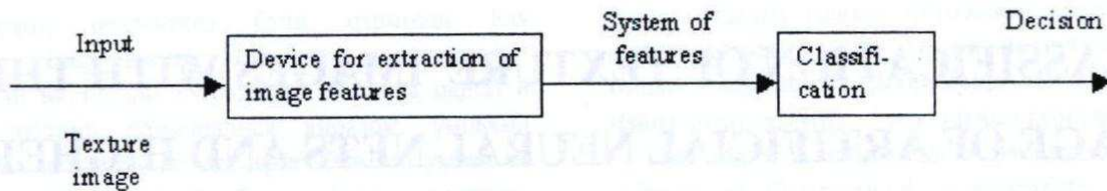


Fig. 1. The generalized scheme of texture classification and recognition.

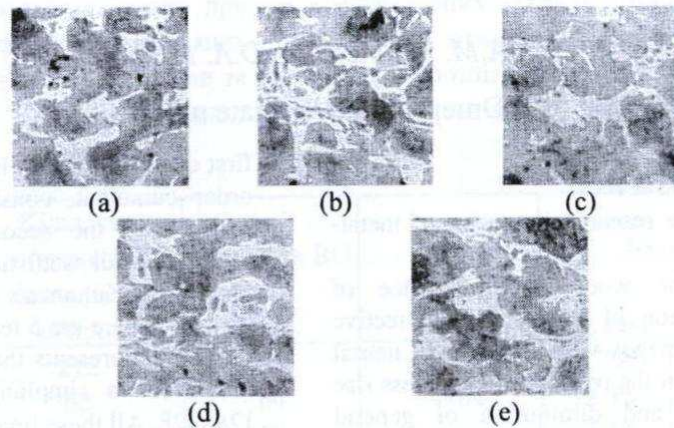


Fig. 2 Textures of different classes

2. The dimension of matrix of estimated cumulants is  $2*n+1$ .
3. Time-series  $y(k)$  is segmented into records of  $m$  samples each, with no overlap; the biased estimates of the third-order cumulants are obtained from each segment on the formula (2) and then averaged.
4. As result the matrix  $C(x, y)$  is calculated where  $(i, j)$  element of matrix is the estimate of  $C_{3y}(i-n-1, j-n-1)$  for  $i, j = 1, \dots, 2*n+1$ .

## 2.2. Spatial gray-level dependence (SGLD) matrices

The essence of one of most widely used methods of description of texture consists in calculation of texture features from spatial gray-level dependence (SGLD) matrices.

The  $(i, j)$ th element of the SGLD matrix is the joint probability that gray levels  $i$  and  $j$  occur in direction at a distance of  $\theta$  pixels apart in an image. Ten texture features, including correlation, entropy, angular second moment, mean, difference entropy, contrast, inverse difference moment, deviation, different angular second moment, different mean were used. The definition of the texture measures are given in the literature [1, 2, 4].

These measures were extracted from each SGLD matrix at pair distance  $d = 5$  and four directions ( $0^\circ, 90^\circ, 180^\circ,$  and  $270^\circ$ ). These features contain information about image characteristics such as homogeneity, contrast, and the complexity of the image.

## 3. Classification

### 3.1. Hamming neural network

Among the different configurations of artificial neural networks there are the ones, in which classification according the principle of training, strictly speaking, doesn't approach either training with the

teacher, or training without the teacher [4]. In such networks the weight coefficients of synapses set up only once before the beginning of work of the network on the basis of the information about processed data, and all training of the network is amounted to this calculation. The Hamming network falls into the class of networks with similar logic of work which is usually used for organization of associative memory. This network is used in case of the binary input images.

Neural net solution of this task can be obtained on the basis of the Hamming architecture. The network has one layer of identical neurons, number of which is equal to quantity of classes. Each neuron is coupled to each the input, which number is equal to the dimension of considered library of images.

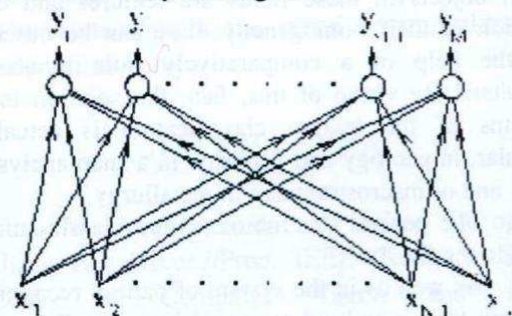


Fig. 3. Hamming neural network

The network choose the sample with minimum of Hamming distance to unknown input signal, therefore as a result one output of the networks appropriating to this sample will be activated.

It is reached by introduction of additional feedbacks between neurons arranged by a principle "lateral braking". Such mechanism is called "Winner Take All" (WTA) [4, 7].



### 3.2. Self-organizing maps

Self-organizing map (SOM) is neural net with straight ties, in which the teaching algorithm without teacher is used. By means of process, which named is self-organize, from outgoing elements SOM will make the topologic conception of initial data.

Algorithm SOM is based on competitive teaching without teacher. It provides reflection, preserving the topology, from space of big dimension on elements of map. Elements of the map, or neurons, usually make a two-dimensional grid. Thus this reflection is the reflection from space of big dimension on flatness. Property of topology preservation of signifies, that SOM distributes the similar vectors of input data on neurons: points, which is near to each other in space of inputs, on near situated elements of map in SOM.

Kohonen's net consist of  $r$  neurons, according to number of classes: to each class corresponds its neuron, and inside out. Input vector  $X$  is given on net input, and on input of each neuron. Its dimension  $n$  depends on task. Each neuron calculates function  $f$  from input vector  $I$ , and as answer of net the number of that neuron is given out (this vector get in class with this number), significance of function  $f$  in which receive maximum (or minimum, as more convenient). So the principle "winner take all" (WTA) is used.

## 4. Realization

### 4.1. Classifier, built on the base of the third-order cumulant

At the first stage of the calculation of texture of informative features the algorithm 1 was used. The size of segments, into which initial array is divided, is selected equal to  $m = 128$ . The size of the matrix of 3-rd order cumulant  $C$  is equated  $2^{*n+1} = 129$ , i.e.  $n = 64$ . The example of such matrix for the class, presented on fig. 2.a, is shown on fig. 4.

As it has been already pointed, that the input signals for Hamming network are binary. Therefore during the transition to the stage of the classification of initial images it is necessary to make them binary. The threshold for binarization was set equal to 0. Thus, if  $C(i, j) \geq 0$ , then that  $B(i, j) = 1$ , if  $C(i, j) < 0$  then  $B(i, j) = 0$ , where  $i, j = 1, \dots, 2^{*n+1}$ ,  $B$  is binary matrix of informative features of texture.

The example of the matrix  $B$  for the class, presented in the fig. 2.a, is shown in the fig. 5.

The fig. 5 demonstrates, that binary feature matrixes for different orientations of textures differ considerably. It is connected, first of all, to the following: for obtaining of the rating matrix of cumulants the initial input was transformed in to the one-dimensional array, and as the image rotates the origin displaces it causes the change of.

It was empirically proved, that the best vector-samples for the Hamming network would be 3-d order cumulants averaged on all four rotations.

Thus, algorithm of the work of the system of texture images recognition (algorithm 2) consists of following steps:

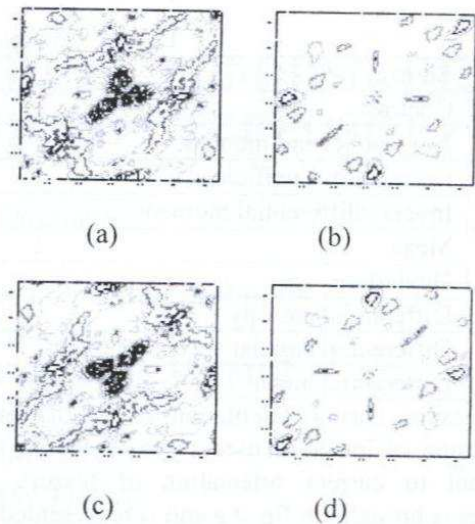


Fig. 4. Third-order cumulants for the textures of the class presented on fig. 2.a in level lines images:  
(a) orientation  $0^{\circ}$ ; (b) orientation  $90^{\circ}$ ;  
(c) orientation  $180^{\circ}$ ; (d) orientation  $270^{\circ}$ ;

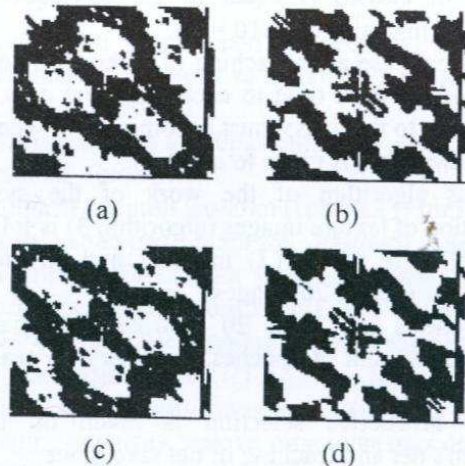


Fig. 5. Third-order cumulants for the textures of the class presented on fig. 2.a in level lines images after binary procedure:  
(a) orientation  $0^{\circ}$ ; (b) orientation  $90^{\circ}$ ;  
(c) orientation  $180^{\circ}$ ; (d) orientation  $270^{\circ}$ ;

1. For each of 20 image the cumulants of 3-rd order are calculated.
2. For each of 5 classes averaged 3-rd order cumulants of all four images of the class is calculated, and then it is binaried with threshold 0.
3. All 5 binary matrixes are processing through the Hamming network as vector-samples.
4. On input of the system of recognition unknown test image (one of the described 20 images) is given, and the system should refer this image to the initial class with minimum error of the classification.

### 4.2. Classifier, built on the base of SGLD features

In elementary phase of practical realization of given classifier is necessary to count the informing features, which in given case are 10 texture features from SGLD matrixes.



Texture features from SGLD matrices, shown on fig. 2.a

Entropy	930.2033
Contrast	3.2671e+06
Angular second moment	2.9109e+03
Correlation Coefficient	8.0294e-015
Inverse differential moment	95.8760
Mean	2.4399e+03
Deviation	1.1409e+07
Differential entropy	-1.5082e+04
Differential angular second moment	2.5254e+05
Differential mean	8.9553e+04

Because during calculation of SGLD brightness description of image is used, type of SGLD matrix is invariant to current orientation of texture. Texture features is brought on fig. 2.a and is represented on table 4.1.

For correcting classification the volume of topographic Kohonen's map exemplarily must excel amount of possible classes in twice. In our case the number of classes is equal to 10, consequently, the Kohonen's map size of  $2 \times 10 = 20$ .

In ideal case after teaching of Kohonen's net only 1 neuron of net must treat to each class and only neuron, conforming to this class must become active when image of unknown class is given to input.

Thus algorithm of the work of the system of recognition of texture images (algorithm 3) is following:

1. Is calculate the SGLD matrices and on 10 texture features for each of 20 images.
2. Kohonen's net from 20 neurons is initialed and maximum amount of epoches, necessary for teaching of net is set.
3. The instructive selection is given on input of Kohonen's net and teaching of net take place.
4. Unknown testing image is given on input of recognition system (one from 20 described above), this testing images must be refered by the system to one or another initial class with minimum mistake of classification by dint of calculation of statistic features of second order and further usage of Kohonen's net.

### 5. CONCLUSION

The accuracy classification of both classifiers is equal to 95-100 %. The operating time of the system was 10 minutes for evaluation of one matrix of features and 2-3 seconds for the classification with help of PC: P5-366 MHz/64M. Such a high speed of work of the classifier is reached by the simplicity of the principle of work of the Hamming network. For increase of processing speed of work of the extractor of features (see fig. 1) the size of the matrix of features decreased down to  $64 \times 64$ , however the accuracy of cumulants classification decreased (down to 70 %) as the speed of calculation increases.

Advantages of the method are the high accuracy of the classification and little operating time of the classifier for the class of the PC.

Disadvantages are insufficient speed of work of the system during feature extraction, and also restriction of the number of outputs of the classifier which are equal to

number of classes. In this situation, when on input of the system is exited with texture image of metal-structure from the 6-th class, the system will still refer it to one of the initial 5.

Thanks to its work principle, the problem of "6-th class" in Kohonen's classifier will be absent. Besides this device of texture classification, built on base of Kohonen's net, works quicker than Hamming's classifier. First of all this reaches at the expense increase of speed of work feature's extractor. In addition to that invariance of SGLD matrix to orientation of texture allows to count up the informing features for each class only onces, where as matrix of estimations of cumulants must be counted 4 times (for each orientation).

The main disadvantage of this algorithm is little number of informing features. The dimension of Kohonen's topographic map must be increased for correcting classification attached to augmentation of amount of classes, but that is gives considerable classifier work time augmentation.

Further researches are directed to removal of disadvantages which were brought above, that can be reached by use, for example, adaptive resonance theory network (ART2).

### REFERENCES

1. Andreev G.A., Bazarskiy O.V. "Analysis and synthesis of random spatial textures" // Foreign radioelectronics. - 1984. - №2. - P. 3 - 34
2. Dhawan A., Chitre Y., "Analysis of mammographic microcalcifications using gray-level image structure features" // IEEE Trans. Med. Imag. - 1996. - №15.
3. Wosserman F. "Neurocomputing". - Moscow: World, 1992.
4. Haralick R.M., Shanmugam K., and Dinstein I. "Textural features for image classification" // IEEE Trans. Syst., Man, Cybern. - 1973. - SMC-3. - P. 610-621.
5. Kosko B. "Neural networks for signal processing". - Prentice Hall, Englewood Cliffs, NJ, 1992.
6. Mendel J.M., Nikias C.L. "Signal processing with higher-order spectra" // IEEE Signal Processing Magazine. - 1993. - №10.
7. Terehov S.A., "Lecture under the theory and applications of artificial neural networks". - Snezhinsk: VNIIT, 1994.
8. Nikias C.L., Raguber M.L. "Bispectral estimation employing to digital signals processing" // IEEE Signal Processing Magazine. - 1987. - №7.



# ОПТИМІЗАЦІЯ КУСКОВО-ЛІНІЙНОЇ ДЕФОРМАЦІЇ ЗНІМКІВ У ЗАДАЧАХ МАШИННОГО СТЕРЕОЗОРУ

М.І.Шлезінгер<sup>1</sup>, Б. Флах<sup>2</sup>, І.В.Онищенко<sup>3</sup>

<sup>2</sup>Institute of Artificial Intelligence, Dresden University of Technology  
Email: bflach@ics.inf.tu-dresden.de

<sup>1,3</sup>Міжнародний науково-навчальний центр ЮНЕСКО/МПІ  
інформаційних технологій та систем.  
пр.Академіка Глушкова 40, Київ 03680  
т.266-62-08 Email: schles@image.kiev.ua

## АНОТАЦІЯ

Розглянуто метод знаходження оптимального монотонного кусково-лінійного відображення відрізка на відрізок. При цьому враховується як ступінь відповідності точок відрізків, так і довжина ламаної. Наведено результати застосування описаного методу при відновленні рельєфу людського обличчя за стереопарою.

## ВСТУП

Дана стаття відноситься до класу робіт з відновлення рельєфу поверхні за стереопарою, що ґрунтуються на використанні паралаксів. Цей напрям було започатковано Г.Л.Гімельфарбом. У його роботі [1] побудовано модель оптичної поверхні, що відтворюється за своїми зображеннями, та модель процесу отримання цих зображень; обґрунтовано зведення задачі стереозору до оптимізаційних задач, зокрема, до задачі динамічного програмування. У роботі [2] запропоновано деякі евристики для звуження класу шуканих поверхонь.

Розв'язування задачі стереозору полягає у знаходженні монотонного відображення скінченної дискретної впорядкованої множини на себе. Єдиним таким відображенням є тотожне відображення. При переході від дискретних множин до неперервних матимемо безліч допустимих відображень, навіть якщо шукатимемо тільки ті з них, що є неперервними та кусково-лінійними. Якщо додатково вимагати, щоб вершини ламаної мали лише цілі координати, то отримаємо скінченну кількість допустимих ламаних.

## 1 ПОСТАНОВКА ЗАДАЧІ

Для заданого натурального числа  $n$  *полем зору* будемо називати відрізок  $T=[0; n]$  дійсної осі. Відрізок  $p$ , вигляду  $[i-1; i] \subset T$ ,  $i=1, \dots, n$  назвемо *i-м пікселом*. *Знімком* будемо називати кусково-постійну

функцію  $g: T \rightarrow C$ , де  $C$  – деяка скінченна множина *кольорів*.  $g$  приймає однакові значення в усіх точках піксела.

Вихідні дані для задачі стереозору задаються двома знімками: лівим  $g_L$  та правим  $g_P$ , визначеними над одним полем зору  $T$ .

Ми шукаємо таке відображення  $f: T \rightarrow T$ , що:

- 1)  $f$  – кусково-лінійне;
- 2) координати вершин ламаної (графіка  $f$ ) цілі числа;
- 3) довжина кожного відрізка ламаної обмежена зверху і знизу; обмеження знизу дозволяє керувати "гладкістю" рельєфа, а обмеження зверху прискорює роботу алгоритму, зменшуючи перебір варіантів.

Відрізок ламаної  $e(i_1, j_1, i_2, j_2)$ ,  $i_1, j_1, i_2, j_2 \in \{0, 1, \dots, n\}$  задається координатами початкової  $(i_1, j_1)$  та кінцевої  $(i_2, j_2)$  точок, причому мають виконуватися наступні обмеження: значення  $i_2 - j_1$  та  $i_2 - j_2$  мають належати деякій множині  $H$ ; обмеження на довжини відрізків ламаної  $\dim_{\min} \leq i_2 - i_1 \leq \dim_{\max}$ ,  $d_{j\min} \leq j_2 - j_1 \leq d_{j\max}$ . Через  $E$  позначимо множину всіх допустимих відрізків. Позначимо через  $S$  множину всіх ламаних  $s$  таких, що:

- 1) кожен відрізок ламаної належить  $E$ ;
- 2) ламана починається у точці вигляду  $(0, k)$  або  $(k, 0)$  і закінчується у точці вигляду  $(n, n-k)$  або  $(n-k, n)$ , де  $0 \leq k \leq n$ .

Під *довжиною*  $l(e)$  відрізка  $e$  розумітимемо його геометричну довжину. Нехай  $\alpha(j, k): T \times T \rightarrow R$  функція, яка кількісно характеризує несхожість кольорів пікселів  $j$  лівого та  $k$  правого знімків. *Якістю*  $q(e)$  відрізка  $e(i_1, j_1, i_2, j_2)$  назвемо число де  $[x]$  означає цілу частину числа  $x$ ,  $d_i = i_2 - i_1$ ,  $d_j = j_2 - j_1$ ,

$$\frac{1}{d_i \cdot d_j} \sum_{n=0}^{d_i \cdot d_j - 1} \alpha([i_1 + n / d_i][j_1 + n / d_j]) + \beta \cdot l(e), \quad (1)$$

$\beta$  – деяка константа. Якість ламаної дорівнює сумі якостей її відрізків. Ми шукаємо таку ламану  $s^* \in S$ , що

$$s^* = \arg \min_{s \in S} \sum_{e \in s} q(e), \quad (2)$$

де сума береться по всіх дужках ламаної  $s$ .



## 2 РОЗВ'ЯЗУВАННЯ ЗАДАЧІ

Позначимо через  $F(i,j)$  якість найкращої ламаної з кінцем в точці  $(i,j)$  і початком у точці вигляду  $(0,k)$  або  $(k,0)$ , де  $0 \leq k \leq k_{\max}$ . Покладемо  $F(0,k)=F(k,0)=0$  для  $0 \leq k \leq k_{\max}$ . Розв'язком задачі буде ламана з кінцем у точці вигляду  $(n,n-k)$  або  $(n-k,n)$ , де  $0 \leq k \leq k_{\max}$ , що має мінімальне значення  $F$ . При розв'язуванні задачі використовується рекурентне співвідношення

$$F(i_2, j_2) = \min_{e(i_1, j_1, j_2) \in E} (F(i_1, j_1) + q(e)). \quad (3)$$

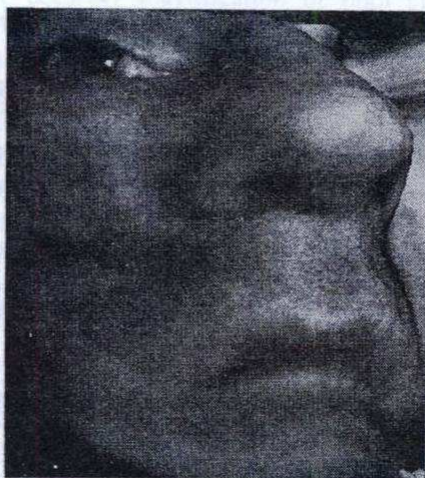
## 3 ОТРИМАНІ РЕЗУЛЬТАТИ

Описаний алгоритм було реалізовано мовою С. На мал.1 показано вихідну стереопару, а на мал.2 отримані результати для таких значень параметрів:  $N = \{-30, -29, \dots, 29, 30\}$ ,  $n=300$ ,  $\text{dimin}=\text{djmin}=3$ ,  $\beta=0.05$ ,  $k_{\max}=20$ ,  $\text{dimax}=\text{djmax}=5$ ; колір точки задається трійкою  $(r, g, b)$ , а функція  $\alpha$  така, що

$$\alpha^2((r_1, g_1, b_1), (r_2, g_2, b_2)) = (r_1 - r_2)^2 + (g_1 - g_2)^2 + (b_1 - b_2)^2$$



Мал.1 Вихідна стереопара



Мал.2 Відновлений рельєф обличчя

## ВИСНОВКИ

Розглянуто метод оптимізації кусково-лінійної функції однієї змінної, який дозволяє знаходити відповідність між точками пари рядків знімків стереопари. Але при цьому не враховується залежність між сусідніми рядками. Врахування такої залежності приводить до задачі оптимізації кусково-лінійної функції двох змінних.

## ЛІТЕРАТУРА

- [1] Гимельфарб Г.Л. "Симметричный подход к задаче автоматических стереоскопических измерений в фотограмметрии" // *Кибернетика*. — 1979. — №2. — С.73-82.
- [2] Гимельфарб Г.Л. "Симметризованное би- и тринокулярное стереозрение: взаимосвязи между теоретическими основами и эвристическими решениями" // *Теоретические и прикладные вопросы распознавания изображений*. — Киев, 1995. — С.4-25.



# КАЛІБРОВКА ЗОБРАЖЕНЬ ДЛЯ ПОРЯДКОВИХ МЕТОДІВ РОЗПІЗНАВАННЯ РЕЛЬЄФУ

Д.І. Рябоконт

Міжнародний Науково-Навчальний Центр Інформаційних Технологій та Систем  
 Проспект Академіка Глушкова 40, Київ-03680. Тел 266-62-08

## АНОТАЦІЯ

В роботі розглядаються методи перетворення стереопари зображень до такого вигляду, щоб пари рядків розгортки могли бути використані як вихідні дані для порядкових алгоритмів розпізнавання рельєфу об'єкта. Описується комплекс, що дозволяє створювати ідеальні зображення для стереоскопа та розпізнавати рельєф на цих зображеннях.

## ВСТУП

Досить легко створити стереопару для побутових потреб. Для цього треба лише фотографувати об'єкт із двох різних позицій і поглянути на пару знімків через стереоскоп. Так можна визначити, які точки знаходяться ближче, а які далі. Для машинної обробки стереопара повинна бути виконана значно ретельніше. Підвищені вимоги до стереопари зумовлені тим, що більшість методів розпізнавання рельєфу [1, 2] оперують із парами рядків розгортки на зображеннях незалежно. Весь рельєф потім складається зі шматочків, що були відновлені на кожному рядку.

Ідеальною стереопарою будемо називати два таких зображення, що для кожної точки на рядку  $j$  першого зображення відповідна точка на другому зображенні знаходиться також у рядку  $j$ . Саме для порядкових методів розпізнавання рельєфу і потрібне перетворення двох зображень до ідеальної стереопари (калібровка зображень).

## ОБРОБКА НЕКАЛІБРОВаних ЗОБРАЖЕНЬ

### 1. Епіполярні лінії

Розглянемо випадок, коли при обробці рядку  $j$  лівого та правого зображення взагалі неможливо знайти правильну відповідність між точками, бо, скажімо, на одному із зображень використовується інший масштаб. Для застосування порядкових алгоритмів необхідно спочатку знайти сімейства ліній на обох зображеннях, котрі будуть використані в процесі розпізнавання.



Рис. 1. Епіполярні лінії

Кожній точці на лівому зображенні ставиться у відповідність лінія на правому зображенні та навпаки. Епіполярна лінія на лівому зображенні – це проекція променя, що прямує від точки поверхні об'єкта до фокуса правої камери на проекційну площину лівої камери (Рис. 1). Точки на зображенні – це перетин променів із проекційною площиною камери. Оскільки промені, що утворюють точки зображення, проходять через фокус камери, всі епіполярні лінії на кожному із зображень перетинаються в одній точці (ця точка може знаходитись і на безмежності).

Епіполярні лінії на парі зображень зручно подати в однорідних координатах, де кожний елемент зображення індексується трьохвимірним вектором  $U=(i,j,1)^T$ . Сімейства епіполярних ліній можна тоді записати у вигляді

$$(U_L)^T F U_R = 0,$$

де  $F$  – матриця  $3 \times 3$ , що має назву *фундаментальна матриця*.

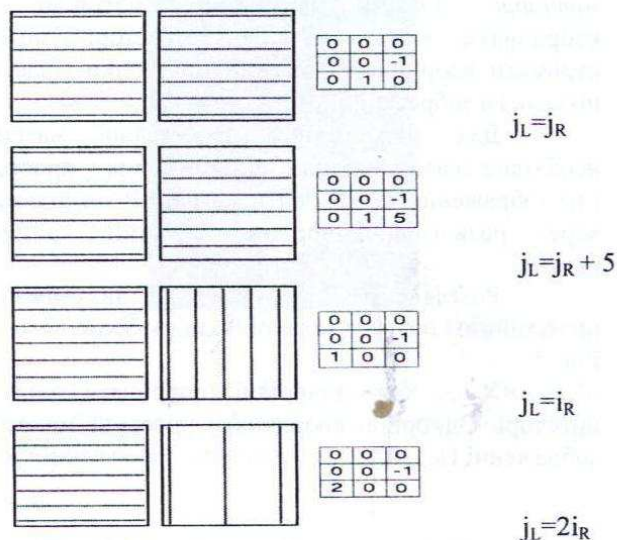


Рис. 2. Сімейства епіполярних ліній



Оскільки всі епіполлярні лінії перетинаються в одній точці, ранг матриці  $F$  повинен дорівнювати двом.

Якщо знайти епіполлярні лінії на зображеннях, перетворити їх так, щоб матриця  $F$  мала вигляд  $\{(0,0,0), (0,0,-1), (0,1,0)\}$ , то можна буде застосувати порядкові методи розпізнавання рельєфу, обробляючи на стереозображеннях рядки розгортки з однаковими індексами.

## 2. 8-точковий алгоритм

Якщо на зображеннях можна вказати відповідність між якимись точками, то фундаментальна матриця знаходиться за допомогою *8-точкового алгоритму* [3]. Оскільки для кожної пари відповідних точок виконується

$$(U_L)^T F U_R = 0,$$

то можна записати наступну рівність.

$$\begin{aligned} & i_L i_R F_{11} + i_L j_R F_{12} + i_L F_{13} + \\ & + j_L i_R F_{21} + j_L j_R F_{22} + j_L F_{23} + \\ & + i_R F_{31} + j_R F_{32} + F_{33} = 0, \\ & \text{rank}(F) = 2. \end{aligned} \quad (1)$$

Якщо маємо сукупність пар відповідних точок, то можна скласти однорідну систему рівнянь для коефіцієнтів матриці  $F$ . Для знаходження всіх коефіцієнтів матриці  $F$  необхідно 8 рівнянь, оскільки матриця знаходиться з точністю до постійного множника.

## 3. Проекційні матриці

Нехай  $X=(x,y,z,1)^T$  однорідні координати точки у просторі,  $U=(i,j,1)^T$  однорідні координати проекції точки на зображенні. Тоді  $X$  та  $U$  пов'язані відношенням

$$(i,j,1)^T = P(x,y,z,1)^T / (xP_{31} + yP_{32} + zP_{33} + P_{34}), \quad (2)$$

де  $P$  – матриця  $3 \times 4$ , що має назву *проекційна матриця*. Знаючи проекційну матрицю та координати якоїсь точки в просторі, можна отримати координати проекції цієї точки (індекси пікселя на зображенні).

Для знаходження проекційної матриці необхідно знати координати шести точок у просторі і на зображенні. Коефіцієнти матриці  $P$  знаходяться через розв'язок однорідної системи лінійних рівнянь.

Розглянемо приклад знаходження проекційних матриць для об'єкта, зображеного на Рис. 3.

$X_0 \dots X_8$  – однорідні координати точок у просторі. Однорідні координати проекцій точок на зображенні  $U_0 \dots U_8$ .

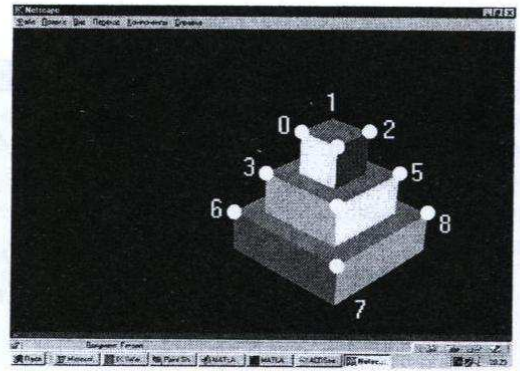


Рис. 3. Об'єкт з відомими просторовими координатами

$$\begin{aligned} U_0 &= (461, 207, 1) & X_0 &= (-1, 1, 6, 1), \\ U_1 &= (520, 233, 1) & X_1 &= (-1, -1, 6, 1), \\ U_2 &= (569, 209, 1) & X_2 &= (1, -1, 6, 1), \\ U_3 &= (406, 276, 1) & X_3 &= (-2, 2, 4, 1), \\ U_4 &= (521, 331, 1) & X_4 &= (-2, -2, 4, 1), \\ U_5 &= (618, 276, 1) & X_5 &= (2, -2, 4, 1), \\ U_6 &= (355, 340, 1) & X_6 &= (-3, 3, 2, 1), \\ U_7 &= (521, 430, 1) & X_7 &= (-3, -3, 2, 1), \\ U_8 &= (664, 342, 1) & X_8 &= (3, -3, 2, 1). \end{aligned}$$

Для точки 0 рівняння (2) дає наступні два рівняння:

$$\begin{aligned} 461(-P_{31} + P_{32} + 6P_{33} + P_{34}) &= (-P_{11} + P_{12} + 6P_{13} + P_{14}), \\ 207(-P_{31} + P_{32} + 6P_{33} + P_{34}) &= (-P_{21} + P_{22} + 6P_{23} + P_{24}). \end{aligned}$$

Для точки 1:

$$\begin{aligned} 520(-P_{31} - P_{32} + 6P_{33} + P_{34}) &= (-P_{11} - P_{12} + 6P_{13} + P_{14}), \\ 233(-P_{31} - P_{32} + 6P_{33} + P_{34}) &= (-P_{21} - P_{22} + 6P_{23} + P_{24}). \end{aligned}$$

І, нарешті, для точки 8:

$$\begin{aligned} 664(3P_{31} - 3P_{32} + 2P_{33} + P_{34}) &= (3P_{11} - 3P_{12} + 2P_{13} + P_{14}), \\ 342(3P_{31} - 3P_{32} + 2P_{33} + P_{34}) &= (3P_{21} - 3P_{22} + 2P_{23} + P_{24}). \end{aligned}$$

Матриця системи однорідних лінійних рівнянь для 12 невідомих  $P_{11}, P_{12}, P_{13}, P_{21}, \dots, P_{34}$  буде:

$$A = \begin{pmatrix} -1 & 1 & 6 & 1 & 0 & 0 & 0 & 0 & 461 & -461 & -2766 & -461 \\ 0 & 0 & 0 & 0 & -1 & 1 & 6 & 1 & 207 & -207 & -1242 & -207 \\ -1 & -1 & 6 & 1 & 0 & 0 & 0 & 0 & 520 & 520 & -3120 & -520 \\ 0 & 0 & 0 & 0 & -1 & -1 & 6 & 1 & 233 & 233 & -1398 & -233 \\ 1 & -1 & 6 & 1 & 0 & 0 & 0 & 0 & -569 & 569 & -3414 & -569 \\ 0 & 0 & 0 & 0 & 1 & -1 & 6 & 1 & -209 & 209 & -1254 & -209 \\ -2 & 2 & 4 & 1 & 0 & 0 & 0 & 0 & 812 & -812 & -1624 & -406 \\ 0 & 0 & 0 & 0 & -2 & 2 & 4 & 1 & 552 & -552 & -1104 & -276 \\ -2 & -2 & 4 & 1 & 0 & 0 & 0 & 0 & 1042 & 1042 & -2084 & -521 \\ 0 & 0 & 0 & 0 & -2 & -2 & 4 & 1 & 662 & 662 & -1324 & -331 \\ 2 & -2 & 4 & 1 & 0 & 0 & 0 & 0 & -1236 & 1236 & -2472 & -618 \\ 0 & 0 & 0 & 0 & 2 & -2 & 4 & 1 & -552 & 552 & -1104 & -276 \\ -3 & 3 & 2 & 1 & 0 & 0 & 0 & 0 & 1065 & -1065 & -710 & -355 \\ 0 & 0 & 0 & 0 & -3 & 3 & 2 & 1 & 1020 & -1020 & -680 & -340 \\ -3 & -3 & 2 & 1 & 0 & 0 & 0 & 0 & 1563 & 1563 & -1042 & -521 \\ 0 & 0 & 0 & 0 & -3 & -3 & 2 & 1 & 1290 & 1290 & -860 & -430 \\ 3 & -3 & 2 & 1 & 0 & 0 & 0 & 0 & -1992 & 1992 & -1328 & -664 \\ 0 & 0 & 0 & 0 & 3 & -3 & 2 & 1 & -1026 & 1026 & -684 & -342. \end{pmatrix}$$



Для існування розв'язку такої системи однорідних лінійних рівнянь, ранг матриці  $A$  повинен бути нижчим за 12. Якщо це не так, то його необхідно понизити. Для цього в сингулярному розкладі матриці  $A$  найменшому власному значенню приписується 0.

#### 4. Знаходження епіполярних ліній за проекційними матрицями

Якщо  $P_L$  – проекційна матриця для першої камери,  $P_R$  – проекційна матриця для другої камери, то сімейства епіполярних ліній можна записати у вигляді

$$(U_L)^T F U_R = X^T (P_L)^T F P_R X = 0, \quad (3)$$

причому добуток  $(P_L)^T F P_R$  має бути антисиметричною матрицею.

При знаходженні  $F$  за проекційними матрицями  $P_L, P_R$  можна перетворити їх так, що

$$P_L := P_L H \quad P_R := P_R H, \quad (4)$$

де матрицю  $H$  необхідно вибрати таким чином, щоб  $P_L$  стала

$$P_L = \{(1 \ 0 \ 0 \ 0), (0 \ 1 \ 0 \ 0), (0 \ 0 \ 1 \ 0), (0 \ 0 \ 0 \ 0)\}.$$

Використовуючи (3) та (4), маємо:

$$(U_L)^T F U_R = X^T (P_L H)^T F P_R H X = \\ = X^T H^T (P_L)^T F P_R H X,$$

внаслідок чого

$$F = \{p_1, p_2, p_3, p_4\}^T \{[p_4, p_1], [p_4, p_2], [p_4, p_3], [p_4, p_4]\}^T$$

де  $p_i$  – вектори-стовпці матриці  $P_R$ .

Можна переконатися, що добуток  $P_L F P_R$  буде антисиметричною матрицею [4].

Знаючи сімейства епіполярних ліній, нескладно перетворити зображення так, щоб фундаментальна матриця мала вигляд

$$F = \{(0,0,0), (0,0,-1), (0,1,0)\}.$$

#### 5. Результати обробки стереозображень

Для двох закріплених одна відносно одної фотокамер досить один раз розрахувати проекційні матриці. Всі знімки, що будуть надалі зроблені такою системою, калібруватимуться одним і тим самим перетворенням. Калібровані зображення будуть придатними для обробки порядковими методами розпізнавання рельєфу.

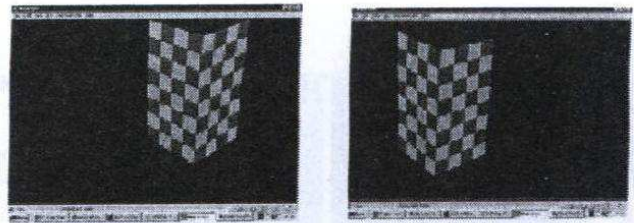


Рис. 4. Об'єкт для калібрування

На Рис. 4 зображена стереопара об'єкта з відомими просторовими координатами. На основі цих зображень знаходяться проекційні матриці та епіполярні лінії для всіх знімків, що будуть зроблені цими ж камерами.

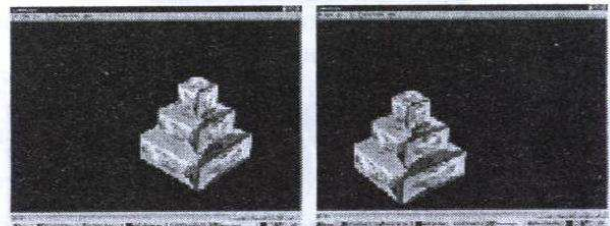


Рис. 5. Об'єкт для розпізнавання

На Рис. 5 зображення стереопара об'єкта, просторову конфігурацію якого необхідно розпізнати. Рядки розгортки на цій стереопарі не відповідають епіполярним лініям. Зображення слід калібрувати перед розпізнаванням.

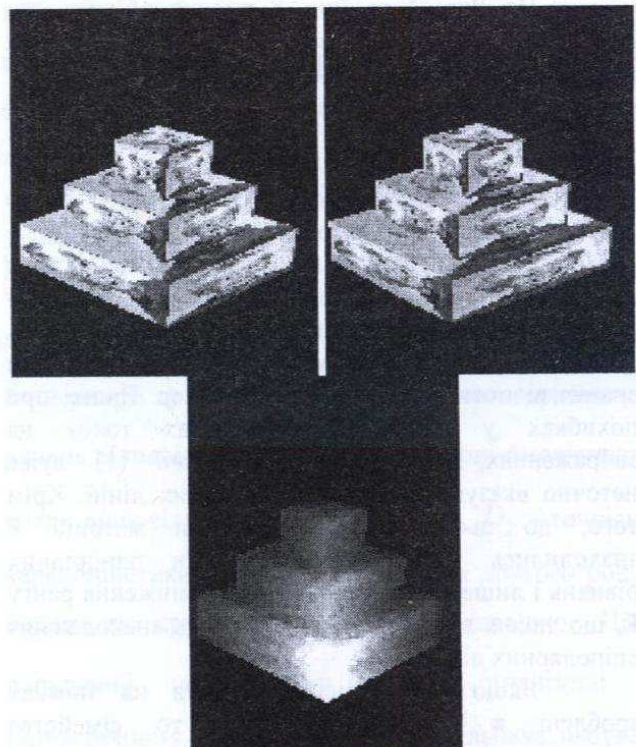


Рис. 6. Калібровані знімки та висотне зображення

На Рис. 6 подані калібровані стереозображення об'єкта та його відновлена просторова конфігурація.



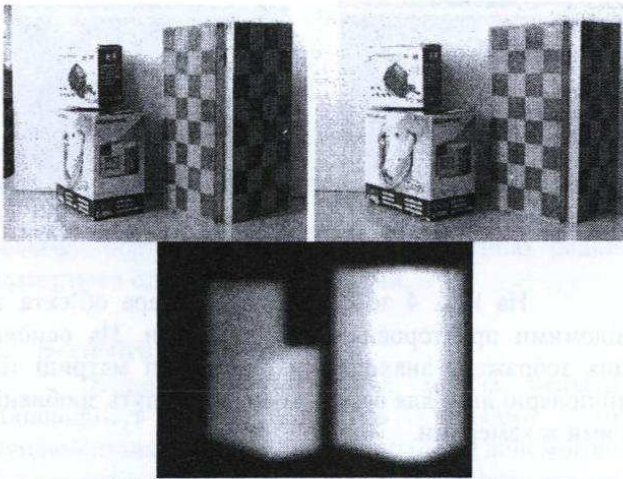


Рис. 7. Об'єкти, сфотографовані цифровою камерою, та відновлений рельєф

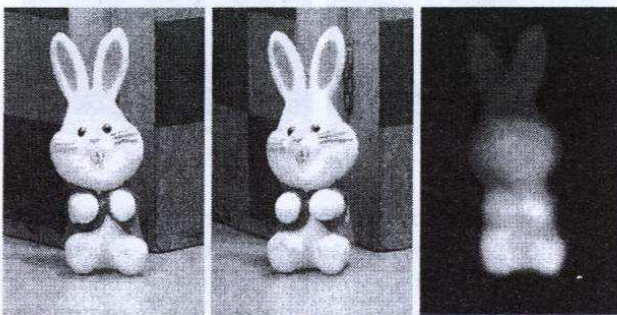


Рис. 8. Об'єкт та його відновлений рельєф

На Рис. 7 та Рис. 8 подано об'єкти, що сфотографовані цифровою камерою, та їх відновлена просторова конфігурація.

## ВИСНОВКИ

Для калібровки зображень було запропоновано два методи. Суть першого – у знаходженні епіполярних ліній за вказаними парами відповідних точок на зображеннях, другого – у знаходженні проєкційних матриць.

Перший метод можна застосовувати без знання відносного розташування камер. Проте, при похибках у вказаних координатах точок на зображеннях, псевдорозв'язок задачі (1) дуже неточно вказує сімейство епіполярних ліній. Крім того, до цього часу коефіцієнти матриці  $F$  знаходились як розв'язок системи однорідних рівнянь і лише потім проводилось пониження рангу  $F$ , що також зумовлювало неточність знаходження епіполярних ліній [4].

Якщо всі елементи об'єкта на знімках зроблені в одному масштабі, то сімейства епіполярних ліній можна було б шукати у вигляді паралельних прямих на кожному із знімків. Задача (1) в цьому випадку значно спрощується. Матриця  $F$  буде автоматично мати ранг 2 та шукатися у вигляді

$$F = \begin{pmatrix} 0 & 0 & F_{13} \\ 0 & 0 & F_{23} \\ F_{31} & F_{32} & F_{33} \end{pmatrix}$$

Матрицю  $F$ , що отримали таким чином можна використовувати як перше наближення і для загального випадку калібровки зображень. Елементи матриці  $F$  слід знаходити ітеративно, на кожному кроці уточнюючи якийсь стовпець чи рядок матриці при фіксованих інших елементах. Обмеження на ранг в такому випадку буде лінійне по елементам  $F_{ij}$  рівняння. Таким чином, на кожному кроці ітерації фундаментальна матриця буде мати ранг 2, а точність знаходження її елементів не погіршуватись.

Калібровка другим способом дозволяє створювати ідеальні стереопари двома фотоапаратами. Для двох камер слід зафіксувати взаємне положення і один раз провести вимір проєкційних матриць, сфотографував об'єкт з відомими просторовими координатами. Всі знімки, що будуть отримані надалі, можна буде перетворювати в ідеальні стереопари.

Автором було створено програмно-апаратний комплекс, що дозволяє розпізнавати рельєф на зображеннях з використанням попередньої калібровки. Знімки робились цифровою камерою, що могла переміщатися по спеціальному штативу. Калібровані зображення оброблялись порядковим алгоритмом розпізнавання рельєфу.

## ЛІТЕРАТУРА

- [1] Г.Л. Гимельфарб. "Симметризованное би- и тринокулярное стереозрение: взаимосвязь между теоретическими основами и эвристическими решениями." Теоретические и прикладные вопросы распознавания изображений. Киев 1995.
- [2] R. Sara. "The Class of Stable Matching for Computation Stereo."
- [3] R. Hartley. "In defence of the 8-point algorithm." In International Conference on Computer Vision, pages 1064–1070, 1995.
- [4] T. Werner. "Practice of 3D Reconstruction from Multiple Uncalibrated Unorganized Images. Czech Pattern Recognition Workshop 2000."
- [5] Г.Л. Гимельфарб. "Симметричный подход к задаче автоматических стереоскопических измерений в фотометрии." Кибернетика 1979.



# АВТОФОКУСУВАННЯ ЕЛЕКТРОННОГО МІКРОСКОПУ ЗА ЗОБРАЖЕННЯМ

*А.С. Краснополюсовський, А.М. Скаковська*

Сумський державний університет  
вул. Римського-Корсакова, 2, м.Суми, 40007  
33-03-90, факс (0542) 33-40-58  
E-mail: pm@ssu.sumy.ua

Розглядається застосування методу функціонально-статистичних випробувань (МФСВ), оснований на оцінці інформаційної здатності системи розпізнавання образів, для класифікаційної настройки (КН) складного об'єкту контролю та управління.

## ВСТУП

Важливою задачею в електронній мікроскопії є автофокусування растрового електронного мікроскопу (РЕМ). Це особливо актуально, наприклад, при автоматичній класифікації об'єктів дослідження, яка вимагає додержання незмінних і однакових умов для прийняття високостовірних рішень. До цього часу невідомі ефективні алгоритми розв'язання цієї задачі. Розглянемо підхід до розв'язання задачі автофокусування РЕМ за зображенням в рамках МФСВ [1,2].

## 1 ТЕОРЕТИЧНІ ТА МЕТОДОЛОГІЧНІ ПОЛОЖЕННЯ ПІДХОДУ

Фокусність зображення досліджуваного зразка залежить від багатьох факторів, які впливають на функціонування електронного мікроскопу. Для РЕМ в якості управляючого контрольованого параметру виступає струм об'єктивної лінзи  $I_L$ . Ідея автофокусування РЕМ полягає в обчисленні інформаційної міри початкового (базового) розфокусованого зображення  $X_0$  відносно

поточного зображення  $X_S$ , яке встановлюється на  $S$ -ому кроці КН, порівнянні її значення з попереднім і прийнятті рішення про зміну управляючого параметру на задану величину. Процес КН продовжується до тих пір, поки значення струму  $I_L$  не забезпечить прийнятної з практичної точки зору фокусності зображення. В якості критерію функціональної ефективності (КФЕ) КН розглянемо інформаційну міру Кульбака, яка є добутком відношення правдоподібності  $\Lambda$  на міру відхилень розподілів ймовірностей. В загальному випадку відношення правдоподібності подамо у вигляді

$$\Lambda = \log_2 \frac{P_i}{P_f} = \frac{p_1 D_1 + p_2 D_2}{p_1 \alpha + p_2 \beta}, \quad (1)$$

де  $P_i = p_1 D_1 + p_2 D_2$ ,  $P_f = p_1 \alpha + p_2 \beta$  - повні ймовірності правильного і неправильного прийняття рішень відповідно;  $p_1, p_2$  - апіорні ймовірності появи відповідних подій;  $\alpha, \beta, D_1, D_2$  - точнісні характеристики: помилки першого та другого роду, перша та друга достовірності відповідно. При допущенні, що  $p_1 = p_2 = 0.5$  (найгірші в статистичному сенсі умови) міра Кульбака набуває вигляду

$$E = 0,5 \log_2 \left( \frac{D_1 + D_2}{\alpha + \beta} \right) [(D_1 + D_2) - (\alpha + \beta)] = \log_2 \left( \frac{2 - (\alpha + \beta)}{\alpha + \beta} \right) [1 - (\alpha + \beta)]. \quad (2)$$



При введенні оцінок  $\alpha = k_1/n^*$ ,  $\beta = k_2/n^*$ , де  $k_1, k_2$  - число реалізацій класів  $X_0$  і  $X_S$  відповідно, які не знаходяться в області своїх класів;  $n^*$  - кількість спостережень, які забезпечують репрезентативність вибіркової послідовності. Тоді формула (2) набуває робочого вигляду

$$E = \frac{1}{n} \log_2 \left\{ \frac{2n + 10^{-r} - [k_1 + k_2]}{[k_1 + k_2] + 10^{-r}} \right\} * [n - (k_1 + k_2)], \quad (3)$$

де  $r$  - число цифр у мантисі значення критерію  $E$ , яке вводиться на випадок появи нульових емпіричних частот. В якості розділяючої гіперповерхні (РГП) у першому наближенні береться гіперсфера, центром якої є еталонний вектор (ЕВ)  $x_0 \in X_0$ .

## 2 ОБРОБКА ЗОБРАЖЕННЯ

Оскільки сучасні ЕОМ дозволяють розрізняти 256 градаций чорно-білого зображення, то його обробку доцільно здійснювати шляхом сканування рецепторного поля, що є природним для РЕМ, і побудови розподілу яскравості для кожного рядка рецепторів. Квантовані в часі криві розподілів яскравості розглядаються як випадкові реалізації зображення  $\{g_{0,n}^{(j)}\}$ , де  $n = \overline{1, n^*}$ ;  $j = \overline{1, J_{\max}}$  - змінна числа рецепторів у рядку.

З метою формування бітового масиву значень реалізацій  $\{x_{0,n}^{(j)}\}$  здійснюється їх квантування на два рівні за формулою

$$x_n^j = \begin{cases} 1, & \text{if } A_H \leq g_n^{(j)} \leq A_B; \\ 0, & \text{if } \text{else} \end{cases}, \quad (4)$$

де  $A_H, A_B$  - нижній і верхній контрольний допуски на рецептори. Оптимізація контрольних допусків в інформаційному сенсі здійснюється безпосередньо в рамках алгоритму КН. За початкові значення  $A_H$  і  $A_B$  приймаються відповідно

$A_H = A_0 - \Delta$  і  $A_B = A_0 + \Delta$ , де  $A_0$  - номінальне (середнє) значення яскравості зображення. Оскільки двобічне поле контрольних допусків вибрано симетричним, то початкове значення  $\Delta$  дорівнює чверті шкали яскравості, тобто 64. Формування бінарного ЕВ базового класу  $X_0$  здійснюється за формулою

$$x_0^{(j)} = \begin{cases} 1, & \text{if } \sum_{n=1}^{n^*} x_{0,n}^{(j)} \geq n^*/2; \\ 0, & \text{if } \text{else}. \end{cases} \quad (5)$$

Аналогічно обробляється поточне зображення  $X_S$ . Отже, в процесі обробки зображень  $X_0$  і  $X_S$  здійснюється формування вхідних даних, які за МФСВ необхідні для обчислення інформаційної міри між базовим і поточним зображеннями. Такими вхідними даними є ЕВ  $x_0, x_s$  і відповідно випадкові реалізації  $\{x_{0,n}^{(j)}\}, \{x_{s,n}^{(j)}\}$ .

Визначення оптимальної в інформаційному сенсі кодової відстані  $d_0^*$  - радіуса РГП, здійснюється за послідовним алгоритмом

$$d_0(k) = d_0(k-1) + \begin{cases} h, & \text{if } E[d(k-1)] < E[d(k)]; \\ 0, & \text{if } E[d(k-1)] \geq E[d(k)], \end{cases}$$

де  $k$  - число прирощень радіуса;  $h$  - крок прирощення. Обчислення КФЕ закінчується при знаходженні екстремального значення  $E^* = \max_{\{d\}} E$ ,

де  $\{d\} = \{1, 2, \dots, d(x_0 \oplus x_s)\}$  - множина радіусів РГП класу  $X_0$ ;  $d(x_0 \oplus x_s)$  - міжцентрова кодова відстань.

Оптимізація контрольних допусків здійснюється після вибору  $d_0^*$  для базового класу  $X_0$ , за алгоритмом

$$\{\delta_{k,i}\} = \{\delta_{k,i}^*\}, \quad \text{якщо} \\ E^*[\{\delta_{k,i}^*\}, d_0^*] = \max_{\{\delta_{e,i}\}} E^*[d_0^*], \quad \text{де } \{\delta_{e,i}\} - \text{система}$$

нормованих полів допусків на рецептори.



### 3 ОПИС АЛГОРИТМУ

Означимо три класи функціонального стану РЕМ:  $X_H^0$ , який відповідає прийнятній фокусності, тобто оцінці "НОРМА";  $X_M^0$ , який відповідає значенню параметру  $I_L$  "МЕНШЕ НОРМИ";  $X_B^0$ , який відповідає значенню параметру  $I_L$  "БІЛЬШЕ НОРМИ". Зрозуміло, що класи  $X_M^0$  і  $X_B^0$  мають розфокусовані зображення різного ступеню. Умовою запуску алгоритму є вихід РЕМ зі стану, що відповідає класу  $X_H^0$ . За базовий клас  $X_0$  приймемо стан РЕМ на момент запуску алгоритму КН. На рис.1 наведено структурну схему алгоритму автофокусування РЕМ в рамках МФСВ.



Рисунок 1.

Формування класу  $X_0$  відбувається таким чином:

1 Зчитується масив значень яскравості зображення базового класу  $X_0$ .

2 Здійснюється вибір початкових допусків  $A_H$  і  $A_B$ .

3 Формується бінарний масив  $BM0[N, J]$  за процедурою (4).

4 Формується ЕВ  $EVO[J]$  шляхом статистичного усереднення масиву  $BM0[N, J]$  за процедурою (5).

Після визначення напрямку зміни струму  $I_L$  шляхом встановлення приналежності класу  $X_0$  до одного із класів  $X_M^0$  або  $X_B^0$  обирається стратегія зміни  $I_L$  на кожному кроці КН.

Поточне зображення  $X_S$  формується аналогічно  $X_0$  при незмінних контрольних допусках. КФЕ обчислюється за формулою (3) в рамках наведених процедур оптимізації параметрів навчання [2].

### 4 РЕАЛІЗАЦІЯ АЛГОРИТМУ

Реалізація розглянутого алгоритму здійснювалася для автофокусування серійного мікроскопу РЕМ-103 виробництва АТ СЕЛМІ (м.Суми, Україна), для якого сталий струм обмотки об'єктивної лінзи змінюється в діапазоні 0.3-2.0 А. При цьому значення струму  $I_N$ , яке відповідає класу  $X_H^0$ , знаходиться в інтервалі 1.08-1.32 А і залежить від багатьох внутрішніх і зовнішніх випадкових факторів функціонування РЕМ. Прирощення струму  $I_L$  на кожному кроці КН здійснювалося шляхом поділу навпіл різниці між розрахованим (номінальним) і поточним його значеннями. На рис.2 наведено зображення досліджуваного зразка: базове розфокусоване зображення (рис. 2,а), зображення після першого кроку настройки (рис. 2,б) і екстремальне сфокусоване зображення (рис. 3,в).



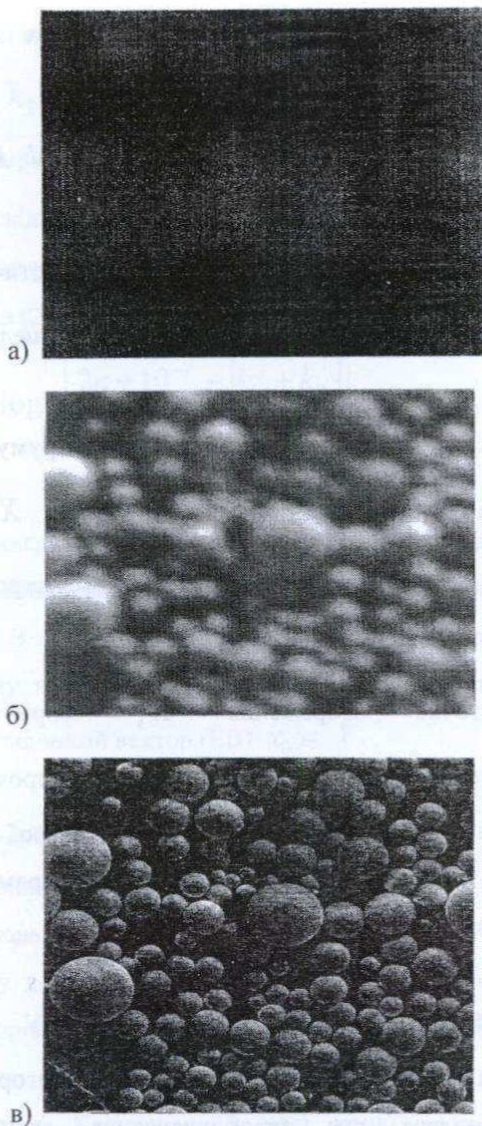
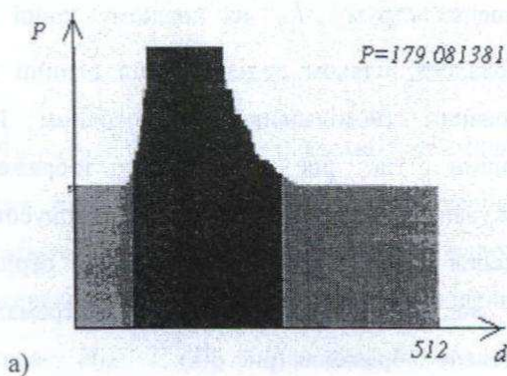
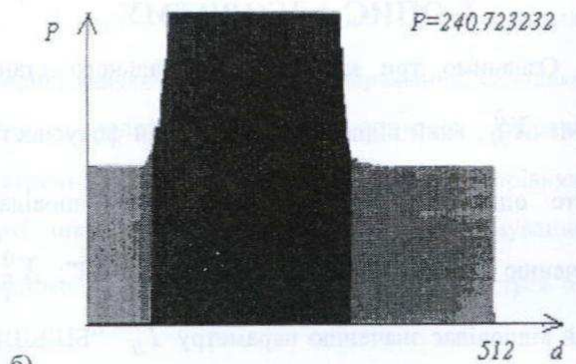


Рисунок 2.

На рис.3,а наведено значення нормованого критерію (3), при  $S=1$  і на рис. 3,б - для екстремального кроку настройки. Тут  $P$  - площа робочої області значень КФЕ, яка обчислюється на кожному кроці КН за умови, що  $D_1, D_2 \geq 0.5$ , і на рис.3 визначається затемненою частиною графіку.



а)



б)

Рисунок 3.

Тривалість автофокусування для розглянутого базового зображення становила 5с, що обумовлено інерційністю управляючої обмотки об'єктивної лінзи.

## ВИСНОВКИ

- 1 Автонастройка складної системи розглядається як ітераційна процедура пошуку максимуму інформаційної міри різноманітності базового початкового образу настройки і образами, що формуються на наступних кроках настройки.
- 2 Обчислення КФЕ на поточному кроці КН здійснюється в рамках алгоритму навчання за МФСВ.
- 3 Автофокусування РЕМ за розглянутим алгоритмом КН є цілеспрямованою ітераційною процедурою, яка здійснюється в межах прийнятної з практичних міркувань оперативності при довільних початкових умовах.

## ЛІТЕРАТУРА

1. Краснополюсовський А.С. Технічна діагностика складних систем за методом функціонально-статистичних випробувань//Обробка сигналів та розпізнавання образів: Праці третьої Всеукраїнської міжнародної конференції, 1996.- С.174-177.
2. Краснополюсовський А.С., Черниш А.В. Оцінка функціональної ефективності навчання системи розпізнавання. Вісник СумДУ, 1997.-№2.- С.112-118.