

On a New Approach to the Automated Detection of Thoracic Organs Diseases Using the Spot Feature in the Analysis of Digital X- Ray Images

S. B. Belhaouair and R. F. Kuleev

Innopolis University, Kazan, Russia

Copyright © 2014 S. B. Belhaouair and R. F. Kuleev. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

This paper presents a new approach to the automated detection of pathologies using chest X-ray imaging. The approach is based on the use of the spot feature. Various methods of interest zones selection and segmentation are explored in the paper, the algorithms of their implementation are proposed. Experimental observations have suggested that preliminary image processing is necessary to enhance the method application effectiveness.

Keywords: automated diagnostics, pulmonary tuberculosis, spot feature, zones of interest selection, segmentation

1. Introduction

Medical diagnostics is the field that relies heavily on information provided by the medical imaging. The most common imaging techniques are X-ray, CT, MRI, and ultrasound. Diagnostic accuracy is an important component of medical science. It is essential to ensuring the most effective and timely treatment, and often plays an important part in determining patient's fate. However, since visual analysis of medical images is hugely affected by the "human factor", an automated diagnostics

is necessary. This requires the development of models, methods, algorithms and relevant software. Pulmonary tuberculosis still remains a major global public health hazard. According to the WHO's (World Health Organization) estimates, there were 8.9 million incident cases of tuberculosis, 1.4 million deaths from tuberculosis in 2013. One of the basic diagnostic techniques for pulmonary tuberculosis is radiography. Diagnostic effectiveness can be substantially increased by the medical evaluation automation to diagnose tuberculosis. This paper presents a new approach to the development of methods of automated pathology detection in medical images. The proposed approach can be used to detect pulmonary lesions of tuberculosis.

In accordance with computer vision terminology, the pathology images are referred to as 'objects' in this paper. The proposed method is based on the selection of objects with the spot feature (fig.1) [1]. The algorithms are based on the assumption that an average brightness of an object and its background are different. The segmentation procedure is used to determine the projection of objects. It is proposed that segmentation is applied only on separate regions, which can contain only one object and its surrounding (background). These regions are often called the zones of interest (candidate regions). Thus, the segmentation stage is preceded by the interest zone selection in the proposed method.

Based on the objects projections identified by segmentation of the zones of interest, it is necessary to recognize and classify the descriptors to detect pathology.

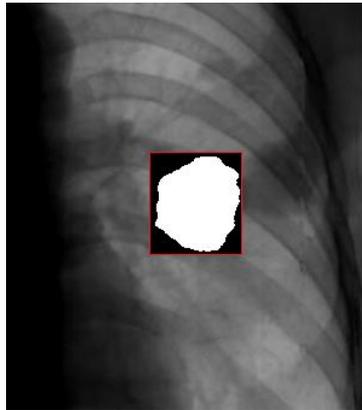


Fig.1. X-ray image spot method segmentation result

2. Mathematical Model

Locally homogenous random fields, in which each object ξ_A is a fragment of a locally homogenous random field with an average value m_A and covariance function K_A , serve as a mathematical model. The application of locally homogenous

random fields in imaging simulation as well as the objects recognition methods using this model are analyzed in works [2] - [7]. It is assumed that for these fields the conditions of Slutsky's ergodic theorem are satisfied. According to the theorem, the arithmetical mean value calculated based on the object's image ξ_A , is an estimate of its average value m_A .

Random variable ξ_z , defined on (Ω, A, P) , that possesses the value from Y , is called pixel and described by coordinates $\mathbf{z} \in Z^2$. A family $(\xi_z)_{z \in Z^2}$ is called a scene. Let us assume that $\omega \in \Omega$, then the implementation $x_z = \xi_z(\omega)$, $\mathbf{z} \in Z^2$ of the random variable ξ_z is called an image of the pixel ξ_z , subsequently, the family $\mathbf{x} = (x_z)$, $\mathbf{z} \in Z^2$ is called an image of the scene .

A set of random variables $\xi_A = (\xi_a)_{a \in A}$ is called the projection of an object, a family $x_A = (x_a)_{a \in A}$ is the image of the object. The square Q on Z^2 the subset of $Fr(Q)$ of points from Q is called the boundary of the square, if each point of $\mathbf{z} \in Fr(Q)$ has at least one neighbor of $Z^2 \setminus Q$. The object ξ_A is considered a bright (or a dark, respectively) spot, if three conditions are met: 1) $\exists Q : A \subset Q \setminus Fr(Q)$, 2) $E\xi_z = m_{Q \setminus A}$, $\mathbf{z} \in Q \setminus A$, and 3) $m_A > m_{Q \setminus A}$ ($m_A < m_{Q \setminus A}$, respectively). The pixels $\xi_{Q \setminus A} = (\xi_z)$, $\mathbf{z} \in Q \setminus A$ are called the neighborhood of the object or the background, and the family $\xi_Q = (\xi_z)$, $\mathbf{z} \in Q$ of random variables is called a zone of interest of the object.

3. Searching for the Zones of Interest

Assume that we know the diameter $d(A)$ of the object ξ_A , the side of the zone of interest l . Let us construct the family of squares $Q(l, \Delta)$ with vertices determined by the coordinates $z = z_0 + i\Delta e_1 + j\Delta e_2$, $i \in Z, j \in Z, e_1 = (1,0), e_2 = (0,1), 1 \leq l \leq l - d(A) - 1$. It is proven [5] that there is $Q(l, \Delta)$ that contains the zone of interest for the object ξ_A .

For the purposes of classification of the set of squares $Q(l, \Delta)$ into the zones of interest and the empty squares it is proposed to use the following algorithm.

Initial data for the algorithm are a rectangular fragment of an image x , an image correlation radius r' , a side of the zone of interest l , a shift Δ , a smoothing radius r , the number s of the fragments of an image, and the sign of the contrast of an object. The center of the square Q with the vertex $\mathbf{z} = (z_1, z_2)$ and the side l is at the point determined by the coordinates $(\lfloor z_1 + l/2 \rfloor, \lfloor z_2 + l/2 \rfloor)$. The algorithm sorts through squares Q from $Q(l, \Delta)$. For the classification of a square Q with the center at \mathbf{a} , the following operations are performed:

- 1) calculating, for the square $B(\mathbf{a}, r)$ with the side $2r + 1$ and the center in \mathbf{a}

$$\bar{x}_a = \frac{1}{|B(\mathbf{a}, r)|} \sum_{z \in B(\mathbf{a}, r)} x_z;$$

2) selecting fragments $Fr_j, 1 \leq j \leq s$, square images with the side $2r + 1$, that adjoins the zone of interest Q from the outer side, and is located not closer than $2r' + 1$ from each other;

3) calculating

$$\bar{x}_{Fr_j} = \frac{1}{|Fr_j|} \sum_{z \in Fr_j} x_z, 1 \leq j \leq s;$$

4) if $\bar{x}_a > \bar{x}_{Fr_j}$ for the bright object ($\bar{x}_a < \bar{x}_{Fr_j}$ for the dark object), $1 \leq j \leq s$, then square Q is considered the zone of interest, otherwise it is a square without an object.

It is proven [5] that the decision rule in p.4 tends if $n \rightarrow \infty$ and $s \rightarrow \infty$ to the Bayes decision rule.

To estimate the effectiveness of the proposed algorithm of searching for the zones of interest, a series of experiments was carried out based on the posterior-anterior chest X-ray images. The results have shown that the zones of interest were indicated for all the pathologies seen in the images. However, the majority of selected squares does not contain objects (pathologies). To enhance the effectiveness of the method of searching for the zones of interest preliminary image processing is necessary. This includes automatic ribs and bones suppression methods. In addition, it is necessary to automate the process of lungs outline detection in order to search for the zones of interest based only on the projection of lungs.

4. Segmentation

4.1. Segmentation by the Spot Feature Method

Segmentation by the spot feature method is one of the segmentation methods of the zones of interest. Let us examine the algorithm of implementation of the method presented in work [1]. The input parameters are a smoothing radius, a quantity s of boundary fragments, and an image correlation radius r' .

- 1) similarly to p.2 of the square classification algorithm used in searching for the zones of interest, the boundary fragments $Fr_j, 1 \leq j \leq s$ are selected ;
- 2) mean values $\bar{x}_{Fr_j} = \frac{1}{|Fr_j|} \sum_{z \in Fr_j} x_z, 1 \leq j \leq s$ are calculated to evaluate mean value of the background $m_{Q \setminus A}$;
- 3) for each pixel $x_z, z \in Q$, a mean value $\bar{x}_z = \frac{1}{|B(z,r) \cap Q|} \sum_{t \in B(z,r) \cap Q} x_t$ is calculated with the center at z and a radius r ;

- 4) if $x_z > \bar{x}_{Fr_j}$ for a lighter object($x_z < \bar{x}_{Fr_j}$ for the dark object, respectively), $1 \leq j \leq s$, then the x_z pixel is part of the object, otherwise, it is part of the background;
- 5) all pixels of indicated objects, located at a distance not exceeding r of scale units from the background, are part of the background.

4. 2. Segmentation by the Mode Method

Initial parameters of the algorithm used for the mode method segmentation [8] are a smoothing radius r , and a width h of the interest zone outline. Algorithm includes the following steps:

- 1) constructing a smoothed with radius r image \bar{x}_Q of the scene in the zone of interest Q with pixel values in the form of $\bar{x}_z = \frac{1}{|B(z,r) \cap Q|} \sum_{t \in B(z,r) \cap Q} x_t, z \in Q;$
- 2) constructing range expansion Fr_h by including pixels that are not a part of the interest zone, but are located at a distance not exceeding h scale units from the zone;
- 3) calculating arithmetic mean $\bar{x}_{Fr} = \frac{1}{|Fr(Q)|} \sum_{t \in Fr(Q)} x_t$ based on range expansion Fr_h with the width h ;
- 4) calculating for each pixel of an image \bar{x}_Q the distance $\lambda_z = |\bar{x}_z - \bar{x}_{Fr}|, z \in Q$ from the average \bar{x}_z based on the boundary of each pixel to the average \bar{x}_{Fr} based on the boundary and their multiplicity $k_i, 0 \leq i \leq |I|$, where $|I| + 1$ is the quantity of image brightness gradations ;
The distance multiplicity $k_i, 0 \leq i \leq |I|$ can be examined in the form of histogram that reflects the dependence of distance multiplicity from its value.
- 5) establishing initial smoothing radius of the histogram $p = 1$;
- 6) distance multiplicity histogram is smoothed with the radius p :

$$\bar{k}_i = \frac{1}{|[0, |I| - 1] \cap [i - p, i + p]|} \sum_{j \in [0, |I| - 1] \cap [i - p, i + p]} k_j ;$$

- 7) determining sets of the indices \bar{M} , and maximums \bar{k}_i

$$\begin{aligned} \bar{M} = & \{i \in [1, |I| - 1]: \bar{k}_{i-q-1} < \bar{k}_{i-q} = \dots = \bar{k}_i = \dots = \bar{k}_{i+q} > \bar{k}_{i+q+1}\} \\ & \cup \{0: \bar{k}_0 = \bar{k}_1 = \dots = \bar{k}_i > \bar{k}_{i+1}\} \\ & \cup \{i \in [1, |I|]: \bar{k}_{i-1} < \bar{k}_i = \bar{k}_{i+1} = \dots = \bar{k}_{|I|-1}\}, q \geq 0, q \in Z; \end{aligned}$$

- 8) if $|\bar{M}| > 2$, then we increase a smoothing radius $p := p + 1$ and go back to step 4 of the algorithm;
- 9) if $|\bar{M}| < 2$, then segmentation threshold cannot be found, we complete the algorithm;
- 10) determining the set of indices \underline{M} minimums \bar{k}_i

$$\begin{aligned} \underline{M} = \{ & i \in [1, |I| - 1]: \bar{k}_{i-q-1} > \bar{k}_{i-q} = \dots = \bar{k}_i = \dots = \bar{k}_{i+q} < \bar{k}_{i+q+1} \} \\ & \cup \{ 0: \bar{k}_0 = \bar{k}_1 = \dots = \bar{k}_i < \bar{k}_{i+1} \} \\ & \cup \{ i \in [1, |I|]: \bar{k}_{i-1} > \bar{k}_i = \bar{k}_{i+1} = \dots = \bar{k}_{|I|-1} \}, q \geq 0, q \in Z; \end{aligned}$$

- 11) calculating the segmentation threshold λ_0 : finding $\lambda_0 \in \underline{M}$, that satisfies the inequation $\min\{\bar{M}\} < \lambda_0 < \max\{\bar{M}\}$;
- 12) pixels $x_z, z \in Q$, for which $\lambda_z \geq \lambda_0$, are part of the object, the rest are the background.

The experiments on segmentation of the zones of interest were conducted on the chest X-ray images using the proposed algorithms. The zones of interest that contain pathology lesions were selected manually. The spot feature segmentation produced lower classification error compared to the mod method.

5. Conclusion

This paper introduces the method of searching for objects using spot features detected in the X-ray images of thoracic organs. The algorithms of the interest zones search and segmentation are presented. Experimental observations have proved the applicability of the proposed approach for the automated pathology detection. However, preliminary image processing is necessary before subsequent analysis of images of thoracic organs.

Acknowledgements. The work has been supported by the Russian Ministry of education and science (agreement: 14.606.21.0002, ID: RFMEFI60614X0002).

References

- [1] S. A. Bourylin, V. B. Fofanov, Attribute of a spot and its application for searching the zones of interest and segmentation. "The 6h German-Russian Workshop "PATTERN RECOGNITION and IMAGE UNDERSTANDING". Workshop proceedings, Novosibirsk 2003, 74 - 77.
- [2] Hamada R. H. and Brahim Belhaouari Samir, "A Computer Aided Diagnosis System for Lung Cancer based on Statistical and Machine Learning Techniques", Journal of Computers, Vol 9, No 2 (2014), 425 - 431, Feb 2014.
<http://dx.doi.org/10.4304/jcp.9.2.425-431>
- [3] Mohamed Meselhy Eltoukhy, Ibrahima Faye, Brahim Belhaouari Samir "A comparison of Wavelet and Curvelet for Breast Cancer Detection in Digital Mammogram," Elsevier, Journal of Computers in Biology and Medicine Journal at February Volume 40, Issue 4, April 2010, Pages 384 - 391.
<http://dx.doi.org/10.1016/j.combiomed.2010.02.002>

- [4] V. B. Fofanov, A. V. Demchenko, R. F. Kuleev, Deciphering multispectral images: methods and results // *Optical Society*, 2007, Vol. 74, № 3, 55 - 59.
- [5] V. B. Fofanov, R. F. Kuleev, An Approach to Estimation of Image Informativeness // *Pattern Recognition and Image Analysis*, 2009, Vol. 19, No. 3, 478 - 483. <http://dx.doi.org/10.1134/s1054661809030134>
- [6] R. M. Alev, S. A. Martynov, V. B. Fofanov, Remarks on Searching Zones of Interest in Locally Uniform Scene // *Pattern Recognition and Image Analysis*, 2011, Vol.21, No. 2, 212 - 215. <http://dx.doi.org/10.1134/s1054661811020027>
- [7] V. B. Fofanov, A. N. Zhiznevskii, Efficiency of Segmenting Zones of Interest in Locally Homogeneous Scenes // *Pattern Recognition and Image Analysis*, 2011, Vol.21, No. 2, 247 - 250. <http://dx.doi.org/10.1134/s1054661811020325>
- [8] J. Prewitt, M. Mendelson, The analysis of cell images // *Ann. N. Y. Acad. Sci.* 1966, Vol. 128, 1035 - 1053. <http://dx.doi.org/10.1111/j.1749-6632.1965.tb11715.x>

Received: October 12, 2014; Published: November 20, 2014