On an Approach to Object Recognition in X-ray Medical Images and Interactive Diagnostics Process

Nikita Chernuhin
Southern Federal University
Rostov-na-Donu, Russia

ABSTRACT

This paper proposes the approach to object detection and edge finding suitable for X-ray medical images and producing object boundaries in vector representation as polygons. Possible applications of this approach to interactive medical diagnosis process are discussed.

Keywords

Image processing, medical imaging, object recognition, edge detection

1. INTRODUCTION

Medical imaging remains one of the most intensively developing branches of digital image processing. This is caused by the rapid development of examination techniques as well as the integration of digital methods of examination.

Many up-to-date methods such as Active Appearance Models [1] require a large set of training images to build classifier or to apply statistical methods. Although such approaches often show brilliant results, they may be redundant on the given narrow classes of images.

Medical studies show that the variability of human bones among the people of given age, sex and constitution is relatively small, i.e. shapes of a healthy human's bones in the particular group are generally the same and mainly vary in scale.

Moreover, it's often almost impossible to obtain a large enough ground-truth training set of X-ray images corresponding to a particular disease within a given sex-age-constituion group.

On the other hand the X-ray examination being somewhat old and efficient method of diagnostics seems to be studied well enough to develop formal criteria of diagnostics. Such criteria include the change of shape, distances and optical density in various anatomical segments.

These criteria can be relatively easily formalized in the terms of an image morphology. The main goal of this research is to propose an approach which will allow to set diagnostics criteria in a declarative manner without using statistical methods.

2. OVERVIEW

The detection process is performed with the *initial template* T — a polygon of n points without self-crossings. This template serves as the base for object detection and also represents the object of interest in its normal healthy state.

Our approach consists of two stages: object detection and edge refinement. During the first stage the initial template has 4 degrees of freedom (two shift components, scale, rotation). We find such transformation of the initial template that results in the best match of template edges with the ones present in the target image.

During the edge refinement stage the template obtains $2n$ degrees of freedom (i.e. each of its points can move independently). The precise edge is then obtained via the active contours method.

3. PREPARATION

A few common steps are used both for processing and for obtaining the initial template. Given a grayscale image X of size $M \times N$ we perform noise filtering by applying a Gaussian blur operator:

$$\tilde{X} = G(\sigma, K) \circ X. \tag{1}$$

Here $G(\sigma, K)$ is the normalized matrix of Gaussian blur operator with standard σ deviation of size

$$2K + 1 \times 2K + 1$$

and $\circ$ denotes the convolution operation.

Gradient matrix $\bar{F}$ each element of which represents gradient vector at the corresponding image point is built by applying standard differential operators to the blurred image $\tilde{X}$. Sobel or Sharr operators were used:

$$\begin{aligned} \bar{F} &= \nabla \tilde{X}, \\ \bar{F}_{ij} &= (f^h_{ij}, f^v_{ij}). \end{aligned} \tag{2}$$

For some operations we also use a Canny-filtered gradient with some threshold.

3.1 Obtaining contour

The initial template T is obtained manually by an expert, following an X-ray image corresponding to a normal state of the anatomical part in target. A template is prepared for each sex-age-constitution group.

We developed a visual tool to simplify the template creation process. This tool makes template points tend to distribute in the areas of local gradient magnitude maxima, disallows self-crossing of a template and controls the density of the template points distribution.

Image gradient components are calculated simultaneously and stored together with the coordinates of template points.

Separate points and point sequences within the template can be labelled with arbitrary labels or with labels corresponding to real anatomical segments. These labels are later used for the formal description of diagnostics criteria as well as for displaying the results of the analysis.

Several templates make a scene which typically correspond to one of the standard X-ray examination projections.

4. RECOGNITION PROCESS

4.1 The contour distance

Shotton et al. [2] proposed a measure of distance between two contours which considered not only the Euclidean distance between contour points, but the orientation of image gradient as well:

$$d_0(E_1, E_2) = \frac{1-\lambda}{|E_1|\tau} \sum_{y \in E_1} EDT^{\tau}(E_2, y) + \qquad (3)$$
$$+ \frac{2\lambda}{\pi |E_1|} \sum_{y \in E_1} \omega_0(\varphi_{E_1}(y), \varphi_{E_2}(V(E_2, y))).$$

In (3) E_1, E_2 are contours, EDT^{τ} is the Euclidean distance transform of E_2 restricted to the maximum distance τ, so that $EDT^{\tau}(E_2, y)$ is the distance from the point y to the closest point of E_2. The ω_0 function returns an acute angle between gradient vectors and $V(E_2, y)$ is the point of E_2 closest to y. Values φ_{E_1} and φ_{E_2} are orientations of the gradient vectors in the given points of corresponding contours.

Parameter $\lambda \in [0, 1]$ affects the measure tendency to geometrical distance or to difference between vector orientations.

Considering an arbitrary class of images it's impossible to predict the exact orientation of the image gradient vector, i.e. we cannot say whether this vector is pointed inside or outside of an object. That is why the ω_0 in (3) compares vector orientations by using an acute angle between them.

However, we noticed that when working with X-ray images this restriction can be easily removed, because in this case we always know if the object of interest is lighter or darker than the background. So our modification of the (3) uses exact angular difference between vector orientations:

$$d(E_1, E_2) = \frac{1-\lambda}{|E_1|\tau} \sum_{y \in E_1} EDT^{\tau}(E_2, y) + \qquad (4)$$
$$+ \frac{\lambda}{\pi |E_1|} \sum_{y \in E_1} \omega(\varphi_{E_1}(y), \varphi_{E_2}(V(E_2, y))).$$

Here ω is the exact angular difference between corresponding vector orientations.

This modification not only improves the detection performance but it is also used for edge refinement.

4.2 Contour simplification

As the detection process time consumption depends linearly on the size of a contour, it is natural to try to reduce the contour size if possible. Paper [2] states that skipping some points during the distance evaluation still allows to perform detection. In our experiment using every fifth or even every tenth point still allowed to find an object in the scene. However, it is obvious that different points of a contour make different contribution to its shape. E.g., points lying on the relatively straight segments clearly make smaller contribution than the ones where the contour changes its direction.

Thus, contour simplification was performed which allowed to eliminate points lying on the straight segments of the initial template. The threshold ε was introduced to denote the maximum allowed deviation of a point from a straight line.

Such approach showed better performance results i.e. it allowed elimination of the bigger amount of points while keeping the recognition quality intact.

4.3 Detection

The detection process is quite straightforward. We introduce a hypothesis space Ξ, consisting of elements that represent spatial transforms of the initial template (shift, scale, rotation).

The detection process can be formalized as a minimization of the distance function in the hypothesis space:

$$d(T, E, \xi) \xrightarrow{\xi \in \Xi} \min \qquad (5)$$

Here T is the initial template, E is the set of points which have non-zero components in Canny-filtered gradient map of a target image. Thus, $d(T, E, \xi)$ is a distance between the initial template T transformed according to the hypothesis ξ and the contour obtained by processing of the target image:

$$d(T, E, \xi) = d(Transform(T, \xi), E), \xi \in \Xi. \qquad (6)$$

Then we build a grid in the Ξ and perform a full-search minimization. To reduce the computational expenses we also add to the grid some constraints based on the considerations of likelihood (e.g. the maximum rotation of a template cannot exceed $\frac{\pi}{4}$ etc.).

After the first step of minimization we build a more dense grid around the found local minima ξ_0. We repeat the minimization process until convergence which was typically achieved on the second iteration in our experiments.

Hypothesis ξ' found during this process is the result of detection stage and it represents such transformation of the initial template when it fits the edges present on the image processed.

We will denote the transformed template with T':

$$T' = Transform(T, \xi'). \tag{7}$$

5. EDGE REFINEMENT

The idea of the edge refinement is based on three assumptions:

- The likelihood of the given point being the part of the target object boundary is greater when the gradient magnitude at the given point is greater.
- The likelihood of the given point being the part of the target object boundary decreases with the distance increase to the closest point of T'.
- The likelihood of the given point being the part of the target object boundary decreases with the difference increase between the orientation of the gradient vector at the given point and that of the gradient vector associated with the nearest point of T'.

Having these assumptions the threshold function μ was introduced:

$$\begin{aligned}\mu(T', x) = \mu_0 \exp(\theta(\frac{1-\lambda}{\tau} EDT^{\tau}(T', x) + \\ + \frac{\lambda}{\pi}\omega(\varphi_{E_1}(y), \varphi_{E_2}(V(E_2, y)))),\end{aligned} \tag{8}$$

where μ_0 and θ are parameters obtained experimentally during the calibration. The actual values are discussed in [3, 6].

The first approach to edge detection [3] was to apply threshold μ to the Canny-filtered gradient map. Points that failed to withstand the threshold were removed from the map while the remaining ones were declared to make the object boundary. Although such approach demonstrated satisfactory visual results, its main disadvantage was to represent the edge as an unordered set of points.

Active contours approach suggested in [4] has since been widely researched and deservedly gained a solid position in image processing. Use of active contours allows to get a somewhat precise boundary of an object and, what is not less important, produces an edge as a polygon.

Well-known drawbacks of active contours method are their weak noise robustness, tendency to stick to the local image force extrema and the need to position the initial contour close to its desired position. We suggest a series of simple steps to minimize these drawbacks.

A transformed template was used as the initial position for an active contour. The image force was represented by image gradient magnitude.

5.1 Gradient filtering

A gradient filtering procedure was introduced to suppress foreign contours on the image. A filtered gradient $\bar{F}'$ is obtained by thresholding the gradient matrix $\bar{F}$ with the μ threshold.

This method appeared to allow to effectively suppress boundaries unrelated to the object of interest. Moreover, it also allows suppressing extraneous contours even if they are situated very close to the target contour because of the precise gradient orientation component in μ. The example of gradient filtering is given in Fig. 1.

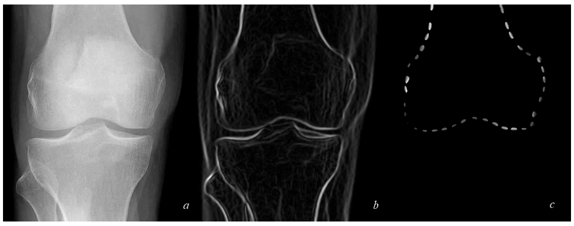

Figure 1: Initial image (a), gradient map (b) and filtered gradient(c).

As shown in Fig. 2 (*a,b*) boundaries of different bones are often situated very close to each other and naturally follow each other's shape. Therefore the use of gradient orientation is crucial in such cases.

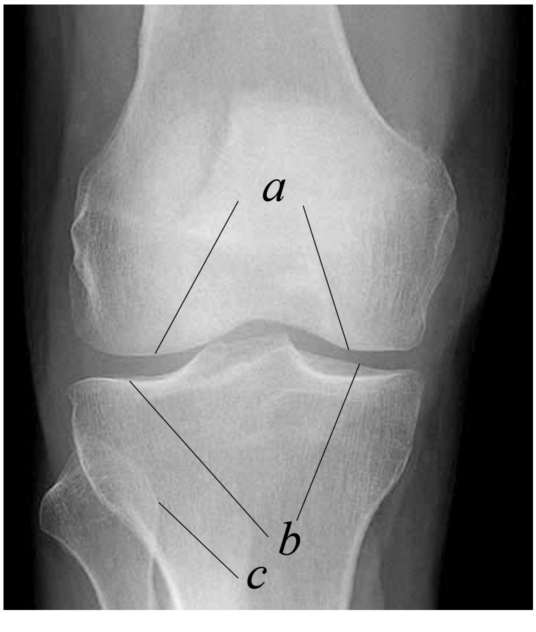

Figure 2: Boundaries of femur (a) are close to the boundaries of tibia (b). A part of fibula proximal part overlaps with tibia (c).

After gradient filtering the use of active contours becomes a much easier task. We used a classical greedy snakes algorithm [5], although the proposed approach does not lay any restrictions on the algorithm used and thus allows the use of any active contours algorithm including the most up-to-date ones [6].

6. CRITERIA FORMALIZATION

Given the initial template T we can easily calculate the mean of its points as

$$c(T) = \frac{1}{n} \sum_{y \in T} y. \tag{9}$$

An average distance from each point of a template to the $c(T)$ is

$$\bar{\rho} = \frac{1}{n} \sum_{y \in T} ||y - c(T)||. \tag{10}$$

So, for any two points x, y of a template distance we can calculate the *relative distance* between them:

$$\rho(x, y) = \frac{||x - y||}{\bar{\rho}}, \forall x, y \in T. \tag{11}$$

The relative distance remains the same during the detection stage. However, it may change during the edge refinement phase when the contour becomes active.

Thus, we can formalize any diagnostic criterion connected with distance change in terms of relative distances. A series of asserts can be made, including:

- An assert of equality of two or more relative distances;
- An assert of value range for a relative distance.

As the resulting boundary is a polygon, a triangulation of an object can be performed. Calculating the mean image intensity inside every triangle allows to estimate the relative optical density change in different parts of the object.

Also a curvature value at each template point can be estimated which gives another opportunity for criteria formalization.

7. CONCLUSION

The suggested approach allows detecting objects and their boundaries in the X-ray medical images. The visual quality of the detection was found to be satisfactory. The representation of the boundaries as polygons allows to set up diagnostic criteria in the form of asserts.

The performance of detection is comfortable for the operator while working with the standard images used in our experiments on modern machines. The full processing of an X-ray projection of a human knee takes about 2-3 seconds on 3.4GHz Core i7 machine.

The proposed approach does not require a large training set of images and time-consuming training process.

Further research of this approach may include finding a better way to solve minimization problem and experiments with up-to-date active contours algorithms.

REFERENCES

[1] Cootes T. F., Edwards G. J., Taylor C. J. Active Appearance Models, *Proc. European Conference on Computer Vision*, Vol. 2, pp. 484–498, 1998.

[2] Shotton J., Blake A., Cipolla R. Multi-scale categorical object recognition using contour fragments, *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 30, Issue 7. pp. 1270–1281, 2008.

[3] Babaev M. V., Pilidi V. S., Cheruhin N. A. A method of edge detection and edge finding for X-ray medical images, *Computer and informational technologies bulletin*, No. 8, pp. 41–45, 2012.

[4] Kass M., Witkin A., Terzopoulos D. Snakes: Active Contour Models, *International Journal of Computer Vision*, Vol 2., pp. 321–331, 1988.

[5] Williams D. J., Shah M. A Fast Algorithm for Active Contours and Curvature Estimation, *CVGIP: Image Processing*, Vol. 55, No 1, Jan., pp. 14–26, 1992.

[6] Chernuhin N. A. A Combined Method of Edge Detection for X-ray Medical Images Using Active Contour Approach, *Scientific Journal of the Kuban State Agrarian University*, Vol. 89, 2013.