A Dynamic Programming Approach for Heart Segmentation in Chest Radiographs

Aarti Raheja, Jason Knapp, Chunsheng Fang
{araheja,jknapp,vfang}@riverainmedical.com

Riverain Technologies
3020 South Tech Boulevard
Miamisburg, OH 45342-4860

Abstract

Chest radiographs are the most routinely acquired exams, which makes their use for diagnosis cost effective. In this paper we present a dynamic programming approach for automated heart segmentation on posterior-anterior (PA) chest radiographs. The goal of the proposed algorithm is to provide an accurate and reproducible method for heart segmentation, which can then be used to detect certain cardiac abnormalities. Our method has several advantages over previous methods, and provides superior performance to previously published results.

Introduction

Heart segmentation in chest radiographs is a challenging task. One major difficulty in segmenting the heart is the low contrast found in the mediastinum and diaphragmatic regions. These areas are difficult to visualize even by radiologists. Other aspects that make the problem challenging include: the significant variation in heart size across patients, the presence of disease in the lungs, and poor breadth holds by patients (leading to lower contrast on the heart boundary). Despite the challenges, development of an automated method for heart segmentation could provide significant clinical value [1].

Several methods have been proposed [1][2][3] for segmenting the heart. Nakamori et al [1] discuss a method to segment the heart by detecting points along the heart boundary, which are then fitted using a Fourier shape model. This method was used in [2] to automatically compute the cardiothoracic ratio (CTR) in 400 radiographs. Out of the 400 radiographs, 20% required manual intervention. It was also shown in [3] that the heart boundaries outlined by four experienced radiologists had a high degree of variability, which is an important result when considering how to assess automatic methods.

Van Ginneken et al [4] discuss several approaches to heart segmentation: active appearance model (AAM), active shape model (ASM) and pixel classification. The individual methods performed comparably well, though significantly better performance was obtained when a hybrid voting scheme was used to combine the three methods. Shape models, such as the ASM, have the drawback that their fitting routine can get caught in local optima [5]. This effect can become quite pronounced when applied to images that differ significantly from those used to build the model. This point is particularly important in our application as abnormal hearts are precisely what we're trying to detect. For this reason, we opted for a different approach.

One important use of heart segmentation is the measurement of the cardiothoracic ratio. The CTR is an important measurement that can imply cardiomegaly (abnormally large heart) [1]. The CTR is defined as the maximum transverse diameter of the heart, divided by the maximum internal diameter of the thoracic cage [6] Research in to methods for automatic CTR extraction has a long history [6]. Later in the paper we show how the CTR can be used for assessing the quality of a heart segmentation. Although the CTR can be computed without segmenting the heart, segmentation is useful as it can help radiologists validate the result. Figure 1 illustrates the idea.

We use an algorithm based on dynamic programming (DP) to segment the heart. DP, an important algorithm in Artificial Intelligence [7], is used in applications such as finding the shortest path within a graph. DP decomposes a complicated problem into simpler sub problems; and, based on Bellman's "Principle of Optimality", the optimal solution to the original problem can be obtained by combining the solutions to each sub problem.

In the proposed algorithm we formulate the DP sub problem in an innovative way. The cost matrix is generated using image information assigning minimum cost to the pixels having heart edge characteristics. The cost matrix is generated in the polar domain since the heart shape is mostly circular. By using this method we allow the shape to vary in regions where enough information is present, but force the shape to be circular in regions of uncertainty.

In the next sections we describe our algorithm based on dynamic programming in detail, followed by a presentation of extensive experimental results and a conclusion.

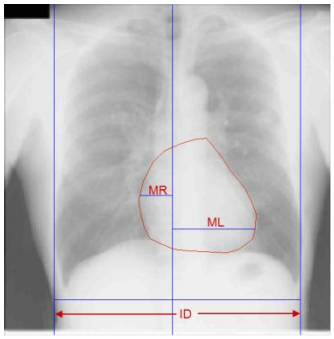


Figure 1: Chest Radiograph with heart outline showing the maximum internal diameter of the thorax (ID) and maximum transverse diameter of the heart that is the sum of maximum right heart width (MR) and maximum left heart width (ML).

Materials and Methods

We used the 247 chest radiographs from the JSRT database to test the method. The JSRT database is available publicly and consists of screen-film images digitized with a 0.175mm pixel size and 2048×2048 image size [8]. The heart annotations for this dataset [9] are available and were used to evaluate our method.
In Figure 2 a flowchart of the method is shown.

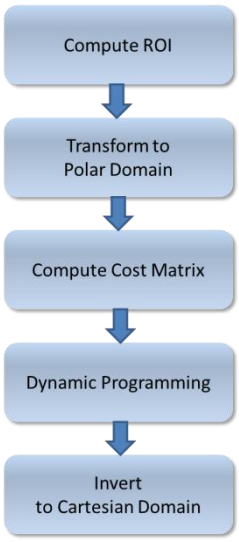


Figure 2: Algorithm Flowchart

Region of Interest around the Heart

We first obtain the ribcage mask and the segmented lung masks from the chest radiographs. This is done using a method developed by Riverain Technologies. The lung masks are then used to detect locations where the air, heart, and diaphragm intersect as shown in Figure 3. These locations are computed based on a curvature detection method as discussed in [10]. The average of these two locations, as shown in Figure 3, is used as the end row value to define an approximate bounding box around the heart region.

The top row of the bounding box, as shown in Figure 3, is selected as the location where the heart and the left lung first meet. The bounding box column locations, as shown in Figure 4, are the locations along each lung mask that are at a maximum distance from the central column.

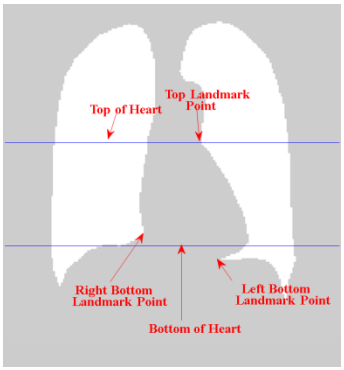


Figure 3: Landmark points computed using curvature information on the lung masks

This bounding box is used to define a center and a radius around the approximate heart region. The center is selected as the midpoint of the bounding box and the radius is selected as half of the distance between the end column locations.

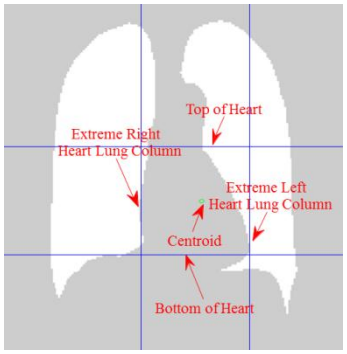


Figure 4: The heart region determined using the lung mask

Polar Transform

The border of the heart is roughly circular. For this reason we apply a polar transform defined in equations (1)-(3) to the approximate heart region.

$$I(x,y) \rightarrow J(r,\theta) \quad (1)$$

$$r = \sqrt{x^2 + y^2} \quad (2)$$

$$\theta = \text{atan}(y,x) \quad (3)$$

where $I(x, y)$ is the image in the Cartesian coordinate system and $J(r, \theta)$ is the image in the polar coordinate system.

The polar transform is applied to the image as shown in Figure 5(a) using the center and radius as defined in the previous section. To ensure all of the heart is included, the radius is multiplied by a factor α. In this paper we selected a α value of 1.5. The polar domain image as shown in Figure 5(b) is used to compute a cost matrix for the purpose of dynamic programming.

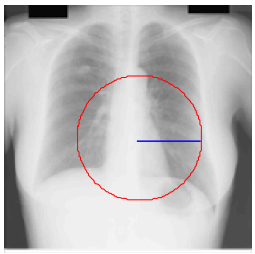


Figure 5.(a): Cartesian system image with center and radius marked for conversion in the polar domain

Figure 5.(b): Polar co-ordinate image expressed in terms of radial distance.

Dynamic Programming

In dynamic programming, the most important part is constructing the cost matrix. Each pixel in the cost matrix is assigned a local cost, where we use low cost values for pixels that have characteristics typical of the heart boundary. The local cost is defined as a linear combination of individual cost images:

$$local\_cost = w_{grad} * C_{grad} + w_{gsc} * C_{gsc} \quad (4)$$

where C_{grad} is the cost based on the gradient magnitude, w_{grad} is the weight assigned to C_{grad}, C_{gsc} is the cost based on a smoothed gray scale image and w_{gsc} is the weight assigned to C_{gsc}. The gradient is calculated by computing the derivative along each column (derivative in the radial direction). The gray scale cost term is defined by first computing a nominal value for the heart-lung border. This is done for each column within a smoothed image. These nominal values are then used to measure each pixel's deviation from the expected border value. Each local cost term is scaled to the unit interval prior to combining.

Given the local cost matrix, the next step is to compute the cumulative cost. The cumulative cost accounts for both the local and transitional costs. The transitional term weights the cost of going from one pixel to the next. The transitional cost we use increases with pixel distance, thus enforcing a smoother result. The total cumulative cost matrix is defined as follows:

$$C(i, 1) = local\_cost(i, 1) \quad (5)$$

$$C(i, j+1) = \min_{-k \le s \le k}\{C(i+s, j) + local\_cost(i, j+1) + T(s)\} \quad (6)$$

where T represents the transition cost. The value "s" is the offset between pixels when going from one column to the next. The value of this offset is not allowed to be larger than a specified value, "k", depending upon the desired path smoothness. The value of k for our experiments was set to 3 pixels.

Pixels outside the lung mask, or those having cost values above a maximum acceptable threshold, are set to the maximum cost value as shown in Figure 6. This causes a straight line to be the optimum path for these regions (circular arc in Cartesian domain).

Figure 6.(a): Original Cost Image

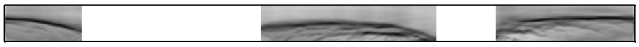
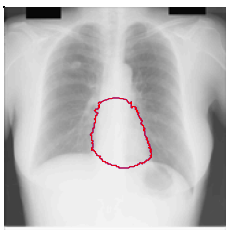


Figure 6.(b): Cost Image with non-air pixels suppressed

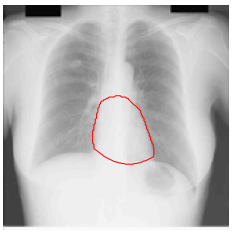
Figure 6.(c): Cost image with pixels having cost values above a maximum acceptable threshold value set to the maximum cost value

Once we obtain the optimal DP solution path, the heart segmentation is obtained by transforming the path to the Cartesian domain, as shown in Figure 7.

(a)



(b)



(c)

Figure 7: (a) The optimum path obtained from dynamic programing solution is converted into (b) Cartesian co-ordinate system to obtain the heart segmentation with (c) some post processing.

Some morphological post processing is applied to make the heart shape smooth and convex, see Figure 7 for an example.

Experiments

We carried out two experiments to validate the proposed method. First, the algorithm output is compared to the manual outlines to evaluate the accuracy of the heart segmentation. In a second experiment, we compared CTR values extracted from the algorithm against those extracted from manual outlines. The specific aim of this experiment was to evaluate if a reliable CTR estimate can be obtained even with a low overlap score.

The overlap score, Ω, between the manually outlined heart boundary and the output of our method is defined in equation (7).

$$\Omega = \frac{TP}{TP+FP+FN} \quad (7)$$

where TP is the true positive area, FP is the false positive area, and FN is the false negative area.
Figure 8 illustrates a summary of the overlap scores obtained by our method.

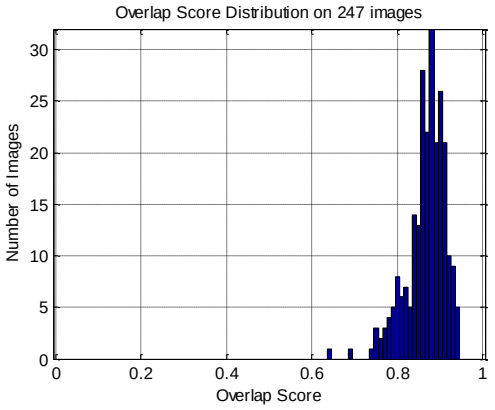


Figure 8: Histogram of overlap scores on 247 chest X ray images. Most of the overlap scores concentrate around 0.87, indicating the high accuracy of our method.

The CTR values are computed by detecting the internal diameter (ID) of the thorax and the transverse diameter of the heart (TD = MR+ML, Figure 1).

$$CTR = \frac{ID}{TD} * 100 \quad (8)$$

The ID value was derived from the ribcage mask. The TD values were computed using the heart mask derived from the algorithm output and the manual outlines.

A relative difference between the CTR values was computed using the above TD and ID values. Figure 9 shows a scatterplot comparing the overlap score with the relative CTR measure. From this plot we can deduce that a good CTR estimate can be obtained even with a low overlap score. An example of such a case is shown in Figure 10. The reason this can occur is that the source of low overlap is generally from the mediastinal and sub diaphragmatic regions, which do not influence the transverse diameter of the heart.

Figure 11 shows the only case with a low overlap score that was not due to the mediastinal or sub diaphragmatic regions. The difficulty here is the fusion of the left lung and colon. This leads to an inaccurate estimate of the left-lower landmark intersection location, which results in significant under segmentation. Fortunately, such an occurrence is rare and is left as an area for future improvement.

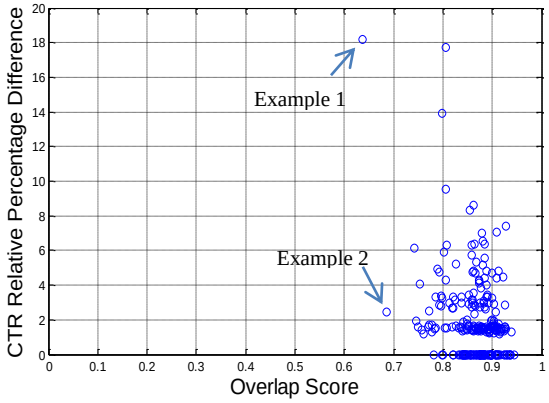


Figure 9: Scatterplot comparing the overlap score with the relative CTR measure

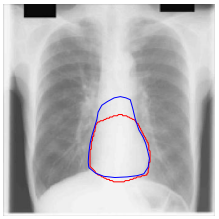


Figure 10: Example 1 with the lowest overlap score of 0.63

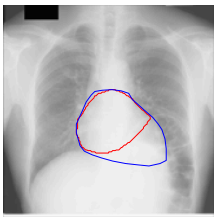


Figure 11: Example 2 having low overlap score, but a more accurate CTR value

Some typical output segmentations are presented in Figure 12. As can be seen, our proposed method captures the actual heart contour fairly accurately in most of the cases.

Discussion

An average overlap score of 0.867 ± 0.046 was obtained from the 247 JSRT images. We find that our method produces outputs that are close to the human observer, while comparing favorably to the other methods discussed in the survey paper [4]. The overlap scores in Table 1 are for the three hybrid methods discussed in [4]. These hybrid

methods make use of multiple methods making them computationally intensive. In addition, these methods are supervised approaches whose outputs might not extend to more atypical cases. By comparison, our method is far less complex and has the advantage of making very few assumptions about the shape of the heart.

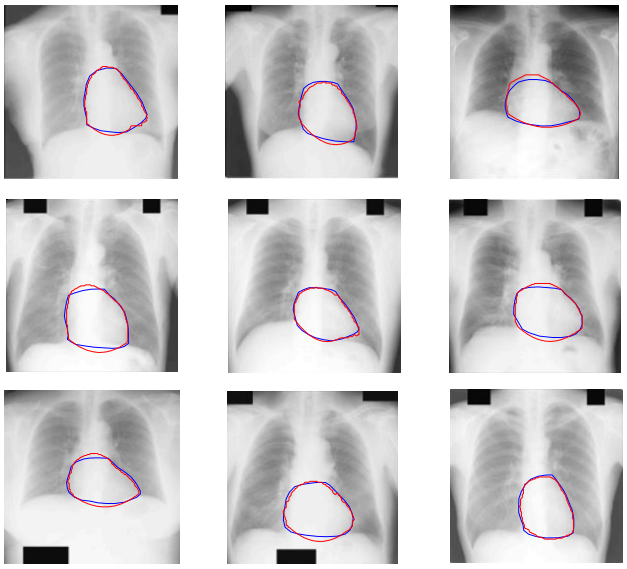


Figure 12: Heart Segmentation. Blue represents user annotation and red represents output of the current method.

Conclusion

We presented an algorithm for segmenting the heart region using dynamic programming. The proposed algorithm provided an accurate and reproducible method for heart segmentation. The presented method makes few assumptions about the heart shape, has a simple implementation, and provides superior performance to previously published results.

Future work will involve the collection of more data, which is needed for further evaluation and the development of strategies for handling outlier cases. Also, additional image features for improving the local cost term will be explored.

References

[1] N. Nakamori, K. Doi, V. Sabeti, and H. MacMohan, "Image feature analysis and computer-aided diagnosis in digital radiography: Automated analysis of sizes of heart and lung in chest images," Med Phys., vol.17: pp. 342-350, 1990.

[2]N. Nakamori, K. Doi, H. MacMohan, Y.Sasaki, and S. Montner, "Effect of heart-size parameters computed from digital chest radiographs on detection of cardiomegaly: Potential usefulness for computer-aided diagnosis," Investigat. Radiol., vol. 26: pp. 546-550, 1991

[3] R. Kruger, J. Townes, D. Hall, S. Dwyer, S. Lodwick. "Automated radiographic diagnosis via feature extraction and classification of cardiac size and shape descriptors," IEEE transaction on Biomedical Engineering., vol. BME-19:pp. 174-186, 1972

[4] B. Van Ginneken, M. Stegmann, M. Loog. "Segmentation of anatomical structures in chest radiographs using supervised methods: a comparative study on a public database", 2004.

[5] T. F. Cootes, C. J. Taylor, D. Cooper, and J. Graham. "Active shape models – their training and application," Computer Vision and Image Understanding., vol. 61(1):pp. 38–59, 1995.

[6] H. Becker, W. Nettleton, P. Meyers, J. Sweeney, Jr CM Nice. "Digital computer determination of a medical diagnostic index directly from chest X-ray images," IEEE Transaction on Biomedical Engineering., vol. BME-11:pp. 67-72, 1964.

[7] Stuart Russell, Artificial Intelligence: A Modern Approach .

[8] J. Shiraishi, S. Katsuragawa, J. Ikezoe, T. Matsumoto, T. Kobayashi, K. Komats, M. Matsui, H. Fujita, Y. Kodera, and K. Doi. "Development of a digital image database for chest radiographs with and without a lung nodule: Receiver operating characteristic analysis of radiologists' detection of pulmonary nodules". AJR., vol. 174:pp. 71-74, 2000.

[9] Image Sciences Institute Research Databases. http://www.isi.uu.nl/Research/Databases/.

[10] S. Muhammad, M Asif and M. R. Asin. "A new approach to corner detection," Computer imaging and vision, vol. 32:pp. 528-533, 2006.

Table 1: Overlap score results compared to a human observer and various methods discussed in [4].

Heart	μ±σ	min	Q1	median	Q3	max
Human Observer	0.878±0.054	0.571	0.843	0.888	0.916	0.965
Dynamic Programming	**0.867±0.0460**	**0.636**	**0.846**	**0.875**	**0.898**	**0.944**
Hybrid Voting	0.860±0.056	0.651	0.833	0.870	0.900	0.959
Hybrid ASM/PC	0.836±0.082	0.430	0.804	0.855	0.889	0.948
Hybrid AAM/PC	0.827±0.084	0.499	0.791	0.846	0.888	0.957