Watershed segmentation using prior shape and appearance knowledge

Ghassan Hamarneh *, Xiaoxing Li

School of Computing Science, Simon Fraser University, Burnaby, BC, Canada V5A 1S6

Abstract

Watershed transformation is a common technique for image segmentation. However, its use for automatic medical image segmentation has been limited particularly due to oversegmentation and sensitivity to noise. Employing prior shape knowledge has demonstrated robust improvements to medical image segmentation algorithms. We propose a novel method for enhancing watershed segmentation by utilizing prior shape and appearance knowledge. Our method iteratively aligns a shape histogram with the result of an improved k-means clustering algorithm of the watershed segments. Quantitative validation of magnetic resonance imaging segmentation results supports the robust nature of our method.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Watershed transformation; Prior shape knowledge; Segmentation; k-Means clustering

1. Introduction

The concept of watersheds is well known in topography. It was first proposed as a potential method for image segmentation in [1]. In [2], watershed transformation was simulated based on an immersion process, enabling an increase in speed and accuracy. Parallel watershed segmentation was later developed [3], offering clear partitions within images.

Watershed transformation has increasingly been recognized as a powerful segmentation process due to its many advantages [4], including simplicity, speed and complete division of the image. Even with target regions having low contrast and weak boundaries, watershed transformation can always provide closed contours. In addition, watersheds typically occur at the most obvious contours of the object, even when oversegmentation is severe. This positioning of contours can offer a stable and accurate initialization for other post-processing and segmentation techniques. A detailed review of algorithms that make use of the watershed transformation for image segmentation can be found in [5].

Oversegmentation and sensitivity to noise continue to plague watershed transformation with respect to medical image data. Typically, the gradient magnitude of the original image is computed before the watershed transformation is applied. Fluctuations in the gradient magnitude image, as well as negative impulse noise being regarded as a local minimum, can result in undesired additional watershed segments. Several methods have been proposed to overcome these drawbacks. Among the most notable is the use of region markers [6–8], in which certain desired local minima are selected as markers, then geodesic reconstruction is applied to fill the other minima to non-minimum plateaus. However, automatic marker selection is a difficult process and in most applications human interaction is required. In addition, if inappropriate markers are selected, geodesic reconstruction can cause important image features to be overlooked, thus dramatically changing the final result.

It has been demonstrated in several medical image segmentation techniques that improved convergence and robustness can be obtained when prior shape knowledge is utilized. In the seminal work of Cootes et al. [9], a training set of shapes represented by landmarks were used to build a point distribution model, which is used to constrain the segmented shapes to lie within an allowable

---

* Corresponding author. Tel.: +1 604 291 3007; fax: +1 604 291 3045.
E-mail addresses: hamarneh@cs.sfu.ca (G. Hamarneh), xli1@cs.sfu.ca (X. Li).
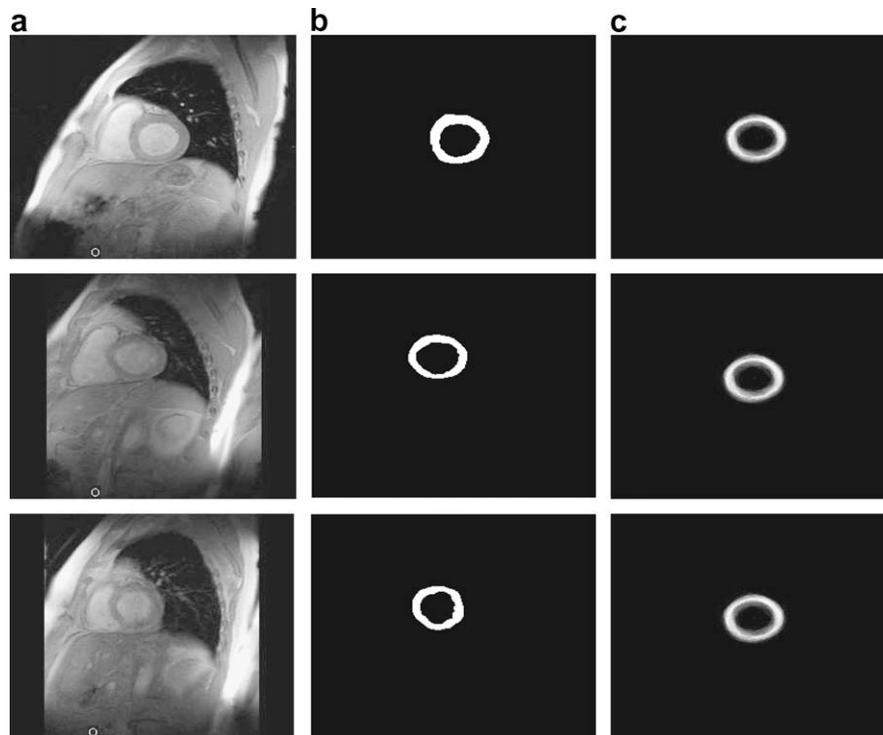
Fig. 1. The training cardiac image data [16,17]. (a) A selection of the grayscale MR images in the training set. (b) The expert-segmented binary masks of the cardiac muscle of each of the images in (a). (c) The shape histograms, constructed from the aligned binary images less the one belonging to the corresponding image in (a). This example is presented for illustration purposes in order to clarify the steps of the algorithm. Refer to Section 3 for more elaborate results and validation.

shape domain. In [12], prior shape models were constructed for level set based shape representations through the use of the signed distance transform. Prior shape knowledge has also been encoded through capturing the variability of Fourier-based shape descriptors in a deformable model framework [13,14]. In [10], the geometrical and physical layers of classical deformable models were complemented with behavioral and cognitive layers that facilitate encoding more complex prior shape and contextual knowledge.

The major detraction toward using prior shape and appearance models is the need to label a training set of images. The set should be large enough to encompass the typical shape and appearance variations one wishes to capture and model. Nevertheless, knowledge about typical shape variations of a structure of interest in medical images can be made available to support the incorporation of prior knowledge. Clearly, by using prior shape models one sacrifices the generality of the method in order to achieve improved robustness only for a specific class of object shapes. So although prior shape models ensure plausible shapes, they may produce incorrect results if the novel data is too different from the training set.

In this paper we propose a novel method for incorporating prior shape and appearance knowledge into the watershed segmentation technique.[1] The method is composed of a training stage and a segmentation stage. In the training stage, a 'shape histogram' and image intensity statistics are used to model prior shape and appearance knowledge, respectively. The segmentation stage is an automatic iterative procedure and consists of four steps: classical watershed transformation, improved k-means clustering, shape alignment, and refinement.

The rest of the paper is organized as follows. Using short-axis, cardiac MR image data as an example, Section 2 discusses the details of the algorithm including modeling prior knowledge (Section 2.1), watershed segmentation (Section 2.2), improved k-means clustering (Section 2.3), current estimate alignment (Section 2.4) and refinement (Section 2.5). In Section 3, we apply our method to segmenting the corpora callosa from a set of 51 brain MR images and provide numerical validation and discussion of the results. In Section 4, we conclude with a summary of our work and possible future directions.

2. Methods

2.1. Modeling shape and appearance knowledge

A training set of gray level images containing the object of interest is collected. In addition, the training set includes data representing expert segmentation of the images. This data is used to form binary images directly relating to the shape of the object of interest. Fig. 1b shows an example subset of binary images, created from the expert segmentation of the cardiac left ventricle wall from axial MRI slices

---

seen in Fig. 1a [16,17]. The binary shape images are then aligned by applying translation, rotation, and scaling factors, in order to normalize the location of the centroid, the direction of the principle axis, and the area of each of the binary shapes, respectively.

A suitable shape model should represent the detailed shape information and be able to model possible shape variations. To this end we provide a modification to the idea of a probabilistic atlas by incorporating additional information derived from the distance transform. First, a *shape histogram* is obtained by adding the values of corresponding pixels of the set of aligned binary shape images. The value at pixel p in the shape histogram, denoted by $SH(p)$, reflects the number of shapes containing p. Shape histograms for the sets of aligned expert-segmented regions are shown in Fig. 1c.

Second, by overlapping and aligning the shape histogram with the instance of the target structure in a new test image, we can construct a probability map in which the value of each pixel represents the probability that the underlying pixel p belongs to the target. Alignment details are described in Section 2.4. The probability map is denoted by $PM(p)$ and defined as

$$PM(p) = \begin{cases} s_2 + s_1 \cdot SH(p) & \text{if } p \in \mathfrak{R}, \\ s_2 \cdot [1 - Dist(p)] & \text{if } p \notin \mathfrak{R}, \end{cases} \quad (1)$$

where $\mathfrak{R}$ is the set of nonzero pixels in the shape histogram. $Dist(p)$ is the normalized two-dimensional Euclidean distance transform for the binary image having $\mathfrak{R}$ as foreground [18]. The two scalars, s_1 and s_2, satisfy $s_1 + s_2 = 1$ and determine the weight of shape information when constructing the probability map (Fig. 2).

In addition to capturing shape variation information from the expert-segmented binary images, we model the *appearance knowledge* from the intensity patch of the object of interest in the training set of gray level images. In medical images, anatomical structures have characteristic intensity features that can be utilized when locating similar structures in new images. Histogram equalization is first applied to the training set of gray level images and then

the mean, M_{app}, and variance, V_{app}, of the intensity patches are calculated and used as two appearance descriptors.

2.2. Watershed transformation

By utilizing prior knowledge our method can cope with the oversegmentation problem of the standard watershed transformation. In this section we present a brief review of the standard watershed segmentation algorithm for the discrete case using topographical distance. The reader is referred to [5] for a detailed treatment.

The lower slope, $LS(p)$, is the maximum slope connecting pixel p in the input image f, with its neighbors of lower altitude, which can be written as

$$LS(p) = \max_{q \in N(p) \cup p} \left(\frac{f(p) - f(q)}{d(p,q)} \right). \quad (2)$$

$N(p)$ is the set of neighbors of pixel p and $d(p,q)$ is the Euclidean distance between p and q. Note that when $q = p$, the right hand side of (2) is forced to be zero; thus we have a lower slope value even when p is a local minimum. Consequently, the cost for moving from pixel p to q is defined as

$$\text{cost}(p,q) = \begin{cases} LS(p) \cdot d(p,q) & \text{if } f(p) > f(q), \\ LS(q) \cdot d(p,q) & \text{if } f(p) < f(q), \\ \frac{1}{2}[LS(p) + LS(q)] \cdot d(p,q) & \text{if } f(p) = f(q). \end{cases} \quad (3)$$

If there exists a path $\pi = (p_0, \dots, p_l)$ from $p_0 = p$ to $p_l = q$, the *topographical distance along π between the two pixels p and q* is expressed as

$$T_f^\pi(p,q) = \sum_{i=0}^{l-1} d(p_i, p_{i+1}) \text{cost}(p_i, p_{i+1}). \quad (4)$$

$T_f^\pi(p,q)$ is the minimum of the topographical distances of all paths linking p and q. Following similar definition, the catchment basin $CB(m_i)$ of a local minimum m_i is defined as the set of pixels which have smaller topographical distances to m_i than any other local minimum. Finally, the set of pixels which do not belong to any catchment basin
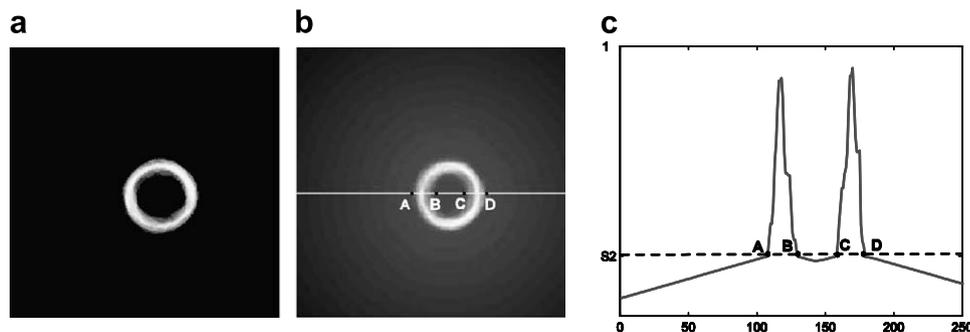
Fig. 2. The shape histogram in (a) is used to generate the probability map in (b). The probability profile values along the line ABCD in (b) are shown in (c). Note that the likelihood of a shape is higher in brighter regions, with the probability decreasing as we move away from the center of the probability map.

are referred to as the watershed pixels. Note that, in practice, the watershed transformation is usually computed on the gradient magnitude image instead of the gray value image. An example of watershed transformation can be seen in Fig. 3b.

2.3. Improved k-means clustering

After performing a standard watershed transformation, we make use of the k-means clustering algorithm to assign the resulting segments into different classes. k-means is a clustering algorithm that assigns N data points to k disjoint subsets, S_j, $j = 1, 2, \ldots, k$, each containing N_j data points, by minimizing the sum-of-squares criterion given by

$$J = \sum_{j=1}^{k} \sum_{n \in S_j} \sqrt{|x_n - \mu_j|^2}, \qquad (5)$$

where x_n is the value of nth data point and μ_j is the mean value of the data points within the cluster S_j.

k-means clustering algorithm can effectively assign the watershed segments into several clusters according to different features. In our algorithm, we jointly utilize two kinds of features for each segment; (i) the mean intensity and (ii) the spatial centroid position of each segment. For the nth segment, these features are denoted by I_n and C_n, respectively. Accordingly, we design an improved k-means clustering scheme with a modified sum-of-squares criterion, given by

$$J' = \sum_{j=1}^{k} \sum_{n \in S_j} \sqrt{|I_n - \eta_j|^2 + \alpha |C_n - \mu_j|^2}, \qquad (6)$$

where η_j and μ_j denote the average values of I_n and C_n, respectively, $\forall$ n belonging to the jth cluster. In (6), note that the values of I_n and C_n are each normalized in advance. By minimizing (6), each resulting cluster of watershed segments will contain segments with similar intensity. Furthermore, the watershed segments within each cluster tend to be in close proximity and the likelihood of assigning non-neighboring segments to the same cluster is reduced. The parameter α determines the weight for penalizing distant segments when clustering. The value of α is chosen to be inversely proportional to the average distance between the two farthest points of an object in the training set (zooming). An excessively large α, for example, can hinder segments belonging to a very thin or poorly connected structure from being grouped together.

To overcome the sensitivity of the k-means algorithm to different initializations, we begin the process by ordering the segments according to their mean intensities and assigning the same number of segments to each of the k clusters.

After clustering, we locate an initial estimate of the target structure by selecting the cluster whose intensity profile
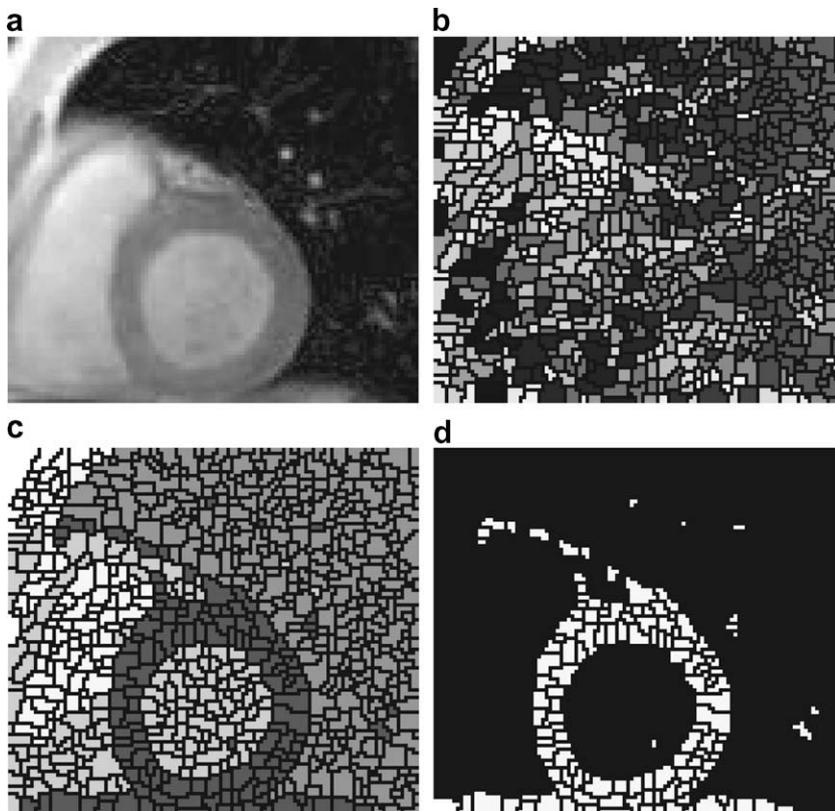
Fig. 3. Calculating the initial estimate of left ventricle segmentation. (a) Original MR cardiac image. (b) Watershed transformation of the image. (c) Coalesced watershed segments resulting from improved k-means clustering. (d) The initial segmentation estimate.
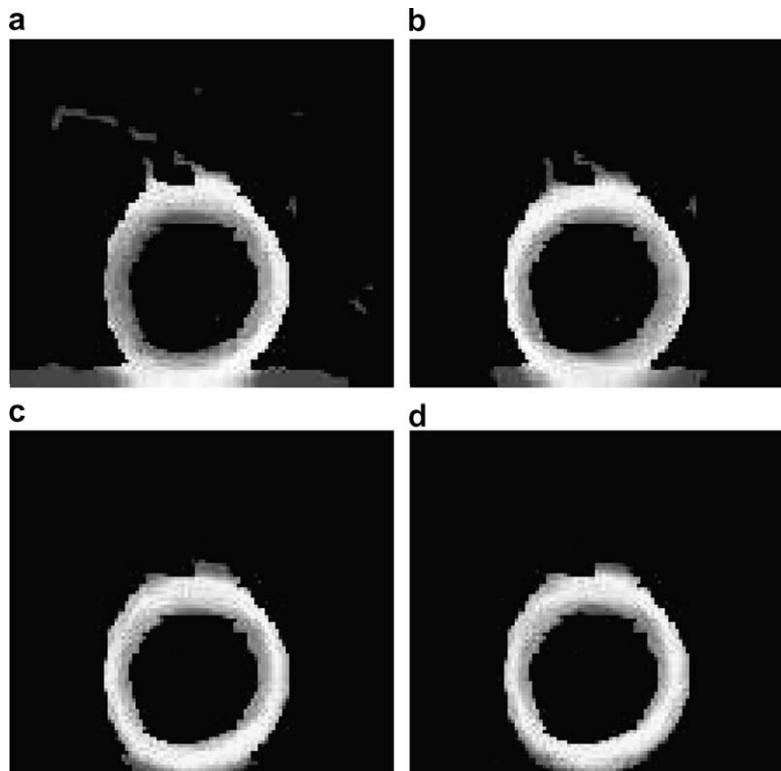
Fig. 4. Iterative shape alignment to the wall of the left ventricle. The intensity gradation represents the prior shape probability map, which is masked by the current estimate of the target. Note how (dark) pixels with lower probability are gradually eliminated and an improved estimate of the target is obtained. (a) Initial alignment estimate. (b) After sixth iteration. (c) After 15th iteration. (d) Convergence after 21st iteration.

is most similar to the appearance knowledge obtained from the training set (Section 2.1). Specifically, after histogram equalization, we calculate the mean and variance of the intensity for each cluster. We then choose the cluster having the most similar intensity statistics to those calculated in the training. This is expressed as finding the jth cluster that minimizes

$$DF(j) = \sqrt{(m_j - M_{app})^2 + (v_j - V_{app})}, \qquad (7)$$

where m_j and v_j denote the mean and variance of the intensity of the jth cluster, respectively. The set of all watershed pixels belonging to this selected cluster is considered to be the initial segmentation estimate and will be iteratively
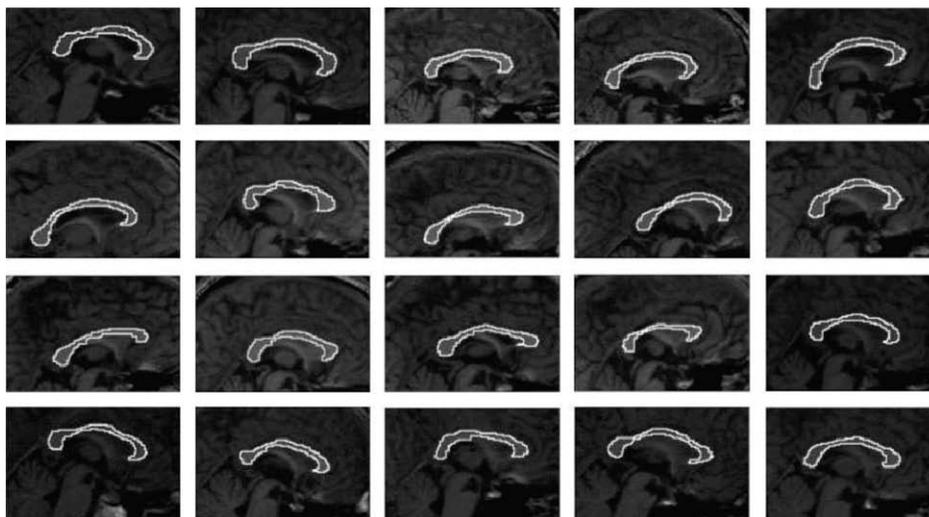
Fig. 5. Examples of corpus callosum segmentation results.

improved (Section 2.4). Fig. 3 demonstrates initializing the first estimate of the region to be segmented by applying our improved *k*-means clustering to a watershed segmentation of an MR image. Note in Fig. 3b how classic watershed segmentation is characterized by excessive oversegmentation. Our clustering technique detects a good first approximation of the region of interest (Fig. 3c), albeit initially including some undesirable segments (upper left corner and along bottom of Fig. 3d). The iterative shape alignment (Section 2.4) will gradually remove these undesired segments which are typically expected to appear.

2.4. Iterative shape alignment

In this section we iteratively merge segments and eliminate undesired ones by aligning the current estimate with the shape histogram. Initially, however, we perform morphological closing operation on the initial estimate (obtained in Section 2.3) to join the segments separated by watershed pixels, the result of which is called *target T*. Now, in the *i*th iteration of the shape alignment step, the area, orientation and centroid of the current target T_i are calculated. Then the shape histogram is scaled, rotated and translated (using bi-cubic interpolation) to be aligned with T_i, and a probability map is constructed accordingly. Finally the pixels in T_i with the lowest probabilities are removed, resulting in an updated target T_{i+1}. The *i*th iteration of this procedure can be expressed by the following steps:

Step 1. *Align:* Resize, rotate and translate the shape histogram to align it with the target

$$\text{SH}_i = R[(A_{T_i}/A_{\Re}) \cdot \text{SH}] + (C_{T_i} - C_{\Re});$$

Step 2. *Probability map:* Construct the probability map PM_i based on SH_i according to (1);

Step 3. *Remove:* The pixels in target T_i having the lowest probability are removed.

$$T_{i+1} = \{p | p \in T_i \cap$$
$$\text{PM}_i(p) > \sigma \max_{q \in T_i}(\text{PM}_i(q)) + (1 - \sigma) \min_{q \in T_i}(\text{PM}_i(q))\}.$$

In the above procedure, SH represents the shape histogram obtained from the training set of images. $A_{\Re}$ and $C_{\Re}$ are the area and centroid of the nonzero region $\Re$ of the shape histogram, respectively. Similarly, A_{T_i} and C_{T_i} are the area and centroid of the target T_i, respectively. By applying the rotation matrix R, the shape histogram will rotate by a degree of $\Theta = \theta_{T_i} - \theta_{\Re}$, where θ_{T_i} and $\theta_{\Re}$ are the orientations of the principle axes of T_i and $\Re$, respectively. σ is the step size associated with pixel removal for each iteration. A small σ generates a more accurate result but requires more iterations for the algorithm to converge.

The algorithm converges when all the remaining pixels have probabilities larger than s_2, which indicates that the shape of the current target is similar enough to the shape

histogram and the desired structure is well segmented. An example illustrating the intermediate results of this process is shown in Fig. 4. In this case, reasonable convergence occurred after 21 iterations. Using the target obtained in the final iteration, we can determine which segments should

Table 1
The mean, maximum, and standard deviation of the Hausdorff distances between our result and expert segmented corpora callosa boundaries

MRI No.	Max error (pixels)	Mean error (pixels)	SD error (pixels)
1	3.6056	0.4451	0.6340
2	4.0000	0.6514	0.7902
3	3.1623	0.5835	0.5784
4	2.0000	0.5212	0.5737
5	2.2361	0.4004	0.5269
6	2.2361	0.5205	0.5441
7	2.0000	0.5148	0.5427
8	2.0000	0.5994	0.5286
9	3.0000	0.4962	0.5723
10	3.6056	0.5975	0.6183
11	2.2361	0.3973	0.5350
12	3.1623	0.6404	0.6677
13	2.8284	0.5402	0.7000
14	2.0000	0.4812	0.5288
15	1.4142	0.3367	0.4777
16	3.6056	0.5143	0.6551
17	4.0000	0.4762	0.6633
18	1.4142	0.4034	0.4954
19	2.0000	0.5200	0.5258
20	1.4142	0.3430	0.4794
21	2.0000	0.4075	0.5302
22[*]	2.0000	0.5234	0.5495
23	2.8284	0.6127	0.6517
24	2.2361	0.4324	0.5364
25	2.2361	0.5491	0.5530
26	3.0000	0.5519	0.6029
27	3.6056	0.5179	0.6969
28	2.0000	0.5019	0.5279
29	2.0000	0.5335	0.5694
30	3.0000	0.5929	0.6114
31	–	–	–
32	2.0000	0.3944	0.5164
33	2.2361	0.4308	0.5503
34	2.0000	0.4907	0.5152
35	1.4142	0.4407	0.5118
36	3.1623	0.5894	0.5674
37	2.8284	0.5002	0.6280
38	3.6056	0.4650	0.6325
39	1.4142	0.4198	0.5035
40	1.4142	0.3758	0.4930
41	2.0000	0.4346	0.5231
42	2.0000	0.4231	0.5252
43	1.4142	0.4716	0.5155
44	2.0000	0.4516	0.5485
45	2.0000	0.5768	0.5212
46	3.6056	0.5109	0.6629
47	1.0000	0.3754	0.4704
48	2.0000	0.3650	0.4889
49	1.4142	0.4785	0.5131
50	2.8284	0.3875	0.5538
51	1.4142	0.5364	0.5319
Average	2.3716	0.4856	0.5648
SD	0.7927	0.0806	0.0678

[*] $k = 4$.
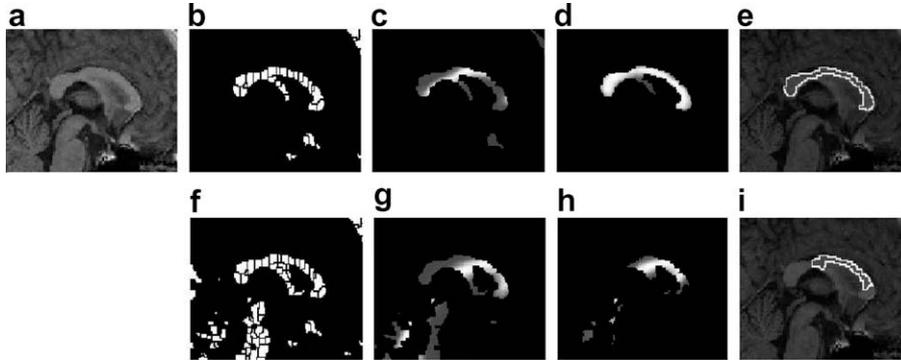
Fig. 6. Effect of initialization. (a) Original MRI. (b) Initialization with $k = 7$. (c)–(d) Intermediate results of iterative alignment based on initialization of (b). (e) Satisfactory segmentation result. (f) Initialization with $k = 4$. (g)–(h) Intermediate results of iterative alignment based on initialization of (f). (i) Unsuccessful segmentation result.

be kept. These segments are merged together to form the segmented result of the desired organ.

2.5. Refining segmentation results

In some areas along the boundary of the segmented result (acquired in Section 2.4), the extracted boundary may fail to follow the true boundary perfectly. This can occur when some of the small segments near the boundary of the object are mis-clustered because their mean intensities are not similar enough to the segments inside the organ. This is remedied by a simple refinement step in which nearby segments are merged only if they are covered by $\Re$ and have mean intensities similar to that of the target structure. Furthermore, the boundary segmentation result exhibits a characteristic staircase effect. This is due to the fact that watershed segments, and consequently the extracted boundary, is a union of *whole* pixel (and not subpixels). A smooth contour with subpixel resolution is easily obtained by initializing an active contour model (snake) [20–22] with the resulting watershed boundary pixels. Subsequent snake deformation iterations according to image gradient and tensile and flexural forces, provides satisfactory results with only a few iterations (see the results in Section 3).

3. Results

The proposed algorithm was tested and validated on the 51 brain MR images in [19], in which expert-segmented corpora callosa for all images in the data set were provided. To ensure cross validation, when segmenting an image, the corresponding expert-segmented result and appearance information were not included in the training set, i.e. a leave-one-out scheme was applied in which 50 MR images with the corresponding expert segmentation were used as the training set for segmenting the remaining single image. The shape histogram is obtained by aligning and adding the expert segmented results. Each test image contains a brain region cropped from the original image with a dimension of 91×141. The cropped brain parts of the 51 MR images mostly cover the region beginning at the calvaria, reach the upper part of the pons and have a fixed aspect ratio. The intensity and contrast of these images vary from one to another and the shape variation of the corpora callosa is considerably large. The original image has been pre-processed by a median filter prior to computing the gradient. Example corpus callosum segmentation results are presented in Fig. 5.

A constant value of $\alpha = 0.5$ for all the MR images was used in k-means clustering. By selecting $\alpha < 1$, we actually put more emphasis on intensity information and use spatial position as additional constraints. We choose $\sigma = 2.5\%$ as the removal step size in each iteration. In general, for less noisy images, σ can be larger since there will not be many small segments spreading around the structure. However, to ensure the accuracy of the results, it should be no larger than 5%. On the other hand, for images with much more noise, σ should be smaller and the algorithm will require more iterations to converge. In our experiment, most cases converged in less than 10 iterations. The two scalars s_1 and s_2 in (1) are selected to be $s_1 = 0.7$ and $s_2 = 0.3$. We always use $s_1 > s_2$ since shape information is the dominant factor.

The results are compared with the corresponding expert segmented ones in the original data set. To demonstrate the accuracy of our method, the maximum, average and standard deviation of Hausdorff distance [23] are calculated and reported for all cases (Table 1). In our earlier work on segmenting the corpus callosum from the same set of mid-sagittal MR images [11], only 26 out of 51 images were successfully segmented with the same set of parameters,
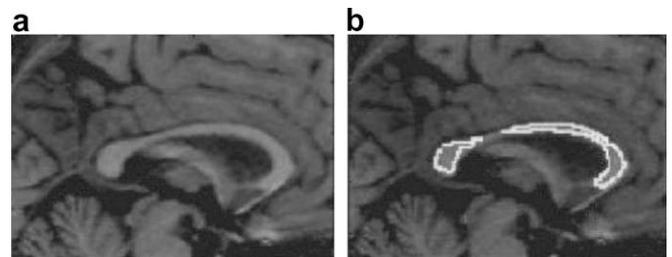
Fig. 7. A failure case. (a) Original image No. 31. (b) Segmentation result.

compared to 49 out 51 in this work. Our results differ very slightly from the expert segmented results.

However, we report two failure cases where a satisfactory result cannot be generated automatically. One failure case is image No. 22. We used an experimental value of $k = 7$ in k-means clustering and obtained satisfactory results for all the images except image No. 22, which had exceptionally low contrast. Consequently, the k-means algorithm did not successfully classify the segments belonging to the corpora callosa into the same cluster. For that case only, a good result is still obtained with k adjusted to $k = 4$. In Table 1, this case is marked with an asterisk (∗) and the maximum, average and standard deviation of Hausdorff distance are all calculated using $k = 4$.

It is important to note that the quality of initialization is crucial for the subsequent iterative shape alignment and in turn depends on the choice of the number of clusters k in the k-means clustering algorithm. The difficulty in determining k is a intrinsic limitation of the k-means algorithm. However, since all training images exhibit the same field of view (or 'zooming' as described in Section 2.3) and brain constituents (white matter, gray matter and cerebrospinal fluid), the intensity distribution was concentrated around certain intensity ranges and a single k value was suitable for all images (except the failure case reported). When the desired target is clustered into different clusters or too many undesired segments are clustered into the same cluster along with the target, the algorithm fails to produce a
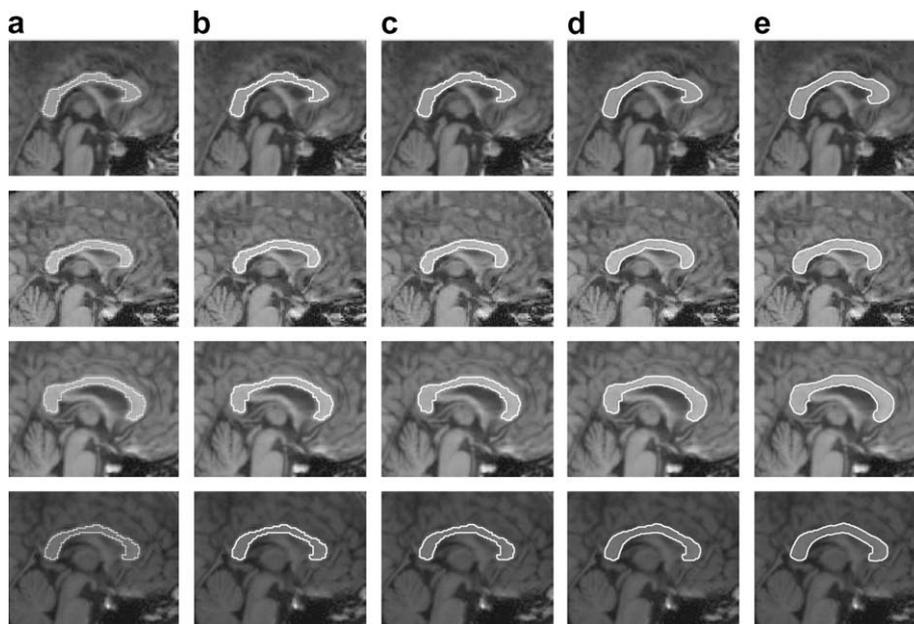
Fig. 8. Obtaining smooth boundaries with subpixel resolution. (a) The results acquired after iterative shape alignment (note the staircase effect in the detected boundaries). (b)–(d) An iterative snake process employed to clean aliased edges and provide boundaries with subpixel resolution. (e) The final segmented result.
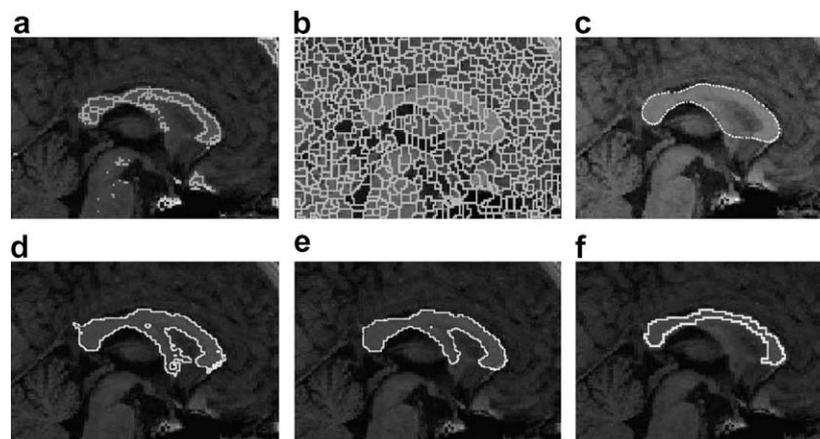
Fig. 9. Qualitative comparison with typical results of other methods. Simple thresholding (a) clearly does not give acceptable results, classical watershed segmentation suffers from over-segmenting the image. Snakes (c), region growing (d), and level set based deformable models (e), leak due the weak boundary. Our proposed method (f) incorporates prior shape and appearance knowledge to merge watershed segments.

correct segmentation. Fig. 6 demonstrates the effect of changing k. In Fig. 6a, we set $k = 7$ and obtain the initialization shown in Fig. 6b. This yields a satisfactory segmentation after seven iterations (Fig. 6e). Setting $k = 4$ for the same image results in a different initialization (Fig. 6f) and 22 iterations for the algorithm to converge to the erroneous result in Fig. 6i. Such a failure due to an improper initialization, which is caused by an improper choice of k, can be detected by comparing the segmentation results with the training set and observing that a large number of iterations is required for convergence. In those cases a different k must be chosen.

Another failure case is image No. 31. The middle part of the corpora callosa in this image is very thin and noisy as shown in Fig. 7a. The watershed transformation was unable to distinguish this part as a separate segment. Consequently, only the two ends of corpora callosa were well segmented and the result is hence broken in the middle, as shown in Fig. 7b.

Active contour models (snakes) was then initialized with the previous results in order to obtain contours with sub-pixel accuracy and only a few snake deformation iterations are executed[2] (see Fig. 8). The error values reported (Table 1) compared our results *prior* to applying snakes to the expert segmented results, simply because expert segmented results are obtained without subpixel accuracy.

In Fig. 9c, we show a typical example of running snakes directly on the original data. The figure demonstrates how snakes leak from areas of weak edges. Fig. 9 presents other qualitative comparisons with other methods (thresholding, classical watershed, region growing, and level set based segmentation[3]). Fig. 9 helps to emphasize the advantage of our algorithm in automatically locating and segmenting the target. This advantage stems from the utilization of prior shape and appearance knowledge while relying on watershed segments derived from the image.

4. Conclusion

In this paper we propose a new approach to improve watershed segmentation by incorporating shape and appearance knowledge. The method overcomes some intrinsic problems with watershed transformation. The oversegmentation problem is handled by clustering and merging appropriate watershed segments. The effect of noise is suppressed by computing the mean intensity of each segment. As a result, even when the input image has a relatively low resolution and low contrast, our algorithm achieves relatively accurate results. Our method is marginally affected by some of the limitations of the k-means clustering algorithm and a failure case is reported. However, this case can be detected by comparing the segmentation result with the shape histogram and adjusting the number-of-clusters parameter in k-means to obtain a satisfactory result.

Future work includes applying our method to other data sets as well as extending it to 3D. Given that the main building blocks of our method (watershed, k-means, and spatial transformations for alignment) are well developed in 3D, we foresee no major difficulties. Segment-level iterative alignment techniques are also being developed to further ensure segments along weak boundaries or thin structures are not lost. Additionally, more advanced appearance models may be used when the intensity variations within the target structures are more complex as in active appearance models [24]. This was not an issue for the corpora callosa segmentation application.

---

[2] Snakes software written in MATLAB (The MathWorks, Inc.) was used. The code is available for download from http://www.cs.sfu.ca/~hamarneh.

[3] Region growing and level set methods are implemented using 'itkConnectedThresholdImageFilter' and 'ThresholdSegmentationLevelSetImageFilter' from the Insight Segmentation and Registration Toolkit, ITK (www.itk.org).

References

[1] S. Beucher, C. LantuTjoul, Use of watersheds in contour detection, in: Proceedings of International Workshop on Image Processing, Real-time Edge and Motion Detection/Estimation, vol. 132, 1979, pp. 2.1–2.12.

[2] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE transactions on pattern analysis and machine intelligence 13 (1991) 583–598.

[3] A.N. Moga, M. Gabbouj, Parallel image component labelling with watershed transformation, IEEE transactions on pattern analysis and machine intelligence. 19 (1997) 441–450.

[4] S. Beucher, F. Meyer, The Morphological approach to segmentation: the watershed transform, in: E.R. Dougherty (Ed.), Mathematical Morphology in Image Processing, Marcel Dekker, New York, 1993, pp. 433–481.

[5] J.B.T.M. Roerdink, A. Meijster, The watershed transform: definitions, algorithms and parallelization strategies, Fundamenta Informaticae 41 (2000) 187–228.

[6] V. Grau, A.U.J. Mewes, M. Alcañiz, Improved watershed transform for medical image segmentation using prior information, IEEE transactions on medical imaging 23 (4) (2004) 447–458.

[7] J.L. Vincent, Morphological grayscale reconstruction in image analysis: application and efficient algorithms, IEEE transactions on image processing 2 (1993) 176–201.

[8] S. Beucher, Watershed, hierarchical segmentation and waterfall algorithm, in: Mathematical Morphology and Its Applications to Image Processing, Kluwer, Dordrecht, The Netherlands, 1994, pp. 69–76.

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

[10] G. Hamarneh, T. McInerney, D. Terzopoulos, Deformable organisms for automatic medical image analysis, in: Proceedings of Medical Image Computing and Computer-Assisted Intervention, in: Lecture Notes in Computer Science, vol. 2208, 2001, pp. 66–75.

[11] T. McInerney, G. Hamarneh, M. Shenton, D. Terzopoulos, Deformable organisms for automatic medical image analysis, Journal of Medical Image Analysis 6 (3) (2002) 251–266.

[12] M. Leventon, W. Grimson, O. Faugeras, Statistical Shape Influence in Geodesic Active Contours, in: Proceedings of the Computer Vision and Pattern Recognition, 2000, pp. 1316–1323.

[13] G. Szekely, A. Kelemen, C. Brechbuehler, G. Gerig, Segmentation of 3D objects from MRI volume data using constrained elastic deformations of flexible Fourier surface models, Medical Image Analysis 1 (1) (1996) 19–34.

[14] G. Hamarneh, T. Gustavsson, Statistically constrained snake deformations, Proceedings of the Systems, Man, and Cybernetics 3 (2000) 1610–1615.

[15] X. Li, G. Hamarneh, Modeling prior shape and appearance knowledge in watershed segmentation, Canadian Conference on Computer and Robot Vision (2005) 27–33.

[16] M.B. Stegmann, R. Fisker, B.K. Ersbøll, Extending and applying active appearance models for automated, high precision segmentation in different image modalities, Proceedings of the 12th Scandinavian Conference on Image Analysis 1 (2001) 90–97.

[17] M.B. Stegmann, An Annotated Dataset of 14 Cardiac MR Images, Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, 2002.

[18] H. Breu, J. Gil, D. Kirkpatrick, M. Werman, Linear time euclidean distance transform algorithms, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (5) (1995) 529–533.

[19] M.E. Shenton, R. Kikinis, F.A. Jolesz, S.D. Pollak, M. LeMay, C.G. Wible, H. Hokama, J. Martin, D. Metcalf, M. Coleman, R.W. McCarley, Abnormalities in the left temporal lobe and thought disorder in schizophrenia: a quantitative magnetic resonance imaging study, New England Journal of Medicine 327 (1992) 604–612.

[20] M. Kass, A. Witkin, D. Terzopoulos, Snakes: active contour models, International Journal of Computer Vision 1 (4) (1988) 321–331.

[21] G. Hamarneh, A. Chodorowski, T. Gustavsson, Active contour models: application to oral lesion detection in color images, IEEE International Conference on Systems, Man, and Cybernetics 4 (2000) 2458–2463.

[22] S. Lobregt, M.A. Viergever, Discrete dynamic contour model, IEEE Transactions on Medical Imaging 14 (1) (1995) 1224.

[23] P.K. Agarwal, M. Sharir, Efficient algorithms for geometric optimization, ACM Computing Surveys 30 (4) (1998) 412–458.

[24] T.F. Cootes, G.J. Edwards, C.J. Taylor, Active appearance models, IEEE PAMI. 34 (6) (2001) 681–685.