

An Accurate Tongue Tissue Strain Synthesis using Pseudo-wavelet Reconstruction-based Tagline Detection

Xiaohui Yuan^a and Cengizhan Ozturk^b and Gloria Chi-Fishman^c

^aDepartment of Computer Science and Engineering, University of North Texas;

^bLab of Cardiac Energetics, NHLBI, National Institutes of Health;

^cPhysical Disabilities Branch, Rehabilitation Medicine Dept., National Institutes of Health

ABSTRACT

This paper describe our work on tagline detection and tissue strain synthesis. The tagline detection method extends our previous work¹⁶ using pseudo-wavelet reconstruction. The novelty in tagline detection is that we integrated an active contour model and successfully improved the detection and indexing performance. Using pseudo-wavelet reconstruction-based method, prominent wavelet coefficients were retained while others were eliminated. Taglines were then extracted from the reconstructed images using thresholding. Due to noise and artifacts, a tagline can be broken into segments. We employed an active contour model that tracks the most likely segments and bridges them. Experiments demonstrated that our method extracts taglines automatically with greater robustness. Tissue strain was also reconstructed using extracted taglines.

Keywords: Tagline, Wavelets, Active Contour

1. INTRODUCTION

To study in vivo organ, e.g. human heart and tongue, tagged magnetic resonance (tMR) imaging is usually employed. The tagged MR imaging technique, initially developed for the assessment of myocardial motion, requires the insertion of magnetic tags that are essentially voids in the images, created by spatially selective pre-saturation pulses. They appear as dark lines on the imaged tissue (hence taglines) in a spatially encoded pattern and move with the tissue in motion, thus serving as fiducial markers for measuring the intrinsic tissue displacement. In our exploratory of human tongue tissue strain, tMRI provides a non-invasive means for studying dynamic physiological deformation within lingual tissues. The human tongue is a complex myoarchitecture and any movement is typically the result of the synergistic actions of several muscles. The myoarchitecture complexity makes it difficult to acquire tagged images and hence many issues have to be dealt with in the post processing step. In this paper, we focus on tagline detection and strain synthesis.

Since late 80's, continuous efforts have been devoted to improve the applicability of tMR imaging and the quality tagline detection. Hand tracing method was firstly used, which is tedious and time-consuming. Several methods were then developed for detecting taglines in the next decades^{3,5,6,9,10} that attempted to tackle the problem from texture and frequency aspects of the tag patten. Chen et al.³ used Markov random fields within the MAP framework for detecting taglines. The MAP estimation starts with an initial solid with less knots and lower spline order from the previous frame and proceeds to ones with more knots and/or higher order to achieve greater accuracy and smoothness. The initial solid must be deployed manually by the user, and the optimal number of knots in the final solid is unknown.

Denney⁶ proposed a tagline detection method based on maximum likelihood (ML) estimation and maximum a posteriori (MAP) hypothesis testing. The tagline identification problem is formulated as a combined parameter estimation and signal detection problem. Using the image intensity model of a tagline along its perpendicular direction as a priori knowledge, the initial tagline candidates are estimated across a region of interest with a snake algorithm. The tagline centers are detected with an MAP hypothesis test, followed by a pruning algorithm to remove disqualified tag segments according to a spatio-temporal continuity criterion.

Further author information: (Send correspondence to Xiaohui Yuan)

E-mail: xyuan@cse.unt.edu, Telephone: 1 940 565 4256, Post mail: P.O. Box 311366, Denton, Texas 76203-1366

Alternatively, the Gabor filter bank demonstrates better frequency selection and enclosure properties.^{3,4,15} The advantages are achieved by employing multiple Gabor filters that span the spectra of the harmonic peak. By varying the size and the orientation of the Gabor filters, the frequency components of taglines are extracted and reconstructed. Qian et al.³ developed a tagline segmentation method based on Gabor filter bank. A tunable set of Gabor filters is selected for an image according to the filter responses. The taglines are thresholded from the merged filter response map. This method requires that a set of filters be constructed for each individual image, and hence is computational expensive. Another challenges lies in the selection of the appropriate filters. The filter set needs to be tuned for taglines that have various curvatures and the existing filter selection methods are subjective and computationally expensive.

Osman et al.⁹ developed a method using a harmonic phase (HARP) image to analyze 2-D strain in tissue. It applies a band pass filter and the first harmonic spectral peak is used to reconstruct tagline image. The sensitivity of the filter limits the application of this method. When using the smaller filters, the output phase image contains primarily the tagline information but the reconstruction consists of mostly straight lines that hardly follow the curving taglines; whereas, when larger filter is applied, the preservation of the curvature is improved. Yet, taglines are obscured due to the interference induced by anatomical edges.

To improve the robustness and accuracy, Yuan et al¹⁶ proposed a pseudo wavelet reconstruction-based tagline detection method. This method maps an image into wavelet domain. The wavelet coefficients are selectively retained or suppressed to minimize the reconstruction of non-tagline information. The advantages of this method is two fold. First, without assumption on tagline model, this method can be applied to various kinds of tMRI images. Second, it improves the robustness and accuracy of the tagline detection.

To reconstruct tissue strain with extracted taglines, two sets of tagged images are usually acquired. One set, referred to as reference images, is acquired at the relaxation stage, where taglines are mostly straight. The other set, referred to as task images, is acquired upon the maximum muscle engagement, and hence the taglines reach the extreme of deformation. Corresponding these two sets of taglines and computing the differences allow a tissue strain map to be constructed. The difficulties of tagline extraction in tagged tongue MR image arise from the subject variability in the task performance and the complex anatomical structure in tongue. Intra-subject task response variability in timing lead to faded taglines and the presence of different tissues (muscular, adipose, connective) and water in the tongue as well as air surrounding the tongue in the oral cavity can degrade the artificially inserted taglines. In addition, the limitations of imaging procedure and task requirement, noise and artifacts are unavoidable in the tagged MR images. A tagline may be broken into segments, which creates issues in matching reference taglines and task taglines and hence result erroneous strain map. Little, however, was discussed in the literature and manual correction has usually been practiced to touch up the extracted results.

The work presented in this paper focus on extracting and recovering taglines distorted by noise and artifacts, and the recovered taglines are used to synthesize tissue strain map. Our method is based on an active contour model (or snake) that is deployed to taglines segmented with pseudo wavelet reconstructed-based method. The control points on our model is described with an internal force and two external forces. Upon a balance is achieved, segments are linked with interpolation.

The remainder of this paper is organized as follows: Section 2 reviews the pseudo wavelet reconstruction-based tagline detection algorithm and presents our snake-based tagline recovery and indexing method. Section 3 presents our method for strain map construction using extracted taglines. Section 4 concludes this article.

2. TAGLINE EXTRACTION

2.1. Pseudo Wavelet Reconstruction-based tagline extraction

The pseudo wavelet reconstruction-based tagline detection method first decomposes images into wavelet coefficients. Let $W_{2^j}^{xy}I(x, y)$ denote a wavelet subband on decomposition level j , where the super script of W denotes the directions of gradient filters and $I(x, y)$ denotes the image. The subbands that retain the horizontal features are the results of applying vertical gradient filters. Hence the extraction of horizontal taglines can be based on the reconstruction with subbands $W_{2^j}^{yy}I(x, y)$ only. Likewise, to extract vertical taglines, $W_{2^j}^{xx}I(x, y)$ consists of mostly vertical features. In addition, intra-scale spatial correlations exist between wavelet coefficients,^{8,17} which are in the form of anatomical structure information with a slightly translation displacement in the 2D wavelet

decomposition. By performing wavelet reconstruction using the decorrelated subbands that contains tagline features, we can extract taglines from MR images.

Let $\lambda, \lambda \in \{H, V, D\}$, denote the subband preference for the pseudo wavelet reconstruction, where H, V , and D denote the subbands with horizontal, vertical, and diagonal features respectively. Given the choice of λ , the incompliant wavelet subbands are replaced with zero matrices and the smoothed subband is substituted with the scaled unity matrix based on the overall average value. The pseudo wavelet reconstruction of the image is formulated as follows:

$$I'(x, y) = \bar{I}U + \sum_{j=1}^J (\omega W_{2^j}^\lambda I) * \chi_{2^j}^\lambda(x, y) \quad (1)$$

where W^λ is the subband that preserves tagline features and ω is the decorrelation function.

$$\omega = \frac{1}{1 + d_{A,B} + |d_{\lambda,A} - d_{\lambda,B}|} \quad (2)$$

where $A, B \in \{H, V, D\}$ and $A \neq B \neq \lambda$. Function $d_{p,q}(x, y) = |\tilde{W}_{2^j}^p(x, y) - \tilde{W}_{2^j}^q(x, y)|$ computes the absolute difference of normalized subband

$$\tilde{W}_{2^j} = \frac{W_{2^j}}{\max |W_{2^j}|} \quad (3)$$

Figure 1 (b) illustrates a result of the pseudo wavelet reconstruction of a tagged MR image with horizontal taglines.

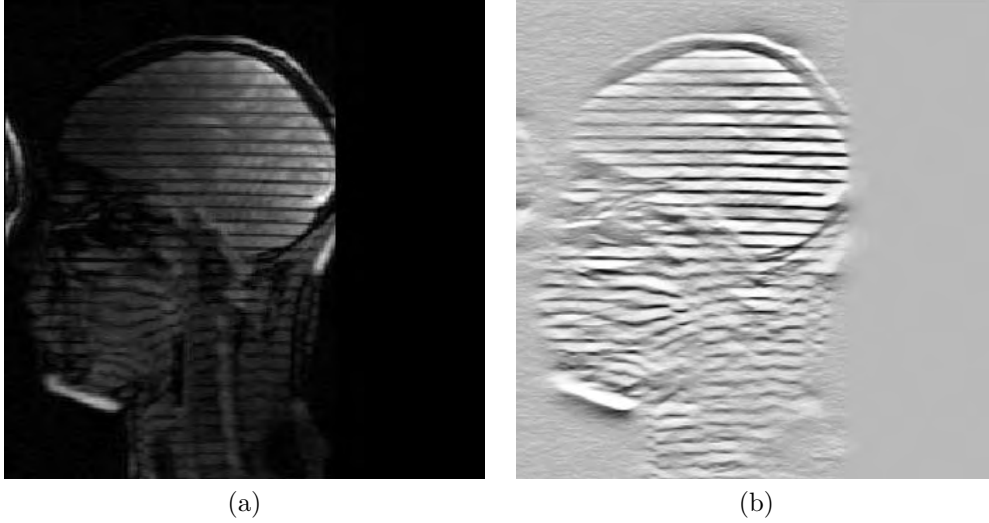


Figure 1. Pseudo wavelet reconstruction of tagline image. (a) is the original image, (b) is the result image with background mostly removed and the taglines enhanced.

As can be seen from Figure 1(b), the dominate color of the reconstructed tagline image is gray and taglines in black. A histogram of the graylevel is shown in Figure 2. A properly selected threshold, denoted by τ , can separate the tagline pixels from the rest. Figure 1 (b) illustrates a result of the pseudo wavelet reconstruction of a tagged MR image with horizontal taglines.

Assume the tagline pixels have a variance of σ_B , and the image has a variance of σ_T . An optimal threshold can be selected by maximizing the separability criterion as follows

$$\tau = \arg \max_{\tau} \frac{\sigma_B^2}{\sigma_T^2} \quad (4)$$

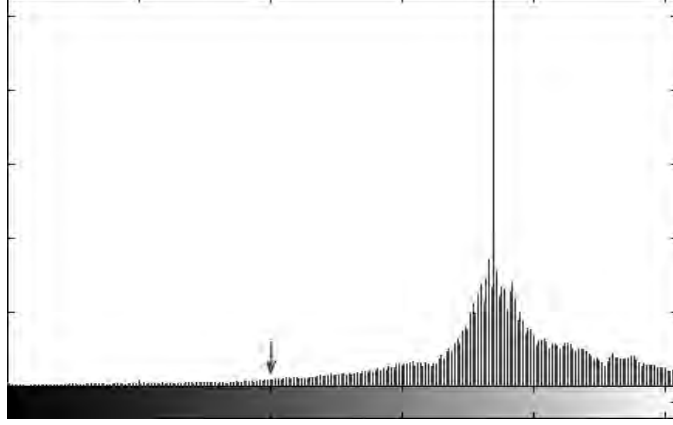


Figure 2. Histogram of a reconstructed tagline image. The color strip shows the grayscale value of each bar. The arrow points at the ideal threshold that separate the tagline pixels from the background.

Since the image variance is functionally independent from τ , the threshold optimization is essentially to maximize the between-class variance.

$$\tau = \arg \max_{\tau} \sigma_B^2 \quad (5)$$

2.2. Tagline Recovery and Indexing with Snakes

In our study, taglines are continuous within the boundary of the anatomical structure. Also, to compute the displacement, the center of the thresholded tagline is usually used, or an erosion is applied to make the width of a tagline to one pixel. In our method, we used projected histogram to cluster tagline pixels and eliminate possible noisy or erroneous pixels. Then we employed a snake-based searching scheme to recover the connectivity and to index taglines.

Let $q(x, y)$ denote a pixel of a tagline on a binary image Q . Assume that tagline pixels in Q have gray value of 1, i.e. $q(x, y) = 1$. By constraining the neighborhood in a window, we can group them into clusters. That is, $\forall q(x, y) \in Q$ can be grouped into M clusters

$$C(j) = \{q(x, y), \exists q(x', y') \in C(j) \text{ and } \|q(x', y') - q(x, y)\| < t\} \quad (6)$$

where $\|\cdot\|$ denote the Euclidean distance, j is the cluster index, and t is the distance threshold.

By projecting all the tagline pixels onto the axis that is perpendicular to the tagline direction, we generate a histogram that encloses the tagline distance. An example is shown in Figure 3.

To avoid erroneous tagline elimination, a regrouping method precedes the noise removal. Let d denote the distance between any two clusters, $C(n)$ and $C(m)$,

$$d_{m,n} = |c(n) - c(m)| \quad (7)$$

where $c(\cdot)$ is the mean of cluster $C(\cdot)$. Given a cluster $C(n)$, its immediate neighbor clusters are denoted by $C(n-1)$ and $C(n+1)$. Cluster $C(n)$ is joined with $C(n+1)$ when it satisfies

$$\begin{cases} d_{n,n+1} < d^* \\ d_{n,n+1} < d_{n-1,n} \end{cases} \quad (8)$$

The threshold d^* is chosen to be one third of the tagline spacing, which can be approximated with the medium tagline cluster distance or obtained from the MR pulse sequence. A median filter is used to remove isolated segments and a skeletonization process follows.

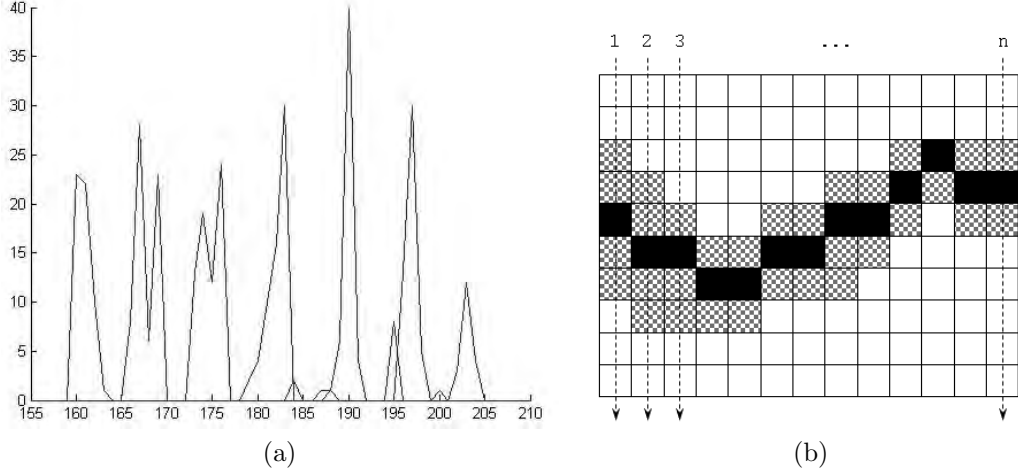


Figure 3. (a) histogram of projected tagline to an axis that is perpendicular to the taglines. Each peak is a group of connected tagline pixels. Tagline pixels are separated into different groups if they are not connected even though they may appear overlap in the histogram. (b) Tagline erosion. The dark pixels are the ones retained in the outcome. The arrow and the number marks the processing order.

At this point, tagline pixels are grouped into clusters. Each cluster represents one line, which corresponds to a peak in the histogram shown in Figure 3. Still, we are facing two issues: first, the index of tagline groups is assigned by tracing pixels column by column from top to bottom. Pixels belonging to one line can be assigned different index if this line is broken somewhere but with significant number of tagline pixels in each group. Second, the indexes in the task image and the reference image may be different, which makes computing point-to-point difference difficult. Re-indexing taglines will register taglines in two sets of images and recover broken ones if exist.

Our method is derived from the snake model that has been widely applied to image segmentation.^{2,7,11-14} Differed from the conventional snakes, the snake in our method represents an open contour. Therefore, each of the internal points has two elastic forces from immediate neighbors and external forces; whereas two end points only have one elastic force from its neighbor. However, we inherit the energy function representation as follows:

$$E(L) = \int_0^L g_b(|\nabla f(L(s))|) ds \quad (9)$$

where ds is the Euclidean element of the length, L is the length of the contour and the function g_b is an edge indicator function that vanishes at object boundaries such as $g_b(|\nabla f|) = \frac{1}{1+\beta|\nabla f|^2}$, and f is the image. Hence, the energy function 9 is actually a new length obtained by weighting the Euclidean elements of length ds by function g_b which contains information of the object boundaries.

Since our tagline images are thresholded binary images, the speed map is generated by mapping the tagline pixels into zeros, and assign ones for the rest of pixels. If a tagline pixel is reached, the control point will be stalled regardless of the forces it bears, i.e. $E(s)$ is forced to be zero. In practice, internal (elastic) and external forces are put upon every control points. Given a location s , the energy is written as follows:

$$E(s) = E_{ext} + E_{int}(s, s-1) + E_{int}(s, s+1) \quad (10)$$

The first component in Equation 10 is the total energy of control point s . The second component is an external force that is enforced equally everywhere in the field and its direction remains the same as always. In our implementation, we initialize our snake on the top row of the image, so all the forces point south are assumed to be positive; and negative forces for the opposite direction. E_{ext} drives all control points downward until they reach a tagline pixel or a balance is formed by combining all forces. The last two terms are the internal forces.

These internal forces exist only on adjacent control points. Here, we simulate the elastic force, i.e., the internal forces are function of distance and the distance between two points can not exceed a limit, denoted by ϕ .

$$E_{int}(s, s') = \text{sgn}(s' - s)\gamma^{|s' - s|} \quad \text{and } |s' - s| < \phi \quad (11)$$

where $\text{sgn}(s' - s)$ is the sign function, γ is a constant that determines the acceleration of the magnitude of the force.

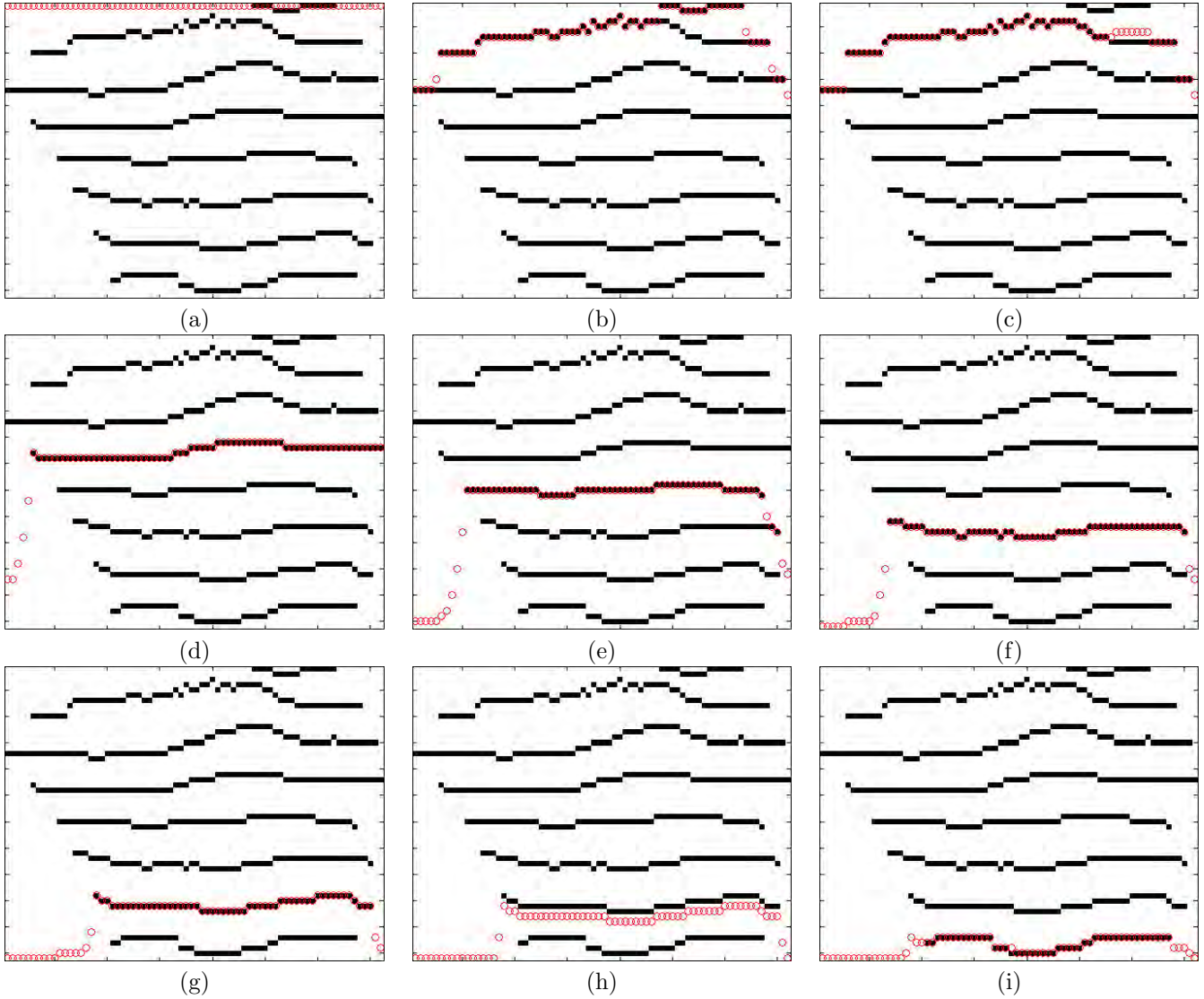


Figure 4. Snapshots of snake algorithm. The black dots are the tagline pixels. The red circles denote the control points on the snake. Images were cropped according to the contour of the tongue. (a) is the initialization of the snake. (b)–(i) are eight images taken during snake evolving.

Figure 4 illustrates snapshots of a running snake. Figure 4(a) shows the initialization. Control points land on the same altitude. A few points are in the tagline pixels with zero speed. So once the program starts running, those points remain in their location, as shown in Figure 4(b).

A problem occurs when the top lines are shorter than the lines underneath. Control points of the snake land on more than one group of tagline pixels. To avoid the incorrect indexing, we cross-verify the current index

with the original label assigned using Equation 7. The primary index is given to the first group label seen by the snake. Any other labels are considered inconsistent. Once the snake stabilized, all the control points with inconsistent labels will be lifted up one pixel. Hence, they land on a non-tagline location. In the next iteration, they can continue explore and will be stopped by the same tagline pixel. Unless their label becomes the primary index, they are lifted. An example is shown in Figure 4(c). Although the second line (from top) has been reached, all the control point revisit that line in the second iteration.

The iteration referred here is defined by a start of moving control point to a balance is reached. At such a balance, the energy function 10 is minimized. To start the next iteration, we need to “re-initialize” the snake. Instead of putting all points in a straight line, we increase their step by one. That is, if they are stalled by taglines, increasing by one put them into a non-tag region so that the summation of forces take over the movement. Figure 4(g) and (h) demonstrate this process. Because of the boundary condition, the control points that reach the bottom remain the location.

Due to quantization error, there are cases that summation of forces is not zero and the control point is bouncing back and forth infinitely trying to find a balancing location. To solve this problem, we gradually reduce the activity of all points. The energy function becomes

$$E(s) = (E_{ext} + E_{int}(s, s - 1) + E_{int}(s, s + 1))(e^{-t} - 1) \quad (12)$$

where t denote the time. Allowing enough time, the energy of a control point reduces to zero.

Each iteration, the primary index of the line is saved and the re-index is performed after all lines are found. The broken lines are identified by examining the labels of two groups of tagline pixels. Groups are joined when there is no overlapped tagline pixels given the index of control points.

The same procedure is applied to the reference tagline images. However, in our experiments, the reference taglines are less affected by the anatomical structures and hence have fewer distortion. The snake method works as to match the index numbers of the reference taglines and the task taglines.

3. TONGUE STRAIN MAP CONSTRUCTION

3.1. Image Acquisition

In our image acquisition, subjects performed gated, repetitive effortful swallowing. For each task response, a 2-D slice was acquired with 1-D SPAMM¹ planar tagging to capture tissue motion at the point of maximum voluntary isometric contraction. We acquired six sets of tagged images for each subject, including sagittal contiguous images with horizontal taglines and axial contiguous images with horizontal and vertical taglines. Three sets of images are the reference images with straight tagline acquired at tongue resting posture; the other three sets are the task images with tagline inserted before tasks are performed and imaged at the peak of performance.

3.2. Computing Strain Maps

By registering and merging the taglines detected in the reference image set and task image set, we formed two 3D grids. We assume the 3D intersection of the grid marks the change in space for the reference points. Linear interpolation was performed to fill the gap between two parallel taglines such that more intersection points were created. Let $[X, Y, Z]^T$ denote a point in the reference grid. it relocates to $[x, y, z]^T$ in the task grid. The displacement field can be built in the Cartesian coordinates as follow:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} u_x(X, Y, Z) \\ u_y(X, Y, Z) \\ u_z(X, Y, Z) \end{bmatrix} \quad (13)$$

where u denotes the displacement field. By converting it into cylindrical coordinates, we have

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} R \cos \theta + u_R \cos \theta - u_\theta \sin \theta \\ R \cos \theta + u_R \cos \theta - u_\theta \sin \theta \\ Z + u_z \end{bmatrix} \quad (14)$$

where $u_R = u_R(R, \theta, Z)$, $u_\theta = u_\theta(R, \theta, Z)$, and $u_Z = u_Z(R, \theta, Z)$ denote the displacement field by the intersection points. The deformation gradient tensor is hence

$$F = \begin{bmatrix} \frac{\partial x}{\partial X} & \frac{\partial x}{\partial Y} & \frac{\partial x}{\partial Z} \\ \frac{\partial y}{\partial X} & \frac{\partial y}{\partial Y} & \frac{\partial y}{\partial Z} \\ \frac{\partial z}{\partial X} & \frac{\partial z}{\partial Y} & \frac{\partial z}{\partial Z} \end{bmatrix} \quad (15)$$

Figure 5 illustrates the sagittal view of the strain map of the 3D volume. The volume was re-sliced into seven slices arranged from lateral to mid-sagittal to lateral. The color depicts the magnitude of the strain. It is shown that the based of the tongue contributes a lot more than center of the tongue. Also the back of the tongue deformed much during the task performance.

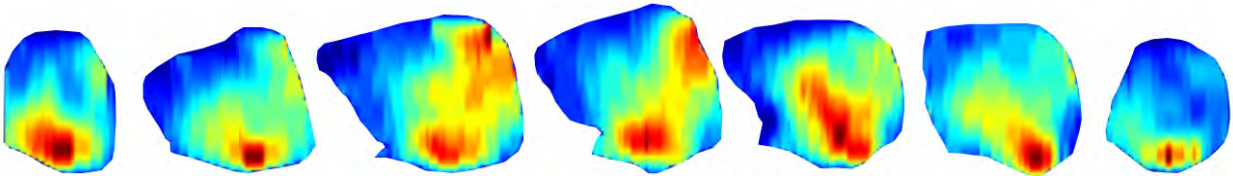


Figure 5. Re-sliced sagittal views of the tongue strain map. Color depicts the magnitude of the strain. From dark red to dark blue, the magnitude decreases.

4. CONCLUSIONS AND FUTURE WORK

In this article, we extend our pseudo wavelet reconstruction-based tagline detection method and improved it by integrating an active contour-based method for tagline recovery and indexing. Our method presents an automatic procedure that requires little user interaction. It has been applied to tagged MR tongue images with a mixture of taglines. The detected taglines were then assembled into 3D grid and strain maps of tongue were reconstructed.

Although 3D strain maps have been computed, visualization is a challenge. Methods reported in literature use vectors to show both direction and magnitude of the strain. In our study, tongue has regions expended and regions compressed. The stain value can therefore be positive or negative. Aggregating this fourth dimension is very important and useful for clinician to understand the synergistic function of tongue muscles. Our next step will focus on 4D visualization of tongue strain.

Acknowledgments

The Physical Disabilities Branch is collaboration between the National Institute of Child Health and Human Development and the Clinical Center, National Institutes of Health. The opinions presented in this report reflect the views of the authors and not those of the National Institutes of Health or the US Public Health Service.

REFERENCES

1. L. Axel and L. Dougherty. MR imaging of motion with spatial modulation of magnetization. *Radiology*, 171:841–845, 1989.
2. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *International journal of Computer Vision*, 22(1):61–79, 1997.
3. Y. Chen and A. A. Amini. A MAP framework for tag line detection in SPAMM data using markov random fields on the B-spline solid. *IEEE Transactions on Medical Imaging*, 21(9):1110–1122, 2002.
4. D. Dunn, W. E. Higgins, and J. Wakeley. Texture segmentation using 2-d gabor elementary functions. *IEEE Transactions on Pattern Recognition and Machine Analysis*, 16(2):130–149, 1994.
5. M. A. Guttman, J. L. Prince, and E. R. McVeigh. Tag and contour detection in tagged MR images of the left ventricle. *IEEE Transactions on Medical Imaging*, 13(1):74–88, March 1994.

6. T. S. Denney Jr. Estimation and detection of myocardial tags in MR image without user-defined myocardial contours. *IEEE Transactions on Medical Imaging*, 18(4):330–344, 1999.
7. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, pages 321–331, 1987.
8. J. Liu and P. Moulin. Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients. *IEEE Transactions on Medical Imaging*, 10(11):1647–1658, 2001.
9. N. F. Osman, E. R. McVeigh, and J. L. Prince. Imaging heart motion using harmonic phase mri. *IEEE Transactions on Medical Imaging*, 19(3):186–202, 2000.
10. Z. Qian, A. Montillo, D. Metaxas, and L. Axel. Segmenting cardiac MRI tagging lines using gabor filter banks. In *Proceeding of the 25th Annual International Conference of the IEEE EMBS*, pages 630–633, Philadelphia, Pennsylvania, October 2003.
11. C. Sagiv, N. A. Sochen, and Y. Y. Zeevi. Integratd active contour for texture segmentation. *IEEE Transactions on Image Processing*, 15(6):1633–1646, 2006.
12. Y. Xiang, A. C.S. Chung, and J. Ye. An active contour model for image segmentation based on elastic interaction. *Journal of Computational Physics*, 219:455–476, 2006.
13. C. Xu and J. L. Prince. Generalized gradient vector flow external forces for active contour. *Signal Processing*, 71:131–139, 1998.
14. C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Transactions on Image Processing*, 7(3):359–369, 1998.
15. J. Yuan, J. Zhang, X. Yuan, and B. P. Buckles. Multi-scale feature identification using evolution strategies. *Image and Vision Computing*, 23(6):555–563, 2005.
16. X. Yuan, C. Ozturk, and G. Chi-Fishman. A pseudo wavelet-based method for accurate tagline tracing on tagged mr images of the tongue. In *Proceeding of SPIE Medical Imaging*, San Diego, CA, February 2006.
17. L. Zhang, P. Bao, and X. Wu. Bybrid inter- and intra-wavelet scale image restoration. *Pattern Recognition*, 36:1737–1746, 2003.