

Detecting Abnormalities in Tympanic Membrane Images

Xianghua Xie^a, Majid Mirmehdi^{a*}, Richard Maw^b, and Amanda Hall^c

^aDepartment of Computer Science, University of Bristol, Bristol BS8 1UB, England

^bDepartment of Otolaryngology, St. Michaels Hospital, Bristol

^cALSPAC, Department of Community Based Medicine, University of Bristol, Bristol, BS8 1TQ

Abstract. We present a method to detect abnormalities in colour tympanic membrane (eardrum) images. The method first uses a geometric snake to localise the regions of interest and normalise the colour pixels based on a reference image. Then, defect detection and localisation in multiscale is performed using a statistical model. Initial experimental results (presented) suggest excellent potential for in-depth progress and validation.

1 Introduction

The tympanic membrane (TM), also known as the eardrum, is a thin membrane that divides the ear canal from the middle ear. The TM vibrates in response to sound, and enables it to be passed along the middle ear bones, through the inner ear and up to the brain (see Figure 1 for an illustration). Abnormalities of the TM include perforations, retractions, otitis media (fluid build up behind the TM) and cholesteatoma (infection within the middle ear cleft). These have various consequences for patients if left untreated, including hearing loss and major infection, and so need to be detected early, particularly in children. Examples of normal and abnormal TM are shown in Figure 1 (note the saturated highlight regions originating from the illumination source). More minor abnormalities of the TM may indicate approaching problems in the future. We have not yet found any work on the detection of TM abnormalities using computer vision techniques.

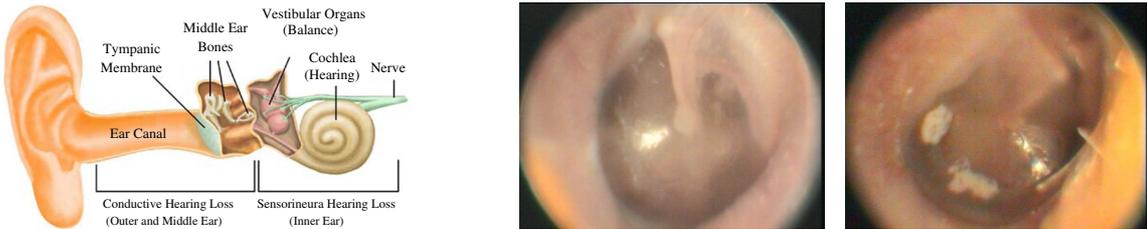


Figure 1. From left: anatomy of ear [1], one normal TM image, and one defective TM image.

In this paper, we present a method to detect TM defects. A geometric snake is first used to localise the regions of interest, presented in section 2. Colour normalisation is then performed based on selected reference TM image, as briefly described in section 3. Finally, defect detection and localisation in multiscale is performed using a statistical model, detailed in section 4. Section 5 presents some initial experimental results.

2 Localising TM using a Geometric GGVF Snake

Deformable contour models or snakes are widely used in image processing and computer vision due to their natural handling of shape variations and independence of operation (once initialised). A hypothesised contour, represented as a curve or a surface, evolves under the influence of internal forces, external image dependent forces, and certain constraints, till it converges on the object(s) of interest. In this paper, we utilise the generalised version of the well-known GVF snake [2]. However, we implement it in a geometric active contour framework using level sets [3] so that the snake enjoys the topological freedom and can be more accurate in representing shapes. This geometric, generalised GVF (GGVF) snake is useful when dealing with boundaries with small gaps.

The snake evolves under internal and external forces, but only the forces in the normal direction of the evolving contours can change the geometry. Thus, a geometric snake can be interpreted as $C_t = [(F_{int} + F_{ext}) \cdot \vec{N}] \vec{N}$, where C is the evolving contour and \vec{N} is the unit inward normal of the contour. The internal force imposes a regularity constraint in the propagation, e.g. the curvature κ . The GGVF field, $\hat{\mathbf{v}}$, then contributes the external force, which is defined as the equilibrium solution of the following vector partial differential equation:

$$\mathbf{v}_t = p(|\nabla f|) \nabla^2 \mathbf{v} - q(|\nabla f|) (\mathbf{v} - \nabla f), \quad (1)$$

where f is the colour edge map, and $p(\cdot)$ and $q(\cdot)$ are monotonically non-increasing and non-decreasing functions respectively, controlling the amount of diffusion. The initial vector field is given by ∇f , i.e. $\mathbf{v}(t = 0) = \nabla f$.

*Email: majid@cs.bris.ac.uk. X. Xie is partly funded by the ORS, UK. The authors wish to thank the ALSPAC Study Team.

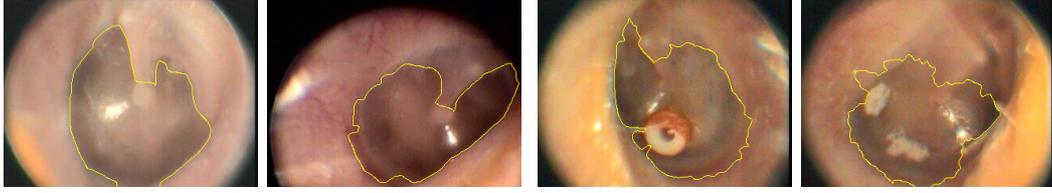


Figure 2. Snake segmentation - from left: two snake segmentation results on normal TM images and two example results on defective TM images.

These two functions are selected such that $p(\cdot)$ gets smaller as $q(\cdot)$ becomes larger with the desired result that in the proximity of large gradients, there will be very little smoothing and the vector field will be nearly equal to the gradient map:

$$p(|\nabla f|) = e^{-(|\nabla f|/K)}, \quad q(|\nabla f|) = 1 - p(|\nabla f|), \quad (2)$$

where K is a constant and acts as a trade-off between field smoothness and gradient conformity. The geometric GGVF snake can be formulated as:

$$C_t = \alpha g(f) \kappa \vec{N} + (1 - \alpha) (\tilde{\mathbf{v}} \cdot \vec{N}) \vec{N}, \quad (3)$$

where α is a constant that balances the contributions between regularity and boundary attraction, and $g(\cdot)$ is a monotonically decreasing function such that $g(f) \rightarrow 0$ as $f \rightarrow \infty$, and $g(f) \rightarrow 1$ as $f \rightarrow 0$. A typical function can be: $g(f) = 1/(1 + f)$. The level set representation of the geometric GGVF snake then takes the form:

$$\phi_t = \alpha g(f) \kappa |\nabla \phi| - (1 - \alpha) \tilde{\mathbf{v}} \cdot \nabla \phi, \quad (4)$$

where ϕ is the level set function with the zero level set corresponding to C . Some snake segmentation results are shown in Figure 2. We discuss possible improvements in section 6.

3 Colour Normalisation

The TM colour can vary substantially from one patient to another, hence, before defect detection, we need to normalise the colour of the TM image under inspection. Here, a simple but effective method, histogram specification [4], is used to modify the values of the TM such that its frequency histogram matches the reference TM distribution. Currently, the highlights in the training samples are removed by simple thresholding to reduce the ambiguity of normal TM regions. However, the highlights in the testing images are left as they are because their shape and distribution can also indicate abnormalities. Also, the highlights may overlap with defective regions. Figure 3 shows colour normalisation examples on a normal and defective TM image.

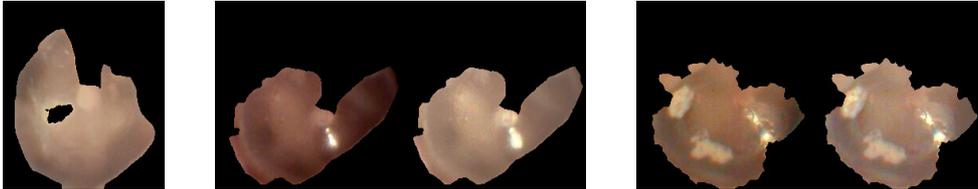


Figure 3. From left: reference TM without highlights, pre and post-normalisation of a normal and a defective TM.

4 Abnormality Detection and Localisation

We apply TM extraction and preprocessing on a small number of normal images as a training stage. By characterising the colour and textural properties of these normal TM region we can perform novelty detection in the testing stage. We consider each TM as a collection of image patches of various sizes, possibly overlapped. We group the patches for a few TM samples into clusters, dependent on the patch size, and describe the clusters using a mixture model. By applying Expectation Maximisation (EM) on the model parameters, we learn *texems* [5], a small set of primitive texture exemplars, encompassing the implicit colour and textural characteristics of the image region. In [5], we developed texems for the detection of defects on ceramic tiles.

4.1 Learning colour texems from normal TM images

The texems are image representations at various sizes that encapsulate the textural or visual primitives of a given image. Each texem, denoted as \mathbf{m} , is defined by a mean, $\boldsymbol{\mu}$, and a corresponding variance matrix, $\boldsymbol{\omega}$, i.e. $\mathbf{m} = \{\boldsymbol{\mu}, \boldsymbol{\omega}\}$. The original TM region \mathbf{T} is broken down into a set of P patches $\mathbf{Z} = \{\mathbf{Z}_i\}_{i=1}^P$, each containing pixels from a subset of image coordinates with a size of $d = N \times N$. The patches may overlap and can be of various sizes. We assume that there exist K texems, $\mathcal{M} = \{\mathbf{m}_k\}_{k=1}^K$, $K \ll P$, for \mathbf{T} such that \mathbf{Z}_i can be generated from \mathbf{m}_k with certain added variations:

$$p(\mathbf{Z}_i|\theta_k) = p(\mathbf{Z}_i|\boldsymbol{\mu}_k, \boldsymbol{\omega}_k) = \prod_{j \in S} \mathcal{G}(\mathbf{Z}_{j,i}; \boldsymbol{\mu}_{j,k}, \boldsymbol{\omega}_{j,k}), \quad (5)$$

where θ_k denotes the k th texem's parameters, $\mathcal{G}(\mathbf{Z}_{j,i}; \boldsymbol{\mu}_{j,k}, \boldsymbol{\omega}_{j,k})$ is a Gaussian distribution with mean $\boldsymbol{\mu}_{j,k}$ and variance $\boldsymbol{\omega}_{j,k}$ at the j th pixel position in the k th texem, and S is the pixel grid. For our mixture model, we assume the following probabilistic model: $p(\mathbf{Z}_i|\Theta) = \sum_{k=1}^K p(\mathbf{Z}_i|\theta_k)\alpha_k$, where the parameters are $\Theta = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_K, \theta_1, \dots, \theta_K)$, and α_k is the *prior* probability of k th texem that is constrained by $\sum_{k=1}^K \alpha_k = 1$. The EM algorithm can be used to find the maximum likelihood estimate of our mixture model parameters. The E-step involves a soft-assignment of each patch \mathbf{Z}_i to texems using Bayes rule:

$$p(\mathbf{m}_k|\mathbf{Z}_i, \Theta^{(t)}) = \frac{p(\mathbf{Z}_i|\mathbf{m}_k, \Theta^{(t)})\alpha_k}{\sum_{k=1}^K p(\mathbf{Z}_i|\mathbf{m}_k, \Theta^{(t)})\alpha_k}, \quad (6)$$

where $\Theta^{(t)}$ denotes intermediate parameters. The M-step then updates the parameters by maximising the log-likelihood. The new estimates are denoted by $\hat{\boldsymbol{\alpha}}_k$, $\hat{\boldsymbol{\mu}}_k$, and $\hat{\boldsymbol{\omega}}_k$ where

$$\begin{aligned} \hat{\boldsymbol{\alpha}}_k &= \frac{1}{P} \sum_{i=1}^P p(\mathbf{m}_k|\mathbf{Z}_i, \Theta^{(t)}), & \hat{\boldsymbol{\mu}}_{j,k} &= \frac{\sum_{i=1}^P \mathbf{Z}_{j,i} p(\mathbf{m}_k|\mathbf{Z}_i, \Theta^{(t)})}{\sum_{i=1}^P p(\mathbf{m}_k|\mathbf{Z}_i, \Theta^{(t)})}, \\ \hat{\boldsymbol{\omega}}_{j,k} &= \frac{\sum_{i=1}^P (\mathbf{Z}_{j,i} - \hat{\boldsymbol{\mu}}_{j,k})(\mathbf{Z}_{j,i} - \hat{\boldsymbol{\mu}}_{j,k})^T p(\mathbf{m}_k|\mathbf{Z}_i, \Theta^{(t)})}{\sum_{i=1}^P p(\mathbf{m}_k|\mathbf{Z}_i, \Theta^{(t)})}. \end{aligned} \quad (7)$$

The E-step and M-step are iterated until the estimations stabilise or the rate of improvement of the likelihood falls below a pre-specified convergence threshold. Then the texem is derived as $\mathbf{m}_k = \{\boldsymbol{\mu}_{j,k}, \boldsymbol{\omega}_{j,k}\}_{j \in S}$.

In order to capture sufficient textural properties, various size of texems are required. However, the computational costs increase rapidly as the size of patches increase. Therefore, instead of generating variable-size texems, we learn fixed size texems in a multiscale image pyramid. This will result in (multiscale) texems with a very small size, e.g. 5×5 . A simple multiscale approach by using a Gaussian pyramid is sufficient.

4.2 Novelty detection and defect localisation

Once the texems are obtained from training images, we can work out the minimum bound of normal samples in each resolution level in order to perform novelty detection. A small set of normal samples (e.g. 4 or 5 only) are arranged within the multiscale framework. The probability of a patch $\mathbf{Z}_i^{(n)}$ belonging to texems in the corresponding n th scale is given by $p(\mathbf{Z}_i^{(n)}|\Theta^{(n)})$ ¹. We then define a novelty score function as the negative log likelihood:

$$\mathcal{V}(\mathbf{Z}_i^{(n)}|\Theta^{(n)}) = -\log p(\mathbf{Z}_i^{(n)}|\Theta^{(n)}). \quad (8)$$

The lower the novelty score, the more likely the patch is normal and vice versa. The distribution in this novelty score space is not necessarily a simple Gaussian distribution. In order to accurately find out the lower bound of the data likelihood, K-means clustering is performed in the novelty score space. The cluster with maximum mean is treated as the *marginal component* (MC), which is characterised by mean $u^{(n)}$ and standard deviation $\sigma^{(n)}$. The maximum novelty score, $\Lambda^{(n)}$, of a patch $\mathbf{Z}_i^{(n)}$ at level n across the training images is then established as: $\Lambda^{(n)} = u^{(n)} + \lambda\sigma^{(n)}$, where λ is a constant, usually selected as 2 or 3. This completes the training stage in which we determine the texems and an automatic threshold for marking new patches as good or defective.

In the testing stage, the image under inspection is again layered into a multiscale framework and patches are examined against the texems. The probability for each patch and its novelty score are then calculated according to (8) and compared to the minimum data likelihood, determined by $\Lambda^{(n)}$, at the corresponding level. Let $Q^{(n)}(x)$ be the novelty score map at the n th resolution level. Then, the potential defect map, $\mathcal{D}^{(n)}(x)$, at level n is:

$$\mathcal{D}^{(n)}(x) = \begin{cases} 0 & \text{if } Q^{(n)}(x) \leq \Lambda^{(n)} \\ Q^{(n)}(x) - \Lambda^{(n)} & \text{otherwise.} \end{cases} \quad (9)$$

We then need to combine the information coming from all the resolution levels to build the certainty of the defect at position (x) . We follow a method described in [6] which combines information from different levels of a multiscale pyramid and reduces false alarms. It assumes that a defect must appear in at least two adjacent resolution levels for it to be certified as such. Using a logical AND, implemented through the geometric mean, of every pair of adjacent levels, resulting in combined maps $\mathcal{D}^{(1,2)}(x)$, $\mathcal{D}^{(2,3)}(x)$, ..., $\mathcal{D}^{(l-1,l)}(x)$, and a logical OR, as the arithmetic mean, we obtain a final map for the defects detected across all the scales:

$$\mathcal{D}(x) = \frac{1}{l-1} \sum_{n=1}^{l-1} \mathcal{D}^{(n,n+1)}(x) = \frac{1}{l-1} \sum_{n=1}^{l-1} [\mathcal{D}^{(n)}(x)\mathcal{D}^{(n+1)}(x)]^{1/2}, \quad (10)$$

where $\mathcal{D}(x)$ contains the joint contribution of all the resolution scales and marks the defects in the test image.

¹Hereafter, the superscription (n) denotes parameters or data at n th scale.

5 Experimental Results

Figure 4 shows a TM example with small areas of Tympanosclerosis and a perforation, and its colour normalisation after snake segmentation. The potentially defective regions detected at each resolution level n , $n = 1, \dots, 4$, are also shown. It can be seen that the texems register good sensitivity at the defective region at different scales. As the resolution progresses from coarse to fine, additional evidence for the defective region is gathered. This evidence is then combined to produce the defect map \mathcal{D} . The final image shows the superimposed defects. As mentioned earlier, the defect fusion process can eliminate false alarms, e.g. see the extraneous false defect regions in level $n = 3$ which disappear after (10).

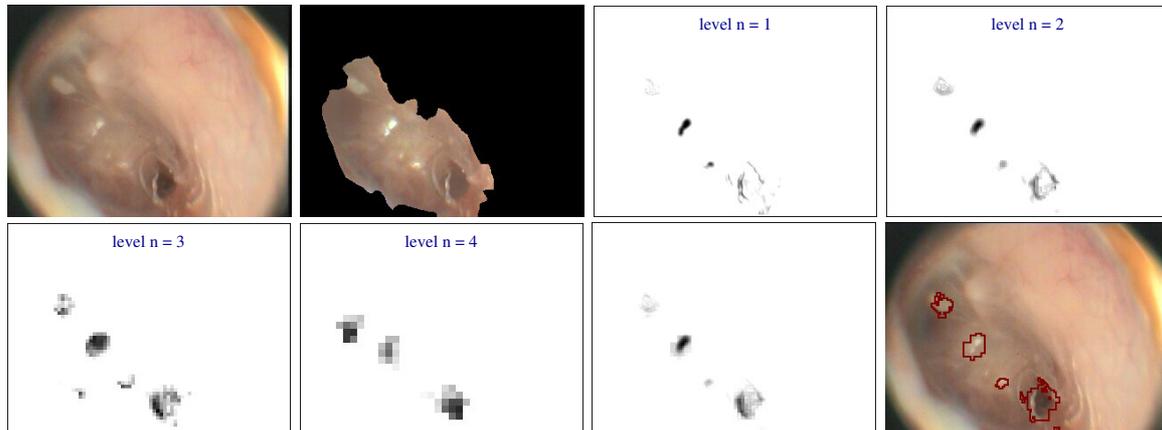


Figure 4. Defect localisation - from top left: initial TM image, after segmentation and colour normalisation, defect probability maps at all 4 scales, fused defect map, and detected defects (and saturated highlights).

Two more examples are shown in Figure 5. The left example shows our method successfully detect and localise more defects. In the other TM image, the defect is much more subtle. Although it does not perfectly localise the defective regions, the method does roughly indicate the regions with abnormalities.

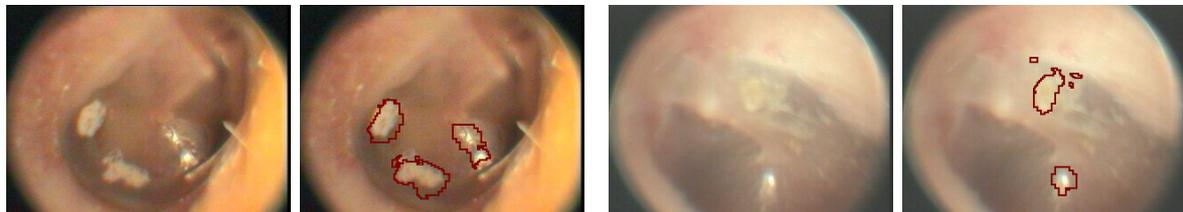


Figure 5. Defect localisation examples - see text for details.

6 Discussion

The results demonstrate the proposed method is a promising approach to detect abnormalities in colour TM images. Clearly, a systematic evaluation against an expert's handlabelled data set is necessary and will be undertaken as part of our future work. Furthermore, we intend to investigate an automatic way of initialising the snake and improving its segmentation by producing a better external force field and incorporating shape prior knowledge, e.g. as in [7]. We will also investigate the possibility of utilising the distribution and characteristics of highlights in the aid of defect detection as those of abnormal TMs' are usually characteristically scattered and skewed.

References

1. H. Rehm, R. Williamson, M. Kenna et al. "Understanding the genetics of deafness." <http://hearing.harvard.edu/info/GeneticDeafnessBookletV2.pdf>, 2003.
2. C. Xu & J. Prince. "Generalized gradient vector flow external forces for active contours." *Sig. Proc.* **71(2)**, pp. 131–139, 1998.
3. J. Sethian. *Level Set Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Materials Science*. CUP, 1996.
4. R. Gonzalez & R. Woods. *Digital Image Processing*. Addison Wesley, 1992.
5. X. Xie & M. Mirmehdi. "Multiscale random colour texture novelty detection using texture exemplars.", 2005. Submitted to BMVC.
6. J. Escofet, R. Navarro, M. Millán et al. "Detection of local defects in textile webs using Gabor filters." *Opt. Eng.* **37(8)**, pp. 2297–2307, 1998.
7. Y. Chen, H. Tagare, S. Thiruvankadam et al. "Using prior shapes in geometric active contours in a variational framework." *IJCV* **50(3)**, pp. 315–328, 2002.