# A Hybrid Feature-Based Segmentation and Classification System for the Computer Aided Self-Diagnosis of Otitis Media

Chuen-Kai Shie, Hao-Ting Chang, Fu-Cheng Fan, Chung-Jung Chen,
Te-Yung Fang and Pa-Chun Wang

*Abstract*— **We propose a novel hybrid otitis media (OM) computer aided detection (CAD) system, designed to aid in the self-diagnosis of various forms of OM. OM is a prevalent disease in both children and adults. Our system is able to differentiate normal ear from acute otitis media (AOM), otitis media with effusion (OME) and the multi-categories of chronic otitis media including perforation, retraction, cholesteatoma, etc. We propose a modified double active contour segmentation method designed for use with otoscope images, and enabled to handle user acquired data. To describe the visual symptoms (e.g., red, bulging, effusion, perforation, retraction, etc.) of otitis media accurately, we extract color, geometric and texture features by grid color moment, Gabor filter, local binary pattern and histogram of oriented gradients. A powerful classification structure based on Adaboost is used to select the most useful features and build a strong classifier. Our system achieves classification accuracy as high as 88.06% and is suitable for real use. In addition, some interesting observations about OM otoscope images are also discussed.**

## I. INTRODUCTION

Otitis media (OM) is any inflammation of the middle ear and can be classified clinically as *acute otitis media* (AOM), *otitis media with effusion* (OME) and *chronic otitis media* (COM). AOM is caused by fluid of the middle ear having an active bacterial infection. OME is defined as inflammation of the middle ear with a collection of fluid that occurs within the middle ear space. COM involves a perforation (hole) in the tympanic membrane along with an active bacterial infection within the middle ear space. Fig. 1. shows a sample image of a normal eardrum and along with the three main OM diagnostic categories.

The goal of this paper is to construct a set of complete image processing algorithms for automated self-diagnosis of OM. AOM is the most prevalent childhood infection, and also one of the most frequent causes of visits to the pediatrician. It is both a costly and significant social burden and leads to many indirect losses every year [1] [2]. Symptoms of AOM are similar to the common cold, and are often difficult for children to effectively describe ear related trauma. If a personal otoscope is available, preliminary diagnosis would be able to be performed at home and medical attention can be sought where appropriate. In addition, OME usually develops after AOM and its symptoms are not obvious to be found out. Parents will be able more effectively analyze possible OME by this device. COM shows a higher prevalence amongst adults. There are various causes and symptoms of COM, making diagnosis more complicated. Moreover, for regions with lacking medical resources, a handy and highly accurate OM CAD system is needed.

In this paper, we have cooperated with seven otolaryngologists of Cathay General Hospital, for the clinical test and execution of our algorithm by 865 expertly diagnosed pictures. The experimental results show that our algorithms achieve high accuracy.

(a) normal eardrum (b) AOM (c) OME (d) COM

Figure 1. Sample (cropped) images from normal ear and the three diagnostic categories of otitis media.

## II. RELATED WORK

Although biomedical image processing has been on the rise in recent decades, there are still only a few researches working on image processing of otitis media. Mironică et al. [3] compare different color descriptors and classification algorithms with color coherence histograms as color descriptor to get the best accuracy of 73.11%. They have concluded that color information alone is not sufficient for accurate classification. Kuruvilla et al. [4] applied "vocabulary" to describe the symptoms which are similar to the terms used by otolaryngologists. They built a tree structure "grammar" to simulate the decision process used by otolaryngologists and then achieved an accuracy rate as high as 89.9%. However, their algorithm only diagnoses AOM and OME, but the more complicated COM is not one of their diagnostic categories. Moreover, it is not necessary to diagnose whether the OM is AOM, OME or COM for OM detection system, since the final diagnosis and treatment still be performed by doctors. OM detection system is responsible for detecting OM so that patients can further take appropriate management.

## III. METHODS

The algorithm has been built using a series of techniques as shown in Fig. 2. First, the acquisition of an image from an

C.K. Shie, H.T. Chang, F.C. Fan and C.J. Chen are with the HTC corporation, New Taipei City Taiwan (e-mail: Chuenkai_Shie, Howard.HT_Chang, Frank.FC_Fan, Jones.CJ_Chen,@ htc.com).

T.Y. Fang is with Department of Otolaryngology, Cathay General Hospital, Taipei, Taiwan (e-mail: yung_11_1@yahoo.com.tw)

P.C. Wang is with Department of Otolaryngology, Cathay General Hospital, Taipei, Taiwan and Department of Public Health, China Medical University, Taichung, Taiwan (e-mail: drtony@seed.net.tw)

978-1-4244-7929-0/14/$26.00 ©2014 IEEE

otoscope. Since the classifier works on the eardrum region, it must be segmented from the source image. We propose a modified double active contour method as our segmentation algorithm, specially designed for otoscope images and described in Section Ⅲ-A. Different diagnostic categories have different key symptoms: redness and bulging are key features of AOM, amber coloring is key for OME and perforation of the tympanic membrane is common for COM. Therefore, we describe those symptoms by color, geometric and texture features. The detailed feature extraction methods are discussed in Section Ⅲ-B. Among those features, some of them play the more important roles in classification, and therefore we do feature selection of key features and use them to build a strong OM classifier by Adaboost [13], discussed in Section Ⅲ-C.

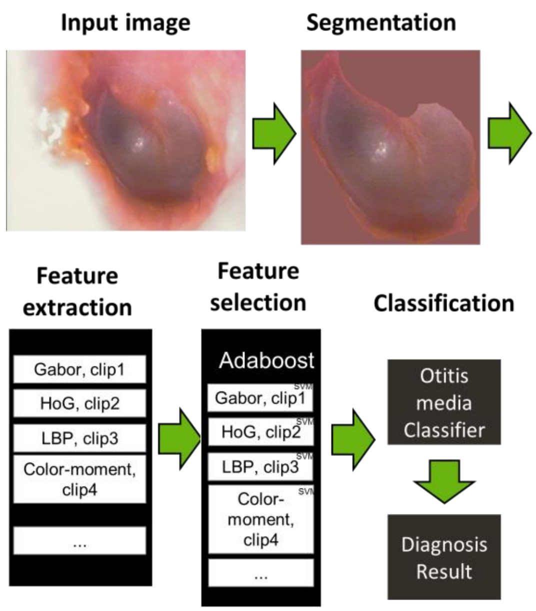



Figure 2. The flowchart of our OM detection system.

## A. Segmentation

Since the visual cues of OM only appear on the tympanic membrane, we need to separate the tympanic membrane from the input otoscope image. However, otoscope images can vary with different instrument and usage. The color and position of the tympanic membrane are not reliable from images captured this way. Therefore, we apply a powerful "active contours" [5] segmentation approach which minimizes an energy function to evolve the active contour and finally terminate on the desired boundary condition. There are two obstacles for accurate performance of this method. First, the algorithm needs an initial curve position and it is difficult to obtain images where the eardrum is centrally located. Second, reflected light makes the color of the ear canal close to bright white, and inadvertently forces the curves of the active contour algorithm to evolve to the incorrect position. Therefore, we have designed a two iterations special active contour algorithm to rectify the problems. The first iteration is to remove the bright ear canal region (if there is no bright ear canal in input image, the first iteration will not work.) The second is to segment the desired eardrum region close to a circle or ellipse in shape. Fig. 3. illustrates our segmentation algorithm. In Fig. 3. (b), the first iteration crops the bright ear canal region. In Fig. 3. (c), the second iteration evolves the curve on the remaining region and segments the desired eardrum. Therefore, it has a great advantage for our method that we can segment the eardrum not at picture center, and do not need user to specify eardrum location. Fig. 4. shows the examples.

(a) input image (b) 1st iteration (c) 2nd iteration

Figure 3. The two iterations of the active contour algorithm for eardrum segmentation.

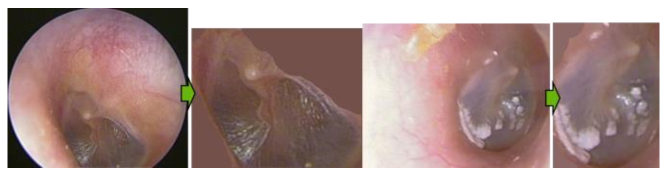


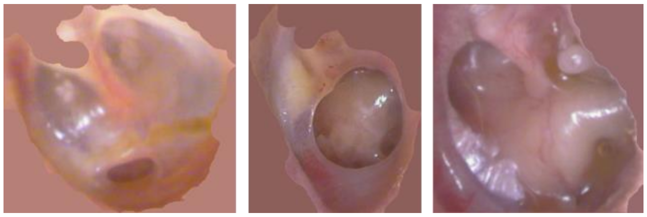
Figure 4. Our method can segment an eardrum not centered locally, and does not require user guidance.

## B. Feature Extraction

Color is one of the most important visual cues for detecting AOM and OME. The color of the normal tympanic membrane is gray and often has varying degrees of translucency. An infected TM becomes red and mostly opaque when an infection is present. OME is caused by the fluid under tympanic membrane, which lends to its opaque and amber appearance. Grid color moment (GCM) [6] is used as our color descriptor. Firstly converting the input image from RGB to HSV color space and dividing the whole image into many grids. Three measures are calculated as color features of the input image: mean, standard deviation and skewness.

The other important visual cues are geometric structures on the tympanic membrane. COM is more complicated and can be categorized into several subcategories: size of the hole of the perforated middle ear area and retraction which is observed in the structure and shape of the tympanic membrane. Fig. 5. illustrates the above three COM symptoms. We apply grids of Histograms of Oriented Gradient (HOG) [7] to represent these geometric features. HOG is performed by dividing the image into small connected regions, named cells, and for each cell compiling a histogram of gradient directions or edge orientations for the pixels within the cell. The combination of these histograms then represents the HOG descriptor. Moreover, we want to describe the texture of tympanic membrane by using local binary pattern (LBP) [8] [9] and Gabor filter [10] [11]. LBP is a powerful texture operator which thresholds the neighborhood of each pixel to label the pixels of an image and uses the result as a binary number. The Gabor filter is a linear filter developed for edge detection.

Frequency and orientation representations of Gabor filters are particularly appropriate for texture representation and discrimination.



(a) small perforation (b) big hole (c) retraction

Figure 5. The three symptoms of COM.

### C. Feature Selection and Classification

After segmentation and visual feature extraction, the final step is to build an accurate OM classifier. We adopt a strong feature selection and training structure shown in Fig. 6. A part of the data is chosen as training data and the remaining data is used as test data. We extract all the visual features (GCM, HOG, LBP, Gabor filter) mentioned in Section Ⅲ-B and name them "low-level features". Then we exploit discriminative learning (i.e., SVM [12]) to train classifiers using the low-level features. Mid-level features (predict probability) are produced by inputting the low-level features of test data. In total, 4 mid-level features are generated as a feature bank. The most effective mid-level features are selected by Adaboost as an OM strong classifier. All data are divided into 6 parts. This structure is executed 6 times for stability and reliable accuracy.

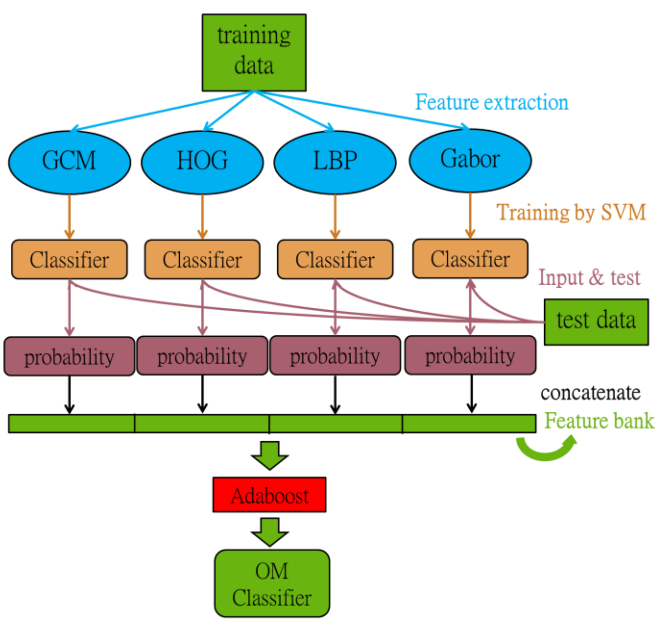



Figure 6. The flowchart of our feature selection and training structure.

## IV. EXPERIMENTAL RESULTS AND DISCUSSION

### A. Database

We have cooperated with seven ear-nose-throat (ENT) physicians of Cathay General Hospital. They have provided 865 pictures with diagnostic labels as a robust training set. The database encompasses almost all OM diagnostic categories: normal; AOM: hyperemic stage, suppurative stage, ear drum perforation, subacute/resolution stage, bullous myringitis, barotrauma; OME: with effusion, resolution stage (retracted); COM: simple perforation, active infection, retraction pocket, cholesteatoma, granulation, post-operative status; OM with ventilation tube, myringitis, and traumatic perforation.

### B. Experimental Results

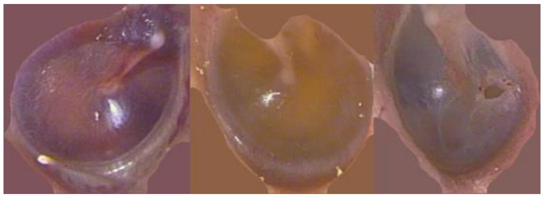
Two experiments were performed to evaluate the performance of our algorithm. We ran our segmentation, feature extraction, feature selection and classification algorithm on the data base mentioned in section Ⅲ-A to evaluate our OM detection system from the first experiment. The second experiment is to test the influence of segmentation and its effect on diagnostic accuracy. Sensitivity and specificity are popular statistical measures of the performance for a classification test. Sensitivity is the percentage of patients who are correctly identified as having OM. Specificity is the percentage of healthy individuals who are correctly identified as not having OM. The F1-measure [14] is a common metric for evaluating the correctness of a classification algorithm. It is defined as the harmonic mean of precision and recall.

Experimental results illustrated in TABLE Ⅰ shows our classification system to achieve 88.06% accuracy and 91.57% sensitivity. Fig. 7. shows that our OM detection system successfully uses color, geometric and texture features to discriminate between infected and normal eardrums. Our segmentation algorithm enhances the accuracy in experiment 2, however the improvement is lower than expected. We infer the reason may be the imperfection of our segmentation results. Boundaries in the region where the malleus connects to ear canal are not always visually obvious. As a result, the malleus of some cases is improperly removed resulting in the loss of some key features, like retraction. Fig. 8. demonstrates perfect and imperfect segmentation results.

The total computation time of test one image is 12.23 seconds including segmentation time (1.92s), feature extraction time (10.2s) and classification time. The program ran on desktop PC with intel i7-3770 CPU and 4GB ram.

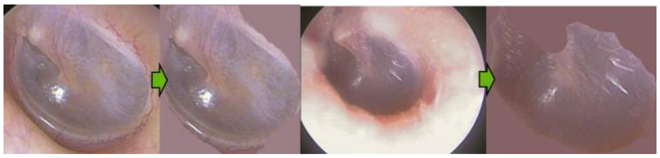
TABLE I. EXPERIMENTS

| **Experiment** | **Measures** | | | |
|---|---|---|---|---|
| | ***Accuracy*** | ***Sensitivity*** | ***Specificity*** | ***$F_1$ Score*** |
| Experiment 1. OM detection system | 88.06% | 91.57% | 79.87% | 0.914 |
| Experiment 2. system without segementation | 83.75% | 90.63% | 67.84% | 0.874 |



(a) AOM (b) OME (c) COM

Figure 7. Our OM detection system can successfully detect the cases, depending on the different color or geometric structure from normal eardrum.



(a) perfect segmentation (b) imperfect segmentation

Figure 8. The perfect and imperfect segmentation results.

### C. Observation and Discussion

We have observed that among all of the visual features, color feature GCM plays the most important role; HOG geometric feature has the second largest influence; while the texture features of the Gabor filter and LBP have a less pronounced influence. We infer that since red color is strongly indicative of AOM, amber color of OME, and the color of COM often appears differently to that of the normal eardrum and other categories of OM, so color is the feature with the highest discriminating power. Geometric features can successfully represent the geometric structure on eardrum such as perforation, cholesteatoma and severe retraction, which are the symptoms of COM, where HOG also has a greater influence. The texture of OM eardrum may be similar to normal eardrum, with only small perforations of the eardrum, so the Gabor filter and LBP have the less influences on the classifier (but still aids in increasing our OM detection accuracy).

We are interested in the false-negative cases, where our system did not detect OM successfully. We conclude two common false-negative types: 1. Small variation, where there is only a small difference in color, shown in Fig. 9. (a); 2. Those with similar structure from non OM cases, as shown in Fig. 9. (b), a small patch of red on the eardrum caused by external force; 3. The case of an improperly segmented TM, as shown in Fig. 9. (c), whose retraction region (red contour) is removed. Some retraction cases like those are difficult to detect and need other methods to calculate the obvious visibility of the malleus. Even now, these cases are still difficult to expertly diagnose by doctors with only one single image.

## V. Conclusions and Future Works

In this paper, we propose a complete OM detection system. Our segmentation method is designed for otoscope images and to be able to properly segment the tympanic membrane. We propose applying color, geometric and texture features to

(a) OME (b) Traumatic perforation (c) retraction (d) the origin image of (c)

Figure 9. Three different types false-negative cases.

characterize the eardrum. A strong OM classifier is then built by Adaboost. Our system achieves high accuracy and sensitivity.

Our future work aims to solve the more difficult cases. Mild symptoms of COM where the malleus is segmented resulting in loss of several key features. As well more features need to be designed to properly address the cases of OM (Fig. 9. (a)(b)) which are seen as normal by the current classification system.

## References


[1] D. W. Teele, J. O. Klein, and B. Rosner, “Epidemiology of otitis media during the first seven years of life in children in greater Boston: a prospective, cohort study,” Journal of Infectious Diseases, vol. 160, no. 1, pp. 83–94, 1989..

[2] American Academy of Pediatrics, “Diagnosis and management of acute otitis media,” Pediatrics, vol. 113, no. 5, pp. 1451–1465, 2004.

[3] Mironica, I., Constantin Vertan, and Dan Cristian Gheorghe. "Automatic pediatric otitis detection by classification of global image features." E-Health and Bioengineering Conference (EHB), 2011. IEEE, 2011.

[4] Kuruvilla, Anupama, et al. "Automated diagnosis of otitis media: vocabulary and grammar." Journal of Biomedical Imaging 2013 (2013): 1.

[5] Chan, Tony F., and Luminita A. Vese. "Active contours without edges." Image processing, IEEE transactions on 10.2 (2001): 266-277.

[6] M. Stricker and M. Orengo, “Similarity of color images,” 1995, pp. 381–392.

[7] Dalal, Navneet, and Bill Triggs. "Histograms of oriented gradients for human detection." Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on. Vol. 1. IEEE, 2005.

[8] Ojala, Timo, Matti Pietikäinen, and David Harwood. "A comparative study of texture measures with classification based on featured distributions." Pattern recognition 29.1 (1996): 51-59.

[9] Ojala, Timo, Matti Pietikainen, and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." Pattern Analysis and Machine Intelligence, IEEE Transactions on 24.7 (2002): 971-987.

[10] Daugman, John G. "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters."JOSA A 2.7 (1985): 1160-1169.

[11] Fogel, Itzhak, and Dov Sagi. "Gabor filters as texture discriminator." Biological cybernetics 61.2 (1989): 103-113.

[12] C.-C. Chang and C.-J. Lin, “LIBSVM: A library for support vector machines,” ACM Transactions on Intelligent Systems and Technology, vol. 2, pp. 27:1–27:27, 2011, software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm.

[13] Y. Freund and R. E. Schapire, “Experiments with a new boosting algorithm,” 1996.

[14] C. van Rijsbergen, Information Retrieval (2nd ed.), Butterworth-Heinemann, Newton, MA, USA, 1979.