Computers in Biology and Medicine 61 (2015) 107–118

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Detecting tympanostomy tubes from otoscopic images via offline and online training

Xin Wang [a], Tulio A Valdez [b], Jinbo Bi [a,*]

[a] Department of Computer Science and Engineering, University of Connecticut, 371 Fairfield Way, Unit 4155, Storrs, CT 06269-4155, USA
[b] Department of Ototaryngology, Connecticut Children's Medical Center, Hartford, CT 06106 USA

ARTICLE INFO

Article history:
Received 24 November 2014
Accepted 23 March 2015

Keywords:
Object detection
Tympanostomy tube
Otoscopic image
Cascaded classifier
Support vector machine

ABSTRACT

Tympanostomy tube placement has been commonly used nowadays as a surgical treatment for otitis media. Following the placement, regular scheduled follow-ups for checking the status of the tympanostomy tubes are important during the treatment. The complexity of performing the follow up care mainly lies on identifying the presence and patency of the tympanostomy tube. An automated tube detection program will largely reduce the care costs and enhance the clinical efficiency of the ear nose and throat specialists and general practitioners. In this paper, we develop a computer vision system that is able to automatically detect a tympanostomy tube in an otoscopic image of the ear drum. The system comprises an offline classifier training process followed by a real-time refinement stage performed at the point of care. The offline training process constructs a three-layer cascaded classifier with each layer reflecting specific characteristics of the tube. The real-time refinement process enables the end users to interact and adjust the system over time based on their otoscopic images and patient care. The support vector machine (SVM) algorithm has been applied to train all of the classifiers. Empirical evaluation of the proposed system on both high quality hospital images and low quality internet images demonstrates the effectiveness of the system. The offline classifier trained using 215 images could achieve a 90% accuracy in terms of classifying otoscopic images with and without a tympanostomy tube, and then the real-time refinement process could improve the classification accuracy by 3–5% based on additional 20 images.

1. Introduction

Tympanostomy tube placement is the most common outpatient surgical procedure in the United States. Each year more than 650,000 children younger than 15 years of age receive tympanostomy tubes [1]. Tympanostomy tubes are commonly inserted because of persistent middle ear fluid affecting hearing, frequent ear infections or ear infections that persist after antibiotic therapy [2]. All of these conditions are composed of the term otitis media, which is second in frequency only to upper respiratory infection as the most common illness diagnosed in children by pediatric health providers. Most children will experience at least one episode of acute otitis media by the age of 3 years and by the age of 6 years, nearly 40% have experienced three or more infections [3]. Otitis media with effusion (OME) can resolve spontaneously, however for patients in which, there is persistent middle ear fluid for at least 3 months with the decrease in hearing, further treatment in the form of tympanostomy tubes insertion may be recommended [2]. Fig. 1 illustrates the placement of the tympanostomy tube on the ear drum.[1]

Since first described in 1954 by Armstrong, tympanostomy tube placement has become the surgical treatment of choice for otitis media [4]. Placement of tympanostomy tubes improves hearing significantly in the presence of otitis media with effusion, reduces the incidence of recurrent acute otitis media (AOM), and provides a mechanism for drainage and administration of topical antibiotic therapy for acute otitis media. The latter has gained significant importance since it allows localized treatment for ear infections rather than systemic antibiotics use. The insertion of tympanostomy tubes involves aspiration of the middle ear fluid leading to instant improvement of hearing thresholds. The tympanostomy tubes are designed to extrude naturally from the tympanic membrane normally within 6 months to 1 year following the procedure.

Following tympanostomy tube placement, regular scheduled follow-ups are recommended by the American Academy of Otolaryngology every 6 months to check the status of the ear tubes [5]. This follow up care is currently performed by an ear, nose and throat

* Corresponding author. Tel.: +1 860 486 1458.
E-mail address: jinbo@engr.uconn.edu (J. Bi).

[1] The left graph of Fig. 1 is downloaded from www.kidshealth.org.nz.

http://dx.doi.org/10.1016/j.compbiomed.2015.03.025
0010-4825/

specialist resulting in increased cost compared to the cost of a visit to a general practitioner [6]. The overall complexity of the procedure is low because it mostly identifies the presence and patency of the tympanostomy tube. If this procedure can be performed by an automated computer program, it can significantly reduce cost as well as enhance the clinical efficiency of both ear nose and throat specialists and general practitioners.

Computer vision techniques have been widely used to automatically detect an object in natural images such as human faces [7–10], cars [11–13] and pedestrians [14–16]. The related techniques are also used to analyze medical images which have largely shown to facilitate the diagnosis and treatment decisions [17–23]. In this study we have developed a computer vision system to predict on an image if a tympanostomy tube is in place and collect feedback to adaptively improve the system classification performance. In this system, we extract a set of image features including RGB intensity features, edge-map based features and other advanced features such as Scale-invariant Feature Transform (SIFT) [24], and Histograms of Oriented Gradients (HOG) features [25]. We use them to learn a discriminative model to determine if an otoscopic image presents a tympanostomy tube. A cascaded classifier is constructed by the support vector machine (SVM) [26] algorithm from labeled image patches. The prediction of this classifier, when applied to a test image, is visualized. A refinement process is designed to refine this trained classifier at the point of patient care according to user feedbacks. Extensive experimental results demonstrate the effectiveness and efficiency of the proposed approach.

2. Related works

Most object detection techniques are composed of two major aspects, feature extraction and model learning, for each of which various methods have been proposed. Readers can consult with [27,28] for an overview and challenges of the field. In the subsequent paragraphs, we briefly review several of the most relevant methods.

Region based and edge-map based features are the two types of widely used features for object detection. For the first type, the features are generated from colors [29–32], or the varying distribution of intensities [33,34,24,25]. Among these features, the SIFT [24] and HOG [25] features are the most widely used nowadays. Other algorithms have considered texture information, such as gray level co-occurrence matrix (GLCM) [35], local binary patterns (LBP) [36] and wavelet texture [37,12]. We explored these features in our system. The edge-map based features are used to capture contour shapes by computing the most representative edge fragments [38–41]. These features are robust to occlusion but are not invariant to illumination conditions. Geometrical shapes and the structure of lines and arcs in an edge map were also studied in [42–44], which facilitated to obtain more reliable features. It has been shown that combining the region based features and edge-map features may lead to robust detection [45,46]. We hence implemented feature extraction methods in both of the categories.

Various machine learning algorithms have been applied to build probabilistic models for prediction of object classes. Typical methods include Bayesian classifiers [47], expectation maximization [48], k-Nearest Neighbor [49], logistic regression [50] and support vector machines (SVM) [26,51]. These methods perform comparably although some may serve certain specific purposes, such as selecting features for use in the model. In this work, an SVM algorithm with linear classifiers was used and served as a good learning model for detecting the tubes. The SVM algorithm solves a quadratic programming optimization problem for a predictive classifier that has an optimal margin to separate different classes of examples.

Besides the two major components discussed above, online learning methods that sequentially take user feedback in building a classifier have been used to detect and track moving objects [52–55]. Our system also consists of a real-time refinement process. Although similar to an online learning method, our refinement targeted at the unseen examples that the offline models fail to predict. After the system highlights the predictions on an image, the users can correct the prediction results on the image. Our system is capable of recording this feedback and retraining the related classifiers.

3. Material and methods

In this section, we describe our approach to automatically predicting if an otoscopic image of ear drum has a tympanostomy tube mounted. Fig. 2 provides an overview of our detection system. In this system, an image is first converted into a feature vector in the feature extraction process. Then a classifier that has been pre-trained and stored in a database (which we call knowledge database) is applied to the image features to predict the presence and location of a tube. The results are then visualized and highlighted in the image. Users can give feedback by correcting the predictions if they are inaccurate. This feedback message is then recorded and used in an automatic process to refine the classifiers stored in the knowledge database. This system is constructed

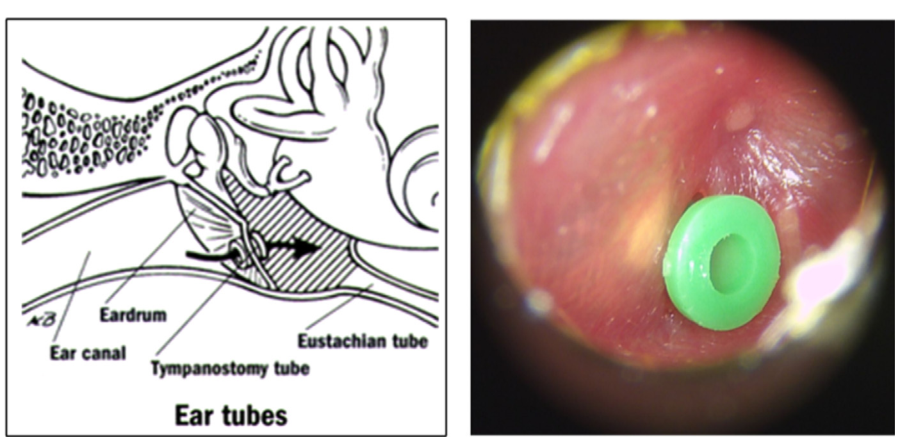

Fig. 1. Illustration of a tympanostomy tube in place and patency in the anterior inferior quadrant of the tympanic membrane: (left) a graph of ear structure and tube and (right) an otoscopic image.