Available online at www.sciencedirect.com
ScienceDirect
Computerized Medical Imaging and Graphics 31 (2007) 587–594
Computerized Medical Imaging and Graphics
www.elsevier.com/locate/compmedimag

A kernel-based approach to categorizing laryngeal images

A. Verikas [a,c,*], A. Gelzinis [a], M. Bacauskiene [a], D. Valincius [a], V. Uloza [b]

[a] *Department of Applied Electronics, Kaunas University of Technology, Studentu 50, LT-3031, Kaunas, Lithuania*
[b] *Kaunas University of Medicine, Kaunas, Lithuania*
[c] *Intelligent Systems Laboratory, Halmstad University, Box 823, S 301 18 Halmstad, Sweden*

Received 5 April 2006; accepted 3 July 2007

Abstract

This paper is concerned with an approach to automated analysis of vocal fold images aiming to categorize laryngeal diseases. Colour, texture, and geometrical features are used to extract relevant information. A committee of support vector machines is then employed for performing the categorization of vocal fold images into *healthy*, *diffuse*, and *nodular* classes. The discrimination power of both, the original and the space obtained based on the kernel principal component analysis is investigated. A correct classification rate of over 92% was obtained when testing the system on 785 vocal fold images. Bearing in mind the high similarity of the decision classes, the correct classification rate obtained is rather encouraging.

Keywords: Laryngeal image; Colour; Texture; Gabor filtering; Co-occurrence matrix; Support vector machine

1. Introduction

The diagnostic procedure of laryngeal diseases is based on visualization of the larynx, by performing indirect or direct laryngoscopy. A physician then identifies and evaluates colour, shape, geometry, contrast, irregularity and roughness of the visual appearance of vocal folds. This type of examination is rather subjective and to a great extent depends on physician's experience. An ability to obtain objective measures of these features would be very helpful for assuring objective analysis of the images of laryngeal diseases and creating systematic databases for education, comparison and research purposes. In addition to the data obtained from one particular patient, information from many previous patients – experience – plays also a very important role in the decision making process. Moreover, the physician interpreting the available data from a particular patient may have a limited knowledge and experience in analysis of the data. In such a situation, a decision support system for automated analysis and interpretation of medical data is of great value. Recent developments in this area have shown that physicians benefit from the advise of decision support systems in terms of increased reliability of the analysis, decreased intra- and inter-observer variability [1–4].

This paper, is concerned with an approach to automated analysis of vocal fold – laryngeal – images aiming to categorize diseases of vocal folds. It is worth noting that due to the large variety of appearance of vocal folds, the categorization task is sometimes difficult even for a trained physician. Fig. 1 presents three examples of laryngeal images. The image placed on the right-hand side of the figure comes from the *diseased* class, while the other two are taken from the *healthy* vocal folds. In this case, the only discriminative feature is the slightly convex vocal fold edges in the upper part of the image coming from the *diseased* class.

A very few attempts have been made to develop computer-aided systems for analyzing vocal fold images. In our previous study [5], a committee of multilayer perceptrons employed for categorizing vocal fold images into three decision classes correctly classified over 87% of test set images.

In this paper, we investigate effectiveness of the kernel-based approach to feature extraction and classification of laryngeal images. We treat the problem as an image analysis and recognition task. To obtain an informative representation of a vocal fold image that is further categorized by a committee of support vector machines, texture, colour, and geometrical features are used. The choice of the feature types was based on the type of

* Corresponding author at: Intelligent Systems Laboratory, Halmstad University, Box 823, S 301 18 Halmstad, Sweden. Tel.: +46 35 167140; fax: +46 35 216724.

E-mail addresses: antanas.verikas@ide.hh.se (A. Verikas), adas.gelzinis@ktu.lt (A. Gelzinis), marija.bacauskiene@ktu.lt (M. Bacauskiene), donatas.valincius@ktu.lt (D. Valincius), uloza@kmu.lt (V. Uloza).

0895-6111/$ – see front matter
doi:10.1016/j.compmedimag.2007.07.003

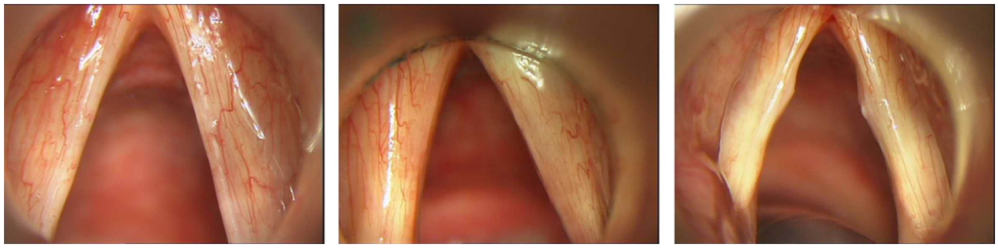

Fig. 1. Three examples of laryngeal images.

information used by the physician when analyzing images of vocal folds. Each member of the committee is devoted for analysis features of a single type. The discrimination power of both, the original and the space obtained based on the kernel principal component analysis is investigated.

The remainder of the paper is organized as follows. In the next section, we briefly describe the data used. Section 3 outlines the analysis techniques employed. Section 4 presents the results of the experimental investigations. Finally, conclusions of the work are given in Section 5.

2. Data

This study uses a set of 785 laryngeal images recorded at the Department of Otolaryngology, Kaunas University of Medicine during the period from October 2002 to December 2003. The internet based archive – database – of laryngeal images is continuously updated. The laryngeal images were acquired during routine direct microlaryngoscopy employing the Moller-Wedel Universa 300 surgical microscope. The 3-CCD Elmo colour video camera of 768× 576 pixels was used to record the images.

2.1. A gold standard

We used the gold standard taken from the clinical routine evaluation of patients. A rather common, clinically discriminative group of laryngeal diseases was chosen for the analysis, i.e. mass lesions of vocal folds. Visual signs of vocal fold mass lesions (colour, shape, surface, margins, size, and localization) are rather typical, clinically evident and descriptive.

Mass lesions of vocal folds could be categorized into six classes namely, *polypus*, *papillomata*, *carcinoma*, *cysts*, *keratosis*, and *nodules*. This categorization is based on clinical signs and a histological structure of the mass lesions of vocal folds. In this initial study, the first task was to differentiate between the *healthy* (*normal*) class and pathological classes and then, differentiate among the classes of vocal fold mass lesions. We distinguish two groups of mass lesions of vocal folds, i.e. nodular – *nodules*, *polyps*, and *cysts*– and diffuse – *papillomata*, *keratosis*, and *carcinoma*– lesions. Thus, including the *healthy* class, we have to distinguish between three classes of images.

Amongst the 785 images available, there are 49 images from the *healthy* class, 406 from the *nodular* class, and 330 from the *diffuse* class. It is worth noting that due to the large variety of appearance of vocal folds, the classification task is difficult even for a trained physician. Fig. 2 presents characteristic examples from the three decision classes considered, namely, *nodular*, *diffuse*, and *healthy*.

3. Methods

To obtain an informative representation of a vocal fold image, colour, texture, and geometrical features are used. The measurement values related to image colour (C), texture (T), and geometry (G) are collected into three separate vectors $\boldsymbol{\psi}_{\mathrm{C}}$, $\boldsymbol{\psi}_{\mathrm{T}}$, and $\boldsymbol{\psi}_{\mathrm{G}}$. Having the measurement vectors, features of the aforementioned three types are then obtained by applying the kernel principal component analysis [6] separately for each of the spaces – spanned by the colour, texture, and geometrical measurements – as explained below.

Thus, having a vector of measurements $\boldsymbol{\psi}$, the feature vector $\boldsymbol{\xi}$ is computed in the following way. Assume that κ is a Mercer ernel [7] and Φ is a mapping of $\boldsymbol{\psi}$ onto a feature space F, such that $\kappa(\boldsymbol{\psi}_i, \boldsymbol{\psi}_j) = \langle \Phi(\boldsymbol{\psi}_i), \Phi(\boldsymbol{\psi}_j)\rangle$, where $\langle \cdot, \cdot \rangle$ stands for the inner product. Let $\tilde{\Phi}(\boldsymbol{\psi}_i)$ denote the centered data point in the feature space F.

$$\tilde{\Phi}(\boldsymbol{\psi}_i) := \Phi(\boldsymbol{\psi}_i) - \frac{1}{M}\sum_{i=1}^{M}\Phi(\boldsymbol{\psi}_i) \qquad (1)$$

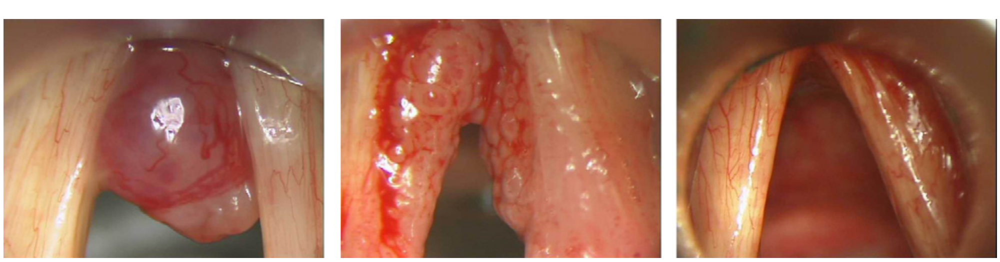

Fig. 2. Images from the *nodular* (left), *diffuse* (middle), and *healthy* (right) classes.