COMPUTER AIDED DIAGNOSIS OF PAROTID GLAND LESIONS USING ULTRASONIC MULTI-FEATURE TISSUE CHARACTERIZATION

Stefan Siebers,* Johannes Zenk,† Alessandro Bozzato,† Nils Klintworth,† Heinrich Iro,† and Helmut Ermert*

*High Frequency Engineering Research Group, Ruhr-University-Bochum, Bochum, Germany; and †Department of Otorhinolaryngology, Head and Neck Surgery, University of Erlangen-Nuremberg, Erlangen, Germany

INTRODUCTION

Among the salivary glands in the head and neck region, the parotid glands are the largest. Located between the outer ear and the lower jaw, they are embedded in the subcutaneous tissue of the face. Parotid gland tumors are removed surgically by partial or complete gland excision. However, although new surgical techniques including nerve monitoring (Timmermann et al. 2004) have been proposed, the intraglandular course of the facial nerve makes parotid surgery a challenging intervention. During surgery, it is intended to preserve the nerve function and to avoid a nerve injury that possibly leads to partial face palsy. Thus, exact and reliable diagnostic methods are desired to determine in advance whether and to what extent a surgical intervention is necessary.

As the main diagnostic method, ultrasound B-mode imaging is used in the majority of cases. Additionally, Doppler sonography, magnetic resonance imaging, palpation and occasional needle biopsies are applied. With B-mode imaging, parotid gland lesions can be detected quiet easily in general. When a lesion is detected, the examiner tries to determine the type of lesion based on its appearance in the image. Common criteria are texture and echogenicity of the lesion, its shape and boundary, the dorsal echo enhancement and compressibility. However, the exact classification and, hence, the diagnosis of malignancy or benignancy is demanding due to the occurrence and diversity of salivary gland tumors (Bozzato et al. 2007) (exemplary B-mode images of parotid gland lesions are shown in Fig. 1). In

Address correspondence to: Stefan Siebers, High Frequency Engineering Research Group, Ruhr-University-Bochum, Universitaetsstr. 150, D-44780 Bochum, Germany. E-mail: stefan.siebers@rub.de

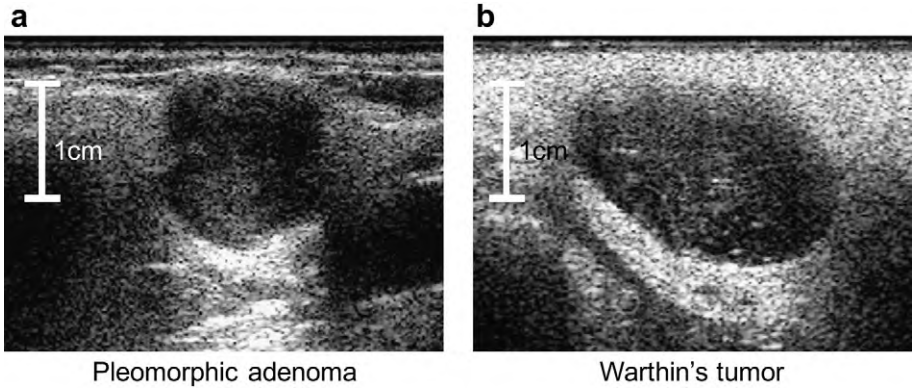

Fig. 1. B-mode images of most common parotid gland lesions. (a) 70-year-old female patient with a pleomorphic adenoma (positive group). (b) 66-year-old male patient with a Warthin's tumor (negative group). Note that pleomorphic adenomas are actually benign lesions but may lead to malignancy in a later stage of disease and are therefore treated as malignant (positive) cases in this work.

addition, tumor-like lesions, as cysts or lymph nodes, are casually mistaken for tumors. Thus, high experience and expertise are required, which only specialized examiners can provide. Since the difficulties in diagnosing salivary gland lesions lead to a large number of false positives and, hence, to many dispensable surgeries, computer aided systems are strongly desired to support and facilitate the physician's diagnosis. For this reason, we developed a sonohistology system for automated ultrasonic tissue characterization aiming at an automated differentiation between cases where surgical treatment is inevitable and such cases that must not necessarily be treated.

As a well researched topic in the past, different approaches to ultrasonic tissue characterization have been applied for various aims. The basic concept, which most approaches have in common, involves the extraction of parameters from ultrasound radio-frequency (RF) or image data to be used as features representing distinctive properties of the investigated biologic tissue. Based on a decision rule, the investigated tissue is then classified by assigning one of several target classes. If one single feature is used, the decision rule is realized by simple thresholding. However, using one feature exclusively is commonly not sufficient to distinguish between different kinds of tissue with the desired accuracy. To account for this, a multi-feature approach is used in this work.

Applications for ultrasonic tissue characterization have been manifold in the past, including prostate cancer (Lizzi et al. 1997b; Feleppa et al. 2001; Scheipers et al. 2003a; Schmitz et al. 1999), atherosclerotic plaques (Bridal et al. 1997; Noritomi et al. 1997), testicular tissue (Jenderka et al. 1999), ophtalmology (Lizzi and Laviola 1975; Lizzi et al. 1987), liver and hepatic diseases (Lang et al. 1994; Lizzi et al. 1997b; Oosterveld et al. 1991), myocardial tissue (O'Donnell et al. 1981), staging of venous thrombosis (Kolecki et al. 1995; Parsons et al. 1993; Siebers et al. 2004), monitoring of thermal therapies (Lemor et al. 2000; Siebers et al. 2006, 2007b), characterization of apoptosis (Kolios et al. 2002), mammary tumors (Chang et al. 2003; Chen et al. 2002), classification of salivary gland tumors (Chikui et al. 2005; Scheipers et al. 2005b; Siebers et al. 2007a) and thyroid alterations (Smutek et al. 2003; Tsantis et al. 2005).

Unlike image-based features, spectral features extracted from RF data have been proven to be superior for ultrasonic tissue characterization (Gaitini et al. 2004; Scheipers et al. 2003b). However, classification rates can be significantly improved in many cases combining spectral and image based features, as it has been shown in several studies (Gaitini et al. 2004; Oosterveld et al. 1991; Scheipers et al. 2003a; Siebers et al. 2007b).

The choice of features used for classification of parotid gland lesions within the study presented in this article was motivated by the criterions applied by physicians during the standard ultrasound examination. Thus, spectral features, image based texture features and additionally contour descriptors were used to account for differences in echogenicity, texture and shape. Using a sequential search strategy, a subset of features was selected to be processed by the classification system. The system's decision rule was realized by a maximum likelihood classifier.

MATERIALS AND METHODS

Data acquisition and preprocessing

Ultrasound echo data were acquired with a Sonoline Elegra digital ultrasound scanner (Siemens Medical Solutions, Erlangen, Germany) and a linear probe (7.5 L40) working at a center frequency of 7.2 MHz. The useful bandwidth of the probe used for spectral analysis was 75% of the center frequency. Data were

captured during the routine examination of patients scheduled to have parotidectomy: 138 patients were included in the clinical study. Over all, 66 patients were male and 72 patients were female. The youngest patient on the day of the examination was 6 years old and the eldest patient was 88 years old. The mean age of the patients was 57.1 years.

The decision whether to operate or to follow-up is based on an initial sonographic imaging. Swollen and protruded lesions are excised, especially when an increase in size is noted. Even benign looking lesions cannot be excluded in general from surgery to confirm or exclude malignancy. A second imaging including the data acquisition was performed at hospital admission right before surgery. Patients with a sonographically assessed space occupying lesion of the parotid gland and the decision of surgical removal were selected. Exclusion criteria were prior surgery in this area and pregnancy. No biopsy was performed before surgery. The time between data acquisition and surgery was 3 days at most. All patients underwent partial or complete parotidectomy under constant facial nerve monitoring. Patient compliance to the procedure was high, as the new method did not extend the normal examination time when applying ultrasound imaging to the head and neck region.

The study was given approval by the local ethics committee. Data acquisitions and conventional diagnostics were performed by two experienced, board certified head and neck sonographers. All patients gave informed consent to participate in the study.

Purpose-made software written in Visual C++ (Microsoft, Redmont, WA, USA) running on a laptop computer was employed to control the ultrasound scanner. All relevant imaging settings were monitored by the software to guarantee a standardized data acquisition. The internal operating system of the Elegra was accessed via telnet to download complex base band ultrasound echo data. Such data were used to reconstruct the original RF data by modulation of a dynamic carrier with known characteristics. The download of data was done via file transfer protocol (FTP). A dump file containing all relevant settings of the Elegra and the probe was also stored along with each frame and used later to compensate for TGC settings.

For each lesion, two orthogonal image frames were recorded. Each frame consists of 360 lines and 2400 samples per line. The approximate size of the images is 5.1 cm in the axial direction and 4 cm in the lateral direction. A single transmit focus was set to a depth of 2 cm, matching approximately the mean depth of lesions observed during this study. The sampling rate and amplitude resolution of the RF echo data was 36 MHz and 12 bits, respectively.

B-mode images were constructed from the baseband data by envelope detection and logarithmic compression. In an initial preprocessing step, the B-mode images were displayed on a laptop computer screen using a purpose-made graphical user interface (GUI). Two experienced physicians manually segmented the lesions in the B-mode images by consensus to define a region-of-interest (ROI). The physicians were blinded to the results of the histology. This initial preprocessing step was crucial for feature extraction to ensure the exclusive usage of signals originating from the lesions and not from surrounding tissue. The GUI as well as all image and signal processing tools were developed using Matlab (The MathWorks, Natick, MA, USA).

For the purpose of comparison, the experienced physicians' diagnoses were also recorded. The conventional diagnosis included the measurement of the tumor in three dimensions and qualitative classification according to the following sonographic criteria (*indicates malignancy): distal phenomenon (distal enhancement/dwindling); sonographic margin definition (sharp, regular, indistinguishable*, irregular*); echogenicity (homogeneous, inhomogeneous*, echo-poor, echo-rich, anechoic); perfusion pattern in power mode (peripheral, central, diffuse) (Bozzato et al. 2007). At the end of the examination, the examiner decided on a suspected ultrasound diagnosis. For this study, the mass was classified as a malignant or a benign lesion according to Table 1. Histopathologic examination results obtained after parotidectomy served as a reference or "gold standard" for both conventional diagnostics and the new computer aided method.

Table 1. Pathohistologic results: Occurrence of different kinds of parotid gland lesions with quantity n during clinical study

Type of tumor	n	Target class
Pleomorphic adenomas	29	
Acinar cell carcinomas	7	
Mucoepidermoid carcinoma	1	
Oncocytic carcinoma	1	
Squamous epithelial carcinoma	2	Ω_{pos}
Sarcomatoid carcinoma	1	
Adenocarcinoma	1	
Adenoid cystic carcinoma	1	
Lymphomas	4	
Metastases	4	
Warthin's tumors	46	
Basal cell adenomas	11	
Canaliculous adenoma	1	
Adenoid Cyst	1	
Lipomas	5	Ω_{neg}
Lymph nodes	9	
Cysts	13	
Nodular fasciitis	1	

The total number of cases was 138. The positive class Ω_{pos} and negative class Ω_{neg} contained 51 and 87 cases, respectively.

Methods: Feature extraction and classification

This section gives a brief overview of the features used for the characterization of parotid glands and motivates their choice. Altogether, three kinds of features have been used in this study: statistical distributions of spectral parameters, textural features and shape descriptive features. Such features have been widely used for ultrasonic tissue characterization. However, only few works took advantage of the possibility to enhance classification rates by combining different kinds of features, *e.g.*, texture and spectral features (Gaitini et al. 2004; Oosterveld et al. 1991; Scheipers et al. 2003a; Siebers et al. 2007b).

Attenuation estimation

The acoustic attenuation is one of the most commonly used parameters for ultrasonic tissue characterization. The choice of the acoustic attenuation for discriminating different kinds of tissue is motivated by findings from fundamental works using transmission ultrasound for the investigation of different kinds of pathological or healthy tissue (Duck 1990). If transmission ultrasound is used, the acoustic attenuation of a tissue sample can be determined by the insertion loss method. However, for *in vivo* ultrasound investigations, predominantly pulse-echo techniques are used since the application of transmission ultrasound is limited to certain organs and regions of the body like the female breast (Jago 1993). Hence, methods for estimating the frequency dependent attenuation from reflected ultrasound signals have been developed in the past. In this work, we applied two different techniques to estimate the frequency dependent attenuation of tissue.

The spatially resolved estimation of spectral attenuation, as well as backscatter parameters as described in the next section, requires the RF data-frame to be subdivided into overlapping segments. The ROI size was chosen as a tradeoff between spatial resolution and quality of spectral estimates. Thus, we set the segment size to 128 samples axially, which facilitates the application of the fast Fourier transform, and 16 lines laterally, spanning an area of approximately 4.6mm^2 (2.7mm $\times$ 1.7mm). These settings correspond with findings in (Chen et al. 1993; Lizzi et al. 1983; Madsen et al. 1984). Overlap was set to 50% in axial and lateral directions. From each segment, a calibrated and averaged power spectrum $\langle S_{korr}(f, z)\rangle$ is calculated. Calibration is achieved by dividing each spectrum by a reference spectrum obtained from a measurement on a single wire phantom at the same depth. Thus, system and diffraction influences are partially eliminated. The wire (polyamide, diameter 60 μm) was spanned perpendicular to the imaging plane and was moved by a stepper motor in the axial direction in increments of 1.4 mm. From each depth, 20 spectra were acquired and averaged to reduce the spectral variance. The resulting spectogram was obtained using a cubic spline interpolation. Measurements on the wire phantom were carried out in a temperature controlled water tank. By adding sodium chloride, the speed of sound of the solution was adjusted to match 1540 m/s [NaCl concentration 53.6 g/L at 20°C; (Chen et al. 1978; Rogers and Pitzer 1982)]. The resulting power spectrum can be expressed as

$$\left\langle S_{korr}(f,z)\right\rangle = B(f)^2 \cdot 10^{-\frac{\alpha(f)}{10\mathrm{dB}}2z}, \tag{1}$$

where $B(f)$ is the frequency dependent acoustic backscatter, $\alpha(f)$ is the frequency dependent acoustic attenuation and $2z$ is the round-trip distance. Within the useful bandwidth, the frequency dependent attenuation can be approximated by a linear model (Bridal et al. 1997; Cloostermans and Thijssen 1983; Kuc and Schwartz 1979)

$$\alpha(f) = \alpha_1 \cdot (f - f_c) + \alpha(f_c), \tag{2}$$

where f_c is the center frequency within the useful bandwidth. The first method used in this work is the multi-narrowband (mnb) method (Cloostermans and Thijssen 1983), which provides as features for tissue characterization the acoustic attenuation as well as the attenuation coefficient that indicates the frequency dependency of the acoustic attenuation. The acoustic attenuation $\alpha(f_j)$ is determined for each discrete frequency f_j within the useful bandwidth as the slope of a regression line fitted to the logarithmic decay of the power spectrum over depth. The coefficient ${\alpha_1}^{mnb}$ is then yielded as the slope of a regression line fitted to $\alpha(f)$.

The second method used in this work is the centroid frequency shift (cfs) method (Fink et al. 1983). This method is based on the assumption of a Gaussian shaped spectral density of the backscattered ultrasound signal. It can be shown that the assumed linear frequency dependency of the acoustic attenuation results in a downshift of the power spectrum's center frequency. Hence, the frequency downshift can be utilized as a measure for the frequency dependent attenuation. The attenuation coefficient ${\alpha_1}^{cfs}$ can be determined by a linear least squares fit to the frequency gradient over depth. For both methods, four consecutive segments in the axial direction were used to estimate the power decay and frequency shift, respectively.

Acoustic backscatter estimation

Estimates for acoustic backscatter as proposed by Lizzi et al. (Lizzi et al. 1997a, 1997b) are also widely used for ultrasonic tissue characterization. Such

parameters describing backscatter can be obtained from diffraction corrected spectra [eqn (1)] after compensation for attenuation effects. To compensate for attenuation, a fixed attenuation is assumed instead of estimated values, since both attenuation and backscatter parameters are used to generate tissue-characterizing features, which should be preferably uncorrelated. The remaining backscatter term can be approximated by a linear least squares fit determined by the slope b^S and intercept b^I. In several works, these parameters have been related to properties of the tissue, *e.g.*, scatterer size, concentration and spacing (Lizzi et al. 1987, 1997a).

Spatial distributions of spectral parameters

Attenuation and backscatter parameters can be utilized to create spatially resolved images of the respective estimates. However, such images still contain too much information to be processed by a classifier. Thus, a reduction of the information contained in a parameter image has to be achieved to characterize the whole lesion by a single value. The most basic way is to take the average of a parameter value for the region-of-interest covering the lesion. However, just taking the average does not account for the inhomogeneous structure of lesions investigated in this work (Fig. 2a and b). The statistical distribution of a parameter is represented by its histogram (Fig. 2c). Several statistical measures from the histograms of parameter values are derived as features representing the imaged lesion. The statistical measures are the mean value, standard deviation, kurtosis, skewness and full width at half maximum. These measures are combined in a feature vector serving as a characteristic signature of the lesion.

Texture features

In this work, 14 features calculated from gray level co-occurrence matrices as proposed by Haralick (Haralick et al. 1973) are used to quantify differences in tissue structure appearing in the B-mode image. Such features take into account transitions of gray values in a discretized image. A gray value transition in an image is defined by two image pixels of gray values k and l, separated by a distance d for a given direction. If the transition (k, l) occurs in the image, the entry (k, l) of the $n_g \times n_g$ co-occurrence matrix $\mathbf{C_d}$, where n_g is the number of discretized gray values in the image, is increased by one. Typically, rectangular ROIs or images are analyzed using co-occurrence matrices for multiple directions to obtain features that are almost invariant against rotation. However, since ultrasonic images suffer from a worse lateral than axial resolution, only the axial direction is commonly considered for the calculation of co-occurrence matrices. This approach allows extending

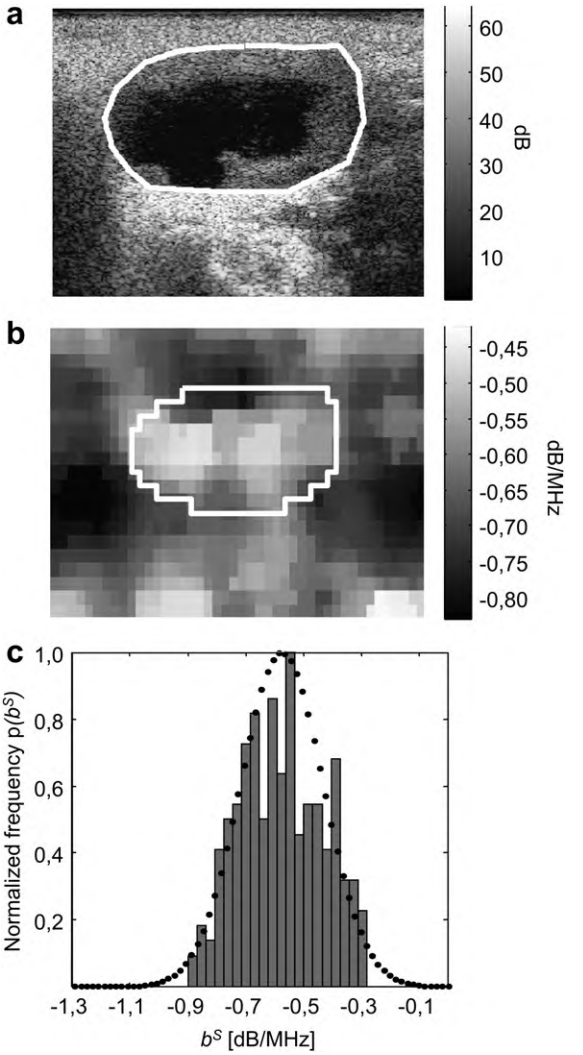

Fig. 2. 84-year-old male patient with a basal cell adenoma. (a) B-mode image with contoured tumor. The hypoechoic region inside the demarcation represents a cystic formation, which is a part of the tumor and a common characteristic for this kind of tumor. (b) Image of parameter b^S (backscatter slope). (c) Histogram of parameter values inside contoured ROI.

the calculation to nonrectangular ROIs as desired here. In general, a manually contoured ROI comprises N_x columns of different lengths. For each column, one co-occurrence matrix $\mathbf{C_{d,x}}$ is determined. The normalized co-occurrence matrix $\mathbf{P_d^{ROI}}$ is then calculated by summing the N_x co-occurrence matrices, divided by the total number of gray value transitions inside the ROI:

$$\mathbf{P_d^{ROI}} = \frac{\sum_{x=1}^{N_x} \mathbf{C_{d,x}}}{\sum_{k=0}^{N_g-1} \sum_{l=0}^{N_g-1} \sum_{x=1}^{N_x} \mathbf{C_{d,x}}}. \quad (3)$$

Texture features are calculated from $\mathbf{P_d^{ROI}}$ according to (Haralick et al. 1973).

Shape descriptors

Besides the echogenicity of a lesion and its textural appearance in the B-mode image, the shape of a lesion in terms of symmetry, boundary regularity or roughness is also commonly utilized as a diagnostic criterion. Differences in shape are quantified in this work by Fourier descriptors. Originally used for hand written character recognition (Granlund 1972; Zahn and Roskies 1972), such descriptors have also been used in medical imaging for characterization of breast lesions using mammography (Shen et al. 1994). Since an automated segmentation of the lesion is not the aim of this work, the manually outlined contours are used for shape analysis. For the calculation of Fourier descriptors, the external boundary line is assumed to be a closed curve in the complex plane. Thus, the curve can be represented by a series of complex numbers. If the curve is supposed to be one single period of the 2π-periodical complex function $c(s)$, a Fourier representation

$$c(s) = \sum_{n=-\infty}^{\infty} C_n e^{jn\frac{2\pi}{L}s}, \tag{4}$$

where s is the curve distance and L is the length of the curve, can be found. The coefficients C_n are given by

$$C_n = \frac{1}{L}\int_0^L c(s)e^{-jn\frac{2\pi}{L}s}ds. \tag{5}$$

Since the curve is noncontinuous and points on the curve are given as discrete values, the discrete Fourier transform is used to calculate descriptors. An inner boundary tracing algorithm (Sonka et al. 1998) was used to yield a uniformly sampled representation of the contour in a four-neighborhood. The coefficients C_k can thus be expressed as

$$C_k = \frac{1}{N}\sum_{n=0}^{N-1} c_n e^{-jkn\frac{2\pi}{N}}. \tag{6}$$

Fourier descriptors are obtained as normalized coefficients

$$\mathcal{F}_k = \frac{|C_k|}{|C_1|}, k \neq 0, 1. \tag{7}$$

Due to the normalization, descriptors $\mathcal{F}_k$ are invariant to scaling. Since coefficient C_0 is omitted and the phase information is removed, descriptors $\mathcal{F}_k$ are also invariant to translation and rotation. A subset of Fourier coefficients with $-10 \leq k \leq 10$, $k \neq 0, 1$ is used as a signature to characterize the lesion's shape. An exemplary case is shown in Figure 3. Coefficients with relatively small values of $|k|$ represent low frequency components of the contour function and, thereby, deviations from a circular shape, whereas coefficients with relatively large values of $|k|$ represent high frequency components and thereby fast changes, *e.g.*, notches or rough edges, in the contour function.

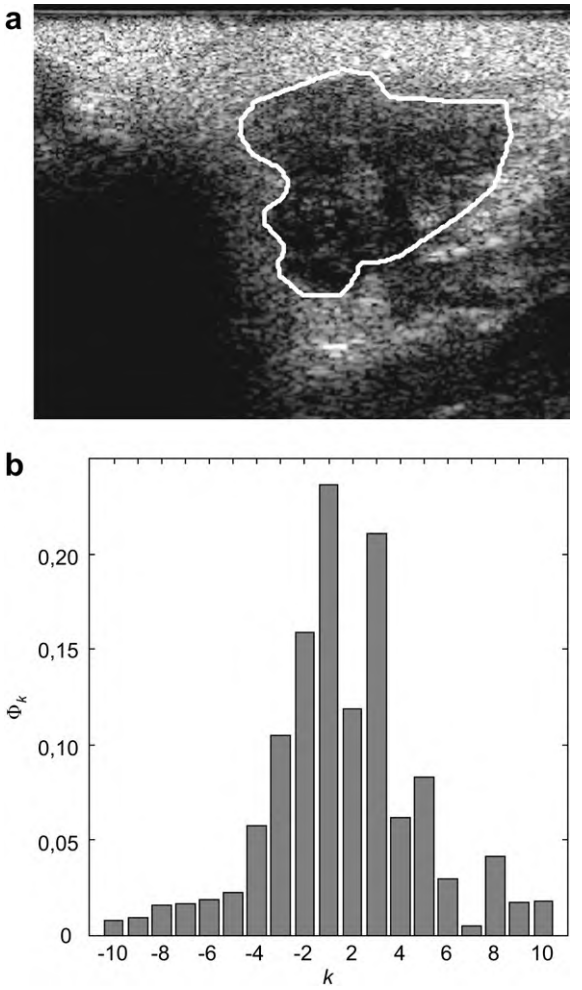

Fig. 3. 52-year-old female patient with an irregular shaped acinar cell carcinoma. (a) B-mode image with contoured tumor. (b) Normalized Fourier descriptors calculated from contour.

Maximum likelihood classification

A supervised classification approach based on a statistical maximum likelihood classifier and leave-one-out cross validation is employed in this work. This methodology is also utilized for heuristic feature selection. According to the pathohistologic results (see Results section), the acquired data sets were subdivided into two groups, Ω_{pos} and Ω_{neg}, comprising all positive and negative cases, respectively.

To yield a decision rule based on a maximum likelihood approach, clusters in feature space are modeled as multivariate normally distributed probability density functions. A class-dependent probability density

function of a cluster in the $N-$ dimensional feature space is given as

$$f(\mathbf{p}|\Omega_q) = \frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma_q|^{\frac{1}{2}}} \cdot e^{-\frac{1}{2}(\mathbf{p}-\overline{\mathbf{p}}_q)^T \cdot \sum_q^{-1} \cdot (\mathbf{p}-\overline{\mathbf{p}}_q)}, \quad (8)$$

where $\mathbf{p}$ is the feature vector to be classified, $\overline{\mathbf{p}}_q$ is the empirical expectation value of cluster q denoting the center of the cluster, and $\mathbf{\Sigma_q}$ is the covariance matrix of cluster q. Taking the logarithm of eqn (8) and eliminating the term $(2\pi)^{N/2}$ yields the discriminant function of the maximum likelihood classifier:

$$d_q(\mathbf{p}) = -\frac{1}{2} \cdot ln(|\Sigma_q|) - \frac{1}{2}(\mathbf{p}-\overline{\mathbf{p}}_q)^T \cdot \sum_q^{-1} \cdot (\mathbf{p}-\overline{\mathbf{p}}_q). \quad (9)$$

In the case of q different classes, the classification of feature vector $\mathbf{p}$ is done by determining the likelihoods associated with each class and assigning the class Ω_m with the maximum likelihood according to the decision rule

$$e^{ML}(\boldsymbol{p}) = \Omega_m \Big| d_m = max(d_n)_{n=1..q} \ \forall \ n \neq m. \quad (10)$$

In the case of a two class problem as in this work (differentiation between positive and negative cases), eqn (10) simplifies to

$$e^{ML}(\mathbf{p}) = \begin{cases} \Omega_{neg} | d_{neg} > d_{pos} \\ \Omega_{pos} | d_{pos} > d_{neg} \end{cases}. \quad (11)$$

Note that a feature vector is classified as positive if d_{neg} equals d_{pos}, since false positives are rather acceptable than false negatives in the case of tumor classification.

Leave-one-out cross validation

The discriminant functions d_{pos} and d_{neg} are calculated for each feature vector $\mathbf{p}$ representing one case in the database. Thus, each case was labeled either positive or negative. Training data included all cases except the one under consideration. As a reference, results from pathohistologic investigations are used (see clinical results subsection). Sensitivity *SE* and specificity *SP* are determined as the fraction of true positive cases by the total number of positive cases and the fraction of true negative cases by the total number of negative cases, respectively. Both can be varied arbitrarily at each other's expense by introducing a scaling factor k to the determination of e^{ML} in eqn (11), thus, shifting the decision threshold. This enables to generate receiver operating characteristic (ROC) curves by plotting sensitivity versus specificity. The area under the ROC curve A_{ROC} is employed as a measure for the classifiers performance, where an area $A_{ROC} = 1$ implies a perfect separation between both classes (Scheipers et al. 2005a).

Feature selection

The selection of appropriate features is a significant step in a classification process since only few features contribute to the classifier's performance. In this work, a stepwise heuristic scheme is carried out to yield a set of features suitable for classification. In the first step, all features are separately evaluated and sorted by the area A_{ROC}. Highly correlated features with a correlation coefficient $\rho > 0.9$ are eliminated from the feature space since they usually do not contribute to the classification performance. The next classification cycle is accomplished with pairwise combinations of the best feature and all remaining features. The best pair of features is then combined with all remaining features and so on. This procedure is repeated until no more increase in A_{ROC} is observed.

RESULTS

Clinical results

The current clinical study comprises 138 cases originating from 138 patients scheduled to have parotid surgery shortly after the diagnosis. Resected lesions were analyzed in the pathohistologic department of the same hospital. Overall, 13 different kinds of benign and malignant tumors or tumor-like lesions occurred during the study. Calcifications were not observed. Physicians closely examined all lesions. Characteristic sonographic features are best seen in the lesion of maximum diameter, which was chosen for data acquisition. Over 80% of lesions were singular occurring alterations of the parotid gland.

The lesion size ranged from 0.81 cm to 3.76 cm in diameter with a mean of 2.11 cm. The number of segments used for spectral estimation ranged between 40 and 1348 per case (mean 303), thus, allowing for reliable determination of features. Cases were subdivided into two classes, Ω_{pos} and Ω_{neg}. Class Ω_{pos} contained all carcinomas, lymphomas and metastases as well as the relatively large group of pleomorphic adenomas. Note that pleomorphic adenomas are actually benign lesions. However, they can lead to malignancy in a later stage of disease and are therefore treated the same as malignant lesions. Class Ω_{neg} contained all remaining benign cases, including the relatively large group of Warthin's tumors. The occurrence of different types of parotid gland lesions during this study as well as the subdivision into two target classes are summarized in Table 1.

The diagnoses of experienced physicians were recorded during the common medical inspections using B-mode ultrasound, manual palpation and occasional Doppler or power Doppler ultrasound. The physicians

typed the inspected lesions according to the World Health Organization (WHO) classification of parotid gland tumors. However, to allow a fair comparison of the conventional diagnoses and the results of the automated classification, the same two-class-scheme was applied as depicted in Table 1, yielding 45 out of 51 correctly typed positive cases and 54 out of 87 correctly typed negative cases. This corresponds to a sensitivity and specificity for the conventional diagnosis of $SE = 0.88$ and $SP = 0.62$, respectively.

Feature extraction and classification

The above described selection algorithm yielded a subset of 10 features. This set was used as a feature vector for the automated classification. The selected features are summarized in Table 2. Two backscatter measures, six attenuation measures, one contour descriptor and one co-occurrence measure were chosen. The increase in classification performance depending on the number of used features is shown in Figure 4a. A classification using the first feature exclusively yielded an area under the ROC curve of $A_{ROC} = 0.73$. Adding the next features successively to the feature vector led to a monotonic increase in A_{ROC}. The best result was obtained using 10 features, achieving an area under the ROC curve of $A_{ROC} = 0.91$. The resulting ROC curve is shown in Figure 4b. The ROC curve can be used to choose an operating point for the classification system. Thus, an arbitrary sensitivity at the expense of specificity, or vice versa, can be achieved. The classification result reached by the experienced physicians using conventional diagnostics is also depicted in Figure 4b. Since the physicians implicitly used a two-class scheme for lesion typing instead of providing malignancy probabilities for each case, no ROC curve could be generated. However, the physicians' results can be compared with the respective points on the ROC-curve. If the operating point is chosen such that the physicians' sensitivity $SE = 0.88$ is achieved, the specificity is $SP = 0.74$. The respective point on the curve for the physicians' specificity $SP = 0.62$ indicates a sensitivity of 0.96. Hence, the results obtained by conventional diagnostics could be significantly exceeded. However, in the given context of typing parotid gland lesions, the system should facilitate the detection of as many positive cases as possible. From the ROC curve, it can be seen

Table 2. Selected features

No.	Feature	Origin	A_{ROC}
1	b^{S}_{μ}	Spectral backscatter	0.73
2	$\mathcal{F}_{-5}$	Shape descriptor	0.62
3	$\alpha_{1,\ \sigma}^{mnb}$	Spectral attenuation	0.67
4	$\alpha_{1,\ fwhm}^{cfs}$	Spectral attenuation	0.51
5	MCC	Texture	0.56
6	$\alpha(f_0)_{\mu}$	Spectral attenuation	0.64
7	$\alpha_{1,\ sk}^{mnb}$	Spectral attenuation	0.59
8	$\alpha_{1,\ fwhm}^{mnb}$	Spectral attenuation	0.57
9	$\alpha_{1,\ ku}^{cfs}$	Spectral attenuation	0.56
10	b^{I}_{ku}	Spectral backscatter	0.52

Indices used with spectral parameters indicate the following measures: mean (μ), standard deviation (σ), full width at half maximum (*fwhm*), skewness (*sk*) and kurtosis (*ku*). Features were chosen by the selection algorithm in the same order as in the table. The last row indicates values of A_{ROC} when the particular feature was used exclusively for classification.

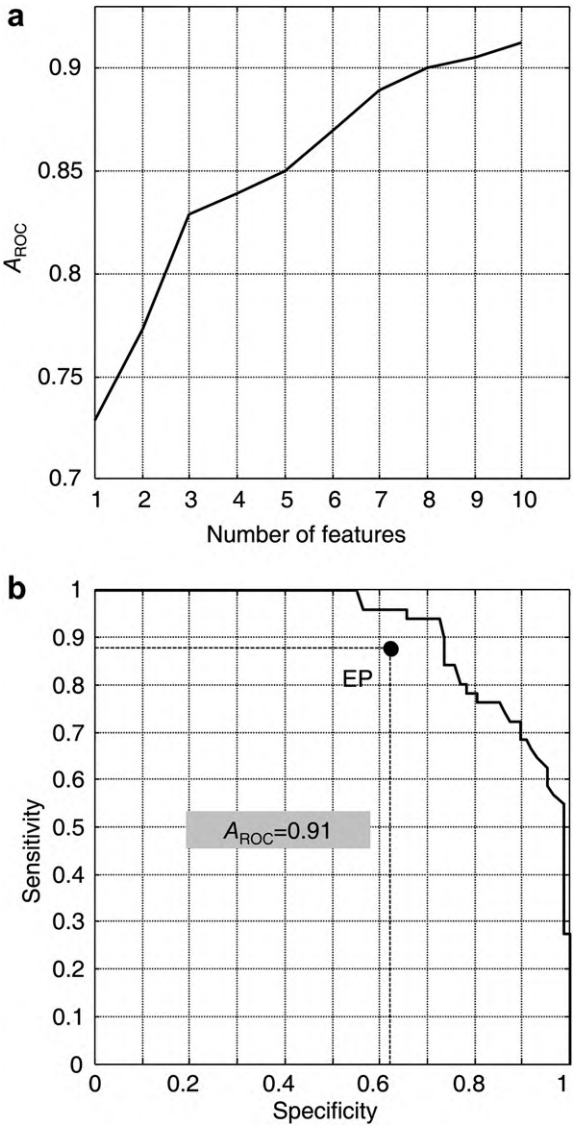

Fig. 4. Classification results. (a) A_{ROC} as a function of feature number. The graph shows the increase in A_{ROC}, starting with the first feature b^{S}_{μ}, and adding features 2–10 successively (see Table 2). (b) ROC-curve using the complete set of 10 features. The area under the ROC-curve A_{ROC} is 0.91. For comparison, the result achieved by experienced physicians (EP) using conventional diagnostics is also marked.

that for a sensitivity $SE = 1$, the remaining specificity is $SP = 0.55$. With a sensitivity of 1, no false negatives occur, thus, all cases classified as negative are true negatives.

DISCUSSION AND SUMMARY

In this article, a sonohistology system for computer assisted characterization of parotid gland lesions is presented. Overall, 138 patients were examined in a clinical study. Although the number of cases is yet too small to draw general conclusions, the results presented here are promising. The proposed computer assisted method achieved classification rates exceeding those of experienced physicians using common diagnostic modalities. Furthermore, several properties of the system might be valuable for prospective applications:

- The system is adjustable to achieve an arbitrary sensitivity at the expense of specificity.
- Except for the demarcation of lesions in the B-mode image the system is user independent.
- The classification rates of the system will possibly increase if a larger training data set is available. If so, the data set may be subdivided into more than two target classes to consider different kinds of benign and malignant lesions.
- The system can be easily expanded by adding other suitable features. In this work, we exclusively used spectral features and their statistical distributions inside the ROIs, textural co-occurrence features and shape descriptors. Other suitable features may be achieved using further ultrasound signal and image processing techniques, as well as different diagnostic imaging modalities or clinical parameters in a multimodal approach.

However, in some aspects the proposed method is yet limited:

- The method may fail if the lesion size is very small. In particular, the methods for attenuation estimation used are restricted to an axial resolution of 6.75 mm. The lesions occurred in this study were large enough to provide reasonable estimates. Another issue exists with shape descriptors, if a very large lesion exceeds the field of view of the transducer and thus the boundary of the lesion can not be determined.
- Although the acquired data have been calibrated to correct partly for diffraction and influences of the ultrasound system, the proposed method may not be system independent, *i.e.*, for a different ultrasound system, building a dedicated database as well as a specific feature selection is necessary.
- Several simplifications have been used in the spectral estimations due to limitations in the clinical setting. In particular, it was not possible to actually measure attenuation and backscatter from resected lesions. The approach for attenuation correction using a fixed attenuation causes a bias in the backscatter estimates. Nevertheless, the features used in this work were apparently suitable to differentiate between the two target classes.
- Compared with the number of diverse tumor and lesion types occurring in the parotid gland, the number of cases included in this study is relatively small. To keep the number of cases in each class reasonably high and to account for the curse of dimensionality, the number of target classes was limited to two, thus only two tumor entities, malignancy and benignity, were considered, leading to rather heterogeneous target classes.

Acknowledgments—This work is an activity of the Ruhr Center of Excellence for Medical Engineering (Kompetenzzentrum Medizintechnik Ruhr, KMR) Bochum. It was supported by the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung), grant No. 13N8079 (2003-2006). Additional support was given by the Ruhr Center for Medical Imaging (zmb ruhr), Bochum, with funding from the European Community and the State Government of North-Rhine-Westphalia, Germany (2007-2008).

REFERENCES

Bozzato A, Zenk J, Greess H, Hornung J, Gottwald F, Rabe C, Iro H. Potential of ultrasound diagnosis for parotid tumors: Analysis of qualitative and quantitative parameters. Otolaryngol Head Neck Surg 2007;137:642−646.

Bridal S, Fornes P, Bruneval P, Berger G. Parametric (integrated backscatter and attenuation) images constructed using backscattered radio frequency signals (25-56 MHz) from human aortae *in vitro*. Ultrasound Med Biol 1997;23:215−229.

Chang R, Wu W, Moon W, Chen D. Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis. Ultrasound Med Biol 2003;29:679−686.

Chen C, Chen L, Millero F. Speed of sound in NaCl, $MgCl_2$, $Na2So_4$ and $MgSo_4$ aqueous solutions as function of concentration, temperature and pressure. J Acoust Soc Am 1978;63:1795−1800.

Chen D, Chang R, Kuo W, Chen M, Huang Y. Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. Ultrasound Med Biol 2002;28:1301−1310.

Chen J, Zagzebski J, Madsen E. Tests of backscatter coefficient measurement using broadband pulses. IEEE Trans Ultrason Ferroelectr Freq Control 1993;40:603−607.

Chikui T, Tokumori K, Yoshiura K, Oobu K, Nakamura S, Nakamura K. Sonographic texture characterization of salivary gland tumors by fractal analyses. Ultrasound Med Biol 2005; 31:1297−1304.

Cloostermans M, Thijssen J. A beam corrected estimation of the frequency dependent attenuation of biological tissues from backscattered ultrasound. Ultrason Imaging 1983;5:136−147.

Duck F. Physical properties of tissue. London: Academic Press Ltd; 1990.

Feleppa E, Ennis R, Schiff P, Wuu C, Kalisz A, Ketterling J, Urban S, Liu T, Fair W, Porter C, Gillespie J. Spectrum-analysis and neural networks for imaging to detect and treat prostate cancer. Ultrason Imaging 2001;23:135−146.

Fink M, Hottier F, Cardoso J. Ultrasonic signal processing for *in vivo* attenuation measurement: Short-time Fourier analysis. Ultrason Imaging 1983;5:117−135.

Gaitini D, Baruch Y, Ghersin E, Veitsman E, Kerner H, Shalem B, Yaniv G, Sarfaty C, Azhari H. Feasibility study of ultrasonic fatty

liver biopsy: Texture vs. attenuation and backscatter. Ultrasound Med Biol 2004;30:1321−1327.
Granlund G. Fourier preprocessing for hand print character recognition. IEEE Trans Comput 1972;21:195−201.
Haralick R, Shanmugam K, Dinstein I. Textural features for image classification. IEEE Trans Syst Man Cybern 1973;SMC-3:610−622.
Jago J. Experimental and theoretical studies of ultrasound computed tomography. Ph.D. thesis, Faculty of Medicine, University of New Castle upon Tyne, 1993.
Jenderka K, Gaertner T, Zacharias M, Heynemann H, Cobet U. System independent tissue typing of human testis and prostate. IEEE Ultrason Symp 1999;2:1377−1380.
Kolecki R, Sigel B, Justin J, Feleppa E, Parsons R, Kitamura H, Machi J, Hayashi J, Taylor P, McGann L. Determining the acuteness and stability of deep venous thrombosis by ultrasonic tissue characterization. J Vasc Surg 1995;21:976−984.
Kolios MC, Czarnota GJ, Lee M, Hunt JW, Sherar MD. Ultrasonic spectral parameter characterization of apoptosis. Ultrasound Med Biol 2002;28:589−597.
Kuc R, Schwartz M. Estimating the acoustic attenuation coefficient slope for liver from reflected ultrasound signals. IEEE Trans Son Ultrason 1979;26:353−361.
Lang M, Ermert H, Heuser L. *In vivo* study of online liver tissue classification based on envelope power spectrum analysis. Ultrason Imaging 1994;16:77−86.
Lemor R, Kleffner B, Tretbar S, Schmitt R. Ultrasound temperature and attenuation monitoring for controlling the laser induced thermo therapy. Acoust Imaging 2000;25:395−400.
Lizzi F, Astor M, Feleppa E, Shao M, Kalisz A. Statistical framework for ultrasonic spectral parameter imaging. Ultrasound Med Biol 1997a; 23:1371−1382.
Lizzi F, Feleppa E, Astor M, Kalisz A. Statistics of ultrasonic spectral parameters for prostate and liver examinations. IEEE Trans Ultrason Ferroelectr Freq Control 1997b;44:935−942.
Lizzi F, Laviola M. Power spectra measurements of ultrasonic backscatter from ocular tissues. IEEE Ultrason Symp 1975;29−32.
Lizzi F, Ostromogilsky M, Feleppa E, Rorke M, Yaremko M. Relationship of ultrasonic spectral parameters to features of tissue microstructure. IEEE Trans Ultrason Ferroelectr Freq Control 1987;34:319−329.
Lizzi FL, Greenebaum M, Feleppa EJ, Elbaum M, Coleman DJ. Theoretical framework for spectrum analysis in ultrasonic tissue characterization. J Acoust Soc Am 1983;73:1366−1373.
Madsen E, Insana M, Zagzebski JA. Method of data reduction for accurate determination of acoustic backscatter coefficients. J Acoust Soc Am 1984;76:913−923.
Noritomi T, Sigel B, Swami V, Justin J, Gahtan V, Chen X, Feleppa EJ, Roberts AB, Shirouzu K. Carotid plaque typing by multiple-parameter ultrasonic tissue characterization. Ultrasound Med Biol 1997;23:643−650.
O'Donnell M, Mimbs JW, Miller JG. Relationship between collagen and ultrasonic backscatter in myocardial tissue. J Acoust Soc Am 1981; 69:580−588.
Oosterveld B, Thijssen J, Hartman P, Romijn R, Rosenbusch G. Ultrasound attenuation and texture analysis of diffuse liver disease: Methods and preliminary results. Phys Med Biol 1991;36:1039−1064.
Parsons R, Sigel B, Feleppa E, Golub R, Kodama I, Loiacono L, Justin J, Swami V, Kimitsuki H, Rorke M. Age determination of experimental venous thrombi by ultrasonic tissue characterization. J Vasc Surg 1993;17:470−478.
Rogers P, Pitzer K. Volumetric properties of aqueous sodium chloride solutions. J Phys Chem Ref Data 1982;11:15−81.
Scheipers U, Ermert H, Sommerfeld H, Garcia-Schürmann M, Senge T, Philippou S. Ultrasonic multi-feature tissue characterization for prostate diagnostics. Ultrasound Med Biol 2003a;29:1137−1149.
Scheipers U, Ermert H, Sommerfeld HJ, Garcia-Schürmann M, Kühne K, Senge T, Philippou S. Ultrasonic tissue characterization for prostate diagnostics: Spectral parameters vs. texture parameters. Biomed Tech 2003b;48:122−129.
Scheipers U, Perrey C, Siebers S, Hansen C, Ermert H. A tutorial on the use of ROC analysis for computer aided diagnostic systems. Ultrason Imaging 2005a;27:181−198.
Scheipers U, Siebers S, Gottwald F, Ashfaq M, Bozzato A, Zenk J, Iro H, Ermert H. Sonohistology for the computerized differentiation of parotid gland tumors. Ultrasound Med Biol 2005b;31:1287−1296.
Schmitz G, Ermert H, Senge T. Tissue-characterization of the prostate using radio frequency ultrasonic signals. IEEE Trans Ultrason Ferroelectr Freq Control 1999;46:126−138.
Shen L, Shen L, Rangayyan R, Desautels J. Application of shape analysis to mammographic calcifications. IEEE Trans Med Imaging 1994;13:263−274.
Siebers S, Geier B, Scheipers U, Vogt M, Mumme A, Ermert H. Classification of venous thrombosis combining ultrasound elastography and tissue characterization. IEEE Ultrason Symp 2004;3:1761−1764.
Siebers S, Scheipers U, Ashfaq M, Hansler J, Frieser M, Strobel D, Hahn E, Ermert H. *In vivo* imaging of coagulated tissue. IEEE Ultrason Symp 2006;1762−1765.
Siebers S, Scheipers U, Grosse D, Gottwald F, Bozzato A, Zenk J, Iro H, Ermert H. Differential diagnosis of parotid gland lesions using spatially fused sonohistologic features. IEEE Ultrason Symp 2007;456−459.
Siebers S, Scheipers U, Hänsler J, Frieser M, Strobel D, Welp C, Werner J, Hahn E, Ermert H. In: Andre M, (ed). Acoustical imaging. New York: Kluwer Academic/Plenum Publisher; 2007b. p. 295−300.
Smutek D, Sara R, Sucharda P, Tjahjadi T, Svec M. Image texture analysis of sonograms in chronic inflammations of thyroid gland. Ultrasound Med Biol 2003;29:1531−1543.
Sonka M, Hlavac V, Boyle R. Image processing, analysis and machine vision. 2nd edition. Pacific Grove, CA: PWS Publishing; 1998.
Timmermann W, Hamelmann W, Thiede A. Schilddrüsenchirurgie: Neuromonitoring zur schonung des nervus recurrens. Dtsch Arztebl 2004;101:A:1341−1345.
Tsantis S, Cavouras D, Kalatzis I, Piliouras N, Dimitropoulos N, Nikiforidis G. Development of a support vector machine-based image analysis system for assessing the thyroid nodule malignancy risk on ultrasound. Ultrasound Med Biol 2005;31:1451−1459.
Zahn C, Roskies R. Fourier descriptors for plane closed curves. IEEE Trans Comput 1972;C-21:269−281.