Ultrasonic Tissue Characterization for the Differentiation of Parotid Gland Tumors

Ulrich Scheipers[1], Stefan Siebers[1], Mohammad Ashfaq[1],
Frank Gottwald[2], Alessandro Bozzato[2], Johannes Zenk[2], Heinrich Iro[2], Helmut Ermert[1]
[1] Institute of High-Frequency Engineering, Ruhr-University Bochum, Bochum, Germany
[2] University Hospital of Otorhinolaryngology, University Erlangen, Erlangen, Germany
Ulrich.Scheipers@ResonantMedical.com

Abstract— **The first ultrasonic tissue characterization system for the computerized differentiation of tumors of the parotid gland is presented. The system is based on a multifeature tissue characterization approach involving spectrum and texture parameters and using fuzzy inference systems as higher order classifiers.**

Baseband ultrasound echo data were acquired during conventional ultrasound imaging examinations of the salivary glands. Several tissue-describing parameters were calculated within numerous small regions of interest in order to evaluate local spectral and textural tissue properties. The parameters were processed by an adaptive network-based fuzzy inference system using the results of conventional histology after parotidectomy as the gold standard. Cases of parotid gland tumors and alterations include basal cell adenomas, monomorphic adenomas, pleomorphic adenomas, adenoid cysts, cysts and canaliculous adenomas. The results of the classification procedure are presented as a numerical score indicating the probability of a certain tumor or alteration for each parotid gland.

In a pilot study, the system was evaluated on 23 cases of benign and malignant parotid gland tumors of patients undergoing parotidectomy. The ROC curve area given as the cross-validation mean and cross-validation standard deviation is A_{ROC}=0.95±0.07 when using four-fold cross-validation over cases and differentiating between various malignant and benign parotid gland tumors as the positive target group and monomorphic adenomas as the negative target group. An exceptional equal error rate of E_{EER}=0.92±0.08 is achieved for the same setup. Some alterations which are of benign nature were counted to the positive group, as they occur too seldom to achieve a high probability for being considered safe if left untreated.

Keywords: computer aided diagnostics, parotid gland, tumor diagnostics, cancer diagnostics, tissue characterization, tissue tying, parametric imaging, head and neck, otorhinolaryngology, fuzzy inference systems, nonlinear classification, salivary glands.

I. INTRODUCTION

The incidence of parotid gland tumors is one of the highest of all incidents of tumors in otorhinolaryngology. In some cases, differentiation between benign and malignant lesions is difficult. In addition, some types of benign parotid gland tumors, such as the pleomorphic adenoma, may become malignant at later stages of the disease. Therefore, the possibility of early detection, differentiation and excision is important. Usually, B-mode medical ultrasound is applied as the main diagnostic modality in order to differentiate among the various types of parotid gland tumors. Two typical examples of tumors of the parotid gland are shown in Figure 1.

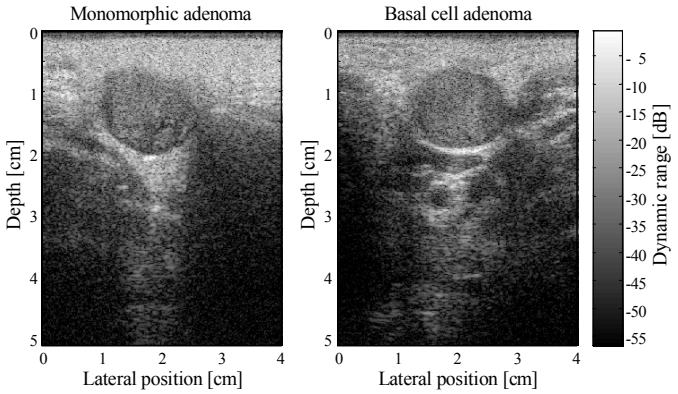

Figure 1. Exemplary B-mode images of a negative case (monomorphic adenoma) on the left hand side and a positive case (basal cell adenoma) on the right hand side. Reliable differentiation between different types of parotid gland tumors and parotid gland alterations requires experienced physicians.

The current methods of diagnostics of parotid gland tumors, including medical ultrasound and MRI, lack accuracy, especially when it comes to differentiate between benign and malignant forms of alterations [1]. Currently, only the histopathologic examination of tissue specimen can provide the information needed for a definite diagnosis. The application of ultrasonic tissue characterization can add additional information to the currently available methods of diagnostics.

As the different parameters used in this approach have a highly nonlinear interdependence, only a nonlinear model is able to combine the parameters and, thus, lead to reliable classification results [2-4]. In this approach, an adaptive network-based fuzzy inference system (FIS) was used [5, 6].

As the information of tissue parameters of different parameter groups (e.g. spectral parameters, first-order texture parameters or second-order texture parameters) is highly uncorrelated, the combination of tissue parameters originating from different parameter groups can lead to better classification results [7-10]. In most cases of computer-aided diagnostics only a combination of different groups of parameters, e.g. spectral and textural parameters, can provide the classification system with enough information to support an accurate and reliable decision.

II. METHODS

A. Data Acquisition

Baseband ultrasound echo data of the parotid gland were captured during the routine examination of the patient using standard ultrasound imaging equipment. Patient compliance to the procedure was high, as the new method does not extend the normal examination time when applying ultrasound imaging of the head and neck region. The proposed system is operator independent, which means that no special knowledge or training is necessary for a successful application of the system.

For the clinical study, a Siemens Elegra digital ultrasound scanner with enabled research interface was used. The linear probe (7.5L40) was set to a center frequency of 7.2 MHz. For each frame, 2400 samples were recorded for 360 lines. The approximate size of the images was 5.1 cm in axial direction and 4 cm in lateral direction, respectively. One transmit focus was applied using a depth of 2 cm. Time-gain compensation was set to a neutral position. The dynamic low pass filter and the frequency of the dynamic local oscillator were set to a constant value. Radio frequency echo data were acquired at 36 MHz and 12 bits. Baseband data were provided by the research interface for download to a personal computer.

Two orthogonal frames per lesion were recorded and every data frame was subdivided into numerous regions of interest (ROI). Each ROI comprises an area of approximately 2.7 mm in axial direction and 3.5 mm in lateral direction. Thus, the smallest region that can be analyzed by the system covers approximately 9.5 mm^2. The ROIs consist of 128 sample points in the axial direction and of 16 scan lines in the lateral direction. The axial and lateral overlaps of the ROIs were 75 % and 50 %, respectively. For attenuation measurements, adjacent ROIs were combined.

ROIs were transformed into frequency domain using Fourier transform on every scan line of the ROI. Before applying the Fourier transform, all ROIs were windowed by a Hamming window of the ROIs length to avoid spectral leakage. The baseband data were compensated for system and depth-dependent effects using the system transfer function over depth as an inverse filter within the effective bandwidth. Using this approach, system effects due to focusing and electro-mechanical characteristics of the transducer can be partly compensated [8, 11].

B. Parameter extraction

Several tissue-describing parameters were calculated for each ROI. The parameters used for classification were calculated from the frequency spectrum and from the time domain, partly before and partly after compression and envelope detection of the baseband data.

Spectrum parameters were calculated after logarithmic compression of the power spectrum. Spectral results of each scan line were averaged to form an estimate of the average power spectrum [8, 12]. The primary set of spectrum parameters consisted of five measures of backscatter calculated for the signal bandwidth. The parameters used in this approach were: axis intercept, slope, midband value, deviation and normalized square deviation of the linear regression spectrum fit [2-4, 13]. Three attenuation parameters were additionally used in this approach: axis intercept, slope and midband value. These three parameters were derived from the frequency dependent attenuation coefficient over depth.

Texture parameters evaluated in this work consist of first and second-order (i.e. cooccurrence) parameters. Texture parameters were calculated after envelope detection of the complex baseband data using rectification. First-order texture parameters consist of eight different estimates of echo amplitude: maximum, minimum, mean, variance, kurtosis, signal to noise ratio, ratio of squares and the full width at half maximum of the gray level histogram.

Cooccurrence parameters were calculated in the spatial domain of demodulated data for different distances [2-4, 14]. Sizes of cooccurrence matrices, i.e. the number of gray levels incorporated, were varied in the range of 16 to 64. The nine second-order texture parameters evaluated in this approach were: angular second moment, contrast, correlation, dimension, inverse difference moment, kappa, peak density, variance and the signal to noise ratio of cooccurrence matrices. The lateral resolution of ultrasound data changes over depth, thus leading to an increase in lateral speckle size for deeper imaging positions. Gray-tone spatial dependency matrices were calculated only in axial direction in order to keep effects of focusing and diffraction as low as possible.

C. Selection of parameters

A preselection of parameters was made by covariance matrix analysis. Parameter vectors that were highly linearly dependent on others were found and discarded. During the selection procedure, the number of tissue parameters was reduced to six using a stepwise selection algorithm based on hypothesis testing. The parameter selection procedure starts by calculating the classification performance of each single tissue parameter using four-fold cross-validation over cases. Four-fold cross validation was performed by dividing the whole number of cases into four datasets. Three datasets were combined and used as the training data while the remaining fourth dataset was used as the test data. Each iteration of the classification system is performed four times with all different combinations of datasets involved. While dividing the whole data into four separate datasets, strict separation between patients was reserved in order to keep the results unbiased. The overall performance of a classification system, that is not dependent on a separation threshold or a cut-off point, can be provided using the area A_{ROC} under the ROC curve as a measure.

After the area under the ROC curve was calculated for every tissue parameter, the tissue parameter with the largest area A_{ROC} was chosen as the first feature of choice. During the next step, this parameter was combined with all other remaining tissue parameters and the parameters of the pair with the largest area A_{ROC} were then chosen as the features of choice. This procedure was repeated until the area A_{ROC} decreases or stalls as the total number of parameters increases.

The number of ROIs used in this work was relatively high. Thus, using up to six tissue parameters for the classification procedure was still a safe approach [15]. The distributions of

the tissue parameters used in the classification procedure could be estimated with a sufficiently low variance, although the number of ROIs that were completely independent of each other may be smaller than the total number of ROIs of this approach, as the ROIs overlap and as several ROIs usually originate from the same case.

Adaptive network-based fuzzy inference systems were used to classify and separate the ROIs into two classes. The fuzzy inference systems used in this work are based on first-order Sugeno-type systems with Gaussian membership functions and up to eight rules to model the feature space [16]. The number of rules was adaptively chosen by the system. Subtractive clustering was used as the initial step in the supervised learning procedure to find natural clusters in the data space.

A hybrid, adaptive-training algorithm based on backpropagation and least-square error estimation with adaptive step sizes was used to train the system. Batch learning was applied to set the width of the Gaussians and to refine their centre position. During the training procedure, the best combination of tissue parameters and the appropriate membership functions were stored in a rule base, which was used during the evaluation procedure and during further application of the system.

D. Post processing

The fuzzy output maps of the network-based fuzzy inference systems were transformed into binary maps applying a separation threshold. The separation threshold is used to separate the quasi-continuous output vectors into two classes or target groups. Usually, the optimum separation threshold can be estimated by the classification system. However, fine-tuning the separation threshold to maximize either sensitivity or specificity is possible. The binary output maps were averaged to provide a mean decision criterion for the complete parotid gland. The decision criterion can be scaled according to probabilities for certain target groups.

III. Clinical Study

During three and a half months, between March and July 2004, ultrasound baseband data of 23 parotid glands originating from 18 patients, were recorded. The parotid glands were contoured in the ultrasound B-mode images using custom software in order to make sure that only signals originating from the glands and not from surrounding tissue were analyzed. All patients were scheduled to have parotid surgery during the following week after acquisition of the ultrasound data.

At the day of the examination, the youngest patient was 25 years old, the eldest 88 years. The mean age of the patients was 65 years while the median age was 70 years. Overall, 11 of the patients were female while 7 of the patients were male. Histopathological examinations following parotidectomy were used as the gold standard.

In this first approach, the system was trained to differentiate between a first target group containing all cases of monomorphic adenoma and a second target group containing all cases of other types of parotid gland alterations. The second target group contained all non-monomorphic adenoma types of parotid gland alterations, which occurred during the clinical study, not making any difference between benign and malignant parotid gland alterations. The first target group was called ‘negative’, as the incident of monomorphic adenoma is considered benign. The second target group was called ‘positive’, as all other types of tumors and alterations of the parotid gland would obviously be considered as malignant. Even those diseases which are actually of benign nature were counted to the second group, as they occur too seldom to achieve a high probability for being considered safe if left untreated.

The baseband datasets were divided into numerous ROIs as described above resulting in a sum of 6,127 negative and 8,444 positive ROIs.

IV. Results

Six tissue describing parameters were automatically chosen by the classification system. Three parameters originate from the second-order texture parameter group: contrast, correlation and dimension. Two spectral parameters were chosen: midband value and slope. In addition, one parameter of the first-order texture parameter group was included: variance. The optimum number of rules of the network-based fuzzy inference system was iteratively found to be two.

The fuzzy inference system was trained using the histological findings as the gold standard or teacher data. The fuzzy inference system provides a fuzzy value for each ROI of the ultrasound dataset. The fuzzy value is a measure of the probability of the ROI belonging to the positive or negative target group.

The area under the ROC curve given as the cross-validation mean and the cross-validation standard deviation is $A_{ROC}=0.95\pm0.07$ when using four-fold cross-validation over cases and differentiating between monomorphic adenoma as the first target group and all other types of parotid gland alterations as the second target group. In two of the four cross-validation cases, exceptional ideal classifications of $A_{ROC}=1\pm0$ were achieved.

Intermediate results were evaluated for comparison to the final classification results. Without the post-processing step, which provides the final score for each case, the area under the ROC curve given as the cross-validation mean and the cross-validation standard deviation is $A^*_{ROC}=0.71\pm0.07$, where the asterisk denotes the intermediate status of the results. In all estimations, the ROC curve area was calculated by continuously varying the separation threshold.

All classification results are presented in Table 1. In addition to the area under the ROC curve A_{ROC}, the standard error of the ROC estimate SE_{ROC} and the equal error rate E_{EER} are shown in the table. The standard error was estimated over the number of ROIs involved in the study. As the ROIs are correlated with each other to an undefined degree, the estimations of standard error tend to underestimate the true error. The equal error rate is the sensitivity or specificity of the system at the operating point where sensitivity equals specificity. For the final system, an equal error rate of $E_{EER}=0.92\pm0.08$ was achieved. Intermediate results yield an equal error rate of $E^*_{EER}=0.66\pm0.05$.

TABLE I. OVERVIEW OF CLASSIFICATION RESULTS

Cross-validation set	Intermediate results estimated over ROIs			Final results estimated over cases		
	A^*_{ROC}	SE^*_{ROC}	E^*_{EER}	A_{ROC}	SE_{ROC}	E_{EER}
1 / 4	0.80	0.01	0.73	0.96	0.00	0.85
2 / 4	0.64	0.02	0.60	0.83	0.01	0.83
3 / 4	0.65	0.01	0.62	1.00	0.00	1.00
4 / 4	0.75	0.01	0.68	1.00	0.00	1.00
Mean ± StdDev	0.71 ±0.07		0.66 ±0.05	0.95 ±0.07		0.92 ±0.08

During the clinical study, diagnoses of an experienced physician were recorded for each case before the ultrasound data of the cases was processed by the classification system. The experienced physician achieved a sensitivity of SE_P=0.67 and a specificity of SP_P=0.79. However, as the database is comparably small, the performance of the classification system can only be compared rudimentary with the results of the experienced physician. At a sensitivity of SE=0.67 the final classification system achieves a specificity of SP=0.96±0.08. At a specificity of SP=0.79, the final classification system achieves a sensitivity of SE=0.98±0.04, respectively.

V. DISCUSSION

Although only a relatively small database consisting of n=23 cases was evaluated in this work, the system for ultrasonic multifeature tissue characterization differentiates between monomorphic adenomas and all other types of alterations of parotid glands with a satisfying grade of accuracy. The area under the ROC curve is A_{ROC}= 0.95±0.07 when using four-fold cross validation methods to evaluate the underlying data. In two of the four cross-validation cases, exceptional ideal classifications were performed. The classification rates will possibly increase if a larger database is available. Several alterations of parotid glands only occurred once during the clinical study, while a total of 14 cases of monomorphic adenoma could be counted.

The classification system seems to have learned the typical characteristics of the monomorphic adenoma sufficiently well. However, a differentiation between the other types of tumors would be interesting. This will be evaluated in the future together with leave-one-out cross-validation tests, when a larger database will be available.

The interobserver variability is a problem of conventional B-mode ultrasound for the differentiation of diverse parotid gland alterations. The interobserver variability is the dependence of the diagnostic results of the experience and expert-knowledge of the conducting physician. A highly experienced physician may be able to detect the features of malignant tumors while a novice physician may not be able to evaluate the echographic properties sufficiently. An ultrasonic tissue characterization system can automate the process of differentiating between various types of parotid gland tumors and hence may assist to reduce the gap in diagnostic results between expert and novice physicians. As the system evaluates characteristics of the ultrasonic echo signal that are not shown in conventional B-mode images (spectral content) or that cannot be visually interpreted by the person operating the ultrasound system (complex textures), the system may also be of great help to the expert giving additional data in small and polymorphic cases.

ACKNOWLEDGMENT

This work is an activity of the Ruhr Center of Excellence for Medical Engineering, Bochum, Germany, supported by the German Federal Ministry of Education and Research (grant 13N8079).

Ulrich Scheipers is now with Resonant Medical Inc., Montreal, Quebec, Canada.

REFERENCES

[1] Izzo L, Sassayannis PG et al. The role of Echo Colour/Power Doppler and magnetic resonance in expansive parotid lesions. J Exp Clin Cancer Res. 2004;23(4):585-92.

[2] Scheipers U, Ermert H et al. Ultrasonic Tissue Characterization for Prostate Diagnostics: Spectral Parameters vs. Texture Parameters, Biomed. Technik 2003;48:122-129.

[3] Scheipers U, Ermert H et al. Ultrasonic Multifeature Tissue Characterization for Prostate Diagnostics, Ultrasound Med Biol 2003;29:1137-1149.

[4] Scheipers U, Koenig K et al. Diagnostics of Prostate Cancer based on Ultrasonic Multifeature Tissue Characterization. IEEE Ultrason Sympos 2004;2153-2156.

[5] Zadeh LA. Knowledge Representation in Fuzzy Logic. IEEE Transactions on Knowledge and Data Engineering 1989;1:89-100.

[6] Jang JSR. ANFIS: Adaptive-Network-based Fuzzy Inference Systems. IEEE Transactions on Systems Man and Cybernetics 1993;23:665-685.

[7] Hartman PC, Oosterveld BJ, et al. Detection and differentiation of diffuse liver disease by quantitative echography. A retrospective assessment. Invest Radiol 1993: 28:1-6.

[8] Thijssen JM. Spectroscopy and Image Texture Analysis. Ultrasound Med Biol 2000;26:41-44.

[9] Schmitz G, Ermert H, Senge T. Tissue Characterization and Imaging of the Prostate using Radio Frequency Ultrasonic Signals. IEEE Trans Ultrason Ferroelec Freq Control 1999;46:126-138.

[10] Siebers S, Welp C, Werner J, Ermert H, Ultrasound-based imaging modalities for thermal therapy monitoring. Biomed Tech 2002;47:438-440.

[11] Huisman HJ, Thijssen JM. Precision and Accuracy of Acousto-spectrographic Parameters. Ultrasound Med Biol 1996;22:855-871.

[12] Lizzi FL, Astor M, Feleppa EJ, Shao M, Kalisz A. Statistical Framework for Ultrasonic Spectral Parameter Imaging. Ultrasound Med Biol 1997;23:1371-1382.

[13] Feleppa EJ, Porter CR, Ketterling J et al. Recent developments in tissue-type imaging (TTI) for planning and monitoring treatment of prostate cancer. Ultrason Imaging 2004;26:163-172.

[14] Valckx FMJ, Thijssen JM, van Geemen AJ, Rotteveel JJ, Mullaart R. Calibrated Parametric Medical Ultrasound Imaging. Ultrason Imaging 2000;22:57-75.

[15] Foley DH. Considerations of Sample and Feature Size. IEEE Transactions on Information Theory 1972;18:618-626.

[16] Sugeno M, Yasukawa T. A fuzzy-logic-based approach to qualitative modeling. IEEE Transactions on Fuzzy Systems 1993;1:7-31.

[17] 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 2005 (in press).