Differential diagnosis of parotid gland lesions using spatially fused sonohistologic features

Stefan Siebers, Ulrich Scheipers, Doris Grosse, Frank Gottwald, Alessandro Bozzato, Johannes Zenk, Heinrich Iro, Helmut Ermert

Angaben zur Veröffentlichung / Publication details:

Siebers, Stefan, Ulrich Scheipers, Doris Grosse, Frank Gottwald, Alessandro Bozzato, Johannes Zenk, Heinrich Iro, and Helmut Ermert. 2007. "Differential diagnosis of parotid gland lesions using spatially fused sonohistologic features." *IEEE Ultrasonics Symposium Proceedings, 28-31 Oct. 2007 ,New York, NY, USA*, 456–59. https://doi.org/10.1109/ultsym.2007.123.

Differential Diagnosis of Parotid Gland Lesions using Spatially Fused Sonohistologic Features

Stefan Siebers[1,3], Ulrich Scheipers[1,3], Doris Grosse[1], Frank Gottwald[2], Alessandro Bozzato[2], Johannes Zenk[2], Heinrich Iro[2] and Helmut Ermert[1,3]

[1]Institute of High Frequency Engineering, Ruhr-University, Bochum, Germany
[2]University Hospital for Otorhinolaryngology, Friedrich-Alexander University Erlangen-Nuremberg, Germany
[3]Ruhr Center of Excellence for Medical Engineering (KMR), Bochum, Germany

stefan.siebers@rub.de

Abstract – **In an ongoing clinical study a sonohistology system is developed and evaluated towards its ability to perform computerized differential diagnosis of parotid gland lesions. First order statistics are used to calculate fused features from spatially resolved parameter images. Thereby, characteristics of patterns representing the type of lesion are quantified. Complex baseband ultrasound data have been acquired during the common examinations of patients who were scheduled to have parotid surgery shortly after the acquisition. Data of benign and malignant parotid-gland alterations originating from 135 patients have been included in the study. For data acquisition, a conventional diagnostic ultrasound scanner controlled by custom software running on a laptop computer was used. Lesions were manually contoured in the B-mode images. Acquired data were stored on an external PC. Fused features were calculated offline. From a large number of fused features, a best performing subset is chosen by a selection algorithm to form a feature vector representing each case. The best feature set was used to classify each case using leave-one-out cross validation. Two different classifiers have been used for comparative reasons: a probabilistic neural network based on radial basis functions, and a maximum likelihood classifier, yielding areas under the ROC-curve of 0.85 and 0.91 with standard errors of 0.04 and 0.03, respectively. The system can be adjusted to reach a sensitivity of 1 to catch all positive cases, leaving a remaining maximal specificity of 0.55. Therefore, the system can be used to optimize treatments of parotid gland lesions and to reduce the number of unnecessary surgical interventions.**

Keywords - tissue characterization; classification; texture features; spectral features; parotid gland; cancer diagnostics; ultrasound; sonography; computer aided diagnosis; decision support systems

I. INTRODUCTION

Sonohistology is an ultrasound based diagnostic method for the computerized differentiation of different kinds of biological tissue. The method takes advantage of the ability of certain textural, spectral and morphological features extracted from ultrasound radiofrequency data as well as demodulated data to quantify different tissue conditions or different types of lesions.

Sonohistology has been used successfully for different tasks in the past. As an imaging modality it has been applied for the visualization of prostate tumors [1], for monitoring of thermal therapies [2] and for the visualization of coronary plaques in intravascular ultrasound (IVUS) imaging [3]. As a tool for differential diagnostics it has been used for staging venous thrombosis [4], and, recently, for the classification of parotid gland lesions [5,6].

In a first approach to use sonohistology for differential diagnostics of parotid gland lesions, classification was performed for small regions of interest (ROI) into which each data frame was subdivided. A classification score for the whole lesion was obtained by averaging the classification results from the ROIs of each case. The score was then used to assign each case to either target group positive (malign) or negative (benign) [5,6]. This approach led to promising results, exceeding the hit rates of experienced physicians using B-mode ultrasound as well as palpation and information taken from the anamnesis. However, this methodology does not necessarily account best for characteristic differences in the global pattern of different kinds of lesions. Since a single score should be found for the entire lesion, it might be of advantage to find features characterizing the specific patterns of lesions. Therefore, in this paper an approach is introduced to generate fused features from spatially resolved parameter images. Instead of an ROI-based classification, a case-dependent classification is performed using two kinds of classifiers for comparative reasons. The classification system can be adjusted to reach a sensitivity of 100%, meaning all positive cases are identified correctly. The remaining specificity in that case is up to 0.55. Therefore, this method could be used to add valuable information when a decision about a proper treatment is required. If a lesion is classified as negative, a surgical treatment could possibly be avoided or needs not to be as extensive as for positive cases. This is of importance, since parotid surgery is a challenging intervention due to the intraglandular course of the facial nerve. A damage of the

facial nerve can decrease facial motion or lead to facial palsy and should be avoided for this reason.

II. Materials and Methods

A. Clinical Study and Data Acquisition

For the clinical study, a Siemens Elegra digital ultrasound scanner was used. The linear probe (7.5L40) was set to a center frequency of 7.2 MHz. All data were captured during the routine examination of patients 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 to the head and neck region.

The internal operating system of the Elegra could be accessed via telnet to download baseband ultrasound echo data. These data could be used to reconstruct the original radiofrequency (RF) data by modulating with a dynamic carrier of known characteristics. This is of importance, since RF data are required for the calculation of spectral features.

Controlling of the Elegra was done using custom made software (SynchroSuite, KMR, Bochum, Germany) running on a laptop computer. All relevant imaging settings were taken care of by the software to guarantee a standardized data acquisition.

The download of data was done via FTP. A dump file containing all relevant parameters of the Elegra was also stored along with each frame and used later to compensate for TGC settings. Two orthogonal frames per lesion were recorded.

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. The single transmit focus was set to a depth of 2 cm. Sampling rate and resolution of reconstructed radio frequency echo data was 36 MHz and 12 bits, respectively.

The tumors were manually contoured in the B-mode images by an experienced physician to define an area of interest. Thus, it was ensured that only signals originating from the lesions and not from surrounding tissue have been used for feature extraction and classification.

During the routine examination, the diagnosis of an experienced physician on the basis of ultrasound B-mode imaging, occasional Doppler imaging, palpation and anamnesis was recorded. Histopathological examinations after parotidectomy were used as the "gold standard". The results of histological examinations and subdivision into different classes (see section C) are shown in table 1.

B. Feature Extraction

For the calculation of fused features, parameter images have been generated as described in [6]. The spatially resolved parameter images were obtained by subdividing each data frame into numerous regions of interest (ROIs). Each ROI spanned an area of about 4.6 mm^2 extending approximately 2.7 mm in the axial direction and 1.7 mm in the lateral direction. 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 50 % each. For each ROI, a set of features was calculated using first and second order statistics [7,8,9] as well as spectral estimates (slope, intercept, midband value, attenuation coefficient) [10,11]. Thus, parameter images have been compiled representing the spatial arrangements of features.

From the ROIs that were enclosed by the manually contoured boundary of the lesion, fused features were calculated using statistical measures of first order. These measures were mean, standard deviation, signal to noise ratio, variance, kurtosis and skewness. A subset of these features was used for the subsequent classification. In addition, Fourier descriptors [12] were applied to account for differences in the shape of lesions. The descriptors were calculated from the manually delineated contour lines. An inner boundary tracing algorithm was used to yield a uniformly sampled representation of the contour in 4-neighborhood. Fourier descriptors were obtained after discrete Fourier transforming the manually drawn contour line.

The choice of a combination of features quantifying the characteristic patterns of lesions as well as their shape is motivated by the physician's common approach to type lesions, where texture and shape as appearing in ultrasound B-mode images are also criteria for diagnosis.

C. Feature Selection and Classification

According to the histological findings, datasets were subdivided into five groups or subclasses for classification. Subclasses Ω_{Spos1} and Ω_{Spos2} contained all positive cases, while subclasses Ω_{Sneg1}, Ω_{Sneg2} and Ω_{Sneg3} contained all negative cases. This division into subclasses was chosen to yield intermediate classification results, which are then used to assign a case to either target group Ω_{pos} or Ω_{neg}. The cases having occurred during this study and the subdivision into classes is outlined in table 1.

Two different classifiers were used to process the data, a maximum likelihood classifier, and a probabilistic neural network based on radial basis functions.

The maximum likelihood (MLH) classifier yields likelihoods d_i for each test vector $\mathbf{c}$ to belong to either subclass 1 to 5:

$$d_i(\mathbf{c}) = -\frac{1}{2}\cdot\ln\left(\det(\mathbf{C}_i)\right) - \frac{1}{2}\cdot(\mathbf{c}-\overline{\mathbf{c}}_i)^T\cdot\mathbf{C}_i^{-1}\cdot(\mathbf{c}-\overline{\mathbf{c}}_i),$$

where $\mathbf{c}_i$ are vectors of training data of subclass i, $\mathbf{C}_i$ are covariance matrices of training data of subclass i, and i =1..5. The test vector is assigned to the class Ω_j with the highest

likelihood $d_j = \max\{d_1, d_2, d_3, d_4, d_5\}$. The likelihoods can be scaled to reach a higher sensitivity at the cost of specificity and vice versa. The ROC-curve in Fig. 1 is compiled that way.

The probabilistic neural network consists of two layers. In the first layer, distances of the test vector to the training vectors of the five subclasses are calculated and weighted by a radial basis function to yield probabilities for the test vectors similarity to any training vector. The second layer sums these probabilities for each class. The test vector is assigned to the class with the highest probability. Again, the probabilities can be scaled as with the MLH-classifier.

To reduce the dimensionality of the feature space and to eliminate features which possess high linear dependencies, a correlation analysis was performed.

Each case was classified separately. Classification was performed by leave-one-out validation over cases. Therefore, for classification of one case, the remaining cases were left out of the training dataset. The area under the receiver operating characteristics curve A_{ROC} was used as a quality measure of the classification [13].

To evaluate each feature separately, a first classification pass was done using one feature at a time. A_{ROC} was determined for each case and used as a performance criterion for each feature. The features were then arranged according to their performance. To find the best performing set of features, a second classification pass started with the active set consisting of the best feature exclusively. After each classification pass the active set was updated with the next feature in the queue. If A_{ROC} increased, the new feature remained in the set, otherwise it was removed again.

TABLE I. OCCURRENCE OF DIFFERENT TYPES OF PAROTID GLAND TUMORS AND ALTERATIONS DURING THE CLINICAL STUDY AND SUBDIVISION INTO SUBCLASSES AND TARGET CLASSES

Type of tumor	***n***	subclass	target class
Pleomorphic adenoma	29	Ω_{Spos1}	Ω_{pos}
Acinus cell carcinoma	7	Ω_{Spos2}	
Other Carcinomas	6		
Lymphoma	4		
Metastasis	4		
Monomorphic adenoma	46	Ω_{Sneg1}	Ω_{neg}
Basal cell adenoma	11	Ω_{Sneg2}	
Cyst	13		
Lymph nodes	8	Ω_{Sneg3}	
Nodular fasciitis	1		
Canaliculous adenoma	1		
Adenoid cyst	1		
Lipoma	4		

III. RESULTS

The ongoing clinical study comprises 135 cases origination from 135 patients so far. The diagnosis of experienced physicians on the basis of common diagnostic modalities during the routine examinations yielded a sensitivity (SE) of 0.88 and a specificity (SP) of 0.61. Apparently, common diagnosis is uncertain, resulting in a large quantity of dispensable surgeries.

The feature selection algorithm yielded 7 features as the best set of features for the MLH classifier and 9 features for the PNN. The results of both classifiers are resumed in table 2.

TABLE II. RESULTS OF EXPERIENCED PHYSICIANS (EP) AND CLASSIFICATION USING PROBABALISTIC NEURAL NETWORK (PNN) AND MAXIUM LIKELIHOOD CLASSIFICATION (MLH).

	EP	PNN	MLH
A_{ROC}	-	0.85	0.91
E_{ROC}	-	0.04	0.03
$SE_{\mathrm{SP=0.61}}$	0.88	0.90	0.98
$SP_{\mathrm{SE=0.88}}$	0.61	0.60	0.80
$SP_{\mathrm{SE=1}}$	-	0.38	0.55

The MLH system performed better than the PNN system. The area under the ROC curve was $A_{\mathrm{ROC}} = 0.91$ for the MLH system and $A_{\mathrm{ROC}} = 0.85$ for the PNN system with standard errors $E_{\mathrm{ROC}} = 0.03$ and $E_{\mathrm{ROC}} = 0.04$, respectively [13,14]. However, considering the current available number of cases in this study, this should not imply a general advantage of the MLH classifier against the PNN or other classifiers.

It can be seen from the ROC curve (Fig. 1), that the classification results exceeded the results of the experienced physicians. Compared to the results of the physicians, the MLH classification achieved a specificity of 0.80 at a sensitivity of 0.88. At a specificity of 0.61, the MLH classification achieved a sensitivity of 0.98

In particular, the system can be adjusted to reach a specific sensitivity due to its quantitative output. Since the system should be used to facilitate or support a therapeutic decision, it should be adjusted to type all positive cases correctly, and therefore reach a sensitivity of 1. At a sensitivity of 1, the remaining specificity is 0.55, i.e. 55 % of the negative cases are classified correctly and could therefore be treated alternatively.

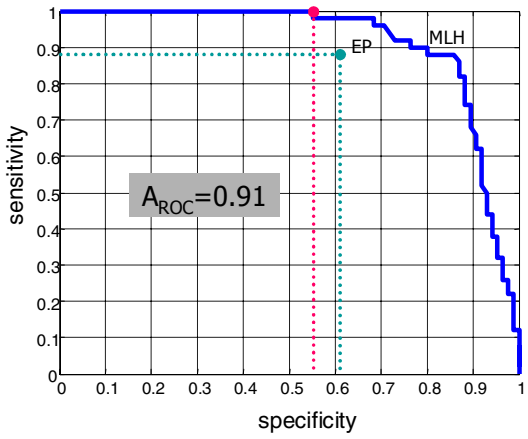

Figure 1. ROC curve of classification result obtained by MLH classification. The area under the ROC curve is $A_{ROC} = 0.91$ with a standard error $E_{ROC} = 0.03$. At a sensitivity of 0.88, specificity is 0.80. At a specificity of 0.61, sensitivity is 0.98. At a sensitivity of 1, the remaining specificity is 0.55.

IV. Discussion and Conclusions

In this paper, a sonohistology system for differential diagnostics of parotid gland lesions is presented. Overall 135 cases originating from 135 patients examined during an ongoing clinical study were considered so far.

Although the number of cases is yet too small to draw general conclusions, the results presented here are promising. The hit rates obtained by the decision support system exceeded the results of experienced physicians using B-mode ultrasound, palpation and information taken from the anamnesis. In addition, the classification system can be adjusted to reach a sensitivity of 1 to detect all positive (malign) cases correctly. The remaining specificity for a sensitivity of 1 is 0.55. This could be of advantage when it comes to making a reasonable decision whether a lesion should be surgically treated or not and to what extent a benign lesion should be treated, respectively. This is of high importance, since parotid gland surgeries can be hazardous due to the intraglandular course of the facial nerve. Therefore, an improved, quantitative diagnostic modality as presented here can be helpful to reduce surgical risks and costs.

The classification rates will possibly increase if a larger database is available. Several alterations of parotid glands occurred only once during this study, others only a few times. However, for an appropriate training of the system, a larger number of each incidence is desirable. If so, evaluation of more subclasses or rather different compositions of subclasses may be beneficial.

Furthermore, the application of second order statistics is well adapted for quantification of characteristics in textural patterns, and therefore might lead to yet another improvement when used to estimate fused features.

Acknowledgements

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). There was additional support by the Ruhr Center for Medical Imaging (zmb ruhr), Bochum with funding from the European Union and the State Government of North-Rhine-Westphalia, Germany (2007).

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

References

[1] U. Scheipers, H. Ermert, H.-J. Sommerfeld, M. Garcia-Schürmann, T. Senge, S. Philippou, "Ultrasonic multifeature tissue characterization for prostate diagnostics", *Ultrasound in Medicine and Biology*, Vol. 29, No. 8, pp. 1137-1149, 2003

[2] S. Siebers, U. Scheipers, M. Ashfaq, J. Hänsler, M. Frieser, D. Strobel, E. Hahn, H. Ermert," In Vivo Imaging of Coagulated Tissue", *Proc. IEEE Ultrasonics Symposium*, pp. 1762-1765, 2006

[3] Perrey, C., Scheipers, U., Bojara, W., Holt, S., Lindstead, M., Ermert, H.: Computerized segmentation of blood and luminal borders in intravascular ultrasound. *IEEE 2004 Ultrasonics Symposium Proceedings*, pp. 1122-1125, 2004

[4] Siebers, S., Geier, B., Vogt, M., Scheipers, U., Mumme, A., Ermert, H.: Classification of venous thrombosis combining ultrasound elastography and tissue characterization. *IEEE 2004 Ultrasonics Symposium Proceedings*, pp. 1761-1764, 2004

[5] U. Scheipers, S. Siebers, F. Gottwald, M. Ashfaq, A. Bozzato, J. Zenk, H. Iro, H. Ermert, "Sonohistology for the Computerized Differentiation of Parotid-Gland Tumors", *Ultrasound Med Biol,* Vol. 31, No. 10, pp. 1287-1296, *2005*

[6] Siebers, S., Scheipers, U., Gottwald, F., Bozzato, A., Mienkina, M., Zenk, J., Iro, H., Ermert, H. : Classification of Parotid Gland Tumors using Sonohistology . *2006 IEEE Ultrasonics Symposium Proceedings*, pp. 645-648, 2006

[7] R.M. Haralick, K. Shanmugam, I. Dinstein, "Textural Features For Image Classification", *IEEE Trans. on Systems, Man, And Cybernetics,* Vol. SMC-3, No. 6, 1973

[8] M. Galloway, "Texture analysis using gray level run lengths", *Computer Graphics and Image Processing,* Vol. 4, pp. 172-179, 1975

[9] K. Laws, "Rapid texture identification", *SPIE Image Processing for Missile Guidance,* Vol. 238, pp. 376-380, 1980

[10] E.J. Feleppa, C.R. Porter, J.A.Ketterling, P. Lee, S. Dasgupta, S., S. Urban, A. Kalisz, "Recent Developments in Tissue-type Imaging (TTI) for Planning and Monitoring Treatment of Prostate Cancer", *Ultrason. Imaging*, Vol. 26, pp. 163-172, 2004.

[11] M. Fink, F. Hottier, J.F. Cardoso, "Ultrasonic signal processing for in vivo attenuation measurement – short time fourier analysis", *Ultrasonic Imaging*, Vol. 5, pp. 117-135

[12] G.H. Granlund, "Fourier Preprocessing for Hand Print Character Recognition", *IEEE Trans. on Computers,* C-21, pp. 195-201, 1972

[13] U. Scheipers, C. Perrey, S. Siebers, C. Hansen, H. Ermert, "A Tutorial on the Use of ROC Analysis for Computer-Aided Diagnostic Systems", *Ultrasonic Imaging*, Vol. 27, pp. 181-198, 2005

[14] J. Hanley, B. McNeil, "The Meaning and Use of the Area under a Receiver Opterating Characteristic (ROC) Curve", Radiology 143, pp. 29-36, 1982