Acta Otolaryngol 2003; 123: 730–734

Colour Texture Analysis for Quantitative Laryngoscopy

JUSTUS F. R. ILGNER[1], CHRISTOPH PALM[2], ANDREAS G. SCHÜTZ[1], KLAUS SPITZER[2], MARTIN WESTHOFEN[1] and THOMAS M. LEHMANN[2]

From the [1]Department of Otorhinolaryngology, Plastic Head and Neck Surgery, University Hospital Aachen and [2]Institute of Medical Informatics, Aachen University of Technology (RWTH), Aachen, Germany

Ilgner JFR, Palm C, Schütz AG, Spitzer K, Westhofen M, Lehmann TM. *Colour texture analysis for quantitative laryngoscopy*. Acta Otolaryngol 2003; 123: 730–734.

Objective—Whilst considerable progress has been made in enhancing the quality of indirect laryngoscopy and image processing, the evaluation of clinical findings is still based on the clinician's judgement. The aim of this paper was to examine the feasibility of an objective computer-based method for evaluating laryngeal disease.
Material and Methods—Digitally recorded images obtained by 90°- and 70°-angled indirect rod laryngoscopy using standardized white balance values were made of 16 patients and 19 healthy subjects. The digital images were evaluated manually by the clinician based on a standardized questionnaire, and suspect lesions were marked and classified on the image. Following colour separation, normal vocal cord areas as well as suspect lesions were analyzed automatically using co-occurrence matrices, which compare colour differences between neighbouring pixels over a predefined distance.
Results—Whilst colour histograms did not provide sufficient information for distinguishing between healthy and diseased tissues, consideration of the blue content of neighbouring pixels enabled a correct classification in 81.4% of cases. If all colour channels (red, green and blue) were regarded simultaneously, the best classification correctness obtained was 77.1%.
Conclusions—Although only a very basic classification differentiating between healthy and diseased tissue was attempted, the results showed progress compared to grey-scale histograms, which have been evaluated before. The results document a first step towards an objective, machine-based classification of laryngeal images, which could provide the basis for further development of an expert system for use in indirect laryngoscopy. *Key words: diagnostic laryngoscopy, electronic imaging, endoscopy, neoplastic larynx disease.*

INTRODUCTION

The quality of laryngeal images visualizing the morphologic changes caused by functional and organic larynx disorders has always been one of the key issues involved in producing a reliable diagnosis and an effective concept for conservative or operative therapy.

Since Garcia and Turck first reported a successful laryngoscopy using sunlight and a tilted, hand-held mirror in the 19th century, many advances in laryngoscopy have been achieved. A major step forward was the introduction of rod endoscopes using a fibre optic lighting system by Hopkins in the 1950s. Twenty years later, von Stuckrad contributed the 90° magnifying laryngoscope to laryngeal imaging. Laryngoscopy by means of modern rod endoscopes has become the gold standard for outpatient diagnostic imaging. Recording laryngoscopic data onto videotape has further facilitated the follow-up documentation of the endolarynx. However, one of the major disadvantages of modern laryngology has always been the lack of suitable tools for normalizing imaging and thus standardizing the evaluation process. Consequently, a conflict remains between advances in imaging and the clinician's judgement, which still relies on applying criteria that are largely subjective.

Concerning digital imaging and image processing, great efforts have been made in the last few decades. In the early 1970s, co-occurrence matrices (CMs) were introduced by Haralick et al. (1) for grey-scale textures. They are defined as a histogram, in which the probability of the simultaneous occurrence of two grey-scale values according to a predefined neighbourhood is stored. Recently, these CMs have been adapted for colour imaging (2, 3). In this paper, we present an evaluation of texture analysis based on CMs with respect to clinical data derived from laryngeal images with and without organic disease in order to examine the feasibility of objective computer-based evaluation of laryngeal disease.

MATERIAL AND METHODS

Image acquisition

Between 13 February, 1999 and 4 January, 2001 we included 16 patients and 19 healthy subjects (controls) in the study. Patients had either been referred by local otolaryngologists for the evaluation of symptomatic, organic laryngeal disorders or directly for larynx surgery due to chronic inflammatory or neoplastic (benign or malignant) larynx disease. After history-taking, all subjects were examined using conventional 90°-angled rod laryngoscopy (Storz 8702 D/S; Storz GmbH, Tuttlingen, Germany) and rigid 70°-angled laryngoscopy (Storz 8200 C) in our outpatient department. The experimental set-up is demonstrated in Fig. 1. Attached to the rod laryngoscope was a charge-coupled device (CCD) video camera (TC 804; Lemke GmbH, Gröbenzell, Germany), whose signal was recorded via an analogue YUV connection to a digital-S video recording machine (digital-S compo-

 ISSN 0001-6489

DOI 10.1080/00016480310000412

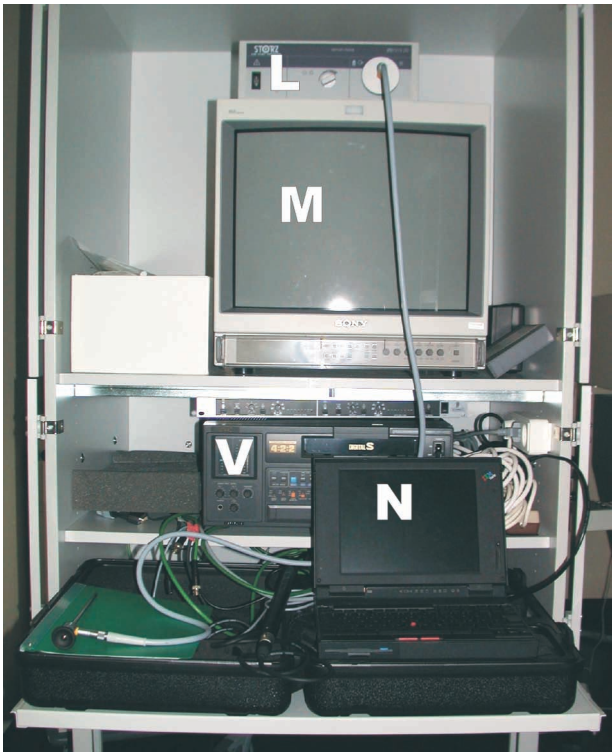

Fig. 1. The videolaryngoscopy set-up, consisting of a xenon light source (L), video monitor (M), digital video cassette recorder (V), notebook computer for storing white balance values (N) and CCD video camera (not pictured).

nent digital 4:2:2; JVC). We employed a 300 W xenon light source (Storz Xenon nova 20131520) in conjunction with a suitable light conductor cable, which remained stationary during the whole examination period. To eliminate ageing artefacts induced by the light source, a standardized white balance excluding all other surrounding light sources in a black box was performed each day before recording. The examiner controlled each recording using a video monitor (Trinitron PVM 2053 MD; Sony) which was switched in parallel to the recorder via another YUV connection.

For each subject, the examiner filled in a clinical questionnaire concerning general parameters such as skin colour, colour of oral mucous membranes, packed cell volume and oral temperature (where applicable). In terms of local parameters, the aspect of the larynx was judged by the clinician with special consideration being paid to the vocal cords, whose aspect was classified using the following criteria: colour (normal, pale or flushed); texture (normal, rough, mucus, singer's nodes, leukoplakia, polyps, papillomas or other organic disorders); movement (normal, reduced or absent); and closure of the vocal cords (normal, posterior gap, anterior gap or combined).

Image processing

Digital image processing and colour texture analysis were programmed in C language and integrated to the Khoros environment (Khoral Research Inc.). For the formal definition of the CMs, we considered the following colour image

$$\mathbf{f}(m,\ n) = (f_R(m,\ n),\ f_G(m,\ n),\ f_B(m,\ n))^T$$

with $0 \le m < M$, $0 \le n < N$ and $m, n, M, N \in \mathrm{IN}$ consisting of the red (R), green (G) and blue (B) colour channels. The corresponding grey-value image $f_g(m, n)$ is derived from the weighted mean of the colour channels at each pixel position $(m, n)^T$. Thus the CMs define a two-dimensional histogram, where the probability, P, of the simultaneous occurrence of two grey or colour values according to a fixed displacement vector $\mathbf{d} = (d_1, d_2)^T$, $d_i \in \mathrm{IR}$ is stored as

$$C_{\mathbf{d}}^{k}(w_1,\ w_2) = P(f_k(n_1,\ m_1) = w_1 \wedge f_k(n_2,\ m_2) = w_2 \\ |m_1 - m_2| = d_1, \quad n_1 - n_2 = d_2)$$

with $k \in \{g, R, G, B\}$.

Each cell of the matrix counts the number of co-occurrences of the two values w_1 and w_2, $w_i \in \mathrm{IN}$, which are defined by their positions on the abscissa and ordinate, respectively. The dimension of the matrix $C_{\mathbf{d}}^{k}(w_1, w_2)$ corresponds to the number of possible grey or colour values in the image. Rotational invariance is obtained by combining the CMs for the eight equidistant angles. In general, homogeneous surface patches are represented near the diagonal, because the same values are neighboured frequently. Inhomogeneity, on the other hand, is represented by a widespread scatterplot (Fig. 2). According to Haralick et al. (1), the matrices themselves are characterized by special features, which refer to the shape and internal structure of the scatterplot. We chose a selection of eight features that have been proven to best describe the entire matrices (4–6).

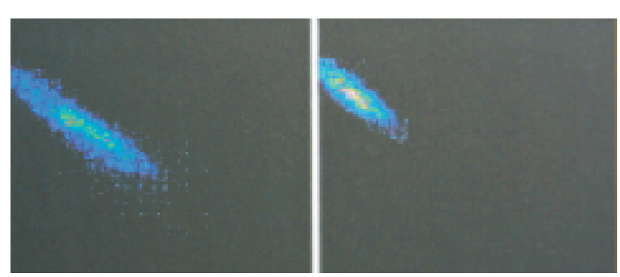

Fig. 2. CMs (C_{RGB}, $d = 1$) for a normal, homogenously textured vocal fold (*right*) and for chronic inflammatory laryngitis at the vocal-fold level (*left*). Differences in the position, shape and structure of the scatterplot were assessed by means of mathematical analysis according to Haralick et al. (1).

The CM features were evaluated by comparing the automatic classification with the clinical findings of two otorhinolaryngologists. The set of 35 images includes 19 healthy individuals and 16 patients with organic diseases. CMs were computed for the area of the vocal cords, which were segmented manually (Fig. 3). The classification was done using the leaving-one-out method for separation of the test and training parts, i.e. each feature vector serves once as a test vector while it is excluded from the training set. Therefore, each training set consists of 34 feature vectors and the classification experiment is repeated for each of the 35 vectors. The five-nearest-neighbour classifier was applied by assigning a test vector to the class that represented the majority of the five nearest vectors in the environment (7).

In particular, classification was computed according to the colour CM features of each of the red, green and blue channel, as well as the grey-value CM and the colour histogram. The last two techniques provide pure intensity texture without colour information and pure colour information without texture dependency, respectively. For all computations, different values of the radius, d, were applied.

RESULTS

The results are shown in Table I. With just 40% correctly classified images, the colour histogram performed rather badly. On the other hand, the best classification results of 74.3% for the grey-value texture analysis and 81.4% for the colour texture analysis were significantly better. Obviously, the intensity patterns were quite different for the different colour channels. For instance, the favoured resolution of the red channel was $d = 1$ whereas the other colour channels showed best results with $d = 8$ or 9 pixels.

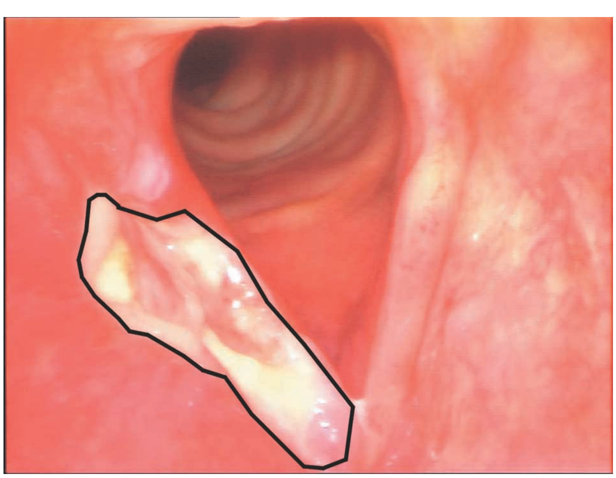

Fig. 3. Clinical picture of a suspect lesion covering the anterior two-thirds of the right vocal fold. Labelling of the lesion by outlining its borders was performed manually.

Overall, the best result of 81.4% was achieved using CM according to the blue channel and $d = 8$. Here, the combination of all the colour channel features performed worse, because in general the increase in the feature space dimension resulted in a decrease in the classification rate. For larger displacements, the increase in information obtained by combining different features is compensated for by this phenomenon, which is called the curse of dimensionality (7). Nonetheless, texture analysis based on the combined CM, C_{RGB}, is superior to both grey and single-colour CMs, within a wide range of displacements $1 < d < 8$. For $d = 5$, which is exactly in the middle of this range, the best classification of 77.1% was obtained.

DISCUSSION

Over the past 20 years, the development of laryngoscopic imaging has branched in many ways and this has led to a substantial increase in the diagnostic information available in order to classify functional and organic disorders with greater reliability. Whilst the mechanisms of functional disorders can be examined further using videostroboscopy, high-speed digital video recordings have facilitated a more subtle tracking of laryngeal movements (8), which can be further visualized by means of videokymography (9–11). With all these techniques, digital image post-processing [e.g. contour enhancement, light intensification (12, 13) or Fourier transformation of light-intensity video images (14)] has led to an increase in diagnostic information. Computer-based evaluation has been utilized by Hess and Ludwigs (15) for contour detection or quantitative measurement of surfaces and distances at the glottic and subglottic levels. Zrunek et al. (16) used digital image processing to measure the surface area of the open glottic space during bilateral electrical stimulation of the posterior cricoarytenoid muscle and at rest. Furthermore, investigations of organic disorders have profited from intra-operative measures, such as contact laryngoscopy or autofluorescence endoscopy (17–19), which have been evaluated further by Arens et al. (20). Intra-operative visualization in turn has been enhanced by virtual endoscopy models created on the basis of spiral CT data (21, 22). Machine-based evaluation aids in laryngoscopy have been applied mostly in a diagnostic setting, where access to laryngeal regions is limited due to the patient's compliance, in order to provide additional quantitative data to support a clinical diagnosis. A standardized, computer-based diagnostic model for the qualitative evaluation of morphologic changes has not yet been introduced.

In this context, laryngeal texture analysis is still dominated by the examiner's judgement, in terms of

Table I. *Classification results based on colour histogram and CMs. C_{RGB} denotes the combined features for the red, green and blue channel by cross-product. Therefore, the feature vector according to C_{RGB} is three times greater than the others*

	Displacement, d										
Features based on	0	1	2	3	4	5	6	7	8	9	10
Colour histogram	0.400										
C_g		0.543	0.543	0.600	0.571	0.686	0.686	0.714	0.686	**0.743**	0.743
C_R		0.714	0.486	0.629	0.629	0.543	0.543	0.600	0.486	0.457	0.457
C_G		0.514	0.514	0.571	0.571	0.571	0.629	0.686	0.714	0.743	0.714
C_B		0.571	0.571	0.600	0.571	0.657	0.686	0.686	**0.814**	0.743	0.771
C_{RGB}		0.629	0.629	0.742	0.686	**0.771**	0.743	0.743	0.771	0.743	0.743
Result		−	+	+	+	+	+	+	−	+/−	−

his/her personal knowledge and skill. Until now, the volume of data acquired and the multitude of variables involved when evaluating laryngeal images have prevented digital imaging from progressing significantly in this field. Whilst improvements in the speed of data processing and storage capacity are about to overcome the methodical obstacles, suitable algorithms, focusing especially on textural and structural aberrations, are yet to be developed. In addition, current concepts of digital evaluation are highly susceptible to adverse collateral factors, such as changes in lighting, visual angle, contrast, patient compliance, etc. This study only focused on a very basic classification, differentiating clinical findings into physiological and pathological states. However, the results obtained by the CM analysis, specifically in the blue channel, represent progress compared to the grey-scale texture analysis of Haralick et al. (1). Furthermore, our results demonstrate the potential of colour-based CMs in relation to laryngeal imaging. With further advances in data processing and storage, the technical obstacles can be overcome. The subject of future research work is the creation of suitable algorithms based on CMs in relation to laryngeal texture and colour that will allow further subclassification of pathological findings. Computer-based evaluation of laryngeal images could play a greater role in the near future as it provides a useful objective aid to the physician's judgement; in addition it could be used by senior clinicians to teach junior colleagues.

ACKNOWLEDGEMENTS

This study was performed as part of research work funded by the "Deutsche Forschungsgemeinschaft" (DFG), project Nos. We 2147/1-1 and SP 538/2-1 and 2-2.

REFERENCES

1. Haralick R, Shanmugam K, Dinstein I. Texture features for image classification. IEEE Trans Syst Man Cybern 1973; 3: 610–21.
2. Palm C, Lehmann TM, Spitzer K. Colour texture analysis of moving vocal cords using approaches from statistics and signal theory. In: Braunschweig T, Hanson J, Schelhorn-Neise P, Witte H, eds. Advances in quantitative laryngoscopy, voice and speech research, Proc. 4th Int. Workshop, Friedrich-Schiller University, Jena, Germany, 2000: 49–56.
3. Hassan H, Ilgner J, Palm C, Lehmann TM, Spitzer K, Westhofen M. Objective judgement of endoscopic laryngeal images. In: Lehmann TM, Palm C, Spitzer K, Tolxdorff T, eds. Advances in quantitative laryngoscopy, voice and speech research, Proc. 3rd Int. Workshop, RWTH Aachen, Aachen, Germany, 1998: 135–42.
4. Gotlieb C, Kreyszing H. Texture descriptors based on co-occurrence matrices. Computer Vision, Graphics and Image Processing 1990; 51: 70–86.
5. Pelikan E. Texturorientierte Segmentierungsmethoden in der medizinischen Bildverarbeitung [PhD Thesis]. Aachen, Germany: Aachen University of Technology, 1995.
6. Bradleyn A, Jackaway P, Lovell B. Classification in scale space—applications to texture analysis. Proc. 14th Conf on Information Processing in Medical Imaging, Ile de Berder, France, 1995: 375–6.
7. Fukunaga K. Introduction to statistical pattern recognition. San Diego, CA: Academic Press, 1973.
8. Miyaji M. Analysis of pathological vocal fold vibrations using high-speed digital imaging. Nippon Jibiinkoka Gakkai Kaiho 1998; 101: 1099–111.
9. Tigges M, Wittenberg T, Mergell P, Eysholdt U. Imaging of vocal fold vibration by digital multi-plane kymography. Comput Med Imaging Graph 1999; 23: 323–30.
10. Svec JG, Schutte HK. Videokymography: high-speed line scanning of vocal fold vibration. J Voice 1996; 10: 201–5.
11. Svec JG, Horacek J, Sram F, Vesely J. Resonance properties of the vocal folds: in vivo laryngoscopic

investigation of the externally excited laryngeal vibrations. J Acoust Soc Am 1999; 108: 1397–407.
12. Maurer D, Hess M, Gross M. High-speed imaging of vocal fold vibrations and larynx movements within vocalizations of different vowels. Ann Otol Rhinol Laryngol 1996; 105: 975–81.
13. Hess MM, Gross M. High-speed, light-intensified digital imaging of vocal fold vibrations in high optical resolution via indirect microlaryngoscopy. Ann Otol Rhinol Laryngol 1993; 102: 502–7.
14. Granqvist S, Lindestad PA. A method of applying Fourier analysis to high-speed laryngoscopy. J Acoust Soc Am 2001; 110: 3193–7.
15. Hess MM, Ludwigs M. Strobophotoglottographic transillumination as a method for the analysis of vocal fold vibration patterns. J Voice 2000; 14: 255–71.
16. Zrunek M, Bigenzahn W, Mayr W, Unger E, Feldner-Busztin H. A laryngeal pacemaker for inspiration-controlled, direct electrical stimulation of the denervated posterior cricoarytenoid muscle in sheep. Eur Arch Otorhinolaryngol 1991; 248: 445–8.
17. Pigott RW. An analysis of the strengths and weaknesses of endoscopic and radiological investigations of velopharyngeal incompetence based on a 20 year experience of simultaneous recording. Br J Plast Surg 2002; 55: 32–4.
18. Zargi M, Smid L, Fajdiga I, Bubnic B, Lenarcic J, Oblak P. Laser induced fluorescence in diagnostics of laryngeal cancer. Acta Otolaryngol Suppl (Stockh) 1997; 527: 125–7.
19. Zargi M, Smid L, Fajdiga I, Bubnic B, Lenarcic J, Oblak P. Detection and localization of early laryngeal cancer with laser-induced fluorescence: preliminary report. Eur Arch Otorhinolaryngol 1997; 254(Suppl 1): S113–6.
20. Arens C, Malzahn K, Dias O, Andrea M, Glanz H. Endoscopic imaging techniques in the diagnosis of laryngeal carcinoma and its precursor lesions. Laryngorhinootologie 1999; 78: 685–91.
21. Triglia JM, Nazarian B, Sudre-Levillain I, Marciano S, Moulin G, Giovanni A. Virtual laryngotracheal endoscopy based on geometric surface modeling using spiral computed tomography data. Ann Otol Rhinol Laryngol 2002; 111: 36–43.
22. Gallivan RP, Nguyen TH, Armstrong WB. Head and neck computed tomography virtual endoscopy: evaluation of a new imaging technique. Laryngoscope 1999; 109: 1570–9.

Submitted September 24, 2002; accepted December 18, 2002

Address for correspondence:
Justus Ilgner
Department of Otorhinolaryngology,
Plastic Head and Neck Surgery
University Hospital Aachen
Pauwelsstrasse 30
DE-52057 Aachen
Germany
Tel.: +49 241 80 88967
Fax: +49 241 80 82523
E-mail: jilgner@ukaachen.de