Biomed Tech 2012; 57 (Suppl. 1) © 2012 by Walter de Gruyter · Berlin · Boston. DOI 10.1515/bmt-2012-4471

Laryngoscopic Image Stitching for View Enhancement and Documentation – First Experiences

Maria Schuster[2], Tobias Bergen[1], Maximilian Reiter[2], Christian Münzenmayer[1], Sven Friedl[1], Thomas Wittenberg[1]

[1] Fraunhofer IIS, Department for Image Processing and Biomedical Engineering, Erlangen, Germany
Email: thomas.wittenberg@iis.fraunhofer.de

[2] Department of Otorhinolaryngology and Head and Neck Surgery, University of Munich, Germany

Abstract

One known problem within laryngoscopy is the spatially limited view onto the hypopharynx and the larynx through the endoscope. To examine the complete larynx and hypopharynx, the laryngoscope can be rotated about its main axis, and hence the physician obtains a complete view. If such examinations are captured using endoscopic video, the examination can be reviewed in detail at a later time. Nevertheless, in order to document the examination with a single representative image, a panorama image can be computed for archiving and enhanced documentation. Twenty patients with various clinical findings were examined with a 70° rigid laryngoscope, and the video sequences were digitally stored. The image sequence for each patient was then post-processed using an image stitching tool based on SIFT features, the RANSAC approach and blending. As a result, endoscopic panorama images of the larynx and pharynx were obtained for each video sequence. The proposed approach of image stitching for laryngoscopic video sequences offers a new tool for enhanced visual examination and documentation of morphologic characteristics of the larynx and the hypopharynx.

1 Introduction

One familiar problem within diagnostic laryngoscopy is the spatially limited view onto the hypopharynx and the larynx through a laryngoscope. Usually, during endoscopy, the physician focuses on the interesting part, e.g. the vocal cords when patients report about voice problems, and zooms in for better visual evaluation. To assess and examine the complete larynx and hypopharynx with the same high spatial resolution, the laryngoscope is moved from one to the other side, and from front to back, and is rotated about its main axis. Hence, the physician obtains a complete view on structures and function. When focusing on special structures, however, the physician might miss other pathologic structures as visual perception is concentrated on the focused structures.

Depending on the speed of the examination, such an "overview" scan relates to a set of consecutive images in the range of dozens to several hundred single image frames, each one depicting a certain area of the larynx and hypopharynx. Even though the laryngoscopic examination can be captured using analogue or digital endoscopic video devices and reviewed at a later point of time, the recorded examination contains a dynamic nature, and the related impression obtained by the physician through the ocular of the laryngoscope or on a video monitor is quite volatile. Nevertheless, in order to capture the laryngoscopic examination with one single, representative image which can furthermore be used for a digital documentation, a panorama endoscopy image can be computed for archiving and enhanced documentation.

The fundamental idea to generate a larger image with a wider view – a so-called panorama image – from a set of temporally and spatially adjacent image frames is certainly well established in computational photography [1,2], and has already been applied within various fields of image based medicine in the past years. Examples of panoramic images within medical applications are e.g. the computation of so-called "virtual slides" [3,4], where a set of several hundreds of adjacent single microscopic views from a scan are computationally fused to yield a large scale (Gigabyte) panoramic micrograph, or "concatenated radiographs" [5] where several sequentially obtained X-ray images are combined to form a larger image. The process of panorama image generation is also widely denoted as "mosaicking".

Within the emerging field of "computer-integrated endoscopy" [6], several approaches and applications have been suggested for the generation of endoscopic panorama images (EPIs). As the computation of EPIs is quite complex and computationally expensive, most of the presented approaches from literature suggest "off-line" solutions, meaning that the calculation is done at some time after the examination. Nevertheless, recently some work has been presented trying to provide real-time EPIs.

For example, *Becker et al.* [7] have evaluated the mosaicking of confocal laser scanning images acquired in the esophagus. *Behrens* [8] has presented an approach for the mosaicking of fluorescence bladder endoscopy based on SIFT features. Both approaches can be considered as "offline" approaches for the generations of EPIs. For the "online" or "realtime" variation also some approaches have been suggested. *Seshamani et al.* [9] have provided a real time endoscopic stitching approach for the assessment of microscopic retinal and catadioptric endometrial images. *Konen et al.* [10] suggested an approach for real-time mosaicking for neuro-endoscopic imagery. [11,12] describe a real time endoscopic stitching approach, based on KLT feature tracking, which has been applied to liver EPIs. Finally, an important contribution to this field is the work by *Mountney and Yang* [13] who suggest a simultaneous location and mapping approach (SLAM) for the real-time generation of EPIs.

Within this work for the first time endoscopic panorama images are generated for the field of laryngoscopy and evaluated on some initial recordings. In contrast to other organs examined, the larynx and hypopharynx are rapidly and constantly moving: The larynx with true and false vocal folds is opening and closing during respiration or phonation, a function that is usually examined during endoscopy. Therefore, next to a moving endoscope, image stitching technique also has to deal with changing morphology of the organs examined.

2 Materials and Methods

2.1 Materials

During the routine consultation at the Department of Otorhinolaryngology, Head and Neck Surgery of the University of Munich, Germany, twenty patients with various clinical findings were examined with a 70° rigid laryngoscope and a digital video recording system (Wolf HRES Endo-Cam 5562). All acquired image sequences were digitally stored in the high spatial resolution mode with 25 frames per second and a spatial resolution of 512 x 384 pixels.

2.2 Methods

The image sequences for all patients were post processed using the freely available image stitching software *Autostitch*[1], which follows the approach presented by Brown and Lowe [14,15]. A main advantage of that approach is its capability to work with an unordered set of images. The pan shot of the larynx is mostly non-linear and due to motion artifacts and difficult lighting conditions, not every single image frame can be processed for a panoramic visualization. The approach itself extracts SIFT features in all input images to determine overlaps. The advantage of SIFT, also for laryngoscopic recordings, is the invariance to rotation and scaling as those occur continuously. Similar feature points are then matched by a k-nearest-neighbor method determining common landmarks in overlapping images. To ensure geometric consistency between the feature point maps, the software applies RANSAC to identify and eliminate outliers. Thus, a set of overlapping images is determined, in which each image has to be transformed regarding the relative geometric projection. To prevent accumulated errors, a Bundle Adjustment is used to determine the final transformation parameters. This panoramic view will be of unsatisfying visual quality in most cases. Considering the different photometric parameters and different lighting conditions, gain compensation adjusts these inhomogeneities. For an optimized rendering, a multi band blending approach is used which considers different frequencies and results in smoother overlaps. Figure 1 depicts a stitched image of the larynx, without and with edge blending.

[1] http://www.cs.bath.ac.uk/brown/autostitch/autostitch.html

2.3 Experiments

All computed endoscopic panorama images were evaluated by two experienced specialists in ENT looking for morphologic abnormalities in the whole image. Specifically, special visual attention was laid onto the regions right and left of the glottis and the vocal folds. Both noted their findings independently.

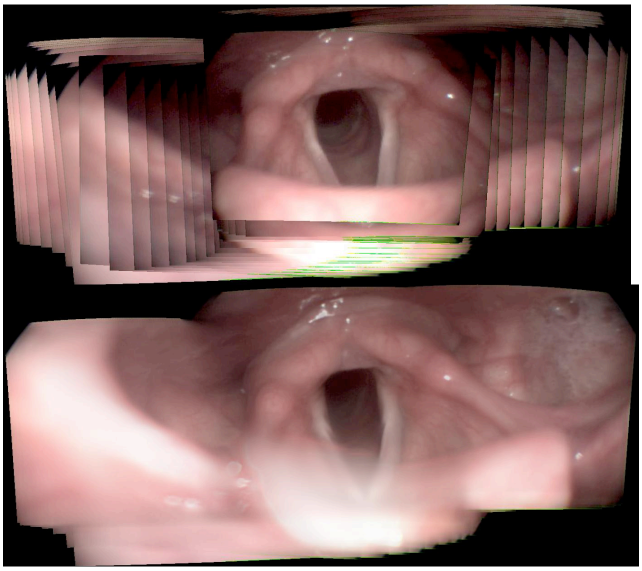

Image 1: Laryngeal panorama image obtained from an image sequence with and without edge blending.

3 Results

For all laryngoscopic image sequences, several panoramic images were computed using varying parameters, as e.g. for the number of single frames from the sequences or the interval between selected frames. This was especially important to avoid motion artifacts, as e.g. from vocal fold vibrations or the opening of the epiglottis. The quality of the obtained panoramic images depends on several factors. Amongst them are the rotation and translational movements of the laryngoscope during the recoding, which is guided by the physician. Also the movements of the patients vocal fold and epiglottis is important, and it was observed that any phonation during the laryngeal examination, which is related to very fast movements of vocal folds, is counter-productive to the mosaicking approach, as the moving vocal folds do not provide steady landmarks to correlate adjacent images with each other. Based on these observations, for twelve of the twenty recordings, meaningful EPIs could be computed.

Next to the findings described and documented directly after the original laryngoscopic examination, both ENT specialists could detect additional morphologic findings on the panorama images. These additional findings include six cysts, four mucosal alterations, one edema, and one saliva pooling. Four of these findings are of clinical relevance; two of them were in accordance with other findings already described before and therapy was adequate. For the other two findings further examinations should have been proposed.

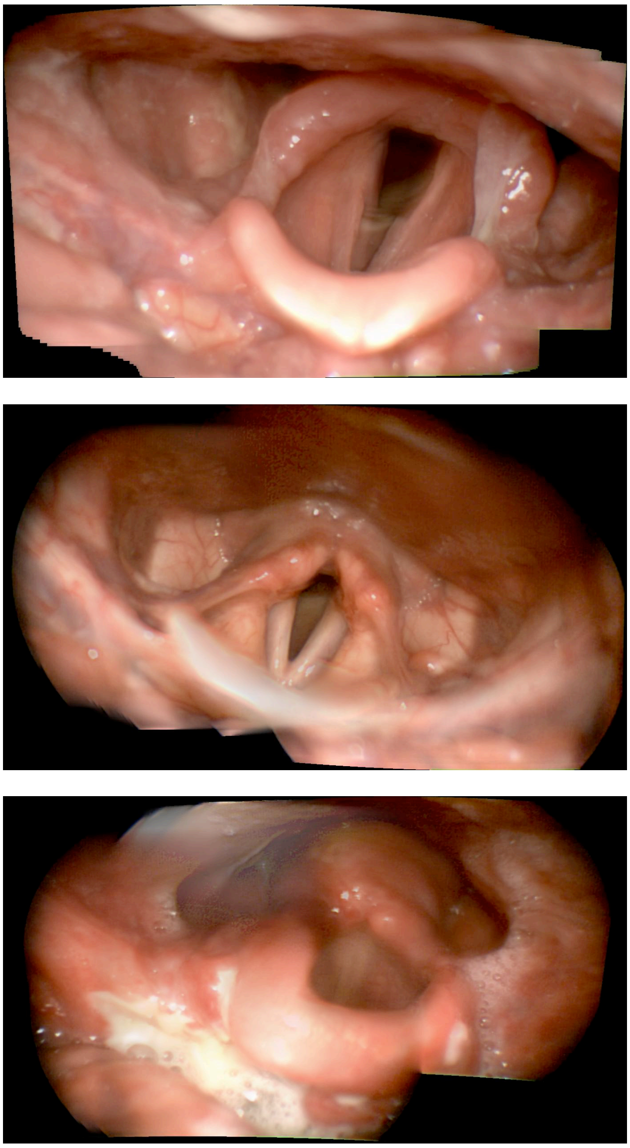

Image 2: Three examples of endoscopic panorama images of the larynx with various findings. Top: acute laryngitis with penetration of saliva into the endolarynx. Center: normal laryngeal morphology in muscle tension dysphonia and small mucosal cyst at the left pharyngoepiglottic plica. Bottom: Inflammation of the Epiglottis and the valleculae with swelling and fibrine plaques.

4 Conclusions

The proposed approach of image stitching for laryngoscopic video sequences offers a new tool for enhanced visual examination and documentation of morphologic characteristics of the larynx and the hypopharynx. In addition to commonly used documentation tools, endoscopic panorama images allow for better overview of laryngeal and hypopharyngeal morphology on a single image. For video sequences with rapidly altering morphologic structures due to movements further post-processing has to be developed. Preliminary results on some video sequences show that next to focused alterations in video sequences, more morphologic alterations can be detected with impact on diagnostic or therapeutic decisions in some cases.

In conclusion, endoscopic panorama imaging may improve diagnostic and therapeutic procedures as more noticeable morphologic alterations can be detected. The impact of the method in everyday clinic situation will be tested on more video sequences.

References

[1] Szeliski R: Image Alignment and Stitching: A Tutorial: Microsoft ResearchTechReportNumber: MSR-TR-2004-92 http://research.microsoft.com/pubs/70092/tr-2004-92.pdf

[2] Chen CY, Klette, R: Image Stitching Comparisons and New Techniques. LNCS, 1999, Vol. 1689/1999, 835

[3] Steckhan D, Paulus D: A quadratic programming approach for the mosaicing of virtual slides that incorporates the positioning accuracy of the microscope stage. Proc's EMBC, pp. 72-77, 2010

[4] Steckhan D, Bergen T, Wittenberg T, Rupp, S: Robust image stitching for virtual microscopy. Proc's 30th EMBC, pp. 4019-4023, 2008

[5] Gooßen A Pralow T, Grigat RR: Automatic Stitching of Digital Radiographies using Image Interpretation: In McKenna S & Hoey J (Eds.). Medical Image Understanding and Analysis 2008, Proc's 12 Ann. Conf. pp. 204-208, University of Dundee, 2008

[6] Wittenberg T, Münzenmayer, C: Computer-integrierte Endoskopie -- Computer Integrated Endoscopy. Endoskopie Heute 24: 271–277,2011.

[7] Becker V, Vercauteren T, von Weyhern C, Prinz C, Schmid RM, Meining A: High-resolution miniprobe-based confocal microscopy in combination with video mosaicing. Gastrointestinal Endoscopy 2007 Nov; 66(5):1001-7

[8] Behrens A: An Image Mosaicing Algorithm for Bladder Fluorescence Endoscopy, Proc's 12th Int. Student Conf. on Electrical Eng., Prague, 2008;

[9] Seshamani S, Lau WW, Hager G: Real-Time Endoscopic Mosaicking. MICCAI (1) 2006: 355-363

[10] Konen W, Naderi M, Scholz M, Endoscopic image mosaics for real-time color video sequences. Proc's Comp. Assisted Radiology & Surgery (CARS), 2007

[11] Bergen T, Ruthotto S, Rupp S, Winter C, Münzenmayer C: Endoscopic Egomotion Computation. Proc's SPIE Medical Imaging 2010: Computer Aided Diagnosis, 13.-18.2.2010, Vol. 7623

[12] Bergen T, Schneider A, Münzenmayer C, Knödgen F, Feussner H, Wittenberg, T Winter C: Echtzeit-Stitching endoskopischer Bilder für eine erweiterte Sicht in chirurgischen Eingriffen, Endoskopie Heute, S. 62, 24.01.2011.

Biomed Tech 2012; 57 (Suppl. 1) © 2012 by Walter de Gruyter · Berlin · Boston. DOI 10.1515/bmt-2012-4471

[13] Mountney P, Yang GZ: Dynamic View Expansion for Minimally Invasive Surgery using Simultaneous Localization And Mapping. 31st Annual Int. Conf. the IEEE EMBS Minneapolis, Minnesota, USA, September 2-6, 2009

[14] Brown M, Lowe DG: Automatic Panoramic Image Stitching using Invariant Features. Int. J. of Comp. Vision, 74(1), pp. 59-73, 2007

[15] Brown M, Lowe DG: Recognising Panoramas. In Proceedings of the 9th Int. Conf. on Comp. Vision (ICCV2003), pp. 1218-1225, Nice, France, 2003