Image stitching of sphenoid sinuses from monocular endoscopic views

T. Bergen[1], P. Hastreiter[2], C. Münzenmayer[1], M. Buchfelder[2], T. Wittenberg[1]

[1] Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany
[2] Department for Neurosurgery, University Clinics Erlangen, Germany

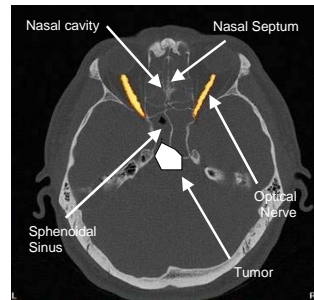Contact: thomas.wittenberg@iis.fraunhofer.de

Abstract:

For operations of the pituitary glands, the most subtle method is an intervention through the paranasal and especially through the sphenoid sinus. To avoid dangerous interference with adjacent organs and nerves, the surgeon has to orient himself in the very small sphenoid cavity and navigate across the hollow space to break through the sellar floor to the pituitary gland above. Especially in reoperations or anatomical variants such as so-called kissing carotids, transsphenoidal surgery is a challenge even in experienced hands. To support such a surgery, various imaging modalities can be applied such as CT, MRI or endoscopy. While pre-operative MRI or CT-data can be used for intervention planning and navigation support, endoscopy can be applied intra-operatively for the examination of surfaces inside the sphenoid sinus. In this work, we present initial experiments and results from real-time panorama-endoscopy of the sphenoid sinus for navigation and orientation support, based on monocular endoscopic sequences of a skull phantom, yielding partial reconstructions of the walls of the sphenoid sinus.
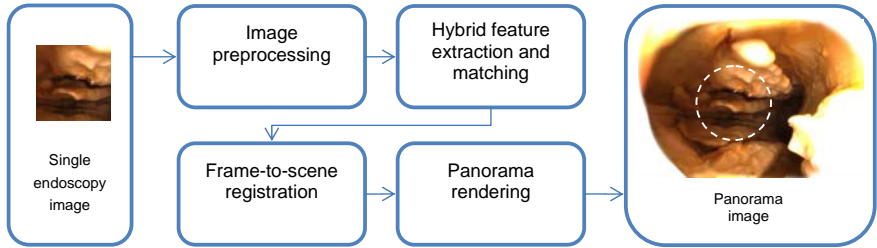
Key words: pituitary surgery, sinus surgery, panorama-endoscopy, stitching, mosaicking, real-time

1   Introduction

The most subtle method for operations of the pituitary glands, such as the removal of tumors or adenomas, is transsphenoidal surgery. This involves the difficulty of maneuvering through the paranasal and especially the sphenoid sinus, a small cavity behind the eyes, to break through the sellar floor and gain access to the pituitary gland. This is a difficult operation due to the risk of damaging adjacent nerves and organs, such as the internal carotid artery. depicts the situation in a CT slice. Various imaging modalities can be applied to support the surgeon, including e.g. CT, MRI or endoscopy. CT and MRI are available in the pre-operative planning phase. The standard imaging modality during the operation is the view through an endoscope. One major aspect of difficulty is the limited field of view provided by the endoscope. To improve orientation and maneuverability for the surgeon, image stitching techniques can be applied to provide an augmented field of view. In this paper, we propose a real-time panorama-imaging approach for navigation and orientation support, based on monocular endoscopic sequences of a skull phantom, yielding partial reconstructions of the walls of the sphenoid sinus. These experiments are based on prior experiences, gained from a 3D recon-

Figure 1: Transnasal approach to the tumor in the pituitary gland, depicted in an axial CT slice of a head.

struction approach from endoscopic views [1]. Further work concerning view enhancement in sinus surgery includes registration techniques for CT/endoscopy registration by Burschka et al. and Mirota et al. [2, 3]. Wise and DelGaudio as well as Palmer and Kennedy provide review articles of computer-assistance in paranasal sinus surgery [4, 5]. Different aspects of navigation and registration of pre- and intra-operative imaging techniques are discussed as a means of facilitating orientation for the surgeon. However, panorama-endoscopy has not yet been considered in the field of sinus surgery.

Figure 2: Workflow of the proposed system. Every video frame is preprocessed. A hybrid feature tracking is applied to align the image with the scene. Finally the panorama is rendered to the screen.

2   Materials and Methods

In this work, we present a system for real-time panorama imaging ("mosaicking") of monocular endoscopic views with application to sphenoid sinuses. The approach is based on a system, which we have published earlier [6] for real-time stitching of the urinary bladder. In this section, the algorithmic components are described. Figure 2 depicts an overview of the proposed approach. In the following sections, we describe all steps in further detail.

Image Preprocessing
The video frames are captured from the camera at a rate of about 30 frames per second. Every video frame is preprocessed to detect the circular mask (aperture) typical for endoscopic recordings. This is achieved by segmenting all non-black pixels from the image and fitting a circular disc to the extracted region. Furthermore, we compensate for lense distortion and inhomogeneous illumination. Endoscopic images usually suffer from a barrel distortion. To reduce this effect, we apply an undistortion filter, computed on the basis of priorly captured images of a checkerboard pattern. We use the undistortion filter provided by the OpenCV software library. Inhomogeneous illumination, i.e. a strong vignetting effect, is caused by the point light source generally used in endoscopy. We compensate for illumination inhomogeneities by applying a high-pass filtering to the input image: a strongly smoothed image is subtracted before passing the frame to the feature tracking module.

Feature tracking and frame registration
This section is based on a recent publication of ours, describing a hybrid tracking approach for real-time stitching during cystoscopy [7]. Here, we briefly describe the essential steps and refer the reader to [7] for details. The tracking module encapsulates both the SURF (Speeded Up Robust Features [8]) and KLT (Kanade-Lucas-Tomasi [9]) tracking algorithms in a multi-threaded implementation. While SURF generates feature descriptors that allow matching features from the current video frame to the global set of all prior features, KLT is more suited for matching features between successive video frames. Consequently, a mosaicking system based on KLT tracking suffers from a drift error, which increases over time. On the other hand, KLT requires less computational time than SURF, making it very suitable for real-time applications. In order to exploit the advantages of both approaches, we combine KLT and SURF tracking to achieve both, a fast processing speed as well as high matching accuracy. Both tracking threads calculate a projective transformation within a RANSAC (RANdom SAmple Consensus) scheme to align the current video frame to the scene, i.e. the panorama coordinate space.

Panorama rendering
Based on the previous image registration step, all processed frames are rendered as a panorama image using OpenGL. The two-dimensional projective transformation is converted to a three-dimensional transform, which maps the respective frame texture to the xy-plane. To reduce visible seams along the edges of a frame, we use a basic alpha blending approach. The alpha channel of each frame is designed as a center weighted function with $\alpha = 1$ in the central image region and linearly decreasing value to the outer image edge. Due to the real-time processing ability, the system is able to dynamically extent the endoscopic view field during the procedure. This motivates our choice to always use the most recent video frame as reference frame placed at the center of the screen surrounded by the panorama. For the final visualization (as depicted in the results section), the panorama is displayed in reference to a coordinate system, defined by the average projective transform of all images to present a panorama image with small global deformation.
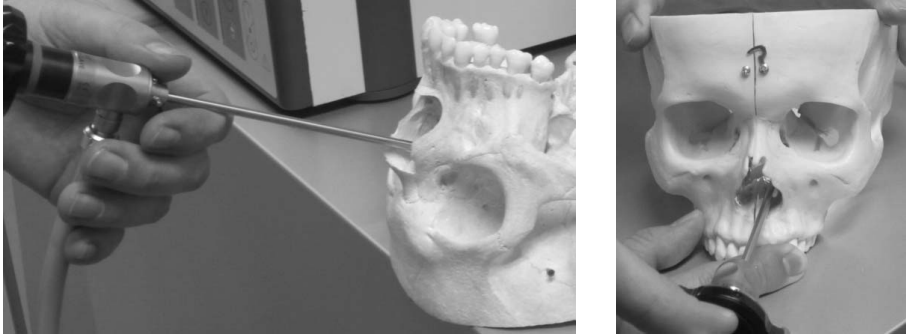
Figure 3: Nasal endoscopy on a real skull phantom (left) and a plastic skull phantom (right)

3   Results

Twenty endoscopic panoramas from monocular image sequences of the sphenoid sinus and the pituitary glands have been obtained of the skull phantoms in real-time (ref.
Figure 3). All panoramas were directly computed during slow manual movements of the endoscope tip in a translational way through the hollows. **Fehler! Verweisquelle konnte nicht gefunden werden.** depicts typical examples of five panoramic images obtained from the pituitary glands of the skull phantom. A white circle approximates the size of one original endoscopic view. Table 1 summarizes further information about the panorama images. *Frames Total* is the total number of image frames considered for the panorama generation. *Frames In Scene* is the number of frames that could be successfully registered to form the panorama. Other frames are omitted either due to insufficient quality (too few corresponding feature points) or due to the fact, that the SURF tracking is not executed with full video frame rate but processes only about every third to fourth frame. In general, this is not of any disadvantage, since it is still sufficient to provide enough overlap between frames for successful stitching. *Scene Features* is the number of SURF feature points,
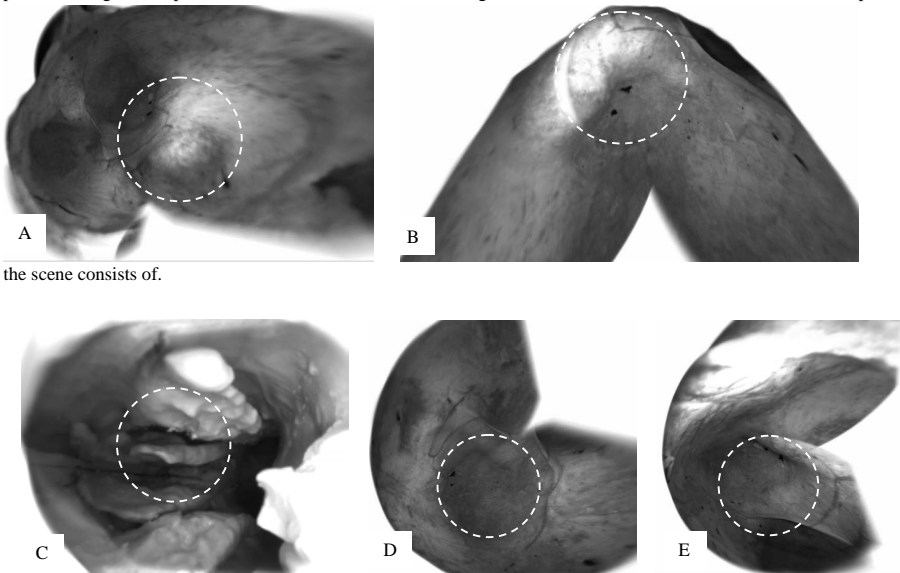
the scene consists of.

Figure 4: Five panorama images of the pituitary glands obtained from real-time stitching. A white circle approximates the size of one original endoscopic view.

Panorama	Frames Total	Frames In Scene	Scene Features
A	773	283	61292
B	541	200	46912
C	701	225	36111
D	447	183	44837
E	532	230	42034

Table 3: *Frames Total*: Number of frames considered for panorama generation. *Frames In Scene*: Number of frames successfully registered. *Scene Features*: Number of SURF features present in panorama.

4   Discussion

The results show that the proposed approach is applicable to the problem of real-time image stitching of the sphenoid sinuses. The generated panorama images from five experiments with a skull phantom have been presented, each consisting of about two to three hundred single video frames. These first experiments show the potential of the approach. Further experiments will be conducted with clinical endoscopic video sequences obtained during transsphenosoidal surgery to validate the method with real patient data.

5   Summary

By applying a real-time image stitching approach to monocular endoscopic images from a skull phantom, an augmented field of view can be provided to the surgeon. We successfully stitched several image sequences obtained by an endoscope and generated panorama views in real-time consisting of several hundred single video frames. This technique has the potential of improving orientation and maneuverability for the surgeon during difficult transsphenosoidal procedures.

6   Literature

[1]   Wittenberg, T., Winter, C., Scholz, I., Rupp, S., Stamminger, M., Bumm, K., Nimsky, C.*, 3-D reconstruction of the sphenoid sinus from monocular endoscopic views: First results*, Proceedings, Gemeinsame Jahrestagung der Deutschen, Österreichischen und Schweizerischen Gesellschaften für Biomedizinische Technik, DGBMT, Zürich, Schweiz (2006)

[2]   Burschka, D., Ming L., Ishii M., Taylor, R. and Hager, G., *Scale-Invariant Registration of Monocular Endoscopic Images to CT-Scans for Sinus Surgery*. In Proceedings of the 7th International Conference on Medical Image Computing and Computer Assisted Intervention, 413–421 (2004)

[3]   irota, D, Wang, H., Taylor, R., Ishii, M. Gallia. G. and Hager, G., *A System for video-Based Navigation for Endoscopic Endonasal Skull Base Surgery,* IEEE Transactions on Medical Imaging, **Vol.31/4**, 963–976, (2012)

[4]   Wise, S. and DelGaudio, J., *Computer-aided surgery of the paranasal sinuses and skull base,* Expert Review of Medical Devices, **Vol. 2/4**, 395–408, (2005)

[5]   Palmer, J. and Kennedy, D., *Historical Perspective on Image-guided Sinus Surgery,* Otolaryngologic Clinics of North America, **Vol. 38/3,** 419–428, (2005)

[6]   Bergen, T., Ruthotto, S., Münzenmayer, C., Rupp, S., Paulus, D. & Winter, C., *Feature-based real-time endoscopic mosaicking,* Proceedings of 6th International Symposium on Image and Signal Processing and Analysis, ISPA, Salzburg, Autria (2009)

[7]   Bergen, T., Nowack, S., Münzenmayer, C., Wittenberg, T., *A hybrid tracking approach for endoscopic real-time panorama imaging*, 17th Annual Conference of the International Society for Computer Aided Surgery, Heidelberg, Germany, International Journal of Computer Assisted Radiology and Surgery, **Vol. 8, Suppl. 1**, Springer, 352-354 (2013)

[8]   Bay, H., Tuytelaars, T., & Gool, L. V., *SURF: Speeded Up Robust Features*, Computer Vision – ECCV 2006, Lecture Notes in Computer Sciencem, **Vol. 3951**, Springer, 404–417 (2006)

[9]   Shi, J., & Tomasi, C., *Good features to track,* Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, 593–600 (1994)