Stabilization of Endoscopic Videos using Camera Path from Global Motion Vectors

Navya Amin, Thomas Gross, Marvin C. Offiah, Susanne Rosenthal, Nail El-Sourani and Markus Borschbach

Competence Center Optimized Systems, University of Applied Sciences (FHDW), Hauptstr. 2, 51465 Bergisch Gladbach, Germany

Keywords: Medical Image Processing, Endoscopy, Global Motion Vectors, Local Motion Estimation, Image Segmentation, Video Stabilization.

Abstract: Many algorithms for video stabilization have been proposed so far. However, not many digital video stabilization procedures for endoscopic videos are discussed. Endoscopic videos contain immense shakes and distortions as a result of some internal factors like body movements or secretion of body fluids as well as external factors like manual handling of endoscopic devices, introduction of surgical devices into the body, luminance changes etc.. The feature detection and tracking approaches that successfully stabilize the non-endoscopic videos might not give similar results for the endoscopic videos due to the presence of these distortions. Our focus of research includes developing a stabilization algorithm for such videos. This paper focusses on a special motion estimation method which uses global motion vectors for tracking applied to different endoscopic types (while taking into account the endoscopic region of interest). It presents a robust video processing and stabilization technique that we have developed and the results of comparing it with the state-of-the-art video stabilization tools. Also it discusses the problems specific to the endoscopic videos and the processing techniques which were necessary for such videos unlike the real-world videos.

1 INTRODUCTION

Invasive diagnostics and therapy makes it necessary for the surgeon or scientist to insert and move the endoscopic devices within the inner parts of the human body. Based on the organ to be diagnosed, the endoscopic procedure is classified as Bronchoscopy (Daniels, 2009), Laryngoscopy (Koltai and Nixon, 1989), Gastroscopy, Laparoscopy (Koltai and Nixon, 1972), Rhinoscopy (University of Tennessee College of Veterinary Medicine, 2012) and Colonoscopy. Visualization of stable videos by physicians is very essential during surgery. But the videos generated during endoscopy are very distorted and shaky due to the camera shakes occurring during insertion of an endoscopic apparatus into the human body, the internal body movements (Breedveld, 2005) like heartbeat, expansion and contraction of the organs as a response to stimuli as well as secretion of body fluids like saliva, mucus, blood etc. Miniaturization of the endoscopic devices also adds to the distortions. This makes it difficult for the surgeon to diagnose or operate. Hence, processing and stabilization of such videos becomes necessary to overcome such problems. Many mechanical video stabilization systems are available for endoscopic video stabilization commercially. These physical means of stabilisation use either a gyroscope or some other self stabilizing hardware device which holds the camera (Chatenever et al., 2000). But the drawback is that they are large, consume a lot of energy and are expensive. Another stabilzation system developed by Canon also stabilizes before the video is converted into digital data requiring larger and heavier camera making it challenging for endoscopic cameras (Canon, 1995). Digital video stabilization system is a better, cheaper and compact alternative for the endoscopic purposes. Among the available state-of-the-art digital video stabilization tools Adobe's video stabilization algorithm, which was subject to continuous development in recent years, is one of the current state-of-the-art stabilization algorithms. Adobe's algorithm is based on the well-known Kanade-Lucas-Tomasi feature tracking (KLT) method, which detects features in a start frame and tracks them over several successive frames. The tracked feature point co-

Amin N., Gross T., C. Offiah M., Rosenthal S., El-Sourani N. and Borschbach M..
Stabilization of Endoscopic Videos using Camera Path from Global Motion Vectors.
DOI: 10.5220/0004688801300137
In *Proceedings of the 9th International Conference on Computer Vision Theory and Applications* (VISAPP-2014), pages 130-137
ISBN: 978-989-758-003-1

ordinates form the trajectory paths, which represent the movements of a video camera. In the subsequent step, these shake-prone trajectories are smoothed with a smoothing method to get a shake-free video. The algorithm includes additional steps that are described in detail in the recent publications (Liu et al., 2011), (Liu et al., 2009) and (Zhou et al., 2013). Another similar method is illustrated in the algorithm which is integrated into YouTube's video editor and is also based on the KLT tracking method (Grundmann et al., 2011). The method is explained in detail in (Grundmann et al., 2012). Microsoft's algorithm worked in a slightly different way as follows (Matsushita et al., 2005), (Matsushita et al., 2006). The motion in the video is estimated by means of the Hierarchical model-based block-motion algorithm. This approach does not track the features across multiple frames to get the trajectories, but the camera path is determined from the movements from every consecutive frame. These motion vectors are used to create a trajectory, which approximate the original camera path and still contains the unwanted camera shake. This path is then smoothed to obtain a shake-free video. Detailed descriptions are found in (Bergen et al., 1992).

Research has been carried out on how well current state-of-the-art stabilization tools and algorithms perform on endoscopic videos. Luo et al (Luó et al., 2010) tested different feature detection algorithms on bronchoscopic videos. Wang et al used their Adaptive Scale Kernel Consensus Estimator to estimate better motion model parameters (Wang et al., 2008). There exist many video stabilization tools - commercial as well as freeware. However, application of these tools for stabilizing endoscopic videos proved to be ineffective: Related work still shows that even high-quality stabilization tools like the Google's video stabilization algorithm by Grundmann et al. (Grundmann et al., 2011) used in Youtube Video Editor and the Adobe Systems algorithm by Liu et al. (Liu et al., 2011) appeared to perform with moderate to unsatisfactory results, although these are the best among a set of state-of-the-art tools (Offiah et al., 2012c), (Offiah et al., 2012a),(Offiah et al., 2012b).

2 STATEMENT OF PROBLEMS

The problems with the stabilization of endoscopic videos are peculiar to the real world videos. In addition to the removal of high-frequency camera shakes from normal videos, it is important that uneven camera panning is aligned so that the output video looks smooth to the viewer. High-frequency camera shakes are also undesirable in endoscopic videos, but due to the area of activity of an endoscope in the human body, sprawling camera movements always exist. Also, in contrast to the real world video, endoscopic recordings include usually no long video sequences, but short scenes with quick image content changes. Even small oscillating movements change the image content to a large extent, because the image objects are very close to the lens. Due to these circumstances current state-of-the-art video stabilization algorithms are not ideal for stabilizing endoscopic videos. No long trajectories for endoscopic videos are available. An algorithm has to be able to compensate the high-frequency camera shake. To ensure a surgeon receives a stable video, the removal of high-frequency shake is the primary task of an endoscopic video stabilization system. Involuntary movements, like movements of body parts should not be removed whether the camera is steady or navigating. The stabilization algorithm by Microsoft deals with the stabilization of videos, which determines the camera movements from frame-to-frame. This approach is better suited for medical videos. Adobe's video stabilization uses the Content Preserve Warp Technique and approximates the camera movements with the information of the tracked feature points. However, this algorithm has significant difficulty in texture-less image regions. Only a few traceable feature points are found to determine an accurate approximation for the image distortion, which can lead to a strong erroneous warp of the image content. Erroneous deformation is completely inadmissible in medical images and thus the algorithm used by Adobe is not favourable for stabilizing endoscopic videos.

Fast movements occur very frequently in endoscopic videos. This is because the objects are located close to the camera lens leading to a very fast change of the image content as well as motion blur. In such frames many feature points are lost and a variety of trajectories break off abruptly for which the tracker must be reinitialized. In comparison to the test recordings of our laboratory, endoscopic videos include much less image features, like edges or texture, resulting in a smaller number of detected feature points. For this reason, long trajectories are normally not detected in case of endoscopic videos. This is a big disadvantage for an endoscopic stabilization algorithm, in which a large number of long trajectories for a clean stabilization is essentially required. The tracker must therefore be re-initialized much more frequently for a stabilized recording. Since the feature detection is the first and crucial step of video stabilization, a maximum number of possible feature points is required for a better image transformation procedure. Multi-frame feature tracking approach to obtain long trajectories is used in the Adobe stabilization algorithm. The long trajectories that such an algorithm demands are not available in case of endoscopic videos. Thus, multi-frame feature tracking approach is not suitable endoscopic videos. Frame-to-frame feature tracking as used in YouTube stabilizer is the solution for this problem which has been used in the current EndoStabf2f algorithm.

3 METHODS FOR VIDEO STABILIZATION

The limitations of the previously proposed methods for video stabilization and the designated problems make it necessary to develop a new algorithm suited for endoscopic videos. The videos are processed to get rid of the meta-data, like text information in the video frame and to determine the shape of the image content if it has a circular, a polygon or a full frame format (no black areas in the video frames). If the video contains a circular output format or unwanted text areas, then these regions are segmented and removed prior to stabilization. These steps fall into the area of image processing and significantly contribute to a better video stabilization.

3.1 Image Preprocessing

Because of different endoscopic types, the quality and output format of the video recordings are significantly different. The endoscopic content in the videos can be circular, rectangular or polygonal (Gershman and Thomson, 2011). Furthermore, endoscopic videos can not only contain image information but also areas with text. These are located either outside, displayed on the edge or within the video frame (Figure 2). The preprocessing step aims at achieving an improved quality and information content of medical images. This does not necessarily mean a visual improvement of the image material but an improvement of the results with respect to subsequent processing steps. For example, preprocessing enhances the quality of the videos to enable an improved and more robust feature detection in the frame. The following describes preprocessing steps which are included in our video stabilization algorithm.

3.1.1 Segmentation of Undesirable Image Regions

Some endoscopic videos contain meta-data like time or year of capture or type of endoscopy inside the video content. Text that appears in the video stays in

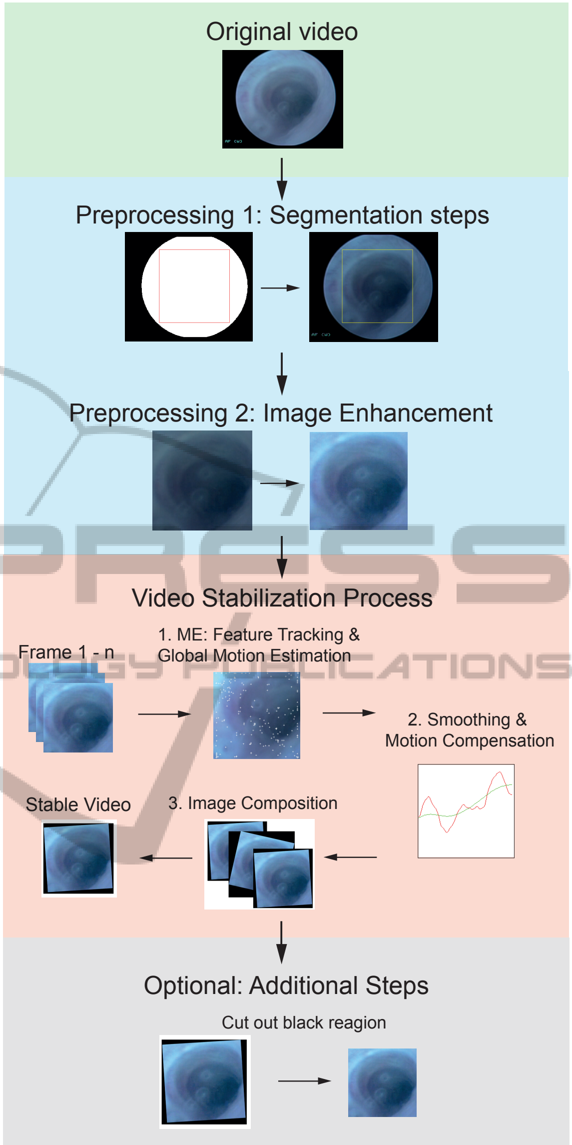

Figure 1: Video Stabilization workflow.

the same position throughout the video which could be a disadvantage for a robust feature tracking. The video in Figure 2 (Original Video Frame) contains text information in the video frame. The strongest features are detected only at this region in the video. For this reason, a segmentation step is required to segment and remove the text information from the video (See figure 2):

1. Create an averaged image from random video frames. This serves as the input image for the binary mask. Many images are used randomly and the image pixels of the video frame are averaged. The areas containing endoscopic image information are brighter. The dark black border around

the video image contains dark to black pixels and hence is darker than the region containing endoscopic content. A better definition for creating the binary mask is guaranteed.

2. Create a binary mask with chosen threshold.
3. Image processing steps: edge detection, dilation, erode, segmentation and selection of the actual endoscopic image content.
4. Reduce the size of the circular mask with x percent to get a tolerance to the blurred edges of the original video frame existing in some cases.
5. Optional: identification of the maximum square area within the video screen for better feature detection.

Finally, one gets a video frame mask showing only the important medical image content. Surrounding regions are excluded. Another algorithm was implemented, which automatically finds the maximum square area within the video frame. (See Figure 2).

3.1.2 Removal of Grid Distortions from Fiberglass-endoscopic-videos

Another automated preprocessing step is the detection of pixel grids in the video. Videos recorded with the fiberglass endoscopes, contain a clearly visible image grid due to the distinctive design of the endoscope, that covers the entire video frame. These highly prominent pixels of the image grid are often detected as robust feature points. Similarly, with the text information, these grid pixels always remain in the same location in the video frame. A robust feature tracking is thus significantly affected. A preprocessing step, using the Fourier transform, and additional image processing operations, remove the grid from the endoscopic video.

3.2 Feature-based Motion Estimation (ME Unit)

The movements of the endoscope are recorded using the KLT algorithm. But unlike the video stabilization algorithm from Adobe and Google, this algorithm does not use any anchor frames to track individual features over many frames to form feature trajectories. Instead, the global and local motion of the camera is determined only by successive frames, using the KLT with our algorithm. Thus, the ME unit determines the Global Motion Vectors (GMV) between two adjacent frames of a video sequence. From this frame-to-frame motion information a trajectory is created, which represents the approximated original camera path. This trajectory still contains the undesirable camera shake. To compensate the shake, a smoothing function is used to remove the jitters and the irregular camera movement. The great advantage of this method is that only two frames are needed to determine the motion from one frame to another. For example, in Adobe's algorithm, 50 or more frames are used to create a trajectory (Liu et al., 2011). As seen in our experiments, such long trajectories do normally not appear in endoscopic videos. The crashing of multi-frame feature tracking algorithm is described in (Gross et al., 2014). The individual steps are described in detail:

3.2.1 Global Motion Estimation

After the preprocessing is completed, the video stabilization process begins. The first frame of the video is read. This serves as the initialization frame for feature detection. The feature points are detected using the KLT method within the previously determined video mask (ROI). The output represents the x/y coordinates of the feature points (FPs) found. Next, the FPs in the second frame are detected. The two sets of feature points, each consisting of x and y coordinates are compared, and only valid feature points (feature point pairs) are used. Feature points which were not found in the previous frame are discarded. The global motion in the image is then calculated by subtracting the feature coordinate values of the first frame from the coordinate values of the second frame (only for valid feature points). In each case, all detected FPs coordinate values (separately for x and y) and motion vectors are stored in a matrix. In the next step, the second frame is used as the initialization frame and in the third frame the FPs are tracked. These steps were repeated until the entire video is processed. Always, one frame serves as the initialization frame and the adjacent frame serves as the tracking frame for motion estimation.

3.2.2 Local Motion Estimation

Sometimes local movements may appear in the video which have nothing to do with the global camera motion. For example, such local movements are the result of the movement of the surgical devices. This would affect the global motion estimation. Therefore, these motions were explored in the second step and were excluded from the calculation of the global motion. From all feature coordinates, a average global mean value is calculated. This value represents the average pixel difference of the current feature point position and its position in the previous frame. From the previously calculated motion vectors for each FP, the Mean Absolute Deviation (MAD) (Sachs, 1984)

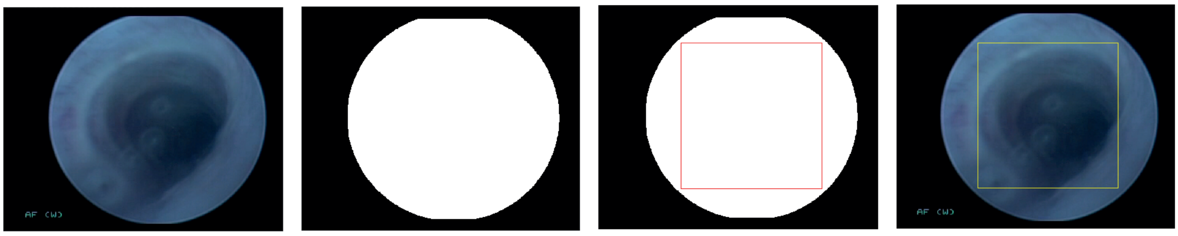

Figure 2: Segmentation workflow. Left: Original image frame. Second left: Video mask with unnecessary regions deleted. Third left: Inscribed square region in the video mask. Right: ROI for feature tracking in the video.

is calculated. This indicates how much a FP moved away from its previous location. This is calculated separately for the x and y component of the motion vectors. The values which differ in + / - the MAD value from the mean of all x and y values are discarded and not used for the subsequent calculation of the global camera motion.

$$MAD = mean_i|x_i - mean_j(x_j)| \quad (1)$$

3.2.3 Case Distinction

- If no FPs are found in the frames:
  In some cases it may occur that no feature points are detected, such as when individual frames contain too few or texture-less regions. In this case, global motion cannot be calculated and the global motion of the frame is set to zero. To use the original frame could be an unsatisfactory solution, since the path of the global movement is interrupted and as a result, the frames begin to jump. Hence, the motion vectors from the previous frame are used to avoid the jump.
- Creation of the camera path:
  The determined and stored global motion vectors are used to create the camera path which represents the movement in the video (see figure 3). As seen, the curve is still very irregular. The fine irregularities of the curve represent the high-frequency camera shake in the video, which has to be removed.

3.3 Smoothing

In the ME step all global motion vectors of the frames are stored, and formed into a global camera path. The task of the smoothing part is to determine a smoother camera path curve to ensure a shake-free video sequence. For smoothing the camera path, a non-parametric kernel smoothing method with the Nadaraya-Watson estimator is used to estimate an unknown regression function (Cai, 2001). In figure 3,

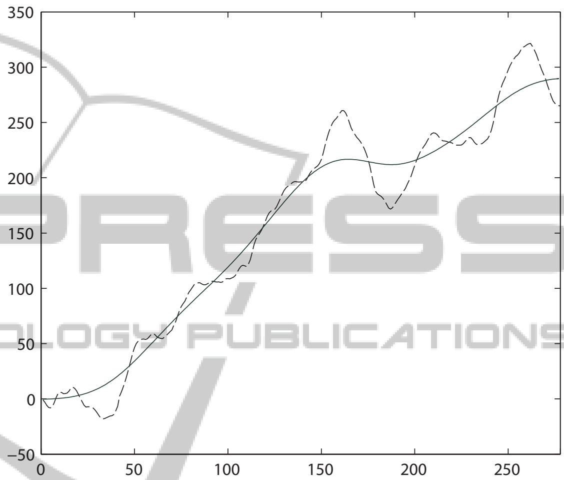

Figure 3: Motion trajectory extracted from a video. The dashed line represents the shaky camera path, the smoothed line the unshaky camera path.

the path of the global motion is shown in comparison to the determined smoothed motion.

3.4 Motion Compensation and Image Composition

To obtain a stabilized video, unwanted movements were compensated by shifting the frames from the detected feature point coordinate values obtained during motion estimation to the the values determined in the smoothing step (Figure 4). Further the existing black regions as shown on figure 4 which are a result of the frame-shifts are further removed to enable better and clean visualization of the stable video. The black regions are cropped out in case of rectangular videos. In case of circular endoscopic videos, instead of the extra black regions there exists a problem of jumping circles as a result of the motion compensation. A stabilized endoscopic video with jumping endoscopic content is not good for visualization. Such videos are further masked by calculating an average mask across the whole video. A largest circle excluding the jumping regions is calculated and the whole video is masked using this mask. Many post-processing tech-

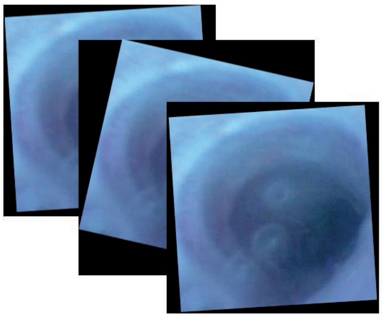

Figure 4: Motion compensated video frame.

niques for video stabilization are applied as per requirements. In case of our endoscopic videos, harsh cropping out of the black regions cause tremendous loss of important information. Hence, the use of this step completely depends on the user and the requirement.

4 EXPERIMENTS

We compare our stabilization algorithm (EndoStabf2f) with Youtube (YT) and Adobe (AE). No other digital stabilizers specifically for endoscopic videos are available. Thus we used the best performing state-of-the-art digital stabilization tools for our comparison experiments (Offiah et al., 2012a). For this, 11 test videos are used:

1. Two real world videos containing forward and backward movement of the camera.
2. Nine endoscopic videos containing different types of endoscopic distortions.

Videos are labelled according to their types (see Appendix). Videos are cropped into different sections to obtain sub-videos containing forward and backward movement, steady scenes, distortions like bubbles, foreground moving objects and circular endoscopic videos trimmed into rectangular regions. Videos are stabilized using frame-to-frame motion estimation using KLT tracker and compensated using perspective transformation.
The black regions in the edges which are a result of the compensation are cropped out in case of rectangular videos. The size of the black region to be cropped is decided on the basis of the maximum shift calculated for the x and the y axis. This trimming is done so as to make the resultant stabilized videos using our algorithm comparable to the other two algorithms. Youtube scales the videos to get rid of the black regions in the stabilized video. Adobe provides an option to apply this cropping. The exisitng black regions would affect the calculated PSNR values. Since the presence of black regions would affect the benchmarking results, these black regions are cropped out in both Adobe and our stabilization algorithm. The stabilized videos are compared across the 2 stabilization algorithms using Inter-frame Transformation Fidelity where Inter-Frame Peak-Signal-To-Noise-Ratio (PSNR) is calculted for every stabilized video.

5 RESULTS AND DISCUSSION

The results of stabilization vary for different types of videos used. The Endostabf2f algorithm is designed specifically for endoscopic videos unlike YT and AD. The stabilization by the YT stabilizer results in compressed video resulting in loss of quality, automatic region of interest selection and scaling leading to immense probable loss of some important medical information. Also, no customization of the stabilization procedure is possible as per user's requirements. Uncompressed, better quality video without any loss of information is the prime necessity of medical endoscopic videos. Thus, YT would not be successful for stabilizing endoscopic videos as per the requirements of the physician. Endostabf2f successfully stabilizes the endoscopic videos fulfilling the above mentioned requirements. We further compared our Endostabf2f with Adobe AE where customization of the stabilization procedure as per requirements is possible. However, the results show that the performance of the stabilization procedure is better for EndoStabf2f (see figure 5) when compared to Adobe AE stabilizer. The videos used for benchmarking are preprocessed for both the algorithms to enable the assessment of motion estimation and compensation. This results in better PSNR values for AE. If the PSNR values are compared without preprocessing, there is a possibility for the values to be much lower due to the existing distortions like grid, moving circle etc. The PSNR values for the grid-removed bronchoscopic videos (g,h,l and j) show that immense body movements in the video affect the quality of video stabilization making it worse. This is not in case of EndoStabf2f. The real world video "Lab video 1" (f) which contains forward and backward movements, is not well stabilized by the AE algorithm resulting in a lower PSNR value. On visualization, the shaky camera motion is not stabilized for a few frames. Similarly, in case of other bronchoscopic and real world videos (a,b,c,d and k) Endostabf2f gave better PSNR values than AE. However, in case of the rhinoscopic video with a steady camera (e) AE performed a bit better

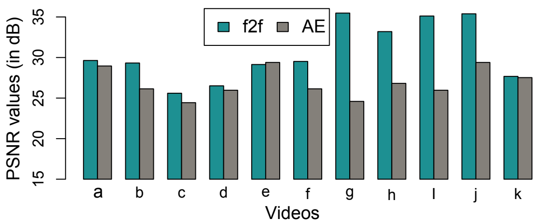

Figure 5: PSNR results for 11 videos.

than the Endostabf2f by 0.4 decibels approximately. This might be because of the available long trajectory for smoothening due to steady camera which usually not frequently available in case of endoscopic videos. Subjective assessment of the stabilized videos is required to affirm the quality of stabilized videos as required by the user. Videos stabilized by AE contain jumps in the frame during scene change. This is extreme in case of endoscopic videos since they contain very frequent scene changes. Hence, visualization of such videos misguides the surgeon prooving to be dangerous. Our algorithm in contrast takes these issues into account and is customized for the purpose of endoscopic video stabilization.

There is scope for further optimization of the EndoStabf2f algorithm with respect to motion estimation using optimized algorithms which is a part of our ongoing research. In addition, the image composition part could be optimized to exclude the black regions after compensation without losing too much of important information.

ACKNOWLEDGEMENTS

This work as a part of the PENDOVISION-project is funded by the German Federal Ministry of Education and Research (BMBF) under the registration identification 17PNT019. The financial project organization is directed by the Research Center Jülich. The work was conducted under supervision of Markus Borschbach.

REFERENCES

Bergen, J. R., Anandan, P., Hanna, K. J., and Hingorani, R. (1992). Hierarchical model-based motion estimation. In *ECCV '92 Proc. of the Second European Conference on Computer Vision*, pages 237–252. Springer-Verlag London, UK.

Breedveld, P. (2005). A new, easily miniaturized steerable endoscope. *IEEE Engineering in Medicine and Biology Magazine*, Volume 24:40–47.

Cai, Z. (2001). Weighted nadaraya-watson regression estimation. In *Statistics and Probability Letters*, volume Volume 51, pages 307–318.

Canon (1995). *Optical image Stabilizer*. Canon U.S.A. Inc., http://www.usa.canon.com/cusa/consumer/standard_display/Lens_Advantage_IS, Accessed:November 5, 2013.

Chatenever, D., Mattsson-Boze, D., and Amling, M. R. (August 1, 2000). Image orientation for endoscopic video displays. *U.S. Patent 6097423 A*.

Daniels, R. (2009). *Delmar's Guide to Laboratory and Diagnostic Tests*. Delmar - Thomson Learning.

Gershman, G. and Thomson, M. (2011). *Practical Pediatric Gastrointestinal Endoscopy (1st Edition)*. Wiley, John & Sons, Incorporated.

Gross, T., Amin, N., Offiah, M. C., Rosenthal, S., El-Sourani, N., and Borschbach, M. (2014). Optimization of the endoscopic video stabilization algorithm by local motion exclusion. In *9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, to appear.*

Grundmann, M., Kwatra, V., Castro, D., and Essa, I. (2012). Calibration-free rolling shutter removal. In *International Conference on Computational Photography*.

Grundmann, M., Kwatra, V., and Essa, I. (2011). Auto-directed video stabilization with robust l1 optimal camera paths. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Koltai, P. J. and Nixon, R. (1972). Laparoscopy-new instruments for suturing and ligation. *Fertility and Sterility*, Volume 23:274–277.

Koltai, P. J. and Nixon, R. (1989). The story of the laryngoscope. *Ear, Nose, & Throat Journal 68*, Volume 7:494–502.

Liu, F., Gleicher, M., Jin, H., and Agarwala, A. (2009). Plane-based content-preservingwarps for video stabilization. In *SIGGRAPH 2009*. ACM New York, NY, USA.

Liu, F., Gleicher, M., Wang, J., Jin, H., and Agarwala, A. (2011). Subspace video stabilization. *ACM Transcations on Graphics*, Volume 30:1–10.

Luó, X., Feuerstein, M., Reichl, T., Kitasaka, T., and Mori, K. (2010). An application driven comparison of several feature extraction algorithms in bronchoscope tracking during navigated bronchoscopy. In *Proc. of the 5th international conference on Medical imaging and augmented reality*, pages 475–484.

Matsushita, Y., Ofek, E., Tang, X., and Shum, H.-Y. (2005). Full-frame video stabilization. In *In Proc. Computer Vision and Pattern Recognition*, pages 50–57.

Matsushita, Y., Ofek, E., Tang, X., and Shum, H.-Y. (2006). Full-frame video stabilization with motion inpainting. Volume 28:1150–1163.

Offiah, M. C., Amin, N., Gross, T., El-Sourani, N., and Borschbach, M. (2012a). An approach towards a full-reference-based benchmarking for quality-optimized endoscopic video stabilization systems. In *ICVGIP 2012*, page 65.

Offiah, M. C., Amin, N., Gross, T., El-Sourani, N., and Borschbach, M. (2012b). On the ability of state-of-the-art tools to stabilize medical endoscopic video sequences. In *MedImage*.

Offiah, M. C., Amin, N., Gross, T., El-Sourani, N., and Borschbach, M. (2012c). Towards a benchmarking framework for quality-optimized endoscopic video stabilization. In *ELMAR, Sep.,2012 Proceedings*, pages 23–26.

Sachs, L. (1984). *Applied Statistics: A Handbook of Techniques*. New York: Springer-Verlag.

University of Tennessee College of Veterinary Medicine, R. (2012). http://www.vet.utk.edu/clinical/sacs/. *internal/mips/rhinoscopy.php*.

Wang, H., Mirota, D., Ishii, M., and Hager, G. D. (2008). Robust motion estimation and structure recovery from endoscopic image sequences with an adaptive scale kernel consensus estimator. In *CVPR*.

Zhou, Z., Jin, H., and Ma, Y. (2013). Content-preserving warps for 3d video stabilization. In *IEEE Conference on Computer Vision and Pattern Recognition*.

APPENDIX

Table 1: List of the videos used for stabilization.

Video	Description
a	Bronchoscopic staboptic video of a rat with circular content
b	Bronchoscopic staboptic video of a rat with rectangular content and moving camera
c	Shaky video of a hippo
d	Human Rhinoscopic 2 with rectangular content and steady camera
e	Human Rhinoscopic 3 with rectangular content and steady camera
f	Lab video 1 with forward and backward movement
g	Bronchoscopic grid removed fibreoptic video of a rat with steady camera
h	Bronchoscopic grid removed fibreoptic video of a rat with moving camera and distortion (Bubbles)
l	Bronchoscopic grid removed fibreoptic video of a rat with forward-backward movement of camera
j	Bronchoscopic grid removed fibreoptic video of a rat with rectangular content and steady camera
k	Shaky video of a tiger with jittery motion