Mosaicing of Endoscopic Placenta Images

Mireille Reeff, Friederike Gerhard, Philippe Cattin, and Gábor Székely

ETH Zurich - Computer Vision Laboratory
Sternwartstrasse 7, CH-8092 Zurich, Switzerland
{reeff,fgerhard,cattin,szekely}@vision.ee.ethz.ch

Abstract: The most effective therapy for the twin-to-twin transfusion syndrome (TTTS) in monochorionic twins is the fetoscopic laser occlusion of the anastomoses between the two fetal circulations. A complete closure of all anastomoses is a prerequisite for the success of the therapy, but the small field of view of the fetoscope makes it difficult for the surgeon to properly locate all the anastomoses. This problem can be overcome by creating a complete image of the placental surface during fetoscopy. In this paper, we present a novel approach for placenta image mosaicing. We evaluate a recently developed interest point detection and feature description method needed to automatically fuse the poor quality fetoscopic images. Moreover, an efficient scheme to improve the accuracy of the mosaic is presented.

1 Introduction

The twin-to-twin transfusion syndrome (TTTS) is a disease of the placenta, which complicates about 15% of monochorionic twin gestations. When two fetuses share the same placenta, vascular anastomoses can develop between their blood circulations. If these anastomoses are unbalanced, blood flows from one fetus (donor) to the other (recipient). Thereby, the recipient has an abnormal increase in blood volume and can die from congestive heart failure, whereas the donor has an insufficient blood supply, leading to absence of amniotic fluid and growth restrictions. If left untreated, this disease carries a high risk for fetal death (80-100%). The most effective therapy for TTTS is fetoscopic laser coagulation of the placental vascular anastomoses ([HDZ$^+$00]). A complete closure of all anastomoses, and in particular the arteriovenous anastomoses, is a prerequisite for the reduction of the perinatal morbidity and mortality rates.

The main problem of this treatment is that the fetoscope has only a small field of view, which makes it very hard for the surgeon to ensure that all the abnormal vascular connections are identified and treated accordingly. The small field of view of the endoscope can be overcome by generating a complete image of the placental surface from the individual fetoscopic images. In practice, the surgeon should pursue the following procedure. The fetoscope is inserted in the amniotic cavity of the recipient twin. Then the surgeon adjusts the focus of the fetoscope and "scans" the placental surface in between the two umbilical cords. Afterwards he removes the fetoscope and calibrates it. To calibrate the endoscope, he needs to take about 15 images of the calibration plate from different viewpoints without

changing the focus of the endoscope. The calibration procedure can also be done before the scanning. Finally, a mosaic is automatically created with the obtained images which will provide the surgeon a map of the region of interest of the placenta. This map can then be used on the one hand to locate all the anastomoses and, on the other hand, for spatial orientation during surgery. Once the anastomoses are identified, they can be coagulated either using a Nd:YAG laser or a diode laser.

Panoramic image mosaicing has been reported extensively in literature, e.g. [BL03], [SS00], [CZ98]. However, these methods are not suitable for fetoscopic images as the endoscope is inserted through a single hole, pivoted around the fulcrum and is usually very close to the curved placental surface. Additionally, its optical system generally has an inclined lens. All these effects induce strong perspective distortions, which are not present with standard cameras.

Although no literature describing the process of mosaicing endoscopic images was found, several research groups are working on the problem of mosaicing microscopic retina images, e.g. [CSRT02a], [CSRT02b], [CBGS06]. However, endoscopic images present additional challenges compared to microscopy. The endoscope used for fetal surgery is only 3.8 mm thick and has a strong fish-eye view. The resulting images suffer from strong lens distortion. Additionally, the fetoscopic images have an inferior quality compared to microscopic images, which makes feature detection more difficult.

In this paper we present a mosaicing method for fetoscopic images which is robust to inhomogeneous illumination caused by the point light source of the endoscope and is fully automatic. The images can be taken freehand without the use of a tracker and without any prior information about the placenta.

2 Methods

For our tests we utilize a progressive frame camera and a 3.8 mm/12° endoscope as commonly used in fetal surgery. To obtain realistic in vitro test images, an ex-vivo placenta was fixed in a hemispherical receptacle submerged in water and filmed with the fetoscope.

The method for creating a mosaic image from a sequence of fetoscopic images of the placenta consists of the following steps. *(1)* First, the endoscope is calibrated to obtain the intrinsic camera parameters used to remove lens distortion in the images. *(2)* Then the interest points are located in each image and their respective feature vectors are matched between pairs of images. *(3)* The transformation matrices from all images to an "anchor" image are estimated and optimized globally. *(4)* The transformations are optimized by minimizing the sum of the squared differences (SSD) of pixel intensities. *(5)* Finally, multi-band blending is used to seamlessly fuse the individual fetoscopic images. These steps are described in more detail in the subsequent paragraphs.

(1) Endoscope calibration: For this purpose we developed a fully automatic, fast and sterilizable calibration tool described in [WRCS06]. In order to cover the working depth, the calibration plate is inclined by 30°. It is important that the calibration is performed in fluid as the focal length changes with the refractive index of the medium. Water was used

Descriptor	mean precision	mean recall
SURF-65	61%	27.5%
SURF-129	67%	25.5%
SIFT-128	68%	19.5%

Table 1: Comparison of different feature descriptors

as it's optical properties reasonably approximate that of the amniotic fluid. The obtained intrinsic camera parameters are used to undistort the images of the fetoscope.

(2) Interest point detection and feature vector matching: This method is described in [CBGS06] in detail. The Fast-Hessian detector introduced in [BTG06] is used in order to identify blob-like interest points $p_{i,1}, ...p_{i,N_i}$ in each image I_i. Their respective feature descriptors are defined by $c_{i,1}, ..., c_{i,N_i}$. We compared the performance of the SIFT descriptor ([BL03]) with the SURF-65 and SURF-129 introduced in [BTG06] on our images with the evaluation method described in [MS05]. Table 1 shows their precision (correct matches/matches detected) and their recall (correct matches/existing matches). The SIFT feature descriptor showed a slightly higher precision than SURF, but it has the lowest recall, meaning that less correct matches are actually detected. The SURF-129 feature descriptor is chosen because it shows the best compromise. Finally, the interest points are matched for each image pair using the RANSAC ([FB81]) method to filter out possible mismatches from the list of putative pairs.

(3) Transformation model estimation: First, the anchor image I_A is selected as the image with the highest number of correspondences with its neighboring images. Then, an affine transformation matrix $\theta_{i,A}$ is estimated for each image. $\theta_{i,A}$ transforms a point x = $(x, y)^T$ of an image I_i to a point x'= $(x', y')^T$ in the coordinate system of I_A as follows: $(x', y')^T = \theta_{i,A}(x, y, 1)^T$. For images without direct overlap with the anchor image this mapping can only be estimated indirectly through a sequence of intermediate images with sufficient overlap. The direct and indirect transformation estimation and the global optimization are described in [CBGS06]. For our images an affine transformation model proved to be sufficient.

(4) Refining the mosaic with SSD: Because of error accumulation, images which have been indirectly mapped to the anchor frame via several images may not be perfectly matched to the final mosaic. In particular, image pairs that do not lie on the consecutive flight path of the camera but on different trajectory loops may have an overlap, but the feature matching algorithm was not able to find any correspondences and thus they do not match perfectly. Figure 1(a) illustrates a typical configuration for the initial mosaic with all connections found by the feature-based algorithm. As an example, the path from image I_w to the anchor image I_A is over 10 connections. This path would be abbreviated to 2 hops if a direct connection between I_w and its neighbor image I_4 would be found. In general, the intensity-based method described below is used to obtain the missing affine transformation matrix $\theta^n_{w,i}$ from image I_w to image I_i. I_w is an image which would need to be processed via a path length of 10 images or more and I_i is a neighbor of I_w having no correspondences to I_w yet. Once this transformation matrix is determined with the intensity-based method, the interest point correspondences between images I_i and I_w can be found. A

point $p_{i,k}$ in I_i is matched to a point $p_{w,m}$ in I_w with the nearest neighbor algorithm restricted to an area of 30 pixels around the point $\theta^n_{w,i}p_{w,k}$. The affine transformation from image I_w to the anchor image is re-estimated. After recalculating these transformations to the anchor image for all the images where new correspondences where found, the global optimization process is re-run.

The intensity-based algorithm minimizes the sum of squared differences (SSD) between the images. As SSD is sensitive to uneven lighting conditions, the images need to be preprocessed before the intensity-based matching method is performed. The image intensities are normalized and possible over-saturated light spots within the image are handled by an illumination correction method as proposed by [You00]. A rough estimate $\theta_{w,i} = \theta^{-1}_{i,A}\theta_{w,A}$ for the transformation parameters can be calculated using the inaccurate initial path that was available from the feature-based algorithm. For time and stability reasons, the image matching process is divided into two stages. First, a global search method, looking for the absolute translation between the images, is applied on a coarse image resolution of 1/10. In this case, genetic algorithms ([Mic99]) were chosen as the global search method using the SSD in the fitness function. Since we already have a rough estimate from the feature-based algorithm, the search can be restricted to finding the global minimum in the area around the initial guess $\theta_{w,i}$. In a second step, the affine transformation matrix is determined using the translation parameters available from the previous step as an initial estimate. This time, we perform the search process on the original fine image resolution and use a local gradient-descent method (Levenberg-Marquardt). This approach requires less computation time than genetic algorithms and iterates considerably faster, especially when the initial guess is close to the effective minimum.

(5) Image warping and blending: A blending algorithm has to be applied to form a seamless mosaic. The same method as described in [CBGS06] is used. In a first step, all the images I_i are transformed to the common coordinate system of the anchor image I_A. In areas were several images overlap, corresponding image pixels often do not have the same intensity and color. The interpolation during the undistortion process at the preprocessing stage and the uneven illumination from the point light source of the endoscope lead to better image quality in the center than at the border. As a consequence an advanced image blending method has to be applied to form the mosaic image. We decided to use the multi-band blending method proposed by Burt [BA83]. It preserves the fine structures on the placenta images and smoothes out low frequency variations caused by irregular illumination.

3 Results

In order to calibrate the endoscope, a series of 12 images of the calibration plate were taken in water. From these images, the intrinsic parameters of the fetoscope were calculated with the calibration algorithm described in Section 2. Typically, a back projection error of 0.6 pixel was achieved. The calibration process (taking the images and obtaining the intrinsic parameters) takes about 1 minute.

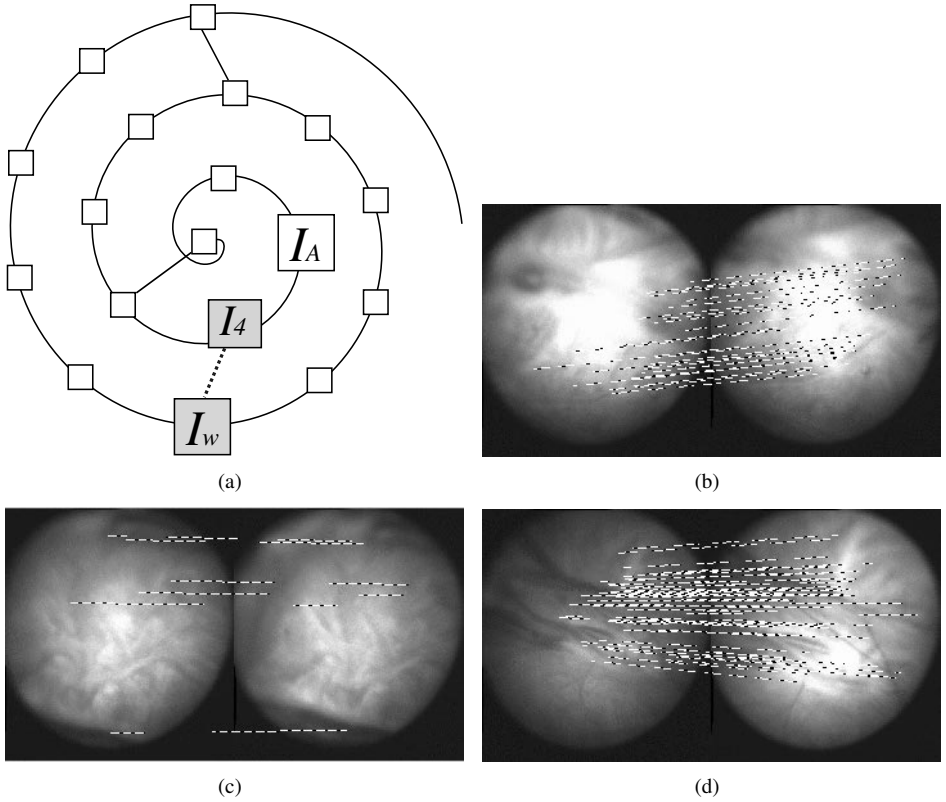


Figure 1: (a) Camera trajectory over the placenta and feature correspondences between the images. (b) Feature correspondences between two images with oversaturation, (c) with poor structure, (d) with different illumination.

To test the method, an image sequence representing 20% of the surface of the placenta was taken. The sequence was recorded freehand with the fetoscope at a frame rate of 15 frames per second. To process as few images as necessary and still have sufficient overlap, every fifth image was undistorted and used to construct a mosaic with the proposed method. We evaluated the robustness of the feature matching step with a set of 40 pairs of overlapping images. For 5 pairs with images presenting very low structural content, the matching method completely or partially failed. It also failed for 2 additional cases, although enough structural content was present. In 33 cases, the method succeeded in finding 100% correct matches, even if in two cases a lot of information was lost because the images were oversaturated (see Figure 1(b)) and in two other cases the images had very low structural content (Figure 1(c)). Figure 1(d) shows a successful matching of 2 images having different illumination conditions. Per correct image pair an average of 44 correspondences (maximum 173, minimum 6) were found. The mean overlap of those images was 68.3% (maximum 77.9%, minimum 58.7%). The time needed to extract the features in each image is on average 0.2s per image. The feature matching step takes on average 3.2s per image pair.

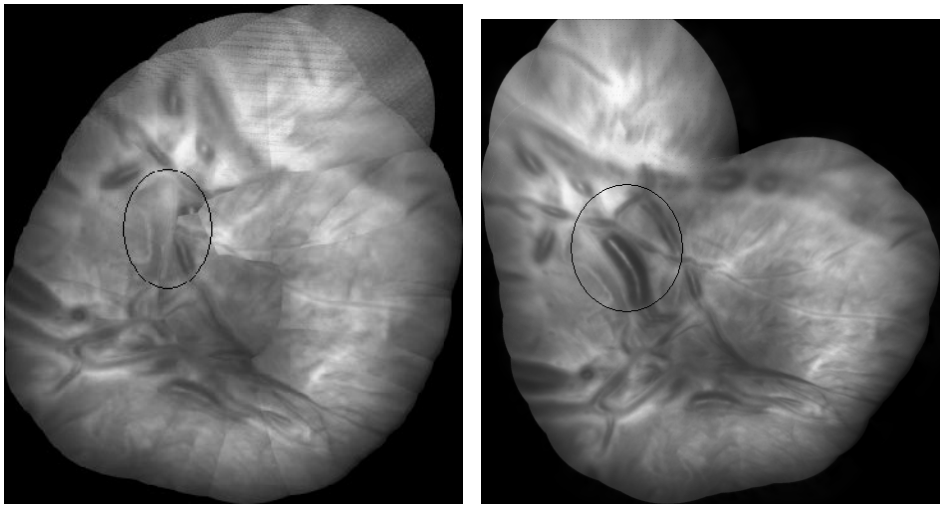
Figure 2(a) shows a mosaic composed of 35 images before the refinement using SSD and before blending. It has in total 1295 correct correspondences. Figure 2(b) shows the same mosaic after refinement with SSD and blending. It has in total 1380 correct correspondences. 85 correspondences have been added between 7 pairs of images. By comparing the two mosaics we can see that the big vessel shown in Figure 2(b) could not be seen in Figure 2(a) due to misregistration. The Matlab implementation of the method required about 60 minutes to build this mosaic, but still leaves room for further speed improvements. The time-consuming part of the mosaicing method is the refinement step using the SSD, which needs about 25 minutes in the Matlab implementation. The C++ implementation of the refinement step needs only 5 minutes. Two other time-consuming parts of the Matlab implementation are the global optimization and the multi-band blending steps.

4 Conclusion

In this paper we described an approach to automatically create a visually correct mosaic of fetoscopic placenta images. We demonstrated that interest point correspondences between pairs of poor quality fetoscopic images can be efficiently found using the SURF feature detector/descriptor. We also proposed an efficient way to find correspondences which could not be found with the feature-based method and improve the accuracy of the mosaic.

The algorithm is currently being tested on images of the entire placenta. The mosaicing method will allow overcoming the problem of the small field of view by creating a map of the entire placental surface on which the surgeon can properly locate all the vascular anastomoses, a prerequisite for the success of the TTTS therapy.

This method could also be applied to other medical fields. In the treatment of stomach cancer for example, it could be used to have a better view of the surface of the stomach when removing cancerous tissue endoscopically. One could also use the panoramic view



(a) (b)

Figure 2: Image mosaic (a) before refinement and (b) after refinement and blending

of a tissue to see the development of a disease in time or after treatment.

5 Acknowledgment

This work has been supported by the CO-ME/NCCR research network of the Swiss National Science Foundation. We thank Dr. med. Tilo Burkhardt and Prof. Dr. med. Roland Zimmermann of the University Hospital in Zurich for providing the endoscope and allowing us to take the images of the ex-vivo placenta.

References

[BA83] Peter J. Burt and Edward H. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, 2(4):217–236, October 1983.

[BL03] Matthew Brown and David G. Lowe. Recognising panoramas. In *Proceedings of the 9th International Conference on Computer Vision*, pages 1218–1225, October 2003.

[BTG06] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. SURF: Speeded Up Robust Features. In *Proceedings of the ninth European Conference on Computer Vision*, May 2006.

[CBGS06] Philippe C. Cattin, Herbert Bay, Luc Van Gool, and Gábor Székely. Retina mosaicing using local features. In *Medical Image Computing and Computer-Assisted Intervention*, October 2006. in press.

[CSRT02a] Ali Can, Charles V. Stewart, Badrinath Roysam, and Howard L. Tanenbaum. A feature-based, robust, hierarchical algorithm for registering pairs of images of the curved human retina. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24, pages 347–364, March 2002.

[CSRT02b] Ali Can, Charles V. Stewart, Badrinath Roysam, and Howard L. Tanenbaum. A feature-based technique for joint, linear estimation of high-order image-to-mosaic transformations: mosaicing the curved human retina. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, volume 24, pages 412–419, March 2002.

[CZ98] David Capel and Andrew Zisserman. Automated mosaicing with super-resolution zoom. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 885–891, June 1998.

[FB81] Martin A. Fischler and Robert C. Bolles. Random Sample Consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the Association for Computing Machinery*, 24(6):381–395, 1981.

[HDZ+00] Kurt Hecher, Werner Diehl, Liza Zikulnig, Monika Vetter, and Bernhard J. Hackeloer. Endoscopic laser coagulation of placental anastomoses in 200 pregnancies with severe mid-trimester twin-to-twin transfusion syndrome. *European Journal of Obstetrics & Gynaecology and Reproductive Biology*, 92:135–139, 2000.

[Mic99] Z. Michlewicz. *Genetic algorithms + data structures = evolution programs*. Springer, 1999.

[MS05] Krystian Mikolajczyk and Cordelia Schmid. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 27(10):1615–1630, 2005.

[SS00] Heung-Yeung Shum and Richard Szeliski. Construction of panoramic image mosaics with global and local alignment. *International Journal of Computer Vision*, 36(2):101–130, February 2000.

[WRCS06] Christian Wengert, Mireille Reeff, Philippe C. Cattin, and Gábor Székely. Fully automatic endoscope calibration for intraoperative use. In *Bildverarbeitung für die Medizin*, pages 419–23. Springer-Verlag, March 2006.

[You00] Ian T. Young. Shading correction: compensation for illumination and sensor inhomogeneities. In *Current Protocols in Cytometry, J.P. Robinson, et al.*, pages 2.11.1–2.11.12. John Wiley & Sons, Inc.: New York, 2000.