CARS 2002 – H.U. Lemke, M.W. Vannier; K. Inamura, A.G. Farman, K. Doi & J.H.C. Reiber (Editors)

Automatic 3D-Reconstruction of the Ocular Fundus from Stereo Images

Michael Wünstel und Hagen Schumann
Fraunhofer Institut für Graphische Datenverarbeitung
Fraunhoferstr. 5
D-64283 Darmstadt
Email: michael.wuenstel@igd.fhg.de, schumann@igd.fhg.de

Abstract

The methods of computer vision allow the 3D-reconstruction of an object from 2D stereo image pairs. The EU-Project Glaucad studies the glaucoma disease regarding appropriate geometric characterizations. Our task within this project is to develop a module for automatic 3D-reconstruction of the ocular fundus out of the uncalibrated stereo images. These reconstructions will be used to analyze the changes of the ocular fundus of glaucoma patients. Using methods of computer vision two main demands turned out to produce precise results: A good image preprocessing to compensate the partly changing and difficult recording conditions and a precise registration as prerequisite for the later calculation steps.

Keywords: Computer Vision, 3D-Reconstruction, Glaucoma

1. Introduction

In Germany more than 500,000 people suffer from the glaucoma disease. There exist several types of this illness. In most cases the disease is accompanied with a constantly increased intraocular pressure. This situation might finally lead to the destruction of the blood vessels and nerves in the eye [1]. Newer approaches for diagnosis try to draw conclusions directly out of the 3D-geometry of the ocular fundus. The goal of the EU-Project Glaucad is to find significant geometric characteristics which refer to the disease [2]. Precondition for this attempt is therefore the ability to reconstruct the 3D geometry of the ocular fundus.

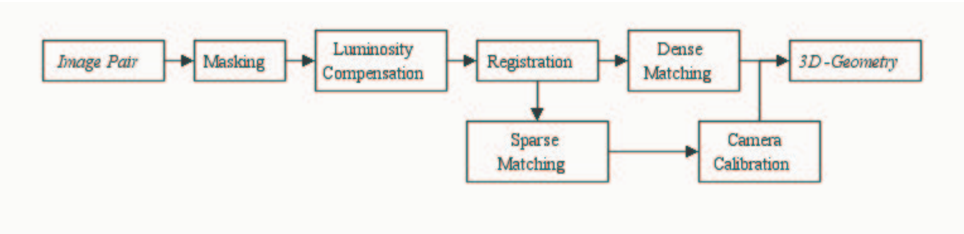
The database in this project consists of more than 15,000 stereo slides. Besides the basic requirements needed to apply methods from computer vision this fact leads to stronger demands concerning the reconstruction module:

- Firstly the whole process has to be fully automatic. Only then such a huge number of pictures can be handled.
- Secondly each process step must include an evaluation of its correctness. Due to algorithmical and numerical restrictions the calculation can lack accuracy in most process steps. If the system recognizes algorithmical problems (e.g. based on poor geometric conditions) or numerical problems an estimation of the usability of the result has to be done.
- Thirdly the algorithms have to be optimized concerning their performance.

2. Method

The basis of the reconstruction are stereoscopic image pairs of the ocular fundus. They were obtained by taking images from two different positions. There is no further information about the camera, neither about its exact positions nor its orientations given. Also the internal camera parameters like the foci are unknown.
The reconstruction process is done in several steps (see Fig. 1). The first main step is the image preprocessing (Masking, Luminosity Compensation). Here the relevant part of the image is detected and the image is processed to compensate varieties in luminosity. The second step registers the images. These two steps have strong influence on the success of the later calculation steps. The third step consists of the dense matching module where the depth information is calculated in form of a disparity map. The fourth step is the calibration step. The goal there is to calculate the camera position out of the image pair. Here we calculate a certain amount of precise point correspondencies (Sparse Matching). Out of this information the camera parameters can be obtained. Due to the difficult geometric situation this calculation lacks reliability and only can be regarded as an approximation. Therefore the output is currently done using standard camera parameters, but the potential procedure is presented. For the future it is planned to optimize these values using bundle block adjustment.

Figure 1: Reconstruction Pipe



2.1 Masking of the Single Images

In the masking step irrelevant parts of the image that would disturb the computation (e.g. patient information, legend) are masked out: To detect the dark background the input images are biniarised. The threshold used for this task is calculated out of the gray value histogram from the originally images using discriminant analysis. Then the pictures are splitted into several coherent regions using a segmentation algorithm [3],[4]. After the detection of the background region the biggest remaining region is identified as the ocular fundus. Thus an image that only contains the interesting part, the ocular fundus, is obtained. Afterwards the results undergo a plausibility check (minimum size, relative size).

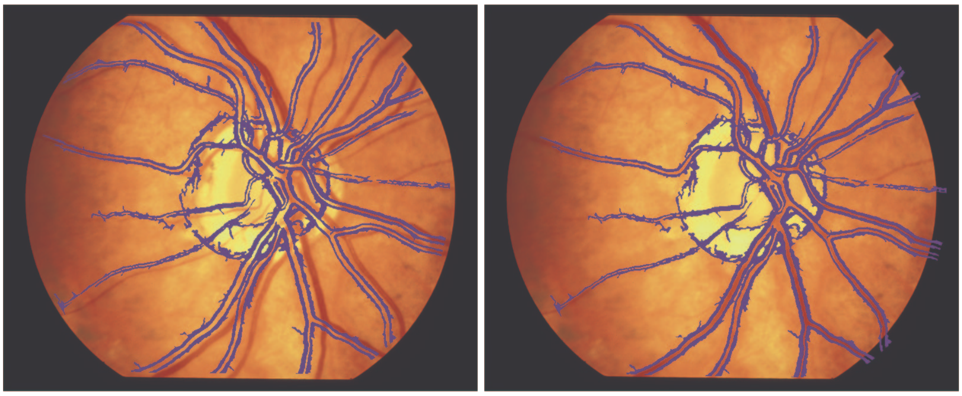
2.2 Luminosity Compensation

The process of taking photographies in the eye underlies difficult conditions. This may lead to images with a bad luminositybalance which can have negative effects on the later calculation. The compensation of luminosity results in much more contrast in weak regions.

2.3 Registration of the Image Pair

An automatic and precise registration of the images is a fundamental condition for the success of the following steps. To achieve this aim the process is splitted into two tasks: In a first step the received masks are registrated. Thus the displacement resulting from the slide framing or scanning process can be eliminated. In a second step the image contents themselves are registrated using contour based registration (see Fig. 2). The significant contours of the images (predominantly blood vessels) are extracted and biniarised. Using a Chamfer-Matching algorithm [5] the images are registrated by minimizing the distance between the two contours. To improve the robustness of this step the small and less distinct contours are eliminated.
One result is an affine transformation that describes the mentioned displacement. This is used to compensate the error of the scanning process and is necessary for a camera calibration. The second result is a transformation that gives the registration of the blood vessels and is used to make the following calculation steps more robust. After the Sparse Matching and the Dense Matching the results have to be transformed back. The assumption of an affine transformation between the image content can be seen as acceptable approximation due to the given special geometric situation (viewing direction near vertical to the ocular fundus, small depth range).

Figure 2: Projection of the contours of the second image before and after the registration



2.4 Dense Matching

The intention here is to find as many point correspondencies of the two images as possible and to detect the relative position in both images. Therefore starting from an image point in one image a template window is guided along a part of the corresponding line in the other image [6],[7]. The image parts are compared using the Normalized-Cross-Correlation (NCC) as measurement. The input are the contour registrated images. Therefore the resulting disparity values have to be transformed to their disparity values in the mask registrated images. The result is a disparity map that encodes the distance of corresponding points in gray values (see Fig. 3).

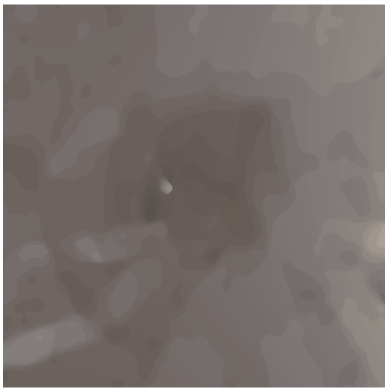


Figure 3: Disparity Map

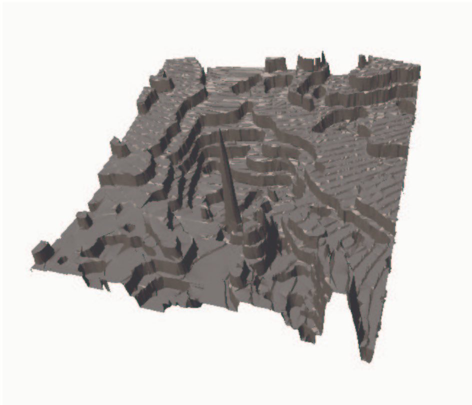


Figure 4: 3D-Reconstruction

2.5 Sparse Matching

At this point some initial point correspondencies are calculated and matched. The algorithm is based on the tracking/matching algorithm by Kanabe et.al. [8]. In a first step a certain number of special feature points in one of the images is detected. Then the corresponding points in the second image are obtained by minimizing the difference of the gray values of a correlation. To improve the robustness for the further calculation the corresponding points are forced to be equally distributed over the image. Here the input are again the contour registrated images. After the sparse matching the positions of the matched points in the mask registrated images have to be calculated.

2.6 Camera Calibration

Using the point correspondencies obtained in the previous step the Fundamental-Matrix can be calculated [9]. To improve the result an iterative process has been implemented that detects weak point correspondencies and excludes them from the calculation. Furthermore special geometric cases are taken into consideration [10].

Generally using this matrix the internal and afterwards the external camera parameters can be calculated. Due to the special geometric situation this problem is ill conditioned [11] and only can be seen as an approximate solution.

2.7 Calculation of the 3D-Geometry

Using the disparity map and the camera parameters a depth map can be obtained. For visualization of the depth information it has to be triangulated (see Fig. 4).

3. Results/Conclusion

We presented a way to reconstruct the 3D-geometry of an object only out of two standard images. In contrast to active systems like laser scanners this technique has the advantage that there is no additional hardware needed to get 3D-reconstructions. Future emphasis of our work will be the increase of the robustness and the calculating speed.

Acknowledgements

This project is supported by European Commission from the project QLG1-2000-00651.

References

1. Norbert Pfeiffer, "Eine neue Glaukomtherapie", *Forschungsmagazin der Johannes Gutenberg-Universität Mainz*, 108-117, 2/95.
2. Axel von Freyberg, Jörg Peters, "Glaucoma Prevention by Computer Aided Diagnostics", *http://www.glaucad.com*, 11. Oct. 2001.
3. H. Samet, "Connected Component Labeling Using Quadtrees", *ACM Journal*, 1981.
4. P.J. Neugebauer, C. Schimpf, "Effiziente Skulpturierung von 3D-Modellen aus Tiefenbildern", *Workshop 3D-Bildverarbeitung*, TAE Esslingen, June 1996.
5. Gunilla Borgefors, "Hierarchical Chamfer Matching: Parametric Edge Matching Algorithm", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1988.
6. Lutz Falkenhagen, "Depth Estimation from Sterescopic Image Pairs Assuming Piecewise Continous Surfaces", *Proc. of European Workshop on combined Real and Synthetic Image Processing for Broadcast and Video Production*, Hamburg, 1994.
7. Lutz Falkenhagen, "Hierarchical Block-Based Disparity Estimation Considering Neighbourhood Constraints", *International workshop on SNHC and 3D Imaging*, Rhodes, Greece, 1997.
8. Jianbo Shi, Carlo Tomasi, "Good Features to Track", *Conference on Computer Vision and Pattern Recognition*, 1994.
9. R. Hartley, A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2000.
10. Philip Hilaire Sean Torr, "Motion Segmentation and Outlier Detection", PhD Thesis, Department of Engineering Science, University of Oxford, Hilary Term, 1995.
11. M. Pollefeys, R. Koch, L. Van Gool, "Self-calibration and metric reconstruction in spite of varying and unknown internal camera parameters", Katholieke Universiteit Leuven, 1997.