ACQUISITION, ANNOTATION AND INTERACTIVE EXPLORATION OF STEREO IMAGES WITH VIRTUAL REALITY

Mohammed Haouach*'**, Karim Benzeroual*'**
Christiane Guinot*'** and Gilles Venturini*
* *Laboratoire d'Informatique, Université François-Rabelais de Tours, 64 avenue Jean Portalis, 37200 Tours, France*

****CE.R.I.E.S., Unité Biométrie et Epidémiologie, 20 Rue Victor Noir, 92521 Neuilly sur Seine, France*

Keywords: Stereoscopic acquisition, Camera calibration, Genetic algorithms, 3D visualization, Image annotation, Hypermedia, Skin relief.

Abstract: We present in the paper a system called Skin3D that integrates all hardware and software to extract information from 3D images of skin. It is composed of a lighting equipment and acquisition-based stereoscopic cameras, a camera calibration using genetic algorithms, virtual reality equipment to restore the images and interact in 3D with them, a set of interactive features to annotate images, annotations and share these 3D hypermedias. We present a comparative study and an application of Skin3D on faces skin.

1 INTRODUCTION

Relief is a complex and important data for many domains. In medicine, numerous methods have been developed in order to acquire relief of various parts of the human body with the aim of discovering information and knowledge. In this paper we are especially interested with the acquisition of a surface, and more precisely the skin. We have conceived a complete and operational system called Skin3D (see an overview in figure 1) which is compound of three main modules: (1) an acquisition module that takes stereoscopic photographs of people with skin problems or specific pathologies, (2) a camera calibration module that estimates the cameras parameters which are necessary for computing 3D information, and (3) a visualization and exploration module which can be used by dermatologists to perform 3D measurements, to create annotations as well as a 3D hypermedia, and to share the extracted knowledge with others.

In this paper, we will detail respectively each module in sections 2, 3 and 4, and we present our motivations and the state of the art for each of them. In section 5 we described the obtained results on the precision of camera calibration and a first example of an annotated 3D hypermedia build on 3D photographs of faces.

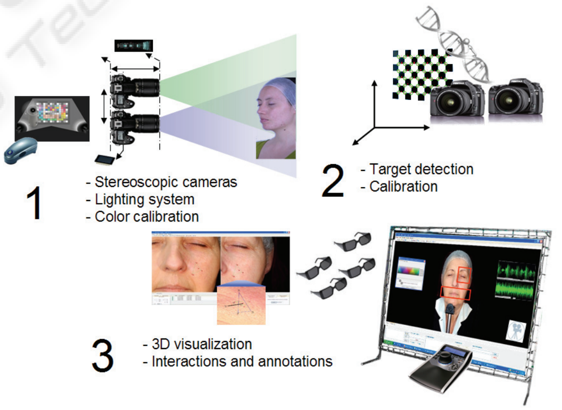

Figure 1: Overview of Skin3D.

2 ACQUISITION AND CALIBRATION MODULE

2.1 Acquisition of Relief

In the skin domain, two types of relief acquisition methods can be distinguished, the so-called active

Haouach M., Benzeroual K., Guinot C. and Venturini G. (2009).
ACQUISITION, ANNOTATION AND INTERACTIVE EXPLORATION OF STEREO IMAGES WITH VIRTUAL REALITY .
In *Proceedings of the International Conference on Health Informatics*, pages 369-374
DOI: 10.5220/0001778803690374

and passive methods (Ben Amor et al. 2005). Actives methods consist in combining an optical sensor with a source of light, like for example Laser scanners, sensors that use structured lights (Salvi, 2004), or profilometry (Rohr and Schrader, 1998). Passives methods rather use one or more images like in (Hernandez and Schmitt, 2003) or like stereophotogrammetry (D'Apuzzo, 2002). In our application, we have three main constraints: to acquire the relief in conjunction with the visual aspect of the skin, to do this with as less constraints as possible for the experimenter and the subject, to keep a high visual fidelity to the real skin. These constraints thus exclude Laser-based systems, methods that need heavy hardware, or systems that do not acquire both the relief and texture. Finally, the last constraint tends to exclude systems that perform a 3D reconstruction, because this process alters the visual quality of the acquired relief compared to a high quality photograph for instance (Ayoub et al. 1998).

For this module of Skin3D, we have conceived an acquisition system on the basis of two cameras assembled together and which are triggered in a synchronized way. We have designed a specific lighting system and we have used an optical sensor to calibrate all graphic devices (cameras, screens, video projectors, etc).

We have used two Pentax K10 reflex cameras with a resolution of 10 megapixels and with 50mm macro objective (ideal for taking pictures of faces). The stereoscopic support allows us to minimize the distance between the two cameras (the natural gap between human eyes is about 6cm) and to maintain parallelism between cameras (or a small converging angle of a few degrees only). These conditions are necessary to ensure a comfortable visualization for the user during the stereoscopic projection without important modifications of the original images.

Taking such stereoscopic pictures requires an adapted lighting system. In our first test we deal mainly with faces, so we have selected specific lights to reveal the skin relief while removing all shadows inherent the shape of the face. This lighting system is compound of two HMI torches and two Pentax AF 540 FGZ flashes.

In order to obtain the highest image fidelity both for the acquisition and visualization, we calibrate all graphic devices by associating them an ICC profile (International Color Consortium). This is performed using an i1 (X-rite) sensor and a standardized color board.

2.2 Calibration with a Genetic Algorithm

Camera calibration is a crucial step in stereovision (Faugeras et al., 1987) because it will determine the accuracy of the acquired relief. It consists in estimating the intrinsic and extrinsic parameters of the cameras (i.e. focal length, distortion, rotation/translation between the two cameras, etc). Numerous methods exist in this context (Tsai, 1987) without a real consensus, even if some algorithms are relatively common (Zhang, 1998). The type of methods we have selected consist in taking pictures of a calibration target with known dimensions, and then to estimate the parameters that minimize a target « reconstruction » error. These methods involve non linear optimization procedures which may have some problems (stability, initial starting point). This has lead researchers to make use of genetic and evolutionary algorithms which are stochastic procedures with less sensitivity. In this context, one may cite for instance (Zhang and Ji, 2001) where a single camera is calibrated, (Cerveri et al. 2001) who use evolution strategies for stereovision, or (Dipanda et al. 2003) who use one camera and a Laser.

We have developed a new calibration method, based on genetic algorithms, and which distinguishes itself from the others on the following points: it is specific to stereovision, it uses the notion of distance between points in its evaluation function (because we want to make precise measurements), it can be applied to several models of objectives (“pin-hole” model but also telecentric model). It proceeds in the following way (see figure 1, step 2): we take pictures of a target of known dimensions and with different orientations, then we detect specific points (corners) on this target in both left and right images. The distances between these points are perfectly known. The objective of our genetic algorithm is to find the set of parameters that minimizes the prediction error (i.e. the difference between known and estimated distances). For this purpose, it uses a population of individual, where each individual is a possible set of parameters. At the beginning, the population is filled with randomly generated individuals where parameters belong to loosely defined intervals. Then parents are selected according to their performance using binary tournament selection, and we recombine these individuals using a crossover operator (either a linear recombination or a discrete uniform crossover) and using a mutation operator (small random noise). The evaluation function takes into account the error between the real and estimated distances, but also other errors computed during the estimation of points (such as the intersection error, see for example (Cerveri et al. 2001)).

The parameters that we consider in the individuals are the 3D transformation between cameras, the focal lengths, the distortion and the decentering of the objectives w.r.t. the CCD. After numerous tests (especially on artificial problems where the parameters to estimate are known), we have kept the following parameters: 1000 individuals in the population, stopping of the algorithm after 30000 generations (one individual is generated per generation) when the population improvement is below a given threshold. The running times for real problems on a standard computer are about 1 minute.

3 VISUALISATION AND INTERACTIVE EXPLORATION

3.1 Virtual Reality

This module of skin3D is very important because once the data are acquired they must be restituted to the expert with the highest possible fidelity and with all interactive tools necessary for knowledge discovery. Using virtual reality is thus necessary in order to visualize the relief in stereoscopy but also to let the expert navigate in the 3D image and, for instance, make annotations. As far as we know systems for exploring and annotating stereoscopic images are rare. We may mention (Zhu, 2007) in ophthalmology but 3D images are basic anaglyphs with no real interactions.

For the stereoscopic visualization of images, we have used two types of projection hardware: on the one hand, standard cathodic screens that alternatively visualize the left and right images using active shuttering glasses, and on the other hand, two video projectors with passive glasses. We have tested two types of video projectors: the first ones (F1+ from Projection Design) use a passive polarization of light (vertical/horizontal polarizations respectively for each left and right images), and the others use F2 video projectors with Infitec filters (equivalent to red/green filters but with higher quality than basic anaglyphs). These projectors have allowed us to project skin 3D images on a 25m2 screen in front of more than 20 people.

As far as navigation is concerned, the user may move along the 3 axis X, Y et Z. All other moves are prohibited (like turning around in 3D) because this would involve a 3D reconstruction that would decrease the visual and photographic quality of the pictures. To perform these moves, the expert may use the mouse or a more specific 3D controller with 6 degrees of freedom (SpacePilot™). Using this controller is very intuitive, and the user may for instance navigate in 3D with his left hand and use the mouse with his right hand for selecting areas.

3.2 Mesuring 3D Information

One may compute the 3D coordinates of a point P on the skin thanks to the parameters estimated by the calibration module of Skin3D. Let Pl denote the projection of P in the left image, and let us suppose that this projected point was selected by the user. In order to compute the 3D coordinates of P, one has to find the point Pr, i.e. the correspondent of Pl in the right image. This is performed using a pattern matching algorithm that tries to maximize the correlation between Pl and Pr. This correlation is computed using the color values of pixels on two small images centered respectively on Pl and Pr (Chambon and Crouzil, 2003). The best candidate is the point that maximizes the correlation and that checks other constraints (for instance, the correspondent of Pr must be, in a symmetrical way, Pl). Then, using Pl and Pr, the 3D coordinates of P are known. The expert may thus measure 3D distances between two selected points. In order to measure depths (or heights), the expert selects 3 points in the image (see figure 2). These 3 points represents a plane, and the distance between this plane and a fourth point can be computed, which results in a height or depth measurement.

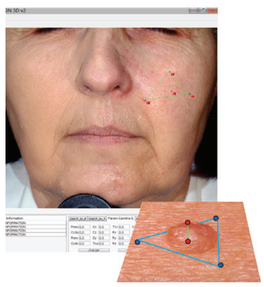

Figure 2: Measuring a depth in the stereoscopic images.

3.3 Selecting Point in 3D

The basic pointing device on a standard display is 2D and is not adapted to stereoscopic visualization because the position of the pointer is relative to the 2D screen and not to the 3D image, which is very confusing for the user. This is the reason why we have defined a 3D stereoscopic pointer in Skin3D. In order to properly position this pointer, one has to compute, for a subset of pixels in the images, the correspondence between the left and right images (see figure 3). We thus obtain a depth map for each of the selected pixels. This map can be used to correctly position the pointer in the 3D visualization. The pointer gives the user the feeling that it is flying just above the skin and that it follows the variation of the pointed relief.

Figure 3: Depth map (right) computed from the image (left) using 1% of the pixels.

3.4 Annotations of Stereoscopic Images

During the exploration of images, the expert may select some regions of interest in order to annotate them. Image annotation is currently the object of many researches, especially for automatic methods. For interactive or manual methods, one may cite for instance the work on VirtualLab (Alfonso, 2005) where microscopic pictures can be annotated, or (Chalam et al. 2006) where images can be annotated using a web interface and with several layers that allows experts to see the evolution of a pathology for instance. In this last work (ophthalmology), the authors mention the possibility to view stereoscopic images but without any details. So as far as we know, systems for annotating stereoscopic images are rare.

Skin3D includes the interactive tools necessary to associate textual or voice annotations to selected areas (see figure 4). For this purpose, the user selects a specific area (wrinkle, specific symptoms, etc.) and may define for this annotation, a title, a text and a recording of his voice. Furthermore, annotations have specific parameters: a name, a color, and a shape. They can be visible or hidden, and specific pointer events can be associated to them (display of the title when the pointer surveys the annotation, automatic zooming with a double click, etc). These annotations are recorded in an XML file (see next section).

Figure 4: Example of annotations in Skin3D.

3.5 Interactive Tour, 3D Hypermedia, Knowledge Sharing in XML

The author of annotations may generate in an intuitive way an interactive guided tour of the 3D picture. For this purpose, he may determine an ordering of the annotations and the corresponding selected areas. Skin3D may then automatically scan these annotations in the specified order, with an adjusted zoom, and with playing the recorded voice. The expert's annotations are turned into an interactive movie. In this way, the expert may underline some facts and present them first, and then he may explain their consequences. Skin3D could thus be used for teaching purposes.

Several image databases exist in dermatology like DermAtlas (Bernard, 2008) or Dermnet (Dermnet, 2007). In Dremnet, it is possible to define links between 2D images. In comparison, Skin3D manages 3D images, and allows the expert to define links between annotations. Each annotation may thus point to several others, either in the same image or in other images. The semantic encoding of hyperlinks is « *annotation@picture*». These links allow the expert to create a graph of relations between images, which may represent for instance relations between pathologies of different subjects.

For each 3D picture, all recorded information (parameters of cameras, annotations, guided tours, hyperlinks, etc) is represented in an XML file. These files can be sent by email in order to obtain complementary information from other experts. The images can thus be annotated in a collaborative way. Finally, as we will describe it in the conclusion, the XML encoding facilitates the evolution of this module in Skin3D.

4 RESULTS

4.1 Experimental Comparison

We have compared our calibration method with a camera calibration toolbox implemented in MatLab (Bouguet 2008). For this purpose, we have used pictures of the calibration target that were taken during the experiment described in the next section. 6 such pictures have been taken with different target orientations, and in order to compare the two approaches, we have used a cross validation technique: each picture is isolated in turn and is used as an unseen test case, while the 5 others are used for learning the parameters. Both methods are evaluated with the same set of detected points, and with the same error measure (difference between real and estimated distances).

Table 1: Evaluation of calibration accuracy using a cross validation technique over 6 pictures (48 points and 82 distances per picture). In underline and italic are presented the results of MatLab's « Camera Calibration ToolBox », and in bold the results of Skin3D.

Images	img1	img2	img3	img4	img5	img6
Mean Error	*142 µm* **45 µm**	*392 µm* **53 µm**	*270 µm* **73 µm**	*943 µm* **41 µm**	*330 µm* **42 µm**	*241 µm* **55 µm**
Std deviation	*86 µm* **36 µm**	*72 µm* **40 µm**	*162 µm* **53 µm**	*83 µm* **35 µm**	*83 µm* **35 µm**	*120 µm* **48 µm**
Max. error	*387 µm* **175 µm**	*559 µm* **192 µm**	*570 µm* **265 µm**	*1147 µm* **165 µm**	*546 µm* **141 µm**	*579 µm* **225 µm**

According to the results, our method has the best accuracy for all images. These results can be explained by the fact that the two methods are very different from each others: classical versus genetic optimization, optimization of parameters for each image versus all images. These results are thus very encouraging and we have planned additional comparative tests.

4.2 Real Study

In order to evaluate our system in a real world application, we have conducted a study involving 18 women from 20 to 65 years old who presented skin specificities. For each woman, we have taken 3D pictures of their face (front and both sides). For some women who presented specific symptoms, we have also taken pictures of their hands and of their back. In order to analyze the pictures, we have presented them to a panel of international dermatologists. They have used Skin3D to visualize the pictures in stereoscopy, to perform 3D measurements and to annotate the pictures. They have defined a guided tour. The possibilities offered by Skin3D (like 3D visualization and annotations) have improved the diagnostic of different skin symptoms by making the identification of specific information easier than in standard photographs.

5 CONCLUSIONS

We have developed the Skin3D system for the acquisition, visualization and interactive exploration of stereoscopic pictures in the domain of dermatology. We have described its 3 main modules. We have defined a new calibration method which, after a first experimental comparison, seems to be efficient and well adapted to our application. We have proposed the use of specific virtual reality hardware in order to visualize stereo images and to navigate through them. A test was performed with success on a large screen. We have developed several ways to perform 3D measurements, annotations and to share the discovered knowledge.

Several perspectives can be derived from this work. As far as the acquisition module is concerned, we want to acquire better cameras with a resolution of 21 Mpixels in order to increase further the accuracy of the system. We want to perform parallel computation of the pattern matching algorithm. We want to improve also the annotation process by adding specific algorithms for region automatic detection. We want to add a search engine for searching specific text in the annotations. Finally, we could study the use of domain ontology in order to help the expert for normalizing the annotations.

REFERENCES

Alfonso, B., 2005. Featured Virtual Lab in their NetWatch section, a publication by the AAAS Vol 308 Science Magazine.

Ayoub, A.F., P. Siebert, K. F. Moos, D. Wray, C. Urquhart, T. B. Niblett, 1998. A vision-based three-dimensional capture system for maxillofacial assessment and surgical planning. British journal of oral & maxillofacial surgery, vol. 36, no5, 353-357

Ben Amor, B., M. Ardabilian, L. Chen, 2005. 3D Face Modeling Based on Structured-light Assisted Stereo Sensor. Image Analysis and Processing Vol.3617, 842-849

Bernard, A., M. Cohen, U. Christoph et MD. Lehmann, 2008. DermAtlas, Johns Hopkins University. www.dermatlas.org.

Bouguet,J., 2008, Camera Calibration Toolbox for Matlab, www.vision.caltech.edu/bouguetj/calib_doc/

Cerveri, P., A. Pedotti et N. A. Borghese, 2001. Combined Evolution Strategies for Dynamic Calibration of

Video-Based Measurement Systems. IEEE Transactions on evolutionary computations, vol.5, No.3. 271-282

Chalam, KV., P. Jain, VA. Shah, Gaurav Y. Shah, 2006. Evaluation of web-based annotation of ophthalmic images for multicentric clinical trials. Indian Journal of Ophthalmology 54, 126-129.

Chambon, S., A. CROUZIL 2003. Dense matching using correlation: new measures that are robust near occlusions. British Machine Vision Conference - BMVC 2003, vol. 1, 143-152.

D'Apuzzo, N., 2002. Modeling human faces with multiimage photogrammetry. Three-Dimensional Image Capture and Applications vol. 4661, 191-197.

Dermnet©, 2007. Interactive Medical Media LLC. www.dermnet.com.

Dipanda, A., S. Woo, F. Marzani, J.M. Bilbault, 2003. 3-D shape reconstruction in an active stereo vision system using genetic algorithms. The journal of pattern recognition society, Vol.36, 2143-2159.

Hernandez Esteban, C., F. Schmitt, 2003. Silhouette and Stereo Fusion for 3D Object Modeling; Computer Vision and Image Understanding Volume 96, Issue 3, 367-392.

Faugeras, O. D., and Toscani, G, 1987. Camera calibration for 3d computer vision. International Workshop on Machine Vision and Machine Intelligence, 240–247.

Rohr, M., K. Schrader, 1998. Fast Optical in vivo Topometry of Human Skin (FOITS) Comparative Investigations with Laser Profilometry. SÖFW Journal 124, 52 - 59

Salvi, J., J. Pagès, J. Batlle, 2004. Pattern codification strategies in structured light systems. Pattern Recognition. Vol. 37, Issue 4, 827-849

Tsai, R. Y., 1987. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal of Robotics and Automation*, 3(4), 323-344.

Zhang, Y., Q. Ji, 2001. Camera Calibration With Genetic Algorithms, IEEE International Conference on Robotics and Automation. 120-130

Zhang, Z., 1998. A Flexible New Technique for Camera Calibration. Technical Report MSR-TR-98-71 of Microsoft Research.

Zhu, Y., 2007. A Java program for stereo retinal image visualization. Computer Methods and Programs in Biomedicine, Vol.85, Issue 3, 214-219.