3D Shape Recovery with no Explicit Video Projector Calibration

Tomislav Pribanic
University of Zagreb
Faculty of Electrical Engineering and Computing
Unska 3
HR - 10000, Zagreb
tomislav.pribanic@fer.hr

Mario Cifrek
University of Zagreb
Faculty of Electrical Engineering and Computing,
Unska 3
HR - 10000, Zagreb
mario.cifrek@fer.hr

Stanislav Peharec
PEHAREC Polyclinic
Biomechanic laboratory
Hotel Histria
HR - 52000, Pula
stanislav.peharec@pu.htnet.hr

ABSTRACT

3D shape reconstruction is a common task in many areas of interests. Yet, very often it involves sophisticated hardware and software. This paper represents cheap and simple system, both in hardware and software sense, for 3D subjects shape reconstruction. It consists of video camera(s) and video projector and is based on well known principle of structured light. A major contribution of the paper is method, which avoids explicit traditional video projector calibration. A system accuracy results are shown along with experimental results of reconstruction for human face and toy mushroom. The system can be easily implemented and adjusted for all kinds of other applications.

Keywords

structured light, 3D reconstruction, video projector calibration

1. INTRODUCTION

A 3D reconstruction of point in space is a task present in variety of areas and applications: entertainment, animation, industrial design, sport/medicine etc. Many applications ask for full 3D shape recovery [Sim03]. Ever developing technology nowadays offers many commercial so-called 3D laser scanners, which routinely carry out 3D reconstruction of partial/full body shape. Besides being rather accurate their major obstacle to wider use is large cost [Bou98]. Many 3D laser scanners use the concept of structured light in order to obtain depth information. The basic idea with structured light is known for quite some time [Sha03]. In brief, structured light approach assumes projection of a certain light pattern (point, stripe, grid, or more complex shape) onto a scene and then uses a CCD camera to observe how it is distorted by objects. If geometry (spatial orientation) between source of light and camera is known it is relatively simple matter to retrieve 3D information. The very same concept can be readily applied with low cost image capture and projection technology, perhaps sacrificing some amount of accuracy. For example, popular image capture and projection combination is off-the-shell CCD camera and inexpensive video-projector. Certain practical obstacle in this case is that camera and projector has to be calibrated, i.e. their spatial orientation has to be found before actual reconstruction. Not to underestimate in any way camera calibration task itself ([Web99], [Pri02]), still projector calibration can be particularly tedious regardless of being usually done by the help of camera(s) [Ram03]. The purpose of this paper is to represent method where explicit projector calibration is avoided. At the same time all aspects of structured light shape recovery are preserved as in the case where traditional projector calibration has been performed [Kon03].

WSCG POSTERS proceedings
WSCG'2004, February 2-6, 2004, Plzen, Czech Republic.

2. MATERIALS AND METHODS

There were two cameras involved and one video projector positioned between them. Both cameras were calibrated for a volume occupying 30 cm × 40 cm × 50 cm (Width × Height × Depth). The calibration volume was chosen in accordance with expected size of experimental objects being reconstructed, for instance face. For the structured light pattern one of the simplest ones was chosen: bright hard edge stripe moving from one side of monitor to another on dark background and thus via video projector scanned throughout the space, in otherwise dimmed ambient light. In fact, we are scanning the planes of light (further in text referred as light plane) in time which all intersect at line passing through the optical center of video projector.

Instead of actual video projector calibration the following was done. Dark flat plate was placed on two positions in space (in our case approximately parallel, but that is no condition). On each position a stripe scanning from one side of plate to another was imaged by pair of cameras. Finally, the object of interest was placed inside calibration volume, scanned and imaged as well. Based on acquired data (images) 3D shape reconstruction was undertaken as further explained below.

For every pixel in the image the change of gray level magnitudes vs. time, as stripe scans over it, was determined. For certain pixels due to lightning conditions change of gray level magnitudes were insignificant and those pixels were excluded from further processing. Also, small changes of gray level values coming from noise were smoothed in accordance with normal practice in image processing [Hei94].

Gray level changes actually show in time domain when stripe edge (light plane) crossed over particular pixel. That moment was determined as follows: minimum and maximum magnitudes were summed and averaged. Next, by the means of linear interpolation frame (time) value (hence, that value is likely to be non-integer) corresponding to average gray level magnitude was calculated [Bou98].

One of the key features when using hard edge stripe is to find out position in space of projected light plane at certain instance of time. In the case of calibrated video projector, based on its parameters and image coordinates of stripe being back projected in space as light plane, it is relatively simple to calculate light plane equation [Har00]. Similarly, for calibrated camera and pixel in context we can calculate line back projected in space [Har00]. Intersection of that line and light plane, at the instance of time when the plane is crossing over, ultimately yields 3D position of pixel in context. Except, in our case we did not calibrate video projector and we do not know its external and internal parameters, meaning that we cannot use it to calculate light plane positions.

Instead we proceeded as follows. The PC software routine, which moves white vertical stripe across the monitor with dark background, was written as simple Windows API (application program interface) function. It is reasonable to assume that scanning was linear in time and repeatable with respect to space/time instances. Next, let us consider again positioning of dark flat plate on two positions in space, scanning and imaging it while on each of those two positions. If we take certain instance of time and let's say first plate position we can observe projected edge images $\mathbf{t_1}$ and $\mathbf{t_2}$ (arose from light plane $\mathbf{p_P}$ and plate intersection) on both camera image planes (Figure 1).

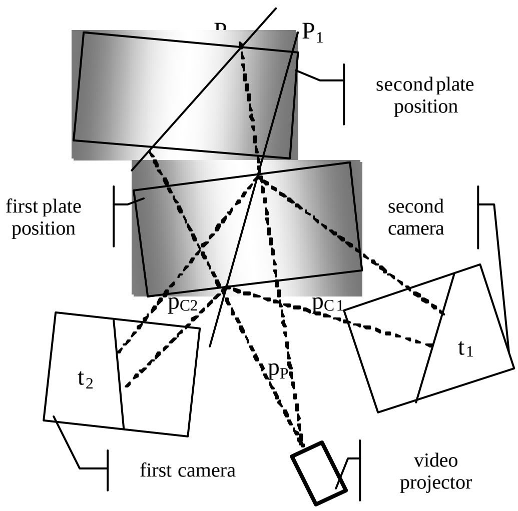

Figure 1. Light plane position determination

Namely, on each cameras image plane edge pixels were found and equations of lines $\mathbf{t_1}$ and $\mathbf{t_2}$ were calculated. Back projecting these lines for calibrated camera gives us planes $\mathbf{p_{C1}}$ and $\mathbf{p_{C2}}$ and their intersection yields equation of line in space $\mathbf{P_1}$ [Har00]. That particular line clearly belongs to light plane $\mathbf{p_P}$ whose equation we seek for. We need one more point on light plane $\mathbf{p_P}$ in order to calculate equation of it. That's exactly what second plate position is for, i.e. determination of space line $\mathbf{P_2}$ in analogues manner (Figure 1). Images of $\mathbf{P_2}$ and corresponding planes were left out for clarity. Therefore, we are able to calculate for each instant of time equation of light plane without explicitly calculating video-projector parameters. An alternative approach to calculate lines $\mathbf{P_1}$, i.e. $\mathbf{P_2}$ is to find corresponding image points on $\mathbf{t}_1$ and $\mathbf{t}_2$ and directly reconstructing. Specifically, corresponding image points can be found through the use of fundamental matrix [Moh96].

3. RESULTS AND DISCUSSION

As an experiment a human face (Figure 2) and mushroom toy were scanned and reconstructed (Figure 3, Figure 4). The shown images are gray level for practical printing purposes of this work. However, color shape reconstruction images are available on request from authors.

Figure 2. Human face while scanned by hard edge stripe; image captured by camera2

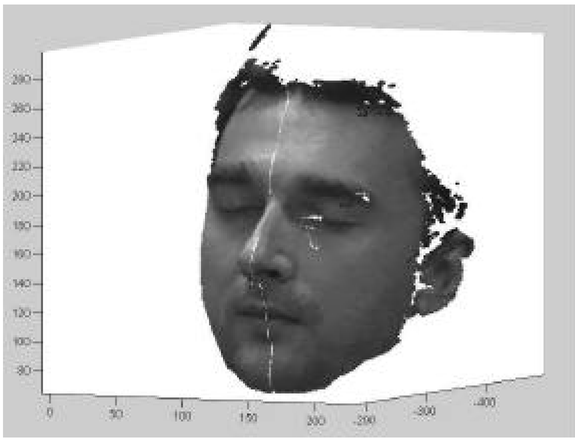

Figure 3. Face shape reconstruction using camera2 and video projector

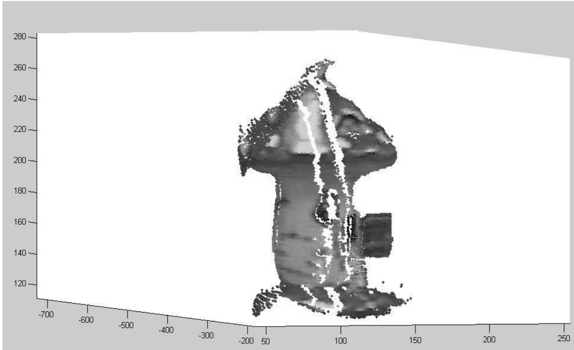

Figure 4. Toy mushroom shape reconstruction using camera1 and video projector

It was already mentioned that for some pixels change in gray level magnitude does not exceed some predefined threshold value. Another similar drawback of the system is possibility that change in gray level magnitude may significantly distinguish from ideally expected one (smooth step edge [Hei94]) when scanning with hard edge stripe (Figure 5). In those cases automatic edge determination is practically impossible. Such pixels are detected from cross-correlation between obtained and ideally expected form of gray level magnitude change. However, extensive experiments have revealed that under some lightning circumstances it is not easy to experimentally determine threshold value for cross correlation algorithm. Thus, next alternative is to use some other light pattern (i.e. sinusoidal) and see if mentioned problem can be solved.

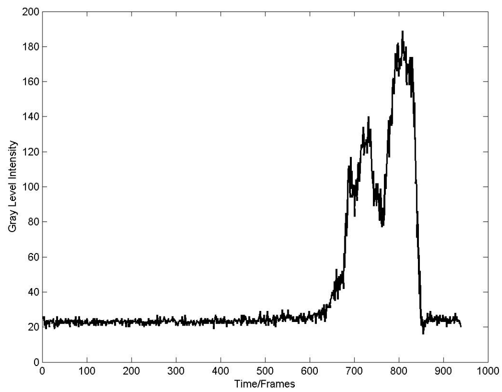

Figure 5. Change of gray level intensity for pixels excluded from further processing

There are numerous ways to test some system accuracy and precision ([Wen92], [Che94]). We have chosen the following. We placed dark plate on additional several positions in space. Each time certain number of image points, approximately occupying field of view inside calibration volume, was detected and its 3D positions were reconstructed. In ideal case all points should lie in plane determined by plate in space. However 3D reconstruction based on photogrammetric approach is subject to numerous sources of errors, primarily speaking general sense and not only when this particular system is in context ([Wen92], [Che94]). Thus, plane positions were calculated from redundant number (>>3) of reconstructed points in least square sense. Specifically, each time were taken around 4000 points for calculation of plane positions. Those points were selected randomly out of all detected image points, within the filed of view and calibration volume, which was around 60000 (Table 1). After the plane equation was found distances of all points to plane were calculated. Mean values, standard deviations (STD) and root means square (RMS) are further computed for obtained distances and representative results are given for three plate (plane) positions, i.e. trials (Table 1). Also, results are given for each camera separately, showing that camera C2 exercises slightly better performance.

Trial	Camera	Mean [mm]	STD [mm]	RMS [mm]	Points Number
1	C1	1.6	1.4	2.1	66136
	C2	1.6	1.3	2.1	59160
2	C1	1.3	1.1	1.7	66170
	C2	1.2	1.0	1.6	59026
3	C1	1.3	1.0	1.6	66064
	C2	1.0	0.8	1.3	59025

Table 1. Accuracy test results

4. CONCLUDING REMARKS

Time consuming and subject to many sources of errors explicit video projector calibration is replaced by simple manipulation of dark plate on two arbitrary positions in space. The only manufacturing prerequisite on plate is flatness, no extra marked (calibration) points on it are needed. Mathematical solution that calculates a light plane position is linear and given in closed form which assures easy and fast implementation. Obtained 3D shape reconstructions are very much satisfactory, not only by visual inspection but rather from undertaken error analysis shown in previous section. It is true that systems which video-projectors calibration parameters are explicitly known need only one camera to obtain 3D information. The same is in our case ones the light plane positions in time are known. Implementing additional camera not only avoids video projector calibration but also enables more points in space to reconstruct. The future work may be directed how the number and orientation of plates positioned throughout the volume influence results, specifically in case of larger calibration volumes.

5. REFERENCES

[Bou98] Bouguet, J-Y., and Perona, P. 3D photography on your desk IEEE Proceedings of International Conference of Computer Vision, Bombay, India, pp. 43-50, 1998.

[Che94] Chen, L., Armstrong, C.W., and Raftopoulos, D.D., An investigation on the accuracy of three-dimensional space reconstruction using the direct linear transformation technique. *J. Biomech*. 27: pp. 493-500, 1994.

[Har00] Hartley, R., and Zisserman, A. Multiple View Geometry in Computer Vision. Cambridge University Press, 2000.

[Hei94] Heijden, F. Image based measurement systems: object recognition and parameter estimation. John Wiley, Chichester, UK, 1994.

[Kon03] Koninckx, T., and L. Van Gool, L. High-speed active 3D acquisition based on a pattern-specific mesh, SPIE's 15th annual symposium on electronic imaging - videometrics , Santa Clara, California, USA pp. 26-37, January 21-22, 2003.

[Moh96] Mohr R., and Triggs B. Projective Geometry for Image Analysis. A Tutorial given at ISPRS, Vienna, Austria, 1996

[Pri02] Pribanic, T., Cifrek M., and Tonkovic, S. Comparison of three different camera calibration types. Proceedings, XIVth Congress of the International Society of Electrophysiology and Kinesiology, June 22-25, Vienna, Austria, pp. 106-107, 2002.

[Ram02] Ramesh Raskar R., and Beardsley P. A self correcting projector, Technical report, Mitsubishi electrical research laboratories, January 2002.

[Sha03] Shape from Structured Light; http://www.prip. tuwien.ac.at/Research/3DVision/struct.html

[Sim03] Simple 3D; http://www.simple3d.com/

[Web99] Webster, J.G. (ed.). Camera Calibration for Image Processing. Wiley Encyclopedia of Electrical and Electronics Engineering. John Wiley, New York, 2, 743-758, 1999.

[Wen92] Weng, J., Cohen, P., and Herniou, M. Camera Calibration with Distortion Models and Accuracy Evaluation. IEEE Transc. On Pattern Anal. and Machine Intelligence, 14, 965-980, 1992.