ilmedia

Technische Universität Ilmenau

Notni, Gunther; Bräuer-Burchardt, Christian; Heist, Stefan; Brahm, Anika;
Kühmstedt, Peter

Irritation-free pattern projection system for accurate 3D face and body scans

Irritation-free pattern projection system for accurate 3D face and body scans

G Notni[1,2]**, C Bräuer-Burchardt**[2]**, S Heist**[2]**, A Brahm**[2]**, P Kühmstedt**[2]
[1]Technical University Ilmenau, Ehrenbergstraße 29, D-98693 Ilmenau, Germany
[2]Fraunhofer IOF Jena, Albert-Einstein-Str. 7, D-07745 Jena, Germany

E-mail: gunther.notni@iof.fraunhofer.de

Abstract. Three dimensional scanning of human bodies or parts of person become more and more a matter of interest especially in medical applications. In this paper a new 3D scanning system is introduced which is absolutely irritation-free based on the structured light illumination by aperiodic sinusoidal fringe patterns in the near infrared range. Hence, it is particularly suitable for continuing human face scanning. The system is based on optical stereo capturing and structured light illumination using the GOBO principle. The measurement principle of the system is explained and examples of face and body scans are given. The measurement accuracy was evaluated by experiments and will be presented and discussed as well as an outlook to future work.

1. Introduction

Three dimensional acquisition of the human face and body is a growing task in medicine, animation, virtual life, and computer games and entertainment. Gestures and facial expressions can be used for communication tasks or machine control. Although gestures and facial expressions can be detected in 2D image sequences yet, 3D data may provide more detailed information which for example can help to increase the power of gesture and expression analysis systems.

Optical scanners based on structured light projection technique are increasingly used for gesture and facial expression detection and recording [1, 2]. This technique provides the advantage of relative high spatial resolution and measurement accuracy. Recently high-speed camera technique provides 3D frame rates up to 1 kHz [3, 4]. The basis for many different fields of application of 3D person scanning is the technological development of hardware and software components for suitable 3D scanning systems. The scanning technology may depend on the fact, whether static or moving objects has to be scanned.

The 3D scanning of moving persons was strongly shifted into the focus of recent research. The 3D scanning of static objects or non-moving persons has its history mainly in medical applications such as computer-assisted tomography (CAT) [5] and nuclear magnetic resonance imaging (NMRI) [6].

An important requirement to 3D scanning systems for a permanent acquisition of human faces is the property of irritation freedom, i.e. the scanning process must not disturb the scanned person. This could be achieved using passive multi-camera systems. However, structured light illumination provides a very higher degree of object resolution and measurement accuracy. The approach for irritation free scanning using structured light illumination is using wavelengths at near infrared (NIR).

---

[1]   To whom any correspondence should be addressed.

2. The pattern projection system

Fringe projection profilometry is one of the established methods of structured light projection (see e.g. [7]), but other structured illumination patterns (binary dot patterns, speckle patterns, aperiodic sinusoidal patterns, see [8]) are also suitable for this technique. This measurement principle works well using wavelengths over 780 nm. However, illumination sources should be selected accordingly or spectral filters must be used in the case of projected pattern generation. The recording process by one or more cameras should be optimized according to the used wavelengths.

The pattern projection system is based on the measurement principle of stereo observation and structured light illumination. It consists of a projection and an observation unit. The observation is done by a stereo-camera arrangement with blocking filters for visible light frequencies. The projection unit projects structured light patterns onto the measurement object which help to identify corresponding points in both cameras.

From the several possibilities for the used pattern for the structured illumination we applied aperiodic sinusoidal fringe patterns [3, 4]. The generation of the pattern sequence (n patterns) was realized by a so-called GOBO projector, that performs the projection by a rotating wheel with radially arranged fringes, illuminated by a high power LED lamp (see figure 1).

The GOBO principle is particularly known in the theatre-industry and means to place a partly transparent slide with a certain pattern is placed before a light source. The name GOBO is derived from "GOes Before Optics", describing that the pattern-depending slide object is positioned between the light source and the lenses. Usually, a GOBO projector contains a light source (e.g. a gas-filled tube or halogen lamp), a light collector (ellipsoidal reflector or condenser), a slide (e.g., made of steel or glass), and an imaging optical system (with fixed focal length or zoom lens) as optical components. For more details concerning the GOBO principle see [3].

The high power LED at 850 nm emission wavelength was used and homogenized with a light funnel made of mirrors with a protected gold coating. The aperiodic sinusoidal patterns were generated using a glass GOBO wheel (according to Heist et al. [3]) on which the pattern was made of a chrome coating. Thus, only the near-infrared radiation is projected into the measurement plane for an irritation-free measurement. Two CMOS cameras, each with a maximum resolution of 4 MPx were arranged in stereo-vision setup with a baseline of approximately 300 mm. For the working frame rate of 360 Hz, a resolution of only 1Mpx (in the image centre) is used. Exposure time is typically 2.7 ms.

**Figure 1.** Principle of the irritation-free pattern projection system
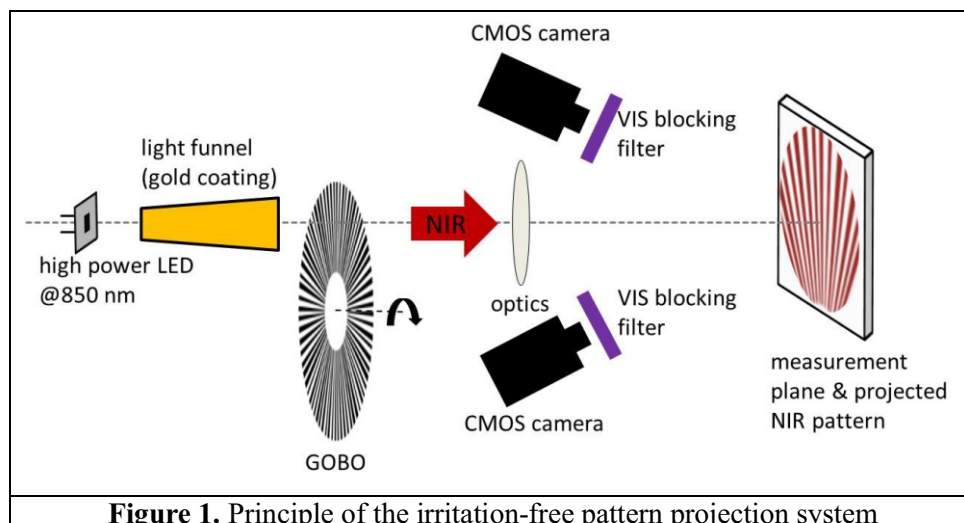
To avoid overexposure and disruptive effects from the visible wavelength range, both cameras were equipped with an additional band-pass filter with cut-on and –off wavelengths of 830 nm and 870 nm, respectively. The working distance of the system can be scaled from 0.5 m to 1.5 m in which

the measurement area varied from 0.3 x 0.3 m^2 to about 1.0 x 1.0 m^2, depending on the optics. The image acquisition was done by a frame grabber via Camera Link connection.

In order to get colour information additionally to the 3D data a colour camera is included in the system. The additional colour camera also has a resolution of 4MPx and can capture images with frame rates up to 90 Hz. Colour data are recorded with a resolution of 4 MPx with lower frame rate than the measurement cameras and transferred via USB 3.0.

The connection between the calculated 3D data and the colour information is realized by a 2D-3D mapping. The corresponding transformation matrix is obtained by inclusion of the colour camera into the calibration process. The complete 3D scanning system is shown in figure 2.
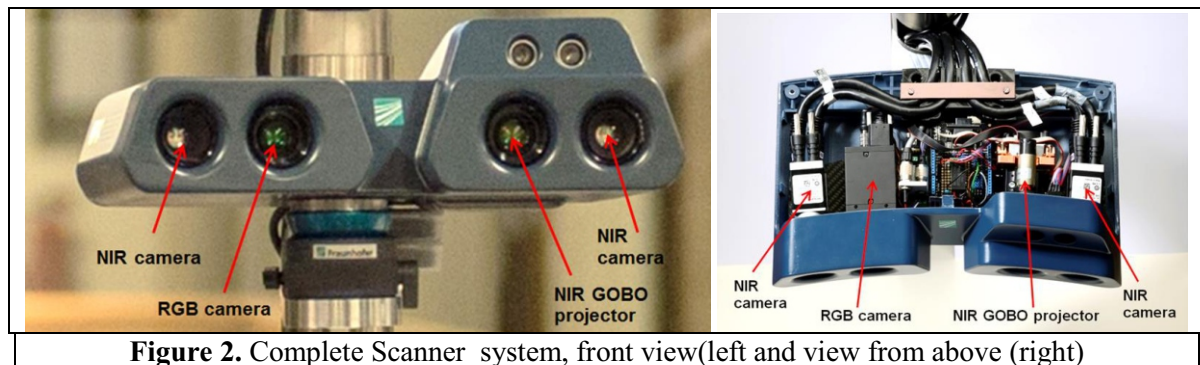
Figure 2. Complete Scanner system, front view(left and view from above (right)

The recorded data are processed to 3D data very quickly using a GPU-processing procedure (NVIDIA Geforce 1070 GTX) and the resulting images are displayed on a monitor in real-time. The latency from image recording to 3D point-calculation is only a few more than 100 ms. Furthermore, displaying the data on the screen needs some more time, depending on the hardware properties.

3. Measurement system evaluation

In order to evaluate the accuracy of the obtained 3D data the 3D scanner was used to perform measurements with certain specimens with known geometric properties. A flatness standard made of ceramic with a size of 300 x 50 x 40 mm³, a ball bar with length between the sphere centre-points of about 200 mm and a sphere with a diameter of 50.1 mm were used as specimens (see figure 3).

In order to quantify the measurement accuracy the characteristic parameters flatness measurement error ΔF, length measurement error ΔL, and probing error P_s were determined similar to the VDI/VDE guidelines [9] concerning optical measuring systems. The measurement volume was set to a size of approximately 0.5 x 0.5 x 0.5 m³. Sequences of either six or ten aperiodic sinusoidal fringe patterns were used. For the determination of each of the characteristic parameters at least six measurements were performed. The characteristic parameters were determined similar to the VDI/VDE guidelines with the supplement of complete noise removal at the determination of the flatness parameter ΔF by subtraction of six times standard deviation σ in order to exclude the influence of noise.

The parameters were defined as follows. Flatness measurement error is the maximum value of flatness deviation ΔF_i at n certain positions i of the flatness standard in the measurement volume.

$$\Delta F = \max_{i=1,\ldots,n}\{\Delta F_i\} \quad \text{with} \quad \Delta F_i = d_i^{\max} - d_i^{\min} - 6 \cdot \sigma_i \tag{1}$$

Here, d_i^{max} is the highest positive and d_i^{min} is the highest negative Euclidean distance of the 3D points from a fitted plane and σ_i are the standard deviations of the Euclidean distances (noise).

Length measurement error is defined as the span of the length measurements of the ball bar:

$$\Delta L = L_{\max} - L_{\min}, \tag{2}$$

where L_{max} is the largest and L_{min} is the smallest of the measured distances between the centre-points of the spheres of the ball bar.

Probing error P_s is defined as the span of the radius measurements of the sphere:

$$P_s = R_{\max} - R_{\min}, \tag{3}$$

where R_{max} is the largest and R_{min} is the smallest of the measured radii of the sphere.

The measurements were performed using both six-image and ten-image sequences. The following characteristic parameters were obtained: $\Delta F = 0.6$ mm, $\Delta L = 2.0$ mm, and $P_s = 0.25$ mm for the six-image and $\Delta F = 0.6$ mm, $\Delta L = 1.3$ mm, and $P_s = 0.08$ mm for the ten-image sequences.

Figure 3. Specimens with known calibrated dimensions for evaluation measurements: flatness standard (left), ball bar (middle), and sphere (right)

Additionally, the noise of the 3D measurement points was defined as the standard deviation σ of the Euclidean distances of the calculated 3D points from a fitted plane at the measurements of the flatness standard. The mean noise value was $\sigma = 0.11$ mm (n=6) and $\sigma = 0.09$ mm (n=10). These values showing the good measurement accuracy compared to 3D scanning systems in the visible light range and/or single shot systems.

4. Measurement examples

Figures 4 and 5 show some examples of the measurements performed by the developed system. The first example shows a female face showing some emotional expression. A subsequently applied software module can detect certain facial points. Figure 4 shows a false colour representation of the 3D data and figure 5 shows the colour mapping of the data obtained by the colour camera onto the 3D data together with certain detected facial points. The colour mapping onto the 3D data allows a representation from an arbitrarily chosen viewing direction onto the face.

The 3D models obtained of these measurements can be used for detection of emotions based on the 3D face data.
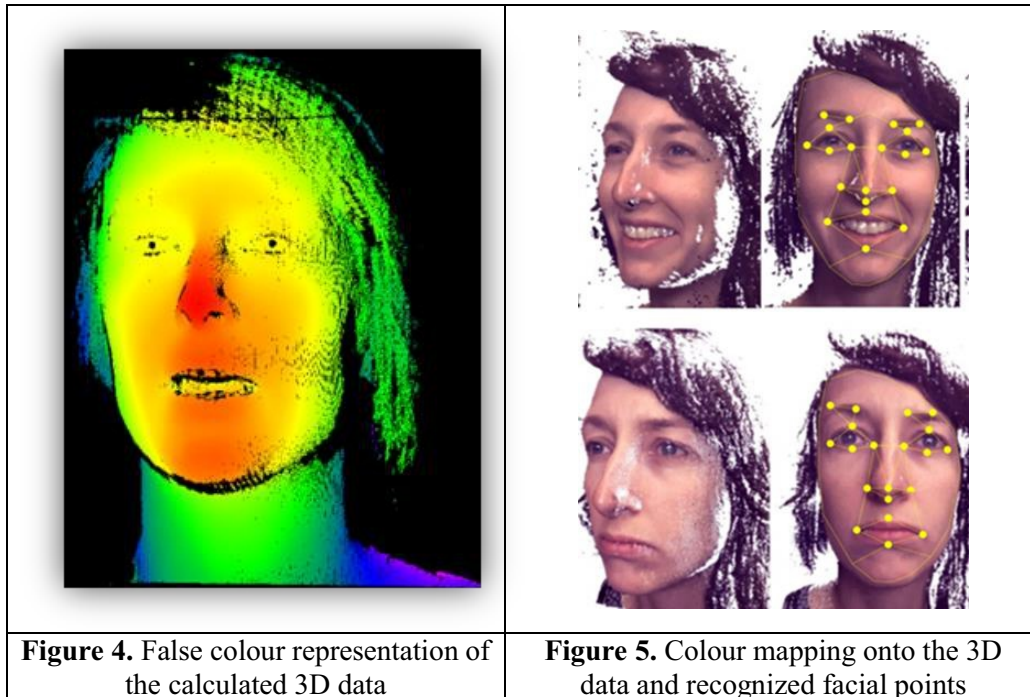
The scanner was also used for quantification of lip-movements in order to support lip-reading [2]. Figure 6 shows an example of the determined trajectories in the three different spatial directions. Note, that lip movement in z-direction cannot be captured using only 2D image sequences.

It could be stated, that using 3D information about 25% more information can be obtained for recognition of the speech than using only 2D image sequences. More detailed information is given by the work of Brahm et al. [2].
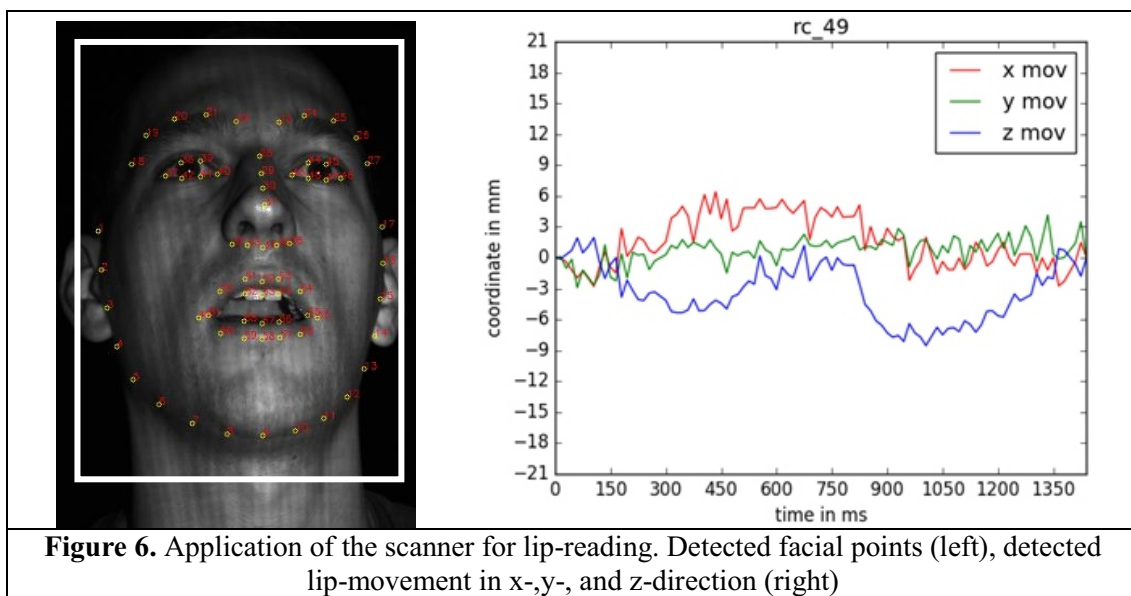
5. Discussion and outlook

The measurement results show that the developed scanning system can acquire robust 3D face images of moving persons, obtained with a very short latency time of little more than 100 ms by processing of

image data of 2 x 1 MPx recorded at 360 Hz. Consequently, a 3D rate of more than 30 Hz is obtained and displayed on the monitor.

Figure 4. False colour representation of the calculated 3D data	**Figure 5.** Colour mapping onto the 3D data and recognized facial points

The projection system is suitable for fast irritation-free 3D detection of moving persons, the acquisition of faces, and body parts with very short latency time. As the accuracy measurements show the system is not designed for high precision measurements of rigid body parts in the lower µm-range. In order to set the focus to highest measurement accuracy a pattern sequence with n≥10 can be used (see e.g. [4]).

Figure 6. Application of the scanner for lip-reading. Detected facial points (left), detected lip-movement in x-,y-, and z-direction (right)

However, if maximum available 3D rate is not necessary and the actual movements are slow, noise can be reduced and robustness can be improved by using a higher number of images per sequence.

Here, a compromise should be found depending on the requirement of the measurement task and the power features of the scanning system.

Future work will be addressed to the further development of 3D processing algorithms, e.g. in order to increase robustness, suppress noise, and decrease the number of outliers (namely false positive detected 3D points). This may for example include optimization of the used aperiodic sinusoidal fringe pattern. For more details concerning these patterns see the work of Heist at al. [4]. Additionally, an appropriate analysis of the origins of outlier production should be performed initially.

Future work will also include the interpretation of facial expressions and body movements like gestures.

6. Conclusion

A new irritation-free pattern projection system for accurate 3D face and body scans was presented in this paper. The suitability of this 3D scanning system for acquisition of the expressions of facial emotions and the support of automated lip-reading could be shown.

The high possible 3D frame rate of 36 Hz with high resolution ($\sigma \approx 100 \mu m$) makes the systems convenient for high-speed and high resolution 3D-face- and body-scans, e.g. at the perpetual 3D acquisition of body parts of moving humans.

References

[1]   Beumier C 1006 3D face recognition IEEE 2006 pp 369-374
[2]   Brahm A, Ramm R, Heist S, Rulff C, Kühmstedt P and Notni G 2017 Fast 3D NIR systems for facial measurement and lip-reading Proc of SPIE Vol. **10220** pp 102200P-1-9
[3]   Heist S, Lutzke P, Schmidt I, Dietrich P, Kühmstedt P and Notni G 2016 High-speed three-dimensional shape measurement using GOBO projection Optics and Lasers In Engineering **87** pp. 90–96
[4]   Heist S, Mann A, Kühmstedt P and Notni G 2014 Array-projection of aperiodic sinusoidal fringes for high-speed three-dimensional shape measurement Opt. Eng. **53** (11) pp 112208-1-12
[5]   Brooks R A and Di Chiro G 1976 Principles of computer assisted tomography (CAT) in radiographic and radioisotopic imaging Physics in Medicine and Biology **21**, No 5 pp 689-732
[6]   Rothwell. W P 1985 Nuclear magnetic resonance imaging Applied Optics Vol. **24**(23) pp 3958-3968
[7]   Salvi J, Fernandez S, Pribanic T and Llado X 2010 A state of the art in structured light patterns for surface profilometry Pattern Recognition **43** pp 2666–2680
[8]   Zuo C, Chen Q, Gu G, Feng S, Feng F, Li R and Shen G 2013 High-speed three-dimensional shape measurement for dynamic scenes using bi-frequency tripolar pulse-width-modulation fringe projection Optics and Lasers in Engineering **51** pp 953–960
[9]   VDI/VDE 2634 2008 Optical 3D-measuring systems VDI/VDE guidelines, Part 2