Research on 3D Sensors - Minimizing the Environmental Impacting Factors

Peter J. Neugebauer

Polygon Technology GmbH
www.polygon-technology.com
Peter.Neugebauer@polygon-technology.de

Abstract: With 3d sensors rigid objects can be nowadays measured very precisely. For measuring faces of living humans some restrictions have to be taken into account. First, the faces are non-rigid objects and hence we have to take care of object movements while scanning. Second, the amount of light projected onto the faces is limited due to eye safety restrictions. Furthermore, a practicable method must be as stable as possible in the presence of uncontrolled lighting conditions in the environment. In this paper we consider active methods which are suitable for measuring human faces and consider the limits arising from eye safety restrictions. As a results a prototype is described which fulfills the above requirements.
Keywords: Active methods for 3d measurement, Eye safety, Building a prototype for 3d face scanning

1 Introduction

The 3d scanning of human faces yields valuable information for verifying persons in a border or entrance control scenario. In general, active methods for 3d acquisition devices are meanwhile well known and very precise for rigid objects ([Wio01]). One of the most popular methods are based on structured light approaches (see fig. 1). They have the advantage that the captured 3d information is precise and dense while the scanning time is still moderate. In order to apply these methods for slightly moving objects, like human faces, we have to take several restrictions into account:

1. The capturing method should be accurate.
2. The capturing method should be very fast.
3. The amount of light projected onto the face must be eye safe.
4. The lighting conditions in the environment should be uncontrolled.

Unfortunately, the above requirements contradict each other. Typically we have the situation that the more accurate a method is, the more capturing time and more light for the active component is needed. In order to be independent of the environmental lighting

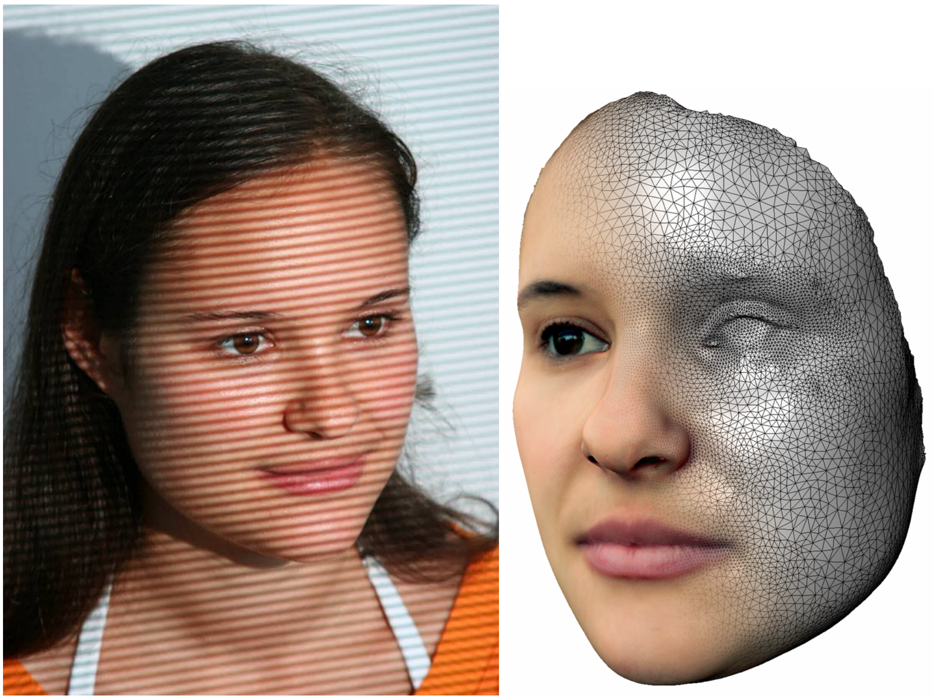

Figure 1: 3D Scanning of human faces with a structured light method.

conditions, the projected light should be as bright as possible, raising the question of eye safety. Eye safety can be achieved by limiting the amount of projected light. Thus, we have to look at the eye safety constraint first. Then, we know how much light is allowed for an active projection. As a consequence we know how much controlling of the environmental lighting conditions is necessary, e.g. by building up a measuring cabin or by blending the daylight with an appropriate facing.

In order to implement a very fast method for 3d scanning, a suitable pattern sequence with only few patterns must be choosen. Of course, the pattern sequence has a direct impact on the scanning speed, accuracy and noise of the acquired range data, and the robustness against object movements.

In order to prevent spoofing scenarios, the 3d scanning should allow a liveness detection of the scanned face. Therefore, a 3d video of the captured face will be very useful.

2 Active methods for 3d scanning

Active methods are very common for 3d measurement. Active methods have an active component, like a projector (or a laser), along with at least one camera to gain the 3d data (see fig. 2). By projecting different stripe patterns onto the face and locating their position in the camera image, depth information can be extracted, exploiting the deformation of the stripes on the curved surface. Typically, multiple patterns are projected sequentially onto the object.

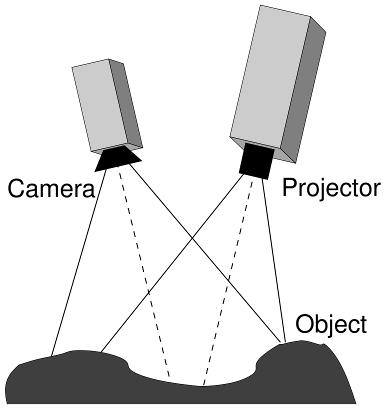

Figure 2: Principle of a 3D Scanner based on the structured light method.

Since the stripe patterns are coded with different brightness, it is essential that the contrast is high enough for each pixel. Pixel without sufficient contrast are not considered reliable. As a consequence the acquisition quality directly depends on the global lighting conditions. Optimal results can only be achieved under controlled lighting conditions with constant light, since the contrast in the individual pixel is influenced by the lighting.

2.1 Structured light pattern sequences

Finding of the best pattern sequence is sometimes called the Holy Grail for structured light methods. In [SPB04] a good overview over different coding schemes is given. When multiple patterns are used for coding we talk about time-multiplexing coding. The important members of this pattern families are binary codes, n-ary codes and Gray code combined with phase shifting methods. Fewer patterns are needed with the so called spatial neighbourhood coding, like non-formal coding, random patterns or the important class of De Bruijn sequences. Fig. 3 shows an example of a color coded De Bruijn sequence.

The methods differs in the number of patterns, the robustness and color versus gray scale coding. Whereas the spatial neighbourhood coding requires only very few patterns, the 3d information is locally easily corrupted due to self-occlusions, outlier pixels or steep surfaces. Color coding reduces the number of patterns but makes the coding scheme more fragile to uncontrolled lighting conditions in the environment.

In [Wol04] it is pointed out that phaseshift methods are in principle a good coding scheme for slightly moving objects, because the measured phase value has some stable properties

Figure 3: Example of a De Bruijn color pattern [ZCS02].

in the presence of movements. The disadvantage of the pure phase shift method is that the phase values are only unique up to a multiple of the wave length. Hence the phase shift method must be combined with e.g. a Gray code method or with a multi-wavelength phase shift method. The multi-wavelength phaseshift methods simply projects a sequence of three different wavelengths onto the object. By combining the different wavelengths, the result of the combination is unique in the measuring volume. [Wol04] introduces a cyclic pattern sequence combined with a phase unwrapping method in order to be stable even in the presence of slight object movements.

The phase shifting algorithms have the important feature that the contrast of the projected pattern is the better the smaller the wavelength is. On the other hand the phaseshift algorithm allows object movements of up to $1/4$ of the wavelength in order to yield a stable result for one phase, meaning that objects movements can be tolerated even better for larger wavelengths.

2.2 Methods suitable for 3d video

For liveness detection a 3d video stream seems to be a valuable feature of a face scanner. Thus, we should be able to extend the 3d measuring method for 3d video. Special methods [HHR03] were developed for real time capturing of the 3d geometry. However, we prefer methods with the ability to update an existing 3d image very quickly, allowing also to capture completely new scenes. Going back to the last subsection, phaseshifting combined with Gray code is not an option, because the incremental update is not robust enough in practice. As stated in [Wol04] cyclic multi-wavelength phaseshifting methods can be a solution for this problem. With these methods the measurement of one phase can be typically done with only four sine patterns. The uniqueness is calculated by a spatial/time phase unwrapping algorithm. By the multi-wavelength approach the generation of enough new uniquely determined points in 3d space is guaranteed, allowing an incremental update of all other points with a phase unwrapping algorithm.

In practice, a single facial 3d image is reconstructed after a sequence consisting of three wavelengths, each consisting of four sine patterns. Then, an update of the slowly changing scene can be done with only four new patterns by taking the last two phase images into account.

3 Eye safetey considerations

The amount of light which can be projected onto the human face is not limited because of technical reasons, but for safety reasons of the scanned persons. The eye safety restrictions are described in the european international norm [IEC01] with the corresponding national norm [DIN01]. In addition, [BGI04] describes the exposition limit values. The norm IEC 60825 describes laser sources with coherent light. LED light sources also fall into this norm because no special norm for them exists yet. The norm provides limiting values for the maximum expositions and a classification scheme into different safety classes. When the accumulated expositions can be accepted for more than 100 sec., the device is classified as class 1 device. If the exposition is only acceptable for at least 0.25 sec. (lid close effect) the device is a class 2 device. In our application class 1M is of great importance. Here, the divergence of the beam rules out the possibility of a damage (see fig. 4). The devices belonging to this class should be safe below 100s of intentional exposure.

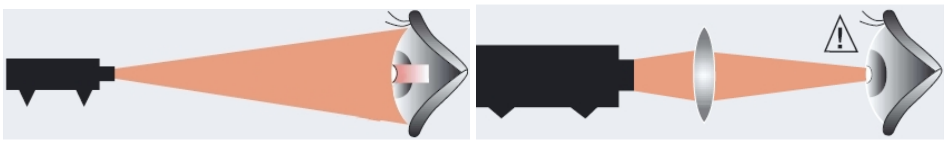

Figure 4: Class 1M classification

Considering the potential damages to the eye several wavelengths can be distinguished:

Blue Light Hazard: Within the wavelength $380nm - 600nm$ *(visible light)* the risk of a photochemical endangerment of retina occurs.

Thermal endangerment of the retina: Might occur at a wavelength of $380nm-1400nm$ *(visible and near infrared light)*.

Cataract and thermal endangerment: Occurs within wavelengths $780nm - 3000nm$ *(infrared light)*.

Endangerment by UV radiation: Concerns the wavelengths $180nm - 400nm$ *(ultraviolet light)*.

The most important properties for our case are the blue light hazard and the thermal endangerment of the retina. For these endangerments the limits are described in [BGI04]. The spectral weighting functions (see fig. 5) directly show the endangerment proportional to the wavelength. They will be used as a weighting function for integrating over the spectrum of the emitted light. Obviously, the blue light is the most dangerous spectrum.

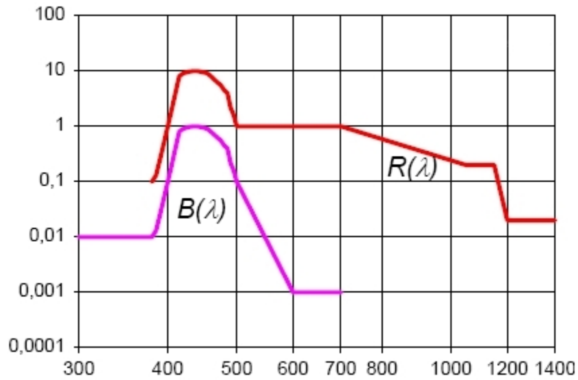

Figure 5: $B(\lambda)$ is a spectral weighting factor belonging to the blue light endangerment. $R(\lambda)$ is a spectral weighting factor belonging to thermal endangerment of the retina.

In order to calculate the danger for the eye, several influence factors must be taken into account:

- Distance to the light source
- Wavelength
- Time of exposure
- Relative size of the source from the eye
- Lenses which might occur between the light source and the eye
- Radiance of the source
- Angle between the viewing ray and the surface normal of the source
- Pulsed or constant radiation

While taking all these factors into account, we come to the conclusion that the most important danger is the blue light hazard for white light LED sources. While reducing the amount of blue light, the exposure limits can be drastically increased by a factor of 10 if necessary. The near infrared spectrum further extends the limit coming from the eye safety constraint.

4 A face scanner prototype

Within the 3d Face project Polygon Technology GmbH developed a face scanner prototype named viSense (see fig. 6). The face scanner fulfills the following requirements:

Resolution in 3D [pixel]	: 640×480
Resolution 2D / color [pixel]	: 1280×960
Field of view [mm]	: 350×260
Capture time [s]	: $0.25sec$
Dimensions [mm]	: $190 \times 130 \times 350$
Weight [g]	: 2300

Figure 6: This is the face scanner prototype viSense developed at Polygon Technology GmbH.

The face scanner is based on the structured light method and consists of a LED projector, a fast gray scale camera for capturing the 3d data and a high resolution color camera for capturing the texture of the face. The maximum exposure time is 3 seconds in a nominal distance of $700mm$ between the user and the device. The exposure limit values for eye safety are $28sec$ of exposure in $10cm$ distance and $230sec$ of exposure in the nominal distance of $70cm$. Hence the device easily fulfills the eye safety constraints.

4.1 Data fitting and interpolation

After obtaining the scan data with the viSense scanner, in a first step jump-edges are detected. A pixel is marked as a jump-edge when adjacent pixel show a significant difference in depth. In practice the underlying surface patch, defined by the direct neighbourhood of a pixel is examined. If the angle between the normal of that patch and the viewing ray exceeds a certain threshold the pixel is marked as jump-edge. Therefore one can also think of jump-edges marking undersampled regions in the range image. That information is also vital for mesh generation, to ensure that areas adjacent in the camera view but not connected in reality remain disconnected in the 3D mesh.

Afterwards small connected regions are discarded by declaring them as invalid. For the detection of small regions connected pixels are counted. A region gets cut if the amount of pixels belonging to it falls below a given limit value. Discarding small connected regions will also affect spikes, since the jump-edges separate the spikes from the surface. That separation leaves them without a sufficiently large neighbourhood (see fig. 7).

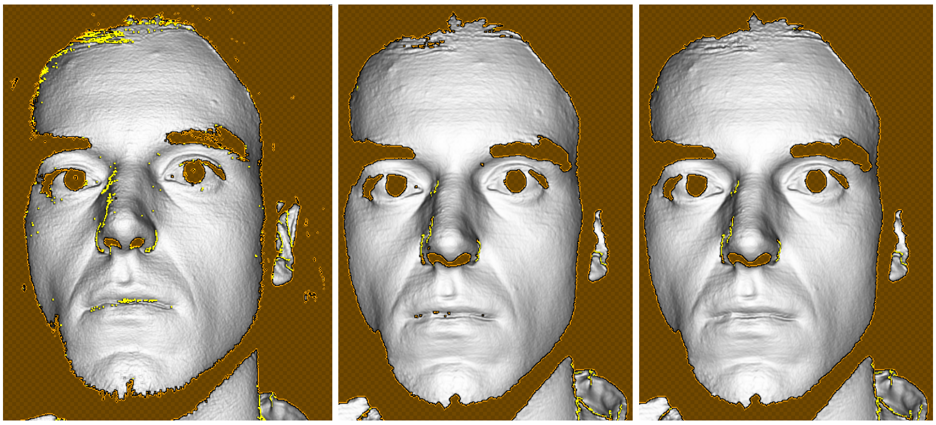

Figure 7: Three different working steps. The left image shows the raw data without pre-processing. Small components exist as noise in the image. After discarding small regions spikes and small components have vanished (middle image). The right image shows the range image after hole-filling.

To determine small holes which can be filled easily and safely the binary morphological closing operator is used. The pixel for which the filling takes place are found by first applying a morphological closing operation on all defined pixels and then subtracting the already defined pixels. The kernel of the closing operation is defined by a circle with a given radius, whereas the anchor point is the center of the circle. Afterwards the hole is iteratively filled pixel by pixel by a linear interpolation from the defined neighbouring pixels, starting and continuing always with the pixels with the maximum number of defined neighbours. Therefore the newly gained neighbourhood information of the preceding step is also used for interpolation.

Fairing and smoothing of the data will not take place in the range image since the amount of smoothing wanted, may differ too much for the different users. The different steps have been integrated with a tried and tested parameter set as fixture in the scanning process.

5 Conclusions

Biometric control requires accurate and fast methods for 3d measurements. Active methods for 3d scanning are an obvious possibility to measure the geometry of the human face. In order to get best results, active methods should use as much light for pattern projection as possible. In addition, movement artifacts could be avoided by choosing a pattern sequence with only few patterns.

As shown in this article, these methods are very well suited for indoor use with more or less controlled lighting conditions of the environment. For outdoor use, the eye safety constraint and the necessary amount of light necessary for abtaining the required contrast for measurement give a sharp limit. Infrared light is an option to extend further the limits for active light measurements.

References

[BGI04] Expositionsgrenzwerte für künstiliche optische Strahlung, Oktober 2004. Berufsgenossenschaftliche Information für Sicherheit und Gesundheit bei der Arbeit.

[DIN01] DIN EN 60825-1:2003-10 (VDE 0837), Sicherheit von Lasereinrichtungen - Teil 1: Klassifizierung von Anlagen, Anforderungen und Benutzer-Richtlinien (IEC 60825-1:1993 + A2:2001), November 2001.

[HHR03] Olaf Hall-Holt and Szymno Rusinkiewicz. Stripe Boundary Codes for Real-Time Structured-Light Range Scanning of Moving Objects. Technical report, Stanford University, 2003. erhältlich unter: http://graphics.stanford.edu/papers/realtimerange/.

[IEC01] IEC 60825-1:1993 + A2:2001-01, Safety of laser products - Part 1: Equipment classification, requirements and user's guide, 2001.

[SPB04] J. Salvi, J. Pages, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*, 37:827–849, April 2004.

[Wio01] Georg Wiora. Optische 3D-Meßtechnik: Präzise Gestaltvermessung mit einem erweiterten Streifenprojektionsverfahren. Ruprechts-Karls-Universität, 2001. Disseration der Naturwissenschaftlich-Mathematischen Gesamtfakultät.

[Wol04] Kai Wolf. Absolute dreidimensionale Vermessung dynamischer Objekte mit Phasenshiftverfahren. Eberhard-Karls-Universität Tübingen, 2004. Disseration der Fakultät für Informations- und Kognitionswissenschaften.

[ZCS02] Li Zhang, Brian Curless, and Steven M. Seitz. Rapid Shape Acquisition Using Color Structured Light and Multi-pass Dynamic Programming. In *The 1st IEEE International Symposium on 3D Data Processing, Visualization, and Transmission*, pages 24–36, June 2002.