Optimized surface extraction from holographic data

Inaugural-Dissertation

zur

Erlangung des Doktorgrades der

Mathematisch-Naturwissenschaftlichen Fakultät

der Heinrich-Heine-Universität Düsseldorf

vorgelegt von

Andrea Thelen

aus Gerolstein

Juni 2006

Aus dem Institut für Lasermedizin
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit Genehmigung der
Mathematisch-Naturwissenschaftlichen Fakultät der
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. P. Hering

Koreferent: Prof. Dr. K. Schierbaum

Tag der mündlichen Prüfung: 26.06.2006

Zusammenfassung

Für die Planung, Simulation und Dokumentation von Eingriffen in der Mund-, Kiefer- und Gesichtschirurgie wird hochaufgelöste, dreidimensionale Information über die Gesichtsoberfläche des Patienten benötigt. Diese kann mittels holographischer Methoden mit extrem kurzen Aufnahmezeiten gewonnen werden, welche die Entstehung von Bewegungsartefakten verhindern. Das Hologramm wird mit einem einzigen Laserpuls von 35 ns Dauer aufgezeichnet und in einem lichtempfindlichen Material gespeichert. Nach einer automatisierten nass-chemischen Entwicklung wird das Hologramm mit Hilfe eines cw-Lasers optisch rekonstruiert. Das während der Rekonstruktion entstehende Lichtfeld, welches eine eins-zu-eins Abbildung des aufgezeichneten Gesichts darstellt, wird schnittbildweise digitalisiert. Digitale Bildverarbeitung überführt eine solche Schnittbildserie in ein dreidimensionales Computermodell. Neben der topometrischen Information wird ebenfalls eine hochaufgelöste, punktgenaue Textur erzeugt, welche zur Texturierung der Computermodelle verwendet wird.

Die vorliegende Arbeit beschäftigt sich mit kontrastbasierten Methoden zur Extraktion der dreidimensionalen Information aus den oben beschriebenen Schnittbildserien, ein Prozess der Oberflächenextraktion genannt wird. Es werden zwölf verschiedene mathematische Operatoren zur Kontrastmessung auf ihre Tauglichkeit für die holographische Gesichtsvermessung hin untersucht, unter anderem der neu entwickelte XSML-Operator, welcher die besten Ergebnisse erzielt. Neben der Wahl des Fokuskriteriums wird auch die Wahl der Größe der lokalen Pixelumgebung, welche für die Kontrastmessung verwendet wird, diskutiert. Wählt man eine feste Umgebungsgröße, stellt dies immer einen Kompromiss zwischen lateraler Auflösung und Robustheit gegenüber Rauschen dar. Ein adaptiver Ansatz, basierend auf einem neu entwickeltem Entscheidungskriterium, passt automatisch die Umgebungsgröße lokal an die Beschaffenheit der Daten an und ermöglicht somit eine optimale laterale Auflösung bei minimaler Benutzerinteraktion.

Die Evaluierung der neu entwickelten Methoden wird qualitativ durch die Begutachtung der Resultate im Hinblick auf die holographische Rekonstruktion aufgezeichneter Gesichtsmodelle und quantitativ mit Hilfe simulierter Datensätze durchgeführt. Da Standardsimulationsmethoden aufgrund der nichtkonstanten Punktabbildungsfunktion der Holographie für diese nicht anwendbar sind, wird eine angepasste Simulationsstrategie entwickelt, welche bemerkenswerte Übereinstimmung mit realen Daten zeigt. Diese hat den zusätzlichen Vorteil, dass eine systematische Untersuchung der beeinflussenden Faktoren ermöglicht wird.

Prinzipiell funktioniert der Prozess der Oberflächenextraktion durch axiales Maximieren der Kontrastwerte. Vier weiterführende Methoden werden präsentiert, welche separat aber auch kombiniert anwendbar sind. Indem eine Gaußkurve an die zu maximierenden Punkte angepasst wird, können Artefakte eliminiert, Rauschen reduziert und vor allem kontinuierliche Oberflächenpunkte gefunden werden, die nicht auf die diskreten Digitalisierungspositionen beschränkt sind. Erstmalig wird eine Auflösung erreicht, die unter der Digitalisierungsschrittweite liegt, welches den gesamten Prozess effektiver macht. Typische Auflösungen für Gesichtesmodelle sind 0,25 mm bei einer Digitalisierungschrittweite von 0,5 mm. Eine andere Methode extrahiert die Oberfläche mit Hilfe eines speziellen Symmetriekriteriums and reduziert damit Artefakte ähnlich wie dies die Gaußanpassung tut, erzeugt jedoch nur diskrete Werte. Sie ist wertvoll, wenn eine schnelle, erste Schätzung der Oberfläche vonnöten ist. Die beiden folgenden Methoden arbeiten mit solch einer ersten Schätzung und verfeinern die Oberfläche in einem zweiten Schritt mittels unterschiedlicher Ansätze. Alle vier Methoden verbessern deutlich die Qualität der extrahierten Oberfläche. Kombiniert kann eine Qualität erreicht werden, welche weitgehend unabhängig vom Rauschanteil in den Daten ist, welches holographische Gesichtsvermessung in einem weiteren Spektrum von Anwendungen ermöglicht.

Die präsentierten Weiterentwicklungen erhöhen enorm die Qualität der erzeugten Computermodelle. Einzelne Haare und Hautfalten, bislang nur sichtbar in der Textur, sind das erste Mal in den Modellen selbst ausgeprägt erkennbar. Es wird gezeigt, dass die erhöhte Robustheit gegen Rauschen hochqualitative Modelle ermöglicht, selbst wenn das Signal-zu-Rausch Verhältnis niedrig ist, wie beispielsweise bei Gesichtsmodellen von Personen mit dunkler Hautfarbe.

Exemplarische Anwendungen holographischer Gesichtsmodelle werden diskutiert. Im medizinischen Bereich bietet die hohe Auflösung die Möglichkeit zur quantitativen Dokumentation und Analyse einer Facelifting Operation. Weiterhin wird die holographische Aufnahme der Schädelkalotte des Neandertalers für wissenschaftliche und Ausstellungszwecke präsentiert. Letztlich werden die Konzepte und erste Schritte für eine forensische Gesichtsrekonstruktion mit Hilfe holographischer Gesichtsmodelle dargelegt. Diese haben den enormen Vorteil, dass sie Weichteilvermessung über den gesamten Gesichtsbereich und nicht nur an ausgewählten anatomischen Punkten ermöglichen. Selbst letztere kann durch hochaufgelöste, holographische Gesichtsmodelle verbessert werden.

Summary

For planning, simulation and documentation of interventions in maxillofacial surgery high resolved soft tissue information of the human face in upright position is needed. This information can be gained by holographic methods, which allow a recording of the whole face in an extremely short time period, so that no movement artifacts occur. The hologram is recorded with a single laser pulse of 35 ns duration and stored in a photosensitive material. After automated wet-chemical processing, the hologram is optically reconstructed with a cw-laser. During the optical reconstruction, a light field, which is a one-to-one three-dimensional representation of the recorded face, emerges at its original position and is digitized into a set of two-dimensional projections. Digital image processing merges these projections into a three-dimensional computer model. Besides the topometric information, a high resolved, pixel precise texture is also extracted from the holographic reconstruction and used for the texturing of the computer models.

This thesis concentrates on contrast based methods for extracting the three-dimensional information from the described set of two-dimensional projections, a process called surface extraction. Twelve different mathematical operators for contrast measurement are evaluated with regard to their suitability for holographic facial measurement, amongst others the newly developed XSML operator, which performs best. Besides the choice of the focus criterion, also the local pixel neighborhood used for contrast measurement is investigated. Choosing a fixed neighborhood size one always has to compromise between achievable lateral resolution and robustness against noise. A multi-scale approach based on a novel decision criterion automatically adapts the neighborhood size locally to the constitution of the data and enables an optimal lateral resolution with minimal user interaction.

All evaluations of new methods are done qualitatively by appraising the results with regard to holographic reconstructions of recorded faces and quantitatively with the aid of simulated data sets. Since standard simulation methods do not work in holography due to the non-constance of the holographic point spread function, an adapted simulation strategy is developed, which shows remarkable accordance with the real data. The simulation holds the additional advantage of making a systematic investigation of influencing factors possible.

In principle, the process of surface extraction is done by axially maximizing the contrast values. Four advanced methods are presented, which can be applied separately or in a combined way. By fitting a Gaussian curve to the points to be maximized, the influence of noise is reduced, artifacts are eliminated and in particular continuous surface values not restricted to the digitization positions are produced. It is the first time that such an approach is reported. With this the resolution of the models is higher than the digitization step size, which makes the whole procedure most effective. A typical resolution for facial models is 0.25 mm gained with a digitization step size of 0.5 mm. A different method extracts the surface with a special symmetry criterion and diminishes artifacts like the Gaussian fit does, but produces discrete values only. It is valuable if a quick, first estimate of the surface is required. The two following methods work with such a first estimate and refine the surface in a second step using different approaches. All four methods highly improve the quality of the extracted surface. Combined they achieve results almost independent from the noise content of the images, which makes holographic facial measurement applicable for a wider range of applications.

The presented advancements improve the quality of the resulting computer models tremendously. Single hairs and wrinkles, earlier only visible in the texture, are for the first time pronounced in the extracted surface itself. It is demonstrated that the improved robustness against noise makes high quality models possible even if the signal-to-noise ratio is low, as it is the case for people with a dark skin color.

Exemplary applications of holographic facial models are discussed. In the medical field, their high resolution enables a quantitative documentation and analysis of face-lifting surgery. Further on the holographic recording of the skull cap of the neanderthal man for scientific and exhibition purposes is presented. Finally the concepts and first steps of forensic facial reconstruction using holographic facial models are explained. They have the tremendous advantage of making soft tissue measurement for the whole facial area possible and not restricting it to anatomical landmarks. Even the soft tissue measurement at anatomical landmarks can be improved through high resolved, holographic facial models.

Contents

Chapter 1

Introduction

> My face is my passport.
>
> *Vladimir Horowitz*

Vladimir Horowitz (1904-1989), an American classical pianist born in Kiev, did most definitely not refer to biometrical data in passports, as he stated this sentence in 1986 to a Soviet official on visit to the USSR. Twenty years later, three-dimensional computer models of the face, which is unique and has an unrepeatable combination of features, are of great interest not only as biometrical characteristics for identification but for a wide range of applications.

He more likely meant that the face is a symbol for the individuality and is a key for judgement by other people in many aspects of life. Considering this it becomes clear, that surgical correction of congenital deformations and reconstructive surgery of tumor or accident patients are very important. Not only an optimal medical result but also an optimal aesthetical outcome is highly desired. This is why precise surgery planning and documentation are indispensable in this field. A necessity for this is highly resolved, three-dimensional information about the face [PCA+02].

Various contact-free, optical measurement systems, for an overview see [CBS00], can in principle be used for this task. But as Bongartz [Bon02], [BGF+03] pointed out, none of them is adequate for facial measurement since either their resolution does not meet the medical requirements or the measuring time is too long, so that the result is affected by motion artifacts caused by changes of the face due to mimic, breathing or heartbeat.

To fulfill the special requirements of three-dimensional measurement of living human faces, a topometry system based on pulsed holography was first implemented successfully by the

Holography and Laser Technology group at caesar (center of advanced european studies and research) in Bonn under the guidance of P. Hering [GBH02]. The main advantage of holography is its capability of capturing a large image volume three-dimensionally with micrometer resolution with only one single, ultra-short laser pulse. In a second step the hologram can be optically reconstructed, digitized and transferred to a three-dimensional computer model. The resulting models are free of motion artifacts and show a realistic appearance because they are recorded in an upright position and contain precisely fitting texture information.

This thesis reports on digital image processing procedures required to extract the topometrical information of the recorded face out of the digitized holographic reconstruction. This problem can be reduced to detecting and maximizing contrast in a set of two-dimensional projections of the real image of the hologram.

This chapter deals first with some aspects about holographic imaging in general, then it discusses similar methods from other fields of image processing, gives an overview over preceding work and ends with an outline of the complete thesis.

1.1 Holographic Imaging

Holography, from the Greek *holos* (whole) and *graphein* (writing), is the science of recording objects in three dimensions. The concept of holography was invented by Gabor in 1948 [Gab48], for which he received the Nobel price in physics in 1971. The discovery was a result of research into improving electron microscopes. He wanted to avoid the poor quality electron lenses by recording an electron wave field on a medium and reconstruct it later optically. Although never really used in electron microscopy, the holographic principle advanced with the invention of the laser in 1960.

The high coherence length of laser light sources made holographic recordings in an off-axis arrangement possible, first demonstrated by Leith and Upatnieks [LU62], [LU63]. The off-axis arrangement shows a lot of advantages in comparison to the less demanding in-line holography proposed by Gabor, like, among others, the separation of the several diffraction orders [MMW73, p. 153]. The development of pulsed lasers allows the recording of living subjects and human portrait holograms [Ans70], [Sie68], [ZS68].

From then on, the holographic principle found its way into uncountable applications, see for example [Gat86], [TW81], and [Tho78] for an overview, covering the whole electromagnetic spectrum down to micro-holography with a wavelength of several nanometers

[ELS+04] or acoustic waves [YFB02].

One important application of holography is interferometric metrology, where two or more wave fields are compared interferometrically, at least one of them must be recorded holographically. This technique makes it possible to measure changes of the phase of the wave field and thus the change of any physical quantity that affects the phase. They range from measurement of vibration modes, over deformation measurement to contour measurement [Kre05, p. 4], [May01].

Apart from these techniques, also the three-dimensional information stored in a single-pulse hologram was utilized for various applications. At Fermilab, a holographic system was used for the measurement of position and thickness of particle traces in a bubble chamber [Ade99], [Har99], since it enables a short recording time combined with a large recording volume. The same principles find nowadays application in systems for cloud particle measurement for example, where holograms are taken out of an airplane with typical speed [FSSS04] or for spray diagnostics [BHK03]. Three-dimensional topometry in holographic particle imaging is done by locating intensity maxima in the reconstructed real image of the hologram. That can be done since the light scattering particles are separated from the non-scattering and thus dark background. Total new problems arise if the three-dimensional information of a continuous object is to be extracted out of the hologram, where an intensity maximization is not necessarily capable of locating the scattering surface. The intensity maximization has to be replaced by a contrast maximization, as was demonstrated for example by [Ste68] and [GMS73].

The procedure of surface extraction stays the same equally if analogue or digital recording material is used. For analogue recording, the hologram is optically reconstructed and then digitized. If a digital recording takes place, the real image is reconstructed numerically. Both procedures lead to a set of two-dimensional projections of the real image as basis for further image processing. For a digital recording, this was recently demonstrated for small objects by Ma et. al. [MWLJ04].

The major advantage of holography, namely the possibility of capturing a large image volume three-dimensionally in an extremely short amount of time, can be used for the three-dimensional measurement of moving objects, where measurement systems working on point by point basis fail. Especially for living human faces, where involuntary movements cannot be avoided, pulsed holography is an ideal tool for three-dimensional measurement [Bon02].

1.2 Shape from focus

The emphasis of this thesis is the digital image processing steps to extract the geometry of the recorded surface from the digitized three-dimensional holographic reconstruction. It can in principle be reduced to the problem of focus detection in a set of images, where each image contains focused as well as unfocused points, a so-called focus series. Recently no work was reported realizing such a procedure in holography. This is why adjacent fields of research, where similar principles are established, were consulted. One of these relevant methods is called *shape-from-focus*.

The acquisition of three-dimensional information out of two-dimensional images has been one of the most important issues in computer vision. These techniques are often referred to as shape-from-x, where x is one of visual cues such as shading, texture, contour, focus, stereo, and motion [YSK99]. Blurring phenomena due to defocusing are among the important cues for depth recovery. *Shape-from-focus*, originally called *depth-from-focus*, is a method where a sequence of images is taken while changing the focus setting in small steps. The focus setting, that locally optimizes image focus, is determined and with it the three-dimensional information about the imaged object.

A lot of work was invested to improve the focus criterion [Jar76], [Kro87], [LG82], [DW88]. Three-dimensional topometry of microscopic objects using this method was demonstrated for firearm bullets and cartridge cases but also for machine tools for inspection purposes [FB97], for material structure analysis in mineralogical research [NNS03] or for micro structures such as a micro-cogwheel [RW96].

Especially in white light microscopy, such an approach, which is based on contrast measurement, is able to compensate for the small depth of focus due to the large magnification by creating both a completely focused image and a depth map containing the three-dimensional information [NN94]. Shape recovery methods not basing on measuring contrast but on intensity variations [AJM98] or wavelet-based methods [LMM95], [KFZS02], [FVdVB+04] are not considered in this thesis.

Many algorithms can be adapted or advanced for the use in holographic facial measurement, as will be demonstrated in the course of this thesis. A crucial advantage of holographic imaging in comparison to the described method of subsequently capturing two-dimensional images with different focus settings is its capability of storing all the three-dimensional information at the same time. This is why also in microscopy holographic recordings are used and the focusing to different layers is done through numerical

reconstruction of the digitally captured holograms. In addition, the holographic images are free of aberrations since no optical lenses are used (see for example [SPI+97],[SJ02], [FGA+05]).

1.3 Preceding work

The first thesis about the holographic facial measurement system was written by Bongartz [Bon02], who covered in his work the hologram recording process, various medical considerations and aspects like the scattering properties of skin, the requirements for the recording material, the wet-chemical processing and the optical reconstruction and digitization with a diffusor screen and a CCD-camera. Surface localization was performed through contrast maximization using the gray value variance as focus criterion. He also introduced structured illumination to enhance the natural contrast of the human skin.

This work was followed by an improvement of the surface localization through an alternative structured illumination, a speckle pattern, by Giel [Gie03]. He also discussed the feasibility of inverse filtering and deconvolution techniques adapted from microscopy for analysis of the digitized real image. Aperture enlargement through mirrors and the concept of hologram perspective, an apparative tool for focus localization, were also part of his work.

Frey [Fre05] switched from a diffusor screen and CCD-camera as instruments for the digitization to a CMOS-scanner and improved the quality of the real image due to image aberration in such a way, that the structured illumination became dispensable and eye-safe recording became possible. Additionally, the lack of structured illumination gave rise to the possibility of extracting the undisturbed intensity information from the real image, which serves as grayscale texture for the resulting three-dimensional computer models. An intensity based surface refinement technique was also part of her work.

Methods alternative to contrast based surface extraction like inverse filtering and deconvolution were not found to be suitable and feasible for hologram tomography. This thesis deals in detail with the possibility of surface extraction through contrast maximization in order to optimize the currently used methods with respect to resolution, robustness against noise and general applicability, and develop new methods.

1.4 Outline of this thesis

Chapter 2 describes briefly the holographic principle, the experimental realization of hologram recording, optical reconstruction and digitization as well as the state of the art surface extractions procedure proposed in the preceding theses, that proved to be applicable and usable in the every day practice of holographic surface measurement.

The following chapter 3 discusses the possibilities of evaluating new surface extraction algorithms and in this context a new method for the simulation of the holographic tomography process is introduced.

Chapter 4 deals with the basics of contrast measurement. Twelve different focus criteria are tested and evaluated, including two newly developed ones. Additionally, chapter 4 demonstrates, how the selection of the neighborhood size for contrast determination is automated through an adaptive algorithm, choosing the smallest possible neighborhood size automatically while maintaining the robustness against noise.

How the actual process of surface extraction is improved is shown in chapter 5. Four newly invented or adapted methods are introduced here, which are applicable separately, or, as is demonstrated in the last section of chapter 5, can be combined to cumulate the improvement.

Chapter 6 presents a method for surface refinement, which is able to eliminate noise while preserving object features. The holographic recording of a test object combined with a quantitative comparison is also included in this chapter.

Various applications from medicine, archaeology and forensic sciences are presented in chapter 7, followed by a conclusion in chapter 8 and appendices on focus criteria and binomial filters.

The algorithms presented in this thesis are implemented as plug-ins to the open source image processing software *ImageJ* created by Wayne Rasband [Ras06]. Despite the usage of basic functionality of ImageJ, all plug-ins were written by myself except if indicated otherwise.

Chapter 2

Holographic Topometry

This chapter deals with the basics of holographic imaging in general (section 2.1) and its experimental realization (section 2.2) with regard to the established setup for facial measurement described in the preceding works [Bon02], [Gie03], [Fre05]. It covers all aspects from the holographic recording setup, the modified recording with structured illumination, the chemical processing, optical reconstruction to the digitization of the real image. These aspects are only covered briefly, since they are described in detail in the theses of Bongartz, Giel and Frey. The concepts of the image processing of the digitized holographic reconstruction and the surface extraction principle are also outlined in this section in a general way. They build the basis for the more sophisticated methods for surface extraction presented in this thesis. The section concludes with the visualization of the gained three-dimensional computer models and a discussion of the spatial resolution of the real image.

2.1 Holographic imaging in theory

In all conventional imaging techniques, such as photography, a picture of a three-dimensional scene is recorded on a light-sensitive surface by a lens. What is recorded is merely the intensity distribution in the original scene. As a result, all information on the relative phases of the light waves from different points or, in other words, information about the relative optical paths to different points of the scene is lost.

The unique characteristic of holography is the idea of recording the complete wave field, that is to say, both the phase and the amplitude of the light waves scattered by an object. Since all recording media respond only to the intensity, it is necessary to convert the phase

information into variations of intensity. This is done by using coherent illumination and using a reference wave incident on the recording medium along with the wave scattered by the object. What is recorded then is the interference pattern produced by these two waves. Accordingly, the processed photographic plate, which is called a hologram, contains information on both phase and amplitude of the object wave. However, this information is in a coded form, and the hologram itself bears no resemblance to the object.

The reason for the success of holography is that the object wave can be generated from the hologram merely by illuminating it once again with the reference wave. To an observer, the reconstructed wave is indistinguishable from the original object wave. The observer sees a three-dimensional image that exhibits all the normal effects of perspective that the object would exhibit if it were still there [Har96, p. 1 ff].

2.1.1 Hologram recording

For a theoretical description, because of the coherence of the used light, the superposition of the object wave O and the reference wave R at the position of the holographic plate has to be considered (see figure 2.1).

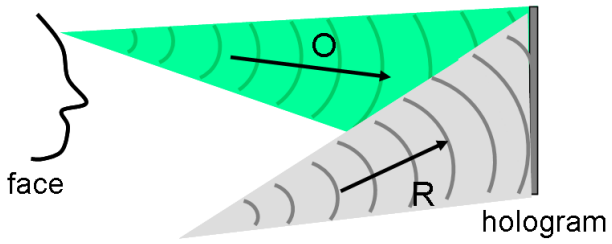

Figure 2.1: Schematic drawing of the hologram recording procedure.

At a point (x, y) of the holographic plate, the electric light field has the following form:

$$E(x, y, t) = O(x, y, t) + R(x, y, t). \tag{2.1}$$

Using light of one frequency and assuming an object which does not move during the recording, the contributing waves can be expressed as $O(x, y, t) = O(x, y) \exp(-i\omega t)$ and $R(x, y, t) = R(x, y) \exp(-i\omega t)$. The intensity in the plane of the hologram simplifies to:

$$I(x, y) = |R(x, y) + O(x, y)| = (R + O)(R + O)^* \tag{2.2}$$

$$= RR^* + OO^* + R^*O + RO^*, \tag{2.3}$$

where * symbolizes the complex conjugation.

The photographic material responds to the total energy per area received during the exposure time and converts it into an optical density. This on the other hand leads to a complex amplitude transmittance τ given by:

$$\tau(x,y) = \frac{E_t(x,y)}{E_e(x,y)}, \tag{2.4}$$

where E_t describes the outgoing light wave immediately behind the holographic plate and E_e is the incident light wave. The complex amplitude transmittance can be expressed in general terms as:

$$\tau(x,y) = T(x,y)\exp(i\phi(x,y)). \tag{2.5}$$

Two cases are distinguished in the recording of a hologram: the amplitude hologram where the phase is constant (ϕ=const.) and the phase hologram, where the amplitude is constant (T=const.) In both cases, the complete information of the wave field is recorded. For the rest of this section, only amplitude holograms are considered.

The real amplitude transmittance T depends on the exposure B (see figure 2.2), which can be expressed, if the intensity does not vary with the time, by the product of the intensity I and the exposure time t_B: $B = It_B$.

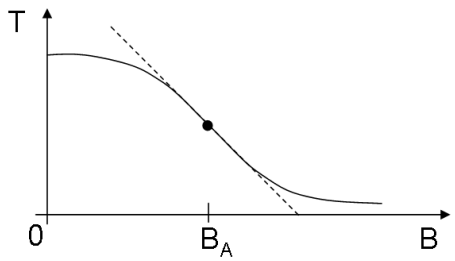

Figure 2.2: Amplitude transmittance of the holographic plate versus exposure B.

The holographic recording is carried out in the linear part of the curve shown in figure 2.2, near the working point B_A. The real transmittance can then be expressed as:

$$T = a - bB = a - bIt_B, \tag{2.6}$$

where the constants a and b depend on the photographic material and on the details of the chemical processing.

After the recording, the holographic plate stores a transmittance distribution [LK03, p. 102ff]

$$T(x,y) = a - bt_B\left(RR^* + OO^* + R^*O + RO^*\right)(x,y). \tag{2.7}$$

2.1.2 Image reconstruction

If illuminating the recorded hologram with a reconstruction wave $C = C(x, y)\exp(-i\omega t)$, the latter is modulated with the complex transmittance $\tau(x, y)$. One gets the complex amplitude of the resulting wave

$$\begin{aligned} A(x,y) &= T(x,y)C(x,y) \\ &= aC - bt_B C\left(RR^* + OO^* + R^*O + RO^*\right) \\ &= \underbrace{\left(a - bt_B\left|R\right|^2\right)C}_{A_0} - \underbrace{bt_B\left|O\right|^2 C}_{\tilde{A}_0} - \underbrace{bt_B R^* OC}_{A_V} - \underbrace{bt_B RO^* C}_{A_R}. \end{aligned} \tag{2.8}$$

Four terms are obtained, which have the following meaning:

A_0 Zeroth order diffraction. The reconstruction wave is multiplied with a constant factor.

$\tilde{A}_0$ Broadened zeroth diffraction order, modulated by $|O|^2$.

A_V Virtual image. This part contains the original object wave O propagating in its initial direction. It only is proportional to the original wave if the multiplied factor is real-valued, otherwise it is a distorted version of O.

A_R Real image. The summand contains the complex conjugate object wave O^*, which can be visualized as the original object wave O travelling back in time. It also is only undistorted in the case of a real-valued factor.

Choosing the reconstruction wave to be a replica of the reference wave ($C = R$), one gets a virtual image term of

$$A_V = bt_B\left|R\right|^2 O, \tag{2.9}$$

which is proportional to the undistorted object wave, while the real image A_R is multiplied with a complex factor and is therefore distorted (see figure 2.3(a)).

Using on the other hand the complex conjugate reference wave as reconstruction wave ($C = R^*$), the real image term has the form

$$A_R = bt_B\left|R\right|^2 O^*, \tag{2.10}$$

which is proportional to the undistorted complex conjugate object wave O^* and a distorted virtual image (see figure 2.3(b)). The real image is a three-dimensional light field, which is

a one-to-one replica of the scattered light distribution of the recorded object. The virtual image exists only in the mind of the observer, who elongates the diffracted wave fronts from the hologram [LK03, p. 104 ff], [EA93, p. 11 ff], [Mei65].

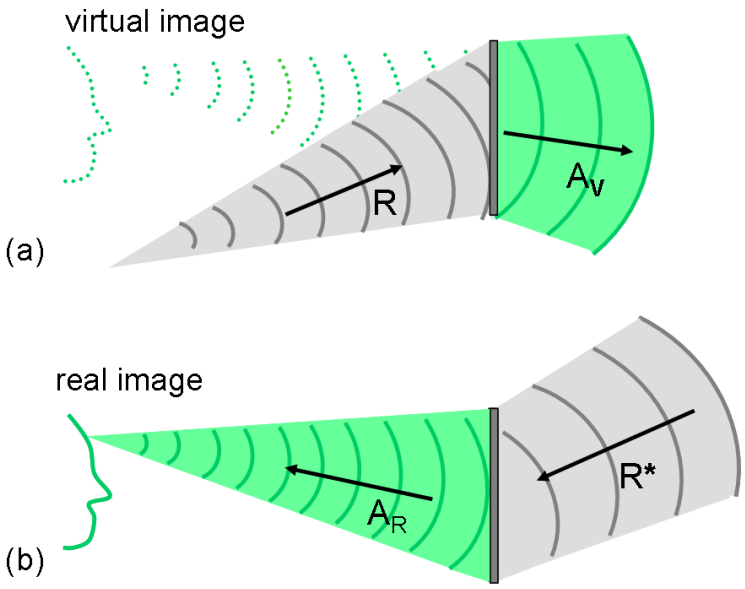

Figure 2.3: Reconstruction with the reference wave R and formation of the virtual image (a). Reconstruction with the complex conjugate reference wave R^* and formation of the real image (b).

2.2 Experimental realization

This sections presents the experimental realization of the principles pointed out in the previous section with regard to portrait holography for three-dimensional facial measurement.

2.2.1 Hologram recording setup

The holographic recording of human faces makes high demands on the used laser system. A short pulse duration is necessary for the recording of living subjects, while at the same time a long coherence length is required. Additionally, the pulse energy has to be high enough to illuminate the whole face in a homogenous way and has on the other hand been adapted to the sensitivity of the recording material.

For the hologram recording a holographic camera GP-2J from Geola was used. It con-

tains a frequency doubled Nd:YLF laser in an oscillator-amplifier arrangement, especially designed for portrait holography.

The master oscillator contains a flash-lamp pumped Nd:YLF crystal and a saturable absorber as passive Q-switch. Single frequency operation is achieved by two additional intra-cavity etalons. The pulses are expanded by a telescope and can be amplified with a flash-lamp pumped Nd:glass amplifier. Two-pass amplification is achieved by inverting the propagation direction of the pulses with a phase-conjugating Brillouin mirror.

The maximum pulse energy after the frequency-doubling in a nonlinear crystal is 2 J at 526.5 nm with a pulse duration of 35 ns. The energy can be adjusted by the voltage of the flash-lamp pumping the glass amplifier.

The pulse is split in three parts, the reference beam and two illumination beams. The reference beam leaves the holographic camera through a pinhole and a concave lens at the back of the camera housing and is led by two ceiling mirrors to the holographic plate over a length of 8.8 meters (see figure 2.4).

The two illumination beams are expanded by concave lenses and led through diffusor plates for a homogeneous illumination of the face (see figure 2.4). The diffusor plates act as extended light sources and make the hologram recording eye-safe. The person is sitting in front of the camera with a distance of approximately 60 cm to the holographic plate. The fraction of the scattered light from the face, which arrives at the holographic plate, is superimposed with the reference beam and forms an interference pattern, which is stored in the photosensitive material.

To record a hologram, a standard size holographic plate (30 $\times$ 40 cm) is put in the holder (see figure 2.4) and the person takes place in front of the camera. The positioning is uncritical, among other things due to the absence of lenses between the person and the holographic plate. After the triggering of the laser and the ultrashort pulse, all the three-dimensional information about the face is stored and the proband can leave immediately [Bon02], [Gie03], [Fre05].

2.2.2 Recording with structured illumination

Aim of the hologram tomography process is the transfer of the three-dimensional information stored in the hologram into a three-dimensional computer model. In order to do so, the hologram is optically reconstructed and digitized into a set of two-dimensional projections, so-called tomograms. The whole process will be described in detail later in

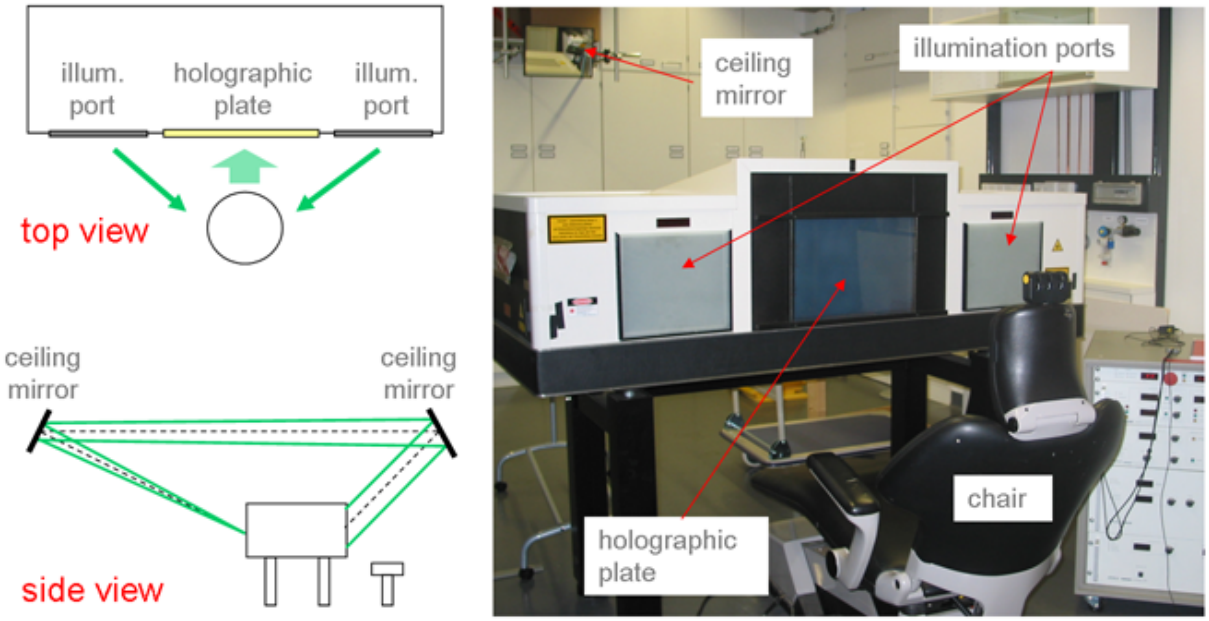

Figure 2.4: Schematic top view, side view and photograph of the holographic camera.

this section. The main challenge in this regard is to distinguish between focused and unfocused sections in the tomograms.

Defocusing is equivalent to blurring. Small object features are visible in focused regions, but invisible in unfocused ones. This criterion is used for the extraction of the three-dimensional information. It works the better if one does not have to rely on the intrinsic feature content of the object only but provides it with a synthetic structure. One way to do this is to spatially modulate the object beam.

This is done by removing the diffusor plates and placing semitransparent slides into the illumination beams. Stripe pattern [Bon02] or statistical speckle pattern [Gie03] can be used. They have the drawback that the recording is not eye-safe anymore. As [Fre05] demonstrated, structured illumination is not necessary for the recording of human faces, since the intrinsic contrast of the skin is sufficient for focus detection.

For artificial objects on the other hand, where eye-safeness is not an issue, structured illumination has to be used if the object does not inherit a structured texture. Figure 2.5 shows two tomograms at the same axial position of a plastic head model (see section 6.2 for a photograph of the model) recorded with homogeneous illumination (a) and with structured illumination (b). Since the object does not have a structured surface, it is impossible to distinguish between focused and defocused image regions (a). Structured illumination provides the object with a synthetic texture and makes therefore focus detection possible (b).

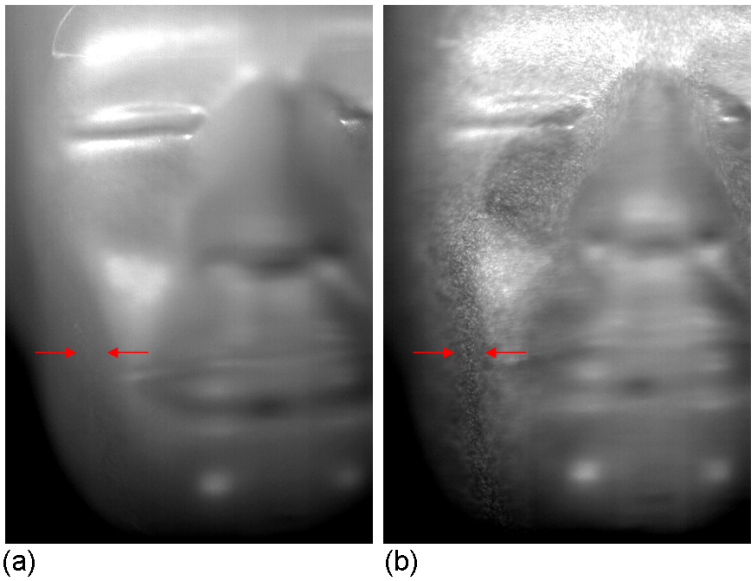

Figure 2.5: Tomogram of a plastic head model recorded with homogeneous illumination (a) and structured illumination (b). The red arrows mark the focused contour. It is nearly invisible if a homogenous object illumination is used (a). Structured illumination makes a distinction between focused and defocused image regions possible (b).

2.2.3 Chemical processing of the hologram

The holographic recording material is a highly sensitive photographic film (silver-halide salt crystals) embedded in gelatine coated onto a glass plate or flexible film. It has a resolution of 3000 lines per mm, which lies above the required resolution of 2300 for typical portrait hologram recording parameters [Bon02].

Depending on the chosen chemical development strategy, phase or amplitude holograms can be generated. Amplitude holograms have a lower diffraction efficiency, since the modulation is realized through absorption in the dark regions. Phase holograms on the other hand work with a phase modulation and are therefore more efficient in respect to the diffraction of light. In the course of this thesis only phase holograms were used.

An automated chemical development can be done using a modified x-ray development machine. Since automated processing in only possible for flexible film, it has to be fixed on a glass plate using vacuum during the recording [Lad04], [LFT+04].

2.2.4 Optical hologram reconstruction

As already mentioned in the previous section, the success of holography can be led back to the fact that the coded three-dimensional information about the recorded object can be decoded through illumination with the reference or complex conjugate reference beam. Since the virtual image cannot be used for topometry purposes, the complex conjugate reference beam has to be chosen and the real image has to be created.

The reference beam used during the recording was a diverging Gaussian beam impinging to the hologram from above. A combination of a beam telescope, a convex lens and a parabolic mirror create the complex conjugate of the reference wave. If it impinges at the hologram from the opposite direction as during the recording, which means from below and from the backside, the real image is created. The used laser is a frequency-doubled Nd:YAG laser with cw single longitudinal mode operation at a wavelength of 532 nm [Gie03].

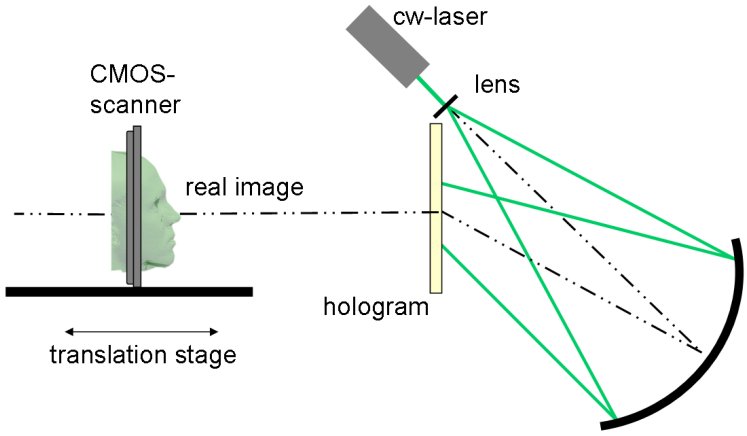

Figure 2.6: Optical reconstruction with the complex conjugate reference beam. The real image appears as a three-dimensional light field at the position of the recorded object.

The wavelength mismatch between the recording and reconstruction wavelength results in a slightly different scaling factor in the lateral and axial direction [Ost88, p. 53], which can be compensated numerically [Bon02].

Details about the creation of the exact complex conjugate reference beam, which is necessary in order to avoid image aberrations, can be found in [Fre05].

2.2.5 Digitization of the real image

The real image, generated through the optical reconstruction, is a one-to-one image of the recorded object and can be used for topography purposes.

A digitization of the real image is the first step towards creating a three-dimensional computer model. It is realized through the recording of a set of two-dimensional projections at different axial positions. Since a two-dimensional image representing a slice through a three-dimensional volume is in medical terms called tomogram, Bongartz named the whole procedure hologram tomography [Bon02].

The projection to 2D can be realized with a diffusor screen and a CCD-camera [Bon02], [Gie03], or with a CMOS document scanner [Fre05], which delivers a much better signal-to-noise ratio. The scanner (CanoScan LiDE 60) is mounted on a linear translation stage (see figure 2.6). A typical data set contains 256 or 512 tomograms, which lead to an inter-slice distance from 0.3 mm to 0.8 mm for the digitization of portrait holograms.

Figure 2.7 shows an exemplary set of tomograms through the real image. In the first image everything is unfocused. The second tomogram shows a sharp image of the nose, the third of the eyes, the fourth of temples and the last of the ears. The three-dimensional information about the face stored in the hologram is now transferred to the set of tomograms.

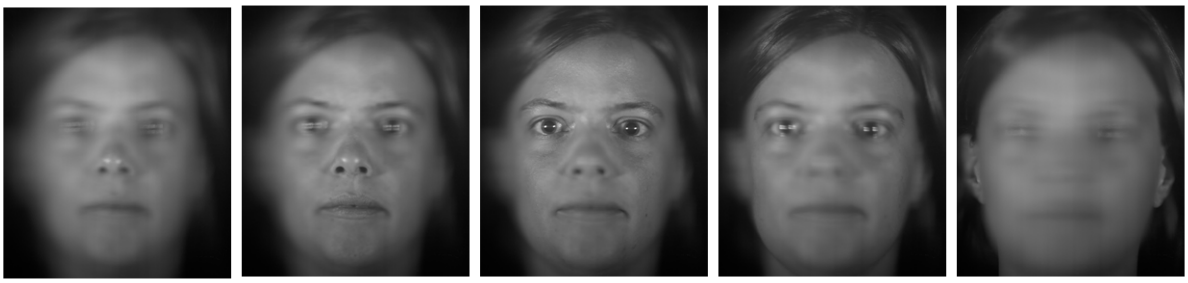

Figure 2.7: Exemplary set of tomograms of the real image. Several regions of the face are imaged sharply, while the other regions are unfocused.

2.2.6 Image processing

The topometry of the recorded object is imitated by the focused image points in the three-dimensional light field, which is the real image. As in every imaging device, the light is directed from the hologram to the focus point in a converging way and diverges

again after the focus. This is why besides the focused points, also unfocused points are visible in each tomogram. A distinction between these two has to be established.

Focused image regions show a higher contrast than unfocused ones. This is why a mathematical operator measuring image contrast, a further on called focus operator, is the optimal tool to realize this task. Since one image point itself cannot have contrast, local pixel neighborhoods of each point are considered for contrast determination (see figure 2.8).

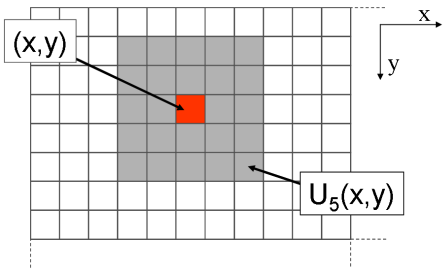

Figure 2.8: Sketch of a local pixel neighborhood $U_5(x, y)$ of the point (x, y) with a size of 5 pixels.

A detailed discussion about focus operators can be found in section 4.1. One example is the gray value variance, used by Bongartz and Giel [Bon02], [Gie03]. The numerical value $F_{xy}(z)$ delivered by the focus operator for each point (x, y) and each tomogram with the position z is further on called focus value. Figure 2.9 illustrates an exemplary tomogram (a) and the corresponding focus values (b). Focused regions produce high focus values while unfocused ones produce low values.

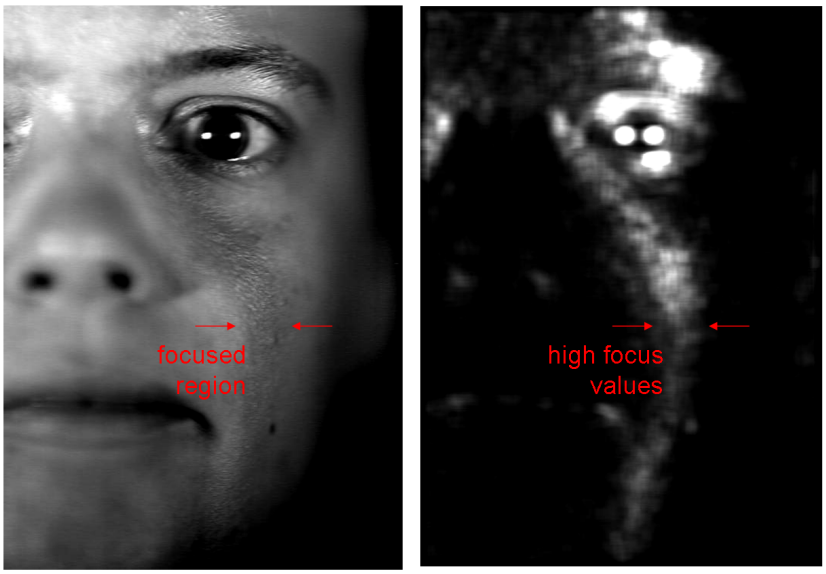

Figure 2.9: Tomogram with a focused contour of the face and unfocused background (a) and the visualization of the calculated contrast values (b). White corresponds to high contrast, black to low.

2.2.7 Surface extraction

In order to extract the surface information of the recorded object, a maximization of the focus values has to be performed. Comparing focus values along the z-axis delivers a z-value for each lateral point (see figure 2.10). Undercuts in the surface are unaccounted for.

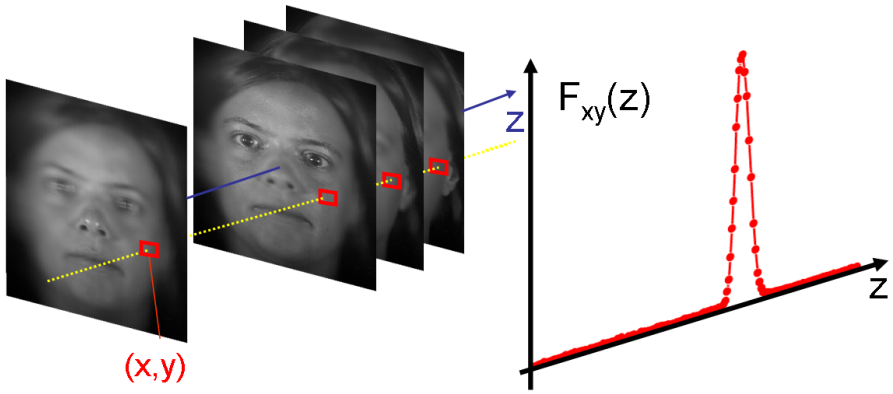

Figure 2.10: For each lateral point (x, y), the focus values $F_{xy}(z)$ are maximized, which leads to the desired z-coordinate.

The extracted surface can be visualized through the so-called height map, which displays the z-values as coded gray values. A gray value of 1 symbolizes the z-value of the first tomogram and so on (see figure 2.11). The height map can be used for postprocessing of the surface. Outlying points can be removed through the usage of a median filter[1], a mean filter[2] can be used for smoothing the surface [Bon02].

In addition to the topometrical data, the procedure can be used to generate an image with an unlimited depth of sharpness by compounding the detected sharp contours from each tomogram into one picture (see figure 2.11), the so-called texture image. It can be used for texturing the resulting computer models. The pixel-to-pixel correspondence, which is inherent to the described process, makes any further registration between the model and the texture expendable. For a more realistic appearance, the grayscale image is dyed with one color. The result can be seen in figure 2.14.

[1]The gray values of a certain window are listed in an ascending order and the medium value is chosen as filtered value.

[2]Each gray value is replaced by the mean value of its neighbors.

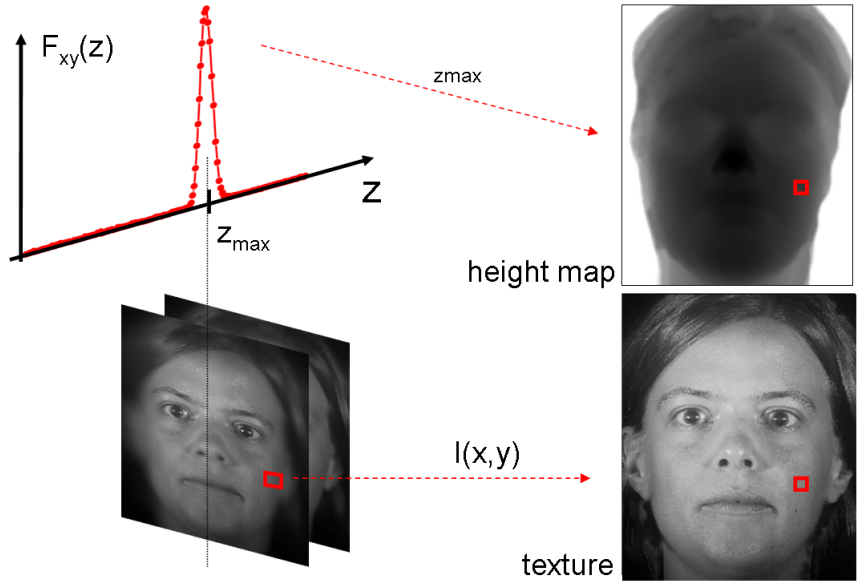

Figure 2.11: Simultaneous extraction of the topometrical information of the face and the corresponding grayscale texture.

At several occasions in this thesis profiles like the one displayed in 2.10 are used for demonstration purposes. They are further on referred to as focus profiles, because they display the focus value $F_{xy}(z)$ against the axial z-value. On other occasions also height profiles are displayed, which are cuts through the height map, where the axial coordinate is displayed against the lateral coordinate. Figure 2.12 shows an exemplary focus profile (a) and a height profile (b) in order to clarify the distinction between those two.

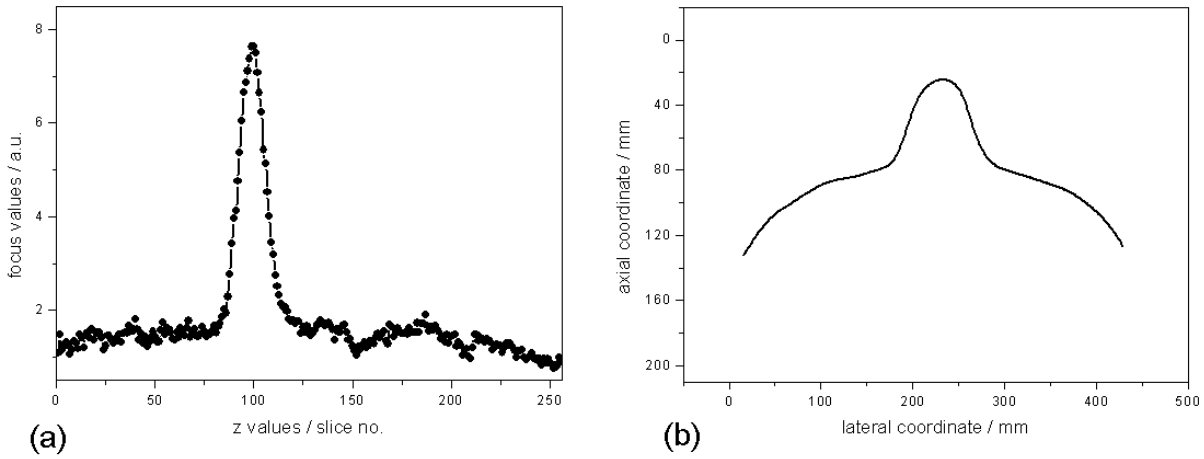

Figure 2.12: Exemplary focus profile (a) and height profile (b).

2.2.8 Surface visualization

To visualize the gained three-dimensional surface interactively, the Virtual Reality Modelling Language (VRML) is used, which is a language for describing interactively three-dimensional objects. VRML is intended to be a universal interchange format for integrated three-dimensional graphics and multimedia and is used in a variety of application areas such as engineering and scientific visualization [VRM97]. It can be displayed with a stand-alone VRML-Viewer or a VRML-Plugin for an internet browser, for example with the *blaxxun Contact* Plugin[3].

[3]http://www.blaxxun.com

A VRML-file is a plain text file starting with `#VRML V2.0 utf8` as indication of the file format. After declarations about illumination, background color and so on, the actual topometry information is given as set of three-dimensional points (x, y, z). In order to create a polygonal surface, the indexes of points sets building one polygon are stated. Such a model, where only the polygon boundaries are displayed, can be seen in figure 2.13(a)[4].

The surface extraction procedure delivers besides the topometrical data also precisely fitting texture information. In order to add the texture to the polygon model correctly, the node positions are projected into the texture image coordinate system. The latter comprises the two coordinates (u, v) both running form zero to one. These projected node coordinates are also written into the VRML-file so that the texture is positioned correctly in the resulting model. Figure 2.14 shows a screenshot of a textured VRML model gained with hologram tomography.

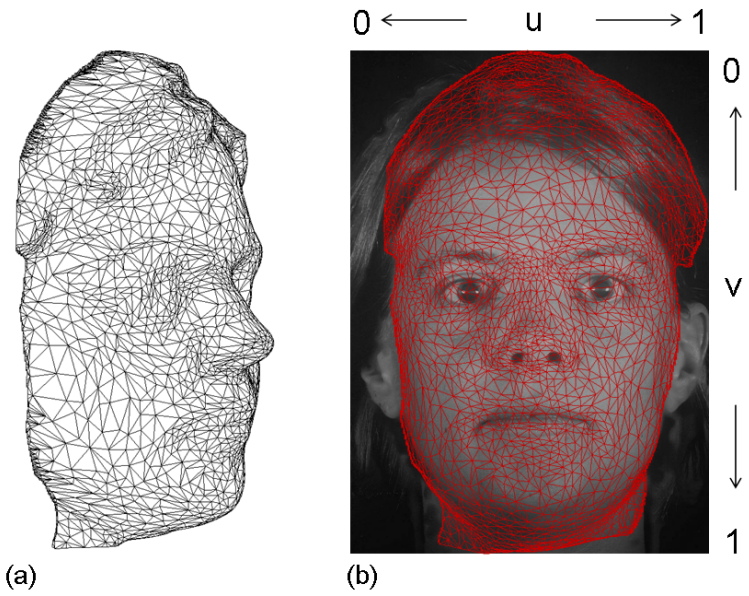

Figure 2.13: Polygonal model (a) and polygonal structure projected into the texture image to guaranty a precisely fitting texture (b).

[4]For visualization purposes, the number of polygons is reduced tremendously. A facial model with the standard digitization parameters (0.17 mm pixel size) has about two million polygons.

Figure 2.14: Resulting three-dimensional computer model with precisely fitting texture.

2.2.9 Spatial resolution of the real image

The spatial resolution of the real image can be estimated by regarding the dimensions of the real image of one single point-like object. Illuminating such an object with coherent light, it acts as a center of a spherical wave. With a planar reference beam, the hologram is a Fresnel zone plate [Rog50]. The optical reconstruction of such a hologram with a Gaussian beam leads to a real image, which can be also described as a Gaussian beam with its focus at the position of the original scattering object point [Bon02]. [Gie03] demonstrated this fact experimentally. The three-dimensional light distribution of the image of a point is called point spread function (PSF). If we consider a point on the optical axis with the coordinates $(x_0 = 0, y_0 = 0, z_0)$ at the distance z_0 from the hologram (see figure 2.15), the PSF can be described by the following equation:

$$h(x, y, z) = \left(\frac{w_0}{w(z - z_0)}\right)^2 \exp\left(-2\frac{x^2 + y^2}{w(z - z_0)^2}\right) \tag{2.11}$$

$$\text{with } w(z - z_0) = w_0\sqrt{1 + \left(\frac{z - z_0}{z_R}\right)^2}, \tag{2.12}$$

where w_0 is the beam waist and z_R the Rayleigh length.

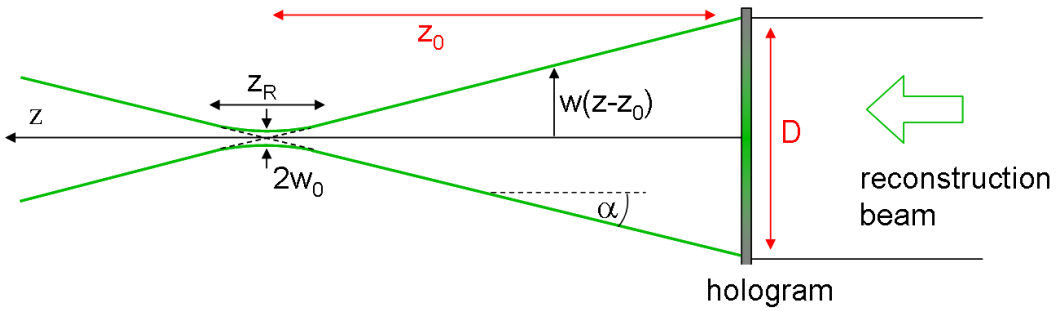

Figure 2.15: The optical reconstruction of a Fresnel zone-plate leads to a Gaussian beam profile with a Rayleigh length z_R and a beam waist of $2w_0$.

Both beam parameters depend on the laser wave length λ, the diameter of the hologram D and the distance of the scattering point to the hologram z_0 according to the following equations [Sie86, p. 664 ff]:

$$w_0 \approx \frac{2 \cdot z_0 \cdot \lambda}{\pi \cdot D} \tag{2.13}$$

$$z_R \approx \frac{4 \cdot \lambda \cdot z_0^2}{\pi \cdot D^2} \tag{2.14}$$

Substituting the wave length with λ=526.5 nm, the diameter D with a typical value of 20 cm and the distance z_0 with 60 cm, w_0 and z_R can be regarded as the lateral and axial

resolution of the real image [Bon02]. The following values are calculated:

$$w_0 \approx 1\mu\text{m} \tag{2.15}$$

$$z_R \approx 6\mu\text{m}. \tag{2.16}$$

These two quantities, $2w_0$ and z_R can be regarded as the lateral and axial resolution of the real image.

If z is much larger than the Rayleigh length z_R, equation 2.12 can be simplified to:

$$w(z - z_0) \approx \frac{w_0}{z_R}(z - z_0) \tag{2.17}$$

which, referring to equation 2.13 and 2.14, can be transformed into:

$$w(z - z_0) = \frac{D}{2 \cdot z_0}(z - z_0) \tag{2.18}$$

The opening angle α of the PSF can be calculated with the aid of the last equation:

$$\tan\alpha = \frac{w_0(z - z_0)}{(z - z_0)} = \frac{D}{2 \cdot z_0}. \tag{2.19}$$

This correspondence will be used later in simulation of the real image projections (see chapter 3).

2.3 Summary

This chapter gave a short introduction to holographic imaging. The concept of hologram recording as storage of the interference pattern of an object and a reference wave was introduced. The possibility of decoding the stored three-dimensional information optically by illuminating the hologram with the complex conjugate reference beam was also discussed.

This general overview was followed by the characterization of the experimental setup, including hologram recording with and without structured illumination, chemical processing of the holograms, optical reconstruction and digitization of the real image. The basic procedure of image analysis and surface extraction was presented. They build the basis for the newly developed and enhanced algorithms for focus detection and surface extraction presented in the course of this thesis. Additionally a short introduction to the visualization of the resulting three-dimensional computer models was given and the spatial resolution of the real image was discussed.

Chapter 3

Simulated hologram tomography

The evaluation of the extracted surface is a critical step in the optimization of different procedures. Several evaluation methods are discussed in the first section of this chapter. In this regard also the simulation of digitized, two-dimensional projections of the real image are discussed. The simulation can act as ground truth for the evaluation. In order to realize such a simulation, a closer examination of the holographic point spread function is necessary, which is also included in this chapter.

3.1 Evaluation of the extracted surface

The evaluation of three-dimensional measurement systems is a widely discussed topic. There are standards for example for measurement systems based on triangulation [VR02] using spheres and planes as reference objects.

In general, the evaluation of the measured three-dimensional data set can be done by comparing the measured surface with its known reference geometry. In order to do so, the two data sets have first to be brought into the same coordinate system by a matching procedure, for example using the *Iterative Closest Point Approach* (ICP) algorithm developed by Besl and McKay [BM92] (see section 6.2 and chapter 7). If the matching is performed, the mean absolute deviation between the measured and the reference surface can be calculated. One always has to take into account that the result of the matching process depends on the used algorithm. It can be negatively influenced by outliers and is therefore critical for the estimation of the error.

Bongartz used an object with plane regions of different slope as a test object for the

evaluation of the holographic topometry system. He made a linear regression along several lines along the surface of the test object and obtained an error in the order of magnitude of the inter-tomogram distance of approximately 0.11 mm [Bon02].

Frey also used a geometrically simple test object, namely two surfaces forming an angle of 90°. She fitted the two surfaces with two planes, estimated additionally the mean deviation at the edge and calculated also an result of approx. the digitization step-size of 0.17 mm [Fre05].

The usage of geometrically simple test objects has the advantage that the reference geometry can be modelled and that the object surface can be accounted error free. It has on the other hand the drawback that, due to the missing intrinsic texture, structured illumination has to be used during the holographic recording. The structured illumination enhances the contrast and simplifies the surface extraction procedure tremendously. Due to this fact it appears critical to transfer the achieved resolution to facial models, where no additional structure is present. Also the complex geometry of the face and its particularities are not accounted for. If using plane reference objects, one risks to produce unrealistic deviations if the surface extraction procedure inherently smoothes the acquired data, like it is the case for contrast measurement with large neighborhoods. For plane objects, surface extraction with inherent smoothing always improves the result (see figure 3.1 (a) and (b)). If on the other hand a more complex geometry is considered, inherent smoothing may downgrade the results due to the abrade of structures (see figure ref 3.1 (c) and (d)). This is why for evaluation purposes a geometry as realistic as possible should be considered.

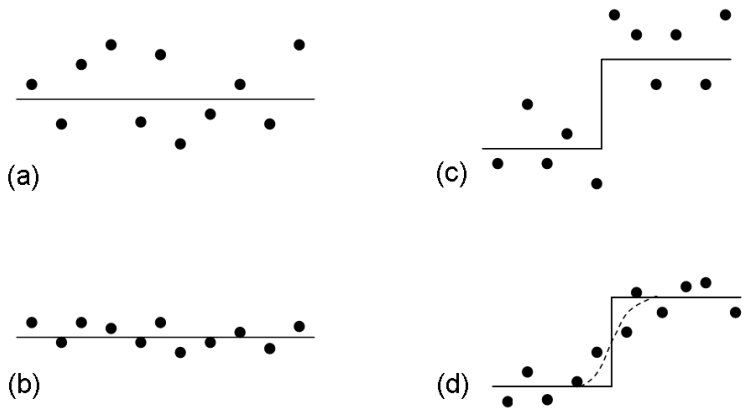

Figure 3.1: If the reference object is plane (a), surface extraction with inherent smoothing always improves the results (b). If a distinctive structure is present (c), inherent smoothing is able to downgrade the results due to the abrasion of structures (d).

A more realistic test object, a face model produced by an automated milling machine, was used by Giel [Gie03]. The nominal resolution of the machine was 10 μm, so he neglected manufacturing errors. He performed the matching process, which correctly should be done considering six degrees of freedom, only for two translational degrees of freedom after pre-aligning the models by hand. An error nearly twice as high as the inter-tomogram distance of 0.5 mm was the result, which can be led back to the more realistic test object or to matching errors. He also used structured illumination.

Ideal would be a real face as test object. No structured illumination would be necessary, since the intrinsic contrast of the skin would be sufficient for focus detection and the geometry would correspond to the actual application. However, in order to produce a reference surface, the face has to be measured with an alternative measuring system, which, inherently, also produces measurement errors. Additionally, the matching process is more complicated due to inevitable changes of the face between the two measurement procedures. This makes a real face as test object unusable.

A similar evaluation was done in this thesis using a plastic face model (see section 6.2) instead of a real face, which at least circumvents the involuntary changes between the two measurements. The missing intrinsic texture made structured illumination necessary, which changes the characteristics of the surface extraction procedure.

Additionally, an alternative to the described evaluation methods was accomplished in this thesis. Instead of using real holographic data for surface extraction, simulated tomograms are used. If the simulation is based upon a height map and texture primarily gained with real holographic data, they can serve as exact reference for the comparison. The simulated data set then has a face-like geometry as well as a true to original texture, which makes structured illumination dispensable. It has the additional advantages that no matching is necessary since the reference and the extracted surface are naturally in the same coordinate system. Disturbing effects like noise can selectively be added, which makes a systematic evaluation of their influence possible and bears the possibility of a deeper understanding of the influencing factors.

Besides these crucial advantages it has to be mentioned that for example image aberrations or systematic digitization artifacts (see [Fre05]) are not taken into account so that the results can only be an approximation to the realistic resolution. However, since the simulated and the real data behave under the procedure of surface extraction qualitatively similar, as is demonstrated in section 3.5, the simulation is a very useful tool for a quantitative comparison of different surface extraction algorithms and procedures. All

results based on the simulated data are fortified with qualitative results from real data.

The simulation of an imaging process always bases on the imaging characteristics of the used imaging system, which are described by the point spread function. This is why the holographic point spread function is discussed in the following section, before the actual simulation is described.

3.2 Holographic point spread function

In classical imaging theory the discussions on imaging performances are usually limited to thin objects, because two-dimensional images of the object are of main interest. For the imaging of three-dimensional objects, the three-dimensional light distribution near the focus is of importance. The characteristics of an imaging system can be described by its point spread function (PSF), which is defined as the intensity distribution of the image of a single, point-like object. For three-dimensional objects it is a function of three coordinates [Gu00, p. 37].

The description of imaging systems by their PSF is commonly simplified by two assumptions. First it is assumed that the shape of the PSF is the same for all object points, which means that the PSF is invariant over the complete imaging range (translational invariance). The second assumption is, that the image can be approximated as an incoherent superposition of intensities of the images of all object points [Mah98, p. 22].

The intensity distribution $I(x, y, z)$ of the three-dimensional, extended image can then be related to the object function $O(x, y, z)$ by the convolution integral, which is defined as

$$I(x, y, z) = \int\int\int h(x - \xi, y - \eta, z - \zeta) \cdot O(\xi, \eta, \zeta) d\xi d\eta d\zeta, \tag{3.1}$$

where $h(x, y, z)$ is the PSF of the imaging system [Yar04, p. 29].

Both assumptions are only partly true for holography: The scattered radiation of the object surface is a coherent superposition of the intensities from the individual object points. In the case of a human face however, the object surface scatters with a fixed but quasi random phase [Gie03]. The translational invariance is the more critical assumption. The fact that the direction as well as the opening angle of the PSF vary for different object points is neglected by [Gie03], but is of great importance for the characteristics of the holographic image and all the following steps, as will be demonstrated in the course of this thesis. According to this variation, the convolution integral (equation 3.1) with an invariant PSF does not describe the correspondence between the image

and the object correctly for holographic imaging, so that alternative models are to be developed. The following paragraphs describe a new, heuristic simulation method. The good qualitative correspondence with the real holographic tomograms and its similar behavior in the process of surface extraction justifies its use for evaluation purposes.

The transformation from an unrealistic, constant PSF to a realistic PSF is done in two steps. First the varying opening angles are taken into account, than the different directions of the symmetry axis of the PSF. As said above, the holographic point spread function is characterized through the image of a point-like object. In section 2.2.9 it was shown that the holographic image of a point-like object can be characterized by a Gaussian beam profile. The Rayleigh length for typical facial holograms is in the order of 6 μm, which is very small compared to the typical distance of 60 cm between the hologram and the face. From now on the Gaussian beam waist is neglected and the PSF is regarded as a cone with a Gaussian intensity distribution.

Figure 3.2 shows schematically two on-axis focus points with the z-coordinates z_{01} and z_{02}. The larger the distance from the focus point to the hologram, the smaller is the opening angle α of the PSF. It can be calculated using equation 2.19 from section 2.2.9:

$$\tan \alpha = \frac{D}{2 \cdot z_0}, \tag{3.2}$$

where z_0 is the distance to the hologram and D its diameter.

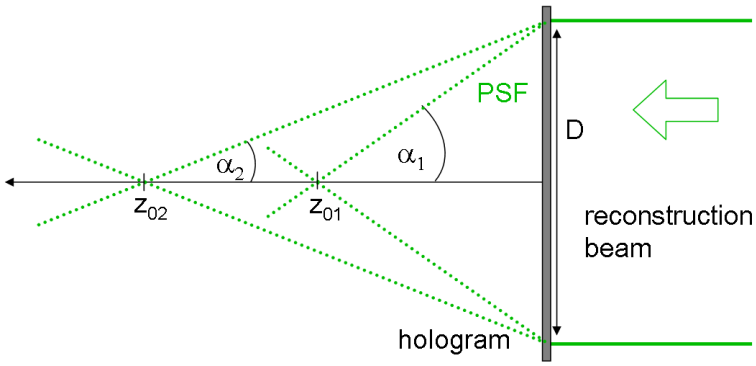

Figure 3.2: The opening angle α_i of the point spread function depends on the distance z_{0i} from the focus point to the hologram.

For off-axis points, the symmetry axis of the PSF (PSF-axis) is not parallel to the z-axis anymore (see figure 3.3).

In that case, the PSF is not only characterized by the opening angle α but additionally by the direction of the PSF-axis. If the center of the coordinate system is transferred to the illumination center of the hologram, the focus coordinate (x, y, z_0) can be expressed

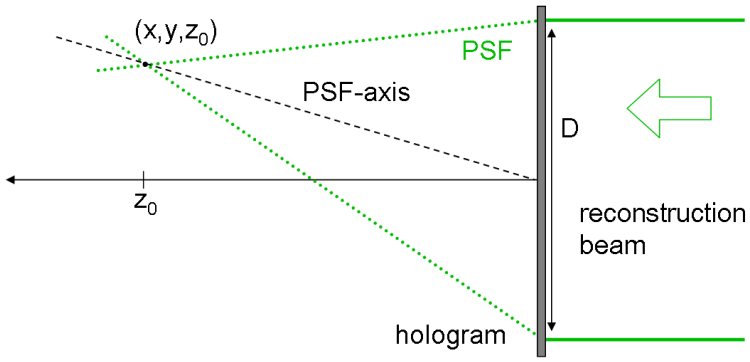

Figure 3.3: For off-axis points, the PSF-axis cannot be assumed to be parallel to the z-Axis.

in spherical coordinates:

$$\rho = \sqrt{x^2 + y^2 + z_0^2} \tag{3.3}$$

$$\theta = \arccos \frac{z_0}{\rho} \tag{3.4}$$

$$\phi = \arctan \frac{y}{x} \tag{3.5}$$

The focus coordinate (x, y, z_0) or alternatively its transformation into spherical coordinates (ρ, θ, ϕ) determines the direction of the PSF-axis. With this a realistic simulation can be performed.

3.3 Simulation with parallel PSF-directions

Exploring how to simulate tomograms of the real image at arbitrary axial positions is the aim of this section. Starting point is a height map and a texture image extracted from a real holographic data set. The height map consists of continuous values (see section 5.2) and is slightly smoothed with a mean filter (filter radius 3).

For a given position z where a tomogram is to be simulated, certain surface points might be in focus. The intensity values of the tomogram and the texture image are the same for these points. During the focusing and defocusing procedure, the focus intensity is spread into a Gaussian profile according to the PSF. For points in the tomogram where no surface point is in focus, the spread intensities of all surface points add up and the sum has to be transferred into the simulated tomogram.

In theory, the contribution of each surface point can be calculated using the characterization of the PSF given in the previous section, but this is a very time consuming calculation. In general it can be said that the larger the lateral distance of a surface point to the exam-

ined position (x, y), the smaller is its contribution. The largest contribution comes from the surface point with the same lateral coordinate (x, y) and its immediate neighbors. The following procedure is performed as an approximation to the correct simulation. The PSF-radius of the surface point with the same lateral coordinate as the considered point (x, y) is calculated correctly. It is now presumed that the intensities from all other surface point are spread with the same PSF-radius. For immediate neighbors this is true, since their z-positions do not vary much. For points far away the error is small since they contribute only a little anyway. This concept has the huge advantage that a set of blurred texture images can be created in advance and can be used for various simulation positions. The blurring is done using a Gaussian filter. For a given z-position and a fixed lateral point (x, y), the PSF-radius of the surface point at (x, y) is calculated and the intensity is set to the value from the same position in the texture image blurred with the calculated radius. Figure 3.4 shows exemplarily a section of the texture image (a) and a version blurred with a radius of 15 pixels (b) and 30 pixels (c). After the set of blurred images is created, the simulation consists only of radius calculation and the transfer of the intensities to the simulated tomogram.

Figure 3.4: Section of the texture file (a), the same section blurred with a Gaussian filter of an exemplary radius of 15 pixels (b) and one with a radius of 30 pixels (c)

The calculation of the PSF-radius is first done for on-axis points, which have a PSF-direction parallel to the z-axis. Then the procedure is repeated for all the other points considering them also to have also PSF-directions parallel to the z-axis. This is an unrealistic assumption and will produce results which are unrealistic, but very helpful for a

systematic analysis of the influencing factors during surface extraction.

Figure 3.5 shows the PSF of an on-axis point. For on-axis points it is only characterized by its opening angle α. The radius of the PSF at the position z can be calculated using the z-coordinate z_0 of the surface point:

$$r = |z - z_0| \cdot \tan \alpha. \tag{3.6}$$

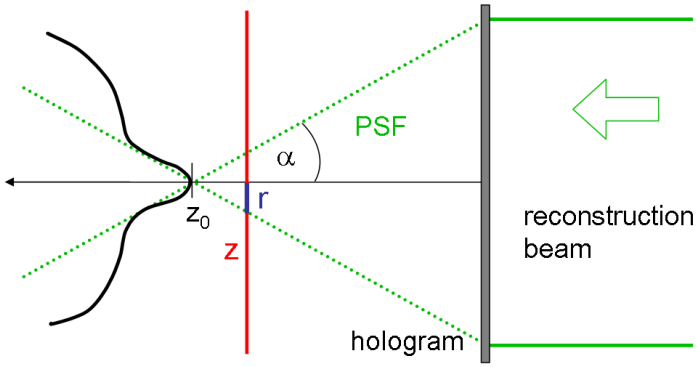

Figure 3.5: Determination of the radius r of the PSF at the position z where the tomogram is to be simulated.

The same procedure is now repeated for every surface point. An exemplary tomogram can be seen in figure 3.7(a).

3.4 Individually calculated PSF-directions

In this section, the more realistic case of PSF-directions not parallel to the z-axis are considered. Apart from the consideration of the direction, the simulation procedure is exactly the same as described above. Especially the calculation of the opening angle α is transferred to this case, which can be done due to the fact that the dimension of the recorded object is in general small compared to its distance to the hologram. The only difference is that the intensity $I(x, y)$ of the blurred texture image is not transferred to the position (x, y) in the simulated tomogram but to the intersection point (ξ, η) of the PSF-axis and the to be simulated tomogram (see figure 3.6)

The direction of the PSF-axis can be characterized by the focus coordinate (x, y, z_0) in Cartesian or as (ρ, θ, ϕ) in spherical coordinates, as mentioned in section 3.2.

The intersection point (ξ, η) lies on the PSF-axis and is therefore determined by the same angles θ and ϕ as mentioned above. Only the radial distance $\tilde{\rho}$ is different. Using the

intercept theorem one gets:

$$\tilde{\rho} = \frac{z}{z_0}\rho. \tag{3.7}$$

Reverse transformation to Cartesian coordinates with the changed radial distance $\tilde{\rho}$ delivers ξ and η:

$$\xi = \frac{z}{z_0}\rho\cos\phi\sin\theta \tag{3.8}$$

$$\eta = \frac{z}{z_0}\rho\sin\phi\sin\theta, \tag{3.9}$$

where ϕ and θ are taken from equation 3.5 and 3.4.

Following this process, there are a number of points in the center of the simulated tomogram where more than one intensity value is to be deposited and points in the outer part where there are none. In the first case the values are averaged, in the latter case the points are filled with the mean value of existing intensities from a neighborhood of three pixels in size.

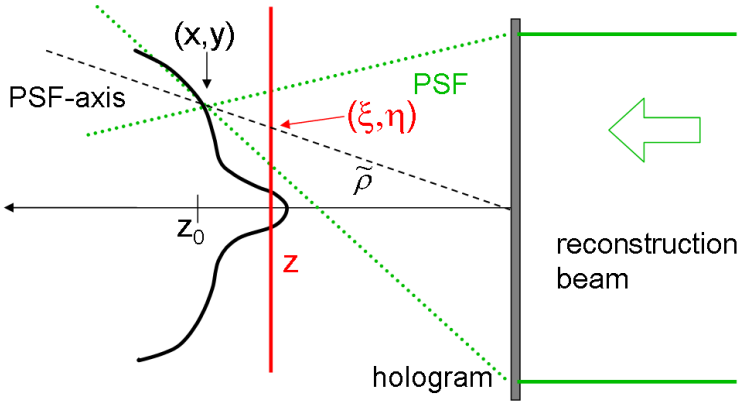

Figure 3.6: Position (ξ, η), to which the blurred intensity of the point (x, y, z_0) has to be transferred during the simulation process of a tomogram at z-value z.

Figure 3.7(b) shows the results in comparison to the real data 3.7(c).

Before discussing the accordance between simulation and reality, a summary of the simulation process is shown in figure 3.8.

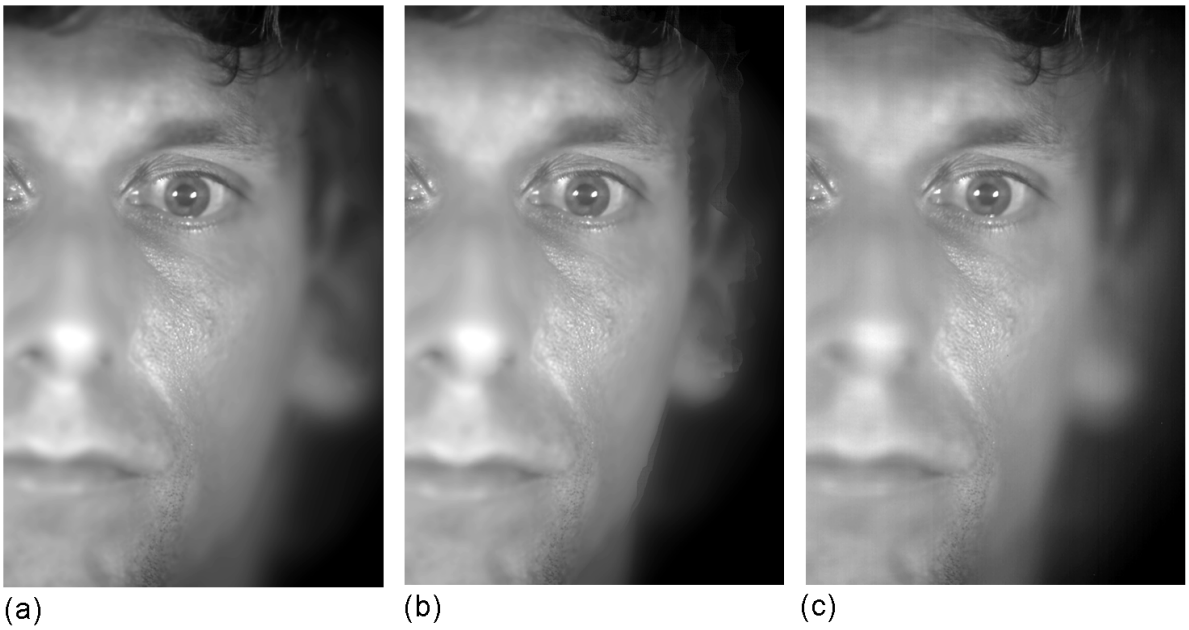

Figure 3.7: Simulated tomograms at the same z-position with parallel PSF-directions (a), individually calculated PSF-directions (b) and the real data (c).

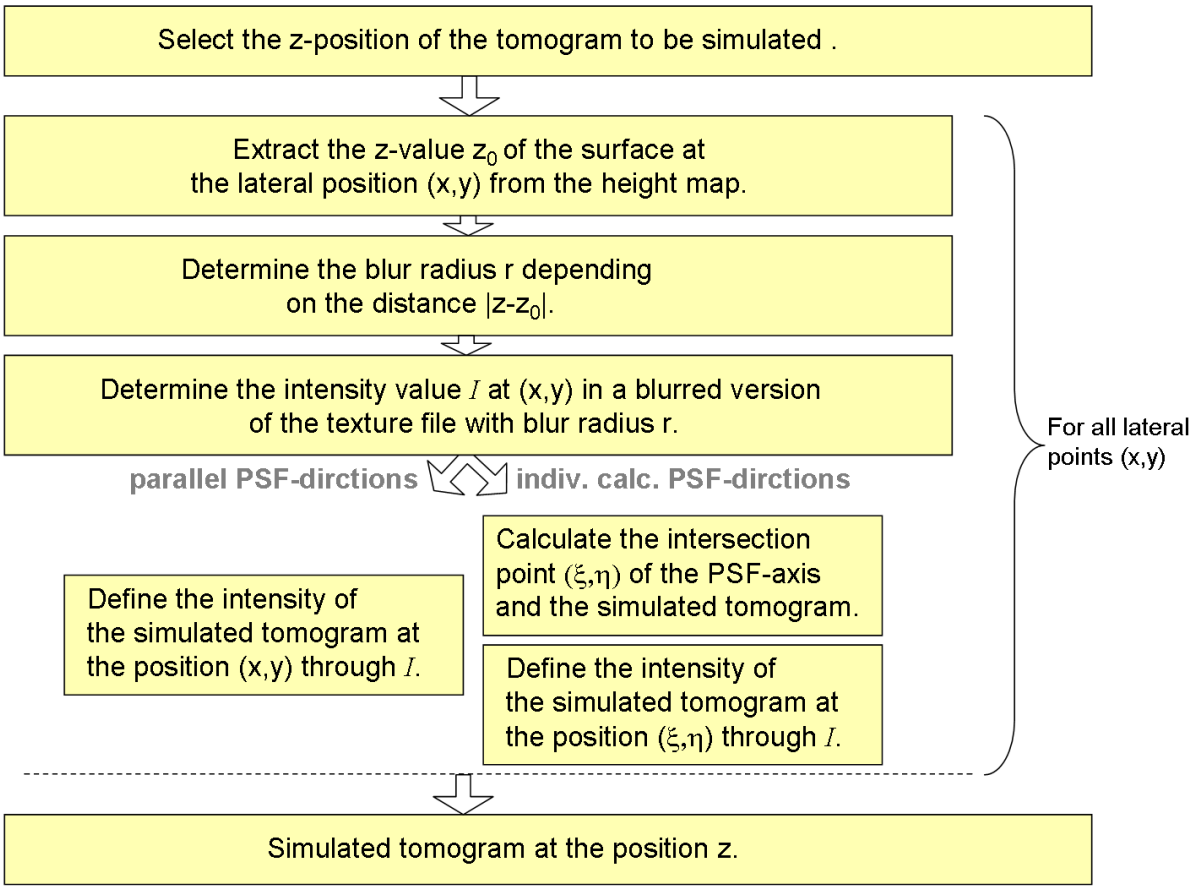

Figure 3.8: Flowchart of the simulation process.

3.5 Comparison between simulated and real data

In order to achieve comparability between the simulated and the real data, the same number of tomograms with the same distance and same lateral resolution as in the original data set were simulated. Noise and digitization artifacts as well as background light were neglected.

The influence of the different PSF-directions can be seen in outlines when comparing figure 3.7(a) and 3.7(b). The blurred image of the ear in 3.7(b) is closer to the center of the image than in 3.7(a). To illustrate this effect further, a horizontal cut through the image volume is regarded. A horizontal cut at a fixed y-coordinate is done by joining all intensity values along the x-axis for all tomograms into an image (see figure 3.9).

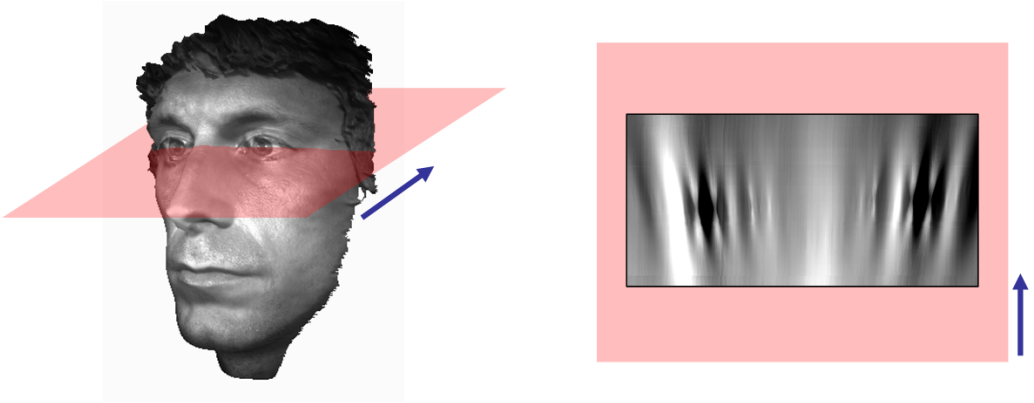

Figure 3.9: Illustration of a horizontal cut through the image volume. The intensities in x-direction at a fixed y-position are joined for all tomograms into an image.

Figure 3.10 shows a horizontal cut at eye-level for the simulated data with parallel PSF-directions (a), individually calculated PSF-directions (b), and the real data (c), averaged over 40 y-values. The red line marks the PSF-axis for a reflex of the eye. At this distance to the center, the difference in direction between the parallel and the individually calculated ones amounts to 3°, at ear-level it is around 8°. A comparison between the three shows qualitatively, that the simulation with the individually calculated PSF-directions is in better accordance with reality than the simulation with the parallel ones.

The qualitative, visible accordance between the simulated and the real tomograms is a first, promising indication for applicability of the simulated data for evaluation purposes. Ultimately however, the simulated and the real data have to behave in a similar way in the procedures of surface extraction. This can be clarified by regarding a focus profile created with an arbitrary focus criterion, for example the SML-operator (see section 4.1

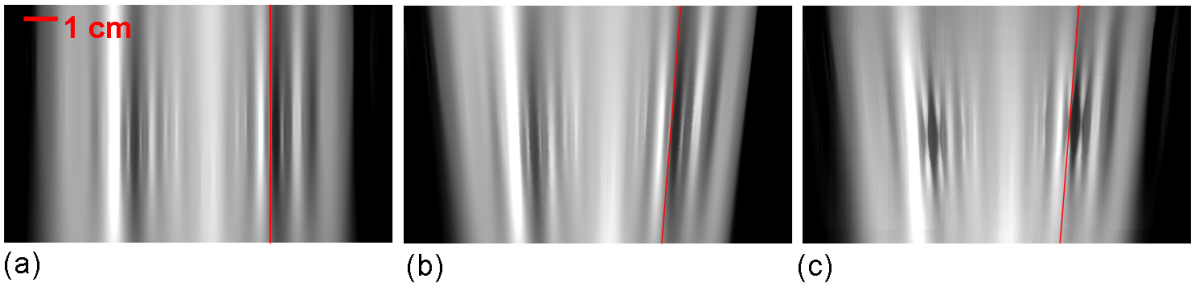

Figure 3.10: Horizontal cut averaged over 40 y-values for the simulated image volume with parallel PSF-directions (a), individually calculated PSF-directions (b), and the real data (c). The simulation with the individually calculated PSF-directions shows a better accordance with the reality than the parallel ones.

for a detailed description), which displays the level of focusing against the z-coordinate. Figure 3.11 shows the focus profile of the real and the simulated data with individually calculated PSF-directions. The peak itself correlates very well in form and height. Far away from the peak, the real data profile has a noise background, which is not present in the simulated profile. Further on noise will be added in order to make the simulation more realistic.

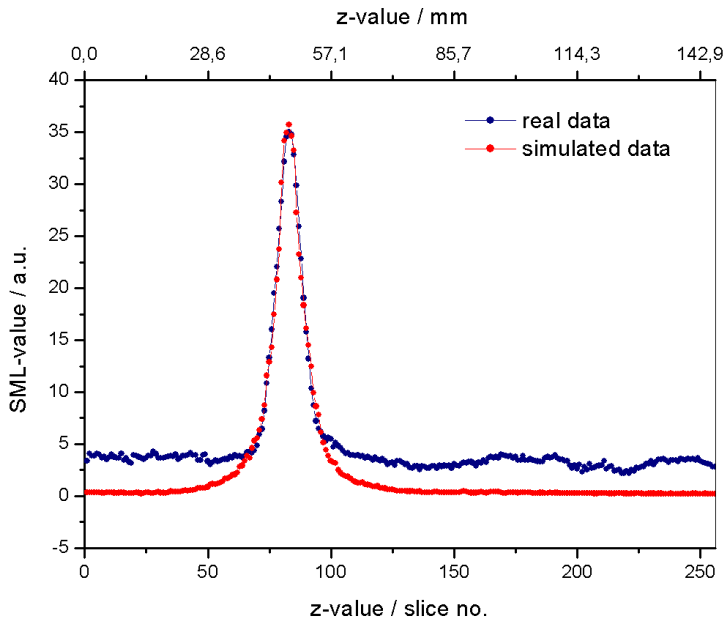

Figure 3.11: Focus profile of the real data (blue) and the simulated data (red). The peak shows in height and form a satisfactory correlation. In the real data profile there is background noise present other than in the simulated data profile.

3.6 Remarks

Some additional remarks concerning the simulation are collected in this section.

3.6.1 Shadowing effects

So far it was considered that the PSF emerges from the whole holographic plate. If the recorded object is of complex geometry, some regions of the object build a barrier for the scattered light of more hidden object points. The scattered light from such points can only reach part of the holographic plate (see figure 3.12). This is called shadowing. During the optical reconstruction process, the diffracted light forming the focus of this point only emanates from the region of the hologram, which was reached during the recording process. The shadowing effect influences slightly the direction of the PSF-axis. It will not

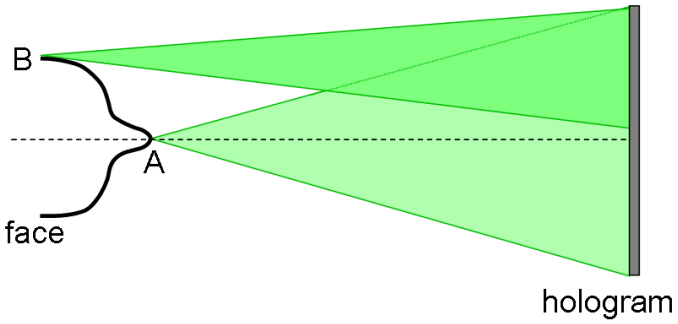

Figure 3.12: During the recording process, the scattered light from point A reaches the whole hologram, the one from point B underlies shadowing effects and reaches only part of the hologram. This leads to different PSF-geometries during optical reconstruction.

be considered any further since the qualitative properties of the imaging process are not changed.

3.6.2 Coordinate transformation

A set of tomograms with a specified distance accomplishes the digitization of the real image. The inter-tomogram distance as well as the absolute distance to the hologram are known quantities. The lateral resolution of the digitized tomogram is also known. These quantities belong to the real world coordinate system, measured in metric units.

Mainly, the surface extraction procedure uses image coordinates, where the lateral distances are measured in pixels and the axial distances in tomogram numbers. Only during

the production of the resulting three-dimensional computer model, the transformation to metric units is performed.

During procedures where the geometry of the PSF, which is a real world quantity, is taken into account, one has to transform the PSF into the image coordinate system, regarding the lateral and axial resolution and the absolute distance to the hologram. In such cases this transformation is performed without further notice.

3.6.3 Addition of noise

The simulation is more realistic if noise is added to the simulated tomograms, as was pointed out in section 3.5. The added noise follows a Gaussian normal distribution. Its extent can be characterized by the standard deviation of the normal distribution measured in gray values. This quantity is further on called noise level. A noise level of two means in this context that the gray value fluctuations follow a normal distribution with a standard deviation of two.

3.7 Summary

In this chapter, several different possibilities for evaluating surface extraction algorithms were discussed. Given that the algorithms are to be optimized for facial measurement, evaluation with the help of test objects that have neither the geometry of the face nor the texture properties of skin are not feasible. To overcome this problem, a simulation of the hologram tomography procedure was performed.

The classical method for image simulation works with an invariant point spread function, which cannot be assumed in holographic imaging. Due to this fact, a new simulation method for hologram tomography was developed. The results showed good correspondence with realistic data sets with little computational effort.

The simulated data set corresponds in geometry and in texture characteristic to the digitized hologram of a real face. Different amounts of noise can be added deliberately to make the simulation more realistic. The simulated data set has an absolute reference surface and is naturally in the same coordinate system, so that no matching is necessary. It can be used for a quantitative analysis of the performance of newly developed algorithms, since it behaves similar to the real data in the procedure of surface extraction.

Chapter 4

Contrast measurement

Chapter 2 gave a quick overview over the procedure of contrast measurement and surface extraction. This chapter discusses in detail the possibilities for contrast measurement. It includes on the one hand a discussion of various mathematical operators, so-called focus criteria, for contrast measurement and on the other hand the choice of the suitable neighborhood size on which these operators work.

The first section of this chapter (section 4.1) introduces twelve different focus criteria and verifies their suitability for holographic surface measurement using the simulation described in chapter 3. After that, the influence of the neighborhood size is discussed and an algorithm is developed, which automatically adapts the neighborhood size to the contrast features of the texture (section 4.2).

4.1 Evaluation of focus criteria

Defocused images inherently have less information than sharply focused images, so most vertebrate animals having eyes with a refractive lens focus to improve the quality of their vision, thereby increasing the information in the images obtained. If one wants to give computer systems the same ability, the used CCD camera needs computer controlled aperture, focal length and focusing distance [Kro87]. This leads directly to the question how the sharpest image can be distinguished from unfocused ones.

Various autofocus algorithms have been proposed in many different applications. Apart from the autofocusing functionality of commercial photo cameras, many applications in robotic and machine vision need high-quality autofocus algorithms for all kind of camera systems [NKCNA01].

Autofocusing is also a fundamental technology for automated biological and biomedical analysis and is indispensable for routine uses of microscopes on a large scale [SDN04], because automated focus systems save time and focus better and more consistently than do humans [FCC+91].

The core of an autofocusing algorithm is a mathematical operator, a focus criterion, working on the pixel representation of the digital image and assigning a quantitative value to a whole image or a specified part. Such a value is a measure for the sharpness of the image or its part and is further on called focus value.

The performance of the different autofocus algorithms depends strongly on the properties of the object to be focused on as well as those of the imaging systems. A lot of research was performed for finding the suitable algorithm for the different kinds of applications. Before going into detail about the general outcome of these studies, the most important focus criteria are mentioned.

Focus criteria deliver a value that indicates the degree of focusing in an image or a region of an image. They are based on the knowledge that defocusing produces a decrease in the high frequency content of the image [LG82]. This can also be seen in the assumption that well-focused images contain more information and detail (edges) than unfocused images [YOJS93].

In the application of holographic topometry, the images under consideration contain focused and unfocused regions. Due to this, not a focus value of the whole image is to be calculated but of each individual point in the image. Though the intensity of one point alone does not inherently have sharpness or not, the point always has to be considered in its local environment. The focus criteria presented in the following section are therefore evaluated in a pixel neighborhood $U_d(x, y)$ of a point (x, y) with a size of $d \times d$ pixels (see section 2.2.6 and appendix A):

4.1.1 Focus criteria

The following focus criteria were evaluated in the course of this work:

- Tenengrad
- Squared Gradient (SquGrad)
- Absolute Gradient (AbsGrad)
- Laplacian
- Sum Modified Laplacian (SML)
- Variance (Var)
- Normalized Variance (NorVar)
- Entropy
- Sum Square Anti Gaussian (SSAG)
- Visibility
- Cross Sum Modified Laplacian (XSML)
- Sum Modified Fourth Order Derivative (SM4D)

A detailed description of the criteria can be found in appendix A. The first ten criteria are taken from literature, the last two are newly developed modifications of the well known SML operator proposed by Nayar [NN94]. The SML operator uses a modification of the Laplacian for contrast measurement. The XSML operator calculates additionally the Laplacian for the diagonal neighbors of a pixel, in the SM4D criterion the fourth order derivative is used.

As mentioned before, finding the best focus criterion depends on the imaging system, on the characteristics of the mapped object and also on the field of application.

Several evaluations of focus criteria can be found in literature. The following paragraph gives a quick overview of the results. Most of the authors discussed among other criteria the Tenengrad operator, the variance with or without normalization, the two versions of the gradient, the Laplacian and some of them also the entropy. The SML-operator as well as the SSAG, visibility and the two newly defined operators were not evaluated.

Krotkov found in his fundamental work [Kro87] that the Tenengrad-operator is the best for autofocusing purposes in robotics, while the Laplacian is very unstable and the entropy shows multimodality. In the same field of application Ng Kuang Chern et. al. [NKCNA01] found the Tenengrad criterion and the variance equally usable.

Ligthart and Groen [LG82] on the other hand found that for autofocusing purposes in microscopy the variance, and slightly better also the normalized variance as well as the squared gradient lead to the best results.

In tissue microscopy, Yeo et. al. [YOJS93] found that the variance performs best in the presence of noise, but suffers from the drawback that it may not locate the best focus. In comparison to this, the Tenengrad function has the sharpest peak.

In microscopic molecular cytogenetic analysis Santos stated the Tenengrad function the best from the criteria above [SOdSV+97], while Sun et. al. [SDN04] found on various biomedical samples that the normalized variance performs best. The same result was gained by Firestone et. al. [FCC+91] for microscopic tissue images.

In the regime of depth from focus applications in microscopy, Subbarao et. al. found the Laplacian performing best [ST98], while for acoustic holographic images Yin et. al. favors the variance.

The former remarks about the outcome of the evaluations make clear that a general solution for the choice of the best criterion in hologram tomography cannot be found in literature. Due to this, a comparative study for this special kind of application was performed.

4.1.2 Suitable focus criteria for hologram tomography

In hologram tomography the topometry of the recorded surface is extracted through maximization of the focus criterion over all tomograms. Therefore the most important quality of the focus criterion should be the similarity of the extracted surface with the original one. Qualities like unimodality or monotonicity at the sides of the peak are not as relevant as in autofocusing. This is why the evaluation of criteria is be done by comparing the original with the extracted surface.

The texture of the recorded object influences tremendously the performance of the focus criterion. Holographic topometry methods aim to the measurement of human faces. This is why the scattering properties and structure of skin should be considered while choosing

the best criterion. Since no test object with simple geometry can simulate the properties of a human face satisfactorily, the original texture of a human face should be used for evaluation purposes.

There is no adequate reference of a recorded human face, so the simulated data set was used for the evaluation (see chapter 3). Critical parameters like noise level or PSF-directions were varied in the simulation and thus their influence can be analyzed in a systematic way.

The twelve focus criteria mentioned in section 4.1.1 were tested on different data sets. A section of a human face was simulated first with parallel PSF-directions and then with individually calculated PSF-directions. Both data sets were tested without additional noise and with added Gaussian noise of a standard deviation of one and two gray levels. A surface was extracted through maximization parallel to the z-axis with the help of all twelve criteria. The resulting height map was compared with the reference and the mean absolute deviation in millimeters was calculated. In order to eliminate the influence of the size d of the neighborhood, the results were averaged over neighborhood sizes of $d = 11$, $d = 13$, $d = 15$, and $d = 17$ pixels. Figure 4.1 shows the results for parallel PSF-directions, figure 4.2 for individually calculated PSF-directions. The mean absolute deviation of the Laplacian was one order of magnitude larger than that of the other criteria and is therefore not displayed.

Examining figure 4.1 and figure 4.2 in detail, the behavior of the criteria can be divided into two groups. For some criteria, like the SM4D, the error increases rapidly if noise is added, other criteria are less sensitive against the addition of noise. The error increases for some criteria tremendously during the change-over from parallel and individually calculated PSF-directions, like for the variance. In order to quantify this different behavior, the normalized difference between the error at noise level two and zero and the difference between parallel and individually calculated PSF-directions averaged over the three different noise levels were calculated. The results can be seen in figure 4.3 and figure 4.4, respectively.

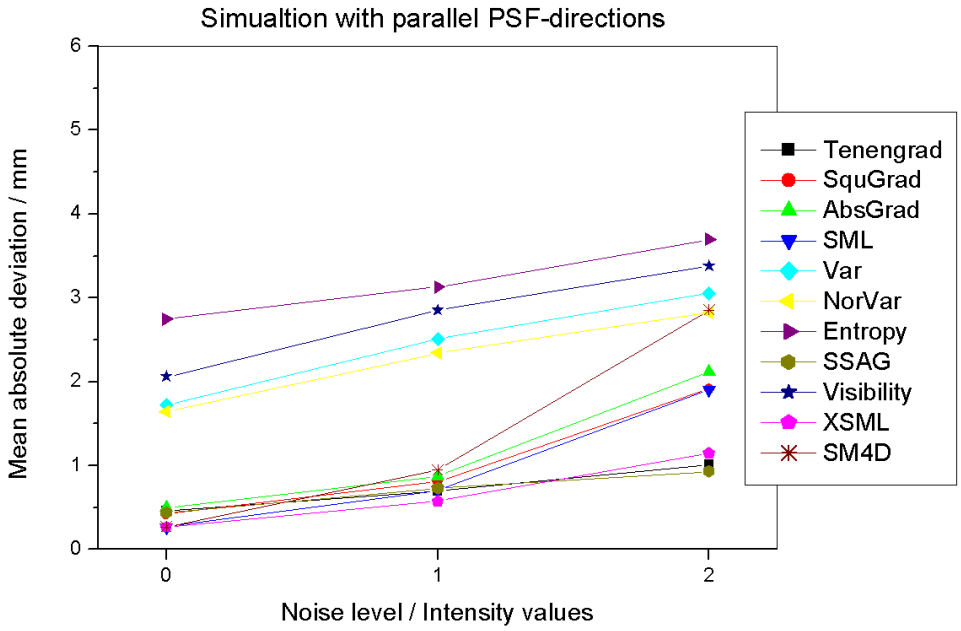

Figure 4.1: Mean absolute deviation between the reference surface and the surface extracted with the respective criterion. The data was simulated with parallel PSF-directions and different noise levels. Noise level zero indicates that no noise was added, whereas a level of one and two means a noise standard deviation was one and two gray levels, respectively.

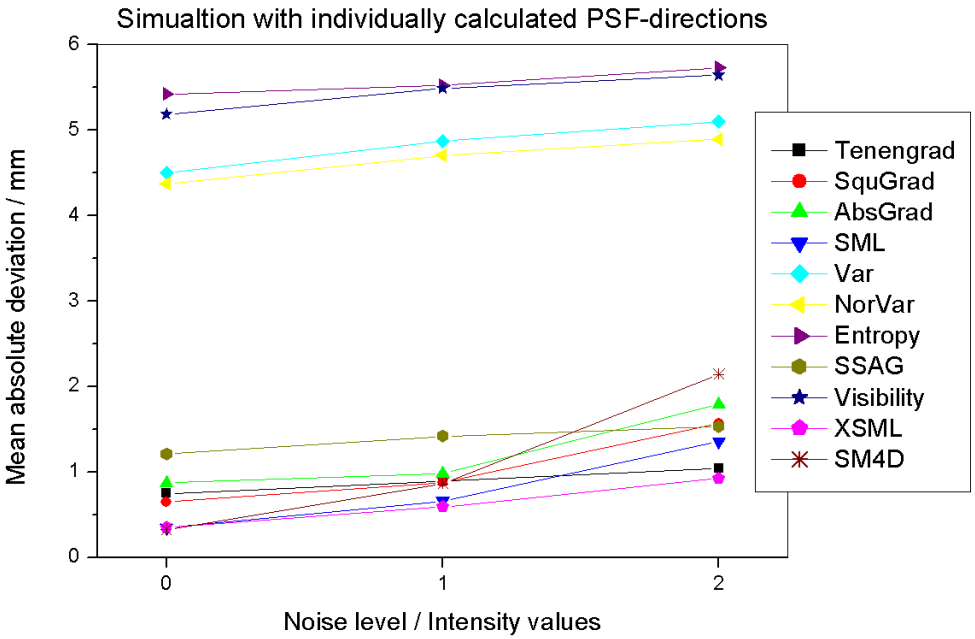

Figure 4.2: Mean absolute deviation between the reference surface and the surface extracted with the respective criterion. The data was simulated with individually calculated PSF-directions and different noise levels. Noise level zero indicates that no noise was added, whereas a level of one and two means a noise standard deviation of one and two gray levels, respectively.

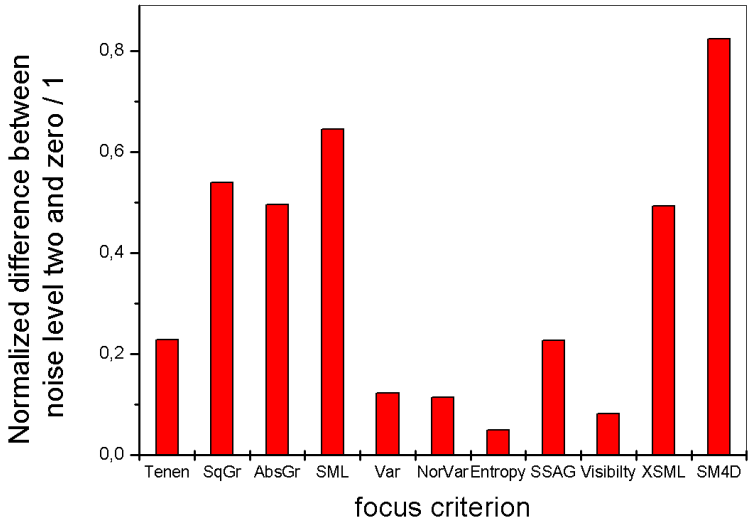

Figure 4.3: Normalized error difference between noise level two and zero for the different focus criteria and individually calculated PSF-directions.

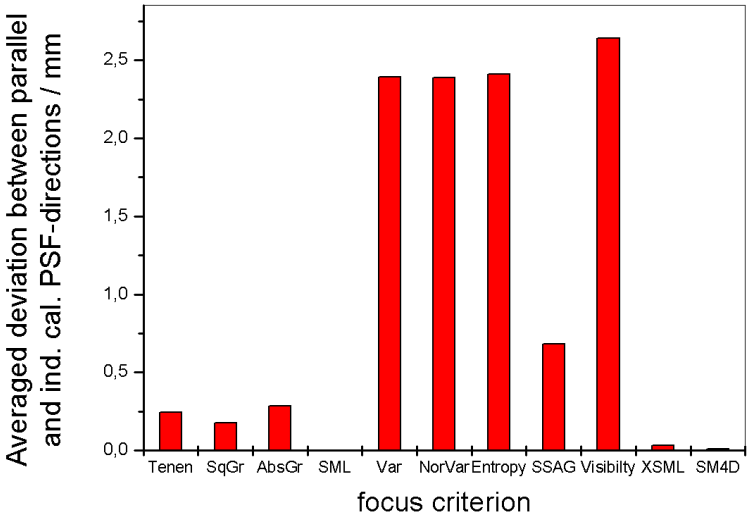

Figure 4.4: Averaged error deviation between parallel and individually calculated PSF-directions for the different focus criteria.

The SM4D operator has the highest sensitivity against noise, followed by the SML, the XSML and the two gradient operators. The other operators are clearly less sensitive. The SM4D bases on the fourth order derivative, the SML and XSML on the second order derivative, the gradient operators on the first order derivative. It can be concluded that the higher the derivative the higher the sensitivity against noise.

All operators are calculated on the same neighborhood size. The ones mentioned above calculate a contrast value only from a small number of points (four to eight) and sum these values over the neighborhood. The other operators take for contrast calculation

more points, some of them the whole neighborhood into account. This makes them less sensitive against noise. The same tendency can be seen while comparing the SML and XSML. The XSML operator takes eight instead of four points into account and reacts less to noise.

Operators working on many points however have the drawback that the mismatch between the direction of image formation and contrast maximization leads to an increasing number of errors. If the operator works locally, influencing factors from neighboring points are lower. The artifacts created by this influence are discussed later in this chapter as artifacts of type two. This is why especially for operators like the variance, normalized variance, visibility and entropy the error increases when switching from parallel to individually calculated PSF-directions.

From figure 4.2 one can generally say that the newly defined XSML operator performs best. This is also true for realistic data where the noise level is approximately one. The only other interesting criterion might be the Tenengrad operator, which has a higher average error but is less sensitive against noise. Figure 4.3 and figure 4.4 show that it is a good compromise between sensitivity against noise and direction mismatch. A qualitative comparison of the XSML and Tenengrad criterion is given in the following paragraph.

The two groups of criteria described above react differently to noise and direction mismatch and produce different kinds of artifacts, which will now be described.

1. If the signal-to-noise ratio is low, delocalized surface points occur. They have arbitrary values and build small spots. Regions of low signal-to-noise ratio in the face are normally regions with steep slope like the wings of the nose (see for example figure 4.5(d)) .

2. The second effect is caused by a non-zero angle between the direction of maximization and the PSF-direction. Especially if points show a high contrast in the texture, this mismatch can lead to artifacts in the extracted surface (see figure 5.18 for a schematic illustration). They occur as regions of continuously increasing or decreasing values in the height map and are attached to "edges". In the human face they can be found on lips, nose, eyebrows, etc. (see for example figure 4.5(c)).

Figure 4.5 shows a section of the texture map for orientation purposes (a) and height maps of the same data set calculated with a neighborhood size of 13 pixels with the XSML-operator (b and d) and with the Tenengrad-operator (c and e). Basis was in the case of (b) and (c) simulated data with a noise level of 1.5, in the case of (d) and (e) one

of 3. The red arrows mark artifacts in the height map. They occur in the height maps generated with the XSML operator mainly at the side of the nose and are of type one while those generated with the Tenengrad operator are mainly of type two. Comparing the results of the XSML operator at the different noise levels it is obvious that the number of artifacts of type one increase. The height maps generated with the Tenengrad operator show nearly the same artifacts of type two, the one with noise level 3 shows also some artifacts of type one but fewer than the XSML-version at the same noise level.

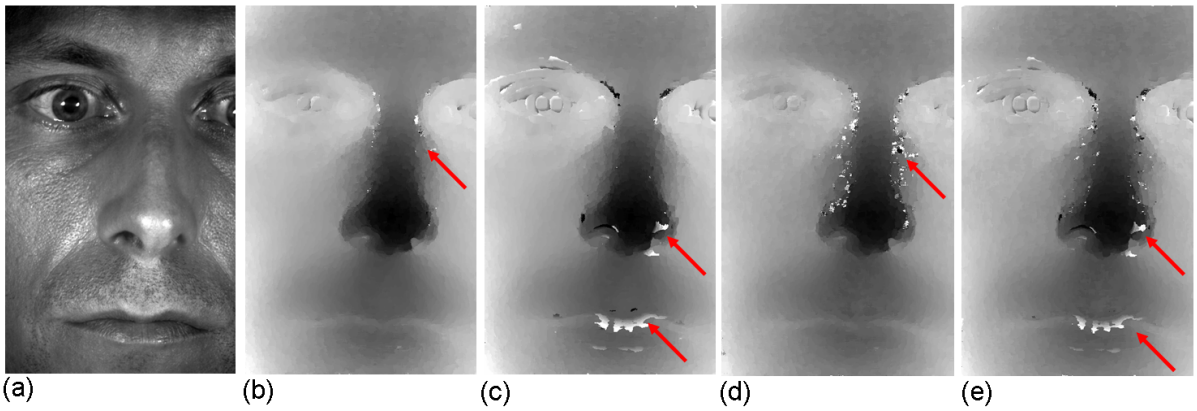

Figure 4.5: Reference height map (a) and height maps extracted with the XSML operator (b and d) and with the Tenengrad operator (c and e) with a neighborhood size of 13 pixels. The standard deviation of the added noise was 1.5 (b and c) and 3 (d and e).

The real data shows the same effects as the simulated data. Figures 4.6 and 4.7 show focus profiles from the real data along the z-direction.

Figure 4.6 demonstrates profiles created with XSML and Tenengrad that are not affected by artifacts. Both show a single maximum whereas the profile generated with the Tenengrad operator shows it more distinctively. This indicates the higher robustness of the Tenengrad operator against noise. The profiles plotted in Figure 4.7 on the other hand are corrupted with artifacts of type two. The profile calculated with the XSML operator has its maximum at the position of the surface. The Tenengrad profile has also a local maximum at this position but additionally a peak caused by the PSF of points from the boundary of the lip to the surrounding skin. This leads to the described type two artifacts.

Artifacts of type one are easier to be dealt with like ones from type two since the former can be avoided by noise filtering or locally bigger neighborhoods (see section 4.2 and chapter 5). Due to this fact, regardless to the noise level present in the data, the newly invented XSML-operator is to be used for surface extraction. Especially if using the holographic topometry system described here, it produces the lowest error.

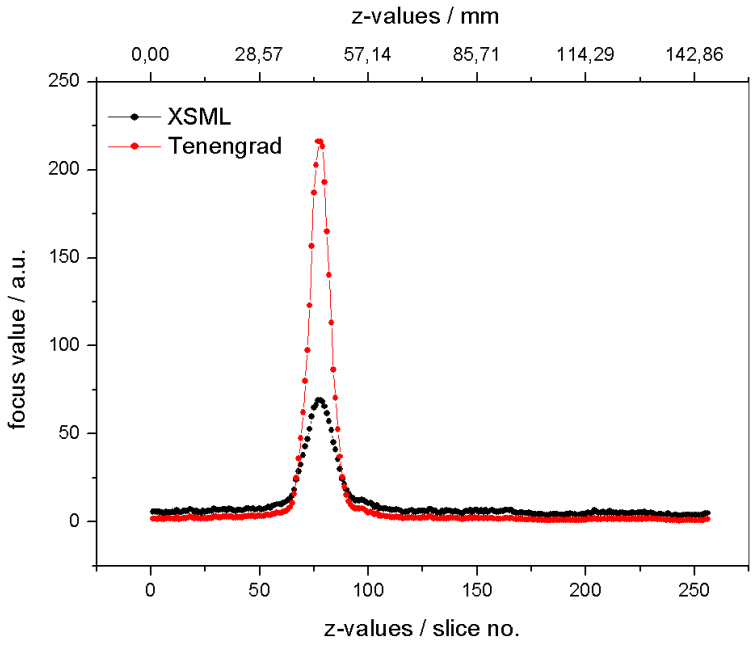

Figure 4.6: Focus profile of real data along the z-direction in the region of the cheek for the XSML and the Tenengrad operator with a neighborhood size of 13 pixels. Both show one single maximum but the profile generated with the Tenengrad operator is more distinctive.

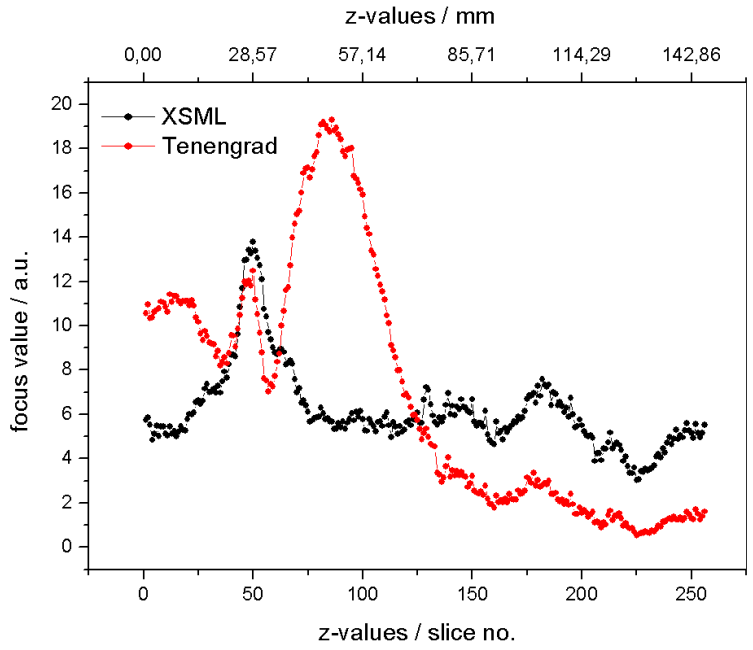

Figure 4.7: Focus profile of real data along the z-direction at the boundary of the lip for the XSML and the Tenengrad operator with a neighborhood size of 13 pixels. The XSML profile shows a pronounced maximum at the position of the surface. The Tenengrad profile has also a local maximum at this position but additionally a peak caused by the PSF of points from the boundary of the lip and the surrounding skin.

4.2 Multi-scale approach

Up to now the optimal choice of a focus criterion was discussed. This section will deal with the optimal choice of the used neighborhood size.

The choice for the neighborhood size is a critical step in the procedure of surface extraction. The lateral resolution of the extracted surface is limited by the size of the neighborhood, since structures smaller than this size cannot be resolved. Thus the neighborhood size should be chosen as small as possible. On the other hand are small neighborhoods very sensitive to noise. The lower the signal-to-noise ratio, the larger the neighborhood should be. Even in the absence of noise, the minimal neighborhood size is limited to a certain extend by the feature size of the texture of the recorded object. With a neighborhood smaller than any feature it is impossible to distinguish between focused or defocused images. When choosing a neighborhood size, one always has to compromise between lateral resolution and robustness against noise.

4.2.1 Influence of the neighborhood size

Figure 4.8 shows the dependency of the mean absolute deviation between the reference surface and the surface extracted out of the simulated data with individually calculated PSF-directions and the chosen neighborhood size. The curve in figure 4.8(a) is based on noiseless data, whereas the one in figure 4.8(b) on data with a noise standard deviation of three. Curve (a) shows the tendency described above. Even in the absence of noise the result improves while increasing the neighborhood size due to the number of texture features included in one neighborhood. If the neighborhood gets to large, small features or pronounced structures cannot be acquired correctly and the error increases. In the case of a high noise level, the error decreases with the neighborhood size due to the increasing robustness against noise. The curves were calculated with the XSML focus criterion.

One solution to this tradeoff is the usage of active illumination, a procedure discussed among others by Noguchi [NN96] in the field of microscopic shape from focus. In doing so, the illumination consists of a defined pattern which superimposes the natural texture of the recorded object. The feature size is therefore constant and the neighborhood size can be adjusted optimally. Additionally, the active illumination leads to a higher signal-to-noise ratio due to the high contrast of the pattern. A similar approach works also in hologram tomography, as was demonstrated in the works of Bongartz [Bon02] and Giel [Gie03], but has the drawback of loosing the eye-safeness while recording a human

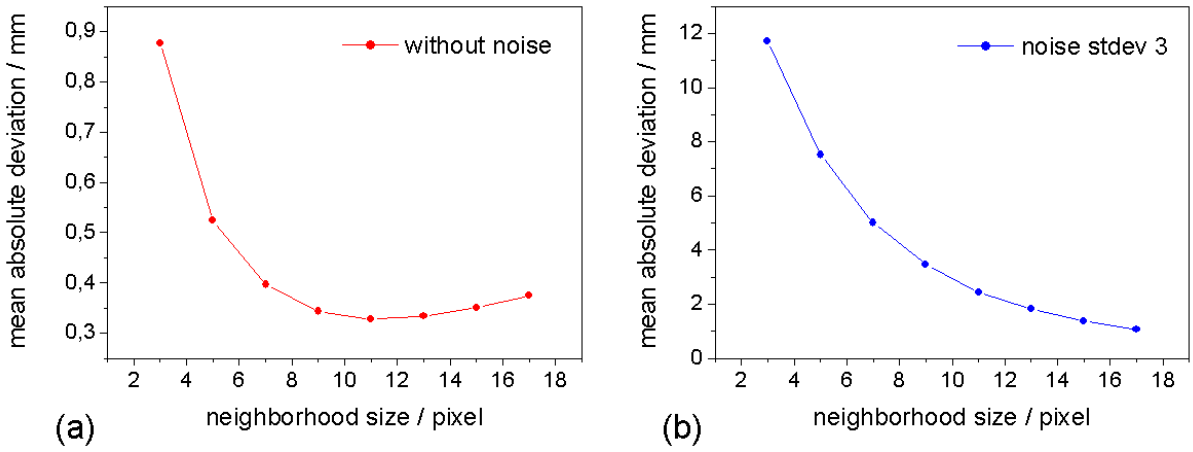

Figure 4.8: Dependency of the mean absolute deviation between the reference surface and the surface extracted from the simulated data with individually calculated PSF-directions and the chosen neighborhood size. The curve in figure (a) is based on noiseless data, whereas the one in figure (b) on data with a noise standard deviation of three. While there is an optimal neighborhood size in (a), in the presence of noise (b) the error decreases with increasing neighborhood size due to the improved robustness against noise. Mind the different scaling of the vertical axis.

face. Frey [Fre05] recommended an eye-save recording without structured illumination and surface extraction based only on the contrast of the skin of the human face. The usage of the intrinsic texture however leads to the difficulties described above.

Figure 4.9 shows exemplarily height maps extracted from real data for the smallest possible neighborhood size of 3 pixels (a) and 17 pixels (b). In regions of high intrinsic contrast, like the mustache of the proband, a high-quality surface can be extracted using a neighborhood size of 3 pixels with a lateral resolution so high that even single hairs can be distinguished. In regions of lower intrinsic contrast however, the extracted surface is affected by noise. Using a neighborhood size of 17 pixels on the other hand allows for a noise resistant surface extraction in regions of lower intrinsic contrast, while the lateral resolution is reduced. The mustache appears as a connected region and especially the boundary to the skin is not pictured in a realistic way. Choosing one fixed neighborhood size, one always has to compromise between noise robustness and lateral resolution. A tremendous improvement would be an adaptive neighborhood size taking the local contrast properties of the texture into account.

Bürkle [Bür04] introduced an algorithm with an adaptive choice of the neighborhood size in the field of microscopic depth from focus. First, the height map is produced with a large neighborhood size, so that the influence of noise or prominent texture features is

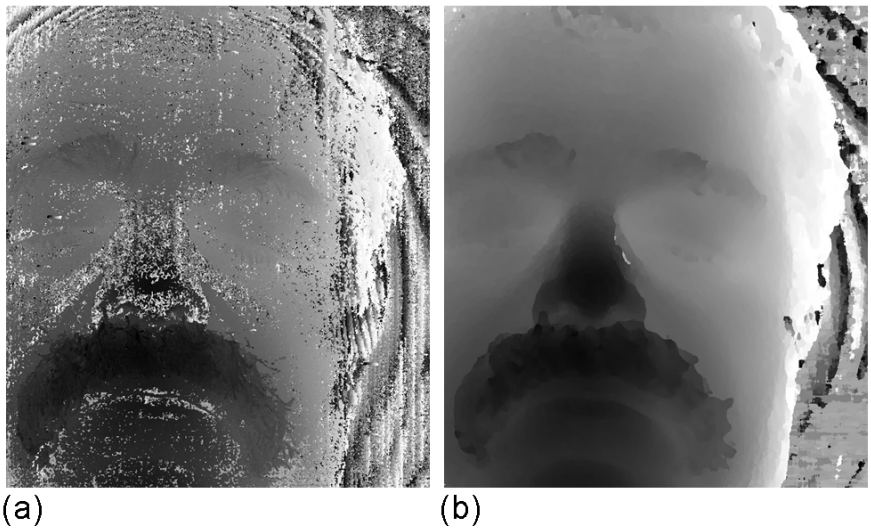

Figure 4.9: Exemplary height maps extracted out of real holographic data with a neighborhood size of 3 (a) and 17 (b) pixels. The small neighborhood allows highly resolved surface extraction in regions of high intrinsic texture contrast, i. e. the regions with hair, while big neighborhood sizes are robust against noise in regions of low intrinsic contrast but reduce the lateral resolution in highly structured regions.

limited. In combination with the localization of the maximum, also a certain z-interval is defined through the set of z-values, whose focus value lies above the maximal value multiplied with a factor γ (see figure 4.10). Then the neighborhood size is reduced and the newly found maximum is accepted only if its z-interval lies within the former. He reports good results gained for plane objects or ones with low declination. Steep edges on the other hand cannot be detected correctly.

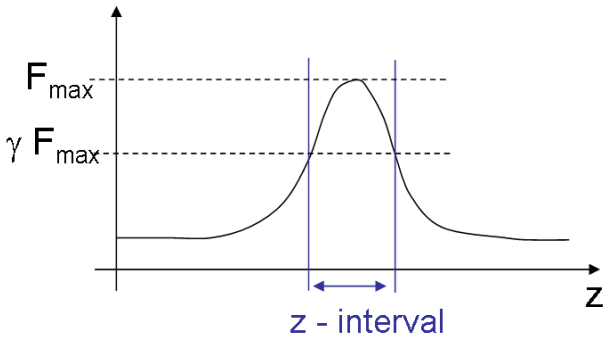

Figure 4.10: Definition of the z-interval proposed by Bürkle [Bür04] for an algorithm with an adaptive neighborhood size

In hologram tomography, the smallest features to be displayed are skin pores and hairs. Skin pores normally do not come along with steep edges, but hairs do. So the algorithm

proposed by Bürkle does not work satisfactorily, as is illustrated in figure 4.11. Part (a) of figure 4.11 displays the section of the face. It contains part of the mustache of the proband as well as skin. (b) and (c) show height maps created with a neighborhood size of 3 and 17 pixels, respectively. (d) shows the results obtained with the adaptive algorithm proposed by Bürkle. In some parts, the lateral resolution is high enough to show single hairs, but the artifacts at the boundary between hairs and skin remain (red arrows). Part (e) shows the result of the adaptive algorithm developed in this thesis. The lateral resolution of the small neighborhood is reestablished and artifacts of big neighborhoods removed while no corruption through noise is visible. In the following section, the newly developed algorithm will be explained.

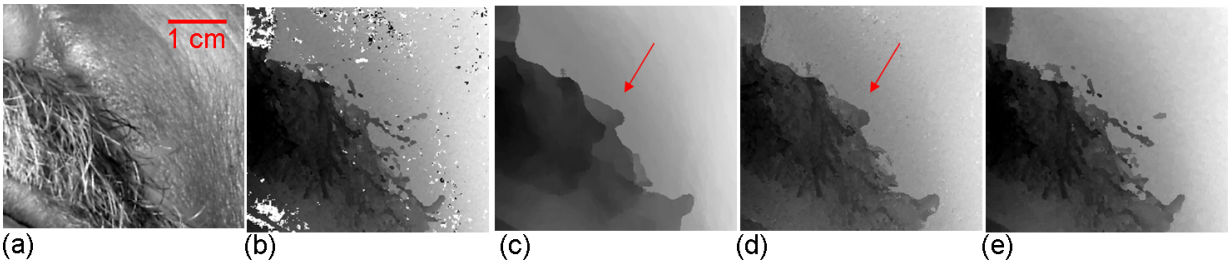

Figure 4.11: Section of a face with part of a mustache (a). Height map produced with a neighborhood size of 3 pixels (b) and 17 pixels (c). (d) illustrates the result created with the adaptive approach proposed by Bürkle [Bür04], (e) the results created with the adaptive approach developed in this thesis. The red arrows in picture (c) and (d) indicate the artifacts evoked by big neighborhood sizes, that are transfered to the result by the adaptive algorithm proposed by Bürkle.

4.2.2 Reliability of extracted surface points

In the approach from Bürkle, extracted surface points with small neighborhood sizes are accepted if their peak lies in the same z-region as the peak of the larger neighborhood. As described above, this leads to problems close to high-contrast regions. To overcome the problem, a different quantity for the reliability, a so-called confidence value c, is defined here.

If some parts of the face have high intrinsic contrast, their focus profile shows a distinctive maximum. If the intrinsic contrast is low, the signal-to-noise ratio is decreased and a distinction between the maximum evoked by the focus and the noise is difficult. The difference between the maximum and the background is therefore a good measure of the intrinsic contrast of the texture and its portability to surface extraction.

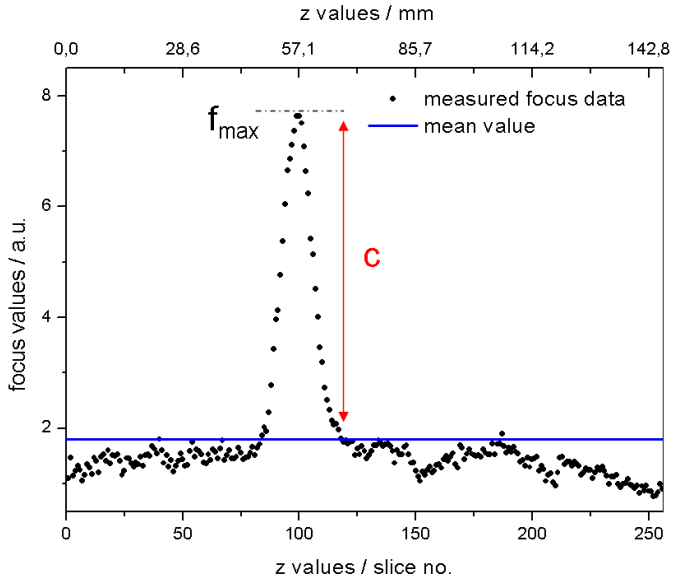

Figure 4.12: Focus profile of one exemplary surface point. The blue line indicates the mean value, the relative height of the maximum f_{max} against this mean is the basis for the definition of the confidence value c.

Figure 4.12 shows the focus profile of one exemplary point. To define the relative height of the maximum, first the mean focus value $\bar{f}$ is evaluated (blue line). The confidence value, depending on the chosen neighborhood size d, is then defined as:

$$c_d(x, y) = \frac{f_{max} - \bar{f}}{f_{max}} \tag{4.1}$$

The normalization leads to confidence values between zero and one, a higher value indicates a more reliable surface point.

The correlation between a high confidence value and a reliable surface point can qualitatively be seen in figure 4.13 for the neighborhood sizes of 3 and 17 pixels. In particular the region with hairs show high confidence values whereas the sides of the nose show low values. The latter regions are more affected by noise.

4.2.3 Adaptive selection of neighborhood sizes

Once a quantitative measure for the reliability of the extracted points is defined, a fusion of height maps created with different neighborhood sizes can be accomplished. Since height maps with the smallest neighborhood provide the highest lateral resolution, all reliable points should be accepted. Only for unreliable ones, height maps created with

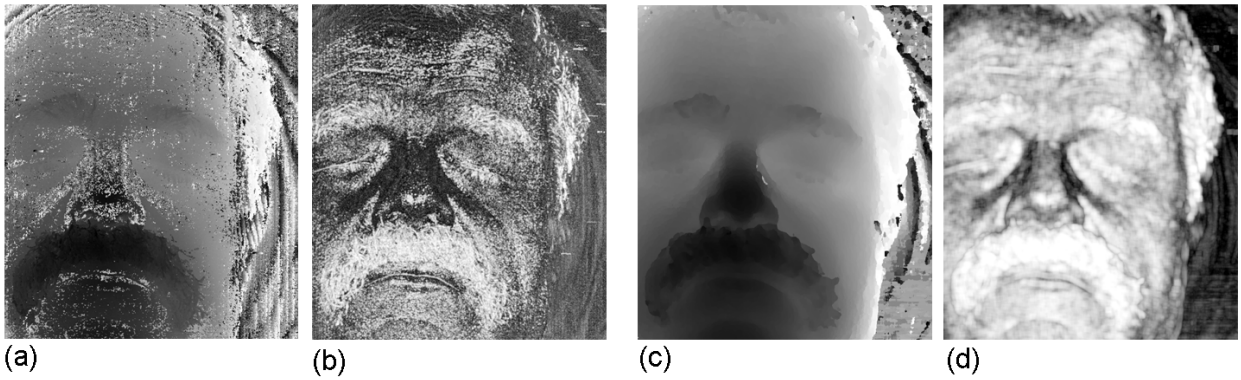

Figure 4.13: Height maps created with a neighborhood size of 3 (a) and 17 (c) pixels, and the corresponding confidence values (b and d). White corresponds to a confidence value of one, black to zero.

bigger neighborhoods should be taken into consideration. In order to distinguish between reliable and unreliable points, a threshold has to be selected.

In order to find a strategy for defining these thresholds, histograms of confidence images for a typical digitized portrait hologram were regarded. Figure 4.14(a) shows exemplarily the confidence image for a neighborhood size of 5 pixels. Figures 4.14(b), (c), and (d) display histograms of confidence images for a neighborhood size of 3, 5, and 11 pixels, respectively. For a small neighborhood size, very few reliable surface points are present. The histogram shows mainly a normal distribution of the unreliable points with a slight asymmetry caused by the reliable points. Increasing the size of the neighborhood, the number of reliable points increases, so that a second peak appears with its center shifting to higher confidence values (see figure 4.14(c) and 4.14(d)).

Unreliable points can be excluded if the threshold is chosen such that the first peak is eliminated. Assuming it to be symmetric, one has to add the width w between the starting point and the maximum to the maximal confidence value and gets the threshold t_d in dependence of the neighborhood size as is indicated in figure 4.15.

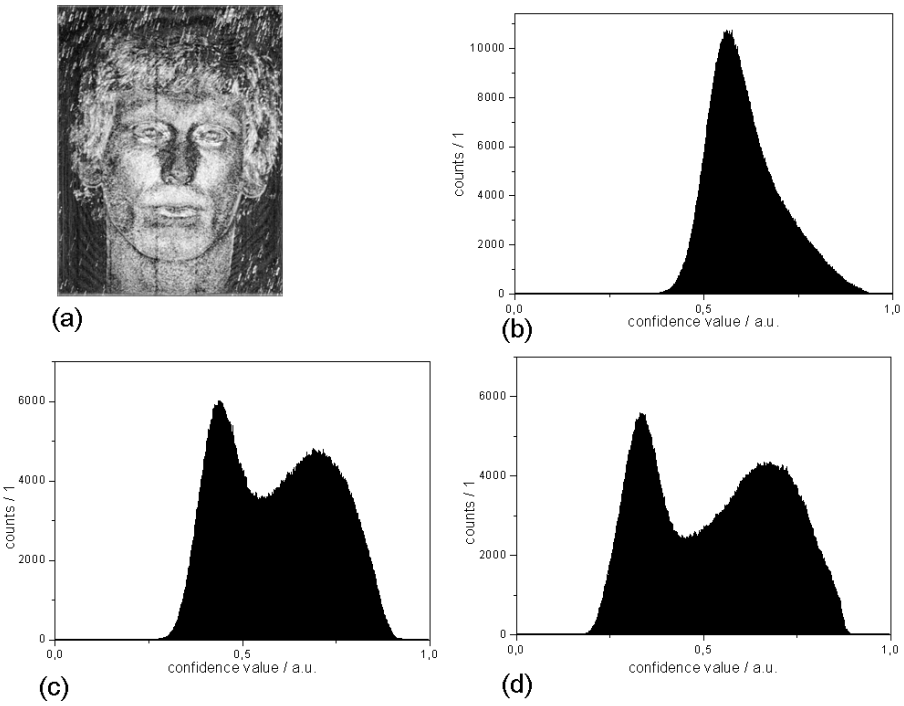

Figure 4.14: Example of a confidence image created with a neighborhood of 5 pixels (a) and histograms of the same with neighborhood sizes of 3, 5, and 11 pixels. An increasing neighborhood size leads to the appearance of a second peak belonging to reliable surface points, while the first peak can be associated with unreliable points.

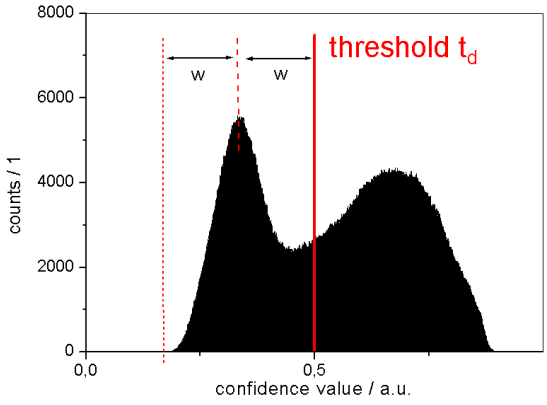

Figure 4.15: Histogram of the confidence image calculated with a neighborhood size of 11 pixels. If one adds the distance w between the starting point and the maximum to the maximal confidence value, a threshold t_d can be determined.

4.2.4 Typical threshold values

Since the determination of the threshold described above depends on the image content, an automated procedure should rely on predefined thresholds. To acquire these, 25 holograms of human faces were investigated. Thresholds for neighborhood sizes of 3, 5, 11, 15, 17, 21, and 23 pixels were build following the procedure described above. The average of the 25 built thresholds as well as an exponential fit with the fit model

$$t(d) = c + A \exp\left(-\frac{d}{s}\right) \tag{4.2}$$

can be seen in figure 4.16.

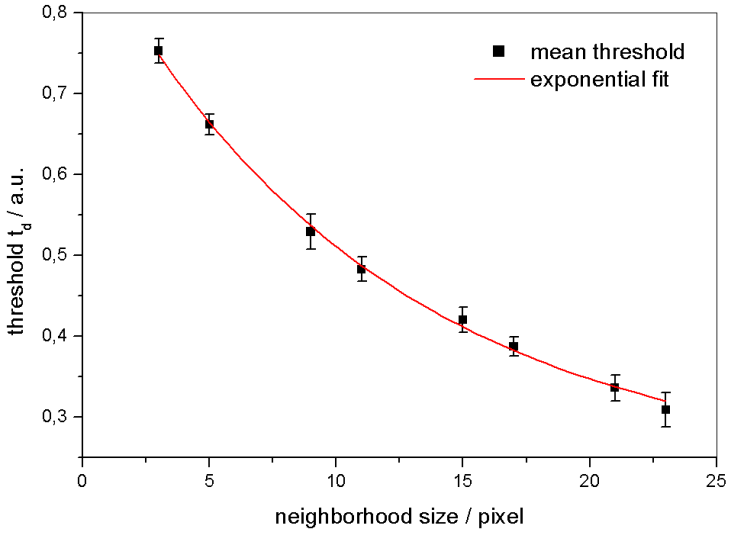

Figure 4.16: Threshold t_d against the neighborhood size. The threshold was averaged over 25 values, the error bar represents the standard deviation. The data was fitted with an exponential fit.

The fit of the 25 investigated holograms led to an exponential decay with the following parameters:

$$t(d) = 0.23 + 0.67 \exp\left(-\frac{d}{11.45}\right) \tag{4.3}$$

With this equation, a threshold can be determined for arbitrary holograms. To compensate for fluctuations in the quality of the real image[1], in the same automated process, two other thresholds curves are generated. They were calculated by fitting an exponential

[1]Different scattering properties of the skin or a different pulse energy influence the quality of the real image.

curve to the addition and substraction of the two standard deviations to the mean value. Such a 2σ-region contains nearly all[2] determined thresholds (see figure 4.17).

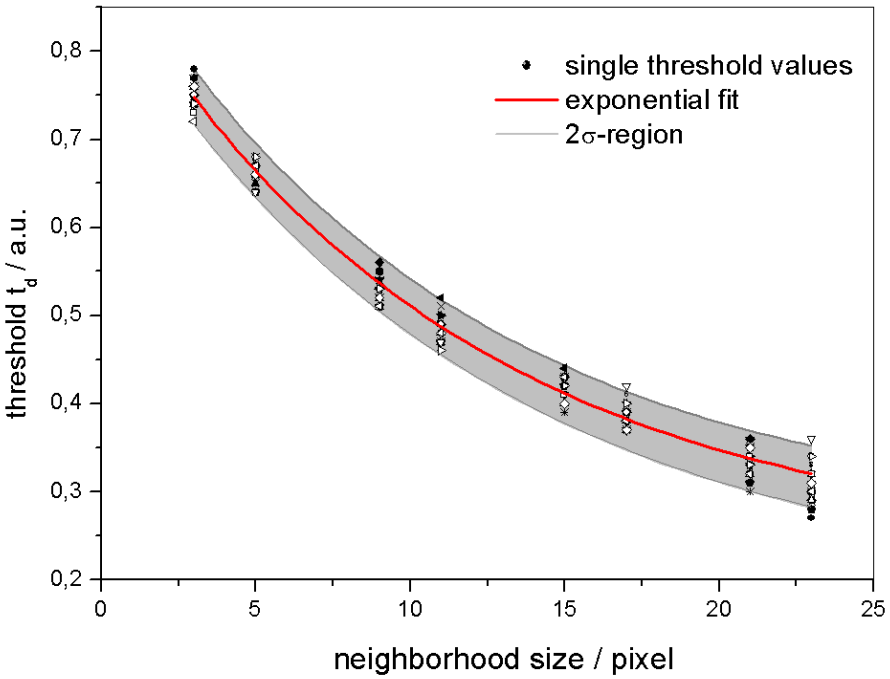

Figure 4.17: All 25 determined thresholds t_d and the 2σ-region against the neighborhood size.

The following parameters were found:

$$t^{-2\sigma}(d) = 0.18 + 0.68 \exp\left(-\frac{d}{11.83}\right) \tag{4.4}$$

$$t^{+2\sigma}(d) = 0.26 + 0.67 \exp\left(-\frac{d}{11.32}\right) \tag{4.5}$$

After the determination of the neighborhood size depending thresholds, the fusion into a height map can take place. First all the surface points generated with the smallest possible neighborhood size are transferred if their confidence value lies above the respective threshold. The procedure is repeated for the remaining points while increasing the neighborhood size. In the example used here, the result of neighborhood sizes from 3 to 17 pixels can be seen in figure 4.18(a). The extracted surface is free of noise artifacts and provides a lateral resolution such that single hairs are visible. Limiting the maximal neighborhood size to 17, there are still points (indicated black in figure 4.18(a)), where no reliable surface values were detected. Most of them belong to the background, but some are found in the central part of the face. Examining the height map and confidence image

[2] theoretically 95%

originating from a neighborhood size of 17 in figures 4.13(c) and 4.13(d), one can see that although the confidence value lies under the specified threshold, most of the surface point present a good approximation to the real surface.

In order to fill these gaps, a height map based on the one generated with the maximal neighborhood size is created by weighted averaging. The confidence values thereby act as weights.

$$h_{av}(x,y) = \frac{\sum_{(\xi,\eta)\in U_{d_{max}}(x,y)} \left[c_{d_{max}}(\xi,\eta) \cdot h_{d_{max}}(\xi,\eta)\right]}{\sum_{(\xi,\eta)\in U_{d_{max}}(x,y)} \left[c_{d_{max}}(\xi,\eta)\right]} \tag{4.6}$$

Since it is highly desired that the model is already separated from the background and this step needs not to be done by hand, only selected undefined points should be filled, the remaining background points should be set to the maximal height value. To perform this selection conveniently, a binary mask showing the undefined surface points is created which can easily be modified in the desired way and used for selective filling. Figure 4.18(b) shows the result, figure 4.19(a) the created binary mask, in which reliable values are displayed in white, the others in black. The manually modified binary mask can be seen in 4.19(b), where the white part is to be filled with the averaged height map, provided that the points were black in the first place. Black points are to be set to the maximum height map value as background.

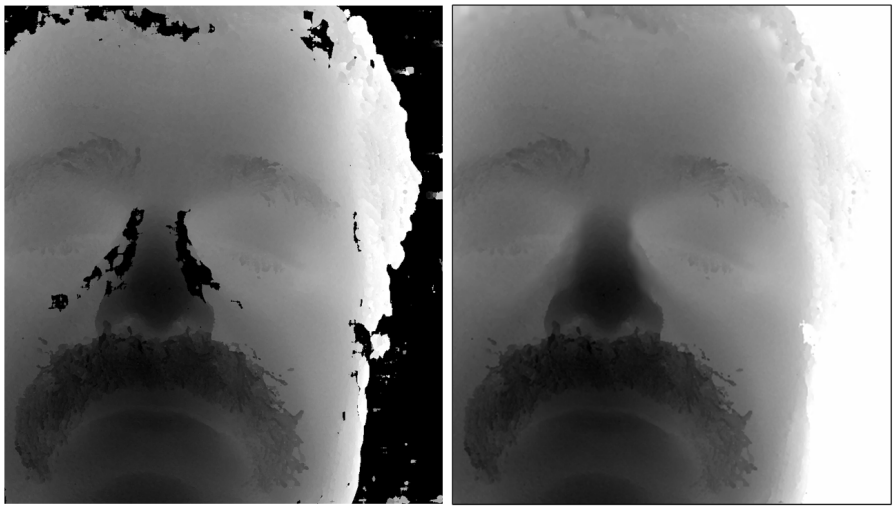

Figure 4.18: Height map gained from the multi-scale approach (a). The black points indicate regions, where no reliable surface points were found. (b) displays the final result, where the background points are set to the maximal height map value and the undefined points inside the face were filled with values created by weighted averaging.

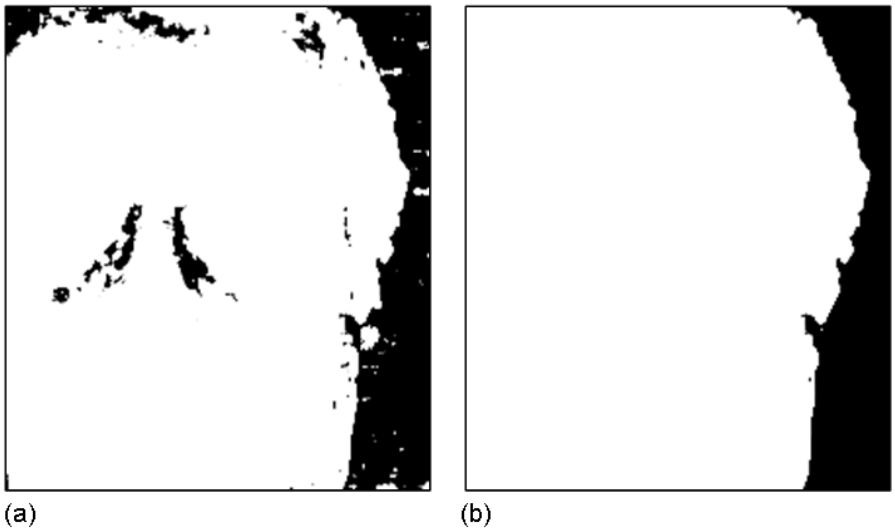

Figure 4.19: Automatically generated binary mask where reliably surface points are shown in white, the others in black (a). (b) shows the manually modified binary mask where the white part is to be filled with the averaged height map, provided that the points were black in the first place. Black points are to be set to the maximum height map value as background.

A summary of the whole process can be seen in figure 4.20 as flow chart. After the generation of the height maps with the desired neighborhood sizes, the average threshold t_d and the 2σ-thresholds according to equations 4.3, 4.4, and 4.5 are calculated. Based on them, three different height maps are generated together with the corresponding binary masks. Normally, the height map generated with the averaged thresholds bears no noise artifacts with the optimal lateral resolution. Only in the case of low quality reconstructions some noise artifacts may be present and the more conservative $+2\sigma$ threshold should be used while for an outstanding high quality reconstructions the lateral resolution can be increased further by choosing the -2σ threshold. After choosing, the corresponding binary image has to be modified as described above. This and the selection of the best suitable height map out of the three are the only steps to be performed with user interaction. All other steps are done automatically. At last the resulting height map is generated on the basis of the modified binary mask.

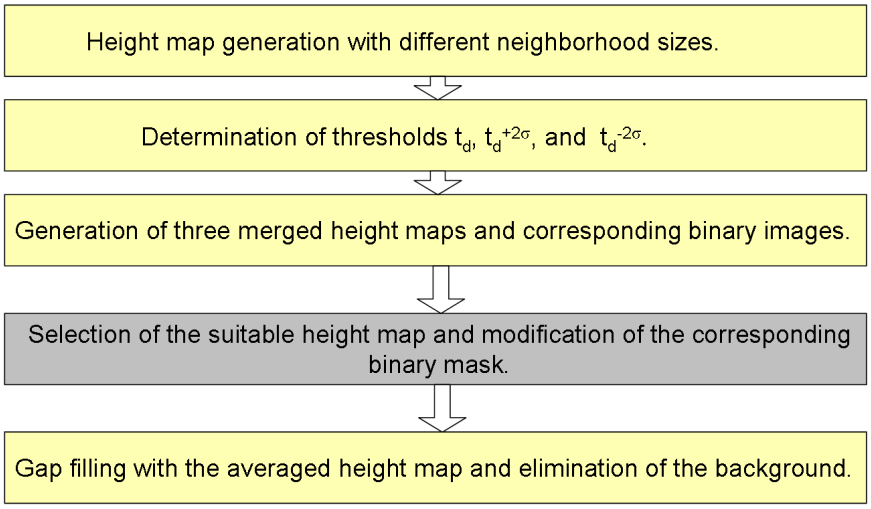

Figure 4.20: Flowchart of the newly developed multi-scale approach for adaptive choice of neighborhood size. Only the step marked in gray needs user interaction. All the other steps are automated.

4.2.5 Results

Using only one neighborhood size, one has to compromise between robustness against noise and lateral resolution. Figure 4.21(a) shows a screenshot of a three-dimensional computer model of the holographically recorded face used as example in this section. A neighborhood size of 11 was chosen. At the side of the nose, noise artifacts are still present.

Part (b) of figure 4.21 shows the result of the new algorithm with adaptive neighborhood size selection. It is free of noise artifacts and shows a higher lateral resolution, which is most obvious at the eyebrows and the mustache.

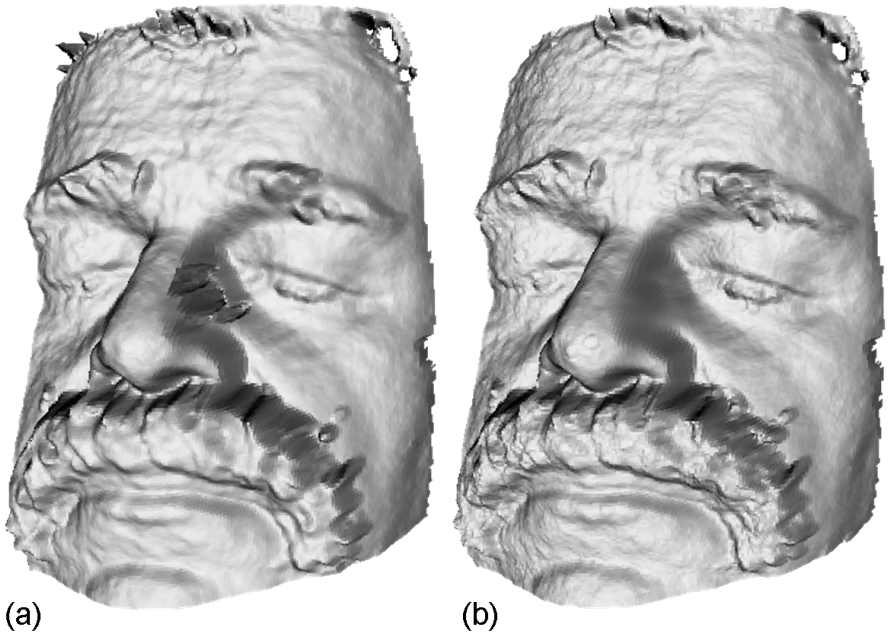

Figure 4.21: Screenshot of the three-dimensional computer model of a holographically recorded face generated with a fixed neighborhood size of 11 pixels (a) and the newly developed algorithm with adaptive neighborhood size selection (b). (b) shows a higher lateral resolution and less noise artifacts.

Another qualitative result can be seen in figure 4.22. It shows in (a) a section of the texture extracted from the holographic data, digitized with an exceptional high resolution[3]. The red rectangle marks the section displayed as three-dimensional model in figure 4.22(b). Single tiny hairs are visible, as is indicated by the green arrow.

[3]Instead of the usual 150 dpi, 600 dpi was used, which corresponds to a pixel size of 0.04 mm instead of 0.17 mm

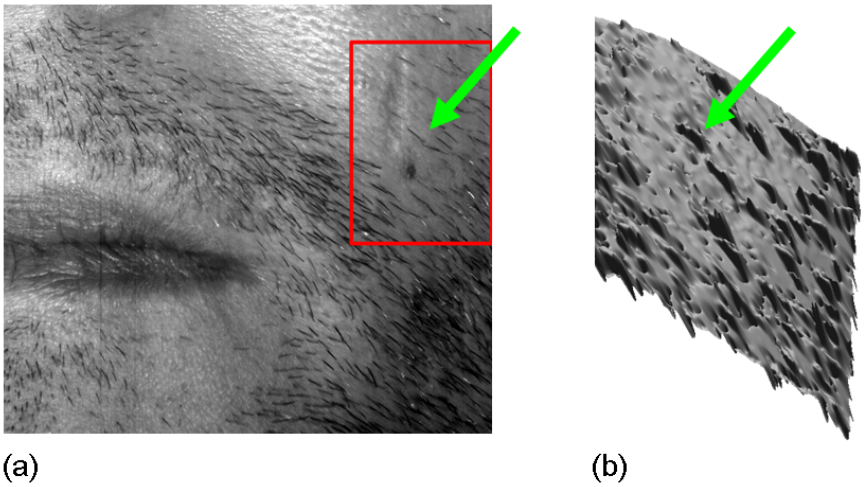

Figure 4.22: Section of the texture extracted out of holographic data digitized with a lateral resolution of 0.04 mm per pixel. A screenshot of the three-dimensional computer model of the red marked section is shown in (b). Tiny single hairs are visible, as is indicated by the green arrow.

In order to quantify the results described above, the adaptive algorithm was applied to simulated data with individually calculated PSF-directions and added noise of a standard deviation of two intensity values. The regarded neighborhood sizes were 5, 9, 11, 15, 17, 21, and 23 pixels. The following table shows the mean absolute deviation δ between the reference height map and the generated height maps with the single neighborhood sizes d mentioned above.

d / pixel	δ / mm
5	5.65
9	1.96
11	1.31
15	0.81
17	0.70
21	0.69
23	0.71

The adaptive approach provided a mean absolute deviation of $\delta_{adaptive} = 0.70$, slightly above the deviation for a fixed neighborhood size of 21 pixels. This on first sight negative result can be ascribed to two causes. First, the reference height map did not show a huge amount of small features, like single hairs. Thus the advantage cannot come to its full potential. Second, reliable surface points localized by small neighborhoods are not effected by noise in a way that huge deviations occur as in the case for unreliable points, but they nevertheless show small fluctuations due to noisy focus profiles. This fluctuations can be dealt with as will be demonstrated in chapter 5 and 6. If one smoothes the surface by using a median and a mean filter with the same parameters for both height maps, the mean deviation changes as follows:

$$\delta_{d=21} = 0.46\ mm \tag{4.7}$$

$$\delta_{adaptive} = 0.38\ mm \tag{4.8}$$

Obviously, smoothing effects the adaptively generated height map in a more positive way than it does for the height map created with a fixed neighborhood size. If the little noise fluctuations are taken care of, the full advantage of the adaptive algorithm comes out.

4.3 Summary

This chapter dealt with optimizing contrast measurement with regard to the used focus criterion and the neighborhood size.

In the first section 4.1 twelve different mathematical operators for contrast measurement were described and evaluated with the aid of a simulated holographic data set. The dependence of their performance on the PSF-geometry as well as the noise level was investigated. It was found that for real holographic data sets created with the experimental setup described in section 2.2 the newly developed XSML operator performed best. It is an advancement of the established SML operator towards a more stable behavior in the presence of noise.

In the following section 4.2 the choice of an optimal neighborhood size was discussed. If using a single neighborhood size, one always has to compromise between lateral resolution and robustness against noise. This can be avoided by adapting the neighborhood size to the contrast properties of the texture. Since such an algorithm from microscopic depth-from-focus did not work well for the extraction of facial models, a modification was created based on an especially defined confidence value. The result was a new algorithm choosing the optimal neighborhood size automatically with minimal user interaction. The resulting models have a lateral resolution such high that single hairs are visible in the height map. The model is at the same time separated from the background.

Chapter 5

Surface extraction

In the previous chapter 4 contrast measurement with regard to the focus criterion and the neighborhood size were optimized. This chapter deals with surface extraction, which is the procedure of maximizing the previously calculated contrast values and combining them to a surface. So far the surface extraction was done by maximizing the focus value over all z-values, as described in section 2.2.7. This standard procedure is the simplest and most established way to extract the surface of the recorded object after contrast measurement.

This chapter will introduce four advanced methods for surface extraction, which may replace or improve the standard maximization procedure.

Section 5.1 presents a method maximizing the focus profile symmetry instead of the focus profile itself. Especially if large neighborhood sizes are used, this is a simple possibility to avoid artifacts caused by the stepwise planar approximation of the surface through the usage of large neighborhoods.

The maximization of the focus profile naturally produces only discrete surface points corresponding to the z-positions of the different tomograms. Section 5.2 discusses several methods for surface interpolation, so that continuous z-values are gained and a resolution below the digitization step size is reached.

Two iterative methods for surface extraction are introduced in section 5.3 and section 5.4. After a first, standard maximization step, one algorithm performs a second maximization along the PSF-directions, which are calculated for each point individually from the first surface estimation. The other algorithm uses neighborhoods for contrast measurement that are tilted so that they are adapted to the slope of the first estimate of the surface.

In contrast to this, the standard procedure works with piecewise planar neighborhoods, which produces errors especially in regions of steep slope.

These four presented methods can all be applied separately or in combination. The improvement they achieve depends strongly on the quality and the characteristics of the data used for surface extraction. In which cases they should be applied and which results can be gained with a combination of methods is discussed in section 5.5.

5.1 Symmetry maximization

In section 4.1 several criteria for the measurement of sharpness were introduced. Since they produce high values for sharp image regions, it is natural to maximize the focus profile in order to extract the surface. Up to now, this was done by simply determining the absolute maximum. While doing so, the information immanent in the characteristics of the whole focus profile is not used.

Since the focus profile is of a Gaussian shape, one way to include this information is fitting a Gaussian curve to the whole peak. Such an approach is introduced in section 5.2. In addition to the benefit of extracting surfaces not restricted to the discrete z-values of the digitization process, it also reduces the influence of noise and improves the quality of the surface, as will be shown in section 5.2.

This section presents a method which also takes the shape of the whole focus peak into account, but is much easier to implement and takes less computational effort. It is on the other hand not capable of producing z-values in between tomogram positions. It can produce high quality surfaces especially if large neighborhood sizes are used. The artifacts produced by the piecewise planar approximation of the surface through large neighborhoods leads to a special kind of artifacts, so-called plateau artifacts, that will be discussed in the following section.

5.1.1 Plateau artifacts

The contrast value of one point is calculated with the aid of a local neighborhood. Lets assume one pixel is in focus at the z-position z_1. The neighborhood of an adjacent pixel with a slightly different z-value, lets say z_2, contains at its rim also the first pixel and has therefore at z_1 an elevated focus profile. This contribution broadens the focus profile of the second pixel. On the same time the contribution of the second pixel also broadens the

profile of the first one. It can be concluded that the bigger the neighborhood size used for the calculation of the focus value, the broader is the focus profile, which agrees with the experimental observance.

If the slope of the surface is steep, single contributions do not necessarily add up to a smooth curve but appear as distinguishable parts, as can be seen in figure 5.1. If some contributions have a higher intrinsic contrast than adjacent pixels, they can become dominant and produce an absolute maximum over the whole region they contribute.

Figure 5.1 shows two focus profiles generated with a neighborhood size of 25 pixels from simulated data with individually calculated PSF-directions. The two chosen points (1 and 2) have a lateral distance of 17 pixels (2.8 mm). The genuine z-values are z_1= 29.7 mm and z_2= 33.6 mm. Determining the absolute maximum delivers an equal z-value of 32.0 mm for both profiles, whose centers are clearly shifted against each other. The equal z-value was found due to one off-focus contribution dominating the peaks. The effect leads to plateaus in the height map not corresponding to the original surface. Figure 5.2 shows an exemplary height profile. The black line shows reference and the red line the profile gained by standard maximization of the XSML focus values. Plateaus, whose size correlates with the neighborhood size, are clearly visible in the red height profile.

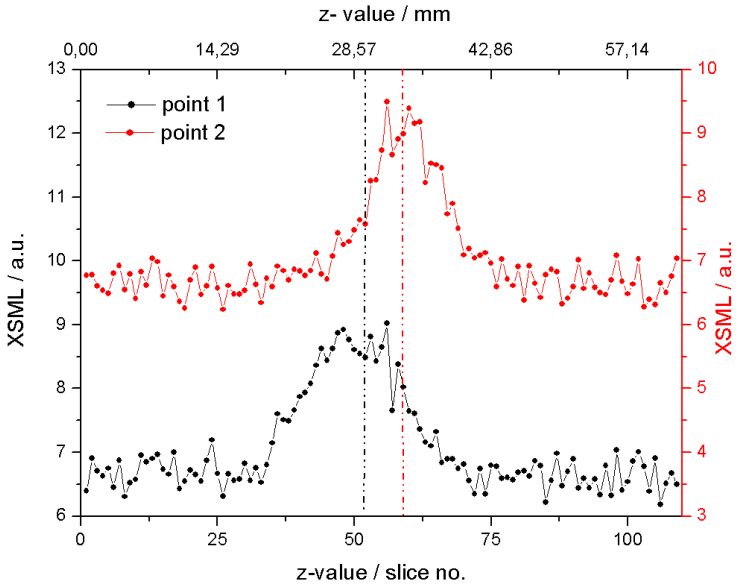

Figure 5.1: Focus profiles of adjacent points with the same absolute maximum but different peak centers. Such dominant off-focus contributions can lead to plateau artifacts. Maximizing focus profile symmetry can eliminate this influence and produce surface points at the peak centers (dashed lines) as will be demonstrated in this section.

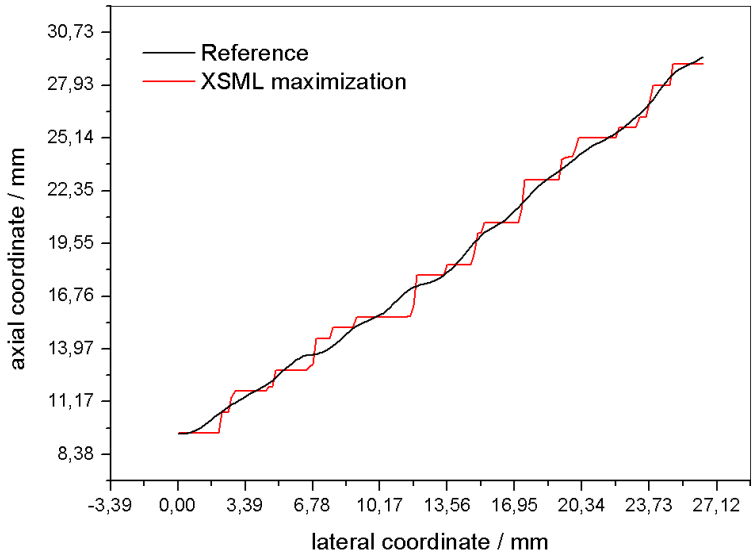

Figure 5.2: Exemplary height profile from the simulated data with individually calculated PSF-directions. The black curve shows the reference profile, the red one the surface extracted with a neighborhood size of 25 pixels. Plateau artifacts are clearly visible.

5.1.2 Focus profile symmetry

The previous remarks showed that regarding the absolute maximum only can lead to errors. Taking the whole peak profile into account, one can gain more reliable information. Figure 5.1 showed that if for example information about the peak center is regarded, plateau artifacts can be avoided. One possible solution to this is fitting a Gaussian curve to the whole peak, as will be considered in section 5.2. Another solution is evaluating the position with the maximal symmetry instead of the maximal focus value.

Symmetry is present if points at the same distance to a central point z in both directions are equal. Due to this, a symmetry measure can be defined as the absolute difference between these two values. The absolute differences are summed over several distances to the center point (see figure 5.3 for an illustration). The summation starts at the immediate neighbors of the center and goes on to a certain number B, which should be chosen in the order of magnitude of the half peak width:

$$S(z) = \sum_{i=1}^{B} |F(z-i) - F(z+i)|\,, \tag{5.1}$$

where $F(z)$ is the focus profile values. Figure 5.4 shows a exemplary peak (a) and the symmetry S defined above (b). At the peak, the S-value shows al local minimum, because the curve is symmetric and the absolute differences between corresponding points on both sides of the center are small. On the flanks of the peak this difference, and with it the

S-value, is high. The local minimum indicates the position, where the peak center is located.

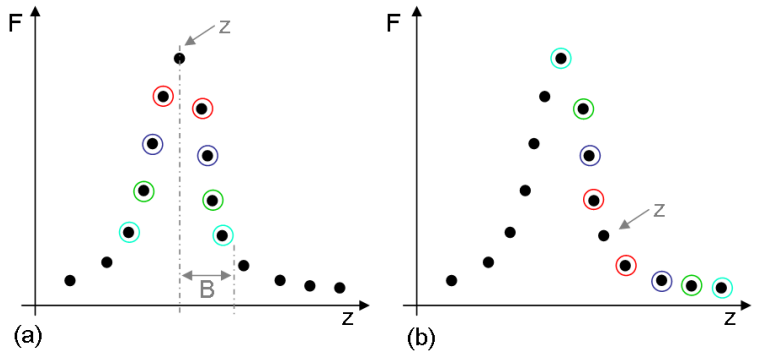

Figure 5.3: Schematic illustration of the points used for the calculation of the symmetry value. Corresponding points are marked by equally colored circles. If the center coincides with the peak center (a), S is low while it is high for chosen centers at the flank of the peak (b).

Since the minimum is only a local and in general not a global minimum, it would be easier to detect if it were transferred once again into an absolute maximum. This can be done by building the second order derivative with regard to z. To minimize the influence of noise, a second summation up to a value of B' is performed, whereupon B' should have half the size of B.

$$S''(z) = \frac{\partial^2 S}{\partial z^2} = \sum_{i=1}^{B'} (S(z-i) + S(z+i) - 2S(z)) \tag{5.2}$$

This leads to a curve as the one plotted in figure 5.4(c). For a symmetric curve, the maximum is at the same position as the one of the original focus profile, for profiles with an asymmetric maximum like the one shown in figure 5.1, the maximum is shifted to the center of the peak (see figure 5.5). In this example B was set to 20 and B' to 10. The maximal position of S" delivered a value in exact correspondence to the reference value. The same is true for the other exemplary focus profile from figure 5.1.

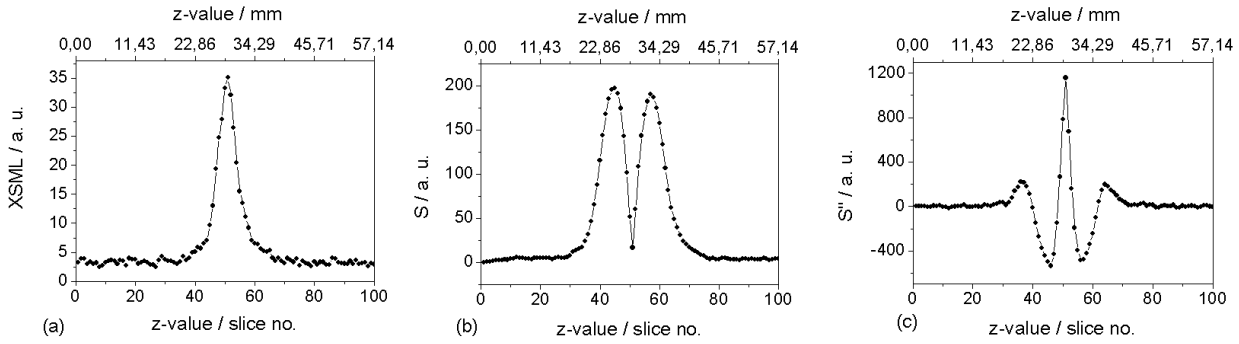

Figure 5.4: Focus profile (a), symmetry curve (b) and its modified second order derivative (c). The latter has its maximum at the same position than (a), in case of profiles with an asymmetric maximum, the maximum of S" is at the center of the peak.

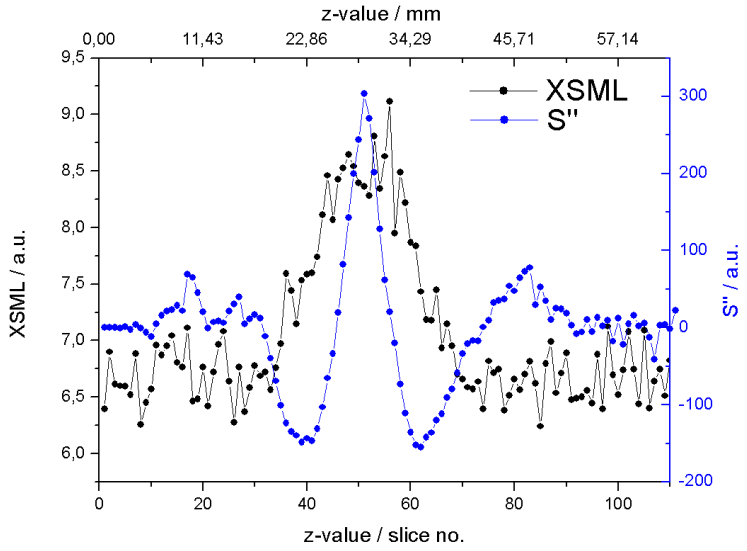

Figure 5.5: Profile from figure 5.1 combined with the S"-curve defined in the text. The S"-curve has its maximum at the center of the focus profile.

5.1.3 Results

To quantify the improvement, simulated data with individually calculated PSF-directions and added noise of one intensity value standard deviation was used.

Since plateau artifacts correlate with the neighborhood size, the dependence between the deviation from the reference surface and the neighborhood size was estimated for the standard maximization of the focus value and the maximization of S". The result can be seen in figure 5.6. For standard maximization the error decreases with increasing neighborhood sizes due to the growing robustness against noise. But for larger neighborhood sizes the error increases once again because of the additional plateau artifacts and the reduced lateral resolution. The S"-maximization performs worse for small neighborhood sizes up to nine pixels. This is caused by its higher sensitivity to noise due to the usage of the second order derivative. For larger neighborhood sizes, the S"-maximization reduces the mean absolute deviation about 0.2 mm.

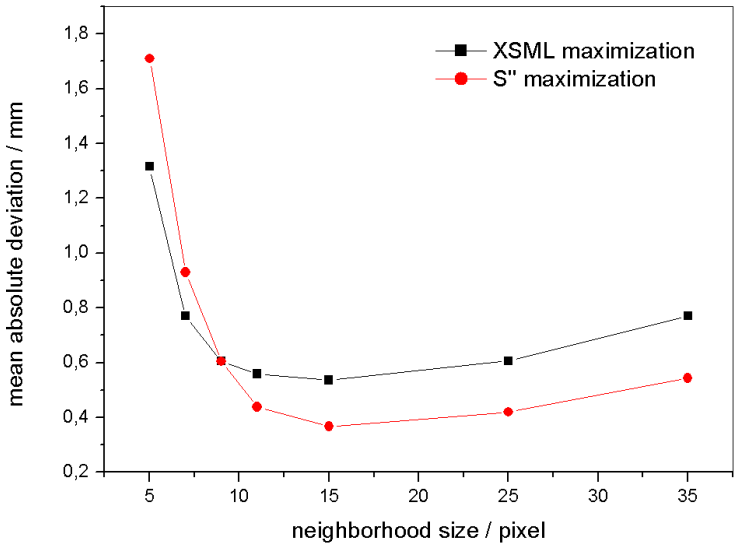

Figure 5.6: Mean absolute deviation against neighborhood size. The symmetry maximization leads to better results for larger neighborhoods.

The reduction of plateau artifacts can be seen in figure 5.7, were an exemplary height profile is displayed. The simulation parameters were the same as mentioned above while the neighborhood size was set to 25. The different profiles were shifted about 3 mm each for better comparability. The red curve, created through standard maximization of the focus profile, shows plateau artifacts. They are reduced through the usage of symmetry maximization (green curve). A Gaussian fit also reduces the plateaus and additionally delivers continuous values (blue curve, see section 5.2).

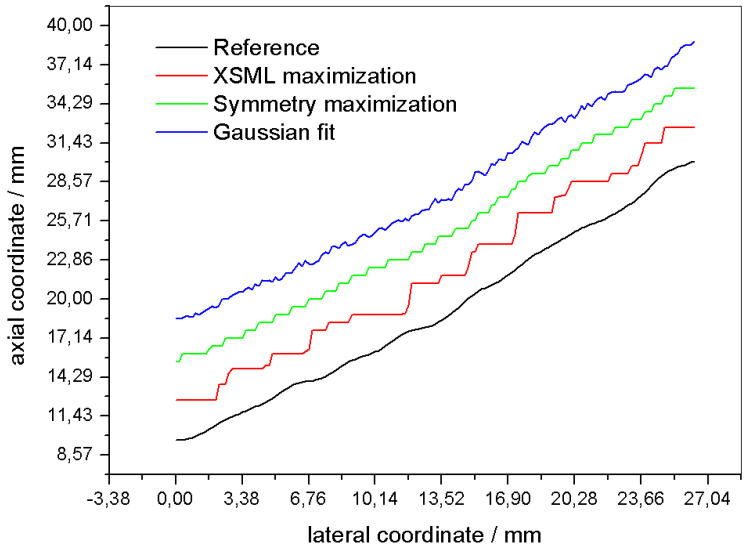

Figure 5.7: Exemplary height profile. The curves are shifted about 3 mm for better comparability. The red curve shows plateau artifacts, which are reduced through the usage of symmetry maximization (green curve). A Gaussian fit also reduces the plateaus and additionally delivers continuous values (blue curve).

Summarizing one can say that the newly invented symmetry maximization is a good tool for surface extraction if big neighborhood sizes are used. The additional computational effort, since no optimization procedure is involved, is low in comparison to the proposed fit in section 5.2. The Gaussian fit on the other hand, besides the reduction of plateau artifacts, is also capable of producing surface points not restricted to the discrete tomogram positions. This subject area is discussed in the following section.

5.2 Surface interpolation

All methods described so far extracted the surface by determining the tomogram numbers for focused image points. Naturally, all emerging z-values belong to the set of discrete z-values chosen for tomogram positions in the digitization step. Thus, the choice of the inter-tomogram distance Δz limits the maximal achievable axial resolution. For discrete z-values, it is not necessarily the case that one tomogram position coincides with the actual focus position $\bar{z}$ (see figure 5.8). The determination of surface points z_{max} through maximization is just an approximation to the real surface position.

To overcome this limitation, the estimated maximal focus value and its surrounding points can be used for the generation of inter-tomogram z-values.

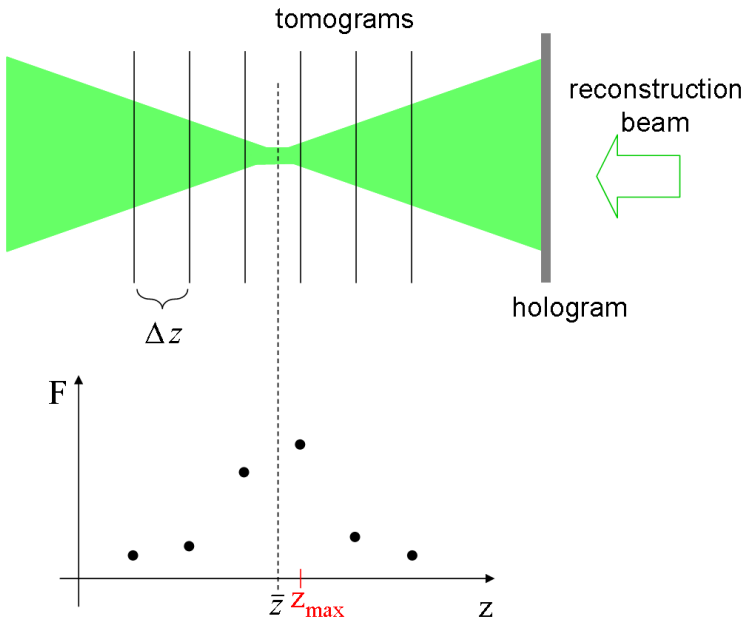

Figure 5.8: Schematic sketch of the reconstruction and digitization of one object point. The tomogram positions not necessarily coincide with the actual focus position $\bar{z}$, so that the estimated z-position z_{max} is only an approximation to the real value.

5.2.1 Methods

This section describes three different methods for surface interpolation. Two established methods, the Gaussian interpolation and parabolic regression, are introduced and the possibility of fitting a Gaussian curve to selected points of the focus profile is discussed.

Gaussian interpolation

Nayar [NN94] proposed a Gaussian distribution to interpolate the focus profile. The focus profile may be expressed as:

$$F(z) = F_p \exp\left(-\frac{(z-\bar{z})^2}{2\sigma^2}\right) \tag{5.3}$$

The interpolation is done using only three focus values F_{m-1}, F_m, and F_{m+1}, that lie on the largest peak of $F(z)$, such that $F_m \geq F_{m-1}$ and $F_m \geq F_{m+1}$ (see figure 5.9).

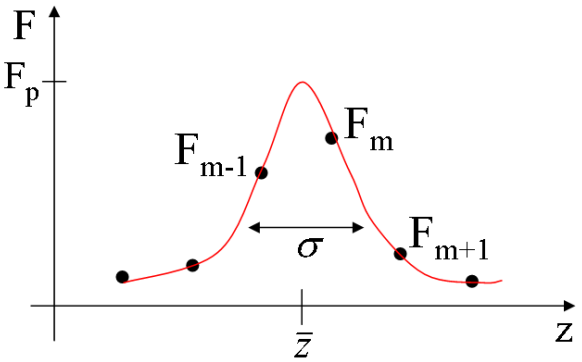

Figure 5.9: Highest peak of the focus profile with the maximal focus value F_m and its two immediate neighbors F_{m-1} and F_{m+1}. The Gaussian curve is characterized by the three parameters $\bar{z}$, F_p, and σ.

Using natural logarithm equation 5.3 can be transferred to

$$\ln F = \ln F_p - \left(\frac{(z-\bar{z})^2}{2\sigma^2}\right). \tag{5.4}$$

By inserting each of the three values F_{m-1}, F_m, and F_{m+1}, and their corresponding z-positions z_{m-1}, z_m, and z_{m+1}, one obtains a set of three equations that can be solved for $\bar{z}$, σ and F_p:

$$\begin{aligned}
\bar{z} &= \frac{(\ln F_m - \ln F_{m+1})\left(z_m^2 - z_{m-1}^2\right)}{2\Delta z\left((\ln F_m - \ln F_{m-1}) + (\ln F_m - \ln F_{m+1})\right)} \\
&\quad - \frac{(\ln F_m - \ln F_{m-1})\left(z_m^2 - z_{m+1}^2\right)}{2\Delta z\left((\ln F_m - \ln F_{m-1}) + (\ln F_m - \ln F_{m+1})\right)} \qquad (5.5) \\
\sigma^2 &= -\frac{\left(z_m^2 - z_{m-1}^2\right) + \left(z_m^2 - z_{m+1}^2\right)}{2\left((\ln F_m - \ln F_{m-1}) + (\ln F_m - \ln F_{m+1})\right)} \qquad (5.6) \\
F_p &= F_m / \exp\left(-\frac{(z_m - \bar{z})^2}{2\sigma^2}\right) \qquad (5.7)
\end{aligned}$$

Thus the determination of the real focus position $\bar{z}$ can be done simply by evaluating equation 5.5 with three focus values and their corresponding z-positions.

Parabolic regression

Scheuermann [SWG99] proposed an alternative method, used as well by Bürkle [Bür04]. He approximated the focus profile with a parabola (see figure 5.10):

$$F(z) = a_0 z^2 + a_1 z + a_2. \tag{5.8}$$

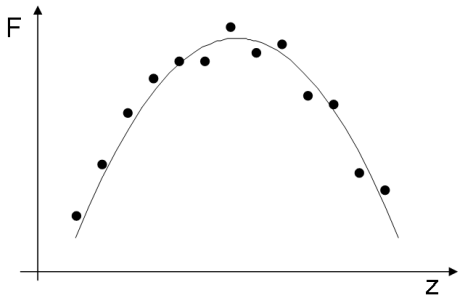

Figure 5.10: Sketch of the focus profile peak and a parabolic regression.

The maximum position can be calculated easily if the coefficients are known using:

$$\bar{z} = -\frac{a_1}{2a_0} \tag{5.9}$$

Such a regression parabola can be calculated by solving an algebraic equation and does not require a numerical optimization method, which reduces the computational effort. For a given set of points $\{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, the optimal parameters $\vec{a} = (a_0, a_1, a_2)$ for a parabola can be determined through:

$$\vec{a}^{\top} = \left(H^{\top} H\right)^{-1} H\vec{b}, \tag{5.10}$$

where H and $\vec{b}$ are defined as follows [Bür04]:

$$H = \begin{pmatrix} x_1^2 & x_1 & 1 \\ \vdots & \vdots & \vdots \\ x_n^2 & x_n & 1 \end{pmatrix} \quad \text{and} \quad \vec{b} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}. \tag{5.11}$$

By solving equation 5.10, the optimal parameters and with them the central position of the peak are calculated and continuous z-values are achieved.

Gaussian fit

An alternative approach is to fit the Gaussian model mentioned in equation 5.3 to a particular chosen number of points of the focus profile. As far as I know, this is the first

time that the usage of a Gaussian fit is reported for the determination of inter-tomogram z-values.

The first question to deal with when applying a fit is the choice of the used data points. The more data points are included in the fit, the higher is the robustness against noise. However, it should be avoided that too many points are far from the maximum, where the shape of the profile is not Gaussian anymore. It was found optimal to first determine the height of the maximum above the average value and consider only points in its upper half. Additionally, the points have to lie in an interval of 20 tomograms before and after the maximal value, to eliminate noisy data points coincidentally lying above the threshold but far away from the maximum. Choosing the fit points like this, the number of points vary with the quality of the focus profile.

Figure 5.11 illustrates an exemplary focus profile, the blue line indicates the mean focus value which is used for the calculation of the threshold (dashed line). The red brackets symbolize the additional axial restriction.

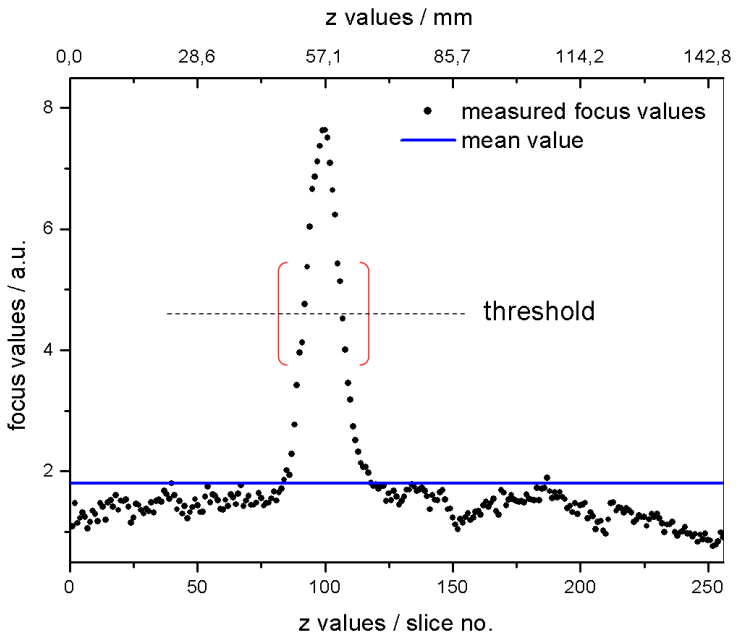

Figure 5.11: Exemplary focus profile. The blue line shows the mean focus value, the dashed line represents a thresholds and the red brackets form an axial boundary around the maximum for points to be considered.

The fitting algorithm was implemented using a downhill simplex method in multi-dimensions [PTVF02, p. 408] proposed by Nelder and Mead [NM65]. Figure 5.12(a) shows the whole profile with the fitted curve, 5.12(b) magnifies the maximum. The deviation between the z-position found by simple maximization and the one from the fit amounts to 0.45 mm.

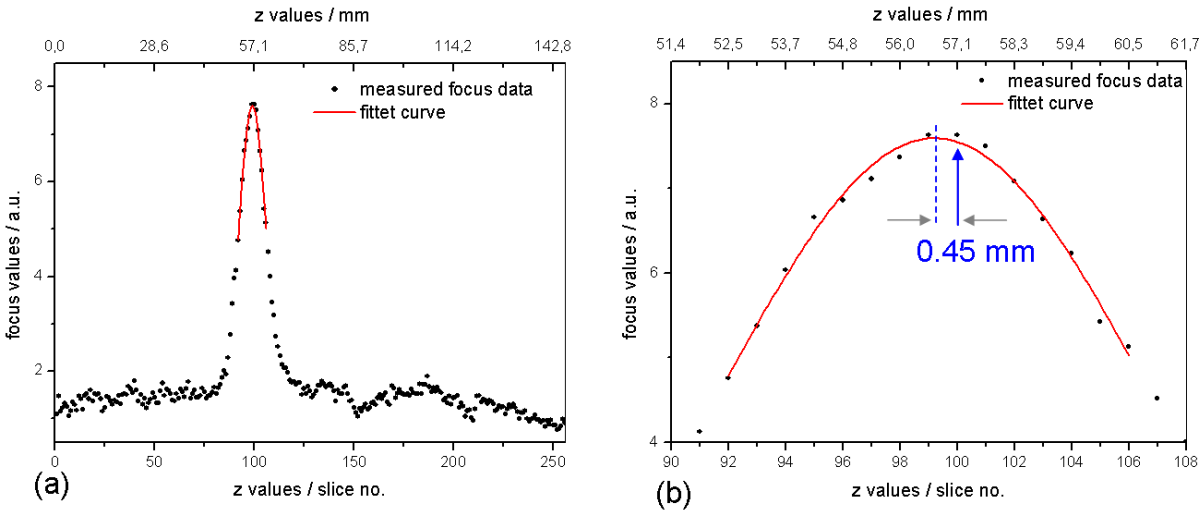

Figure 5.12: Whole focus profile from figure 5.11 with the fitted curve (a) and magnification of the peak region (b). The blue arrow indicates the z-position found by maximization, the blue dashed line the value $\bar{z}$ determined by the fit. The difference amounts to 0.45 mm.

It was found that for points with an extremely small signal-to-noise ratio, the fit produces z-values lying outside the digitized z-range. If only a small number of points are used for the fit, it can deliver arbitrary results, as can be seen in figure 5.13(a). Only the four points marked with the red ellipse are used for fitting. No reliable surface point could be found. If one smoothes the focus profile in advance with a binomial filter (see appendix B), the result could be improved, as can be seen in figure 5.13(b).

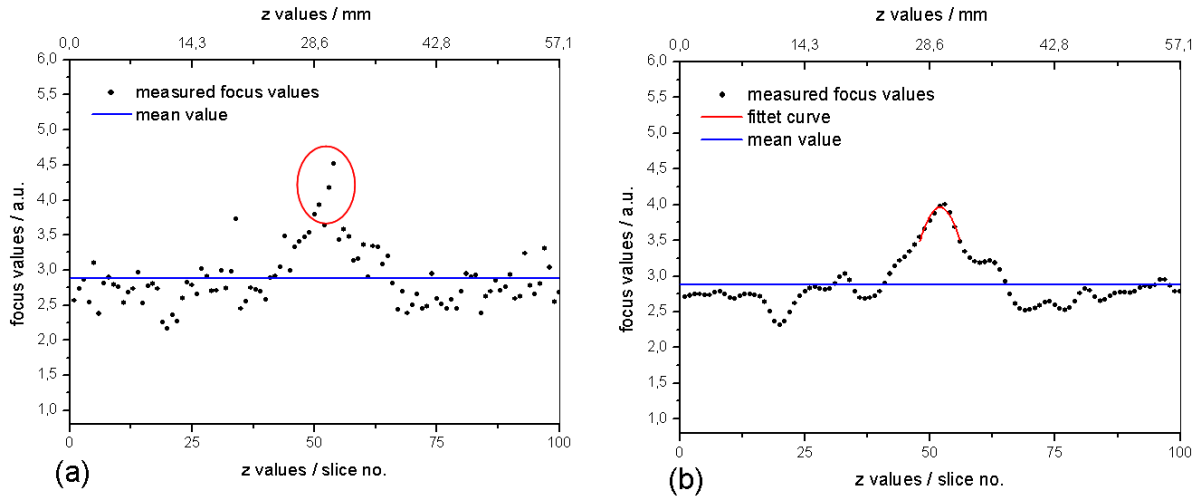

Figure 5.13: Focus profile with a low signal-to-noise ratio (a). Only the four marked points are used for fitting. Part (b) shows a smoothed version of the same profile, where the fit was able to produce a realistic z-value.

Since the smoothing operation needs computation and reliable focus profiles should not be manipulated before fitting, only noisy profiles should be smoothed. For the decision of the smoothing parameters, once again the the confidence value defined in section 4.2 was used. For a confidence value c above 0.5, no smoothing was applied, below 0.5 the filter order n (see appendix B for further information about the filtering process) is increased from zero ($c = 0.5$) to 30 ($c = 0$) in a linear way:

$$n = \begin{cases} / & \text{if} \quad c > 0.5 \\ 60 \cdot (0.5 - c) & \text{if} \quad c < 0.5 \end{cases} \tag{5.12}$$

5.2.2 Comparison of the different methods

The section above introduced three different methods for extracting a surface not limited to digitization positions. All three were tested with simulated data with individually calculated PSF-directions and four different noise levels. Figure 5.14 shows their mean absolute deviation against the added noise standard deviation. Additionally, the mean

absolute deviation of the surface gained by simple maximization of the XSML-values is plotted. The surface extraction was done with a constant neighborhood size of 15 pixels using the XSML focus criterion in all cases. The Gaussian interpolation introduced by Nayar was accomplished using F_m, F_{m-3}, and F_{m+3}, which was found to perform best. For the parabolic regression, following Scheuermann, an interval of eleven points centered on the maximum value was applied. The Gaussian fit was used as described in the previous section. Since the fitting problems for profiles with low signal-to-noise ratio are equal for fitting a Gaussian model and parabolic regression, the profile smoothing described above was used for both methods.

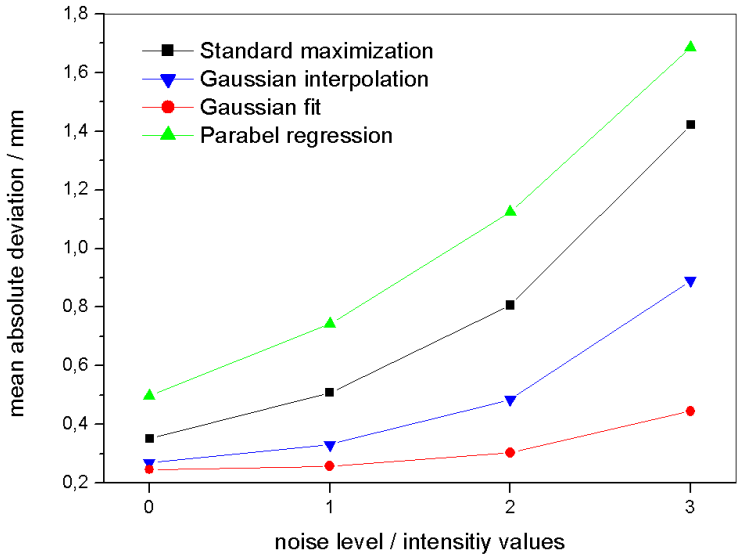

Figure 5.14: Mean absolute deviation against the noise level for the three different methods obtaining inter-tomogram values and the standard maximization method.

Obviously, the parabolic regression performs worst. Scheuermann also reported about systematic errors resulting from the different shapes of the regression parabola and the focus profile that leads to a systematic deviation from the object shape. At several points the calculated parabola has a center lying outside the digitization range. A systematic error due to the used model does not occur in the remaining two methods using the Gaussian model as basis.

The Gaussian fit performs better than the interpolation method. The latter is based upon three focus values exactly met by the curve, while the fit method takes up to twenty points into account and finds the curve by minimizing the overall deviation. That makes the fit more accurate and more stable in the presence of noise.

Since the optimization during the fit is an iterative procedure and the average number of

iterations for a holographic data set of typical parameters is approximately 140, the computational effort is much higher than that for the simple calculation in the interpolation approach. It was found experimentally, that the computation time for the fit is approximately four times as high as the one for the surface extraction by simple maximization. In contrast to this, the additional computation time for the interpolation approach can be neglected. It should be mentioned that faster algorithms as the downhill simplex could be implemented, like for example Powells direction set method [PTVF02, p. 412], which is surely faster but more difficult to implement.

Ultimately, the kind of application determines the best approach. In the case of hologram tomography, where the computation is not as relevant as for example in real time applications, the Gaussian fit approach is optimal since it produces a more accurate surface. If computation time is to be minimized, the Gaussian interpolation is a good compromise.

5.2.3 Improvement through the Gaussian fit

Besides the effect of creating surface points in between the digitization positions, this approach has additional positive effects on the extracted surface, as will be demonstrated in this section.

Figure 5.15 shows screenshots of three-dimensional computer models without any post-processing like smoothing generated with a neighborhood size of nine pixels from real holographic data with and without Gaussian Fit. The surface improvement through the elimination of the quantization steps is clearly visible.

In order to analyze the influence of the fit on the quality of the surface, the expected improvement through the avoidance of the discrete steps is estimated.

Restricting the values for an extracted surface of an originally continuous model to discrete values, a quantization error cannot be avoided.

In order to estimate the theoretical quantization error δ, a planar surface is assumed. For a continuous value z lying in the interval

$$z \in \left[z_0 - \frac{1}{2}\Delta z, z_0 + \frac{1}{2}\Delta z \right] \tag{5.13}$$

around a discrete value z_0 the quantization error is $|z - z_0|$. In order to calculate the mean error one has to integrate over all z-values in the interval mentioned above:

$$\delta = \frac{1}{\Delta z} \left(\int_{z_0 - \frac{1}{2}\Delta z}^{z_0 + \frac{1}{2}\Delta z} |z - z_0| \, \mathrm{d}z \right) \tag{5.14}$$

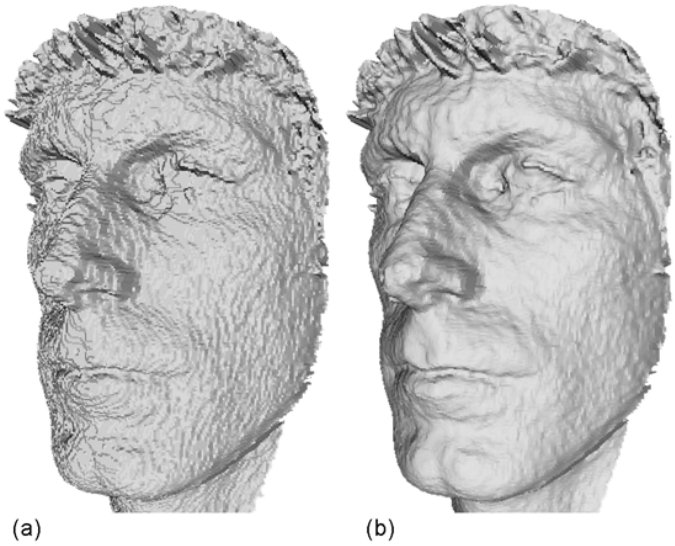

Figure 5.15: Screenshot of a three-dimensional computer model with (b) and without Gaussian fit (a). The step like structure in (a) is not present in (b).

The integral can be simplified to:

$$\delta = \frac{1}{\Delta z}\left(\int_{z_0-\frac{1}{2}\Delta z}^{z_0}(z_0-z)\,\mathrm{d}z+\int_{z_0}^{z_0+\frac{1}{2}\Delta z}(z-z_0)\,\mathrm{d}z\right) \tag{5.15}$$

$$= \frac{1}{\Delta z}\left(\left[z_0 z-\frac{z^2}{2}\right]_{z_0-\frac{1}{2}\Delta z}^{z_0}+\left[\frac{z^2}{2}-z_0 z\right]_{z_0}^{z_0+\frac{1}{2}\Delta z}\right) \tag{5.16}$$

$$= \frac{1}{4}\Delta z \tag{5.17}$$

The average error produced by the quantization is therefore $\frac{1}{4}$ of the inter-tomogram distance. This result was verified on the continuous height map used for all the simulation in this thesis and was found to be 0.249 times the inter-tomogram distance, which corresponds to $\delta = 0.142$ mm. To calculate this value, the continuous height map was quantized simply by choosing the nearest tomogram position for every continuous point. The difference between the continuous and the quantized height map was calculated.

If no other effects occur, the expected improvement in the height maps generated by the Gaussian fit should lie around the value of $\delta = 0.142$ mm.

The improvement through the fit in comparison to the simple maximization can be seen figure 5.14. The numeric values for the deviation δ are additionally given in the following table:

Noise level	δ without Gaussian fit	δ with Gaussian fit
0	0.35	0.24
1	0.51	0.27
2	0.81	0.30
3	1.42	0.45

Obviously, especially for noisier data, the improvement in much higher than the expected 0.142 mm. It can be concluded that the Gaussian fit does more than simply determining surface points between the discrete tomogram positions. It also reduces plateau effects as will be demonstrated later on and obviously also reduces the influence of noise.

Especially if the multi-scale approach proposed in section 4.2 is used, small fluctuations due to noise are present. They can be eliminated using a Gaussian fit. Figure 5.16 shows a height profile in real data along the cheek of the proband attained with the multi-scale approach. The curve generated with the Gaussian fit shows less fluctuation due to noise than the one created with simple maximization.

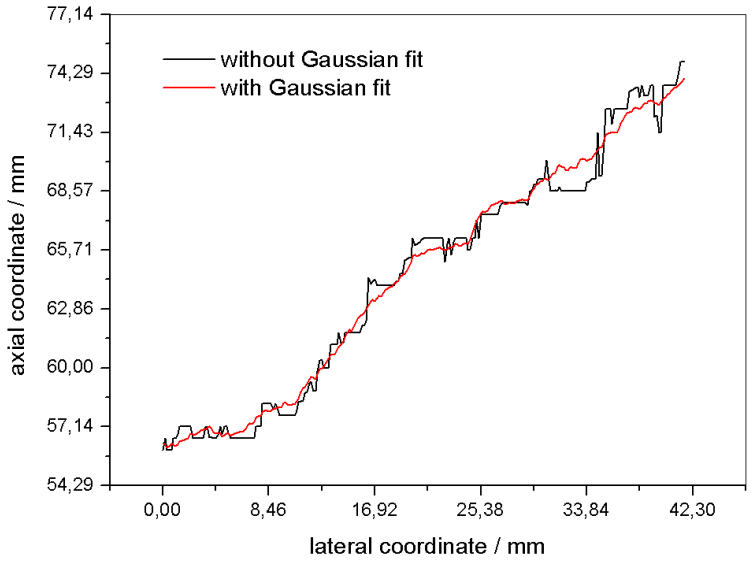

Figure 5.16: Height profile from real data along the cheek of the proband generated with and without a Gaussian fit.

Plateau artifacts were already described in section 5.1.1. If large neighborhoods are used for surface extraction, regions of steep slope tend to build plateaus of equal height values with a size in the same order of magnitude as the neighborhood size. Figure 5.7 shows the elimination of these artifacts through the usage of the Gaussian fit.

Another way to visualize the artifacts is regarding face contours, which are points of

equal height value. If plateau artifacts are present, the contours are clustered[1]. Since the Gaussian fit reduces plateau artifacts, it should also improve the shape of the contours. Figure 5.17 shows a contour created with a neighborhood size of 25 pixels through focus profile maximization (a) and with a Gaussian fit[2] (b). The contour in (b) shows less clustering than the one in (a), which is an additional proof for the reduction of plateau artifacts.

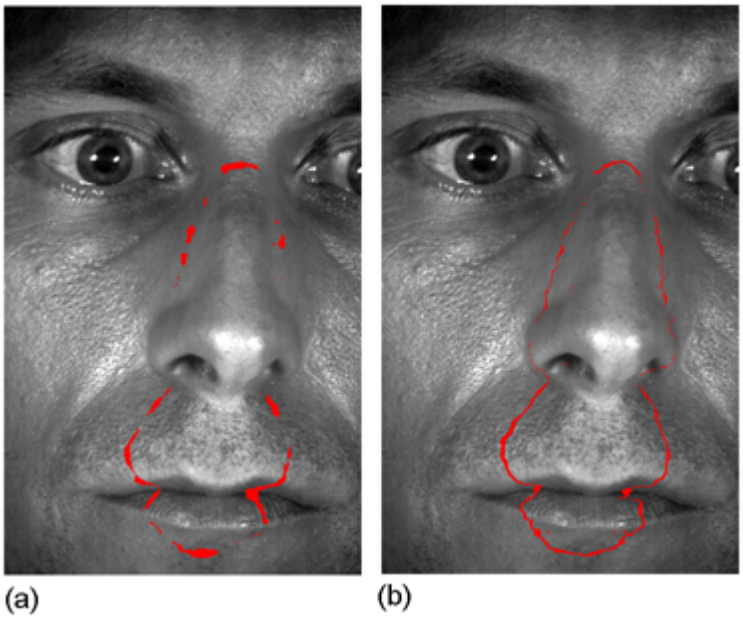

Figure 5.17: Contour created with focus maximization (a) and with Gaussian fit (b). The latter shows less plateau artifacts.

Taking this additional improvement through the Gaussian fit into account, the Gaussian fit is a computational costly but excellent tool for producing a surface not restricted to the discrete tomogram positions on the one hand and reducing the influence of noise and plateau artifacts on the other hand. It showed better performance than the methods reported in literature.

[1] Frey [Fre05] referred to the same artifacts as clustering effect.

[2] The continuous values were rounded to integers to create the contour.

5.3 Maximization along the PSF-directions

The previous two sections 5.1 and 5.2 introduced new methods for locating the maximum of a focus profile. So far, the focus profiles were regarded as series of points parallel to the z-axis. This section deals with a new method for defining the points belonging to one focus profile, which takes the characteristics of the holographic imaging process into account.

As already stated in section 3.2, the characteristics of the holographic imaging process can be described through the point spread function (PSF). Different than in microscopy, where the PSF is constant for all imaged points, this is not true for holography. The position of an imaged point in respect to the optical axis determines the direction of the PSF, which means its angle to the optical axis. The distance to the holographic plate determines the opening angle. If shadowing effects occur, the PSF is additionally characterized by the shape of the recorded object (see section 3.2).

In the procedures introduced so far, the maximization of the focus values was done by finding the maximal z-position for a fixed lateral position (x, y), which is equivalent to a maximization parallel to the z-axis, or the optical axis, respectively. Since the image formation, described by the PSF, takes places at an angle to the z-axis, errors in surface localization occur. For focus points in the off-axis region, it cannot be guaranteed that focused and unfocused images of one and the same point are compared. It is possible that unfocused images of regions with a high intrinsic contrast are compared with focused regions of lower intrinsic contrast, which leads to a false estimation of the surface. Figure 5.18 illustrates neighboring points with high and low intrinsic contrast. While maximization along the z-axis produces good results for on-axis points, for off-axis points false surface positions are located.

A similar problem is known in the depth-from-focus methods in microscopy, where it can be corrected by the use of telecentric lenses [WN95]. Such a procedure is in principle impossible for holographic imaging, because the holographic reconstruction process can be deemed to be a simultaneous imaging with different focal length for each imaged point and telecentric lenses work only for one fixed focal length.

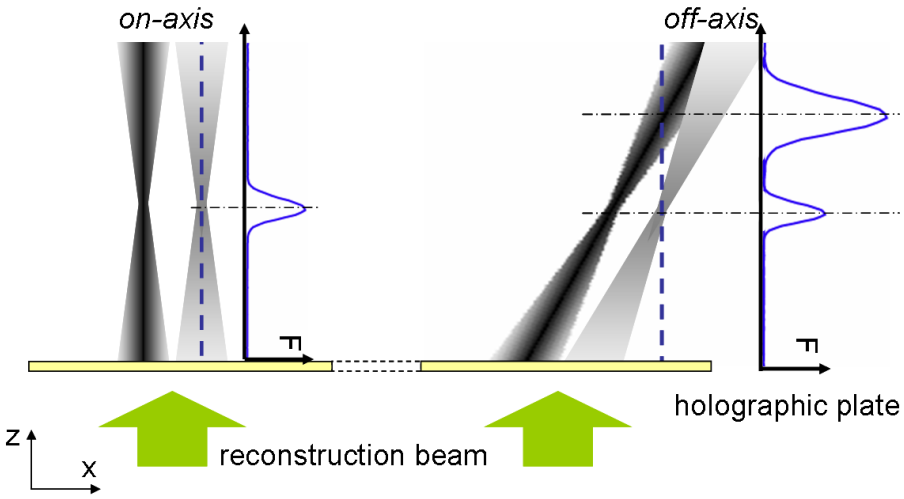

Figure 5.18: Schematic illustration of neighboring points with high and low intrinsic contrast. The surface extraction fails in the off-axis region, if it is done parallel to the optical axis.

5.3.1 Previous work

Maximizing the focus values not along the z-axis but along the PSF-direction, which is different for every image point, is the solution to the problem described above. In principle, this direction is determined through the position of the focused image point and the illumination center of the hologram. Since the position of the focused image point is not known in advance, an iterative algorithm was developed in the course of my diploma thesis [The03], [TBG+05]. Starting from a focus position estimated parallel to the z-axis, the axis of the PSF is calculated using a first guess focus position and the illumination center of the hologram for each lateral point individually. A second maximization is performed by comparing focus values along this line, where only points in a certain interval around the first guess are taken into account (see figure 5.19).

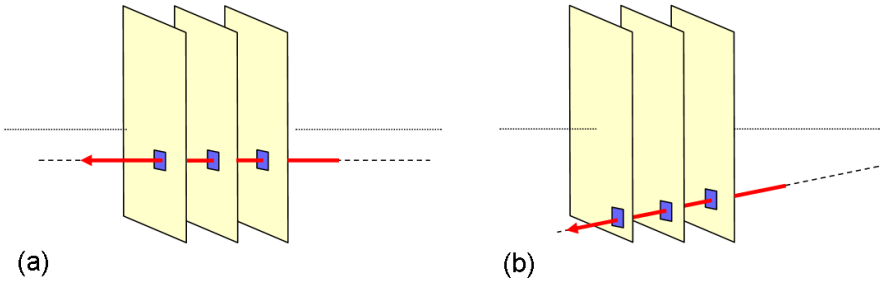

Figure 5.19: Maximization through comparison of focus values parallel to the z-axis (a) and along the iteratively determined PSF-axis (b).

The improvement of the extracted surface through this algorithm was demonstrated for test objects as well as for human faces. Nevertheless it is only an approximation to reality, since it does not take shadowing effects into account. If shadowing effects occur, the scattered light can reach only a part of the hologram, due to the geometry of the scattering surface (see section 3.2 for a detailed explanation). This is why the PSF originates only from the part of the hologram reached during the recording process. An estimation of the PSF-axis through the illumination center of the hologram is in such cases only an approximation.

The algorithm has the other drawback, that the illumination center of the hologram has to be known. It can be determined or estimated, but in each case the procedure can be a source of errors and is inconvenient.

In the following section an advancement of the procedure is introduced, which takes shadowing effects into account and does not need ab-initio knowledge about the reconstruction geometry and the illumination center of the hologram.

5.3.2 Improved algorithm

In principle, the algorithm works in the same way as the one described above, only that after the first step of maximization parallel to the z-axis, the direction of the PSF is not calculated through the first guess focus point and the illumination center of the hologram, but is extracted form the image volume itself. In order to do so, substacks are created which contain sections of the tomograms in an interval before and after the estimated z-position of the central lateral point (see figure 5.20). Since the focused contour present in the substack and its unfocused versions show a correlation in intensity, they can be used for a gray-value based registration. In the registration process the central slice is used as anchor and the adjacent slice is translated in such a way that the mean square intensity difference between the two slices is minimal. This transformed slice acts as reference for the next one until the whole substack is aligned.

An algorithm proposed by Thévenaz et. al. [TRU98], using a coarse-to-fine iterative strategy based on a modification of the Levenberg-Marquardt minimization algorithm, is used for the registration step[3].

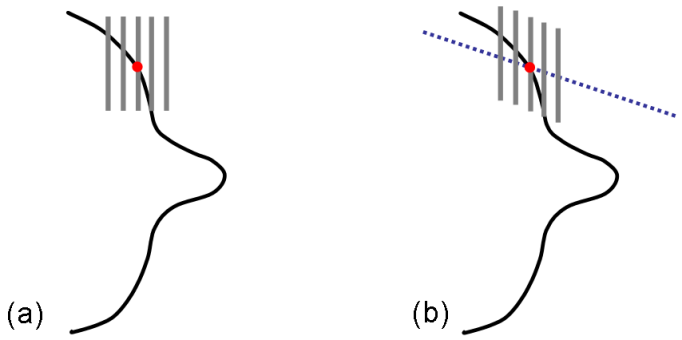

Figure 5.20: Schematic drawing of the a substack centered at a specified point at the first estimate of the surface (a). After the registration procedure, the slices of the substack are sheared in respect to the PSF-axis.

The algorithm delivers for each slice translation quantities Δx and Δy, which can be of sub-pixel precision due to the used spline representation in the algorithm. These values are averaged over all slices in the substack and used as characterization of the PSF-direction. Figure 5.21 shows exemplarily three images of one substack. The upper row displays the original tomogram sections, where the focused and defocused image of the eye is at different lateral positions. The lower row shows the aligned substack, where the slices are

[3]The algorithm was not implemented by myself but the implementation of Thévenaz et. al. [TRU98], available as ImageJ plugin *stackreg*, was used.

translated in such a way that the image of the eye is at the same lateral position. The average translation per slice determines the PSF-direction in this region of the data set.

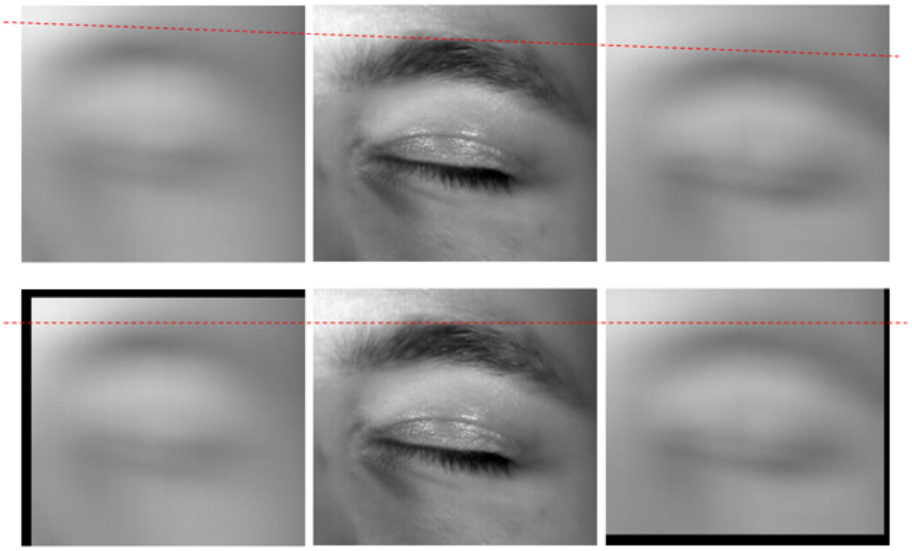

Figure 5.21: Three exemplary slices of the substack, where the one in middle is the anchor for the registration. The upper row shows the original substack, the lower row the registered version. Pixels without information due to the translation are marked in black. The dashed line connects the upper boundary of the eyebrows. In the upper row it is tilted while it is horizontal in the lower row.

To save computational time, substacks are not created for every lateral point but overlapping substacks are generated centered on a mesh with a distance of 50 pixels in each direction. The resulting directions were interpolated in a linear way, so that for each lateral point (x, y) an individual direction is defined.

Figure 5.22 summarizes the process schematically. For each point on a grid of 50×50 pixels (red dot) a substack with a specified lateral and axial size is defined around the first estimate surface point at this grid position (a) and extracted (b). The substack is aligned using a gray-value based registration (c). The PSF-direction of the center point (red dot) is defined as the resulting direction of the alignment (b). This is done for each point on the grid. For points in-between the PSF-direction is determined through linear interpolation. The maximization procedure for surface extraction is then performed along the PSF-direction.

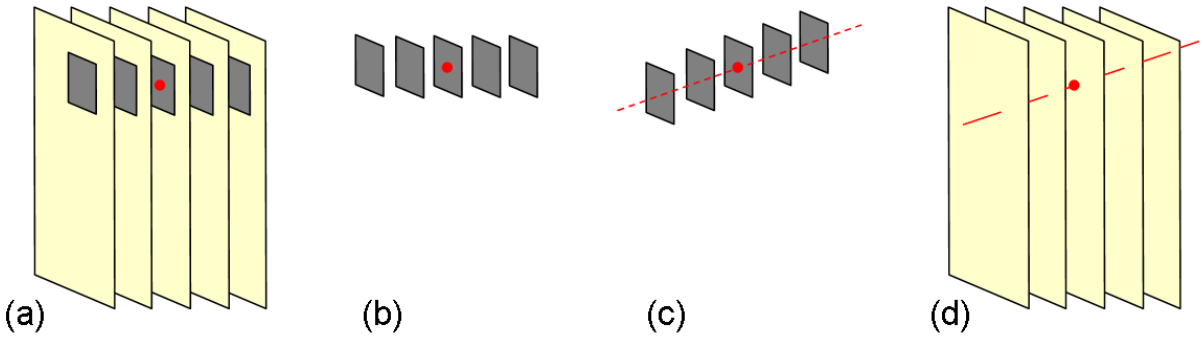

Figure 5.22: Schematic presentation of the procedure for one exemplary grid point. Around the first estimate surface point a substack is defined (a) and extracted (b). It is aligned using a gray-value based registration (c). The PSF-direction of the point is defined as the direction resulting from the alignment (d).

5.3.3 Evaluation of the PSF-directions

Before using this new algorithm it will be demonstrated that the registration strategy is able to determine the PSF-direction satisfactorily. A stack of simulated tomograms with individually calculated PSF-directions is a good reference for the evaluation of the extracted PSF-directions. Since the synthetically generated PSF-axes are directed to one central point, the mean angular deviation between the ideal and the extracted direction can be calculated. This quantity can be used for a qualitative comparison between different parameters of the algorithm like the lateral and axial substack size.

The calculated directions are illustrated through their projection into the xy-plane (see figure 5.23). The position of each slice center is projected into the xy-plane for every slice in the substack. The center substack point is drawn as a red dot. With this, the different directions can be illustrated two-dimensionally.

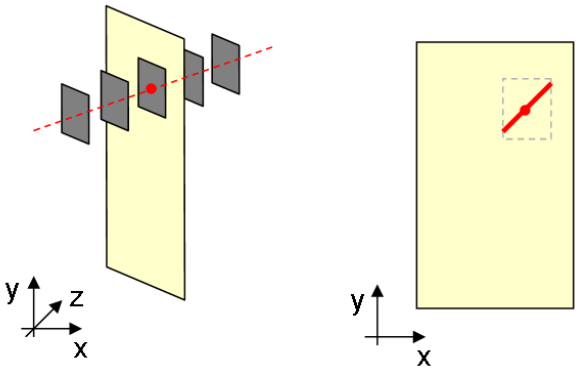

Figure 5.23: The calculated PSF-direction is illustrated through its projection into the xy-plane.

Figure 5.24(a) illustrates the extracted directions for a section of the face. It shows the projections explained above with the height map as background for better orientation. While all directions point to one center in the case of the simulated data set (figure 5.24(a)), the directions of the real data show different behaviors due to shadowing effects, as can be seen in figure 5.24(b).

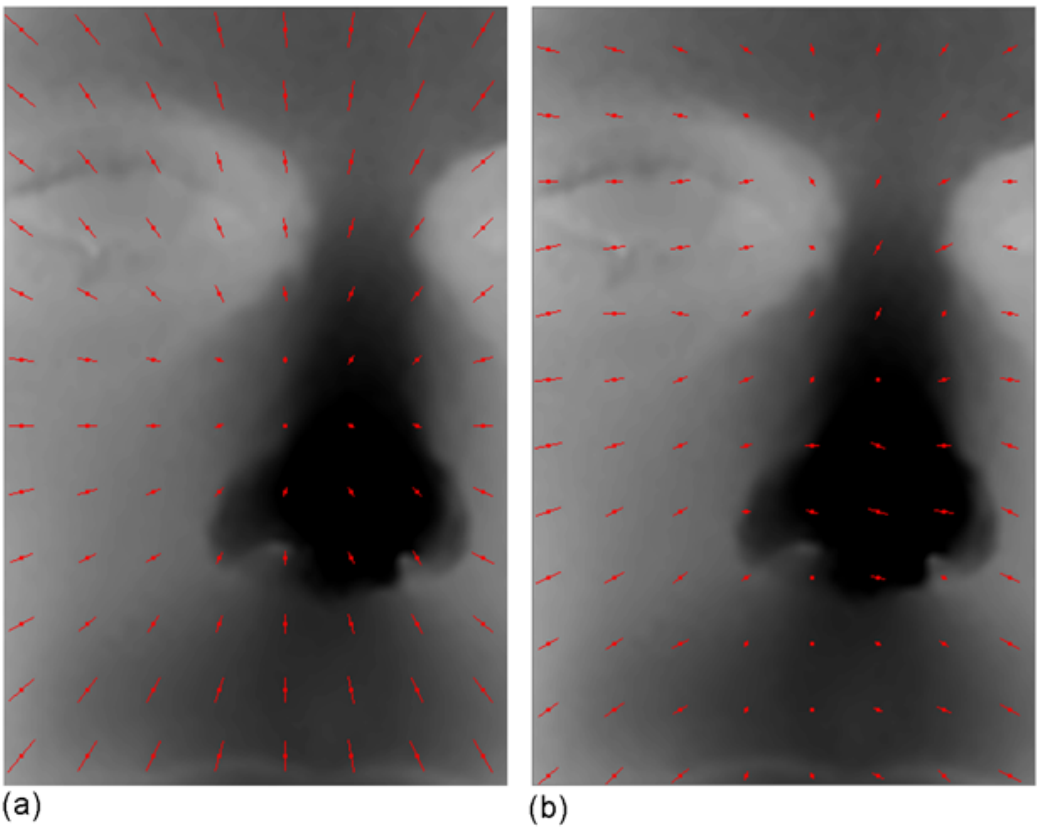

Figure 5.24: Extracted PSF-directions for simulated data (a) and real data (b). While in the first case the directions point all to one center, the real data directions show a different behavior due to shadowing effects.

The angle-component in the xy-plane of the extracted PSF-axis can be defined as:

$$\alpha = \arctan \frac{\Delta y}{\Delta x}, \tag{5.18}$$

where Δx and Δy are the average translation amounts per tomogram.

In order to optimize parameters, the mean angular deviation was calculated in two ways:

- for different lateral sizes of the substack with constant axial size
- for different axial sizes with a constant lateral size

Since the calculated deviation between different parameters was considerably small, an overall optimization was abandoned.

The results for the variation of the axial substack size can be seen in the following table:

lateral size / pixel	axial radius / no. slices	mean angular dev. / degree
150	10	5.7
150	15	5.3
150	20	5.1
150	25	5.2

An axial radius of 20 seems to be optimal. It corresponds to a substack of a total number of $2 * 20 + 1 = 41$ slices. The following table shows the results for the variation of the lateral substack size:

lateral size / pixel	axial radius / no. slices	mean angular dev. / degree
50	20	5.9
100	20	5.0
150	20	5.1
200	20	5.3

According to these values, a lateral substack-size of 100 pixels is optimal. These parameters were used during the following calculations. The mean angular deviation from the ideal direction to the direction parallel to the z-axis, which corresponds to $\Delta x = 0$ and $\Delta y = 0$, is 53 °. The mean angular deviation between the extracted and ideal directions is therefore around 10% of the deviation in case of maximization along the z-axis. Even if the PSF-direction is not met exactly, a much better approximation than using simply parallel directions is achieved.

Another possibility to illustrate the accordance between the calculated PSF-direction and the realistic ones is a horizontal cut through the real image volume (see section 3.5 for an explanation of horizontal cuts). Figure 5.25 shows two horizontal cuts through real holographic data sets at the level of the eye-reflexes (a) and the outer tip of the ear (b). The red lines indicate the extracted PSF-directions. Their correspondence to the visible part of the PSF is satisfactory.

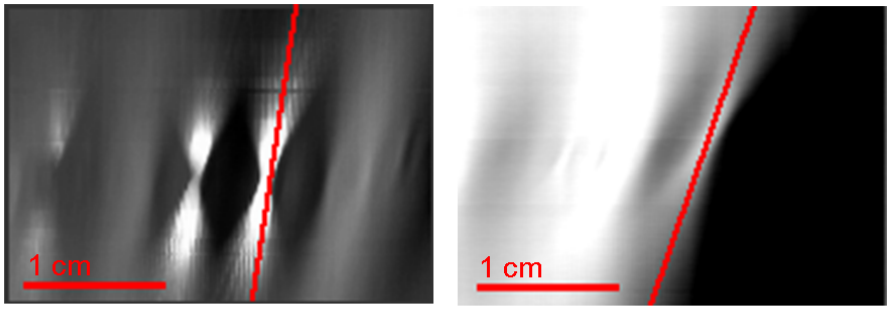

Figure 5.25: Horizontal cut through real image volumes at the reflexes in the eye (a) and the outer tip of the ear (b). The red lines indicate the extracted PSF-directions.

5.3.4 Results

After the correspondence between the calculated PSF-directions and the real directions was quantified, the algorithm was tested. It contains, in accordance with [TBG+05], the following steps:

- Surface extraction through maximization parallel to the z-axis is performed. A large neighborhood size should be used in order to avoid noise artifacts and smoothing is applied to achieve a good approximation to the real surface.
- Substacks with a lateral size of 100 × 100 pixels centered around a mesh of points with a distance of 50 pixels in each direction and 20 tomograms in the $+z$- and $-z$-direction are created and registered. The mean displacement in x- and y-direction Δx and Δy is calculated. The displacements are linearly interpolated between the mesh points, so that an individual PSF-direction is determined for each lateral point.
- For a chosen neighborhood size, the surface is extracted through maximization along the determined PSF-direction in an interval of 20 tomograms before and after the first z-value estimation.

To quantify the improvement of the extracted surface, the algorithm was once again tested on simulated data with individually calculated PSF-directions for different levels of added noise. The novel XSML-operator was used as focus measure and the results of the four neighborhood sizes 11, 13, 15, and 17 pixels were averaged (see figure 5.26). The error produced through maximization along the PSF-direction is smaller than the one through maximization parallel to the z-axis, especially for higher noise levels. This result cannot

be assigned to the other maximization direction alone, but also to the restriction of the newly found values to an interval around the initial surface.

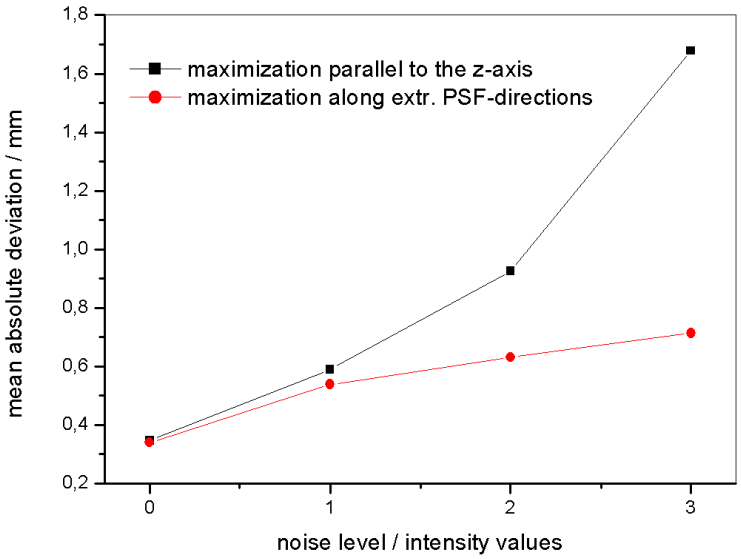

Figure 5.26: Absolute mean deviation against noise level for surfaces extracted through maximization parallel to the z-axis (black) and along the PSF-direction (red). The latter produce a smaller deviation.

Applied to realistic data, an improvement is also visible. Figure 5.27 shows for example a section of the height profile in the area of the cheek. The surface extracted through maximization along the PSF-directions is smoother and shows less artifacts than the one extracted through maximization parallel to the z-axis.

As outstanding results in the off-axis regions as reported in [TBG+05] could not be observed, which caused a detailed investigation into the differences. In [The03] and [TBG+05], the gray value variance was used as focus criterion, which is more sensitive to artifacts caused by an angle mismatch between the direction of image formation and maximization of the focus values. Since the less sensitive XSML criterion is used for the present results, the improvement is not as obvious. Figure 5.28 shows a focus profile of an off-axis point at the ear for both maximization methods and two focus criteria, the variance (a) and XSML (b). Both black curves show two maxima. The first maximum corresponds to an unfocused image of a region with high intrinsic contrast, while the second belongs to the focused point with lower intrinsic contrast. If the maximization is done parallel to the z-axis and the variance is used as focus criterion, the false maximum is dominant, in the case of the XSML criterion it is much less pronounced and nearly invisible. Maximization along the PSF-directions eliminates this false peak, which leads to an improvement of the extracted surface in the case of the used variance but to no

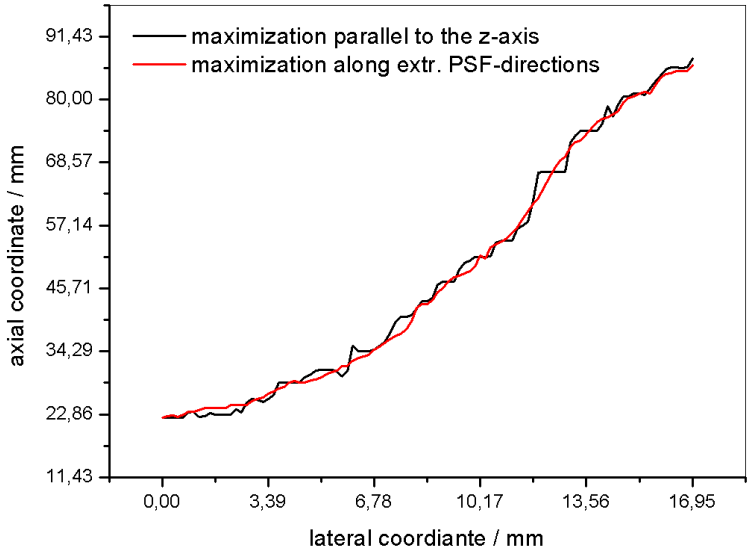

Figure 5.27: Exemplary height profile along a line crossing the lip for maximization parallel to the z-axis (black) and along the extracted PSF-directions (red). The latter is smoother and less effected by artifacts.

difference for the XSML criterion, since the found point was correct in the first place.

Summarizing one can say that the improved algorithm is capable of extracting the PSF-directions from the digitized image volume itself without ab-initio knowledge about the reconstruction geometry. At the same time, shadowing effects are accounted for, which was not the case in the algorithm presented in [The03] and [TBG+05]. The quality of the surface is improved. If for some reason focus criteria have to be used that are very sensitive for an angle mismatch between PSF-direction and the direction of maximization, like for example the gray value variance or the Tenengrad criterion, the improvement achieved through the newly developed algorithm is higher than if less sensitive criteria like the XSML are used. In the latter case, especially if the signal-to-noise ratio is relatively low, the improvement does not justify the additional computational effort.

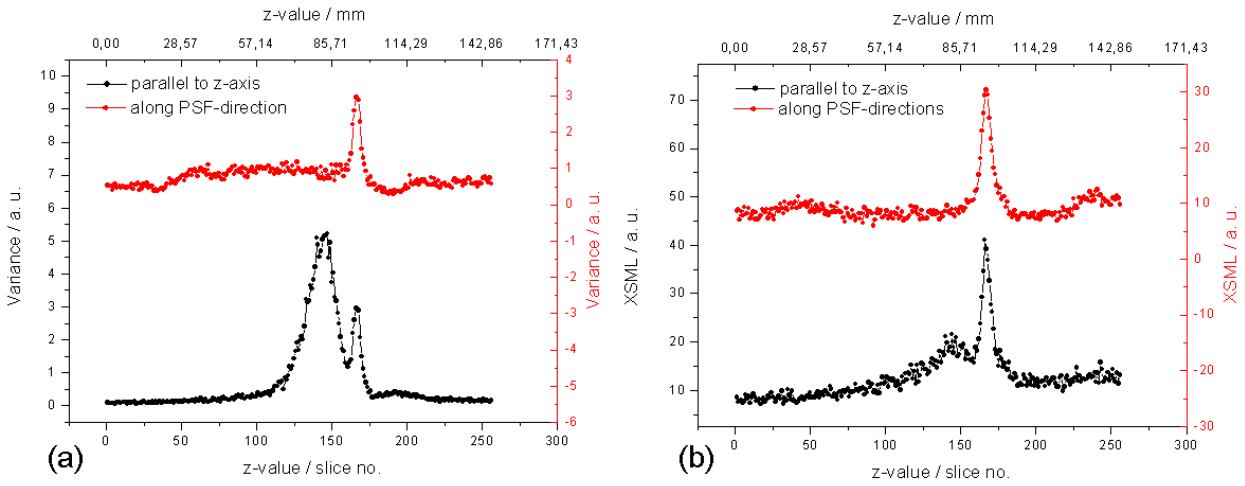

Figure 5.28: Focus profiles for an off-axis point at the ear for both maximization directions and the variance as focus criterion (a) as well as the XSML criterion (b). (a) shows a false peak, which is eliminated through maximization along the PSF-axis. In (b) the false peak is much lower in the first place, so that the improvement is less obvious.

5.4 Maximization using tilted neighborhoods

In this section another method is presented, which uses a standard maximization as a first, initial step and refines the extracted surface in an additional evaluation step. It is based on the concept of *Focused Image Surface* (FIS) invented by Subbarao et. al. [SC95] in the field of microscopic shape-from-focus. He defines the FIS as the surface created through the focused image points. The FIS is therefore, for ideal imaging, the one-to-one correspondence to the surface of the imaged object.

Image acquisition takes place with a planar image sensor and the surface detection works with local neighborhoods in the respective image planes. Following Subbarao, this corresponds to a piecewise planar approximation of the FIS, which leads naturally to errors. He proposes a piecewise linear approximation of the FIS through an iterative algorithm. A first estimate of the surface is made in the traditional way. After that, a small cubic volume centered at the initial estimate of the FIS is extracted from the image sequence. In this volume, planes of arbitrary orientation are considered and the one which contains the highest focus values is estimated through an optimization algorithm. The initial estimates of the position and orientation of the FIS are used as starting values during the search, which is done by gradient descent [SC95]. The found planes are combined and build a piecewise linear approximation of the FIS. Choi improved the method through the usage of Lagrange polynomials as approximation of the FIS instead of planes [CY00] and Asif improved the optimization process through the usage of neural networks [AC01]. Recently it was modified by Ahmad in such a way that the search is performed over the whole image volume and not only over small cubes [AC05].

Since the optimization process is time consuming and sometimes unstable in the presence of noise, an alternative approach based on the same principle is proposed in this thesis.

Surface regions with a steep slope relative to the holographic plate are challenging in hologram tomography. Due to their angle, only a small amount of scattered light reaches the hologram and they suffer from shadowing effects. Additionally, the piecewise planar approximation through the selection of planar neighborhoods effects them more than other regions. The focus profiles show in general a low signal-to-noise ratio. Figure 5.29 shows exemplarily two focus profiles, one from the cheek (black), the other one from the side of the nose (red) of the simulated data with parallel PSF-directions. Since shadowing effects etc. are not accounted for in the simulation, the tremendous difference in the peak height is only caused by the very narrow focused contours in the case of steep slopes.

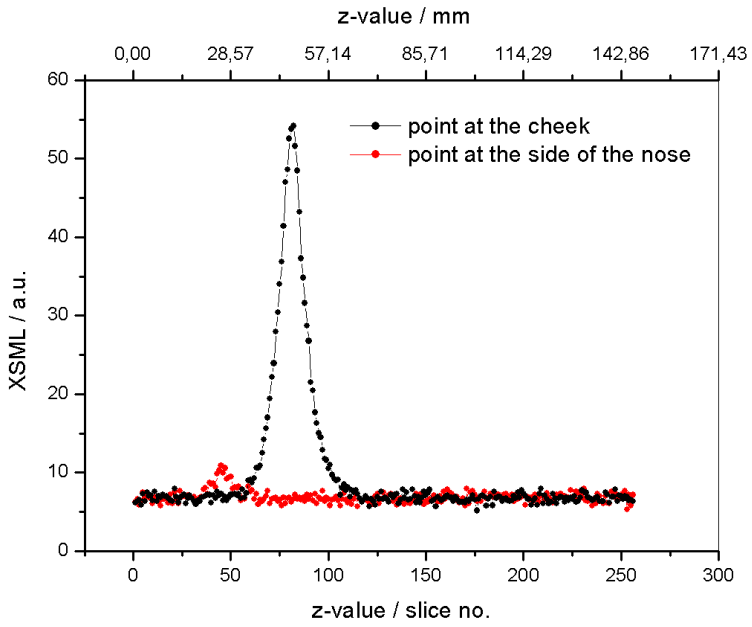

Figure 5.29: Exemplary focus profiles from points at the cheek and the side of the nose. The latter shows a significant lower peak height due to very narrow focused contours in case of steep slopes.

The method proposed by Subbarao is now modified in such a way that for each surface point the slope of a tangent plane is determined by the first surface estimation. In the method of Subbarao, the parameters of the tangent plane are used as a starting point to find the optimal plane. In the algorithm proposed in this thesis no optimization is performed, but the parameters of the tangent plane are used to calculate the focus values with tilted neighborhoods. After that the surface extraction is done by a maximization of the focus values in the common way. Figure 5.30 illustrates schematically the idea. In the case of planar neighborhoods, especially in regions with a steep slope, the focused contours present in a planar image are very narrow. Due to this the peak of the focus profile is relatively low. If the used neighborhoods are tilted in such a way that they build a tangent plane to the surface, the difference between adjacent z-positions is much higher since the whole neighborhood contains completely focused or completely unfocused image points, respectively. Figure 5.31 illustrates the same idea three-dimensionally. The tilted neighborhoods are magnified for a better visibility.

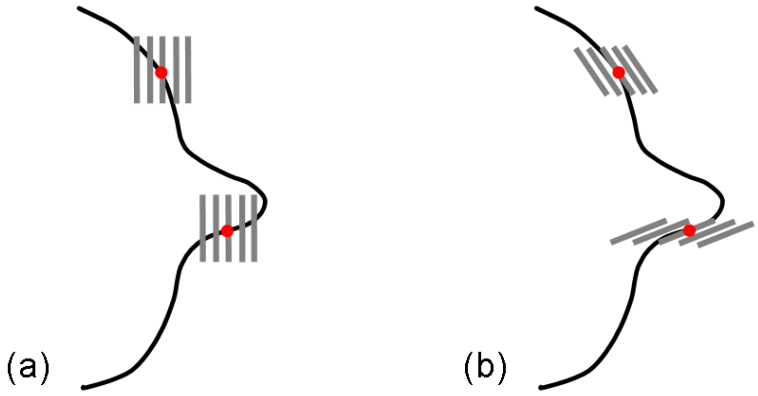

Figure 5.30: Schematic drawing of the neighborhoods used for calculating the focus value lying in the image plane (a) and overlapping different image planes (b). The slope of the tilted neighborhood corresponds to the slope of the tangent plane of the first surface estimate.

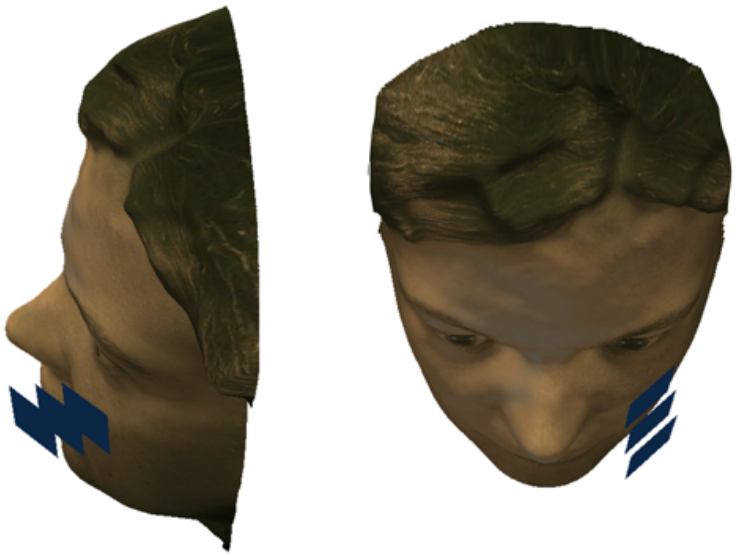

Figure 5.31: Three-dimensional illustration of the neighborhoods used for calculating the focus value. They are magnified for better visibility.

5.4.1 Detailed description of the algorithm

Starting from the first estimate of the surface, which can be obtained using a large neighborhood size to overcome noise artifacts and the symmetry approach to avoid plateau artifacts, the tangent plane for every surface point (x_0, y_0) is characterized through the gradient in x and y direction.

$$s_{x_0} = \frac{h(x_0 + \Delta, y_0) - h(x_0 - \Delta, y_0)}{2\Delta} \tag{5.19}$$

$$s_{y_0} = \frac{h(x_0, y_0 + \Delta) - h(x_0, y_0 - \Delta)}{2\Delta} \tag{5.20}$$

It is found optimal to choose half the size of the used neighborhood size as Δ.

The XSML-value is calculated in two steps (see appendix A). First the modified Laplace value is determined for every point and then these values are summed over a neighborhood. This summation, which is normally done over a neighborhood completely lying in one tomogram, is modified in such a way that it is performed over the same lateral range, but for varying tomograms corresponding to the determined slope.

The neighborhood U_d of size d centered around (x_0, y_0) at an arbitrary z-value $\tilde{z}$ is determined as:

$$U_d(x_0, y_0, \tilde{z}) = \{(x, y, [\tilde{z} + (x - x_0)\, s_{x_0} + (y - y_0)\, s_{y_0}]) : |x - x_0| \leq d \wedge |y - y_0| \leq d\}, \tag{5.21}$$

where [ ] represents the nearest integer value. The summation is now performed over tilted neighborhoods determined according to equation 5.21. Figure 5.32 shows a horizontal cut through the image volume at the bridge of the nose (see section 3.5 for the explanation of a horizontal cut). The first estimate of the surface is drawn in red. Five exemplary neighborhoods with a distance of $\Delta\tilde{z}$=10 are drawn in white. Their slope corresponds to that of the surface at the central point.

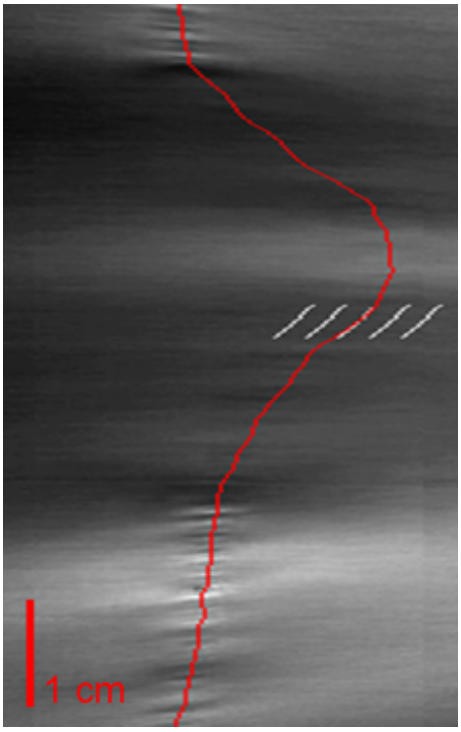

Figure 5.32: Horizontal cut through the image volume at the bridge of the nose. The first estimate of the surface is drawn in red, exemplary neighborhoods with a distance of $\Delta\tilde{z}$=10 are drawn in white. Their slope corresponds to that of the surface at the central point.

5.4.2 Results

The proposed algorithm was applied to simulated data with individually calculated PSF-directions for different levels of added noise. Neighborhood sizes of 11, 13, 15, and 17 pixels were used and the results were averaged. The dependence between the mean absolute deviation and the added noise can be seen in figure 5.33. The tilted neighborhoods reduce the error significantly.

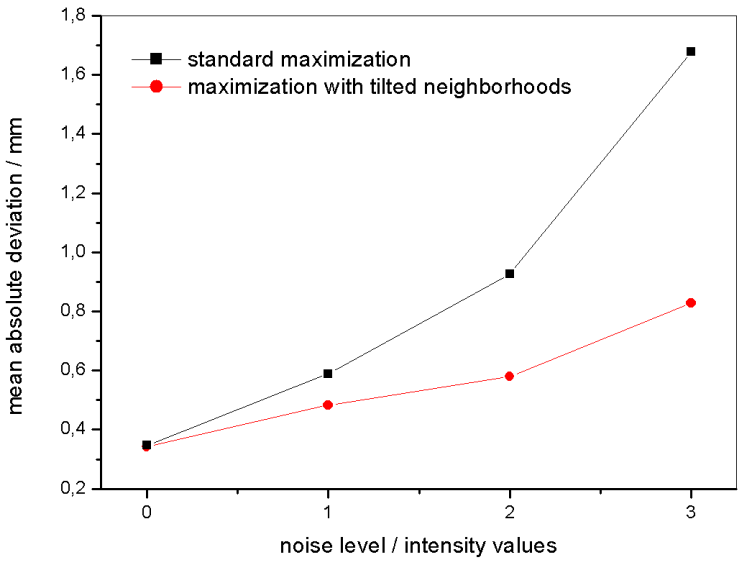

Figure 5.33: Mean absolute deviation against noise level for standard maximization with plane neighborhoods and with tilted neighborhoods for simulated data with individually calculated PSF-directions.

Also in real data an improvement is visible. Figure 5.34 shows height maps of a section of the face for standard maximization and maximization with tilted neighborhoods, in both cases 13 pixels in size. The sides of the nose, which have naturally a steep slope, show noise artifacts due to a low signal-to-noise ratio (a). They are reduced if tilted neighborhoods are used. Figure 5.35 shows an exemplary focus profile in this part of the face. The profile calculated with tilted neighborhoods shows a better signal-to-noise ratio which leads to a reduction of noise artifacts.

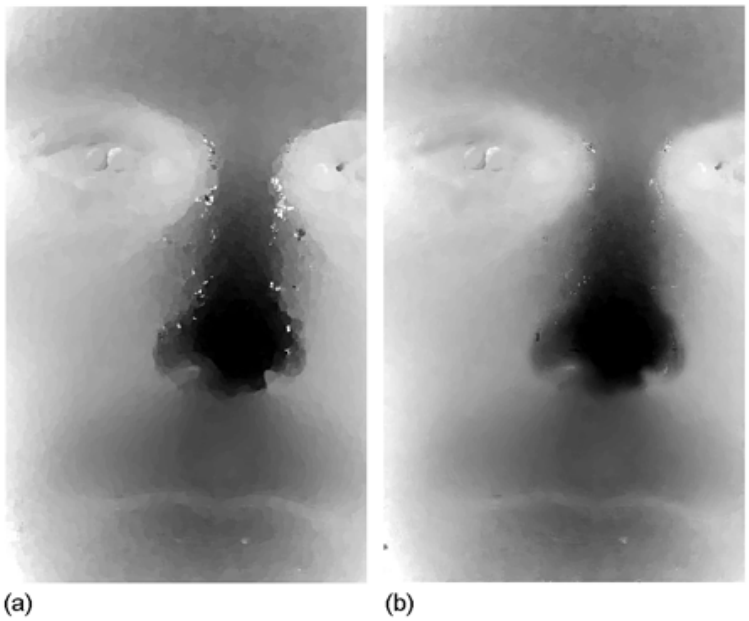

Figure 5.34: Height map of a section of the face for a real data set gained with a neighborhood size of 13 pixels through standard maximization (a) and maximization with tilted neighborhoods (b). Noise artifacts at the sides of the nose are reduced by the latter.

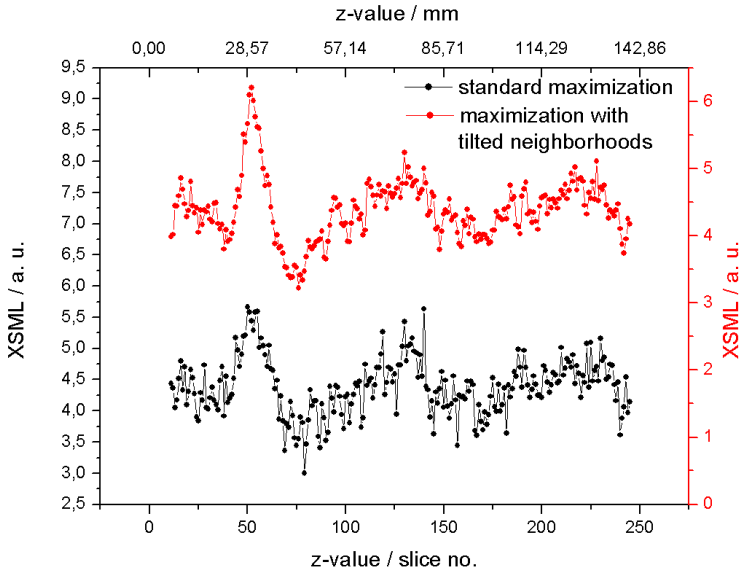

Figure 5.35: Focus profile from a real data set for an exemplary point on the side of the nose for standard maximization (red) and maximization with tilted neighborhoods (black). The red curve shows a better signal-to-noise ratio.

5.5 Combination of methods

In this chapter several independent methods for surface extraction were proposed, which are alternative approaches to the common maximization of the focus profile parallel to the z-axis.

If the focus profile symmetry is maximized instead of the focus profile itself, especially for large neighborhood sizes, plateau artifacts are eliminated. The same is true for fitting the focus profile with a Gaussian curve, which requires a higher computational effort, but produces on the other hand continuous values.

Both methods deal with the determination of optimal z-position by a given focus profile. In addition to this, two methods were suggested dealing with the choice of the points belonging to a focus profile. The first one regards focus profiles along estimated PSF-directions, the other method tilts the neighborhoods, so that they are tangential to the surface estimated in the first iteration step. These two methods are independent and can be combined, as is indicated in figure 5.36.

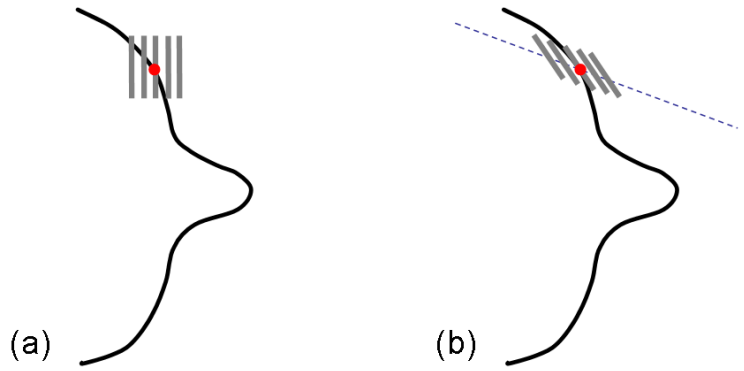

Figure 5.36: Schematic drawing of the neighborhoods used for calculating the focus value. They can be chosen parallel to the z-axis and lying in the planes of the tomograms (a) or be considered along the PSF-direction and tilted in a way that they are tangential to the first estimate of the surface (b).

The initial surface extraction should be done with large neighborhoods to reduce the influence of noise. Using the symmetry approach helps to avoid plateau artifacts.

The focus profile produced by a combination of the two iterative algorithms can then be fitted with a Gaussian curve, so that all methods can be merged.

Figure 5.37 shows the results gained with all three methods separately and their combination for simulated data with individually calculated PSF-directions and different levels

of added noise. A constant neighborhood size of 15 pixels was used and 20 tomograms before and after the first estimate of the surface were considered.

In comparison to the results gained with simple maximization, the first two methods improve the result nearly equally, while the usage of the Gaussian fits leads to a clearly bigger improvement. With the combination of the methods, a tremendous reduction in the absolute deviation between the extracted and the reference surface can be achieved. It lies under 0.25 mm for all noise levels, which is lower than half the inter-tomogram distance of 0.571 mm used for the simulation.

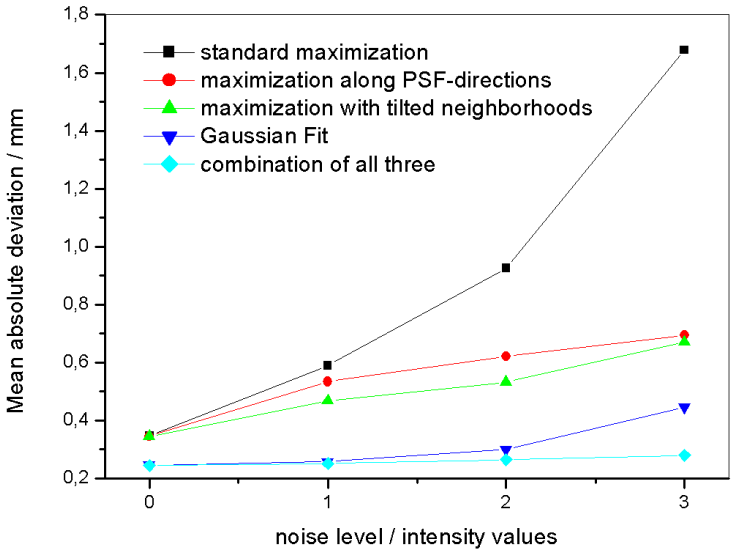

Figure 5.37: Mean absolute deviation against added noise level for the three discussed methods separately and their combination for simulated data with individually calculated PSF-directions and a constant neighborhood size of 15 pixels.

5.6 Summary

After the detailed discussion about the tools for measuring the amount of focus in the tomograms in chapter 4, this chapter dealt with alternatives to the standard procedure of surface extraction introduced in section 2.2.7.

Section 5.1 introduced a newly invented method for maximizing focus profile symmetry instead of maximizing the focus profile itself. Especially for large neighborhood sizes this leads to a quality improvement of the surface.

The following section 5.2 discussed established methods, Gaussian interpolation and parabolic regression, for surface interpolation and additionally a Gaussian fit for the same

task. As far as I know, it is the first time that such an approach is reported. It was found that the Gaussian fit is superior to the two other mentioned methods. It is not only able to create surface points not restricted to the discrete tomogram positions, it also reduces artifacts in the extracted surface and lowers the influence of noise.

Section 5.3 presented an advancement to previous work, where, based on a first estimation of the surface by standard methods, the direction of the holographic point spread function, which is different for each focus point, is extracted automatically from the digitized images. No ab-initio knowledge about the reconstruction geometry is necessary. The achieved improvement stayed behind the observed one in previous work. This can be lead back to the different focus criteria used. The previously used variance is more sensitive to a mismatch between the maximization direction and the direction of image formation than the novel XSML criterion. This is why the improvement through the new algorithm is not as effective as was expected. Nevertheless was the algorithm capable of improving the quality of the extracted surface.

The algorithm presented in section 5.4 is a modification of a known method. The piecewise constant approximation of the surface, which is inherent to the summation of contrast values over a plane neighborhood, is replaced by the summation over a tilted neighborhood, whose slope corresponds to the slope of the tangent plane. The latter is gained by first estimation of the surface through standard maximization. This method increases the signal-to-noise ratio especially in regions of steep slope and thereby improves the quality of the extracted surface.

The combination and interaction of the presented novel algorithms was discussed in section 5.5. The maximization of the symmetry of the focus profile is a suitable method if computation time is limited, for example for real time applications, or if a quick but reliable first estimate of the surface is required as starting point for two-step algorithms. The combination of the remaining three algorithms was compared with the results gained by their separate application. The best performance was achieved by the Gaussian fit. For a noise level corresponding to the noise present in the real holographic digitizations, the performance difference between the Gaussian fit and the combination of all three methods is low. For higher noise levels on the other hand the combination produces results nearly independent of the noise, which is a remarkable quality and can become very important if the methods are transferred to other fields of application.

The following chapter discusses a method for dealing with the still remaining surface fluctuations.

Chapter 6

Surface refinement and evaluation

A lot of improvements through newly developed algorithms were reported in the previous chapters. However, despite sophisticated surface extraction methods, little noisy surface fluctuations cannot be avoided entirely. This chapter presents a method for surface smoothing proposed by Tasdizen [TWBO03], which eliminates noisy fluctuations while preserving characteristic features of the face.

6.1 Anisotropic diffusion of normals

All image data captured by technical devices suffers from noise. This is why one of the most relevant uses of image processing is for smoothing and denoising images. Often denoising of images is done with a low-pass filter, which reduces noise, but also blurs sharp features and details, such as edges. The problem of preserving features while smoothing away noise has been studied extensively in computer vision. Anisotropic diffusion introduced by Perona and Malik [PM90] has been very successful in dealing with this problem in a wide range of images [TWBO03].

6.1.1 Anisotropic diffusion in images

The essential idea of the diffusion approach is to embed an image characterized through the intensity distribution $I_0(x, y)$ in a family of derived images $I(x, y, t)$ obtained by convolving the original image with a Gaussian kernel $G(x, y; t)$ of variance t (see figure 6.1):

$$I(x, y, t) = I_0(x, y) \star G(x, y; t). \tag{6.1}$$

Larger values of t correspond to images with a coarser resolution.

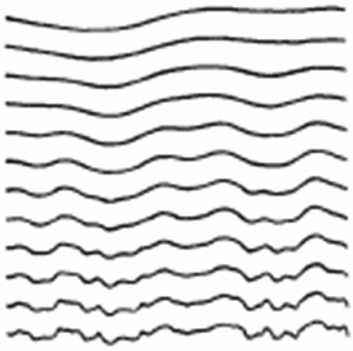

Figure 6.1: A family of 1-D signals I(x,t) obtained by convolving the original one (bottom) with Gaussian kernels whose variance t increases from bottom to top. Source: [PM90]

Assuming the space to be homogeneous and isotropic, the Gaussian smoothing can be described as a solution to the classical isotropic diffusion equation:

$$\frac{\partial I(x,y,t)}{\partial t} = c \cdot \mathrm{div}\left[\nabla I(x,y,t)\right] = c \cdot \triangle I(x,y,t), \tag{6.2}$$

with the constant diffusion coefficient c.

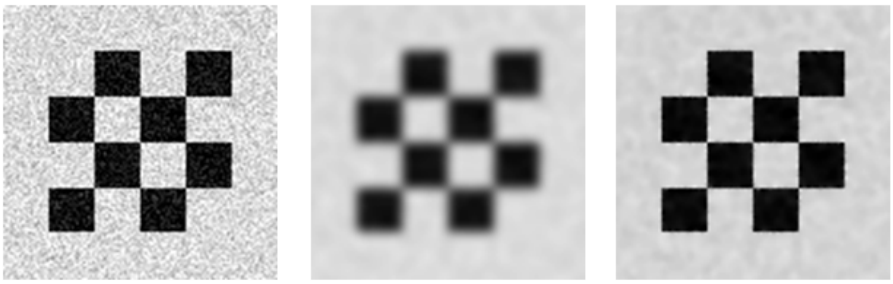

Figure 6.2: Noisy image (a), isotropically smoothed version (b) and version smoothed with anisotropic diffusion after 10 iterations (c).

Figure 6.2 shows as an example a noisy image (a) and the isotropically smoothed version (b). Through isotropic smoothing the boundaries become blurry. In order to invent a smoothing operation which respects boundaries, Perona and Malik considered an anisotropic diffusion equation

$$\frac{\partial I(x,y,t)}{\partial t} = div\left[c(x,y,t) \cdot \nabla I(x,y,t)\right], \tag{6.3}$$

with a non-constant diffusion coefficient $c(x,y,t)$. Ideally the coefficient should be set to one in the interior of the separated regions and zero at the boundaries. The blurring

would then take place separately in each region with no interaction between regions. Since the positions of the boundaries are not known, Perona and Malik suggested a so-called edge stopping function

$$c(x, y, t) = c\left(\|\nabla I\|\right) = \exp\left(-\frac{\|\nabla I\|^2}{2\mu^2}\right), \tag{6.4}$$

where μ is a positive, free parameter that controls the level of contrast of edges that effects the smoothing process. For $\|\nabla I\| \ll \mu$ the coefficient approaches one, while it approaches zero for $\|\nabla I\| \gg \mu$. Edges are generally associated with large image gradients, and thus diffusion across edges is stopped while regions that are relatively flat undergo smoothing [PM90], [TWBO03]. The differential equation 6.3 can numerically be solved in an iterative way. Figure 6.2(c) shows the result after ten iterations.

In order to use the proposed anisotropic diffusion for holographically gained surface models, one could apply the procedure to the extracted height map. A similar approach was followed by [TW03], who found that the anisotropic diffusion fails in this case. This is due to the fact that it was designed to preserve edges, which are discontinuities in the height map. They occur only at vertical cliffs and are therefore not very common in human faces. In human faces one is interested in ridges, which are continuous, but not differentiable features (see figure 6.3).

Figure 6.3: Schematic drawing of an edge and a ridge.

6.1.2 Anisotropic diffusion of normals

The generalization of anisotropic diffusion in images to three-dimensional surfaces was proposed by Tasdizen [TWBO03]. Since not the discontinuities in the height map but the discontinuities of its derivatives have to be preserved, the anisotropic diffusion process was carried on to the normal vectors of the surface. In analogy to equation 6.3 one gets a set of differential equations for the normal vectors:

$$\frac{\partial \vec{N}}{\partial t} = \left(\mathrm{Id} - \vec{N} \otimes \vec{N}\right) \nabla \cdot \left[c\left(\|\nabla \vec{N}\|\right) \nabla \vec{N}\right], \tag{6.5}$$

where the operator $\left(\mathrm{Id} - \vec{N} \otimes \vec{N}\right)$ enforces the normal vectors to remain unit length. The diffusion coefficient now no longer depends on the gradient, but on a second order quantity

of the height map, the total curvature

$$\kappa = \|\nabla \vec{N}\|, \tag{6.6}$$

where the norm denotes the Frobenius norm of a 3×3 matrix. It is the square root of the sum of the squares of all the matrix elements. The diffusion coefficient is in analogy to equation 6.4 expressed as:

$$c\left(\|\nabla \vec{N}\|\right) = \exp\left(-\frac{\kappa^2}{2\mu^2}\right). \tag{6.7}$$

Tasdizen showed that equation 6.5 is best to be solved in a two-step process. First the normal vectors are decoupled form the height map and smoothed, then the height map is refitted to the normal vectors (see figure 6.4).

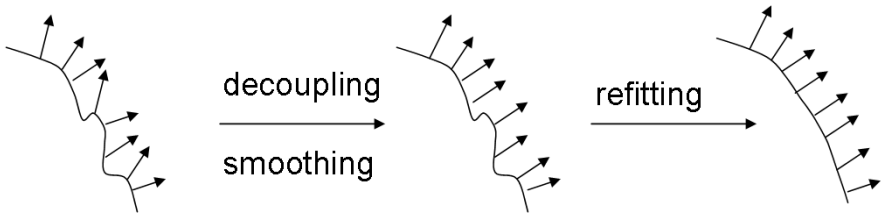

Figure 6.4: Equation 6.5 is solved in two steps. First the normal vectors are decoupled from the height map and smoothed. Then the height map is refitted to the smoothed normal vectors.

A flowchart of the whole algorithm is shown if figure 6.5. The normals are processed with a fixed number of 25 iterations following equation 6.5. This limitation avoids evolving the normals too far away from the height values. Then the height map is refitted until the discrepancy measure between the new normals and the normals of the height ceases to decrease. The overall algorithm repeats these two steps n times. This number of overall iterations n and the number μ from equation 6.7 are the two free parameters of the algorithm [TW03].

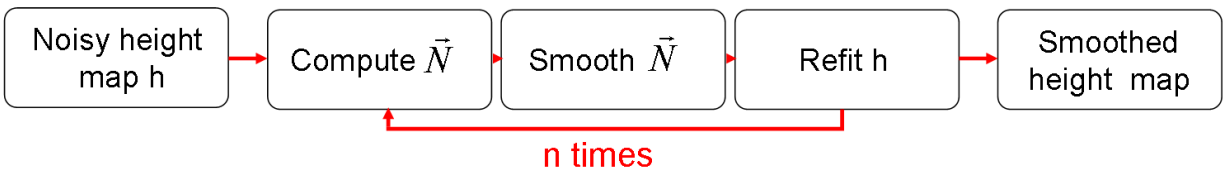

Figure 6.5: Flowchart of the whole smoothing process using anisotropic diffusion of normals.

6.1.3 Results

This section will present some results gained with anisotropic diffusion of normals. First a test objects is used, prerequisites of the input data are estimated and afterwards models gained by hologram tomography are dealt with.

Test model

The algorithm was implemented and first tested with a simple pyramid with added noise as test object. It can be seen in figure 6.6(a). When comparing the isotropic smoothed version 6.6(b) and the result of the anisotropic diffusion of normals with μ=0.1 after 15 iterations, it can be seen that the ridges of the pyramid are more pronounced in the latter. To further illustrate the process, the projections of the surface normal vectors and the total curvature are displayed in figure 6.7 after the first, the fourth and the fifteenth iteration. The normal vectors and the total curvature are transformed towards these for the ideal noiseless object through the anisotropic diffusion process.

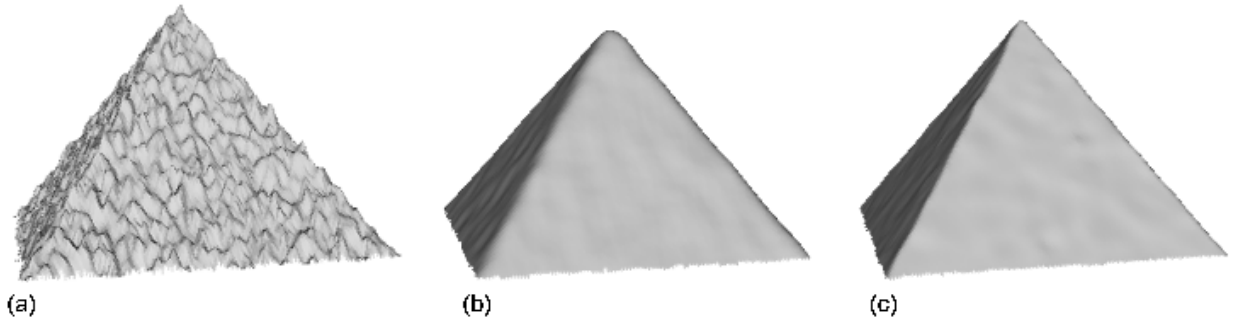

Figure 6.6: 3-D model of the test object with added noise (a), the result gained with isotropic diffusion (b) and anisotropic diffusion of normals with μ=0.1 after 15 iterations (c). The ridges of the pyramid are more pronounced in the latter.

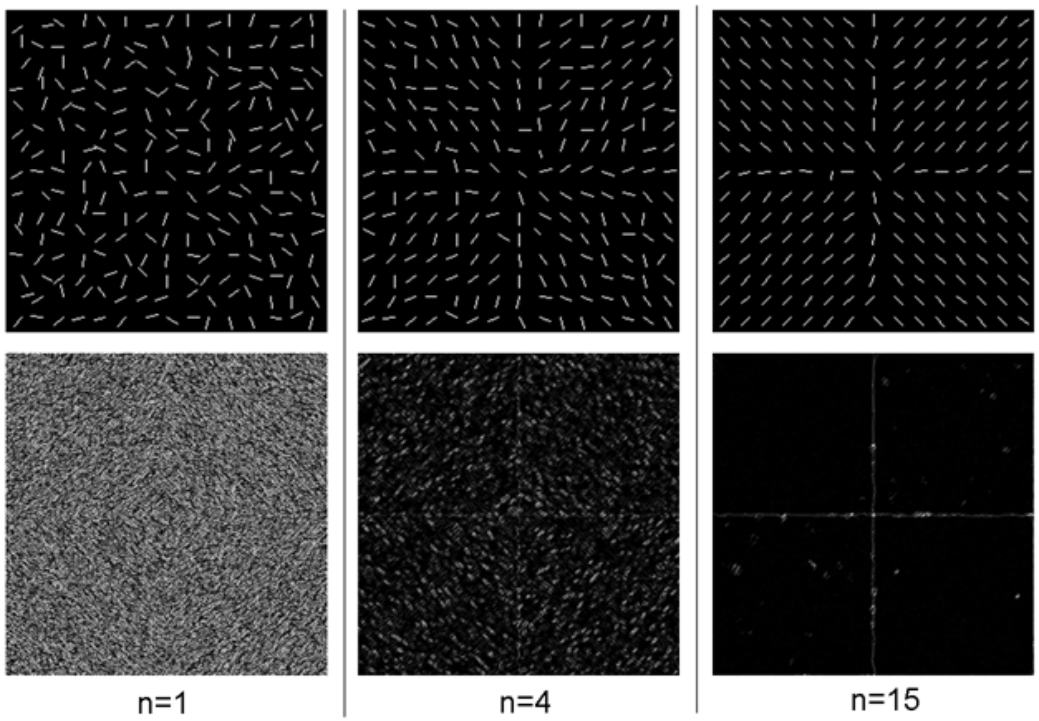

Figure 6.7: Projection of the surface normals (upper row) and the total curvature (lower row) for μ=0.1 after the first, the fourth, and the fifteenth iteration of the described algorithm.

Prerequisites

Except for the interpolation and fitting techniques presented in the last chapter, hologram tomography inherently produces height maps with discrete values. These step like surfaces cannot be smoothed with anisotropic diffusion of normals, as is demonstrated in figure 6.8. It shows a noiseless model with discrete height values. Anisotropic diffusion behaves in an unstable way and the model is contaminated with artifacts.

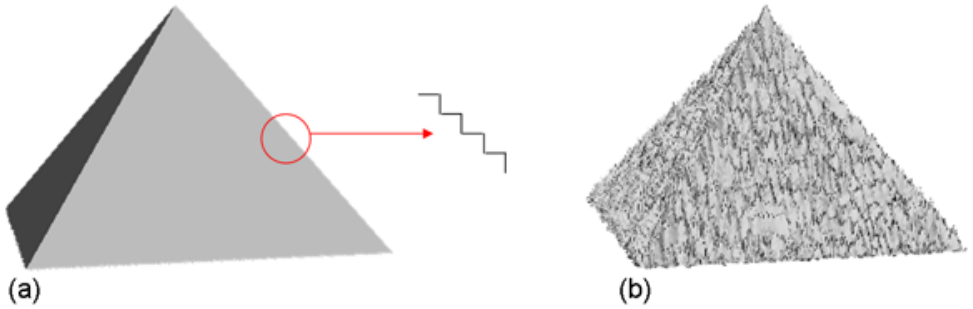

Figure 6.8: Noiseless test model with discrete height values (a) and the result of the anisotropic diffusion process (b), which behaves in an unstable manner.

This is why anisotropic diffusion of normals only works in combination with continuous models, which can be generated with the methods described in section 5.2.

Application to face models gained with hologram tomography

The first problem to be solved when applying a method like the one described above to new kinds of data is determining the best parameters.

Figure 6.9 shows the results gained with anisotropic diffusion of normals for a fixed number of 20 iterations and different values for the μ-parameter. The first model, which was produced with μ=0.1, shows fine structures in great detail, but also some remaining noise. In the model gained with μ=0.25, the fine structures are nearly as visible as in the first but the remaining noise is eliminated. In the last model on the other hand, produced with μ=0.4, the surface is very smooth but the visibility of the fine structures is affected.

Different numbers of iterations with a fixed μ-value (μ=0.25) were compared in figure 6.10. The visibility of the fine structures of the face is unmodified, while the smoothness of the remaining parts varies. Between 20 and 50 iterations, nearly no difference is apparent.

Since similar results were achieved with other facial models, it can be summarized that the number of iterations should be chosen between 10 and 20. Concerning the μ-value,

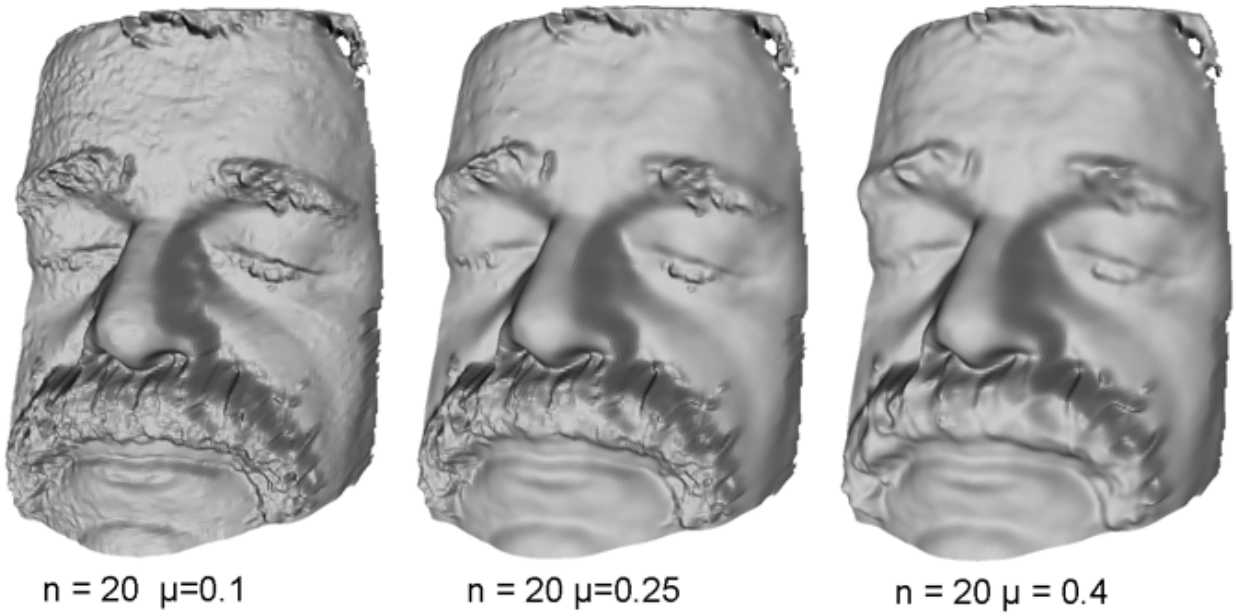

Figure 6.9: Results gained with anisotropic diffusion of normals for a fixed number of iterations (n=20) and three different μ- values.

good results were gained with μ from 0.2 to 0.3, depending on the noise content of the data set[1].

With the parameters specified, a comparison between isotropic smoothing and smoothing through anisotropic diffusion of normals could be made. Figure 6.11 shows the comparison between the raw data (a) gained with the multi-scale approach in combination with the Gaussian fit, the isotropic smoothed version (b) and the surface smoothed with anisotropic diffusion of normals with μ=0.25 after 20 iterations (c). Although the smoothness of the unfeatured parts of the face is nearly the same, the small features are tremendously obliterated in (b) while they are as pronounced as in the raw data in (c).

[1]Although the presented algorithm going back to Tasdizen [TWBO03] was implemented by myself, the results for the facial models were gained with the open source implementation of the author himself, available as part of the software toolkit ITK at http://www.itk.org. Different than the self implemented version, this one is multi-threaded so that a double processor computer could be used efficiently. With it, the computation time was shorter than 10 minutes.

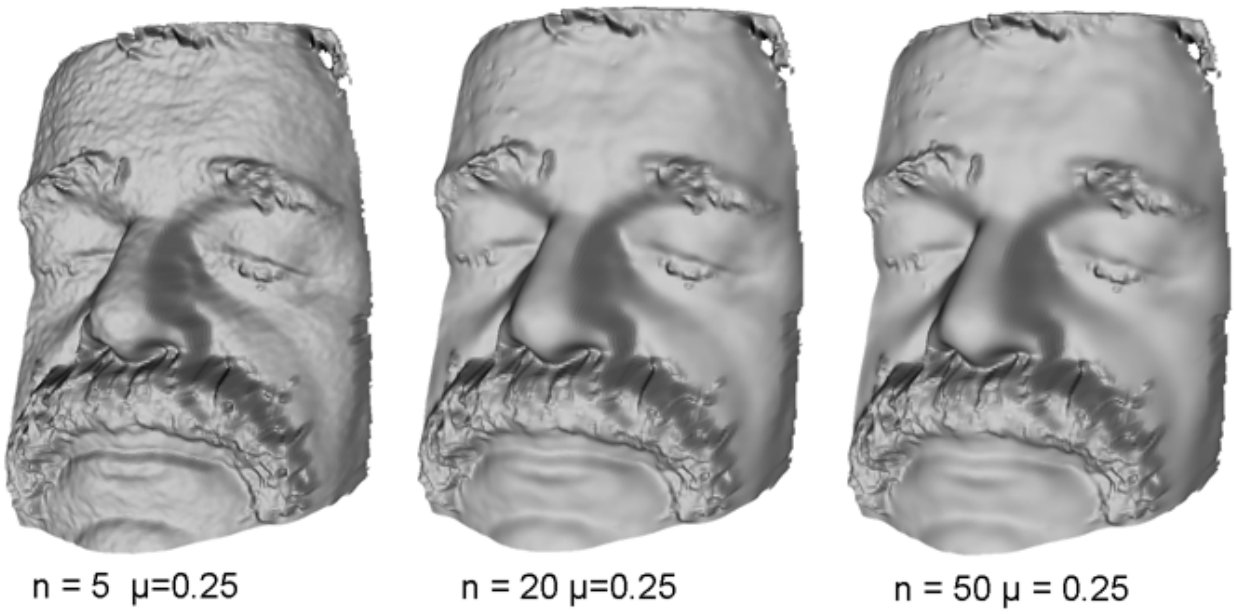

Figure 6.10: Results gained with anisotropic diffusion of normals for a fixed μ parameter of 0.25 and three different numbers of iterations.

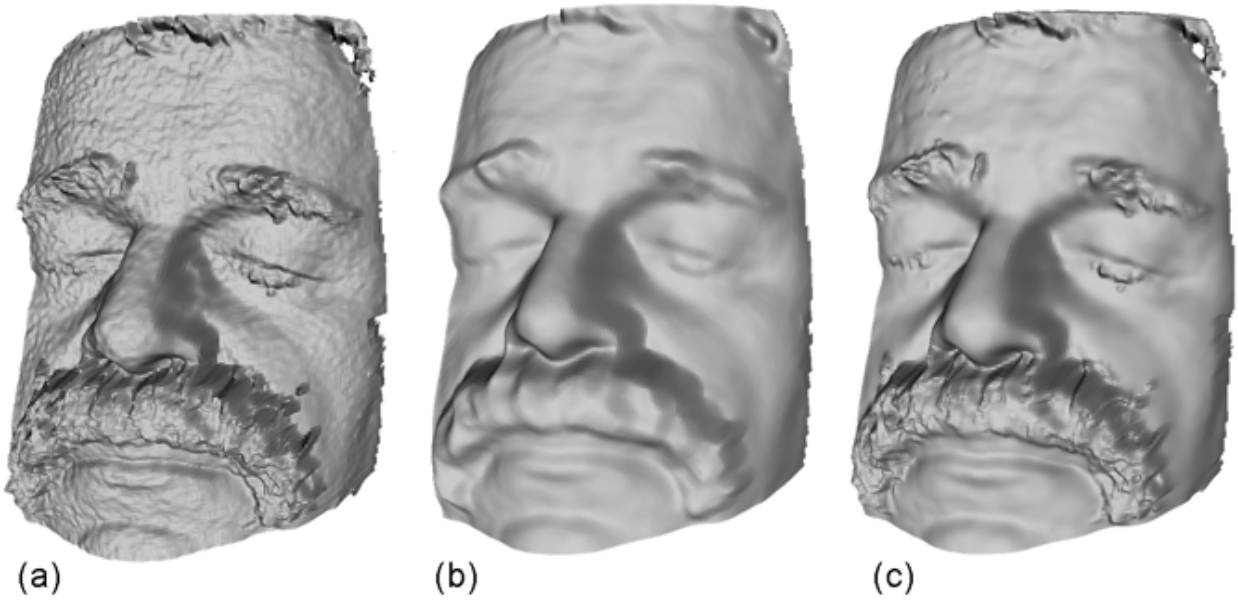

Figure 6.11: Comparison between isotropic and anisotropic smoothing. (a) shows the raw data, (b) the isotropically smoothed version, and (c) the surface smoothed with anisotropic diffusion of normals with μ=0.25 after 20 iterations.

Another example can be seen in figure 6.12, where the chin region of another proband is shown. (a) pictures the raw data, (b) the isotropic smoothed version and (c) the one gained with anisotropic diffusion of normals with the same parameters used in the previous example (n=20, μ=0.25). Height profiles of the same holographic data set can be seen in figure 6.13, where the drawing colors correspond to the colored bounding boxes in figure 6.12. The advantages of the anisotropic diffusion show clearly in regions with high curvature.

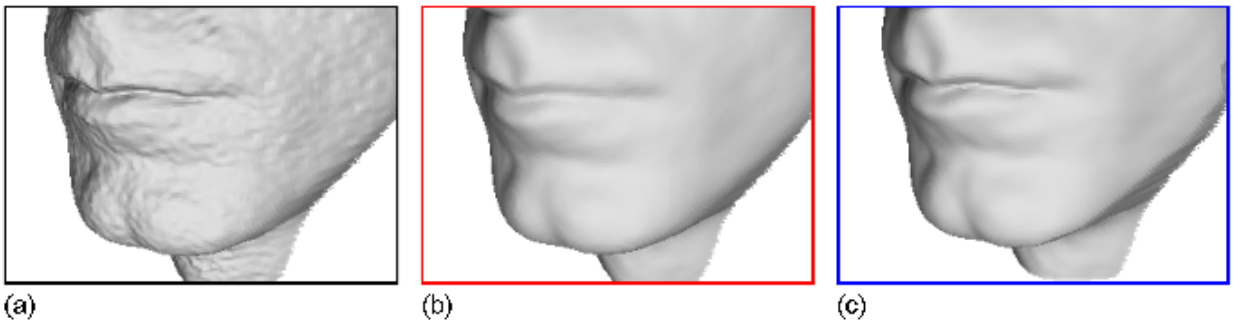

Figure 6.12: Chin of a proband isotropically smoothed (b) and smoothed with anisotropic diffusion of normals with μ=0.25 after 20 iterations (c) in comparison to the raw data (a). The colored bounding boxes correlate with the drawing colors in figure 6.13.

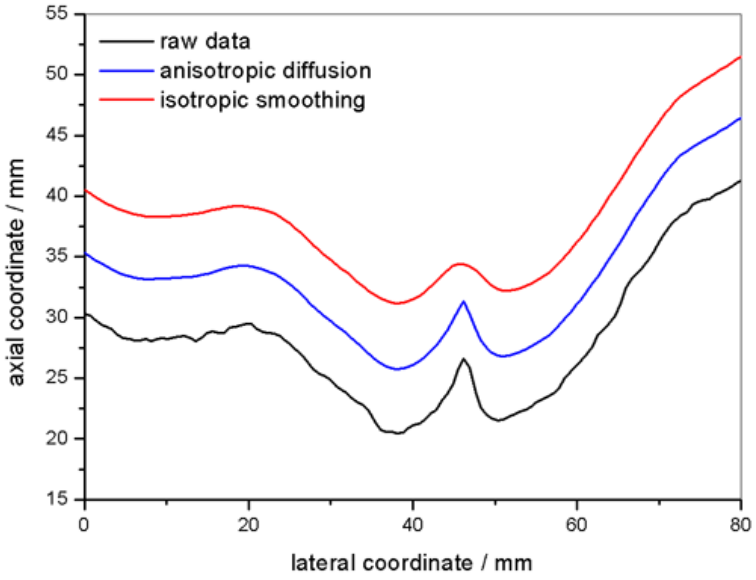

Figure 6.13: Raw height profile along a line crossing the lip, the isotropically smoothed and the anisotropically smoothed version. The drawing colors correspond the to the colored bounding boxes in figure 6.12. The profiles were shifted about 5 mm each for better comparability.

6.2 Evaluation of the extracted surface

Up to now, the proposed methods were evaluated in a qualitative way by comparing focus profiles and height maps from real holographic reconstructions and in a quantitative way by using the simulated data sets.

Now the algorithms described in this thesis are applied to a real holographic reconstruction of a face-like object. The extracted surface is compared with a reference surface gained with an alternative topometry method.

The test object was an anatomical model originally used to train orthopaedic surgeons on specific surgical techniques. It was a head model called spine surgery simulator from creaplast[2] (see figure 6.14).

Figure 6.14: Photograph of the plastic model used as test object.

As reference acted a high precision CT-scan done with the computer tomograph SOMATOM Sensation 4 from Siemens. Slice thickness was 1 mm, the voxel size was 0.5 mm $\times$ 0.5 mm $\times$ 1.0 mm. The head model and the CT-data were kindly provided by Filip Schutyser from medicim medical image computing[3].

Since the model did not have a surface rich in contrast, the hologram was recorded using a speckle pattern as structured illumination. The digitization was performed with a scanner resolution of 150 dpi, which corresponds to a lateral pixel size of 0.169 mm and an inter-tomogram distance of 0.498 mm.

Since the texture properties were totally different from the natural texture of human

[2]See http://creaplast.littoralmedia.com/indexgb.htm
[3]see http://www.medicim.com

skin due to the used structured illumination, the first step was finding an optimal focus criterion. In order to do this, once again a simulation was performed. Analogue to the simulation of a human face, it was found that the novel XSML operator worked best, although the difference in performance between the individual criteria was lower than for real faces. In a second step, the multi-scale approach proposed in section 4.2 was applied. Since the XSML focus criterion is not sensitive to a mismatch between the direction of image formation and focus maximization, it was not deemed necessary to use the algorithm considering the PSF-directions. Given that the structured illumination delivers a high contrast pattern and with that a naturally higher signal-to-noise ratio than that of human faces, also the maximization using tilted neighborhoods was set aside. The usage of the Gaussian fit, however, was essential for producing a continuous surface and allowing for anisotropic diffusion, which was applied with the parameters n=50 and μ=0.1.

The result can be seen in figure 6.15. Part (a) shows the reference surface from the CT-scan, part (b) shows the model gained by the combination of methods described above out of a digitized holographic reconstruction.

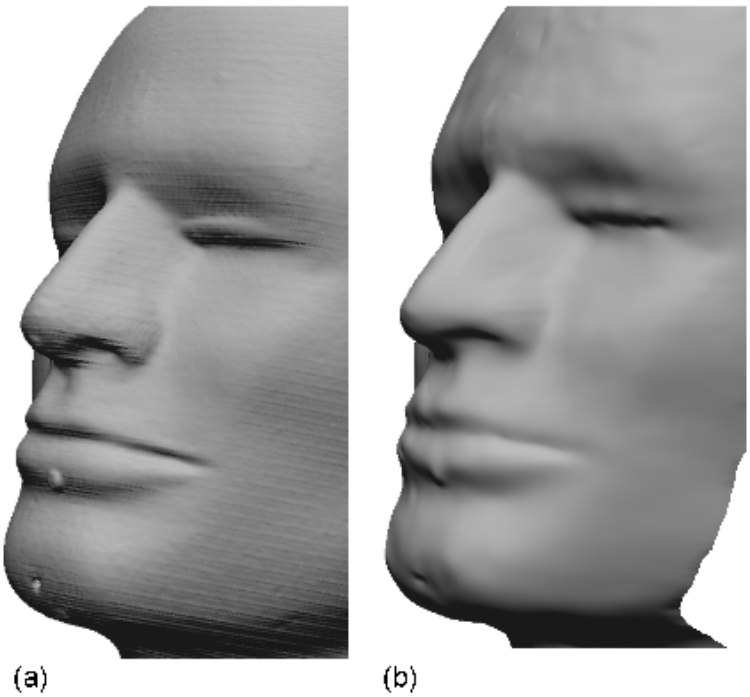

Figure 6.15: Reference surface from a CT scan (a) and holographical model gained with a combination of newly invented methods in combination with anisotropic diffusion (b).

In order to realize a quantitative comparison, the two surfaces had to be registered (see section 3.1). This was done with the software RapidForm[4] using rigid surface registration.

[4]see http://www.rapidform.com

Three-dimensional translations and rotations were performed until the total deviation between the two surfaces was minimal. After registration, the absolute difference between the two models was calculated and displayed as a color map (see figure 6.16). The average deviation was 0.28 mm. Major deviations appeared in the prominent features of the face like the lips, eyes and around the nose.

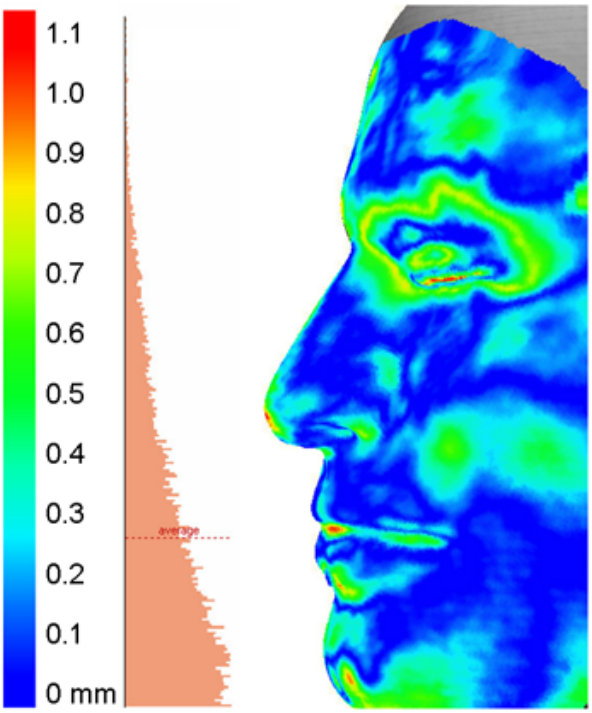

Figure 6.16: Deviation in mm between the reference surface (CT) and the holographically gained surface coded into a color map. A histogram of the deviations is plotted next to the reference color bar.

The reference surface was extracted from a CT scan with a slice thickness of 1 mm. The holographic reconstruction on the other hand provided a continuous surface due to the used Gaussian fit. Just because of the discrete nature of the CT reference surface, a minimal deviation of $\frac{1}{4} \cdot 1.0$ mm $= 0.25$ mm was expected (see section 5.2.3). A comparison to the calculated mean deviation of 0.28 mm shows that the two surfaces are in extraordinary accordance. The deviation apart from the expected discretization error is only approximately 0.03 mm, which is more than a factor of ten smaller than the inter-tomogram distance of 0.498 mm.

6.3 Summary

This chapter introduced a well-known method based on anisotropic diffusion of normals for surface smoothing while obtaining characteristic features of the face. It was demonstrated that the method can only be applied to holographic data sets if continuous surface information is at hand. This is why this smoothing method has to be combined with one of the surface interpolation methods described in section 5.2, preferably the Gaussian fit. Considering this, the anisotropic diffusion of normals is a capable tool for surface refinement in hologram tomography. Optimal smoothing parameters were also discussed in this chapter.

Additionally, a comparison between a holographically gained surface of a face-like test object and a corresponding surface produced with a computer tomograph was performed. The holographic surface was extracted with a combination of the novel algorithms presented in this thesis. The calculated mean deviation between the two surfaces was 0.28 mm. Taking into account that the CT surface had a discretization step size of 1 mm, a minimum deviation of 0.25 mm was expected only due the discretization error. Considering this, the two surfaces are in extraordinary accordance, with a resulting error of only approximately 0.03 mm, which lies by more than a factor of ten under the inter-tomogram distance of 0.498 mm.

Chapter 7

Fields of application

Medical applications were the main motivation for the development of the holographic topometry system [Bon02]. Various clinical cases were reported in the works of Giel [Gie03] and Frey [Fre05] and many recent publications confirm the growing acceptance of the holographic facial measurement system for medical applications [Bri05], [Koc05], [FBT+05], [FTHH06], [TFH+06a], [Löf06], [BGZ+06], [TFH+06b].

In order to achieve optimal results in surgery, accurate information about the skull and the soft tissue thickness is essential. A three-dimensional computer model of the skull can be generated with a computer tomograph, three-dimensional, high-resolved information about the face is added by means of hologram tomography. Apart from the improved possibility to predict the outcome of an individual operation, if complete information about bone and soft tissue is at hand, three-dimensional face measurement can also be used for a quantitative documentation of surgery and is therefore important for the scientific improvement of surgical procedures.

Giel reported about the documentation of a complex jaw surgery performed by C. U. Fritzemeier [Gie03], where three-dimensional skull data (CT) and holographical surface data were used for a quantitative analysis of the surgery. In section 7.1, a clinical case is presented where a face lifting surgery is documented three-dimensionally. An alternative field of medical application is the generation of three-dimensional facial models for the adaption of epitheses [Fre05].

In addition to the medical application, holography can also be used for three-dimensional documentation and shape measurement of archeological findings. Although a fast recording is not necessary in this context, the precisely fitting texture and the possibility to transform the recorded hologram into a white light copy makes hologram tomography an

interesting tool for archaeology. An example will be shown in section 7.2.

Three-dimensional computer models of living human faces are also of great interest in forensic science, where facial reconstructions of skull findings are needed to identify dead people. A database of holographically recorded faces could be a crucial step for an automation of the facial reconstruction process. First steps towards such a data base are presented in section 7.3.

7.1 Surgery documentation

Documentation of plastic surgery is normally done with analog or digital photographs. The main differences in pre- and post-operative photographs are not only caused by the surgery but often date from a different mimic, light conditions during the recording, make-up or hair style (see figure 7.1). An objective evaluation of a face lifting surgery is only possible, if three-dimensional information is called in.

Figure 7.1: Manipulative photographic documentation of a face lifting surgery. The proband has a different mimic expression, hair style and make up and the photograph is recorded with different illumination. Source: [FBT+05]

A three-dimensional documentation of a face lifting surgery with means of hologram tomography was for the first time done in cooperation with C. U. Fritzemeier [FBT+05]. A pre- and post-operative photograph can be seen in figure 7.2. The patient was holographically recorded directly before and six weeks after surgery. This period is considered long enough for the operative tumescences to disappear. Screenshots of the resulting three-dimensional computer models can be seen in figure 7.3.

Figure 7.2: Photograph of the patient before (a) and after (b) the face lifting surgery. The photographs were provided by C. U. Fritzemeier.

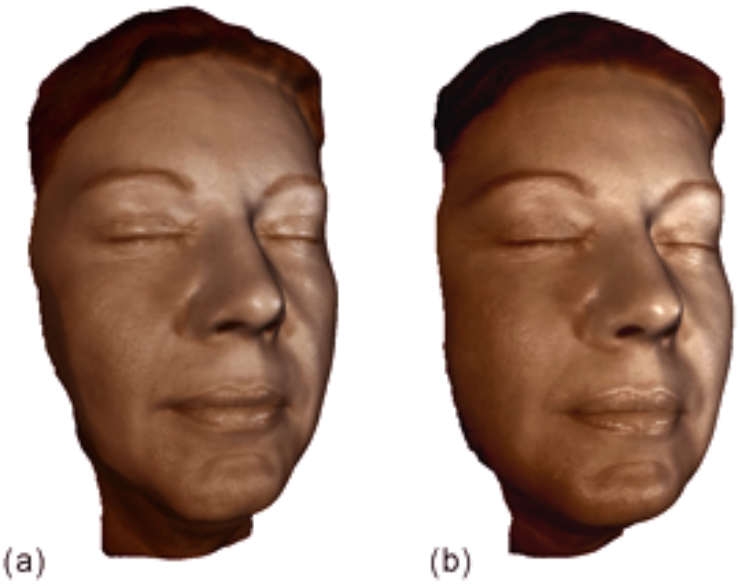

Figure 7.3: Three-dimensional computer models gained with hologram tomography before (a) and after (b) surgery.

In order to quantify the deviations metrically, the two resulting models had to be registered. Like mentioned in section 6.2, the registration was done with the help of RapidForm. Different than before, only those parts of the face were used for the registration process that were not effected by the surgery, mainly the back of the nose and the lower forehead. The result is displayed in figure 7.4. Metric deviations are coded in a color map. Noticeable is the fact that the average deviation is only approximately 0.5 mm. Small objective changes lead to remarkable changes in the subjective impression of a face. An objective evaluation can therefore only be done with high-resolved, three-dimensional computer models, as was demonstrated here.

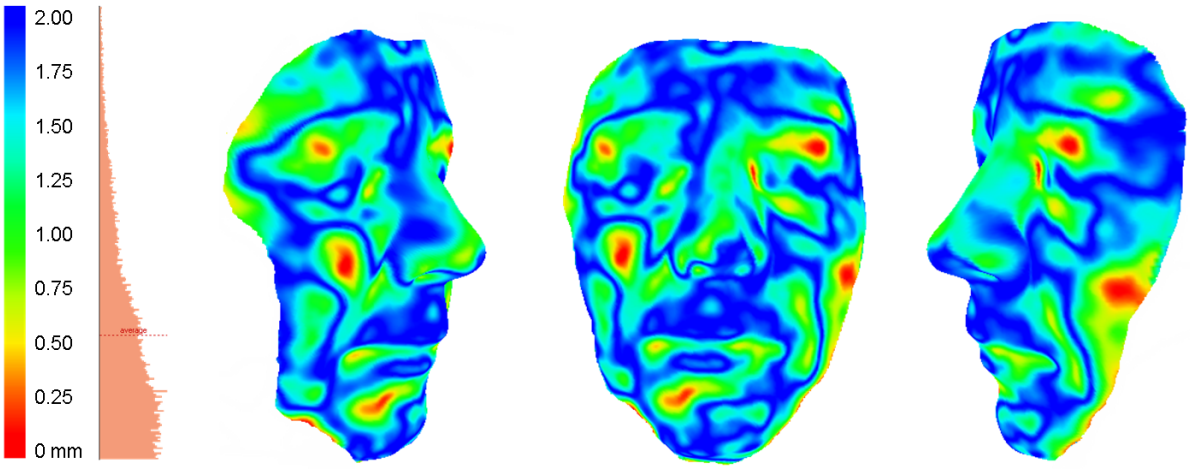

Figure 7.4: Quantitative comparison of the pre- and postoperative three-dimensional face model. The deviation is coded in a color map. A histogram of the deviations is plotted next to the reference color bar.

7.2 Cultural heritage

Cultural heritage, like archeological findings, are of great interest to the scientific community and the public. Since they require a special environment for maintaining their current condition, only few people get the chance for a detailed survey of the findings.

In order to allow for a detailed impression without endangering the findings, holograms are an ideal tool. White light holograms, which can be copied into as many reproductions as desired, serve as three-dimensional substitutes for the finding and can be displayed to the public or used for research purposes at arbitrary places. This has been demonstrated for the bog body of Husbäke [FBG+03], [Abb02], [Hol02]. Another example, where also the gained topometrical information was of great scientific interest, can be found in [Fre05]. In that case, the detailed information gained about the form of the scull, in particular, the strongly developed supraorbital ridges of a bog finding in northern Germany brought new and quantitative evidence into the controversial discussion about the sex of the bog body.

The documentation of cultural heritage presented in this section deals with the Neanderthal man, discovered in a limestone quarry of the Neander Valley near Düsseldorf in 1856. The finding contained the skull cap, two femora, three bones from the right arm and two from the left, and several other bones or bone fragments (see figure 7.5). It serves as type specimen for the *Homo neanderthalensis*. The workers who recovered this material originally thought it to be the remains of a bear. They gave the material to amateur naturalist Johann Karl Fuhlrott, who turned the fossils over to anatomist Hermann Schaafhausen. The discovery was announced in 1857. The skeleton is displayed nowadays in the Rheinisches Landesmuseum in Bonn [ST02].

Figure 7.5: Skeleton of the neanderthal man lying in front of the holographic camera.

In 2006, 150 years after the discovery, the finding has a long history of heated international debate. It marked the beginning of research into human evolution that is still continuing unabated today. Three-dimensional computer models with precisely fitting texture as well as display holograms contribute material for further research and educational assistance for the festivities of the anniversary.

In addition to holograms displaying all bone findings, especially the skull cap was of great interest. In order to create an all around model, several holograms from different viewing angles were recorded and optically reconstructed. Figure 7.6(a) and 7.6(b) show two of these views. In order to combine them, a registration of the overlapping region was performed using RapidForm. The result can be seen in figure 7.6(c). A challenging task thereby is the combination of the texture information, given that the illumination during the holographic recording is done through two diffuse but nevertheless directional light sources. This leads to an inhomogeneous brightness distribution and a brightness mismatch between different views, which can in outlines be seen in figure 7.6(c). A sophisticated method for texture brightness corrections should be object of further research.

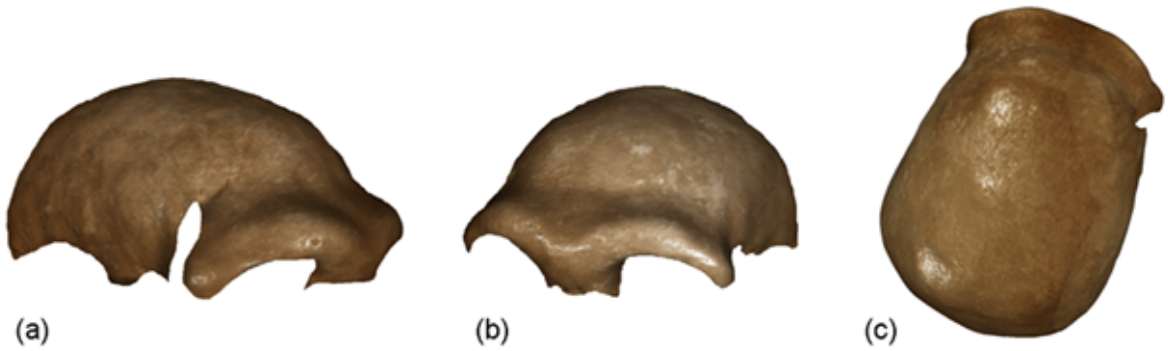

Figure 7.6: Screenshots of three-dimensional computer models of two separately recorded views of the skull cap of the neanderthal man (a and b) and a screenshot of the merged, textured all-around model (c).

7.3 Forensic applications

Forensic facial reconstruction is the process of creating the face of an unidentified individual from their skeletal remains. Facial reconstruction methods can be divided into two-dimensional and three-dimensional methods.

Two-dimensional methods are portrait drawings, which usually require the collaboration of an artist, or computer based methods like for example computer-superposition [Str04].

In the case of three-dimensional methods one can also distinguish between computer based methods or sculpturing methods, needing an artist for accomplishment.

The most critical problem in facial reconstruction is the data used as average soft tissue thickness for different categories of age, sex, body build and ethnical origin. The data lacks in accuracy and is limited in range.

As Giel [Gie03] and Frey [Fre05] already pointed out, a highly resolved, holographically gained data set of the face in combination with a CT data set of the skull provides a basis for soft tissue measurement. It could exceed classical soft tissue databases, which only contain soft tissue thicknesses at a small number of anatomical landmarks [Hel84, p. 52], in accuracy and number. Another crucial advantage of the holographically gained surface data is the fact that it captures the face in a sitting position. Since the appearance of a face changes enormously between a sitting and a lying position, for recognition purposes the sitting position is the natural choice.

First steps for the buildup of such a database are undertaken. This section will demonstrate the principles of using a holographic and a CT data set for soft tissue thickness estimation with the software RapidForm.

Basis is the holographic and the CT data set, as they can be seen in figure 7.7(a) and 7.7(b). A critical step is registering the holographic data set to the skull, since they normally have no common regions. In order to fulfill this task, additionally to the skull, also the facial surface was extracted from the CT data. It represents only a rough estimate of the facial surface and suffers deviations in all regions where the soft tissue moves due to gravity if the person is lying. This is why only regions with little soft tissue movement are considered for the registration between the holographic and CT facial surface model. These are mainly the back of the nose and the region of the lower forehead. A detailed investigation about the accuracy of this registration method, which has to be done in order to appraise the results of the soft tissue measurement, would exceed the scope of this work. The result of the registration process can be seen in figure 7.7(c).

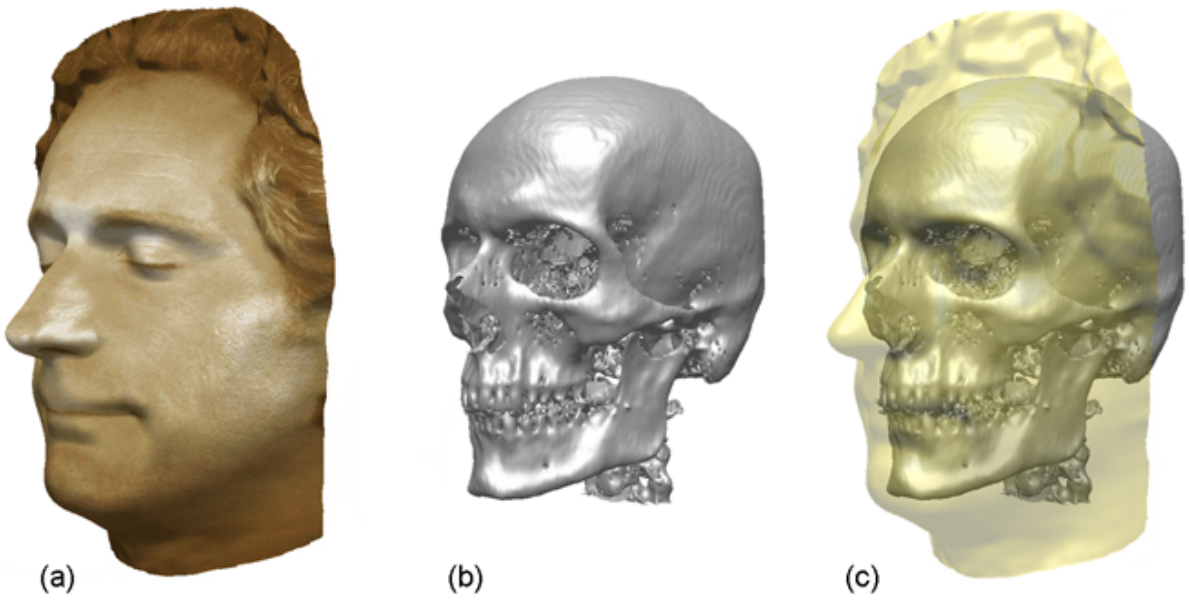

Figure 7.7: Holographically gained model of the face (a) and CT data set of the skull (b). A combination of the two is displayed in (c).

Once the registration is done, the soft tissue thickness can be measured perpendicular to the skull surface at arbitrary anatomical landmarks (see figure 7.8(a)) or for the whole skull simultaneously (see figure 7.8(b)). The latter figure displays the soft tissue thickness as a color coded map.

Also the investigation of geometrical cuts is possible. Figure 7.9 shows a cut in z-direction (a) and one in x-direction (b). The direction of the measurement is illustrated through the color coded displacement vectors.

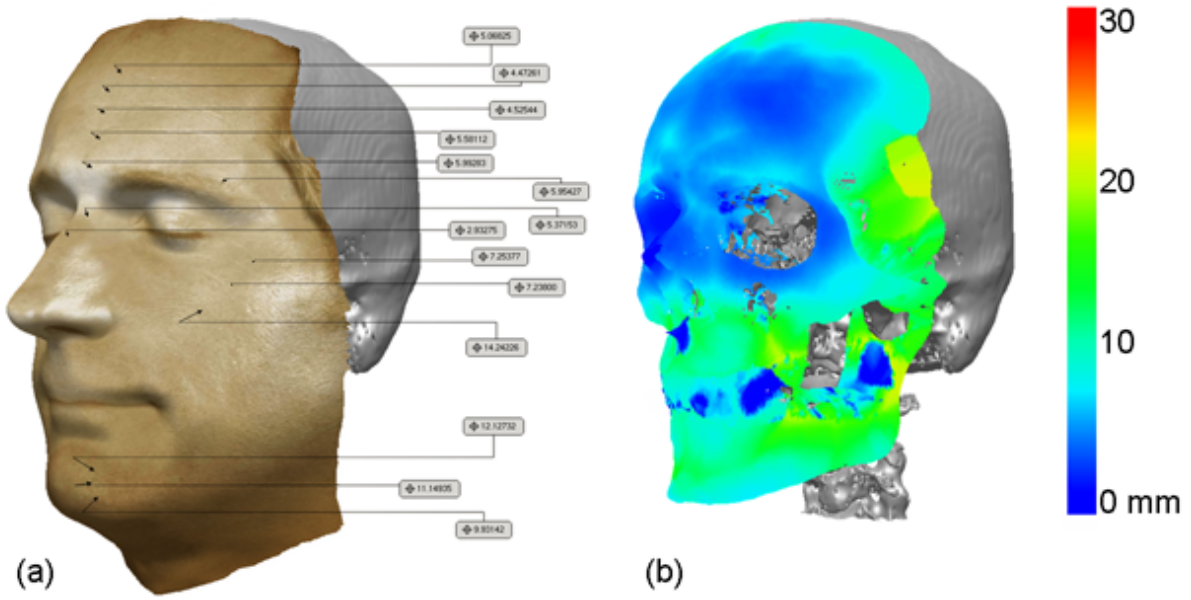

Figure 7.8: Soft tissue thickness at arbitrary anatomical landmarks (a) and color coded for every point on the skull (b).

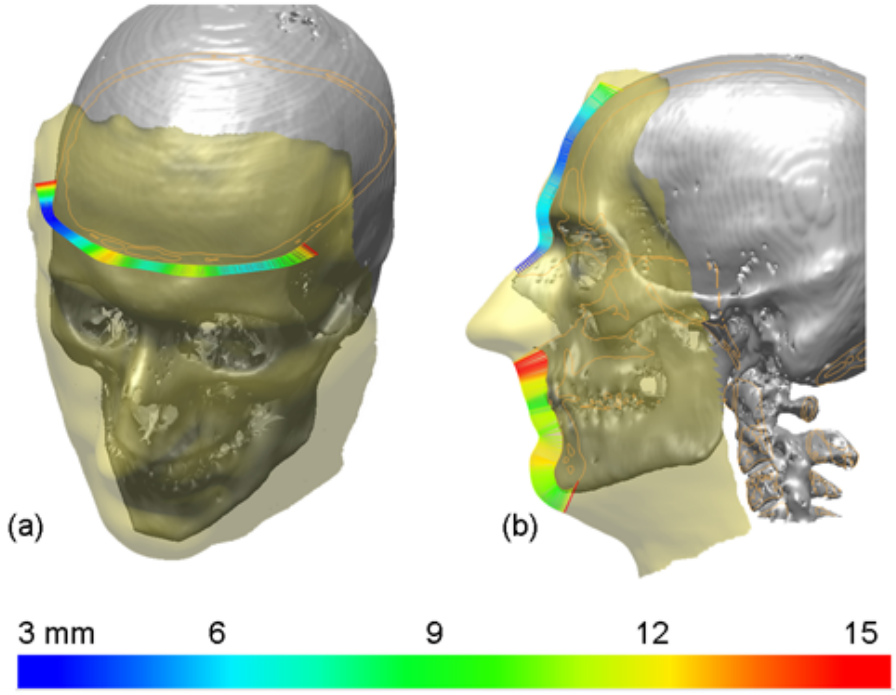

Figure 7.9: Cut in z-direction (a) and x-direction (b) with color coded displacement vectors.

After this general overview over using holographically gained facial models in combination with CT data for soft tissue measurement, the first steps for the buildup of a database are discussed. For this purpose, 18 holograms of probands with existing low-dose CT scans of the head were made. The CT data sets were kindly provided by Frank Prieels and Rik van Achten. They had a voxel size of 0.5 × 0.5 × 0.6 mm. The selected group of probands consisted of nine males and nine females, who were between 20 and 25 years old. They were all of the same ethnic origin (Europid). Figure 7.10 shows exemplarily the holographic and CT data set of one female and one male proband.

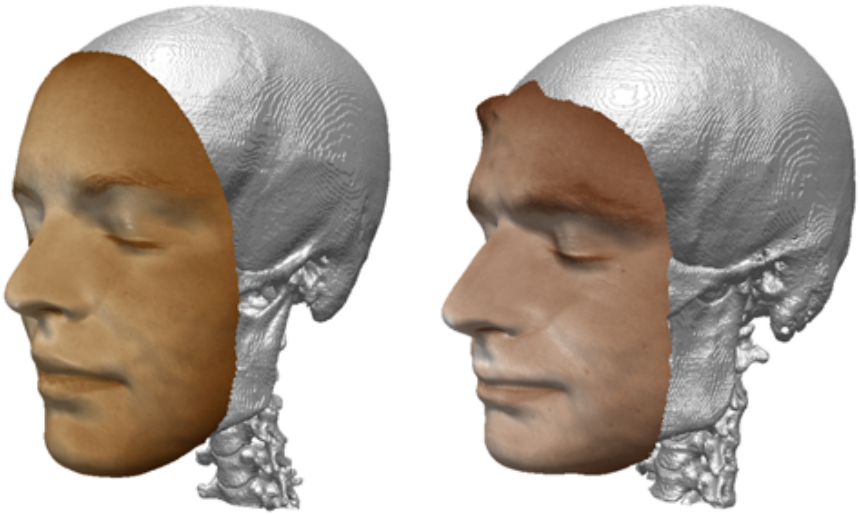

Figure 7.10: Exemplary holographic and CT data sets of one female (left) and one male (right) proband.

The procedure used for the registration between the skull and the holographically gained surface was already described above. The holographical face model (see figure 7.11(a)) is registered to the facial surface extracted from the CT data. Only the region marked in red is used for registration (see figure 7.11(b)). This region is little effected by soft tissue deformation due the different positions during the CT (lying position) and the holographic recording (sitting position) and can therefore be assumed to be equal in both models.

Classical soft tissue databases consists of averaged soft tissue thicknesses at anatomical landmarks, for example the 34 Helmer points [Hel84]. In contrast to former work, where the soft tissue thickness was estimated with cadavers, the database of Helmer was created with living people using the Ultra-Impulse-Echo-Technique. This technique measures the soft tissue thickness with ultrasound by positioning a sensor with a contact gel at the chosen anatomical landmarks. It has the drawback that the sensor is in direct contact with the soft tissue, which might lead to deformations.

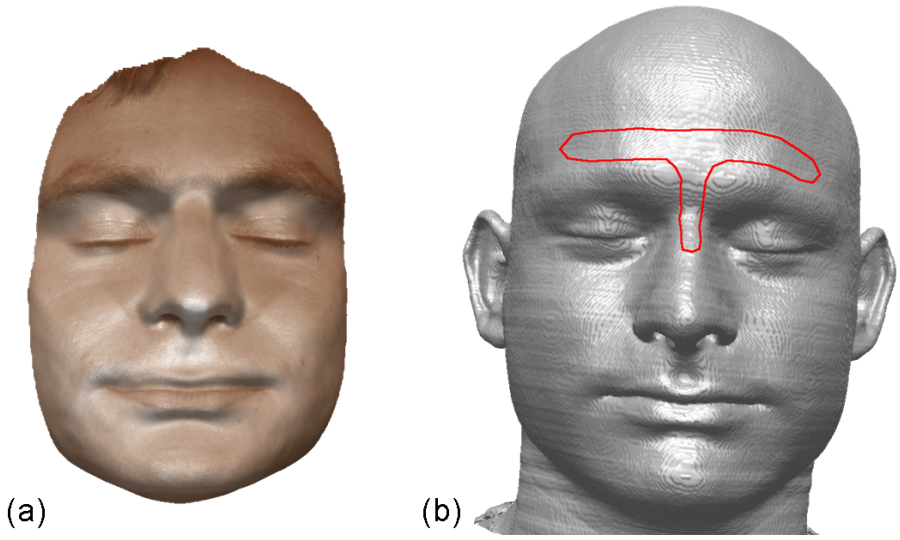

Figure 7.11: Holographically gained facial model (a) and facial model extracted from the CT data with the region used for the registration marked in red (b).

The holographic method for facial measurement delivers contact free information of living people in a sitting position. In addition to this, the gained soft tissue thickness information is not restricted to anatomical landmarks but is accessible for the whole face. If a large database containing holographic facial and CT skull models would be available and a face is to be reconstructed belonging to the skull of an unknown dead person, the following procedure could be performed. Every skull in the corresponding category of the database could be transformed into the skull of the unknown person using an elastic transformation. If the same transformation would be applied to the face, the result would be a facial model of the person in the database if their skull would be of the shape of the one of the unknown person. If one would repeat this procedure for every data set in the corresponding category and would average the results, one would get the average face corresponding to the skull of the unknown person in an automated process without directly measuring any soft tissue thickness. A demonstration of such a procedure would exceed the scope of this work. Separate theses will be written on that topic.

Just to demonstrate the possibilities for soft tissue measurement, three Helmer points were selected and the soft tissue thickness was measured for all 18 probands. The registration step between the holographically gained data and the surface extracted from the CT data is a critical step, as explained above. Due to this, three points in the region of the face used for the registration (nasal bone and lower forehead) were chosen. In that region the registration can be assumed to be correct. The Helmer points five (Glabella), six (Nasion) and eight (Rhinion) were chosen (see figure 7.12).

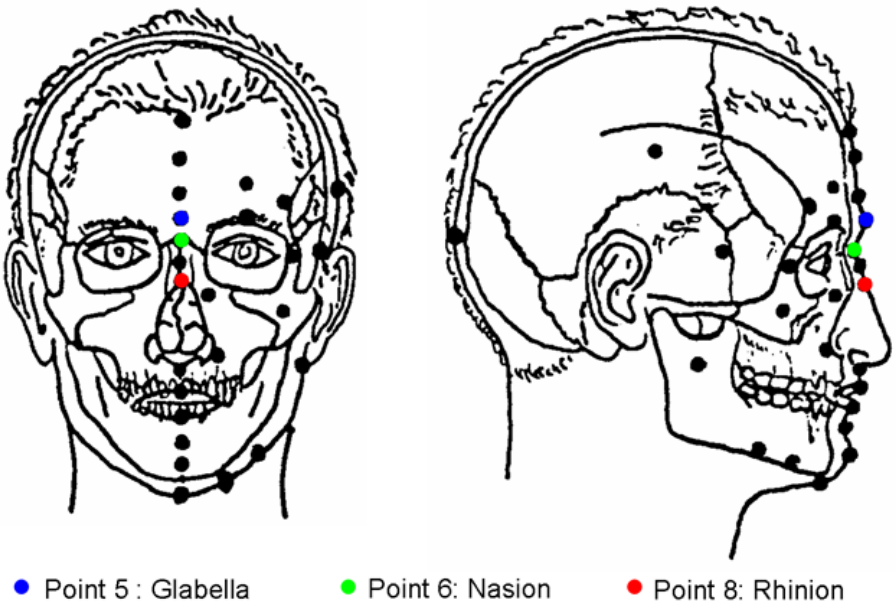

Figure 7.12: Schematic drawing of the face and marking of the Helmer points. The colored points were reevaluated with the help of 18 data sets consisting of holographic facial models and CT skull models. Source [Hel84]

Helmer found that the soft tissue thicknesses for different probands did not follow a normal distribution. This is the reason he used the median for characterization of the average value. He additionally stated a 95 % confidence interval (CI) for the values. The following tabular shows the soft tissue thickness results gained in this study and their median value:

	Male			Female		
	Point 5	Point 6	Point 8	Point 5	Point 6	Point 8
	4.7	6.7	1.6	7.1	7.2	2.8
	7.0	8.5	1.7	6.2	7.1	2.3
	5.6	9.2	2.7	4.9	7.3	1.6
	5.2	7.9	3.1	4.4	6.4	2.3
	5.4	8.4	2.3	6.9	6.7	2.1
	5.5	8.7	2.2	5.0	6.5	2.1
	5.2	8.0	2.2	5.3	7.4	2.9
	5.7	7.5	2.2	5.1	5.7	2.4
	5.5	6.9	2.3	7.0	6.7	2.3
Median	5.5	8.0	2.2	5.3	6.7	2.3

Helmers results are displayed in the following table [Hel84]:

	Male		Female	
Point	Median	CI (95%)	Median	CI (95%)
5	5.7	5.5-6.5	5.5	5.0-5.5
6	8.2	7.0-8.9	6.9	6.0-7.0
8	2.3	1.0-2.8	2.3	1.5-2.7

The results gained in this study differ only slightly from the results from Helmer. The values for point five and six are 0.2 mm smaller than Helmers values, the ones for point eight are 0.1 mm smaller for the male and equal for the female probands. All of them lie easily in the 95% confidence interval stated by Helmer. For a quantitatively appraisal of the results, several error sources should be considered:

- Resolution of the CT data
- Variation in the CT surfaces through different segmentation parameters
- Registration between the holographic and CT face models
- Statistical deviations due to the small number of probands
- Anatomical localization of the Helmer points

A detailed error analysis would exceed the scope of this work. It is just to be mentioned that even if the measured soft tissue thickness of the anatomical landmarks cannot reach the resolution of alternative methods, mainly due the limited resolution of the CT scan, the fact that soft tissue information of the whole face is accessible is to be considered a tremendous advantage, since it can lead to a fully automated reconstruction of the whole face.

The buildup of a broad database for facial reconstruction includes faces from various ethnic origin. The question arises if the holographic facial measurement system, which works with the backscattered light from the face, works equally good for different skin colors. To demonstrate the general applicability of the system, for the first time a person with dark skin color was recorded (see figure 7.13). The newly developed algorithms allowed the creation of a model with a remarkably good quality, despite the reduced amount of backscattered light. The fact that the quality of the facial models does not depend on the color of the skin broadens the applicability of the holographic system for forensic and medical applications.

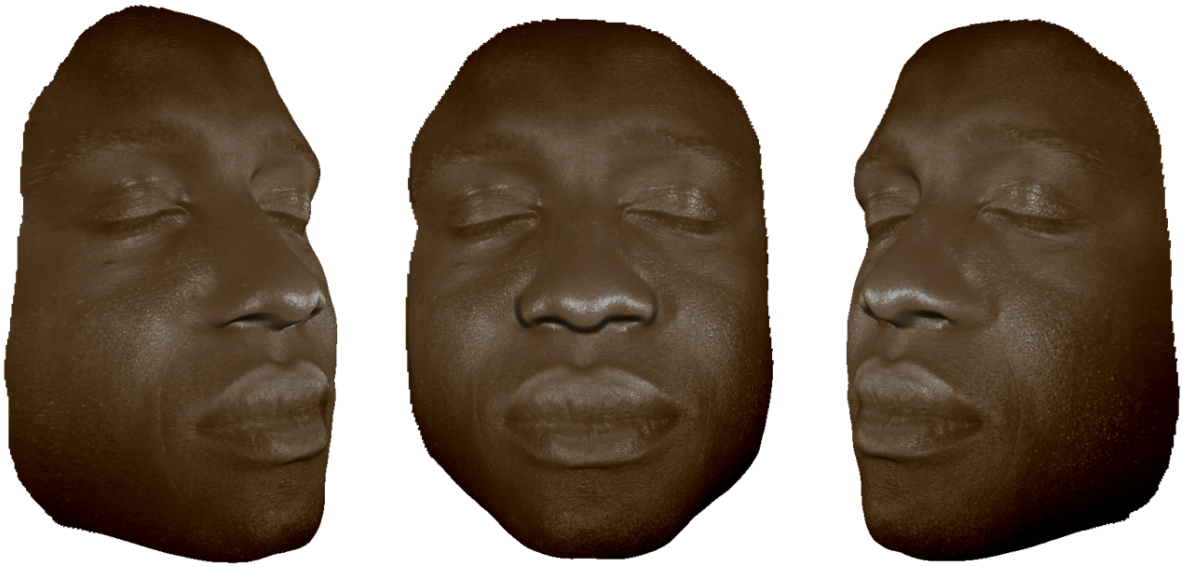

Figure 7.13: Screenshots of the three-dimensional computer model of a holographically recorded test person with dark skin color.

7.4 Summary

Three different kinds of applications for the holographically gained three-dimensional face models were presented in this chapter.

In the field of medical applications, the improved surface extraction procedures allow for quantitatively analyzing the subtle distinctions before and after a face lifting surgery.

The high-resolved, precisely textured computer models are of great value for the documentation and scientific analysis of archeological findings like the skull cap of the neanderthal man. The most realistic impression gained through the textured computer models or the hologram itself give rise to a wide distribution of unique findings without endangering them.

Also the advantages of holographic facial models in the field of forensic facial reconstruction was pointed out. A database containing skull data from a CT and face data from a hologram would exceed classical soft tissue measurement. First steps to build such a database including the possibilities of soft tissue measurement with the software RapidForm were presented. Additionally it was shown that high quality facial models could be created independent of the skin color of the proband.

Chapter 8

Conclusion

This chapter summarizes the methods and results presented in this thesis and discusses future developments.

8.1 Contrast based surface extraction

The general theme of this thesis was the extraction of three-dimensional information from a focus series. A focus series is a set of images where each one contains focused and unfocused contours. Localizing all contours in focus one gains three-dimensional information about the surface of the imaged object.

Such focus series can be produced with classical imaging devices through subsequently varying the focus settings, as for example in the established *shape-from-focus* method in white light microscopy. Another method to produce a focus series is holography. In holography, the whole three-dimensional information about the recorded object is stored by superimposing the back-scattered light coherently with a reference wave. The amplitude and the phase information of the scattered light field are stored as an interference pattern in the photosensitive emulsion of a holographic plate. The information can be decoded by illuminating the hologram with the complex conjugate reference beam, which produces the so-called real image, a three-dimensional light field with one-to-one correspondence to the recorded object.

This kind of acquisition of three-dimensional data has the tremendous advantage of storing the high-resolved information in an extremely short amount of time of the 35 ns laser pulse duration. No other device for three-dimensional imaging is capable of storing such

a huge amount of highly resolved data in such a short amount of time. However, the only way to get access to the stored information is reconstructing the hologram. This is done either optically, if the hologram was stored on an analogue recording material, or numerically, if it was recorded with a digital recording device. The latter produces two-dimensional projections of the recorded light field at arbitrary axial positions. The optical reconstruction creates a light field, which has to be digitized into a set of two-dimensional projections. One way or the other, the result is a set of images containing focused and unfocused regions similar to the focus series gained through subsequent recording. Investigations on the surface extraction out of focus series, in this thesis only regarded in relation to the holographic recording of faces, are of great interest in all kind of holographic applications, covering areas from acoustic holography to digital holographic microscopy, as well as adjacent fields of research in computer vision.

To make the proposed methods feasible for all different kinds of focus series, their general performance was investigated in detail. General characteristics of imaging systems, like image noise or imaging geometry were taken into account. In order to do so, simulated focus series were produced, where noise content or imaging geometries could be influenced in a controlled way. It was the first time that a realistic simulated focus series showing a holographically recorded face was reported. Since the simulation could not be done with standard methods due to the non-constance of the holographic point spread function, a new simulation method was developed. The simulated data set gave a ground truth for the quantitative comparison of different surface extraction algorithms and allowed for a systematic analysis of influencing factors.

Basis of the surface extraction is measuring the amount of sharpness present in individual images. For this purpose, twelve different focus criteria were tested for their suitability for holographic facial measurement. The novel XSML operator, based on but outperforming the established SML operator, was found to perform best. Besides the choice of the focus criterion, also an optimization of the neighborhood size used for contrast measurement is necessary. While choosing a constant neighborhood size, one always has to compromise between lateral resolution and robustness against noise. A multi-scale approach based on a newly developed decision criterion adapts the neighborhood size to the conditions of the data and allows for an optimal lateral resolution while minimizing the influence of noise. The resulting models have such a high lateral resolution that three-dimensional information about single hairs is available.

Improvements of the actual procedure of surface extraction were also developed during this thesis. The most profound innovation was the introduction of a Gaussian fit on the

focus profile. It was able to reduce the influence of noise and eliminate artifacts created by the usage of pixel neighborhoods for contrast measurement. Primarily it was also able to produce continuous surfaces not restricted to the discrete digitization positions. It is the first time that holographic facial models with a resolution below the digitization step size were presented. This Gaussian fit worked better for producing continuous surfaces than established methods like Gaussian interpolation or parabolic regression.

Another novel method for surface extraction introduced in this thesis is the maximization of focus profile symmetry instead of maximizing the focus profile itself. It is able to eliminate artifacts nearly equally as good as the Gaussian fit but it is unable to produce continuous values. Nevertheless, if a quick first estimation of the surface is desired, it is extremely helpful.

A further development concerning previous work was also performed, dealing with the fact that the image formation, which is characterized through the holographic point spread function, does take place along different directions, depending on the position of the focus relative to the optical axis. Optimal results are reached if the maximization of the focus profile is performed in this direction, and not, as in standard procedures, parallel to the optical axis. A determination of these directions can be realized by extracting substacks around an initial estimate of the surface and using a gray value based registration procedure. This method works without ab-initio knowledge about the imaging system, also for regions underlying shadowing effects. Outstanding improvements can be achieved if either the recorded volume is large, which inherently leads to large angles of the point spread function, or the used focus criterion is very sensitive to a mismatch between the directions of image formation and focus maximization. For the standard application of holographic facial measurement, particulary if the novel XSML focus criterion is used, the appliance of this new method is helpful but not obligatory to achieve good results.

A fourth, advanced technique for surface extraction using tilted neighborhoods for contrast calculation was demonstrated. They are tilted in such a way that their slope corresponds to the slope of a first estimate of the surface. Especially in regions of steep slope, the signal-to-noise ratio and at the same time the quality of the extracted surface are improved. In facial models mainly the sides of the nose and the outer regions of the cheeks gain from this procedure. If objects of a more complex geometry with a great variation in slopes are recorded, the full advantage of this method comes out.

With a combination of the surface extraction methods described earlier, a deviation between the extracted and the reference surface below 0.25 mm for every considered amount

of added noise could be achieved. This is lower than half the inter-tomogram distance of 0.571 mm.

Despite carefully accomplished surface extraction procedures, some remaining noisy fluctuations cannot be avoided. A known method for surface smoothing while maintaining characteristic surface features was implemented and the optimal parameters for the smoothing of holographic models were discussed.

Besides the ongoing qualitative evaluation of the newly developed methods with the help of focus profiles and height maps from holographically recorded data sets and the quantitative evaluation through the simulated data sets, an evaluation with a face-like test object was performed. The determined resolution of the holographic model was in all cases at least at half the inter-tomogram distance.

The last chapter reported about different applications for holographically gained three-dimensional models. For the first time a face lifting surgery was documented three-dimensionally using hologram tomography. Due to the small absolute changes, high-resolved information is necessary to overcome the subjective photographic documentation. In the field of archaeology, precisely textured computer models of the skull cap of the neanderthal man were presented. Additionally, the concept and first steps of forensic facial reconstruction using holographic facial models were explained.

Summarizing it can be said that the novel algorithms introduced here deliver solutions for a wide range of challenges in surface extraction from focus series. In regard to the creation of holographically gained facial models an optimized procedure was presented which works with minimal user interaction. The improved resolution of the models leads to new kinds of applications, like the mentioned documentation of a face-lifting surgery.

8.2 Outlook

During the last years the increasing computational power and memory capacity of even standard computers helped three-dimensional data to make its way into a wide range of applications. Especially three-dimensional facial models gained great interest in various fields. Besides the previously mentioned medical and forensic applications, three-dimensional facial models are also of great interest for security purposes or animated movies.

Some future developments of the holographic system for facial measurement would fur-

ther consolidate its position among competing products for three-dimensional topometry. Apart from the fully automated hologram recording procedure, the already heightened automation for the chemical processing of the hologram and the automated image processing and surface extraction presented in this thesis, the optical reconstruction of the hologram and digitization process could be enhanced. At the moment, the optical fine alignment of the holographic plate to the complex conjugate reference beam is done manually. An automated adjustment as well as a faster digitization with reduced digitization noise would improve the manageability of the system and the quality of the resulting models considerably.

One of the outstanding advantages of the holographic system for facial measurement is the precisely fitting gray scale texture, inherently present in the digitized data. An even more realistic impression of the computer models would be achieved, if the texture had true color information. If additionally to the holographic recording digital color photographs are taken quasi-simultaneously, the precisely fitting gray scale texture could be used for an automated determination of the transformation matrix between the digital photograph and the holographic model. Additional precise color information could be transferred to the three-dimensional model and would enlarge its applicability in many fields, like the medical, three-dimensional documentation of the progressive course of swelling after a surgery, where supplementary color information could correlate the swelling to redness distribution.

Using smaller wavelengths for the holographic recording a whole different area of applications could be entered. Pulsed holography with x-rays for example would make three-dimensional imaging of living cells in their natural environment possible.

Appendix A

Focus criteria

This chapter gives a detailed description of the focus criteria tested in chapter 4. Additionally, their computational cost is determined. The focus criteria presented here are evaluated in a pixel neighborhood $U_d(x, y)$ of a point (x, y) with a size of $d \times d$ pixels (see section 2.2.6):

$$U_d(x, y) = \{(\xi, \eta) | \ |\xi - x| \leq d \wedge |\eta - y| \leq d\} \tag{A.1}$$

The gray value at a point (x, y) is indicated as $I(x, y)$.

Tenengrad Since the quality of focus affects edge characteristics, it is natural to use an edge detector for computing the quality of focus. A method investigated by Tenenbaum [Ten70], described in [Kro87], uses the Sobel edge operators to approximate gradients in horizontal and vertical direction. The Sobel operator estimates the gradient in one direction while the image is smoothed in the perpendicular direction [Jäh02, p. 350]. Using the Sobel operator in x and y direction can be realized by convolving the image with the following convolution kernels:

$$S_x = \frac{1}{4}\begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \qquad \text{and} \qquad S_y = \frac{1}{4}\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \tag{A.2}$$

The focus criterion is stated as the sum of the gradient magnitude over the neighborhood $U_d(x, y)$:

$$F_{Tenengrad}(x, y) = \sum_{(\xi,\eta)\in U_d(x,y)} (S_x \star I(\xi, \eta))^2 + (S_y \star I(\xi, \eta))^2 \tag{A.3}$$

Originally, only values $S(x, y)$ lying above a pre-defined threshold T were summed. Also other focus criteria, i. e. the SML criterion invented by Nayar [NN94], work with

such thresholds. Since threshold selection inevitably requires heuristic choices and depends on image content, following Krotkov [Kro87], no thresholds were considered in this thesis.

Squared Gradient Instead of using Sobel operators to compute the approximation of the gradient, it can be calculated by just using intensities of adjacent pixels:

$$F_{SquGrad}(x,y) = \sum_{(\xi,\eta)\in U(x,y)} \left[I\left(\xi+1,\eta\right) - I\left(\xi,\eta\right)\right]^2 + \left[I\left(\xi,\eta+1\right) - I\left(\xi,\eta\right)\right]^2 \quad \text{(A.4)}$$

Absolute Gradient By squaring, the former criterion enhances the larger gradients over the small ones. Using the absolute difference instead of the square avoids this effect [Jar76]. The criterion can be described as:

$$F_{AbsGrad}(x,y) = \sum_{(\xi,\eta)\in U(x,y)} \left|I\left(\xi+1,\eta\right) - I\left(\xi,\eta\right)\right| + \left|I\left(\xi,\eta+1\right) - I\left(\xi,\eta\right)\right| \quad \text{(A.5)}$$

[Kro87] calls an equivalent criterion Sum-Modulus-Difference.

Laplacian A function of second order derivative (Laplacian) suppresses more strongly the lower frequencies than a first order derivative (Gradient) [LG82]. In the case of discrete two-dimensional data the Laplacian can be expressed as follows:

$$\begin{aligned} F_{Laplacian}(x,y) \quad = \quad & \sum_{(\xi,\eta)\in U(x,y)} \left[I\left(\xi+1,\eta\right) + I\left(\xi-1,\eta\right) - 2\cdot I\left(\xi,\eta\right) + \right. \\ & \left. I\left(\xi,\eta+1\right) + I\left(\xi,\eta-1\right) - 2\cdot I\left(\xi,\eta\right)\right] \end{aligned} \quad \text{(A.6)}$$

Sum Modified Laplacian Nayar [NN94] invented a focus criterion based on the Laplacian with some modification. He does not take the immediately adjacent pixels into account for calculating the second order derivative but introduces an additional variable called s and calculates the second order derivative for example in x-direction as follows:

$$\frac{\partial^2 I(x,y)}{\partial x^2} = I(x+s,y) + I(x-s,y) - 2\cdot I(x,y) \quad \text{(A.7)}$$

This variable s can be adjusted to the characteristic size of texture elements in the recorded scene[1]. If the second derivatives in x and y directions have opposite signs they tend to cancel each other and the Laplacian may have an unstable behavior. To overcome this problem the absolute values of the derivatives in x- and y-direction are added. Nayar calls this the modified Laplacian (ML). Since the modified values are added over a neighborhood, he calls the result Sum-modified-Laplacian (SML):

$$\begin{aligned} F_{SML}(x,y) \quad = \quad & \sum_{(\xi,\eta)\in U(x,y)} \left|I\left(\xi+s,\eta\right) + I\left(\xi-s,\eta\right) - 2\cdot I\left(\xi,\eta\right)\right| + \\ & \left|I\left(\xi,\eta+s\right) + I\left(\xi,\eta-s\right) - 2\cdot I\left(\xi,\eta\right)\right| \end{aligned} \quad \text{(A.8)}$$

[1] It was set to $s = 3$ for all computations in this thesis.

Variance The variance function can be used since well-focused images are expected to have greater variations in gray levels than those that are blurred:

$$F_{Var}(x,y) = \frac{1}{N^2} \sum_{(\xi,\eta)\in U(x,y)} \left[I\left(\xi,\eta\right) - \mu\right]^2, \tag{A.9}$$

where μ is the mean gray value given by:

$$\mu = \frac{1}{N^2} \sum_{(\xi,\eta)\in U(x,y)} I\left(\xi,\eta\right) \tag{A.10}$$

Normalized Variance Dividing the variance by the mean gray value compensates for changes in the average image brightness.

$$F_{NorVar}(x,y) = \frac{1}{\mu^2}\frac{1}{N^2} \sum_{(\xi,\eta)\in U(x,y)} \left[I\left(\xi,\eta\right) - \mu\right]^2, \tag{A.11}$$

μ is the mean gray value as defined in equation A.10.

Entropy The gray value histogram of a blurred image tends to a uniform distribution, whereas i. e. a sharply focused edge has a bimodal histogram. The population entropy E is a measure of the uniformity of the histogram. Let $P(I)$ denote the frequency of the occurrence of gray value I. Define the histogram entropy by

$$F_{Entropy}(x,y) = E = -\sum_{I} P(I)\ln P(I) \qquad \text{with} \qquad P(I) \neq 0 \tag{A.12}$$

E takes its maximum value if all $P(I)$ are equal, and minimum value if $P(I)$ is zero for all but one value of I. By this definition, a blurred edge image will have greater entropy than the sharp edge image, so the criterion is to minimize E [Kro87]. All the other criteria presented here have to be maximized.

Sum Square Anti Gaussian This function, abbreviated with SSAG, was invented by Yim [YB98]. Defocusing can be modelled as blurring with a Gaussian filter. A highpass filtered image can be computed as the difference between the original image and the image filtered with a Gaussian blur filter. [YB98] calls such an image anti-Gaussian filtered and indicates the resulting intensities with $I_a(x,y)$. The sum over the square of this values defines a focus criterion:

$$F_{SSAG}(x,y) = \sum_{(\xi,\eta)\in U(x,y)} \left[I_a\left(\xi,\eta\right)\right]^2 \tag{A.13}$$

Visibility This feature is inspired from the human visual system and is defined as

$$F_{VI}(x,y) = \sum_{(\xi,\eta)\in U(x,y)} \frac{\left|I\left(\xi,\eta\right) - \mu\right|}{\mu^{\alpha+1}} \tag{A.14}$$

where μ is the mean intensity value (see A.10) and α is a visual constant ranging from 0.6 to 0.7 [LKW02]. For the analysis in this thesis it was set to 0.65.

Cross Sum Modified Laplacian (XSML) Since preliminary studies showed good results using the SML-operator proposed by Nayar [NN94], this and the following criterion, which are newly developed, follow a similar principle.

The Laplacian takes only direct neighbors in the x- and y-directions into account. Additionally, also the diagonal neighbors could be included, weighted with a factor of $\frac{1}{\sqrt{2}}$ to compensate for their larger distance to the central point. Figure A.1 illustrates the different schemes.

$$\begin{aligned} F_{XSML}(x,y) = \sum_{(\xi,\eta)\in U(x,y)} & |I(\xi+s,\eta)+I(\xi-s,\eta)-2\cdot I(\xi,\eta)| + \\ & |I(\xi,\eta+s)+I(\xi,\eta-s)-2\cdot I(\xi,\eta)| + \\ & \frac{1}{\sqrt{2}}|I(\xi+s,\eta+s)+I(\xi-s,\eta-s)-2\cdot I(\xi,\eta)| + \\ & \frac{1}{\sqrt{2}}|I(\xi-s,\eta+s)+I(\xi+s,\eta-s)-2\cdot I(\xi,\eta)| \end{aligned} \quad \text{(A.15)}$$

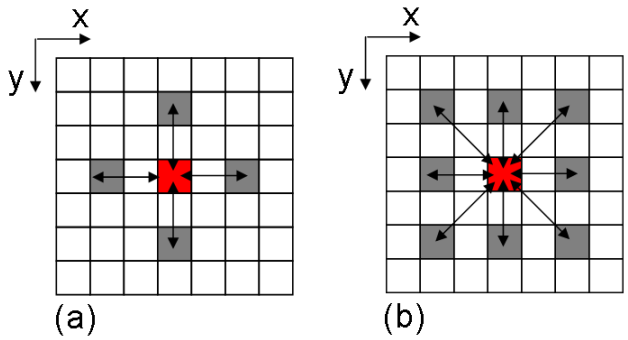

Figure A.1: Sketch of the pixel scheme used for computing the modified Laplacian (a) and the newly developed cross modified Laplacian (b).

Sum Modified Fourth Order Derivative Instead of using the second order derivative with a specified step-size s, the fourth order derivative could be used:

$$\frac{\partial^4 I}{\partial x^4} = I(x+2s,y)-4\cdot I(x+s,y)+6\cdot I(x,y)-4\cdot I(x-s,y)+I(x-2s,y) \quad \text{(A.16)}$$

This leads, with analogue modifications as in the SML-operator, to a focus criterion abbreviated with SM4D:

$$F_{SM4D}(x,y) = \sum_{(\xi,\eta)\in U(x,y)} \left|\frac{\partial^4 I(\xi,\eta)}{\partial \xi^4}\right| + \left|\frac{\partial^4 I(\xi,\eta)}{\partial \eta^4}\right| \quad \text{(A.17)}$$

Computational costs

Aside from the entropy, all focus criteria sum values over a neighborhood of $d \times d$ pixels. This can mathematically be expressed as a convolution with a kernel

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \tag{A.18}$$

here exemplarily displayed for d=3. This leads to a number of additions for an image containing N pixels of $\#\text{add} = d^2N$. Significant computational advantages may be realized in cases involving a separable kernel. Separable kernels permit the replacement of the 2D operation by two successive 1D operations [SOS00]:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \tag{A.19}$$

This reduces the number of additions to $\#\text{add} = 2dN$. This simplification can be made for all criteria except the ones using the mean gray value of the same neighborhood, which are the variance, normalized variance and the visibility. For these criteria the computational cost adds up to $\#\text{add} = 2dN + d^2N$, which is a factor $\frac{d+2}{2}$ larger than that of the remaining criteria. The difference in the number of operations aside from the summation over the neighborhood can be neglected.

Appendix B

One-dimensional binomial filter

The general purpose of filtering is noise reduction. This is normally done by smoothing and, more specifically, using a appropriate low-pass filter. To accomplish the desired noise reduction, each point is examined with respect to its neighboring points, and, if its intensity value is found to differ sharply from those of its neighbors, this difference may be reduced by an adjustment of the value of the point. It is replaced by a weighted average of the neighboring points [SOS00].

The filter used for one-dimensional smoothing in this thesis is the binomial filter. The weights for the weighted averaging are set proportional to the binomial coefficients. The weighted averaging can be expressed with the following convolution kernels [SOS00]:

$$B_2 = \frac{1}{2^2}\begin{bmatrix} 1 & 2 & 1 \end{bmatrix} \tag{B.1}$$

$$B_4 = \frac{1}{2^4}\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \end{bmatrix} \tag{B.2}$$

$$\vdots \tag{B.3}$$

$$B_n = \frac{1}{2^n}\left[\binom{n}{n} \quad \binom{n}{n-1} \quad \cdots \quad \binom{n}{0} \right] \tag{B.4}$$

The order n of the filter has always to be even. Odd numbers are rounded to the next even number. In comparison to a moving average filter working without weighted averaging, the binomial filter shows a flatter amplitude response, which leads to less filtering artifacts. The binomial filter is for large n proportional to a Gaussian filter.

Bibliography

[Abb02] A. Abblott. Hologram technique to flush out features of bog body. *Nature*, 418:908, 2002.

[AC01] M. Asif and T. Choi. Shape from focus using multilayer feedforward neural network. *IEEE Transactions on Image Processing*, 10:1670–1675, 2001.

[AC05] M. B. Ahmad and T. Choi. A heuristic approach for finding the best focused shape. *IEEE Transactions on Circuits and Systems for Video Technology*, 15:566–574, 2005.

[Ade99] M. Aderholz et. al. A measurement of the holographic minimum-observable beam branching ratio in the FERMILAB 15-ft bubble chamber. *Nuclear Instruments and Methods in Physics Research A*, 421:1–11, 1999.

[AJM98] N. Asada, H. Jujiwara, and T. Matsuyama. Edge and depth from focus. *International Journal of Computer Vision*, 28:153–163, 1998.

[Ans70] D. A. Ansley. Techniques for pulsed laser holography of people. *Applied Optics*, 9:815, 1970.

[BGF+03] J. Bongartz, D. Giel, S. Frey, A. Thelen, and P. Hering. Hochauflösende dreidimensionale Gesichtsprofilvermessung mit kurzgepulster Holographie. In *Physikalische Methoden der Laser und Medizintechnik*, pages 51 – 56. VDI Verlag, 2003.

[BGZ+06] R. Bakanas, G. A. Gudaitis, S. J. Zacharovs, D. B. Ratcliffe, S. Hirsch, S. Frey, A. Thelen, N. Ladriere, and P. Hering. Ultrafast holographic topometry of the face for medical applications. *to appear in: Journal of Optical Technology*, 2006.

[BHK03] J. Burke, C. F. Hess, and V. Kebbel. Digital holography for instantaneous spray diagnostics on a plane. *Particle and Particle Systems Characterization*, 20:183–192, 2003.

[BM92] P. J. Besl and N. D. McKay. A method for registration of 3d shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:239–269, 1992.

[Bon02] J. Bongartz. *Hochauflösende dreidimensionale Gesichtsprofilvermessung mit kurzgepulster Holographie.* PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät der Heinrich-Heine-Universität Düsseldorf, 2002. http://www.ub.uni-duesseldorf.de/home/etexte/diss/diss_files/192.pdf.

[Bri05] U. Brinkmann. Mobile holography system records 3-d portraits for planning surgery. *Laser Focus World*, 41:28, 2005.

[Bür04] A. Bürkle. *Optische 2,5D-Rekonstruktion Mikroskopischer Objekte.* Logos Verlag, Berlin, 2004.

[CBS00] F. Chen, G. M. Brown, and M. Song. Overview of three-dimensional shape measurement using optical methods. *Optical Engineering*, 39:10–22, 2000.

[CY00] T. Choi and J. Yun. Three-dimensional shape recovery from focused image surface. *Optical Engineering*, 39:1321–1326, 2000.

[DW88] T. Darrell and K. Wohn. Pyramid based depth form focus. *Proc. CVPR*, pages 504–509, 1988.

[EA93] J. Eichler and G. Ackermann. *Holographie.* Springer Verlag, Berlin, 1993.

[ELS+04] S. Eisebitt, J. Lüning, W. F. Schlotter, M. Lörgen, O. Hellwig, W. Eberhardt, and J. Stöhr. Lensless imaging of magnetic nanostructures by x-ray spectro-holography. *Nature*, 432:885–888, 2004.

[FB97] P. L. Fernando and J. Beyerer. Datenfusion zur Gewinnung hochwertiger Bilder in der automatischen Sichtprüfung. *at - Automatisierungstechnik*, 45:480–489, 1997.

[FBG+03] S. Frey, J. Bongartz, D. Giel, A. Thelen, and P. Hering. Ultrafast holographic technique for 3D in situ documentation of cultural heritage. *Proc. SPIE 5146*, 194, 2003.

[FBT+05] C. U. Fritzemeier, H. G. Bull, A. Thelen, J. Bongartz, P. Hering, S. Frey, S. Hirsch, and N. Ladriere. Hologramm-Tomographie als Qualitätskontrolle bei ästhetischen Gesichtsoperationen. *Magazine for Aesthetic Surgery*, 3:14–25, 2005.

[FCC+91] L. Firestone, K. Cook, K. Culp, N. Talsania, and K. Preston. Comparison of autofocus methods for automated microscopy. *Cytometry*, 12:195–206, 1991.

[FGA+05] P. Ferraro, S. Grilli, D. Alfieri, S. De Nicola, A. Finizio, G. Pierattini, B. Javidi, G. Coppola, and V. Striano. Extended focused image in microscopy by digital holography. *Optics Express*, 13:6738–6749, 2005.

[Fre05] S. Frey. *Three-dimensional facial measurement by portrait holography and texture-based focus detection.* PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät der Heinrich-Heine-Universität Düsseldorf, 2005. http://www.ub.uni-duesseldorf.de/home/etexte/diss/diss_files/1166.pdf.

[FSSS04] J. P. Fugal, R. Shaw, E. W. Saw, and A. V. Sergeyev. Airborne digital holographic system for cloud particle measurement. *Applied Optics*, 43:5987–5995, 2004.

[FTHH06] S. Frey, A. Thelen, S. Hirsch, and P. Hering. Shape measurement using pulsed optical holography. *Laser Technik Journal*, 1:42–45, 2006.

[FVdVB+04] B. Forster, D. Van de Ville, J. Berent, D. Sage, and M. Unser. Extended depth-of-focus for multi-channel microscopy images: A complex wavelet approach. In *Proceedings of the Second 2004 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, volume 65, pages 660–663, Arlington VA, USA, 2004.

[Gab48] D. Gabor. A new microscopic principle. *Nature*, 161:777–778, 1948.

[Gat86] J. W. C. Gates. The influence of holography on measurement technology. *Journal of Physics E-scientific Instruments*, 19:998–1007, 1986.

[GBH02] D. Giel, J. Bongartz, and P. Hering. Ultrafast holographic topometry for medical applications. In *International Symposium on Photonics in Measruement*, pages 333–338. VDI Berichte 1694, 2002.

[Gie03] D. Giel. *Hologram tomography for surface topometry.* PhD thesis, Mathematisch-Naturwissenschaftliche Fakultät der Heinrich-Heine-Universtität Düsseldorf, 2003. http://www.ub.uni-duesseldorf.de/home/etexte/diss/diss_files/598.pdf.

[GMS73] A. D. Gara, R. F. Majkowski, and T. T. Stapleton. Holographic system for automatic surface mapping. *Applied Optics*, 12:2172–2179, 1973.

[Gu00] M. Gu. *Advanced Optical Imaging Theory.* Springer-Verlag, Berlin, 2000.

[Har96] P. Hariharan. *Optical Holography. Principles, Techniques and Applications.* Cambridge University Press, 1996.

[Har99] G. G. Harigel. Measurement of bubbles in a superheated liquid. *Nuclear Instruments and Methods in Physics Research A*, 421:12–22, 1999.

[Hel84] R. Helmer. *Schädelidentifizierung durch elektronische Bildmischung. Zugleich ein Beitrag zur Konstitutionsbiometrie und Dickenmessung der Gesichtsweichteile.* Kriminalistik Verlag, Heidelberg, 1984.

[Hol02] C. Holden. A face for german bog man. *Sciences*, 297:1271, 2002.

[Jäh02] B. Jähne. *Digitale Bildverarbeitung.* Springer-Verlag, Berlin Heidelberg, 2002.

[Jar76] R. A. Jarvis. Focus optimization criteria for computer image processing. *Microscope*, 24:163–171, 1976.

[KFZS02] J. Kautsky, J. Flusser, B. Zitova, and S. Simberova. A new wavelet-based measure of image focus. *Pattern Recognition Letters*, 23:1785–1794, 2002.

[Koc05] G. Koch. Surgical planning and documentation in the 21st century. *Biophotonics*, 12:86, 2005.

[Kre05] T. Kreis. *Handbook of Holographic Interferometry.* Wiley-VCH, Weinheim, 2005.

[Kro87] E. Krotkov. Focusing. *International Journal of Computer Vision*, 1:223–237, 1987.

[Lad04] N. Ladriere. Optische und chemische Aspekte der hochauflösenden, vollautomatischen Hologrammentwicklung. Master's thesis, Fachbereich Photoingenierwesen und Medientechnik an der Fachhochschule Köln, 2004.

[LFT+04] N. Ladriere, S. Frey, A. Thelen, S. Hirsch, J. Bongartz, D. Giel, and P. Hering. Ultraschnelle holographische Gesichtsprofilvermessung mit vollautomatischer Hologrammentwicklung. In Hartmann, Kohl-Bareis, Hering, Lonsdale, Bongartz, and Buzug, editors, *"Aktuelle Methoden der Laser- und Medizintechnick"*, pages 272–274. VDE-Verlag, 2004.

[LG82] G. Ligthart and C. A. Groen. A comparision of different autofocus algorithms. *Proc. of IEEE Int. Conf. on Pattern Recognition*, pages 597–600, 1982.

[LK03] W. Lauterborn and T. Kurz. *Coherent Optics*. Springer-Verlag, Berlin, Heidelberg, 2003.

[LKW02] S. Li, J. T. Kwok, and Y. Wang. Multifocus image fusion using artificial neural networks. *Pattern Recognition Letters*, 23:985–997, 2002.

[LMM95] H. Li, B. S. Manjunath, and S. K. Mitra. Multisensor image fusion using the wavelet transform. *Graphical Models and Image Processing*, 57:235–245, 1995.

[Löf06] J. O. Löfken. Holographischer Blick aufs Gesicht. *Physik Journal*, 1:14, 2006.

[LU62] E. N. Leith and J. Upatnieks. Reconstructed wavefronts and communication theory. *Journal of the Optical Society of America*, 52(10):1123–1130, 1962.

[LU63] E. N. Leith and J. Upatnieks. Wavefront reconstruction with continuous-tone objects. *Journal of the Optical Society of America*, 53:1377–1381, 1963.

[Mah98] V. Mahajan. *Optical Imaging and Aberrations*. SPIE, Bellingham, Washington, 1998.

[May01] F. Mayinger. *Fundamentals of Holography and Interferometry*, chapter 4, pages 27–50. Springer-Verlag, 2001.

[Mei65] R.W. Meier. Magnification and third-order aberrations in holography. *Journal of the Optical Society of America*, 55:987–992, 1965.

[MMW73] E. Menzel, W. Mirande, and I. Weingärtner. *Fourier-Optik und Holographie*. Springer-Verlag, Wien, 1973.

[MWLJ04] L. Ma, H. Wang, Y. Li, and H. Jin. Numerical reconstruction of digital holograms for three-dimensional shape measurement. *Journal of Optics A: Pure and Applied Optics*, 6:396–400, 2004.

[NKCNA01] N. Ng Kuang Chern, Poo Aun Neow, and M. H. Ang. Practical issues in pixel-based autofocusing for machine vision. In *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, pages 2791–2796, Seoul, Korea, 2001.

[NM65] J. A. Nelder and R. Mead. A simplex-method for function minimization. *Computer Journal*, 7:308–313, 1965.

[NN94] S. K. Nayar and Y. Nakagawa. Shape from focus. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16:824–830, 1994.

[NN96] M. Noguchi and S. K. Nayar. Microscopic shape from focus using a projected illumination pattern. *Mathematical and Computer Modelling*, 24:31–48, 1996.

[NNS03] M. Niederöst, J. Niederöst, and J. Scucka. Automatic 3d reconstruction and visualization of microscoptic objects from a nanoscopic multifocus image sequence. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XXXIV-5/W10, 2003.

[Ost88] J. Ostrowki. *Holografie - Grundlagen, Experimente und Anwendungen.* Teubner Verlagsgesellschaft, Leipzig, 1988.

[PCA+02] M. A. Papadopoulos, P. K. Christou, A. E. Athonasiou, P. Boettcher, H. F. Zeilhofer, R. Sader, and N. A. Papadopulos. Three-dimensional craniofacial reconstruction imaging. *Oral Surgery Oral Medicine Oral Pathology*, 93:382–393, 2002.

[PM90] P. Perona and J Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[PTVF02] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C++.* Cambridge University Press, 2002.

[Ras06] W. Rasband. ImageJ. Image processing and analysis in java, 2006. http://rsb.info.nih.gov/ij/.

[Rog50] G. L. Rogers. Gabors diffraction microscopy: The hologram as a generalized zone-plate. *Nature*, 166:236–237, 1950.

[RW96] G. Rönneberg and H. Weißmantel. Mikrostrukturen im Fokus. *FM Zeitschrift für Elektronik, Optik und Mikrosystemtechnik*, 104:715–720, 1996.

[SC95] M. Subbarao and T Choi. Accurate recovery of three-dimensional shape from image focus. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 17:266–274, 1995.

[SDN04] Y. Sun, S. Duthaler, and B. J. Nelson. Autofocusing in computer microscopy: Selecting the optimal focus algorithm. *Microscopy Research and Technique*, 65:139–149, 2004.

[Sie68] L. D. Siebert. Large scene front-lighted hologram of a human subject. *Proceedings of the Institude of Electrical and Electronics Engineers*, 56:1242–1243, 1968.

[Sie86] A. E. Siegmann. *Lasers*. University Press, Oxfort, 1986.

[SJ02] U. Schnars and W. P. O. Jüptner. Digital recording and numerical reconstrution of holograms. *Measurement Science and Technology*, 13:R85–R101, 2002.

[SOdSV+97] A. Santos, C. Ortiz de Solorzano, J. J. Vaquero, J. M. Pena, N. Malplica, and F. Del Pozo. Evaluation of autofocus functions in molecular cytogenetics analysis. *Journal of Microscopy*, 188:264–272, 1997.

[SOS00] M. Seul, L. O'Gorman, and M.J. Sammon. *Practical Algorithms for Image Analysis: Discription, Examples, and Code*. Cambridge University Press, 2000.

[SPI+97] B. W. Schilling, T. C. Poon, G. Indebetouw, B. Storrie, K. Shinoda, Y. Suzuki, and M. H. Wu. Three-dimensional holographic fluorescence microscopy. *Optics Letters*, 22:1506–1508, 1997.

[ST98] M. Subbarao and J. Tyan. Selecting the optimal focus measure for autofocusing and depth-from-focus. *IEEE Transactions of Pattern Analysis and Machine Iintelligence*, 20:864–870, 1998.

[ST02] R. W. Schmitz and J. Thissen. *Neandertal. Die Geschichte geht weiter.* Spektrum Akademischer Verlag, Heidelberg, Berlin, 2002.

[Ste68] K. A. Stetson. Holographic surface contouring by limited depth of focus. *Applied Optics*, 7:987–989, 1968.

[Str04] H. Stratomeier. Methods of forensic facial reconstruction. Master's thesis, Academy of Visual Arts Maastricht, 2004.

[SWG99] T. Scheuermann, G. Wiora, and M. Graf. Topographical maps of microstructures generated by depth-from-focus techniques. In *Handbook of Computer Vision and Applications*, pages 387–410. Academic Press, 1999.

[TBG+05] A. Thelen, J. Bongartz, D. Giel, S. Frey, and P. Hering. Iterative focus detection in hologram tomography. *Journal of the Optical Society of America A*, 22:1176–1180, 2005.

[Ten70] J. M. Tenenbaum. *Accomodation in Computer Vision.* PhD thesis, Stanford University, 1970.

[TFH+06a] A. Thelen, S. Frey, S. Hirsch, L. Ladiere, and P. Hering. Ultra-fast facial topometry using pulsed holography. *Proc. SPIE 6081*, 60810F, 2006.

[TFH+06b] A. Thelen, S. Frey, S. Hirsch, N. Ladriere, and P. Hering. Ultrafast holographic topometry of the face for medical applications. *Medical Laser Application*, 21:9–14, 2006.

[The03] A. Thelen. Holographische Topometrie. Master's thesis, Mathematisch-Naturwissenschaftliche Fakultät der Rheinischen Friedrich-Wilhelms-Universität Bonn, 2003.

[Tho78] B. J. Thompson. Applications of holography. *Reports on Progress in Physics*, 41:634–669, 1978.

[TRU98] P. Thèvenaz, R. E. Ruttimann, and M. Unser. A pyramid approach to subpixel registration based on intensity. *IEEE Transactions on Image Processing*, 7:27–41, 1998.

[TW81] B. A. Tozer and J. M. Webster. Holography as a measuring tool. *Journal of Photographic Science*, 28:93–98, 1981.

[TW03] T. Tasdizen and R. Whitaker. Feature preserving variational smoothing of terrain data. In *IEEE Workshop on Variational Geometric and Level Set Methods in Computer Vision*, 2003.

[TWBO03] T. Tasdizen, R. Whitaker, P. Burchard, and S. Osher. Geometric surface smoothing via anisotropic diffusion on normals. *ACM Transactions of Graphics*, 22:1012–1033, 2003.

[VR02] VDI / VDE-Richtlinien. Opitsche 3d-Messsysteme: Blatt 2: Bildgebende Systeme mit flächenhafter Antastung, 2002. VDI/VDE 2634.

[VRM97] VRML97. The virtual reality modelling language: ISO / IEC 14772-1:1997, 1997. http://www.web3d.org/x3d/specifications/vrml/.

[WN95] M. Watanabe and S. K. Nayar. Telecentric optics for constant magnification imaging. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:1360–1365, 1995.

[Yar04] L. Yaroslavsky. *Digital Holography and Digital Image Processing: Principles, Methods, Algorithms*. Kluwer Academic Publishers, Dordrecht, 2004.

[YB98] C. Yim and A. C. Bovik. Mulitresolution 3-d range segmentation using focus cues. *IEEE Transactions on Image Processing*, 7:1283–1299, 1998.

[YFB02] R. M. Yin, P. J. Flynn, and S. L. Broschat. Position-dependent defocus processing for acoustic holography images. *International Journal of Imaging Systems and Technology*, 12:101–111, 2002.

[YOJS93] T. T. E. Yeo, S. H. Ong, Jayasooriah, and R. Sinniah. Autofocusing for tissue microscopy. *Image and Visison Computing*, 11:629–639, 1993.

[YSK99] N. Yokoya, T. Shakunaga, and M. Kanbara. Passive range sensing techniques: Depth from images. *IEICE Transactions on Information and Systems E82D*, 3:523–533, 1999.

[ZS68] G. E. Zech and L. D. Siebert. Pulsed laser reflection holograms. *Applied Physics Letters*, 13:417–418, 1968.

Acknowledgements

I would like to thank sincerely all people who have contributed to the holographic project and gave support for this thesis.

In the first place I would like to thank Prof. Dr. Peter Hering for the opportunity to work with the holography group, his support, academic guidance, and his confidence in my abilities.

I am also grateful to Prof. Dr. K. Schierbaum for taking over the part of the second referee.

I wish to acknowledge the support of the `caesar` foundation in Bonn for the opportunity to work in an excellent interdisciplinary environment with good working conditions.

In this context I want to cordially thank my colleagues Dr. Susanne Frey, Sven Hirsch and Natalie Ladriere for their steady support and for taking over many routine jobs while I was writing this thesis. Especially I would like to thank Sven Hirsch for the productive discussions and his numerous ideas. Natalie Ladriere I thank for her encouragement and help in many ways. I would also like to thank my former colleagues Prof. Dr. Jens Bongartz and Dr. Dominik Giel for their pioneering work on holographic facial measurement as well as all members of the *Holography and Laser Technology* group for the friendly atmosphere and the help they always provided.

Many collaborations from medical, forensic and archaeological science contributed to the development of the holographic system by arising new challenges and giving the opportunity to demonstrate the potential of the holographic system, which was greatly appreciated.

In particular I would like to thank Prof. Dr. Dr. Claus Udo Fritzemeier who made the evaluation of the face lifting surgery possible as one of a great variety of medical applications he supported since the beginning of the holographic project. I would also like to express my gratitude to Prof. Dr. Dr. Hans-Florian Zeilhofer and many of his coworkers, especially Dr. Dr. Katja Schwenzer and Thoralf Seewald, for establishing the holographic system at the University hospital in Basel and their valuable support and feedback.

For the unique possibility to work with as famous archaeological findings as the neanderthal man I would like to express my thanks to Dr. Michael Schmauder and all his coworkers form the Rheinisches Landesmuseum in Bonn and to Dr. Ralf Schmitz.

Among all the productive collaborations in forensic science I would like to show special appreciation to Frank Prieels, who provided CT data for the first attempt on soft tissue measurement.

Additionally I would like to show gratitude to Prof. Dr. Thomas Vetter from the University of Basel and his coworkers for a productive discussion, Burkahrt Wassmann from ObjektScan GmbH for his support and the establishment of contacts to test the holographic method for industrial applications.

My special thanks I express to Filip Schutyser from medicim for providing the test object mentioned in section 6.2 as well as the reference CT data set.

Last but not least I sincerely thank my parents for their encouragement over many years and would like to highly acknowledge the many helpful revisions suggested by Mascha Moeller-Herrmann as well as her support and encouragement.

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 09.06.2006

(Andrea Thelen)