# TOWARDS ROBUST 3D FACE RECOGNITION FROM NOISY RANGE IMAGES WITH LOW RESOLUTION

O. EBERS, T. EBERS, T. SPIRIDONIDOU, M. PLAUE, P. BECKMANN, G. BÄRWOLFF, AND H. SCHWANDT

Abstract. For a number of different security and industrial applications, there is the need for reliable person identification methods. Among these methods, face recognition has a number of advantages such as being non-invasive and potentially covert. Since the device for data acquisition is a conventional camera, other advantages of a 2D face recognition system are its low data capture duration and its low cost. However, the recent introduction of fast and comparatively inexpensive time-of-flight (TOF) cameras for the recording of 2.5D range data calls for a closer look at 3D face recognition in this context. One major disadvantage, however, is the low quality of the data aquired with such cameras. In this paper, we introduce a robust 3D face recognition system based on such noisy range images with low resolution.

## 1. Introduction

There is a number of applications that require the identification of humans. Examples include the authentification for a computer application or access control for high-security areas like an airport control tower. Face recognition systems are well suited for the task of human identification as they require less cooperation by the user than an iris or fingerprint scan. It is natural, robust and unintrusive, and the user is not required to remember any passwords or codes [2]. While the automatic face recognition on 2D images has been a research issue for several years, the recent development of 3D sensors has resulted in a considerable interest in methods for face recognition on range images.

In this project, we explored the state of the art of 3D face recognition and analyzed the advantages and disadvantages of several methods in regard to our project goals. Our work resulted in the development of a real-time system for the processing of three-dimensional data that is specialized on pattern recognition tasks. The algorithms we chose to implement were modified according to the project's needs and were reinvestigated and recombined.

The result of our work is a general development platform for 3D pattern recognition, specially designed for 3D face recognition on noisy and low-resolution data. In this context the platform can be extended for the recognition of any kind of 3D objects and it can be easily enhanced by the supplementary processing of two-dimensional intensity data.

---

*Date*: October 27, 2008.

2000 *Mathematics Subject Classification.* 68T45, 68U10.

*Key words and phrases.* 3D face recognition, time-of-flight camera, range data denoising, pattern recognition, pattern matching.

This project was funded by the European Regional Development Fund (ERDF).

In order to develop a face recognition system based on range images—for example acquired with the new 3D sensor type of time-of-flight (TOF) cameras—one has to turn particular attention on the quality of the data since such data is still very noisy and biased [23, 55]. For this reason our main goal was the development of algorithms that improve low-quality range data and process it efficiently and in real-time. Furthermore, our 3D face recognition system is constructed modularly, and can thus be easily adapted to data of higher quality obtained by other sensors.

To deal with low-quality range data, one has to (a) calibrate the imaging system with this particular application in mind, and (b) employ a pre-processing step that filters and smoothes the image data to achieve a quality suitable for feature extraction. The pre-processing algorithms have to account for the particular characteristics of the range data at hand since for example the noise model of a TOF sensor differs from the usual Gaussian white noise model assumed for the majority of standard denoising methods.

After acquiring and pre-processing the data, one wishes to extract discriminant and robust features. Again, it is important to consider the special nature of the data which for example forbids a robust calculation of the curvature. In particular, we have considered three features: the surface normals (or Gaussian map), the local binary pattern (LBP) and facial profiles (1D cross sections of the face).

The final face recognition task can then be accomplished by the usual classification methods such as Principal Component Analysis (PCA [4]), the Linear Discriminant Analysis (LDA [47]) or the Modified Linear Discriminant Analysis (MLDA [36]).

## 2. Related Work

While there exists extensive work on 2D face recognition, 3D face recognition is still a comparatively new research field. As has been shown in several experimental surveys [1, 14, 15, 32], in particular multi-modal approaches combining 2D and 3D features give results that surpass those of a simple 2D system. One main disadvantage of a face recognition system using range images, however, is the high cost of an industrial high resolution 3D scanner that is often needed to aquire the data. Most of the 3D face recognition work published until today use such laser or structured-light scanners [40, 63]. One cost-effective way to record range data is of course stereographic imaging [18]. However, it is well-known that such systems require a robust solution for the correspondence problem [26] and precise calibration. The also comparatively inexpensive time-of-flight imaging systems on the other hand have been used in a substantial number of application areas such as automated production [39, 46] or automotive applications [49, 58, 57, 67], while there are little studies that investigate the feasibility of TOF imaging for more complex recognition tasks like facial recognition. One major problem that arises with the use of cost-efficient 3D imaging systems like TOF is the low quality and resolution of the data. The main goal of our project was the implementation of a software pipeline capable of processing such data in real-time which will be described in the following sections. Although the system is taylored for 3D face recognition from TOF range images, it can be easily modified for other object recognition tasks based on low-quality data.

**Denoising of 2.5D Data.** The first processing step in our pipeline aims at the removal of noise present in typical range sensor data. The denoising of 3D data

and range data is a wide research field, and the choice of the appropriate denoising method depends on the noise and data characteristics. Typical denoising methods include the median filter [20], the moving least-squares method [43] and anisotropic diffusion [64], especially anisotropic smoothing of point sets [37] and surface meshes [30].

The wavelet transform is widely used for the purpose of image denoising and has been found to be a high-performance tool. For example, Cai et al. provide a useful MATLAB® framework [12] we used in our work (a detailed description can be found in [61] and [35]). In [62], a good introduction of complex wavelet transforms and their applications can be found.

**Point-to-Point Registration.** Another crucial step is the face registration which aims at detecting the exact face position and attempts to align the face with a position suitable for recognition tasks, which is usually the frontal view.

For the coarse registration, one common practice is to identify the position of three significant local features, for example the pupils and the nose tip. Afterwards, the features are mapped onto the corresponding features of a reference face by an affine map consisting of a rotation and a translation (cf. [66]). The parameters of this map express the feature points' relation to the corresponding points in the reference face. Via the affine map determined in this way, all data points are subsequently transformed to realize the coarse alignment along a position that is common for all faces in the database.

Common algorithms for fine alignment on the other hand is the family of Iterative Closest Point algorithms (ICP) which try to minimize the Hausdorff-distance (or one of its various relatives) between surfaces, and the Thin Plate Spline algorithm (TPS) [42]. Chen et al. [17] and Besl et al. [6] use ICP for scan registration during 3D model creation. In this context, ICP can be used for fine face alignment by fitting the face data onto the reference face. An exhaustive overview about ICP algorithms is provided by Rusinkiewicz et al. [59].

An interesting variant of the aforementioned (rigid) ICP-based registration was proposed by Bronstein et al. [9, 10], who used the Gromov–Hausdorff distance to compute inter-facial embeddings with minimal metric distortion, thereby enhancing the registration toolbox with the ability to match faces with different expressions against each other.

As a generalization to the Hausdorff distance, which is usually expressed as a min–max problem of the maximal distance of two sets (using the metric of their common embedding metric space), the Gromov–Hausdorff distance minimizes the maximal *inner-metric distortion* among all common ambient metric spaces and all possible embedding mappings, thus rendering the Gromov–Hausdorff distance independent of isometries. Since the computation of the functional as described here is intractable, the authors propose a discretization of the Gromov–Hausdorff distance in terms of mutual inter-surface embeddings, thus minimizing the metric distortion while embedding one surface onto the other and vice versa.

Once this distance functional and its corresponding embeddings are computed, the resulting distance value can be used directly for registration tasks by interpreting it as a similarity measure between faces. Moreover, the resulting embeddings carry an optimal inter-facial point-to-point correspondence regardless of the actual facial expressions involved. Still, as in the case of ICP, the process relies heavily on a previous rough initialization of a few feature points.

**Face Recognition.** For the final task of face recognition for 2.5D images or 3D models, three main methodologies can be identified: shape matching, feature-based and image-based techniques. A detailed overview on face recognition methods is provided in [2].

The first group consists of algorithms that iteratively try to map a 3D point cloud or a 3D mesh onto a reference point cloud or reference mesh, respectively [3, 5, 9, 10, 11, 19, 65]. The shape matching methods can be seen as pattern recognition methods without feature extraction. A test pattern is directly compared with the reference pattern, a feature extraction does not take place. The similarity measure—which is often implemented via a correlation measure—can be optimized by using a sufficiently large number of training samples. These approaches demand an extensive computational effort and an accurate point-to-point data registration and assume the existence of many correspondences between the reference model and the test data.

The modus operandi of feature-based methods correponds mostly to that of shape matching. However, with pattern matching, not the whole data is processed but appropriate subsets. For example, particular regions (eye, forehead, cheek, nose) or the nose profile of the face could be detected, extracted and processed [5, 16, 25, 42, 45]. Like shape matching methods, the feature-based methods demand a robust image registration, since the features are selected during a pre-processing step without the possibility to change their value later on.

Image-based methods attempt to extract the face data subset significant for face recognition with the aid of statistical learning techniques and without any human interaction. With feature-based methods, there are no or at least less pre-processing steps involved as is the case with image-based methods: All of the image information is used for statistical analysis. This methods have been very successful in the context of 2D face recognition [4, 47]. Since the TOF sensor data is a 2D range distribution and can thus technically be viewed as a conventional 2D image, it does not surprise that these techniques are also applied in this context. Introductions in state-of-the-art techniques of statistical learning and statistical pattern recognition can be found in [7, 21, 33].

In our approach, we use the statistical learning techniques with Local Binary Patterns (LBP [31, 50]) and surface normals, thereby proposing a combination of a feature-based and an image-based method: There is less information lost with this technique, since the whole image and not some preselected region is used for the feature extraction and subsequent classification. As a statistical learning method, we used the Principal Component Analysis (PCA [4]), the Linear Discriminant Analysis (LDA [47]) and the Modified Linear Discriminant Analysis (MLDA [36]) for classification.

As an alternative feature-based approach, we used different profiles of the face (see e.g. [51]) for classification via the Pearson coefficient.

## 3. The General Setup

Our project, funded by the European Regional Development Fund (ERDF), was concerned with the processing of facial biometric data in the context of the description of pedestrian movement. To obtain data for crowd movement models that account for the position of individuals [27] (in contrast to a crowd fluid [28]) it is necessary to identify those individuals with unintrusive biometric techniques. A

more specific application would be the analysis of commuter behaviour in public transportation: the usual systems available at the time of writing of this article only count passengers without recognizing individuals changing means of transportation.

To achieve this task of processing and classifying individual biometrical information we developed a 3D face recognition system that is able to cope with low-quality range data. This software platform was implemented as a toolbox for the MATLAB® scripting language.

Multiple modules for pre-processing and the actual face recognition were implemented and tested separately (cf. figure 1). For almost every module, we have developed and implemented alternative approaches to adjust the system to different application requirements. Depending on operating conditions and available capacities, the user can choose from a variety of individual modules and algorithms. The software contains conventional methods for 3D face recognition as well as unique and novel ideas.

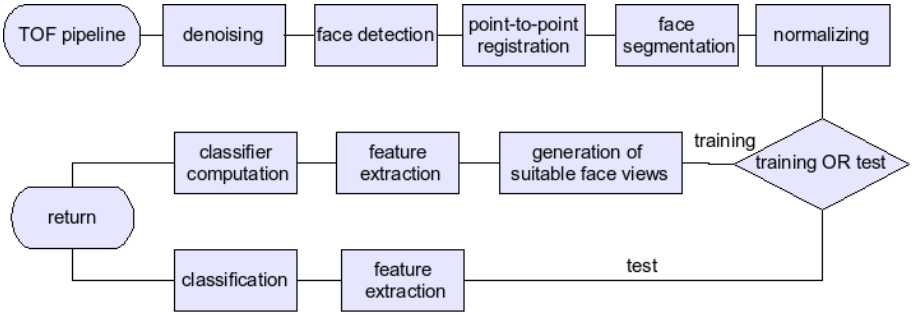



FIGURE 1. Software pipeline

As a main result of this project we implemented a robust real-time face recognition system from an innovative multi-modal approach that accounts for the typical characteristics of low-quality data obtained with a TOF sensor by combining 3D and 2D techniques that can deal with low-resolution images and little preliminary pre-processing capacities.

Since we would like to compare the performance of the system for data obtained from various sources, we implemented a simulation pipeline to emulate different noise chracteristics and resolution (see figure 2). The simulation pipeline features additional modules and algorithms for gradually degrading the pre-processed and comparatively noise-free laser scanner data from the Gavab database towards the data quality of a realistic cost-effective real-time ranging system. This simulation served as an important tool for assessing the sensor's requirements like resolution and signal-to-noise ratio.

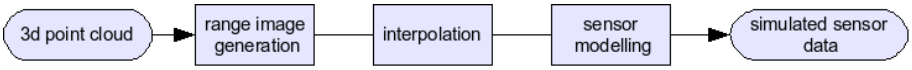



FIGURE 2. Simulation of low-quality sensor data

In figure 1, the flow chart of the final system is illustrated.

The pipelines can be divided into four main steps which will be described in the following: noise simulation, pre-processing, feature extraction and classification.

### Step 1: Simulation of Low-Quality Range Data

In order to assess the effectiveness of the implemented denoising algorithms for data acquired from different types of range sensors, we decided to simulate the noise characteristic of a standard continous-wave TOF sensor. Such a ranging system provides a number of advantages:

- It is less expensive than most other 3D imaging systems,
- suitable for real-time applications,
- not intrusive and suitable for covert operation (since it works with infrared light),
- space saving,
- also delivers 2D intensity data,
- delivers spatiotemporal data.

Since the TOF sensor independently delivers the geometry of an object as well as the intensity of the reflected infrared light, one could in principle also obtain information about the surface reflectivity of that object (at that particular wave length). A quite exhaustive survey about optical time-of-flight measurement is given in [38].

Unfortunately, as can be seen from figure 3, data acquired with standard TOF sensors available at the time of writing of this article is subject to significant systematic and random errors (noise). The random errors are mainly caused by

- shot noise,
- quantization noise,
- phase jitter and modulation errors;

whereas to name some reasons for systematic errors, among these are:

- Nonlinear characteristics (saturation effect),
- dark current,
- edge and movement artifacts due to mixed phases,
- reflecting surfaces,
- scattered light.

It is also for this reason that the simulation of noisy range data is necessary to account for future improvements in the image quality of such systems.

**Noise Characteristic of a TOF Range Sensor.** The precise technical method by which the optical signal received by a TOF sensor is measured and processed depends on the particular camera model, see e.g. [54, 60]. However, we suspect the noise characteristic to be very similar for the various systems since most of them use the phase-shifting technique, for which can be shown—assuming fixed environmental conditions and similar base characteristics for each pixel of the optical chip—that the standard deviation $\sigma$ of the range measurement is reciprocal to the amplitude $A$ of the optical signal [23]: $\sigma = \frac{1}{A}$.

In particular, the most popular 4-phase-shifting technique essentially leads to estimating the angular component of a random variable with values in $\mathbb{R}^2$; the mean length of this variable is given by the signal amplitude. This fact can be used to simulate the noise pattern of a TOF sensor by generating a pseudo-random point in the plane with normal distribution (by the Box–Muller method [8] for example)

and extracting the angular component. Further assuming a uniform reflectivity of the facial surface (i.e. skin), the signal amplitude can be estimated by taking the component of the surface normal projected onto the optical axis. Finally, this value has to be multiplied by the luminosity of the infrared LEDs at this point which can be estimated by considering the inverse-squared-distance law as well as a realistic directional characteristic for the emitter [52]. This approach leads to a simple but effective model for simulating the random noise characteristic of a generic TOF sensor.

It should be noted, however, that systematic errors are not accounted for. These errors are either difficult to control or can be eliminated by a careful sensor calibration [34, 41]; denoising algorithms are not of much use to eliminate this problem.

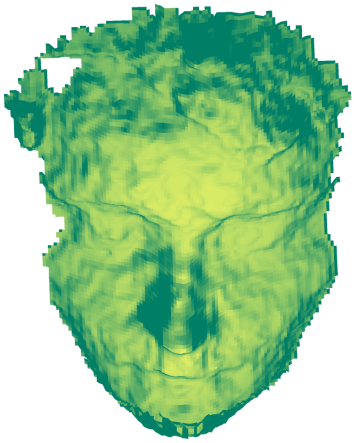


FIGURE 3. Facial surface obtained from a SwissRanger 3000 range image

### STEP 2: PRE-PROCESSING

To improve the face recognition on range images, a pre-processing step is needed which includes denoising, face detection, segmentation and registration.

**Denoising of Range Data.** The performance of the methods used to remove noise from range images depends on the sensor features and environmental conditions like the distance to the observed object, the reflectivity of the object's surface as well as the angle of incidence of the illumination. For the denoising step of our system, we implemented wavelet denoising and normalized convolution.

*The Discrete Wavelet Transform.* For a 2D signal $\mathbf{I}(x, y)$, the discrete wavelet transform (DWT) provides a multi-scale signal representation. The DWT is computed by a high-/low-pass filter bank, iteratively applied on the low-pass signal output of the previous stage. The multi-scale signal representation is then a collection of the resulting sub-band output coefficients. The inverse discrete WT is calculated by an iteratively applied synthesis filter bank.

The standard DWT is a powerful and non-redundant tool of signal processing, with four major drawbacks (as discussed in [61]):

- Oscillations in the neighborhood of singularities,

- lack of translational invariance,
- lack of directional invariance,
- missing phase information.

In order to remedy those drawbacks, a number of variants of the standard DWT has been developed. One example of such a generalization of the DWT is the complex wavelet transform (CWT).

*The Complex Wavelet Transform.* The CWT is computed in a way similar to the DWT, but with a complex-valued scaling function $\phi_c(t)$ and a complex-valued base wavelet $\psi_c(t)$:

$$\phi_c(t) = \phi_{re}(t) + i\phi_{im}(t) \tag{3.1a}$$

$$\psi_c(t) = \psi_{re}(t) + i\psi_{im}(t), \tag{3.1b}$$

where the indices $re$ and $im$ label the real and the imaginary part respectively. After projecting the signal onto the basis functions $2^{\frac{i}{2}}\psi(2^i t - n)$, one can calculate the wavelet coefficients

$$d_c(j,n) = d_{re}(j,n) + id_{im}(j,n) \tag{3.2}$$

with magnitude

$$|d_c(j,n)| = \sqrt{d_{re}^2(j,n) + d_{im}^2(j,n)} \tag{3.3}$$

and phase

$$\arg(d_c(j,n)) = \arctan\left(\frac{d_{im}(j,n)}{d_{re}(j,n)}\right). \tag{3.4}$$

A redundant form of the CWT is the complex dual-tree wavelet transform (CDTWT), which is discussed by Kingsbury and Selesnick in [35] and [61]. We will briefly describe the CDTWT in the next section.

*The Complex Dual Tree Wavelet Transform.* The complex dual-tree WT of a 2D signal is obtained by the parallel computing of four conventional critically-sampled separable 2D DWTs. The transform is therefore four times as expensive compared with a DWT. On the other hand, the complex dual-tree WT can provide—based on a certain design of the upper and lower high-pass and low-pass filters—a nearly ideal shift invariance and directional selectivity in two or more dimensions, as opposed to the critically-sampled discrete WT (see [35] and [61] for details.) The sub-band output of the upper discrete WT can be interpreted with this filter design as the real part, the sub-band output of the lower discrete WT as the imaginary part of a complex dual-tree wavelet transform.

As the actual wavelet denoising algorithm, we used the implementation of soft thresholding for 2D signal denoising as described in [12]. In this method, new values $w_{new}$ for wavelet coefficients $w$ for all scales and sub-bands are computed via the equations (3.5). In the first step, we delete the frequencies the coefficients of which lie below a certain threshold $T$. In the next step, the remaining coefficients are scaled. In this way, the effect of small values in the high-frequency sub-bands of the reconstructed 2D signal is decreased.

$$w_{new} = \max(\mathrm{abs}(w) - T, 0), \tag{3.5a}$$

$$w_{new} = \frac{w_{new}}{w_{new} + T} * w. \tag{3.5b}$$

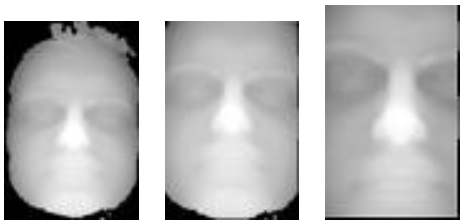
The new coefficients $w_{new}$ are used for an inverse wavelet transformation to reconstruct the 2D signal.

*Normalized Convolution.* Another approach for the adaptive denoising of an image $I$ with given confidence values $C$ is the normalized convolution [22, 56]. One advantage of this method is the fact that it can be easily generalized to data of arbitrary dimension since the denoised image $I'$ can be written invariantly in the form

$$I' = \frac{g * (C \cdot I)}{g * C} \ . \tag{3.6}$$

Here, $g$ is a suitable filter mask – a Gaussian kernel, for example. The multiplication is to be understood componentwise and $*$ denotes convolution. Using the inverse amplitude of the optical signal as a confidence measure, the normalized convolution can be used to filter TOF image sequences that are represented by 3D data.

**Segmentation and Registration.** To eliminate the non-facial parts of an image like the neck or shoulders, we invoked a multi-stage segmentation process, as shown in figure 4. First, the images were sliced into several regions by using a range threshold. Afterwards, the particular parts were analyzed morphologically and the face-like regions were chosen for further processing. In the next step, the nose tip was found for each of the detected faces. Around the nose tip, we set a sphere of fixed radius $r = 14\,\mathrm{cm}$ which proved to deliver the best recognition rates after some testing. We will call the partial surface cropped by the interior of this sphere the sphere-cropped image. Afterwards, a rectangle-shaped region from the corresponding 2.5D image was cut. This rectangle-cropped region was resized to $80 \times 50$ pixels in order to meet the classifier feature vector length. The individual steps of this multi-stage segmentation are illustrated in figure 4.



(a) (b) (c)

FIGURE 4. Multi-stage segmentation.
(a) Threshold-based segmentation, (b) segmentation by a ball of radius $r = 14\,\mathrm{cm}$ around the nose tip (sphere-cropped), (c) rectangle-cropped face region.

Before the segmentation process, an ICP algorithm was invoked to match the facial surface with a reference template.

### Step 3: Feature Extraction

In this step, the denoised, segmented and registered faces are passed to the feature extraction part of the pipeline. In particular, we extracted three features: surface normals, the local binary pattern and profiles.

**Surface Normals.** One of the earliest ideas in 3D face recognition is the use of curvature as a discriminant feature [25]. Unfortunately, at least the Gaussian curvature (which encodes valuable intrinsic geometric information) is highly susceptible to noise, as shown in figure 5 (visualized with JavaView© [53]). The Gaussian map (the distribution of surface normals) depends on the first derivatives of the parametrization and is thus more robust than the curvature which is encoded in the derivative of the Gaussian map [13]. Figure 6 shows the three cartesian components of the surface normals. (However, since the normal vectors are normalized, they actually represent 2-dimensional data on the unit sphere.) As one can easily see, the components correspond to a standard gray-valued 2D image that one would obtain from a human face if the reflectivity of the skin and the lighting were uniform. Thus, by using the Gaussian map as a feature one overcomes one of the most serious problems in 2D face recognition, namely varying illumination.

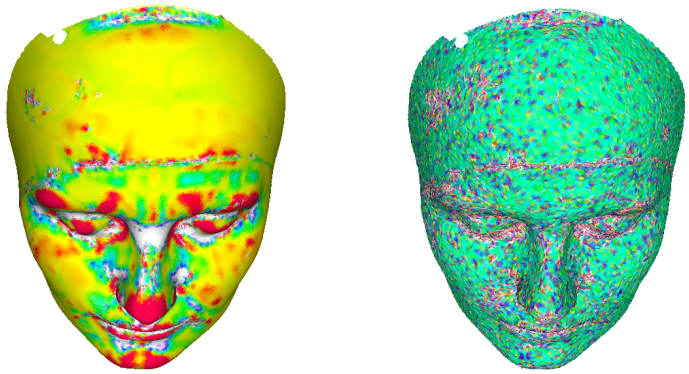
A major concern, however, is a suitable representation of the sphere-valued data. As a scalar function, the Gaussian curvature is independent of a specific parametrization of the surface, and being an intrinsic feature it is even invariant against isometric transformations of the facial surface. For the representation of the distribution of surface normals on the other hand, one has to agree on a specific coordinate system like spherical or stereographic coordinates. It is not a trivial problem to decide which coordinate system is optimal for a specific recognition task at hand. In this study, we have tested polar stereographic coordinates. This means that the surface normals $\mathbf{N} = (N_1, N_2, N_3)$ were projected onto the complex plane via the map

$$\pi(\mathbf{N}) = \frac{1}{1 - N_3} (N_1 + iN_2) \tag{3.7}$$

and represented by polar coordinates $(N_3, \arg \pi(\mathbf{N}))$.

**Local Binary Pattern (LBP).** Originally, the LBP approach was developed for the description and recognition of 2D textures [50]. The next step towards efficient face recognition using LBP was done by the authors of [31] with 3D Local Binary Patterns (3D LBP). They refined the feature value by adding three levels. However, they still used the histogram comparison, which is not the optimal comparison method. We applied classifiers and compared the results of the different conventional LBP and the different 3D LBP levels with the performance of other features.

To compute the LBP value, we first calculated the differences between the gray-scale values of a point and its $P$ neigbours. For $P$, the values 8, 16 or 24 can be chosen. The radius of the neighborhood defines an additional parameter $R$. In a second step, the signs of the intensity differences are binarily coded as 0 for a difference value of less than 0, and 1 for other difference values. Finally, the resulting binary digits are collected in a clockwise fashion to represent a binary number, which is then written as a decimal number. In this way, each point is mapped to an LBP value.



(a) High resolution image obtained with a structured-light 3D scanner and color-coded Gaussian curvature...

(b) ...after adding low power Gaussian noise

FIGURE 5. The computation of Gaussian curvature is susceptible to noise

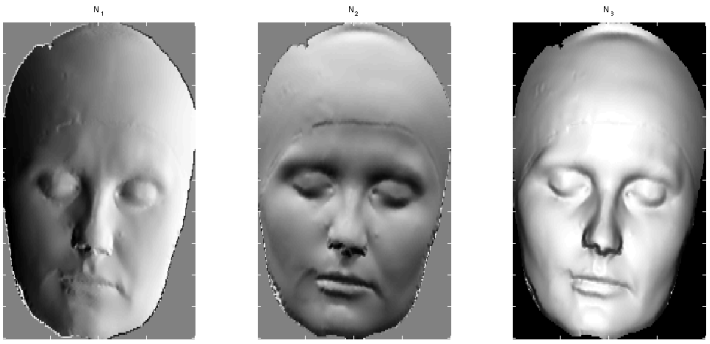


FIGURE 6. Cartesian components of surface normals

The authors of [50] also describe an alternative technique that is rotation-invariant. With this technique, each binary LBP number is shifted until the minimal value is reached. To increase the efficiency of the histogram comparison, one can encode all rare features with a single value. (Frequent features are edges, curvature lines or homogenous regions.) This variant of an LBP is called the uniform LBP. The texture operator for the general case based on a circularly symmetric neighborhood of $P$ members on a circle of radius $R$ will be denoted as $\mathrm{LBP}^{riu2}_{P,R}$ as in [50].

For the computation of 3D LBP values not only the signs of the gray-scale differences to the neighboring pixels are coded, but also the values themselves. Four bits are used to keep track of a difference value: One bit for the sign and three bits for the value. A difference less than 7 is coded directly, greater differences are coded as 7. If $P$ neighbors are considered, a $4 \times P$-sized table is used to list the four 3D LBP levels. The four $P$-bit binary numbers in the columns are decimally coded

and represent the four 3D LBP layers. A detailed description of this method can be found in [31]. The 3D texture operator will be denoted as $3DLBP_{layer\ i}$ with $i \in \{1, \ldots, 4\}$. Some examples of different LBP feature images are shown in figure 7.

**Reference Texture Features.** The authors of [29] evaluated the recognition performance of different gradient images and recommended the use of the Horizontal Sobel Operator (HSO) and the Large Horizontal Gradient operator (LHG) as effective texture descriptors. We ranked the performance of LBP and surface normals by using these descriptors and the original range images. The LHG describes the relative range differences along the horizontal direction in a range of five pixels and is computed via the filter mask $(\mathrm{LHG}_{ij}) = \begin{pmatrix} -1 & 0 & 0 & 0 & 1 \end{pmatrix}$. The HSO detects vertical edges and is calculated via the filter mask

$$(\mathrm{HSO}_{ij}) = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} . \tag{3.8}$$

Figure 7 shows example images of the examined texture descriptors.

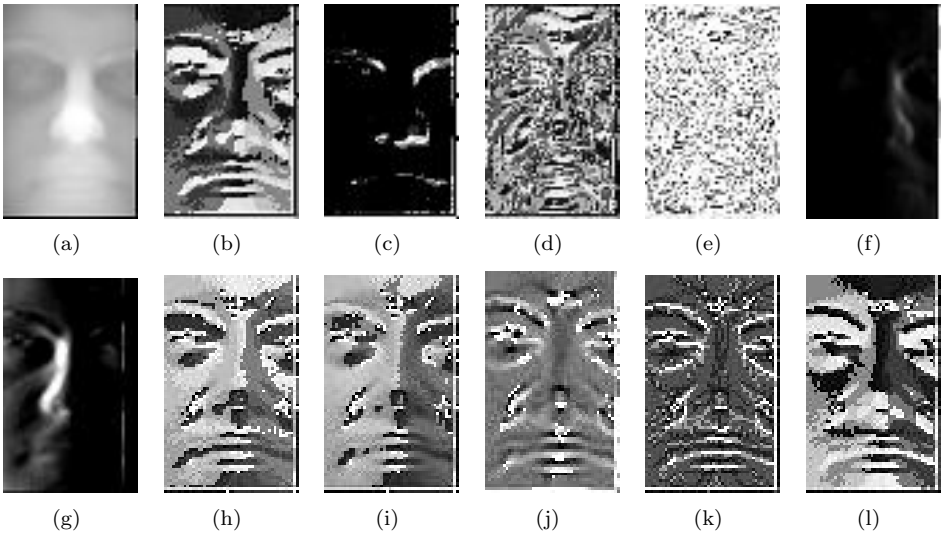



FIGURE 7. Feature images. a: range image, b: $3DLBP_{layer\ 1}$, c: $3DLBP_{layer\ 2}$, d: $3DLBP_{layer\ 3}$, e: $3DLBP_{layer\ 4}$, f: Horizontal Sobel Operator, g: Horizontal Gradient Large, h: $LBP^{u2}_{8,1}$, j: $LBP^{u2}_{16,1}$, i: $LBP^{riu2}_{16,2}$, k: $LBP^{ri}_{8,1}$, l: $LBP^{riu2}_{8,1}$

**Facial Profiles.** The extraction of vertical profiles from range data can be seen as a quite natural approach if you think of the fact that humans are quite capable of identifying a person by their facial silhouette only. In addition, templates of profile data need very little memory and the computation of their degree of correlation is very quick.

Here, four profiles (the vertical central profile, horizontal nose-crossing profile, horizontal root-of-the-nose-crossing profile and the horizontal forehead-crossing profile) have been investigated. The main difficulty consists of extracting the vertical central profile. However, two methods can be used to achieve this: (a) by finding the vertical symmetry plane of the face or (b) by aligning the face with a reference face to assure a vertical position and subsequent detecting the nose tip.

Figure 8 illustrates the vertical symmetry plane, and figure 9 shows the extracted profiles.

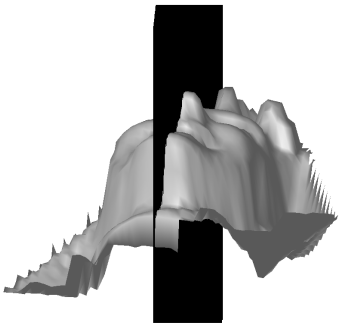


FIGURE 8. Symmetry plane of a face

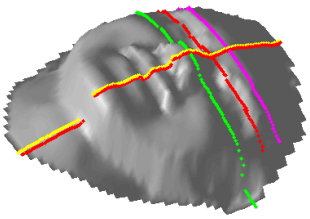


FIGURE 9. Vertical profiles: by the symmetry plane detection method (yellow), by detecting the nose tip (red) — horizontal profiles: nose-crossing profile (green), root-of-the-nose-crossing profile (red), forehead-crossing profile (pink)

### STEP 4: CLASSIFICATION

In a pre-processing step each 2D image of the $N$ training samples is converted into a 1D vector $\mathbf{x}_i$ with $m$ components ($m$ = image height $\times$ image width in pixels) by successively appending the rows of the image. In this way, we obtain a data set $(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) \in (\mathbb{R}^m)^N$.

**Principal Component Analysis.** One of the classifiers most commonly used as a reference is the PCA technique. PCA is a method for dimensional reduction that computes a linear projection operator $\mathbf{W}_{PCA}$ which maximizes the determinant of the total scatter matrix $\mathbf{S}_T$ of the $N$ image samples $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N$:

$$\mathbf{W}_{PCA} = \underset{\mathbf{W}}{\operatorname{argmax}} |\mathbf{W}^T \mathbf{S}_T \mathbf{W}| \tag{3.9}$$

with

$$\mathbf{S}_T = \sum_{k=1}^{N} (\mathbf{x}_k - \boldsymbol{\mu})(\mathbf{x}_k - \boldsymbol{\mu})^T \ , \tag{3.10}$$

where $\boldsymbol{\mu}$ is the mean of all samples. The optimal projection matrix $\mathbf{W}_{PCA}$ is comprised of the eigenvectors to the largest eigenvalues of the total scatter matrix $\mathbf{S}_T$. The images corresponding to these eigenvectors are called eigenfaces [4].

The new feature vectors $\mathbf{y}_k$ are given by

$$\mathbf{y}_k = \mathbf{W}_{PCA}^T \mathbf{x}_k. \tag{3.11}$$

**Linear Discriminant Analysis.** With the PCA approach, the considered total scatter matrix $\mathbf{S}_T$ includes not only the between-class scatter, which is useful for classification, but also the within-class scatter. In contrast to PCA, the LDA computes the projection matrix $\mathbf{W}_{LDA}$ in such a way that the ratio of the between-class scatter and the within-class scatter is maximized:

$$\mathbf{W}_{LDA} = \underset{\mathbf{W}}{\operatorname{argmax}} \frac{|\mathbf{W}^T \mathbf{S}_B \mathbf{W}|}{|\mathbf{W}^T \mathbf{S}_W \mathbf{W}|} \ , \tag{3.12}$$

where $\mathbf{S}_W$ and $\mathbf{S}_B$ describe the within-class scatter and the between-class scatter matrices, defined by

$$\mathbf{S}_W = \sum_{j=1}^{C} \sum_{k=1}^{N_j} (\mathbf{x}_{jk} - \boldsymbol{\mu}_j)(\mathbf{x}_{jk} - \boldsymbol{\mu}_j)^T \tag{3.13}$$

and

$$\mathbf{S}_B = \sum_{j=1}^{C} N_j (\boldsymbol{\mu}_j - \boldsymbol{\mu})(\boldsymbol{\mu}_j - \boldsymbol{\mu})^T \ . \tag{3.14}$$

Here $C$ is the class number, $\boldsymbol{\mu}_j$ is the the mean of the class $X_j$ and $N_j$ is the number of training samples in class $X_j$.

**Modified Linear Discriminant Analysis.** With conventional LDA, the Fisher optimization criterion is in a quotient form. As was shown in [4] and in [36], the quotient form can cause a numerical problem due to an insufficient number of training images for each person in the database. To avoid this problem, the fisherface approach was proposed in [4]. The dimensionality is reduced twofold: first without regarding the between-class differences using PCA for projecting the data set onto a subspace of dimension $m - C$, and second using an LDA projection. Another method was proposed by the authors of [36]; a modified Fisher optimization criterion in the deduction form:

$$\mathbf{W}_{MLDA} = \underset{\mathbf{W}}{\operatorname{argmax}} |\mathbf{W}^T (\mathbf{S}_B - \alpha \mathbf{S}_W) \mathbf{W}| \ . \tag{3.15}$$

Here $\alpha$ is the adjusting parameter for the weighting of the within-class scatter matrix $\mathbf{S}_W$ relative to the between-class scatter matrix $\mathbf{S}_B$.

**Pearson Coefficient.** This method is ideally suited for the comparison of the facial profiles mentioned above.

Correlation indicates the strength and direction of a linear relationship between two random variables (see e.g. [44] for details). The most commonly known correlation coefficient is the Pearson product-moment correlation coefficient $p(X, Y)$, which is obtained by dividing the covariance of the two variables by the product of their standard deviations. The correlation coefficient is thus defined as:

$$p(X,Y) = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{cov}(X,X)\cdot\operatorname{cov}(Y,Y)}} = \frac{\operatorname{cov}(X,Y)}{\sqrt{\operatorname{var}(X)}\cdot\sqrt{\operatorname{var}(Y)}}\ , \tag{3.16}$$

where $\operatorname{cov}(X, Y)$ is the covariance of $X$ and $Y$:

$$\operatorname{cov}(X,Y) = E((X - E(X))\cdot(Y - E(Y))). \tag{3.17}$$

The coefficient takes values between $-1$ and $+1$. A maximal positive linear relationship is given if the coefficient is $+1$, a maximal negative linear relationship is given by a coefficient of $-1$. A vanishing coefficient implies no linear relationship of the features.

## 4. Experiments and First Results

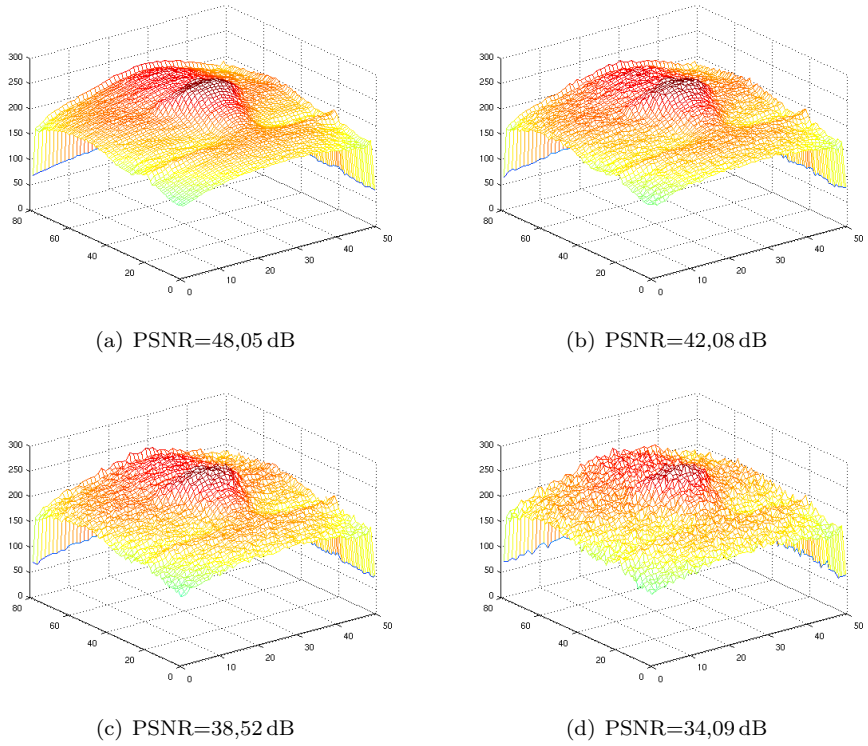
We tested our face recognition algorithms with (a) the original laser scanner data and (b) scanner data with added Gaussian noise for different values of the peak signal-to-noise ratio (PSNR, figure 10), (c) scanner data with added Gaussian noise that were subsequently denoised with the discrete WT method, (d) TOF data.

**Database Description.** The range images we used for our experiments were generated from 2.5D scans of human faces, contained in GavabDB, a database provided by the Gavab group. A short description of this database can be found in [48]. The GavabDB range images are in average of the size of $180 \times 120$ pixels. We used frontal view images with neutral and non-neutral facial expressions. We also created a database containing 2.5D TOF data sequences of 24 human faces acquired with a SwissRanger SR-3000 camera. The size of the TOF images was $176 \times 144$ pixels, the distance between the test subject and the TOF camera was approx. 40 cm. For the experiments we used $50 \times 80$ face segments, as shown in figure 11.

**Image Denoising.** According to our test data, both the discrete WT and CDTWT provided a robust performance for image denoising, but the CDTWT showed better numerical results. An example for CDTWT-based image denoising is presented in figure 12, with Gaussian noise added. Even a visual image comparison shows the high performance of CDTWT; this was observed repeatedly during the experiments. However, the choice of an appropriate noise dependent threshold value is crucial for this approach since an inaccurate threshold value can produce data corruption.

Since it is possible to estimate the noise of a TOF sensor via the amplitude of the optical signal [23], the denoising performance of the WT and the CDTWT can be warranted by choosing the optimal threshold value for each noise level as was done in our experiments.

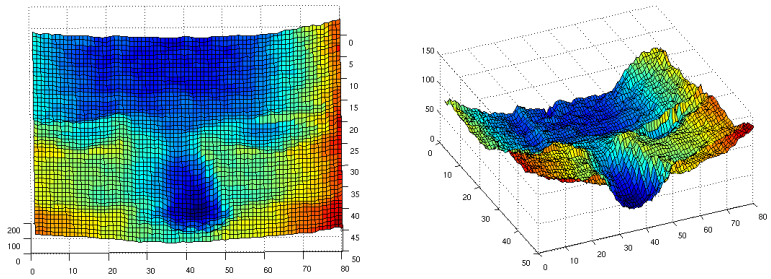
The TOF data sequences were simply filtered by averaging ten subsequent frames.



(a) PSNR=48,05 dB (b) PSNR=42,08 dB

(c) PSNR=38,52 dB (d) PSNR=34,09 dB

FIGURE 10. Face data after adding Gaussian noise for different PSNRs

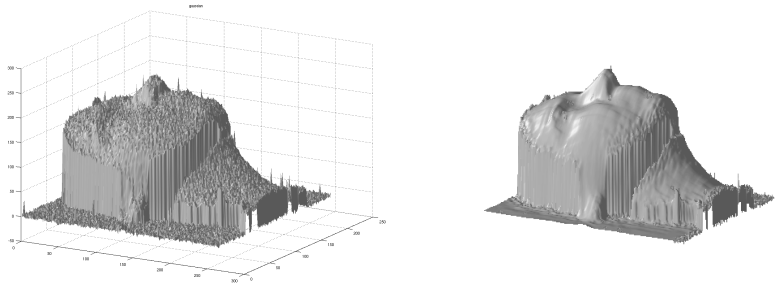


(a) TOF range image (b) TOF range image

FIGURE 11. TOF range image

**Point-to-Point Registration.** For face registration, we examined two different variants of the ICP algorithm: The first method aligns the face data with a reference face. In this way, 85% of the faces in the database could be rotated to a frontal and upright view.

The second method mirrors the 2.5D face data about the vertical axis and moves both the original face and the mirrored face against each other via ICP to achieve



(a) Original image, Gaussian noise added (b) ...after wavelet denoising

FIGURE 12. Complex dual tree wavelet transform denoising

an upright position. This technique assumes a mirror symmetry of the face data and works well on 82% of the faces whose angle of inclination does not exceed 30°.

To achieve an improved rotation invariance, we used additional images of different face views for the classifier training.

**Classifier Training.** The chosen feature vector length of $80 \times 50$ pixels corresponds to the resolution of a standard face image acquired by a TOF range camera (after segmentation). It is also necessary to restrict the resolution considering the limited computational resources of realistic applications.

To ensure rotation invariance (at least for small angles) and to address the problem of an insufficient number of training images for the LDA and the MLDA classifier, twelve additional images were generated for two unregistered training samples (one with a neutral and one with a smiling expression). To this end, the rectangle-cropped and sphere-cropped faces were rotated by $-5$, $-2.5$, $+2.5$, $+5$ degrees about all three axes.

The classifier training of the TOF data was processed on 20 randomly picked frames of the data sequence.

**Recognition on Original Laser Scanner Data.**

*Comparison of Different Image Features and Classifiers.* Table 1 shows the best recognition rates (recognition rates for two training samples) for PCA, LDA and MLDA classifiers. Depth values (range images) and $3\text{DLBP}_{layer\ 1}$ images were used as classifier features. For the MLDA classifier, the adjusting parameter $\alpha$ and for the PCA, the number of eigenvalues were investigated.

| | test image | | | test sequence | | |
|---|---|---|---|---|---|---|
| feature | PCA | LDA | MLDA | PCA | LDA | MLDA |
| depth (range image) | 47 | 47 | 47 | 47 | 47 | 50 |
| $3\text{DLBP}_{layer\ 1}$ | 87 | 85 | **88** | **92** | 85 | 88 |

TABLE 1. Recognition rate for neutral face expression in % (PCA, LDA and MLDA with 180 eigenvectors, MLDA with $\alpha = 6$)

Table 2 and table 3 show the recognition rates for different texture features and linear classifiers applied on images with a neutral face expression. The different LBP variants perform significantly better than the range and gradient images. The best MLDA results for a single-view test image were obtained with $\mathrm{LBP}^{riu2}_{8,1}$ (88%), for a test sequence with $\mathrm{3DLBP}_{layer\ 1}$ (92%). PCA and LDA classifiers showed a recognition rate of 92% for a single test image using the $\mathrm{LBP}^{riu2}_{8,1}$ texture feature.

| | test image | | test sequence | |
|---|---|---|---|---|
| feature | sphere-cropped | rectangle-cropped | sphere-cropped | rectangle-cropped |
| depth (range image) | 45 | 47 | 60 | 47 |
| $\mathrm{3DLBP}_{layer\ 1}$ | **78** | 87 | **92** | **88** |
| $\mathrm{3DLBP}_{layer\ 2}$ | 33 | 32 | 45 | 38 |
| $\mathrm{3DLBP}_{layer\ 3}$ | 58 | 77 | 72 | 90 |
| $\mathrm{3DLBP}_{layer\ 4}$ | 15 | 43 | 13 | 53 |
| Horizontal Sobel Operator | 32 | 40 | 43 | 50 |
| Large Horizontal Gradient | 42 | 50 | 52 | 55 |
| $\mathrm{LBP}^{u2}_{8,1}$ | 77 | 83 | 87 | 87 |
| $\mathrm{LBP}^{u2}_{16,1}$ | 70 | 82 | 83 | 87 |
| $\mathrm{LBP}^{riu2}_{16,2}$ | 63 | 77 | 80 | 80 |
| $\mathrm{LBP}^{ri}_{8,1}$ | 48 | 77 | 65 | 85 |
| $\mathrm{LBP}^{riu2}_{8,1}$ | 75 | **88** | 88 | **88** |

TABLE 2. Recognition rate for neutral face expression for different features in % for a test image and for a test sequence (MLDA with 60 eigenvectors and $\alpha = 8$)

*Recognition from Surface Normals.* Table 4 displays the recognition rates for the MLDA classifier using the polar stereographic representation of the Gaussian image. Obviously, the angular component is not robust against varying facial expressions. However, it should be noted that this data was not phase-unwrapped, which is not a trivial procedure for 2D data [24].

*Comparison of Different Profile Features.* With regard to the reference data for facial profile recognition we proceed on the assumption that during a real enrolment phase of the face recognition system only data of high quality is chosen as a reference pattern. For this reason, we singled out each person's best profiles. Afterwards, we conducted 427 comparisons per person and per profile type. The results are displayed in table 5. Particularly good results with profiles are achieved when the person is looking down. In this case, horizontal root-of-the-nose-crossing profiles and horizontal nose-crossing profiles obtained recognition rates of over 60%, whereas vertical central profiles allowed for a recognition rate of 58%. Generally, vertical central profiles and horizontal root-of-the-nose-crossing profiles show the best results. The horizontal forehead-crossing profile is the weakest feature and thus does not qualify for a discriminant recognition of faces. Furthermore, the profile recognition rates seem to be stable against facial expressions, compared with upward-oriented neutral faces.

| | test image | | test sequence | |
|---|---|---|---|---|
| feature | PCA | LDA | PCA | LDA |
| depth (range image) | 45 | 47 | 45 | 47 |
| $3DLBP_{layer\ 1}$ | 87 | 85 | **90** | 85 |
| $3DLBP_{layer\ 2}$ | 25 | 28 | 40 | 48 |
| $3DLBP_{layer\ 3}$ | 72 | 77 | 80 | 87 |
| $3DLBP_{layer\ 4}$ | 27 | 35 | 18 | 47 |
| Horizontal Sobel Operator | 38 | 45 | 48 | 55 |
| Large Horizontal Gradient | 43 | 50 | 50 | 55 |
| $LBP^{u2}_{8,1}$ | 88 | 83 | 88 | 88 |
| $LBP^{u2}_{16,1}$ | 85 | 85 | 87 | 88 |
| $LBP^{riu2}_{16,2}$ | 77 | 80 | 80 | 87 |
| $LBP^{ri}_{8,1}$ | 72 | 80 | 80 | 87 |
| $LBP^{riu2}_{8,1}$ | **92** | **92** | **90** | **92** |

TABLE 3. Recognition rate for neutral face expression for different features in % for a test image and for a test sequence for PCA and LDA with 180 eigenvectors and the rectangle-cropped face segmentation

| | $N_3$ | $\arg\pi(\mathbf{N})$ |
|---|---|---|
| Neutral | 87 | 90 |
| Smiling | 90 | 67 |

TABLE 4. Recognition rate for neutral and smiling face expression for surface normals in polar stereographic representation for MLDA (rectangle-cropped face segmentation)

| | Profile | | | |
|---|---|---|---|---|
| Face direction | Hor. nose-root | Hor. forehead | Hor. nose | Vert. central |
| Down | **60%** | 48% | **63%** | **58%** |
| Up | 37% | 26% | 37% | 32% |
| Frontal 1 | 41% | 13% | 21% | 51% |
| Frontal 2 | 37% | 13% | 21% | 49% |
| Frontal 1 and 2 | 39% | 13% | 21% | 50% |
| Facial expression | 38% | 0% | 26% | 53% |
| Laughter | 36% | 13% | 22% | 48% |
| Smile | 38% | 12% | 25% | 47% |
| on average | 41% | 20% | 31% | 48% |

TABLE 5. Profile recognition rates

*Comparison of Different Features.* Table 6 displays the best recognition results for faces with neutral and smiling expression for different features.

**Recognition on Noisy and Denoised Laser Scanner Data.** Table 7 illustrates our test results on simulated noisy data before and after applying the denoising

| | Neutral | Smiling |
|---|---|---|
| PCA from LBPs | 92 | 82 |
| MLDA for surface normals | 87 | 90 |
| Profiles | 51 | 48 |

TABLE 6. Best recognition rates of the introduced methods for faces with neutral and smiling expression

| | | image features | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | surface normals | | | | LBPs | | | | depth | | | |
| | classifiers | PCA | | MLDA | | PCA | | MLDA | | PCA | | MLDA | |
| PSNR | test seq. | a | b | a | b | a | b | a | b | a | b | a | b |
| 34,09 | noisy | 2 | 3 | 3 | 3 | 15 | 23 | 7 | 10 | 48 | 47 | 45 | **57** |
| | denoised | 48 | 35 | **58** | 33 | 52 | 47 | 45 | 43 | 38 | 40 | 47 | 53 |
| 38,52 | noisy | 3 | 7 | 5 | 5 | 25 | 48 | 12 | 27 | 48 | 47 | 45 | **57** |
| | denoised | 50 | 38 | **58** | 38 | 48 | 50 | 48 | 42 | 40 | 38 | 47 | **58** |
| 42,08 | noisy | 27 | 28 | 37 | 48 | 57 | **73** | 37 | 60 | 48 | 45 | 47 | 57 |
| | denoised | 53 | 38 | **58** | 38 | 48 | 52 | 50 | 45 | 38 | 38 | 47 | 53 |
| 48,05 | noisy | 75 | 83 | 78 | 85 | 72 | **87** | 72 | **87** | 48 | 45 | 45 | 55 |
| | denoised | 52 | 38 | **57** | 38 | 47 | 52 | 50 | 47 | 40 | 38 | 47 | 53 |

TABLE 7. Recognition rate for smiling (test seq. a) and neutral (test seq. b) face expressions for different PSNRs in % for PCA and MLDA with 180 eigenvectors and $\alpha = 6$

algorithms. During the classifier training we used high-quality images and trained only on scan sequences with little noise. During the recognition process we assumed the use of a low-quality sensor for mobile application. Our results show that some of the features like the surface normals and the LBPs are highly dependent upon noise and fail, whereas range images work properly and seem to be noise independent.

After the denoising process, the derived features achieve recognition rates similar to those obtained with the range images, independent of the noise level due to the image corruption by denoising. To sum up, for data with little noise the derived features are more discriminant than range images. For noisy images, we need to investigate other denoising algorithms and/or redesign the computation of the derived features.

**Recognition on TOF images.** Table 8 shows the best recognition rates (for one frame and for a sequence of 20 frames) obtained with the MLDA classifier on TOF images. Depth values (range images), $\mathrm{LBP}^{riu2}_{8,1}$ and surface normals were used as classifier features. The best recognition rates for single test images and test sequences are the same (83%).

## 5. Discussion and Future Work

Within the scope of this research project, we investigated, implemented and tested current methods for 3D face recognition. The developed system consists of multiple optional and mandatory modules for denoising and interpolation, classification training and the final face recognition. Since most parts of the system are

| feature | test image | test sequence |
|---|---|---|
| depth (range image) | 79 | 79 |
| $\mathrm{LBP}^{riu2}_{8,1}$ | 71 | 79 |
| surface normals | **83** | **83** |

TABLE 8. Recognition rate for TOF data and neutral face expression in % (MLDA with 180 eigenvectors and $\alpha = 6$, test sequence of 20 randomly picked frames)

interchangeable, it serves as an excellent test environment for the investigation of new methods.

For the classification and recognition tasks, a number of features was tested. The Gaussian map showed high recognition rates even though the representation (one stereographic coordinate) should be considered far from optimal and will probably be improved in the future.

The local binary pattern—which originally has been introduced as a means of describing textures by encoding the gray-value distribution in the neighborhood of a pixel—also gave very good results and proved the usefulness of this type of feature for images consisting of range data.

The facial profiles appear to be a robust feature against facial expressions like e.g. smiling. Although the recognition rates achieved with the current system are comparatively low, a further investigation to improve the results are in order since the processing of profiles would allow for a significant reduction in computing time and memory capacity that is often needed for realistic applications. For example, it would be interesting to check if a more suitable segmentation or different measures of profile correlation improve the recognition rates.

Although curvature data obtained from noisy range data by standard techniques is unreliable, a further investigation of more robust methods for extraction of this feature is of interest.

We achieved robust and real-time recognition performance (up to 92% for single features) from unregistered and low-quality data, which is very good compared with the currently existing methods. We expect these results to be improved by our future work which will include the combination of several features and the use of non-linear classifiers.

## References


[1] A. F. Abate, M. Nappi, S. Ricciardi, G. Sabatino, *Multi-modal Face Recognition by Means of Augmented Normal Map and PCA*, in: *IEEE Int. Conf. on Image Proc.* (2006).

[2] L. Akarun, B. Goekberk, A. Salah, *3D Face Recognition for Biometric Applications*, in: *13th European Signal Processing Conference (EUSIPCO)* (Antalya, Turkey, 2005), URL `http://www.cmpe.boun.edu.tr/~gokberk/eusipco2005.pdf`.

[3] H. Alt, L. J. Guibas, *Discrete Geometric Shapes: Matching, Interpolation, and Approximation: A Survey*, Tech. Rep. B 96-11 (1996), URL `citeseer.ist.psu.edu/alt96discrete.html`.

[4] P. N. Belhumeur, J. P. Hespanha, D. J. Kriegman, *Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection*, IEEE Trans. Pattern Analysis and Machine Intelligence, **Vol. 19, No. 7** (1997), 711–720, URL `http://cs.gmu.edu/~kosecka/cs803/pami97.pdf`.

[5] B. BenAmor, K. Ouji, M. Ardebilian, L. Chen, *3D Face Recognition by ICP-Based Shape Matching*, in: *The second International Conference on Machine Intelligence (ACIDCA-ICMI'2005)* (2005), URL `http://liris.cnrs.fr/publis/pdf/Chen-2005_liris1963.pdf?id=19%63`.

[6] P. J. Besl, N. D. McKay, *A Method for Registration of 3-D Shapes*, in: *IEEE Transactions on pattern analysis and machine intelligence, Vol. 14, No. 2* (1992), 239–256, URL `http://www.informatik.uni-bonn.de/~schulz/amr-prak/beslmckay.%pd`.

[7] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer Science+Business Media, LLC, 2001).

[8] G. E. P. Box, M. E. Muller, *A Note on the Generation of Random Normal Deviates*, Ann. Math. Stat., **29** (1958), 610–611.

[9] A. M. Bronstein, M. M. Bronstein, R. Kimmel, *Efficient Computation of Isometry-Invariant Distances Between Surfaces*, SIAM J. Sc. Comp., **28** (2006), 1812–1836, URL `http://www.cs.technion.ac.il/~ron/PAPERS/SIAM06.pdf`.

[10] A. M. Bronstein, M. M. Bronstein, R. Kimmel, *Generalized Multidimensional Scaling: A Framework for Isometry-Invariant Partial Surface Matching*, Proc. National Academy of Sciences (PNAS), **103** (2006), 1168–1172.

[11] A. M. Bronstein, M. M. Bronstein, R. Kimmel, *Expression-Invariant Representations of Faces*, IEEE Trans. Image Proc., **16** (2007), 188–197.

[12] S. Cai, K. Li, I. Selesnick, *Matlab Implementation of Wavelet Transforms: Introduction*, URL `http://taco.poly.edu/WaveletSoftware/index.html`.

[13] M. do Carmo, *Differential Geometry of Curves and Surfaces* (Prentice Hall, 1976).

[14] K. I. Chang, K. W. Bowyer, P. J. Flynn, *Multi-modal 2D and 3D Biometrics for Face Recognition*, in: *Proc. IEEE Int. Workshop on Analysis and Model. of Faces and Gestures* (2003).

[15] K. I. Chang, K. W. Bowyer, P. J. Flynn, *An Evaluation of Multimodal 2D+3D Biometrics*, IEEE Trans. Pattern Analysis and Machine Intelligence, **27** (2005), 619–624.

[16] K. I. Chang, K. W. Bowyer, P. J. Flynn, *Multiple Nose Region Matching for 3D Face Recognition under Varying Facial Expression*, IEEE Trans. Pattern Analysis and Machine Intelligence, **28** (2006), 1695–1700.

[17] Y. Chen, G. Medioni, *Object Modeling by Registration of Multiple Range Images*, in: *Proceedings of the 1991 IEEE International Conference on Robotics and Automation*, vol. 3 (1991), 2724–2729, URL `http://ieeexplore.ieee.org/iel2/347/3640/00132043.pdf?arnumbe%r=132043`.

[18] N. Chiba, H. Hanaizumi, *Three-Dimensional Face Recognition System*, in: *SICE Annual Conf. in Sapporo* (2004).

[19] T. Dey, J. Giesen, S. Goswami, *Shape Segmentation and Matching with Flow Discretization*, in: M. S. F. Dehne, J.-R. Sack (Ed.), *Proc. Workshop Algorithms and Data Strucutres (WADS 03), LNCS 2748*, vol. Volume 2748/2003 of *Lecture Notes in Computer Science* (Springer Berlin / Heidelberg, 2003), 25–36, URL `citeseer.ist.psu.edu/article/dey03shape.html`.

[20] C. Dorai, J. Weng, Jain, A. K., C. Mercer, *Registration and Integration of Multiple Object Views for 3D Model Construction*, IEEE Trans. Pattern Analysis and Machine Intelligence

(1998), 83–89, URL http://ieeexplore.ieee.org/iel3/34/14286/00655652.pdf?arnumbe%r=655652.

[21] R. O. Duda, P. E. Hart, D. G. Strok, *Pattern Classification* (John Wiley & Sons, Inc., 2001).

[22] M. Frank, M. Plaue, U. Köthe, F. A. Hamprecht, *Denoising of Continous-Wave Time-of-Flight Depth Images using Confidence Measures* (2007), submitted.

[23] M. Frank, M. Plaue, H. Rapp, U. Köthe, B. Jähne, F. A. Hamprecht, *Theoretical and Experimental Error Analysis of Continous-Wave Time-of-Flight Range Cameras* (2007), submitted.

[24] D. C. Ghiglia, M. D. Pritt, *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software* (Wiley, 1998).

[25] G. G. Gordon, *Face Recognition Based on Depth Maps and Surface Curvature*, in: *Proc. SPIE*, vol. 1570 (1991), 234–247, URL http://www.vincent-net.com/gaile/papers/SPIE_sandiego/spie_sa%ndiego.pdf.

[26] W. E. L. Grimson, *From Images to Surfaces* (MIT Press, 1981).

[27] D. Helbing, I. Farkas, T. Vicsek, *Simulating Dynamical Features of Escape Panic*, Nature, **407** (2000), 487–490.

[28] L. F. Henderson, *The Statistics of Crowd Fluids*, Nature, **229** (1971), 381–383.

[29] T. Heseltine, N. Pears, J. Austin, *Three-Dimensional Face Recognition Using Surface Space Combinations*, ICIP04, **II** (2004), 1421–1424, URL http://www-users.cs.york.ac.uk/~nep/tomh/3DFaceRecUsingSurfac%eSpaceCombinations-BMVC.pdf.

[30] K. Hildebrandt, K. Polthier, *Anisotropic Filtering of Non-linear Surface Features*, Comp. Graph. Forum, **23** (2004), 391–400.

[31] Y. Huang, Y. Wang, T. N. Tan, *Combining Statistics of Geometrical and Correlative Features for 3D Face Recognition*, ICPR04, **3** (2006), I: 330–333, URL http://www.visionbib.com/bibliography/people890.html\#TT65507%.

[32] M. Hüsken, M. Brauckmann, S. Gehlen, C. von der Malsburg, *Strategies and Benefits of Fusion of 2D and 3D Face Recognition*, in: *Proc. IEEE CVPR* (2005).

[33] A. K. Jain, R. P. W. Duin, J. Mao, *Statistical Pattern Recognition: A Review*, in: *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22 (2000).

[34] T. Kahlmann, F. Remondino, H. Ingensand, *Calibration for Increased Accuracy of the Range Imaging Camera Swissranger™*, in: *Proc. ISPRS* (2006), URL http://www.photogrammetry.ethz.ch/general/persons/fabio/kahlm%ann_etal_ISPRSV2006.pdf.

[35] N. G. Kingsbury, *Complex Wavelets for Shift Invariant Analysis and Filtering of Signals*, Applied and Computational Harmonic Analysis, **10** (2002), 234–253, URL http://www-sigproc.eng.cam.ac.uk/~ngk/publications/ngk_ACHApa%p.pdf.

[36] H. Kong, X. Li, J.-G. Wang, E. K. Teoh, C. Kambhamettu, *Discriminant Low-dimensional Subspace Analysis for Face Recognition with Small Number of Training Samples*, IEEE Transactions on system, man and cybernetics, Part B, URL http://www.bmva.ac.uk/bmvc/2005/papers/165/HK_BMVC05.pdf.

[37] C. Lange, K. Polthier, *Anisotropic Smoothing of Point Sets*, Comp. Aid. Geom. Des., **22** (2005), 680–692, URL http://page.mi.fu-berlin.de/polthier/articles/pointSet/points%etFairing.pdf.

[38] R. Lange, *3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD Technology*, Ph.D. thesis, University of Siegen (2000), URL http://deposit.ddb.de/cgi-bin/dokserv?idn=960293825.

[39] T. Ledermann, J. Pannekamp, *3-D-Wahrnehmung in der Robotik*, interaktiv 2.2006, **2** (2006), 24–25.

[40] C. Li, A. Barreto, *An Integrated 3D Face Expression Recognition Approach*, in: *Proc. ICASSP* (2006).

[41] M. Lindner, A. Kolb, *Lateral and Depth Calibration of PMD distance sensors*, in: *Proc. ISCV* (2006), 524–533, URL http://www.cg.informatik.uni-siegen.de/data/Publications/2006%/isvc2006.pdf.

[42] X. Lu, *3D Face Recognition Across Pose and Expression*, Ph.D. thesis, Michigan State University (2006), URL http://www.msu.edu/~lvxiaogu/thesis/thesis_3DFace_Lu.htm.

[43] B. Mederos, L. Velho, L. H. de Figueiredo, *Robust Smoothing of Noisy Point Clouds*, in: N. Press (Ed.), *Geom. Des. and Comp.: Seattle 2003* (2004), URL http://w3.impa.br/~boris/seattle2003.pdf.

[44] B. Meffert, O. Hochmuth, *Werkzeuge der Signalverarbeitung* (Pearson Studium, 2004).

[45] A. S. Mian, D. Mathers, M. Bennamoun, R. Owens, G. Hingston, *3D Face Recognition by Matching Shape Descriptors*, in: *Proc. IVCNZ '04* (2004), 23–28, URL `http://www.postgraduate.uwa.edu.au/__data/page/110739/3Dfacer%ecogmatchingshape.pdf`.

[46] K. U. Modrich, *Industrielle Bildverarbeitung für automatisierte Produktionen*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim (2007), 25–31.

[47] B. Moghaddam, T. Jebara, A. Pentland, *Bayesian Face Recognition*, in: *Pattern Recognition, Vol. 33, No. 11, pps. 1171-1782* (2000), URL `http://courses.csail.mit.edu/6.869/handouts/MerlTR2000-42%20B%ayesianFacReco.pdf`.

[48] A. B. Moreno, A. Sanchez, *GavabDB: a 3D Face Database*, in: *Proc. 2nd COST275 Workshop on Biometrics on the Internet* (Vigo (Spain), 2004), 75–80, URL `http://gavab.escet.urjc.es/recursosen.html`.

[49] T. Oggier, B. Büttgen, F. Lustenberger, G. Becker, B. Rüegg, A. Hodac, *SwissRanger SR3000 and First Experiences Based on Miniaturized 3D-TOF Cameras*, Tech. rep., CSEM, IEE, Fachhochschule Rapperswil Switzerland (2005).

[50] T. Ojala, M. Pietikaeinen, *Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns*, IEEE Trans. Pattern Analysis and Machine Intelligence, **24** (2005), 971–987, URL `http://www.mediateam.oulu.fi/publications/pdf/6.pdf`.

[51] G. Pan, Y. Wu, Z. Wu, *Investigating Profile Extracted from Range Data for 3D Face Recognition*, in: *Proc. IEEE Int. Conf. on Systems, Man and Cybernetics* (2003).

[52] M. Plaue, *Analysis of the PMD Imaging System*, Tech. rep., IWR, Univ. of Heidelberg (2006).

[53] K. Polthier, *JavaView - Interactive 3D Geometry and Visualization* (1999-2006), URL `http://www.javaview.de/`.

[54] H. Rapp, *Experimental and Theoretical Investigation of Correlating TOF Camera Systems*, Master's thesis, IWR, University of Heidelberg (2007).

[55] H. Rapp, M. Frank, F. A. Hamprecht, B. Jähne, *A Theoretical and Experimental Investigation of the Systematic Errors and Statistical Uncertainties of Time-of-Flight Cameras*, Int. J. Accounting, Auditing and Performance Evaluation (accepted).

[56] J. Restle, M. Hissmann, F. Hamprecht, *Nonparametric Smoothing of Interferometric Height Maps Using Confidence Values*, Opt. Eng., **43** (2004), 866–871.

[57] T. Ringbeck, B. Hagebeuker, *A 3D Time Of Flight Camera For Object Detection*.

[58] T. Ringbeck, T. Möller, B. Hagebeuker, *Multidimensional Measurement by Using 3-D PMD Sensors*.

[59] S. Rusinkiewicz, M. Levoy, *Efficient Variants of the ICP Algorithm*, in: *Proceedings of the Third Intl. Conf. on 3D Digital Imaging and Modeling* (2001), 145–152.

[60] R. Schwarte, Z. Xu, H. Heinol, J. Olk, B. Buxbaum, *New Optical Four-Quadrant Phase-Detector Integrated into a Photogate Array for Small and Precise 3D Cameras*, in: *Proc. SPIE*, vol. 3023 (1997), 119–128.

[61] I. W. Selesnick, R. G. Baraniuk, N. G. Kingsbury, *The Dual-Tree Complex Wavelet Transform*, IEEE Signal Processing Magazine, **123** (2005), 123–151, URL `http://ieeexplore.ieee.org/iel5/79/33042/01550194.pdf`.

[62] P. Shukla, *Complex Wavelet Transforms and their Applications*, Ph.D. thesis, University of Strathclyde, Glasgow, UK (2003), URL `http://www.commsp.ee.ic.ac.uk/~pancham/MPHIL_THESIS.pdf`.

[63] H. J. Song, K. H. Sohn, *Face Recognition Using Two Different 3D Sensors*, in: *Proc. ISPACS* (2004).

[64] T. Tasdizen, R. Whitaker, P. Burchard, S. Osher, *Geometric Surface Smoothing via Anisotropic Diffusion of Normals*, Proceedings of the Conference on Visualization, **P4** (2002), 125 – 132, URL `http://portal.acm.org/citation.cfm?id=602117`.

[65] R. Veltkamp, M. Hagedoorn, *State of the Art in Shape Matching*, URL `citeseer.ist.psu.edu/veltkamp99stateart.html`.

[66] D. M. Weinstein, *The Analytic 3-D Transform for the Least-Squared Fit of Three Pairs of Corresponding Point*, Tech. rep., Department of Computer Science, University of Utah (1998), URL `http://www.sci.utah.edu/publications/dmw98/UUCS-98-005.pdf`.

[67] J. Wikander, *Automated Vehicle Occupancy Technologies Study: Synthesis Report*, Tech. rep., Texas Transportation Institute (2007).