Using line features for 3D face registration

Fabian Brix

18.06.2013

Abstract

In this bachelor thesis we attempt to modify the existing face registration pipeline for the morphable face model of Prof. Thomas Vetter by using a registration algorithm developed by PD Marcel Lthi at the University of Basel. ALTERNATIVE: In this bachelor thesis we discuss the construction of a face registration pipeline using an algorithm based on a vector-valued gaussian process and at the same time attempting to ensure registration quality through the use of contours marking important parts of the face - referred to as line features.

The algorithm is capable of mapping any two shapes on to one another. All that is needed is a set of corresponding points on the two shapes. Different constraints to the displacement field can be applied through regularisation.

The aim of this bachelor thesis is more specifically to apply this general algorithm for point correspondences to scanned face data, that is to implement feasible registration of face scans onto the mean face of the morphable model. In order to achieve this we mark important parts of the face meshs not only with point landmarks, but also structures and organs (eyebrows, eyes, ears) with lines - line features - and thereby to create further correspondences for the algorithm to perform better by. Instead of using sparse points of key features points of the face we mark complex features, e.g. the eyes, with contour lines - line features in order to create further correspondences

These line features are marked by hand using bzier curves on three 2D images to the front, left and right of the 3D face. In order to utilize them, however, they have to be projected on to the computed mesh of the face that was recorded by a 3D scanner. These meshs have holes in the region of the eyes and the ears rendering the projected line features useless at first. This thesis first gives an overview over the morphable model and the face registration pipeline, then goes on to obtaining 3D points from the 2D line features, to explain the theory behind the general algorithm and in the main part discusses the problems and solutions we encountered trying to optimize the algorithm for and without line features for the face registration process.

Chapter 1

Introduction

1.1 Problem Statement

So es bizzeli alles schriebe 1. Use Gaussian Processes - 2. Use Line Features =¿ prepare for Gaussian Process Regression In this bachelor thesis Implement 3D face registration using Gaussian Processes and Line Features. One part of the problem is to sample equidistant 3D points from 2D line features marked on images of a 3D face scan. These line features should then be used as an additional input to a registration algorithm which is based on Gaussian Process Regression. The aim is to build a pipeline which starts off with the raw scan data as well as the landmarks and line features. The feature points are used to register the mean face of the MM/BFM (Basel Face Model) on to/with the raw scan thereby obtaining a fully defined and textured 3D model representation of the face in 3D. Registration is the technique of aligning to objects using a transformation, in this case the registration is performed by adding displacements to every points in the mean face model. A model is represented as vector N*d. What is a model? A vector representation of a 3D scan? For the morphing a Posterior Shape Model is used in combination with a Gaussian Process. Image registration is a process of aligning two images into a common coordinate system thus aligning.

(gaussian process + line features for accurate, reproducable registration)

1.2 Review Literature

2. Definition of terms (morphable model, 3D face registration, Gaussian Process regression, posterior shape models) 3. Review of literature (papers)

Chapter 2

Model building

2.1 Building a model from registered face data

Model building is our motivation. We are going to great lengths to build a model In this chapter we explain how to go about building a 3D face model from 3D face scans. 2.3 Building a Model from the registered data (short) cite morphable model, give a short overview how a model is built A set of faces parametrized by coefficients a, set of Textures parametrized by coefficients b Fit multivariate normal distribution to data set, based on average of faces and textures. Build covariance matrices over differences between the mean and face samples in surface and texture. =¿ two distributions. Perform PCA to get orthogonal basis system In the MM three subspaces are morphed independently

Eine Gruppe (G,) heisst abelsch [abelian] oder kommutativ wenn ab = ba gilt fr alle a,b G.

2.2 Prerequisite Data

image with landmarks and line features a short overview what data we have given

Facial Scans: face scans given as point clouds The data we have given is a set of about 300 face scans that have had a set of key points marked. Furthermore important and detailed regions like the eyes, ears and lips have been marked by contour lines known as line features. The scans have been obtained with a scanner. The surface is very detailed, however the eyes and the nostrils are not recorded. From these scans we want to create fully textured 3D faces, which can be used to build a new face model.

Mean Face: The mean face has been derived from a collection of 100 male and 100 female 3D face models.

2.3 Finding Correspondences

WE WANT POINT TO POINT CORRESPONDENCE BETWEEN THE TWO FACES in general: point to point correspondence between to images Are scans already in semantical correspondence? No semantical correspondence FINDING CORRESPONDENCE IS EXACTLY THE AIM OF REGISTRATION =¿ HAVING SAME POINTS AS CLOSE TO ONE ANOTHER AS POSSIBLE Now in order to obtain a 3D representations of the face we need to transform the mean face so that it fits a particular 3D face scan. To find the transformation, however, we first have to find feature points in both 3D representations which correspond to the same semantical structure. Previous work has shown that point landmarks are not sufficient to preserve the level of detail which is imminent in the regions of the eyes, ears and lips and that the computed transformations are not able to preserve these regions. For this reason, additional line features have been introduced. In order to relate these

How registration works so far

What we want to change

Chapter 3

Gaussian Processes in 3D Face Registration

As described briefly in the introduction, the first of our two objectives is to build a face registration pipeline. In achieving this we use an algorithm which can handle arbitrary shapes for the registration of the 3D faces. It is derived from a stochastic process, more specifically a vector-valued Gaussian process or Gaussian random field. In this chapter we deal with the theory necessary for understanding the functionality of the registration pipeline. To begin with, we recapitulate the definition of stochastic processes and extend it to the definition of Gaussian processes. In the next step, we then delve into Gaussian process regression and finish by applying a vector-valued Gaussian process to our problem of 3D face mesh registration.

3.1 Stochastic Processes

In probability theory a stochastic process consists of a collection of random variables $\{X(t)\}_{t\in\Omega}$ where Ω is an index set. It is used to model the change of a random value over time. The underlying parameter time is either real or integer valued. In our case, however, we use a collection of vector-valued random variables with indices in the $\mathbb{R}^3$ space because we want to model deformations of the vectors on the faces' surfaces. This generalization of a stochastic process - which can handle multidimensional vectors - is called a random field. Defining the random variables on an index set in an n-dimensional space, allows for spatial correlation of the resulting values, which is an important aspect of the algorithm discussed later on.

3.2 Gaussian Processes

A Gaussian process is a stochastic process in which each finite collection $\Omega_0 \subset \Omega$ of random variables has a joint normal distribution. More formally, we define the collection of random variables $\{X(t)\}_{t\in\Omega}$ to have a d-dimensional normal distribution if the collection $\{X(t)\}_{t\in\Omega_0}$ - for any finite subset Ω_0 - has a joint $d \times |\Omega_0|$-dimensional normal distribution with mean $\mu(\Omega_0)$ and covariance $\Sigma(\Omega_0)$. If $\Omega \nsubseteq \mathbb{R}$ holds, the process is a Gaussian random field, which holds true for our case, because we use an index set $\Omega \subseteq \mathbb{R}^3$. In the further proceedings the term "Vector-valued Gaussian Processes" will be used to refer to Gaussian random field" will be used to refer to Gaussian random fields.

An alternative way of viewing a Gaussian process is to consider it as a distribution over functions. This allows us to look for inference in the space of these functions given a dataset, specifically to find the deformation function given a 3D face mesh. Each random variable now yields the value of a function $f(x)$ at a location $x \in \mathcal{X}$ in the index set of possible inputs. We now denote the index set by $\mathcal{X}$ to stress that we are ceasing to discuss Gaussian processes defined over time. In this function-space view a Gaussian Process at location x is thus $f(x) \sim GP(\mu(x), k(x,x'))$ defined by its mean $\mu : \mathcal{X} \to \mathbb{R}$ and covariance $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ functions which in turn are defined over the set of input vectors. With $\mu(\mathcal{X}) = (\mu(x))_{x\in\mathcal{X}}$ and $\Sigma(\mathcal{X}) = (k(x,x'))_{x,x'\in\mathcal{X}}$ we obtain the full distribution of the process $GP(\mu(\mathcal{X}), \Sigma(\mathcal{X}))$. For the purpose of simplifying calculations we may assume that every random variable has zero mean without a loss of generality. When modeling a deformation field with a Gaussian process this circumstance implies that the expected deformation is itself zero.

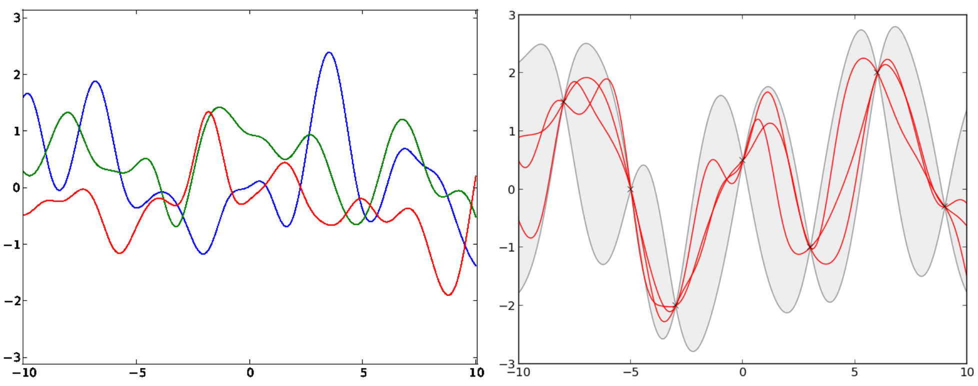
Covariance Functions The key feature of a Gaussian Process is its covariance function also known as "kernel". It specifies the covariance $\mathbb{E}[f(x)f(x')]$ between pairs of random variables for two input vectors x and x', allowing us to make assumptions about the input space by defining the spatial co-dependency of the modelled random variables. Note that when assuming zero mean we can completely define the process' behaviour with the covariance function.
A simple example of a covariance function is the squared exponential covariance function, defined by $cov(f(x), f(x')) = k(x,x') = \exp(-\frac{(x-x')}{2l^2})$. (derivation Rasmussen et al. p.83) *still to be continued and refined...*

It is possible to obtain different prior models by using different covariance functions. In our case, we use a stationary (x-x, invariant to translation), isotropic exponential covariance function - Squared Exponential Covariance Function (p. 38)

Gaussian Process Prior The specification of the covariance function implies that a GP is a distribution over functions. To illustrate this one can draw samples from a prior distribution of functions evaluated at any number of points, X_*. The Gaussian Process Prior is solely defined by the covariance matrix made up of the covariances of the input points.

$$\Sigma(X_*) = \begin{bmatrix} k(x_{*1}, x_{*2}) & \cdots & cov(f(x_{*1}, f(x_{*n}) \\ \vdots & \ddots & \vdots \\ cov(f(x_{*n}, f(x_{*1}) & \cdots & cov(f(x_{*n}, f(x_{*n}) \end{bmatrix} \in \mathcal{M}^{|X_*|\times|X_*|} \tag{3.1}$$

A sample is a random Gaussian vector $f_* \sim \mathcal{N}(0, \Sigma(X_*))$ containing a function value for every given input point. Plotting random samples above their input points is a nice way of illustrating that a GP is indeed a distribution over functions, see figure 3.1. The GP Prior forms the basis for inference in Gaussian Process Regression.



(a) Samples drawn from the prior (b) Samples after a few observations

Figure 3.1: In figure a) three functions have been drawn from the GP prior, each has a sample size of 1000 points. Figure b) again shows three functions, but this time the prediction incorporates the information of random values observed at seven points in the input space

Vector-valued Gaussian Processes In order to use Gaussian processes to model deformation fields of three dimensional vectors as intended, there is the need for a generalization of the above definition from the function-space view. The random variables $X_1, X_2, \ldots, X_k, \ldots, X_n$ are now d-dimensional vectors, yielding a covariance function of the form $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}^{d\times d}$ and $k(x, x') = \mathbb{E}[X_k(x)^T X_k(x')]$. *Should this paragraph be continued?*

3.3 Gaussian Process Regression

The task of registering two 3D face meshs can be treated as a regression problem in which the goal is to predict the deformation of all floating mesh points, given the displacement of the landmarks present in both meshs. Trying to fit an expected function - be it linear, quadratic, cubic or nonpolynomial - to the data is not a sufficiently elaborated approach to our problem. Using a Gaussian Process disposes of the need to describe the data by a specific function type, because the response for every input point is now represented by a normally distributed random value, in turn governed by the specification of the covariance function.

Key assumption: data can be represented as a sample from a multivariate gaussian distribution P

Regression Problem Assume a training set $\mathcal{D} = \{(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)\}$ where $x \in \mathbb{R}^d$ and y is a scalar output or target. Later on, in the case of the training set consisting of landmarks, a Vector-valued Gaussian Process must be used, because y is then also a vector $y \in \mathbb{R}^d$. The task is now to infer the conditional distribution of the targets for yet unseen inputs and given the training data $p(\mathbf{f}_*|\mathbf{x}_*, \mathcal{D})$

Noise-free Prediction First we assume the observations from the training data to be noise-free so that we can fix the training data to these observations $\mathbf{y}$ without complicating the model. The joint prior distribution with training $\mathbf{f}$ and test $\mathbf{f}_*$ outputs indicated is the following:

$$\begin{bmatrix}\mathbf{f}\\ \mathbf{f}_*\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}\Sigma(X) & \Sigma(X, X_*)\\ \Sigma(X_*, X) & \Sigma(X_*)\end{bmatrix}\right) \tag{3.2}$$

We can now obtain the posterior samples illustrated in 3.1 b) by conditioning the above joint Gaussian prior distribution on the observations $\mathbf{f}_*|\mathbf{f} = \mathbf{y}$ which results in the following distribution:

$$\mathbf{f}_*|X_*, (X, \mathbf{f}) \sim \mathcal{N}\left(\Sigma(X_*, X)\Sigma(X)^{-1}\mathbf{f}, \Sigma(X_*) - \Sigma(X_*, X)\Sigma(X)^{-1}\Sigma(X, X_*)\right) \tag{3.3}$$

Later on, we will extend this definition to 3-dimensional inputs and outputs.

Prediction with Gaussian Noise Model In most real world applications as is the case for the problem we will look into later, however, observations from the training data are not free of noise. The landmarks clicked on the 3D face meshs, for example, can never be marked at the exact same feature location. These circumstances call for the incorporation of a noise model. We specify a simple additive i.i.d Gaussian noise model $y = f(x) + \varepsilon$ where $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ for every input vector x. In section **??** the variances will be varied for every sole landmark. For now it is enough to add the variance of the noise model to the covariance of the training.

$$\begin{bmatrix}\mathbf{y}\\ \mathbf{y}_*\end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix}\Sigma(X) + \sigma^2\mathcal{I}_{|X|} & \Sigma(X, X_*)\\ \Sigma(X_*, X) & \Sigma(X_*)\end{bmatrix}\right) \tag{3.4}$$

The distribution - now conditioned on the noisy observations - is thus

$$\mathbf{y}_*|\mathbf{f} = \mathbf{y} \sim \mathcal{N}\left(\overline{\mathbf{y}}_*, \Sigma(\mathbf{y}_*)\right) \tag{3.5a}$$

where the mean depends on the observed training targets

$$\bar{\mathbf{y}}_* = \Sigma(X_*, X)\left(\Sigma(X) + \sigma^2\mathcal{I}_{|X|}\right)^{-1}\mathbf{y} \tag{3.5b}$$

whilst the covariance depends only on the input points

$$\Sigma_* = \Sigma(X_*) - \Sigma(X_*, X)\left(\Sigma(X) + \sigma^2\mathcal{I}_{|X|}\right)^{-1}\Sigma(X, X_*) \tag{3.5c}$$

Conclusion, how does this help us to proceed?

3.4 Application to 3D Face Meshs

deformation field bold instead of calligraphic? We strive to predict a deformation field $\mathcal{D} : \mathcal{M} \subset \mathbb{R}^3 \to \mathbb{R}^3$ which assigns a displacement vector to every vertex in a triangulated mesh. The mesh in question is called a floating or moving mesh $\mathcal{M}$. Adding the displacement field to the moving mesh should then render a mesh corresponding as closely as possible to a target mesh $\mathcal{T}$ and thereby performing a registration / should provide an accurate mapping on to the target mesh. The mean mesh of the Morphable Model serves as the moving mesh which we want to register with multiple triangulated meshs of scanned target faces.

Reference Mesh Prior ***Prior consists of smooth deformations of the mean face*** As defined by the deformation field the output the regression problem is in $\mathbb{R}^3$ calling for the use of a Vector-valued Gaussian Process with random variables $u \subseteq \mathbb{R}^3$. After the reference and target have been aligned **??** a Vector-valued Gaussian Process can be initialized by defining the prior over all vertices of mean mesh. For this purpose the covariance function has to be redefined to handle 3-dimensional vectors.

$$k\left(\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}, \begin{bmatrix}x'_2\\x'_2\\x'_3\end{bmatrix}\right) = xy^T \in M^{3\times3} \tag{3.6}$$

Each covariance entails 9 relationships between the different components of the vectors, yielding a 3×3 matrix. The covariance matrix then grows to become:

$$\Sigma_{\mathcal{X}} = \begin{bmatrix} k(\mathbf{x}_1, \mathbf{x}_1) & \cdots & k(\mathbf{x}_1, \mathbf{x}_n) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_n, \mathbf{x}_1) & \cdots & k(\mathbf{x}_n, \mathbf{x}_n) \end{bmatrix} \in M^{3n\times3n} \tag{3.7}$$

The mean mesh is defined by a set of vectors $\mathcal{X} \in \mathbb{R}^3$ and a set of landmarks $L_{\mathcal{M}} = \{l_1, \cdots, l_n\} \subset \mathbb{R}^3$. The mean vector μ is made up of the component-wise listing of vectors so that is has dimensionality $3n$. Setting the whole mean vector to zero, as discussed before, implies a mean deformation of zero and makes perfect sense in this setting, because we are modelling deformations of the mean face surface. The prior distribution over the mean face mesh is therefore defined as

$$\mathcal{D} \sim \mathcal{N}(\mathbf{0}, \Sigma_{\mathcal{X}}) \tag{3.8}$$

meaning that a deformation field can be directly drawn as a sample from the prior distribution of the vertices of the mean mesh. ***show two or three samples of prior here, next to mean mesh***

Reference Mesh Posterior The target landmarks also consist of a set $L_{\mathcal{T}} = \{l_{\mathcal{M}1}, \cdots, l_{\mathcal{M}N}\} \subset \mathbb{R}^3$. Fixing the prior output to the deformation vectors $\mathcal{Y} = \{\mathbf{t} - \mathbf{m} | \mathbf{t} \in L_{\mathcal{T}}, \mathbf{m} \in L_{\mathcal{M}}\}$ defined by the distance between the reference and target landmarks and assuming additive i.i.d Gaussian noise the resulting posterior distribution is

$$\begin{bmatrix} \mathcal{Y}_\varepsilon \\ \mathcal{D} \end{bmatrix} \sim \mathcal{N}\left(\mathbf{0}, \begin{bmatrix} \Sigma(\mathcal{Y}) + \sigma^2 \mathcal{I}_{3|\mathcal{Y}|} & \Sigma(\mathcal{Y}, X_*) \\ \Sigma(X_*, \mathcal{Y}) & \Sigma(X_*) \end{bmatrix}\right) \tag{3.9}$$

Is this a correct definition for the distribution?

The deformation model is now rendered fixed at certain landmark points in the target mesh and the goal is to find valid deformations through the set of fixed targets, analogous to the case of eq. 3.5a. The posterior model is defined as the joint distribution of all mean mesh points and the mean landmarks, conditioned on the output deformation vectors for every mean landmark with added noise.

$$\mathcal{D}|\mathcal{X} \rightarrow \mathcal{Y}_\varepsilon. \tag{3.10}$$

e now have defined a distribution over our mean face fesh. The variance of the gaussian kernel can thereby be described as a smoothing parameter P ***mean is now max aposteriori solution***

Sampling the conditional distribution creates deformed 3D surfaces of the mean mesh which are fixed at the target landmarks. ***show images of mean, prior and posterior with added landmarks***

3.5 Fitting & Optimization

Due to the fact that the posterior model is fixed at the target landmarks it is possible to perform the registration by drawing a shape from the posterior model. The sample, however, has to be optimized according to the shape of the target face. In order to find the deformation corresponding to the optimal fit d_* a linear optimization with the posterior process as a constraint is be employed/ regularization term. (small lambda) a bit of the posterior mean)

$$d_* = \underset{d \in \mathcal{D}}{\mathbf{arg\ min}} \quad L[O_{\mathcal{T}}, O_{\mathcal{M}} \circ d] + \lambda R[d] \tag{3.11}$$

Minimizing a loss function L - mean square distance for example - on the target and the deformed mean provides a feasible deformation field. D denotes the space of possible deformations in the posterior model. The information needed is held by the covariance matrix of the posterior process. However, using Mercer's theorem (in short what it does, yada yada yada) a basis of functions corresponding to all possible deformations can be extracted. The kernel is thus described as a linear model of these basis functions. ***whole matrix or simple kernel? refer to paper "a unified formulation of statistical model fitting and non-rigid registration***

$$k(x, x') = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \phi_i(x') \tag{3.12}$$

λ_i are the eigenvalues and ϕ_i the eigenvectors of K. They denote the deformation directions while the eigenvalues . . . We are looking for a finite linear combination of eigenvectors that form a deformation field with $\exists \alpha_1 \cdots \alpha_n \in \mathbb{R}$ as linear parameters.

$$f(x) = \sum_{i=1}^{n} \alpha_i \lambda_i \phi_i(x) \tag{3.13}$$

f GP(0, K) we take our gaussian process ¿ f — x=y, ***ask Marcel for a helping hand with the theory?***

Next, we want to minimize residuals for the whole of our surface according to optimal parameter values α_i

$$\underset{\alpha \in \mathbb{R}^n}{\textbf{arg min}} \quad \Sigma_{x_i \in \mathcal{T}_R} (f(x_i) - \phi_T(x_i))^2 \tag{3.14}$$

where $f(x_i)$ is the deformation function and $\varphi_T(x_i)$ returns the nearest point on the target mesh. Yields the overall loss function Φ_L

$$\Phi_L(f(x_i) - \varphi_T(x_i)) \tag{3.15}$$

The eigen vectors - which are deformation vectors defining a deformation for every model vertex - of the covariance matrix define a basis space? Shape Modell =¿ select best eigenvectors via PCA in order to simplify computation.

=¿ Vorstellen wie wenn mehrere Wellbleche durch die Target"_"landmarks gelegt werden und dann mit bestimmten parametern alpha zwischen ihnen interpoliert wird Alternative way to understand basis functions for gaussian process: sample from the GP(0, K) and then build a linear model from the functions, f(x) = sum(i, n) alpha(i) si(x) Posterior Distribution of Landmarks Defining the Gaussian Process Posterior Distribution - Landmarks (Referenz deformieren From Gaussian Processes to Shape Models =¿ by selected principal components of the covariance matrix

3.6 Robust Loss Functions

robust against outliers the Alignment of the mean face mesh and the target mesh causes overlaps on the forehead, the side of the head and the neck. Using a simple Mean Square error between the reference and target mesh for optimization penalizes the overlapping regions with a strong gradient and therefore causes strong distortions. Our approach to tackling this problem was to try out a range of different robust estimators, namely the Tukey, Huber, and Fair estimators. (table with formulas?) The advantage is that these estimators are less sensitive to outliers, reducing the artefacts of registration considerably. However, as can be seen from the formulas, these techniques require finding appropriate parameters first which produce reasonable/acceptable visual results Fair

$$\rho(x) = c^2 \left[\frac{|x|}{c} - log(1 + \frac{|x|}{c}) \right] \tag{3.16a}$$

$$\psi(x) = \frac{x}{1 + \frac{|x|}{c}} \tag{3.16b}$$

Huber

$$\rho(x) = \begin{cases} \frac{x^2}{2} & \text{if } |x| < k \\ k(|x| - \frac{k}{2}) & \text{if } |x| \geq k \end{cases} \tag{3.17a}$$

$$\psi(x) = \begin{cases} x & \text{if} \\ ksgn(x) & \text{if} \end{cases} \tag{3.17b}$$

Tukey

$$\rho(x) = \begin{cases} \frac{c^2}{6} \left(1 - \left[1 - \left(\frac{x}{c} \right)^2 \right]^3 \right) & \text{if } |x| \leq c \\ \frac{c^2}{6} & \text{if } |x| > c \end{cases} \tag{3.18a}$$

$$\psi(x) = \begin{cases} x \left[1 - \left(\frac{x}{c} \right)^2 \right]^2 & \text{if} \\ 0 & \text{if} \end{cases} \tag{3.18b}$$

Rigid Alignment This part is the theoretical part, alignment can follow in the Pipeline Therefore, we first have to perform a rigid transformation to align the meshs according to the feature/ landmark points.

Chapter 4

Registration Pipeline using Line Features

4.1 Definition of Line Features

draw bzier curves

4.2 Why use Line Features for Registration

4.3 Sampling 3D Points from 2D Line Features

Arc Length Parametrization

In theory

In practice

Projection on to the Mesh

How registration worked so far?

Proposed ”Solution”

4.4 Preparing the Mean Mesh

rendering, marking line features, projecting back possible, because we know direction

4.5 Rigid Mesh Alignment

simple rigid transformation of the scanned face onto the mean, transformation computed from landmark vectors.

4.6 Prior Model

what to say here? describe programme?

4.7 Posterior Model

what to say here? describe programme?

4.8 Fitting

4.9 Optimizing the Loss Function

4.10 Varying the Variances