

# Masterthesis Medientechnologie

mit einem eventuell  
ganz langen Untertitel

vorgelegt von

**Oliver Kalbfleisch**

vorgelegt an der

TECHNISCHEN HOCHSCHULE KÖLN

CAMPUS DEUTZ

FAKULTÄT FÜR INFORMATIONEN-,

MEDIEN- UND ELEKTROTECHNIK

im Studiengang

MEDIENTECHNOLOGIE

Erster Prüfer: Prof. Dr. Peter Silie

Technische Hochschule Köln

Zweiter Prüfer: Prof. Dr. Maria Musterprof

Technische Hochschule Köln

Köln, im August 2017

**Adressen:** Max Mustermann  
Musterstraße 1  
12345 Musterstadt  
max@mustermann.de

Prof. Dr. Peter Silie  
Technische Hochschule Köln  
Institut für Informatik  
Steinmüllerallee 1  
51643 Gummersbach  
peter.silie@th-koeln.de

Prof. Dr. Maria Musterprof  
Technische Hochschule Köln  
Institut für Informatik  
Steinmüllerallee 1  
51643 Gummersbach  
maria.musterprof@th-koeln.de

# Contents

<b>1</b>	<b>Einleitung</b>	<b>4</b>
<b>2</b>	<b>Description of hand in digital space</b>	<b>5</b>
2.1	Physiological structure of the human hand . . . . .	5
2.1.1	Constraints in Hand Motion . . . . .	6
2.2	Kinematics . . . . .	8
2.2.1	Forward Kinematics . . . . .	8
2.2.2	Inverse Kinematics . . . . .	9
2.2.3	FABRIK Algorithm . . . . .	12
2.3	Digital hand models . . . . .	14
2.3.1	1 . . . . .	14
<b>3</b>	<b>Tracking in real space</b>	<b>15</b>
3.1	General tracking technologies . . . . .	15
3.1.1	Optical tracking . . . . .	15
3.1.2	Magnetic tracking . . . . .	16
3.1.3	Accoustic tracking . . . . .	17
3.2	Hand tracking Systems . . . . .	17
3.2.1	Sensor based approach . . . . .	17
3.2.2	optical approach . . . . .	18
<b>4</b>	<b>Fazit</b>	<b>20</b>
	<b>Abbildungsverzeichnis</b>	<b>21</b>
	<b>Tabellenverzeichnis</b>	<b>22</b>
	<b>Literaturverzeichnis</b>	<b>25</b>

# Kurzfassung

Hier folgt die Kurzfassung.

# **Abstract**

Hier folgt die Kurzfassung auf Englisch.

# 1 Einleitung

The standard interface between human and computer has for long years been mouse and keyboard. But with the advance of technology, new interfacing methods were developed in the last few years.

Touch technology for interfacing with mobile devices and desktop computers has become a reliable technology and has been integrated into our everyday lives. Advances in capabilities of CPU as well as GPU hardware has build a foundation for the use of advanced AR and VR Technology. 3D and Stereoscopic rendering can now be accomplished even by mobile hardware (with some limitations)without the need of specifi Hardware. For An intently immersive experience VR Googles are used to explore digitally created worlds. But with this level of immerivenes, a touch device or even a mouse and keyboard setuop is rather hindering the user experience. First attempts of solving this problem came with the introduction of tracked controllers for the interaction with the digital world, but these can also only supply a fration of the capabilities of our natural interaction devices, namely the human hands.

## 2 Description of hand in digital space

Tracking of the human hand has always been a challenging Problem. In comaprison to other larger bodyparts like the Arm or the head, the human hand itself contains a large variety of smaller parts, namely bones and muscles. These components have to be taken into acount when trying to replicate the natural motion of the hand in digital space.

### 2.1 Physiological structure of the human hand

[13] describes the human hand as "an articulated structure with about 30 degrees of freedom [which] changes shape in various ways by its joint movemnnents."

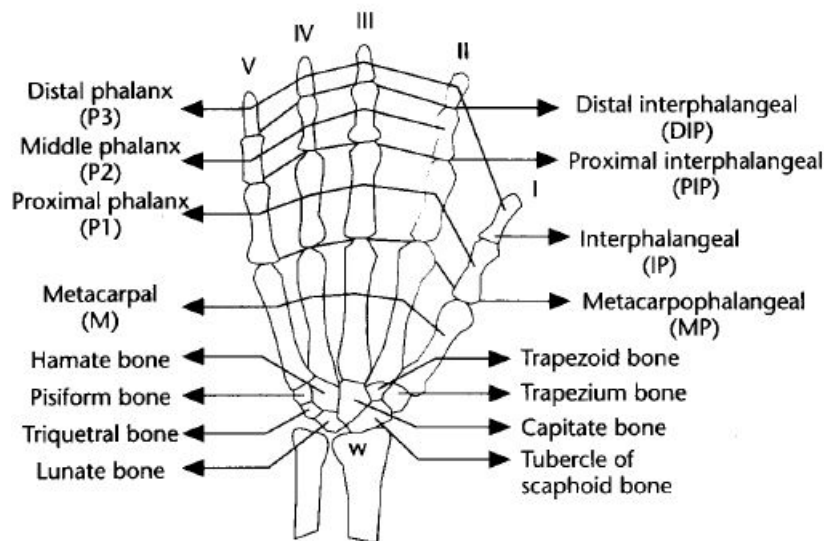


Figure 2.1: Bone structure of the human left hand ([13])

All of the hand components are connected to at least one neighboring component via a joint. Teh joints affect the position of the connected components. To describe the movement of the hand components, we can use the roation angles of the joints to correlate to a specific position.

To do so, we define a local coordinate system for each of the exiting hand joints. By doing so, we achieve a sequence of rotaions in the local coordinate systems of the joints. Such a sequence can then be used to describe a specific movement and/or position of

a component. Not all of the joints in the human hand have equal degrees of freedom. Their functionality can be classified in the amount of DOFs (Degrees of freedom)[11]

- 1 DOF
  - A joint movement that can perform a **flexion** or **twist** in one direction
- 2 DOF
  - A joint movement that can perform **flexion** in more than one direction (**directive**)
- 3 DOF
  - A joint movement that permits simultaneous **directive** and **twist** movements. (**spherical**)

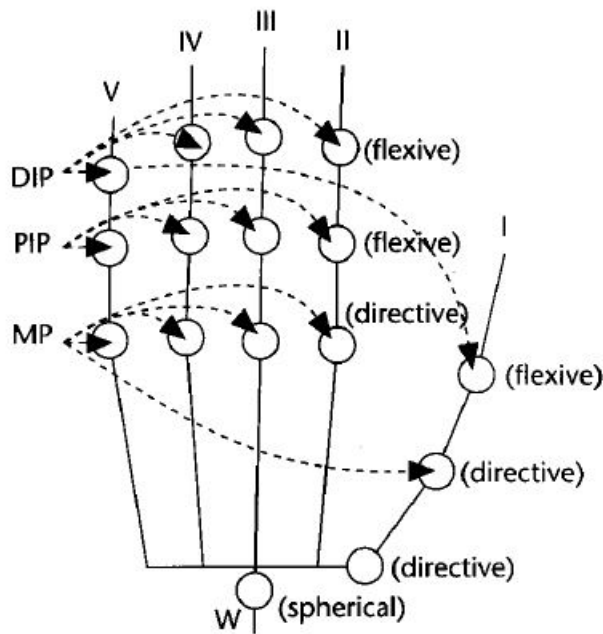


Figure 2.2: Representation of the DOFs of the human hand

When looking at the DOFs displayed in Figure 2.2, each finger (II-V) sums up to 4 DOFs and the thumb to 5 DOFs. Also considering 6 DOFs for the rotation and position of the whole and itself, the result gets us to 27 DOFs for the human hand.

### 2.1.1 Constraints in Hand Motion

A full usage of all the declared DOFs would lead to a large amount of possible combinations. Since the hand is not only made up of bones but also Muscles and the skin, we can impose some constraints ([4]) to the movement of the joints. Ling, Wu and Huang ([14]) proposed the following classification for the constraints:



- **Type I constraints**

- A constraint that limits the range of finger motions based on hand anatomy

- **Type II constraints**

- A constraint that the position of the joints during finger movement

- **Type III constraints**

- A constraint that limits position based on natural hand motions

The **Type I** and **Type II** constraints rely on the physiological and mechanical properties of the hand. **Type III** constraints are results of common and natural movements and can be differing from person to person. As these movements are to some degree similar for everyone, a broad grouping can be applied. The curling of the fingers at the same time when forming a fist is way more natural than curling each finger by itself. Here the motion of the hand is quite similar between different persons, but the constraints cannot be described in a mathematical form.

A **Type I** constraint example would be that the position of the fingertip is limited by the length of the other finger segments and thereby can only reach as far as the combined length.

An example for **Type II** constraints would be that, for your fingertip to touch your hand palm, all joints in the finger have to be bent to achieve this position. The following inequalities can be used to describe these constraints:

**Type I:**

$$\begin{aligned} 0^\circ &\leq \Theta_{MP\_flex} \leq 90^\circ \\ 0^\circ &\leq \Theta_{PIP\_flex} \leq 110^\circ \\ 0^\circ &\leq \Theta_{DIP\_flex} \leq 90^\circ \\ -15^\circ &\leq \Theta_{MP\_abduct/adduct} \leq 15^\circ \end{aligned} \tag{2.1}$$

A further constraint that is specific to the middle finger is, that this finger's MP normally does not abduct and adduct much. Therefore we can infer an approximation and thereby remove 1DOF from the model:

$$\Theta_{MP\_abduct/adduct} = 0^\circ \tag{2.2}$$

The same behavior can be seen in the combination of hand parts labeled W (the connection point between hand and lower arm). This approximation also eliminates one DOF on the connected thumb:

$$\Theta_{W\_abduct/adduct} = 0^\circ \tag{2.3}$$

Since the DIP, PIP and MP joints of our index, middle, ring, and little fingers only have 1DOF for flexion, we can further assume that their motion is limited to movement in one plane.

### Type II:

The **Type II** constraints can be split into interfinger and intrafinger constraints. Regarding intrafinger constraints between the joints of the same finger, human hand anatomy implies that to bend the DIP joints on either the index, middle, ring or little fingers, the corresponding PIP joints of that finger must also be bent. The approximation for this relation [18] can be described as :

$$\Theta_{DIP} = \frac{2}{3}\Theta_{PIP} \quad (2.4)$$

Interfinger constraints can be imposed between joints of adjacent fingers. Interfinger constraints describe that the bending of an MP joint in the index finger forces the MP joint in the middle finger to bend as well.

When combining the constraints described in the above equations, the starting number 21 DOF's of the human hand can be reduced to 15. Inequalities for these cases, obtained through empiric studies, can be found in [13].

## 2.2 Kinematics

The preceding sections gave an overview of how we can describe a model of the human hand and introduced some limiting constraints. With the model and the constraints, we can now start to build a kinematic system for the animation of the model.

Kinematic systems contain so called *kinematic chains*, which consist of a *starting point* or *root*, kinematic elements like *joints*, *links* and an *endpoint*, also called *end effector*. Applied to the human hand, the whole hand model represents the kinematic system. This system contains several *kinematic chains*, namely the fingers of the hand with the fingertips being the *end effectors* of each of these chains.

As we begin to move our hands, the states of the kinematic chains begin to change. Joint angles and end effector positions are modified until the end position is reached. To represent the new position and angle dataset of our physical hand with our kinematic system, two major paths for achieving a solution can be taken.

### 2.2.1 Forward Kinematics

Forward Kinematics (FK) uses the knowledge of the new angles and positions after the application of known transformations to the kinematic chain. The data of the *joints* and *links* between the *root* and the *end effector* is then used to solve the problem of finding the *end effector's* position.

We can denote the existing end effectors relative position to an origin as  $s_1, \dots, s_k$ . The  $s_i$  position is the result of the combination of all the joint angles in the corresponding kinematic chain. Respectively, we define the target position of the end effectors as  $t_1, \dots, t_i$ ,

with  $t_i$  beeing the target positiojn for the end effector  $s_i$ . The required positional change for the end effector can now be described as  $e_i = t_i - s_i$ . In systems with more than one end effector, like our hand system, the components can be written as vectors.

$$\begin{aligned}\vec{s} &= (\mathbf{s}_1, \dots, \mathbf{s}_n)^T \\ \vec{t} &= (\mathbf{t}_1, \dots, \mathbf{t}_n)^T \\ \vec{e} &= \vec{t} - \vec{s}\end{aligned}\tag{2.5}$$

As the vector components of  $\vec{s}$  are reults of the chain joint angles  $\theta_1, \dots, \theta_n$  and therefore are effected by them, we define

$$\begin{aligned}\vec{s}_i &= f_i(\theta) \\ \vec{s} &= f(\theta)\end{aligned}\tag{2.6}$$

With  $\theta$  beeing the column vector  $\theta = (\theta_1, \dots, \theta_n)^T$ . The second vector equation displayed in (2.6) is also callled the *Forward Kinematics*(FK) solution.

The advantage of an FK solution is that there is always an unique solution to the problem. In consequence, this approach is commonly used in the field of robotics, where the information on the chain elements is easily available.

The tracking of the human hand and all if its chain components is rather complicated. Therefore a solution which takes a known position of the *end effektor* and calculates the parameters for the rest of the cain would be more desirable.

### 2.2.2 Inverse Kinematics

Inverse Kinematics (IK) is a method for computing the posture via estimating each individual degree of freedom in order to satisfy a given task [2, S. 14]

The concept of *Inverse Kinematics* (IK) already describes it's principle in it's name. It takes the reversed approach in comparison to the FK principle in chapter 2.2.1. Instead of knowing the states of the chain elements and calculating the resulting position of the *end effector*, we take the position of the *end effector* and try to retrieve the possible states of the other chain elements.

$$\theta = f^{-1}(\vec{s}_d)\tag{2.7}$$

The result of this equation is the vector  $\theta$  for which the values of  $\vec{s}$  coincide to the desired configuration  $\vec{s}_d$ . In the case of an optimal result, this configuration would have the same position values as the target positions.

The main problem with this method occurs in the calulation of the  $f^{-1}$  function, due to it beeing a highly non linear operator which is not easily invertible. Some of the approaches

that are used to counter this problem will be displayed later on in this chapter.

In contrary to having a unique solution with the FK approach, the IK approach can end at the point of not finding a suitable solution. Figure 2.3 displays three possible outcomes for the IK approach.

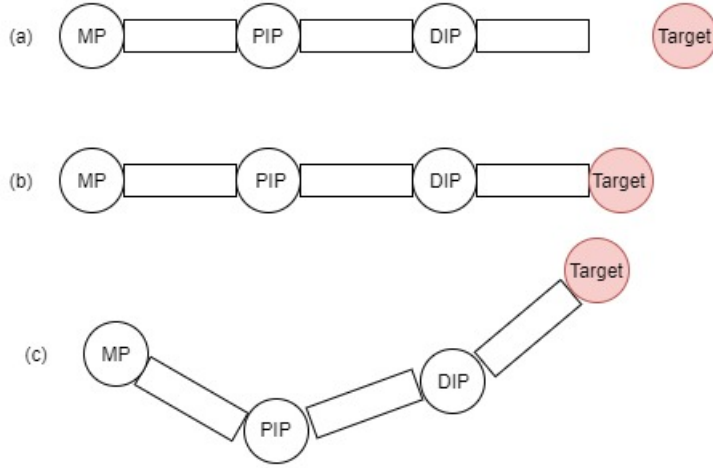


Figure 2.3: Possible solution for an IK problem of a human finger:

- (a) The given target position of the end effector can not be reached. (b) The given target can only be reached by one solution. (c) The target position can be reached with multiple different solutions.

## Nonlinear Programming

### Jacobian Inverse

One common approach to solve the IK problem is the utilization of a Jacobian Matrix and an iterative calculation process. This matrix contains the partial derivatives of the chain systems relative to the end effector  $\mathbf{s}$ . When using the Jacobian, a linear approximation of the IK problem will be applied for solving. The approximation's components model the end effector's relative movement to changes in transitions of the systems link translations and joint angles. Therefore, the resulting function is dependent on the joint angles  $\theta$  values and can be defined as

$$J(\theta)_{ij} = \left( \frac{\partial \mathbf{s}_i}{\partial \theta_j} \right)_{ij} \quad (2.8)$$

with  $i=1, \dots, k$  and  $j=1, \dots, n$ . Further readings on methods for the calculation of Jacobian matrices can be found in [16]. Based on definition (2.8), an entry for the  $j$ -th rotational joint would be calculated as follows:

$$\frac{\partial \mathbf{s}_i}{\partial \theta_j} = \mathbf{v}_j \times (\mathbf{s}_i - \mathbf{p}_j) \quad (2.9)$$

where  $\mathbf{p}_j$  is the position of the joint, and  $\mathbf{v}_j$  is the unit vector pointing along the current axis of rotation for the joint. Taking the derivative of definition (2.6) with respect to time

gives the basic equation for forward kinematics that describes the velocities of the end effectors:

$$\dot{\vec{s}} = J(\theta)\dot{\theta} \quad (2.10)$$

Now having all the values for the angles, the *end effector* position and the target positions, we can compute the resulting Jacobian matrix. Thereafter we seek an update value  $\Delta\theta$  for incrementing the current joint values:

$$\theta_{new} = \theta_{curr} + \Delta\theta \quad (2.11)$$

The idea here is that the chosen value for  $\Delta\theta$  should lead to the resulting  $\Delta\vec{s}$  being approximately equal to  $\vec{e}$  from (2.5). The  $\vec{e}$  can be approximated by:

$$\Delta\vec{s} \approx J(\theta_{curr})\Delta\theta \quad (2.12)$$

Using this approximation we can reformulate the FK problem as  $\vec{e} = J\Delta\theta$  and therefore our inverse kinematics problem from 2.7 can be expressed as  $\Delta\theta = J^{-1}\vec{e}$ .

The problem we run into with this solution is the construction of the inverse Jacobian matrix. The Jacobian  $J$  may not be square or invertible. In the case of it being invertible, the result may only work inferiorly because of it being nearly singular. Being singular means that no change in joint angle values may achieve the desired end effector position as an outcome.

### Jacobian Transpose

One approach to calculating the value of  $\Delta\theta$  without having to calculate the inverse of  $J$  is done by replacing the inverse with the transpose of  $J$ .

$$\Delta\theta = \alpha J^T \vec{e} \quad (2.13)$$

Of course the transpose and the inverse of  $J$  are not the same thing. When using the theorems displayed in [1, 23] we can show that:

$$\langle JJ^T \vec{e}, \vec{e} \rangle = \langle J^T \vec{e}, J^T \vec{e} \rangle = \|J^T \vec{e}\|^2 \geq 0 \quad (2.14)$$

Under a sufficiently small  $\alpha > 0$  the updated angles from 2.11 will change the end effector positions by approximately  $\alpha JJ^T \vec{e}$ . They also state that the value of  $\alpha$  can be calculated by minimising the new value of the error vector  $\vec{e}$  after each update.

$$\alpha = \frac{\langle \vec{e}, JJ^T \vec{e} \rangle}{\langle JJ^T \vec{e}, JJ^T \vec{e} \rangle} \quad (2.15)$$

### Jacobian Pseudo Inverse

Instead of calculating the normal inverse of the Jacoboian, which can lead to the problems described before, we can use the so called *pseudo-inverse*[6] for the calculation. The *pseudoinverse* is defined for all matrices  $J$ , even ones which are not square or not of full row rank.

$$\Delta\theta = J^\dagger \vec{e} \quad (2.16)$$

The *pseudoinverse* represents the best possible solution for  $J\Delta\theta = \vec{e}$  in respect to least squares. It has the favourable properties thath if the  $\vec{e}$  Vector lies within the range of the  $J$

#### 2.2.3 FABRIK Algorithm

A more recent approach in the kinematics field is the **FABRIK** algorithm proposed by Aristidou and Lasenby[3].As shown in section 2.2.2, these approaches all depend on computational intensive matrix operation like calculating an inverse and may have problems with matrix singularity.

The **FABRIK** algorithm does not depend on these matrix operations as is solves for the position of a point on a line to retrieve the new joint positions. This is done in an forward and also inverse solving approach, iterating these steps until the calculated position converges towards the target position from the tracking data.

Figure 2.4 illustrates the steps that are contained in one iteration step. The chain joints are denoted as  $\mathbf{p}_i$  with the distance  $\mathbf{d}_i$  being  $|\mathbf{p}_{i+1} - \mathbf{p}_i|$ . The target point for the end effector is denoted as  $\mathbf{t}$ . Step (a) displays the starting pont of the iteration. The joint positions are taken from either a previous iteration or from an initial calibra-

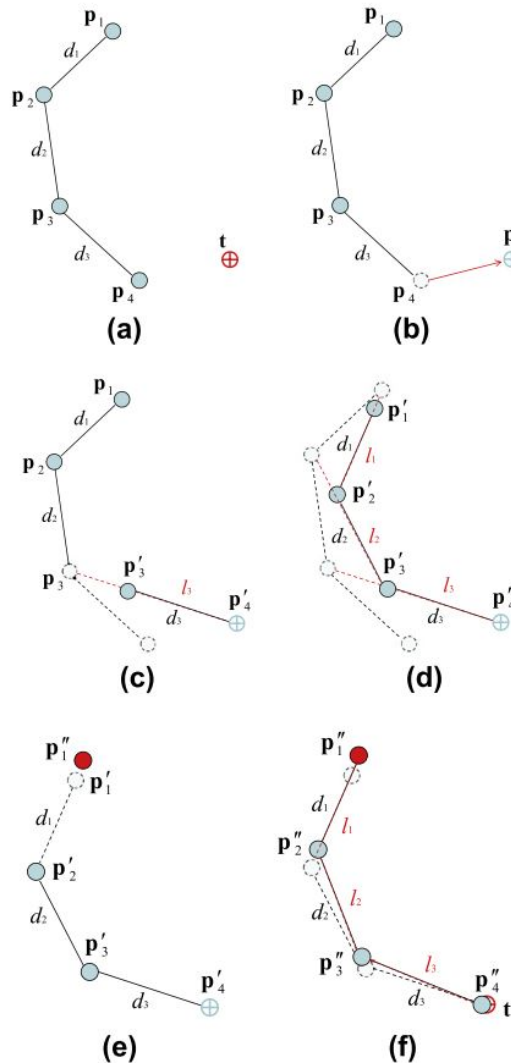


Figure 2.4: Forward and backward calculation steps for one iteration of the FAB-

tion.

But before calculations can begin, the algorithm has to check whether the intended target point  $\mathbf{t}$  is reachable for the end effector. This is done by measuring the distance between the root of the kinamtic chain and the target point  $\mathbf{t}$ . This value is then compared with the sum of the distances  $\mathbf{d}_i$ .

$$dist_{t,d_1} < \sum_{k=1}^i d_i \quad (2.17)$$

The inverse calculation step is the first step, which is displayed in **(b)** and **(c)**. The calculation is started at the end effector, moving to the root of the chain. If the summed distance is greater, then the  $\mathbf{t}$  is within the reach of the system and the calculation can continue, otherwise the calculation has to be aborted and the error has to be manged otherwise. Assuming this requirement to be met, we can now begin with the first calculation. Therefore we assume that the new position  $\mathbf{p}'_n$  with  $n=4,...,1$  is equal to  $\mathbf{t}$ .

$$\mathbf{p}'_n = \mathbf{t} \quad (2.18)$$

From this new point, we can construct a line that goes through  $\mathbf{p}'_n$  and  $\mathbf{p}_{n-1}$ .

$$\begin{aligned} A &= \mathbf{p}'_n \\ B &= \mathbf{p}_{n-1} \\ \mathbf{l}_{n-1} &= \overline{AB} \end{aligned} \quad (2.19)$$

The resultng position of the new  $\mathbf{p}'_{n-1}$  point is located in this line with the distance of  $\mathbf{d}_{n-1}$  from  $\mathbf{p}'_n$  (see **(c)**).

$$\mathbf{p}'_{n-1} = \mathbf{p}'_n + \left( \frac{\overline{AB}}{|\overline{AB}|} \cdot \mathbf{d}_{n-1} \right) \quad (2.20)$$

Consecutively, this is done with the remaining joints until the root joint is reached(see **(d)**). This finishes the first halv of the iteration step. With the calculated positions, we now perform a forard calculation, starting from the root until we reach the end effector. Since the root of the system normally does not move from it's initial position, we have to reset the root joint to this value before starting to calculate the new positions of the subsequent joints(see**(e)**).

Analogous to the procedure in the inverse step, we construct the lines between the points and determine the new positon values of the joints. The end result of this step is shown in **(f)**. At this point, we can decide if the result position of the end effector is appropriate in comparison to the value of  $\mathbf{t}$ . A simple threshold value for this case could be the position difference between these two points.

## **2.3 Digital hand models**

### *2.3.1 1*



## 3 Tracking in real space

The previous sections provided information on the structure of the human hand and how to represent it in the digital space. To apply the algorithms presented in Section 2.2 to our digital hand model, we need input data from our real world representative.

### 3.1 General tracking technologies

The methods for gaining positional data can be roughly categorized into two major groups, glove based methods and vision based methods. Glove based methods have already been in development since the 1980's [5] and have since then resulted in several solution attempts. Sturman and Zeltzer gave a survey on the existing tracking methods in their paper [20].

They distinguish between two areas of tracking, first the 3D positional tracking of the hand (and also other bodyparts) without regard to the hands shape and secondly the tracking of the hand shape with glove technologies.

These tracking technologies presented are still applied today in modern tracking solutions [19, 22]. They account for solutions based in optical tracking based on marker detection, magnetic detection via measurements of an artificial magnetic field [17] and acoustic measurements via triangulation of ultrasonic pings.

#### 3.1.1 Optical tracking

The components for an optical tracking systems are several cameras for object detection and some kind of tracking characteristic of the object to be tracked. These characteristics can be either artificially applied ones like active flashing infrared LED on key tracking positions of the body or infrared reflective markers.

A series of cameras positioned around the tracking subject will then track these markers inside their visual fields. The second method uses a single camera to capture the silhouette image of the subject, which is analyzed to determine positions of the various parts of the body and user gestures.

The image data is supplied to special software which correlates the marker positions in the multiple viewpoints and uses the different lens perspectives to calculate a 3D coordinate for each marker. These image interpretation and correlation tasks require computationally costly operations. The marker tracking is also prone to errors through variation in lighting of the scene, material reflection properties and also marker occlusion as the trackers are moved. Also most of the systems rely on several tracking cameras for

### 3 Tracking in real space

a complete coverage of the tracking space. This leads to a higher system complexity in terms of setup and calibration.

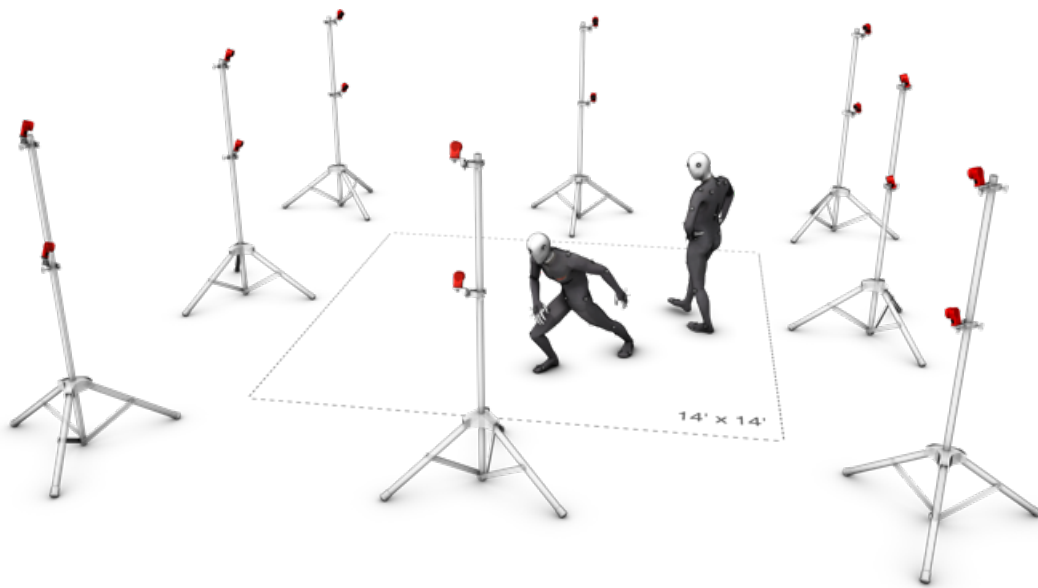


Figure 3.1: Example for an optical tracking system with passive infrared reflector markers [15]

#### 3.1.2 Magnetic tracking

The usage of a magnetic field for position tracking is a relatively uncomplicated technique. The earth already provides us with a magnetic field to orient on. Similar to the optical trackers, magnet field sensors can be placed at key tracking positions. These sensors measure field strength in 3 orthogonally oriented axes to produce a 3D vector of the unit's orientation with respect to the excitation. As the earth's magnetic field is prone to changes based on geographical location, data corrections have to be applied to the measurements.

Another solution is to generate a local magnetic field via a multicoil source unit. The source unit coils are energized in sequence and the corresponding magnetic field vector is measured in the sensor unit. With three such excitations, you can estimate the position and orientation of the sensor unit with respect to the source unit.

The downside of this technology is that ferromagnetic and conductive material in the surrounding environment will have an effect on the generated magnetic field. These distortion effects will form small unwanted secondary source units through the induction of small Foucault currents by the main source's magnetic field. Magnetic fields have the property of having an inverse cubic falloff as a function of distance from the source, which limits the range of operation for the system. Position resolution in the radial direction from source to sensor depends on the gradient of the magnetic field strength, and thus the positional jitter grows as the fourth power of the separation distance.

In comparison to other tracking technologies, the magnetic field tracking solution has the convenience of not suffering from line of sight problems from tracker occlusion. The

magnetic fields are capable of passing through the human body. Also the sensor size for measuring magnetic fields is rather small, giving the trackers a small volume. Furthermore, only one source unit is needed for the tracking of multiple sensor units.

#### 3.1.3 *Acoustic tracking*

The principles of acoustic tracking are very similar to those of the optic tracking technologies. Instead of using lightwaves, the systems utilize acoustic pulses of ultrasonic wavelengths to time the time of flight between emitter and sensor for range measurement. To get a good measuring result from the systems, the used acoustic transducers have to be as omnidirectional as possible, so that the signal can be detected no matter how the emitter is positioned or oriented in the tracking volume. For the speakers to achieve a wide beam width, their size has to be small.

To be able to build the microphones into the tracker, they can only have active surfaces a few millimeters in diameter. This leads to a reduction in range as the efficiency of acoustic transducers is proportional to the active surface area. Also acoustic systems can have problems with ambient noises occluding the signal. This becomes even more critical when using such a system outdoors.

Soundwaves travel at a much slower speed than lightwaves which brings benefits and downsides with it. Soundwaves can be reflected from objects, producing so echos which arrive at the receiving sensor at a later point in time. Here the slower speed can be beneficial as we can await the first sound occurrence to arrive at the sensor and filter out all later reflections from the data. The reflections of the previous pulse also have to be subsided before a new measurement can be made, lowering the update rates of the system. The air the soundwaves travel through is also a limiting factor as humidity, air pressure and air currents can influence the travelling soundwaves. In comparison to the optical systems, the acoustic systems are not as prone to occlusion errors since soundwaves have a better ability to bend around obstacles than lightwaves.

Most of these downsides can be addressed with a combination of these systems with another form of tracking like in [8].

## 3.2 Hand tracking Systems

In comparison to the tracking of the whole human body, the tracking of the human hands with their maximum of 27 DoF's each in a small space is rather difficult to handle.

#### 3.2.1 *Sensor based approach*

One popular approach for tracking the hand movements is via a sensor based approach.

#### 3.2.2 optical approach

Optical approaches use properties of the human hand to hand to estimate the current position. The evaluation of the hand pose is usually split up into two steps. The first one is the registration of the "real" hand. This is mostly done by at least one camera which is aimed at the hand, supplying a "continuous" data stream to the evaluation Software. In the second step, the evaluation software then tries to find key spots in the provided images. This data is then used for the search of the matching digital pose. The definition of these key spots can be achieved by several techniques.

An example is the usage of some kind of marking for finger segments as displayed in [7, 9, 21]. In the method described by Wang and Popovic [21], a specially colored glove is used. The glove is segmented into ten segments, colored randomly from a pool of ten distinct colors (see Figure 3.3). The color pool was



Figure 3.2: Color glove setup used by Wang and Popovic [21]

limited by the color distinguishing capabilities of the camera used in the experiment. The setup used was engineered to fit the standard home user capabilities, so when using a higher grade camera more color variation could be possible. They used more large pattern areas on the glove to minimize the risk of occlusion problems which could appear with smaller patch areas. The pattern is created by selecting twenty seed triangles on a 3-D hand model that are maximally distant from each other. The remaining triangles are assigned into patches based on their distance to each of these seeds. Each patch is assigned one of the ten colors at random. The jagged boundaries between the patches are artifacts from the low triangle count of the used hand model.

Algorithms that use other characteristics for tracking like texture or shading [12] rely on an accurate pose from a previous frame to constrain the pose search in the current frame. When using bare hand pose estimation two different hand poses can map to similar images. This can lead to an inaccurate pose estimation and the breakdown of the algorithm if the wrong estimation is made. The benefit of the color glove is that the unique patch pattern, hand poses always map to distinguishable images, simplifying the search process. Therefore it can effectively recover itself at each frame without the need of a previous frame.

With the colored glove as a tracking feature, the images from the camera are prepared for a database sampling step. The database that was used consisted of 18,000 finger configurations. A distance metric between two configurations was defined as the root mean square (RMS) error between the vertex positions of the corresponding skinned hand models. With this distance metric, a low dispersion low-dispersion sampling was used to draw

a uniform set of samples from the overcomplete collection of finger configurations. The selected configurations are rendered at various 3-D orientations using a (synthetic) hand model at a fixed position from the (virtual) camera. The camera image was compared then compared to the database images. The comparison is done via a nearest neighbor lookup using a Hausdorff-like distance[10] and penalizing the distance to the closest pixel of the same color in the other image. The resulting pose from the database can then be applied to the digital hand model.

The method presented by Fredriksson and Ryen [9] also uses a color coded glove. In contrast to the fully colored glove described before, the used glove only has colored fingerstips, color markers for the palm and a colored band for the wrist. The palm markers and the wrist band are used to retrieve the 3D position of the hand. The wrist marker is furthermore used to define the boundingbox of the tracking area in a calibration process. Defining a bounding greatly simplifies further image analysis operations by taking away a large amount of unneeded image data. After determining position and orientation of the hand in 3d by utilizing wrist band and palm marker position, the system tracks the grip angle and the lateral tilt angle for each finger. The

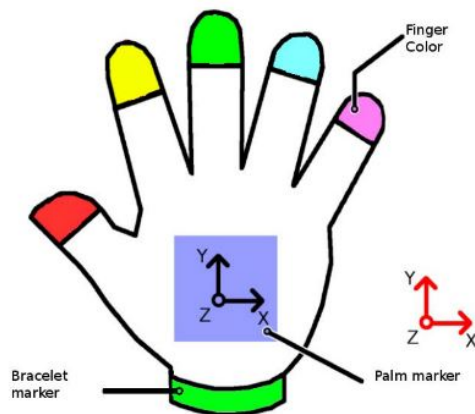


Figure 3.3: Color glove setup used by Fredriksson and Ryen [9]

grip angle is defined as the curling of the finger towards the palm. The second angle is defined as the lateral tilt angle which measures the spread of the fingers. The grip angle is calculated by comparing the density of finger pixels around an estimated knuckle line. From the density, an estimation for the finger's Y position around an origin position at the knuckle's base is made. This value is then transformed into an angle value using an inverse tangent operation.

One downside of this tracking method is that it does not yet incorporate the movement of the thumb. The thumb has other movement possibilities and therefore needs a different algorithmic approach. Also, the approach only focuses on the tracking of the hand data and does not yet implement a digital counterpart like the previously described system.

## 4 Fazit

## List of Figures

2.1	Bone structure of the human left hand ([13]) . . . . .	5
2.2	Representation of the DOFs of the human hand . . . . .	6
2.3	Possible solution for an IK problem of a human finger: (a)The given target position of the end effector can not be reached. (b) The given target can only be reached by one solution.(c) The target position can be reached with multiple different solutions. . . . .	10
2.4	Forward and backward calculation steps for one iteration of the <b>FAB-RIK</b> algorithm.(a) initial position of the system.(b)End effector is moved to target position.(c) Deterimne position of next joint on constructed line.(d)repeat until root is reached. (e)move root joint to initial position. (f) repeat calculation in reverse direction [3] . . . . .	12
3.1	Example for an optical tracking system with passive infrared reflector markers [15] . . . . .	16
3.2	Color glove setup used by Wang and Popovic [21] . . . . .	18
3.3	Color glove setup used by Fredriksson and Ryen [9] . . . . .	19

## List of Tables



# Bibliography

- [1] Robust Control of Robotic Manipulators. In: *IFAC Proceedings Volumes* 17 (1984), Nr. 2, S. 2435–2440. [http://dx.doi.org/10.1016/S1474-6670\(17\)61347-8](http://dx.doi.org/10.1016/S1474-6670(17)61347-8). – DOI 10.1016/S1474-6670(17)61347-8. – ISSN 1474-6670
- [2] ANDREAS ARISTIDOU AND JOAN LASENBY: Inverse Kinematics: a review of existing techniques and introduction of a new fast iterative solver.
- [3] ARISTIDOU, Andreas ; LASENBY, Joan: FABRIK: A fast, iterative solver for the Inverse Kinematics problem. In: *Graphical Models* 73 (2011), Nr. 5, S. 243–260. <http://dx.doi.org/10.1016/j.gmod.2011.05.003>. – DOI 10.1016/j.gmod.2011.05.003. – ISSN 15240703
- [4] BADLER, Norman ; MANOOCHHEHRI, Kamran ; WALTERS, Graham: Articulated Figure Positioning by Multiple Constraints. In: *IEEE Computer Graphics and Applications* 7 (1987), Nr. 6, S. 28–38. <http://dx.doi.org/10.1109/MCG.1987.276894>. – DOI 10.1109/MCG.1987.276894. – ISSN 02721716
- [5] BOLT, Richard A.: Put-that-there. In: THOMAS, James J. (Hrsg.): *Proceedings of the 7th annual conference on Computer graphics and interactive techniques*. New York, NY : ACM, 1980. – ISBN 0897910214, S. 262–270
- [6] DAHMEN, Wolfgang ; REUSKEN, Arnold: *Numerik für Ingenieure und Naturwissenschaftler*. Zweite, korrigierte Auflage. Berlin Heidelberg : Springer-Verlag Berlin Heidelberg, 2008 (Springer-Lehrbuch). <http://dx.doi.org/10.1007/978-3-540-76493-9>. – ISBN 9783540764939
- [7] DUCA, Florin ; FREDRIKSSON, Jonas ; FJELD, Morten: Real-time 3d hand interaction: Single webcam low-cost approach. In: *Proceedings of the Workshop at the IEEE Virtual Reality 2007 Conference; Trends and Issues in Tracking for Virtual Environments*, 2007, S. 1–5
- [8] FOXLIN, Eric ; HARRINGTON, Michael ; PFEIFER, George: Constellation. In: CUNNINGHAM, Steve (Hrsg.): *Proceedings of the 25th annual conference on Computer graphics and interactive techniques*. New York, NY : ACM, 1998. – ISBN 0897919998, S. 371–378
- [9] FREDRIKSSON, Jonas ; RYEN, Sven B. ; FJELD, Morten: Real-time 3D hand-computer interaction. In: TOLLMAR, Konrad (Hrsg.): *Proceedings of the 5th*

- Nordic conference on Human-computer interaction building bridges*. New York, NY : ACM, 2008. – ISBN 9781595937049, S. 133
- [10] HUTTENLOCHER, D. P. ; KLANDERMAN, G. A. ; RUCKLIDGE, W. J.: Comparing images using the Hausdorff distance. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 15 (1993), Nr. 9, S. 850–863. <http://dx.doi.org/10.1109/34.232073>. – DOI 10.1109/34.232073. – ISSN 01628828
  - [11] KOREIN, James U.: *A geometric investigation of reach: Zugl.: Pennsylvania Univ., Diss. : 1984*. Cambridge, Mass. : MIT Press, 1985 (ACM distinguished dissertations). – ISBN 0262111047
  - [12] LA GORCE, Martin de ; PARAGIOS, Nikos ; FLEET, David J.: Model-based hand tracking with texture, shading and self-occlusions. In: *IEEE Conference on Computer Vision and Pattern Recognition, 2008*. Piscataway, NJ : IEEE Service Center, 2008. – ISBN 978-1-4244-2242-5, S. 1–8
  - [13] LEE, Jintae ; KUNII, T. L.: Model-based analysis of hand posture. In: *IEEE Computer Graphics and Applications* 15 (1995), Nr. 5, S. 77–86. <http://dx.doi.org/10.1109/38.403831>. – DOI 10.1109/38.403831. – ISSN 02721716
  - [14] LIN, John ; WU, Ying ; HUANG, T. S.: Modeling the constraints of human hand motion. In: *Proceedings, Workshop on Human Motion*. Los Alamitos, Calif : IEEE Computer Society, 2000. – ISBN 0-7695-0939-8, S. 121–126
  - [15] OPTITRACK: *flex13MocapVolume.png (680×375)*. <http://www.optitrack.com/public/images/flex13MocapVolume.png>. Version: 2017
  - [16] ORIN, David E. ; SCHRADER, William W.: Efficient Computation of the Jacobian for Robot Manipulators. In: *The International Journal of Robotics Research* Vol.3 (1984), Nr. No.4. – ISSN 0278-3649
  - [17] RAAB, Frederick ; BLOOD, Ernest ; STEINER, Terry ; JONES, Herbert: Magnetic Position and Orientation Tracking System. In: *IEEE Transactions on Aerospace and Electronic Systems* AES-15 (1979), Nr. 5, S. 709–718. <http://dx.doi.org/10.1109/TAES.1979.308860>. – DOI 10.1109/TAES.1979.308860. – ISSN 0018-9251
  - [18] RIJPKEMA, Hans ; GIRARD, Michael: Computer animation of knowledge-based human grasping. In: THOMAS, James J. (Hrsg.): *Proceedings of the 18th annual conference on Computer graphics and interactive techniques*. New York, NY : ACM, 1991. – ISBN 0897914368, S. 339–348
  - [19] ROLLAND, Jannick P. ; BAILLOT, Yohan ; GOON, Alexei A.: A SURVEY OF TRACKING TECHNOLOGY FOR VIRTUAL ENVIRONMENTS: Center for Research and Education in Optics and Lasers (CREOL) University of Central Florida,

- Orlando FL 32816. In: *(Fundamentals of wearable computers and augmented reality* (2001)
- [20] STURMAN, D. J. ; ZELTZER, D.: A survey of glove-based input. In: *IEEE Computer Graphics and Applications* 14 (1994), Nr. 1, S. 30–39. <http://dx.doi.org/10.1109/38.250916>. – DOI 10.1109/38.250916. – ISSN 02721716
- [21] WANG, Robert Y. ; POPOVIĆ, Jovan: Real-time hand-tracking with a color glove. In: *ACM transactions on graphics (TOG)* Bd. 28, 2009, S. 63
- [22] WELCH, G. ; FOXLIN, E.: Motion tracking: No silver bullet, but a respectable arsenal. In: *IEEE Computer Graphics and Applications* 22 (2002), Nr. 6, S. 24–38. <http://dx.doi.org/10.1109/MCG.2002.1046626>. – DOI 10.1109/MCG.2002.1046626. – ISSN 02721716
- [23] WOLOVICH, W. ; ELLIOTT, H.: A computational technique for inverse kinematics. In: *The 23rd IEEE Conference on Decision and Control*, IEEE, 1984, S. 1359–1363

# Eidesstattliche Erklärung

Ich versichere, die von mir vorgelegte Arbeit selbständig verfasst zu haben.

Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten oder nicht veröffentlichten Arbeiten anderer entnommen sind, habe ich als entnommen kenntlich gemacht. Sämtliche Quellen und Hilfsmittel, die ich für die Arbeit benutzt habe, sind angegeben.

Die Arbeit hat mit gleichem Inhalt bzw. in wesentlichen Teilen noch keiner anderen Prüfungsbehörde vorgelegen.

Gummersbach, xx. August 2016

Max Mustermann