Lehrstuhl für Steuerungs- und Regelungstechnik
Technische Universität München
Univ.-Prof. Dr.-Ing./Univ. Tokio Martin Buss

Data acquisition and motion interpolation for a multimodal medical training environment

Dipl.-Inform. (Univ.) Thomas Pröll

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. sc.techn. (ETH) Andreas Herkersdorf

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing./Univ. Tokio Martin Buss

2. Univ.-Prof. Dr. rer. nat. Rüdiger Westermann

Die Dissertation wurde am 20.04.2006 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Infromationstechnik am 08.08.2006 angenommen

Data acquisition and motion interpolation for a multimodal medical training environment

Modern, medical image generating methods like CT or MRI provide an insight into the soft tissue structures of the knee joint. However, these techniques are mostly used for acquiring still pictures, because they are either not fast enough or not precise enough to guarantee a profound, three-dimensional live observation of deformable human internals. Neither the functionality, nor the shape of the components during fast movements can be acquired that way. The same drawbacks occur when biological structures are dissected.

In recent medical literature, the shape, texture and function of almost every part of the human body is described. As these descriptions are either based on dissected material or on those image generating methods, the description of the functionality and interaction, as well as the description of the deformation of the shape during normal operation are mostly qualitative assumptions of the authors, but not quantified or measured facts.

In this work, a quantitative measurement is performed for a more profound data acquisition. Opposed to the known, slice wise measurements, a three-dimensional technique is applied that allows new insights into the movement and deformation of the knee joint components. Moreover, most publications consider only one rotational freedom of the knee joint - in most cases the flexion movement. In this work, all three rotational degrees of freedom are considered, providing new insights into the knee joint mechanics. The results of these measurements extend the current knowledge about the human knee joint, which proves the significance of three-dimensional measurements.

This measured data is merged to a three-dimensional knee joint model that can display the knee joint internals in any arbitrary posture. This is used to set up an interactive graphical display that allows the user to move a virtual knee joint and observe the movements and deformations of the knee joint internals in real time. This animation can be helpful for improving the understanding of the internal knee joint mechanics and can improve the medical education.

Datenakquisition und Bewegungsinterpolation für eine multimodale medizinische Trainingsumgebung

Moderne, bildgebende Verfahren wie das CT oder MRT ermöglichen einen Einblick in das Innere von biologischen Strukturen, wie etwa dem menschlichen Kniegelenk. Sie werden überwiegend dafür verwendet, Einzelaufnahmen zu erstellen, da sie entweder nicht schnell genug oder nicht präzise genug sind, um eine fundierte, drei dimensionale Messung von verformbaren, menschlichen Komponenten in Echtzeit zu erstellen. Damit kann aber weder die Funktion eines Körperteils, noch die Veränderung seiner Form während einer schnellen Bewegung erfasst werden. Auch Sektionen eignen sich für eine solche Datenakquisition kaum, da beim Freilegen der zu untersuchenden Strukturen die mechanischen Eigenschaften des Gelenkes verändert werden.

Trotz dieser Schwächen in der Datenakquisition werden die Komponenten des menschlichen Körpers in heutiger Literatur in ihrer Form, Struktur und Funktion beschrieben. Da diese Beschreibungen auf Material aus Sektionen oder auf Material von bildgebenden Verfahren basieren, sind die Beschreibungen der Funktionalität oder der Deformation qualitative Betrachtungen der Autoren, jedoch keine quantifizierten oder gemessenen Fakten.

In dieser Arbeit wird eine quantifizierende Messung am menschlichen Kniegelenk durchgeführt. Im Gegensatz zu den bekannten, schichtweisen Messungen wird hierbei eine dreidimensionale Technik verwandt, die neue Einblicke in die Bewegung und Deformation der Komponenten des Kniegelenks erlaubt. Während die meisten Publikationen wenige Komponenten des Knies in nur einem rotatorischen Freiheitsgrad untersuchen, werden hier alle erkennbaren Komponenten in allen drei rotatorischen Freiheitsgraden untersucht.

Die gemessenen Daten werden in einem dreidimensionales Kniemodell zusammengefasst, das damit das Kniegelenk in jeder möglichen Stellung grafisch dargestellt werden kann. Mit diesem Modell wurde ein interaktives, grafisches Display aufgebaut, das es dem Benutzer erlaubt, ein virtuelles Kniegelenk zu bewegen und die resultierenden Bewegungen und Deformationen im Knie in Echtzeit zu beobachten. Dieses Display kann Laien und medizinischen Novizen dabei helfen, das Kniegelenk besser zu verstehen und kann damit einen positiven Einfluss auf die medizinische Ausbildung haben.

Contents

Notations

Abbreviations

General abbreviations

ACL	Anterior Cruciate Ligament
AR	Augmented Reality
CT	Computer Tomography
DoF	Degrees of Freedom
FEM	Finite Element Method
fps	frames per second
GHz	Gigahertz
HU	Hounsfield Units
LCL	Lateral Collateral Ligament
MCL	Medial Collateral Ligament
MRI	Magnetic Resonance Imaging
PCL	Posterior Cruciate Ligament
PDE	Partial Differential Equation
RBF	Radial Basis Function
RF	Radio Frequency
SLERP	Spherical Linear intERPolation
SQUAD	Sperical QUADrangle
VR	Virtual Reality

Physical Units

MB	Megabytes = 1024 Bytes
mS	Millisievert
N	Newton
Nm	Newtonmeter
nT	Nanotesla

Medical Directions

ant.	anterior
ext.	external
int.	internal
post.	posterior
rot.	rotation
trans.	translation
val.	valgus
var.	varus

Conventions

Scalars, Vectors, and Matrices

Scalars are denoted by upper and lower case letters in italic type. *Vectors* are denoted by lower case letters in boldface type, as the vector $\boldsymbol{x}$ is composed of elements x_i. *Matrices* are denoted by upper case letters in boldface type, as the matrix $\boldsymbol{M}$ is composed of elements M_{ij} (i-th row, j-th column).

x	scalar
$\boldsymbol{x}$	vector
$\boldsymbol{X}$	matrix
$f(\cdot)$	scalar function
$f'(\cdot)$	derivate of $f(\cdot)$
$\boldsymbol{f}(\cdot)$	vector function
$\boldsymbol{X}^T$	transposed of matrix $\boldsymbol{X}$
$\boldsymbol{X}^{-1}$	inverse of matrix $\boldsymbol{X}$

Symbols

$\boldsymbol{q}$	A quaternion
v	The interpolated value
t	Control variable for SLERP
h	Control variable for SQUAD
$\boldsymbol{s}$	Quaternion for SQUAD calculation
m	Weight for interpolations
l	Exponent for the weighted averages
(x, y)	The cartesian coordiantes in 2D
(x, y, z)	The cartesian coordinates in 3D
(w, x, y, z)	The four values of a quaternion
i, j, k	The three imaginary units of a quaternion
$\boldsymbol{p}$	Three DoF for the quaternion
θ	One DoF for the quaternion
ω	The angle between two quaternions and the center of the sphere
Φ	A radial basis function (RBF)

List of Figures

List of Tables

1 Introduction

In modern, western countries the average life span increases since decades. This poses new problems to the health care system, since elderly people suffer different illnesses and pathologies than younger people, for example diabetes, weakened cardiovascular system, decreased eyesight and hearing as well as defective knee and hip joints.

The other challenge for the national health care systems are new trend sports. More and more people are performing high velocity sports like inline skating, snowboarding or mountain biking. While this trend is supposed to reduce the weaknesses of the cardiovascular system, other injuries like fractures and joint injuries occur more often.

The third change that has to be taken into account is the growing cost of the health care. Diabetes medicine, pace makers, hearing aids, knee- and hip prosthesis are technical achievements of the last few decades, increasing the quality of life for those who require them, but also increasing the cost. Until today it is possible to provide modern medical treatment to everyone who requires it, but this will not be possible in the future, if the newer treatments are getting more and more expensive.

Thus, a further challenge to the medical system will be the cost reduction without reducing the quality of the treatment. For reducing those cost, it has to be determined how much money is spent on which occasion. Furthermore, one has to estimate whether research in those pathologies or injuries is likely to save money. This is done in section 1.1.

The author concludes that research about joints and research about education is likely to save money. As a consequence, the Knee Joint Simulator that is described in section 1.3 was developed. The graphical display that is developed in this thesis is a part of this simulator. Accordingly, the state of the art of the medical graphical animations is shown in section 1.4. In sections 1.5 and 1.6, the challenges from the medical point of view and the challenges from the technical point of view are described. The last three sections explain, why the current approaches do not apply to the problem of the Knee Joint Simulator. As a contrast to the state of the art approaches, the approach of this work is explained in section 1.7. The outline of the thesis, that explains this approach in detail is given in section 1.8.

1.1 Costs in Medicine

The factor “cost” in the medical sector is threefold:

- The *direct cost* is the amount of money that has to be spent to cure the patient. This contains the expenses for diagnosis, therapy and rehabilitation.

- The *indirect cost*, which includes all expenses that are related to the illness or pathology, but which are not necessary for the cure. Examples for indirect cost are loss of salary, longer retirement times or a reduced gross national product.

- The psychosocial factor, is also called *intangible cost*. It is a burden for the affected people and includes pain, fear and reduced quality of life as results of the illness or pathology.

In order to reduce the total cost, all three factors have to be taken into account. The direct cost can be calculated quite easy by adding the expenses of the sick-funds.

In literature, there are different approaches to estimate the indirect cost of the different pathologies by quantifying the loss of resources as a result of disability, unemployability and premature death. The results of these approaches show huge differences, due to the different evaluation standards. Thus, only *objective* losses, like lost years of productivity or lost years of life should be considered here.

The intangible cost can not be expressed as a monetary amount at all, since they can not be measured. Fear, pains or frustration influence the life quality of the patients, their relatives and friends. This is certainly a negative aspect of illnesses and could be called a "cost", but it can not be expressed in a currency. Only the absolute number of individual cases can give an impression of the national impact of a certain disease.

In 2002, an analysis of the national health care system in Germany showed, that the most *direct cost* is spent for the group of cardiovascular diseases (35.4 billion Euros), followed by alimentary diseases (31.1 billion Euros), and by musculo-skeletal illnesses (25.2 billion Euros) [2].

The major part of *indirect cost* is spent for the large group of injuries and poisoning; for this group 1.3 million years of productivity and 2.7 million years of life time were lost in 2002. 794,000 years of productivity and 1.7 million years of life time are lost by illnesses of the musculo-skeletal system. Number three in this list are psychiatric illnesses, which caused 737,000 lost years of productivity and 1.9 million years of life time.

To quantify the intangible cost, the *number of cases* is estimated for Germany. In 2002, the largest number of cases occurred in the group of external forces, i.e. injuries, poisoning, burning and intrusion of objects (17,051,165 cases). The second largest group were the cardiovascular illnesses (2,803,282 cases), while the third largest group with 1,360,633 cases were problems with the musculo-skeletal system [3].

Under all three aspects, musculo-skeletal illnesses are playing an important role:

- Thirdmost cause of direct cost
- Thirdmost number of patients
- Thirdmost cause of decreased life time
- Secondmost cause of decreased productivity

Additional to the significance of illnesses, musculo-skeletal *injuries* are also very common [4] [5], so the focus of this work is set on this. As the musculo-skeletal system is too complex to be analyzed in a single thesis, one manageable part is chosen. Moreover, methods are used that can be transferred to other body parts.

83% of the musculo-skeletal illnesses are illnesses of the spine. Pains in the back are the main reason for huge direct and indirect cost. Due to the complexity of the spinal system, 10 different spinal illnesses were classified in the report of the "Statistisches Bundesamt" in Germany [3]. Although a detailed research of this part of the musculo-skeletal illnesses is necessary, it is still too complex to be treated in one single project.

The knee joint is the part of the musculo-skeletal system that is affected most often by osteoarthritis; from all 338,198 cases of osteoarthritis that were diagnosed 2002 in Germany, 168,705 cases were located in the knee joint, followed by 139,309 cases of hip osteoarthritis.

The knee joint is not only affected by osteoarthritis most often, but also by luxations. Luxations of the knee joint were diagnosed 72,670 times in 2002, while the luxations of the shoulder on the second place were diagnosed only 24,688 times. In addition to that, the knee joint is exposed to further threads that do not have a detailed medical ICD-10 coding (like arthritis), so quantification fails here [3].

The fact that the knee joint is (one of) the most endangered joints was also found by Heim et al. [4], Steinbrück [5] and Riener et al. [6]. It is very likely that the health care system can benefit from any improvement of the diagnosis, therapy or rehabilitation of knee joint illnesses and injuries. However, the knowledge about this joint is still very limited; the deformation characteristics of soft tissue knee joint components, and even the relative bone movements are still a discussion topic (see 2.7.2). This medical relevance together with the economic relevance qualify the knee joint as the center of this work.

This work provides new methods of researching the biomechanical properties of knee joints, quantifying them and make them suitable for simulation. This simulation can improve the understanding of the measured biomechanics, and is supposed to improve the medical education.

Certainly, these findings can be generalized (to a certain extend) to other joints, like the shoulder, elbow, ankle and even much smaller ones like the fingers. However, the future target is the research of the most complicated and the most vulnerable joint complex - the spine.

1.2 Knee joint diagnostics as a challenge for medical education

The focus of this work is the quality improvement of economic diagnostic methods of knee joint injuries. There are three possibilities of knee joint diagnostics:

- Image generating methods (x-ray, CT, MRI)
- Diagnostic arthroscopy
- Functional knee joint tests

CT and X-ray scans are exposing the patient to dangerous radiation and should be avoided if possible. The average inhabitant of Germany is exposed to 2 Millisievert (mS) from 1.5 x-ray scans per year. More x-ray scans are only performed in Japan, what suggests that some of those scans can be replaced by other methods.

This theory is supported by the report "Bedarfsgerechtigkeit und Wirtschaftlichkeit" [7], where the experts found that complicated manual diagnostic and therapeutic measures (like chiropractics) are often underprovided, while automatic, image generating measures are often overprovided. In agreement with that report, Regulla [8] found that too much x-ray scans are performed in Germany.

Replacing those hazardous methods with MRI scans can reduce the risk for the patients, but is not a promising approach for cost reduction as it is even more expensive. Thus the two alternatives (diagnostic arthroscopy and functional knee joint tests) are discussed here.

1.2.1 Arthroscopy

The term *arthroscopy* has its origin in the Greek language, where "arthro" means "joint" and "scopie" means "to look at". Looking at or into joints can be done in several ways, however, "arthroscopy" is generally done by "keyhole surgery", which causes smaller wounds and shorter recovery times than conventional surgery methods.

Therefore, a needle is inserted into the knee joint, and liquid is pumped into the joint space. The pressure of the liquid widens the knee joint in all three dimensions. After that, a small camera, a light source and instruments can be inserted through small holes. The pressure of the liquid ensures a wider working space, so that a visualization and manipulation of the knee joint internals are possible (see figure 1.1).

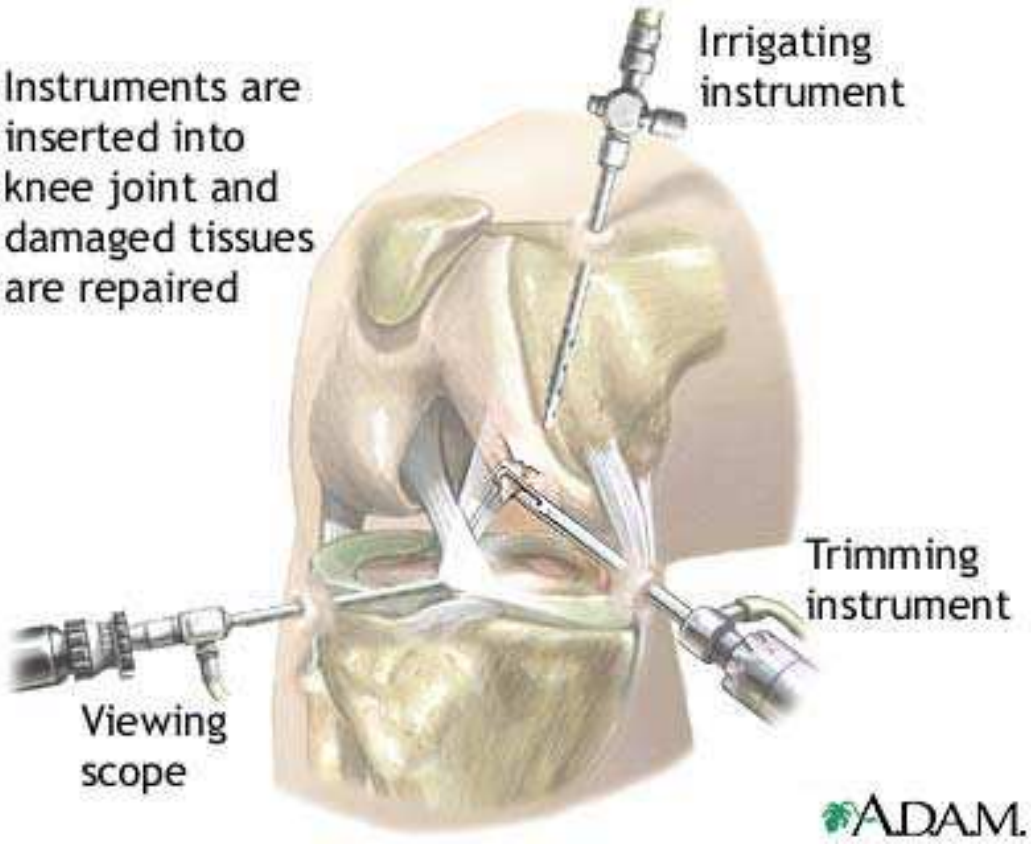


Figure 1.1: Animation of a keyhole surgery from "ADAM"

According to Bitzler [9], this method provides a more detailed insight into the human knee than any other kind of diagnosis. Small pathologies like the partly ruptured ligaments or menisci can be observed directly, while they might be invisible on the CT or MRI scans due to the limited resolution.

Lawson [10] and Noble [11] agree that these advantages lead to a (too) broad usage of this technique. Bitzler [9] showed that 22.5% of all knee joint surgeries are performed for diagnostic reasons only, what was also found by Kleinmann [12], who found that percentage "stunningly high" ("erstaunlich hoch").

If the diagnostic arthroscopy would be a cheap and harmless method, this high grade of utilization would be appreciated. However, there are severe disadvantages, under the two aspects of cost and risk. The direct cost of one diagnostic arthroscopy, together with

the loss of productivity is around 1,000 Euro, while a MRI scan costs around 200 Euro. Moreover, diagnostic arthroscopy bears the same risks as a common surgery does, for example infections, inflammations and anesthesia risks.

1.2.2 Functional knee joint tests

As the diagnostic arthroscopy is not cheaper than image generating methods, and as it bears further risks, the functional knee joint tests have to be examined. A medical doctor uses a functional knee joint tests by manually applying forces to the patient's lower leg in a specified way. Due to the applied forces, the patient's knee joint is moved in a characteristic manner.

As the internal knee joint mechanics are changed by most pathologies, the responding movement of an injured knee joint is different to that of a physiological knee joint. If, for example, a ligament is torn, it can neither transmit any forces any more, nor can it contribute to the knee joint stability. Thus, torn ligaments can be recognized by immoderate flexibility in certain directions. Other pathologies can be recognized by snapping or scratching movements and sounds, or simply by the pain that is expressed by the patient. For further details see appendix A and appendix B.

The difference between a physiological knee joint and a pathological knee joint is subtle in many cases. A further difficulty for the correct diagnosis is the fact that the difference between two *physiological* knee joints of two different humans is larger than the difference between one physiological and one pathological knee joint of the same person. Thus, the medical doctor needs much experience and has to compare the flexibility of a patient's injured knee joint with his healthy one.

In opposite to the increased flexibility, some meniscus pathologies can be identified by different signs. Pain that is expressed by the patient or snapping movements or sounds from inside the knee joint that occur when the lower leg is passively moved along a certain trajectory by the physician are examples for such signs.

As these trajectories are sometimes very complicated (for example for the McMurray test), unexperienced physicians might miss the correct trajectory and get no adequate feedback, although the pathology exists. As with the ligament pathologies, the physician needs much experience for a definite diagnosis.

The low-tech-approach of functional knee joint examination looks old-fashioned and prone to errors. However, Gurtler et al. [13], Rubinstein et al. [14] and many more agree that the functional knee joint tests can be used for a very reliable diagnosis, if they are performed in a correct and careful manner. As these tests can be performed very quickly and as no special equipment is needed, the cost is significantly lower than the cost of any other knee joint diagnosis method.

The only drawback of the functional knee joint tests is that the medical doctors have to be trained very well to ensure a high quality of the diagnosis. This can only be assured by a profound education, so the focus is set on the theoretical and practical education of medical doctors.

1.2.3 Theoretical education

The word "theory" derives from the Greek "theorein", which means "to look at". The theoretical education is a knowledge transfer, where the student is mostly a passive receiver

of information, e.g. the student is *listening to* or *looking at* an information source. This information source can be a teacher, but also a book or a video tape.

The theoretical education in the medical sector should provide scientific backgrounds, instructions and principles to the students. The theories that are provided by this education influence the practical work of the students later.

There were always (and there still are) errors in the theoretical education. In the middle ages, the theoretical knowledge of human anatomy was plain wrong in many cases, which led to obscure practical treatments like bloodletting or blood transfusions between animals and humans.

Due to better research, such basic flaws have been eliminated. However, thirty years ago, the scientific society still agreed that the menisci are unnecessary relicts of evolution, just like the appendix. As a consequence, they were removed completely, whenever a pathology occurred. Just as most of these patients suffered osteoarthritis several years after the operations, the load distribution capabilities of the menisci were discussed.

Until today, the knee joint is not really understood. The function of the patella (knee cap), the cruciate ligaments and even the whole movement of the knee joint are still topic of scientific discussions. The fact that artificial replacements show severe problems in mobility and are not very durable might be just one consequence of this weak understanding.

As the theoretical knowledge of the intact knee joint is not really profound until today, the knowledge of the pathologies and the damage mechanisms is even worse. Thus, the theoretical education can not always explain everything with a satisfying precision, and in some cases, the explanations might even be plain wrong.

Another weakness of the theoretical education is the presentation of the information. In most books, the movement and deformation of the menisci is depicted by two drawings in full extension and full flexion. Students can imagine how the menisci behave between those two flexion angles, but it is hardly possible to estimate how they behave in such a complicated, multi-dimensional movement like the McMurray test.

One common approach to improve the theoretical education are multimedia or VR applications, which are known to improve the understanding of complicated processes [15]. However, as it is already mentioned above, the knowledge of the knee joint kinematics is not very profound, as they have some serious limitations like the consideration of only one degree of freedom (DoF), the negligence of soft tissue deformations, low precision measurement and many more. Multimedia applications based on such flaws have only a limited ability to improve the theoretical education.

In this work, a precise, multidimensional investigation of soft and hard tissue knee components is performed and an adequate multimedia application that is suitable for education purposes is developed on that basis.

1.2.4 Practical education

The word “practice” is derived from the Greek “prattein”, which means “to act”. In opposite to the theoretical education, the student is the main actor of the scene and is interacting with the object of examination. The teacher has a more passive part by supervising the student or by giving hints.

The practical education should prepare the student for the work as a physician. Generally, this is done with artificial plastic phantoms, cadavers, animals, probationers or real

patients. Plastic phantoms are the cheapest, but most unrealistic training possibility, cadavers and animals are more realistic, but harder to get, thus more expensive. The training with probationers and patients is very realistic, but rare. Errors are very dangerous, so only very experienced students get the possibility to train with them.

For the education of functional knee joint tests, plastic phantoms are available. One example from "Wisdomking" [16] can be seen in figure 1.2. Ligaments and menisci are represented by rubber bands, where the bones are implemented as hard plastic objects. As rubber has a different elasticity as the soft tissue, and as the friction between the objects is different, this plastic model hardly behaves like a knee joint. Thus, these passive phantoms can only give a general idea of the knee joint mechanics.

Figure 1.2: A plastic model of a knee joint from "Wisdomking"

Training with cadaver knee joints is more realistic for simulating healthy knee joints, since elasticity, friction and other physical parameters are realistic. Pathologies like torn ligaments or menisci can be produced easily by cutting the structures with a scalpel. However, as this is done, other knee joint components like the capsule might also get damaged, what changes the knee joint behavior. Thus, not all pathologies can be simulated in a realistic manner. Moreover, it is possible to feel the different knee joint behaviors, but it is not possible to estimate if a patient would express pain whenever a certain knee joint test is performed.

After the first practical experiences are made, every medical student can try to perform knee joint tests on healthy participants of the same course. The teacher can only demonstrate the test and ensure the correct carriage and the correct trajectories of the movements, when the students try it (see figure 1.3). The judgment, whether the move-

ments are performed with the correct amount of force can hardly be given by the teacher and is up to the two unexperienced probationers.

Figure 1.3: An experienced teacher treats a healthy student

At the end of the practical education, when the students acquired enough experience with healthy knee joints, they are allowed to assist an experienced physician by the treatment of real patients.

The education on the patient is the most realistic training a novice can get. Unfortunately, this kind of training is scarce:

- Most novices get their training in hospitals. However, many people who get a knee joint injury in everyday life will visit their family doctor or a local orthopedic, so they are not available for education in hospitals.

- The education is scheduled. When the knee joint is the current topic of the anatomical course, it is very unlikely that many patients with many different knee joint injuries show up in the hospital. Thus, there will be a certain time between the theoretical education and the practical training on the patient.

- Functional knee joint tests of injured knee joints are causing pain in most cases. Thus, the patient will not allow all interested students of a course to touch the aching knee.

Thus, it takes years until a novice has collected enough experience to perform knee joint tests without the supervision of a more experienced physician. One consequence of this is the fact that the education is either very expensive or not profound enough, what could explain the immense use of image generating methods.

To overcome the problems of unavailable patients, scarce pathologies, the "Knee Joint Simulator" is developed. This multimodal approach should provide a training possibility that is always available, that can simulate many different pathologies and that gives a realistic force feedback to the novices.

It is certainly very unlikely that the training on the real patient can be completely replaced by any kind of simulator, however the students might be better prepared before touching the first human patient.

1.3 Project context: The Knee Joint Simulator

The basic problem of the practical education is to transfer the teacher's subjective impressions of movements and forces to the novices. The snapping of a meniscus pathology can neither be simulated by passive phantoms, nor can it be taught by explaining "how it feels". The novice *has to feel* it himself.

As passive phantoms fail in giving a realistic impression, Virtual Reality (VR) seems to be a promising approach to implement a more realistic behavior. Many parameters of a virtual knee joint can be influenced and adjusted by the programmer, so it seems to be easy to simulate individual healthy and pathological knee joints.

However, if it is about grasping and touching objects, the *immersion*[1] of pure VR-approaches is rather low until today: The current VR-hardware for displaying *kinesthetic*[2] impressions like the CyberGrasp glove [17] is able to simulate small forces on the fingertips, but it is not possible to give a force impression to other parts of the hand (e.g. the palm), which does not provide a satisfactory impression.

However, the *tactile*[3] feedback is even worse. Simulating the fine texture of the human skin with hairs, wrinkles and all other irregularities is not possible at all. Thus, simulating a human knee joint with *haptic*[4] VR displays provides even a less realistic impression than passive plastic phantoms do.

Thus, a mechatronic approach is chosen for the Knee Joint Simulator that combines the advantages of VR systems and passive phantoms. The user interface is a passive shank with a skin-like surface to provide a realistic tactile impression. The embedded artificial bone structures enable the user to locate important landmarks like the tibia plateau, the ankle or the shin-bone by *palpation*[5]. Thus, this artificial shank provides better tactile impression than most available passive phantoms, and certainly a much better tactile impression than all pure VR approaches.

The same techniques are used to produce an adequate thigh. Together with a complete mannequin lying on a couch, these passive phantoms provide a realistic medical training environment (figure 1.4).

[1] A simulation that offers complete immersion can not be distinguished from the reality by the user

[2] Kinesthetic sensations are felt by movements or strains in muscles, joints or tendons, for example by lifting weights

[3] Tactile sensations are felt by pressure receptors of the skin, for example to distinguish rough surfaces from smooth ones

[4] Haptic feedback is general force feedback and inhibits kinesthetic and tactile feedback. Some publications do not distinguish between haptic and kinesthetic impressions

[5] Palpating a body part means examining by touching or pressing

In opposite to the known, passive phantoms, the knee joint characteristic (movement of the shank relative to the thigh) is implemented with an active, mechatronic system. An industrial robot Stäubli RX90 [18] is used and controlled by a Personal Computer (PC).

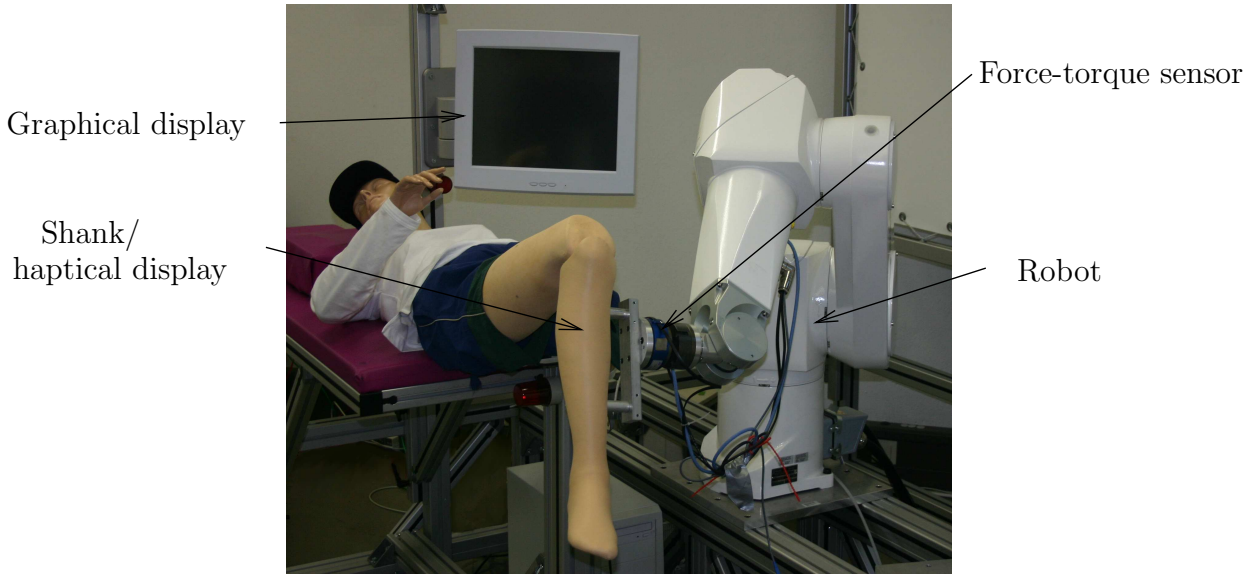


Figure 1.4: This knee joint simulator should provide a realistic training environment

Two different modes that are called the "active mode" and the "passive mode" are implemented for the knee joint simulator control. In the first, "active mode", the robot moves the shank along the trajectories of the different functional knee joint tests. Thus, the user is guided by the robot and gets a first impression of how a certain test should look and feel like.

For the second, the "passive mode", a real knee joint is simulated that reacts in a realistic manner whenever the user applies forces to it. Therefore, a 6-DoF force-torque-sensor has to be mounted between the robot and the shank. Initially, the robot does not move at all, but holds the shank in a neutral starting position. If the user touches the leg and applies forces to the shank, these forces are measured by the force-torque-sensor. Those values are sent to a biomechanical model, which is used to estimate how a real knee would react under this load. Afterwards, the robot is used to move the shank into the calculated position (figure 1.5). If this mechanism works without noticeable delay (around 1000 position adjustments per second), the user gets the impression that the movements of the lower leg are caused directly by the applied force. More detailed information about the haptic display can be found in [19].

Additional to that haptic display, other modalities are implemented for improving the immersion. Loudspeakers are mounted in the thigh and in the mannequin's pillow, so pathological sounds from the knee joint can be simulated as well as the patient uttering pain. The decision, whether the patient utters pain or not in a certain situation is taken by a neuronal net that is trained with unsharp statements taken from the literature ("severe pain", "slightly painful"). The acoustic display is explained in [20] and [21].

The third modality is the graphical display, which has also two functionalities. The first one is combined with the "active mode" of the haptic display. When the robot is displaying the trajectory of a functional knee joint test, the user can watch a video of an

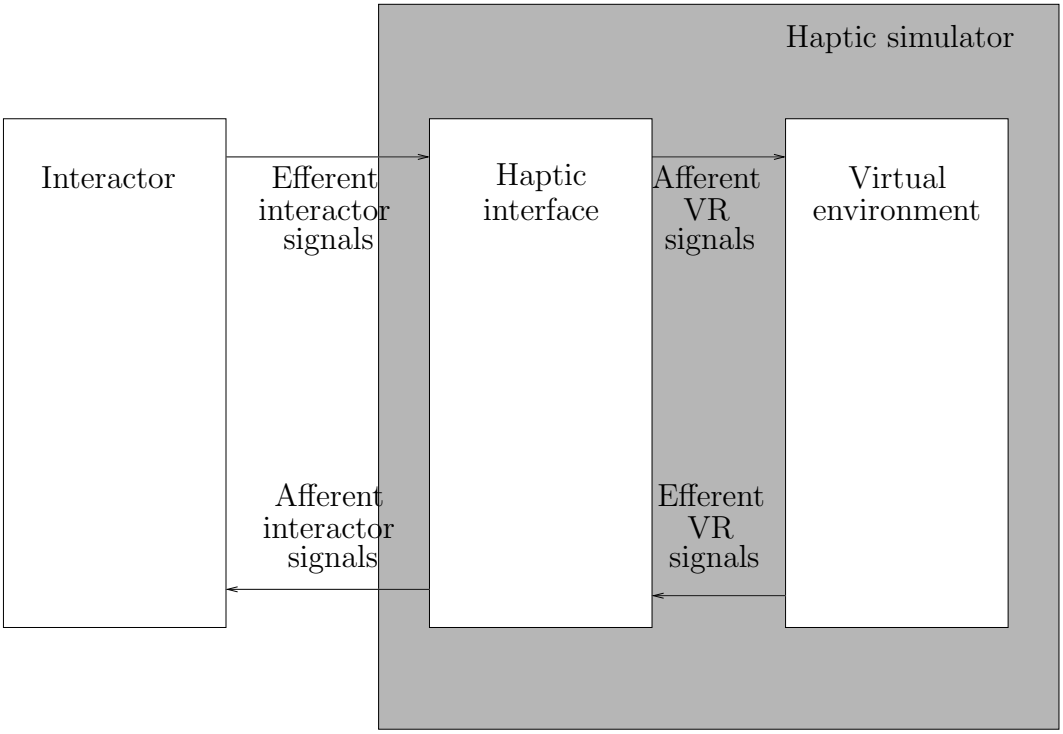


Figure 1.5: Interaction between the interactor, haptic interface and virtual environment

experienced physician who performs exactly the same movements with a patent's leg. The second mode of the graphical display is related to the "passive mode" and should enable the user to look inside the virtual knee and watch the behavior of the internal structures. This interactive mode of the graphical display is the focus of this thesis.

1.4 State of the art

Multimedial software approaches are very common to describe and depict human anatomy, surgeries and treatments. These software packages are developed with state-of-the-art software tools that are also used in the entertainment sector. To allow a correct classification of this work, a variety of available medical animations are described here.

Medical multimedia software packages can be classified into dynamic approaches that allow a manipulation of the knee joint, and in static approaches that do not. Moreover, two dimensional approaches have to be distinguished from three dimensional approaches, as the third dimension permits much more interaction.

Another focus of this chapter will be the state of the art in the scientific field. These approaches are not available as software packages, but may or may not be used to implement dynamic, three dimensional multimedia software - like the graphical display of the Knee Joint Simulator. Thus, an overview of these techniques is given.

1.4.1 Available medical animations

Plenty of software packages deal with all kinds of human anatomy. Some medical atlases describe the whole human body, others are limited on one single body part. This overview

will treat only software that have to do with the knee joint, but as the techniques are used for other parts of the body in a similar way, it can also be considered as a general overview.

The software package "Bodyworks 6.0" from Mindscape [22] is a multimedia anatomy atlas that describes the whole human body with images, animations, 3D objects, text and voice information. Unfortunately, the closeup view of the knee joint can only be seen from the front and the back (see figure 1.6).

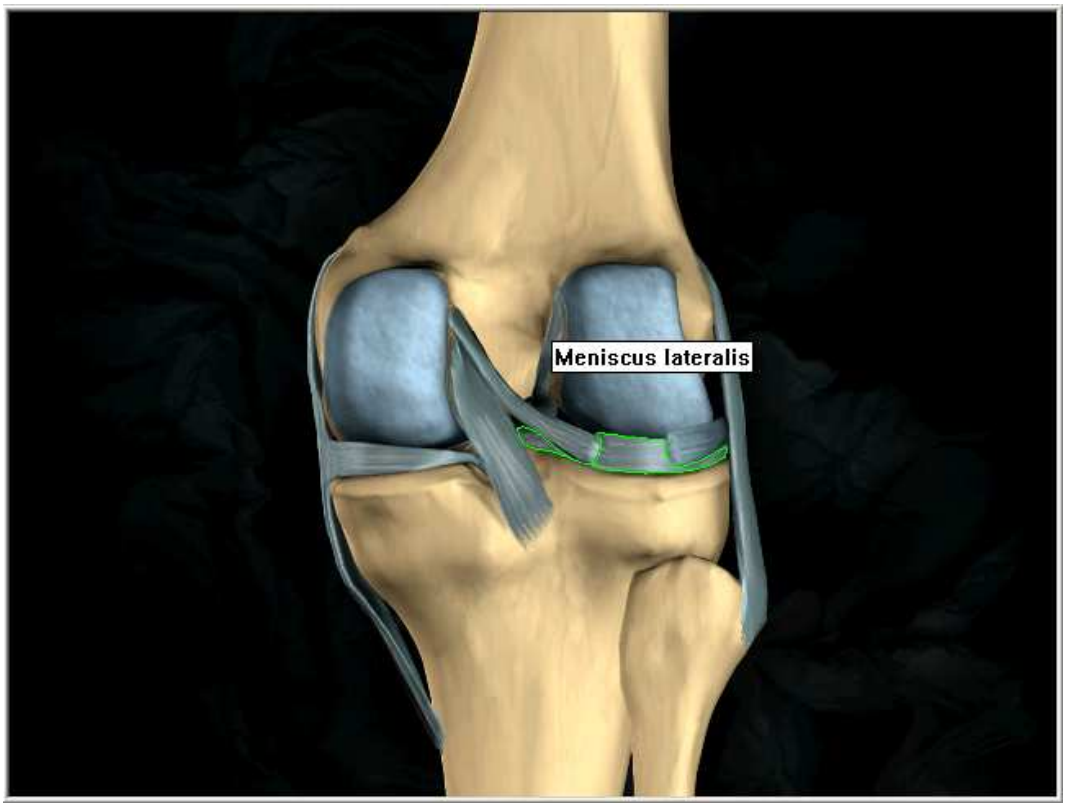


Figure 1.6: The anatomical atlas from Mindscape shows and explains the knee joint

Pointing the mouse on certain knee joint components provides the user with informations about this structure. Additionally, the user can zoom in, but the graphical resolution is very limited. 3D interaction like the free choice of perspective is not possible, which proves the software to be two dimensional only, in spite of its three dimensional appearance. Changing the knee joint flexion or influencing other knee joint parameters in also not possible, which proves that this software is a *static* approach.

A very similar approach is available from "Swarm Interactive Inc." [23] (see figure 1.7). Again, a right knee joint can be seen from the front and from behind. This knee joint is hand drawn, but shows all major soft tissue components as the menisci, and the six largest ligaments.

Additional informations are presented, if the mouse cursor is moved over the different knee joint components; this increased interactivity significantly improves the descriptive nature of this tool.

Nevertheless, zooming or rotating the knee joint is not possible, the precise shape and location of the components remains unknown, as they can only be seen partly. Thus, this is again a *static* 2D approach.

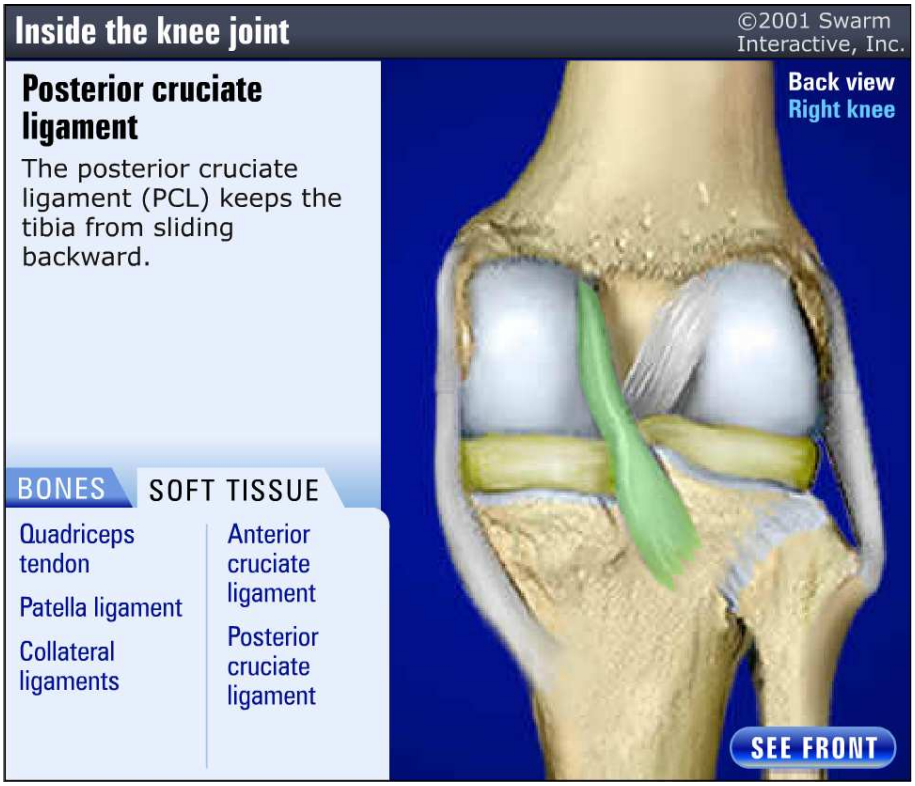


Figure 1.7: The interactive anatomy atlas of "Swarm Interactive Inc.".

The company "Nucleus medical art" [24] produces more sophisticated animations of knee joint surgeries which show deformations. In the animation that can be seen in figure 1.8, the ruptured part of a meniscus are removed with keyhole surgery.

Figure 1.8: The cutting animation of "Nucleus medical art"

The animation consists of only around five different images per second. For a fluent animation, several intermediate images are missing between two subsequent images, thus a blending or morphing algorithm is used for the transition in between. The movement of the surgical instruments could be simulated that way.

The meniscus was drawn in two different versions: Once with a torn region present and once with the torn region removed. Blending between those two versions could be

considered as a "meniscus deformation". As the clip can be played forth and back, the user is able to manipulate the knee joint in a one dimensional manner. Thus, this can be considered as a (very simple) *dynamic* simulation.

Like the anatomy atlases that are mentioned above, this clip also gives a three dimensional graphical impression. However it is a pure two dimensional animation as the user can not influence the perspective, lightning or other 3D aspects.

The "Montana Spine Center" [25] shows different animations on their "Patient Information Resource" web page. These clips should provide anatomical knowledge about different joints (hand, hip, knee and many more). The different components of these joints, the possible pathologies and the different therapies are explained.

These animations are hand drawn (as it is very usual in the medical sector), and comparable to the animations provided by "Nucleus medical art", as it can be seen in image 1.9. In these figures, a torn meniscus is sewed, and the meniscus deformation is - again - implemented by blending between two subsequent images. Like the cutting animation, the sewing animation is a simple *dynamic* simulation.

Figure 1.9: Sewing the meniscus as it is explained in the video clip of the "Montana Spine Center"

Those two clips were 2D only, slightly dynamic but hardly interactive; the animation could be played and paused. The speed of the animation and the zoom factor could be influenced with other tools only. However, this technique is able to depict cutting, sewing, moving and has the capability to show complete surgical operations. This is not possible with the static approach of the anatomy atlases.

The first three dimensional approach that is considered here is the tool "Interactive Knee - Surgery" from "Human Anatomy Inc." [26]. It is part of a multimedia anatomy atlas that can display three dimensional models of organs and other body parts (see figure 1.10).

Its overwhelming feature is the completeness of all components. Not only bones, menisci and ligaments can be displayed, but also muscles, fat, blood vessels, cartilage and other tissue of the human body. For improving the clarity, the user can choose which tissue should or should not be displayed. Moreover, the user can rotate and translate the whole scene and zoom into details. However, no joints can be moved or no tissue can be deformed. The possibility of showing or hiding certain materials is a feature of the three dimensional approach and not a manipulation of the knee joint. Thus, this is a *static* 3D approach.

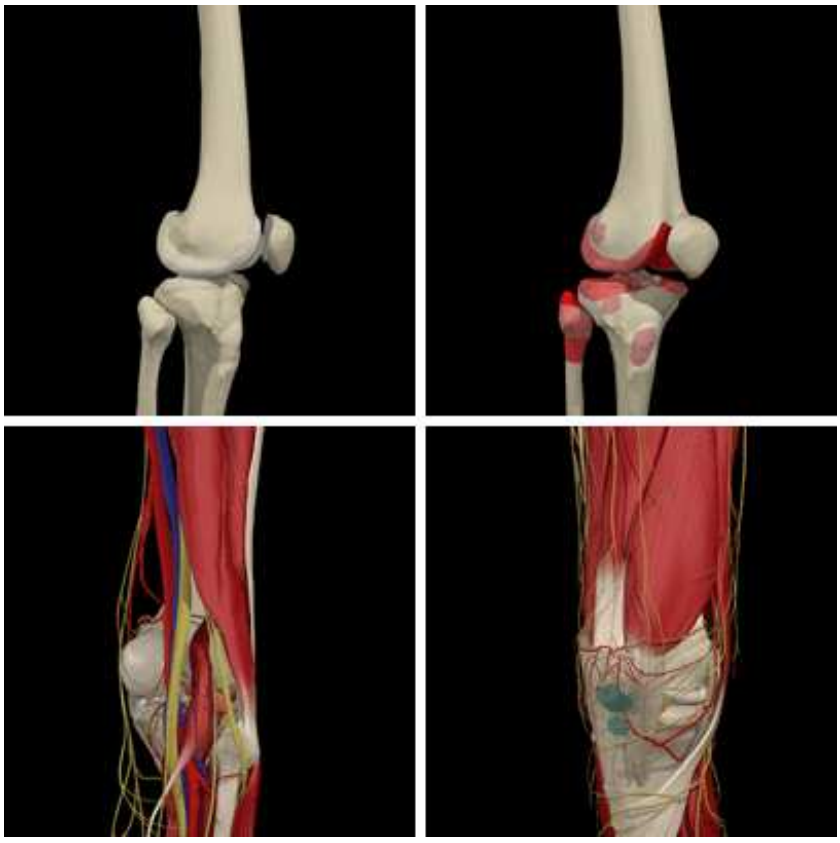


Figure 1.10: Images taken from the three dimensional anatomy atlas of “Human Anatomy Inc.”. The number of displayed knee joint components can be influenced by the user.

Other anatomy atlases like "The Visible Human Project" [27] or many other projects that are based on the same data set like the products of the "Center for Human Simulation" [28] share the same drawback of being static. Thus, medical multimedia products are available in 2D and in 3D, but they are hardly interactive and hardly dynamic.

1.4.2 More dynamic approaches in science

In commercial products, the user had only very limited possibilities to manipulate the knee joint. Cutting and sewing was only one dimensional, i.e. the operation could be performed or not. The user had no choice, where to cut or where to place the needle.

If this technique would be used for the Knee Joint Simulator, the user would only have the choice if the knee joint should be flexed or not. Arbitrary flexion angles and rotational axis can not be implemented with that. As this is not satisfactory, more sophisticated methods have to be evaluated.

The challenge of the dynamic manipulation is not the visualization, but consistency of the locations of the knee joint components. If the position or orientation of the lower leg is changed, the other structures will change their positions and orientations, too[6]. Thus, a system is needed that describes the movement of each knee joint component, depending of the angle between thigh and lower leg.

There are four different approaches that can be used for this task:

- Finite Element Method (FEM)
- Mass-spring-damper-systems
- Particle Systems
- Keyframe interpolation

The Finite Element Method (FEM)

The Finite Element Method is a numerical method to solve partial differential equations (PDE) approximately. In engineering sciences, it is used for calculating electromagnetic scenarios, heat transportation scenarios or scenarios of force transmission. Simulating the force transmission in the knee joint can solve the problem of how the knee components move in different load conditions. Simulating the force transmission in soft tissue components can solve the problem of how those components deform under load. Thus, it is suitable for simulation inter-object and intra-object correlations.

For calculating the solution of a finite element approach, all involved components have to split up in simple, smaller parts, the so-called finite elements. These parts are so primitive that their behavior can be described with a set of linear equations. As the finite elements are influencing each other, a set of boundary and transition conditions have to be described. Certainly, other equations have to be used for calculating the heat transportation capabilities of an element than for calculating the force distributing properties.

[6] Deformable components will even change their shape

For a detailed description of a complicated geometrical 3D object, up to 10^7 unknown variables have to be solved for the resulting linear equation system. With common hardware, this can not be done in real time, what is an enormous drawback for time dependent applications. Nevertheless, there are many scientific approaches that try to model the knee joint by the Finite Elements Method, for example Schreppers [29], Meaken [30], Blankevoort [31] and Maldonado [32].

There are several drawbacks of these approaches that make it unsuitable for education purposes; Meaken was only researching a two-dimensional model of the meniscus (see figure 1.11 left), while Blankevoort's research was done with an idealized three dimensional model without menisci and without friction (see figure 1.11 right). The focus of Maldonado's work is the strains in the humerus in different load conditions, thus soft tissue is not explicitly considered.

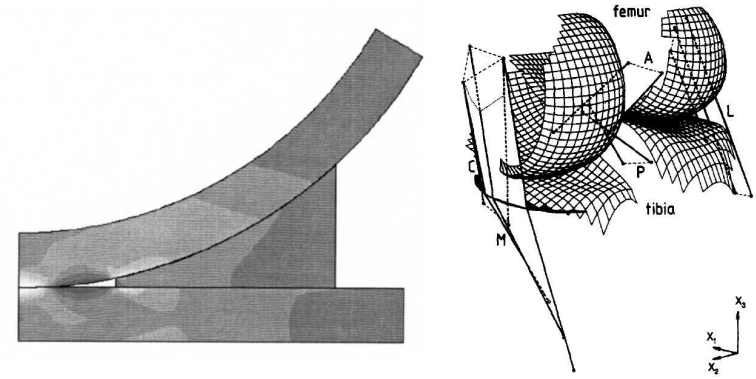


Figure 1.11: Meakin's two-dimensional FEM approach (left) and Blankevoort's simplified three-dimensional model (right)

As all of the scientific simulations of the knee joint ignore relevant parts or constraints, the resulting movements will be correct in a small range of motion only. Extending FEM to the needs of the Knee Joint Simulator is discussed in section 1.6.

Mass-Spring-Damper-Models

The mass-spring-damper-models like [33] or [34] are more macroscopic than the FEM approach. Thus, a distinction has to be made between rigid objects and deformable objects. Rigid objects are not split up in microscopic particles, but represented according to their real geometric appearance. Intra-object processes are ignored, only inter-object processes are relevant, e.g. if collisions appear or if two objects are connected (see figure 1.12).

Deformable objects are treated similar as rigid objects, as they can also collide or be connected with others. Additionally to this, forces can deform the object. This deformation can be simulated, if the deformable structures are represented by a set of zero dimensional vertexes which have a defined weight. These vertexes are interconnected via edges. Edges are defined as springs and dampers, sometimes with additional viscous elements. The location of the vertexes correspond to the geometry of the object. The connection scheme is not standardized, i.e. every software implementation uses different algorithms. Most

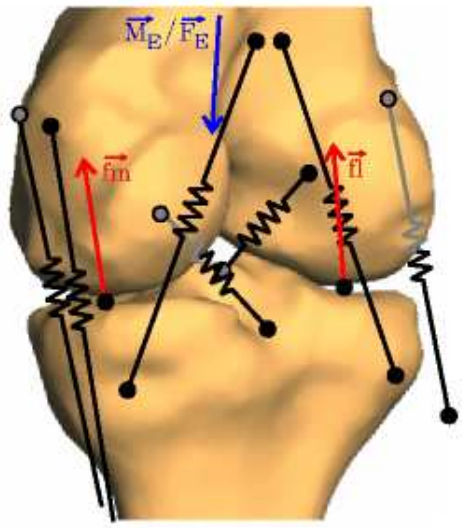


Figure 1.12: Wismans' knee joint model with springs

authors like Gohl [35] and Kuhn et. al. [36] use the surface mesh (which was created by the Delaunay triangulation [37]) for their simulations.

If a force is applied to one of the vertexes, this vertex experiences an acceleration. Thus it changes its current speed and location, depending on its inertia and the damped connections to other masses. As those connections transfer certain forces to the surrounding vertexes, a punctual force will affect a certain area - just like in real deformable objects. The given parameters for damping spring constant and mass influence the characteristic of this force distribution.

Interactions of rigid objects that are implemented as mass-spring-damper systems are much faster than the FEM approach, however, even soft tissue deformation can be much faster since the equation systems are less complex and less equations have to be calculated, as only the surface vertexes are relevant. This increased speed has its disadvantages. In general, mass-spring-damper systems are less exact than FEM systems and more heuristic. The parametrization of FEM is done by material constants, whereas the mass of the vertexes, the spring constant of the spring and the damping of the damper is adjusted heuristically to give the "best" result.

Since they are quite easy to set up, mass-spring-damper models are quite common. Nevertheless, all implementations in literature neglect either the friction, the menisci, the cartilage, the capsule or combinations of them. Some of the most basic approaches do not even consider the real shape of the bones, but replace them with geometric primitives like cylinders, cones or spheres [38] [39]. These animations resemble only vaguely real knee joints, which have a much more "organic" appearance (see figure 1.13).

Due to this simplification, the knee joint dynamics can be calculated very easily. Collision detection is very simple with perfect geometrical primitives, thus the resulting contact forces and movements can be estimated in real time. Thus, this approach offers the possibility to complete interactivity. Changing the knee joint angle, rotating the knee joint, zooming and hiding objects is theoretically possible.

Unfortunately, the behavior of these approaches is not very realistic. Due to the simplified geometry, the virtual knee joint can not behave exactly as a real knee joint.

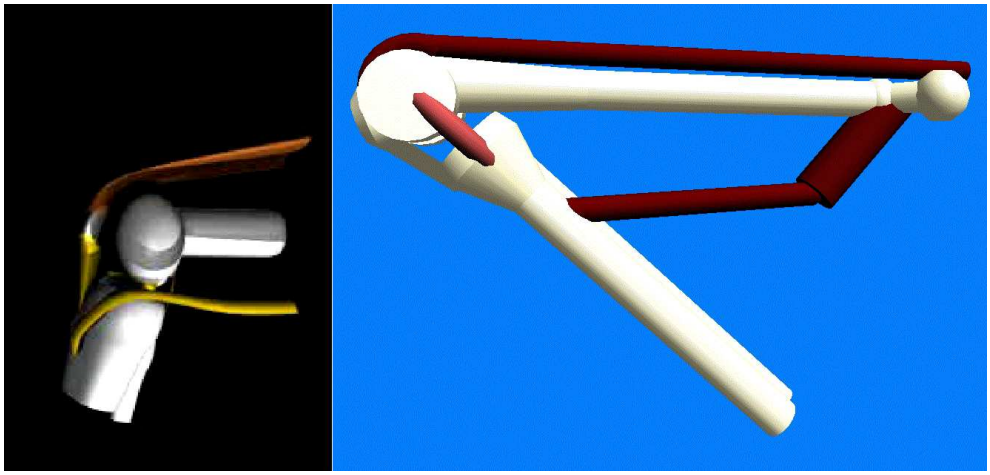


Figure 1.13: Two animated knee joints based on geometry primitives: Rob Kroeger (left) and Riener et al. (right)

A more realistic approach was done by Wismans [34] as it was is already shown in figure 1.12. Rigid bones were connected by springs that simulate the ligaments. However, as major parts are missing in this approach, a realistic animation seems unlikely.

Particle Systems

Particle Systems are quite similar to the mass-spring-damper models, as they use zero dimensional vertexes. However, the edges are not defined as spring and dampers. Each particle interacts with each other by a physical function. This interaction may simulate the gravitational attraction if the particle system is set up for the simulation of a galaxy (Appel [40]). In this case, the functional interaction allows a very free motion of all implemented particles.

Other approaches simulate liquids (Miller et al. [41] and gasses (Vapilon et al. [42] and need other functional systems that implement distraction between the particles, but still allow a very free motion.

If the particles are set up with an orientation and if the functional between those particles is set up to simulate a grid, then deformable or rigid bodies can be simulated, according to Szeliski [43].

However, the advantages of the particle system is the fact that the particles can change their position, so their focus is set on simulating amorphous materials like fire, smoke, sand, snow. Simulating deformable materials like the menisci that do not require the free motion of all particles is possible, but not efficient. Implementing rigid bodies like the bones would be a pure waste of calculation time.

Keyframe interpolation

Video games are very interactive as the user can influence almost every detail in the game play. The variety of used techniques is large; physical simulations are used to calculate how the surrounding reacts on an explosion, how a race car reacts on a pothole and how an airplane reacts on a damaged wing. Artificial intelligence is used to calculate the responses

of enemies against the user's actions and neuronal networks are learning the preferences of a human chess player to adjust the calculations for the next move.

As video games are also dealing with real time graphics, some strategies might be applied for the knee joint simulator, too. One of the most important techniques in this sector is called 'keyframe interpolation' and is used by artist for drawing animated cartoons [44]. Instead of drawing dozens of frames for a single scene, only a few frames are drawn by the artist; the missing frames are generated by the interpolation algorithm (this is also called "in-betweening").

This keyframe interpolation can be done in multiple ways. The easiest way of interpolating a stick-man's motion in 2D is to interpolate its joint angles from the first drawn image *linearly* to the second image.

However, if linear interpolations are used, interpolating more than two images will give discontinuous motions at each sampling image. Thus, more sophisticated approaches are common that use smooth interpolation algorithms or energy constraints [45].

Another problem of the keyframe interpolation is that interpolations are always estimations and can not provide exact values. The feet of a walking stick-man will either sink into the floor or float above it. Thus, keyframe interpolations are often combined with physical approaches and collision detections.

In computer games, a related technique is the "facial animation" (Bonamico et al. [46] or Li et al. [47]) that is used to blend between two facial expressions, e.g. from neutral to angry, happy or sad (see figure 1.14).

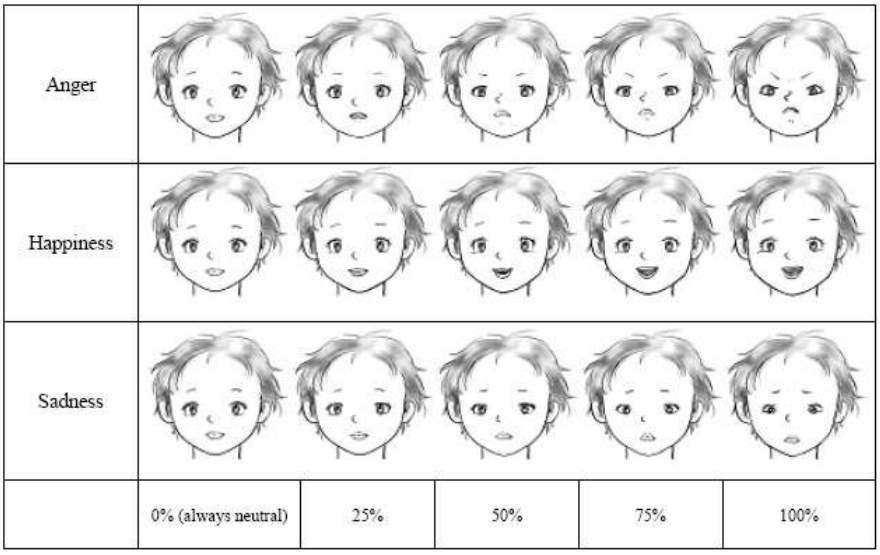


Figure 1.14: Blending from the neutral emotion to other emotions can be useful for cartoon drawing and computer games

With this technique, the interaction with virtual humans in computer games becomes more realistic; thus, this technique can be used to for interactive systems. However, keyframe interpolation and facial animation is only one dimensional in most cases. The stick-man can walk forth and back and the emotion in the face of the girl can slide between neutral and sad.

This is not sufficient for the knee joint animation, as the knee can not only swing forth and back. At least three DoF should be interpolated to allow a thorough examination. Multi dimensional keyframe interpolation would be necessary, but these techniques are not yet fully developed. Sloan et al. [48] and Rose et al. [49] developed a two dimensional keyframe interpolation for computer games, but a three dimensional approach for computer games is not known until now (see figure 1.15).

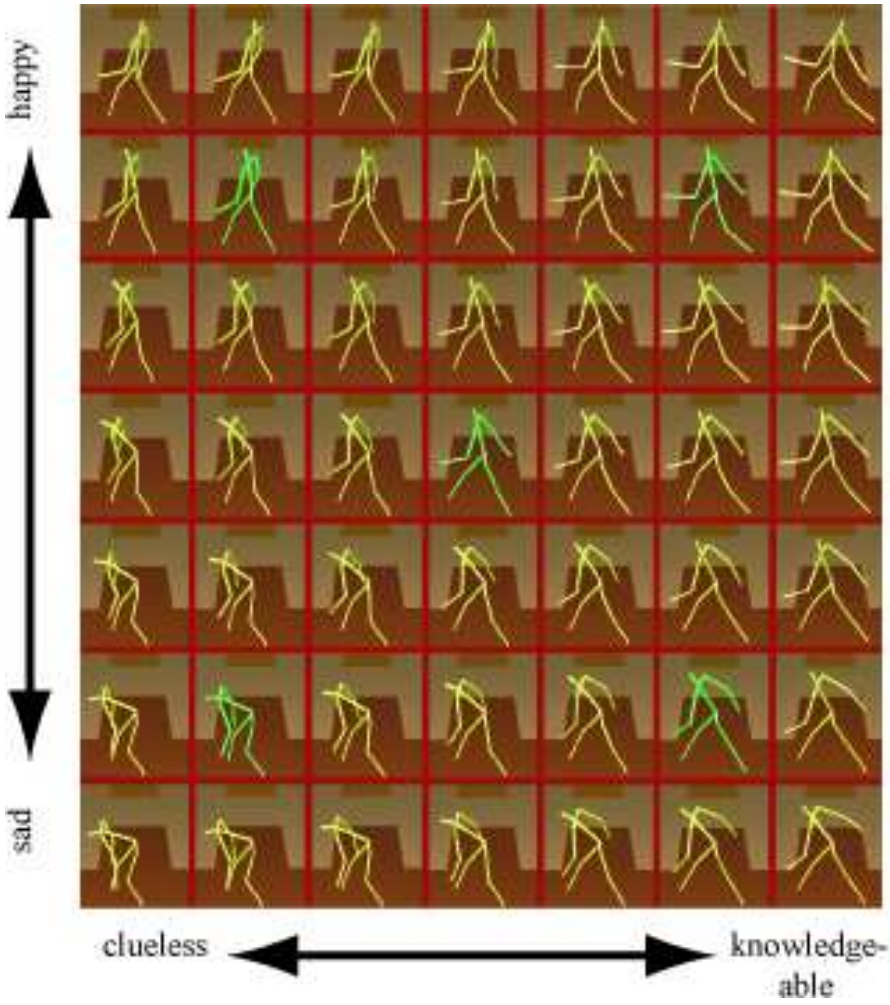


Figure 1.15: Only five figures were drawn manually, the rest was interpolated by the interpolation algorithm of Rose et. al.

Another aspect of the keyframe interpolation that has to be considered is the target group. Animated comic strips are neither used for displaying the reality, nor are they bound to physical correctness. If the foot of a comic character penetrates the floor because of an interpolation error, the artist can easily provide an additional keyframe. This is not possible for the knee joint animation, since acquiring knee joint poses is extremely time consuming, what makes the acquisition of further data hardly possible. The data acquisition process in chapter 2 turned out to be one of the hardest challenges of this approach.

1.5 The challenge from the medical point of view

In section 1.4, different methods for generating moving pictures with deformations are described. Each of those methods has its advantages and disadvantages, which have to be considered carefully whether this method can fulfill the medical requirements.

1.5.1 Interaction with the animation

Giving the user the possibility to explore a virtual knee joint is the most obvious innovation of the Knee Joint Simulator. This exploration includes the manipulation of the knee joint's flexion angle and its other degrees of freedom. Thus, influencing the knee joint posture is one of the most basic interactions the graphical interface has to offer.

Another challenge for the animation is the free choice of the point of view. In still images, the photographer or artist can choose one fixed point of view which allows a thorough examination of all components. However, watching the knee joint internals in motion is difficult, because most of them have a central position, so other knee components may block the user's view in certain knee joint postures. Thus, a fixed perspective is not appropriate. Another possible approach to avoid occlusions in dynamic animations is to give the user the possibility make occluding parts transparent or invisible. Further interactions that should be offered by the graphical display are standard operations like zooming, positioning of light sources or influencing the colors of the knee joint components for manual contrast improvements.

2D animations can only provide zooming and influencing the knee joint posture in one degree of freedom. Changing the perspective or influencing the visibility of certain components can not be realized with this technique, so it can not be used in this work. Static 3D models avoid these two drawbacks, but changing the knee joint pose is still not possible.

Thus, interaction is only possible with physical based models or keyframe interpolation on six degrees of freedom.

1.5.2 Macroscopic and microscopic precision

One of the hardest challenges from the medical point of view is the high grade of precision that is required for an educational tool. The term "precision" can be subdivided in "macroscopic precision" and "microscopic precision".

Macroscopic precision would be the correct appearance at the first glimpse. If even unexperienced users see obvious flaws like severe errors in the displayed positions and orientations of the knee joint components, macroscopic precision is not given. Failure in the macroscopic precision is fatal for an educational tool.

Microscopic precision is needed for closeup views: Imprecise details like small interpenetrations between two objects or simplified geometries will be visible to the skilled user and might cause confusion. However, unexperienced users can ignore those details and can be satisfied with a correct macroscopic impression.

Most of the very impressive animations that are known from movies are not based on real world data, but on pure fiction. Thus, the macroscopic precision is limited, as the laws of physics are often ignored. Moreover, these animations often also fail in the microscopic precision. As the perspective can not be chosen by the viewer, obvious errors can be cut out, hidden in shadows or touched up after the generation of the movie. In this

project, touching up is not possible and hiding information would contradict the idea of an educational tool.

Animations that are created by geometrical primitives can offer a high grade of microscopic precision, as the simple geometry allows a fast collision detection, but neither the appearance of the knee joint components, nor their positions or orientations can provide a good macroscopic precision.

FEM (and precise mass-spring-damper systems) have the same microscopic precision, but they can also offer a higher grade of macroscopic precision since they can represent the bones as structures of complex, organic geometries. The calculations have a more realistic input, which theoretically improves the macroscopic precision. In practice, the grade of macroscopic precision is still not clear as major knee joint components are neglected.

If keyframe interpolation is used to interpolate measured data, macroscopic and microscopic precision depend on the quality and amount of the given data and on the quality of the chosen interpolation algorithm. However, since every interpolation bears small errors, microscopic precision seems to be the greater challenge, because even those small errors can cause interpenetrations on laminar contact scenarios.

As a result, only the physics-based approaches or the keyframe interpolation have the potential to fulfill the medical requirements. As it is discussed in subsection 1.6.1, FEM and mass-spring-damper-models can not be used in this project, so the keyframe interpolation remains. For this approach, the input data has to be acquired, which is a challenge from the medical point of view and is discussed here.

1.5.3 Research of data about the knee movement

To set up a three dimensional animation which provides a high grade of realism, the current knowledge of the knee joint movements is not sufficient. In most studies, the characteristic movement of knee joint parts are measured from CT or MRI slices. Unfortunately, those measurements are limited to only one or two knee joint components, which prevents the complete animation of all components.

Another aspect which makes it impossible to simulate the knee joint from the data of other studies is the fact that most of those values are obtained from the two dimensional slices. Three dimensional animations can not be obtained from these two dimensional measurements, since they ignore one DoF. The third aspect is that most studies have researched only the knee joint flexion, omitting all other movement possibilities.

As a literature research can not provide enough reliable or miscible data to set up a complete, realistic and precise model of knee joint movements in multiple degrees of freedom, a profound data acquisition has to be done for generating this animation.

1.6 The challenge from the technical point of view

In the last section it is described what medical doctors or students expect from a perfect graphical animation. These requirements prohibit the usage of different technical approaches. However, there are further restrictions from the technical point of view that are discussed here.

1.6.1 Real time animation

The knee joint simulator or the stand alone animation should be interactive, so whenever the user moves the virtual shank, an immediate response of the animation is expected. As several objects in high 3D resolution have to be moved and deformations of soft tissue components have to be calculated, the computational requirements are high and not all approaches are quick enough to provide an immediate response.

From the medical requirements, only sophisticated mass-spring-damper approaches, FEM models and the keyframe interpolation are candidates for generating the animation. Tests show that three dimensional FEM calculations need several hours to calculate a single pose. Even if this could be optimized by factor 1000 (e.g. by triangle reduction, parallel processing or code optimization), a fluent animation is still not possible. Thus, the FEM strategy was not embarked further on.

Today's hardware is also limiting the performance of mass-spring-damper approaches. As the computational effort increases in an exponential manner with the number of triangles, the speedup by code optimizations or parallel processing is only limited. Tests [35] showed that a Pentium 4 (1.2 GHz, 512 MB RAM) with a professional NVidia Quadro 4 graphics card can not calculate collisions between a punctual object and a deformable object with more than 20,000 vertexes, 60,000 edges and 40,000 triangles fluently at 25 fps.

In the scenario of the knee joint simulation, no punctual, but laminar contact situations occur. Thus, the collision detection has to compare the position of several thousand triangles of one object with several thousand triangles of the other object.

This setup can be calculated with today's hardware, if both objects would be non-deformable. In these cases, cascaded bounding boxes can be applied [50] which simplify the calculations significantly. However, since at least one of the colliding objects is deformable, this technique can not be used. Thus, the calculation on modern PC is only possible, if both objects are reduced to a few hundred triangles.

As every soft tissue component needs around 20,000 vertexes for a realistic appearance and as at least six soft tissue components[7] have to be calculated, a fluent animation is not possible today.

1.6.2 Data acquisition

As it is mentioned above, the knowledge of human knee joints is not sufficient to implement a graphical display which meets the given demands. Thus, a data acquisition process has to be performed.

The result of this data acquisition process is a detailed knowledge of the behavior of all knee joint components during all possible knee joint movements. This knowledge should include precise quantifications of locations, orientations and shapes. The technical challenge is the measurement of those parameters with a high precision, without changing the knee joint properties.

As mechanical measurement methods of knee joint internals require direct access to those components, the destruction of surrounding tissue is necessary. This destruction would alter the knee joint properties, so non-invasive methods have to be chosen, what is explained in chapter 2. In this work, an image generating method is used.

[7] At least 2 menisci, 2 cruciate ligaments, LCL and MCL. Modeling the capsule, the cartilages of further ligaments might be necessary to obtain realistic results

A three dimensional image generating method like CT or MRI produces a stack of slices of a certain resolution (256x256 pixels in most cases). In general, the slice thickness and the distance between two slices causes a worse resolution in this dimension. Due to this anisotropic resolution, the voxels[8] are not cubes, but cuboids.

In most cases, the object orientation and the voxel orientation in the scanning setup is optimized to minimize this effect, i.e. the object's largest dimension is oriented perpendicular to the high resolution slices. For studying the ligaments which are in a vertical position, a horizontal orientation of the slices is ideal.

In this study, all internal knee joint components have to be scanned. As the menisci are more or less perpendicular to the ligaments, no voxel orientation is ideal for both components. So, the slice thickness and the distance between two slices has to be minimized to obtain a nearly isotropic resolution, thus to have equal precision in all three dimensions.

In this work, a resolution around 512x512x180 voxels is chosen, which leads to a voxel size of $0.29x0.29x0.7$ mm and a size of more than 100 MB for each data set, which is outstanding. However, post processing such precise scanning is a challenge for every step of image processing, as common hard- and software are driven to their limit.

The advantages of the high resolution are visible in the 3D models that are available after the post processing and reconstruction. As multiple scans are performed and as those precise 3D models can be measured very precisely, all necessary parameters can be extracted to allow a keyframe interpolation between the scanned discrete poses.

1.6.3 The graphical display as part of the Knee Joint Simulator

The haptic interface is using a biomechanical model to determine the rotations and translations of the lower leg relative to the thigh that occur due to external forces. The output of this biomechanical model is a transformation matrix that can be used as the input for the graphical display. However, it is not possible to use this matrix directly to display the tibia position and orientation.

The haptic display is simulating a different knee joint with a different geometry. Moreover, some degrees of freedom are scaled to provide a better haptic impression. If these values would be used directly for the graphical display, neither microscopic nor macroscopic precision would be given as collisions and interpenetrations of knee joint components would occur frequently. Thus, only the three rotational parameters are accepted from the biomechanical model so that the three remaining translation parameters can be chosen freely to ensure a realistic knee joint posture.

1.6.4 The graphical display as a stand alone tool

When using the stand alone tool, the user wants to influence all 6 DoF of the knee joint position and orientation directly. As no constraints like a biomechanical model is between the user's input and the graphical output, the user could provoke even very unrealistic postures like too large rotational angles or too large translations. The visualization of such unrealistic postures is not suitable for an educational tool. Again, the input data has to be treated in a certain way to fit to the properties of displayed knee joint. In general, the input possibilities have to be limited or scaled to a reasonable extend. In this work a

[8] Pixels are the smallest units in a 2D data sets. Voxels are the same in three dimensional sets.

heuristic approach is chosen; a natural extrapolation was implemented, that extends the measured values by 10%, what gives a realistic impression.

1.7 This approach

As the physically based approaches miss the real time requirement and show many other drawbacks, the keyframe interpolation technique is chosen. Therefore, it is necessary to acquire a set of anatomical data that can be interpolated.

For this data acquisition, a human cadaver knee joint is scanned with an image generating method in many different load situations. Translational load is not applied, as only rotational parameters can be accepted from the biomechanical model (see section 1.6.3) and to achieve a reasonable amount of necessary scans.

Moreover, if only three rotational DoF are accepted as user input or as input from the biomechanical model of the Knee Joint Simulator, the three other DoF can be used to generate a realistic output of the graphical display. Thus, the tibia *orientation* is directly determined by the user or the biomechanical model of the haptics, while the *position* of the tibia is chosen according to the knee joint characteristics. Other knee joint components like the patella have to be treated in another way. The user does not influence neither their positions not their orientations directly, so all six DoF have to be calculated, depending of the tibia orientation.

Soft tissue deformations are yet another challenge, since they extend the six DoF problem of orientations and positions for further n dimensions. Different deformation algorithms exist which determine this n to be between 3 and 3000 for each soft tissue component. In this approach, 13 landmarks are placed on each deformable component, which allows a warping of the whole object with the well known "Thin Plate Splines" algorithm. Thus, the deformation problem of this approach has $n = 13$ dimensions.

1.8 Thesis outline

This center of this thesis is the human knee joint. For describing it, medical terms are used throughout this thesis. Readers from other faculties than medicine may not know the meaning of those terms and are advised to read Appendix A first.

The chapter 1 of this thesis is the introduction that explains the motivation of the project and the importance and the state of the art of current graphical animations. The description of the requirements from the technical and medical point of view shows that neither state of the art animations nor the state of the art knowledge of the knee joint is sufficient for a realistic animation. Finally, the approach of this work is described that consists of a data acquisition and interpolation.

As the common knowledge of human knee joints is limited in many aspects, a data acquisition has to be performed. In chapter 2, different image generating methods are compared, a scanning setup is designed and the data post processing is explained in detail. The resulting data set is the most complete and detailed data base of a human knee joint.

The acquired data only represents distinct postures of the knee joint, what is not sufficient for the Knee Joint Simulator. As smooth transitions are needed between those measured postures, an interpolation algorithm is developed that can handle positions and

deformations in the Euclidean space as well as rotations in the Hamilton space. Common approaches for computer animations are limited in different ways, so in chapter 3 different basic mathematical approaches are used, that are normally not used for animation purposes.

Realistic real time deformations are needed for animating detailed representations of soft tissue components in the knee joint. Today's hardware and algorithms allow only a certain compromise of the three quality requirements "realism", "real time animation" and "detailed geometric appearance". In chapter 4 the different possibilities of deforming objects, including the "Thin Plate Splines" algorithm are explained and evaluated.

This thesis demonstrates new ways of analyzing movements and deformation of biological material. These techniques can be transferred to other research fields; some of those are described in chapter 5. Moreover, the limits of this technique make also clear, in which research fields it should *not* be transfered.

2 Data acquisition and processing

2.1 Introduction

Modern commercial simulators for the entertainment sector simulate vehicles like racing cars or airplanes in a very realistic manner. External parameters like wind or rain, as well as internal parameters like the gear ratios or suspension parameters influence the behavior of the vehicles. Thus, for setting up such simulators, a realistic model of the vehicle has to be set up that describes the interaction possibilities of the different internal and external components. Moreover, those different components have to be parametrized in a correct way.

In the case of a racing car simulator, the behaviors of the different characteristics of the shock absorbers and the springs of different cars are known from literature or the manufacturer's data sheets. If those values are used in the correct model, the combination of them can predict the behavior of a whole suspension element.

Another way to simulate the suspension element is the simulation of the oil inside the shock absorber with a particle system (see section 1.4). This is a more microscopic approach, so the used model has to be another than the one that is mentioned above. Thus, the decision which model to use depends partly on the requirements like precision or speed, as it has already been mentioned in Section 1.4, but also on the kind of information that is to be used.

For the human knee joint, no manufacturer's data sheets are available. Data from literature and other data acquisition possibilities show flaws that are explained in the next subsections. In 2.2 it will be determined that positions and orientations of the knee joint components have to be measured in different postures. Section 2.3 describes the relevant image generating methods that are able to acquire those positions and orientations. Section 2.4 describes the used setup for the data acquisition, whereas the post processing of the acquired data is presented in section 2.5 and analyzed in section 2.6. The sections 2.7, 2.8 and 2.9 describe the knee joint behavior from three different, macroscopic perspectives. Finally, the soft tissue deformations are analyzed in section 2.10 and a conclusion is given in section 2.11.

2.2 Parameters to measure

2.2.1 Physical properties of knee joint components

Physical parameters of knee joint components like shapes, friction coefficients, or tensile moduli can be used for FEM or mass-spring-damper models. Apart from the drawbacks that have already been mentioned in Section 1.4, the acquisition of those parameters is challenging.

None of the current research articles provides all necessary values to simulate a complete knee joint. Even worse, none of the current research articles provides all necessary values to describe more than one component of the knee joint. Even if all available values of all research articles are merged together, some parameters (e.g. the friction coefficient between the patella and the femur) have never been measured.

If the missing values could be measured, the result is still very doubtful, as the merged values are most probably not compatible. One of the most basic parameters, the tensile modulus of cartilage was measured by Fithian et al. [51], who found more than 300% difference between two individuals. Tissakht et al. [52] agrees with the individual 300% difference, but found values that were approximately half as large as the values of Fithian et al. Thus, two factors avoid the merging of results from literature: The individual variation, that is definitely too large to be ignored and differences in the measurement setups.

Biological objects are not as regular as most technical objects that are constructed by humans. Even longish bones like the femur or the tibia are not straight tubes, but are significantly bent with an irregular radius. Due to that irregular shape, two femurs from two different individuals can neither be aligned in a parallel nor in a concentric manner.

As the measurement procedures are not standardized, it is very likely that different research groups measure the tensile modulus of the cartilage at different locations of the bone, where a different thickness, a different curvature and a different direction of the force cause systematic measurement deviations. Thus, in addition to the individual differences, the non-standardization leads to further errors.

As there are so many inconsistencies and unknowns in today's knowledge about these parameters, a reliable data acquisition is necessary for using FEM or mass-spring-damper systems in this project. Unfortunately the huge number of necessary parameters and the difficult measurement setup make this approach very unlikely.

For the measurements, at least eight different materials (3 cartilages, 2 menisci, 4 ligaments, 1 capsule) have to be taken into account. For each of those eight materials, numerous of different parameters (shape, friction coefficient, tensile modulus) have to be measured. Depending on the material, these parameters are not constant over the whole surface, i.e. different regions of the same meniscus have different tensile moduli, requiring even more measurements.

In addition to the large number of parameters that have to be measured, most of those parameters are direction dependent and non-linear. The condyli are generally loaded with pressure orthogonal to their surface. The ligaments are generally loaded with tension along their longitude. For both materials, the non-linearity is relevant, but the direction dependence is not. However, the menisci can be loaded with anterior, posterior or medial/lateral pressure or even torsion, these measurements have to be performed in full extend.

Due to the natural muscle tone, the stiffness of the capsule and the knee joint geometry, the unloaded knee joint bears internal forces and torques. These internal forces of the neutral posture have to measured for an adequate parametrization of the neutral posture of the model. Thus, disassembling the knee joint for easier measurements changes the load situation, which might change the shapes of certain components and their friction coefficients.

Keyhole measurements can avoid the disassembly of the knee joint, but are more complicated; due to the high number of parameters to be measured and the complicated measuring approach, the duration of the measurements is a critical factor in respect to the degeneration.

Thus, most FEM based approaches are set up with the data from literature, although the reported range of each parameter is significant. Additionally, none of the different approaches consider all relevant knee joint components, which almost certainly influences the result of the calculation. The final uncertainty is added as different implementations of FEM solvers find different solutions, even if the same model was used.

As a result, some parameter combinations lead to a reasonable behavior with a certain FEM implementation, but some do not. Thus, the parameter combinations are not chosen according to scientific methods, but more or less by trial and error.

2.2.2 Positions and orientations of knee joint components

As reliable physical parameters are not available and as a profound and comprehensive data acquisition for FEM or mass-spring-damper systems is unlikely, another model for estimating the positions and shapes of each knee joint component for every possible posture has to be found.

The physical approach can be called a "bottom-up" approach, where acquired, basic, physical parameters are used with known, basic, physical laws to calculate macroscopic movements and deformations. The approach presented here is based on acquired macroscopic postures. From these postures, movements and deformations one can estimate and interpolate.

The acquisition of the macroscopic postures can be done with medical imaging methods. These approaches use different physical effects to measure the water content or the x-ray attenuation for every microscopic volume unit (voxel) of a body part. The calculated value of each volume unit can be displayed as a gray value. Depending on which volume units are displayed, different slices in different orientations can be represented. As different materials have different water concentrations or different x-ray attenuation, the different materials are shown in different gray tones.

This approach overcomes the main problems of the physical approach that has been mentioned above, as the complete data acquisition process can be performed without the destruction of the knee joint. Thus, the problem of merging data from different knee joints, measurement setups or literature sources does not occur here. Moreover, uncertainties due to unconsidered knee joint components or different solver implementations cannot occur here, as the postures are *acquired* and not just the result of error prone calculations.

However, there are several uncertainties in this approach, too, which are explained in the next section.

2.3 Medical imaging

In the medical sector, non-invasive image generating methods are very common. These methods give a precise insight into the human body and allow the measurement of internal parts of the body. For seizing the internals of a knee joint, two different methods have to be taken into account: CT scans and MRI scans.

2.3.1 CT scans

The term "CT" is an abbreviation of "computer tomography" which is derived from Greek: tomós (slice) and gráphein (to write), describing an image generating technique that is capable of showing slices of the scanned object.

The scanned object is located between the x-ray source and the x-ray detector. X-rays that are emitted from the source pass through the scanned object and being more or less absorbed, depending on the material properties of the object. After the attenuation of the x-rays is recorded, the x-ray source and the x-ray detector are rotated around the scanned object (about one degree in most cases) and the next measurement is performed. When 360 degrees of x-ray scans are completed, a three-dimensional dataset of voxels can be computed. Unlike normal x-ray photos, distinct slices of any orientation can be derived from this dataset in seconds. The gray value of each volume element is represented by a 12 bit value, thus $2^{12} = 4096$ different values can be distinguished. These values are called the "Hounsfield units" (HU), a scale between -1000 HU (which represents air) and 3096 HU (complete attenuation). Water has zero HU, most voxels of bones are between 100 HU and 1500 HU, ligaments between 60 HU and 200 HU, cartilage around 50 HU to 150 HU, the menisci between 50 HU and 200 HU, while metals have more than 1000 HU.

CT scanning is a very quick way of data acquisition and can be performed in less than one minute, depending on the scanner technique, the desired resolution and the scanned area. The disadvantage of the CT technique is the low contrast for soft tissue and the harmful x-ray radiation.

2.3.2 MRI scans

The atomic nuclei of every macroscopic body consist of neutrons and protons. As the protons can exist in spin states, the nuclei have a magnetic moment. As these magnetic moments are not ordered, they compensate each other and the macroscopic objects do not show macroscopic magnetism. However, if the magnetic moments of the nuclei are aligned by a powerful external magnetic field B_0 (modern MRI scanners use up to three Tesla), they superpose and the macroscopic object shows a macroscopic magnetism of several Nanotesla (nT).

A second, pulsed magnetic field (RF pulse) is set up perpendicular to the existing magnetic field B_0. If the correct pulse frequency is chosen, the spins of hydrogen nuclei are influenced by this excitation field and start to precess, i.e. their axis of gyration are not aligned along B_0 any more but start to rotate. The pulse frequency is called the "Larmor frequency" and must be chosen according to the strength of the magnetic field and the material that is exposed to that magnetic field. This effect is quite sensitive, i.e. if the pulse frequency is slightly different to the Larmor frequency, no precession will occur at all.

After the excitation, the pulsed magnetic field is turned off. The precession does not stop at once, so the resulting macroscopic magnetism can be measured by electromagnetic induction with the coils that produced the RF pulse. Thus, these coils are also called "receiving coils". The intensity of the macroscopic magnetic field depends on the number of precessing nuclei, but decreases fast due to two effects:

- As the perpendicular forces vanish, the axis of gyration will again be aligned parallel to B_0. The effect of restoring the parallel alignment is called the "longitudinal relaxation".

- The synchronized rotation of the axis of gyration is getting more and more disturbed. Sooner or later, the axis might still rotate, but if it is not synchronized any more, the magnetic fields compensate and no macroscopic magnetism can be measured perpendicular to B_0. This is called the "transversal relaxation".

The time of the longitudinal relaxation is called T1 (typical between 0.2s and 2s), while the time for the transversal relaxation is called T2 (typical between 20ms and 200ms). If the axes of gyration are aligned parallel to B_0, no precession takes place any more, so $T1 > T2$ due to physical reasons.

T1 and T2 are two parameters that describe the kind of material that is scanned, while the intensity of the signal describes the density of hydrogen. Thus, three different material properties can be measured with the MRI technique.

For image generating purposes, T1, T2 and intensity have to be obtained from every voxel of the object to be scanned. This can be obtained by a two step approach. First, a linearly changing magnetic field is established. This is the first of the three "gradient fields". As the Larmor frequency of all nuclei depends on the magnetic field they are exposed to, the RF pulse can only excite a thin slice perpendicular to the direction of the gradient field. Thus, this gradient field is called the "slice selection gradient".

In the second step, a further gradient (the so-called "readout gradient" or "measurement gradient") is established orthogonal to B_0 and to the slice selection gradient field. This is done after the RF pulse and the slice selection gradient are turned off, but still before the inducted current is measured in the receiving coils.

Like the slice selection gradient, the readout gradient is a magnetic field that does not change over time, but its magnetic force changes linearly in space. Thus, this field is not able to cause a precession of the nuclei, but it can influence the Larmor frequency of those nuclei that are already excited. Thus, each row of the chosen slice inducts its current into the receiving coils with another frequency. This can be analyzed with the Fourier transformation to obtain the amplitude and position information.

With this kind of measurement, the content of hydrogen of a volume element of an object can be estimated. If the detector is moved around the object (like in a CT scanner), a three-dimensional dataset can be acquired.

MRI scans are especially suitable to display soft tissue, since it contains a reasonable amount of water. However, bones are clearly visible, too. Another advantage is the absence of dangerous radiation, so repeated in vivo measurements are possible without doubt.

There are also several disadvantages of MRI measurements, for example the long duration of the scans. Another problem is the surrounding natural and artificial magnetic field that causes distortion in magnetic fields. To avoid the influence of

external fields to the scanning setup, MRI scanners are constructed as shielded tubes. However, as the small volume of these tubes pose problems for claustrophobic people and as they forbid the scanning of huge objects, open MRI scanners are also available. As they are not shielded, the distortion has to be compensated by scanning calibration objects to recalculate the results, but this method has its limitations, too [53].

2.3.3 Conclusions about medical imaging

To produce a fluent animation, the acquired postures have to be interpolated. Each acquired posture is a control point of the inter- and extrapolation function. Except for these given control points, inter- and extrapolation algorithms are usually not precise. Moreover, the error grows with the distance to the next control point. Thus, the most desirable setup for data acquisition provides as many control points as possible. Moreover, the given control points should cover an area that is as large as possible, since extrapolation is always less precise as interpolation.

As soft tissue components have to be simulated, too, MRI scans would be ideal. Some knee joint components have a very small size in one dimension, e.g. the cartilage is flat, ligaments are thin. For a precise depiction of the whole knee joint, all three dimensions should have ideal resolution (e.g. 0.2mm x 0.2mm x 0.2mm).

Tests show, that this ideal scanning setup is not possible. The scan of a single posture in a resolution of 512x512 pixels and more than 150 slices takes several hours with a modern MRI scanner. This kind of data acquisition is hardly possible if the MRI scanner should also be used for other purposes than research. Moreover, the tests showed that volunteers have severe problems not to move their legs over this period of time, especially if the posture is stressed, not comfortable or even painful. Movements will lead to movement artifacts [54], which spoil the whole data set. Cadaver knees show severe signs of degeneration if multiple scans are performed over several hours.

To overcome these problems, MRI scans are performed in different lower resolutions (512x512x30, 256x256x140, 256x256x70, 128x128x120) and post-processed as described in section 2.5. In this post-processing task, the bad resolution turns out to be fatal: If only 30 or 70 slices are chosen, the identification of vertical structures like the cruciate ligaments fails, as most of their geometry is lost 'between' the slices. Increasing the slice thickness avoids those gaps, but since parts of the ligaments are thinner than those slices, their contrast is poor and the recognition still fails.

If the resolution of the other dimensions is decreased in favor of the slice distance, the resolution in these dimensions is not sufficient any more, e.g. the cartilage appears only one voxel thick or vanishes completely. A *minimum* resolution of 256x256x256 is necessary for successful image post-processing, while the *desirable* resolution is 512x512x512. Neither of these resolutions can be acquired in an reasonable amount of time with an MRI scanner.

As in vivo measurements are probably not possible with the available MRI scanners due to the limited working space (closed scanners) or distortions, (open scanners) and as in vitro measurements are not possible due to the long scanning times, CT scans of cadaver knees were chosen for this project.

2.4 Scanning setup

For the data acquisition, a CT Scanner (Siemens Somatom Sensation 16) is used. A wooden frame is used to keep the position of the thigh fixed and to apply forces of different directions to the shank. Thus, it is possible to influence the flexion angle, the internal and external rotation and the varus-valgus stress. Additional, an anterior force can be applied (drawer test).

As the bones of the cadaver knee are slightly curved, reproducible and unequivocal straight axis for a precise measurement are not available; however, the data acquisition was done in flexion angles around 0°, 30°, 60°, 90° and 120°. In each of these five

flexion angles, five different load scenarios are applied and scanned: Two of those scenarios are set up by applying a torque of approximately 10Nm to enforce an internal and and external rotation. Two further CT scans are performed by applying 15Nm in varus and valgus direction. The final load scenario for each flexion angle was the natural pose, which is the knee joint without any applied load. These 25 resulting scans are used for the database for the animation.

For evaluation purposes, five further scans are performed around 30 degrees of flexion: One scan is carried out with a load of 100N to enforce an anterior translation, the four other scans are performed by 15Nm varus or valgus load and 10Nm internal or external rotation at the same time.

The forces and torques are applied to the bones by connecting them with nylon strings to hanging weights. The quantity of torque can be easily influenced by adding or removing weights (see figure 2.1). However, the absolute values of forces and torques are not very reliable, since the nylon strings are redirected by rollers without roller bearings, so the friction is unknown.

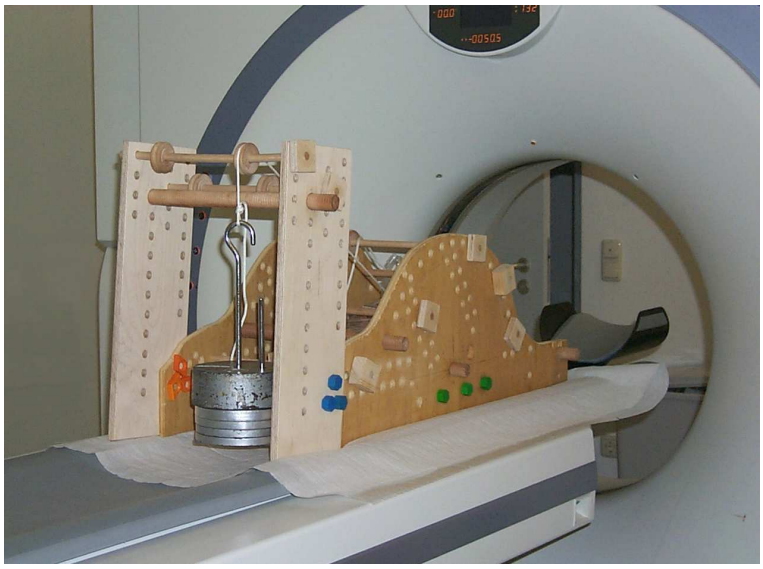


Figure 2.1: This wooden frame is used to adjust the angles of the cadaver knee.

When a force is applied to the shank in a certain direction, its following movement is not exactly in the same direction. In this setup, it is not possible to ensure a pure movement in varus or valgus direction without an inner or outer rotation due to the internal knee mechanics. The differences in the amount of applied forces and the different, non-orthogonal directions of movement lead to certain problems that will be examined in detail in chapter 3.

In general, CT scans are not suitable to depict soft tissue because the attenuation is low and the soft tissue provides only low contrast in the data set. To compensate this drawback, different contrast agents are evaluated. A liquid contrast agent which appears pure white in the scans is found not to be appropriate, as its gray value is too close to the gray value of the bone structures. As those two materials can hardly be distinguished, the

image post-processing is extremely difficult and time consuming. Thus, newer tests are made with air injected into the knee joint. Air does not block x-rays and appears as pure black in the scans (-1000 HU), thus it can easily be distinguished from other materials. Moreover, the gaseous material is better distributed inside the knee joint, so tinier details can be recognized.

For a further image improvement, the radiation dose is increased to a level that gives the best soft tissue contrast. This dose can only be applied to a cadaver knee, because the damage to living tissue is intolerable. As the radiation is increased, longer scanning times occur; nevertheless, the scanning procedure of the 30 different postures takes less than 2 hours and can be done in one evening. There are different advantages of these short scanning times, e.g. there is no freezing and thawing between the measurements, and all measurements are done by the same team.

Finally, a tiny scanning area is chosen; if a very big area is covered by a resolution of 512 x 512, each voxel is quite large. Quick overview scans prove that *only* the relevant parts and *all* relevant parts are visible on the chosen dataset. The resulting resolution of the data set is 512x512x180 voxels, which is a voxel size of 0.29x0.29x0.7mm. This is one of the best resolutions that can be found in literature, it exceeds the minimum requirements and nearly fulfills the ideal requirements for the desired resolution that has been mentioned above.

2.5 Image post-processing

As 30 different data sets with different postures of the same knee joint are acquired, the first data extractions with common 2D methods show inexplicable results. The movements of the components seem unsteady and unpredictable. This can be explained by measurement errors that are in the same magnitude as the measured data. These errors are intrinsic in the common measurement procedures and are widely underestimated:

- The *positioning device* that is used to force the knee joint into certain positions consists of wood, plastics or other materials that are compatible to CT and MRI scanners. These devices bend under serious load, leading to errors of unknown extend.

  If a scale is attached to the device to measure the flexion angle, a single rotation axis is assumed. As real knee joints do not have a single rotation axis, the measurement of this scale is imprecise.

- The *scanning resolution* bears errors, since continuous shapes are divided in discrete voxels. Whenever a voxel is measured exactly on the border of two materials, the so-called "partial volume effect" occurs. As the size of the voxel is larger than the width of the border, it consists partly of both materials. As the voxel is represented by only one gray value, the average gray value is chosen.

  Together with the aliasing effect, the border between two materials is a transition of the gray values and several voxels wide. A precise localization is not possible.

- The *slice wise measurement* of features is very common in the medical sector. The stack of CT slices is leafed through, until the slice is found, that shows the feature best. The measurement is done in only this slice. This very common method is possibly the main source of errors:

- The choice, which slice to use for the measurement is subjective. If the same measurement is performed by another human operator (or on another scanned object), another slice might be chosen, leading to different results.
- Many features can not be measured at all, since they can not be shown on a single slice, e.g. volume measurements, 3D distances.
- Since the measurement is done in slices that are not necessarily orthogonal to the feature's axis, parallax errors may occur. If a normal tube would be scanned and measured that way, it would appear not exactly round but oval. These errors may be interpreted as movements or deformations of objects, even if two identical scenarios are compared.

It should be noted here, that these errors do not occur on their own, but they add up instead. The parallax error sometimes is very severe, since e.g. the cruciate ligaments might be located in acute angles to these projection planes[1]. If it causes the borders of the meniscus to be 4 voxels wide, this border is up to

2.8mm wide. If the knee joint is flexed from extension to the full range of motion, the movement of the medial posterior meniscus horn is 2mm to 5mm according to literature [1]. As these measurements are performed in 30 degree steps, the error due to the partial volume effect can be larger than the real meniscus deformation.

In this project, measures are taken against each of the possible sources of error. An extremely high resolution is chosen to minimize the partial volume effect, automatic tools are used for edge detection and measurements to avoid human estimation errors. For avoiding the effects of the slice wise measurement, all translations and rotations are measured in the three-dimensional objects that can be reconstructed from the CT raw data.

As explained later, this reconstruction process bears many advantages for minimizing several error sources. Unfortunately, the three-dimensional geometries of the knee joint components can not be acquired automatically from the data sets. This requires manual image processing for every slice[2], supported by state-of-the-art tools. The main task is called the "segmentation" (see subsection 2.5.3) and is about allocating every voxel to a virtual material, the three-dimensional reconstructions are built of.

2.5.1 Data window

The first tool that is important for the manual processing of a CT slice is the so-called "data window". The original CT raw data has a variety of 4096 different gray values (Hounsfield units), but the human eye is able to distinguish only around 35 of them. Most of the available information can not be recognized for manual image post-processing.

The data window is a technique to make the 12-bit gray value depth usable for human operators. The core idea of this technique is the definition of a spectrum of gray values that is most interesting

[1] In the first, slice wise evaluation of this data set, the medial meniscus seemed to move in anterior direction when the flexion angle was increased from 90 degrees to 120 degrees. In reality, it moved slightly to the posterior direction.

[2] With around 180 slices in each of the data sets, the challenge is the processing of 5400 single hi-res images

for the current action. All darker values are displayed as pure black and all lighter values are displayed as pure white. Thus, not 4096 values, but only several hundreds are presented to the user.

For the segmentation of the menisci, this data window contains all values between 50 and 200 Hounsfield units; thus the user is confronted with 150 gray values. This is still three times as much as a human can distinguish, but as it can be seen in figure 2.2, it is clearly more detailed than the original view.

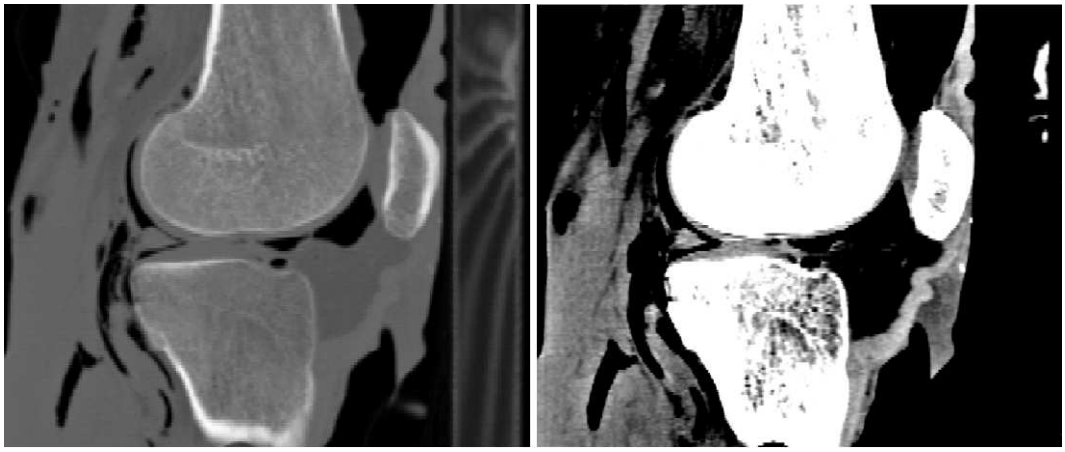


Figure 2.2: The knee joint in sagittal view: The original image (left) shows all (un-)necessary details, but lacks contrast for manual segmentation. The same image with a data window for menisci and cartilage provides more contrast (right).

The correct use of the data window seems to be the most important Post-processing tool for manual segmentation. Other techniques like the different filters were just used for (semi-)automatic segmentation.

2.5.2 Image filtering

To ease (semi-)automatic image processing, filtering algorithms can be used to improve the image quality. For CT data, two different filtering approaches prove to be helpful: Smoothing algorithms (lowpass) and edge detection algorithms (highpass).

In some situations of the manual segmentation, the human operator expects an edge from his/her anatomical knowledge, but cannot find its correspondent in the CT data set. In such cases, an edge detection filter like the "Sobel filter" can be useful, because it has the capability of finding edges independent from the data window.

The Sobel filter output for a certain voxel is the sum of the three local gray value gradients of the three dimensions. As the gradients have vectorial character and are orthogonal to each other, the three-dimensional Sobel filter can be described as:

$$sobel(x, y, z) = \sqrt{f'_x(x, y, z)^2 + f'_y(x, y, z)^2 + f'_z(x, y, z)^2}$$

The result of the Sobel filter is difficult to survey. Edges are detected, but they are either several voxels thick or relocated for several voxels (see figure 2.3). Thus, the filtered data set cannot be used for segmentation any more, as the segmentation error grows significantly.

However, it can be used for giving the human operator an impression of the edge topology, which helps him/her with the manual segmentation task.

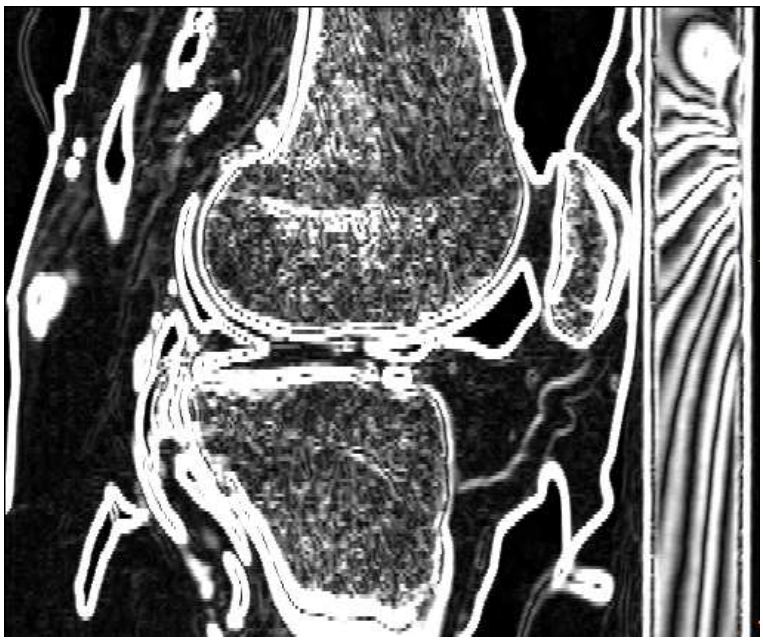


Figure 2.3: The Sobel filter detects all edges, but the result cannot be used for precise segmentation any more.

Like the output of edge detecting filters is helpful in only very special cases, a close look at the smoothing filters reveals different drawbacks, too. As smoothing is always done by equalizing the gray value of neighboring voxels, sharp edges get blurred. This can be observed after using the median filter and the Gauss filter, as it is done in figure 2.4. Although they approach this problem in different ways, both filters seem to reduce the size of the meniscus and the thickness of the cartilage by approximately 3 pixels in each slice. This poses a severe problem, since the deformation of the menisci (that should be measured) is of the same magnitude. With that behavior, the filter algorithms can add significant errors to the measured data. Thus using filtering algorithms for segmentation issues is still discussed controversially in literature.

2.5.3 Segmentation

The most exhaustive part of the image post-processing is called the segmentation. All voxels belonging to the same structure have to be marked, thus assigned to be part of the same component or material. If this task is carried out manually, a human operator colors the shape of an object with the computer mouse or with a graphical tablet. In figure 2.5, the femur is segmented and marked with white color. If the segmentation is made in all slices, i.e. if each voxel of the object is marked, then the 3D object can be reconstructed.

This process is commonly underestimated. The acquired CT data have a very high resolution and a very good contrast due to the enormous radiation that can be used. The high resolution is necessary to display all thin structures like the ligaments or the cartilages. Unfortunately, the segmentation of the high resolution slices is very time consuming. The

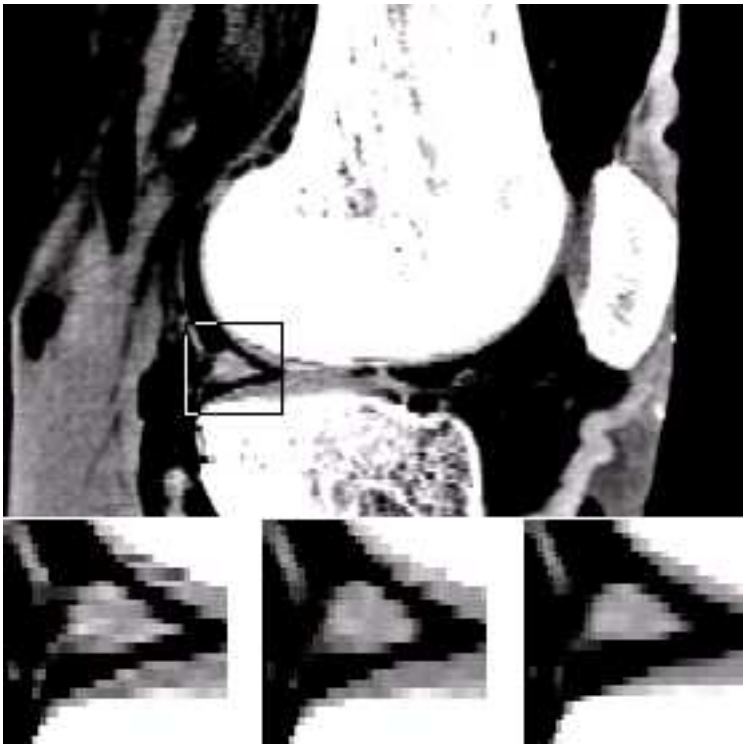


Figure 2.4: The meniscus in closeup: Unfiltered (left), median filter (center) and Gauss filter (right).

manual segmentation of a 512x512 slice will take 60 minutes - but as approximately 180 slices are acquired for each posture, the full segmentation takes around one month.

The support of automatic tools provides only a limited advantage. In spite of the high radiation, the cartilages of the femur and the tibia appear only a few voxels thick or are even invisible in some areas. On the one hand, the segmentation process has to be performed very precisely, because missing a single voxel reduces the thickness of the cartilage significantly. On the other hand, the segmentation process has to be performed in a flexible manner to complete the parts that are invisible due to the low contrast. These two targets cannot be achieved by automatic segmentation tools, so manual work is still essential.

However, manual segmentation alone is not sufficient. Due to the partial volume effect, the borders between two materials are generally not sharp edges, but transitions of the gray values. Even with a careful choice of the displayed data window, the human operator is not able to distinguish the subtle changes of the gray values, as there are still more than 100 of them. The consistent, objective choice of a border within the blurred edge is not possible for the human operator. Voxels of a certain gray value will randomly be assigned to one of the two materials. As this segmentation error occurs to a different extent by the segmentation of each slice, a stack of segmented slices will show rugged borders like in figure 2.5.

To avoid this kind of segmentation error, semi-automatic tools have to be used that draw the segmentation line precisely according to the absolute gray values. Additionally, the segmentation results can be corrected by controlling and editing them in the perpendicular cutting planes.

Figure 2.5: This femur was manually segmented in axial direction and shows rugged borders in sagittal and coronar views.

Several different segmentation algorithms are tested for segmenting the CT datasets, but only different region growing algorithms show acceptable results if they are used together with smoothing algorithms. Nevertheless, even the best results of automatic segmentation tools require major manual corrections.

In section 2.5, the potential error sources have been explained. The segmentation reduces the errors due to the partial volume effect, as automatic tools are oriented on a gray value threshold and are not confused by the blurred edges.

The 3D reconstruction that follows the segmentation process is not susceptible to the parallax errors, as the orientations of the objects can be determined exactly. Moreover, the limitations of the slice wise measurements is not given, as the surface area, the volume, 3D distances and 3D angles can be measured.

2.5.4 Measurement of translations and rotations

After the triangulation (either with the Delaunay triangulation [37] algorithm or with the marching cubes [55] algorithm), 30 different 3D representations of the knee are available. Each of these solids has its own coordinate system and scale, which has to be corrected for comparable measuring results. This is done by aligning all femurs.

In the second step, the movements of the knee components have to be determined. Therefore, the representation of the unloaded knee joint in full extension is chosen as the "neutral position". All measurements of translations, rotations or deformations are performed in relation to this neutral position.

The coordinate system is constructed around the femur of the neutral position by determining two prominent locations: The contact point with the lateral collateral ligament (LCL) and the contact point with the medial collateral ligament (MCL). These two points are considered to be located on the x-axis of the new coordinate system. The y-axis is set up to penetrate the femur in the middle of the median groove of the patellofemoral joint, and the z-axis is aligned with the longitudinal axis of the femur (figure 2.6).

The measurement itself is done by aligning each component of each reconstructed knee joint posture with its corresponding representation of the neutral position. Thus, the transformation matrix that leads to the alignment describes the movement precisely.

An essential task in these operations is the surface alignment of the femurs (to put them into the same coordinate system) and the surface alignment of all other knee components (to measure the transformation matrix). As these objects are reconstructed from different segmentations, they have slightly different shapes, different numbers of triangles and vertexes. Thus, the alignment is done by iteratively minimizing the euclidean distance between each vertex of one surface to the other surface. For the femur alignment, the largest root mean square error is 0.39mm, the smallest error is 0.14mm and the average error over all 29 surface alignments is 0.27mm, which is in the same dimension as the voxel size[3].

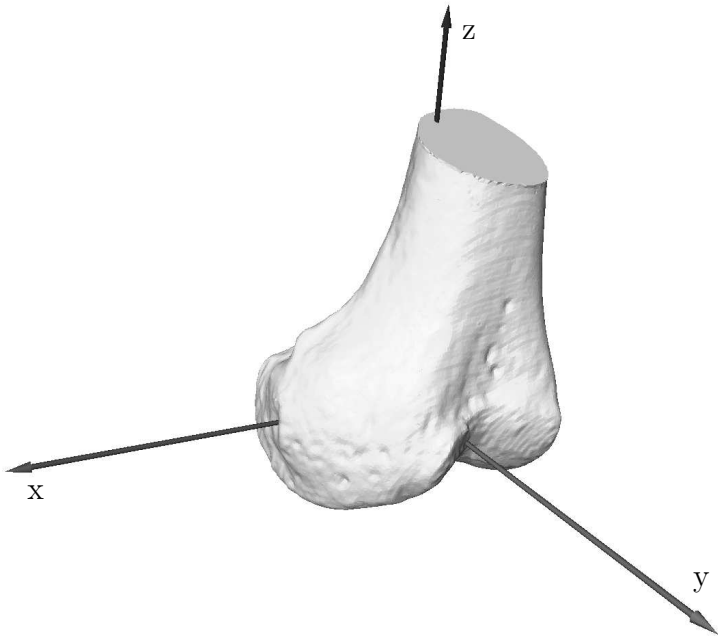


Figure 2.6: The coordinate system that is used throughout this work

Due to the more complex shape of the tibia, the alignment quality is slightly worse (best root mean square error: 0.21mm, worst: 0.49mm, average: 0.32mm). However, the macroscopic view reveals a quite precise match, as can be seen in figure 2.7.

From these transformation matrices, the *most exact* rotations and translations of every scanned pose can be calculated. This three-dimensional alignment is an iterative, converging method; if the same alignment is performed multiple times from different starting

[3]If the root mean square error was larger than 0.3mm, the segmentation was checked again, and segmentation errors were corrected

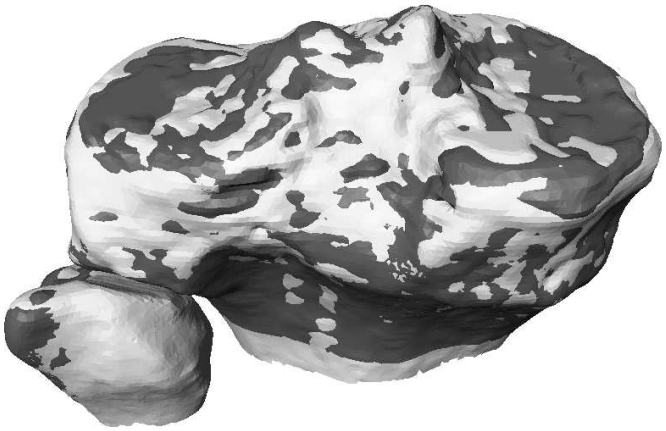


Figure 2.7: The tibia and fibula of 90 degrees flexion (dark) are matched to the tibia and fibula of zero degrees flexion (light). The root mean square error is 0.3mm and the spotted pattern comes from the slightly different segmentations.

points, the maximum difference between the resulting flexion angles is 0.5 degrees. The error of the manual, two-dimensional measurement (or estimation) is around ten times as much, which is not acceptable for the measurement of the varus/valgus or internal/external rotation of the knee joint.

2.5.5 Landmark positioning

The data acquisition process for rigid objects consists of the measurement of the change of the position (three cartesian coordinates) and the orientation (three Euler angles) for each of the 30 postures. The expenses for acquiring these six values with a reasonable precision are immense.

For deformable objects, the data acquisition is even more challenging:

- The contrast on the CT scans is moderate, so the segmentation is more challenging and the segmentation errors are more severe than for hard tissue.

- The measurement of the position and orientation, which is the main challenge for rigid objects, has to be performed for deformable objects, too. The surface alignment algorithm that is successfully used for rigid objects does not work reliably due to the deformations.

- For perfect measurements of deformations, the movement of each vertex has to be traced. As a single meniscus has more than 3000 vertexes, and as reliable automatic tools do not exist, this amount of data cannot be acquired within reasonable time.

As it is not possible to acquire the exact deformation, other techniques have to be applied. A common approach for three-dimensional image warping is the use of Thin Plate Splines [56]. This algorithm estimates the deformation of a complete three-dimensional object, if the movement of only few distinct points ("landmarks") is given. It is commonly used for facial animations (interpolating faces), medical applications (matching organs), in art science (merging different laser scans of the same statue) and in many other applications.

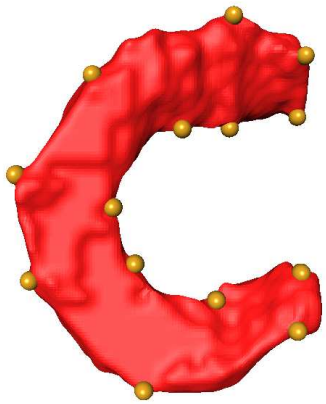


Figure 2.8: The landmarks are set on the lateral meniscus

In this case, the choice of the landmark position is not as easy. There are only three prominent points (described in 2.10) that are ideal for landmark positioning, which is not sufficient for a precise warping. Adding further landmarks can only be done inrelation to those three precise landmarks with the human sense of proportions. These landmarks may bear errors in their exact locations, but they provide information about the macroscopic shape. As different warping methods are tested, it turns out that more landmarks provide a better interpolation in spite of the uncertainties of their exact locations.

However, as the placement of the landmarks has to be performed manually by a human operator, huge numbers of landmarks do not lead to inaccuracies, but to fatal errors: If one landmark is forgotten in one of the 25 used representations, the Thin Plate Splines algorithm does not work at all. If the landmarks are placed in the right locations, but in an incorrect order, the warping causes disastrous results. Thus, the best result for meniscus deformation with human landmark placement is achieved with 13 landmarks placed on every meniscus.

2.6 Analysis of the acquired data

The data sets are acquired with modern equipment and extracted by experienced human operators. However, if the measured knee joint happens to show an unusual topology, a pathology or if a systematic error was overseen in the data acquisition process, the measured data is useless. Thus, a profound validation is performed and the precision of the new methods is evaluated.

	Medial anterior horn	Medial posterior horn	Lateral anterior horn	Lateral posterior horn
Excursion (literature)	6-8mm	2-5mm	9-18mm	8-12mm
Excursion (this work)	7mm	5mm	10mm	9mm

Table 2.1: Comparison with literature [1]: Absolute values of meniscal excursion

2.6.1 Validation

The knee joint does not show any signs of degeneration or pathology, neither in the CT scans, nor from the outer appearance. Nevertheless, the knee joint could still be abnormal, due to injuries that are invisible in the scans, individual variation or freezing. Thus, one has to check if this very knee joint is representative, which can be done in several ways.

The most interesting point is the validation of the femur movement relative to the tibia. This movement and shape of the interacting areas was first described by Weber and Weber [57] in 1836. As the movement of the femur is a combined rotational and translational movement, and as the surfaces are not simple geometrical shapes, the controversy about the correct description is still going on. However, most recent publications like Pinskerova [58] can be confirmed by the two-dimensional measurements that were performed for this thesis. [58] describes the femoral condyli by inscribing two circles of different diameters. The first, smaller circle is inscribed in the posterior part. The region, where the circle and the border of the condylus match is called the "flexion facet". Accordingly, a larger circle is inscribed in the distal and anterior part of the condylus and is called the "extension facet". If the same, slice wise measurement is performed on the knee joint of this project, both facets can be identified, and their radii and their behavior is in accordance with [58].

The next validation is carried out with the two-dimensional measurements of the menisci by Thompson et al. [1] and the posterior cruciate ligament by Nakagawa et al.[59].

Thompson [1] et al. chooses a "midcondylar, parasagittal plane along the tibial plateau" and measures the excursions of all four meniscal horns (both menisci, each anterior and posterior), as the knee joint is flexed from extension to complete flexion. His measurements are performed in the 2D slices of a MRI dataset. Five different knee joints are used for his work to achieve a higher reliability. The maximum and minimum of the measured values are compared to the results of this work in table 2.1. The excursion ratios of the horns of the menisci are also calculated and compared with the values of [1] in table 2.2. While the absolute values of the excursion are in the same interval as the values from [1], the ratios are at the lower level; the medial excursion ratio is even slightly lower than the minimum that is found by [1]. As only five knee joints are considered in Thompson's work and as the results of this work differ only slightly, the deviations are considered to be acceptable.

Another soft tissue deformation that is verified by literature is the posterior cruciate ligament (PCL). The most recent and detailed work was produced by Nakagawa et. al. [59], who takes six cadaver knees, twenty living Japanese knees and thirteen Caucasian knees into account. The living knees are scanned with an MRI scanner in different load situations, where flexion angles of up to 160 degrees are reached.

Although many knee joints have been scanned for evaluating different scanning resolutions in section 2.3.3 of this work, none of them came close to this extreme flexion angle. The best flexion angle was achieved with an 28-year old male volunteer, who finished at

	Med. excursion ratio: Post:Ant	Lat. excursion ratio: Post:Ant	Differential excursion ratio: Medial:Lateral
Minimum ratio	1:1.6	1:1.1	1:1.5
Maximum ratio	1:3.5	1:1.6	1:3.8
This work	1:1.5	1:1.2	1:1.6

Table 2.2: Comparison with literature [1]: Excursion ratios

Estimated angle	0°	30°	60°	90°	120°
Flexion in 2D	5°	35°	46°	95°	117°
Flexion in 3D	0.0°	32.0°	63.5°	98.3°	117.4°
Rotation in 3D	0.0°	0.3°	6.6°	11.3°	16.9°
Varus/Valgus in 3D	0.0°	0.9°	0.2°	−0.9°	−1.9°

Table 2.3: Measurements of the flexion angle differ up to 17 degrees

138 degrees of flexion. However, the lengths and angles of the PCL acquired by [59] are comparable to the values acquired in this work between 0 and 120 degrees of flexion. As the relative bone movement and the soft tissue movement of the measured knee joint behave according to literature, the knee joint was considered as representative.

2.6.2 Comparing 2D and 3D measurements

As mentioned in section 2.5, 2D measurements Bear system immanent errors. These errors can be quantified relative to the three-dimensional measurement. Table 2.3 shows the values that are estimated with three different techniques.

The 'estimated angle' is the value that is displayed by the scale of the wooden positioning device. This scale is known to be susceptible to the deformation of the wooden frame and to simplify the complicated flexion movement to one single rotational axis.

The second row of the table ("Flexion in 2D") shows the measurements taken from the sagittal, 2D CT slices as it is common in literature. The relevant error sources are the parallax error, and the partial volume effect.

The third to fifth row represent the flexion angle, the internal or external rotation and the varus/valgus rotation calculated from the transformation matrix after the surface alignment.

The three-dimensional measurement is supposed to provide the most precise value, as the mentioned errors are avoided or minimized. If this value is accepted to be the "real" flexion angle, the error of the two-dimensional measurement is up to 17 degrees, while the maximum error of the scale is less than 9 degrees.

2.7 Knee movement: Tibial view

The data acquisition, post-processing and the validation of the correctness and precision are fundamental to the animation of the knee joint. However, as the three-dimensional measurements of distances, angles and contact points are very intuitive and may bear new aspects for understanding the kinematics, these measurements are performed as well. Thus,

several interesting aspects are discovered by these measurements, that are not related to the interpolation.

2.7.1 Contact points

One of the new aspects that can be visualized in 3D is the position of the contact point between the tibia and the femur. Therefore, the minimal distance between the femur and the tibia (zero in many cases) is marked in the sagittal view and in the coronar view of the data set. The intersection of these markings was visible as a more or less elliptic region in the axial view (see figure 2.9, right). The center of gravity of these regions were estimated and marked.

One of the findings is the fact that the contact points between femur and tibia do not simply move forth an back when the knee joint is flexed, as the two-dimensional evaluation suggests it. Instead, when the contact points are reconstructed in three dimensions and marked on the tibia plateau, they follow a curved path, more or less parallel to the curved border of the tibia plateau as can be seen in figure 2.9 (left).

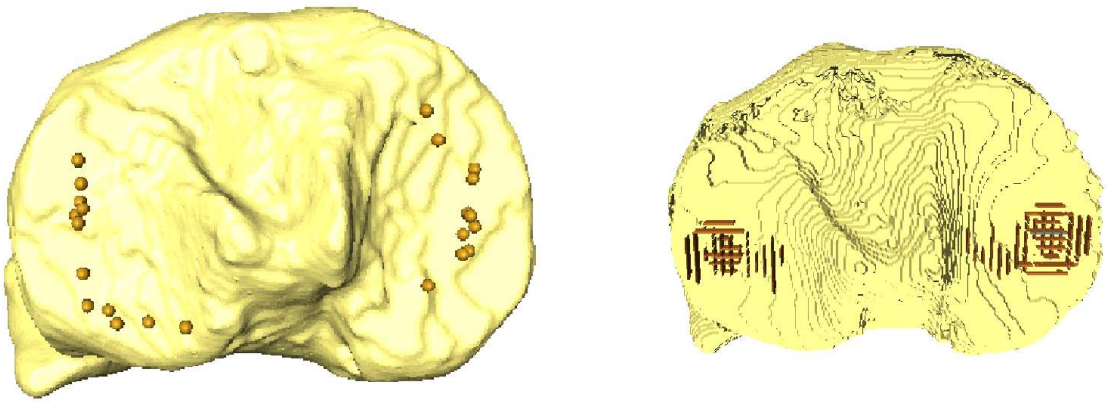


Figure 2.9: The contact points between femur and tibia are found by marking the smallest distances on sagittal and coronar slices (right). The center of the generated grid is calculated and is shown for representative flexions and rotations (left)

To confirm this interesting aspect, four different knee joints that have previously been scanned in two or three different poses are segmented, reconstructed and the contact points are determined as described above. The resulting contact points are merged together and compared to the previous results. Even this error prone method of merging the data from different knee joints showed a similarly curved path.

The evaluation of the contact points for the internal and external rotations, the varus and valgus rotations and the translation are performed in a similar manner, and it turns out that even those contact points are on (or at least very close to) the same path.

2.7.2 “Screw home”

One characteristic movement of the knee joint is the so-called “screw home” movement: Different authors in literature agree that the lower leg perfors an internal rotation if flexed

Unloaded flexion angle	0.0°	32.0°	63.5°	98.3°	117.4°
Applied external rotation	4.5°	16.2°	20.6°	23.8°	27.2°
Resulting flexion	−7.8°	27.7°	61.4°	97.8°	117.4°
Applied internal rotation	−13.4°	−13.7°	−13.4°	−8.4°	−7.7°
Resulting flexion	−3.7°	29.0°	63.3°	97.9°	117.5°

Table 2.4: 10Nm internal or external torque do not change the varus or valgus orientation in a significant way.

from full extension. Iwaki [60] describes a 5 degrees "screw home" movement between −5° and 5° of flexion, Johal [61] finds 10 degrees "screw home" between 0° and 20° and Li [62] finds around four degrees between 0° and 30°, where Wretenberg [63] finds only two degrees.

However, some authors do not agree with the screw-home theory. Between 30° and 60°, Wretenberg finds five degrees of *external* rotation, while Pinskerova [58] and Iwaki [60] hardly find any rotation at all between 5° and 45°. The source of these different results can be found in the different coordinate systems and in the error sources that are already described in section 2.5. All of these approaches set up arbitrary coordinate systems, measure arbitrary angles and "assume" that these angles correspond to the flexion angle of the tibia. As no standardized guideline for the setup of the coordinate system exists, the values of the different publications are sometimes misleading and hardly ever comparable, as already described (for a mechanical system) in [64].

To demonstrate the effect of incomparable values, the values of this work are compared to the values of Wretenberg's work, [63] who uses the contact points between femur and tibia for quantifying the screw-home measurements[4]. [63] connects the two contact points of the two condyli with a straight line and measures the angle of this line.

Using this approach with the acquired data set, an internal rotation is found between 0° and 30°, and an external rotation is found between 30° and 60°. The values of the rotations are even larger in this work (4.6° and 9.4°) than the results of [63] . However, when the movement of the lower leg is measured by the surface alignment (table 2.3) the results are completely different: Between 0° and 30° almost no rotation occurs, while with increasing flexion from 30° to 120°, the internal rotation grows uniformly to nearly 17°. Thus, no "screw home" can be found in 3D, although the 2D values correspond to literature. It seems that the complicated articulating surfaces forbid an easy conclusion from contact points to the orientation, as is tried in [63].

Most of the other measurements from literature suffer similar inaccuracies. The three-dimensional measurement is more reproducible and unequivocal.

2.7.3 Other DoF

As three-dimensional measurements show surprising results in the common flexion movement, the same measurement is used for varus/valgus and internal/external rotation scenarios.

Table 2.4 shows the smallest rotational freedom in full extension and the highest rotational freedom in flexion, which is consistent with the experience of orthopedics. Moreover,

[4]The same contact points have been acquired in section 2.7.1 in this work

Unloaded flexion angle	0.0°	32.0°	63.5°	98.3°	117.4°
Applied varus	3.5°	4.8°	3.9°	2.7°	3.5°
Resulting flexion	0.4°	33.0°	64.6°	99.5°	118.6°
Applied valgus	−2.0°	−1.2°	−1.4°	−3.5°	−5.7°
Resulting flexion	0.3°	33.1°	63.6°	98.6°	117.9°

Table 2.5: 15Nm varus or valgus stress do not significantly change the internal or external rotations.

Load	none	int.+val.	ext.+val.	int.+var.	ext.+var.	anterior
Flexion	32.0°	23.8°	19.5°	31.4°	25.7°	32.6°
Ext./int. rotation	0.3°	−7.8°	17.2°	−5.2°	4.5°	9.9°
Var./val.	0.9°	−1.5°	−1.4°	4.1°	5.0°	0.8°

Table 2.6: The knee joint in some very special load situations.

the table suggests that applying internal or external torque can influence the flexion angle. The flexion angle of the neutral posture (around zero degrees) is decreased by nearly eight degrees, but the influence vanishes in higher flexion angles. The internal rotation seems to have more influence on the flexion angle than the external rotation.

The same table can be set up for the varus and valgus loaded knee joint (table 2.5). Contrary to the internal and external rotation, the varus or valgus stress can be applied to a knee joint without causing significant changes in the flexion angle[5]. Similar to table 2.4, the largest freedom in varus/valgus direction occurs around the largest flexion angle.

The last data sets that are acquired from this knee joint are some special load cases, that are all scanned around 30 degrees of flexion: Combinations of 15Nm varus or valgus load together with 10Nm of internal or external torque. In addition to these four cases, an anterior translation with 100N is measured (table 2.6).

These values are quite interesting, because the combined forces lead to more varus/valgus movement, but to less internal and external rotation than the isolated forces. The anterior translation leads to nearly 10 degrees of internal rotation.

2.7.4 Results of the tibial view

Different forces in different directions are applied to the lower leg in a systematic way. Due to the internal knee joint mechanics and the geometry of the knee joint, an isolated force in a certain direction always leads to movements in the other DoF. However, those movements are either systematic or negligible.

The use of the three-dimensional measurements reveal some details that cannot be found by using the common two-dimensional techniques. If these results can be confirmed in further research projects, investigations should be made for an explanation from the mechanical properties of the knee joint.

[5] or the rotational angle

2.8 Knee movement: Femoral view

After the evaluation of the contact path on the tibia plateau in section 2.7.1, the contact points of the flexion are transferred to the femur surface, as can be seen in figure 2.10. Here, the points on each condylus are located on a common plane; each of those planes is inclined to the coronar plane of approximately 81 degrees and around 97 degrees (lateral) or 90 degrees (medial) to the axial plane.

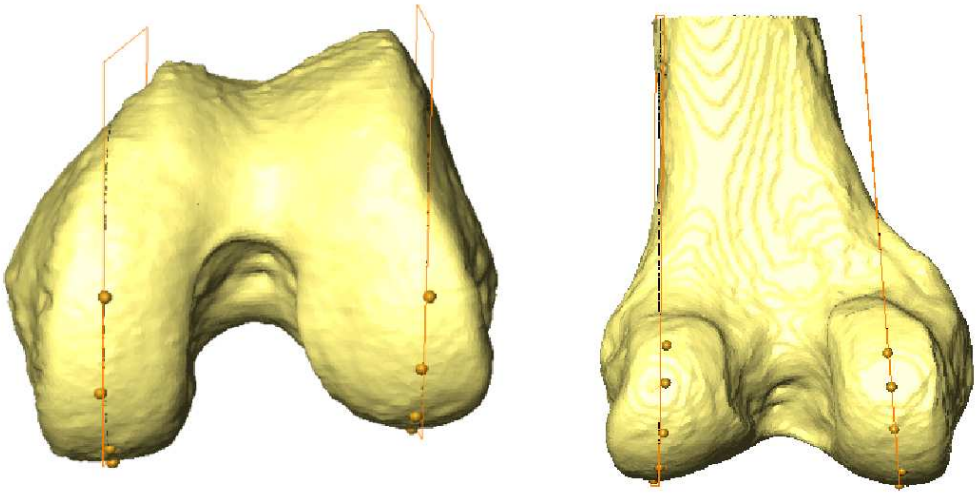


Figure 2.10: The cutting planes through the contact points. Here, the parallel perspective is used to stress that the planes on the right images are indeed perpendicular to the drawing plane.

One controversy in literature is the description of the femur condyli: While Bugnion [65] describes them as a helix, Weber [57] and Pinskerova [58] construct them as the arcs of two circles with different radii. However, it is a common agreement that the medial and the lateral condylus (or at least their sagittal slices) are generally different. This general difference cannot be found if the cutting planes are chosen according to the location of the contact points as proposed above. In fact, the lateral condylus is smaller, but the differences in curvature are minimal. As similar cutting planes were chosen for another femur, these results were confirmed.

With this geometry of the femur, the curved contact path on the tibia plateau can be explained. As the planes of the contact points are not parallel, the distance between two corresponding contact points change with the flexion angle. They reach their maximum distance in full extension, which decreases with growing flexion.

2.9 Knee movement: Patellar view

The patella is connected to the tibia by the patellar tendon and therefore follows its flexion movements. However, as the patella slides over the femoral condyli, and as the posterior part of the patella has a different geometry than the tibia plateau, the movement of the patella and the tibia differ.

Unloaded tibial flexion angle	0.0°	32.0°	63.5°	98.3°	117.4°
Patellar flexion angle	0.0°	29.4°	49.1°	70.3°	76.7°
Patellar int./ext. rotation	0.0°	1.5°	−2.8°	−1.6°	0.7°
Patellar varus/valgus	0.0°	0.1°	4.9°	3.3°	7.5°

Table 2.7: Patella movement in different knee flexion angles.

According to the tibial measurement, all six DoF of the patella are measured. Since there is not yet an established terminology to describe the patellar orientation in the three rotational DoF, the DoFs of the patella are defined according to the tibial orientations: The patellar flexion angle, the patellar varus/valgus and the patellar internal or external rotation.

As the patellar orientation of the unloaded, extended knee joint is defined as the neutral position, it has zero degrees of rotation in each of the three rotational DoF.

Table 2.7 shows the patella movement of the pure knee joint flexion without additional varus/valgus or internal/external rotation of the tibia. The corresponding values of the tibia are shown in table 2.3. The patellar movement is quite confusing, as neither the patellar varus/valgus rotation nor the internal/external rotation grows uniformly with the patellar or tibial flexion.

One part of the explanation of the odd patellar movement can be found in the femoral topology. As can be seen in figure 2.11, the contact area between the femur and the patella changes with the flexion angle. Between 60 and 120 degrees of tibial flexion, the patella moves over the distal part of the femoral cartilage that comes into contact with the tibia plateau, too. Between 0 and 60 degrees of flexion, the anterior part of the femoral cartilage is used, that is never in contact with the tibia plateau.

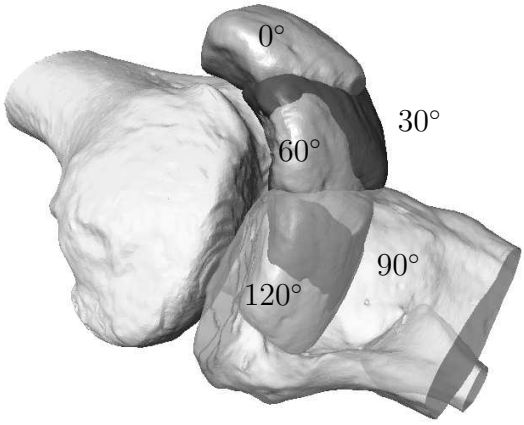


Figure 2.11: The movement of the patella with changing tibial flexion. The tibia of zero degrees of flexion is made transparent for a clearer view.

Average tibial flexion angle	0°	33°	64°	99°	117°
Average patellar flexion angle	1°	30°	52°	71°	78°
Patellar varus rotation at tibial varus rotation	1.9°	4.5°	7.0°	7.0°	10.9°
Patellar valgus rotation at tibial valgus rotation	−0.9°	0.3°	1.8°	2.0°	4.5°
Patellar varus/valgus freedom	2.8°	4.2°	5.2°	5.0°	6.4°
Patellar int./ext. rotation at tibial varus rotation	−2.4°	−3.6°	−3.3°	−2.6°	0.4°
Patellar int./ext. rotation at tibial valgus rotation	−0.7°	0.6°	−0.8°	−1.5°	1.8°
Patellar int./ext. freedom	1.7°	4.2°	2.5°	1.1°	1.4°

Table 2.8: The tibia was forced to varus/valgus rotation, and the orientation of the patella followed. No system can be found for the corresponding internal or external patellar rotation.

These two regions of the femoral condyli have very different topologies, and the most obvious difference is the curvature. These different topologies may explain why the patellar behavior changes around 60 degrees of tibial flexion.

The changes in the behavior can be seen in the other dimensions, too. Table 2.8 shows the behavior of the patella, if the tibia is varus/valgus stressed. The corresponding values for the tibia can be found in table 2.5. The patella follows the movement of the tibia and performs a patellar varus/valgus movement. Additionally to that movement, an internal or external patellar rotation occurs. The difference between the patellar varus and the patellar valgus rotation of the patella grows monotonously with the increasing flexion angle of the tibia, however no obvious description can be found for the erratic behavior of the third patellar DoF, the internal/external rotation. Unfortunately, this erratic error can not be neglected, because the patella rotates up to 4.2 degrees in this DoF.

Another measurement is carried out with the internal and external rotation of the tibia in different tibial flexion angles (the tibial values can again be seen in table 2.4). The patellar results of these measurements are shown in table 2.9. Most striking in the analysis of the patella orientation in those postures is the fact that the internal or external rotation of the tibia influences the internal or external rotation of the patella *less* than the varus/valgus movement of the patella.

Moreover, the internal tibial rotation does not necessarily have the opposite effect on the patella as the external tibial rotation. The clearest example of this fact can be seen around 30 degrees of flexion; both rotations (internal and external) of the tibia lead to an *internal* rotation of the patella. This situation is graphically depicted in figure 2.12.

2.9.1 Results of the patellar view

The patella movement is not as predictable as the tibia movement. The flexion angle of the patella more or less follows the tibia movement. This angle changes very fast in the beginning due to the small radius of the femoral cartilage, but slows down when the patella slides over the distal part of the femoral cartilage.

Average tibial flexion angle	−4°	30°	63°	98°	117°
Average patellar flexion angle	11°	28°	45°	68°	77°
Patellar int./ext. rotation at tibial int. rotation	−1.5°	−2.6°	−3.0°	−1.7°	2.2°
Patellar int./ext. rotation at tibial ext. rotation	−4.7°	−5.5°	−2.6°	−3.6°	−0.1°
Patellar int./ext. rotation without tibial int./ext. rotation	0.0°	1.5°	−2.8°	−1.6°	0.7°
Patella varus/valgus at tibial int. rotation	3.7°	0.1°	2.3°	2.2°	5.0°
Patella varus/valgus at tibial ext. rotation	0.3°	3.0°	4.2°	5.2°	7.7°
Patella varus/valgus without tibial int./ext. rotation	0.0°	0.1°	4.9°	3.3°	7.5°

Table 2.9: Patella movement in internal or external tibial rotation scenarios.

The varus movement of the tibia leads to a varus movement of the patella, and the tibial valgus movement leads to a patellar valgus movement. These patellar varus/valgus movements are small in smaller flexion angles and grow monotonously with the growing flexion angle.

The internal and external rotations of the tibia lead to changes of the patella orientation and so do the varus/valgus movements of the tibia. A systematic mechanism can be found in neither of the movements.

2.10 Soft tissue deformation

The relative bone movement can be described and validated by six DoF parameters; these parameters can be estimated with satisfying precision and reproducibility. Soft tissue like the menisci can be moved in those six DoF as well, but it also shows deformations. The surface matching method used for the measurement of the bone movement does not work with these deformed surfaces any more, and other kinds of measurement become quite imprecise, too. As the coordinates of the thirteen landmarks from 2.5.5 are neither intuitive, nor objective or reproducible, other approaches from literature are considered to describe the meniscus deformation.

Thompson et al. [1] chose a “midcondylar, parasagittal plane along the tibial plateau” for all measurements, like the one reconstructed in figure 2.13. This kind of measurement makes sense if the knee joint is extended (left image), because the most anterior and posterior parts of the meniscus can be found in these images. If the same slice is viewed at 120 degrees of flexion (right image), neither the most anterior nor the most posterior part of the meniscus can be found in this slice any more. Thus, this two-dimensional measurement determines the distance between two points: a prominent feature that can be identified very easily (in extension), and an erratic part of the meniscus that happens to be on the same slice of another CT scan (in flexion). There is no deep meaning in this kind of measurement - any measurement on any other slice gives equally important values.

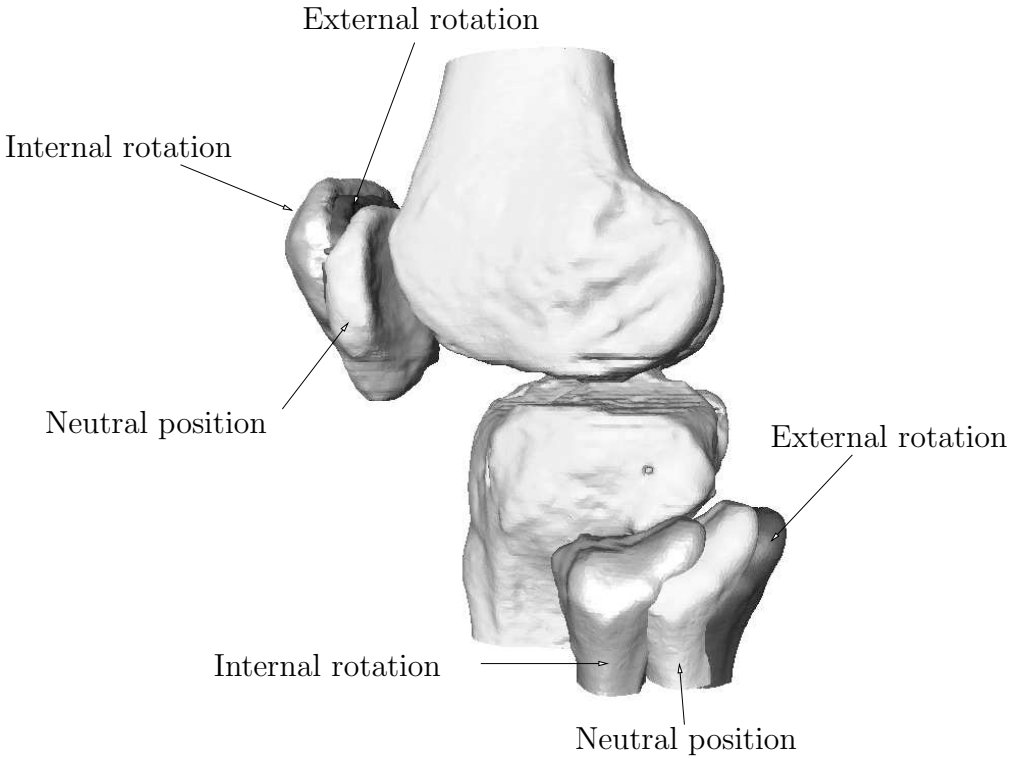


Figure 2.12: The movement of the patella around zero degrees of tibial flexion and internal/external rotation. The tibial movements were neglected to provide a better view.

As the midcondylar, sagittal cutting plane cuts the meniscus of the fully flexed knee in approximately 45 degrees, the errors due to partial volume effects and parallax errors can be significant. As other approaches in literature have similar disadvantages and produce equally weak results, another approach is tried here.

As explained in section 2.6, the measurement of the relative bone movements in three-dimensional space shows improvements in precision and objectivity compared to the measurement in the two-dimensional space. Due to this advantage, the meniscal movements are measured in three-dimensional space, too. However, measurements in 3D also have their pitfalls. An obvious approach of quantifying the deformation of a meniscus would be the observation of the most lateral or medial voxel, as it seems to wander like the contact point between femur and tibia (see figure 2.13). However, this is misleading. If measurements in three-dimensional space rely on canonical cutting planes like "axial", "sagittal" or "coronar", severe measurement errors occur, if one of these cutting planes is tilted for a few degrees. If measurements are done that way, the measurement errors turn out to be larger than the measured data, so the results show no trend at all. Thus, a kind of "coordinate system" has to be found that depends only on the features of the knee joint.

The two ends of the meniscus are unmovable, because they are connected to the tibia plateau. These two points (figure 2.14: point "A" and point "B") can be easily located on any three-dimensional reconstruction. Two other points that can readily be found are the contact points of the cruciate ligaments (ACL or PCL) to the tibia plateau (the contact point of the PCL is called "C" in figure 2.14). For measuring the lateral meniscus, "C" is

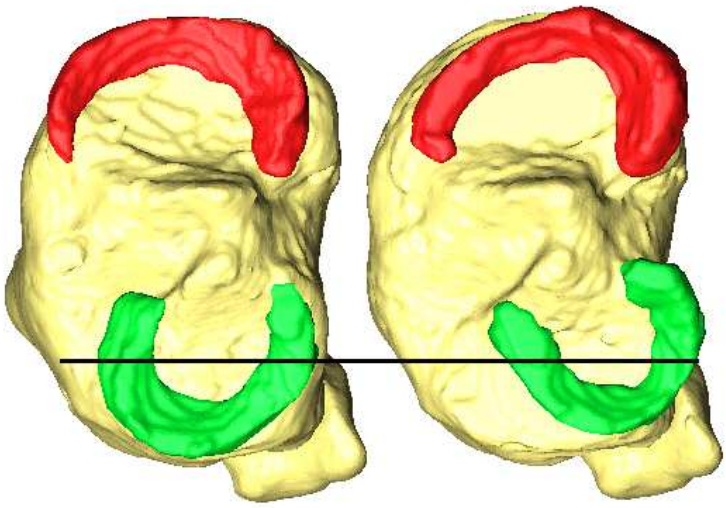


Figure 2.13: A sagittal slice is represented by the black line: While significant regions are found at 0 degrees of flexion (left), the same slice does not show prominent features at 120 degrees of flexion (right).

Distances	AC	BC	AB
Lateral meniscus	2.2cm	2.2cm	2.3cm
Medial meniscus	3.2cm	1.9cm	3.2cm

Table 2.10: Distances between the three relevant locations on the tibia plateau

set on the contact point of the posterior cruciate ligament PCL, for measuring the medial meniscus, the contact point of the anterior cruciate ligament ACL is chosen. These contact points on the tibia plateau can be located more easily in the reconstructions of other knee joint postures, if the distances are kept constant (table 2.10).

For estimating the deformation of the meniscus, a point on the meniscus surface is chosen, that is most distant from the contact point of the PCL (called "X" in figure 2.14). If the surfaces of the meniscus representations are smooth, the angles AXC and BXC are measured and found to be good representations of the meniscal deformation.

This technique of choosing point "X" has several advantages. As the distance between "C" and "X" is quite large, smaller errors by choosing "C" are not severe. Moreover, since the distance is so large, the most distant point of the meniscus surface can be estimated precisely. Nevertheless, the measurements are performed multiple times by multiple people, but the measuring error is negligible. The results of those measurements can be seen in table 2.11.

If additional points on the surface of the meniscus were considered, the description of the deformation might be more precise; however, no further points were found that are as reproducible able as the point called "X".

2.11 Conclusion

30 high resolution CT scans of a cadaver knee joint are performed, and the hard and soft tissue movements are measured according to the known guidelines from different publica-

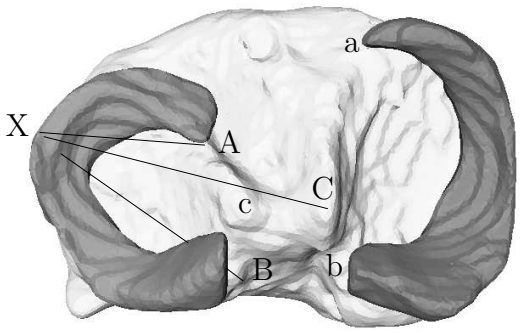


Figure 2.14: The meniscus is connected to the tibia at the points A and B. The posterior cruciate ligament is connected to the tibia at C. X is the point on the meniscus surface, that has the largest distance to C. Small letters are the measuring points for the medial meniscus.

Lateral meniscus	0.0°	32.0°	63.5°	98.3°	117.4°
Angle AXC	17°	17°	19°	25°	27°
Angle BXC	26°	26°	25°	23°	22°
Medial meniscus	0.0°	32.0°	63.5°	98.3°	117.4°
Angle AXC	43°	44°	38°	32°	32°
Angle BXC	20°	16°	16°	12°	12°

Table 2.11: The movement of the menisci described as changing angles

tions. As the values of the measured knee joint correspond to the values from literature, the knee joint is considered to be representative.

These scans are segmented and reconstructed to obtain three-dimensional objects. Three-dimensional measurements are carried out for the soft and hard tissue, which does not only deliver more precise results; it also creates the possibility of measuring distances and angles that cannot be measured in the two-dimensional slices. This again creates the opportunity to look at old controversies from a new point of view and to contribute to the discussions. Thus, the three-dimensional reconstruction is a promising approach to improve the understanding of the knee joint mechanics.

The advantages of the three-dimensional measurement are increasingly known. Thus, in recent publications this measurement method Has been gaining popularity. Still, the advantages of this method are far from state of the art. To the knowledge of the author, automatic surface alignment is not used in other publications, reducing the precision of those results.

Moreover, due to the difficult segmentation process, most publications use only a very limited number of different postures, DoF or knee joint components. The success of this method is strongly related to the availability of (semi-)automatic segmentation algorithms, which is still a challenge for computer science. Thus, one target of the future work is the development and the establishment of more powerful tools for segmentation and measurement.

Finally, as this technique is not yet established, standardized measurement guidelines are not yet available. Most of the publications introduce their own coordinate systems, allowing no direct comparison with other publications. With deformations, this is even worse. To the knowledge of the author, this work presents the first reproducible guideline for measuring meniscus deformation in 3D space.

Due to the high precision and the objectivity, the technique of performing three-dimensional measurements in reconstructed biological structures is complicated, but rewarding. Apart from the knee joint, most other joints can be researched this way; due to the very complicated setup and due to its importance, the research of the spine should be explicitly encouraged here.

3 Data interpolation and extrapolation

3.1 Introduction

Measurements are rarely performed as singular actions. In most cases, multiple measurements are performed over one or more dimensions. If enough measurements are performed in one dimension, (e.g. "time"), different statistical evaluations like median, average, standard deviation can be calculated or possible patterns or trends can be searched. A common example for such a measurement and the according evaluations is the measurement of a patient's body temperature over several days to estimate the development of the illness.

The measurement of the body temperature is generally performed every morning. For a more intuitive depiction of the development, the discrete temperature values of the last days are connected to a continuous graph over time. The connection of these points is called the "interpolation" and bears intrinsic errors: The body temperature changes during the day, which is not taken into consideration by simply connecting discrete, measured points.

Many different mathematical approaches are known for interpolation tasks, which provide different properties, like different grades of smoothness or different compromises between global and local interpolation capabilities. If details like the behavior in infinity, the standard deviation, the average curvature about the measured system are known, the mathematical interpolation method can be chosen accordingly. However, none of those methods is able to generate more information than the given, measured values (which are called the "control points"). Thus, generated values between control points are only estimates, so they are faulty in most cases. An optimization, that is done by a clever choice of the interpolating function and its parametrization is only able to minimize these errors.

The task of the interpolation is to estimate values between some given control points. The term "between" is easy to visualize in the one dimensional case and means that there are control points on both side of the value to estimate. If more dimensions are involved, the term "in the convex hull" is more applicable.

Opposed to the interpolation, the task of the extrapolation is to estimate values outside the convex hull of the given control points. In the one dimensional case, all given control points are either larger or smaller than the point to estimate. Commonly, extrapolation is done by simply retaining the last control point, by averaging several control points, by retaining the last gradient, or by using further derivates. As a consequence, extrapolation errors are generally more significant than interpolation errors. This effect is very obvious, if one tries to extrapolate a patient's body temperature[1].

Interpolation and extrapolation techniques are important for this work. At the data acquisition process in chapter 2, 30 three-dimensional representations of the same knee joint in different postures have been created. For displaying every possible posture between the given postures interpolation techniques have been chosen to calculate the positions and

[1] If the last derivate is chosen for the extrapolation task, boiling blood is predicted for patients with growing fever, freezing blood is predicted for patients with sinking body temperature and the influence of death to the body temperature is completely ignored

orientations of all knee joint components. However, this kind of interpolation is significantly more challenging than the simple example of a patient's body temperature.

A patient's temperature is measured every morning. Interpolation and extrapolation have to be performed in the one-dimensional

space called "time". Accordingly, for interpolating the knee joint movement in flexion direction, the one-dimensional space "flexion angle" would be sufficient. However, first differences become obvious here: The measurements of the body temperature are equidistant, i.e. precisely 24 hours passed between two measured values and deviations are negligible. In the case of the knee joint's flexion angles, the data acquisition process has been optimized for equidistant measurements, too. However, the precise analysis showed, that deviations from the equidistant measurements are *not* negligible. As many interpolation algorithms require equidistant measurements, the *scattered data* problem occurs, which is considered in section 3.2.

A further difference to the interpolation of the body temperature values is the use of more than one dimension. The knee joint has not only been measured in different flexion angles, but also in internal and external rotation and under varus and valgus load. In those dimensions, the scattering was even more severe. As figure 3.6 shows, the control points are arbitrarily grouped around every 30 degrees in flexion angle, which is frequently called "clustering".

Another crucial challenge for the interpolation algorithm is the uncommon nature of the values that have to be interpolated. The body temperature is a value in Euclidean space. Thus, mathematical operations can be used in their common context. This also applies for the locations of knee joint components, but it does not apply for their orientations, as orientations are *not* values of an Euclidean space, as is discussed in section 3.3. The last challenge is the real time requirement that is explained in section 3.4.

After all requisites for the interpolation algorithm are known, the state of the art of the inter- and extrapolation algorithms is explained in section 3.5. As it turns out, the "Weighted Averages Algorithm" is superior, thus it is parametrized for this project in section 3.6. The inter- and extrapolation results are shown in section 3.7 and concluded in section 3.8.

For the following explanations, the shape of the unloaded, extended knee joint are considered as the neutral position. The movements of the tibia, fibula and patella have been measured in relation to this neutral position. The corresponding transformations are calculated for every component and every load scenario.

3.2 Interpolation of scattered data

As common, polynomial interpolation algorithms (and many more) are based on linear interpolation algorithms, they inherit that these univariate algorithms need a kind of 'order' in their input values. Linear interpolation means to draw a connection between the input value number n and the input value number $n+1$. An *unsorted* cloud of 2D or 3D values cannot be interpolated that way, because for an arbitrary starting point number n, the corresponding $n+1$-th value can not be determined[2].

[2]In the common, one-dimensional case, an "unsorted" cloud does not exist.

When acquiring transformations from CT data sets, the establishment of an equal spacing of all orientations is tried, e.g. the flexion angles are increased in five 30 degree steps. Table 2.3, table 2.4 and table 2.5 show how bad this equal spacing works.

There are several ways to solve this problem. One of the possibilities is to ignore these inaccuracies, as is done in almost all[3] medical publications on human knee joints. Ignoring the measurement errors seems inappropriate for this work, since accepted errors would be larger than the values that should be displayed. Another - frequently suggested - approach is to heuristically find an analytical function that fits the measured points in an acceptable manner; the fine tuning can be done by the least squares method. Unfortunately, the knowledge about the knee joint movement is not yet consolidated to

a sufficient extent. The heuristic definition of this type of multidimensional function with such faint knowledge is difficult and leaves doubts (which might be the reason, why this approach is only frequently suggested but never published).

In this project, the problem with scattered data interpolation is solved by a careful choice of the multivariate, grid-less interpolation algorithm. As only few interpolation approaches are known that fulfill these requirements, the most prominent ones are discussed in section 3.5.

3.3 Interpolation of rotations

The interpolation of translations is a very common and well understood technique. However, the two most evident knee movements are the flexion and the internal/external rotation. Unfortunately, there is a fundamental difference between rotations and translations. One aspect where this difference opens up very intuitively is the repetition of one of those operations: A repeated translation will move an object further and further away, while a repeated rotation will place the object in its original position sooner or later[4]. As translations are defined in euclidean space, but rotations are not, they cannot be interpolated as straight forward.

As the object orientation can be described by the position of two points, one may be tempted to avoid this problem by interpolating the positions of two points and calculating the resulting rotations. This approach is misleading, as can be easily demonstrated by trying to rotate a staff 180 degrees around its center: Suppose, the two points, whose positions are to be interpolated are located at the opposite ends of the staff. Due to the 180 degrees rotation, the source posture of the staff looks like its target posture. In linear interpolation of the two end points, both points will take the direct path from their source locations to the target locations. This direct path is in the center line of the staff. Thus, when the user tries to interpolate between zero and 180 degrees, the staff will shrink to zero length and grow again to its real length.

There are multiple possible representations for rotations; the most common are Euler angles, rotation matrices and quaternions. The three representations differ in their algebra, and the algebras offer different possibilities to treat the problems that are mentioned above. Thus, all three approaches are considered in detail here.

[3] Measurement errors are generally not mentioned in these publications

[4] At least for some cases

3.3.1 Euler angles

The most common representation for rotations are Euler angles (roll, pitch and yaw). This representation is used on airplanes and ships and is supposed to be intuitive since their rotational axes represent the vehicle's geometrical axes. Euler angles have severe drawbacks: Apart from the singularity called "gimbal lock" which is explained by Dam et. al. [66] or by Barr et. al. [67], they are not suitable for interpolation. As the rotational axes are not independent from each other, the rotation around one often changes others, too.

In the two-dimensional case, the Euler angle algebra is easier to handle. Only one rotational axis, perpendicular to the 2D plane is defined. Rotations around this axis can be added and subtracted according to the common algebra in Euclidean spaces. This is done in most 2D approaches, e.g. most keyframe interpolations or facial animations. However, as the Knee Joint Simulator has more than one rotational axis, this simplification does not apply here.

3.3.2 Rotation matrices

Rotation matrices are 3×3 matrices, where the norm of each row or column vector is 1. The row or column vectors span an orthogonal space, which is rotated relative to the orthogonal space of the identity matrix. This representation is very common in computer graphics and robotics, since a cascade of rotations can easily be calculated by multiplying the matrices of each cascade layer.

The naive approach of interpolating between two rotations is the interpolation of the matrix component by component. However, it should be clear that these interpolated matrices do not fulfill the requirements of normed rows and columns in most cases; so this approach is misleading, too.

The reduction to simple rotations around an axis that is perpendicular to a two-dimensional plane, is possible with rotation matrices by freezing the values of the appropriate row and column and performing the rotation operations with the two remaining rows and columns. Contrary to the Euler angles, an interpolation is still not possible in this 2D mode, which proves that the different angular representations *do* not only have different algebras, but also different capabilities and weaknesses[5]

3.3.3 Quaternions

Euler angles and rotation matrices are not suitable for interpolation, since they do not offer adequate arithmetic operations. Euler angles do not have a defined multiplication operation, and rotation matrices do not have a proper addition operation. Thus, even the simple linear interpolation algorithm can hardly be implemented.

If even sophisticated interpolation algorithms are to be implemented, a representation for rotations has to be found that allows at least

basic algebraic operations. There are more examples for such spaces, for example [69], but the most common is the quaternion space that was defined by Hamilton [70]. As the quaternions are used in this work, a short introduction is given here.

[5] Eberly [68] published a way of linear interpolation with rotation matrices. However, it is neither intuitive nor performant. More sophisticated algorithms are not known.

Quaternions are a four-dimensional extension of the complex numbers with one real component w and three imaginary components x, y, z:

$$\boldsymbol{q} = [w, xi, yj, zk] = [w, \boldsymbol{p}] = [\cos\theta, \boldsymbol{p}\sin\theta] = e^{\boldsymbol{p}\theta}$$

Quaternions bear a certain redundancy, as four values represent only three degrees of freedom. If these four-dimensional vectors are normed ($|(w, xi, yj, zk)| = 1$) they can be seen as the three-dimensional surface of a four-dimensional sphere, allowing a basic algebra. The imaginary parts are the square root of -1, just like in the case of the complex numbers $i^2 = j^2 = k^2 = -1$, but the product of two imaginary parts results in the third one:

$$ij = -ji = k \qquad jk = -kj = i \qquad ki = -ik = j$$

The multiplication of quaternions is associative, but not commutative, just like the matrix multiplication. The multiplicative inverse is the same as the conjugate: $\boldsymbol{q^{-1}} = \bar{\boldsymbol{q}} = [w, -\boldsymbol{p}]$. The addition of two quaternions is associative and commutative and defined as: $\boldsymbol{q} + \boldsymbol{q'} = [(w + w'), (\boldsymbol{p} + \boldsymbol{p'})]$.

Two points that are located on opposite locations of the quaternion sphere represent the same rotation: $[w, xi, yj, zk] = [-w, -xi, -yj, -zk]$, which should be considered for software implementations, as it might cause problems for some operations that require the quaternions to be on the same hemisphere. One example of these functions is mentioned in section 3.5.3, where the distance between two points on the surface of the quaternion sphere is important.

This algebra has some special properties that might make the implementation of some algorithms impossible, but also allows calculations that are impossible in euclidean space. In opposite to the euclidean space, in quaternion space a triangle with three 90 degrees angles is possible - which makes the use of Pythagoras' theorem impossible. Another special characteristic is that the distance between two rotations is defined. As two rotations can be considered as points on the surface of the sphere, they set up a triangle with the center of the sphere. Thus, the angle of the triangle at the center point can be seen as the distance of two rotations.

The interpolation of two rotations can be performed quite easily on the quaternion sphere. If a line is drawn between two points on the surface, and if the line is also on the surface, every point on this line is an interpolated rotation. For achieving direct interpolation, the line has to be the shortest possible connection (great arc, with the same center point as the sphere). This kind of interpolation is called "SLERP" (Spherical Linear intERPolation):

$$SLERP(\boldsymbol{q}_1, \boldsymbol{q}_2, t) = \frac{\sin((1-t)\omega)}{\sin\omega}\boldsymbol{q}_1 + \frac{\sin(t\omega)}{\sin\omega}\boldsymbol{q}_2$$

where ω is the angle between the two quaternions and t is the interpolation factor between 0 and 1. The angle ω can be calculated as: $\omega = 2\arccos(w_1 \cdot w_2)$, where the "$\cdot$" is the common scalar product of vectors.

Figure 3.1 depicts a SLERP interpolation on a sphere (note that a three-dimensional sphere is shown here, but quaternions set up a *four*-dimensional sphere). It is clearly visible that the interpolation is not smooth at the data points. SLERP shares this drawback with the linear interpolation in euclidean space. The algebra of the quaternion space allows the calculation of Bézier splines, which are called SQUAD in quaternion space:

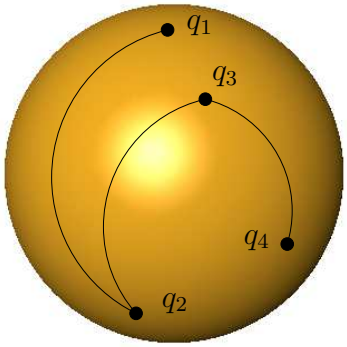


Figure 3.1: The SLERP algorithm causes unsmooth behavior and unexpected accelerations due to sudden changes of the direction at each intermediate control point q_2 and q_3.

$$SQUAD(\boldsymbol{q}_i, \boldsymbol{q}_{i+1}, \boldsymbol{s}_i, \boldsymbol{s}_{i+1}, h) = SLERP(SLERP(\boldsymbol{q}_i, \boldsymbol{q}_{i+1}, h), SLERP(\boldsymbol{s}_i, \boldsymbol{s}_{i+1}, h), 2h(1-h))$$

where

$$\boldsymbol{s}_i = \boldsymbol{q}_i \exp(-\frac{\log(\boldsymbol{q}_i^{-1}\boldsymbol{q}_{i+1}) + \log(\boldsymbol{q}_i^{-1}\boldsymbol{q}_{i-1})}{4})$$

If the equations of SLERP and SQUAD are analyzed, the correspondence to linear interpolation (for SLERP) or Bézier splines (for SQUAD) becomes evident. Like in euclidean space, SQUAD splines are constructed with the help of the linear algorithm. So, even sophisticated algorithms can be implemented in quaternion space. However, if this algebra is not sufficient, the curved space can be mapped into an exponential map, which is an Euclidean space again. Calculations can be made on this exponential map and the results can be mapped back to the quaternion sphere.

The mapping from the quaternion sphere is done by the logarithm function:

$$\log \boldsymbol{q} := [0, \theta \boldsymbol{p}]$$

This definition of the logarithm function has some characteristics of the logarithm function of the real case, for example: $\log[1, 0i, 0j, 0k] = [0, 0i, 0j, 0k]$. The logarithm maps a quaternion from the four-dimensional quaternion sphere to the three-dimensional euclidean space, because w is set to zero, and (xi, yj, zk) can be treated like a common, three-dimensional vector.

For mapping a three-dimensional vector back onto the quaternion sphere, the exponential function has to be used:

$$\exp \boldsymbol{q} := [\cos \theta, \sin \theta \boldsymbol{p}]$$

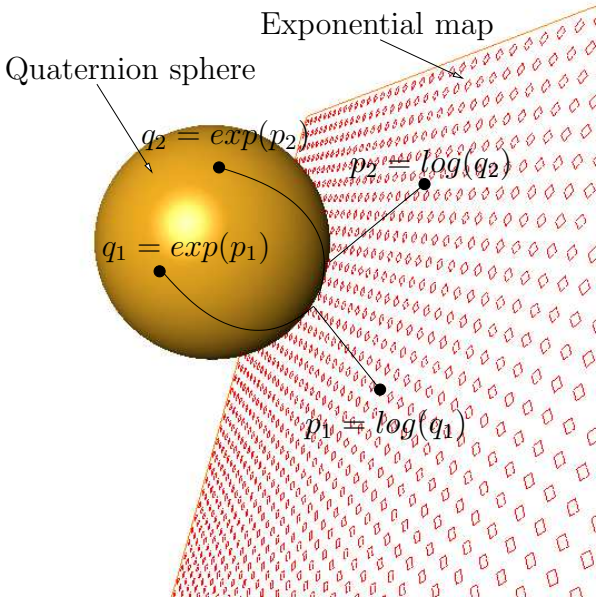


Figure 3.2: The quaternion space is a 4D sphere, and its exponential map is a 3D space. To allow an illustration, the dimensions of the sphere and the exponential map are reduced by one dimension.

Unfortunately, this mapping bears distortions and preserves only *distances through* and *directions from* the contact point. Moreover, the distances through the contact point have to be shorter than π .

The contact point between the exponential map and the quaternion sphere is always at $[1, 0, 0, 0]$, so if another tangent point $\boldsymbol{q}_t = [w, x, y, z]$ is needed, it is necessary to rotate the whole scenario around $\boldsymbol{q}^{-1} = [w, -x, -y, -z] = [-w, x, y, z]$ before applying the logarithm. After mapping the result of the euclidean calculations back to the quaternion sphere, it must be rotated again around $\boldsymbol{q} = [w, x, y, z]$.

By applying $\exp \boldsymbol{q}$ and $\log \boldsymbol{q}$, some rotations can be transferred from the quaternion space to euclidean space and back. Thus, some calculations can be performed in a very flexible manner. This mapping is necessary in section 3.6.4, where the algebraic possibilities of only one space are not sufficient.

3.4 Real time requirements

The knee joint simulator is an interactive tool for educational purposes. When the user moves the artificial shank, graphical feedback is expected immediately. This poses certain problems to today's hardware.

Even simple operations like zooming, rotating or translating require the recalculation of each vertex position, the recalculation of the visibility of each vertex (Bresenham's algorithm [71]) and the recalculation of the shading (Phong's algorithm [72]). As the three-dimensional reconstruction of a knee joint has more than 500,000 vertexes and more than one million triangles, even those simple tasks take around two seconds on high-end personal computers (Pentium IV, 1GB RAM DDR, NVidia Quadro 4).

Because the interpolation and the deformation (to come in chapter 4) require additional calculation power, the complexity of the objects is a severe problem and has to be reduced. If the number of triangles is reduced, their size increases

and the visual quality of curved surfaces decreases. Thus, a compromise has to be found that allows a fluent animation and precise geometries at the same time.

By reducing the number of triangles to different levels, the performance of the system could be estimated in different setups. By this heuristic approach, 50,000 triangles, resulting in 25,000 vertexes, are found to provide 25 frames per second (fps) on the hardware that is mentioned above.

3.5 Interpolations: State of the art

In the one-dimensional interpolation case, all control points are given as x_i and are assigned to their value v_i. All those points $v_i(x_i)$ can be intuitively connected to a curve in 2D space.

If the interpolation is done in two-dimensional space, all control points are given as (x_i, y_i), with their corresponding value v_i. All those points $v_i(x_i, y_i)$ can be intuitively connected to a curved surface in 3D space.

In this project, at least three-dimensional interpolation is necessary, because the data points are acquired according to the tibia flexion, rotation and varus/valgus orientation. An intuitive description like the curve or the surface in the one- or two-dimensional case is not possible here. To simplify the explanations, the interpolation algorithms will be considered for the two-dimensional case.

As shown above, the interpolation algorithm has to meet a series of constraints. Triangulation, Radial Basis Functions and Weighted Averages are the most promising candidates to meet these constraints and are presented in this section.

3.5.1 Triangulation

This kind of interpolation is a two-step approach: First, the given control points (x_i, y_i) are connected to triangles (in the two-dimensional case), then the value v_i for the interpolation is searched on the triangulated surface. The triangulation can be done in many different ways, but different triangulation algorithms show different strengths and weaknesses. Commonly, the Delaunay triangulation, is used for interpolation (and three-dimensional object reconstruction, see section 2.5.4). Its strength is that the calculated triangles have the largest possible minimal angle. Degenerated triangles with zero degree angles that are known from other triangulation algorithms do not occur here[6].

For calculating the Delaunay triangulation, the Voronoi diagram of the given dataset has to be calculated. For every given point $(x_i, y_i) \quad i = 1...N$, a region is defined that is closer to this point than to any other point $(x_j, y_j) \quad j \neq i$.

In figure 3.3, the given points are depicted as black circles, whereas the borders between those points are depicted as solid lines. The Delaunay triangulation is performed by connecting *those* points that share a common border in the Voronoi diagram. Thus, the triangulation is depicted with the dotted lines.

[6]Moreover, the circle that circumscribes any constructed triangle does not contain another point of the dataset, which is a nice feature for interpolation.

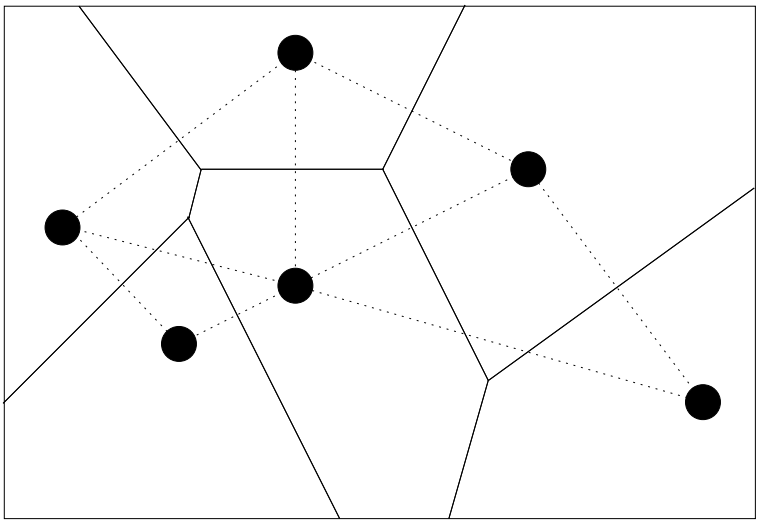


Figure 3.3: The solid lines are the Voronoi diagram. Connecting the points that share common borders, gives the Delaunay triangulation (dotted lines).

The triangulation can be performed offline, as the given control points do not change any more. Calculating the interpolated value v_i of a certain point (x_i, y_i) on the triangulated surface is the second part of the algorithm and has to be performed in runtime. Therefore, an algorithm has to search the triangle of the surface which contains the requested point (x_i, y_i). If this triangle is found, a linear interpolation algorithm estimates the requested value v_i from the value of the three corners of the triangle. Thus, the interpolation with triangulation is very

local and has discontinue derivates on the borders of the triangles[7].

The capability of offline preprocessing (triangulation) and the high speed of this algorithm are very welcome characteristics in the application for the knee joint simulator. It is already transferred to the general n-dimensional case and to the quaternion space, thus it seems to be a good approach.

On the other hand, the local character turns out to be a large drawback if only few[8] control points are given. Moreover, an extrapolation is not possible with this algorithm, which poses a problem if the knee joint should be displayed in postures outside the convex hull of the given control points. As the same restrictions apply for a similar approach, the Natural Neighbor Interpolation [73], this technique is not described here.

3.5.2 Radial basis functions

A radial basis function calculates a corresponding value $v(x, y)$ for each point (x, y) by combining an univariate basis function $\phi(\sqrt{(x_k - x)^2 + (y_k - y)^2})$ and a low order polynomial

[7] The rarely used Clogh-Tocher Method is an extension of the triangulation that uses cubic interpolation instead of linear interpolation for the calculations on the triangles to overcome these discontinuities.

[8] 25 control points are available for the knee joint animation

	ϕ	Constraint	Parameter
Thin Plate Splines	$r^{2a}\log(r^2)$	$a \in \mathbb{N}$	$M > l$
Polynomials	r^a	$a > 0, a \notin 2\mathbb{N}$	$M > \frac{a}{2}$
Gauss	$\exp(r^{-2})$		$M \geq 0$
Multiquadrics	$(1+r^2)^{\frac{a}{2}}$	$a > 0$, $a \notin 2\mathbb{N}$	$M \geq 0$

Table 3.1: Some known radial basis functions ($r = \sqrt{(x-x_i)^2 + (y-y_i)^2}$).

function $g(x, y)$[9]:

$$v(x,y) = \sum_{j=1}^{N} a_j g(x,y) + \sum_{k=1}^{M} m_k \phi(\sqrt{(x_k - x)^2 + (y_k - y)^2}) \tag{3.1}$$

The interpolation capabilities of this equation are obvious, because for every coordinate (x, y), there is a result from this equation. If the RBF ϕ is continuous, the summation of the different RBFs is also continuous, thus the result of the equation is a continuous surface.

The quality of the interpolation depends on the choice of the radial basis function. This choice seems easy, since almost every continuous function can be used as a RBF, but in practice problems occur. Because RBFs are very new interpolation techniques, several characteristics are not yet known in detail for every candidate. Thus, for most applications the mathematical methods from table 3.1 are used.

For successful calculation, the function $\phi(\sqrt{(x_k - x)^2 + (y_k - y)^2})$, the function $g(x, y)$ and all M control points (x_k, y_k) have to be given for this equation. Solving the equation means finding all N parameters a_j and all M parameters m_k. As there are $N + M$ linear equation systems, the corresponding matrices can be set up, as shown in section 4.4.2.

The advantages of the radial basis functions are the simple extension possibilities to three or more dimensions and the preprocessing stage of solving the linear equation system, which is desirable for this thesis as the online calculation is fast. Moreover, the interpolation is quite global and extrapolation is possible, too. In this work, it turned out that the "Polynomials" tend to overshoot, while the "Thin Plate Splines" perform quite nicely. Unfortunately, the transfer to the quaternion space for interpolating the rotations is not successful, due to the incompatible metrics.

3.5.3 Weighted averages

The weighted averages method is also known as the "inverse distance weighted method" or "Shepard methods" and has the common form:

$$v(x,y) = \sum_{i=1}^{N} m_i(x,y) v_i = \sum_{i=1}^{N} \frac{f_i(x,y)}{\sum_{j=1}^{N} f_j(x,y)} v_i \tag{3.2}$$

The function $m_i(x, y)$ is the *weight* or influence that the given control point (x_i, y_i) has on the arbitrary point (x, y) that is to be calculated. One requirement for the sum of these weights is the norming: $\sum_{i=1}^{N} m_i(x, y) = 1$. If the sum is larger or smaller, the interpolating

[9]This polynomial is not necessary for all applications

function would not meet the given data points. To simplify the choice of $m_i(x, y)$, any interpolating function $f_i(x, y)$ is allowed, if it is normed afterwards (see equation (3.2)).

Shepard originally used the reciprocal distance:

$$f_i(x, y) = \frac{1}{distance((x, y), (x_i, y_i))} = \frac{1}{\sqrt{(x - x_i)^2 + (y - y_i)^2}}$$

The smoothness of the interpolation can be influenced by using the l-th power of $f_i(x, y)$. The complete term of the original Shepard interpolation (together with equation 3.2) looks like:

$$v(x, y) = \sum_{i=1}^{N} \frac{h_i(x, y)}{\sum_{j=1}^{N} f_j(x, y)} = \sum_{i=1}^{N} \frac{\frac{1}{(\sqrt{(x-x_i)^2+(y-y_i)^2})^{l_i}}}{\sum_{j=1}^{N} \frac{1}{(\sqrt{(x-x_j)^2+(y-y_j)^2})^{l_j}}} \tag{3.3}$$

The euclidean distance $\sqrt{(x - x_i)^2 + (y - y_i)^2}$ is zero if the value $v(x, y)$ of a given data point (x_i, y_i) is to be calculated, so the numerator and the denominator of the Shepard function approach infinity at this point. However, if numerical problems are omitted, it can be shown that the resulting interpolated surface is continuous.

The choice of the parameter l_i has to be chosen according to the desired properties of the interpolation for the point (x_i, y_i). A general disadvantage is that the first partial derivates are zero at each given data point and the interpolated surface shows "flat spots". If large values like $l_i = 10$ are chosen, these flat spots are extended to huge "terraces", so the constructed surface consists of flat plateaus that are connected by steep walls to obtain a continuous surface (see figure 3.4).

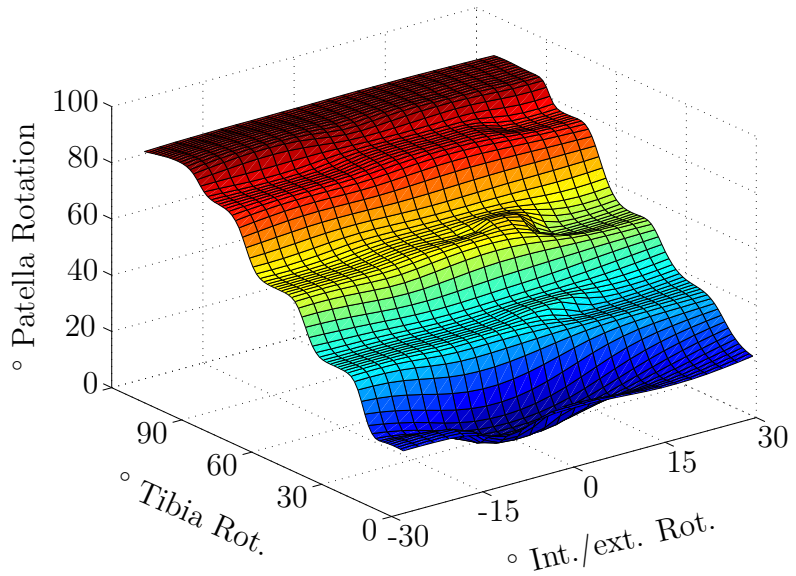


Figure 3.4: A Shepard function with exponent $l_i = 4$: Large exponents l_i cause "terracing" behavior. This surface shows the patella orientation depending on the tibial flexion and the tibial internal/external rotation at 0 degrees varus/valgus.

Interpolations with small values like $l_i = 0.5$ show the opposite behavior: The major part of the surface is found around the average of all given data points, only the data points themselves peak out of this surface, as can be seen in figure 3.5.

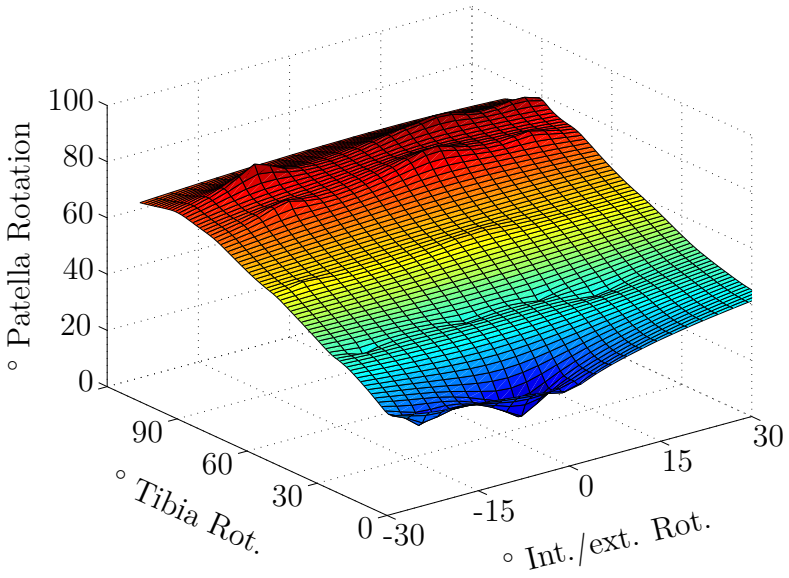


Figure 3.5: A Shepard function with exponent $l_i = 1$: Small exponents cause peaks out of the averaged surface.

The principal idea of the "weighted averages" is that distant control points should have less influence on the calculation of the interpolated value than close control points. This leads to a problem, if the data points are not uniformly distributed, but grouped in so-called "clusters". In the example of this project, there are basically five data clusters: around 0°, 30°, 60°, 90° and 120° of knee joint flexion. In each of these clusters, five data points are packed closely together, while each of these clusters is quite distant to the next. Figure 3.6 depicts this situation: The data points (circles) are gathered around five flexion angles (light gray planes). Some of those data points have a slightly higher or lower flexion angle, so they are located in front or behind this gray plane (black or dark gray color). The two-dimensional expansion of the gray planes depict the dimensions for the varus/valgus movement and the internal/external rotation.

Suppose, the value $v(x, y)$ of an arbitrary point (x, y) that happens to be located close to one of the planes has to be calculated. The distances to the five control points that are located around this plane are relatively short compared to the distances to all other data points. Thus, the interpolated value practically only depends on those five data points that are gathered around the plane. If only these five data points have to be taken into account, the weighted averages algorithm works perfect: The closest of those five control points is *significantly* closer than the other four control points. Thus, this control point has *significantly* more influence on the calculated result than the other points.

If the value $v(x, y)$ of an arbitrary point (x, y) is calculated, that is located far away from the planes, e.g. half way between two planes, problems occur. Suppose, (x, y) is located around 45° of flexion (see figure 3.7). The distances to each of the given control points of the 30° plane or the 60° plane are approximately the same. Even if (x, y) is moved to

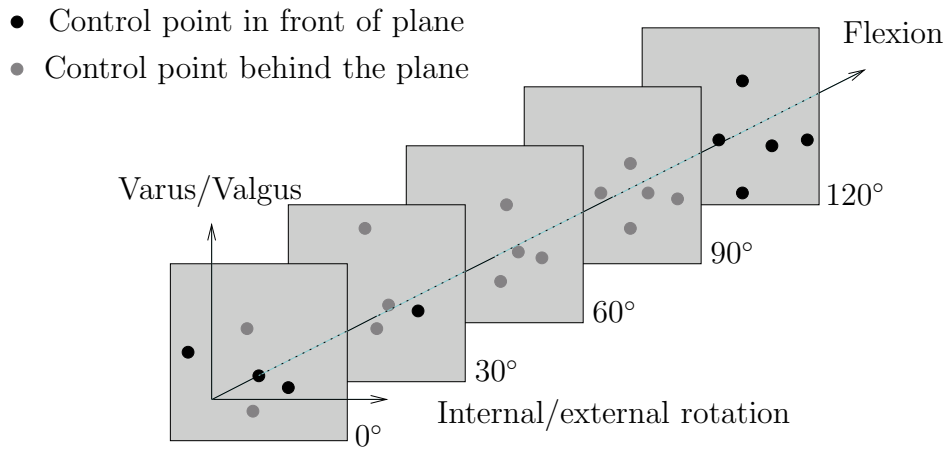


Figure 3.6: Clusters of data points are located around the gray planes representing discrete flexion angles.

locations with more or less internal or external rotation, the relation of the distances does not change *significantly*.

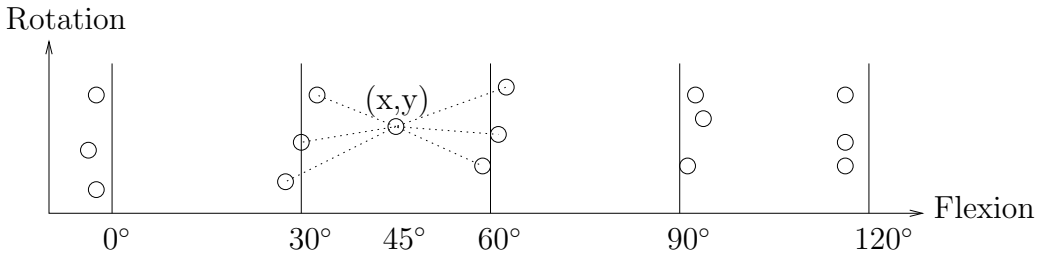


Figure 3.7: The given control points around 60 and 30 degrees of flexion all have approximately the same influence on the value $v(x, y)$. Changing the rotation for (x, y) has hardly any influence on $v(x, y)$.

The result of this effect is that the interpolated value $v(x, y)$ hardly changes around 15°, 45°, 75° or 105°, no matter what internal or external rotation is applied (the same applies to the varus/valgus movement). The reason for this is, that the flexion angle is measured in 30 degree steps (which results in large distances between the control points), but the other rotations are measured in much smaller steps (which results in small distances between the control points)[10].

To compensate this clustering, the rotational dimension is scaled by a factor 2 and the varus/valgus dimension is scaled by a factor of 6 for reasonable calculation of the distance. It is noted here, that this approach is suitable for this work, since the clustering is well known and in a uniform manner. Irregular clustered data cannot be treated this way.

Thus, the clustering problem could be avoided in this work. In addition, the weighted averages have plenty other advantages, for example the natural extrapolation or the easy extension to multiple dimensions. Even the transfer to quaternion space can easily be done, as the quaternion distance metric can be used. The interpolation is global and thus showed plausible behavior even with few given control points.

[10]The typical internal/external rotation is 15 degrees (half of the stepping of the flexion angle), while the typical varus/valgus angle is around five degrees (one sixth of the stepping of the flexion angle)

Although the calculation is a one-step approach, and an off-line optimization is not possible (as it was possible with the radial basis functions or the triangulation), it is computationally inexpensive and performs very fast. As there are quite a lot of ways to calculate weighted averages, and as alternative approaches can influence the characteristic of the animation, a few different methods are considered for this project.

Weight calculations

There are many different possibilities to calculate the weights. Shepard's original suggestion from 1968 is still very popular, however many approaches try to improve its properties. The most recent approaches from Robert Renka [74] [75] [76] use different weight calculations which try to localize the Shepard algorithm by considering only the nearest n data points. As these algorithms are published as Fortran77 code on the ACM algorithm collection [77], they are compared with the original Shepard's proposal. It turns out that Renka's approaches are optimized for larger data sets of several hundred control points, where the use of $n = 25$ control points is considered to be "local". Thus, these approaches do not apply here.

Other approaches like a more local interpolation, damping or different weight functions are published in [78] of [79], not showing significant advantages, but just add a computational overhead in this work.

3.6 Parameterizations

The evaluation of the method with different scaling factors and exponential parameters for the Shepard function is carried out by interpolating or extrapolating a knee joint posture that is known from the data acquisition process but taken out of the given data set. After the interpolation or the extrapolation, the estimated orientation of the patella is compared with the true orientation of the patella.

This evaluation process is performed with the patellar *orientation* only and *not* with the patellar *position*. The interpolation and extrapolation of the *positions* do not require the quaternion calculation, so they can be done with more sophisticated algorithms than the Shepard function. Thus, the parametrization is optimized for the rotation inter- and extrapolation.

The optimizations of the parameters are performed by comparing the average interpolation and extrapolation errors and the standard deviations. Due to the challenging data acquisition (see chapter 2), only very few interpolation or extrapolation results can be compared to acquired data. Thus, statistical methods like the "standard deviation" are not without doubt.

3.6.1 Parametrization of the extrapolation

For estimating the extrapolation capabilities, all 25 postures of the data set are used as input for the Shepard Algorithm. The desired output were those four postures acquired by a varus/valgus load combined with an internal/external rotation (see section 2.4). These four values are clearly outside the convex hull of the given data points and thus give rise to an extrapolation problem.

Average error	No scaling	1-2-6-scaling	1-4-12-scaling
$l_i = 4$	0.089	0.090	0.104
$l_i = 3$	0.084	0.089	0.106
$l_i = 2$	0.084	0.103	0.132
$l_i = 1$	0.133	0.161	0.189

Table 3.2: Average extrapolation error for different exponents and scaling factors

Standard deviation	No scaling	1-2-6-scaling	1-4-12-scaling
$l_i = 4$	0.018	0.016	0.025
$l_i = 3$	0.021	0.019	0.035
$l_i = 2$	0.025	0.023	0.036
$l_i = 1$	0.016	0.015	0.019

Table 3.3: Standard deviation of the extrapolation error for different exponents and scalings

Three different scaling methods for the different dimensions were examined:

- No scaling
- Double internal/external rotation and six times varus/valgus rotation (1-2-6)
- Four times internal/external rotation and twelve times varus/valgus rotation (1-4-12)

For each of the topologies with different scalings, different exponents l_i can be evaluated. The calculated orientation of the patella is compared to the true orientation from the data acquisition process. To get an objective quality measure, the distance between those two quaternions is calculated. This distance correlates with the extrapolation error. As four different extrapolated points are taken into account, the average error and the standard deviation are calculated (table 3.2).

The extrapolation error seems to be reduced by an increasing exponent l_i, but the scaling seems to have a negative influence. However, the standard deviation has to be taken into account, too (table 3.3).

Table 3.3 needs a closer look, as its low standard deviation is striking. For this explanation, the behavior of the Shepard interpolation with low exponents has to be considered, as can be seen in figure 3.5: The majority of the interpolated values are very close to the arithmetic average, only the given data points peak out. This explains how certain parameterizations have a low standard deviation but also a large average error in table 3.2.

Apart from the phenomena for $l_i = 1$, the standard deviation decreases with increasing exponent, and the 1-2-6-scaling gives the best standard deviation.

3.6.2 Parametrization of the interpolation

For estimating the extrapolation error, four knee joint postures are known that are outside the convex hull of the 25 given knee joint postures. Thus, four extrapolations can be compared with corresponding real postures.

Average error	No scaling	1-2-6-scaling	1-4-12-scaling
$l_i = 4$	0.099	0.063	0.047
$l_i = 3$	0.090	0.066	0.059
$l_i = 2$	0.084	0.077	0.083
$l_i = 1$	0.099	0.107	0.118

Table 3.4: Average interpolation error for different exponents and scalings

Standard deviation	No scaling	1-2-6-scaling	1-4-12-scaling
$l_i = 4$	0.021	0.032	0.024
$l_i = 3$	0.019	0.021	0.009
$l_i = 2$	0.011	0.009	0.018
$l_i = 1$	0.033	0.040	0.049

Table 3.5: Standard deviation of the interpolation error for different exponents and scalings

Unfortunately, there are no additional scans for the estimation of the *interpolation* error. The only possible approach to giving an estimation about the interpolation capabilities is to take one posture out of the dataset and perform the interpolation with only 24 data points. This bears some errors, because the 24-point interpolation will bring different results than the 25-point interpolation. Due to the averaging character of the Shepard algorithm, this error has its maximum at the location where the control point is missing. Thus, the error estimation with only 24 given data points is a reasonable error *over*estimation.

Not all of the 25 given data points can be used for estimating the interpolation error. If the point in question is on the surface of the convex hull, its removal from the data set will cause a shrinking of the convex hull. Thus, the determination of the point would be an extrapolation.

The unloaded knee joints postures of 30, 60 and 90 degrees of flexion are found to be clearly in the convex hull of the data points. Thus, the calculated orientations of these three control points are used for the estimation of the interpolation error by calculating their distance to the real orientation on the quaternion sphere. From the three resulting distances, the mean value and the standard deviation are calculated again (see table 3.4 and table 3.5).

According to table 3.4, the scaling seems to have a positive effect for high exponents like $l_i = 4$ or $l_i = 3$, but no clear effect for $l_i = 2$ and even a negative effect for $l_i = 1$. The lowest standard deviation occurs with the highest exponent and the highest scaling factors. According to the estimation of the extrapolation error, the standard deviation of the interpolation has to be considered, too.

The standard deviation of the interpolation error shows a non-uniform behavior, however the biggest and smallest exponents $l_i = 4$ and $l_i = 1$ show the worst behavior. The evaluation of the standard deviation with different scalings does not lead to clear conclusions.

3.6.3 Overall parametrization

As the same algorithm is used for interpolation and extrapolation, the parameters have to be chosen to fulfill both tasks with adequate inter- and extrapolation quality. As the

evaluation of the inter- and extrapolation led to different results, a compromise has to be found for the four targets:

- *Large exponent* l_i and *large scaling factor* for reducing the average interpolation error.
- *Medium exponent* l_i and *any scaling factor* for reducing the standard deviation of the interpolation error.
- *Large exponent* l_i, but *low scaling factor* for reducing the average extrapolation error.
- *Large exponent* l_i and *medium scaling factor* for reducing the standard deviation of the extrapolation.

The compromise of those requirements is clearly a large exponent ($l_i = 3$ or $l_i = 4$). The scaling is a bit more challenging, since no clear preference can be determined.

However, if the curves of the different inter- and extrapolations are visualized, the effect of the different scaling methods becomes clear. In figure 3.8 the patellar flexion angle is shown as a function of the tibial flexion and the tibial internal/external rotation with zero degrees varus/valgus. No scaling at all (top row) leads to the "clustering effect", where the data clusters around 0°, 30°, 60°, 90° and 120° cause a vivid local interpolation behavior, but hardly influence the flexion angles between those clusters.

At different flexion angles, tibial internal or external rotation does not show influence on the patella orientation at all, which is not satisfactory.

The strong 1-4-12-scaling leads to a very vivid interpolation over the whole inter- and extrapolation area, but causes huge gradients between the given control points. If the graphical animation is implemented using this scaling method, it shows unnatural positive and negative accelerations if the knee is moved with a constant velocity. Thus, this scaling method proved not to be usable for this project.

The medium 1-2-6-scaling is the scaling that is most plausible from the clustering pattern of the control points. Moreover, the diagram shows the smoothest behavior and the animation acts pleasingly to the subjective viewer. A closeup of the curves with the exponents $l_i = 3$ and $l_i = 4$ shows, that both produce flat spots - the so-called "terracing effect". In direct comparison, the exponent $l_i = 4$ shows larger terraces and steeper gradients between them. This difference was not obvious, when the animation was looked at, so the exponent $l_i = 3$ and the medium scaling factor are chosen.

3.6.4 Shepard's algorithm on the quaternion sphere

The weight calculation and the calculation of the complete interpolation can easily be done using equation (3.3). However, this is only valid in euclidean spaces - on the quaternion sphere problems occur using the quaternion addition. Additions are defined in quaternion space, but the sum of two unit quaternions is generally not a unit quaternion any more. Thus, the interpretation of the result of the quaternion addition as a rotation is not possible.

Equation (3.3) cannot be implemented directly in unit-quaternion space, the exponential map is useful. Section 3.3.3 describes that the exponential map bears distortions for lengths or directions that have a different source than the contact point of the sphere and the exponential map. This distortion is a severe problem for the calculation of the weighted averages, since the weight calculation depends on the distance between an arbitrary rotation $v(x, y)$ and all given rotations $v(x_i, y_i)$. Thus, the contact point of the exponential

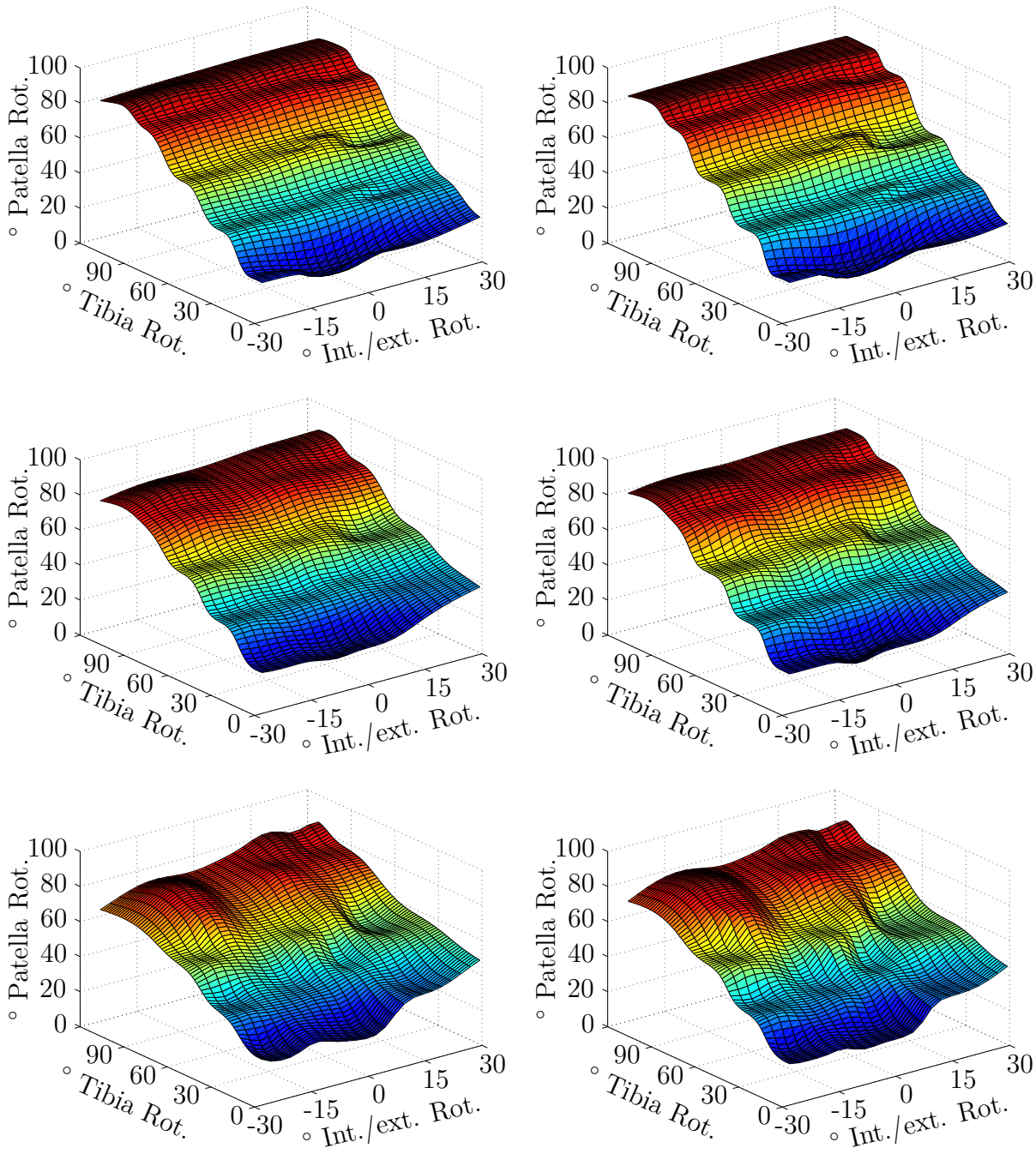


Figure 3.8: The surfaces of the left column were created with the exponent $l_i = 3$, the curves of the right column had the exponent $l_i = 4$. The top row has no scaling, the middle row has a 1-2-6-scaling and the lowest row shows the results of a 1-4-12-scaling.

map has to be at the point $v(x, y)$. Unfortunately, this rotation $v(x, y)$ is the result and thus unknown at the beginning of the calculation. To overcome this problem, iterative approaches are known. A more or less random tangent point tp_1 is chosen, $v_1(x, y)$ is calculated on the exponential map and mapped back onto the sphere. As $tp_1 \neq v_1(x, y)$, distortions occur. However, due to the nature of these distortions [80], $v_1(x, y)$ is closer to the correct value $v(x, y)$ than tp_1. This procedure can be repeated with $tp_{n+1} = v_n(x, y)$, leading to an even better $v_{n+1}(x, y)$. If the difference between $v_n(x, y)$ and $v_{n+1}(x, y)$ is small enough, the algorithm can terminate:

```
Given:
25 quaternions q[1]..q[25]
25 floats m[1]..m[25] with m[1]+m[2]+..+m[25]=1

Variables:
4D quaternions: P, TP, E
3D vector: p
Double: error=100
Integers: i

TP={0, 0, 0, 1}
while(|error|>0.001)
  error=0
  for(i=0; i<25; i++)
    P=mult(q[i], inverse(TP))
    p=log(P)
    error=error+p*m[i]
  E=exp(error)
  TP=mult(E, Q)
print(TP)
```

It is evident that this iterative procedure is time consuming, but it has not been possible to find an inverse Shepard method on the quaternion space that would allow a direct calculation.

3.7 Results

As the commonly known interpolation functions are not suitable for interpolating scattered data, triangulation, radial basis functions or Shepard methods have to be used. Most of these approaches have never been implemented for interpolating three-dimensional data on quaternion spaces. The Shepard methods proved to be the most promising approach, as most of their drawbacks (terracing and clustering) can be avoided by using scaling methods and by carefully choosing the correct parameterization.

Different implementations and weight calculations have been compared under the aspect of smoothness, speed and precision of interpolation and extrapolation. The optimal parameter configuration has been determined and verified by the subjective impression.

However, due to the small number of given control points, these calculations are not free of doubt.

3.8 Conclusion and outlook

Scattered control points are very common, when control points are the results of measurements. In combination with measured values that are in non-Euclidean spaces, the interpolation is challenging. Although the theoretical background of quaternions and interpolation algorithms are well understood, combinations of them are rare. In fact, the iterative approach that has been presented in section 3.6.4 is hardly elegant. Further research can be made in this sector, probably not with quaternions any more, but with different representations of rotations.

Many rotations - like skidding cars or capsizing boats - have only one relevant rotational axis. The same applies to interactive characters of computer games, which have only one relevant axis in each joint. The Euler angle representation of such simple rotations can be treated like other values in an Euclidean space. Thus, sophisticated interpolation algorithms are not necessary for these applications. More complicated rotations around different axes mainly occur in robotics and aeronautics. In those applications, interpolations are rarely used and may be covered by the SLERP and SQUAD algorithms that have been presented in section 3.3.3.

To the knowledge of the author, multidimensional interpolations of scattered rotations are neither commonly used nor desperately wanted. However, for this work they are inevitable. Accordingly, they are needed for other approaches that try to simulate biomechanical systems, as the data acquisition processes (thus their results) are similar. Simulating the mechanics of the hip, shoulder or the elbow seems to be an appropriate target; simulating more complicated and filigree structures like the palm or the spine might drive the algorithm's precision to its limits or beyond.

4 Deformations

4.1 Introduction

Deformations are very common in everyday life: clothes deform with every movement, metal plates are bent to the shape of a car and can easily be re-bent to any other shape in case of an accident. An egg-shell deforms if hit with a spoon and a pile of peas also changes its shape, if a spoon is inserted. As deformations are very common, one might think that their simulation should pose no problem today. However, there are still major challenges.

The four examples above are four completely different physical phenomena. Clothes that deform due to the movement of the wearer deform in an elastic manner, i.e. the deformation can be reversed by releasing the applied forces. The microscopic background of this phenomenon is the molecular structure of the macroscopic object which does bend due to the applied forces, but does not break. This elastic effect can be seen in most soft materials like rubber, skin or foam.

Opposed to this elastic deformation, the plastic deformation of metal plates can be explained by partial breaking, shifting and reconnecting of the intermolecular bondings. As the molecular bonding pattern is altered, the deformation cannot be undone by releasing the external forces, thus it is not reversible.

The third example of the egg shell deformation is explained by the breaking of molecular bonds without reconnecting again. Millions of breaks in the molecular structure result in a macroscopic destruction of the egg-shell. Of course, releasing the external forces does not heal the damaged shell, so these breaks and cracks are not reversible, either.

The fourth kind of deformation is the amorphous deformation, that can be seen in a pile of peas, sand, MoS_2 lubricant or fire. The deformation occurs by changing the relative position of objects [1]. These objects do not have static inter-object connections like the molecular bonds in the three other kinds of deformations. Thus, the amorphous deformation is not reversible. Moreover, it is highly flexible, as every object can change its position freely. As an example for this flexibility, a pile of sand can be chosen, where a grain of sand may be pushed through the center without severe effects on the neighboring grains.

In fact, these four basic models of deformations show immense overlappings. If a metal spring is located between two objects which are pushed against each other, the spring deforms elastically, as it was designed by the producer. However, if more force is applied, the spring can bend plastically and even break. If even more force is applied, the spring can break multiple times and the remains can interact as an amorphous mass. Thus, all four kinds of deformations can be applied to a metal spring and to most other objects - at least to a certain extent.

[1] The size of these objects can be molecular (like the molecules in a flame) or macroscopic (like the peas in a pile)

Until today, no simulating system can simulate all kinds of deformations, as has been described in section 1.4.2. FEM for example simulates the molecular bondings, that can change their length and other parameters, but they cannot break. A cutting simulation or a breaking simulation require a new setup of the FEM mesh, which is only possible in experimental FEM simulations. A particle system on the contrary can easily simulate broken or cut objects, as it mainly consists of objects without static bonds. However, it has only very limited capabilities to simulate elastic deformations.

In this chapter, an adequate deformation system is found for simulating a meniscus. The challenges for this task are described in section 4.2. As the requirements for the deformation algorithm are known, different state of the art approaches are considered in section 4.3 and the best one is explained in detail in section 4.4. The results of the approach are presented in 4.5 and concluded in 4.6.

4.2 Problem definition

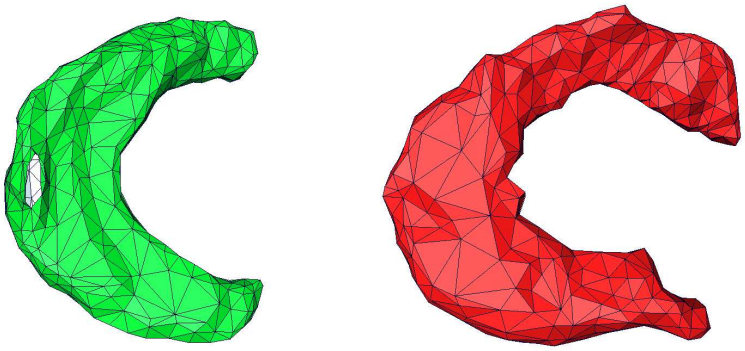
Soft tissue deformations can be calculated in many different ways. All those approaches differ in many aspects like the nature of the deformation, calculation speed and precision of the deformations. For the knee joint simulator, a real-time elastic deformation is needed that has enough precision to handle the laminar contact situation. As it has already been described in section 1.4.2, approaches based on physical models (FEM or mass-spring-damper models) are not suitable for this task.

Moreover, the data acquisition process of section 2 does not provide the necessary input for realistic physical simulations directly. Nevertheless, the positions and orientations of all knee joint components are provided, which is sufficient for interpolating and extrapolating the posture of the rigid knee joint components (see chapter 3). Via multiple optimizations, the shapes of the soft tissue components are also available from this data acquisition process. The approach of this chapter is the interpolation of those shapes, similar to the interpolation of the positions and orientations of rigid objects. However, the interpolation of shapes is not as straight forward as the interpolation of translations or orientations. For improving the understanding of the specific problems, only the two menisci that are shown in figure 4.1 are used throughout this chapter. Different approaches are proposed to transform one of those menisci to the shape of the other one by linear interpolation.

4.2.1 Simple mesh mapping approaches

The six translational and rotational DoF of an instance of a rigid object can easily be interpolated to the six DoF of another instance of the same object, as a correspondence of those parameters is clearly given: The x-coordinate of the first instance corresponds to the x-coordinate of the second instance and so on. Interpolating deformations or shapes is more challenging, because the shape can be represented in many ways (triangulated surface, volume and many more) and each representations is defined by much more than only six parameters.

For interpolating shapes, a suitable representation has to be chosen and a suitable set of parameters identified. In the following subsections, some approaches are presented, but only few are adequate for the problem.



Meniscus a Meniscus b

Figure 4.1: Two menisci that are used for depicting interpolation methods

Interpolation of Vertex Positions

To interpolate between two meshes, the suggestion is frequently made of interpolating the positions of their vertexes, edges or triangle normals. This does not seem possible at the first glance, because the manual segmentation and the Delaunay triangulation lead to a different number of vertexes, edges and triangles. However, at the second glance, it is possible to use pruning algorithms to trim the meshes to an equal number of vertexes, edges and triangles. This has been done with the two menisci in figure 4.1; however, their topology is still too different to allow an unequivocal assignment between two vertexes, edges or triangles. However, if mis-assignments occur, such simple interpolations do not provide valid, closed surfaces.

Nearest Surface Mapping

As the assignments of corresponding vertexes, edges or triangle normals pose such problems, a frequent suggestion is to compute the minimal distance of each vertex of one object to the *surface* of the other object. Thus, a vertex is not assigned to another vertex, but to a cartesian coordinate on the surface of the other object. Now, this approach is not symmetric, i.e. the assignment from surface "a" to surface "b" is different to the assignment from surface "b" to surface "a". In figure 4.2, an interpolation process with this algorithm is shown in six steps.

The vertexes of the meniscus "b" from figure 4.1, are mapped onto the nearest point on the surface of meniscus "a". This approach has several drawbacks: As meniscus "b" is significantly larger due to imprecise segmentation, the anterior horn (top) gets squeezed enormously by this approach. Many triangles are mapped onto the same region, which results in zero degree angles, thus degenerated triangles. This is most obvious in the 80% state in figure 4.2.

As has been described above, this approach is not symmetric. Mapping the vertexes of meniscus "a" to the surface of meniscus "b" leads to another result than the mapping of the vertexes from meniscus "b" to meniscus "a". As meniscus "b" is larger than meniscus "a", a squeezing effect occurs. Inverting the mapping direction, i.e. mapping from meniscus "a" to meniscus "b" can avoid this stuffing of triangles, because meniscus "a" is smaller. However, as the *nearest* point is chosen to be the corresponding partner of each triangle,

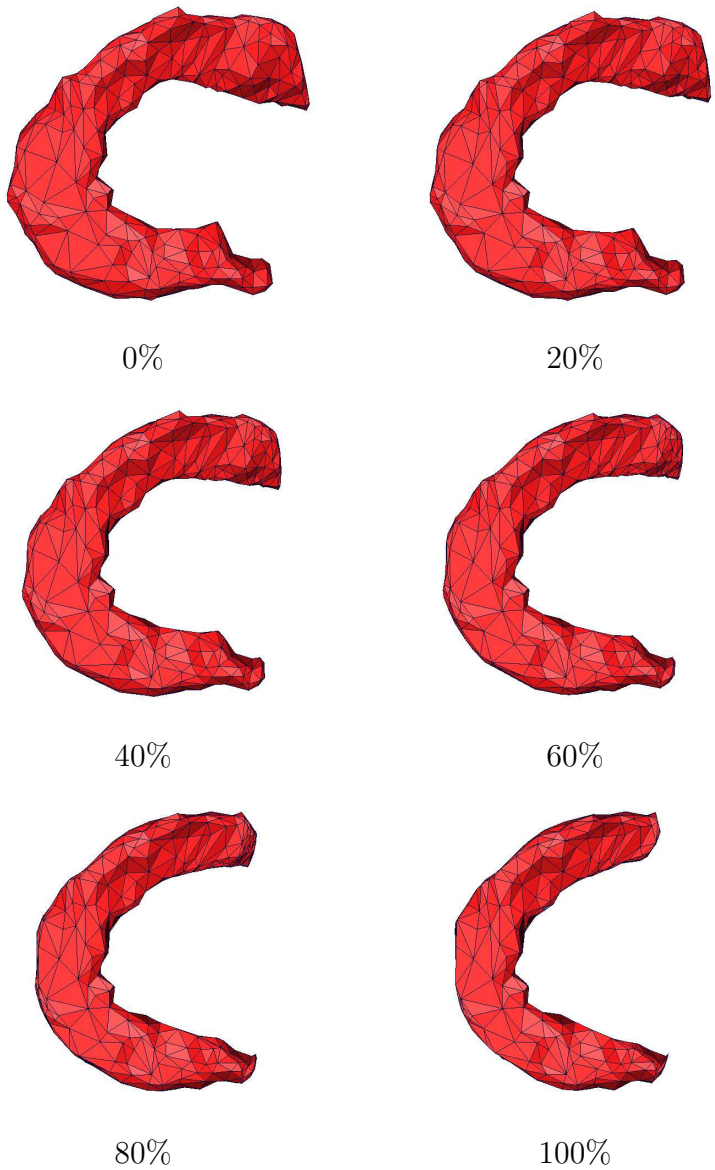
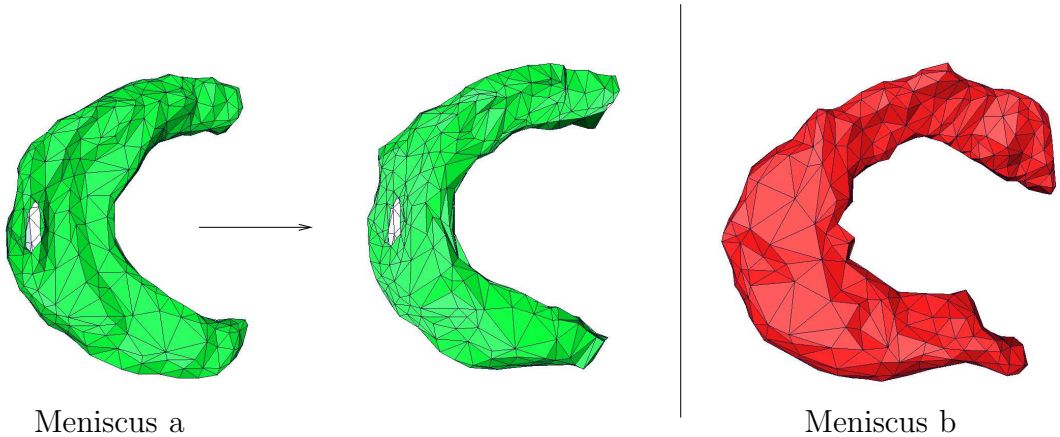


Figure 4.2: Simple interpolation between two menisci

the small geometry of meniscus “a” cannot grow to the size of meniscus “b”. The result of this approach can be seen in figure 4.3.



Meniscus a Meniscus b

Figure 4.3: Growing of geometries fails with the “Nearest Surface Mapping” approach. Here, the meniscus “a” is interpolated towards the geometry of meniscus “b”, but does not show significant changes.

Neither the deformation from meniscus “a” to meniscus “b”, nor the opposite direction is acceptable for the desired application. As the different mesh topologies pose severe problems to the simple, straight forward deformation approaches, more sophisticated deformation functions have to be considered in the following sections.

4.3 State of the art: Deformation functions

The known transformation matrices are able to represent the rigid body transformations, like translations and rotations, but also dilatations in all three dimensions. These operations (called the “Procrustes transformations”) can be combined in many ways. However, these Procrustes deformations are always global. Unfortunately, soft tissue like the knee joint components can be deformed locally due to punctual contact and thus cannot be described by such global functions.

Deformation functions are comparable with the Procrustes transformations, as they are not based on triangulated meshes or any other graphical representation of virtual objects. Deformation functions are rules of how to deform a certain region of space. All points in this region follow the same rule, no matter whether they are connected to triangles, cubes, hexagonal grids or if they are not connected at all.

Opposed to the Procrustes transformations, more sophisticated deformation functions also offer a local component as described by Sederberg et al. [81], who virtually embed the objects into such a cuboid. Figure 4.4 shows the model of a telephone receiver[2] that is embedded into a cuboid. The cuboid is split in two parts and the generated control points are translated.

After the translation, Bézier splines can be set up according to the start, control and end points of each edge. The embedded geometry can be deformed according to that spline.

[2]Note that the telephone receiver is smoothly deformed, but the bounding box is not. This is only done for emphasizing the location of the control points

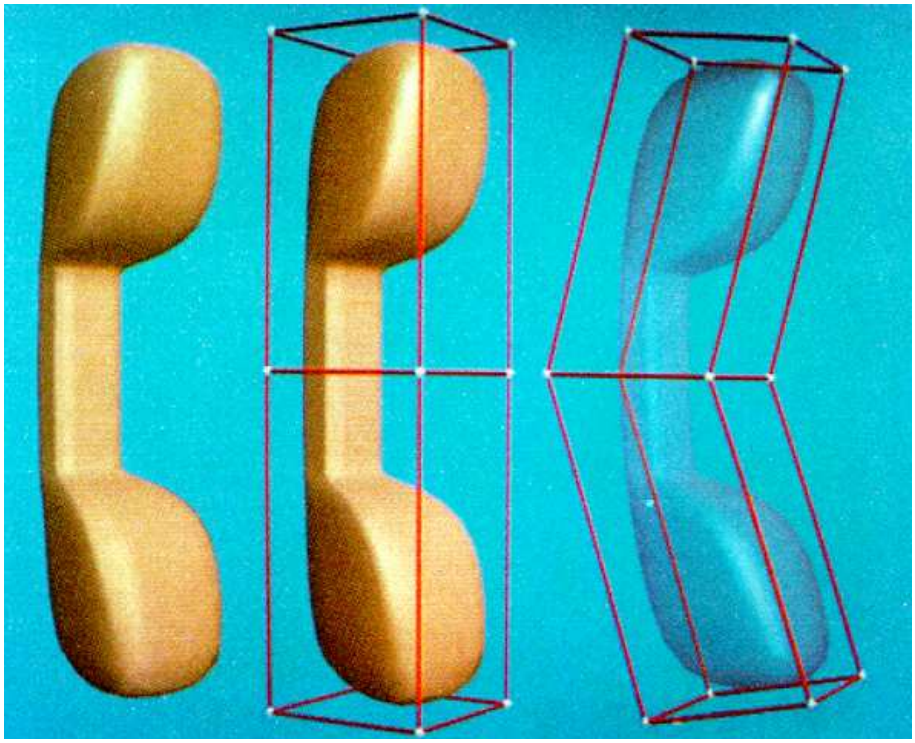


Figure 4.4: By adding few control points to the cuboid, Bézier splines can be used for global deformation

More local deformations can be achieved by adding more control points in the region of interest.

Barr [82] uses a two step approach: on the one hand global functions to bend, twist, taper, compress or expand the object, on the other hand "tangent transformations" and "normal transformations" for local deformation. Terzopoulos [83] combines conventional super-ellipsoids with locally deformable splines.

There are many more approaches to deform objects with the help of a set of parameters. They all share the same drawback concerning this application: They are optimized for an interactive, intuitive deformation. Depending on the necessary parameter set, they fulfill this target more or less with equal quality. These techniques do not apply here directly, as the user does not microscopically interact with all control points, but in a much more macroscopic way with the virtual shank. Transferring this macroscopic interaction to the necessary microscopic parameters is not trivial. Moreover, these algorithms allow a totally free deformation of the soft tissue components. There is no intrinsic method to avoid *unrealistic* deformation of the soft tissue components, like zero thickness.

Thus, the control points of the splines have to be parametrized carefully to ensure a realistic meniscus deformation. This parametrization has to be calculated from the shank position and orientation, as this is the one and only user input. Such parameterizations can be performed either by collision detection or based on the data that has been acquired by the data acquisition process in chapter 2. Both approaches are discussed in the next subsections.

4.3.1 Parametrization with collision detection

As deformation functions require the user to influence the control points, it is frequently suggested to take contact points between two objects as such control points. In this case, the femur and the tibia collide with the meniscus and cause deformations or control point displacements. Estimating the contact points requires a collision detection algorithm.

Normally, collision detection algorithms are used in combination with physical models like mass-spring-damper models, as already

explained in section 1.4. These slow physical approaches can be replaced by the much faster deformation functions. However, section 1.4 has shown, that collision detection algorithms themselves are far too slow to calculate laminar contact situations of soft tissue components in real time.

A further problem can be explained with the example of Sederberg's telephone receiver: A collision detection algorithm can detect contact points between an object and the telephone receiver. These points cannot be used for Sederberg's algorithm, as they are on the surface of the telephone receiver, but not on the bounding box. On the other hand, detecting contact points between the bounding box and the colliding

object is easy, but those points do not have any physical relevance.

4.3.2 Parametrization by acquired data

As the geometries of the soft tissue components have been acquired in chapter 2, it seems obvious to parametrize the deformation functions to match the 25 acquired shapes in the 25 corresponding knee joint postures and to provide smooth transitions between them.

As explained above, analytical deformation functions deform the space and everything in that space follows the deformation. This deformation is generally done with multiple control points. Thus, for interpolating the deformed objects that have been acquired before, the number of control points has to be defined and their positions have to be calculated for every acquired shape.

The number of control points to be defined depends on the required precision. The more control points are used, the more precise and detailed deformations can be simulated. However, there is an upper limit.

As the shapes are the result of manual segmentation and discretization, the positions of all voxels or vertexes contain a certain random factor. If one voxel or vertex is regarded in two different reconstructions of the same object, its position might have changed, although it did not move in reality. Accordingly, two neighboring voxels or vertexes that behave identically may appear as if they moved independently from each other. Thus, for a very precise interpolation, the number of control points has to be in the same dimension as the number of voxels or vertexes. This is not possible here due to the real time requirements, so an approach with less control points is inevitable, but will only be a compromise.

After the number of control points has been defined, their positions need to be calculated. The core problem for this is that the shapes of the deformed object are given, but the control points are on a virtual bounding box, that is not given.

4.3.3 Conclusion about deformation functions

Deformation functions are powerful tools for interactive and intuitive deformation of geometrical, virtual objects. Thus, if they are implemented e.g. in CAD tools, they can be used by designers and engineers to create organic geometries.

However, they have severe general drawbacks. The main problem is that they are computationally too expensive[3] for this work. Moreover, intuitive and interactive deformations are not needed here, but the inverse problem - that *is* needed here - has no solution. Thus, a parametrization of the splines is not possible. Due to those drawbacks, deformation functions cannot be used in this work. Accordingly, the menisci from figure 4.1 cannot be deformed and presented here.

4.4 State of the art: Surface warping

The surface warping approach is divided into three steps. The first step is the placement of so-called "landmarks" on all reconstructed surfaces, which can be performed off-line. The positions of corresponding landmarks are interpolated in the second step. This step is time critical, as it has to be performed on-line. The last step is the warping of the meniscus surface according to the interpolated landmark position. These three steps are explained in this section.

4.4.1 Landmark positioning

All approaches that offer local deformation have to divide the surface or the volume in subdivisions. The direct vertex, edge, or triangle wise interpolation that has been mentioned above divides the surface into several vertexes, edges or triangles. The deformation functions use a number of control points of splines to divide the surface into certain regions.

Generally, the more subdivisions are generated, the more detailed deformations can be implemented. However, neither of the suggestions above could use the maximum number of possible subdivisions available from the data acquisition process due to the real time requirements.

The family of surface warping methods also uses subdivisions of surfaces, but less than the already mentioned methods. In this approach, the geometry of the menisci is reduced to a set of landmarks that are placed on the surface by a human operator. Each of the 25 reconstructed representations of a soft tissue structure was marked with the same number of landmarks (see section 2.5.5).

In contrast to the approach of interpolating corresponding vertexes, edges or triangles, corresponding landmarks are easy to define on the surfaces of all different reconstructions. Thus, the landmark positions can be interpolated between all acquired postures, as it is depicted in figure 4.5.

The landmark positions can be interpolated much better than depicted in figure 4.5, if the weighted averages algorithm is used, which has already been described in chapter 3 and is superior to the linear interpolation as the discontinuities at the control points are avoided. Moreover, the landmark positions of the other degrees of freedom can be interpolated, which is not possible with the linear interpolation.

[3]Barr states that there are "simpler mathematical models" with "reduced computational needs"

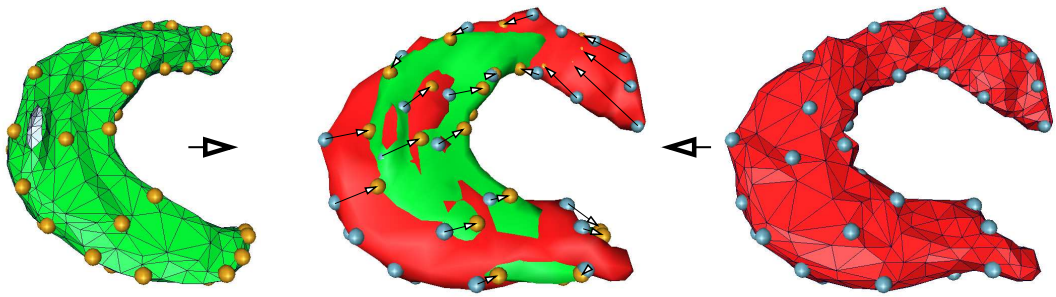


Figure 4.5: Here, the positions of the landmarks of two menisci are interpolated linearly

As those landmark positions determine the macroscopic or global deformations, the detailed topography of the meniscus has to be deformed accordingly. Therefore, the so-called surface warping methods can be used.

4.4.2 Thin Plate Splines - an overview

The very common warping methods use one reference mesh and deform it to fulfill certain needs. This deformation is achieved by changing the positions of the vertexes, the lengths of the edges and the size of the triangles. However, the topological structure remains the same.

A common approach for image warping problems are the "Thin Plate Splines", that were used by Fred Bookstein [56] for two-dimensional interpolation, but which can easily be extended to the three-dimensional warping. In any case, Thin Plate Splines require two sets of landmarks. The first set of landmarks is placed on the source object, i.e. each landmark is one point from the surface that is to be deformed. In this work, the meniscus in the extended, unloaded posture (without varus/valgus stress or internal/external rotation) is used as the source object.

The second set of landmarks are points on the surface of the target geometry. As each landmark in the first set is assigned to its corresponding landmark in the second set, the dislocation of those dedicated points is given. The calculation of the dislocation of the other points of the surface corresponding to the dislocation of the landmarks is the task of the Thin Plate Splines.

The Thin Plate Splines are radial basis functions and have already been mentioned in section 3.5.2. As all radial basis functions, the Thin Plate Splines can interpolate scattered data, which is needed for spanning up a surface from arbitrary landmarks.

The name "Thin Plate Splines" is derived from their characteristic of minimizing the bending energy of the interpolating surface, which resembles a thin metal plate that is bent around the control points[4]:

$$E_{bending} = \iint_{R^2} ((\frac{\partial^2 z}{\partial x^2})^2 + 2(\frac{\partial^2 z}{\partial x \partial y})^2 + (\frac{\partial^2 z}{\partial y^2})^2) dx dy$$

[4]Gravity, in-plane deformations and other real-world parameters of metal plates are omitted here

In the two-dimensional case, the plate is located in the (x, y) plane and the data points cause the curvature in the z direction:

$$v(x, y) = \phi(r) = r^2 \log r^2 \quad r = \sqrt{(x^2 + y^2)}$$

The value of this function is zero in the origin of the cartesian coordinate system and on the circle with the radius $r = 1$. The maximum of the surface is on the circle with a radius $r = \frac{1}{\sqrt{e}}$, and there is no global minimum, since the function approaches $-\infty$ with growing distance from the origin.

Note that this function is a so-called "fundamental solution of the biharmonic equation $\nabla^2\phi = 0$" [56]. As the Laplace operator produces different expressions in different dimensions, the function $U(r)$ has to be adapted to the dimension of the problem. In one-dimensional space, $\phi(r) = |r|^3$, the two-dimensional space is already explained with $\phi(r) = r^2 \log r^2$, while in three-dimensional problems, $\phi(r) = |r|$, if cartesian coordinates are used. The following explanations will cover the two-dimensional case, as it is the most catchy.

4.4.3 Thin Plate Splines - warping

To set up a warping scenario with the Thin Plate Splines, M landmarks $(x_1, y_1), (x_2, y_2), ..., (x_M, y_M)$ have to be given on the surface of the source object in the euclidean space[5]. The function $\phi(r)$, which is a function of the distance between a point (x, y) to the origin of the coordinate system $(0, 0)$, can be written as a function of the distance between two arbitrary points (x_i, y_i) and (x_j, y_j):

$$\phi(r_{ij}) = ((x_i - x_j)^2 + (y_i - y_j)^2) \log((x_i - x_j)^2 + (y_i - y_j)^2)$$

With this RBF, equation (3.1) can be set up:

$$\boldsymbol{v}(x, y) = \begin{bmatrix} x' \\ y' \end{bmatrix} = a_1 + a_x x + a_y y + \sum_{i=1}^{M} m_i \phi(|(x_i, y_i) - (x, y)|) \tag{4.1}$$

As already mentioned in section 3.5.2 the parameters a_1, a_x, a_y and all M parameters m_i need to be calculated. As this is a linear equation system, the matrix notation is handy. First, the $M \times M$ matrix $\boldsymbol{K}$ is defined from the RBFs[6]:

$$\boldsymbol{K} = \begin{bmatrix} 0 & \phi(r_{12}) & ... & \phi(r_{1M}) \\ \phi(r_{21}) & 0 & ... & \phi(r_{2M}) \\ ... & ... & ... & ... \\ \phi(r_{M1}) & \phi(r_{M2}) & ... & 0 \end{bmatrix} \quad M \times M$$

The locations of the landmarks can be written as matrices, too. Remember that two sets of landmarks are necessary: One set for the target object in matrix $\boldsymbol{Y}$ (see below) and the landmarks of the source object in matrix $\boldsymbol{P}$:

$$\boldsymbol{P} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ ... & ... & ... \\ 1 & x_n & y_n \end{bmatrix}, \quad M \times 3;$$

[5] In opposite to the interpolation of rotations, euclidean space is sufficient for deformation tasks. Thus, radial basis functions can be applied here

[6] In Euclidean space, $U(r_{ii})$ is zero and $U(r_{ij}) = U(r_{ji})$

The linear equation system is set up with the matrix $\boldsymbol{L}$:

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{K} & \boldsymbol{P} \\ \boldsymbol{P}^T & \boldsymbol{0} \end{bmatrix}, \quad (M+3) \times (M+3)$$

where $\boldsymbol{0}$ is a 3×3 matrix of zeros. For including the target landmarks, a matrix $\boldsymbol{V}$ is set up and extended to matrix $\boldsymbol{Y}$:

$$\boldsymbol{V} = \begin{bmatrix} \boldsymbol{v}_1, & \boldsymbol{v}_2, & ..., & \boldsymbol{v}_M \end{bmatrix}^T, \quad M \times 2$$

$$\boldsymbol{Y} = \left[\begin{array}{c|ccc} \boldsymbol{V} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{array}\right]^T, \quad (M+3) \times 2$$

If a vector $\boldsymbol{m} = (m_1, ..., m_M)$ is defined, then the linear equation system can be completed and solved:

$$\boldsymbol{L} \left[\begin{array}{c|ccc} \boldsymbol{m} & a_1 & a_x & a_y) \end{array}\right]^T = \boldsymbol{Y}$$

or

$$\boldsymbol{L}^{-1}\boldsymbol{Y} = \left[\begin{array}{c|ccc} \boldsymbol{m} & a_1 & a_x & a_y) \end{array}\right]^T$$

With $\boldsymbol{m}$ and a_1, a_x and a_y, all parameters are known to interpolate with the equation (4.1).

This warping function consists of two parts: The affine part $(a_1 + a_x x + a_y y)$ that defines the behavior of $f(x, y)$ at infinity and the sum of functions $(\sum_{i=1}^{M} m_i \phi(|(x_i, y_i) - (x, y)|))$ which was parametrized for a bounded and asymptotically flat interpolation. A more detailed introduction with more theoretical background is given in [56].

The extension to the three-dimensional case is straight forward, including an adaption of the RBF to $\phi(r) = |r|$ instead of $\phi(r) = r^2 \log r^2$. Then, it can be set up like:

$$\boldsymbol{v}(x, y, z) = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = a_1 + a_x x + a_y y + a_z z + \sum_{i=1}^{M} m_i \phi(|(x_i, y_i, z_i) - (x, y, z)|) \qquad (4.2)$$

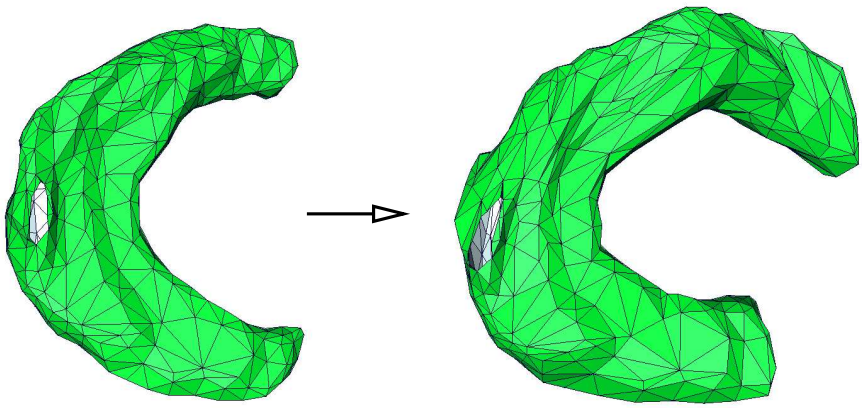
and its solution works with several changes in the matrix dimensions like:

$$\boldsymbol{L}^{-1}\boldsymbol{Y} = \left[\begin{array}{c|ccc} \boldsymbol{m} & a_1 & a_x & a_y & a_z) \end{array}\right]^T$$

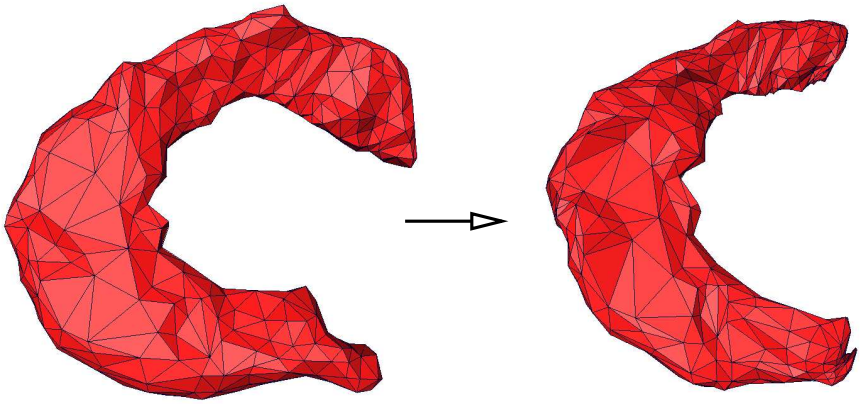
$$\boldsymbol{Y} = \left[\begin{array}{c|ccc} \boldsymbol{V} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \end{array}\right]^T, \quad (M+4) \times 3$$

4.5 Results

As already considered in section 1.5.2, the quality of the animations can be judged from a macroscopic and from a microscopic point of view. In rigid bodies, the quality of the interpolation can be determined in many different ways, as demonstrated in section 3.6.1 and 3.6.2. Similar approaches for estimating the interpolation quality are not applicable for soft tissue, thus other methods have to be used.



Meniscus a



Meniscus b

Figure 4.6:
Meniscus “a” warped to the surface of meniscus “b” (top). The opposite warping direction is depicted below

4.5.1 Subjective comparison

First, the two menisci from figure 4.1 are deformed by the Thin Plate Splines. As this approach is not symmetrical, figure 4.6 shows the result of meniscus “a” , as it is transformed to meniscus “b” in the top row, while the result of the opposite deformation can be seen in the same figure below.

The interpolation between those shapes is possible without degenerated triangles or other drawbacks of the previously mentioned methods. Moreover, the microscopic appearance of the transformed object remains intact and only the macroscopic appearance is adjusted. This is a welcome feature, because many segmentations show flaws in the closeup view, whereas their macroscopic appearance is acceptable. Due to this characteristic, the menisci of the neutral knee joint position (extended, unloaded position) were segmented very carefully to provide a realistic and pleasant appearance, even in the warped postures.

Apart from the two examples, the Thin Plate Splines are implemented in the graphical display; there, no linear interpolation, but the weighted averages interpolation algorithm

(see chapter 3) is used. The interpolation of the deformation in three DoF shows a smooth characteristic without any unexpected jumps or static behavior. This is not surprising, as the deformation is based on the position interpolation of the rigid objects, that also showed a nice behavior. Thus, further consideration is not about the *interpolation* of the landmarks, but about the *deformation* of the whole geometries.

As shown above, the Thin Plate Splines can warp the meniscus geometry in a pleasing way. However, this has to be verified with the new, three-dimensional interpolation, too. Therefore, the virtual knee joint has been directed to different postures known from the data acquisition process. Thus, the real, segmented meniscus geometries can be compared with the generated geometries from the animation.

In figure 4.7, four interpolated menisci geometries are opposed to four segmented and reconstructed menisci geometries. The four cases show the knee joint in 32.0°, 63.5°, 98.3° and 117.4° flexion without any additional varus/valgus stress or internal/external rotation. Of course, the other segmented and reconstructed postures have also been considered, but as they are quite similar to the given examples they are not explicitly shown here.

As can easily be seen in figure 4.7, the segmented and reconstructed menisci on the left side show rugged surfaces due to segmentation errors. The menisci on the right side of figure 4.7 are deformed versions of the menisci in the natural, extended pose. As an extremely precise segmentation was carried out to generate these "master" menisci, their surfaces are very smooth, and so are the surfaces of the warped versions.

Apart from the smoother surfaces of the warped menisci, direct comparison shows that the macroscopic deformation is simulated correctly. In low flexion angles, the semi-circled shape of the lateral meniscus is stretched around 45 degrees to the lateral anterior direction[7]. As the contact point between the femoral condyli and the tibial plateau slides backward in increasing flexion angles, the menisci are also stretched in this direction.

Although both menisci are deformed in a similar manner, the deformation of the lateral meniscus is more obvious than the deformation of the medial meniscus. This can be explained by the smaller size of the lateral meniscus and by the smaller geometry of the lateral femoral condylus.

4.5.2 Quantification of the warping error

The subjective quality of the meniscus deformation is satisfactory. Until today, there are no objective standards for the quality of deformations, as no common metric can be applied here. Measuring the distance of each warped vertex to the segmented and reconstructed surface and calculating the average error is misleading, as those surfaces suffer significant local segmentation errors. However, in section 2.10, a method for comparing meniscus geometries in three-dimensional space was introduced. These measurements - based on measuring two angles to estimate the overall deformation - have been performed on the real, segmented menisci and on the simulated, warped menisci. The results of these measurements can be seen in table 4.1 and table 4.2.

Table 4.1 and table 4.2 show that the differences between the real meniscus and the warped meniscus are hardly noticeable in the macroscopic view. Thus, not only the subjective, but also the objective quality of the warping is satisfactory.

[7] Accordingly, the medial meniscus is stretched in anterior medial direction in low flexion angles

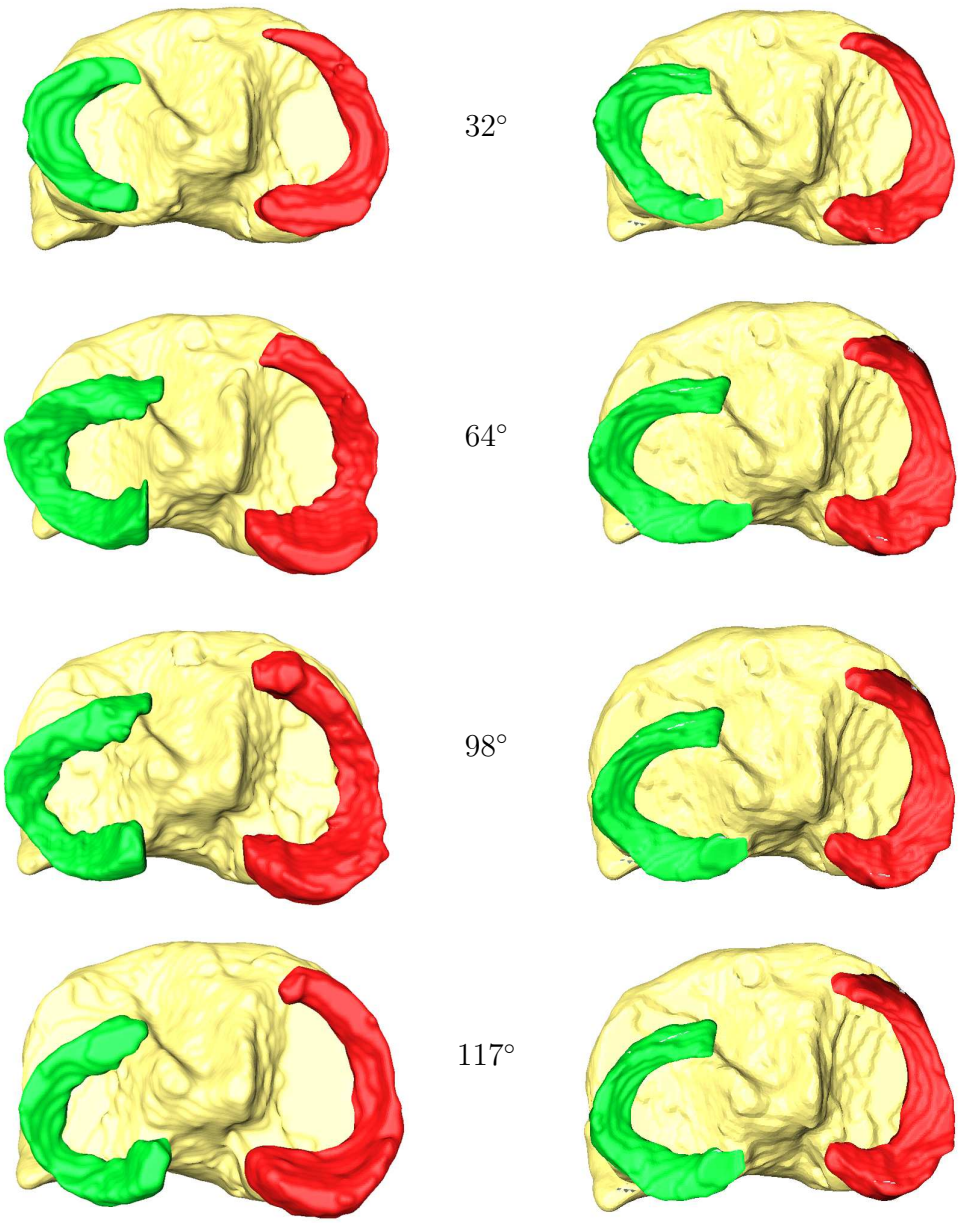


Figure 4.7: The segmented menisci (left) are compared with their corresponding animated representations (right)

Real lateral meniscus	32.0°	63.5°	98.3°	117.4°
Angle AXC	17°	19°	25°	27°
Angle BXC	26°	25°	23°	22°
Warped lateral meniscus	32.0°	63.5°	98.3°	117.4°
Angle AXC	18°	19°	25°	26°
Angle BXC	25°	25°	24°	19°

Table 4.1: The measured and the warped lateral meniscus

Real medial meniscus	32.0°	63.5°	98.3°	117.4°
Angle AXC	44°	38°	32°	32°
Angle BXC	16°	16°	12°	12°
Warped medial meniscus	32.0°	63.5°	98.3°	117.4°
Angle AXC	42°	38°	33°	32°
Angle BXC	16°	15°	14°	13°

Table 4.2: The measured and the warped medial meniscus

4.5.3 Warping errors and their sources

In chapter 3, the movement of the rigid objects has been developed and parametrized to minimize its interpolation errors. In the current chapter, the deformation of menisci has been developed. In both cases, the interpolation errors are considered to be minimal. However, when these two techniques are combined in the setup of the interactive tool, certain flaws can occur in several knee joint postures. One of the more severe examples is shown in figure 4.8, where the femur penetrates the meniscus completely.

The menisci have a large contact area with the tibia plateau and the femoral condyli. Other parts that are not in contact with the bones are at least very close to the bones. Thus, even small interpolation errors in only one DoF can lead to penetrations of those surfaces. Such interpenetrations will occur with any interpolation algorithm without collision detection and can be accepted if they remain microscopic errors and do not (macroscopically) spoil the whole simulation.

If severe interpenetrations like in figure 4.8 occur frequently, mechanisms for avoiding such postures have to be considered. As three-dimensional collision detection of laminar contact areas and deformable materials is computationally very expensive, and as the real time capabilities are absolutely necessary, this approach is not promising.

A better way to enhance the animation is by reducing the interpolation error. Deformations are calculated by a cascade of imprecise methods, leading to an error accumulation. To minimize the interpolation error, the most significant error source in this cascade has to be located. Generally, this error source can be:

- Segmentation errors
- Errors in estimating the landmark positions
- Number of landmarks
- Landmark interpolation error

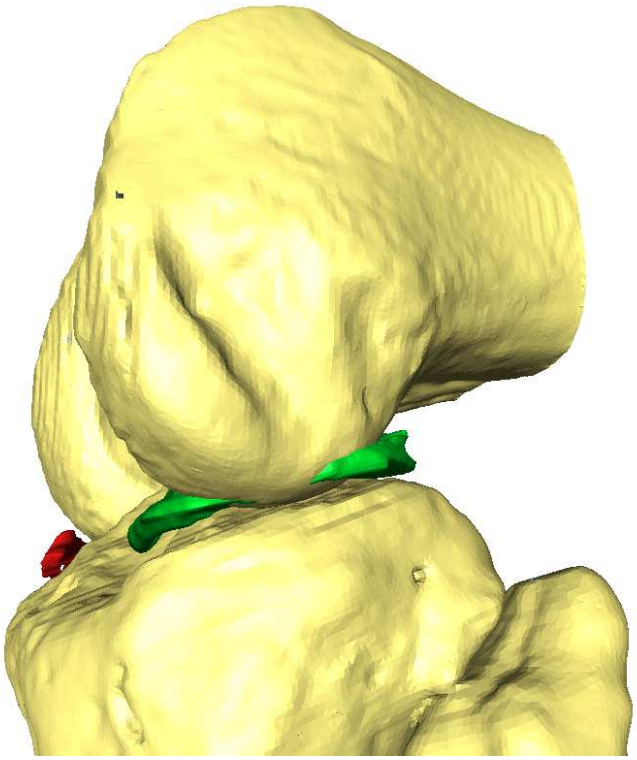


Figure 4.8: Where interpolation or extrapolation errors sum up, severe macroscopic errors can occur

The first errors occur when the human operator places the landmarks on the surfaces of the different reconstructions. As soft tissue contrast is rather low in CT scans, the reconstructed models are not free of errors. One of the side effects is an incorrect placement of landmarks on those surfaces. One example is displayed in figure 4.5: As the horns of the menisci are connected to the cruciate ligaments, and as both materials have similar properties in the CT scanner, a definite border between the horn of the meniscus and the ligament cannot be found. While trying so, the human operator may over-interpret light shadows or noise in the CT data, finding borders that do not exist. Thus, the horns of the menisci look different in different segmentations.

The next source of errors results from the procedure of landmark placement itself. There are not enough locations on the menisci that can be clearly identified on all 25 reconstructions. In most cases, the anterior horn, the posterior horn and the most lateral/medial [8] voxel are said to allow a clear identification. This is not correct for the horns of the menisci due to the segmentation errors mentioned above.

The term "most lateral or medial point of the meniscus" is also imprecise, as this definition depends on the definition of the sagittal plane, which is *not* unequivocal. Moreover, a surface element of a real meniscus that is on the most lateral/medial position in full extension is very unlikely to be the most lateral/medial surface element in flexion. Therefore, the assignment of these virtual landmarks to the real surface fails.

Thus, the first three landmarks were placed on the points "A", "B" and "X" that have been described in section 2.10

[8] depending whether the lateral or medial meniscus is considered

(see figure 2.14). The rest of the landmarks have been distributed over the surface in relation to those three basic landmarks. It should be clear that this positioning of the landmarks is error prone, as it basically relies on the human sense of proportions.

The next source of errors is the low number of landmarks. Like every interpolation technique, the Thin Plate Spline Surface Interpolation behaves better with more control points. Unfortunately, the placement has to be made manually, thus the number of manageable landmarks is limited. If 20 landmarks are set on each of the 25 reconstructions, 500 manual assignments have to be done for each soft tissue component. This number proved to be beyond the limit. Generally, errors in the assignment of landmarks lead to a destruction of the interpolated surface. The best approach in this work has only 13 landmarks on each meniscus.

The last possible error, the "landmark interpolation error" is twofold: First, the landmarks that define the target geometry for the Thin Plate Splines are calculated by the Weighted Averages Algorithm, which has been explained in section 3.5.3. Like every interpolation algorithm, all values (except those of the control points) are only approximative. These errors have already been measured and are not significant by themselves. In figure 4.8, the position of the tibia, the position of the meniscus and the deformation of the meniscus are influenced by this approximation error, thus these errors may sum up to a significant amount.

The second aspect of the "landmark interpolation error" is the second interpolation algorithm - the Thin Plate Splines - which adds further errors. However, as the surface is closely bound to the given landmarks, only very small errors are expected here.

4.6 Conclusion and outlook

In this chapter, a method for a three-dimensional object deformation has been developed. It is a unique combination of the known "Thin Plate Splines" warping algorithm and the "keyframe interpolation" (see section 1.4.2). The keyframe interpolation is used with the "Weighted averages" interpolation algorithm. To the knowledge of the author, neither of these combinations have ever been implemented and evaluated before. Accordingly, a interpolation-based deformation algorithm for a three-dimensional object has never been tried before.

The deformation is not really a free-form deformation. 13 landmarks have been placed on all reconstructions of one meniscus, and all landmarks can be moved in three dimensions; thus, the three-dimensional meniscus can be deformed in $13 * 3 = 39$ dimensions. Although this deformation has only been implemented for the menisci, the deformation of ligaments and other soft tissue components can be implemented in the same way.

This approach interpolates or mixes different given shapes. Thus, it generally differs from most other deformation approaches, as they are optimized for interactive deformations.

This algorithm can be used for all applications, where multiple objects have to be morphed, mixed or interpolated. Such applications can be found especially in the entertainment sector, e.g. in the development of animated films or computer games.

However, there are also more serious applications. Let us suppose a number of thick and thin, tall and short, old and young individuals, which are scanned with a 3D-laser scanner. Corresponding landmarks are placed on reproducible positions, such as joints. Any individual of this dataset can be interpolated to be thicker, thinner, older, younger,

taller or smaller - and any combination thererof[9]. Thus, the technique of mixing individuals might find its application in anthropology; the technique of mixing faces might be useful for crime science, e.g. for producing identikits.

[9]Interpolating different postures like standing, sitting or kneeling is possible, but other techniques like the skeletal animation are more useful.

5 Conclusion and future work

5.1 Resum of this work

The basic idea of this thesis looks easy and straight forward: If the internal topology of a human knee joint is known for a couple of postures, it should be possible to draw conclusions as to

other postures. This really is possible for humans, neuronal networks and expert systems, if the conclusion has to be done only for qualitative expressions like "With increasing flexion angle, the contact point between femur and tibia slides in posterior direction".

If it is about quantification, this conclusion fails as human operators turned out to be very imprecise by managing 6 DoF. The simple task to place one virtual femur coinciding over the other was performed with up to 5 degrees of positioning error against the automatic positioning with the iterative least square positioning. More complicated tasks like the free placement of bones in certain angles certainly showed even more severe errors.

Thus, if an experienced medical doctor has a clear picture of a knee joint posture in his or her mind, the presentation to the students will still be error prone. Moreover, it is very likely that the imaginary knee joint posture in the doctor's mind is not precise and adds a few further degrees and millimeters of additional error.

As it turned out in this thesis, even small errors are important, as the knee joint moves only slightly in some degrees of freedom. The obvious idea that humans can deduce valid knee joint postures from their knowledge is wrong, as they cannot ensure the necessary precision.

Human knowledge of the knee joint is very limited, as the common data acquisitions are either only qualitative observations or only two-dimensional measurements. A more comprehensive, three-dimensional data acquisition is necessary and has been carried out for the first time.

This data acquisition process shows the limits of state of the art image generating methods. Three dimensional geometries cannot be acquired in an automatic way, and even the slice wise acquisition offers low contrast, low resolution, is time consuming and/or hazardous. The flaws of these image generating methods may be acceptable for their common use in diagnostic environment, but they pose a real problem if a detailed knowledge base is to be generated.

Due to the low image quality, the segmentation process is extremely time consuming. The menisci were chosen to be the focus of this work, because they were the soft tissue components that could be identified best. In most cases, their shape was defined by the laminar contact against femur and tibia. Other soft tissue components like the collateral ligaments are not even visible in all scans, as they are embedded in other soft tissue and as the contrast agent did not reach that remote area.

Because of these difficulties, the knowledge base consists of only one knee joint, does not cover all knee joint components, ignores translational forces and ignores pathologies.

The second topic of this work is the interpolation of the given postures. Although keyframe interpolation is a very common method, most of these approaches are limited to interpolating two-dimensional postures, avoiding three-dimensional rotations, thus avoiding the quaternions. Moreover, no approach is known where soft tissue deformation is implemented based on keyframe interpolation, which makes this work, the implementation and the quantification of the results especially unique.

From the medical point of view, this animation provides the first interactive multi-DoF virtual knee joint with deformable menisci. The displayed knee joint postures are faulty in some cases, but still superior to the known simulations. Moreover, this work demonstrated the capabilities and the limitations of current medical simulations.

Moreover, this work demonstrates that three-dimensional measurements are superior to the two-dimensional, slice wise measurements that are still common today. Due to the flaws of these measurements, most current descriptions of the tibia movement, femur geometry and meniscus deformations are imprecise or even wrong.

5.2 Future work

5.2.1 Data processing

Three-dimensional measurements prove to be important for scientific, medical work. On the other hand, the data acquisition, segmentation and post-processing task is very laborious. This seems to be the main reason why three-dimensional measurements are not common today.

Human operators can quickly identify the different components of a CT or MRI dataset due to their model based knowledge. If the geometry of the knee joint is known, the human operator searches the fibula on the lateral side and the patella on the anterior side of the dataset. This approach is more sophisticated than the simple evaluation of gray values or the search for highest gradients in the data set. Computer scientists try to copy this knowledge based approach, but a breakthrough is not yet probable.

Another approach to simplify the data post-processing would be the improvement of the common CT or MRI techniques. The CT scans that are performed in this work emit extremely high radiation, which can only be tolerated with cadaver experiments. As the desired resolution was achieved but the contrast still remained low in spite of the huge radiation, the problem seems to be related to the principle of x-ray attenuation.

The solution of this problem seems to be the development of stronger magnetic fields in the MRI scanners which allow shorter scanning times. Latest fMRI (functional MRI) scanners show impressive enhancements, but are still not good enough for acquiring a whole high-res dataset in reasonable time. However, this problem may be solved in several years.

5.2.2 Medical simulations

During this work, the developments of further medical simulators were considered that follow the idea of the Knee Joint Simulator, providing tactile feedback via passive plastic phantoms, kinesthetic feedback via a controlled actuator and graphical feedback. The most interesting idea was the birth simulator, where medical students and midwives can

train the basic methods of obstetrics. In opposite to the knee joint simulator, the birth simulator has time critical components, as some obstetrical methods should not be used too early or too late. Both errors can be lethal for the newborn or the mother, making the birth simulator more relevant than the knee joint simulator.

Due to the more complicated data acquisition process, the challenge of the birth simulator is significantly higher than that of the knee joint simulator. Haptic and tactile data acquisition are not in the scope of this work, so only the graphical data acquisition is considered here. The last profound data acquisition of the birth process was carried out by Sellheim in 1898 [84]. Sellheim used x-ray scans to estimate the fetus' position and posture, but this way of data acquisition is forbidden today, as x-rays proved to be especially hazardous to children.

If a three-dimensional data acquisition is needed to create a knowledge base for interpolation techniques, as it is done in this work, CT scans are also forbidden, since they expose mother and child to an even higher dose of x-rays. So either fast, low resolution MRI scans or ultrasound techniques are available, without the use of contrast agents. The resulting image quality is most probably even worse than the CT scans of this work. Thus, this kind of data acquisition is not very promising as long as MRI scanners are so slow.

Even if a data set of a complete birth process could be acquired, and even if it was possible to acquire such datasets from pathological cases, the other drawbacks of the interpolation approach would still be relevant: The contact situation between mother and child is even more laminar than the contact situation between meniscus and bones. Collisions and interpenetrations will be more frequent and more severe. A real time collision detection is even more unrealistic as the computational costs rise extremely if both colliding geometries are deformable. Moreover, active forces due to labors cannot be simulated by interpolation.

5.3 Conclusion

Every simulation depends on the model of the simulated system. Most precise models lead to a very realistic simulation, as is the case with flight simulators. The geometry of every single part is precisely known from the CAD construction plans, the behavior of the whole system was simulated in FEM, wind channel and measured in real flights. With this immense knowledge, flight simulators are the most precise simulations available, simulating even very uncommon situations like stalled motors, defective wings or loss of oil pressure for a complete series of airplanes. Even experienced pilots take advantage of simulator training.

The transfer from technical systems to biological systems is obvious: If technical simulators can be used to improve the education of technical staff, medical simulators can improve the education of medical staff.

However, the transfer is not that easy. Biological systems are not known to the same extent as technical systems. Neither the geometries of the different components, nor the behavior of the complete system is known in detail. Simulators that are based on such rudimentary knowledge cannot show the same usability as flight simulators do. In fact, if the knowledge base is too small, these simulations may be completely useless, even for novices.

Thus, there are enormous differences between different medical trainers, depending on the techniques used and on the required knowledge. Virtual endoscopy can be implemented

with mass-spring-damper systems, as only punctual contact situations occur. The required knowledge about spring constants and damper coefficients can easily be obtained. These systems provide nearly realistic graphic and haptic feedback and are useful for the training of students and young doctors.

Other simulations that cannot be implemented with punctual contact situations do not reach the same level of realism. Other techniques have to be used, that are not yet suitable for product development, but still require a certain amount of pure research.

A Medical background of the knee joint

This work is read by people of the medical sector, but also by computer scientists and engineers. As the knee joint is researched mainly in the medical sector, the corresponding nomenclature is used in this work. Thus, descriptions of locations, directions or components are made using medical terms that may not be known to the technically oriented people. Therefore, a profound introduction to the medical nomenclature is given here.

A.1 The knee joint: Components and topology

In this section, the basic knee joint components are listed. The precise explanation of the functionality is given in section A.2, as it requires the unequivocal definition of movements and directions, which are given in section A.1.3.

A.1.1 Nomenclature of the rigid knee joint components

The knee joint consists of four bones and numerous soft tissue components. The thighbone is called *femur* and is the longest bone of the human body. Its lower part is split up in two convex disks (called the *condyli*), which are the femoral part of the knee joint and allow a combined rolling and sliding motion of the knee joint. The shaft of the femur connects these condyli to the pelvis, thus the whole body weight is transferred through this bone.

The *tibia* (also called "shinbone") is the counterpart of the femur, as its upper part is more or less a flat plane, which is the tibial part of the knee joint. The tibia connects the knee joint to the ankle, but it does not carry the whole body weight. The *fibula*, which is also called the "calf-bone" lies parallel to the tibia, so the load is distributed between these two. In many cases, both bones are treated like a single one, so the combination is called the "lower leg" or "shank".

The fourth bone of the knee joint is the knee cap or *patella*. It is located in front of the femur and is always in contact with it. However, it does not carry the weight of the body.

Each contact area between two bones is more or less flexible and is called a joint of its own. Thus, there are three joints in the human knee, called the *tibiofemoral*, the *patellofemoral* and the *tibiofibular* joint. As the movement of the tibiofibular joint is hardly noticeable, it is neglected in most studies.

A.1.2 Nomenclature of the soft tissue components of the knee

Each joint is covered by cartilage, that is slightly compressible, thus deformable, but no special names are given to the different pieces of cartilage.

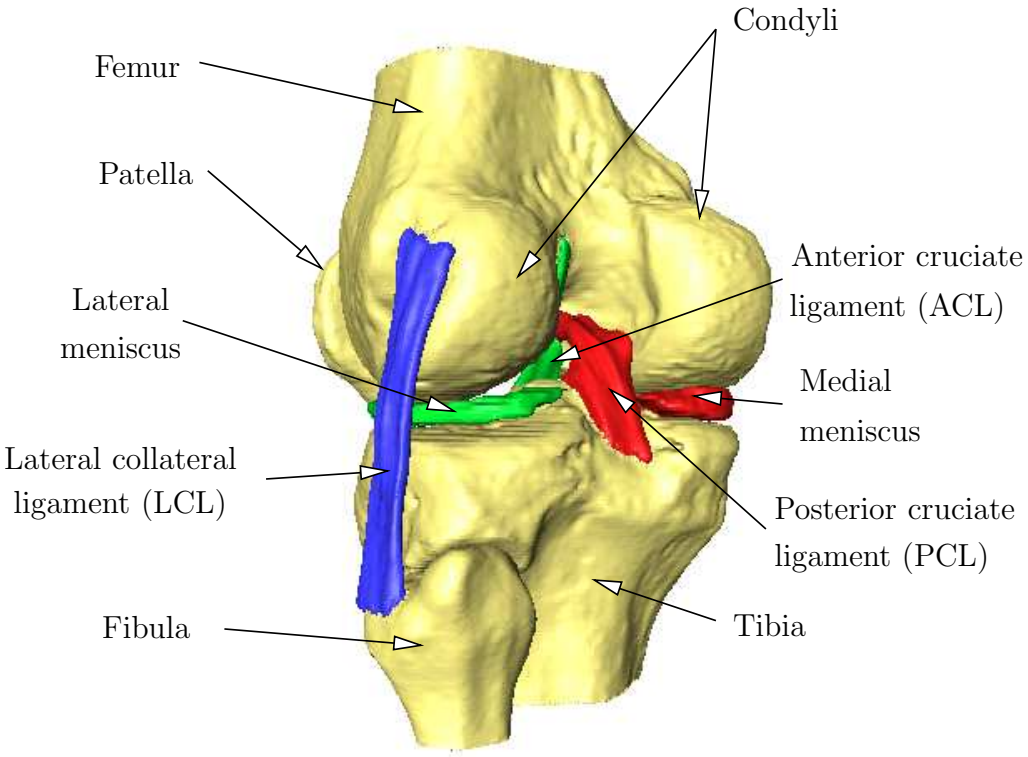


Figure A.1: The main components of the left human knee joint, seen from the left rear

Two menisci are located between the femur and the tibia (see image A.2). A meniscus is a c-shaped object that is made of collagen. The two ends of each meniscus are called its *horns* and are connected in the middle of the tibia plateau. The rest of the menisci can slide over the left or the right part of the plateau.

The geometry of the menisci is optimized to fit the surface of the condyli, thus the cross sections of the menisci are triangular, forming a concave geometry on the tibia plateau.

The cruciate ligaments are located at the center of the knee joint and connect the femur to the tibia. Viewed from the side, the ligaments seem to cross each other as one starts from the front side of the tibia and ends at the back of the femur, while the other cruciate ligament runs vice versa.

Two other ligaments are called the collateral ligaments as they connect the femur and the tibia on one side and the femur and the fibula on the opposite side. Those four ligaments provide the knee joint with stability in all directions.

Two further ligaments not considered in this thesis are the patellar ligaments, that connect the patella to the tibia and the femur, and which link the patella movement to the tibia movement.

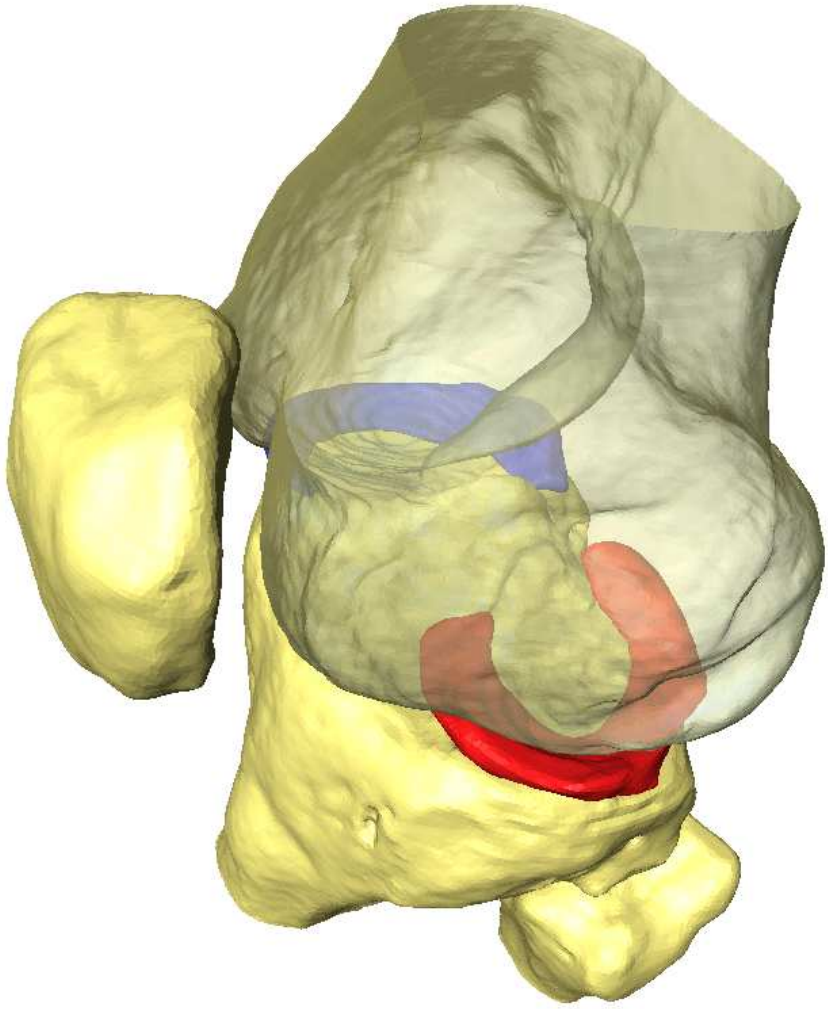


Figure A.2: The transparent femur reveals the view on the menisci.

A.1.3 Topology of the knee joint

In section A.1.1, and A.1.2 the knee joint components are introduced. To describe their precise locations in different knee joint postures, medical terms have to be explained, since terms like “left”, “right”, “front” and “back” are not clear any more whenever the knee joint moves. An overview of these terms can be seen in figure A.3.

The *flexion angle* is the largest degree of freedom in the knee joint. As it will be seen later, the measurements of these angles are not consistent in literature, but most authors suggest a maximal flexion angle between 130 and 150 degrees.

The *internal* or *external* rotation is much more limited with approximately +/-15 degrees. The term “internal tibial rotation” (or “external femoral rotation”) applies, if the tibia and fibula are rotated inwards around the centerline of the tibia. So, even if the knee joint is flexed, the rotational axis is defined by the tibia.

The third degree of freedom which was not yet mentioned is called the *varus* or *valgus* position. This degree of freedom is the tiny rotation that occurs when forces are applied in the direction of the ankle. Moving the shank “inwards” causes a “bow-leg” position,

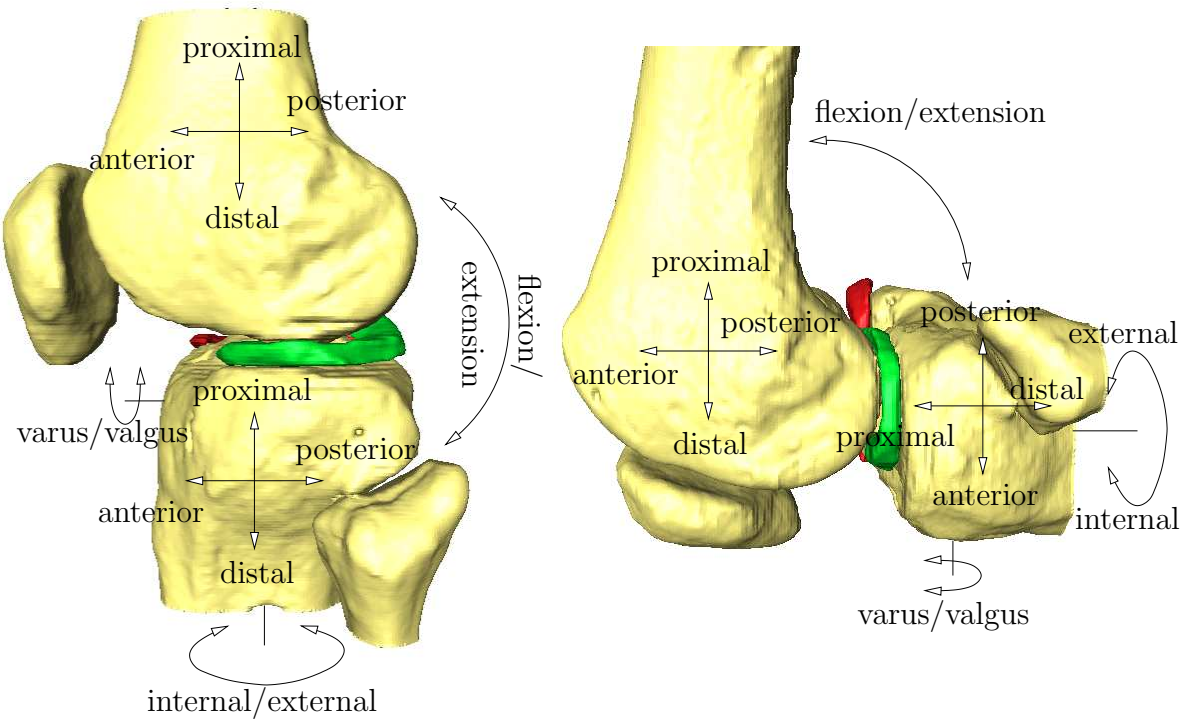


Figure A.3: The medical terminology of rotations and translations. The reader is facing the lateral side of the knee joint, so the medial side is not visible.

which is called the "varus" position. Likewise, pushing the shank outwards leads to a "knock-knee" position, which is called the "valgus" position. The varus/valgus movement is limited to approximately +/-5 degrees.

The three translational degrees of freedom are quite similar: The *anterior* direction points to the front of the body part, so the shinbone is the anterior side of the shank. The *posterior* direction is the opposite, so the posterior part of the shank is called the calf.

The *medial* part describes the inner part, while the *lateral* part is located towards the (out-)side of the body. The fibula is always located at the lateral side of the shank, whether looking at a left or a right leg. The location of the fibula allows medical doctors to identify the side of the leg at the first glance on MRI or CT slices or on images like figures A.1, figure A.2 and figure A.3. Here, a left knee joint is used.

The last translational degree of freedom is the *proximal* or *distal* direction. Proximal is often translated with "towards the torso", while distal is the opposite. This definition might be misleading, if the knee joint is not extended.

If the knee joint of a standing person is flexed around 90 degrees, the lower leg points backward, and the calf points upward, thus "towards the torso". However, the calf is the *posterior* part of the tibia, not the *proximal*.

Maybe it is more precise to define proximal as "situated toward the origin of the body part".

The directions of cutting planes are *axial*, *sagittal* or *coronar*. An axial slice shows one region of the body in its anterior-posterior and in its medial-lateral extend. All axial slices are parallel and are located perpendicularly to the proximal-distal dimension.

The sagittal and the coronar slices are perpendicular to each other and also perpendicular to the axial slices: Sagittal slices show the part of the body in proximal-distal and in anterior-posterior dimension and are perpendicular to the medial-lateral dimension. With this description, only one dimension remains for the coronar dimension.

A.2 Functionality of the knee joint components

The knee joint has often been called an error of nature, as it consists of many different components making this highly loaded joint prone to defects. However, over the last years this opinion has been more and more revised as knowledge of the functionality of the knee joint components increases.

As already mentioned in the introduction, the menisci were considered to be unnecessary, just like the appendix or the tonsils.

Only as most people who had a complete meniscus *resection*[1], suffered from osteoarthritis several years later, this theory was negated. Today it is a known fact that the osteoarthritis was caused by overloading the cartilage between the femur and the tibia. This overload is caused by the geometry of the femur and the tibia: While the femur condyli are convex, the tibia plateau is nearly flat. The cartilage is not able to deform enough to ensure a sufficient load distribution.

The meniscus has vast load distribution capabilities, as its triangular profile fits perfectly between the femur and tibia (see A.4). The convex shape of the femur presses all parts of the meniscus outwards radially. As both horns are fixed on the tibia plateau, the pressure of the femur is transformed into a radial tension of the meniscus. As collagen has an enormous tensile strength and as 90% of its collagen fibres are located radially, the meniscus can easily cope with this pressure.

Without the cushioning effect of the menisci, the resulting contact area is a very small region, leading to a high local pressure.

Other functions of the meniscus like lubrication, shock absorption and others were also suggested [85].

The femoral condyli are convex, but the curvature is not constant. In 90° of flexion, the radius of the femoral contact area is much smaller than the radius of the contact area in full extension. Even the center of rotation changes with different flexion angles. As the menisci are able to slide over the tibial plateau to a certain extend, they can compensate the different radii and the wandering contact points and ensure a proper weight distribution in all knee joint postures.

The cruciate ligaments are located in the center of the knee joint, between the menisci (see figure A.1). The anterior cruciate ligament (ACL) connects the anterior part of the tibia plateau to the posterior part of the femur. The posterior cruciate ligament (PCL) also connects the tibia to the femur, but from the posterior part of the tibia to the anterior part of the femur. Thus, these ligaments cross each other which explains their name.

There are several possible purposes of the cruciate ligaments. While most authors see them as stabilizing components, others also make them responsible for the "screw home" movement 2.7.2 [2].

[1] complete or partial excision

[2] Nakagawa [59] suggested that the posterior cruciate ligament is never loaded in everyday life.

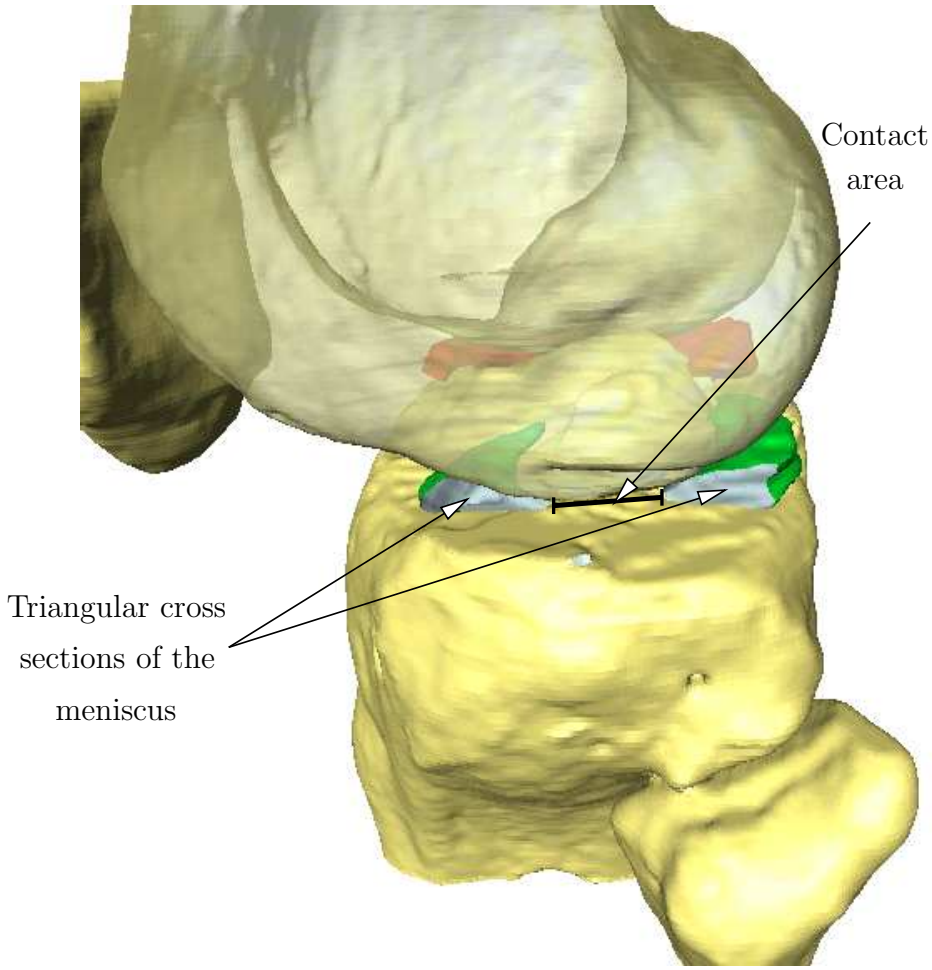


Figure A.4: The triangular cross-sections improve the congruency of the bony surfaces and improve the load distribution.

One common symptom for a broken cruciate ligament is an uncertain feeling while going down the stairs: Without the stability of these ligaments, the femur slides over the tibia plateau a little more whenever the foot is set on the lower stair and the weight of the body pushes the shank forward. This movement is stopped by the patellar ligaments, probably in conjunction with the stabilizing effect of the menisci. This pathology of the ACL or PCL can be detected by the so-called 'drawer-test', where the shank is fixed in horizontal position and the lower leg is pushed in anterior and posterior directions in 90° flexion.

The last knee components that are in the scope of this work are the medial (tibial) collateral ligament (MCL) and the lateral (fibular) collateral ligament (LCL). While the MCL connects the femur to the tibia on the medial side, the LCL connects the femur to the fibula on the lateral side. However, these two ligaments limit the lateral and medial (varus/valgus) movement of the lower leg. An injured MCL or LCL can be diagnosed with the so-called 'varus-valgus-test', where the shank if fixed and the lower leg is pressed inwards or outwards in a slightly flexed position.

A.3 Knee joint anatomy and knee joint tests

Every knee joint component has one or more functions ensuring the knee joint functionality. If a knee joint component is damaged severely, it cannot fulfill its task, so the functionality of the whole knee joint is affected. One example for this kind of pathology would be a completely ruptured LCL or MCL, which would affect the knee joint stability in varus or valgus direction. The so-called "varus-valgus-test" can be performed, with the shank pressed alternatingly in varus and valgus direction to estimate the stability in those directions. If the injured knee joint is significantly more flexible in one direction than the healthy leg, the diagnosis is unequivocal.

If those ligaments are only slightly ruptured, the varus/valgus stability is still given, but the patient will experience pain if the ruptured ligament is loaded. However, this diagnosis is not unequivocal.

If a varus load to the knee joint leads to pains, the LCL is stretched, but the medial meniscus is squeezed. Pains will occur, no matter which one of these components is damaged. In many cases, the patient cannot tell on which side the pain occurs, as the whole knee joint hurts. Thus, further tests have to be performed that load only one of those components. For further details see appendix B.

The cruciate ligament stabilizes the knee joint against anterior/posterior movements. Thus, the function knee joint tests for those ligaments provoke an anterior or posterior translation of the tibia.

This so-called "drawer test" can be done in 90 degrees of flexion. The medical doctor pushes the shank forth and back, feels the movement and the characteristic stop of the movement. If the movement of the injured leg is wider than the movement of the healthy leg and if the movement stops less abrupt, the PCL or the ACL is ruptured, depending whether it was a posterior or an anterior movement.

For both diagnostic approaches a comparison to the healthy knee joint is necessary. Absolute values do not apply, because the difference between two healthy knee joints of two different persons may be larger than the difference between an injured and a healthy knee joint of the same person. A young girl may have a very flexible ACL, thus the stopping characteristic of the drawer test is very soft. Without the comparison to the healthy knee joint this can be misinterpreted as a rupture.

A.3.1 Conclusion: Quality and quantity

The human knee joint is a very complex system with many components and complicated geometries and non-linear physical parameters. All those knee joints are assigned to certain stabilizing tasks, but in many cases, there is no consensus as to the *quantity*. For example, Nakagawa [59] states, that the PCL is never loaded in everyday life, while others [86] find different results.

Thus, the mechanics of the knee joint are not yet researched in detail. However, the knowledge is sufficient to determine the *quality* of the knee joint behavior under distinct load situations: A squeezed meniscus will be pressed outwards. Its radius and internal strain will increase, leading to pain if the meniscus is damaged.

Thus, this qualitative knowledge is sufficient to found the functional knee joint tests. It is *not* sufficient for the graphical animation, as *quantitative* values are needed for a realistic simulation.

B Common manual knee joint tests

The central aspect of the Knee Joint Simulator is the teaching of functional knee joint tests. These tests are performed by a medical doctor who applies certain forces in determined directions. The effects like occurring pains, snapping sounds or movements or increased or decreased mobility, enable the medical doctor to estimate the location and severeness of the injury and thus give a diagnosis.

According to [19], the haptic part of the Knee Joint Simulator allows the user to train the Anterior Drawer Test, the Pivot Shift Test, the Varus-Valgus Test and the McMurray Test. However, Riener [87] describes 86 different tests in seven different families. These families are described here to give a first impression about knee joint tests to readers without medical background. Thus, these explanations are simplified from different sources like [87] [38] [88], [89], [90] and are for explanation purposes only.

B.1 Tests in rotational directions

These eleven tests can be performed in active and passive manner, i.e. either the patient or the medical doctor applies the force to the injured knee joint. The movements are very simple, as the knee joint is moved in only one defined rotational DoF, e.g. only in flexion directions *or* only in varus direction *or* only to an internal rotation.

Some of these tests require the patient in a sitting position, but most of the tests are performed with the patient lying down. One example for this family is the "Passive Internal Rotation of the Knee Joint with a 45° flexion in the hip joint". To perform this test, the patient lies on the back with his hip flexed at 45°. The flexion angle of the knee joint is 90°, so that the heel rests on the couch. The medical doctor stands on the lateral side, stretches the ankle joint to its maximum extension and rotates the lower leg internally. The medical doctor fixates the knee joint with his other hand.

If the patient experiences pain, or if hyper- or hypomobility occurs (compared with the healthy knee joint), the test is positive and a pathology has occurred. Possible pathologies are:

- Illness on the meniscotibial ligament (pain)
- Lesion of the lateral meniscus (pain)
- Lesion or Rupture of the posterolateral capsule (pain or hypermobility)
- Lesion of the ACL (pain or hypermobility)

B.2 Tests in translational directions

These 17 tests are performed in an active or passive manner, mostly with a lying patient and the knee joint at 90° of flexion. Only 2 DoF of the knee joint are tested, because the proximal DoF of the lower leg is too tiny to be noticed. Thus, only the "Drawer Tests" with an anterior or posterior movement and medial or lateral shift tests are performed here. Some of the tests require an internal or external rotation.

One example for this family is the test of the lateral dislocation at 90° of flexion. The patient lies on his back, with his hip flexed at 45°, the knee joint is flexed at 90°. One hand of the medical doctor applies lateral force on the lower leg shortly distal of the cleft of the knee joint. The other hand applies a counterforce on the thigh, shortly proximal of the joint.

If the patient experiences pain, or if the flexibility is more significant than in the other knee joint, the test is positive and one of the following pathologies has occurred:

- Lesion of the lateral or medial meniscus
- Corpus liberum
- Injury at the MCL
- Osteoarthritis

This test was chosen as an example, as it can be used together with the test from section B.1 to give a first diagnosis about a damaged lateral meniscus. For a more proven diagnosis, further tests from section B.5 need to be applied.

B.3 Lachmann tests

The seven different Lachmann tests mainly determine the following pathologies:

- Lesion or over-stretching of the ACL
- Osteoarthritis

The various shades of the Lachmann tests are very different in respect to the patient's position, the doctor's position and the way of grabbing the leg. The force is sometimes applied by the patient, sometimes by the doctor, however, it is always applied in anterior direction. Thus, the Lachmann tests are often classified as "Tests in translational direction" (see section B.2).

B.4 Pivot-Shift-Tests

Riener [87] found 13 different Pivot-Shift-Tests in anterior direction and three in posterior direction. According to the direction, the tests identify ruptures, lesions or general instabilities of the anterior or posterior cruciate ligaments (ACL or PCL).

Like the "Lachmann tests" (see section B.3), all anterior/posterior Pivot-Shift-Tests apply only slightly different forces, but require different positions of the patient and the physician. One example is the "Lemaitre-Test", which is performed with the patient lying on his back. The foot of the leg is rotated in medial direction to force the extended shank to the internal rotation. The physician's other hand applies a medial force to the knee joint. Under this load configuration, a flexion and extension movement is performed.

This test is positive, if the lateral part of the tibia plateau subluxes in the extension movement and repones in the flexion movement. If this occurs, the ACL is probably ruptured.

B.5 Meniscus tests

23 different meniscus tests are known, which is not astounding, since there are many different pathologies of the menisci. Due to that variety, it is hardly possible to find a general description of this family of knee joint tests. As the McMurray test is one of the most complicated and as it is mentioned in the thesis, it is described here. Note, that the Knee Joint Simulator is able to teach the movement actively; novices can touch the shank, follow the movements and let themselves be guided by the robot. This is much more intuitive than descriptions like these:

The patient lies on his back, with hip and knee joints flexed at 90 degrees. One hand of the medical doctor palpates the cleft of the joint, the other one holds the heel. While the position of the knee joint is not changed, the heel is rotated to the internal or external direction. At the same time, the knee joint is extended. This movement can be repeated after releasing the internal or external force and increasing the flexion angle again. Thus, a repeatedly performed McMurray test looks like the physician drawing a circle with the lower leg - however, the internal/external movement is overlooked, but essential.

For testing the lateral meniscus, the shank has to be rotated to the internal direction; accordingly, for testing the medial meniscus, the shank has to be rotated to the external direction.

B.6 Further tests

Nine further tests are described in [87], seven of which are resistance tests, where the medical doctor immobilizes parts of the leg and the patient applies forces against this barrier. If pains occur or if not enough force can be applied by the patient, a pathology is present. These tests - like all tests with active patient cooperation - are not implemented in the Knee Joint Simulator.

The remaining two tests are tests of the patella tendon reflex. This test is very popular: The patient sits with the shank hanging freely in 90° degrees of flexion. The medical doctor uses a hammer to hit the patellafibular tendon, which normally causes an extension reflex. This knee joint test is implemented in the Knee Joint Simulator.

Bibliography

[1] W.O. Thompson, F.L. Thaete, F.H. Fu, and S.F. Dye. Tibial meniscal dynamics using three-dimensional reconstruction of magnetic resonance images. *Am J Sports Med*, 19:210–210, 1991.

[2] Statistisches Bundesamt. Kostennachweis der Krankenhäuser. http://www.destatis.de, 2003.

[3] Statistisches Bundesamt. Diagnosedaten der Krankenhauspatientinnen und -patienten. http://www.destatis.de, 2002.

[4] D. Heim, A. Weymann, U. Loeliger, and P. Matter. Epidemiology of winter sport injuries. *Z Unfallchir Versicherungsmed.*, Suppl(1):16–31, 1993.

[5] K. Steinbrück. Epidemiology of sports injuries – 25-year-analysis of sports orthopedic-traumatologic ambulatory care. *Sportverletz Sportschaden*, 13(2):38–52, 1999.

[6] Robert Riener and Rainer Burgkart. A survey study for the development of virtual reality technologies in orthopaedics. In *9th Annual Medicine Meets Virtual Reality (MMVR) conference*, pages 407–409, 2001.

[7] G. Fischer, A. Kuhlmey, K. Lauterbach, R. Rosenbrock, F. Schwartz, P. Scriba, and E. Wille. Bedarfsgerechtigkeit und Wirtschaftlichkeit. http://www.epilepsie-online.de, 2000/2001. Band III: Über- Unter- und Fehlversorgung.

[8] D.F. Regulla and H. Eder. Patient exposure in medical x-ray imaging in europe. *Radiat Prot Dosimetry*, 114(1-3):11–25, 2005.

[9] E.M. Bitzer, H. Dörning, and F.W. Schwartz. *Der Erfolg von Operationen aus Sicht der Patienten*. Gmünder Ersatzkasse, Gottlieb Daimler Strasse 19, 73529 Schwäbisch Gmünd, Germany, 1998.

[10] G.M. Lawson and R.W. Nutton. A prospective audit of knee arthroscopy: A study of the accuracy of clinical diagnosis and therapeutic value of 325 knee arhtroscopies. *JR Coll Surg Edin*, 40(2):135–137, April 1995.

[11] J. Noble. Unnecesary arthroscopy. *J Bone Joint Surg*, 74:797–798, November 1992.

[12] K.-H. Kleinmann, B. Markefka, and G. Holfelder. Zusammenfassung der Potentialermittlung für Hüft- und Kniegelenkoperationen in Allgemeinkrankenhäusern der Bundesrepublik Deutschland. *Orthopädie Mitteilungen*, pages 445–448, May 1996.

[13] R.A. Gurtler, R. Stine, and J.S. Torg. Lachman test revisited. *Contemp Orthop*, 20:145–154, February 1990.

[14] R.A. Rubinstein, K.D. Shelbourne, J.R. McCarroll, C.D. VanMeter, and A.C. Rettig. The accuracy of the clinical examination in the setting of posterior cruciate ligament injuries. *Am J Sports Med*, 22:550–557, July-August 1994.

[15] B.S. Gerber, I.G. Brodsky, K.A. Lawless, L.I. Smolin, A.M. Arozullah, E.V. Smith, M.L. Berbaum, P.S. Heckerling, and A.R. Eiser. Implementation and evaluation of a low-literacy diabetes education computer multimedia application. *Diabetes Care*, 28:1574–1580, July 2005.

[16] WisdomKing.com, Inc, 4015 Avenida de la Plata, Unit 401, Oceanside, CA-92056. *Funtional Model of the Knee Joint.*

[17] Immersion Cooperation, 801 Fox Lane, San Jose, CA 95131 USA. *CyberGrasp.*

[18] Stäubli Faverges, F-74210 Faverges, France. *Stäubli RX 90.*

[19] Martin Frey. *Regelungstechnische und biomechanische Beiträge für den Einsatz haptischer Simulatoren in der medizinischen Ausbildung und Rehabilitation.* PhD thesis, Eidgenössische Technische Hochschule Zürich, 2004.

[20] C. Michas, M. Frey, and R. Riener. Akustisches Display zur Darstellung der bei einer Knieuntersuchung auftretenden Geräusche. *Biomedizinische Technik / Biomedical Engineering*, 48:190–191, 2003.

[21] C. Michas and H. Yang. Aufbau eines akustischen Displays zur Darstellung der bei einer Knieuntersuchung auftretenden Geräusche". Technical report, Department of Automatic Control Engineering, TU München, Theresienstr. 90, 80333 München, Germany, 2002.

[22] Mindscape Germany GmbH, Dunantstr. 8, D-52064 Aachen, Germany. *BodyWorks 6.0.*

[23] Swarm Interactive, 9503 NW 52 Court, Sunrise, FL 33351, USA. *Inside the knee joint.*

[24] Nucleus Medical Art, 1275 Shiloh Road N.W., Suite 3130, Kennesaw, Georgia 30144, USA. *Knee Injury - Torn Meniscus (Cartilage) with Arthroscopic Surgery.*

[25] Montana Spine Center, 500 W. Broadway, 3rd Floor, Missoula, MT 59802, USA. *Animation of sewing the edges of a torn meniscus.*

[26] Human Anatomy, Inc., 4500 9th Avenue NE Suite 300-6, Seattle, WA 98105, USA. *Interactive Knee.*

[27] National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD 20894, USA. *The Visible Human Project.*

[28] Center for Human Simulation, University of Colorado, UCHSC at Fitzsimons, Center for Human Simulation, Mail Stop F435, P.O. Box 6508, Aurora, CO 80045, USA. *The Visible Human thigh & knee.*

[29] GJ Schreppers, AA Sauren, and A Huson. A numerical model of the load transmission in the tibio-femoral contact area. *Proc Inst Mech Eng [H]*, 204:53–59, 1990.

[30] Judith R. Meakin, Nigel G. Shrive, Cyril B. Frank, and David A. Hart. Finite element analysis of the meniscus: the influence of geometry and material properties on its behaviour. *The Knee*, 10:33–41, March 2003.

[31] L. Blankevoort and R. Huiskes. Validation of a three-dimensional model of the knee. *Journal of Biomechanics*, 29:955–961, July 1996.

[32] Zully M. Maldonado, Jorn Seebeck, Markus O. W. Heller, Doris Brandt, Pierre Hepp, Helmut Lill, and Georg N. Duda. Straining of the intact and fractured proximal humerus under physiological-like loading. *Journal of Biomechanics*, 36:1865–1873, December 2003.

[33] P.H. McCrea, J.J. Eng, and A.J. Hodgson. Linear spring-damper model of the hypertonic elbow: reliability and validity. *J Neurosci Methods*, 128:121–128, 2003.

[34] J. Wismans, F. Veldpaus, J. Janssen, A. Huson, and P. Struben. A three dimensional mathematical model of the knee joint. *Journal of Biomechanics*, 13(8):677–685, 1980.

[35] Martin Gohl. Entwicklung und Evaluierung einer Multimodalen Simulationstechnik zur Darstellung von Deformierbaren Objekten. Technical report, Department of Automatic Control Engineering, TU München, Theresienstr. 90, 80333 München, Germany, June 2004.

[36] Doug L. James and Dinesh K. Pai. Accurate real time deformable objects. In *Proceedings of SIGGRAPH 1999*. ACM, 1999.

[37] D. T. Lee and B. J. Schachter. Two algorithms for constructing a delaunay triangulation. *International Journal of Parallel Programming*, pages 219–242, June 1980.

[38] Robert Riener, Jens Hoogen, Rainer Burgkart, Martin Buss, and Günther Schmidt. Development of a multi modal virtual human knee joint for education and training in orthopaedics. *Stud Health Technol Inform*, 81:410–416, 2001.

[39] Rob Kroeger. The ACL-deficient knee.

[40] A.W. Appel. An efficient program for many-body simulations. *J. Scientific and Statistical Computing*, 6(1):85–103, 1985.

[41] G. Miller and A. Pearce. Globular dynamics: A connected particle system for animating viscous fluids. *Computer & Graphics*, 13(3):305–309, 1989.

[42] A. Vapillion, A. Luciano, A. Habibi, and Y. Duroc. A physical model of turbulent fluids. In *Computer Animation and Simulation '95*. Springer Verlag, 1995.

[43] R. Szeliski and D. Tonnesen. Surface modeling with oriented particle systems. In *Computer Graphics, Proc. SIGGRAPH '92*, pages 185–194. ACM, 1992.

[44] Jean-Christophe Nebel. Keyframe interpolation with self-collision avoidance. Computer Science Department, Glasgow, G128QQ, UK.

[45] Andrew Witkin, Kurt Fleischer, and Alan Barr. Energy constraints on parameterized models. In *Computer Graphics, Proc. SIGGRAPH '87*. ACM, 1987.

[46] C. Bonamico, M. Costa, F. Lavagetto, and R. Pockaj. Real-time mpeg-4 facial animation with 3D scalable meshes. *Signal Processing: Image Communication*, 17(9):743–757, October 2002.

[47] Yan Li, Feng Yu, Ying-Quing Xu, Eric Chang, and Heung-Yeung Shum. Speech-driven cartoon animation with emotions. Technical report, Microsoft Research, Microsoft Research Center China, 3F Sigma Center, No49 Zhichun Road, Beijing 100080, China, 2001.

[48] Peter-Pike Sloan, Charles F. Rose, and Michael F. Cohen. Shape and animation by example. Technical report, Microsoft Research, Microsoft Cooperation, One Microsoft Way, Redmond, WA 98052, July 2000.

[49] Charles Rose, Michael F. Cohen, and Bobby Bodenheimer. Verbs and adverbs: Multidimensional motion interpolation. *IEEE Computer Graphics and Applications*, 18(5):32–41, /1998.

[50] Gino Van Den Bergen. *Collision detection in interactive 3D environments*. Morgan Kaufmann Publishers, Elesevier, 500 Sansome Street, Suite 400, 2004.

[51] D. C. Fithian, M. A. Kelly, and V. C. Mow. Material properties and structure-function relationships in the menisci. *Clin. Orthop*, 252:19–31, 1990.

[52] M. Tissakht and A. M. Ahmed. Tensile stress-strain characteristics of the human meniscal material. *J. Biomech*, 28:411–422, 1995.

[53] Dennis Mah, Michael Steckner, Alexandra Hanlon, Gary Freedman, Bart Milestone, Raj Mitra, Himu Shukla, Benjamin Movsas, Eric Horwitz, Pasi Väisänen, and Gerald E. Hanks. MRI simulation: Effect of gradient distortions on three dimensional prostate cancer plans. *Int. J. Radiation Oncology Biol. Phys*, 53(3):757–765, July 2002.

[54] T. Krings, M. H. T. Reinges, S. Erberich, S. Kemeny, V. Rohde, U. Spetzger, M. Korinth, K. Willmes, J. M. Gilsbach, and A. Thron. Functional MRI for presurgical planning: problems artefacts, and solution strategies. *J. Neurol. Neurosurg. Psychatry*, 70:749–760, 2001.

[55] William Lorenson and Harvey Cline. Marching cubes: A high resolution 3d surface construction algorithm. In *Computer Graphics, Proc. SIGGRAPH '87*, pages 163–170. ACM, 1987.

[56] F.L. Bookstein. Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(6):567–585, 1989.

[57] W. Weber and E. Weber. *Mechanik der menschlichen Gehwerkzeuge*. Dietrich Verlag, Göttingen, Germany, 1836.

[58] V. Pinskerova, H. Iwaki, and M.A. Freeman. The shapes and relative movements of the femur and tibia at the knee. *Orthopäde*, 29 Suppl 1:S3–S5, Jun 2000.

[59] S. Nakagawa, P. Johal, V. Pinskerova, T. Komatsu, A. Sosna, A. Williams, and M.A.R. Freeman. The posterior cruciate ligament during flexion of the normal knee. *J Bone Joint Surg Br*, 86:450–456, Apr 2004.

[60] H. Iwaki, V. Pinskerova, and M.A. Freeman. Tibiofemoral movement 1: the shapes and relative movements of the femur and tibia in the unloaded cadaver knee. *J Bone Joint Surg Br*, 82:1189–1195, Nov 2000.

[61] P. Johal, A. Williams, P. Wragg, D. Hunt, and W. Gedroyc. Tibio-femoral movement in the living knee. A study of weight bearing and non-weight bearing knee kinematics using 'interventional' MRI. *Journal of Biomechanics*, 38:269–276, February 2005.

[62] X.M. Li, B. Liu, B. Deng, and S.M. Zhang. Normal six-degree-of-freedom motions of knee joint during level walking. *Journal of Biomechanical Engineering*, 118:258–261, May 1996.

[63] Per Wretenberg, Dan K. Ramsey, and Gunnar Nemeth. Tibiofemoral contact points relative to flexion angle measured with MRI. *Clinical Biomechanics*, 17:477–485, July 2002.

[64] Stephen J. Piazza and Peter R. Cavanagh. Measurement of the screw-home motion of the knee is sensitive to errors in axis alignment. *Journal of Biomechanics*, 33:1029–1034, August 2000.

[65] Édouard Bugnion. *Le Mécanisme du Genou*. Lausanne, 1892.

[66] Erik Dam, Martin Koch, and Martin Lillholm. Quaternions, interpolation and animation, May 1998.

[67] Alan H. Barr, Bena Currin, Steven Gabriel, and John F. Hughes. Smooth interpolation of orientations with angular velocity constraints using quaternions. *Computer Graphics*, 26(2):313–320, 1992.

[68] David Eberly. Rotation representations and performance issues. Technical report, Magic Software, Inc., http://www.magic-software.com, January 2002.

[69] Calin Belta and Vijay Kumar. New metrics for rigid body motion interpolation. In *Proc. Ball 2000 Symp. Commemorating the Legacy, Works and Life of Sir Robert Stawell Ball*, 2000.

[70] W.R. Hamilton. *Lectures on Quaternions: Containing a Systematic Statement of a New Mathematical Method*. Hodges and Smith, Dublin, 1853.

[71] J.E. Bresenham. Algorithm for computer control of a digital plotter. *IBM Systems Journal*, 4(1):25–30, 1965.

[72] B.T. Phong. *Illumination for Computer-Generated Images*. PhD thesis, Department of Computer Science, University of Utah, Salt Lake City, July 1973.

[73] R. Sibson. A brief description of natural neighbor interpolation. In *Interpreting Multivariate Data*, pages 21–36. Wily, Chichester, 1981.

[74] Robert J. Renka. Algorithm 790: Cshep2d: cubic shepard method for bivariate interpolation of scattered data. *ACM Trans. Math. Softw.*, 25(1):70–73, 1999.

[75] Robert J. Renka and Ron Brown. Algorithm 791: Tshep2d: cosine series shepard method for bivariate interpolation of scattered data. *ACM Trans. Math. Softw.*, 25(1):74–77, 1999.

[76] Robert J. Renka. Algorithm 772: Stripack: Delaunay triangulation and voronoi diagram on the surface of a sphere. *ACM Trans. Math. Softw.*, 23(3):416–434, 1997.

[77] Tim Hopkins. ACM collected algorithms. http://www.netlib.org/toms.

[78] Detlef Ruprecht and Heinrich Müller. Free form deformation with scattered data interpolation methods. In *Geometric Modelling (Computing Suppl. 8)*, pages 267–281, Wien, 1993. Springer Verlag.

[79] Carlos Zuppa. Error estimates for modified local shepard's interpolation formula. *Appl. Numer. Math*, 49(2):245–259, 2004.

[80] Samuel R. Buss and Jay P. Fillmore. Spherical averages and applications to spherical splines and interpolation. *ACM Trans. Graph.*, 20(2):95–126, 2001.

[81] Thomas W. Sederberg and Scott R. Parry. Free-form deformation of solid geometric models. In *Computer Graphics, Proc. SIGGRAPH '86*, pages 151–160. ACM, 1986.

[82] Alan H. Barr. Global and local deformations of solid primitives. In *Computer Graphics, Proc. SIGGRAPH '84*, pages 21–30. ACM, 1984.

[83] Demetri Terzopoulos and Dimitri Metaxas. Dynamic 3d models with local and global deformations: Deformable superquadrics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(7):703–714, July 1991.

[84] Hugo Sellheim. *Die Beziehungen des Geburtskanals und des Geburtsobjekts zur Geburtsmechanik*. Verlag von Georg Thieme, Leipzig, Germany, 1906.

[85] V. Vedi, A. Williams, S. J. Tennant, E. Spouse, D.M. Hunt, and W.M. Gedroyc. Meniscal movement. An in-vivo study using dynamic MRI. *J Bone Joint Surg Br*, 81:37–37, Jan 1999.

[86] G. Li, E. Most, L.E. DeFrate, J.F. Suggs, T.J. Gill, and H.E. Rubash. Effect of the posterior cruciate ligament on posterior stability of the knee in high flexion. *Journal of Biomechanics*, 37:779–783, 2004.

[87] Robert Riener. *Techniken der Multimodalen Virtuellen Realität in der Medizin*. Habilitationsschrift, Department of Automatic Control Engineering, TU München, Theresienstr. 90, 80333 München, Germany, 2002.

[88] F. Baumgartl and G. Thiemel. *Untersuchung des Kniegelenks*. Georg Thieme Verlag Stuttgart, 1993.

[89] F.H. Fu, C.D. Harner, and K.G. Vince. *Knee surgery*. Williams & Wilkins, Baltimore, Maryland, 1994.

[90] J.N. Insall (Edt.). *Surgery of the knee*. Churchill Livingstone, New York, 1984.