Navigation for PDT in the paranasal sinuses using virtual views

L.E. (Laura) Visser

MSc Report

Committee:
Dr.ir. F. van der Heijden
Dr. F.J. Siepel
Ir. E. Molenkamp
R.L.P. van Veen, PhD
Dr. M.B. Karakullukcu

ROBOTICS
AND
MECHATRONICS

UNIVERSITY OF TWENTE.

MIRA CTIT
BIOMEDICAL TECHNOLOGY
AND TECHNICAL MEDICINE

Contents

1   Introduction

Cancer in the paranasal sinuses is a serious disease. Successful surgical or radiation treatment in which the tumor will not recur is a challenge. Not only is a wide variety of tumors covered by the description sinus malignancies, each with their own biological characteristics and prognosis[1], but these tumors are also growing in regions where critical structures are very nearby. Because of the presence of these delicate structures, like the orbit, carotid arteries and the skull base, in many cases the margins that can be achieved with surgery and radiation therapy are not sufficient. Incomplete removal increases the chance for the tumor to recur. Due to these difficulties the recurrence-free survival rate after five years is less than 50% [2].
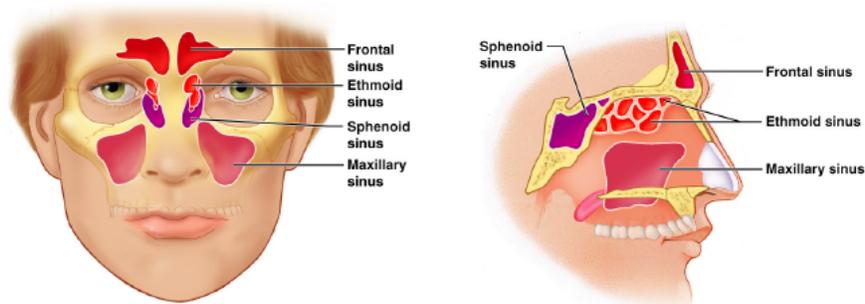
Figure 1: The different paranasal sinuses [3]

Photo Dynamic Therapy (PDT) is a relatively new and lesser-known method, which can extend the margins after surgery and thereby potentially decreases the chance of recurrence. In this treatment, a photosensitizing agent which is localized in the tumor, is activated by a light source. A cytotoxic process occurs resulting in cell death[1], [4]. It can be used to increase the margins after surgery by 5-10mm, thereby destroying the leftovers of the tumor[5]. Radiation is not desirable to repeat in this area, since it poses a high risk for complications, in particular for damage to the optical structures[6]. PDT however, is repeatable, and does not compromise other treatments as radiation or chemotherapy[7].

Without a 3D view of the surgical scene, as is the situation in current endoscopic procedures, the task of navigation is difficult[8], [9]. In order to avoid the critical structures and to enable placing the light source at the most effective location, precise navigation is necessary, and good orientation of the surgeon in the complex cavities is very important. During an open surgery, the view is very intuitive and orientation is obvious. However, when using an endoscope, this information is not directly visible, and the line of sight of the endoscope can be blocked by smoke or bleeding. This increases the task workload for the surgeon of the surgery or treatment.

In order to provide information next to the endoscopic view, during a PDT treatment, the Anthony van Leeuwenhoek Hospital (AvL) uses a Brainlab navigation system. Usually, a small endoscope, for example a 0°, 4 mm Karl Storz rigid endoscope, is used for surgery in the nasal cavities, the visual output of which is shown on a screen. On another screen the output of the electromagnetic navigation system from

Brainlab is given. The usual view of this system consists of the coronal, sagittal, and axial planes of a preoperative CT. In these planes the positions of those surgical tools which are equipped with a sensor are highlighted as well.

Usually the surgeon has to mentally reconstruct the 3D area of the surgery with use of the planes from the navigation system [8], which is a very demanding task, especially for less experienced surgeons [10], [11].

A more intuitive solution to reduce the difficulty of navigation for the surgeon during PDT treatment is to bring back the 3D view of the scene, combined with some information like the nature of surrounding tissue, position of delicate structures and the planned target location for the light source. Different studies support the fact that task workload of a functional endoscopic surgery (FESS) decreases when some sort of augmented reality is used [9]–[11].

The aim of this assignment is to improve navigation ability of the surgeon during PDT through generating a view of the 3D representation of the surgical scene. In order to provide this improvement the design of a system to create a virtual 3D-model based view will be presented. Five criteria or subgoals for the design are identified. The first is for the registration error to be below or equal to that of the current navigation system. Secondly the registration is demanded to take a comparable amount of time to the current registration. The third goal is real-time performance. Fourth is for the system to improve the intuitiveness of the representation of the data during the treatment. And last, the navigation must be based on the position of the endoscope in order to be independent of other tools during the treatment. Successful realization of this design is a step forward towards an endoscopic procedure with the navigation and orientation advantages of an open surgery.

The design, presented in this report, is an electromagnetic guided system that tracks the endoscope and gives a real-time virtual representation of the surgical scene in 2D and a non-real-time representation on a 3D screen. Aspects of the design include real-time tracking of the endoscope, intuitive visual feedback of the position of the endoscope to the surgeon and visible information about critical structures during the treatment.

This report describes the proposed design of the navigation system. It addresses the functional architecture, the implementation and the realization. Experiments to test the quality of the system are reported, and the resulting discussion and conclusion are given. The functional architecture of the system, which can be found in section 2, clarifies the principle of operation, and describes the tasks that need to be fulfilled. It includes the interactions between these tasks as laid down in a functional block diagram. In section 3, the implementation is given, it describes how the tasks and system blocks are made to happen, and how the timing of the process and activation of system blocks is done. As such, the implementation consist of the main loop from input to output, and the needed underlying processes, such as calibration, registration and visualization. In section 4, the realization of the design is presented. Here, the physical realization in terms of hardware components and coding of the software is defined. In order to quantify the quality of the system some experiments are con-

ducted. This is described in section 5. The methods of these experiments consist of visual inspection, error measures, Monte Carlo analyses and real-time performance tests. Results are presented and their interpretation is discussed. In section 7 recommendations for improvements of the system are discussed. Finally, in section 8 the aim of this assignment is reviewed in order to conclude to what extend the resulting system contributes to intuitive navigation during PDT treatment of the nasal cavities.

2 The functional design of a system for navigation in the paranasal sinuses

The principle of operation is defined in this section, and the tasks that need to be done. In order for a system to improve the ability of navigation of a surgeon during PDT, it needs to fulfill some criteria. Tasks that the system should be able to perform follow from the criteria. The proposed design combines these tasks and takes the literature study that is performed beforehand into account.

2.1 From literature research: a novel system for navigation in the sinuses

The literature research in appendix A shows that already some techniques exist to acquire real-time 3D data during a sinus treatment. Most 3D acquisition techniques make use of an algorithm to extract the position and orientation of the camera from the endoscopic video. The depth is also estimated using the video stream, often enhanced with information from preoperative CT or MRI data. Also, quite a few navigation systems for tracking surgical tools inside the sinuses are on the market.

An addition to the current techniques and systems can be made by enabling direct tracking of the endoscopic camera, and using that information as the basis of navigation.

Visualization of the 3D model is often done by placing an overlay on the endoscopic video, or in some cases a non real-time virtual reality model is created for learning purposes.

A new way of displaying all information is a real-time virtual view of the surgical scene, seen from the position and orientation of the real endoscope. Disadvantages of endoscopic video, such as occlusion by smoke or bleeding, are thereby avoided and extra information regarding critical structures can be added. The tumor and the ideal light source location can also be shown in the virtual view during the treatment. Displaying the virtual view in 3D is a novelty in this area and would make the view more intuitive.

Concluding, a novel system for enhancing the view of the surgeon during a sinus treatment will consist of real-time tracking of the endoscope and a virtual view from the viewpoint of the real endoscope containing visual information about the tissue. This view would ideally be displayed in 3D.

2.2 Criteria for a navigation system in the paranasal sinuses

In order to make an improvement to the current situation, the specifications of the new system should meet the accuracy of the current system at least and some new features should be added.

Currently an electromagnetic (EM) tracking system is used for tool tracking during the treatment. In figure 2 an example of such a system, its coordinate system and the

volume in which tracking is most accurate, can be seen. The current system shows the position of pre-calibrated tools real-time in three planes on the CT scan, an example of this situation can be seen in figure 3.
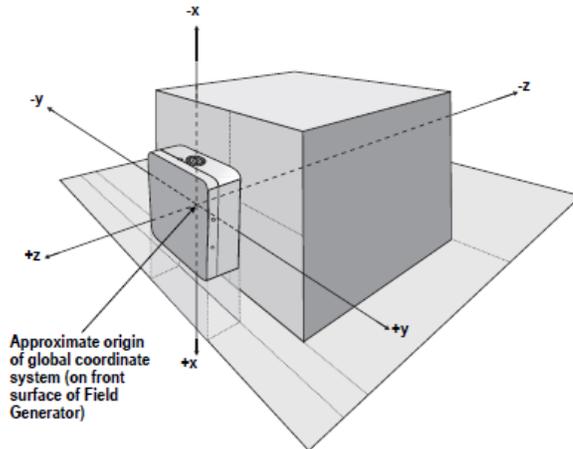
Figure 2: Electromagnetic field generator and its coordinate system and ideal tracking volume

The generator creates a magnetic field, wherein the sensors or tools with integrated sensors can be tracked. All tools and sensors are pre-calibrated with respect to the EM generator. That means that the system gives the position and orientation of the tool tip, or a point on a known location on the surface of the sensor encapsulation.

This system requires registration of the CT data to the patient in the EM coordinate system before surgery. For initialization, some initial points are touched and after that, the tool is used for surface registration.

The new design should have an accuracy that is comparable to that of the current system. At this moment the navigation system can reach an accuracy of less than 1.5mm. This means that the reprojection error of the registration from the patient to the pre-operative CT data is below 1.5mm. The reprojection error is the average of the distance between the points in the CT data and the projections of the measured points onto the CT data. Ideally they would match perfectly, but due to distortions of the magnetic field and other error sources, for example from pinpointing for registration, there is a difference. The smaller this difference is, the better the registration is assumed to be.

The registration at the start of the surgery should not take considerable longer when there is made use of the new system. Registration with the current system takes about 5 minutes. The new system should take comparable time, or even better it should be possible to make a registration during the original registration.

For a PDT treatment usually the light source is attached to the endoscope, with the bulb diffuser about 15 mm ahead of the tip. Ideally, no other tools need to enter the nasal cavity, so the navigation should be done in some way using the position of the endoscope.
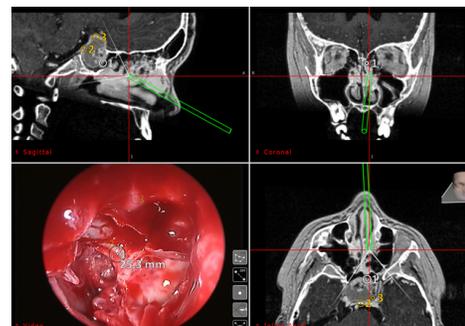
Figure 3: Example of the display during sinus surgery

The overall aim of this project is to display more information in a more intuitive way and thereby making an improvement to the current situation.

Concluding, the criteria for the proposed system are the following:

- Registration error is comparable to the error of the existing system
- Registration should not take more time than the current registration procedure and is preferably possible to be combined with current registration
- The system functions real-time
- Information is shown in a more intuitive way than three planes
- Navigation is based on the position of the endoscope rather than that of a pre-calibrated tool so that only the endoscope needs to enter the cavities

2.3 Needed functionality for a system for navigation in the paranasal sinuses

The designed system should fulfill all criteria as they are defined in section 2.2. The block diagram in figure 4 shows on a high level what actions define the functionality of the system and what data streams run in it.
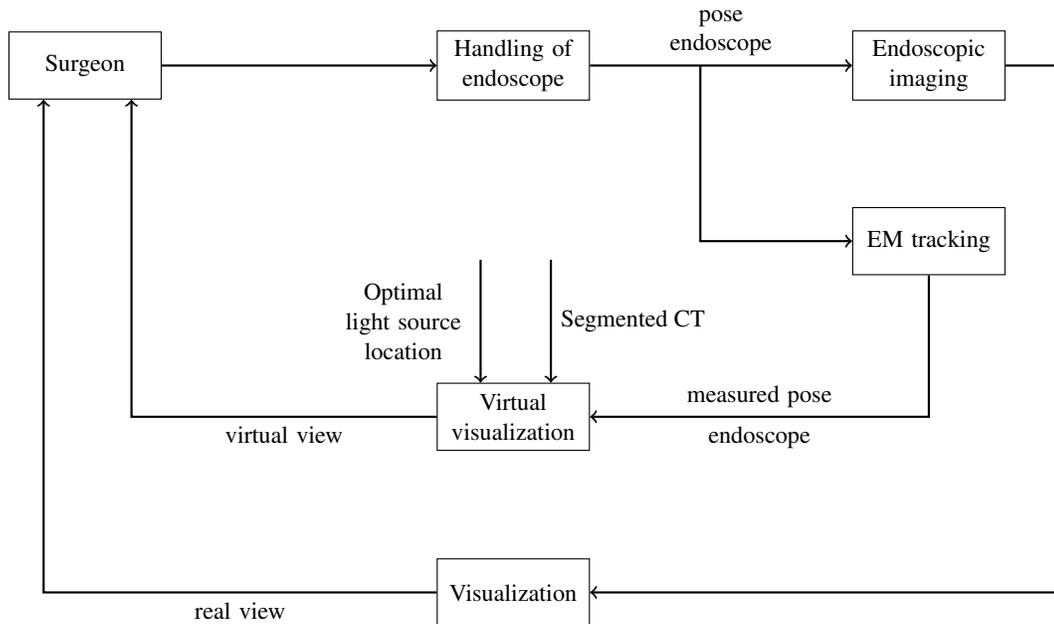
Figure 4: High level block diagram of the functionality of the system

With this diagram, the tasks the system should be able to perform on a functional level are defined. In order to create a system that shows an intuitive view of the real-time tracked endoscope data, these tasks need to be fulfilled according to the diagram:

1 Real time tracking of the endoscope, this includes:
   – Calibration of the camera with respect to the coordinates of the tracking system
   – Registration of the tracking tool with the coordinates of the CT data
2 Intuitive visual feedback of the position of the endoscope to the surgeon, this includes:
   – Creating a virtual view of the scene
3 Visible information about critical structures during the treatment, this includes:
   – Segmentation of critical structures
   – Creation of a virtual 3D model of the head of the patient and the segmented structures inside it

2.4 Functional architecture of a system for navigation in the paransal sinuses

In order to provide the functionality that is demanded in section 2.3, some choices are made regarding the design of the system. This section describes the different possibilities and choices that were made, supported by the literature research that was conducted at the beginning of the project. The literature research can be found in appendix A.

2.4.1 Real time tracking

Real time tracking of the endoscope will be realized using an EM system. Another option was the use of an optical system. Due to two reasons the EM system is chosen: first of all this is the system that is used by the surgeons in the AvL hospital already, so it means that the least additional hardware is introduced in the operation room (OR). Second, this system can easily be read out and it is intuitive to manipulate and make simulations with. Some experiments are conducted on the optical system in order to make a comparison. Information about this system can be found in appendix B. In the literature research can be read that the accuracy of both systems is similar.

Other options of tracking the position of the endoscope in the nasal cavities are based on the video stream of the endoscope itself. These are the so called "volume from view" algorithms. Some options are for example: shape from shading [12]–[14]and structure from motion [15]–[17]. Other techniques are discussed in the literature research in appendix A.3.1 and A.3.2.

It is chosen to work with a tracker system in order to increase accuracy and to have information about the position and orientation of the endoscope when it is not inside the patient. In this way the system does not have to initialize again if the endoscope moves in and out of the patient during a treatment.

2.4.2 Intuitive visual feedback for the surgeon

An intuitive view can be achieved in many different ways. One is for example to augment the endoscopic view with additional information from a virtual view, as described in [18]. Another option is to only show the virtual view from the same viewing point as the endoscope. A third option is to show only a part of the patients head, and from an outside perspective, the movements of the tool inside it. The decision is made to investigate the latter two options. This choice is mainly made because there are less steps required in order to create this, and the functionality of merging the real and virtual endoscopic view can always be added later on if required.

2.4.3 Visible information about critical structures

Critical structures can be shown in various ways as well. Different options are an overlay on the endoscopic images, specific landmarks and distinctive colors for specific tissue in the rendered view [19]–[22]. In this design the choice is made to make a spacial segmentation in the CT data files of the patient. These structures are displayed in different colors and combined with the main patch in the rendered view. All anonymous CT and MRI data is available in the standard Digital Imaging and Communications in Medicine (DICOM) format.

2.4.4 Ideal light source location

In parallel with this project, another project concerning navigation during PDT is running in the AvL hospital. In that research a model is developed for reflection of light inside the nasal cavities, in order to determine the amount of light that reaches a specific surface. With this information the best location for the light source can be defined. The place where the most light hits the tumorous tissue and as little as possible light reaches the healthy tissue is considered the optimum. This location is calculated using a Matlab script and a DICOM data matrix. The location can be imported in the designed system, since it works on the same CT coordinate system. This location can be made visible in the virtual view.

2.4.5 3D visualization

With only a 2D image available, the 3D situation of a scene can mostly be reconstructed in a human head, but some clues are missing in order to make the image appear to be 3D. One important clue is that in the real world each eye sees a slightly different image, this is called stereo parallax. Other missing clues are: movement parallax (if the head moves, different images are seen by the eyes), accommodation (the lenses of the eyes focus on the object where the owner of the eyes wants to look at) and convergence (the eyes converge on the object where the owner of the eyes wants to look at) [23]. 3D visualization techniques mostly focus on bringing back the stereo parallax to the user. Different options are[23], [24]:

- A display combined with the use of glasses:
  - **Stereo anaglyph**
    The differently colored left and right image are superimposed on each other. Typically the contrasting colors are blue and cyan. With corresponding colored glasses for the left and the right eye, the image that is intended for each eye can be seen. Prolonged use of this technique can cause headaches [25].
  - **Polarization**
    The light rays that compose the image are polarized. Light that is polarized in one orientation and light that is polarized in another orientation is superimposed. With the use of glasses that filter the light, the different images for the left and the right eye can be obtained. Disadvantage is the loss of half of the brightness of the image.
  - **Active shutter technique**
    The left and the right image are displayed sequentially. The glasses of the user shut the right and the left eye in matching frequency with the display. These glasses need some sort of power supply and are more heavy than the other two options.
- Autostereoscopic display (without the use of glasses): The left and right image are interlaced per pixel and an extra layer is placed in front of the display. Options are:
  - **Parallax barriers**
    The layer consist of precisely spaced slits. When the user is in a certain position, the slits make it possible to see a different set of pixels with each eye.
  - **lenticular lenses**
    The layer consist of lenses that magnify a different set of pixels when seen from a different angle. When the user is in a certain position one set of pixels reaches the left eye and the other set reaches the right eye, allowing them to see a different image.
- **Head mounted display**
  Two displays are mounted in a headset, showing a different image to each eye. The disadvantage is that the user cannot decide when to look at the real world or at the screen. Either the real world is not visible, or in the case of see-through glasses, the virtual view is always seen against the background of reality.

Since the system will be used in the OR, having an extra set of glasses for the surgeon is not preferred. This might make the vision of the original endoscopic video less clear and therefore might introduce a risk during a surgical treatment. The same holds for a head-mounted display.

The ideal solution would be an autostereoscopic display for multiple users. However, in order to test the functionality of the system, and to limit the costs, a single user display is sufficient.

(a) Lenticular lenses   (b) Parallax barrier
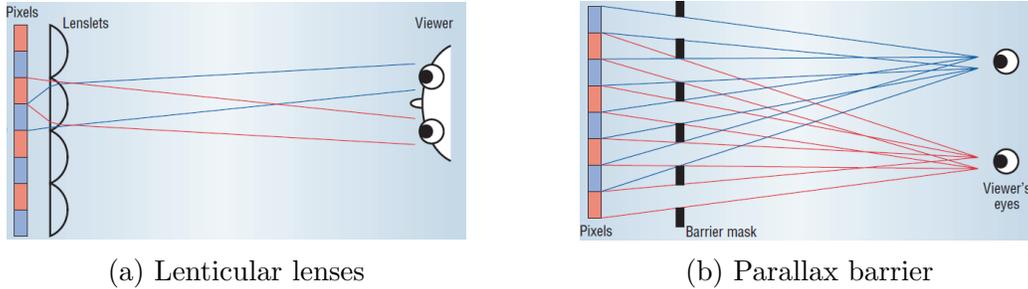
Figure 5: Auto stereoscopic display options [23]

The working principle of the two options for auto stereoscopic displays is given in figure 5.

2.5 Final architecture

The final architecture consist of parts that follow from the requirements and the design choices that are explained in the previous sections. The final design consist of the following parts:

- *Segmentation*
  First the air is distinguished from tissue and later the eyes, optical nerve and the tumorous tissue are separated from the other tissue

- *Camera calibration*
  In this step the camera is linked to the EM sensor that is attached to it, using a checkerboard pattern to define the world coordinates, and taking several images of it, combined with EM measurements. Also, the camera specifications, or 'intrinsics' matrix K, of the endoscopic camera is defined.

- *Point registration*
  An initial registration is made using pinpointing on the model of the patient

- *Surface registration*
  Fine tuning of the registration is performed with a surface registration on a patch of the model of the patient

- *Real time camera positioning*
  Real time EM data and camera images enter the system. The calibration and registration matrices are used to calculate the position and orientation of the camera in the CT coordinate system.

- *Displaying the rendered view*
  Position, orientation and intrinsics of the camera are used together with the model of the patient's head and the segmentation to render a virtual view

13

The different tools, sensors and other relevant parts of the system all have their own coordinate system. In order to be able to talk about the architecture, these different coordinate systems need to be defined first. The coordinate systems of the relevant parts of the system are the following and can be seen in figure 6:

- CS: **Camera Sensor; sensor attached to the endoscope**
  When the EM system is used to find the position and orientation of the endoscopic camera, this camera needs to be equipped with a sensor
- HS: **Head Sensor; sensor attached to the head of the patient**
  The head of the patient should be tracked with a sensor as well, since in case it moves, the images should still be correct. If the head would not been tracked, the it can move with respect to the CT data, thereby destroying the registration of CT data with the EM system coordinates.
- TS: **Tool sensor; sensor of the pre-calibrated tool**
  A pre-calibrated tool, is used for registration of the patient to its CT data via the EM system. This tool is more accurate than the endoscopic tip, because there is no calibration step involved which introduces errors.
- C: **Camera**
  The endoscopic camera that is used throughout the procedure
- EM: **Electromagnetic beacon**
  The reference of the electromagnetic tracking system
- CT: **CT images of the patient**
  The virtual view is based on CT data and is displayed in CT coordinates
- W: **Checkerboard pattern**
  The checkerboard pattern defines the world coordinate system
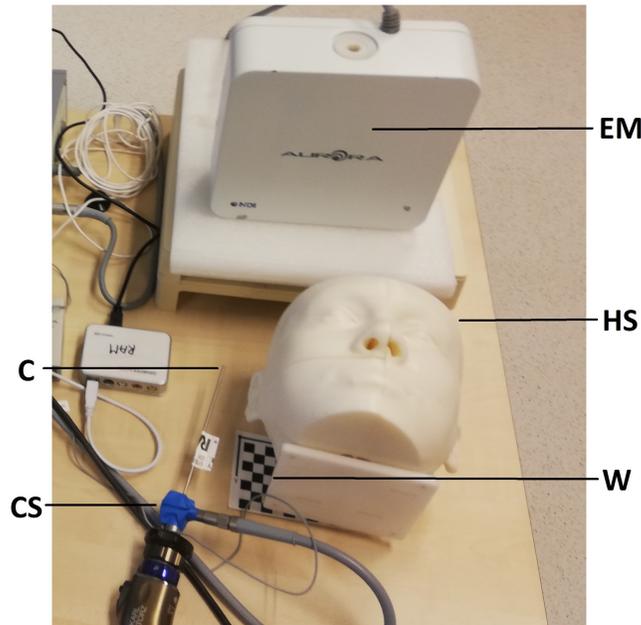
Figure 6: Setup of the system

2.5.1 The transformation matrix

Different coordinate systems can be related to each other via rotation and translation of the origin. A certain point P can be expressed in frame b: P_b, to find this point in system a, P_a has to be determined. If the rotation from b to a, ^aR_b, and the translation of the origin: $^aP_{b,org}$ are known, the new coordinates can be obtained. In figure 7 a 2D example of the transformation of a point P in coordinate system A to the same point expressed in coordinate system B can be seen. The new coordinates are calculated as
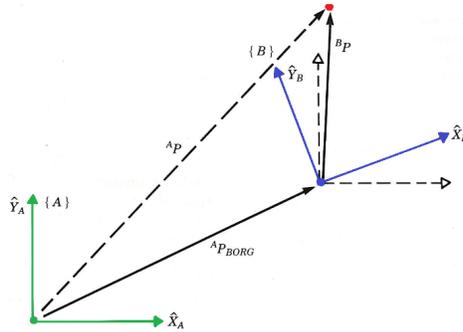
Figure 7: 2D transformation from coordinate system A to coordinate system B

follows:

$$^aP = \ ^aR_b \ ^bP + \ ^aP_{b,org}$$

If the coordinates are expressed as homogeneous coordinates, which means that a 1 is appended at the end of the vectors, the rotation and translation can be combined in one matrix, the so called transformation matrix:

$$^aT_b = \left[\begin{array}{ccc|c} & & & \\ & ^aR_b & & ^aP_{b,org} \\ & & & \\ \hline 0 & 0 & 0 & 1 \end{array}\right]$$

The transformation with use of ^aT_b looks like:

$$\begin{bmatrix} ^aP \\ 1 \end{bmatrix} = \ ^aT_b \begin{bmatrix} ^bP \\ 1 \end{bmatrix}$$

Some important aspects of transformation matrices are the following:

- To find the transform from c to a, when only the transforms ^aT_b and ^bT_c are known, the transformation matrices can be multiplied in order to find the desired transformation: $^aT_c = \ ^aT_b \ ^bT_c$
- The inverse of the transformation matrix gives the inverse relation between the coordinate systems: $^aT_b^{-1} = \ ^bT_a$

If in this report the notation ^aP_b is used, the origin translation $^aP_{b,org}$ is meant.

2.5.2 Relating transformation matrices for the system

The system should be able to display the virtual endoscopic view from the same position and orientation as the real endoscopic view. The only real-time data available during surgery is the 'real' endoscopic view, the EM tracker data of the sensor attached to the endoscope and the data from the head sensor. The desired result is the CT data, seen from the position at which the camera 'looks' at the scene, or the camera coordinate system. The resulting transformation matrix is therefore: $^{CT}T_C$.

In order to realize this functionality an architecture of the system is made that shows how all different coordinate systems will be linked to each other and how the input is being processed and leads to the desired output. In the graph in figure 8 all coordinate systems and their relating transformation matrices are given. The blue arrows indicate data that can be measured directly, the black arrows are for data that can be derived from previous matrices and real-time measurements. The data on the green arrow is available from Matlab's camera calibration. The red arrows indicate transformations that are not directly available. In order to find $^{CS}T_C$ the calibration procedure is needed and for $^{HS}T_C$ the registration procedure must be performed.
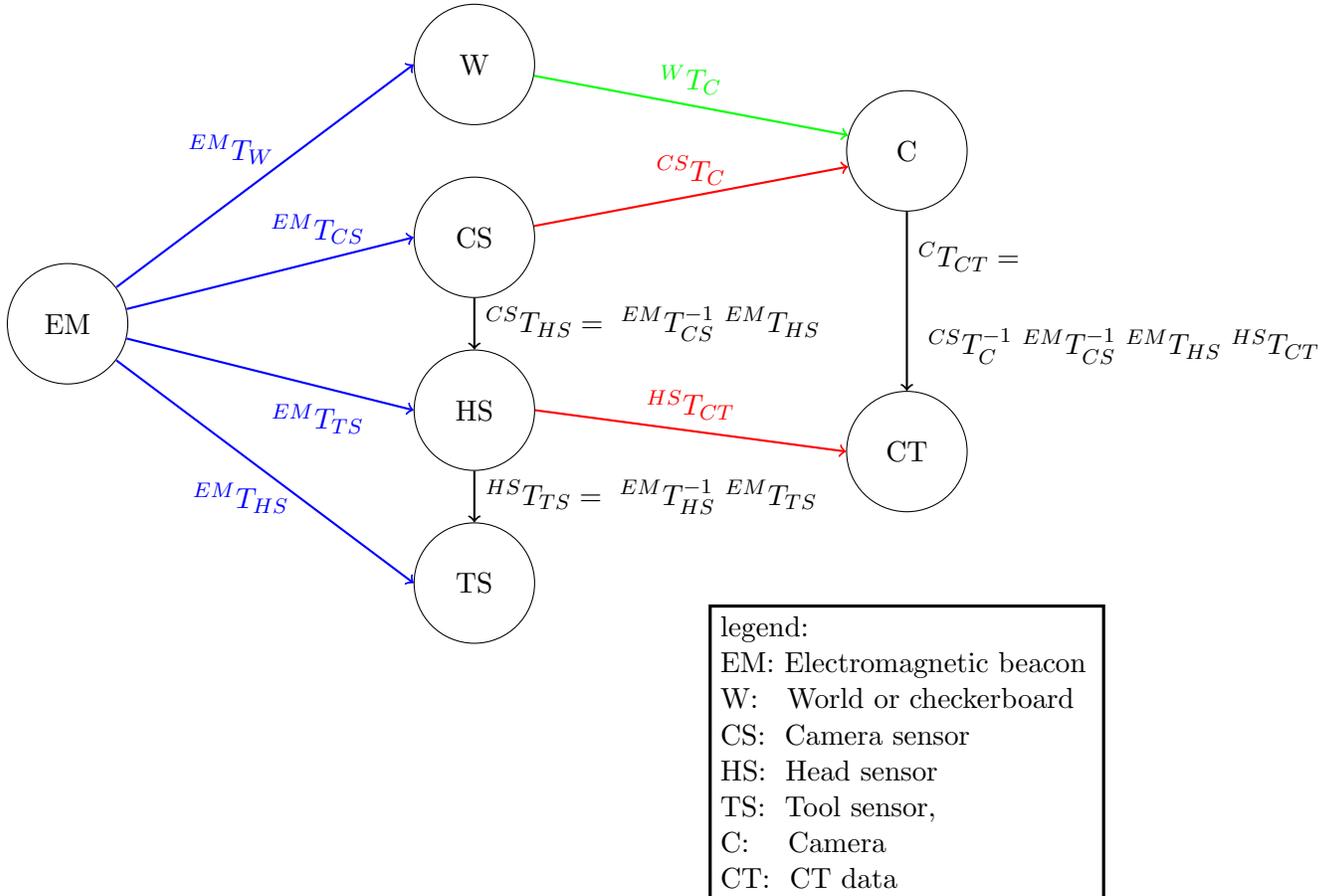
Figure 8: Diagram of the coordinate systems and their relations

2.5.3 Functional block diagram

All parts and coordinate transformations come together in one diagram: the functional block diagram in figure 9. It shows what actions the system takes (blocks) on what data streams (arrows).

Definition of world coordinates

Checkerboard

Pinpointing corners on checkerboard

EM Measurement

Corners in EM coordinates $^{EM}\boldsymbol{p}_m$

Camera

Positioning of camera

Image Acquisition

Images of checkerboard $\boldsymbol{F}_n$

EM measurement

Transform CS during images $^{EM}T_{CS_n}$

Camera Calibration

CT image

Air segmentation

Pinpointing annotation

Registration points in CT coordinates $^{CT}\boldsymbol{r}_i$

Phantom

Pinpointing with probe

Registration points in EM coordinates $^{EM}\boldsymbol{s}_i$

Point Registration

The head sensor in EM coordinates $^{EM}T_{HS}$

Initial registration $^{CT}T_{HSinit}$

Surface segmentation

Definition of registration surface in CT

Surface Registration

Surface scan with probe

EM coordinates when moving over surface

Registration parameters $^{CT}T_{HS}$

Calibration parameters $^{c}T_{CS}$

Real time EM data camera sensor $^{EM}T_{CS}(t)$

Real time EM data head sensor $^{EM}T_{HS}(t)$

Main co-ordinates calculation loop

Position and orientation of endoscopic camera in CT coordinates $^{CT}T_c(t)$

Calibration parameters K

Segmentation of critical structures

3D matrix definition of head

Definition of head model

3D visualization of virtual endoscopic view

Figure 9: Block diagram of the functionality of the system

3 Implementation of the system

This section describes how the tasks and system blocks are executed. The implementation consists of transformations from the EM sensor to the camera coordinate system and the visualization of this data in order to create a virtual view that matches the endoscopic view. Different transformations are made to find the real-time position and orientation of the endoscope. The main loop combines all underlying processes and links the different parts of the system to each other. First this main loop is described and then all underlaying processes are addressed.

3.1 Real time virtual camera positioning: main loop from input to output

When the coordinates from the EM sensors enter the computer, this information has to be transformed into the position and orientation of the tip of the endoscope in camera coordinates. With the help of the calibration and registration steps all bits of information are available, and only a number of calculations needs to be performed in order to find the camera coordinates necessary for a virtual visualization.

When all pre-processes are finished, as can be seen in the block diagram in figure 9, the input of the system consists of:

- Calibration parameters ${}^{C}T_{CS}, K$
- Registration parameters ${}^{CT}T_{HS}$
- Real time EM data camera sensor ${}^{EM}T_{CS}(t)$
- Real time EM data head sensor ${}^{EM}T_{HS}(t)$

One real-time input is the matrix ${}^{EM}T_{CS}(t)$, it gives the camera sensor transform in EM coordinates. To find the transform of the endoscopic tip the camera calibration can be used:

$$^{EM}T_C(t) =^{EM} T_{CS}(t) \; ^{C}T_{CS}^{-1}$$

Another real-time input is the matrix ${}^{EM}T_{HS}(t)$, together with the registration ${}^{CT}T_{HS}$ the coordinates of the camera can be transformed to the CT coordinate system:

$$^{CT}T_C(t) =^{CT} T_{HS} \; ^{EM}T_{HS}(t)^{-1} \; ^{EM}T_C(t)$$

The resulting transform is the desired ${}^{CT}T_C$, which contains the position and orientation of the camera in the CT data coordinate system. This can be used to match the virtual view to the real view of the endoscopic camera. A visualization protocol is designed, which is described in section 3.2.3.

In pseudo code the main calculation is in algorithm 1.

3.2 Underlying processes

The main calculation loop makes use of inputs: calibration parameters ${}^{C}T_{CS}$ and registration parameters ${}^{CT}T_{HS}$. It gives a transformation matrix as output. These inputs

---

Algorithm 1 Main Loop

1: **Initialize** 3D head model
2: **connect** to plusserver
3: **while** stopbutton $\neq$ pressed **do**
4:    **input** from plusserver:
5:         $^{EM}T_{CS}(t)$
6:         $^{EM}T_{HS}(t)$
7:    $^{EM}T_C \leftarrow {}^{CT}T_{HS} \, {}^{EM}T_{HS}^{-1} \, {}^{EM}T_{CS} \, {}^{CS}T_C$
8:    $^{CT}T_{C,ortho} \leftarrow$ **orthogonalize** $^{CT}T_C$
9:    **Output** for visualization $^{CT}T_{C,ortho}$
10: **end while**

---

have to be defined beforehand, and the transformation matrix has to be converted into a visualization afterwards. This section describes the procedures that are used for these pre- and post-processes of the main loop.

Initially a 6-step procedure was developed in order to identify and link all coordinate systems. And to find the position and orientation of the endoscope expressed in the desired coordinates. This procedure started with estimating the transformation from the sensor on the endoscope to the tip: $^{CS}T_t$. This transformation would be used throughout all steps where the endoscope was involved. This however, means a reduction in accuracy in comparison with the use of a pre-calibrated tool of the EM system. It also meant that the registration was dependent on the calibration. Two new, separated, step-by-step procedures were created in which the pre-calibrated tool (TS) is used whenever possible. The new 3-step procedure for calibration is given in section 3.2.1 and the procedure for registration is given in section 3.2.2. The original procedure can be found in appendix C.

3.2.1 Calibration: from EM sensor to camera coordinate system

This procedure describes how to find the transformation from the information of the sensor on the endoscope (CS) to the position and orientation of the endoscopic camera (C).

Step-by-step procedure for camera calibration

1. $^W\boldsymbol{T_{EM}}$ **registration**:
   Registration of the EM coordinate system to the world coordinate system
2. $^W\boldsymbol{T_C}$ **callibration**:
   Camera calibration to register the camera coordinate system to the world coordinate system
3. $^{CS}\boldsymbol{T_C}$ **estimation**:
   Estimation of the transformation from the endoscopic sensor to the camera coordinates using the calculated transforms

Step 1: $^{W}T_{EM}$ registration

The world coordinate system is defined by a checkerboard. In order to link this information to the EM system, the corners of the checkerboard are touched with the pre-calibrated tool and these positions and orientations are saved. An algorithm to calculate the transformation from EM to W is used, for example Procrustes from Matlab [26]. This transformation is found by estimating a linear(translation, reflection, orthogonal rotation, and scaling) fit between the two input datasets. In this case the to sets of data are ^{W}P; corner points in checkerboard coordinates and ^{EM}P; corner points in EM coordinates. The transform is hence:

$$^{W}T_{EM} = Procrustes(^{W}P, ^{EM}P)$$

Step 2: $^{W}T_{C}$ calibration

A number of images of the checkerboard is taken using the endoscope, and the transformation of the sensor on the endoscope is stored simultaneously. The camera calibratior is started in order to find the corners in the images, and calculate $^{W}T_{C}$ [27].

Step 3: $^{CS}T_{C}$ estimation

As written in the previous step, when the pictures are taken also the EM position of the sensor of the endoscope is stored in the matrices $^{EM}T_{CS}(i)$, where i indicates the ith image. With the use of the previously calculated matrices this can be transformed into the position of the tip in world coordinates using the following formula:

$$^{W}T_{CS}(i) = ^{W}T_{EM} \, ^{EM}T_{CS}(i)$$

The transformation matrix to go from endoscopic tip coordinates to the camera coordinates can then be determined for every image:

$$^{C}T_{CS}(i) = ^{C}T_{W} \, ^{W}T_{CS}(i)$$

In order to find $^{C}T_{CS}$, the rotation and translation of the $^{C}T_{CS}(i)$ matrices are averaged.

These steps are translated into pseudocode in algorithm 2.

Algorithm 2 Calibration protocol

1: **input** registration of EM system to checkerboard:
2:        $^{W}P_{CH}$
3:        $^{EM}P_{CH}$
4: $^{W}T_{EM} \leftarrow procrustes(^{W}P_{CH}, ^{EM}P_{CH})$
5: **input** Take images from checkerboard:
6:        $^{EM}T_{CS}(i)$
7:        $F(i)$
8: $^{W}T_{C} \leftarrow Matlab\ camera\ calibration(F(i))$
9: **for** $i \leftarrow 1\ to\ 20$ **do**
10:    $^{W}T_{CS}(i) \leftarrow^{W} T_{EM}\ ^{EM}T_{CS}(i)$
11:    $^{CS}T_{C}(i) \leftarrow^{W} T_{CS}^{-1}(i)\ ^{W}T_{C}$
12: **end for**
13: $^{CS}T_{C} \leftarrow average(^{CS}T_{C}(i))$

3.2.2   Registration

The registration procedure links the coordinate system of the electromagnetic head sensor (HS) to the coordinate system of the CT scan (CT), the result is the transformation matrix: $^{CT}T_{HS}$.

Why using a head sensor

A model of the patients head is created with use of the pre-operative CT-data and the endoscopic tip will be registered to this model. As a part of this procedure, the real face of the patient or the phantom is touched with the endoscope. With one sensor this would be a rigid registration, depending on the location of the head or phantom in world coordinates. But during surgery or treatment, the head might move with respect to the world and the registration might lose accuracy. This problem can be avoided by adding an extra sensor to the system, and defining the registration matrix via this sensor. This sensor can be attached to the head and move along with it during surgery.

Point registration versus surface registration

In general two methods of registration are commonly applied in medical equipment, point-based registration [28] and point cloud based registration [29]. This is also called surface registration.

The first is the easiest to implement. Several points on the surface, that needs to be registered, are predefined. This can be done beforehand by the program or programmer, or the points can be selected in the OR by, the surgeon. In order to make a registration, the tool tip should be used to touch each of the points on the real surface and the EM orientations and positions must be stored. The position that should be touched can be determined by an estimation of the surgeon, but also skin markers or fiducial markers can be used[30], [31]. Then an algorithm is used to estimate the best matrix $^{CT}T_{HS}$, for example again Procrustes. This transform matches the tool

tip positions in the EM-coordinate system, with respect to the head sensor, to the points on the CT-model.

The second method is surface registration, a somewhat more cumbersome method to implement, but also more accurate. For this method a point cloud is defined, consisting of a part of the surface of the CT-model. The tool tip is moved around on this surface, and the data of its orientation and position in EM coordinates, with respect to the head sensor, is stored. Then an Iterative Closest Point (ICP)-algorithm is used to find the best $^{CT}T_{HS}$ to match the two point clouds. ICP-algorithms are used to minimize the distance between two point clouds. This is done by iteratively matching every point in the moving cloud with the closest point in the reference cloud and estimating the transformation that is needed to project the moving cloud onto the reference. This is done until the distance between the point clouds is below a set maximum.

Combination of forces

The best results can be gotten from ICP-algorithms if there exists a reasonable initial transformation [32]. In order to satisfy this condition and get the best possible registration, a combination of both point registration and surface registration is made, consisting of three steps:

Step-by-step procedure for registration

1. **Point registration $^{CT}T_{HS,init}$:**
   Make an initial transformation between the CT and the EM system using four predefined points
2. **Surface registration $^{CT}T_{CT,init}$:**
   Iterate the ICP-algorithm in order to find an accurate match between the surface and the points registered by moving the EM tool on a (predefined) part of the surface
3. **Total registration $^{CT}T_{HS}$:**
   Calculation of the transformation from the head sensor coordinates to the CT coordinates

These steps are translated into pseudocode in algorithm 3.

3.2.3 Visualization

Visualization of the data is done by first segmenting CT and MRI data of a patient, and then generating a patch of the head, as well as a filling on the inside of the head. In order to show an intuitive view, the calculations of the previous section are used to create a virtual endoscope with the same view as the real endoscope. Displaying this view can be real-time in 2D on a computer screen, or afterwards in 3D on the screen of a designated mobile phone. Another option is to visualize a part of the 3D head model from an external point of view and show the endoscope's position and orientation in this view.

Algorithm 3 Registration protocol

1: **initialize** 3D head model
2: **connect** to plusserver

3: Point registration

4: **input** reference points: $^{CT}P_{CS}(i)\ i \in [1,4]$
5: **for** $i \leftarrow 1\ to\ 4$ **do**
6:    **input** from plusserver:
7:         $^{EM}T_{CS}(i)$
8:         $^{EM}T_{HS}(i)$
9:    $^{HS}T_{CS}(i) \leftarrow^{EM} T_{HS}^{-1}(i)\ ^{EM}T_{CS}(i)$
10:   $^{HS}P_{CS,init}(i) \leftarrow^{HS} T_{CS}(i)[1:3,4]$
11: **end for**
12: $^{CT}T_{HS,init} \leftarrow procrustes(^{CT}P_{CS},^{HS}P_{CS})$

13: Surface registration

14: **initialize** Surface patch on 3D head model: pointcloud $^{CT}P_{CS}(n);\ n \in patch$
15: **while** $j < n$ **do**
16:    **input** from plusserver:
17:         $^{EM}T_{CS}(j)$
18:         $^{EM}T_{HS}(j)$
19:    $^{CT}T_{CS,init}(k) \leftarrow^{CT} T_{HS,init}\ ^{EM}T_{HS}^{-1}(k)\ ^{EM}T_{CS}(k)$
20:    $^{CT}P_{CS,init}(k) \leftarrow^{CT} T_{CS,init}(k)[1:3,4]$
21:    $j++,\ k++$
22: **end while**
23: $^{CT}T_{CT,init} \leftarrow ICPalgorithm(^{CT}P_{CS},^{CT}P_{CS,init})$
24: $^{CT}P_{CS} \leftarrow^{CT} T_{CT,init}\ ^{CT}P_{HS,init}$
25: *show* registration on 3D head model
26: **if** User is satisfied **then**
27:    $^{CT}T_{CS} \leftarrow^{CT} T_{CT,init}\ ^{CT}T_{HS,init}$
28:    **output** $^{CT}T_{HS}$
29: **else**
30:    **if** More points required **then**
31:       $j \leftarrow 1$, go back to 15
32:    **else if** New surface registration required **then**
33:       $j,k \leftarrow 1$, go back to 15
34:    **else if** New point registration required **then**
35:       go back to 5
36:    **end if**
37: **end if**

Segmentation

Air segmentation is performed in order to make a differentiation between the head of the patient and the air around it, as well as the cavities inside it. Further segmentation is done to emphasize critical structures and visualize anatomical landmarks of the bone structure. Typically both CT and MRI data are available for a sinus surgery and thus also for the succeeding PDT treatment.

A CT scan determines relative densities and expresses this in numbers on the Hounsfield scale, in Hounsfield Units (HU). Distilled water at standard pressure and temperature is defined as 0 HU. Air is defined as -1000 HU and soft tissue is around $100 - 300$ HU. By using these numbers a segmentation of the CT data can be made by applying a threshold to the data. In figure 10 the difference can be seen between one slice of the original data and one slice where the distinction between air and tissue is made. The edges now show a sharp differentiation which can be used to generate a patch.
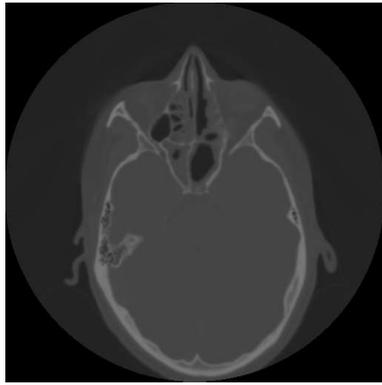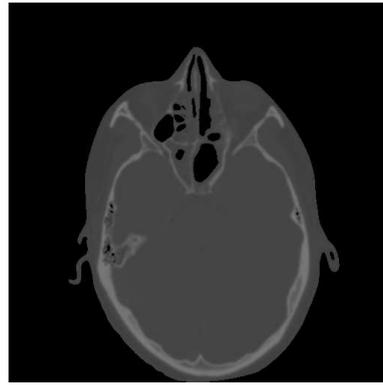
(a) Original data   (b) Threshold is applied

Figure 10: Segmentation of air and tissue with use of a threshold value for the Hounsfield units

For the MRI images, on basis of visual inspection of the surgeon, a segmentation can be made as well. The structures that are chosen to be segmented from this data are: eyes from the CT, carotid arteries and tumor from the MRI.

Patch and filling of the model

The edge between air and tissue is used to generate a patch. The inside voxels of the head are filled in order to give a realistic as possible view of the scene.

Ideal light source location

When the ideal position and orientation of the light source is available it can be written as ${}^{CT}T_{light,ideal}$. If the light source is attached to the endoscope a transform can be determined to relate the camera and the light source: ${}^{C}T_{light}$. Together with the information of the position of the camera ${}^{CT}T_C$, the distance $D(t)$ and difference in

orientation $^{light,ideal}R_{light}(t)$, between the desired transform and the current transform can be calculated:

$$^{CT}T_{light}(t) = \,^{CT}T_C(t)\,^CT_{light}$$

$$D_{light}(t) = ||^{CT}P_{light,ideal} - \,^{CT}P_{light}(t)||_2$$

$$^{light,ideal}R_{light}(t) = \,^{CT}R_{light,ideal}^{-1}\,^{CT}R_{light}(t)$$

These values can be shown on screen in order to inform the surgeon about the proximity of the ideal location.

Virtual endoscopic viewpoint: the pinhole camera model

The virtual endoscopic viewpoint should match the real endoscopic viewpoint, but also give additional information if possible. Information of the real camera is used to define the parameters of the virtual camera.

The so-called pinhole model of an optical camera is used. An image of this model can be seen in figure 11. The camera coordinate system (C) is based on this model. The origin is the focal point of the camera, this is the pinhole or camera center. The z-axis is the optical axis by definition. The x-axis is usually the horizontal axis and the y-axis is vertical. If the image is considered to be a matrix, the x values are the rows and the y values are the columns.
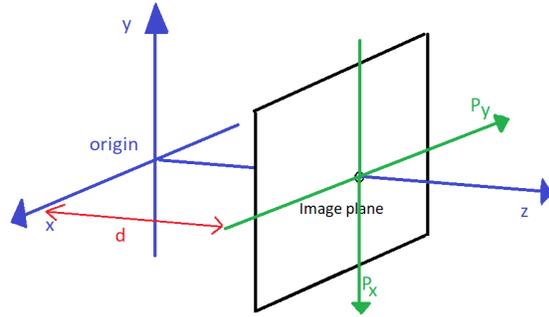
Figure 11: Pinhole camera model
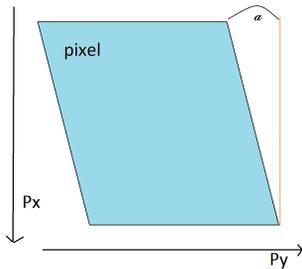
Figure 12: Skew α

The z-axis intersects the image plane orthogonally at distance d, this is the focal distance, thus the image plane is parallel with the xy-plane. If the x and y axis are not orthogonal to each other, a skew coefficient is also included. In figure 12 can be seen what the effect of skew is.

The distance between neighboring pixels is called the pitch and has symbol Δ. Focal distance is expressed in Δ. Other relevant information about the camera are the nonlinear distortions of the lens. Tangential distortion occurs when the lens is not mounted parallel to the image plane, radial distortion is caused by lens imperfections. The position and orientation in the world coordinate frame, rather than in the camera system, give information about the camera as well.

25

Intrinsics

Part of the intrinsics of a camera are combined in the matrix K and include focal length, the principal point and the skew coefficient. The other intrinsics are distortion parameters. An ideal pinhole camera does not have a lens and therefore it does not include distortion. However, the endoscopic camera does have a lens, so distortion parameters are also relevant.

In order to generate a rendered view that matches the real endoscopic view as accurately as possible, all camera intrinsics should be applied to the model. On the other hand, in order to have a view that shows as much and as clear information about the real situation as possible, distortion and skew might be deformations that should not be included. The focal length and the principal point define how wide the angle of view is, that means that these parameters decide what will fit in the image and what will not. To match the view, this should be the same for rendered and real, but maybe it is preferable to see more (or less) than what is on the actual endoscope's display. Because of these considerations, the intrinsics of the camera are calculated, but application is limited to adaptation of the camera angle.

Extrinsics

The extrinsics of the camera viewpoint consist of the position and orientation of the camera in another frame than the camera system coordinates. From camera calibration in Matlab the transformation from C to the checkerboard, or W frame can be found. The calculations of the calibration then make it possible to get the transformation from the camera sensor(CS) to the camera (C). Finally when the registration parameters are applied, the position of the camera in CT coordinates can be found this gives the matrix: ${}^{CT}T_{C,ortho}$. This gives the position and orientation of the virtual camera in the CT coordinate system.

Outside viewpoint: 2D visualization of the endoscope

The extrinsics of the camera are used to position a model of the endoscope within the virtual view. The head is partially shown, in order to be able to see the cavities on the inside. The model of the endoscope is based on the real dimensions of the 0°, 4 mm Karl Storz rigid endoscope.

3D visualization

Since autostereoscopy was chosen to be the most suitable solution for 3D in the OR, an autostereoscopic display was obtained from the company Dimenco [33]. This company is at this point in time one of the only companies investing in 3D technologies. Larger autostereoscopic displays for medical use are under development.

For this project a phone is made available with a layer, consisting of lenticular lenses, on the screen and software to display 3D images and movies.

4 Realization

In order to realize the system, hardware components and programming solutions are chosen, to fulfill the desired tasks. This section describes the physical realization of the building blocks of the system.

4.1 Hardware components

The physical setup consist of several parts: the EM tracker system, a laptop with the Matlab code on it, a phantom of a patients head, an endoscope with camera connected to the laptop via a framegrabber from Terratec, an EM tool, a head sensor and a sensor that is attached to the endoscope. A checkerboard pattern is necessary for the calibration with respect to the world coordinate system. The table where the setup is placed on may not consist of too much metal, since this disturbs the magnetic field of the tracker system. In case this is unavoidable, some plastic spacers are used to move the EM field further away from the table. In figure 13 the combination and connection of the different parts can be seen. All relevant parts will be discussed in the following sections.
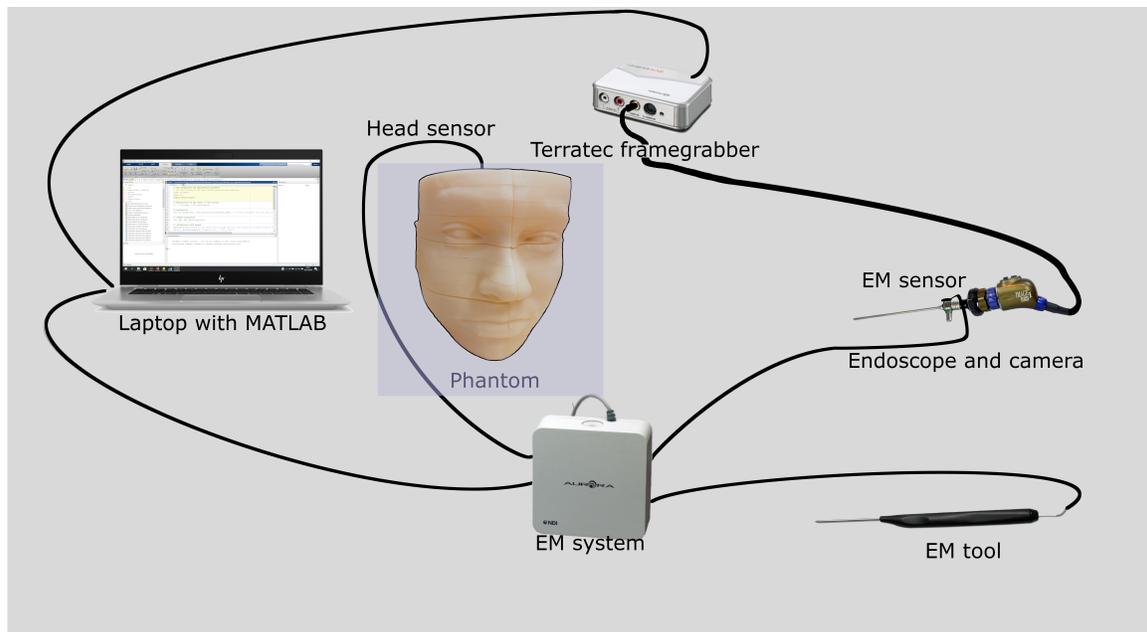
Figure 13: Setup diagram

4.1.1 Phantom

In order to test the code during the process, and to perform experiments, a phantom was created especially for this research at the very beginning of the project. DICOM data of an anonymous patient of the AvL was used for this purpose.

By using the Hounsfield numbers as described in section 3.2.3 a segmentation of the

CT data could be made, differentiating tissue from cavities.

The segmentation was loaded into software called meshmixer, in order to create a 3D model of the head, excluding the cavities [34]. 3D printing is only possible when the model consist of one part, no loose pieces can be involved. Also, if any loose parts exist his is caused by inaccuracies of the model, since there are no floating parts in anyone's head. With use of Netfabb the model was tested for print-ability and the loose parts were detected and deleted.

To make the phantom useful for testing the code, it is an advantage to be able to look inside it, in order to visually match the real and the virtual views. To simplify this, the model was vertically cut in half and horizontally cut in 14 slices, which were all printed separately. Four recesses were made into the model, to be able to put it on four poles and keep the parts together.

Figure 14: The 3D printed phantom

The model is 3D printed using Acrylonitrile Butadiene Styrene plus (ABSplus) P430 in the color ivory. This material is right for an in-vivo experiment because of its light scattering properties, which are comparable to those of mucosal tissue, as measured, on white ABS,with a setup in the AvL hospital [35].

Figure 14 shows a picture of the phantom. In appendix D.2 a manual can be found for creating smooth, evenly spaced, slices of a model with the use of meshmixer.

4.1.2 Checkerboard

The world coordinate system is defined by a 9x6 corners checkerboard pattern with 10x10 mm squares. One direction needs an odd number of squares and one an even, because Matlab can then easily define the x and y directions. This is printed and stuck to a plastic spacer in order to prevent it from moving during the experiments.

4.1.3 Head sensor

This sensor can be stuck to the phantom or the patient's head. It is a small sensor in order to make it easy to attach it to the the head. This sensor is wrapped in a thin plastic thread and the location that it sends to the EM sensor system is at the tip of this thread.

4.1.4 Sensor attached to the endoscope

The sensor that is attached to the endoscope is a standard 6 degree of freedom (DOF) sensor from the EM system. The coordinate system associated with this sensor can be seen in figure 15 [36].

4.1.5 Framegrabber

In order to establish the connection between the endoscope tower and the computer, a "framegrabber" or analog-to-digital converter was used. This is the grabster AV300 from TerraTec, which was connected using a composite cable. It has a framrate up to 29 fps. With software from TerraTec and Matlab code, the "grabbed" images could be loaded into the design.

4.2 Code of the software

This section shows how the different procedures that are described in the implementation are realized in code.

4.2.1 Programming language

As is mentioned before, Matlab is used for all the code parts of the system. This is the most convenient choice, since some parts like the connection protocol with the plusserver and a camera calibration toolbox, were already available in Matlab. Also, code for additional algorithms could easily be found and implemented as will be described in the next sections.

A large part of the designed system consist of matrix calculations of transformations on datasets such as the CT data. Matlab is specifically good at matrix manipulation, its linear algebra routines are powerful and simple to use with often only very few code. For example, matrix calculations with inverses can elegantly be coded, the original implementation $inv(A) * B$ is suggested to be replaced with the faster: $A \backslash B$.

Figure 15: Coordinates of the 6DOF EM sensor, all dimensions are in mm

4.2.2 Patch and filling of the model in Matlab

By looking at the Hounsfield units, with Matlab a segmentation in the original data can be made differentiating outside air and inside cavities from tissue. The edge between air and tissue is used to generate a patch. To minimize computational load in the programs, the number of faces of this patch is downsized with Matlab's algorithm Reducepatch [37]. Since a patch is only a thin layer, with in and outside looking the same, going "through a wall" does not really give

Figure 16: 3D filling of the head model

the idea of looking inside some tissue. To resolve this problem, the algorithm vol3d is used [38].

Dependent on the normalized Hounsfield value a specific shade of a chosen color scheme is used to fill the voxels inside the patch. In figure 16 the model of the head can be seen from above.

4.2.3 Segmentation of critical structures

The segmentation of the chosen structures is done with the program 3Dslicer [39]. This program is able to automatically segment bone, but other sorts of tissue have to be manually segmented. In figure 17 a screenshot of the program, with a slice of a MRI scan with roughly segmented eyes can be seen.
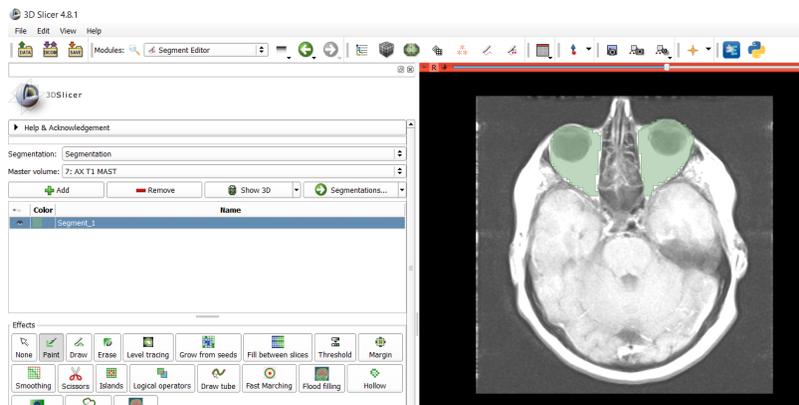
Figure 17: Screenshot of segmentation with 3D slicer

The data of the MRI and the CT scan need to be matched in order to be able to show this information on the same screen, this is also done with 3dSlicer. The segmentation is stored in a stl file, just like the original greyvalues, which is loaded into Matlab to define the CT-model of the critical structures in the head.

The visualization of the segmented tissue is also done by generation of a patch. These patches have different colors from the head patch, in order to emphasize their presence and importance during a treatment. In figure 17 a combination can be seen of the segmentation of the eyes and the tumor, against the background of a transparent patch of the total head.

4.2.4 Calibration

The realization of the calibration protocol follows the steps that are described in section 3.2.1.

Checkerboard registration

First the checkerboard pattern needs to be registered, therefore, the user is asked to fill in how many squares there are in the x and y direction. Then the user is prompted
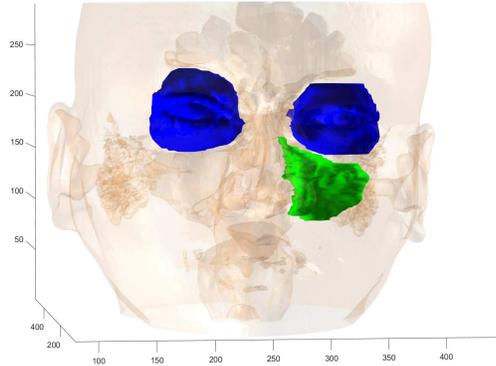
Figure 18: Patches of the segmentation of the eyes and tumor in Matlab

to put the tooltip on the first corner and click "ok". When this is done the user is asked to do the same for the second corner and so on until the 54th(if a checkerboard with 9x6 corners is used). When taking the positions of the checkerboard corners, an average of 50 measurements is taken in order to minimize the positioning error. Matching of the EM data with the world data, that is; the defined coordinates of the checkerboard corners, is performed with the Matlab algorithm Procrustes. This algorithm minimizes the sum of squared errors between the original points and the transformed points after registration: $^{W}P_{reg} = ^{W}T_{EM} \; ^{EM}P_{reg}$.

Taking images of the checkerboard

The next step is to take images from the checkerboard with the endoscopic camera, while saving the position of the sensor on the endoscope. For this the user is asked to hold the endoscope in a specific position, and press "ok". At that moment a snapshot of the endoscopic video is taken, using the framgrabber and the Matlab plug-in, and the position of the sensor is stored in a vector. Before starting to take images, the user has to enter the amount of images that will be taken. When all images are done, the user is asked to start the camera calibration app of Matlab and enter the images in there. When the calibration is performed, usually some images are left out because they are not good enough.

Combination of inputs

All pieces of information are combined in the final part of the code, where the actual calibration calculation takes place. The user has to enter the numbers of the images that are used for camera calibration. The tracker data should only be used from the instances of which there is a picture as well. The transformations from the camera coordinates to the camera sensor coordinates are calculated via the world coordinate system defined by the checkerboard.

The points $^{W}P_{def}$ are the defined corner coordinates in the world coordinate system of

31

the checkerboard. The mean squared error gives the distance between the individual origins of $^{CS}T_C(i)$ and their average. If all calculated matrices lay far away from the average, the average is less likely to be correct. If only one value is added or deleted, the mean changes a lot. A high RMSE has therefore less impact if more measurements are taken. Adding or removing one measurement, even if it has a large differentiation from the average, has less effect on the outcome.

4.2.5 Registration

Point registration

For the initial point registration, four points are chosen on the surface of the patch of the head. For accuracy it is recommended to choose feature points of the face, like eye corners, mouth corners, sides of the nostrils etcetera. This is in order to be able to point at the exact same location on the phantom as is indicated on the patch. In figure 19a an example of four points is shown. When the user wants to define the points himself a separate program can be started and the user can indicate as many points on the patch as he desires for the registration. The predefined points on the CT patch are matched to the stored positions of the tool in EM coordinates with use of the Matlab algorithm Procrustes.
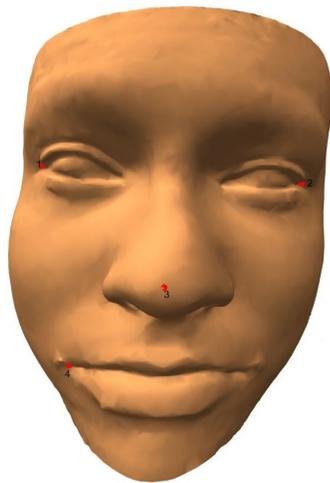
(a) Initialization of surface registration with 4 points
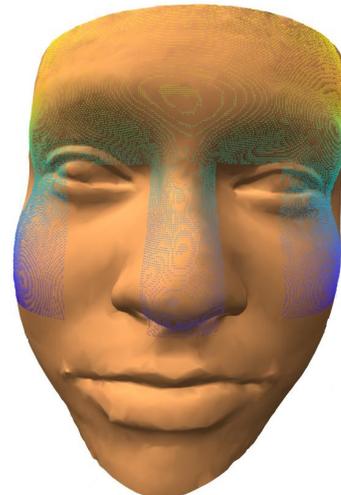
(b) Part of the patch that is used for surface registration

Surface registration

When the point registration is done, the surface registration is started. For this a separate patch is loaded that only consists of a specific outside part of the face. This patch can be seen in image 19b. The user moves the tool around on this area of the head of the patient and the positions of the toolsensor $^{EM}T_{CS}$ are stored in a vector, just like described in the pseudocode. Then the patch and the stored transforms have to be matched using an ICP algorithm. Different algorithms exist with different

functionalities. For example the Matlab algorithm pointCloud finds a rigid transform without scaling, as well as rigidICP but ICP_finite has more options. All three, and a combination of the rigid and the scaled algorithms were tested and compared in order to get the most accurate results. These results were very much equal for affine transformations, so the ICP_finite algorithm is used with the option "affine" [40].

4.2.6 Virtual endoscopic viewpoint

The virtual view can be seen as the view through a pinhole camera, and as such it has intrinsics and extrinsics that can be set.

Intrinsics

The only part of the camera intrinsics that can be easily influenced in Matlab is *camva*, the viewing angle of the camera. To determine this value, the intrinsics that follow from the camera calibration in section 3.2.1 are used. In figure 20 a schematic representation of the camera intrinsics is given. It can be seen that the viewing angle α depends on the focal length expressed in pixels: $d\Delta$ and the image size $N\Delta$. Where Δ is again the pitch, or distance between pixels in mm. This distance is assumed to be equal for the x and y direction, so $\Delta_x = \Delta_y = \Delta$.
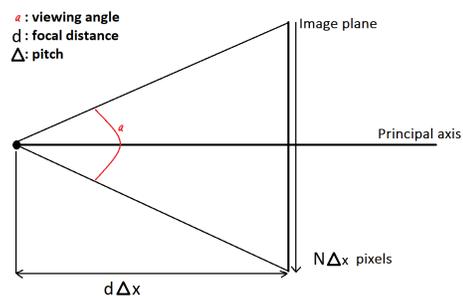
Figure 20: Schematic representation of camera angle α in x-direction

To calculate the viewing angle a simple application of Pythagoras' theorem in both directions suffices:

$$\alpha_i = tan^{-1}\left(\frac{1/2N_i\Delta}{d_i\Delta}\right) = tan^{-1}\left(\frac{N_i}{2d_i}\right) \qquad i = x, y$$

In Matlab only one value can be set for *camva*, so the maximum of α_x and α_y is used.

Extrinsics

In Matlab a pinhole camera model is present of which the extrinsics can implicitly be set. This is done via the vectors *campos*, *camtarget* and *camup*. These respectively hold the position of the camera, the target where the camera is "looking at" and the vector that defines which axis is pointing upwards.

The position and orientation of the camera is found via the procedure as described in section 3.1 and given in $^{CT}T_C$. In order to apply this information to the camera model in Matlab, the resulting *camtarget* and *camup* need to be defined.

The line going through *campos* and *camtarget* defines the camera axis along which the view is oriented. This corresponds to the orientation of the camera in the CT

scene. Starting from the "zero" orientation, the axes of the camera correspond to the axes of the ct view. In this case $-y$ is pointing upwards and the viewing direction is z. So, the starting values are $camup = [0 \ -1 \ 0]$ and $camtarget = [0 \ 0 \ 1]$ with $campos = [0 \ 0 \ 0]$. In order to create the virtual view that matches the real view, the orientation and translation are applied to the zero position, as follows:

$campos \leftarrow \ ^{CT}T_{C,ortho} \ [0\ 0\ 0\ 1]^T$
$camtarget \leftarrow \ ^{CT}T_{C,ortho} \ [0\ 0\ 1\ 1]^T$
$camup \leftarrow \ ^{CT}T_{C,ortho} \ [0\ \text{-}1\ 0\ 1]^T$

4.2.7 Main loop for real-time 2D endoscopic view

In Matlab a virtual camera is defined and updated with the new extrinsics from the main calculation loop. With the virtual endoscopic viewpoint protocol the 2D view of the virtual camera is shown real-time.

4.2.8 Main loop for real-time 2D visualization of the endoscope in the model

The main calculation loop is used to position the model of the endoscope in the 3D model of a part of the head of the patient. The viewpoint is kept equal throughout the treatment, only the endoscope moves in this virtual view.

4.2.9 Main loop for 3D visualization

For generation of a 3D video, a recording is made of all the transformations of the camera sensor and the head sensor during a user-specified period of time. After the treatment or experiment, a movie can be created that can be played on a 3D screen.

In order to generate 3D data, two images are made, one for the left eye and one for the right. This is done by moving the position of the virtual camera and its *camtarget*. To create 3D content some rules of thumb apply in the case you are taking images of a normal scene. These consist of placing the two camera positions eye distance apart and let the camera axes, the lines through *campos* and *camtarget*, cross on the position of the object that should appear "on screen". Objects that are further away appear to go into the screen, closer objects seem to float in front of it. For endoscopic surgery these rules do not apply, since the field of view is very narrow and the distance between the camera and the objects of interest is very small. The solution is to place the virtual camera's apart by only a small distance and letting the axes cross at some point closer to the camera position. Visual inspection is performed for different settings, and the best looking one is with the distance between the positions 1 voxel and the intersection at 5 voxels from the positions.

The virtual view is shown on the 3D monitor of a mobile phone, which can handle multiple variants of 3D content. The options are right-left(RL) and left-right(LR), top-bottom and bottom-top formats. The two images that are taken with the virtual cameras are stuck together in RL format and stored in a MP4 file during the session. It can be displayed using the 3D media player of the mobile phone afterwards.

5 Experiments

In order to asses the quality of the delivered design, some experiments are conducted. These experiments are: visual inspection, assessment of error measures, Monte Carlo analyses and real-time performance tests.

5.1 Goal

The goal of the assignment is to make a protocol of programs that together form an accurate and intuitive navigation platform for a surgeon during a PDT treatment of the nasal cavities. In order to quantify when the system is accurate, some subgoals are defined with respect to five area's of interest:

1 Accuracy and limitations of tool tracking
2 Accuracy of the registration of an electromagnetic sensor to CT coordinates
3 Accuracy of camera calibration
4 Quality of visualization
5 Real time character of the solution

5.1.1 Accuracy and limitations of tool tracking

The specifications of the EM system define the accuracy of the positions and rotations of the sensors, dependent on the distance from the EM generator. Metal objects in the field can cause a distortion of the signal. To be able to keep track of the precise location of the endoscopic tip, somewhere on the endoscope, which contains a lot of metal, a sensor needs to be attached. At first the extra distortion of the endoscope needs to be determined and then it must be minimized; the associated goal is therefore:

Subgoal 1

Minimizing the influence of the distortion introduced by the attachment of a sensor to an endoscope, which contains a lot of metal

5.1.2 Accuracy of the registration of an electromagnetic sensor to CT coordinates

The accuracy of the current system is about 1,5 mm for registration of the EM tool to the CT coordinates. Other navigation systems for the nasal cavities also have sub 2 mm accuracy[41]. The proposed system should be comparable to existing systems, therefore the subgoal on this area of interest is:

Subgoal 2

Sub 2 mm accuracy for the registration from electromagnetic sensor to CT coordinates

5.1.3 Accuracy of camera calibration

Camera calibration is the procedure of finding the transformation matrix between the camera coordinates and the EM sensor that is attached to the endoscope. The error

for this calibration should be as small as possible, in order to make the total system as accurate as possible. Therefore the demand is sub 1 mm accuracy:

Subgoal 3

Sub 1 mm accuracy for the camera calibration transformation matrix from camera coordinates to camera sensor coordinates

5.1.4 Quality of visualization

The goal for the quality of the visualization is twofold. At one hand the requirement is that the rendered view gives a more clear visualization of the situation than the real view, at the other hand it should be matched accurately to the real viewpoint.

The first part seems to be entirely dependent of the opinion of the user, but can at the same time be quantified by the amount of information that is added to the virtual scene. The second part is dependent on the accuracy of the registration and calibration and can therefore be quantified using the criteria of subgoals 3 and 4 as well as a new criterium of the combined protocol. Since the system will be used in the OR and be looked at by surgeons, the visual comparison of the real and the endoscopic view also can give an indication of the quality.

Subgoal 4a

Improvement of the clarity of the situation in the virtual views compared to the real view

Subgoal 4b

Accurate match between the real endoscopic output and the virtual view

For the 3D view an extra requirement arises. The virtual 3D view of the scene should be generated in a way that the depth of the cavities is realistic. It must give an intuitive representation of the real situation.

Subgoal 5

Intuitive representation of the depth of the cavities when using the 3D visualization

5.1.5 Real time character of the solution

If the solution is real-time, the images appear on the screen without any delay. Because of some delay in the EM system, on which this solution has no influence, and the delay the Matlab code causes, this is not possible. But a maximum can be set to the amount of time one calculation and rendering loop in Matlab might cost, where every loop produces a new image on the screen. The time needed for human visual processing of an image is about 150ms [42], so a loop-time of less than 150ms will suffice.

Subgoal 6

Maximum mean computation time of 0.15 seconds for one iteration of the loop from input to output in Matlab

The registration of the CT scan with the EM system has to be done during surgery, since the patient or phantom is needed to do this. Since the surgery should not be slowed down considerably by the use of this system, a limit for the total time of the registration is set:

Subgoal 7

Total time necessary for registration of the CT data to the head sensor coordinate system less than 5 minutes

5.2 Method

5.2.1 Setup

The hardware components of the setup are described in the section 4.1. The setup consist of: the EM tracker system, a laptop with the Matlab code on it, a phantom of a patients head, an endoscope with camera, connected to the laptop via a framegrabber from Terratec, a checkerboard pattern, an EM tool, a head sensor and a sensor that is attached to the endoscope. In order to check whether the subgoals are met, the following experiments are conducted, with in brackets which subgoal can be tested with this particular experiment:

- Visual inspection (4a, 4b, 5)
- Error measures (2,3)
- Monte Carlo analysis (1,2,3,4b)
- Time performance tests (6,7)

All experiments are carried out in the simulated OR of the University of Twente.

5.2.2 Statistical data processing

All experiments, except the visual inspection, can somehow be quantified. This section describes how the errors are calculated and how they can be interpreted. Also, the input errors for the Monte Carlo Analyses are discussed.

Error measures

When a registration is performed from one coordinate system onto another, the fiducial registration error (FRE) can be calculated. This is a measure that is often used to indicate the accuracy of the registration. It is the root mean square error in alignment of the fiducial points in one coordinate system and the projections from the points in the other coordinate system onto the first. This means that a measure for the difference between reference points and mapped measured points using the estimated mapping is calculated.

If the reference points in coordinate system a are $^{a}P(i)$ and the measured points in

coordinate system b are $^bP(i)$ and the estimated mapping is $^aT_{b,est}$, the estimated reference point positions are:

$$^aP_{est}(i) =^a T_{b,est}\ ^bP(i)$$

Then the FRE for n points is calculated as:

$$FRE = \sqrt{\frac{i=1}{n}\sum_1^n (^aP_{est}(i) -\ ^aP(i)).^2}$$

For all calculations with vectors P the dot-square means the element-wise square of the vector.

Another error measure is the root mean squared error (RMSE) itself. The estimated matrix for camera calibration, is an average of all of the outcomes. With the RMSE of these matrices with respect to the average, a measure for accuracy of the estimate can be given. When it is small, the measurements are consistent and the resulting matrix is very likely to be correct. Whereas if this number is larger, there must be some larger errors in the measurements, causing the resulting matrix to be less accurate. However, the more pictures are taken into account, the less influence the inaccuracy of the individual matrices have as long as the mean error is zero, as will be addressed further in section 5.3.2. If the average matrix ^aT_b and all individual matrices $^aT_b(i)$ are known, when focusing only on the translation, the RMSE for n matrices is:

$$RMSE_{translational} = \sqrt{\frac{1}{n}\sum_{i=1}^n (^aP_b -\ ^aP_b(i)).^2}$$

When for example in a simulation, multiple times an estimate is made of the same matrix for which the true one is known, the RMSE can be calculated in exactly the same way as for the average and the individual estimates. Only now the result gives a measure for the average distance from the true translation.
Until now, only the translation from the transformation matrix is taken into account. When the focus is on the part ^aR_b of the transformation matrix, the change in rotation between the estimate and the real matrix can be found by multiplying with the inverse:

$$^{b,true}R_b =\ ^aR_{b,true}^{-1}\ ^aR_b$$

By transforming this matrix into an euler angle $^{b,true}E_b$ the root mean squared error can be calculated again for n matrices:

$$RMSE_{rotational} = \sqrt{\frac{1}{n}\sum_{i=1}^n (^{b,true}E_b).^2}$$

Error measures for calibration

The FRE for registration of the checkerboard, which is the world coordinate system, with the EM coordinate system, can be calculated using the known locations in world coordinates of the checkerboard corners $^wP_{ch}$ and the reprojected measured points:

$$^WP_{ch,est}(i) =^W T_{EM,est} \; ^{EM}P_ch(i)$$

Then the FRE for k checkerboard corners is calculated as:

$$FRE_{wPch} = \sqrt{\frac{i=1}{k}\sum_1^k (^wP_{ch,est}(i) - \; ^wP_ch(i)).^2}$$

The FRE of the ^WT_C matrices and the estimated matrices using the $^{CS}T_C$ estimate, can also be calculated because there are measurements available for ^WT_C. When ^WT_C are the matrices that are measured and calculated with the camera calibration app, and $^WT_{C,est}(i) =^W T_{EM,est} \; ^{EM}T_{CS,est}(i) \; ^{CS}T_C$ from the camera calibration procedure, the FRE for the n pictures is:

$$FRE_{wPe} = \sqrt{\frac{i=1}{n}\sum_1^n (^wP_{C,est}(i) - \; ^wP_C(i)).^2}$$

The RMSE of the individual resulting matrices $^{CS}T_C(i)$ with respect to the averaged matrix, $^{CS}T_C$, can be calculated. Only the translation is taken into account:

$$RMSE_{csPc} = \sqrt{\frac{1}{n}\sum_{i=1}^n (\; ^{CS}P_C - \; ^{CS}P_C(i)).^2}$$

With a simulation, also the true $^{CS}T_{C,true}$ can be compared to the estimated $^{CS}T_C$ for every m iterations of the simulation.

$$RMSE_{csPc} = \sqrt{\frac{1}{m}\sum_{i=1}^m (^{CS}P_{C,true} - \; ^{CS}P_C(i)).^2}$$

$$RMSE_{csRc} = \sqrt{\frac{1}{m}\sum_{i=1}^m (^{C,true}E_C).^2}$$

Error measures for registration

The point-registration of the phantom to the CT data is performed as described in section 3.2.2, four points are registered in order to initialize the surface registration. The FRE is calculated for the chosen points on the model of the CT data: $P_{CT,ini,ref}$ and the projection of the measured points $P_{HS,ini}$ as points $P_{CT,ini}$. In this case g is

the scaling factor between the model and the EM points. This is used in order to get the error in mm instead of voxel size. The calculations are:

$$P_{CT,ini}(i) = {}^{CT}T_{HS,ini} \; P_{HS,ini}(i)$$

$$FRE_{ctPhs,ini} = \frac{1}{g}\sqrt{\overline{\begin{matrix} i=1 \\ \frac{}{4} \end{matrix} \sum_{i=1}^{n}(P_{CT,ini,ref}(i) - P_{CT,ini}(i))^2}}$$

Surface registration gives an error measure as well, since it tries to minimize the distance between the projection of the measured points and the surface. The error that results from this algorithm can usually be set to a maximum, in order to specify for how long the algorithm should continue with iterating, before accepting the transform. This is usually set to 0.001 mm. This error measure is however not so meaningful since there is no unique match of the measured points and the given surface. The surface consists of many more points than the registration, and since scaling by IPC is allowed, the points can easily be matched to a wrong part of the surface and still be within limits. This is typically caused by an initialization that is far from accurate.

When a simulation is done, the points on the surface where the measured points should be matched to are however known as P_{CT}. For a simulation the registration error between the target locations and the reprojections of those can be calculated. This is called the target registration error (TRE) and computed in the same way as the FRE. The errors should be calculated in mm, so the scaling factor from the head sensor to the CT system should be taken into account. This can be calculated with $\sqrt{det({}^{HS}T_{CT})}$.The points $P_{CT,ref}(i)$ are known from the simulation.

$$P_{CT}(i) = {}^{CT}T_{HS} \; P_{HS}(i)$$

$$TRE_{ctPhs} = \sqrt{det({}^{HS}T_{CT})}\sqrt{\frac{1}{n}\sum_{i=1}^{n}(P_{CT,ref}(i) - P_{CT}(i))^2}$$

With a simulation, also the true ${}^{CT}T_{HS}$ can be compared to the estimated ${}^{CT}T_{HS}$ for every m iterations of the simulation.

$$RMSE_{ctPhs} = \sqrt{det({}^{HS}T_{CT})}\sqrt{\frac{1}{m}\sum_{i=1}^{m}({}^{CT}P_{HS,true} - {}^{CT}P_{HS}(i)).^2}$$

$$RMSE_{ctRhs} = \sqrt{det({}^{HS}T_{CT})}\sqrt{\frac{1}{m}\sum_{i=1}^{m}({}^{CT,true}E_{HS}).^2}$$

Error measures for calibration and registration combined
When a simulation is performed, the RMSE for the difference between the true positions of the camera in the CT coordinate system ${}^{CT}T_{C,true}(t)$ and the estimation

$^{CT}T_C(t)$ can be calculated over the discrete time period [0,T], again using the scaling from CT to mm:

$$RMSE_{ctPc} = \sqrt{det(^{HS}T_{CT})}\sqrt{\frac{1}{T}\sum_{t=0}^{T}(^{CT}P_{C,true}(t) - {}^{CT}P_C(t)).^2}$$

Monte Carlo Analyis

Different sorts of errors negatively influence the correctness of the camera calibration matrix $^{CS}T_C$ as well as the registration matrix $^{CT}T_{HS}$. Monte Carlo analysis is performed in order to quantify how an error somewhere in the inputs can affect the output. The error sources and their magnitude are given in table 1, these are considered to be equal in x, y and z direction therefore only one value is given for translational errors and rotational errors per source. For the EM system this is based on the specifications in the documentation[43].
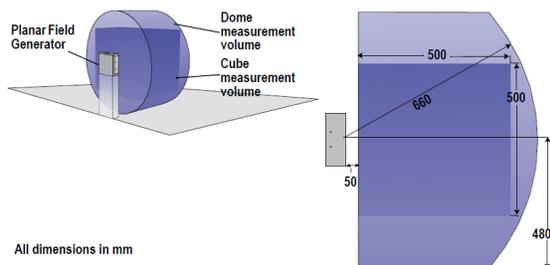
Figure 21: Magnetic filed generator measurement volume [43]

The magnitude for the head sensor is chosen to be within the best tracking volume with respect to the EM system since it can be placed in this volume and during the registration it will not move. The dome volume in figure 21 shows the total measurement volume and the cube close to the generator gives the most accurate tracking volume. The sensor that is attached to the endoscope might escape the cube, so the error magnitude is chosen to be the less accurate one in the specifications.

Furthermore, whenever it was possible there is made use of the fact that more measurements give more accurate results. The new error measure can then be calculated using the central limit theorem for the sample average. When a stochastic variable is distributed with (μ, σ) its average over n measurements will be close to the distribution $(\mu, \frac{\sigma}{\sqrt{n}})$. The errors are assumed to be distributed $(0, \sigma)$ originally, where the σ is given in the reference guide for the EM tracker [43]. When there are more measurements performed, the calculation of the resulting error is given in table 1. For calibration, the checkerboard corner measurements are all performed 50 times. For point registration, there are 10 data points per point on the surface. The head sensor does not move during the registration phase, so the average over the 4 times 10 measurements for the point registration, and over the 1000 measurements for the surface registration is taken.

The camera calibration tool in Matlab gives errors as output, these are round up to the estimated values in the table for error source 1.3.

In order to be able to validate the errors that are input for the Monte Carlo analysis,

also the FRE of the checkerboard registration, and the total FRE of the calibration is calculated during the simulation. The average FRE should equal the resulting FRE of the original calibration step. For registration no such test is available but since the input errors are based on the same assumptions and calculations they are assumed to be correct as long as the errors for calibration are.

When the images are taken from the checkerboard, the endoscope is held by hand. Because this is not a ideal situation and highly dependent on the steadiness of the user, the error sources are chosen to be quite high. The value is based on the resulting FRE from camera calibration, different values are tested in order to make sure the FRE of the simulation is comparable to the FRE of the real calibration.

Table 1: Error sources

Transform	nr	associated errors	translational (mm)	angular (degrees)
$^{EM}T_W$	1.1	Pinpointing error with tool on corners of checkerboard	0.5	n/a
	1.2	EM measurement error	$\frac{0.48}{\sqrt{50}} = 0.07$	n/a
$^{W}T_C$	1.3	Localization of the corners with Matlab's camera calibrator error	0.40	0.005
$^{EM}T_{CS}$	1.4	EM system measurements error and keeping hand still (locations of photo's)	3	2
$^{CT}T_{HS,ini}$	2.1	Pinpointing error with endoscope on surface of phantom	2 voxels	n/a
	2.2	EM measurements error endoscope sensor point registration	$\frac{0.48}{\sqrt{10}} = 0.15$	$\frac{0.3}{\sqrt{10}} = 0.1$
	2.3	EM measurements error head sensor point registration	$\frac{0.48}{\sqrt{40}} = 0.08$	$\frac{0.3}{\sqrt{40}} = 0,05$
$^{CT}T_{CT,ini}$	3.1	EM measurements error endoscope sensor surface registration	0.48	0.3
	3.2	EM measurements error head sensor surface registration	$\frac{0.48}{\sqrt{1000}} = 0.015$	$\frac{0.3}{\sqrt{1000}} = 0.01$

For a Monte Carlo analysis, typically a number of different inputs is defined and varied according to known distributions. The result of the simulation is the distribution of the outcome. If the true outcome is known, a distribution of the error can be drawn from the outputs of the simulation. For both the camera calibration and the surface registration a protocol is developed for generating a known "true" outcome and applying the errors to it in order to find the error distribution of the outcome of the RMSE of the resulting matrices compared to the true ones.

Monte Carlo analysis for camera calibration

For camera calibration the analysis protocol is as follows:

1. Load the raw data that was input for a camera calibration
2. Perform a camera calibration on it to get the original $^{CS}T_C$ matrix
3. Calculate the ground truth(gt) by going backwards from $^{CS}T_C$; using these sub-steps:
   - assume $^{C}T_W(i)$ and $^{EM}T_W$
   - $^{EM}T_{CS,gt}(i) = {}^{EM}T_W\ {}^{W}T_C(i)\ {}^{C}T_{CS}$
   - $^{W}T_{CS,gt}(i) = {}^{W}T_{EM}\ {}^{EM}T_{CS,gt}(i)$
   - $^{CS}T_{C,gt}(i) = {}^{CS}T_{W,gt}(i)\ {}^{W}T_C(i)$
4. Loop for N times:
   - apply errors to the ground-truth of the different data sources
   - calculate the new $^{CS}T_C$
5. calculate the average error between the ground truth of $^{CS}T_{C,gt}$ and the estimated ones
6. calculate the average FRE and compare it to the original FRE of the dataset

Monte Carlo analysis for registration

First the endoscope was used for registration, the Monte Carlo Analysis for that situation can be found in appendix D. Another option for the registration, is to use a pre-calibrated tool instead of the endoscope. This tool has a smaller tip, decreasing the inaccuracy for pinpointing, and does not need to take the error of the $^{CS}T_C$ matrix into account. To see the total impact of the errors on the resulting position during surgery, another Monte Carlo analysis is performed. The steps of this registration are completely the same as for the alternative registration, only the use of $^{CS}T_C$ is omitted, which causes the coordinates to stay in the sensor domain instead of being transformed onto the camera domain.

1. Load the raw data that was input for a registration
2. Perform a registration on it to get the original $^{CT}T_{HS}$ matrix
3. Calculate the ground truth(gt) of the initializations by going backwards from $^{CT}T_{HS}$ using these sub-steps:
   - Assume $^{CT}T_{HS,ini}$
   - $^{CT}T_{CT,ini} = {}^{CT}T_{HS}\ {}^{CT_{ini}}T_{HS}^{-1}$
   - Assume the chosen points on the patch and combine with a random rotation to get: $^{HS}T_{C,ini}$
   - Assume the position of the head sensor and apply a random rotation to it to get: $^{HS}T_{EM,ini}$
3. Calculate the ground truth(gt) of the ICP part by going backwards from $^{CT}T_{HS}$; using these sub-steps:
   - Calculate $^{CT}T_{CT_{ini}} = {}^{CT}T_{HS}\ {}^{CT_{ini}}T_{HS}^{-1}$
   - Choose K random positions on the patch and assign a random orientation to them to get $^{CT}T_C$
   - $^{HS}T_C = {}^{CT_{ini}}T_{HS}^{-1}\ {}^{CT}T_{CT_{ini}}\ {}^{CT}T_C$
   - Assume the position of the head sensor and apply a random rotation to it

to get: $^{HS}T_{EM}$

4 Loop for M times:
  - Apply errors to the ground-truth of the different data sources
  - Calculate the new initialization
  - Calculate the new ICP registration
  - Combine initialization and registration to the transform matrix: $^{CT}T_{HS}$
  - Calculate the RMSE between the original, 'true' $^{CT}T_{HS}$ and the new one, for both the orientation and the position

5 Calculate the total RMSE between the ground truth and the estimated transformation matrices

Monte Carlo analysis of calibration and registration combined

The steps of the total impact analysis are given below. In this case the data from the sensor is assumed to be correct, in order to see purely the impact of the registration and calibration errors.

1 Calculate the average error on $^{CS}T_C$ and $^{CT}T_{HS}$ using the other analysis programs
2 Load some data of a session with the endoscope as well as a calibration matrix and a registration matrix
3 Apply the matrices to get the groundtruth of $^{CT}T_C$ for every datapoint
4 Loop for K times:
  - Apply the errors to $^{CS}T_C$ and $^{CT}T_{HS}$
  - Calculate all the estimated $^{CT}T_C$
  - Calculate the RMSE between the original data points and the estimated ones
5 Calculate the average RMSE between the ground truth and the estimated data

5.2.3   Real time performance tests

Different modes can be chosen for the rendering of images with openGL in Matlab [44]. The modes are: software rendering, hardware rendering and basic hardware rendering. All three modes are tested for speed using visual inspection, the one with the most smooth result is used for the rest of the tests. In Matlab the time a specific part of the algorithm costs can be recorded using the tic-toc functionality [45]. This feature is embedded in the main loop of the real-time navigation code. It records the time from the start of the main loop, until the view is updated at the end of the loop.

The different parts of the code can be analyzed with Matlab as well, using the profiler. This program times the programs run and shows how the spent time is divided over the subprograms or parts of the code.

Also, visual inspection can be used in order to determine whether the video stream is perceived to be smooth or not.

5.3 Results

In this section the results of the system are given, by both the visual outcomes of the programs as well as the numerical outcomes of the Analysis.

5.3.1 Visual outcomes
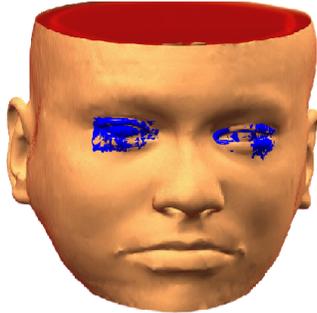
Figure 22: The outside of the patch of the head model
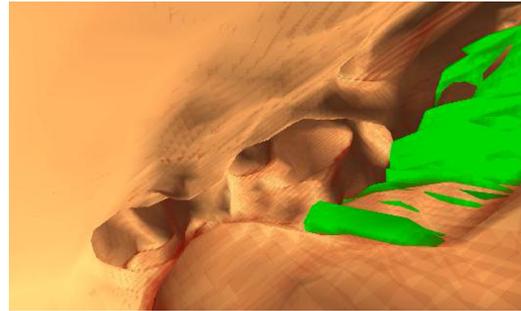
Figure 23: The nasal cavity of the head model

In image 22 the virtual view of the outside of the patch of the head is shown, and in 23 a nasal cavity is shown. The eye orbits and the tumor are segmented. In these cases the endoscopic tool stays within the area that is segmented to be hollow or outside the patient. If the endoscope ends up being at a place where there was tissue at the time of the CT scan, the rendered view will show that the "inside" of the patient is reached. This inside is filled as described before with voxels with colors according to the Hounsfield number. This can be seen in figure 24 in which at the left the endoscope goes through the tissue and at the right the cavities of the patient are still visible.

During surgery the real endoscopic view is displayed next to the rendered view, as shown in figure 25.

The other option, where a part of the head model is visualized as well as the endoscope during the treatment, can be seen in figure 26.
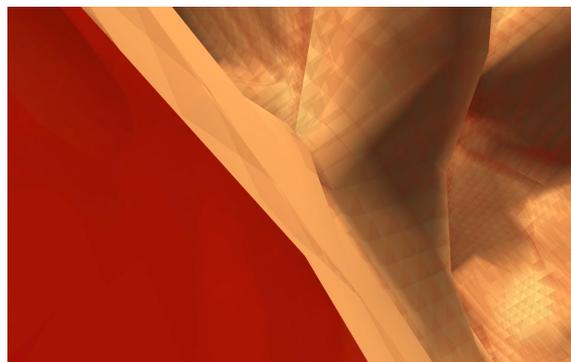
Figure 24: Endoscope goes partly through the wall of the cavity
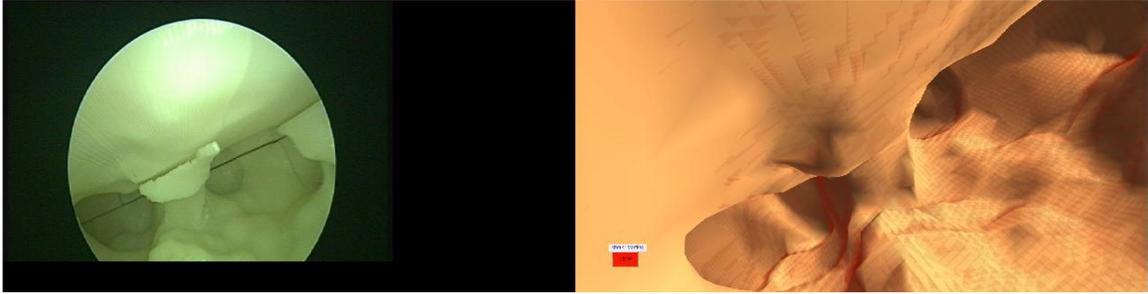
Figure 25: The real and rendered endoscopic view combined
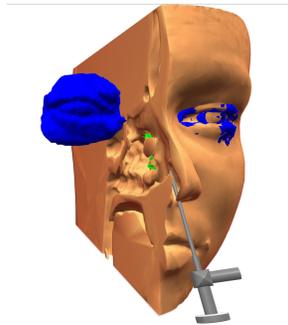
Figure 26: Virtual view from outside the model, with visualization of the endoscope

Point registration can be seen in figure 27, in this case some points are already registered by the user and some still have to be. The registration points chosen by the user are red at first, and when they get processed they will turn green.

Surface registration is shown in figure 28. The initialization step is basically equal to the point registration, only the user cannot choose which points to use. The figure shows the points of the registration according to the initial registration in red and according to the eventual registration in black.
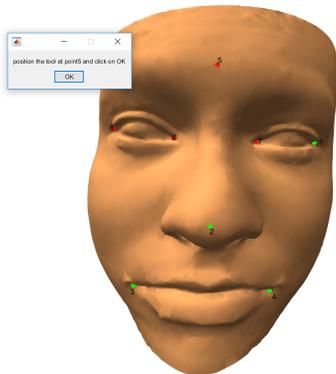
Figure 27: Point registration with 7 points
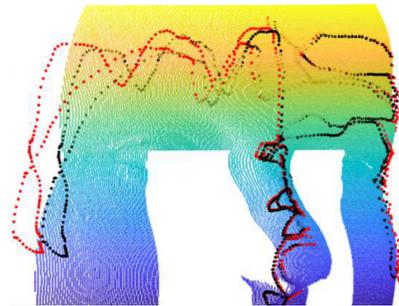
Figure 28: Surface registration with 1000 points

For 3D visualization the two images intended for the left and the right eye are put next to each other, and the 3D software of the mobile phone will use this to alternate the pixels on screen to match the 3D coating that is applied to the screen. In figure 29 the two images can be seen that are the input for the 3D view.
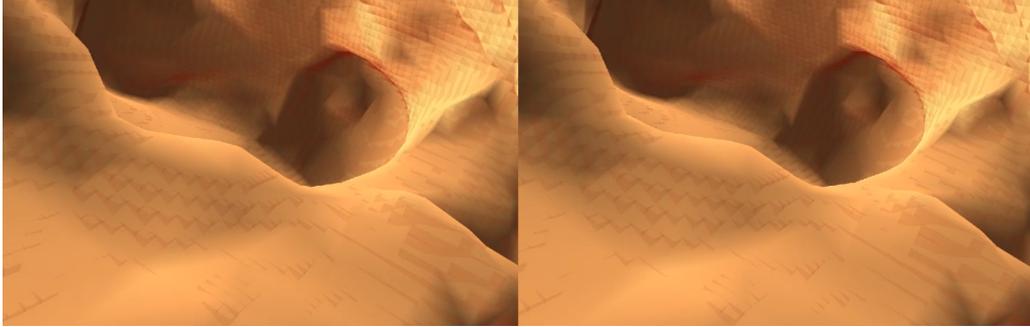
Figure 29: Two images, from a slightly shifted viewpoint, create the 3D view

5.3.2   Numerical results

Error measures for calibration

Camera calibration is performed on a dataset with 135 images and corresponding EM sensor measurements. Two cases are taken into account when making a camera calibration. One is with the use of the plastic coated sensor and the other is with the sensor encapsulated by a thread and separated from the endoscope using a Styrofoam spacer. All error measures are calculated as described in section 5.2.2.

The RMSE of the calibration matrix with respect to the average of the matrices is for the translations:

$$RMSE_{csPc,plasticsensor} = 7.3389 \ mm$$

$$RMSE_{csPc,threadsensor} = 33.5864 \ mm$$

For all other experiments, the plastic coated sensor is used on the endoscope because it minimizes the distortion.

The FRE of the checkerboard registration is $FRE_{wPch} = 0.8600 \ mm$. The FRE of the positions of the $^{W}T_{C}$ matrices is $FRE_{wPc} = 7.6420 \ mm$.

The outcomes of the simulation will be described in section 5.3.2 about the Monte Carlo Analysis.

Error measures for registration

For point registration, the FRE is calculated of the four initialization points that are used to give an initialization of the surface registration algorithm. This is: $FRE_{ctPhs} = 1.9482 \ mm$

Monte Carlo analysis

The Monte Carlo Analyses have been performed as explained in section 5.2.2. The error sources are applied separately and simultaneously in order to find their influence on the average error. The number of iterations of the algorithm is chosen to be 1000. The results of the analysis can be found in table 5.

Camera calibration takes 135 pictures of the checkerboard into account. All errors in the table are the RMSE between the resulting matrix and the true matrix, $^{CS}T_C$ for calibration and $^{CT}T_{HS}$ for registration unless mentioned otherwise. Monte Carlo analysis is also applied to registration with the endoscope and with different parameters. Results of these simulations can be found in appendix D.

In the figures 30a and 30b the distribution of the resulting errors of the calibration matrix $^{CS}T_C$ are shown. For every interval the bar shows how many iterations the RMSE fell within this interval.

(a) Translational errors

(b) Rotational errors
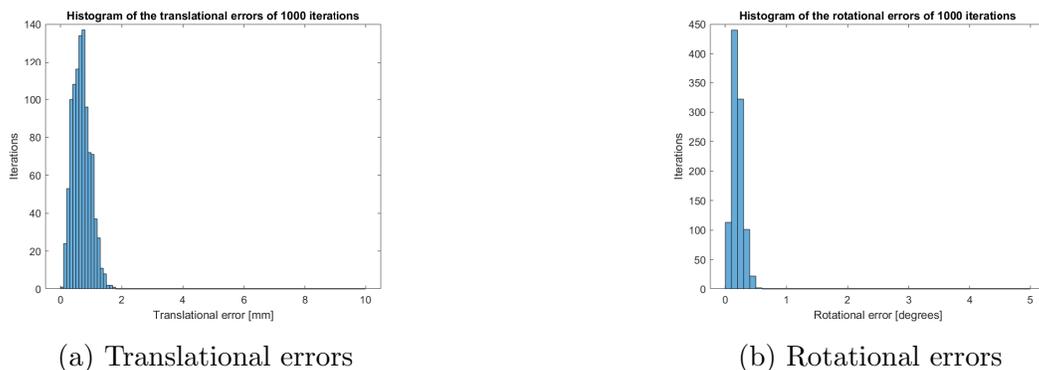
Figure 30: Resulting errors in calibration for 1000 iterations of the Monte Carlo algorithm

In the figures 31a and 31b the distribution of the resulting errors of the registration matrix $^{CT}T_{HS}$ are shown.

(a) Translational errors

(b) Rotational errors
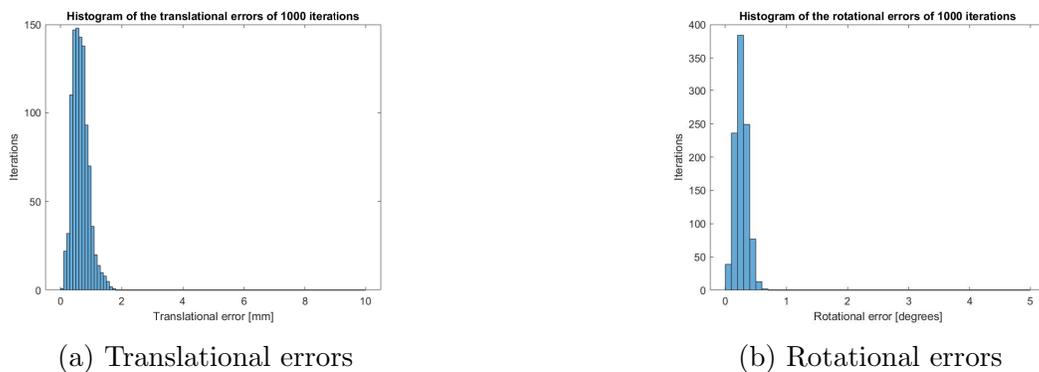
Figure 31: Resulting errors in registration for 1000 iterations of the Monte Carlo algorithm

48

Table 2: Resulting errors from Monte Carlo simulation

nr	associated errors	translational impact (mm)	angular impact (degrees)
1.1	Pinpointing error checkerboard	0.4028	0.1424
1.2	EM measurement error	0.0523	0.0186
1.3	Matlab's camera calibrator error	0.0538	0.0398
1.4	EM system measurements error and keeping hand still	0.5354	0.1390
1.5	**Calibration errors combined**	**0.6889**	**0.2001**
	FRE of the calibration data	7.6420	
	FRE of the calibration simulation	8.2787	
2.1	Pinpointing error with endoscope	0.4643	0.2235
2.2	EM measurements error endoscope sensor point registration	≈ 0	≈ 0
2.3	EM measurements error head sensor point registration	≈ 0	≈ 0
3.1	EM measurements error endoscope sensor surface registration	0.4000	0.1457
3.2	EM measurements error head sensor surface registration	0.0142	0.0052
4	**Registration errors combined**	**0.6550**	**0.2696**
	TRE of the registration simulation	0.5450	

An analysis of the total impact the calibration and registration errors have on the output of the system is performed, with all errors applied. In the figures 32a and 32b the distribution of the resulting errors of the virtual camera extrinsics matrix $^{CT}T_C(t)$ are shown. The average RMSE of the camera position is 1.9689 mm.
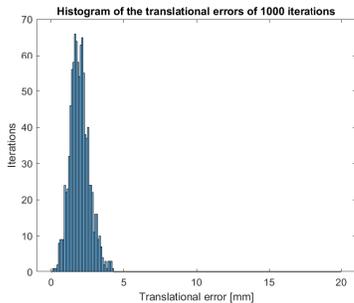
(a) Translational errors
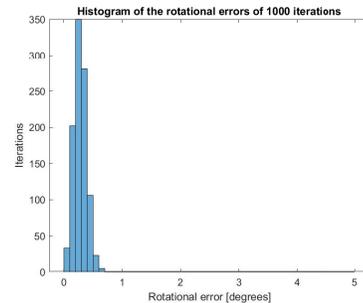
(b) Rotational errors

Figure 32: Resulting errors in the extrinsics matrix for 1000 iterations of the Monte Carlo algorithm

Time performance tests

Basic hardware rendering is used for the time performance tests since this gave the most smooth image sequences.

The average time that was spend in one iteration of the main loop of the real-time navigation, during a session of in total 4 minutes, was 0.0336 seconds, according to Matlabs tic-toc function. However, visual inspection showed that the video stream was not behaving smoothly. This can be due to the fact that the EM system actually does not send as many input signals as the documentation claims, or the fact that these signals are not interpreted fast enough.

Further research, by use of the profiler tool of Matlab gives the result that can be seen in figure 33. This shows that the time consuming part of the code is in the receiving of the data from the EM system. However, the draw command for the images is not shown in this output and that is suspected to take the largest amount of time.

When a recording of the data is made, where no image rendering takes place during the movement of the endoscope, in the resulting video the images are not following one another smoothly.

Even though the video is not completely smooth, the delay does not increase or build up, and the system in considered to function mostly real-time.

Profile Summary

Function Name	Calls	Total Time	Self Time*
...sion_sensor_tTsreg_2sensor_sTc_faster	1	127.869 s	22.400 s
ndi_read_igtl_2sensors	2625	87.902 s	1.461 s
igtlReceivePosition	5250	82.784 s	1.663 s
...ition>ReadOpenIGTLinkMessageHeader	10037	80.006 s	12.208 s
java.io.DataInputStream (Java method)	616272	59.656 s	59.656 s
igtlReceivePosition>socketReadUint64	30111	10.318 s	7.646 s

Figure 33: Output of the profiler tool from Matlab when used on the real-time visualization program

The tic-toc function of Matlab is also used in order to find the total time, one registration procedure takes. This is also dependent on the speed of the user, touching the four initialization points and starting the next step of the code. The total time that was used for registration, averaged over 10 registrations is 2 minutes.

50

5.4 Discussion of the experiments

A system is realized that tracks the endoscope and generates a rendered view that matches the actual view and adds information to it. The resulting system is tested against the goals set in section 5.1. This section will discuss the outcomes of the experiments.

5.4.1 Experimental results versus goals

Per goal is discussed whether the demand is met or if there is more research necessary in order to reach this.

Subgoal 1: Minimizing the influence of the distortion introduced by the attachment of a sensor to an endoscope, which contains a lot of metal

The FRE of the calibration matrix is rather large, about 7mm. This is mostly influenced by the sensor on the endoscope at the time of the pictures, as this is the only input that has an uncertain amount of distortion. To minimize the impact of this inaccuracy, more images and measurements are taken into account. When using 135 images, the Monte Carlo analysis for calibration shows that the impact that the input errors have on the output is below 1 mm. This goal is thereby reached. The plastic coated sensor, which has the least distortion, is used, instead of the thread sensor.

Subgoal 2: Sub 2 mm accuracy for the registration from electromagnetic sensor to CT coordinates

The TRE of the point registration is 0.5450 mm and thereby below 2 mm. Also the RMSE resulting from the Monte Carlo analysis, indicating the influence from the error sources on the output matrix is 0.6550 mm, thus below 2 mm. This goal is achieved.

Subgoal 3: Sub 1 mm accuracy for the camera calibration transformation matrix from camera coordinates to camera sensor coordinates

The total impact of the input errors on the calibration matrix is 0.6889mm as was found with the Monte Carlo analysis. This is below the 1 mm demand from the subgoal, so this requirement is met.

Subgoal 4a: Improvement of the clarity of the situation in the virtual view compared to the real view

The intuitive view of the situation is completely lost with the use of the original system as shown in appendix A. This can be regained by the use of the proposed system. Because the virtual view has the same orientation and position as the endoscopic view, this is very intuitive. And because it is, in contrast to the real endoscopic view, not blocked by smoke or bleeding, it adds information to the surgical scene. Also, the segmentation of the critical structures provides information to the surgeon by showing in a clear way, which places should be avoided. The segmentation of the tumorous tissue can help the surgeon with the positioning of the light source for PDT.

The outside viewing point with the endoscope shown in the model does give a more clear idea of where the endoscope is than the three slices, but since the head model might occlude the position of the endoscopic tip, the virtual endoscopic view is considered more helpful.

Subgoal 4b: Accurate match between the real endoscopic output and the virtual view

The total impact the input errors have on the output is 1.9689 mm. This error is also below the 2mm accurate for current systems, so this is considered to be satisfactory.

Also, the virtual and the real endoscopic view show a good match when placed next to each other. This implies that the system can be used to augment the vision of the surgeon.

Subgoal 5: Intuitive representation of the depth of the cavities when using the 3D visualization

In the 3D view, two images are shown, one for the left eye and one for the right eye. Visual inspection shows that the depth of the video is perceived to be realistic to the user.

Subgoal 6: Maximum mean computation time of 0.15 seconds for one iteration of the loop from input to output in Matlab

The outcomes of the Matlab function for timing imply that the real-time character of the solution is within the proposed limits, the main computation loop takes 0.0336 seconds. This can be considered to be implemented successfully, however the images appear not completely smooth on the screen. So extra effort needs to be put into investigation of the causes of the non-smooth video and into minimizing these problems.

Subgoal 7: Total time necessary for registration of the CT data to the head sensor coordinate system less than 5 minutes

This is considered to be achieved since the average time for a registration is 2minutes.

5.4.2   Limitations

The limitations of the system lay mostly in the applicability in the OR. Some extra steps are needed in order to calibrate the camera and register the EM sensor to the camera output. There has been no test of the system in the OR, so it is difficult to say how much it would influence the workflow of a PDT treatment. Also, since there will be some new software in the OR and even new hardware in the case of a 3D screen, approval of the medical board is needed.

At this moment there has not been the possibility of an actual integration with the perfect place calculation for the light source during PDT treatment.

Another limitation is the rendering speed of the program. Due to the relatively large

computational load of working in Matlab, the program is not as fast that the visualization goes flawless. During surgery it can be deemed unsatisfactory that the images are not always following each other smoothly. Also, the initialization of the figures takes some time for the real-time visualization.

When taking the pictures for the camera calibration, the endoscope is held by hand by the user. Due to this, in case of a not so steady-handed person, the position that is received by the sensor might slightly differ from the position the image is taken from, since there is a small delay between those two. This results in a larger FRE error and a larger RMSE as well.

At this moment the sensor is attached to the endoscope using tape. When a new endoscope is used, the sensor has to be placed again and might be in a different position or orientation than before. So for every surgery, since the endoscope needs to be sterilized after surgery, a new camera calibration has to be performed. When using tape the sensor is only separated from the endoscope with the plastic outside of the sensor, which makes it very close to the metal of the endoscope. This can still cause distortions even though it is less than with the thread sensor.

6   Discussion of the system

The aim of this assignment was : "to improve navigation ability of the surgeon during PDT through generating a view of the 3D representation of the surgical scene". In this report a system is presented that makes the generation of such a view possible. Some criteria were set at the beginning of the project, in order to make sure the system is an improvement of the current situation. These requirements were reviewed in the experiments and translated into subgoals. The system is discussed on basis of the criteria that were set in section 2.2.

Registration reprojection error is comparable to the error of the existing system

This requirement is met since the reprojection error and the error for registration from the CT data to the EM head sensor is within the limit of 2 mm. The experiments showed that the camera calibration error is below 1 mm as well. The total error of the system is also around 2 mm.

Registration should not take more than the current registration procedure and is preferably possible to combine with Brainlab registration

Since the registration is much alike the original registration, it is very likely possible to combine these two. Experiments showed that the registration procedure can be completed within 2 minutes.

The system functions real-time

The main calculation loop of the system can be completed within 0.04 seconds, so the system functions within the time limits that were demanded in the experiments. The video does not seem to be completely smooth, but there is no build up in delay of the frames.

Information is shown in a more intuitive way than three planes

Research confirms that a virtual view is more intuitive than the representation in three planes. The 2D virtual view provides extra information about critical structures and is not obstructed by smoke or bleeding and therefore gives a more clear view than the endoscope. The 3D visualization is even more intuitive because it returns the depth perception of the cavities to the surgeon.

Navigation is based on the position of the endoscope rather than that of a pre-calibrated tool

With the use of the calibration protocol and the registration protocol combined, the navigation is based on the position of the endoscope.

So the system realizes real-time tracking of the endoscope, gives intuitive visual feedback of the position of the endoscope to the surgeon and provides visible information about critical structures during the treatment. Therefor it fulfills the aim of the assignment.

7 Recommendations

Some future work can be done in order to improve the system. This includes performing experiments in the real OR, on a real patient rather than on a phantom. When this is done, the system can also be reviewed in terms of ease of use for the surgeon. For example, questionnaires can be held in order to review the possible reduction in task workload the surgeon experiences. Even cadaver studies could be performed in order to address inattentional blindness of the surgeon due to focus on the highlighted area's and indicated targets, as was suggested in [46].

Integration of the system with the ideal location calculation for the light source during PDT can be realized in the future.

Rendering speed can be minimized by the use of a different programming language or the study of parallelization in Matlab.

Further investigation into the calibration and registration protocol is necessary to improve the total error on the resulting real-time output. For example, calibration performance improvement can be reached when the endoscope is held by a holder instead of by hand. And registration accuracy can be higher when a more accurate phantom or a real patient is used.

In order for this method to be applicable to any surgery, without necessarily having to calibrate again, the sensor should be attached to the endoscope using a holder instead of tape.

The rendered view of the scene can be improved when different rendering techniques or another programming language is used.

The image acquisition from the endoscope can be made simpler and probably faster via a digital connection so that the use of a framegrabber is not needed anymore.

An interesting option to investigate in future research is overlaying the virtual endoscopic video on the real endoscopic video. In that way the real endoscopic video gives more information, but the surgeon does not need to look at two screens.

8   Conclusion

The system presented in this reports fulfills the aim of improving navigation ability of the surgeon during PDT trough generating a virtual view of a 3D representation of the surgical scene. It realizes real-time tracking of the endoscope, gives intuitive visual feedback of the position of the endoscope to the surgeon and provides visible information about critical structures during the treatment.

The system consists of several protocols and subsystems that are created in Matlab, to give the desired representation of the surgical scene. First a calibration protocol is defined, which makes the link between the camera coordinates and the coordinate system of the EM sensor attached to the camera. The next part consist of registration of the CT coordinates to the sensor coordinate system of an EM sensor attached to the phantom. The calibration and registration together provide all information that is necessary to track the position of the endoscopic camera in CT coordinates real-time.

A visualization is made, consisting of a patch of the head and the critical structures, seen from a virtual camera with the same specifications as the real camera. In order to create a virtual view, a segmentation is made between air and tissue, and of the critical structures. Also the filling of the patch is defined for the 3D model. The virtual view is based on the CT coordinate system. A virtual camera is positioned in the CT coordinates according to the real-time EM data and the calibration and registration matrices. The viewpoint of this camera is therefore influenced by the extrinsics, which are position and orientation of the real endoscopic camera. From the intrinsics of the real camera, the viewing angle is taken into account for the virtual view.

A 3D view is created by positioning two virtual cameras very near to each other and thereby creating the different images for each eye. These images are then interlaced on an auto stereoscopic 3D screen.

Outcomes of experiments to qualify the system, show that most of the criteria are met. Navigation is based on the position of the endoscope rather than on a pre-calibrated tool. The camera calibration error is below 1 mm, the registration error is below 2 mm. These values are within the set limits and comparable with the current system. The total error is around 2 mm, which is also comparable to the error in current systems. Registration resembles the current registration protocol and does not take longer than 5 minutes, thus it does not have a large impact on the treatment. The system functions real-time and all calculations of one loop from input to output are performed within 0.04 seconds. The images of the virtual video are not following each other completely smoothly, some positions seem to be skipped. The system shows an intuitive view of the scene by displaying it from the same viewing point as the real endoscope and adding extra visual information to it about the critical structures. In the 3D view, also depth perception is brought back to the surgeon.

The realization of the proposed system is a step towards navigation for PDT that brings back a 3D view of the surgical scene. This view has the intuitiveness of an open surgery and displays additional information regarding critical structures and light source position.

References

[1] L. Caesar, T. E. van Doeveren, I. B. Tan, A. Dilci, R. L. van Veen, and B. Karakullukcu, "The use of photodynamic therapy as adjuvant therapy to surgery in recurrent malignant tumors of the paranasal sinuses", *Photodiagnosis and Photodynamic Therapy*, vol. 12, no. 3, pp. 414–421, 2015, ISSN: 18731597. DOI: 10.1016/j.pdpdt.2015.06.001.

[2] I. Ganly, S. G. Patel, B. Singh, D. H. Kraus, P. G. Bridger, G. Cantu, A. Cheesman, G. D. Sa, P. Donald, D. M. Fliss, P. Gullane, I. Janecka, and S.-e. Kamata, "Craniofacial resection for malignant paranasal sinus tumors: Report of an international collaborative study", no. July, pp. 575–584, 2005. DOI: 10.1002/hed.20165.

[3] S. Vertebral, "The Axial Skeleton Eighty bones segregated into three regions Skull Vertebral column Bony thorax Bones of the Axial Skeleton", 2006.

[4] D. D. Ruysscher, T. J. Dougherty, and S. L. Marcus, "Feature Articles Photodynamic Therapy", pp. 1734–1742, 1992.

[5] B. Karakullukcu, K. V. Oudenaarde, M. P. Copper, W. M. C. K. Robert, M. Wildeman, and I. B. Tan, "Photodynamic therapy of early stage oral cavity and oropharynx neoplasms : an outcome analysis of 170 patients", pp. 281–288, 2011. DOI: 10.1007/s00405-010-1361-5.

[6] S. S., R. K.T., V. M., W. J., and R. J., "Intra-arterial cisplatin and concomitant radiation therapy followed by surgery for advanced paranasal sinus cancer", *Archives of Otolaryngology - Head and Neck Surgery*, vol. 130, no. 8, pp. 948–955, 2004, ISSN: 0886-4470. DOI: 10.1001/archotol.130.8.948LK.

[7] C. Hopper, "Photodynamic therapy : a clinical reality in the treatment of cancer",

[8] M. S. Nosrati, R. Abugharbieh, J. M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, and G. Hamarneh, "Simultaneous Multi-Structure Segmentation and 3D Nonrigid Pose Estimation in Image-Guided Robotic Surgery", *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 1–12, 2016, ISSN: 1558254X. DOI: 10.1109/TMI.2015.2452907.

[9] B. J. Dixon, H. Chan, M. J. Daly, A. D. Vescan, I. J. Witterick, and J. C. Irish, "The effect of augmented real-time image guidance on task workload during endoscopic sinus surgery", *International Forum of Allergy and Rhinology*, vol. 2, no. 5, pp. 405–410, 2012, ISSN: 20426976. DOI: 10.1002/alr.21049.

[10] B. J. Dixon, M. J. Daly, H. Chan, A. Vescan, I. J. Witterick, and J. C. Irish, "Augmented image guidance improves skull base navigation and reduces task workload in trainees: A preclinical trial", *Laryngoscope*, vol. 121, no. 10, pp. 2060–2064, 2011, ISSN: 0023852X. DOI: 10.1002/lary.22153.

[11] L. Li, J. Yang, Y. Chu, W. Wu, J. Xue, P. Liang, and L. Chen, "A novel augmented reality navigation system for endoscopic sinus and skull base surgery: A feasibility study", *PLoS ONE*, vol. 11, no. 1, pp. 1–17, 2016, ISSN: 19326203. DOI: 10.1371/journal.pone.0146996.

[12] L. Lei, J. Li, M. Liu, X. Hu, and Y. Zhou, "Shape from Shading and Optical Flow used for 3-Dimensional Reconstruction of Endoscope Image", *Otolaryngology - Open Journal*, vol. 1, no. 1, pp. 20–23, 2015, ISSN: 24704059. DOI: 10.17140/OTLOJ-1-106.

[13] G.-h. Wang, J.-q. Han, and X.-m. Zhang, "Three-dimensional reconstruction of endoscope images by a fast shape from shading method", *Measurement Science and Technology*, vol. 20, no. 12, p. 125 801, 2009, ISSN: 0957-0233. DOI: 10.1088/0957-0233/20/12/125801.

[14] A. Kumar, Y. Y. Wang, C. J. Wu, K. C. Liu, and H. S. Wu, "Stereoscopic visualization of laparoscope image using depth information from 3D model", *Computer Methods and Programs in Biomedicine*, vol. 113, no. 3, pp. 862–868, 2014, ISSN: 01692607. DOI: 10.1016/j.cmpb.2013.12.013.

[15] S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. H. Taylor, and G. D. Hager, "Image-based navigation for functional endoscopic sinus surgery using structure from motion", vol. 9784, p. 97840V, 2016, ISSN: 16057422. DOI: 10.1117/12.2217279.

[16] S. Rattanalappaiboon, T. Bhongmakapat, and P. Ritthipravat, "Fuzzy zoning for feature matching technique in 3D reconstruction of nasal endoscopic images", *Computers in Biology and Medicine*, vol. 67, pp. 83–94, 2015, ISSN: 18790534. DOI: 10.1016/j.compbiomed.2015.09.021.

[17] T. Wittenberg, C. Winter, and I. Scholz, "3-D Reconstruction of the Sphenoid Sinus from Monocular Endoscopic Views: First results", *Informatik.Uni-Erlangen.De*, no. Bmt, pp. 3–5, 2006.

[18] A. Hattori, S. Naoki, and O. Nobuyoshi, "The development of a Real-Time Image-Guided Surgery System for Stereo-Endoscopic Sinus Surgery", *Studies in health technology and informatics*, vol. 142(5), pp. 112–6, 2009.

[19] D. Hutchison, J. C. Mitchell, X. Pan, and Y. Fan, "Visualization Techniques for Augmented Reality in Endoscopic Surgery", vol. 1, pp. 404–414, 2016, ISSN: 0302-9743. DOI: 10.1007/978-3-319-43775-0.

[20] R. U. Thoranaghatte, J. G. Giraldez, and G. Zheng, "Landmark based augmented reality endoscope system for sinus and skull-base surgeries.", *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2008, pp. 74–7, 2008, ISSN: 1557-170X. DOI: 10.1109/IEMBS.2008.4649094.

[21] G. P. Moustris, S. C. Hiridis, K. Deliparaschos, and K. Konstantinidis, "Evolution of autnomous and semi-autnomous robotic surgical systems: a review of the literature", *The international journal of medical robotics + computer assisted surgery : MRCAS*, vol. 7, no. April, pp. 375–392, 2011, ISSN: 1478-596X. DOI: `10.1002/rcs`.

[22] S. Kapakin, "The paranasal sinuses: Three-dimensional reconstruction, photo-realistic imaging, and virtual endoscopy", *Folia Morphologica (Poland)*, vol. 75, no. 3, pp. 326–333, 2016, ISSN: 16443284. DOI: `10.5603/FM.a2016.0006`.

[23] A. Neil, "Autostereoscopic", pp. 31–36, 2005.

[24] D. Hamill, S. Banerjee, and G. Verghese, "Introduction to 3DTV and 3D-Cinema", no. 1 992, pp. 3–4, 2001.

[25] L. Lipton, "Foundations of the stereoscopic cinema - a study in depth", *Annals of Physics*, vol. 54, p. 258, 1969.

[26] *Procrustes*. [Online]. Available: `https://nl.mathworks.com/help/stats/procrustes.html` (visited on 03/22/2017).

[27] Matlab, *Camera calibrator*. [Online]. Available: `https://nl.mathworks.com/help/vision/ug/single-camera-calibrator-app.html` (visited on 04/01/2017).

[28] J. M. Fitzpatrick, J. B. West, and C. R. Maurer, "Predicting Error in Rigid-Body Point-Based Registration", vol. 17, no. 5, pp. 694–702, 1998.

[29] G. Krell, N. Saeid Nezhad, M. Walke, A. Al-Hamadi, and G. Gademann, "Assessment of Iterative Closest Point Registration Accuracy for Different Phantom Surfaces Captured by an Optical 3D Sensor in Radiotherapy", *Computational and Mathematical Methods in Medicine*, vol. 2017, 2017, ISSN: 17486718. DOI: `10.1155/2017/2938504`.

[30] C. R. Maurer, J. Michael Fitzpatrick, M. Y. Wang, R. L. Galloway, R. J. Maciunas, and G. S. Allen, "Registration of head volume images using implantable fiducial markers", *IEEE Transactions on Medical Imaging*, vol. 16, no. 4, pp. 447–462, 1997, ISSN: 02780062. DOI: `10.1109/42.611354`.

[31] A. Mert, E. Knosp, S. Wolfsberger, G. R. Sutherland, and L. S. Gan, "Advanced Cranial Navigation", *Neurosurgery*, vol. 72, no. suppl_1, A43–A53, 2013, ISSN: 0148-396X. DOI: `10.1227/NEU.0b013e3182750c03`.

[32] S. Rusinkiewicz and M. Levoy, "Efficient Variants of the ICP Algorithm", *3-D Digital Imaging and Modeling*, pp. 145–152, 2001. DOI: `10.1109/IM.2001.924423`.

[33] *Dimenco*. [Online]. Available: `https://www.dimenco.eu/` (visited on 12/01/2018).

[34] *Meshmixer*. [Online]. Available: `http://www.meshmixer.com/` (visited on 10/18/2016).

[35] S. G. Brouwer de Koning, E. J. M. Baltussen, M. B. Karakullukcu, B. Dasht-bozorg, L. A. Smit, R. Dirven, B. H. W. Hendriks, H. J. C. M. Sterenborg, and T. J. M. Ruers, "Toward complete oral cavity cancer resection using a handheld diffuse reflectance spectroscopy probe", *Journal of Biomedical Optics*, vol. 23, no. 12, p. 1, 2018. DOI: `10.1117/1.jbo.23.12.121611`.

[36] NDI, "Aurora 6DOF Reference , 25 mm Disc , Standard data sheet", vol. 2, no. 610066, 2000.

[37] Matlab, *Reduce patch*. [Online]. Available: `https://nl.mathworks.com/help/matlab/ref/reducepatch.html` (visited on 04/03/2017).

[38] M. file exchange, *Vol3d*. [Online]. Available: `https://nl.mathworks.com/matlabcentral/fileexchange/22940-vol3d-v2`.

[39] Slicer, *3Dslicer*. [Online]. Available: `https://www.slicer.org/` (visited on 04/01/2017).

[40] Matlab, *ICP finite*. [Online]. Available: `https://nl.mathworks.com/matlabcentral/fileexchange/24301-finite-iterative-closest-point` (visited on 10/02/2018).

[41] G. Widmann, R. Stoffner, and R. Bale, "Errors and error management in image-guided craniomaxillofacial surgery", *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology*, vol. 107, no. 5, pp. 701–715, 2009, ISSN: 10792104. DOI: `10.1016/j.tripleo.2009.02.011`.

[42] D. Fize, C. Marlot, and Thorpe, "Speed of processing in the human visual system.", *Nature*, vol. 381, no. 6582, pp. 520–522, 1996, ISSN: 0028-0836. DOI: `10.1038/381520a0`.

[43] NDI, "Aurora V2 User Guide", vol. 3, no. July, 2012.

[44] Matlab, *OpenGL Matlab*. [Online]. Available: `https://nl.mathworks.com/help/matlab/ref/opengl.html` (visited on 09/10/2018).

[45] ——, *Tic Matlab*. [Online]. Available: `https://nl.mathworks.com/help/matlab/ref/tic.html` (visited on 04/01/2017).

[46] B. J. Dixon, M. J. Daly, H. Chan, A. D. Vescan, I. J. Witterick, and J. C. Irish, "Surgeons blinded by enhanced navigation: The effect of augmented reality on attention", *Surgical Endoscopy and Other Interventional Techniques*, vol. 27, no. 2, pp. 454–461, 2013, ISSN: 09302794. DOI: `10.1007/s00464-012-2457-3`.

A Literature Study: Augmented reality for FESS

A.1 Introduction

Endoscopic guidance is typically used when cancer sinus surgery is performed. The procedure is then called Functional Endoscopic Sinus Surgery (FESS).

Endoscopic surgery belongs to the group of minimally invasive surgeries (MIS). With the introduction of MIS, the risk of infection has diminished due to minimal incisions, or no incisions in case of endonasal surgery. This leads to faster recovery times for patients. However, a shortcoming of endoscopic procedures is the loss of direct 3D perspective of the surgical scene [1]. The endoscope's very limited field of view makes it challenging for the surgeon to maneuver surgical tools. Particularly for sinus surgery where critical structures like the optical nerve and the carotid artery are close to the region of surgery, this holds a risk of incidental damage and therefore accuracy is crucial. The recurrence rate of sinus cancer is high, partly because it is never completely certain if the whole tumor is removed during surgery. This is because endoscopic images do not show a clear difference between normal and tumorous tissue. A digital view of the surgical scene can add such information.

Before surgery, a surgeon normally takes a look at a patient's preoperative CT scan. One of the objectives is to identify the critical structures, for example, the optical nerve and carotid artery. With this information, a rough road-map or envision of the surgical scene is made and a plan for the surgical strategy that should be applied. A disadvantage is that this information is only in the mind of the surgeon, so the quality of the surgery is highly dependent on the surgeon remembering the information correctly. Again, a digital view could provide such information by graphically displaying the path to be followed by the endoscope.

The usual FESS screen view is the camera's output of the endoscope and an axial, sagittal and coronal view of the CT scan, augmented with the positions of some of the surgical tools for navigation. An example of a display can be seen in figure 34. An endoscope mostly gives a plain 2D image, as can also be seen in the lower left section of this figure.
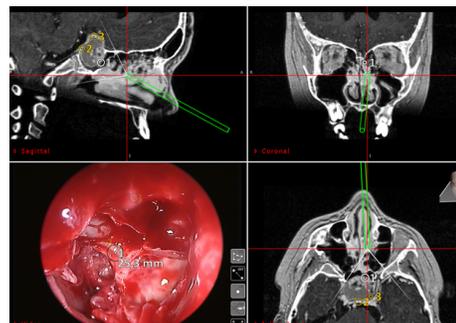
Figure 34: Display during FESS of manufacturer Scopis

Endoscopic sinus surgery would be much easier if at least some of the 3D view could be retrieved and the important structures could be made visible, even though they are obscured by other material. To get the digital information about the position of these critical structures at the surgery scene, a segmentation of the CT data can be made beforehand. The view of the surgeon can be augmented by combining endoscopic images and segmented CT data. A 3D version of this augmented endoscopic video (i.e. augmented reality: AR) can help the surgeon navigate. The surgeon would have a better

sense of the planned path for the tools and the structures to be avoided. This increases the feeling of certainty during the procedure. This would not only make it easier for the surgeon but would also decrease the risk of the procedure for the patient. Additionally, more certainty for the surgeon could shorten the duration of the procedure. The introduction of an augmented view can also lead to more accurate sinus surgeries with surgeons experiencing less task workload [2] [3].

An example of a technique that is often used in sinus surgery research, to regain 3D information from the 2D endoscopic video, is Structure from Motion (SfM). This algorithm utilizes the assumption that pixels in an image belonging to a video stream don't differ much from the pixels in the previous image. Specific features in the image, such as corner-points, are tracked trough time.With this information, an estimation of the position of those points and the camera movement can be made.
In general endoscopic surgery, a technique that is often used is Shape from Shading (SfS). This method is based on the fact that the amount of light reflected by a surface, is determined by the orientation of the surface, the position and sort of the light source and the position of the observer or camera.

The aim of this research is to describe the state of the art of techniques that enhance the view of the surgeon during FESS. In order to get clear what the critical structures in the sinuses are exactly, some attention will be paid to the anatomy. Also, the normal, clinically accepted surgical procedure will be addressed, in order to get more insight into the problems that arise with it. Analysis of augmented reality surgeries, or the use of a 3D view instead of a 2D view in FESS, is interesting to get an idea of the possible gains of such technologies. Finally, the state of the art of sinus surgery data acquisition and AR visualization is described. Also, research is currently conducted into augmenting the surgeon's view for surgery of other body parts. Techniques used in other fields might also be applied in the field of FESS, so they are also described.

A.2   Materials and Methods

A systematic literature search, in the database Scopus, was performed. The selected database was Scopus because it is known for its great technological content. Search terms were:

1 "**Maxillectomy endoscopy**"(124 hits)
2 "**3-Dimensional OR three-dimensional AND reconstruction AND endoscopy**"
  (649 hits)
3 "**Augmented reality endoscopy**"(176 hits)
4 "**3D AND endoscopy AND (sinus OR sinuses)**"(105 hits).

Based on the advice of the surgeon of the AvL also two search terms containing FESS, were included, these were: "**functional endoscopic sinus surgery" AND "augmented reality**" (36 hits) and "**FESS" AND "augmented reality AND NOT flywheel**" (9 hits), the NOT flywheel was added because FESS is also known as Flywheel Energy Storage Systems.

Also, a cross-reference literature search was performed, the citations in the papers regarding augmented reality in a sinus surgery procedure were inspected and another 36 papers were selected.

First, a selection on basis of titles was made. For the first search term, case studies of rare diseases were left out, as well as articles about sinusitis since both subjects are not in the scope of the research. In the area of 3D techniques, a lot of progress is made in the last couple of years. For the second search term, this means that "old" articles should not be included because they don't give a fair view of the most recent developments. In this case, an article is considered old if the publication date is before 2013. Older articles are left out of the research.

The abstract of the papers that were left was read and articles were included if at least one of the following elements was present: **endoscopic sinus surgery**, **three-dimensional visualization in a surgery procedure**, **augmented reality in a surgery procedure**.
Papers are excluded if they concentrate on solving the problem of deformation of organs when the patient undergoes endoscopic surgery in comparison to the pre-operative CT data. Since the structures in the sinus area are rigid, these algorithms do not need to be explored within the scope of this research. Other papers could be excluded because the focus was on surgical results on airflow through the nasal cavities, or on the three-dimensional characteristics of the anatomical structures without a discussion of the technical background. A number of articles could not be accessed due to reading-right issues and is therefore excluded from this report.
In total 49 articles were selected about three-dimensional visualization or augmented reality in a surgery procedure, six additional articles providing background on anatomy

and ten about endoscopic sinus surgery. In figure 35 the numbers of papers found and included after the different iterations in the process can be seen.

Figure 35: Chart of the number of papers included per iteration

A.2.1 Manufacturers and systems

There are already some systems for navigation for FESS on the market.These systems are often used as a basis for further research. A summary of the manufacturers and the features of their systems is given in table 3.

Table 3: features of the system per manufacturer

featues	Optical navi	Magnetic navi	fusion pre-operative and intra-operative	Additional instruments
Brainlab	X	X	X	
Medtronics		X	X	X
Scopis	X	X	X	X
Fiagon	X		X	X
Stryker	X			
Claronav	X			
Healfore	X	X		
Olympus				X
MASMEC Biomed	X		X	X
Sonowand*	X		X	X
NDI	X			X
Parsiss	X		X	X
Anke	X		X	

* sales closed

A.2.2 Separation of the topics

The literature can be divided into literature about **sinus surgery** and **general surgery** of other parts of the body. Most information is about sinus surgery since that is the main subject of the research. The sinus surgery part consist of background information, analysis and state of the art 3D AR techniques for endoscopic sinus surgery.

Background information consist of information about the anatomy of the sinuses and the endoscopic surgical procedure to remove a tumor in this area. The main goal of gathering literature about this topic is to get familiar with the anatomy of the sinuses in order to be able to identify important landmarks and structures and to gain insight into the challenges of a surgical procedure performed in the sinuses.

Some **Analysis** is already done for AR systems that are used during FESS. The objective of looking into this topic is to identify the gains for the surgeon and the patient that could possibly be achieved with and AR view. The performed analyses are a comparison between 2D and 3D endoscopic visualization in clinical studies, the effect of augmented reality on inattentional blindness and task workload for the surgeons in cadaveric studies and the effect of the augmented view on the duration of the procedure in a phantom study.

State of the art of 3D AR endoscopic sinus surgery can be divided into the topics data acquisition and visualization. **Acquisition** consists of methods to gather the necessary information and **visualization** is about displaying this data in an intuitive way. The general surgery part only consists of the topic **state of the art 3D AR techniques for general surgery**, again divided into data **acquisition** and **visualization**.

A.3   Results

In this section, the results of the systematic literature research will be described. Per article, the most relevant information for the research topic will be addressed. The results are categorized as described in methods and materials.

A.3.1   Sinus surgery

The most important part of the research is about sinus surgery, so this will be the most elaborate section of the results. Articles are categorized as described in the methods and materials section.

Background information

Articles that contain information that is necessary to understand the structures of the sinuses and the difficulties of a sinus surgery are addressed in this section.

Anatomy

The paranasal sinuses consist of a number of different cavities, the shape of these cavities can differ from person to person. The most important cavities with their names can be seen in 36. A lot of research is performed to get insight in these personal differences but also into the similarities.
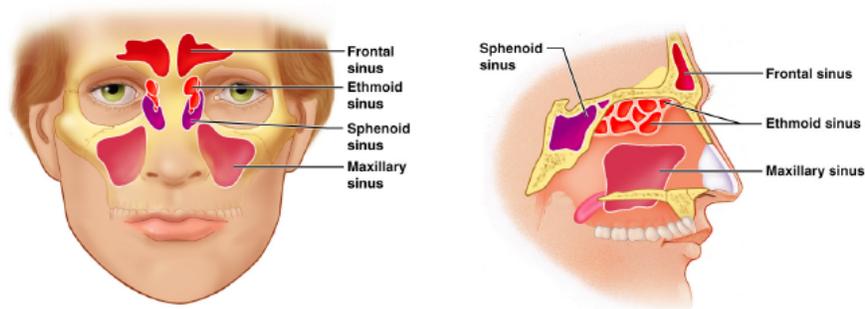
Figure 36: Paranasal sinuses [4]

In [5] the anatomy of the anterior area of the splenoid sinus is described. Observations and measurements of fifteen adult cadaveric skulls and twenty bleached adult dry skulls are taken under a microscope. This resulted in an elaborate description of the sinus area along with the definition of different types of common anatomy. The sphenoid sinus ostium can have a round shape, an oval shape or an ovoid shape. The anatomy of sphenopalatine foramen was in some cases a round shape, an oval shape, a triangular shape, or an irregular shape. Where the oval shape was the most common. Most skulls had one sphenopalatine foramen, and some of them had a secondary foramen, there was even a side with three foramen. The anatomical features of sphenopalatine artery and its main branches are also addressed. The two possibilities were sphenopalatine artery branching before exiting sphenopalatine foramen and branching after exiting sphenopalatine foramen. According to the branching site of

sphenopalatine artery, the authors of the article classified the branching pattern into four different types as can be seen in figure 37a. The posterior nasal septum artery is the end branch of the sphenopalatine artery. The posterior nasal septum artery is by the authors classified into three types. In type I, the stem of posterior nasal septum artery exited the sphenopalatine foramen and branched outside the sphenopalatine foramen. In type II, the posterior nasal septum artery itself entered the sphenopalatine foramen and branched after exiting the sphenopalatine foramen. In type III, the posterior nasal septum artery branched before entering sphenopalatine foramen. Schematics of these types can be seen in 37b.
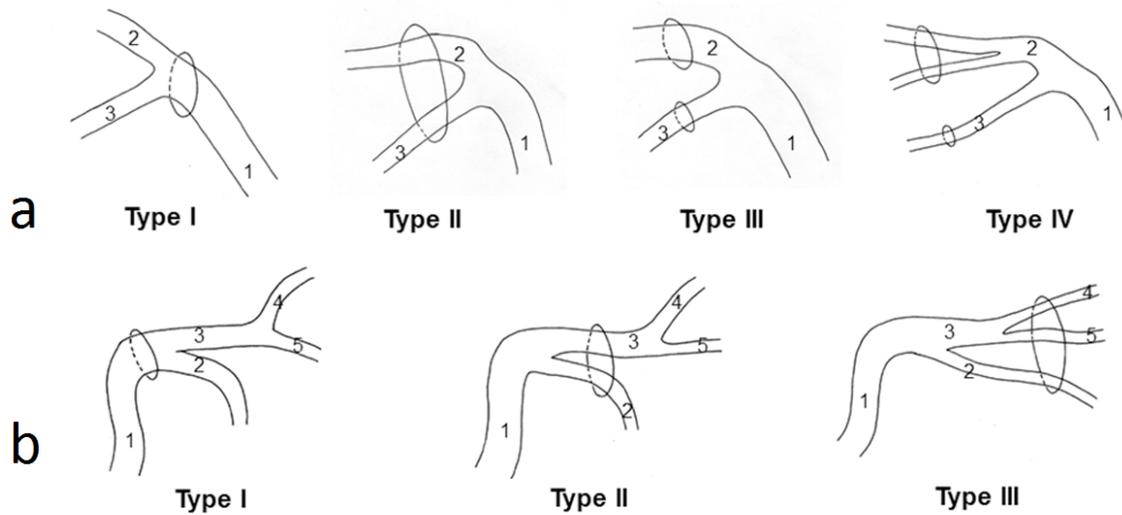
Figure 37: a.sphenopalatine artery b. posterior nasal septum artery [5]

A lot of research into the anatomy of specific cavities in the sinuses is available. For example, in [6] the volume of the maxillary sinus is addressed. Nineteen adult patients underwent a CT scan and with software, the volume of the maxillary sinus was determined. The result was that 64 % of the volume of the maxillary sinus is inferior to the lateral insertion of the inferior turbinate on the lateral nasal wall.

The location of the pterygopalatine fossa and it's relationship to the structures in the sellar region is addressed in [7]. A number of important structures have a relatively stationary position and can thus be located in the sellar region during a surgery, using known landmarks. These structures are the anterior opening of the sphenopalatine foramen, the ptherygoid canal, the palatovaginal canal and the foramina rotundum.

In [8] the anatomical features of the sphenoid sinus are discussed. Various distances in this sinus, as well as the relationship with other important structures such as the optical canal and carotid artery, are addressed. Also, the prescience or absence of an Onodi cell and its influence on the distances is researched.

Statistical shape models of the erectile tissue in the superior, middle and inferior turbinate are presented in [9]. This is done with use of an automatic segmentation technique which made it possible to process large data sets.

Besides the anatomy of the cavities, the different kind of tumors that can be found there are of importance for surgery as well. In [10] an overview is given of the different kinds of cancer that are located in the area of the sinuses, these are osteoma, inverted papilloma and Juvenile angiofibroma of the nasopharynx.

Surgical procedure
In order to get familiar with the surgical procedure for resection of a tumor from one or more of the sinuses, some papers and books are addressed here.

In [10] a comprehensive overview is given for a complete endoscopic sinus surgery. All steps including patient preparation, pre-operative scans, preparation of the surgical site, the surgical procedure itself, complications, results and post-operative care are described. A classification of all possible endoscopic sinus surgeries is given and elaborately explained. The classes are uncinectomy (infundibulotomy), maxillary sinus fenestration/maxillary sinus surgery, anterior ethmoidectomy, posterior ethmoidectomy, sphenoid sinus fenestration/sphenoid sinus surgery, frontal sinus drainage/frontalsinus surgery type I-III, and pansinus surgery. The technical equipment chapter is partly about virtual and 3D endoscopes, navigation and robotics.

In order to access the different cavities of the endonasal sinuses, certain paths are made with the endoscope. In [11] a summary is given of the different sinuses and the paths that lead to them.
Not all paths can be followed without modifying the anatomy of the patient; sometimes some bone has to be drilled away. For example, a large percentage of the maxillary sinus lays below the lateral insertion of the inferior turbinate on the lateral nasal wall, the nasolacrimal canal prevents endoscopic visualization and access to the anterior maxillary sinus wall. So resection of the lateral nasal wall to the inferior meatus and anteriorly to include the nasolacrimal canal is required for proper access to the maxillary sinus[6].
To get a better idea of the anatomical changes a sinus surgery can make, endoscopic sinus surgical procedures are correlated to CT scans, by means of color coded maps on CT images of twelve common sinus surgeries in [12]. Moreover, some very intuitive 3D-PDFs are described in [13]. In this paper, an explanation is given about obtaining these models from CT data. The models themselves give a good insight in which parts are removed in which order during the extended endoscopic endonasal transsphenoidal approaches of a sinus surgery.

A large number of sinus surgeries are revision surgeries, navigating on anatomical landmarks is difficult because they might be changed in an earlier surgery. These revision surgery procedures are the topic of [14]. This book indicates the technologies for revision surgery as well as the, most often changed, anatomy in these cases. Also, the surgical instruments that are used and choices of anesthetics are included.

All classes of sinus surgery, as mentioned in [10], are addressed in case of a revision surgery.

In [15] also different endoscopic strategies are described with respect to recurrent sinonasal inverted papilloma. This paper states that the aggressive endoscopic resection of the tumor is a method of surgery that can be performed as an acceptable alternative to the open approach.

In [16] a description is given of sinus surgeries and outcomes for six different patients. This paper gives an oversight of the procedures, how they are usually performed and what outcomes can be expected. Five patients had an inverted papilloma confined to the maxillary sinus and one had a solitary primary extramedullary plasmocytoma (EMP) in the inferior meatus. In one case, it was a primary surgery, in five other cases it was a revision surgery.

Nowadays, tracking inside a patient's head to enhance the surgeon's view is applied in many hospitals. This is most commonly done with the use of an optical (e.g. infrared) system or with an electromagnetic system. In [17] a comparison is made between the two and positive and negative aspects of both are described. The writers conclude that an image guided navigation system is very useful in enhancing the view of the surgeon in complex surgeries. A positive aspect is that complications and damage to surrounding tissue can be avoided. The negative aspect is the price, not many hospitals may be able to afford such a system, but it might be mandatory in the future.

Analysis

The introduction of navigation in endoscopic sinus surgery brought improvements for the confidence of the surgeon, the idea of a 3D orientation as well as an increased precision of the surgical approach [18]. With the next step, augmenting the view of the surgeon even more, hopes are that more improvements can be made. This section describes the analyses of the improvements that are realized so far.

Comparison 2D and 3D

A review of three different endoscopes is described in [19]. Endoscopes of the following types are compared: 2D High Definition(HD), 3D Standard Definition(SD) and 3D HD. The stereoscopic depth perception was significantly the best in the 3D HD endoscope followed by the 3D standard definition one. The brightness, however was the best in the 2D endoscope. This is due to the fact that the 3D view made use of polarized light to get a 3D effect, which more than halves the brightness of the image. The 3D HD endoscope scored the best for comfort during surgery. For the surgical procedures that were tested it reduced the operation time slightly. The biggest problem of the 3D endoscope is the prize, a system brings currently about €100000 of initial costs.

Effect of augmented reality
When an augmented reality view is introduced at the surgery scene the question arises whether this is beneficial for the patient, the surgeon or the hospital. The recurrence rate of tumors in the sinus is high [14]. With better accuracy, the certainty that the whole tumor is removed could be increased, causing the recurrence rate to decrease as well. FESS is a very intensive procedure for the surgeon and can also take a rather long time. So research is also focused on improvements for the task workload of the surgeon and the duration of the procedure, to decrease the overall costs. Note that all the experiments that are described in this section are pre-clinical, on cadavers or on a phantom.

Lists of the negative and positive aspects of the AR view in FESS are below. In the rest of this section, the different points will be explained and supported by literature.

negative aspects:
- inattentional blindness increases
- too busy surface mesh which makes it distracting
- inexperienced surgeon benefit more that experienced surgeons

positive aspects:
- median distance to target point gets smaller
- task workload for the surgeon decreases
- confidence of the surgeon increases
- duration of the surgery decreases

In [20] otolaryngology trainees, fellows and staff surgeons performed an endoscopic navigation exercise on a single cadaveric specimen. The subjects were randomized, stratified to balance the level of experience, to either a standard endoscopic view (control group) or an AR view consisting of an endoscopic video fused with anatomic contours. The fusion was 75% endoscopic view and 25% contours overlay, in order to equal the sight of both groups, the control group also got an endoscopic view with opacity of only 75%. The outcome measure was the detection rate of one critical complication, which was an injured nerve, and a foreign body, in this case a screw. The outcome of this research showed a significant difference in the detection rates. This supports the hypothesis of inattentional blindness caused by the AR view. Inattentional blindness means that all the extra info causes the surgeon to be "blind" for alarming situations in clear sight. The AR view consists of too much additional information, demanding a lot of attention, this is distracting the surgeon. But another outcome is that the median distance to the target point is significantly smaller in the AR group. The authors note that the extreme events which were tested here, are rare and focus is also a good thing so this article is not an essay against AR, but the perceptual limitations of the task-focused end-user have to be considered. Therefore, in the opinion of the authors, it is essential that physicians play a role in the development of AR systems

A comparison between the task workload for an endoscopic sinus surgery with AR and one without is performed in [2]. Otolaryngology fellows and residents performed an endoscopic exercise of localizing landmarks on a cadaveric head, one time only with the endoscope and one time with three-planar guidance and a virtual view with visible anatomical contours. A NASA Task Workload Index questionnaire was filled out. This questionnaire asks the surgeon about the amount and gravity of the mental demand, physical demand, temporal de-

Figure 38: NASA task load index scores for each domain. $*P < .05$ calculated by Wilcoxon signed rank test [2]

mand, performance, effort and frustration [21]. Results of this questionnaire can be seen in figure 38. Also, an extra questionnaire regarding confidence was filled out. The AR exercise got better scores on both the NASA and the extra questionnaire. Also, the mean deviation from the target points was smaller in the AR exercise. Remarks are that all participants are likely to be rather inexperienced surgeons and they had to perform both the exercises, which is very likely to make it easier the second time whether AR is used or not.

A comparable experiment is described in [22]. In this setup also otolaryngology fellows and residents performed ESS on cadaveric heads. The difference is that they had to perform the surgery for one-half of the head with a conventional method and for the other half with and additional augmented reality view. The participants were randomized in the order in which they had to perform these two parts of the exercise. The task workload was again evaluated with the NASA task workload questionnaire.

Also, an extra questionnaire and an open interview were used to evaluate the findings of the surgeons. The augmented reality system, ART-IGS, had a 3D virtual view that was dynamically updated by the tip position of the instruments and also provided feedback by means of auditory proximity alerts. An example of the screen can be seen in figure 39. The NASA questionnaire revealed that the additional augmented reality view decreased the task workload. Feedback on the method, however included the opinion that the surface mesh was sometimes too busy with unnecessary information what made it confusing for the surgeons.
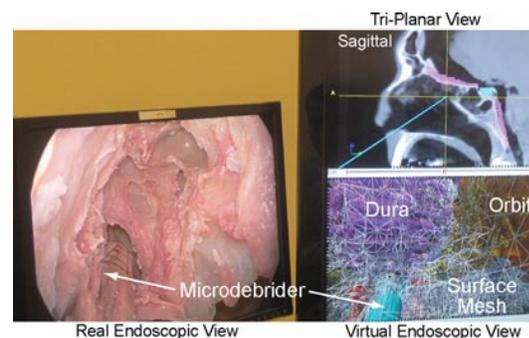
Figure 39: Parallel virtual view allowed fast referencing of the position and trajectory of the tracked ablative instrument. Virtual endoscopy enabled visualization of critical structures lying behind the visible surface [22]

72

The participants also mentioned that au-
ditory alerts were probably the best since haptic and visual alerts could be too dis-
tracting. For this method the critical structures had to be contoured manually, this
takes a lot of time but can also be seen as a valuable training exercise for students.

Another task workload experiment is described in [3], this time also the duration
of the AR procedure is compared with that of a normal one. In this paper, a self-
developed augmented reality method is evaluated. In this method, a rendered view of
the pre-operative CT or MRI is used as background and the endoscopic image is semi-
transparently fused onto it. Participants had to indicate specific anatomical markers
on the phantom of a head. Both the duration of the procedure and the task workload
decreased. An important note is that inexperienced surgeons had more benefit from
this system than experienced surgeons.

**State of the art of 3D Augmented reality techniques in endoscopic sinus
surgery**

Calibration and registration are the start of the augmenting process. Since the endo-
scopic images are often used as a basis for the view, this information needs to be as
accurate as possible. The camera needs to be calibrated to guarantee a minimal error
in the video images and the video data needs to be registered to the other available
data in order to match the coordinate systems.

Pre-operative and interoperative information is necessary to enable a 3D augmented
reality view in sinus surgery. There are different choices of the sort of information
that is used and different methods of acquiring this vital information. Almost all
methods make use of pre-operative CT data, but a reconstruction of the region of
interest using MRI is also possible. The 3-dimensional information can be derived
from the 2D endoscopic images. This can be done "sparse" which means that not all
of the pixels are translated to a 3D model but only a part of them. A model based
on information of the images before the last one is used to fill in the rest of the 3D
image. It can also be done "dense" in which case the whole image is translated to 3D.
This costs a lot of computation power and is, therefore, hard to perform real-time.
Nowadays there are also 3D endoscopes on the market which are small enough to fit in
the nose, a comparison of the quality of the images is for example made in [19]. These
endoscopes are however very expensive and not available at all manufacturers, so the
opportunity to make use of them is limited by the hospital's budget and supplier. In
many papers, a combination of the pre-operative CT data and an algorithm to get 3D
images out of the 2D endoscopic video is used to create a 3D virtual reality view for
endoscopic surgery. Some further details of the methods will be given in the section
"acquisition".

After acquiring the information, a method of visualizing it in an intuitive way needs to
be conducted. Conventionally the visualization of augmented reality in sinus surgery
is done with a three-display view of the 2D sight along the three dimensions (coronal,

axial and sagittal) from the 3D CT data, augmented with the pose of the instruments. The visualization of the endoscopic images is mostly 2D. To enhance the view sometimes the 2D acquired images are translated into 3D images. The augmented view consist mostly of the endoscopic video overlaid with some extra information. This information can, for example, be CT data of landmarks that are pre-operatively chosen by the surgeon. An elaborate summary of the methods will be given in the section "visualization".

Calibration and registration

The endoscopic camera always plays a role, since it is the purest and real-time observation of the scene. It is very important that the camera is proper calibrated when the video stream of the camera is used in combination with other data. Registration of one sort of data to another sort of framework also plays a large role in the final accuracy.
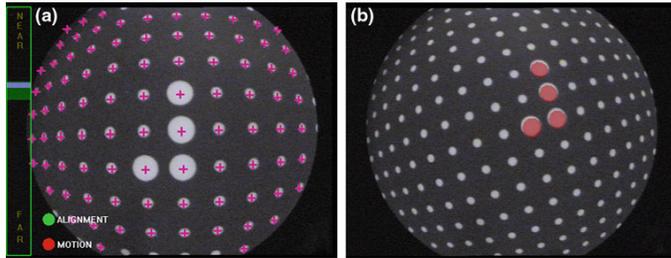
Figure 40: a) Endoscope calibration process b) overlay of the bigger dot markers of the calibration pattern for test purposes [23]

Camera calibration is usually performed with a checkerboard, but in [23] another option is presented. With the use of a dot pattern, the endoscopic camera can be very accurately calibrated. The pattern consists of a raster with a lot of small dots and four big dots. After calibration, the surgeon can check the accuracy and start over if it is not satisfying. In figure 40 an example of a calibration procedure with the use of this system can be seen.

Initial registration of tracking algorithms (initial pose estimation of the camera that matches the reality best, in a known coordinate system) is not always very accurate. In [24] a novel registration method is presented, which tries to optimize the similarity metric between the endoscopic image and a predicted image that is constructed via rendering of CT images. This method takes physical constraints, such as collisions between endoscope and tissue, into account. The registration time is on average 4.4 seconds, which is too slow for a real-time application, but it can serve as an initial registration for another method.

Another method for accurate registration of endoscopic images to CT data is presented in [25]. First, the feature points of the image are extracted using Scale Invariant Feature Transform(SIFT), the motion between the images is estimated, and then the feature points are reconstructed. At the start of the algorithm, these points are registered to a segmented CT image and with this registration, a tracking method
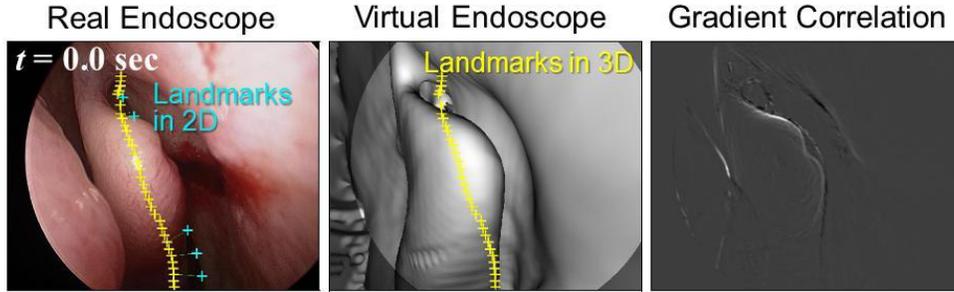
Figure 41: Result of the registration algorithm on image sequence from a patient study [24]

is initialized. The results are very promising although it is not tested against an objective truth. The time to process a frame-pair is 2 minutes, so it is not useful for real-time tracking but could be used to enhance the accuracy of other tracking methods.

Acquisition

The next step is to get the 3D information from a 2D endoscopic image, this process is also called "volume from view". Possibly pre-operative data like CT or MRI images can help to improve this information. A technique that is used often is Structure from Motion, SfM. This is an example of a method that can both be applied sparse and dense. This method is used in [26], [27] and [28]. The algorithm makes use of the assumption that pixels in an image that belong to a video stream don't differ much from the pixels in the previous image. The sparse implementation works with specific features of the image, such as corner points, these are tracked through time. Using this information, an estimation of the position of those points and of the camera movement can be made.

In [26] the structure from motion algorithm is used to generate a 3D point cloud from a sequence of images. This point cloud is then registered to a 3D mesh of a part of the patient's sinuses, using the Iterative Closest Point (ICP) algorithm. This 3D mesh is created using pre-operative CT data. With a reasonable initial registration, the results of the average registration error are 1.21 mm when the endoscope is pointed at erectile tissue and 0.91 mm for non-erectile tissue. The computation time is less than 7 seconds for 0.5 seconds of video images, about 15 frames, which is not real-time

In [27] a fuzzy zoning technique is introduced to reduce the computation time of the structure from motion algorithm. The new technique was about 2.6 times faster than the normally used, SIFT, technique, but it still can't perform real-time.

In [29] a combination is used of pre-operative CT-data and the structure from motion algorithm applied to inter-operative endoscopic images in order to estimate a

depth-map of the image. A regression function is used that is trained per patient and therefore assures a higher accuracy than a general function.

In [30] another method is presented: photo-consistency. This method works on the assumption that the illumination of an object is independent of the viewing direction. The intensity difference of one voxel which appears in two or more endoscopic video images is measured to calculate a cost function to align the images.

A very different technique for augmented 3D navigation is discussed in [31]. The endoscope is equipped with a laser, the beam crosses the endoscopic field of view and hits a structure, with the use of known and calibrated geometry this spot can be extracted real-time. The spatial distance of the spot in the patient relative to the distal tip of the 0° endoscope along the optical axis can be calculated.

Visualization
Apart from the depth information acquisition of the image, visualization of the 3D information is important to get the information to the viewer, the surgeon. This may also be done in several ways.

In [23] a technique is presented that contains visualization of the surgical equipment and silhouettes of the target structures, along with a distance indication from the pointer to the structure. Surgeons who used the system tell that it is close to the normal workflow and can make an improvement in safety and efficiency of this kind of surgery.

In [32] the CT data is used to create a 3D model of the head and overlay this on the endoscopic video during surgery. This superimposed model can be switched on and of using a foot-switch during surgery.

A very simplistic method of visualizing augmented reality is given in [33]. The surgeon has to choose specific landmarks in either MR or CT volumetric data. These landmarks are overlaid on the video stream. To ease the registration of the landmarks to the patients head a Dynamic Reference Base is attached to it. To register the pre-operatively chosen landmarks to the video stream a pair-point algorithm and surface matching are used.

The authors of [34] use a rather comparable method, also some landmarks are selected by the surgeon in the pre-operative CT data. These are overlaid on the endoscopic view during surgery. Some extra features of this system are the colors that can be chosen, the distances to the landmarks that are displayed and the alert system, which makes the landmarks turn red in case of proximity.

Sometimes a virtual reality view is only constructed for learning purposes, then it does not need to update interactively with the endoscope. For example in [35] two different software packages are used to create 3D reconstructions of CT and MRI data in order to be able to perform virtual endoscopy.

A.3.2 General endoscopic surgery

In general endoscopic surgery, augmented reality is also used to enhance the view of the surgeon. The methods that are applied to other body parts can possibly also be used for the sinuses. This section describes the state of the art of the AR methods used in endoscopic surgeries apart from sinus surgeries.

State of the art of 3D Augmented reality techniques in general endoscopic surgery

The acquisition of the information necessary for augmented reality are mostly 2D endoscopic images, sometimes combined with pre-operative CT or MRI data. The visualization of the surgical scene is usually showed as 2D endoscopic video. The images are translated to a 3D view sometimes. This is all very similar to the acquisition and visualization for sinus surgeries. An important difference is that the structures of the sinus cavities are rigid, while size and volume of other body parts can differ trough time. In [36] a very broad outline of the steps that are needed to create an AR view is given. This paper states that the techniques nowadays are advanced enough to be applied in medicine but that improvement is still necessary to let the surgeons have even more advantage of the AR displaying methods. The outline of this section will be exactly the same as for the last section about sinus surgery since the process of enhancing the view of the surgeon is more or less the same.

Calibration and registration
Camera calibration is normally easily performed with the use of a checkerboard pattern. But due to vibration and refocusing of an endoscopic camera, re-calibration during surgery might be necessary. In [37] this challenge is solved by a continuous framework that checks the performance of the calibration. When the update criteria are met a particle filter is activated to re-calibrate. The epipolar constraints are used as the weighting function.

In [38] a mapping algorithm to register 2D images to 3D volumes is given. This is done for 2D information presented in slices, such as CT or MRI images, so it can be called "slice to volume registration". This method is capable of capturing in-plane transformations, such as a heartbeat for images of a heart. It makes use of Markov random fields (an explanation of this mathematical method can be found in [39]), based on intensity and independent of the metric. A comparable precision to that in other state of the art methods is reached.

Acquisition
In [40] a survey is performed of all existing techniques to go from a 2D image to a

3D one. There are twelve different depth clues identified, and the main conclusion is that combining them gives the most accurate results. Also, the use of multiple images for the acquisition of depth cues gives more accurate results than the use of only one image.

In [41] a comparative validation study of different techniques of 3D reconstruction for laparoscopic surgery is done. In this paper, the focus is on "single-shot" techniques, so there is no movement of the endoscope needed. Validated methods are stereoscopy, Structured Light and Time of Flight. The conclusion is that stereoscopy currently is the only feasible real-time solution to obtain 3D information during surgery. All methods had problems concerning robustness in case of contamination of the lens by smoke or bleeding.

In methods for different body parts than the sinuses, a technique that is often used is shape from shading(SfS). This method is based on the fact that the amount of light reflected by a surface is determined by the orientation of the surface, the position and sort of the light source and the position of the observer or camera. The technique is most often applied dense, to the whole image.

The shape from shading algorithm is used in [42]. In order to obtain the absolute grey gradient field, which is necessary for the SfS algorithm, a technique is used that is called "optical flow". A grey card is introduced, this can be used to calibrate the relationship between the light source intensity and the camera response function. With the use of this relationship the inverse of the image intensity can be obtained to balance the intensity of the original image. This compensated image can then be used to obtain the absolute grey gradient. This, however, is done with only one image, it is not yet applicable to a video stream in real-time.

Another paper about the shape from shading algorithm is [43]. In this article, a very fast version of the algorithm is presented, which makes use of a new image irradiance equation, based on better assumptions than usual. A perspective projection from the camera is assumed, as well as a light source which is close to the object instead of a single point at infinity. The results indicate that this method is faster and more accurate than existing techniques.

The shape from shading method is used to make a comparison to the method in [44]. This paper presents a method for 3D image synthesis from 2D laparoscopic images. The depth information from CT data is incorporated in the 2D laparoscopic image. A number of algorithms for feature detection and matching and also camera tracking are used. This depth map generation is compared to the that of the shape from shading method. A stereo image, thus creating a 3D effect is created with use of the depth map and the image of the video.

Different algorithms are combined to arrange a 3D panorama visualization of the scene with an endoscopic video in [45]. SfS is used to obtain a depth map, Speed Up Robust Features (SURF) is used for feature detection, Binary Robust Independent Elementary Features (BRIEF) to find matching features and the ICP algorithm to stitch the consecutive frames together. The ICP algorithm is also used in [46], to align different frames. In this paper, Simultaneous Localization and Mapping (SLAM) is used with extra information from the endoscopic camera in order to reduce the difficulty of the estimation.

Sometimes, most often in surgeries of body parts that can undergo drastic anatomic changes such as the abdomen, an inter-operative CT image is made, using a CT cone beam arm. The transformation of the organs between the original data and the new data is obtained. But roll, translation along the image axis of the endoscope, zoom and focus can still cause errors in the registration of the CT image to the endoscopic images. The writers of [47] aim to solve these problems by introducing a local formulation of shading constraints, they claim that a dense shape from shading algorithm cannot be used because there is no constant albedo throughout the scene. The method makes uses of piecewise constant albedo and light intensity on different patches that are separated from the original images using a watershed algorithm. In this paper, the position of the endoscope is estimated rather that the shape of the organs, which makes the writers introduce the term "pose-from-shading".

In [48] a method for 3D non-rigid pose estimation for kidney surgery is presented. This method can segment multiple important structures and estimate their pose in the 3D space, intra-operatively. It can correct for non-rigid deformations and also for changing camera parameters, such as zoom or focus. The pose estimation does not need any correspondence between points from the pre-operative and intra-operative data, the user can simply click on structures that are of his interest. This method works with a local optimization framework what makes is rather dependent on the initialization. Since there is a large variability in patients, the first frames from the video data were used to train random forest models for that specific patient. This is not so applicable in real-time since the training of the models consumes a lot of time, but the method is very parallelizable which could make it faster.

Visualization

The way of visualizing the augmented reality view is a big influence to the usefulness of the system for the surgeon. In [49] four different techniques are described and compared for advantages and disadvantages. The techniques are transparent overlay, virtual window, random dot mask and the ghosting method. A random dot mask is a pattern of small dots through which the underlying structure can be seen. The ghosting method calculated the importance of the camera image and then determines the transparency of the pixel according to that. The table below gives the advantages and disadvantages of the different systems.

Table 4: Comparison of different AR visualization modes in [49]

Techniques	Advantage	Disadvantage
Transparent overlay	Simple to operate; Applying to both simple and complex virtual structures	Causing virtual structure floating over the real surface
Virtual window	Maintaining correct spatial relationship between real and virtual structures	Requiring user to define a region of interest; Missing the shape and color of some part of the real surface; Applying only to simple and small virtual structures
Random-dot mask	Maintaining correct spatial relationship; Solving the problem of real surface removal	Causing the feeling of clutter; Missing some part of virtual structures; Applying only to simple and small virtual structures
The ghosting method	Applying to both simple and complex virtual structures; Maintaining the sense of the correct spatial relationship	Fail when the real image does not have sufficient features; Requiring some image processing techniques
The depth-aware ghosting method	Providing clearer depth information of complex virtual structures	Revealing usefulness only in complex virtual structures

A virtual reality view and an augmented reality view are synchronized in [50] in order to obtain depth info of objects within the view of the endoscope. The VR and AR views are overlaid on the endoscopic image. The endoscope and other surgical instruments are also displayed in the virtual reality view.

A.4 Discussion

The aim of the paper was to describe the current status of the state of the art systems augmenting the view of the surgeon during FESS. The algorithms and methods could be designed for sinus or other types of surgery. Solutions have not yet led to clinical applied inter-operative used systems. This is mainly due to the time complexity of the algorithms, which is too high for real-time computation. Also, most algorithms are not at the stage yet to be tested clinically, due to the newness of the techniques.

The research is very pre-clinical and because of the novelty of the subject, no standardized methods of evaluation are available. This makes it difficult to compare the quality of different techniques.

Most common methods for gathering depth information from a single image are Structure from Motion and Shape from Shading. Both algorithms, as well as other methods, face robustness challenges in case of lens occlusion by smoke or bleeding. Other challenges are for example; removal of structures during surgery or in case of organ surgery, the non-rigid transformations of tissue. A difficulty of the SfS method is the non-uniqueness of the solution when the Lambertian reflection of the surface is not known, which is generally the case for sinus surgery. The main problem with the SfM algorithm is the real-time constraint, but this can possibly be solved by using pre-operative information like CT or MRI as an extra source of depth information.

As mentioned before, a simple way of acquiring 3D info of the scene is to make use of a 3D endoscope. However, there aren't many manufacturers selling this equipment small enough to be applicable in the sinus area and these endoscopes are very expensive. Augmentation can consist of a lot more info than just a 3D view, so solely using a 3D endoscope does not cover the need of extra information.

Displaying the acquired 3D augmented info is done in several ways. For example, a (semi-)transparent overlay of the structures on the endoscopic video, or an overlay of just some specified landmarks. Also, a complete Virtual Reality (VR) view is developed for virtual surgery. Some other described techniques are virtual window, random dot mask and the ghosting method.
A challenge is to balance the amount of info with the visibility of the original endoscopic view. Too much information in one image can lead to distraction of the surgeon or occlusion of important structures.

Something lacking in most cases is 3D visual feedback for the surgeon in the Operating Room (OR). Furthermore, the endoscopic video seems to be the base of most visualization techniques, even when a VR view is available. This VR view is most often used for training young surgeons-to-be but is not used in the OR itself.

Despite the challenges and problems mentioned, an augmented view for endoscopic sinus surgery can positively change the field of work. In the few analytical types of research conducted, the task workload for the surgeon decreased, the duration of the procedure decreased and the accuracy improved. However, experienced surgeons are still required and one must be aware of problems such as inattentional blindness. The research about Augmented Reality for FESS is not conducted in clinical environments and mostly performed with young, inexperienced surgeons. For this group the workload clearly decreases with use of the systems, but it remains indecisive whether this would decrease for all groups of surgeons.

A.5 Conclusion

Techniques to enhance the view of the surgeon during FESS can be divided into several categories. The goal of enhancing the view is to regain the intuitive 3D view like in open sinus surgery and to provide extra information. Therefore, 3D data has to be acquired and displayed.

The easiest but also the most cost-intensive way is for the hospital to purchase a 3D endoscope.

Another option is to acquire images from a 2D endoscope and to transform these into 3D information. For sinus surgery, techniques that are described to achieve this are, Structure from Motion (SfM) [26], [27] [28], combining CT data with SfM [29] , photo-consistency [30] and an endoscope equipped with a laser beam [31]. Of all techniques used in other surgical procedures that can possibly be applied to sinus surgery, Shape from Shading (SfS) is most often used [42] [43]. This method can also be applied using a local formulation of the shading constraint, to estimate the pose of the endoscope. This is called pose from shading [47].

Other techniques are depth map generation with use of pre-operative CT data [44], structured light and time of flight [41] or a local framework with per patient trained random forest models [48]. Also, SfM can be extended with feature detection and matching using Speeded Up Robust Features (SURF) and Binary Robust Independent Elementary Features (BRIEF) and combining images with the Iterative Point algorithm [45] [46].

With only the acquisition of 3D info, the surgeon does not have an enhanced view yet, the way of visualizing the data is also important. This is also described in several ways, for example silhouettes [23], a 3D overlay of CT data on the endoscopic video [32], showing landmarks on the endoscopic video [33] [34] or a VR view of a sinus surgery [35]. For visualization, techniques from other surgical fields can also be used for sinus surgery. Methods discussed in this research are transparent overlay, virtual window, random dot mask and the ghosting method [49] and the combination of an AR and a VR view [50].

In conclusion, many techniques exist to enhance the view of the surgeon, during sinus surgery or other sorts of surgery. Improvement can still be made since the techniques are not yet clinically applied and inter-operatively used. They are not able to perform real-time and lack 3D visual feedback during surgery. All methods had problems concerning robustness in case of endoscopic lens contamination.

A.6 A novel system for navigation in the sinuses

This literature research shows that already some techniques exist to acquire real-time 3D data during a sinus treatment. Also quite a few navigation systems for tracking tools inside the sinuses are on the market. Most techniques make use of an algorithm to extract the position and orientation of the camera from the endoscopic video. The depth is also estimated using the video stream, often enhanced with information from pre operative CT or MRI data. An addition to the current techniques can be made

by enabling direct tracking of the endoscopic camera, and using that information as the basis of navigation.

Visualization of the 3D model is often done by placing an overlay on the endoscopic video, or in some cases a non real-time virtual reality model is created for learning purposes. A new way of displaying all information is a real-time virtual view of the surgical scene, seen from the same position and orientation of the real endoscope. Disadvantages of endoscopic video, such as occlusion by smoke or bleeding, are thereby avoided and extra information regarding critical structures can be added. Displaying the virtual view in 3D is a real novelty in this area and would make the view even more intuitive.

Concluding, a novel system for enhancing the view of the surgeon during a sinus treatment will consist of real-time tracking of the endoscope and a virtual view from the viewpoint of the real endoscope containing visual information about the tissue. This view would ideally be displayed in 3D on a 3D screen.

A.7 References

References

[1] M. S. Nosrati, A. Amir-Khalili, J. M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, R. Abugharbieh, and G. Hamarneh, "Endoscopic scene labelling and augmentation using intraoperative pulsatile motion and colour appearance cues with preoperative anatomical priors", *International Journal of Computer Assisted Radiology and Surgery*, vol. 11, no. 8, pp. 1409–1418, 2016, ISSN: 18616429. DOI: `10.1007/s11548-015-1331-x`.

[2] B. J. Dixon, M. J. Daly, H. Chan, A. Vescan, I. J. Witterick, and J. C. Irish, "Augmented image guidance improves skull base navigation and reduces task workload in trainees: A preclinical trial", *Laryngoscope*, vol. 121, no. 10, pp. 2060–2064, 2011, ISSN: 0023852X. DOI: `10.1002/lary.22153`.

[3] L. Li, J. Yang, Y. Chu, W. Wu, J. Xue, P. Liang, and L. Chen, "A novel augmented reality navigation system for endoscopic sinus and skull base surgery: A feasibility study", *PLoS ONE*, vol. 11, no. 1, pp. 1–17, 2016, ISSN: 19326203. DOI: `10.1371/journal.pone.0146996`.

[4] S. Vertebral, "The Axial Skeleton Eighty bones segregated into three regions Skull Vertebral column Bony thorax Bones of the Axial Skeleton", 2006.

[5] S. Wang, J. Zhang, L. Xue, L. Wei, Z. Xi, and R. Wang, "Anatomy and CT reconstruction of the anterior area of sphenoid sinus", *International Journal of Clinical and Experimental Medicine*, vol. 8, no. 4, pp. 5217–5226, 2015, ISSN: 19405901.

[6] N. Tanna, J. D. Edwards, H. Aghdam, and N. Sadeghi, "Transnasal endoscopic medial maxillectomy as the initial oncologic approach to sinonasal neoplasms: the anatomic basis.", *Archives of otolaryngology–head & neck surgery*, vol. 133, no. 11, pp. 1139–1142, 2007, ISSN: 1041892X. DOI: `10.1016/S1041-892X(08)79122-4`.

[7] Y. Cheng, H. Xu, Y. Chen, S. Xu, H. Yu, S. Zhang, Y. Feng, and G. Zhao, "Location of Pterygopalatine Fossa and its Relationships to the Structures in Sellar Region.", *The Journal of craniofacial surgery*, vol. 26, no. 6, pp. 1979–1982, 2015, ISSN: 1536-3732 (Electronic). DOI: `10.1097/SCS.0000000000001899`.

[8] S. H. Hwang, Y. H. Joo, J. H. Seo, J. H. Cho, and J. M. Kang, "Analysis of sphenoid sinus in the operative plane of endoscopic transsphenoidal surgery using computed tomography", *European Archives of Oto-Rhino-Laryngology*, vol. 271, no. 8, pp. 2219–2225, 2014, ISSN: 14344726. DOI: `10.1007/s00405-013-2838-9`.

[9] A. Sinha, S. Leonard, A. Reiter, M. Ishii, R. H. Taylor, and G. D. Hager, "Automatic segmentation and statistical shape modeling of the paranasal sinuses to estimate natural variations", vol. 9784, pp. 1–8, 2016, ISSN: 16057422. DOI: `10.1117/12.2217337`.

[10] R. K. Weber and W. Hosemann, "Comprehensive review on endonasal endo-scopic sinus surgery.", *GMS current topics in otorhinolaryngology, head and neck surgery*, vol. 14, Doc08, 2015, ISSN: 1865-1011. DOI: `10.3205/cto000123`.

[11] a. I. Moral, M. E. Kunkel, K. Tingelhoff, M. Rilk, I. Wagner, K. W. G. Eich-horn, F. Bootz, and F. M. Wahl, "3D endoscopic approach for endonasal si-nus surgery", *Annual International Conference of the IEEE Engineering in Medicine and Biology - Proceedings*, pp. 4683–4686, 2007, ISSN: 05891019. DOI: `10.1109/IEMBS.2007.4353385`.

[12] J. J. Ludwick, K. H. Taber, S. Manolidis, A. Sarna, and L. A. Hayman, "A computed tomographic guide to endoscopic sinus surgery: axial and coronal views.", *Journal of computer assisted tomography*, vol. 26, no. 2, pp. 317–22, 2002, ISSN: 0363-8715. DOI: `10.1097/00004728-200203000-00026`.

[13] M. Mavar-Haramija, A. Prats-Galino, J. a. J. Méndez, A. Puigdelívoll-Sánchez, and M. de Notaris, "Interactive 3D-PDF Presentations for the Simulation and Quantification of Extended Endoscopic Endonasal Surgical Approaches", *Jour-nal of Medical Systems*, vol. 39, no. 10, 2015, ISSN: 1573689X. DOI: `10.1007/s10916-015-0282-7`.

[14] S. E. Kountakis, J. B. Jacobs, and J. Gosepath, Eds., *Revision_sinus_surgery[1]*. Springer, 2008, ISBN: 9783540789307. [Online]. Available: `http://download.springer.com/static/pdf/434/bok%5C%3A978-3-540-78931-4.pdf?originUrl=http://link.springer.com/book/10.1007/978-3-540-78931-4%5C&token2=exp=1483370214~acl=/static/pdf/434/bok%5C%253A978-3-540-78931-4.pdf?originUrl=http%5C%`.

[15] F. Lian and H. Juan, "Different endoscopic strategies in the management of recurrent sinonasal inverted papilloma.", *The Journal of craniofacial surgery*, vol. 23, no. 1, e44–8, 2012, ISSN: 1536-3732. DOI: `10.1097/SCS.0b013e318241dae7`.

[16] P. Eloy, N. Mardyla, B. Bertrand, and P. Rombaux, "Endoscopic endonasal medial maxillectomy: Case series", *Indian Journal of Otolaryngology and Head and Neck Surgery*, vol. 62, no. 3, pp. 252–257, 2010, ISSN: 00195421. DOI: `10.1007/s12070-010-0076-7`.

[17] Z. M. Samarakkody and B. Abdullah, "The use of image guided navigational tracking systems for endoscopic sinus surgery and skull base surgery: A re-view", *Egyptian Journal of Ear, Nose, Throat and Allied Sciences*, vol. 17, no. 3, pp. 133–137, 2016, ISSN: 20900740. DOI: `10.1016/j.ejenta.2016.07.005`.

[18] M. Cartellieri, F. Vorbeck, and J. Kremser, "Comparison of six three-dimensional navigation systems during sinus surgery.", *Acta oto-laryngologica*, vol. 121, no. 4, pp. 500–4, 2001, ISSN: 0001-6489. DOI: `10.1080/00016480120004`.

[19] T. Albrecht, I. Baumann, P. K. Plinkert, C. Simon, and S. Sertel, "Three-dimensional endoscopic visualization in functional endoscopic sinus surgery", *European Archives of Oto-Rhino-Laryngology*, vol. 273, no. 11, pp. 3753–3758, 2016, ISSN: 14344726. DOI: `10.1007/s00405-016-4040-3`. [Online]. Available: `%22http://dx.doi.org/10.1007/s00405-016-4040-3`.

[20] B. J. Dixon, M. J. Daly, H. Chan, A. D. Vescan, I. J. Witterick, and J. C. Irish, "Surgeons blinded by enhanced navigation: The effect of augmented reality on attention", *Surgical Endoscopy and Other Interventional Techniques*, vol. 27, no. 2, pp. 454–461, 2013, ISSN: 09302794. DOI: 10.1007/s00464-012-2457-3.

[21] *TLX @ NASA Ames - NASA TLX Paper_Pencil Version*. [Online]. Available: https://humansystems.arc.nasa.gov/groups/TLX/tlxpaperpencil.php (visited on 03/17/2017).

[22] B. J. Dixon, H. Chan, M. J. Daly, A. D. Vescan, I. J. Witterick, and J. C. Irish, "The effect of augmented real-time image guidance on task workload during endoscopic sinus surgery", *International Forum of Allergy and Rhinology*, vol. 2, no. 5, pp. 405–410, 2012, ISSN: 20426976. DOI: 10.1002/alr.21049.

[23] C. Winne, M. Khan, F. Stopp, E. Jank, and E. Keeve, "Overlay visualization in endoscopic ENT surgery", *International Journal of Computer Assisted Radiology and Surgery*, vol. 6, no. 3, pp. 401–406, 2011, ISSN: 18616410. DOI: 10.1007/s11548-010-0507-7.

[24] Y. Otake, S. Leonard, A. Reiter, P. Rajan, J. H. Siewerdsen, G. L. Gallia, M. Ishii, R. H. Taylor, and G. D. Hager, "Rendering-based video-CT registration with physical constraints for image-guided endoscopic sinus surgery", vol. 9415, 94150A, 2015. DOI: 10.1117/12.2081732.

[25] D. J. Mirota, S. Member, H. Wang, S. Member, R. H. Taylor, M. Ishii, G. L. Gallia, and G. D. Hager, "A System for Video-based Navigation for Endoscopic Endonasal Skull Base Surgery", no. c, pp. 1–15, 2011.

[26] S. Leonard, A. Reiter, A. Sinha, M. Ishii, R. H. Taylor, and G. D. Hager, "Image-based navigation for functional endoscopic sinus surgery using structure from motion", vol. 9784, p. 97840V, 2016, ISSN: 16057422. DOI: 10.1117/12.2217279.

[27] S. Rattanalappaiboon, T. Bhongmakapat, and P. Ritthipravat, "Fuzzy zoning for feature matching technique in 3D reconstruction of nasal endoscopic images", *Computers in Biology and Medicine*, vol. 67, pp. 83–94, 2015, ISSN: 18790534. DOI: 10.1016/j.compbiomed.2015.09.021.

[28] T. Wittenberg, C. Winter, and I. Scholz, "3-D Reconstruction of the Sphenoid Sinus from Monocular Endoscopic Views: First results", *Informatik.Uni-Erlangen.De*, no. Bmt, pp. 3–5, 2006.

[29] A. Reiter, S. Leonard, A. Sinha, M. Ishii, R. H. Taylor, and G. D. Hager, "Endoscopic-CT: learning-based photometric reconstruction for endoscopic sinus surgery", vol. 9784, p. 978 418, 2016, ISSN: 16057422. DOI: 10.1117/12.2216296.

[30] M. S. Chen, G. Gonzalez, and R. Lapeer, "Intra-operative registration for image enhanced endoscopic sinus surgery using photo-consistency", *Studies in Health Technology and Informatics*, vol. 125, pp. 67–72, 2007, ISSN: 09269630.

[31] J. H. Nagel, "Live endoscopic video streams augmented for 3D-navigation", *IFMBE Proceedings*, vol. 25, no. 6, 2009, ISSN: 16800737. DOI: `10.1007/978-3-642-03906-5`.

[32] A. Hattori, S. Naoki, and O. Nobuyoshi, "The development of a Real-Time Image-Guided Surgery System for Stereo-Endoscopic Sinus Surgery", *Studies in health technology and informatics*, vol. 142(5), pp. 112–6, 2009.

[33] R. U. Thoranaghatte, J. G. Giraldez, and G. Zheng, "Landmark based augmented reality endoscope system for sinus and skull-base surgeries.", *Conference proceedings : ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, vol. 2008, pp. 74–7, 2008, ISSN: 1557-170X. DOI: `10.1109/IEMBS.2008.4649094`.

[34] G. P. Moustris, S. C. Hiridis, K. Deliparaschos, and K. Konstantinidis, "Evolution of autnomous and semi-autnomous robotic surgical systems: a review of the literature", *The international journal of medical robotics + computer assisted surgery : MRCAS*, vol. 7, no. April, pp. 375–392, 2011, ISSN: 1478-596X. DOI: `10.1002/rcs`.

[35] S. Kapakin, "The paranasal sinuses: Three-dimensional reconstruction, photo-realistic imaging, and virtual endoscopy", *Folia Morphologica (Poland)*, vol. 75, no. 3, pp. 326–333, 2016, ISSN: 16443284. DOI: `10.5603/FM.a2016.0006`.

[36] J. Marescaux, M. Diana, and L. Soler, "Augmented reality in minimally invasive surgery", *Journal of Gastroenterology and Hepatology Research*, vol. 2, no. 5, pp. 555–560, 2013. DOI: `10.6051/j.issn.2224-3992.2013.02.175`.

[37] B. H. Yang and F. L. Lian, "Continuous self-calibration of stereo camera based on particle filter framework for endoscopic videos", *Proceedings of the IEEE International Conference on Industrial Technology*, vol. 2016-May, pp. 1476–1481, 2016. DOI: `10.1109/ICIT.2016.7474977`.

[38] E. Ferrante and N. Paragios, "Non-rigid 2D-3D medical image registration using markov random fields.", *Medical image computing and computer-assisted intervention : MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 16, no. Pt 3, pp. 163–70, 2013, ISSN: 03029743. DOI: `10.1007/978-3-642-40760-4_21`.

[39] Y.A. Rozanov, *Markov Random Fields*. Springer-Verlag New York, 1982. DOI: `10.1007/978-1-4613-8190-7`.

[40] Q. Wei, "Converting 2D to 3D: A survey", *International Conference, Page (s)*, vol. 7, no. December, p. 37, 2005.

[41] L. Maier-Hein, a. Groch, a. Bartoli, S. Bodenstedt, G. Boissonnat, P.-L. Chang, N. T. Clancy, D. S. Elson, S. Haase, E. Heim, J. Hornegger, P. Jannin, H. Kenngott, T. Kilgus, B. Muller-Stich, D. Oladokun, S. Rohl, T. R. dos Santos, H.-P. Schlemmer, a. Seitel, S. Speidel, M. Wanger, and D. Stoyanov, "Comparative

validation of single-shot optical techniques for laparoscopic3D surface reconstruction", *IEEE Trans. Medical Imaging*, vol. 33, no. 10, pp. 1913–1930, 2014, ISSN: 1558-254X. DOI: `10.1109/TMI.2014.2325607`.

[42] L. Lei, J. Li, M. Liu, X. Hu, and Y. Zhou, "Shape from Shading and Optical Flow used for 3-Dimensional Reconstruction of Endoscope Image", *Otolaryngology - Open Journal*, vol. 1, no. 1, pp. 20–23, 2015, ISSN: 24704059. DOI: `10.17140/OTLOJ-1-106`.

[43] G.-h. Wang, J.-q. Han, and X.-m. Zhang, "Three-dimensional reconstruction of endoscope images by a fast shape from shading method", *Measurement Science and Technology*, vol. 20, no. 12, p. 125 801, 2009, ISSN: 0957-0233. DOI: `10.1088/0957-0233/20/12/125801`.

[44] A. Kumar, Y. Y. Wang, C. J. Wu, K. C. Liu, and H. S. Wu, "Stereoscopic visualization of laparoscope image using depth information from 3D model", *Computer Methods and Programs in Biomedicine*, vol. 113, no. 3, pp. 862–868, 2014, ISSN: 01692607. DOI: `10.1016/j.cmpb.2013.12.013`.

[45] "Panorama visualization", *Three-dimensional endoscopic visualization in functional endoscopic sinus surgery Atul Kumar, Yen-Yu Wang, Kai-Che Liu, Wan-Chi Hung, Shih-Wei Huang, Wen-Nung Lie, Ching-Chun Huang*, 2015.

[46] D. Burschka, M. Li, R. Taylor, and G. D. Hager, "Scale-Invariant Registration of Monocular Stereo Images to 3D Surface Models", in *Proceedings of 2004 IEEElRSJ Internalional Conference on Intelligent RobOts and Systems*, 2004, pp. 2581–2586, ISBN: 0780384636.

[47] S. Bernhardt, "Computer-Assisted and Robotic Endoscopy", vol. 9515, pp. 59–68, 2016. DOI: `10.1007/978-3-319-29965-5`.

[48] M. S. Nosrati, R. Abugharbieh, J. M. Peyrat, J. Abinahed, O. Al-Alao, A. Al-Ansari, and G. Hamarneh, "Simultaneous Multi-Structure Segmentation and 3D Nonrigid Pose Estimation in Image-Guided Robotic Surgery", *IEEE Transactions on Medical Imaging*, vol. 35, no. 1, pp. 1–12, 2016, ISSN: 1558254X. DOI: `10.1109/TMI.2015.2452907`.

[49] D. Hutchison, J. C. Mitchell, X. Pan, and Y. Fan, "Visualization Techniques for Augmented Reality in Endoscopic Surgery", vol. 1, pp. 404–414, 2016, ISSN: 0302-9743. DOI: `10.1007/978-3-319-43775-0`.

B Optical tracking system

Figure 42: Vector vision compact 4

Besides EM tracking, another option is available; optical tracking. With this system, the tools are equipped with reflectors (small infrared reflecting balls) rather than with sensors. The system has two camera's which have infrared (IR) pass filters in front of the lenses and IR LEDs around the lenses. The light is reflected into the camera's and the position of the balls is detected, thereby the position and orientation of the tools can be determined. The advantage of an optical tracker is that there is no distortion caused by the amount of metal in the endoscope.
At the Utwente a quite old system is used, the "Vector vision compact 4" from manufacturer Polaris, which can be seen in figure 42. Disadvantage of this system is that it is very difficult to manipulate the code from within and calibration of tools with the system is cumbersome and in various situations appeared not to be successful at all. Tool calibration with the system is necessary since the output is only the position and orientation of the tool, the internals about the position of the reflectors are not allowed to be accessed for the user. This implies also that the definition of a new tool, for example if the endoscope would be equipped with reflectors, is not easy or even not possible to be made. Tests for point registration with one of the standard probes of the system resulted in reprojection errors of more than 10mm. Due to the inconvenience of this system, and the fact that in the AvL the EM system is currently used for PDT in the sinuses, it is decided to not investigate further in the use of the optical tracking system.

C Calibration registration procedure with $^{CS}T_t$

In this section the original calibration and registration step-by-step plan is given, when there was made use of the $^{CS}T_t$ matrix troughout the procedure. Steps that are equal to the ones in the new procedure are omitted.

1 **$^{CS}T_t$ registration**:
   Registration of the endoscopic tip with respect to the sensor attached to the endoscope in EM coordinates
2 **$^{W}T_{EM}$ registration**:
   Registration of the EM coordinate system to the world coordinate system
3 **$^{W}T_C$ callibration**:
   Camera calibration to register the camera coordinate system to the world coordinate system
4 **$^{t}T_C$ estimation**:
   Estimation of the transformation from the endoscopic tip to the camera coordinates using the calculated transforms
5 **$^{CT}T_{HS}$ registration**:

Registration of the tip of the endoscope in the EM coordinate system, with respect to a reference sensor, to the CT coordinate system

6 **From input to output; all transformations combined**:
Combination of all information gathered in the previous steps in order to get the real-time position of the endoscope tip in camera coordinates

C.0.1 Step 1: $^{CS}T_t$ registration

Some surgical tools have a build-in sensor and are pre-registered to the EM system, but the endoscope is not one of them. So if we attach a sensor to it, the transformation from this sensor to the position and orientation of the tip still have to be determined. We define two matrices: $^{EM}T_{CS}$ is the position and location of the sensor in EM coordinates, this is the direct output of the EM system, and $^{EM}T_t$, the location and orientation of the tip of the endoscope, this one is not known. In order to find this transform, the following formula is used: $^{EM}T_t =^{EM} T_{CS} \, ^{CS}T_t$. To find $^{CS}T_t$, some code is written to register tooltip and sensor to each other.

This code contains a registration step using the EM pointer tool, a 3D registration cube, and the sensor attached to the endoscope. The cube acts as a reference object with respect to which the tool-tip position is defined. The original EM pointer tool outputs exactly where it's tip is in the EM coordinate system, and what the orientation of the tool is. Positioning the tool in the holes in the cube gives the transformation matrices:$^{EM}T_{TS,h}(i)$, with h_i is the i-th hole in the cube. Because we position the tool in the same orientation in the same holes, $^{EM}T_{TS}$ and $^{EM}T_t$ are assumed to be equal.The sensor on the endoscope outputs it's position and orientation as well, when the endoscope is positioned in the same holes, giving the matrices: $^{EM}T_{CS,h}(i)$. When doing this, not only the position of the tool and the endoscope should match for the corresponding reference hole, but also the orientation. This can be established by orienting the upside of both the tool and the endoscope towards the number indicated next to the hole. Then the transforms can be calculated as follows:

$$^{CS,h}T_{TS,h}(i) =^{CS,h} T_{TS,h}(i) = \; ^{EM}T_{CS,h}(i)^{-1} \, ^{EM}T_{TS,h}(i)$$

The rotation and translation of these transforms are averaged to obtain the transformation matrix $^{CS}T_{TS} \approx \; ^{CS}T_t$.

C.0.2 Step 2: $^{W}T_{EM}$ registration

The world coordinate system is defined by a 9x6 checkerboard with 10x10 mm squares. In order to link this information to the EM system, code is written which asks you to touch the corners of the checkerboard with the tip of the endoscope. Using the $^{CS}T_t$ registration, the position of these corners in EM coordinates is determined. The Procrustes algorithm from Matlab is used in order to find the transformation.

$$^{W}T_{EM} = \text{Procrustes}(^{W}P, ^{EM}P)$$

C.0.3 Step 3: $^{W}T_C$ calibration

Equal to the step in the new procedure.

C.0.4 Step 4: ^tT_C estimation

As written in the previous step, when the pictures are taken also the EM position of the sensor of the endoscope is stored in the matrices $^{EM}T_{CS}(i)$, where i indicates the ith image. With use of the previously calculated matrices we can now transform this into the position of the tip in world coordinates using the following formula:

$$^WT_t(i) =^W T_{EM} \; ^{EM}T_t(i)$$

The transformation matrix to go from endoscopic tip coordinates to the camera coordinates can than be determined for every image:

$$^CT_t(i) =^C T_W \; ^WT_t(i)$$

In order to find ^CT_t, the rotation and translation of the $^CT_t(i)$ matrices are averaged.

C.0.5 Step 5: $^{CT}T_t$ registration

The registration step is mostly equal to the new procedure. Only the endoscopic tip is used instead of the tool tip. Therefor in order to find the registration, also the transformation from the tool sensor to the tool tip needs to be applied.

The tool tip is moved around on the registration surface, and the data of it's orientation and position in EM coordinates with respect to the reference sensor $^{CS}T_{HS}$, is stored. This is transformed into $^tT_{HS} =^t T_{CS} \; ^{CS}T_{HS}$. Then an Iterative Closest Point (ICP)-algorithm is used to make the best $^{CT}T_{HS}$ to match the two point clouds.

C.0.6 Step 6: From input to output; all transformations combined

When the coordinates from the EM sensors enter the computer, this information has to be transformed into the position and orientation of the tip of the endoscope in camera coordinates. With help of the preceding steps all bits of information are available, and only a number of calculations needs to be performed in order to be able to find the camera coordinates necessary for a nice virtual visualization. One input is the matrix ^eT_s with the sensor transform in EM coordinates, the transform of the endoscopic tip can be found with the $^{CS}T_t$ transformation:

$$^{EM}T_t =^{EM} T_{CS} \; ^{CS}T_t$$

Another input is the matrix $^{EM}T_{HS}$ from the reference sensor. Together with the registration $^tT_{HS}$, this can transform the coordinates of the endoscope tip to the CT coordinate system:

$$^{CT}T_t =^{CT} T_{HS} \; ^{EM}T_{HS}^{-1} \; ^{EM}T_t$$

The last step is to transform into the camera coordinates by applying the matrix ^tT_C:

$$^{CT}T_C =^{CT} T_t \; ^tT_C$$

The resulting transform gives the position and orientation of the camera in the CT data in order to match the real view of the endoscopic camera. A visualization protocol is designed using this information.

Table 5: Error sources

Errors		transl (mm)	ang (°)
Calibration			
1.1	Pinpointing error with endoscope on corners of checkerboard	0.5	-
1.2	EM measurement error	0.48	-
1.3	Localization of the corners error	0.4	0.005
1.4	EM system measurements error (locations of photo's)	0.7	0.3
Registration			
2.1	EM measurements error endoscope sensor	0.7	0.3
2.2	EM measurements error head sensor	0.48	0.3
2.3	Pinpointing error with endoscope on surface of phantom	3	-
3.1	EM measurements error endoscope sensor	0.7	0.3
3.2	EM measurements error head sensor	0.48	0.3
4	error from the $^{CS}T_C$ registration	..	..

D Monte Carlo analysis for registration with use of the endoscope

In the final design the registration is performed using a pre-calibrated tool of the EM system. Before this was done, the endoscopic tip was used for this, just like it was used for the registration itself. The protocol for the registration with use of the endoscope is mostly the same as for the registration with the tool, only the information from the calibration is taken into account, thereby introducing an extra source of errors. The error sources as initially used for this analysis are in table 5. The error for the camera calibration can only be filled in when that analysis is done.

D.1 Results of Monte Carlo analysis with use of the endoscope and different parameters

The Monte Carlo Analyses have been performed as explained before. The error sources are applied separately and simultaneously in order to find their influence on the behavior of the programs and to find the average error that is introduced. For the camera calibration N, the number of iterations of the algorithm, is chosen to be 1000, for the registration, due to the computational load of the ICP algorithm and time constraints, M is chosen to be 100. For the camera calibration 20 pictures of the checkerboard are taken.

Table 6: Error in the output with N=1000, M=100 and number of images is 20

Error source	Translational error	angular error
Calibration		
all sources zero	0.0	0.0
1.1	0.2535	0.0.1078
1.2	0.5529	0.2556
1.3	0.1558	0.0654
1.4	0.3738	0.1158
1 (total)	0.8903	0.3808
Registration		
all sources zero	1.5308	0.4854
2.1	1.8563	0.6877
2.2	1.7499	0.5760
2.3	2.15	0.9554
2 (subtotal)		
3.1	2.3919	0.6922
3.2	2.4198	0.6588
4	2.0079	0.5594
3 (subtotal)	2.5715	0.7685
2 and 3 (total)	2.6797	0.9711

Because the endoscope consist of a large amount of metal, the signal of the sensor that is attached to it, might be prone to a bigger distortion. If the translational part of the error (1.4) is taken to be 1.7mm and the rotational 0.7, the new outcomes are in table 7. This also gave a FRE that compared to the real FRE during the measurements.

More images and measurements could be taken to improve the performance. In the

Table 7: The effect of an enlarged distortion of the sensor data from the endoscope for N=1000 and 20 images

Error source	Translational error	angular error
Calibration		
all sources zero	0.0	0.0
1.1	0.2535	0.1078
1.2	0.5529	0.2556
1.3	0.1558	0.0654
1.4	0.9204	0.2774
1 (total)	1.1954	0.4523

Table 8: Error in the output with N=1000 and number of images is 80 and enlarged distortion parameters

Error source	Translational error	angular error
Calibration		
all sources zero	0.0	0.0
1.1	0.2862	0.1089
1.2	0.2734	0.1084
1.3	0.0726	0.0304
1.4	0.3972	0.1252
1 (total)	0.5879	0.2017

table below, the results are shown for a calibration with use of 80 images and the parameters as set before.

D.2 Discussion of old results

The analysis shows that camera calibration with 20 images of a $9x6$ checkerboard, using the new parameters, gives an average translational accuracy of 1.19 mm. This is slightly larger than the demanded 1 mm accuracy. It can be seen that the distortion of the sensor that is attached to the endoscope is the factor that influences this error most. In order to decrease it's impact, more measurements could be taken. This is tested by taking 80 pictures and the results are in table 8. Note that due to this being a different data set, the other errors might also differ from the original ones. Now the error is below the demanded 1 mm.

E Manual for making smooth slices in a 3D stl model

Manual for making smooth slices with Meshmixer

1. Download and install meshmixer: http://www.meshmixer.com/download.html
2. Import your stl file:
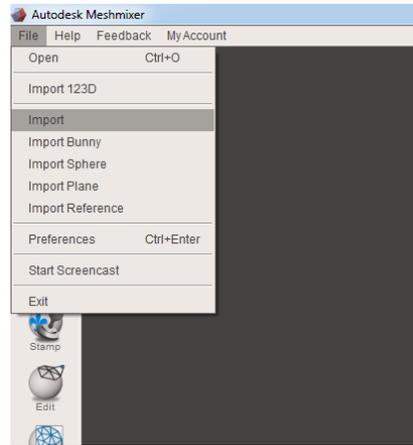
3. Import Plane, choose append to add it to your model:
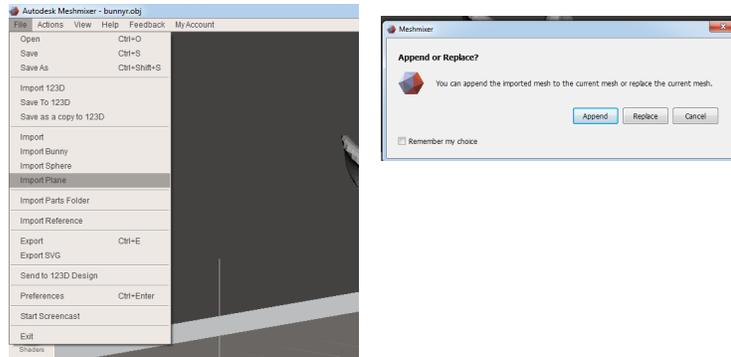
4. Resize the plane to make it visible and larger than your model:
Select plane object in the Object browser, click **edit**, click **transform**, change the x and z coordinates. If "uniform scaling" is switched on you can also change the scale of x, then the other coordinates will be scaled accordingly. The blue lines indicate the format of the resized plane. Click **accept** if you are satisfied.
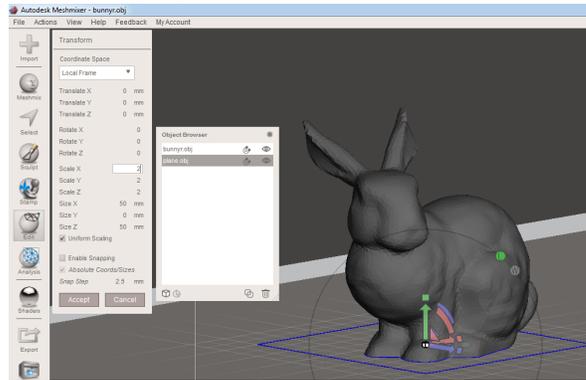
5. Move the plane to the height of the bottom of the slice you want to have, say height "a" in the picture:

Select plane object in the Object browser, click **edit**, click **transform**, translate y to height a. Click **accept** if you are satisfied.
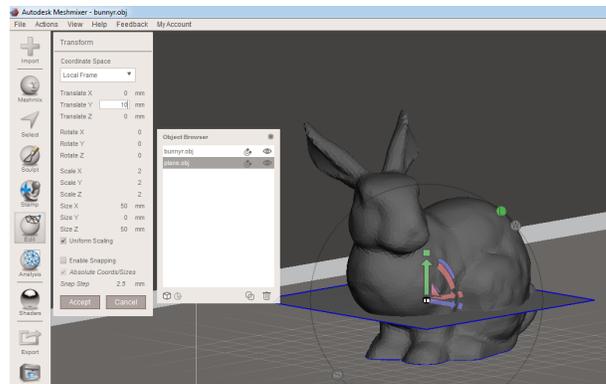
6. Create a pivot on the plane:
   Select plane object in the Object browser, click **edit**, click **create pivot**. Click on the plane, click on **drop pivot** click on **done** if you are satisfied. A orange ball should appear on the plane when you click on it and it should turn grey when you click on drop pivot.
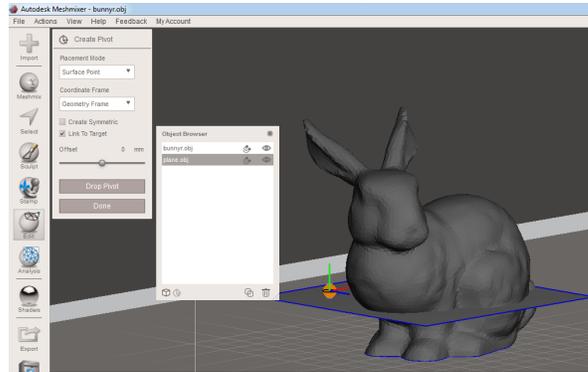
7. Delete part below the slice:
   Select your model in the Object browser, click **edit**, click **plane cut**. Click on the grey ball of the pivot plane. Set cut type to **disregard half** and fill type to **remeshed** fill. The translucent part will be deleted; this is indicated with the large blue arrow. If you want exactly the opposite part to be deleted, click the arrow and it will change direction. If you are satisfied click **accept**.
   If you can't see the big blue arrow it might be handy to turn around your model, press the ALT key and click and move your mouse to do so.
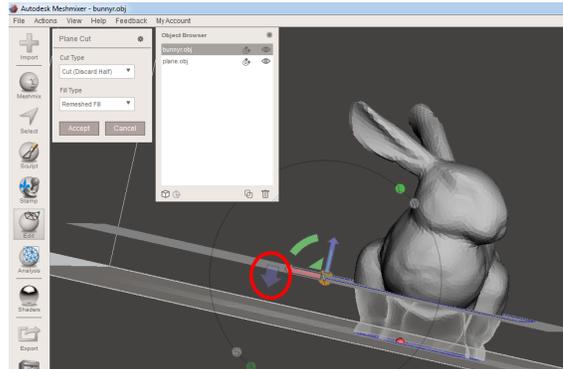
8. Move the plane to the height of the top of the slice you want to have, say height "b" in the picture:

Select plane object in the Object browser, click **edit**, click **transform**, translate y to height h. Remember that you deleted the bottom of the model, so the plane is now at zero again. You only have to move it up for the distance between a and b. Click **accept** if you are satisfied.
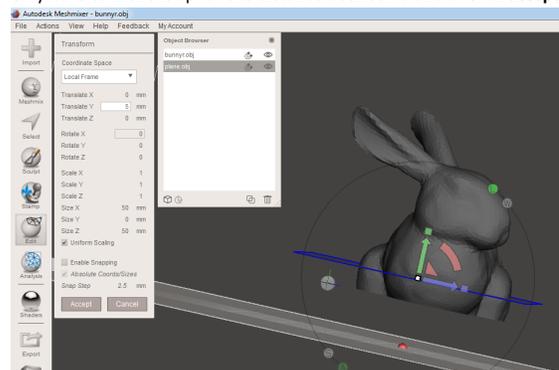
9. Delete part above the slice:

Select your model in the Object browser, click **edit**, click **plane cut**. Click on the grey ball of the pivot plane. Set cut type to **disregard half** and fill type to **remeshed** fill. The translucent part will be deleted; this is indicated with the large blue arrow. You probably want exactly the opposite part to be deleted, so click the arrow and it will change direction. If you are satisfied click **accept**.
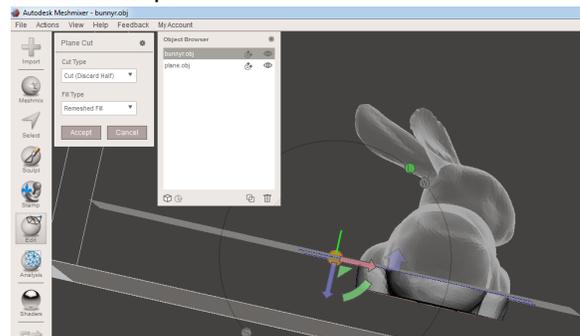
10. Your smooth slice of height a-b is ready!

    Select plane object in the Object browser and delete it by clicking the small trashcan symbol in the right bottom corner of the Object browser.

    Select your model in the Object browser, click **analysis** and then **units/dimensions,** you can see that your model has the desired height.
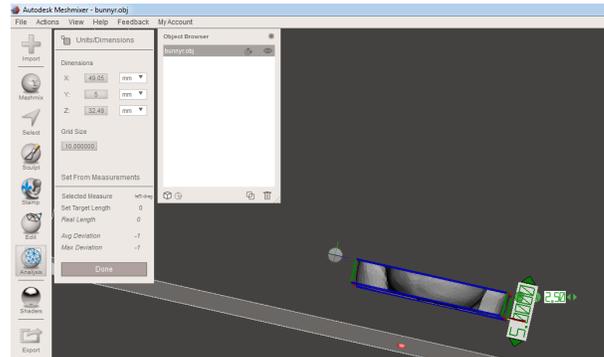
11. This program does not check whether the model is printable. Export the model as stl file and make it printable with for example Netfabb. This also has a very good online service where you can just upload your model and it will fix it: https://service.netfabb.com