

BACHELOR OPDRACHT

AN AUTOMATED DAVID
LASER SCANNER FOR
FACIAL ACQUISITION

Niek Moonen

TNW/EWI
SIGNALS AND SYSTEMS

EXAMENCOMMISSIE
DR. IR. L.J. SPREEUWERS
MSC C. VAN DAM
DR. R. LUTTGE

DOCUMENTNUMMER
EWI SAS 2012-003

15-02-2012

1. Introduction

“Face recognition has become one of the most active research fields in computer vision due to increasing demands in commercial and law enforcement applications.” There has been much research on two dimensional face recognition, simply because the acquisition of 2D faces is easy and widely used in practice. This research trend is gradually shifting towards the 3D research plane, since it is promising to overcome the difficulties and limitations associated with face recognition in 2D space. Thus the need for 3D face acquisition rises. [1]

3D face acquisition techniques can be divided into two classes, active and passive. Active techniques apply external lights such as a laser beam or a structured light pattern. Passive techniques use multiple intensity images of a scene usually without active illumination.

The passive techniques are less intrusive and more covert, which is why it is mostly developed for commercial and law enforcement use. Active techniques are mostly used in controlled environment, like research labs.

The most commonly used active acquisition techniques are the laser scanning and structured light projection techniques. They both use the same triangulation principle to determine where a point in the picture is in 3D space. The difference between them is that the laser range scanner uses one laser line and multiple images, while the structured light projection uses a projected pattern and requires only one or a few images.

Completely automated laser scanners are very expensive. The DAVID laser scanner is nearly a factor 10 cheaper and it promises reasonable results. However the laser projection line is moved manually. The DAVID scanner starter kit consists of a webcam, laser line projector, DAVID software and calibration backgrounds.

This report focuses on the creation of an automated three dimensional laser scanning device from the starter kit and determining the quality and accuracy of the results. There will be a brief theoretical explanation of a laser scanner in general and the DAVID laser scanner will be described, then the construction process will be described and after that there will be an evaluation of the scanner. But first the assignment description will be given.

The assignment consists of 2 parts. The first part is the automation of the scans. The second part is evaluation of the scans. The following main research questions need to be answered:

- Create a mechanical structure to control the DAVID laser scanner.
- What is the quality of the 3D scans?
- What is the accuracy of the 3D scan?

To answer these questions first some background information about 3d face acquisition techniques is given. Then the building process of the 3D scanner setup is described, followed by an evaluation of the 3d scan results. But before the evaluation is done, a few preliminary experiments are performed.

2. Background

As shown in the introduction, we divide the 3d image acquisition techniques in two major categories, namely passive and active techniques. First stereo vision is described, since it's a good example of a passive acquisition technique. Then the active techniques laser scanning and structured light scanning are explained.

Passive acquisition technique: Stereo Vision[1]

The depth information of the 3d image is recovered from two 2d images that have slightly different viewpoints. So the points in one image will be shifted in the other. But the technique requires a calibration step. When the calibration is done and the images are taken, building a 3d model requires performing two major tasks: matching and reconstruction. There are 3d model based matching and structured light assisted matching techniques. Using structured light makes stereo vision an active 3d image acquiring technique. A 3d model based matching technique may try to minimize the amount of curvature between the evolving 3d face and the 3d model. This introduces inaccuracies which are a problem of model based stereo vision.

The calibration step is made easier if two identical stereo cameras are used, since they have the same intrinsic parameters. The intrinsic parameters are camera characteristics such as focal length, coordinates of the principal point and size of pixel in terms of length and width. But the extrinsic parameters still need to be estimated. This is the relative position and rotation of the cameras. Another advantage of using two identical cameras at the same time is that the capturing is very fast, while there are two images taken at the same time.

In general it can be said that stereo vision has better in plane accuracy than laser scanners, but for the accuracy in depth it's the other way around. But Boehnen en Flynn[6] showed that a laser scanner has a higher in plane accuracy than stereo vision. Which is contradictory with the fact that a laser scanner hides details within stripes whereas stereo gives access to pixels and hence should be able to reveal more details.

Active acquisition technique: Laser scanning and structured light projection

As mentioned in the introduction, the techniques are quite similar. They both use the same triangulation principle. But the difference between them is that the laser range scanner uses one laser line and multiple images, while the structured light projection uses a profile of lasers and requires only one shot.

The triangulation principle is common to both techniques. It is a way of mapping the x - and y -coordinates of the image to the X -, Y - and Z -coordinates of the world. This principle is displayed in figure 1. The depth is recovered by applying the following equation.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{b}{(f \cdot \cot(\theta) - x)} \cdot \begin{bmatrix} x \\ y \\ f \end{bmatrix}$$

But for this to work, the parameters f , b and θ need to be known a priori.

f =focalpoint, b =distance of the laser from the camera's optical axis, θ =intersection angle. These parameters are calculated in the calibration process.[1]

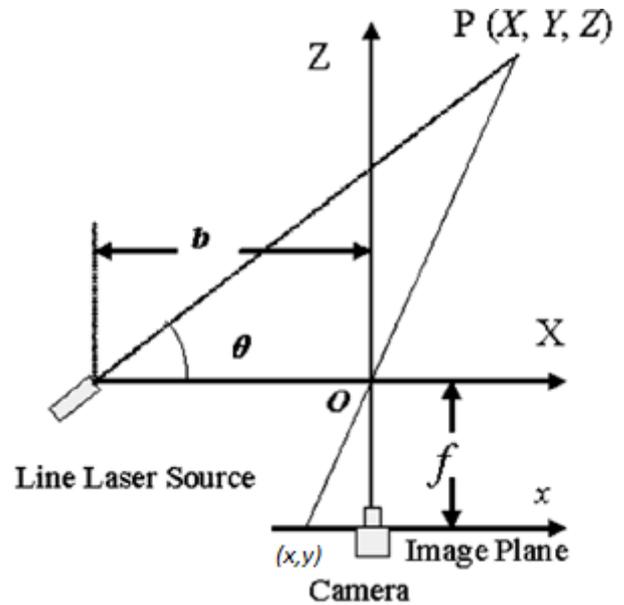


Figure 1: Basic geometry of active triangulation from laser scanning [1]

This is an important step in the acquisition of 3d objects by means of a laser scanner. In this step the relation between the image coordinates and the world coordinates is determined. A pattern with known and marked xyz-points is used for determining this relation. If the relation is known, the processing of the acquired data is fast and of low cost. This makes it useful as a demonstration model.

The accuracy of the depth measurement is in the millimetre area and the field of view is up to several meters. This is ideal for 3d face acquisition. Faces are of an average size of 20 centimetres and the depth accuracy of micro- to millimetre is more than enough, because faces can change several millimetres or more over time. So the field of view and the depth range of triangulation are sufficient for scanning faces. This is an advantage of using these active techniques.[1]

On the other hand, laser scanning needs the cooperation of the subject/object that is being scanned. Other, usually passive, techniques for 3d face acquisition are more covert and thus easier to implement for example surveillance. This isn't a requirement for this project. However should we consider the safety of the person being scanned. So the laser needs to be safe for the eyes.

Another problem with using a laser range scanner is the duration of the scan. A cooperating subject has to stay still for up to 40 seconds depending on the technique and accuracy of the acquisition. Photometric stereo vision can acquire a 3d face model in under a second by means of a high speed camera and short light flashes from different angles. The duration of the laser range scanner can be shortened by using a camera which can capture more frames per second.

3. DAVID Laser scanner

The DAVID laser scanner consists of four main parts. A webcam, a laser line projector, the DAVID software and calibration backgrounds of different sizes. *Note: "the figures 2 to .. are examples taken from the manual found on the website of DAVID, the values and settings shown are not used!"*

The Webcam

The webcam used is a "Logitech quickcam 9000 Pro". It has a native resolution of 1600x1200 pixels. But there are multiple resolutions and frames per second to choose from. Ranging from 1600x1200 to 640x480 pixels and 30fps to 5fps. Unfortunately the latest webcam driver (2.x) is not ideal for use with DAVID. The latest driver has fewer camera setting options and the settings don't show any values, so it will be hard to create a stable experiment environment. So we installed the older (1.x) version of the driver.

The Laser

The laser is class 1 laser module producing red light with 650nm. It has an adjustable focus and its power supply is a battery.

Calibration backgrounds

The package contains cardboard panels with calibration patterns in different sizes. The four calibration panels have grid sizes of 30mm, 60mm, 120mm and 240mm. They are suitable for scanning objects of sizes between approximately 50mm and 300mm.

Software

When using the DAVID laser scanner software, you are guided through three steps. The camera calibration, 3D laser scanning and 3D shape fusion.

The camera calibration

In figure 2 the calibration step of the DAVID software is shown. First the camera used is selected with the desired resolution and frames per second. Also the scale of the calibration background has to be entered. This also means it's possible to create a larger version of the calibration pattern.

The camera settings are important for a successful camera calibration. Figure 3 shows an image of the background taken with the ideal camera settings. Notice how bright the image is and the contrast with the black dots.

Also important for the calibration is the 90 degree angle of the background. If this is not exactly 90 degrees, a scan of a sphere for example will result in something more egg-shaped.

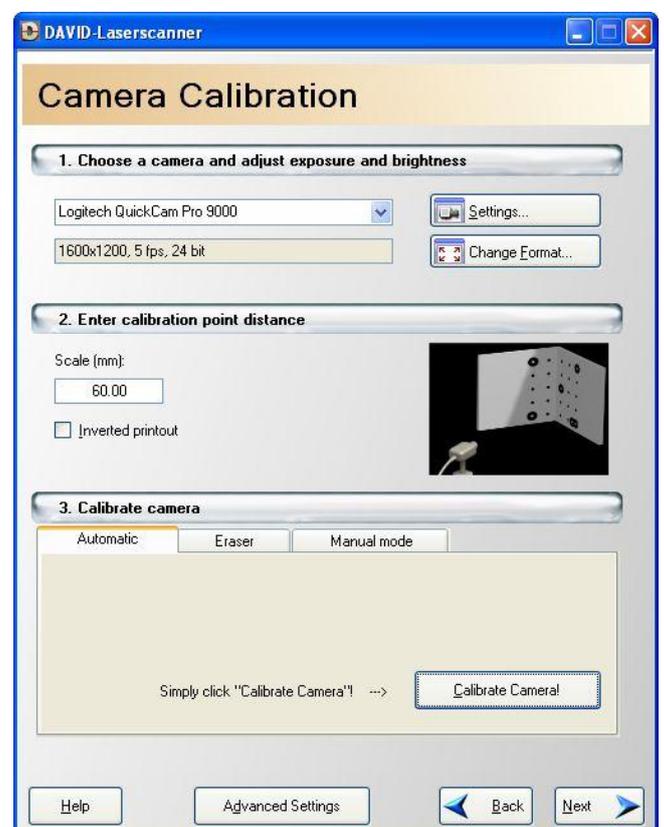


Figure 2: Calibration step of DAVID laser scanner software

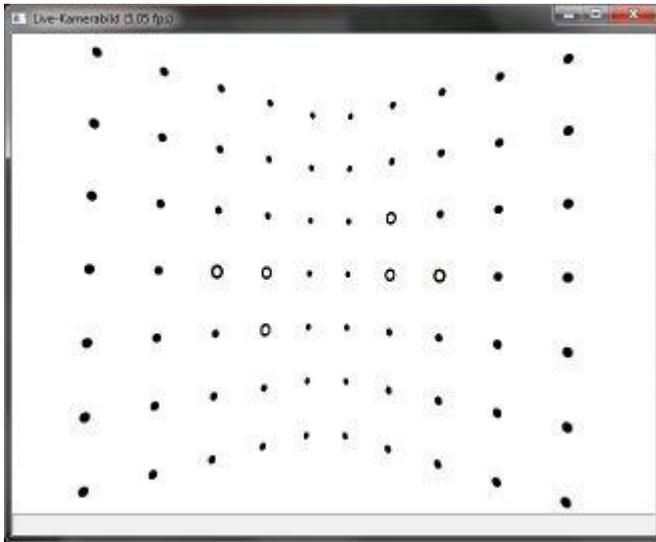


Figure 3: Image of calibration background taken with ideal camera settings

3D laser scanning

Figure 4 shows the 3D laser scanning step. It consists of 3 separate sub steps. Scanning the object, grabbing the texture and post processing of the scan.

For clarification, figure 5 shows the object that is being scanned. The calibration pattern is different, since the images are taken from the DAVID manual that was made using the first version of the background pattern.

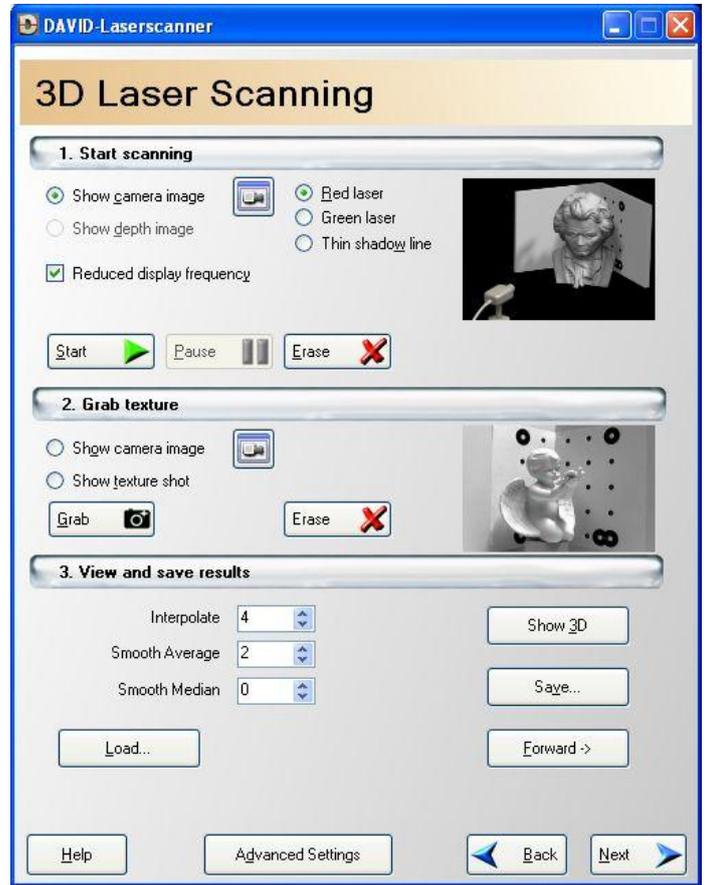


Figure 4: 3D laser scanning step of DAVID laser scanner software



Figure 5: scanning object

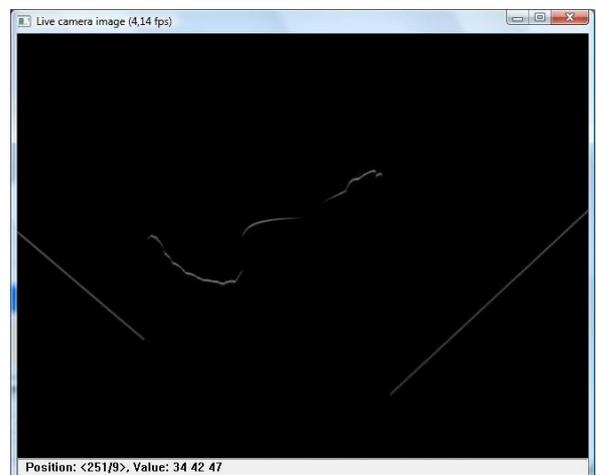


Figure 6: Example of good live camera image when scanning

Before scanning the object, the camera settings need to be set. Figure 6 shows an example of good live camera image for scanning. Notice that it is completely dark and only the laser line is clearly visible. Here the laser line is white, but this is due to a very low saturation setting. In figure 7 the adjustable camera settings are shown. These values were not used to create figure 5, the image is an example showing the different adjustable settings. In chapter 5 “setup setting determination”, the used settings are shown.

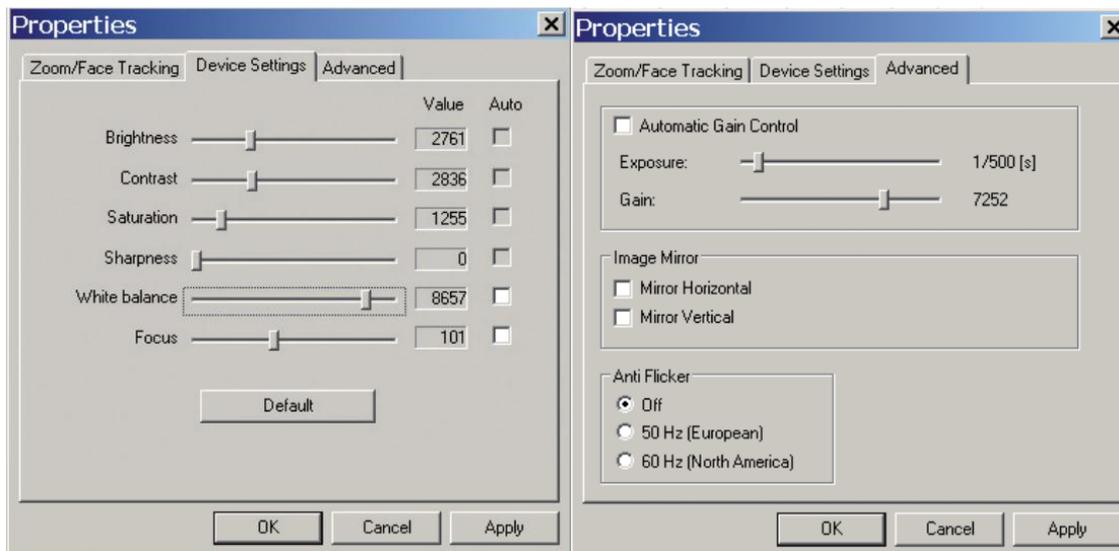


Figure 7: adjustable camera settings (Logitech camera driver version 1.x)

As soon as the scan is started, the live camera image is replaced with the scan result shown in figure 8. This is a depth map of the image, but the colours are repeating themselves. So this it's difficult to interpret the depth image. For example, in figure 8 the purple spot on the left of the image (the right wing of the angel) lies in a different depth plane in comparison to the purple spot on the bottom (the right knee of the angel) of the scan.

Figure 8 also shows large gaps in the scan result. This can have multiple causes. For example the laser line could not have been detected, the intersection angle could have been too low, the laser was moved to quickly over that part of the object or the computer was not able to keep up the calculations with the live feed. If the computer is not able to keep up the "reduced display frequency" option is recommended. It's shown in figure 4. It lets the scan result only be updated once every second. Depending on the hardware, this can increase the scanning speed.

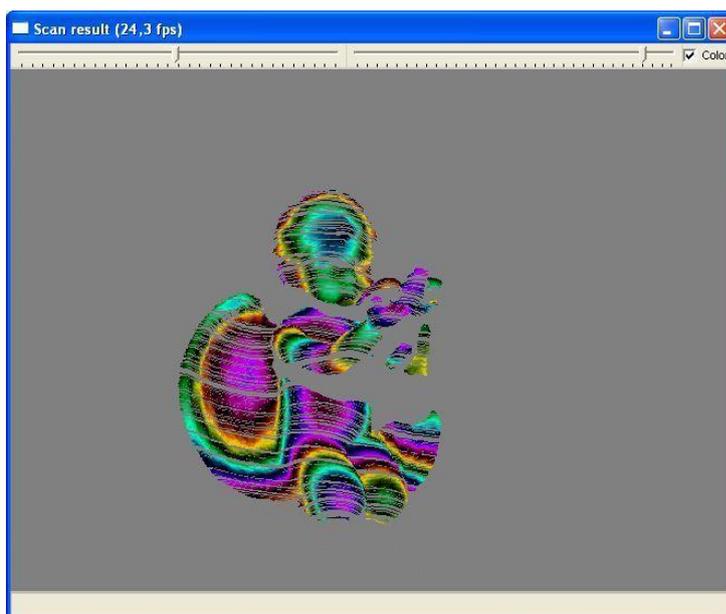


Figure 8: Scan result of scanning the object shown in figure 5

It is optional to grab a texture from the 3D scan. The “Grab Texture” button grabs a camera shot that will be used to texturize (i.e. colorize) the 3D scan. To get a high quality texture, the illumination needs to be uniform and bright. No reflections, shades or shadows should be visible in the image. For acquiring the texture image, different camera settings are needed and are set in the same way as shown by figure 7. But they are set with different image criteria.

Also it's optional to post process the data. Figure 4 shows it's possible to interpolate, smooth median and smooth average. These options will generate extra data points from the original captured set. This will enhance the image quality artificially.

3D shape fusion

Shape fusion is used to combine multiple scans into an all-round model. For example, when scanning a head, only part of that head is visible to the camera. So if a head is scanned from the front, the side of the head is not visible in the scans. By acquiring multiple scans that are rotated compared to each other and then fusing them, it's possible to create an all-round model. This is shown in figure 9.



Figure 9: Shape fusion of a Beethoven bust. (Top image) The different scans that are being fused. (Lower images) The fusion result, from the front and back.

4. Building the setup

Our first task is to create a mechanical structure to control the DAVID laser scanner. First a realisation of the setup will be sketched and after stating the design criteria, the different main components are discussed and a full description of the acquisition process is described.

First intuitive realisation of the demonstration model:

For creating the setup a few different components were made available:

- The starter-kit of the DAVID laser scanner
 - o Laser line projector
 - o Webcam, Logitech quickcam 9000 Pro
 - o Calibration pattern background
 - o David laser scanner software
- A XYZ-table controllable by computer
- Dedicated computer

A sturdy table was used as a platform where the XYZ-table could rest. The platform table is also used as a workspace with a computer that is dedicated to the demonstration model. This is not shown in figure 10, since the picture was taken in an early stage of development.

The distance between the camera and face/object that is being scanned is about 30-50 cm's. The camera's height is such that it is positioned lower than the mounted laser and almost at the same height as the face/object. This is done to avoid as much gaps as possible in the scans. The camera isn't attached to the table where the xyz-table is positioned on. The vibrations in the moving laser might distort the image captured by the camera. Instead it's mounted on a tripod that is adjustable in height. This is done for easy switching between scanning of objects and persons.

The setup has the following design criteria:

- the maximum duration of the scan should be around 30 seconds
- the setup should be user friendly
- the setup must be able to scan a human face
- a face should cover approximately 50% of the image
- the camera should not be attached to anything that is in contact with the xyz-table

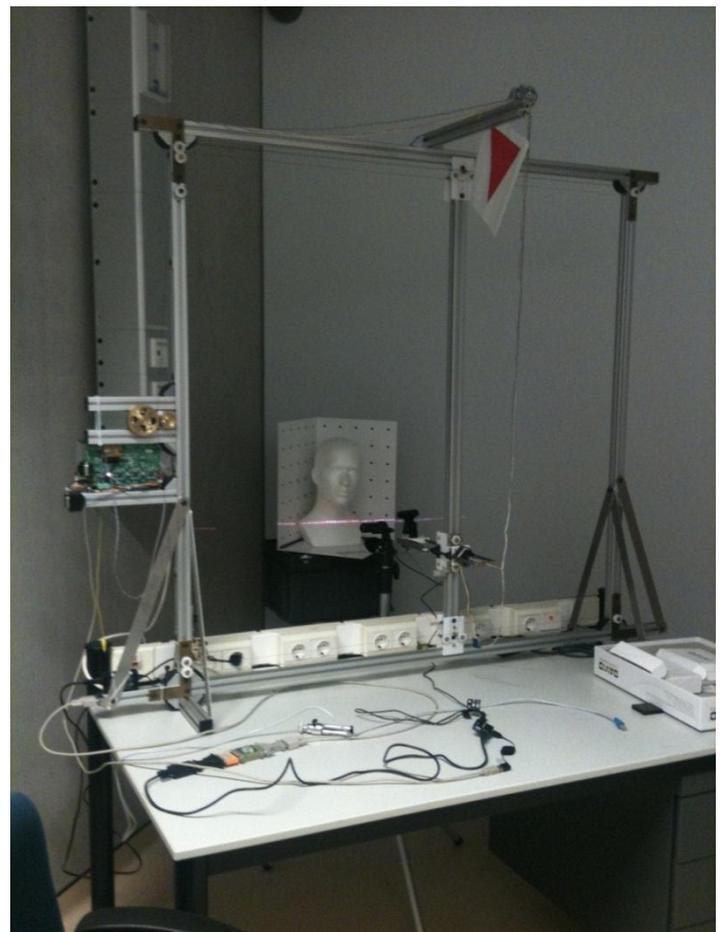


Figure 10: XYZ-table. x-axis is from left to right, y-axis is up and down, and the z-axis is towards and away from the scanning object.

XYZ-table

The XYZ-table was designed at the University of Twente by M. Verhoeve. The XYZ-table is a construction that is able to position a mounted object in 3 dimensions within certain boundaries. This is to be seen in figure 2. The structure is almost 2 by 2 meters and within these parameters the object can be moved autonomously through the use of matlab.

The setup should be as easy and as simple as possible, so almost everyone can use it without knowing much about the setup. A graphical user interface was made to aid in the control of the XYZ-table. Specific instructions were written in it to clarify the GUI. Before explaining how this was done, the available software is discussed.

Scanning speed

When scanning faces it is assumed that people can't sit still for a very long time and any movements of the head would result in inaccuracies of the scans. So we set the goal to scan a person's face within 30 seconds. The GUI is created with an option to easily control the speed with which the laser is being moved. But only the speed in the Y-axis is controllable, since this is the only axis that has an influence on the scan. The Y-axis from the table is defined as the upward and downward movement of the laser. As shown in figure 2. The speed is also presumed to have a certain influence on the accuracy of the scans. The faster the scan, the less accurate the results will be. Here the assumption is made when moving the laser at a higher speed less data points are captured.

Vibrations

The vibrations emanating from the movement of the xyz-table could introduce noise or other inaccuracies in the scans. In the mechanical realization of the setup there are no (quantitative) indications this is the case.

Webcam

The webcam used is a "Logitech quickcam 9000 Pro". It has a native resolution of 1600x1200 pixels. It has to be positioned at approximately 50cm from the calibration background. This is done so a face will cover approximately 50% of the image. Also it is mounted on a tripod, so the vibrations of the XYZ-table wont effect the camera.

Resolution

There are multiple resolutions and fps to choose from. Ranging from 1600x1200 to 640x480 pixels. We are going to use the resolution of 800x600 pixels, since it's half the native resolution. According to the moderators of the David laser scanner forum, when you don't use factors of the native resolution, then it's possible that your scans are compromised with artefacts and distortions. The native resolution is 1600x1200 pixels, but it seems it can only record on 10fps, which probably means the scan has to be slowed down to get a good quality. If a scan is done on full speed at 10 fps the images might show gaps, since there are fewer images grabbed and therefor the spacing between the laser line in following images will be larger. This is why we choose to use 800x600 at 30 fps. Here full speed is the speed needed to scan a face within 30 seconds. Since this is the goal, we have to determine the trade-off between resolution and scanning speed. This will be done in the evaluation part of the report.

Frames per second (FPS)

Not only the frames per second of the camera are of importance in our setup. The software is showing signs of not being able to cope with the speed of the scans. While in scanning mode the resulting depth

image is being shown with the fps it is being generated. The fps shown here is always below the fps that was set on the camera. For example when scanning on 800x600 with 30 fps, the software shows a frame rate drop to around 20 fps. But the DAVID-software has a function called: “reduced display frequency”. David will update the scan results window only once per second. Depending on the hardware, this can increase the scanning speed. In other words this should compensate for this frame rate drop, but it still has to be confirmed. For now we assume that it does compensate for the loss in frames and that our images aren’t compromised.

Software

Two software packages are used in the creation of the setup. Matlab is used to control the xyz-table, to analyse the results and to control the software of the DAVID laser scanner. Next is described how the DAVID software is controlled by Matlab and then the communication from Matlab to the xyz-table is explained. How the results are analysed is described in the experiments and will not be described here.

David-laser scanner software

David-laser scanner software is controllable through communication ports. It has an option in advanced settings to listen to a predetermined communication port. So the desktop has been equipped with an internal PCI-card with two interconnected com-ports. This has been done to let Matlab communicate with the DAVID-software. It should also be possible to do this virtual, but it was easier to do by hardware.

Matlab is now able to send commands to the david software. But only to a certain extend. The calibration step of the camera still has to be manually completed or skipped. Skipping is only possible if the calibration object and camera have not been disturbed. This is not recommended, since the calibration has a major influence on the scan.

Matlab: Communication to xyz-table

A graphical user interface (GUI) was created, while one of the goals was to create a scanner that is operable by any member of the faculty group. Figure 11 shows the GUI.

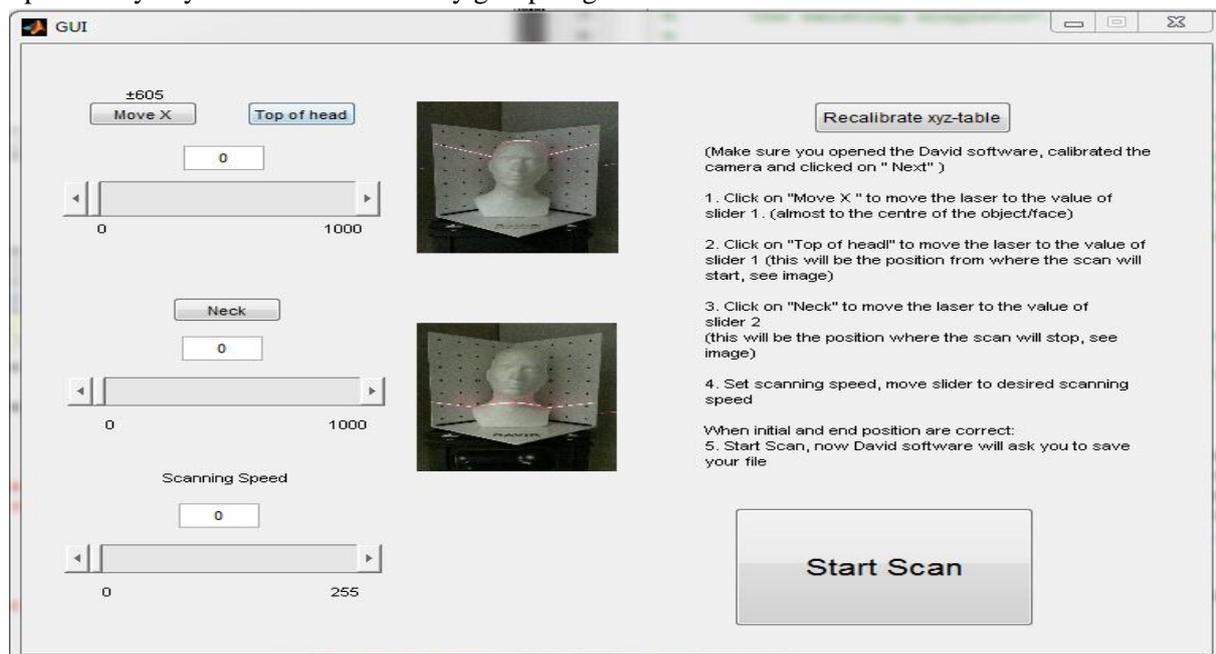


Figure 11: Graphical User Interface, designed to control the xyz-table with ease and start a scan automatically

The button “Move X” is used to position the laser approximately in front of the scanning object. This is done by setting the top slider to a certain value, in our setup approximately 605, and pressing “Move X”. The “Top of head” and “Neck” buttons are used to move the laser line to the position from where the scan should respectively start and end. For the “Top of head” button the value in the top slider determines its position. For the “Neck” button this is done by the value in the middle slider.

“Recalibrate xyz-table” is used when there’s a problem with the table, but in practice it’s not used. The slider of “scanning speed” will determine how quickly the laser is moved from top to bottom.

“Start Scan” will send a command to the DAVID software to start scanning and will send a command to the XYZ-table to move the laser from the position set with “Top of head” to the position set with “Neck”. When the scan is done the DAVID software will be instructed by the GUI to save the scan. The name of the scan and destination will have to be put in manually. The design criteria “the setup should be user friendly” is met.

Laser

The DAVID laser scanner has a few different lasers available in the store. The main differences are the colour, the power and power supply. On the discussion forum of DAVID it was explained that a green laser should theoretically give better results. Since the camera has 2 green receptors to every blue and red receptors. So it would be more capable to detect green lasers. We still used a red one since it was complementary to the starter kit

In practice a laser with an external power supply is more convenient. Since that power supply could be computer controlled, which can save energy when the laser is automatically turned off when not used. Also the switch on the laser used in the setup is sturdy and when pushing it, it is possible to move the laser or change the intersection angle unintentionally. The effect is not very big, but it might disturb the focus of the laser and causes the laser line to widen. This should have a negative effect on the quality of the scans. The thinner the line, the fewer pixels in the image map to the same position on the object.

Laser angle

The laser is mounted under an angle that is sufficiently high for the DAVID software to recognise it. It’s suggested in the manual and on the forum of the laser scanner that the intersection angle should be as high as possible to ensure accurate measurements. In chapter 5 “Setup setting determination” it is shown that there is a limit. Figures 14-16 display the extreme angles and one angle in between these extremes.

Acquisition process

The laser is mounted on the XYZ-table in the way shown in figure 15, but should not be switched on. The DAVID laser scanner software should be started first, then the GUI in matlab. By starting the GUI the XYZ-table will automatically calibrate itself. When this is completed, the camera needs to be calibrated through the DAVID software. How this is done, is describe in chapter 3 at “the camera calibration”. The settings for a good calibration depend strongly on the surrounding illumination. So no set of settings is given, instead an image is provided as an example of how the calibration image should look like. The settings need to be adjusted until the image resembles the example.

After calibration the object is put in front of the background and the laser is switched on. In chapter 3 at 3D laser scanning it is shown how to set the camera to get good quality scans. In chapter 5 the settings used for our experiments are determined. Earlier in this chapter is shown how the GUI is used to acquire scans. When the first scan is done, only the GUI is needed to make the next scan. Except the name and destination of the scans have to be put into the DAVID software. The scans are saved without doing any form of post processing, even though DAVID is able to do this.

conclusion

Our goal to create a mechanical structure to control the DAVID laser scanner is achieved. Also all the design criteria stated at the beginning of this chapter were met.

5. Setup setting determination

There are many factors that can influence the quality of a scan. We are not able to examine every factor independently due to a time limit. However some factors can be evaluated based on assumptions made about the quality of a scan. We assume the fewer gaps appear in a scan the better the quality of the scan. Also the less background noise there is, the better the scan. Based on these assumptions a set of camera settings is determined and the influence of certain factors is determined. The factors examined are surrounding light and intersection angle.

Moving speed of the laser

As mentioned above, the goal is to do a full facial scan within 30 seconds. This means the laser has to move across the entire face (at least once) within the time limit. This is done by setting the speed of the xyz-table to a high enough value. The value was determined to be 30. The higher the value gets, the faster the laser will move.

The resolution, exposure and frames per second settings have to be adapted to the speed. The exposure setting needs to be fast enough to capture a laser line that is as little stretched as possible and the frames per second will have to be high enough to capture the movement of the laser. The FPS and resolution are directly connected. At 1600x1200 pixels the maximum fps is 10 and at 800x600 pixels the fps is 30.

As a preliminary experiment we tried to find the trade-off between speed and quality. Here the quality of a scan was determined by looking at the gaps present in the scans. See figure 12 for an example of a good scan and see figure 13 for a bad scan. In chapter 3 it is explained how to interpret the images. We found that for our purposes the setting of 800x600 pixels at 30 fps was the better choice. As explained in the previous chapter, we only tested 1600x1200 and 800x600, since they have an even scaling factor.

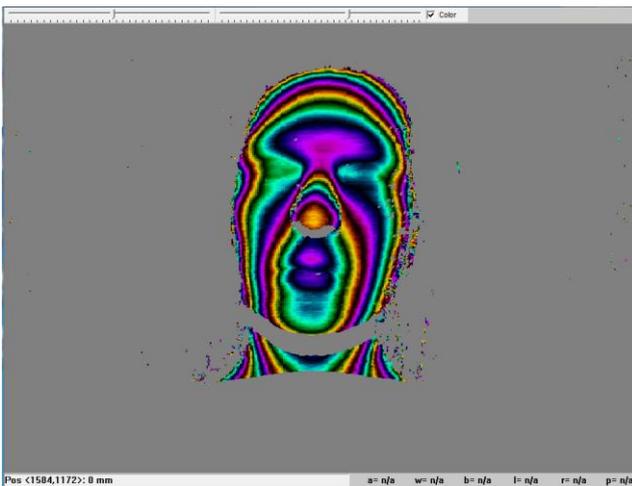


Figure 12: Good scan result

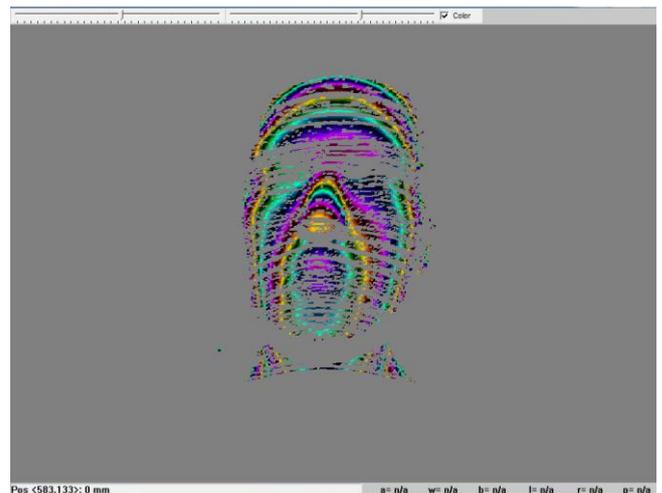


Figure 13: Bad scan result

Intersection angle

The intersection angle is the angle made by the laser with the image plane. The influence of the angle is determined by evaluating the scans with the assumption: “the fewer gaps appear in a scan the better the quality of the scan”. In the manual of the laser scanner was stated: the bigger the intersection angle the better the results.

We tested this statement by scanning a bust with the laser positioned in different angles. Figures 14 and 16 show the most extreme angles. After comparing the results, in the same fashion as was done in the preliminary speed experiment, we conclude that the angle shown in figure 15 is sufficient to get good results.

No Noise settings

The webcam has 8 different parameters that can be set. Brightness, contrast, saturation, sharpness, white balance, focus, exposure and gain. As mentioned before, the scans need to have as few gaps as possible and preferably no background noise at all. Figures 17 and 18 display results with and without background noise. The images show a point cloud of a sphere with radius 10.9mm that was scanned. In figure 17 only points on the sphere were detected. In figure 18 there is also a lot of noise visible. The dense concentration of points in the middle is the sphere, but for an overview of the noise the image is zoomed out compared to figure 17.

We believe the noise occurs due to reflection of the laser line. If the wrong settings are used, these reflections are detected.

Again the settings were found by means of trial and error. The following settings were found to be adequate for our scans:

- brightness: 0
- contrast: 7000
- saturation: 1000
- sharpness: 0
- white balance: 0
- focus: 100
- exposure: 1/200 [s]
- gain: 4500

The settings show a high contrast setting and a low brightness setting. The DAVID manual[7] and the DAVID wiki suggest that this is needed for good scans. Brightness of the image can also be altered by exposure and gain. Contrast is set very high to ensure a good laser detection.

The camera settings and laser angle determined in this chapter are used for the scans in the following experiments. But first is explained which experiments are done and why.

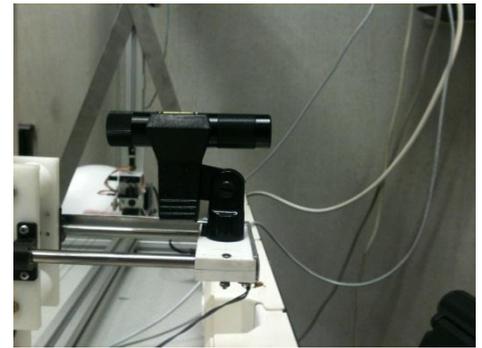


Figure 14: lowest possible angle

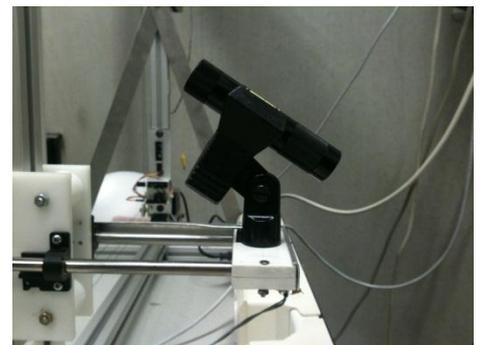


Figure 15: Angle chosen

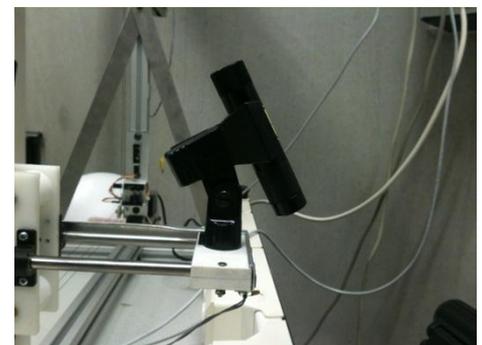


Figure 16: highest possible angle

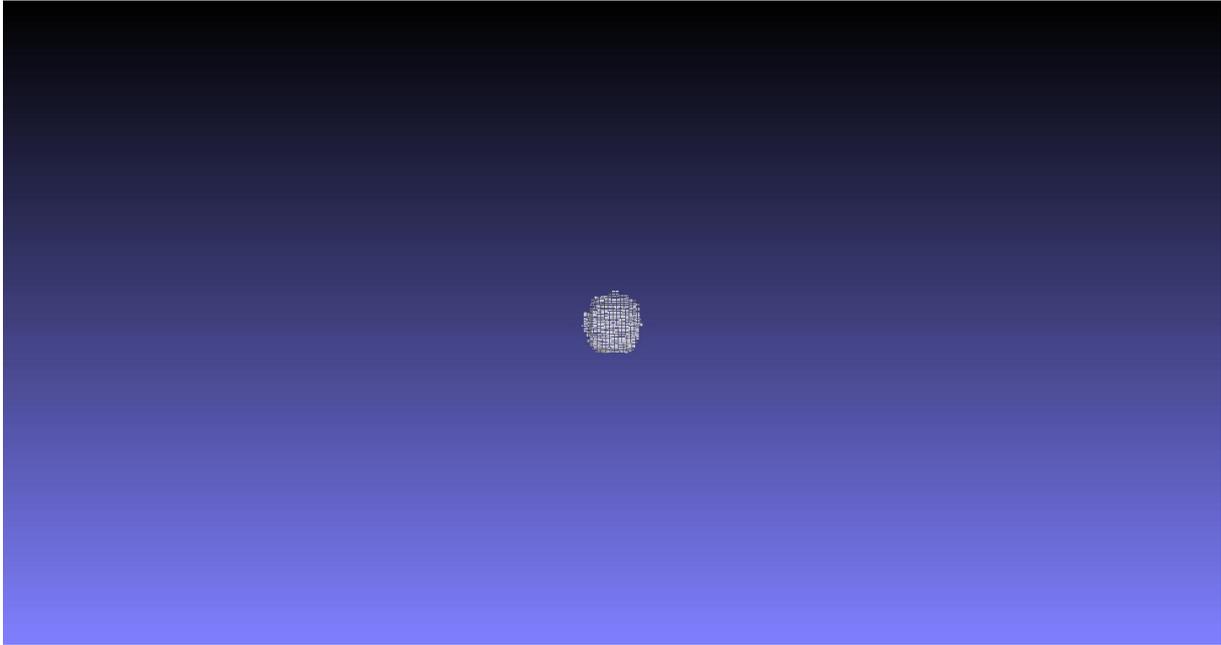


Figure 17: result without background noise

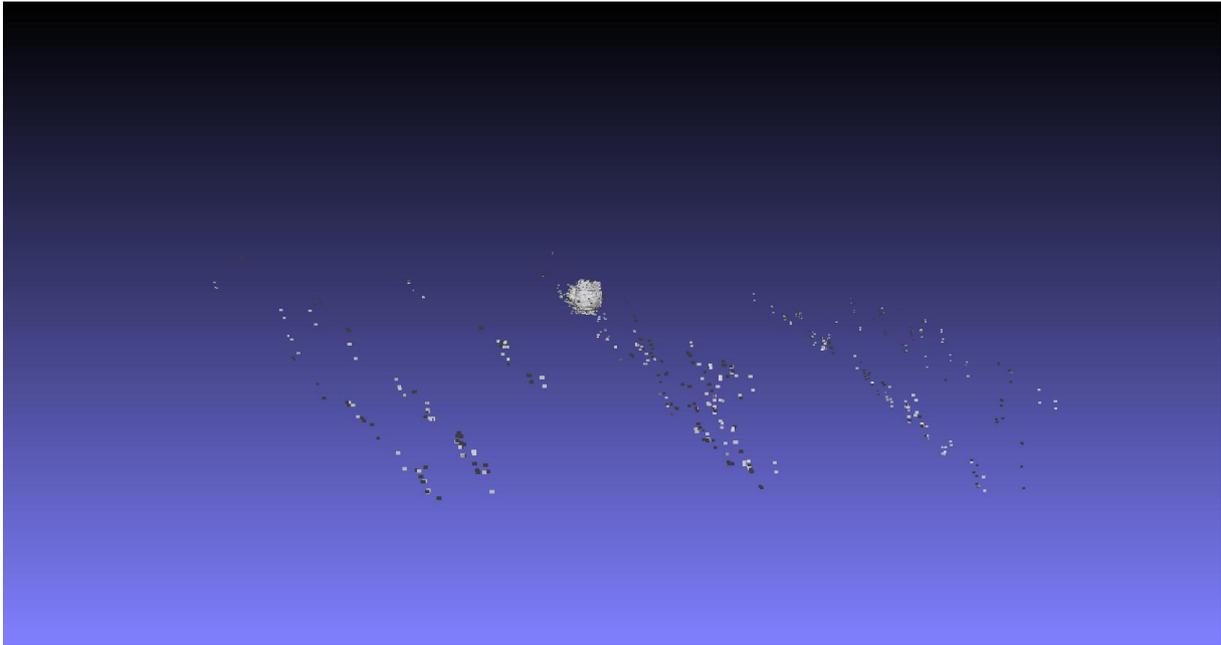


Figure 18: result with background noise

6. Evaluation of the laser scanner

At the introduction we have stated the following research questions that apply to the evaluation of the scanner:

- What is the quality of the 3D scans?
- What is the accuracy of the 3D scan?

To answer these questions, we first define the accuracy and quality of a 3d scan in measurable quantities. In our case we have chosen to use the absolute accuracy and point density, since these parameters are directly related to the quality of a 3d scan. The point density is approximately the resolution. This can be determined in different directions. With the absolute accuracy we are able to determine the smallest observable feature that is measured. Also the error in different parts of the scan will be analysed, to determine if certain parts of the image are more accurately acquisitioned.

Our goal is to get an absolute accuracy in the order of 1mm in every direction, since the scanner is going to be used for face acquisition. Every direction means in x, y and z direction. The x and y directions are directly linked to the resolution of the camera. So it's more convenient to use the point density to determine the quality of the scan in these directions. Our goal is to distinguish features of 1mm or smaller, so the density has to be 1 pixel/mm or higher. We define that 95% of all the 3D reconstructed points should have this accuracy for it to be a good scan, since this means that two times the standard deviation in the distance error is 1mm.

We have done three experiments to determine if this accuracy was achieved. The first experiment was determining the distance and its error from an object to the calibration screen. From this we determine the absolute accuracy in z direction. The second experiment is determining the point density in x and y direction. This is calculated from one of the scans of the first experiment. The third experiment is done to determine if the error in accuracy is constant over the entire scanning space. Here the scanning space is the space in which an object can be placed and the 3d scanner is able to acquire an image of it.

First the experiments will be explained in further details, then the results will be shown and discussed.

Experiment 1: Distance measurement

This experiment is for determining the distance and its error from an object to the calibration screen. Our goal is to get an absolute accuracy in the order of 1mm in every direction, since the scanner is going to be used for face acquisition.

For this experiment, the ball from an old computer mouse was used as scanning object. The assumption is made that it's "perfectly" round shaped. Why a sphere was used is explained later on. The object has to be positioned at different positions in depth. A heavy rail with a ruler on the side is used along which an object can slide. The ball was placed on top of the sliding object, so the ball could be moved in depth. The experiment is sketched in figure 19 and the real setup is shown in figure 20. Also the coordinate system used is shown in figure 20.

In figure 19 is shown that the scan object is a mounted sphere and it is moved towards or moved away from the camera along the rails. Steps of approximately 1 mm can be made, since the ruler used has a scale with the accuracy of a millimetre. The reading error of the ruler is approximately 0.4mm. From the data points collected in the scan, the distance to the background (d , in figure 11) can be calculated in multiple ways.



Figure 19: sketch of the experiment.

One could determine for every single data point the distance to the background. The distance to the background is chosen as the distance to y-axis displayed in figure 20.

The points are saved in xyz-coordinates with the origin in the middle of the calibration background. Because the ruler is positioned precisely on the 45 degree line, the distance from a data point to the background is $d = \sqrt{x^2 + z^2}$.

That distance will be different for every reconstructed 3D point since the centre of the sphere is not in the origin. Even if it would have been in the origin, this method is not accurate. If you think of a sphere as an infinitely stacked pack of circles with

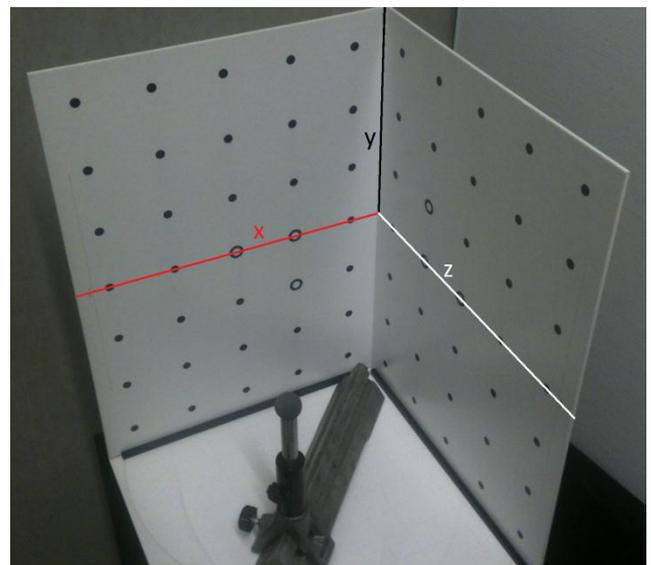


Figure 20: definition of the coordinate system

different radii and the circles are stacked in the direction along the y-axis. One can now easily see that the distance to the y-axis is different for every circle. So this method is not accurate enough and is dismissed.

Another method is always use the same point on the sphere and determine its distance to the y-axis. At first instance the point farthest away from the y-axis was chosen. But determining which point this was, appeared to be difficult.

Primarily noise was a problem, since we determine the distance to the y-axis from the x- and z-coordinate. In figure 21 an example is given. The left image is taken alongside the z-axis looking perpendicular to the x-plane, in the coordinate system defined by figure 20. The right image is taken directly from the front, from where the camera is positioned.

The green circle shows a data point that can be treated as an outlier and should not form a problem. The red dot represents the problem and the blue dot represents the probably correct point. The blue one is more likely to be the correct point, since it is almost in the middle of the scan. But the distance of the red dot to the y-axis is larger and is therefore incorrectly chosen.

Another problem is that it's possible that the point farthest away from the y-axis is not always the same one, since only a limited set of data is created. Not every point on the sphere is captured in the scan, and therefore small differences in the distance calculation might occur.

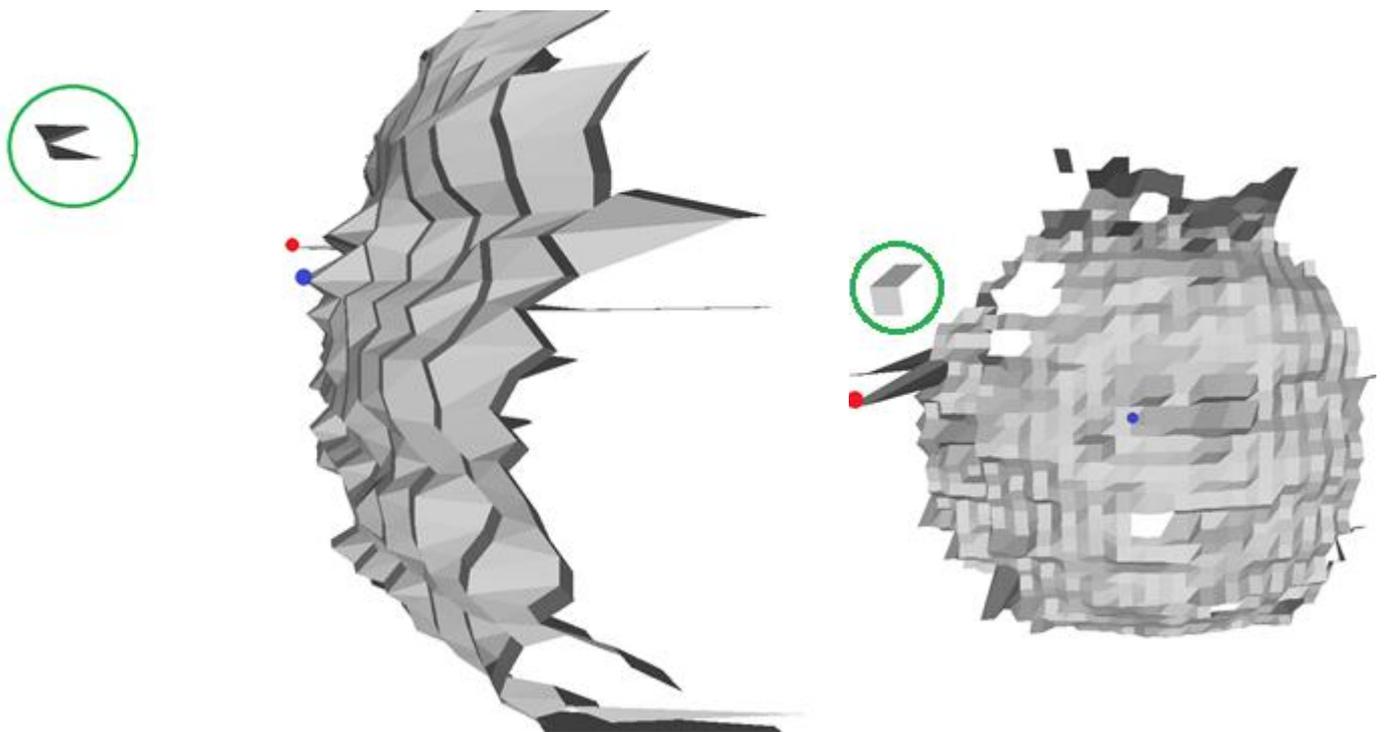


Figure 21: a mesh representation of example scan of a sphere

Another point in the sphere that should always be the same is the centre of the sphere. The coordinates of the centre can be determined from the entire collection of data points. This way, assuming there is only a small number of outliers, the outliers don't have a huge influence on the calculation of the centre. The calculation can be done in multiple different ways, we chose for a least squares fit and determine its centre. The fitting can be done algebraically or geometrically.[4][5] Both would try to minimize the following sum:

$$\sum_{i=1}^m e_i^2 = \sum_{i=1}^m [(x_i - x_c)^2 + (y_i - y_c)^2 + (z_i - z_c)^2 - r^2]^2$$

x_i , y_i and z_i are the data and the x_c , y_c and z_c are the sphere's centre coordinates and r is the radius.

The advantage of doing it geometrically is that a function was already provided for matlab. The disadvantage was that the radius could not be fixed. The function would estimate the radius of the sphere that was being fitted. An error in that estimation would also have an influence in the estimation of the sphere's centre coordinates. There are also several other known disadvantages of algebraic fitting, these can be found in reference 4. So this method is discarded and geometric fitting is used for the least squares fit. This means the sum stated above is minimized by iteration with a fixed radius of 10.9mm. The diameter of the mouse ball has been measured using a digital slide rule.

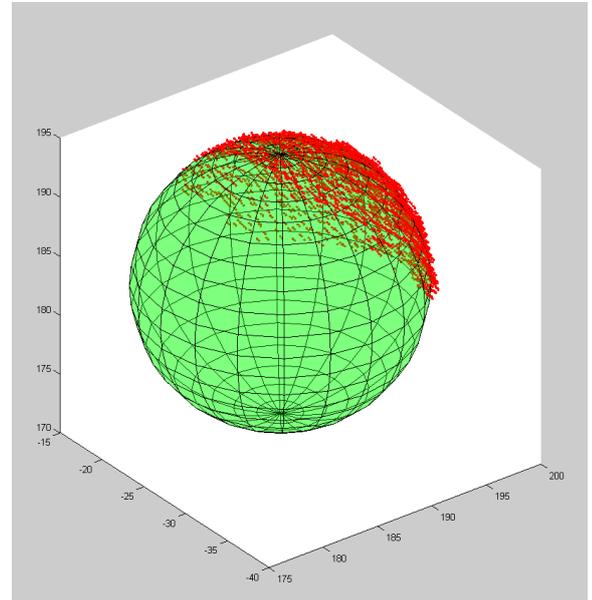


Figure 22: Least sphere fit

Figure 22 shows an example of the least squares fit of a sphere to the data. The red dots are the dataset acquired from the scanner and the sphere is the calculated best fit. From the calculated coordinates of the spheres centre the distance to the y-axis is determined.

The dataset is fully used in the fitting process. It is assumed no outliers are present in the scans. This assumption can be made as is made clear in the preliminary experiments .

Results

The sphere was approximately positioned at 26cm from the background. The sphere was moved to approximately 25cm in 10 steps. The results of the distances calculated using the sphere fitting process are shown in table 1. The reference distances are generated by setting one reference to the same value as the measurement and then adding or subtracting multiples of 1mm. This represents the steps that were made between every different scan.

The results in table 1 are from scanning the sphere at 800x600pixels with 30 frames per second at a speed of 30. This setting was used, since it represents the scanning condition that is needed when scanning faces. In chapter 5 is explained why we used these settings.

Table 1: results of distance measurement experiment, with a fixed radius of 10.9 mm

Reference distances (mm)	3d measurement (mm)	Δ (mm)
261.1795	261.1905	-0.0110
260.1795	260.1875	-0.0081
259.1795	259.2749	-0.0954
258.1795	258.3500	-0.1705
257.1795	257.2596	-0.0802
256.1795	256.1795	0.0000
255.1795	255.3107	-0.1312
254.1795	254.2613	-0.0819
253.1795	253.2424	-0.0629
252.1795	252.3926	-0.2131
251.1795	251.5481	-0.3687

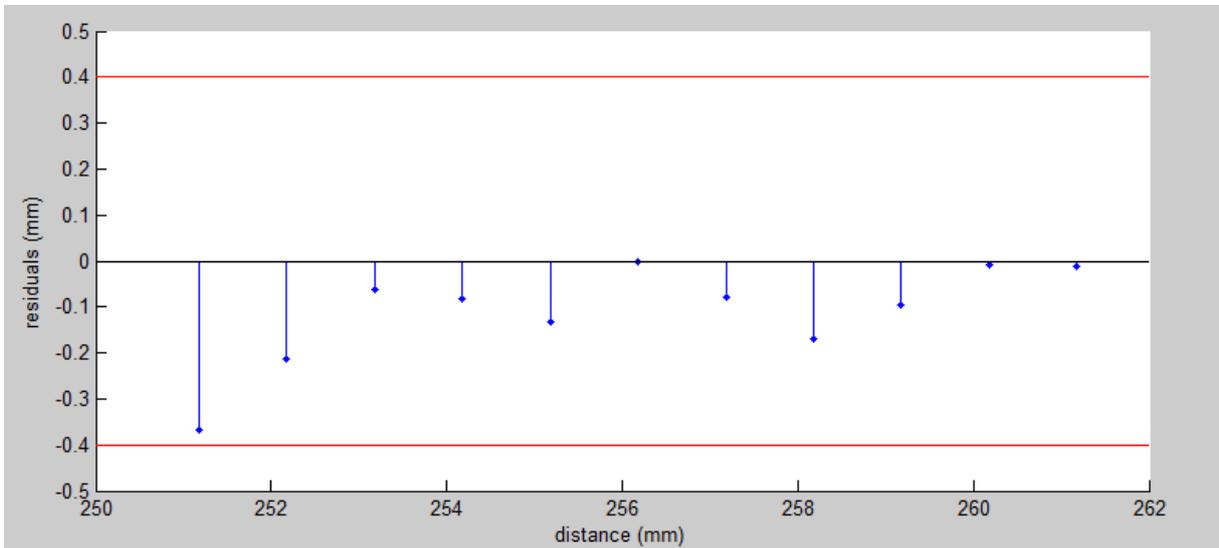


Figure 23: Residuals, Δ . Red lines are the reading error.

The delta depicts the error in the measurements, since it's *reference distance – measurement*. The error doesn't exceed the 0.4mm reading error that is present in the reference model. This is seen in Figure 23, since it shows the results of table 1 with the reading error. The error seems random. However all the errors are negative, this could be a coincidence but it's probably a systematic error. The error doesn't seem to depend on the distance, since the value is fluctuating randomly.

Even though the exact absolute accuracy is not determined, it can be said that it's smaller than 0.4mm. This is the absolute accuracy in the depth of view and our goal was this to be smaller than 1mm. Which means for face acquisition the scanner's absolute accuracy is high enough.

Experiment 2: point density

The experiment is for determining the point density in x and y direction. The point density is approximately the resolution and this will be used to answer the question: “What is the quality of the 3D scans”. The goal is to have a point density that is higher than 1 pixel/mm. We first estimated the point density by means of a calculation and then measured the point density.

The x- and y-axis are the width and height respectively on the image taken with the camera. The camera resolution is set on 800x600pixels, which gives 800 pixels in x direction and 600 pixels in y direction. The camera is set in such a way that the calibration background almost fills the entire image. Approximately 100 pixels were not used to display the y-axis of the background. So approximately 500 vertical pixels are used to display the y-axis.

The worst possible density is then found at the y-axis of the background, since this is the furthest away from the camera. The background is printed on A2 format and is placed sideways. So the length of y-axis, which is the height of the background, is the width of an A2 paper. Therefore the density is $500/420 \approx 1.19$ pixels/mm. But since the object is put in front of the background, we now approximate the best possible density for an object that is located at the middle of the baseplate and is parallel to the image plane. Figure 24 shows that the object is located at 210mm from the background. Figure 25 shows an approximation for the maximum height of that object. Here there is assumed the entire object is displayed in the image.

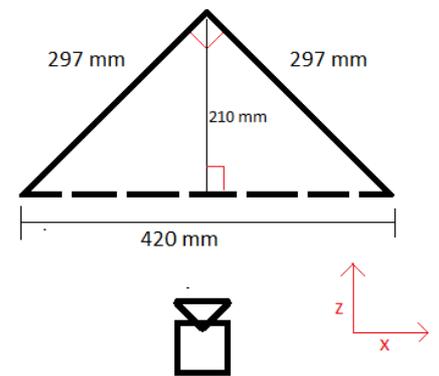


Figure 24: sketch of topview of the calibration background

The red distance shown in figure 25 is calculated in the following way:

$$\alpha = \sin^{-1}\left(\frac{210}{500}\right) \approx 0.4334 \text{ rad}$$

$$2 \cdot \sin(\alpha) \cdot (500 - 210) \approx 243.6 \text{ mm}$$

Since this is the maximum height of the object that is put in front of the background at 210mm from the y-axis, the density will approximately be: $500/243.6 \approx 2.05$ pixels/mm.

So for the y direction the goal will probably be achieved, since the density is in the worst case 1.19 pixels/mm and for an object the density will be around 2.05 pixels/mm. The calculations also show that the closer an object is to the camera the higher the resolution becomes.

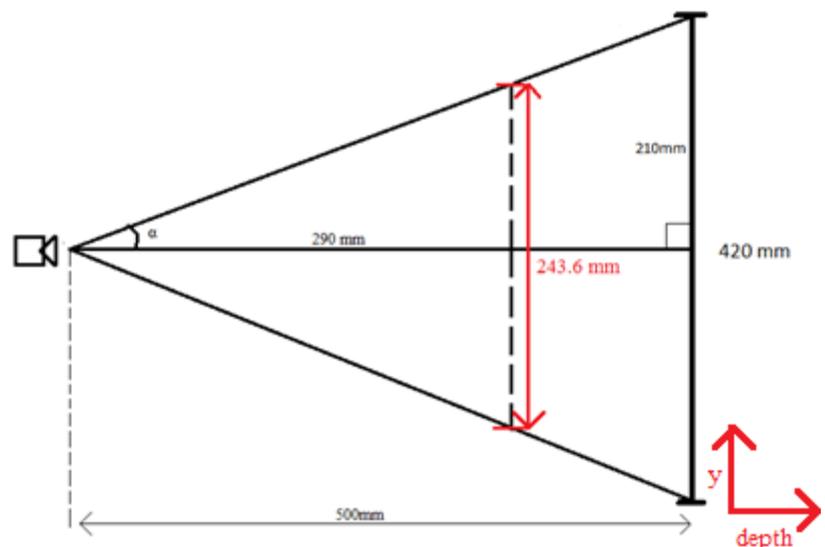


Figure 25: a sketch of the sideview of the calibration background

The density calculation in the x-direction is explained with figure 24.

Notice the coordinate system is defined differently compared to the coordinate system defined in figure 20 and thus different from the coordinate system in the scans.

The object/face will be located at approximately 210mm from the background and at this distance over 800pixels approximately 420mm are displayed. Which gives a density of $800/420 \approx 1.90$ pixels/mm.

Rapidform Explorer[8] is used for determining the point density. The distance between two neighbouring points is measured alongside the different axes. So the distance is measured in x, y and z direction. This will give an indication of the resolution in the different directions. However the coordinate system in the scans are different than the system defined in figure 24. The coordinate system of the scans is displayed in figure 20. Therefore the density in the x-direction as defined in figure 24, is not directly found in the results. However the density in y-direction is.

Results

For calculation of the point density a randomly picked scan from the absolute accuracy experiment was used. This means the scan was made of a sphere with a radius of 10.9mm, positioned at approximately 25.5cm from the background. The camera resolution was set on 800x600pixels with a frame rate of 30 and the laser speed was set to 30.

In figure 2 it is shown that several distances are measured across the entire dataset. The points were chosen at random. Figure 26 shows the data points with the measurements. Keep in mind the coordinate system used in the results is the system defined in figure 20 not the coordinate system shown in figure 26.

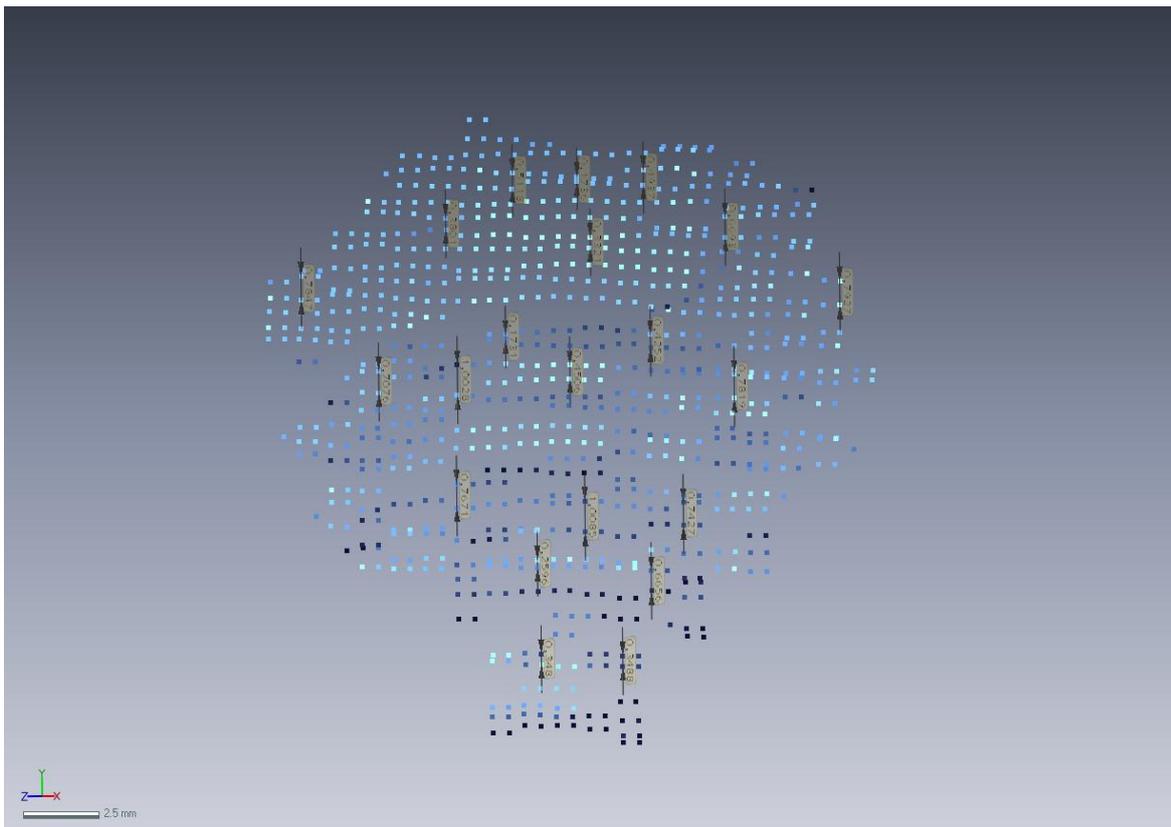


Figure 26: dataset loaded in rapidform explorer[8] with distances between neighbouring points. (the x and z are switched in the coordinate system.)

The example shows the distances aligned with the y-axis. The density was calculated from the average distance. Two neighbouring points have a certain separation and by choosing a few different couples

an average distance can be found. From the average distance, the density is calculated by: $\frac{1}{\bar{a}} = \text{density}$. The same was done in x and z direction and table 2 shows the results.

Table 2: Results of point density calculation

	Number of points used	Average distance (mm)	Standard deviation (mm)	Density (points/mm)
x-axis alignment	32	0.3821	0.0755	2.6173
y-axis alignment	21	0.5628	0.2522	1.7768
z-axis alignment	38	0.3576	0.2779	2.7961

The density in y-direction is found to be 1.7768 points/mm. Our estimated density was in worst case 1.19 pixels/mm. If the object was put in front of the background at 210mm the density was approximately 2.05pixels/mm. The measured value is in between the two calculated values. The object used was a sphere and thus is curved towards the background. The density becomes lower if moved towards the background, so this explains why the measured value is lower than the calculated value for an object that is parallel to the image plane placed at 210mm from the background.

The estimated density in x-direction cannot directly be compared to the calculated densities. However since these densities are far larger than the estimated density, it can be said that it's very likely that our goal is achieved.

Experiment 3: error consistency check

The third experiment is done to determine if the error in the absolute accuracy is constant over the entire scanning space. Here the scanning space is the space in which an object can be placed and the 3d scanner is able to acquire an image of it. Our goal is that 95% of all the 3D reconstructed points should have an accuracy of 1mm.

For the error consistency check a sphere of much larger proportions was used. The diameter is 30cm what is representative for a face. The object is scanned with the camera set to 800x600 pixels with 30 frames per second and the laser moving speed was set to 30. As explained before these settings are realistic for scanning faces. This is emphasized since one could get better scan results by slowing down the laser movement and increasing the resolution of the camera, but then a person should be able to sit still for more than 30 seconds. Which is uncomfortable and difficult to realise.

Figure 27 shows the sphere that was used. Since it is 30cm in diameter, half a sphere is used and it is made from Styrofoam. The resulting data is processed in a similar manner as done in experiment 1. The centre of the sphere is again calculated with a geometric fit, with a known radius of 15.0cm. If the centre is known, the distance to each data point is calculated and the radius is subtracted from this:

$$d = \sqrt{(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2} - r^2.$$

This is the distance from a data point to the sphere, which is a measure for the error. Then the average error, standard deviation, variance, maximum absolute error and minimum absolute error can be calculated.



Figure 27: Scanning object for experiment 3

Results

Two scans under the same conditions have been performed to determine if there's a pattern in the error. Also the data points have been assigned into quadrants to determine if the error is location dependent. The quadrants are formed in such a way that every quadrant has exactly the same number of data points. This is to be seen in figure 28.

The quadrants were fitted locally as well as globally. Locally fitted means that the centre of the sphere is found by only using the data points from the quadrant that is being reviewed. Globally fitted means that the centre of the sphere is found by using the entire dataset. The error is the distance between a data point and the reference sphere, and this is calculated by :

$d = \sqrt{(x - x_c)^2 + (y - y_c)^2 + (z - z_c)^2 - r^2}$. So the centre coordinates have an influence on the error. The results are shown in tables 3 and 4.

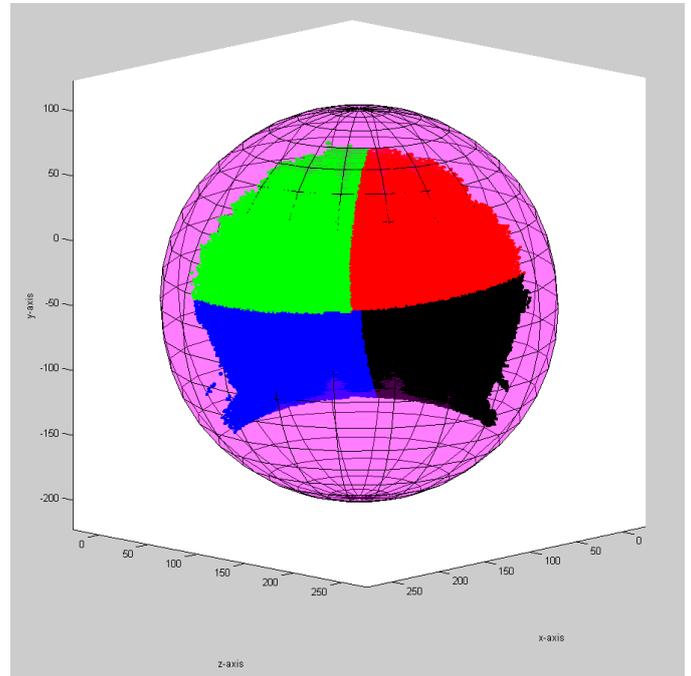


Figure 28: The four quadrants, green = 1, red = 2, black = 3 and blue =4

Table 3: results of the first scan

Scan 1	Avg. Error (mm)	Std. dev. (mm)	Max. Error (mm)	Min. Error (mm)
Total sphere	-0.0020	0.9371	5.2313	0.0000
Quadrant 1 local fit	0.0008	0.3073	1.3580	0.0000
Quadrant 1 global fit	0.0368	0.7077	2.1644	0.0001
Quadrant 2 local fit	0.0037	0.3708	1.6607	0.0000
Quadrant 2 global fit	0.0137	0.6616	2.0848	0.0000
Quadrant 3 local fit	-0.0015	0.6908	3.6745	0.0000
Quadrant 3 global fit	0.0331	1.0944	5.1030	0.0001
Quadrant 4 local fit	-0.0017	0.6769	3.6110	0.0000
Quadrant 4 global fit	-0.0916	1.1687	5.2313	0.0000

Table 4: results of the second scan

Scan 2	Avg. Error (mm)	Std. dev. (mm)	Max. Error (mm)	Min. Error (mm)
Total sphere	-0.0004	0.8953	4.2123	0.0000
Quadrant 1 local fit	0.0007	0.3078	1.3754	0.0000
Quadrant 1 global fit	0.0514	0.7008	2.1106	0.0000
Quadrant 2 local fit	0.0032	0.3738	1.5985	0.0000
Quadrant 2 global fit	0.0223	0.6517	2.2113	0.0000
Quadrant 3 local fit	-0.0006	0.6314	3.3427	0.0001
Quadrant 3 global fit	0.0290	1.0387	4.1579	0.0000
Quadrant 4 local fit	0.0000	0.6092	3.4855	0.0000
Quadrant 4 global fit	-0.1044	1.0938	4.2123	0.0000

The average errors of the local fits and the total sphere are very small, this is to be expected since this error is minimized in the fitting.

The standard deviation of the global fits as well as the local fits show that the spread of the error in quadrants 1 and 2 are smaller than the spread of the error in quadrants 3 and 4. Also the maximum absolute error is bigger in quadrants 3 and 4 in comparison to quadrants 1 and 2. So the upper hemisphere is scanned with less noise and therefore is scanned more accurately.

Our goal is to get an absolute accuracy in the order of 1mm. The standard deviations calculated over the total sphere are around 0.9mm. This means that approximately 68% of the points acquired are scanned with an accuracy of almost 1mm. Our goal was not achieved.

It was explained in chapter 3 that the post processing in the DAVID is available. DAVID can interpolate, smooth median and smooth average. These options will generate extra data points from the original captured set and smooth out outliers. When done correctly, this will lower the standard deviation. How much influence it will have is not determined, but is recommended for future work.

6. Conclusion

The DAVID laser scanner is a cheap alternative to produce 3 dimensional scans with reasonable results for facial acquisition. The DAVID laser scanner starter kit consists of a webcam, laser line projector, DAVID software and calibration backgrounds.

The following research questions and goal were stated:

- Create a mechanical structure to control the DAVID laser scanner.
- What is the quality of the 3D scans?
- What is the accuracy of the 3D scans?

Mechanical structure

The research goal was achieved since a mechanical structure was created. A XYZ-table is used to move the laser line projector. The XYZ-table is controlled via a GUI in matlab that also controls part of the DAVID software. An object or person is placed in front of the calibration background and scanned automatically with use of the GUI.

The setup had the following design criteria:

- the maximum duration of the scan should be around 30 seconds
- the setup should be user friendly
- the setup must be able to scan a human face
- a face should cover approximately 50% of the image
- the camera should not be attached to anything that is in contact with the xyz-table

All of the criteria were met except for “the setup must be able to scan a human face”. A person is not able to sit in front of the background.

Analysis of the DAVID laser scanner

The goal was an absolute accuracy in the order of 1mm in every direction. The x and y directions are x and y axes in the camera image and the z direction is the depth. So it's more convenient to use the point density to determine the quality of the scan in x and y. Our goal is to distinguish features of 1mm or smaller, so the density has to be 1 pixel/mm or higher. The goal of an absolute distance accuracy in depth is 1 mm. We also defined that 95% of all the 3D reconstructed points should have this accuracy for it to be a good scan, since this means that two times the standard deviation in the distance error is 1mm.

The scans made with the setup were analysed in three experiments. Experiment one was for determining the distance and its error from an object to the calibration screen. A sphere with a radius of 10.9mm was positioned at different distances and scanned. With a least squares fit the centre coordinates of the sphere were calculated and used to calculate the distance to the y-axis of the background. The results of the experiment show that the accuracy of the 3D scanner is below 0.4mm in the depth.

In experiment two the point density in x and y direction in the image plane was calculated and measured. It was measured by randomly measuring the aligned distance between nearest neighbouring points. $\frac{1}{d} = \text{density}$. The density in y-direction is found to be 1.78 points/mm. Our estimated density was in worst case 1.19 pixels/mm. If the object was put in front of the background at 210mm, the

density was approximately 2.05pixels/mm. For the x-direction in the image plane it was shown that it's very likely that the point density is higher than our goal.

The third experiment is done to determine if the error in the absolute accuracy is constant over the entire scanning space. A sphere with a radius of 15.0mm was scanned and with a least squares fit the centre of the sphere was calculated. The error is the distance between a data point and the reference sphere. The reference sphere is the sphere with the calculated centre coordinates and an exact radius of 15.0mm. The standard deviation calculated over the total sphere is approximately 0.9mm. This means that approximately 68% of the points acquired are scanned with an accuracy of almost 1mm. Since 95% was needed for a good scan, we concluded that the scans are reasonable. Also because the image could still be enhanced by means of post processing the data.

We conclude that the quality of the 3D scans was high enough for our standards in 3D face acquisition. The accuracy of the 3D scanner is not good enough according to our criteria. But we recommend further work, since it has displayed the potential to give good results.

Future work

It was mentioned that post processing is available in the DAVID software, but not used. Since it should have a positive effect on the accuracy, we recommend that the influence of post processing is examined. It might be possible to enhance the images to even be good enough for facial acquisition according to our standards.

A better camera and slowing down the scanning speed could also increase the quality and accuracy of the results. We recommend these steps taken first, since they most likely have a great influence on the results.

Since facial acquisition was the intend, we recommend that a background is made were people can position themselves in front of. Also the entire setup is very large and takes up a lot of space. If a more portable setup could be created, then it would benefit the usability of the DAVID laser scanner.

7. References

- [1] S. Huq, B. Abidi, S.G. Kong and M. Abidi, A survey on 3d modelling of human faces for face recognition, The university of Tennessee, 2007 Springer
- [2] Mark F. Hansen, Gary A. Atkinson, Lydon N. Smith, Melvyn L. Smith, 3D face reconstructions from photometric stereo using near infrared and visible light, University of the west of England, 2010 Elsevier.
- [3] M. Verhoeve, XYZ-table, Utwente, 2009
- [4] Sung Joon Ahn, Wolfgang Rauh, Hans-Jürgen Warnecke, Least-squares orthogonal distances fitting of circle, sphere, ellipse, hyperbola and parabola, Fraunhofer Institute for Manufacturing Engineering and Automation (IPA), 2000, Elsevier
- [5] Christoph Witzgall, Geraldine S. Cheok, Anthony J. Kearsley, Recovering spheres from 3D Point Data, National Institute of Standard and Technology,
- [6] C. Boehnen and P. Flynn. Accuracy of 3D scanning technologies in a face scanning context. Proc. of 5th Int'l Conf. on 3D digital imaging and modelling, 2005
- [7] David Laser scanner manual
- [8] Rapidform Explorer, <http://www.rapidform.com>

Appendix A

Manual for making a 3D scan with the automated DAVID laser scanner

1. Switch on computer, make sure power is supplied to the XYZ-table and the usb-stick with the DAVID software is attached to the computer.
2. Start DAVID software
3. Start “run.exe” and wait until the XYZ-table is done calibrating
4. Calibrate the camera as described in chapter 3
5. Put object or person in front of the background
6. Set the camera settings as described in chapter 3
settings that worked for us were:
 - brightness: 0
 - contrast: 7000
 - saturation: 1000
 - sharpness: 0
 - white balance: 0
 - focus: 100
 - exposure: 1/200 [s]
 - gain: 4500
7. Use the GUI as explained in chapter 4 to start scanning.
 - set x-position
 - set top y-position
 - set bottom y-position
 - Start Scan
8. Save the scan