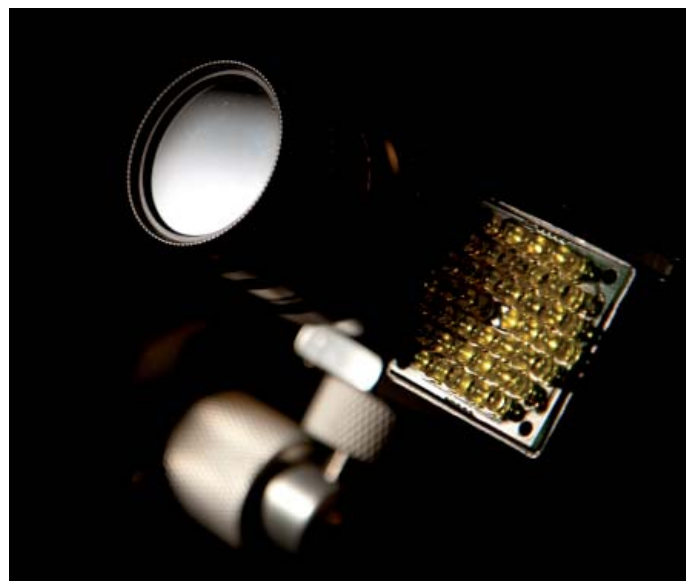


Performance of a one-camera and a three-camera system



Authors

Christer Ahlström, VTI

Tania Dukic, VTI

Erik Ivarsson, Smart Eye

Albert Kircher, VTI

Bosse Rydbeck, Smart Eye

Matias Viström, Saab Automobile



Virtual Prototyping and
Assessment by Simulation

www.vipsimulation.se

Preface

Head and Eye Behaviour Measurement and Visualisation in Simulators – VisualEyes is a collaborative project between Smart Eye AB, Saab Automobile AB, VTI, Pixcode AB and the Skaraborg Institute within the competence centre *Virtual Prototyping and Assessment by Simulation* (ViP). The project has three main goals:

- Evaluation of the influence of factors such as wearing glasses and participant's age on gaze tracking system performance for a one-camera and a three-camera system
- Development of a self-initialized visual attention detection module based on a one-camera system.
- Development of a real-time visualization system for gaze direction.

This report describes the first part of the project, the study comparing a one-camera system and a three-camera system. The study was conducted in the Saab driving simulator in 2009 and was financed by the ViP competence centre.

Participants from Saab Automobile were Arne Nåbo (project leader) and Matias Viström (test leader).

Participants from VTI were Tania Dukic (project leader at VTI), Christer Ahlström (data analysis) and Albert Kircher (experimental design).

Participants from Smart Eye were Martin Krantz (project leader at Smart Eye), Erik Ivarsson (test set up, data analysis), Bosse Rydbäck (data analysis) and Henrik Otto (data analysis).

Participant from Pixcode was Henrik Bergström (test set up, data collection).

Participants from the Skaraborg Institute were Hans Wedel, Ingela Krantz and Per Nordin (experimental design).

Göteborg, January 2010

Tania Dukic

Quality review

Peer review was performed on 2010-01-07 by Katja Kircher (VTI) and on 2010-02-02 by Marcus Nyström and Kenneth Holmqvist (Lund University Humanities Lab).

Christer Ahlström has made alterations to the final manuscript of the report. The ViP Director Lena Nilsson examined and approved the report for publication on 2010-03-15.

Table of contents

Executive summary	5
1 Introduction	7
1.1 Goals of the project.....	7
1.2 Hypotheses.....	7
1.3 Eye movement data.....	8
2 Method.....	10
2.1 Simulator and measurement setup	10
2.2 Requirements for participants to be selected for the study	11
2.3 Procedure and instructions to participants	11
2.4 Data collected	12
2.5 Study design	13
2.6 Expected validity of the study	13
2.7 Data analysis	13
2.7.1 Data collection, profiles, world models and data logging	13
2.7.2 Calibration and quality	14
2.7.3 Performance indicators	14
2.7.4 Statistical analysis	16
3 Results.....	17
3.1 Availability.....	17
3.2 Accuracy	20
3.3 Precision	21
3.4 Summary of results.....	27
4 Discussion	29
4.1 Methodology	29
4.1.1 Experimental design	29
4.1.2 Head-eye coordination.....	30
4.1.3 Data processing.....	31
4.1.4 Availability, accuracy and precision	32
4.2 Results.....	32
5 Guideline.....	34
5.1 Minimum requirements	34
5.1.1 Driver characteristics	34
5.1.2 Camera locations.....	35
5.1.3 Vehicle.....	35
5.1.4 Physical parameters	35
5.2 Single-camera systems.....	36
5.3 Multi-camera systems.....	36
References	38
Appendix 1: Information to participants and informed consent	
Appendix 2: Experiment protocol	

Performance of a one-camera and a three-camera system

By Christer Ahlström (VTI), Tania Dukic (VTI), Erik Ivarsson (Smart Eye), Albert Kircher (VTI), Bosse Rydbeck (Smart Eye) and Matias Viström (Saab)

Swedish National Road and Transport Research Institute (VTI)
Olaus Magnus väg 35
SE-581 95 Linköping
Sweden

Executive summary

Driving and operating a vehicle is to a great extent a visual task. In driver behaviour studies it is therefore important to be able to measure where the driver is looking. Today this can be done unobtrusively and remotely in real-time with camera based eye tracking. The most common remote eye tracking systems use multiple cameras in order to give satisfactory results. However, promising results using only one camera has recently emerged on the market. The main objective of this study is to compare eye tracking systems with one and three cameras, respectively, during various measurement conditions.

A total of 53 participants were enrolled in the study. Data from the two eye trackers were acquired and analysed in terms of availability, accuracy and precision. The results indicate that both availability and accuracy are affected by many different factors. The most important factors are the number of cameras that is used and the angular distance from straight ahead. In the central region (straight ahead) both one-camera and three-camera systems have a high degree of accuracy and availability, but with increasing distance from the central region, the results deteriorate. This effect falls harder upon the one-camera system. Interestingly, there were no significant effects when wearing glasses in either availability or accuracy. There was however an interaction effect between distance and glasses.

Advantages with a one-camera system are that it is cheaper, easier to operate and easier to install in a vehicle. A multi-camera system will, on the other hand, provide higher availability and accuracy for areas that are far from the road centre. A one-camera system is thus mostly suitable for in-vehicle applications such as systems that warn drivers for sleepiness or distraction while multi-camera solutions are preferable for research purposes.

1 Introduction

Driving and operating a vehicle is to a great extent a visual task (Wierwille, 1993). In driver behaviour studies, it is therefore very important to be able to measure eye movements (Coughlin, Reimer, & Mehler, 2009; Kircher, 2007; Lee, Young, & Regan, 2009). Traditionally this has been done by applying electrodes around the driver's eyes (electrooculography, EOG) or by using video recording combined with manual annotation of interesting events, but recently camera based systems have been developed that can monitor the driver's head and eye movements unobtrusively in real-time (Duchowski, 2007).

The most common camera systems use multiple cameras in order to give satisfactory signal quality. However, recent developments show promising results using only one camera (Smith, Shah, & Lobo, 2003). Multi-camera systems cover a larger head rotation envelope with a higher degree of accuracy, while one-camera systems are cheaper, easier to operate and easier to install in a vehicle. One-camera systems are today available from several eye tracking manufacturers such as Smart Eye (Smart Eye AB, Gothenburg, Sweden) and Seeing Machines (Seeing Machines, Canberra, Australia).

The motive for carrying out this project is to make head and eye gaze measurement systems more adapted and usable in simulator environments and to establish guidelines for when a particular system (one-camera or many-camera) is needed.

1.1 Goals of the project

The *Head and Eye Behaviour Measurement and Visualisation in Simulators – VisualEyes* project has three main goals:

1. Evaluation of the influence of factors such as wearing glasses and participant's age on gaze tracking system performance for a one-camera and a three-camera system
2. Development of a self-initialized visual attention detection module based on a one-camera system (classified).
3. Development of a real-time visualization system for gaze direction.

This report covers the first part of the project and will describe the performance of gaze tracking systems based on a one-camera system as compared to a three-camera system. A number of factors that are likely to affect the systems, such as glasses, age, gender, skin colour, wrinkle and make-up will also be investigated.

1.2 Hypotheses

The following hypotheses were defined for the present experiment to compare the one-camera system to the three-camera system. All hypotheses might not be possible to answer but they were formulated as relevant for the project. The terms availability and accuracy will be defined in Section 2.7. Dots are specific gaze targets that the participants look at during the measurements, see Section 2.1.

1. The one-camera system has a lower accuracy and availability compared to the three-camera system over the total test envelope (covering 180 degrees horizontally and 110 degrees vertically).
2. The three-camera system has higher accuracy for points outside the central area compared to the one-camera system. The central area is defined as the area spanned by nine dots located straight ahead in a direction where the driver is likely to aim his/her gaze for most of the time during naturalistic driving.
3. Accuracy (of the gaze and head direction towards the LEDs) will decrease with increasing distance from the centre points in both systems.
4. Presence of glasses on participants' faces lead to lower accuracy /availability compared to without glasses.
5. High degree of wrinkle is influencing the accuracy/availability: the higher the presence of wrinkles the lower the accuracy/availability.
6. Presence of make-up will deteriorate accuracy/availability compared to non make up.
7. The participant characteristics skin colour and unshaved does not influence accuracy and availability on either system.

1.3 Eye movement data

The eye movement data consist of a sequence of points that have two or three spatial dimensions and that are represented as $\langle x, y, z, t \rangle$. For eye tracking systems from Smart Eye, the x-value describes the movement in horizontal direction with negative values indicating eye movements to the right and positive values indicating eye movements to the left, because the coordinate system is seen through the "eyes" of the cameras. The y-value describes the movement in vertical direction, with negative values indicating downward movements and positive values indicating upward movements. The z-value is only available in some systems. The t-value indicates time, meaning that a set of $\langle x, y, z, t \rangle$ -values reflect the eye path. These time varying points are obtained for the head direction as well as for the gaze direction, and the directions are encoded as unit vectors. For this reason, these vectors are often referred to as head vectors and gaze vectors, respectively. Note that $\langle x, y, z \rangle$ are given as gaze direction angles and not as absolute coordinates (Antisleep User Manual; Smart Eye Pro User Manual).

The eye path, also called scan path, basically consists of fixations (pauses over informative regions of interest), saccades (rapid movements between fixations) and smooth pursuits (fixations on moving objects) (Salvucci & Goldberg, 2000). An illustration of fixations and saccades can be found in *Figure 1*.

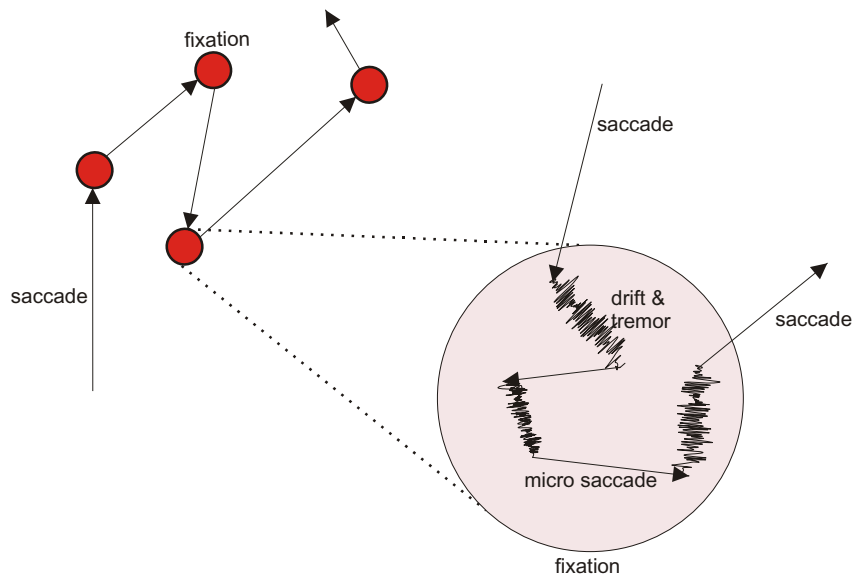


Figure 1 Illustration of fixations and saccades.

2 Method

The one-camera system that is used in this project is the Smart Eye AntiSleep system (SmartEye AB, Gothenburg, Sweden). The system uses a single standard camera of VGA resolution together with IR flash illuminators. The camera and the IR flashes are mounted in a compact unit but can be placed separately if required. The IR illuminators and filters are tuned to frequencies with minimum interference of outdoor light. This means that the system uses its own light, making it highly robust to all natural illumination conditions in automotive applications.

Smart Eye AntiSleep measures the driver's head position and orientation, gaze direction and eyelid opening at a rate of 60 Hz. All delivered measurements have confidence values based on the estimated quality of the measurements. The system detects generic and person-specific facial features and maps them to a generic 3D head model. The head model is then quickly adapted to the driver in real-time. This driver initialisation procedure is fully automatic and invisible to the driver.

The three-camera system that is used in this project is called Smart Eye Pro (SmartEye AB, Gothenburg, Sweden). Smart Eye Pro is a head and gaze tracking system well suited for the demanding environment of a vehicle cockpit. The system measures the participant's head pose and gaze direction in full 3D along with eye lid opening and pupil dilation. The system can be used with up to six cameras with different lenses, allowing for a very large field of view.

2.1 Simulator and measurement setup

The data collection took place in the General Motors Europe Driving Simulator at Saab in Trollhättan. A total of 41 dots at predetermined fixed locations were used as gaze targets (14 in the cockpit of the car and 27 outside the vehicle). The dots located in the cockpit were represented by LEDs while the dots on the outside were projected on the cylindrical projection screen. The dots located inside the car are illustrated in Figure 2.

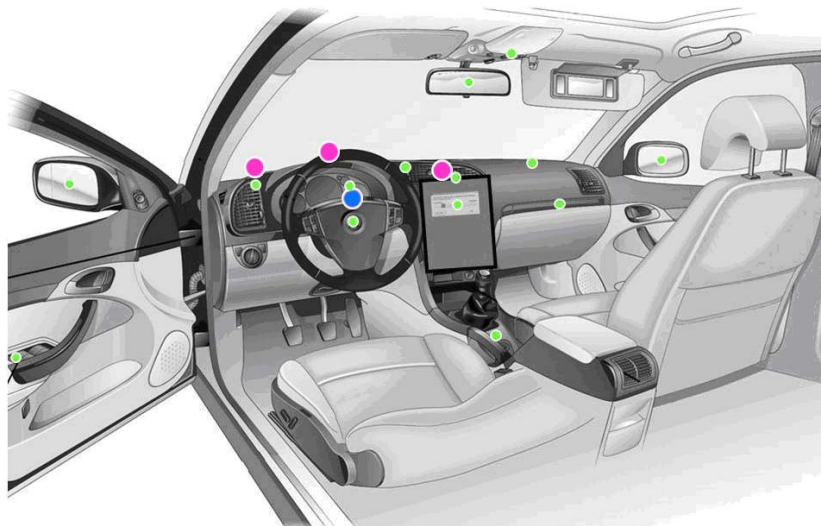


Figure 2 The layout of the dots inside the cockpit (green dots: LED; pink dots: camera locations for three-camera system; blue dot: camera location for one-camera system).

The centre point (CP) dot will always be activated before any of the other 40 dots in a randomized sequence according to CP dot–Dot1–CP dot–Dot2–CP dot–Dot3–... The dots are blinking to make it easier for the participant to locate the activated dot. When a blinking dot has been identified, the participant presses a button to indicate that he or she is looking at the dot. The dot lights up for 2 seconds until the next dot is activated. It is important that the participant don't look away while the dot is lit.

2.2 Requirements for participants to be selected for the study

Saab employees aged from 30 to 60 years, 50% males and 50% females, were included in the study. Participants over 60 years of age were excluded, since after this age there is a risk for cataract development without the participants having noticed it. The participants should have normal visual acuity without wearing glasses, however, people who use contact lenses or reading glasses were allowed to participate. The participants should also not wear piercings in the face (eyebrows, nose, and lips), however, earrings were allowed. Different hair lengths were allowed, but participants with long hair were asked to put their hair up so that it wouldn't cover their face. Also, the body length must lie between 155 and 193 cm to fit into the range of the camera.

The participants must have had a driving license for passenger cars for at least five years. Furthermore they must agree that data (video, pictures and hard coded data) will be collected and stored, and answer a number of questions before the driving sessions. They must as well agree to not discern the procedure during the trial to other potential participants, since knowing the procedure beforehand may influence gaze behaviour and readiness.

2.3 Procedure and instructions to participants

The participants were scheduled over the phone and were at the same time asked about the required characteristics; milage driven per year, length of their hair, if they were using glasses or lenses and if wearing piercings in the face. Participants with long hair were asked to bring e.g. a rubber band to prevent it from covering the ears. Those using either glasses or contact lenses were asked to wear contact lenses the day of the test.

Upon arrival to the simulator the participants were informed about the study whereafter they signed a consent form. Each participant was documented by two photos (one with and one without glasses) facing the camera, see Figure 3, and two photos from the right side. Further characteristics that were documented was the age and length of the participant, and whether the participant had beard, make-up or earrings.

Sitting in the simulator the participants were first given the opportunity to adjust the seat. The steering column was not allowed to be adjusted since the camera was mounted there. The participants were introduced to the camera setup and asked not to cover the cameras or their faces with their hands, and to not lean out of the car.

The following instruction was given to all participants:

“Look at the blinking dot, focus your gaze at it and press the “+”-button on the steering wheel. The dot will then shine, without blinking, for two seconds. Keep your gaze on the dot during these two seconds. When the light is put out another dot will start blinking whereafter you repeat the procedure. You will go through all dots twice, once with the glasses and once without the glasses. After each session you will be asked to hold a chessboard in front of your face in order to calibrate the eye tracking system”.

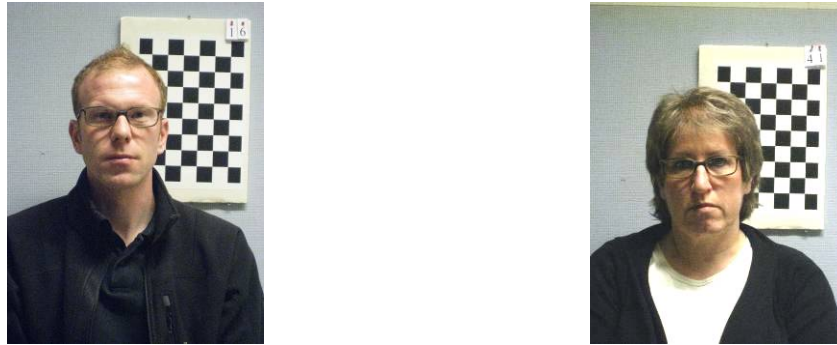


Figure 3 Male (left) and female (right) participants wearing the glasses used in the study.

Mobile phones were asked to be turned off before the session. At the beginning of the test, a procedure for calibrating the three-camera system was carried out. The participants were asked to first slowly sweep their head and gaze from left to right and then fix their gaze upon the left and right cameras respectively. This was repeated for the second run. The participants were left alone for the actual test procedure. When the test was finished a short unstructured debriefing was held, after which the participants left the simulator facilities.

No particular instructions were given about how to direct the head while looking at the blinking dots.

2.4 Data collected

The following data were collected via a questionnaire from 53 participants:

- Gender
- Age
- Length
- Facial hair (none, moustache, beard, heavy beard)
- Make-up (yes/no)
- Earrings (yes/no)
- Skin colour (pale white/white/brown/dark brown)
- Eyelid opening (small, medium, large)
- Number of years with a valid drivers license
- Driven mileage per year
- If the driver uses contact lenses or glasses
- If the driver had facial piercing
- Length of hair (Short or Long).

2.5 Study design

Data from the one-camera system and the three-camera system were recorded in parallel, and each participant performed the experiment with and without glasses (Figure 3) procured by Smart Eye. The glasses had rather heavy frames and were the same model for each participant. The order (with or without mock-up glasses) was randomized for participants. Also the sequence of the dots was randomized for all test persons.

The main factors of the study were one- vs. three-camera system and with/without glasses (within design - each driver drives in both conditions). Additional factors are mentioned above.

2.6 Expected validity of the study

Given the very diverse face characteristics of people, having a study which evaluates the influence of different facial features on the performance of the two systems is intrinsically complicated, and limiting the number of factors is also difficult. Because of the within-individual-design of the factor “use of fake glasses” it is expected that the test will give valid evidence as to the importance of this variable. The effect of age, gender and skin fairness will also be tested. The age distribution will be recorded by age in years but also in age groups, such as 30–39, 40–49, and 50–60 in order to analyse the three age groups as a factor. A similar approach can be used for the skin fairness. Note that each additional factor can cause the study results to be less reliable (less statistical power), because of the limited number of participants and because of dependent variations.

A number of factors that possibly affects system performance in real driving have been eliminated in the study. These are: sunglasses, hat, scarf and similar items, hair obscuring the face and ears, movements of the participant while driving, varying light conditions, and vehicle vibrations and varying temperature.

2.7 Data analysis

Several processing steps were conducted to clean up the raw data before the analysis.

2.7.1 Data collection, profiles, world models and data logging

An uncompressed video data stream was collected from each camera for both eye tracking systems, together with metadata including the 3D-position of the active gaze target. From this data a three-camera system profile was generated for each participant. A three-camera system profile consists of sets of images from each camera with facial features marked. The markings are used to build a 3D head model. For optimal performance the image sets, called poses, should be spread out and separated in order to cover most of the participant’s head movements during the test.

A Smart Eye World Model is used, being the same in both the one-camera and the three-camera system. This model contains the 3D-positions of the gaze targets. A special world model with four 3D-gaze targets is used in the three-camera system for calibrating the eye model for each participant.

The video stream, the profiles and the world models are fed into the three-camera system and the one-camera system and text logs are generated. The logs contain gaze

origin (corresponding to the pupil centre in the three-camera system case and to the iris centre in the one-camera system), gaze direction, gaze direction quality and the position of the current active gaze target. All positions and directions are in 3D.

To calculate the one-camera gaze in the VisualEyes setup we utilized three different measured entities:

1. The position and orientation of the participant's head.
2. The position of the participant's iris.
3. The position of the glints relative the irises. (Glint = reflections of IR-illumination in the cornea. The IR-illumination has known 3D-position.)

All of these entities influence the accuracy of the system in different ways:

- The error in the head orientation measurement increases as the participant turn his or her head away from the centre camera and the initialization pose. The component of the error increases along the direction of the rotation and is likely to be larger than the component transversal to the rotation.
- Since the upper and lower eyelids often obscure the iris upper and lower edges, respectively, the error of the iris position measurement is normally bigger in the vertical direction than the error in the horizontal direction.
- The error in the glint position is likely to be uniform in the vertical and horizontal directions and is probably not affected by head rotation, as long as the glint is well inside the iris perimeter, where the curvature of the cornea changes rapidly.

2.7.2 Calibration and quality

In order to compensate for the offset in the one-camera system, the median difference from the centre gaze target, i.e. the CP dot, is subtracted from the mean differences of all the other gaze targets. No calibration for the scale factor is performed in this analysis at the present stage (this could probably be performed in a post processing stage in the current experiment, but the quality of the results is not known). Only log entries corresponding to the activated dot are used in the calculations. Further, only log entries with nonzero gaze direction quality are included.

2.7.3 Performance indicators

The following definitions were used within the project for data analysis:

- Availability per participant per system: The total number of logged gaze entries per test person, for all gaze targets, having a nonzero gaze direction quality, divided by the maximum number of possible gaze log entries per test person, for all gaze targets.
- Availability per gaze target per system: The total number of logged gaze entries per gaze target, for all individuals, having a non-zero gaze direction quality, divided by the maximum number of possible gaze log entries per gaze target, for all test persons.

- Accuracy: the difference in angle between the mean of all tagged gaze log entries for a specific gaze target and the corresponding true gaze direction. The accuracy is expressed in degrees.
- Precision: the variability in angle between the tagged gaze log entries for a specific gaze target. The precision is expressed in degrees and the variability is quantified using the standard deviation of the gaze log entries.

Availability, which is related to the robustness of the system, is measured as the total number of logged gaze entries divided by the maximum number of logged gaze entries (i.e. the tracking ratio). If the system is able to log an entry, it will be classified as available regardless if the participant is looking at the correct target or not. For example, if the participant is staring at the CP dot while he or she is supposed to look at dot c5, c5 will get a very high availability value.

Accuracy reflects how close a series of measurement are to a reference value, while precision indicates the degree to which repeated measurements, under unchanged conditions, show the same results (Taylor, 1999), see Figure 4. Here the measured gaze directions are compared with the corresponding vectors between the gaze origin and the active gaze target, i.e. the true gaze. The mean of the differences between the measured gaze direction vectors and the true gaze vectors, split in horizontal and vertical components, are then calculated and expressed in degrees for each gaze target. The orthogonal horizontal and vertical components are then combined to give the mean Euclidean distance between the measured gaze direction and the true gaze direction. This provides a measure of accuracy, which is expressed in degrees.

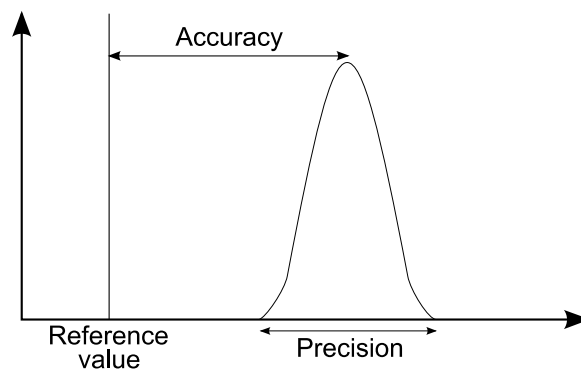


Figure 4 Accuracy indicates how close the measurements are to a true reference value while precision indicates the repeatability of the measure.

2D-maps for accuracy and availability were calculated based on the availability per gaze target and the accuracy per gaze target. Median values were used to aggregate across participants to avoid exaggerating the effect of outliers.

Since data are only available for a discrete set of irregularly sampled coordinates (the dots), the rest of the 2D map was interpolated with triangle-based cubic interpolation. An example can be found in Figure 5. The axes represent the horizontal and vertical gaze direction, respectively, and each of the targets is illustrated as black dots in this 2D space. The colours represent the accuracy of the system, ranging from blue (high accuracy) to red (low accuracy). For example, the centre gaze target is coded in blue which means that the eye tracking system provides results that are about zero degrees wrong (high accuracy). In contrast, some dots in the peripheral areas are coded in red which means that they can be more than 30° wrong (low accuracy). The green ellipses

represent the borders of three different regions that are used in the subsequent statistical analysis. The central area includes dots inside the inner ellipse, the middle region includes dots between the two ellipses and the outer region includes dots outside the larger ellipse. The ellipses have the horizontal radius r and the vertical radius $2r/3$, where $r = 20^\circ$ in the inner ellipse and $r = 120^\circ$ in the outer ellipse.

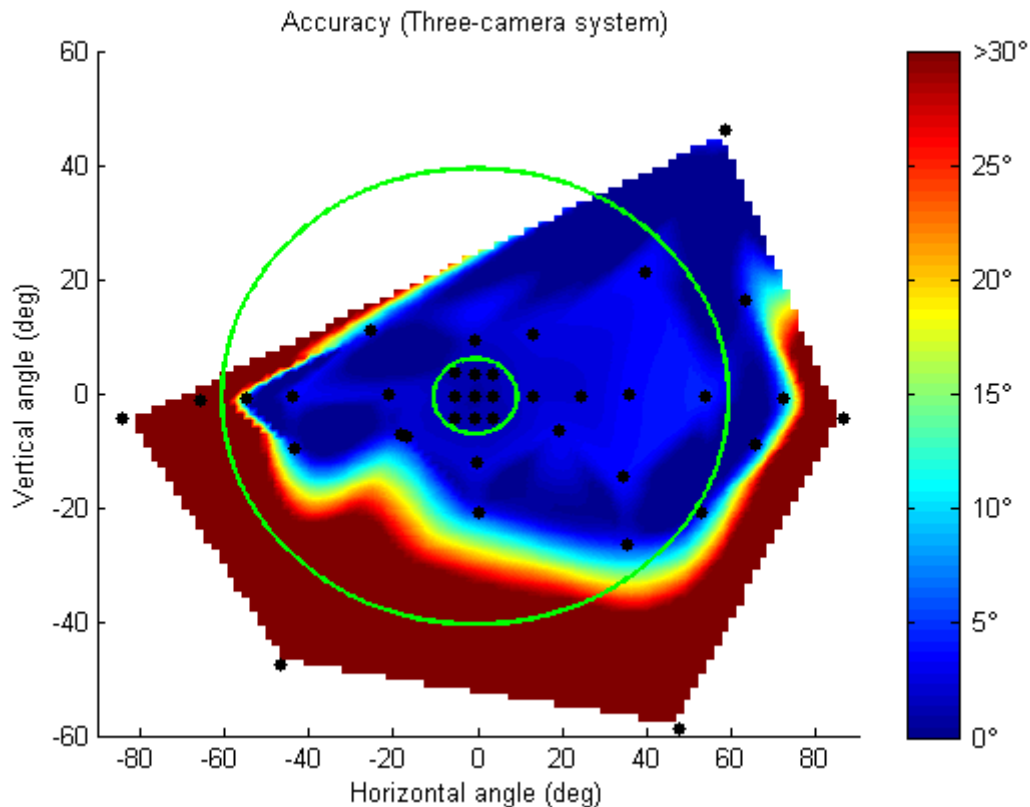


Figure 5 2D map of the accuracy in three-camera system.

2.7.4 Statistical analysis

Multiway (n-way) analysis of variance (ANOVA) was used to test the effects of multiple factors on the mean of the accuracy and availability measures per gaze target. This test compares the variance explained by factors to the left over variance that cannot be explained. The factors are glasses (on/off), eye tracking system (three-camera system versus one-camera system), distance from centre (continuous factor or discrete factor according to the elliptical regions in Figure 5), gender (male/female) and age (30–39, 40–49 and 50–60 years). Note that participant is not included as a factor since it entails missing factor combinations and leads to terms that do not have full rank. The chosen ANOVA model includes the main effects of each individual factor as well as interactions at all levels of the factors glasses, eye tracking system and distance from centre. The significance level was set to five percent. Possible significant differences are further analyzed with a multiple comparison test. Here the significance value was set to $\alpha = 0.05$ and Tukey's honestly significant difference criterion was used.

3 Results

Out of the 53 participants, four were excluded from further analysis due to technical problems during data collection. During data collection, some additional problems were encountered:

- Incomplete fixation during the 2 seconds that the gaze target was lit. The participants shifted their gaze from the target before the logging was completed.
- Erroneous camera calibration (8 occurrences). This only relates to the three-camera system.
- Bad initialisation (9 occurrences). This relates to the one-camera system only.

The collected data can be seen as a matrix with the 53 participants as rows and the 41 dots as columns. The cells belonging to low quality data were excluded from further studies.

The final 49 participants (24 female, 25 male) had a mean \pm standard deviation (s.d.) age of 44 ± 8 years (range 32–59), five wore contact lenses, nine wore make-up, seventeen had long hair and six of them had either a moustache or a beard. The distributions of participants in the two groups that started without glasses and with glasses, respectively, were similar (Table 1).

Table 1 Characteristics of participants analysed.

	Number	Female (Male)	Age span	Mean age (s.d.)
Group started without glasses	24	9 (15)	33-49	45 (8.6)
Group started with glasses	25	15 (10)	32-48	45 (7.8)

3.1 Availability

ANOVA results related to availability are summarized in Table 2. The factors distance, age and gender show significant effects on availability. Especially, the availability decreases with the distance from the centre region, see Figure 6. The interaction between distance and the eye tracking system that is used is also evident, where availability decreases more with distance for the one-camera system as compared to the three-camera system. There is also a significant interaction between glasses and distance, and in Figure 6, where it can be seen that availability decreases when the participant is wearing glasses, especially outside the centre region.

Table 2 Results from the ANOVA analysis for availability.

Source	Sum Sq.	d.f.	Mean Sq.	F	p
Glasses	2150.5	1	2150.5	2.21	0.14
Eye tracker	2289	1	2289	2.35	0.13
Distance (continuous)	4450935.8	1	4450935.8	4566.11	0
Age	58900.1	2	29450	30.21	0
Gender	17612.1	1	17612.1	18.07	0
Glasses*eye tracker	7.6	1	7.6	0.01	0.93
Glasses*distance	11345.9	1	11345.9	11.64	0
Eye tracker*distance	344318.6	1	344318.6	353.23	0
Glasses*tracker*distance	52.4	1	52.4	0.05	0.82
Error	6175213.8	6335	974.8		
Total	11359214.7	6345			

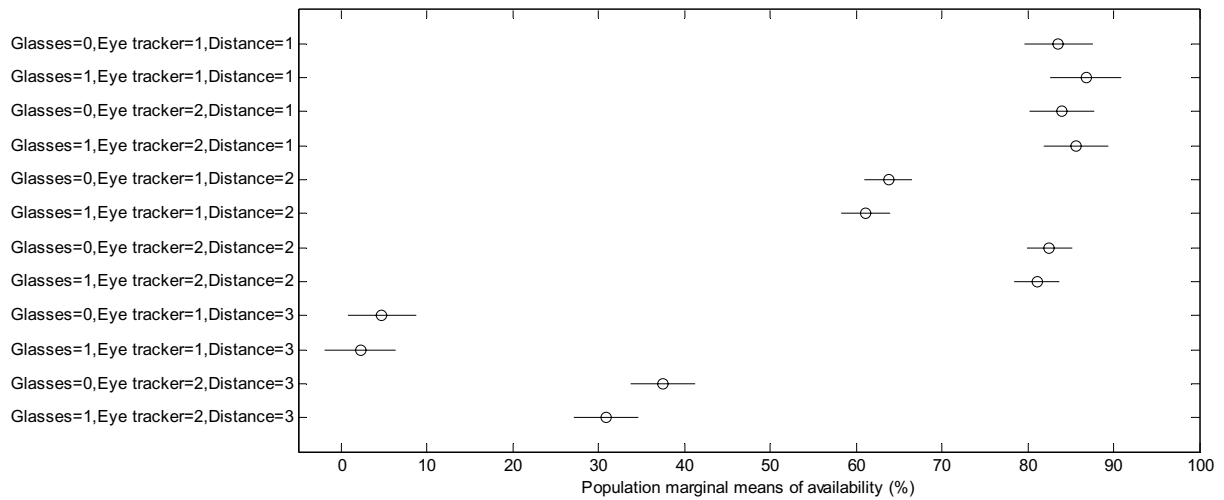


Figure 6 Results from multiple comparison test showing confidence intervals for availability for the three factors glasses (no = 0, yes = 1), eye tracker (One camera = 1, three-camera system = 2) and distance (centre region = 1, middle region = 2, outer region = 3).

The overall availability of the three-camera system and the one-camera system was 72% and 54%, respectively, see Table 3. The availability in the different regions clearly decreases with the distance from the centre dot, especially for the one-camera system. However, the difference when the participant wears glasses is strikingly small. Availability values per participant when they do not wear glasses and when they do wear glasses is also illustrated in Figure 7. The missing values are due to the quality issues already mentioned. Also note the large variability amongst the participants.

Table 3 Overall availability (centre <20°, middle 20°–120°, outer >120°).

	No glasses				Glasses			
	All	Centre	Middle	Outer	All	Centre	Middle	Outer
One-camera system	53.9	84.0	63.8	5.0	52.2	86.9	60.8	2.1
Three-camera system	71.6	84.1	82.1	37.2	69.6	85.7	80.7	30.7

Figure 8 and Figure 9 show 2D maps of accuracy and availability for the one-camera and the three-camera system. Figure 8 provides maps when the participants do not wear glasses while Figure 9 is derived from data with glasses. The three-camera system has high accuracy and availability in a larger area compared to the one-camera system (blue regions). This is expected since multiple cameras have a larger coverage. The small island in the three-camera system measurements where the availability is somewhat lower as compared to the surroundings originates from two dots located inside the cockpit of the car.

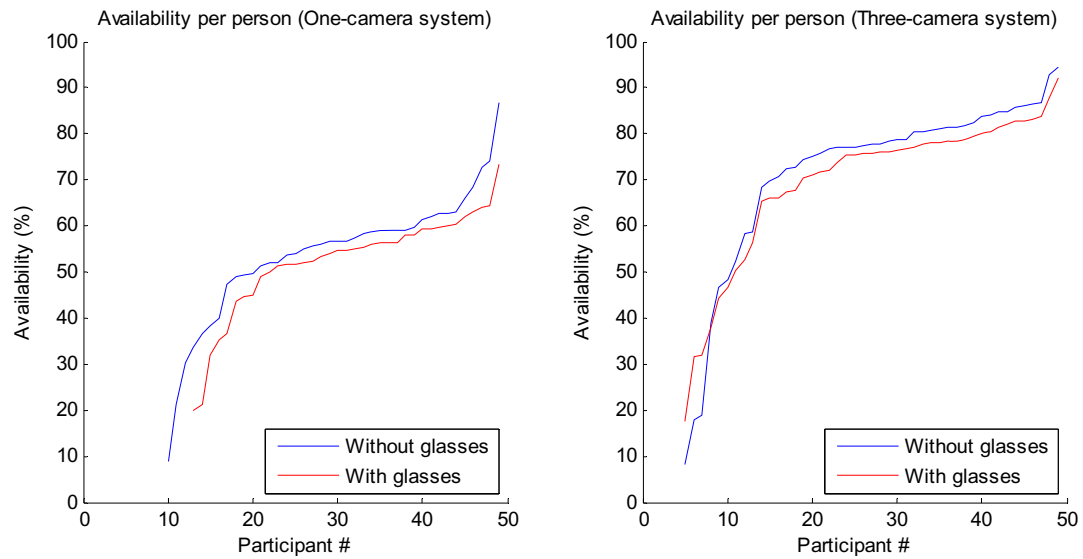


Figure 7 Availability per participant, sorted by increasing availability.

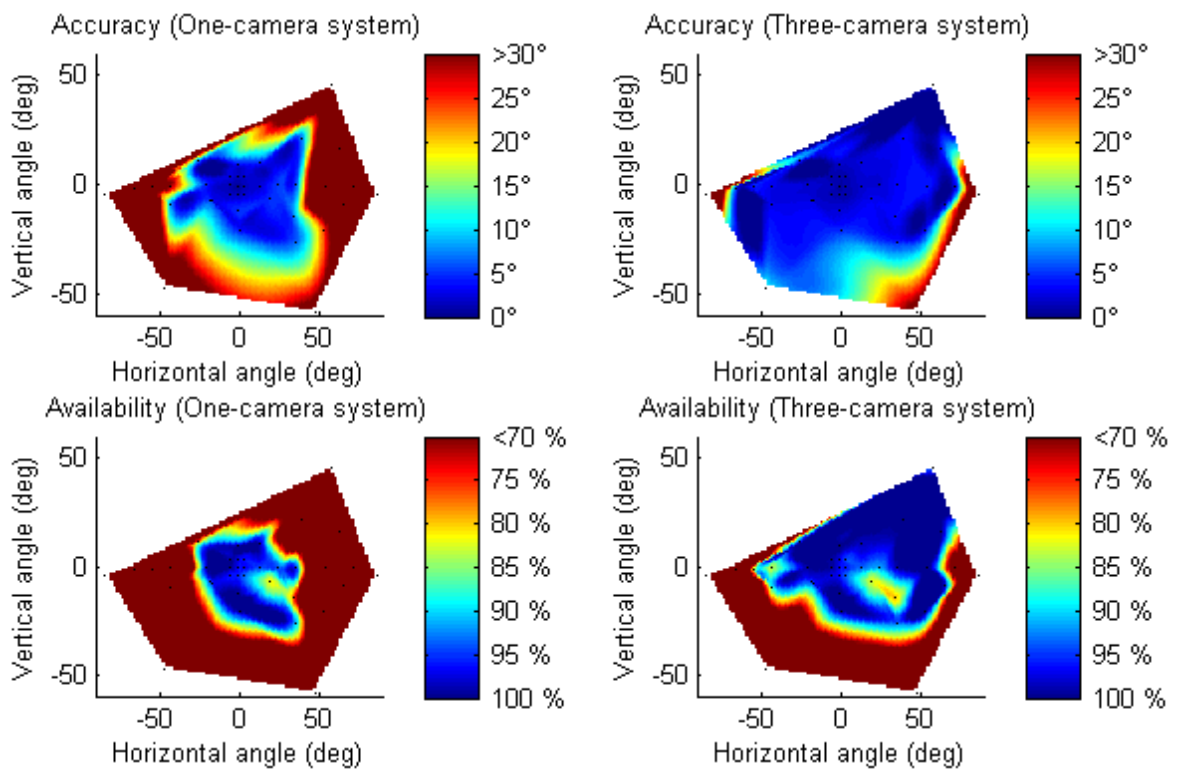


Figure 8 2D-maps of accuracy and availability when not wearing glasses.

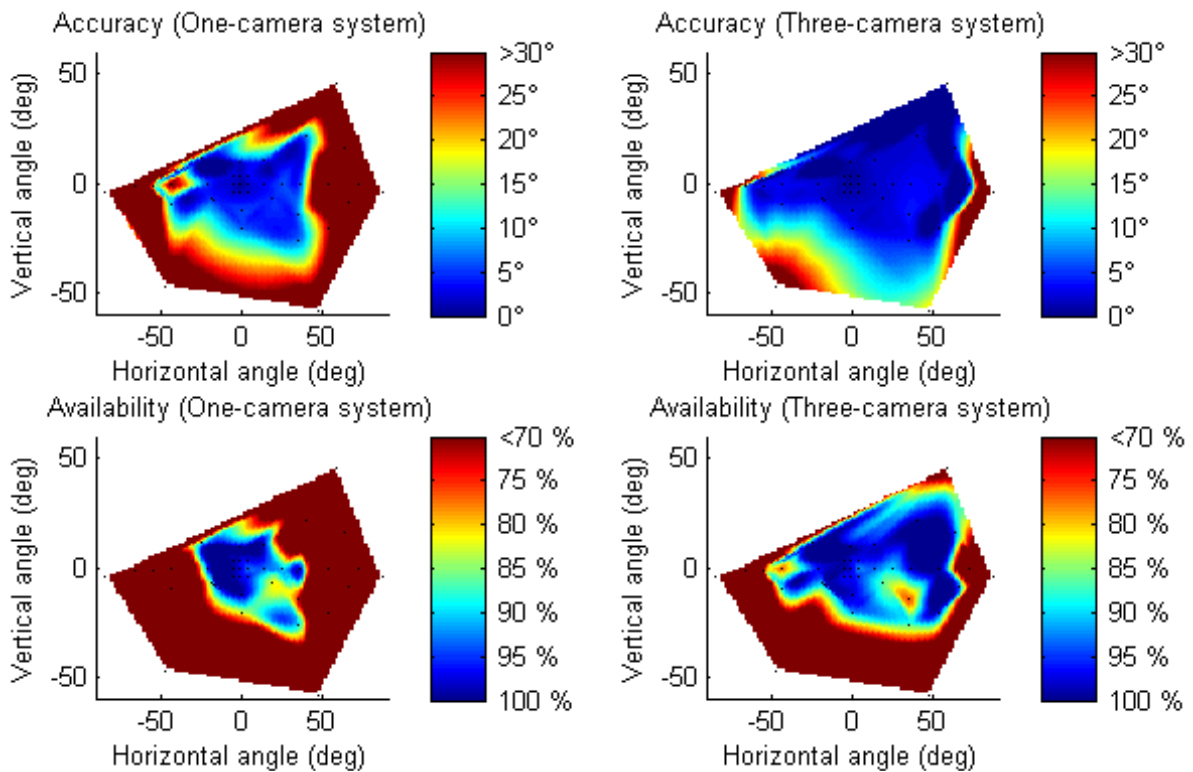


Figure 9 2D-maps of accuracy and availability when wearing glasses.

3.2 Accuracy

ANOVA results related to accuracy are summarized in Table 4, where it can be seen that most factors and interactions show significant differences. The exceptions are glasses and gender. To find out where the differences were located a multiple comparison test was performed for the factors glasses, eye tracking system and distance (Figure 10). The main differences are due to the eye tracking system that is used and the distance from the centre region. It can also be seen that the accuracy for the one-camera system is as low as 60° for the outer region. This is probably due to the participant looking straight ahead at the wrong dot (contrary to the instructions), resulting in a large measurement error. This is supported by Figure 6, where availability is very low for the one-camera system in the outer region.

Table 4 Results from the ANOVA analysis for accuracy.

Source	Sum Sq.	d.f.	Mean Sq.	F	p
Glasses	97.1	1	97.1	1.54	0.21
Eye tracker	7432.3	1	7432.3	118.07	0
Distance (continuous)	179716.2	1	179716.2	2854.91	0
Age	663.5	2	331.8	5.27	0
Gender	11.7	1	11.7	0.19	0.67
Glasses*eye tracker	389.5	1	389.5	6.19	0.01
Glasses*distance	410.1	1	410.1	6.51	0.01
Eye tracker*distance	89935.6	1	89935.6	1428.69	0
Glasses*tracker*distance	767.6	1	767.6	12.19	0
Error	259794.2	4127	62.9		
Total	481493.6	4137			

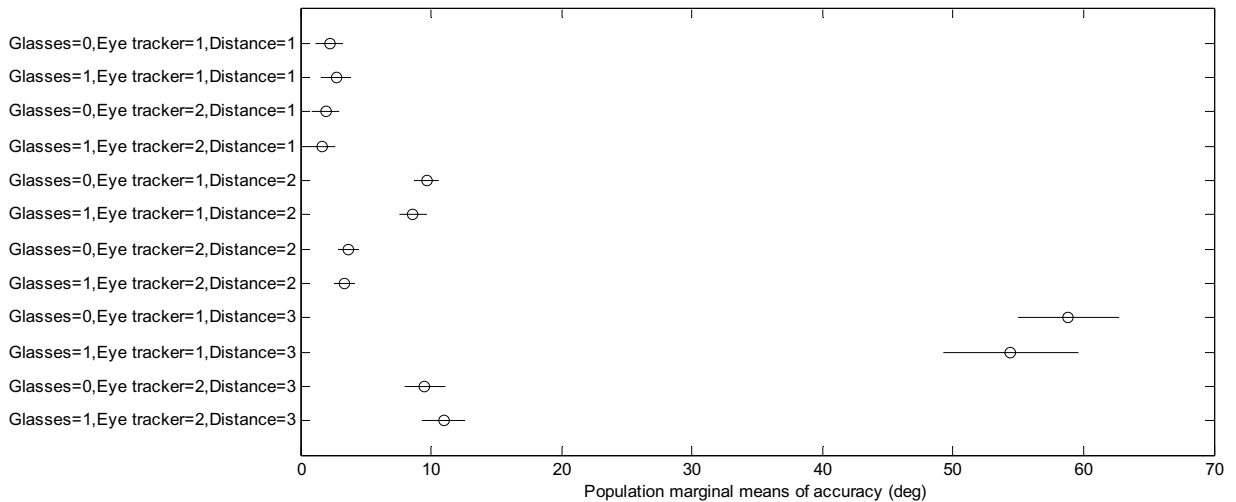


Figure 10 Results from multiple comparison test showing confidence intervals for accuracy for the three factors glasses (no = 0, yes = 1), eye tracker (One camera = 1, three-camera system = 2) and distance (centre region = 1, middle region = 2, outer region = 3).

3.3 Precision

Precision is measured as the horizontal and vertical standard deviation of the gaze direction for each dot. Similarly, accuracy is divided in a horizontal and a vertical component in this section, where each direction indicates the average distance between the reference dots and the measured gaze direction. Note that accuracy has been treated as the Euclidean distance in 2D space in previous sections as compared to the separate horizontal and vertical components used here. Also note that precision is reported as standard deviation values that have been averaged across participants. This means that precision is not really the same as standard deviation in this report but rather an average of several standard deviation measures.

Figure 11 illustrates the precision of the one-camera and the three-camera system, respectively, when the participants are not wearing glasses. In the central area the precision is similar between the two systems, but the three-camera system performs better in peripheral areas. The centre dot has lower precision than its neighbouring dots, a result which is probably due to the experimental setup. Since each participant looks at the centre dot in between looking at every other dot, the standard deviation is prone to increase. Results have been omitted from dots in cases where more than 85% of the data across participants have insufficient quality. This is why some of the ellipses are missing.

Figure 12 contains the same information about precision as Figure 11, but accuracy has been included in the figure as well. As before, both accuracy and precision are comparable for the two systems in the central area whereas the three-camera system performs better in the periphery. More details as well as a summary of how much data that was finally used in the accuracy and precision calculations are summarized in Table 5. The effect of glasses on precision is illustrated in Figure 13 and Figure 14. No systematic changes in precision due to glasses versus no glasses could be found.

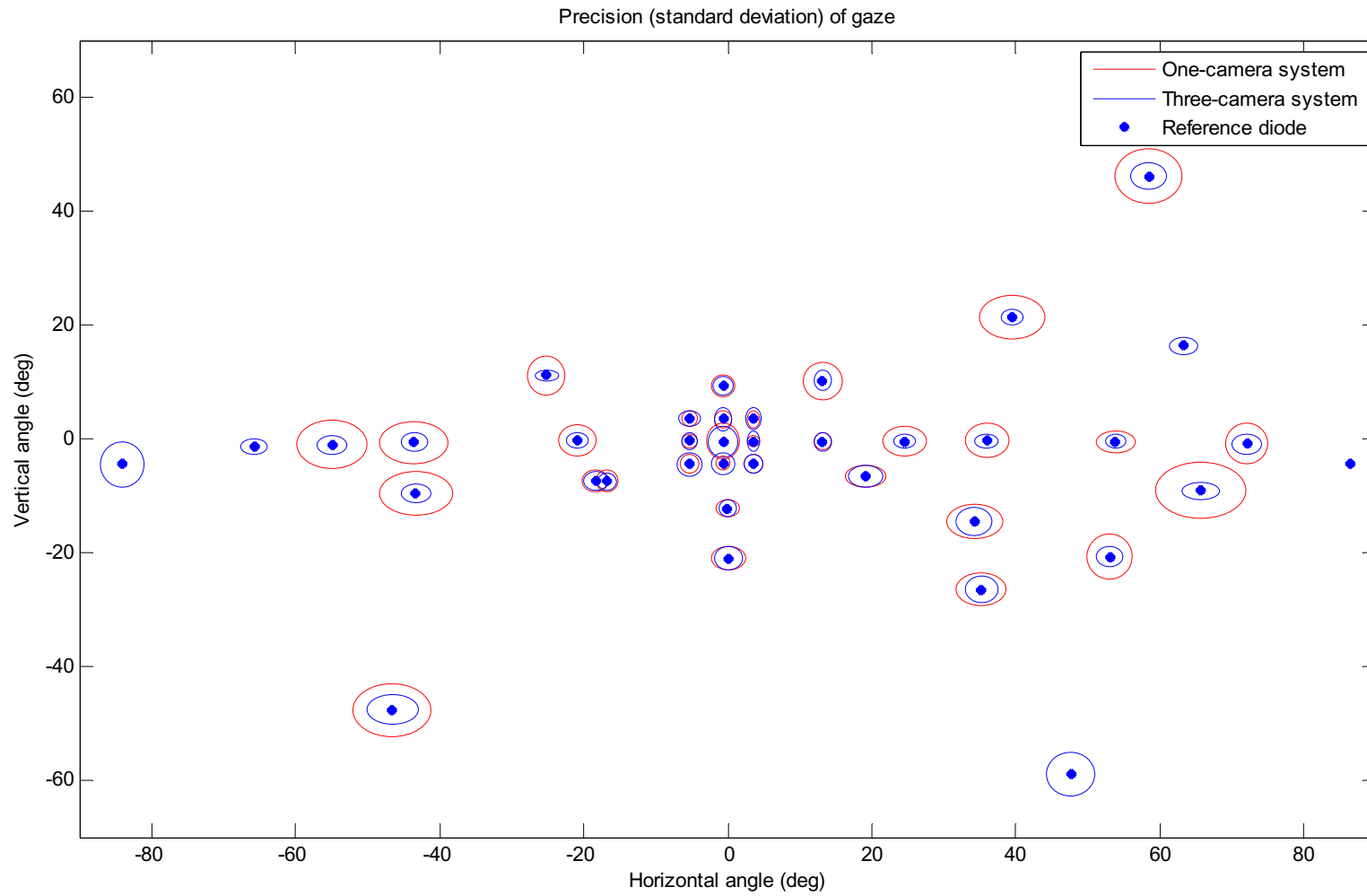


Figure 11 Precision of the measurements for each dot, where the ellipses indicate the average standard deviation across participants (no glasses). Results are omitted in cases where the precision calculations are based on less than 15% of the data.

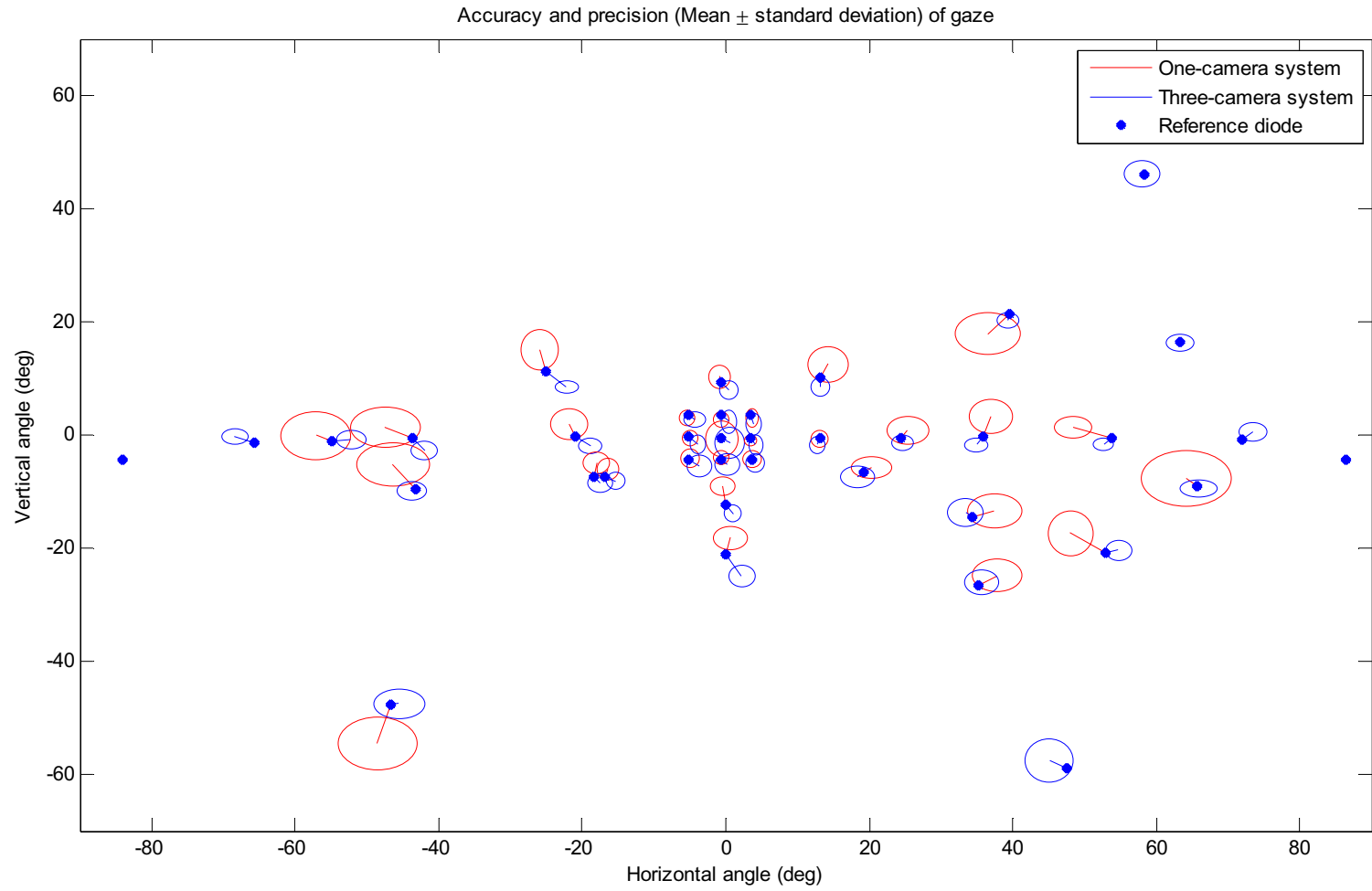


Figure 12 Precision and accuracy of the measurements for each dot, where the ellipses indicate the average standard deviation across participants (no glasses) and the lines indicate the accuracy (distance from the reference point). Results are omitted in cases where the precision calculations are based on less than 15% of the data.

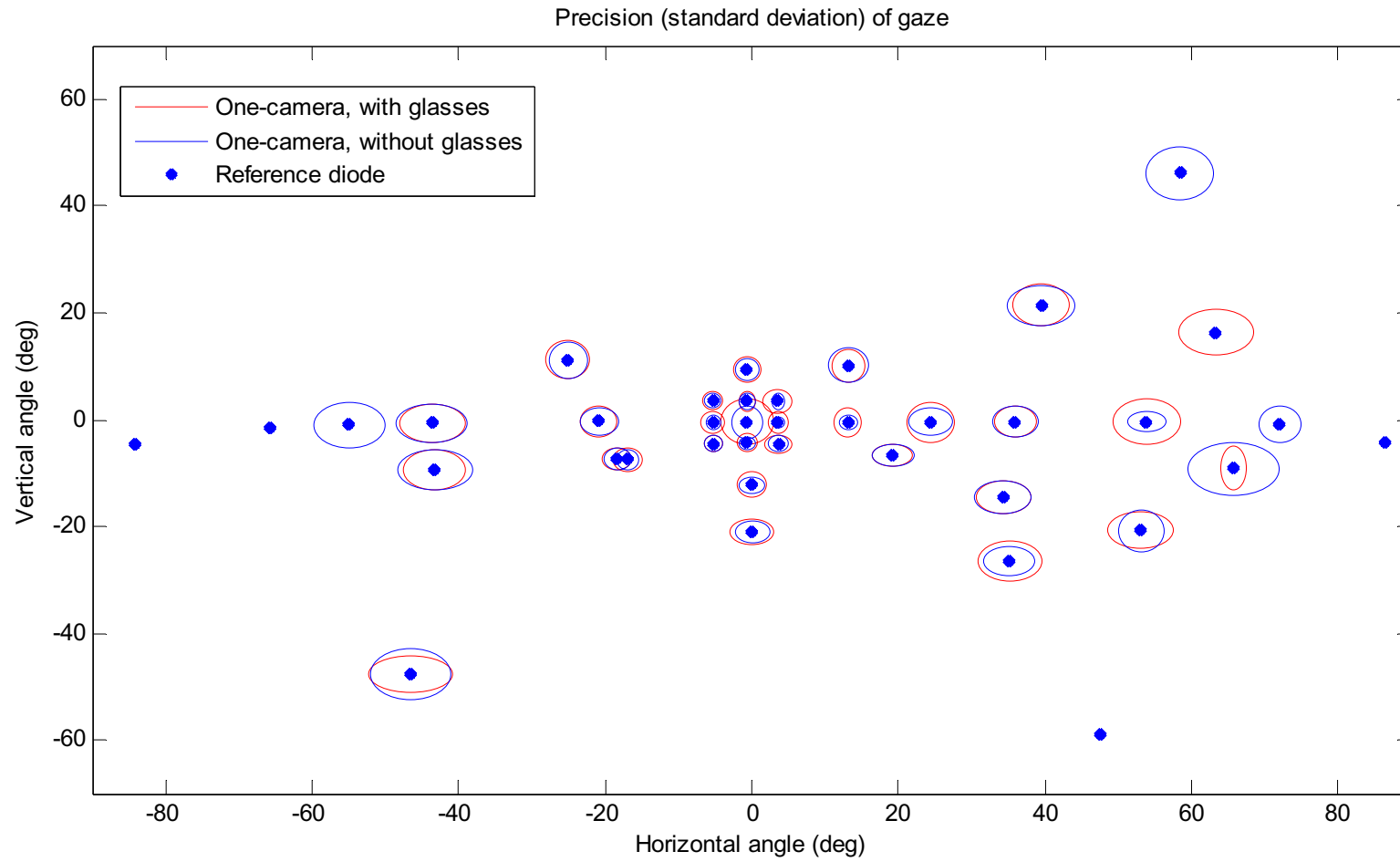


Figure 13 Precision for the one-camera system when the participants are either wearing glasses or not. The ellipses indicate the average standard deviation across participants (no glasses). Results are omitted in cases where the precision calculations are based on less than 15% of the data.

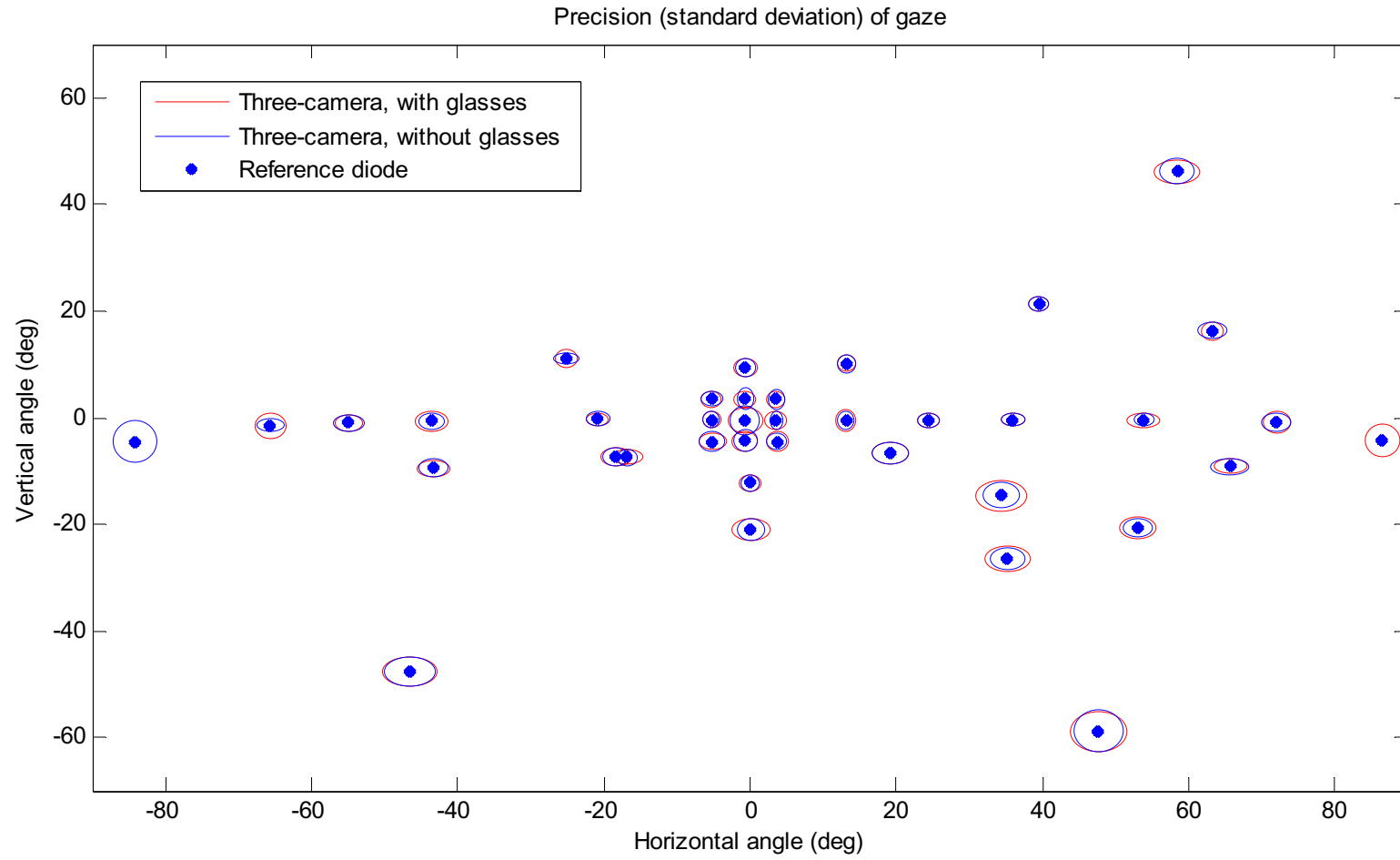


Figure 14 Precision for the three-camera system when the participants are either wearing glasses or not. The ellipses indicate the average standard deviation across participants (no glasses). Results are omitted in cases where the precision calculations are based on less than 15% of the data.

Table 5 Horizontal and vertical accuracy and precision for each dot and for one-camera and three-camera system, respectively. Also included in the table are the percentages of participants (amount data) that the accuracy and precision calculations are based upon.

Dot	One-camera system			Three-camera system			
	Amount data (%)	Horizontal (deg)	Vertical (deg)	Amount data (%)	Horizontal (deg)	Vertical (deg)	
Central area	c0	100	0.11±2.22	-0.27±3.14	100	1.18±1.99	-0.87±2.78
	c10	86	0.21±1.28	0.39±1.64	96	1.50±1.71	-1.06±1.99
	c15	86	-0.00±1.04	0.42±1.22	94	0.83±1.64	-0.90±1.96
	c16	86	-0.00±1.30	0.25±1.58	94	0.43±1.31	-0.57±1.60
	c17	80	0.37±1.03	-0.19±1.33	92	1.38±1.16	-1.21±1.60
	c18	88	-0.15±0.84	-0.51±0.97	96	0.65±0.93	-1.22±1.74
	c19	88	-0.04±1.11	-0.73±1.41	96	1.05±1.12	-1.10±2.09
	c20	84	-0.18±1.11	-0.57±1.34	92	1.01±1.46	-0.91±1.39
	c21	88	0.10±0.84	-0.43±1.55	94	0.36±1.09	-1.55±1.92
Intermediate region	c3	16	-2.22±4.87	0.88±4.26	70	2.65±2.01	0.25±1.63
	c4	58	-3.87±4.85	2.00±3.64	84	1.61±1.83	-1.94±1.64
	c5	78	-0.81±2.60	3.99±3.48	92	2.92±1.65	-2.61±1.02
	c6	86	-0.96±2.58	2.17±2.66	90	2.06±1.61	-1.52±1.37
	c7	86	-0.16±1.53	0.91±1.93	94	1.06±1.34	-1.51±1.64
	c8	92	1.06±2.73	2.29±3.23	94	0.01±1.20	-1.73±1.73
	c9	86	-0.08±1.15	-0.08±1.48	94	-0.36±1.11	-1.22±1.57
	c12	24	-5.46±2.62	1.94±1.89	90	-1.22±1.34	-1.19±1.19
	c14	86	0.34±1.48	1.37±1.92	92	1.49±1.31	-0.64±1.48
	c22	92	0.88±3.00	1.33±2.50	98	0.06±1.51	-0.92±1.32
	c23	70	1.03±3.02	3.60±3.02	92	-1.03±1.62	-1.35±1.23
	c25	88	1.08±2.84	0.83±1.94	96	-0.86±2.40	-0.66±1.96
	c30	56	-5.04±3.11	3.42±3.94	100	1.69±1.88	0.35±1.73
	c32	46	-3.04±4.55	-3.53±3.71	88	-0.20±1.45	-1.16±1.30
	c33	86	3.18±3.82	1.17±2.99	98	-0.78±2.49	0.82±2.45
	c34	90	0.31±1.79	2.39±1.97	96	0.81±1.70	-1.06±1.67
c36	86	-0.45±1.71	3.17±1.54	94	0.98±1.27	-1.63±1.48	
c37	88	0.47±2.35	2.84±2.12	92	2.13±1.87	-3.95±1.99	
c38	90	2.61±3.47	1.59±2.82	98	0.47±2.30	0.45±2.21	
c40	76	-3.09±5.11	4.29±3.83	86	-0.48±2.04	-0.34±1.62	
Outer region	c1	8			10		
	c2	6			46	-2.74±1.86	1.05±1.32
	c11	8			94	-0.00±1.97	-0.13±1.50
	c13	4			8		
	c24	12			68	1.45±2.00	1.46±1.71
	c29	24	-1.51±6.24	1.44±4.95	86	0.22±2.60	-0.38±1.55
	c31	12			70	-0.42±2.41	0.07±2.35
	c35	20	-1.94±5.42	-7.00±4.68	32	1.02±3.57	0.12±2.63
c39	8			30	-2.50±3.37	1.34±3.86	

3.4 Summary of results

Based on the presented results, the seven hypotheses can be verified or falsified according to:

1. The one-camera system has a lower accuracy and availability compared to the three-camera system over the total test envelope (covering 180 degrees horizontally and 110 degrees vertically).
2. *The accuracy was higher for the three-camera system (Table 4 and Figure 8) but there was no significant difference in availability (Table 2). The mean differences in Table 3 may seem large, but they are not significant on the five percent level. Note however that there is an interaction effect between eye tracking system and distance (Table 2), indicating that availability decreases more with increasing distance from the centre region for the one-camera system as compared to the three-camera system.*
3. The three-camera system has higher accuracy for points outside the central area compared to the one-camera system.

There are significant differences in accuracy between the two eye tracking systems in the two regions outside the central area (Figure 10), and Figure 8 reveals that the three-camera system has higher accuracy outside the centre region.

4. Accuracy (of the gaze and head direction towards the dots) will decrease with increasing distance from the centre points in both systems (1 and 3 cameras).

There is a significant difference in accuracy both for the continuously measured distance (Table 4) and for the discrete distance (Figure 10) from the centre point. This decrement in accuracy as the distance from the centre point increases is also illustrated in Figure 8.

5. Presence of glasses on participants' face lead to lower accuracy /availability compared to without glasses.

There was no significant effect of glasses on accuracy or availability. There were however significant interaction effects between glasses and distance for both accuracy and availability. Especially, it can be seen that availability decreases with glasses in the peripheral region (Figure 6).

6. High degree of wrinkle is influencing the accuracy/availability: the higher the presence of wrinkle the lower the accuracy/availability.

The amount of wrinkles was not measured in the study so this hypothesis can not be answered. However, there is a significant effect of age on accuracy which might be related to wrinkles.

7. Presence of make-up will deteriorate accuracy/availability compare to non make up.

There were not enough participants wearing make-up to answer this hypothesis.

8. The participant characteristics skin colour and unshaved does not influence accuracy and availability on either system.

The diversity in skin colour characteristics amongst the participants was not varied enough to be tested in the study.

4 Discussion

The motive for the VisualEyes project was to evaluate eye tracking systems based on one or three cameras in terms of accuracy, availability and precision. The results show that both accuracy and availability decreases with the distance from the centre region, and that the decrease is larger for the one-camera system as compared to the three-camera system.

4.1 Methodology

A standardized test of eye tracking equipment has been developed. A set of gaze targets have been defined and a number of performance indicators have been identified. This section discusses some limitations with this methodology that should be considered in future studies.

4.1.1 Experimental design

There are so many important and interesting factors that need to be controlled for when evaluating and comparing eye tracking systems. When it comes to driver characteristics, it is important to have proper face visibility to obtain a high quality tracking. If the face is temporarily covered, for example by the driver's hands while eating, drinking or talking on the mobile phone, it is important that tracking is instantly re-established as soon as the face becomes visible again. Such performance comparisons were not feasible in the present experimental setup. More permanent coverage of the face from headbands, high collars, caps, scarves etc. would also have been interesting to test for, but such a study setup would have been too extensive.

Some of the driver characteristics that were taken into account include skin colour, facial hair and makeup. However, due to the relatively small number of participants in the study, it was still difficult to test for these factors. For example, not enough people were wearing makeup to be able to test for this factor. Similarly, the presence of different skin colours and facial hair were not diverse enough to perform any statistical tests. Consequently, hypotheses six and seven could not be answered.

Glasses are another factor which affects the tracking quality. This is partly due to the frames which might block the view of the cameras and partly due to the glass which might give obscuring reflections from the IR illumination. Flat lenses leads to large reflections, but the probability that the reflection is directed straight towards the camera is low. With a convex lens, the size of the reflection is smaller. However, the reflection obscures the camera's view of the eye more often. If the lens has different curvature on the front and back surface, the reflections are still small as in the convex lens case. However, there will be twice as many reflections since the reflections from the front and the back surface will not coincide. In this study, a pair of mock-up glasses with convex lenses was used by all participants. This might have been an unfortunate choice since the results showed no statistical differences due to glasses. A different study approach would have been to use fewer participants but with a larger variety of different glasses. In hindsight, this would have been a better approach since glasses is one of the key parameters that we intended to investigate.

Suggestion for future studies: Limit the number of factors by a more homogenous study population.

4.1.2 Head-eye coordination

The head direction vectors were not analysed in this study, mostly because the experiment was not designed to investigate head directions. For example, the participants' head/eye movements are not interacting in a way that is similar to an actual driving situation. When looking at a dot for two seconds in the lab, the head and eye direction is usually aligned while in real life the head and the eyes are not necessarily linearly correlated (Collins & Barnes, 1999). For example, it has been indicated that the head is being aligned with the expected reorientation of the car in respect to the environment whether the reorientation is based on visual or remembered information, and that the eyes are controlled by the vestibulocollic reflex during these head movements (Land, 1992; Proudlock & Gottlob, 2007). Besides this, little research on the topic has been made, and the following statements should be considered as suggestions based on the authors' observations and experience. The orientation of the head/nose vector compared to the gaze vector for a driver in a real driving situation is influenced by a number of factors, including:

- Fixation history/Preceding fixation
- Hysteresis
- Spontaneous or planned fixation
- Knowledge of the next fixation
- The duration of the fixation
- The angular separation between two consecutive fixations
- The mobility of the human eye relative to the head
- How much the nose obscures the field of view.

The influence of these factors can be exemplified as follows:

- For *small saccades* and short fixations **no head rotation** is needed. For example, a short glance in the interior rear view mirror while driving can be accomplished without moving ones head.
- For *large saccades*, e.g. a quick glance in the outer right rear view mirror, the gaze shift is so large that it implies a rotation of the head, since it would be very uncomfortable for the eyes otherwise. As we know that we soon will redirect our gaze back onto the road again, we do not rotate the head all the way to the mirror, but just enough to be comfortable.
- When performing fixation with **long duration and a large deflection** from the road, we sometimes rotate the head all the way to the gaze target. For example while waiting for an entry slot at a crossroad.
- The position of the head/nose vector during a fixation is dependant on the **previous/next fixation**. If we change our gaze back and forth between the right outer rear view mirror and the interior rear view mirror, the head/nose vector, when looking at the interior rear view mirror, is likely to be to the right of the mirror. If we instead change our gaze back and forth between the left outer rear view mirror and the interior rear view mirror, the head/nose vector is likely to the left of the interior rear view mirror, when we look at it.

- During a *spontaneous fixation* with a large deflection angle the first thing that is happening is a saccade relative to the head, followed by a head rotation to avoid an uncomfortably large eye deflection relative to the head.
- If we instead analyse a *planned fixation* with large deflection angle outside the comfort zone, it can be observed that we prepare for the fixation by rotating the head prior to the saccade. If we, for instance, want to adjust the radio, it can be observed that the head/nose vector is prepositioned somewhere in between the road and the radio. From that position we can let the gaze switch back and forth between the road and the radio, while performing the task.

There is no simple relation between the head/nose vector and the gaze vector. If the head/nose vector is to be used as an indication of the driver's gaze, as a fallback when the gaze measurement fails for one reason or the other, it is not possible to utilize a direct mapping of the head/nose vector versus the ground truth gaze vector. There is, however, a lot of useful information in the head/nose vector:

- If the head/nose direction relative to the road is greater than the envelope of the eye deflection comfort zone, the gaze direction is not focused on the road in that instant.
- If the head/nose vector is directed towards the road, the gaze is most probably looking at the road as well, possibly with short intermissions of fixations with small deflections, e.g. to the inner rear view mirror.
- If the head/nose vector is rotated relative to the road, but still in the comfort zone, the gaze direction is probably not aimed at the road for more than part of the time.

To extract more information from the head/nose vector, further studies in a naturalistic setting is needed. The experimental setup used in this study resembles a car standing still where the participants do not have to apprehend other road users, traffic situation, traffic signals or other information. The fixations are performed in a planned manner with no other visual tasks interfering and "unlimited" time allocated for finding the gaze target, before the data logging begins. For a useful comparison between head/nose direction and ground truth gaze direction a setup which mimics a real driving situation more closely is needed.

Suggestion for future studies: The gaze direction measurements are adequate, but for head direction measurements, the participants head direction and gaze direction should be aligned (aim the head in the direction of the target).

4.1.3 Data processing

To find data of low quality, a threshold of ten degrees was applied to the accuracy results. This means that all results where the measurement error was ten degrees or more were marked as low quality candidates. Each of these candidates was then investigated manually and if the low accuracy value was due to a problem in the measurements this value was omitted from further analysis. Such problems include:

- Incomplete fixation
- Bad initialization
- Measurement error

- Initialized without eyes
- Reflections
- Bad camera calibration.

It might be questionable if all of these problems are valid causes for omitting data from the study. For example, ‘incomplete fixations’ should definitely be removed since the participant did not follow the instructions. On the other hand, ‘bad initialization’ or ‘measurement error’ reflects the robustness of the system rather than a problem with the experimental setup.

In the statistical tests, the factors glasses, eye tracking system, distance from centre dot, gender and age were used. The choice of factors is based on the available data set and therefore deviates from more logical choices based on the hypotheses. For example, skin colour, makeup and facial hair should have been included as factors, but as already stated, the collected data set does not allow these factors to be tested since too few participants in the experiment wore makeup, were bearded etc. The selection of participants should have controlled for these factors, but for practical reasons only Saab employees were enrolled in the study and in this population it was hard to fulfil all criteria.

Suggestions for future studies: Exclude data that are of low quality because the participant did not follow the instructions. Do not remove data that are of low quality due to eye tracking malfunction. Effects of these latter errors should have an impact on the results.

4.1.4 Availability, accuracy and precision

There are a few peculiarities related to the three performance indicators that are used to assess the eye tracking systems. The most important thing is to remember that the three should always be interpreted as a whole and not as three standalone indicators.

Suggestion for future studies: Complement the suggested performance indicators with an overall indicator that takes both availability and accuracy into account. For example, combine availability with accuracy so that a measurement is available only when certain accuracy is achieved.

4.2 Results

Most of the results are very intuitive. According to the statistical tests, both accuracy and availability deteriorates with distance from the centre dot, especially for the one-camera system. This means that with more cameras, you obtain higher accuracy, availability and precision over a larger area. Besides these expected results, there is also one very surprising finding, namely that glasses does not significantly affect the tracking performance. Since wearing glasses is often claimed to give poor tracking results in practise, this is rather strange. The only explanation we can find is that the mock-up glasses were not realistic enough and that the subjects’ head movements were unrealistic in the static test environment.

Availability is measured as the amount of logged gaze entries divided by the maximum amount of possible gaze entries. This is measured during the two seconds when the participant is supposed to look at a particular target dot. During this time, it is assumed that the participant is really looking at the dot and not somewhere else. This means that

it is possible to cheat the system. By looking straight ahead, where the eye tracking system has better performance, instead of at a peripheral dot where one is supposed to look, it is possible to get high marks on availability. At the same time, by doing so, accuracy will indicate large errors. This is what happens for the one-camera system in the outer region in Figure 10, where it can be seen that the system has an accuracy of about 60° . What is really happening is that when the system is able to measure anything at all, the participant is not looking at the target dot but straight ahead instead. Similarly, accuracy (and precision) is calculated based on data with a certain quality. This means that only high quality data is used in the actual calculations. If the system is unable to measure the gaze direction for most of the time, the calculations will be very unreliable. It would have been possible to punish such unreliable data with a weighting function, but this was not done in this study.

As a consequence of the way that gaze is measured in the one-camera system, see section 2.7.1, one would expect that the precision error in the central area of the measurement range would be combined by a uniform glint error, an approximately uniform head rotation error and an iris position error being larger in the vertical direction than in the horizontal direction. As the head is panned horizontally away from "straight forward" we would get an increased error in the head rotation and consequently in the gaze. The horizontal component of this error is likely to be larger than the vertical component, which can be seen in Figure 11 – Figure 14.

The operational range of the one-camera system is about $\pm 30^\circ$ horizontally and about $\pm 20^\circ$ vertically. This means that it is not possible to measure eye movements in peripheral areas with a one-camera system. This fact has important implications, for example, it is not possible to assess eye movements directed towards the left and right rear view mirrors or towards the middle console. The operational range of the three-camera system is wider, about $\pm 55^\circ$ horizontally and about $\pm 35^\circ$ vertically. Once again, choose the system that fit your needs. When wearing glasses, the operational range shrinks with a few degrees, especially in the horizontal direction.

5 Guideline

The present “Guideline” will advise researchers undertaking measurements with camera systems in vehicles and simulators for gaze direction calculations. It is composed by three different parts where the first part addresses general requirements independent of the number of tracking cameras. It is followed by two parts where specific requirements for single camera and multi-camera systems are presented. For specific requirements related to hardware and software the reader is referred to the Smart Eye manuals for both Anti Sleep (single-camera system) and Smart Eye pro (multi-camera system).

5.1 Minimum requirements

Several requirements need to be fulfilled in order to perform eye tracking measurements. A minimal setup required for eye tracking is listed below where aspects related to driver characteristics, physical parameters and camera locations are defined.

5.1.1 Driver characteristics

Constant face visibility: To obtain high quality tracking, a constant level of face visibility should be ensured. Clothes covering the face such as headbands, high collars, caps, sunglasses, strong make-up, piercing on the face, scarf will deteriorate tracking quality. The same thing applies to temporary cover of the face which occurs during eating, drinking or talking in a hand held mobile phone, or otherwise obscuring the camera view with the hands or other objects. All kinds of face obstruction will momentarily affect tracking quality until the face is fully visible again.

Glasses: Sunglasses can affect the image content in the eye region. Some glasses are so dark that eye features are hard, or even impossible, to track. Other sunglasses are almost totally transparent in the near infrared spectrum. Unfortunately there is no simple way of knowing which, other than testing. Glasses in general will affect the gaze tracing negatively by three different mechanisms:

- The frames will sometimes obscure the eye corners, and thereby degrade the head tracking quality.
- The (doubly) curved lenses can exhibit (double) reflexes of the IR flashes. These reflexes can sometimes, at certain head orientations, coincide with the iris and pupil and thereby degrade the gaze tracking quality. In part this can be avoided by intelligent positioning of the cameras and the IR flashes.
- The optical properties of the lenses will distort the perceived eye geometry and introduce nonlinearities. The impact of this mechanism is, however, often negligible.

Eye colour: The larger the contrast of the eyes the better the tracking will be. Blue eyes have been observed to give the best contrast both between iris and the white part of the eye and iris and pupil. However, all eye colours give enough contrast to allow tracking with good quality.

Length of the driver: Drivers who are extremely short or extremely long will be problematic since it is difficult to adjust the cameras’ visual angle so that the driver’s face is visible in the camera. A range within 155 to 193 cm has been used in prior field studies.

5.1.2 Camera locations

Cameras should be placed in a low position to catch the drivers' eyes and head position. If several cameras are used, they should be positioned to cover the area from the left hand side behind the B-pillar to the right hand side just beyond the right side rear view mirror. Locations in the middle console, close to the left rear view mirror and close to the right rear view mirror have been identified as satisfactory camera locations. Due to head movements and vehicle geometry there are not too many degrees of freedom to position the cameras in optimal locations. Notice that camera location should primarily be decided in accordance with the purpose of the measurements.

For one-camera systems, the camera is usually positioned either on the steering wheel column or close to the centre rear view mirror. The location should be chosen so that the camera can see the driver's eyes as good as possible, i.e. in a position right in front of the driver where it doesn't block the driver's view.

5.1.3 Vehicle

To install tracking equipment in a vehicle, a certain amount of space is required. Power source, computers and wiring composed the major parts of the system and will put requirement on a certain level of temperature and humidity that should be fulfilled to assure quality in the measurements. Note that due to the temperature requirements, it may be beneficial to mount data acquisition hardware inside the car instead of in the trunk.

5.1.4 Physical parameters

Light conditions, both inside the vehicle and on the outside, may be an issue when recording eye movements with cameras. Since the cameras are using powerful IR flashes attached on the side of the camera, disturbances from strong sunshine is minimal. The IR flashes might, however, add reflexions when using glasses. The use of IR-flashes might also disturb other camera-based systems in the car or simulator.

Vibrations might disturb the system. It is important to fixate the camera on a solid support and to assure that the brackets of the cameras are strong enough. This issue is important in field studies as well as when moving base simulators are used. Temperature changes may disturb the system (because the camera position will shift a bit due to the expansion of the cockpit material). By automatically recalibrating the head model regularly this problem can be remedied. When the cameras have been securely locked, they have to be carefully focused. The focus should be optimised for the glints in the eye. Normally there is no further need for refocusing, as long as the operating distance remains roughly the same.

Prior to the initial set-up, one has to define the so called "head box", i.e. the volume in which one wants to be able to track head and gaze. By choosing optics with the correct focal distance and by aligning the camera, a cone angle and orientation is defined, giving a cone in which the head is fully visible. In the single-camera case this cone, truncated by the focus span, constitutes the headbox. In the multi-camera case, the intersections of these truncated cones define the headbox.

Since any 3D triangulation requires at least two cameras, more than one camera should always be used if exact depth information (i.e. distance from the camera) is needed. A one-camera system estimates the depth information based on generic facial dimensions.

5.2 Single-camera systems

Application: A single camera system is most suitable as part of a warning system alerting sleepy or distracted drivers. In those applications you're interested in the blink behaviour of the driver and whether the driver is looking at the road or not. Additional information about single-camera systems includes:

- Limited field in which measurements are accurate. (The gaze direction should not be too far from the camera.)
- The camera is usually positioned in the steering wheel region.
 - If the camera is placed on the steering wheel column, one has to consider that the column can be adjusted by the driver, in which case a new world coordinate system definition might be necessary.
 - If the camera is placed behind the steering wheel, the camera will be intermittently obscured by the hands and/or the wheel spokes, while turning the wheel.
- 3D head position, with depth information based on generic facial dimensions, is mapped into a 3D world model. The same applies to gaze.
- Fully automatic initialisation. The eyes, nose and mouth have to be visible for gaze direction calculations. At least the nose and the mouth have to be visible to calculate head direction. During large head rotations, it is beneficial for the accuracy and availability if the participant's ears are visible.
- Automatic gaze calibration is theoretically possible but is yet to be implemented. Offline statistical gaze calibration can be performed manually in post processing.

Advantages: Single-camera systems are cheaper, easier to operate and easier to install in a vehicle compared to multi-camera systems.

Disadvantages: The level of accuracy is lower than in a multi-camera system when it comes to areas far from the road centre. No gaze calibration is implemented yet.

5.3 Multi-camera systems

Applications: Multi-camera systems are mostly used for research purposes. Here, it is of interest to know in detail what the driver is looking at. For that purpose the higher level of precision of multi-camera systems is essential. With multi-camera systems, larger head rotations can be covered as well (up to 360°). Application areas are for example studies on driver behaviour in traffic. Additional information about multi-camera systems includes:

- The camera placement is flexible and can be adapted to both space/packing limitations as well as any specific measuring task. (E.g. extra camera on the centre console for optimal accuracy in that region or extra camera positioned next to right outer rear view mirror for optimal accuracy in that region.)

- More cameras give a higher degree of accuracy and a wider field of measurement, but as well higher cost and a system that is more difficult to install.
- The position of the IR flashes has to be defined in the set-up parameters. (In the one-camera system, the positions are predefined.)
- True 3D head position and gaze are mapped into the 3D world model.
- In order for the gaze tracking to function properly at least one eye has to be visible in at least two cameras.
- The cameras position relative to each other needs to be calibrated by a chessboard calibration. It is good practise to perform this prior to each measurement session, since the cameras can be affected by both thermal and mechanical perturbation. Due to the use of 3D triangulation, the multi-camera systems are inherently more sensitive to spatial perturbations than single-camera systems.
- The generation of the personal profile is not yet fully automatic, and often requires a little hands-on to be optimal.

Advantages: A larger head rotation envelope is covered with high degree of accuracy, availability and precision.

Disadvantages: The cost of the system is higher than for a one-camera system and the installation require some work. The system is not fully automatic.

References

- Antisleep User Manual. Smart Eye AB, Gothenburg, Sweden.
- Collins, C. J. S., & Barnes, G. R. (1999). Independent control of head and gaze movements during head-free pursuit in humans. *Journal of Physiology-London*, 515(1), 299–314.
- Coughlin, J. F., Reimer, B., & Mehler, B. (2009). *Driver Wellness, Safety & the Development of an AwareCar: AgeLab White Paper*.
- Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and practice*. London: Springer-Verlag.
- Kircher, K. (2007). *Driver distraction: A review of the literature*. Linköping, Sweden: VTI (Swedish National Road and Transport Research Institute).
- Land, M. F. (1992). Predictable Eye Head Coordination during Driving. *Nature*, 359(6393), 318–320.
- Lee, J. D., Young, K. L., & Regan, M. A. (2009). Defining Driver Distraction. In M. A. Regan, J. D. Lee & K. L. Young (Eds.), *Driver Distraction: Theory, effect and mitigation* (pp. 31–40). London: CRC Press, Taylor & Francis Group.
- Proudlock, F. A., & Gottlob, I. (2007). Physiology and pathology of eye-head coordination. *Progress in Retinal and Eye Research*, 26(5), 486–515.
- Salvucci, D. D., & Goldberg, J. H. (2000). Identifying fixations and saccades in eye tracking protocols. Paper presented at the Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL, USA.
- Smart Eye Pro User Manual. Smart Eye AB, Gothenburg, Sweden.
- Smith, P., Shah, M., & Lobo, N. D. (2003). Determining driver visual attention with one camera. *Ieee Transactions on Intelligent Transportation Systems*, 4(4), 205–218.
- Taylor, J. R. (1999). *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*. New York: University Science Books.
- Wierwille, W. W. (1993). Demands on driver resources associated with introducing advanced technology into the vehicle. *Transportation Research C*, 1(2), 133–142.

Information to participants and informed consent

VIP – VisualEyes

Testledare: _____

Informerat samtycke

Undertecknad har tagit del av den skriftliga och muntliga informationen angående studien ”*VisualEyes - en studie med mätning av ögonrörelsedata i simulatorn*” och accepterar att delta på angivna villkor:

Testledaren har beskrivit testet för mig och jag förstår vad som förväntas av mig. Jag har fått svar på eventuella frågor . Jag accepterar att mätdata, bild- och inspelningarna av mig från mättillfällena kan användas vid presentationer av studien, och kommer att sparas och användas för forskning. Jag vet att jag har rätt att när som helst avbryta studien utan närmare förklaring.

Notera att insamlade rådata kan spåras till dig, eftersom vi filmar dig. Vi publicerar dock ingenting som kan spåras till dig om du inte uttryckligen ger oss tillstånd till det. All rådata kommer att lagras i en databas hos Saab. De andra företag och institut som deltar i projektet har tillgång till databasen. De som kommer att ha tillgång till data är forskare som deltar i projektet, men ingen utanför. Slutsatser dragna på statistisk nivå från en sammanvägning av data kommer att publiceras. Dina svar och dina resultat kommer att behandlas så att inte obehöriga kan ta del av dem. Insamlad data kommer att separeras från ditt namn och dina personuppgifter. Vi vill ej att du pratar om studiens upplägg till andra innan hela studien är slutförd

Datum: _____

Underskrift: _____

Namnförtydligande: _____

Arne Nåbo
Projektledare

Experiment protocol

Before the experiment begins the experiment leader must:

- Explain the study and give the informed consent form to the participants. This must be signed by the participants.
- If the participant has long hair, this should be tied in a way not to obscure facial features and so that the ears are visible.
- Take two pictures of the participant with mock-up glasses and two without mock-up glasses.
- Record the following data for each participant (with examples below):

Unique number assigned to participant	04
Gender	M
Age	45
Length (measure!)	169 cm
Facial hair (for males) (none, moustache, beard, heavy beard...)	heavy beard
Make-up (yes/no)	no
Earrings (yes/no)	one, left side
Skin colour (pale white/white/brown/dark brown)	white
Eyelid opening (small, medium, large)	medium

Experiment leader must instruct participants to **avoid** the following during the experiment:

- placing hands in front of camera and have his/her hands on the face
- eating or drinking
- fast / sweeping head movements
- talking on mobile phones
- unnatural driving position
- wear any hats, scarves, or similar items
- use chewing gum during the drive
- change seat or steering wheel position while the experiment is ongoing

A protocol has to be kept for each driver, in order to note all particular occurrences which may be of importance for the data analysis. One page per participant, identification by participant number. Drivers are allowed to adjust the seat position in the simulator, but they must not adjust the steering wheel position, as one camera is mounted on the steering column. Adjustments when the experiment is ongoing must be avoided. The right focus of all cameras has to be checked when the participant has adjusted the seat before starting the experiment. Mock-up glasses must fit the participants. A measuring tape will be used by the test leader to measure the body length of the participant. The pictures taken from each participant are as well part of the protocol. File name for each picture is the unique number assigned to each participant, and "front" or "side" to indicate the profile. Temperature inside the simulator cab should be measured and recorded once. It is assumed that this is constant for all participants, thus only one measurement when starting the experiment is necessary (not for all participants).

Each participant will perform the experiment two times: once with and once without glasses. Participants will wear mock-up glasses in the first or second run according to the sequence specified. The order should anyway be recorded in the protocol. A short brake is allowed between the passes (5 minutes).



ViP

Virtual Prototyping and Assessment by Simulation

ViP is a joint initiative for development and application of driving simulator methodology with a focus on the interaction between humans and technology (driver and vehicle and/or traffic environment). ViP aims at unifying the extended but distributed Swedish competence in the field of transport related real-time simulation by building and using a common simulator platform for extended co-operation, competence development and knowledge transfer. Thereby strengthen Swedish competitiveness and support prospective and efficient (costs, lead times) innovation and product development by enabling to explore and assess future vehicle and infrastructure solutions already today.



Virtual Prototyping and
Assessment by Simulation

Centre of Excellence at VTI funded by Vinnova and ViP partners

VTI, Saab Automobile, Scania, Volvo Trucks, Volvo Cars, Bombardier Transportation, Swedish Transport Administration,
Dynagraph, HiQ, Pixcode, SmartEye, Swedish Road Marking Association

www.vipsimulation.se

Olaus Magnus väg 35, SE-581 95 Linköping, Sweden – Phone +46 13 204000