The comparison of accuracy and precision of eye tracking: GazeFlow vs. SMI RED 250.

Document version 1.1, August, 2013
SIMPLY USER, User Experience Lab
Kraków, 6 August, 2013

Abstract

The report describes the results of a research comparing the Accuracy and Precision
of the GazeFlow eye tracking software based on the image from webcams with the SMI
RED 250 device, a standard eye tracker using infrared light to track the position of an eye.
The measurement of Accuracy and Precision was taken using the method suggested by
Tobii Technology.

The conclusions obtained show sufficient results in the area of accuracy and very high
results in the area of precision of the GazeFlow software. This means a prospect for
commercialization of the software for commercial marketing research purposes,
as well as controlling a computer with eyesight, meaning a non-contact computer interface.

Table of contents

1.   Introduction

1.1.   Eye tracking devices

Eye tracking devices (eye trackers) have in recent years been extremely popular research tools. Dynamic progress and the increase of accessibility of devices (price drop) have made eye tracking research commonly used for commercial purposes. Most of all the eye tracking research has made its way into marketing research methods and website usability research.

The most widespread eye tracking devices are currently the ones using infrared light to track pupils. Among the stationary eye trackers available on the market and used in commercial research there are two prevalent solutions coming from Tobii Technology and SMI Vision companies.

What is more, during the last two years the first commercial solutions appeared, which use webcams for eye tracking. Most often this kind of service is available online, and as with infrared-based trackers, it is used for marketing and website research.

At the moment there are three solutions available, YouEye (www.youeye.com), Gazehawk (www.gazehawk.com) and EyeTrackShop (www.eyetrackshop.com). Among the research practitioners the solutions using webcams are criticized for low accuracy and significant discrepancies of the results in comparison with solutions based on infrared light (see: Aga Bojko[1]).

It is important to note that eye tracking devices, especially those based on widely available solutions, have a huge potential for being used as a tool for controlling machines. Eye tracking may become one of the next means for human – computer interaction, thanks to which it will be possible to control and interact solely with eyesight. Mostly, such solutions may be used in entertainment, for instance in games to control a character and in general gameplay, however the most promising area of application is neurorehabilitation, i.e. the prospect of using eyesight controlled computers to interact with people who lost the ability to communicate or use devices in a standard way.

The goal of the report is the comparison of the results and data obtained for two solutions: an eye tracker based on infrared light and software using a webcam.

An eye tracker based on infrared light is a standard solution of the SMI company, RED 250 is a device with high resolution equipped with software enabling efficient commercial research.

---

[1]Blog: http://rosenfeldmedia.com/books/eyetracking/blog/the_truth_about_webcam_eye_tra/

The second solution is GazeFlow, an authoring solution of Szymon Deja, based on the analysis of the optical flow of webcam image. The software which currently can be used as a standalone solution, is intended for further commercialization

1.2. Comparative tests

The comparative tests have employed the solution suggested by Tobii Technology described in the Accuracy and Precision test method for remote eye trackers report (Tobii, 2011). The method suggested by Tobii is aimed to objectify and create an opportunity to compare devices manufactured by different suppliers. The accuracy and precision of all devices manufactured by Tobii Technology is assessed according to this method.

The devices are assessed in two areas using this method: Accuracy and Precision.

Accuracy is the reading of an average difference between the position of a stimulus and the measured position of an eye. Precision is the ability of the device to repeat a reliable measurement.

The matrix shows all possibilities and relations of accuracy and precision of the measurements of eye positions made by the eye tracking device.

High accuracy High precision	High accuracy Low precision
Low accuracy High precision	Low accuracy Low precision

Only the systems with high accuracy and precision deliver reliable and adequate measurements of the position of an eye on the screen. This means that on the basis of the systems' readings we get the information on the actual position of an eye and the measurement is repeatable. A good measurement of accuracy is considered to be an average smaller than $0.8°$ under ideal conditions (measurement taken with a fixed head, in lighting conditions of circa 300 lux).

A good measurement of precision is considered to be an average precision smaller than $0.5°$ under ideal conditions.

The level of required accuracy of the equipment for eye tracking depends largely on the type of research and the kind of the analyzed stimuli. The smaller the analyzed stimuli are, or if the process of reading is also considered, the higher the requirements concerning accuracy and precision are. In case of the marketing materials commonly used in eye tracking analyses, or website researches, especially in case of remote researches regarding the free-following of stimuli by the people examined, the requirements are more liberal.

In the Accuracy and Precision Test suggested by Tobii (2011) the results of different experimental conditions are being compared, among others:
- ideal conditions
- different eye angle
- different lighting conditions
- different head positions

In case of the following research special attention was paid to different head positions due to a significant effect of a head position on the measurements taken by software utilizing webcams.
In the subject literature there is a discussion regarding taking the measurements on one dominant eye versus an average measurement for both eyes (see Tobii, 2011).
The following research employs binocular measurement.

2. Methodology

2.1. The persons examined

The experiment involved 30 people – 13 men and 17 women. The average age of the subjects was 27 (the youngest person was 21 and the oldest 39). The education of the persons examined was as follows: 14 people held a high school diploma, 12 people had master's degree and 4 had bachelor's degree. The summary of the subjects' demographic data is included in Table 1. The persons examined represented a random sample of the population between 20-40 years of age. During the selection process people with visual impairment or wearing glasses were excluded from the research due to the requirement of employing eye tracking equipment using infrared light.

No.	Sex	Age	Education
B1	M	25	bachelor
B2	M	28	high school
B3	W	28	master
B4	W	39	high school
B5	W	23	bachelor
B6	W	38	high school
B7	M	31	high school
B8	W	25	master
B9	W	28	master
B10	W	25	high school
B11	W	23	high school
B12	M	29	master
B13	W	30	master
B14	M	26	high school
B15	W	23	bachelor
B16	W	39	master
B17	M	27	high school
B18	M	37	high school
B19	W	24	high school
B20	W	23	bachelor
B21	M	29	master
B22	W	30	master
B23	M	22	master
B24	W	23	master
B25	W	38	master
B26	M	25	master
B27	M	25	high school
B28	M	21	high school
B29	W	24	high school
B30	M	28	high school

Table 1. The summary of the subjects' demographic data.

2.2. The time and place of the research

The research took place in a laboratory of the Simple User company between 13-17 June, 2013. The room was arranged in a way which allowed for stable and controlled conditions for conducting the experiment. The subjects sat at a desk in front of a screen with a remote eye tracking device SMI RED 250 and 4 webcams. The remote eye tracker was placed under the screen, and webcams were located both above and under the screen. Image 1 shows the monitor on which the experimental procedure was presented to subjects. The monitor was located around 65-70 cm in front of the subjects' eyes (the standard distance recommended in remote eye tracking research).

Image 1. The arrangement of webcams.

During two variants of the procedure, in order to fix a subject's head in one place, a chinrest located 70 cm from the monitor was used. Image 2. shows the location of a subject's head on the chinrest.
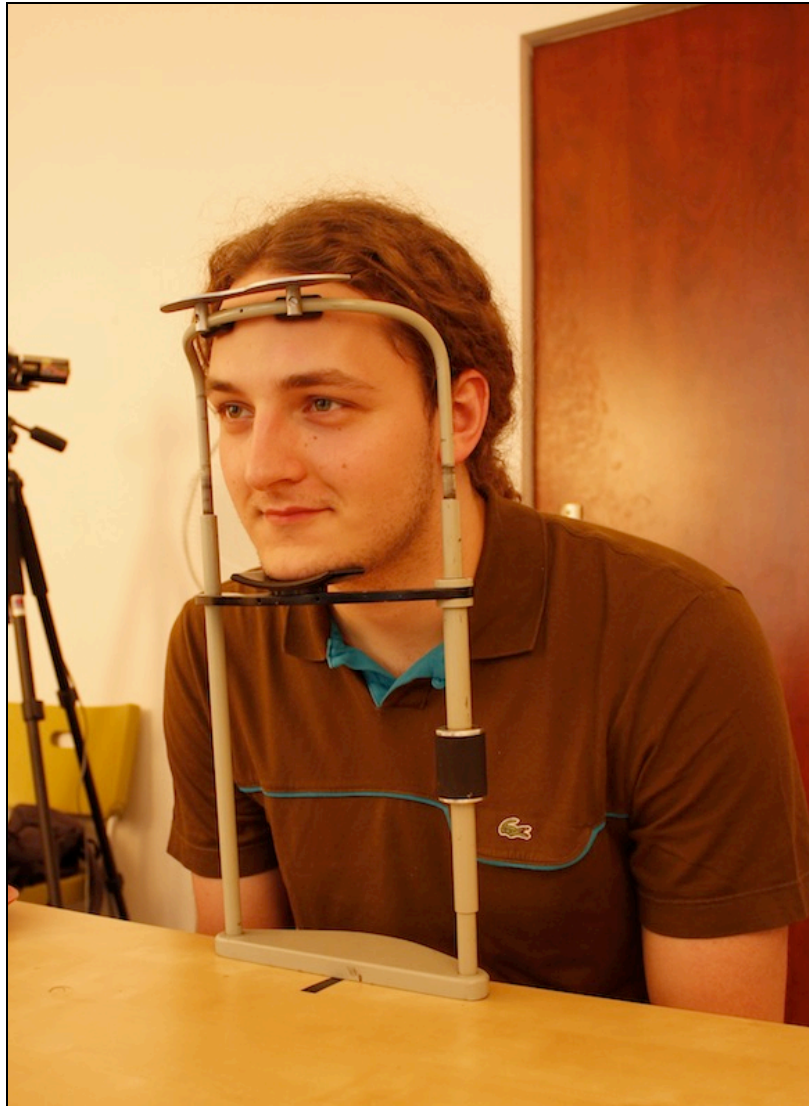

Image 2. Fixing the head in one place using chinrest.

All experiments were conducted in controlled lighting conditions. The lighting consisted of two softbox lamps. Illuminance in the room was circa 350 lux. During the experiment the lighting conditions were not manipulated. The research was conducted under the supervision of a qualified technician and a researcher.

2.3. Testing equipment

The research used remote eye tracker SMI RED 250 Hz with the iView X 2.8 system (installed on a dedicated laptop being an integral part of the device). Experimental stimuli were being displayed on a 22'' Dell LCD screen with a resolution of 1680x1050.

In addition, eyes were also tracked by the GazeFlow software using webcams.
The research employed 4 differently arranged webcams:

1.   Microsoft Lifecam 5000 (CAM_0) – placed above the screen
2.   Logitech pro 9000 (CAM_1) – placed under the screen
3.   Logitech HD Pro Webcam C920 (CAM_2) – placed above the screen
4.   Microsoft Playstation PsEye (CAM_3) – placed under the screen

The webcam parameters are: resolution of 640 x 480 and 30 fps frame rate.

The experimental procedure responsible for stimuli presentation was controlled through a dedicated software using iViewX SDK on a RF711 Samsung laptop with a 17'' screen.
The computer parameters:
-      IntelCore i7-2630QM 2 GHz processor
-      6 GB RAM
-      Windows 7 Home Premium (64-bit version) operational system
-      17,3'' (16:9, LCD) screen size
-      1600 x 900px resolution
-      32 bit color depth

The software was connected with the SMI RED 250 device through Wi-Fi. Image and data from the webcams, as well as data from the eye tracker, were recorded simultaneously.

To analyze the results a dedicated software was used as well, which calculated the Accuracy and Precision on the basis of formulae suggested in the article describing the Tobii method (Tobii, 2011).

2.4. Experimental procedure

2.4.1. Instructions

The experimental procedure was initiated by informing the subjects about the goal of the research, which was watching the displayed images and website dumps. The subjects were also informed about the necessity of following closely with their eyes the center of the moving point while the calibration of both eye trackers took place (the SMI device and the GazeFlow software). The instructions for subjects included detailed information on the way of moving and changing position of their heads during the calibration procedure, as well as using the feedback displayed on the screen. During the procedure the subjects were not given any additional information apart from a reminder to closely follow the point during calibration.

2.4.2. Calibration

The experiment consisted of nine experimental procedures. Each of the procedures included a combination of calibration, validation and presenting stimuli. In seven procedures the subjects had freely positioned heads, while in two they were fixed on chinrests (see Image 2).

The research employed a double calibration of the devices due to the fact of using two solutions. Below is a description of calibration for particular solutions:
1.   SMI RED 250 calibration – 9-point calibration using a moving white and red point against a grey background; 4-point validation using the same point;
2.   GazeFlow calibration – 10- to 30-point calibration (depending on the efficiency of the procedure) using a displayed red, pulsing point against a grey background; validation using a green point.

Additionally, in procedures with a freely kept head, the head movement was taken into account in the calibration process (calibration with a moving head), or the lack thereof (calibration with a fixed head).
Calibration accounting for the moving head was made on the basis of visual feedback information (calibration with a moving head). The subjects were presented with a yellow point with an arrow in different positions, indicating the direction of a head shift. After obtaining the desired head position the point changed its color to red and another

calibration point was displayed. During the introductory procedure the subjects had a chance to test this style of calibration.

During the experimental procedure the subjects were shown feedback information in case of an incorrect head position (in a form of a red head imitation), the subjects' task was to correct their head position in order to receive positive feedback.

2.4.3. Experimental stimuli

The research employed images from commonly accessible free databases, having a balanced salience, as well as website dumps of the most popular websites in Poland (according to the Gemius ranking).
The display of stimuli was preceded by the presentation of a fixation point against a background, whose color was the average of all colors of a particular image in order to balance illumination of the presented experimental stimuli.

2.4.4. Variants of the experimental procedure

All experimental procedures consisted of the following stages:

1.   Instructions
2.   Head initialization
3.   SMI calibration
4.   SMI validation
5.   GazeFlow calibration
6.   GazeFlow validation
7.   Stimuli presentation
8.   GazeFlow validation

The table below shows all the variants of experimental procedures used in the research.

Procedure Proper name	Fixed head calibration	Free head calibration	Chinrest	Types of stimuli and the time of display
Procedure 1 Into	_	YES	_	_
Procedure 2 HeatMapSet1	_	YES	_	10 photos 5 sec with the point of fixation
Procedure 3 HeatMapSet2	YES	_	_	10 photos 5 sec with the point of fixation
Procedure 4 HeatMapWWW1	YES	_	_	5 screenshot website 9 sec with the point of fixation
Procedure 5 ReadText	_	YES	_	reading text aloud
Procedure 8 Glass_2_HeatMap_Set4	YES	_	_	10 photos 5 sec with the point of fixation
Procedure 9 Glass_3_HeatMap_www2	_	YES	_	5 screenshot website 9 sec with the point of fixation
Procedure 6 Statyw_WalidationColor	_	_	YES	_
Procedure 7 Statyw_HeatMapPodstawkaSet3	_	_	YES	10 photos 5 sec with the point of fixation

Table 2. Conditions for all experimental procedures employed in the research.

3.   Results

The analysis of results was done according to the Accuracy and Precision measurement methods published by Tobii (2011). According to the methodological guidelines only the results of tests with proper calibration were used in the analysis. The best level of calibration for the GazeFlow software was reached in procedure no. 4 HeatMapWWW1, however, for the SMI RED 250 (henceforth referred to as SMI) device, in procedure no. 8 Glass_2_HeatMap_Set4. The table including the percentage data for calibration accuracy is presented in Appendix 1. Please pay attention to the lower level of the SMI eye tracker calibration, which may have been caused by the difficulties in calibrating this device due to error and delays in displaying the calibration points.

Below is the data for all four webcams used in the research. For each camera the measure of Accuracy and Precision was calculated for GazeFlow and SMI. The tables show a detailed comparison of results. Results of the procedures with free and fixed head are presented separately.
All the results shown are expressed as a degree of deviation between the point on the screen and the eye position.

3.1. The results of the Accuracy measurements for procedures with a freely kept head

The tables show the results of the Accuracy measurements for GazeFlow and SMI in those procedures, where the subjects had a freely kept head, which is a standard approach in researches with a remote eye tracker.

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_0		0,94488479	1,070045699	0,956618552	0,674499497
	Into	1,026403714	1,058630381	0,889618905	0,623168333
	HeatMapSet1	1,039534917	0,984184792	0,879858292	0,598922042
	HeatMapSet2	0,806336957	1,078214609	1,116748739	0,677186783
	HeatMapWww1	0,860038	0,881491905	0,844814238	0,646282333
	ReadText	0,980167529	1,418166941	0,954960118	0,749193588
	Glass_2_Heat Map_Set4	0,84559465	0,99993015	0,9970957	0,72923725
	Glass_3_Heat Map_WWW2	1,084348235	1,161597353	1,023253	0,736735471

Table 3. The comparison of average values of Accuracy on the X and Y axis for cam 1 (CAM_0)
in procedures with a freely kept head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_1		1,09356004	1,176071508	0,939732234	0,662699089
	Into	1,33037145	1,3028553	0,8438387	0,61381745
	HeatMapSet1	1,143760048	1,200061429	0,828500048	0,608644238
	HeatMapSet2	0,948090826	1,153010304	1,026425783	0,646012261
	HeatMapWww 1	0,934483611	1,122464778	0,845122444	0,6821635
	ReadText	1,058331111	1,354267111	1,009833333	0,647873556
	Glass_2_Heat Map_Set4	1,033065222	1,031861722	1,015450833	0,737642556
	Glass_3_Heat Map_WWW2	1,215207133	1,139263533	1,070994	0,724743733

Table 4. The comparison of average values of Accuracy on the X and Y axis for cam 2 (CAM_1) in procedures with a freely kept head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_2		1,031819868	1,156535076	0,955698528	0,667993042
	Into	1,307013762	1,11921319	0,890103286	0,619831429
	HeatMapSet1	0,94959672	1,24239868	0,84591404	0,60782996
	HeatMapSet2	0,841246174	1,242968348	1,116748739	0,677186783
	HeatMapWw w1	0,975540143	0,994712143	0,844814238	0,646282333
	ReadText	1,095747824	1,349947941	0,916929647	0,721256647
	Glass_2_Heat Map_Set4	1,046537905	1,056947857	1,055455429	0,711300238
	Glass_3_Heat Map_WWW2	1,0596775	1,08471	1,0376175	0,727056

Table 5. The comparison of average values of Accuracy on the X and Y axis for cam 3 (CAM_2) in procedures with a freely kept head in GazeFlow (WebCam) and SMI

| Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_3		1,066661694	1,221480898	0,889155694	0,653861592
	Into	1,332723	1,385011182	0,923244818	0,566182
	HeatMapSet1	1,0975465	1,166807167	0,882522583	0,5409855
	HeatMapWww1	0,854962	1,2553629	0,7372276	0,6221437
	ReadText	0,873405143	1,167121714	0,924754	0,725906
	Glass_3_HeatMap_WWW2	1,085828444	1,099141556	0,997456778	0,890734556

Table 6. The comparison of average values of Accuracy on the X and Y axis for cam 4 (CAM_3)
in procedures with a freely kept head in GazeFlow (WebCam) and SMI

The graph shows the summary of the average Accuracies for all cams for both solutions.
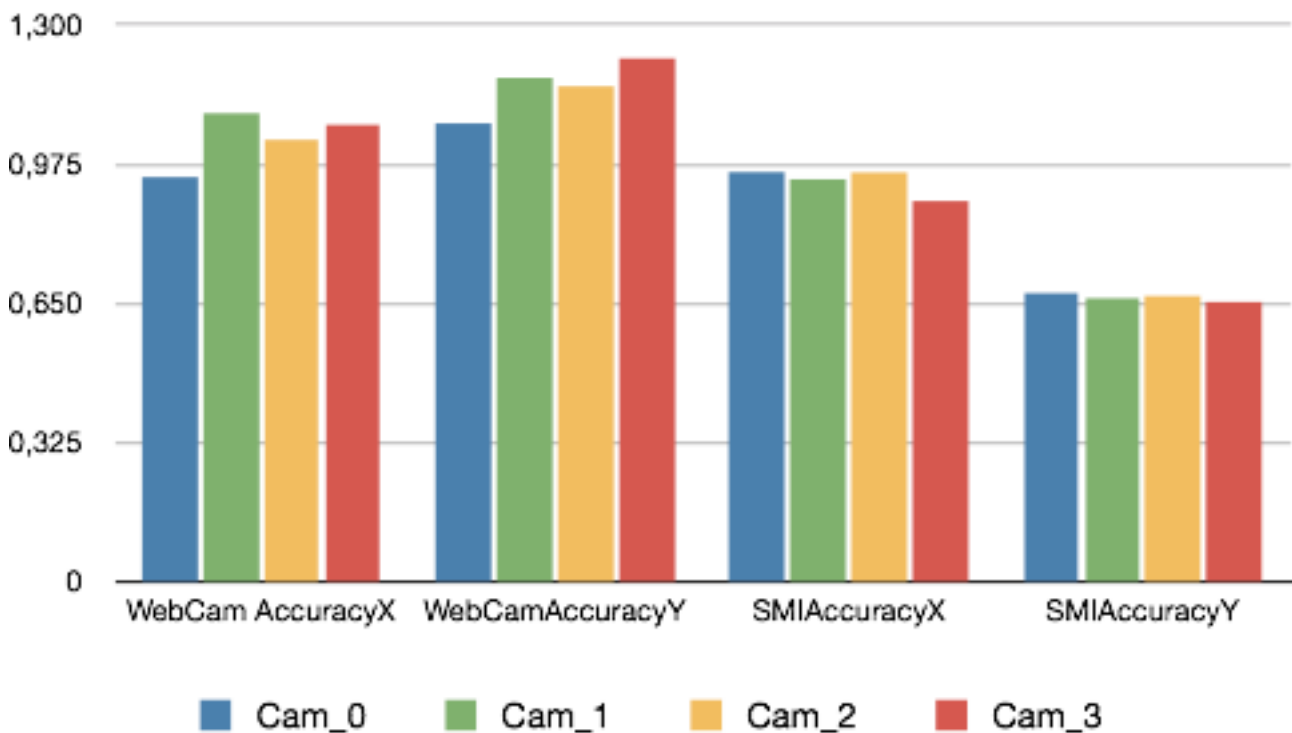
Graph 1. The comparison of the average measurements of Accuracy for conditions where the head was kept freely for GazeFlow (WebCam) vs SMI

Summarizing the test results one has to pay attention to the fact that the results
of the Accuracy measurements are slightly better for the SMI device's the Accuracy

measurements on the X axis (horizontal), whereas significantly better for the Accuracy measurements on the Y axis.

In case of GazeFlow the best Accuracy results were obtained for cam 1 (CAM_0), i.e. the one placed in the central position above the screen. In all cases better Accuracy results were obtained (lower value of deviation expressed in degrees) from cameras placed above the screen.

3.2. The results of the Accuracy measurements for procedures with a head fixed on a chinrest.

The tables gather the results for all cameras and procedures, where the subjects had their heads fixed on a chinrest.

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_0		0,838121	0,99749246	0,96628342	0,78371292
	Statyw_Wali dationColor	1,011000957	1,04781587	0,902361609	0,733661783
	Statyw_Heat MapPodstaw kaSet3	0,690852889	0,95462437	1,020735333	0,826349074

Table 7. The comparison of average values of Accuracy on the X and Y axis for cam 1 (CAM_0) in procedures with a fixed head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_1		1,032267118	1,082964804	0,958015275	0,779070451
	Statyw_Wali dationColor	1,259673667	1,154732667	0,887455208	0,725882
	Statyw_Heat MapPodstaw kaSet3	0,830127963	1,019171148	1,020735333	0,826349074

Table 8. The comparison of average values of Accuracy on the X and Y axis for cam 2 (CAM_1) in procedures with a fixed head in GazeFlow (WebCam) and SMI

| Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_2		0,87373802	1,061464714	0,967781347	0,777237367
	Statyw_Wali dationColor	1,111964	1,122387435	0,906260522	0,731583739
	Statyw_Heat MapPodstaw kaSet3	0,662999654	1,007571538	1,022203615	0,817623269

Table 9. The comparison of average values of Accuracy on the X and Y axis for cam 3 (CAM_2) in procedures with a fixed head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_3		0,924271957	1,02620287	0,900610174	0,713605565
	Statyw_Wali dationColor	1,0330376	1,0555321	0,8393879	0,6650874
	Statyw_Heat MapPodstaw kaSet3	0,840606077	1,003641923	0,947704231	0,750927231

Table 10. The comparison of average values of Accuracy on the X and Y axis for cam 4 (CAM_3) in procedures with a fixed head in GazeFlow (WebCam) and SMI
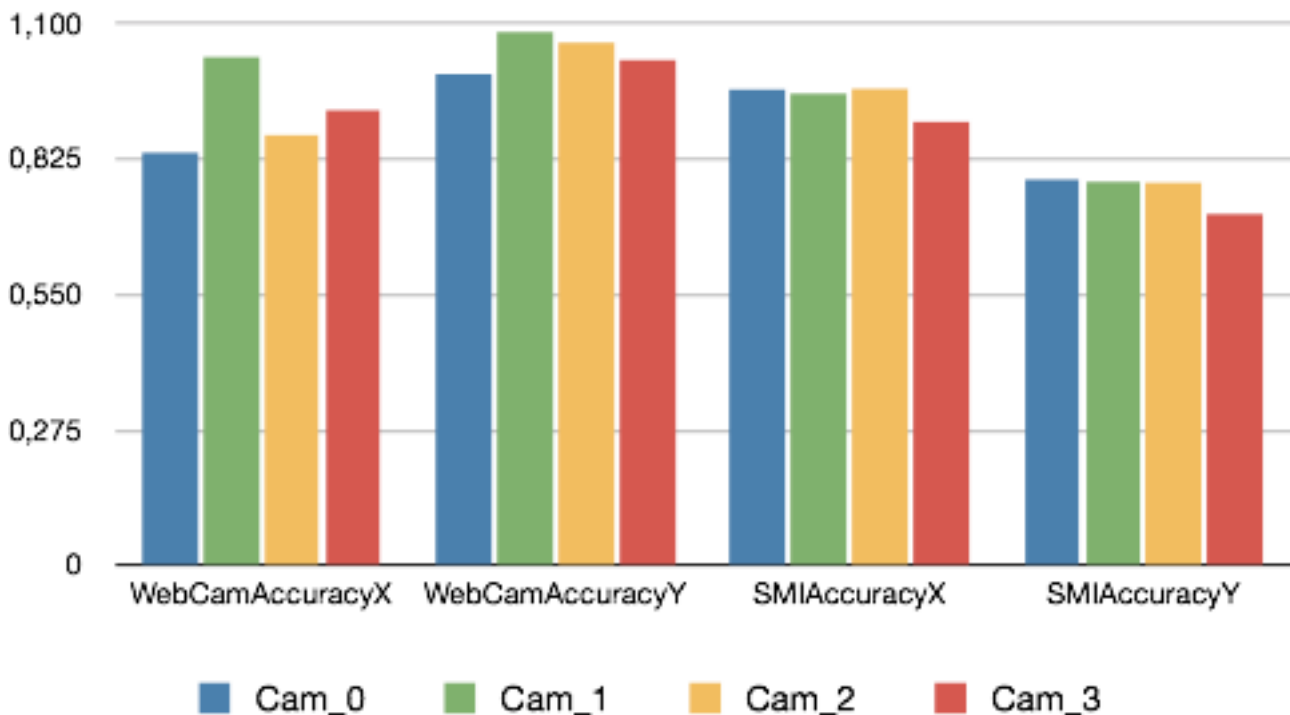
Graph 2. The comparison of the average measurements of Accuracy for all conditions, where the head was fixed for GazeFlow (WebCam) vs SMI

The results in procedures with a fixed head are lower in comparison with procedures with a freely kept head, which indicates a higher accuracy of the solutions.

In case of procedures with a fixed head the results for Accuracy on the X axis in GazeFlow are comparable to the results obtained for the SMI device. It is surprising because one can expect, that in case of the fixed head condition, being close to ideal, the results for a device using infrared light to track eyes should be better.

This may mean a high accuracy of GazeFlow, comparable to SMI under conditions close to ideal, or one should consider the possibility of an artifact stemming from a non-optimal position of the head for conditions of calibration and conducting the test in a position which the subjects were in, and whose heads were placed on chinrests, as well as a significant increase of the angle on the X axis for stimuli presented on peripheries and the position of an eye.

The Accuracy measurement results, i.e. the measurement of the eye position, are higher than the ones indicted as desirable under ideal conditions ($< 0.8°$) for procedures with a free head, especially for GazeFlow, however in procedures with a fixed head the Accuracy measurement in both cases meets the desired value.

3.3. The results of the Precision measurements for procedures with a freely kept head.

The tables below contain the results of the Precision measurements for the GazeFlow software and SMI, in cases where the subjects' heads were freely kept. The results are presented separately for each webcam.

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_0		0,247711692	0,257332406	0,275402364	0,269548643
	Into	0,321631857	0,27725681	0,27531081	0,267714524
	HeatMapSet1	0,244856208	0,267781917	0,275097333	0,273258292
	HeatMapSet2	0,22354787	0,264544348	0,236990609	0,23589013
	HeatMapWww1	0,223061286	0,239552524	0,246522286	0,272730143
	ReadText	0,221118529	0,214340824	0,331550647	0,312985118
	Glass_2_HeatMap_Set4	0,21870405	0,2405562	0,2948133	0,2579744
	Glass_3_HeatMap_WWW2	0,284292353	0,292902	0,284605647	0,278365353

Table 11. The comparison of average values of Precision on the X and Y axis for cam 1 (CAM_0)
in procedures with a free head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_1		0,255058645	0,26032226	0,277273282	0,266271242
	Into	0,3202907	0,3233968	0,27721435	0,2717916
	HeatMapSet1	0,267960571	0,25107749	0,279729238	0,281761762
	HeatMapSet2	0,232992783	0,249868826	0,23701813	0,236692609
	HeatMapWww1	0,251469778	0,248067222	0,259000889	0,283906722
	ReadText	0,207442778	0,189246444	0,353862889	0,268749
	Glass_2_Heat Map_Set4	0,230153889	0,244691111	0,289369667	0,248405111
	Glass_3_Heat Map_WWW2	0,246616067	0,281303067	0,297095533	0,281368067

Table 12. The comparison of average values of Precision on the X and Y axis for cam 2 (CAM_1)
in procedures with a free head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_2		0,239130458	0,262665701	0,279189674	0,270538861
	Into	0,304370286	0,308151286	0,27601381	0,26883619
	HeatMapSet1	0,23742324	0,2546572	0,30429952	0,28784724
	HeatMapSet2	0,207789	0,273790174	0,236990609	0,23589013
	HeatMapWww1	0,221426429	0,23847081	0,246522286	0,272730143
	ReadText	0,211077765	0,203051176	0,315797471	0,300766059
	Glass_2_Heat Map_Set4	0,236553667	0,277771238	0,292944667	0,25489519
	Glass_3_Heat Map_WWW2	0,257648625	0,274757938	0,29071175	0,281076688

Table 13. The comparison of average values of Precision on the X and Y axis for cam 3 (CAM_2)
in procedures with a free head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_3		0,297889673	0,269753941	0,294017408	0,277324408
	Into	0,301578182	0,268378555	0,299079909	0,275885727
	HeatMapSet 1	0,310925917	0,279415333	0,280986083	0,259226917
	HeatMapWw w1	0,3004628	0,2573095	0,2555388	0,2620733
	ReadText	0,244551429	0,244879286	0,340354571	0,307726
	Glass_3_He atMap_WW W2	0,314626111	0,291727222	0,311919	0,296512778

Table 14. The comparison of average values of Precision on the X and Y axis for cam 4 (CAM_3) in procedures with a free head in GazeFlow (WebCam) and SMI

The graph shows the summary of the Precision measurement results of all cameras for both solutions.
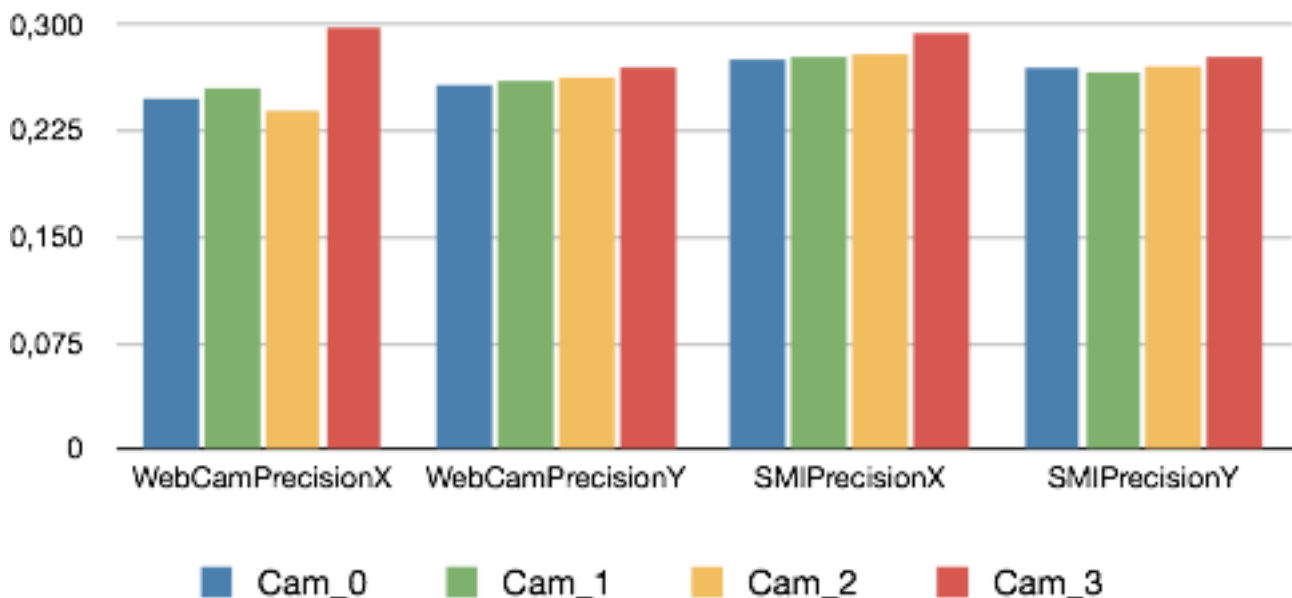
Graph 3. The comparison of the average measurements of Precision for all conditions where the head was free for GazeFlow (WebCam) vs SMI

In case of the Precision measurements, i.e. the repeatability of measurements, comparable results were obtained for both solutions. In precision measurement on the X axis, turning off cam 4 (CAM_3), GazeFlow gets even slightly higher results than SMI. One has to pay attention to the fact that both solutions obtain measurement results indicated as the proper level of precision, even under conditions with a freely kept head. With respect to repeatability of measurements, both solutions meet the criterion of reliability.

3.4. The results of the Precision measurements for procedures with a head fixed on a chinrest.

As it was the case previously the tables present the Precision measurement results for both solutions under conditions, where the head of a subject was placed on a chinrest.

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_0		0,194331558	0,22589244	0,30659356	0,29482526
	Statyw_Wali dationColor	0,188116326	0,207922304	0,325743609	0,298495478
	Statyw_Heat MapPodstaw kaSet3	0,199626015	0,241200333	0,290280556	0,291698778

Table 15. The comparison of average values of Precision on the X and Y axis for cam 1 (CAM_0) in procedures with a fixed head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_1		0,193082251	0,227918227	0,303640804	0,292440588
	Statyw_Wali dationColor	0,195894667	0,225785942	0,318671083	0,293275125
	Statyw_Heat MapPodstaw kaSet3	0,190582326	0,229813593	0,290280556	0,291698778

Table 16. The comparison of average values of Precision on the X and Y axis for cam 2 (CAM_1) in procedures with a fixed head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_2		0,17964051	0,231512714	0,305915571	0,29281851
	Statyw_WalidationColor	0,186365522	0,217215435	0,324216348	0,297404826
	Statyw_HeatMapPodstawkaSet3	0,173691462	0,244160308	0,289726423	0,288761385

Table 17. The comparison of average values of Precision on the X and Y axis for cam 3 (CAM_2) in procedures with a fixed head in GazeFlow (WebCam) and SMI

Camera	Name of the procedure	WebCam average AccuracyX	WebCam average AccuracyY	SMI average AccuracyX	SMI average AccuracyY
Cam_3		0,268014826	0,270949696	0,31937587	0,284119478
	Statyw_WalidationColor	0,2443794	0,2346711	0,3379877	0,2817945
	Statyw_HeatMapPodstawkaSet3	0,286195923	0,298856308	0,305059077	0,285907923

Table 18. The comparison of average values of Precision on the X and Y axis for cam 4 (CAM_3) in procedures with a fixed head in GazeFlow (WebCam) and SMI

The graph shows the total comparison of the average measurements of Precision for the X and Y axis in both devices.
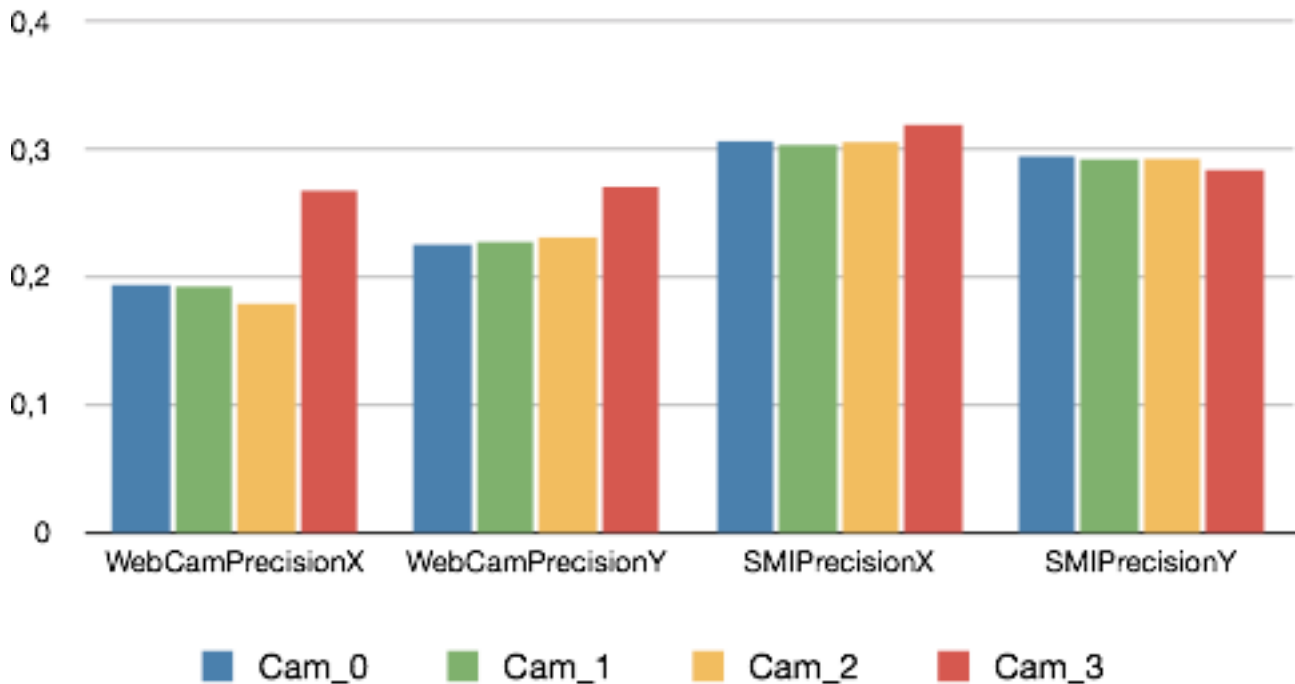
Graph 4. The comparison of the average measurements of Precision for all conditions where the head was fixed for GazeFlow (WebCam) vs SMI

The results for conditions close to ideal, i.e. with a subject's head fixed, indicate a high reliability, especially for the GazeFlow software. This means that under these conditions the software returns repeatable measurements even to a greater degree than the SMI eye tracker.

4.   Conclusions

The results obtained in the Accuracy and Precision tests show high capability of making reliable and accurate measurements with the GazeFlow software which tracks eyesight on the basis of image from webcams. The software's Accuracy measurements ($< 0.9°$-$1.0°$) are, especially under conditions where the head is kept freely, higher than indicated in the literature as the desired level of Accuracy, however they are comparable to the device using infrared light. In procedures with a fixed head the Accuracy measurement meets the desired value.

On the other hand, the repeatability (Precision measurement) of results meets all criteria necessary to consider the software as reliable when pitched against devices employing other solutions, in this case infrared eye trackers.

One has to pay attention to the fact that different results were obtained for different types of cameras and their arrangement. The recommended position of a camera, for which more accurate and precise results were obtained, is above the screen.

The GazeFlow software can be successfully used in marketing and website researches, where one has to indicate the level of obtained Accuracy of the eye tracker.
The capabilities of the software in areas of accuracy and precision allow for its successful use as a tool for controlling computers through eyesight and application in entertainment or rehabilitation. To this extent the software is ready for commercialization.

5.   References

Tobii Technology, *Accuracy and precision test method fo remote eye trackers. Test Specification* Vertion: 2.1.1. February, 2011

6.   The list of appendices

Table 1. GazeFlow's validation of calibration

Table 2. SMI's validation of calibration

3. Heatmaps

4. Scanpath

Appendices

Table 1. GazeFlow's validation of calibration

Procedure	GazeFlow % of incorrect calibration	GazeFlow % of correct calibration
1_Into	24%	76%
2_WalidationHeadForce	20%	80%
3_HeatMapSet1	23%	77%
4_HeatMapSet2	7%	93%
5_HeatMapWww1	22%	78%
6_ReadText	40%	60%
7_Statyw_WalidationColor	16%	84%
9_Statyw_HeatMapPodstawkaSet3	10%	90%
Glass_2_HeatMap_Set4	13%	87%
Glass_3_HeatMap_WWW2	14%	86%
Suma	20%	80%

Table 2. SMI's validation of calibration

Procedure	SMI % of incorrect calibration	SMI % of correct calibration
1_Into	38%	62%
2_WalidationHeadForce	100%	0%
3_HeatMapSet1	29%	71%
4_HeatMapSet2	17%	83%
5_HeatMapWww1	33%	67%
6_ReadText	23%	77%
7_Statyw_WalidationColor	23%	77%
9_Statyw_HeatMapPodstawkaSet3	15%	85%
Glass_2_HeatMap_Set4	9%	91%
Glass_3_HeatMap_WWW2	30%	70%
Suma	31%	69%