GSET Somi: A Game-Specific Eye Tracking Dataset for Somi

Hamed Ahmadi[1], Saman Zad Tootaghaj[1], Sajad Mowlaei[1], Mahmoud Reza Hashemi[1], Shervin Shirmohammadi[1,2]

[1]*Multimedia Processing Laboratory (MPL), School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran*

[ha.ahmadi | s.tootaghaj | smowlaei | rhashemi | sshirmohammadi]@ut.ac.ir

[2]*Distributed and Collaborative Virtual Environments Research Laboratory (DISCOVER Lab), School of Electrical Engineering and Computer Science, University of Ottawa, Ottawa, Canada*

shervin@discover.uottawa.ca

ABSTRACT
In this paper, we present an eye tracking dataset of computer game players who played the side-scrolling cloud game Somi. The game was streamed in the form of video from the cloud to the player. This dataset can be used for designing and testing game-specific visual attention models. The source code of the game is also available to facilitate further modifications and adjustments. For collecting this data, male and female candidates were asked to play the game in front of a remote eye-tracking device. For each player, we recorded gaze points, video frames of the gameplay, and mouse and keyboard commands. For each video frame, a list of its game objects with their locations and sizes was also recorded. This data, synchronized with eye-tracking data, allows one to calculate the amount of attention that each object or group of objects draw from each player. As a benchmark, we also show various attention patterns could be identified among players.

CCS Concepts
• Computing methodologies →Artificial intelligence→ Computer vision→ Computer vision problems→ Interest point and salient region detections

Keywords
Eye-tracking dataset; visual attention model; perceptual video coding, game dataset

1. INTRODUCTION
Eye-tracking datasets present recordings of the users' gaze and visual focus while those users view visual objects such as images, videos, games, etc. One of the main applications of such datasets is the development of visual attention models. Such models are quite useful in fine-tuning a specific application to provide a higher quality of experience for its users. In other words, by knowing a priori which region(s) in the visual subject the user will likely be paying more attention to at a given time, the application can be configured to provide a higher quality for that region, potentially leading to a higher quality of experience for the user.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
MMSys'16, May 10-13, 2016, Klagenfurt, Austria
© 2016 ACM. ISBN 978-1-4503-4297-1/16/05…$15.00.
DOI: http://dx.doi.org/10.1145/2910017.2910616

Naturally, the development of such visual attention models is very application-specific, because the specific tasks of an application drives the attention of its users [1, 2] and since each application has its own distinct required tasks, it requires a dedicated visual attention model. As an intuitive example, while users of a video conferencing application are likely to pay attention to the speaker's face during a conference, the same users in a video broadcasting application of a volleyball match are likely to pay more attention to the ball and to the players' actions rather than their faces. In other words, even though both of the above examples are video streaming, their visual attention models are quite different. As such, visual attention models must be developed for specific applications.

The application we are targeting in this paper is gaming, specifically side-scrolling games, represented by the game Somi. Our dataset can be used by researchers and practitioners to develop visual attention models specific to this game and its genre. Such visual models can then be used for a variety of purposes. For example, they can be used in Cloud Gaming (CG), which allows players to play games streamed in the form of video from the cloud over the Internet to their devices, instead of buying discs or waiting for lengthy downloads [3]. One of the challenges of CG is its heavy bandwidth requirement needed to create the necessary high-quality gaming experience that players expect. It is known that CG requires a network connection with an average bandwidth of 5Mbps per player to provide interactive gaming services with a resolution of 720p at 30fps [3]. To reduce bandwidth, visual attention models can be used to identify regions of interest in a game frame and to encode those regions with higher quality in the video, while encoding other regions with lower quality, with successful and effective results [4-6]. Visual attention models can also be used to determine where game players might look at. Such a priori knowledge enables game specific systems, such as cloud gaming, serious game design [7] and level difficulty adjustment [8], to optimize their performance. For instance, in [9], game objects' positioning was done based on visual attention in order to adjust the difficulty of the game. This is important because improper difficulty level provokes anxiety in a discouragingly hard game or apathy in a boringly easy game [10].

Our dataset can also be used to better understand game players' cognitive abilities such as eye-hand coordination, decision making, and following directions in the context of side-scrolling games. Recently, a similar research [11] has studied the eye-hand coordination skill among gamers of the First Person Shooter (FPS) video game genre.

In this paper, we present our dataset of game players' eye tracking of the side-scrolling game Somi, plus a benchmark of various attention patterns that can be identified among players. We hope that by releasing this dataset and its benchmark, other researchers can use them to design their own visual attention models and to advance both the science and technology in this area. It should be noted that since each game has its own unique game logic and design, a model developed based on Somi alone cannot necessarily be used for other games. However, the presented dataset and benchmark make significant contributions by fulfilling the following important roles: 1- Providing a more realistic and modern dataset compared to existing game-specific eye tracking datasets, as shown in Table 1. Specifically our dataset provides all at once: HD resolution, gameplay instead of game watching, recording of mouse and keyboard inputs, recording of game objects' locations, losslessly recorded videos, and a large number of subjects; 2- Advancing the body of knowledge for the understanding and development of game-specific visual attention models; 3- Providing a collection methodology and benchmark that can be applied to other games in the side-scrolling genre, allowing future investigations which might reveal some common attention patterns among video games of this genre, since most of them have a similar game logic and design as Somi.

1.1 Related Work

Visual attention has already been used in video compression. A complete review of visual attention based video compression approaches is provided in [12] and [13]. Such approaches rely on two key aspects: accuracy of the visual attention model in predicting gaze points, and the efficiency of incorporating the model into the video encoder. But such approaches are not efficient for game videos. On the one hand, misprediction of gaze points has two undesired consequences. First, the player's region-of-interest (ROI) is encoded at lower quality and hence his/her perceived quality is damaged. Second, non-ROI regions are encoded at high quality causing bandwidth to be squandered. On the other hand, even assuming that the model is one hundred percent accurate, not all methods of incorporating the model into the video encoder result in the same coding efficiency.

Similarly, numerous eye-tracking datasets have been contributed to the community [14], but they are hardly suitable for the game context. Therefore, we conducted an eye-tracking experiment, recording participants' gaze data while playing a game titled "My Beautiful Doll, Somi". Table 1 compares the existing game-related eye-tracking datasets with ours.

As can be seen, the eye-tracking data in CRCNS [15] and DIEM [16] was collected during watching rather than playing. However, during gameplay players pay attention to the regions which are important for the accomplishment of their current activity, so gaze points of a player during gameplay could be significantly different from merely watching the game. Furthermore, since it is almost impossible for a participant to keep his/her head fixed while playing a game, using a remote eye-tracker, which is unobtrusively installed in front of the display and remotely tracks the subject's eyes, is more realistic than using a chin rest or head mount eye tracker which restrains the subject's movements. Another point to mention is the resolution of the videos in [17] and [18]. Considering that current cloud gaming companies offer HD resolution, coupled with the observation that the performance

Table 1. Comparison of the game-related eye-tracking datasets. ● and ○ symbols indicate existence, and non-existence, respectively.

	GSET Somi	[17]	[18]	CRCNS [12]	DIEM [13]
Collected while playing	●	●	●	○	○
Collected while watching	○	○	○	●	●
Game video	●	●	●	●	○
Game video trailer	○	○	○	○	●
#Subjects	84	5	21	8	-
#Videos	135	24	27	-	4
Resolution	720p	680x480	680x480	680x480	Varying
Video format	Raw	Raw	H.264/AVC	MPEG-1	-
Eyes	Both	Right	-	-	-
Eye-tracker	Remote	Chin rest	Chin rest + Head mount	Chin rest	-

Figure 1. A snapshot of Somi.

of the same perceptual video compressor might differ on SD and HD resolutions [12], there is a need for a dataset containing HD resolution game videos. Furthermore, game videos must be recorded in a lossless manner to provide a fair basis for comparing the performance of the attention based video compression methods. Finally, larger number of participants allows researchers to investigate the difference among attention patterns of different game players. The above are the main reasons we created our own dataset called GSET Somi: Game-Specific Eye Tracking for Somi.

The rest of this paper is organized as follows. The next section explains the data collection methodology including the information about video game, eye-tracking device, environment, and participants. Section 3 presents the results of our preliminary analyses on the collected data which reveals the influence of some new factors, such as skill level, on visual attention patterns. Section 4 gives the information about dataset availability to be used by other researchers and its copyright. Finally, the paper is concluded in section 5.

2. DATA COLLECTION PROCEDURE

2.1 Video Game

We conducted our own eye-tracking experiment by recording participants' gaze data while playing a side-scrolling game, titled "My Beautiful Doll, Somi". This game is about Somi, a doll, which is in love with her owner Sara. Somi has to confront her enemies and collect love symbols to prove her love to Sara. This game is a combination of "Rail Shooters" and "Scrolling Shooters" game genres. Somi has been developed by means of "GameMaker: Studio" game engine. It includes seven main levels plus several bonus levels and can be configured to run at 720p resolution. Figure 1 shows a screenshot from Somi.

There are three kinds of enemies in this game. The player is responsible for controlling Somi to aim and shoot at the enemies, jump over obstacles and collect hearts and shells, obtaining scores in the process. The player is allowed to opt his/her gun among three available guns: Uzi, Shotgun and Bazooka. Powerful guns score more. However, the more destructive a gun is, the more it costs. So in order to obtain a higher score, the player needs to find a balance between cost and earning.

2.1.1 Object Information

Each level of Somi has a set of game objects. Figure 2 shows the game objects in the first level of Somi, excluding the background. These objects enter the scene, move around and exit during the gameplay based on the game logic and user inputs. Therefore, developing object-based visual attention models which require matching gaze points with game objects would be challenging. In order to develop such model, we implemented a networking interface for the game video which sends the size and location of game objects in each frame over a TCP or UDP connection. More specifically, it listens on a configurable port number and whenever a client connects, it starts sending object information to that client. For the sake of efficient bandwidth utilization, it avoids sending information of fixed objects such as the health indicator. This information is stored in a separate file and must not be neglected later in object-based analysis.

In our design, there are three kinds of data packets sent over the connection: frame packet, object packet, and state packet. These packet types start with ‘F’, ‘O’, and ‘S’, respectively. At the start of each frame, a frame packet including the frame number is sent over the connection. ‘F’ packets can be used to align the eye tracker data with the video frames. The first frame with eye-tracking data is the one captured just after the player clicks the ‘Start’ button located on the game screen. For each object rendered in that frame, an object packet is then sent. The object packet includes an encoded string like $OY_1Y_2Y_3X_1X_2X_3X_4H_1H_2H_3W_1W_2W_3W_4$, where $Y_1Y_2Y_3$ and $X_1X_2X_3X_4$ are the top and left distance of the object from the top left corner of the screen, respectively. $H_1H_2H_3$ and $W_1W_2W_3W_4$ are also the height and width of the object. It should be noted that the size allocated to these properties must match the resolution in which the video game runs. For example, a resolution of 1280x720 (HD 720) requires four and three characters for left and top properties, respectively. If the size of an object does not change during its appearance in the scene, the $H_1H_2H_3W_1W_2W_3W_4$ part can be omitted. At the start of each frame, the game is in one or more specific states. For each state, a state packet including the state identifier is sent over the connection. These packets are only sent if the game states have changed since the last transmission of the state packets.

2.2 Eye Tracker Device

In the real world, players can move their heads freely while playing. Since it is unrealistic or even impossible to ask a player to keep his head fixed while playing a game, we did not use any chin rest trackers. Neither did we use heavy head-mounted eye-trackers. Instead, we carried out the experiment via a Tobii X2-30 Compact Eye Tracker system with a sampling rate of 30 Hz [19]. This system has a large head movement box that allows a test participant to move his/her head freely and naturally during a test

Table 2. Demographic profile of the participants

Gaming Experience				
Bad	Poor	Good	Fair	Excellent
23.08%	24.36%	28.21%	16.67%	7.69%
Monthly Game Play				
<= 5	6 – 10	11 - 20	21 - 30	> 30
8.97%	48.72%	25.64%	12.82%	3.85%
Gaming Platform (Already Played on)				
PC	Console	Tablet	Mobile Phone	
80.77%	23.08%	35.9%	73.08%	
Genres (Already Played)				
Running	Side Scrolling	Sports		
56.41%	42.31%	71.79%		
Endless Platform	Shooter	Shoot'em Up		
42.31%	67.95%	32.05%		

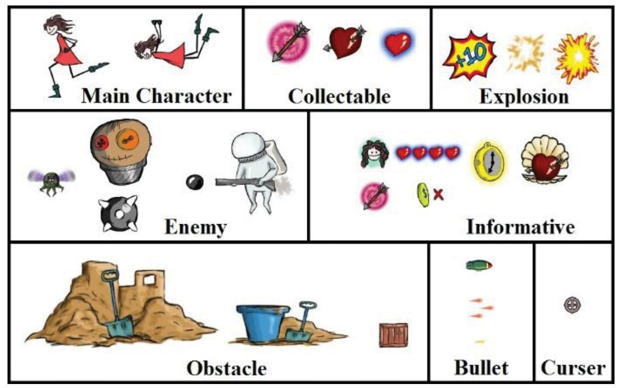

Figure 2. Somi’s game objects in the first level.

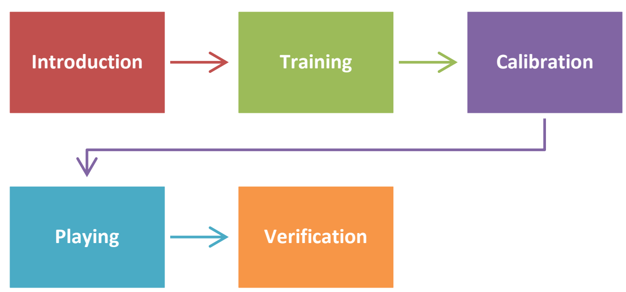

Figure 3. The five steps of eye-tracking data collection.

session. The device’s accuracy is reported to be 0.4° on average at a 60 cm distance with 9-point calibration [20].

2.3 Environment

The experiments were conducted under ambient light, i.e., sources of light that are already available naturally (e.g. the sun light through windows) or artificial light (e.g. normal electric lighting of a room). Participants sat in front of a 20” LED Samsung monitor at an average distance of 60 cm.

2.4 Participants

Table 2 shows the demographic profile of the participants. About eighty subjects, ranging from 19 to 30 years old, participated in our experiment. Approximately, twenty three percent of them were female. The bad, poor, good, fair and excellent are the users’ self-reported score about their gaming experience in general for all the games they had ever played. Each person played at least one session, sometimes two or three, leading to a total of 135 sessions.

2.5 Data Collection Procedure

Each session comprised five steps: Introduction, Training, Calibration, Playing, and Verification. Figure 3 shows the order of these steps. The calibration and playing steps together took seven minutes on average, while all the other steps together took ten minutes. The subjects who played more than once skipped the introduction and training steps in their next sessions. Let us know go over the details of each step.

In the first step, we gave participants a brief introduction to the game. We taught them how to jump, fire and change their guns. We also showed them several screenshots highlighting the enemies, obstacles and collectable items in the game. In the second step, we asked them to freely play the game, providing them an opportunity to get to know it. The third step was to run a 3x3 point grid calibration. This procedure has been recommended by the eye-tracker's manufacturer and helps it to perform more accurately. After the calibration step, the game screen was shown to the participants and they started playing while their gaze data plus their mouse and keyboard strokes were being recorded. Meanwhile, the game video frames were captured by means of a third-party application configured to store the video using lossless compression. The duration of playing was fixed to three minutes for all the participants. In the last step, for each participant, we determined whether the gaze data was valid or not. A participant's gaze data was considered valid, if more than ninety percent of the gaze records were valid. A gaze record would be reported as valid by the eye tracker device, if it were able to detect both eyes simultaneously. When the device was unable to detect either of the eyes, it would report that eye's gaze report as invalid.

3. SAMPLE RESULTS AND BENCHMARKS

In this section, we show three instances of how our dataset can be used, producing benchmark results that can be used for other researchers. Please note that we will not present the full design and development of a visual attention model from this dataset, because that would be another full paper of its own and would not fit here.

First, we postulate that game-specific visual attention models not only should take game logic and design into the account, but also should consider players' skill levels as an influencing factor on attention patterns. To provide evidence for this hypothesis, we first measure the amount of attention per game object for each player. To do so, we calculate the percentage of gaze points which land in the bounding box of each object. Note that we expand the bounding boxes by 16 pixels in all four directions, to account for the precision of the utilized eye tracker. Then, we determine players' skills based on their game scores, categorize them into three groups of beginners, intermediates, and experts and calculate the average amount of attention per object in each group of skill level. Table 3 shows the score range for each group. It should be noted that determining players' skills based on their gained scores is a valid process, because it has been shown that game score is influenced by factors such as player's skill, network latency, and network jitter, with skill being the most influential factor in game score across the great majority of genres, possibly all [21].

Table 3. Participants were categorized into three groups

Skill Level	Score Range
Beginner	score <= 1000
Intermediate	1000 < score <= 6000
Expert	6000 < score

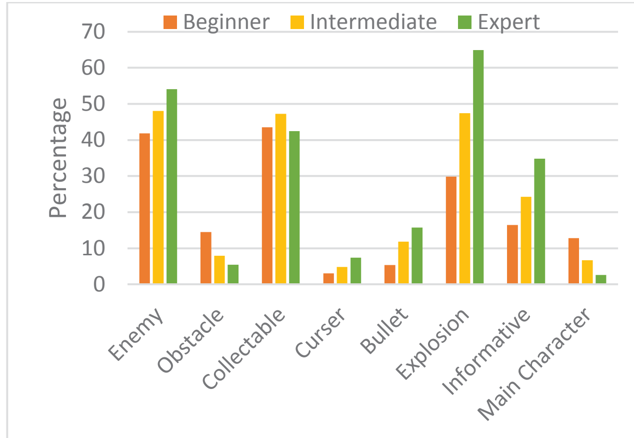

Figure 4. Average attention per category.

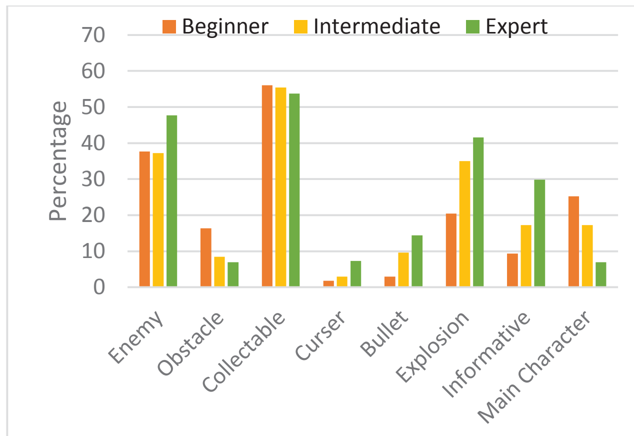

Figure 5. Average attention per category in Jumping state.

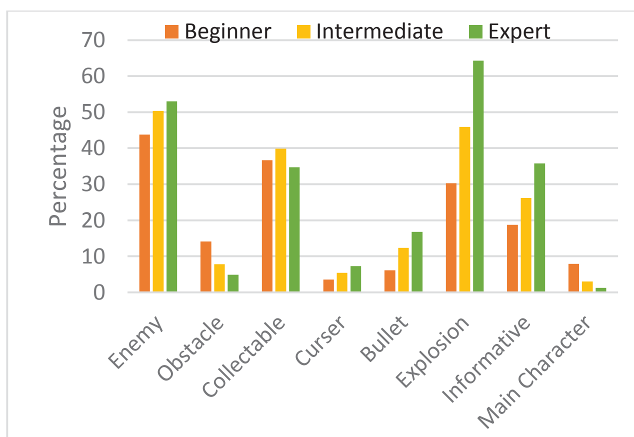

Figure 6. Average attention per category in Running state.

Since in our experiments not only delay and jitter but also hardware and software are the same for all players, we can confidently depend on player's skill as the only factor affecting the score.

Second, in order to have a concise view of the gathered data, we group game objects into the eight categories as shown in Figure 2 and measure the amount of attention per category. Figure 4 illustrates the average attention per category for each skill level. This figure shows that a player's skill level significantly affects his attention behavior. As an instance, we see that experts paid approximately two times more attention to informative game objects as what beginners did. Intuitively, this behavior can be justified by the fact that experts tend to maximize their scores by tracking game events through reading informative objects. This figure not only shows the attention pattern for each skill level, but also reveals the shift in the players' attention as their skill improves. For example, as players get accustomed to the game environment, they manage to jump over obstacles without having to look at them, allowing them to score more by either killing more enemies or collecting more items. In fact, the more they play, the better they estimate when an appeared obstacle reaches them. Therefore, players with higher skills simply keep track of time in their mind and press the jump key after the estimated time has elapsed. As can be seen, Figure 4 vividly highlights this attention shift in obstacle category as the player's skill level increases.

Third, we analyze the attention patterns in different game states. To illustrate this, we consider two states for Somi: Jumping and Running states which are mutually exclusive. Figure 5 and Figure 6 chart the average attention per category in Jumping and Running states, respectively. As can be seen, the attention patterns differ in these states. For example, players pay more attention to the main character while jumping. This behavior is vividly more likely amongst beginners. As another instance, we see that collectable items draw more attention in the Jumping state compared to other states. In addition to the game logic, this pattern is partially due to the fact that in this game, collectable items usually appear on high altitudes. So game design coupled with game logic affect where players look. Inversely, by analyzing attention patterns, game designers can learn how to guide a player's attention to a specific game target. For example, consider an adventure game in which the player engages in an interactive story driven by exploration and puzzle-solving. In such a game, if the player does not find the decoding messages within some tolerable time to solve the game puzzles, s/he will give up and leave the game, something that is not desirable for the game company. But, they can simply avoid this kind of customer churn by bringing the decoding messages in the attention zone of a player based on the player's skill level.

It should be noted that the number of skill levels and score boundaries as well as the object categories and game states that have been used here are not necessarily the optimal choices. We have used 3 skill levels just as an example. Optimizing these factors in order to obtain an accurate game-specific game attention model is something that the presented dataset can be used for in future research.

4. AVAILABILITY AND FORMAT

The dataset and the Somi video game are available for free as long as they are used for non-commercial and academic purposes. Instructions on how to obtain them are posted at our website[1]. At the time of writing this paper, the size of the dataset is about one Terabyte.

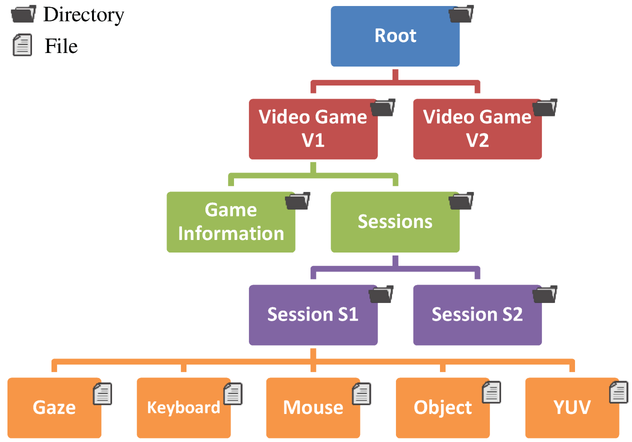

Figure 7. The Directory Structure of GSET.

The source code of the companion video game is also available for download. In order to modify it, a "GameMaker Studio" development license is required.

4.1 Directory Structure

For each video game, there is a specific subdirectory in the root directory of GSET with the same name as that video game. Therefore, currently GSET contains a subdirectory, named Somi, which includes all the gathered information regarding the video game Somi. Figure 7 shows the hierarchical structure of GSET. In this figure, files and directories have been differentiated by two distinct symbolic icons. The video game subdirectory itself has two subdirectories: Game Information and Sessions. The former includes the common information among all eye-tracking sessions of the video game such as the size and location of the fixed game objects in the recorded part of the video game which have been stored in .csv file format. Table 4 shows what columns each csv file in this dataset contains. The latter contains a subdirectory per each eye-tracking session. Eye-tracking information, mouse and keyboard events, size and location of the dynamic game objects in the video game and an approximately 5000-frame YUV video file are stored in separate files. Some sample and significantly shortened (120 frames) videos are also available for a pre-trial of the dataset. The offset of these clipped frames from the start of the video has been stored in a separate file in the session subdirectory. To obtain the whole video frames, please follow the instructions on the aforementioned website. The session subdirectory also contains a file including the demographics of the participant in that session.

5. CONCLUSION

In order to facilitate the development of more reliable and accurate game-specific visual attention models, we present in this paper our eye tracking dataset for the game Somi. Compared to existing datasets, ours has the following features at once: HD resolution, collection during gameplay instead of watching the game videos, recording of mouse and keyboard inputs, recording of game objects' locations, and a large number of subjects. The dataset can be used to recognize the different attention patterns among players and to come up with novel visual attention models.

[1] http://www.eecs.uottawa.ca/~shervin/gaze

Table 4. Columns of the CSV files

File	Column Name	Description
Gaze.csv	FrameNumber	The number of the frame
	Synchronized°	Whether SDK and eye tracker synchronized
	LocalTime°	Eye tracker clock
	ConvertedTime°	SDK clock
	LeftGazePoint2D.X	The left eye's position from left
	LeftGazePoint2D.Y	The left eye's position from top
	RightGazePoint2D.X	The right eye's position from left
	RightGazePoint2D.Y	The right eye's position from top
	LeftPupilDiameter	Diameter of the left eye's pupil
	RightPupilDiameter	Diameter of the right eye's pupil
	LeftValidity°	Whether left eye is found
	RightValidity°	Whether right eye is found
Keyboard.csv	FrameNumber	The number of the frame
	Type	Keyboard event (Down, Up, …)
	KeyCode	The code of the key
	Shift	Whether Shift key is held down
	Alt	Whether Alt key is held down
	Ctrl	Whether Ctrl key is held down
Mouse.csv	FrameNumber	The number of the frame
	Type	Mouse event (Down, Up, …)
	Button	Mouse button
	X	Mouse position from left
	Y	Mouse position from top
	Delta	Mouse movement
Object.csv	FrameNumber	The number of the frame
	Object.Code	The code number of the object
	Object.X	Object position from left
	Object.Y	Object position from top
	Object.H*	Height of the object
	Object.W*	Width of the object

* Reported only for an object whose size changes during the gameplay
° For more information refer to [22]

6. REFERENCES

[1] A. Borji and L. Itti, "Defending Yarbus: Eye movements reveal observers' task," *Journal of vision,* vol. 14, p. 29, 2014.

[2] A. L. Yarbus, *Eye movements during perception of complex objects*: Springer, 1967.

[3] S. Shirmohammadi, M. Abdalla, D. T. Ahmed, K.-T. Chen aka ShengWei Chen, Y. Lu, and A. Snyatkov, "Introduction to the Special Section on Visual Computing in the Cloud: Cloud Gaming and Virtualization," *Circuits and Systems for Video Technology, IEEE Transactions on,* vol. 25, pp. 1955-1959, 2015.

[4] H. Ahmadi, S. Z. Tootaghaj, M. R. Hashemi, and S. Shirmohammadi, "A game attention model for efficient bit rate allocation in cloud gaming," *Multimedia Systems,* vol. 20, pp. 485–501, 2014.

[5] H. Ahmadi, S. Khoshnood, M. R. Hashemi, and S. Shirmohammadi, "Efficient bitrate reduction using a Game Attention Model in cloud gaming," in *Haptic Audio Visual Environments and Games (HAVE), 2013 IEEE International Symposium on*, 2013, pp. 103-108.

[6] M. Hemmati, A. Javadtalab, A. A. N. Shirehjini, S. Shirmohammadi, and T. Arici, "Game as Video: Bit Rate Reduction through Adaptive Object Encoding," presented at the NOSSDAV'13, Oslo, Norway, 2013.

[7] Z. Yusoff and A. Kamsin, "Game Rhetoric: Interaction Design Model of Persuasive Learning for Serious Games," in *Learning and Collaboration Technologies*, ed: Springer, 2015, pp. 644-654.

[8] L. Jie and J. J. Clark, "Video game design using an eye-movement-dependent model of visual attention," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM),* vol. 4, p. 22, 2008.

[9] G. A. Koulieris, G. Drettakis, D. Cunningham, and K. Mania, "An automated high-level saliency predictor for smart game balancing," *ACM Transactions on Applied Perception (TAP),* vol. 11, p. 17, 2014.

[10] D. Johnson and J. Wiles, "Effective affective user interface design in games," *Ergonomics,* vol. 46, pp. 1332-1345, 2003.

[11] D. G. Gozli, D. Bavelier, and J. Pratt, "The effect of action video game playing on sensorimotor learning: evidence from a movement tracking task," *Human movement science,* vol. 38, pp. 152-162, 2014.

[12] J.-S. Lee and T. Ebrahimi, "Perceptual video compression: A survey," *Selected Topics in Signal Processing, IEEE Journal of,* vol. 6, pp. 684-697, 2012.

[13] Z. Chen, W. Lin, and K. N. Ngan, "Perceptual video coding: challenges and approaches," in *Multimedia and Expo (ICME), 2010 IEEE International Conference on*, 2010, pp. 784-789.

[14] S. Winkler and S. Ramanathan, "Overview of Eye tracking Datasets," in *QoMEX*, 2013, pp. 212-217.

[15] J. L. Teeters, K. D. Harris, K. J. Millman, B. A. Olshausen, and F. T. Sommer, "Data sharing for computational neuroscience," *Neuroinformatics,* vol. 6, pp. 47-55, 2008.

[16] P. K. Mital, T. J. Smith, R. L. Hill, and J. M. Henderson, "Clustering of gaze during dynamic scene viewing is predicted by motion," *Cognitive Computation,* vol. 3, pp. 5-24, 2011.

[17] R. J. Peters and L. Itti, "Beyond bottom-up: Incorporating task-dependent influences into a computational model of spatial attention," in *Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on*, 2007, pp. 1-8.

[18] A. Borji, D. N. Sihite, and L. Itti, "What/where to look next? Modeling top-down visual attention in complex interactive environments," *Systems, Man, and Cybernetics: Systems, IEEE Transactions on,* vol. 44, pp. 523-538, 2014.

[19] (2015/08/08). *Tobii X2 Eye Tracker Portable Lab* [Online]. Available: http://www.tobii.com/en/eye-tracking-research/global/products/hardware/tobii-x2-30-eye-tracker/

[20] (2015/08/08). *Tobii X2-30 Eye Tracker Accuracy and Precision Test Report* [Online]. Available: http://www.tobii.com/Global/Analysis/Training/Metrics/Tobii%20X2-30%20Eye%20Tracker%20Accuracy%20and%20Precision%20Test%20Report.pdf

[21] M. Dick, O. Wellnitz, and L. Wolf, "Analysis of factors affecting players' performance and perception in multiplayer games," in *Proceedings of 4th ACM SIGCOMM workshop on Network and system support for games*, 2005, pp. 1-7.

[22] (2013, 2015/12/30). *Tobii Analytics SDK Developers Guide* [Online]. Available: http://www.acuity-ets.com/downloads/Tobii%20Analytics%20SDK%20Developers%20Guide.pdf