Eye-Gaze Control of Virtual Agents Compensating Mona Lisa Effect

Hsueh-Han Wu[1], Hironori Mitake[1], Shoichi Hasegawa[1]

[1] Information and Communications Engineering, Tokyo Institute of Technology

Abstract: The uses of digital signage with a virtual agent have gradually increased nowadays. Like real people do eye-contact to keep others' attention or to have communication, we also expect the virtual agent to exert the functions of gaze on the viewers. However, the 2D display will lead to the inaccurate visual perception of gaze direction. This phenomenon is referred to as the "Mona Lisa" effect, which the gaze of a portrait appears to follow observers no matter what viewing angle they stand.

Our goal is to design a system that the virtual agent makes eye contact with observers in the necessary timing. We consider that there are mainly 2 situations while we look at the virtual agents. One is standing still to look, and another is looking while passing by the display. We conducted the experiments in these 2 situations separately. By controlling iris position and head rotation of the stimuli, we have participants to judge whether they feel the eye contact with the virtual agent.

Keywords: Gaze perception, Mona Lisa gaze effect, Multiparty communication, Situated Interaction.

1. Introduction

Due to the demand for development of digital signage by companies, the global digital signage market keeps growing recently[1]. In addition, among various purposes of digital signage, digital signage with virtual agents have been gradually developed. They serve as virtual concierges to provide guide information or attract customers' attention. In other cases, they are used as the advertisement of certain character or mascot promotion. With the applications of virtual agents, they offer customers more engaging experience, which not only allow customers to retrieve new information easily but also can drive the increasing tech-savvy shoppers to purchase decisions[2].

To allow virtual agents have realistic social interaction with people, not only verbal communications but non-vocal expression such as gestures, eye-contact, and facial expressions plays an important role in affecting the interaction. In this research, we focus on the eye gaze cues and try to solve the problems that happens in designing the eye gaze patterns. According to Kleinke's research review[3], eye gaze includes the functions as follows: (a) provide information, (b) regulate interaction, (c) express intimacy, (d) exercise social control, (e) facilitate service and task goals. It is well-established that eye gaze is an influential non-verbal cue and performs various functions in the face to face communication. Nevertheless, when we use a 2-D flat display to show a virtual agent, the visual illusion, Mona Lisa effect, may appear and severely influence us on perceiving the correct eye gaze direction. It is a perceptual phenomenon that when we look at a portrait, no matter on which vantage points we stand, the looker in the portrait is always staring at us. On the other hand, when the looker changes the eye gaze direction, it would be hard for us to find an suitable vantage point to reach the eye-contact. This inaccurate eye gaze perception will interfere with the communication between we humans and the virtual agents.

Regarding the visual illusion that happens in the 2D pictures, several studies accounted for its consequences with different perspectives. Maruyuma et al.[4] carried experiments with various stimuli to confirm the existence of the Mona Lisa effect. Besides facial pictures, they also used schematic face picture with no other facial parts. They revealed the importance of the face frame for a portrait in the experiment. If the face frame is taken out, the Mona Lisa effect did not appear. Goldsteins[5] investigated how pictures in an angle is perceived. The spatial layout of 3-D array pictures remain relatively constant as viewing angle changes whereas judgements of their

*連絡先: 東京工業大学工学院情報通信系
〒226-8503 神奈川県横浜市緑区長津田町 4259 R2-20 長谷川晶一研究室
E-mail: wu.h.ae@m.titech.ac.jp

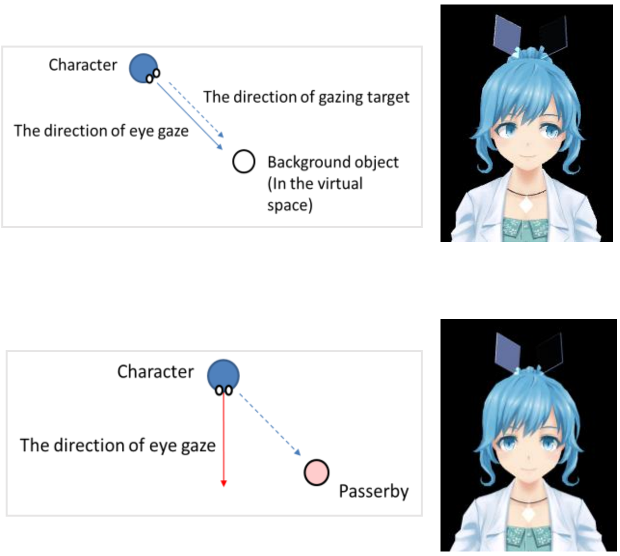

Fig. 1. (Up) If we don't want the character to see any observer, making it to see the background objects in the virtual space. (Down) If we want the character to see the observer (in any position), keeping the direction of the eye gaze frontal.

orientation relative to the observer varied. They showed Mona Lisa effect as the "Differential Rotation effect (DRE)". Their experiment results show that the spatial layout of a photograph remains unaffected while the perceived orientations changed dramatically with different viewing angle. Todorović[6] explains the Mona Lisa effect by comparing 3-D head and 2-D portrait. A real static 3-D head has fixed environmental head orientation and gaze direction even if the observer moves the position. For a 2-D portrait, however, their head orientation and the gaze direction relative to the observer are generally fixed, but they would change according to the environment.

Although researchers used different ways to explain the Mona Lisa effect, we can find that this visual illusion mainly comes from the distortion when the image transforms from 2D pictured space to the physical space. Based on this point, the Mona Lisa effect is inevitable when we try to develop the virtual agents in a traditional flat display. Therefore, our goal in this research is to design the eye gaze patterns for virtual agent to makes eye contact with observers in the necessary timing. We consider that there are mainly 2 situations while we look at the virtual agents. One is standing still to look, and another is looking while passing by the display. We conducted the experiments in these 2 situations separately. By controlling the stimuli of the iris position and the head rotation, we aim to find out what amount a character's head has to move in order to let the observers perceive the eye contact from the virtual agent.

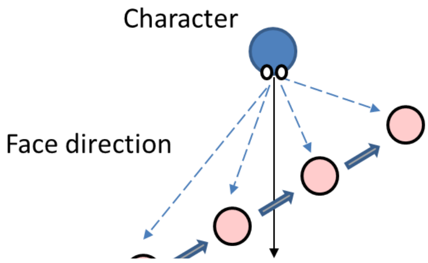

Fig. 2. In the situation of the passerby passing through the display, we keep the eye of the character frontal but head rotates relative the observer's position.

2. Related Work

Several studies [7][8](K. Misawa et al, 2012; S. Al Moubayed, 2012) managed to negate the Mona Lisa effect by utilizing the display of a 3D projection surface for embodied conversational agents or telepresence system. The experimental results showed that subjects could judge eye gaze direction accurately from the 3D projection surface compared to the traditional flat display. Nevertheless, the use of the customized display may take great cost, and it also leads to some limitation that the whole bodies of the characters are hard to be shown on the display. For immediate utilization of the virtual character or showing more various performance, the uses of traditional display still occupy in a great portion. Therefore, it is still necessary for the developers to specially design eye gaze patterns in the 2D display.

In our study, we would like to control the eye gaze based on the Mona Lisa effect and makes the observers feel the character's eye contact in the necessary timing. According to the studies on the basis of Todorović[9] , Maruyama and Endo[10], and Wollaston[11], eye gaze direction not only depends on the iris position within the sclera but also varies by the orientation of the head. Therefore, we also utilize the eye gaze angle and the head turn as control variable to adjust the eye gaze direction in this research. Continued with our prior research, we mainly divide the situations of seeing the virtual agent into the static situation and the situation passing by the digital signage. In the stationary situation, when we want the virtual agent to look at us, we keep the face and eye gaze direction frontal toward the screen. If we want the character not to see anyone, we make the eye gaze see a dummy object in the background of the virtual space (See Fig.1). Nevertheless, this eye gaze pattern causes the side effect that when we want the virtual agent to look at the person in the moving status, the eye gaze behavior would stay still, which is unnatural in the situation if someone attempts to stare at a moving target. In this situation, our prior research proposed the method that eyes keep straight toward the screen but the head turn follows observer vantage points (See Fig2.). According to Todorović[6], the head orientation (κ) in one direction can be balanced by the eye turn (λ) in the opposite direction, which keeps the observer-related gaze (γ) direction at zero, and that means "the looker in the image is looking at me (With "me" located at the origin of the observer-centered coordinate system)". Such a condition can be expressed by the equation "$\gamma = \kappa + \lambda = 0$". As for the situation in the pictorial space, because of the Mona Lisa effect, observers can always feel the eye gaze if the looker in the image is staring in the frontal direction orthogonal to the picture. That is the reason why we keep the eye turn opposite to the head turn in order match the eye contact. To confirm whether the methods above are effective or not, we conducted some experiment and tried to solve some questions we had for the eye gaze generation.

3. Experiment 1

In the experiment 1, we would like to know whether the 2D image of the rotated face can be perceived as giving eye contact only at a certain viewing angle. On the basis of our proposed method (See Fig2.), we expected that the head orientation points at the observer's vantage point with the iris facing straight toward the screen.

3.1. Subjects

5 students from the Tokyo Institute of the Technology participated in the experiment. The range of their age is from 20 to 30.

3.2. Experimental Setup

The 19" LCD monitor was placed on the desk and the participant was seated 1 meter away from the monitor. 3 viewing positions were tested in this experiment, the first one is the straight-ahead direction orthogonal to the flat plane of the monitor, and then successively increasing 30° for the next 2 positions (See. Fig.3.).

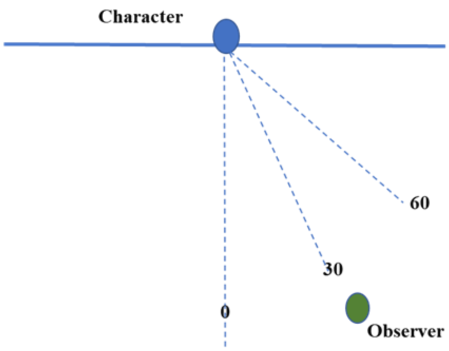

Fig. 4. The result of the experiment 1 shown in 3 bar chart. The letters A ~ E represent each participant.

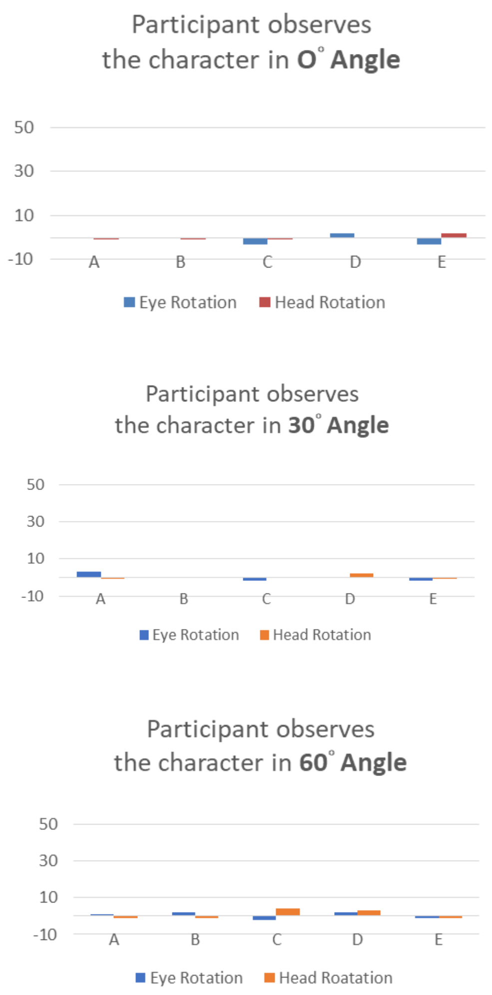

Fig. 3. The installation of the experiment 1. Participants took the experiment on the three positions by turning 30° successively.

3.3. Stimuli

The stimulus was the toon-shaded 3D character with different variation of the head orientation and the iris displacement. Since Mona Lisa effect originally happens in a realistic portrait, we would like to see whether eye gaze perception could change by seeing such a non-photorealistic 3D character. In the current and the sequent experiment, the character was all controlled by the Unity engine. In the first experiment, the head orientation could be rotated from -90° (left side) to 90° (right side), and the iris was moved in angular displacement, which the center of the eye was the original point. We applied the method of adjustment to let participants to adjust the head turn and iris position on their own.

3.4. Procedure

The participant was seated on the one of the marked positions randomly. When the experiment started, the head of the character repeated rotating from 60° to 0° (The same side as the observer). The participant had to adjust the destination of the head turn and the iris position while it kept moving. For example, in the initial destination point, the head turn of the character was 0° and the iris displacement was also 0° (back to frontal face). The participant could adjust the destination of the head turn to 10° and the angular displacement of the iris to 5°. As the participant perceived the eye-contact at the moment that the character's head and the iris turn to the target position, pressing the key to stop the rotation and finish the trial. The program then would record the final head turn and the iris's angular position. Afterwards, the participant repeat the same experimental procedure for the other 2 vantage points.

3.5. Result and Discussion

According to the results (See Fig. 4)we obtained, not only head rotation but when the eye displacement was nearly 0° mostly matched eye contact. This result didn't reach our expectation, which we thought that eye gaze would be perceived with the head orientation turning to the participant. The Mona Lisa effect is quite powerful that the character faces straight-ahead direction or slightly moves a little mostly does eye contact with the observer.

4. Experiment 2

Another question we want to know is that when the passerby is passing through the digital signage, to what extent the head should turn to him/her? Although we know that the character is able to keep eye contact with the observer as long as the eye gaze keeps in the frontal, we didn't know what magnitude of the head turn mostly fits eye contact.

4.1. Apparatus

In the experiment 2, we turned to use a flat mirror display (80.5cm x 148cm) which was similar to normal digital signage showing 2D virtual character model in nearly human size. To allow the character's head and iris to move with an observer, a Kinect device (Microsoft Kinect Sensor V2) was positioned behind the mirror display to track one's movement. The participants stood 100 cm away from the screen to observe the character model. For the moving test trial, the participants would stand on the left side of the display and walk toward the right side of the display in 30° direction (See Fig.5), and the path was also marked on the floor. In addition, we conducted the experiment in the bright room instead of the dark room to allow the mirror display to reflect the objects of the real world, which made the screen look like 3D space.

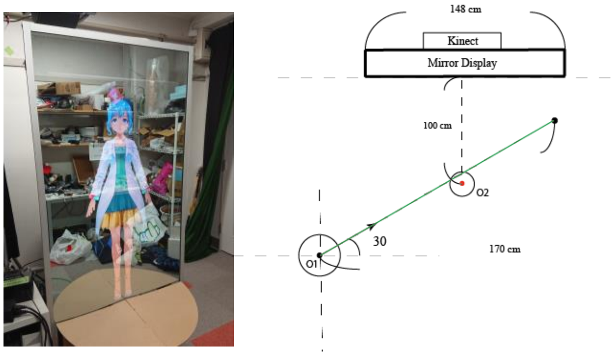

Fig. 5. (Left) The mirror display used for the experiment. (Right) The top view of the room arrangement. In the moving test trial, participants start from the O1 point and walk through the green line to observe the character model. For the standing test trial, participants stand on the O2 point to observe the target.

4.3. Stimuli

The stimulus we used was still the same as the experiment 1, and it would be shown with 2 different moving patterns. The deviation of the head turn is fixed in the horizontal plane and controlled by the relative angle from the center of the character model toward the observer. The eyes, on the other hand, have two ways of moving pattern. For the first, eyes move with the same angle as the head turn. Second, eyes keep in the frontal direction toward the display while the head orient rightward or leftward. (See Fig.6) Both eye moving patterns are also fixed in the horizontal plane as well.

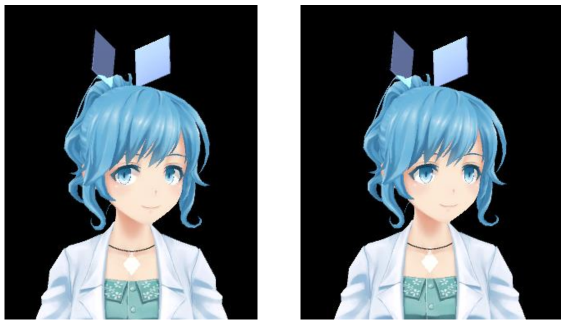

Fig. 6. Two stimuli in the experiment. The character's head would turn in response to an observer's position. The left image shows that the iris move at the same angular position with the head turn. In the right image, the irises almost stay in the direct gaze but the head can rotate rightward or leftward.

4.4. Procedure

In each trial, the participants would see two stimuli: eyes kept in the same direction as the frontal face, and the stimulus that eyes kept in the ahead gaze but the head might orient rightward or leftward. In addition, participants had to see these two stimuli in moving status and fixed status separately. At first, one of the stimuli and observing condition would be randomly decided for the participants. Assuming that the experiment starts in moving status to observe the stimulus which eye direction follows the movement of frontal face, and then the next steps are as follows:

1. Show different head turn amplitude by 5 different multipliers from the biggest one to the smallest one for the participants (1.5x, 1.0x, 0.5x, 0.1x, 0.0x), and using number 1 to 5 to represent each multiplier. 1.0x makes the head turn angle the same as the coordinate of the real world. The participants pass by the screen once for each multiplier and answers the number that mostly fits eye-contact after checking 5 different head turn amplitude.
2. Give the mouse to the participants and let themselves adjust the head turn amplitude while walking forth and back in front of the display. (Rolling the mouse wheeler once to increase or decrease 0.1 x) The participants start to adjust the parameter from the multiplier he/she choose in the last step and keeps adjusting it until he/she feels that the eye-contact always matches while he/she is walking.

When checking another stimulus, the participant repeated the

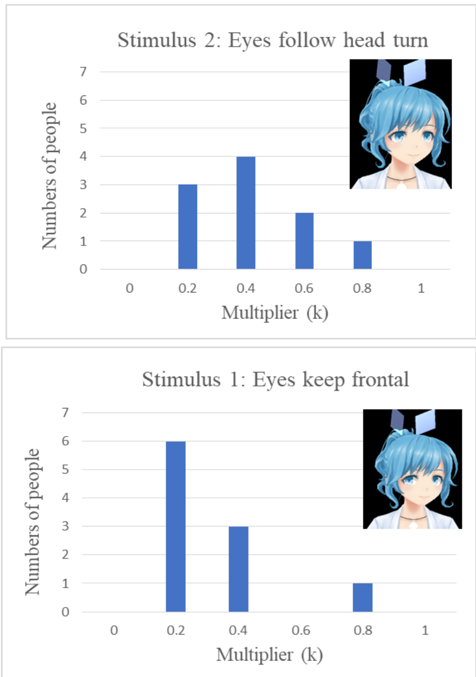

Fig. 7. The results of two stimuli in the condition of walking through the display.

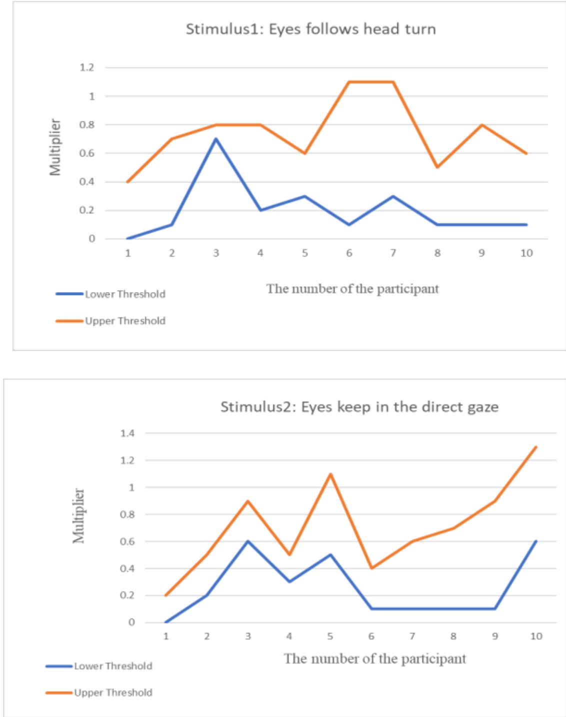

Fig. 8. The results of two stimuli in the condition of standing in a fixed position ahead the display.

above steps again.

In the next condition, observing in the fixed position, the participants can only see the character by moving upper half of the body. The steps of the experiment are also the same. However, in the step2, there is a difference that the participants are required to find the upper threshold and the lower threshold that eye-contact occurs instead of finding the pinpoint. In addition, if the participants observe the character in a long time, the visual fatigue might appear and influence the result. We would pay attention to this case and allow the participants take a rest if this case happens.

Finally, when the participants checked the two conditions, they would be asked to answer which stimulus appears to match eye-contact.

4.4. Results and Discussion

We had 10 students conduct the experiment. The figure 7 shows the results of observing in walking status. The smaller the multiplier is, the narrower the head turn amplitude is. As the histogram shows, values of multiplier in the stimulus 1 (character's eyes follow head turn) tend to be scattered, and the values of multiplier in the stimulus 2 (Character's eyes stay in the direct gaze) mostly gather in the range of 0.1~0.2. We originally estimated that the result of the general situation (stimulus 1) would tend to be smaller values because of Mona

Lisa effect, and then the values of stimulus 2 are expected to be bigger ones. However, the result is contrary to the expectation despite the data of stimulus 2 are more centralized. According to participants' comment, some people felt that the small head turn of the stimulus was more natural because eye-contact would mismatch in the larger moving amplitude. Besides, some people also feel that it is not natural to watch people in this way.

For another condition, observing the participant in the fixed position, we expected that the perception of the gaze range in the stimulus 1 would be narrower and near the zero value because Mona Lisa effect makes the frontal face eye-contact most intensive. On the other hand, we anticipate our method different from the stimulus 1 and observers should perceive the gaze under a higher value, which means that head turn will become larger when observers shake upper bodies to watch the character's gaze. For the results shown in the figure 8, however, the stimulus 2 that eyes keep in the direct gaze didn't really match our expectation. There are still some participants feeling that less head turn does match eye-contact. In addition, the range of gaze perception is not very stable. This is because every individual may have different perception, or perhaps the method of the experiment itself is not very objective. We should figure out if there is some noise that influence the experiment result.

5. Conclusion and Future Work

Seeing the results from the 2 experiments above. The common points are that observers do more perceive eye contact with the head rotating slightly. Despite the powerful Mona Lisa effect, it is still better to make the head move when the observer is passing by the display. However, there could be some interference or other reasons that made the results. For example, the participants sometimes went over the range and made Kinect device unable to detect them. The method of adjustment might also not be good in our case, which adjusting the parameter back and forth could influence people's perception gradually. Besides, the walking distance and speed of foot pace should also be changed like the situation that passersby go through the digital signage. Therefore, in the next step we aim to conduct a more accurate experiment to confirm the results again. Apart from the toon shaded character, we will also test other types of the character such as photorealistic model and geometric primitive model. Afterwards, we will conduct the experiment in the VR space to compare the eye gaze perception when the character is in the 3D space. Such experiment will help us to confirm whether our system can be widely used for different kinds of virtual agents.

6. Acknowledgements

This paper is based on author's mater thesis which is supported by Hasegawa Laboratory. We also thank our laboratory members and those who came to participate in the experiments.

7. References

[1] ltd, Research and Markets. "Global Digital Signage Market Analysis & Forecasts 2012-2016 & 2017-2022 ".2018.

[2] INSIGHTS. "Virtual Assistants: The Next Generation of Digital Signage ". *https://insights.samsung.com/*. Retrieved 2015-05-19.

[3] Kleinke, C. L. 1986. Gaze and eye contact: a research review. Psychological Bulletin Vol. 100 No. 1, 78-100.

[4] Maruyama K, Endo M, Sakurai K, 1985 "An experimental consideration on Mona Lisa gaze effect" Tohoku Psychologica Folia, 44, 109-121

[5] Goldstein, E. B. (1987). Spatial layout, orientation relative to the observer, and perceived projection in pictures viewed at an angle. Journal of Experimental Psychology: Human Perception and Performance, 13, 256–266.

[6]S. Al Moubayed, J. Edlund, and J. Beskow. Taming mona lisa: communicating gaze faithfully in 2d and 3d facial projections. ACM Transactions on Interactive Intelligent Systems, 1(2):25, 2012.

[7]Kana Misawa, Yoshio Ishiguro, Jun Rekimoto, LiveMask: A Telepresence Surrogate System with a Face-Shaped Screen for Supporting Nonverbal Communication, Interaction 2012, pp.41-48, (2012).

[8] Todorović, D. (2005) Geometrical basis of perception of gaze direction. Vision research 46 (2006) 3549-3562

[9] Todorovic, D. (2009)The effect of face eccentricity on the perception of gaze direction, Perception, 38 (2009), pp. 109-132

[10] Maruyama, K., & Endo, M. (1983). The effect of face orientation upon apparent direction of gaze.Tohoku Psychologica Folia, 42, 126–138.

[11] W. H.Wollaston, 1824, "On the Apparent Direction of Eyes in a Portrait,", Philosophical Transactions of the Royal Society of London, 114, 247-256.