Coding gaze tracking data with chromatic gradients for VR Exposure Therapy

Bruno Herbelin
Aalborg University Esbjerg,
6700 Esbjerg, Denmark
bh@aaue.dk

Helena Grillon
École Polytechnique Fédérale de Lausanne,
1015 Lausanne, Switzerland
helena.grillon@epfl.ch

Pablo De Heras Ciechomski
Visualbiotech Sarl,
1015 Lausanne, Switzerland
pablo@visualbiotech.ch

Daniel Thalmann
École Polytechnique Fédérale de Lausanne,
1015 Lausanne, Switzerland
daniel.thalmann@epfl.ch

Abstract

This article presents a simple and intuitive way to represent the eye-tracking data gathered during immersive virtual reality exposure therapy sessions. Eye-tracking technology is used to observe gaze movements during virtual reality sessions and the gaze-map chromatic gradient coding allows to collect and use these important information on the subject's gaze avoidance behavior. We presents the technological solution and its relevance for therapeutic needs, as well as the experiments performed to demonstrate its usability in a medical context. Results show that the gaze-map technique is fully compatible with different VR exposure systems and provides clinically meaningful data.

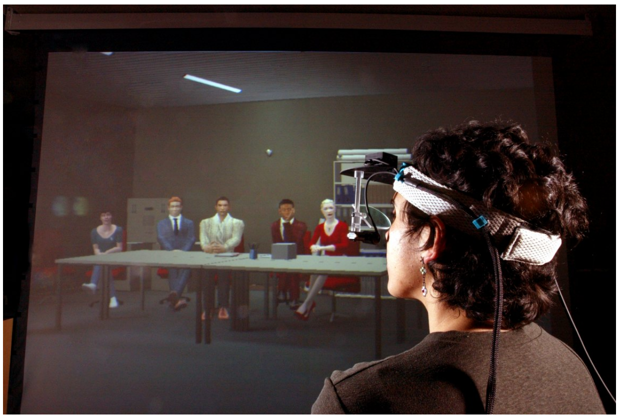


Figure 1. Subject wearing an eye-tracking device while facing a virtual assembly *(Photo Alain Herzog.)*

1 Introduction

It is well known that one of the defensive behaviors present in phobic people is gaze avoidance of the feared stimuli [1]. More specifically, in the case of social anxiety disorders, this translates itself in an avoidance of salient facial features (eyes, nose, mouth). Horley [6] observed that the gaze behaviors of social phobics show a characteristic 'eye to eye' avoidance. Gaze behavior analysis is therefore of high interest for psychiatrists working on the treatment of phobias with Cognitive and Behavioral Therapy (CBT).

Eye-tracking systems can be used to observe such behaviors. However, usual eye-tracking equipments only provide 2D gaze point coordinates on the recorded video images of the subject's view during exposure. Analyzing such data is very interesting, but extremely laborious as it is essentially based on human interpretation and video annotation. Finding an automatic and reliable way to observe and quantify the avoidance in gaze behaviors would offer many opportunities for the assessment and diagnosis of anxiety disorders.

A first step towards a solution is to use Virtual Reality (VR) for therapeutic exposure sessions. Compared to classical *in-vivo* exposure therapy, Virtual Reality Exposure Therapy (VRET) has many advantages such as on-demand simulation of any situation and dynamic control of the content. Moreover, according to our experience with the VR treatment of social phobia [5], the simulation context is much more appropriate to behavioral observation while preserving the efficiency and validity of the usual CBT procedures. For example, in an earlier experiment with gaze tracking, we have shown that a simple map of the patients gaze targets on the scene was already an interesting tool for therapists [2]. Recently, we have confirmed the clinical validity of this tool with social phobic subjects [3].

Although encouraging, these experiments only performed eye-tracking with a static point of view of the simulated scene. This way, the 2D eye-tracking points are shown on the image seen by the subject. However, this is not appropriate for VR immersion: during the exploration of a virtual environment (VE), the 2D coordinates of the eye-tracker have to be coupled with the subject's moving view. One solution is to operate directly in the 3D space, for example by computing geometric factors expressing the angular deviation between the gaze vector and a point of interest [17]. However, the resulting data are purely numerical and abstract, hence the therapists' preference for the former visual solution.

We propose a compromise which consists in obtaining gaze target coordinates on the surface of the 3D objects. This can be done by performing 3D picking at the tracking coordinates on the perspective view. Our gaze-map chromatic gradient coding system uses color picking to obtain numerical gaze information during immersion and represents the results in an intuitive and visual manner.

First, we present the various works related to eye-tracking and gaze behavior analysis for therapy in section 2. In section 3, we relate to our previous observations to analyze the issues to resolve for an optimal gaze tracking. We then describe the key elements of our implementation in section 4. Finally, we present the tests made under different VR exposure conditions for social phobia therapy in section 5 before discussing our results and concluding.

2 Related work

As of today, many studies have been conducted regarding the use of VR in the treatment of social phobia [12, 13, 4, 7, 9], all leading to the conclusion that VR immersion seems adequate for such treatments. However, our aim in this paper is not to demonstrate this hypothesis but to provide researchers and therapists with a new diagnosis and assessment tool.

Eye-tracking consists in following eye movements and computing gaze direction with a computer system. This technology really became usable in the late nineties [22] and today's commercial products usually track pupil and corneal reflection with a video camera placed on the head or close to it.

Various experiments were conducted on its applications to gaze-controlled simulations [10, 18] or interactive multimodal systems [19, 8]. Recently, Prendinger et al. [14] proposed an eye-based 'infotainment' presentation system in which 3D agents present product items. Their system uses real-time eye movements to adapt the presentation to the user. However, as they pre-define rectangular areas of interest on screen, the user cannot change the point of view and the characters have to be static.

To perform gaze analysis during a VR experiment, the solution originally developed for aviation was to integrate tracking cameras directly in the Head Mounted Displays (HMD). Experiments conducted by Renaud et. al [17] lead to very detailed analysis of the behavioral dynamic of users' visual exploration in a VE. They base their results on the numerical estimation of Gaze Radial Angular Deviation (GRAD), geometrically obtained from the line of sight's vector and a point of interest in space. This method allowed them to demonstrate gaze avoidance toward spiders in arachnophobic patients [16], and to analyze visual centers of interest to sexual stimuli [15].

Lange et al. [11] conducted a study for arachnophobia. They used eye-tracking to determine the differences in visual behavior between phobics and non-phobics. They conclude that phobics scan the environment as part of defensive behavior. Smith [20] worked with 46 socially anxious and non-socially anxious subjects to determine their gaze behavior toward disgust-faces versus happy faces. The author concludes that socially anxious individuals have a tendency to present delayed disengagement from social threat.

3 Initial results

The gaze-map solution we opted for was motivated by the results to our former experiments, as summarized in this section. For our experiments with gaze tracking, we combined the Polhemus VisionTrak^{TM} eye-tracking device with a 6 degrees of freedom Ascension MotionStar^{TM} magnetic sensor. Subjects wore these devices on the head and, after a brief calibration procedure, were free to move while their gaze target was computed on the screen. They were then exposed to a 3D scene displayed in front of them on a $3 \times 2.3m$ back-projection screen (figure 1).

3.1 Precision issues

According to our reliability and precision tests carried out on the eye-tracking equipment [2], gaze can be used to interact with virtual humans and to analyze the visual interest of a subject for different body parts. More specifically, we estimated the reliability of eye-tracking data by measuring how far the tracked points were from a point supposedly looked at. When a subject is fixing a point on screen for two minutes, 80% of the eye-tracking data are in an area centered on that point and covering 13% of the screen width, 60% are in an area covering 6.5%, and only 30% in a small 3% area.

Figure 2. Low LOD picking meshes animated with the humanoid skeleton.

These results are not surprising as it is known that the accuracy of the eye-tracking technology is not perfect, and that the eyes are not static when fixing a point. A filtering has to be performed on the data in order to average the tracking points and eliminate the eye saccades. One solution suggested by the above results is to consider that a point is looked at when it is relatively close to the point measured by eye-tracking. For example, by considering all points located inside an area covering 6.5% of the screen width from the supposedly looked at point, we ensure 60% chances of covering the point actually looked at.

3.2 Therapeutic needs

During a clinical experiment with eight social phobic patients following a full VRET treatment [3], we validated the efficiency of VR exposure to treat social phobia. Moreover, the observation of gaze behavior before and after treatment consisted in a very promising tool which allowed us to conclude that there was a noticeable improvement in eye contact avoidance after therapy.

However, the visual observation of the cloud of gaze points on the static view of the scene only provides a qualitative indication of the behavioral changes. Furthermore, it fully depends on the therapist's interpretation. These observations were essentially intended to determine if "the virtual character's face is much more looked at after the end of the treatment than before treatment", or if the talking characters "are more looked at after treatment than before" (p.111). We therefore considered that the gaze analysis should focus on a precise and quantitative measurement of the gaze target position directly on the objects of interest (the virtual humans).

4 Gaze-map for virtual humans

In order to satisfy the therapeutic requirements for gaze analysis of social phobic subjects, we consider the tracking of gaze targets on virtual humans only (as opposed to other environment objects). This section presents the implementation and the technological choices motivated by our design decisions.

4.1 Implementation

Picking allows to identify which object is visible at a point on the image. It is easy to integrate into a real-time 3D engine (potentially doing stereoscopic rendering), and it follows the user's point of view. We did not retain the 3D polygon picking technique because its precision is limited by the mesh's level of segmentation. Contrary to this suggestion, texture-based color picking offers more flexibility regarding the mesh complexity (it works even with multiple levels of details), and can be performed on a texture intentionally designed to represent a map of the object parts.

The implementation of color picking only requires basic OpenGL features. On request, the program performs a hidden rendering of a specially colored version of the object to be tracked and "the application simply reads back the pixel under the cursor" [21, p.508]. In fact, rendering one single pixel is enough, and the cost in performance is negligible.

The integration within the rendering of our animated virtual humans was done in order to keep all their rendering features (real time skinning on skeleton animation, Levels Of Detail (LOD), textures). To improve the performance, we used a simplified version of the mesh for the picking humans. This was possible because the rendering to the picking buffer is independent from the final visual rendering. As the same skeleton is used in both cases, the animation is the same and the mesh coverage is almost the same (see figure 2). In order to have accurate color picking, we turn off every color modulating step to render the picking buffer, and use a non-lossy format for the picking texture file. The color rendering is performed individually for each humanoid present in the scene at the time of its rendering. We therefore obtain separate picking information for each character.

4.2 Hue-Saturation gaze-map

The main idea behind the gaze-map chromatic gradient coding is to consider the hue and saturation color components as the coordinates of the 2D point on the surface of a 3D object. This is obtained by mapping a color gradient texture figuring vertical hue (H) variations and horizontal saturation (S) changes. Figure 3 shows how this texture is mapped on the 3D model of a human to cover its entire body. The H values are low in the feet and high for the head, the S values are high on the left and low on the right. Moreover, it is quite easy for a designer to perform this front-view texture mapping (symmetric for the back). We equally introduced optional U-V mapping distortions on the face to have more detail on facial regions as for

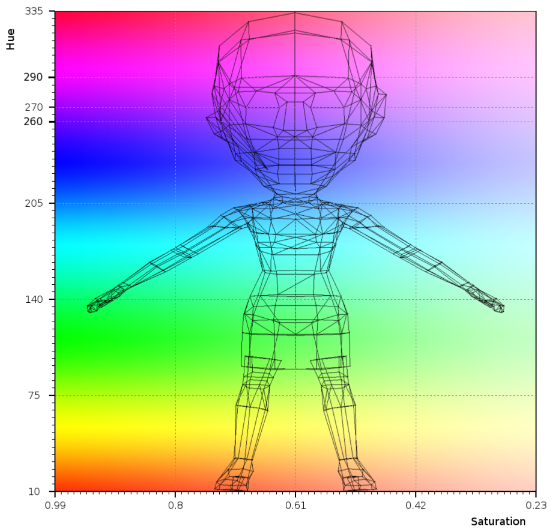


Figure 3. Gaze-map chromatic gradient coding on a humanoid mesh (front view).

eyes and mouth. Note that the rendering is done in *RGB*, but the conversion to *HSV* is simple. We avoided low saturation in the color gradient since for $S = 0$, the color is white for any value of H, and the conversion would introduce artifacts.

Basically, reading the color of a point on the picking humanoid provides immediate correspondence with a precise location on its surface by referring to the UV mapping shown in figure 3.

4.3 Area picking approximation

According to our observations on the reliability of eye-tracking data, we needed to compensate for the low precision and the instability of the gaze target. As suggested in section 3.1, a simple way to filter the data is to consider the average position of all points located in an area surrounding the eye-tracking point.

We enlarged the picking area to a size corresponding to the eye-tracking precision by extending the picking algorithm to support square regions centered around the picking point. To perform a fast OpenGL render-to-texture, the dimensions should be a power of two ($D = 2^n$ pixels with $n \geq 0$). The degree n has to be chosen according to the desired gaze picking reliability. For example, with a screen resolution of $800 \times 600px$, a picking area of size $D = 2^5 = 32px$ covers points lying between 4% and $\sqrt{2}.D = 5.6\%$ of the screen width, whereas an area of $64px$ covers 8% to 11%.

To calculate the gaze-map coordinates of the center of a picking area, the *HVS* colors are simply averaged for all

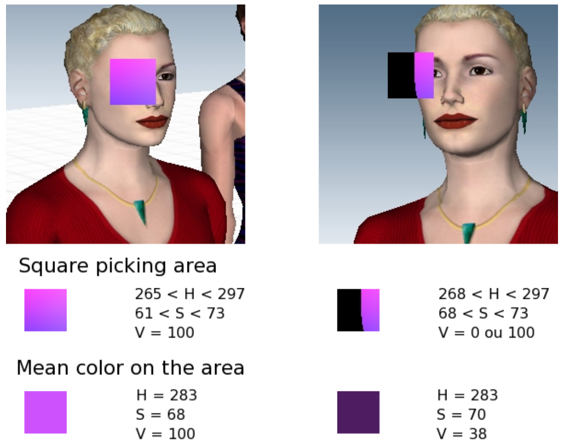


Figure 4. Approximation on a large picking area.

points within it. Moreover, as the background is cleared to black before picking each humanoid, only the value component (V) of the *HSV* color is affected when averaging on an area containing background. Figure 4 shows how the H and S picking coordinates are preserved when the picking is partially outside a virtual human. V can be described as the percentage of pixels of the 3D model inside the picking area ($V = 0\%$ is background). As a consequence, it can be considered as a tolerance factor to the picking, with $V = 100$ meaning the picking is done inside the model, and $V = 0$ meaning the picking is outside.

One special case have to be considered though: when a part of a character is in front of another one (e.g. a hand in front of the trunk), the color average will not be able to specify which part is picked, and the resulting *HSV* coordinates may even be outside the character's body. The picking area shall therefore remain relatively small in order to avoid the occurrences of this particular case. However, considering the a picking frequency at 30Hz for a session of several minutes, the amount of data collected compensates for the rare occurrence of such error.

4.4 Gaze-map data interpretation

As seen before, therapists need to have a quantitative measure of the attention given to each virtual human over an exposure session. They also need the gaze distribution on various body parts (such as the face).

The attention given to each character is easy to determine; the V component can be used to decide if a character is looked at. Various features can be computed based on V, the simplest is the average over time which estimates the percentage of the session duration spent looking at a character.

The distribution of gaze on body parts can be obtained from the gaze-map picking data by segmenting the H and S values into slices. Taking advantage of the linearity of the texture gradient, we made nine sections in regular intervals of H to identify the body parts: feet, knees, thighs, hips, torso, shoulders, neck, mouth, eyes, and hair. The S component of the gaze-map provides lateral information on the body, where $S > 0.61$ corresponds to the left part and $S < 0.61$ to the right. In addition, in order to obtain a quantitative estimation of the visual attention on each virtual human's face, the following features can be computed:

- $\Delta_H = H - 260$: Vertical difference to the center of the face. $\Delta_H = 0$ when the subject is looking straight at a virtual human, $\Delta_H > 0$ when looking above the eyes and $\Delta_H < 0$ when below.
- $\Delta_S = S - 0.61$: Horizontal difference to the center of a virtual human. $\Delta_S = 0$ is middle, $\Delta_S < 0$ when looking on the left side, and $\Delta_S > 0$ on the right.
- $d = \sqrt{\Delta_{H_n}^2 + \Delta_{S_n}^2}$: Distance to the face of a virtual human. Normalized values H_n and S_n are obtained by dividing H and S by the extrema of the $H - S$ map (figure 3).

The criteria $d \lesssim 0.15$ determines if a point is inside the face (computed by considering $H = 290$ and $S = 0.51$ are the limits for the face). This allows to determine if a subject is looking inside the face of a virtual human or not.

5 Experiments and results

We conducted three experiments to verify that our solution is usable during VR immersion and satisfies the therapists' needs. For each one, we simulated a typically feared situation for social phobic: public speaking in front of an assembly. The virtual humans in the scene were all animated to show interest in the subject's talk (simulation of behaviors such as looking at the subject, blinking the eyes and changing posture). Additionally, one of the characters gave verbal encouragements from time to time (manually triggered).

5.1 First experiment: tracking with HMD

Our first objective was to verify that we could perform gaze tracking during immersion. The typical VR condition chosen for this validation was immersion with HMD. We used a rather low cost setup consisting of a pair of Virtual I/O i-glasses™ equipped with an InterSense™ tracker. Although we did not use an eye-tracking device, this experiment provided the necessary conditions to prove our point: the picking technique should be robust to the HMD camera movements. We made the hypothesis that the visual attention in a low field of view display would be mainly around the center of the screen. The picking area was set very large to cover 16% of the screen width (figure 5). Therefore, the picking data shall indicate the changes of attention with head movements (instead of eye movements).

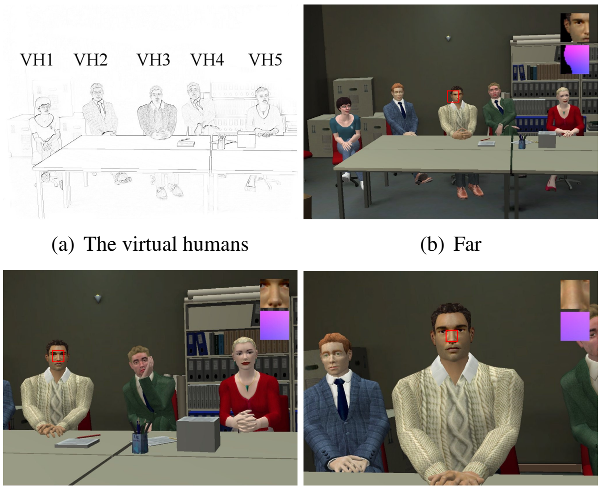
Figure 5. View of the 3D scene for the 1st experiment (the head up display was off during the sessions).

We exposed 130 non-phobic subjects to a virtual environment figuring an assembly of five characters facing them. According to their preference, they had to simulate an examination, a job interview, or a professional meeting for a few minutes.

Our technique operated well in the HMD condition, providing time-stamped raw HSV data for every humanoids. This allowed us to compute interesting features on subjects' head movements during immersion. We used the V component in combination with the identification of the virtual human to count the number of times a subject turned towards the different characters during the session. We used the distance d to determine when a subject was facing a character and to measure the duration and the frequency of these face-to-face phases. We could observe that, with the HMD, people were not naturally inclined to turn the head to face people, but also that this behavior was reinforced by social phobia tendencies. A detailed interpretation of these data is given in [5].

5.2 Second experiment: gaze-map with eye tracking

The goal of this little experiment was to validate the reliability of the gaze-map technique when used with the eye-tracking device; if our estimation of the gaze tracking reliability is correct, we should be able to detect which virtual human and which part of it a subject is looking at. The distance to and the size of the objects of interest have a strong influence on the gaze-map accuracy: we cannot obtain as precise information on very small targets (a character situated far in the scene) as on large ones (close-up). To experiment with various target sizes, three subjective



(a) The virtual humans (b) Far (c) Medium (d) Close

Figure 6. Camera points of view for the 2nd experiment (the head up display was off during the sessions).

Table 1. Distribution of gaze as % of picking per body part (2nd experiment).

Body parts	Far	Medium	Close
Hair	7.6	0.0	2.2
Eyes	28.2	43.6	57.1
Mouth	33.6	28.3	21.0
Neck	9.7	6.6	2.2
Shoulders	15.9	14.8	11.8
Torso & arms	2.6	6.0	4.5
Hips & hands	1.0	0.6	1.2
Thighs	0.3	0.0	0.0
Knees	0.3	0.0	0.0
Feet	0.8	0.0	0.0

Figure 7. 2D representation of gaze targets for a phobic subject (3rd experiment).

positions toward a virtual assembly were selected: far, medium and close views (figure 6).

The test was performed on non-phobic subjects whom we asked to look at the characters in the eyes during two minutes. The eye-tracking system was used (as in section 3) and the picking area was intentionally set slightly low to stress the precision limits ($D = 32px$ for $800 \times 600px$ displays). Table 1 summarizes a typical set of data (25 years old male subject).

Concerning the need for therapists to automatically establish the distribution of gaze targets over virtual humans, an average of V over the exposure session represents the intensity of gaze on each character. The complement to one of the total V for every humanoids estimates the gazes spent on the background.

Our second requirement was to automate the analysis of gaze interest for the different body parts. Table 1 shows that each distance configuration allows a different level of accuracy in gaze target detection. A close up on a character allows to observe gaze differences between hair, eyes and mouth, whereas in the far view, results remain at the level of head, body and legs.

Finally, in order to verify that the values were actually correct and sufficiently reliable, we compared the points looked at by the subjects with the ones we observed. First, using the think-aloud testing protocol, we could continuously confirm that the gaze location verbally expressed corresponded to the the picking area visible in a head-up display (up-right corner in figures 6.b to 6.d). Second, we asked the subjects to summarize their behavior after each session and obtained a good match between the expressed gaze targets and the gaze-map data. For instance, the distribution of V over the five characters (10%, 18%, 22%, 10% and 11%) corresponded to what the subject related: "I have successively looked at each person for the same lapse of time, then came back on the central character for a longer period". As the subjects were not phobic, their answers were considered trustworthy.

5.3 Third experiment: comparing classical and gaze-map data

The objective of this last validation check was to confirm that therapists could use the gaze-map data in the same way as in the former validated eye-tracking sessions (using 2D points on screen). We recorded both 2D and gaze-map data in some of the public speaking sessions performed during our former study with social phobic patients [3]. The hypothesis to verify here is that the newly obtained data are at least as valuable from the therapeutic point of view, if not better.

Table 2. Distribution of gaze as % of picking per virtual character (3^{rd} experiment).

Character	Non-Phobic	Phobic
VH0	0.58	0.27
VH1	0.00	0.00
VH2, 3	0.00	0.01
VH4	0.01	0.02
VH5	0.02	0.02
VH6, 7	0.02	0.06
VH8	0.00	0.02
Background	0.35	0.53

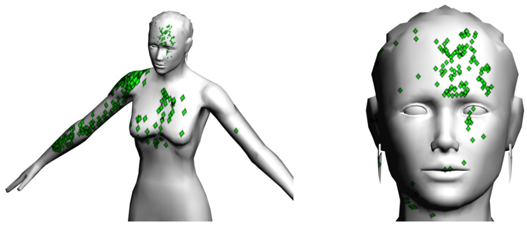
Table 3. Distance to the central character (3^{rd} experiment).

Distance	Non-Phobic	Phobic
Δ_H	-9.68	-39.62
Δ_S	0.04	0.09
d	0.17	0.44

The set-up was the same as in section 3. Subjects were asked to simulate a discussion in a bar with a person recently met (figure 7). The sessions performed with the phobic subjects were guided by their therapist (controlling the virtual human).

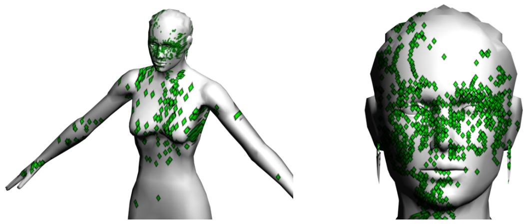
Figures 7 and 8 show an example of the results obtained with a phobic subject. The first is a traditional 2D representation of the gaze values on the projected scene. The second is a mapping of gaze map data on the main character (VH0). We can easily see that the results are identical and observe the same bias on the two representations: the subject looked at the forehead or on the left side of the face, but avoided the character's eyes. The analysis of gaze behavior over the exposure session is immediate with the gaze-map data. For comparison, the 3D gaze-map data of the non-phobic subject are shown in figure 9. Table 2 shows the repartition of gaze per characters. The proportion of time spent looking at the different characters in the scene was much higher for the non-phobic subject as for the phobic one (average V 65% v.s. 47%). This difference is even larger for VH0. In table 3 we can also see that the distance to the center of the face is equally much smaller for the control subject as for the phobic subject who was looking mainly bellow the eyes ($\Delta_H < 0$).

The 3D visualization of gaze-map data (figures 8 and 9) could been used for a qualitative estimation of the behavior by the therapist, and also as a tangible element to show to the patient. The factors derived from the data provided



(a) Upper body (b) Close up on the face

Figure 8. Gaze-map representation of gaze targets on the 3D model for a phobic subject (3^{rd} experiment).



(a) Upper body (b) Close up on the face

Figure 9. Gaze-map representation of gaze targets on the 3D model for a non-phobic subject (3^{rd} experiment).

quantitative estimation of the avoidance (tables 2 and 3).

6 Conclusion

We introduce a simple solution to the problem of eye-tracking data representation and analysis in the context of VR immersion. Firstly, whereas classical eye-tracking data recording systems provide 2D gaze point coordinates relative to the user view, gaze-map picking gather data directly in the 3D scene. This allows our technique to record all the gaze points during a session when a user is freely exploring a VE (e.g. immersed with HMD). Secondly, this technique exploits the properties of color picking on a hue-saturation gradient to efficiently provide robust and meaningful measurements. Chromatic gradient coded data can be obtained on multiple moving and deforming meshes – *e.g.* skinned character in different levels of details. Finally, when used with an eye-tracking device, the gaze-map technique allows to compute statistics on user's visual interest for the objects in a scene or for some specific parts of them.

Throughout experiments in the context of VRET of social phobia, we could satisfy the needs for therapists to characterize the subject's gaze behavior relatively to the feared stimuli. Our results show that the technique provides information on the gaze distribution over the characters and over their body parts which are as valuable for the therapist as the classical 2D gaze target coordinates on screen. Moreover, the computation of numerical factors and the assessment of data significance is very intuitive and explicit. However, in order to validate this system as a diagnosis tool for therapists, more extensive research on a large cohort should be undertaken. The gaze-map data shall also be complemented with our work on other behavioral factors (blinks, pupil dilation).

For a general application, the implementation of picking and gaze-map chromatic gradient coding could be extended to all the objects of the virtual environment with limited influence on performances.

Acknowledgments

We would like to thank Dr. Francoise Riquier for her competence and patience as psychiatric expert, Mireille Clavien for her great design work, and Jan Ciger for his technical support.

References

[1] A. P. Association. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV)*. American Psychiatric Publishing, Washington DC, 4 edition, 1994.

[2] J. Ciger, B. Herbelin, and D. Thalmann. Evaluation of gaze tracking technology for social interaction in virtual environments. In *Proc. Second International Workshop on Modelling and Motion Capture Techniques for Virtual Environments (CAPTECH)*, CH, 2004.

[3] H. Grillon, F. Riquier, B. Herbelin, and D. Thalmann. Virtual reality as therapeutic tool in the confines of social anxiety disorder treatment. *International Journal on Disability and Human Development*, 5(3):243–250, 2006.

[4] S. Harris, R. L. Kemmerling, and M. North. Brief Virtual Reality Therapy for Public Speaking Anxiety. *Cyberpsychology & behavior*, 5(6):543–550, Dec. 2002.

[5] B. Herbelin. *Virtual Reality Exposure Therapy for Social Phobia*. PhD thesis, École Polytechnique Fédérale de Lausanne, Lausanne, CH, 2005.

[6] H. Horley, L. Williams, C. Gonsalvez, and E. Gordon. Social phobic do not see eye to eye: A visual scanpath study of emotional expression processing. *Journal of Anxiety Disorders*, 17(:):33–44, 2003.

[7] L. K. James, C.-Y. Lin, A. Steed, D. Swapp, and M. Slater. Social anxiety in virtual environments: Results of a pilot study. *Cyberpsychology & behavior*, 6(3):237–243, June 2003.

[8] M. Kaur, M. Tremaine, N. Huang, J. Wilder, Z. Gacovski, F. Flippo, and C. Mantravadi. Where is “it”? event synchronization in gaze-speech input systems. In *5th international conference on Multimodal interfaces, ICMI2003*, November 2003.

[9] E. Klinger, S. Bouchard, P. Legeron, S. Roy, F. Lauer, I. Chemin, and P. Nugues. Virtual reality therapy versus cognitive behavior therapy for social phobia: a preliminary controlled study. *Cyberpsychology & behavior*, 8(1):76–88, 2005.

[10] C. Krapichler, M. Haubner, R. Engelbrecht, and K. Englmeier. VR interaction techniques for medical imaging applications. *Computer Methods and Programs in Biomedicine*, 56(1):65–74, April 1998.

[11] W. Lange, K. Tierney, A. Reinhardt-Rutland, and P. VivekanandaSchmidt. Viewing behaviour of spider phobics and non-phobics in the presence of threat and safety stimuli. *British Journal of Clinical Psychology*, 43(3):235–243, 2004.

[12] M. North, S. North, and J. Coble. Virtual reality therapy: an effective treatment for the fear of public speaking. *International Journal of Virtual Reality*, 3(3):2–7, 1998.

[13] D.-P. Pertaub, M. Slater, and C. Barker. An experiment on fear of public speaking in virtual reality. *Studies in health technology and informatics*, 81:372–378, 2001.

[14] H. Prendinger, T. Eichner, E. André, and M. Ishizuka. Gaze-based infotainment agents. In *ACE ’07: Proceedings of the international conference on Advances in computer entertainment technology*, pages 87–90, New York, NY, USA, 2007. ACM Press.

[15] P. Renaud, G. Albert, S. Chartier, M. Bonin, P. DeCourville-Nicol, S. Bouchard, and J. Proulx. Mesures et rétroactions psychophysiologiques en immersion virtuelle: le cas des réponses oculomotrices et sexuelles. In *IHM ’06: Proceedings of the 18th International Conferenceof the Association Francophone d’Interaction Homme-Machine*, pages 175–178, New York, NY, USA, 2006. ACM Press.

[16] P. Renaud, S. Bouchard, and R. Proulx. Behavioral avoidance dynamics in the presence of a virtual spider. *IEEE Trans Inf Technol Biomed*, 6(3):235–43, 2002.

[17] P. Renaud, J.-F. Cusson, S. Bernier, J. Décarie, S.-P. Gourd, and S. Bouchard. Extracting perceptual and motor invariants using eye-tracking technologies in virtual immersions. In *Proceedings of the IEEE International Workshop on Haptic Virtual Environments and their Applications (HAVE’2002)*, pages 73–78, Nov. 2002.

[18] M. Rizzo, J. Moon, M. Wilkinson, K. Bateman, J. Jermeland, and T. Schnell. Ocular search of simulated roadway displays in drivers with constricted visual fields. *Journal of Vision*, 2(7):162, 2002.

[19] L. Sibert and R. Jacob. Evaluation of eye gaze interaction. In *CHI 2000 Conference on Human Factors in Computing Systems*, April 2000.

[20] J. D. Smith. *Social Anxiety and Selective Attention: A Test of the Vigilance-Avoidance Model*. PhD thesis, Florida State University, 2005.

[21] M. Woo, J. Neider, and T. Davis. *OpenGL Programming Guide*. Addison Wesley Longman, Reading, MA, second edition, 1997.

[22] G. Yang, L. Dempere-Marco, X. Hu, and A. Rowe. Visual search: psychophysical models and practical applications. *Image and Vision Computing*, 20(4):273–287, April 2002.