LETTER

Saccade Information Based Directional Heat Map Generation for Gaze Data Visualization

Yinwei ZHAN†, Yaodong LI†, Nonmembers, Zhuo YANG††, Member, Yao ZHAO††, and Huaiyu WU†††, Nonmembers

SUMMARY Heat map is an important tool for eye tracking data analysis and visualization. It is very intuitive to express the area watched by observer, but ignores saccade information that expresses gaze shift. Based on conventional heat map generation method, this paper presents a novel heat map generation method for eye tracking data. The proposed method introduces a mixed data structure of fixation points and saccades, and considers heat map deformation for saccade type data. The proposed method has advantages on indicating gaze transition direction while visualizing gaze region.

key words: eye tracking, gaze analysis, heat map, visualization

1. Introduction

Eye tracking research and application are actively undertaken [1], [2]. By using eye tracker device, raw data like time stamp, pupil center and confidence can be obtained.

Visualization is an intuitive technique for eye tracking data analysis. There are three main visualization methods: scan path, heat map and AOI (area of interest). Scan path can clearly reflect paths and order of saccades, while heat map and AOI mainly reflect spatial distribution of attention.

Eye tracking data pretreatment and parameterization are required before visualization [2]. Eye tracking data pretreatment is mainly used to remove the error data in the eye tracking data, and combine several gaze points that are close in time and space into one fixation point. Parameterization refers to the adjustment of the structure of the eye tracking data to make it more convenient for further processing and analysis. Eye tracking data can be parameterized into fixation points, saccades and scan paths, as shown in Fig. 1.

Conventional heat map generation methods do not reflect the shift of gaze. This paper extends our previous work [3] and comprehensively presents a pathbreaking study on saccade information based directional heat map generation that reflects directionality of gaze shift during dynamic eye movement, as shown in Fig. 2. The proposed method is useful for a number of scientific visualization purposes in gaze tracking experiments and applications like autism detection, study efficiency assessment and advertisement design evaluation.

2. Related Works

Eye tracking data acquisition systems are built upon pupil detection algorithms. Commercial products like Tobii and SMI are typical eye trackers with proprietary software. Pupil platform[4] contains a complete eye tracking device and software for eye tracking, fixation point detection and visualization. PyGaze[5] is a gaze detection tool that can work with supported eye tracking devices.

Eye tracking data analysis mainly includes signal de-noising, fixation and saccade detection, as well as focused target detection and degree of attention [2]. Researchers [6], [7] proposed novel mathematical models to improve eye tracker accuracy.

For eye tracking data visualization, Blascheck et al. [8] gave a comprehensive review and compared large number of eye tracking data visualization techniques. They partic-
ularly mentioned the difference between static and dynamic attention maps while heat map is a kind of attention maps. Špakov et al. [9] proposed linear, sinusoidal and Gaussian models based heat map generation that considering parameters like sensitivity, brightness and hiding level. Many subsequent studies are based on this approach. Bildnaut’s heat map generation method [10] discussed color mapping. Exclusive attributes of eye trackers are not under consideration during visualization.

3. Directional Heat Map Generation Method

The proposed heat map method for eye tracking data in this work is compose of three steps: (1) gaze data pretreatment; (2) parameterization; (3) heat map generation. The input data are some gaze points \( G = (g_1, g_2, \ldots, g_n) \), each \( g_i \) includes the gaze point position \((x, y)\), the timestamp \(t\), and the detection confidence \(c\).

Before processing the data, first determine the difference between fixation and saccade. Fixation point is defined as the point at which the fixation time exceeds 100 ms. Fixation region consists of foveal region and blending region. Foveal region of fixation point is around 5°, and blending region is between 15° and 30°, see Fig. 3. Saccade time is about 30 to 120 ms, and the range is 1° to 40° [2].

3.1 Pretreatment

First of all, we need to determine the size of the display area and remove the part of the gaze data that exceeds the display area. If the range of stimuli used when collecting the gaze point is larger than the display area, then the gaze point beyond the area where the stimuli are located is removed. At the same time, ignore the less confidence gaze point.

Next, merge the gaze points. We set up a variable-length array for storing gaze points that need to be merged. When the distance between the first gaze point \( o \) in the array and the current gaze point \( g_i \) is less than a certain threshold \( \gamma \) and their time difference is less than the threshold \( \tau \), the current gaze point is pushed into the array, and the average value of all points in the array is taken as the current position of the gaze point. Then compare with the next gaze point until the distance between the two exceeds the threshold. At this time, all the points in the array are merged into a single point, and the array is cleared.

3.2 Parameterization

Parameterized frame data contains coordinates \((x, y)\), timestamps \(t\), duration \(d_c\), confidence \(c\), and other information, where the meaning of coordinates varies with the type. The coordinates represent the position of the fixation point in the fixation point type and the vectors of the saccade in the saccade type. The parameterized data has one more boolean flag is fixation to indicate whether the type of the frame data is the type of fixation point than the original data. In actual operation, gaze data pretreatment and parameterization are performed synchronously. See Algorithm 1 for computation details.

3.3 Heat Map Generation

First, we need to build a heat map \( f_{\text{base}} \) for a single fixation point as a basis. We use the model given in [10] to generate this heat map. The heat map at all times is generated after deformation and displacement on the basis of this image. The main reason for this is that the range of angles at which a person is watching is generally fixed, and doing so can greatly reduce the number of calculations. In order for the heat map to reflect as much as possible the observer’s actual area of interest, we need to determine the gaze region based on the fixation point information.

The gaze region is a circle of radius

\[
R = d \tan \frac{\alpha}{2}
\]

where \( \alpha \) is the angle of view and \( d \) is the distance between the eye and the object of view, see Fig. 3. It should be noted

**Algorithm 1 Pretreatment and Parameterization**

*Input*: gaze points \( G = (g_1, g_2, \ldots, g_i, \ldots, g_n) \)

*Output*: fixation points \( P = (p_1, p_2, \ldots, p_j, \ldots, p_m) \)

1: procedure PRETREATMENT PARAMETERIZATION(G)

2: \( v \leftarrow \text{new dynamic array} \)

3: \( o \leftarrow g_0 \)

4: for all \( g_i \in G \) do

5: if \( g_i \) in screen range

6: and \( c_i > \text{confidence threshold} \) then

7: if distance\((o, g_i) < \gamma \)

8: and time\_difference\((g_{i-1}, g_i) < \tau \) then

9: push \( g_i \) into the array \( v \)

10: else

11: \( p_j \leftarrow \text{merge point}(v) \)

12: clear the array \( v \)

13: \( o \leftarrow g_i \)

14: push \( g_i \) into the array \( v \)

15: end if

16: end if

17: end for

18: end procedure

![Fig. 3 Gaze region. Dark orange is the foveal region, light orange is the blending region. \( \alpha \) is the angle of view, \( d \) is the distance from the eye to the fixation plane, and \( R \) is the radius of the circular fixation area.](image-url)
that the units here are pixels. The following situation is similar.

The following is a visualization method for saccade data. The heat map of a saccade is a distortion from the fixation heat map. Here, we are deforming the original circular pattern into an ellipse. Specifically speaking, it is to zoom the units here are pixels. The following situation is similar.

The following is a visualization method for saccade data. The heat map of a saccade is a distortion from the fixation heat map. Here, we are deforming the original circular pattern into an ellipse. Specifically speaking, it is to zoom data. The heat map of a saccade is a distortion from the fixation heat map. Here, we are deforming the original circular pattern into an ellipse. Specifically speaking, it is to zoom in two directions first, then rotate the corresponding angle, and then pan. See Algorithm 2 for computation details.

We assume that saccade vector \( \mathbf{v} = [v_x, v_y] = P_{j-1}P_j \).

The scaling factor along the \( x \) axis and the \( y \) axis are \( S_x \) and \( S_y \) respectively. The lengths of the major and minor axes of the scaled ellipse are \( a_{\text{major}} = 2S_xR \) and \( a_{\text{minor}} = 2S_yR \) respectively. Here we hope that when \( ||\mathbf{v}|| = 0 \), the shape of the saccade will return to the circle of the fixation point, namely \( a_{\text{major}} = 2R \) and \( a_{\text{minor}} = 2R \); when \( ||\mathbf{v}|| \to +\infty \), the shape of the saccade is close to a line, that is \( a_{\text{major}} \to +\infty \) and \( a_{\text{minor}} \to 0^+ \).

Since the major axis represents the direction of the saccade vector, the major axis must be linear with \( ||\mathbf{v}|| \). The major axis can then be deduced as

\[
a_{\text{major}} = ||\mathbf{v}|| + 2R. \tag{2}
\]

We can think of the minor axis having a negative exponential relationship with \( ||\mathbf{v}|| \), that is \( a_{\text{minor}} = 2R \times e^{-k||\mathbf{v}||} \), \( k \) being a constant. At this point we consider the case of \( ||\mathbf{v}|| = 2R \), as shown in Fig. 4. Assume that minor axis length is \( 2\lambda R \) at this time. Then there must be \( 0 < \lambda < 1 \). So we can get \( k = \lambda^{-1/2}R \), and the minor axis is

\[
a_{\text{minor}} = 2R \sqrt{||\mathbf{v}||^2 - 2R^2}, \quad \lambda \in (0, 1). \tag{3}
\]

It can be seen that the smaller the \( \lambda \), the shorter the corresponding minor axis and the stronger the directionality.

The rotation angle of the ellipse is

\[
\theta = \arccos \frac{v_x}{||\mathbf{v}||}. \tag{4}
\]

Then pan the rotated image to the middle point of \( v \). Use \( T \) to represent the transformation matrix, which contains the operations of scale, rotation, and pan. As show in Fig. 5 the deformation result is

\[ \frac{c}{S_xS_y}T \mathbf{f}_{\text{base}}. \tag{5} \]

After obtaining the distorted heat map, we obtain the new image \( \mathbf{h}_k \) by summing up the continuous \( \delta \) images. Duration of accumulated images is typically 300 to 1000 ms, that depends on frame rate of eye tracker and \( \delta \) value. Find the maximum value in the image \( \mathbf{h}_k \) and divide the values of all the pixels by the maximum value, multiply by 255, and then round them. The value is in the interval \([0, 255]\) so that a single channel heat map is obtained. We can map it to the color space, to generate the general color heat map. We can also add a transparency property to it, let it superimpose it into the original scene, so that we can see more intuitively the object that the observer pays attention to. See Algorithm 2 for computation details.

### 4. Heat Map Generation Results

The proposed method is implemented by Python using OpenCV and SciPy library. We compare proposed method with conventional method. Figure 6 shows the heat map generated in the real scene. As shown in Fig. 7, when the
Fig. 6 Generated heat map. (a) The stimulus; (b) the conventional method; (c) the proposed method.

Fig. 7 Generated heat map. Left: the fixation point moves slowly; right: the fixation point moves rapidly.

fixation point moves slowly, there is no significant difference. When the fixation point moves rapidly, the points in conventional method are scattered and lack continuity. The proposed method can clearly show the trajectory of the fixation point.

For a long time static heat map, the heat map generated by this method is not much different from the heat map generated by conventional method. The proposed method has advantages in short term dynamic scenarios.

5. Conclusion

Proposed method in this work considers combination of fixation point and saccadic data structures, and add heat map deformation model for saccade type data. For eye tracking data, proposed heat map shows good directivity. Scientific visualization tasks and experiments in eye tracking can benefit from this work.

In future we will study parameter selection. Parameters such as $\alpha$ could be relevant to eye movement tasks. Different eye tracking applications based on proposed method will also be considered.

Acknowledgments

This research is supported by Science and Technology Planning Project of Guangdong Province (No. 2017B010110007 and 2017B010110015), Natural Science Foundation of Guangdong Province (No. 2018A030313802), National Natural Science Foundation of China (No. 61872365) and Beijing Natural Science Foundation (No. 4182065).

References


