1. Introduction

We can assume that when people attempt to get information of an object in which they are interested, they naturally gaze at it. Hence if we can get information about people’s gaze directions, we can also get information about objects in which they are interested. Information about people’s interests is very useful in many scenes of our life. For example, we can advertise items to a person who is interested in them, and predict a shoplifter who gaze at not goods but security cameras or store clerks. Therefore gaze information is very useful in our life.

There are many ways to obtain gaze direction. One of them is a method of measuring movement of eyeball directly. For example, Ville et al. use EMR-9 as eyetracker produced by NAC to measure gaze direction. Konno et al. installed camera under the display and observed user’s eye directly to get gaze direction. By using the method, we can gain gaze direction accurately. However, a range of space that we can obtain gaze direction is limited because we have to measure human’s eyeball directly. Therefore if many people are in a wide range of space like town, it is difficult to get their gaze information. In addition, we can use the method only a situation that observer is observed eyeball movement actively. This situation is not suitable for the purpose of crime prevention.

On the other hand, we can gain gaze direction indirectly by estimation method without measuring eyeball directly. This method uses images taken by a surveillance camera like security camera. In this way, a range of space that we can obtain people’s gaze information is wider than the method of measuring eyeball directly. In addition, the number of observer we can get gaze information at the same time is larger than the way of measuring eyeball directly. Moreover, we can gain gaze information without observer’s consciousness. Therefore, this way is suitable for the purpose of advertisement and crime prevention.

However, if we try to gain gaze information, it is hard to observe the motion of eyeball by the way of using a surveillance camera. Hence many researches define head direction as gaze direction. Certainly, people tend to turn their face to a direction on which they try to concentrate, but the definition is not correct. If we hypothesize that head direction and gaze direction is always equal, we have to consider that people try to set gaze direction without moving their eyeball. Therefore, We observed three situation, to checked whether head and gaze direction is not always equal. First, we gaze at front direction of the body Figure 1, second right angle of 45 degree Figure 2, third right angle of 90 degree Figure 3. We can confirm from these figures that gaze direction is sum of eyeball direction and head direction, so head and gaze direction is equal when we gaze at front direction of the body, but it is not equal when we gaze at right angle of 45 and 90 degree. Therefore, head direction and gaze direction are not always equal. Hence, we conclude that this estimation method, regard head direction as gaze direction, is not rigorous.

In our research, we define a gaze estimation model which is considered relationship between head direction and gaze direction. By using the model, it is possible to estimate more accurate.

Many studies have reported that head direction and gaze direction relate to each other when people gaze a target. Stahl et al. have reported that when sitting people gaze at a fixed target which is displayed in their visual field, a bias of their head and eyeball movement keep linear relationship. In addition, Okada et al. indicate that walking people’s head and eyeball relation is similar to sitting people’s one. Moreover, Fang et al. point out the similar relation in the case of gazing at a moving target.

In addition, relation between the movement of eyeball and head is affected by neural control. When people shift their
gaze direction from one target to another target, eyeball rapidly moves in a direction of new target, and a little later, head moves slowly in the same direction. It is called sac-cade. Once the gaze reaches the target, head moves in the same direction continuously but eyeball moves in an opposite direction. These relation of movement is called eye-head coordination.

In our research, we propose a gaze estimation model which is considered these relationship between eye and gaze. We observe subjects who behave naturally, to correct data of head and gaze directions, and we estimate gaze direction by using these data. Finally, we evaluate an accuracy of estimation result by comparing the result of our proposed method and the method which regard head direction as gaze direction with measured gaze direction, and confirm that accuracy of estimation is improved and our proposed model is effective.

We explain about eye-head coordination particularly in section2, propose a method to estimate gaze direction by using head movement in section3, and evaluate an accuracy of estimation in section4.

2. Eye-head coordination

Human beings shift their gaze direction by moving eyeball and head. Then, eyeball and head motion is not simple, for example they move linearly in the direction of new target at the same speed, but complex like Figure 4. As you can see from this graph, In the start phase of shift, eyeball move in the direction of new target rapidly. This motion, called sac-cade, is happened in order to catch a image of the gaze target in the center of the retina. After that, head move in the same direction a little later just as if head chase eyeball motion. The speed of head motion is slower than it of eye motion. In the end phase of shift, head move in the same direction continuously but eyeball move in the opposite direction. It is neural control of eyeball movement, called vestibulo-ocular reflex (VOR). If eyeball move together with head or body, a visual field shake. Therefor it is made unclear. To prevent this phenomenon, eyeball is controlled to move in an opposite direction by neural nervous system.

![Fig. 4 Eye-head coordination](image)

People’s head and eyeball have composite motion, like the above one. This relation is called eye-head coordination. Human beings do it reflexively. Therefor we assumed that we can estimate gaze direction accurately by considering this coordination. We explain about methods in next section.

3. Method

In this section, we proposed a method of gaze estimation considering the eye-head coordination. We propose a model which can estimate gaze direction accurately by using head direction in each frame of sequence. The model enable us to gaze estimation considering the eye-head coordination. By inputting the head directions to the model, we gain estimation gaze directions as output.

3.1 Gaze estimation model

Gaze direction and head direction in Figure 4 shows us the phenomenon, at first gaze start to shift to the target, after a while, head start to move in the same direction. Head speed is slower than gaze one. This relation, head motion is pulled by gaze motion with time lag, looks like of elasticity of a spring. Checking gaze and head directions which are actually measured, we can observe behavior that gaze starts to move earlier than head in many timing. Figure 5

![Fig. 5 Relation between gaze and head](image)

Therefor, we consider the model that two balls are connected by a spring Figure 6. We regard one ball, orange color, as gaze direction, and the other, blue ball, as head direction. If we grasp an orange ball and move rapidly, then a blue ball does not move at first due to elasticity of a spring. A short time later, a blue ball is pulled by a spring and start to move. Moreover, if we stop moving an orange ball, a blue ball also stop a little later. These movement is the same as a relation between gaze direction and head direction, as you can see in Figure 4. Hence we can carry out gaze estimation which is considered eye-head coordination by using this model.

We represent these balls relation by equation of motion. We defined continuous and discrete system of chest, head and gaze direction as below.

(1) $h(t) \cdots$ A function of direction of head to chest
(2) $g(t) \cdots$ A function of direction of gaze to chest
(3) $e(t) \cdots$ A function of direction of eyeball to head
(4) $H = [h_1, h_2, \ldots, h_t]$
(5) $G = [g_1, g_2, \ldots, g_t]$
(6) $E = [e_1, e_2, \ldots, e_t]$

We define natural length of a spring as $l$, stretch of a
spring as $\Delta x$. Moreover, we consider a damper which is in proportion to velocity. We regard Coefficient of viscosity as $\lambda$. The equation of motion of a blue ball is as follow.\

粘性抵抗: $\lambda x'(t)$

\[
\begin{align*}
F &= m h''(t) = k \{g(t) - h(t) - l\} - \lambda h'(t) \\
\Longleftrightarrow g(t) &= a h(t) + b h'(t) + c h''(t) + d
\end{align*}
\]

(1)

Direction of eyeball to head is the difference between gaze direction and head direction. Therefore, we can transform 1 to 3.

\[
e(t) = g(t) - X(t)
\]

(2)

\[
e(t) = a h(t) + b h'(t) + c h''(t) + d
\]

(3)

We defined this linear equation as gaze estimation model. Input is head direction and output is estimation gaze direction.

3.2 Define coefficients of gaze estimation model

The gaze estimation model, I proposed in the last section, can be changed into matrix 4 by using series of sequence which is constructed $T$ images.

\[
\begin{bmatrix}
g_1 \\
\vdots \\
g_T
\end{bmatrix} = 
\begin{bmatrix}
h_1 & h'_1 & h''_1 & 1 \\
\vdots & \vdots & \vdots & \vdots \\
h_T & h'_T & h''_T & 1
\end{bmatrix}
\begin{bmatrix}
a \\
b \\
c \\
d
\end{bmatrix}
\]

(4)

We define matrix $A$ 5, and calculate pseudo inverse matrix $A^+$. We can calculate coefficients $a, b, c, d$ by multiplying both side of equation by $A^+$ from left side.

\[
A = 
\begin{bmatrix}
h_1 & h'_1 & h''_1 & 1 \\
\vdots & \vdots & \vdots & \vdots \\
h_T & h'_T & h''_T & 1
\end{bmatrix}
\]

(5)

\[
\begin{bmatrix}
a \\
b \\
c \\
d
\end{bmatrix} = A^+ 
\begin{bmatrix}
g_1 \\
\vdots \\
g_T
\end{bmatrix}
\]

(6)
Human eyeball always move slightly if they gaze at one point. It is called involuntary eye movement during fixation. There is a possibility of including outliers in the series of sequence. If we calculate coefficients without considering outliers, the result of calculation is affected by them. Therefore, we remove outliers in advance by following the steps described below, and after that, we calculate coefficients.

1. Pick up $n$ frames randomly from a series of sequence, and create matrix $A$.
2. Calculate coefficients as described above and coefficients calculated in (2) step.
3. Calculate estimation gaze directions by using head directions and coefficients calculated in (2) step.
4. Count the number of frames that the error between ground truth, it is measured in actuality, and estimation gaze direction are under $x$, and define frames the error is over $x$ as outliers.
5. Keep frames of outliers if the value of count which is counted in (4) step is higher than before.
6. Repeat steps (1) to (5) $m$ times, and after that, decide frames of outliers. 
7. Calculate coefficients as described above by using frames those are defined as outlier in step (6).

We decide coefficients those are calculated in final step as coefficients of gaze estimation model. We use the model to carry out estimation. In our research, we set $n = 5$, $m = 1000000$, $x = 10$.

4. Experiment

5. Conclusion

In our research, we defined gaze estimation model which is considered relation between eyeball and head movement, called eye-head coordination, and carry out gaze estimation by using the model. When we shift our gaze direction, at first, eyeball start to move rapidly to a new target. After that, head move in the same direction with slow speed. After eyeball catch the target, head move continuously but eyeball move in the opponent direction. Those movement of eyeball and head looks like two balls’ movement connected by a spring. Therefore, we define a motion equation of the relation between two balls as a model to carry out gaze estimation. By inputting head directions to the model, we gain gaze estimation direction as output. We evaluate difference, both in the case of using our proposed method and regarding head direction as gaze direction, from the ground truth. In horizontal direction, the result of gaze estimation direction’s accuracy is higher than the method regarding head direction as gaze direction. On the other hand, in vertical direction, the accuracy is not improved. Therefore, at this point in time, we can regard head direction as gaze direction in vertical direction.

We would like to point out the above two future work. First, we analyze the similarity of regression coefficients. We calculate coefficients vary from scenes to scenes now. However, if we can find the similarity of them between individuals or between behaviors, we can fix coefficients and estimate in real time. Therefore, we need to make an experiment under the situation that subjects are instructed in detail.

Second, we consider discretizing gaze signal before calculating coefficients. Saccade is occurred momentarily, and involuntary eye movement during fixation is occurred regardless of human’s intention. Therefore, we discretize gaze signal to left these movement out of consideration.

We solve these problems, and make possible to estimate gaze direction more accurately.

References