Vergence eye movements elicited by non-disparity factors in 2D realistic movies

Iijima, Atsuhiko[1,2]; Kiryu, Tohru[3]; Ukai, Kazuhiko[2*]; Bando, Takehiko[1*]
[1] Div. Integrative Physiol., Grad. Sch. Med. & Dent. Sci., Niigata University,
1-757 Asahimachi-dori, Chuo-ku, Niigata, 951-8510 Japan
[2] Dept. Appl. Phys., Sch. Sci. & Tech., Waseda University,
3-4-1 Okubo, Shinjyuku-ku, Tokyo, 169-8555 Japan
[3] Div. Inform. Sci., Grad. Sch. Schi. & Tech., Niigata University,
8050 Ikarashi-2, Nishi-ku, Niigata, 950-2181 Japan
[1,2]a-iijima@med.niigata-u.ac.jp, [2*]ukai@waseda.jp, [3]kiryu@eng.niigata-u.ac.jp,
[1*]bando@adm.niigata-u.ac.jp

Keywords: vergence eye movement, disparity, optic flow, 3D/2D, realistic motion picture, pictorial cues.

Abstract. We have analyzed characteristics of vergence eye movements while subjects watch three-dimensional (3D) and two dimensional (2D) movies, and found that 2D images without binocular disparity sometimes evoked convergence similar to those found under 3D condition. In this study, we investigated factors that drove convergence other than binocular disparity. We presented a roller coaster movie, which was taken by the cameras set on the top of the coaster for comparison of vergence under different movie conditions (3D or 2D). The motion vectors localized to their gaze points were calculated, and the optic flow was estimated based on flame-by-flame analyses of images. The traces of vergence eye movements under 2D condition consisted with the changes in the optic flow. It is suggested that most of factors that drove convergence in 2D movies are closely related to the optic flow.

Introduction

Vergence eye movements, such as convergence and divergence are used in our daily life to detect objects in three-dimensional space. The two eyes disjunctively move, and the movement is relatively slow. In real visual conditions, vergence is driven by binocular disparity, near vision (proximal), image blur, and other depth cues in order to fuse the two images from the right and left eyes and to perceive depth.

Several studies have looked at vergence eye movement, and several vergence models based on experimental approaches have been developed (Krishnan and Stark, 1977; Erkelens and Collewijn, 1985; Semmlow *et al*., 1986; Pobuda and Erkelens, 1993; Hung *et al*., 1994; Schor, 1979). Most of the previous studies investigating vergence used a random dot stereogram (Erkelens and Collewijn, 1985; Pobuda and Erkelens, 1993) or a simple visual pattern for stimulation (Krishnan and Stark, 1977; Semmlow *et al*., 1986). Vergence while viewing a real-motion movie has yet to be fully investigated.

In this study, we used real movie stimuli projected on a large screen. Under controlled binocular disparity movie conditions, human subjects were shown 3D- and 2D- movies, respectively. Eye movement during the movies was analyzed using binocular video-based oculography. Binocular disparity and motion components of the movies were also analyzed using image processing techniques. The aims of the study were to understand, (1) the time course of vergence in humans watching a real movie, (2) the differences in vergence variations with or without binocular disparity, and (3) effects on vergence with non-disparity vergence driving cues and how multi cues work.

Methods

Measurements of eye and head motion. We measured eye movement and pupillary response with a binocular video oculography (ET-60-L, Newopto, Japan), which included compact CCD cameras with infrared sensitivity situated in head-mounted goggles. The goggles had openings in front of the eyes and half-mirrors that enabled the subjects to watch anything within their visual field, and we could simultaneously monitor their eye movement.

Eye images (NTSC video signals, 30fps) from the CCD cameras were captured via an image digitizing board and entered into a PC system (LabVIEW 8.2, National Instruments, USA). The PC calculated eye position and pupil diameter using binary images in which the pupil area was displayed in black and the other areas in white. Eye position data (horizontal and vertical) were calculated by tracking the center of the pupil coordinates. Vergence eye movement was calculated by subtraction of the right eye position from the left eye position.

Because the subject sat on a chair without any head-fixation apparatus, head motion and eye position were simultaneously measured for an accurate analysis of gaze points on the screen. Head movement was measured using an electro-magnetic sensor technique (Liberty, Polhemus, USA).

Movie projection. Two liquid crystal projectors (TH-L795J, 700 ANSI lumen, Panasonic, Japan) were used to show movies on an 80-inch (64 × 48 inch) transmissive screen, located 2.0m in front of the subject and with an image field consisting of a 44×35 deg arc.

We used a 200-sec long live-action movie. The movie field-sequentially had binocular disparities, and a movie splitter (STG-3D, Nisho-electronics, Japan) divided the movie into pictures for the right/left eyes. When the split mode was off, 2D movies were available. The movie was filmed using two cameras set on top of a roller coaster car. The angle between the two cameras was constant but the binocular disparities on the screen varied depending on the distance between the camera and object. The movie gave the perspective from the front view of a roller-coaster. In this way, the subjects were able to experience virtually the riding atmosphere on the coaster.

Movie component analysis. Motion and disparity components of the movie were calculated using PC software we developed on MATLAB ver.7.1 (The MathWorks Inc., USA). Our software computed horizontal/vertical (pan and tilt) and zoom (expansion and concentration) components in the movie as camera parameters. Using a pattern-matching algorithm, we first obtained local motion components for each block, which was divided into 352 × 288 parts. Next, we computed the whole screen motion as Global Motion Vectors (GMV) from integration of all the local motion components. The GMV had three components (pan, tilt, and zoom) representing camera motion (Wang and Adelsen, 1994). The zoom component was generated especially when the lens of the camera was closing in on an object or the camera itself was moving forward, which may be equivalent to radial expanding optic flows. We developed other software on MATLAB to obtain the numerical values of binocular disparities within the region where subjects focused on the screen. We calculated gaze point binocular disparity (GPBD) with the gaze point data and the binocular disparity data. We defined the ROI for calculating GPBD (5×5 (deg)). The area size was necessary for conducting calculations and on the ground of the visual field of the subject.

Experimental protocol. Eight young healthy subjects (males 4/ females 4; age range 19 to 36 yrs) with normal or corrected-to-normal visual acuity and oculomotor function for their age participated in the experiments. The experimental room was under controlled illumination (10 lux) and temperature (22°C). All subjects signed a consent form voluntarily after full explanation of the experimental methods and procedures and were free to withdraw at any time. The research followed the tenets of the Declaration of Helsinki and was approved by the human experimentation committee of Niigata University School of Medicine.

After scotopic adaptation (approximately 10 min), 5-min rest period with no movie projection was established pre and post movie exposure. The roller coaster movie was 200 sec in duration. Subjects were randomly exposed to the 2D and 3D movies that were shown at designated times (from 9 AM to 12 PM). Before and after the movie presentation, eyes were calibrated using five fixation targets in crossed positions that were projected on the screen.

Results

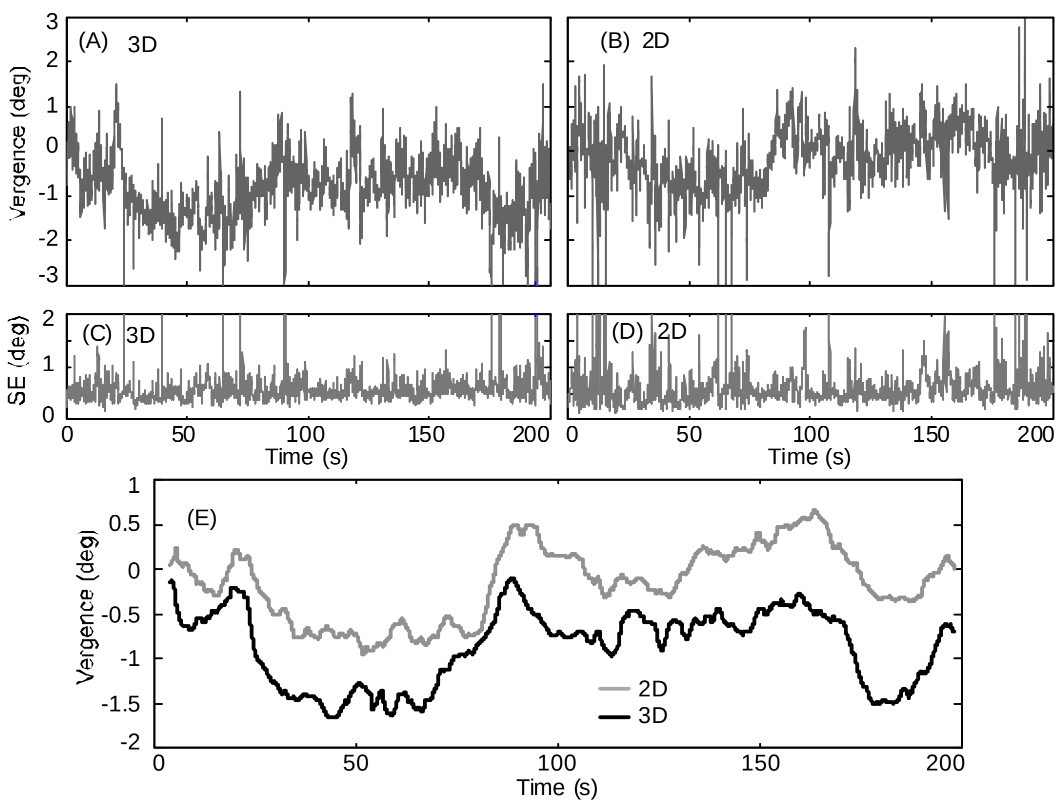

Fig.1 Vergence eye movements under 3D and 2D conditions.
(A): mean 3D vergence; (B): mean 2D vergence;
(C): SE of 3D vergence; (D): SE of 2D vergence;
(E): trend of vergence with a low-pass filter (cutoff 1Hz).

The average vergence eye movements of all subjects while watching the movie on 3D mode and 2D mode, respectively, are shown in Fig.1A and Fig.1B We defined the vergence angle as follows. Zero (deg) means that the subject fused an object on a screen. Positive and negative values mean that the cross point of their eye spot was in front of and on the backside of the screen. When vergence angle was positive and negative, eyes were relatively convergence and divergence from the screen position, respectively. The SEs for the mean vergence data are shown in Fig.1C and Fig.1D for 3D and 2D vergence, respectively. The both 3D and 2D shapes of vergence eye movements were almost parallel, but different from baseline. The range of the amplitude of vergence was wider for 3D movies than for 2D. The baselines of the vergence traces were relatively negative, because negative (uncrossed) disparity was major in this movie from start to finish. Fig.1E shows the trend of the vergence from Fig.1A and Fig.1B.

We show the subtraction of 2D vergence from 3D vergence, defined as follow called V_sub (1) and mean value of GPBD (2) in Fig.2A and Fig.2B, respectively.

$$\text{V\_sub} = \text{Vergence\_3D} - \text{Vergence\_2D} \quad (1)$$

$$\text{GPBD\_mean} = \text{GPBD\_positive} + \text{GPBD\_negative} \quad (2)$$

The cross correlation between V_sub and GPBD_mean shown in Fig.2C, clearly represents the effect of binocular disparities on vergence. In this analysis, the cross correlation in each 3s-time window was calculated for every 1 sec. Meshed areas show correlation coefficients over 0.7.

Moving on 2D vergence, we analyzed the effect of the zoom component of global motion vector to vergence system. Figure 3A shows the 2D vergence trace and Fig.3B shows the strength of the radial expanding component of the GMV. Figure 3C shows the result of cross correlation between 2D vergence and the optic flow element. The way of this correlation analysis was same as Fig.2C and meshed area also means the correlation coefficients were over 0.7. In this analysis, there were scenes with inability to calculate cross correlation

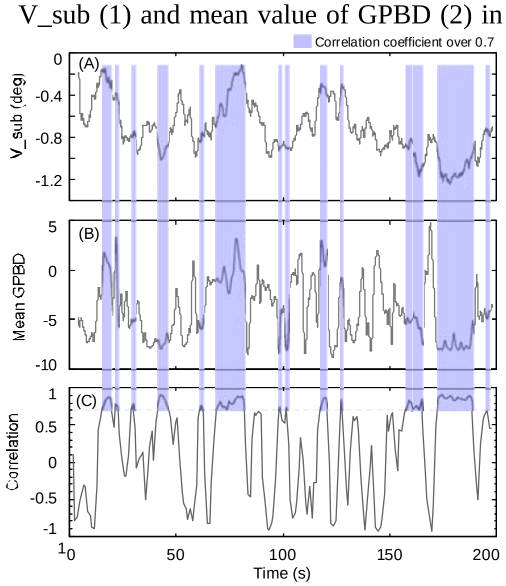

Fig.2 Relationship between V_sub and GPBD. Time course of V_sub (A) and mean GPBD (B). (C): Cross correlation value between V_sub and mean GPBD. Color bars mean correlation coefficient was over 0.7.

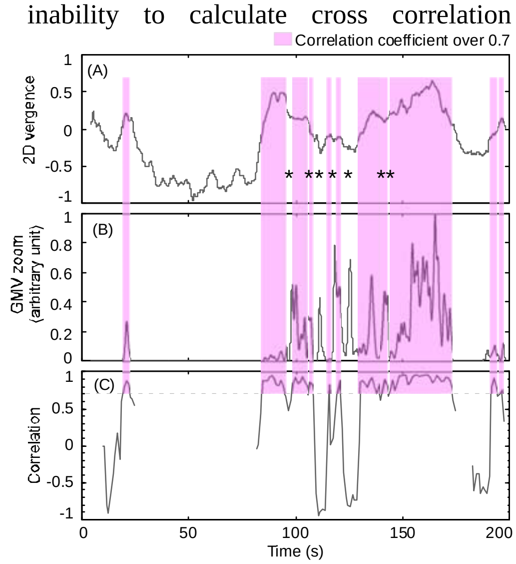

Fig.3 Relationship between 2D vergence and GMV_zoom. Time course of 2D vergence (A) and the strength of GMV_zoom (B).
(C): Cross correlation value between 2D vergence and GMV_zoom. Color bars mean correlation coefficient was over 0.7.

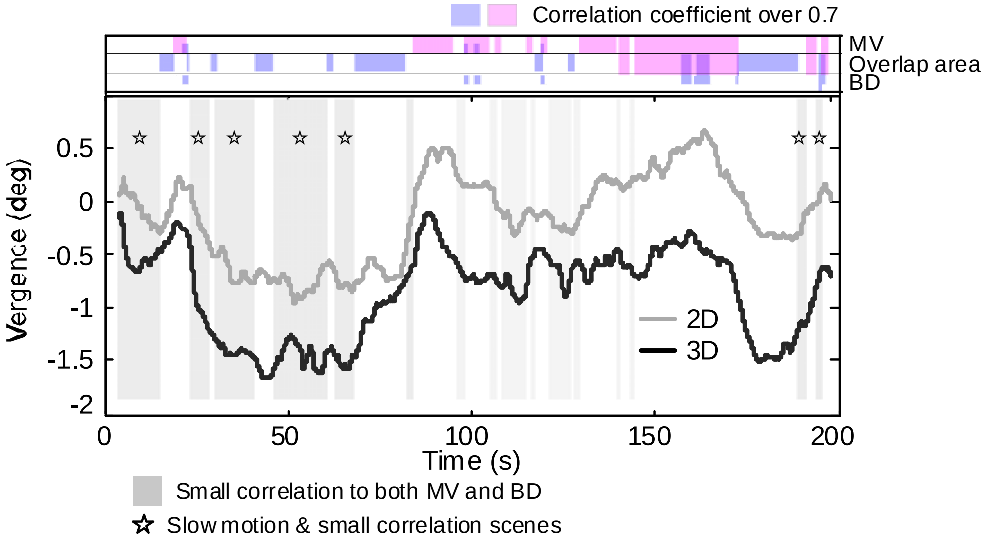

Fig.4 How did the two factors work for driving vergence?

because of the scenes in which the power of the radial expanding factor was zero. Asterisks indicate the scenes without correlation to the radial expanding component although the component was not zero.

Summing up the cross correlation analyses, binocular disparity and optic flow were identified as the driving factors of vergence. The top chart in Fig.4 indicates that which cue affected vergence system. Gray bar means highly-correlated scenes to disparity factor and motion factor, respectively and scenes with double-dense colored over lapping area might be controlled by the double factors. In contrast, non-colored scenes, pointing with half-tone dot meshing in the bottom figure were with small correlation to both motion and disparity components. Star marks in Fig.4 especially indicate the scenes without dynamic motion in the meshing groups. Almost all scenes were conveyed using binocular disparity and optic flow.

Discussion

Stereoscopic 3D movies with large binocular disparity are expected to evoke vergence, while 2D movies, which have no binocular disparity, should not. We found, however, that vergence was also dynamically changing in subjects viewing 2D films. In this study, subjects viewing both 3D and 2D films had almost similar variations in vergence throughout the movie presentation (Fig.1). Subjects were shown a roller coaster movie that was 200 sec in duration. From the start of the film to 80 sec and from 160 sec to the end point, vergence slanted toward divergence. On the other hand, from 80 to 160 sec, vergence shifted to the convergence side and dynamically varied. Normally, a real approaching object is necessary to elicit convergence eye movements for fusing on the retinas. To explore the vergence observed during both 3D and 2D movies, we focused on dynamic characteristics such as binocular disparity and radial optic flow to determine their potential to evoke vergence. To do this, we proposed a model in which vergence is elicited by dynamic visual input to clarify the driving factor behind vergence during 2D movie.

Since binocular disparity was one of the easier factors to evoke vergence, we began by calculating the local disparity for areas at which the subjects gazed. To next isolate and investigate the effects of binocular disparity on vergence eye movement in 3D conditions, we subtracted 2D vergence from 3D vergence and analyzed the results in Fig.2. From this stance, 2D vergence included anything that elicited vergence, except binocular disparity. We, therefore, dealt with 2D vergence as a control of 3D vergence especially for binocular disparity.

Under 2D conditions, there was no binocular disparity information at all. So what drove vergence under these conditions? Depending on if the conditions were with or without disparity, the baselines of the vergence angle differed between 3D and 2D (Fig.1). In a static 2D picture representing a real scene, there were several guides for depth perception. These cues were not necessarily equal to vergence driven factors, but they could potentially elicit vergence (Enright, 1987a; Enright, 1987b). In contrast to these static cues, optic flow is considered a dynamic cue for evoking vergence eye movement. Optic flow creates a self-in-motion sensation, especially with radial optic flow, which is achieved through forward motion scenes or zoom-in scenes (Cornilleau-Peres and Gielen, 1996). Radial expanding optic flow is identical to the zoom component of GMV, since both represent the global motion of a viewer or a camera. Under special conditions, radial optic flow stimuli evoke very short latency convergence eye movements (Busettini *et al*., 1996; Busettini *et al*., 1997; Yang *et al*., 1999; Miles, 1999). Since our roller coaster movie showed continuous forward motion, successive radial expanding optic flows arose and might have elicited continuous vergence.

We investigated the relationship between 2D vergence and the strength of the radial expanding component of the GMV in Fig.3. It is obvious that vergence angles from the 2D films were on the convergence side when the GMV was in use. On the other hand, when the power of the radial expanding component was weak, vergence angles were reduced. It seemed to release from the convergence tonus. Thus, radial expanding image flow might cause the dynamic variations of convergence in 2D through a continuous input of optic flow information that refreshes the previous information, allowing the vergence system to be maintained without oscillation.

From the cross correlation results in Fig.4, binocular disparity and optic flow were identified as the driving factors of vergence during movie presentation. The upper bars clearly shows when these cues work. Almost all scenes were conveyed using binocular disparity and optic flow. Small correlation areas described as meshed bars overlaying vergence traces might not be affected by both binocular disparity and optic flow component. Star marks represent that the power of optic flow was almost zero in these non-explainable scenes. Focusing on the vergence manners before and after the star mark areas, continuity of vergence was maintained. This could be explained by two types of theory; one is the existence of the remaining effect of the previous status and the other is another cue except binocular disparity and motion vector works for vergence system. Since star mark scenes were included in motionless scenes, vergence driving cues other than motion and binocular disparity components might operate. When visually inspected, these scenes indeed included expressions on perspective and objects in the far distance. Although the amplitudes of 2D vergence were smaller than for 3D vergence, the trends and variation patterns were similar, implying that a virtual 3D inner world can be created using non-binocular disparity cues, to reproduce binocular disparity information.

In this study, we found that vergence driven cues such as binocular disparity, optic flow, and perspective factors evoked vergence eye movements. In our framework for the interaction of vergence driven cues, the common cues for both 3D and 2D conditions, optic flow and perspective factors, could work under any conditions, while binocular disparity alone should work under only 3D conditions. Binocular disparities in the movie were determined on only the distances between objects and the cameras. In a 2D movie, the viewer estimates the distances to objects and can enable this perception of distance without any difficulty. The viewer might feel the binocular disparity information indirectly, even in non-disparity 2D conditions. Several vergence driven cues are available for conveying the virtual real world on a screen. Multi cues may work in an averaged or in a coordinated manner (Buthoff and Mallot, 1988; Howard and Rogers, 1995).

Conclusion

Based on exposure to movies containing the same content but varied disparity (3D and 2D), we confirmed how vergence driven factors create vergence under both 3D and 2D conditions. Two main cues for vergence occurrence, radial optic flow and binocular disparity, may work concertedly. Radial optic flow may thoroughly elicit dynamic vergence, even in non-disparity conditions. Binocular disparity may play several roles. For example, it may fix the baseline of the vergence variations. Another role may be its function as a strong and precise cue to catch an object in the foveal vision and for space recognition. For depth perceptions from this kind of an attraction-type movie that created the strong sensation of self-motion and speed experienced on a real roller coaster attraction, optic flow and binocular disparity may account for most of the segments of vergence, whereas the remaining unexplained segments of vergence may be driven by other pictorial cues. Especially driving motion that gives a sense of speed, continuous picture flow could sustain the previous driving effect to the successive vergence eye movement.

Acknowledgements

This study was subsidized the Japan Keirin Association through its Promotion funds from KEIRIN RACE and was supported by the Mechanical Social Systems Foundation and the Ministry of Economy, Trade and Industry.

This study was also supported by Tsukada medical foundation, Nakayama Foundation for Human Science and NS foundation for Sensory Science (A.I).

References

Busettini C., Miles F.A. and Krauzlis R.J. (1996) Short-latency disparity vergence responses and their dependence on a prior saccadic eye movement. J Neurophysiol., 75(4), pp.1392-410.

Busettini C., Masson G.S., and Miles F.A. (1997) Radial optic flow induces vergence eye movements with ultra-short latencies. Nature, 390(6659), pp.512-5.

Buthoff H.H. and Mallot H.A. (1988) Integration of depth modules: stereo and shading. J optical society America, A5, pp.1749-1758.

Cornilleau-Peres V. and Gielen C.C. (1996) Interactions between self-motion and depth perception in the processing of optic flow. Trends Neurosci., 19(5), pp.196-202.

Enright J.T. (1987a) Art and the oculomotor system: perspective illustrations evoke vergence changes. Perception, 16(6), pp.731-46.

Enright J.T. (1987b) Perspective vergence: oculomotor responses to line drawings. Vision Res., 27(9), pp.1513-26.

Erkelens C.J. and Collewijn H. (1985) Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. Vision Res., 25(11), pp.1689-1700.

Howard I. and Rogers B.J. (1995) Binocular vision and stereopsis. Oxford, UK: Oxford University Press, pp.427-460.

Hung G.K., Ciuffreda K.J., Semmlow J.L. and Horng J.L. (1994) Vergence eye movements under natural viewing conditions. Invest Ophthalmol Vis Sci., 35(9), pp.3486-3492.

Krishnan V.V. and Stark L. (1977) A heuristic model for the human vergence eye movement system. IEEE Trans Biomed Eng., 24(1), pp.44-49.

Miles F.A. (1999) Short-latency visual stabilization mechanisms that help to compensate for translational disturbances of gaze. Ann N Y Acad Sci., 871, pp.260-271.

Pobuda M. and Erkelens C.J. (1993) The relationship between absolute disparity and ocular vergence. Biol Cybern., 68(3), pp.221-228.

Schor C.M. (1979) The relationship between fusional vergence eye movements and fixation disparity. Vision Res., 19, pp.1359-1367.

Semmlow J.L., Hung G.K. and Ciuffreda K.J. (1986) Quantitative assessment of disparity vergence components. Invest Ophthalmol Vis Sci., 27(4), pp.558-564.

Wang J.Y.A., and Adelsen E.H. (1994) Representing moving images with layers. IEEE Trans Image Processing., 3(5), pp.625-638