Time-Frequency Structure of Image Motion Vectors under Vection-Induced Cybersickness

Tohru Kiryu, Eri Nomura*

Graduate School of Science & Technology
Niigata University
8050 Ikarashi-2, Niigata, Japan
kiryu@bc.niigata-u.c.jp

Astuhiko Iijima, Takehiko Bando*

Graduate School of Medical and Dental Sciences,
Niigata University
757 Asahimachi-dori-1, Niigata, Japan

*Center for Transdisciplinary Research, Niigata University

Abstract

Virtual reality (VR) is now applying to biomedical engineering, and penetrating rehabilitation engineering. However, there is a problem called cybersickness which would prevent the applications of VR. We have investigated the influences of vection-inducing images on the autonomic regulation by analyzing biosignals, and using the image motion vectors to quantify video image scenes. The results showed a specific frequency band (0.3–2.5 Hz) of motion vectors, which were concentrated in time and on a screen, would induce cybersickness.

1 Introduction

Development of digital imaging technology by computer graphics is producing many image formats, resolutions, frame rates, in addition to conventional factors such as screen size, brightness, color, and display devise. Contrary to the wider flexibility and commercialization, digital imaging technology is sometimes spreading unexpected visual stimuli. For broadcasting video images, Harding (1998) analyzed video image factors that evoked photosensitive seizures and recently produced a screening instrument. He pointed out that flashing and some specific patterns had bad influences on the brain. Practical problems are also emerging especially in applications of the virtual reality (VR) or the virtual environment (VE). The VR is now applying to biomedical engineering, and penetrating rehabilitation engineering (Kenyon *et al.,* 2004), but, there are some problems in relation to unpleasant sensations and aftereffects due to visual stimuli. Stanney *et al.* (1999) reviewed human factor issues in VEs. Nichols *et al.* (2002) pointed out health and safety implications of VR to make recommendations regarding the future direction of VR. However, details in relation to cybersickness are still unknown.

Cybersickness is an unpleasant sensation under visually-inducing illusions of self-motion. It has been reported that the mismatch between the visual system and vestibular system causes cybersickness (sensory conflict theory) (Hettinger *et al.,* 1990). Unpleasant sensation has been assessed by autonomic-nervous-activity-related indices estimated from biosignals including heart rate, blood pressure, finger pulse volume, respiration rate, skin condition, and gastric myoelectrical activity (Cowings *et al.,* 1990; Peters, *et al.,* 2000; Gianaros *et al.,* 2001). There is a large difference in time-scale between autonomic regulation and sensory systems. Actually, sensory activities work within a few tens of milliseconds, whereas autonomic nervous regulation takes several seconds (Figure 1).

In this paper, we mainly measured a subject's electrocardiogram (ECG), blood pressure, and respiration, while he/she was watching vection-inducing video images. We quantified vection-induced video images by the motion vectors that are used in image data compression. As a similar

parameter, So *et al.* (2001) proposed a metric for quantifying virtual scene movement by the spatial velocity of computer graphics images. We have analyzed biosignls and motion vectors by time-frequency representations, and investigated the relationship between the autonomic nervous regulation and the motion vector components (Kiryu *et al.*, 2002).

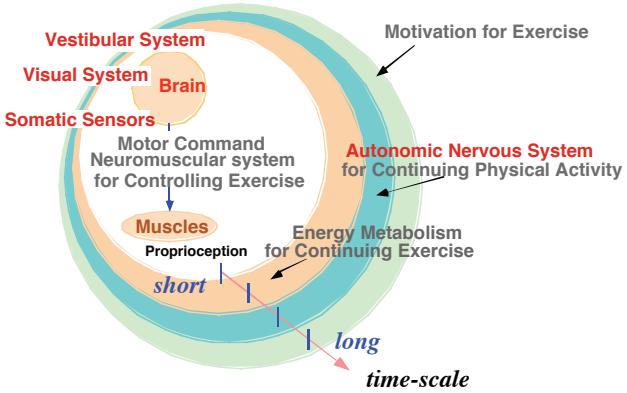

Figure 1: Several time-scales in biosignals

2 Methodology

2.1 Experiments

The experimental subjects were 22 healthy young males and 3 young females (22.4 ± 1.0 yrs. old). They were informed of the risks involved in advance, and their ECG, blood pressure, and respiration were monitored while they were sitting down in the chair during the experiments. The ECG was measured on the chest and the continuous blood pressure was measured by a tonometry method. The respiration was measured with sensors around the chest and the abdomen. These biosignals were sampled at the rate of 1000 Hz with 12-bit resolution. The video image was back-projected onto an 80-inch screen by XGA video projector with over 2500 ANSI lumens. The distance between a subject and the screen was 2 meters and the illumination was 10 lux. All of the biosignals were adjusted at a sampling frequency of 30 Hz by cubic spline interpolation and resampling technique, because the frame rate of the video image was 30 frame/s.

For the 18-min-long vection-inducing video, ten sections of different sports: for example, the video camera mounted on the handlebars of a mountain bike produced unexpected camera motion. In particular, we selected the image of mountain bike and bike race sections.

2.2 Biosignal Processing

We studied autonomic nervous activity (ANA) by calculating the power at limited frequency ranges of blood pressure, respiration, and R-R interval by the continuous wavelet transform (CWT) with the Gabor function as a mother wavelet.

For determining the intervals where cybersickness occurs, we set up threshold levels for the low-frequency (LF) power and the high-frequency (HF) power of R-R interval time-series (Figure 2). That is, after estimating the averaged LF and HF power components during a resting period, we determined the cybersickness intervals based on the following ANA-related conditions (Kiryu *et al.*, 2004): the LF component is greater than the 120% of averaged LF component, LF120, and the HF component is less than the 80% of averaged HF component, HF80. Then, we determined

the onset of unpleasant sensations that could be triggered, tracing the time-series of the LF component backwards in time to find out the local minimum of the LF component time-series.

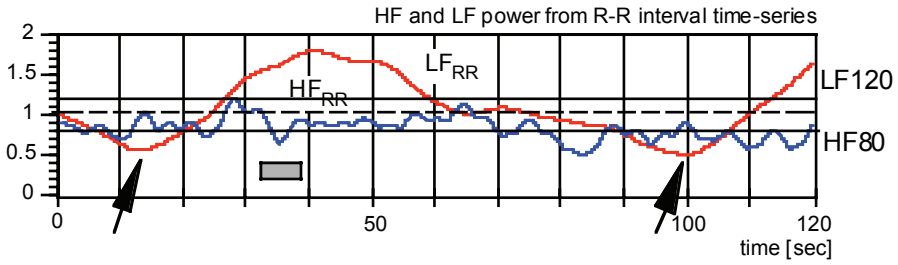

Figure 2: Determination of cybersickness interval (gray region) and trigger point.

2.3 Quantification of Video Images by Motion Vectors

We investigated, as a candidate of the stressor in the vection-inducing video images, the image motion vectors. We estimated the local motion vectors (LMVs) at each region of a screen for an image size of 352×288 pixels. Note that the whole screen was divided into 25 (5×5) regions. Moreover, the motion vectors were estimated in the block of 8×8 pixels and then averaged in each region to obtain the averaged LMVs. In addition, a global motion vector (GMV), which represents camera motion, was estimated from LMVs by a bottom-up approach (Jinzenji *et al.*, 1998). Using a frame-by-frame sliding window with a 1 s interval, we estimated the time-series of the correlation coefficients γ between three GMV components (zoom, pan, and tilt) and each LMV components (up/down and right/left) in 25 regions of the screen. Moreover, the time-varying behavior of each GMV component was analyzed by the CWT.

To trace the similar time-frequency structure of GMVs in the local time of each section, we used the similarity function given by

$$\zeta = \cos^2\theta = \left[\frac{(\boldsymbol{g}_0 \cdot \boldsymbol{g})}{\|\boldsymbol{g}_0\| \cdot \|\boldsymbol{g}\|}\right]^2 .$$

The reference power vector $\boldsymbol{g}_0$ was composed of the powers of GMVs averaged around trigger points over 3 s. The components of the arbitrary power vector, $\boldsymbol{g}$, were the powers over other 3-s intervals. The time-series of the similarity was estimated by sliding the 3-s interval by 1 s.

3 Results

3.1 Questionnaire and Property of Image Sessions

Figure 3 shows the difference of the total score in the SSQ (Simulator Sickness Questionnaire) (Kennedy *et al.* 1993) for each subject before and after watching the 18-min-lomg vection-inducing video. The difference of the total score increased for 8 of sampled 11 subjects. This means that subjects felt uncomfortable due to the vection-inducing video image. Among image sections the mountain bike section showed the highest number of subjects who felt uncomfortable for a simple questionnaire (unpleasant, neutral, or not) and the highest number of trigger points normalized by the elapsed time (Figure 4).

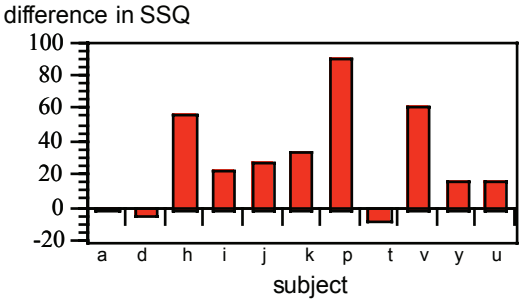

Figure 3: Difference of the total score in the SSQ for each subject.

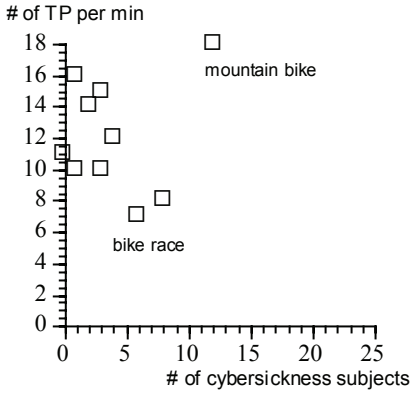

Figure 4: Scatter graph between the number of subjects who felt ncomfortable and the number of trigger points per minute.

Thus we select the mountain bike session as a typical cybersickness-inducing video image and the bike race session as a weaker-cybersickness video image.

The correlation coefficients γ between the three GMV components and each LMV component represented the region where and when the camera motion related the motion of each scene. In the mountain bike section, | γ | was high at the center and distant views for the combinations of pan-(right/left) components and tilt-(up/down) components, whereas it was low at the near views (Figure 5). On the other hand, the high-| γ |-regions were distributed over the screen and in the local time, and only appeared in the tilt-(up/down) component.

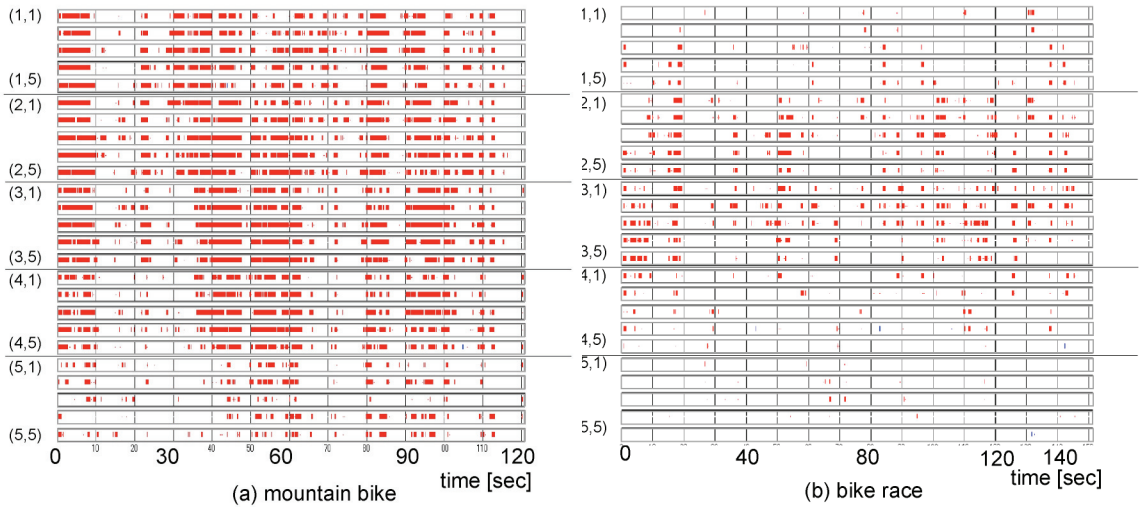

Figure 5: Time-series of intervals with | γ | > 0.7 at 25 regions (from top-left (1, 1) at distant view to bottom-right (5, 5) at near view) on a screen. Note that γ is the correlation coefficient between pan and right/left components for (a) the mountain bike section and between tilt and up/down components for (b) the bike race section.

Moreover, the intervals with high | γ | were relatively long in the mountain bike section. The zoom components did not present significant correlation with LMV component except for zoom-(up/down) components around the above intervals at the distant view. .

3.2 Time-Distribution of Trigger Points and the Similarity

Figure 6 shows an example of the time-varying behavior of biosignals and ANA-related indices. The subject quitted watching video due to unpleasant sensation over 6 min. From top to bottom, respiration at nose, R-R interval, mean blood pressure, perspiration at arm, HF component from respiration, LF component from blood pressure, HF component from R-R interval, and LF component from R-R interval. The gray zone indicates the cybersickness interval and virtical lines show the trigger points. During cybersickenss interval biosignals varied differently from other intervals and in turn changes appeared in the ANA-related indices.

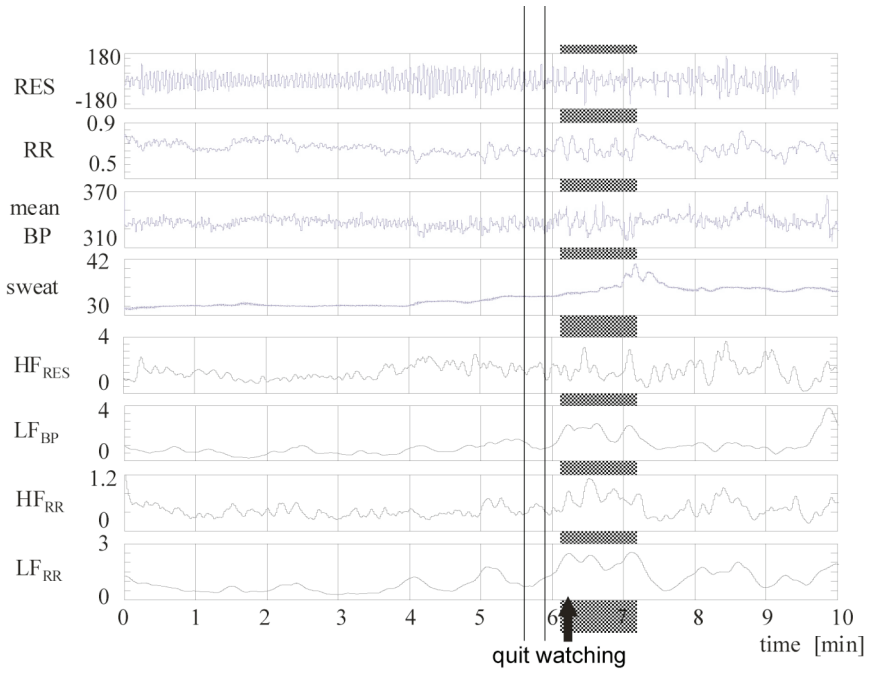

Figure 6: Time-series of biosignals and ANA-related indices

We used two pairs of ANA related indices, (LF120, HF80), for determining the cybersickness intervals: LF120 and HF80 estimated from R-R interval, and HF80 from respiration and LF120 from blood pressure. Figure 7(a) shows the time-distribution of 36 trigger points for 25 subjects at each 10-s segment in the mountain bike session. The results showed that the cybersickness intervals were around 61–70 s and 91–100 s segments and the time-distribution of the trigger points showed two peaks around these segments. Note that the two peaks were located around the high-| γ |-regions. On the other hand, there were 18 trigger points in the bike race section and the time-distribution did not match the high-| γ |-regions (Figure 7(b)).

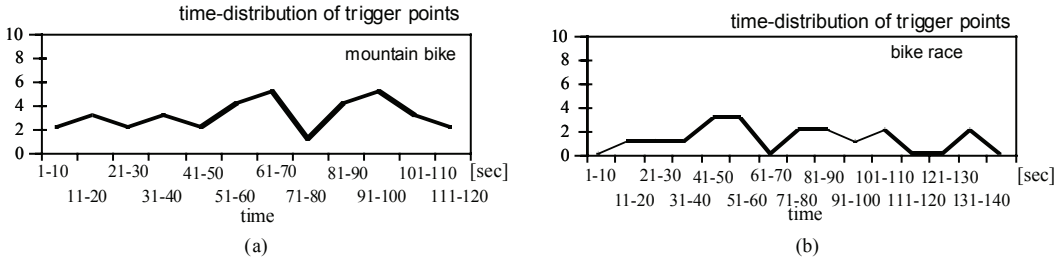

Figure 7: Time-distribution of trigger points in (a) the mountain bike and (b) bike race sections

Figure 8(a) demonstrates that time-distribution of γ at 10-s segment between the pan and right/left components in the mountain bike section. Note that the location of LMV was the center of the screen, (3,3). The result showed that the time-distribution of γ was similar to that of trigger points. On the other hand, there was not so similar time-distribution of γ between the tilt and up/down components in the bike race section (Figure 8(b)).

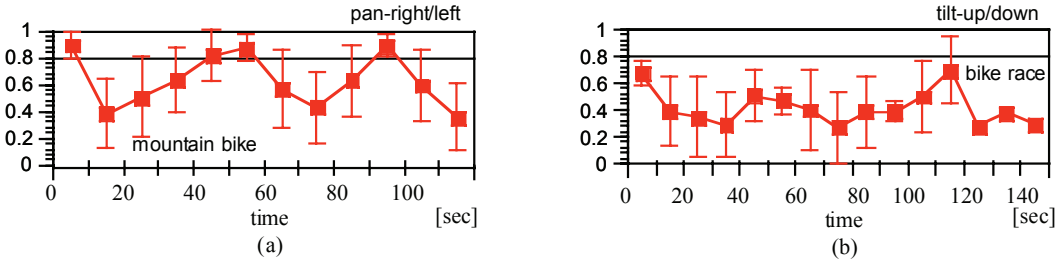

Figure 8: Time-distribution of γ at 10-sec segment between the pan and (right/left) components, and the tilt and (up/down) components in (a) the mountain bike and (b) bike race sections

Figure 9(a) shows the time-varying behavior of the similarity function referring to the two trigger points, 65–68 and 91–94 s, estimated in the mountain bike section. The similarity showed 1.0 at the trigger points, and decreased at around 30 and 80 s where there was no camera motion (see Figure 5(a)). As a result, the overall behavior of the similarity based on the time-frequency structure of GMVs was quite similar to the time-distribution of trigger points estimated from biosignals (see Figure 7(a)). On the other hand, the overall behavior of the similarity function was not so similar to the time-distribution of trigger points in the bike race section (Figure 9(b)).

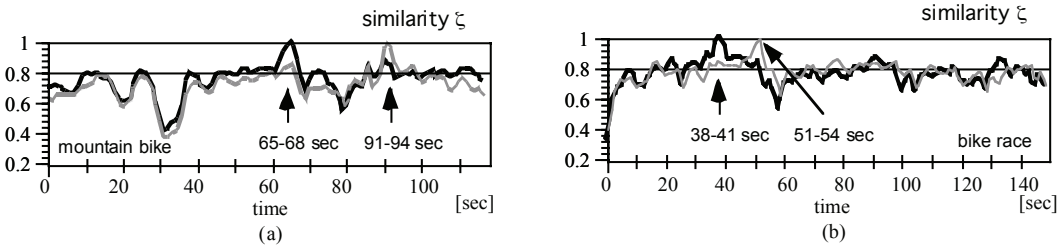

Figure 9: Time-distribution of similarity function in (a) the mountain bike and (b) bike race sections

3.3 Time-Frequency Structure of Motion Vectors in relation to Cybersickness

Although there were individual differences for cybersickness intervals, specific time-frequency structures of motion vectors around the two peaks of the trigger points appeared in the mountain bike section were statistically high potential to cause cybersickness. Checking the details of the time-frequency structure of GMVs over 0.5 in the normalized power around trigger points, the time-frequency structure included low frequencies ranging 0.3–2.5 Hz in three components (Figure 10(a)). Regarding the bike race section, the time-frequency structure included 2.5–8 Hz in three components (Figure 10(b)).

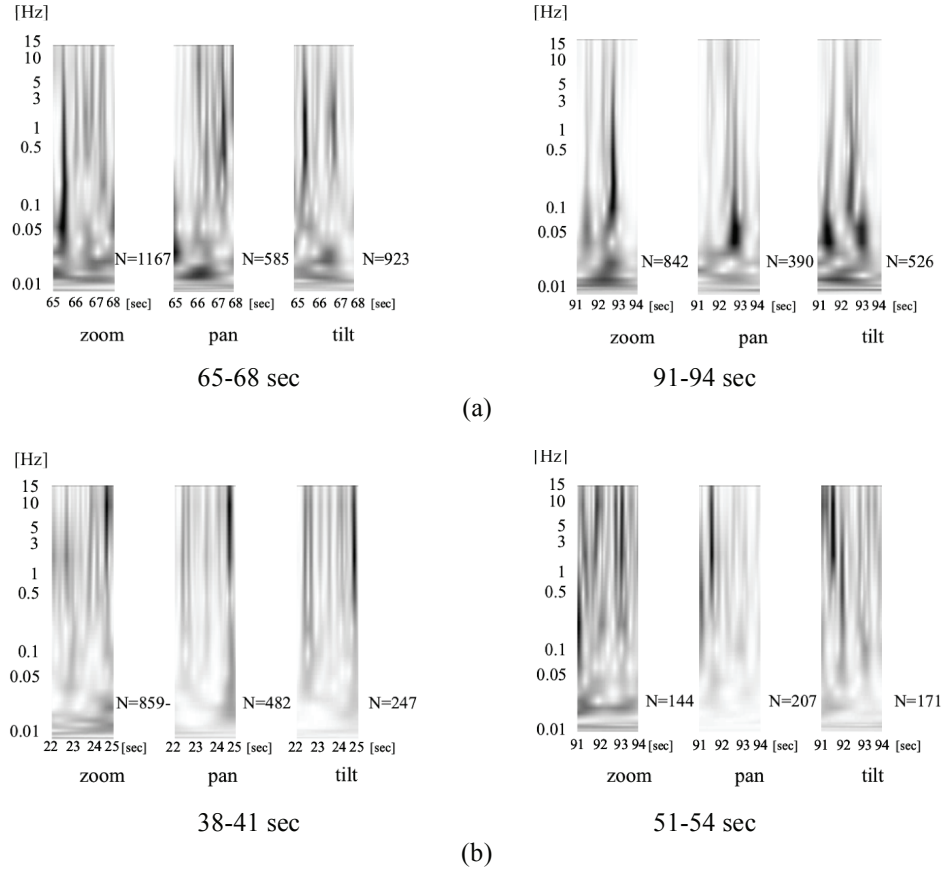

Figure 10: Time-frequency representation of GMVs around trigger points in (a) the mountain bike and (b) bike race sections.

4 Discussion

4.1 Featuring Vection-Induced Video Images

In general, vection is primarily dependent on motion in peripheral vision (Webb & Griffin, 2002). In the mountain bike section, due to the high-$|\gamma|$-regions in the distant and center views (Figure 5(a)), the vection most likely occurred. In addition, the time-varying behavior of γ was similar to those of the time-distribution of trigger points and the similarity (Figures 7(a), 8(a), and 9(a)). In this sense, the mountain bike section had the effective spatial and temporal behavior in the pan and tilt components for inducing cybersickness.

4.2 Cybersickness

It has been reported that the mismatch between the visual and vestibular systems disturbs the autonomic regulation when vection-inducing images are viewed (Yates & Miller, 1998). During unpleasant sensations, HF power is expected to decrease and LF power to increase in general (Grillot *et al.,* 1995). Thus, we set up LF120 and HF80 for the ANA-related conditions. Since there might be a time delay in the autonomic regulation (Wood *et al.,* 2000), we determined the triggered onsets, tracing the time-series of the LF component backwards in time (Figure 2). Further experiments should be needed to assess the cybersickness intervals and trigger points in terms of different types of ANA-related indices.

The influences of vection-inducing images are caused by many different factors. One of the factors is video content that links to the experience of each subject. That is, even though the visual stimuli was the same, the influence on the autonomic regulation depends on individual experiences and conditions at that time. Vection-inducing images made by random-dot pattern would reduce the experience related bias and the influence of the motion vectors can be solely studied.

When camera motion is difficult to follow, subjects feel unpleasant sensation. Thus, for example, off-centered vection-inducing video images with certain exposure time could enlarge the sensory conflict and in turn disturb the autonomic regulation. Moreover, the exposure time should be discussed in relation to the cybersickness intervals.

4.3 Motion Vecotrs

Although there are differences between cybersickness and motion sickness, it was reported that motion sickness was increased by around 0.5 Hz horizontal translational oscillation (Golding *et al.,* 1997). The motion vectors alone do not present all of the features, but the time-frequency representation allowed us to predict the candidates that might disturb the autonomic regulation. In this sense, the time-frequency structure of motion vectors around trigger points should be further investigated. Using vection-inducing images synthesized by random-dot pattern images with the same motion vectors as the real images, several combinations of GMV components and frequency ranges would reveal the detailed factors for cybersickness.

5 Conclusion

Prevention of cybersickness is a key point to apply virtual reality to telemedicine and rehabilitation for a long time use. Cybersickness is related to factors in both images and the conditions of users. We studied influences of vection-induced images in the relationships between the autonomic-nervous-activity-related indices and the motion vectors of video images. Autonomic nervous activity was evaluated from R-R interval, blood pressure, and respiration.

According to the time-varying behavior of motion vectors, the high correlation coefficients between local and global motion vectors and the time-frequency structure of global motion vectors could be used for predicting cybersickness intervals. However, we have not yet concluded whether the unpleasant feeling was caused by the contents of the vection-inducing images or the structure of the video image scene. Further studies are required to clarify the levels of sensory activity on autonomic regulation to block cybersickness.

Acknowledgement

This study is carried out under the Standard Authentication R&D Program, "Standardization of Assessment Method for Visual Image Safety," promoted by the Ministry of Economy, Trade and Industry in Japan.

References

Cowings, P. S., Naifeh, K. H., Toscano, & W. B. (1990). The stability of individual patterns of autonomic responses to motion sickness stimulation. *Aviat Space Environ Med*, 61 (5), 399-405.

Gianaros, P. J., Quigley, S., Mordkoff, J. T., Stern, & R. M. (2001). Gastric myoelectrical and autonomic cardiac reactivity to laboratory stressors, *Psychophysiology*, 38 (4), 642-652.

Golding, J. F., Finch, M. I., & Stott, J. R. (1997). Frequency effect of 0.35-1.0 Hz horizontal translational oscillation on motion sickness and the somatogravic illusion. *Aviat. Space Environ. Med*., 68 (5), 396-402.

Grillot, M., Fauvel, J. P., Cottet-Emard, J. M., Laville, M., Peyrin, L., Pozet, N. & Zech, P. (1995). Spectral analysis of stress-induced change in blood pressure and heart rate in normotensive subjects. *J. Cardiovasc. Pharmacol.,* 25 (3), 448-452.

Harding, G. F. A. (1998). TV can be bad for your health. *Nature Medicine*, 4 (3), 265-267.

Hettinger, L. J., Berbaum, K. S., Kennedy, R. S., Dunlap, W. P., & Nolan, M. D. (1990). Vection and simulator sickness. *Mil Psychol*, 2 (3), 171-181.

Jinzenji, K. Watanabe, H. & Kobayashi, N. (1998). Global motion estimation for static sprite production and its application to video coding. *IEEE ISPAC'98*, 328-332.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *Int. J. Aviat. Psycholo.*, 3 (3), 203-220.

Kenyon, R. V., Leigh, J. & Keshner, E. A. (2004). Considerations for the future development of virtual technology as a rehabilitation tool. *J. NeuroEng. Rehab.,*1 (13).

Kiryu, T., Nanbo, Y., Kobayasi, N., & Bando, T. (2002). Relationship between Motion Vectors of Vection-Induced Image and Multivariate Biosignals under Visual Tasks. in *Proc. 4th International Workshop on Biosignal Interpretation*, 517-520, Como, Italy.

Kiryu, T., Yamada, H., Jimbo, M., & Bando, T. (2004). Time-Varying Behavior of Motion Vectors in Vection-Induced Images in Relation to Autonomic Regulation. in *Proc. 26th Annu. Int. Conf. IEEE/EMBS*, 2403-2406, San Francisco, CA.

Nichols S. & Patel, H. (2002). Health and safety implications of virtual reality: a review of empirical evidence. *Appl Ergon*, 33 (3), 251-271.

Peters, K., Darlington, C. L. & Smith, P. F. (2000). The effects of repeated optokinetic stimulation on human autonomic function. *J Vestib Res*, 10 (3), 139-142.

So, R. H., Ho, A., & Lo, W. T. (2001). A metric to quantify virtual scene movement for the study of cybersickness: Definition, implementation, and verification. *Presence*, 10 (2), 192-215.

Stanney, K. M., Kennedy, R. S., Drexler, J. M., & Harm, D. L. (1999). Motion sickness and proprioceptive aftereffects following virtual environment exposure, *Appl Ergon*, 30 (1), 27-38.

Webb N. A. & Griffin, M. J. (2002). Optokinetic stimuli: motion sickness, visual acuity, and eye movements. *Aviat. Space Environ. Med*., 73 (4), 351-358.

Wood, S. J., Ramsdell, C. D., Mullen, T. J., Oman, C. M., Harm, D. L., & Paloski, W. H. (2000). Transient cardio-respiratory responses to visually induced tilt illusions. *Brain Res. Bull*., 53 (1), 25-31.

Yates B. J. & Miller, A. D. (1998). Physiological evidence that the vestibular system participates in autonomic and respiratory control. *J Vestib Res*, 8 (1), 17-25.