

# A CONSIDERATION ON A DETERMINATION METHOD FOR SCREEN SHAKE WITH HISTOGRAMS IN MOTION VECTORS

Mei Kodama<sup>†, ††</sup>

<sup>†</sup> Graduate School of Integrated Arts and Sciences, Hiroshima University

<sup>††</sup> Information Media Center, Hiroshima University

## ABSTRACT

Recently, we have used many videos by various image display devices. In particular, since an opportunity of viewing contents in a large size display or a mobile device has been increased, it is one of the important problems to prevent visually induced motion sickness (VIMS). Then, in this study, in order to decrease an adverse effect on a human body due to VIMS, a determination method of screen shake in consideration of high-speed processing is proposed. The histogram is calculated by using motion vectors which are obtained by a simple block matching method. The screen shake state is determined by using the degree of similarity to grasp the temporal change of the histogram. The effectiveness of the proposed method is considered by the simulation experiments. By the experimental results, it is revealed that the pseudo motions affect the values of the degree of similarity and the simple proposed method can determine the state of screen shake.

## 1. INTRODUCTION

4K Test Broadcasting started last year in CS (communication satellite) broadcasting of Japan, and it also began on June this year in a part of BS channels of Japan. High-definition TV (e.g., 2K, 4K, and 8K) has become popular. Furthermore, the higher resolution proceeds in video distribution services by using on-demand type. Thus, we realize that many kinds of high-definition monitors adapted to 4K or 8K will be sold and a viewing opportunity by an advanced high-definition display will be increased in the future. However, in a viewing opportunity by a large screen display or a super/ultra high-definition TV as described above, there are some important problems to prevent visually induced motion sickness (VIMS), a photosensitive seizure, and an asthenopia by 3D TV [1]– [14]. They had been studied as a biological safety problem. For example, since a measure against flicker occurrence is needed to prevent a photosensitive seizure, warning information (e.g., “Please turn on the light in your room.”) is displayed as a telop during broadcasting. However, there is a limit to a measure against VIMS in the same manner as a photosensitive seizure. That’s because an external factor is ineffective. Therefore, an automatically image processing technique is required to prevent it, and a systematic mechanism of the prevention is needed. In particular, we could not ignore the problem the influence for children. It is one of the important problems that have to be solved. Moreover, there are

concerns about VIMS for not only a large screen shake of the moving images, but also a small screen shake of them in 4K video in addition to 2K issues. Therefore, since digital cameras and digital video movies adapted to HDTV and UHDTV becomes popular, it is important to improve a occurrence of a camera shake during photographing, or a defect of a camera work by an amateur. For example, a screen shake occurs under weather conditions (e.g., the wind blows hard.), or under traffic conditions (e.g., it is used in the train or on the pedestrian over-bridge.). We could not ignore the slight shake for VIMS.

Then, the extraction methods had been studied to prevent VIMS [4]–[14]. However, there is a problem in a practical processing time with them due to the additional processing to the motion analysis for VIMS. For example, the global motion vectors (GMV), which are anew generated in the special domain, are able to be used to analyze the detailed motion of the screen [9][13][15][16]. On the other hand, simple motion vectors are calculated by a general motion estimation method (e.g., a block matching (BM) method), which is used in a generic coder. The motion estimation method with GMV incurs the calculation load grater than a general method. Therefore, this paper proposes a determination method of screen shake for VIMS by using motion vectors with BM method to realize real-time processing of the motion analysis. It uses the histograms of motion vectors, and it is the point to analyze the motion of the screen caused by the camera shake by using the change of the histogram distribution. We consider the effectiveness of the determination method with histograms in the simulation experiments.

## 2. A DETERMINATION METHOD OF SCREEN SHAKE

The determination method with the global motion compensation had been studied to analyze the detailed camera motions [9]. However, the real-time processing is extremely difficult in theses methods. There is a problem due to not adapting to the real time processing to determine a video scene including the danger of VIMS [1][2].

By the way, the searching method with spacial information of a query video and the retrieval method with motion vectors of it had been studied as a high-speed video searching method [17][18]. Those retrieval method used the characteristics of the similarity of motion information. Then, considering the reduction of the processing time, we propose a determination method with histograms of motion

vectors based on a BM method in a general motion compensation processing. This method can determine the feature of motion vectors (MV) by reversely using the principle of the similarity determination. In other words, we can determine general video scenes, which do not include the scene corresponding to screen shake by the temporal correlation of the histograms.

First, the principle of the scene determination for screen shake is described. Fig.1 illustrates the sample of the change of histogram of motion vectors. As shown in Fig.1(i), in the case of moving pictures, which do not include scene change, the histogram of MV in frame  $n$  is generally similar to that of MV in frame  $n + 1$  due to the series of image motion periods. On the other hand, as shown in Fig.1(ii), in the case of moving pictures, which include scene change or screen shake information (e.g., hand shake), it is considered that the histogram distribution changes from the left figure to the right figure due to hand shake. Moreover, as illustrated in Fig.2, we consider that the histogram change in one of the motion directions of MV is a smooth and continuous curve in time domain, when the sequence do not have screen shake information (Fig.2(i)). However, the change is discontinuous, when the sequence has that information (Fig.2(ii)).

Thus, we use the change of the distribution of MV ob-

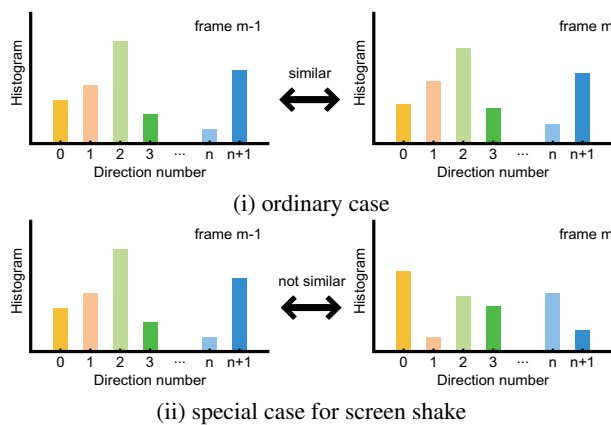


Fig. 1. Comparison of the degree of similarity in histograms

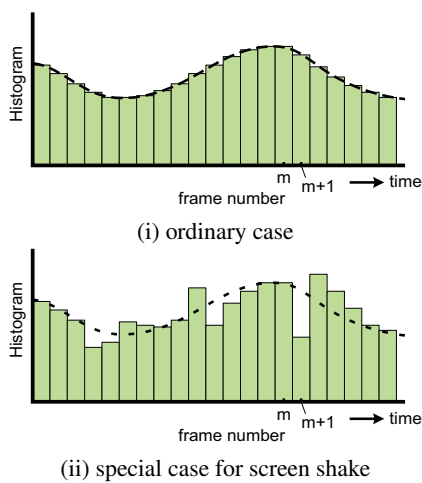


Fig. 2. Comparison of temporal change in motion direction histograms

tained by each frame, and propose a determination method by using the change of the similarity between the histograms continuing in time. Next, Fig. 3 shows the determination flow of screen shake. The procedure of this processor is indicated in the following steps.

1. MV information based on a BM method is calculated. Where the block size is  $16 \times 16$  pixels, and the vector search range per frame is  $\pm 32$  pixels.
2. The histogram of MV is calculated in each motion direction type and the two special motion types. The first special type shows that motion vectors are not able to be calculated, and it is also defined as non-mv type( $D_0$ ). The second special one shows that motion vectors are zero, and it is also defined as zero-mv type( $D_1$ ). The motion direction type(md type) has motion vectors for each direction. These types are numbered starting from  $D_2$ , e.g., in order of right in the horizontal direction(HD), left(HD), upward in the vertical direction(VD), and downward(VD). In particular, a pair of motion direction histograms(PMDH) in a motion direction and its opposite motion direction is utilized in order to extract screen shake information eas-

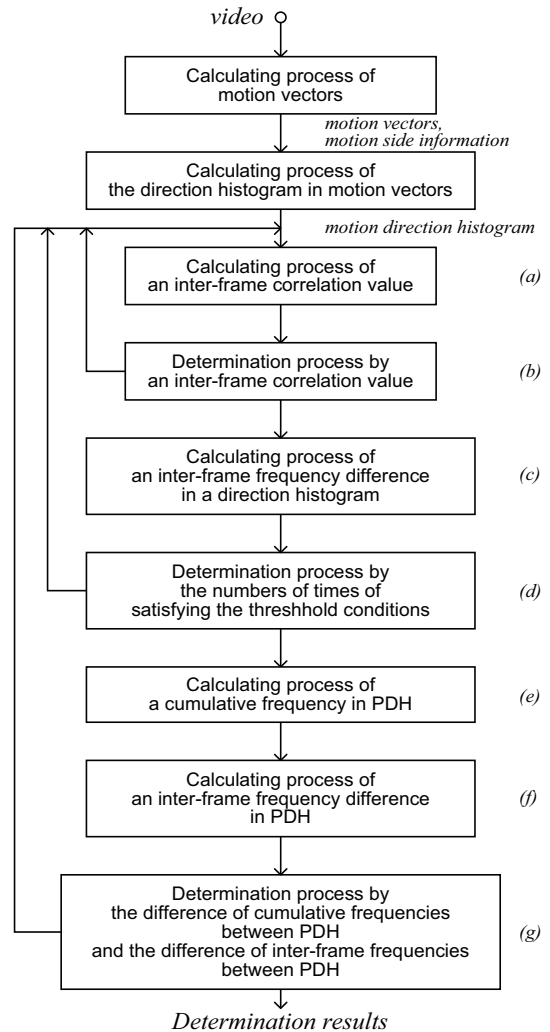


Fig. 3. Determination flow of screen shake

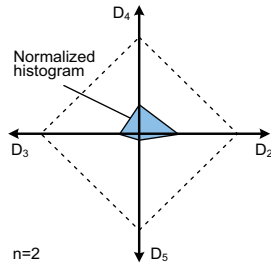


Fig. 4. Sample of normalized motion direction histogram

ily. That's because the change appears in PMDH due to screen shake. After calculating process of the histogram, the values are normalized in all types. Fig.4 shows a sample of the normalized motion direction histograms in md types. Where the number of md types is defined as  $2^n$ , and  $n = 2$ . In this study, we focus on the histogram of the motion directions simply. The other methods are available in the same manner [18].

- The degree of similarity is used as an inter-frame correlation value. It is calculated by the histograms continuing in time to check whether the scene has screen shake information. This processing is shown by Fig.3 (a). This study uses the determination method with *Bhattacharraya* distance(BD)[19]. The candidate video scene is determined by using the threshold in the distance of BD in Fig.3 (b). Where the threshold value ( $Th_{bh}$ ) is set to 0.1. *Bhattacharraya* coefficients ( $\rho_{m+1}$ ) between frame number  $m$  and  $m + 1$  are calculated by Eq.(1). Where  $H_{m+1}(x)$  and  $H_m(x)$  denote the histogram of motion directions in current frame and that of motion directions in the previous frame, respectively, and  $h_{m+1}(x)$  and  $h_m(x)$  denote the normalized motion direction histograms, respectively.  $x$  shows each bin for the histogram. Next, the distance( $d_{m+1}$ ) is given by Eq.(1). The candidate scene of screen shake is narrowed down as the first step according to the condition of Eq.(1).

$$\left\{ \begin{array}{l} h_m(x) = \frac{H_m(x)}{\sum_x H_m(x)}, \quad h_{m+1}(x) = \frac{H_{m+1}(x)}{\sum_x H_{m+1}(x)} \\ \rho_{m+1} = \sum_x \sqrt{h_m(x) h_{m+1}(x)} \\ d_{m+1} = \sqrt{1 - \rho_{m+1}} \geq Th_{bh} \end{array} \right. \quad (1)$$

- If the correlation values are used only, there are some problems of false detection by using the previous processing (b) in the special scenes. For example, since the histogram changes when new objects move in a scene, the correlation value changes. Moreover, there is the same problem when objects existing at a near distance move sharply or largely. Therefore, it is found that a determination method by using only BD as a correlation value is insufficient to detect the screen shake. Then, an inter-frame frequency difference(IFFD) in a

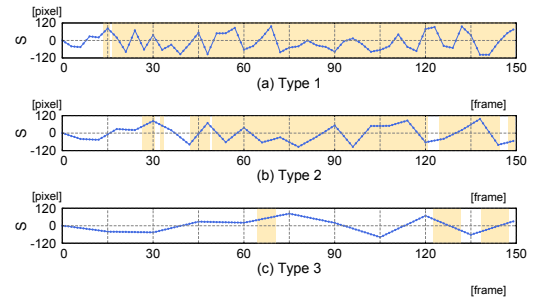


Fig. 5. Samples of used white noise

whole motion direction histogram is used to detect it as illustrated in Fig.3 (c). Next, the determination method uses the numbers of times of satisfying the threshold conditions in IFFD as the second step in Fig.3 (d). When the normalized value of the histogram for  $D_j$  is defined by  $h_m(D_j)$ , the conditions are shown by  $|h_m(D_j) - h_{m+1}(D_j)| \geq Th_{if}$ .

- A cumulative frequency(CF) in PMDH and IFFD in PMDH are calculated in time domain as illustrated by Fig.3 (e) and (f). The scene of screen shake(the caution scene) is detected as the third step by the determination method using both CF in PMDH and IFFD in PMDH as shown by Fig.3 (g). We focus on the relationship between PMDH caused by screen shake. If motion vectors have screen shake information in a motion direction, a negative correlation between PMDH can be used as a correlation value. For example, if the frequency of the histogram in the right motion direction increases due to the shake, the frequency in the left motion direction decreases. At this time, if the histogram changes due to screen shake, the difference of CF between PMDH would be small. Therefore, the caution scene is finally determined by the conditions(i.e.,  $|d1| \leq th_{cf}$  and  $d2 \times d3 < 0$ ). Where  $d1$  denotes the difference of CF between PMDH in time domain,  $d2$  denotes the difference between PMDH in the right motion direction, and  $d3$  denotes the difference between PMDH in the left motion direction in time domain. The results of the caution scene by using the third step are output as the final determination results.

### 3. EXPERIMENTS

First, experimental conditions for extracting the caution scenes for screen shake are defined in this section. Next, the features of determination values, which are used in the proposed method, are evaluated. Lastly, we consider the determination results and the effectiveness.

#### 3.1. Experimental conditions

The experimental conditions are indicated to evaluate the proposed method. First, the swing images are generated by using white noise defined. The samples of white noise are illustrated in Fig.5. Here, the same noise is produced by

a random function. Additionally, the three kinds of white noises (Type1, Type2, and Type3) in time domain are defined by the same noise. On the other hand, the usage case of original images without the white noise is shown as Type 0. In this study, these noises have only the horizontal noises to evaluate it simply. Furthermore, the simulation images are generated by adding these noises to the original images. The caution frame determined as screen shake is shown by the background in yellow color in Fig.5. Provided that the number of increasing or decreasing times is grater than 2 within a second and the size of shake is larger than 15 pixels at the same time as a caution scene, which is used in this

study. Next, 3 test sequence are used (s1:bus, s2:mobile and calendar, s3:flower garden, all sequences: 150 frames), and the original images and the swing images are the same size of  $504 \times 280$  pixels. The search range in a BM method is set to  $\pm 32$  pixels, and the number of the direction group (DG) is set to 4 ( $2^n$ :  $n = 2$ ). In a BM method, a full search method is used. Here, this paper does not consider the norm of MV to be easy to manage motion direction histograms. As the histogram information, it uses each direction group, the group of zero vectors, and the group of non-mv type. Table 1 shows the rates of motion direction type for each sequence. Here,  $G_1$ : the group of non-mv type,  $G_2$ : the group of zero-mv type, and  $G_3$ : the group of md types. Type 0 indicates original sequence without white noise. Type 1–3 indicate shake sequence with each white noise in Fig.5. Of course, when the rate of  $G_3$  is smaller, the accuracy of detecting the caution scene becomes unstable in the proposed method due to lack of motion direction information. In Table 1, it is found that the rates are grater than about 45 %.

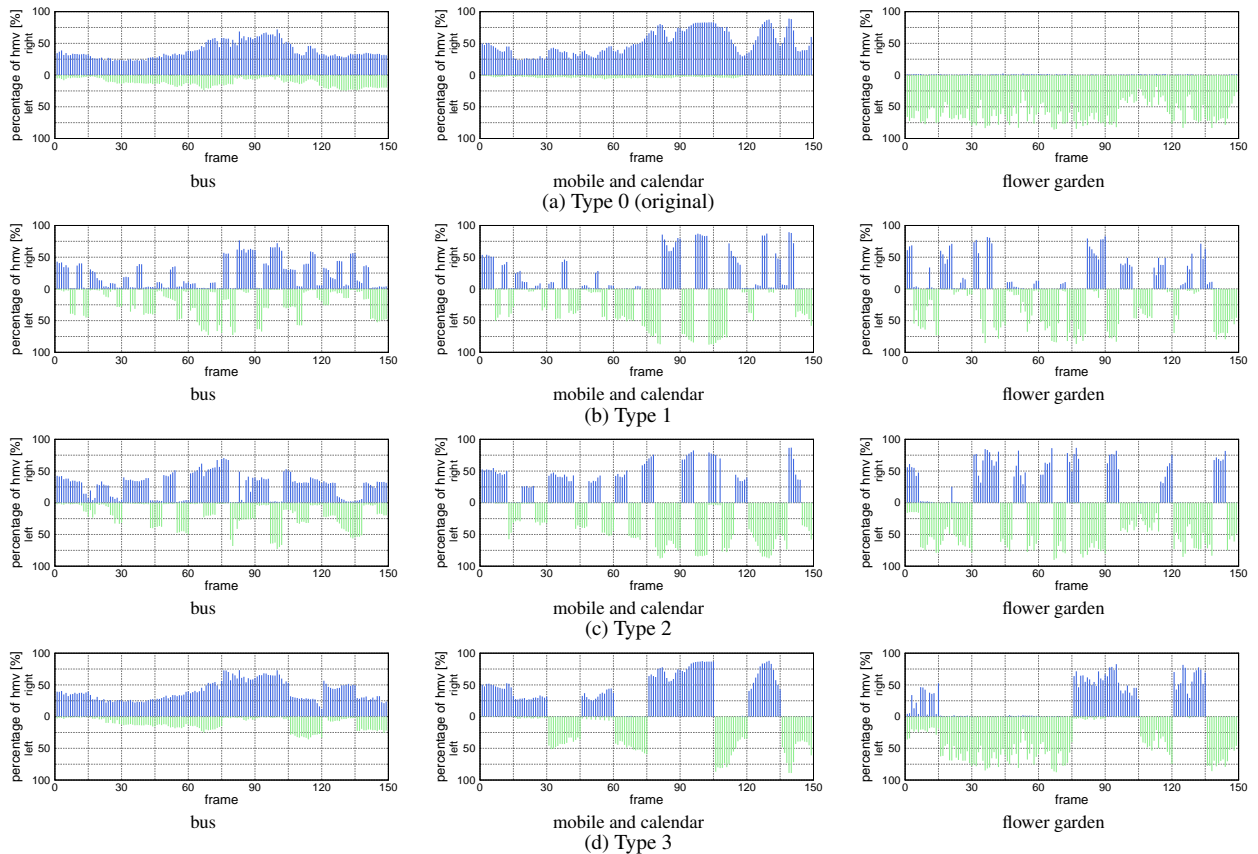
**Table 1. Average rates of all groups and noise type [%]**

sequence	noise type	$G_1$	$G_2$	$G_3$
s1	0	45.528	2.028	52.444
s1	1	53.467	0.434	46.099
s1	2	48.286	2.335	49.379
s1	3	46.099	1.088	52.813
s2	0	44.567	0.607	54.826
s2	1	54.152	0.250	45.598
s2	2	44.904	0.657	54.439
s2	3	44.982	0.372	54.646
s3	0	37.178	0.682	62.140
s3	1	50.075	1.049	48.877
s3	2	37.009	0.659	62.331
s3	3	36.915	3.360	59.725

### 3.2. Evaluation of determination values

As determination values, we observe the distribution of motion direction in the horizontal motion in time domain and *Bhattacharayya* distance in order to understand the change due to white noise.

Fig. 6 shows the occupation rates of the horizontal MV in the original images and the swing images in bus sequence.



**Fig. 6. Occupation rates of motion vectors in horizontal directions of right and left for each sequence**

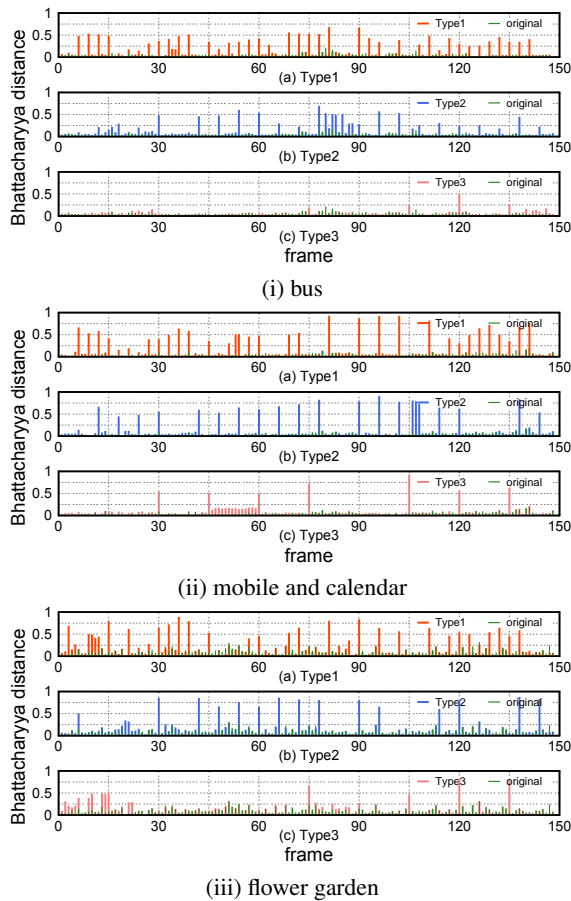


Fig. 7. Comparison of Bhattacharyya distance for each sequence

It is found that there is the correlation in temporal domain in the original images in Fig.6(a). On the other hand, we can find the lack of the continuous distribution in the rate in Fig. 6 (b)–(d). These characteristics are obtained by the determination processing of Fig. 3 (b) in the proposed method.

Next, Fig.7 shows the change of *Bhattacharyya* distance for each swing images comparing to the original images in flower garden. It is found that the values of the original images are extremely smaller than those values of swing images in the sequences of bus, and mobile and calendar. It is indicated that the values of the original images (flower garden) are a little larger than those values of them (the other sequences). The reason is that the images of flower garden have noncontinuous motion vectors in the part of objects.

Therefore, it is found that the determination results derived by the determination method with the first step only have the erroneous results for the original sequences.

### 3.3. Considerations

As shown in sec.3.2, there is a problem of detecting the scene without screen shake in the determination method with the first step. Of course, if we control the threshold value in BD, the detected scenes change. However, if this value is set to larger than the BD value of original images, e.g., 0.2(flower garden), extracted scenes could not be narrowed down. Then, this paper focuses on the correlation

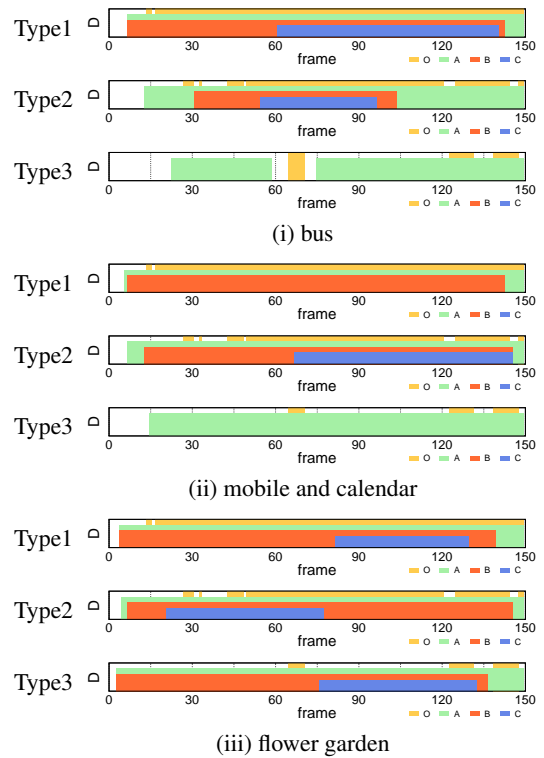


Fig. 8. Results of caution scenes for each determination step

relationship between the motion direction histograms in the horizontal direction. The normal scene which does not have screen shake could be eliminated by the relationships in the second step and the third step.

The results of caution scenes by each determination step are illustrated by Fig.8. Where ‘O’ denotes the results obtained by only white noises as shown in Fig.5. ‘A’, ‘B’ and ‘C’ denote the determination results for 1st step, 2nd step, and 3rd step, respectively.

It is observed that the distance values are larger than 0.25 in swing images, and the values change largely in Fig. 7. Additionally, there are the values which are larger than 0.1 in the original sequences. Then, the threshold value is set to 0.25 in order not to become lower than the maximum value of BD in the original sequence. Here, the caution scene is not found in the proposed method with each 1st step, 2nd step and 3rd step for the original images. On the other hand, as shown in Fig.8, the caution scenes are detected by each determination step in the scene having the noise of each Type 1–3. In particular, it is found that the different caution scenes are detected according to the width of screen shake.

In 1st step, the determination scene results include the defined caution scenes in all sequences except for Type 3 in bus as illustrated by green color in Fig. 8. Here, since the values of BD are smaller than 0.25 for around 70 frame in Type 3 (bus) of Fig. 7, the caution scene cannot be detected. Hence, the setting of the threshold value is the further problem. In addition, there is room for reconsidering the definition of the caution scene. In this sequence, since the bus moves to left quickly, it is difficult to detect the slow screen shake in Type 3 due to hiding in the motion direction his-

rogram. However, it is found that this method can detect the fast that in Type 1.

In 2nd step, it is found that almost caution scenes, which are obtained by 1st step, are extracted in Type 1 due to the fast motion by screen shake as illustrated by orange color in Fig. 8. However, this step has the same problem as 1st step in the slow screen shake as Type 3.

In 3rd step, the final determination results are narrowed down from the results in 2nd step as illustrated by blue color in Fig. 8. In mobile and calendar, it seems that this step is not effective since the rate of md type ( $G_3$ ) in Type 1 is less than that in the other cases. Therefore, it is necessary to set the appropriate search range of BM in this sequence. Consequently, these scenes have the characteristics of screen shake, thereby the results by each step represent the priority of the caution scene. However, there is a problem of false detection due to the change of the histogram by the part of moving objects. This problem will be considered in the further study.

Next, the results of precision rates for screen shake are shown in Table 2 as the efficiencies of each determination step. It seems that precision rate decreases by using 3rd step in Table 2. This problem is also one of future issues, and we need to consider the intensity of VIMS caused by screen shake. Furthermore, the quantitative evaluation for swing (e.g., frequency domain analysis, the relationship between the distance and the frequency in the moving objects ) is required. Finally, it is revealed that the proposed method is effective as a determination method in this study.

Next, we evaluate the processing time in the proposed method. The simulation machine with Intel Core i7-3770K CPU 3.50GHz and 16GB memory is used. The result of the average time is  $1.846 \times 10^{-4}$  sec per frame in whole 1st, 2nd and 3rd steps. Furthermore, since a full search method is used as a BM method in this study, it takes about 1 sec per frame to calculate motion vectors in the software simulation. However, since MV in a BM can be obtained from generic decoder(e.g., MPEG-2 decoder, MPEG-4 AVC decoder, etc.[20]–[23]), this processing time can be extremely shortened. Consequently, it is revealed that the proposed method can secure a real-time property since the results are sufficiently smaller than decoding processing. In addition, it is necessary to extend the search range for detecting screen shake fast and precisely. Therefore, a BM method for improving the processing time is also one of further issues.

#### 4. CONCLUSIONS

In this study, the determination method with motion direction histograms of MV as a determination method for screen shake was proposed to realize the real-time processing. In particular, the processing time of the global motion compensation method was extremely greater than that of BM method. Therefore, the proposed method could shorten the determination time. By the experiments, it was revealed that the proposed method realized the extraction for screen shake caused by VIMS. The evaluation by using the practical swing images is required. There are some issues, e.g., a BM method for improving the processing time, an improved

**Table 2. Results of precision rate [%]**

sequence	noise type	1st step	2nd step	3rd step
s1	1	100.00	94.81	59.26
s1	2	100.00	58.65	40.38
s1	3	75.00	0.00	0.00
s2	1	100.00	94.81	0.00
s2	2	100.00	98.08	71.15
s2	3	100.00	0.00	0.00
s3	1	100.00	92.59	35.56
s3	2	100.00	98.08	37.50
s3	3	100.00	62.50	37.50

algorithm of the proposed method considering the precision of detecting caution scenes, other determination values and the definition method of caution scene. Moreover, we need to consider the characteristics of the camera shake for 2K and 4K TV.

#### ACKNOWLEDGMENT

This work was partly supported by Kayamori Foundation of Information Science Advancement.

#### REFERENCE

- [1] S. Chiba: “Health Hazard by Video Image”, Journal of ITE, Vol. 56, No. 6, pp. 926–927 (2002). (in Japanese)
- [2] H. Ujike: “Visually Induced Motion Sickness”, Journal of ITE, Vol. 61, No. 8, pp. 1122–1124 (2007). (in Japanese)
- [3] H. Watanabe, H. Ujike: “The Activity of ISO/Study Group on “Image Safety” and Three Biological Effect”, Proc. of Second International Symposium on Universal Communication (ISUC '08), pp. 210 – 214 (2008).
- [4] H. Ujike, T. Yokoi, S. Saida: “Effects of Virtual Body Motion on Visually-Induced Motion Sickness”, Proc. of 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEMBS '04), Vol. 1, pp. 2399 – 2402 (2004).
- [5] N. Sugita, M. Yoshizawa, A. Tanaka, K. Abe, T. Yambe, S. Nitta, S. Chiba: “Evaluation of the Effect of Visually-Induced Motion Sickness Based on Pulse Transmission Time and Heart Rate”, Proc. of SICE 2004 Annual Conference, Vol. 3, pp. 2473 – 2476 (2004).
- [6] I. Tsubaki, T. Morita, T. Saito, K. Aizawa: “An Adaptive Video Stabilization Method for Reducing Visually Induced Motion Sickness”, Proc. of IEEE International Conference on Image Processing 2005, Vol. 3, pp. III - 497–500 (2005).
- [7] T. Kiryu, G. Tada, H. Toyama, A. Iijima: “Integrated Evaluation of Visually Induced Motion Sickness in Terms of Autonomic Nervous Regulation”, Proc. of 2008 30th Annual International Conference of the

- IEEE Engineering in Medicine and Biology Society, pp. 4597 – 4600 (2008).
- [8] N. Sugita, M. Yoshizawa, M. Abe, A. Tanaka, S. Chiba, T. Yambe, S. Nitta: “Evaluation of Adaptation to Visually Induced Motion Sickness by Using Physiological Index Associated with Baroreflex Function”, Proc. of 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 303 – 306 (2007).
- [9] A. Tanaka, N. Sugita, M. Yoshizawa, M. Abe, T. Yambe: “Dynamic Characteristics between the Subjective Score of Motion Sickness Discomfort and Video Global Motion”, Proc. of 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 1368 – 1369 (2010).
- [10] T. Kiryu, A. Iijima: “A Multi-Timescale Autonomic Regulation Model for Interpreting Visually Induced Motion Sickness”, Proc. of 2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE), pp. 254 – 257 (2014).
- [11] S. Wibirama, K. Hamamoto: “Investigation of Visually Induced Motion Sickness in Dynamic 3D Contents Based on Subjective Judgment, Heart Rate Variability, and Depth Gaze Behavior”, Proc. of 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4803 – 4806 (2014).
- [12] N. Kobayashi, R. Iinuma, Y. Suzuki, T. Shimada, M. Ishikawa: “Using Bio-Signals to Evaluate Multi Discomfort in Image Viewing - Balancing Visually Induced Motion Sickness and Field of View -”, Proc. of 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6198 – 6201 (2015).
- [13] I. Tsubaki, T. Morita, K. Aizawa, T. Saito: “The Analysis of Oscillatory Motion of Videos Affected by Camera Shaking for Visually Induced Motion Sickness”, IEICE A, Vol. 89, No. 3, pp. 262–267 (2006).
- [14] H. Ujike: “Developing an Evaluation System of Visually Induced Motion Sickness for Safe Usage of Moving Images”, Synthesiology English edition, Vol. 5, No. 3, pp. 139–149 (2012).
- [15] Y. Kellaer, A. Averbuch: “Fast Gradient Methods Based on Global Motion Estimation for Video Compression”, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 13, No.4, pp.300–309 (2003).
- [16] S.-C. Cheng, W.-K. Huang: “A Fast Global Motion Estimation for Moving Objects Segmentation Using Moment-Preserving Technique”, Proc. of IEEE International Symposium on Circuits and Systems, Vol.4, pp.3455–3458 (2005).
- [17] M. Kodama: “Video Matching Methods Using Spatial Characteristics in MSP Search System: Enabling Society with Information Technology”, Springer-Verlag Tokyo, pp.38–48 (2002).
- [18] M. Kodama: “A Consideration of A High-Speed Video Identification Method Using Histogram Information of Coded Motion Vectors”, Journal of IIEEJ, Vol.43, No.4, pp.499–507 (2014).
- [19] J. Kazama, et. al: “A Bayesian Similarity Measure for Large-scale Calculation of Distributional Similarities”, Trans. on IPSJ, Vol.52, No.12, pp.3349–3362 (2011).
- [20] ISO/IEC 13818-2, Information Technology - Generic Coding of Moving Pictures and Associated Audio Information: Video, Recommendation H.262 (1995).
- [21] ISO/IEC 14496-10, Information Technology - Coding of Audio-Visual Objects - Part 10: Advanced Video Coding (2006).
- [22] D. Marpe, T. Wiegand, G. J. Sullivan: “The H.264/MPEG4 Advanced Video Coding Standard and Its Applications”, IEEE Communications Magazine, Vol.44, No.8, pp. 134–143 (2006).
- [23] ISO/IEC 23008-2, Information Technology - High Efficiency Coding and Media Delivery in Heterogeneous Environments – Part 2: High Efficiency Video Coding (2013).