Feeling Wind: An Interactive Haptization System for Motion Rendering in Video Contents using SPIDAR

Anusha Jayasiri*
Precision and Intelligence Laboratory
Tokyo Institute of Technology, Japan

Kenji Honda†
Tokyo University of Marine Science and Technology
Japan

Katsuhito Akahane‡
Precision and Intelligence Laboratory
Tokyo Institute of Technology, Japan

Makoto Sato§
Precision and Intelligence Laboratory
Tokyo Institute of Technology, Japan

ABSTRACT

We present a method to feel the movement of objects in object rich image sequences using SPIDAR-G haptic device. This method addresses two major drawbacks of our previous research in object motion rendering in an image sequences. On the one hand, it changes the role of the user from passive to active by enabling the user to select the desired object of which she or he needs to feel the movement. On the other hand, it reduces the impact of background noise on haptic feedback by limiting the sensible region for motion force calculation to an area of a specified size around the point selected by the user. This paper presents the details of the proposed method and some preliminary results of an experimental evaluation involving real users. Experimental results show that having haptic feedback enhances the user experience with videos. It further reveals that the proposed system has smooth and immersive interaction with object rich video contents.

Index Terms: H.5.1 [Information Interfaces and Presentations]: Multimedia Information Systems—Artificial, augmented, and virtual realities; H.5.1 [Information Interfaces and Presentations]: Multimedia Information Systems—Video; H.5.2 [Information Interfaces and Presentations]: User Interfaces—Haptic I/O;

1 INTRODUCTION

Unarguably, the virtual and mixed reality technologies play a vital role in modern multimedia applications. Among them, Haptization technologies enable users to interact with objects in a virtual world. With the recent developments of high technology multimedia systems such as three dimensional televisions, users' viewing experience has been enhanced significantly by pictures in natural color and true dimensions. However, these systems still rely on only two out of five basic senses of human nature in creating user experiences. Apart from the senses of seeing and hearing, there exists feeling, smelling and tasting. Therefore, we believe that the users are more likely to get an enhanced viewing experience if they are provided with the capability of feeling the movement of objects in the video [9] [10].

This research is an attempt to find a method for video users to interact with multiple objects in a video to feel their movement. Rather than feeling the collective movement of all objects in the scene passively we intend to enable users to actively interact with the scene by selecting a desired object and feel the movement of that particular object. For example, figure 1 shows an image sequence of moving leaves. A user may prefer to tap on a particular leave in the video and in response user can feel the movement of the leave rather than feeling the movement of whole sprig. In other words, when the user is watching this video, they can see the leaves moving, but cannot experience the wind. However, by using this system, user can feel the impact of the wind rather than just seeing the leaves moving.

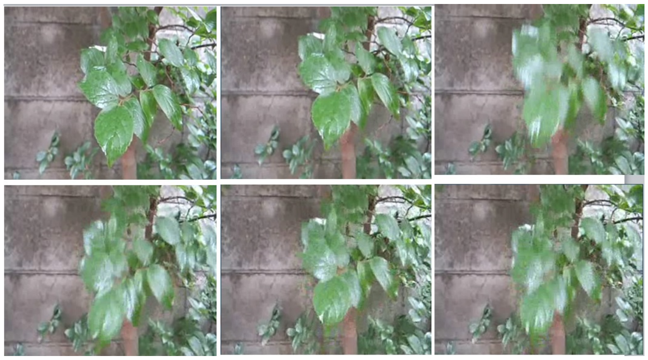


Figure 1: An image sequence of moving leaves

There exist few whole body sensation devices to enrich video viewing. For example, the wearable tactile jacket introduced by Lemmens et al provides tactile stimuli to the viewer's body [12]. Apart from that, Surround Haptic is also a new tactile technology introduced by the Disney research, which allows users to feel digital contents directly on their body, from all possible directions using a small number of vibratory actuators [1]. However, despite their success in providing full body sensation, these devices limit users' interactions with the video. For example, the wearable jacket lets users only to passively feel the video. In our approach, we use the SPIDAR-G haptic device shown in figure 2, which is a grip type, tension-based six degrees of freedom; three degrees of freedom for translation, three degrees of freedom for rotation and grasp enabled force-feedback device [11]. The advantage of using the device is that it enables the users to point and select desired objects in the video and feel their movement actively through the gripping hand.

SPIDAR stands for SPace Interface Device for Artificial Reality. There is a family of SPIDAR devices and the SPIDAR-G, which we are here using, has a grip and it is attached to 8 strings. Each string is connected to a motor and an encoder at one end and to the grip at the other end. The feedback force is determined by the tension of each string generated by the motor, which is transformed to the user's hand through the grip. The device can be connected to a computer using a USB cable to get a high definition force feedback sensation to the user's hand [3]. Apart from that the motion of objects could not include only translation but also rotation. In order to represent those three degrees of rotation and translation in future,

---

*e-mail: anusha@hi.pi.titech.ac.jp
†e-mail:khonda@kaiyodai.ac.jp
‡e-mail:kakahane@hi.pi.titech.ac.jp
§e-mail:msato@pi.titech.ac.jp

we need such kind of 6DOF single-point force feedback device.

Figure 2: The haptic device: SPIDAR-G

Even though, the SPIDAR system has been used in various types of virtual reality applications, it has not been adequately used for haptization in video media [17]. Recently, we have proposed a haptization system [9][10] that enables users to feel the 2D motion force of moving objects in a video using the haptic device SPIDAR-G. To do so, we proposed two haptic motion rendering methods: linear gain controller and nonlinear gain controller method. Experimental evaluation of the two methods has attained the conclusion that using a non-linear gain controller method is more effective than using a linear gain controller method for object motion rendering as that method enables the user to get a continuous feeling of the movement of objects in the video.

However, the previous system had certain limitations. First of all, the videos used to evaluate our methods consisted only one object. The system was not effective for object rich image sequences because of high background noise. It made feature point selection very complex and this complexity affected adversely to the objective of the research. The technique we used to calculate the motion was the optical flow method, in which the optical flow is the apparent visual motion that we experience as we move through the world. This lead to an incorrect evaluation of the velocity as the magnitude and direction of the average velocity of objects in the scene differs from that of a single object in the scene. Apart from that, the user could not interact with the object actively; he just felt the movement of the entire scene.

This paper elaborates how we attempt to improve our previous system to meet the objective of our research by overcoming the above limitations. We propose to limit the velocity calculation to an area around the haptic interface point, pointed by the user. This leads to active user interaction as the user feels the movement of and around the selected object only. Since the background off the selected area is omitted for motion estimation, background noise is also limited. Another advantage of this system is that the user can feel the movement of any desired place on the scene regardless it is a foreground object or a part of the background. Figure 3 shows the demonstration setup of the proposed system. User can see the objects in the video through the computer display and simultaneously, he can get the feeling of the object movement into his hand through the SPIDAR-G haptic device.

This paper is organized as follows: section 2 presents related work for point based haptic interaction with video media. Section 3 presents our proposed approach for object motion rendering, which addresses the limitations of our previous system. Section 4 presents the results of the experimental evaluation involving real users' feedback regarding their viewing experience of video with and without haptic feedback. Section 5 concludes the work and section 6 presents future work.

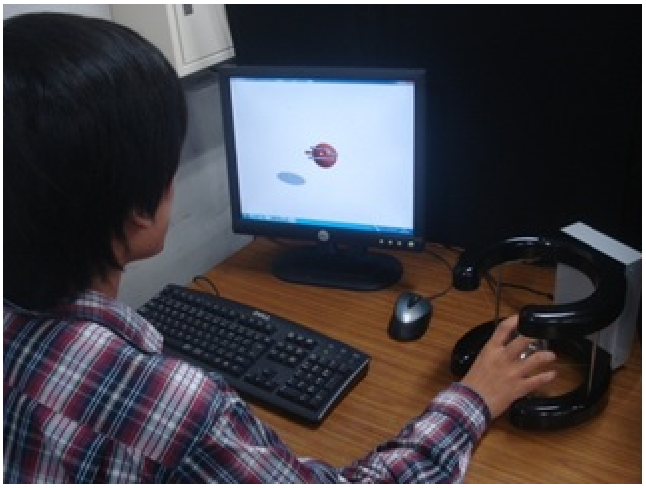


Figure 3: System demonstration

2 RELATED WORK

This section summarizes few existing research on haptic interaction with video media and using SPIDAR for image haptization.

Dinder et al have introduced the concept of haptic motion for the first time and they discussed a method to compute haptic structure and motion signals for 2D video -plus-depth representation [6]. It enables the viewer to navigate the scene and get the experience of the geometry of objects in the scene as well as the forces related to moving objects in the scene using PHANToM haptic interface. They used depth-map information to get 3D video representation. They have modeled the total force as the sum of static and dynamic forces. While the static force is related to the geometry, material and the surface properties of an object, dynamic force related to the relative motion between the object and the haptic interaction point. Since their research, dynamic force for haptic motion related to the relative motion between the object and the haptic interaction point and in that case if the user moves the haptic interface point towards an accelerating object , force experience increased and otherwise it decreased.

Cha et al have proposed a touchable 3D video system which provides haptic interaction with objects in a video scene through the PHANToM force feedback device [5]. They used Depth Image-Based Haptic Representation (DIBHR) and the system enables to physically explore the video content and feel various haptic properties such that texture, height map and stiffness of the scene.

O'Modhrain et al have discussed how haptic interaction can enhance and enrich the viewer's experience in broadcast content [15]. They proposed a touch TV project with the use of 2 DOF force feedback gaming joystick called Gravis Xterminator Force and remote control handset to generate haptic cues for cartoons and live sports broadcastings, which adds greater sense of immersion. They have proposed *presentation interaction* which allows the relocation of a character's rendered position in a scene of a created cartoon without altering the structure of the narrative in any way [16].

Even though, the SPIDAR interface has been used in various types of virtual reality systems, it has not adequately been used for haptization in video media [17]. We have proposed a haptization system [10][9] that enables users to feel the 2D motion force of moving objects in a video using the haptic device SPIDAR-G. We proposed two methods for object motion rendering, one using a linear gain controller and another using a nonlinear gain controller. In the linear gain controller, we used a gain controller *k* to control the feedback force within the sensible region for all velocity levels linearly. Here, *k* is a fraction of the maximum force output level of the SPIDAR-G for better sensation given by the maximum velocity of the video frame at a time T. In the nonlinear gain controller, we used a nonlinear function to control the feedback force within the sensible region for all velocity levels. Furthermore, we evaluated those two methods with the participation of real users and we concluded that the nonlinear gain controller method is more effective than linear gain controller method for object motion rendering. Moreover, in section 3.2.2, we explain the details of nonlinear gain controller method. Also we discussed the drawbacks existing in the previous system in section 1.

Further to that, SPIDAR-G system has been used in the context of images. Liu et al [13] have proposed a 2D image haptization system which provides the users with sense of touch on an image with local deformations using the SPIDAR-G haptic interface. The haptization of 2D image system has used the color haptic parameter mapping and the penalty method to render the haptic feedback. Local deformation was applied by using an image distortion to the local area of the pointer.

He further extends his research on 2D images to 3D image haptization with local deformations by using depth representation of image [14]. In the depth image haptization system, a triangle polygon mesh is extracted and built from the depth image for the haptic rendering. Revised proxy graph algorithm is used to calculate the contact force between the contents and the user's finger. Local deformation is applied by moving the corresponding vertices towards the vector of proxy and endpoint.

3 PROPOSED APPROACH

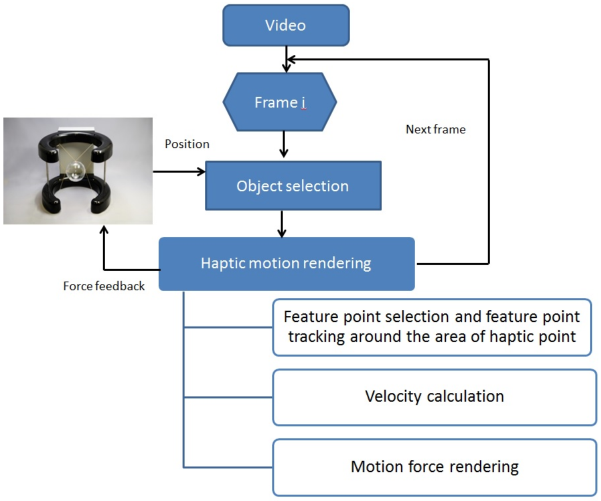


Figure 4: Proposed approach

As illustrated in figure 4, the proposed method has two major parts namely object selection and haptic motion rendering. The first step is object selection. In that step, user can select an object or a part of an object in the image using SPIDAR-G. The next step is haptic motion rendering which involves feature points selection and tracking, velocity calculation and motion force rendering. Feature points selection and tracking involves identifying good features and tracking those feature points from frame to frame within a selected area around the haptic interface point. Velocity calculation involves calculating the velocity by getting the average velocity of selected feature points of the selected object. Finally to render the motion force, we use the nonlinear gain controller method from our previous research. The following sections broadly describe each of these sub processing.

3.1 Object selection

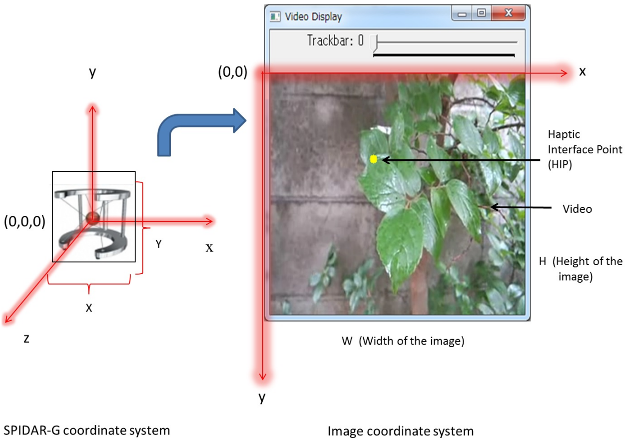


Figure 5: Haptic coordinate system and image coordinate system

In order to touch the image through the haptic device, it is necessary to identify the properties of the input image frame and the hardware limitations of the SPIDAR-G device.

As shown in figure 5, input image is a color image with a width of W pixels and a height of H pixels. Workspace of the SPIDAR-G device is, X mm of width, Y mm of height and Z mm of depth. Since the coordinate systems of the image and the SPIDAR-G device are not matching, it is mandatory to map the SPIDAR-G coordinates to image coordinates, in order to identify the corresponding position of the pointed object on the image.

Equation (1a) and (1b) shows the mapping of SPIDAR-G haptic device coordinates into image coordinates.

$$w = \frac{W}{2} + (\frac{W}{X} * x) \tag{1a}$$

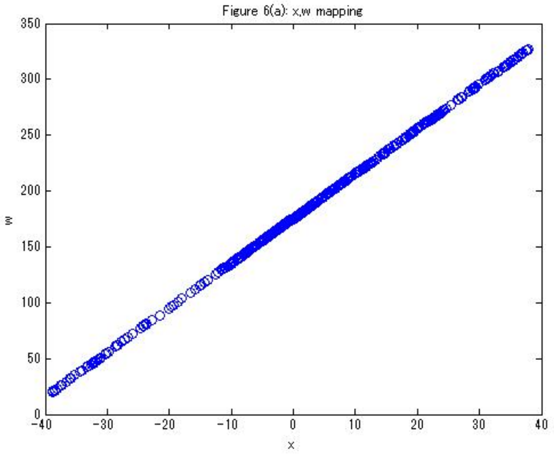
$$h = \frac{H}{2} + (\frac{H}{Y} * (-y)) \tag{1b}$$

Here w and h are the coordinates of image pixel and x and y are the coordinate values of the SPIDAR-G device in the 3D environment along the X and Y axis. However, for most parts of this paper, we only use the XY coordinate plane. In other words, we often do not consider the z values of SPIDAR-G device. However, in the future, we use that value, which allow user to change the region of interest around the haptic pointer to feel different motions.

Figure 6 shows this mapping for the video mentioned in figure 1 which has 594 frames. The width of the image frame is 352 and the height is 288. Figure 6(a) shows the mapping of x coordination of SPIDAR-G to w coordinates of the image frame. Figure 6(b) shows the mapping of y coordinates of SPIDAR-G to h coordinates of image.

3.2 Haptic motion rendering

This section contains how to manipulate velocity information of feature points on the selected object and map them to the haptic force generation of the SPIDAR system.



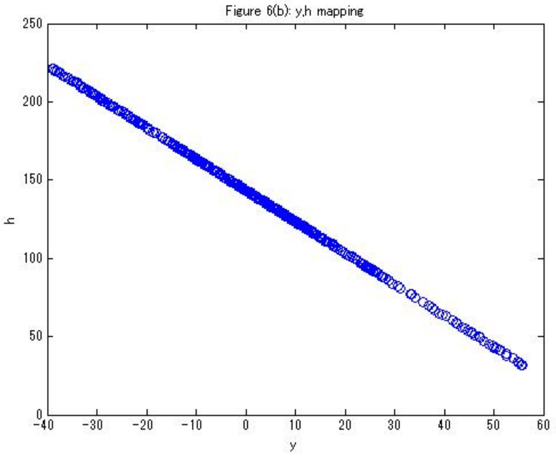


Figure 6: Mapping of haptic coordinates to image coordinates

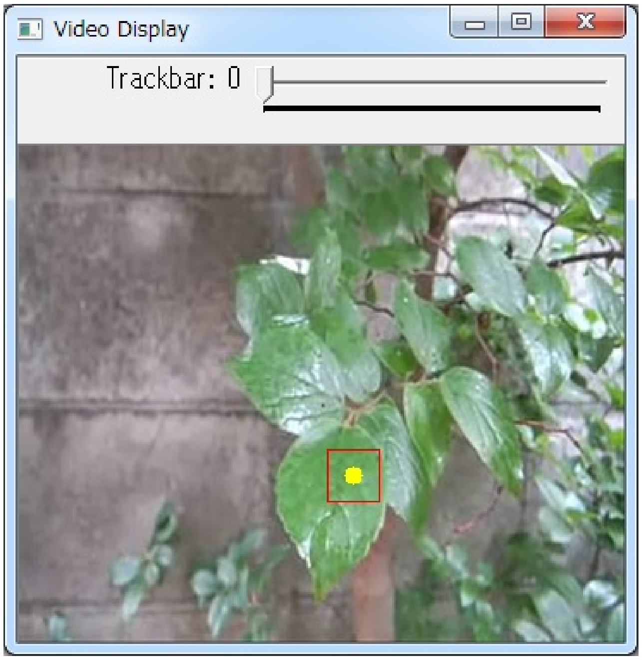


Figure 7: Area selection around the haptic interface point

3.2.1 Feature point selection and tracking

As shown in figure 7, we limit the area of the feature point selection to an area of a specified size around the haptic interface point. This limits the haptic feedback only to the selected object rather than the entire image. Consequently, this eliminates the one drawback of the past system by avoiding the background noise. We used the Shi and Thomasi algorithm for feature point selection and the pyramid luces kanade algorithm for feature point tracking [2] [4] [7].

Figure 8 shows the result of the selected feature points around an area of a haptic point of a randomly selected frame of the video. In this figure, red point shows the image point corresponding to the haptic interface point and the blue points show the feature points selected around 30 x 30 area around the haptic interface point. According to this figure, we can conclude that the feature points are coherent to the area around the haptic point and the distribution of those feature points are close and around the haptic interface point. Hence those feature points are good features to use in the optical flow algorithm.

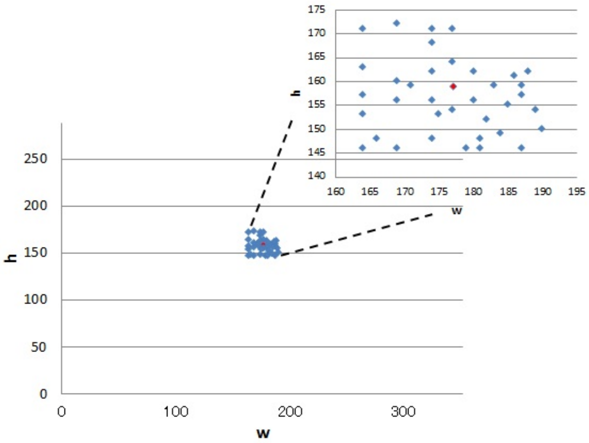


Figure 8: Features around the haptic interface point

Figure 9(b) shows the flow of points in a two consecutive frames in an image sequence. '*' shows the position of the feature points in the first frame and the 'o' shows the position of that feature points in the next frame. Notably, these image points are within the area of 30x30. Figure 9(c) shows the data about optical flow of the above feature points. The length of the arrow shows the magnitude of the optical flow and the direction of the arrow shows the direction of the optical flow.

3.2.2 Velocity calculation

We use the technique called optical flow which shows the object movement in the video [8]. As shown in figure 9, the length of a particular feature point in two subsequence frames at time t and $(t+\Delta t)$ represents the flow deatails.

We use velocity of a feature point to estimate the motion. The velocity of a feature point is calculated using the equation (2).

$$\vec{V_i}(t) = \frac{p_i(t+\Delta t) - p_i(t)}{\Delta t} \qquad (2)$$

Here $p_i(t)$ and $p_i(t+\Delta t)$ represent the position of a particular feature point in two consecutive frames. Those feature within the area of 30x30 squares around the haptic interface point.

If there are N feature points which have a motion , then the velocity is given by the average velocity of feature points, as shown in equation (3).

Figure 9(a): An image frame

Figure 9(b): Flow of points in two consecutive frames

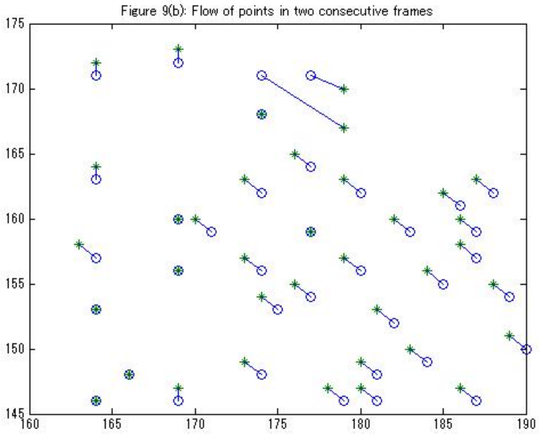


Figure 9(c): Optical Flow

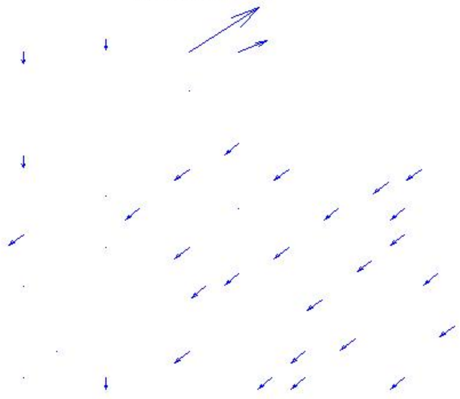


Figure 9: Features distribution in two consecutive frames

$$\vec{V_f}(t) = \frac{1}{N}\sum_{i=1}^{N}\vec{V_i}(t) \tag{3}$$

3.2.3 Motion force rendering

In this section, we discussed how force is rendered in relating to the selected object. To generate force feedback to user, through SPIDAR-G device we use velocity mapping approach. We use the nonlinear gain controller method proposed in [9][10] for this purpose. The nonlinear gain controller method reduces haptic jitter by reducing force for high velocities and increasing for low velocities. This enables a smooth and realistic sensation to user.

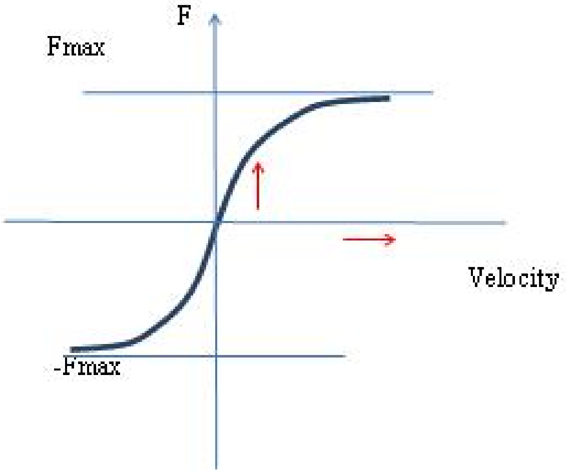


Figure 10: Purpose of two proposed methods

In the nonlinear gain controller method, we used a nonlinear function to map the velocity into force. The resulting feedback force to sense the motion of objects is shown in equation 4.

$$F(t) = f(\vec{v(t)}) \tag{4}$$

Our criteria in selecting a nonlinear function was the ability of avoiding haptic jitter by bringing down the feedback force into the sensible region of SPIDAR-G. As shown in figure 10, due to the S-shape behaviour, the sigmoid function proved to be a good candidate. However, as the velocity needs to be zero when changing the moving direction, we got an additional requirement such that the selected sigmoid function needs to go through the origin. Therefore, we selected the inverse tangent function, of which the corresponding candidate as shown in equation (5).

$$\vec{F(t)} = \frac{2\times F_{max}}{\pi} tan^{-1}(\alpha \times \vec{v(t)}) \tag{5}$$

Here α is chosen as 0.01.

Resulting velocity and force feedback of the SPIDAR-G for the image sequence calculated using equation (3) and (5), is shown in figure 11. Here, the resulting force is limited to 1N, it is within the sensible region of SPIDAR-G and hence user can smoothly feel the movement of the objects in the video.

4 EXPERIMENTAL EVALUATION

We evaluated the usefulness of haptization for video viewing with our proposed method to real users by evaluating the system with and without haptic feedback. The users are experienced users of the SPIDAR system and we used different videos of windy environments such as a moving trees, leaves, bushes and flowers. Those videos are 2D image sequences. Moreover, the camera is not moving in these videos.

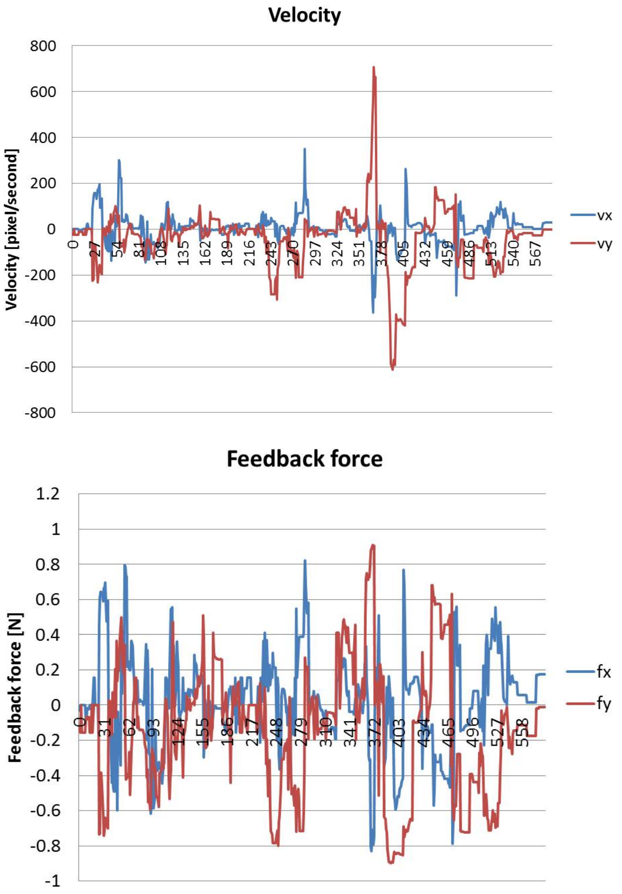


Figure 11: Changes of velocity and feedback force

For each user, we randomly assigned a video and conducted the experiment. After the experience with the image sequence with and without haptic feedback, we asked the users to rate the system based on their experience using a questionnaire. Here, we evaluated the system using three aspects: interactivity, sensitivity and reality. Interactivity evaluates whether the user can interact with the objects in the video based on their desires. Sensitivity evaluates whether the system enables getting a real feeling of the scene. Reality evaluates whether user feels that he or she as better involved in the scene, rather than as they appear to be imagined. In the questionnaire, users were asked to rate their experience for each aspect in both situations. (i.e. with and without haptic feedback) in a scale of 'Bad', 'Poor', 'Average', 'Good' and 'Very good'. The weighted average values of user's responses in each aspect with and without haptic feedback have been plotted in the graph in Figure 12.

We define the weighted average values as shown in equation (6), where R_{Bad}, R_{Poor},$R_{Average}$,R_{Good} and $R_{Verygood}$ represent the number of responses on the scale. We applied this to both cases of with and without haptic feedback.

$$WAV = \frac{(R_{Bad}*1)+(R_{Poor}*2)+(R_{Average}*3)+(R_{Good}*4)+(R_{Verygood}*5)}{Total responses} \tag{6}$$

According to figure 12, it is clear that all the three aspects with haptic feedback outperform without haptic feedback. Moreover, the factors of interactivity and sensitivity with haptic feedback have a significant difference from the case of without haptic feedback. This shows that the ability of users to interact with the desired objects in the video and the ability of using a third sense to create their experience with the video have a significant impact on the final viewing experience.

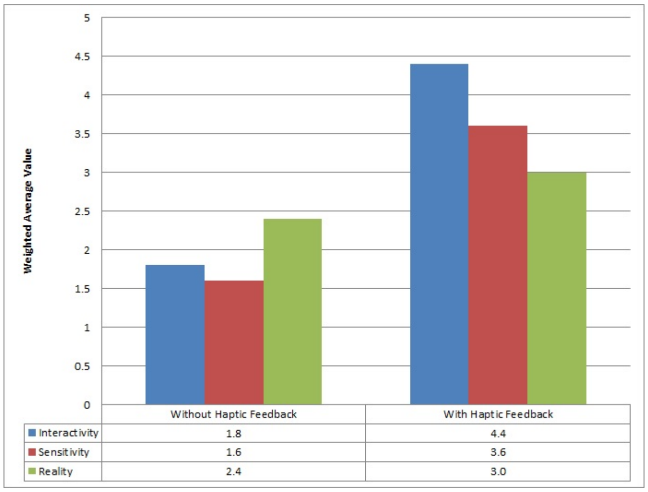


Figure 12: Distribution of responses of three aspects

5 CONCLUSION

This research is an extension to our previous research enabling users to interact with videos. Here we attempt to overcome the drawbacks identified in our previous approach; (1) Higher background noise in the case of video with object rich environments (2) users being passive in feeling the movement of objects. As a solution, we propose enabling the user to point a particular object of which she or he prefers to feel the movement. The effect of background noise has been reduced by limiting the sensible region for motion force calculation to an area of specified size around the selected point by the user. The feature point selection and tracking has been done using the Shi and Thomasi method and the Pyramidal Lucas Kanade method. Furthermore, for motion force generation, we use the nonlinear gain controller method, proposed in our previous research.

Testing this system with real users revealed that the system enhances the viewing experience along the dimensions of interactivity, sensitivity and reality.

6 FUTURE WORK

Despite its manifestation of enhanced viewing experience for object rich videos, the method has its limitations and hence, ample room for further improvements exists.

At first, it is highly necessary to test the system with more real users to attain a more concrete conclusion about the usefulness of the system. Even though we are generating a 2D feedback force through SPIDAR-G, the haptic pointer has the capability of working in 3-Dimensional space. Hence we may use the 3D coordinates of haptic pointer in which the third dimension will be used to change the region of interest around the haptic interface point. Changing the ROI along the depth coordinates enables users to better select different objects in the video. For example, if the video scene contains a moving tree with fruits, the user can either feel the movement of the whole tree or the movement of a single fruit. We also intend to enable users to do some time deformation to an image sequence by either slowing down or fastening the timing of the sequence.

Finally, the motion of the objects are not only translational but also rotational. We believe that it would be interesting and highly necessary to improve this method to incorporate rotational features as 3D technologies are becoming increasingly popular. Since SPIDAR-G is a 6 DOF device, we hope that this could be implemented through SPIDAR-G.

ACKNOWLEDGEMENTS

The authors wish to thank all the participants of the experiemnt. This work was supported in part by Grants for Excellent Graduate Schools, MEXT,Japan.

REFERENCES

[1] Surround haptics :immersive tactile experiences. `http://www.disneyresearch.com/project/surround-haptics-immersive-tactile-experiences/`. Accessed: 2012-10-03.

[2] F. Abdat, C. Maaoui, and A. Pruski. Real time facial feature points tracking with pyramidal lucas-kanade algorithm. In *Proceedings of the 17th IEEE International Symposium on Robot and Human Interactive Communication*, 2008.

[3] K. Akahane, S. Hasegawa, Y. Koike, and M. Sato. A proposal of a high definition haptic rendering for stability and fidelity. In *Proceedings of the 16th International Conference on Artificial Reality and Telexistence Workshop*, pages 162–167, 2006.

[4] J.-Y. Bouguet. *Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the Algorithm*. Intel Corporation, Microprocessor Research Lab, 2000.

[5] J. Cha, M. Eid, and A. E. Saddik. Touchable 3d video system. *ACM Transactions on Multimedia Computing, Communications and Applications*, 5(4), 2009.

[6] N. Dindar, A. M. Tekalp, and C. Basdogan. Immersive haptic interaction with media. In *Proceedings of the Visual Communications and Image Processing*, Huangshan, China, 2010.

[7] B. K. P. Horn and B. G. Schunck. Determining optical flow. *AI(17)*, (1-3):185–203, 1981.

[8] R. A. J. Sommer and D. Grest. Sparse optical flow for 3d objects. 2008.

[9] A. Jayasiri, K. Akahane, and M. Sato. Feeling the motion of object in a dynamic image sequence through haptic interface. In *Proceedings of the International Conference on Advances in ICT for Emerging Regions (ICTer2012)*, pages 19–26, December,2012.

[10] A. Jayasiri, K. Akahane, and M. Sato. Haptic rnedering of dynamic image sequence using string based haptic device spidar. In *Proceedings of the Joint Virtual Reality Conference of ICAT-EGVE-EuroVR (JVRC‘12)*, pages 9–15, October,2012.

[11] S. Kim, Y. Koike, and M. Sato. Tension based 7 dofs force feedback device: Spidar-g. *Transactions On Control, Automation, and Systems Engineering*, 4(1):9–16, 2002.

[12] P. Lemmens, F. Crompvoets, D. Brokken, J. van den Eerenbeemd, and G.-J. de Vries. A body-conforming tactile jacket to enrich movie viewing. Third Joint Eurohaptics Conference and Symposium on Haptic Interfaces for Virtual Enviornment and Teleoperator Systems, 2009.

[13] X. Liu, K. Akahane, M. Isshiki, and M.Sato. Design and implementation of an image haptization system. pages 260–262. 3DSA 2010 proceedings, 2010.

[14] X. Liu, A. Jayasiri, K. Akahane, and M. Sato. Proposal of a hapitzation system for depth images with local deformation. pages 147–150. 3DSA 2012 proceedings, 2012.

[15] S. O‘Modhrain and I. Oakley. Touch tv: Adding feeling to broadcast media. In *Proceedings of the 1st European Conference on Interactive Television: from Viewers to Actors*, pages 41–47, UKBrighton, UK, 2003.

[16] S. O‘Modhrain and I. Oakley. Adding interactivity:active touch in broadcast media. In *Proceedings of 12th International Symposium on Haptic Interfaces Virtual Enviornment and Teleoperator System (HAPTICS‘04)*, pages 293–294, 2004.

[17] M. Sato. A string-based haptic interface spidar. In *ISUVR2006*, pages 1–5, Yanjin, China, 2006.