Joint Virtual Reality Conference of ICAT - EGVE - EuroVR (2012)
R. Boulic, C. Cruz Neira, K. Kiyokawa, and D. Roberts (Editors)

Haptic Rendering of Dynamic Image Sequence Using String based Haptic Device SPIDAR

Anusha Jayasiri, Katsuhito Akahane, Makoto Sato

Precision and Intelligence Laboratory, Tokyo Institute of Technology, Japan

Abstract

This paper concerns how to associate haptic signals with a dynamic image sequence or in other words a video to feel the haptic motion. Nowadays, there is a significant evolvement in haptic technologies and they are being used in a wide range of application areas. With the invention of digital multimedia and immersive displays, significance of exploring new ways of interacting with video media has grown up. However, the incorporation of haptic interface technology into a dynamic image sequence is still in its infancy. Rather than just seeing and hearing a video, viewers' experience can be further enhanced by letting them feel the movement of the objects in the video through haptic interface, as it is an additional sensation to seeing and hearing. The objective of this research is to use haptic interface technology to interact with a dynamic image sequence and enable the viewers to feel the motion force of objects in the video beyond passive watching and listening. In this paper, we have discussed how to feel the haptic motion, which is computed from frame to frame calculation of velocity using optical flow. For haptic motion rendering, we have evaluated two methods using a gain controller and using a non-linear function to identify a better method. To interact with the video we used the string based haptic device, SPIDAR, which provides a high definition force feedback sensation to users.

1. Introduction

Usually, haptic rendering is fully integrated with virtual reality (VR) worlds. This enables users to touch, feel and manipulate virtual objects through haptic interface [BS01]. Haptic interface can generate mechanical signals, which stimulate those human kinaesthetic and touch channels [HACH*04]. Nowadays, there is a significant evolvement in haptic technologies and they are being used in a wide range of application areas in training [BDK*04] [WWZ09], education [SLM*08], entertainment and 3D interaction [MJS04].

With the invention of digital multimedia and immersive displays, significance of exploring new ways of interacting with video media has grown up. For example, in three-dimensional television (3D-TV), viewers can see objects in true dimensions and natural color, providing a natural viewing experience with advanced audio technologies. The Viewers' experience can be further enhanced if viewers can get the feeling of the movement of the objects in the video or in other words a dynamic image sequences through haptic interface, as it is an additional sensation to seeing and hearing.

However, the incorporation of haptic interface technology into video media is still in its infancy. As Dinder et al [DTB10] discusses, there are 3 types of haptic effects which are cooperating with haptic interaction with video media. Those are haptic structure, haptic texture and haptic motion. Haptic structure refers to the touching or getting the feeling of the geometry of an object in the video scene. Haptic texture refers to the rendering of surface properties such that roughness of various objects in the video scene. Haptic motion refers to the rendering forces related to the moving objects in the scene. In this research we address the haptic motion part. Consequently, this research is an attempt to use haptic technology to interact with a dynamic image sequence with the objective of enabling the viewers to feel the motion force of objects in the video beyond simply passive watching and listening.

To achieve this, two methods using a gain controller and using a non-linear function have been proposed and evaluated using a haptic device SPIDAR, which is developed by Sato laboratory, Tokyo Institute of Technology [KKS02]. SPIDAR, which stands for Space Interface Device for Artificial Reality, is a string-based haptic device and can be used in various types of virtual reality applications. Different versions of SPIDAR systems are currently available and for this research we used the SPIDAR-G system shown in Figure 1,

DOI: 10.2312/EGVE/JVRC12/009-015

Figure 1: *SPIDAR-G haptic device*

which is a grip type, tension based 6 degrees of freedom force feedback device. By connecting this device to a personal computer, it provides a high definition force feedback sensation to the user [AHKS06]. As a result, the viewer can feel the movement of the objects in the video into his/her hand through the grip of SPIDAR-G rather than just seeing it as illustrated in Figure 2.

Figure 2: *Haptic Interaction Scenario*

This paper discusses how to associate haptic signals with video media to feel the haptic motion. The paper is organized as follows: section 2 presents related work in haptic rendering with video media and using SPIDAR for image haptization. Section 3 presents our proposed approach includes feature point selection, feature point tracking, motion estimation and haptic motion rendering. We further elaborate on our proposed approach for haptic motion rendering using a gain controller and a non-linear function. Section 4 presents both analytical and experimental evaluation of the comparison of above two methods and section 5 discuss the conclusion and future work.

2. Related work

Incorporation of haptic technology into video media is not adequately researched in the haptic rendering field. This section summarized few of them.

Dinder et al [DTB10] 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 which 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.

O'Modhrain et al [OO03] have discussed how haptic interaction can enhance and enrich the viewer's experience in broadcast content. They proposed a touch TV project with the use of Gravis Xterminator Force and remote control handset to generate haptic cues for cartoons and live sports broadcastings which adds greater sense of immersion.

Cha et al [CKO*05] [CKH06] [CES09] have proposed a touchable 3D video system which provides haptic interaction with objects in a video scene. As a result it enables the viewers to actively touch a video scene through a PHANToM force feedback device and enable to physically explore the video content and feel various haptic properties such that texture, height map and stiffness of the scene. They have introduced depth image-based haptic representation (DIBHR) method to add haptic surface properties of the video media.

Kim et al [[KCL*07] have proposed a 3DTV system which enables not only enjoying a high-quality 3D video in real time, but also user can experiencing various user-friendly interactions such as free viewpoint changing, composition of computer graphics and haptic display

Even though SPIDAR has not been used for haptization in video media, it has been used in the context of images. Liu et al [LAIS10] have proposed a haptization system which provides the users with sense of touch on an image with local deformations using SPIDAR haptic interface In this research, user can interact with the image contents using SPIDAR-G haptic device.

3. Proposed approach

The proposed approach for this research can be illustrated using a block diagram as in Figure 3.

3.1. Feature point selection

Feature point selection is an important task of any computer vision and image processing application. Since feature selection is the starting point of many computer vision algorithms, the performance of the subsequent algorithm as well as the overall performance of the process basically depends on it.

Feature point selection finds which points are good to track. For example corners or good textures may be good feature points of an image. There are various methods exist for feature selection such as Harris, Canny, Sobel etc. Among those methods we used Shi and Thomasi algorithm

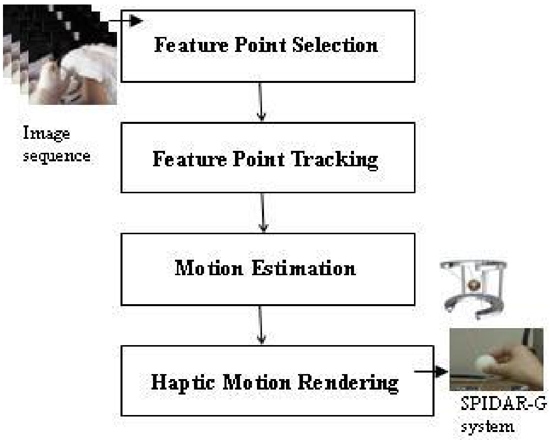


Figure 3: *Block diagram of the proposed approach*

for feature points selection in the image sequence.This algorithm is more efficient in interesting point detection and hence, the processing time of the overall process becomes less [AMP08]. This algorithm is based on the assumption that the brightness of a pixel does not change as it is tracked from frame to frame.

3.2. Feature point tracking

Feature point tracking identifies above features reliably from frame to frame and those features are tracked using feature matching. Also we need to measure motion of the objects between two frames in an image sequence without any prior knowledge about the content of those frames. To address the above concerns, we used optical flow technique [BK08].

Optical Flow is the distribution of apparent velocities of movement of brightness patterns in an image. Optical flow arises from relative motion of objects and the viewer [HS81]. The way an object moves when it is seen or followed in a video or sequence of images is known as optical flow [JSG08]. There are two types of optical flow methods namely dense optical flow and sparse optical flow. In dense optical flow methods, it associates velocity with every pixel in an image. Horn-Schunck method and block matching method are examples for this type of optical flow [BK08]. In practice, calculating dense optical flow is not easy because of the high computational cost. Alternatively, sparse optical flow techniques calculate velocities only on the points which have certain desirable properties.Lucas-Kanade method is an example for this type of optical flow [BK08] [JSG08].

We used Pyramid Lucas-Kanade Algorithm, which is a pyramidal implementation of the Lucas-Kanade feature tracker [Bou00]. At first in this technique, it solves optical flow at the top layer of the pyramid and then use the resulting motion estimates as the starting point for the next layer down. It continues going down the pyramid in this manner until it reach the lowest level. Therefore by using this method, it can track faster and longer motions [BK08]. This has less computation and therefore it could be easily adapted for real time applications [AMP08].

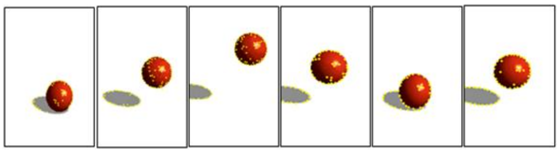


Figure 4: *Feature point selection with the Shi and Thomasi algorithm and tracking of those points in the subsequent frames using the Pyramid Lucas-Kanade method*

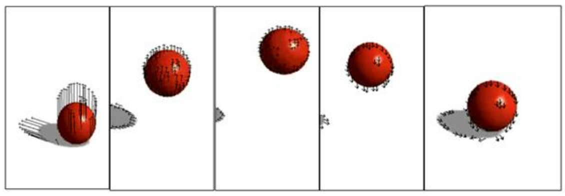


Figure 5: *Results of optical flow*

We tested with real image sequences and synthetic image sequences. We identified that, real image sequences have lots of background noices due to the object richness of the environment. As a result, it makes an unnecessary feature points and this complexity affects adversly to the objective of our research; i.e. haptic rendering of moving objects. Therefore, as this initial stage, we used a synthetic image sequence of a bouncing ball , which includes a single moving object and does not have any other objects in the background as the main source for the experiments and discussion of this research. Figure 4 shows the obtained results for feature point selection using Shi and Thomasi algorithm and tracking of those points using the Pyramid Lucas-Kanade method for the above metion synthetic image sequence of the bouncing ball. Figure 5 shows the results of the optical flow. In that figure, the direction of each arrow represents the direction of optical flow and the length of each arrow represents the magnitude of the optical flow.

3.3. Motion estimation

This section explains how we calculate the motion of an object in the image frame.

We use velocity of a feature point to estimate the motion. As shown in Figure 6, the position of a feature point in two subsequent frames at time t and $t+\Delta t$ can be represented

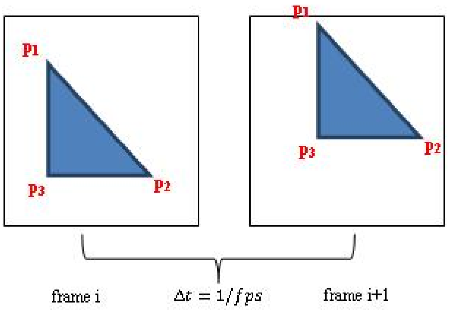


Figure 6: *Consecutive frames of an image sequence*

as $p_i(t)$ and $p_i(t+\Delta t)$. The velocity of the feature point can then be calculated using the equation 1.

$$\vec{v_i(t)} = \frac{p_i(t+\Delta t) - p_i(t)}{\Delta t} \tag{1}$$

Here, if there are N feature points in a frame, then the velocity of each frame is given by the average velocity of feature points in the image frame, as shown in equation 2.

$$\vec{v_f(t)} = \frac{1}{N}\sum_{i=1}^{N}\vec{v_i(t)} \tag{2}$$

The Figure 7, 8 and 9 shows the changes of the position, velocity and acceleration of the ball in each frame during each run.

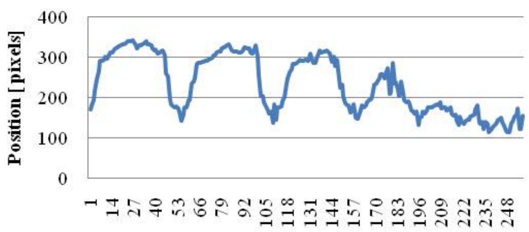


Figure 7: *Position changes of bouncing ball*

Bouncing ball is a good example for position, velocity and acceleration change against time because these graphs convey lots of information about speeding up, slowing down, rising or falling of the ball.

According to the Figure 7, we can easily recognize the occasions when the ball is falling and rising. As shown in Figure 8, when the ball is falling, the velocities on the graph are shown as negative values and the velocity is increasing. On the other hand, when the ball is rising, the velocities on the graph are shown as positive values and the velocity is

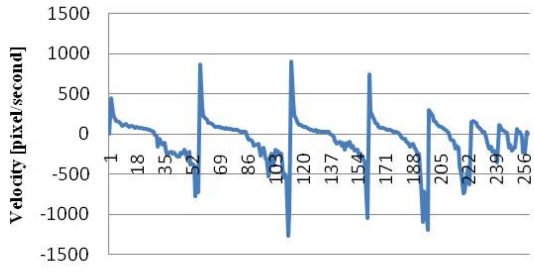


Figure 8: *Velocity changes of bouncing ball*

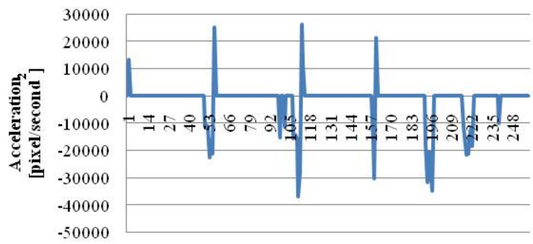


Figure 9: *Acceleration changes of bouncing ball*

decreasing. At the top of each bounce the velocity is zero because the ball changes its moving direction.

As the ball falls towards the floor its velocity increases and just before it hits the floor, its velocity becomes maximum.. Immediately after leaving the floor, i.e. at the start of the upward journey, its velocity is maximum and in it is in the upward direction. As the ball rises towards its highest position, its velocity approaches zero.

The Figure 7, 8 and 9 complies with the physics of a bouncing ball [Cro99] and hence we can conclude that our proposed method is accurate.

3.4. Haptic motion rendering

Haptic motion rendering means rendering of forces related to the moving objects in the scene. In this section we explain how we calculated forces based on the above velocity changes in the dynamic image sequence.

We used SPIDAR-G haptic device shown in Figure 1. This device is 6 degrees of freedom; 3 degrees of freedom for translation, 3 degrees of freedom for rotation enabled force-feedback device. This device has a grip and this grip 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 users hand through the grip.

However, high velocity produced high force and low velocity produced low force lead to an unrealistic sensation. To overcome this problem and to get a realistic sensation to the user we need to reduce force for high velocities and increase the force for low velocities. For this purpose, we proposed two alternative methods using a gain controller and a non-linear function to select the one performs better.

3.4.1. Method using a gain controller

Using the gain controller method, the feedback force is calculated from equation 3. This enables user to get the feeling of the movement of the object .

$$\vec{F(t)} = k \times \vec{v(t)} \tag{3}$$

Here k is a gain controller.

$$k = \frac{F_{max}}{v_{max}(T)} \tag{4}$$

Calculation of k is done as in equation 4 to control the feedback force within a sensible region for all velocity levels. In other words, the purpose of the gain controller k is to increase the feedback force for weak changes in velocity and decrease the feedback force for strong changes in velocity.

Here F_{max} is the maximum force output level of the SPIDAR-G for better sensation for this application. $V_{max}(T)$ is the maximum velocity of the dynamic image sequence at a time T ,which can be expressed as from equation 5.

$$v_{max}(T) = \Big\{ |v_{max}(t)| \qquad 0 \le t \le T \tag{5}$$

We analysed the pattern of k value using the previously mentioned image sequence of the bouncing ball. The Figure 10 shows the results of the changing k value for the image sequence.

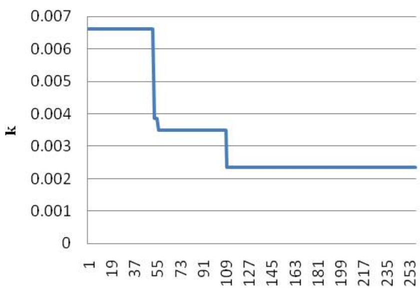


Figure 10: *Performance of the k value*

3.4.2. Method using a non-linear function

Similar to the previous method, the purpose of using a non-linear function is to maintain the feedback force in the sensible region by decreasing the feedback force for high velocities and increasing the feedback force or low velocities. Figure 11 illustrate this idea in a graphical form.

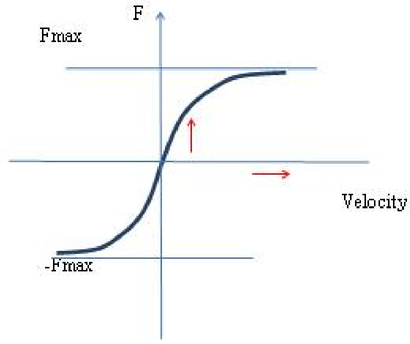


Figure 11: *Purpose of the non-linear function*

Using a non-linear function, the feedback force to sense the motion of objects is calculated from equation 6.

$$\vec{F(t)} = f(\vec{v(t)}) \tag{6}$$

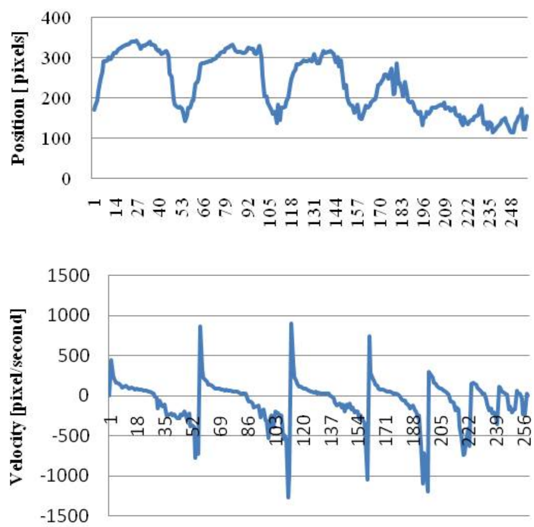
To achieve the above objective, we tested with different sigmoid functions and selected the simple sigmoid function of inverse tangent function and the corresponding conversion is shown by equation 7.

$$\vec{F(t)} = \frac{2 \times F_{max}}{\pi} tan^{-1}(\alpha \times \vec{v(t)}) \tag{7}$$

Here α is chosen as 0.01.

4. Evaluation

4.1. Performance analysis



Force feedback of the SPIDAR-G for the image sequence calculated using equation 3 and 7. Notably, resulting force

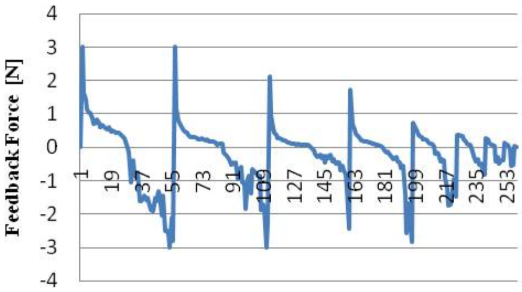


Figure 12: *Resulting feedback force using gain-controller method*

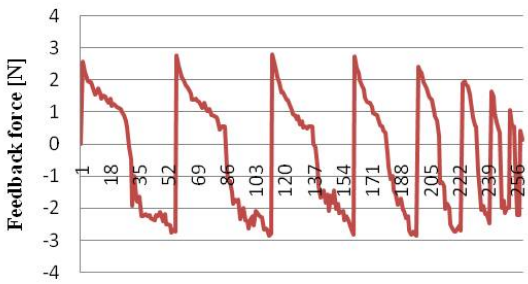


Figure 13: *Resulting feedback force using a non-linear function method*

is within the sensible region of the SPIDAR-G and hence user can get a more realistic feeling of the movement of the objects in the video.

We can compare the resulting feedback force using two methods: a gain controller and non-linear function with respect to the position and velocity changes by using Figure 12 and 13 . It is clear from these figures, non-linear function method outperforms the gain controller method, since it increases the feedback force for low velocities than the gain controller method and decreases the feedback force for high velocities reasonably within the sensible region of the SPIDAR-G haptic device. In other words, user can feel the movement of the objects in the video even for smaller changes of the velocity. As a result user can get a continuous and realistic feeling of the movement of the ball even when the ball is in the air.

4.2. Experimental evaluation

For further verification, which method gives better feeling about the movement of the objects in the video, we conducted an experimental evaluation for both methods with the participation of real users. For each user we conducted the experiment using two methods, i.e gain controller method and the non-linear function method. We used a questioner to get the users feedback for each method. In the questioner the user was asked to rate the each method in a scale of 'Bad', 'Poor', 'Average', 'Good' and 'Very good' based on their feeling of movement of the object.The results are shown in the Figure 14 and 15 .

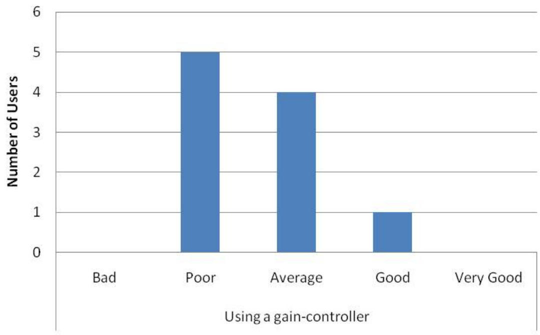


Figure 14: *Users' responses when using a gain controller method*

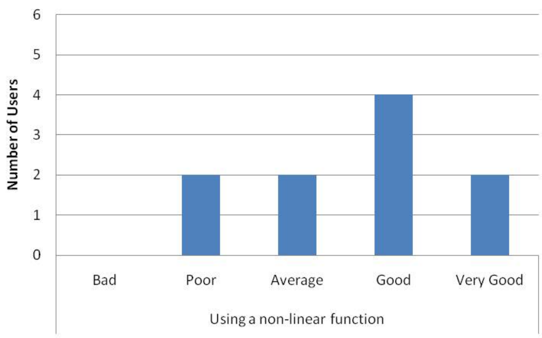


Figure 15: *Users' responses when using a non-linear function method*

We used the weighted average of the user's responses to compare the outcome using two methods. When the gain controller method is using, the average response is 2.6 and When the non-linear function method is using, its value is 3.6. It is clear that , using a non-linear function gives better feeling of the movements of the object for the synthetic image sequence of the bouncing ball. This complies with our previous result.

5. Conclusion and future work

In this research we concern how to associate haptic signals with a dynamic image sequence or in other words a video to feel the haptic motion. To achieve the above objective, we experimentally evaluated two methods i.e. the gain controller method and the non-linear function method, with the objective of identifying the better one. Based on the feedback of real users, we can conclude that using a nonlinear function is more effective than using a gain controller because user can get a continuous feedback force and hence, can get the continuous feeling of the movement of the objects in the video. This was revealed by both analytical and experimental evaluations.

We have identified certain limitations of this research, which affect the quality of the final outcome. In our future research, we expect to adress those limitations and improve the outcome of the system.

First, the videos we used had only one moving object. However, real videos are objects rich enviornments with multiple moving objects. Thus, it is highly necessary to research on how to interact with multiple objects in a object rich enviornments. Further, real image sequences include lots of background noise, which can be affect the feedback force of the rendered object. We further expect to improve our method to eliminate such background noise and improve the output.

Second, we belive that it would be interesting and highly necessary to improve this method to render 3D motion from 2D image sequence as 3D technologies will increasingly become popular in the future. In consequece it will allow the users to get the 6 degrees of freedom force back based on rotation and translation of the motion of objects.

Third, since SPIDAR-G is a grip type haptic device, user can feel the movement of the object only through the grip to gripping hand. However this is not adequate to a realistic sensation. Therefore, another line of the future work is to improve the system by identification of a suitable device, which enables the user to feel the movements of the objects to the whole body .

References

[AHKS06] AKAHANE K., HASEGAWA S., KOIKE Y., SATO M.: 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* (2006), pp. 162–167. 2

[AMP08] ABDAT F., MAAOUI C., PRUSKI A.: 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

[BDK*04] BASDOGAN C., DE S., KIM J., MUNIYANDI M., KIM H., SRINIVASAN M. A.: Haptics in minimally invasive surgical simulation and training. *Proceedings of the IEEE Computer Graphics and Applications 24*, 2 (2004), 56. 1

[BK08] BRADSKI G., KAEHLER A.: *Learning OpenCV*. O'Reilly Media, Inc, Gravenstein Highway North, Sebastopol, CA 95472, 2008. 3

[Bou00] BOUGUET J. Y.: *Pyramidal Implementation of the Lucas Kanade Feature Tracker Description of the Algorithm*. Intel Corporation, Microprocessor Research Lab, 2000. 3

[BS01] BASDOGAN C., SRINIVASAN M. A.: Haptic rendering in virtual environments. *Handbook of Virtual Environments: Design, Implementation, and Applications* (2001), 117–134. 1

[CES09] CHA J., EID M., SADDIK A. E.: Touchable 3d video system. *ACM Transactions on Multimedia Computing, Communications and Applications 5*, 4 (2009). 2

[CKH06] CHA J., KIM S., HO Y.: 3d video player system with haptic interaction based on depth image-based representation. *IEEE Transaction on Consumer Electronics, Vol 52*, 2 (2006). 2

[CKO*05] CHA J., KIM S., OAKLEY I., RYU J., LEE K. H.: Haptic interaction with depth video media. *Advances in Multimedia Information Processing - PCM 2005* (2005). 2

[Cro99] CROSS R.: The bounce of a ball. *American Journal of Physics 67*, 3 (1999), 222–227. 4

[DTB10] DINDAR N., TEKALP A. M., BASDOGAN C.: Immersive haptic interaction with media. In *Proceedings of the Visual Communications and Image Processing* (Huangshan, China, 2010). 1, 2

[HACH*04] HAYWARD V., ASTLEY O. R., CRUZ-HERNANDEZ M., GRANT D., ROBLES-DE-LA-TORRE G.: Haptic interfaces and devices. *Journal of Sensor Review 24*, 1 (2004), 16–29. 1

[HS81] HORN B. K. P., SCHUNCK B. G.: Determining optical flow. *AI(17)*, 1-3 (1981), 185–203. 3

[JSG08] J. SOMMER R. A., GREST D.: Sparse optical flow for 3d objects. 3

[KCL*07] KIM S., CHA J., LEE S., RYU J., HO Y.: 3dtv system using depth image-based video in mpeg-4 multimedia framework. 3DTV Conference. 2

[KKS02] KIM S., KOIKE Y., SATO M.: Tension based 7 dofs force feedback device: Spidar-g. *Transactions On Control, Automation, and Systems Engineering 4*, 1 (2002), 9–16. 1

[LAIS10] LIU X., AKAHANE K., ISSHIKI M., SATO M.: Design and implementation of an image haptization system. In *3DSA2010(International Conference on 3D Systems and Applications)* (2010), pp. 260–262. 2

[MJS04] MORRIS D., JOSHI N., SALISBURY K.:. In *Haptic Battle Pong: High-Degree-of-Freedom Haptics in a Multiplayer Gaming Environment* (Proceedings of Experimental Gameplay Workshop at Game Developers Conference (GDC)'04, 2004). 1

[OO03] O'MODHRAIN S., OAKLEY I.: Touch tv: Adding feeling to broadcast media. In *Proceedings of the 1st European Conference on Interactive Television: from Viewers to Actors* (UK-Brighton, UK, 2003), Brighton. 2

[SLM*08] SATO M., LIU X., MURAYAMA J., AKAHANE K., ISSHIKI M.: A haptic virtual environment for molecular chemistry education. *Transactions on Edutainment I, LNCS 5080* (2008), 28–39. 1

[WWZ09] WU J., WANG D., ZHANG Y.: Virtual fixture based haptic rendering of handwriting. In *Human-Computer Interfaces, and Measurement Systems* (Hong Kong, China, 2009), IEEE International Conference on Virtual Environments, pp. 16–21. 1