Development of the High Frame Rate 3D System

Yoshihiko Kuroki
Sony Corporation, Technology Development Group

Abstract--* *Prototypes of a 240fps single lens 3D camera, a 240fps 4Kx2K projector and a 240fps LCD display have been developed. These high frame rate (HFR) equipment bring extremely high quality of motion images. The camera has the following major advantages: 1) comfortable 3D image viewing by the absence of vergence-accommodation conflict, 2) simple and accurate control of the zoom / focus of a single lens, 3) 3D/2D compatibility - without 3D glasses, normal 2D images can be viewed.

I. INTRODUCTION

In order to realize the best motion images, we first measured the resolution limit of the human eye for moving objects and found that an object moving at a speed of 60 dps (degrees per second) is the maximum that can be resolved even at a very low spatial frequency of 7.5 cpd (cycles per degree). We then made psychophysical evaluations of the test images from 60fps (frames per second) to 480fps and found that a frame rate of around 250fps is the perception limit beyond which there is no blur and jerkiness. For the cinema and TV, 240fps is the ideal frame rate because 240 is a simple multiple of 24 and 60 [1][2]. Some studies related to the jerkiness have been made [3][4] and now it is known that high frame rate can reduce jerkiness.

We have developed prototypes of a 240fps frame rate and HD resolution camera with 3D capability [5], a 240fps 4Kx2K projector [6] and a 240fps LCD monitor. The combination of the 240fps 3D camera and the 240fps display produces natural and high-quality motion images.

II. PERCEPTION OF BLUR AND JERKINESS

When we see a moving object on a hold-type display of 60Hz vertical frequency, even if the image is crisp, we perceive blur corresponding to 1/60 second because the retinal image slips as the eyes smoothly follow the trajectory of movement of the object, while the image keeps its position on the display while the frame is displayed. A high speed shutter and impulse-type display reduce such blur perception but images displayed at a common frame rate such as 60fps often cause jerkiness, especially in fixated viewing. A high frame rate therefore can improve picture quality by minimizing both blur and jerkiness. Figures 1 and 2 show that the perception limit for both blur and jerkiness is around 250 fps. When we consider the 24fps frame rate used in cinema, 240fps is an ideal frame rate because it provides the highest motion quality using straightforward image processing as 240 is a common multiple of 24 and 60.

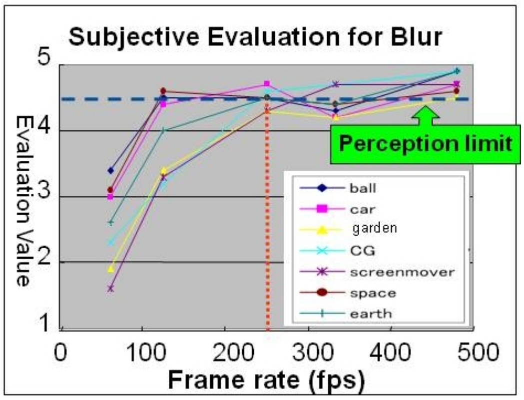

Fig. 1. Results of the Psychophysical Evaluation for Blur

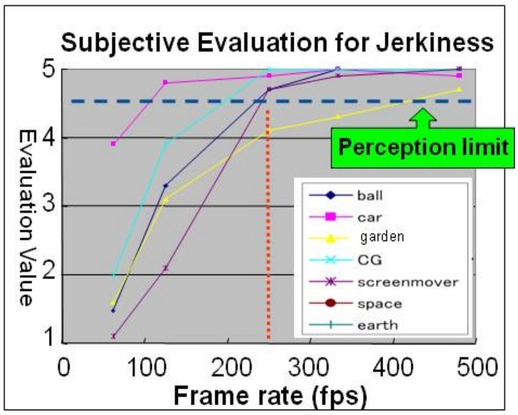

Fig. 2. Results of the Psychophysical Evaluation for Jerkiness

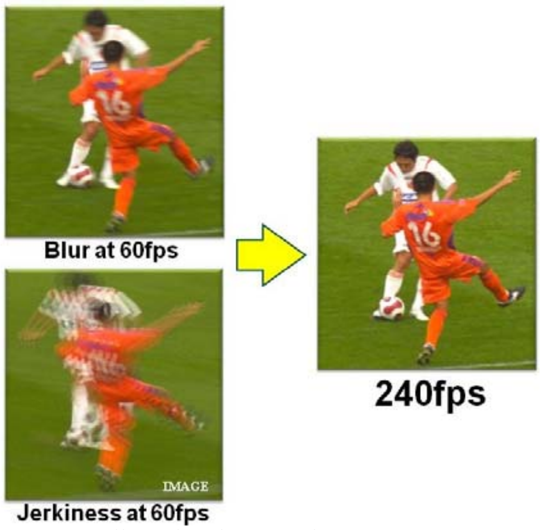

Fig. 3. The effects of 240fps' image for Blur and Jerkiness

III. 240 FPS SINGLE LENS 3D CAMERA PROTOTYPE

The first prototype of our 240fps camera was 2D and it produced good motion images with even more depth impression than at 60fps in most cases, because each frame was clearer than at 60fps. It is thought that the images of 240fps have more effective pictorial depth cues such as occlusion, shading, texture gradient, linear perspective, aerial perspective, and so on.

But there still seemed to be something missing. The flying football or tennis ball was hard to visualize in open space because of less pictorial cues. So in order to bring more reality to high frame rate motion images, we developed a 3D capability. The type of 3D optics which we chose was the single lens 3D system because it enables the 240fps camera to capture images which can be viewed with very natural depth impression. Our concept of the single lens 3D system took its hint from the conventional system shown in Figure 4.

A. Conventional idea of single lens 3D system

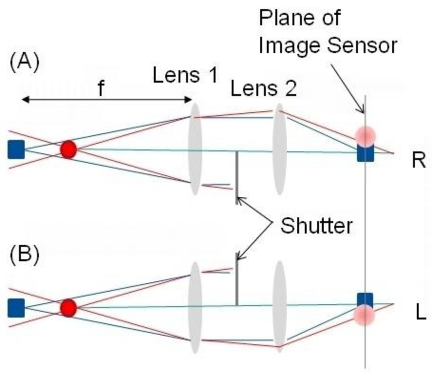

Fig. 4. Conventional single lens 3D system

In Figure 4, when the optics is focused on the rectangular object on the optical axis, the rays, after passing through Lens 1, become parallel, and then are converged by Lens 2, forming the image in the position of the optical axis at the image sensor. The circular object is in the inner defocus area, so the image position is behind the image sensor, resulting in a blurred image at the image sensor. When the shutter blocks half the rays as in Figure 4(A), the blurred image is formed at a position displaced to the upper direction. When the shutter blocks the other half as in Figure 4(B), the displacement of the blurred image is downward. When these two images are viewed binocularly, a 3D image is perceived. Conventional video camera uses two liquid crystal shutters operating alternately at the (interlaced) field frequency. This means that pairs of disparity images are captured with a time lag of one field duration. This time lag causes artifacts in the depth dimension in motion images. For example, when the camera follows a soccer player running across the scene, if the left eye's image is at the field time previous to the right eye's, the leg moving to the right in the picture locking on to the player appears to be farther than the body of the player, and the other leg moving in the opposite direction appears to be nearer, resulting in a motion image which looks unnatural.

B. Simultaneous- capture single lens 240fps 3D system

Thus we developed an optical system for the single lens 3D camera which captures the left and right images simultaneously, as shown in Figures 5 and 6.

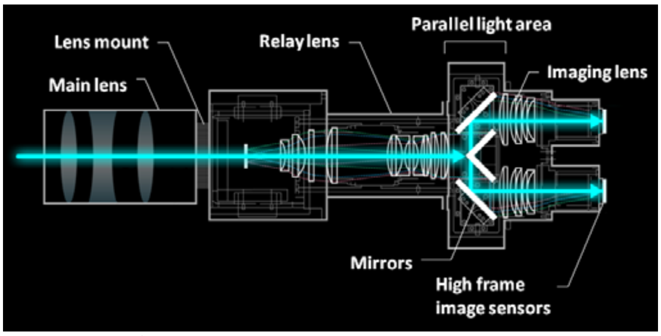

Fig. 5. Optics of HFR 3D Camera

Fig. 6. Appearance of HFR 3D Camera Prototype

The basic mechanism for capturing 3D images is the same as the single lens 3D camera with shutters, which we described previously. The key difference between these two systems is the use of mirrors instead of shutters. The use of mirrors realizes the simultaneous capturing of left and right images, and it also enables capturing brighter images than with the shutter method because there is no loss of light. Shutters result in the loss of half of the light. Therefore, the proposed system brings higher quality of 3D motion images by avoiding the time-lag effect and also doubles the sensitivity.

We measured the effective L-R viewpoint separation of this camera by measuring the distance perpendicular to the optical axis that the camera needs to be moved to the right to keep an object in the same position in the left image as it was in the right image before moving. The results obtained using two different lenses are as follows:

・Lens a: Fujinon HAe10x10, F1.8, f =100mm, the shooting distance is 6.5m and the effective separation was 20mm.

・Lens b: Fujinon HA42x13.5, F2.8, f=100mm, the shooting distance is 6.5m and the effective separation was 12mm.

We predicted that the effective separation corresponds to the distance between the centers of gravity of the half-disks of the pupil as shown in Figure 7.

The values of this distance calculated using f/F as the diameter of the pupil, where f is the focal length and F is the F number, were 23.6mm and 15.2mm, respectively, which is in fair agreement with the measured results.

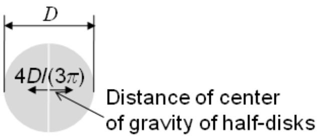

Fig. 7. Distance of center of gravity of half-disks of pupil

The images captured by the HFR 3D Camera are shown in Figures 8-9.

Fig. 8. Sample image shot by the HFR 3D Camera

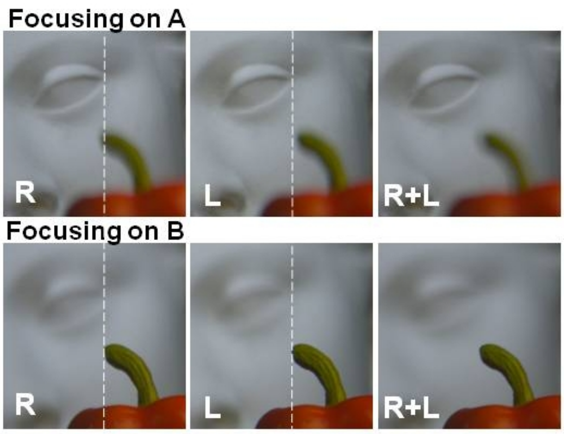

Fig. 9. Close-up of images shot by the HFR 3D Camera and the averaged images

The image R+L in each row in Figure 8 is the frame-averaged image of the right and left images marked R and L, respectively. The upper row images show that when the camera is focused on A in Figure 9, a plaster statue, it is in almost exactly the same position in the right and left images, whereas the image of B, a paprika, is defocused and is in a different positions in the right and left images. However, the difference is on the order of the blur width so that the frame-averaged image, R+L, appears to be the same as an ordinary 2D image. The lower row images show the case of focusing on B, the paprika, in which case B is in almost exactly the same position in the right and left images, and the image of the statue is defocused and is in a different position of the order of the blur width. This means that when these left and right images are viewed with a 3D display, for example a polarization-modulated display using polarized glasses, 3D images are perceived, whereas when we view the same display without wearing the polarized glasses, we see ordinary 2D images.

C. Human sensitivity to binocular disparity in depth perception

The important technical background of the single lens 3D camera is the high sensitivity to disparity of the human eyes. As shown in Figure 10, human eyes are able to perceive depths with as small disparity as in the order of arc seconds. This acuity is much higher in comparison to the usual sensitivity to spatial frequency of contrast which is in the order of arc minutes. This means that when we see an HD image with horizontal 2k resolution at a 3H distance, the spatial resolution limit of the human eyes is almost equal to the pixel size, but the sensitivity to disparity is on the order of sub-pixels.

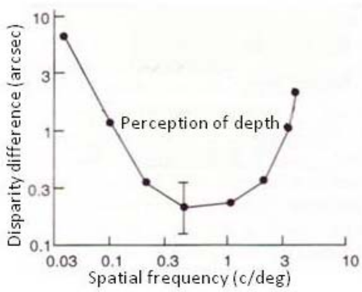

Fig. 10. Human sensitivity to disparity (modified from [6])

D. Block diagram of the HFR 3D Camera

The block diagram of the complete HFR 3D camera incorporating the optical design described above is shown in Figure 11.

The left and right image sensors are 240Hz frame rate, HD resolution and color-filtered CMOS sensors. For recording, we used a custom built solid state disc recorder.

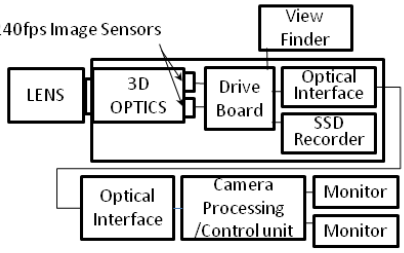

Fig. 11. Block diagram of HFR 3D Camera

IV. 240FPS 4KX2K SXRD PROJECTOR PROTOTYPE

A 240fps 4Kx2K(4096H x 2160V) resolution SXRD projector prototype was developed as shown in Figure 12. Evaluation of the blur for three frame rates: 60fps, 120fps and 240fps showed that 240fps brings the highest performance for avoiding blur [7].

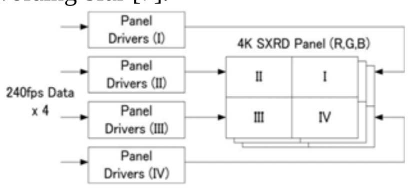

Fig. 12. Block diagram of the interface of the customized 240fps 4kx2k projector

V. 240FPS LCD MONITOR PROTOTYPE

240Hz LCD TVs are now popular in consumer products and they perform much better on blur and jerkiness than conventional 60Hz LCD TVs. These new TVs can create frames from the usual signal source, such as 60Hz or under, using frame interpolation. As the interpolated frame is calculated using the motion vectors from sequential frames, when pixels of a part of an object in a frame are occluded in another frame, the calculation of the motion vectors becomes difficult. To avoid this difficulty, we developed a 240fps source-inputting circuit for the LCD monitor, as shown in Figure 13. The four 60fps input signals are in parallel and are produced from the original 240fps source of our camera or disc recorder. Tests by displaying 240fps images captured with our 240fps camera showed extremely good quality.

VI. OVERALL EVALUATION BY VIEWERS

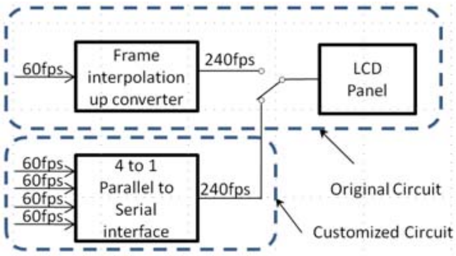

Fig. 13. Block diagram of the interface of the customized 240fps LCD Monitor

We viewed the 240fps motion images captured by our 240fps 2D or 3D camera of scenes such as soccer and tennis games, aerial shooting, golf, a boat and a pan shot of a flower as shown in Figure 14. The soccer, tennis and aerial scenes were captured by the 240fps 3D camera and the others by the 240fps 2D camera.

Fig. 14. Sample scenes shot by the 240fps 3D/2D Camera

The 240fps 2D contents were shown using a 240fps 2D display, a 240fps 4Kx2K projector, or a 240fps HD LCD monitor, and the 240fps 3D contents were shown using a 240fps 3D display system, a twin stack of the 240fps 4Kx2K projectors with circular polarized filters, or a 240fps HD LCD monitor with line-by-line circular polarized filters. The viewers were 10 adults with normal visual acuity. The viewing distance was 3H and the ambient condition was a dark room. The observers used circular polarized glasses when they viewed the 3D contents. These 2D or 3D contents were compared with simulated 60fps (1/60sec-shuttered) contents by repeating 4 times either the average of 4 frames or every fourth frame.

All observers recognized the difference between 240fps and 60fps and they reported that 240fps was clearly desirable for images, 2D or 3D, compared to the conventional 60fps [8].

VII. DISCUSSION

High frame rate motion images tend to bring good quality of depth perception because these images have richer information including depth cues with more precise edges than conventional images. High frame rate is also expected to be beneficial for increasing the spatial resolution for the following reason. The stimuli on the retina are accumulated over a duration of about 20msec. This phenomenon is known as Bloch’s Law [9] as described by the following.

$$I \cdot T = \mathrm{k} \qquad (1)$$

where I is the intensity of the stimulus, T is the displaying time and k is a constant. This law means that images of about 4 frames in 240fps are accumulated on the retina. Therefore, the spatial resolution of 4k60fps motion images can be closely matched by 2k240fps and that of 8k60fps images by 4k240fps.

VIII. CONCLUSION

Prototypes of a 240fps HD resolution 3D camera, a 240fps 4Kx2K projector and a 240fps LCD display were developed. These high frame rate systems produce natural and extremely high quality of motion image.

ACKNOWLEDGMENT

The author thanks Sony colleagues for fruitful discussion.

REFERENCES

[1] Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu and S. Yoshimura, “Improvement of Motion Image Quality by High Frame Rate”, *SID Symposium Digest*, 37, pp.14-17, 2006

[2] Y. Kuroki, T. Nishi, S. Kobayashi, H. Oyaizu, S. Yoshimura, “A psychophysical study of improvements in motion-image quality by using high frame rate”, *Journal of the SID*, Vol. 15, No. 1, pp. 61-68, 2007

[3] T. Kurita, “A Consideration on Motion-Image-Quality Improvement of LCD-TVs”, *SID Symposium Digest*, 57.2, 2009

[4] M. Sugawara, K. Omura, M. Emoto, Y. Nojiri, “Temporal Sampling Parameters and Motion Portrayal of Television”, *SID Symposium Digest*, P.30, 2009

[5] Sony press release 2009. 10. 1, http://www.sony.net/SonyInfo/News/Press/200910/09-117E/index.html

[6] Ian P. Howard, Brian J. Rogers, *Binocular Vision and Stereopsis*, Oxford Psychology Series No. 29, Oxford University Press, 1995

[7] T. Hayashi, Y. Kuroki, “Technology Study for a 4K x 2K High-Frame-Rate (240 fps) Projector”, *SID Symposium*, 8.4L, 2009

[8] Y. Kuroki, “Improvement of Motion Image Quality by Using High Frame Rate from Shooting to Displaying”, *IDW DES*, 1-4, 2009

[9] J. A. J. Roufs, “Dynamic properties of vision. I. Experimental relationships between flicker and flash thresholds”, *Vision Research*, 12, pp. 261-278, 1972