ISSN 1345-3041

Mar.2013 / Vol.141

MITSUBISHI ELECTRIC

ADVANCE

Latest Video Technologies for Various Applications

MITSUBISHI ELECTRIC
Changes for the Better

Mar. 2013 / Vol. 141
MITSUBISHI ELECTRIC ADVANCE
Latest Video Technologies for Various Applications

• **Editorial Inquiries**
Kazuhiro Oka
Corporate Total Productivity Management
& Environmental Programs
Fax +81-3-3218-2465

• **Product Inquiries**
Administration Department
Advanced Technology R&D Center
Fax +81-6-6497-7289

Mitsubishi Electric Advance is published on line quarterly (in March, June, September, and December) by Mitsubishi Electric Corporation.

CONTENTS

Technical Reports

Overview

Author: *Kenichi Tanaka**

Imaging technologies have been evolving with the advancement of television technologies. Digital television has become a device for managing video and data using various basic imaging technologies, such as for digital image processing, image compression, video recording, digital video transmission, IP networks, and wireless transmission. The latest imaging technologies for consumer products have contributed to 3D TV, laser TV and other high-quality digital TVs. Meanwhile, these basic imaging technologies are also utilized by a growing number of other applications including digital signage, visual information systems for public transportation, and surveillance systems. In addition, analysis technologies for digital images are required for products that need advanced sensing technologies such as in factory automation. This issue presents developed technologies used in these applications.

**General Manager, Advanced Technology R&D Center*

Evaluation Technology for Image Quality - From Liquid Crystal Display to Printer

Authors: *Sayaka Kobayashi** and *Nobuhiko Yamagishi**

1. Introduction

The image quality of display devices has always been classified and evaluated based on the brightness, tone and other physical quantities. However, viewers perceive such factors in an integrated manner, and thus, a subjective evaluation needs to be considered in combination with the physical quantities for the development of viewer-friendly display images. Mitsubishi Electric has been using an evaluation technique that combines the physical quantities and subjective evaluation to improve the image quality of not only 2D TVs but also 3D TVs and printed photographs.

2. Evaluation of Perception of Brightness with Respect to LCD TV

A display with a higher brightness generally provides better visibility and gives the impression of higher image quality. Therefore, as the performance of LCD panels and other devices improves, the value of the display luminance increases. As a result, although the image quality of LCD TVs has improved, viewers sometimes feel dazzled.

The perception of brightness is significantly dependent on the viewer's age. When young people (20s) and elderly people (60s and 70s) are exposed to different-sized round patterns on a screen, the threshold of brightness for feeling dazzled was examined for each pattern size and at various ambient luminances. The results are shown in Fig. 1.

At a visual angle of 25° on the display pattern, no difference was observed among ages, whereas at a visual angle of 1°, the young people felt dazzled at a lower luminance. Other studies have reported that the TV viewing distance of children is 0.8 times that of adults. Based on these results, a brightness control technology (*Katei Gashitsu* mode) has been developed and introduced in the "REAL MX60 series" LCD TVs released in 2006 and later models. In the *Katei Gashitsu* mode, dazzling is suppressed by automatically controlling the luminance according to the ambient brightness and the viewer's age (elderly person, youth or child).

Another consideration is the increasing need to reduce power consumption for global environmental conservation. However, if the luminance is simply lowered to reduce the power consumed by LCD TVs, of which a large portion is attributed to the LCD backlight, viewers feel that the screen is too dark. Accordingly, we have evaluated: (1) the preferable luminance level and (2) the lowest acceptable luminance by using TV sets with adjustable luminance levels.

As shown in Fig. 2, the perception of brightness was evaluated by means of a comparison method using multiple TV sets. Based on the experimental results, an energy saving technology (Fig. 3) has been introduced in the products, in which the backlight is controlled by the average level of the video input signal within the preferable luminance range that is not perceived as being too dark.

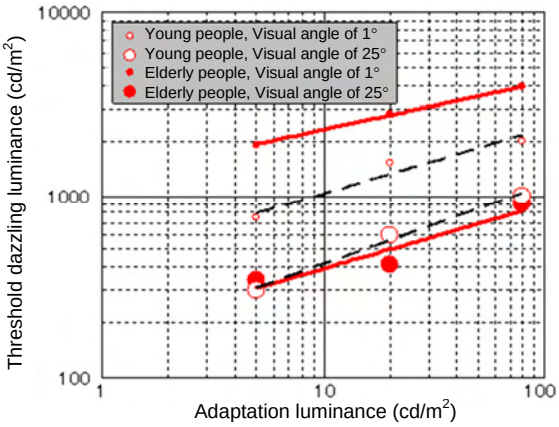


Fig. 1 Evaluation of threshold dazzling luminance

Fig. 2 Evaluation of perceived brightness by the comparison method

3. Evaluation of Image Quality of 3D TVs

For evaluating the image quality of 3D TVs, a new evaluation index – the stereoscopic effect – has been added to the conventional evaluation parameters. General 3D TVs use the stereopsis provided by the binocular disparity derived from different images for the

*Advanced Technology R&D Center

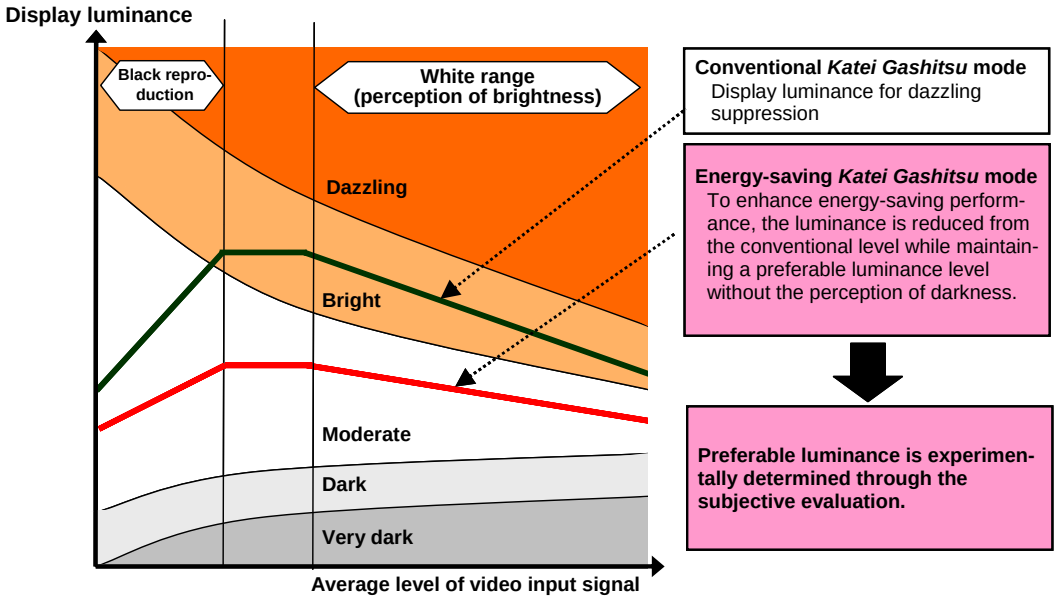


Fig. 3 Display brightness control based on the perception of brightness

left and right eyes; thus, the larger the binocular disparity, the greater the perceived image extrusion or retraction.

3.1 Evaluation of eye-friendliness of stereoscopic vision

An excessive amount of disparity does not provide stereoscopic vision. Therefore, we have evaluated the relationship between the amount of disparity and the eye-friendliness of the stereoscopic vision. A condition with an appropriate amount of disparity to achieve stereoscopic vision is referred to as "binocular fusion," and a condition with an excessive amount of disparity to cause unclear vision as "binocular separation," and we measured the threshold amounts of disparity for the binocular fusion and separation by using pictures of cups that offer either a sense of extrusion or retraction (Fig. 4).

Evaluation of the threshold disparity for binocular fusion was started from a state of binocular separation with a large disparity. The subject reduced the amount of disparity to find a disparity angle at which stereoscopic vision was achieved. For binocular separation, evaluation was started from a state of obtained stereoscopic vision, and the amount of disparity was increased to find a disparity angle at which the stereoscopic vision was lost. The results for the extrusion case are shown in Figs. 5 and 6, which indicate a difference between the threshold disparities for fusion and separation, and it is also noted that once the binocular fusion is achieved, it tends to be difficult to separate. In addition, Figs. 5 and 6 show that the binocular fusion and separation are widely distributed, indicating large individual differences in obtaining stereoscopic vision. The results for the retraction case also showed a similar tendency.

We have also evaluated the time required to perceive a fused stereoscopic image. A 3D image, as shown in Fig. 7, was created with an extruded part (A) and a retracted part (B). The subject alternately viewed the extracted and retracted parts, and the time required to obtain the fused image was measured. The results shown in Fig. 8 indicate that it takes a longer time for subjects 41 years or older to perceive stereoscopic vision. Based on these results, when watching 3D videos at home, it is desirable to adjust the stereoscopic effect according to individual differences. Thus, we have introduced the disparity adjustment function to MDR1 series LCD TVs released in 2010 and RDT233WX-3D LCD Display Monitors released in 2011.

3.2 Luminance measurement with 3D glasses

When the viewing brightness is measured for a 3D TV that requires a pair of 3D glasses, the 3D glasses need to be attached to a luminance meter and the left

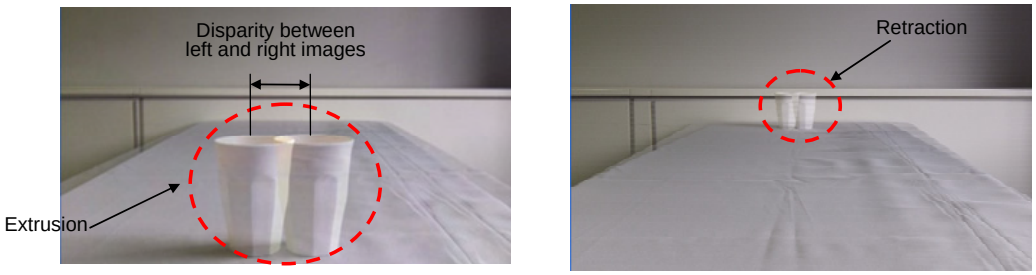


Fig. 4 Pictures used for the evaluation of binocular fusion/separation thresholds

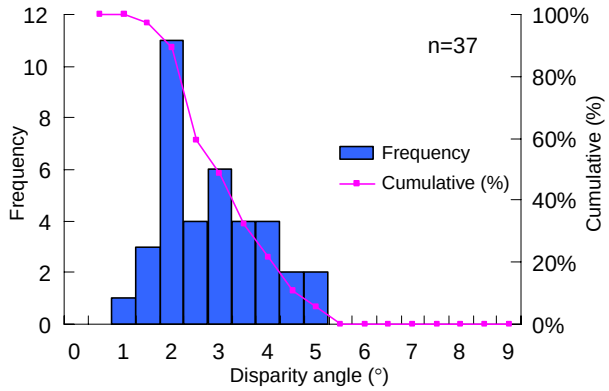


Fig. 5 Distribution of threshold disparity angles for binocular fusion and cumulative frequency percent (Extrusion)

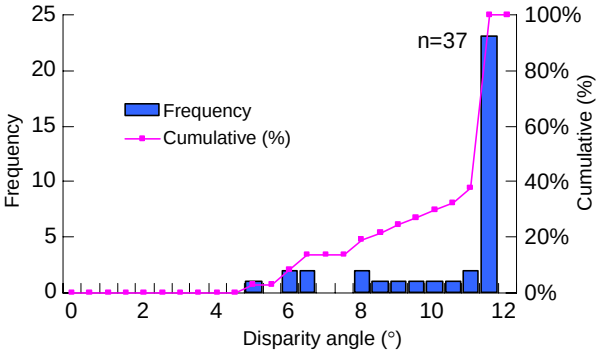


Fig. 6 Distribution of threshold disparity angles for binocular separation and cumulative frequency percent (Extrusion)

and right images need to be separated. However, the lenses of most 3D glasses are smaller than the objective lens of the luminance meter, and thus the following problems arise.

Unseparated light containing both left and right images enters through the outside of the 3D glasses lenses. If the unseparated light is shut off to prevent its entrance, the amount of light entering the luminance meter is not sufficient for accurate measurement.

To solve this problem, we have developed a 3D glasses adapter as shown in Fig. 9, which can be attached to the objective lens of the luminance meter. The 3D glasses adapter has a small inlet opening of 20 mm in diameter, making it possible to measure only the light entering through the 3D glasses lenses.

However, since the inlet opening is smaller than the objective lens of the luminance meter, the amount of light entering into the meter decreases. In addition, when the distance to the measuring object is changed, the focus angle and the amount of incident light also changes accordingly. Consequently, we measured the luminance with and without the 3D glasses adapter (the 3D glasses lenses were not attached) and derived the ratio between them (correction coefficient) for various luminance levels of measured objects as well as at various measuring distances. The resulting coefficients show no dependence on the luminance of the measured objects, but vary according to the distance. Therefore, the luminance without the 3D glasses adapter can be given by the measurement with the adapter multiplied by an appropriate correction coefficient determined beforehand as a function of the measuring distance. This has led to establishing an environment for performing accurate luminance measurements through the 3D glasses lenses.

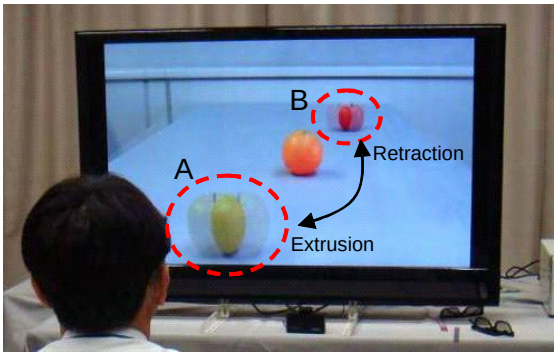


Fig. 7 Evaluation of time required to perceive 3D images by moving the viewpoint

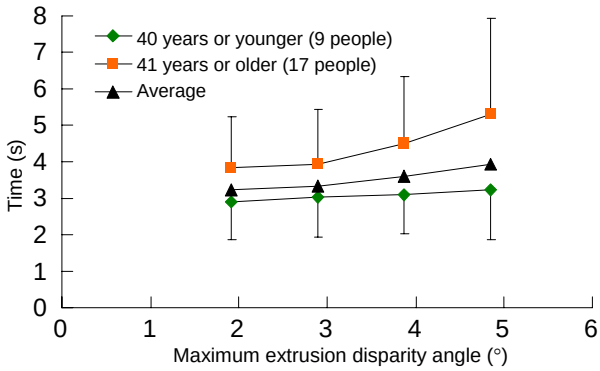


Fig. 8 Time required for binocular fusion by age

Fig. 9 3D glasses adapter

4. Image Quality Evaluation of Digital Printing Kiosk

Digital printing kiosks provide a photo printing service from the user's digital camera data. The data may need image quality correction, and preferable correction is added to photos such as low-contrast photos taken in a dark place or photos with too much backlighting.

4.1 Subjective evaluation by paired comparison method

For image quality correction, the brightness, contrast and other parameters may be calculated. The amount of correction of each parameter changes the brightness, vividness of color, and hence the overall impression of the photo.

Preferable amounts of correction were evaluated by means of the paired comparison method, using various sample data that gives different impressions depending on the correction amounts (Fig. 10). For each sample data, photos with different correction amounts were printed, and all combinations of paired photos were randomly presented to the subjects (total of 20 people, male and female, aged 20s to 40s). Then, they entered their preference for each pair on a 1–7 grade scale and the correction amount with the highest score was determined from the sample data. From all sample evaluation results, the correction amount giving the highest overall satisfaction was finally selected and adopted for the product.

4.2 Image quality correction and classification of contents of digital photos

Portrait photos are corrected in a different way from scenery photos, as desirable corrections include improving the face brightness and smoothing the skin appearance. In addition, some scenery photos (non-portrait photos) offer an unclear image because of excessive backlighting or insufficient overall brightness, and thus a case-by-case correction is required. In order to grasp the characteristics of the photo data to be printed, we focused on the luminance, one element for the evaluation of display image quality, and examined the luminance information of various portrait and non-portrait photos that need correction. As shown in Fig. 11, a photo with a person is corrected according to the degree of backlighting in the photo based on the luminance difference between the face area and the background. Non-portrait photos are classified into four types (three types with backlight and one type with direct light) according to the luminance distribution pattern of the image, and then the image quality is improved by correcting the backlight while retaining the overall contrast balance.

5. Conclusion

This paper described examples of image quality evaluation and the functions that were developed based on the evaluation results. Mitsubishi Electric offers a wide range of imaging products, and thus various evaluation technologies for the image quality are often required. We will continue to pursue a human-friendly and hence a higher image quality.

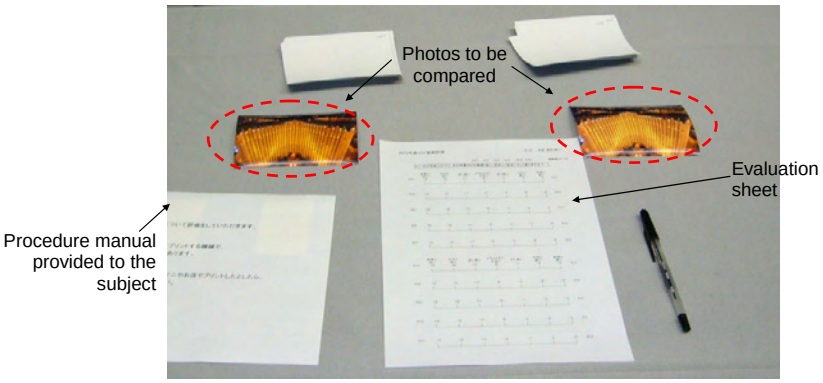


Fig. 10 Comparative evaluation of impression with various correction amounts

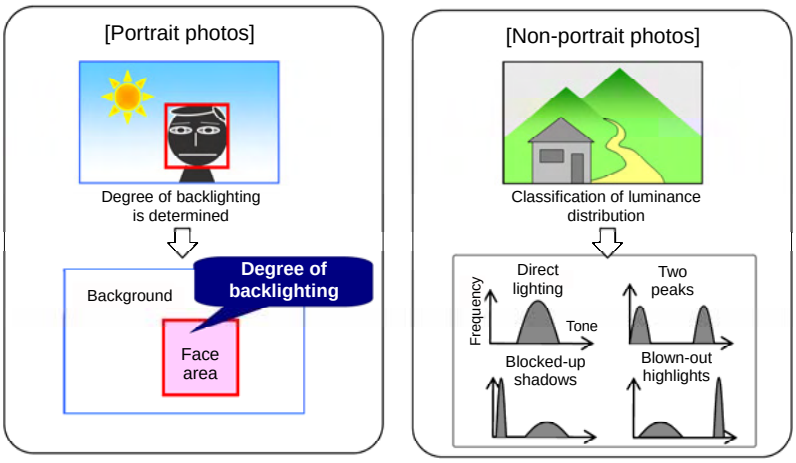


Fig. 11 Classification of digital photos by luminance characteristics

Home Network Technology for AV Appliance Interoperability

Authors: *Masaaki Shimada** and *Satoko Miki**

1. Introduction

We have developed home network technologies for interoperability among audio video (AV) appliances. The newly developed high-speed and stable delivery algorithm has achieved a high-quality image display for random access playback without fluctuations in the image display intervals. We have also developed a remote control technology that enables high-speed control of AV appliances linked through a comfortable user interface.

2. AV Appliance Interoperability

AV appliance interoperability refers to the ability of multiple appliances interconnected on a home network to share/control contents, and is roughly divided into "AV device interoperability" and "mobile device interoperability." AV device interoperability makes it possible to enjoy video contents that have been recorded in the living room, on the LCD TV in the bedroom, while mobile device interoperability allows AV devices to be controlled from a mobile device (Fig. 1).

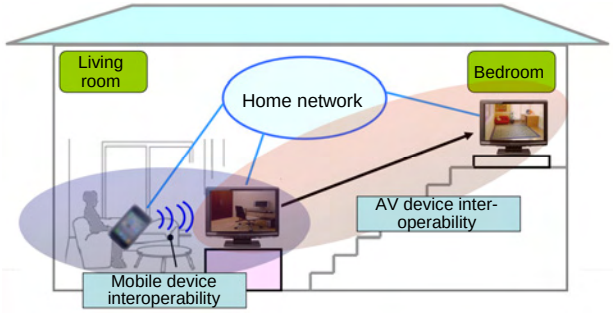


Fig. 1 Conceptual diagram of AV appliance interoperability

3. Technologies for AV Device Interoperability

3.1 Basic sequence of video delivery process

This time, we have built a video delivery middleware system for embedded devices based on the Digital Living Network Alliance (DLNA) guideline. Figure 2 shows the basic operation sequence of this system. Once the video delivery device is connected to the home network, interconnection between AV devices is automatically established. The user retrieves a list of stored video contents from the video display device, and selects the desired content. Once selected, the video display device uses the HTTP GET method to instruct the video delivery device to start playback of the selected content. Subsequently, the video delivery device sends a 'playback started' response as an HTTP response to the video display device, followed by the transmission of AV streams as HTTP response packets. The video display device decodes the received AV streams to play the AV content.

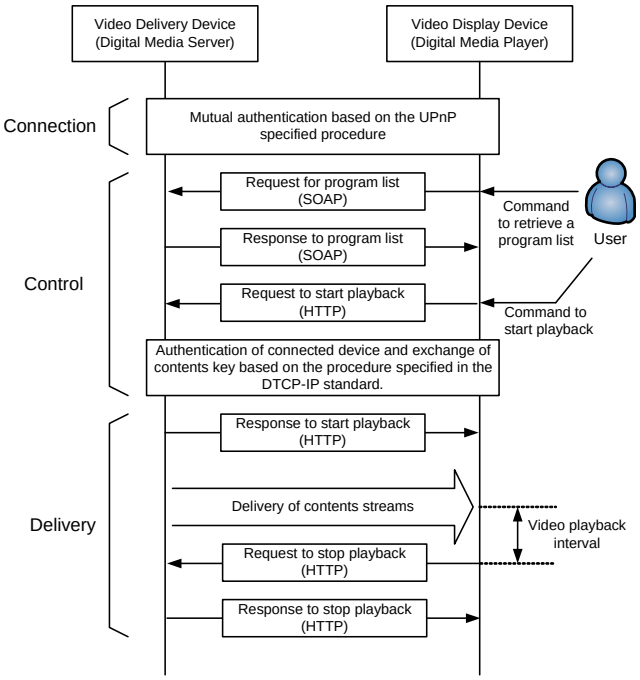


Fig. 2 Basic sequence for video delivery

3.2 Video delivery algorithm

3.2.1 Current issues

For video delivery over the network, it takes a long time to execute the sequence of playback start-up processes, from "delivery start instruction" to "video stream delivery." For the random access playback such as fast forwarding and fast backward, the playback start-up processes are repeatedly executed, and thus the number of displayable video frames per unit time is reduced, and smooth video display is disrupted. In addition, in some rare cases of delivery contents that have a special data rate, it takes a longer time than usual to start the video stream delivery, leading to fluctuations in the image display intervals and possible degradation of the display quality.

*Advanced Technology R&D Center

3.2.2 Automatic adjustment of transfer buffer size

This time, we have developed a dynamic video delivery control algorithm for the random access playback mode, where the status of the display device is estimated, and the transfer rate from the delivery buffer is controlled based on the estimated playback status and the data rate of the delivery content. With this control, video data is provided to the video display device at an appropriate transfer rate and timing for the playback status even in the random access playback, no fluctuations occur in the displayed image and a high-speed, stable video stream delivery can be achieved. Figure 3 shows the display intervals measured at the video display device before and after the application of this control method.

The data in Fig. 3 is the switching intervals between two consecutive images measured on each test sample in the random access playback mode. A larger display interval means a less smooth video playback, and a larger up/down variance or fluctuation of the display interval means a lower display quality.

As shown in Fig. 3, applying this algorithm to the random access playback reduces the video delivery time by about 25% in terms of the average display interval, as well as the standard deviation of the display intervals from 0.18 to 0.06. As such, it has been demonstrated that the application of this system reduces the fluctuation in the display intervals and achieves a high-quality image display.

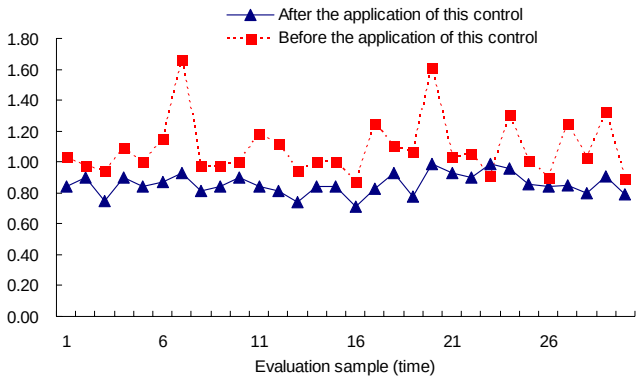


Fig. 3 Measurement results for display interval (Random Access Playback)

4. Technologies for Mobile Device Interoperability

4.1 Remote control technology

Mobile devices such as smart phones and tablet devices can be intuitively and easily operated using a touch panel. Recently, we have been receiving requests for the capability to control AV devices in a similar manner to mobile devices. However, a TV's processing capacity is lower than that of mobile devices, and thus it has been difficult to ensure the same operational performance as mobile devices. Accordingly, we have built a remote control system that requires no modification of TV hardware and provides the users with satisfactory TV control capability by using a mobile device.

Figure 4 illustrates the block diagram of the newly developed software architecture.

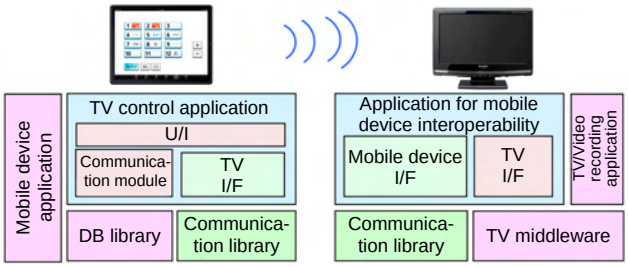


Fig. 4 Software architecture for mobile device interoperability

A new application program for the interoperability is installed in both the mobile device and TV set. The application is used to control intercommunication using the communication library. When the control from the mobile device is started, mutual interface specifications are exchanged to determine the functional compatibility. When a control command is executed, the mobile device uses the abovementioned interface specifications to transmit the command to the TV, which is then interpreted by the TV/Video recording application program. In this manner, channel selection, volume and image quality adjustment, and other functions are remotely controlled from the mobile device.

4.2 Speed-up of response process

In order to improve the response speed between the system modules, we have developed a method to reduce the waiting time between those modules. First, when a control command from the mobile device and a conventional TV software program are executed in parallel, the priority between the command and the TV program is determined to ensure the functionality of the TV. Second, given that various types of control commands from the mobile device require different processing times, an appropriate waiting time sufficient for each command type is determined, by not simply setting the same response waiting time to all operation commands. This modification has reduced the response time between modules and improved the speed of the response process. To evaluate the performance of this method, the response waiting time of the mobile device was examined (refer to Table 1).

As shown in Table 1, it has been confirmed that this method provides the mobile device with a response waiting time equivalent to that of the conventional infrared remote controller (0.02 s on average).

Table 1 Measurement results for response waiting time of mobile device

Action	Response waiting time (s)
Request for connection	0.02
Disconnection	0.02
Audio volume change	0.02
Channel change	0.02
Input change	0.02
Timer recording setup	0.23
Retrieval of a program list of timer recording	0.37

5. Conclusion

We have developed a middleware system for embedded devices to realize AV appliance interoperability. We intend to enhance the home network functionality as well as to develop new functions to improve the convenience of users.

References

(1) C. Morita, et al.: Network Technology for LCD-TVs, Mitsubishi Denki Giho Vol. 85, No. 3, 187–190 (2011) (in Japanese).
(2) S. Akatsu, et al.: IPTV Technology, Mitsubishi Denki Giho Vol. 82, No. 12, 755–758 (2008) (in Japanese).

Backlight Control Technology for Liquid Crystal Display TVs

Authors: *Hideki Yoshii** and *Masaaki Hanai***

1. Introduction

Liquid Crystal Display Televisions (LCD TVs) sometimes suffer from motion blurring caused by the hold-type display (a still image is displayed for one frame) and/or low contrast due to the backlight system. This paper describes the prevention of motion blurring by means of an impulse-type display, and the enhancement of contrast by controlling the backlight brightness based on the image information in multiple sections of the screen.

2. Motion Blurring of LCD TV

2.1 Hold-type display

A moving image on the TV is produced by continuously displaying images at the frame rate of 1/60 s. The hold-type display refers to the method of displaying each frame for 1/60 s.

Figure 1 shows a conceptual illustration of the hold-type display of a circle moving upward from the bottom. The response time of the LCD is assumed to be fast enough for the frame rate of 60 images per second. While the viewer's line of sight is continuously moving upward with time, a real object (not a TV) is always present in the viewer's line of sight, but the image in the hold-type display remains at the same display position for 1/60 s. As a result, the viewer recognizes the circle at a position backward from the line of sight, and perceives the object with a tail (motion blurring).

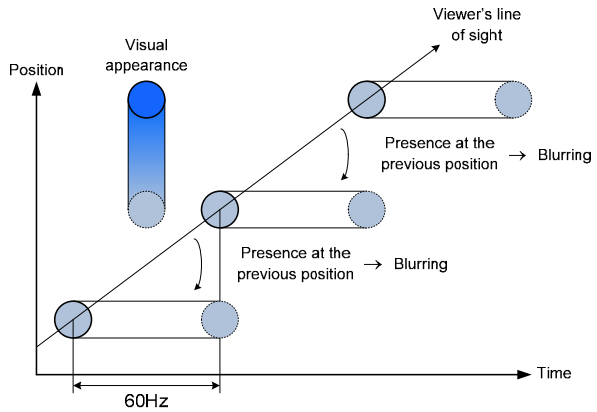


Fig. 1 Blurring in hold-type display

2.2 Double rate driving

Figure 2 illustrates the same image as shown in Fig. 1 in the double-rate driving mode. Interframe images are estimated from a sequence of images at the frame fate of 60 frames per second, and then displayed at the rate of 120 frames per second. Consequently, while the circle is still present backward from the viewer's line of sight, the amount of blurring is halved from the level at the rate of 60 frames per second. If interframe images are further estimated in the same manner, the display rate can be 240 frames per second, and the amount of blurring would be further halved.

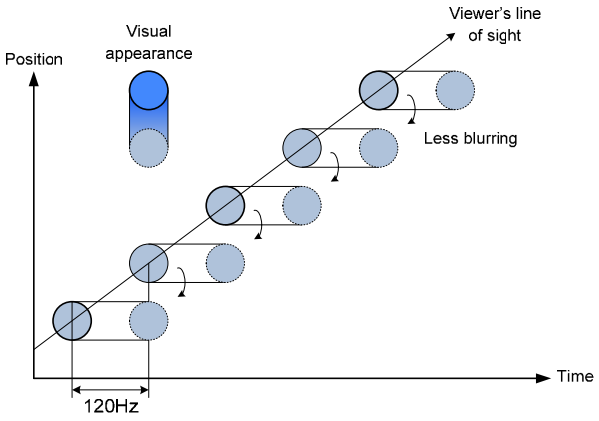


Fig. 2 Reduction in blurring by double-rate driving

As depicted in Fig. 2, while motion blurring can be reduced by the double-rate driving, it is true only when interframe images are correctly estimated from the images at the frame rate of 60 frames per second.

In reality, as shown in Fig. 3, images can be estimated fairly accurately in such relatively simple cases where the background of the moving circle is uniform, whereas if a still object is present in the background of the moving circle (a tree in Fig. 3), or the moving object has a complicated shape or pattern, the estimation may be incorrect, resulting in a deformed or blurred image.

2.3 Impulse-type display

Figure 4 depicts the concept of the impulse-type display. The LCD acts in the same manner as the hold-type display, but the display time is made shorter by reducing the on-time of the backlight. While the viewer's line of sight moves upward with time, the circle at a position backward from the line of sight is recog-

*Advanced Technology R&D Center **Kyoto Works

nized only during the on-time of the backlight and is invisible during the off-time, resulting in a reduced amount of blurring.

Unlike the double-rate driving, it is not necessary for the impulse-type display to estimate images to increase the frames, and even for a complicated image, an effect to reduce the blurring is always ensured in a stable manner without any incorrect estimation.

In the above explanation, the response speed of the LCD is assumed to be fast enough for the frame rate of 60 frames per second, and the motion blurring occurs in the hold-type display mode. In reality, however, it takes time for the LCD to respond, as illustrated in Fig. 5. Therefore, when the backlight on-time is shortened for the impulse-type display, as shown in Fig. 5, it is turned on at the longest time elapsed after the LCD has begun to respond. With this technique, the blurring due to the LCD's slow response speed (image lag) is also improved.

The whole screen of the LCD does not respond at once, but in a sequential manner from the top to the bottom (in some cases, from the bottom to the top). Therefore, as shown in Fig. 6, the backlight area needs to be divided into multiple sections, which are turned on in sequence according to the response of the LCD (sequential turn-on).

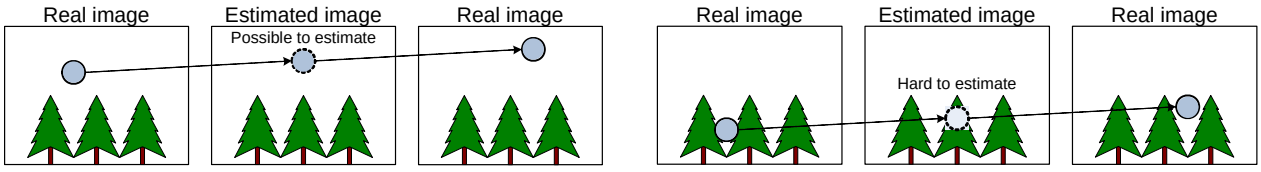
For such an impulse-type display, the backlight needs to have a fast on/off speed, and the backlight area needs to be divided into individually controllable multiple sections. Recently, LCD modules with a backlight that is fast enough for turning on/off and section controllable such as an LED backlight have become relatively easily available. The on-time of the backlight for impulse-type display is shortened and hence the brightness is lower than for the hold-type display. Nevertheless, sufficient brightness is becoming ensured even for the impulse-type display by means of transmittance improvement of the LCD and brightness enhancement of the backlight.

3. Contrast Ratio of LCD TV

An LCD TV displays an image by activating the backlight and controlling the transmittance of the LCD pixels in front of the backlight. This causes the following problems: the backlight remains on even when a black image is displayed, and light leakage from the backlight prevents the LCD from displaying a fully dark image even if the transmittance of the LCD is reduced. That is, the ratio of white to black (= contrast ratio) cannot become higher.

A conventional backlight control to overcome this problem is to increase the brightness when the image (scene) is bright and reduce it when dark. This method increases the brightness ratio of the bright scene to dark scene (dynamic contrast ratio). However, the contrast ratio within one image (static contrast ratio) depends on the performance of the liquid crystal panel and this backlight control method cannot increase the static contrast ratio.

Consequently, the backlight area is divided into multiple sections as shown in Fig. 7, and the brightness of each backlight section is controlled based on the image. This increases not only the dynamic but also the static contrast ratios. An example of such backlight module shown in Fig. 7 is equipped with LEDs on both sides, and divided into 2 horizontal and 8 vertical sec-



(a) Relatively monotonous and easy-to-estimate image (b) Complex and hard-to-estimate image

Fig. 3 Image estimation for double-rate driving

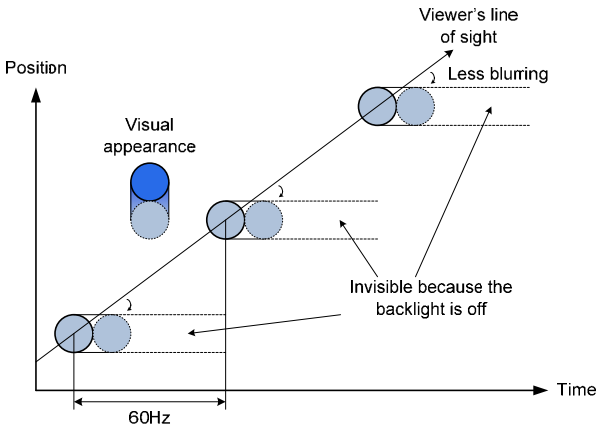


Fig. 4 Display in impulse-type display

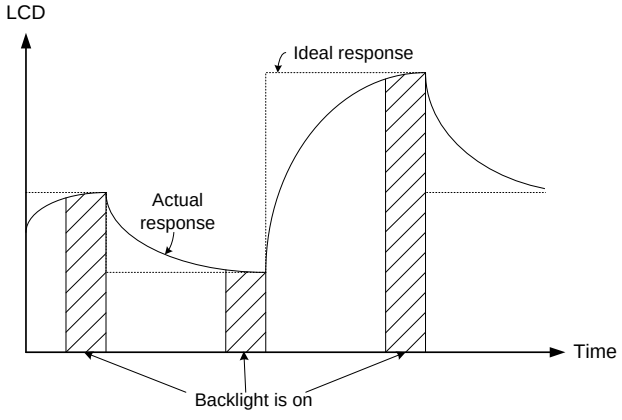


Fig. 5 Turn-on timing of backlight

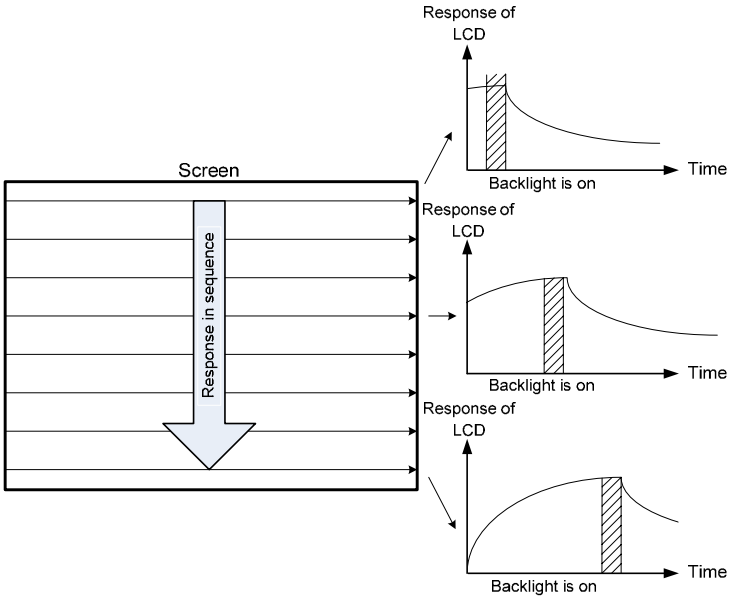
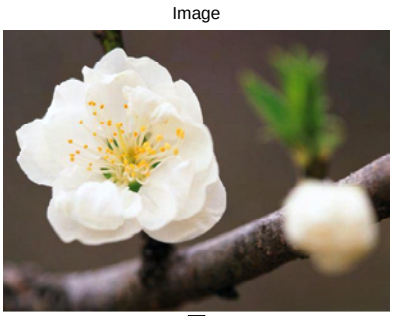


Fig. 6 Sequential turn-on of backlight



Backlight

50%	30%
80%	40%
90%	40%
100%	60%
90%	70%
60%	60%
40%	40%
20%	30%

Fig. 7 Backlight with local dimming control

tions. As a result of the vertical division into 8 sections, the above-described sequential turn-on control is implemented for the impulse-type display, which improves the motion blurring while the control is not too complicated because of the small number of sections, that is, a good balance is achieved between the cost and the image quality. If the cost is further reduced in the future, the number of sections could be increased for even more effective control.

4. Future Prospects

In response to the evolution of devices and changes in the market demands, we have pioneered the application of technologies, which so far could not be used, to improve the image quality. Based on technical trends and consumer behavior, we will continue to develop technologies in preparation for the timely commercialization of the products.

Reference

(1) Someya: Evaluation method and its standardization for motion blurring of liquid crystal display, ITE Journal, Vol. 60, No. 4, 510–515 (2006) (in Japanese).

Optical Pickup for Multilayer Blu-ray Disc

Authors: *Hironori Nakahara** and *Masayuki Omaki***

1. Introduction

Multilayer Blu-ray discs (BDs) have been commercially available since 2010 to meet the needs for video recording of long TV programs. The recording capacity of the multilayer BD was enhanced by increasing the number of layers.(1) However, in developing an optical pickup for the multilayer BD, suppression of layer crosstalk in the sensor system has been an issue.(2) (3) To overcome this problem, a one-beam push–pull method with a holographic optical element (HOE) has been developed.

2. Sensor System

Figure 1 shows an example optical layout of the optical pickup for BDs. A blue-violet laser with a wavelength of 405 nm and an objective lens with a numerical aperture of 0.85 are used. The laser beam from the semiconductor laser is reflected by the polarized beam splitter, and then converted into a parallel beam by the collimator lens. The collimator lens can be moved back and forth to compensate for spherical aberration due to the cover layer thickness. The linearly polarized laser beam is converted into a circularly polarized beam by the quarter-wave plate, and is then focused on the recording layer of the optical disc by the objective lens. The beam, which has been reflected by the recording layer and then polarized by the quarter-wave plate, passes through the polarized beam splitter and reaches the sensor system. The sensor system includes an HOE, a cylindrical lens, and a photodetector. The HOE has been newly developed to realize the one-beam push–pull method. It is arranged between the polarized beam splitter and the cylindrical lens. It should be noted that the cylindrical lens generates a focusing error signal using the astigmatic method.

Figure 2 illustrates a schematic diagram of the laser beam diffraction in which the beam is diffracted by the grooves on the disc. The grooves are periodically arranged like a diffraction grating, and thus the focused laser beam is not only reflected by the disc but is also diffracted by the grooves. While some of the diffracted laser beams pass through the objective lens and return toward the optical pickup side, when the phase of the diffracted beam is equal to that of the reflected beam, the optical intensity is increased in the area at which the diffracted beam arrives; on the contrary, if the diffracted and reflected beams are in reverse phase, the optical intensity is reduced in that area. This variation in the optical intensity is registered by the photodetector, generating a tracking error signal.

Figure 3 illustrates a schematic diagram of the holographic optical element (HOE). The HOE is divided into three areas, where Area 1 is a vertical diffraction grating, and Areas 2 and 3 are horizontal diffraction gratings. The sizes of these areas are almost equal to the spot size of the laser beam, and Area 1 fits the spot size of the ±1st order diffraction light generated by the grooves on the disc. The laser beam spot on the HOE moves in the horizontal direction in the figure as the objective lens moves in the radial direction for tracking purposes.

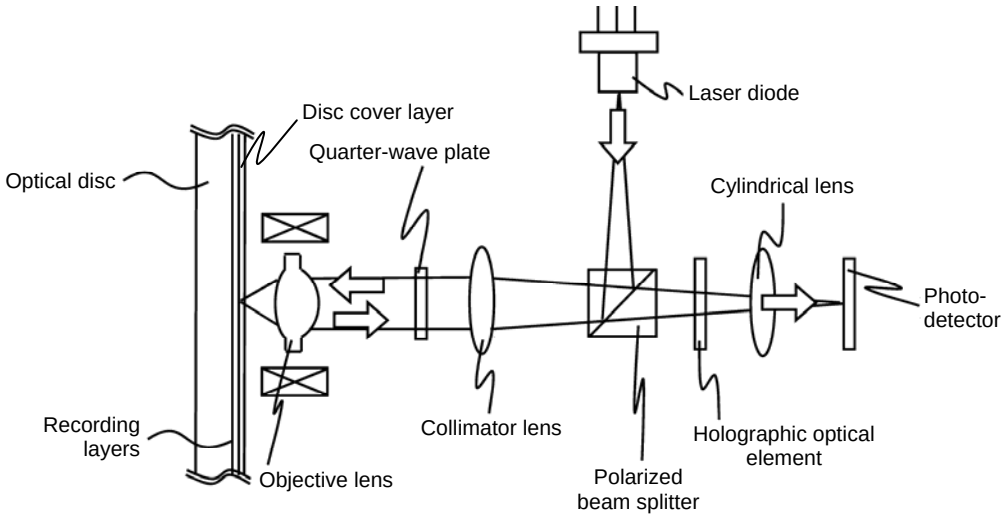


Fig. 1 Schematic layout of BD optical pickup

*Advanced Technology R&D Center **Kyoto Works

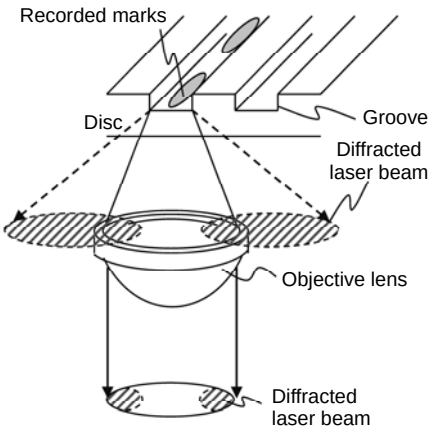


Fig. 2 Schematic diagram of disc grooves

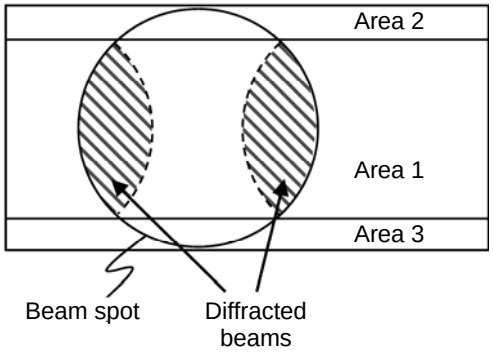


Fig. 3 Schematic diagram of holographic optical element

Figure 4 illustrates a schematic diagram of the one-beam push–pull method. The sensor system uses a general eight-segment ("A" through "H") photodetector. The group of four detectors "A" through "D" is referred to as the "main detector," and each pair of detectors "E" and "F", and "G" and "H" is referred to as a "sub-detector."

The laser beam that passes through the HOE reaches the main detector, while the beams vertically diffracted in Areas 2 and 3 reach the sub-detectors.

To generate a focusing error signal and a tracking error signal, the general astigmatic method and the one-beam push–pull method are used, respectively. The tracking error signal is generated by the signals output from the main detector and sub-detectors, and the operation $(A + D) - (B + C)$ for the main detector gives the push–pull signal. In the meantime, the beam spots on the sub-detectors move vertically in the figure when the objective lens moves in the radial direction, and the operation $(E - F) + (G - H)$ for the sub-detectors gives the lens error signal that indicates the shift amount of the objective lens. At this time, the laser beam on the main detector also moves in the vertical direction in Fig. 4, and thus an offset is generated in the push–pull signal. This offset can be canceled using the lens error signal by the operation $(A + D) - (B + C) - k[(E - F) + (G - H)]$ with an appropriately selected gain k.

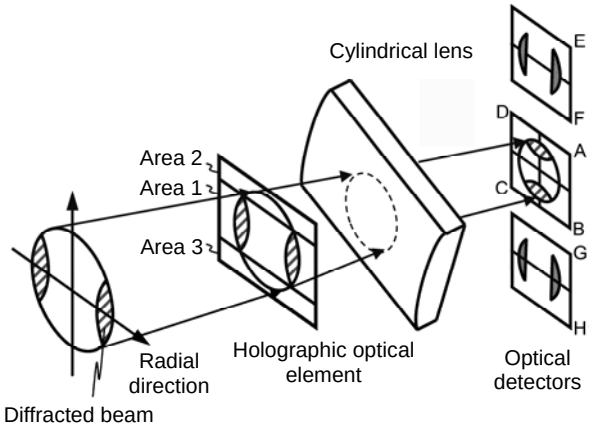


Fig. 4 Schematic diagram of one-beam push–pull method

3. Signal Simulation

For the dual-layer BD, the amount of stray light that is irradiated to the detectors was numerically simulated. The total amount of stray light, which is represented by "$A + B + C + D + k \times (E + F + G + H)$," was calculated as the percentage of the amplitude of the push–pull signal of the main detector. The calculated amount of stray light was 5% and 1% for layer 0 and layer 1, respectively. This simulation shows that the stray light target is 10%, and the calculated amount is below the target. The one-beam push–pull method makes it possible to freely set the laser beam power at the sub-detectors, and thus the influence of stray light can be reduced by setting the beam power at a higher level than the stray light.

4. Experimental Results

Figure 5 shows the measurement results of the tracking error signal obtained by the one-beam push–pull method. The measured amounts of stray light were 3% and 5% for layer 0 and layer 1, respectively. They are suppressed to a similar level to that of the simulation results. During the experiments, the tracking control also operated in a stable manner.

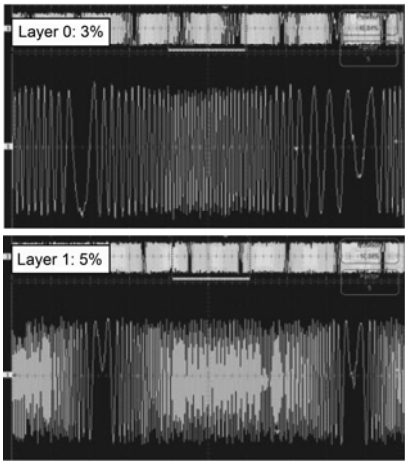


Fig. 5 Tracking error signal of one-beam push–pull method

5. Conclusion

It has been experimentally confirmed that the amount of stray light can be suppressed to 5% or less, and this technology enables the realization of an optical pickup for the multilayer Blu-ray discs.

References

(1) Blu-ray Disc Association: White Paper Blu-ray Disc Format General 2nd Edition (2010).
(2) T. K. Kim., et al.: Blu-ray Disc Pickup Head for Dual Layer, Jpn. J. Appl. Phys. 44, 3397–3401 (2005).
(3) A. V. D. Lee., et al.: Drive Considerations for Blu-ray Multi-Layer Discs, Jpn. J. Appl. Phys. 46, 3761–3764 (2007).

3-D Sensing Technology for Industrial Robots

Author: *Yukiyasu Domae**

We have developed a 3-D sensing technology suitable for industrial robots. The motion stereo method is capable of 3-D measurement using only one camera, and thus is a useful and cost-effective method; however, its application to industrial robots has been impractical because of the difficulty in synchronizing the timing between taking an image and capturing the position. We propose an improved motion stereo method that has overcome the problem by using images asynchronously captured during the motion. This method achieves stable 3-D measurement with an S/N ratio improved by 16 times and a noise level reduced by 94% compared to the conventional method.

1. Introduction

Stereo measurement generally uses two cameras. However, for cost efficiency and convenience of installation, there is a method called "motion stereo," which uses only one movable camera to achieve stereo measurement. Some of the techniques for this method include: generating a temporal image sequence to stably associate the images with each other based on the tracking of the temporal image sequence,[1] [2] and multiple-baseline stereo[3] realized by using a single camera.[4] For the motion stereo method, it is necessary to synchronize the timing for the camera to take an image and the timing for capturing the position of the automatic positioning stage or robot. If the stage or robot is halted when an image is taken by the camera, synchronization would be easy; however, a trade-off arises between the accuracy of the position and the duration of the halt. Consequently, we have established a problem setting with two viewpoints where the image-taking timing and position-capturing timing are synchronized, and multiple asynchronous viewpoints in which images are taken from the two viewpoints while moving between the synchronized viewpoints. This paper presents a new motion stereo method best suited for this problem setting.

2. Motion Stereo

Figure 1 shows a model of the motion stereo. WCS and CCS stand for World Coordinate System and Camera Coordinate System, respectively. Let $\mathbf{x}_f$ and $\mathbf{x}_l$ be the respective projections of a certain point $\mathbf{X}$ in the 3-D space on the two images taken from different positions, then the following holds:

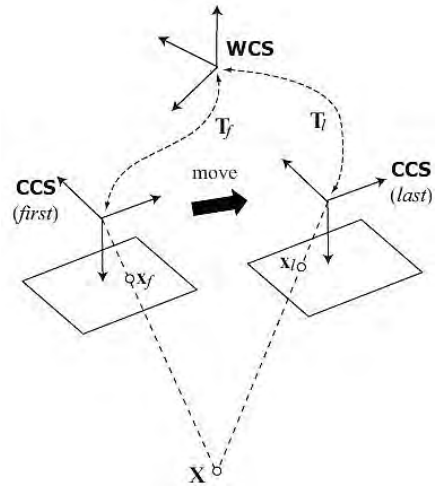


Fig. 1 Model of motion stereo

$$\begin{cases} s_f \mathbf{x}_f = \mathbf{A}\mathbf{P}\mathbf{T}_f \mathbf{X} \\ s_l \mathbf{x}_l = \mathbf{A}\mathbf{P}\mathbf{T}_l \mathbf{X} \end{cases}$$

where s_f and s_l are scalars, $\mathbf{A}$ is the camera intrinsic matrix, $\mathbf{P}$ is the projection matrix, and $\mathbf{T}_f$ and $\mathbf{T}_l$ are the external camera parameters representing the position and posture, respectively, of the viewpoints. In order to solve these equations for $\mathbf{X}$, the values of $\mathbf{A}$, $\mathbf{T}_f$ and $\mathbf{T}_l$ need to be known. In the case of motion stereo, $\mathbf{A}$ is always the same and is easily calculated; and the coordinates of the robot or translation stage can be used as $\mathbf{T}_f$ and $\mathbf{T}_l$. By specifying the WCS to be, for example, the base coordinate system of the robot, accurate $\mathbf{T}_f$ and $\mathbf{T}_l$ are determined if the camera's image-taking timing coincides with the coordinates-capturing timing.

3. Model of Hand-Eye Motion

The shortest path between two viewpoints for efficient travel is a straight line. The translation stage and industrial robot provide linear translational motion, and thus we use a linear motion between the two viewpoints. In addition, since the stereo measurement yields the most reliable results when the cameras are aligned parallel and on the same level, the direction of the translational motion is arranged parallel to the image plane. Images are assumed to be taken at a constant interval, and since an industrial robot is unable to travel between two points at a constant speed, and instead

**Advanced Technology R&D Center*

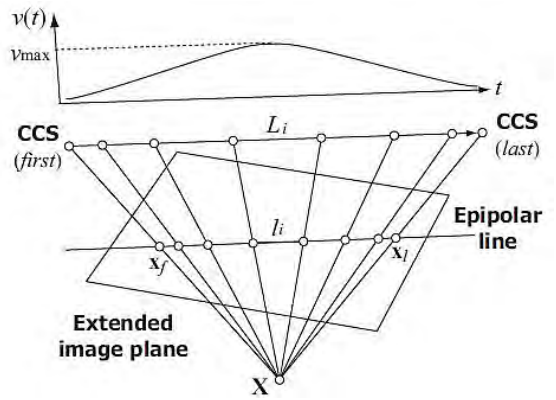


Fig. 2 Model of hand-eye motion

travels at a smoothly varying speed, the model illustrated in Fig. 2 can be postulated.

The projection x_f of point X on the image taken from the first viewpoint, and x_l from the last viewpoint satisfy the same epipolar constraint on the same plane (here, it is called an extended image plane). Let the projection at the first viewpoint $x_f = x_1$, the projection on the i-th image be x_i, and the projection at the last viewpoint $x_l = x_k$, $i = 1, 2, \ldots, k$, then the following equation holds:

$$\mathbf{x}_1^T(\mathbf{t} \times (\mathbf{x}_i + \mathbf{t})) = 0$$

4. Tracking Stability

Successively searching for correspondence between multiple images is, if the sampling rate is high, similar to point tracking. Therefore, based on the deviations of the point tracking trajectory and epipolar constraint of the model, defective correspondences are removed. The index for statistically evaluating the degree of the deviations, called "tracking stability," is defined as:

$$T_i = T_{li} T_{\theta i}$$

where T_{li} and $T_{\theta i}$ are the parameters related to the positional and angular displacements, respectively. Please see Reference (5) for details. The tracking stability is 1 for ideal tracking, and as the tracking becomes unstable, it approaches 0. If the variance of either positional displacement or angular variance exceeds the threshold value, T_i becomes zero, and thus any point that finally yields $T_i = 0$ is removed as defective.

5. Experiment

In the experiment, a camera was attached to the tip of the robot arm, a cylinder was measured by three types of motion stereo methods, and the performance of each measurement was examined. Figure 3 shows the experimental system, where the robot arm has six degrees of freedom with a repetitive positioning accuracy of 0.05 mm; the basic specifications of the camera are: VGA, 15 fps, lens focal length of 6 mm; and the lens distortion, intrinsic and external camera parameters are calibrated.

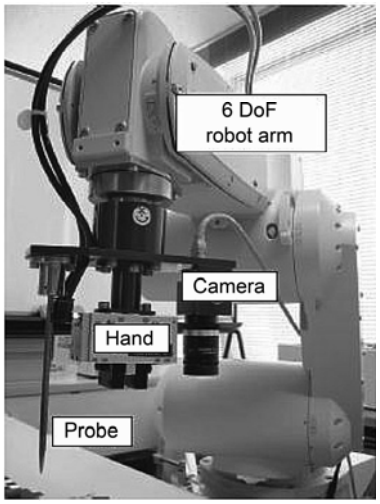


Fig. 3 Robot system for the experiment

The 3-D measurement results are shown in Figs. 4 and 5. It can be seen that the traditional stereo method generates much noise, whereas no noise is observed with the proposed method. It was indicated that the stereo method does not work well even for an area with textures. The proposed method is given an overwhelmingly high evaluation in terms of the S/N ratio: noises were almost completely removed; although the signal level was reduced by 7.7%, 94% of noise was removed.

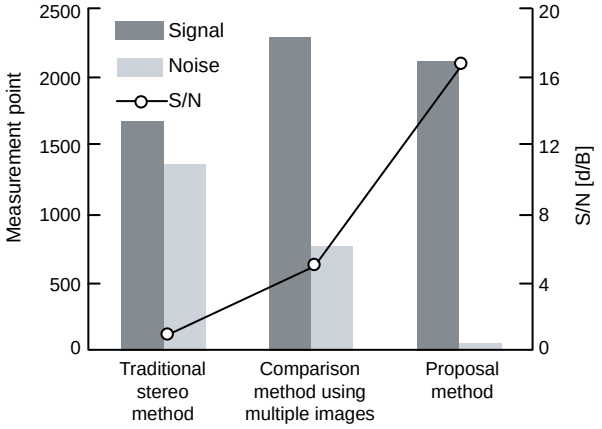


Fig. 4 Measurement results of each method

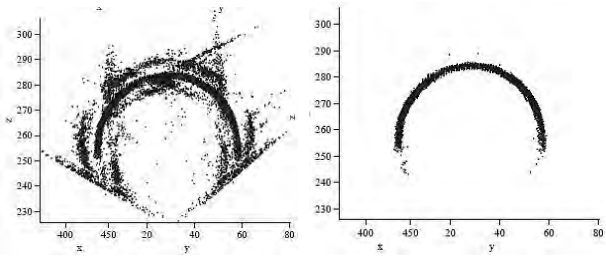


Fig. 5 Measurement results for a cylindrical object

In addition, the average error of the points that were removed as defective was 1.2 mm, indicating that the proposed method achieves higher accuracy than the conventional method.

6. Conclusion

We have confirmed that the motion stereo method using the travel between two viewpoints provides improved measurement performance, with a high S/N ratio and without increasing the number of stopping times.

References

(1) M. Yamamoto: Determining three-dimensional structure from image sequences given by horizontal and vertical moving camera, IEICE Transactions, H69-D, 11, 12–20 (1986) (in Japanese).
(2) R. C. Bolles, H. H. Baker, D.H. Marimout: Epipolar-plane image analysis an approach to determining structure from motion, Int. J. Computer Vision 1(1), 7–55 (1987).
(3) M. Okutomi, T. Kanade: A multiple-baseline stereo, Proc. IEEE CVPR, 63–69 (1991).
(4) T. Sato, M. Kanatsugu, N. Yokoya, H. Takemura: Dense 3-D reconstruction of an outdoor scene by hundreds-baseline stereo using a hand-held video camera, Int. J. Computer Vision, 47(1–3), 119–129 (2002).
(5) Y. Domae, H. Okuda, H. Takauji, S. Kaneko and K. Sumi: Motion stereo including tracking stability analysis for eye-in-hand systems, J. Japan Society for Precision Engineering, 77(1), 90–96 (2011) (in Japanese).

Technologies for Next-Generation Video Coding Standard

Authors: *Kazuo Sugimoto** and *Shun-ichi Sekiguchi***

High Efficiency Video Coding (HEVC), the next-generation video coding standard, is expected to prevail as the high compression technology for future video applications. This paper describes the HEVC standardization trends, and the technologies developed by Mitsubishi Electric to facilitate standardization. Future prospects for the standardization of HEVC are also discussed.

1. Trends of HEVC Standardization

The target of HEVC standardization is to substantially surpass the compression efficiency of the existing MPEG-4 AVC/H.264 international standard,[1] which currently provides the highest compression performance. The expected complexity of HEVC must be an acceptable level for the hardware technology when supported products are put on the market. In the standardization activities, proposed technologies recognized as being effective in terms of both performance and complexity are adopted in the Working Draft (WD) specifications. The following is an outline of WD-4, the specifications as of July 2011.[2]

1.1 Extension of coding block and hierarchical structure

WD-4 specifies that the screen is divided into square blocks called Largest Coding Units (LCUs), and the coding is performed for each LCU. The LCUs are then divided into sub-blocks called Coding Units (CUs), which are arranged in a hierarchical quadtree structure based on the image content and motion. CUs are further divided into Prediction Units (PUs), which can vary in size and shape. For each PU, intra frame prediction or inter frame motion-compensated prediction is performed. The prediction error signal is then orthogonally transformed and quantized for each Transform Unit (TU), which is a hierarchical quadtree subdivision of CU according to the distribution of prediction error power. Figure 1 depicts an example of the relationship between the CU, PU and TU. This technology has already demonstrated its ability to achieve efficient bit allocation based on the video signals with locally varying characteristics, and to significantly contribute to enhanced compression performance for high-definition video signals.[3]

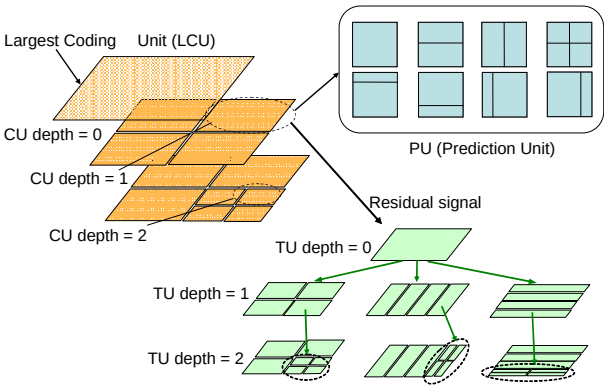


Fig. 1 Hierarchical block structure in WD-4

1.2 Intra prediction

Intra angular prediction is adopted in which prediction samples are generated by using spatially adjacent pixel values as the prediction reference value and switching multidirectional spatial prediction modes. A local change in the image signal can be properly handled by providing the maximum 33 angular prediction modes for each PU size (4 × 4–64 × 64), a direct-current (DC) prediction mode that uses the average value of prediction reference samples as the prediction image, and a planar prediction mode that is effective for areas with a flat gradation.

1.3 Motion-compensated prediction

For the motion compensation prediction, efficient coding of the motion parameters is adopted in addition to the expansion and adaptation of the prediction block size. Compared to the conventional video coding technology, a much higher degree of freedom is provided for the allocation of areas where the motion parameters need to be accurately described or can be omitted. Figure 2 shows the prediction mode that is available for the motion compensation prediction. In addition to the prediction mode where a CU is divided into equally sized rectangular PUs for the motion prediction, a mode of horizontal or vertical asymmetric division at the ratio of 1:3 is also available. This allows the shape of the prediction block to be changed in various ways according to the contents of the image, and thus offers a more accurate motion prediction.

*Mitsubishi Electric Research Laboratories **Information Technology R&D Center

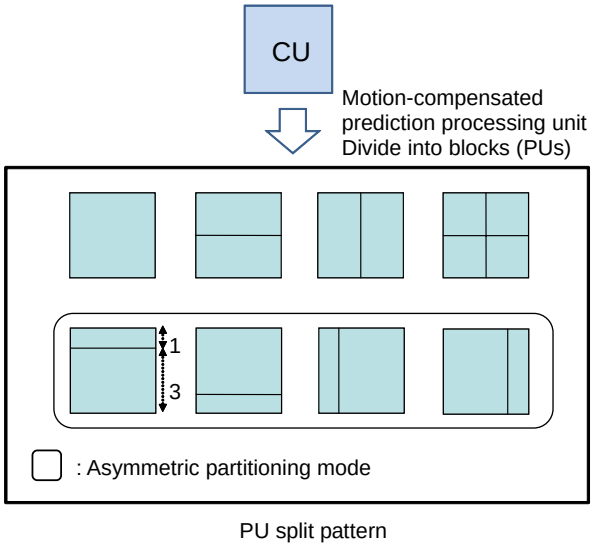


Fig. 2 PU split modes for motion compensation

1.4 Loop filter

For improved rate-distortion performance and subjective image quality, WD-4 has adopted the loop filter configuration by cascading: (a) the deblocking filter that reduces blocking noise generated at orthogonally transformed block boundaries; (b) the sample adaptive offset that adds an adaptive pixel-level offset based on the local signal characteristics to compensate for a difference between the original signal and the (local) decoded signal containing a coding distortion; and (c) the adaptive Wiener filter that compensates for coding distortion by successively designing the Wiener filter, which minimizes the average mean square error between the original signal and the output of the sample adaptive offset, and incorporating the loop filter mechanism for directly coding and transmitting the filter coefficients.

2. Technologies Developed by Mitsubishi Electric for Standardization of HEVC

Since the AVC/H.264 standard was established in 2003, Mitsubishi Electric has been developing high-efficiency encoding algorithms. The following sections describe some of the key technologies proposed for the standardization of HEVC.

2.1 Smoothing of intra predicted samples

In the intra prediction, prediction samples are generated by repeating reference samples along the prediction direction, and thus if the luminance changes in the prediction direction, an accurate prediction cannot be obtained. Accordingly, Mitsubishi Electric has developed a technique to improve the prediction performance in the DC prediction mode by applying the vertical adaptive smoothing filter at the boundary between the prediction and reference samples, as well as another technique to improve the prediction performance in the vertical and horizontal directions by adding to the prediction signal a value proportional to the pixel value change in the reference samples along the prediction direction, and thus increasing the inter pixel correlation at the block boundary (Fig. 3).[4] These techniques together improve the encoding efficiency by about 1%, and have been both adopted in the WD.

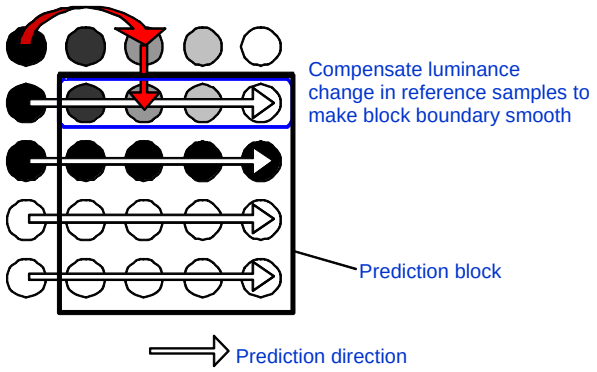


Fig. 3 Intra prediction smoothing for intra horizontal prediction

2.2 Prediction of motion in cooperation with decoder

There is a technology known as the direct motion vector, where instead of directly encoding the motion vector that represents the inter frame motion, the direct encoding is substituted by the motion vector information held by the decoder. Mitsubishi Electric has developed a technique for the encoder and decoder to work together for adaptively selecting the most reliable candidate as the direct motion vector by using multiple motion vectors of the blocks close to the prediction block. It has been confirmed that this technology reduces the amount of coded bits by 1–2%.[5]

2.3 High-efficiency fixed length coding

Mitsubishi Electric has developed an entropy coding scheme, Probability Interval Partitioning Entropy Codes/Variable-to-Fixed Length (PIPE/V2F). PIPE/V2F is a fixed length code, yet it achieves a lossless high compression performance.[6]

This technology has achieved a high coding efficiency by designing an optimum fixed length coder for each of the multiple probability intervals, then adaptively switching the fixed length coders based on the occurrence probability of input symbols. It has been confirmed that the coding efficiency is improved by 3–4% from the level of the Context Adaptive Variable Length Coding (CAVLC) adopted in the WD-4.

3. Future Prospects

Following the conventional coding architecture up to MPEG-4 AVC/H.264, the algorithm of WD-4 has demonstrated an improvement of the compression efficiency by more than 40% in terms of the

rate-distortion performance. Additional discussions are in progress for expanding the standard including: coding tools customized for certain screen contents such as the PC desktop screen, scalable coding for videos with different resolutions in a single bitstream, and multiview video coding for 3D free-view videos.

HEVC is a key technology for the distribution of next-generation high-quality video contents. We will continue to contribute to its standardization to be completed in 2013, followed by the expansion of the standard.

References

(1) ISO/IEC 14496-10: Information technology, Coding of audio-visual objects – Part 10: Advanced video coding (2003).
(2) B. Bross, et al.: "WD4: Working Draft 4 of High Efficiency Video Coding," JCTVC-F803 (2011).
(3) S. Sekiguchi, et al.: "A novel video coding scheme for Super Hi-vision," PCS (2010).
(4) A. Minesawa, et al.: "Vertical and horizontal intra predictions considering luminance change in the direction of prediction," FIT (2011) (in Japanese).
(5) Y. Itani, et al.: "A Method of Implicit Direct Vector Derivation with Decoder-Side Motion Estimation," IEICE Transactions, Vol. J94-D, No. 12 (2011) (in Japanese).
(6) K. Sugimoto, et al.: "A novel high efficiency fixed length coding for video compression based on symbol probability estimation," VCIP (2011).

MITSUBISHI ELECTRIC CORPORATION