An HDTV-Compatible 3DTV Broadcasting System

Namho Hur, Gwangsoon Lee, Woongshik You,
Jinhwan Lee, and Chunghyun Ahn

We introduce a high-definition three-dimensional television (3DTV) broadcasting system that is fully compatible with the existing transmission system of high-definition television (HDTV). Specifically, we developed high-definition 3DTV broadcasting subsystems including a 3DTV camera, 3DTV video multiplexer and demultiplexer, 3DTV receiver, and 3DTV outdoor broadcast van. To verify the developed subsystems, we performed experimental services of 3DTV broadcasting during the 2002 FIFA World Cup Korea/Japan. According to our subjective evaluation test, 88% of 273 viewers rated the perceived depth of 3DTV as "Good," and 36% of the viewers preferred 3DTV to other digital broadcasting services.

Keywords: 3DTV, 3DTV broadcasting, 3DTV camera, 3DTV video multiplexer/demultiplexer, 3DTV receiver.

I. Introduction

With the rapid development of digital broadcasting technologies in recent years, more sensational and realistic broadcasting services are strongly required. Among them, three-dimensional television (3DTV) broadcasting is being considered as one of the promising next-generation multimedia services following high-definition television (HDTV). The inevitable requirements of realistic broadcasting services have lead to extensive research in the direction of 3DTV. We conducted research in high-definition 3DTV during the years 2000 to 2002 in order to keep pace with these irresistible research trends. With this 3DTV project, we developed a fundamental 3DTV broadcasting system. This newly developed 3DTV format can also be viewed on a regular two-dimensional HDTV, when using a 3DTV-compatible decoder and selecting only one of the stereoscopic views, thus enabling the transmission of 3DTV content to a much larger audience over both a terrestrial broadcasting frequency channel and a terrestrial/satellite broadband network.

In Europe, as a result of the RACE DISTIMA (DIgital STereoscopic IMaging & Applications) project [1], a real-time transmission experiment of stereoscopic video, via asynchronous transfer mode (ATM) at 10 Mbps, was carried out to link the laboratories of KPN Research in Leidschendam, The Netherlands, and the laboratories of Deutsche Telekom in Berlin, Germany, in 1994.

Furthermore, it is reported that standard-definition 3D MPEG-2 video transmission over a broadband network and broadcast channels was done in Canada in 2002 [2]. Similar to our system, they used the existing HDTV infrastructure with a multiplexer, demultiplexer, and slightly modified MPEG-2

Manuscript received Apr. 18, 2003; revised Feb. 2, 2004.

This work was supported by the Ministry of Information and Communication of Korea under the title of "A study on the 3D-image broadcast-repeat demonstration service."

Namho Hur (phone: +82 42 860 6568, email: namho@etri.re.kr), Gwangsoon Lee (email: gslee@etri.re.kr), Woongshik You (email: wsyou@etri.re.kr), Jinhwan Lee (email: jinhwan@etri.re.kr), and Chunghyun Ahn (email: hyun@etri.re.kr) are with Broadcasting System Research Group, Digital Broadcasting Research Division, ETRI, Daejeon, Korea.

decoder in order to deliver standard-definition 3D video to a very large audience. The major differences lie in the multiplexing scheme and the resolution of the images, i.e., the top-bottom multiplexing scheme and the standard-definition images.

In Japan, high-definition 3DTV live broadcasting was done during the Nagano Winter Games in February, 1998 by transmitting the left and right images separately to Tokyo through a communication satellite, N-star, at 45 Mbps [3].

But in our case, we performed high-definition 3DTV live-broadcasting during the 2002 FIFA World Cup Korea/Japan by sending only one single, horizontally-multiplexed stereoscopic image through a terrestrial network at 45M bps and a communication satellite, Koreasat-3, at 155 Mbps [4]. In some venues, the stereoscopic images were displayed on a 300-inch silver screen in order to provide the audience with a strong impression of both a sense of reality and a sensation of depth. However, we did not perform experimental services of the terrestrial digital television (ATSC) via a terrestrial digital broadcasting frequency channel for several reasons concerning commercial broadcasting rights, although the 3DTV broadcasting system is compatible with the HDTV system. We produced 3DTV content by field production using three stereoscopic cameras and a 3DTV outdoor broadcast (OB) van at the World Cup stadiums.

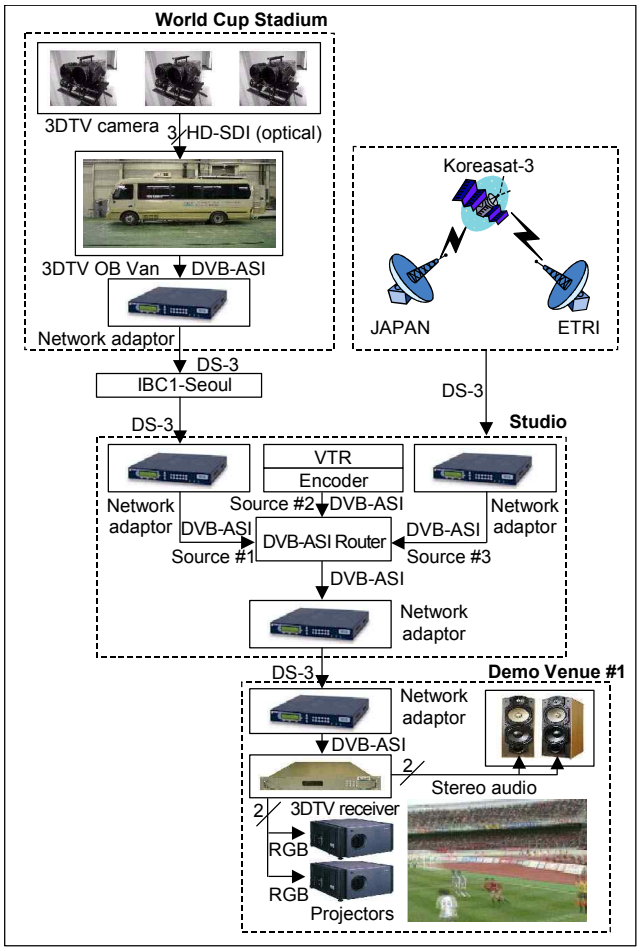


Fig. 1. Configuration of a 3DTV broadcasting system.

II. System and Technologies

1. Overview of the 3DTV Broadcasting System

Figure 1 shows the simple configuration of an experimental 3DTV broadcasting system covering from a World Cup stadium to a venue. This experimental system was used to verify the developed 3DTV broadcasting subsystems during the 2002 FIFA World Cup Korea/Japan. The communication satellite, Koreasat-3, was used for high data-rate (155 Mbps) transmission experiments between Korea and Japan. Five first-round World Cup soccer matches were directly captured with the use of three toed-in stereoscopic cameras and a 3DTV OB van. We placed one camera at the highest, unilateral camera position possible and two cameras at the pitch level on the reverse side of the stadium. A pair of left and right images captured by the 3DTV camera was horizontally multiplexed by using the developed 3DTV video multiplexer. To be specific, the multiplexer reduces the original sampling rate for the left and right images down to its half-rate in the horizontal direction and then rearranges the re-sampled left and right images such that they fit into a single standard video frame, i.e., in a *side-by-side* format. Hereafter, we call the reduction of the original sampling rate down to its half-rate as *decimation*, the opposite concept of interpolation. Accordingly, this horizontal multiplexing scheme enables us to deal with the stereoscopic images as a single HDTV image with the sacrifice of a slightly lower picture quality. The multiplexed images of the three stereoscopic cameras were sent to the 3DTV OB van, located at the TV compound, over single-mode (SM) optical fibers.

In the OB van, we chose the most comfortable stereoscopic image among the three multiplexed images by monitoring the perceived depth in real-time, and then inserted a computer generated text or graphics into the selected stereoscopic image. Next we encoded the final stereoscopic image and sent only one transport stream (TS) to the studio through IBC1-Seoul (International Broadcasting Center 1 in Seoul), as shown in Fig. 1. Note that from the studio there were 3 different 3DTV sources: source one from a World Cup stadium, source two from a VTR and an encoder, and source three from Japan via the Electronics and Telecommunications Research Institute (ETRI) in Daejeon. According to the program schedule, the selected 3DTV signal from the three resources was distributed from the studio to the venues over a terrestrial DS-3 network specified in ITU-T Rec. G.703 [5]. At the venues, the 3DTV

receiver decoded and demultiplexed the signal sequentially. The recovered stereoscopic images were displayed on a large silver screen with beam projectors. With the aim of advertising different digital broadcasting technologies, HDTV, data broadcasting, and satellite broadcasting technologies were also exhibited along with 3DTV in some venues. The number of participants at the venues was 570,668.

2. 3DTV Camera

Figure 2 shows the geometry of a toed-in stereoscopic camera consisting of a pair of lenses and charge-coupled devices (CCDs), where f is the focal length of the zoom lenses, b is the baseline between cameras, O is the origin of a three-dimensional object space, D_v is the convergence distance, and 2β is the convergence angle of the camera. Let (X_{cl}, Y_{cl}) and (X_{cr}, Y_{cr}) represent the location of projected points on the left and right imaging devices, respectively. By using the simple geometric relations of triangles, we obtain the corresponding image points (X_{cl}, Y_{cl}) and (X_{cr}, Y_{cr}) of point $P(x_0, y_0, z_0)$ in the three-dimensional object space as follows:

$$X_{cl} = -f \tan\left(\tan^{-1}\left(\frac{b + 2x_0}{2z_0}\right) - \beta \right),$$

$$X_{cr} = +f \tan\left(\tan^{-1}\left(\frac{b - 2x_0}{2z_0}\right) - \beta \right),$$

$$Y_{cl} = \frac{-fy_0}{z_0 \cos\beta + [(b/2) + x_0]\sin\beta},$$

$$Y_{cr} = \frac{-fy_0}{z_0 \cos\beta + [(b/2) - x_0]\sin\beta}.$$

Let us define the horizontal and vertical disparities by $d_x \equiv X_{cl} - X_{cr}$ and $d_y \equiv Y_{cl} - Y_{cr}$. Note that the vertical disparity is normally non-zero if $x_0 \neq 0$ or $\beta \neq 0$. Besides, the vertical disparity becomes larger as the baseline increases, the convergence angle increases, or the focal length increases [6]. But in the case of capturing soccer matches, z_0 is dominant compared with baseline b and x_0, i.e., $\|z_0\| >> \max\{\|b\|, \|x_0\|\}$. Hence, we assume that the vertical disparity is small enough. From now on, we will consider only the horizontal disparity for the depth control of 3DTV. If the above condition is satisfied, then the approximated horizontal disparity, d_x, is given by

$$d_x \equiv X_{cl} - X_{cr} \approx 2f \tan\beta.$$

Note that the horizontal disparity becomes a controllable quantity simply by adjusting the angle β for a fixed f. This explains the operational principle of the toed-in stereoscopic camera.

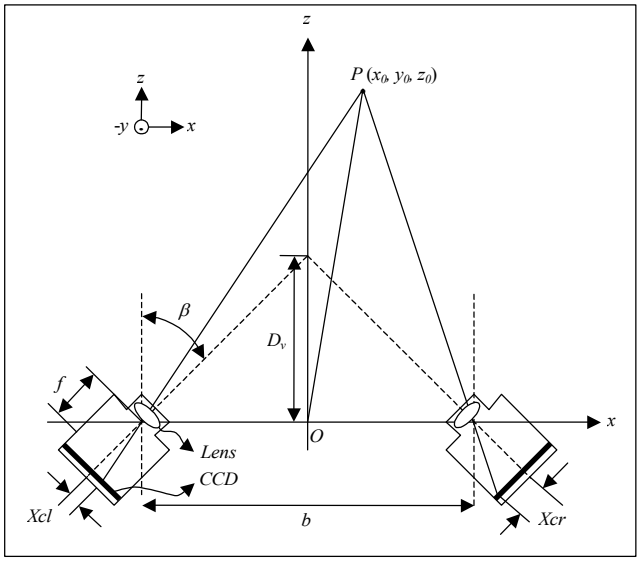


Fig. 2. Geometry of a toed-in stereoscopic camera.

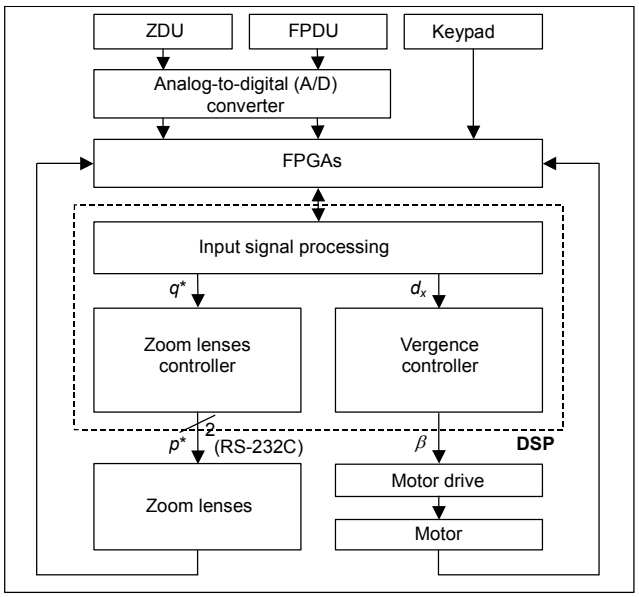


Fig. 3. Block diagram of a 3DTV camera controller.

Figure 3 shows the block diagram of a 3DTV camera controller. This board processes command signals from three user interfaces such as the zoom demand unit, focus position demand unit, and Keypad. In this paper, we consider how to control zoom lenses simultaneously, how to compensate an inherent mismatch between zoom lenses, and how to adjust the horizontal disparity d_x with β. First of all, we synchronized the zoom lenses by using the dedicated RS-232C serial communication protocol provided by the zoom lens maker. To this end, the entire range of zoom was subdivided by 256 or 512 levels. Details of the protocol and its error boundary will

not be addressed here for simplicity. In the follow up to this paper, we will explain the remaining control problems in more detail.

We observed an inherent mismatch of errors between zoom lenses. Most errors were caused by the different characteristics of the electrical and optical mechanisms, especially mismatched zooming patterns and time-varying image centers over the whole dynamic range. The above phenomenon for image centers has been introduced in [7] by stating that the image centers shift by changing the focus and zoom of a lens. This is due to the misalignment of the lens components. In this paper, however, we assume that such an error is mainly caused by the mismatched zooming patterns. That is, we regard that the error component, due to the time-varying image centers, constitutes the above assumed error. Furthermore, we assume that there is no error in focus in either zoom lenses.

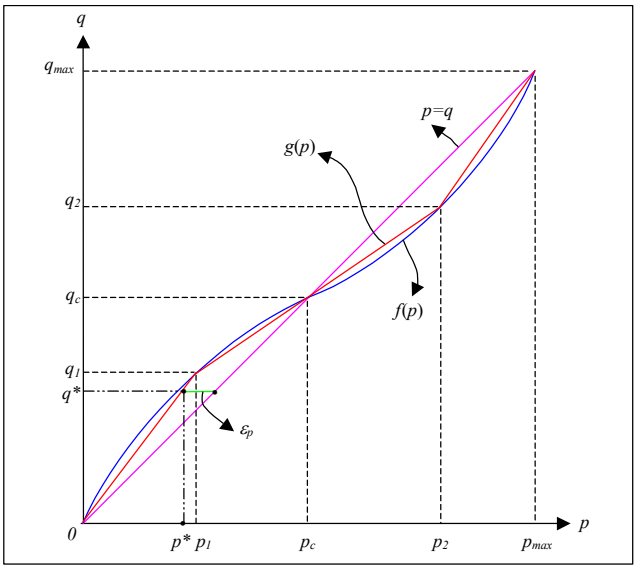


Fig. 4. Proposed scheme of compensating a zooming error with a piece-wise linear function.

Figure 4 shows the proposed scheme of compensating for a zooming error caused by the mismatched zooming patterns using a simple piece-wise linear function. From Fig. 4, p and q denote the zoom position command and the corresponding actual value of zoom position, respectively. We estimate the unknown zoom pattern $f(p)$ as $g(p)$ by using the actual measurements at several specified points including (0,0), (p_1, q_1), (p_2, q_2), and $(p_{\max}, q_{\max})$. Here, we assume that both $f(p)$ and $g(p)$ pass through the center point (p_c, q_c), i.e., $f(p_c) = q_c$ and $g(p_c) = q_c$. Then, for a given desired or actual zoom position q^*, using the input signal processing shown in Fig. 3, the zoom lens controller calculates the zoom position command p^* by $p^* = g^{-1}(q^*)$ or $p^* = q^* - \varepsilon_p$ if either g^{-1} or ε_p is available. The zoom position commands calculated for both zoom lenses are sent separately to the corresponding zoom lenses via RS-232C. This approach enables us to achieve identical zoom positions for both zoom lenses, although the zooming characteristics of both zoom lenses are not the same because of their differing zooming patterns. Besides, another feature of the proposed method is that the unknown zooming patterns are approximated using a few measurements. The following piece-wise linear function was applied for the approximation of the unknown zooming pattern $f(p)$:

$$g(p) = \begin{cases} \dfrac{q_1}{p_1} p & \text{if} \quad 0 \le p \le p_1, \\ \dfrac{q_2 - q_1}{p_2 - p_1}(p - p_1) + q_1 & \text{if} \quad p_1 \le p \le p_2, \\ \dfrac{q_{\max} - q_2}{p_{\max} - p_2}(p - p_2) + q_2 & \text{if} \quad p_2 \le p \le p_{\max}. \end{cases}$$

Hereafter, we address the problem of controlling the horizontal disparity between stereoscopic images. The horizontal disparity of a stereoscopic camera is a very important factor affecting 3D depth perception and visual fatigue [3]. Normally, depth perception is strongly affected by the relationship between convergence and accommodation. Therefore, it is best if the horizontal disparity not exceed a certain limit. For example, an object with zero disparity in a scene is perceived as if it is positioned exactly on the screen when the stereoscopic images are observed. Such distance and the point resulting in zero disparity are called *the convergence distance* and *the convergence point*. Based on the above observations, we know that by increasing or decreasing the horizontal disparity of a specific object in the object space, the perceived distance, or depth, of the specific object can be adjusted. We mentioned previously that the horizontal disparity d_x is controlled by changing the convergence angle 2β, the angle between the optical axes of the cameras. Therefore, a numerical algorithm for transforming the given horizontal disparity command into the convergence angle command should be executed to adjust the perceived depth. This numerical algorithm is executed at the vergence controller implemented on a digital signal processor, as shown in Fig. 3. Fortunately, there are several methods for estimating the horizontal disparity. Some researchers employ cepstrum-like techniques as a method of extracting disparity information from the stereoscopic images. Some utilize a fine rectification method such that the horizontal disparity can be found by searching the corresponding point lying on the epi-polar line, which is generated by the fundamental matrix and a point of

interest on one image [8]. In our case, we adopted an interactive method of adjusting the horizontal disparity command by way of an audio intercom between a director in the OB van and the camera operators in the stadium. In the case of 3DTV field production, a more reliable method is to directly monitor the horizontal disparity and to choose the best one among several cameras in real-time, prior to the transmission. We would say that such a method is an easy way of preventing a sudden disparity change, and thereby providing viewers with more comfortable 3D depth perception.

For the experiment, we developed various 3DTV cameras by using 2/3-inch HDTV camera heads and HDTV zoom lenses, as shown in Table 1. Four stereoscopic cameras were employed for the experimental services of 3DTV broadcasting during the 2002 FIFA World Cup Korea/Japan. Three toed-in cameras were utilized for the transmission of five first-round soccer matches. One parallel-axis camera was used to capture the *KODO* performance in Niigata, Japan for the *Korea-Japan High Data Rate Satellite Communication Experiment*. Note also that a stereoscopic camera with a bi-prism adaptor was developed for special application such as close-up shots, i.e., capturing close objects within 2-3m. We improved the bi-prism adaptor by correcting chromatic aberration. This improved bi-prism adaptor was also applied to a commercial camcorder for use as a 3D camcorder. In this case, we made a 3D camcorder simply by attaching the developed bi-prism adaptor in front of the lens of the commercial camcorder.

Table 1. Camera heads and zoom lenses used for high-definition 3DTV cameras.

Type	Camera head	Zoom lens
Toed-in	DXC-H10(1035i)	HA36 × 10.5BERD
	DXC-H10(1035i)	HA20 × 7.8BERD
	HDL-37E(1080i)	HA17 × 7.8BERD
Parallel-axis	HDL-40(1080i)	HA10 × 5BERD
Bi-prism	DXC-H10(1035i)	HA15 × 8BERD

3. 3DTV Video Multiplexer

Multiplexing a pair of stereoscopic images into one image is an efficient method for solving a possible synchronization problem that may happen during stereoscopic image processing. By treating separate stereoscopic images as a single image, both the effort and equipment required for recording, compression, transmission, and reception are greatly reduced, compared with processing the stereoscopic images separately. Accordingly, a major advantage of such multiplexing is the backward compatibility with existing HDTV transmission systems. Figure 5 shows the structure of a multiplexed frame, which was obtained using a frame decimation scheme. Frame decimation is a process of resizing the left and right images by half in the horizontal direction and rearranging the resized images in a new single frame, as explained earlier in section II.1. The colored region denotes the area of active video.

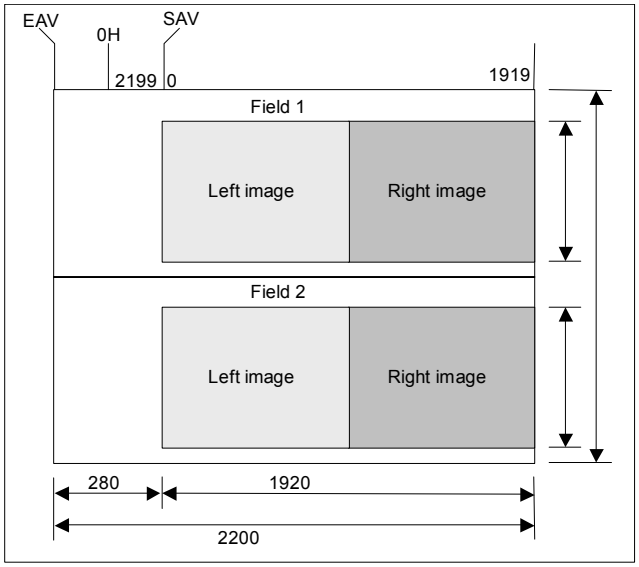


Fig. 5. Structure of a multiplexed frame-by-frame decimation.

Figure 6 shows the block diagram of a 3DTV video multiplexer incorporating the frame-decimation and frame-multiplexing algorithms. If we utilize High Definition Serial Digital Interface (HD-SDI) input channels, then serial digital data streams are first converted to two 20-b parallel data sets. With the start of active video (SAV) identified by the timing reference (TR) codes, 3FF/000/000 and H=0 in XYZ field [10], the multiplexer starts to write and read two input-data sets to and from the buffers for synchronization. Next, the read input data from the buffers are multiplexed by using frame decimation and frame multiplexing. In the case of using analog input channels, the multiplexer digitizes RGB or YPbPr signals in 8-b mode. Then, the multiplexer synchronizes the left and right frames with the horizontal and vertical sync signals detected from the digitized G or Y signals. The remaining process after color formatting and 4:2:2 bit-resolution conversion is the same as the previous one for HD-SDI inputs. The multiplexer outputs the multiplexed frame in both HD-SDI and analog signal format.

Note that there is another multiplexing scheme similar to frame multiplexing in field-sequential mode. Frame multiplexing is a process of synthesizing a new frame with odd and even fields from the left and right frame, alternatively. This field-sequentially multiplexed frame is frequently used to support a field-sequential mode in a 3DTV display. When

viewed with shutter glasses, one can perceive depth from the display. But this method will result in a lot of flicker because the frame rate is lower than 55 Hz, the binocular critical flicker frequency [9]. Correspondingly, for flicker-free stereoscopic 3D, considering the critical flicker frequency, a more sophisticated processing technique is required to increase the field frequency up to at least 110 Hz.

We have implemented a 3DTV video multiplexer by using commercial chips for analog-to-digital (A/D) or digital-to-analog (D/A) conversion, parallel-to-serial conversion (P/S), and serial-to-parallel conversion (S/P). A number of video and signal-processing algorithms run on high-speed and high-capacity field-programmable gate arrays (FPGAs) at the clock frequency of CLK=74.1758 MHz.

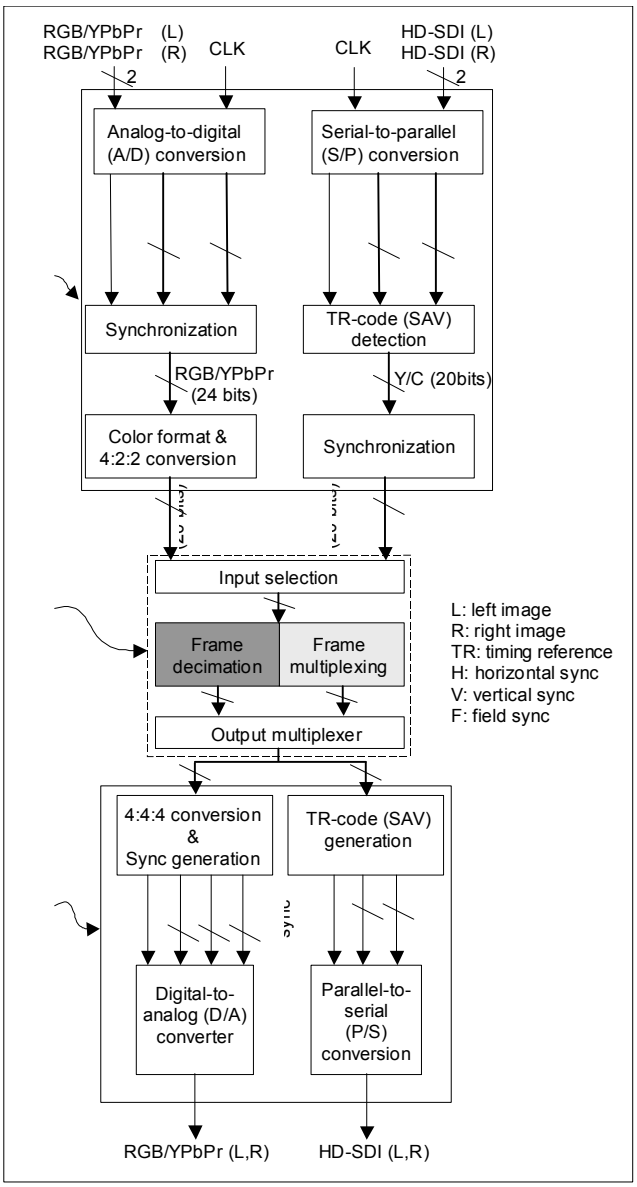


Fig. 6. Block diagram of a 3DTV video multiplexer.

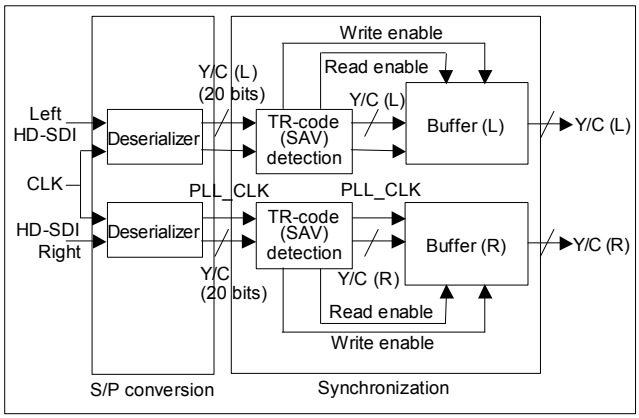


Fig. 7. HD-SDI input processing block.

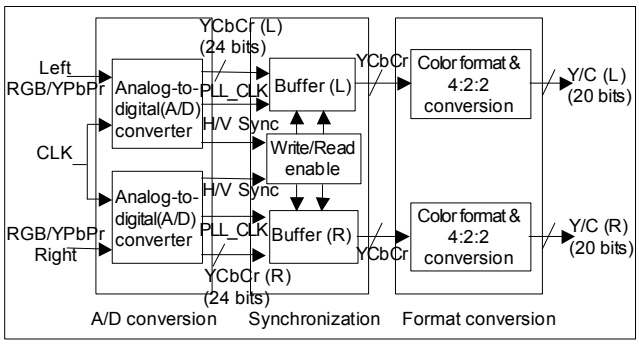


Fig. 8. Analog input processing block.

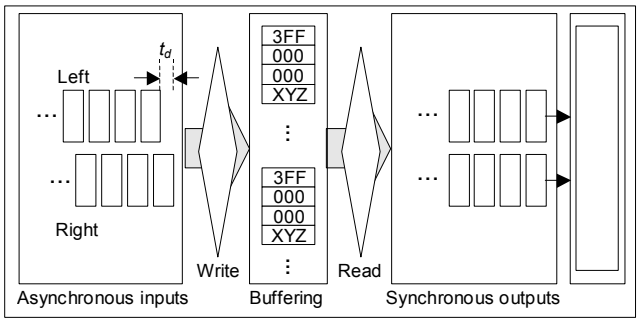


Fig. 9. Proposed scheme of synchronizing left and right input data streams with a small time delay t_d.

Figures 7 and 8 show the input processing blocks for both HD-SDI and analog inputs, respectively. We see from Fig. 7 that HD-SDI inputs are de-serialized into 10-b luminance Y and 10-b color difference C, and then Y/C with a PLL-CLK (Phase Locked Loop – Clock) are fed into the TR-code (SAV) detection block. Details of the synchronization block are shown in Fig. 9. The proposed scheme is a very efficient method of synchronizing asynchronous data streams with a small time delay t_d. If SAV is detected at the TR-code (SAV) detection block, then the block writes and reads left and right data streams to and from the buffers. Next, the synchronized data streams are sent to the 3DTV multiplexing block, as shown in

Fig. 9. However in the case of analog inputs, each analog signal is digitized in 8-b with the A/D converter, as shown in Fig. 8. Unlike the HD-SDI case, two data streams are synchronized with Horizontal/Vertical Sync signal instead of SAV. In this case, the synchronized data streams are sent to the 3DTV multiplexing block through the color format and 4:2:2 conversion block.

The 3DTV video multiplexer can treat both RGB and YPbPr formats. If the input is RGB, the color format conversion to YCbCr is done at the color format conversion block, as shown in Fig. 8. Luminance Y and color difference C are obtained using a 4:4:4 to 4:2:2 bit-resolution conversion, as shown in Fig. 10(a). Note that there is no actual change in the luminance Y.

Figure 11 shows the 3DTV video output processing block of the 3DTV video multiplexer. The multiplexer outputs the multiplexed frame in both HD-SDI and analog formats. For HD-SDI output, TR-codes are regenerated in the TR-code generation block by using H, V and F sync signals. The P/S conversion block serializes the 20-b Y/C data into HD-SDI format. In the case of analog output, bit-resolution conversion from 4:2:2 to the 4:4:4 format is done with interpolation, as shown in Fig. 10(b). Color format conversion is carried out again if RGB is required. The converted signals through the 4:4:4 Conversion & Sync Generation Block becomes analog signals using the D/A converter. Also the composite sync (C sync) signal is provided for external use.

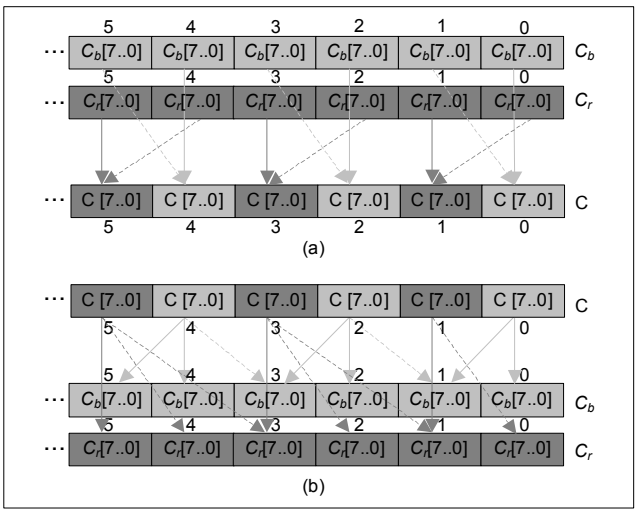


Fig. 10. Bit-resolution conversion: (a) 4:4:4 to 4:2:2, (b) 4:2:2 to 4:4:4.

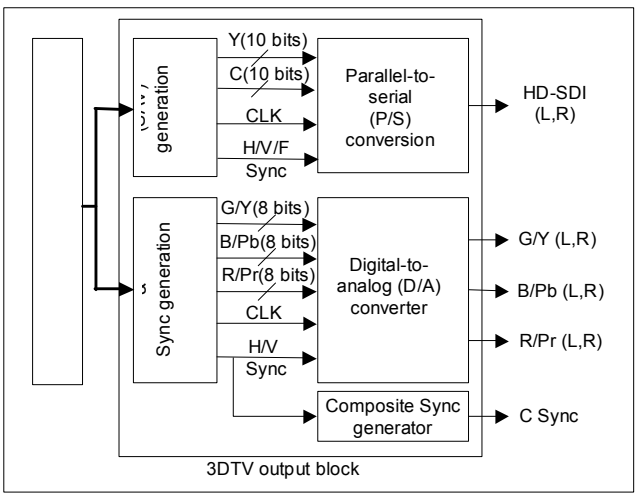


Fig. 11. 3DTV video output processing block.

4. 3DTV Receiver

Figure 12 shows the block diagram of a 3DTV receiver consisting of an MPEG-2 decoding block and a 3DTV video demultiplexing block. The former block performs the same function as the professional HDTV set-top box, and the latter block demultiplexes a multiplexed frame. Note from Fig. 12 that the 3DTV receiver can be applied to both 2D and 3D applications, since the 3DTV receiver can decode a regular 2D HDTV signal by switching off the 3DTV video demultiplexing block. The MPEG-2 decoding block is composed of a 3DTV/HDTV TS input, an HD MPEG-2 decoder, and a processor. The processor is used for initialization, maintenance, and control of the 3DTV receiver. A terrestrial digital broadcasting radio frequency (RF) signal or base band TS Digital Video Broadcasting Asynchronous Serial Interface (DVB-ASI) can be fed into the 3DTV/HDTV TS input. Either the demodulated RF or the other base-band TS is decoded by the HD MPEG-2 decoder. The decoded video signals are available as both HD-SDI and analog format at the end of the 3DTV/HDTV output signal driver. If the input is a 3DTV signal, then the decoded video signal is fed into the 3DTV video demultiplexing block. Note from Fig. 12 that another multiplexed 3DTV video signal can be also fed into this block. This is to expand the flexibility of the 3DTV receiver, since the 3DTV receiver can be used as a stand-alone 3DTV video demultiplexer.

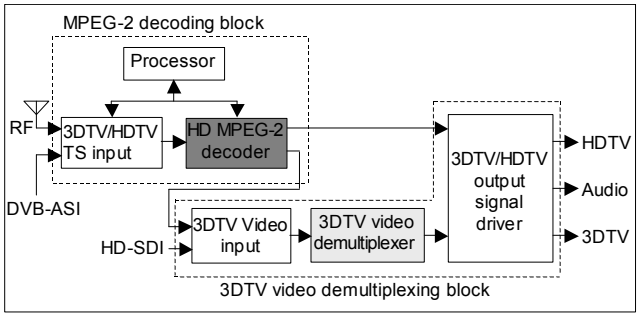


Fig. 12. Block diagram of a 3DTV receiver.

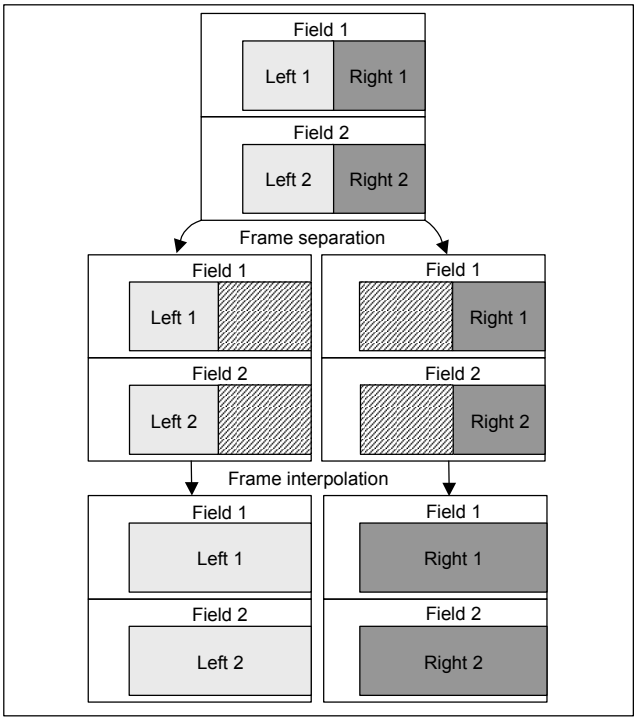


Fig. 13. Process of demultiplexing a multiplexed frame.

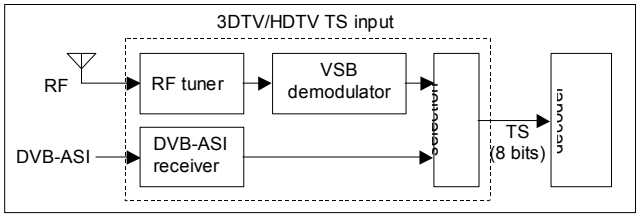


Fig. 14. 3DTV/HDTV TS input block and HD MPEG-2 decoder block of the 3DTV receiver.

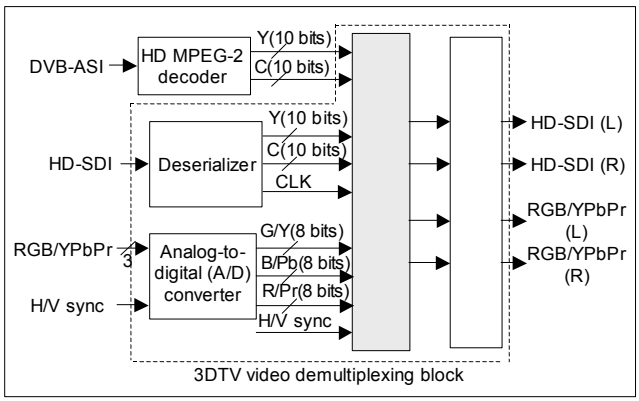


Fig. 15. 3DTV video demultiplexing block of the 3DTV receiver.

Figure 13 shows the process of demultiplexing a multiplexed frame by "frame separation" and "frame interpolation." From Fig. 13, we see that the multiplexed frame is separated into left and right double-frames during the frame separation step, and then interpolated to the original resolution using a proper interpolation filter during the frame interpolation step. The demultiplexed stereoscopic images are directly sent to left and right beam projectors.

Figure 14 shows the 3DTV/HDTV TS input block and the HD MPEG-2 decoder block of the 3DTV receiver. At the 3DTV/HDTV TS input block, the RF signal received through the RF tuner is demodulated with the vestigial sideband demodulator to a base-band DVB-ASI signal. Note that there is another base-band input signal apart from the RF. This signal is the output of a network adaptor installed at a venue and is transmitted from the studio over a DS-3 network to the venue, as shown in Fig.1. The DVB-ASI receiver inputs the base band TS with a differential line receiver. Besides, by using a completely integrated PLL clock synchronizer, the DVB-ASI receiver recovers the timing information required for data refinement. Then, the input selection block must select one of these two base-band signals. The selected base-band signal is sent to the HD MPEG-2 decoder in an 8-b parallel data format. The HD MPEG-2 decoder was implemented with a professional decoding chipset. The decoding chipset separates video and audio streams, and then decodes the video stream compressed in MPEG-2 MP@HL (main profile at high level) and the audio stream compressed in AC-3. The transport packet parser in the decoder handles a transport stream of up to 80 Mbps. The audio interface supports both analog stereo and the digital S/PDIF (Sony/Philips Digital Interface) format.

Figure 15 shows the 3DTV video demultiplexing block of the 3DTV receiver. This block is used to demultiplex a number of multiplexed 3DTV signals. Note from Fig. 15 that the output of the decoder (20-b Y/C), the deserialized (via S/P conversion) HD-SDI (20-b Y/C), and the digitized analog RGB/YPbPr signal can be demultiplexed at the 3DTV video demultiplexer. As explained previously in Fig. 13, the 3DTV demultiplexing algorithms running on high-speed FPGAs separate and interpolate the multiplexed frame by "frame separation" and "frame interpolation" sequentially. The 3DTV/HDTV output signal driver outputs the video signal in HD-SDI format and analog format for compliance with both digital and analog HDTV or beam projectors.

III. Experimental Broadcasting

We developed several subsystems required for the experimental services of 3DTV broadcasting and built an experimental broadcasting system by using all of the developed subsystems shown in Fig. 1. This system was used as a test bed to evaluate the performances of the developed 3DTV broadcasting subsystems. According to the contracts with the host broadcast services, we captured five first-round matches

using three developed stereoscopic cameras in real-time, although the 3DTV OB van can handle at least seven such stereoscopic cameras. There was a delicate time delay between camera signals with the optical-to-electrical conversion at the OB van, so we utilized a frame synchronizer as a simple method of synchronizing the asynchronous stereoscopic cameras. Together with the frame synchronizer we used a portable digital video switcher to select the best camera signal among the three signals in relation to the perceived depth. For the quality of depth, we employed an audio intercom as an auxiliary device for communicating between the director and the cameramen. The audio intercom was very useful in adjusting the horizontal disparity of each camera interactively because the program director, staying at the 3DTV OB van, was able to ask for an adjustment to the cameramen promptly. This prevented an abrupt change in 3D depth. Therefore, direct communication between a director and the cameramen is very important in producing high-quality stereoscopic images, especially in the case of live field-productions.

Since each OB van was at a very long distance from the stereoscopic cameras, we used SM optical fibers, with a wavelength of 1310 nm, as a means for transmitting the multiplexed 3DTV signals to the 3DTV OB van. This is a necessary step for the preservation of signal strength. The reason is that in the case of using a 75 Ω coaxial BNC cable, the transmission of an HD-SDI signal is limited to within 200 meters [10].

Overlaying a computer-generated text or graphics with the selected stereoscopic images, we compressed the final stereoscopic image with an embedded audio at 40 Mbps using an HD MPEG-2 encoder. The compression was done in an MPEG-2 4:2:0 MP@HL format. Then we handed over the TS to the studio via IBC1-Seoul, as shown in Fig. 1.

At the studio, we received the transport streams coming from the World Cup Stadiums and passed them through the DS-3 to the DVB-ASI converters. We call this stream a live stream. Apart from the live streams, there was another stream, this one coming from ETRI. This was the signal transmitted originally from Japan via Koreasat-3. The *KODO* performance held in Niigata, Japan, was captured and sent to ETRI. Recall from Fig. 1 that there was a local 3DTV source. This was to provide stereoscopic images when the live stream was not available. We used a VTR and encoder set as a local transport stream generator. Along with the 3DTV signal, we sent HDTV and data broadcasting signals to the demonstration venues. This was planned to provide participants at the venues with an equal chance to compare digital broadcasting services. By using the 8×8 DVB-ASI router, we chose a proper 3DTV signal from the three resources according to the program schedule. The output of the DVB-ASI router was converted to DS-3 again with a network adaptor and then distributed to all the venues over a DS-3 network.

For this experiment, ten venues in all were constructed in the following places: the main press center at the international media center, Seoul; Sangam Peace Park, Seoul; Yeouido Park, Seoul; and various locations in Busan, Daegue, Daejeon, Ulsan, Gwangju, Incheon, and Seogwipo. We sent HDTV and data broadcasting signals to each venue. But we sent the 3DTV signal to all the venues except for the ones in Ulsan, Gwangju, Incheon, and Seogwipo. We used a pair of internally polarized beam projectors with reflective liquid crystal D-ILA (Digital Direct Drive Image Light Amplifier) devices. The brightness and contrast ratio of this projector are 5000 ANSI lumens and 1000:1, respectively. Owing to an internal polarization scheme, this projector has a major advantage in that no extra effort is needed for attaching a polarization film in front of its lens. As a result, a small amount of brightness is saved. For 3DTV, the screen size used at two of the venues was 300 inches, but the screen size used at the other four venues was 120 inches. Note that the height of the 300-inch screen is 3.7 m for those with an aspect ratio of 16:9. One well-known rule for determining an optimal viewing distance is 2.5 H, where H denotes the screen height [11]. According to this rule, we set the optimal viewing distance to 9 m. Since a front projection scheme was used, the projection distance of the beam projectors should be greater than the optimal viewing distance. By leaving enough space for viewers, we set the projection distance between the beam projector and screen at 19 m. To obtain such a projection distance, we employed a 1.5:1 fixed focus lens, GL-M4015S, for the beam projector. Similarly the same rule was applied to the case of the 120-inch screens. At each venue, a network adaptor converted the DS-3 signal to a 40 Mbps base band TS. This base band TS was decoded and demultiplexed by the 3DTV receiver sequentially, as explained earlier in section II. 4. Then the left and right images were displayed on either a 300-inch or 120-inch silver screen with stereo audio.

To evaluate the performance of the experimental 3DTV broadcasting system objectively, we compared the colors between the stereoscopic images, and measured each image quality in terms of its peak signal-to-noise ratio (PSNR). Furthermore, we performed a subjective evaluation test with regard to the perceived depth and preference for 3DTV.

Figure 16 shows the results of a color comparison between the 520-th scan-lines inside the yellow circles on the stereoscopic image. Figure 16(a) is the multiplexed image using frame decimation, while Fig. 16(b) shows the color plots of red, green, and blue. In order to compare the colors, we have chosen a specific region, such as the inside of the yellow circle, because these regions have a similar horizontal disparity. The 520th scan-lines lie in the yellow circles on the left and

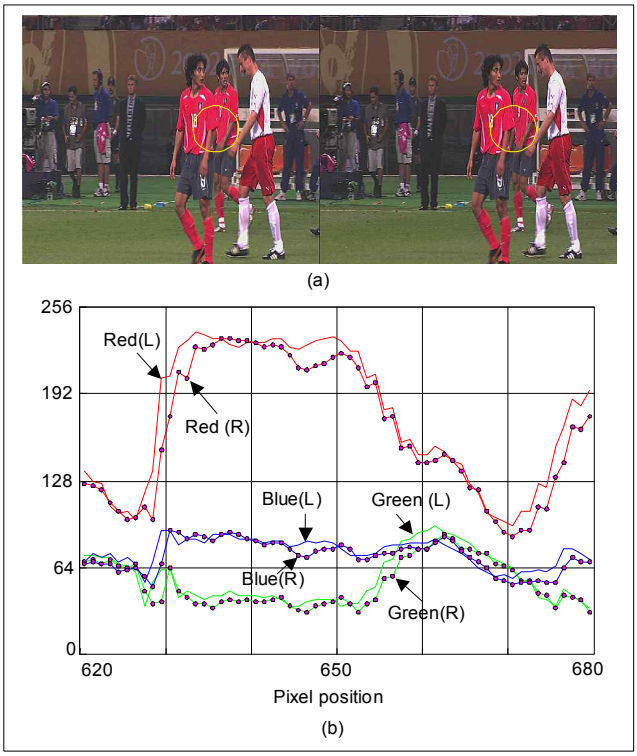


Fig. 16. Color comparison between the 520th scan-lines inside the yellow circles on the stereoscopic image: (a) side-by-side image, (b) color plots.

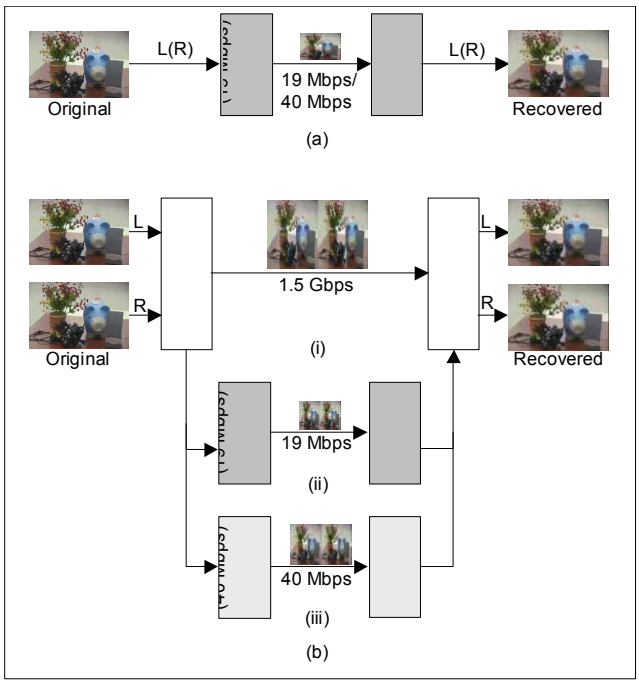


Fig. 17. Several experimental setups for PSNR comparison: (a) HDTV cases, (b) 3DTV cases (i) without encoding, (ii) with 19 Mbps encoding, and (iii) with 40 Mbps encoding.

right images, respectively. We obtained the color plots, shown in Fig. 16(b), from the 620th pixel position to the 680th pixel position on the 520th scan-lines. The horizontal and vertical scales denote the pixel position on the scan-lines and the color value of between 0 and 255. From Fig. 16(b), one can see that the color of the left and right images matches well.

Figure 17 shows the several experimental setups used for PSNR comparison. For this comparative study of PSNRs, we considered two bit-rate cases, 19 Mbps and 40 Mbps for HDTV, while three bit-rate cases, 19 Mbps, 40 Mbps, and 1.5 Gbps were considered for 3DTV. Note that the bit-rate of 1.5 Gbps means that the stereoscopic images are multiplexed and demultiplexed without encoding. In the case of HDTV, a single left-or-right image was encoded without multiplexing. Then the encoded image was decoded by a professional HDTV receiver. We calculated the PSNR for the recovered left-or-right image, as shown in Fig. 17(a). In the case of 3DTV, on the other hand, we applied the developed 3DTV video multiplexer before encoding and then applied the developed 3DTV receiver to obtain the stereoscopic images, as shown in Fig. 17(b). For all cases, the assigned bit-rate for the audio was fixed at 384 Kbps. Based on the above conditions, Table 2 shows the PSNRs of the left and right images of five cases. From Table 2, note that the reduction of the PSNR is not so severe using the developed 3DTV broadcasting subsystems,

Table 2. Comparison results of PSNR.

		PSNR (dB)	
		L	R
HDTV	19 Mbps	33.78	34.13
	40 Mbps	34.15	34.35
3DTV	1.5 Gbps	35.47	35.71
	19 Mbps	32.53	33.75
	40 Mbps	33.20	34.05

compared with HDTV. At 19 Mbps, the PSNRs of the left image are 32.53 dB for 3DTV and 33.78 dB for HDTV. Similarly, the PSNRs of the right image at 19 Mbps are 33.75 dB for 3DTV and 34.13 dB for HDTV. Comparing the PSNRs of the left images at 19 Mbps, we see that the PSNR of 3DTV is lower than that of HDTV by 1.25 dB. But, the PSNR of 3DTV increases with the increased bit-rate. Note that the PSNR of 3DTV at 40 Mbps is similar to that of HDTV at 19 Mbps. This is because 20 Mbps is assigned to the left and right images for 3DTV encoding.

Figure 18 shows the results of our subjective evaluation in regard to the depth perception of 3DTV and viewer preference for 3DTV over other digital broadcasting services. For this subjective evaluation test, we used a question sheet including the following questions: (i) What do you feel about 3D depth?

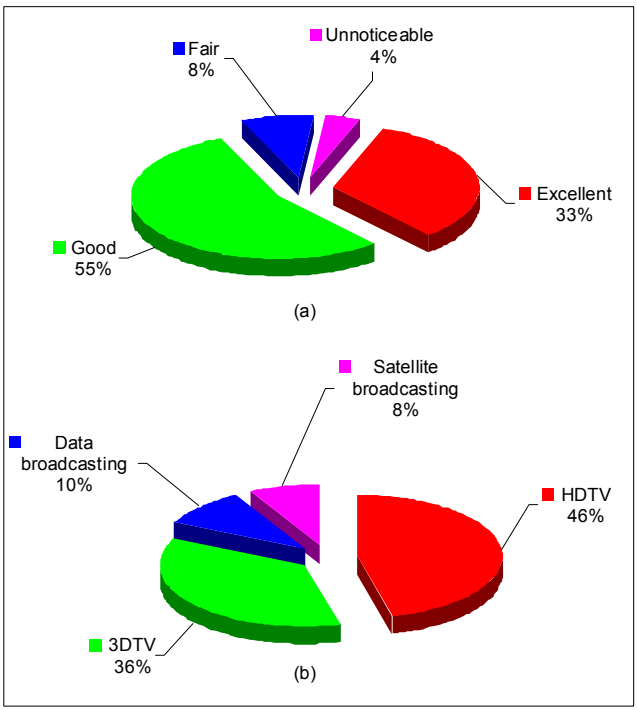


Fig. 18. Results of a subjective evaluation: (a) depth perception of 3DTV, (b) preference among digital broadcasting schemes.

(the choices were Excellent, Good, Fair, and Unnoticeable); (ii) Which is the most interesting service among digital broadcasting services? (the choices were HDTV, 3DTV, Data broadcasting, and Satellite broadcasting). In the choices of question (i), "Unnoticeable" may be interpreted as either a *bad* rating or *poor* rating, according to the rating scale of the double-stimulus continuous quality-scale method described in ITU-R Recommendation 500-10 [12]. In this experiment, we gathered test results from 273 viewers, 177 males and 96 females. 9% of the viewers were under 20 years of age, 48% of the viewers were between 20 and 40, and 43% of the viewers were above 40. All of the viewers were non-experts who were not directly related with 3D depth perception and digital broadcasting services. Note from Fig. 18(a) that 33% of the 273 viewers rated the perceived depth of 3DTV as "Excellent." Note also from Fig. 18(b) that 36% of the viewers rated 3DTV as the most interesting digital broadcasting service. According to Fig. 18, we would say that 88% of 273 viewers evaluated the depth of 3DTV as "Good" or better and 36% of the viewers preferred 3DTV to other digital broadcasting services.

IV. Concluding Remarks

We have developed 3DTV broadcasting subsystems that are fully compatible with the existing HDTV broadcasting infrastructure, except for the display and camera. With a terrestrial and satellite DS-3 network, we have verified the developed subsystems by performing experimental services of 3DTV live broadcasting during the 2002 FIFA World Cup Korea/Japan. The PSNR of a 3DTV image at 40 Mbps was comparable with that of an HDTV image at 19 Mbps. Interestingly, 88% of 273 viewers involved in the subjective evaluation experiment rated the perceived depth of 3DTV as "Good" or better and 36% of the viewers preferred 3DTV to other digital broadcasting services. Accordingly, we would say that 3DTV has a possibility of being one of the promising next-generation multimedia services following HDTV. Based on the developed stereoscopic-image processing techniques and its subsystems, a multiview stereoscopic broadcasting will be investigated in future studies.

Acknowledgment

The authors would like to thank the anonymous reviewers whose comments improved this paper.

References

[1] RACE 2045–DISTIMA, http://www.tnt.uni-hannover.de/plain/project/eu/distima/.

[2] G. Gagnon, S. Subramaniam, and A. Vincent, "3-D MPEG-2 Video Transmission over Broadband Network and Broadcast Channels," *Stereoscopic Displays and Virtual Reality Systems VIII*, *Proc. SPIE*, vol. 4297, 2001, pp. 290-298.

[3] B. Javidi and F. Okano, "Three-Dimensional Video and Display: Devices and Systems," *Critical Review CR76*, 2001.

[4] N. Hur, C. Ahn, and Chieteuk Ahn, "Experimental Service of 3DTV Broadcasting Relay in Korea," *Three-Dimensional TV, Video, and Display II*, *Proc. SPIE,* vol. 4864, 2002, pp. 1-13.

[5] ITU-T Recommendation G.703, *Physical/Electrical Characteristics of Hierarchical Digital Interfaces,* 2001.

[6] A. Woods, T. Docherty, and R. Koch, "Image Distortions in Stereoscopic Video Systems," *Stereoscopic Displays and Applications IV*, *Proc. SPIE*, vol. 1915, 1993, pp. 36-48.

[7] R.G. Wilson and S.A. Shafer, "What is the Center of the Image?" *J. Opt. Soc. Am. A*, vol. 11, no. 11, 1994.

[8] O. Faugeras, *Three-Dimensional Computer Vision: a Geometric Viewpoint*, MIT Press, 1999.

[9] M. Okui, A. Hanazato, F. Okano, and I. Yuyama, "A Study on Scanning Methods for a Field-Sequential Stereoscopic Display," *IEEE Trans. CSVT*, vol. 10, no. 2, 2000, pp. 244-253.

[10] ANSI/SMPTE 292M-1996, *Television—Bit-Serial Digital Interface for High-Definition Television Systems*, 1996.

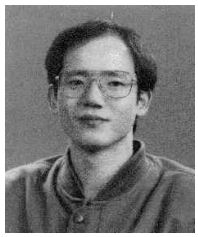
[11] W.A. Ijsselsteijn, H. de Ridder, and J. Vligen, "Subjective Evaluation of Stereoscopic Images: Effects of Camera Parameters and Display Duration," *IEEE Trans. CSVT*, vol. 10, no. 2, 2000, pp. 225-233.

[12] ITU-R Recommendation BT.500-10, *Methodology for the Subjective Assessment of the Quality of Television Pictures*, 2000.

Namho Hur received the BS, MS, and PhD degrees in electrical and electronic engineering from Pohang University of Science and Technology (POSTECH), Pohang, Korea, in 1992, 1994, and 2000. He is currently with the Digital Broadcasting Research Division, Electronics and Telecommunications Research Institute (ETRI), Daejeon, Korea. As a research scientist, he spent a year with Communications Research Centre Canada (CRC) from 2003 to 2004. His main research interests are ac motor drives, control theory and its application to power electronics, high-performance power converter/inverter systems, three dimensional television (3DTV) broadcasting systems, and perceptual requirements of stereoscopic multiview video systems.

Gwangsoon Lee obtained the MA in electronics from the Kyungpook National University in Korea. He joined Electronics and Telecommunications Research Institute (ETRI) in 2001, and he is currently with the Broadcasting System Research Group. His research interests include the 3DTV broadcasting system and the integration of broadcast and telecommunications in DMB systems.

Woongshik You received the BS and MS degrees in computer engineering from Chungnam National University, Daejeon, Korea, in 1997 and 2000. In 1998 he joined Electronics and Telecommunications Research Institute (ETRI), where he works on the research of digital broadcasting technologies. His research interests include the 3DTV broadcasting system and digital CATV system.



Jinhwan Lee received the BS degree in communication engineering from Hankuk Aviation University of Korea in 1987 and the MS degree in communication engineering from the Information and Communications University (ICU) of Korea in 2000. He joined the Electronics and Telecommunications Research Institute (ETRI) in Daejeon, Korea, in 1989. He is currently engaged in the development of a Digital Multimedia Broadcasting (DMB) system. His main research interests are digital video compression and video coding for broadcasting.

Chunghyun Ahn obtained his PhD in remote sensing & GIS from the Chiba University of Japan. He joined Electronics and Telecommunications Research Institute (ETRI) in 1996, and he is currently with the Broadcasting System Research Group. He has interests in the fields of integration of remote sensing and GIS, 3DTV broadcasting systems, and the integration of broadcast- and location-based services in DMB systems.