The adaptation of 3D stereoscopic video in MPEG-21 DIA

Man Bae Kim\textsuperscript{a,\ast}, Jeho Nam\textsuperscript{b}, Woonhak Baek\textsuperscript{a}, Jungwha Son\textsuperscript{a}, Jinwoo Hong\textsuperscript{b}

\textsuperscript{a}Department of Computer, Information, and Telecommunication, Kangwon National University, 192-1 Hyojo2-dong, Chuncheon 200-701, South Korea

\textsuperscript{b}Broadcasting Media Research Department, Electronics and Telecommunications Research Institute, 161 Gajeong-Dong, Yuseong-Gu, Daejeon 305-350, South Korea

Abstract

In this paper, we present descriptions for 3D stereoscopic video conversion that can be utilized in MPEG-21 digital item adaptation. Our description tools provide a functionality of description of user characteristics and terminal capabilities. In particular, the description of user’s display presentation preferences supports 3D stereoscopic conversion of 2D video (2D/3D video conversion) as well as 2D conversion of 3D stereoscopic video (3D/2D video conversion) so that users can enjoy either 2D or 3D stereoscopic video they prefer. Furthermore, we investigate the effects of the type of parallax, the range of depth, and the maximum interval of a delayed frame on the perception of 3D depth. The syntax of the descriptions and their elements is represented in eXtensible Markup Language schema, and their feasibility has been examined and verified in our experiments, where MPEG test sequences are adopted for 2D/3D video conversion. On the contrary, ordinary stereoscopic test video is used for 3D/2D video conversion. © 2003 Elsevier B.V. All rights reserved.

Keywords: MPEG-21; Digital item adaptation; 3D stereoscopic video; Stereoscopic conversion

1. Introduction

The vision for MPEG-21 is to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities [2]. ‘Digital Item’ is a structured digital object including a standard representation, identification, and metadata, which is a fundamental unit of distribution and transaction within MPEG-21 multimedia framework.

Digital item adaptation (DIA) is one of main MPEG-21 parts. The goal of the DIA is to achieve interoperable transparent access to multimedia contents by shielding users from network and terminal installation, management and implementation issues. This will enable the provision of network and terminal resources on demand to form user communities where multimedia content can be created and shared, always with the agreed/contracted quality, reliability and flexibility, allowing the multimedia applications to connect diverse sets of users, such that the quality of the user experience will be guaranteed [9,11].

As shown in Fig. 1, the combination of resource adaptation and descriptor adaptation produces a newly adapted (modified) Digital Item. DIA Tools store all the necessary data of descriptor information.
Then the descriptor adaptation transforms the input descriptor, $D$, to the output one. As well, the input resource, $R$ is modified into the output resource. Output $D$ and $R$ are then delivered to a user that has requested the adapted Digital Item.

There could be a variety of Digital Items that need to be managed in the DIA framework. Among them, we are concerned primarily with the adaptation of three-dimensional (3D) stereoscopic video that is well fitted into the concept of the DIA.

While two-dimensional (2D) video has been a general media so far, 3D video has been also introduced in the field of information and telecommunications. The stereoscopic image and video are easily found at many Internet sites, DVD titles, etc. The fields of stereoscopic image and video are composed of acquisition, compression, processing and display. A stereoscopic image is acquired from a stereoscopic camera. The camera is generally equipped with functions controlling the distance between two cameras, convergence angle, etc., besides focus and zooming. These are required to acquire stereoscopic images suitable to users’ preferences such as parallax type and 3D depth of the image. As well, the size of the stereoscopic image is twice as much as that of a monoscopic image. Without an appropriate compression, the size is so huge that the transmission and storage of the stereoscopic image is almost impossible. It is well known that video compression schemes such as MPEG, H26X utilize the spatial and temporal redundancy of neighboring images. For stereoscopic video compression, the disparity existing between left and right images is also utilized to remove the stereoscopic redundancy. Processing can include filtering, enhancement, and so forth. For display, the left and right images need to be presented to both human eyes. For a user to experience the perception of 3D depth, special 3D viewing devices are needed. They are generally composed of stereoscopic and auto-stereoscopic systems. The former needs special 3D eye glasses, while no eye glasses are needed for auto-stereoscopic display systems such as 3D monitor and 3D TV. Anaglyph glasses can be used to view stereoscopic images in anaglyph mode.

As mentioned above, the stereoscopic video is produced using a stereoscopic camera with a pair of left and right cameras. Unlike the stereoscopic video, the 3D stereoscopic conversion of 2D video (2D/3D video conversion) makes it possible for users to feel 3D depth perception from ordinary 2D video data. Compared with general stereoscopic images acquired from a stereoscopic camera, an essential difference is that the stereoscopic conversion is to generate a stereoscopic image from a single 2D image. On the other hand, 2D video can be extracted from 3D stereoscopic video acquired from a stereoscopic camera and this conversion is called 3D/2D video conversion.

When users wish to enjoy 3D depth perception from an ordinary 2D video data such as TV, VCD, DVD, etc., the input 2D video is converted to 3D stereoscopic video (i.e., in the sense of resource adaptation in MPEG-21 DIA due to user’s display presentation preferences) and delivered to them. On the contrary, if 3D stereoscopic video is the input to a user who does not want any 3D depth perception (or unfortunately, the user’s terminal does not support 3D viewing capability), the input 3D stereoscopic video is converted to 2D video in order to adapt the user’s presentation preferences and the terminal capabilities. As well, users need to send information of terminal capabilities, such as monitor type and rendering format of the stereoscopic video, for 2D/3D and/or 3D/2D video conversion.
This paper is organized as follows. Section 2 presents the overall architecture of DIA for 3D stereoscopic video conversion. We introduce the basic methodology for the stereoscopic video conversion in Section 3. In Section 4, we propose our DIA descriptions for the stereoscopic video conversion. Experimental results will be presented in Section 5. The feasibility of our descriptions will be examined based upon the experimental results. Section 6 outlines this paper and our future plans.

2. DIA architecture

We propose efficient description tools for MPEG-21 DIA, in which associated video resources are converted from 2D video to 3D stereoscopic video or vice versa. Fig. 2 illustrates how both 2D/3D and 3D/2D video conversions are implemented in DIA framework. Using either a monoscopic camera or a stereoscopic camera, 2D video or 3D stereoscopic video is acquired, respectively. These raw data needs to be encoded using any general encoding schemes if necessary [7]. The role of the DIA consists of 2D/3D video conversion in the case of 2D input video as well as 3D/2D video conversion for 3D stereoscopic video. Users request their adaptation information which is stored in DIA description tools.

The proposed DIA description tools for 3D stereoscopic video are categorized into two description schemes (DS) that functionally belong to Presentation Preferences and Terminal Capabilities in Usage Environment description tools for MPEG-21 DIA [11]. The 2D/3D video conversion-related Presentation Preferences enable users to enjoy their best 3D experiences. If users select their 2D/3D video conversion-related preferences, the adapted DI containing converted video resources will be then delivered. In the Terminal Capabilities, the 3D related characteristics of the terminal, especially related to 3D capability is proposed.

Our proposed DIA description tools are independent of any specific coding/compression standard schemes, because all elements defined in our description are essential to characterize the presentation and viewing capabilities of 3D stereoscopic video. Hence, the proposed descriptions can support most of types of 3D stereoscopic video that have been widely used in industrial applications as well as academic research.

3. Method for stereoscopic video conversion

In this section, we describe the methodology underlying the stereoscopic video conversion. In
2D/3D video conversion, the stereoscopic image for each image frame should be generated. Usually, the motion and image information is used, from which depth information can be obtained. The depth might be defined on pixel or block basis. Using such information, it can be controlled by user preferences such as parallax type and range of depth. Among the parallax type, a negative parallax is easily carried out by simply moving a convergence point further. By the convergence point, we mean a 3D point where the optical axes of two cameras intersect together.

Depth can be derived from the properties of an image. The human visual system (HVS) uses many psychological depth cues to disambiguate the relative positions of objects in a 3D scene, including linear perspective, shading and shadowing, aerial perspective, interposition, texture gradient and color. For instance, bright-colored objects will appear to be closer than dark-colored objects, and we can perceive detail more easily in objects that are closer to us. As objects become more distant, the texture becomes blurred. One of stereoscopic image generation methods is to represent and integrate the two depth cues. Another method is to make use of a previous (delayed) image frame if object and/or camera movement exists. Then, a previous image and a current image become a stereoscopic image. If they are appropriately displayed to both human eyes, a user can perceive 3D depth. In this method, the selection of one of the previous images plays an important role in the perception of 3D depth.

In Section 3.1, we explain one of the 2D/3D video conversion schemes. Similar to this approach, a variety of methods have been proposed [1,3–5,12]. The principal method is to utilize a previous frame, with which a current frame forms a stereoscopic image. The stereoscopic conversion will focus on such method. The 3D/2D video conversion is relatively simple and presented in Section 3.2.

3.1. Conversion from 2D video to 3D stereoscopic video

Unlike general stereoscopic images acquired from a stereoscopic camera, an essential difference is that the stereoscopic conversion is to generate stereoscopic images from 2-D image sequence. This technology requires motion and image analysis. There have been many efforts in the stereoscopic processing of 2D video. We review some of those methods related to the stereoscopic conversion. Among them, the simplest one is to use a current image and a delayed image chosen from an image sequence for stereoscopic images. A modified time difference (MTD) method has been presented [12]. A dedicated digital video processor to perform stereoscopic conversion of NTSC video was developed. This system senses the direction of horizontal movements and adaptively adjusts delays while selecting an appropriate eye to receive a delay image. Spatio-temporal interpolation method has been proposed [1]. Using the range of spatio-temporal sampling density, the stable stereoscopic perception can be achieved. Computed image depth (CID) method is to compute the depth of a given image and then to generate two perspective-projected images that are displayed to left and right eyes [5].

We presented a stereoscopic conversion scheme that is performed on MPEG compressed video [3,4]. Since the motion vector field (MVF) can be obtained either from MV data in MPEG compressed domain or from any motion estimation methods, an image type analysis is carried out in order to determine the image type of the current image. The result is provided to the stereoscopic image generation (Fig. 3). Note that the image type is classified into two categories: (i) a static image having no movement, a non-horizontal motion image in which a camera and objects move in the non-horizontal direction, and (ii) a horizontal motion image in which a camera and objects move in the horizontal direction. In this paper, we will deal with the case of the horizontal motion, which is sufficient to verify that the stereoscopic video conversion is feasible.

For a horizontal motion image, a stereoscopic image is composed of a current image and a delayed image, where each eye sees the same visual field, but at different times. The delay factor that chooses a previous frame plays an important role in the stereoscopic perception. The delay factor determines one of the previous images according to its value.
There are two methods of selecting an appropriate eye to receive a previous image. The first one is to determine the current image as the left image and the previous image as the right image, which is referred to as a mode A. The second one is for determining the current image as the right image and the delay image as the left image, which is referred to as a mode B. When stereoscopic images of the horizontal motion image are formed, it is important to correctly determine a mode A or a mode B.

The movement of the camera and/or the object existing in the image is classified into eight types of motion as shown in Table 1 that describes motion type, its associated left/right image selection, and perceived 3D parallax. There are the motion directions of the camera and/or the object according to each motion type. The left image and the right image are adaptively determined according to the classified motion type. The cases where the motion types are classified into the mode A are 🗼 camera left motion, 🗼 object right motion, 🗼 camera right motion and object right motion, and 🗼 camera left motion and object right motion. The cases where the motion types are classified into the mode B are 🗼 camera right motion, 🗼 object left motion, 🗼 camera right motion and object left motion, and 🗼 camera left and object left motion. In the mode A, the current image and the delay image are viewed to the left eye and the right eye, respectively. On the other hand, in the mode B, the current image and the delay image are viewed to the right eye and the left eye, respectively. For the image with left object motion, previous image and current image are viewed to left and right eyes (images), respectively. Then, the 3D depth of the object will be in the negative parallax, so that it will be perceived between the monitor screen and human eyes. The motion and image analysis algorithms are methodology-dependent scheme, so that they will not be covered in this paper.

3.2. Conversion from 3D stereoscopic video to 2D video

This conversion requires the extraction of 2D video from 3D stereoscopic video. A simple method is to choose either the left video or the right video given a stereoscopic video. Besides

---

**Table 1**

<table>
<thead>
<tr>
<th>Motion type</th>
<th>Camera motion</th>
<th>Object motion</th>
<th>Left image</th>
<th>Right image</th>
<th>Parallax type</th>
</tr>
</thead>
<tbody>
<tr>
<td>🗼</td>
<td>Right</td>
<td>None</td>
<td>Previous</td>
<td>Current</td>
<td>NP</td>
</tr>
<tr>
<td>🗼</td>
<td>Left</td>
<td>None</td>
<td>Current</td>
<td>Previous</td>
<td>NP</td>
</tr>
<tr>
<td>🗼</td>
<td>None</td>
<td>Right</td>
<td>Current</td>
<td>Previous</td>
<td>Object NP, BG ZP</td>
</tr>
<tr>
<td>🗼</td>
<td>None</td>
<td>Left</td>
<td>Previous</td>
<td>Current</td>
<td>Object NP, BG ZP</td>
</tr>
<tr>
<td>🗼</td>
<td>Right</td>
<td>Right</td>
<td>Current</td>
<td>Previous</td>
<td>Object ZP, BG PP</td>
</tr>
<tr>
<td>🗼</td>
<td>Right</td>
<td>Left</td>
<td>Current</td>
<td>Previous</td>
<td>Object NP, BG NP</td>
</tr>
<tr>
<td>🗼</td>
<td>Left</td>
<td>Left</td>
<td>Current</td>
<td>Current</td>
<td>Object ZP, BG PP</td>
</tr>
<tr>
<td>🗼</td>
<td>Left</td>
<td>Right</td>
<td>Current</td>
<td>Previous</td>
<td>Object ZP, BG PP</td>
</tr>
</tbody>
</table>

NP = negative parallax, PP = positive parallax, ZP = zero parallax, BG = background.
Fig. 4. Schematic diagram of StereoscopicVideoConversion DS.

```xml
<complexType name="StereoscopicVideoConversionType">
  <sequence>
    <element name="From2DTo3DStereoscopic" minOccurs="0">
      <complexType>
        <sequence>
          <element name="ParallaxType">
            <simpleType>
              <restriction base="string">
                <enumeration value="Positive"/>
                <enumeration value="Negative"/>
              </restriction>
            </simpleType>
          </element>
          <attribute name="DepthRange" type="mpeg7:zeroToOneType " use="optional"/>
          <attribute name="MaxDelayedFrame" type="nonNegativeInteger" use="optional"/>
        </sequence>
      </complexType>
    </element>
    <element name="From3DStereoscopicTo2D" minOccurs="0">
      <complexType>
        <sequence>
          <element name="LeftRightInterVideo">
            <simpleType>
              <restriction base="string">
                <enumeration value="Left"/>
                <enumeration value="Right"/>
                <enumeration value="Intermediate"/>
              </restriction>
            </simpleType>
          </element>
        </sequence>
      </complexType>
    </element>
  </sequence>
</complexType>
```

Fig. 5. The syntax of StereoscopicVideoConversion DS in XML.
them, a synthesized video can be generated from the left and right video and delivered to users with the better quality.

4. Proposed DIA descriptions

This section presents two main DS for 3D stereoscopic video conversion, which are primarily related to Display Presentation Preferences and Terminal Capabilities that are currently defined by DIA description tools in [11]. We propose two DS; StereoscopicVideoConversion DS and StereoscopicVideoDisplay DS. The former is currently adopted under Display Presentation Preferences, and the latter under Display of Terminal Capabilities in [8].

4.1. Stereoscopicvideoconversion DS

Fig. 4 depicts a XML schema diagram of the proposed StereoscopicVideo Conversion DS that is under DIA/UserCharacteristics/PresentationPreferences/Display/ node tree as defined in [8,11]. As well, Fig 5 shows the syntax of the StereoscopicVideoConversion described in XML. Subsequently, its semantics is explained in Table 2.

Note that From2DTo3D stereoscopic supports the video conversion from 2D video to 3D stereoscopic video. In particular, a user can describe her/his own display presentation preferences by specifying the following three elements: ParallaxType, DepthRange, and MaxDelayedFrame.

Table 2
The semantics of StereoscopicVideoConversion DS

<table>
<thead>
<tr>
<th>Name</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>StereoscopicVideoConversionType</td>
<td>Tool that describes a User’s preference for stereoscopic video conversion.</td>
</tr>
<tr>
<td>From2DTo3D stereoscopic</td>
<td>Describes parameters to perform 2D to 3D stereoscopic video conversion.</td>
</tr>
<tr>
<td>ParallaxType</td>
<td>Describes the type of parallax including negative and positive parallaxes.</td>
</tr>
<tr>
<td>DepthRange</td>
<td>Describes the range of 3D depth perceived by the user. The range is the distance between monitor screen and objects in 3D. Its value is varied at [0.0,1.0], where 0.0 and 1.0 indicate the lowest and highest of the depth, respectively.</td>
</tr>
<tr>
<td>MaxDelayedFrame</td>
<td>Describes the maximum interval of a delayed frame.</td>
</tr>
<tr>
<td>From3D StereoscopicTo2D</td>
<td>Describes parameters to perform 3D stereoscopic to 2D video conversion.</td>
</tr>
<tr>
<td>LeftRightInterVideo</td>
<td>Describes the video between the left and right video of the stereoscopic video that is preferred by the user when converting 3D stereoscopic video into 2D video.</td>
</tr>
</tbody>
</table>

Fig. 6. Negative and positive parallaxes.

Fig. 7. Varying ranges of 3D depth for positive parallax.

ParallaxType represents the type of the parallax being composed of positive parallax and negative parallax [6]. This description can be used by the
resource adaptation of 3D stereoscopic video in order to deliver the perception of 3D depth. In the negative parallax, the 3D depth is perceived between the monitor screen and human eyes. On the contrary, the 3D depth is perceived behind the monitor screen in the positive parallax as illustrated in Fig. 6. In (a), a left image $I_L$ and a right image $I_R$ are displayed to left and right eyes, respectively, where an object is perceived in the negative parallax. By switching the two images like in (b), the object is perceived in the positive parallax.

DepthRange indicates the range of 3D depth perceived by the user and defined as the distance between the monitor screen and the object in 3D. It applies identically to the positive and negative parallaxes. Fig. 7 illustrates its concept in the positive parallax. Suppose that a point A is the convergence point. The range of depth is $D_A$. For a point B, it is $D_B$. So, comparing to a point B, the greater depth is perceived at a point A. By shifting either a left or a right image appropriately, the range of 3D depth can be varied. The amount of DepthRange is varied at [Min, Max], which can be normalized to [0,1]. For the positive and negative parallaxes, shifting the right image to the left direction, the range of depth increases. On the contrary, shifting it to the right direction decreases the range of depth.

One of the stereoscopic conversion schemes is to make use of a delayed (previous) image. Suppose that the image sequence is \{$I_K, I_{K-1}, I_{K-2}, \ldots$\} and $I_K$ is the current frame. One of the previous frames, $I_K-i (i > 1)$ is chosen. Then, a stereoscopic image consists of $I_K$ and $I_{K-i}$. If the current and previous images are appropriately

![Schematic diagram of StereoscopicVideoDisplay DS.](image)

![Syntax of the StereoscopicVideoDisplay DS.](image)
presented to both human eyes, then the user feels the 3D stereoscopic perception. \textit{MaxDelayed-Frame} determines the amount of \( i \) value. Thus, the larger it is, the more depth the user feels.

\textit{From3DStereoscopicTo2D} indicates that 3D stereoscopic video is converted to 2D video. \textit{LeftRightInterVideo} needs to be specified when 3D stereoscopic video is converted to 2D video (i.e., 3D/2D video conversion). Since the stereoscopic image is generally composed of the left image \( I_L \) and the right image \( I_R \), a user can prefer an either (single) image that we wish to receive. Hence, a user might select a single video between \( I_L \) and \( I_R \). As well, a certain type of image processing techniques can be performed on a pair of \( I_L \) and \( I_R \), thus producing a synthesized image with the better quality.

4.2. Stereoscopic video display DS

Fig. 8 depicts a XML schema diagram of the proposed \textit{StereoscopicVideoDisplay} DS that is expected to be under \textit{DIA/TerminalCapabilities} node tree \cite{8,11}. It has the following two elements: \textit{DisplayDevice} and \textit{RenderingFormat}. Subsequently, the XML syntax of the \textit{StereoscopicVideoDisplay} DS is shown in Fig. 9.

The semantics of the \textit{StereoscopicVideoDisplay} are described in Table 3. \textit{DisplayDevice} is related to the characteristics of the display device regarding

![Fig. 10. Shows the different 3D depth of the interlaced images as MaxDelayedFrame is varied at [2,10].](image)

<table>
<thead>
<tr>
<th>Table 3</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The semantics of StereoscopicVideoDisplay DS</strong></td>
<td></td>
</tr>
<tr>
<td>Name</td>
<td>Definition</td>
</tr>
<tr>
<td>DisplayDevice</td>
<td>Describes if the display is a monoscopic display or stereoscopic display. The stereoscopic display device includes both autostereoscopic and non-autostereoscopic display devices</td>
</tr>
<tr>
<td>RenderingFormat</td>
<td>Describes the type of rendering formats the display is capable of including Interlaced, Progressive, Sync-Double, Page-Flipping, Anaglyph (Red-Blue, Red-Cyan, Yellow-Blue, and Red-Green).</td>
</tr>
</tbody>
</table>
the capability of 3D stereoscopic viewing. It should be either monoscopic or 3D stereoscopic display device. The monoscopic device means the monitor without 3D viewing capability. For the monoscopic device, 3D stereoscopic video needs to be converted to 2D video that its current display capability supports. The user may have one of the following 3D Rendering Format: Interlaced, Sync-Double, Page-Flipping and Anaglyph, where the Anaglyph is composed of Red-Blue, Red-Cyan, Red-Green and Yellow-Blue. The encoded video needs to be adapted to one of such rendering formats. By Interlacing we mean orderly presentation; first the odd scan-line 1,3,5,7,... is presented followed by the even scan-line 2,4,6,8. Page-flipping means alternately showing the left and right eye images on the screen. Sync-Double arranges the left and right eye images up and down. This is sometimes called top/bottom format.

5. Experiments

In this section, we present the experimental results being composed of two experiments, whose purpose is to verify the feasibility of our proposed DIA descriptions in terms of a resource conversion from 2D video to 3D stereoscopic video as well as from 3D stereoscopic video to 2D video. In the 2D/3D video conversion, we use MPEG elementary stream (ES) test sequences, which are Flower and Garden, Mobile and Calendar, and Fun. Their motion types are camera-right, camera-left and object-left/object right, respectively. In our experiments, the selection of an appropriate motion type is manually made. As well, a stereoscopic video sequence is tested for 3D/2D video conversion.

We developed two software modules which are 2D/3D video conversion and 3D/2D video conversion. Fig. 10 shows the interlaced images of MPEG Fun sequence, where a merry-go-round is moving to the right. Therefore, the motion type is object-right, \( \odot \) of Table 1. According to Table 1, a current image and a previous image are viewed to the left and right eyes, respectively. Since the image scene has negative parallax, it is perceived between the monitor screen and their eyes. It is observed that as the MaxDelayedFrame increases from one to four, the disparity between the left and right images accordingly increases as observed, thus resulting in the greater 3D depth [6].

Fig. 11 shows the MPEG Flower and Garden sequence, where the camera is moving to the right. Depending upon the selection of negative or positive parallaxes, the scene will be viewed in front of the monitor screen or behind the screen.

Fig. 12 illustrates the effect of DepthRange. The first test sequence is MPEG Fun. The DepthRange is varied from low to high. It is observed that the DepthRange increases in negative parallax. The second test sequence is MPEG Mobile and Calendar being viewed in positive parallax. Similarly, it is shown that the range of depth increases as the DepthRange is varied from low to high.

Fig. 11. Interlaced Images of MPEG Flower and Garden sequence. (Top) negative parallax and (Bottom) positive parallax.
Fig. 12. (Top) MPEG fun test sequence with varying DepthRange. (Bottom) MPEG Mobile and Calendar test sequence with varying DepthRange. The type is positive parallax.
For LeftRightInterVideo, the stereoscopic test video is used. Fig. 13 shows the stereoscopic test video, the left video, and the right video. Although the generation of the intermediate video was not carried out, any sophisticated methods can generate such intermediate video.

The verification for StereoscopicVideoDisplayDS is not necessary because DisplayDevice and RenderingFormat are hardware systems already available in the consumer markets.

6. Conclusion and future works

In this paper, we have presented StereoscopicVideoConversionDS as well as StereoscopicVideoDisplayDS. In the former, From2DTo3Dstereoscopic has three elements and From3DStereoscopicTo2D has a single element. To verify our DIA description tool, we have performed the experiments and presented our experimental results verifying its feasibility suitable to MPEG-21 DIA framework. Based upon the results, we have proved that our proposed StereoscopicVideoConversionDS and StereoscopicVideoDisplayDS, and all their elements correctly work. The proposed descriptions are useful for representing the users’ 3D stereoscopic adaptation as well as providing them with the efficient 3D stereoscopic viewing.

Our future works are composed of two parts. One is to develop more descriptions related to users’ Presentation Preferences to achieve the improved 3D depth perception. The other is to propose the type of display devices in more detail because advanced 3D devices are currently being developed.

References


Fig. 13. Stereoscopic test sequence: Original interlaced stereoscopic video (Top), left video (Middle), and right video (Bottom) are shown.


