

# Television in 3-D: What Are the Prospects?

BY LEVENT ONURAL, *Senior Member IEEE*

*3DTV Project Coordinator, Professor*

*Bilkent University*



(Graphic artist: Erdem Yucel)

High-quality 3-D video is regarded by the general public as the ultimate viewing experience. The objective is well-understood and has been depicted in many popular fiction movies: the target is a magical and somewhat mysterious optical replica of an object that is visually indistinguishable from its original (except perhaps in size). Such 3-D video technology will have many applications; and more than that, it will have a significant social impact by deeply affecting our daily lives. Although the goal is clear, and its impact is somewhat predictable, there is still a long way to go before we will have widespread commercial high-quality 3-D products. However, researchers are working with increasing interest on various elements of 3-D video technologies.

## I. HISTORICAL PERSPECTIVE

It is instructive to note that 3-D counterparts of photography, cinema, and TV are not new: indeed, stereoscopic 3-D versions of these common visual media are almost as old as their 2-D counterparts. Stereoscopic 3-D pictures were

shown in 1838, and 3-D photography was invented in 1839. The first 3-D cinema was available in the early 1900s. Various forms of early 2-D TV were developed in the 1920s, and by 1929 the stereoscopic 3DTV was also known.

However, although the 2-D versions of photography, cinema, and TV have flourished, the 3-D counterparts have almost disappeared since their peak around 1950. The reason for this failure is primarily because of the choice of technology: stereography. However, this was not quite a choice, but rather a technically imposed necessity. Stereoscopic 3-D video is based primarily on human perception, and it is quite easy to create. Two parallel classical 2-D video streams are produced by a pair of cameras mimicking the two human eyes which see the environment from two slightly different angles. One stream is shown to the left eye, and the other one to the right eye. Common means of separation of the right-eye and left-eye views are goggles with colored or polarization filters. Although technologically quite simple, stereoscopy causes eye fatigue within minutes after the viewing starts. This is a feeling similar to motion sickness, experienced by the viewer, and caused by a mismatch of various perception cues received by the brain from different sensory signals. Recently, with the adoption of digital technologies in all aspects of motion picture production, some of the technology-induced factors which result in eye fatigue

have been eliminated. As a consequence, it seems that stereoscopic 3-D movies will be the norm in the neighborhood movie theaters in a couple of years. However, some intrinsic eye fatigue factors are naturally still present, and will always be there as a consequence of using stereoscopy as the chosen 3-D technology.

## II. STEREOSCOPIC 3-D DISPLAYS

Stereoscopic 3-D displays are similar to known 2-D displays: a vertical screen or a monitor provides the two videos simultaneously. In contrast, autostereoscopic monitors are novel display devices. No special goggles are required to watch stereoscopic 3-D if autostereoscopic displays are used. Covering the surface of a regular, preferably high-resolution, digital video display device by a vertical or slanted lenticular sheet, and driving these monitors by so-called interzigged video, one can deliver the two different scenes to the left and the right eyes of the viewer, provided that the viewer stays in the correct position. Head-tracking autostereoscopic display devices have been developed to avoid the viewer position constraint. Serving many users at the same time is still a challenge. Barrier technology is another way of achieving autostereoscopy: electronically generated fence-like optical barriers coupled with properly interzigged digital pictures generate the two different views required. It is possible to provide many more views than the two views of the classical stereoscopy, using the autostereoscopic approach with slanted lenticular sheets. Up to nine views are common, creating a horizontal parallax with a viewing angle of about  $20^\circ$ . Classical stereoscopy with two views does not yield parallax as the head is moved.

## III. HOLOGRAPHIC 3-D DISPLAYS

Stereoscopic 3-D, whether more conventional as in stereoscopic cinema,

or more modern as in autostereoscopic monitor viewing, will never yield the ultimate experience as outlined at the beginning. Instead, such a desirable result can only be achieved if the technology is based not on perception, but on physical principles. Contrary to popular belief, all optically sensitive devices, including cameras and our eyes, do not really see the environment or the objects in those environments. Instead, what we see is only the light which enters the pupils of our eyes. This light carries the information about the scene. It is processed by the visual system and the brain, and thus we perceive our environment. Therefore, if the light which fills a given 3-D space can be recorded with all its physical properties, and then recreated from the recording to fill the space in the absence of the original objects, any optical device embedded in this re-created light will see the original scene; the original space-filling light and the re-created light will be physically indistinguishable. If the same light enters our pupils, we see the same thing. This is the basic principle of holography, which is a technique known since 1948. As expected, as the fidelity of the recorded and re-created light increases, the success of the created ghostlike objects and environments will also increase. Holographic cinema, digital holography, and holographic 3DTV are still in their infancy. However, there is no doubt that the obstacles will be eliminated and the ultimate 3-D viewing will become common as the underlying optics and electronics techniques mature.

Given the basic principle of holographic recording and reconstruction as above, the technique called integral imaging can also be classified as a form of holography. Integral photography has been known since 1905. The basic principle is to record the incidence angle distribution of incoming light at every point of recording, and then regenerate the same angular illumination distribution by proper back projection. The same effect is

achieved in classical holography by recording the phase and amplitude information simultaneously, instead of the intensity-only recording in conventional photography. The phase information is recorded using interference, and therefore, holographic recordings require coherent light (lasers). Intensity recording, such as that used in common optical film or digital photography, loses the direction distribution information.

## IV. CAPTURE AND TRANSMISSION OF 3-D CONTENT

We have so far primarily discussed the display end of the 3DTV technologies. Surely, an end-to-end operation requires not only the display, but also the capture and the transmission of the 3-D content. Some techniques have already been implied as we discussed the displays. For example, stereoscopic 3DTV requires a stereoscopic camera, which is nothing but two cameras rigidly mounted side by side. Recording in integral imaging is achieved using microlens arrays, whereas the recording in holography uses coherent light and interference. However, current research in 3DTV is targeting a much more general scenario.

In futuristic 3DTV systems, it is envisioned that the techniques for 3-D capture and 3-D display will be totally decoupled from each other. Rather, a capture device will provide enough information for the computerized synthesis of the 3-D scene. The synthesized scene will be similar to the 3-D graphics commonly used today in computer animations. However, instead of synthetic data, the 3-D information will be created from the real-life scene. Many techniques exist today for the capture of the 3-D scenes. A common technique is based on shooting the scene simultaneously from different angles using many conventional 2-D cameras. There are arrays with 128 cameras discussed in the literature. However, acceptable quality 3-D can be captured by using a

much lower number of cameras, especially if the scene is not too complex.

The synthesized 3-D video, created from the data provided by the capture unit, can then be either transmitted or stored. Luckily, 3-D scenes carry much less information than one may initially think, and therefore, such a 3-D video is highly compressible. To that end, special purpose compression techniques are needed. Such techniques are reported in the literature and the research is continuing. Transmission of such data is not too different than transmission of classical video. For example, video streaming techniques which are commonly used over the internet can easily be adapted to the 3-D case. However, such adaptation requires care as the usability of incomplete 3-D video data is totally different than the usability of incomplete 2-D video and packet losses are common in video streaming.

None of the displays can immediately show the 3-D video if the received data is abstract 3-D as explained above. Therefore, a translation of the received 3-D video data to the required driving signals of a specific 3-D display device is needed, and this can be quite difficult in some cases. Signal processing techniques are being utilized and developed for this purpose. Such a decoupled approach will provide complete interoperability by enabling the display of the content on totally different display devices with different technologies and capabilities. For example, it may be possible to feed the same video stream to a high-end holographic display device, a low-end stereoscopic 3-D monitor, or even a regular 2-D monitor. Each display device will receive the same content, but will have a different signal processing interface for the necessary data conversion.

In the near future, however, it looks as if multiview video will be the

common mode of 3DTV delivery. A large amount of 2-D video data, captured in parallel from an array of cameras shooting the same scene from different angles, will be directly coded by exploiting the redundancy of data, and then streamed to the receiver. The display at the receiving end will then create the autostereoscopic 3-D scene from the received data. Standardization activities for such a 3DTV operation are well underway under the ISO MPEG and ITU umbrella.

Applications of 3-D video techniques are endless. In addition to the obvious TV applications, there are countless others. Applications to gaming, defense, medicine, art, science, transportation, and culture are just a few. ■

### Acknowledgment

The 3DTV Project is funded by the European Commission within FP6 under Grant 511568.