With an Eye on the Third Dimension

The MEDIAGROUP is breaking new ground in 3-D video

Perception of three-dimensionality, in other words stereoscopic vision, does not arise in the eye but in the brain. Thus, it is sufficient to provide appropriate signals for the primary visual cortex in order to create a perception of 3-D for the viewer. Today such signals can be computed from almost any original image. A concept, which a few years ago would have been pure utopia, has become the latest MEDIAGROUP research project: the three-dimensional representation of conventional video images

Spatial reproduction of sound recordings has a long tradition. In 1940 the first movie with an early Version of Stereo sound was running. Various subsequent attempts were made to improve spatial reproduction which led finally to 5.1 surround sound. Meanwhile, some projects have included additional height information - we have truly arrived in the world of virtual 3-D-acoustics! In contrast to sound, reproducing three-dimensional video is much more difficult. Development is a long way from being as advanced as for sound even though the first screening of a 3-D movie also stretches back to 1940. While over the decades high-quality 3-D movies with elaborate effects have established a niche in the cinema, acceptance of 3-D TV has so far remained behind expectations. Putting on 3-D glasses in a cinema is just different from sitting in front of the expensive new television without moving and whilst wearing special glasses. Even then it only demonstrates its advantages at a predetermined distance and in the rare event of a 3-D TV broadcast!

Autostereoscopy as the way out

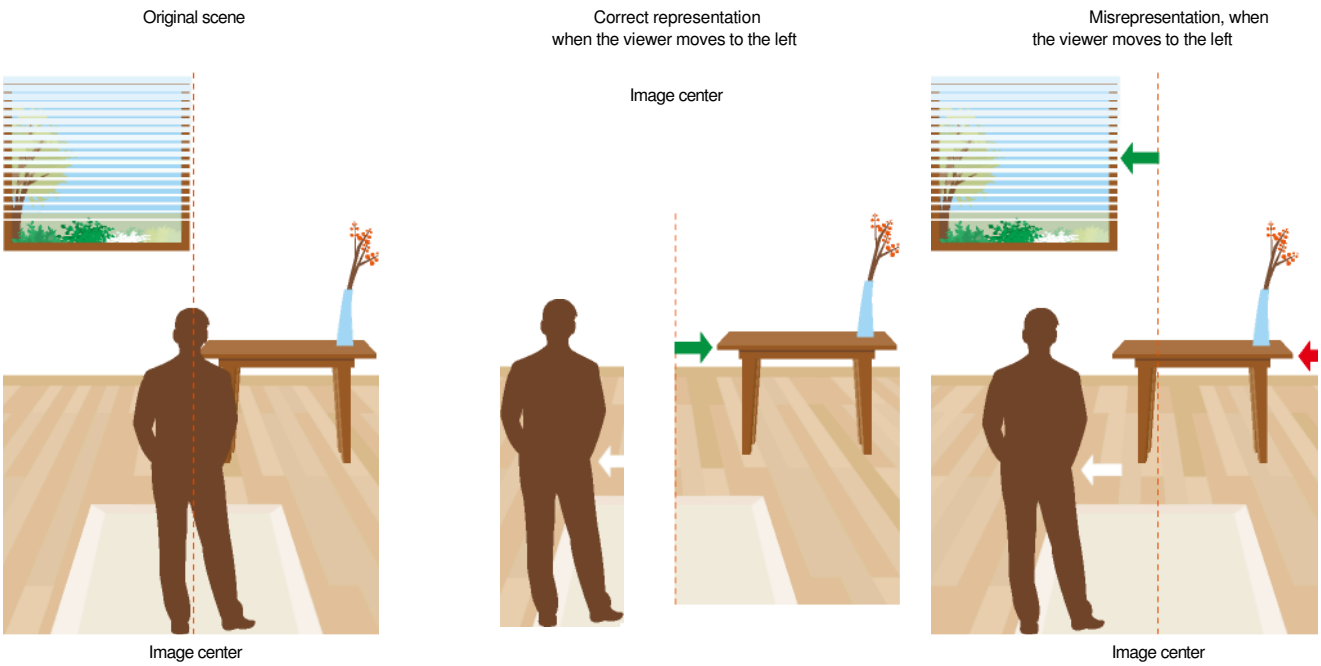
In order to help 3-D technology achieve general acceptance in the living room, as a first step the annoying but still necessary glasses have to disappear. Meanwhile, a few auto-stereoscopic technologies are available which do not require glasses. These are mainly Systems with a barrier filter in front of, or a lenticular film directly on, the screen surface. Their common principle is to draw each of the spectator's eyes to different pixels. This method works, but it is always active. It can not be switched off, not even when viewing 2-D content. Thus, 2-D content will look worse than on a Standard HD screen, with darkened Images and the risk of illegible subtitles - not the best preconditions for establishing a new technology. The problem would be defused if a method of presenting 2-D Images in good quality on a 3-D screen could be developed. Exactly such a procedure was found by the MEDIAGROUP in 2010. And the iVu algorithm by Natural View Systems, the youngest member of the MEDIAGROUP, can do even more: It computes a spatial Image from every video frame, as if it were a natural 3-D Image.

Stereoscopic vision

Any viewer recognizes immediately whether a person in shot is standing in the foreground or in the background. This is valid even for bad black and white originals, such as half-tone newspaper Images. Spatial arrangements are not determined solely from size differences. Perspective lines, exposure, shadows, changes in colour - all these are part of our unconscious spatial perception. In this process the eyes serve as sensors delivering Signals which the brain processes into a 3-D Image. In a manner comparable with psycho-acoustics in audio, several masking and reinforcing effects are involved. With the help of a suitable Image analysis that takes the rules of perception into account, let's call it psycho-optic evaluation in order to keep the analogy to psycho-acoustics, a 3-D representation can be generated from any Image, even from the worst newspaper picture. To display it on a TV screen, each individual pixel must be distributed at least twice, spatially separated and colour differentiated. The iVu algorithm performs exactly this Image analysis and processing.

Psycho-optic challenge

For deployment in domestic television sets the analysis and processing must be fast. Therefore, each pixel is divided into its basic colour constituents (RGB), which are displayed on screen as adjacent sub-pixels. At the full-HD resolution of 1,920 x 1,080 pixels, this results in six million sub-pixels to be calculated up to sixty times per second. This only works with fast processors and a very effective algorithm. However, it must also work very precisely in order to avoid the common problems of 3-D representations. One very disruptive effect is spatial misrepresentation which occurs if the viewer moves.



In reality it can be observed, that the furthest visible background moves with the viewer, for example the rear wall of a hall. If the viewer moves one step to the left while looking at the wall, it appears, as if the wall would also move to the left with the viewer.

In contrast, a table in the foreground between the viewer and the wall moves right, i.e. opposite to the viewer's direction of movement. With an incorrect 3-D representation both foreground and background move in the same direction.

In this case, foreground and background are shifted unnaturally against each other. If the image analysis fails during 3-D conversion then, for example, the foreground does not move in the opposite direction to the viewer as it should, but moves in the same direction instead. Another miscalculation manifests itself in the misinterpretation of an object's depth level. For example, individual bricks in an absolutely vertical wall may be interpreted differently. In the processed version this can lead to the impression that some bricks are protruding. It took more than ten years of research and development of the iVu algorithm to avoid all these misinterpretations. Today it presents a natural-looking video image which remains stable even as a freeze frame.

Free focusing?

The effectiveness of a video produced in real 3-D is highly dependent on the shooting situation. In order to achieve a natural 3-D image, the two cameras required should be mounted rigidly at the interpupillary distance. In reality the separation is often increased by a certain angle which leads to spectacular effects. However, the viewer is compelled to focus on exactly the same point as the camera. Contrary to our natural 3-D perception, the viewer no longer has the choice of focussing on the background, foreground or in the middle. In the long term viewing such 3-D material results in irritations and eyestrain. If the cameras are moved too far apart and the object in focus is too far forward, out of the screen, then double images develop in the background during replay. An object in the background is represented once to the left of the foreground object and once again to the right. Here in the relatively new field of 3-D production the necessary experience to produce aesthetically pleasing and compatible images consistently is lacking.

Michael Steglich, developer and founder of Natural View Systems GmbH, and Rainer Weiß, managing director of Natural View Systems GmbH, during our discussion on the basic principles of 3-D television and the iVu algorithm.

One for all

None of these problems occurs if iVu computes the 3-D image from a 2-D video. On the contrary, the TV experience is more relaxed, since the viewers have the impression that they can focus on different pixels at different depth levels. The effect is not limited to the best seat centred in front of the screen, but works in a comparatively broad range of approximately 160° from the screen. Even better, the iVu system dispenses with the biggest disadvantages of 3-D TV in the living room: The 3-D glasses otherwise necessary and the lack of 3-D content! Watching lavish 3-D movies with spectacular out of screen effects and enjoying unobtrusively vivid and eye-friendly 2-D content as well will be possible with the same hardware in the near future. And those who really want to can still watch TV in classic 2-D mode. But who listens to one loudspeaker when a stereo system is playing mono?

Read more on www.stagetec.com/stages