Light field technology enables production options

Light field representation

Conventional cameras capture a 2D representation of a scene; they achieve very high spatial sampling (recording the light intensity at each scene point sampled by the sensor). This provides a photorealistic representation of the scene but all the creative decisions, such as shot composition, focus, depth of field, and so on, must be made at the point of capture and are baked into every frame. It is not possible to alter these elements post-capture without going through a tedious post-production process and without risking degrading the realism of the shot. Besides, composition with computer-generated (CG) assets is not straightforward as conventional images are 2D and lack depth information.

In contrast, a light field is a 4D representation describing light rays travelling through space at all locations and in all directions (Levoy & Hanrahan, 1996; Gortler et al., 1996). A light field captures both the spatial properties of the scene (the light intensity emitted by each scene point sampled by the camera, akin to conventional camera) and the directional properties of the light emitted by the scene (the light intensity as a function of the direction in which light is reflected off surfaces in the scene). This richer representation provides added flexibility and creative control as some of the creative decisions that usually have to be made at the point of capture can now be postponed to post-production – without compromising on realism. For example, if the director may decide to recompose the shot, alter the camera path or change the aperture and focus in post.

This new representation is also ideally tailored to immersive content such as virtual reality, where the user can move and refocus content (Google Jump, GoPro Odyssey). The scene is captured using multiple cameras and stereoscopic content is produced in post by texture mapping the captured images onto a 3D model of the scene. These approaches used either a camera mounted on a gantry (Levoy & Hanrahan, 1996) or more recently a handheld camera (Davis, Levoy & Durand, 2012). In these approaches, each frame effectively captures a 2D slice of the 4D light field. Due to the need to physically move the camera around the scene, these approaches are limited to static scenes. They use image-based rendering techniques to synthesize new views by interpolating the information from the captured views. As they do not explicitly capture the scene geometry, a large number of camera views (typically around 100) is required in order to densely sample the light field.

Multi-view stereo capture methods extend light field acquisition to the video domain by using a network of synchronised cameras distributed around the scene to simultaneously acquire 2D slices of the light field from multiple viewpoints. A 3D representation is then extracted using multi-view stereo reconstruction methods (Dittrich et al., 2004; Starck & Hilton, 2007) and the 3D model is used as a geometric proxy to optimise blending of the different input views for photorealism (Bruehler et al., 2001).

Initial multi-camera light field setups were based on dense multi-camera arrays (i.e., a large number of cameras in close arrangement, as in Wilburn et al., 2005). Recent research has explored the use of wide-baseline acquisition setups to reduce the number of required cameras (Starck & Hilton, 2007; Guillenmaut & Hilton, 2011) and enable use in outdoor settings (Battani et al., 2010; Kim et al., 2012, Mustafa et al., 2011). These methods are more practical than their dense camera array counterparts but do not model all the viewpoint-dependent detail and there are limitations to the viewpoints that can be synthesised due to the sparse acquisition setup.

With the advent of computational photography and light field cameras (Kg et al., 2005), it has become possible to capture a light field using a single sensor. Light field cameras (e.g., Raytrix, Lytro) insert a lenslet array between the main lens and the camera sensor to multiplex the 4D light field onto a 2D image sensor. Each micro lens separates the bundle of incoming light rays depending on their incoming directions, thus capturing the directional properties of incoming light at this particular location in the array. The complete lenslet array enables formation of an image which captures both spatial and directional information of the light field.

These cameras effectively trade some of the spatial resolution against angular resolution. Consequently, they require significantly more pixels than a conventional camera to achieve an equivalent final image resolution. Light field cameras are now being extended to compact stereoscopic 360° systems capable of producing photorealistic novel views and refocussing content (Google Jump, GoPro Odyssey). However, the viewpoint remains limited to a relatively small volume, only suitable for limited head motion. Very recently, Lytro has introduced the first professional cinematic light field camera called the Lytro Cinema, which was premiered at NAB 2016. They are also introducing the Lytro Immerge which is dedicated to cinematic and virtual reality (VR) production.

A major advance enabled by light field capture is the ability to postpone the need to make hard decisions on shot composition, aperture and focus from capture time to render time. This removes the burden of having to get every element of the shot perfect at the time of capture. For example, it becomes unnecessary to perfectly track and maintain focus on small and rapidly moving objects since it is possible to refocus the shot at render time.

Some examples of visual effects that are enabled by this technology are:

- The light field representation enables extraction of depth information to facilitate visual effects. Depth extraction is made possible by the richer representation which samples the scene from a large number of viewpoints defined by the main lens aperture (in the case of a light field camera) or the different cameras in the setup (in the case of a camera array). The availability of depth as well as the multi-view appearance information allows the synthesis of new viewpoints by image-based rendering or texture mapping using the capture images, a process called novel view synthesis or free-viewpoint video.

A light field is a 4D representation describing light rays travelling through space at all locations and in all directions.
Current workarounds include filming actors on green screen with two- or four-camera rigs positioned in front of, or to one side of, the performer. However, the illusion breaks down when the user moves within the VR scene, revealing the secret of the trick—that the video is playing back on a 2D plane or a crude 3D mesh, which only looks correct from a restricted range of perspectives.

In contrast, light field technology directly captures a 4D representation and therefore naturally lends itself to immersive content production as it becomes possible to adjust the rendered content as a function of the user’s position and orientation in the scene, while preserving the realism of the captured scene. Commericially available ‘outside-in’ camera arrays (i.e. with cameras surrounding the scene and pointing inward) are capable of producing content elements that can be viewed from multiple viewpoints, but they can only reproduce the surrounding environment from one point of view. Outside-in light fields are often recorded in a chromakey studio to isolate the actors, or props, which can then be composited with a separately captured real background or CG background.

‘inside-out’ camera arrays, such as commercially available compact stereoscopic 360° camera rigs, offer a limited range of potential viewpoints. More research is required to develop light field capture systems to support a larger range of viewpoints.

Light field challenges

Light fields have the potential to revolutionise the way we create and consume content, however, they also introduce a new set of challenges. Firstly, the light field representation is significantly larger than conventional image/video representations due to the increased dimensionality. This poses some challenges in terms of data storage and transmission. An important aspect to ensure accuracy of the technology will therefore be standardisation as well as the development of efficient compression techniques tailored specifically to light field data.

An additional challenge relates to the development of creative tools dedicated to the light field editing process. There is currently a lack of software tools for light field post-production. Finally, giving the user the ability to freely explore and move in a scene introduces new challenges in terms of storytelling and narratives for VR content production.

Increasing creative freedom and flexibility

Whether based on multiple camera setups or microarray technology, light fields now offer content-makers a whole new degree of freedom and flexibility to create visual effects.

References


