Coded Exposure HDR Light-Field Video Recording

David C. Schedl, Clemens Birklbauer, and Oliver Bimber*
Johannes Kepler University Linz
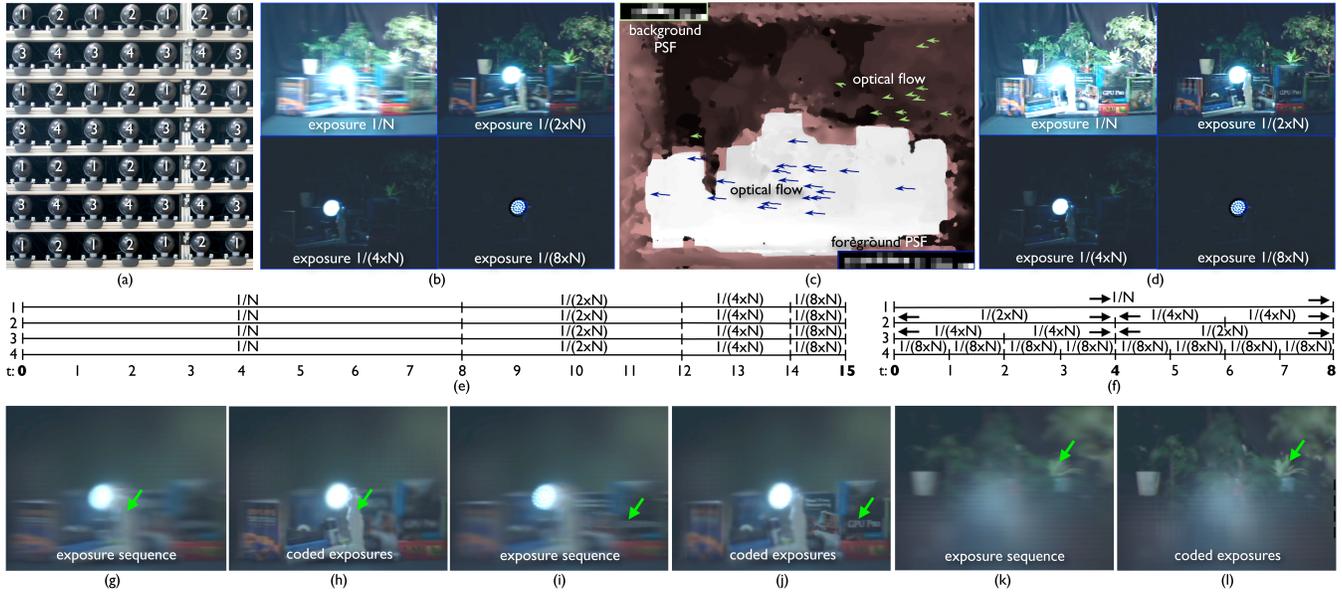
Figure 1: *HDR light-field recording during camera movement: Instead of recording complete exposure sequences per perspective of a camera array (e), we encode exposure times in each repeating camera quad-tuple (a,f). By computing a composite depth map from all exposures, segmenting it into depth layers, and tracking the optical flow of scene features in each depth layer (c), we derive depth-individual PSFs that are used for local motion deblurring (b,d). Finally, we interpolate deblurred exposure images that have not been recorded from particular perspectives. Combining them results in HDR images for each camera perspective. These images can then be used for tone-mapped light-field rendering (g-l). In contrast to recording regular exposure sequences (g,i,k), our approach (h,j,l) reduces motion blur. Figures g-l compare tone-mapped wide-aperture (i.e., shallow depth-of-field) light-field renderings for different focus settings (indicated with the green arrows). Note, that remaining artifacts in out-of-focus regions are due to directional under-sampling (i.e., too few cameras in the array).*

Our Approach

Capturing exposure sequences for computing high-dynamic range (HDR) images causes motion blur in case of camera movements. This also applies for light-field cameras, such as camera arrays. Images composed from multiple blurred HDR light-field perspectives are also blurred.

To reduce motion blur, we encode four exposure times in each repeating camera quad-tuple of a camera array (assuming relatively small camera-baselines): Instead of recording all exposures sequentially for all perspective cameras (cf. fig. 1e), we apply the spatio-temporal exposure pattern shown in figures 1a,f. This reduces the overall recording interval, but also prevents from recording all exposures at all perspectives, and therefore constrains the smallest synthetic aperture to be at least 2×2. Within each camera quad-tuple, the shortest and longest exposures are captured only from one perspective each (1 and 4, respectively), while the two medium exposures are recorded interleaved from two perspectives (2 and 3).

For each of the four different exposures (cf. fig. 1b) that are recorded at varying perspectives, we compute a depth map. Since these depth maps vary locally in quality (due to motion blur for the higher exposures, and due to low SNR for the lower exposures), we compile them to a single composite depth map based on a depth confidence criteria (cf. fig. 1c). This composite depth map is then segmented into different depth layers. For each depth layer we track

features in the low exposure images. The optical flow of these features allows us to determine the local point-spread function (PSF) that causes motion blur in higher exposures. These PSFs are used for motion deblurring (i.e., deconvolution) of each recorded exposure image at all camera perspectives (cf. fig. 1d).

By shifting the PSFs before deconvolution (as illustrated by the arrows in figure 1f), we can receive three subframes for each recording interval at all perspectives of each camera quad-tuple. Thus, for subsequent recording intervals of an HDR light-field video (coded exposures and deconvolution shifts as illustrated in figure 1f), we obtain the following exposure images at times $0, 4, 8$: $1/N$ for perspective 1, $1/(2 \times N)$ and $1/(4 \times N)$ for perspectives 2 and 3, and $1/(8 \times N)$ for perspective 4. Compared to classical exposure sequencing (cf. fig. 1e), this leads to a $\times 3.75$ higher frame rate. With the recorded and deblurred exposure images we finally compute an enhanced composite depth map to interpolate deblurred exposure images for all camera perspectives that have not been directly recorded.

Figures 1g-l illustrate tone-mapped, wide-synthetic-aperture (i.e., shallow depth-of-field) images rendered from a light field that was recorded during camera motion. Figures 1g,i,k show the results at different focus settings (green arrows indicate selected in-focus region) with regular exposure sequences for each camera perspective (cf. fig. 1e). Figures 1h,j,l show the same images recorded with our coded exposures (cf. fig. 1f) and computed with the technique outlined above.

*firstname.lastname@jku.at