



<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper presented at *Data Compression Conference (DCC 2018), Snowbird, Utah, US, March 27 - March 30, 2018.*

Citation for the original published paper:

Ahmad, W., Vagharshakyan, S., Sjöström, M., Gotchev, A., Bregovic, R. et al. (2018)
Shearlet Transform Based Prediction Scheme for Light Field Compression

In:

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:miun:diva-33356>

Shearlet Transform Based Prediction Scheme for Light Field Compression

Waqas Ahmad^{*}, Suren Vagharshakyan⁺, Mårten Sjöström^{*}, Atanas Gotchev⁺,
Robert Bregovic⁺, and Roger Olsson^{*}

^{*}Mid Sweden University
Holmgatan 10
Sundsvall, 85352, Sweden
{waqas.ahmad,marten.sjostrom}@miun.se

⁺Tampere University of Technology
Korkeakoulunkatu 10
Tampere, 33720, Finland
{suren.vagharshakyan,
atanas.gotchev}@tut.fi

Abstract: Light field acquisition technologies capture angular and spatial information of the scene. The spatial and angular information enables various post processing applications, e.g. 3D scene reconstruction, refocusing, synthetic aperture etc at the expense of an increased data size. In this paper, we present a novel prediction tool for compression of light field data acquired with multiple camera system. The captured light field (LF) can be described using two plane parametrization as, $L(u, v, s, t)$, where (u, v) represents each view image plane coordinates and (s, t) represents the coordinates of the capturing plane. In the proposed scheme, the captured LF is uniformly decimated by a factor d in both directions (in s and t coordinates), resulting in a sparse set of views also referred to as key views. The key views are converted into a pseudo video sequence and compressed using high efficiency video coding (HEVC). The shearlet transform based reconstruction approach, presented in [1], is used at the decoder side to predict the decimated views with the help of the key views.

Four LF images (Truck, Bunny from Stanford dataset, Set2 and Set9 from High Density Camera Array dataset) are used in the experiments. Input LF views are converted into a pseudo video sequence and compressed with HEVC to serve as anchor. Rate distortion analysis shows the average PSNR gain of 0.98 dB over the anchor scheme. Moreover, in low bit-rates, the compression efficiency of the proposed scheme is higher compared to the anchor and on the other hand the performance of the anchor is better in high bit-rates. Different compression response of the proposed and anchor scheme is a consequence of their utilization of input information. In the high bit-rate scenario, high quality residual information enables the anchor to achieve efficient compression. On the contrary, the shearlet transform relies on key views to predict the decimated views without incorporating residual information. Hence, it has inherit reconstruction error. In the low bit-rate scenario, the bit budget of the proposed compression scheme allows the encoder to achieve high quality for the key views. The HEVC anchor scheme distributes the same bit budget among all the input LF views that results in degradation of the overall visual quality. The sensitivity of human vision system toward compression artifacts in low-bit-rate cases favours the proposed compression scheme over the anchor scheme.

References

- [1] S. Vagharshakyan, R. Bregovic, and A. Gotchev, "Light field reconstruction using shearlet transform," IEEE transactions on pattern analysis and machine intelligence, 2017.

Acknowledgment: The work in this paper was funded from the European Unions Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 676401, European Training Network on Full Parallax Imaging.