



<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[†], Roger Olsson*
*Mid Sweden University, Sundsvall, Sweden [†]Tampere University of Technology, Tampere, Finland

Introduction

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. In this work, we present a novel prediction tool for light field compression. The compression scheme follows two stages:

- Selection and coding of key views.
- Reconstruction of non-key views.

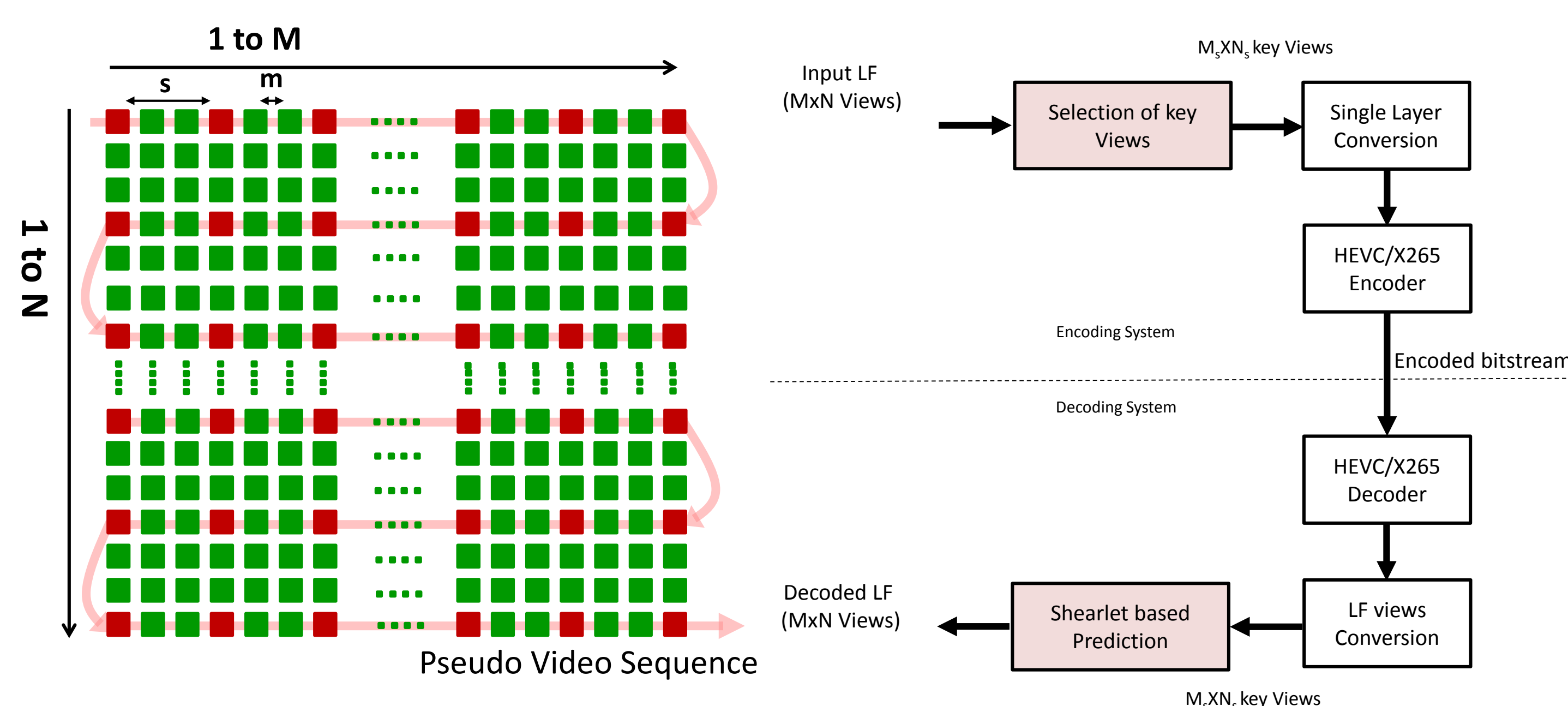


Figure 1: Proposed compression scheme, (a) key frame selection process (key frames in red) and (b) compression system.

Proposed compression scheme

The captured Light Field (LF) is uniformly decimated by factor s in both directions, resulting into a sparse set of $M_s \times N_s$ views. The sparse set of views are converted into a pseudo video sequence and compressed using High Efficiency Video Coding (HEVC). At the decoder side, non-key views (de-selected views) are predicted using shearlet transform. The directional sensitivity of shearlet transform allows to construct desirable frequency domain tiling. The non-key views reconstruction in 4D full parallax case is interpreted as multiple 3D horizontal and vertical parallax reconstructions. The Epipolar Plane Image (EPI) representation of each 3D parallax LF is used to reconstruct densely sampled LF from the sparse set of views.

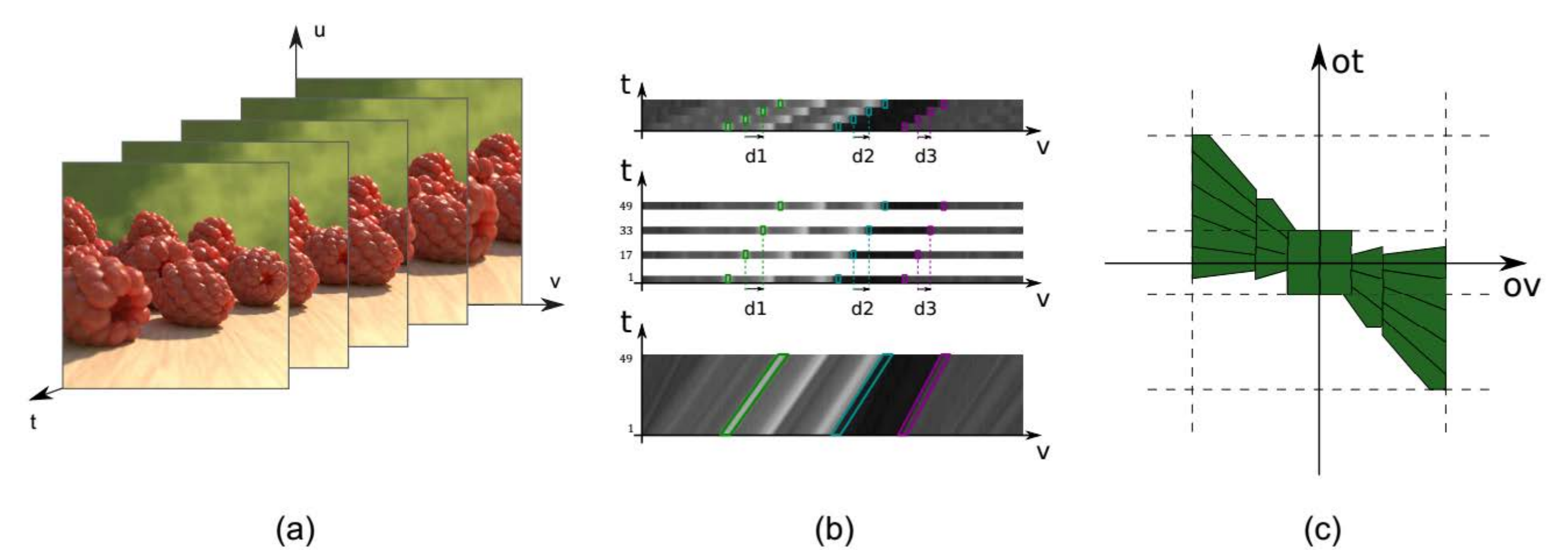


Figure 2: (a) Parameterization of captured LF images, (b) interpretation of input images as decimation of densely sampled light field, (c) frequency plane tiling using shearlet transform required for reconstruction of densely sampled light field.

Results and Discussion

The experimentation was performed on two light field data sets: Stanford dataset and High Density Camera Array (HDCA) dataset. The rate-distortion analysis for the proposed compression scheme shows significant compression efficiency in low bit-rate scenarios as compared to the anchor compression scheme. However, the anchor performs better in high bit-rates. The sensitivity of human vision system towards the compression artifacts in low bit-rates favours the proposed compression scheme over the anchor.

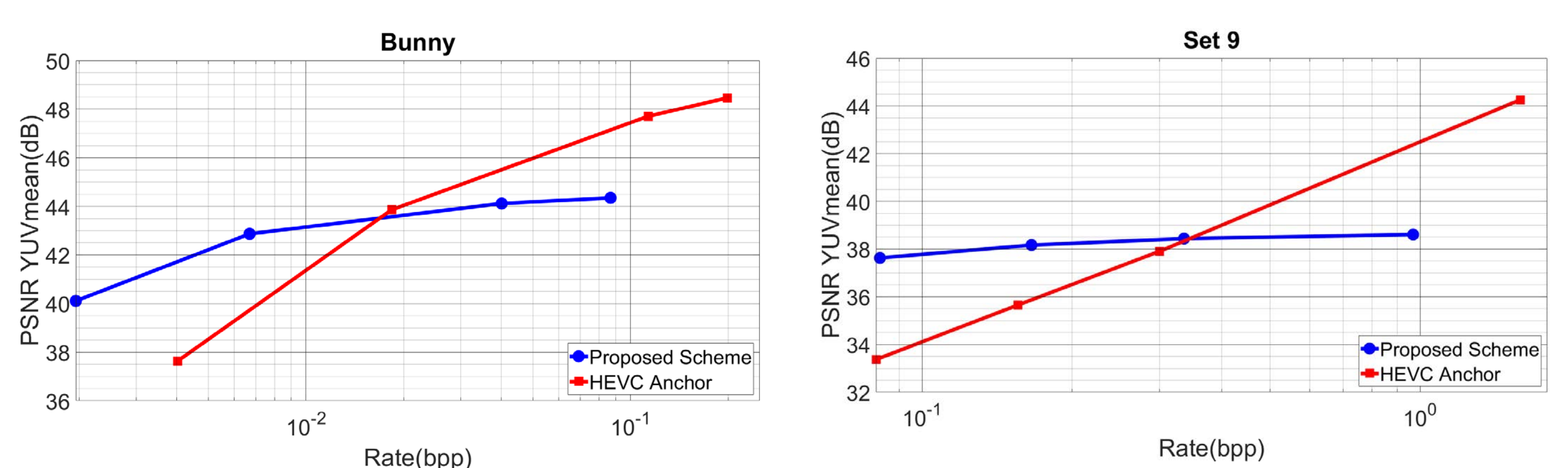


Figure 3: Rate Distortion analysis of proposed compression scheme with reference HEVC video compression standard on Stanford and HDCA LF images.

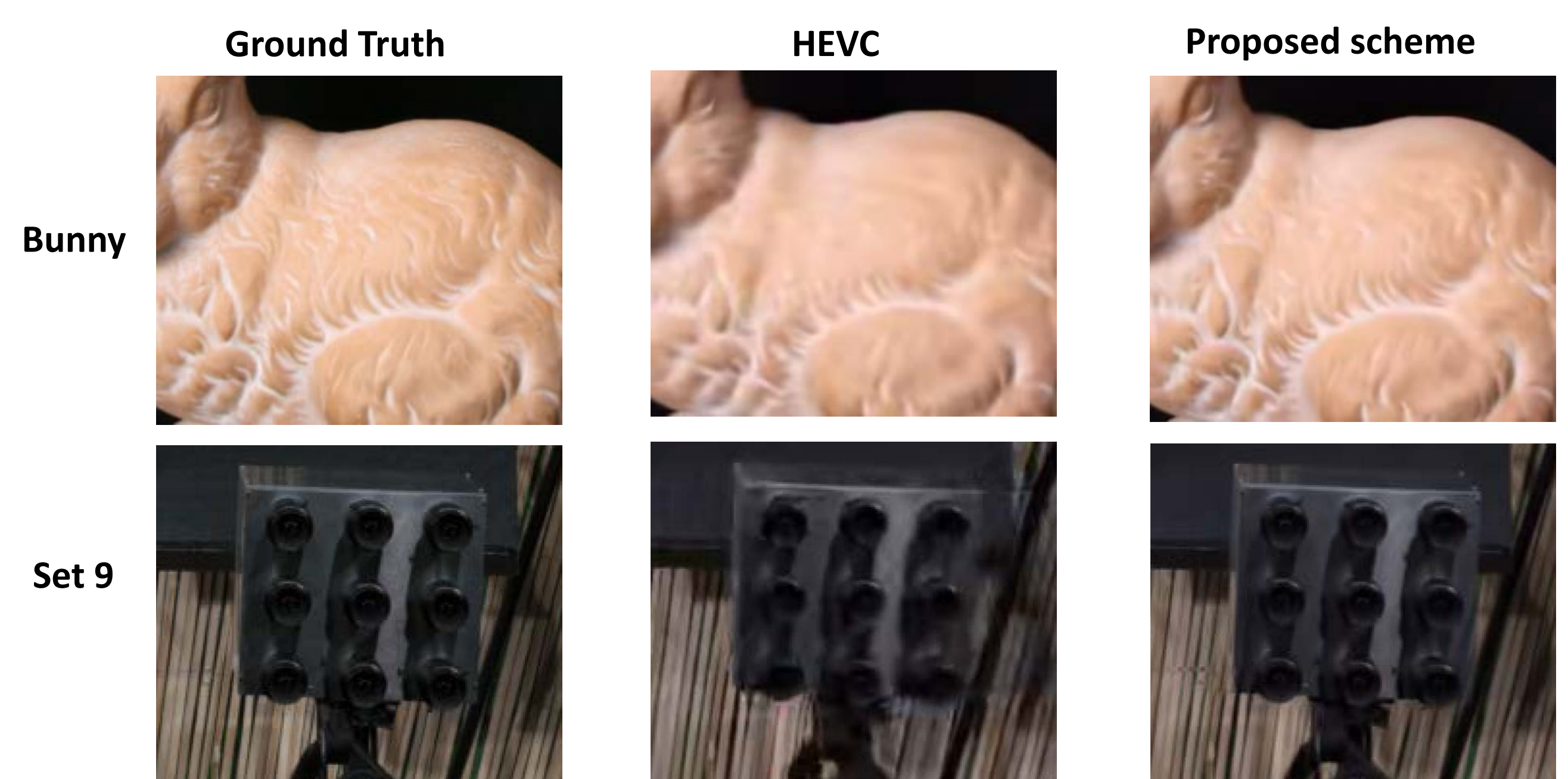


Figure 4: Subjective analysis of proposed compression scheme with reference HEVC video compression standard in low bit-rate scenario.