Beyond Flicker, Beyond Blur: View-coherent Metameric Light Fields for Foveated Display

Prithvi Kohli* UCL
David R. Walton† UCL
Rafael Kuffner dos Anjos‡ University of Leeds
Anthony Steed§ UCL
Tobias Ritschel¶ UCL

Abstract

Ventral metamers, pairs of images which may differ substantially in the periphery, but are perceptually identical, offer exciting new possibilities in foveated rendering and image compression, as well as offering insights into the human visual system. However, existing literature has mainly focused on creating metamers of static images. In this work, we develop a method for creating sequences of metameric frames, specifically light fields, with enforced consistency along the temporal, or angular, dimension. This greatly expands the potential applications for these metamers, and expanding metamers along the third dimension offers further new potential for compression.

Index Terms: Computing methodologies—Computer graphics—Graphics systems and interfaces—Perception; Computing methodologies—Computer graphics—Image compression

1 Introduction

We consider the application of light field [12, 22] streaming [18] as presumably required in future VR and AR display. However, transmitting the entire 4D light field would require a prohibitively large bandwidth [4,6,9,15]. Foveated rendering [17] and display [20] of light fields are ways to steer bandwidth to those areas attended by foveal vision and reduce bandwidth in the periphery. However, simply reducing spatial detail in the periphery leads to blur.

For single images, Walton et al. [21] have recently proposed a method based on ventral metamers [7] that shows details with the right statistics in the periphery without having to transmit the detailed signal. However, applying this method straight away to light fields either introduces flicker or a persistent noise pattern when moving between views, both of which degrade the experience [10, 11].

To remove these effects and achieve view coherence, we extend [21] by additionally compressing the angular dimension of the light field in such a way that we retain the perceptually important information in the periphery: the specific spatio-angular orientations (epipolar lines) of the light field [1].

2 Background

The method of Walton et al. [21] reduces image data in the periphery to the mean and variance of spatial features, localised in orientation and scale, over spatial pooling regions. These regions grow with distance from the fixation point of an observer. Noise is matched to these statistics to synthesise a metamer [21]. However, applying this method independently to each spatial slice of a light field using a 3D block of noise introduces large amounts of flickering. Using the same 2D noise for each spatial slice instead introduces the "shower door" effect, where the persistent noise pattern does not move with the flow of the scene [10].

[21] could be naïvely extended to 3D, matching statistics over pooling volumes in a 3D pyramid. However, this also introduces flickering since it still does not preserve the slant structure of the light field [1]. Alternatively, [21] could be applied to motion-coherent noise [10]. This resolves the "shower door" effect, but some flickering persists, as it is not possible to fully control the angular frequency content using this strictly 2D (spatial-only) metamer approach.

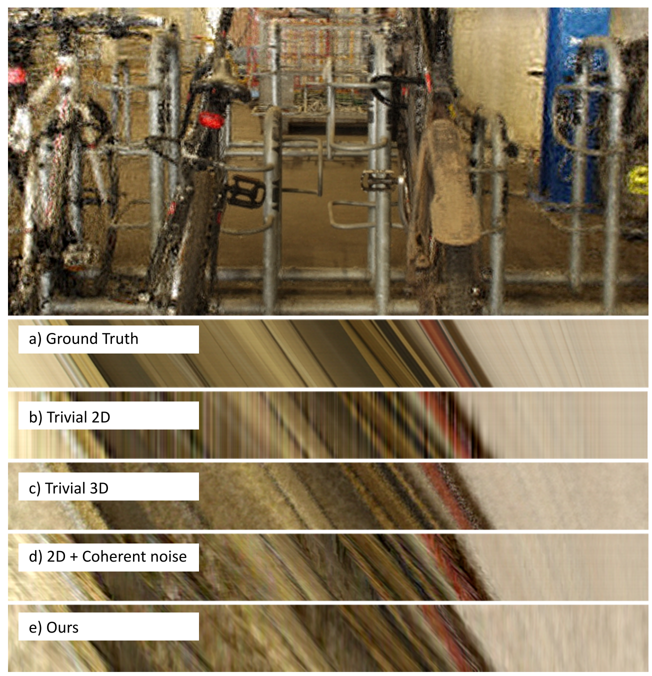

Figure 1: A spatial slice of a light field, metamerized using our approach (top). The four rows below show an epipolar slice through the same result by different techniques.

Instead, our method introduces noise, but follows motion and generates features that preserve both the spatial statistics of the input, as well as its epipolar structure, without any flicker.

3 Method

The aim of our method is to extract compact information from a light field on a server, and decode it to give a perceptually identical field on the client.

We consider a discrete light field parameterised by a 2D spatial coordinate and a 1D angular coordinate [2] in YCrCb colour space [21]. On the server, we first decompose the light field into different spatio-angular scales and orientations using a 3D steerable pyramid decomposition in the Fourier domain [5] with horizontal and vertical spatial orientations, and an angular orientation.

Second, we find motion as optical flow of the steerable pyramid bands, which is possible because these bands are the filtered partial

*email: prithvi.kohli.19@ucl.ac.uk
†email: david.walton.13@ucl.ac.uk
‡email: r.kuffnerdosanjos@leeds.ac.uk
§email: a.steed@ucl.ac.uk
¶email: t.ritschel@ucl.ac.uk

derivatives of the light field [13, 19]. The optical flow represents the spatio-angular orientations, or epipolar lines, of the light field [8].

Next, we pool statistics of the spatial bands of the steerable pyramid over 3D pooling regions in the spatio-angular domain. We compute mean and variance using a spatially-varying lowpass filter [21]. The spatial bandwidth of the filter falls off according to the receptive field sizes investigated by [7]. The angular bandwidth depends on the relation of the spatial and angular sampling density of the light field and falls off in a different rate [16].

To compress the angular bands of the steerable pyramid in a way that retains the spatio-angular orientations [1], we blur optical flow using the same lowpass filter [16]. This is because we can derive angular bands from the spatial bands via the optical flow constraint. As an optimisation, we only compute optical flow at the top level of the steerable pyramid and downsample this, rather than computing the flow at every level of the pyramid.

Pooling essentially removes information that is not perceived in the periphery, resulting in compression. To be clear, the dimensions of the steerable pyramid responses are larger than the input image. However, the pooling preserves only lower frequencies of the statistics maps, allowing for effective compression. By keeping second order statistics (variance) we avoid blurring the pyramid levels themselves. These pooled data are transmitted to the client.

On the client, we then generate pixel noise coherent with the blurred optical flow of the light field using the method of Kass and Pesare [10]. The coherent noise is decomposed into a 3D steerable pyramid and the spatial bands are matched to the pooled spatial statistics [21]. Using these matched bands and the blurred optical flows, we compute the angular band at every level of the pyramid via the optical flow constraint to match the spatio-angular orientations (epipolar lines) of the original light field. The final metameric light field is synthesized by reconstructing the pyramid.

4 Results

We test our method on light fields from the MPI Light Field Archive [2]. All light fields consist of 101 view samples with an image resolution of 960×720. We encourage readers to view the results in the accompanying video. Please note that the brightness fluctuations in the "bikes" light field are part of the reference. We note that our approach produces temporally coherent, flicker-free metamers without the "shower door" effect. These results show to our knowledge the first example of coherent light field metamers.

4.1 Compression

To demonstrate how our method can reduce the filesize of light fields, we implement a basic compression system based on [21]. We quantise and apply a cortical remapping to the statistics and optical flow maps, before saving them as JPEG images. We find that we achieve improved angular/temporal-coherence compared to the method of Walton et al. whilst storing slightly fewer values.

5 Limitations

Our current 3D pooling does not take into account abrupt changes in spatial features (e.g. disocclusion) or motion (e.g. fast-moving video). This may lead to perceptual mismatches in the periphery in these cases.

We do not use any kind of sophisticated model (e.g. [3, 11, 16]) to inform us of how to correctly pool both spatial features statistics and optical flow over angle/time such that we are able to accurately match the peripheral perception of motion and optimise compression.

6 Conclusion

We presented a method for producing metamers of 3D light fields which avoids both flicker and shower door. We implemented a compression method to show how these light fields could be efficiently transmitted in a server-client setting.

The current approach is focused on 3D light fields exclusively, but there is a possibility to create a similar approach for 4D light fields in the future, for example by establishing spatio-angular coherence via ray flow [14].

Acknowledgments

This work was funded in part by the EPSRC/UKRI project EP/T01346X/1.

References

[1] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, 2(2):284–299, Feb 1985.

[2] V. K. Adhikarla, M. Vinkler, D. Sumin, R. Mantiuk, K. Myszkowski, H.-P. Seidel, and P. Didyk. Towards a quality metric for dense light fields. In *CVPR*, 2017.

[3] K. Amano, M. Edwards, D. R. Badcock, and S. Nishida. Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *Journal of Vision*, 9(3):4–4, 03 2009. doi: 10.1167/9.3.4

[4] C.-L. Chang, X. Zhu, P. Ramanathan, and B. Girod. Light field compression using disparity-compensated lifting and shape adaptation. *IEEE Trans. Image Proc.*, 15(4):793–806, 2006.

[5] C. A. Delle Luche, F. Denis, and A. Baskurt. 3d steerable pyramid based on conic filters. In *Wavelet Applications in Industrial Processing*, vol. 5266, pp. 260–8, 2004.

[6] M. Flierl and B. Girod. Multiview video compression. *IEEE Signal Processing Magazine*, 24(6):66–76, 2007.

[7] J. Freeman and E. P. Simoncelli. Metamers of the ventral stream. *Nature Neuroscience*, 14(9):1195–1201, 2011.

[8] B. K. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1):185–203, 1981.

[9] X. Jiang, M. Le Pendu, R. A. Farrugia, and C. Guillemot. Light field compression with homography-based low-rank approximation. *IEEE J Selected Topics in Signal Processing*, 11(7):1132–45, 2017.

[10] M. Kass and D. Pesare. Coherent noise for non-photorealistic rendering. *ACM Trans. Graph. (Proc. SIGGRAPH)*, 30(4):1–6, 2011.

[11] B. Krajancich, P. Kellnhofer, and G. Wetzstein. A perceptual model for eccentricity-dependent spatio-temporal flicker fusion and its applications to foveated graphics. *ACM Trans. Graph.*, 40, 2021.

[12] M. Levoy and P. Hanrahan. Light field rendering. In *Proc. SIGGRAPH*, pp. 31–42, 1996.

[13] B. D. Lucas and T. Kanade. An iterative image registration technique with an application to stereo vision. In *Proc IJCAI*, pp. 674–9, 1981.

[14] S. Ma, B. M. Smith, and M. Gupta. 3d scene flow from 4d light field gradients. In V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, eds., *Computer Vision – ECCV 2018*, pp. 681–698. Springer International Publishing, Cham, 2018.

[15] M. A. Magnor, A. Endmann, and B. Girod. Progressive compression and rendering of light fields. In *Proc. VMV*, pp. 199–204, 2000.

[16] S. P. McKee and K. Nakayama. The detection of motion in the peripheral visual field. *Vision research*, 24(1):25–32, 1984.

[17] X. Meng, R. Du, J. F. JaJa, and A. Varshney. 3d-kernel foveated rendering for light fields. *IEEE Trans. Vis. and Comp. Graph.*, 27(8):3350–3360, 2020.

[18] I. Peter and W. Straßer. The wavelet stream: Progressive transmission of compressed light field data. In *IEEE Visualization*, pp. 69–72, 1999.

[19] E. P. Simoncelli and W. T. Freeman. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Proc. ICPC*, vol. III, pp. 444–447. Washington, DC, Oct 23-26 1995.

[20] Q. Sun, F.-C. Huang, J. Kim, L.-Y. Wei, D. Luebke, and A. Kaufman. Perceptually-guided foveation for light field displays. *ACM Trans. Graph. (Proc. SIGGRAPH Asia)*, 36(6):1–13, 2017.

[21] D. R. Walton, R. K. D. Anjos, S. Friston, D. Swapp, K. Akşit, A. Steed, and T. Ritschel. Beyond blur: Ventral metamers for foveated rendering. *ACM Trans. Graph. (Proc. SIGGRAPH 2021)*, 40(4), 2021.

[22] G. Wu, B. Masia, A. Jarabo, Y. Zhang, L. Wang, Q. Dai, T. Chai, and Y. Liu. Light field image processing: An overview. *IEEE J of Selected Topics in Signal Processing*, 11(7):926–54, 2017.