Epipolar Shift Compensated Light Field Video Quality Metric

Nusrat Mehajabin, Dan Jin, Rui Yao, Thomas Dykstra, Mahsa Pourazad, and Panos Nasiopoulos

Presented by:

Nusrat Mehajabin

PhD Candidate,

Electrical and Computer Engineering

University of British Columbia

Vancouver, Canada

nusratm@ece.ubc.ca

Research

- Light Field (LF) Compression
- LF Refocusing
- LF Quality Metrics
- LF view synthesis

Education

- Ph.D. 2018-present; UBC
- M.S. 2013-2015; DU
- B.Sc. 2009-2013; DU

Work

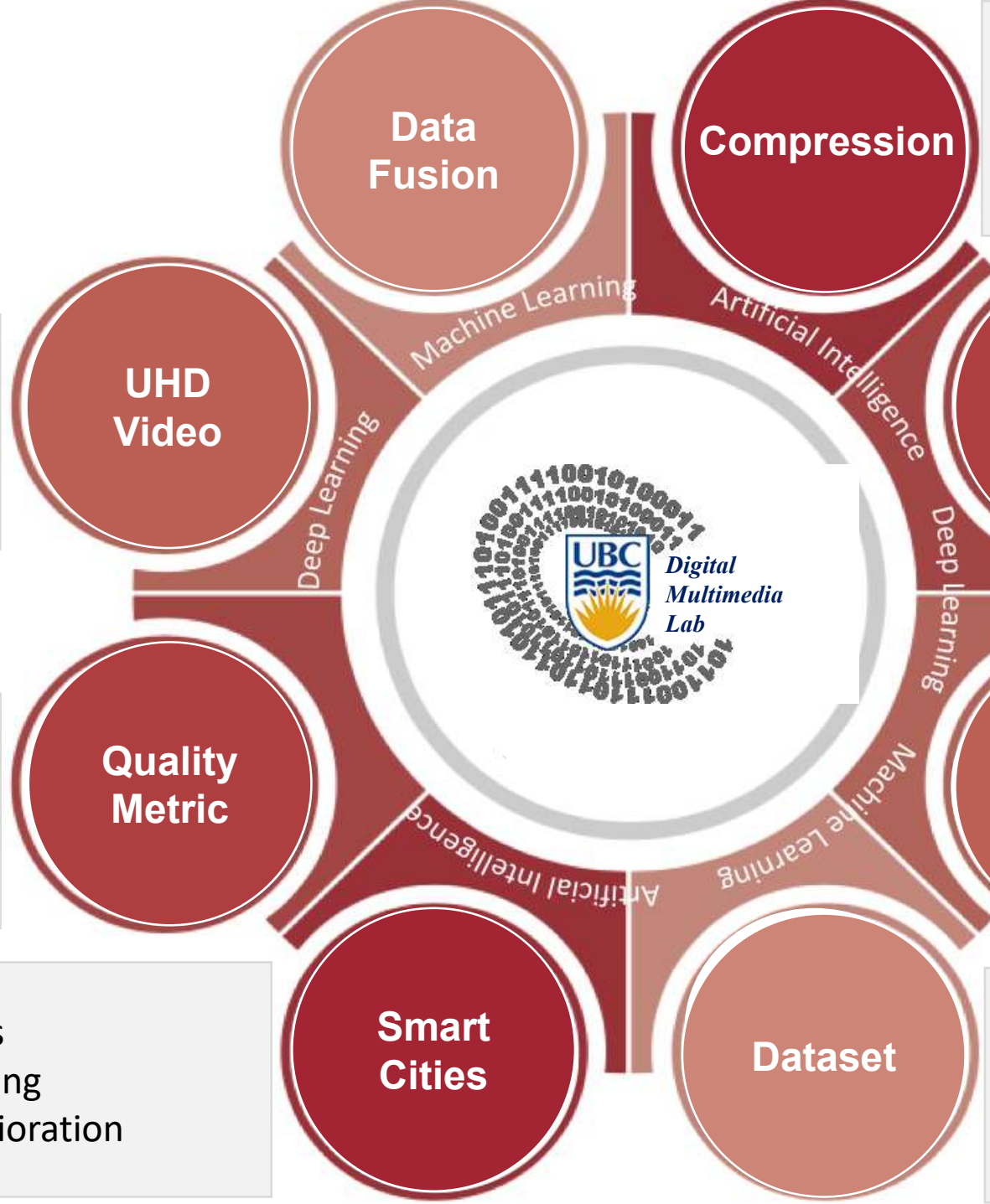
- Research Assistant; DML
- Teaching Assistant; UBC
- Lecturer; DU; BRAC U

Skills

- Deep Learning
- UI design and front-end development
- Software development

Interests

- Internet
- Movies
- Travelling
- Cooking



- 8K
- 4K
- Light Field
- Immersive video

- 4K, 8K, LF capture
- Calibration
- Rendering

- Tone Mapping
- Inverse Tone Mapping

- DL based video quality
- LF spatial and angular quality

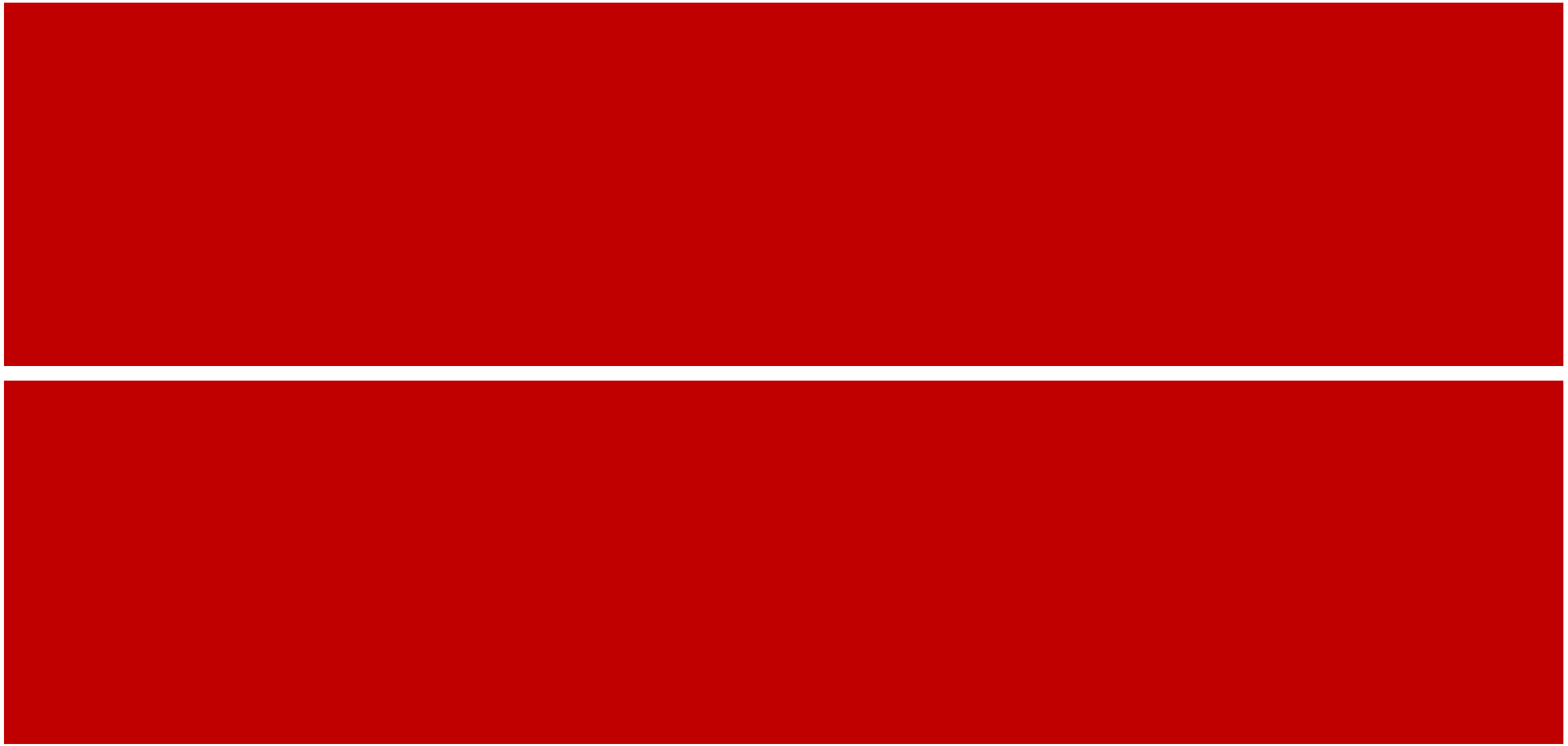
- Fall Detection
- Medical Image processin

- Street Signs
- Street Parking
- Road Deterioration

- 4K and 8K pairs
- 2K and MOS pairs
- Light field

Every light ray, from every point in space - towards every direction

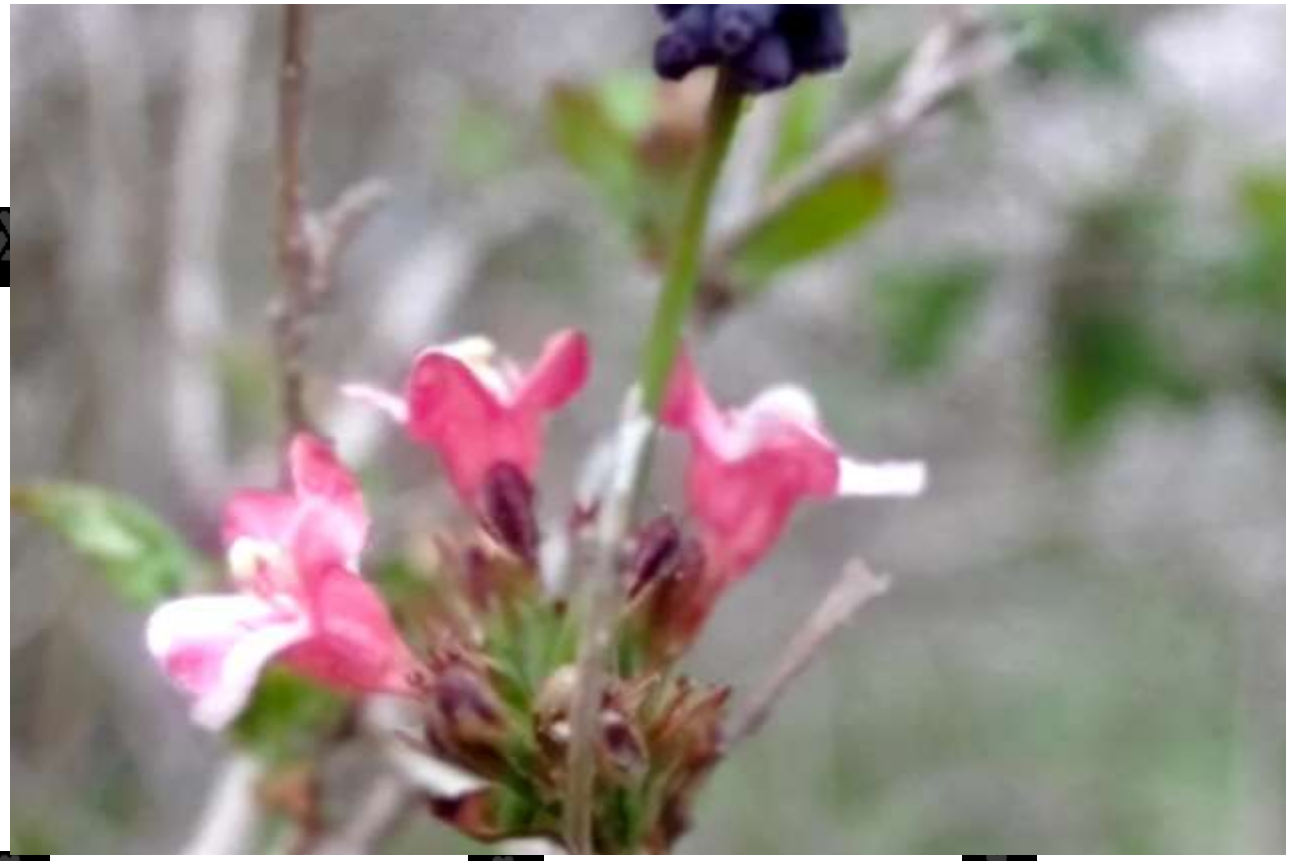
Light Field (LF) Capture



Captured Content (Camera array)

LF can be refocused after capture

LF: Perspective Change



Light Field Applications

Autonomous driving

- Tesla
- Waymo
- Ford
- Uber

Remote surgery

- Intuitive Surgical
- Hansen Medical
- Diligent Robotics
- Medrobotics

Security & surveillance

- Bosch
- Honeywell
- Cisco
- Milestone

AR/VR applications

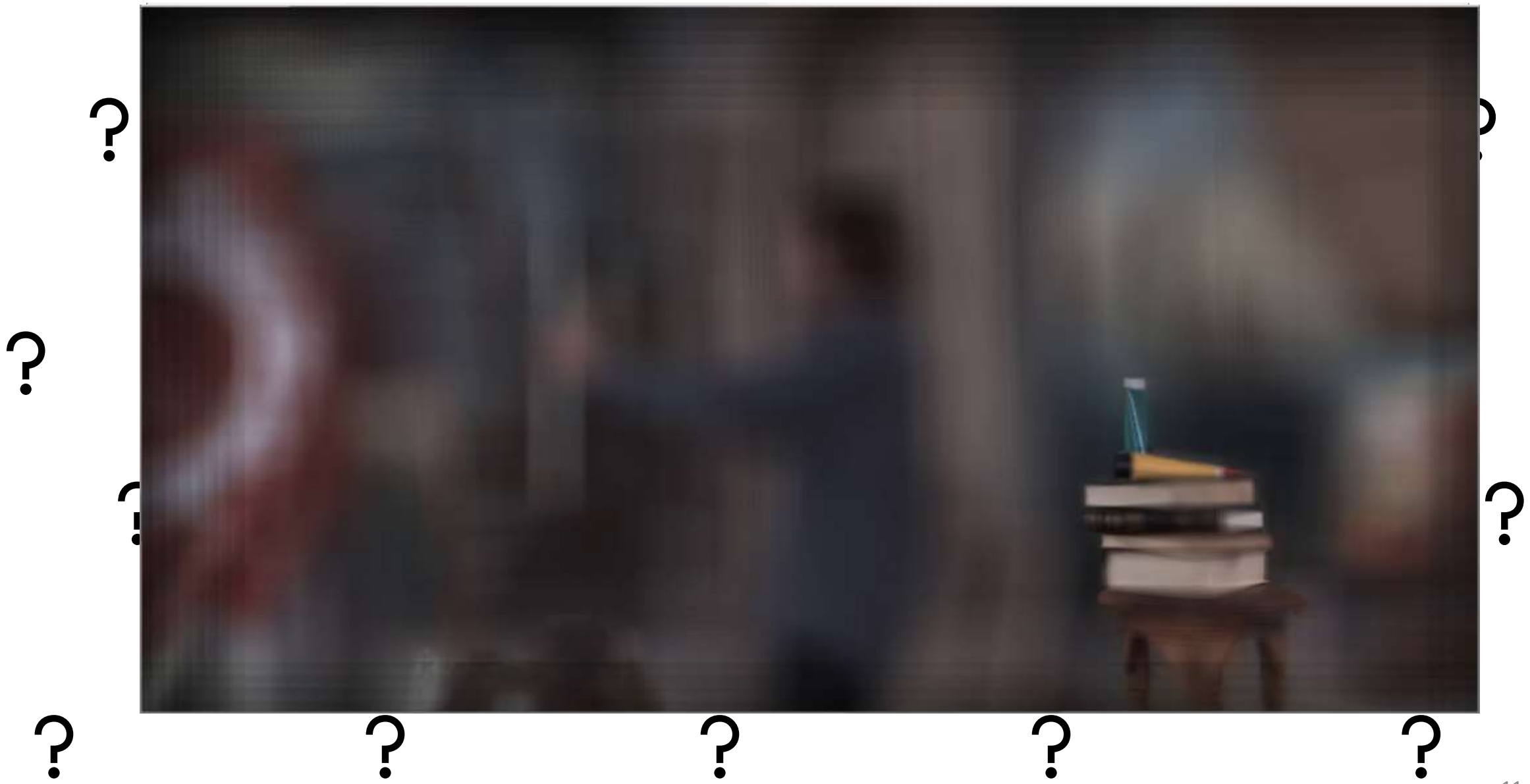
- Magic Leap
- Microsoft
- Apple
- Meta

Training & education

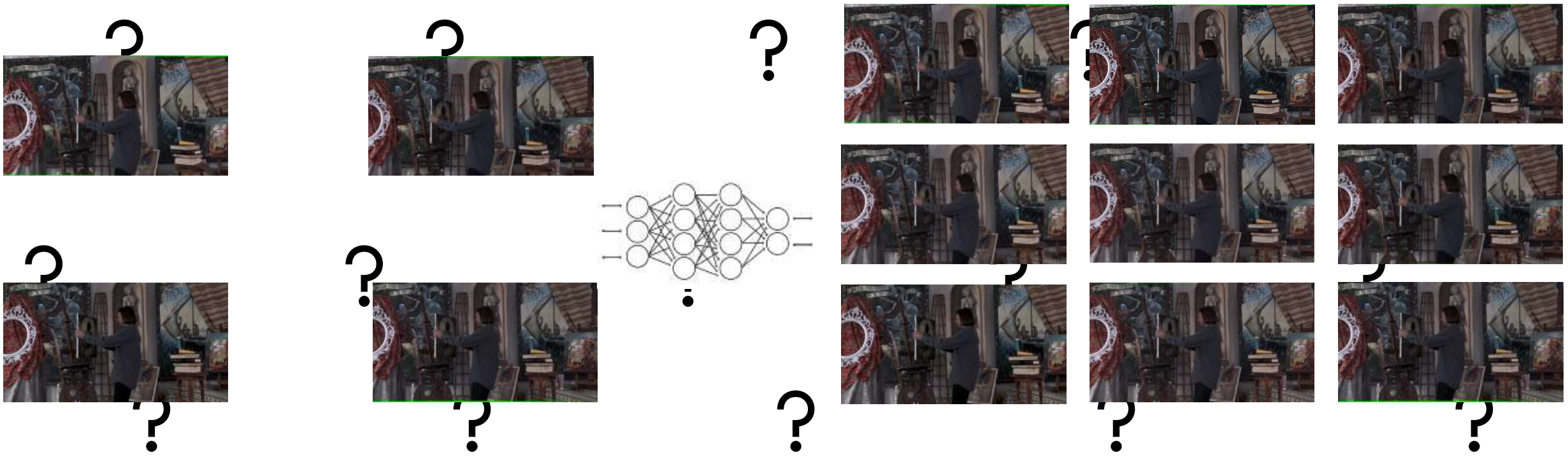
- Zoom
- Microsoft
- Udemy
- Meta

Process the LF

Refocused on the “books”

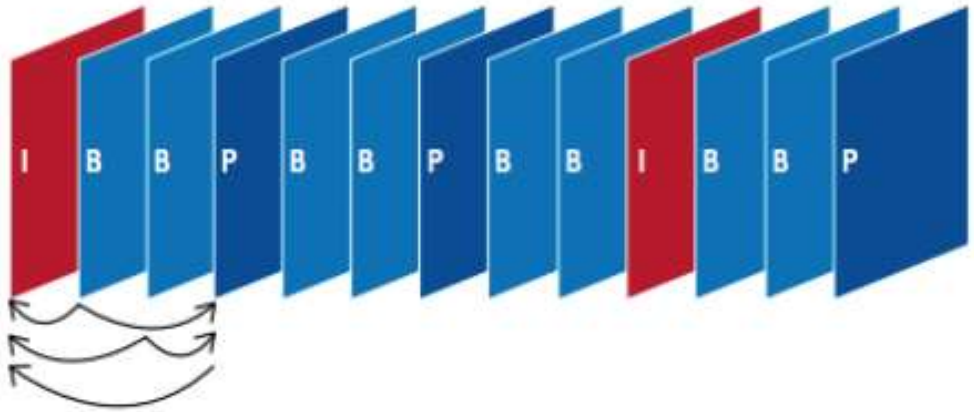


Synthesize Novel Views in between

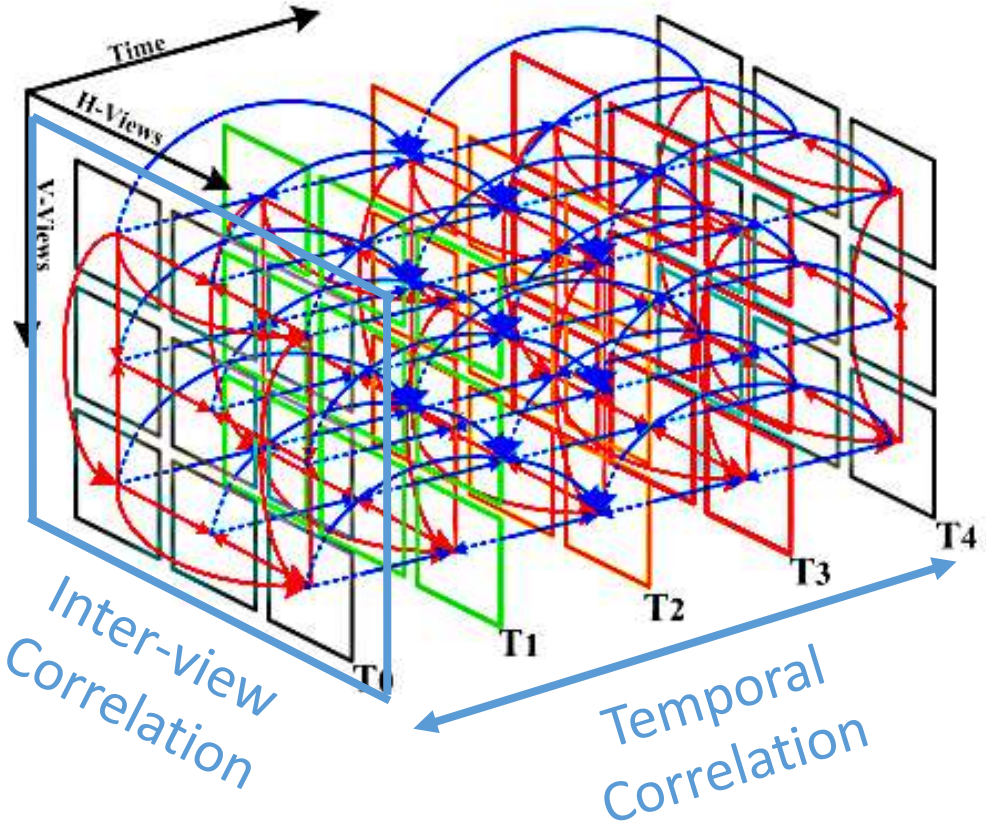


Compression

Conventional Video



Light Field Video



Best quality

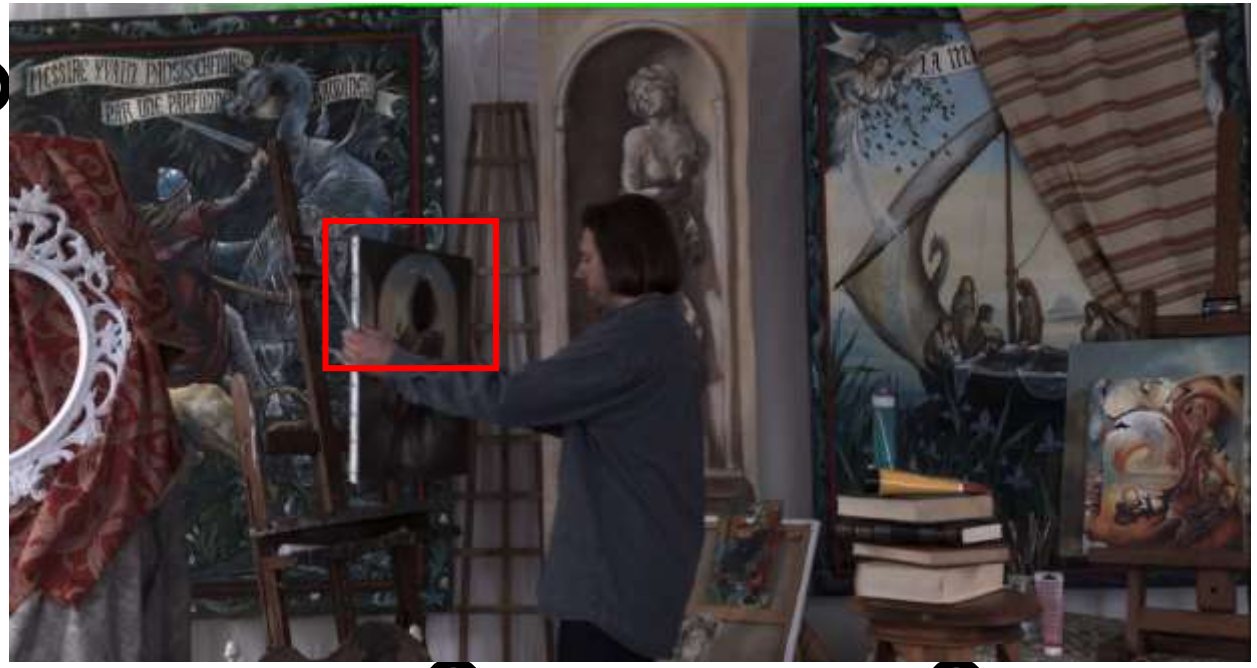
Avg. quality

Low quality

?

?

?



?

?

?

?

?

Our Objective:

Determine Quality of Processed LF Compared to Original



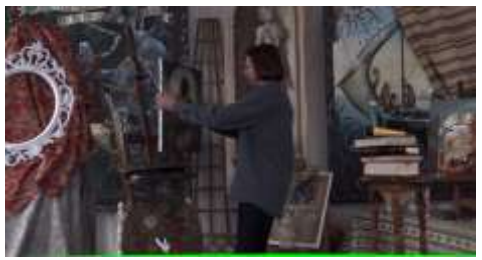


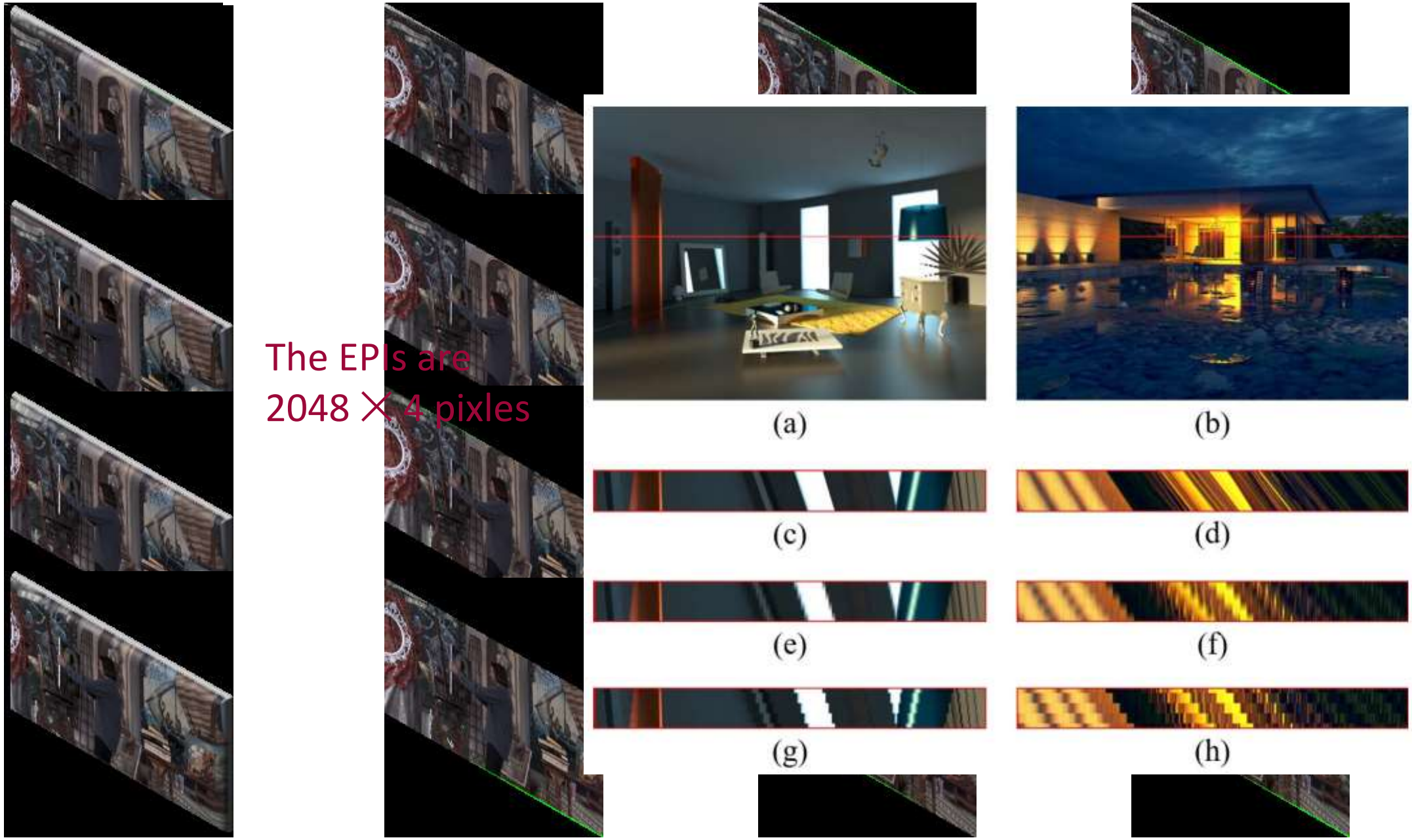


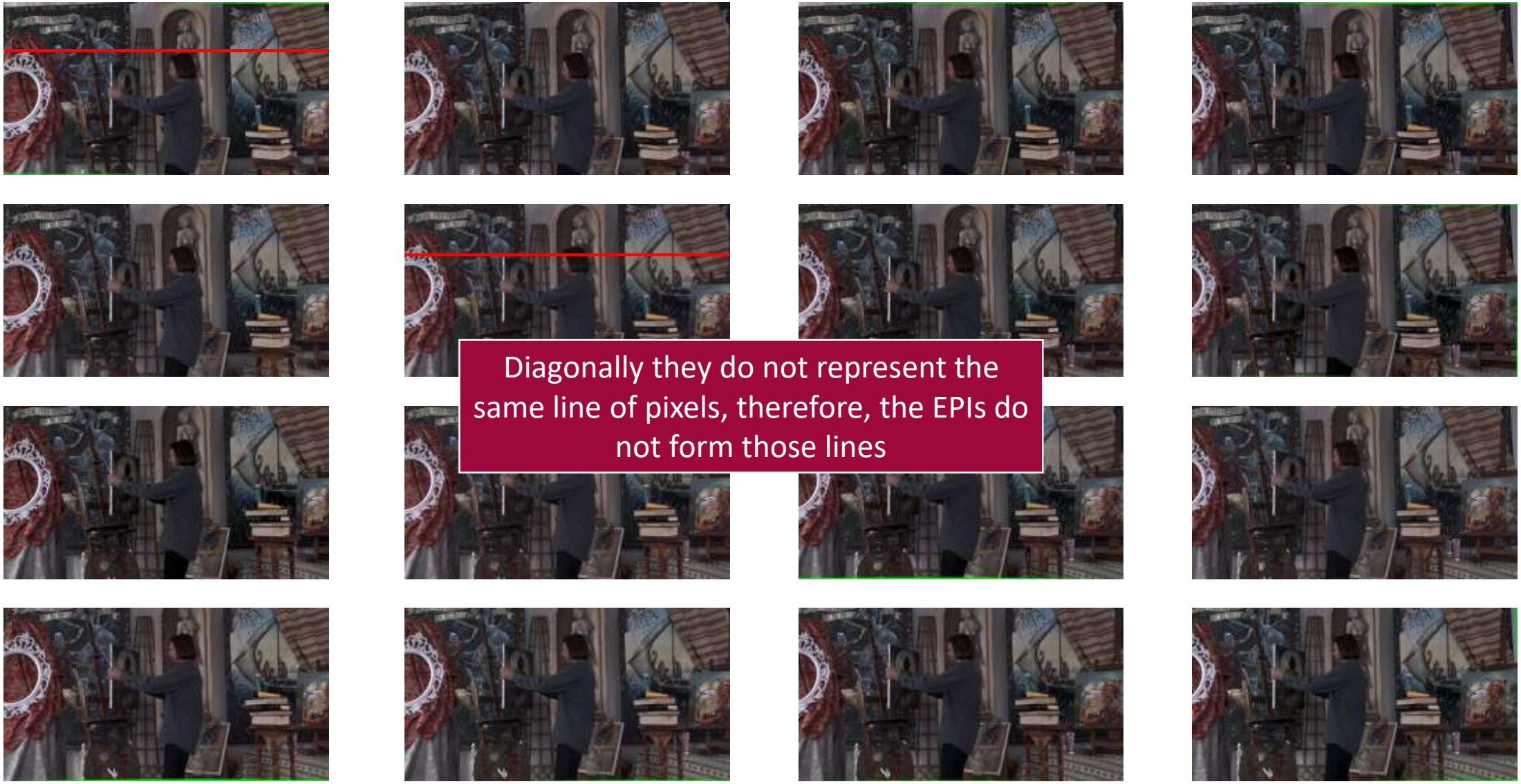


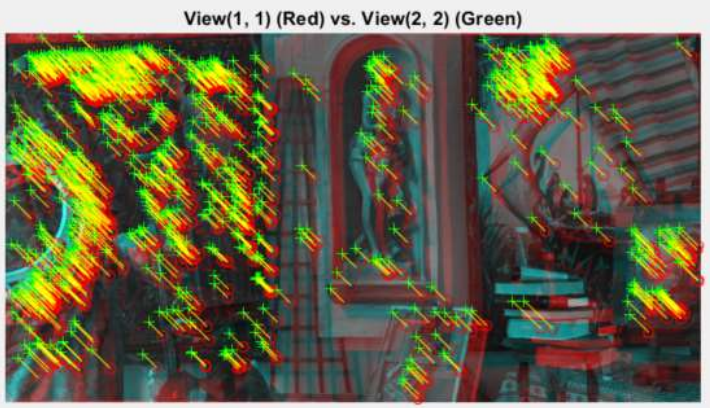


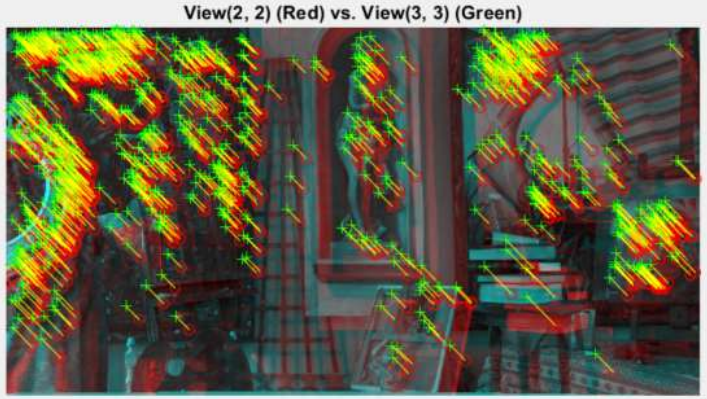


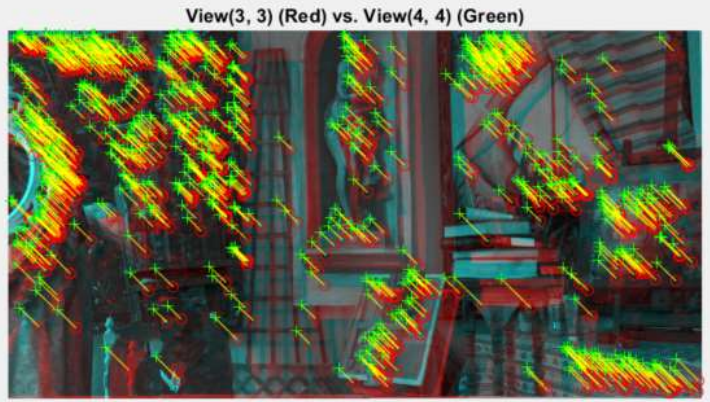




The EPIs are
2048 × 4 pixles
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)


Diagonally they do not represent the same line of pixels, therefore, the EPIs do not form those lines







We find the corresponding rows by using SIFT

Spatial Consistency

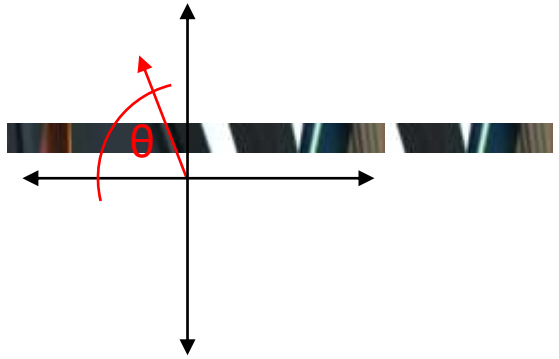
Original

Distorted (Compressed)

	QP45	QP40	QP35	QP30	QP25
Horizontal EPI PSNR	30.6437	33.1357	35.6426	37.8914	39.9501
Vertical EPI PSNR	30.9127	33.4326	35.9023	38.0744	40.0443
Diagonal EPI PSNR	29.2132	31.3736	32.9539	33.7960	35.3868
Average PSNR	**30.2565**	**32.6473**	**34.8329**	**36.5873**	**38.4604**

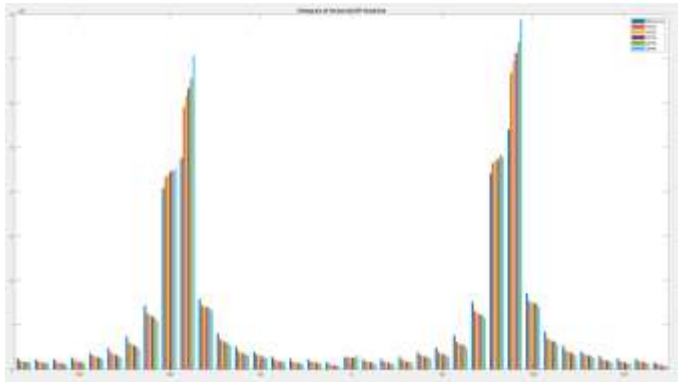
	QP45	QP40	QP35	QP30	QP25
Horizontal EPI SSIM	0.8945	0.9351	0.9605	0.9746	0.9831
Vertical EPI SSIM	0.8804	0.9250	0.9531	0.9688	0.9782
Diagonal EPI SSIM	0.8528	0.9002	0.9282	0.9407	0.9556
Average SSIM	**0.8759**	**0.9201**	**0.9473**	**0.9614**	**0.9723**

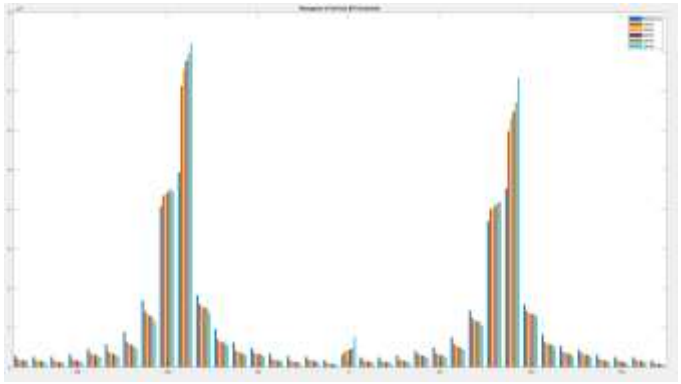
Angular Consistency

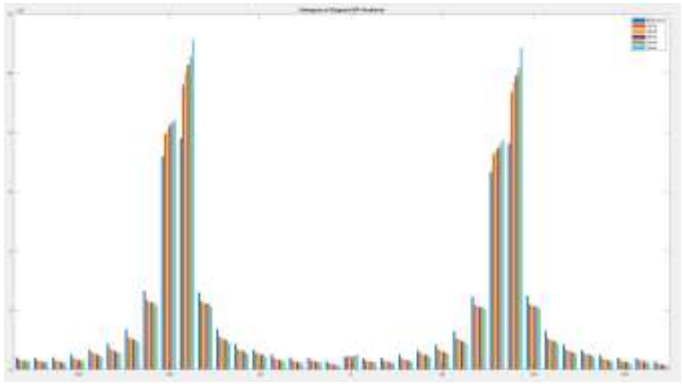


Original

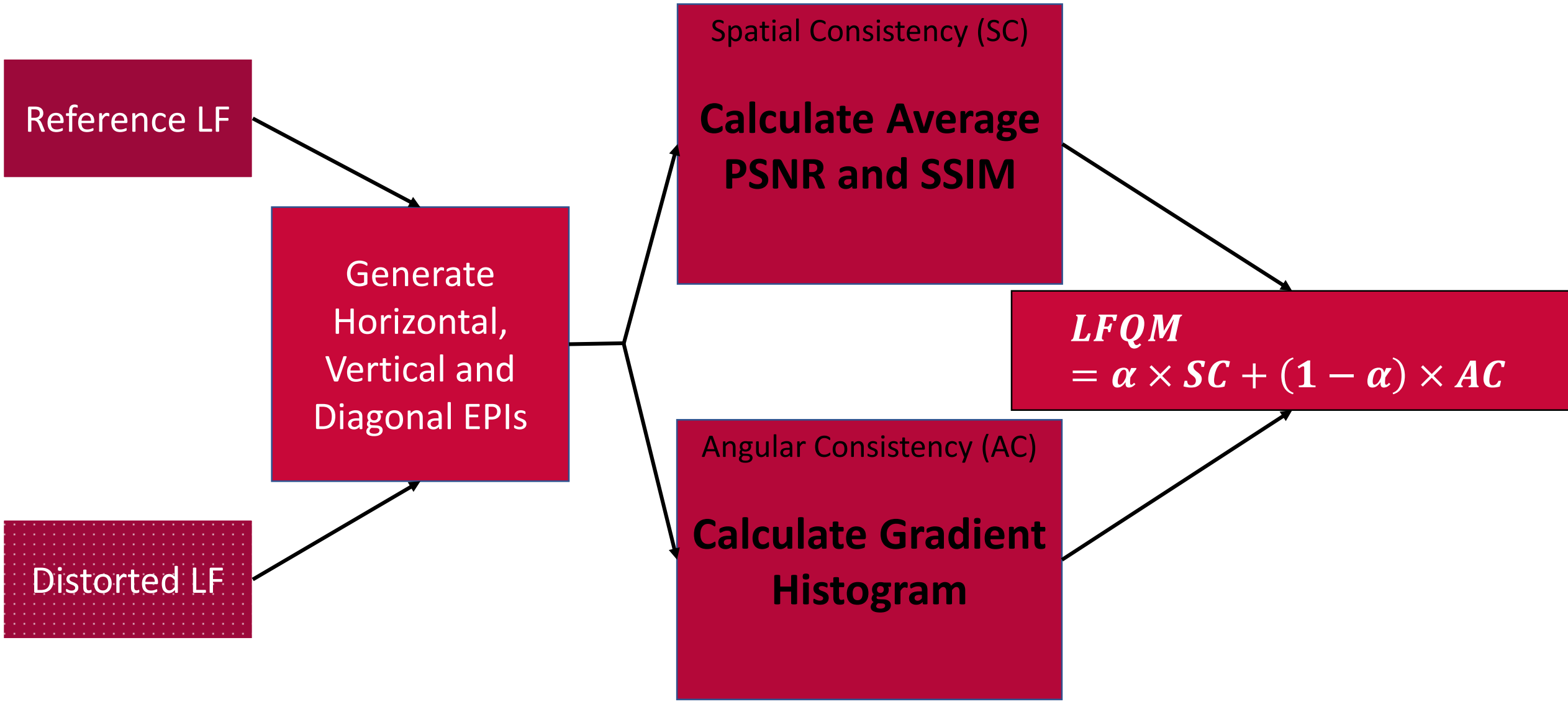
Distorted (Compressed)







Horizontal, vertical and diagonal EPI gradient histogram



*α is chosen depending on the distortion operation

Conclusion

- We developed a new LF quality metric
- First method to considers spatial and full angular consistency of LF
- Future goals
  - Test the metric for other operations such as synthesis and refocusing
  - Subjective tests to validate the objective metric

References

[1] M. Levoy and P. Hanrahan, “Light Field Rendering,” In Proceedings of the 23rd annual conference on Computer graphics and interactive techniques, 1996, Aug 1, pp. 31-42.

[2] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, “Light Field Photography with a Hand-held Plenoptic Camera,” 2005 (Doctoral dissertation, Stanford University) .

[3] C. Zhang, G. Hou, Z. Zhang, Z. Sun, and T. Tan, “Efficient auto-refocusing for light field camera,” Pattern Recognit., vol. 81, pp. 176–189, Sep. 2018, doi: 10.1016/j.patcog.2018.03.020.

[4] C. Conti, L. D. Soares, and P. Nunes, “Light field coding with field-of-view scalability and exemplar-based interlayer prediction,” IEEE Trans. Multimed., vol. 20, no. 11, pp. 2905–2920, 2018, doi: 10.1109/TMM.2018.2825882.

[5] N. Mehajabin, M. T. Pourazad, and P. Nasiopoulos, “An Efficient Pseudo-Sequence-Based Light Field Video Coding Utilizing View Similarities for Prediction Structure,” IEEE Trans. Circuits Syst. Video Technol., 2021, pp. 1-16, doi: 10.1109/TCSVT.2021.3092282.

[6] N. Sabater et al., “Dataset and Pipeline for Multi-view Light-Field Video,” IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work., vol. 2017-July, pp. 1743–1753, 2017, doi: 10.1109/CVPRW.2017.221.

[7] S. L. P. Yasakethu, C. T. E. R. Hewage, W. A. C. Fernando, and A. M. Kondoz, “Quality analysis for 3D video using 2D video quality models,” IEEE Trans. Consum. Electron., vol. 54, no. 4, pp. 1969–1976, 2008, doi: 10.1109/TCE.2008.4711260.

[8] P. Joveluro, H. Malekmohamadi, W. A. C. Fernando, and A. M. Kondoz, “Perceptual video quality metric for 3D video quality assessment,” 3DTV-CON 2010 True Vis. - Capture, Transm. Disp. 3D Video, pp. 1–4, 2010, doi: 10.1109/3DTV.2010.5506331.

[9] V. K. Adhikarla, M. Vinkler, D. Sumin, R. K. Mantiuk, K. Myszkowski, and P. Didyk, “Towards a quality metric for dense light fields,” in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2, 2017, pp. 58–67.

[10] Y. Fang, K. Wei, J. Hou, W. Wen, and N. Imamoglu, “Light Filed Image Quality Assessment by Local and Global Features of Epipolar Plane Image,” 2018 IEEE 4th Int. Conf. Multimed. Big Data, BigMM 2018, pp. 1-6, doi: 10.1109/BigMM.2018.8499086.

[11] P. Paudyal, F. Battisti, S. Member, M. Carli, and S. Member, “Reduced Reference Quality Assessment of Light Field Images,” IEEE Trans. Broadcast., vol. 65, no. 1, pp. 152–165, 2019.

[12] L. Shi, W. Zhou, Z. Chen, and J. Zhang, “No-Reference Light Field Image Quality Assessment Based on Spatial-Angular Measurement,” IEEE Trans. Circuits Syst. Video Technol., vol. 30, no. 11, pp. 4114–4128, 2020, doi: 10.1109/TCSVT.2019.2955011.

[13] D. G. Lowe, “Object recognition from local scale-invariant features,” Proc. IEEE Int. Conf. Comput. Vis., vol. 2, pp. 1150–1157, 1999, doi: 10.1109/iccv.1999.790410.

[14] G. Tech, Y. Chen, K. Müller, J. R. Ohm, A. Vetro, and Y. K. Wang, “Overview of the multiview and 3D extensions of high efficiency video coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 26, no. 1, pp. 35–49, 2016, doi: 10.1109/TCSVT.2015.2477935.

Thank You

Questions