Motion-Plane-Adaptive Inter Prediction in 360-Degree Video Coding

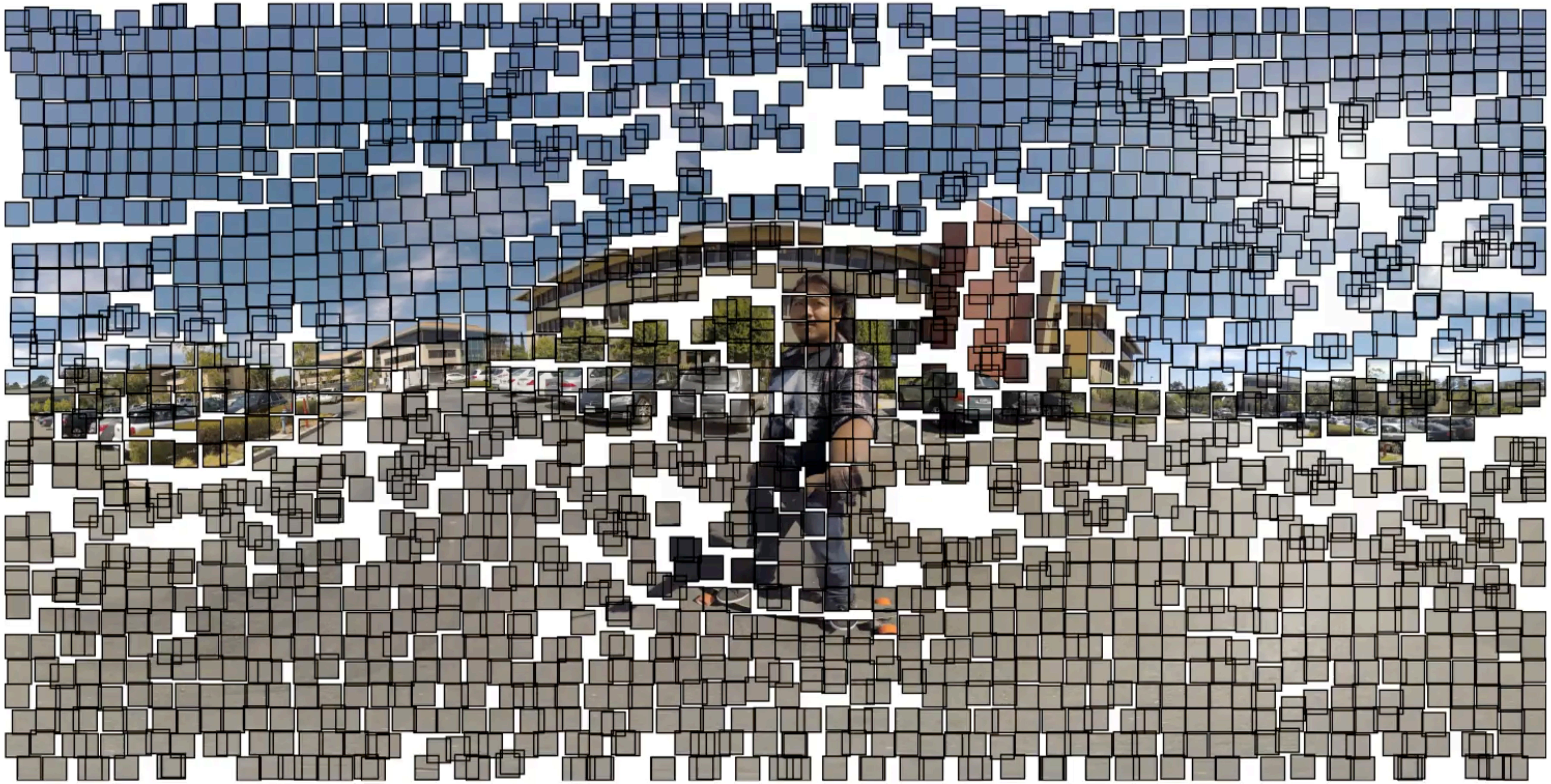
Andy Regensky
andy.regensky@fau.de
Chair of Multimedia Communications and Signal Processing

FAU
Friedrich-Alexander-Universität
Technische Fakultät

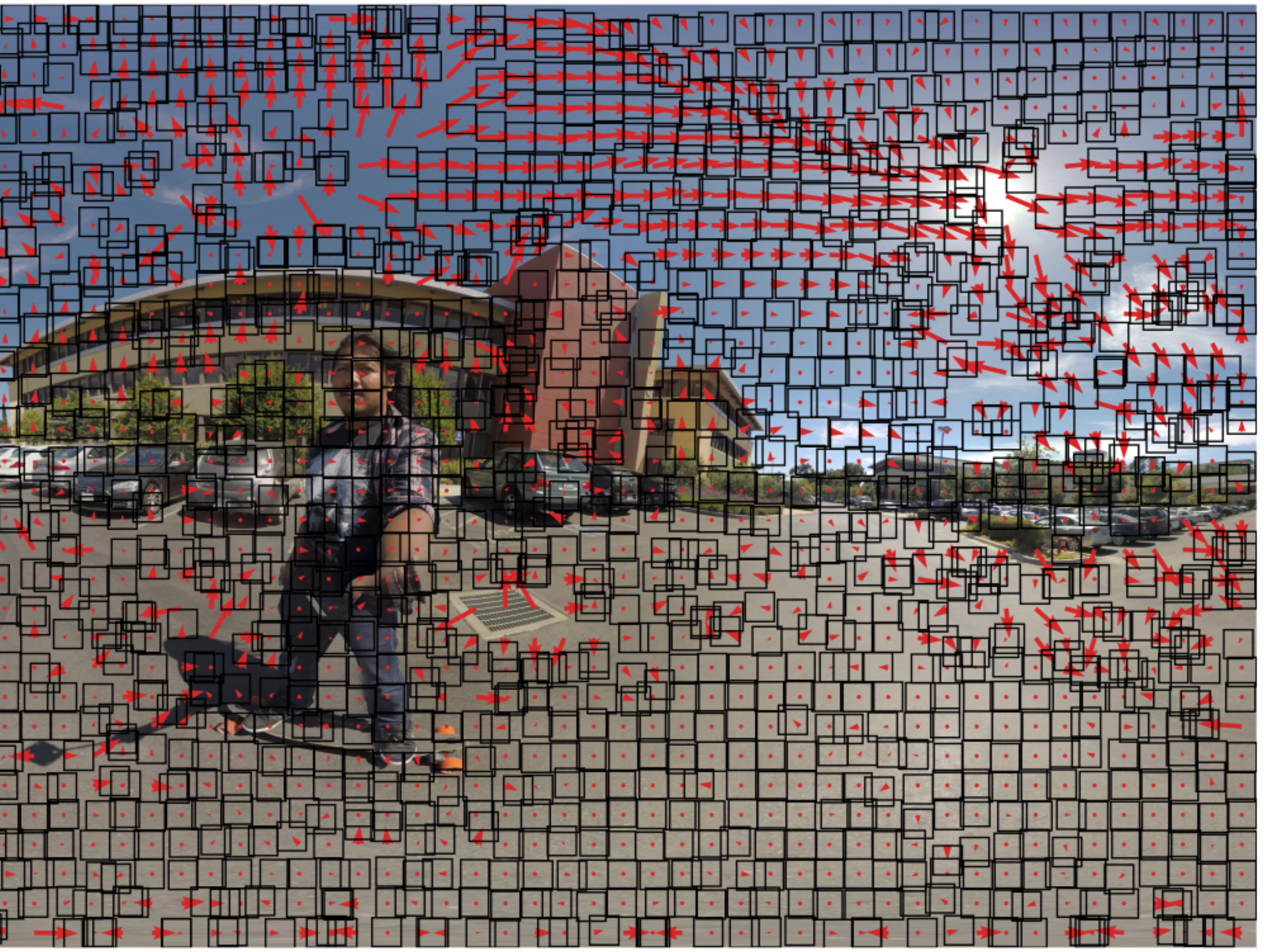
LMS

Motivation

Motivation



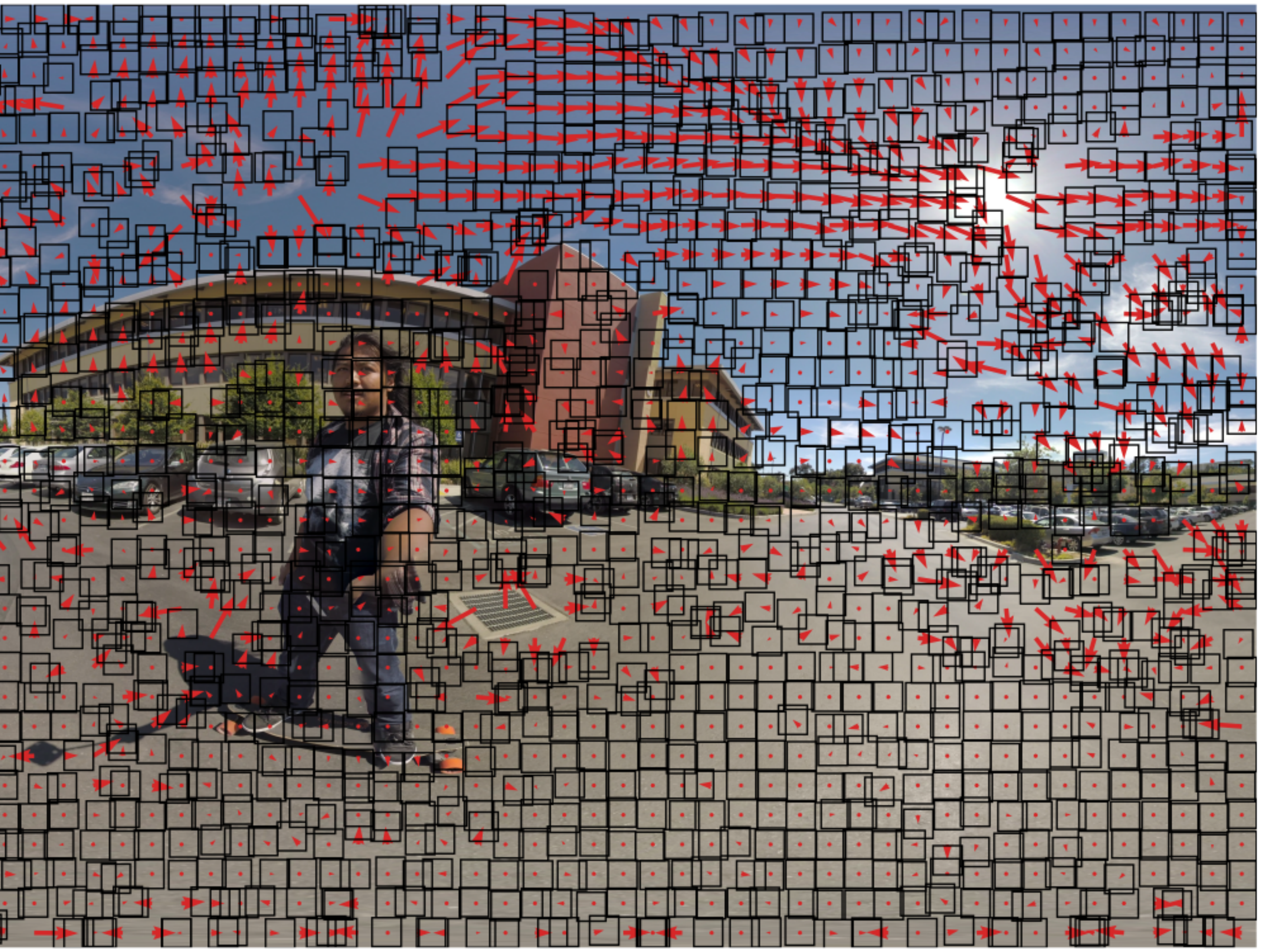
Motivation



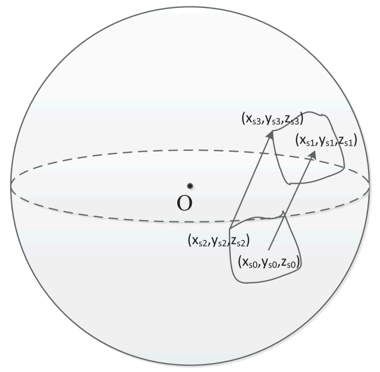
$$\boldsymbol{I}_{\text{pred}}(\boldsymbol{p}) = \boldsymbol{I}_{\text{ref}}\left(\boldsymbol{m}_{\text{t}}(\boldsymbol{p}, \boldsymbol{t}_{ij}^{*})\right) \;\; \forall\, \boldsymbol{p} \in \mathcal{B}_{ij}$$

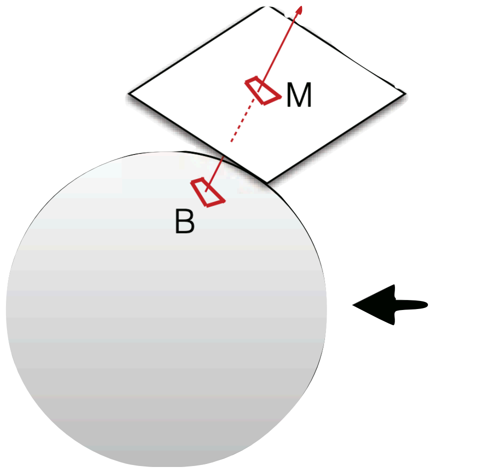
$\boldsymbol{I}_{\text{pred}} \in \mathbb{R}^{U \times V}$	:	predicted image
$\boldsymbol{I}_{\text{ref}} \in \mathbb{R}^{U \times V}$	:	reference image
$\boldsymbol{p} \in \mathbb{R}^2$	:	pixel coordinate
$\boldsymbol{t}_{\boldsymbol{ij}}^{*} \in \mathbb{R}^2$	:	motion vector for block (i, j)
$\mathcal{B}_{ij}$	:	set of pixel coordinates within block (i, j))
$\boldsymbol{m}_{\text{t}} : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}^2$	:	translational motion model with $\boldsymbol{m}_{\text{t}}(\boldsymbol{p}, \boldsymbol{t}) = \boldsymbol{p} + \boldsymbol{t}$

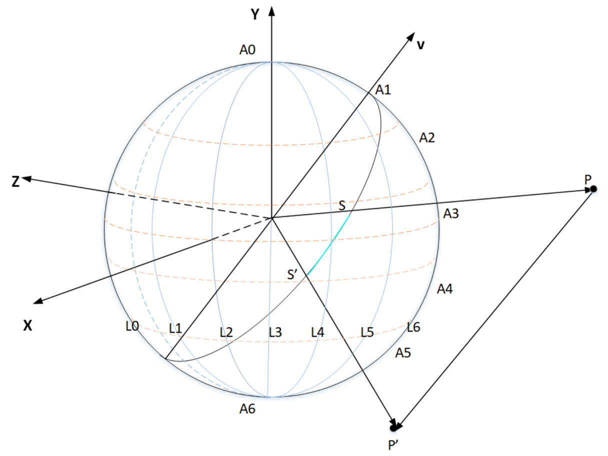
Motivation



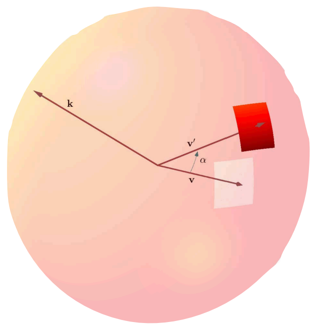
$$\boldsymbol{I}_{\text{pred}}(\boldsymbol{p}) = \boldsymbol{I}_{\text{ref}}\left(\boldsymbol{m}_{\text{t}}(\boldsymbol{p}, \boldsymbol{t}_{ij}^{*})\right) \ \forall \ \boldsymbol{p} \in \mathcal{B}_{ij}$$

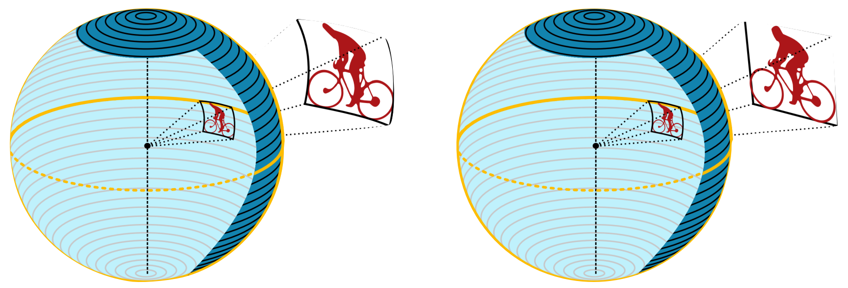






- 3D-Translational: *Li et al.*, *Wang et al.*
- Tangential: *De Simone et al.*
- Rotational: *Vishwanath et al.*
- Geodesic: *Vishwanath et al.*
- Tangential + Rotational: *Marie et al.*





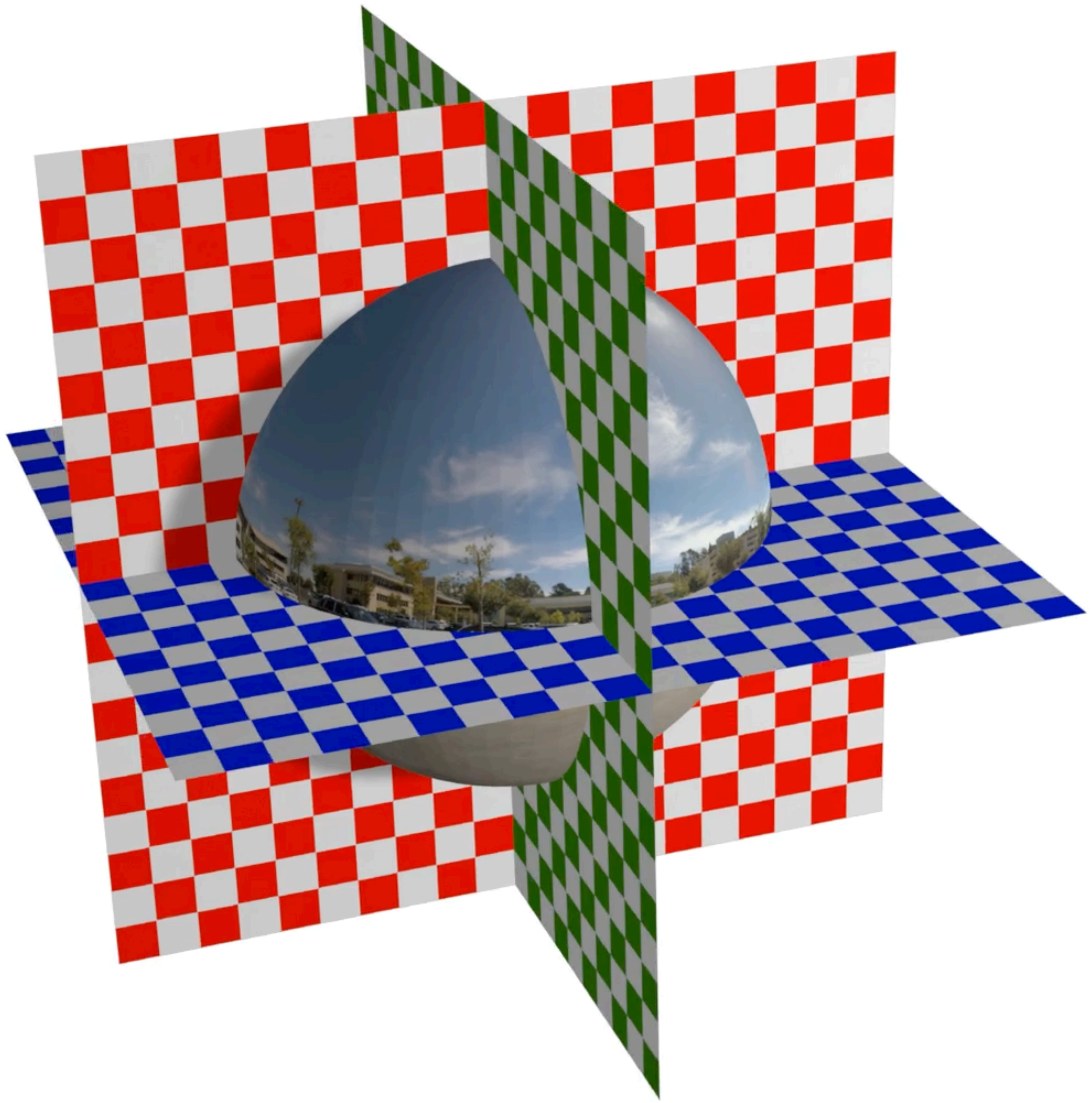
Li et al., "Projection Based Advanced Motion Model for Cubic Mapping for 360-Degree Video," in Proc. IEEE International Conference on Image Processing (ICIP), 2017, pp. 1427–1431.
Li et al., "Advanced Spherical Motion Model and Local Padding for 360° Video Compression," IEEE Trans. Image Process., vol. 28, no. 5, pp. 2342–2356, May 2019.
Wang et al., "A New Motion Model for Panoramic Video Coding," in Proc. IEEE International Conference on Image Processing (ICIP), 2017, pp. 1407–1411.
Wang et al., "Spherical Coordinates Transform-Based Motion Model for Panoramic Video Coding," IEEE J. Emerg. Sel. Top. Circuits Syst., vol. 9, no. 1, pp. 98–109, Mar. 2019.
De Simone et al., "Deformable Block-Based Motion Estimation in Omnidirectional Image Sequences," in Proc. IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), 2017, pp. 1–6.
Vishwanath et al., "Rotational Motion Model for Temporal Prediction in 360 Video Coding," in Proc. IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), 2017, pp. 1–6.
Vishwanath et al., "Rotational Motion Compensated Prediction in HEVC Based Omnidirectional Video Coding," in Proc. Picture Coding Symposium (PCS), 2018, pp. 323–327.
Vishwanath et al., "Motion Compensated Prediction for Translational Camera Motion in Spherical Video Coding," in Proc. IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), 2018, pp. 1–4.
Marie et al., "Rate-Distortion Optimized Motion Estimation for on-the-Sphere Compression of 360 Videos," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021, pp. 1570–1574.

Outline

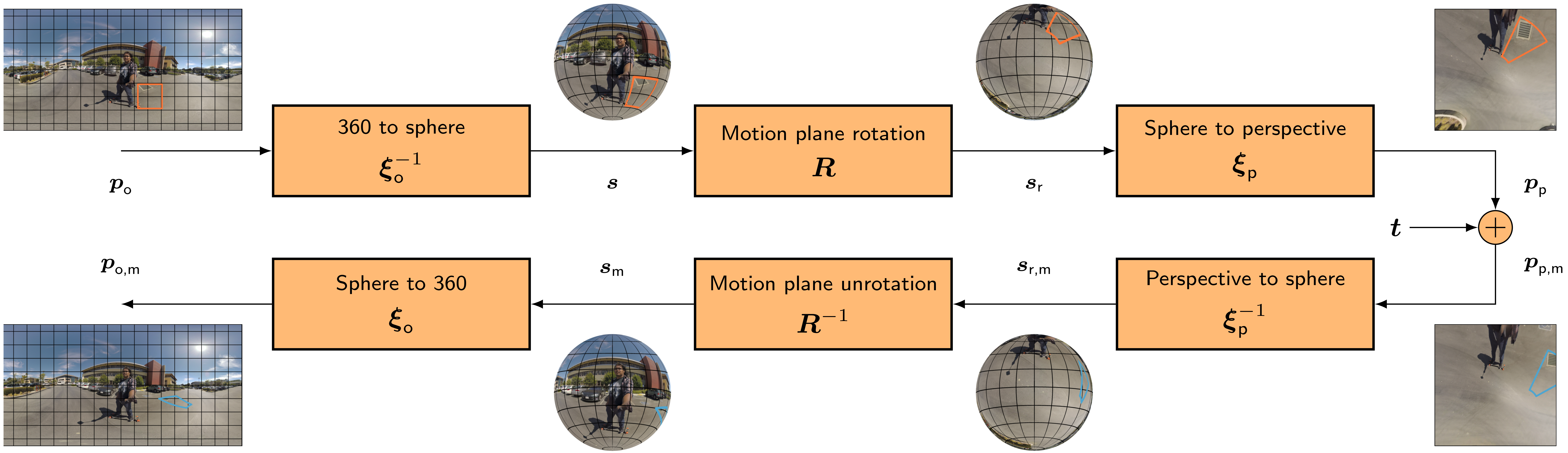
1. Motion-Plane-Adaptive Inter Prediction
   a. Motion-Plane-Adaptive Motion Model (MPA)
   b. Adapted Motion Vector Prediction (MPA-MVP)
2. Performance Evaluation
3. Conclusion

Motion Planes

- Virtual perspective camera
- Set of three motion planes
  - Left/Right
  - Front/Back
  - Bottom/Top
- Each CU selects motion plane leading to minimum RD-cost



Motion-Plane-Adaptive Motion Model



Motion plane reprojection

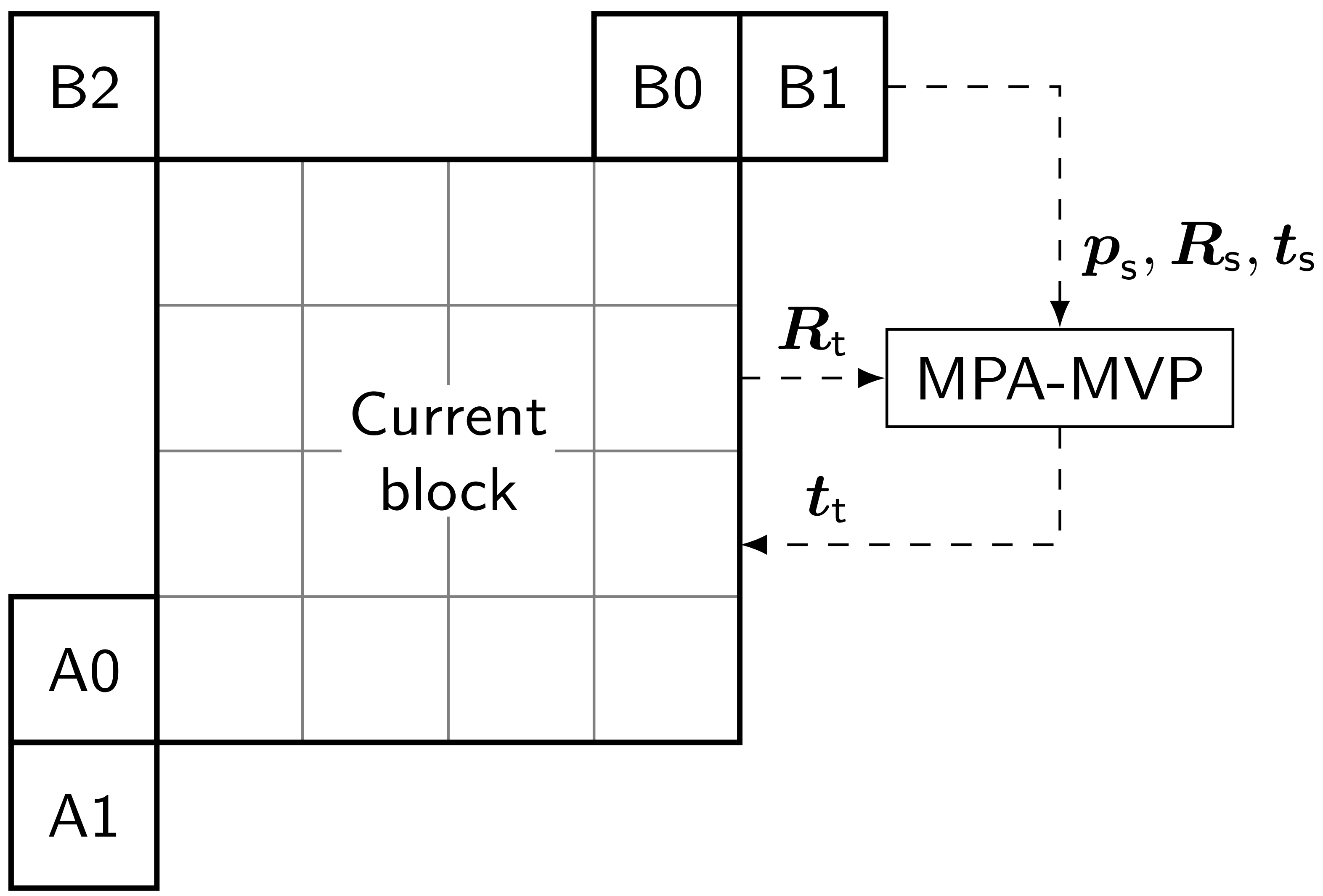
$$\zeta_{\boldsymbol{R}}(\boldsymbol{p}_{\mathrm{o}}) = \xi_{\mathrm{p}}(\boldsymbol{R}\xi_{\mathrm{o}}^{-1}(\boldsymbol{p}_{\mathrm{o}}))$$

Motion-plane-adaptive motion model

$$\boldsymbol{m}_{\mathrm{mpa}}(\boldsymbol{p}_{\mathrm{o}}, \boldsymbol{t}, \boldsymbol{R}) = \zeta_{\boldsymbol{R}}^{-1}(\zeta_{\boldsymbol{R}}(\boldsymbol{p}_{\mathrm{o}})+\boldsymbol{t})$$

Regensky et al., “Motion-Plane-Adaptive Inter Prediction in 360-Degree Video Coding,“ arXiv:2022.03323, Feb. 2022.

Motion-Plane-Adaptive Motion Vector Prediction



- MPA to MPA:

$$\boldsymbol{m}_{\mathrm{mpa}}(\boldsymbol{p}_{\mathrm{s}}, \boldsymbol{t}_{\mathrm{s}}, \boldsymbol{R}_{\mathrm{s}}) \overset{!}{=} \boldsymbol{m}_{\mathrm{mpa}}(\boldsymbol{p}_{\mathrm{s}}, \boldsymbol{t}_{\mathrm{t}}, \boldsymbol{R}_{\mathrm{t}})$$

$$\downarrow$$

$$\boldsymbol{t}_{\mathrm{t}} = \zeta_{\boldsymbol{R}_{\mathrm{t}}}(\boldsymbol{m}_{\mathrm{mpa}}(\boldsymbol{p}_{\mathrm{s}}, \boldsymbol{t}_{\mathrm{s}}, \boldsymbol{R}_{\mathrm{s}})) - \zeta_{\boldsymbol{R}_{\mathrm{t}}}(\boldsymbol{p}_{\mathrm{s}})$$

- Similar for MPA $\leftrightarrow$ translational
- Also applicable to affine MVP and merge candidate construction

Li et al., “An Efficient Four-Parameter Affine Motion Model for Video Coding,” IEEE Trans. Circuits Syst. Video Technol., vol. 28, no. 8, pp. 1934–1948, Aug. 2018.
Zhang et al., “An Improved Framework of Affine Motion Compensation in Video Coding,” IEEE Trans. Image Process., vol. 28, no. 3, pp. 1456–1469, Mar. 2019.
Bross et al., “Versatile Video Coding (Draft 10), JVET-T2001-v2,” in Proc. 20th Meeting of the Joint Video Exploration Team (JVET), 2020, pp. 1–521.

Outline

1. Motion-Plane-Adaptive Inter Prediction
   a. Motion-Plane-Adaptive Motion Model (MPA)
   b. Adapted Motion Vector Prediction (MPA-MVP)
2. Performance Evaluation
3. Conclusion

Experimental Setup

- VTM-14.2 as baseline
- Random access configuration
- 32 frames of 10 grayscale sequences
- QPs 22, 27, 32, 37
- Evaluation procedure according to JVET 360-CTC

ITU/ISO/IEC, “VVC VTM Reference Software VTM-14.2,” 2021. [Online]. Available: https://vcgit.hhi.fraunhofer.de/jvet/VVCSoftware_VTM/-/tree/VTM-14.2.

Browne et al., “Algorithm Description for Versatile Video Coding and Test Model 14 (VTM 14), JVET-W2002-v1,” in Proc. 23rd Meeting of the Joint Video Exploration Team (JVET), 2021, pp. 1–111.

Hanhart et al., “JVET Common Test Conditions and Evaluation Procedures for 360° Video, JVET-L1012,” in Proc. 12th Meeting of the Joint Video Exploration Team (JVET), 2018, pp. 1–7.

Bossen et al., “VTM Common Test Conditions and Software Reference Configurations for SDR Video, JVET-T2010,” in Proc. 20th Meeting of the Joint Video Experts Team (JVET), 2020, pp. 1–2.

G. Bjontegaard, “Calculation of Average PSNR Differences between RD-curves, VCEG-M33,” in Proc. 13th Meeting Video Coding Experts Group, 2001, pp. 1–5.

Ye et al., “Algorithm Descriptions of Projection Format Conversion and Video Quality Metrics in 360Lib Version 12, JVET-T2004-v2,” in Proc. 20th Meeting of the Joint Video Experts Team (JVET), 2020, pp. 1–65.

ITU/ISO/IEC, “360Lib Software 360Lib-12.0,” 2020. [Online]. Available: https://jvet.hhi.fraunhofer.de/svn/svn_360Lib/tags/360Lib-12.0/.

Motion Model Comparison

BD-Rate in % for different motion models based on WS-PSNR with respect to VTM 14.2

	MPA	3DT	TAN	ROT	GED	TAN+ROT
SkateboardInLot	**-1.06**	+0.13	+0.78	+0.79	-0.13	+0.88
ChairliftRide	**-1.22**	-0.16	+0.10	+0.42	-1.19	+0.13
KiteFlite	+0.03	+0.08	+0.06	**-0.02**	+0.05	+0.29
Harbor	+0.27	+0.16	+0.31	+0.07	+0.13	+0.15
Trolley	+0.04	+0.16	-0.02	+0.23	+0.15	**-0.04**
GasLamp	-0.08	-0.10	**-0.74**	-0.47	-0.57	-0.64
Balboa	**-1.87**	-0.07	+0.59	+0.71	-0.53	+0.59
Broadway	**-1.64**	+0.11	+0.78	+0.50	-0.61	+0.52
Landing2	-2.92	-1.03	-1.72	+0.25	**-3.50**	-0.99
BranCastle2	-0.58	-0.22	+0.03	+0.18	**-0.62**	+0.15
Average	**-0.90**	-0.09	+0.02	+0.27	-0.68	+0.10

MPA-MVP=OFF

Li et al., "Projection Based Advanced Motion Model for Cubic Mapping for 360-Degree Video," in Proc. IEEE International Conference on Image Processing (ICIP), 2017, pp. 1427–1431.
Li et al., "Advanced Spherical Motion Model and Local Padding for 360° Video Compression," IEEE Trans. Image Process., vol. 28, no. 5, pp. 2342–2356, May 2019.
Wang et al., "A New Motion Model for Panoramic Video Coding," in Proc. IEEE International Conference on Image Processing (ICIP), 2017, pp. 1407–1411.
Wang et al., "Spherical Coordinates Transform-Based Motion Model for Panoramic Video Coding," IEEE J. Emerg. Sel. Top. Circuits Syst., vol. 9, no. 1, pp. 98–109, Mar. 2019.
De Simone et al., "Deformable Block-Based Motion Estimation in Omnidirectional Image Sequences," in Proc. IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), 2017, pp. 1–6.
Vishwanath et al., "Rotational Motion Model for Temporal Prediction in 360 Video Coding," in Proc. IEEE 19th International Workshop on Multimedia Signal Processing (MMSP), 2017, pp. 1–6.
Vishwanath et al., "Rotational Motion Compensated Prediction in HEVC Based Omnidirectional Video Coding," in Proc. Picture Coding Symposium (PCS), 2018, pp. 323–327.
Vishwanath et al., "Motion Compensated Prediction for Translational Camera Motion in Spherical Video Coding," in Proc. IEEE 20th International Workshop on Multimedia Signal Processing (MMSP), 2018, pp. 1–4.
Marie et al., "Rate-Distortion Optimized Motion Estimation for on-the-Sphere Compression of 360 Videos," in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2021, pp. 1570–1574.

Rate-Distortion Analysis

BD-Rate for MPA in % with respect to VTM-14.2

	PSNR		WS-PSNR		S-PSNR-NN		PSNR-DYN-VP0		PSNR-DYN-VP1	
MPA-MVP	OFF	ON	OFF	ON	OFF	ON	OFF	ON	OFF	ON
SkateboardInLot	-0.97	**-1.91**	-1.06	**-1.64**	-1.02	**-1.59**	+0.14	**-0.73**	-0.15	**-0.31**
ChairliftRide	-1.56	**-3.05**	-1.22	**-2.41**	-1.27	**-2.32**	-0.58	**-1.74**	-1.54	**-3.74**
KiteFlite	**-0.15**	-0.03	+0.03	**-0.02**	+0.03	**-0.05**	+0.10	+0.04	+0.15	**-0.33**
Harbor	+2.80	+1.40	+0.27	**-0.15**	+0.17	**-0.17**	+0.01	**-0.23**	-0.11	**-0.28**
Trolley	+0.11	**-0.38**	+0.04	+0.01	+0.04	+0.01	+0.06	+0.20	+0.17	+0.24
GasLamp	+1.18	**-1.27**	-0.08	**-0.43**	-0.18	**-0.46**	+0.01	**-0.12**	-0.07	**-0.47**
Balboa	-2.05	**-3.46**	-1.87	**-3.24**	-1.84	**-3.19**	-1.79	**-3.49**	-2.65	**-4.09**
Broadway	-1.63	**-3.12**	-1.64	**-3.11**	-1.61	**-3.09**	-2.37	**-3.95**	-3.19	**-5.40**
Landing2	-3.37	**-3.96**	-2.92	**-3.40**	-2.89	**-3.42**	-3.07	**-3.47**	**-4.01**	-3.46
BranCastle2	-0.74	**-1.42**	-0.58	**-1.19**	-0.62	**-1.19**	-0.89	**-1.11**	-0.80	**-1.28**
Average	-0.64	**-1.72**	-0.90	**-1.56**	-0.92	**-1.55**	-0.84	**-1.46**	-1.22	**-1.91**

Inter Predicted Frames

Broadway, Frame 16, QP32

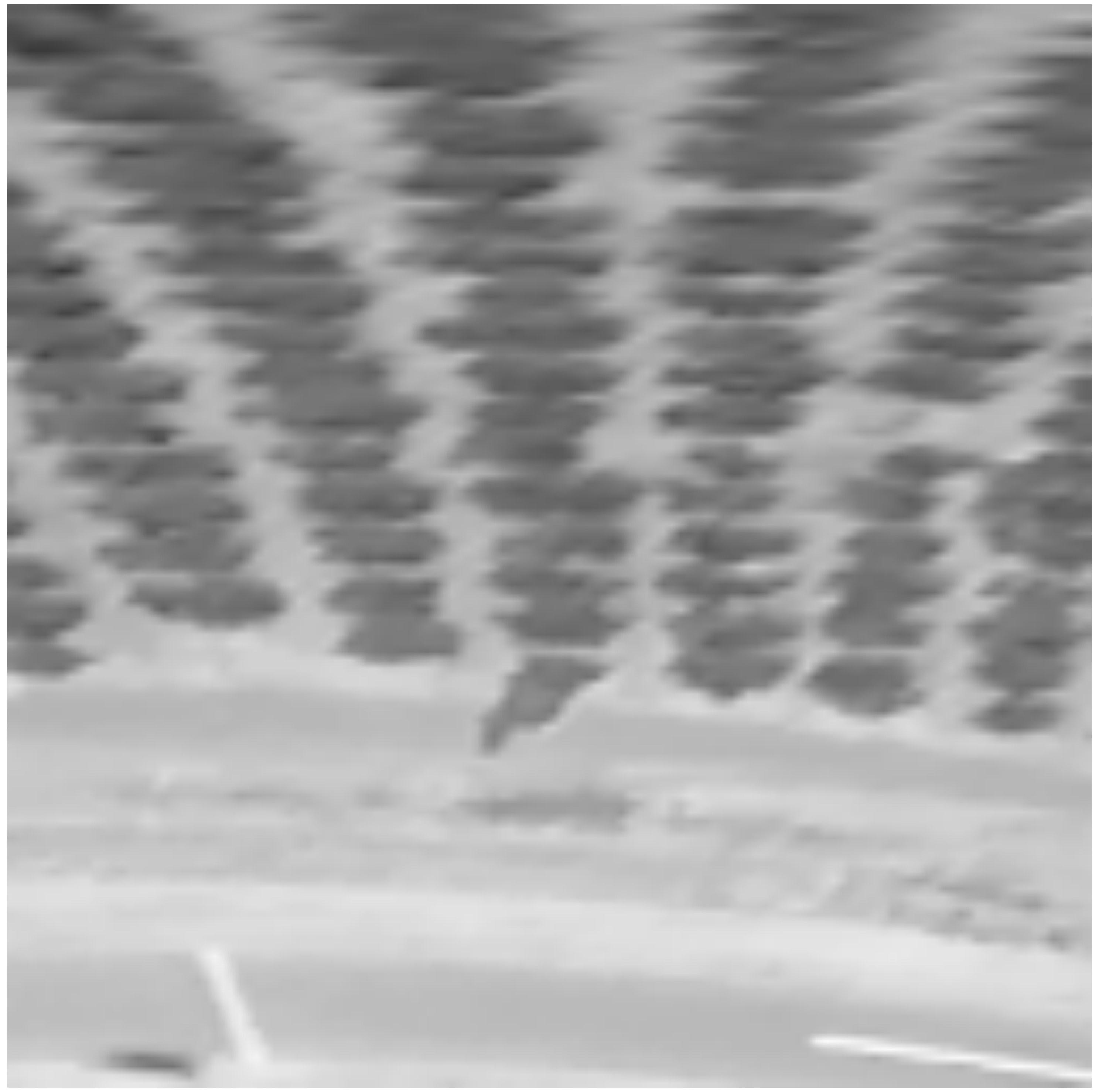
Original

VTM-14.2

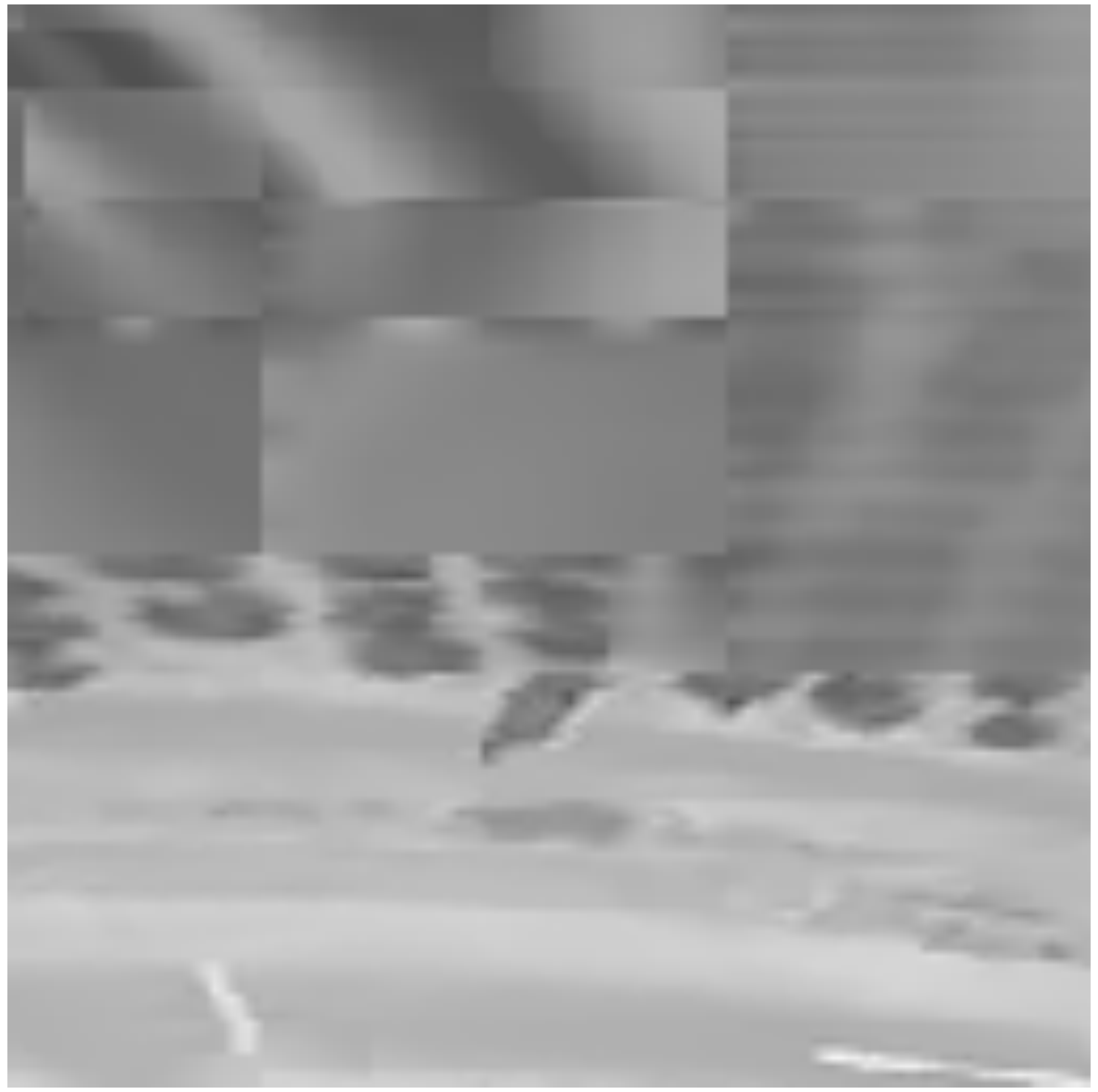
MPA + MPA-MVP

Inter Predicted Frames

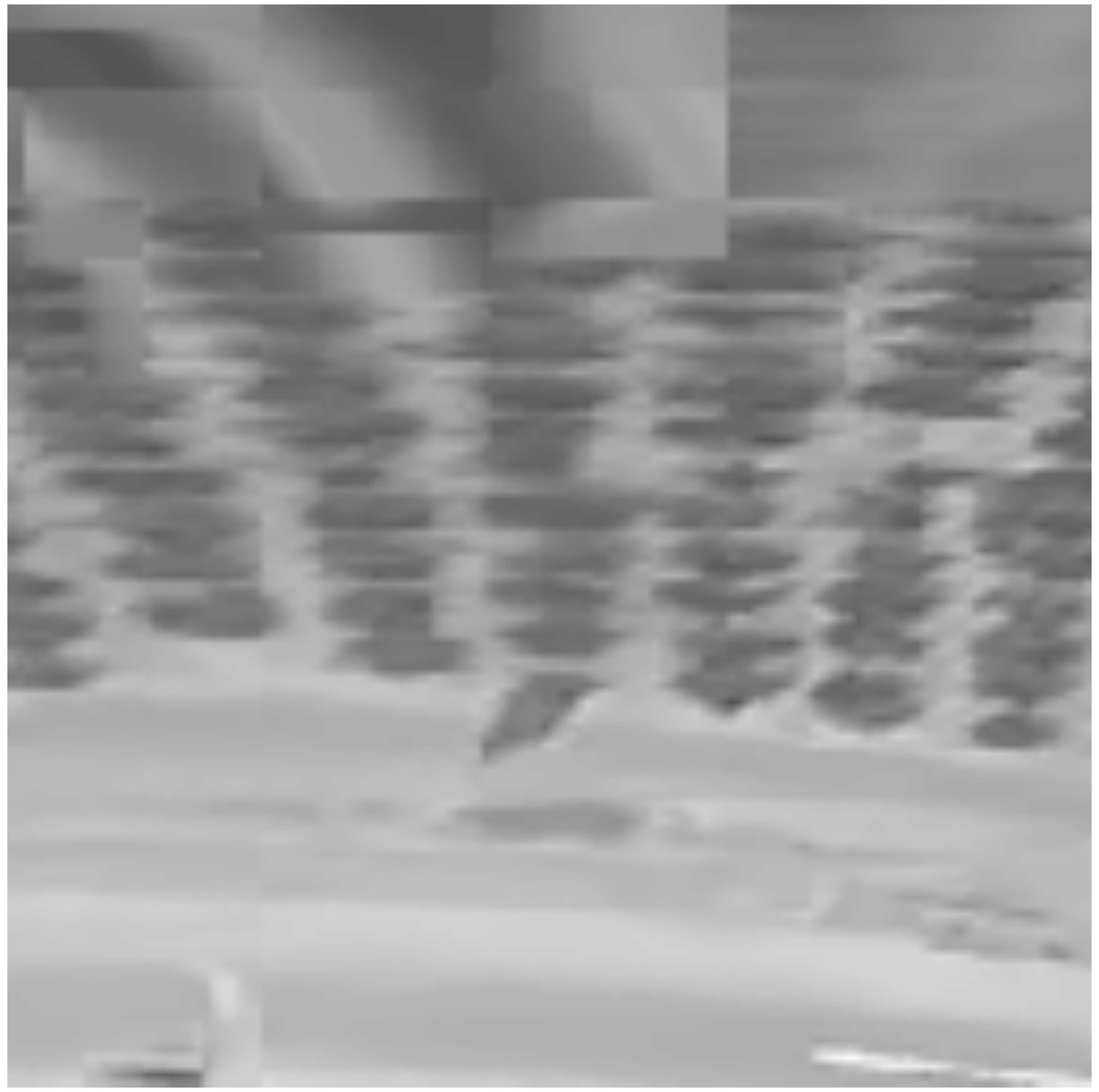
Landing2, Frame 4, QP32



Original



VTM-14.2



MPA + MPA-MVP

Outline

1. Motion-Plane-Adaptive Inter Prediction
   a. Motion-Plane-Adaptive Motion Model (MPA)
   b. Adapted Motion Vector Prediction (MPA-MVP)
2. Performance Evaluation
3. Conclusion

Conclusion

- MPA among most competitive 360-degree motion models
- MPA-MVP with additional rate savings of more than 0.60% on average
- Combined rate savings of 1.56% based on WS-PSNR on average

- Future work
  - Multi-model inter prediction
  - Affine MPA

Thank you!