Trinity College Dublin
The University of Dublin

V-SENSE

A Spatio-Angular Binary Descriptor for Fast Light Field Inter View Matching

Dr Martin Alain (alainm@tcd.ie), Prof Aljosa Smolic
IEEE International Conference on Image Processing 2020
Session 3D-02 Light-Field Image Processing and Compression

Outline

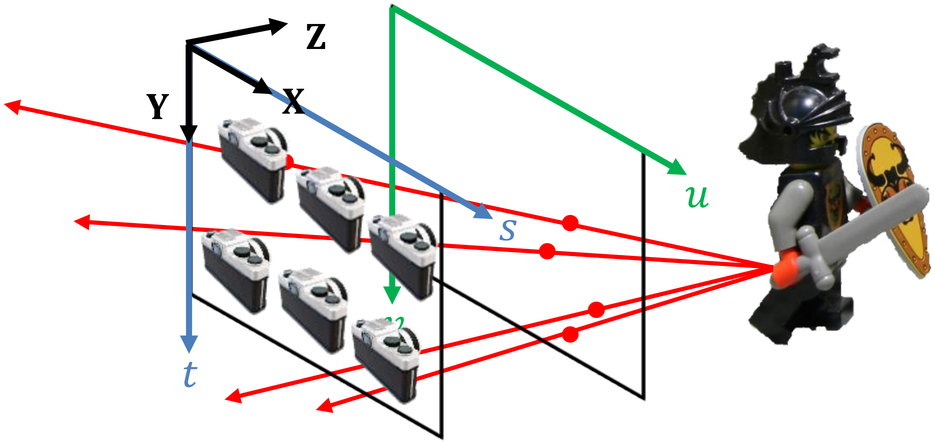
- Introduction
- Proposed binary descriptor for light fields
- Experiments and results
- Conclusion

Outline

- **Introduction**
- **Proposed binary descriptor for light fields**
- **Experiments and results**
- **Conclusion**

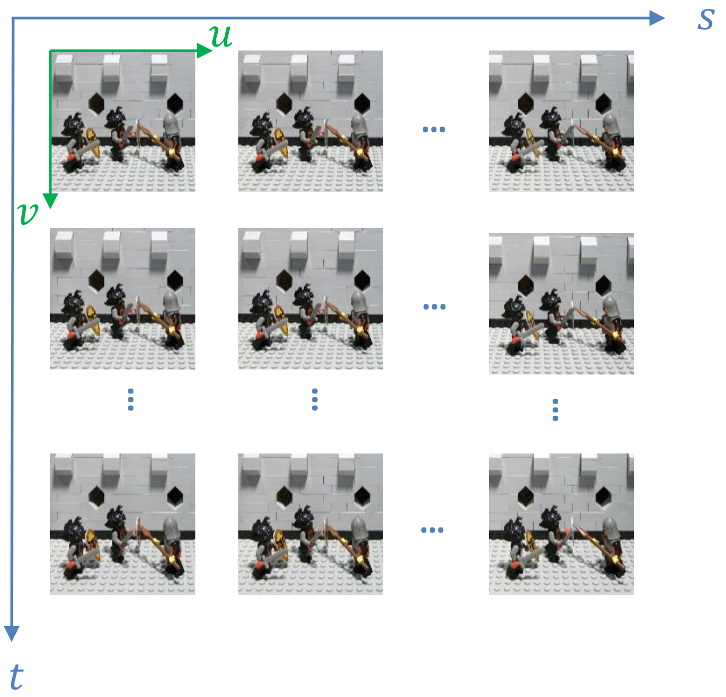
Light field imaging

2 parallel planes parameterization



the camera plane
Angular dimensions

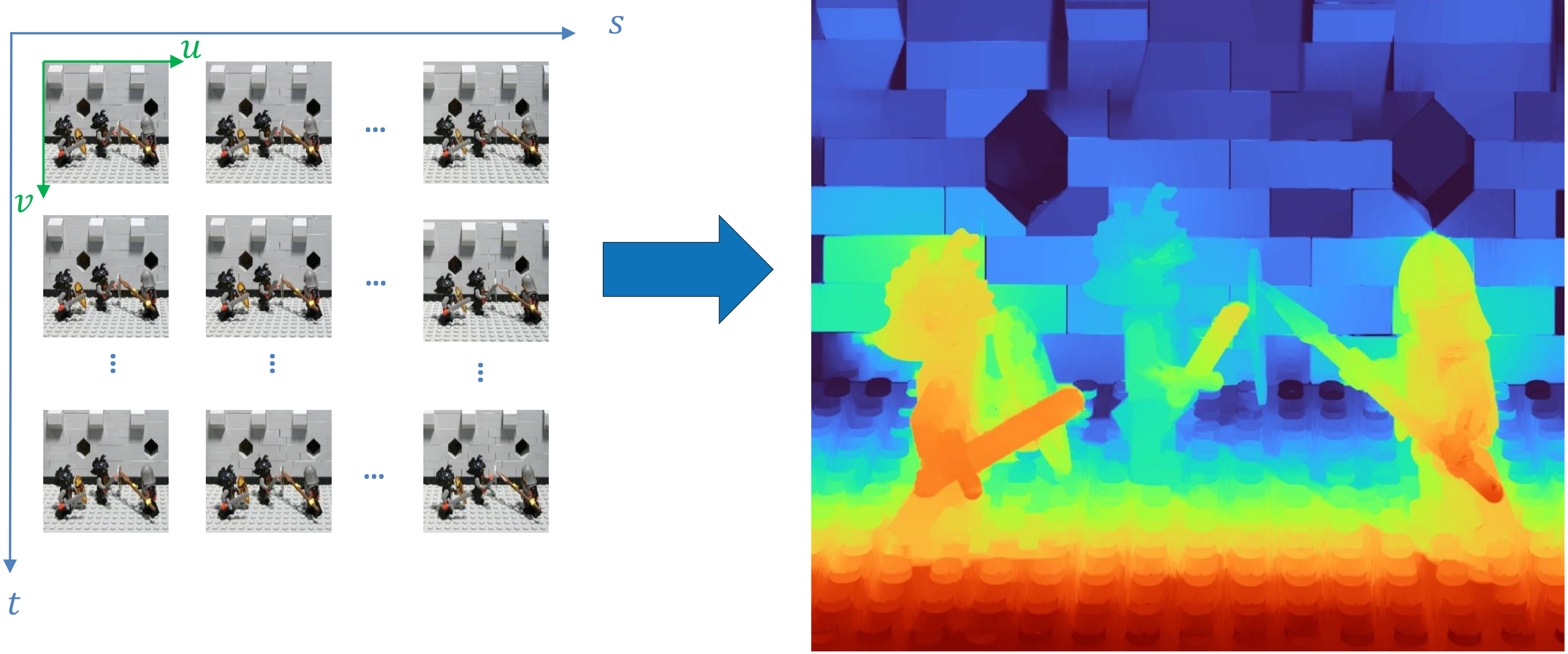
the sensor / image plane
Spatial dimensions



The light field $L = L(s, t, u, v)$

Light field applications

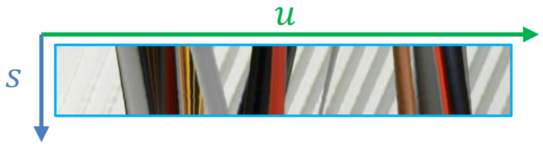
Disparity / Depth estimation

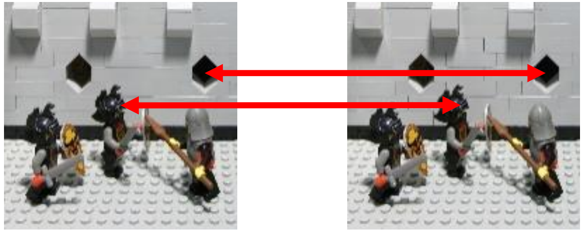


Light field applications

Disparity / Depth estimation

- **EPI line slope**
- **Defocus cues**
- **Inter-view matching**
- **+ Learning-based**



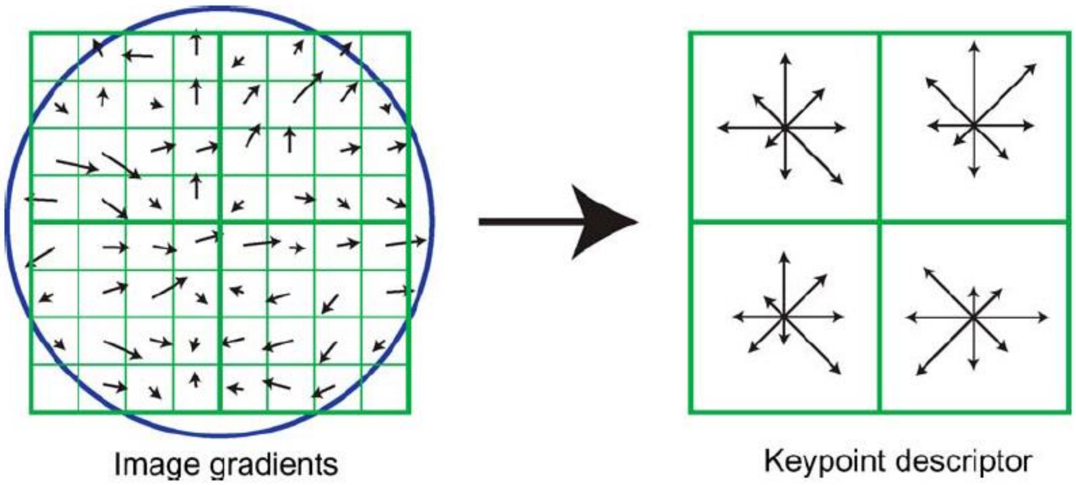


Chen, Yang, Martin Alain, and Aljosa Smolic. "Fast and accurate optical flow based depth map estimation from light fields." *Irish Machine Vision and Image Processing Conference (IMVIP)*. 2017

Descriptors

2D pixel descriptor

- **SIFT [1]**
- **DAISY [2]**
- **BOOM [3]**



[1] Lowe, David G. "Object recognition from local scale-invariant features." *Proceedings of the seventh IEEE international conference on computer vision*. Vol. 2. Ieee, 1999.

[2] Tola, Engin, Vincent Lepetit, and Pascal Fua. "Daisy: An efficient dense descriptor applied to wide-baseline stereo." *IEEE transactions on pattern analysis and machine intelligence* 32.5 (2009): 815-830.

[3] Schaffner, Michael, et al. "Towards edge-aware spatio-temporal filtering in real-time." *IEEE Transactions on Image Processing* 27.1 (2017): 265-280.

Outline

- Introduction
- **Proposed binary descriptor for light fields**
- Experiments and results
- Conclusion

SABOM

Spatio-Angular Binarised Orientation Maps

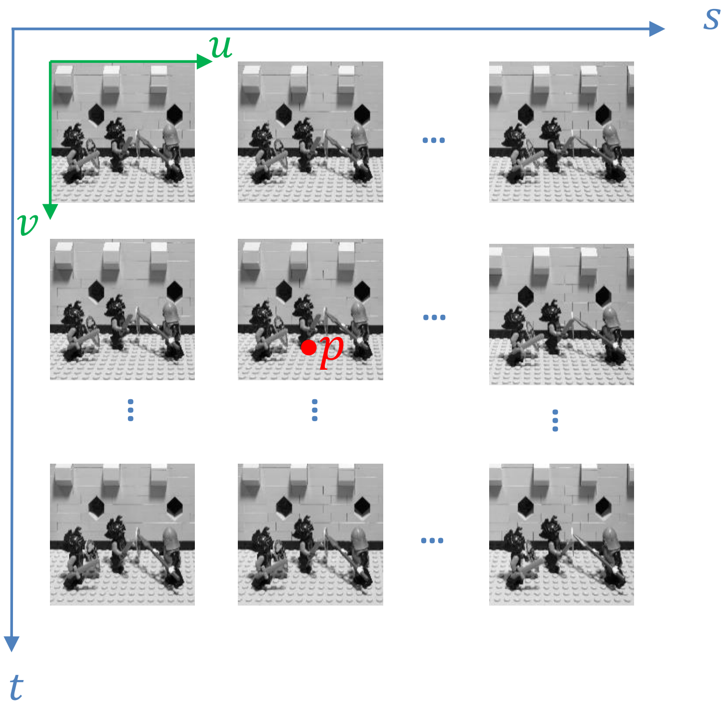
- **Extension of BOOM for light fields**
- **Exploits the 4D light field gradient**
- **Application to fast disparity map estimation**

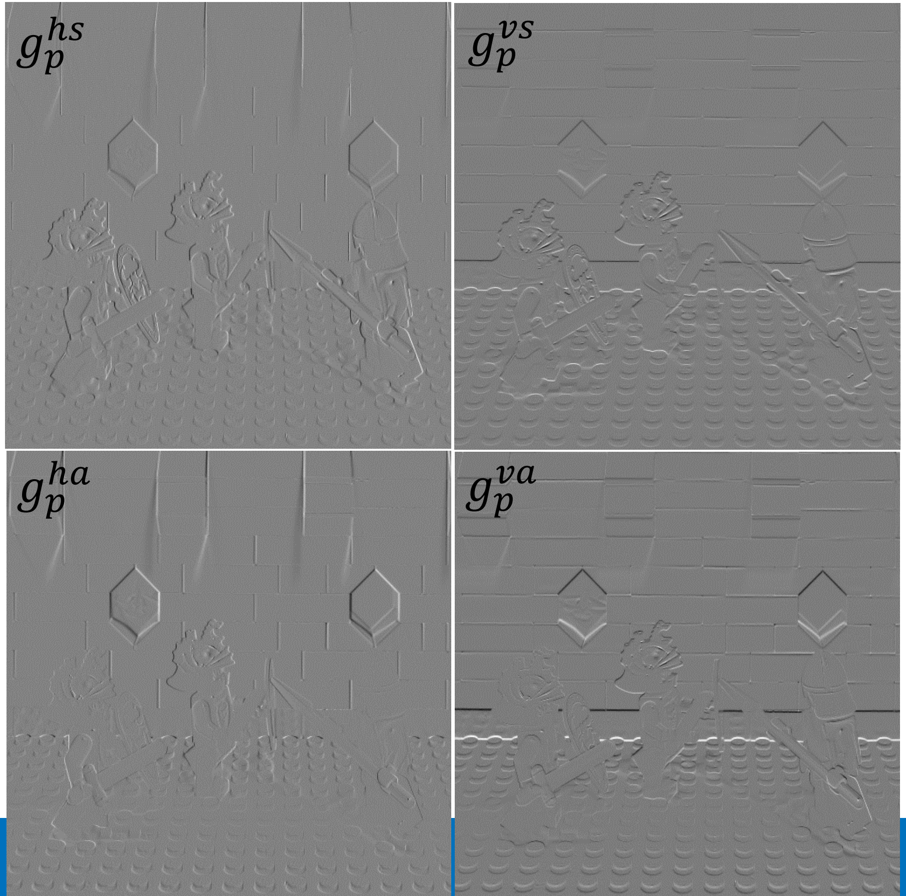
SABOM

4D Gradient computation

$$\mathbf{g}_p = [\mathbf{g}_p^s \; \mathbf{g}_p^a] = [g_p^{hs} \; g_p^{vs} \; g_p^{ha} \; g_p^{va}]$$

4D — 2D — Scalar

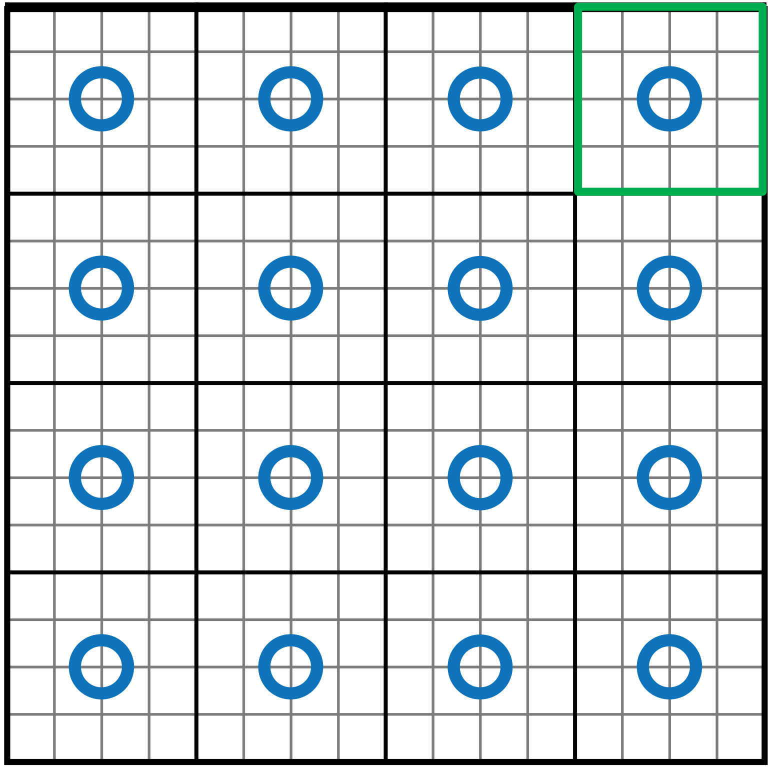




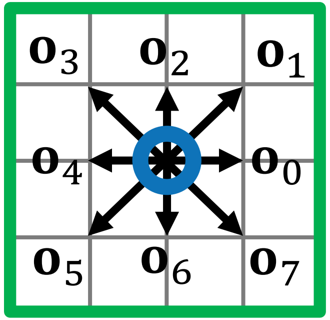
SABOM

16 × 16 Input patch $\mathbf{P}$

16 spatial bin centers



8 spatial orientations



4 × 4 Pixel block $\mathcal{B}^s$

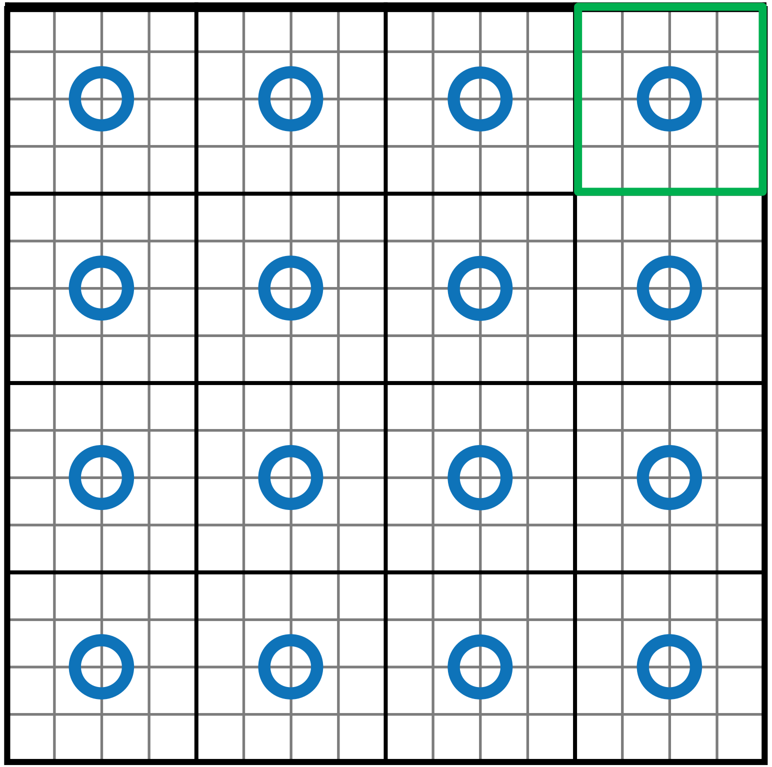
$$r_{kl}^s = \sum_{p \in \mathcal{B}_l^s} \max\left(0, \left\langle \mathbf{o}_k \middle| \mathbf{g}_p^s \right\rangle\right)$$

$$k = 0 \dots 7, l = 0 \dots 15$$

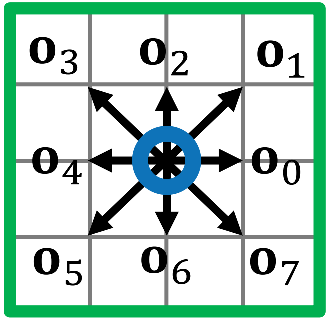
SABOM

16 × 16 Input patch **P**

16 spatial bin centers



8 spatial orientations



4 × 4 Pixel block $\mathcal{B}^s$

$$r_{kl}^s = \sum_{p \in \mathcal{B}_l^s} \max\left(0, \langle \mathbf{o}_k | \mathbf{g}_p^s \rangle\right)$$

$$k = 0 \dots 7, l = 0 \dots 15$$

Naive binarisation:

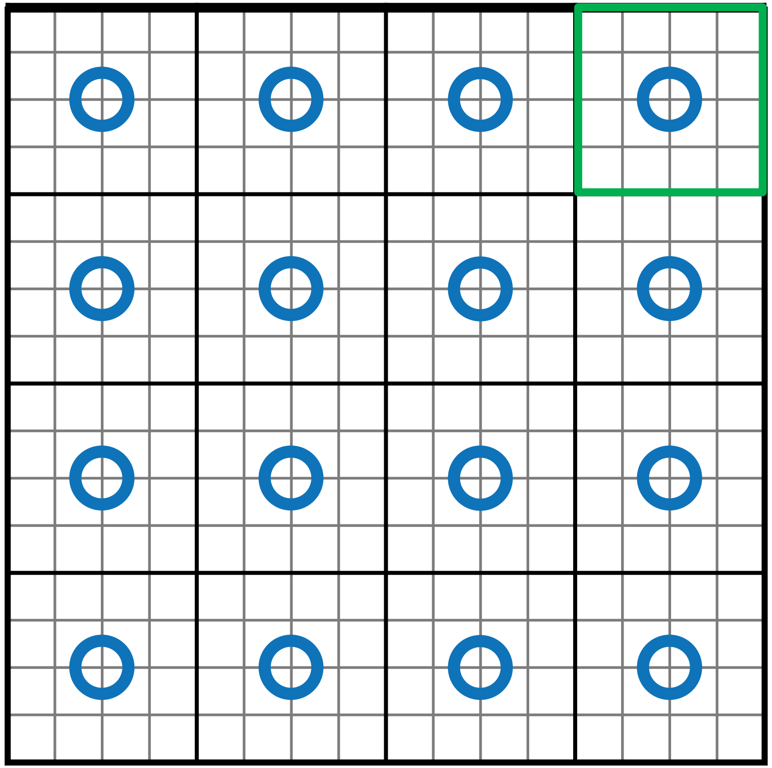
$$b_{kl}^s = \frac{r_{kl}^s}{s^s} > \epsilon$$

$$s^s = \sum_{p \in \mathbf{P}} \left\| \mathbf{g}_p^s \right\|_2$$

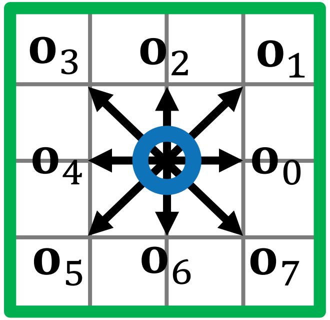
SABOM

16 × 16 Input patch **P**

16 spatial bin centers



8 spatial orientations



4 × 4 Pixel block $\mathcal{B}^s$

$$r_{kl}^s = \sum_{p \in \mathcal{B}_l^s} \max\left(0, \langle \mathbf{o}_k | \mathbf{g}_p^s \rangle\right)$$

$$k = 0 \ldots 7, l = 0 \ldots 15$$

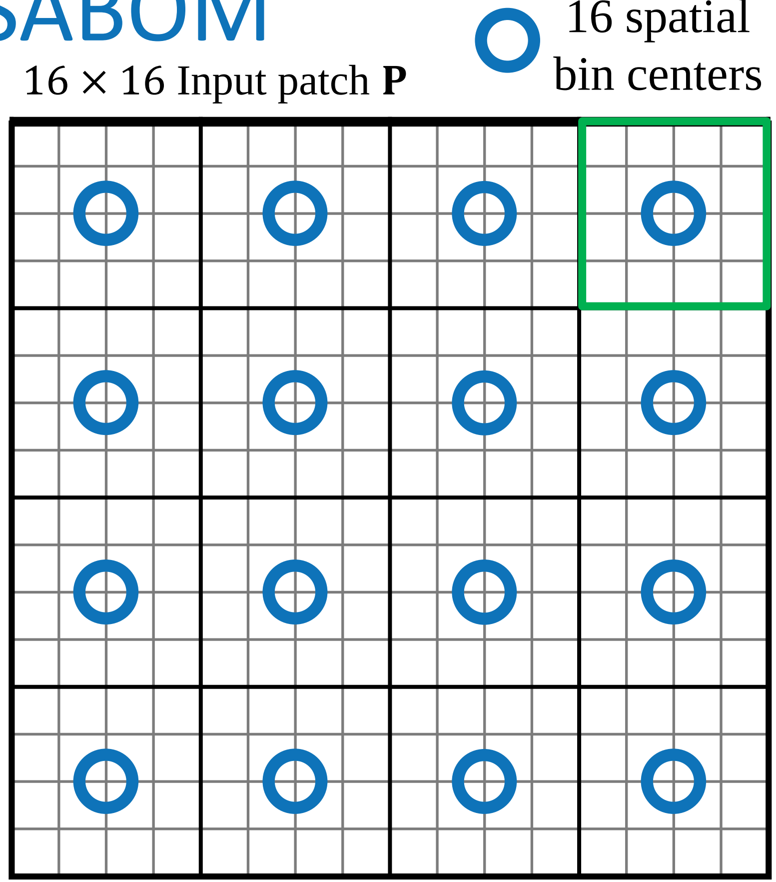
$$s^s = \sum_{p \in \mathbf{P}} \alpha \cdot \left\| \mathbf{g}_p^s \right\|_\infty + \beta \cdot \left\| \mathbf{g}_p^s \right\|_1$$

Seol, Changkyu, and Kyungwhoon Cheun. "A low complexity Euclidean norm approximation." *IEEE Transactions on Signal Processing* 56.4 (2008): 1721-1726.

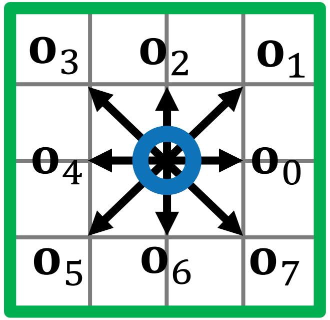
SABOM

16 × 16 Input patch **P**

16 spatial bin centers



8 spatial orientations



4 × 4 Pixel block $\mathcal{B}^s$

$$r_{kl}^s = \sum_{p \in \mathcal{B}_l^s} \max\left(0, \langle \mathbf{o}_k | \mathbf{g}_p^s \rangle\right)$$

$$k = 0 \ldots 7, l = 0 \ldots 15$$

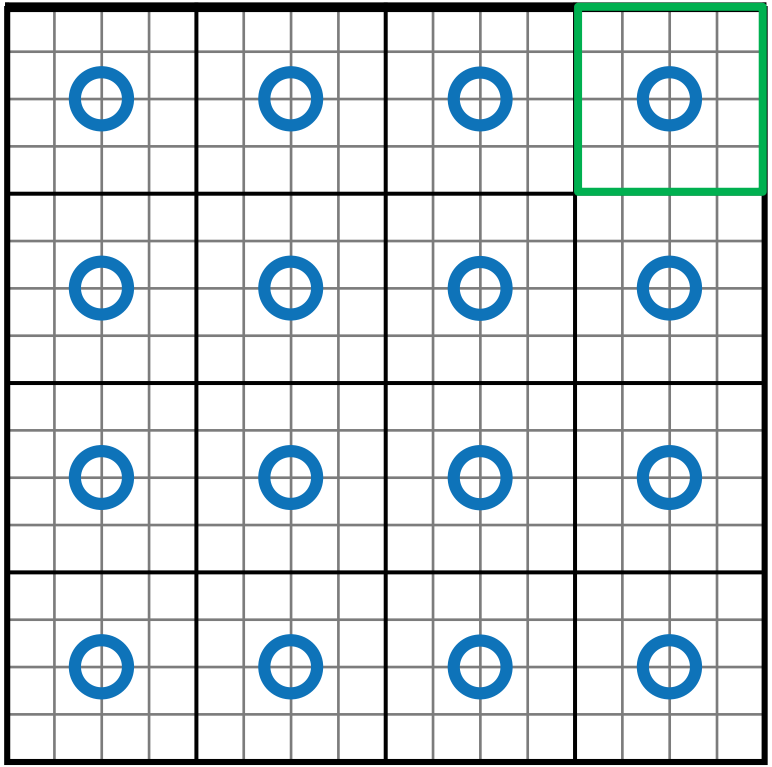
$$s^s = \sum_{p \in \mathbf{P}} \alpha \cdot \|\mathbf{g}_p^s\|_\infty + \beta \cdot \|\mathbf{g}_p^s\|_1$$

$$b_{kl}^s = r_{kl}^s \cdot \theta > s^s, \theta = \frac{1}{\varepsilon}$$

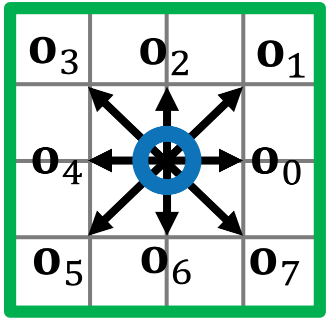
SABOM

16×16 Input patch $\mathbf{P}$

16 spatial bin centers



8 spatial orientations



4×4 Pixel block $\mathcal{B}^s$

$$r_{kl}^s = \sum_{p \in \mathcal{B}_l^s} \max\left(0, \langle \mathbf{o}_k | \mathbf{g}_p^s \rangle\right)$$

$$k = 0 \ldots 7, l = 0 \ldots 15$$

$$s^s = \sum_{p \in \mathbf{P}} \alpha \cdot \left\| \mathbf{g}_p^s \right\|_\infty + \beta \cdot \left\| \mathbf{g}_p^s \right\|_1$$

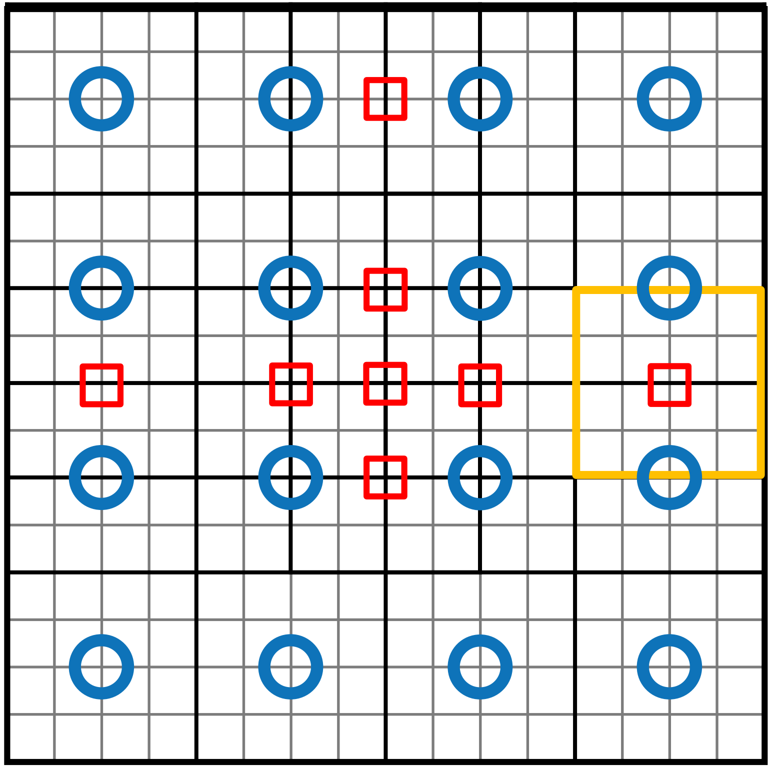
$$b_{kl}^s = \begin{cases} r_{kl}^s \cdot \theta_0^s > s^s, \text{if } k \text{ is even} \\ r_{kl}^s \cdot \theta_1^s > s^s, \text{if } k \text{ is odd} \end{cases}$$

$16 \times 8 = 128$ bits

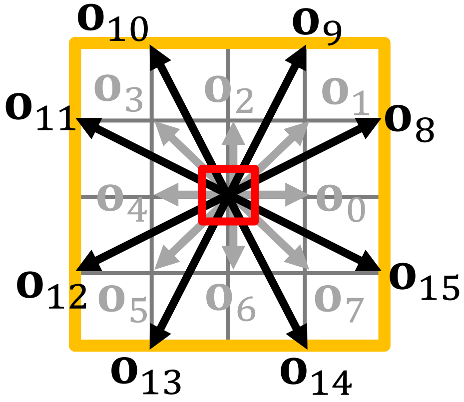
SABOM

16×16 Input patch $\mathbf{P}$

8 angular bin centers



16 angular orientations



4×4 Pixel block $\mathcal{B}^a$

$$r^a_{kl} = \sum_{p \in \mathcal{B}^a_l} \max\left(0, \langle \mathbf{o}_k | \mathbf{g}^a_p \rangle\right)$$

$$k = 0 \ldots 15, l = 0 \ldots 7$$

$$s^a = \sum_{p \in \mathbf{P}} \alpha \cdot \|\mathbf{g}^a_p\|_\infty + \beta \cdot \|\mathbf{g}^a_p\|_1$$

$$b^a_{kl} = \begin{cases} r^a_{kl} \cdot \theta^a_0 > s^a, \text{if } k \text{ is even and } k < 8 \\ r^a_{kl} \cdot \theta^a_1 > s^a, \text{if } k \text{ is odd and } k < 8 \\ r^a_{kl} \cdot \theta^a_2 > s^a, \text{if } k \geq 8 \end{cases}$$

$8 \times 16 = 128$ bits

SABOM

Implementation details

- **C++ implementation based on OpenCV (github.com/V-Sense)**
- **CPU parallelisation over light views with OpenMP**
- **Use of integer arithmetic, no division**
- **Efficient dense scan implementation, re-using gradient response computed for neighbours, gradient is precomputed for the whole image**
- **Distance between descriptor computed using** `__popcount64` **instruction**
- **Binarisation implemented with bitshift**
- $\alpha = 6, \beta = 2$
- $\theta_0^s = 512, \theta_1^s = 256, \theta_0^a = \theta_1^a = \theta_2^a = 512$

SABOM

Application to disparity map estimation: LF CPM + PF

Schaffner, Michael, et al. "Towards edge-aware spatio-temporal filtering in real-time." *IEEE Transactions on Image Processing* 27.1 (2017): 265-280.

Chen, Yang, Martin Alain, and Aljosa Smolic. "Fast and accurate optical flow based depth map estimation from light fields." *Irish Machine Vision and Image Processing Conference (IMVIP)*. 2017.

SABOM

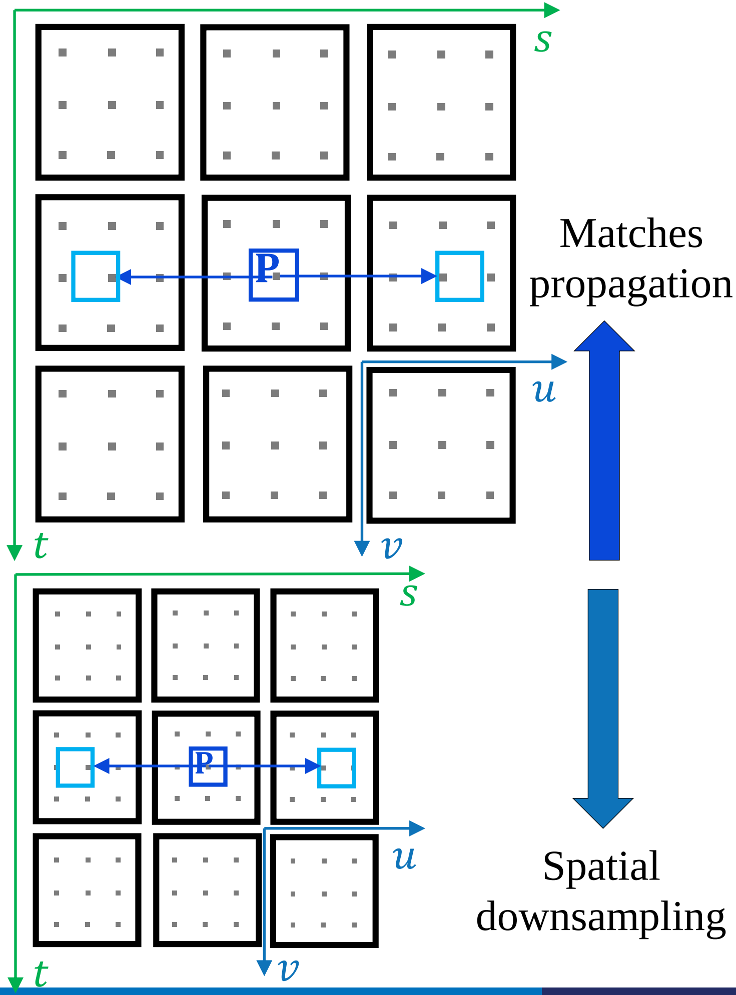
Application to disparity map estimation: LF CPM + PF

CPM

- **Gaussian pyramid computed for all views**
- **Starting from coarse level, matches between descriptors are computed on a sparse grid**
- **Estimates are propagated to the next level**
- **At finest level, foward backward check is performed**
- **Check threshold θ_{desc} is used to remove outliers**

PF

- **Apply permeability filter to obtain dense disparity map from sparse matches**



Outline

- Introduction
- Proposed binary descriptor for light fields
- **Experiments and results**
- Conclusion

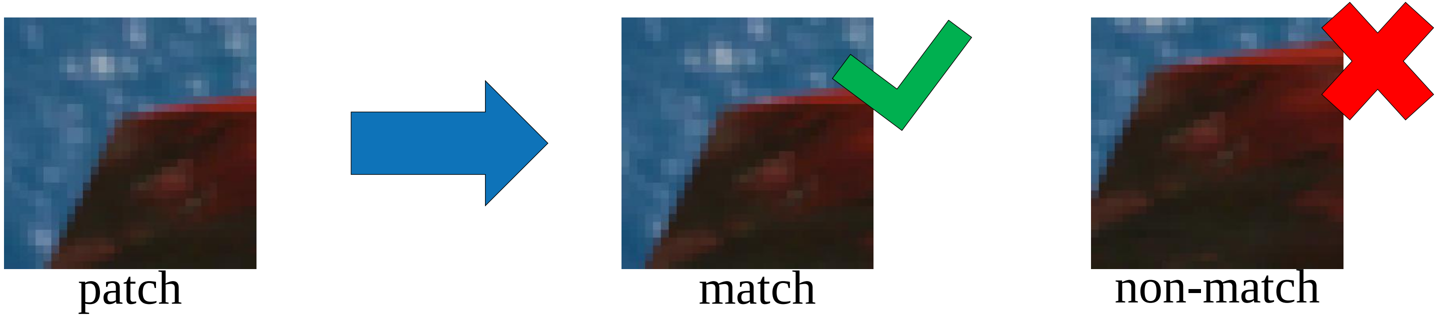
Datasets

- **HCI benchmark - lightfield-analysis.uni-konstanz.de**
  - 24 light fields
  - Disparity range $[-3.6, 3.5]$
- **INRIA Dense - clim.inria.fr/Datasets/InriaSynLF/index.html**
  - 39 light fields
  - Disparity range $[-5.5, 5.3]$
- **Stanford Gantry - lightfield.stanford.edu**
  - 12 light fields

ROC performance

Procedure

- **Split all light fields views in 32x32 patches and add AWGN with** $\sigma_n = 0.1$
- **Create pair of ground truth matches using ground truth disparity maps**
- **Create a non-match pair for each match by randomly sampling a patch within a 5 pixels radius around the ground truth match**

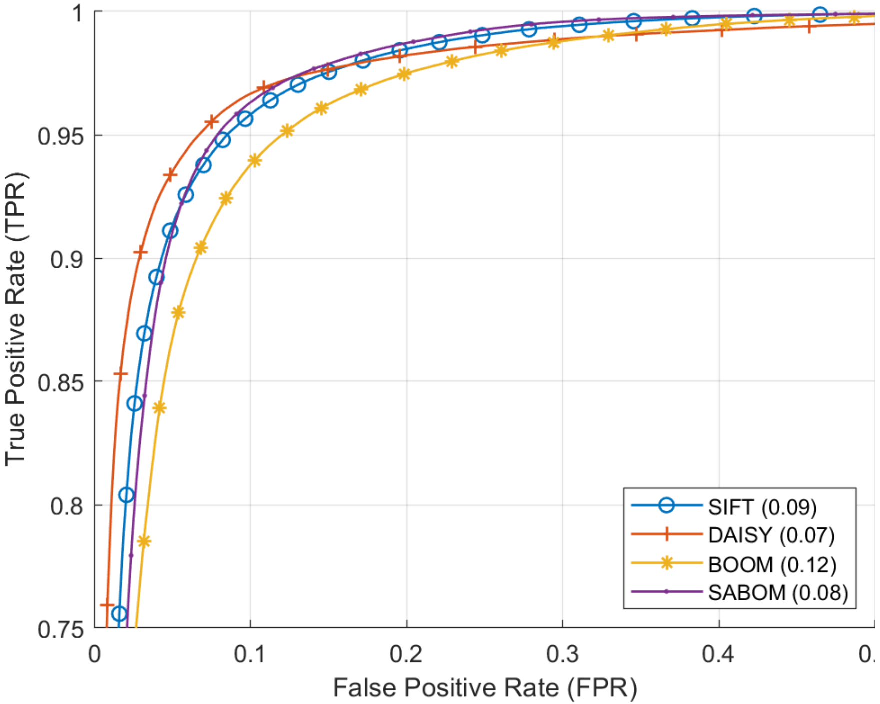


- **Compute descriptor distance between pairs of patches**
- **Use a sweeping threshold to to compute ROC performance curve**

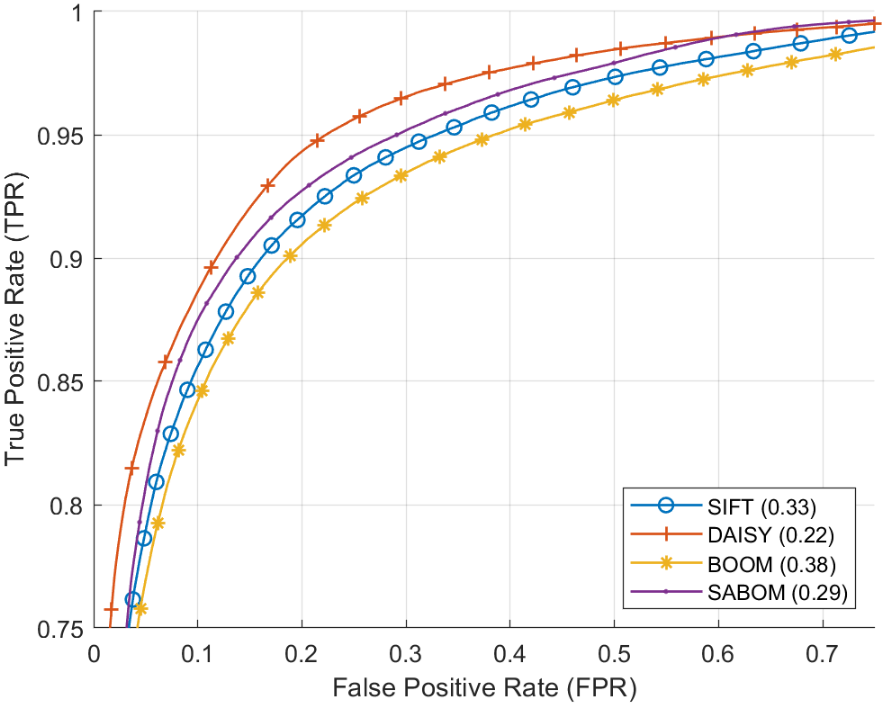
ROC performance

ROC curves

HCI dataset



INRIA Dense dataset



ROC performance

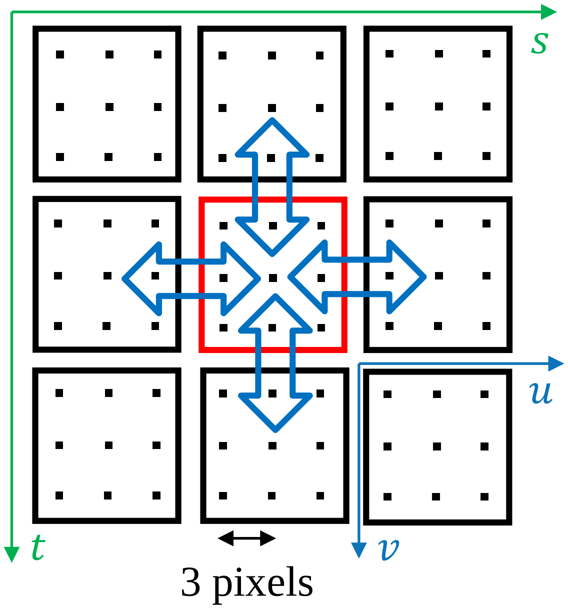
Computation time

Descriptor	Single descriptor (μs)	Patch distance (ms)
SIFT	0.635	0.152
DAISY	0.254	0.132
BOOM	0.130	0.020
SABOM (proposed)	0.176	0.021

Disparity map estimation

Procedure

- Estimated on the center view, using 3x3 surrounding views
- CPM grid spacing: 3 pixels
- Sweep θ_{desc} range so that matching density ranges from 0 to 100%
- Quality evaluated using MSE



Disparity map estimation

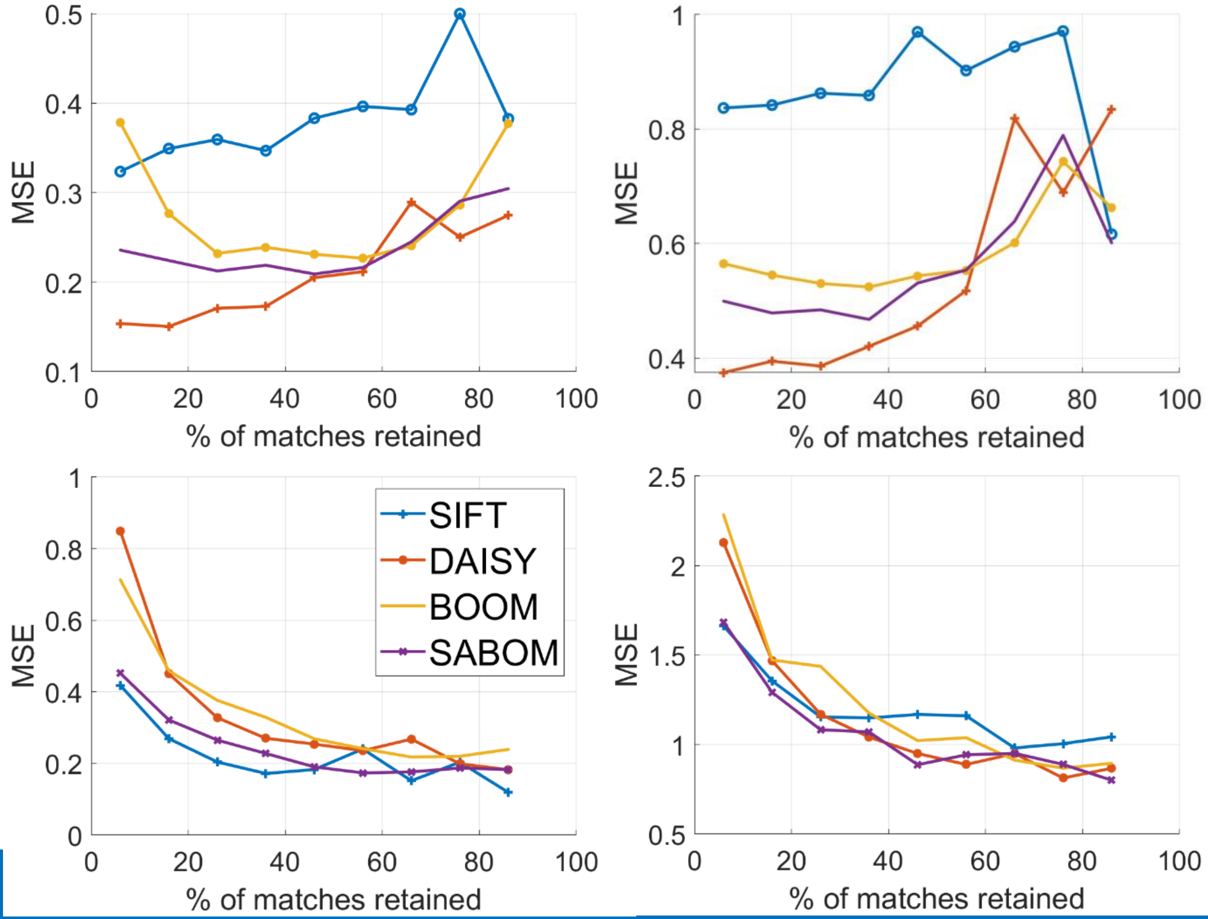
Accuracy

HCI CPM

INRIA CPM

HCI CPM + PF

INRIA CPM + PF



Disparity map estimation

Computation time (in seconds)

	CPM Pyramid Descriptor				CPM Matching	PF
	Level 2	Level 1	Level 0	Total		
SIFT	0.131	0.496	1.952	2.579	0.527	0.569
DAISY	0.054	0.219	0.843	1.116	0.423	
BOOM	0.025	0.099	0.382	0.520	0.325	
SABOM	0.059	0.236	0.564	0.893	0.367	

Disparity map estimation

Qualitative results

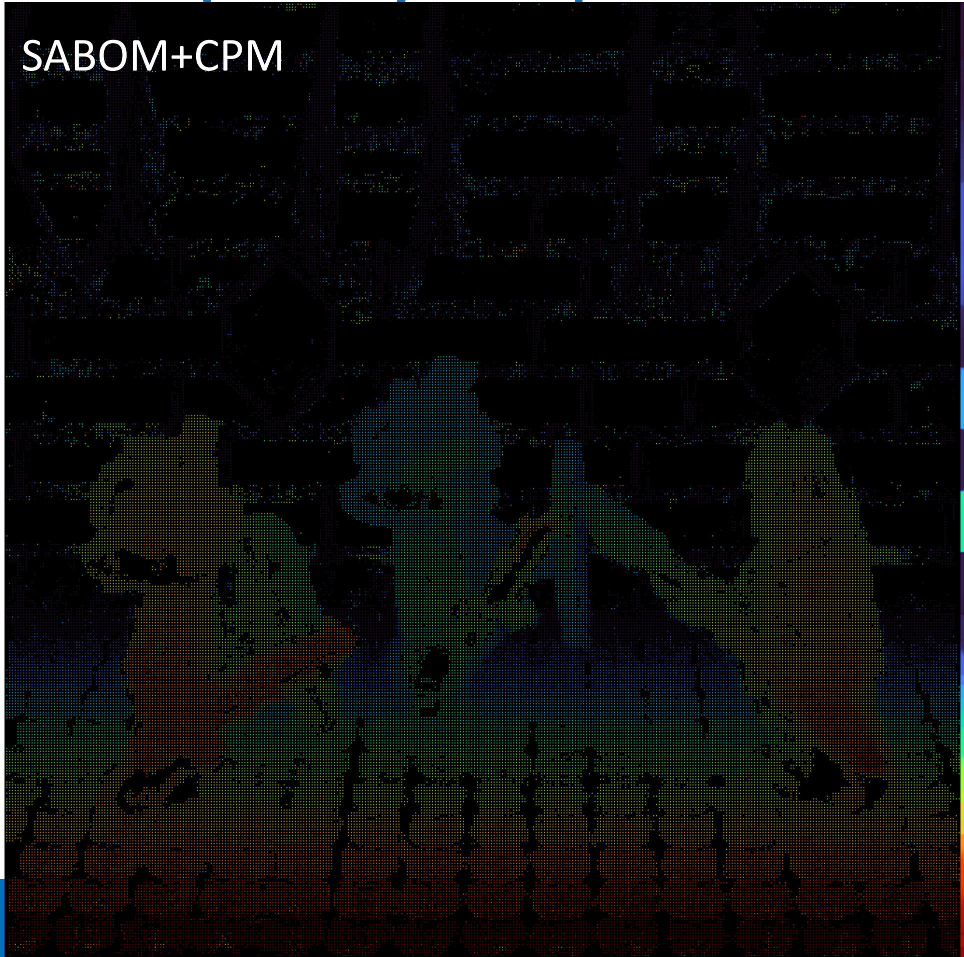
More results online:

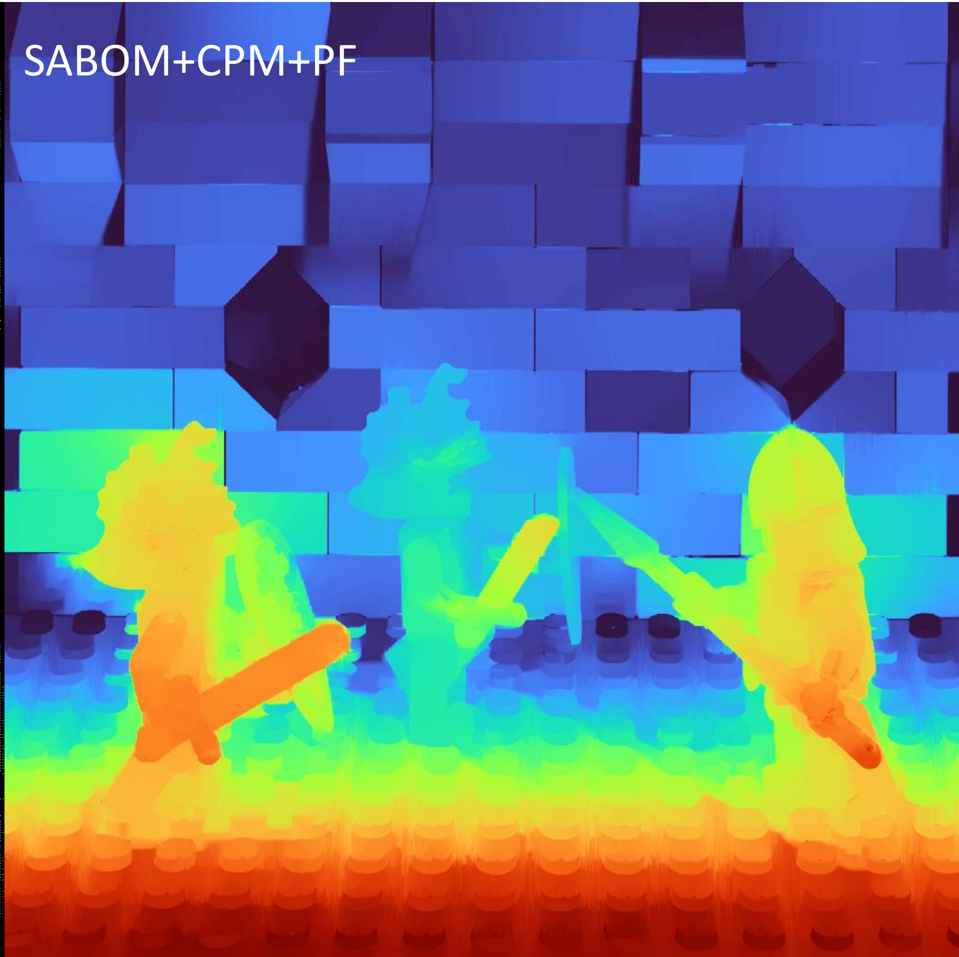
v-sense.scss.tcd.ie/research/a-spatio-angular-binary-descriptor-for-fast-light-field-inter-view-matching

Code available:

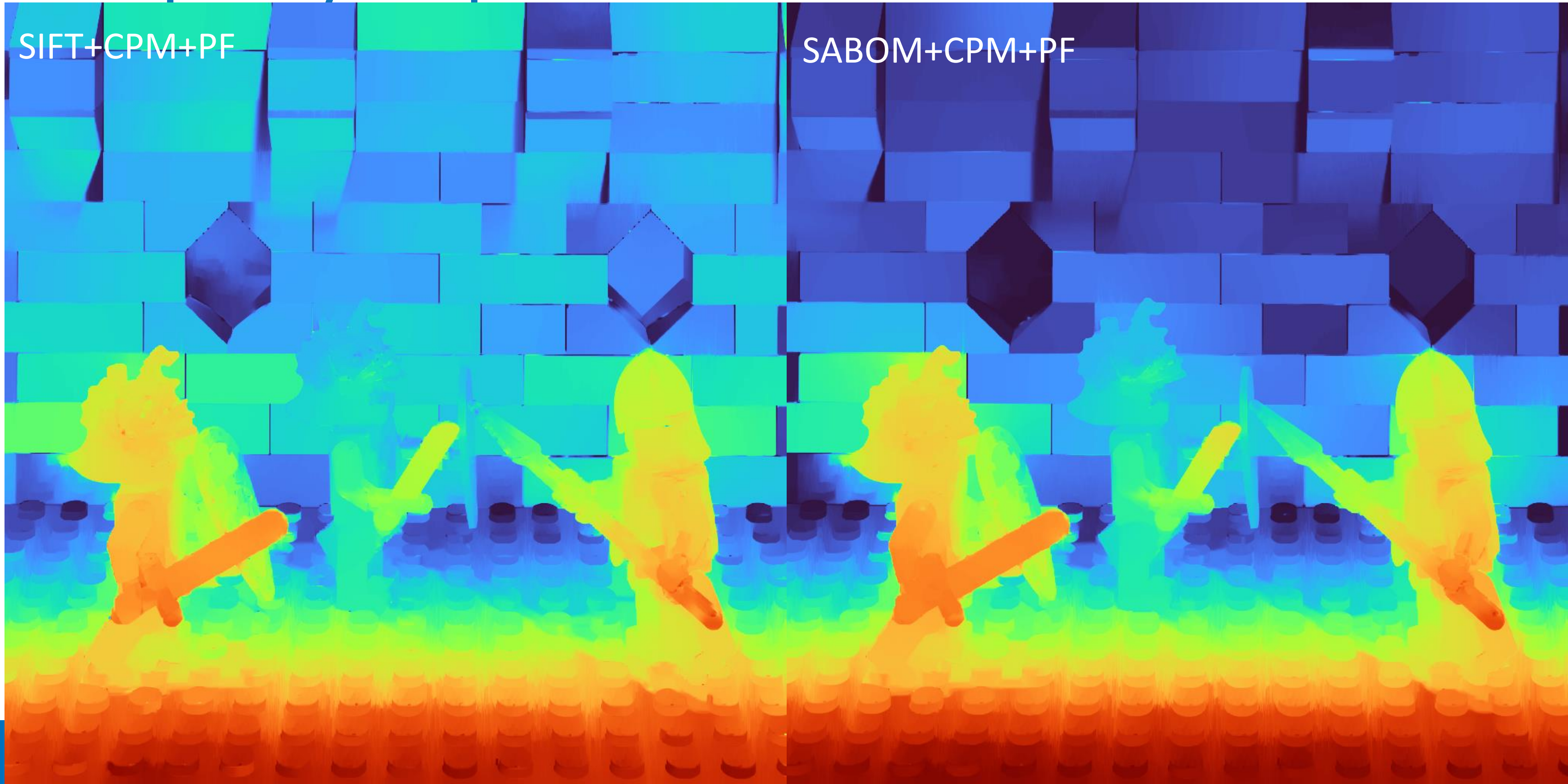
github.com/V-Sense

Disparity map estimation





Disparity map estimation



Outline

- Introduction
- Proposed binary descriptor for light fields
- Experiments and results
- **Conclusion**

Conclusion

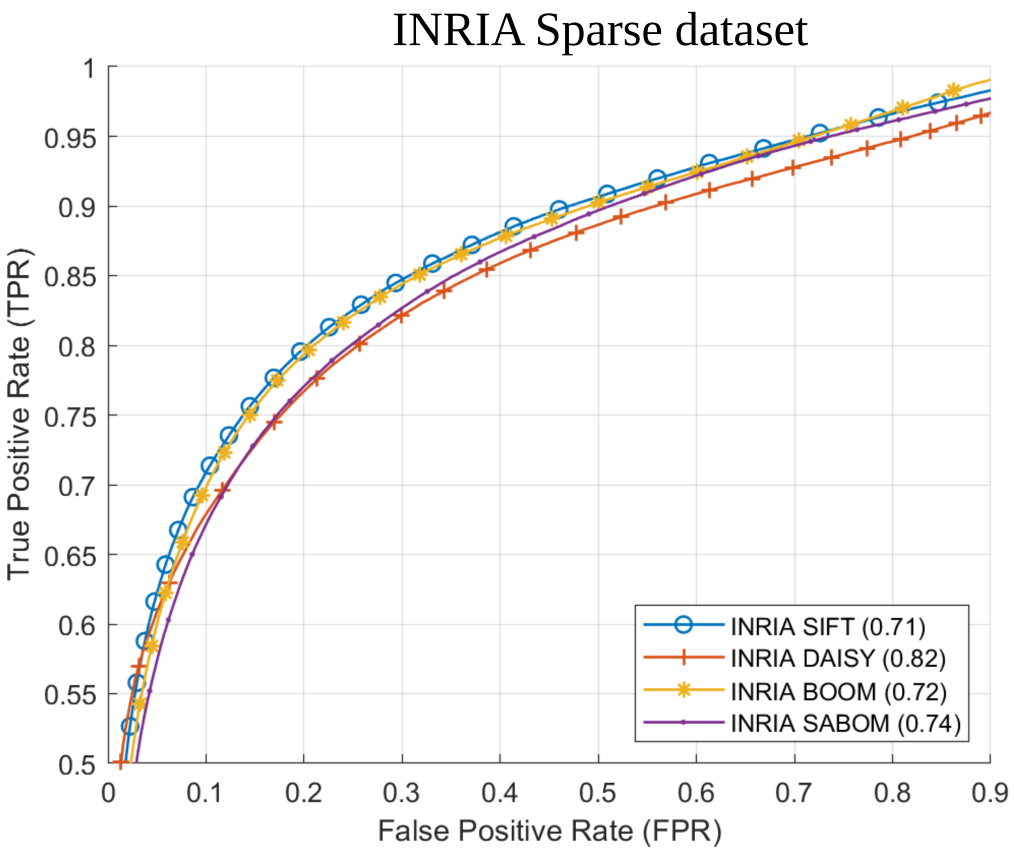
Summary

- A novel binary descriptor is introduced for fast and accurate light field inter fiew matching
- The light field spatial and angular gradients are exploited
- Experiments demonstrate improved performance compared to existing descriptors in terms of ROC performance and when applied to disparity estimation

Conclusion

Limitations

- Not suited for sparsely sampled light fields



Conclusion

Future work

- Address limitations
- Use in existing framework for more accurate disparity estimation
- Include in other applications
  - LFBM5D
  - Light field colour correction

Dąbała, Łukasz, et al. "Efficient Multi-image Correspondences for On-line Light Field Video Processing." *Computer Graphics Forum*. Vol. 35. No. 7. 2016.

Alain, Martin, and Aljosa Smolic. "Light field denoising by sparse 5D transform domain collaborative filtering." *2017 IEEE 19th International Workshop on Multimedia Signal Processing (MMSP)*. IEEE, 2017.

Alain, Martin, and Aljosa Smolic. "Light field super-resolution via LFBM5D sparse coding." *2018 25th IEEE international conference on image processing (ICIP)*. IEEE, 2018.

Grogan, Mairead and A. Smolic, “L2 based colour correction for light field arrays”, in Proc. ACM CVMP, 2019.

Matysiak, Pierre, et al. "High quality light field extraction and post-processing for raw plenoptic data." *IEEE Transactions on Image Processing* 29 (2020): 4188-4203.

Many Thanks!

- alainm@tcd.ie

- https://v-sense.scss.tcd.ie/research/light-field-imaging/
- https://v-sense.scss.tcd.ie/research/a-spatio-angular-binary-descriptor-for-fast-light-field-inter-view-matching/
- https://github.com/V-Sense/