Event cameras

presented by Emanuela Haller
Bucharest Computer Vision Reading Group

Contents

- Introduction to event cameras
- Semi-dense 3D structure estimation [1]
- Optical flow and intensity estimation [2]

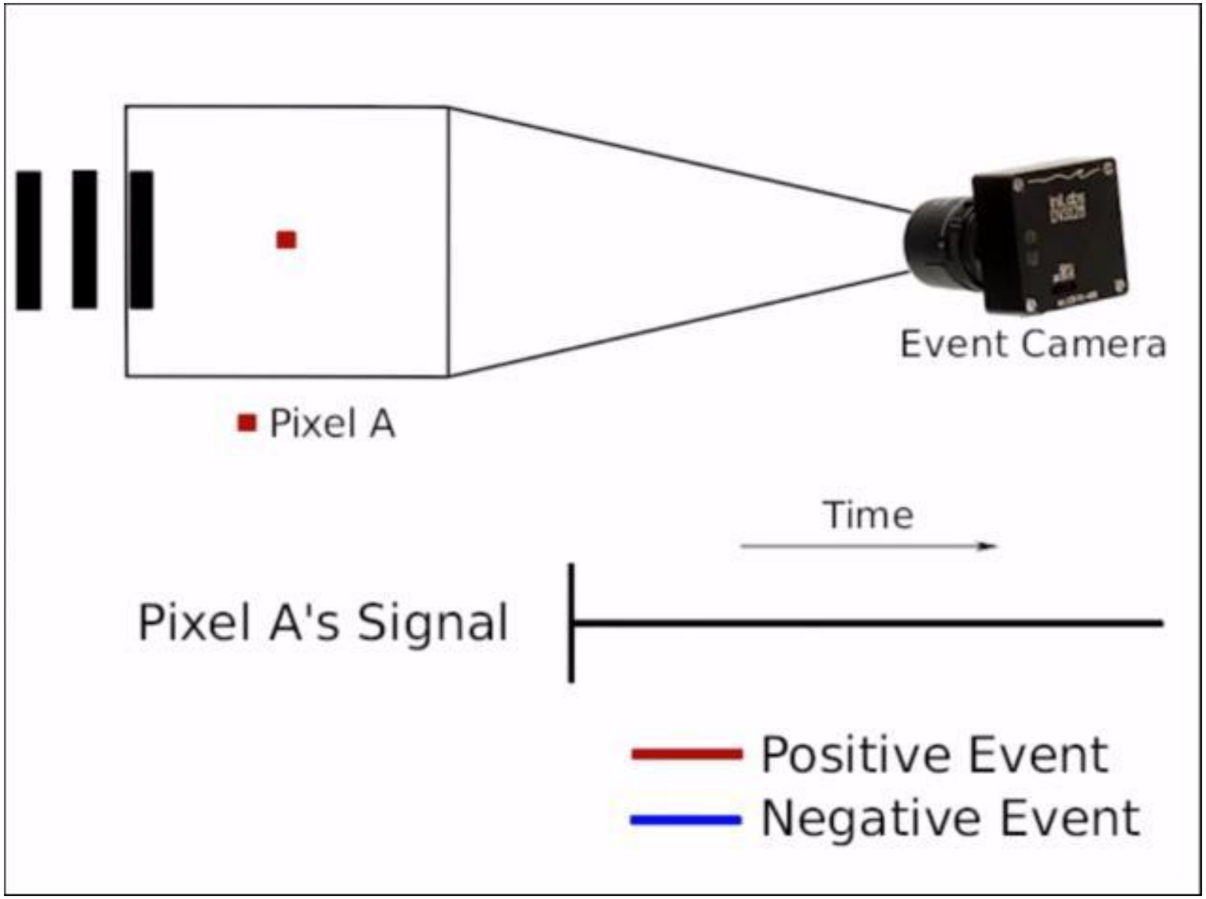
[1] H Rebecq, G. Gallego and D. Scaramuzza. "EMVS: Event-based Multi-View Stereo", BMVC 2016.

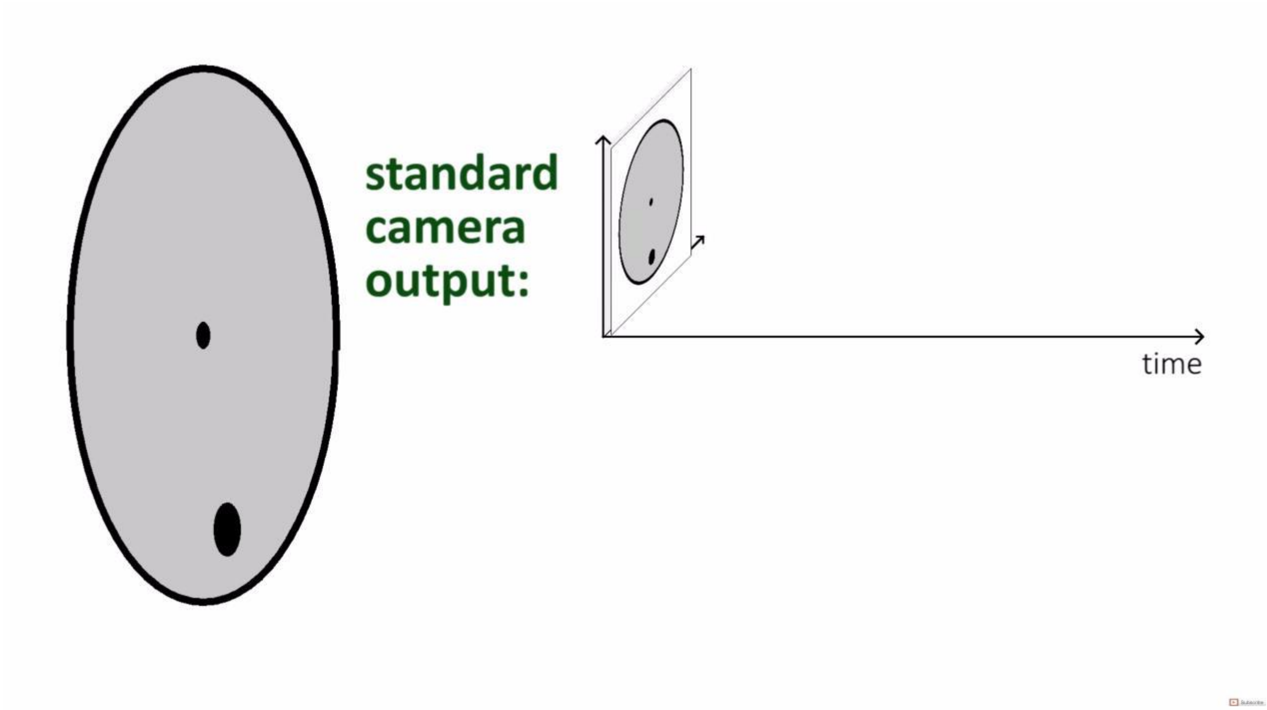
[2] P. Bardow, A.J. Davison and S. Leutenegger. "Simultaneous Optical Flow and Intensity Estimation from an Event Camera", CVPR 2016.

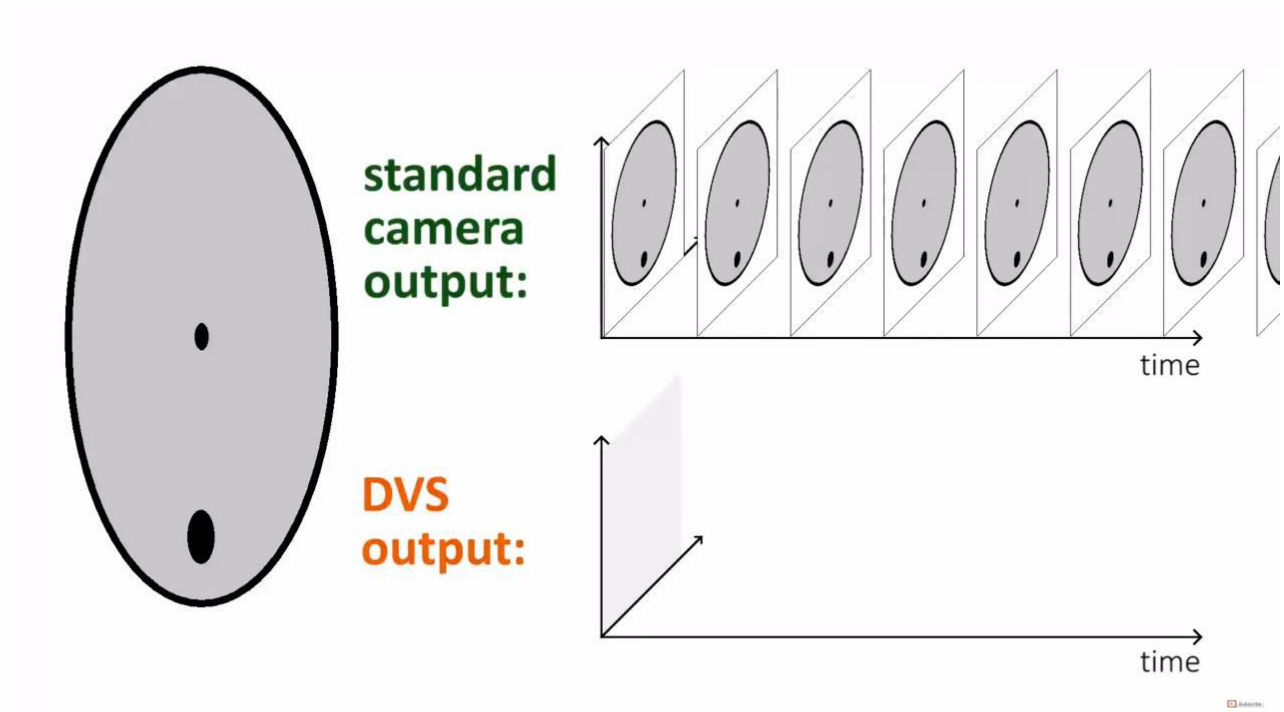
Event cameras

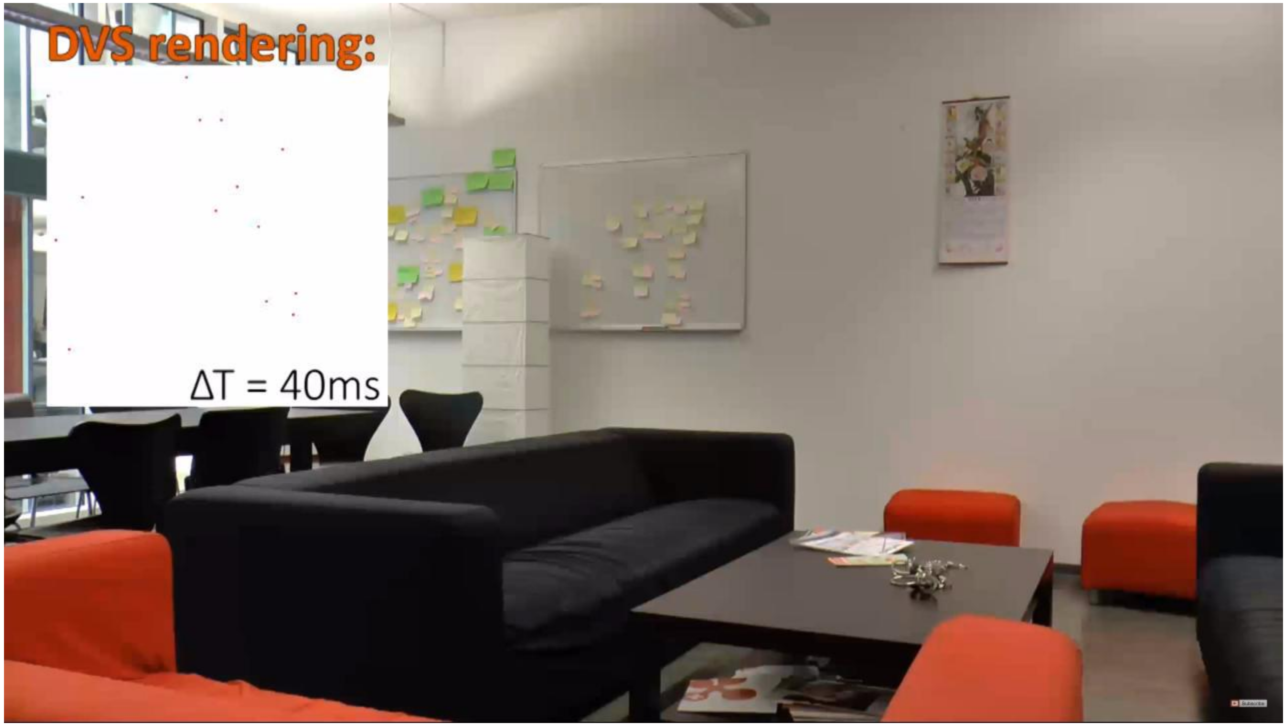
- Silicon retina
- Pixel-level changes
- Asynchronous events (microsecond resolution)
  - OFF events – encode decreasing brightness
  - ON events – encode increasing brightness

- DVS – Dynamic Vision Sensor


Event Camera
Pixel A
Time
Pixel A's Signal
Positive Event
Negative Event


standard
camera
output:
time


standard
camera
output:
time
DVS
output:
time


DVS rendering:
ΔT = 40ms

DVS

- No redundant information
- Low latency
- High dynamic range
- No motion blur
- Reduced power consumption

- Paradigm shift

- DVS128

  (Dynamic Vision Sensor)

array size	128 x 128
pixel size	40 x 40 μm
dynamic range	120 dB
latency	15 μs
bandwidth	1 M events / sec
price	3000 CHF

[3] P.Lichtsteiner, C. Posch and T. Delbruck. "A 128x128 120dB 15μs latency asynchronous temporal contrast vision sensor". JSSC 2008

http://inilabs.com/

- DAVIS240

(Dynamic and **Active-pixel Vision Sensor**)

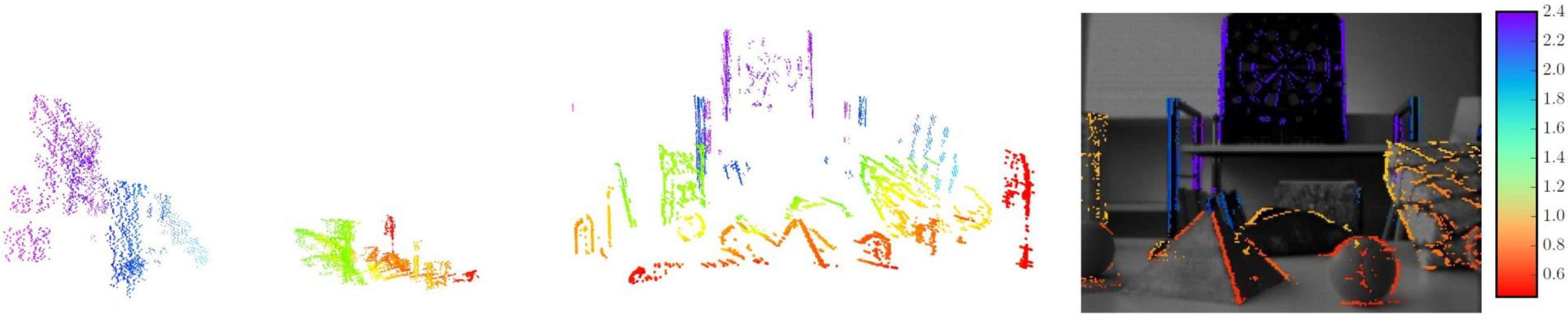
array size	240 x 180
pixel size	18.5 x 18.5 μm
dynamic range	130 dB
latency	3 μs
bandwidth	3 M events / sec
ASP(active pixel sensor)	55 dB, 30 fps

[4] C. Brandli, R. Berner, M. Yang, S.C. Liu and T. Delbruck. "A 240x180 130dB 3μs Latency Global Shutter Spatiotemporal Vision Sensor". JSSC 2014

http://inilabs.com/

EMVS: Event-based Multi-View Stereo [1]

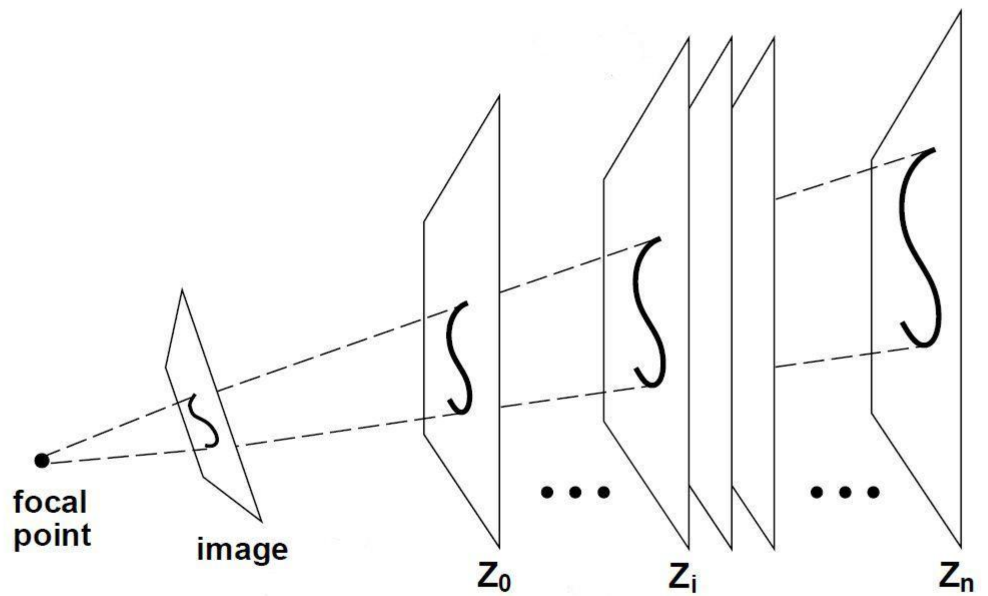
- Semi-dense 3D structure



(a) Side view. (b) Front view. (c) Projection on a frame.

- MVS – the problem of 3D structure estimation, of a static scene, from a collection of images taken from known viewpoints

A Space-Sweep Approach to True Multi-Image Matching [5]



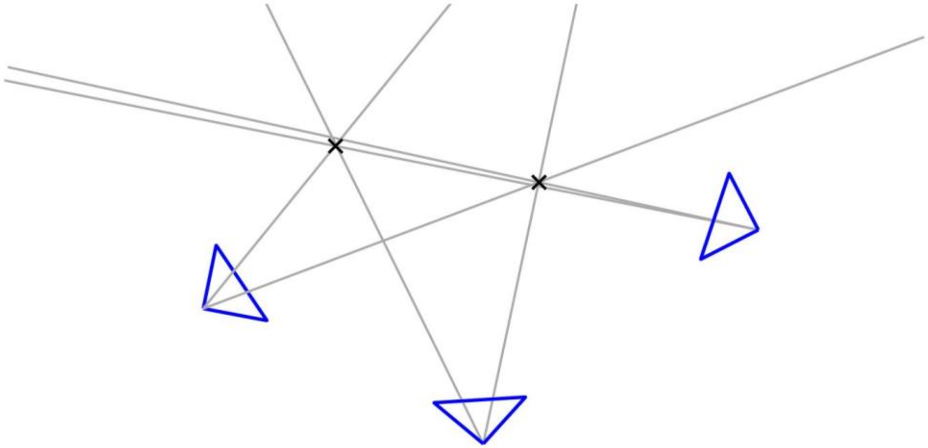
- Binary edge images

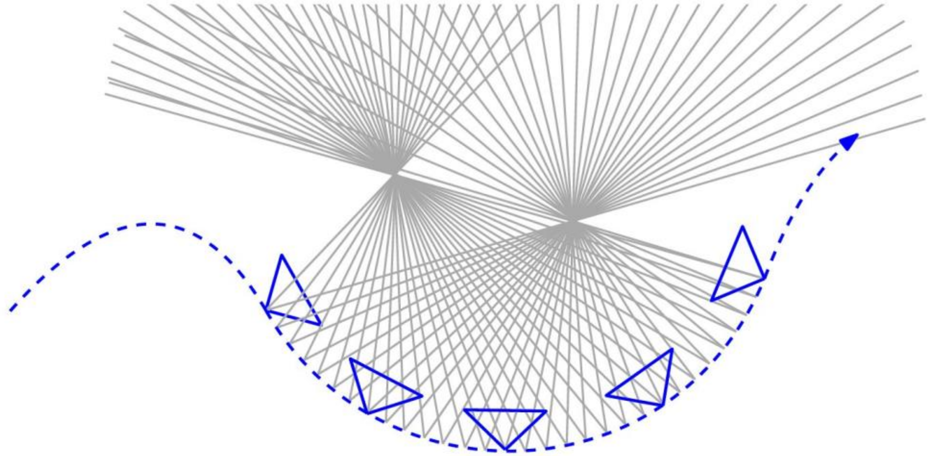
[5] R.T. Collins. "A space-sweep approach to true multi-image matching". CVPR 1996

Event-based Space-Sweep

1. Back-project event locations as rays through a defined volume of interest
2. Record the number of rays that pass through each voxel
3. Determine locations of 3D points

Frame-based vs. Event-based

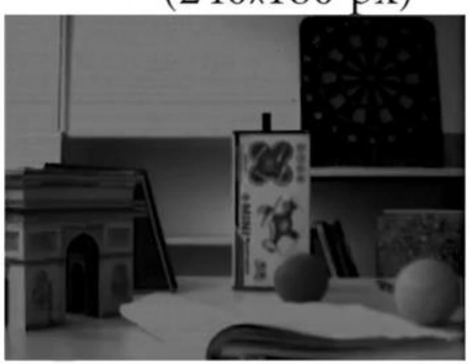


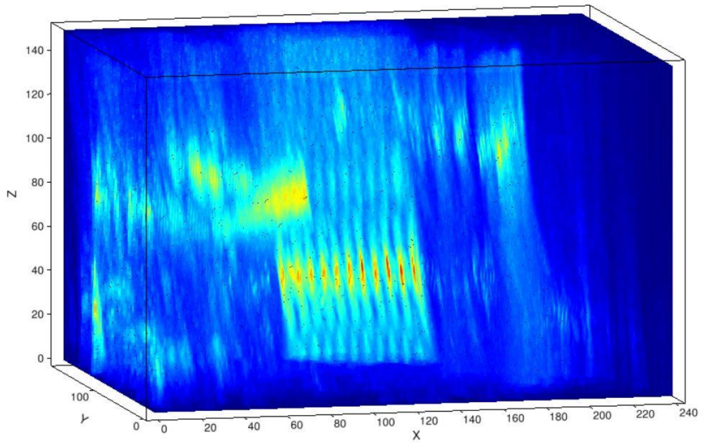


Volume of interest - Disparity Space Image (DSI)

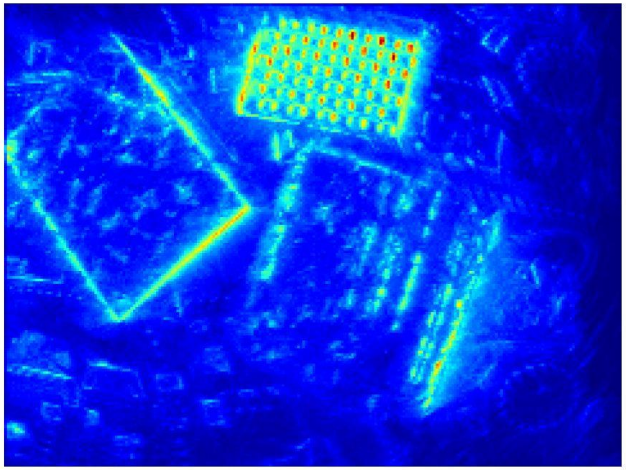
240 x 180 x 100 voxels

(240x180 px)

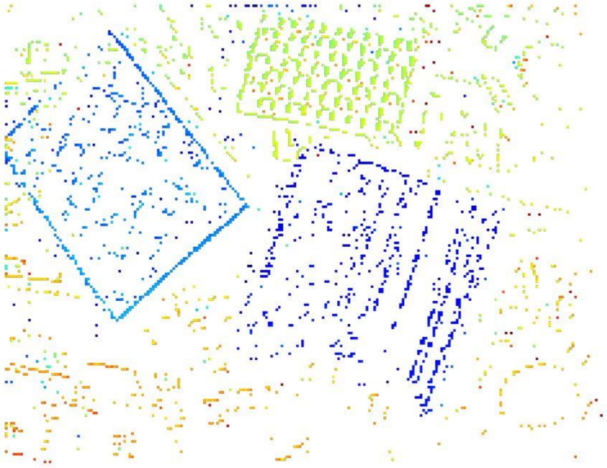




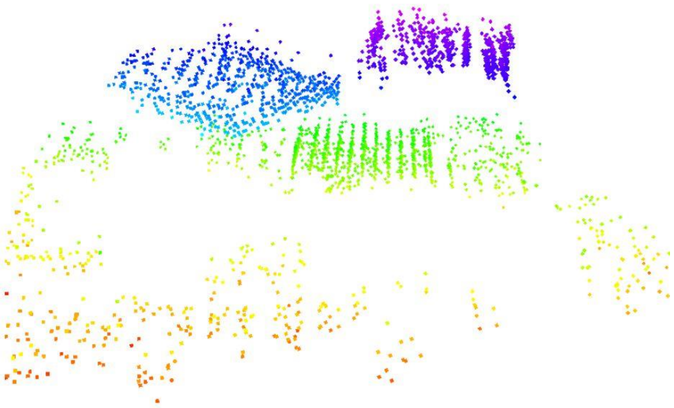
(a) Ray density DSI $f(\mathbf{X})$.



(b) Confidence map.



(c) Semi-dense depth map.



(d) 3D point cloud.

Camera positions

- Motorized linear slider
- Visual odometry algorithm – SVO [6]

[6] C. Forster, M. Pizzoli and D. Scaramuzza. “”SVO: Fast semi-direct monocular visual odometry”. ICRA 2014

Table 1: Depth estimation accuracy in the synthetic datasets ($N_z = 100$)

	Dunes	3 planes	3 walls
Depth range	3.00 m	1.30 m	7.60 m
Mean error	0.14 m	0.15 m	0.52 m
Relative error	4.63%	11.31%	6.86%

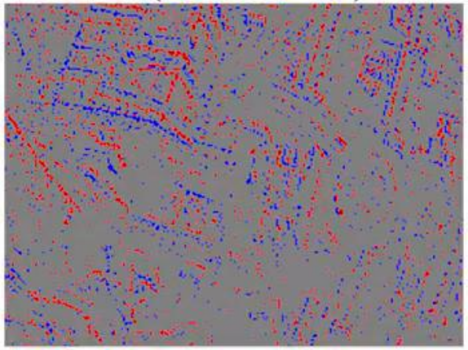
Table 2: Depth estimation accuracy in the HDR experiment

	Close (distance: 23.1 cm)		Far (distance: 58.5 cm)	
Illumination	Mean error	Relative error	Mean error	Relative error
◦ constant	1.22 cm	5.29%	2.01 cm	4.33%
◦ HDR	1.21 cm	5.25%	1.87 cm	3.44%

Large-scale reconstruction with hand-held DAVIS

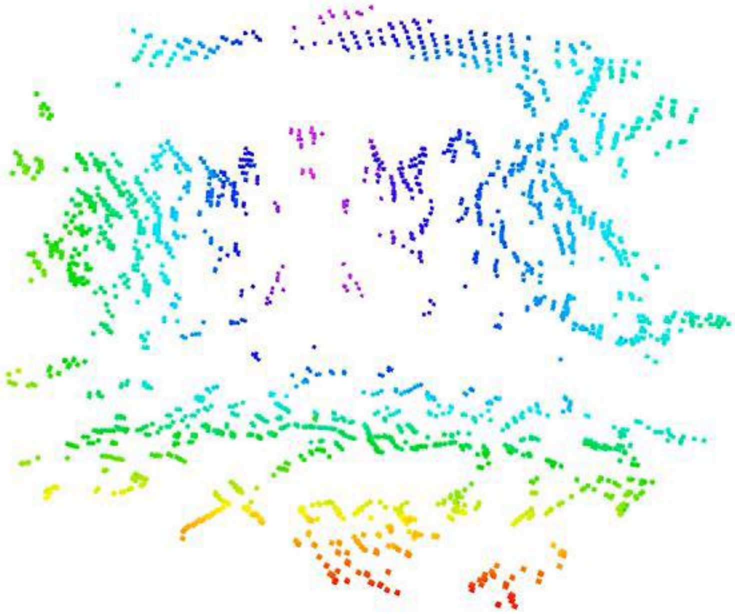
Frames (240x180 px)

Events (ON, OFF)



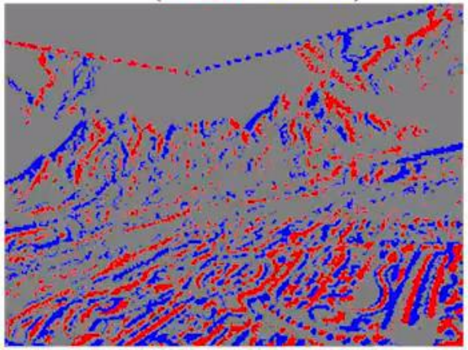
Real-Time

3 walls dataset (synthetic)

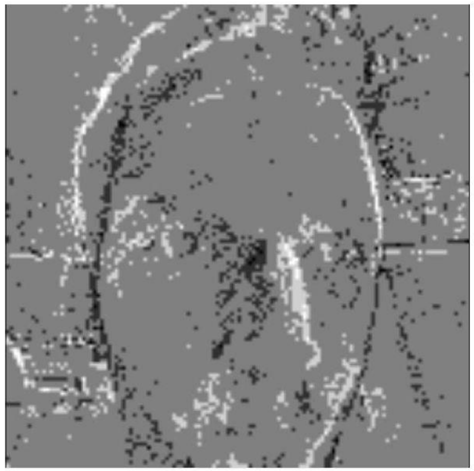


Frames (240x180 px)

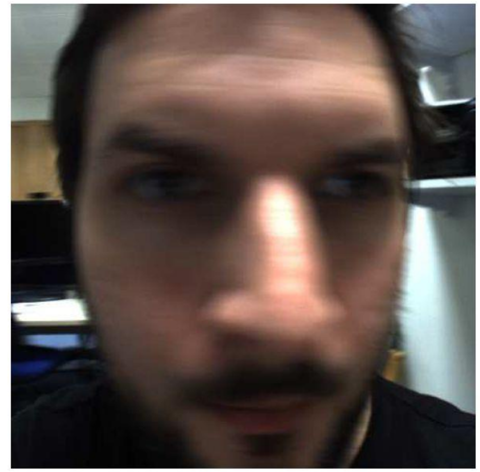
Events (ON, OFF)



Simultaneous Optical Flow and Intensity Estimation from an Event Camera[2]

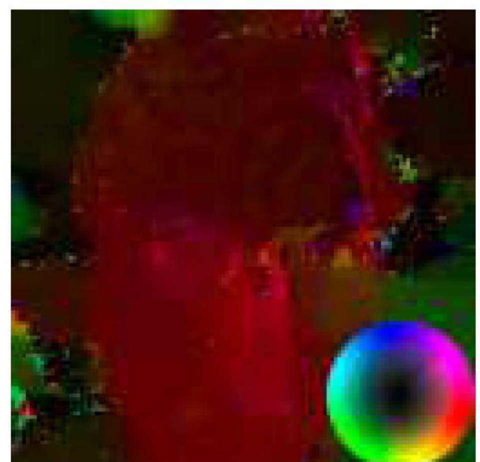


(a) Raw event camera output



(b) Standard camera image

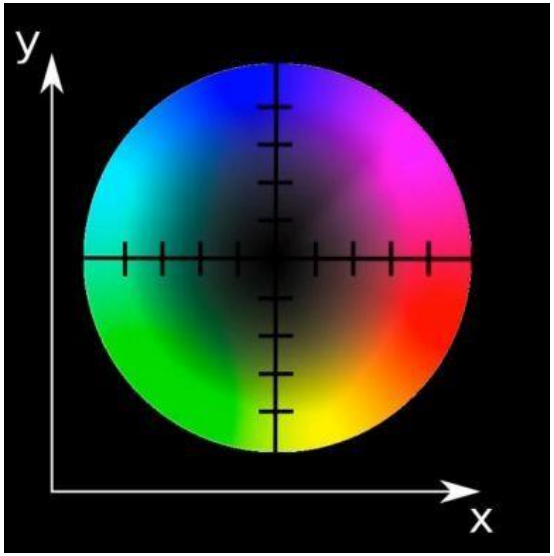
(c) Intensity estimate from events



(d) Optical flow from events

Optical flow

- Apparent motion of brightness patterns



Problem formulation

Events: $e_i = (x_i, t_i, \rho_i)$

$x_i \in \Omega$ - position of the event

t_i - timestamp

$\rho_i \in \{-1, 1\}$ - polarity

Event fired if: $\left|L(x,t) - L\left(x, t_p(x,t)\right)\right| \geq \theta$

$L(x,t)$ log intensity at pixel x, at time t

$t_p(x,t)$ time of the previous event

Goal: u - velocity

I - image

Brightness constancy: $I(x + \delta_t u, t + \delta_t) = I(x,t)$

δ_t time discretisation

Spatial smoothness of the flow

$$\lambda_1 \|u_x\|_1$$

Temporal smoothness of the flow

$$\lambda_2 \|u_t\|_1$$

Intensity smoothness

$$\lambda_3 \|L_x\|_1$$

Temporal consistency

$$\lambda_4 \|\langle L_x, \delta_t u\rangle + L_t\|_1$$

Event data term

$$\sum_{i=2}^{|P(x)|} \|L(x, t_i) - L(x, t_{i-1}) - \theta\rho_i\|_1$$

P(x) set of all events fired at x

No-event data term

$$\lambda_5 h_\theta \left(L(x,t) - L\big(x, t_p(x,t)\big)\right)$$

$$h_\theta(x) = \begin{cases} |x| - \theta & , \text{if } |x| > \theta \\ 0 & , \text{otherwise} \end{cases}$$

Prior image constraint

$$\left\|L(x, t_1) - \hat{L}(x)\right\|_2^2 \qquad \hat{L} - \text{prior image}$$

$t_1 -$ first event timestamp at x, in current window, or minimum of T

Discretisation

Ω => regular pixel grid of size MxN
T => K cells, each of length δ_t μs

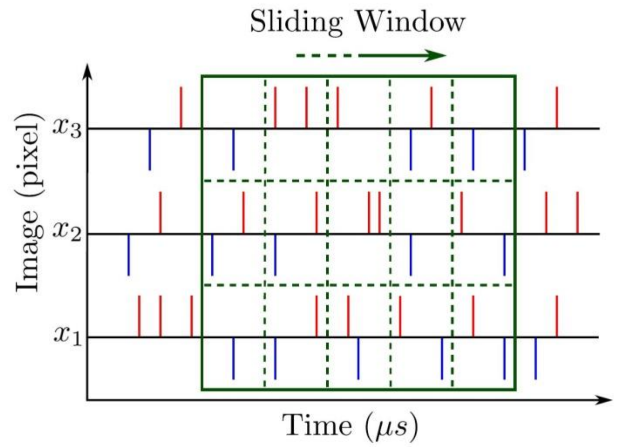


Figure 3: The sliding window (green box) bins the incoming positive events (red bars) and negative events (blue bars) into a regular grid (dashed lines). When the minimisation converges, the window is shifts to the right.

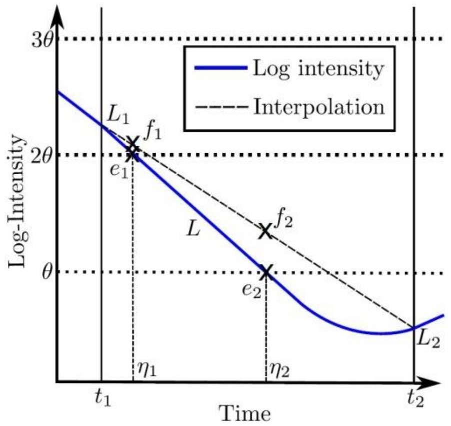


Figure 4: Approximation of the intensity for two given events e_1 and e_2 between two intensity estimates L_1 and L_2. For the discrete data term, we use the linear approximations f_1 and f_2 at the time of each event η_1 and η_2.

Optimisation

$$\min_{u,L} \int_\Omega \int_T \left(\lambda_1 \|u_x\|_1 + \lambda_2 \|u_t\|_1 + \lambda_3 \|L_x\|_1 + \lambda_4 \|\langle L_x, \delta_t u\rangle + L_t\|_1 + \lambda_5 h_\theta\left(L - L(t_p)\right)\right) dtdx$$
$$+ \int_\Omega \left(\left(\sum_{i=2}^{|P(x)|} \|L(t_i) - L(t_{i-1}) - \theta\rho_i\|_1 \right) + \lambda_6 \left\|L(x, t_1) - \hat{L}(x)\right\|_2^2 \right) dx$$

- preconditioned primal-dual algorithm [7]
- Legendre-Fenchel transform [8]

[7] T.Pock and A. Chambolle. “Diagonal preconditioning for first order primal-dual algorithms in convex optimization”. ICCV 2011

[8] A. Handa, R. A. Newcombe, A. Angeli and A.J. Davison. “Applications of the Legendre-Fenchel transformation to computer vision problems”

Experiments

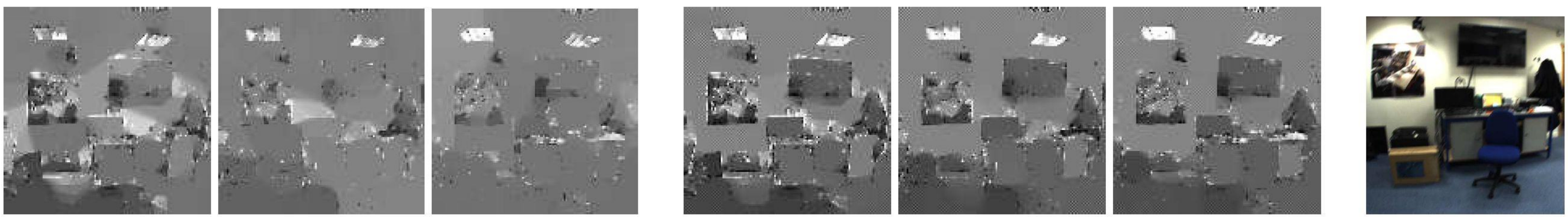
DVS128 => 128 x 128

K = 128, $\delta_t = 15\text{ms}$

$\theta = 0.22$

$\lambda_1 = 0.02,\ \lambda_2 = 0.05, \lambda_3 = 0.02, \lambda_4 = 0.2, \lambda_5 = 0.1, \lambda_6 = 1$

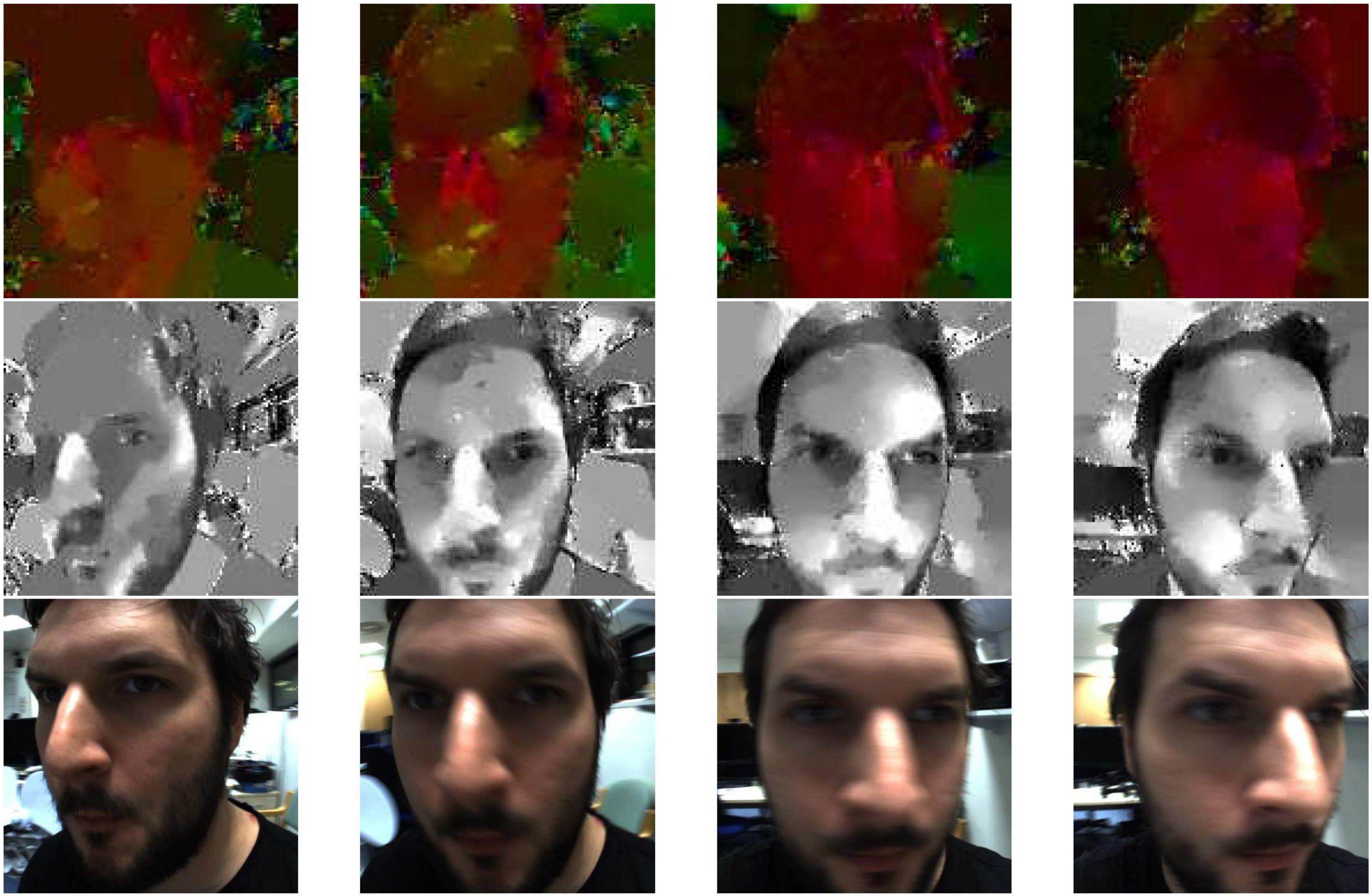
- Benefits of simultaneous estimation



(a) Without optical flow (b) With optical flow (c) Camera images

Experiments

DVS128 => 128 x 128
K = 128, $\delta_t = 15\text{ms}$
$\theta = 0.22$
$\lambda_1 = 0.02,\ \lambda_2 = 0.05, \lambda_3 = 0.02, \lambda_4 = 0.2, \lambda_5 = 0.1, \lambda_6 = 1$



Experiments

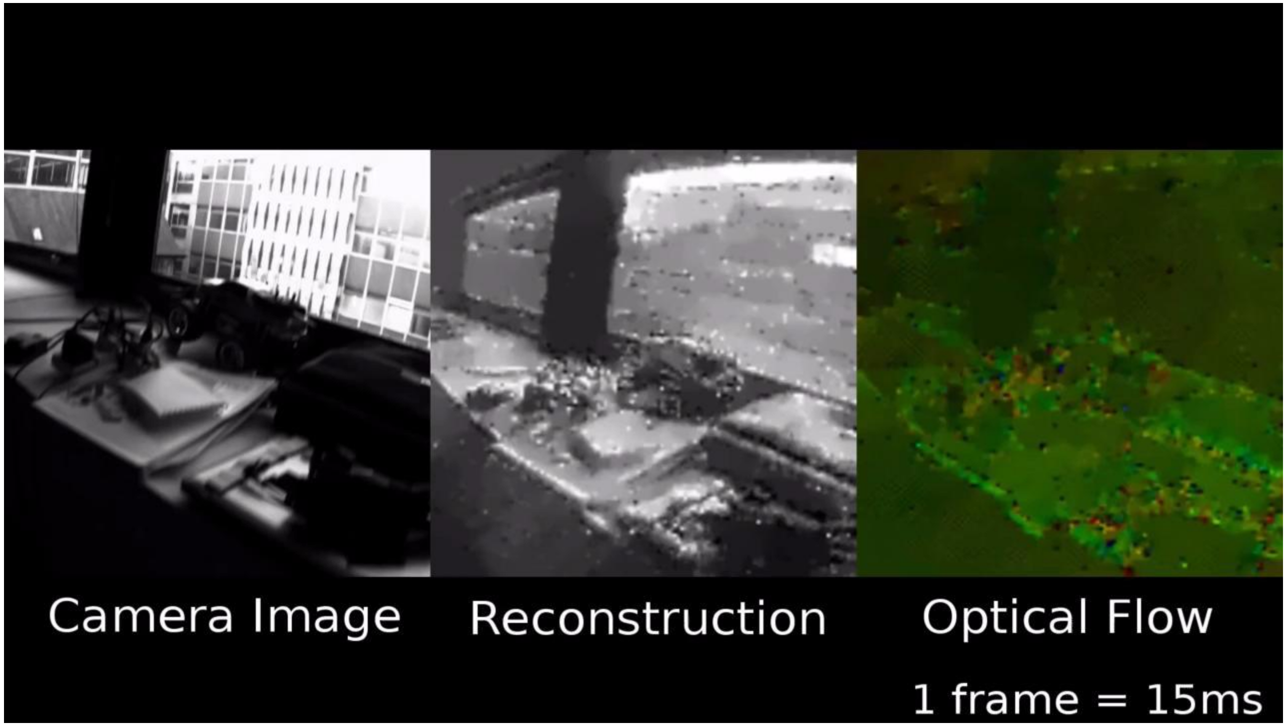
DVS128 => 128 x 128

K = 128, $\delta_t = 15\text{ms}$

$\theta = 0.22$

$\lambda_1 = 0.02,\ \lambda_2 = 0.05, \lambda_3 = 0.02, \lambda_4 = 0.2, \lambda_5 = 0.1, \lambda_6 = 1$

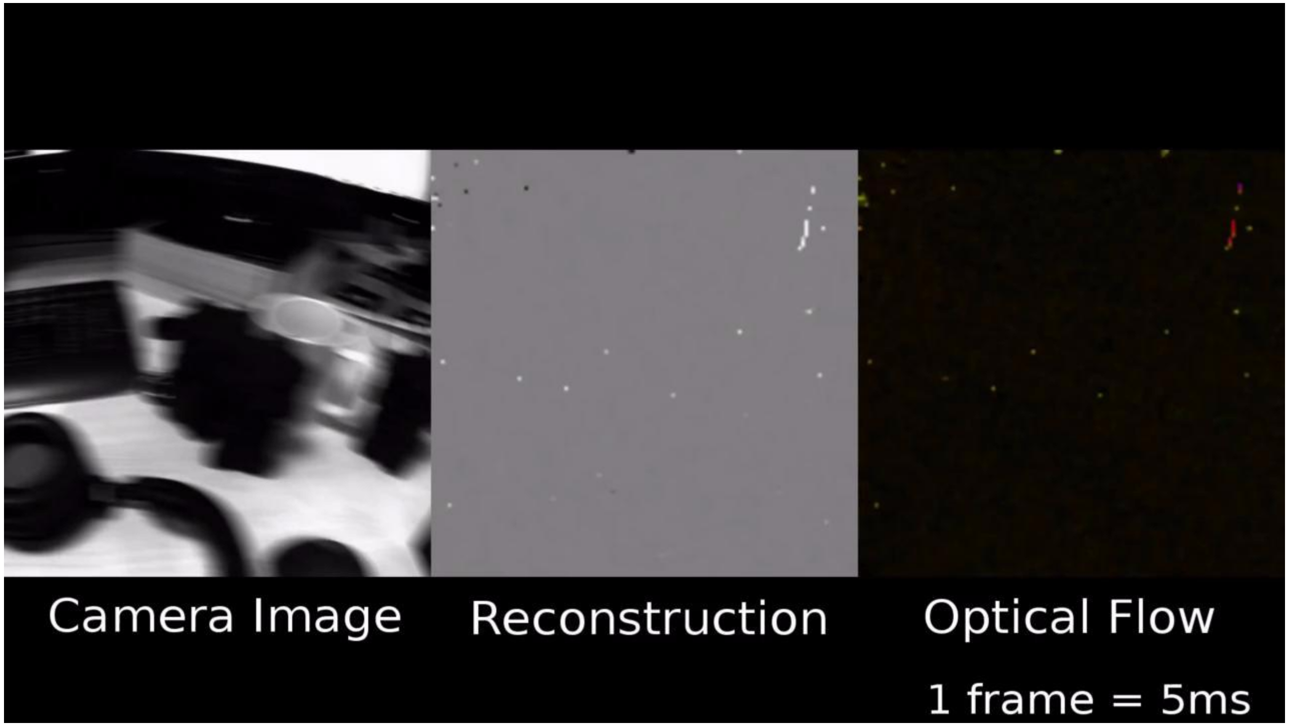
- High Dynamic Range



Experiments

DVS128 => 128 x 128
K = 128, $\delta_t = 4\text{ms}$
$\theta = 0.22$
$\lambda_1 = 0.02,\ \lambda_2 = 0.05, \lambda_3 = 0.02, \lambda_4 = 0.2, \lambda_5 = 0.1, \lambda_6 = 1$

- Rapid motion



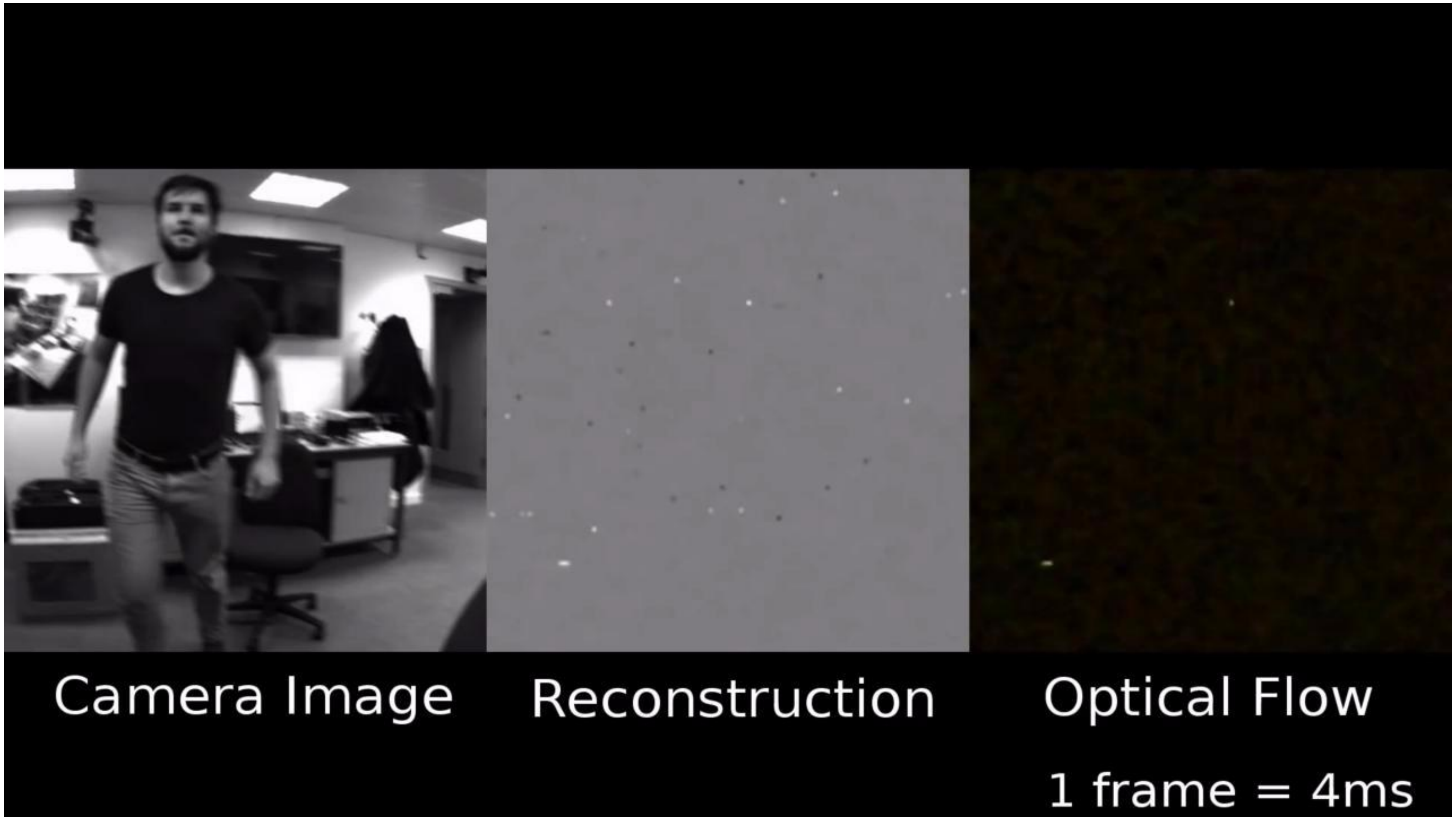
Experiments

DVS128 => 128 x 128

K = 128, $\delta_t = 7\text{ms}$

$\theta = 0.22$

$\lambda_1 = 0.01,\ \lambda_2 = 0.05, \lambda_3 = 0.01, \lambda_4 = 0.2, \lambda_5 = 0.1, \lambda_6 = 1$



References

[1] H Rebecq, G. Gallego and D. Scaramuzza. “EMVS: Event-based Multi-View Stereo”, BMVC 2016

[2] P. Bardow, A.J. Davison and S. Leutenegger. “Simultaneous Optical Flow and Intensity Estimation from an Event Camera”, CVPR 2016

[3] P.Lichtsteiner, C. Posch and T. Delbruck. “A 128x128 120dB 15μs latency asynchronous temporal contrast vision sensor”. JSSC 2008

[4] C. Brandli, R. Berner, M. Yang, S.C. Liu and T. Delbruck. “A 240x180 130dB 3μs Latency Global Shutter Spatiotemporal Vision Sensor”. JSSC 2014

[5] R.T. Collins. “A space-sweep approach to true multi-image matching”. CVPR 1996

[6] C. Forster, M. Pizzoli and D. Scaramuzza. “”SVO: Fast semi-direct monocular visual odometry”. ICRA 2014

[7] T.Pock and A. Chambolle. “Diagonal preconditioning for first order primal-dual algorithms in convex optimization”. ICCV 2011

[8] A. Handa, R. A. Newcombe, A. Angeli and A.J. Davison. “Applications of the Legendre-Fenchel transformation to computer vision problems”