THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL

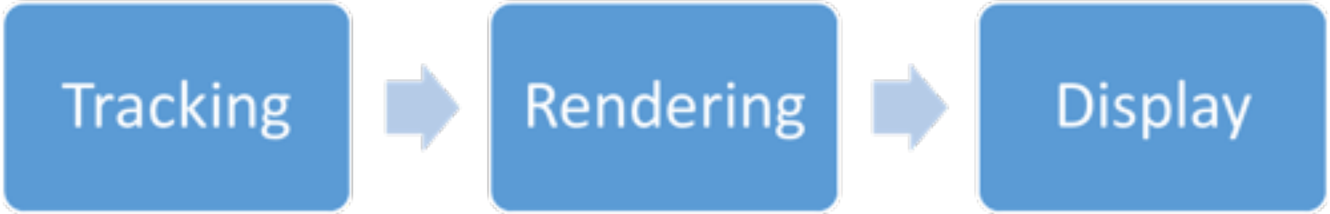
ISMAR 2016
September 19 - 23 | Merida, MEXICO

STEVENS INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

Towards kHz 6-DoF Visual Tracking Using an Egocentric Cluster of Rolling Shutter Cameras

Akash Bapat[1], Enrique Dunn[1,2] & Jan-Michael Frahm[1],
UNC Chapel Hill, USA[1],
Stevens Institute of Technology, USA[2]

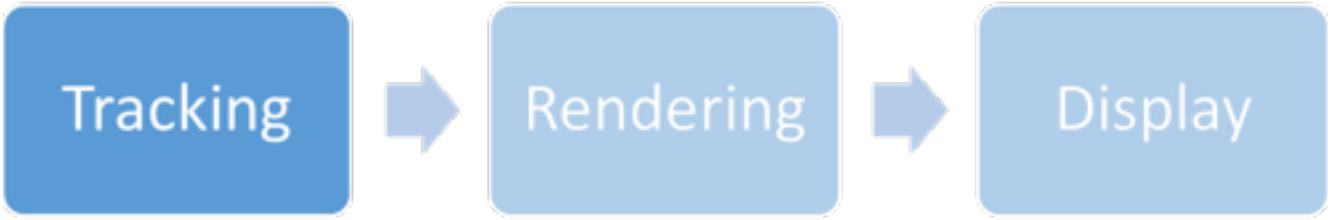
AR/VR System Components



Tracking → Rendering → Display

Image source : http://static3.gamespot.com/uploads/original/1179/11799911/2786939-hololens7.jpg

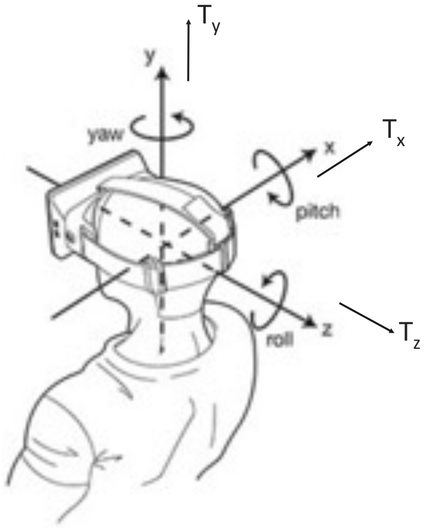
AR/VR System Components



Tracking

Rendering

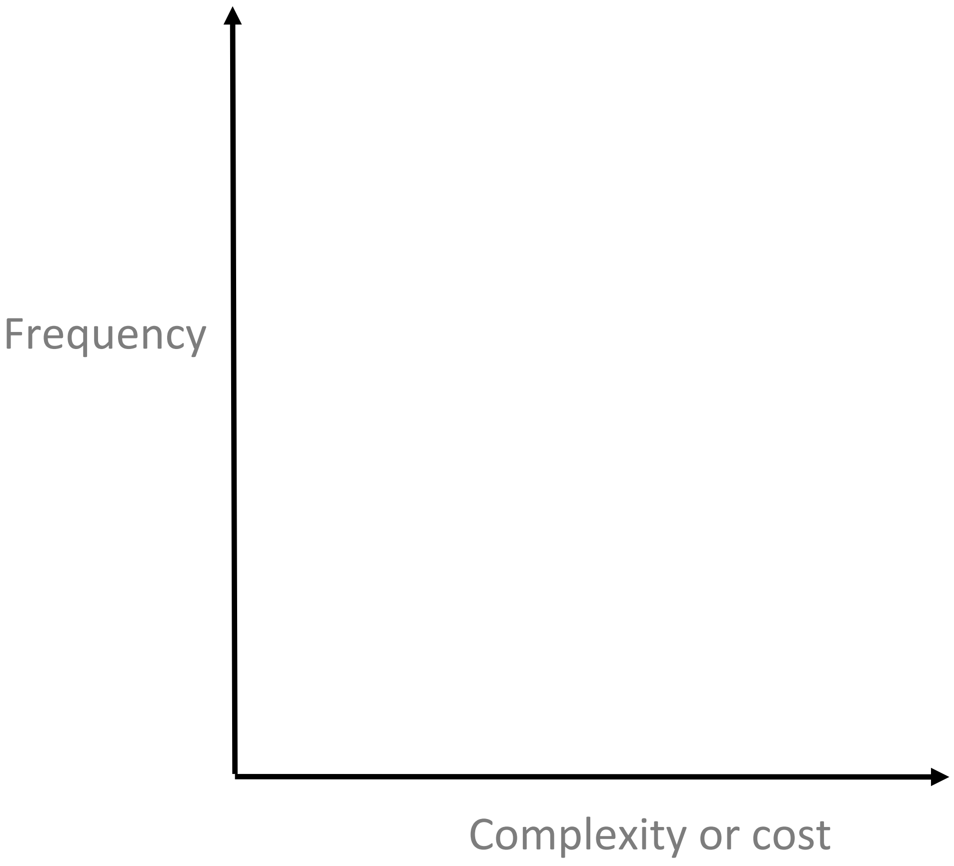
Display



Find 6-DoF head pose[1]

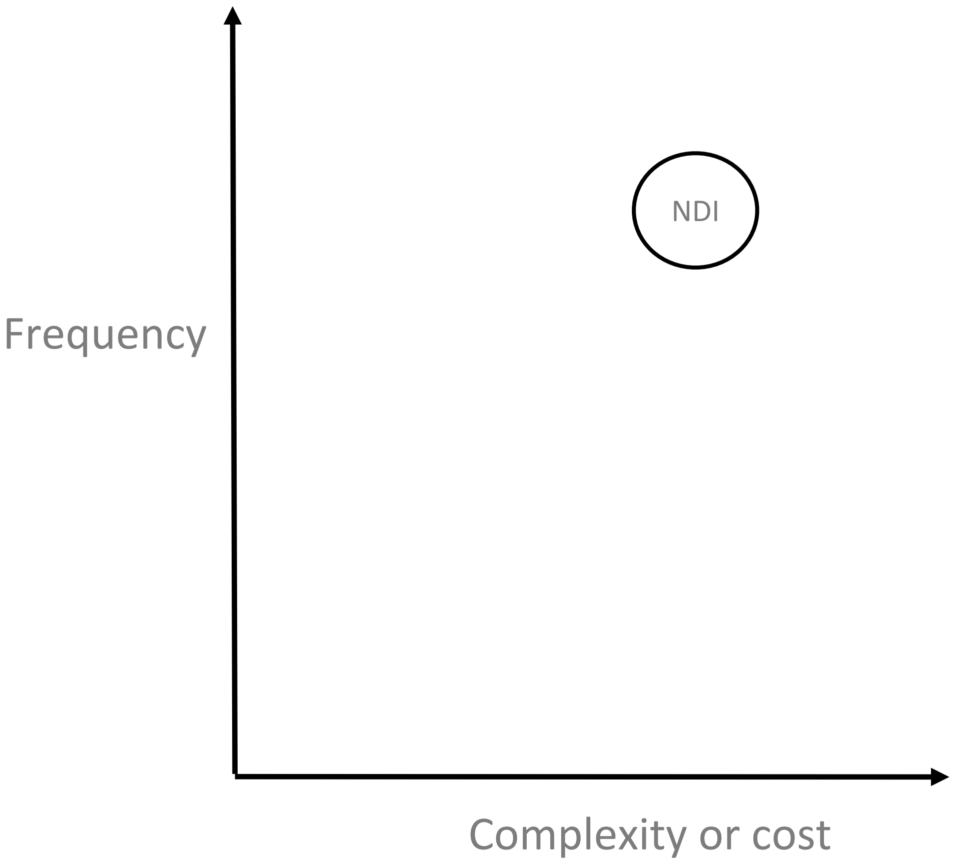
[1] LaValle etal, Head Tracking for Oculus Rift

Tracking Systems



1. http://certus.ndigital.com
2. http://vrshoot.ru/sites/default/files/oculus-rift-dk2-cam.jpg
3. http://i.imgur.com/64LA9Kv.jpg
4. http://cdn.pocketnow.com/wp-content/uploads/2011/04/gyro.jpg
5. LaValle etal, Head Tracking for Oculus Rift

Tracking Systems

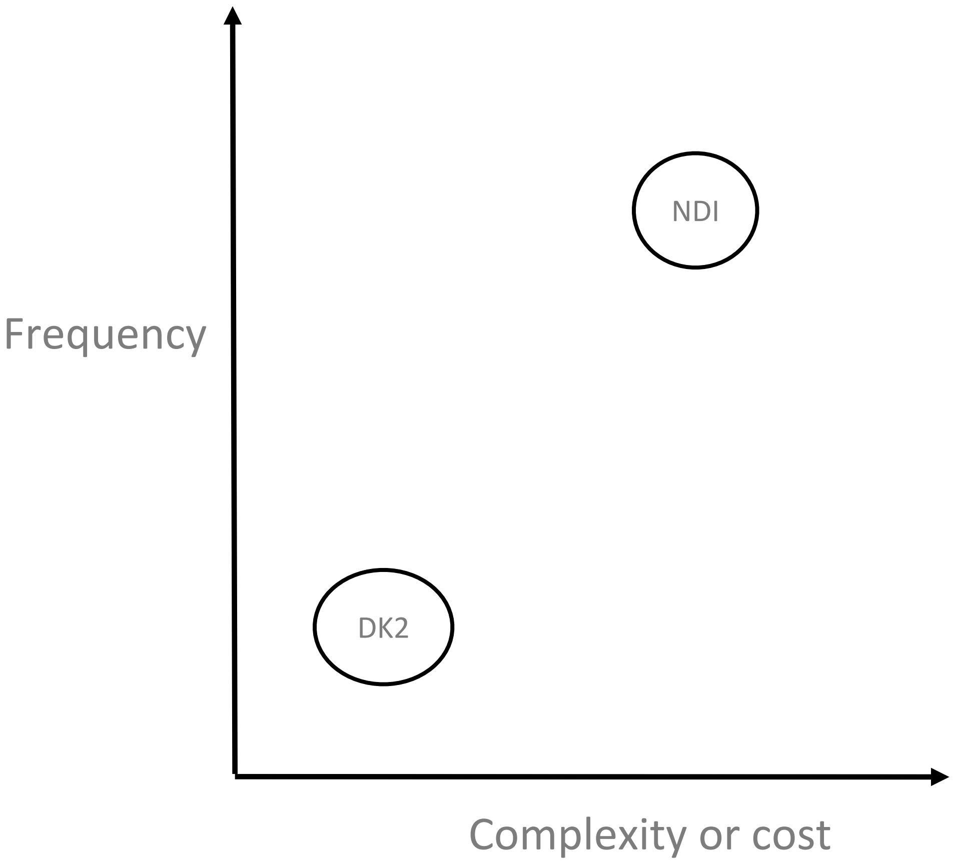


NDI Optotrak Certus[1]

- Tracking frequency ≈ 5kHz
- State of the art

1. http://certus.ndigital.com
2. http://vrshoot.ru/sites/default/files/oculus-rift-dk2-cam.jpg
3. http://i.imgur.com/64LA9Kv.jpg
4. http://cdn.pocketnow.com/wp-content/uploads/2011/04/[illegible]ro.jpg
5. LaValle etal, Head Tracking for Oculus Rift

Tracking Systems



Frequency

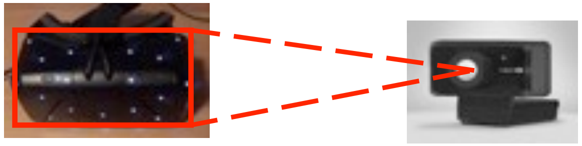
NDI

DK2

Complexity or cost

Oculus Rift DK2

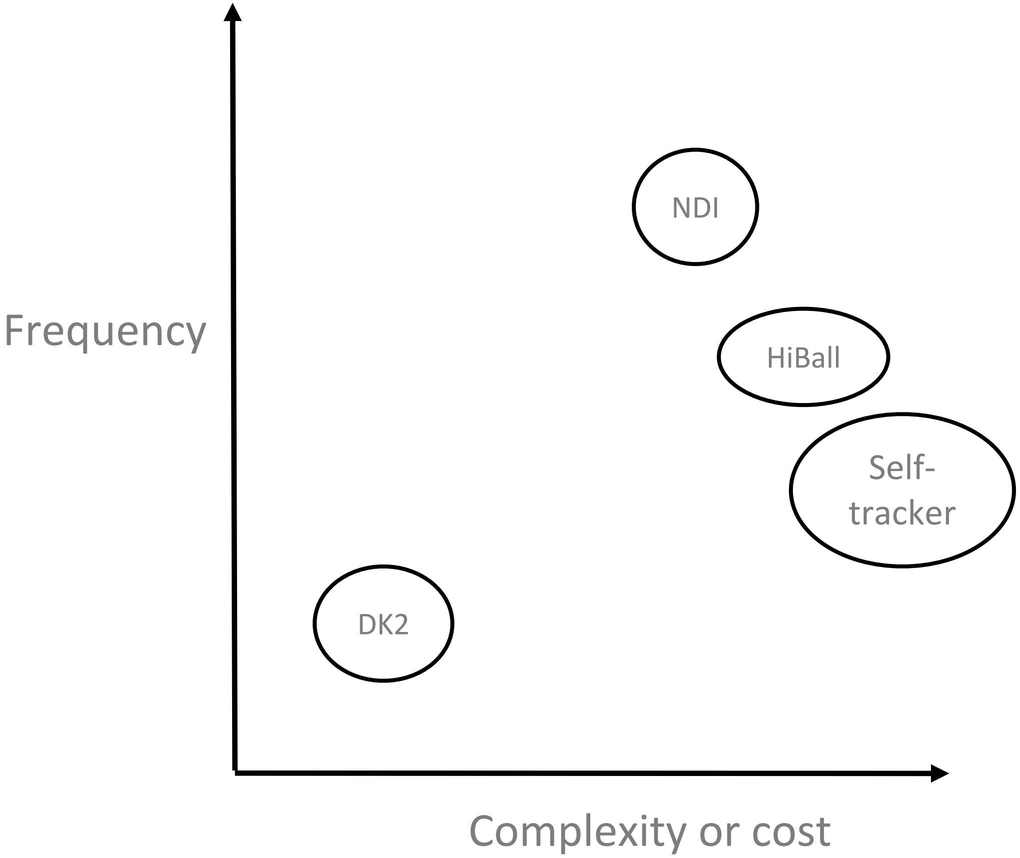
Orientation tracking
IMU at 1000 Hz [4,5]



Positional tracking [2,3]
Camera fps = 60 fps

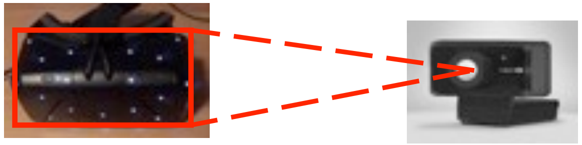
1. http://certus.ndigital.com
2. http://vrshoot.ru/sites/default/files/oculus-rift-dk2-cam.jpg
3. http://i.imgur.com/64LA9Kv.jpg
4. http://cdn.pocketnow.com/wp-content/uploads/2011/04/gyro.jpg
5. LaValle etal, Head Tracking for Oculus Rift

Tracking Systems



Oculus Rift DK2

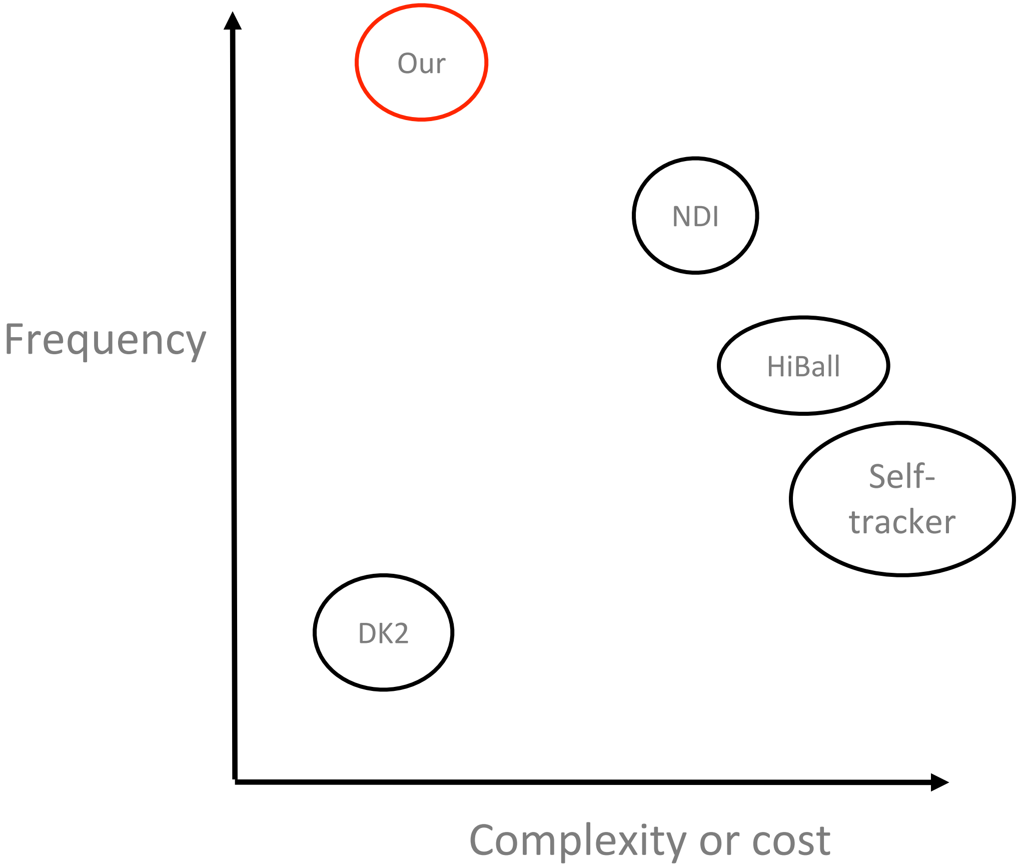
Orientation tracking
IMU at 1000 Hz [4,5]



Positional tracking [2,3]
Camera fps = 60 fps

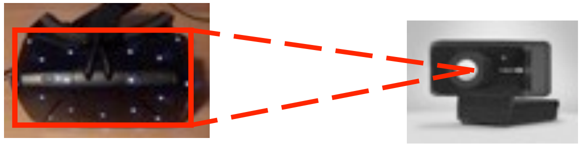
[1] http://certus.ndigital.com
[2] http://vrshoot.ru/sites/default/files/oculus-rift-dk2-cam.jpg
[3] http://i.imgur.com/64LA9Kv.jpg
[4] http://cdn.pocketnow.com/wp-content/uploads/2011/04/gyro.jpg
[5] LaValle etal, Head Tracking for Oculus Rift

Tracking Systems



Oculus Rift DK2

Orientation tracking
IMU at 1000 Hz [4,5]



Positional tracking [2,3]
Camera fps = 60 fps

1. http://certus.ndigital.com
2. http://vrshoot.ru/sites/default/files/oculus-rift-dk2-cam.jpg
3. http://i.imgur.com/64LA9Kv.jpg
4. http://cdn.pocketnow.com/wp-content/uploads/2011/04/gyro.jpg
5. LaValle etal, Head Tracking for Oculus Rift

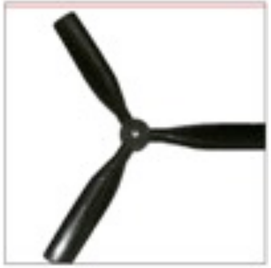
Rolling Shutter

Rolling shutter capture[1]

[1] http://jasmcole.com/2014/10/12/rolling-shutters/

Rolling Shutter

- Row-by-row acquisition of linescan snapshots at slightly different times



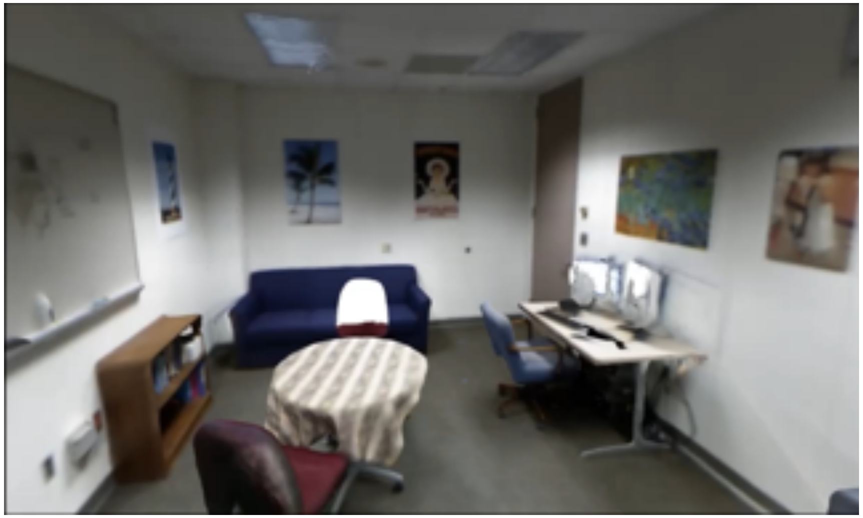
Rolling shutter capture[1]

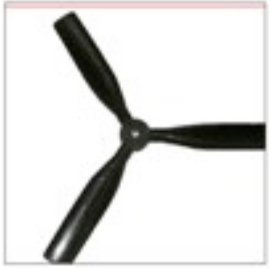
[1] http://jasmcole.com/2014/10/12/rolling-shutters/

Rolling Shutter

- Row-by-row acquisition of linescan snapshots at slightly different times

- Stream of row-images



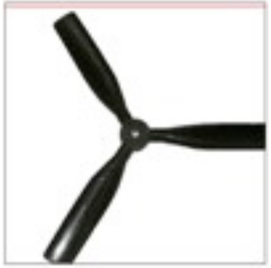


Rolling shutter capture[1]

[1] http://jasmcole.com/2014/10/12/rolling-shutters/

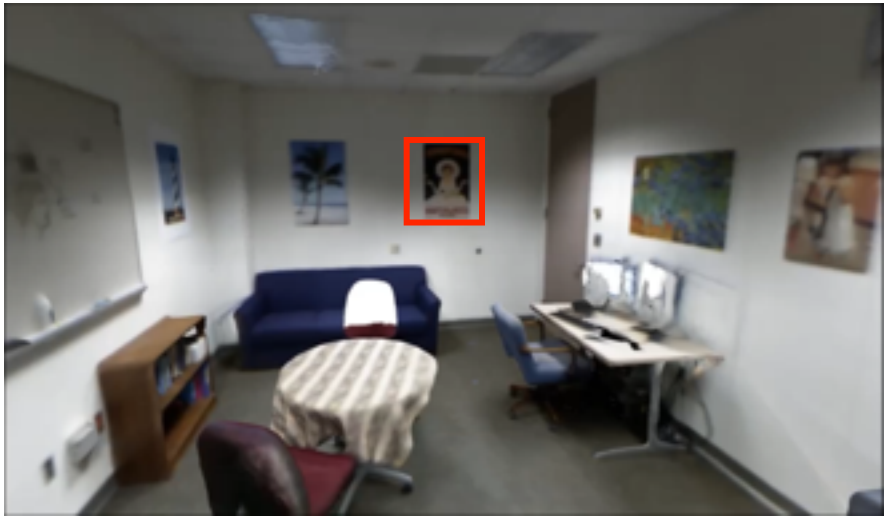
Rolling Shutter

- Row-by-row acquisition of linescan snapshots at slightly different times



Rolling shutter capture[1]

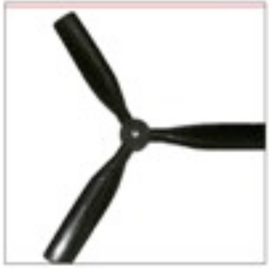
- Stream of row-images



[1] http://jasmcole.com/2014/10/12/rolling-shutters/

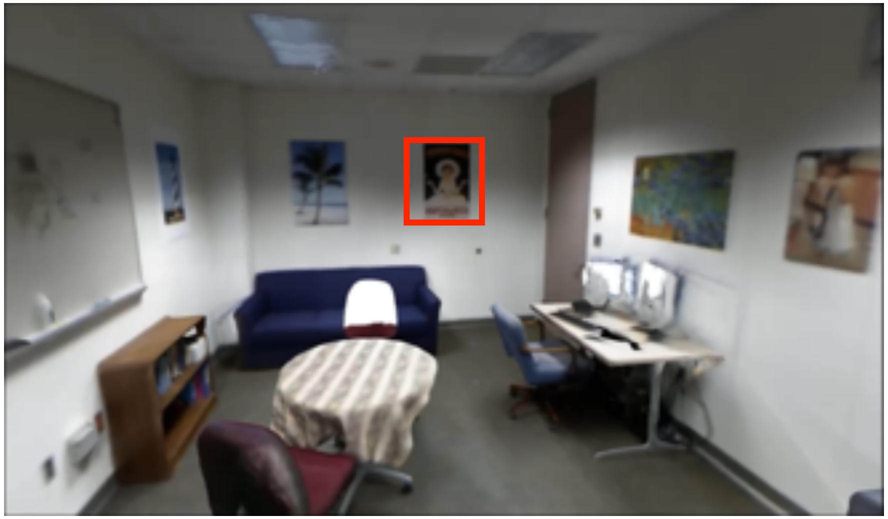
Rolling Shutter

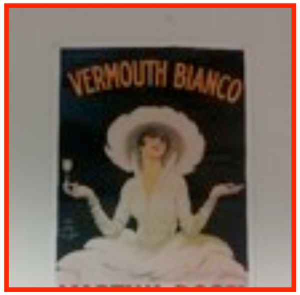
- Row-by-row acquisition of linescan snapshots at slightly different times

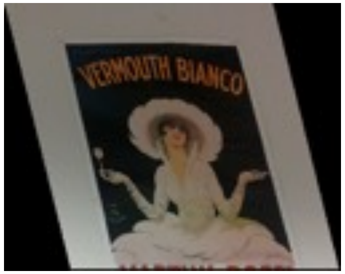


Rolling shutter capture[1]

- Stream of row-images







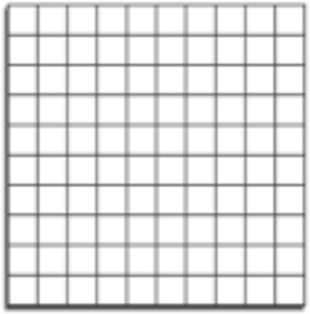
[1] http://jasmcole.com/2014/10/12/rolling-shutters/

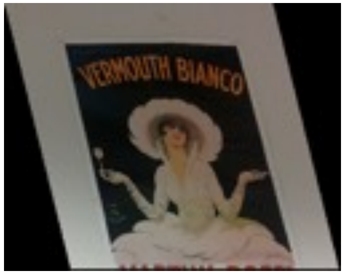
Stream of Rows

- Frequency of row-samples
  - F = FPS * Height

    = 120 * 720 > 80kHz

Rolling shutter[1]





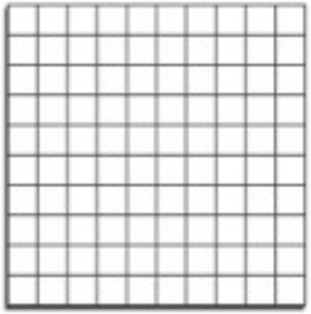
[1] http://www.diyphotography.net/everything-you-wanted-to-know-about-rolling-shutter/

Stream of Rows

- Frequency of row-samples
  - F = FPS * Height

    = 120 * 720 > 80kHz

Rolling shutter[1]



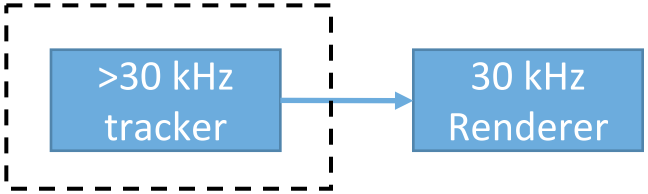
[1] http://www.diyphotography.net/everything-you-wanted-to-know-about-rolling-shutter/

Our Tracker

1. http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html

Our Tracker

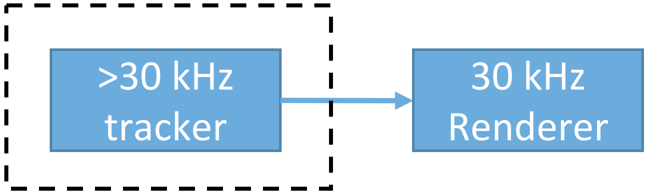
- Enabling component for rendering
  1. Zheng et al., 2014
  2. Lincoln et al., 2016



[1] http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html

Our Tracker

- Enabling component for rendering
  1. Zheng et al., 2014
  2. Lincoln et al., 2016

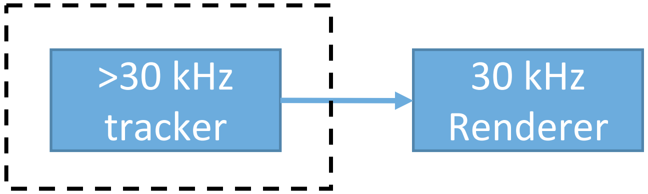


GoPro[1]

[1] http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html

Our Tracker

- Enabling component for rendering
  1. Zheng et al., 2014
  2. Lincoln et al., 2016



>30 kHz tracker

30 kHz Renderer

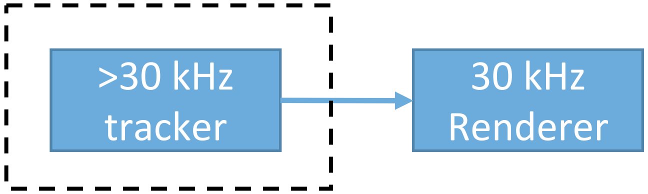
- Use commodity cameras

GoPro[1]

[1] http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html

Our Tracker

- Enabling component for rendering
    1. Zheng et al., 2014
    2. Lincoln et al., 2016



>30 kHz tracker

30 kHz Renderer

- Use commodity cameras

GoPro[1]

- Tracking frequency
    - Up to 80 kHz
- Break frame-rate barrier
    - Process each row of image

[1] http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html

Related Work

Artifact visualisation[1]

[1] http://jasmcole.com/2014/10/12/rolling-shutters/

Artifact visualisation[1]

Related Work

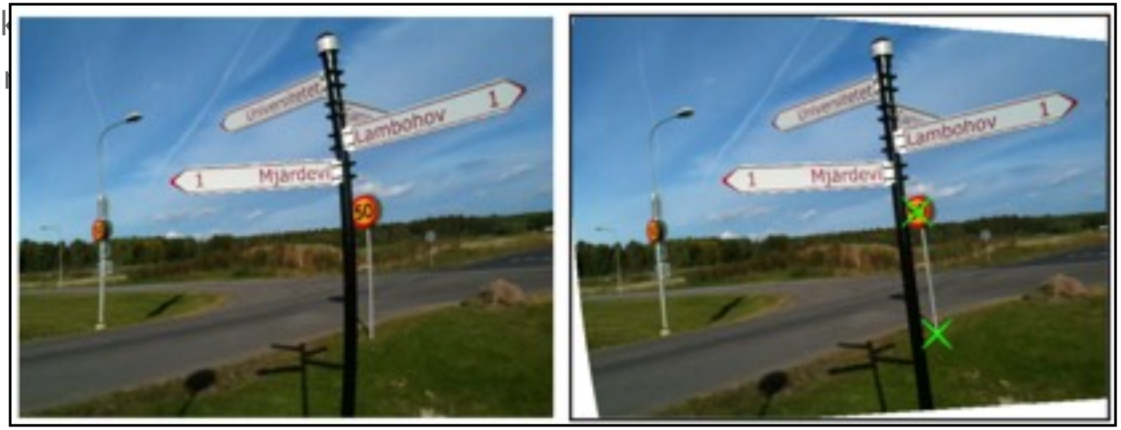
- Removing rolling shutter
  - Forssen et al., CVPR 2010
  - Track points using KLT tracker, estimate rotation
  - Parametrize intra-frame rotation as spline

[1] http://jasmcole.com/2014/10/12/rolling-shutters/

Related Work

Artifact visualisation[1]

- Removing rolling shutter
  - Forssen et al., CVPR 2010
  - Track
  - Para



[1] http://jasmcole.com/2014/10/12/rolling-shutters/

Related Work

Artifact visualisation[1]

- Removing rolling shutter
  - Forssen et al., CVPR 2010
  - Track points using KLT tracker, estimate rotation
  - Parametrize intra-frame rotation as spline

- Geometric problems
  - Multi-view stereo: Saurer et al., ICCV 2013
  - Adapt plane sweep stereo for rolling shutter

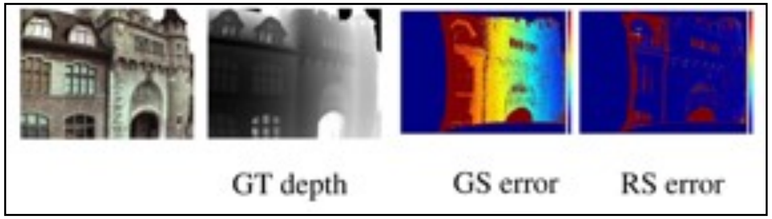
[1] http://jasmcole.com/2014/10/12/rolling-shutters/

Related Work

Artifact visualisation[1]

- Removing rolling shutter
  - Forssen et al., CVPR 2010
  - Track points using KLT tracker, estimate rotation
  - Parametrize intra-frame rotation as spline

- Geometric problems
  - Multi-view stereo: Saurer et al., ICCV 2013
  - Adapt plane sweep stereo for rolling shutter



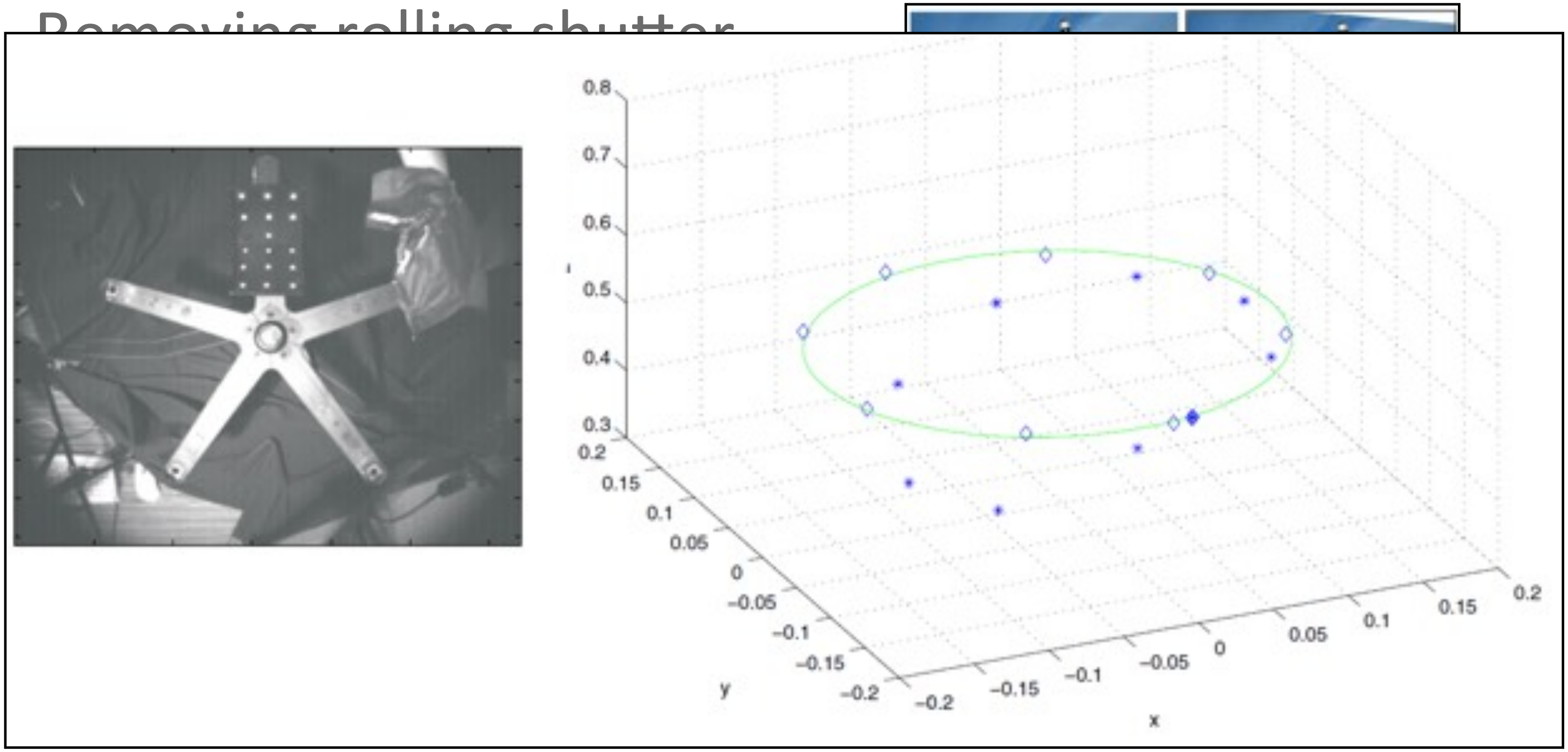
- Velocity estimation
  - Ait-Aider et al., ICVS 2006
  - Solve for pose and velocity using bundle adjustment
  - Use 2D-3D correspondence, similar to camera calibration

[1] http://jasmcole.com/2014/10/12/rolling-shutters/

Related Work

Artifact visualisation[1]

- Removing rolling shutter



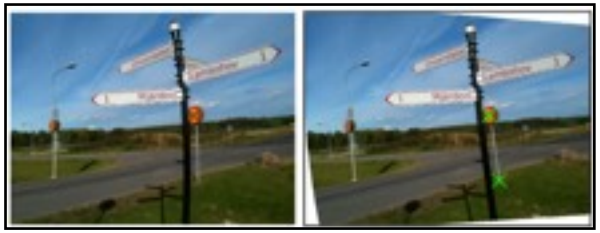
- Solve for pose and velocity using bundle adjustment
- Use 2D-3D correspondence, similar to camera calibration

[1] http://jasmcole.com/2014/10/12/rolling-shutters/

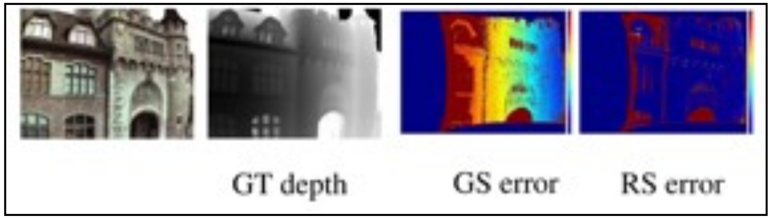
Artifact visualisation[1]

Related Work

- Removing rolling shutter
  - Forssen et al., CVPR 2010
  - Track points using KLT tracker, estimate rotation
  - Parametrize intra-frame rotation as spline



- Geometric problems
  - Multi-view stereo: Saurer et al., ICCV 2013
  - Adapt plane sweep stereo for rolling shutter



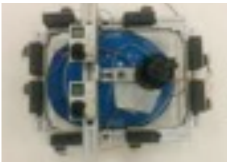
- Velocity estimation
  - Ait-Aider et al., ICVS 2006
  - Solve for pose and velocity using bundle adjustment
  - Use 2D-3D correspondence, similar to camera calibration

[1] http://jasmcole.com/2014/10/12/rolling-shutters/

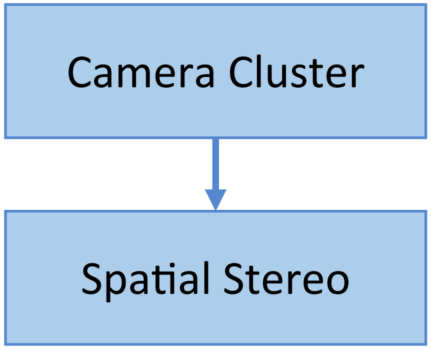
Approach Overview

Approach Overview

Camera Cluster

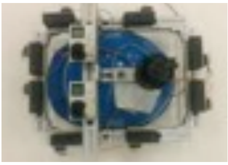


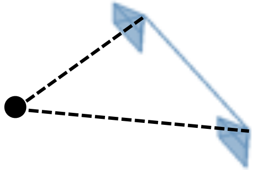
Approach Overview



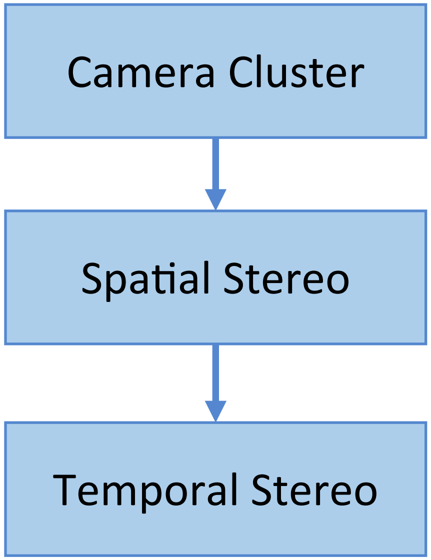
Camera Cluster

Spatial Stereo





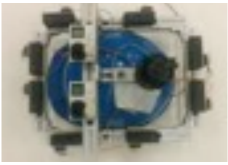
Approach Overview

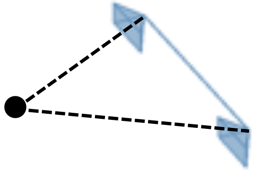


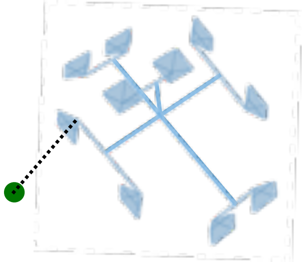
Camera Cluster

Spatial Stereo

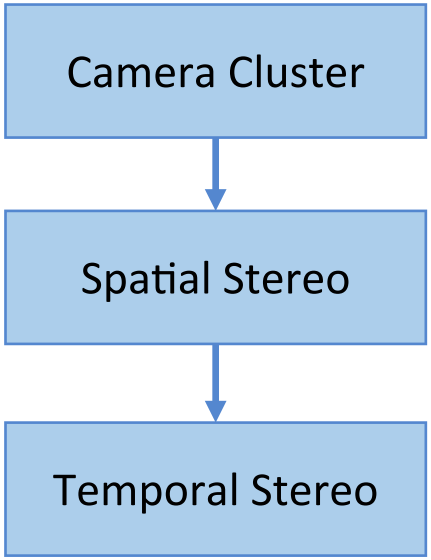
Temporal Stereo





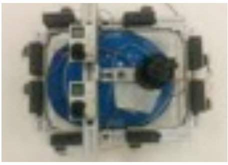


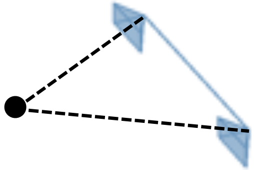
Approach Overview

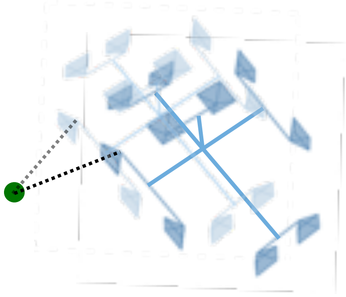

Camera Cluster

Spatial Stereo

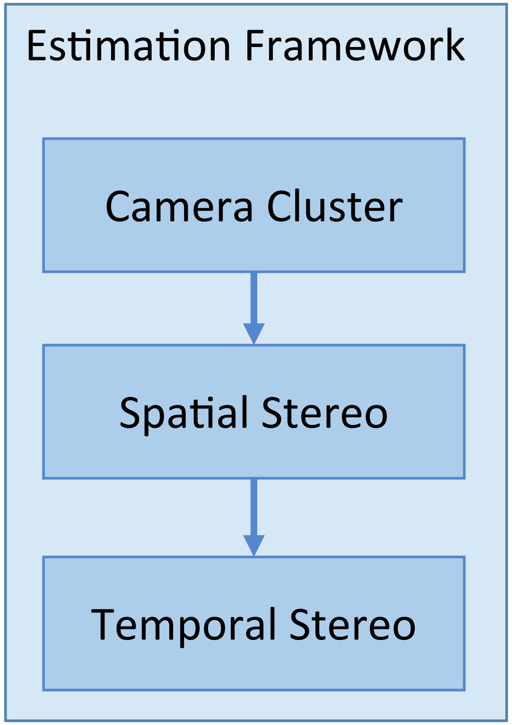
Temporal Stereo







Approach Overview

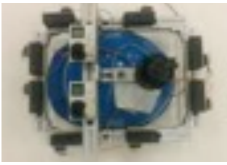


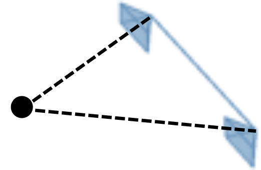
Estimation Framework

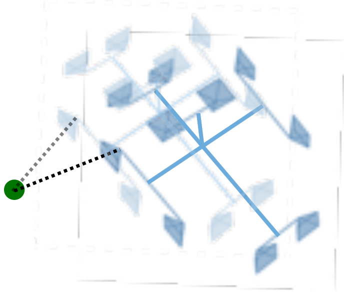
Camera Cluster

Spatial Stereo

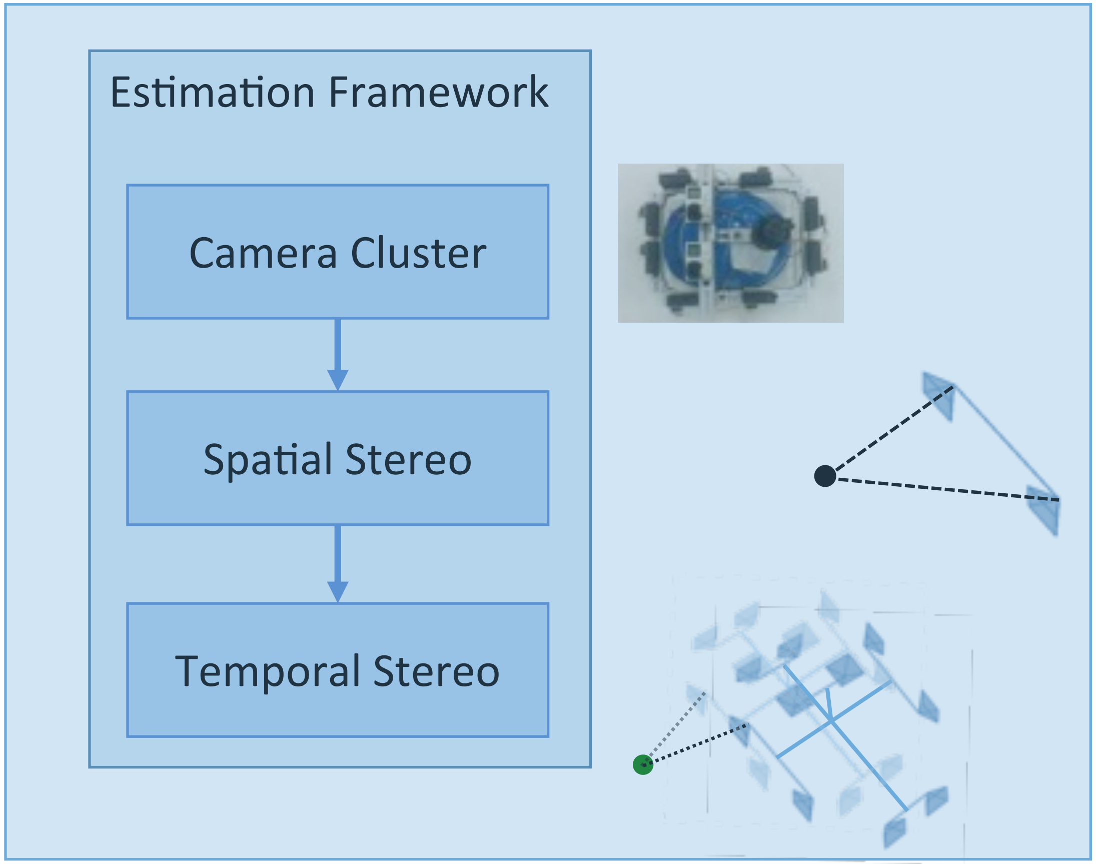
Temporal Stereo







Approach Overview



Estimation Framework

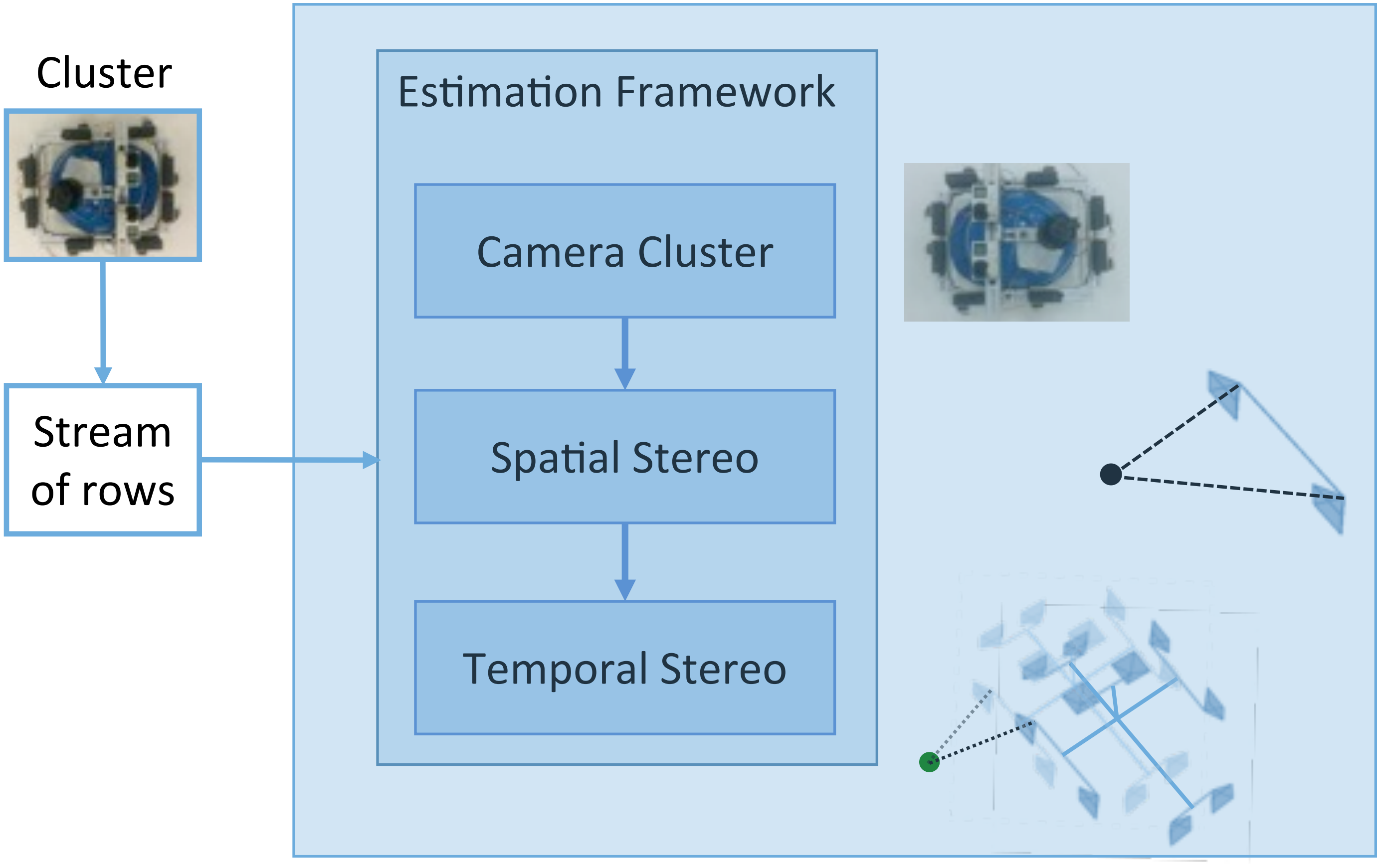
Camera Cluster

Spatial Stereo

Temporal Stereo

Our approach

Approach Overview



Cluster

Stream of rows

Estimation Framework

Camera Cluster

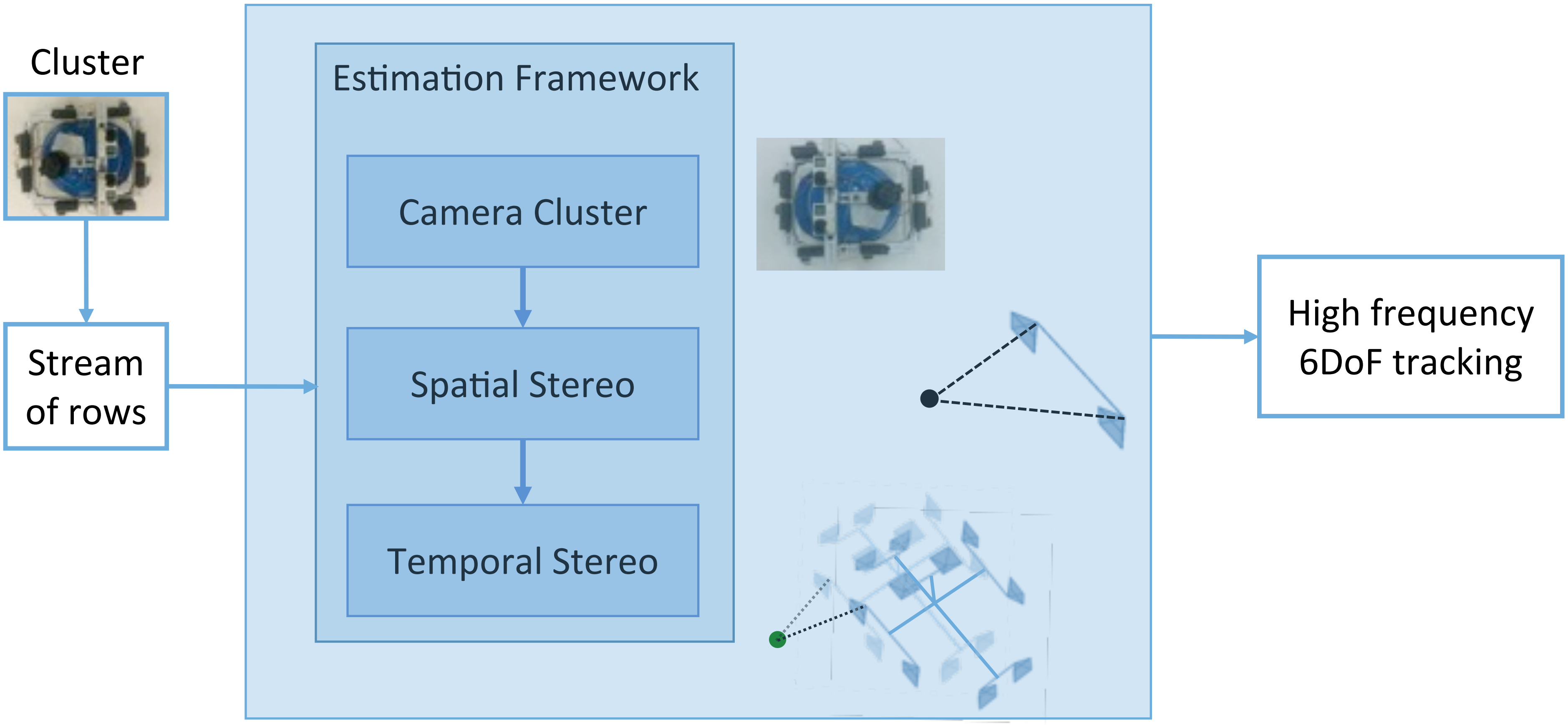
Spatial Stereo

Temporal Stereo

Input

Our approach

Approach Overview



Cluster

Stream of rows

Estimation Framework

Camera Cluster

Spatial Stereo

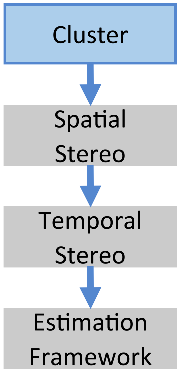
Temporal Stereo

High frequency 6DoF tracking

Input

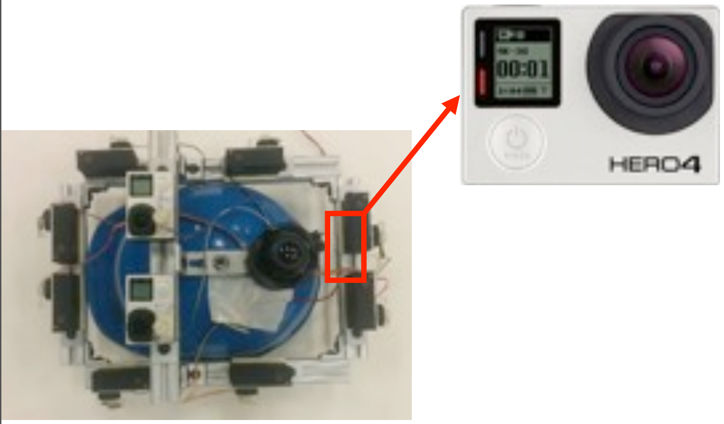
Our approach

Output



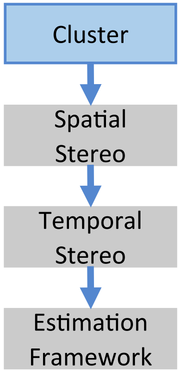
Our Cluster

GoPro[1]



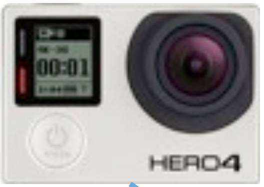
[1] http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html

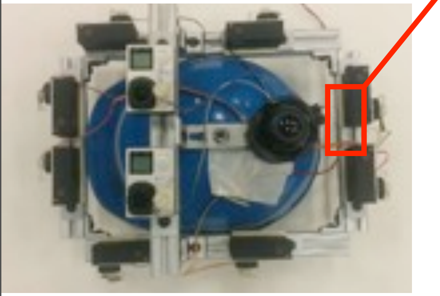
[2] http://users.erols.com/njastro/faas/image001/CCD2.jpg



Our Cluster

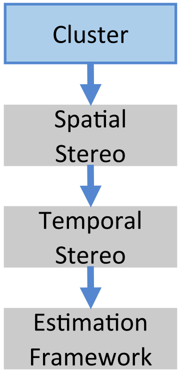
GoPro[1]





Dynamic sensor with periodic movement[2]

[1] http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html
[2] http://users.erols.com/njastro/faas/image001/CCD2.jpg



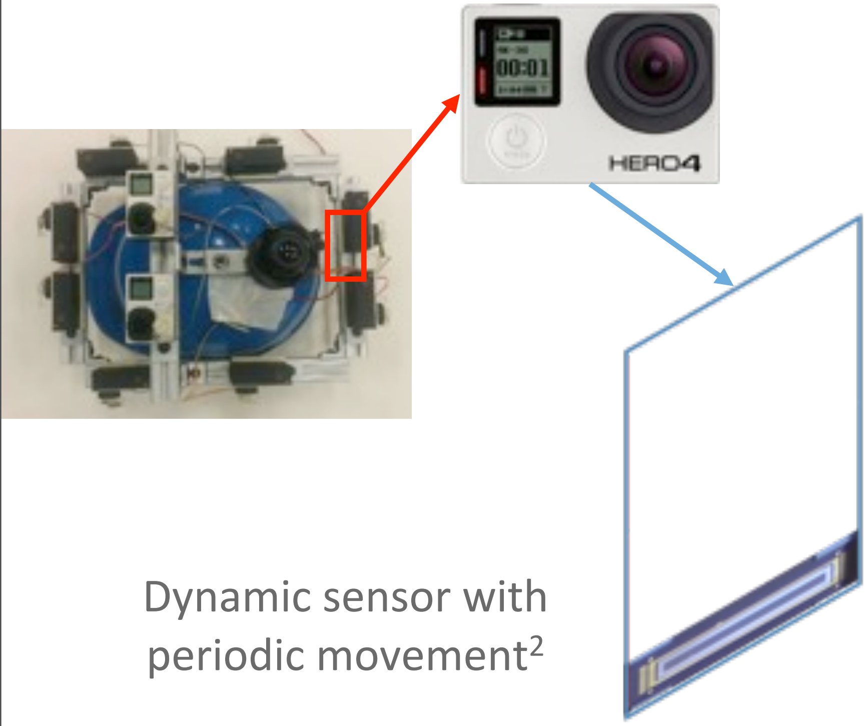
Cluster

Spatial Stereo

Temporal Stereo

Estimation Framework

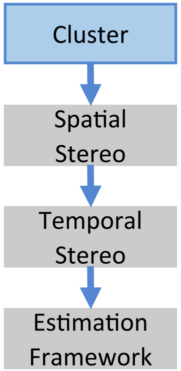
Our Cluster



GoPro[1]

Dynamic sensor with periodic movement[2]

[1] http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html
[2] http://users.erols.com/njastro/faas/image001/CCD2.jpg



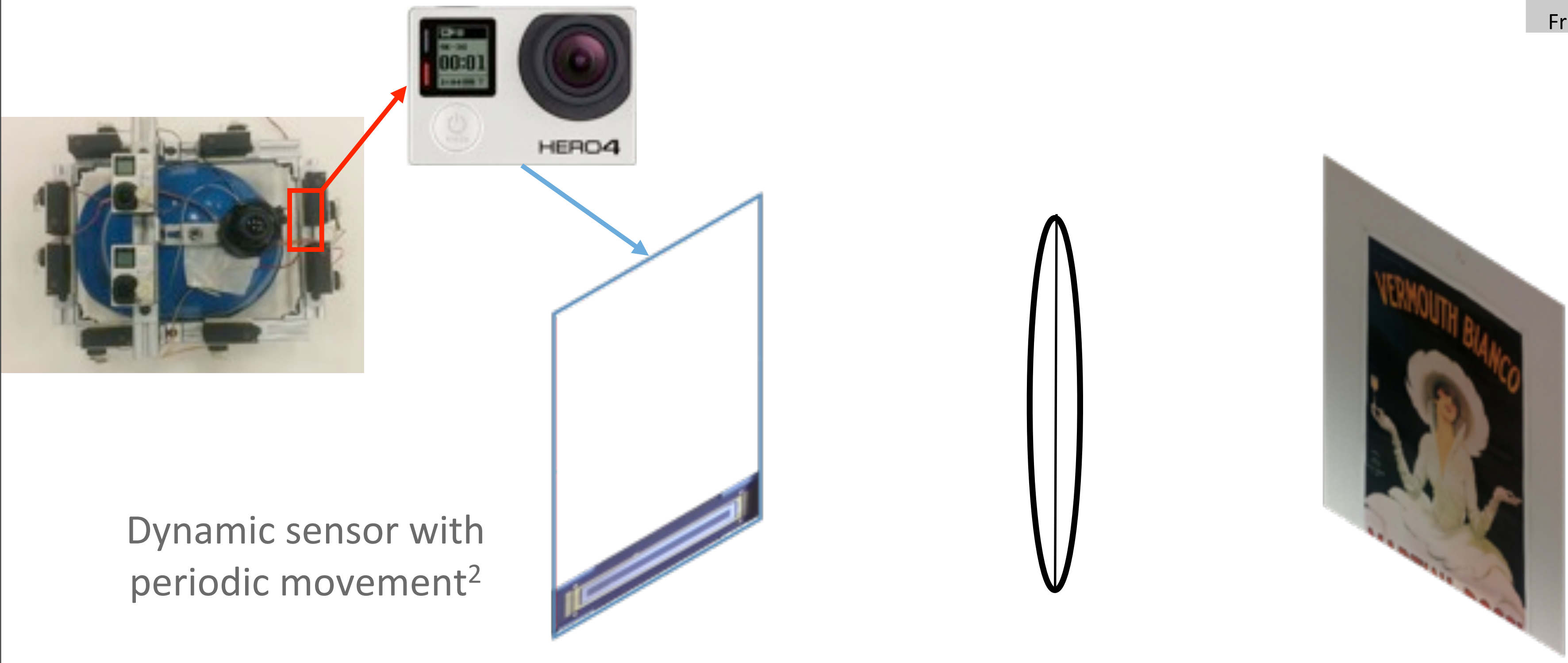
Cluster

Spatial Stereo

Temporal Stereo

Estimation Framework

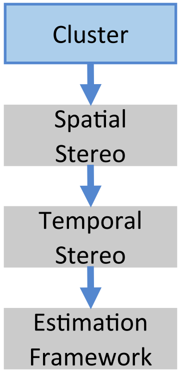
Our Cluster



GoPro[1]

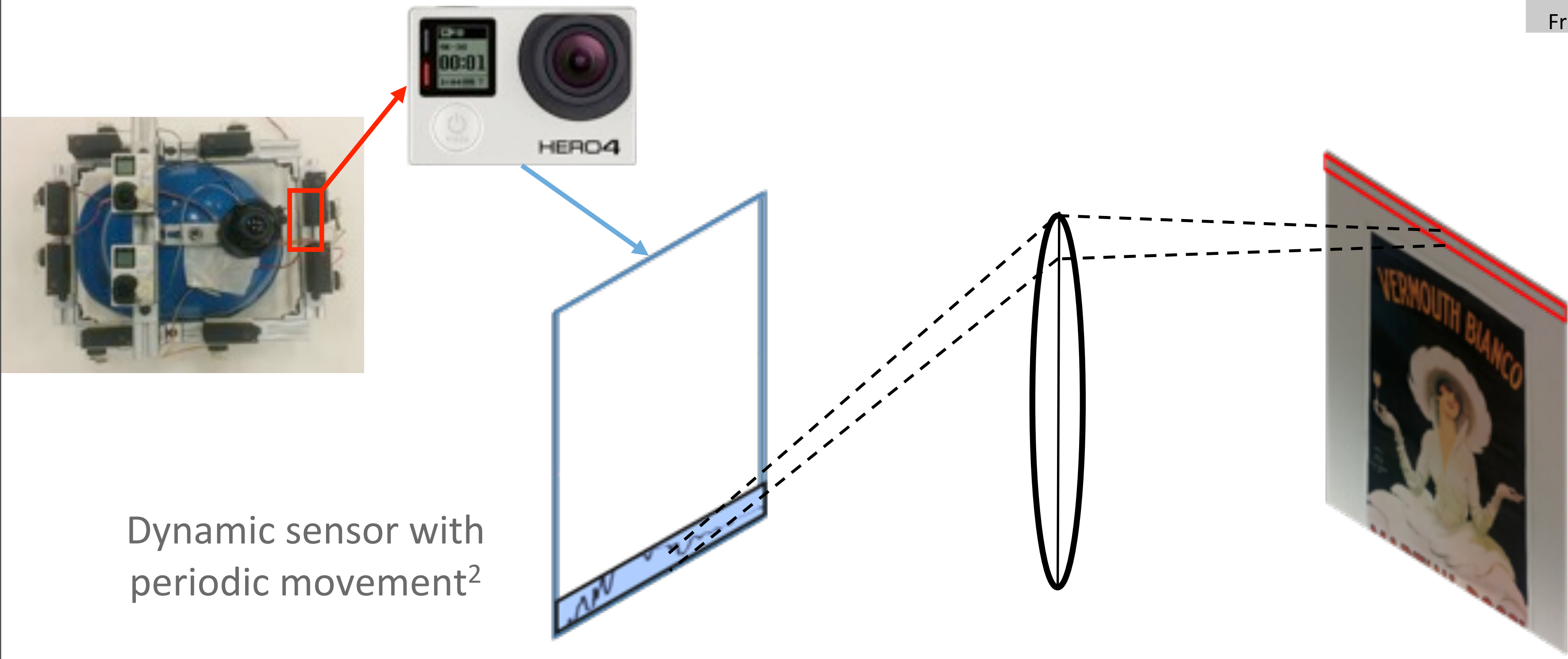
Dynamic sensor with periodic movement[2]

1. http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html
2. http://users.erols.com/njastro/faas/image001/CCD2.jpg



Our Cluster

GoPro[1]



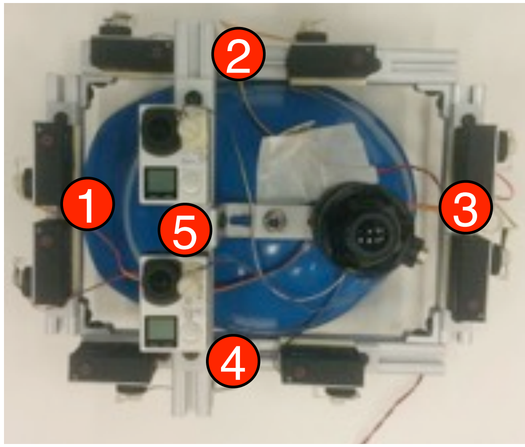
Dynamic sensor with periodic movement[2]

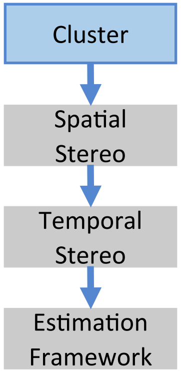
Small vertical FoV

1. http://shop.gopro.com/hero4/hero4-black/CHDHX-401.html
2. http://users.erols.com/njastro/faas/image001/CCD2.jpg

Our Cluster

●





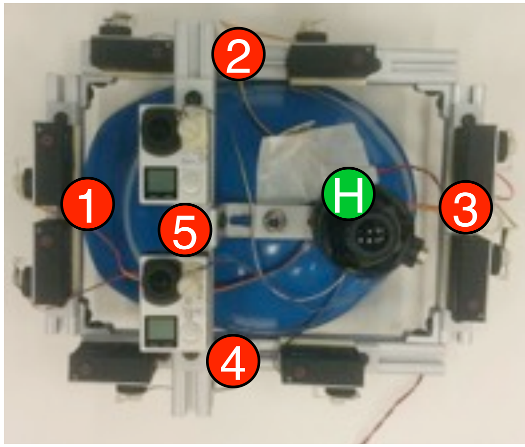
Cluster

Spatial Stereo

Temporal Stereo

Estimation Framework

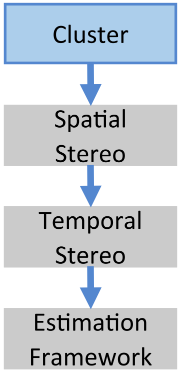
Our Cluster



● N stereo-pairs of rolling shutter camera, N=5

- Precalibrated intrinsic and extrinsics
- Temporal sync
- Known history of motion

H Hi-Ball for ground truth


Cluster

Spatial Stereo

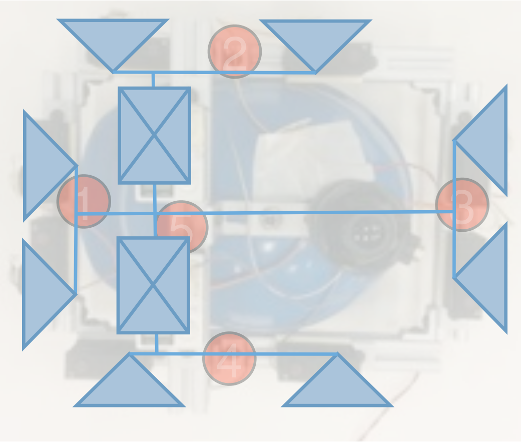
Temporal Stereo

Estimation Framework

Our Cluster

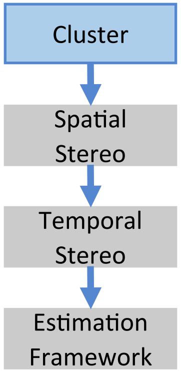
● N stereo-pairs of rolling shutter camera, N=5

- Precalibrated intrinsic and extrinsics
- Temporal sync
- Known history of motion



Geometry of cluster

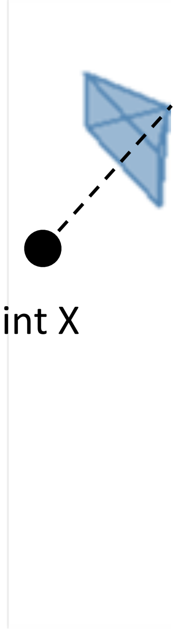
Ⓗ Hi-Ball for ground truth

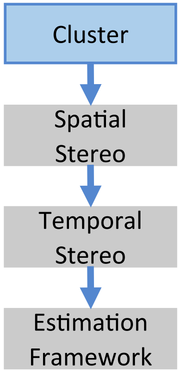


Motion Model

3D point

$$X^{t+1}_{cam_i} = M_{cam_i} V^1 \boxed{X^t_{cam_i}}$$

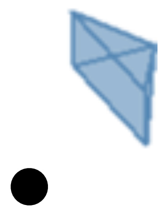


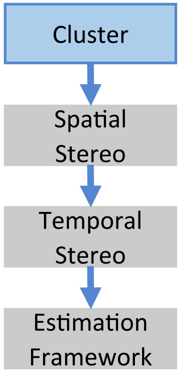

Cluster
Spatial Stereo
Temporal Stereo
Estimation Framework

Motion Model

3D point

$$X^{t+1}_{cam_i} = M_{cam_i} V^1 \boxed{X^{t}_{cam_i}}$$



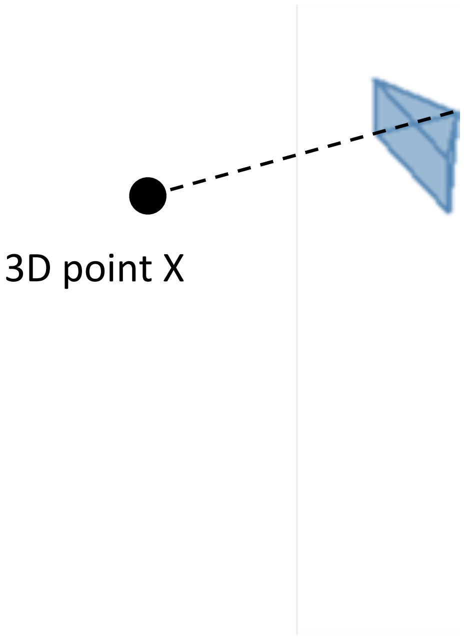


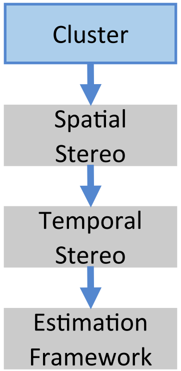
Motion Model

3D point

$$X^{t+1}_{cam_i} = \boxed{\boldsymbol{M}_{cam_i}} \boxed{X^{t}_{cam_i}}$$

Motion of camera





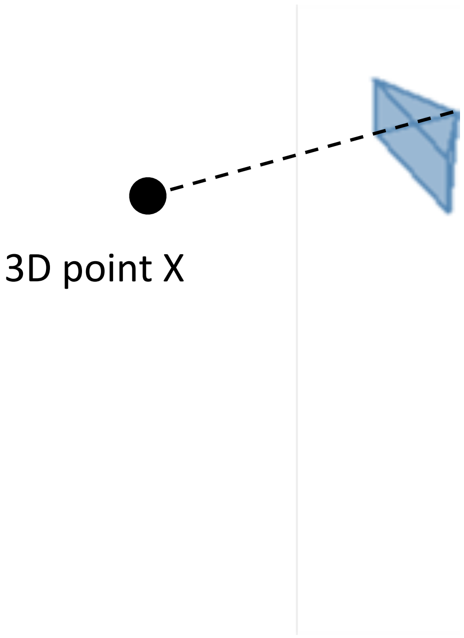
Motion Model

Moved 3D point

3D point

$$\boxed{X^{t+1}_{cam_i}} = \boxed{M_{cam_i}}\ \boxed{X^{t}_{cam_i}}$$

Motion of camera



Cluster

Spatial Stereo

Temporal Stereo

Estimation Framework

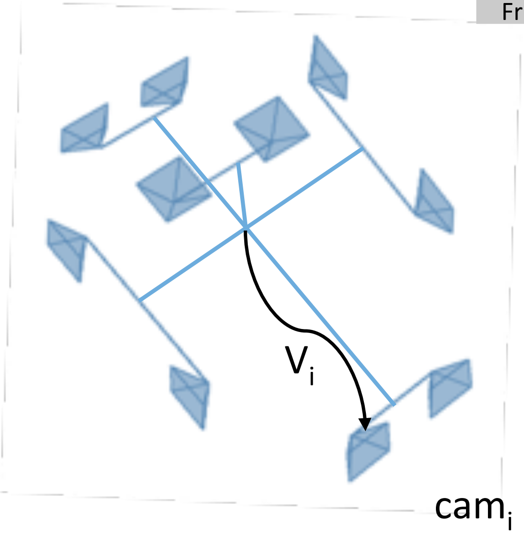
Motion Model

Moved 3D point

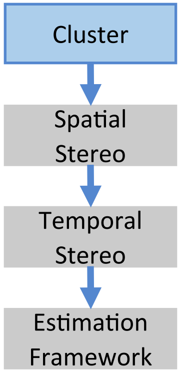
3D point

$$X^{t+1}_{cam_i} = M_{cam_i} X^{t}_{cam_i}$$

$$X^{t+1}_{cam_i} = V_i \, M_{cluster} \, V_i^{-1} \, X^{t}_{cam_i}$$



RS cluster



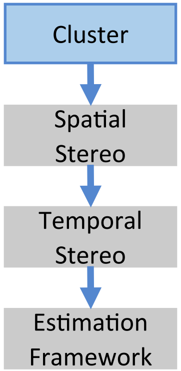
Linearized Motion Model

$$\boldsymbol{X}_{cam_i}^{t+1} = \boldsymbol{M}_{cam_i} \quad \boldsymbol{X}_{cam_i}^{t}$$

$$\boldsymbol{X}_{cam_i}^{t+1} = \boldsymbol{V}_i \ \boldsymbol{M}_{cluster} \ \boldsymbol{V}_i^{-1} \ \boldsymbol{X}_{cam_i}^{t}$$

$$\boldsymbol{M}_{\boldsymbol{cluster}} = \begin{bmatrix} R(\theta_x, \theta_y, \theta_z) & T \\ 0 & 1 \end{bmatrix}$$

Linearized Motion Model



Cluster

Spatial Stereo

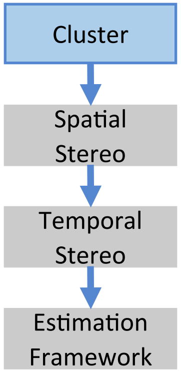
Temporal Stereo

Estimation Framework

$$\boldsymbol{X}_{cam_i}^{t+1} = \boldsymbol{M}_{cam_i} \boldsymbol{X}_{cam_i}^{t}$$

$$\boldsymbol{X}_{cam_i}^{t+1} = \boldsymbol{V}_i \, \boldsymbol{M}_{cluster} \, \boldsymbol{V}_i^{-1} \, \boldsymbol{X}_{cam_i}^{t}$$

$$\boldsymbol{M}_{cluster} = \begin{bmatrix} R(\theta_x, \theta_y, \theta_z) & T \\ 0 & 1 \end{bmatrix} \Rightarrow \boldsymbol{dM} = \begin{bmatrix} 1 & -\theta_z & \theta_y & \delta T_x \\ \theta_z & 1 & -\theta_x & \delta T_y \\ -\theta_y & \theta_x & 1 & \delta T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Cluster

Spatial Stereo

Temporal Stereo

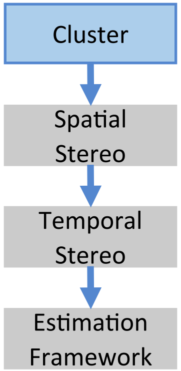
Estimation Framework

Linearized Motion Model

$$\boldsymbol{X}_{cam_i}^{t+1} = \boldsymbol{M}_{cam_i} \quad \boldsymbol{X}_{cam_i}^{t}$$

$$\boldsymbol{X}_{cam_i}^{t+1} = \boxed{V_i} \boldsymbol{M}_{cluster} \boxed{V_i^{-1}} \boldsymbol{X}_{cam_i}^{t}$$

$$\boldsymbol{M}_{cluster} = \begin{bmatrix} R(\theta_x, \theta_y, \theta_z) & T \\ 0 & 1 \end{bmatrix} \Rightarrow \boldsymbol{dM} = \begin{bmatrix} 1 & -\theta_z & \theta_y & \delta T_x \\ \theta_z & 1 & -\theta_x & \delta T_y \\ -\theta_y & \theta_x & 1 & \delta T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Cluster

Spatial Stereo

Temporal Stereo

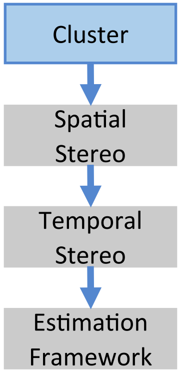
Estimation Framework

Linearized Motion Model

$$X^{t+1}_{cam_i} = M_{cam_i} X^{t}_{cam_i}$$

$$\boxed{X^{t+1}_{cam_i}} = \boxed{V_i} M_{cluster} \boxed{V_i^{-1}} \boxed{X^{t}_{cam_i}}$$

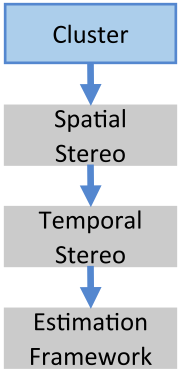
$$M_{cluster} = \begin{bmatrix} R(\theta_x, \theta_y, \theta_z) & T \\ 0 & 1 \end{bmatrix} \Rightarrow dM = \begin{bmatrix} 1 & -\theta_z & \theta_y & \delta T_x \\ \theta_z & 1 & -\theta_x & \delta T_y \\ -\theta_y & \theta_x & 1 & \delta T_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$



Motion Estimation

$$X^{t+1}_{cam_i} = V_i \, M_{cluster} \; V_i^{-1} \, \boxed{X^{t}_{cam_i}}$$

$$X^{t}_{cam_i} = [0 \quad y \quad d \quad 1]^T$$



Cluster

Spatial Stereo

Temporal Stereo

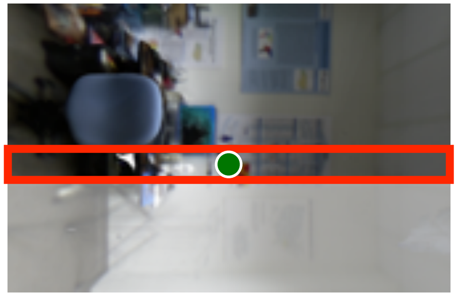
Estimation Framework

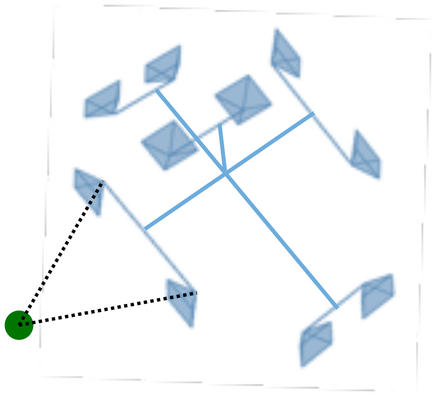
Motion Estimation

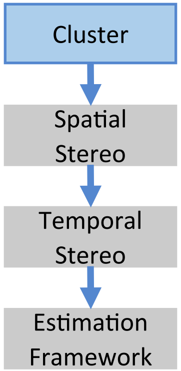
$$X^{t+1}_{cam_i} = V_i \; M_{cluster} \; V_i^{-1} \; \boxed{X^{t}_{cam_i}}$$

$$X^{t}_{cam_i} = [0 \quad y \quad d \quad 1]^T$$

Stereo in space







Cluster

Spatial Stereo

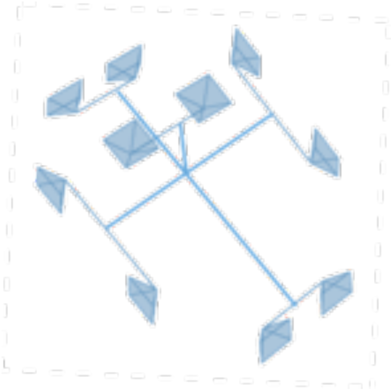
Temporal Stereo

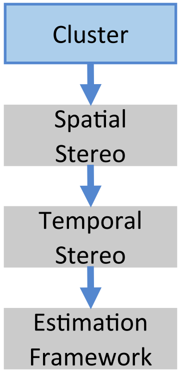
Estimation Framework

Motion Estimation

$$X_{cam_i}^{t+1} = [\Delta x \quad y+\Delta y \quad d+\Delta d \quad 1]^T$$

$$\boxed{X_{cam_i}^{t+1}} = V_i \; M_{cluster} \; V_i^{-1} \; X_{cam_i}^{t}$$

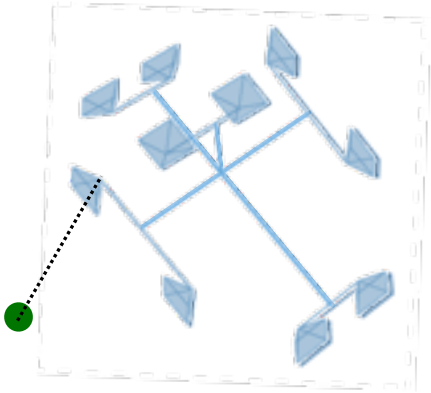


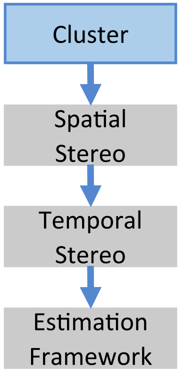


Motion Estimation

$$X^{t+1}_{cam_i} = [\Delta x \quad y + \Delta y \quad d + \Delta d \quad 1]^T$$

$$\boxed{X^{t+1}_{cam_i}} = \boldsymbol{V}_i \, \boldsymbol{M}_{cluster} \quad V_i^{-1} \quad \boldsymbol{X}^{t}_{cam_i}$$





Cluster

Spatial Stereo

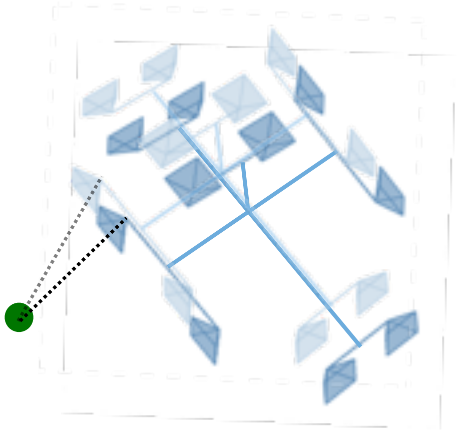
Temporal Stereo

Estimation Framework

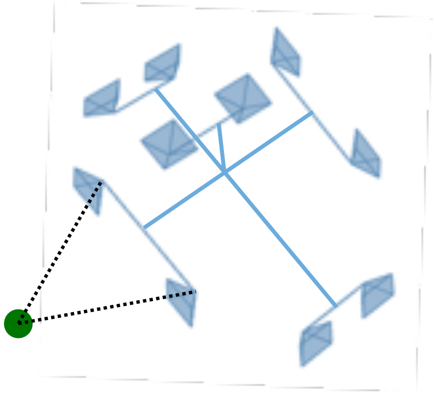
Motion Estimation

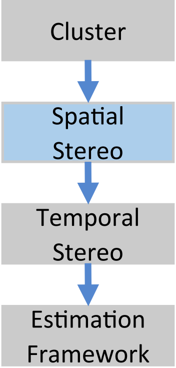
$$X^{t+1}_{cam_i} = [\Delta x \quad y+\Delta y \quad d+\Delta d \quad 1]^T$$

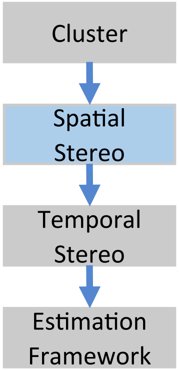
$$\boxed{X^{t+1}_{cam_i}} = V_i \; M_{cluster} \; V_i^{-1} \; X^{t}_{cam_i}$$



Stereo Estimation

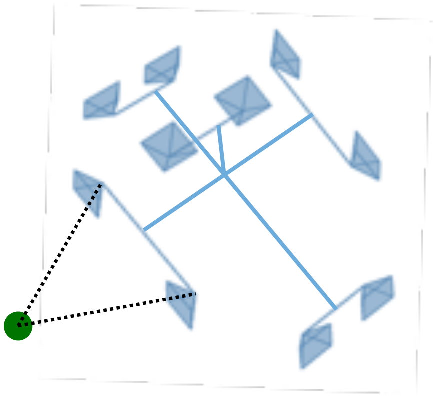


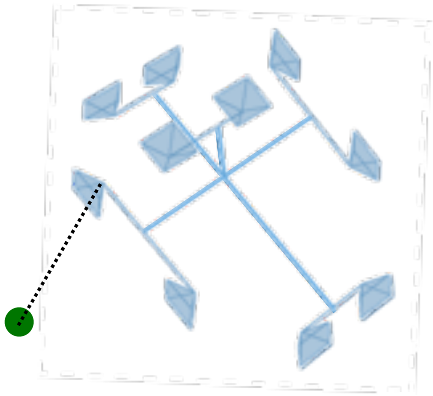




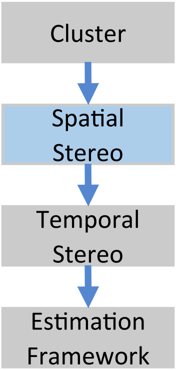
Stereo Estimation

- Spatial stereo s_d
  - Use stereo-pair of cameras
  - Measure pixel disparity
- Temporal stereo s_t
  - Measure small shifts in pixel
  - Compare row at time t and t-k



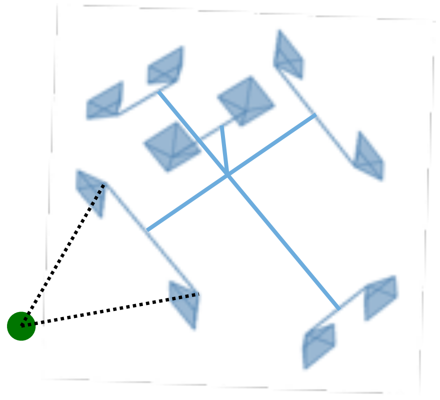


t1 → t2

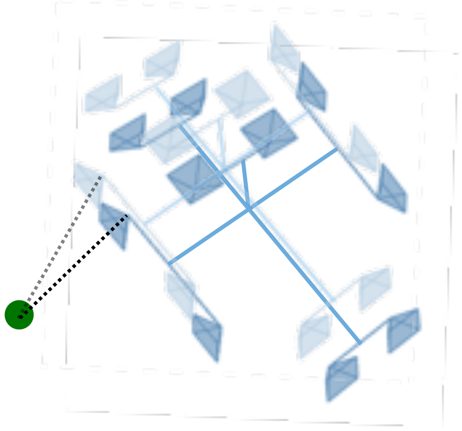


Stereo Estimation

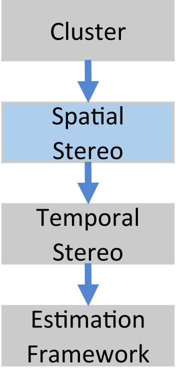
- Spatial stereo s_d
  - Use stereo-pair of cameras
  - Measure pixel disparity



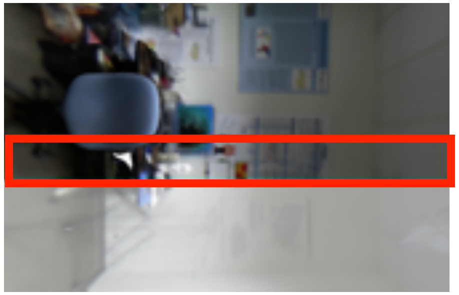
- Temporal stereo s_t
  - Measure small shifts in pixel
  - Compare row at time t and t-k

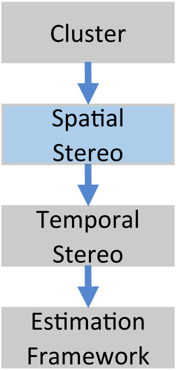


t1 ⟶ t2

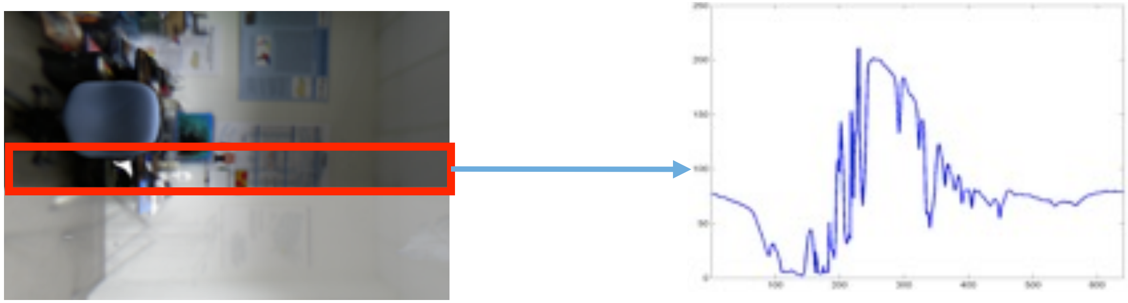


Binary Row Descriptor



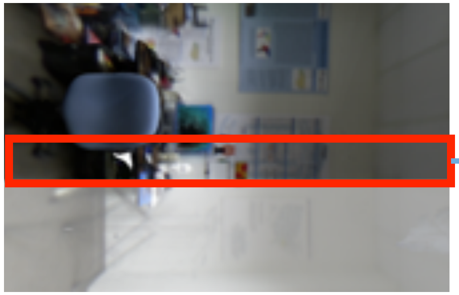


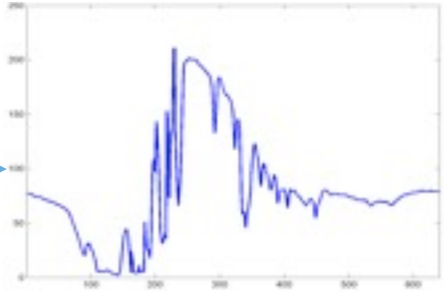
Binary Row Descriptor



Visualization of row-image (296th)

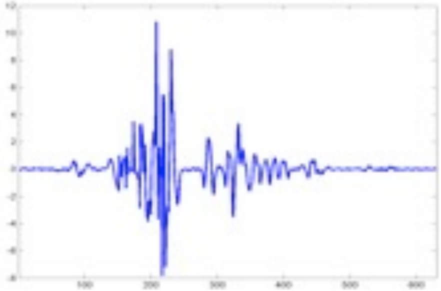
Binary Row Descriptor





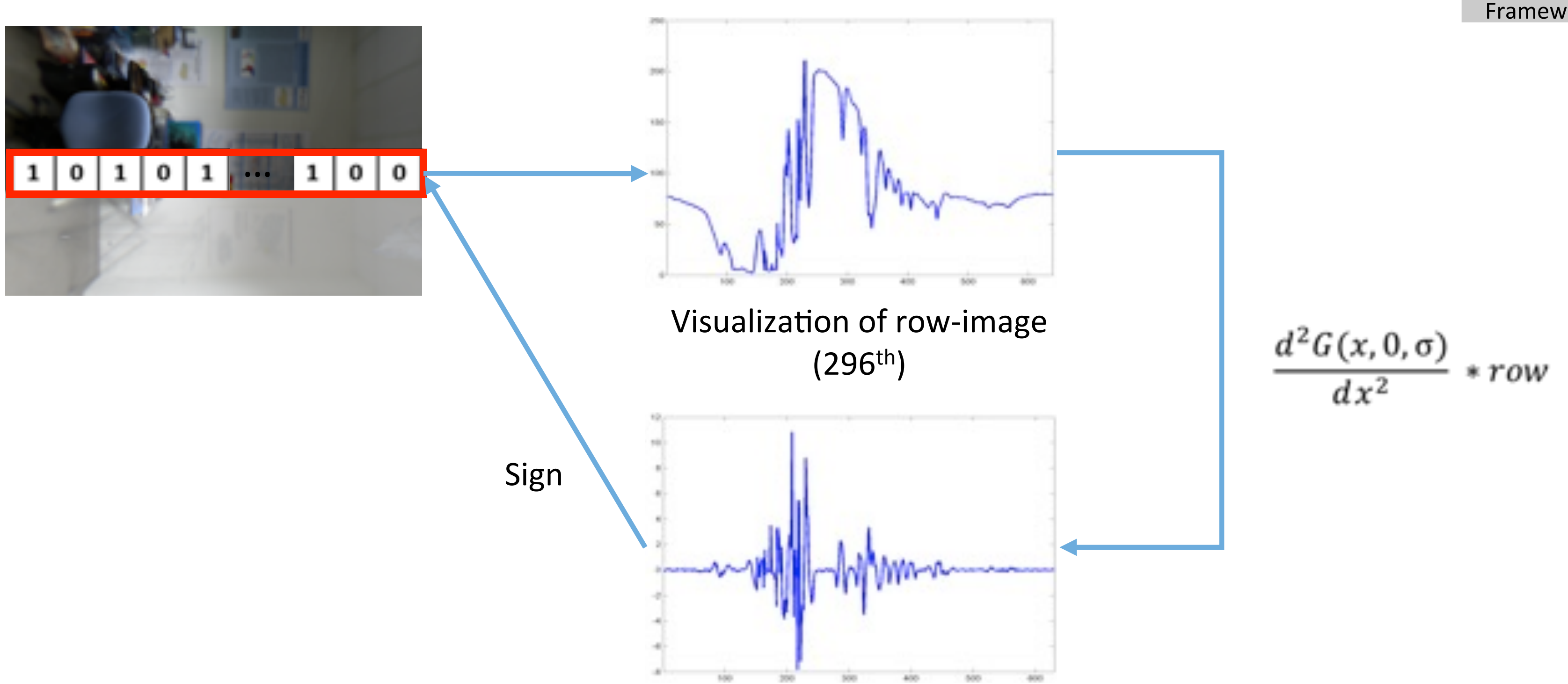
Visualization of row-image (296th)

$$\frac{d^2G(x,0,\sigma)}{dx^2} * row$$



2nd derivative

Binary Row Descriptor



Visualization of row-image (296th)

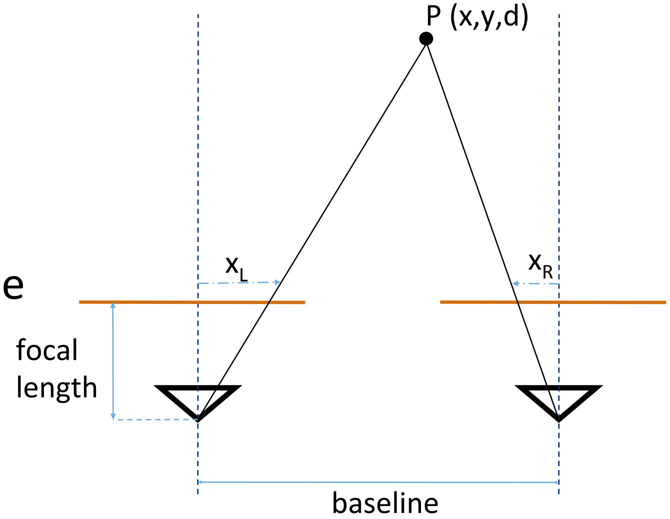
$$\frac{d^2G(x,0,\sigma)}{dx^2} * row$$

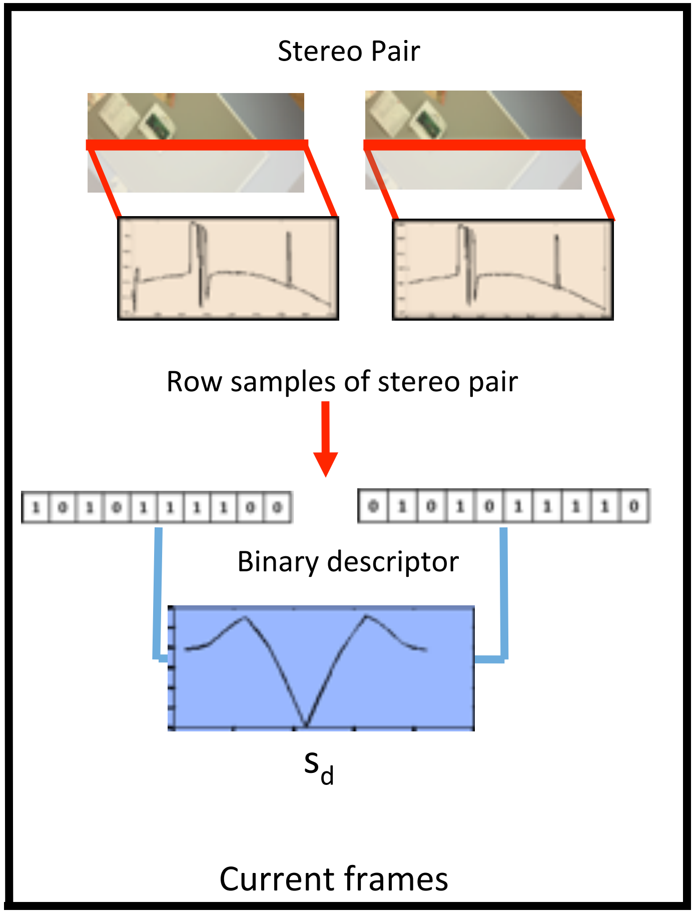
2nd derivative

Spatial Stereo : Measure Disparity

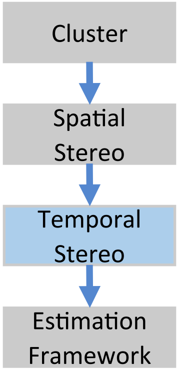
- Compare binary descriptor of rows of stereo cameras
- Fast hamming cost matching
- $depth = \frac{focal * baseline}{disparity}$

Stereo in space





Temporal Stereo: Measure Shift s_t

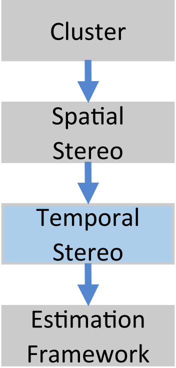

Cluster

Spatial Stereo

Temporal Stereo

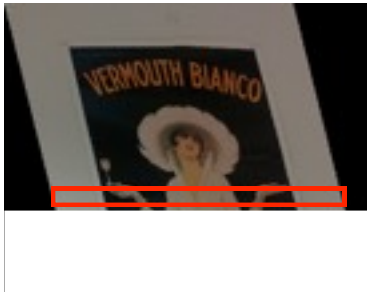
Estimation Framework

- Different regions in space are captured at different timestamps
- Stereo in time
  - Need snaps of same space but at different timestamps

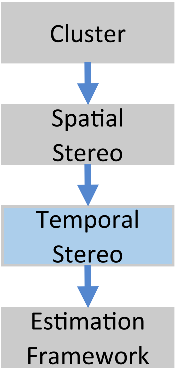


Temporal Stereo: Measure Shift s_t

- Different regions in space are captured at different timestamps
- Stereo in time
  - Need snaps of same space but at different timestamps

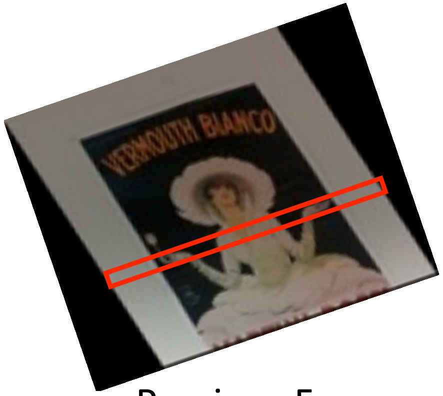


Current time

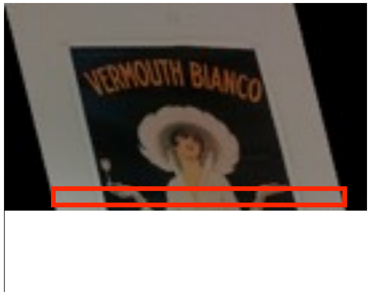


Temporal Stereo: Measure Shift s_t

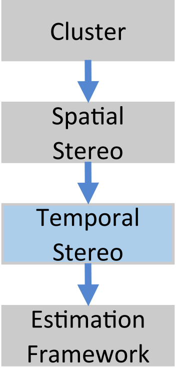
- Different regions in space are captured at different timestamps
- Stereo in time
  - Need snaps of same space but at different timestamps



Previous Frame

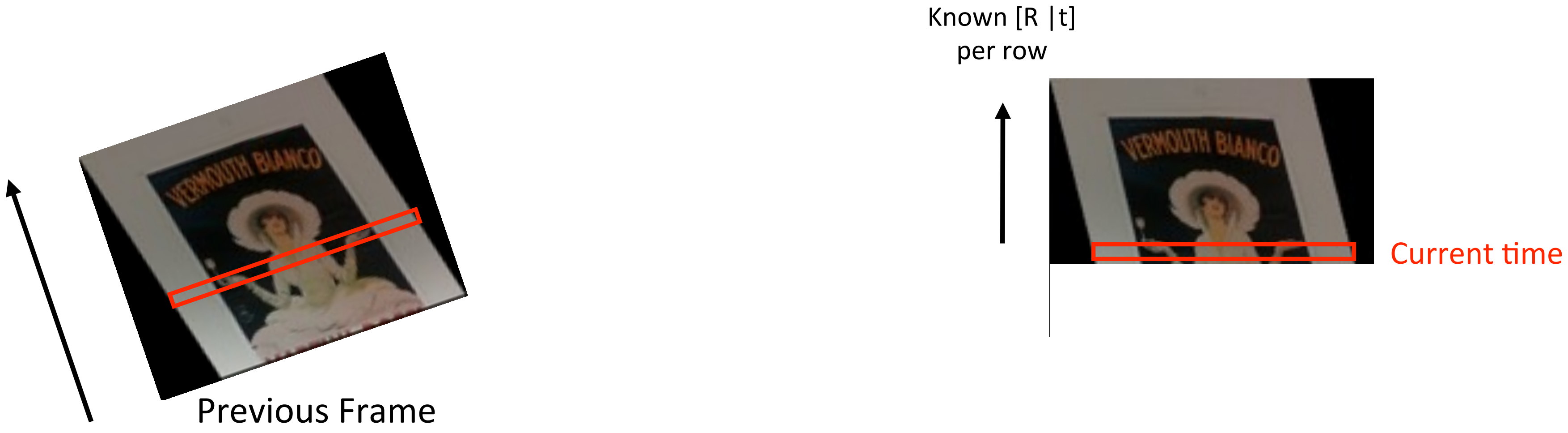


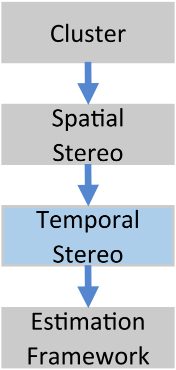
Current time



Temporal Stereo: Measure Shift s_t

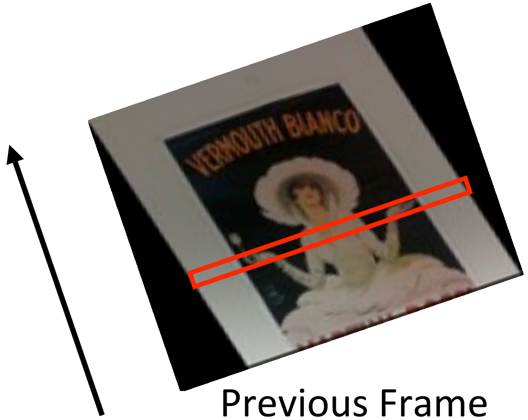
- Different regions in space are captured at different timestamps
- Stereo in time
  - Need snaps of same space but at different timestamps



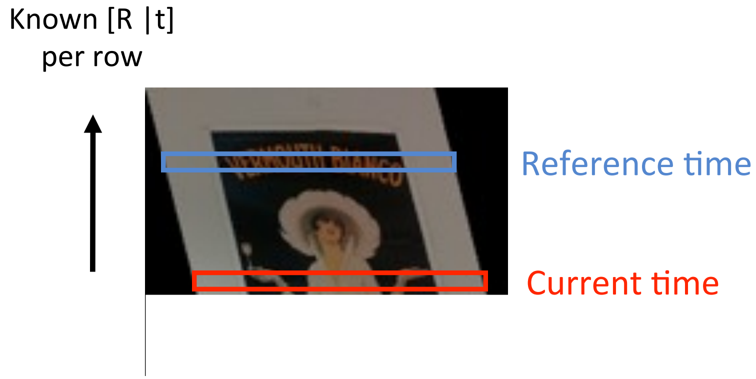


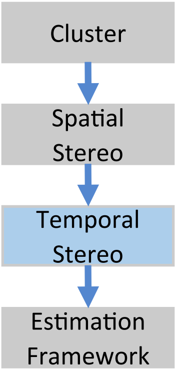
Temporal Stereo: Measure Shift s_t

- Different regions in space are captured at different timestamps
- Stereo in time
  - Need snaps of same space but at different timestamps



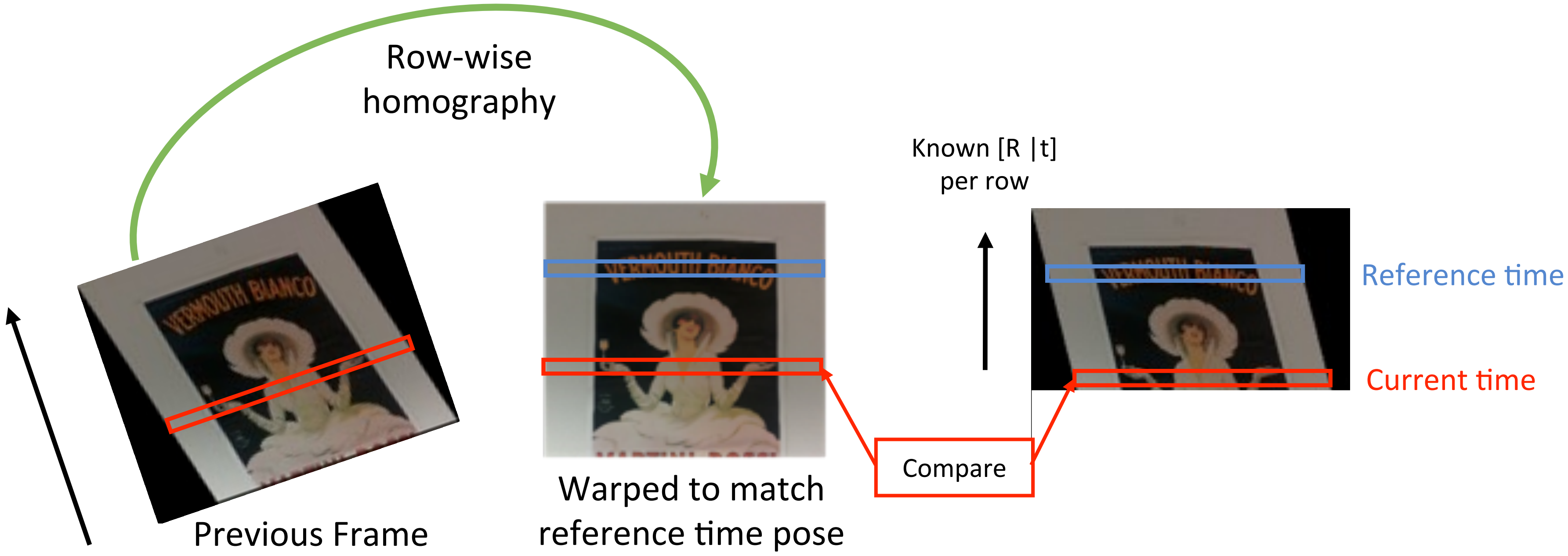
Previous Frame

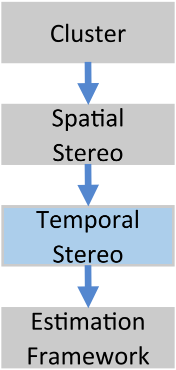




Temporal Stereo: Measure Shift s_t

- Different regions in space are captured at different timestamps
- Stereo in time
  - Need snaps of same space but at different timestamps





Cluster

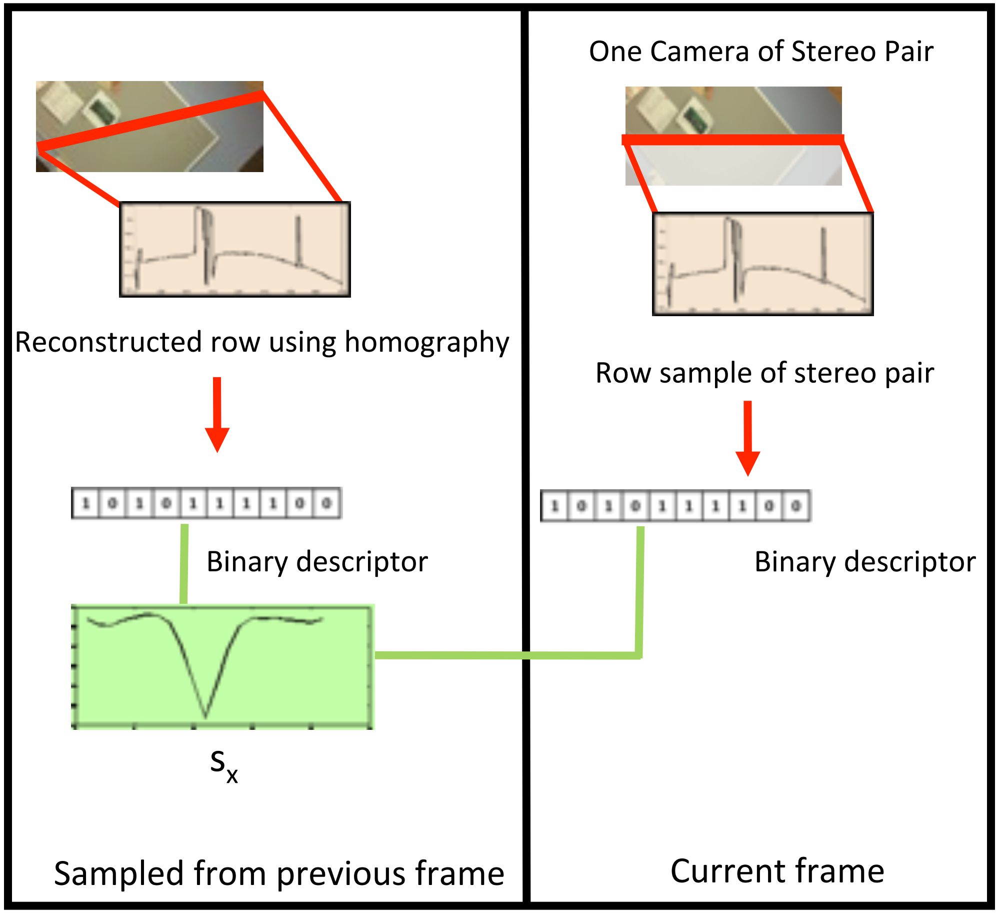
Spatial Stereo

Temporal Stereo

Estimation Framework

Temporal Stereo: Measure Shift s_t

- Reconstruct row using per- row homography
- Use binary descriptors and hamming distance



Reconstructed row using homography

1 0 1 0 1 1 1 1 0 0

Binary descriptor

S_x

Sampled from previous frame

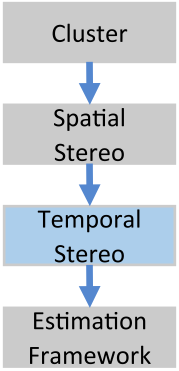
One Camera of Stereo Pair

Row sample of stereo pair

1 0 1 0 1 1 1 1 0 0

Binary descriptor

Current frame



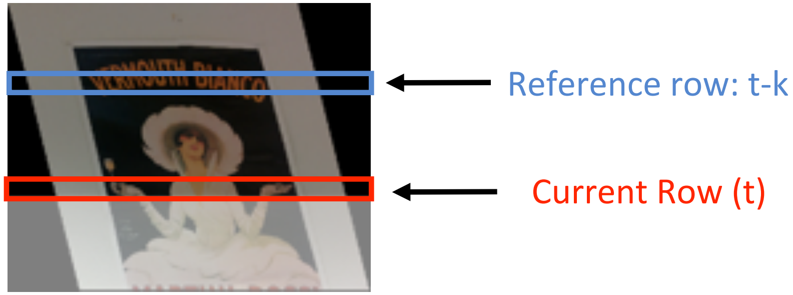
Cluster

Spatial Stereo

Temporal Stereo

Estimation Framework

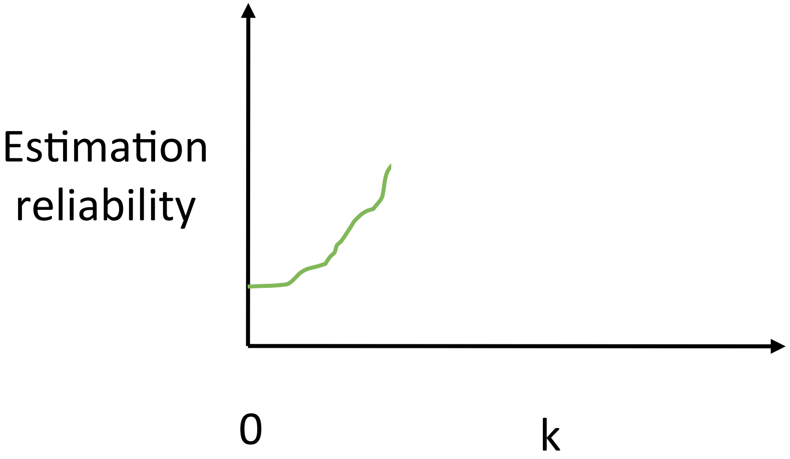
Adaptive Reference



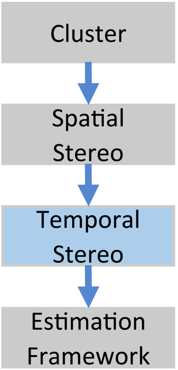
Reference row: t-k

Current Row (t)

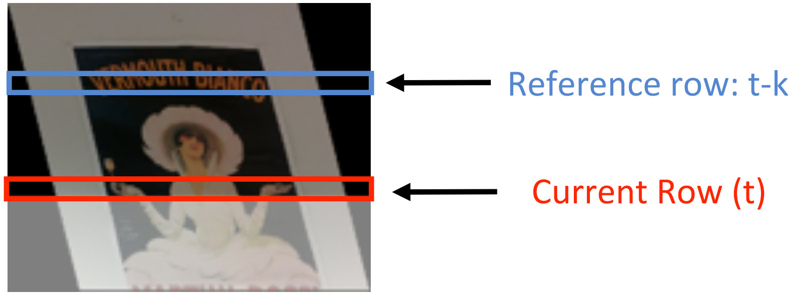
Current frame



Adaptive Reference

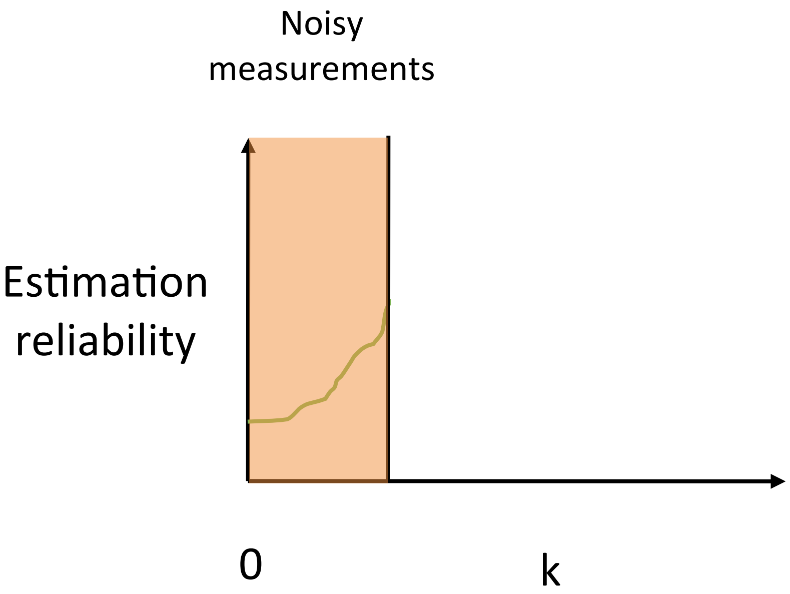


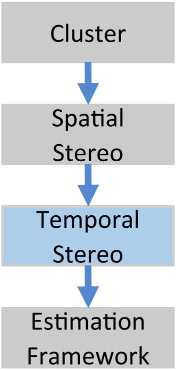
- Leave sufficient motion
  - Row-to-row motion
  - Interpolation & pixel measurement noise



Current frame

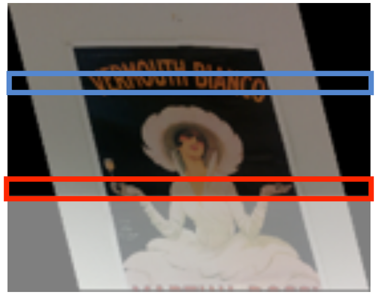
- Satisfy small-motion assumption
  - Reference should not be far away





Adaptive Reference

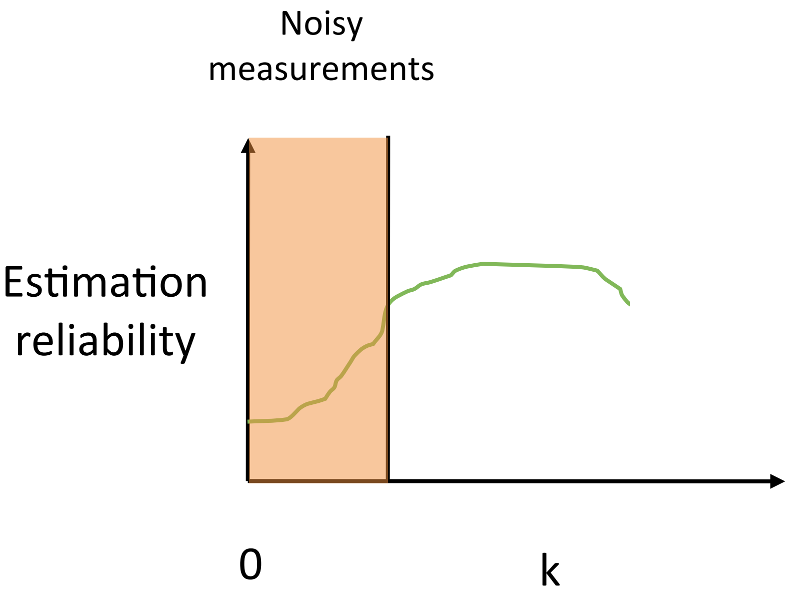
- Leave sufficient motion
  - Row-to-row motion
  - Interpolation & pixel measurement noise

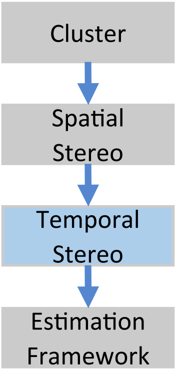

Reference row: t-k

Current Row (t)

Current frame

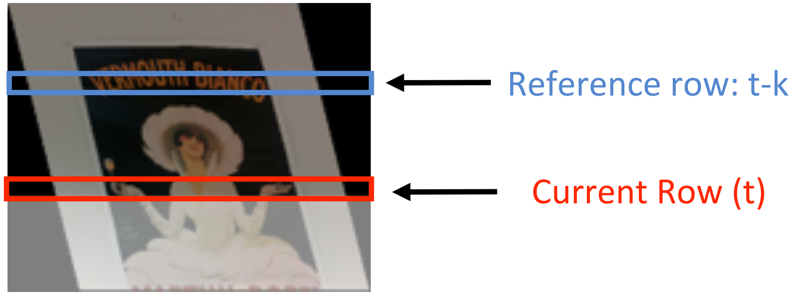
- Satisfy small-motion assumption
  - Reference should not be far away





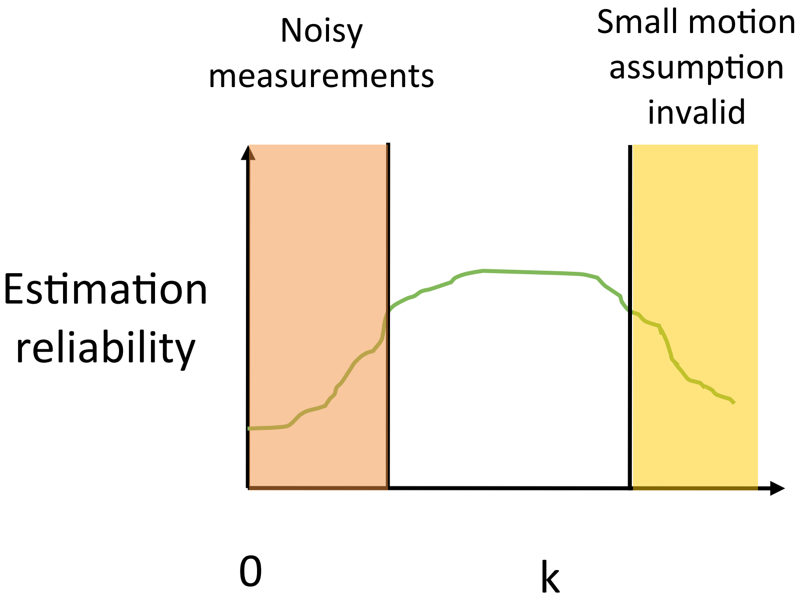
Adaptive Reference

- Leave sufficient motion
  - Row-to-row motion
  - Interpolation & pixel measurement noise



Current frame

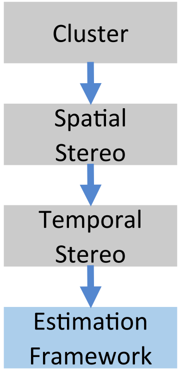
- Satisfy small-motion assumption
  - Reference should not be far away



Confidence Scores

High confidence

Low Confidence



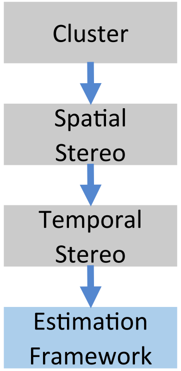
1. Quality of minimum: c_{PKR}
   - Is the valley unique?
   - Use Peak ratio (PKR)[1]

2. Temporal consistency : c_t
   - Consistent shifts in time
   - Penalize sudden changes

3. $c_{i,t} = c_{PKR} * c_t$

[1] X. Hu and P. Mordohai. A quantitative evaluation of confidence measures for stereo vision, PAMI 2012

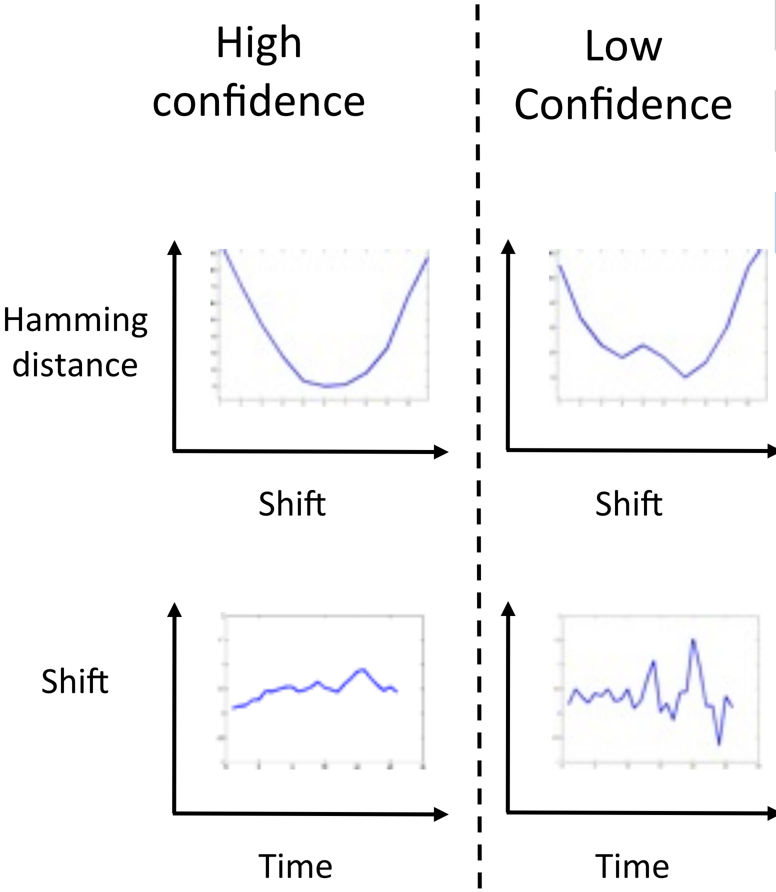
Confidence Scores



Cluster → Spatial Stereo → Temporal Stereo → Estimation Framework

1. Quality of minimum: c_{PKR}
   - Is the valley unique?
   - Use Peak ratio (PKR)[1]

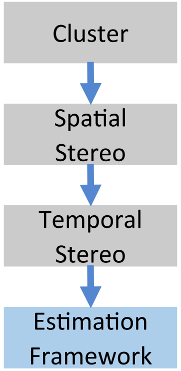
2. Temporal consistency : c_t
   - Consistent shifts in time
   - Penalize sudden changes



3. $c_{i,t} = c_{PKR} * c_t$

$$\mathbf{C}(t) = \begin{bmatrix} c_{1,t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_{n,t} \end{bmatrix}$$

[1] X. Hu and P. Mordohai. A quantitative evaluation of confidence measures for stereo vision, PAMI 2012



Motion Estimation

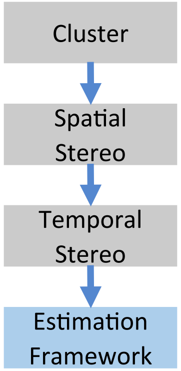
$$X^{t+1}_{cam_i} = [\Delta \mathrm{x} \quad \mathrm{y} + \Delta y \quad \mathrm{d} + \Delta d \quad 1]^T$$

Stereo in time

$$\boxed{X^{t+1}_{cam_i}} = V_i \, M_{cluster} \, V_i^{-1} \quad \boxed{X^{t}_{cam_i}}$$

$$X^{t}_{cam_i} = [0 \quad \mathrm{y} \quad \mathrm{d} \quad 1]^T$$

Stereo in space



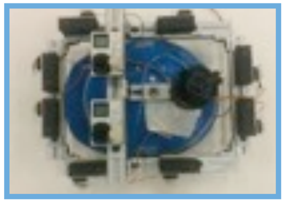
Weighted Linear System

- $X_{cam_i}^{t+1} = V_i \, M_{cluster} \, V_i^{-1} X_{cam_i}^{t}$
  - One equation from each camera
  - Use more cameras for robustness

$$C\,A \begin{bmatrix} \delta T_x \\ \delta T_y \\ \delta T_z \\ \theta_x \\ \theta_y \\ \theta_z \end{bmatrix} = C\;B$$

Approach Summary

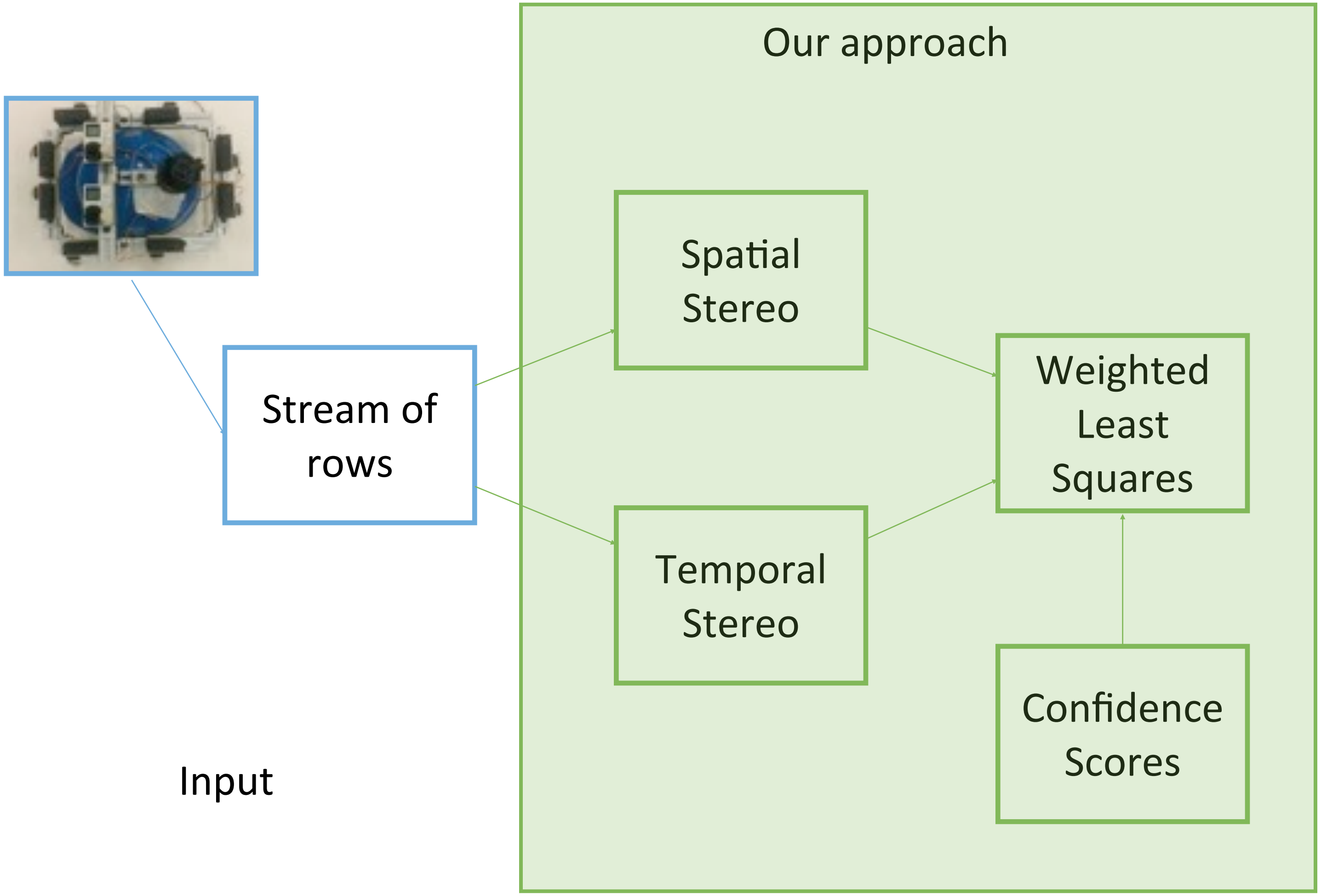
Approach Summary



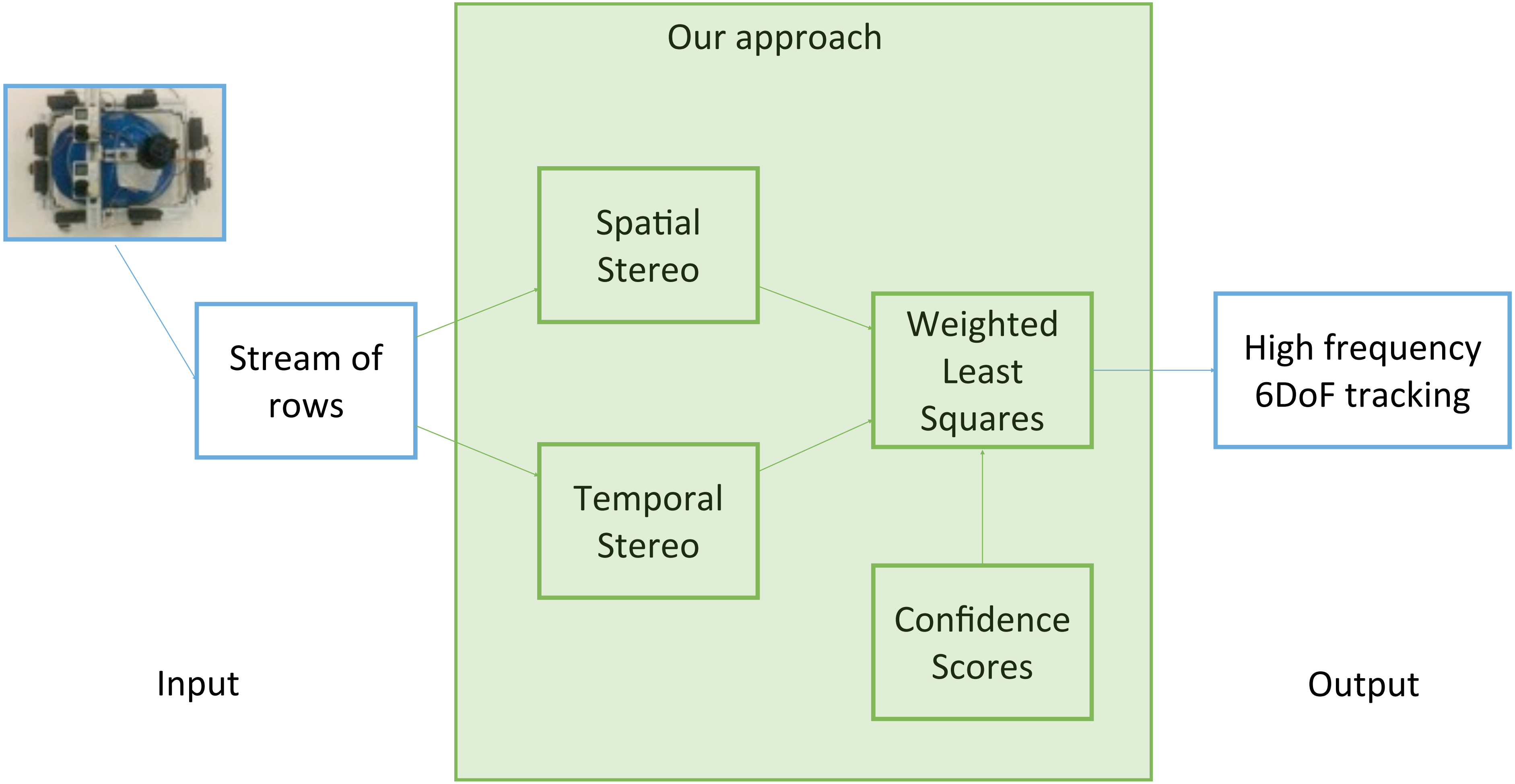
Stream of rows

Input

Approach Summary



Approach Summary



Our approach

Spatial Stereo

Temporal Stereo

Weighted Least Squares

Confidence Scores

Stream of rows

High frequency 6DoF tracking

Input

Output

Simulator

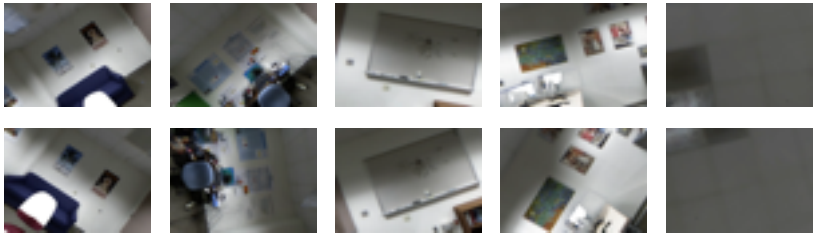
- Developed in OpenGL+Qt and Unity3D
  - Support for Hi-Ball tracker data
  - Motion Blur

Simulator

- Developed in OpenGL+Qt and Unity3D
  - Support for Hi-Ball tracker data
  - Motion Blur

Experiments

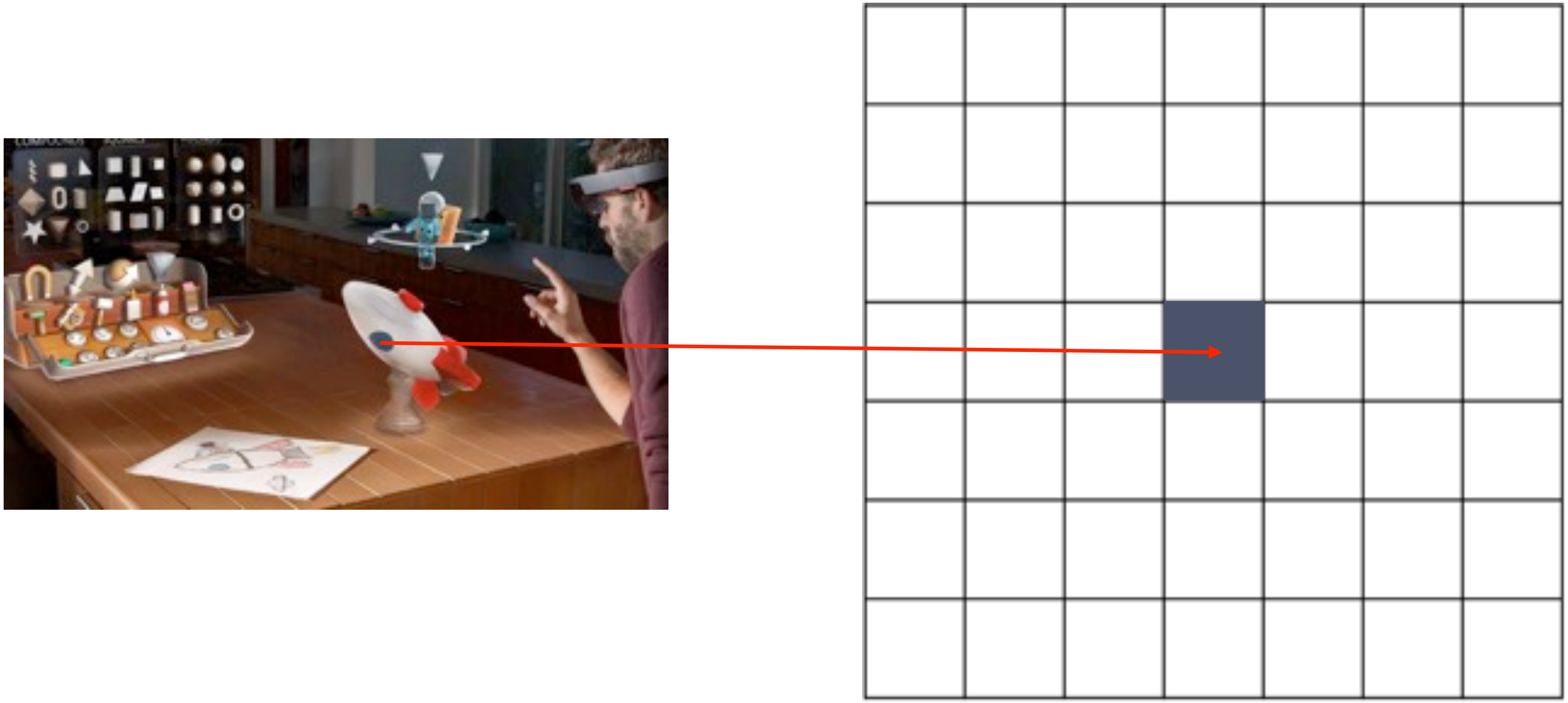
- Camera specs:
  - 640 × 480 pixels
  - FoV = 60°
  - Stereo pairs = 10
  - 120 fps



- Quantify errors in terms of display pixel errors
  - Display specs : HTC Vive
    - 120° FoV
    - 1080 × 1200 pixels per eye
    - Point 1m in front

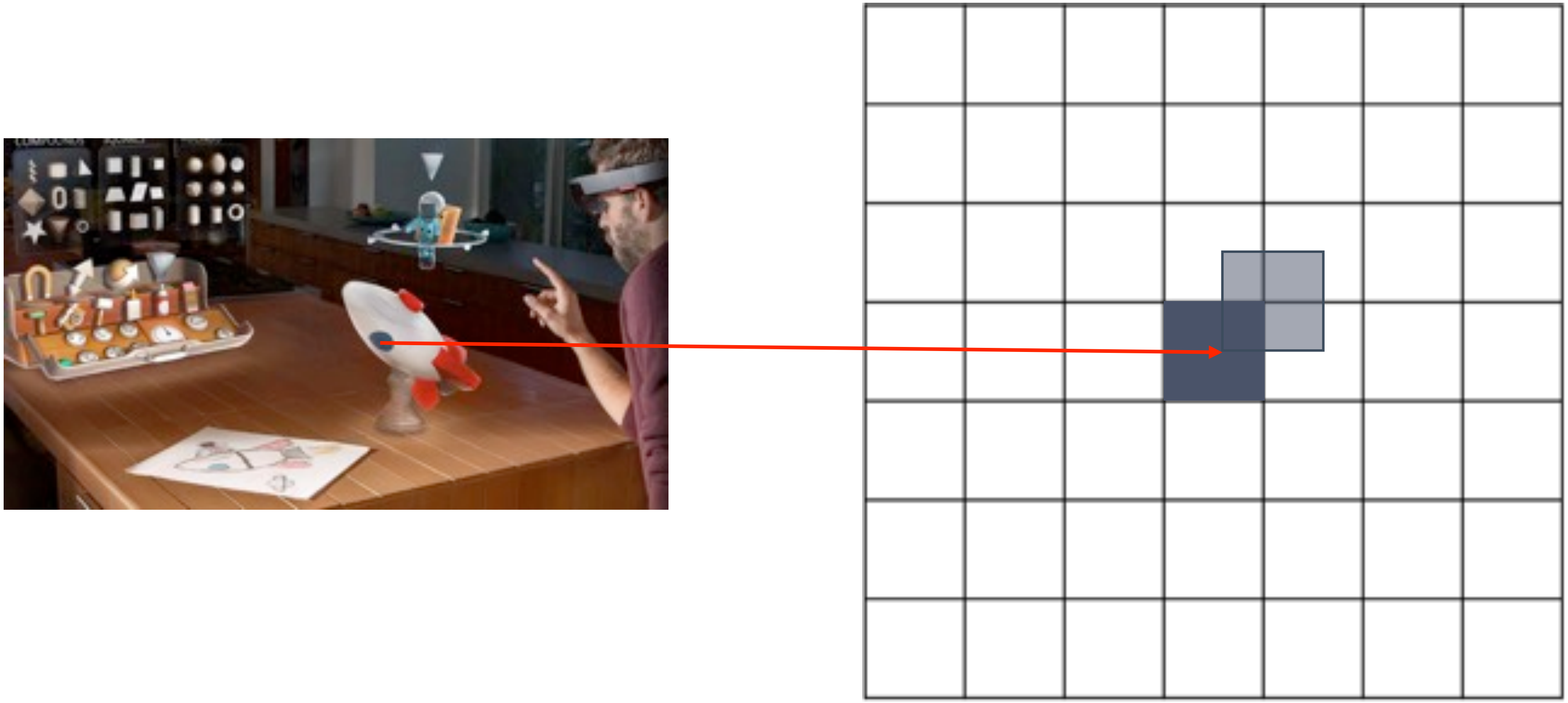
1.Mingsong Duo et al, Exploring high-level plane primitives for indoor 3d reconstruction with a hand-held rgb-d camera, ACCV 2012

Display Pixel Error



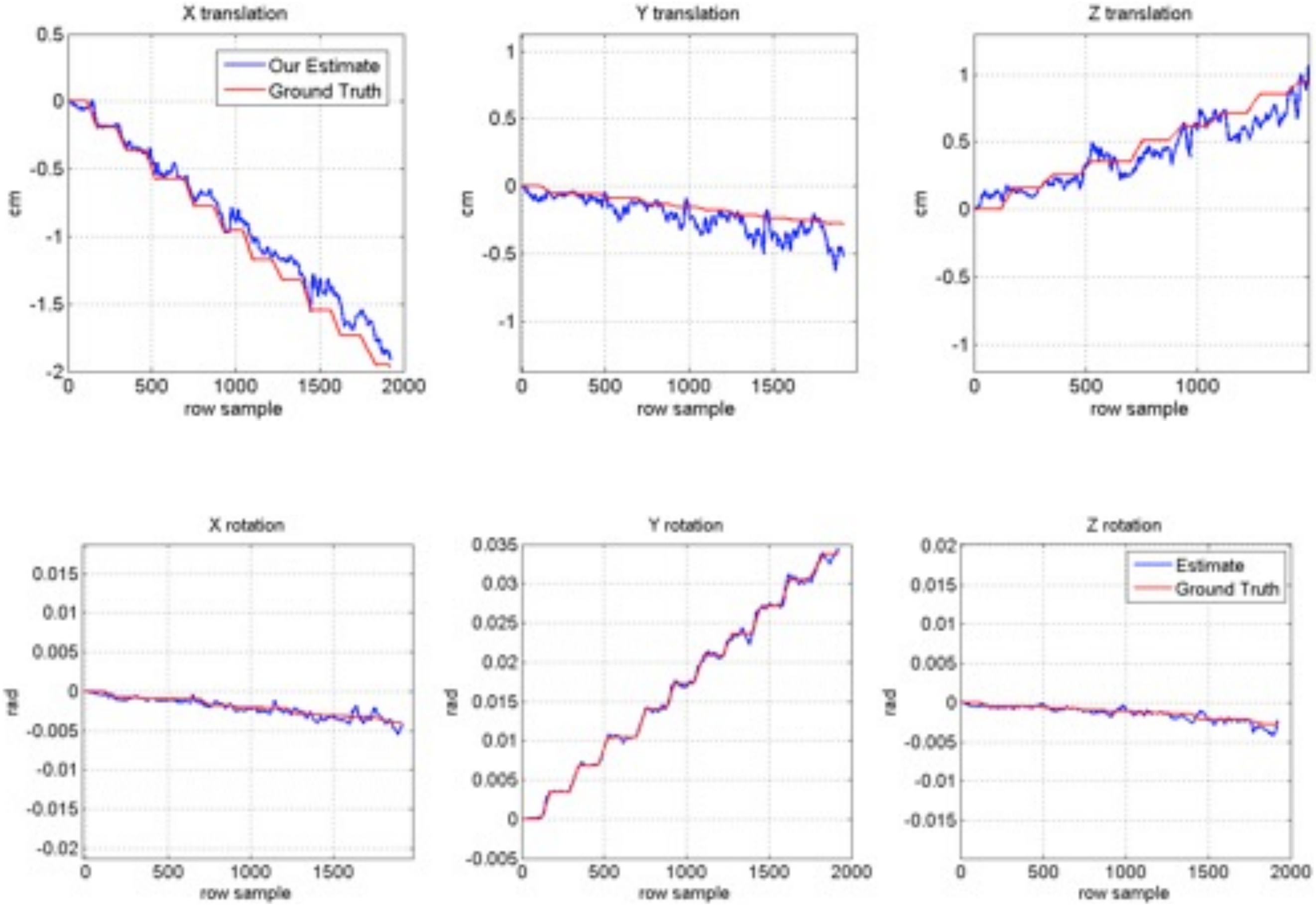
Pixels shown in HMD

Display Pixel Error



Pixels shown in HMD

Experiments : Real motion[1] in Synthetic Room

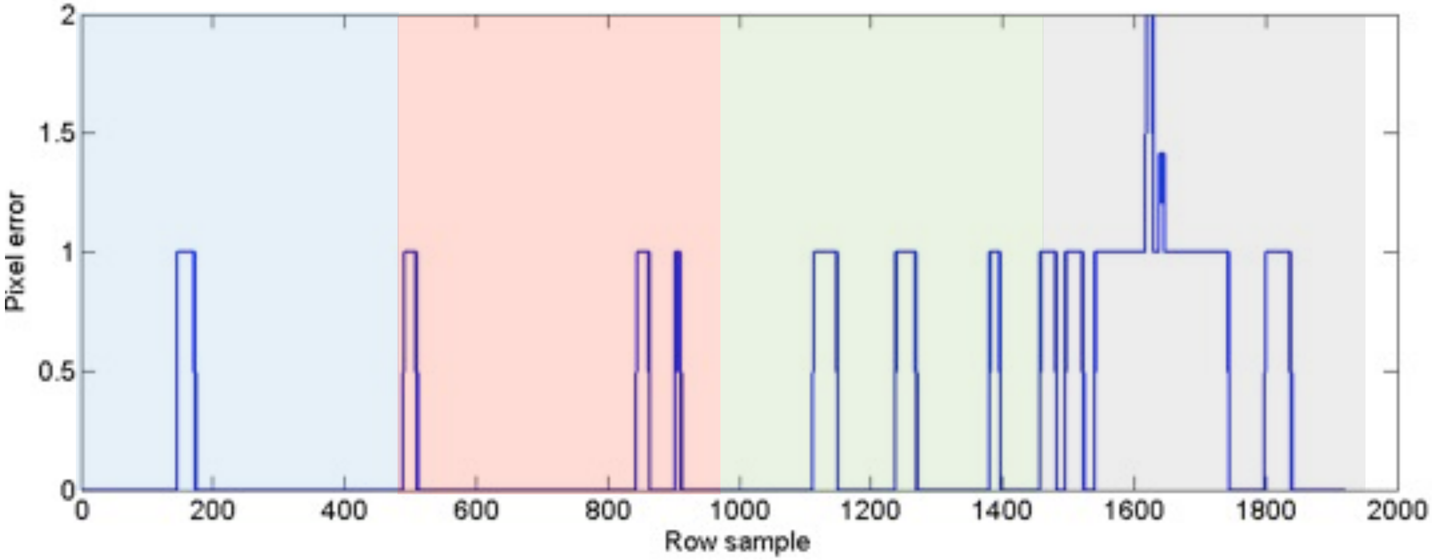


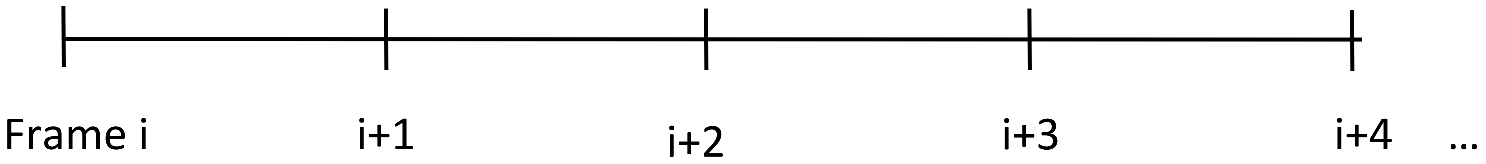
f= 56.7kHz

[1] Hi-Ball tracker data

Experiments : Real motion[1] in Synthetic Room

f= 56.7kHz




Frame i

i+1

i+2

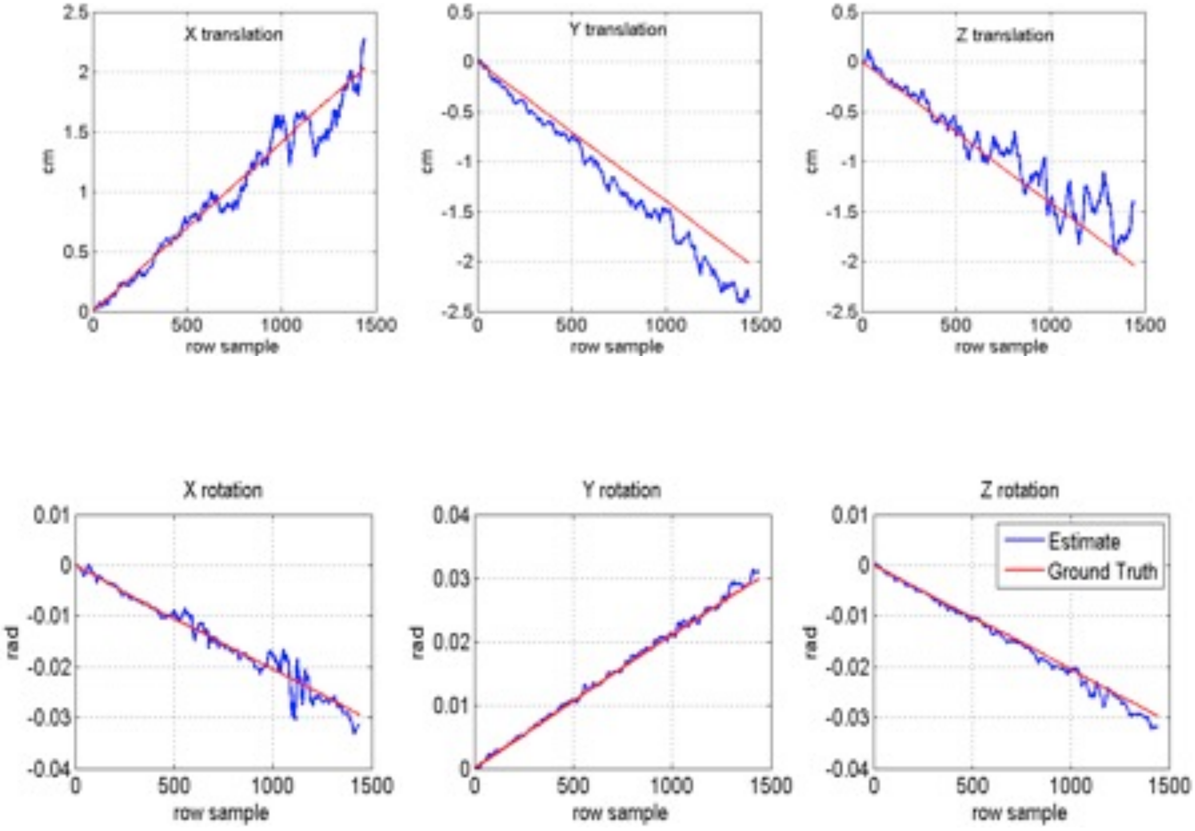
i+3

i+4

…

[1] Hi-Ball tracker data

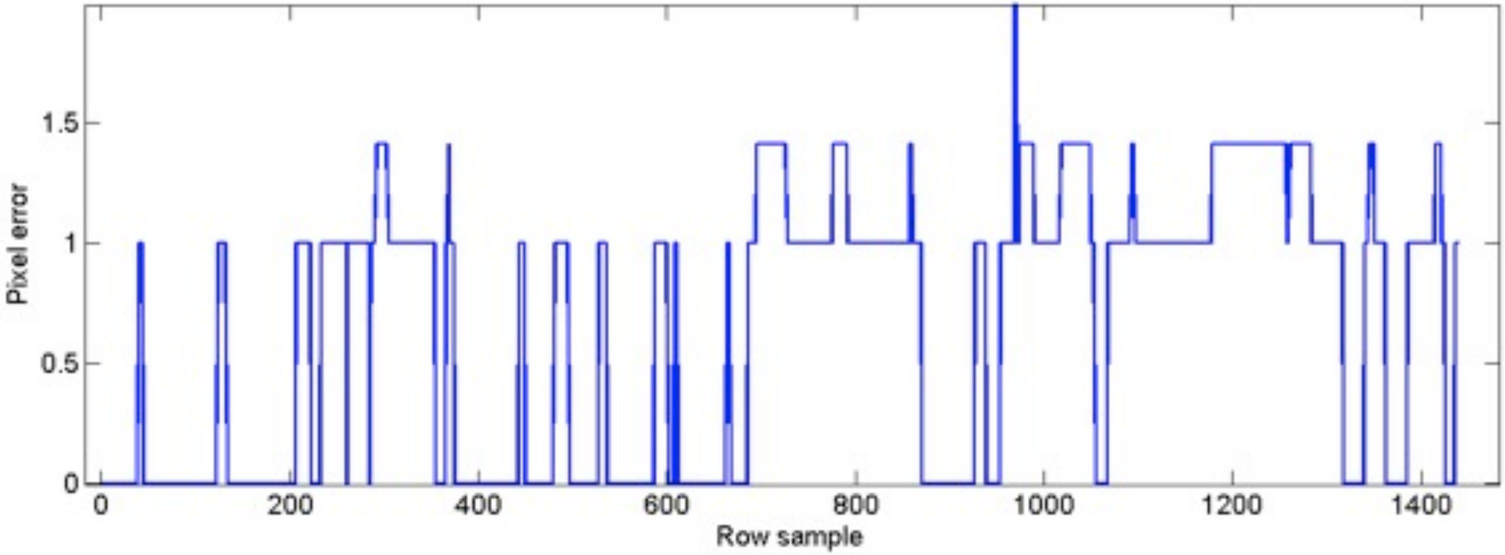
Experiments: Synthetic Large Motion



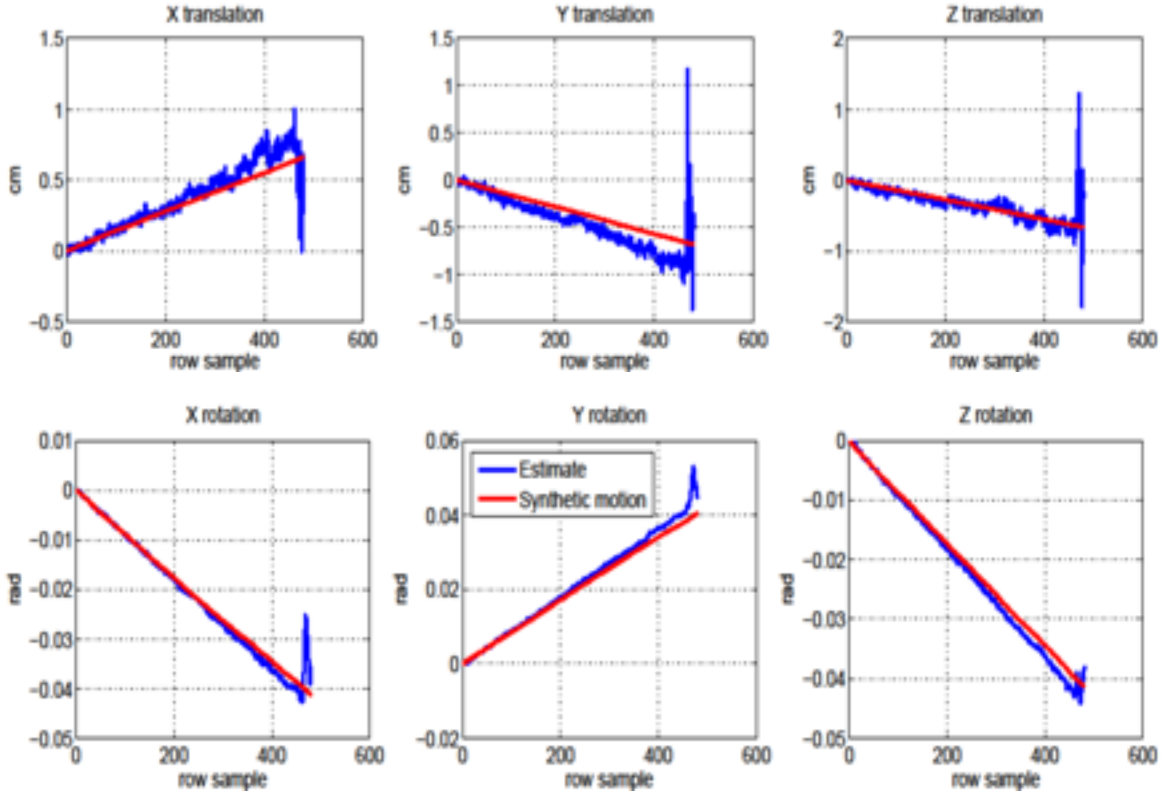
f= 56.7kHz
v= 1.4 m/s
ω= 120 deg/s

Experiments: Synthetic Large Motion

f= 56.7kHz



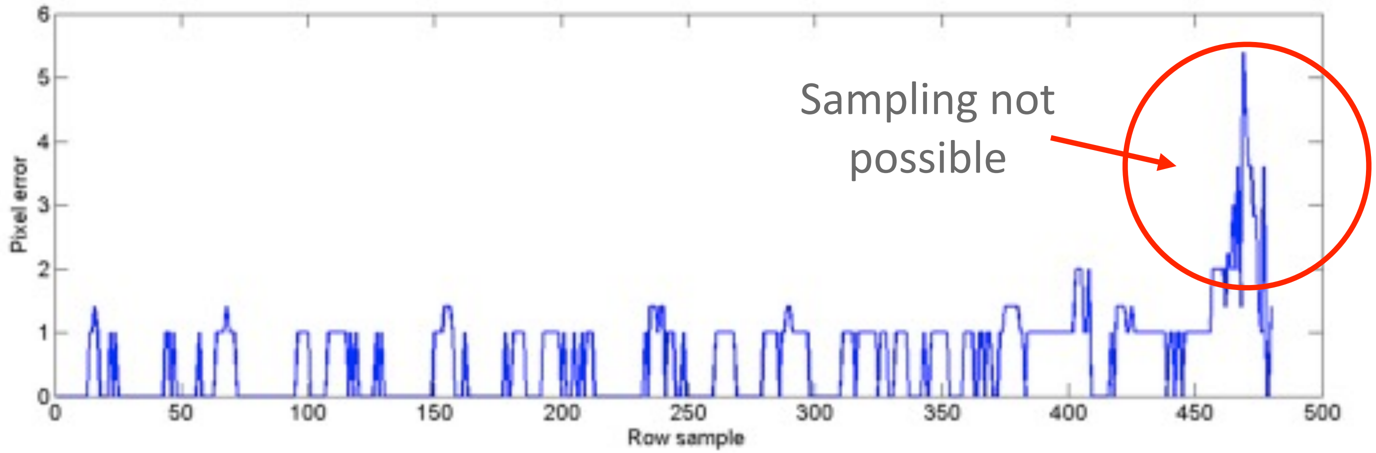
Experiments : Synthetic Extreme Motion

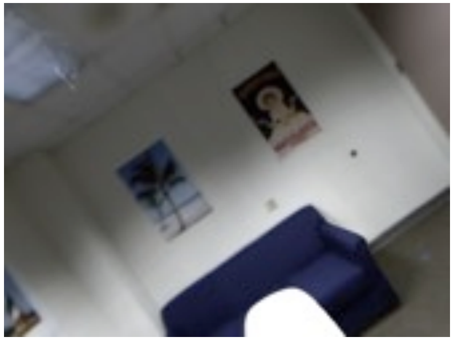


f= 56.7kHz
v= 1.4 m/s
ω= 500 deg/s

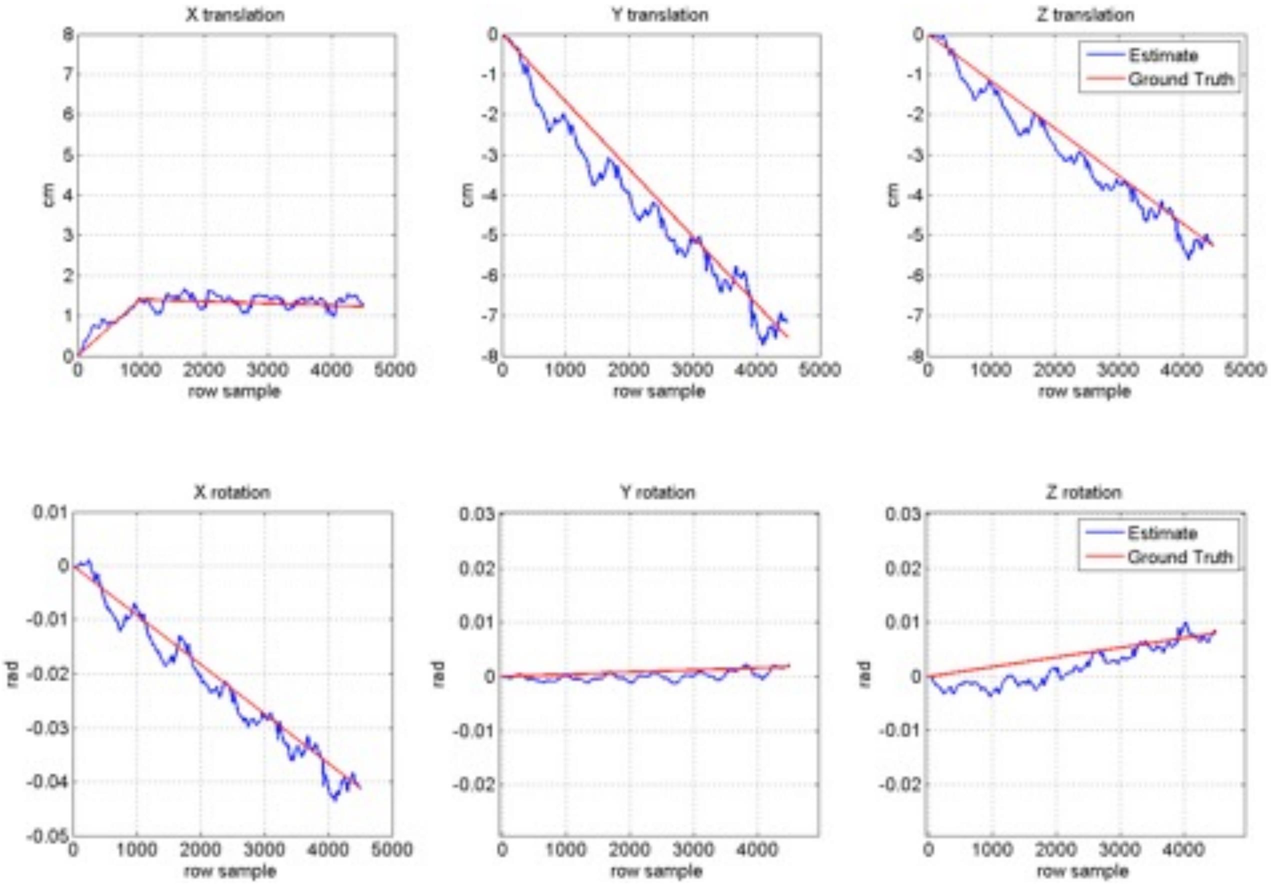
Experiments : Synthetic Extreme Motion

f= 56.7kHz
v= 1.4 m/s
ω= 500 deg/s





Results for Real Imagery



f= 80.4kHz

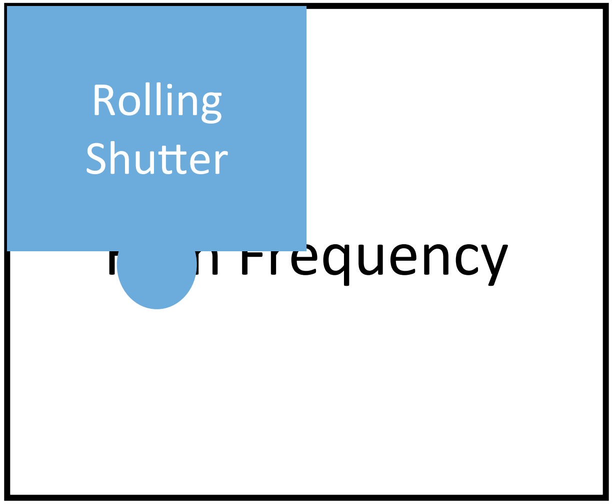
Conclusion

- High frequency visual tracker
  - Up to 80 kHz
  - Off-the-shelf cameras

High Frequency

Conclusion

- High frequency visual tracker
  - Up to 80 kHz
  - Off-the-shelf cameras

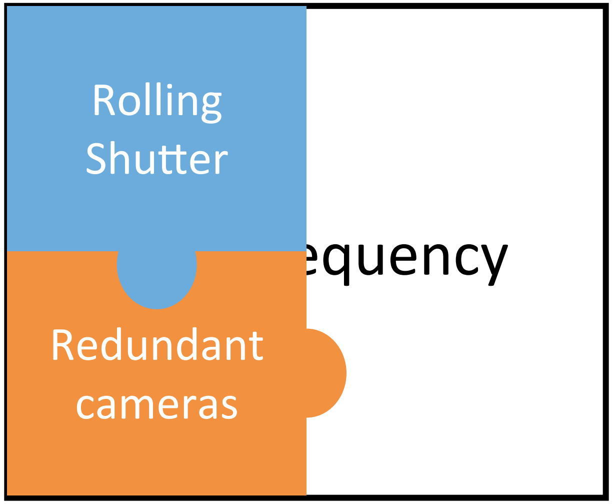


Rolling Shutter

Frequency

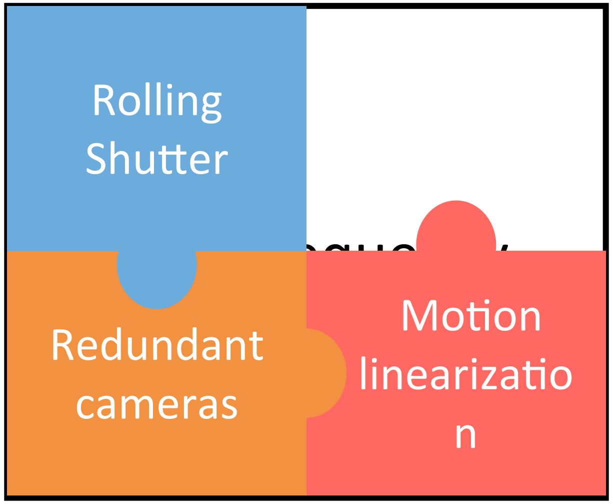
Conclusion

- High frequency visual tracker
  - Up to 80 kHz
  - Off-the-shelf cameras



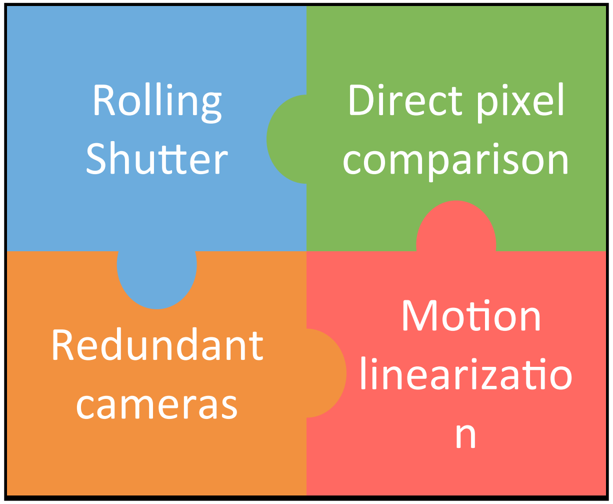
Conclusion

- High frequency visual tracker
  - Up to 80 kHz
  - Off-the-shelf cameras



Conclusion

- High frequency visual tracker
  - Up to 80 kHz
  - Off-the-shelf cameras



Rolling Shutter

Direct pixel comparison

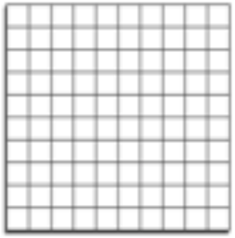
Redundant cameras

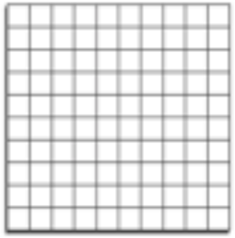
Motion linearizatio n

Thank you!

Contact: akash@cs.unc.edu

Camera tracking : The spectrum





1-D camera[1]

Global Shutter	Rolling Shutter as 1-D sensor	1-D sensor
+ No distortion	– Distortion	+ No distortion
+ Drift correction	+ Drift correction	– No drift correction
– High cost	+ Cheap	+ High Cost
– Low fps ≈ 200Hz	+ High fps, 56 kHz	+ High fps, till 87 kHz
– Higher noise	+ Low noise	– Higher noise