Video Stabilization and Rectification for Handheld Cameras

Chao Jia

Ph.D. Defense

Committee Members:
Prof. Alan Bovik
Prof. Constantine Caramanis
Prof. Brian L. Evans (supervisor)
Prof. Donald Fussell
Prof. Todd Humphreys

2014-04-04

Outline

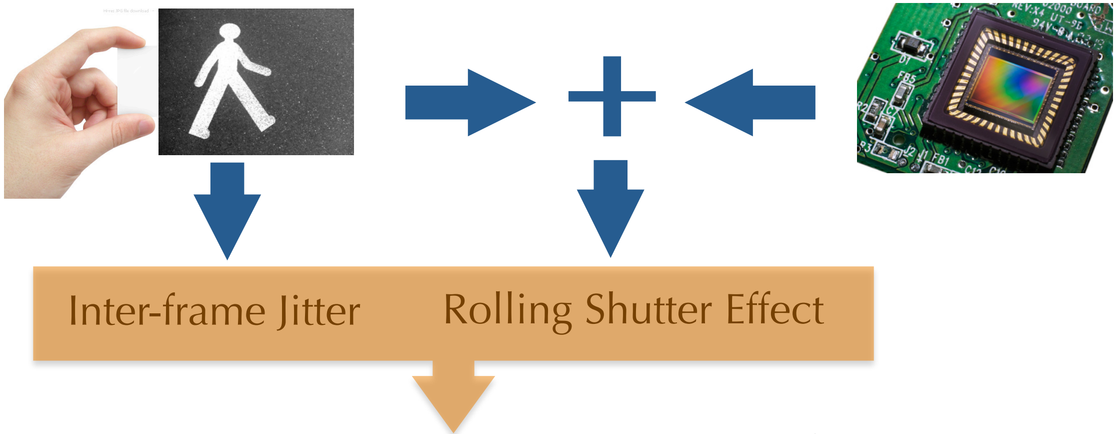
- Introduction to Video Stabilization and Rectification
- Camera-Gyroscope Calibration for Motion Estimation
- Offline 3D Rotation Smoothing
- Online (Real-Time) 3D Rotation Smoothing
- Conclusion

Introduction

Video Recording by Handheld Cameras (Smart phones, Tablets…)

Compactness
Everywhere & Anytime
Easy Sharing
Better User Experience (touchscreen)
⋮

Poor Video Quality Sometimes …



Camera Motion

CMOS Sensor

Inter-frame Jitter

Rolling Shutter Effect

Handheld Camera Video Stabilization & Rectification

CMOS Sensor & Rolling Shutter Effect

- CMOS sensors in almost every smart phone camera
  - Lower power consumption
  - Faster data throughput

Each row is captured under ***a different camera pose*** *if the camera is moving.*

Rolling Shutter Effect Under Fast Camera Panning

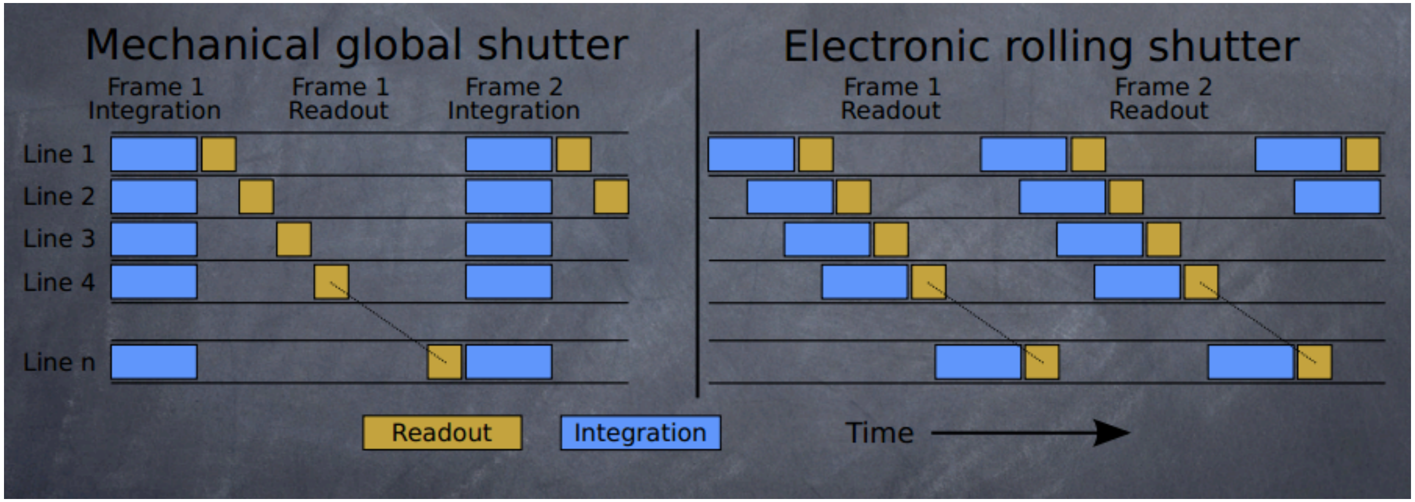
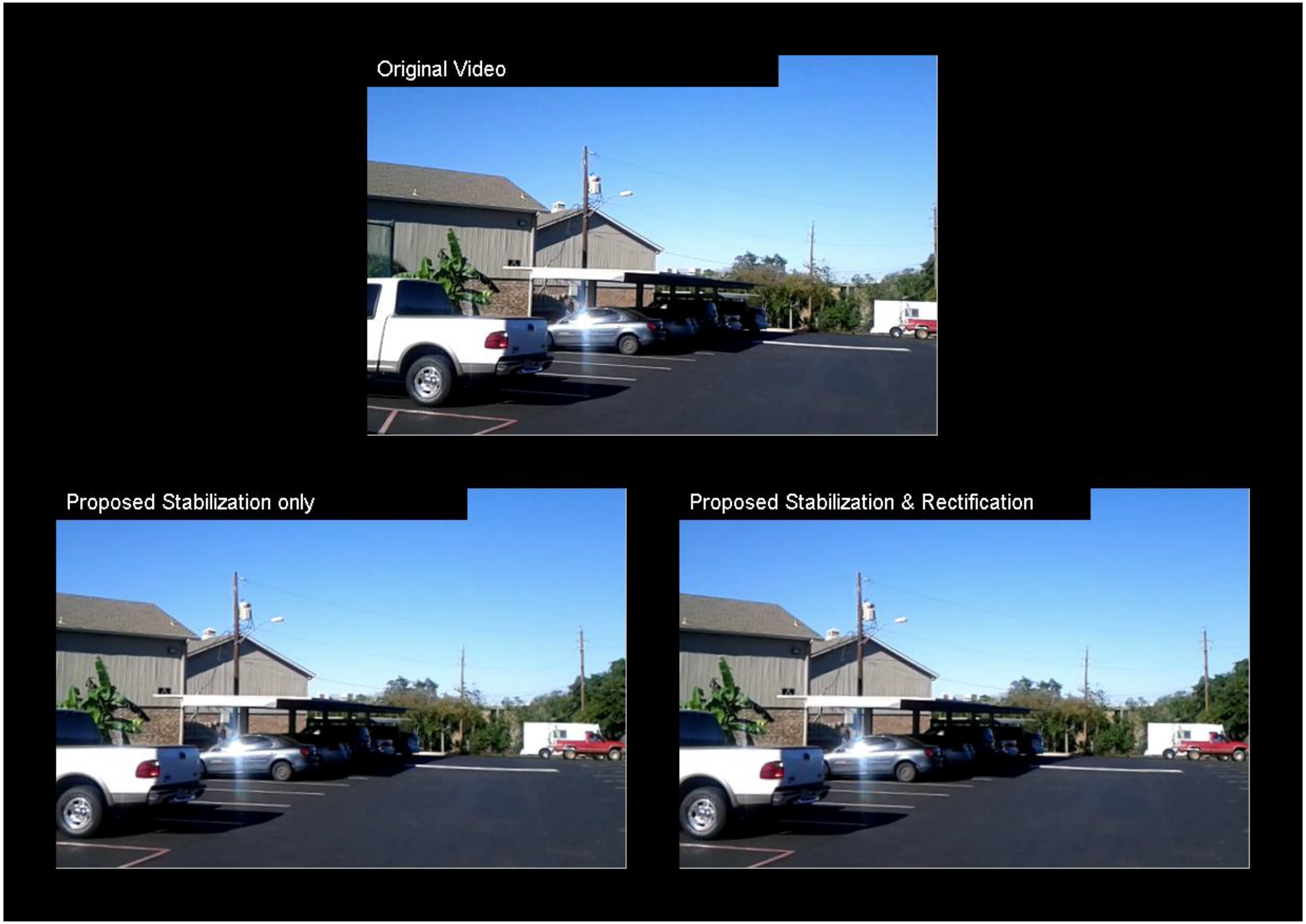


Image from P.E. Forssén, E. Ringaby, J. Hedborg CVPR 2012 Tutorial

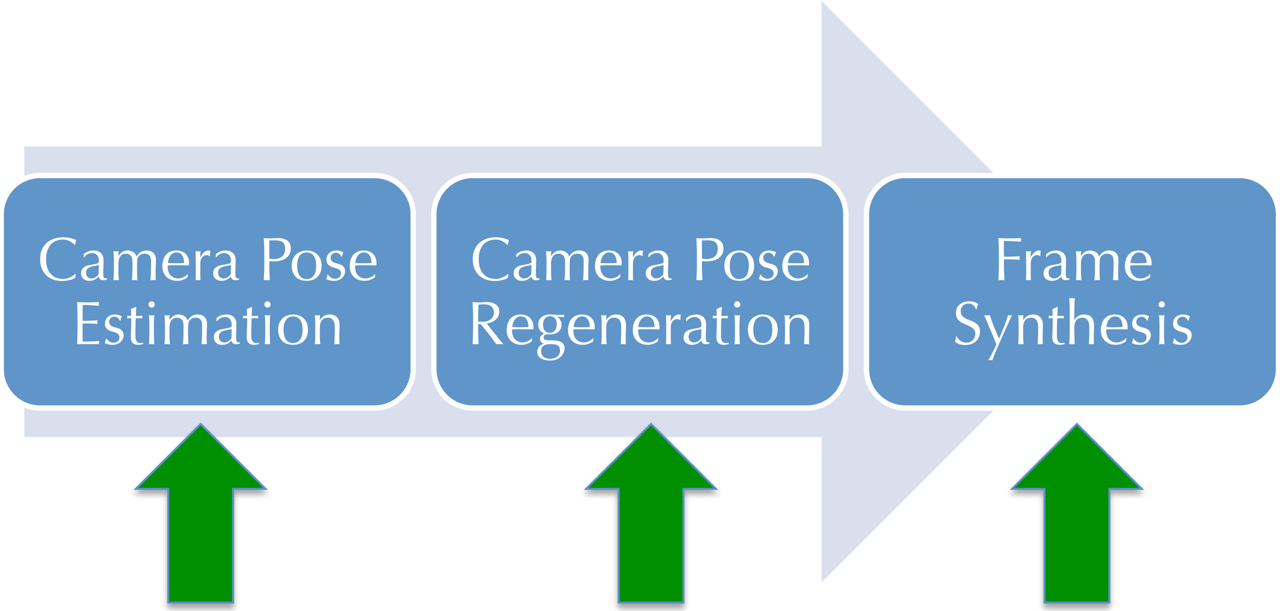
Rolling Shutter Rectification is Necessary



High frequency camera shake will cause non-rigid rolling shutter distortion

Video Stabilization & Rectification

- **Stabilization**: Removing unwanted jitter (**inter**-frame correction)
- **Rectification**: Removing rolling shutter effect (**intra**-frame correction)



Camera Pose Estimation

Camera Pose Regeneration

Frame Synthesis

Camera pose for each row

Fix a unique pose/frame
Motion smoothing

Apply the correction

Motion Model Selection

- 2D motion: apparent pixel displacements

Translation → Similarity → Euclidean → Affine → Projective

Degrees of Freedom

- 3D real camera motion

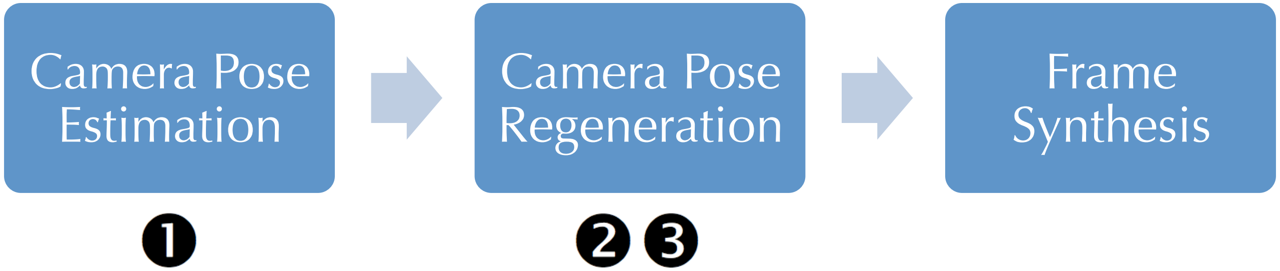
Rotation Full (Rotation + Translation)

Motion Model	Estimation Complexity	Smoothing Effectiveness	Correction Complexity
2D	high	low	low
3D Full	high	high	high
3D Rotation	low (using gyro)	high	low (projective transform)

- No approximation in 3D rotational stabilization (proposed method)
  - We are not assuming pure camera rotation
  - Translation is kept as is, and not smoothed

Thesis Contributions

❶ Online Camera-Gyroscope Calibration & Synchronization
❷ Offline 3D Rotation Smoothing
❸ Online (real-time) 3D Rotation Smoothing



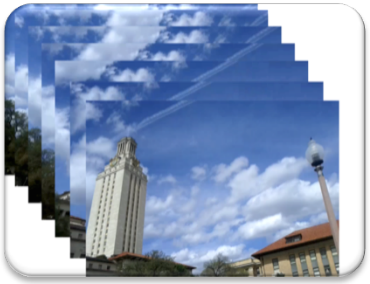
Thesis Statement

For handheld cameras with CMOS sensors, videos can be rectified and then stabilized either online or offline, with the camera motion estimated directly from gyroscope readings after effective sensor calibration.

Outline

- Introduction to Video Stabilization and Rectification
- Camera-Gyroscope Calibration for Motion Estimation
- Offline 3D Rotation Smoothing
- Online (Real-Time) 3D Rotation Smoothing
- Conclusion

Camera Motion (Rotation) Estimation



Vision-based (feature points/pixel intensities)

- Not robust to large moving objects, motion blur, etc.
- Highly complicated for rolling shutter camera

Gyro-based [Karpenko report 2011]

- High sampling frequency (suitable for rolling shutter)
- Independent of video quality
- Sensor calibration & synchronization needed

Parameters to Estimate

Intrinsic parameters f, c_x, c_y

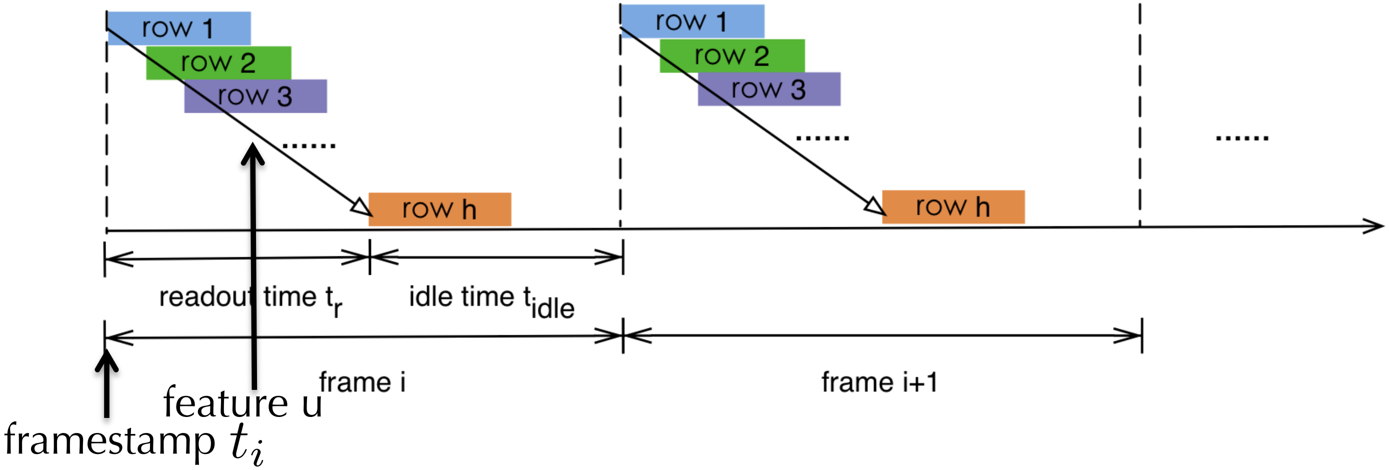
RS readout time t_r

bias $\mathbf{b}_g$

relative orientation $\mathbf{q}_c$

delay (use gyro as reference) t_d

- Rows are captured sequentially in rolling shutter cameras



$$t(\mathbf{u}, i) = t_i + t_d + t_r \frac{u_y}{h}$$

Gyro-camera Calibration

- Previous work on camera-gyro calibration
  - Simultaneous Localization and Mapping (SLAM)
    - Calibrate both gyro & accelerometer
  - Projective matching
    - Assume pure rotation
    - Calibrate only gyro

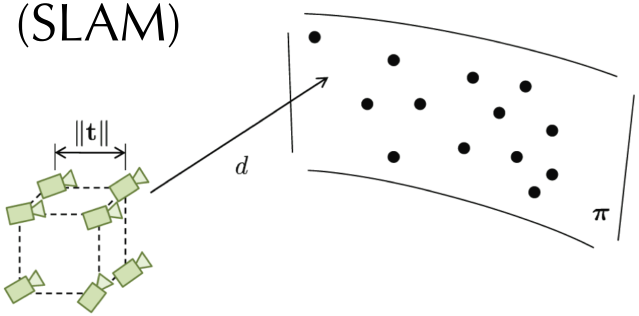


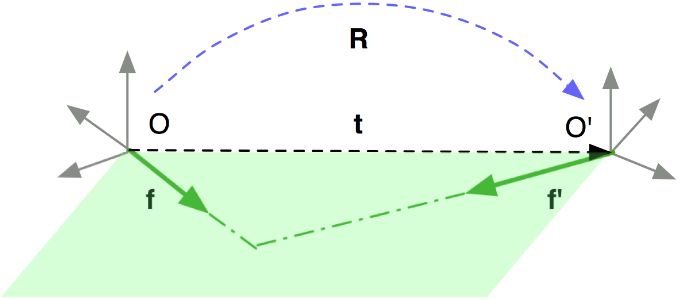
Image from [Hwangbo IJRR 2011]

Method	Computational Complexity	General Motion?	Online?
SLAM-based	high	✓	✓
Projective matching	low	✗	✗

- What makes good gyro-camera calibration for video stabilization
  - Online
  - Simple (no need to estimate translation or scene structure)
  - General motion (non-zero translation)

Coplanarity Constraint

- How can we get rid of translation?



Epipolar Constraint $\det[\mathbf{t}|\mathbf{f}|\mathbf{R}\mathbf{f}'] = 0$

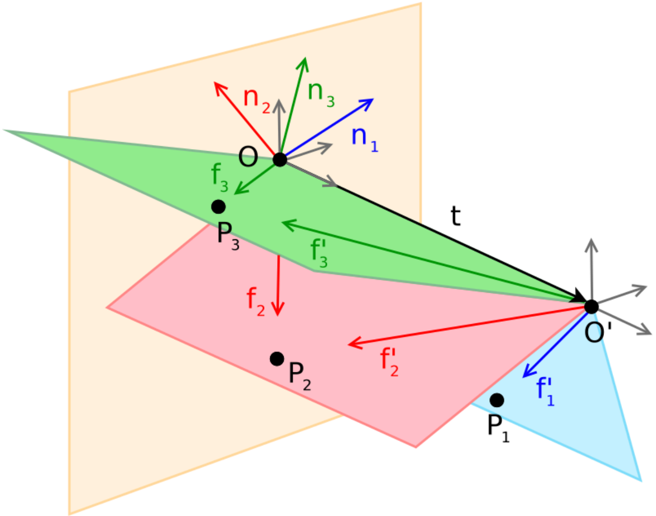


Image from [Kneip ECCV 2012]

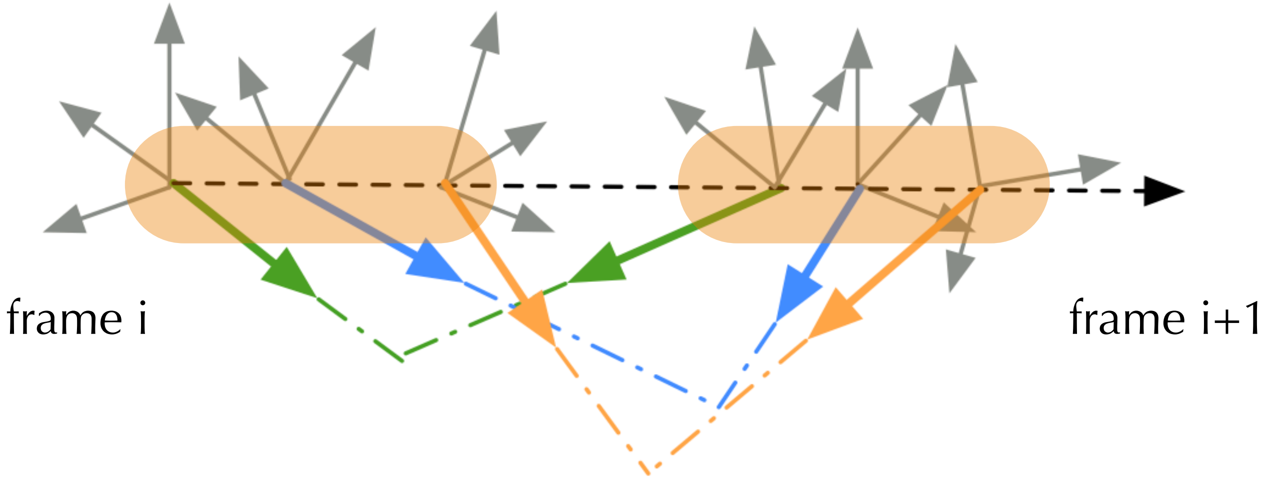
$$\det[(\mathbf{f}_1 \times \mathbf{R}\mathbf{f}'_1)|(\mathbf{f}_2 \times \mathbf{R}\mathbf{f}'_2)|(\mathbf{f}_3 \times \mathbf{R}\mathbf{f}'_3)] = 0$$

Coplanarity constraint of global shutter camera

Constraint for rotation only!

Rolling Shutter Coplanarity Constraint

Assume translation direction is constant



Pick any reference time

$$\det[(\mathbf{R}_1\mathbf{f}_1 \times \mathbf{R}_1'\mathbf{f}_1')|(\mathbf{R}_2\mathbf{f}_2 \times \mathbf{R}_2'\mathbf{f}_2')|(\mathbf{R}_3\mathbf{f}_3 \times \mathbf{R}_3'\mathbf{f}_3')] = 0$$

Coplanarity constraint of rolling shutter camera

EKF-based Online Calibration & Synchronization

- State Prediction
  - Copy state variables except for gyro bias (random walk process).

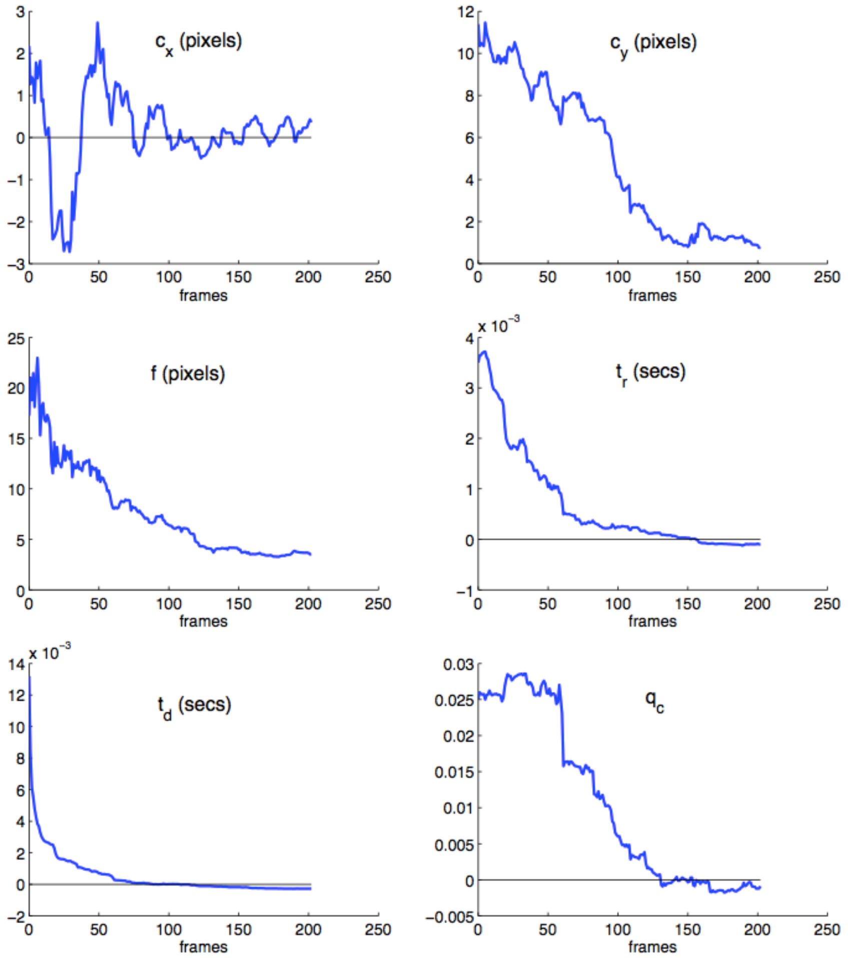
- State Update
  - Use coplanarity constraints as “implicit” measurements

    All variables in the state can be related to the constraint

- Run time: 7fps on 2.3GHz PC
  - Matlab implementation
  - Not necessary to run on every pair of adjacent frames

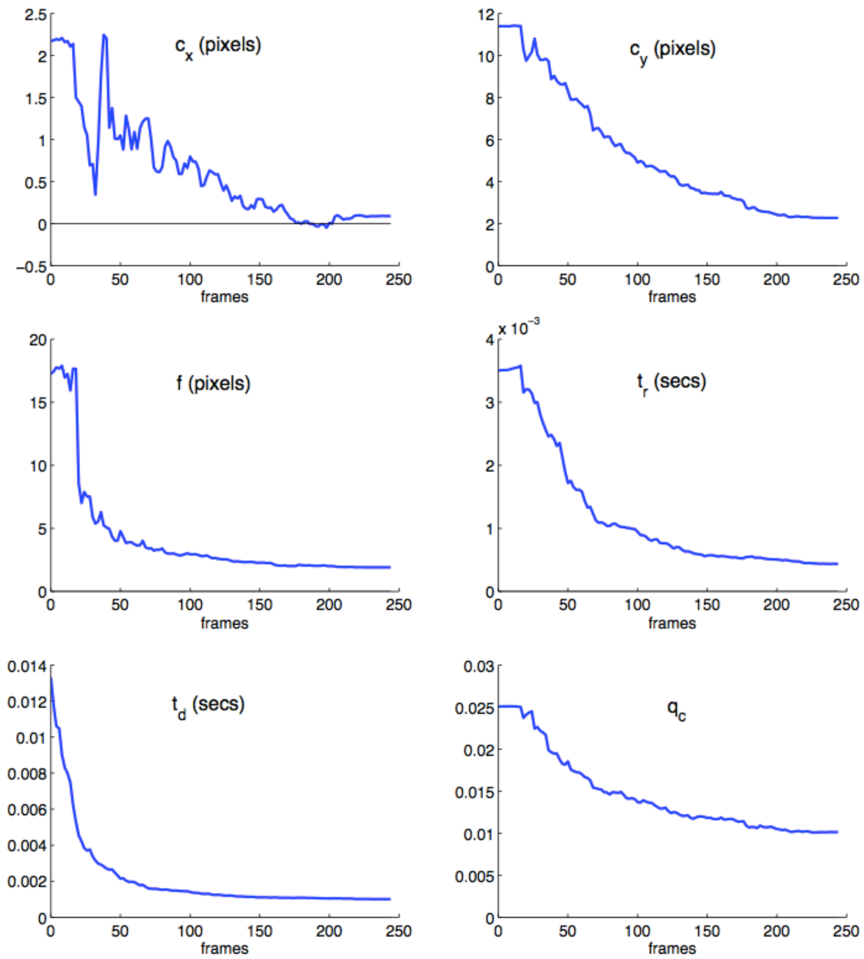
Running sequence (random rotation & translation)

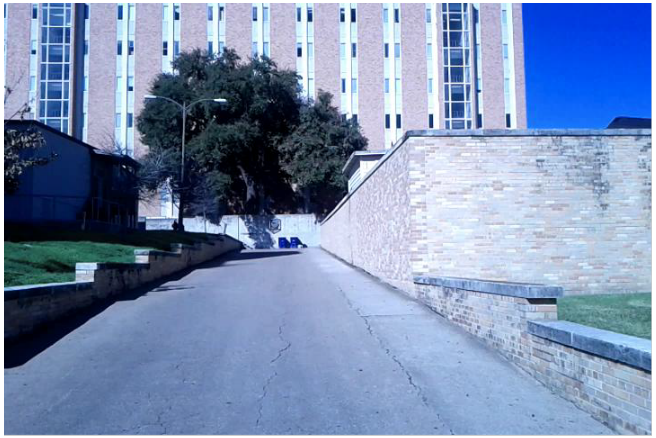
- Estimate Error



Panning sequence (almost zero translation)

- Estimate Error



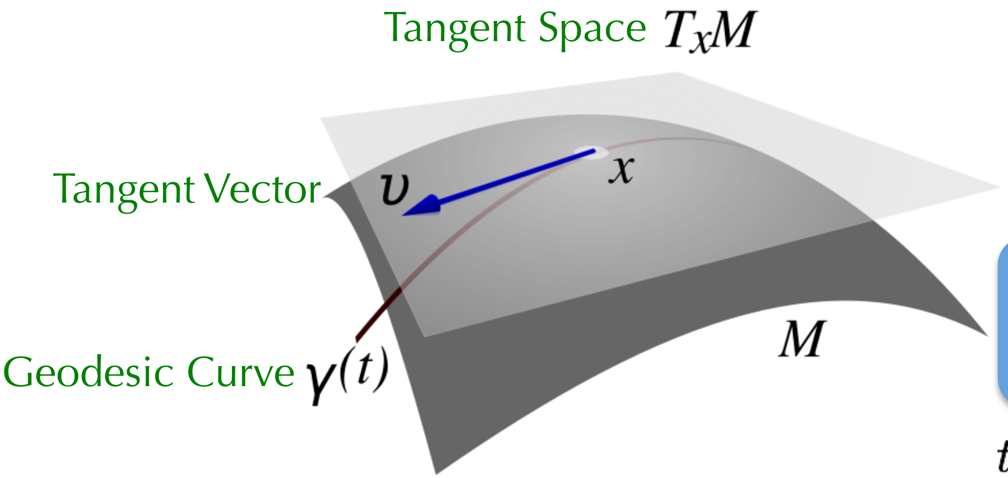


Outline

- Introduction to Video Stabilization and Rectification
- Camera-Gyroscope Calibration for Motion Estimation
- Offline 3D Rotation Smoothing
- Online (Real-Time) 3D Rotation Smoothing
- Conclusion

3D Rotation Matrix sequence

- Manifold of 3D Rotation Matrices
  - Special Orthogonal Group $\mathbf{SO(3)}$
  - An embedded submanifold in $\mathbb{R}^9$ (dimension = 3)



Generalization of Straight Line in Euclidean Space

$$t \in [0,1] \mapsto \gamma(t) = \exp_x(t \cdot v)$$

- Minimizing Geodesic & Geodesic Distance

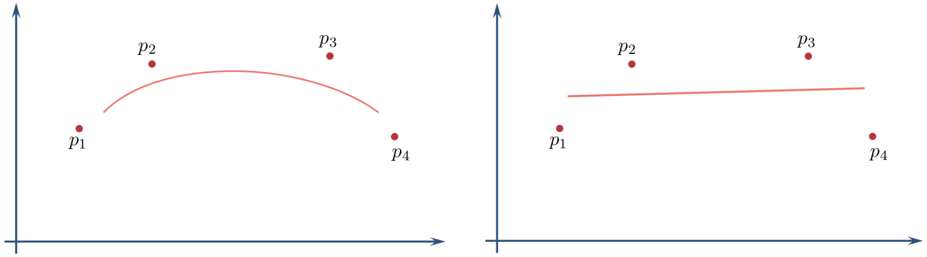
$$d_g(\mathbf{R}, \mathbf{R}') = ||\text{logm}(\mathbf{R}^{-1}\mathbf{R}')||_F$$

3D Rotation Sequence Smoothing

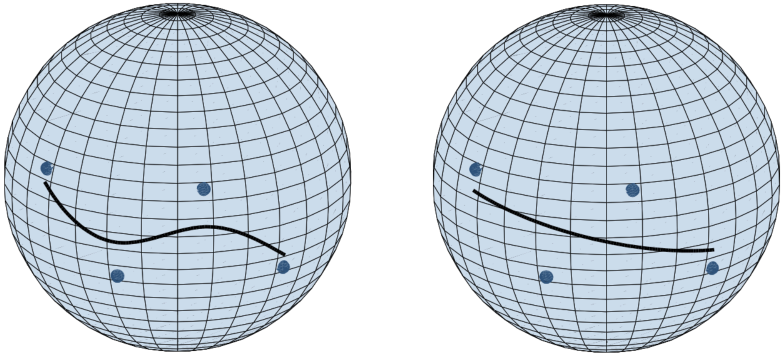
- Offline Motion Smoothing

$$\{\tilde{\mathbf{R}}_n\} \rightarrow \{\mathbf{R}_n\}$$

- Motion Smoothing as Discrete Curve Fitting



Regression on $\mathbb{R}^2$



Regression on $\mathbb{S}^2$

A Balance between Fitting and Smoothness

Motion Smoothing as Regression

- First-Order Difference Penalty

$$\min_{\{\mathbf{R}_n\}} \sum_{n=1}^{N} d_g^2(\mathbf{R}_n, \tilde{\mathbf{R}}_n) + \alpha \sum_{n=1}^{N-1} d_g^2(\mathbf{R}_n, \mathbf{R}_{n+1})$$

Regression using 2D Motion Model

[Grundmann CVPR 2011]

- Convex optimization in Euclidean space
- 2D Motion

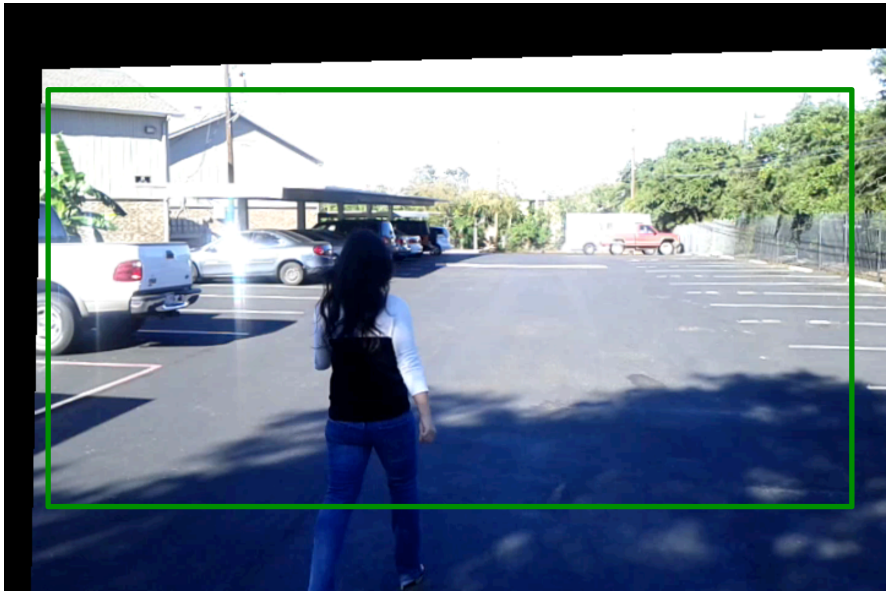
Local Low-pass Filtering on Manifold

[Hanning IWMV 2011]

- 3D Rotation
- Not globally optimal
- Cannot include additional constraints (unless ad-hoc)

Constrained Motion Smoothing

- Inevitably some pixels are not visible after view change



$$\begin{bmatrix}\tilde{u}_{ij}\\ \tilde{v}_{ij}\end{bmatrix} = g\left(\mathbf{K}\tilde{\mathbf{R}}_i\mathbf{R}_i{}^{\mathrm{T}}\mathbf{K}^{-1}\begin{bmatrix}u_{ij}\\ v_{ij}\\ 1\end{bmatrix}\right)$$

Correction by image warping

$$\begin{cases}0 \leq \tilde{u}_{ij} \leq w\\ 0 \leq \tilde{v}_{ij} \leq h\end{cases}, \forall \begin{bmatrix}u_{ij}\\ v_{ij}\end{bmatrix} s.t. \begin{cases}c_1 \leq u_{ij} \leq c_2\\ d_1 \leq v_{ij} \leq d_2\end{cases}$$

Hard Constraint: All of the pixels in the cropped new frame should be visible in the original frame.

Constrained Manifold Regression

- Constraint approximation: a geodesic convex set on the manifold

$$d_g(\mathbf{R}_i, \tilde{\mathbf{R}}_i) \leq r_0, \forall i$$

Geodesic balls centered at the original rotation

- Constrained Motion Smoothing

$$\min_{\{\mathbf{R}_i\}} \sum_{i=1}^{N} \frac{1}{2} d_g^2(\tilde{\mathbf{R}}_i, \mathbf{R}_i) + \alpha \sum_{i=1}^{N-1} \frac{1}{2} d_g^2(\mathbf{R}_i, \mathbf{R}_{i+1}), \text{ s.t. } \{\mathbf{R}_i\} \in \Omega$$

$$\Omega = \Omega_1 \times \Omega_2 \times \cdots \times \Omega_N$$

$$\Omega_i = \{\mathbf{R}_i \in \mathbf{SO(3)} : d_g(\mathbf{R}_i, \tilde{\mathbf{R}}_i) \leq r_0\}$$

- Variable: sequence of rotation matrices

Optimization

- Manifold Optimization

Optimization on Euclidean space [Bertsekas 1999]

- Orthogonality as additional constraints
- Non-convex objective + Non-convex set (Poor guarantee of convergence or optimality)

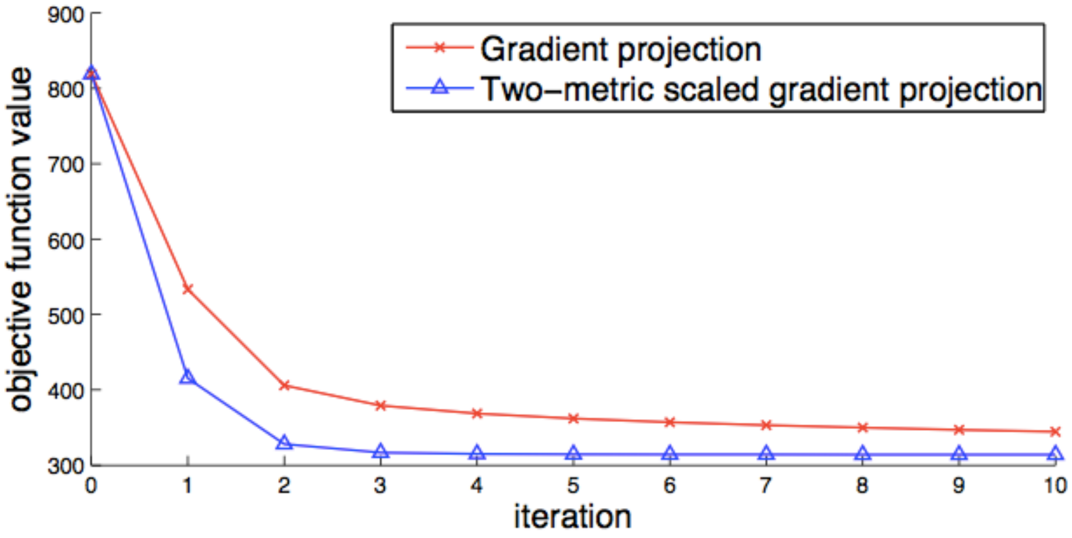
Optimization on manifold (Proposed)

- Geodesic-convex
- Gradient-related iterative algorithms (Similar convergence properties)
- *Gradient & Hessian computation needs Riemannian geometry*

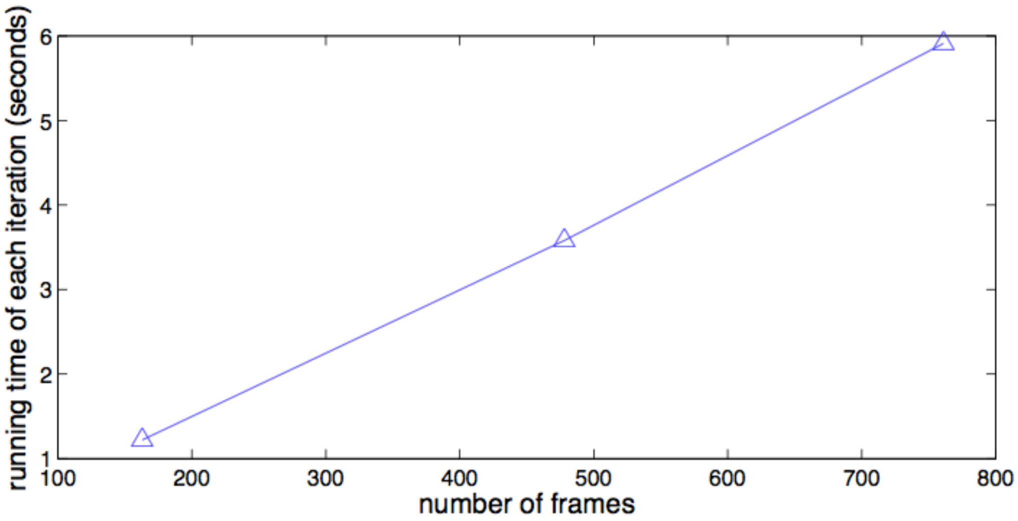
- From Unconstrained to Constrained

Unconstrained	**Constrained on a convex set**
steepest gradient descent	gradient projection
scaled gradient descent (Newton, Quasi-Newton, Congjugate Gradient)	two-metric projection

Convergence Analysis



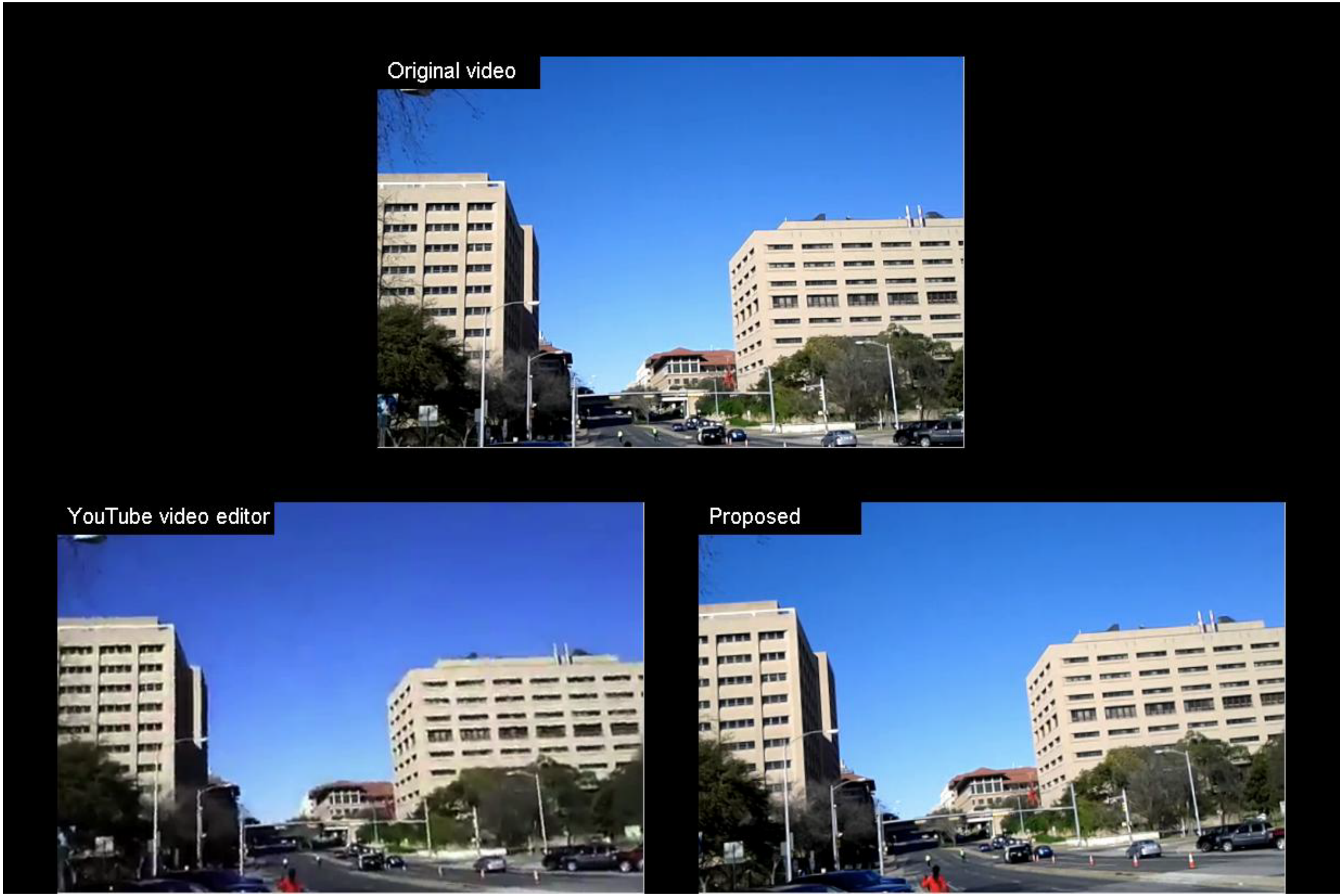
- Convergence comparison for a video with 478 frames



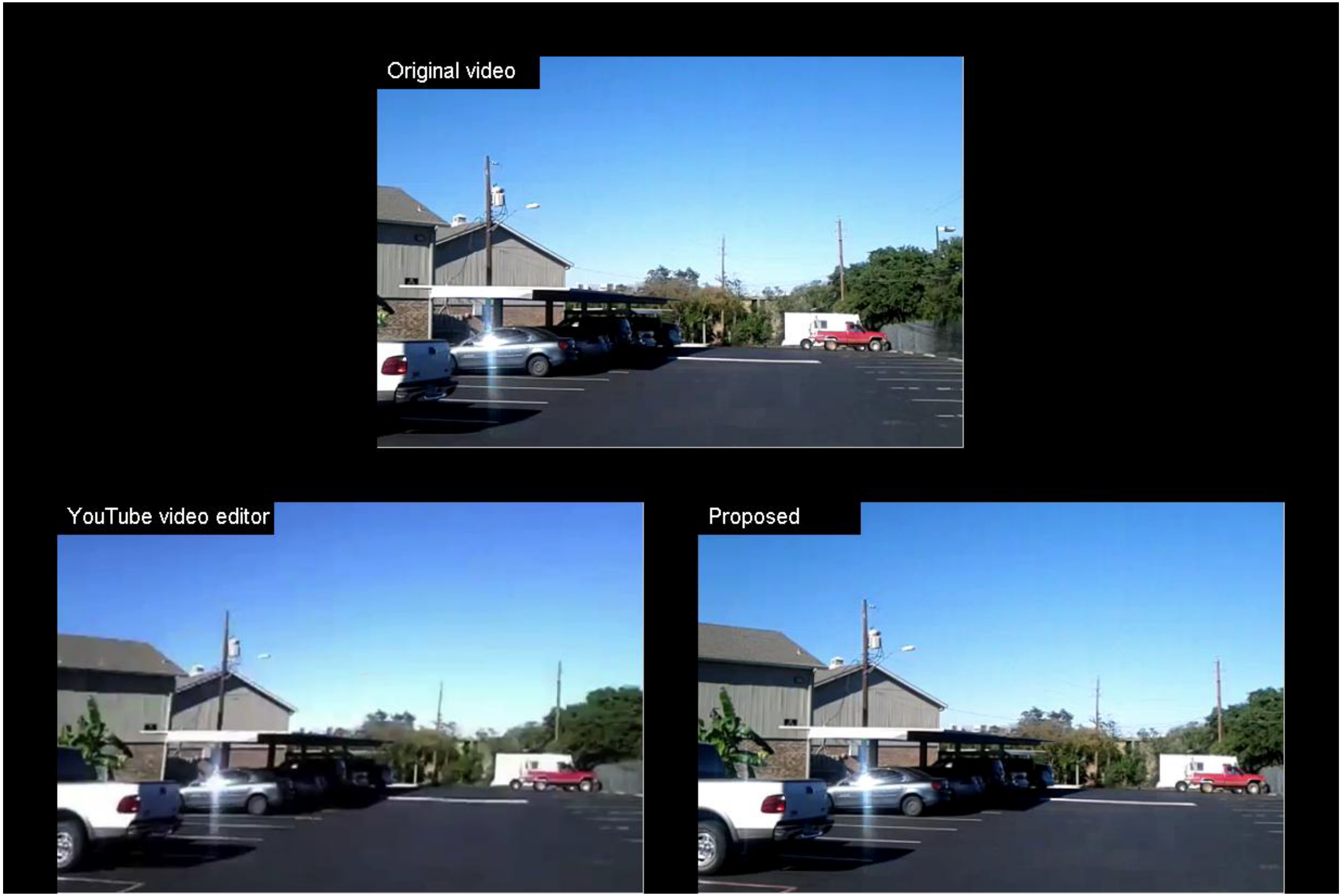
- Run time of each iteration of 2-metric scaled gradient projection
- Matlab implementation @2.3GHz
- ~7.5ms /iteration/frame

~39.0ms /iteration/frame for gradient projection

Experimental Results – vs. YouTube



Experimental Results – vs. YouTube



Outline

- Introduction to Video Stabilization and Rectification
- Camera-Gyroscope Calibration for Motion Estimation
- Offline 3D Rotation Smoothing
- Online (Real-Time) 3D Rotation Smoothing
- Conclusion

Online (Real-Time) Motion Smoothing

- Why online video stabilization
  - Real-time delivery: video conferencing, broadcasting, etc.
  - Improved user experience: WYSIWYG
  - More efficient compression

- Classical approaches for 2D motion models
  - (1st order low-pass) IIR filtering
  - Kalman filtering with constant-velocity (CV) model

- Extension to 3D rotation smoothing
  - Euclidean space → SO(3) manifold
  - Ad-hoc projection for black-border constraint

IIR-like 3D Rotation Smoothing

- First-Order IIR filtering $\hat{\boldsymbol{\theta}}_k = \alpha\hat{\boldsymbol{\theta}}_{k-1} + (1-\alpha)\boldsymbol{\theta}_k$ SO(3) ✗

$$\hat{\boldsymbol{\theta}}_k = \underset{\boldsymbol{\theta}}{\operatorname{argmin}}\, \alpha||\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{k-1}||^2 + (1-\alpha)||\boldsymbol{\theta} - \boldsymbol{\theta}_k||^2$$

SO(3) ✓

Euclidean distance → Geodesic distance

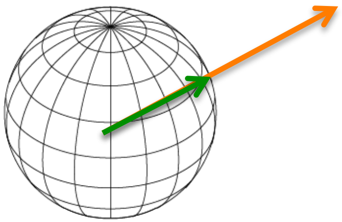
$$\hat{\mathbf{R}}_k = \underset{\mathbf{R}}{\operatorname{argmin}}\, \alpha d_g(\mathbf{R}, \hat{\mathbf{R}}_{k-1})^2 + (1-\alpha) d_g(\mathbf{R}, \mathbf{R}_k)^2$$

spherical linear interpolation (SLERP)

- Ad-hoc projection

$$\hat{\mathbf{R}} = \mathbb{P}(\hat{\mathbf{R}}^*) = \mathbf{R}\text{expm}(\beta^*\text{logm}(\mathbf{R}^{-1}\hat{\mathbf{R}}^*))$$

Move closer to the original rotation
if necessary for black-border constraint



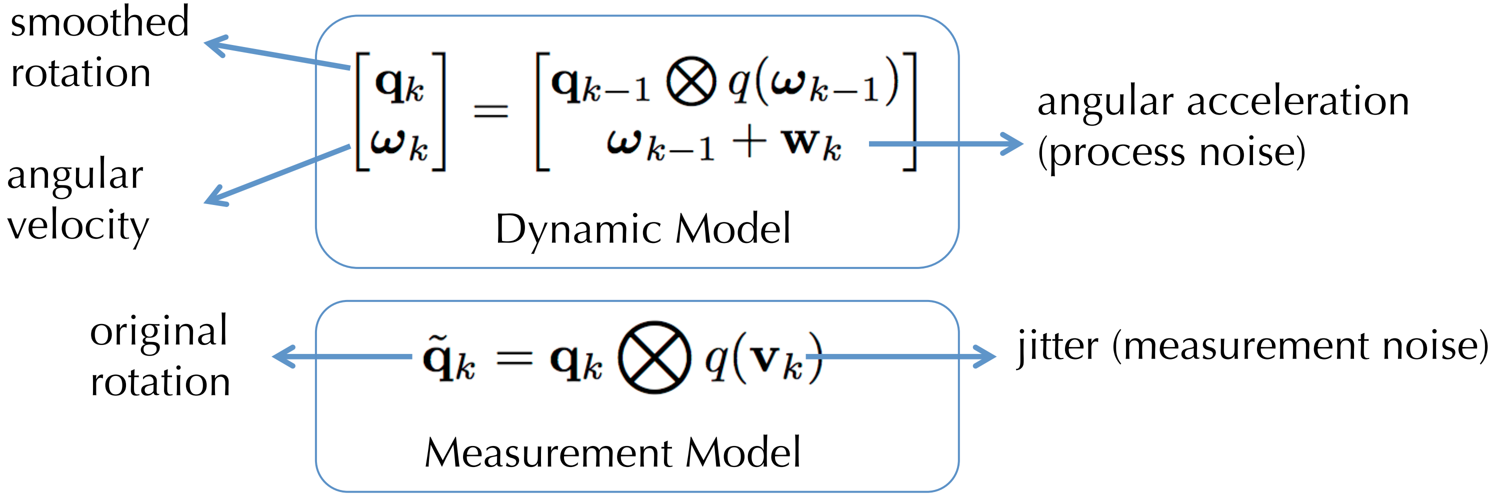
Algorithm IIR-like 3D Rotation Smoothing

```
Input: $\mathbf{q}_1, \cdots, \mathbf{q}_K$ (original rotations)
Output: $\hat{\mathbf{q}}_1, \cdots, \hat{\mathbf{q}}_K$ (smoothed rotations)
$\hat{\mathbf{q}}_1 = \mathbf{q}_1$
for $k = 2$ to $K$ do
    $\hat{\mathbf{q}}_k = \text{slerp}(\mathbf{q}_k, \hat{\mathbf{q}}_{k-1}, \alpha)$
    $\hat{\mathbf{q}}_k \leftarrow \mathbb{P}(\hat{\mathbf{q}}_k)$
end for
```

1.54ms/frame

UKF-based 3D Rotation Smoothing

- Constant-Velocity Model (widely used in target tracking)



smoothed rotation

angular velocity

$$\begin{bmatrix}\mathbf{q}_k\\ \boldsymbol{\omega}_k\end{bmatrix} = \begin{bmatrix}\mathbf{q}_{k-1}\bigotimes q(\boldsymbol{\omega}_{k-1})\\ \boldsymbol{\omega}_{k-1}+\mathbf{w}_k\end{bmatrix}$$

angular acceleration (process noise)

Dynamic Model

original rotation

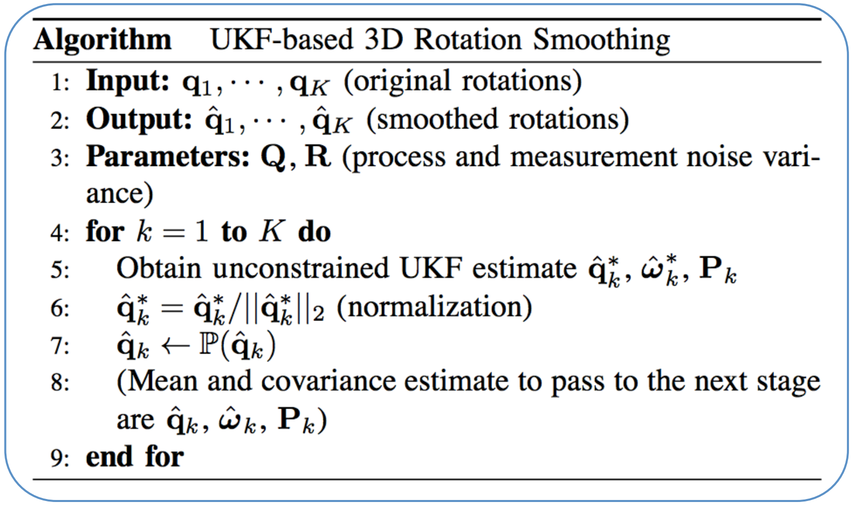
$$\tilde{\mathbf{q}}_k = \mathbf{q}_k \bigotimes q(\mathbf{v}_k)$$

jitter (measurement noise)

Measurement Model

- Hard to solve on SO(3)
- Nonlinear on Euclidean space
- Solved approximately by unscented Kalman filter (UKF)

6.97ms/frame

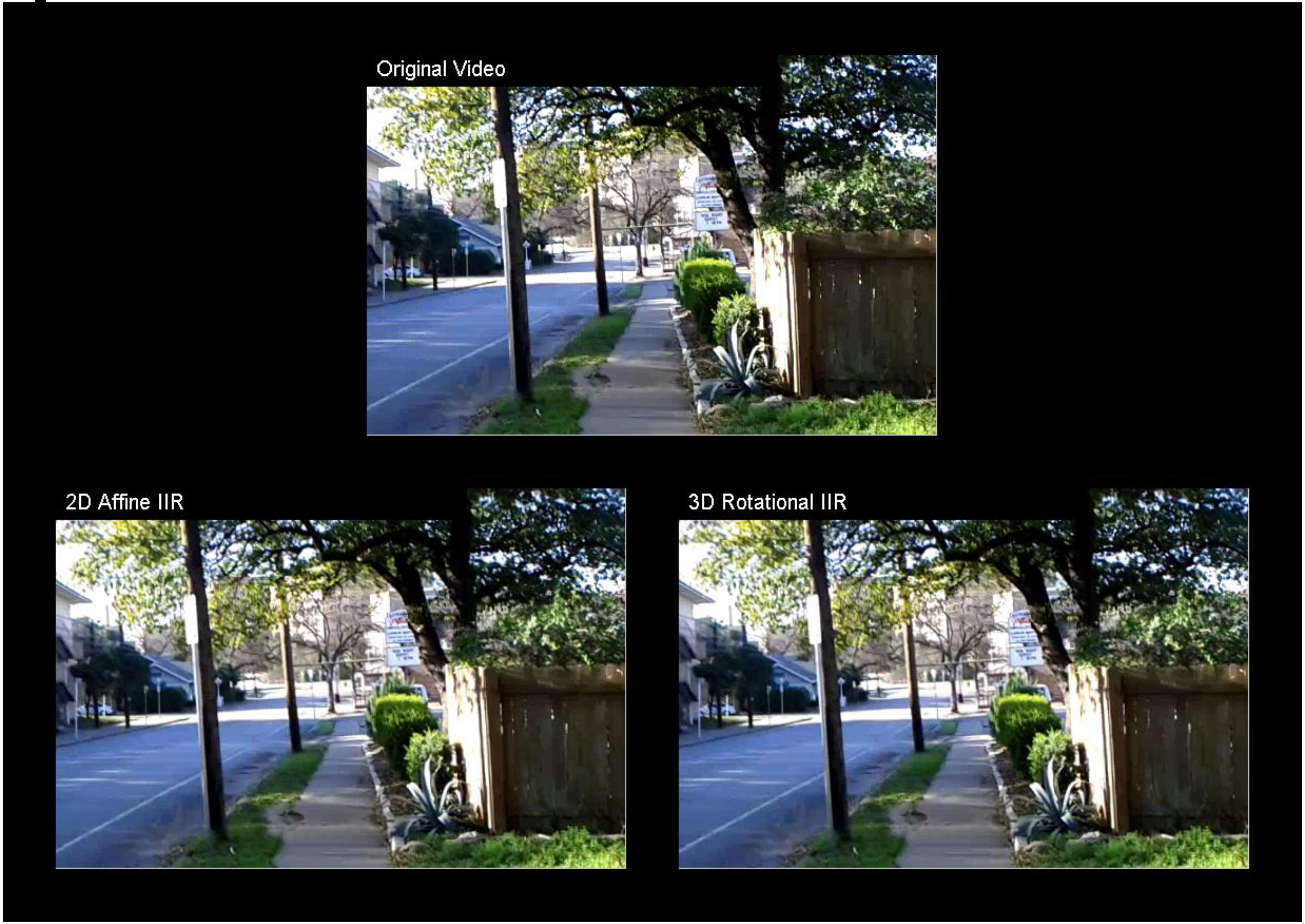


Algorithm UKF-based 3D Rotation Smoothing

```
Input: $\mathbf{q}_1, \cdots, \mathbf{q}_K$ (original rotations)
Output: $\hat{\mathbf{q}}_1, \cdots, \hat{\mathbf{q}}_K$ (smoothed rotations)
Parameters: $\mathbf{Q}, \mathbf{R}$ (process and measurement noise variance)
for $k = 1$ to $K$ do
    Obtain unconstrained UKF estimate $\hat{\mathbf{q}}_k^*, \hat{\boldsymbol{\omega}}_k^*, \mathbf{P}_k$
    $\hat{\mathbf{q}}_k^* = \hat{\mathbf{q}}_k^* / ||\hat{\mathbf{q}}_k^*||_2$ (normalization)
    $\hat{\mathbf{q}}_k \leftarrow \mathbb{P}(\hat{\mathbf{q}}_k)$
    (Mean and covariance estimate to pass to the next stage are $\hat{\mathbf{q}}_k, \hat{\boldsymbol{\omega}}_k, \mathbf{P}_k$)
end for
```

Experimental Results – 2D vs. 3D KF

Experimental Results – 2D vs. 3D IIR



Conclusions

- Video Stabilization & Rectification for Handheld Cameras
  - CMOS image sensors (rolling shutter effects)
  - Equipped with gyroscopes

Camera-Gyroscope Calibration

- Online calibration & synchronization
- No need to estimate translation

Offline 3D Rotation Smoothing

- Stabilization as regression on manifold
- Convex approximation of constraint
- Manifold optimization

Online 3D Rotation Smoothing

- IIR-like smoothing
- UKF (unscented Kalman Filter)-based smoothing
- Ad-hoc projection for constraint

Thanks!

References

- [Bertsekas 1999] D. Bertsekas, *Nonlinear Programming*, Athena Scientific, 1999.
- [Civera JFR 2010] J. Civera, O.G. Grasa, A.J. Davison, and J.M.M. Montiel, "1-point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry," *Journal of Field Robotics*, vol. 27, issue 5, pp. 609-631, Sep. 2010
- [Gallier 2012] J. Gallier, *Notes on Differential Geometry and Lie Groups*. University of Pennsylvania, 2012.
- [Grundmann CVPR 2011] M. Grundmann, V. Kwatra, and I. Essa, "Auto-directed video stabilization with robust L1 optimal camera paths," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Jun. 2011.
- [Hanning IMWV 2011] G. Hanning, N. Forslow, P.-E. Forssen, E. Ringaby, D. Tornqvist, and J. Callmer, "Stabilizing cell phone video using inertial measurement sensors," in *Proc. IEEE Intl. Workshop on Mobile Vision*, Nov. 2011.
- [Hwangbo IJRR 2011] M. Hwangbo, J.-S. Kim, and T. Kanade, "Gyro-aided freature tracking for a moving camera: fusion, auto-calibration and GPU implementation," *Intl. Journal Robotics Research*, vol. 30, no. 14, pp. 1755-1774, 2011.
- [Karpenko report 2011] A. Karpenko, D. Jacobs, J. Baek, and M. Levoy, "Digital video stabilization and rolling shutter correction using gyroscopes," *Tech. Rep., Stanford University*, Mar. 2011.
- [Kneip ECCV 2012] L. Kneip, R. Siegwart, and M. Pollefeys, "Finding the exact rotation between two images independently of the translation," in *Proc. European Conf. Computer Vision*, Oct. 2012.
- [Morimoto CVPR 1997] C. Morimoto, and R. Chellapa, "Fast 3D stabilization and mosaic construction," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Jun. 1997.
- [Pollefeys IJCV 2004] M. Pollefeys, L. Van Gool, M. Vergauwen, F. Verbiest, K. Cornelis, J. Tops, and R. Koch, "Visual modeling with a hand-held camera," *Intl. Journal of Computer Vision*, vol. 59, no. 3, pp. 207-232, 2004.

Personal Publications

- Journal Papers
  - C. Jia and B.L. Evans, “Real-time Motion Smoothing for Video Stabilization via Constrained Multiple-Model Estimation”, *IEEE Trans. on Circuits and Systems for Video Technology*, to be submitted.
  - C. Jia and B.L. Evans, “Online Camera-Gyroscope Auto-Calibration for Cellphones”, *IEEE Trans. on Image Processing*, In revision.
  - C. Jia and B.L. Evans, “Constrained 3D Rotation Smoothing via Global Manifold Regression for Video Stabilization”, *IEEE Trans. on Signal Processing*, Accepted with minor revision.
  - V. Raveendran, P. Bhamidipati, X. Luo, X. Huang and C. Jia, “Mobile Multipath Cooperative Network for Real-time Streaming”, Signal Processing: Image Communication, Vol 27, Feb 2012

- Conference Papers
  - C. Jia, Z. Sinno and B.L. Evans, “Real-Time 3D Rotation Smoothing for Video Stabilization”, *Asilomar Conf. on Signals, Systems, and Computers*, Nov. 2-5, 2014, Pacific Grove, CA, to be submitted.
  - C. Jia and B.L. Evans, “Online Calibration and Synchronization of Cellphone Camera and Gyroscope”, *Proc. IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Dec. 3-5, 2013, Austin, TX.
  - C. Jia and B.L. Evans, “3D Rotational Video Stabilization using Manifold Optimization”, *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, May. 26-31,2013, Vancouver, Canada. **(Google Travel Grant Winner)**
  - C. Jia and B.L. Evans, “Probabilistic 3D Motion Estimation for Rolling Shutter Video Rectification from Visual and Inertial Measurements”, *Proc. IEEE Int. Workshop on Multimedia Signal Processing (MMSP)*, Sep. 17-20, 2012, Banff, Canada. **(Top 10% Best Paper Award)**
  - C. Jia and B.L. Evans, “Patch-based Image Deconvolution via Joint Modeling of Sparse Priors”, *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, Sep. 11-14, 2011, Brussels, Belgium.