Robust Frame-to-Frame Camera Rotation Estimation in Crowded Scenes

Fabien Delattre [1] David Dirnfeld [1] Phat Nguyen [1] Stephen Scarano [1] Michael J. Jones [2] Pedro Miraldo [2] Erik Learned-Miller [1]

[1] *University of Massachusetts Amherst* [2] *Mitsubishi Electric Research Laboratories (MERL)*

Problem statement

Goal

Estimate the 3D camera rotation given two consecutive video frames using vision.

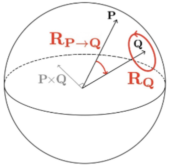
Challenges

- Pixel motions can be caused by camera rotation and translation, but also by moving objects and noise in the correspondences between the consecutive frames.
- Therefore, motion estimation algorithms need to be run using a robust estimation algorithm such as RANSAC.

Contributions

- We propose a novel generalization of the Hough transform on SO(3) for camera rotation estimation.
- We introduce a new dataset, BUsy Street Scenes Dataset (BUSS).

Finding the compatible rotations



Using perspective projection

The problem of finding the compatible rotations with a single optical flow vector boils down to finding the set of rotations that map P to Q. This set can be obtained by composing a single rotation about P×Q with any rotation about the vector Q.

Using Longuet-Higgins motion model

The rotation component of the Longuet-Higgins motion model express the optical flow v in terms of

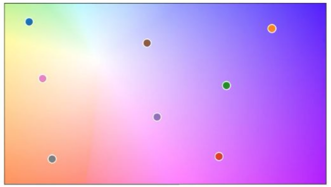
$$\mathbf{v}(x,y)=\begin{pmatrix} A\left(\frac{xy}{f}\right)-B\left(\frac{f^2+x^2}{f}\right)+Cy \\ A\left(\frac{f^2+y^2}{f}\right)-B\left(\frac{xy}{f}\right)-Cx \end{pmatrix}$$

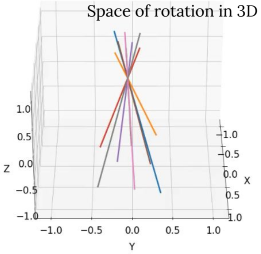
the rotation velocities A,B,C. Each equation describes a plane in the rotations space. Thus, the system of equations describes the intersection of the two planes, which is the line of compatible rotations.

Hough transform on SO(3)

An optical flow vector has two free parameters, and the space of rotation is 3D, therefore there is a 1d manifold of compatible rotations.

Pure rotation flow case:




Space of rotation in 3D

In the pure rotation case, all the lines of compatible rotations intersect in a single point in the rotation space which is the camera rotation that has generated the optical flow.

General case:

Optical flow vectors on far away points can be considered as generated by camera rotation only. Optical flow vectors on distant objects give coherent rotations, while optical flow on moving or nearby points don't give coherent rotations.

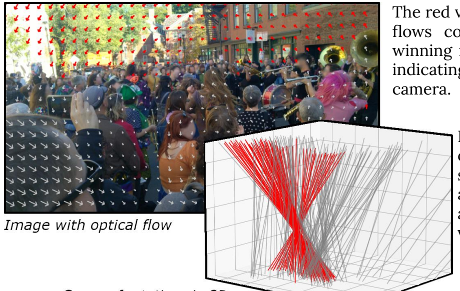


Image with optical flow

Space of rotations in 3D

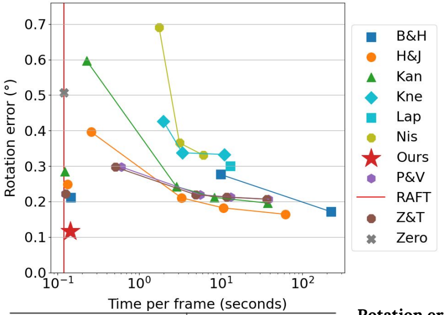
The red vectors show optical flows compatible with the winning rotation estimation, indicating the rotation of the camera.

Each line shows the one-dimensional set of rotations that are compatible with a single optical flow vector.

BUsy Street Scenes Dataset (BUSS)

We introduced a new dataset composed of 17 video sequences of about 10 seconds each at 30 fps in full HD resolution (1920x1080). The videos are taken from a handheld mobile phone in crowded city streets with synchronized gyroscope data.

Results



	Rotation err. (°)					Time per frame (seconds)				
Continual methods										
# iters.	N/A	1	25	100	500	N/A	1	25	100	500
B&H [7]	0.21	0.28	0.17	—*	—*	0.14	9.92	226.01	—*	—*
H&J [20]	0.25	0.40	0.21	0.18	0.16	0.13	0.26	3.30	10.72	62.09
Kan [30]	0.28	0.60	0.24	0.21	0.20	0.12	0.23	2.86	8.37	37.54
L&R [39]	0.30	—	—	—	—	13.07	—	—	—	—
P&V [58]	0.22	0.30	0.22	0.21	0.21	0.12	0.61	5.62	13.27	38.76
Z&T [75]	0.22	0.30	0.22	0.21	0.21	0.13	0.51	4.93	11.82	36.50
Ours	0.12	—	—	—	—	0.14	—	—	—	—
Discrete methods										
# iters.		500	5K	50K			500	5K	50K	
Kneip [34]		0.69	0.36	0.33			1.63	3.00	6.04	
Nistér [55]		0.43	0.34	0.33			1.86	3.26	10.91	

On BUSS, our method is 50% more accurate than the best methods for similar runtimes, and more accurate than any methods that have "reasonable" run times. On the IRSTV dataset (dataset that is mostly composed of static scenes), our method is on par with the other methods with respect to accuracy and speed.

Rotation error vs. run time on BUSS

Methods run with RANSAC appear on a line, with different numbers of RANSAC iterations at each point. Standalone points do not use RANSAC. The run time of the continual methods includes the run time of the optical flow computation.

Correspondence to: Fabien Delattre (fdelattre@umass.edu)