F16 16-720 Computer Vision:

Real-time Homography Estimation and Augmented Rendering

Carnegie Mellon University

Chao-Ming Yen, John Ryan, Xidong Pi

Introduction

Motivations and Goals:

❑Estimate the transformation matrix of target object in real time and augment some funny new images onto the original object in video

❑Need to track the target object in video efficiently and accurately

❑Investigate different tracking algorithms and compare their efficiency and accuracy

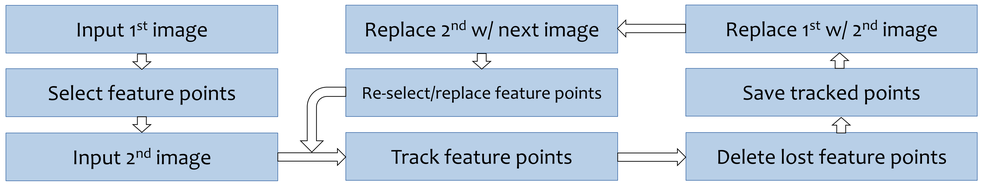
Methodologies & Algorithms

I. Specify the Target Object

II. Tracking

❑ Lucas-Kanade Inverse Compositional (LK-IC) Tracker

❑ SURF/MinEigen Features + Kanade-Lucas-Tomasi (KLT) Tracker



❑ Conditional Lucas-Kanade (C-LK) Tracker (Lin & Lucey 2016)

$$\min_{\nabla\mathcal{T}(\mathbf{0})} \sum_{n\in\mathcal{S}} \|\Delta\mathbf{p}_n - \mathbf{R}[\mathcal{I}(\mathbf{p}_n \circ \Delta\mathbf{p}_n) - \mathcal{T}(\mathbf{0})]\|_2^2$$

$$\text{s.t.} \quad \mathbf{R} = \left(\begin{bmatrix} \nabla\mathcal{T}(\mathbf{x}_1) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \nabla\mathcal{T}(\mathbf{x}_D) \end{bmatrix} \begin{bmatrix} \frac{\partial\mathcal{W}(\mathbf{x}_1;\mathbf{0})}{\partial\mathbf{p}^\top} \\ \vdots \\ \frac{\partial\mathcal{W}(\mathbf{x}_D;\mathbf{0})}{\partial\mathbf{p}^\top} \end{bmatrix} \right)^\dagger$$

$$\min_{\nabla\mathcal{T}(\mathbf{0})} \left\| \Delta\mathbf{p} - \left(\nabla\mathcal{T}(\mathbf{0}) \quad \frac{\partial\mathcal{W}(\mathbf{x};\mathbf{0})}{\partial\mathbf{p}} \right)^\dagger \mathcal{I}(\mathbf{p}) - \mathcal{T}(\mathbf{0}) \right\|_2^2$$

N pixels → $2N$ degrees of freedom

predefined

predicts $\Delta\mathbf{p}$ conditioned on $\mathcal{I}(\mathbf{p}) - \mathcal{T}(\mathbf{0})$

III. Homography Estimation & Augmented Rendering

❑ Transformation Matrix + Alpha Blending

Experiment Results

❑ LK-IC

❑ KLT

❑ C-LK

❖ Running time & Accuracy

For tracking 100 frames in the sample video

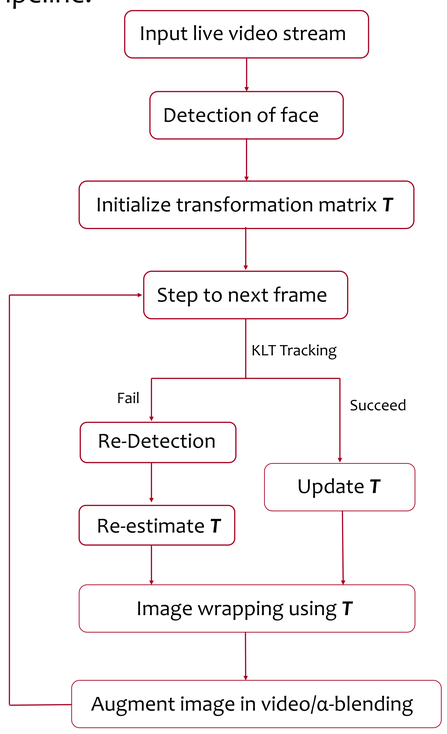
- LK-IC: 41.6582 s
- KLT: 6.6989 s
- C-LK: 8.5127 s

- The accuracy for KLT and C-LK tracker is obviously much better than LK-IC
- C-LK uses 6 pre-trained regressors, which is in turn more robust.

Real-time Demo

Face Detection + Tracking + Homography Estimation + AR

❑ Pipeline:



❑ Examples:

Conclusion & Future Work

❑ Kanade-Lucas-Tomasi (KLT) & Conditional Lucas-Kanade (C-LK) Tracker works much better, both in running time and accuracy of result.

❑ In case of low-frame-rate AR, the C-LK tracking might out-perform the other LK algorithms due to its higher tolerance for frame-to-frame shift. The KLT tracker of the above framework could be replaced with C-LK tracker in the future.