Multi-Sensors Fusion and Tracking

Multi-Face Tracking using Kanadi Lucas-Tomasi Algorithm

F. Legrand, O. Agbohoui, V. Kodipaka

Centre Universitaire Condorcet - Master 2 in Computer Vision - 2019/2020

December 12, 2019

UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ

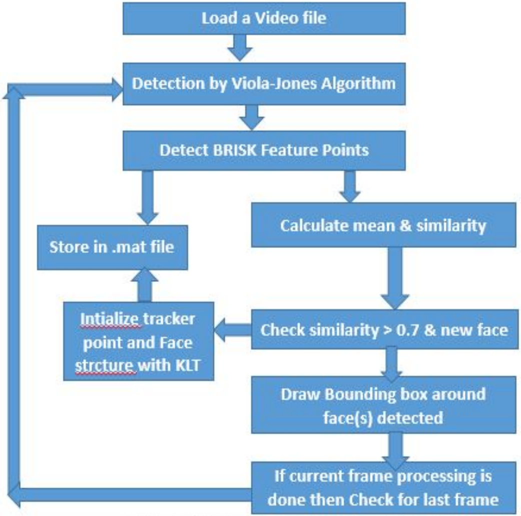
Introduction

Introduction

Overview

- Methodology
  - Viola Jones Detection
  - Kanade Lucas-Tomasi Tracking
- Implementation
- Results
- Conclusion

Introduction



Load a Video file

Detection by Viola-Jones Algorithm

Detect BRISK Feature Points

Store in .mat file

Calculate mean & similarity

Intialize tracker point and Face strcture with KLT

Check similarity > 0.7 & new face

Draw Bounding box around face(s) detected

If current frame processing is done then Check for last frame

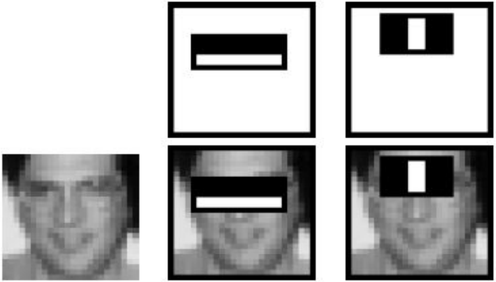
BLOCK REPRESENTATION OF THE PROJECT

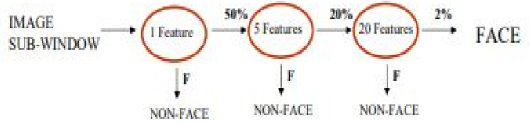
Methodology - Viola Jones Detection

Viola-Jones method can be splitted into 4 main steps:

- Haar features comparison and differentiation
- Combination into integral image
- Adaboost enhanced classification
- Attentional cascade discarding most of the negative classification
- BRISK: applied after the VJ method

Methodology - Viola Jones Detection

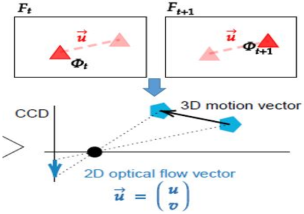




Methodology - KLT

- Brisk features as input for tracker
- Detection throughout the window
  (Centred around features)
- Pyramidal Fashion
- Compute the similarity
- Integrate window
  Pixel -> Neighbours (Height and Width) -> Area

Methodology - KLT



$$\boxed{I_x u + I_y v = -I_t}$$

Optical Flow base:

$$\frac{d}{dt} I(x(t), t) = \nabla I^T \left(\frac{dx}{dt} \right) + \frac{\delta I}{\delta t} = 0.$$

Brightness Consistency Equation:

$$\nabla I(x', t)^T v + \frac{\delta I}{\delta t} = 0$$

$$E(v) = \int_{W(x)} |\nabla I(x', t)^T v + I_t(x', t)|^2 dx'.$$

$$E(b) = \int_{W(x)} |\nabla I^T b + I_t|^2 dx'$$

Methodology - KLT

Translation Motion and Affine Motion:

$$E(b) = \int_{W(x)} |\nabla I(x')^T S(x')p + I_t(x')|^2 dx'$$

$$S(x)p = \begin{pmatrix} x & y & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x & y & 1 \end{pmatrix} \begin{pmatrix} p_1 & p_2 & p_3 & p_4 & p_5 & p_6 \end{pmatrix}^T.$$

$$M(x) = \int_{W(x)} \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} dx'$$

$$b(x,t) = -M(x)^1 \begin{pmatrix} \int I_x I_t dx' \\ \int I_y I_t dx' \end{pmatrix}$$

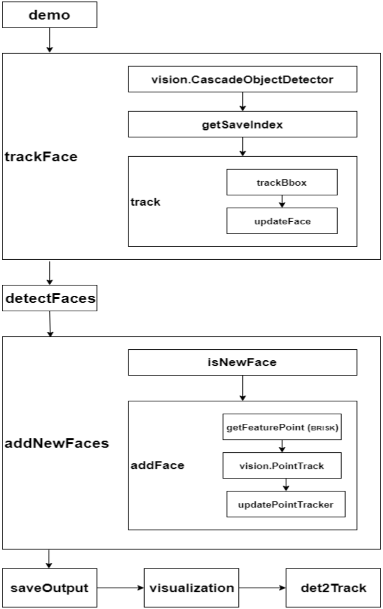
$$I_x u + I_y v = -I_t \quad \Rightarrow \quad \begin{bmatrix} I_x & I_y \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = -I_t$$

Minimization, Multi-scale estimation and Local Smoothness:

$$\hat{u} = (A^T A)^{-1} A^T b$$

$$A^T A = \begin{bmatrix} \sum I_x^2 & \sum I_x I_y \\ \sum I_x I_y & \sum I_y^2 \end{bmatrix}$$

Implementation



Implementation - Viola-Jones

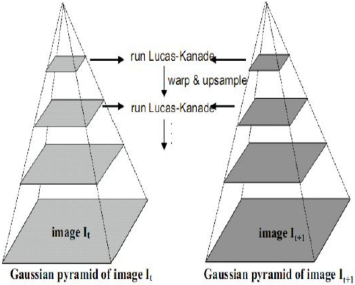
- Usage classification and reduction for Frontal Face Detection
- Algorithm loads pre-trained classification model
- Features encoded and stored in .xml files
- 12 Haar models associated to different regions of the face

$$\mathrm{Similarity} = \left(\frac{\varphi}{\Psi 1} + \frac{\varphi}{\Psi 2} \right) \big/ 2$$

- BRISK: applied after the bounding box is computed

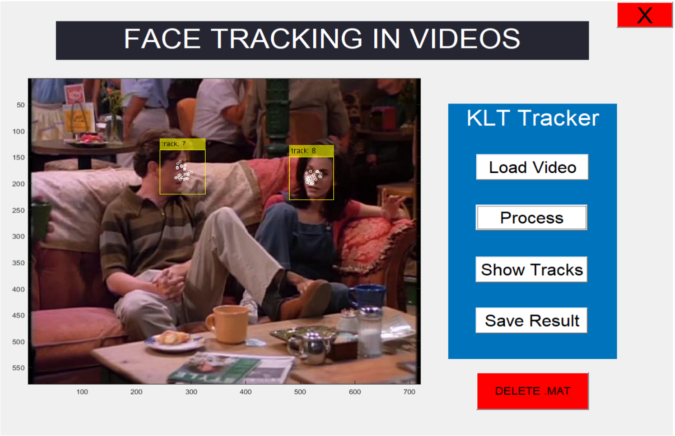
Implementation - KLT

- Neighbour area increase = complexity increases (31x31)
- Max iteration search for each frame
  (each point conversion rate of 10)
- Bidirectional error (tracking error) - T0 and T1 frames
- Pyramid - downsampling - displacement
  increased - between frames - computation



Results

- Detection by VJ
- Brisk Features
- Tracking by KLT



Conclusion

- Multi-Face Applications
- Frame Rate Range
- Object Motion
- Limitations - Computational complexity

References

- Viola, Paul Jones, Michael. (2004). Robust Real-Time Face Detection. I
- Leutenegger, Stefan Chli, Margarita Siegwart, Roland. (2011). BRISK: Binary Robust invariant scalable keypoints.
- Ranganatha S, Y P Gowramma, Development of Robust Multiple Face Tracking Algorithm and Novel Performance Evaluation Metrics for Different Background Video Sequences
- Tomasi, Carlo Kanade, Takeo. (1999). Shape and Motion from Image Streams: a Factorization Method?
- Equations are referred from lecture slides of Mr. Marc Blanchon, (Centre Universitaire Condorcet MSCV2 2019-2020), VISUAL TRACKING: Multi-Sensor Fusion and Tracking.

THANKS