Cleaning up after a Face Tracker **False Positive Removal**

Makarand Tapaswi¹, Cemal Çağrı Çörez², Martin Bäuml¹ Hazım Kemal Ekenel², Rainer Stiefelhagen¹

¹Computer Vision for Human Computer Interaction, Karlsruhe Institute of Technology, Germany ²Smart Interaction, Mobile Intelligence, and Multimedia Technologies Lab, Istanbul Technical University, Istanbul, Turkey

Contributions

- Automatically discard false positive face tracks while keeping true positive tracks intact
- Generic (confidence-based) and domain-specific (context-based) cues for identifying false positive face tracks
- Evaluation on a database of over 11000 tracks from 2 diverse TV series



Motivation

Face detection and tracking is a prerequisite for person identification in TV series. Errors during detection such as false positive face tracks affect the results but are typically ignored.

In this work, we propose a set of cues both generic and domain-specific to tackle the problem of false positive face track detection while minimizing the removal of true positive tracks.



Context based cues

facial location heat maps

In a TV series video frame, where do faces appear?



relative size

Faces tend to appear at the same depth in video frames



Confidence based cues

skin color Simple, but effective method to detect false positive faces





facial feature point (landmark) localization

Detect face landmarks using the nine point detector [1]. Compute average confidence score for each track



[1] M. Everingham, J. Sivic, and A. Zisserman. "Hello! My name is... Buffy" - Automatic naming of characters in TV video. In British Machine Vision Conference (BMVC), 2006.



animation

Face tracks are animate objects, false positives are not





Results

Classification Results

— Correctly detected false positive tracks Wrongly classified true positive tracks

| | BBT | BUFFY | TOTAL |
|-----------------|------------------------|------------------------|-------------------------|
| #Tracks | 4684 | 6759 | 11443 |
| #FPFT | 903 | 997 | 1900 |
| Skin Confidence | 361 / <mark>23</mark> | 284 / <mark>65</mark> | 645 / <mark>88</mark> |
| Facial Features | 127 / <mark>11</mark> | 170 / <mark>71</mark> | 297 / <mark>82</mark> |
| Animation | 105 / <mark>29</mark> | 134 / <mark>107</mark> | 239 / <mark>136</mark> |
| Facial Location | 408 / <mark>105</mark> | 229 / <mark>22</mark> | 637 / <mark>127</mark> |
| Relative Size | 218 / <mark>6</mark> | 275 / <mark>42</mark> | 493 / <mark>48</mark> |
| Combined | 690 / <mark>129</mark> | 754 / <mark>218</mark> | 1444 / <mark>347</mark> |
| MOTA Before | 68.2 | 69.0 | 68.6 |
| MOTA After | 75.1 | 72.8 | 74.0 |

Different detectors / trackers

ROC curve

| Detector | | BBT-1 | BUFFY-5 |
|-----------------------------|----------|---------------------|-----------------------|
| Haar + KLT | #Tracks | | 760 |
| | #FPFT | — | 94 |
| | Combined | | 51 / <mark>22</mark> |
| MCT + Assoc. Based | #Tracks | 584 | |
| | #FPFT | 53 | _ |
| | Combined | 41 / <mark>4</mark> | |
| MCT + Particle Filter | #Tracks | 704 | 963 |
| | #FPFT | 79 | 150 |
| | Combined | 59 / <mark>3</mark> | 113 / <mark>40</mark> |



Contact

tapaswi@kit.edu , cccorez@gmail.com

Project page

http://simitlab.itu.edu.tr



Acknowledgment: This work was supported by the Quaero Programme, funded by OSEO; TUBITAK within the CHIST-ERA project CAMOMILE, project no. 112E176; and a Marie Curie FP7 Integration Grant within the 7th EU Framework Programme.

KIT – University of the State of Baden-Wuerttemberg and National Research Center of the Helmholtz Association



