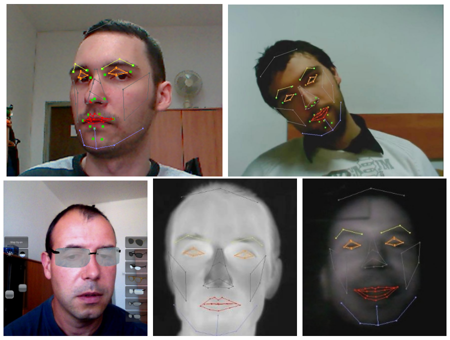
High-performance face tracking

Nenad Markuš, Miroslav Frljak, Igor S. Pandžić, Jörgen Ahlberg, Robert Forchheimer [*]



1 Introduction

Face tracking is an well studied field. Nevertheless, it is still a challenge to make a robust and efficient face tracker, especially on mobile devices. This extended abstract briefly describes our implementation of a high-performance face tracking system. The main characteristics of our approach are that the tracker is fully automatic and works with the majority of faces without any manual initialization. It is robust, resistant to rapid changes in pose and facial expressions, does not suffer from drifting and is modestly computationally expensive. The tracker runs in real-time on mobile devices.

2 The face tracking algorithm

We are interested in the global position and rotation of the face and its expressions, i.e. facial action. We address the problem as fitting a 3D deformable mesh to the face of the user. Candide [Ahlberg 2001] is a simple wireframe face model. To modify the geometry, the model implements a set of action parameters that enable dynamic changes in facial geometry, like eye blinking and lip movement. Thus, for each frame of the video sequence, we are interested in estimation of rotation, translation and action parameters that best fit the face of the user.

The tracker uses a Viola-Jones detector to determine the position of the face in the initial frame. Next, it proceeds to roughly fit the 3D model and select the characteristic points on the face (we denote these as feature points). The feature points are tracked from one frame to another using normalized cross-correlation. The obtained data is passed to the information form of extended Kalman filter (EIF) [Bar-Shalom et al. 2001]. The filter recursively estimates the rotation, translation and action parameters.

In real-world conditions, errors occur in the computation of feature point correspondences between two consecutive frames due to illumination variation, noise, occlusions, etc. These errors accumulate over time and the system will lose track of the face in long video sequences. To solve the problem of drifting, we use a textured face model, similar to [Ingemars and Ahlberg 2007]. The texture is extracted from the initialization frame, using the Candide face mesh. The main idea is to use the currently estimated rotation, translation and action parameters to transform the wireframe face model vertices and render the face texture in the frame. Next, the feature point correspondences between this rendered model and the actual face in the original frame are determined. The EIF is used to refine the estimated model parameters. This correcting step is sufficient to stop the drift problem.

–	CPU	FPS
PC	2GHz Intel Core 2 Duo T7200	30
iPad	dual-core 1GHz Cortex-A9	30
iPhone 4S	dual-core 1GHz Cortex-A9	30
iPhone 4	1GHz Cortex-A8	20
HTC	1GHz Scorpion	12

3 Tracking performance

In order to evaluate the performance of our face tracking implementation, we conducted a series of tests under various conditions. The quality of the system can be best assessed from the following video: `http://youtu.be/7ziXA4ZSRSA`.

The images show snapshots from various tracking sessions. These include tracking from streams supplied from ordinary webcams as well as professional hardware (near-infrared and thermal cameras). One of the images depicts a virtual eyewear try-on application running on a mobile platform. The table shows the frame rates achieved by the tracker on various devices. The image stream was supplied from the built-in camera. Thus, the frame rates are bounded by the 30 FPS camera input rate.

4 Ongoing work

Ongoing work focuses on the possibility of using randomized forests [Breiman 2001] for detection of characteristic points on the face (for example, eye or eyebrow corners). We hope that this research will result in the improvement of tracker initialization as well as increase in tracking quality and stability. The approach could help combat the observed tracker drift problem and even make the previously described correcting step redundant.

5 Acknowledgements

The core tracking algorithm has been developed by Visage Technologies AB, Linköping, Sweden and work continues in collaboration with the company. This research is partialy supported by the Ministry of science education and sports of the Republic of Croatia, grant nr. 036-0362027-2028 ”Embodied Conversational Agents for Services in Networked and Mobile Environments” and by the ACROSS project, 285939 FP7-REGPOT-2011-1.

References

AHLBERG, J. 2001. Candide-3 – an updated parameterised face. Tech. rep., Linköping University.

BAR-SHALOM, Y., LI, X. R., AND KIRUBARAJAN, T. 2001. *Estimation with applications to tracking and navigation*. John Wiley and Sons.

BREIMAN, L. 2001. Random forests. *Journal of Machine Learinng Research*.

INGEMARS, N., AND AHLBERG, J. 2007. Feature-based face tracking using extended kalman filtering. In *Swedish Symposium on Image Analysis*.

LEFVRE, S., AND ODOBEZ, J.-M. 2010. View-based appearance model online learning for 3d deformable face tracking. In *Int. Conf. Computer Vision Theory and Applications*.

[*] e-mail: {nenad.markus, miroslav.frljak, igor.pandzic}@fer.hr, {jorgen.ahlberg, robert}@isy.liu.se