Affordable Yet Robust and Precise Face Tracking using USB Cameras with Application to Designing Hands-Free User Interfaces

Gorodnichy, Dimitry; Roth, Gerhard

National Research Council Canada
Institute for Information Technology

Conseil national de recherches Canada
Institut de technologie de l'information

NRC · CNRC

*Affordable Yet Robust and Precise Face Tracking using USB Cameras with Application to Designing Hands-Free User Interfaces **

Gorodnichy, D., and Roth, G.
October 2002

* published in Proceedings of the ACM Conference on Software and Technology of Human-Computer Interfaces. October 27-30, 2002. NRC 45856.

Canada

Affordable yet robust and precise face tracking using USB cameras with application to designing hands-free user interfaces

Dmitry O. Gorodnichy and Gerhard Roth
Computational Video Group, IIT, NRC
Ottawa, Canada K1A 0R6
www.cv.iit.nrc.ca

CHALLENGE OF DESIGNING FACE TRACKING BASED UI

The applications of vision-based face tracking to HCI are evident. Face tracking based program control can be used as a hands-free alternative and/or extension to conventional pointing devices such as mouse, joystick, track pad or track ball. This can be used, for example, to switch the focus of attention in windows environment. Vision-based perceptual user interfaces can also be used to control commercial computer games, immersive 3D worlds and avatar-like computer-generated communication programs. For users with physical disabilities, this technology offers a way of controlling an on-screen cursor by moving their heads. Finally, face tracking has applications in security industry, where it is a prerequisite for the next step of face recognition, in video-conferencing, where it can be used to correct the gaze direction, and also in video-coding and content-based image retrieval.

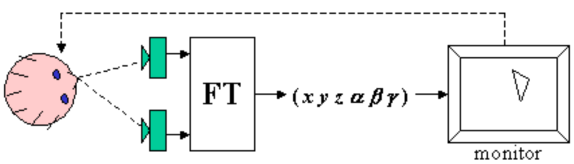


Figure 1: Face tracking based user interface.

Designing face tracking based user interfaces, such as illustrated in Figure 1, which use a videocamera to track user's face position in 3D in order to convert it to a position of a cursor or another virtual object in 2D screen, may also seem to be a straightforward problem to resolve. However, as researchers in the computer vision community know [4], it is just the opposite. The problem is that for such interfaces to be useful, face tracking has to be fast, affordable and, most importantly, precise and robust. In particular, the precision should be sufficient to control a cursor, while the robustness should be high enough to allow a user the convenience and the flexibility of head motion. Unless dedicated hardware or structured environment (e.g. markings on the user's face) are used, these goals are very difficult to achieve. Even with recent advances in hardware and falling camera prices, the obtained vision-based solutions still do not exhibit the required precision and robustness.

Reviewing vision-based solutions, two classes of face tracking approaches are considered [5]: (global) image-based and (local) feature-based. Image-based approaches use global facial cues such as skin colour, head geometry and motion. They are robust to head rotation and scale and do not require high quality images. These approaches however lack precision and therefore cannot be used to control the cursor precisely. In order to achieve precise and smooth face tracking, feature-based approaches are used [3]. These approaches are based on tracking individual facial features and theoretically can track faces with pixel-size precision, at least. In practice however they do not, as tracking of the local facial features is not robust to the head's rotation, scale and the changes of facial expressions. Feature-based approaches usually also require high-resolution cameras, which makes them less affordable.

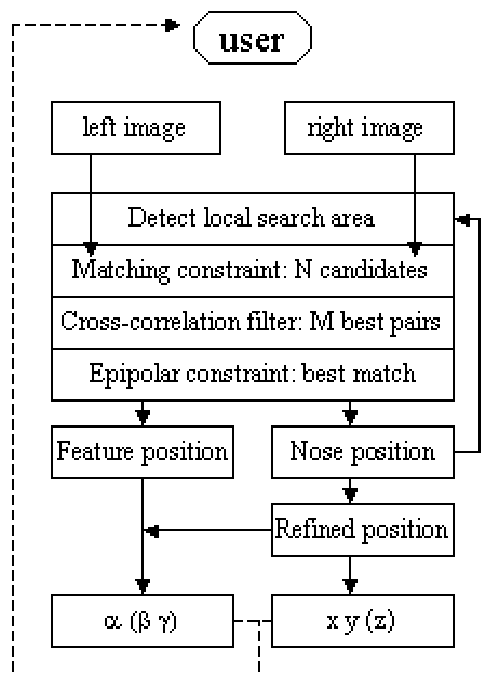


Figure 2: Overview of the tracking procedure for the face tracking based user interfaces.

Perceiving with two eyes

While for humans it is definitely easier to track objects with two eyes than with one eye, in computer vision, face tracking is usually done with one camera only. This is due to the fact that, unlike in human brains, the relationship between the images observed by two arbitrary video cameras, in many cases, is not known.

A few authors do use stereo for face tracking [2]. They however use the second camera mainly for the purpose of acquiring the third dimension rather than making tracking more robust, precise or affordable. The problem is that, in order to track a face in 3D, the tracking should be robust to the 3D motion of the head, which, as mentioned above, is not the case for the feature-based approaches. In addition, most used stereotracking approaches rely on dedicated hardware with precalibrated high resolution cameras, and hence are not very affordable.

CONTRIBUTION OF OUR WORK

In this demo, we present techniques which allow one to build an affordable face tracking system which tracks faces both precisely and robustly. First, we use the convex-shape nose feature proposed in [1], which can be tracked with subpixel accuracy at all times, regardless of face orientation and expression. Second, we use recent advances in projective vision theory to compute the relationship between any two cameras. This relationship, represented by the fundamental matrix, is naturally obtained while observing a face with both cameras. It provides an extra constraint which makes face tracking more robust.

The proposed techniques allow us to build precise face tracking based user interfaces with the aid of generic USB cameras. We show the **Nouse** *'Use your* **n***ose as m***ouse***'* 2D hands-free user interface we have designed. We also show the experimental results of our 3D face tracking technique which combines convex-shape nose tracking with stereo tracking. The robustness of our face tracking technique, the binary code of which can be downloaded from our website, is such that the rotations of a head of up to 40 degrees in all three axes of rotation can be tracked. Figure 2 shows the overview of our face tracking technique as applied to designing hands-free user interfaces and Figures 3-5 show a few hands-free interactive programs which is based on the proposed technique.

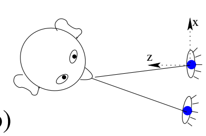
a)  b) 

Figure 3: A user plays an aim-n-shoot game, pointing the direction of shooting with his nose. Very slight rotations of head left and right are sufficient to cover the entire 180^o range of aiming a turret.

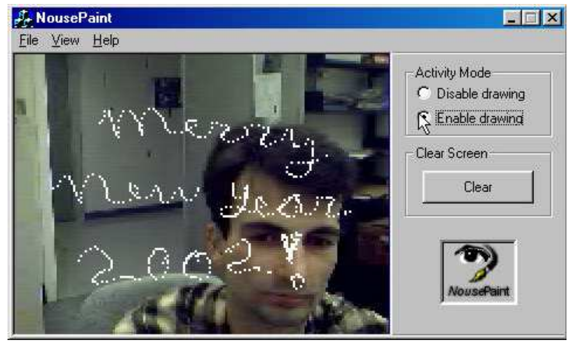


Figure 4: The robustness and precision of the tracking technique is such that it allows one to write with the nose not causing any discomfort or fatigue.

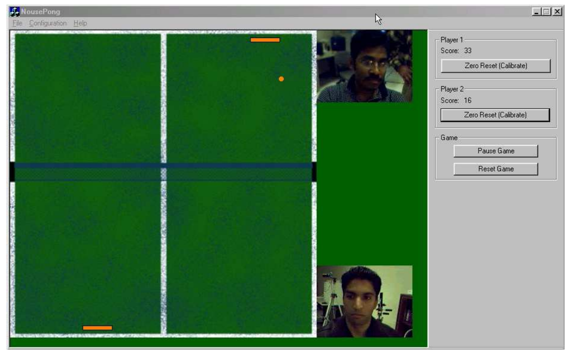


Figure 5: Two users are playing a pong game using their heads to bounce a ball.

REFERENCES

1. D.O. Gorodnichy. On importance of nose for face tracking. In *FG 2002*, Washigton D.C., May 2002.

2. R.Newman, Y. Matsumoto, S. Rougeaux, and A. Zelinsky. Real-time stereo tracking for head pose and gaze estimation. In *FG 2000*, 2000.

3. K. Toyama. Look, Ma – no hands! Hands-free cursor control with real-time 3D face tracking. In *PUI'98*, San Fransisco, November 1998.

4. M. Turk, C. Hu, R. Feris, F. Lashkari, and A. Beall. TLA based face tracking. In *VI 2002 (available online www.visioninterface.org/vi2002)*, Calgary, May 2002.

5. M. Yang, N. Ahuja, and D. Kriegman. Detecting faces in images: A survey. *IEEE TPAMI*, 24(1):34–58, 2002.