# Face on the Fly: How to Track Faces

Catherine Herold, Stéphane Gentric, Nicolas Moënne-Loccoz
Morpho, Safran Group
catherine.herold, stephane.gentric@morpho.com

*A great challenge for border-control gates is to maintain good biometric performances, even with less cooperative subjects. We present here a novel method to simultaneously track and estimate the 3D head pose of people walking in an authentication gate (Fig. 1), from video sequences acquired by a multi-view system. A 3D head model is adapted to each individual at the beginning of the sequence and is then used during the tracking process to generate images of the face under any pose.*

**From 3D face to 2D views**. A main issue concerning tracking is the way to describe the face. Due to appearance changes under large pose variations, using 2D descriptors from the initial frame may lead to tracking failure. Given a 3D pose $x$ to evaluate, we render the corresponding views of the 3D model (Fig. 2) and compare them to the observations.

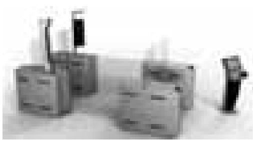

Fig 1. Authentification system

Fig 2. Synthesized views

**3D Multiview tracking by particle filter**. In a Bayesian framework, tracking consists of estimating the density of the current state $p(x_t|y_{0:t})$ in a recursive way, given the observations $y_{0:t}$. This density is approximated by a set of $N$ pose samples $x^{(i)}$, called particles, affected by a weight $w^{(i)}$. The weight of a particle is defined as a likelihood score between the corresponding synthesized views and the observations, leading to the following density approximation: $p(x_t|y_{0:t}) = \sum_{i=1}^{N} w^{(i)}x^{(i)}$. A multi-layer sampling approach leads to an improvment of the tracker performances.

**Results**. Our 3D model-based algorithm is compared to two particle filter methods using 2D-patches extracted from the first frame. With the first one, the weights are computed with a color histogram distance, while the second one uses a ZNCC score. Both of them show tracking failures after a few frames due to appearance changes. In our algorithm, the likelihood is computed with evolving descriptors adapted to the observations, thus improving the accuracy.

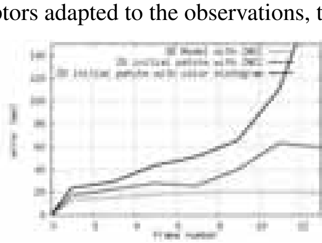

Fig 3. Error with/without 3D model views

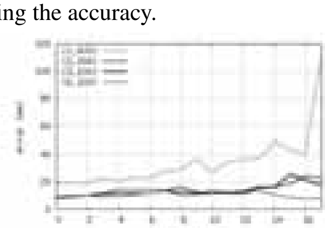

Fig 4. Results with ($N_L$ layers, $N$ particles)

Fig. 4 illustrates the tracking performances for different parameters, all of them needing similar computation times. Multi-layer methods outperform the one-layer method due to an improved state-space exploration. The remaining error is about $1cm$ in the 3D space for the best parametrization.