Automatic identification of facial clues to lies

Tomas Pfister and Matti Pietikäinen

Computers can now detect micro-expressions that reveal emotions people try to conceal.

Micro-expressions are very short involuntary facial expressions that reveal emotions people try to hide (see Figure 1). They can be used for lie detection and are used by trained officials at US airports to detect suspicious behavior. For example, a terrorist trying to conceal a plan to commit suicide would very likely show a very short expression of intense anguish. However, human recognition accuracy is very low; even highly trained human detectors are notoriously inaccurate, achieving a recognition accuracy of only 47%. This performance makes an automatic computer detector very attractive.

We have developed the first system for recognizing real, spontaneous facial micro-expressions. We have developed a temporal interpolation method that enables multiple kernel learning and other machine learning algorithms to classify micro-expressions even with a normal 25 frames per second (fps) camera (see Figure 2 for an illustration of the method). Our system achieves very promising results that compare favorably to human micro-expression recognition accuracy. We also have the first publicly available database of micro-expressions.

The major challenges in recognizing micro-expressions are twofold. The first is involuntariness. How can we get human training data for our algorithm when the expressions are involuntary? We cannot rely on actors as they cannot act out involuntary expressions. The second is short duration. The implication of this is there are a very limited number of frames available using normal cameras, making recognition very challenging.

To obtain samples of involuntary expressions, we conducted an induced emotion suppression experiment where subjects were recorded as they attempted to suppress their facial expressions while watching 16 emotion-eliciting film clips. They were told that experimenters would be watching their face and that if their facial expression leaked so that the experimenter correctly guessed which clip they were watching, they would be asked to fill in a dull 500-question survey. This induced 77 micro-expressions in six subjects, recorded with a camera taking 100fps. To solve the problem of the expressions’ short duration, we devised a temporal interpolation method that interpolates each micro-expression to a larger number of frames. This method enables us to detect micro-expressions even with a standard camera.

To address the large variations in the spatial appearances of micro-expressions, we crop and normalize the face geometry according to the eye positions from a Haar eye detector and the

Continued on next page
Figure 2. An example of a facial micro-expression (above left) being interpolated through graph embedding (above right); the result from which spatiotemporal local texture descriptors are extracted (lower right), enabling recognition with multiple kernel learning.

Figure 3. Spatial normalization of faces. Facial feature points derived from an active shape model are mapped onto a model face.

Figure 4. Temporal interpolation method. The video is mapped onto a curve along which a new video is sampled.

Figure 5. Subject cross-validation results. MKL: Multiple kernel learning. RF: Random forests classifier. TIMn: Temporal interpolation to n frames. SVM: Support vector machine.

In an ideal case, spontaneous micro-expression recognition would work with standard cameras without special hardware. An illustrative comparison of results using different machine learning algorithms shows that TIM enables high recognition accuracy even when using a standard 25fps frame rate. In the detection phase we distinguish micro-expressions from other facial data (see Figure 5). The best results were achieved by combining the ‘random forest’ classifier combined with TIM to 20 frames using 25fps data, yielding 78.9% accuracy.

In summary, we have described the first system to successfully recognize spontaneous facial micro-expressions. We use temporal interpolation to counter short video lengths and an induced emotion suppression experiment to induce spontaneous micro-expressions in human subjects. Our system achieves very promising results that compare favorably to human micro-expression recognition accuracy. Future work includes expanding the micro-expression corpus to more participants, comparing our system to the performance achieved by trained humans on our dataset and enabling real-time recognition.

The authors are grateful to Infotech Oulu and the Academy of Finland for their funding.

Continued on next page
Author Information

Tomas Pfister and Matti Pietikäinen
Machine Vision Group
University of Oulu
Oulu, Finland

Tomas Pfister received a BA in computer science from the University of Cambridge, UK. He also completing his PhD at the Visual Geometry Group at the University of Oxford, UK. His research interest is human-centered computer vision, in particular affect recognition from facial features and sign language recognition.

References