Face Detection and Recognition in Video-Streams

Jannik Boll Nielsen and Rasmus Larsen

DTU-Informatics, Technical University of Denmark, Richard Petersens Plads, Building 321, DK-2800 Lyngby, Denmark

Extended Abstract

In this work we demonstrate how the Viola-Jones face detector [1] can be combined with a person specific active appearance model [3] and used for automated annotation of video streams. Here we understand annotation as identification of the first frame and last frame of a persons face continuous appearance in a video stream, as well as detection, segmentation, and identification of the face in each frame. The Viola-Jones face detector is based on simple sums and differences of image pixels within rectangular areas. A main advantage of these features is that they can be computed in constant time using the so-called integral image in which a pixel value is the sum of all pixel values above and to the left of that pixel in the original image. A boosting classifier is then employed, where each of the weak classifiers in the sequence of classifiers is constrained to be based on a single of the features described above. In each step of the sequence of classifiers the number of false negatives is minimized. In this way a cascade of classifiers is constructed in which at each stage if a sub-windowed is rejected as not being a face no further processing is done, the positives and false positives are send to the next stage. This make for a computationally very efficient face detector. When a face is detected in a sequence of frames within certain motion limits a sub-sequence is defined. From this sequence a series of frames uniformly sampled over time are selected for segmentation using a person-specific active appearance model. The active appearance model uses a face model based on a truncated principal component model of combined variation of shape as defined by a set of and texture as defined by the sampled intensity values across the face.

We demonstrate the face detection and recognition scheme on a series of Danmarks Radio (Danish Broadcasting Corporation) game shows featuring actor and talk show host Jarl Friis Mikkelsen. In a series of video streams face sub-sequences are successfully detected and classified as being either the actor in question or not. The computational implementation is based on the OpenCV implementation of the Viola-Jones face detector and the publically available active appearance model software from DTU Informatics [3].

1. P. Viola and M. Jones. Robust real-time face detection. International Journal of Computer Vision (IJCV) 57(2) 137-154, 2004.
2. M. B. Stegmann, B. K. Ersbøll, and R. Larsen. FAME – a flexible appearance modelling environment. IEEE Trans. on Medical Imaging, 22(10):1319-1331, 2003
3. T.F.Cootes, G.J. Edwards and C.J.Taylor. "Active Appearance Models", IEEE PAMI, Vol.23, No.6, pp.681-685, 2001