Privacy Protecting, Real-time Face Re-recognition

T. Niederberger[1], R. Hegner[1], A. Hartmann[2] and G. M. Schuster[1]

[1]HSR Hochschule für Technik Rapperswil, Switzerland
[2]ADEC Technologies AG Eschenbach, Switzerland

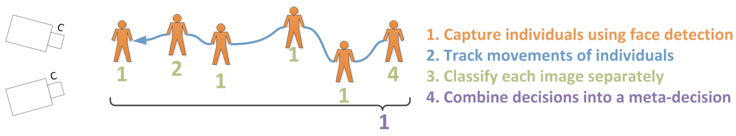

Figure 1: *Solution strategy for face re-recognition*

Abstract

We present a novel system for recognizing human individuals walking past a depth camera that is compatible with privacy protecting laws. The system is developed to support the statistical analysis of movement patterns in indoor spaces. The system is able to re-recognize previously seen individuals but is also capable of recognizing that an individual has not been seen before. The system is designed in a privacy protecting way and does not rely on previously collected training data but rather collects data during run-time. The proposed system processes each image of an individual separately, but we also present a new approach that is based on combining several decisions into a single meta-decision in order to enhance classification performance.

Categories and Subject Descriptors (according to ACM CCS): I.4.9 [Image Processing and Computer Vision]: Applications—

1. Introduction

Recognizing people by means of a camera has long been a goal in the scientific community. In particular face recognition has received significant attention by researchers and a wide variety of approaches have been proposed [ZCPR03]. In this work we propose a new system for re-recognizing individuals walking by a depth camera. With the term "re-recognizing" we want to emphasize that the system is not designed for large scale face recognition but rather for re-recognizing individuals that passed an associated camera shortly before. Such an approach is less privacy invasive than "traditional" face recognition. Therefore the system can be used in sensitive places where a face recognition system is not allowed due to privacy protecting laws. The system does not rely on data that was collected and processed offline. Instead, data treatment is done during run-time: collection, preprocessing, feature extraction and training. We present a system design and a processing pipeline implementation for such a system. We also present preliminary data and discuss future improvements.

2. Main contribution

Our main contribution is a system for face re-recognition that protects the privacy of an individual. The system can track the movement of people and runs in real-time. Furthermore, the system is able to identify new individuals and learn during run-time. Finally, we describe how multiple decisions can be combined into an improved meta-decision.

3. System design and privacy

In designing the system we put a high emphasis on privacy protection. For this reason we store the data for classification (depth data) only in a transient memory and delete it after a couple of minutes. In persistent memory we store only abstract movement descriptions and hence it is not possible to identify a person at a later point in time.

DOI: 10.2312/egp.20151034

4. Classification approach

The classification approach is guided by the idea of capturing many pictures from an individual in only a few seconds. Each image is processed separately and classified by a classification algorithm. The classifications from all the images are then combined into a single meta-decision (figure 1). We put a high emphasis on speed and the system is capable of processing more than 20 frames per second on a standard personal computer (Intel Xeon quad-core 64 bit, 3.2GHz).

We start by applying a face detection algorithm on all of the foreground regions. We use a trained cascade classifier [VJ01] on depth data. For each detected face we separate it from the background using thresholding and locate the nose. The detected nose is then used to establish a new coordinate system. In the new coordinate system the face is registered to a predefined standard face using the iterative closest point (ICP) algorithm. In a parallel process the detected faces and foreground regions are assigned to tracks such that we can reconstruct where an individual was located at a specific time.

After preprocessing we start the feature extraction procedure. Our system is mainly based on distance measurements that can be measured easily once the face region has been identified. Features such as the face profile, the head size and the tallness have proven to be among the most useful features for the task at hand. Well known features from the literature (such as histograms of oriented gradients or local binary patterns) were also tested but could not improve the performance of the system. After feature extraction we classify the image using a random forest classifier [Bre01]. A separate classifier is trained for each individual (one-vs-all approach) and each image is tested against every classifier. If several classifiers signal on a specific image the class with the highest confidence is selected. If none of the classifiers accepts an image the image is assumed to represent a new individual.

The system is tracking the movement of a person over several frames in order to facilitate combining several decisions to a meta decision. The tracking is essential because otherwise it is not possible to know which face stems from which individual when several individuals are in the field of view. The meta decision is generated by taking a majority vote of all images assigned to an individual. If the majority of the decisions concludes that the individual has not been seen before a new classifier is trained, the individual is assigned to that class and the classifier is added to the system.

5. Experimental results

In order to show the performance of our system we collected a data set of 19 individuals that we use for training and validation. Figure 2 shows the performance of the system depending on the number of classes (individuals). As expected the accuracy decreases when the number of individuals increases; however the drop in performance is rather small. With 19 individuals we achieve an average recognition performance of 67%. Furthermore these results are for a single image and figure 3 shows that combining several images to a meta-decision leads to a significant increase in performance.

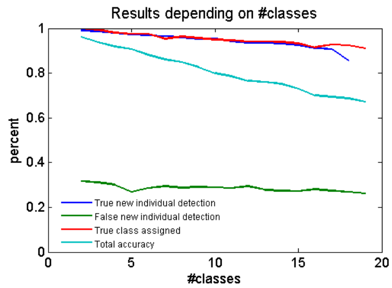

Figure 2: *Classification performance depending on the number of classes (individuals)*

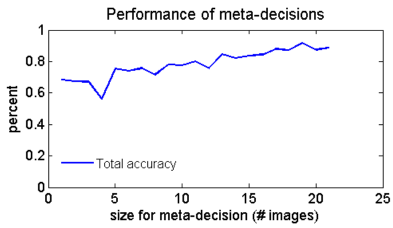

Figure 3: *Classification accuracy depending on how many images are combined into a meta-decision. Accuracy increases from 67% to 88% when using 21 images.*

6. Future work

In the future we will introduce a new algorithm for face tracking. This is expected to improve face registration performance and also provide additional features for classification. We will also collect a larger and more diverse data set in order to further verify the performance of the system. An enhanced system will support multiple cameras that capture an overlapping field of view. This will increase the amount of data available for learning and classification.

References

[Bre01] BREIMAN L.: Random forests. *Machine learning 45*, 1 (2001), 5–32. doi:10.1023/a:1010933404324. 2

[VJ01] VIOLA P., JONES M.: Rapid object detection using a boosted cascade of simple features. In *Proc. CVPR '01* (2001), vol. 1, pp. 511–518. doi:10.1109/CVPR.2001.990517. 2

[ZCPR03] ZHAO W., CHELLAPPA R., PHILLPS P., ROSENFELD A.: Face recognition: A literature survey. *ACM Computing Surveys 35*, 4 (Dec. 2003), 399–458. doi:10.1145/954339.954342. 1