On-line Learning a Person Model from Video Data

Peter M. Roth
Graz University of Technology
Inffeldgasse 16/II, 8010 Graz, Austria
Email: pmroth@icg.tu-graz.ac.at

Horst Bischof
Graz University of Technology
Inffeldgasse 16/II, 8010 Graz, Austria
Email: bischof@icg.tu-graz.ac.at

Theoretical Background

Recently we have proposed the Conservative Learning Framework [2] for unsupervised learning a person detector from video data. The main idea is to minimize the manual effort when learning a classifier and to combine the power of a discriminative classifier with the robustness of a generative model. Starting with motion detection an initial set of positive examples is obtained by analyzing the geometry (aspect ratio) of the motion blobs. If a blob fulfills the restrictions the corresponding patch is selected. Negative examples are obtained from images where no motion was detected. Using these data sets a first discriminative classifier is trained using an on-line version of AdaBoost [1]. In fact, applying this classifier all persons are detected (we got a quite general model) but there is a great number of false positives. Thus, we apply a generative classifier, i.e., robust PCA [3] to verify the obtained detections and to decide if a detected patch represents a person or not. As we have a huge amount of data (video stream) we can be very conservative and use only patches for (positive or negative) updates if we are very confident about our decision! Thus, most of the patches are not considered at all. Applying this update rules an incrementally better classifier is obtained without any user interaction needed!

Description

To demonstrate the functionality and the power of the approach the video is subdivided in several parts. First, the method is motivated by showing a typical scene (a corridor in a public building) that may be used for learning. Next, the main step, the on-line learning, is demonstrated. Therefore, detection results obtained by applying different classifiers that where estimated during the learning process are compared. Whereas for the initial classifier a great number of false positives is detected the final one yields a perfect detection result. Examples of incrementally better detection results during the learning process are shown in Fig. 1. In addition, the learning process is demonstrated. Thus, the update rules described above are visualized using different colors for the bounding boxes of the detections: green - positive update, red - negative update, white - detection, but no update.

Finally, the classifiers obtained by Conservative Learning are evaluated on different sequences. In addition, it is shown that a classifier that was trained on one setup can be re-trained and applied for a detection task on a completely different setup. Examples of further results shown in the video are depicted in Fig. 2. As can be seen from Fig. 2(c) the same framework can also be applied for learning, e.g., a car detector.

(a) (b) (c)

Fig. 1
Conservative Learning: (a) initial classifier, (b) 100 frames processed, (c) 1200 frames processed.

(a) (b) (c)

Fig. 2
Detection results obtained by classifier learned by Conservative Learning.

Acknowledgments

This work has been supported by the Austrian Joint Research Project Cognitive Vision under projects S9103-N04 and S9104-N04 and by the Austrian Federal Ministry of Transport, Innovation and Technology under P-Nr. I2-2-26p VITUS2. In addition this work has been sponsored by the MISTRAL Project which is financed by the Austrian Research Promotion Agency (www.ffg.at) and the EU FP6-507752 NoE MUSCLE IST project.

References

[1] Helmut Grabner and Horst Bischof. On-line boosting and vision. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, 2006. to be published.

[2] Peter Roth, Helmut Grabner, Danijel Skočaj, Horst Bischof, and Aleš Leonardis. On-line conservative learning for person detection. In *Proc. IEEE Workshop on VS-PETS*, pages 223–230, 2005.

[3] Danijel Skočaj and Aleš Leonardis. Weighted and robust incremental method for subspace learning. In *Proc. IEEE Intern. Conf. on Computer Vision*, volume II, pages 1494–1501, 2003.