

11/2014

Virtual and Real World Adaptation for Pedestrian Detection



Pedestrian detection is currently a key component of systems present in many markets such as the automotive, surveillance and multimedia industries. However, this application involves a tiresome collection and annotation of large amount of data required for its training. Therefore, we propose a novel method to reduce the efforts involved in the mentioned training process by the use of synthetic data that is automatically collected from a video game. Our method allows the adaptation of a detector model trained with synthetic data to successfully work with real world data.

Pedestrian detection is of paramount interest for many applications, e.g., Advanced Driver Assistance Systems, Intelligent Video Surveillance and Multimedia systems. Most promising pedestrian detectors rely on appearance-based classifiers trained with annotated data, i.e., images where the location of the pedestrians is annotated by setting a box around each of them. However, the required annotation step represents an intensive and boring task for humans, making it worth to minimize their intervention in this process by using computational tools like realistic virtual worlds, i.e., video games. The reason for using these kinds of tools lies in the fact that they allow for the automatic generation of precise and rich annotations of visual information.

Nevertheless, the use of this kind of data comes with the following question: can a pedestrian appearance model learnt with virtual-world data work successfully for pedestrian detection in real-world scenarios? To answer this question, we conducted different experiments that suggest a positive answer.

We found that the pedestrian detectors trained with virtual-world data do not perform as well as the real-world based ones. This problem is called dataset shift and happens even when training with a real-world model of a specific domain (e.g., beach) and then using it in a different domain (e.g., mountain). Accordingly, we have designed different domain adaptation techniques to face this problem; all of them are integrated into one same framework (V-AYLA). We have explored different methods to train a domain adapted pedestrian detector by collecting a few pedestrian samples from the target domain (real world) and combining them with many samples of the source domain (virtual world). The extensive experiments we present show that pedestrian detectors developed within the V-AYLA framework do achieve domain adaptation.

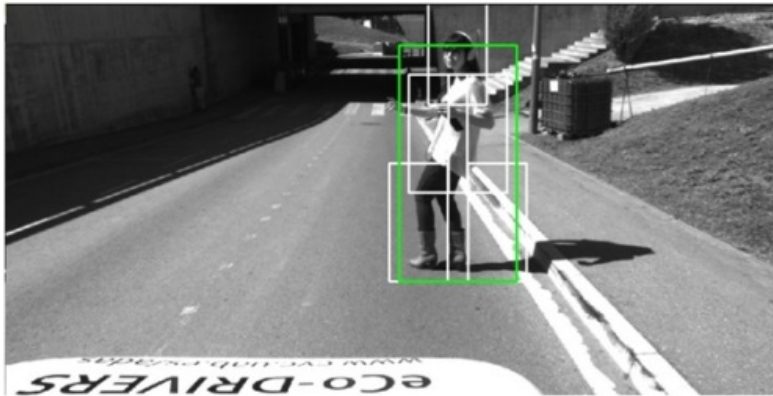


Figure: Video frame of a virtual-world sequence (left) with its corresponding automatic annotation (right).

The results presented on this work not only end with a proposal for adapting a virtual-world pedestrian detector to the real world, but also it goes further by pointing out a new methodology that would allow the system to adapt to different situations, which we hope will provide the foundations for future research in this unexplored area.

This research has been done at the Advanced Driver Assistance Systems (ADAS) group at the Computer Vision Center (CVC). This work has been supported by Spanish MICINN projects TRA2011-29454-C03-01 and TIN2011-29494-C03-02.

Top left figure: Video frame of our virtual-world pedestrian detector adapted for working in the real world.

David Vázquez

dvazquez@cvc.uab.es

References

Vázquez, David; Marín, Javier; López, Antonio M.; Ponsa, Daniel; Gerónimo, David. Virtual and Real World Adaptation for Pedestrian Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence 34(4): 797–809. 2014. doi: 10.1109/TPAMI.2013.163.

Xu, Jiaolong; Ramos, Sebastian; Vázquez, David; López, Antonio. Domain Adaptation of

Deformable Part-Based Models. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2014. doi: 10.1109/TPAMI.2014.2327973.

[View low-bandwidth version](#)