

UCSD Pedestrian Dataset

Antoni Chan (`abchan - AT - ucsd.edu`)
Department of Computer Science, City University of Hong Kong

Nuno Vasconcelos (`nuno - AT - ece.ucsd.edu`)
Dept. of Electrical and Computer Engineering, University of California, San Diego

Feb 28, 2013

This is the UCSD pedestrian dataset used in [1] for motion segmentation, and [2, 3, 4] for crowd counting. If you use this dataset, please reference [1] and/or [4].

1 Dataset Video Format

The dataset contains video of pedestrians on UCSD walkways, taken from a stationary camera. There are currently two viewpoints available. All videos are 8-bit grayscale, with dimensions 238×158 at 10 fps. The original video is 740×480 at 30 fps, and is available on request.

The `video` directory contains the videos for the two scenes. Each scene is in its own directory `vidX` where `X` is a letter (e.g. `vidf` and `vidd`), and is split into video clips of length 200 named `vidXY_33_ZZZ.y`, where `Y` is the video number and `ZZZ` is the clip number. The clips for each video are continuous, and there may be a small jump when moving between videos. Finally, each video clip is saved as a set of `.png` files, with file name `vidXY_33_ZZZ_fFFF.png`, where `FFF` is the frame number. Examples from each scene are presented in the left column of Figure 1.

2 Motion Segmentation Experiments

The `vidf` scene was used in the motion segmentation experiment of [1]. In particular, `vidf1_33_000.y` and `vidf1_33_007.y` are the “sparse traffic” and “heavy traffic” pedestrian scenes in Figures 11a and 11c of [1]. In addition, `vidf1_33_007.y` was used to train the model for segmenting the remaining `vidf` video. These results are available online [5, 6].

3 Motion Segmentation for Crowd Counting

The motion segmentations used for crowd counting in [2, 4] are located in the `segm` directory. The segmentation files use the same naming convention as the videos, except the extension is now `.segm` instead of `.y`. The `vidf` scene was segmented into 2 motions, while the `vidd` scene was segmented into 4 motions. Table 1 presents the semantic meaning for each of motion segments and their grayscale value in the segmentation image. Example segmentations are in the right column of Figure 1.

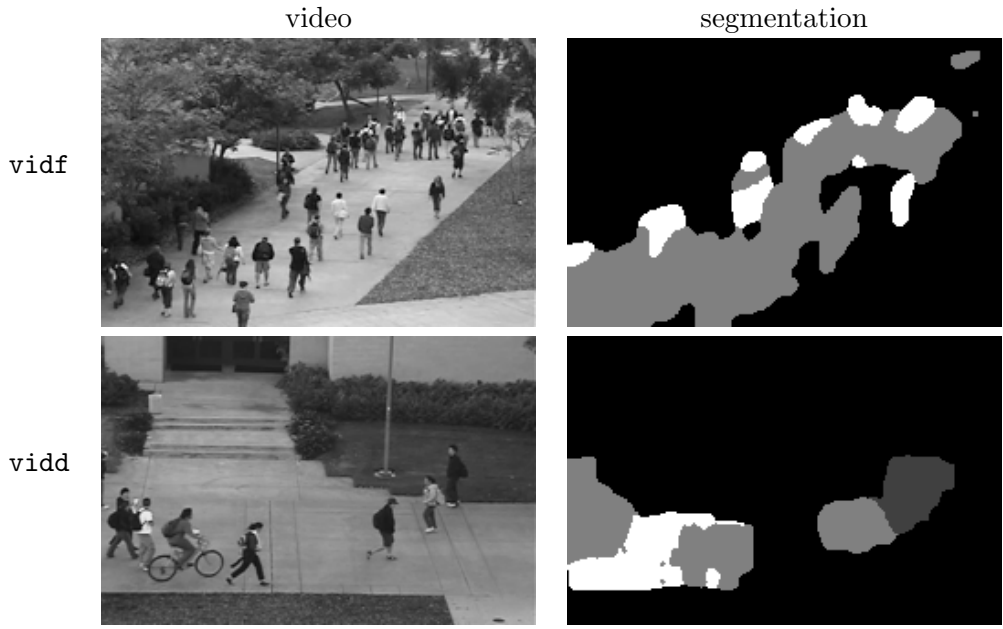


Figure 1: Example frames and segmentations from the dataset.

scene	grayscale values
vidf	0=no motion, 128=right (away), 255=left (towards)
vidd	0=no motion, 64=left slow, 128=right slow, 196=left fast, 255=right fast

Table 1: Motion segmentation values

4 Acknowledgments

The collection of this database was funded by US National Science Foundation Award IIS-0534985 and CCF-0830535, NSF IGERT Award DGE-0333451, and by the Research Grants Council of Hong Kong Special Administrative Region of China under grant CityU 110610 (9041552). Part of the ground-truth was annotated by Jeffrey Cuenco and Zhang-Sheng John Liang.

5 History

- 2012/02/28 - added vidd, and segmentations.
- 2008/05/22 - initial version (vidf only)

References

- [1] A. B. Chan and N. Vasconcelos, “Modeling, Clustering, and Segmenting Video with Mixtures of Dynamic Textures,” *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 30(5), pp. 909-926, May 2008.
- [2] A. B. Chan, Z. S. J. Liang, and N. Vasconcelos, “Privacy Preserving Crowd Monitoring: Counting People without People Models or Tracking,” In *IEEE Conference on Computer Vision and Pattern Recognition*, June 2008.

- [3] A. B. Chan and N. Vasconcelos, "Bayesian Poisson Regression for Crowd Counting," In *IEEE Intl Conf. on Computer Vision*, Kyoto, Sept 2009.
- [4] A. B. Chan and N. Vasconcelos, "Counting People with Low-Level Features and Bayesian Regression," *IEEE Trans. on Image Processing*, vol. 21(4), pp. 2160-77, April 2012.
- [5] <http://www.svcl.ucsd.edu/projects/motiondytex>
- [6] <http://visal.cs.cityu.edu.hk/research/motiondytex/>