SABS - Stuttgart Artificial Background Subtraction Dataset

The SABS dataset was created for pixel-wise evaluation of the performance of background models for background subtraction. It covers 9 typical challenges of background subtraction that occur in the context of video surveillance. By using artificially generated data, we were able to provide high quality segmentation ground-truth annotation.

The dataset is publicly available on the project page:
http://www.vis.uni-stuttgart.de/index.php?id=sabs

If you publishing research based on the SABS dataset, please include a citation to:

```
@inproceedings {cvpr11brutzer,
    author = {{Brutzer, S. and H{\"o}ferlin, Benjamin and Heidemann, Gunther}},
    title = {{Evaluation of Background Subtraction Techniques for Video Surveillance}},
    year = {{2011}},
    booktitle = {{Computer Vision and Pattern Recognition (CVPR)}},
    pages = {{1937-1944}},
    publisher = {{IEEE}}
}
```

Experimental Setup

The experimental setup for the evaluation of background models with respect to the 9 typical challenges listed in the paper is specified in table 1. Further details on the background subtraction methods and their parameters we used for evaluation can be found in the paper.

To account for foreground/background blending due to image rasterization, we use a "*don't care*" border around the ground-truth foreground objects. Pixels on this border are not considered for evaluation. The "*don't care*" border is simply calculated from b/w ground-truth images (threshold at 0.001 for pixel values scaled between 0..1; alpha channel is ignored) by applying a morphological *erode* operation with a square structuring element of 3 pixels length. All pixels that left unchanged by this operation are considered for evaluation.

For the evaluation of shadow removal capabilities of background models, we only considered shadow-pixels with a luminance distance of more than five.

Table 1: Setup of the conducted experiments: Note that frame indices start from 0, while pixel indices begin with the index 1. The start and end pixels of the cropped test area are both included in the evaluation.

Experiment	Training Data	Training Frames	Test Data	Test Frames	Cropped Test Area (x_1/y_1):(x_2/y_2)	Ground-Truth Frame Range (Shadow Mask Range)
Basic	NoForegroundDay	0:800	Basic	0:599	-	801:1400
Bootstrap	Bootstrap	0:0	Bootstrap	1:1400	-	1:1400
Dynamic Background	NoForegroundDay	0:800	Basic	0:599	(100/200):(380/560)	801:1400
Light Switch	NoForegroundNight	0:800	LightSwitch	0:599	-	801:1400
Darkening	NoForegroundDay	0:800	Darkening	0:1399	-	1:1400
Shadow	NoForegroundDay	0:800	NoCamouflage	0:599	-	1:600 (0:599)
Camouflage	NoForegroundDay	0:800	Camouflage	0:599	(400/150):(600/500)	1:600
NoCamouflage (used for comparision only)	NoForegroundDay	0:800	NoCamouflage	0:599	(400/150):(600/500)	1:600
Noisy Night	NoForegroundNightNoisy	0:800	NoisyNight	0:599	-	801:1400
Video Compression (MPEG4)	MPEG4_40kbps MPEG4_80kbps MPEG4_160kbps MPEG4_320kbps MPEG4_640kbps	0:800	MPEG4_40kbps MPEG4_80kbps MPEG4_160kbps MPEG4_320kbps MPEG4_640kbps	0:599	-	801:1400

Evaluation Measures

In the paper we use binary classification measures for evaluation, such as Precision, Recall and F1-Score. However, we did not mention that we calculate these measures per sequence instead of per frame. A sequence-based measure was used to circumvent issues with division-by-zero, if a single frame does not contain any foreground objects. In the unlikely case of an empty foreground pixel set (summed over a whole sequence), we suggest to set precision to 1, since no irrelevant “objects” have been retrieved.

In detail, calculation of precision and recall requires prior summation of all classification events (true positively classified pixels, etc.) over the whole sequence.