# Supplementary Material

Chen-Ping Yu[1] Hieu Le[1] Gregory Zelinsky[1,2] Dimitris Samaras[1]

Computer Science Department[1], Psychology Department[2]
Stony Brook University, Stony Brook NY 11790
{cheyu, hle, samaras}@cs.stonybrook.edu[1], gregory.zelinsky@stonybrook.edu[2]

## 1. Run time and Accuracy of the Sub-sampling effect

In this section we offer additional evidence of the efficiency and robustness of our method, by evaluating an important advantage of PGP: It allows subsampling only a subset of the edges of the graph in order to estimate the critical values that define whether two super-pixels are similar or dissimilar without any drop in performance. As we have discussed in Sec. 3.1 of the main paper, the method fits WMM over the distribution of edge weights where the $L_p$ distance statistics are shown to be Weibull distributed in theory. Therefore we can also work with a subset of the edge weights (by uniformly random sampling) and the inherent distribution would remain Weibull, while the WMM fit should (in theory) remain similar. This is especially useful for larger videos with HD frames and/or longer duration, where the number of edges in the spatio-temporal graph is very large, and by fitting WMM over a subset of those edges would offer even higher computational efficiency.

To evaluate the effect of sub-sampling on segmentation and run-time performance, we conducted the following experiment: we use the setting of NLS optimization with motion feature with 400 superpixels per frame as the base-line PGP setting, and we progressively sub-sample at 100%, 75%, 50%, and 25% of the total number of edges. For instance, sample rate of 25% means using a random subset of one quarter of the edges for fitting the WMMs. Note that these sub-sampling rates from 400 superpixels exactly correspond to the number of edges if starting with 100, 200, or 300 superpixels per frame, thereby allowing a direct comparison with Figure 3 and 4 of the main manuscript.

The results of the experiment are shown in Figure 1, where the solid lines are the sub-sampled results (ss) that correspond to the sub-sample rate of the x-axis, and the dotted lines are the full-sample results (fs) that correspond to the number of superpixels (in parenthesis) of the x-axis labels. The plots of the segtrackV2 dataset indicate sub-sampling barely affects any of the measured segmentation performance as compared to the base-line setting, while the fluctuation is more pronounced for Chen's dataset ojust for the case of 1/8 sub-volume processing. The result is especially apparent in the segtrackV2 plots, where sub-sampling results in much higher 3d segmentation performance than starting at a lower number of superpixels across the four quantitative metrics. This result suggests that one can segment to a higher number of initial superpixels in order to achieve better boundary preservation, and then still use only one quarter of the data for optimization to achieve the same level of accuracy at approximately 50% of the original run time.

| Sampling Rate | Full (1.0) | 0.75 | 0.5 | 0.25 |
|---|---|---|---|---|
| Run time (sec) | 345.02 | 279.5 | 221.6 | 168.9 |

Table 1. Run-time analysis on the effect of sub-sampling on a 85-frame video (Chen's dataset), the base-line (full) setting is using 400 initial superpixels per frame, 1/4 sub-volume processing, NLS optimization with motion feature. The sub-sampling allows PGP to further reduce run-time by over 50%, while maintaining the segmentation performance of the base-line setting (see Figure 1). The experiment is done on a Xeon X3470 at 2.93 Ghz with 32 Gb. The maximum memory usage was recorded at 4 Gb.

Table 1 shows the run time analysis on performing sub-sampling at the levels of 0.75, 0.5, and 0.25 as compared to full use of the data (1.0), excluding the superpixel extraction time which takes about 1 second per frame. The run time decreased by more than 50% from the original full processing mode when working with one quarter of the edge weights. In our implementation, the result of sub-sampling is deterministic as we compute the linearly spaced vector from the full set of edges, such that the sub-sampled edges are the edges at linearly equally spaced points between edge 1 and edge n, where n is the ending edge of the graph. This is done to eliminate any randomness of the algorithm that may be undesirable for practical applications.

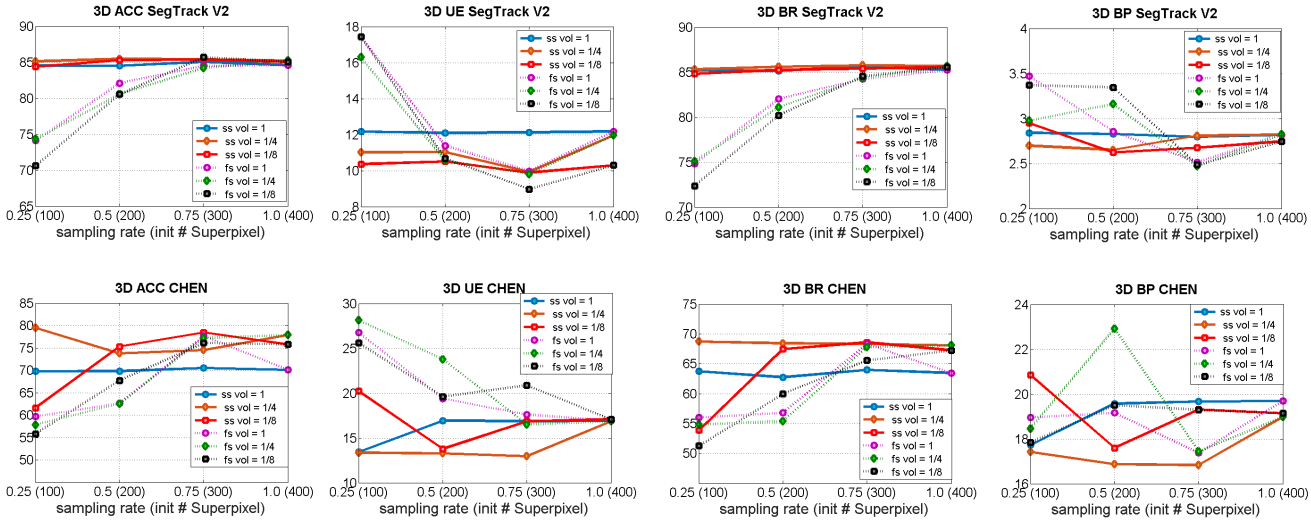


Figure 1. Effect of sub-sampling with respect to 3D segmentation performance. The baseline setting being used here is 400 initial superpixels per frame, 1/4 sub-volume processing, NLS optimization with motion feature (red dotted line). The dotted lines are without any sub-sampling (fs) at 100, 200, 300, and 400 superpixels per frame (x-axis labels in the paranthesis); the solid lines are with sub-sampling at 0.25, 0.5, 0.75 and 1.0 from the edge sets of the baseline. The effects are also analyzed under different sub-volume processing (full (1), 1/4, and 1/8). Note that sub-sampling at 0.25 rate of the original 400 superpixel corresponds to running non sub-sampling on 100 superpixels per frame, therefore making the solid lines and the dotted lines directly comparable. The plots show that sub-sampling barely affects the performance in segtrackV2 dataset, and only observing an effect for the 1/8 sub-volume processing setting in Chen's dataset.