

# A Comprehensive Performance Evaluation of Deformable Face Tracking “In-the-Wild”

Grigorios G. Chrysos · Epameinondas Antonakos\* ·  
Patrick Snape\* · Akshay Asthana · Stefanos Zafeiriou

Received: date / Accepted: date

## 1 Introduction

In the supplementary material, we have included all the curves as derived per detector/tracker. We follow the same experimental sections as in the main paper and include those additional curves. Since for each method all 4 landmark localisation techniques are provided, there is a single unified legend for all the figures for each landmark localisation method, given in 1.



Fig. 1: The common legend for all the landmark localisation techniques used in the paper.

Additionally we have prepared a video per experiment, illustrating the visual comparison of the top 5 performing methods in each category. The links are presented in the Table 1.

---

\* E. Antonakos and P. Snape contributed equally and have joint second authorship.

G. Chrysos · E. Antonakos · P. Snape · S. Zafeiriou  
Department of Computing, Imperial College London, 180 Queen’s Gate, London SW7 2AZ, UK  
E-mail: {g.chrysos, e.antonakos}@imperial.ac.uk  
          {p.snape, s.zafeiriou}@imperial.ac.uk

A. Asthana  
Seeing Machines Ltd., Level 1, 11 Lonsdale St, Braddon, ACT, Australia, 2612  
E-mail: a.asthana@seeingmachines.com

Experiment 1	<a href="https://www.youtube.com/watch?v=6bzgmsWgK20">https://www.youtube.com/watch?v=6bzgmsWgK20</a>
Experiment 2	<a href="https://www.youtube.com/watch?v=peQYzqgG2UA">https://www.youtube.com/watch?v=peQYzqgG2UA</a>
Experiment 3	<a href="https://www.youtube.com/watch?v=RXo9hZAaQVQ">https://www.youtube.com/watch?v=RXo9hZAaQVQ</a>

Table 1: Links for the videos illustrating the outcomes of the top performing methods in each experiment.

Each image of a video is composed of the original video frame, plus a cropped result for each of the top 5 methods per category. A black frame signifies a missing detection, as shown for HOG-SVM in Fig. 2.

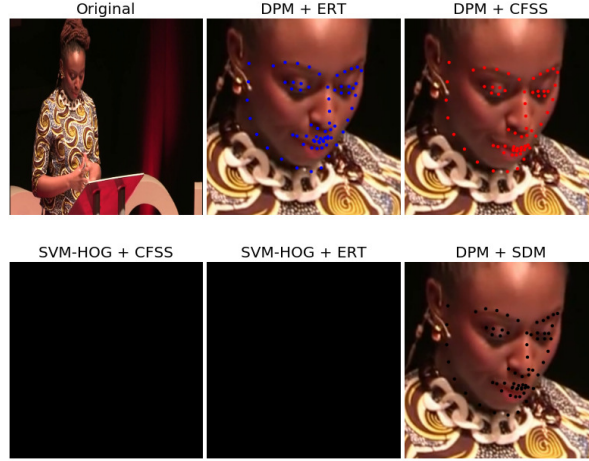


Fig. 2: Arbitrary frame from the supplementary videos.

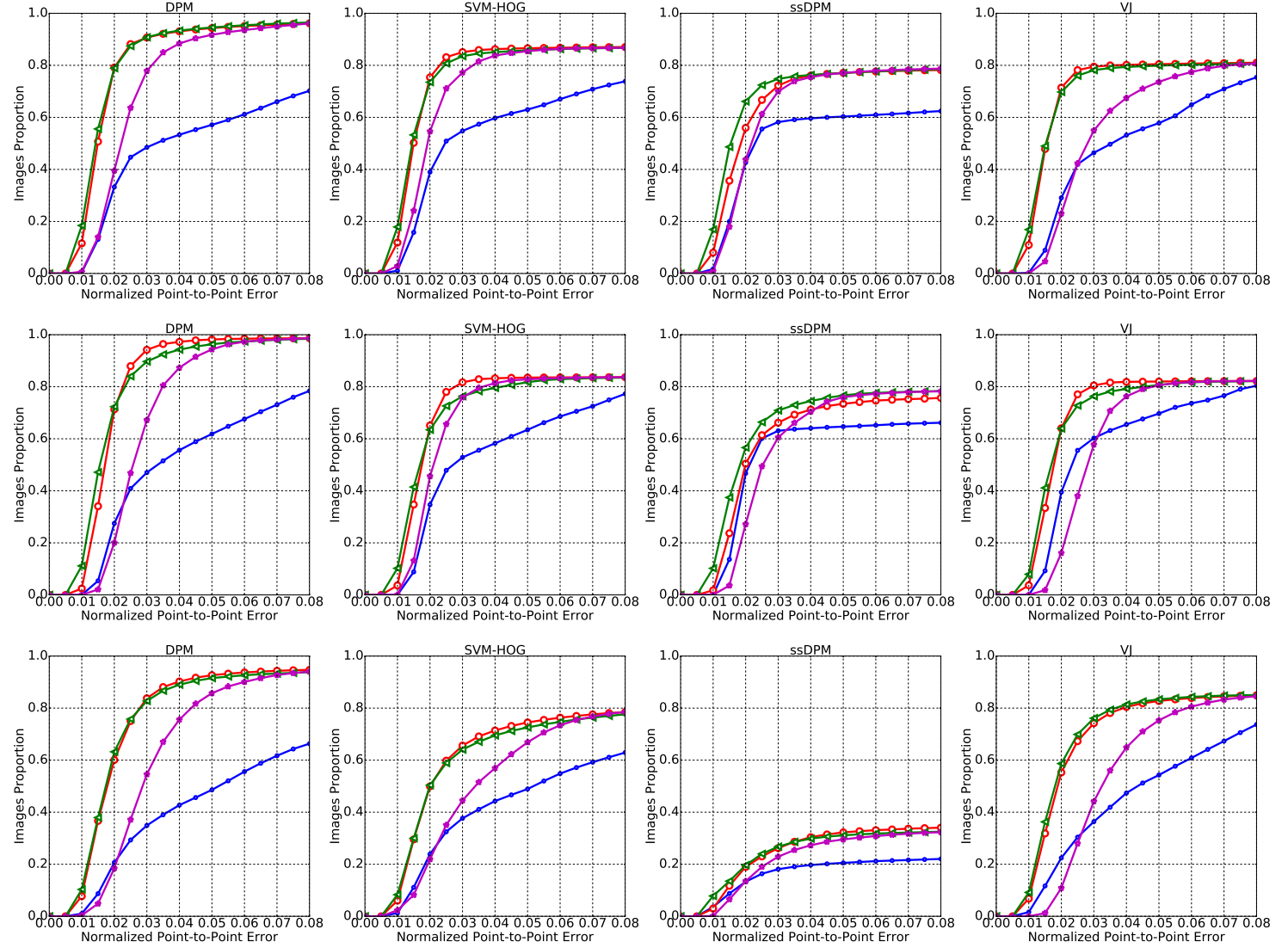


Fig. 3: This resembles the full set of curves as produced in the *Experiment 1* of the main paper. Rows are Category 1, 2 and 3 respectively. Each column is labelled by the figure header representing the work of [Felzenszwalb et al(2010)Felzenszwalb, Girshick, McAllester, and Ramanan], [King(2015)], [Zhu and Ramanan(2012)] and [Viola and Jones(2004)] respectively.

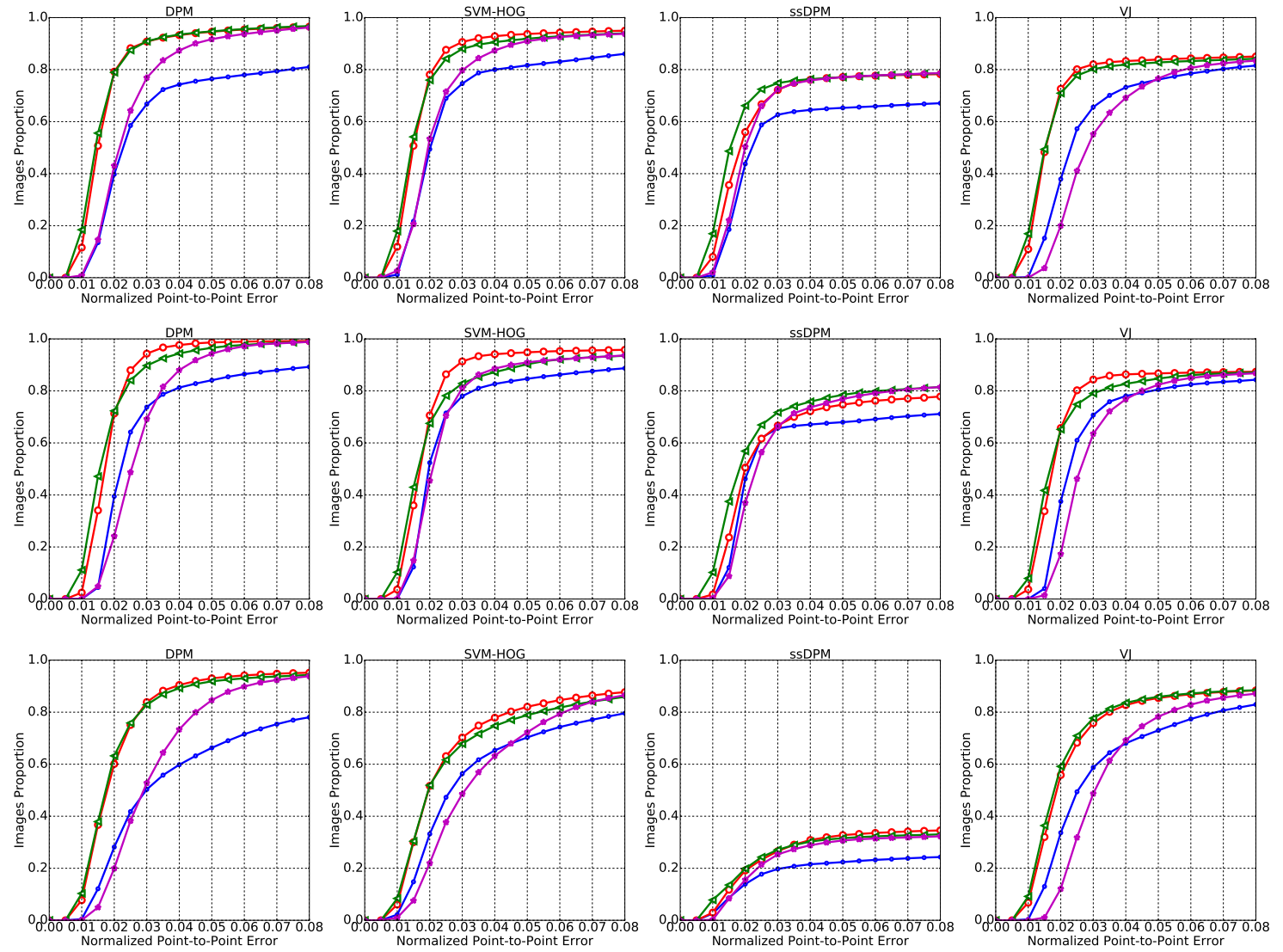


Fig. 4: This resembles the full set of curves as produced in the *Experiment 2* of the main paper. Rows are Category 1, 2 and 3 respectively. Each column is labelled by the figure header representing the work of [Felzenszwalb et al(2010)Felzenszwalb, Girshick, McAllester, and Ramanan], [King(2015)], [Zhu and Ramanan(2012)] and [Viola and Jones(2004)] respectively.

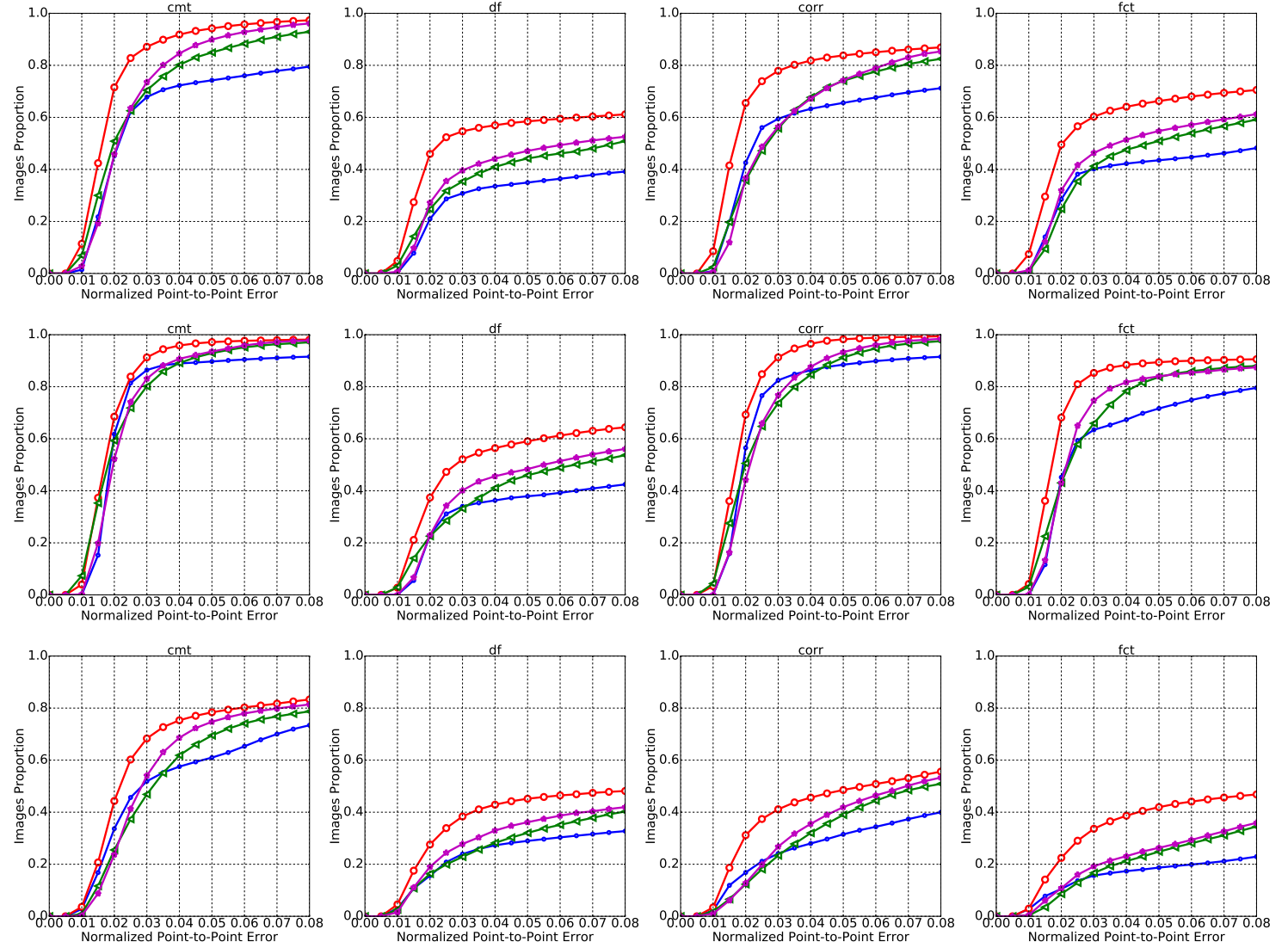


Fig. 5: This resembles the first part of the full set of curves as produced in the *Experiment 3* of the main paper. Rows are Category 1, 2 and 3 respectively. Each column is labelled by the figure header representing the work of [Nebehay and Pflugfelder(2015)], [Sevilla-Lara and Learned-Miller(2012)], [Danelljan et al(2014)Danelljan, Häger, Khan, and Felsberg] and [Zhang et al(2014a)Zhang, Zhang, and Yang] respectively.

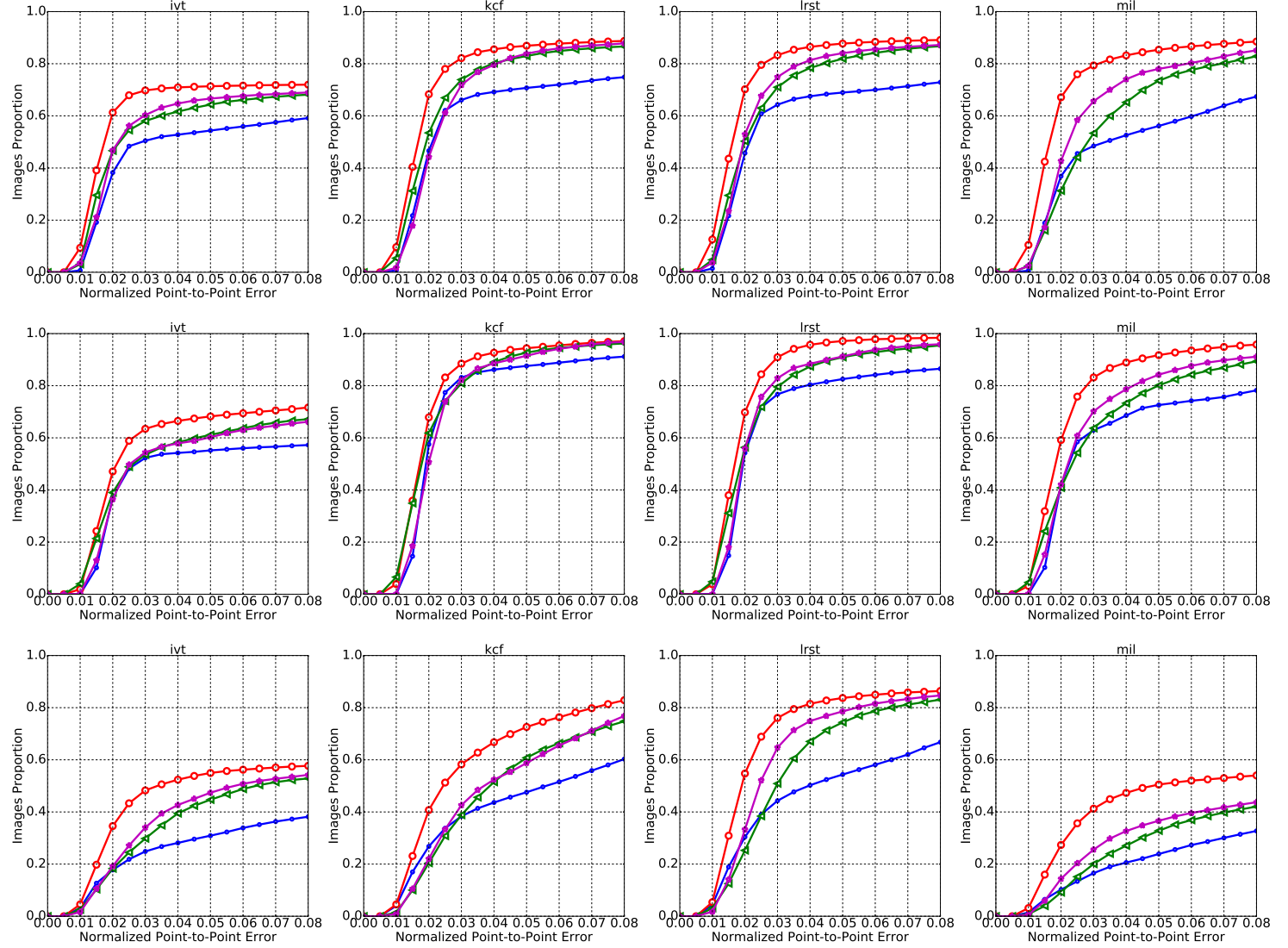


Fig. 6: This resembles the second part of the full set of curves as produced in the *Experiment 3* of the main paper. Rows are Category 1, 2 and 3 respectively. Each column is labelled by the figure header representing the work of [Ross et al(2008)Ross, Lim, Lin, and Yang], [Henriques et al(2015)Henriques, Caseiro, Martins, and Batista], [Zhang et al(2014b)Zhang, Liu, Ahuja, Yang, and Ghanem] and [Babenko et al(2011)Babenko, Yang, and Belongie] respectively.

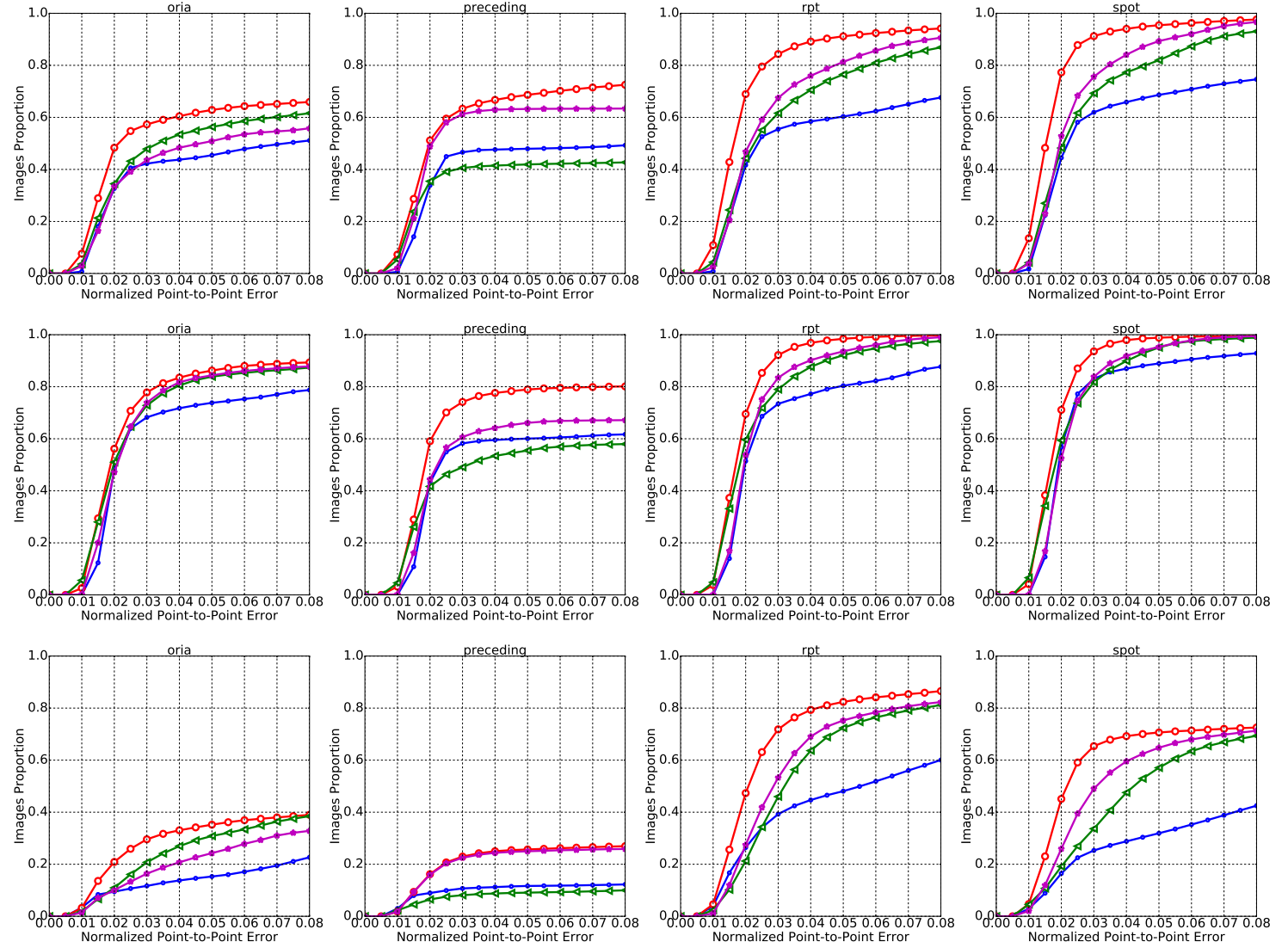


Fig. 7: This resembles the third part of the full set of curves as produced in the *Experiment 3* of the main paper. Rows are Category 1, 2 and 3 respectively. Each column is labelled by the figure header representing the work of [Wu et al(2012)Wu, Shen, and Ling], the baseline tracker (denoted as PREV in the main paper), [Li et al(2015)Li, Zhu, and Hoi] and [Zhang and van der Maaten(2014)] respectively.



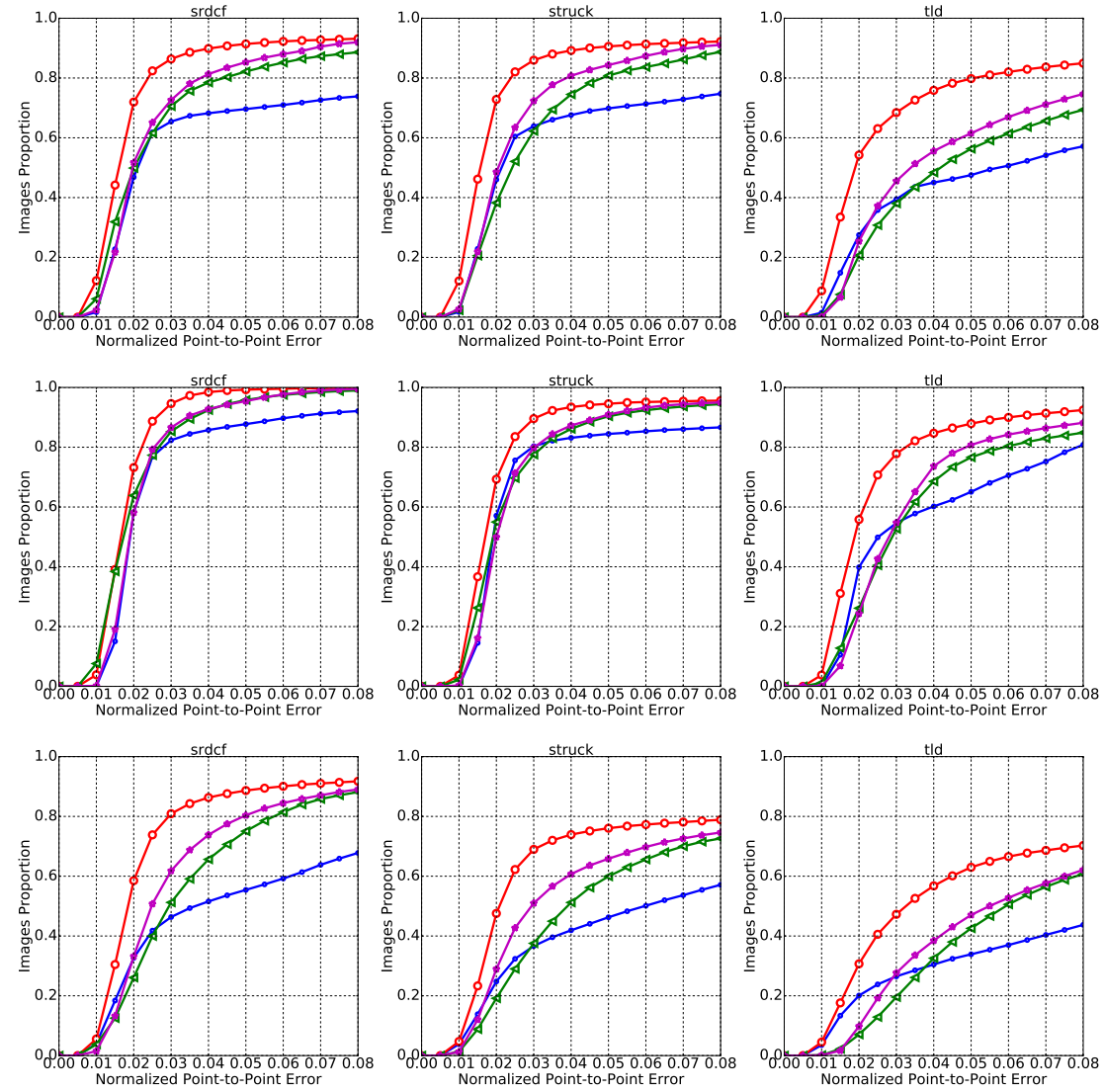


Fig. 8: This resembles the fourth part of the full set of curves as produced in the *Experiment 3* of the main paper. Rows are Category 1, 2 and 3 respectively. Each column is labelled by the figure header representing the work of [Danelljan et al(2015)Danelljan, Häger, Shahbaz Khan, and Felsberg], [Hare et al(2011)Hare, Saffari, and Torr] and [Kalal et al(2012)Kalal, Mikolajczyk, and Matas] respectively.



## References

- [Babenko et al(2011)]Babenko, Yang, and Belongie] Babenko B, Yang MH, Belongie S (2011) Robust object tracking with online multiple instance learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* 33(8):1619–1632, DOI 10.1109/TPAMI.2010.226
- [Danelljan et al(2014)]Danelljan, Häger, Khan, and Felsberg] Danelljan M, Häger G, Khan FS, Felsberg M (2014) Accurate scale estimation for robust visual tracking. In: *Proceedings of British Machine Vision Conference (BMVC)*
- [Danelljan et al(2015)]Danelljan, Häger, Shahbaz Khan, and Felsberg] Danelljan M, Häger G, Shahbaz Khan F, Felsberg M (2015) Learning spatially regularized correlation filters for visual tracking. In: *IEEE Proceedings of International Conference on Computer Vision (ICCV)*, pp 4310–4318
- [Felzenszwalb et al(2010)]Felzenszwalb, Girshick, McAllester, and Ramanan] Felzenszwalb PF, Girshick RB, McAllester D, Ramanan D (2010) Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* 32(9):1627–1645
- [Hare et al(2011)]Hare, Saffari, and Torr] Hare S, Saffari A, Torr PH (2011) Struck: Structured output tracking with kernels. In: *IEEE Proceedings of International Conference on Computer Vision (ICCV)*, IEEE, pp 263–270
- [Henriques et al(2015)]Henriques, Caseiro, Martins, and Batista] Henriques JF, Caseiro R, Martins P, Batista J (2015) High-speed tracking with kernelized correlation filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* 37(3):583–596
- [Kalal et al(2012)]Kalal, Mikolajczyk, and Matas] Kalal Z, Mikolajczyk K, Matas J (2012) Tracking-learning-detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* 34(7):1409–1422
- [King(2015)] King DE (2015) Max-margin object detection. *arXiv preprint arXiv:150200046*
- [Li et al(2015)]Li, Zhu, and Hoi] Li Y, Zhu J, Hoi SC (2015) Reliable patch trackers: Robust visual tracking by exploiting reliable patches. In: *IEEE Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR)*, pp 353–361
- [Nebahay and Pflugfelder(2015)] Nebahay G, Pflugfelder R (2015) Clustering of Static-Adaptive correspondences for deformable object tracking. In: *IEEE Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE
- [Ross et al(2008)]Ross, Lim, Lin, and Yang] Ross DA, Lim J, Lin RS, Yang MH (2008) Incremental learning for robust visual tracking. *International Journal of Computer Vision (IJCV)* 77(1-3):125–141
- [Sevilla-Lara and Learned-Miller(2012)] Sevilla-Lara L, Learned-Miller E (2012) Distribution fields for tracking. In: *IEEE Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp 1910–1917
- [Viola and Jones(2004)] Viola P, Jones MJ (2004) Robust real-time face detection. *International Journal of Computer Vision (IJCV)* 57(2):137–154
- [Wu et al(2012)]Wu, Shen, and Ling] Wu Y, Shen B, Ling H (2012) Online robust image alignment via iterative convex optimization. In: *IEEE Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp 1808–1814
- [Zhang et al(2014a)]Zhang, Zhang, and Yang] Zhang K, Zhang L, Yang MH (2014a) Fast compressive tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* 36(10):2002–2015
- [Zhang and van der Maaten(2014)] Zhang L, van der Maaten L (2014) Preserving structure in model-free tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)* 36(4):756–769
- [Zhang et al(2014b)]Zhang, Liu, Ahuja, Yang, and Ghanem] Zhang T, Liu S, Ahuja N, Yang MH, Ghanem B (2014b) Robust visual tracking via consistent low-rank sparse learning. *International Journal of Computer Vision (IJCV)* 111(2):171–190
- [Zhu and Ramanan(2012)] Zhu X, Ramanan D (2012) Face detection, pose estimation, and landmark localization in the wild. In: *IEEE Proceedings of International Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp 2879–2886