



UvA-DARE (Digital Academic Repository)

Tubelets: Unsupervised Action Proposals from Spatiotemporal Super-Voxels

Jain, M.; van Gemert, J.; Jégou, H.; Bouthemy, P.; Snoek, C.G.M.

DOI

[10.1007/s11263-017-1023-9](https://doi.org/10.1007/s11263-017-1023-9)

Publication date

2017

Document Version

Other version

Published in

International Journal of Computer Vision

[Link to publication](#)

Citation for published version (APA):

Jain, M., van Gemert, J., Jégou, H., Bouthemy, P., & Snoek, C. G. M. (2017). Tubelets: Unsupervised Action Proposals from Spatiotemporal Super-Voxels. *International Journal of Computer Vision*, 124(3), 287-311. <https://doi.org/10.1007/s11263-017-1023-9>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Tubelets: Unsupervised action proposals from spatiotemporal super-voxels

Mihir Jain, Jan van Gemert, Hervé Jégou, Patrick Bouthemy, Cees G.M. Snoek

1 Tubelet proposal evaluation on MSR-II and UCF101

Super-voxel grouping	MABO	Recall	#Proposals
Single grouping function			
<i>Motion</i>	19.4	0.4	999
<i>Color</i>	16.1	0	3398
<i>Texture</i>	14.4	0	2739
<i>Size</i>	20.3	0	6555
<i>Fill</i>	26.5	2.2	6482
<i>Motion+Size+Fill</i>	24.0	0	4969
<i>Texture+Size+Fill</i>	22.1	1.1	5222
<i>All-but-motion</i>	19.4	0	4610
<i>All</i>	21.0	0	4584
Multiple grouping functions			
Union set, Φ	29.0	2.2	21644

Table 1 Super-voxel groupings with *video segmentation* on MSR-II. Numbers are low without combination.

Super-voxel grouping	MABO	Recall	#Proposals
Single grouping function			
<i>Motion</i>	27.2	1.5	740
<i>Color</i>	29.4	1.5	676
<i>Texture</i>	26.2	1.7	616
<i>Size</i>	26.5	0	1047
<i>Fill</i>	29.0	0.4	1036
<i>Motion+Size+Fill</i>	28.8	2.1	862
<i>Texture+Size+Fill</i>	30.5	0.4	862
<i>All-but-motion</i>	31.1	1.5	791
<i>All</i>	30.1	1.7	790
Multiple grouping functions			
Union set, Φ	34.8	3.0	4218

Table 2 Super-voxel groupings with *segmentation of iMotion maps* on MSR-II. Here the performance is slightly better than in Table 1.

Super-voxel grouping	MABO	Recall	#Proposals
Single grouping function			
<i>Motion</i>	33.7	19.8	420
<i>Color</i>	28.9	12.3	805
<i>Texture</i>	27.0	10.3	670
<i>Size</i>	27.9	8.0	1318
<i>Fill</i>	28.6	9.9	1285
<i>Motion+Size+Fill</i>	33.9	18.7	1062
<i>Texture+Size+Fill</i>	29.8	12.2	1164
<i>All-but-motion</i>	31.0	15.3	1023
<i>All</i>	34.4	20.3	1004
Multiple grouping functions			
Union set, Φ	39.0	27.5	4794

Table 3 Super-voxel groupings with *video segmentation* on training set of UCF101. The *iMotion* based functions perform better. The union of the five selected grouping functions, Φ , increases the MABO and Recall.

Super-voxel grouping	MABO	Recall	#Proposals
Single grouping function			
<i>Motion</i>	24.8	10.2	167
<i>Color</i>	23.3	8.0	153
<i>Texture</i>	22.0	5.4	141
<i>Size</i>	22.6	5.2	246
<i>Fill</i>	21.4	4.6	245
<i>Motion+Size+Fill</i>	24.7	9.5	204
<i>Texture+Size+Fill</i>	22.4	6.0	215
<i>All-but-motion</i>	23.4	7.7	190
<i>All</i>	25.1	10.4	196
Multiple grouping functions			
Union set, Φ	27.0	13.1	1002

Table 4 Super-voxel groupings with *segmentation of iMotion maps* on the training set of UCF101. Again the union set, Φ , improves on MABO and Recall.