Prepared July 31, 2022

Updated February 28, 2024

Trevor Jackson Darrell
748 Sutardja Dai Hall, UC Berkeley CA 94720
trevor@eecs.berkeley.edu

I. Education:

School	Degree	Date
University of Pennsylvania.	B.S.E	1988
MIT	MS	1991
MIT	Ph.D.	1996

II. Experience:

Employer	Position	Beginning	Ending
Univ. of Pennsylvania	Sys. Prog./Administrator	1985	1988
Cold Spring Harbor Lab	Teaching Assistant	1994	1994
Stanford University	Visiting Researcher	1994	1994
Stanford University	Visiting Instructor	1997	1997
Interval Research Corp.	Research Staff	1996	1999
MIT	Assistant Professor	1999	2003
MIT	Associate Professor	2003	2008
ICSI	Group Leader	2008	2014
UC Berkeley CS	Associate Adj. Professor	2008	2011
UC Berkeley CS	Professor-in-Residence	2011	present

III. Professional Service:

Activity	Beginning	Ending
Conference on Computer Vision and Pattern Recognition (CVPR)		
Area Chair	2001, 2008, 2017, 2019, 2020, 2021, 2022	
Program Chair	2010	

International Conference on Computer Vision (ICCV) Area Chair	2005, 2015	
International Conference on Multimodal Interfaces (ICMI)		
Program Committee	2000, 2002	
Advisory board member	2003	2011
Program Chair	2003	
General Chair	2004	
Area Chair	2005, 2006	
European Conference on Computer Vision (ECCV) Area Chair	2016, 2018, 2022	
Intl. Conference on Machine Learning (ICML) Area Chair	2016. 2018	
Intl. Conference on Learning Representations (ICLR) Area Chair	2016, 2020	
Conference on Perceptual User Interfaces (PUI) Program Chair	2001	
Conference on Ubiquitous Computing (UBICOMP) Program Committee	2002	
NSF Panels (CISE, SBIR, HCI, ITR, CAREER, AI, NRI, …)	2002	present
DARPA Information Science and Technology Advisory Study Group (ISAT): Member	2003	2007
Member, Second Appointment	2011	2014
Future Command Study	2003	2004
Embedded Humans Study	2004	2005
Trusted Deployable Adaptive Systems Study	2005	2006
Data Taming Study	2005	2006
Co-Chair, Adaptive and Interactive Representations Quick Reaction Study	2006	
Chair, Exploitation of Persistent Operational Surveillance (EXPOSE) Study	2007	2008
Chair, TACT Study	2012	2013
Chair, TRUST Study	2013	2014
Neural Information Processing Systems (NIPS) Conference: Senior Program Committee / Area Chair	2004	
Senior Area Chair	2018	
Editor, *CACM Special Issue on Perceptive Multimodal Interfaces*	2004	
AAAI conference; Program Committee, 2004	2004	
Associate Editor, *Artificial Intelligence Journal*	2004	2011
Associate Editor, *IEEE Trans. Pattern Analysis and Machine Intelligence*	2005	2011
Second term	2015	2017
DARPA Computer Sciences Futures Study (Junior Faculty), Member	2006	

IV. Honors and Awards

Outstanding paper award, International Conference on Multimodal Interfaces, 2005

Outstanding paper award, International Conference on Multimodal Interfaces, 2006

Best paper award, Internet Vision Workshop, 2008

Keynote Speaker, ACM Intelligent User Interfaces Conference: "Image Recognition Interfaces," February 2009.

Best Cognitive Robotics Paper Award, International Conference on Robotics and Automation 2013

Best Open Source Software Award, ACM Multimedia 2014

Best Paper Honorable Mention, Computer Vision and Pattern Recognition 2015

Best Paper, North American Association for Computational Linguistics, 2016

Everingham Award, ICCV 2017

V. Postdocs

Current

Name	Dates of Appointment	PhD Granting Institution	Curr. Position
David Chan	2024-	UCB	Postdoct.

Previous

Name	Title	Employer	Position
Konrad Tollmar	Postdoctoral Lecturer	Lund University	Assoc. Prof.
David Demirdjian	Postdoctoral Researcher	Toyota Research	Research Sci.
Raquel Urtasun	Postdoctoral Researcher	TTI-C	Asst. Prof.
Mario Fritz	Postdoctoral Researcher	MPI	Junior Prof.
Mathieu Salsmann	Postdoctoral Researcher	TTI-C	Asst. Prof.
Mario Christoudias	Postdoctoral Researcher	EPFL	Postdoc
Brian Kulis	Postdoctoral Researcher	Ohio State University	Asst. Prof.
Kate Saenko	Postdoctoral Researcher	B.U.	Assoc. Prof.
Sergio Guadarrama	Postdoctoral Researcher	Google	Research Engineer
Daniel Goehring	Postdoctoral Researcher	HU Berlin	Asst. Prof
Lorenzo Riano	Postdoctoral Researcher	Bosch NA	Sr. Research Eng.
Jiashi Feng	Postdoctoral Researcher	NUS	Asst. Prof

Stefanie Jegelka	Postdoctoral Researcher	MIT	Asst. Prof
Marcus Rohrbach	Postdoctoral Researcher	FAIR	Research Sci.
Ryan Ferrell	Postdoctoral Researcher	BYU	Asst. Prof.
Ross Girshick	Postdoctoral Researcher	FAIR	Research Sci.
Philipp Krahenbuhl	Postdoctoral Researcher	UT Austin	Asst. Prof.
Oscar Beijiboom	Postdoctoral Researcher	Startup	CTO
Zeynep Akata	Postdoctoral Researcher	U. Amsterdam	Asst. Prof.
Fisher Yu	Postdoctoral Researcher	ETHZ	Asst. Prof.
Huijuan Xu	Postdoctoral Researcher	Penn State	Asst. Prof.
Anna Rohrbach	Postdoc & Rsrch. Sci.	TU Darmstadt	Prof.
Angjoo Kanazawa	Postdoctoral Researcher	UCB	Asst. Prof.
Xihui Liu	Postdoctoral Researcher	Hong Kong University	Asst. Prof.
Roi Herzig	Postdoctoral Researcher	IBM	Research Scientist

VI. Students

2005	Gregory Shakhnarovich	MIT EECS	Doctor of Philosophy	Learning Features for Visual Classification
2005	Ali Rahimi	MIT EECS	Doctor of Philosophy	Learning to Transform Time Series with a Few Examples, Oct. 2005
2006	Louis-Philippe Morency	MIT EECS	Doctor of Philosophy	Dialogue Context and Visual Gesture Recognition

2006	Kevin Wilson	MIT EECS	Doctor of Philosophy	Learning Uncertainty Models for Audiovisual Speech Source Localization in Real-World Environments
2006	Kristen Grauman	MIT EECS	Doctor of Philosophy	Matching sets of features for efficient retrieval and recognition,
2006	Leonid Taycher	MIT EECS	Doctor of Philosophy	Statistical methods for dynamic visual processing
2008	Sy Bor Wang	MIT EECS	Doctor of Philosophy	Communication Error Detection Using Facial Expressions
2009	Kate Saenko	MIT EECS	Doctor of Philosophy	Image Sense Disambiguation: A Multimodal Approach
2009	Tom Yeh	MIT EECS	Doctor of Philosophy	Interacting with Computers using Images for Search and Automation
2009	Mario Christhoudias	MIT EECs	Doctor of Philosophy	Probabilistic Models for Semi-Supervised

				Learning and Coding
2009	Ariadna Quattoni	MIT EECS	Doctor of Philosophy	Transfer Learning Algorithms for Image Classification
2010	Ashley Eden	Computer Science PhD	Doctor of Philosophy	Finding Lost Children
2014	Sergey K. Karayev	Computer Science PhD	Doctor of Philosophy	Anytime Recognition of Objects and Scenes
2014	Hyun Oh Song	Computer Science PhD	Doctor of Philosophy	Learning with Parsimony for Large Scale Object Detection and Discovery
2014	Yangqing Jia	Computer Science PhD	Doctor of Philosophy	
2015	Ning Zhang	Computer Science PhD	Doctor of Philosophy	Visual Representations for Fine-grained Categorization

2016	Judy Hoffman	EECS	Doctor of Philosophy	Adaptive Learning Algorithms for Transferable Visual Recognition
2016	Judith F. Hoffman	Electrical Eng & Comp Sci PhD	Doctor of Philosophy	Adaptive Learning Algorithms for Transferable Visual Recognition
2016	Jonathan L. Long	Computer Science PhD	Doctor of Philosophy	Understanding and Designing Convolutional Networks for Local Recognition Problems
2017	Jeffrey D. Donahue	Computer Science PhD	Doctor of Philosophy	Transferrable Representations for Visual Recognition
2017	Deepak Pathak	EECS	Doctor of Philosophy	
2018	Chelsea B. Finn	EECS	Doctor of Philosophy	Learning to Learn with Gradients
2019	Yang Gao	Computer Science PhD	Doctor of Philosophy	End to End Learning in Autonomous Driving Systems

2019	Deepak Pathak	Computer Science PhD	Doctor of Philosophy	Learning to Generalize via Self-Supervised Prediction
2019	Evan G. Shelhamer	Computer Science PhD	Doctor of Philosophy	Local and Adaptive Image-to-Image Learning and Inference
2019	Lisa Anne M. Hendricks	Electrical Eng & Comp Sci PhD	Doctor of Philosophy	Visual Understanding through Natural Language
2020	Coline Devin	Computer Science PhD	Doctor of Philosophy	Compositionality and Modularity for Robot Learning
2020	Eric S. Tzeng	Computer Science PhD	Doctor of Philosophy	Adapting Across Domains by Aligning Representations and Images
2020	Sayna Ebrahimi	Mechanical Engineering PhD	Doctor of Philosophy	Mechanical Behavior of Materials at Multiscale Peridynamic Theory and Learning-based Approaches

2020	Ronghang Hu	Computer Science PhD	Doctor of Philosophy	Structured Models for Vision-and-Language Reasoning
2020	Xin Wang	Computer Science PhD	Doctor of Philosophy	The Design of Dynamic Neural Networks for Efficient Learning and Inference
2020	Samaneh Azadi	Computer Science PhD	Doctor of Philosophy	Visual Content Creation by Generative Adversarial Networks
2021	Vijay Govindarajan	Electrical Eng & Comp Sci PhD	Doctor of Philosophy	Adaptive Prediction and Planning for Safe and Effective Autonomous Vehicles
2021	Huazhe Xu	Computer Science PhD	Doctor of Philosophy	Learning Predictive Models for Efficient Policy Learning
2022	Zhuang Liu	Computer Science PhD	Doctor of Philosophy	Efficient and Scalable Neural Architectures for Visual Recognition

2022	Dequan Wang	Computer Science PhD	Doctor of Philosophy	Learning to Generalize in Dynamic Environments
2022	Parsa Mahmoudieh	Computer Science PhD	Doctor of Philosophy	Multi-task Policy Learning with Minimal Human Supervision
2023	Shizhan Zhu	Computer Science PhD	Doctor of Philosophy	Implicit Modeling for 3D Applications
2023	Medhini Gulganjalli Narasimhan	Computer Science PhD	Doctor of Philosophy	Multimodal Long-Term Video Understanding
2023	Tete Xiao	Computer Science PhD	Doctor of Philosophy	Scalable Representations for Vision and Robotics
2023	Dong Huk S. Park	Computer Science PhD	Doctor of Philosophy	Vision and Language Understanding Through Generative Modeling
current	Xudong Wang	Electrical Eng & Comp Sci PhD	Doctor of Philosophy	

current	Evonne Ng	Electrical Eng & Comp Sci PhD	Doctor of Philosophy	
current	Sheng Shen	Computer Science PhD	Doctor of Philosophy	
current	Marissa I. Ramirez de Chanlatte	Applied Science & Tech PhD	Doctor of Philosophy	
current	Olivia G. Watkins	Computer Science PhD	Doctor of Philosophy	
current	Suzie Petryk	Computer Science PhD	Doctor of Philosophy	

VII. Committee Service

2001-2006 MIT EECS Graduate Admissions

2002 MIT LCS and AI lab Merger Committee

2010-2017 UC EECS Graduate Admissions

2018-present UC EECS Graduate Matters

2021-2023 UC EECS External Relations

2023-present UC Berkeley COI

VIII. Publications

Refereed Journal articles (through 2022; for updates see Google Scholar):

1. Darrell, T., & Wohn, K. (1990). Depth from focus using a pyramid architecture. Pattern Recognition Letters, 11(12), 787–796. http://doi.org/10.1016/0167-8655(90)90032-W

2. Pentland, A., Scherock, S., Darrell, T., & Girod, B. (1994). Simple range cameras based on focal error. Journal of the Optical Society of America A, 11(11), 2925. http://doi.org/10.1364/JOSAA.11.002925

3. Darrell, T., & Pentland, A. P. (1995). Cooperative robust estimation using layers of support. IEEE Transactions on Pattern Analysis and Machine Intelligence, 17(5), 474–487. http://doi.org/10.1109/34.391395

4. Johnson, M. P., Maes, P., & Darrell, T. (1994). Evolving Visual Routines. Artificial Life, 1(4), 373–389. http://doi.org/10.1162/artl.1994.1.4.373

5. Darrell, T. J., Essa, I. A., & Pentland, A. P. (1996). Task-specific gesture analysis in real-time using interpolated views. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(12), 1236–1242. http://doi.org/10.1109/34.546259

6. Maes, P., Darrell, T., Blumberg, B., & Pentland, A. (1997). The ALIVE system: wireless, full-body interaction with autonomous agents. Multimedia Systems, 5(2), 105–112. http://doi.org/10.1007/s005300050046

7. Wren, C. R., Azarbayejani, A., Darrell, T., & Pentland, A. P. (1997). Pfinder: real-time tracking of the human body. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19(7), 780–785. http://doi.org/10.1109/34.598236

8. Darrell, T., Gordon, G., Harville, M., & Woodfill, J. (2000). Integrated Person Tracking Using Stereo, Color, and Pattern Detection. International Journal of Computer Vision, 37(2), 175–185. http://doi.org/10.1023/A:1008103604354

9. Ackerman, M., Darrell, T., & Weitzner, D. J. (2001). Privacy in Context. Human–Computer Interaction, 16(2–4), 167–176. http://doi.org/10.1207/S15327051HCI16234_03

10. Darrell, T., & Covell, M. (2001). Correspondence with cumulative similarity transforms. IEEE Transactions on Pattern Analysis and Machine Intelligence, 23(2), 222–227. http://doi.org/10.1109/34.908973

11. Taycher, L., & Darrell, T. (2002). Range Segmentation Using Visibility Constraints. International Journal of Computer Vision, 47(1–3), 89–98. http://doi.org/10.1023/A:1014533505864

12. Demirdjian, D., & Darrell, T. (2002). Using Multiple-Hypothesis Disparity Maps and Image Velocity for 3-D Motion Estimation. International Journal of Computer Vision, 47(1–3), 219–228. http://doi.org/10.1023/A:1014502126337

13. Bentley, F., Tollmar, K., Demirdjian, D., Koile, K., & Darrell, T. (2003). Perceptive presence. IEEE Computer Graphics and Applications, 23(5), 26–36. http://doi.org/10.1109/MCG.2003.1231175

14. Fisher, J. W., & Darrell, T. (2004). Speaker association with signal-level audiovisual fusion. IEEE Transactions on Multimedia, 6(3), 406–413. http://doi.org/10.1109/TMM.2004.827503

15. Demirdjian, D., Ko, T., & Darrell, T. (2005). Untethered gesture acquisition and recognition for virtual world manipulation. Virtual Reality, 8(4), 222–230. http://doi.org/10.1007/s10055-005-0155-3

16. Grauman, K., & Darrell, T. (2005a). Efficient image matching with distributions of local invariant features. In 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) (Vol. 2, pp. 627–634 vol. 2). http://doi.org/10.1109/CVPR.2005.138

17. Grauman, K., & Darrell, T. (2005b). The pyramid match kernel: discriminative classification with sets of image features. In Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1 (Vol. 2, p. 1458–1465 Vol. 2). http://doi.org/10.1109/ICCV.2005.239

18. Christoudias, C. M., Morency, L.-P., & Darrell, T. (2006). Non-parametric and light-field deformable models. Computer Vision and Image Understanding, 104(1), 16–35. http://doi.org/10.1016/j.cviu.2006.06.001

19. Wilson, K. W., & Darrell, T. (2006). Learning a Precedence Effect-Like Weighting Function for the Generalized Cross-Correlation Framework. IEEE Transactions on Audio, Speech, and Language Processing, 14(6), 2156–2164. http://doi.org/10.1109/TASL.2006.872601

20. Taycher, L., Fisher III, J. W., & Darrell, T. (2007). Combining object and feature dynamics in probabilistic tracking. Computer Vision and Image Understanding, 108(3), 243–260. http://doi.org/10.1016/j.cviu.2006.11.022

21. Rahimi, A., Recht, B., & Darrell, T. (2007). Learning to Transform Time Series with a Few Examples. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(10), 1759–1775. http://doi.org/10.1109/TPAMI.2007.1001

22. Quattoni, A., Wang, S., Morency, L. P., Collins, M., & Darrell, T. (2007). Hidden Conditional Random Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(10), 1848–1852. http://doi.org/10.1109/TPAMI.2007.1124

23. Morency, L.-P., Sidner, C., Lee, C., & Darrell, T. (2007). Head gestures for perceptual interfaces: The role of context in improving recognition. Artificial Intelligence, 171(8), 568–585. http://doi.org/10.1016/j.artint.2007.04.003

24. Rahimi, A., Morency, L.-P., & Darrell, T. (2008). Reducing drift in differential tracking. Computer Vision and Image Understanding, 109(2), 97–111. http://doi.org/10.1016/j.cviu.2006.12.004

25. Saenko, K., Livescu, K., Glass, J., & Darrell, T. (2009). Multistream Articulatory Feature-Based Models for Visual Speech Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(9), 1700–1707. http://doi.org/10.1109/TPAMI.2008.303

26. Kapoor, A., Grauman, K., Urtasun, R., & Darrell, T. (2009). Gaussian Processes for Object Categorization. International Journal of Computer Vision, 88(2), 169–188. http://doi.org/10.1007/s11263-009-0268-3

27. Stone, Z., Zickler, T., & Darrell, T. (2010). Toward Large-Scale Face Recognition Using Social Network Context. Proceedings of the IEEE, 98(8), 1408–1415. http://doi.org/10.1109/JPROC.2010.2044551

28. Miller, S., Berg, J. van den, Fritz, M., Darrell, T., Goldberg, K., & Abbeel, P. (2012). A geometric approach to robotic laundry folding. The International Journal of Robotics Research, 31(2), 249–267. http://doi.org/10.1177/0278364911430417

29. Chung, S., Mario Christoudias, C., Darrell, T., Ziniel, S. I., & Kalish, L. A. (2012a). A Novel Image-based Tool to Reunite Children With Their Families After Disasters. Academic Emergency Medicine, 19(11), 1227–1234. http://doi.org/10.1111/acem.12013

30. Chakrabarti, A., Xiong, Y., Sun, B., Darrell, T., Scharstein, D., Zickler, T., & Saenko, K. (2014). Modeling Radiometric Uncertainty for Vision with Tone-Mapped Color Images. IEEE Transactions on Pattern Analysis and Machine Intelligence, 36(11), 2185–2198. http://doi.org/10.1109/TPAMI.2014.2318713

31. Chu, V., McMahon, I., Riano, L., McDonald, C. G., He, Q., Martinez Perez-Tejada, J., … Kuchenbecker, K. J. (2015). Robotic learning of haptic adjectives through physical interaction. Robotics and Autonomous Systems, 63, Part 3, 279–292. http://doi.org/10.1016/j.robot.2014.09.021

32. Darrell, T., Ferrari, V., Jurie, F., & Lepetit, V. (2015). Introduction to the CVIU Special Issue on "Parts and Attributes. Comput. Vis. Image Underst., 138(C), 85–. http://doi.org/10.1016/j.cviu.2015.07.001

33. Song, H. O., Girshick, R., Zickler, S., Geyer, C., Felzenszwalb, P., & Darrell, T. (2015). Generalized Sparselet Models for Real-Time Multiclass Object Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence, 37(5), 1001–1012. http://doi.org/10.1109/TPAMI.2014.2353631

34. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2016). Region-Based Convolutional Networks for Accurate Object Detection and Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(1), 142–158. http://doi.org/10.1109/TPAMI.2015.2437384

35. Guadarrama, S., Rodner, E., Saenko, K., & Darrell, T. (2016). Understanding object descriptions in robotics by open-vocabulary object retrieval and detection. The International Journal of Robotics Research, 35(1–3), 265–280. http://doi.org/10.1177/0278364915602059

36. Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-End Training of Deep Visuomotor Policies. Journal of Machine Learning Research, 17, 1–40. Retrieved from http://www.jmlr.org/papers/volume17/15-522/15-522.pdf

37. Shelhamer, E., Long, J., & Darrell, T. (2016). Fully Convolutional Networks for Semantic Segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence, PP(99), 1–1. http://doi.org/10.1109/TPAMI.2016.2572683

38. Song, H. O., Fritz, M., Goehring, D., & Darrell, T. (2016). Learning to Detect Visual Grasp Affordance. IEEE Transactions on Automation Science and Engineering, 13(2), 798–809. http://doi.org/10.1109/TASE.2015.2396014

39. Guadarrama, S., Rodner, E., Saenko, K., Darrell, T. (2016), Understanding object descriptions in robotics by open-vocabulary object retrieval and detection, The International Journal of Robotics Research 35 (1-3), 265-280. https://doi.org/10.1177%2F0278364915602059

40. Hoffman, J., Pathak, D., Tzeng, E.,, Long, J., Guadarrama, S., Darrell, T. (2015), Large scale visual recognition through adaptation using joint representation and multiple instance learning, The Journal of Machine Learning Research 17 (1), 4954-4984. http://www.jmlr.org/papers/v17/15-223.html

41. Simon, M., Rodner, E., Darrell, T., Denzler, J. (2018), The whole is more than its parts? From explicit to implicit pose normalization, IEEE TPAMI, Published online (early access) 18 December 2018, 14pp. https://doi.org/10.1109/TPAMI.2018.2885764

42. Hu, H.-N., Yang, Y.-H., Fischer, T., Darrell, T., Yu, F., and Sun, M. (2022). Monocular quasi-dense 3d object tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence. 8pp [link]

43. Liu, Z., Wang, H., Zhou, T., Shen, Z., Kang, B., Shelhamer, E., and Darrell, T. (2022). Exploring simple and transferable recognitionaware image processing. IEEE Transactions on Pattern Analysis and Machine Intelligence. 8pp. [link]

44. Ebrahimi, S., Petryk, S., Gokul, A., Gan, W., Gonzalez, J. E., Rohrbach, M., and Darrell, T. (2021). Remembering for the right reasons: Explanations reduce catastrophic forgetting. Applied AI Letters, 8pp, 2(4):e44. [link]

45. Hendricks, L. A., Rohrbach, A., Schiele, B., Darrell, T., and Akata, Z. (2021). Generating visual explanations with natural language. Applied AI Letters, 8pp, 2(4):e55. [link]

46. Kim, J., Rohrbach, A., Akata, Z., Moon, S., Misu, T., Chen, Y.-T., Darrell, T., and Canny, J. (2021). Toward explainable and advisable model for self-driving cars. Applied AI Letters, 8pp, 2(4):e56. [link]

47. Watkins, O., Huang, S., Frost, J., Bhatia, K., Weiner, E., Abbeel, P., Darrell, T., Plummer, B., Saenko, K., and Dragan, A. (2021). Explaining robot policies. Applied AI Letters, 8pp, 2(4):e52. [link]

Refereed Conference Papers (through 2022; for updates see Google Scholar)::

1. Batchelder, N., and Darrell, T. (1987), "Psfig - A Ditroff preprocessor for PostScript figures," *USENIX Conference Proceedings*, Phoenix, AZ, pp. 31-42, June 1987.

2. Darrell, T., & Wohn, K. (1988). Pyramid based depth from focus. In *Computer Society Conference on Computer Vision and Pattern Recognition, 1988. Proceedings CVPR '88* (pp. 504–509). http://doi.org/10.1109/CVPR.1988.196282

3. Pentland, A., Darrell, T., Turk, M., & Huang, W. (1989). A simple, real-time range camera. In *, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1989. Proceedings CVPR '89* (pp. 256–261). http://doi.org/10.1109/CVPR.1989.37858

4. Darrell, T., Sclaroff, S., & Pentland, A. (1990). Segmentation by minimal description. In *Proceedings, Third International Conference on Computer Vision, 1990* (pp. 112–116). http://doi.org/10.1109/ICCV.1990.139506

5. Darrell, T., & Pentland, A. (1991). On the representation of occluded shapes. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91* (pp. 728–729). http://doi.org/10.1109/CVPR.1991.139802

6. Darrell, T., & Pentland, A. (1991). Robust estimation of a multi-layered motion representation. In *, Proceedings of the IEEE Workshop on Visual Motion, 1991* (pp. 173–178). http://doi.org/10.1109/WVM.1991.212810

7. Darrell, T., & Pentland, A. (1992). Against Edges: Function Approximation with Multiple Support Maps. In J. E. Moody, S. J. Hanson, & R. P. Lippmann (Eds.), *Advances in Neural Information Processing Systems 4* (pp. 388–395). Morgan-Kaufmann. Retrieved from http://papers.nips.cc/paper/462-against-edges-function-approximation-with-multiple-support-maps.pdf

8. Darrell, T., & Pentland, A. (1993). Space-time gestures. In *, 1993 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1993. Proceedings CVPR '93* (pp. 335–340). http://doi.org/10.1109/CVPR.1993.341109

9. Darrell, T., & Simonecelli, E. (1993). `Nulling' filters and the separation of transparent motions. In *, 1993 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1993. Proceedings CVPR '93* (pp. 738–739). http://doi.org/10.1109/CVPR.1993.341150

10. Johnson, M., Maes, P., & Darell, T. (1994). Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems. In *Artificial Life IV: Proceedings of the Fourth International Workshop on the Synthesis and Simulation of Living Systems* (pp. 373–389). Boston, MA: MIT Press.

11. Darrell, T. J., & Pentland, A. P. (1994). Classifying Hand Gestures with a View-Based Distributed Representation. In J. D. Cowan, G. Tesauro, & J. Alspector (Eds.), *Advances in Neural Information Processing Systems 6* (pp. 945–952). Morgan-Kaufmann. Retrieved from http://papers.nips.cc/paper/832-classifying-hand-gestures-with-a-view-based-distributed-representation.pdf

12. Pentland, A. P. & Darrell, T. (1994). Visual perception of human bodies and faces for multi-modal interfaces. In *ICSLP-1994* (pp. 543-546).

13. Essa, I. A., Darrell, T., & Pentland, A. (1994). Tracking facial motion. In *, Proceedings of the 1994 IEEE Workshop on Motion of Non-Rigid and Articulated Objects, 1994* (pp. 36–42). http://doi.org/10.1109/MNRAO.1994.346257

14. Pentland, A., Darrell, T., Essa, I., Azarbayejani, A., & Sclaroff, S. (1994). Visually guided animation. In *, Proceedings of Computer Animation '94* (pp. 112–121). http://doi.org/10.1109/CA.1994.324000

15. Maes, P., Blumberg, B., Darrell, T., Pentland, A., & Wexelblat, A. (1995). Modeling Interactive Agents in ALIVE. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2* (pp. 2073–2074). San Francisco, CA, USA: Morgan Kaufmann Publishers Inc. Retrieved from http://dl.acm.org/citation.cfm?id=1643031.1643178

16. Darrell, T., Essa, I. A., & Pentland, A. (1995). Correlation and Interpolation Networks for Real-time Expression Analysis/Synthesis. In G. Tesauro, D. S. Touretzky, & T. K. Leen (Eds.), *Advances in Neural Information Processing Systems 7* (pp. 909–916). MIT Press. Retrieved from http://papers.nips.cc/paper/999-correlation-and-interpolation-networks-for-real-time-expression-analysissynthesis.pdf

17. Darrell, T., & Pentland, A. P. (1995). Attention-driven Expression and Gesture Analysis in an Interactive Environment. Presented at the Intl. Workshop on Automatic Face and Gesture Recognition, Zurich, Switzerland. http://doi.org/10.1.1.48.2033

18. Maes, P., Darrell, T., Blumberg, B., & Pentland, A. (1995). The ALIVE system: full-body interaction with autonomous agents. In *Computer Animation '95., Proceedings.* (pp. 11–18, 209). http://doi.org/10.1109/CA.1995.393553

19. Darrell, T., & Pentland, A. (1996). Active gesture recognition using partially observable Markov decision processes. *ICPR 1996* (pp. 984-988)

20. Darrell, T., & Pentland, A. (1996). Active Gesture Recognition using Learned Visual Attention. In D. S. Touretzky & M. E. Hasselmo (Eds.), *Advances in Neural Information Processing Systems 8* (pp. 858–864). MIT Press. Retrieved from http://papers.nips.cc/paper/1079-active-gesture-recognition-using-learned-visual-attention.pdf

21. Darrell, T., Maes, P., Blumberg, B., & Pentland, A. P. (1996). A Novel Environment for Situated Vision and Behavior. In M. S. Landy, L. T. Maloney, & M. Pavel (Eds.), *Exploratory Vision* (pp. 319–331). Springer New York. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4612-3984-0_13

22. Darrell, T. (1998). A radial cumulative similarity transform for robust image correspondence. In *1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1998. Proceedings* (pp. 656–662). http://doi.org/10.1109/CVPR.1998.698674

23. Darrell, T., Gordon, G., Harville, M., & Woodfill, J. (1998). Integrated person tracking using stereo, color, and pattern detection. In *1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1998. Proceedings* (pp. 601–608). http://doi.org/10.1109/CVPR.1998.698667

24. Darrell, T., Gordon, G., Woodfill, J., Baker, H., & Harville, M. (1998). Robust, real-time people tracking in open environments using integrated stereo, color, and face detection. In *, 1998 IEEE Workshop on Visual Surveillance, 1998. Proceedings* (pp. 26–32). http://doi.org/10.1109/WVS.1998.646017

25. Darrell, T., Gordon, G., Woodfill, J., & Harville, M. (1998). A virtual mirror interface using real-time robust face tracking. In *Third IEEE International Conference on Automatic Face and Gesture Recognition, 1998. Proceedings* (pp. 616–621). http://doi.org/10.1109/AFGR.1998.671016

26. Darrell, T. (1999). Example-Based Image Synthesis of Articulated Figures. In M. J. Kearns, S. A. Solla, & D. A. Cohn (Eds.), *Advances in Neural Information Processing Systems 11* (pp. 768–774). MIT Press. Retrieved from http://papers.nips.cc/paper/1504-example-based-image-synthesis-of-articulated-figures.pdf

27. Gordon, G., Darrell, T., Harville, M., & Woodfill, J. (1999). Background estimation and removal based on range and color. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* (Vol. 2, p. 464 Vol. 2). http://doi.org/10.1109/CVPR.1999.784721

28. Covell, M. M., & Darrell, T. J. (1999). Dynamic occluding contours: a new external-energy term for snakes. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.* (Vol. 2, p. 238 Vol. 2). http://doi.org/10.1109/CVPR.1999.784635

29. Harville, M., Rahimi, A., Darrell, T., Gordon, G., & Woodfill, J. (1999). 3D pose tracking with linear depth and brightness constraints. In *The Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999* (Vol. 1, pp. 206–213 vol.1). http://doi.org/10.1109/ICCV.1999.791219

30. Covell, M. M., Rahini, A., Harville, M., & Darrell, T. J. (2000). Articulated-pose estimation using brightness- and depth-constancy constraints. In *IEEE Conference on Computer Vision and Pattern Recognition, 2000. Proceedings* (Vol. 2, pp. 438–445 vol.2). http://doi.org/10.1109/CVPR.2000.854875

31. Darrell, T., Fisher, J. W., III, Viola, P. A., & Freeman, W. T. (2000). Ausio-visual Segmentation and "The Cocktail Party Effect." In *Proceedings of the Third International Conference on Advances in Multimodal Interfaces* (pp. 32–40). London, UK, UK: Springer-Verlag. Retrieved from http://dl.acm.org/citation.cfm?id=645524.656793

32. Fisher III, J. W., Darrell, T., Freeman, W. T., & Viola, P. A. (2001). Learning Joint Statistical Models for Audio-Visual Fusion and Segregation. In T. K. Leen, T. G. Dietterich, & V. Tresp (Eds.), *Advances in Neural Information Processing Systems 13* (pp. 772–778). MIT Press. Retrieved from http://papers.nips.cc/paper/1898-learning-joint-statistical-models-for-audio-visual-fusion-and-segregation.pdf

33. Shakhnarovich, G., Lee, L., & Darrell, T. (2001). Integrated face and gait recognition from multiple views. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001* (Vol. 1, p. I-439-I-446 vol.1). http://doi.org/10.1109/CVPR.2001.990508

34. Rahimi, A., Morency, L. P., & Darrell, T. (2001). Reducing drift in parametric motion tracking. In *Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings* (Vol. 1, pp. 315–322 vol.1). http://doi.org/10.1109/ICCV.2001.937535

35. Demirdjian, D., & Darrell, T. (2001). Motion estimation from disparity images. In *Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings* (Vol. 1, pp. 213–218 vol.1). http://doi.org/10.1109/ICCV.2001.937520

36. Darrell, T., Demirdjian, D., Checka, N., & Felzenszwalb, P. (2001). Plan-view trajectory estimation with dense stereo background models. In *Eighth IEEE International Conference*

on Computer Vision, 2001. ICCV 2001. Proceedings (Vol. 2, pp. 628–635 vol.2). http://doi.org/10.1109/ICCV.2001.937685

37. Demirdjian, D., & Darrell, T. (2001). Using multiple-hypothesis disparity maps and image velocity for 3-D motion estimation. In *IEEE Workshop on Stereo and Multi-Baseline Vision, 2001. (SMBV 2001). Proceedings* (pp. 121–128). http://doi.org/10.1109/SMBV.2001.988770

38. Taycher, L., & Darrell, T. (2001). Range segmentation using visibility constraints. In *IEEE Workshop on Stereo and Multi-Baseline Vision, 2001. (SMBV 2001). Proceedings* (pp. 37–43). http://doi.org/10.1109/SMBV.2001.988761

39. Wilson, K., Checka, N., Demirdjian, D., & Darrell, T. (2001). Audio-video Array Source Separation for Perceptual User Interfaces. In *Proceedings of the 2001 Workshop on Perceptive User Interfaces* (pp. 1–7). New York, NY, USA: ACM. http://doi.org/10.1145/971478.971500

40. Shakhnarovich, G., Fisher, J. W., & Darrell, T. (2002). Face Recognition from Long-Term Observations. In A. Heyden, G. Sparr, M. Nielsen, & P. Johansen (Eds.), *ECCV 2002* (pp. 851–865). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/3-540-47977-5_56

41. Darrell, T., Tollmar, K., Bentley, F., Checka, N., Morency, L.-P., Rahimi, A., & Oh, A. (2002). Face-Responsive Interfaces: From Direct Manipulation to Perceptive Presence. In G. Borriello & L. E. Holmquist (Eds.), *UbiComp 2002: Ubiquitous Computing* (pp. 135–151). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/3-540-45809-3_10

42. Demirdjian, D., & Darrell, T. (2002). 3-D Articulated Pose Tracking for Untethered Diectic Reference. In *Proceedings of the 4th IEEE International Conference on Multimodal Interfaces* (p. 267–). Washington, DC, USA: IEEE Computer Society. http://doi.org/10.1109/ICMI.2002.1167005

43. Fisher, J. W. III, & Darrell, T. (2002). Probabalistic Models and Informative Subspaces for Audiovisual Correspondence. In A. Heyden, G. Sparr, M. Nielsen, & P. Johansen (Eds.), *Computer Vision — ECCV 2002* (pp. 592–603). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/3-540-47977-5_39

44. Wilson, K., Rangarajan, V., Checka, N., & Darrell, T. (2002). Audiovisual Arrays for Untethered Spoken Interfaces. In *Proceedings of the 4th IEEE International Conference on*

Multimodal Interfaces (p. 389–). Washington, DC, USA: IEEE Computer Society. http://doi.org/10.1109/ICMI.2002.1167026

45. Demirdjian, D., Tollmar, K., Koile, K., Checka, N., & Darrell, T. (2002) Activity maps for location-aware computing. *Applications of Computer Vision* (p. 70). IEEE Workshop on Applications of Computer Vision. (WACV) http://dx.doi.org/10.1109/ACV.2002.1182159

46. Morency, L. P., & Darrell, T. (2002). Stereo tracking using ICP and normal flow constraint. In *16th International Conference on Pattern Recognition, 2002. Proceedings* (Vol. 4, pp. 367–372 vol.4). http://doi.org/10.1109/ICPR.2002.1047472

47. Morency, L. P., Rahimi, A., Checka, N., & Darrell, T. (2002). Fast stereo-based head tracking for interactive environments. In *Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002. Proceedings* (pp. 390–395). http://doi.org/10.1109/AFGR.2002.1004185

48. Morency, L. P., Rahimi, A., & Darrell, T. (2002). Fast 3D model acquisition from stereo images. In *First International Symposium on 3D Data Processing Visualization and Transmission, 2002. Proceedings* (pp. 172–176). http://doi.org/10.1109/TDPVT.2002.1024057

49. Oh, A., Fox, H., Van Kleek, M., Adler, A., Gajos, K., Morency, L.-P., & Darrell, T. (2002). Evaluating Look-to-talk: A Gaze-aware Interface in a Collaborative Environment. In *CHI '02 Extended Abstracts on Human Factors in Computing Systems* (pp. 650–651). New York, NY, USA: ACM. http://doi.org/10.1145/506443.506528

50. Rahimi, A., & Darrell, T. (2002). Bayesian network for online global pose estimation. In *IEEE/RSJ International Conference on Intelligent Robots and Systems, 2002* (Vol. 1, pp. 427–433 vol.1). http://doi.org/10.1109/IRDS.2002.1041427

51. Shakhnarovich, G., & Darrell, T. (2002). On probabilistic combination of face and gait cues for identification. In *Fifth IEEE International Conference on Automatic Face and Gesture Recognition, 2002. Proceedings* (pp. 169–174). http://doi.org/10.1109/AFGR.2002.1004151

52. Wilson, K. W., & Darrell, T. (2002). Audio-video array source localization for intelligent environments. In *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)* (Vol. 2, p. II-2109-II-2112). http://doi.org/10.1109/ICASSP.2002.5745051

53. Rahimi, A., & Darrell, T. (2003). Location Estimation with a Differential Update Network. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in Neural Information Processing*

Systems 15 (pp. 1073–1080). MIT Press. Retrieved from http://papers.nips.cc/paper/2341-location-estimation-with-a-differential-update-network.pdf

54. Taycher, L., Iii, J., & Darrell, T. (2003). Recovering Articulated Model Topology from Observed Rigid Motion. In S. Becker, S. Thrun, & K. Obermayer (Eds.), *Advances in Neural Information Processing Systems 15* (pp. 1335–1342). MIT Press. Retrieved from http://papers.nips.cc/paper/2182-recovering-articulated-model-topology-from-observed-rigid-motion.pdf

55. Morency, L. P., Rahimi, A., & Darrell, T. (2003). Adaptive view-based appearance models. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings* (Vol. 1, p. I-803-I-810 vol.1). http://doi.org/10.1109/CVPR.2003.1211435

56. Demirdjian, D., Ko, T., & Darrell, T. (2003). Constraining human body tracking. In *Ninth IEEE International Conference on Computer Vision, 2003. Proceedings* (pp. 1071–1078 vol.2). http://doi.org/10.1109/ICCV.2003.1238468

57. Grauman, K., Shakhnarovich, G., & Darrell, T. (2003b). Inferring 3D structure with a statistical image-based shape model. In *Ninth IEEE International Conference on Computer Vision, 2003. Proceedings* (pp. 641–647 vol.1). http://doi.org/10.1109/ICCV.2003.1238408

58. Shakhnarovich, G., Viola, P., & Darrell, T. (2003). Fast pose estimation with parameter-sensitive hashing. In *Ninth IEEE International Conference on Computer Vision, 2003. Proceedings* (pp. 750–757 vol.2). http://doi.org/10.1109/ICCV.2003.1238424

59. Koile, K., Tollmar, K., Demirdjian, D., Shrobe, H., & Darrell, T. (2003). Activity Zones for Context-Aware Computing. In A. K. Dey, A. Schmidt, & J. F. McCarthy (Eds.), *UbiComp 2003: Ubiquitous Computing* (pp. 90–106). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-39653-6_7

60. Siracusa, M., Morency, L.-P., Wilson, K., Fisher, J., & Darrell, T. (2003). A Multi-modal Approach for Determining Speaker Location and Focus. In *Proceedings of the 5th International Conference on Multimodal Interfaces* (pp. 77–80). New York, NY, USA: ACM. http://doi.org/10.1145/958432.958449

61. Ko, T., Demirdjian, D., & Darrell, T. (2003). Untethered Gesture Acquisition and Recognition for a Multimodal Conversational System. In *Proceedings of the 5th International Conference on Multimodal Interfaces* (pp. 147–150). New York, NY, USA: ACM. http://doi.org/10.1145/958432.958461

62. Grauman, K., Shakhnarovich, G., & Darrell, T. (2003a). A Bayesian approach to image-based visual hull reconstruction. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings* (Vol. 1, p. I-187-I-194 vol.1). http://doi.org/10.1109/CVPR.2003.1211353

63. Checka, N., Wilson, K., Rangarajan, V., & Darrell, T. (2003). A Probabilistic Framework for Multi-modal Multi-Person Tracking. In *Conference on Computer Vision and Pattern Recognition Workshop, 2003. CVPRW '03* (Vol. 9, pp. 100–100). http://doi.org/10.1109/CVPRW.2003.10099

64. Fisher, J.W., & Darrell, T. (2003) Learning cross-modal appearance models with application to tracking. In *Proceedings, 2003 International Conference on Multimedia and Expo* (pp. II-13-16). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=1221541&isnumber=27437

65. Morency, L. P., Sundberg, P., & Darrell, T. (2003). Pose estimation using 3D view-based eigenspaces. In *IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 2003. AMFG 2003* (pp. 45–52). http://doi.org/10.1109/AMFG.2003.1240823

66. Tollmar, K., Demirdjian, D., & Darrell, T. (2003). Gesture + Play Exploring Full-Body Navigation for Virtual Environments. In *Conference on Computer Vision and Pattern Recognition Workshop, 2003. CVPRW '03* (Vol. 5, pp. 47–47). http://doi.org/10.1109/CVPRW.2003.10046

67. Christoudias, C. M., Morency, L.-P., & Darrell, T. (2004). Light Field Appearance Manifolds. In T. Pajdla & J. Matas (Eds.), *ECCV 2004* (pp. 481–493). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-24673-2_39

68. Rahimi, A., Dunagan, B., & Darrell, T. (2004b). Tracking People with a Sparse Network of Bearing Sensors. In T. Pajdla & J. Matas (Eds.), *ECCV 2004* (pp. 507–518). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-24673-2_41

69. Grauman, K., & Darrell, T. (2004). Fast contour matching using approximate earth mover's distance. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004* (Vol. 1, p. I-220-I-227 Vol.1). http://doi.org/10.1109/CVPR.2004.1315035

70. Rahimi, A., Dunagan, B., & Darrell, T. (2004a). Simultaneous calibration and tracking with a network of non-overlapping sensors. In *Proceedings of the 2004 IEEE Computer Society*

Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004 (Vol. 1, p. I-187-I-194 Vol.1). http://doi.org/10.1109/CVPR.2004.1315031

71. Yeh, T., Tollmar, K., & Darrell, T. (2004). Searching the Web with mobile images for location recognition. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004* (Vol. 2, p. II-76-II-81 Vol.2). http://doi.org/10.1109/CVPR.2004.1315147

72. Morency, L.-P., & Darrell, T. (2004). From Conversational Tooltips to Grounded Discourse: Head poseTracking in Interactive Dialog Systems. In *Proceedings of the 6th International Conference on Multimodal Interfaces* (pp. 32–37). New York, NY, USA: ACM. http://doi.org/10.1145/1027933.1027940

73. Saenko, K., Darrell, T., & Glass, J. R. (2004). Articulatory Features for Robust Visual Speech Recognition. In *Proceedings of the 6th International Conference on Multimodal Interfaces* (pp. 152–158). New York, NY, USA: ACM. http://doi.org/10.1145/1027933.1027960

74. David Demirdjian, D., Wilson, K., Siracusa, M., & Darrell, T. (2004). Real-time audio-visual tracking for meeting analysis. In *Proceedings of the 6th international conference on multimodal interfaces* (pp. 331-332). Retrieved from: https://pal.sri.com/wp-content/uploads/publications/calo/2005/p331-demirdjian.pdf

75. Checka, N., Wilson, K. W., Siracusa, M. R., & Darrell, T. (2004). Multiple person and speaker activity tracking with a particle filter. In *IEEE International Conference on Acoustics, Speech, and Signal Processing, 2004. Proceedings. (ICASSP '04)* (Vol. 5, p. V-881-4 vol.5). http://doi.org/10.1109/ICASSP.2004.1327252

76. Grauman, K., Shakhnarovich, G., & Darrell, T. (2004). Virtual visual hulls: Example-based 3D shape inference from a single silhouette. In *Proceedings of the 2nd Workshop on Statistical Methods in Video Processing*. (pp. 26-37). Springer-Verlag. Available at: http://www.cs.utexas.edu/~grauman/papers/grauman_et_al_smvp2004.pdf

77. Lee, C., Lesh, N., Sidner, C. L., Morency, L.-P., Kapoor, A., & Darrell, T. (2004). Nodding in Conversations with a Robot. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 785–786). New York, NY, USA: ACM. http://doi.org/10.1145/985921.985935

78. Taycher, L., Iii, J. W. F., & Darrell, T. (2004). Combining Simple Models to Approximate Complex Dynamics. In D. Comaniciu, R. Mester, K. Kanatani, & D. Suter (Eds.), *Statistical Methods in Video Processing* (pp. 94–104). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-30212-4_9

79. Tollmar, K., Demirdjian, D., & Darrell, T. (2004). Navigating in Virtual Environments Using a Vision-based Interface. In *Proceedings of the Third Nordic Conference on Human-computer Interaction* (pp. 113–120). New York, NY, USA: ACM. http://doi.org/10.1145/1028014.1028033

80. Tollmar, K., Yeh, T., & Darrell, T. (2004). IDeixis – Searching the Web with Mobile Images for Location-Based Information. In S. Brewster & M. Dunlop (Eds.), *Mobile Human-Computer Interaction - MobileHCI 2004* (pp. 288–299). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-28637-0_25

81. Yeh, T., Tollmar, K., & Darrell, T. (2004). IDeixis: Image-based Deixis for Finding Location-based Information. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (pp. 781–782). New York, NY, USA: ACM. http://doi.org/10.1145/985921.985933

82. Arandjelovic, O., Shakhnarovich, G., Fisher, J., Cipolla, R., & Darrell, T. (2005). Face recognition with image sets using manifold density divergence. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 1, pp. 581–588 vol. 1). http://doi.org/10.1109/CVPR.2005.151

83. Christoudias, C. M., & Darrell, T. (2005). On modelling nonlinear shape-and-texture appearance manifolds. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 2, pp. 1067–1074 vol. 2). http://doi.org/10.1109/CVPR.2005.255

84. Demirdjian, D., Taycher, L., Shakhnarovich, G., Grauman, K., & Darrell, T. (2005). Avoiding the "streetlight effect": tracking by exploring likelihood modes. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1* (Vol. 1, p. 357–364 Vol. 1). http://doi.org/10.1109/ICCV.2005.41

85. Grauman, K., & Darrell, T. (2005a). Efficient image matching with distributions of local invariant features. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 2, pp. 627–634 vol. 2). http://doi.org/10.1109/CVPR.2005.138

86. Grauman, K., & Darrell, T. (2005). The pyramid match kernel: discriminative classification with sets of image features. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1* (Vol. 2, p. 1458–1465 Vol. 2). http://doi.org/10.1109/ICCV.2005.239

87. Morency, L.-P., Sidner, C., Lee, C., & Darrell, T. (2005). Contextual Recognition of Head Gestures. In *Proceedings of the 7th International Conference on Multimodal Interfaces* (pp. 18–24). New York, NY, USA: ACM. http://doi.org/10.1145/1088463.1088470

88. Quattoni, A., Collins, M., & Darrell, T. (2005). Conditional Random Fields for Object Recognition. In L. K. Saul, Y. Weiss, & L. Bottou (Eds.), *Advances in Neural Information Processing Systems 17* (pp. 1097–1104). MIT Press. Retrieved from http://papers.nips.cc/paper/2652-conditional-random-fields-for-object-recognition.pdf

89. Rahimi, A., Darrell, T., & Recht, B. (2005). Learning appearance manifolds from video. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 1, pp. 868–875 vol. 1). http://doi.org/10.1109/CVPR.2005.204

90. Saenko, K., Livescu, K., Siracusa, M., Wilson, K., Glass, J., & Darrell, T. (2005). Visual speech recognition with loosely synchronized feature streams. In *Tenth IEEE International Conference on Computer Vision (ICCV'05) Volume 1* (Vol. 2, p. 1424–1431 Vol. 2). http://doi.org/10.1109/ICCV.2005.251

91. Taycher, L., Fisher III, J. W., & Darrell, T. (2005). Combining Object and Feature Dynamics in Probabilistic Tracking. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 2 - Volume 02* (pp. 106–113). Washington, DC, USA: IEEE Computer Society. http://doi.org/10.1109/CVPR.2005.102

92. Yeh, T., Grauman, K., Tollmar, K., & Darrell, T. (2005). A Picture is Worth a Thousand Keywords: Image-based Object Search on a Mobile Platform. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (pp. 2025–2028). New York, NY, USA: ACM. http://doi.org/10.1145/1056808.1057083

93. Quattoni, A., Collins, M., & Darrell, T. (2005). Conditional Random Fields for Object Recognition. In L. K. Saul, Y. Weiss, & L. Bottou (Eds.), *Advances in Neural Information Processing Systems 17* (pp. 1097–1104). MIT Press. Retrieved from http://papers.nips.cc/paper/2652-conditional-random-fields-for-object-recognition.pdf

94. Quattoni, A., Collins, M., & Darrell, T. (2005). Incorporating Semantic Constraints into a Discriminative Categorization and Labelling Model. In *Tenth IEEE International Conference on Computer Vision Workshops (ICCVW'05)* (pp. 1877–1877). http://doi.org/10.1109/ICCV.2005.256

95. Saenko, K., Livescu, K., Glass, J., & Darrell, T. (2005). Production domain modeling of pronunciation for visual speech recognition. In *Proceedings. (ICASSP '05). IEEE International*

Conference on Acoustics, Speech, and Signal Processing, 2005. (Vol. 5, p. v/473-v/476 Vol. 5). http://doi.org/10.1109/ICASSP.2005.1416343

96. Taycher, L., III, J. W. F., & Darrell, T. (2005). Incorporating Object Tracking Feedback into Background Maintenance Framework. In *Seventh IEEE Workshops on Application of Computer Vision, 2005. WACV/MOTIONS '05 Volume 1* (Vol. 2, pp. 120–125). http://doi.org/10.1109/ACVMOT.2005.63

97. Wilson, K., & Darrell, T. (2005). Improving audio source localization by learning the precedence effect. In *Proceedings. (ICASSP '05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.* (Vol. 4, p. iv/1125-iv/1128 Vol. 4). http://doi.org/10.1109/ICASSP.2005.1416211

98. Yeh, T., & Darrell, T. (2005). Doubleshot: An Interactive User-aided Segmentation Tool. In *Proceedings of the 10th International Conference on Intelligent User Interfaces* (pp. 287–289). New York, NY, USA: ACM. http://doi.org/10.1145/1040830.1040901

99. Grauman, K., & Darrell, T. (2006). Unsupervised Learning of Categories from Sets of Partially Matching Image Features. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* (Vol. 1, pp. 19–25). http://doi.org/10.1109/CVPR.2006.322

100. Morency, L.-P., Christoudias, C. M., & Darrell, T. (2006). Recognizing Gaze Aversion Gestures in Embodied Conversational Discourse. In *Proceedings of the 8th International Conference on Multimodal Interfaces* (pp. 287–294). New York, NY, USA: ACM. http://doi.org/10.1145/1180995.1181051

101. Morency, L.-P., & Darrell, T. (2006). Head Gesture Recognition in Intelligent Interfaces: The Role of Context in Improving Recognition. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (pp. 32–38). New York, NY, USA: ACM. http://doi.org/10.1145/1111449.1111464

102. Taycher, L., Demirdjian, D., Darrell, T., & Shakhnarovich, G. (2006). Conditional Random People: Tracking Humans with CRFs and Grid Filters. In *2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)* (Vol. 1, pp. 222–229). http://doi.org/10.1109/CVPR.2006.83

103. Wang, S. B., Quattoni, A., Morency, L. P., Demirdjian, D., & Darrell, T. (2006). Hidden Conditional Random Fields for Gesture Recognition. In *2006 IEEE Computer Society*

Conference on Computer Vision and Pattern Recognition (CVPR'06) (Vol. 2, pp. 1521–1527). http://doi.org/10.1109/CVPR.2006.132

104. Christoudias, C. M., Saenko, K., Morency, L.-P., & Darrell, T. (2006). Co-Adaptation of Audio-visual Speech and Gesture Classifiers. In *Proceedings of the 8th International Conference on Multimodal Interfaces* (pp. 84–91). New York, NY, USA: ACM. http://doi.org/10.1145/1180995.1181013

105. Grauman, K., & Darrell, T. (2007). Approximate Correspondences in High Dimensions. In B. Schölkopf, J. C. Platt, & T. Hoffman (Eds.), *Advances in Neural Information Processing Systems 19* (pp. 505–512). MIT Press. Retrieved from http://papers.nips.cc/paper/3030-approximate-correspondences-in-high-dimensions.pdf

106. Quattoni, A., Collins, M., & Darrell, T. (2007). Learning Visual Representations using Images with Captions. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). http://doi.org/10.1109/CVPR.2007.383173

107. Grauman, K., & Darrell, T. (2007). Pyramid Match Hashing: Sub-Linear Time Indexing Over Partial Correspondences. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). http://doi.org/10.1109/CVPR.2007.383225

108. Morency, L. P., Quattoni, A., & Darrell, T. (2007). Latent-Dynamic Discriminative Models for Continuous Gesture Recognition. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). http://doi.org/10.1109/CVPR.2007.383299

109. Urtasun, R., & Darrell, T. (2007). Discriminative Gaussian Process Latent Variable Model for Classification. In *Proceedings of the 24th International Conference on Machine Learning* (pp. 927–934). New York, NY, USA: ACM. http://doi.org/10.1145/1273496.1273613

110. Kapoor, A., Grauman, K., Urtasun, R., & Darrell, T. (2007). Active Learning with Gaussian Processes for Object Categorization. In *2007 IEEE 11th International Conference on Computer Vision* (pp. 1–8). http://doi.org/10.1109/ICCV.2007.4408844

111. Quattoni, A., Collins, M., & Darrell, T. (2007). Learning Visual Representations using Images with Captions. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1–8). http://doi.org/10.1109/CVPR.2007.383173

112. Yeh, T., Lee, J., & Darrell, T. (2007). Adaptive Vocabulary Forests for Dynamic Indexing and Category Learning. In *2007 IEEE 11th International Conference on Computer Vision* (pp. 1–8). http://doi.org/10.1109/ICCV.2007.4409053

113. Wang, S. B., Demirdjian, D., & Darrell, T. (2007). Detecting Communication Errors from Visual Cues During the System's Conversational Turn. In *Proceedings of the 9th International Conference on Multimodal Interfaces* (pp. 323–326). New York, NY, USA: ACM. http://doi.org/10.1145/1322192.1322248

114. Morency, L.-P., & Darrell, T. (2007). Conditional Sequence Model for Context-Based Recognition of Gaze Aversion. In A. Popescu-Belis, S. Renals, & H. Bourlard (Eds.), *Machine Learning for Multimodal Interaction* (pp. 11–23). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-78155-4_2

115. Saenko, K., & Darrell, T. (2007). Object Category Recognition Using Probabilistic Fusion of Speech and Image Classifiers. In A. Popescu-Belis, S. Renals, & H. Bourlard (Eds.), *Machine Learning for Multimodal Interaction* (pp. 36–47). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-540-78155-4_4

116. Christoudias, C. M., Urtasun, R., & Darrell, T. (2008). Unsupervised feature selection via distributed coding for multi-view object recognition. In *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008* (pp. 1–8). http://doi.org/10.1109/CVPR.2008.4587615

117. Quattoni, A., Collins, M., & Darrell, T. (2008). Transfer learning for image classification with sparse prototype representations. In *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008* (pp. 1–8). http://doi.org/10.1109/CVPR.2008.4587637

118. Saenko, K., & Darrell, T. (2008). Unsupervised Learning of Visual Sense Models for Polysemous Words. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in Neural Information Processing Systems 21* (pp. 1393–1400). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/3389-unsupervised-learning-of-visual-sense-models-for-polysemous-words.pdf

119. Stone, Z., Zickler, T., & Darrell, T. (2008). Autotagging Facebook: Social network context improves photo annotation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08* (pp. 1–8). http://doi.org/10.1109/CVPRW.2008.4562956

120. Urtasun, R., & Darrell, T. (2008). Sparse probabilistic regression for activity-independent human pose inference. In *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008* (pp. 1–8). http://doi.org/10.1109/CVPR.2008.4587360

121. Urtasun, R., Fleet, D. J., Geiger, A., Popović, J., Darrell, T. J., & Lawrence, N. D. (2008). Topologically-constrained Latent Variable Models. In *Proceedings of the 25th International Conference on Machine Learning* (pp. 1080–1087). New York, NY, USA: ACM. http://doi.org/10.1145/1390156.1390292

122. Yeh, T., & Darrell, T. (2008). Dynamic visual category learning. In *IEEE Conference on Computer Vision and Pattern Recognition, 2008. CVPR 2008* (pp. 1–8). http://doi.org/10.1109/CVPR.2008.4587616

123. Yeh, T., Lee, J. J., & Darrell, T. (2008). Photo-based Question Answering. In *Proceedings of the 16th ACM International Conference on Multimedia* (pp. 389–398). New York, NY, USA: ACM. http://doi.org/10.1145/1459359.1459412

124. Yeh, T., & Darrell, T. (2008). Multimodal Question Answering for Mobile Devices. In *Proceedings of the 13th International Conference on Intelligent User Interfaces* (pp. 405–408). New York, NY, USA: ACM. http://doi.org/10.1145/1378773.137884

125. Yeh, T., Lee, J. J., & Darrell, T. (2008). Scalable classifiers for Internet vision tasks. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08* (pp. 1–8). http://doi.org/10.1109/CVPRW.2008.4562958

126. Christoudias, C. M., Urtasun, R., Kapoorz, A., & Darrell, T. (2009). Co-training with noisy perceptual observations. In *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009* (pp. 2844–2851). http://doi.org/10.1109/CVPR.2009.5206572

127. Geiger, A., Urtasun, R., & Darrell, T. (2009). Rank priors for continuous non-linear dimensionality reduction. In *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009* (pp. 880–887). http://doi.org/10.1109/CVPR.2009.5206672

128. Frampton, M., Fernández, R., Ehlen, P., Christoudias, M., Darrell, T., & Peters, S. (2009). Who is "You"?: Combining Linguistic and Gaze Features to Resolve Second-person References in Dialogue. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics* (pp. 273–281). Stroudsburg, PA, USA: Association for Computational Linguistics. Retrieved from http://dl.acm.org/citation.cfm?id=1609067.1609097

129. Quattoni, A., Carreras, X., Collins, M., & Darrell, T. (2009). An Efficient Projection for L_1, ∞ Regularization. In *Proceedings of the 26th Annual International Conference on Machine Learning* (pp. 857–864). New York, NY, USA: ACM. http://doi.org/10.1145/1553374.1553484

130. Fritz, M., Bradski, G., Karayev, S., Darrell, T., & Black, M. J. (2009a). An Additive Latent Feature Model for Transparent Object Recognition. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22* (pp. 558–566). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/3808-an-additive-latent-feature-model-for-transparent-object-recognition.pdf

131. Fritz, M., Bradski, G., Karayev, S., Darrell, T., & Black, M. J. (2009b). An Additive Latent Feature Model for Transparent Object Recognition. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22* (pp. 558–566). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/3808-an-additive-latent-feature-model-for-transparent-object-recognition.pdf

132. Brian Kulis, B. & Darrell, T. (2009). Learning to hash with binary reconstructive embeddings**.** In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22 (NIPS 2009)* (pp. 1042-1050). http://papers.nips.cc/paper/3667-learning-to-hash-with-binary-reconstructive-embeddings

133. Saenko, K., & Darrell, T. (2009a). Filtering Abstract Senses From Image Search Results. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 22* (pp. 1589–1597). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/3860-filtering-abstract-senses-from-image-search-results.pdf

134. Yang, A.Y., Maji, S., Christoudias, C.M., Darrell, T., Malik, J. & Sastry, S.S. (2009) Multiple-view object recognition in band-limited distributed camera networks (pp. 1-8). *Third ACM/IEEE International Conference on Distributed Smart Cameras*. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5289410&isnumber=5289340

135. Yeh, T., Lee, J. J., & Darrell, T. (2009). Fast concurrent object localization and recognition. In *IEEE Conference on Computer Vision and Pattern Recognition, 2009. CVPR 2009* (pp. 280–287). http://doi.org/10.1109/CVPR.2009.5206805

136. Salzmann, M., Ek, C. H., Urtasun, R., & Darrell, T. (2010). Factorized orthogonal latent spaces. In *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS)*. Sardinia, Italy: JMRL. Retrieved from http://www.jmlr.org/proceedings/papers/v9/salzmann10a/salzmann10a.pdf

137. Christoudias, C. M., Urtasun, R., Salzmann, M., & Darrell, T. (2010). Learning to Recognize Objects from Unseen Modalities. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010* (pp. 677–691). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-15549-9_49

138. Friedland, G., Vinyals, O., & Darrell, T. (2010). Multimodal location estimation. In, *Proceedings of the 18th ACM international conference on Multimedia* (pp. 1245-1252). http://dl.acm.org/citation.cfm?id=1874197

139. Saenko, K., Kulis, B., Fritz, M., & Darrell, T. (2010). Adapting Visual Category Models to New Domains. In K. Daniilidis, P. Maragos, & N. Paragios (Eds.), *Computer Vision – ECCV 2010* (pp. 213–226). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-15561-1_16

140. Fritz, M., Saenko, K., & Darrell, T. (2010). Size Matters: Metric Visual Search Constraints from Monocular Metadata. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 23* (pp. 622–630). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/4104-size-matters-metric-visual-search-constraints-from-monocular-metadata.pdf

141. Jia, Y., Salzmann, M., & Darrell, T. (2010). Factorized Latent Spaces with Structured Sparsity. In J. D. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel, & A. Culotta (Eds.), *Advances in Neural Information Processing Systems 23* (pp. 982–990). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/3953-factorized-latent-spaces-with-structured-sparsity.pdf

142. Miller, S., Fritz, M., Darrell, T., & Abbeel, P. (2011). Parametrized shape models for clothing. In *2011 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 4861–4868). http://doi.org/10.1109/ICRA.2011.5980453

143. Shyr, A., Darrell, T., Jordan, M., & Urtasun, R. (2011). Supervised hierarchical Pitman-Yor process for natural scene segmentation. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2281–2288). http://doi.org/10.1109/CVPR.2011.5995647

144. Kulis, B., Saenko, K., & Darrell, T. (2011). What you saw is not what you get: Domain adaptation using asymmetric kernel transforms. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 1785–1792). http://doi.org/10.1109/CVPR.2011.5995702

145. Owens, T., Saenko, K., Chakrabarti, A., Xiong, Y., Zickler, T., & Darrell, T. (2011). Learning object color models from multi-view constraints. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 169–176). http://doi.org/10.1109/CVPR.2011.5995705

146. Karayev, S., Fritz, M., Fidler, S., & Darrell, T. (2011). A probabilistic model for recursive factorized image features. In *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 401–408). http://doi.org/10.1109/CVPR.2011.5995728

147. Janoch, A., Karayev, S., Jia, Y., Barron, J. T., Fritz, M., Saenko, K., & Darrell, T. (2011). A category-level 3-D object dataset: Putting the Kinect to work. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)* (pp. 1168–1174). http://doi.org/10.1109/ICCVW.2011.6130382

148. Jia, Y., & Darrell, T. (2011). Heavy-tailed Distances for Gradient Based Image Descriptors. In J. Shawe-Taylor, R. S. Zemel, P. L. Bartlett, F. Pereira, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 24* (pp. 397–405). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/4183-heavy-tailed-distances-for-gradient-based-image-descriptors.pdf

149. Farrell, R., Oza, O., Zhang, N., Morariu, V. I., Darrell, T., & Davis, L. S. (2011). Birdlets: Subordinate categorization using volumetric primitives and pose-normalized appearance. In *2011 International Conference on Computer Vision* (pp. 161–168). http://doi.org/10.1109/ICCV.2011.6126238

150. Tuytelaars, T., Fritz, M., Saenko, K., & Darrell, T. (2011). The NBNN kernel. In *2011 International Conference on Computer Vision* (pp. 1824–1831). http://doi.org/10.1109/ICCV.2011.6126449

151. Yangqing Jia, Y., Mathieu Salzmann, M., & Darrell, T. (2011) Learning cross-modality similarity for multinomial data. In *IEEE International Conference on Computer Vision, 2011* (pp. 2407-2414). http://doi.ieeecomputersociety.org/10.1109/ICCV.2011.6126524

152. Song, H. O., Fritz, M., Gu, C., & Darrell, T. (2011). Visual grasp affordances from appearance-based cues. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)* (pp. 998–1005). http://doi.org/10.1109/ICCVW.2011.6130360

153. Wang, P. C., Miller, S., Fritz, M., Darrell, T., & Abbeel, P. (2011). Perception for the manipulation of socks. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 4877–4884). http://doi.org/10.1109/IROS.2011.6095013

154. Saenko, K., Karayev, S., Jia, Y., Shyr, A., Janoch, A., Long, J., … Darrell, T. (2011). Practical 3-D Object detection using category and instance-level appearance models. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 793–800). http://doi.org/10.1109/IROS.2011.6095000

155. Eden, A., Christoudias, C. M., & Darrell, T. (2011). Finding lost children. In *2011 IEEE Workshop on Person-Oriented Vision (POV)* (pp. 7–12). http://doi.org/10.1109/POV.2011.5712362

156. Karayev, S., Baumgartner, T., Fritz, M., & Darrell, T. (2012). Timely Object Recognition. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25* (pp. 890–898). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/4712-timely-object-recognition.pdf

157. Vinyals, O., Jia, Y., Deng, L., & Darrell, T. (2012). Learning with Recursive Perceptual Representations. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 25* (pp. 2825–2833). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/4747-learning-with-recursive-perceptual-representations.pdf

158. Hoffman, J., Kulis, B., Darrell, T., & Saenko, K. (2012). Discovering Latent Domains for Multisource Domain Adaptation. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), *Computer Vision – ECCV 2012* (pp. 702–715). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-33709-3_50

159. Jia, Y., Huang, C., & Darrell, T. (2012). Beyond spatial pyramids: Receptive field learning for pooled image features. 2012 IEEE Conference on Computer Vision and Pattern Recognition (pp. 3370-3377). http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6248076&isnumber=6247647

160. Song, H. O., Zickler, S., Althoff, T., Girshick, R., Fritz, M., Geyer, C., … Darrell, T. (2012). Sparselet Models for Efficient Multiclass Object Detection. In A. Fitzgibbon, S. Lazebnik, P. Perona, Y. Sato, & C. Schmid (Eds.), *Computer Vision – ECCV 2012* (pp. 802–815). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-33709-3_57

161. Virtanen, S., Jia, Y., Klami, A., & Darrell, T. (2012). Factorized Multi-Modal Topic Model. In *Proceedings of the Twenty-Eighth Conference on Uncertainty in Artificial Intelligence (UAI2012)*. Catalina Island, United States. Retrieved from http://arxiv.org/abs/1210.4920

162. Xiong, Y., Saenko, K., Darrell, T., & Zickler, T. (2012). From pixels to physics: Probabilistic color de-rendering. In *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 358–365). http://doi.org/10.1109/CVPR.2012.6247696

163. Zhang, N., Farrell, R., & Darrell, T. (2012). Pose pooling kernels for sub-category recognition. In *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3665–3672). http://doi.org/10.1109/CVPR.2012.6248364

164. Althoff, T., Song, H. O., & Darrell, T. (2012). Detection Bank: An Object Detection Based Video Representation for Multimedia Event Recognition. In *Proceedings of the 20th ACM International Conference on Multimedia* (pp. 1065–1068). New York, NY, USA: ACM. http://doi.org/10.1145/2393347.2396384

165. Chu, V., McMahon, I., Riano, L., McDonald, C. G., He, Q., Perez-Tejada, J. M., … Kuchenbecker, K. J. (2013a). Using robotic exploratory procedures to learn the meaning of haptic adjectives. In *2013 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 3048–3055). http://doi.org/10.1109/ICRA.2013.6631000

166. Hoffman, J., Rodner, E., Donahue, J., Darrell, T., & Saenko, K. (2013). Efficient Learning of Domain-invariant Image Representations. In *Conference Proceedings International Conference on Learning Representations (ICLR) 2013.* Scottsdale, Arizona. Retrieved from http://arxiv.org/abs/1301.3224

167. Donahue, J., Hoffman, J., Rodner, E., Saenko, K., & Darrell, T. (2013a). Semi-supervised Domain Adaptation with Instance Constraints. In *2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 668–675). http://doi.org/10.1109/CVPR.2013.92

168. Jia, Y., Vinyals, O., & Darrell, T. (2013b). On Compact Codes for Spatially Pooled Features (pp. 549–557). Presented at the Proceedings of The 30th International Conference on Machine Learning. Retrieved from http://www.jmlr.org/proceedings/papers/v28/jia13.html

169. Girshick, R., Song, H. O., & Darrell, T. (2013). Discriminatively Activated Sparselets (pp. 196–204). Presented at the Proceedings of The 30th International Conference on Machine Learning. Retrieved from http://www.jmlr.org/proceedings/papers/v28/girshick13.html

170. Guadarrama, S., Krishnamoorthy, N., Malkarnenkar, G., Venugopalan, S., Mooney, R., Darrell, T., & Saenko, K. (2013). YouTube2Text: Recognizing and Describing Arbitrary Activities Using Semantic Hierarchies and Zero-Shot Recognition. In *2013 IEEE International Conference on Computer Vision* (pp. 2712–2719). http://doi.org/10.1109/ICCV.2013.337

171. Guadarrama, S., Riano, L., Golland, D., Go¨hring, D., Jia, Y., Klein, D., … Darrell, T. (2013). Grounding spatial relations for human-robot interaction. In *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems* (pp. 1640–1647). http://doi.org/10.1109/IROS.2013.6696569

172. Jia, Y., Abbott, J. T., Austerweil, J., Griffiths, T., Darrell, T., Burges, C. J. C., … Weinberger, K. Q. (2013). Visual Concept Learning: Combining Machine Vision and Bayesian Generalization on Concept Hierarchies. Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/5205-visual-concept-learning-combining-machine-vision-and-bayesian-generalization-on-concept-hierarchies.pdf

173. Jia, Y., & Darrell, T. (2013). Latent Task Adaptation with Large-Scale Hierarchies. In *2013 IEEE International Conference on Computer Vision* (pp. 2080–2087). http://doi.org/10.1109/ICCV.2013.260

174. Zhang, N., Farrell, R., Iandola, F., & Darrell, T. (2013). Deformable Part Descriptors for Fine-Grained Recognition and Attribute Prediction (pp. 729–736). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2013/html/Zhang_Deformable_Part_Descriptors_2013_ICCV_paper.html

175. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E., & Darrell, T. (2014). DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition. (p. 647–655) In *Proceedings of the 31st International Conference on Machine Learning* (Vol. 32). Beijing, China: Retrieved from http://www.jmlr.org/proceedings/papers/v32/donahue14.pdf

176. Feng, J., Jegelka, S., Yan, S., & Darrell, T. (2014). Learning Scalable Discriminative Dictionary with Sample Relatedness (pp. 1645–1652). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2014/html/Feng_Learning_Scalable_Discriminative_2014_CVPR_paper.html

177. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation (pp. 580–587). Presented at the

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2014/html/Girshick_Rich_Feature_Hierarchies_2014_CVPR_paper.html

178. Goehring, D., Hoffman, J., Rodner, E., Saenko, K., & Darrell, T. (2014). Interactive adaptation of real-time object detectors. In *2014 IEEE International Conference on Robotics and Automation (ICRA)* (pp. 1282–1289). http://doi.org/10.1109/ICRA.2014.6907018

179. Hoffman, J., Darrell, T., & Saenko, K. (2014). Continuous Manifold Based Adaptation for Evolving Visual Domains. In *2014 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 867–874). http://doi.org/10.1109/CVPR.2014.116

180. Hoffman, J., Guadarrama, S., Tzeng, E. S., Hu, R., Donahue, J., Girshick, R., … Saenko, K. (2014). LSDA: Large Scale Detection through Adaptation. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 27* (pp. 3536–3544). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/5418-lsda-large-scale-detection-through-adaptation.pdf

181. Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama. S., Darrell, T., Trevor Darrell (2014). Caffe: Convolutional Architecture for Fast Feature Embedding. In *Proceedings of the 22Nd ACM International Conference on Multimedia* (pp. 675–678). New York, NY, USA: ACM. http://doi.org/10.1145/2647868.2654889

182. Karayev, S., Fritz, M., & Darrell, T. (2014). Anytime Recognition of Objects and Scenes (pp. 572–579). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2014/html/Karayev_Anytime_Recognition_of_2014_CVPR_paper.html

183. Long, J. L., Zhang, N., & Darrell, T. (2014). Do Convnets Learn Correspondence? In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 27* (pp. 1601–1609). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/5420-do-convnets-learn-correspondence.pdf

184. Song, H. O., Girshick, R., Mairal, J., Harchaoui, Z., & Darrell, T. (2014). On learning to localize objects with minimal supervision. In *Proceedings of the 31st International Conference on Machine Learning* (Vol. 32). Beijing, China: JMRL. Retrieved from http://www.jmlr.org/proceedings/papers/v32/songb14.pdf

185. Song, H. O., Lee, Y. J., Jegelka, S., & Darrell, T. (2014). Weakly-supervised Discovery of Visual Pattern Configurations. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems 27* (pp. 1637–1645). Curran Associates, Inc. Retrieved from http://papers.nips.cc/paper/5284-weakly-supervised-discovery-of-visual-pattern-configurations.pdf

186. Zhang, N., Paluri, M., Ranzato, M., Darrell, T., & Bourdev, L. (2014). PANDA: Pose Aligned Networks for Deep Attribute Modeling (pp. 1637–1644). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2014/html/Zhang_PANDA_Pose_Aligned_2014_CVPR_paper.html

187. Freytag, A., Rodner, E., Darrell, T., & Denzler, J. (2014). Exemplar-Specific Patch Features for Fine-Grained Recognition. In X. Jiang, J. Hornegger, & R. Koch (Eds.), *Proc. GCPR* (pp. 144–156). Springer International Publishing. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-11752-2_12

188. Zhang, N., Donahue, J., Girshick, R., & Darrell, T. (2014). Part-Based R-CNNs for Fine-Grained Category Detection. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.), *Computer Vision – ECCV 2014* (pp. 834–849). Springer International Publishing. Retrieved from http://link.springer.com/chapter/10.1007/978-3-319-10590-1_54

189. Guadarrama, S., Rodner, E., Saenko, K., Zhang, N., Farrell, R., Donahue, J., & Darrell, T.. (2014). Open-vocabulary Object Retrieval (pp. 1-9). Presented at Robotics Systems and Science 2014. Retrieved from https://people.eecs.berkeley.edu/~nzhang/papers/rss14_openvoc.pdf

190. Donahue, J., Anne Hendricks, L., Guadarrama, S., Rohrbach, M., Venugopalan, S., Saenko, K., & Darrell, T. (2015). Long-Term Recurrent Convolutional Networks for Visual Recognition and Description (pp. 2625–2634). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Donahue_Long-Term_Recurrent_Convolutional_2015_CVPR_paper.html

191. Feng, J., & Darrell, T. (2015). Learning The Structure of Deep Convolutional Networks (pp. 2749–2757). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2015/html/Feng_Learning_The_Structure_ICCV_2015_paper.html

192. Girshick, R., Iandola, F., Darrell, T., & Malik, J. (2015). Deformable Part Models are Convolutional Neural Networks (pp. 437–446). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Girshick_Deformable_Part_Models_2015_CVPR_paper.html

193. Hoffman, J., Pathak, D., Darrell, T., & Saenko, K. (2015). Detector Discovery in the Wild: Joint Multiple Instance and Representation Learning (pp. 2883–2891). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Hoffman_Detector_Discovery_in_2015_CVPR_paper.html

194. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully Convolutional Networks for Semantic Segmentation (pp. 3431–3440). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2015/html/Long_Fully_Convolutional_Networks_2015_CVPR_paper.html

195. Mrowca, D., Rohrbach, M., Hoffman, J., Hu, R., Saenko, K., & Darrell, T. (2015). Spatial Semantic Regularisation for Large Scale Object Detection (pp. 2003–2011). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2015/html/Mrowca_Spatial_Semantic_Regularisation_ICCV_2015_paper.html

196. Pathak, D., Krahenbuhl, P., & Darrell, T. (2015). Constrained Convolutional Neural Networks for Weakly Supervised Segmentation (pp. 1796–1804). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2015/html/Pathak_Constrained_Convolutional_Neural_ICCV_2015_paper.html

197. Shelhamer, E., Barron, J. T., & Darrell, T. (2015). Scene Intrinsics and Depth From a Single Image (pp. 37–44). Presented at the Proceedings of the IEEE International Conference on Computer Vision Workshops. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2015_workshops/w10/html/Shelhamer_Scene_Intrinsics_and_ICCV_2015_paper.html

198. Tzeng, E., Hoffman, J., Darrell, T., & Saenko, K. (2015). Simultaneous Deep Transfer Across Domains and Tasks (pp. 4068–4076). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2015/html/Tzeng_Simultaneous_Deep_Transfer_ICCV_2015_paper.html

199. Venugopalan, S., Rohrbach, M., Donahue, J., Mooney, R., Darrell, T., & Saenko, K. (2015). Sequence to Sequence - Video to Text (pp. 4534–4542). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://www.cv-foundation.org/openaccess/content_iccv_2015/html/Venugopalan_Sequence_to_Sequence_ICCV_2015_paper.html

200. Finn, C., Tan, X. Y., Duan, Y., Darrell, T., Levine, S., & Abbeel, P. (2016) Deep Spatial Autoencoders for Visuomotor Learning (pp 512-519). Published in the International Conference on Robotics and Automation (ICRA). http://dx.doi.org/10.1109/ICRA.2016.7487173

201. Hoffman, J., Gupta, S., Leong, J., Guadarrama, S., & Darrell, T. (2016). Cross-modal adaptation for RGB-D detection. In *2016* IEEE International Conference on Robotics and Automation (ICRA) (pp. 5032–5039). http://doi.org/10.1109/ICRA.2016.7487708

202. Murali, A., Garg, A., Krishnan, S., Pokorny, F. T., Abbeel, P., Darrell, T., & Goldberg, K. (2016). TSC-DL: Unsupervised trajectory segmentation of multi-modal surgical demonstrations with Deep Learning. In 2016 IEEE International Conference on Robotics and Automation (ICRA) (pp. 4150–4157). http://doi.org/10.1109/ICRA.2016.7487607

203. Gao, Y., Beijbom, O., Zhang, N., & Darrell, T. (2016). Compact Bilinear Pooling (pp. 318-326). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Gao_Compact_Bilinear_Pooling_CVPR_2016_paper.pdf

204. Hu, R., Xu, H., Rohrbach, M., Feng, J., Saenko, K., & Darrell, T. (2016). Natural Language Object Retrieval (pp. 4555-4564). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Hu_Natural_Language_Object_CVPR_2016_paper.pdf

205. Andreas, J., Rohrbach, M., Darrell, T., & Klein, D. (2016). Neural Module Networks (pp. 39–48). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern

Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Andreas_Neural_Module_Networks_CVPR_2016_paper.html

206. Hendricks, L.A., Venugopalan, S., Rohrbach, M., Mooney, R., Saenko, K., & Darrell, T. (2016). Deep Compositional Captioning: Describing Novel Object Categories Without Paired Training Data (pp. 1–10). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Hendricks_Deep_Compositional_Captioning_CVPR_2016_paper.html

207. Andreas, J., Rohrbach, M., Darrell, T., & Klein, D. (2016). Learning to Compose Neural Networks for Question Answering (pp. 1545-1554). Presented at North American Assoc. Computational Linguistics (NAACL). Retrieved from http://www.aclweb.org/anthology/N16-1181

208. Pathak, D., Krahenbuhl, P., Donahue, J, Darrell, T., and Efros., A., (2016). Feature learning by inpainting, (pp. 2536-2544). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from: https://arxiv.org/abs/1604.07379

209. Hoffman. J., Gupta, S., Darrell, T. (2016). Learning with side information through modality hallucination, (pp. 826-834), Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from: https://doi.org/10.1109/CVPR.2016.96

210. Gao, Y., Hendricks, L.A., Kuchenbecker, K.J., & Darrell, T. (2016). Deep Learning for Tactile Understanding from Visual and Haptic Data (pp. 536-543). In 2016 IEEE International Conference on Robotics and Automation (ICRA). Available at: https://arxiv.org/abs/1511.06065

211. Burka, A., Hu, S., Helgeson, S., Krishnan, S., Gao, Y., Hendricks, L.A., Darrell, T., & Kuchenbecker, K.J., (2016). Proton: A Visuo-Haptic Data Acquisition System for Robotic Learning of Surface Properties (pp. 439-445). In 2016 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI). https://doi.org/10.1109/MFI.2016.7849467

212. Burka, A., Hu, S., Helgeson, S., Krishnan, S., Gao, Y., Hendricks, L.A., Darrell, T., & Kuchenbecker, K.J., Design and implementation of a visuo-haptic data acquisition system for

robotic learning of surface properties, (pp. 350-352). Proc. IEEE Haptics Symposium Retrieved from: https://www.alexburka.com/pubs/Proton_Haptics16_wip.pdf

213. Hendricks, L.A., Akata, Z., Rohrbach, M., Donahue, J., Schiele, B., & Darrell, T. (2016). Generating Visual Explanations (pp. 3-19). Presented at the European Conference on Computer Vision. Retrieved from https://doi.org/10.1007/978-3-319-46493-0_1

214. Rohrbach, A., Rohrback, M., Hu, R., Darrell, T., & Schiele, B. (2016). Grounding of Textual Phrases in Images by Reconstruction (pp. 817-834). Presented at the European Conference on Computer Vision. Retrieved from: https://arxiv.org/abs/1511.03745

215. Hu, R., Rohrback, M., & Darrell, T. (2016) Segmentation from Natural Language Expressions. (pp. 108-124). Presented at the European Conference on Computer Vision. Retrieved from: https://arxiv.org/abs/1603.06180

216. Donahue, J., Krähenbühl, P., & Darrell, T. (2017). Adversarial Feature Learning. ICLR 2017 Conference Presentation. Retrieved from http://arxiv.org/abs/1605.09782

217. Venugopalan, S., Hendricks, L.A., Rohrback, M., Mooney, R., Darrell, T., & Saenko, K. (2017). Captioning Images with Diverse Objects (pp.5753-5761). Presented at the Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Retrieved fromhttp://openaccess.thecvf.com/content_cvpr_2017/papers/Venugopalan_Captioning_Images_With_CVPR_2017_paper.pdf

218. Hu, R., Rohrbach, M., Andreas, J., Darrell, T., & Saenko, K. (2017). Modeling Relationships in Referential Expressions with Compositional Modular Networks (pp. 1115-1124). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://openaccess.thecvf.com/content_cvpr_2017/papers/Hu_Modeling_Relationships_in_CVPR_2017_paper.pdf

219. Xu, H., Gao, Y., Yu, F., & Darrell, T. (2017). End-to-End Learning of Driving Models from Large-Scale Video Datasets (pp. 2174-2182). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from https://arxiv.org/abs/1612.01079

220. Pathak, D., Girschick, R., Dollár, P., Darrell, T., & Hariharan, B. (2017). Learning Features by Watching Objects Move (pp. 2701-2710). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from: https://arxiv.org/abs/1612.06370

221. Tzeng, E., Hoffman, J., Saenko, K., & Darrell, T. (2017) Adversarial Discriminative Domain Adaptation (pp 7167-7176). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://openaccess.thecvf.com/content_cvpr_2017/papers/Tzeng_Adversarial_Discriminative_Domain_CVPR_2017_paper.pdf

222. Azadi, S., Feng, J., & Darrell, T. (2017) Learning Detection with Diverse Proposals (pp. 7149-7157). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://openaccess.thecvf.com/content_cvpr_2017/papers/Azadi_Learning_Detection_With_CVPR_2017_paper.pdf

223. Hu, R., Andreas, J., Rohrbach, M., Darrell, T., & Saenko, K. (2017). Learning to Reason: End-to-End Module Networks for Visual Question Answering (pp.804-813) Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://openaccess.thecvf.com/content_ICCV_2017/papers/Hu_Learning_to_Reason_ICCV_2017_paper.pdf

224. Simon, M., Gao, Y., Darrell, T., Denzler, J., & Rodner, E. (2017). Generalized Orderless Pooling Performs Implicit Salient Matching (pp. 4960 – 4969). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://openaccess.thecvf.com/content_ICCV_2017/papers/Simon_Generalized_Orderless_Pooling_ICCV_2017_paper.pdf

225. Hendricks, L.A., Wang, O., Schechtman, E., Sivic, J., Darrell, T., & Russell, B. (2017). Localizing Moments in Video with Natural Language (pp. 5803-5812). Presented at the Proceedings of the IEEE International Conference on Computer Vision. Retrieved from http://openaccess.thecvf.com/content_ICCV_2017/papers/Hendricks_Localizing_Moments_in_ICCV_2017_paper.pdf

226. Devin, C., Gupta, A., Darrell, T., Abbeel, P., & Levine, S. (2017). Learning Modular Neural Network Policies for Multi-Task and Multi-Robot Transfer (pp. 2169-2176). In 2017 IEEE International Conference on Robotics and Automation (ICRA). Available at: https://arxiv.org/abs/1609.07088

227. Zhu, J.-Y., Zhang, R., Pathak, D., Darrell, T., Efros, A.A, Wang, O., & Schechtman, E. (2017). Toward Multimodal Image-to-Image Translation (pp.465-476). In Advances in Neural Information Processing Systems. Retrieved from https://papers.nips.cc/paper/6650-toward-multimodal-image-to-image-translation.pdf

228. Pathak, D., Mahmoudieh, P., Luo, G., Agrawal, P., Chen, D., Shentu, Y., Shelhamer, E., Malik, J., Efros, A., Darrell, T. (2018) Zero-shot Visual Imitation. ICLR 2018 Conference Presentation. Retrieved from: https://arxiv.org/abs/1804.08606

229. Yu, F., Wang, D., Shelhamer, E., & Darrell, T. (2018). Deep Layer Aggregation (pp. 2403-2412). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://openaccess.thecvf.com/content_cvpr_2018/papers/Yu_Deep_Layer_Aggregation_CVPR_2018_paper.pdf

230. Hu, R., Dollár, P., He, K., Darrell, T., & Girshick, R. (2018). Learning to Segment Every Thing (pp. 4233-4241). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Retrieved from http://openaccess.thecvf.com/content_cvpr_2018/papers/Hu_Learning_to_Segment_CVPR_2018_paper.pdf

231. Azadi, S., Fisher, M., Kim, V.G, Wang, Z., Shechtman, E., & Darrell, T. (2018). Multi-Content GAN for Few-Shot Front Style Transfer (pp. 7564-7573). Presented in the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Referenced from http://openaccess.thecvf.com/content_cvpr_2018/papers/Azadi_Multi-Content_GAN_for_CVPR_2018_paper.pdf

232. Xu, X., Chen, X., Liu, C., Rohrbach, A., Darrell, T., & Song, D. (2018). Fooling Vision and Language Models Despite Localization and Attention Mechanism (pp. 4951-4961). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Referenced from http://openaccess.thecvf.com/content_cvpr_2018/papers/Xu_Fooling_Vision_and_CVPR_2018_paper.pdf

233. Park, D.H., Hendricks, L.A., Akata, Z., Rohrbach, A., Schiele, B., Darrell, T., & Rohrbach, M. (2018). Multimodal Explanations: Justifying Decisions and Pointing to the Evidence (pp.8779-8788). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Referenced from http://openaccess.thecvf.com/content_cvpr_2018/papers/Park_Multimodal_Explanations_Justifying_CVPR_2018_paper.pdf

234. Devin, C., Abbeel, P., Darrell, T., & Levine, S. (2018). Deep Object-Centric Representations for Generalizable Robot Learning (pp. 7111-7118). In 2018 IEEE International Conference on Robotics and Automation (ICRA). Referenced from https://arxiv.org/pdf/1708.04225 https://doi.org/10.1109/ICRA.2018.8461196

235. Hendricks, L.A., Hu, R., Darrell, T., & Akata, Z. (2018). Grounding Visual Explanations (pp.264-279). Presented at the Proceedings of the European Conference on Computer Vision (ECCV). Retrieved from http://openaccess.thecvf.com/content_ECCV_2018/papers/Lisa_Anne_Hendricks_Grounding_Visual_Explanations_ECCV_2018_paper.pdf

236. Hu, R., Andreas, J., Darrell, T., & Saenko, K. (2018). Explainable Neural Computation Via Stack Neural Module Networks (pp. 53-69). Presented at the Proceedings of the European Conference on Computer Vision (ECCV). Reference from http://openaccess.thecvf.com/content_ECCV_2018/papers/Ronghang_Hu_Explainable_Neural_Computation_ECCV_2018_paper.pdf

237. Kim, J., Rohrbach, A., Darrell, T., Canny, J., & Akata, Z. (2018). Textual Explanations for Self-Driving Vehicles (pp. 563-578). Presented at the Proceeding of the European Conference on Computer Vision (ECCV). Referenced from http://openaccess.thecvf.com/content_ECCV_2018/papers/Jinkyu_Kim_Textual_Explanations_for_ECCV_2018_paper.pdf

238. Wang, X., Yu, F., Dou, Z.-Y., Darrell, T., & Gonzalez, J.E. (2018). Skipnet: Learning Dynamic Routing in Convolutional Networks (pp. 409-424). Presented at the Proceedings of the European Conference on Computer Vision (ECCV). Referenced from http://openaccess.thecvf.com/content_ECCV_2018/papers/Xin_Wang_SkipNet_Learning_Dynamic_ECCV_2018_paper.pdf

239. Xiao, C., Deng, R., Li, B., Yu, F., Liu, M., & Song, D. (2018). Characterizing Adversarial Examples Based on Spatial Consistency Information for Semantic Segmentation (pp.217-234). Presented at the Proceedings of the European Conference on Computer Vision (ECCV). Referenced from http://openaccess.thecvf.com/content_ECCV_2018/papers/CHAOWEI_XIAO_Characterize_Adversarial_Examples_ECCV_2018_paper.pdf

240. Hendricks, L.A., Burns, K., Saenko, K., Darrell, T., & Rohrbach, A. (2018). Women Also Snowboard: Overcoming Bias in Caption Models (pp. 793-811). In European Conference on Computer Vision. Referenced from https://doi.org/10.1007/978-3-030-01219-9_47

241. Fried, D., Hu, R., Cirik, V., Rohrback, A., Andreas, J., Morency, L.-P., Berg-Kirkpatrick, T., Saenko, K., Klein, D., & Darrell, T. (2018). Speaker-Follower Models for Vision-and-Language Navigation (pp. 3314-3325). In Advances in Neural Information Processing Systems. Referenced from https://papers.nips.cc/paper/7592-speaker-follower-models-for-vision-and-language-navigation.pdf

242. Luo, Y., Xu, H., Li, Y., Tian, Y., Darrell, T., & Ma, T. (2019). Algorithmic Framework for Model-Based Deep Reinforcement Learning with Theoretical Guarantees. ICLR 2019 Conference Presentation. Retrieved from: https://arxiv.org/abs/1807.03858

243. Liu, Z., Sun, M., Zhou, T., Huang, G., & Darrell, T. (2019). Rethinking the Value of Networking Pruning. ICLR 2019 Conference Presentation. Retrieved from https://arxiv.org/abs/1810.05270

244. Azadi, S., Olsson, C., Darrell, T., Goodfellow, I., & Odena, A. (2019). Discriminator Rejection Sampling.]. ICLR 2019 Conference Presentation. Retrieved from https://arxiv.org/abs/1810.06758

245. Burda, Y., Edwards, H., Pathak, D., Storkey, A., Darrell, T., & Efros, A.A. (2019). Large-Scale Study of Curiosity-Driven Learning. ICLR 2019 Conference Presentation. Retrieved from https://arxiv.org/abs/1808.04355

246. Schonfeld, E., Ebrahimi, S., Sinha, S., Darrell, T., & Akata, Z. (2019). Generalized Zero-and Few-Shot Learning Via Aligned Variational Autoencoders (pp. 8247-8255). Presented in the Proceedings of the IEEE Conference on Computer Vision and Patter Recognition. Referenced from http://openaccess.thecvf.com/content_CVPR_2019/papers/Schonfeld_Generalized_Zero-_and_Few-Shot_Learning_via_Aligned_Variational_Autoencoders_CVPR_2019_paper.pdf

247. Yin, Z., Darrell, T., Yu, F. (2019). Hierarchical Discrete Distribution Decomposition for Match Density Estimation (pp. 6044-6053). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Referenced from http://openaccess.thecvf.com/content_CVPR_2019/papers/Yin_Hierarchical_Discrete_Distribution_Decomposition_for_Match_Density_Estimation_CVPR_2019_paper.pdf

248. Park, J.S., Rohrbach, M., Darrell, T., & Rohrbach, A. (2019). Adversarial Inference for Multi-Sentence Video Description (pp. 6598-6608). Presented at the Proceedings of the IEEE Conference on Computer Vision and Patter Recognition. Referenced from http://openaccess.thecvf.com/content_CVPR_2019/papers/Park_Adversarial_Inference_for_Multi-Sentence_Video_Description_CVPR_2019_paper.pdf

249. Wang, X., Yu, F., Wang, R., Darrell, T., Gonzalez, J.E. (2019). TAFE-Net: Task-Aware Feature Embeddings for Low Shot Learning (pp. 1831-1840). Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Referenced from http://openaccess.thecvf.com/content_CVPR_2019/papers/Wang_TAFE-Net_Task-Aware_Feature_Embeddings_for_Low_Shot_Learning_CVPR_2019_paper.pdf

250. Devin, C., Geng, D., Abbeel, P., Darrell, T., and Levine, S. (2019). Compositional plan vectors. Advances in Neural Information Processing Systems, 32. 8pp [link]

251. Ebrahimi, S., Sinha, S., and Darrell, T. (2019). Variational adversarial active learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), volume 4. 8pp [link]

252. Gao, H., Xu, H., Cai, Q.-Z., Wang, R., Yu, F., and Darrell, T. (2019). Disentangling propagation and generation for video prediction. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 9006–9015. [link]

253. Kang, B., Liu, Z., Wang, X., Yu, F., Feng, J., and Darrell, T. (2019). Few-shot object detection via feature reweighting. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 8420– 8429. [link]

254. Laielli, M., Smith, J., Biamby, G., Darrell, T., and Hartmann, B. (2019). Labelar: a spatial guidance interface for fast computer vision image collection. In Proceedings of the 32nd annual ACM symposium on user interface software and technology, pages 987–998. [link]

255. Pathak, D., Lu, C., Darrell, T., Isola, P., and Efros, A. A. (2019). Learning to control self-assembling morphologies: a study of generalization via modularity. Advances in Neural Information Processing Systems, 32, 8pp. [link]

256. Saito, K., Kim, D., Sclaroff, S., Darrell, T., and Saenko, K. (2019). Semi-supervised domain adaptation via minimax entropy. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 8050–8058. [link]

257. Sinha, S., Ebrahimi, S., and Darrell, T. (2019). Variational adversarial active learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 5972–5981. [link]

258. Hu, H.-N., Cai, Q.-Z., Wang, D., Lin, J., Sun, M., Krahenbuhl, P., Darrell, T., and Yu, F. (2019). Joint monocular 3d vehicle detection and tracking. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 5390–5399. [link]

259. Park, D. H., Darrell, T., and Rohrbach, A. (2019). Robust change captioning. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 4624–4633. [link]

260. Wang, D., Devin, C., Cai, Q.-Z., Kr¨ahenbu¨hl, P., and Darrell, T. (2019). Monocular plan view networks for autonomous driving. In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 2876–2883. IEEE. [link]

261. Hu, R., Rohrbach, A., Darrell, T., and Saenko, K. (2019). Language-conditioned graph networks for relational reasoning. In Proceedings of the IEEE/CVF international conference on computer vision, pages 10294– 10303. [link]

262. Wang, D., Devin, C., Cai, Q.-Z., Yu, F., and Darrell, T. (2019). Deep object-centric policies for autonomous driving. In 2019 International Conference on Robotics and Automation (ICR), pages 8853–8859. IEEE. [link]

263. Ebrahimi, S., Meier, F., Calandra, R., Darrell, T., and Rohrbach, M. (2020). Adversarial continual learning. In European Conference on Computer Vision, pages 386–402. Springer, Cham. [link]

264. Herzig, R., Bar, A., Xu, H., Chechik, G., Darrell, T., and Globerson, A. (2020). Learning canonical representations for scene graph to image generation. In European Conference on Computer Vision, pages 210–227. Springer, Cham. [link]

265. Hu, R., Singh, A., Darrell, T., and Rohrbach, M. (2020). Iterative answer prediction with pointer-augmented multimodal transformers for textvqa. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 9992–10002. [link]

266. Kim, J., Moon, S., Rohrbach, A., Darrell, T., and Canny, J. (2020). Advisable learning for self-driving vehicles by internalizing observation-to-action rules. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 9661–9670. [link]

267. Li, R., Jabri, A., Darrell, T., and Agrawal, P. (2020). Towards practical multi-object manipulation using relational reinforcement learning. In 2020 ieee international conference on robotics and automation (icr), pages 4051–4058. IEEE. [link]

268. Luo, Z., Guillory, D., Shi, B., Ke, W., Wan, F., Darrell, T., and Xu, H. (2020). Weakly-supervised action localization with expectation maximization multi-instance learning. In European conference on computer vision, pages 729–745. Springer, Cham. [link]

269. Materzynska, J., Xiao, T., Herzig, R., Xu, H., Wang, X., and Darrell, T. (2020). Something-else: Compositional action recognition with spatial-temporal interaction networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1049–1059. [link]

270. Narasimhan, M., Wijmans, E., Chen, X., Darrell, T., Batra, D., Parikh, D., and Singh, A. (2020). Seeing the un-scene: Learning amodal semantic maps for room navigation. In European Conference on Computer Vision, pages 513–529. Springer, Cham. [link]

271. Park, J. S., Darrell, T., and Rohrbach, A. (2020). Identityaware multi-sentence video description. In European Conference on Computer Vision, pages 360–378. Springer, Cham. [link]

272. Shi, B., Hoffman, J., Saenko, K., Darrell, T., and Xu, H. (2020). Auxiliary task reweighting for minimum-data learning. Advances in Neural Information Processing Systems, 33:7148–7160. [link]

273. Tzeng, E., Devin, C., Hoffman, J., Finn, C., Abbeel, P., Levine, S., Saenko, K., and Darrell, T. (2020). Adapting deep visuomotor representations with weak pairwise constraints. In Algorithmic Foundations of Robotics XII, pages 688–703. Springer, Cham. [link]

274. Wang, X., Yu, F., Dunlap, L., Ma, Y.-A., Wang, R., Mirhoseini, A., Darrell, T., and Gonzalez, J. E. (2020). Deep mixture of experts via shallow embedding. In Uncertainty in artificial intelligence, pages 552–562. PMLR. [link]

275. Wen, C., Lin, J., Darrell, T., Jayaraman, D., and Gao, Y. (2020). Fighting copycat agents in behavioral cloning from observation histories. Advances in Neural Information Processing Systems, 33:2564–2575. [link]

276. Xu, J., Xu, H., Ni, B., Yang, X., Wang, X., and Darrell, T. (2020). Hierarchical style-based networks for motion synthesis. In European conference on computer vision, pages 178–194. Springer, Cham. [link]

277. Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., Madhavan, V., and Darrell, T. (2020). Bdd100k: A diverse driving dataset for heterogeneous multitask learning. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 2636–2645. [link]

278. Zhou, Y., Wang, X., Jiao, J., Darrell, T., and Yu, F. (2020). Learning saliency propagation for semi-supervised instance segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 10307–10316. [link]

279. Cao, J., Wang, X., Darrell, T., and Yu, F. (2021). Instanceaware predictive navigation in multi-agent environments. In 2021 IEEE International Conference on Robotics and Automation (ICR), pages 5096–5102. IEEE. [link]

280. Chen, Y., Liu, Z., Xu, H., Darrell, T., and Wang, X. (2021). Meta-baseline: Exploring simple meta-learning for few-shot learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 9062–9071. [link]

281. Du, Y., Watkins, O., Darrell, T., Abbeel, P., and Pathak, D. (2021). Auto-tuned sim-to-real transfer. In 2021 IEEE International Conference on Robotics and Automation (ICR), pages 1290–1296. IEEE. [link]

282. Guillory, D., Shankar, V., Ebrahimi, S., Darrell, T., and Schmidt, L. (2021). Predicting with confidence on unseen distributions. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 1134–1144. [link]

283. Lambeta, M., Xu, H., Xu, J., Chou, P.-W., Wang, S., Darrell, T., and Calandra, R. (2021). Pytouch: A machine learning library for touch processing. In 2021 IEEE International Conference on Robotics and Automation (ICR), pages 13208–13214. IEEE. [link]

284. Levi, E., Xiao, T., Wang, X., and Darrell, T. (2021). Rethinking preventing class-collapsing in metric learning with margin-based losses. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 10316–10325. [link]

285. Li, B., Wang, Y., Zhang, S., Li, D., Keutzer, K., Darrell, T., and Zhao, H. (2021). Learning invariant representations and risks for semisupervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1104–1113. [link]

286. Liu, Z., Xu, Z., Wang, H.-J., Darrell, T., and Shelhamer, E. (2021). Anytime dense prediction with confidence adaptivity. In International Conference on Learning Representations. 8pp [link]

287. Narasimhan, M., Rohrbach, A., and Darrell, T. (2021). Clip-it! language-guided video summarization. Advances in Neural Information Processing Systems, 34:13988–14000. [link]

288. Ng, E., Ginosar, S., Darrell, T., and Joo, H. (2021). Body2hands: Learning to infer 3d hands from conversational gesture body dynamics. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 11865–11874. [link]

289. Pang, J., Qiu, L., Li, X., Chen, H., Li, Q., Darrell, T., and Yu, F. (2021). Quasi-dense similarity learning for multiple object tracking. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pages 164–173. [link]

290. Reed, C. J., Metzger, S., Srinivas, A., Darrell, T., and Keutzer, K. (2021). Selfaugment: Automatic augmentation policies for selfsupervised learning. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 2674–2683. [link]

291. Saito, K., Kim, D., Teterwak, P., Sclaroff, S., Darrell, T., and Saenko, K. (2021). Tune it the right way: Unsupervised validation of domain adaptation via soft neighborhood density. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 9184–9193. [link]

292. Shi, B., Dai, Q., Hoffman, J., Saenko, K., Darrell, T., and Xu, H. (2021). Temporal action detection with multi-level supervision. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 8022–8032. [link]

293. Xu, H., Chen, B., Gao, Y., and Darrell, T. (2021). Zero-shot policy learning with spatial temporal reward decomposition on contingencyaware observation. In 2021 IEEE International Conference on Robotics and Automation (ICR), pages 10786–10792. IEEE. [link]

294. Yue, X., Zheng, Z., Zhang, S., Gao, Y., Darrell, T., Keutzer, K., and Vincentelli, A. S. (2021). Prototypical cross-domain self-supervised learning for few-shot unsupervised domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 13834–13844. [link]

295. Zhao, S., Wang, Y., Li, B., Wu, B., Gao, Y., Xu, P., Darrell, T., and Keutzer, K. (2021). epointda: An end-to-end simulation-to-real domain adaptation framework for lidar point cloud segmentation. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 35, pages 3500–3509. [link]

296. Zhu, S., Ebrahimi, S., Kanazawa, A., and Darrell, T. (2021). Differentiable gradient sampling for learning implicit 3d scene reconstructions from a single image. In International Conference on Learning Representations. [link]

297. Wang, G., Liu, Z., Hsieh, B., Zhuang, S., Gonzalez, J., Darrell, T., and Stoica, I. (2021). sensai: Convnets decomposition via class parallelism for fast inference on live data. Proceedings of Machine Learning and Systems, 3:664–679. [link]

298. Watkins, O., Gupta, A., Darrell, T., Abbeel, P., and Andreas, J. (2021). Teachable reinforcement learning via advice distillation. Advances in Neural Information Processing Systems, 34:6920–6933. [link]

299. Xiao, T., Reed, C. J., Wang, X., Keutzer, K., and Darrell, T. (2021). Region similarity representation learning. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 10539– 10548. [link]

300. Wang, X., Huang, T. E., Liu, B., Yu, F., Wang, X., Gonzalez, J. E., and Darrell, T. (2021). Robust object detection via instance-level temporal cycle confusion. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 9143–9152. [link]

301. Xiao, T., Singh, M., Mintun, E., Darrell, T., Doll´ar, P., and Girshick, R. (2021). Early convolutions help transformers see better. Advances in Neural Information Processing Systems, 34:30392–30400. [link]

302. Bar, A., Wang, X., Kantorov, V., Reed, C. J., Herzig, R., Chechik, G., Rohrbach, A., Darrell, T., and Globerson, A. (2022). Detreg: Unsupervised pretraining with region priors for object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 14605–14615. [link]

303. Chen, D., Wang, D., Darrell, T., and Ebrahimi, S. (2022). Contrastive test-time adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 295–305. [link]

304. Herzig, R., Ben-Avraham, E., Mangalam, K., Bar, A., Chechik, G., Rohrbach, A., Darrell, T., and Globerson, A. (2022). Objectregion video transformers. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 3148–3159. [link]

305. Narasimhan, M., Ginosar, S., Owens, A., Efros, A. A., and Darrell, T. (2022). Strumming to the beat: Audio-conditioned contrastive video textures. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pages 3761–3770. [link]

306. Ng, E., Joo, H., Hu, L., Li, H., Darrell, T., Kanazawa, A., and Ginosar, S. (2022). Learning to listen: Modeling non-deterministic dyadic facial motion. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 20395–20405. [link]

307. Park, J. S., Shen, S., Farhadi, A., Darrell, T., Choi, Y., and Rohrbach, A. (2022). Exposing the limits of video-text models through contrast sets. In Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 3574–3586. [link]

308. Petryk, S., Dunlap, L., Nasseri, K., Gonzalez, J., Darrell, T., and Rohrbach, A. (2022). On guiding visual attention with language specification. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 18092–18102. [link]

309. Reed, C. J., Yue, X., Nrusimha, A., Ebrahimi, S., Vijaykumar, V., Mao, R., Li, B., Zhang, S., Guillory, D., Metzger, S., et al. (2022). Self-supervised pretraining improves self-supervised pretraining. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pages 2584–2594. [link]

310. Shen, Z., Liu, Z., Liu, Z., Savvides, M., Darrell, T., and Xing, E. (2022). Un-mix: Rethinking image mixtures for unsupervised visual representation learning. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 36, pages 2216–2224. [link]

311. Liu, Z., Mao, H., Wu, C.-Y., Feichtenhofer, C., Darrell, T., and Xie, S. (2022). A convnet for the 2020s. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 11976–11986. [link]

Non-refereed Conference Proceedings:

1. Urtasun, R., Quattoni, A., Lawrence, N., & Darrell, T. (2008). Transferring Nonlinear Representations using Gaussian Processes with a Shared Latent Space. Snowbird workshop abstract. Retrieved from http://dspace.mit.edu/handle/1721.1/41517

2. Salzmann, M., Ek, C. H., Urtasun, R., & Darrell, T. (2010). FOLS : Factorized Orthogonal Latent Spaces. Presented at the Snowbird Learning Workshop. Utah. April 2010. Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2%3A462379&dswid=-8345

3. Brosh, E., Friedmann, M., Kadar, I., Yitzhak Lavy, L., Levi, E., Rippa, S., Lempert, Y., Fernandez-Ruiz, B., Herzig, R., and Darrell, T. (2019). Accurate visual localization for automotive applications. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, 6pp [link]

4. Herzig, R., Levi, E., Xu, H., Gao, H., Brosh, E., Wang, X., Globerson, A., and Darrell, T. (2019). Spatio-temporal action graph networks. In Proceedings of the IEEE/CVF international conference on computer vision workshops, 6pp. [link]

5. Shen, X., Batkovic, I., Govindarajan, V., Falcone, P., Darrell, T., and Borrelli, F. (2020). Parkpredict: Motion and intent prediction of vehicles in parking lots. In 2020 IEEE Intelligent Vehicles Symposium (IV), pages 1170–1175. IEEE. [link]

Books:

1. Darrell, T., Maes, P., Blumberg, B., & Pentland, A. (1996). A Novel Environment for Situated Vision and Behavior. Landy, M.S., Maloney, L.T., & Pavel, M. (Eds.), *Exploratory Vision: the Active Eye*, 319- 331. Springer-Verlag Series in Perception Engineering.

2. Darrell, T., Fisher, J., & Wilson, K., (2006) Geometric and statistical approaches to audiovisual segmentation for untethered interaction. Kuppevelt, J.V., Dybkjaer, L., Bernsen, N.O. (Eds.), *Advances in Natural Multimodal Dialog Systems (Text, Speech and Language Technology*), 175-194. Springer.

3. Shakhnarovich, G., Darrell, T., & Indyk, P. (2006). *Nearest-Neighbor Methods in Learning and Vision: Theory and Practice (Neural Information Processing)*. The MIT Press.