

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/275085326>

A new approach for real-time target tracking in videos

Article in SPIE Newsroom · January 2013

DOI: 10.1117/2.1201301.004648

CITATION

1

READS

78

3 authors, including:



Barak Fishbain

Technion – Israel Institute of Technology

104 PUBLICATIONS 2,826 CITATIONS

SEE PROFILE

A new approach for real-time target tracking in videos

Barak Fishbain, Dorit Hochbaum, and Yan Yang

A generic, graph-theory-based tracking method in videos helps solve the problem of tracking.

Video segmentation is fundamentally challenging, but it has many computer vision applications. The goal of video segmentation is to isolate a target object from its background across a sequence of frames. It is a 3D problem that incorporates the dimension of time. That makes the input an order of magnitude larger than 2D problems, so computational efficiency is a major challenge. Time-efficient tracking of algorithms allows for tailoring targets of interest in real-time and tracing these objects in large visual data sets. These capabilities are useful in law enforcement, analyzing shoppers' behavior, sports broadcasting (having a camera zoom on a ball at all times), and many more applications.

Tracking algorithms in the literature are categorized into three main classes, the active-contour approach,^{1,2} statistical and stochastic methods,^{3,4} and graph-theory-based tracking.^{5,6} The active-contour approach uses continuous models coupled with consistency constraints to delineate a target object's boundary. Since digital images or videos are innately discrete, these methods introduce errors when converting the discrete input to a certain continuous function and when converting the continuous output back to a discrete solution. Statistical and stochastic schemes, by comparison, rely heavily on iterative steps that are computationally intense and neither guarantee optimal solution nor the same output over different runs with the same input data. The third approach, which we employ, represents the issue as a graph problem.

Until recently, graph formulations incorporated a set of motion-consistency-constraints, meaning the object's location in the current frame is constrained to appear close to its location in the previous frame, shifted by the estimated motion. This approach is vulnerable to occlusions and often results in computationally complex problems. In contrast to existing graph-theoretic methods, our approach represents motion as a feature (like color or position), which applies neither heuristics



Figure 1. Representing frames from surveillance sequences taken from the Context Aware Vision using Image-based Active Recognition (CAVIAR) data set. CAVIAR is a project of the European Commission's Information Society Technology program.

nor approximations. A further distinction of our algorithm is that we estimate motion as part of the tracking process, while the other approaches estimate it by optical flow algorithms. While considered to be most accurate, the convergence of optical-flow methods is slow, and often thousands of iterations are necessary for solving the problem. Our approach incorporates the significantly less computationally intensive MPEG-4 motion estimation schemes⁷ while achieving high-quality tracking.

We devised a generic, graph-theory-based tracking scheme in videos. Our method casts the tracking problem as a variant of the normalized cut (NC') problem⁸ and solves it with Hochbaum's Pseudo Flow (HPF) algorithm,⁹ which was shown to be most efficient for the current problem.¹⁰ The robustness of our approach allows for incorporating computationally less intensive block-matching, motion-estimation schemes, such as the ones used within the MPEG-4 video coding framework.¹¹ Although block-matching techniques generate noisy and coarse motion fields, we used them because they have two advantages. First, they have faster computation times, and there is a broad variety of off-the-shelf software and hardware components for performing this task that can easily be incorporated into the segmentation scheme. Second, if the videos are already compressed, then the motion information is inherent in their compressed form and is available from the video encoder.¹² In that case, there is no need to apply any motion estimation algorithm. These two advantages enable real-time execution.

Continued on next page

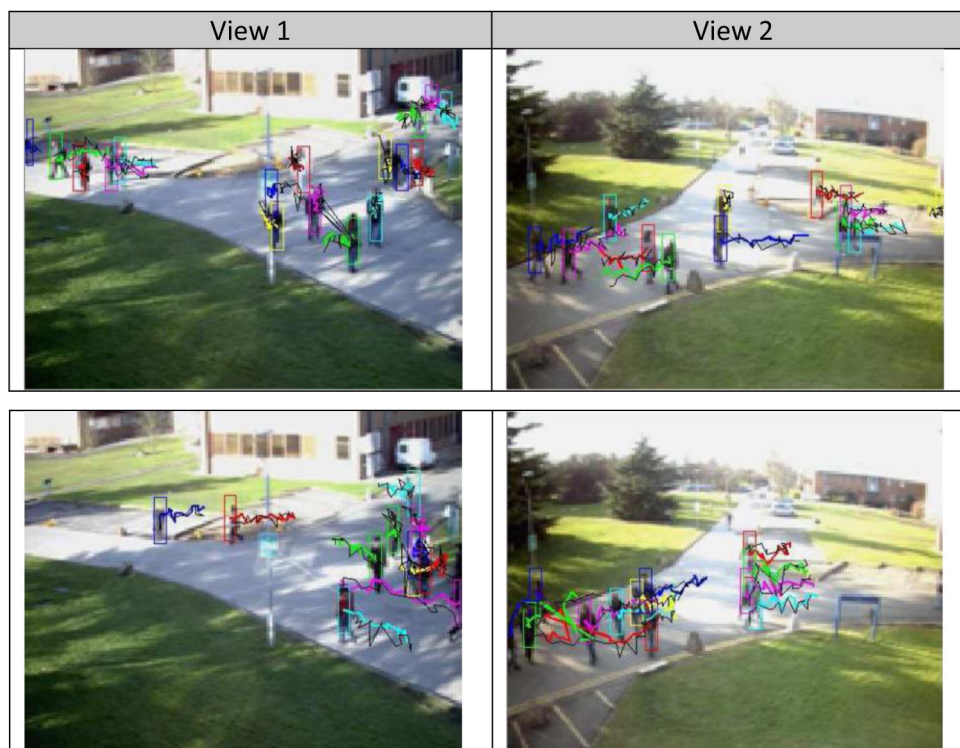


Figure 2. The resulting trajectories for the PETS09 (tasks S2.L1).

Our evaluation of the method on standard (see Figures 1 and 2) and non-standard benchmark videos showed that the integration of motion into the segmentation process provided a superior segmentation output. For quantitative comparison of our method's performance to the state-of-the-art tracking algorithms,^{13–15} we employed the multi object tracking (MOT) measures¹⁶ for precision (MOTP) and accuracy (MOTA). As our algorithm does not tag or identify the different objects, this quantity is omitted in the MOTP calculations. As a base for the quantitative comparison, the PETS2009 S2.L1 and S2.L2 benchmarks (see Figure 2) were arbitrarily selected. Our method presents comparable results in significantly shorter running times. The fact that we can use an MPEG-4 motion field, which is coarse and noisy, indicates that our algorithm is robust, even though that is not a measurable quantity. Along with the time efficiency of the algorithm, this robustness makes the suggested scheme a suitable choice for many video segmentation applications.

Going forward, we plan to exploit the robustness and efficiency of the HPF algorithm in other vision problems. Another line of research is to implement our tracking scheme on a computer chip so the processing will be done on the camera itself. The latter will allow broadcasting of only the target of interest from the camera.

Author Information

Barak Fishbain

Technion - Israel Institute of Technology
Haifa, Israel

Barak Fishbain is an assistant professor. Prior to his arrival there, he was associate director at the Integrated Media Systems Center (IMSC), Viterbi School of Engineering, University of Southern California (USC). He did his postdoctoral studies at the department of Industrial Engineering and Operations Research (IEOR) at the University of California, Berkeley. His research focuses on enviromatics, a new research field that aims to devise mathematical programming methods for machine understanding of trends and behaviors of built and natural environments. This includes environmental distributed sensing (distributed air and water quality monitoring), safety and traffic data realization, and structural sensory networks.

Continued on next page

Dorit Hochbaum and Yan Yang

Industrial Engineering & Operations Research
University of California, Berkeley
Berkeley, CA

Dorit Hochbaum is a full professor and chancellor chair with research interests in approximation algorithms and design and analysis of computer algorithms and discrete and continuous optimization. Her recent work focuses on efficient techniques for network flow-related problems, ranking, data mining, and image segmentation problems. She is the author of more than 140 papers. In 2004 he received an honorary doctorate of Sciences of the University of Copenhagen for her work on approximation algorithms. In 2005 she was named an Institute for Operations Research and the Management Sciences (INFORMS) fellow.

Yan Yang holds a BS from Cornell University in the department of Applied and Engineering Physics, a dual major in the department of Operations Research and Industrial Engineering, and an MS in industrial engineering and operations research from the University of California, Berkeley. He is currently pursuing a PhD in the same department. His studies includes combinatorial optimization, data mining and machine learning, bioinformatics, and video tracking.

References

1. T. Brox, A. Bruhn, and J. Weickert, *Variational motion segmentation with level sets*, **Euro. Conf. on Computer Vision**, pp. 471–483, Springer, 2006.
2. R. Goldenberg, R. Kimmel, E. Rivlin, and M. Rudzsky, *Behavior classification by eigendecomposition of periodic motions*, **Pattern Recognit.** **38** (7), pp. 1033–1043, 2005.
3. M. Black, *Combining intensity and motion for incremental segmentation and tracking over combining intensity and motion for incremental segmentation and tracking over long image sequences*, **Euro. Conf. on Computer Vision**, pp. 485–493, 1992.
4. M. M. Chang, A. M. Tekalp, and M. I. Sezan, *Simultaneous motion estimation and segmentation*, **IEEE Trans. Image Process.** **6** (9), pp. 1326–1333, 1997.
5. A. Bugeau and P. Pérez, *Track and cut: simultaneous tracking and segmentation of multiple objects with graph cuts*, **J. Image Video Process.**, pp. 3:1–3:14, 2008.
6. N. Papadakis and A. Bugeau, *Tracking with occlusions via graph cuts*, **IEEE Trans. Pattern Analysis and Machine Intell.** **33** (1), pp. 144–157, 2011.
7. S. Zhu and K.-K. Ma, *A new diamond search algorithm for fast block-matching motion estimation*, **IEEE Trans. Image Process.** **9** (2), pp. 287–290, 2000.
8. D. S. Hochbaum, *Polynomial time algorithms for ratio regions and a variant of normalized cut*, **IEEE Trans. Pattern Analysis and Machine Intell.** **32** (5), pp. 889–898, 2010.
9. D. S. Hochbaum, *The pseudoflow algorithm: a new algorithm for the maximum-flow problem*, **Operation Res.** **56** (4), pp. 992–1009, 2008.
10. B. G. Chandran and D. S. Hochbaum, *A computational study of the pseudoflow and push-relabel algorithms for the maximum flow problem*, **Operation Res.** **57** (2), pp. 358–376, 2009.
11. S. Zhu and K.-K. Ma, *A new diamond search algorithm for fast block-matching motion estimation*, **IEEE Trans. Image Process.** **9** (2), pp. 287–290, 2000.
12. B. Fishbain, L. Yaroslavsky, and I. Ideses, *Real-time stabilization of long-range observation system turbulent video*, **J. Real-Time Image Process.** **2** (1), pp. 11–22, 2007.
13. M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. V. Gool, *Markovian tracking-by-detection from a single, uncalibrated camera*, **IEEE Int'l. Workshop Perf. Eval. of Tracking and Surveillance**, 2009.
14. M. D. Breitenstein, F. Reichlin, B. Leibe, E. Koller-Meier, and L. V. Gool, *On-line multi-person tracking-by-detection from a single, uncalibrated camera*, **IEEE Trans. Pattern Anal. and Machine**, p. 1820–1833, 2010. doi:10.1109/TPAMI.2010.232
15. J. Yu, D. Farin, and B. Schiele, *Multi-target tracking in crowded scenes*, **Pattern Recognition**, pp. 406–415, 2011.
16. K. Bernardin and R. Stiefelhagen, *Evaluating multiple object tracking performance: The clear mot metrics*, **J. Image and Video Process.**, p. 246309, 2008. doi:10.1155/2008/246309