

STUDIES IN INFORMATICS AND CONTROL

With Emphasis on Useful Applications of Advanced Technology

ISSN 1220-1766

First published: 1992

Published and distributed by ICI Publishing House

Home

About SIC

Editorial Organization

Current Issue

Past Issues

Subscriptions

Guidelines for authors

Search

Studies in Informatics and Control (SIC) is a quarterly publication for all those involved in the field of Information Technology. SIC is published by the **National Institute for Research and Development in Informatics, ICI Bucharest** and has an editorial board whose members referee submissions. Their decisions are based upon the relevance, clarity, and value of the articles submitted.

CORE SUBJECTS

Studies in Informatics and Control covers current topics relating to innovative research and practice in Information Technology: integration of IT with control; IT use in control and management systems; advanced automatic control; application of IT in socio-economic systems and manufacturing processes.

CONTENTS REMARKS

The journal welcomes original, applied and previously-unpublished papers that report on activities that have successfully connected applied research with practical application. In addition to full-length articles, the journal also accepts book reviews, special issues and scientific events.

TARGET AUDIENCE

The articles published in *Studies in Informatics and Control* are of interest both to practicing information technology professionals and to university and industry researchers.

PUBLISHING DETAILS

The journal has been published both in print and online by the Romanian Institute for R&D in Informatics, *ICI Bucharest* under the auspices of the Romanian Academy. It has been published continuously since 1992 and has a worldwide circulation. SIC reviews papers within approximately three months of submission. Every manuscript is reviewed by at least two anonymous referees and accepted articles are published upon receiving the final versions. Publishing intervals would exceptionally be longer than three months.

Short name in **ISI WoK**: STUD INFORM CONTROL

Publisher: National Institute for R&D in Informatics, ICI Bucharest

INDEXING AND ABSTRACTING

SIC is indexed and/or abstracted by several indexing and abstracting services in Information Systems Applications, Computer-based Control, Electrical Engineering, Database Management, etc.

It has been included within the coverage of **Inspec** database (since 1993), **Doaj** (since 1996), **Google Scholar** (since 2006) and Philadelphia-based Institute of Scientific Information (**ISI**) / **Thomson Scientific-SCIEx** (since 2008).

SIC is also included in **SCPIO**.

SIC indicators in Journal Citation Reports:

Impact Factor/3 years in JCR 2010: **0.671**

Impact Factor/3 years in JCR 2011: **0.578**

Where to spot recent articles of the journal:

- [ICI](#)
- [Directory of open access journals](#)
- [The IEE](#)
- [Information Systems Journals](#)

All articles are copyrighted and may be reproduced without permission for non-profit use provided credit is given to SIC.

This publication is partly financed by The Romanian Ministry of Education and Research

Contact: ICI Bucharest, Publishing Dept., 8-10 Averescu Blvd., 011455 Bucharest, ROMANIA e-mail: sic@ici.ro, fax: +40 21 3161030

STUDIES IN INFORMATICS AND CONTROL

With Emphasis on Useful Applications of Advanced Technology

Volume 21 · Issue 4 · 2012

**A Quarterly Journal
An International Editorial Board
Publication Commencement Date: December 1992**

This publication is partly financed by The Romanian Ministry of Education and Research

© National Institute for Research & Development in Informatics

CONTENTS

ARTICLES

Mehmet Ali Ilgin, Surendra M. Gupta

[Physical Programming: A Review of the State of the Art](#)

Horia-Nicolai Teodorescu

[Taylor and Bi-local Piecewise Approximations with Neuro-Fuzzy Systems](#)

Constanța Zoie Rădulescu, Marius Rădulescu

[A Decision Support Tool Based on a Portfolio Selection Model for Crop Planning under Risk](#)

Jaroslav Michalco, Pavol Navrat

[Arrangement of Face-to-Face Meetings using Social Media](#)

Oscar Esquivel-Flores, Hector Benitez-Pérez, Paul Méndez-Monroy

[Scheduling Strategy Using Frequency Transition for a Helicopter Simulation as a Network Control System Approximation](#)

Sándor Tihamér Brassai, Barna Iantovics, Călin Enăchescu

[Optimization of Robotic Mobile Agent Navigation](#)

Mihai Făgădar-Cosma, Marwen Nouri, Vladimir-Ioan Crețu, Mihai Victor Micea

[A Combined Optical Flow and Graph Cut Approach for Foreground Extraction in Videoconference Applications](#)

M. Helaimi, M. Benghanem, B. Belmadani

[An Improved PI^λ Controller for Resonant Inverter Induction Heating Systems under Load and Line Variations](#)

Gisli Thorsteinsson, Andrei Niculescu

[Using Mobile Technology for Problem Need Identification in School-aged Children Environment](#)

Valentina Tudor, Alexandru Sipiță

[Providing Online Information Services to Small Farm Owners](#)

Mihai Talmaciu

[Determination of the Optimally Efficient Sets in Special Classes of Graphs](#)

[AUTHOR INDEX](#)

A Combined Optical Flow and Graph Cut Approach for Foreground Extraction in Videoconference Applications

Mihai Făgădar-Cosma

Alcatel-Lucent Bell N.V.

Copernicuslaan 50, 2018 Antwerp, Belgium

"Politehnica" University of Timișoara, Department of Computer Science

Bd. Vasile Pârvan 2, 300223 Timișoara, Romania

Vladimir-Ioan Crețu

"Politehnica" University of Timișoara, Department of Computer Science

Bd. Vasile Pârvan 2, 300223 Timișoara, Romania

Marwen Nouri

Alcatel-Lucent Bell Labs, Centre de Villarceaux

Route de Villejust, 91620 Nozay, France

Mihai Victor Micea

"Politehnica" University of Timișoara, Department of Computer Science

Bd. Vasile Pârvan 2, 300223 Timișoara, Romania

Abstract: Immersive videoconferences have added a new dimension to remote collaboration by bringing participants together in a common virtual space. To achieve this, the conferencing system must extract in real-time the foreground from each incoming video stream and translate it into the shared virtual space. The method presented in this paper differentiates itself in the sense that no prior training or assumptions on the video content are used during foreground extraction. A temporally coherent mask is created based on motion cues obtained from the video stream and is used to provide a set of hard constraints. Based on these constraints, a graph cut algorithm is employed to produce the pixel-accurate foreground segmentation. The obtained results are evaluated using a state-of-the-art perceptual metric to provide an objective assessment of the method accuracy and reliability. Furthermore, the presented approach makes use of parallel execution in order to achieve real-time processing capabilities.

Keywords: Foreground extraction, videoconference, optical flow, graph cut, motion segmentation, real-time video processing

[Full text](#)

CITE THIS PAPER AS:

Mihai Făgădar-Cosma, Marwen Nouri, Vladimir-Ioan Crețu, Mihai Victor Micea, A Combined Optical Flow and Graph Cut Approach for Foreground Extraction in Videoconference Applications, *Studies in Informatics and Control*, ISSN 1220-1766, vol. 21 (4), pp. 413-422, 2012.

1. Introduction and Related Work

In the last decade videoconferencing has gained a lot of momentum, supported by the introduction of fixed and mobile broadband Internet and the availability of affordable and easy to use video capture hardware. Having achieved the desiderate of real-time audio, video and document sharing, the next step in videoconferencing is to deliver an immersive experience by gathering participants into a common virtual space that further enhances collaboration options [1]. At the root of this concept rests the ability of the conferencing system to accurately extract foreground information from each incoming video stream and use it to populate the virtual space which is shared with all participants. The perceived quality of foreground segmentation represents a key aspect for achieving a true immersive experience, further accentuated by the fact that foreground segmentation is in itself an ill-posed problem [2]. The implementation of an immersive videoconferencing system must therefore rely on a real-time, automatic foreground extraction algorithm, capable of handling monocular video sequences that may exhibit illumination changes, multimodal and cluttered backgrounds, camera noise and video compression artifacts.

Foreground / background segmentation has been an active research area of video sequence processing, with many algorithms and methods being developed [3-6]. The most accurate results are obtained by methods that rely on dedicated setups involving stereoscopic [7] or multiple cameras [8, 9]. While highly robust, these methods are not feasible for common videoconferencing scenarios which use off-the-shelf or integrated monocular webcams.

In the field of monocular object segmentation, the majority of encountered algorithms rely on background subtraction [4] based on an empty image of the scene provided during the initialization stage. In [10] the foreground layer is extracted by combining background subtraction with color and contrast cues. The key concept revolves around background contrast attenuation, which reduces contrast in the background layer while preserving it around object boundaries. The method proposed in [11] uses a known and stationary background image and a frontal human body detector in order to perform an initial segmentation of the person in the scene. The result is subject to a coarse to fine segmentation process [12] that relies on a GMM model of foreground and background pixels to provide the input for an unsupervised graph cut segmentation [13, 14]. In addition, a self-adaptive initialization level sets scheme is applied in order to find the most salient edges along the person's contour. The major drawback of these otherwise accurate methods is the requirement for an initial clean background image. This cannot be satisfied in videoconference scenarios, since people are usually in the scene starting from the first frame and the number of potential backgrounds is virtually infinite.

Another approach is to replace the need for an initial background image with a learning model trained using manually labeled video sequences. Criminisi et al. [15] have adapted stereoscopic approaches to monocular video by using a probabilistic framework to fuse motion, color and contrast cues with spatio-temporal (S-T) priors generated during training phase. The accuracy of this method is similar to the one in [7], except for cases when foreground color distribution resembles the one in the background or when there is insufficient motion. Further improvements described in [16] have replaced the Hidden Markov Model with tree-based classifiers trained on ground-truth segmentations that imitate depth masks used in stereoscopic vision. The classifiers operate on motion information encoded in the form of motions (motion descriptors similar to textons, which encode texture information). This method allows a better segmentation of the foreground which is closest to the camera, being able to discard background motion. Despite their relatively high accuracy, both methods can be prohibitive due to the need to calibrate the learned priors for different types of scenes using manually labeled sequences.

A third way of addressing the foreground segmentation problem takes the form of constraints placed on the nature and position of foreground objects. Kim et al. [17, 18] propose an algorithm which targets the part of the MPEG-4 standard related to object-based video compression and handling. The algorithm performs S-T motion segmentation by combining a low-complexity spatial technique with a marker extraction and update process followed by a region growing phase. The low complexity and relative accuracy of the approach makes it suitable for use in mobile devices, but the a priori assumption that the foreground object is placed in the center of the frame limits the number of applicable scenarios. For videoconferencing, segmentation needs to take into account extra movements, other than only those related to head and torso. For example, the system must handle cases in which a person uses hand gestures and body language in order to support the presentation of a topic or to show an exhibit.

The foreground extraction method proposed in the present paper eliminates the need for initial training as well as any a priori assumptions or knowledge related to the nature of the observed scene. Starting from accurate motion cues obtained through aggregation of dense and sparse optical flow information [19], the system builds a temporally coherent mask (TCM) of foreground detected through motion. The temporal coherence of the mask in absence of motion is achieved through the use of image statistics, similar to other methods encountered in the state-of-the-art [3, 18]. To obtain the final pixel-accurate segmentation, a heuristic approach combines the TCM and sparse optic flow information in order to generate the hard foreground and background constraints for a graph-cut algorithm. The accuracy and reliability of the obtained results are evaluated using the state-of-the-art perceptual objective metric described in [20]. The proposed approach supports parallelization, enabling it to achieve real-time execution capabilities.

References

1. KAUFF, P., O. SCHREER, **An Immersive 3D Video-conferencing System using Shared Virtual Team User Environments**, Proc. 4th International Conference on Collaborative Virtual Environments, ACM, 2002, pp. 105-112.
2. GRADY, L., M. P. JOLLY, A. SEITZ, **Segmentation from a Box**, Proceedings of IEEE International Conference on Computer Vision, 2011, pp. 367-374.
3. BOUWMANS, T., F. E. BAF, B. VACHON, **Statistical Background Modeling for Foreground Detection: A Survey**, Handbook of Pattern Recognition and Computer Vision, W.S. Publishing, 2010, pp. 181-199.
4. PICCARDI, M., **Background Subtraction Techniques: a Review**, Proc. of IEEE International Conference on Systems, Man and Cybernetics, 2004, pp. 3099-3104.

5. ZAPPELLA, L., X. LLADO, J. SALVI, **Motion Segmentation: a Review**, Proc. 11th International Conference of the Catalan Association for Artificial Intelligence, IOS Press, 2008, pp. 398-407.
6. ZHANG, D., G. LU, **Segmentation of Moving Objects in Image Sequence: A Review**, Circuits, Systems, and Signal Processing, vol. 20, 2001, pp. 143-183.
7. KOLMOGOROV, V., A. CRIMINISI, A. BLAKE, G. CROSS, C. ROTHER, **Bi-Layer Segmentation of Binocular Stereo Video**, Proceedings of 2005 IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, vol. 2, 2005, pp. 407-414.
8. DEVINCENZI, A., L. YAO, H. ISHII, R. RASKAR, **Kinected Conference: Augmenting Video Imaging with Calibrated Depth and Audio**, Proceeding of ACM Conference on Computer supported cooperative work, ACM, 2011, pp. 621-624.
9. KIM, H., R. SAKAMOTO, I. KITAHARA, T. TORIYAMA, K. KOGURE, **Robust Foreground Segmentation from Color Video Sequences Using Background Subtraction with Multiple Thresholds**, Technical report of IEICE. PRMU vol. 106, 2006, pp. 135-140.
10. SUN, J., W. ZHANG, X. TANG, H.-Y. SHUM, **Background Cut**, Proc. 9th European conference on Computer Vision, Springer-Verlag, 2006, pp. 628-641.
11. LIU, Q., H. LI, K. N. NGAN, **Automatic Body Segmentation with Graph Cut and Self-adaptive Initialization Level Set (SAILS)**, J. Vis. Comun. Image Represent. 22 (2011) 367-377.
12. LI, H., K. N. NGAN, Q. LIU, **FaceSeg: Automatic Face Segmentation for Real-time Video**, Trans. Multi. vol. 11, 2009, pp. 77-88.
13. BOYKOV, Y., M. P. JOLLY, **Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images**, Proceedings of the 8th IEEE International Conference on Computer Vision, 2001, pp. 105-112.
14. BOYKOV, Y., V. KOLMOGOROV, **An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision**, IEEE Trans. Pattern Anal. Mach. Intell., vol. 26, 2004, pp. 1124-1137.
15. CRIMINISI, A., G. CROSS, A. BLAKE, V. KOLMOGOROV, **Bilayer Segmentation of Live Video**, Proc. 2006 IEEE Conference on Computer Vision and Pattern Recognition, IEEE Computer Society, vol. 1, 2006, pp. 53-60.
16. YIN, P., A. CRIMINISI, J. WINN, I. A. ESSA, **Bilayer Segmentation of Webcam Videos Using Tree-Based Classifiers**, IEEE Trans. Pattern Anal. Mach. Intell. 33, 2011, pp. 30-42.
17. KIM, J., H.-J. LEE, T.-H. LEE, M. CHO, J.-B. LEE, **Hardware/Software Partitioned Implementation of Real-time Object-oriented Camera for Arbitrary-shaped MPEG-4 Contents**, Proc. 2006 IEEE/ACM/IFIP Workshop on Embedded Systems for Real Time Multimedia, IEEE Computer Society, 2006, pp. 7-12.
18. KIM, J., J. ZHU, H.-J. LEE, **Block-Level Processing of a Video Object Segmentation Algorithm for Real-Time Systems**, CORD Conference Proceedings, 2007, pp. 2066-2069.
19. FAGADAR-COSMA, M., V. I. CRETU, M. V. MICEA, **Dense and Sparse Optic Flows Aggregation for Accurate Motion Segmentation in Monocular Video Sequences**, Proc. 2012 International Conference on Image Analysis and Recognition, Springer LNCS, vol. 7324, 2012, pp. 208-215.
20. DRELIE-GELASCA, E., T. EBRAHIMI, **On Evaluating Video Object Segmentation Quality: A Perceptually Driven Objective Metric**, IEEE Journal of Selective Topics Signal Process., vol. 3, 2009, pp. 319-335.
21. NIKULIN, M. S., **Hellinger Distance**, Encyclopaedia of Mathematics, Springer, 2001.
22. BOYKOV, Y., O. VEKSLER, R. ZABIH, **Fast Approximate Energy Minimization via Graph Cuts**, IEEE Transactions Pattern Anal. Mach. Intell., vol. 23, 2001, pp. 1222-1239.
23. ROTHER, C., V. KOLMOGOROV, A. BLAKE, **"GrabCut": Interactive Foreground Extraction using Iterated Graph Cuts**, ACM Trans. Graph., vol. 23, 2004, pp. 309-314.
24. BOYKOV, Y., G. FUNKA-LEA, **Graph Cuts and Efficient N-D Image Segmentation**, Int. J. Comput. Vision, vol. 70, 2006, pp. 109-131.
25. BOYKOV, Y., G. FUNKA-LEA, **Optimal Object Extraction via Constrained Graph-cuts**, Int. J. Comput. Vision, 2004.
26. LI, Y., J. SUN, C.-K. TANG, H.-Y. SHUM, **Lazy Snapping**, ACM Trans. Graph., vol. 23, 2004, pp. 303-308.
27. KUMAR, P., K. SENGUPTA, S. RANGANATH, **Real Time Detection and Recognition of Human Profiles Using Inexpensive Desktop Cameras**, Pattern Recognition, International Conference on, 2000, pp. 1096-1099.
28. SUNDARAM, N., T. BROX, K. KEUTZER, **Dense Point Trajectories by GPU-accelerated Large Displacement Optical Flow**, Proc. 11th European conference on Computer vision: Part I, Springer-Verlag, 2010, pp. 438-451.
29. HE, Z., F. KUESTER, **GPU-based Active Contour Segmentation using Gradient Vector Flow**, Proceedings of the 2nd International Conference on Advances in Visual Computing, Springer-Verlag, 2006, pp. 191-201.
30. VINEET, V., P.J. NARAYANAN, **CUDA Cuts: Fast Graph Cuts on the GPU**, 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008, pp. 1-8.
31. HORPRASERT, T., D. HARWOOD, L. DAVIS, **A Statistical Approach for Real-time Robust Background Subtraction and Shadow Detection**, ICCV Frame-Rate WS, 1999, pp. 1-19.
32. JABRI, S., Z. DURIC, H. WECHSLER, A. ROSENFELD, **Detection and Location of People in Video Images using Adaptive Fusion of Color and Edge Information**, International Conference on Pattern Recognition, 2000, pp. 627-630.