BACKGROUND SUBTRACTION USING SPATIO-TEMPORAL CONTINUITIES

Srenivas Varadarajan[1], Lina J. Karam[1] and Dinei Florencio[2]

[1]Image, Video, and Usability (IVU) Lab
School of Electrical, Computer & Energy Engineering
Arizona State University, Tempe, Arizona, USA

[2]Microsoft Research
One Microsoft Way, Redmond, Washington, USA

svarada2@asu.edu, karam@asu.edu, dinei@microsoft.com

ABSTRACT

We present a novel scheme for dynamically recovering a background image from consecutive frames of a video sequence based on spatial and temporal continuities. The proposed algorithm applies a boundary-level spatial continuity constraint in order to detect and correct ghosting, which corresponds to incorrectly classified foreground regions due to fast moving objects. The proposed method can be applied successfully to sequences with deformable foreground objects and non-uniform motion. Simulation results show that the extracted background, when used for foreground detection, results in a higher performance in terms of recall and precision as compared to existing popular schemes.

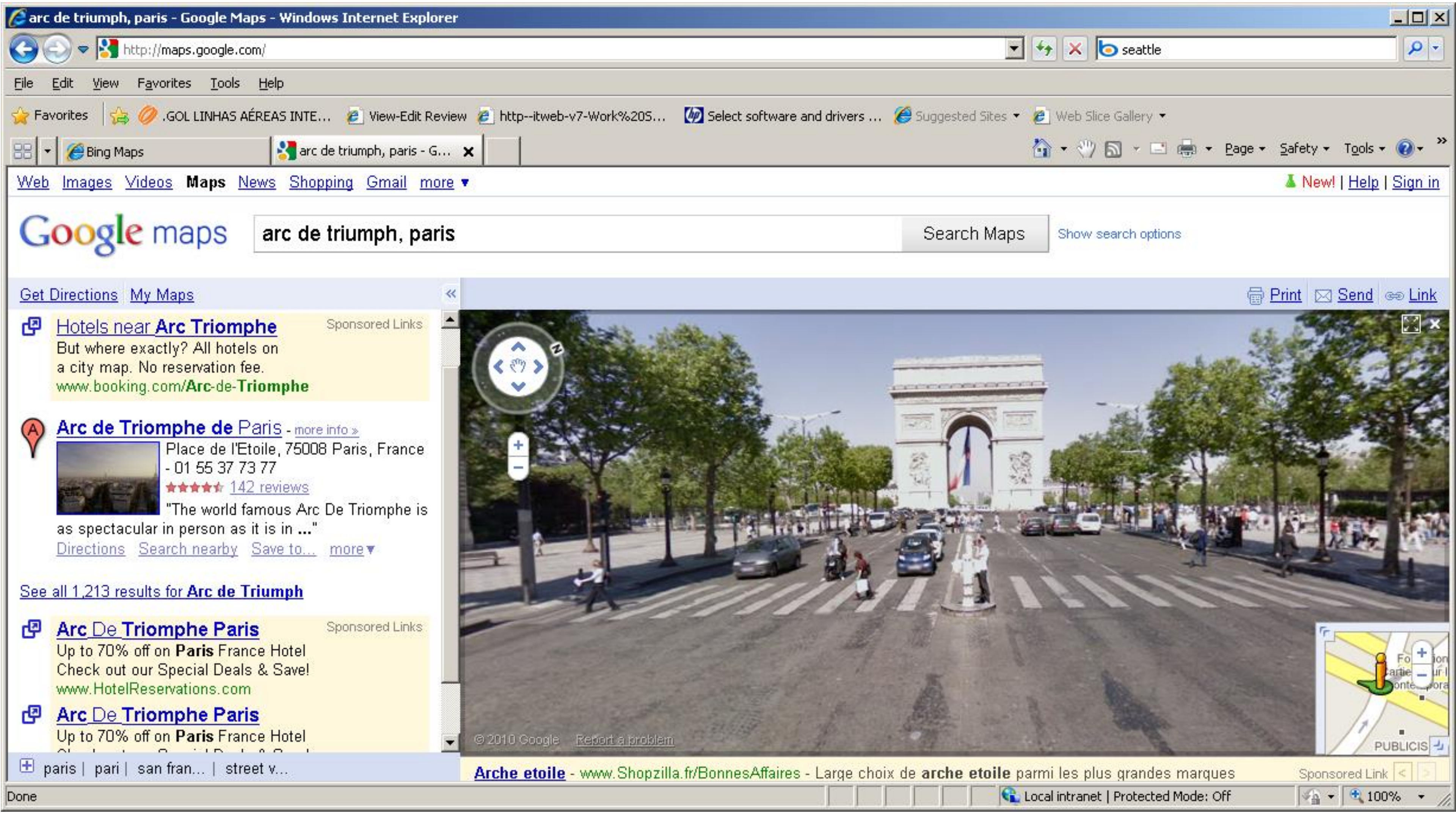
Outline

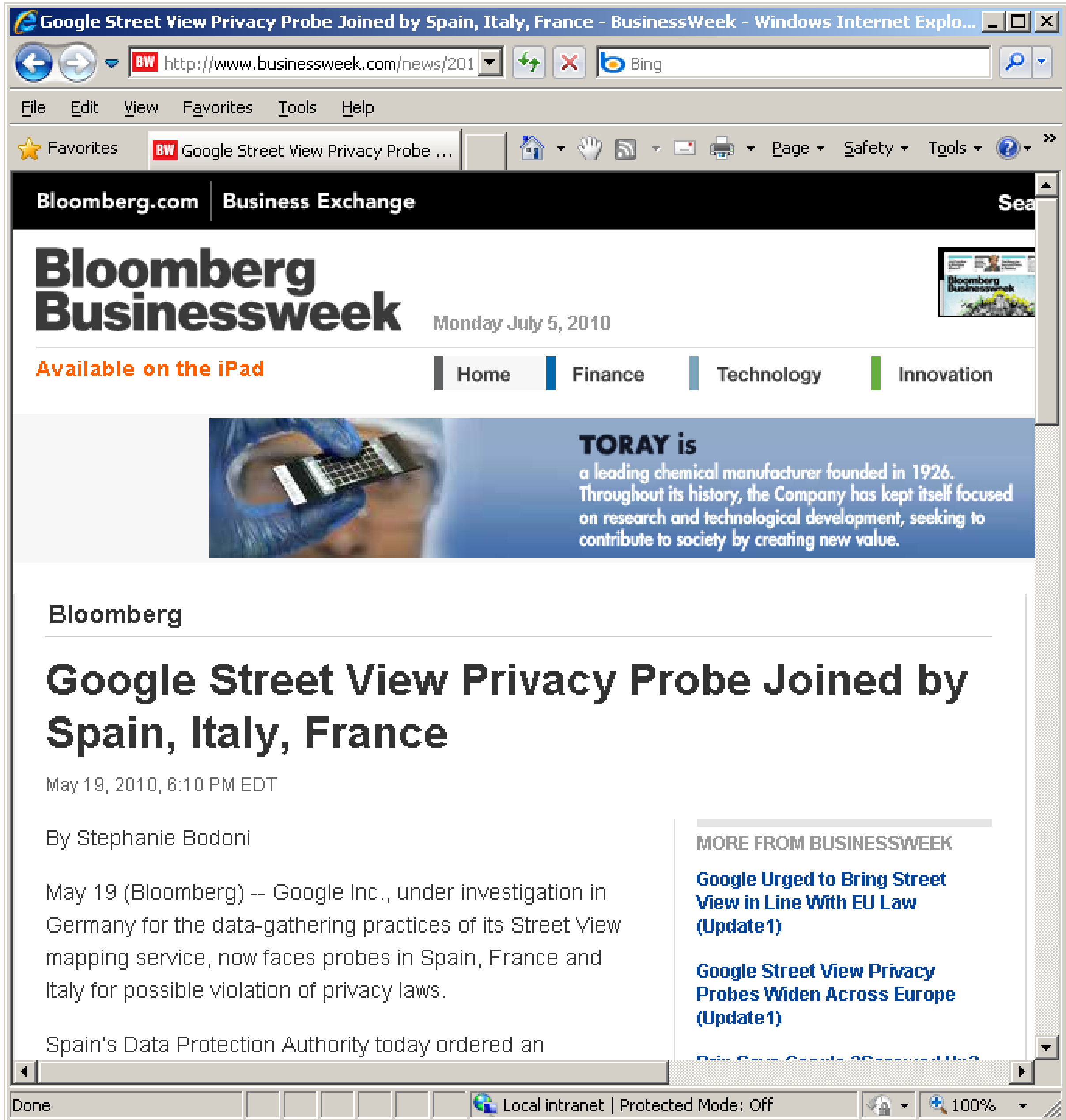
1. Motivation / Applications
2. Introduction
    - Existing Approaches for Foreground Detection
    - Automatic Occlusion detection
3. Proposed Algorithm
4. Simulation Results and Analysis
5. Conclusion

Motivation 1: Background Replacement

Motivation 2: 3D effects

Motivation 3: Privacy




Google Street View Privacy Probe Joined by Spain, Italy, France - BusinessWeek - Windows Internet Explo...

http://www.businessweek.com/news/201

Bloomberg.com | Business Exchange

Bloomberg Businessweek

Monday July 5, 2010

Available on the iPad

Home | Finance | Technology | Innovation

Bloomberg

Google Street View Privacy Probe Joined by Spain, Italy, France

May 19, 2010, 6:10 PM EDT

By Stephanie Bodoni

May 19 (Bloomberg) -- Google Inc., under investigation in Germany for the data-gathering practices of its Street View mapping service, now faces probes in Spain, France and Italy for possible violation of privacy laws.

Spain's Data Protection Authority today ordered an

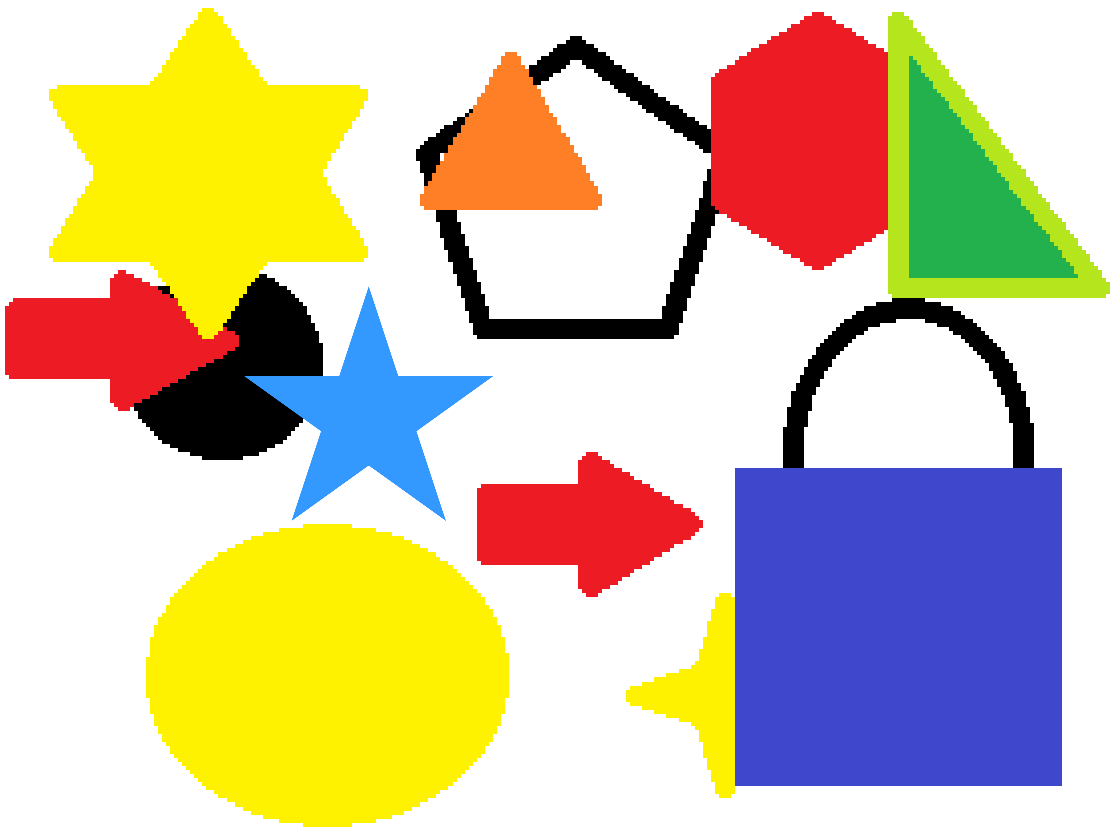
MORE FROM BUSINESSWEEK

Google Urged to Bring Street View in Line With EU Law (Update1)

Google Street View Privacy Probes Widen Across Europe (Update1)

2. Introduction

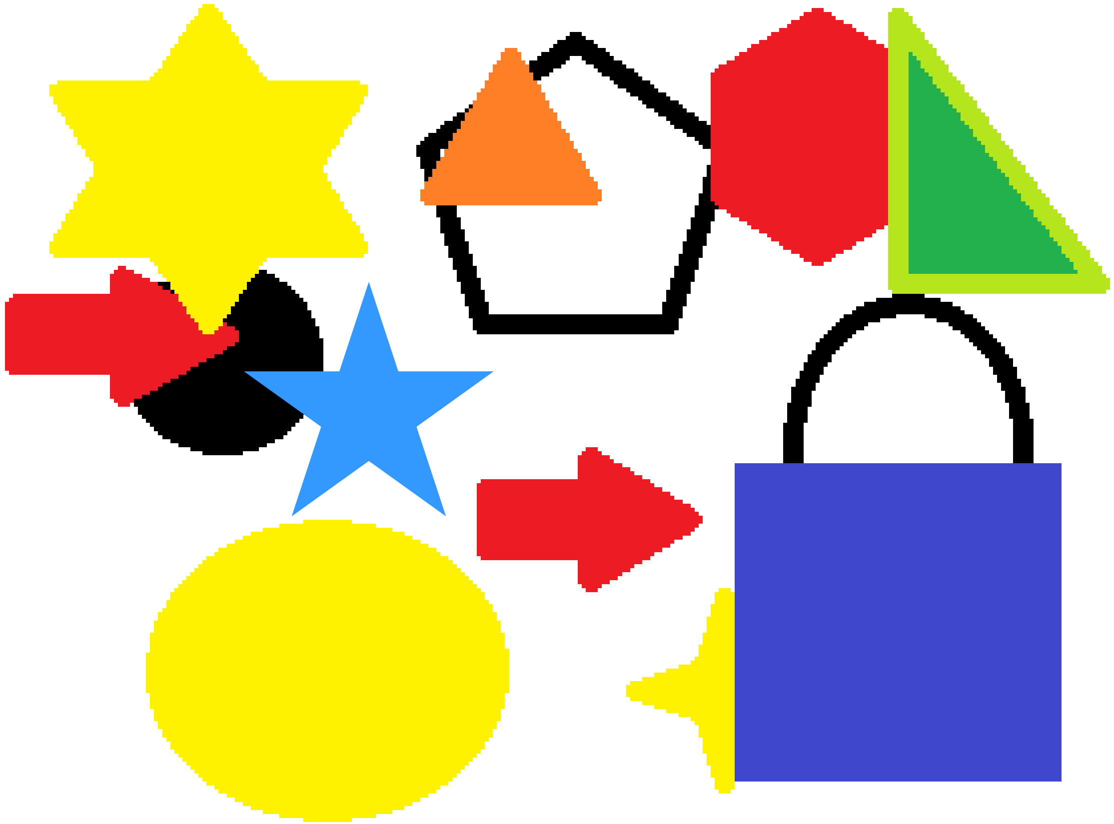
Can you tell foreground vs. background?



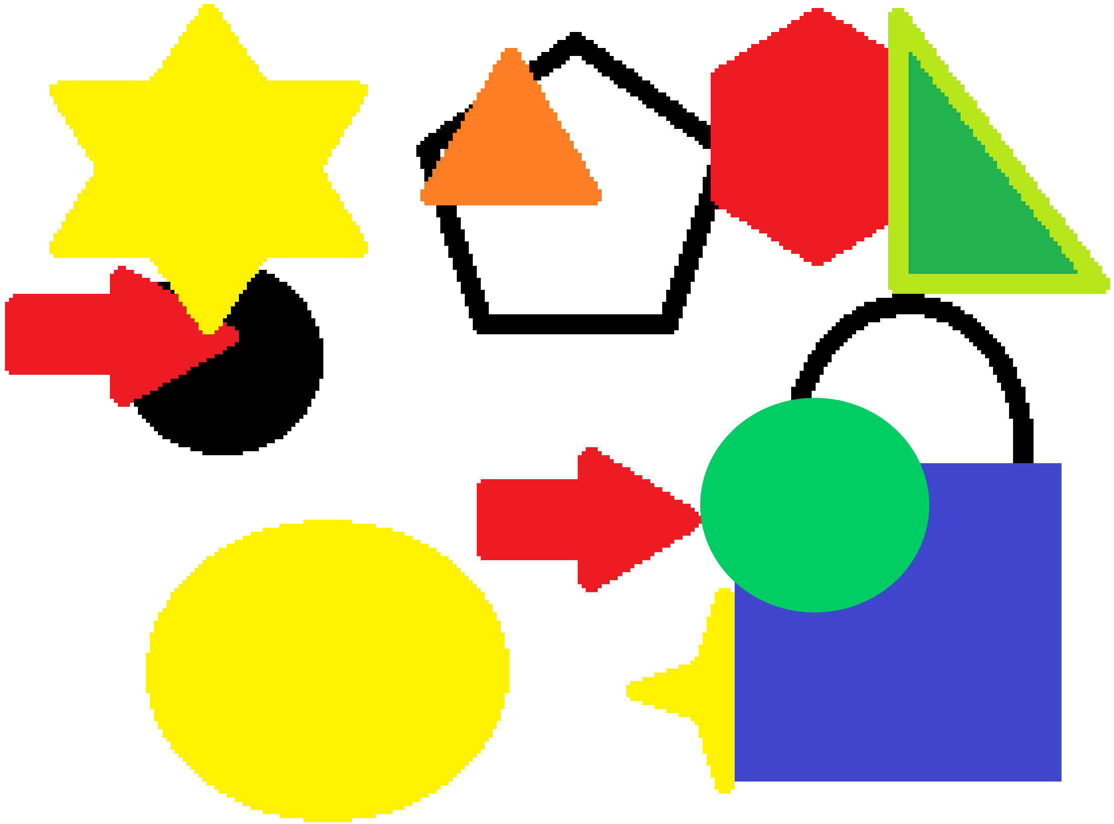
Existing Approaches for Foreground Detection

Approach	Description	Drawbacks
Model-based Stauffer et al. [1] Elgammal et al. [2]	Parametric or non-parametric models are fitted to the background and/or to the foreground pixels. Depending on the deviation from these models, a pixel is classified as foreground or background	Optimal number of Gaussians and the learning rate cannot be set a priori for all situations
Feature-based Z. Wu et al. [6]	A feature like color, edges, motion and texture is used for classification	Edge and texture-based approaches fail when the background is smooth
Pixel-differencing-based S.Varadarajan et al.[12]	The difference of co-located pixels in adjacent frames is compared with a threshold in order to detect the foreground objects	Ghosting and Foreground-Aperture problems due to fast moving and large uniform foreground objects, respectively
Other Approaches Xun Xu et al. [4] F. El Baf et al. [5]	Loopy belief propagation [4] and those based on fuzzy integrals which combine a set of features	1. Smoothness assumption fails at object transitions in the background 2. High complexity multi-pass algorithms

Can you tell foreground versus bkgnd?



Can you tell foreground versus bkgnd?

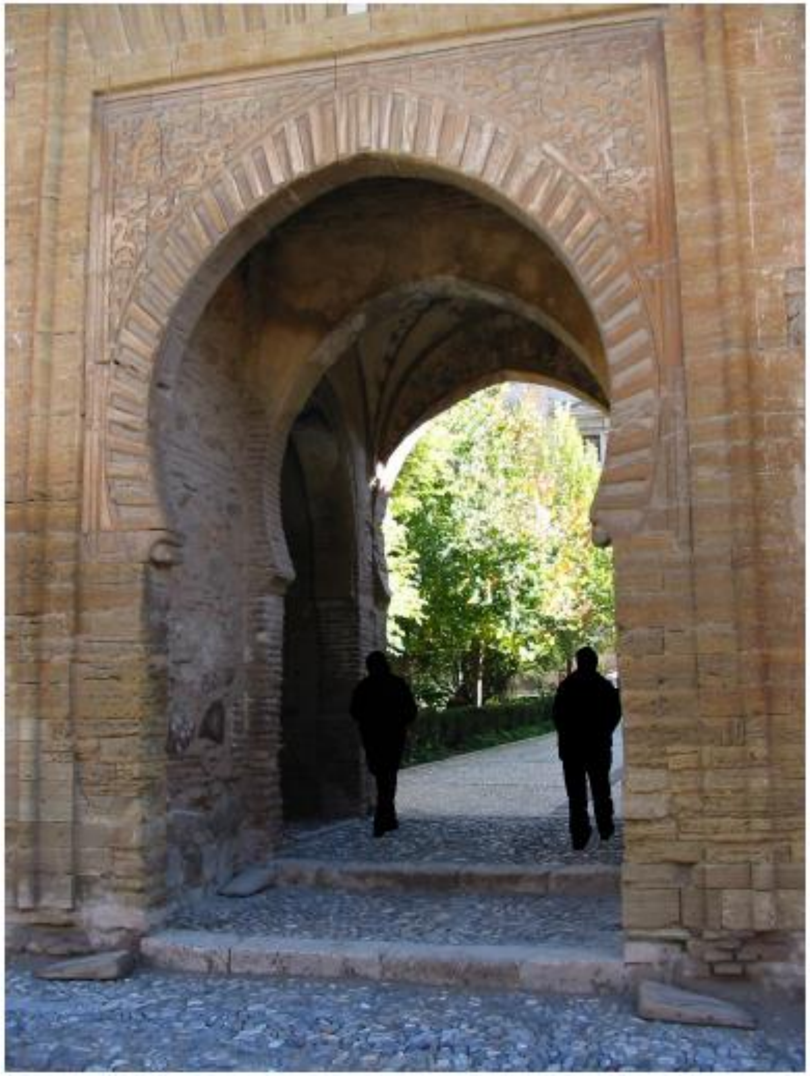


Algorithm for automatically detecting which object is in front

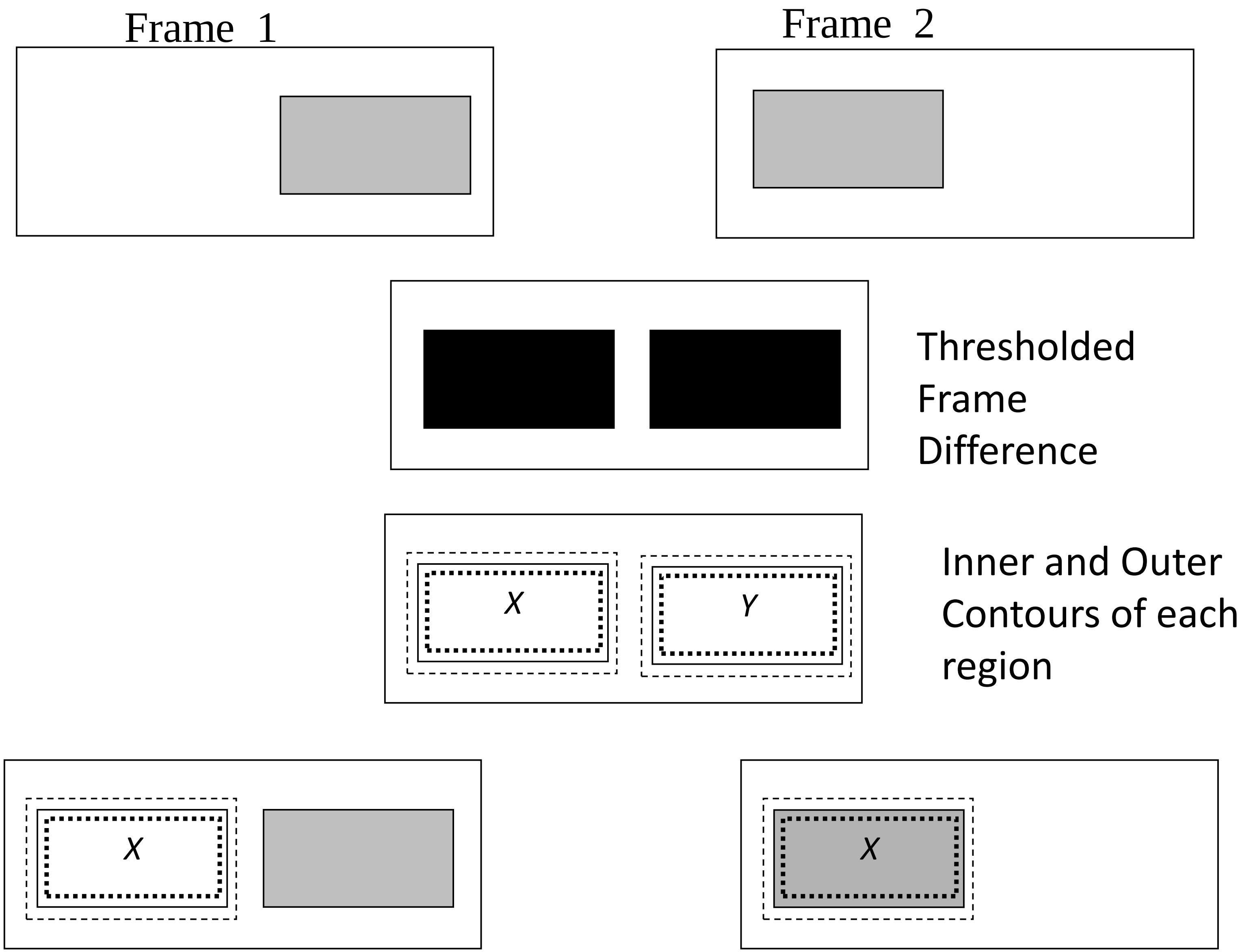
[1] C. Herley, "Automatic occlusion removal from minimum number of images", *ICIP 2005.*

The inputs

Occluded areas



Unoccluded


Frame 1
Frame 2
Thresholded
Frame
Difference
X
Y
Inner and Outer
Contours of each
region
X
X

3. Proposed Algorithm

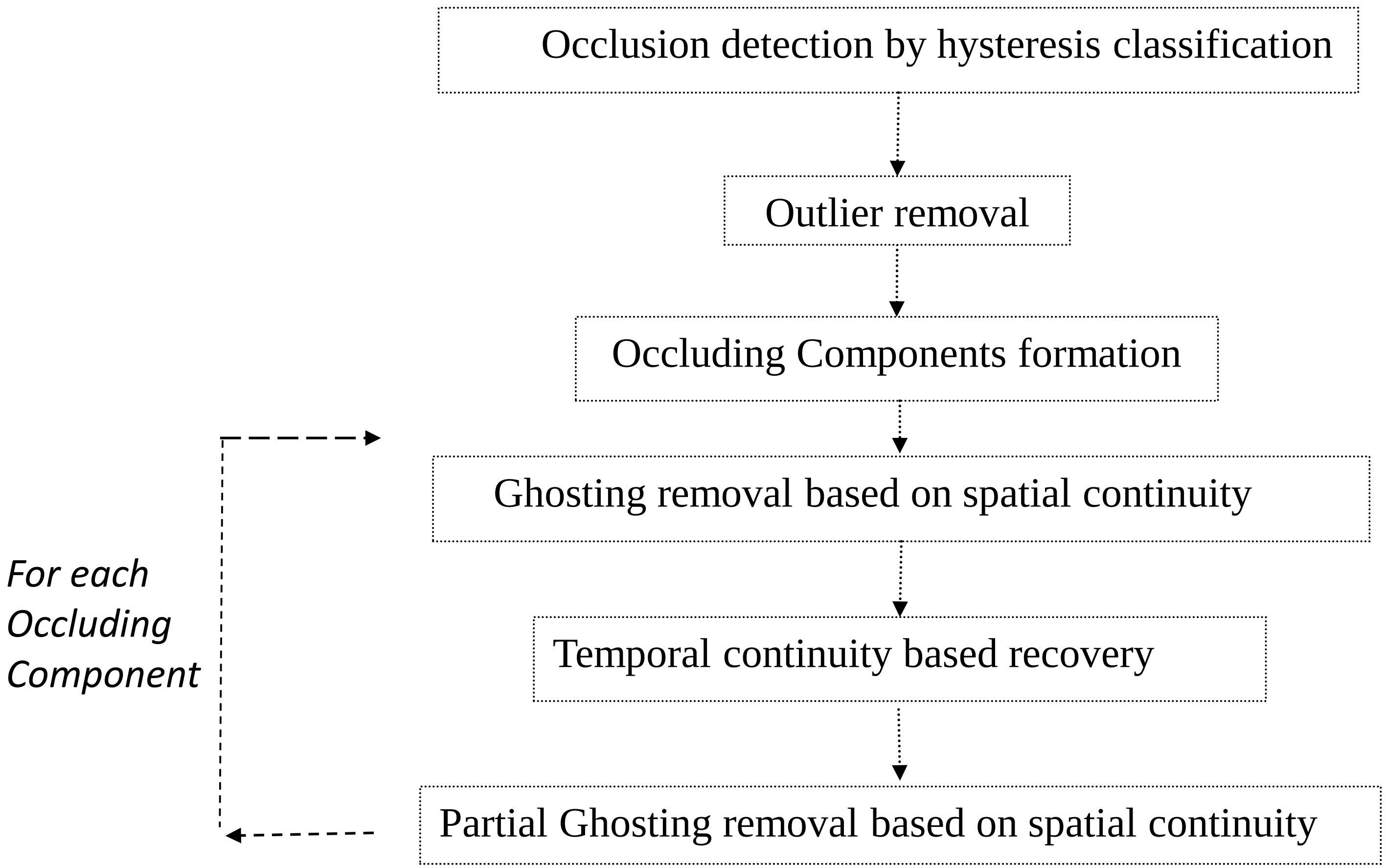
Assumptions

• A **static background** and a **moving foreground** is assumed. Both the background and foreground may consist of several objects.

• Since the background is static, it exhibits temporal continuity, i.e., co-located background pixels in adjacent frames have similar values.

•We do NOT assume background is unoccluded most of the time.

Proposed Spatio-Temporal Continuity-based Background Subtraction algorithm



Foreground-Background Hysteresis Classification [12]

- Initial foreground guess based on initial *N* frames.
- First *N* consecutive frames are low pass filtered. (*N* = 5 in our implementation). Then, each color pixel is classified as
  - Strong Foreground (*SF*),
  - Weak Foreground (*WF*), or
  - Background (*B*)

$$C_{x,y,1} = \begin{cases} SF, & if \ \|P_{x,y,1} - P_{x,y,n}\|_{L_1} > t_2 \ for \ any \ n = 2, \ldots, N \\ WF, & if \ t_1 < \|P_{x,y,1} - P_{x,y,n}\|_{L_1} < t_2 \ for \ any \ n = 2, \ldots, N \\ B, & else \end{cases}$$

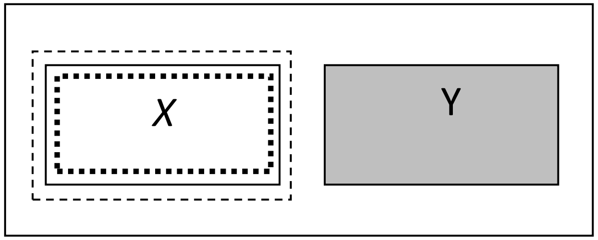
Where t_1 and t_2 ($t_1 < t_2$) correspond to a low threshold and a high threshold ($t_1 = 3$ and $t_2 = 20$ in our implementation).

- Followed by outlier removal and incorporating WF into neighboring SF.

Occluding Components Formation

- Each moving foreground object corresponds to an Occluding Component (OC) and will be treated independently of others.

Ghosting Detection and Removal



Assume a fast moving object:

- **X** be the position of the box in the previous frame and **Y** be its position in the current frame. Both these regions are detected as background on pixel-differencing.

Ghosting Detection:

Compute a spatial discontinuity metric:

$$D = \frac{1}{\# B_k^{in}} \sum_{(x,y) \in B_k^{in}} \left\| p_{x,y,n} - p_{ClosestOut} \right\|_{L_1}$$

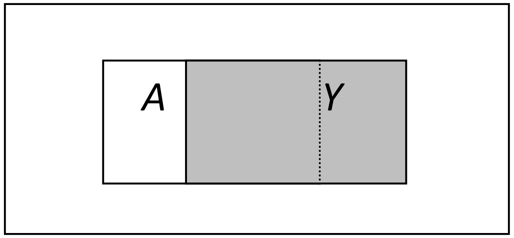
where

$\# B_k^{in}$ the number of pixels in the inner boundary of the OC

Ghosting Detection:

If $D < 5$, the region is recognized as a background blob and replaced with the pixels of the current frame.

Partial-Ghosting Detection and Removal



Assume:

- There is a partial overlap of a foreground object across two successive frames.
- ***A*** is the portion of the background uncovered by the object but detected as foreground (ghost), while ***Y*** is the actual position of the object.

Partial-Ghosting Detection:

- The spatial continuity criterion is applied at a pixel level instead of the entire object's boundary.
- The vertical and horizontal boundaries of each OC in the current frame are located by horizontal and vertical scans of the foreground mask, respectively.
- For a horizontal boundary pixel located at ***(x,y)***, a horizontal discontinuity metric is computed as follows:

$$D_H(x,y,n) = \left\| p_{x-1,y,n} - p_{x+1,y,n} \right\|_{L_1}$$

- For a vertical boundary pixel located at *(x,y)*, a vertical discontinuity metric is computed as follows:

$$D_V(x,y,n) = \left\| p_{x,y-1,n} - p_{x,y+1,n} \right\|_{L_1}$$

Partial-Ghosting Removal :

If the computed D_H or D_V is below a threshold (equals 3 in our implementation), then the pixel is considered as a background pixel and recovered.

Temporal Continuity Based Recovery

Issues not solved by a system based only on spatial continuity:

- The assumption of a smooth background fails when the background contains sharp transitions
- These sharp background edges may coincide with edges of the foreground object, in which case the spatial continuity constraint fails.

Temporal Continuity Based Recovery:

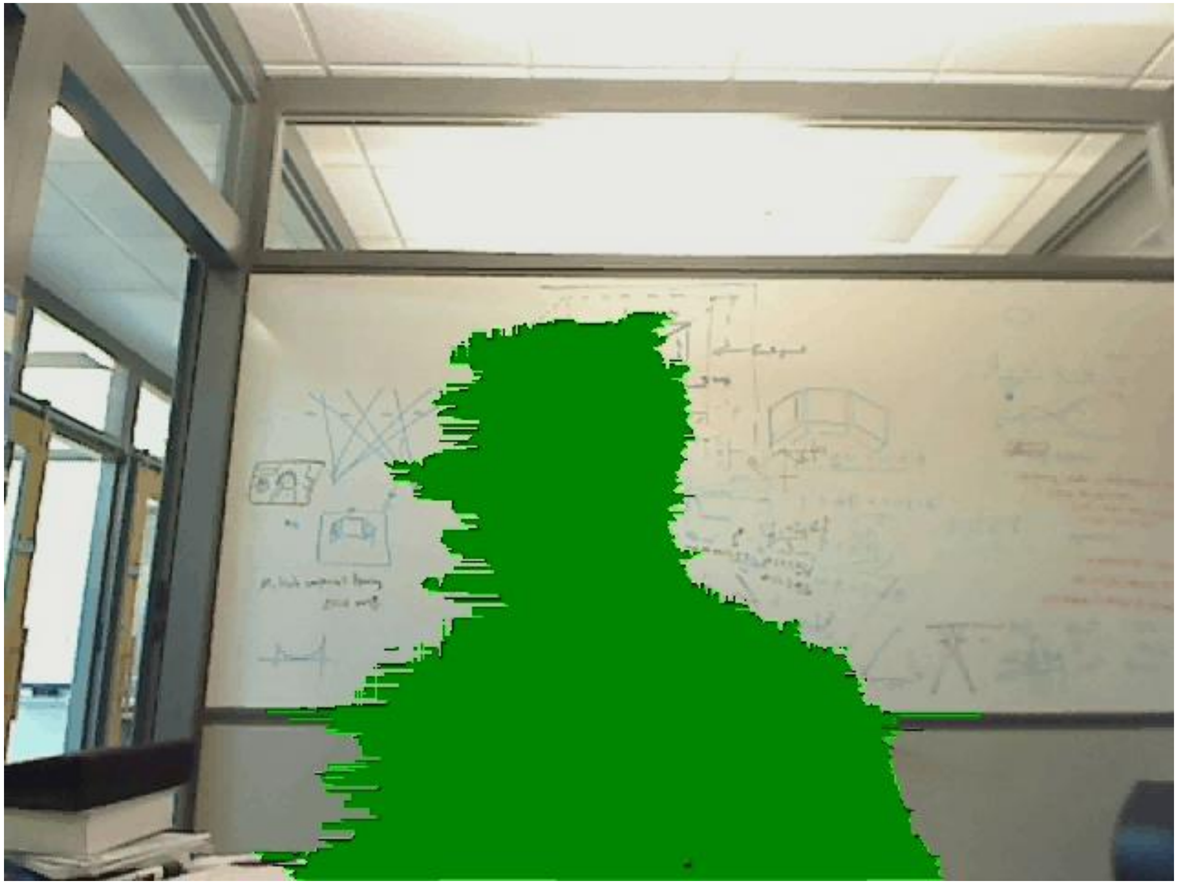
Before the Partial Ghosting Removal step, the boundaries of the occluding foreground regions are updated by exploiting a temporal continuity constraint as follows:

$$C_{x,y,n} = \begin{cases} B, & if \left\| P_{x,y,n} - P_{x,y,n-m} \right\|_{L_1} < t_3 \; for \; all \; m = 1, \dots, M \\ SF, & else \end{cases}$$

where (***x,y***) denotes the location of a boundary pixel of an OC in the current frame ***n*** .

In our simulation, the threshold t_3 = 6 and M = 4

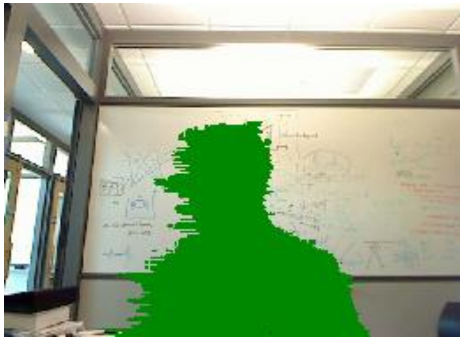
4. Simulation Results and Analysis



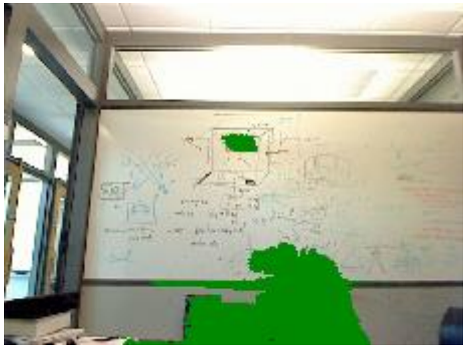
Progressive background extraction using spatio-temporal continuity for the 640x480 "Office" sequence

(a) Original Frame 25

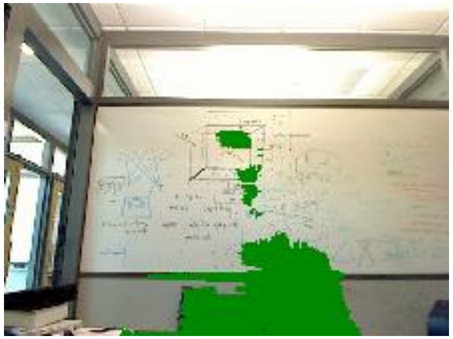
(b) Original Frame 88



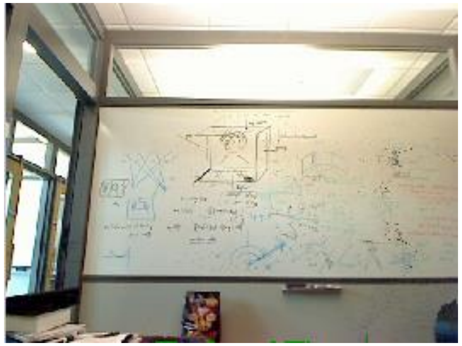
(c) Initial background
after 6 frames



(d) Background after 27 frames



(e) Background after 28 frames
(blob removed)



(f) Extracted background
after 63 frames

Performance Metrics for Foreground Detection

$$\text{Recall} = \frac{\text{Number of pixels correctly detected in the foreground}}{\text{Total number of pixels in the foreground given by Ground Truth}}$$

$$\text{Precision} = \frac{\text{Number of pixels correctly detected in the foreground}}{\text{Total number of pixels detected in the image as foreground}}$$

Performance Evaluation : 640x480 Office sequence

	Frame 30		**Frame 50**		**Frame 70**	
Method	**Rec.**	**Prec.**	**Rec.**	**Prec.**	**Rec.**	**Prec.**
MoG[2]	0.226	0.756	0.530	0.761	0.619	0.735
Global Motion Comp. [9]	0.910	0.297	0.925	0.289	0.688	0.238
Block Motion Parameters [12]	0.778	0.912	0.957	0.619	0.975	0.423
Proposed Method	0.986	0.927	0.993	0.772	0.956	0.692

Performance evaluation : 176 x 144 Hall Monitor sequence

	Frame 40		Frame 50		Frame 60	
Method	**Rec.**	**Prec.**	**Rec.**	**Prec.**	**Rec.**	**Prec.**
MoG[2]	0.531	0.558	0.599	0.549	0.586	0.554
Global Motion Comp. [9]	0.710	0.339	0.590	0.342	0.483	0.363
Block Motion Parameters [12]	0.726	0.692	0.736	0.693	0.770	0.691
Proposed Method	0.700	0.711	0.719	0.685	0.768	0.686

Analysis of Simulation Results

• The proposed algorithm consistently yields higher recall and precision rates.

• Mixture of Gaussians [2], absorbs foreground pixels into the background model.

• The loss of precision in the method based on global motion parameters[9] is due to background recovery at a block level instead at a pixel level.

• The method of background recovery based on motion parameters [12], performs well on the Hall Monitor sequence, but not on the Office sequence due to non-uniform motion.

5. Conclusions

Conclusion

- A new approach for background estimation/subtraction allowing:
  - Complete ghosting removal based on boundary-level spatial continuity constraints
  - Partial ghosting removal based on pixel-level spatio-temporal continuity constraints.
- Robust to non-uniform as well as uniform motion.
- Resilient to pauses in motion.
- Handles well deformable foreground objects and background clutter.

Questions ?

References

[1] C. Herley, "Automatic occlusion removal from minimum number of images," *IEEE International Conference on Image Processing*, vol. 2, pp. 1046-1049, 2005.
[2] C. Stauffer and W. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no.8, pp. 747-757, Aug 2000.
[3] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," *Proceedings of the 6th European Conference on Computer Vision-Part II*, pp.751-767, 2000.
[4] Xun Xu and T.S. Huang , "A Loopy Belief Propagation approach for robust background estimation", *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1 – 7, 2008.
[5] F. El Baf, T. Bouwmans and B. Vachon, "A fuzzy approach for background subtraction," *IEEE International Conference on Image Processing*, pp.2648 – 2651, 2008.
[6] Z. Wu, J. Bu and C. Chen, "Detection and location of people in video streams by fusion of color, edge and motion information," *IEEE International Conference on Image Processing*, 2002.
[7] K.-P. Karmann and A. Brandt, "Moving object recognition using and adaptive background memory," *Time-Varying Image Processing and Moving Object Recognition*, V. Cappellini, ed., 1990.
[8] H. Yuan, Y. Chang and Y. Ma, "A moving objects extraction method based on true motion information," *International Conference on Visual Information Engineering*, pp. 638 – 642, 2008.
[9] M. Unger, M. Asbach and P. Hosten, "Enhanced background subtraction using global motion compensation and mosaicing," *IEEE International Conference on Image Processing*, 2008.
[10] M. Ling and X. Mei, "Automatic segmentation of moving objects in video sequences based on spatio-temporal information," *International Conference on Communications, Circuits and Systems*, 2007.

[11] C-M Mak and W-K Cham, "Fast video object segmentation using Markov random field," *IEEE Workshop on Multimedia Signal Processing*, pp.343 – 348, 2008.
[12] S. Varadarajan, L.J. Karam and D. Florencio, "Background recovery from video sequences using motion parameters," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009.
[13] D. Florencio and R. Schafer, "Decision-based median filter using local signal statistics," Proc. of VCIP, 1994.
[14] D. Florencio, R, Schafer, "Post-sampling aliasing control for natural images," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 1995.
[15] Y. Rui, D. Florencio, W. Lam, and J Su, "Sound source localization for circular arrays of directional microphones," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2005.
[16] D. Florencio, "Motion sensitive pre-processing for video," *IEEE International Conference on Image Processing*, 2001.
[17] D. Florencio, "Image de-noising by selective filtering based on double-shot pictures," *IEEE International Conference on Image Processing*, 2005.
[18] C. Zhang, Z. Yin, D. Florencio, "Improving depth perception with motion parallax and its application in teleconferencing," *Proc. of MMSP*, 2009.
[19] D. Florencio, R. Armitano, R. Schafer, "Motion transforms for video coding, in Proc. of ICIP, 1996.
[20] R. Armitano, D. Florencio, R. Schafer, "The motion transform: a new motion compensation technique," in Proc. of ICASSP, 1996.
[21] S Varadarajan, LJ Karam, D Florencio, "Background subtraction using spatio-temporal continuities," *Proc. of EUVIP, 2010*.