Leibniz Universität Hannover

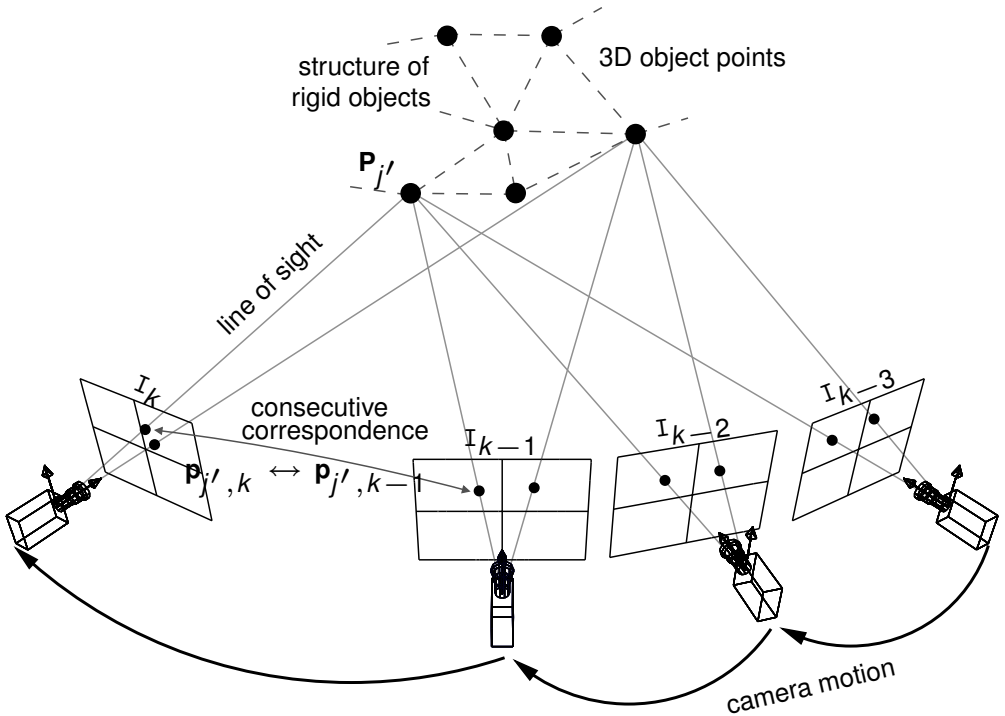
Occlusion Handling for the Integration of Virtual Objects into Video

Kai Cordes, Björn Scheuermann,
Bodo Rosenhahn, Jörn Ostermann

Institut für Informationsverarbeitung
Leibniz Universität Hannover

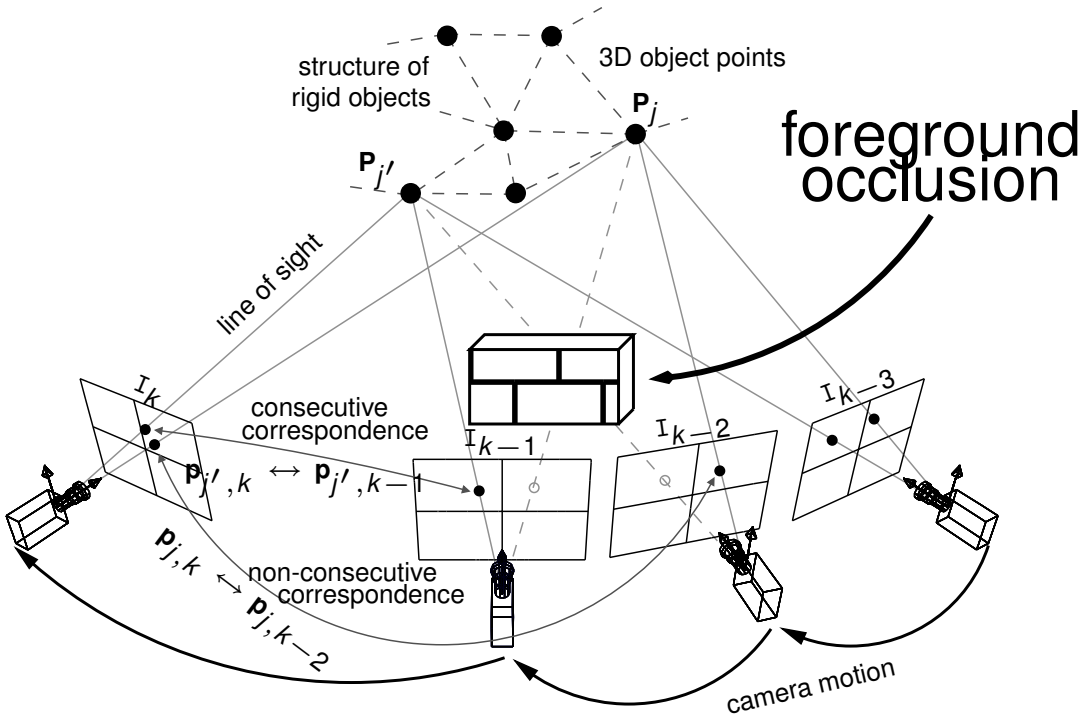
VISAPP 2012 - International Conference on Computer Vision Theory and Applications

Structure and Motion Recovery (SAM)



- Feature detection
- Correspondence analysis
- Outlier elimination
- Bundle adjustment

Structure and Motion Recovery (SAM)



- Occlusion

⇒ Trajectories lost

⇒ Error-prone object points

⇒ Drift

Integration of Virtual Objects into Video

Camera Motion Estimation

- High-Accuracy needed: Reprojection Errors < 0.25 pel [a]
- Usage of non-consecutive correspondences [b]

[a] Hillman et al., ICIP 2010
[b] Cordes et al., ISVC 2011

Occlusion of Foreground Objects

- Objects are integrated by composition
- Problem: augmented objects should be occluded

Integration of Virtual Objects into Video

Camera Motion Estimation

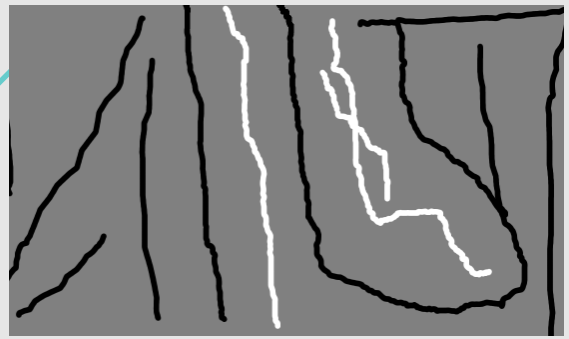
- High-Accuracy needed: Reprojection Errors < 0.25 pel [a]
- Usage of non-consecutive correspondences [b]

[a] Hillman et al., ICIP 2010
[b] Cordes et al., ISVC 2011

Occlusion of Foreground Objects

- Objects are integrated by composition
- Problem: augmented objects should be occluded

Interactive Segmentation:

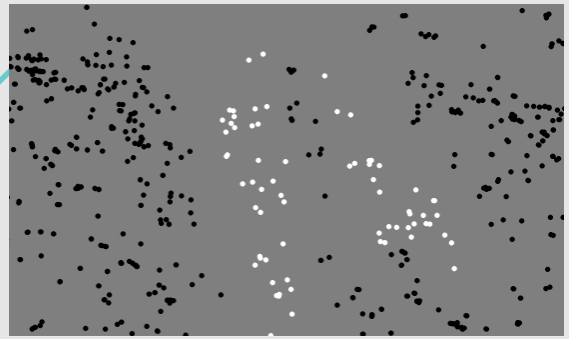
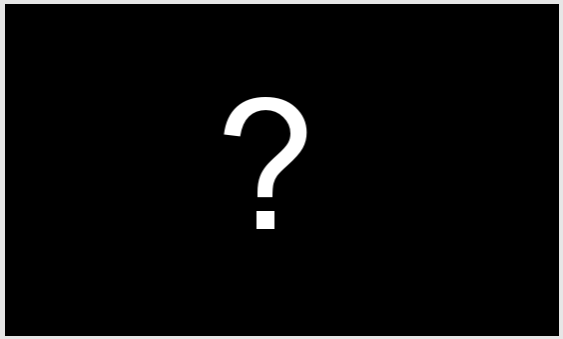

Boykov et al., ICCV 2001

Objectives

→ Minimize Energy: $E(f) = \sum_{p \in \mathcal{I}} D_p(f_p) + \sum_{(p,q) \in \mathcal{N}} V_{p,q}(f_p, f_q)$

- User Input:
  - Mark foreground and background with strokes
  - → Initialization of minimum cut / maximum flow

Automatic Segmentation:

Proposed, VISAPP 2012

Objectives

- Automatic Foreground Segmentation
  - → Generate images with foreground/background knowledge
  - → Use segmentation for compositing
- Extract occlusion Information
  - → Establish non-consecutive correspondences (KLT + SIFT)
  - → Find occluded object points and their locations in the images

Outline

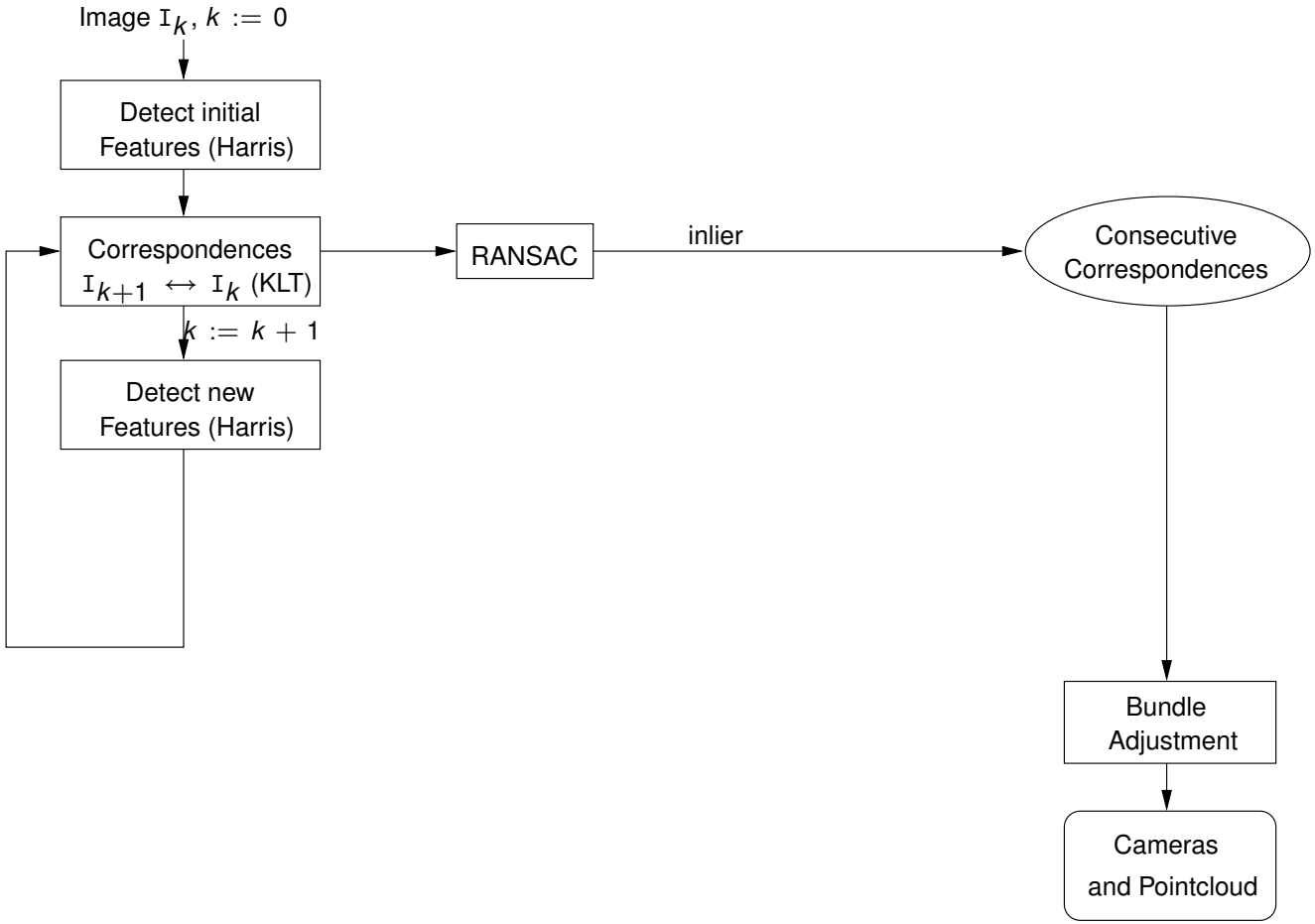
Feature Tracking in SAM Recovery

Automatic Foreground Segmentation

Experimental Results

Conclusion

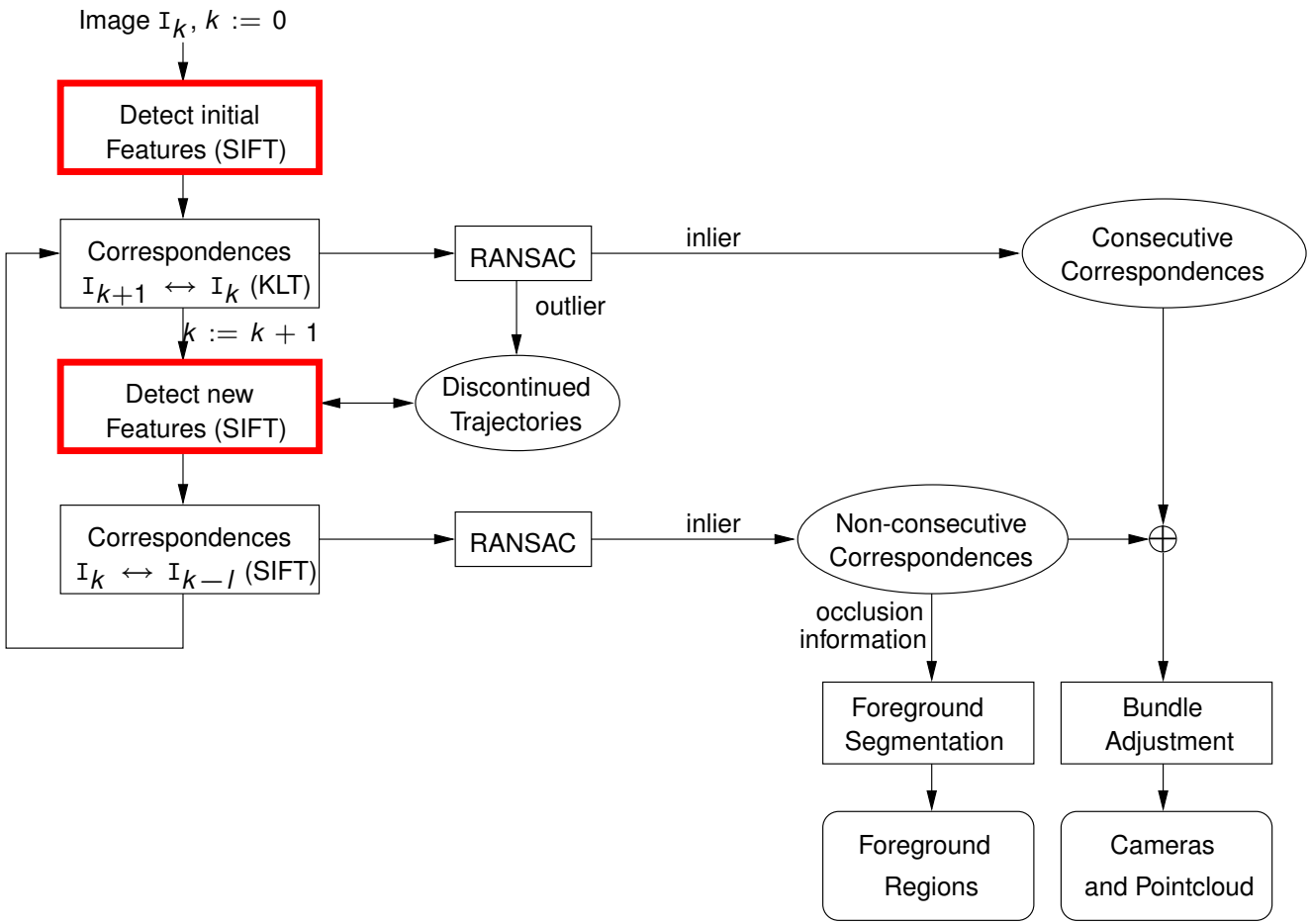
Workflow



SAM pipeline

- Standard SAM
- SIFT detection
- KLT tracking (consecutive)
- SIFT matching (non-consec.)
- Occlusion information
- Segmentation

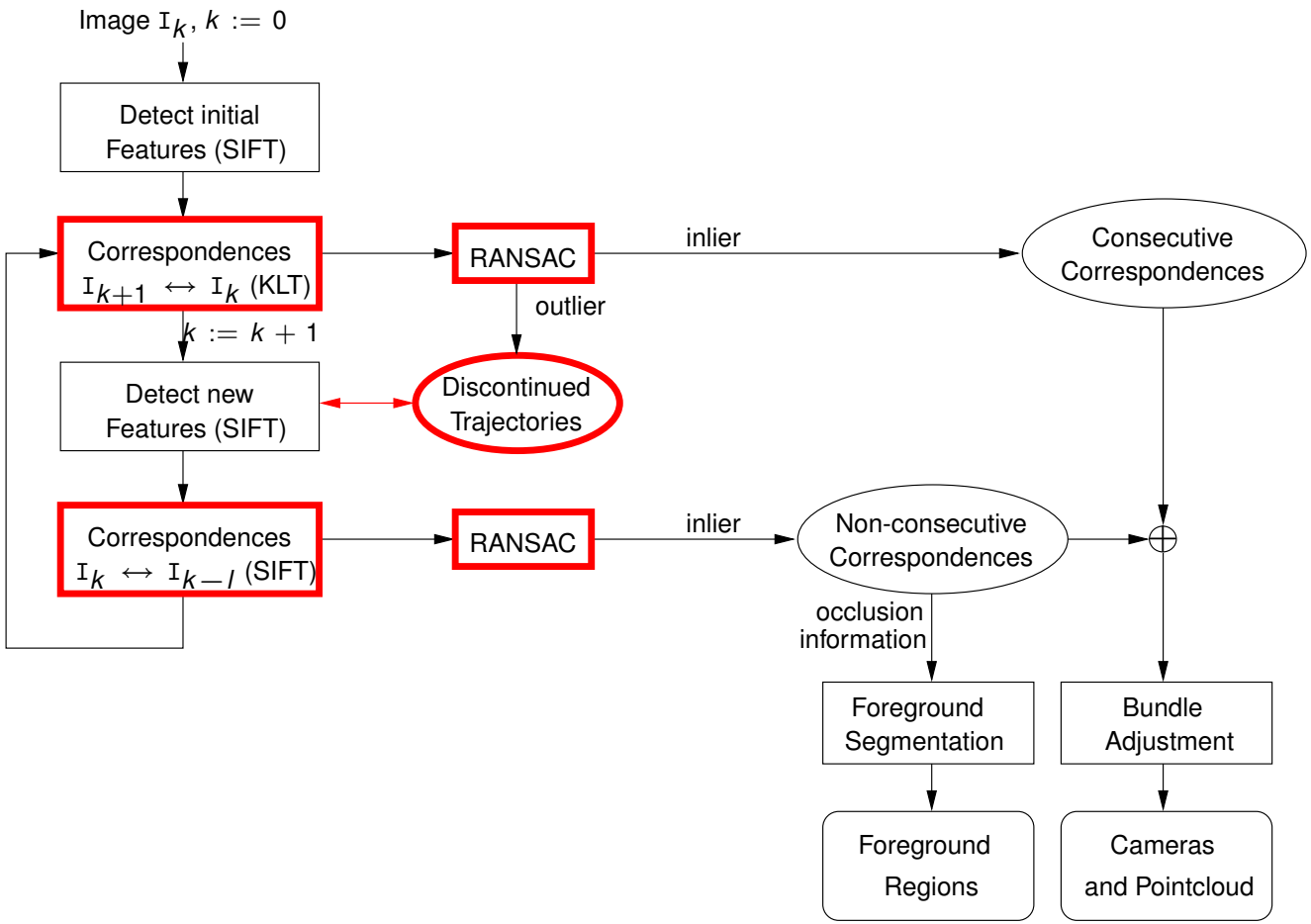
Workflow



SAM pipeline

- SIFT detection
- KLT tracking (consecutive)
- SIFT matching (non-consec.)
- Occlusion information
- Segmentation

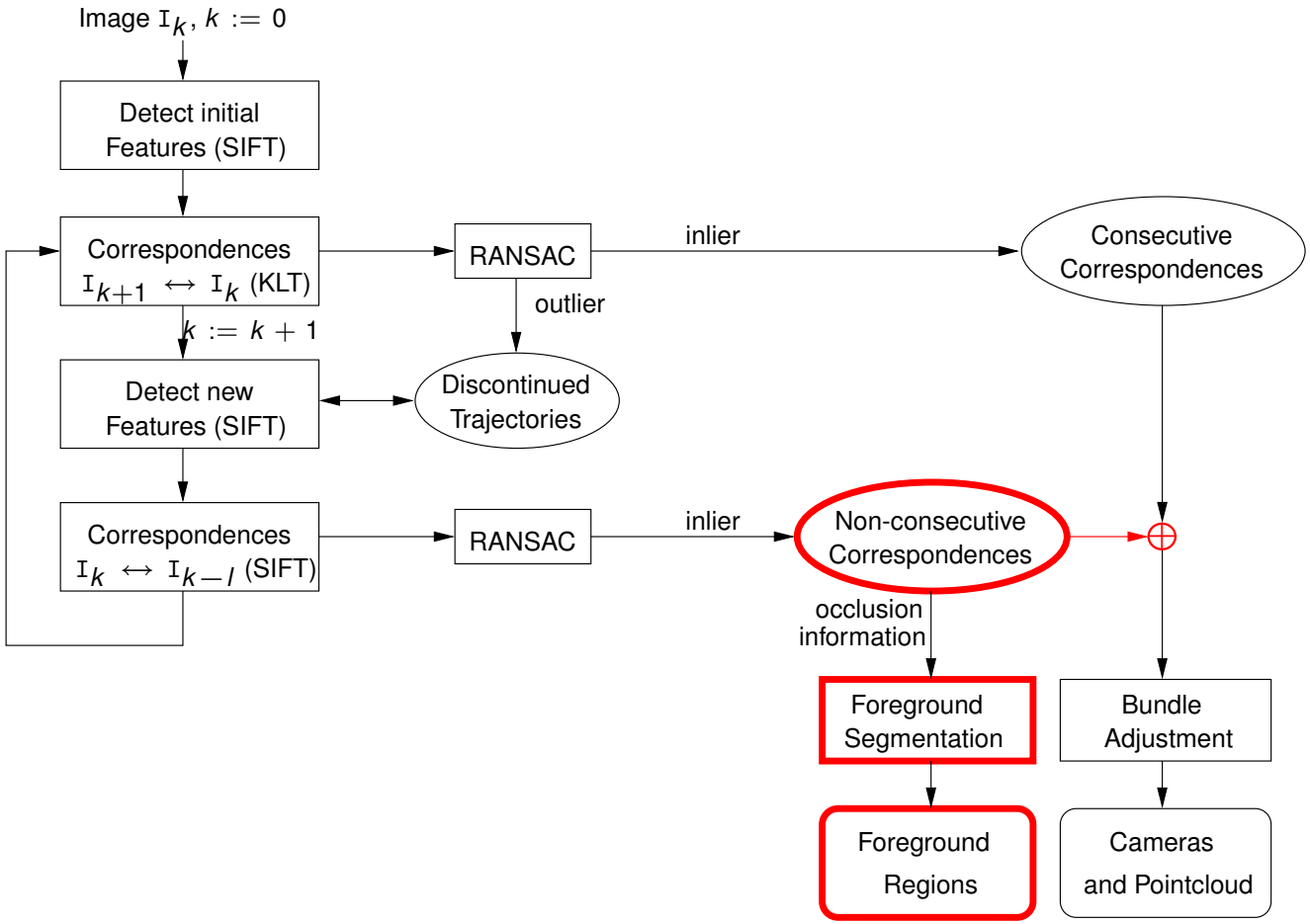
Workflow



SAM pipeline

- SIFT detection
- KLT tracking (consecutive)
- SIFT matching (non-consec.)
- Occlusion information
- Segmentation

Workflow



SAM pipeline

- SIFT detection
- KLT tracking (consecutive)
- SIFT matching (non-consec.)
- Occlusion information
- Segmentation

Outline

Occlusion Information

Non-consecutive correspondences

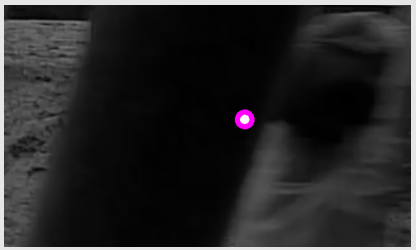
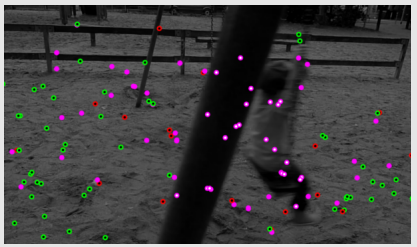
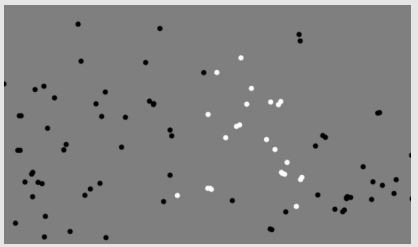
- $\mathbf{p}_{j,k} \leftrightarrow \mathbf{p}_{j,k-l-1}$ induces trajectory:

  $$\mathbf{t}_j^* = (\mathbf{p}_{j,k}^{visible}, \mathbf{p}_{j,k-1}^{occluded}, \ldots, \mathbf{p}_{j,k-l}^{occluded}, \mathbf{p}_{j,k-l-1}^{visible}, \mathbf{p}_{j,k-l-2}^{visible}, \ldots)$$

$\rightarrow$ $\mathbf{p}_{j,k-1}^{occluded} = \mathtt{A}_{k-1}\mathbf{P}_j^*, \ldots, \mathtt{A}_{k-l}\mathbf{P}_j^*$ (Reprojection)

$\rightarrow$ Object point $\mathbf{P}_j^*$ is

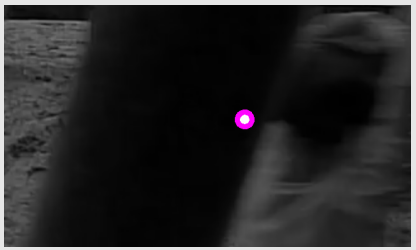
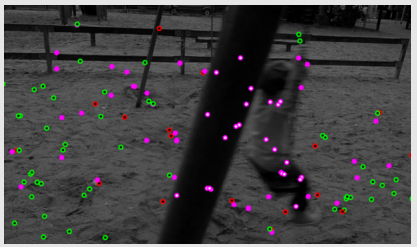
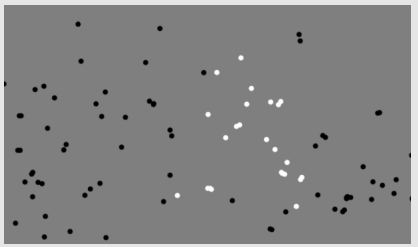
- visible in the current image $\mathtt{I}_k$
- occluded in the images $\mathtt{I}_{k-1}, \ldots, \mathtt{I}_{k-l}$.

Occlusion Information

Similarity Constraint to Verify Occlusion

- Non-Consecutive Correspondences induced by:
  1. **Occlusion**
  2. Scene content temporarily leaves field of view
  3. Repeated Texture
  4. Noise, e.g. motion blur
- Use neighborhood color histogram of $\mathbf{t}_j^*$ - features

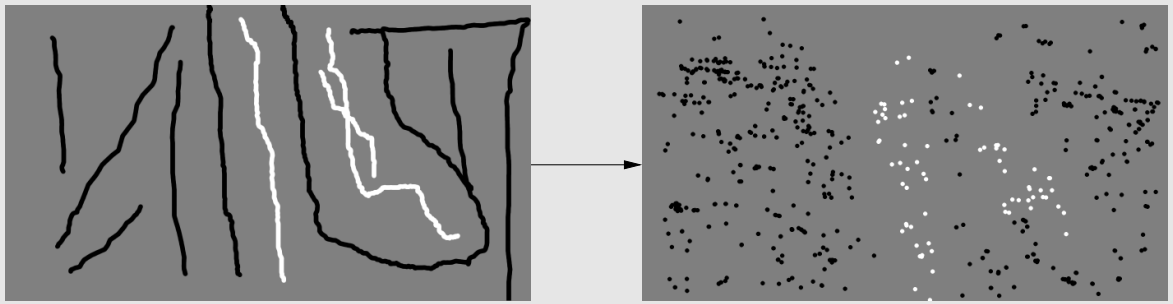
Automatic Segmentation

Use Graph-Cut based approach

- User interaction: manually drawn strokes
- Overhead: less than a second for 3 MPixels

Scheuermann et al., EMMCVPR 2011

Strokes derived automatically

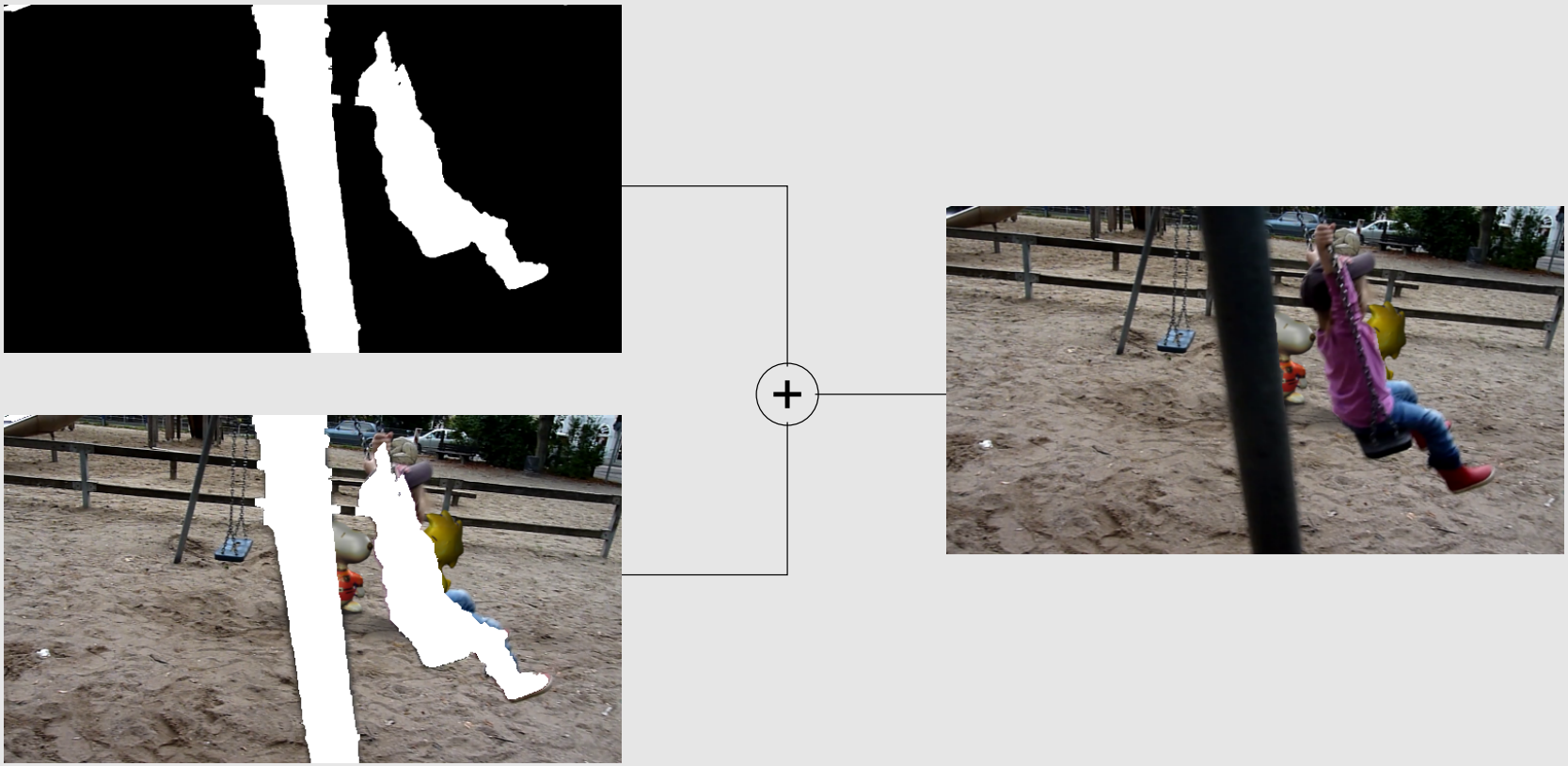


Compositing

Combine Input and Augmented Sequence

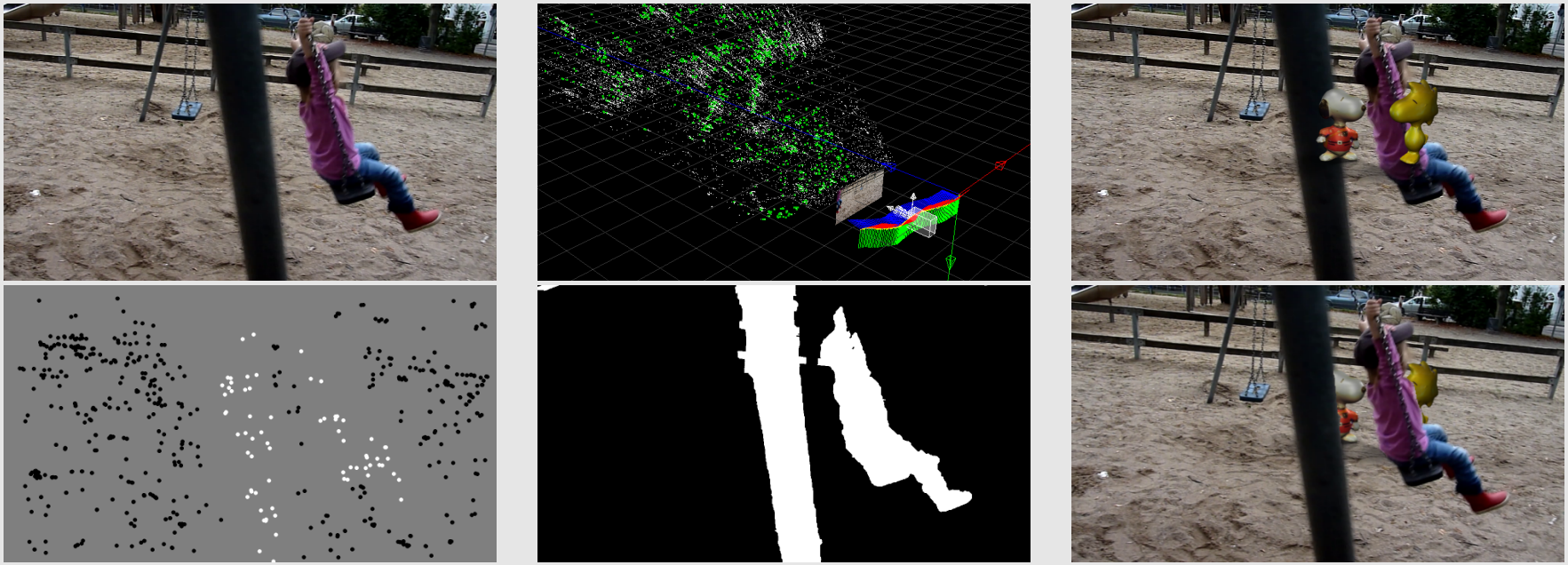
Compositing

Combine Input and Augmented Sequence



Outline

Playground Sequence



Video at:

`http://www.youtube.com/watch?v=RuSsIMfMlCE`

Extended Structure and Motion Recovery

- Combination of KLT and SIFT
- Extract Occlusion points

$\Rightarrow$ Automatic foreground segmentation

Realistic results

$\Rightarrow$ Highly accurate reconstruction

$\Rightarrow$ Integration and occlusion of virtual objects convincing

Future Work:

- Use temporal constraint for segmentation
- Use complementary feature types to maximize distribution

http://www.tnt.uni-hannover.de/staff/cordes/

Conclusion

Extended Structure and Motion Recovery

- Combination of KLT and SIFT
- Extract Occlusion points

$\Rightarrow$ Automatic foreground segmentation

Realistic results

$\Rightarrow$ Highly accurate reconstruction

$\Rightarrow$ Integration and occlusion of virtual objects convincing

Future Work:

- Use temporal constraint for segmentation
- Use complementary feature types to maximize distribution

http://www.tnt.uni-hannover.de/staff/cordes/

Conclusion

Extended Structure and Motion Recovery

- Combination of KLT and SIFT
- Extract Occlusion points

$\Rightarrow$ Automatic foreground segmentation

Realistic results

$\Rightarrow$ Highly accurate reconstruction

$\Rightarrow$ Integration and occlusion of virtual objects convincing

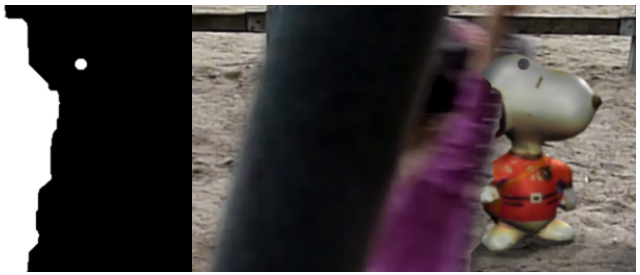
Future Work:

- Use temporal constraint for segmentation
- Use complementary feature types to maximize distribution

`http://www.tnt.uni-hannover.de/staff/cordes/`

Backup: limitations



Limitations

- Errors resulting from a misleading segmentation
- Left: no occlusion information in the fence, but visually similar
- Right: point is correct, but boundary smeared (motion blur)

→ Use temporal constraints for segmentation