Using Shape Priors to Regularize Intermediate Views in Wide-Baseline Image-Based Rendering

Cédric Verleysen[1], T. Maugey[2], P. Frossard[2], C. De Vleeschouwer[1]

[1] ICTEAM institute, UCL (Belgium) ; [2] LTS4 lab, EPFL (Switzerland)

cedric.verleysen@uclouvain.be

Organization of the presentation

- Context
- Current limitations in model/image-based rendering
- Contributions in wide-baseline rendering
  - Match epipolar segments based on virtual prior
  - Determine prior about the virtual shapes
- Experimental validation

Organization of the presentation

- Context
- Current limitations in model/image-based rendering
- Contributions in wide-baseline rendering
  - Match epipolar segments based on virtual prior
  - Determine prior about the virtual shapes
- Experimental validation

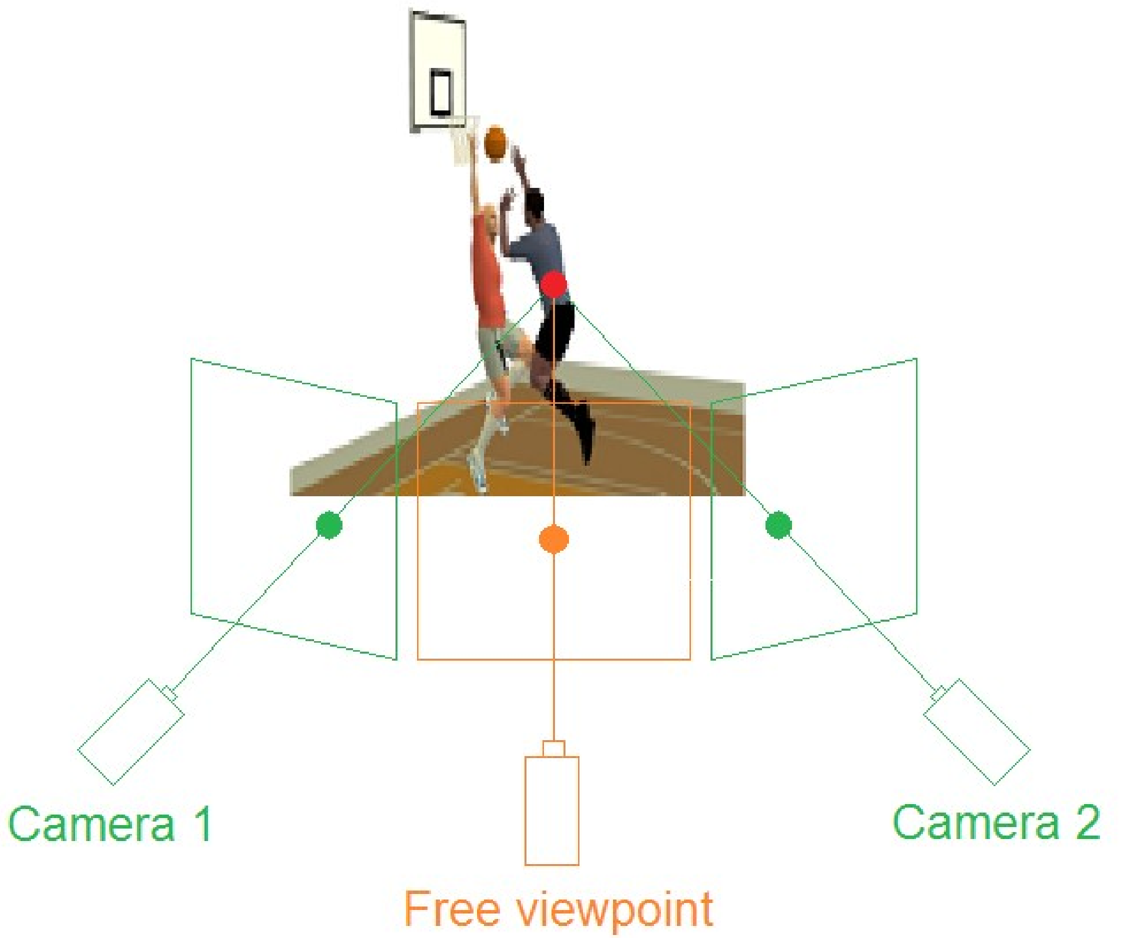
Conventional video production fixes the viewpoint to one of the cameras that capture the scene

Weaknesses of existing solution

Main issues: Restricted camera coverage (cameras on the same side, outside the scene)

Non-smooth transitions between fixed cameras.

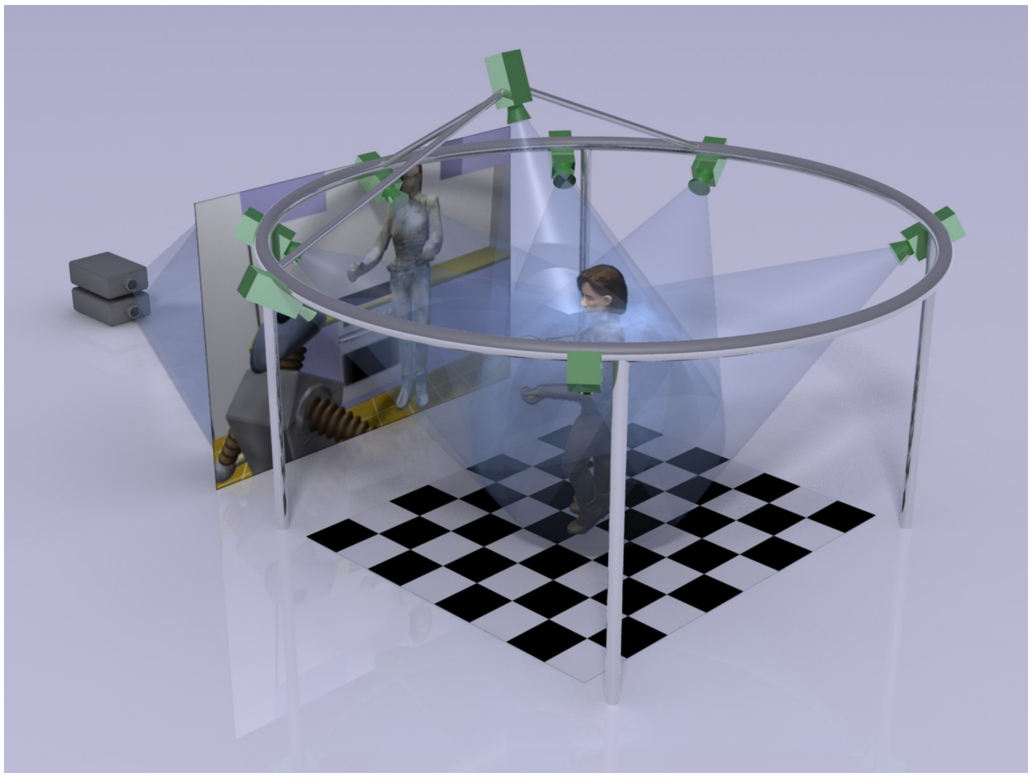
What is free viewpoint rendering ?

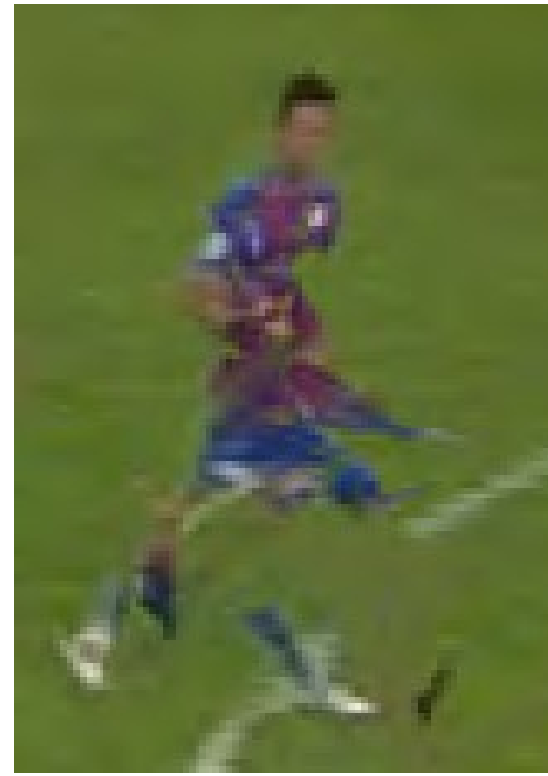


Organization of the presentation

- Context
- Current limitations in model/image-based rendering
- Contributions in wide-baseline rendering
    - Match epipolar segments based on virtual prior
    - Determine prior about the virtual shapes
- Experimental validation

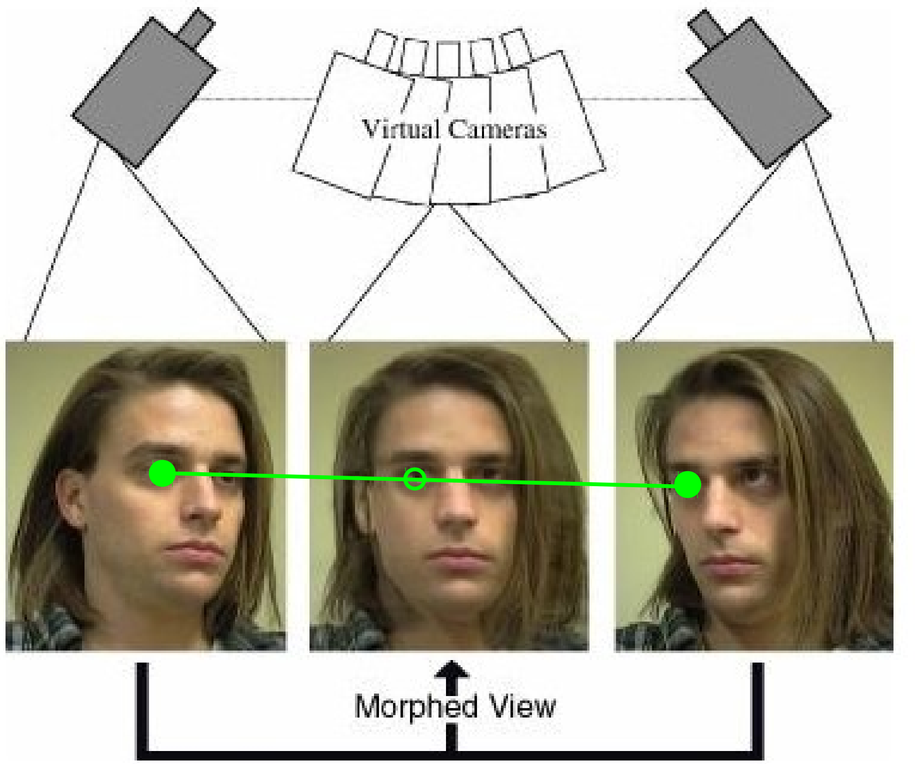
Model-based rendering [1-4] projects an estimated 3D model on the virtual camera





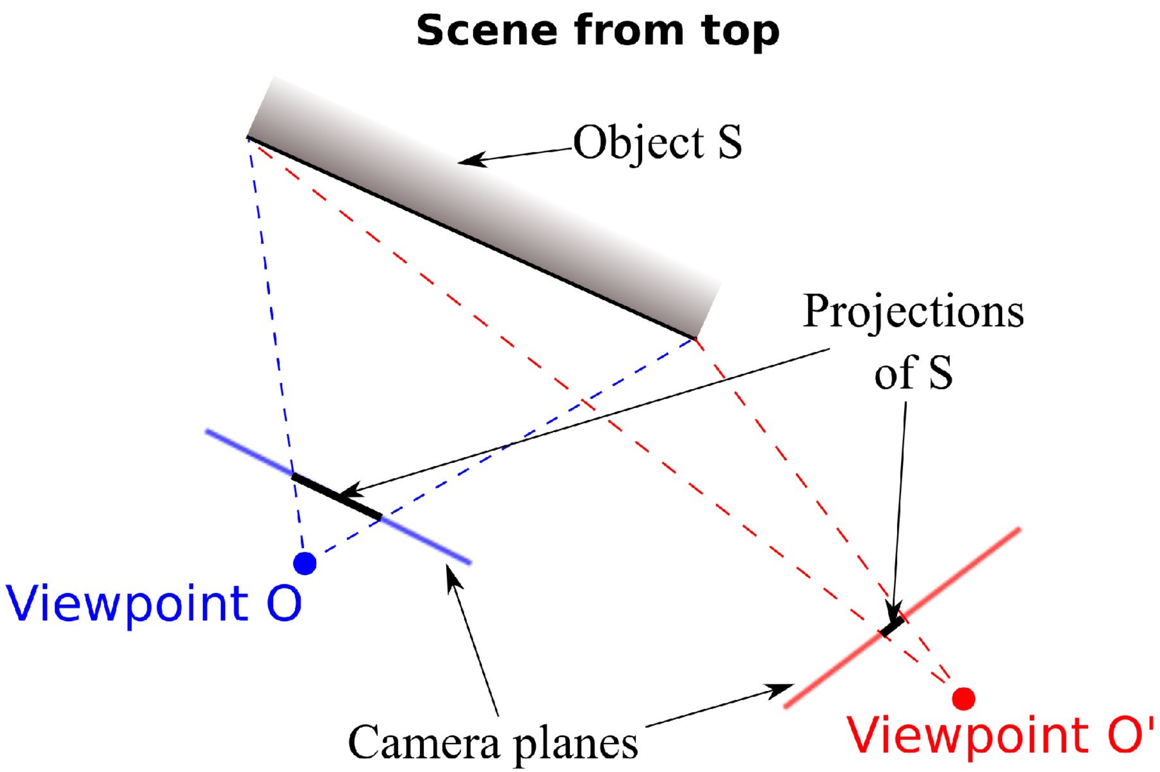
Realistic virtual view (accurate 3D model) only possible with dense camera coverage [5] or far views

Image-based rendering [6] interpolates intermediate virtual images based on dense matches



Assumption : one-to-one pixel correspondences [7-8]

In wide-baseline, several pixels in one basis image can correspond to one pixel in the other basis view (foreshortening effect)



State-of-the-art stereo matching [9-11] makes the assumption of one-to-one correspondences

Also, because there is no correspondence in the occluded areas, the virtual view has important holes

Filling the holes is an ill-posed problem and can lead to unrealistic virtual views

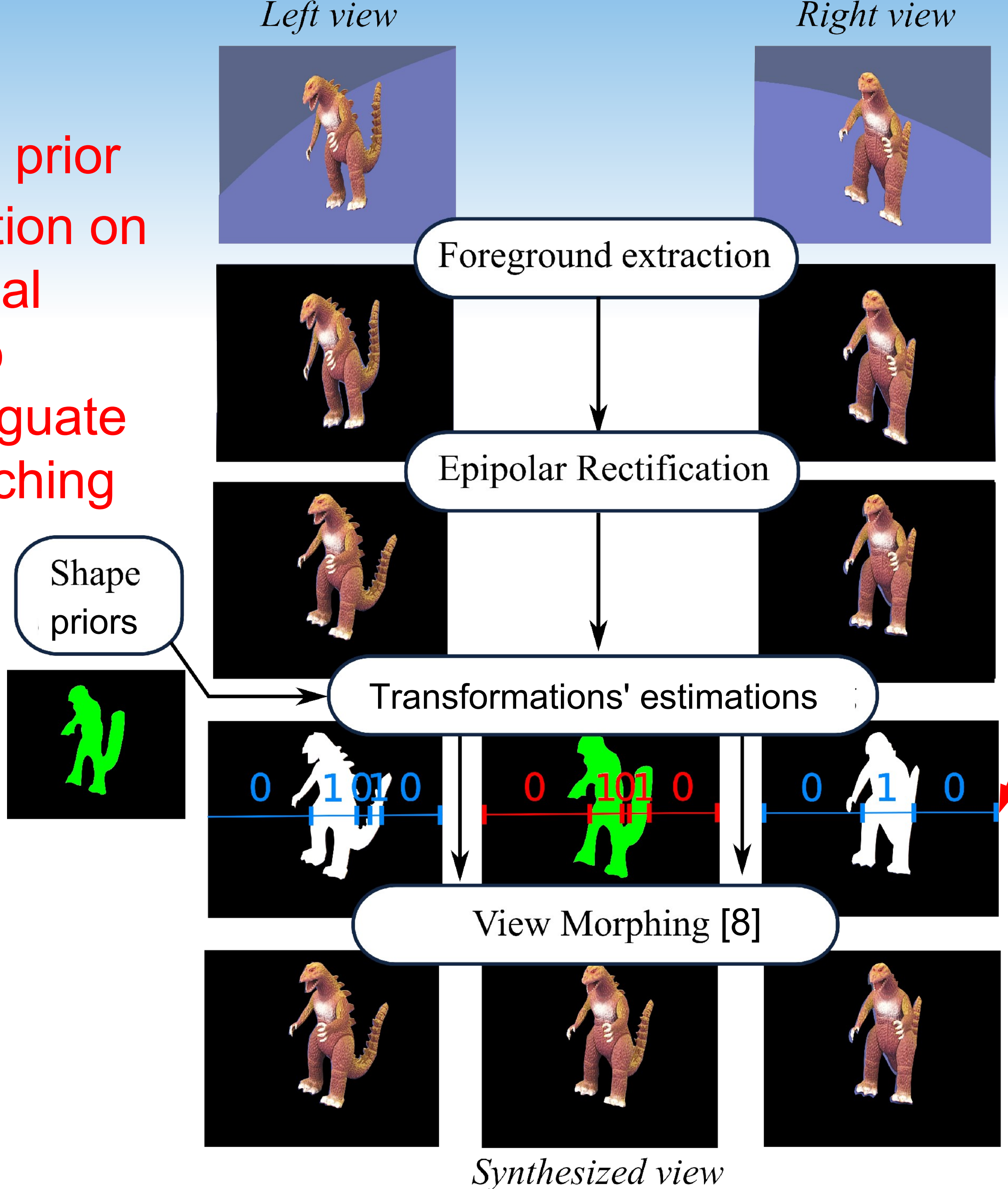
Organization of the presentation

- Context
- Current limitations in model/image-based rendering
- Contributions in wide-baseline rendering
  - Match epipolar segments based on virtual prior
  - Determine prior about the virtual shapes
- Experimental validation

Organization of the presentation

- Context
- Current limitations in model/image-based rendering
- Contributions in wide-baseline rendering
    - Match epipolar segments based on virtual prior
    - Determine prior about the virtual shapes
- Experimental validation

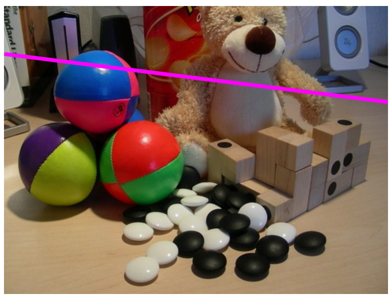
2) Add prior information on the virtual views to disambiguate the matching



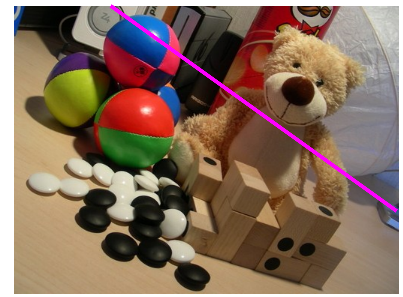
1) Consider epipolar foreground segments to tolerate foreshortening effect

What is an epipolar line segment ?

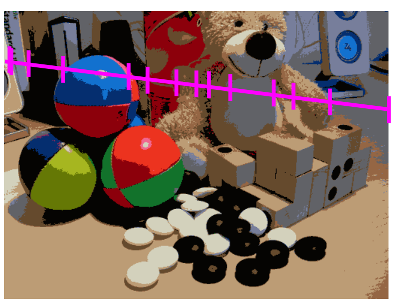
- New (and simple) image representation that describes the
  - Geometric constraints between the calibrated cameras



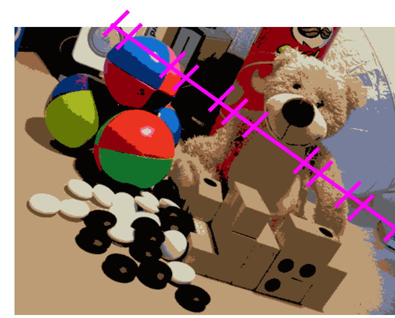
Epipolar geometry [12]



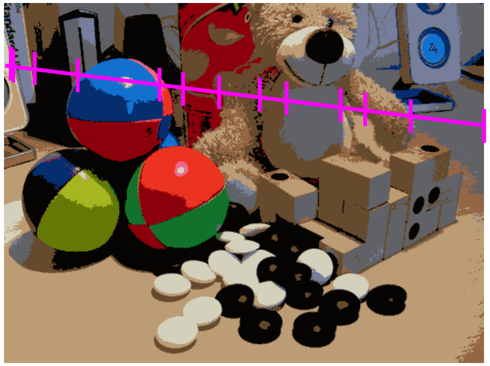
- Semanticat content of the observed scene



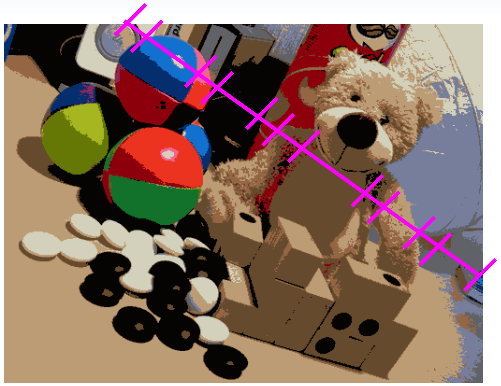
(Color) segmentation [13]

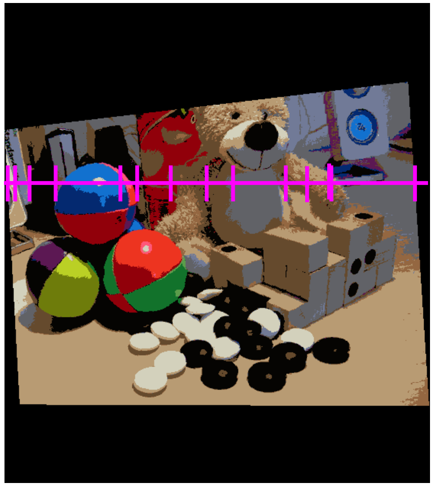


Advantages of epipolar line segments ?

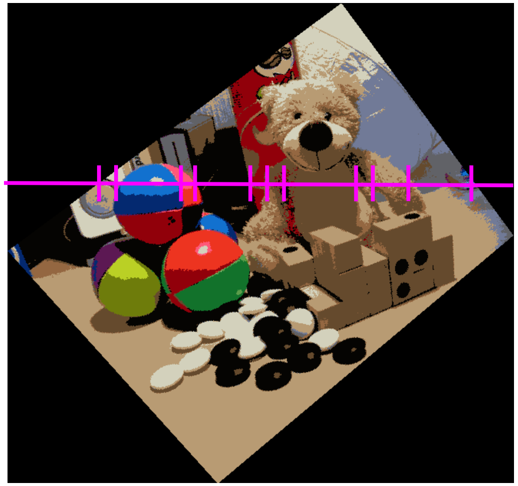


Model the foreshortening effect

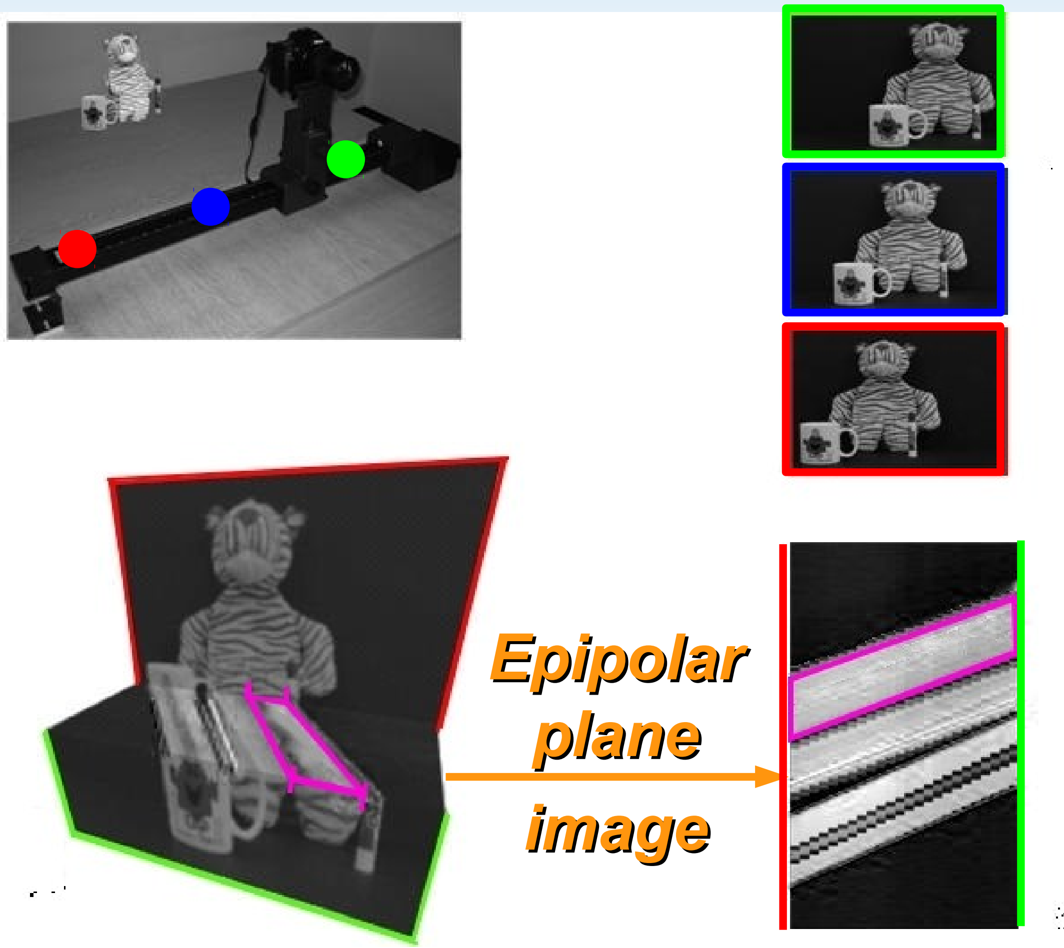




1D matching problem



What happens to the epipolar line segments when the viewpoint changes?



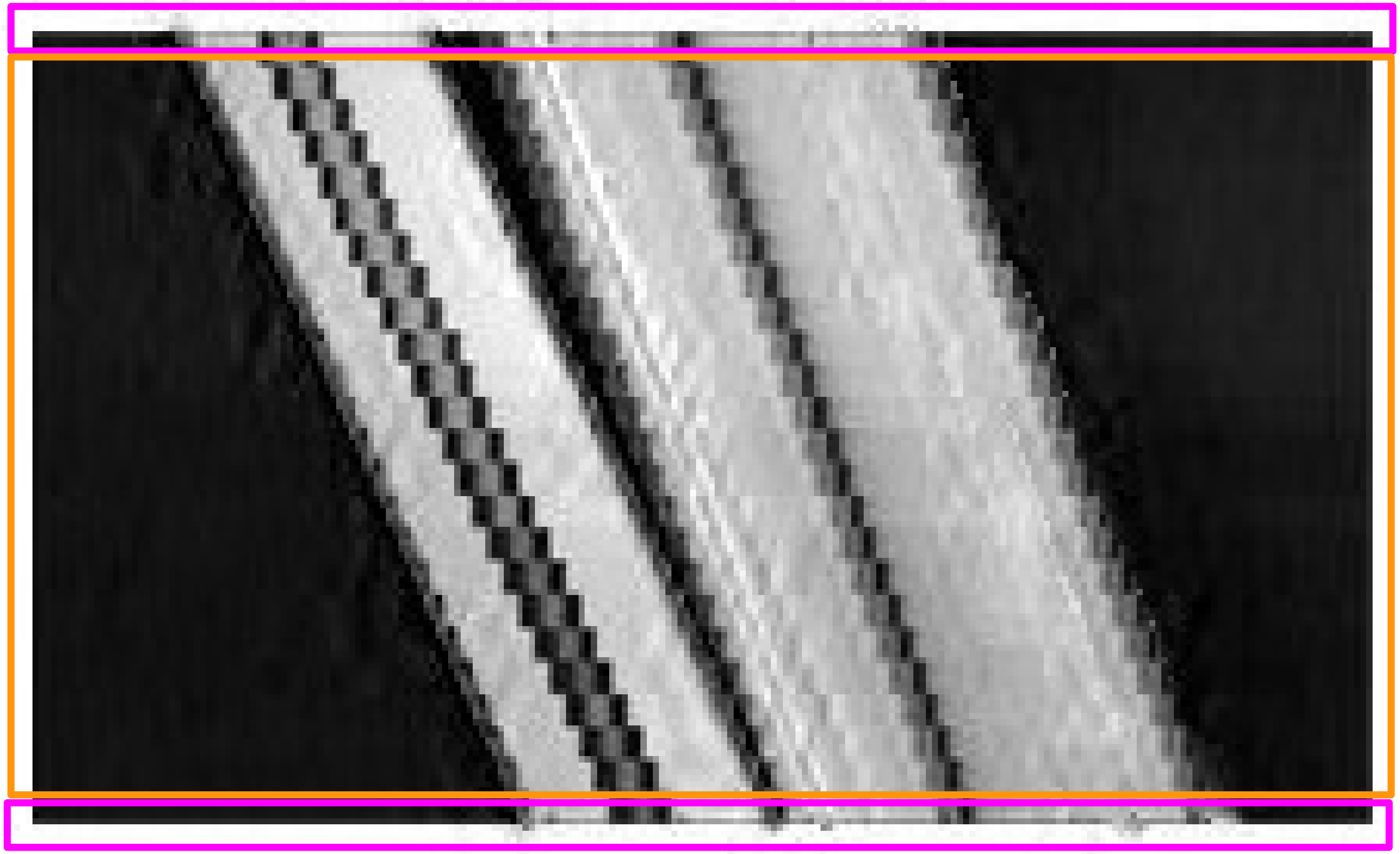
Epipolar line segments undergo a 1D translation and 1D scaling

View interpolation problem :

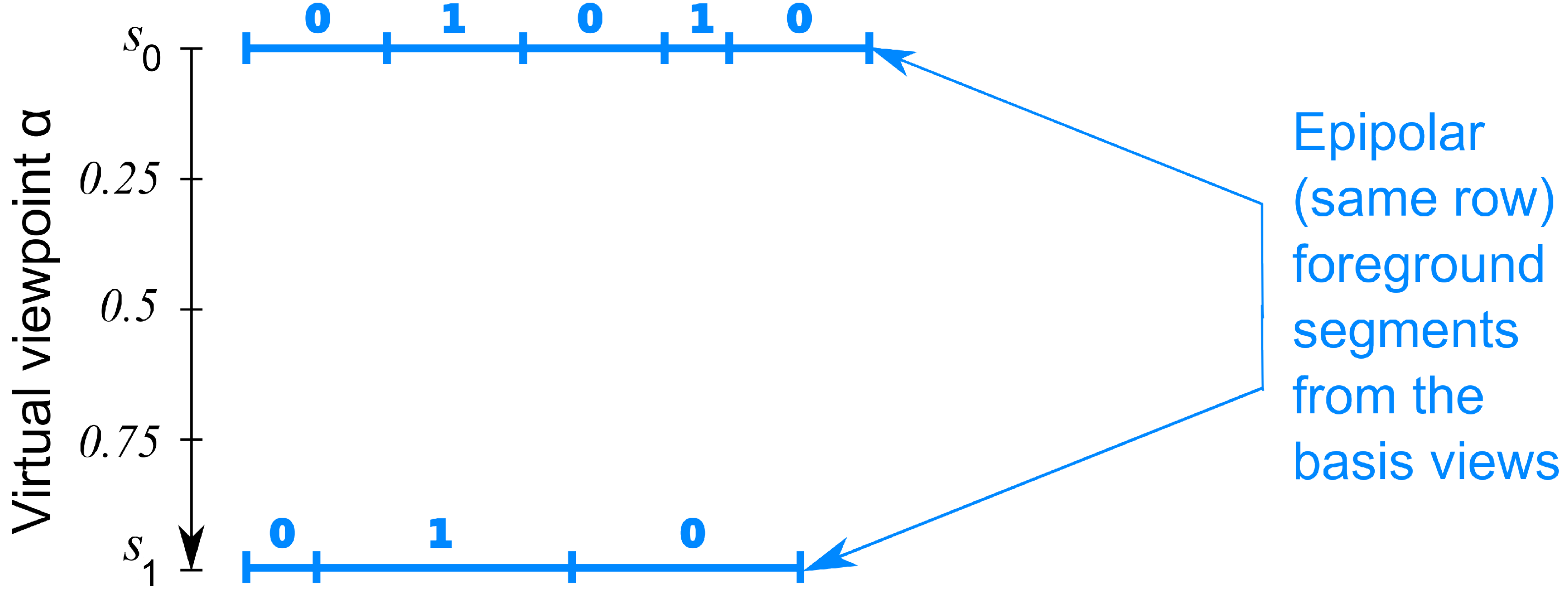
Given two epipolar sequences

View interpolation problem :

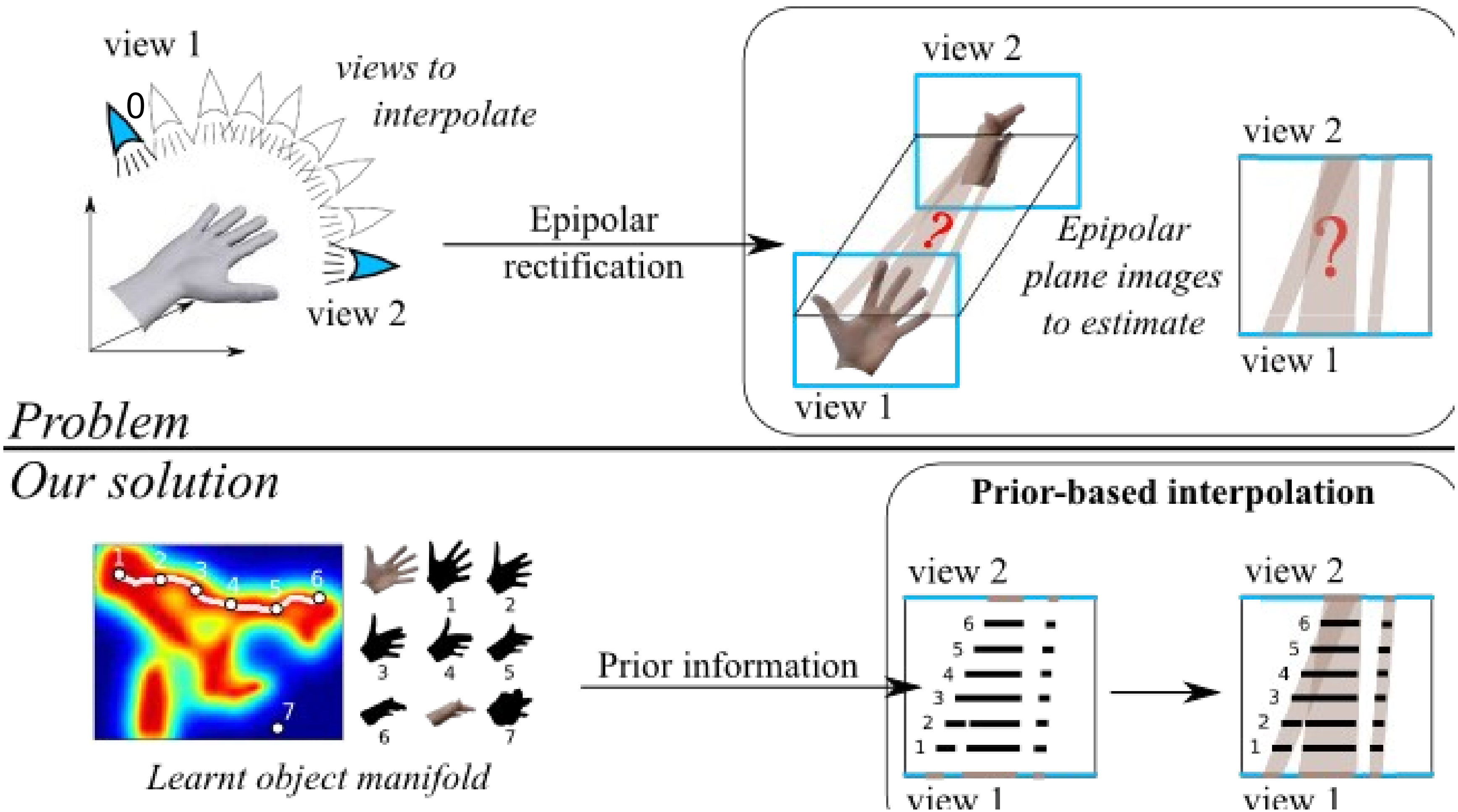
Given two epipolar sequences, reconstruct the EPI



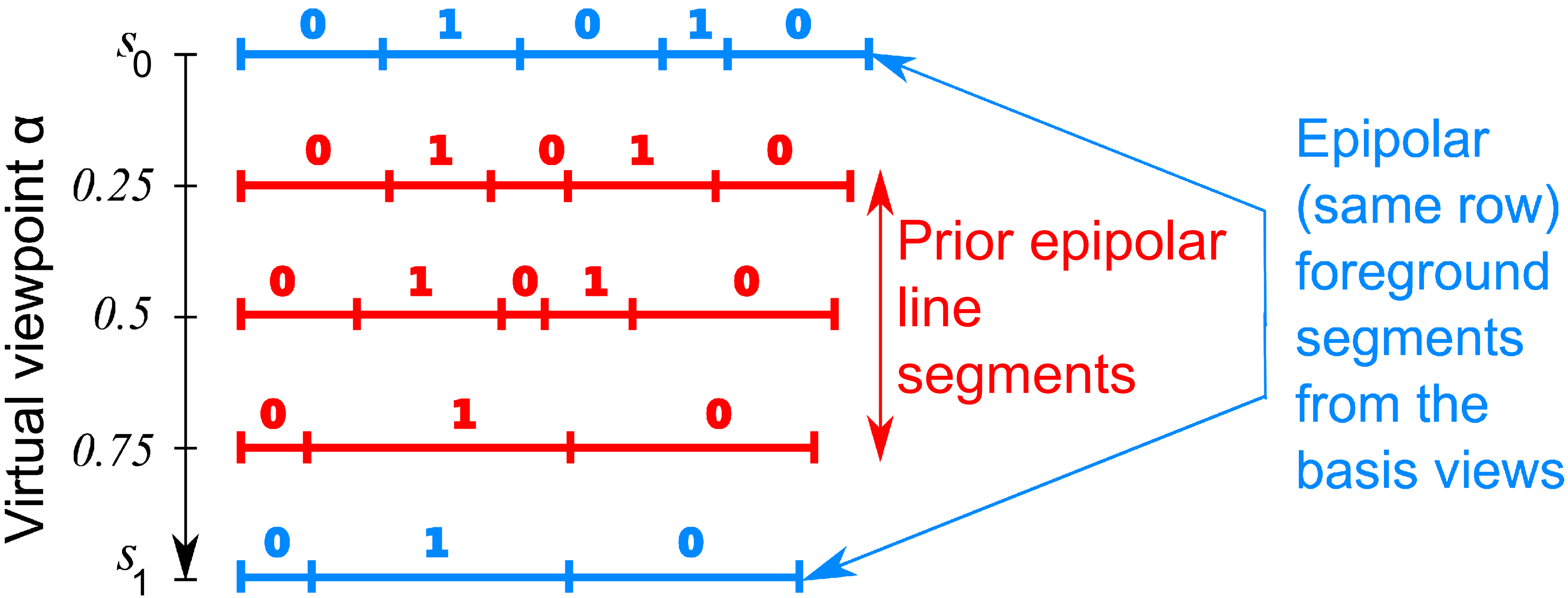
Due to occlusions, EPI reconstruction is an ill-posed problem



We propose to regularize the reconstruction so that the silhouettes of the reconstructed object tends to belong to the object's manifold.

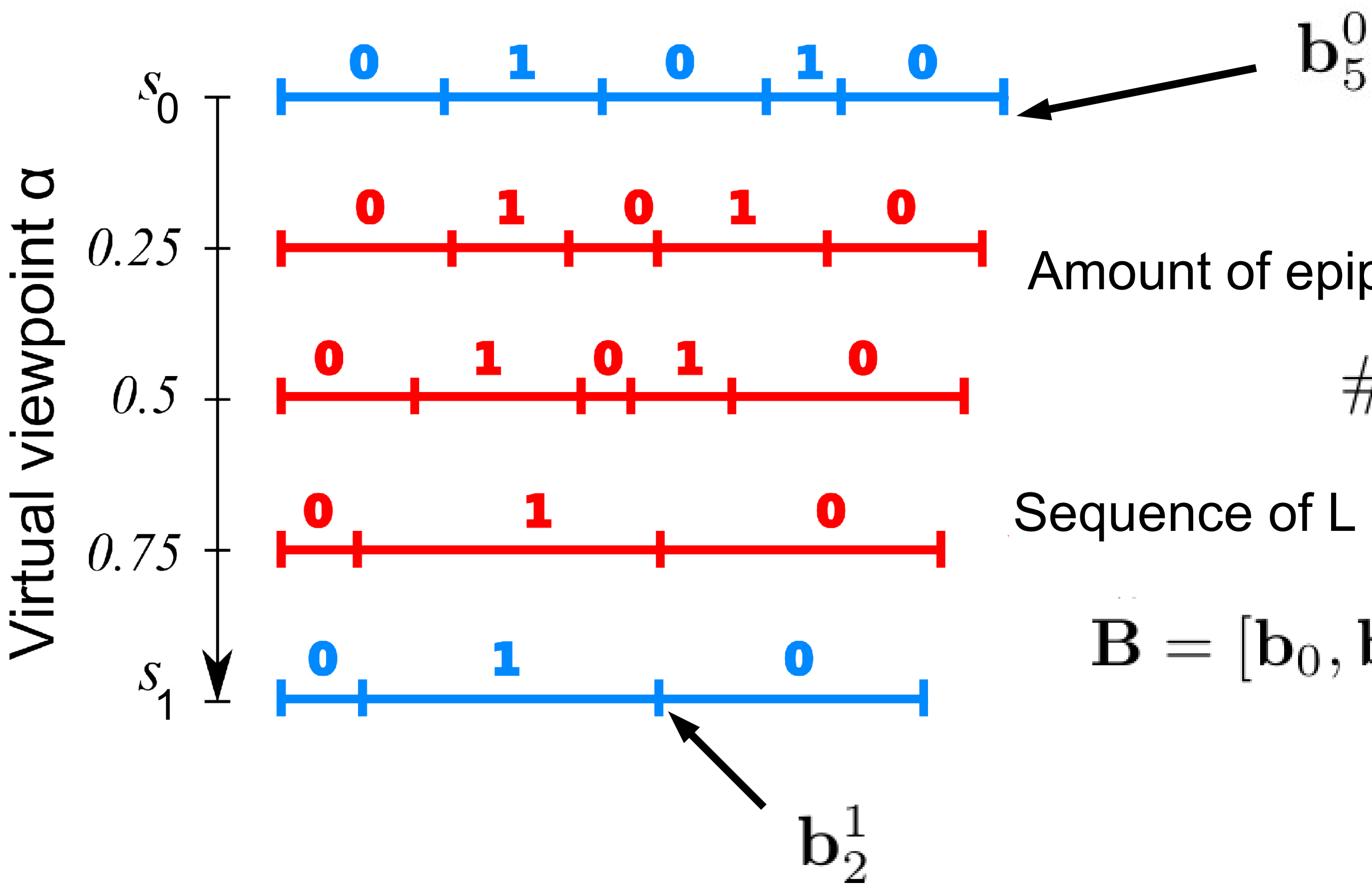


Imagine we have a prior knowledge about the epipolar line segments in the intermediate views...



Constrain (regularize) the possible transformations of the epipolar line segments

Instead of determining the segments' transformations, we estimate the displacement of their borders



Amount of epipolar segments:

$$\#\mathbf{S}$$

Sequence of L epipolar borders:

$$\mathbf{B} = [\mathbf{b}_0, \mathbf{b}_1, \cdots, \mathbf{b}_{\#\mathbf{S}}]$$

Step 1 : Indentify, in the basis views, the matched epipolar borders and the occluded ones

$$\min_{\mathbf{U}\in\{0,1\}^{L_0\times L_1}} \sum_{i=0}^{L_0}\sum_{j=0}^{L_1} \boxed{\mathbf{U}_{ij}\cdot\mathbf{D}_{ij}} + \lambda\cdot\boxed{\mathbf{C}_{ij}(\mathbf{1}-\mathbf{U}_{ij})}$$

Cost of matching $\mathbf{b}_i^0$ with $\mathbf{b}_j^1$

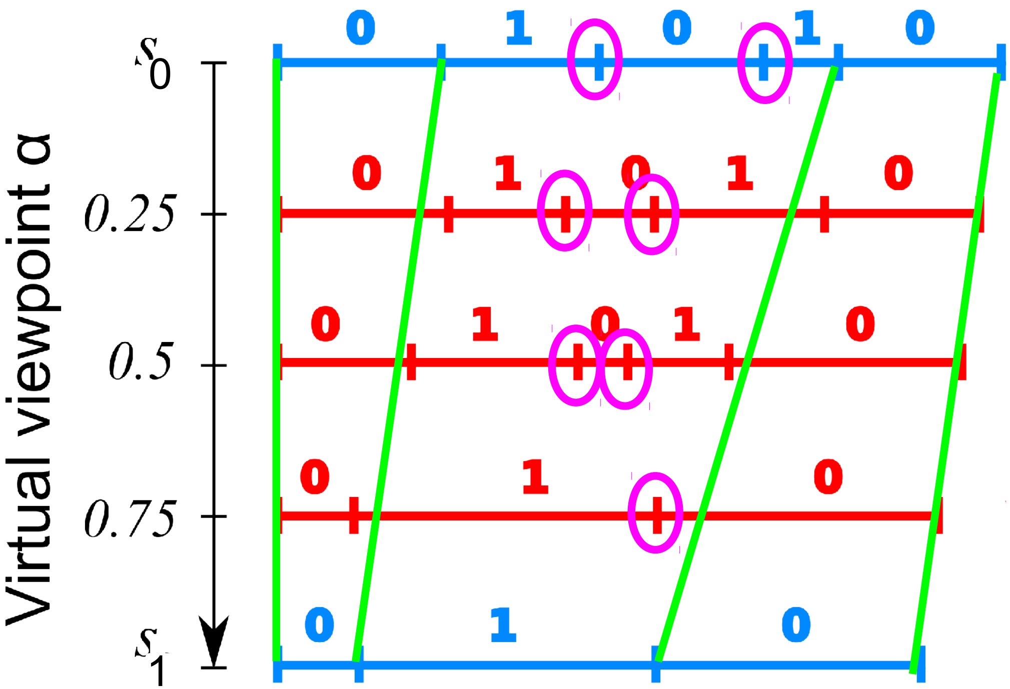
Cost of considering $\mathbf{b}_i^0$ (or $\mathbf{b}_j^1$) as occluded

with the indicator matrix $\mathbf{U} \in \{0,1\}^{L_0\times L_1}$

$$\text{s.t.} \begin{cases} \sum_{i=0}^{L_0} \mathbf{U}_{i,j} \leq 1 & \forall j \in \{0,\cdots,L_1\} \\ \sum_{j=0}^{L_1} \mathbf{U}_{i,j} \leq 1 & \forall i \in \{0,\cdots,L_0\} \end{cases}$$

➡ Solved by dynamic programming (Needleman and Wunsch algorithm [14])

Step 2 : Based on the matched epipolar borders, determine the prior borders that are occluded

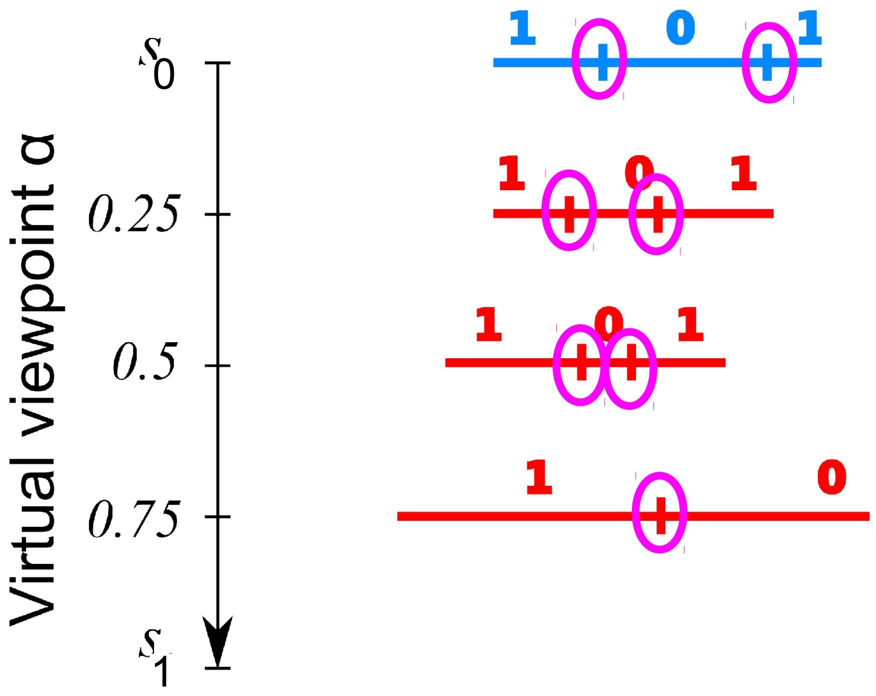


Prior borders are considered as occluded if their L1 distance to any of the interpolated path is higher than a threshold T.

T is set to an arbitrary (very) small value (*e.g.* 5% of the width of the interpolated image).

➡ The set of occluded prior borders is corrupted by matched prior borders

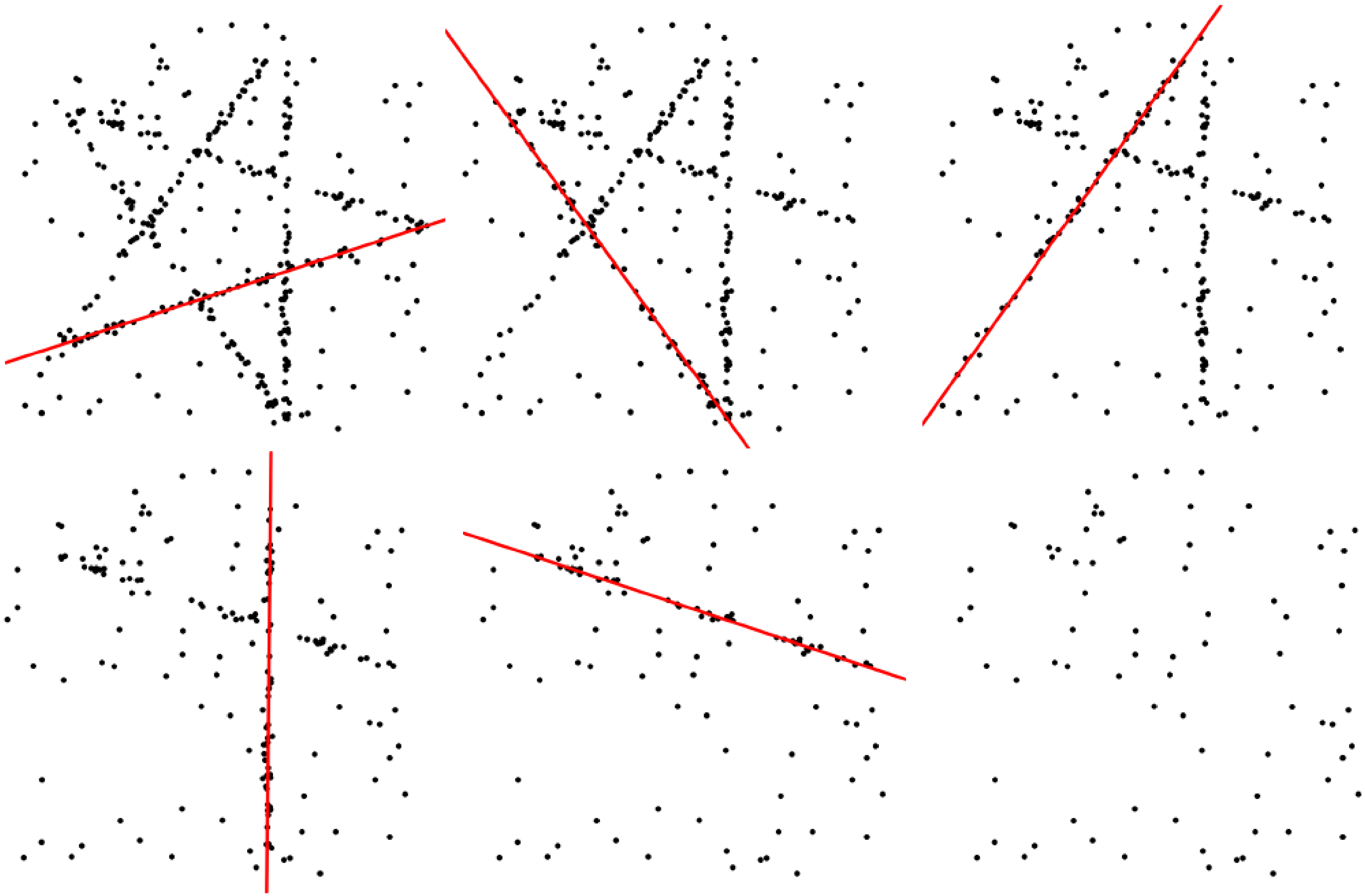
Step 3 : Estimate the trajectories of the occluded borders



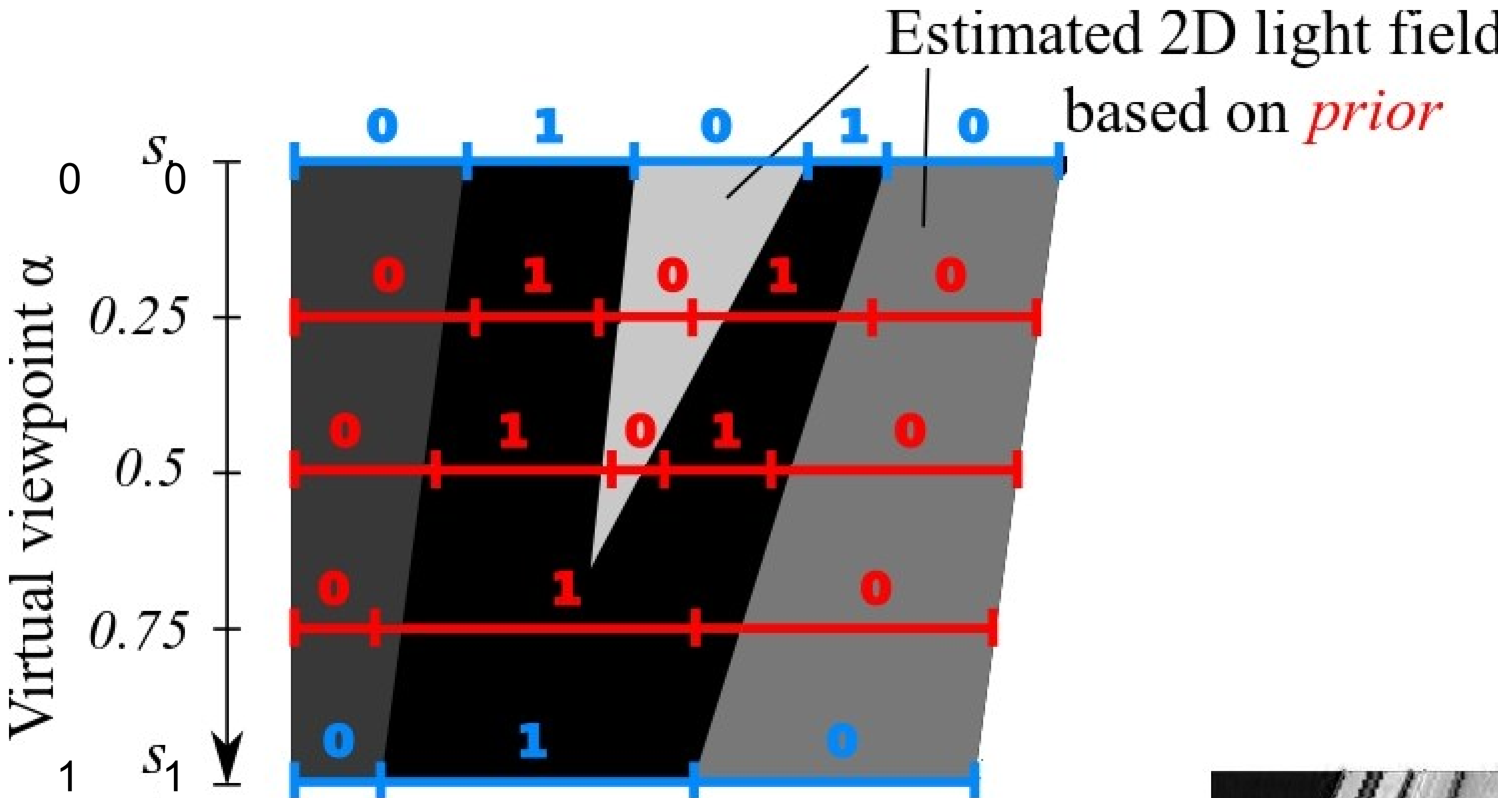
From the corrupted set of occluded prior borders

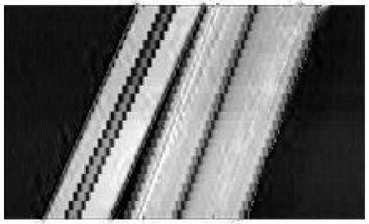
Apply RANSAC sequentially on borders with the same foreground/background transition

Sequential RANSAC [15]



These steps enables to estimate the how epipolar line segments transform in-between the views





Organization of the presentation

- Context
- Current limitations in model/image-based rendering
- Contributions in wide-baseline rendering
  - Match epipolar segments based on virtual prior
  - Determine prior about the virtual shapes
- Experimental validation

Defining shape priors in the intermediate views

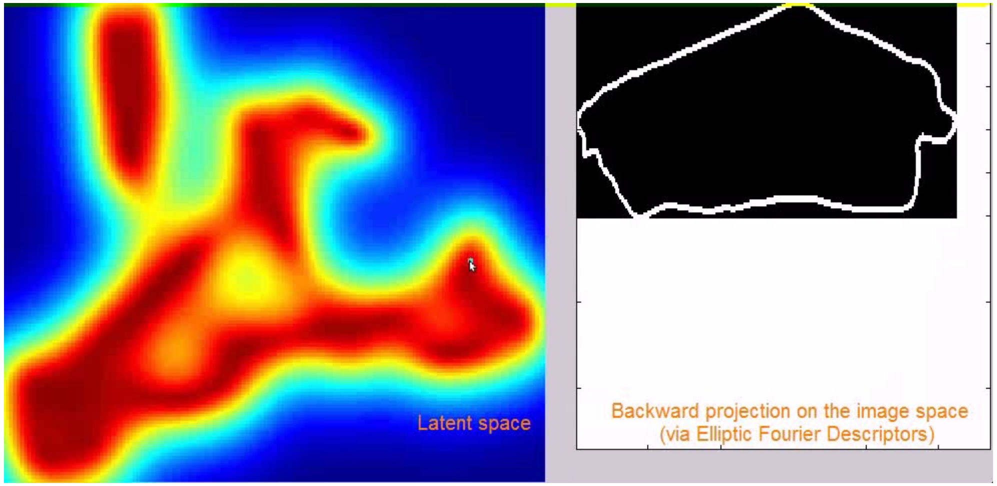
1) Learn a low-dimensional space representative of the shape of the analyzed object

2) Project the shapes of the basis views onto the low-dimensional space

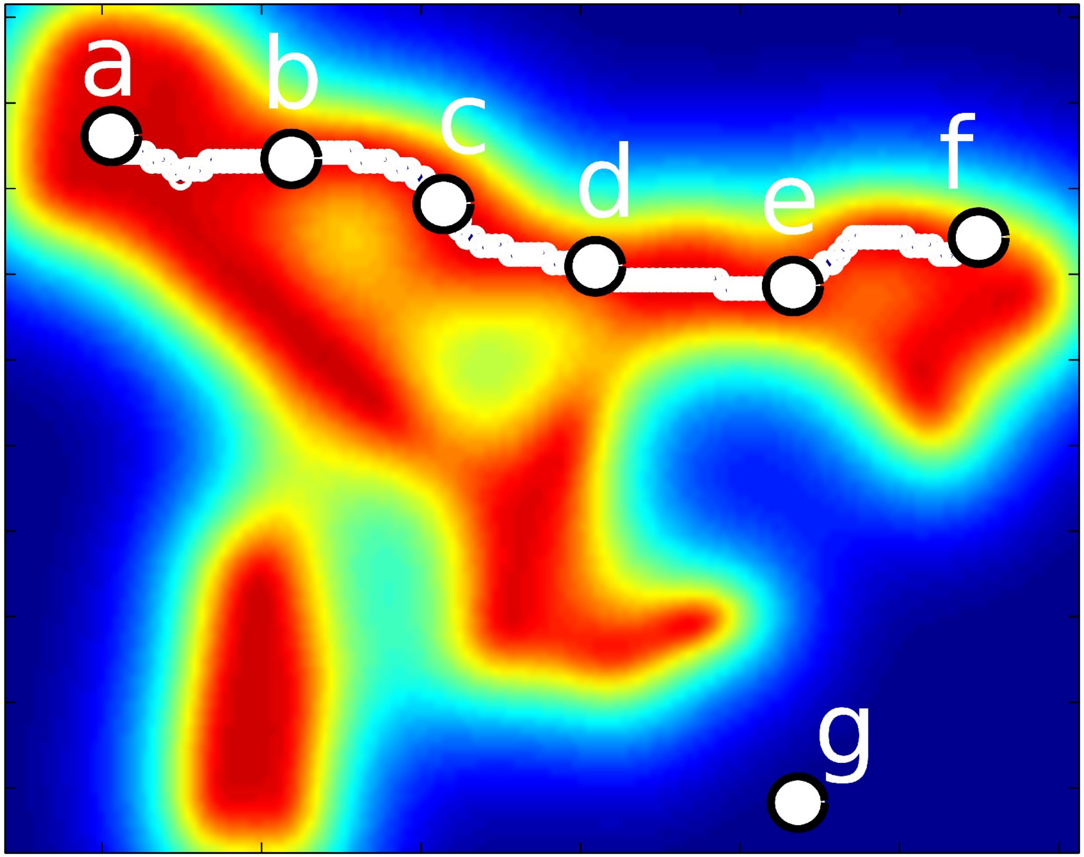
3) Interpolate the prior intermediate shapes in this low-dimensional latent space and project them back onto the (high-dimensionality) shape space

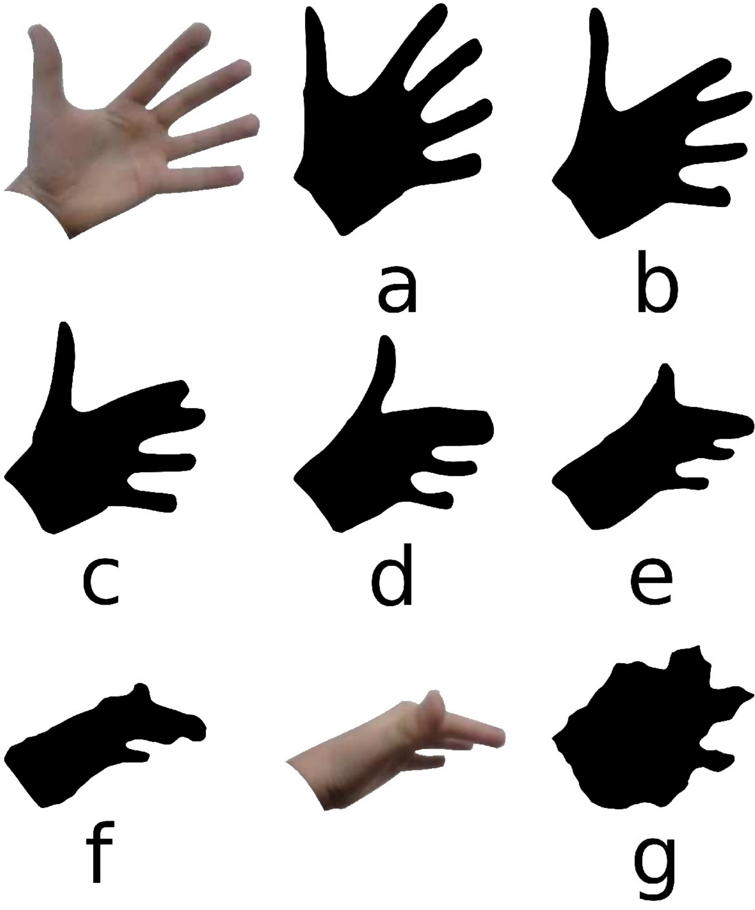
Learn a low-dimensional space representative of the shape of the analyzed object

- Description of a given shape of an object based on Elliptic Fourier shape Descriptors (EFD) [16]
- Learn the latent space of the object by non-linear dimensionality reduction (by GPLVM [17]) of the EFD of multiple object's occurences

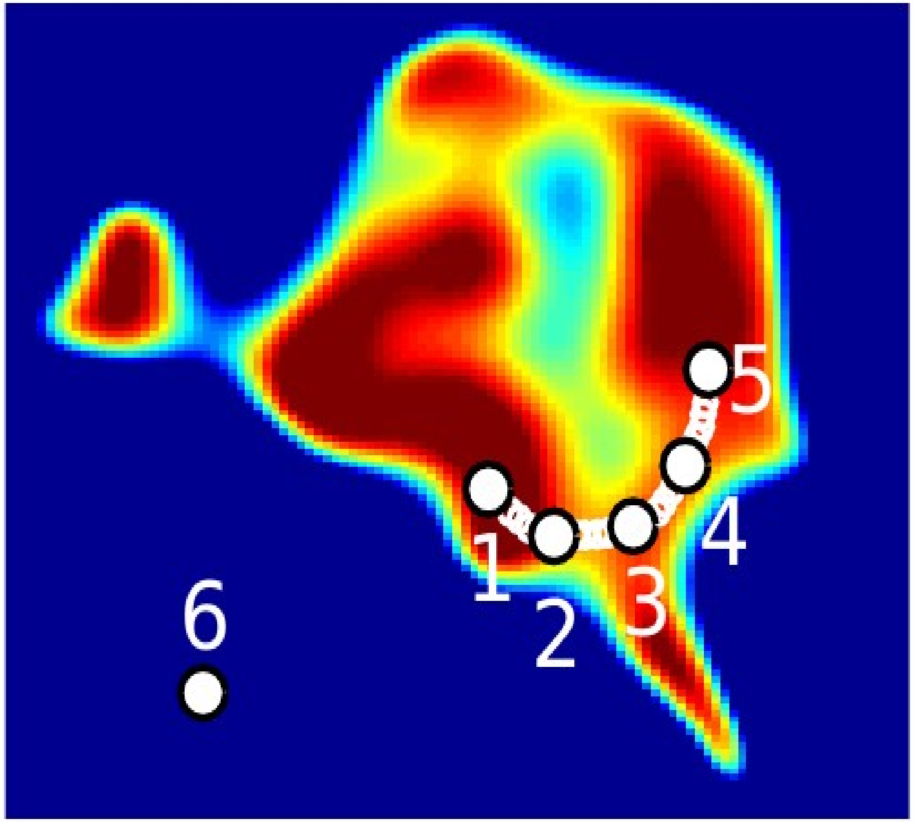


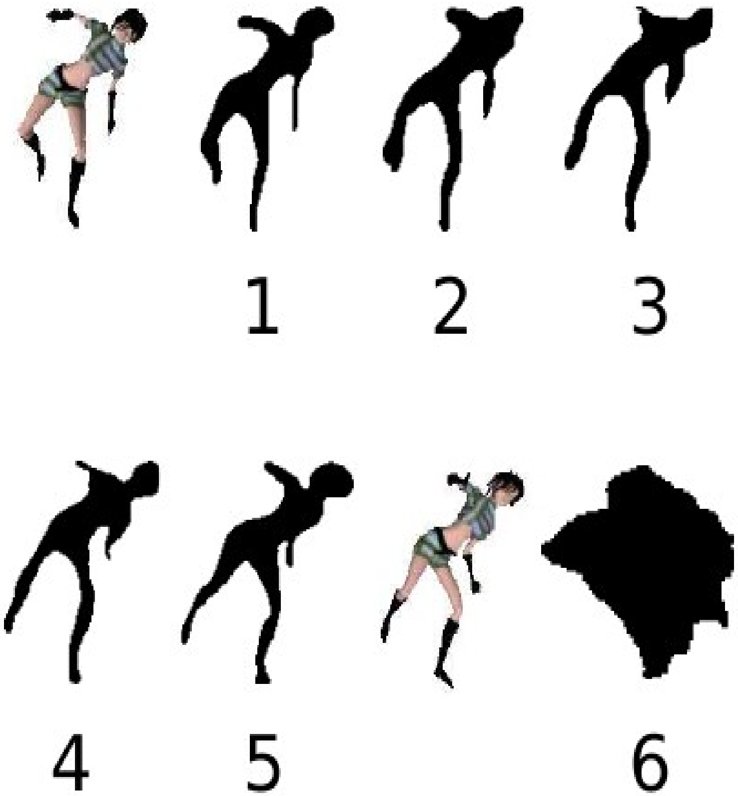
Interpolate [18] the prior intermediate shapes in this low-dimensional latent space and project them back





Interpolate [18] the prior intermediate (virtual) shapes in this low-dimensional latent space and project them back





Organization of the presentation

- Context
- Current limitations in model/image-based rendering
- Contributions in wide-baseline rendering
  - Match epipolar segments based on virtual prior
  - Determine prior about the virtual shapes
- Experimental validation

Result based on Zitnick et al., SIGGRAPH'04 [11]

Dynamic programming [14] on epipolar foreground segments

Proposed method (50 harmonics for EFD and 6 intermediate priors to regularize the matching)

Proposed method (50 harmonics for EFD and 6 intermediate priors to regularize the matching)

$\alpha = 0$ $\alpha = 0.25$ $\alpha = 0.5$ $\alpha = 0.75$ $\alpha = 1$

Conclusion

- Free viewpoint rendering enables to virtually navigate across a scene
- The foreshortening effect, as well as the occlusions, make the state-of-the-art methods fail in wide-baseline configurations
- We propose to match foreground segments between two very different viewpoints to tolerate the foreshortening effect
- We propose to add prior information on the intermediate views to disambiguate the ill-posed matching
- This allows to determine realistic vanishing and appearing trajectories of occluded parts
- Perspective : regularize based on color segments instead of foreground segments

Some references

(1) A. Laurentini, “The visual hull concept for silhouette-based image understanding”, in IEEE PAMI, 16(2):150-162, 1994.

(2) W. Matusik, C. Buehler, R. Raskar, S. J. Gortler, and L. McMillan, “Image-based visual hulls”, in Proc. of the ACM GRAPHITE, pages 369–374, 2000.

(3) J. Carranza, C. Theobalt, M. A. Magnor, and H.-P. Seidel, “Free-viewpoint video of human actors”, in ACM PAMI,22(3):569–577, 2003.

(4) M. Germann, T. Popa, R. Keiser, R. Ziegler, and M. Gross, “Novel-view synthesis of outdoor sport events using an adaptive view-dependent geometry”, in Computer Graphics Forum, volume 31, pages 325–333. Wiley Library, 2012.

(5) S. M. Seitz, “Image-based transformation of viewpoint and scene appearance”, in PhD thesis, Citeseer, 1997.

(6) H.-Y. Shum, S.-C. Chan, and S. B. Kang, “Image-based rendering”, in Springer Science+ Business Media, 2007.

(7) N. Inamoto and H. Saito, “Virtual viewpoint replay for a soccer match by view interpolation from multiple cameras”, in IEEE TM, 9(6):1155–1166, 2007.

(8) S. M. Seitz and C. R. Dyer, “View morphing”, in ACM GRAPHITE, pages 21–30, 1996.

(9) S. Mattoccia, “Accurate dense stereo by constraining local consistency on superpixels”, in IEEE ICPR, pages 1832-1835, 2010.

Some references

(10) A. Klaus, M. Sormann, and K. Karner, “Segment-based stereo matching using belief propagation and a self-adaptif dissimilarity measure”, in ICPR, volume 3, pages 15–18, 2006.

(11) C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, “High-quality video view interpolation using a layered representation”, in PAMI, volume 23, pages 600–608, 2004.

(12) R. Hartley and A. Zisserman, “Multiple view geometry in computer vision”, in Cambridge university press, 2003.

(13) C. Verleysen and C. De Vleeschouwer, “Learning and Propagation of Dominant Colors for Fast Video Segmentation”, in ACIVS, pages 657-668, Springer, 2013.

(14) I. J. Cox, S. L. Hingorani, S. B. Rao, and B. M. Maggs, “A maximum likelihood stereo algorithm”, in CVIU, 63(3):542-567, 1996.

(15) R. Toldo and A. Fusiello, “Robust multiple structures estimation with j-linkage”, in ECCV, pages 537-547, Springer, 2008.

(16) F. P. Kuhl and C. R. Giardina, “Elliptic fourier features of a closed contour”, in CGIP, 18(3):236–258, 1982.

(17) N. D. Lawrence and J. Quinonero-Candela, “Local distance preservation in the GPLVM through back constraints”, in ACM NIPS, pages 513–520, 2006.

(18) E. W. Dijkstra, “A note on two problems in connexion with graphs”, in Numerische mathematik, 1(1):269–271, 1959.

Thank you very much for your attention...

Questions?