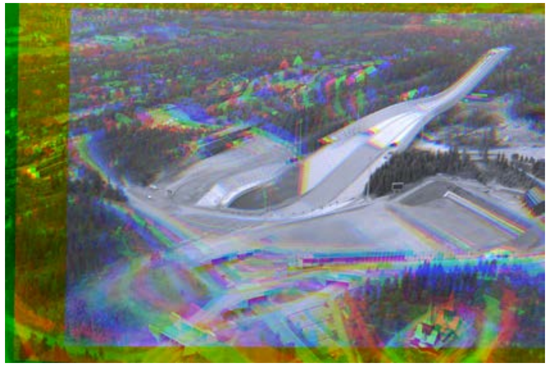
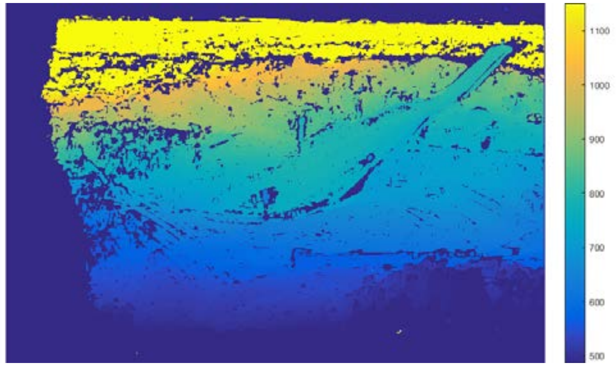
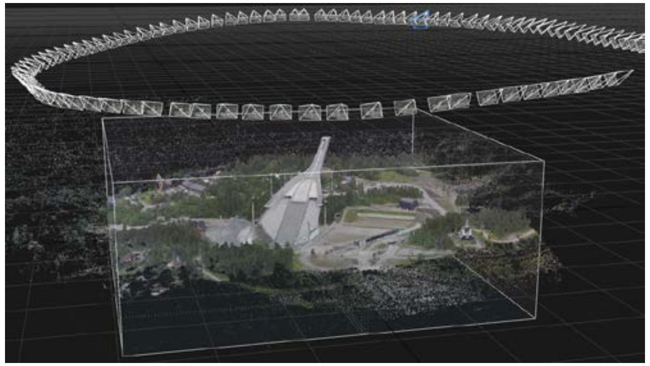
  

Lecture 8.3
Multiple-view stereo

Trym Vegard Haavardsholm

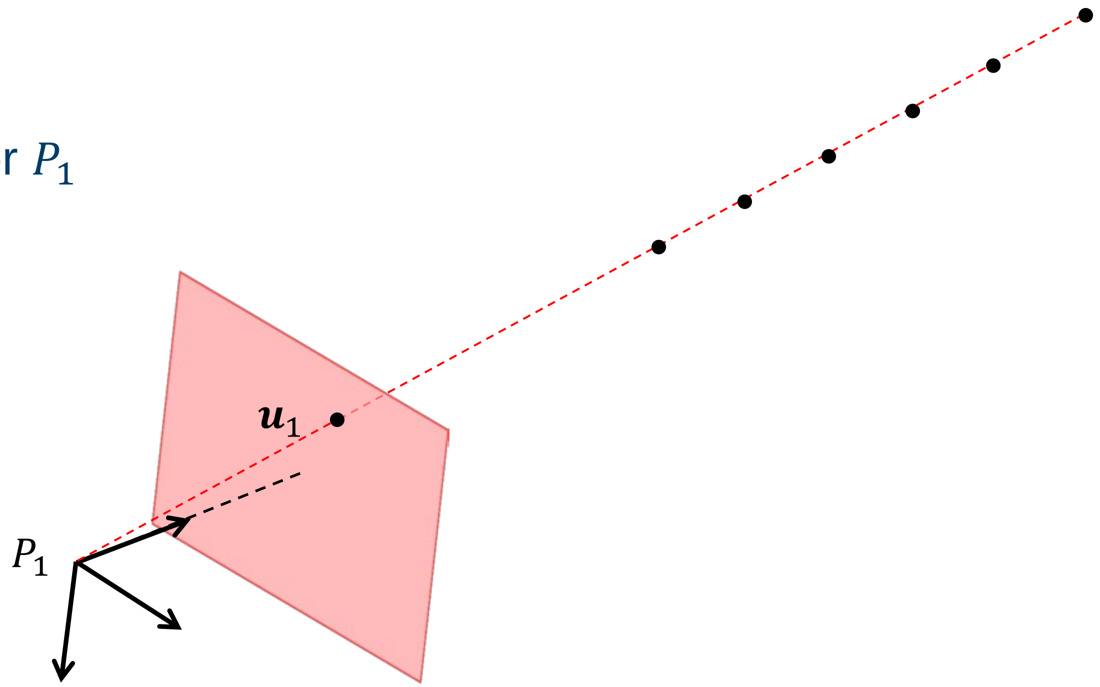
Several slides from Svetlana Lazebnik,
Steve Seitz, Noah Snavely and Rob Fergus

World geometry from correspondences

	Structure (scene geometry)	Motion (camera geometry)	Measurements
Pose estimation	Known	Estimate	3D to 2D correspondences
Triangulation, Stereo	Estimate	Known	2D to 2D correspondences
Reconstruction, Structure from Motion	Estimate	Estimate	2D to 2D correspondences

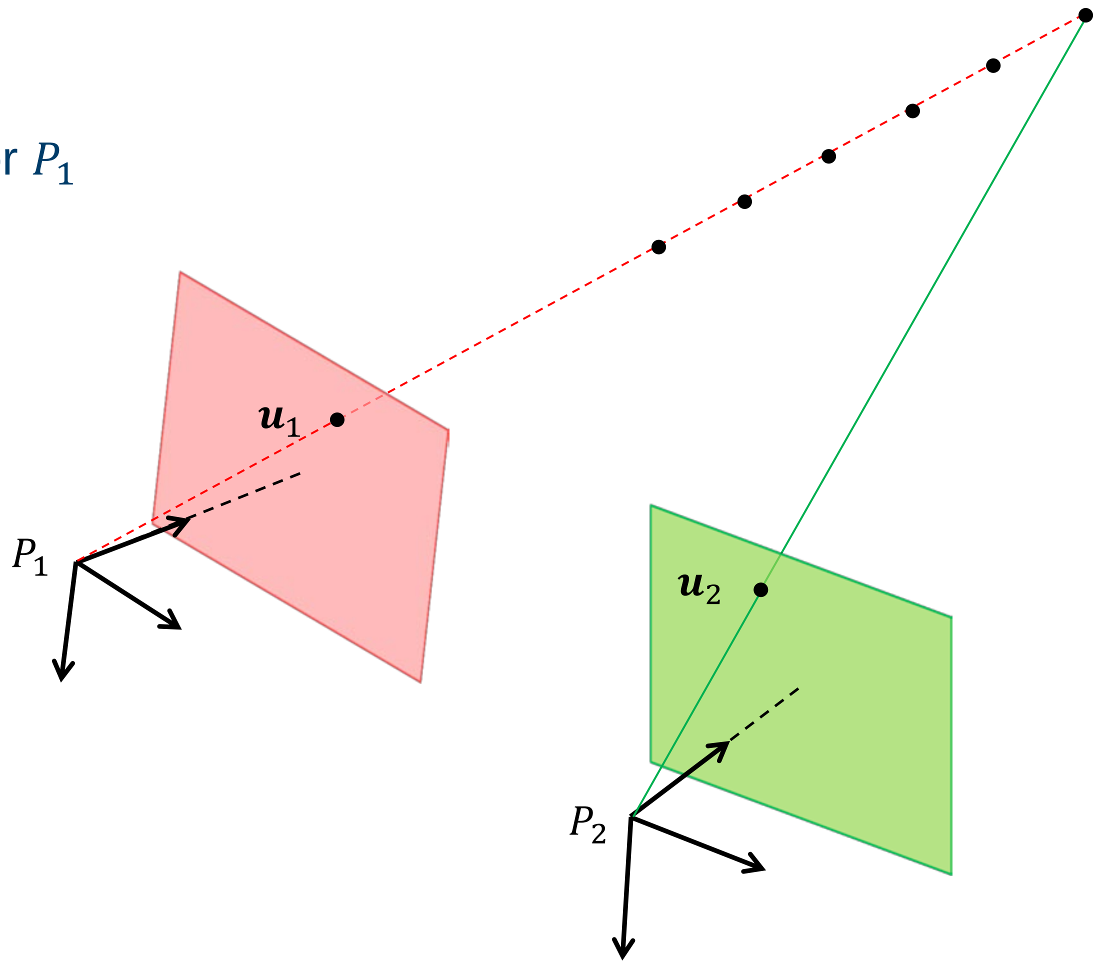
Depth from images

- Any point on the ray has image $\boldsymbol{u}_1$ for P_1



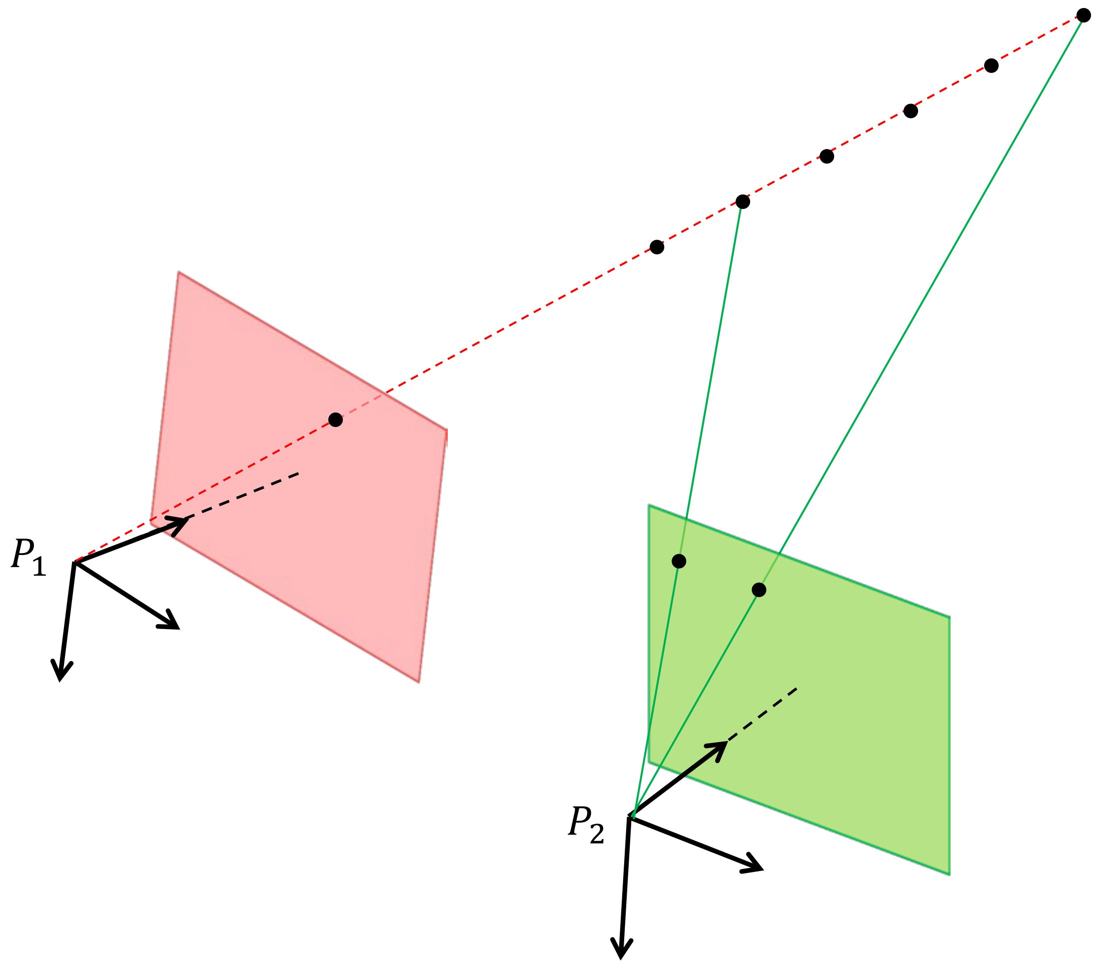
Depth from images

- Any point on the ray has image $\boldsymbol{u}_1$ for P_1
  - Add a second view!



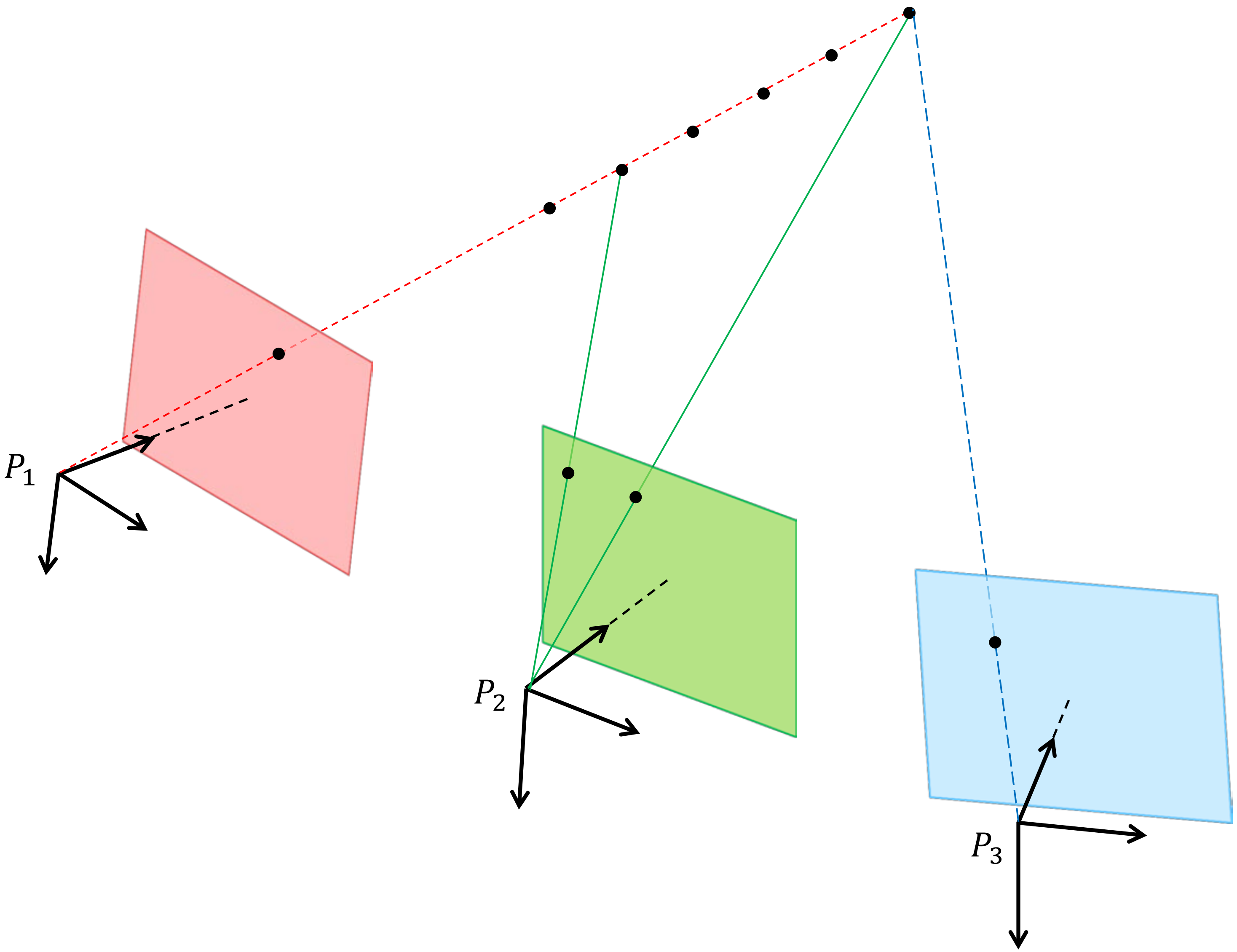
Depth from images

- But it is not always possible to find correct correspondences



Multiple-view stereo

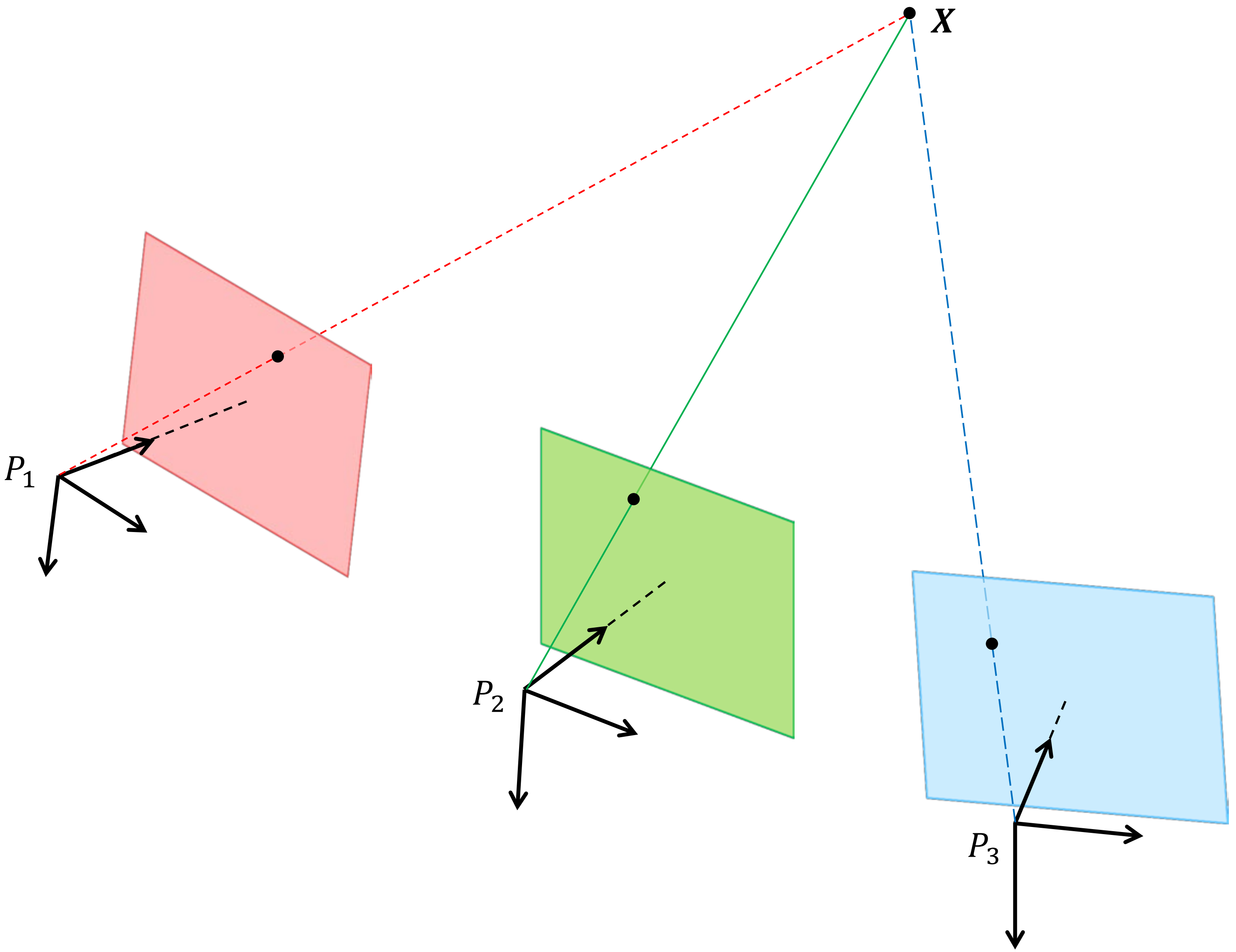
- But it is not always possible to find correct correspondences
  - Add more views!



Multiple-view stereo

- But it is not always possible to find correct correspondences
  - Add more views!

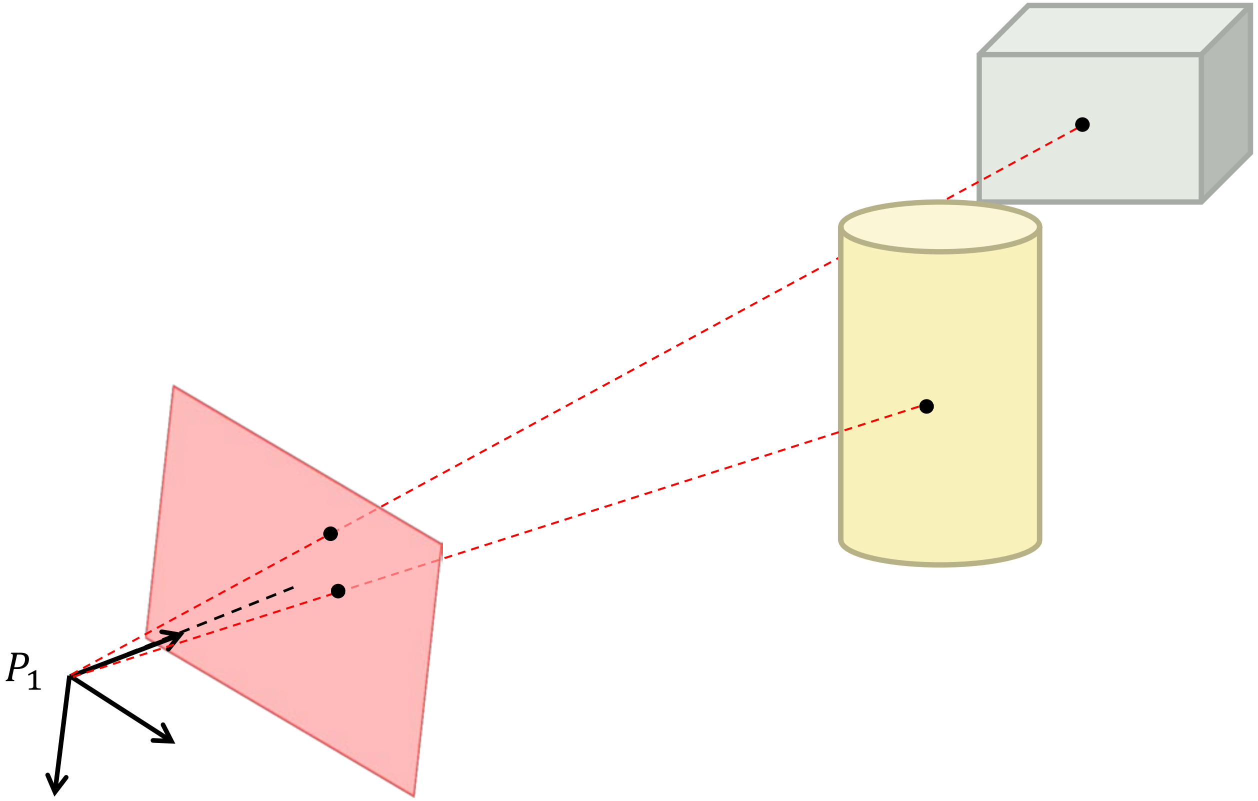
- Multiple views:
  - can be used to verify correspondences



Multiple-view stereo

- But it is not always possible to find correct correspondences
  - Add more views!

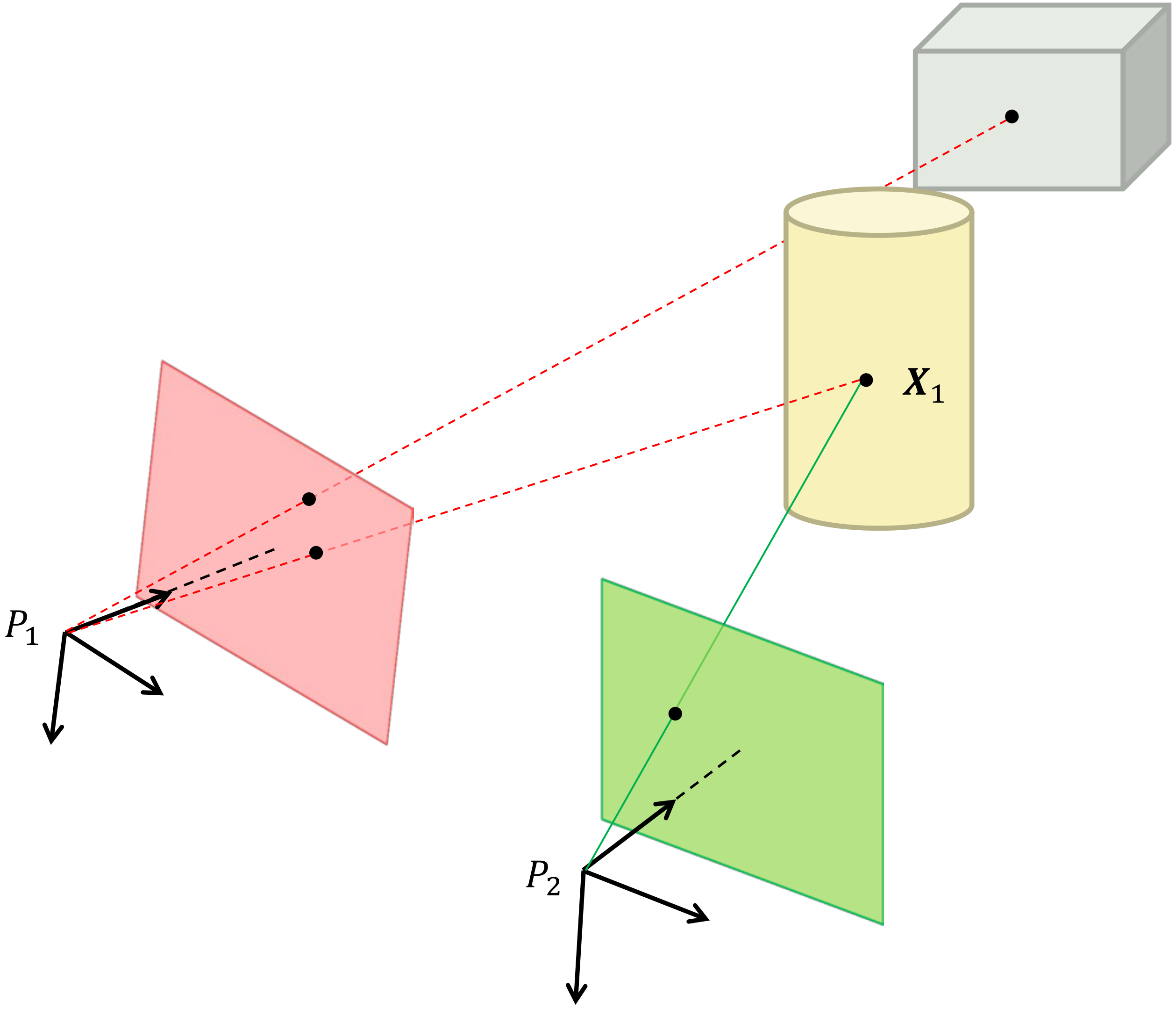
- Multiple views:
  - can be used to verify correspondences



Multiple-view stereo

- But it is not always possible to find correct correspondences
  - Add more views!

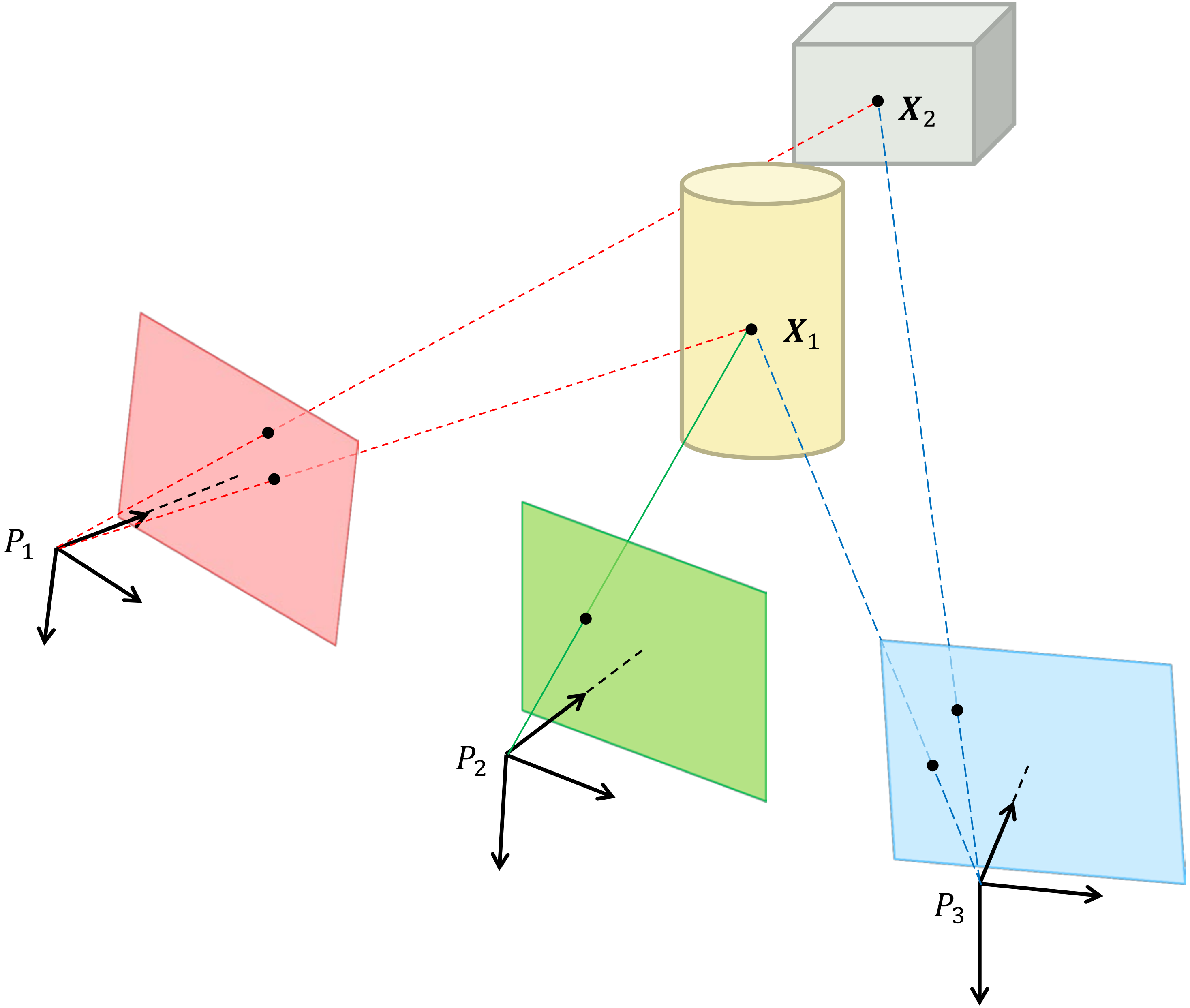
- Multiple views:
  - can be used to verify correspondences



Multiple-view stereo

- But it is not always possible to find correct correspondences
  - Add more views!

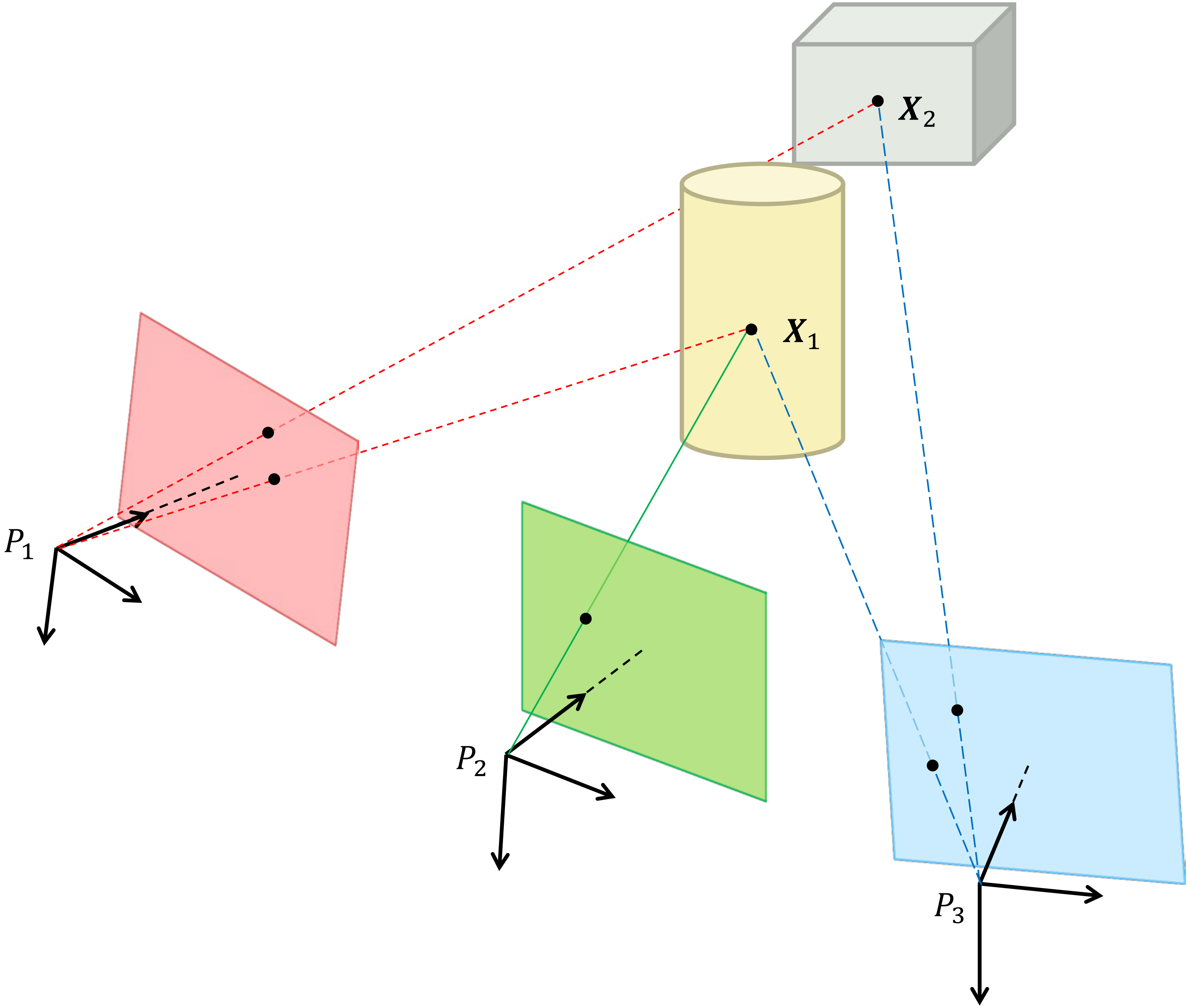
- Multiple views:
  - can be used to verify correspondences
  - can make reconstruction more robust to occlusions



Multiple-view stereo

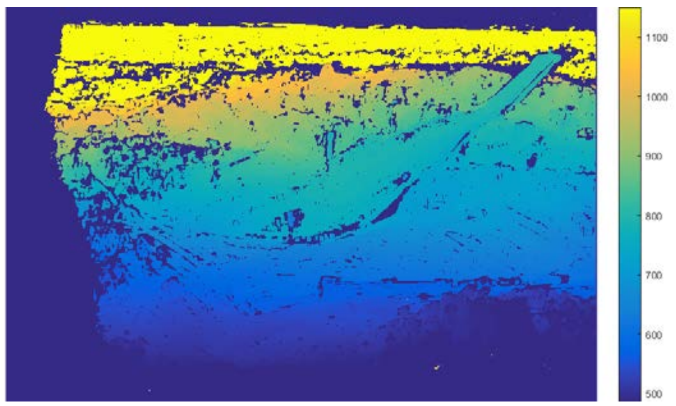
- But it is not always possible to find correct correspondences
  - Add more views!

- Multiple views:
  - can be used to verify correspondences
  - can make reconstruction more robust to occlusions
  - can be used to infer free-space and solid volumes

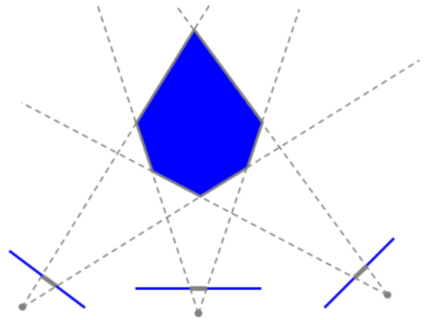


Approaches to multiple-view stereo

- Multiple-view stereo depth maps
  - Fusion of point clouds



- Volumetric stereo
  - Volume carving on discrete grids

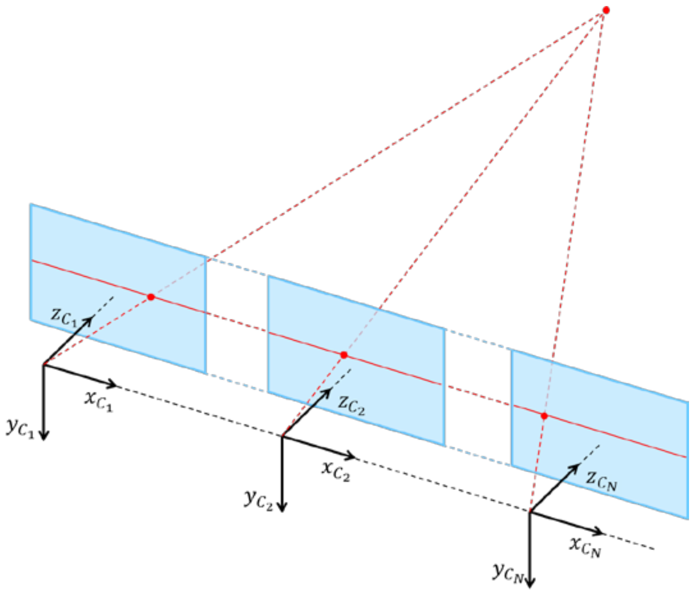


- Surface Expansion
  - Dense feature matching

Depth maps

- Multiple-baseline stereo
  - Rectification of several cameras onto common plane
  - Problems with wide baselines and distortions after rectification

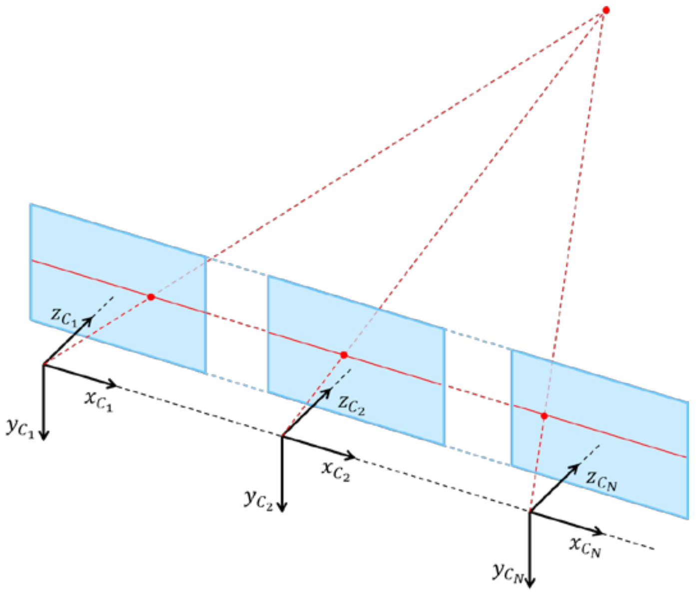
www.ptgrey.com



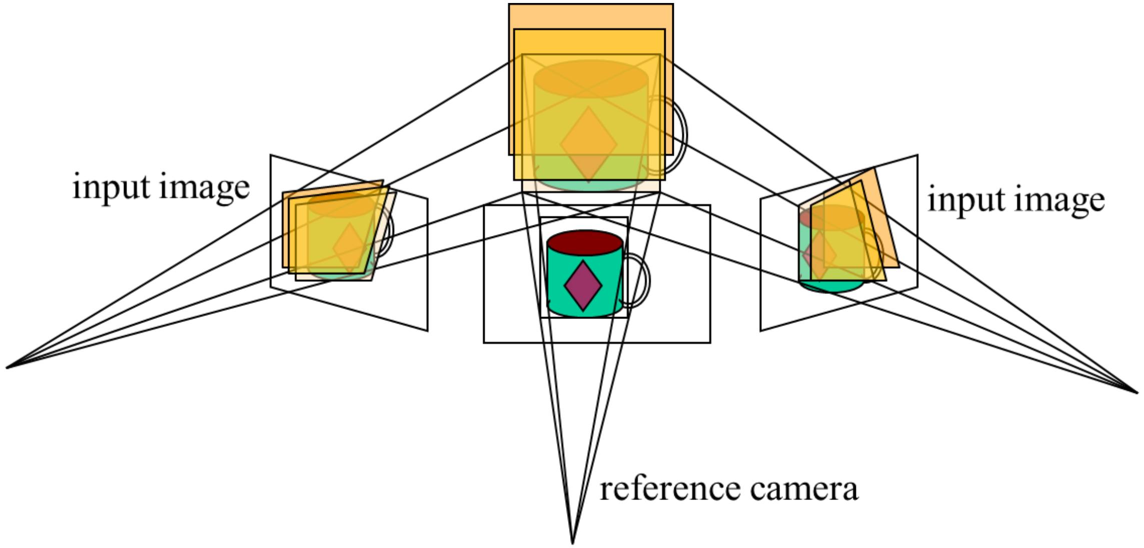
Depth maps

- Multiple-baseline stereo
  - Rectification of several cameras onto common plane
  - Problems with wide baselines and distortions after rectification

www.ptgrey.com



- Plane sweep stereo
  - Choose a reference view
  - Sweep family of planes at different depths with respect to the reference camera



Plane sweep

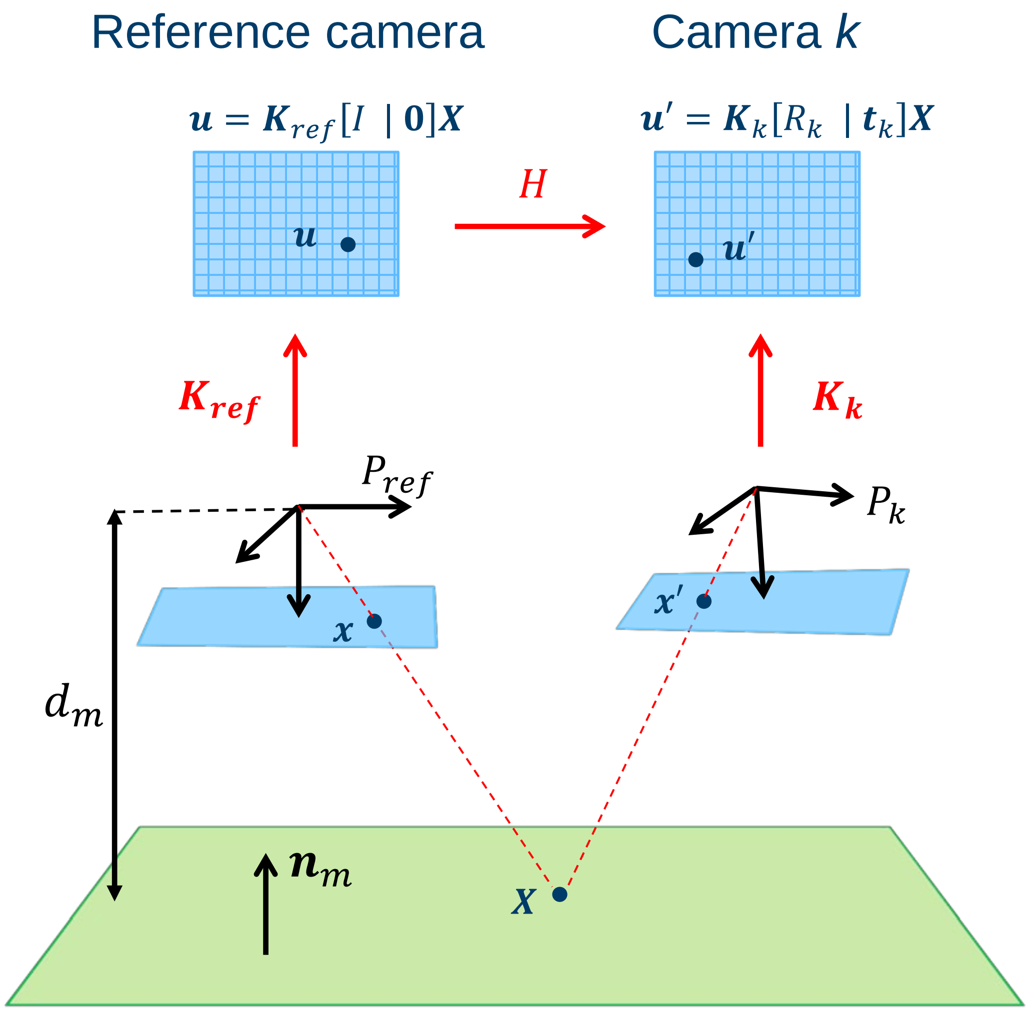
- The family of depth planes
  in the coordinate frame of the reference view

$$\Pi_m = \begin{bmatrix} \boldsymbol{n}_m^T & -d_m \end{bmatrix}$$

- The mapping from the reference camera P_{ref} onto the plane Π_m and back to camera P_k is described by the homography induced by the plane Π_m

$$H_{\Pi_m, P_k} = K_k \left(R_k - \boldsymbol{t}_k \boldsymbol{n}_m^T / d_m \right) K_{ref}^{-1}$$

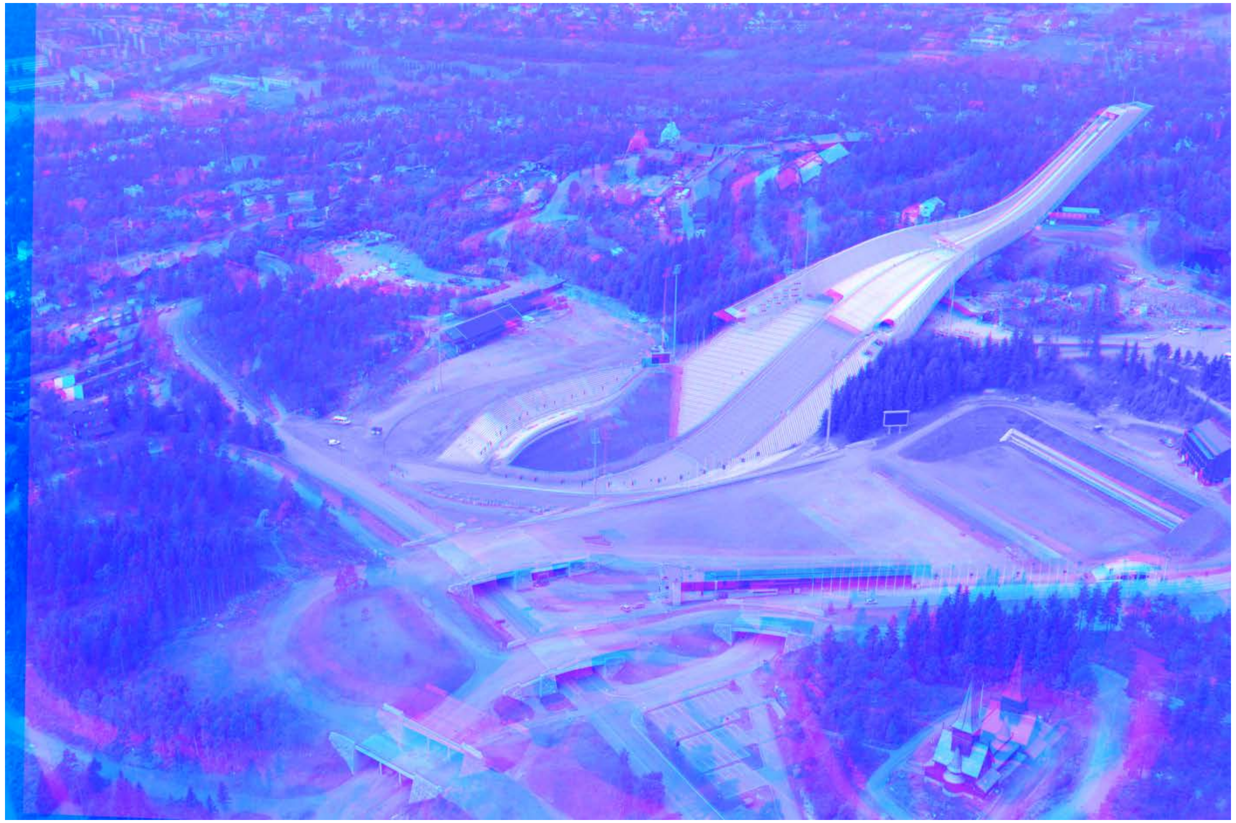
- The mapping from P_k to P_{ref} induced by Π_m is the inverse homography H_{Π_m, P_k}^{-1}

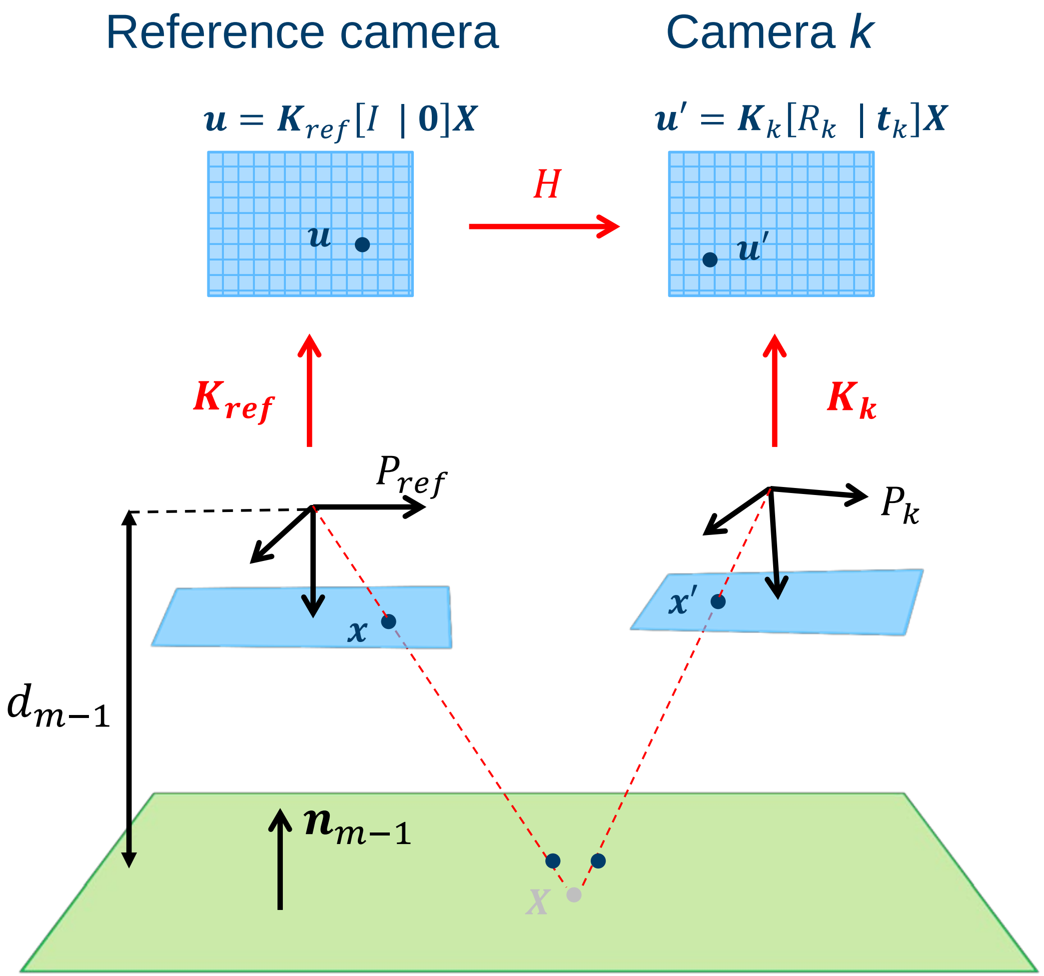


Robert Collins, A Space-Sweep Approach to True Multi-Image Matching, CVPR 1996.
D. Gallup, J.-M. Frahm, P. Mordohai, Q. Yang and M. Pollefeys, Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions, CVPR 2007

Plane sweep

- Sweep planes at different depths

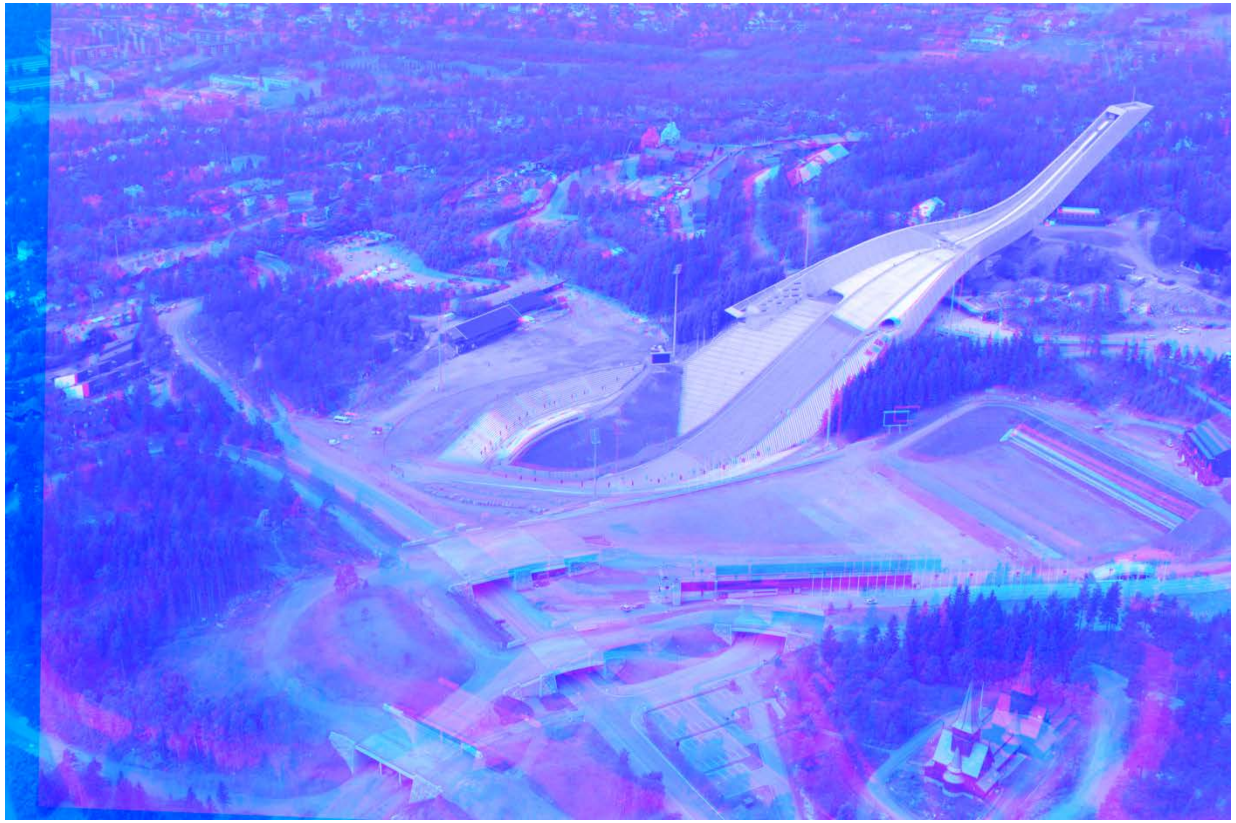


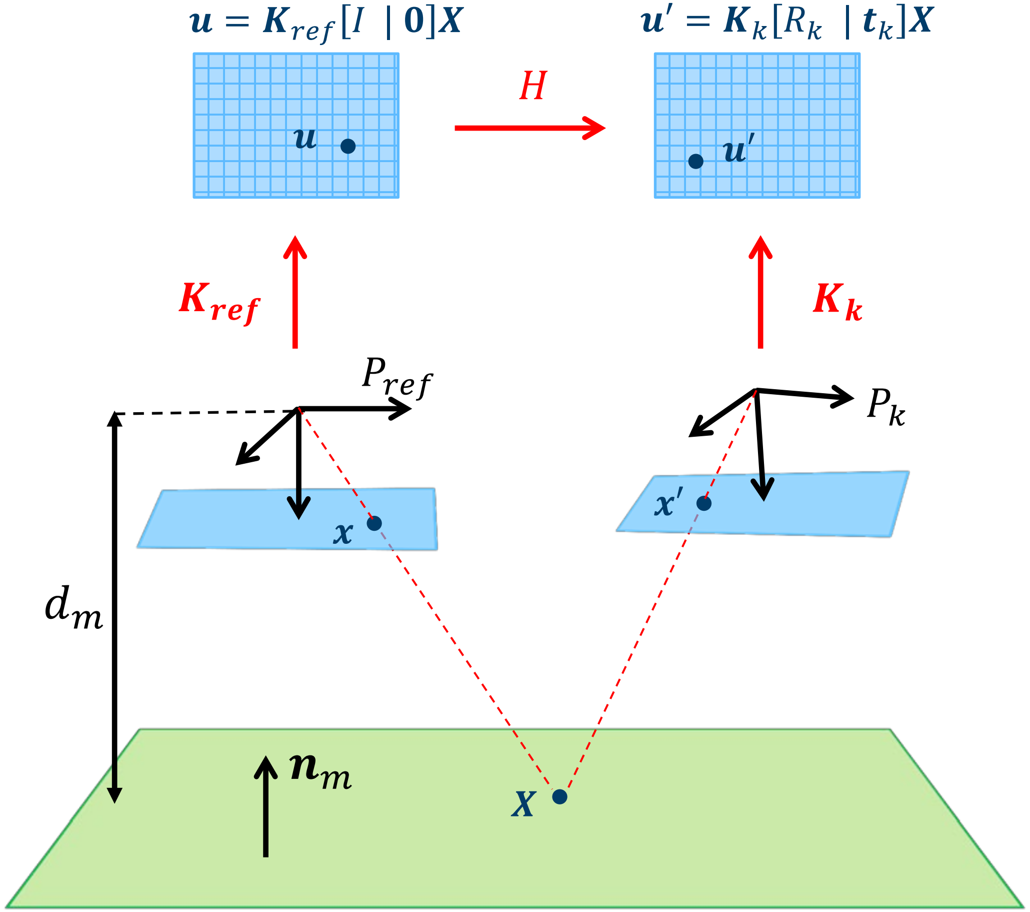


Robert Collins, A Space-Sweep Approach to True Multi-Image Matching, CVPR 1996.
D. Gallup, J.-M. Frahm, P. Mordohai, Q. Yang and M. Pollefeys, Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions, CVPR 2007

Plane sweep

- Sweep planes at different depths

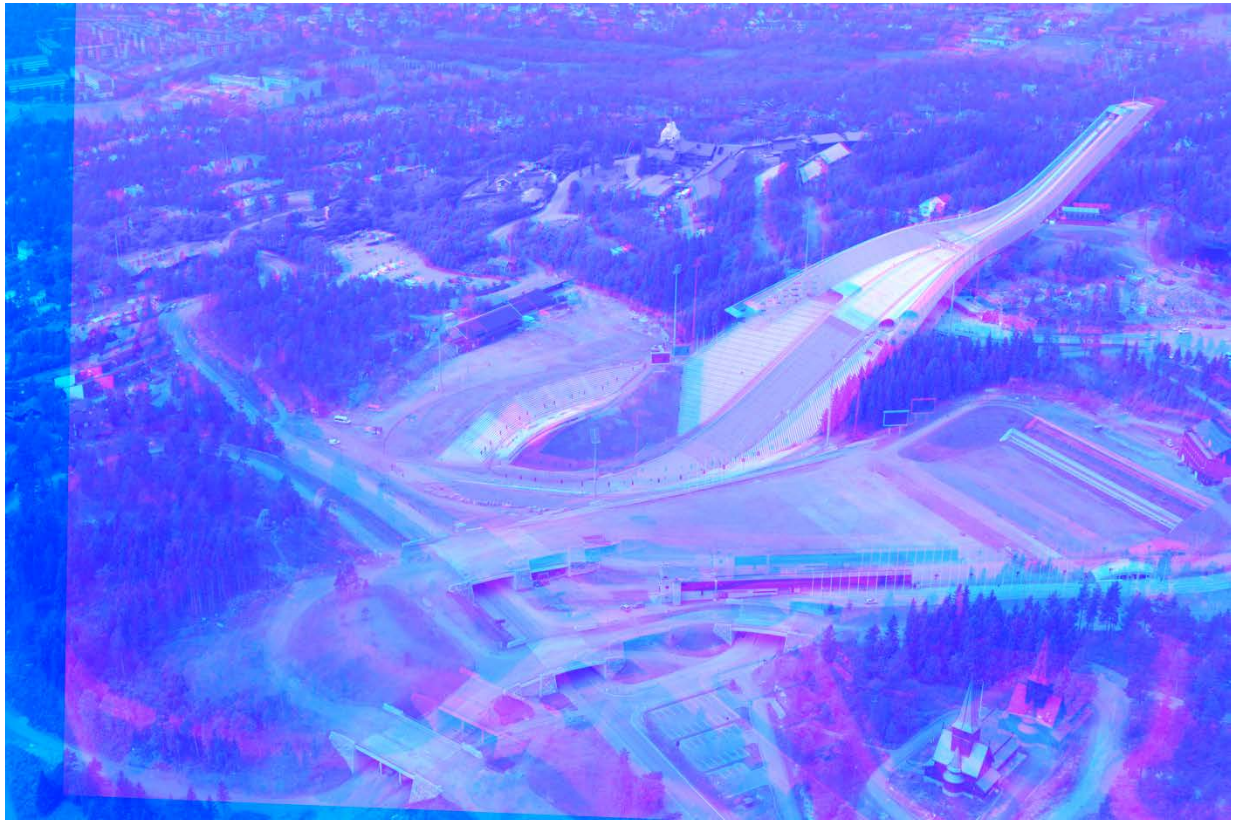


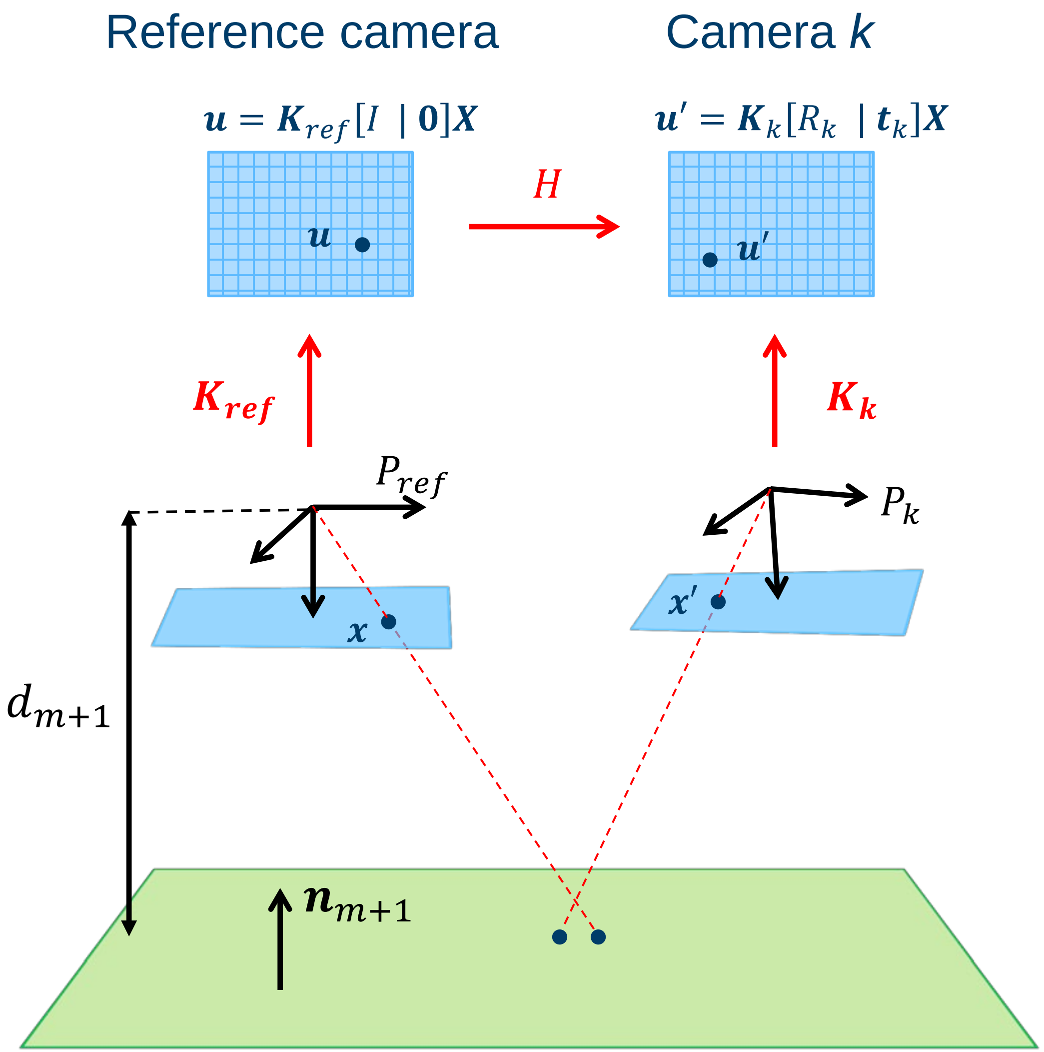


Robert Collins, A Space-Sweep Approach to True Multi-Image Matching, CVPR 1996.
D. Gallup, J.-M. Frahm, P. Mordohai, Q. Yang and M. Pollefeys, Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions, CVPR 2007

Plane sweep

- Sweep planes at different depths





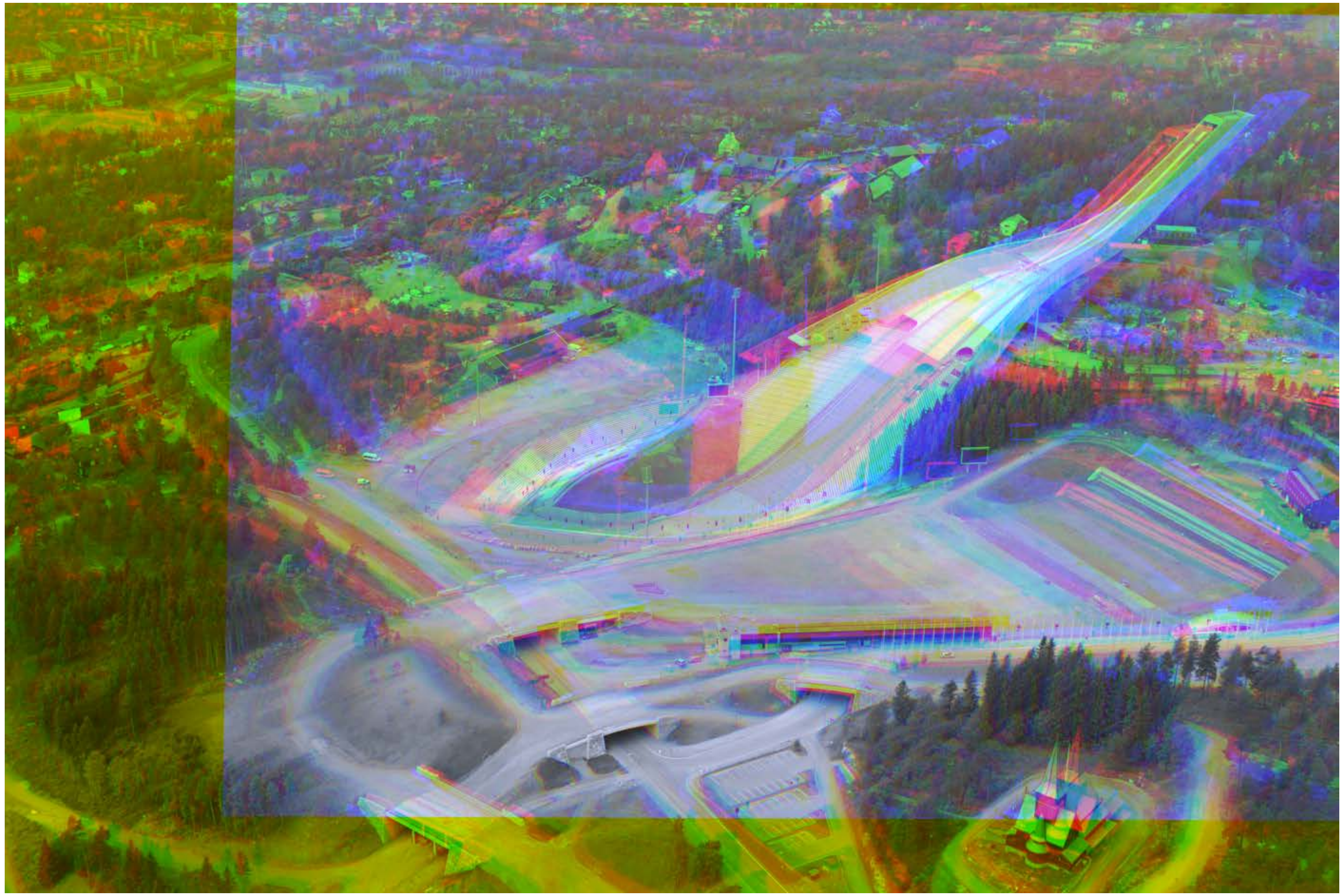
Robert Collins, A Space-Sweep Approach to True Multi-Image Matching, CVPR 1996.
D. Gallup, J.-M. Frahm, P. Mordohai, Q. Yang and M. Pollefeys, Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions, CVPR 2007

Plane sweep for multiple views

Red:

Green:

Blue:



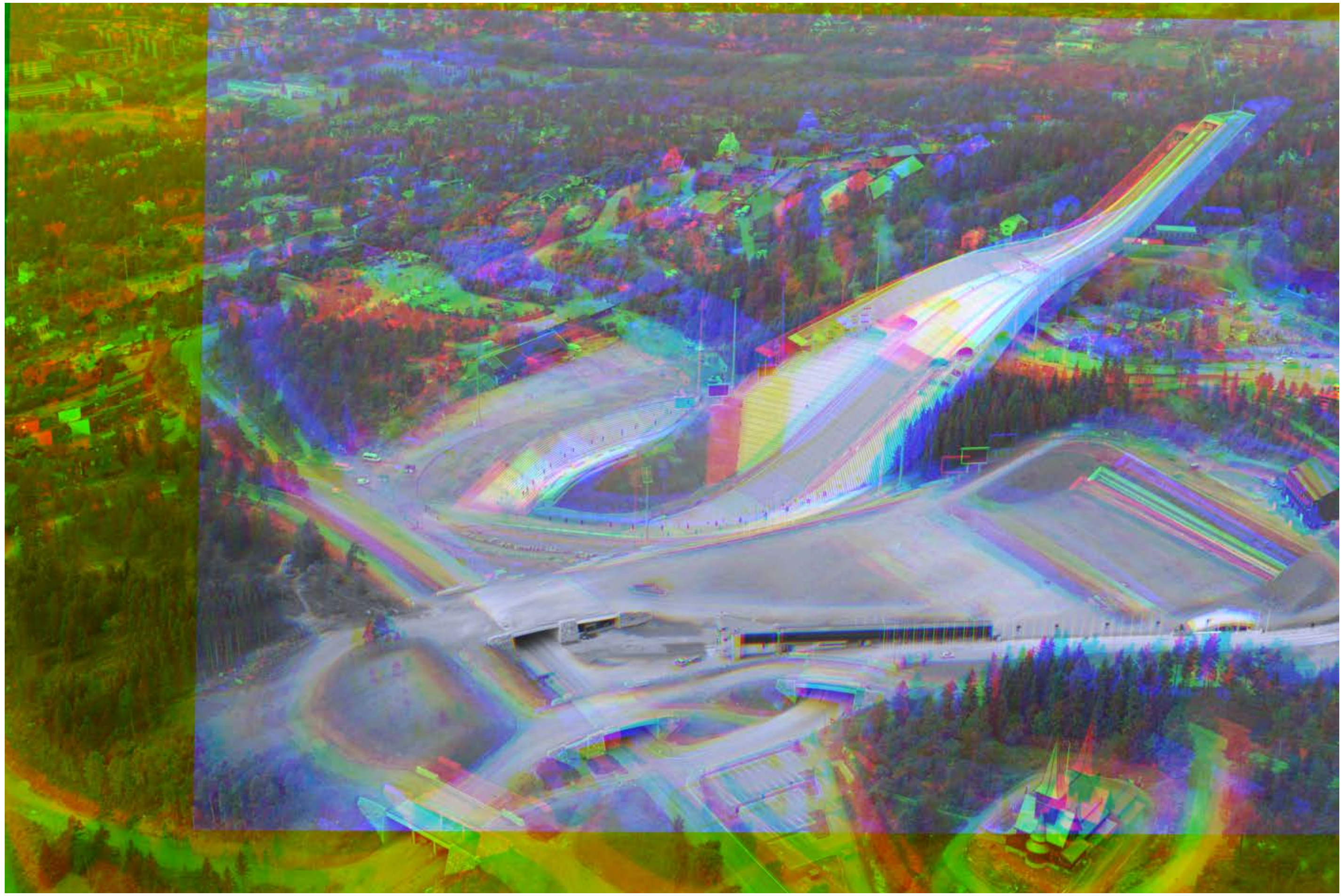
d_m = 520 meter

Plane sweep for multiple views

Red:

Green:

Blue:



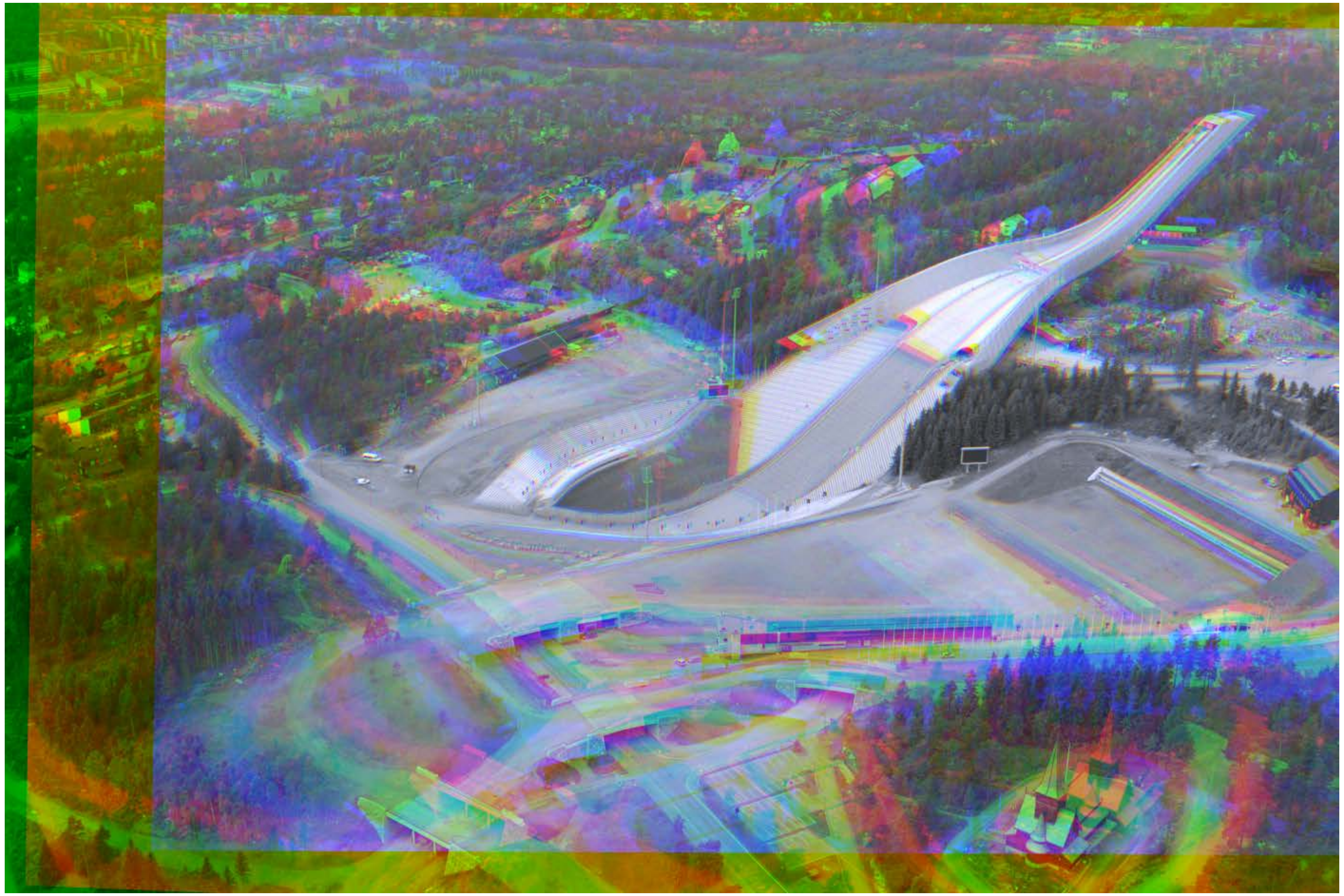
d_m = 583 meter

Plane sweep for multiple views

Red:

Green:

Blue:



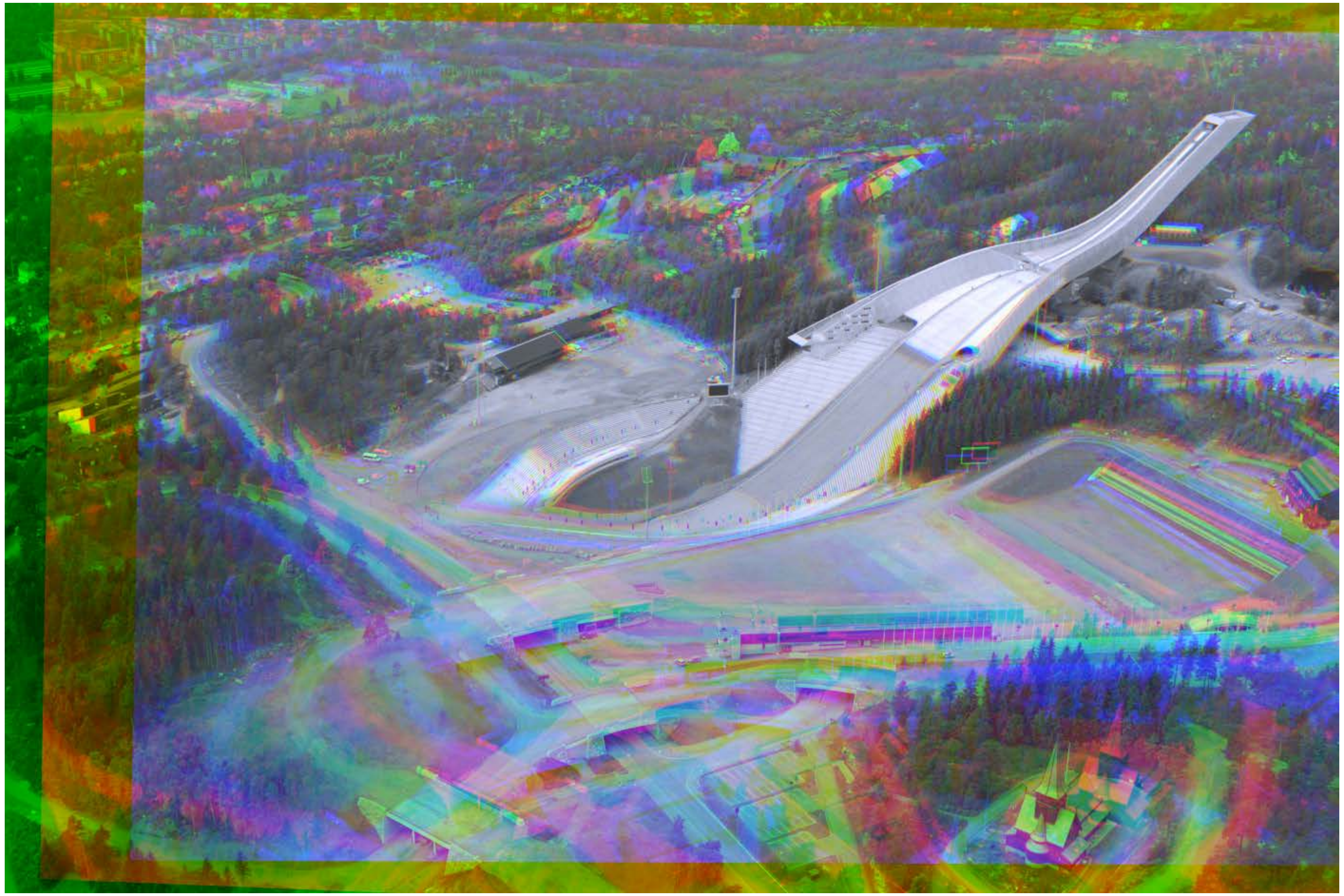
d_m = 706 meter

Plane sweep for multiple views

Red:

Green:

Blue:



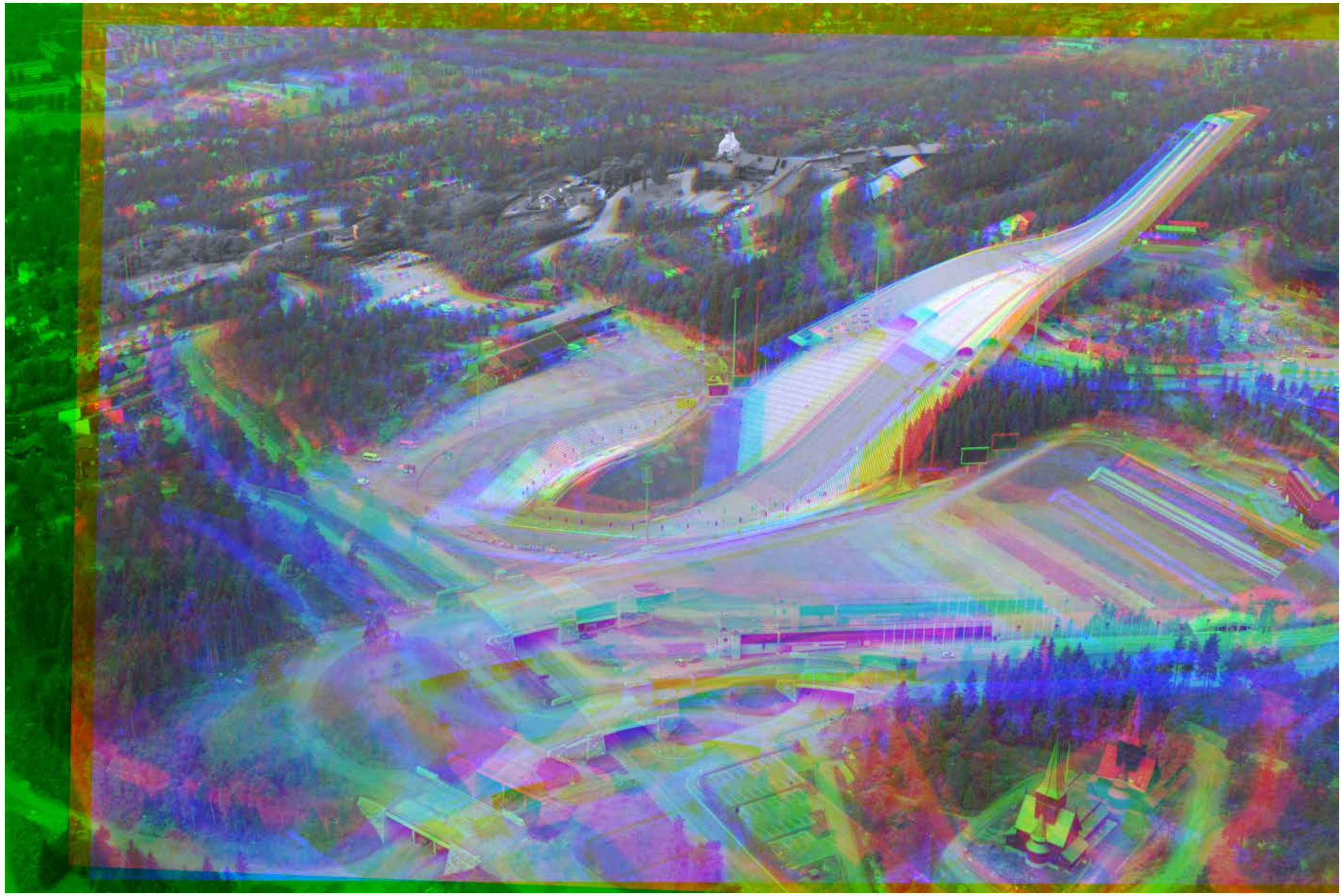
d_m = 790 meter

Plane sweep for multiple views

Red:

Green:

Blue:



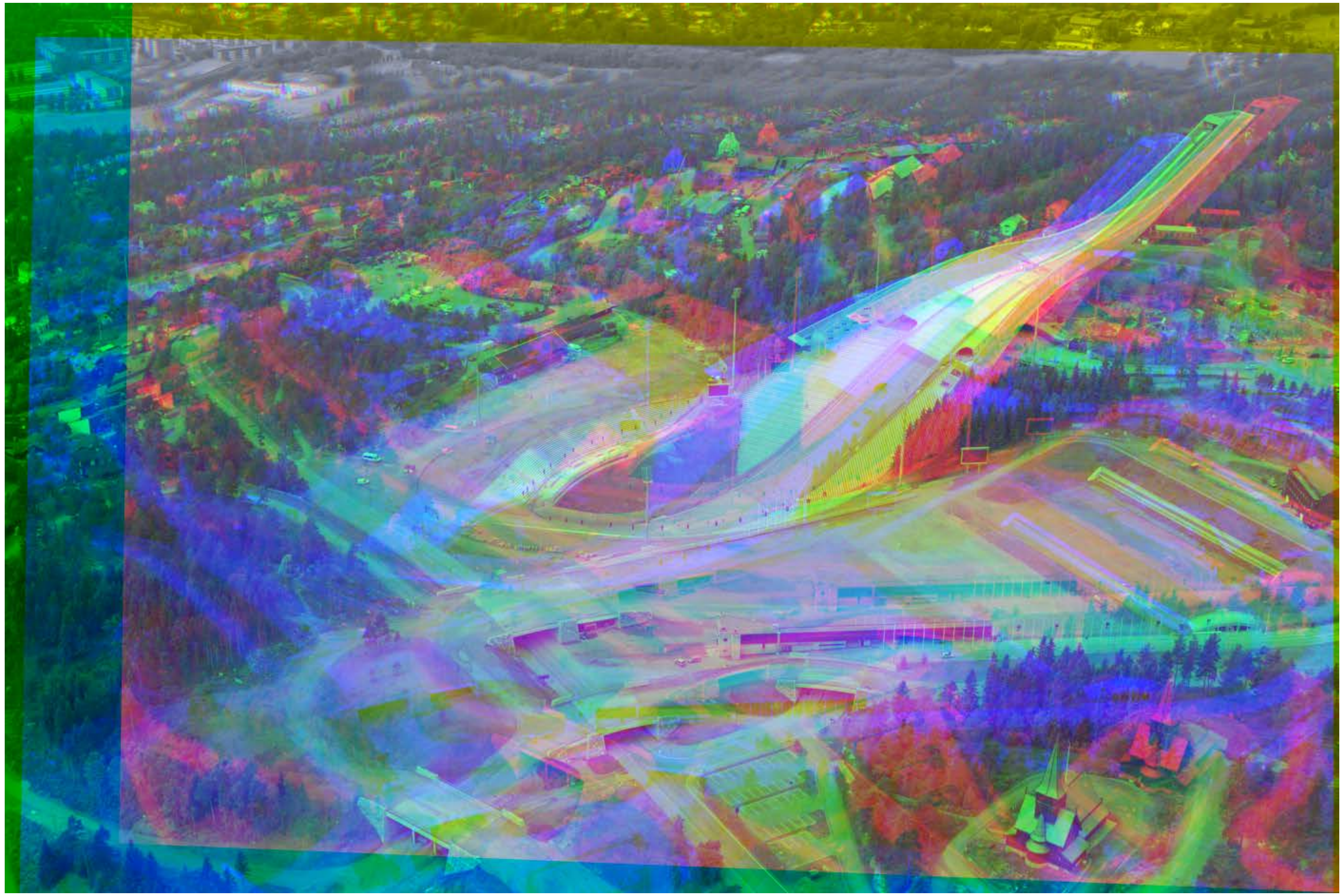
d_m = 1026 meter

Plane sweep for multiple views

Red:

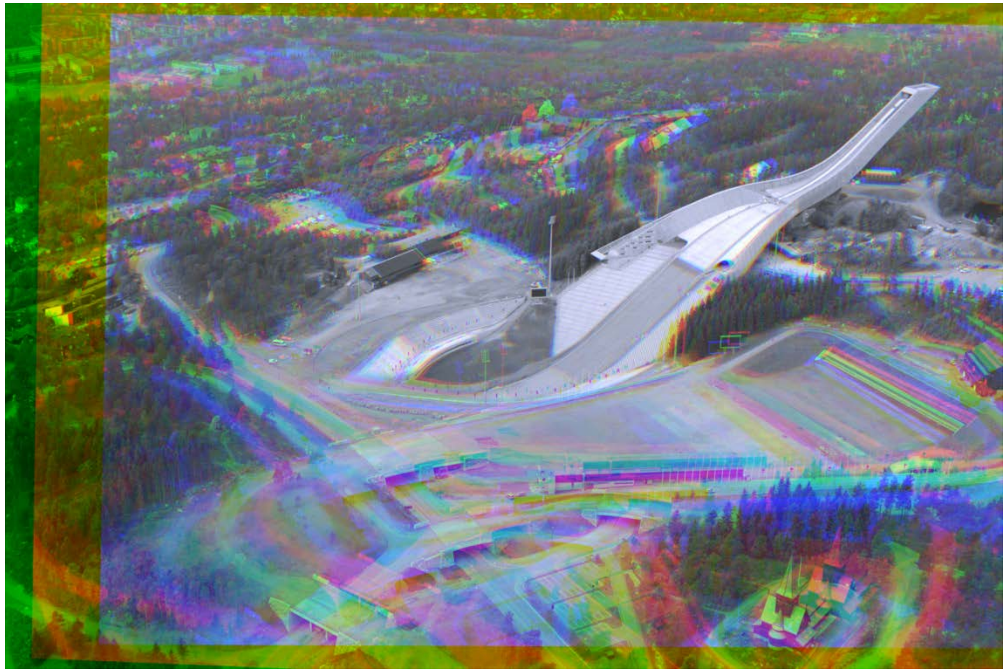
Green:

Blue:



d_m = 2168 meter

Plane sweep stereo



1. Map each target image I_k to the reference image I_{ref} for each depth plane Π_m with the homography $H^{-1}_{\Pi_m, P_k}$ giving the warped images $\breve{I}_{k,m}$

2. Compute the similarity between I_{ref} and each $\breve{I}_{k,m}$
   - Zero Mean Normalized Cross Correlation (ZNCC) between small patches *W* around each pixel

3. Compute the figure-of-merit for each depth plane by combining the similarity measurements for each image *k*

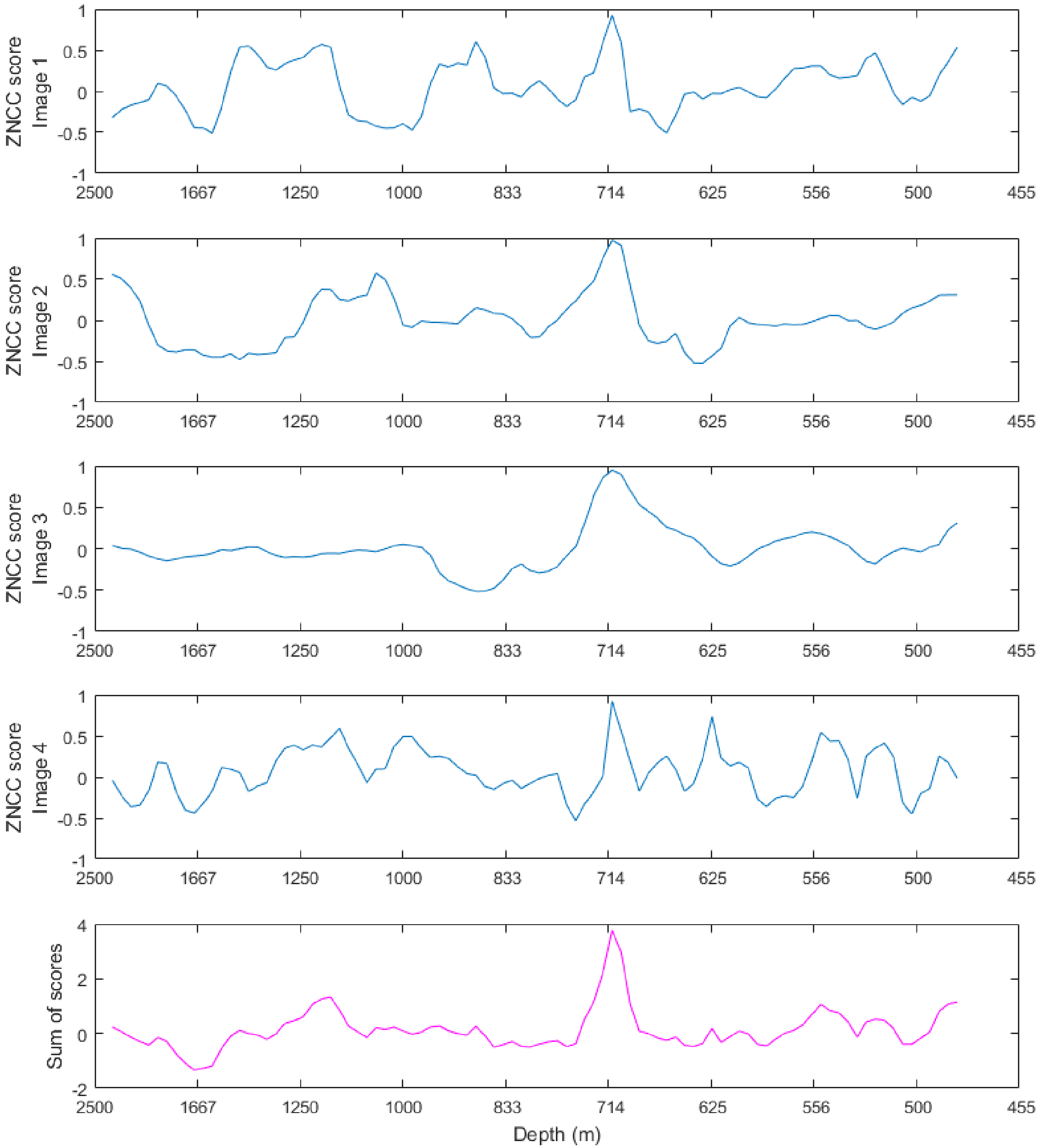
$$M(u,v,\Pi_m) = \sum_k \underset{W}{ZNCC}(I_{ref}, \breve{I}_{k,m})$$

4. For each pixel, select the depth plane with best fit

$$\tilde{\Pi}(u,v) = \arg\max_m M(u,v,\Pi_m)$$

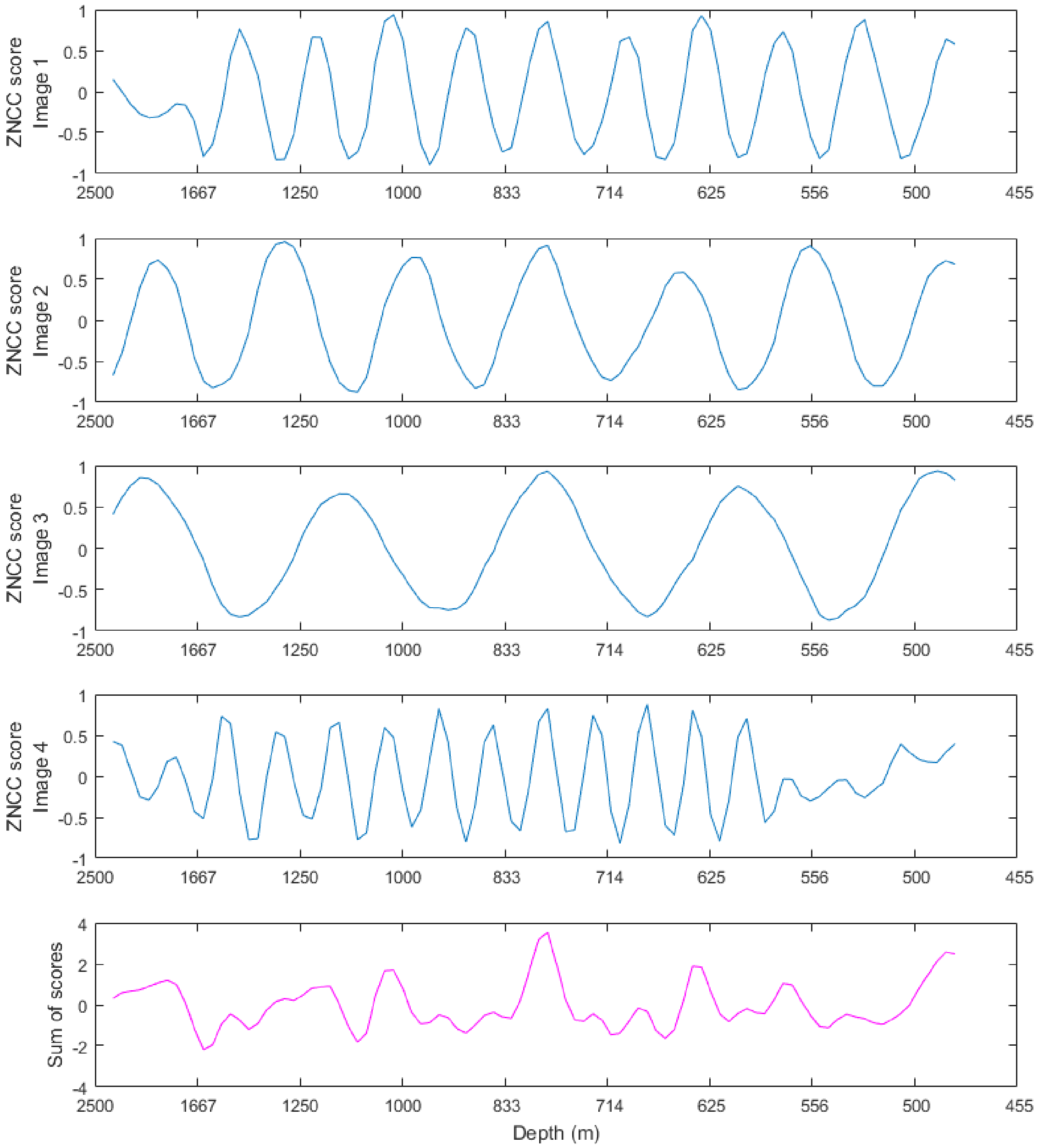
Plane sweep stereo example

- ZNCC scores for different depths and *k*



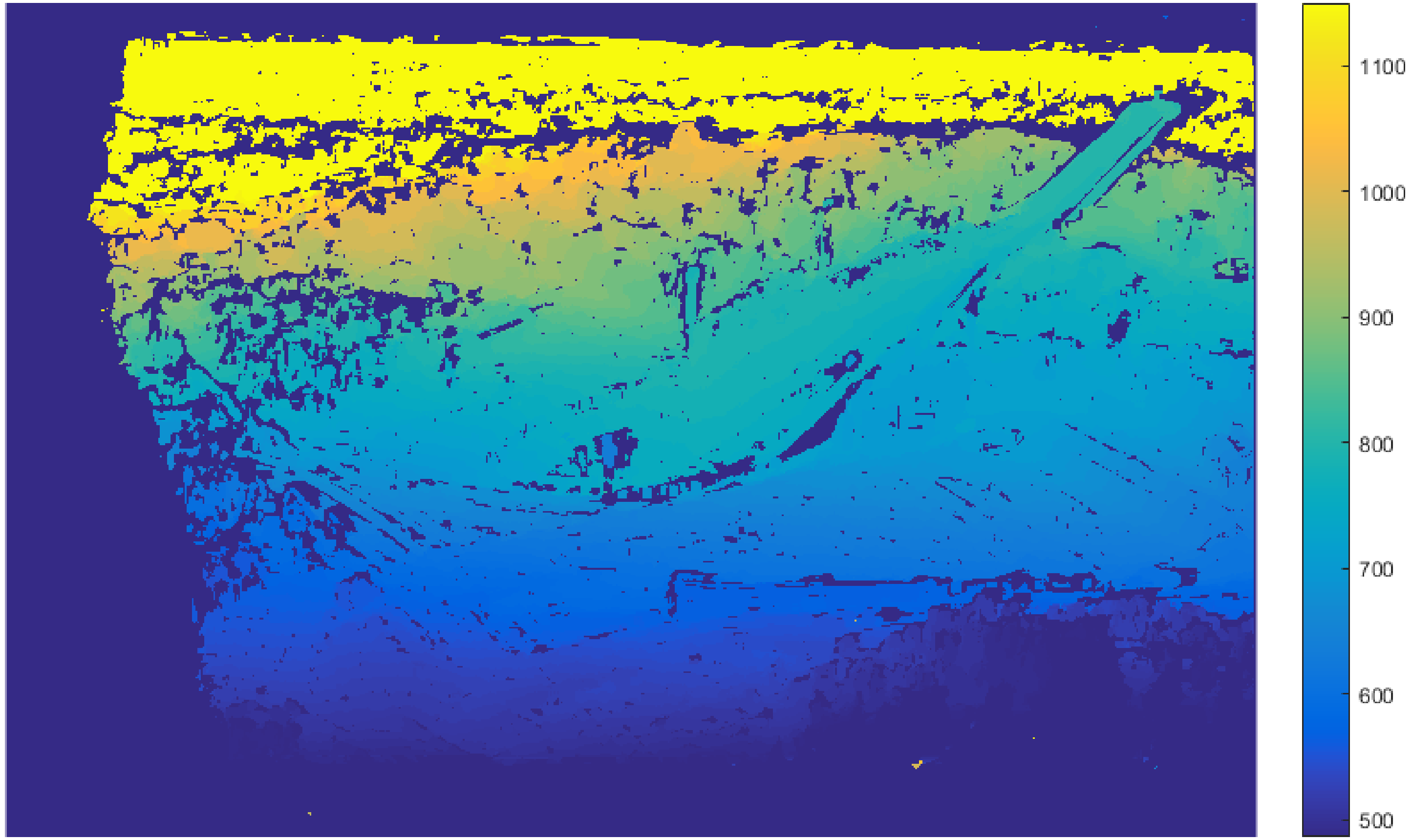
Plane sweep and ambiguities

- Multiple views can resolve ambiguities in difficult areas!



Plane Sweep

- Depths with mean ZNCC score above a threshold $\alpha = 0.8$



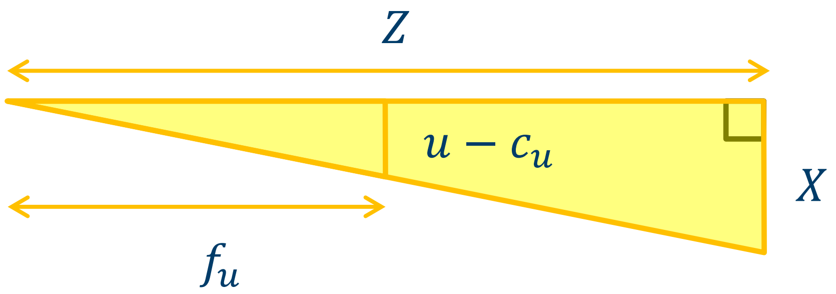
From plane sweep depths to 3D point clouds

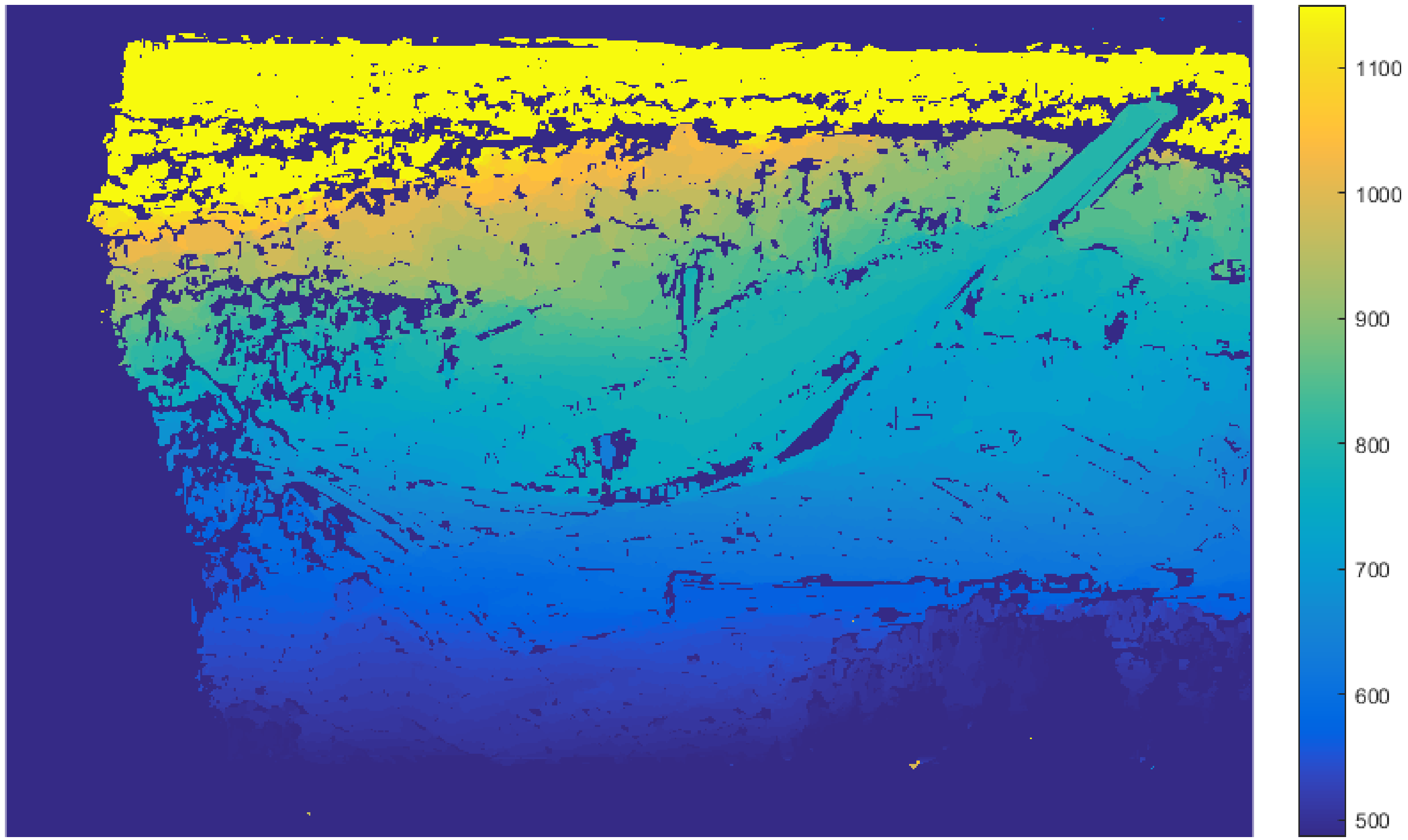
- For fronto-parallel sweep

$$Z_m = d_m$$

$$X = (u - c_u)\frac{Z_m}{f_u}$$

$$Y = (v - c_v)\frac{Z_m}{f_v}$$





3D point cloud

- Removed points where $Z > 1000$
- Monochrome image used as texture

Plane sweep through oriented planes

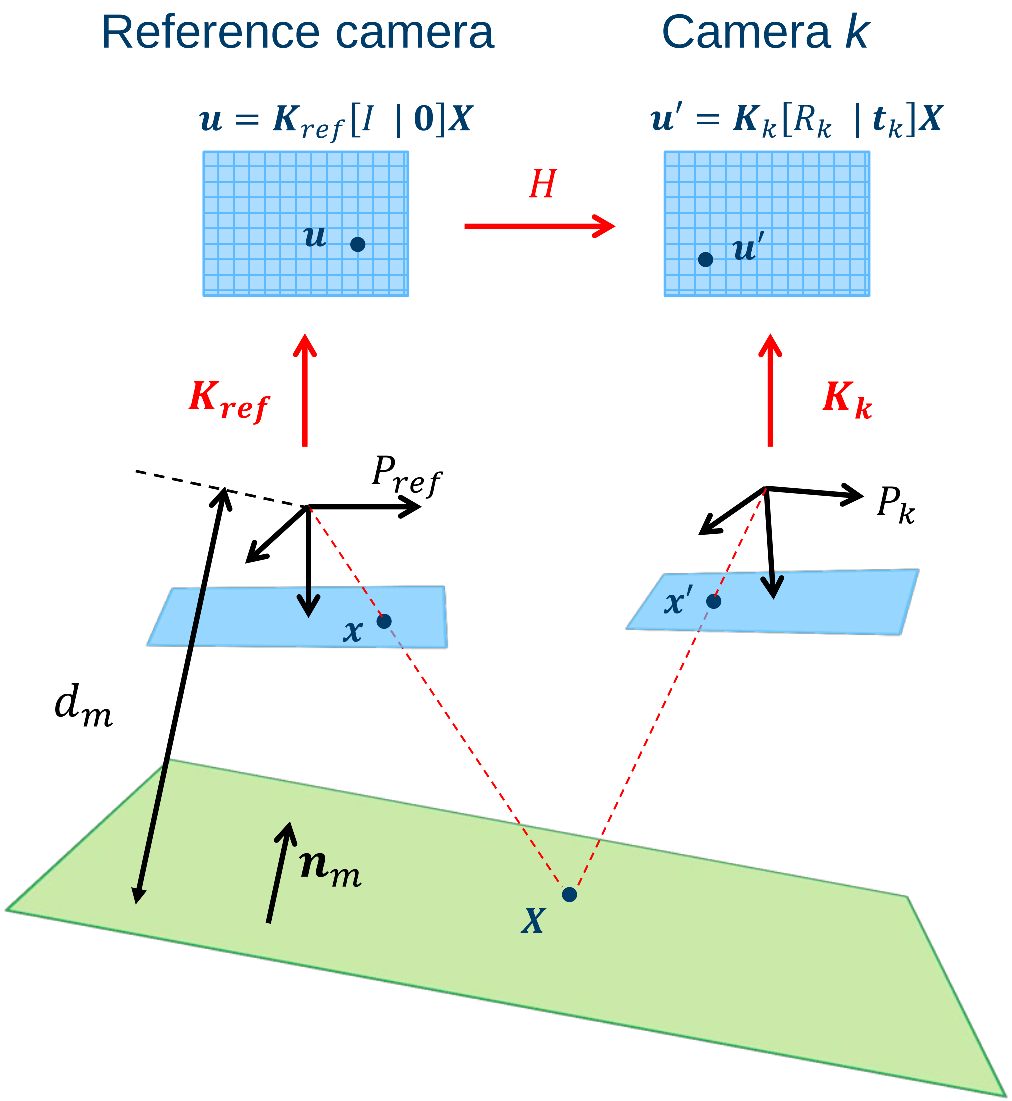
- Fronto-parallel

$$\boldsymbol{n}_m = \begin{bmatrix} 0 & 0 & -1 \end{bmatrix}^T$$

$$Z_m(u,v) = d_m$$

- Other plane orientations

$$Z_m(u,v) = \frac{-d_m}{\begin{bmatrix} u & v & 1 \end{bmatrix} K_{ref}^{-T} \boldsymbol{n}_m}$$



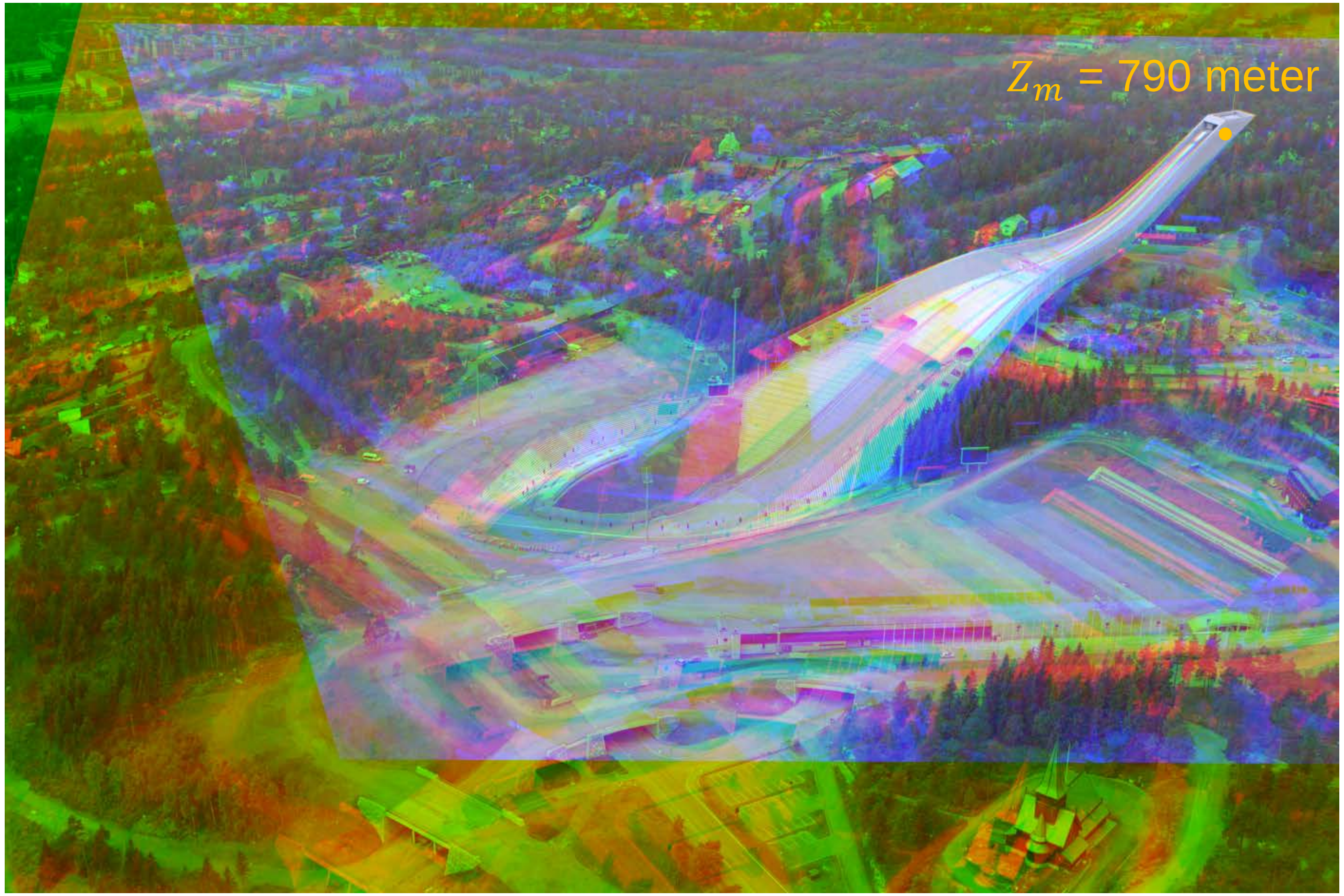
Robert Collins, A Space-Sweep Approach to True Multi-Image Matching, CVPR 1996.
D. Gallup, J.-M. Frahm, P. Mordohai, Q. Yang and M. Pollefeys, Real-Time Plane-Sweeping Stereo with Multiple Sweeping Directions, CVPR 2007

Plane sweep with ground normal

Red:

Green:

Blue:



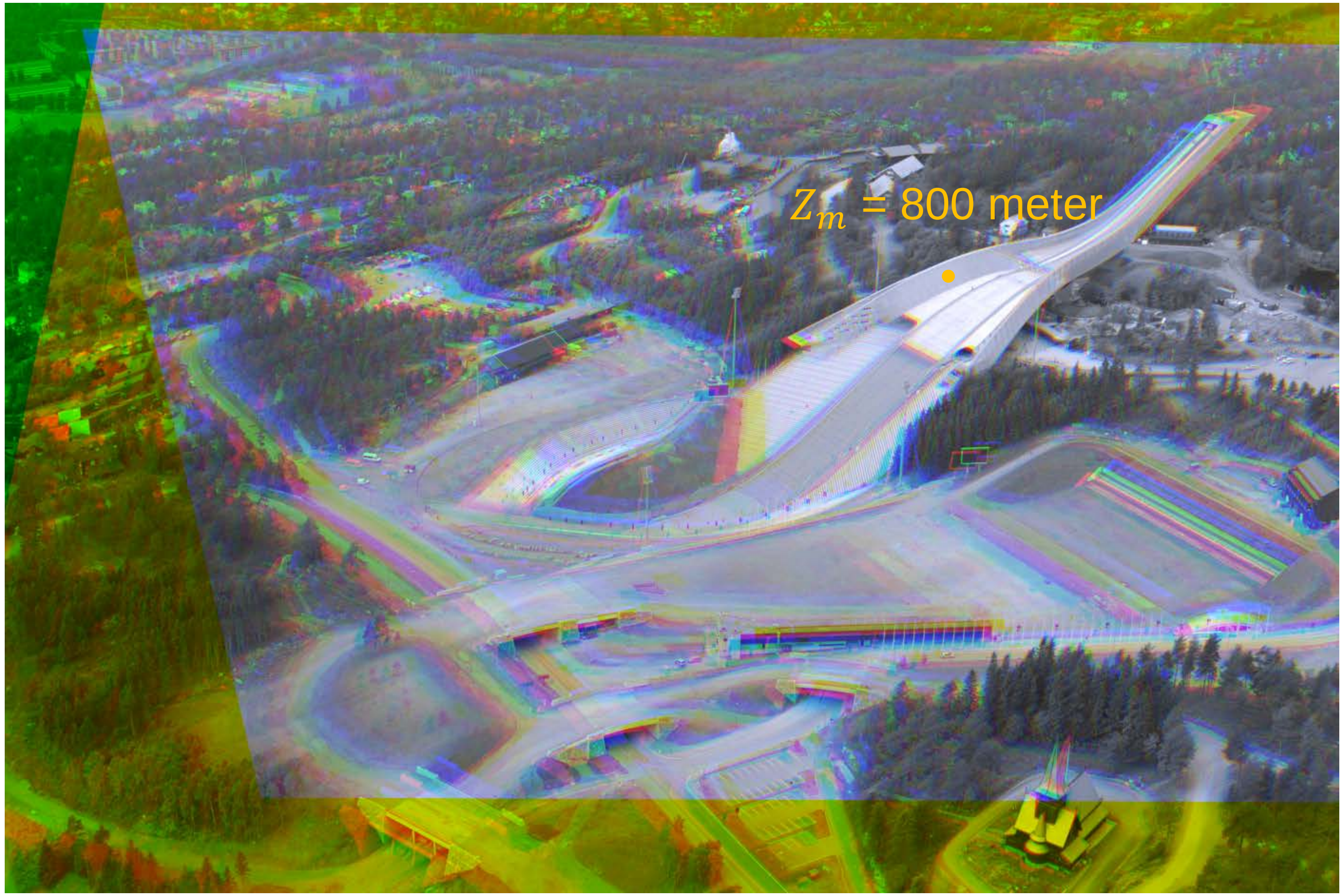
d_m = 200 meter below reference camera

Plane sweep with ground normal

Red:

Green:

Blue:



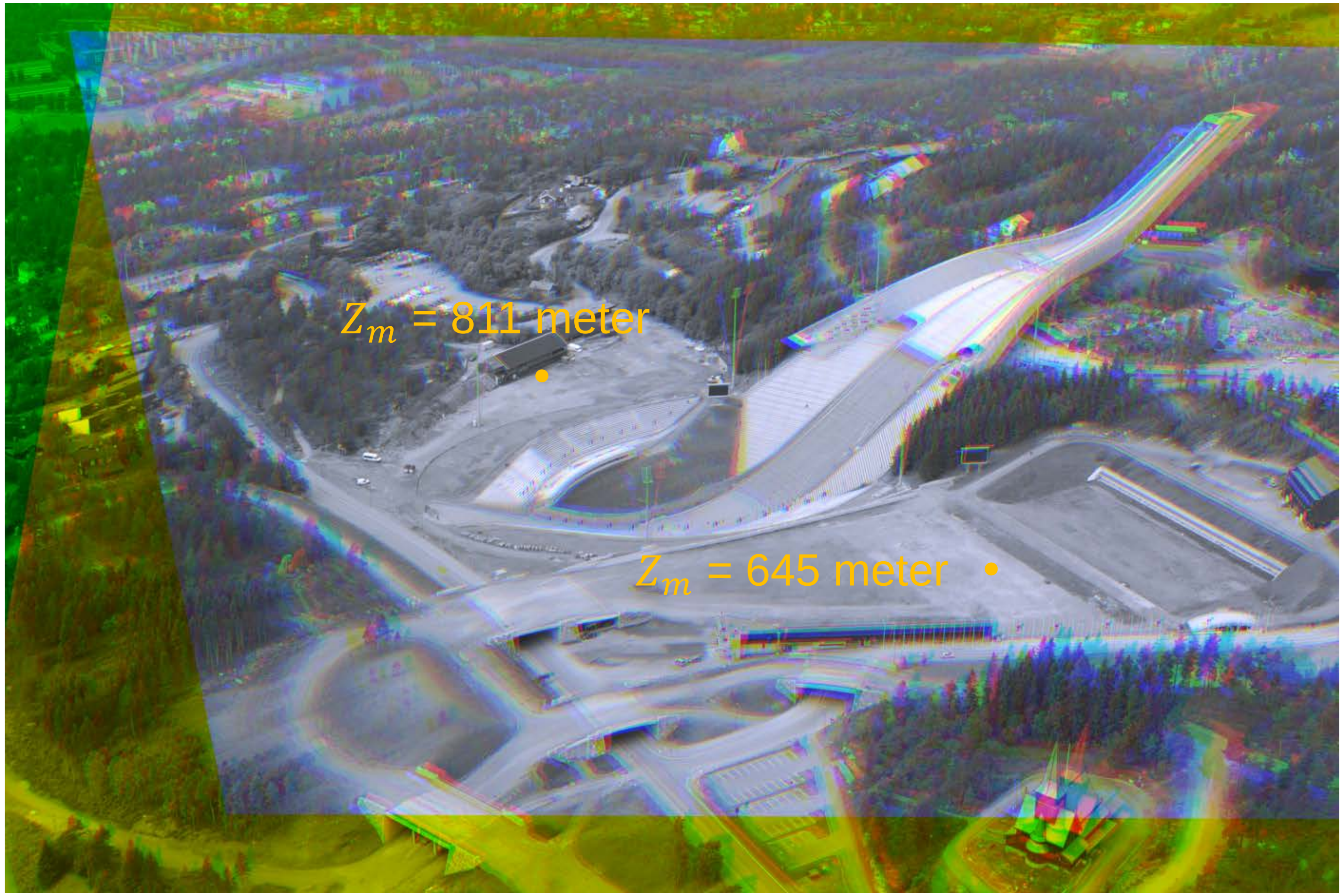
d_m = 261 meter below reference camera

Plane sweep with ground normal

Red:

Green:

Blue:



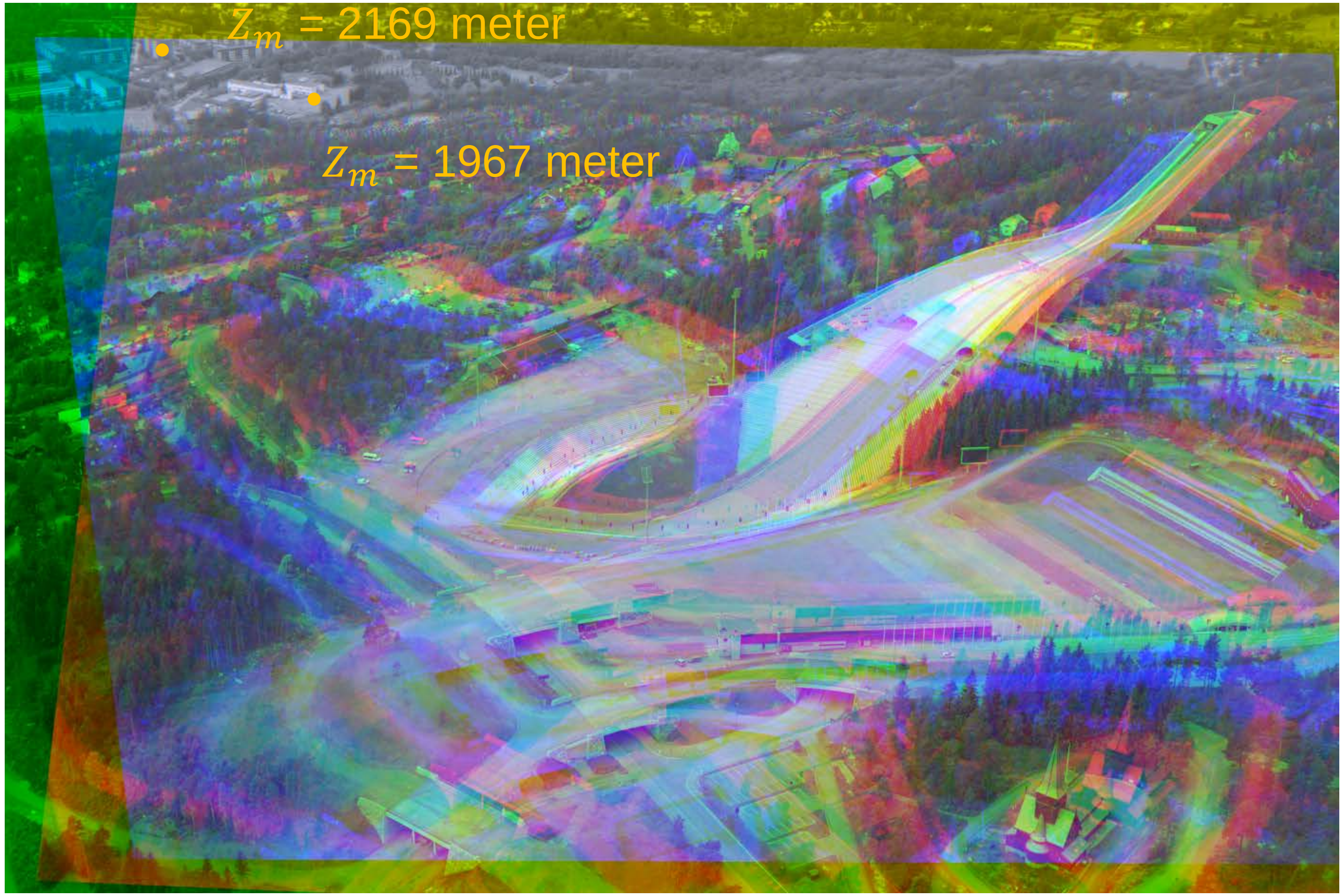
d_m = 298 meter below reference camera

Plane sweep with ground normal

Red:

Green:

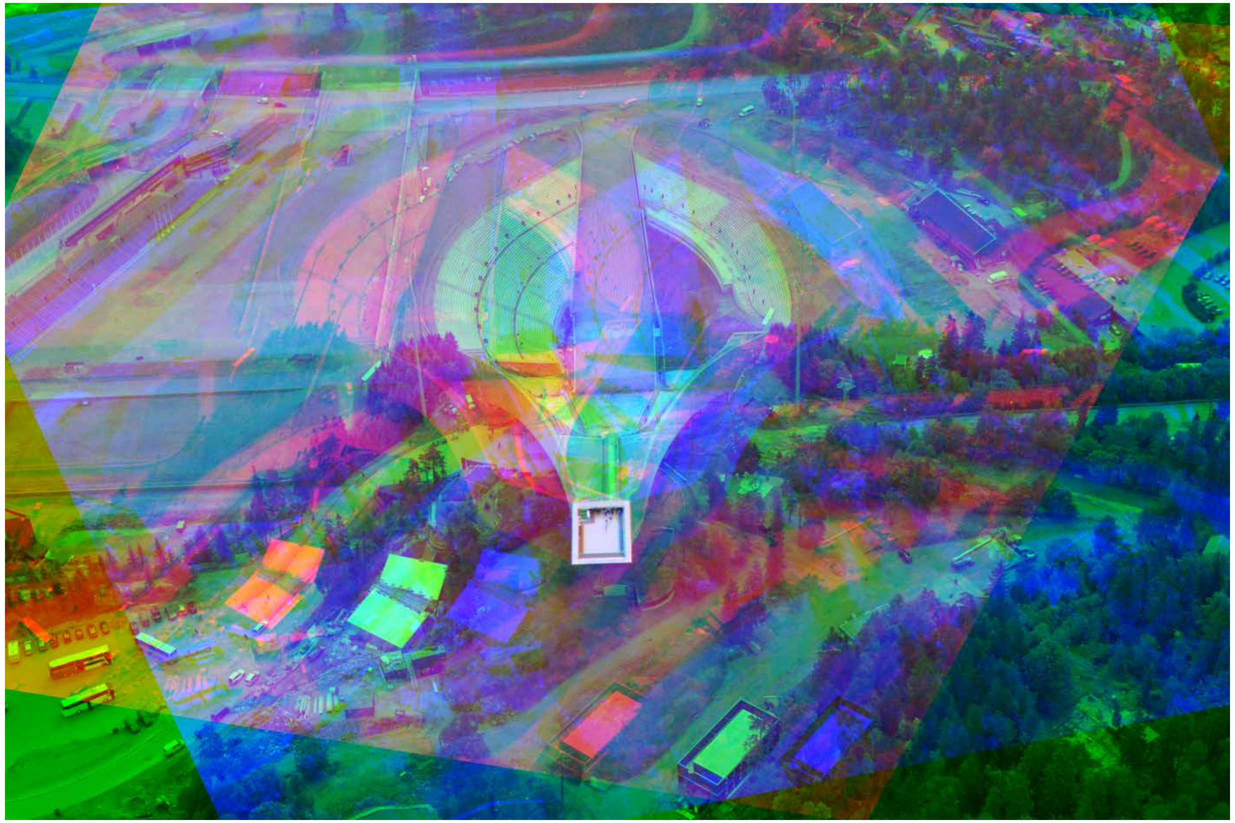
Blue:



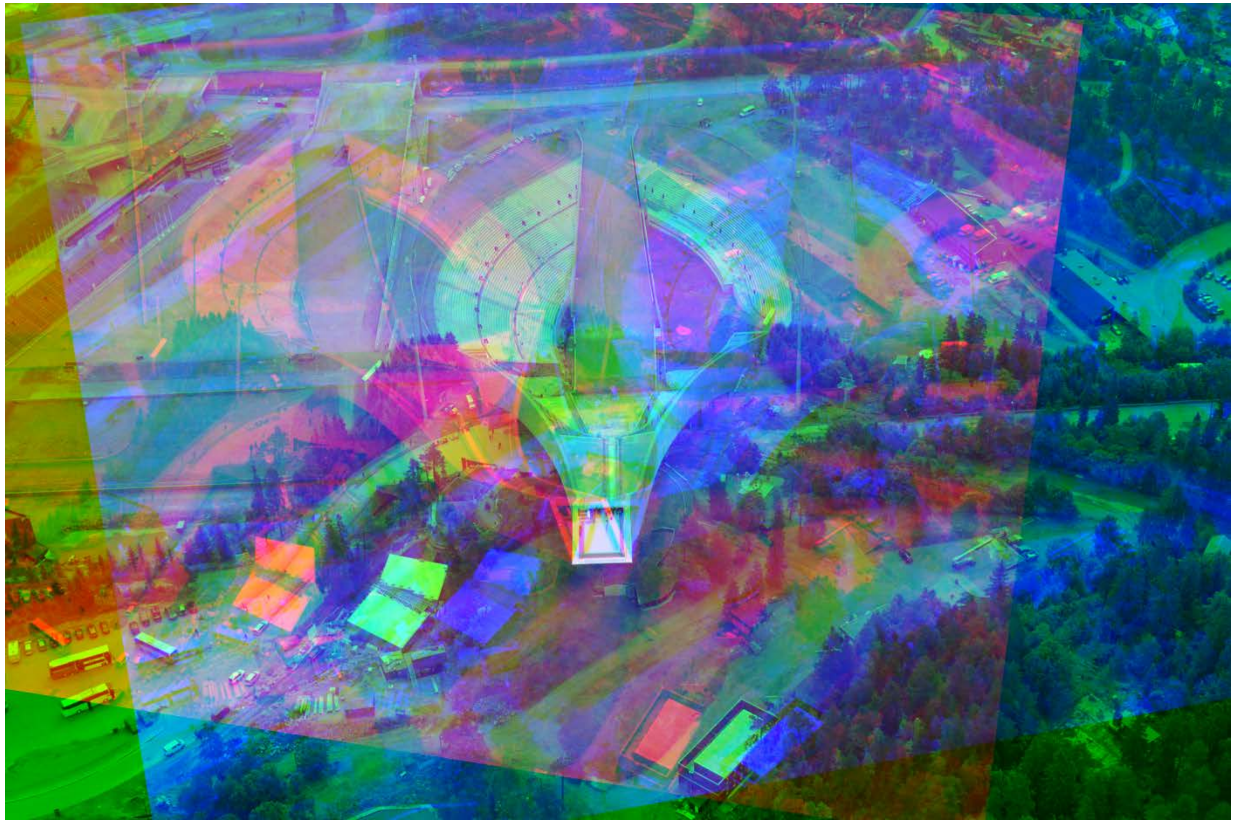
d_m = 471 meter below reference camera

Plane sweep with ground normal compared to fronto-parallel

Ground normal

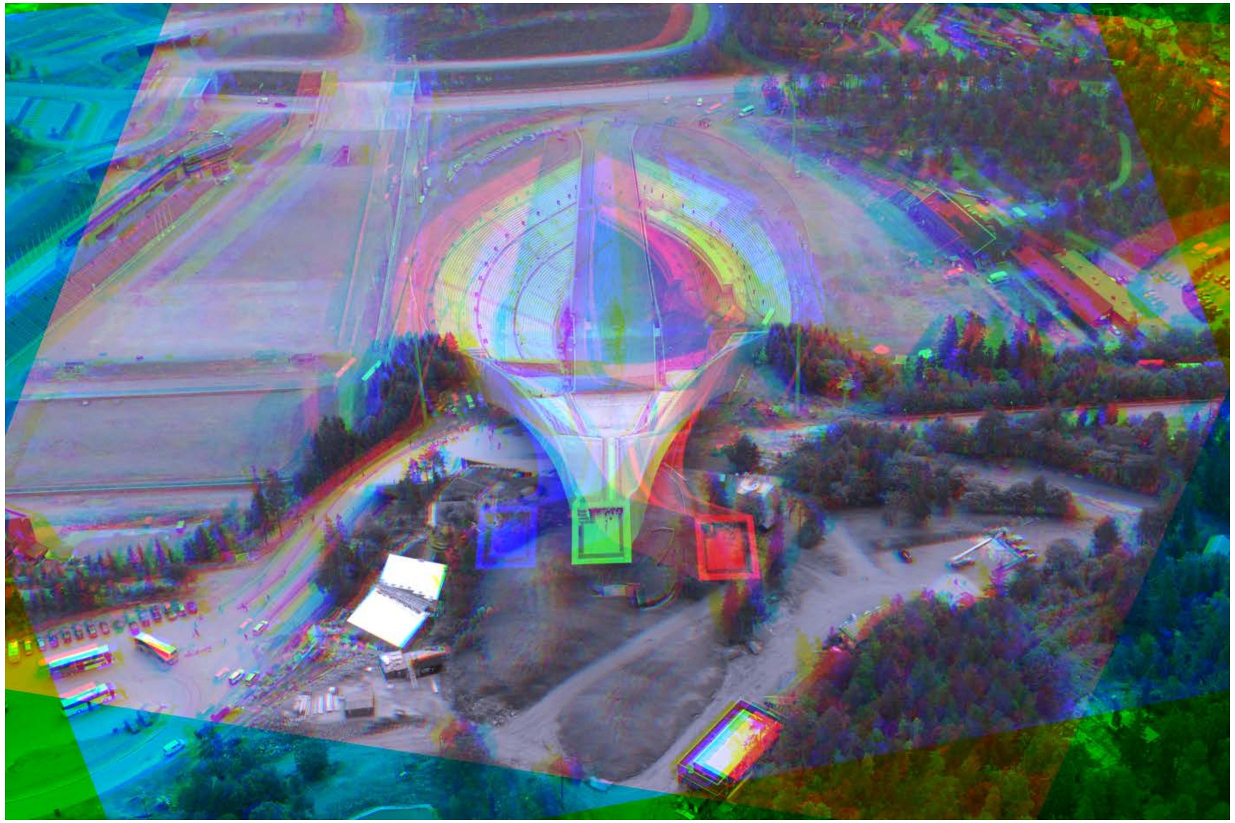


Fronto-parallel

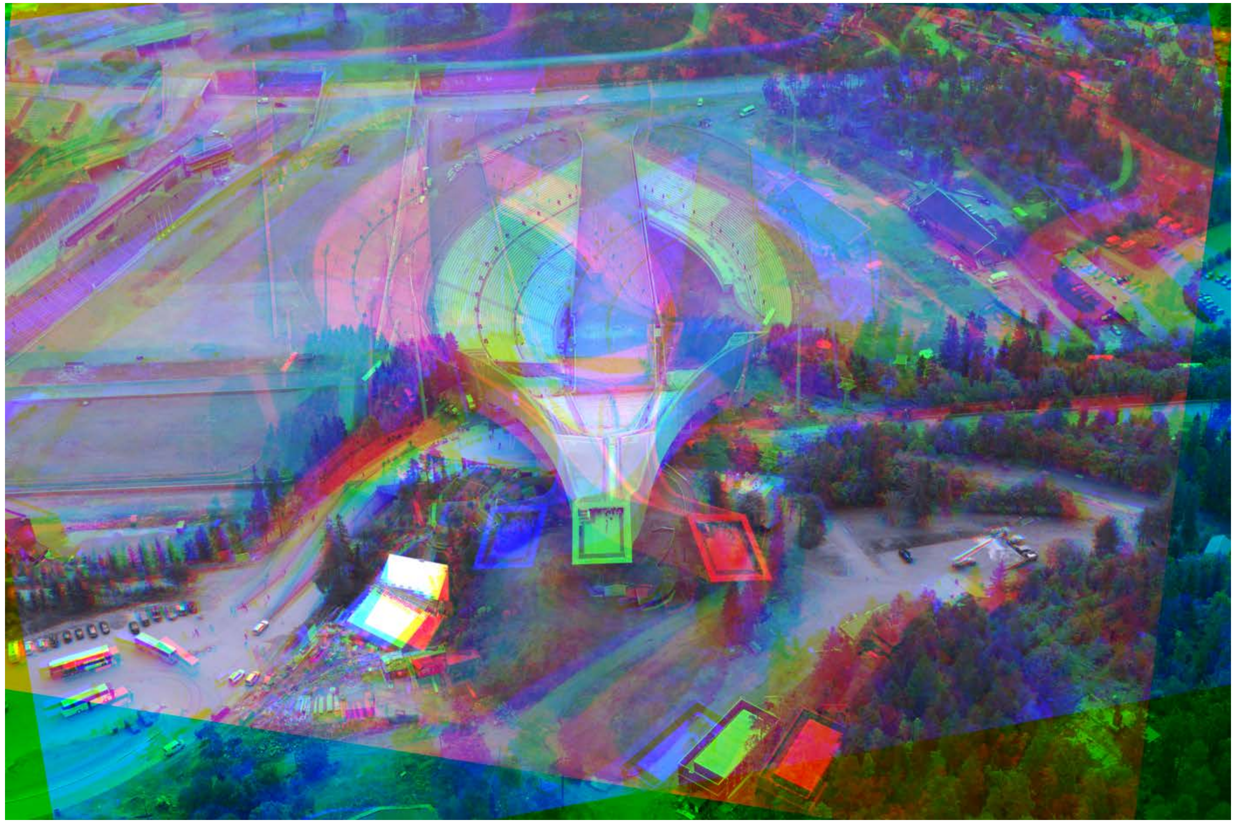


Plane sweep with ground normal compared to fronto-parallel

Ground normal

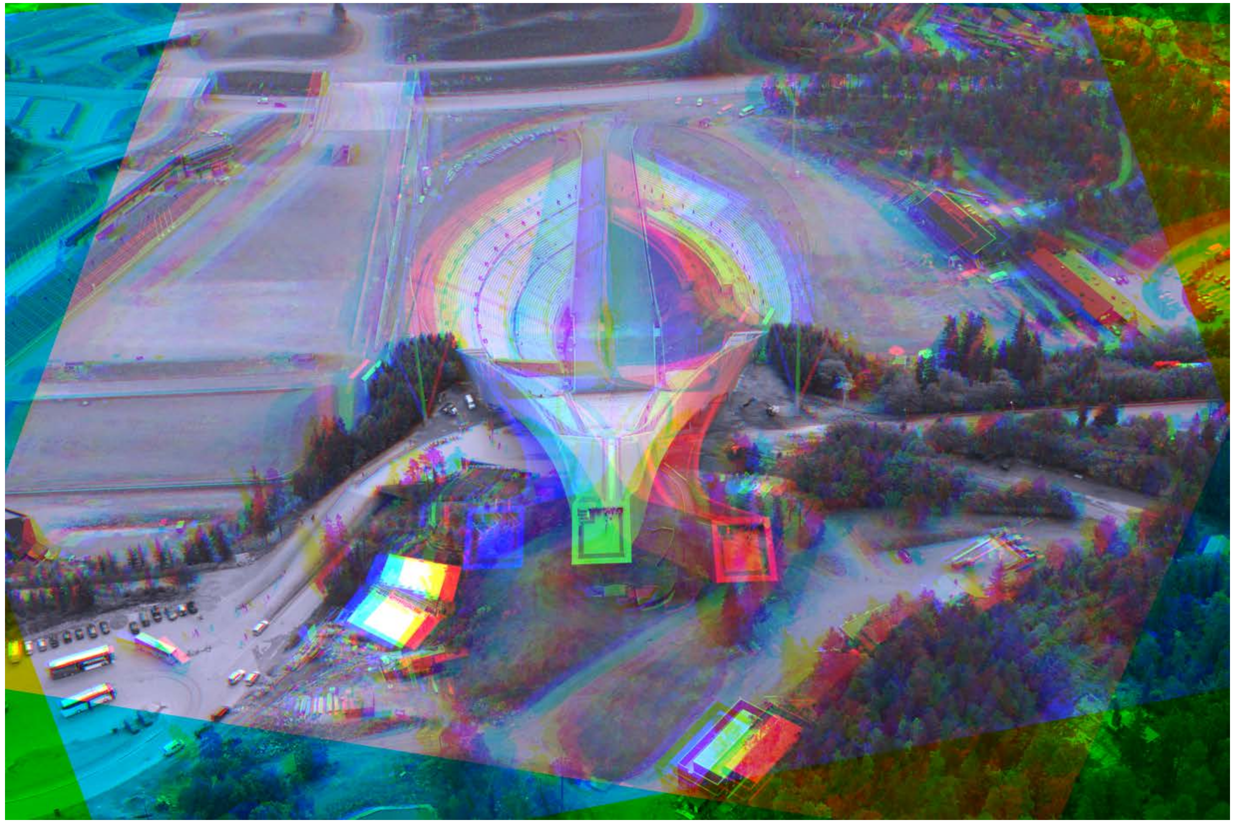


Fronto-parallel

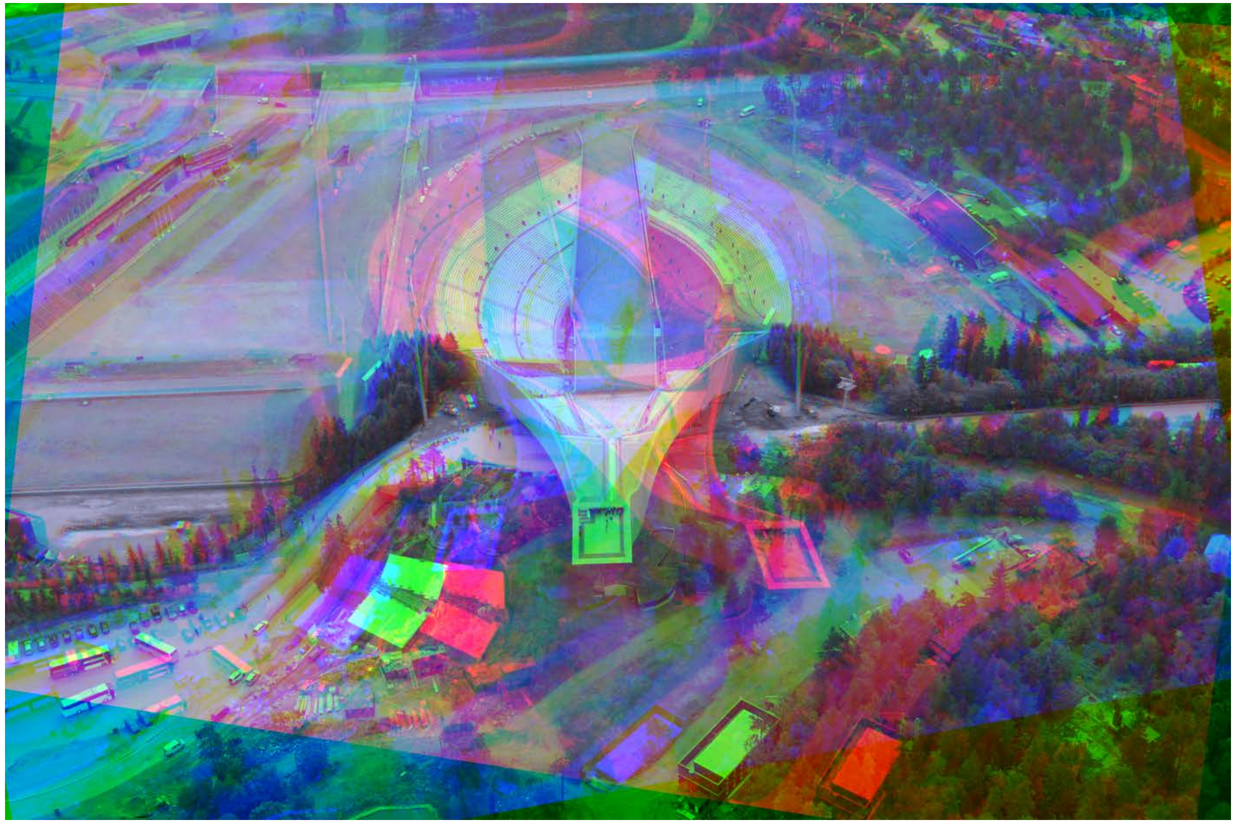


Plane sweep with ground normal compared to fronto-parallel

Ground normal

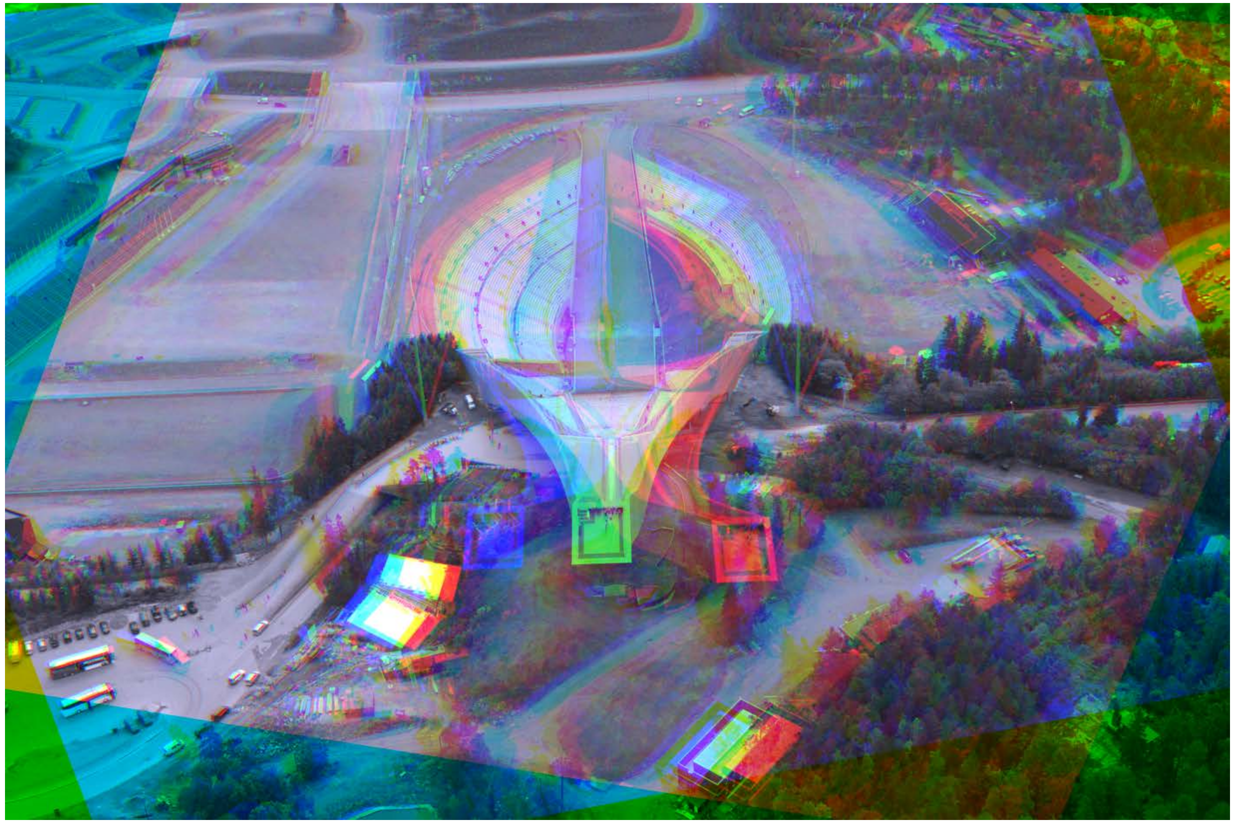


Fronto-parallel

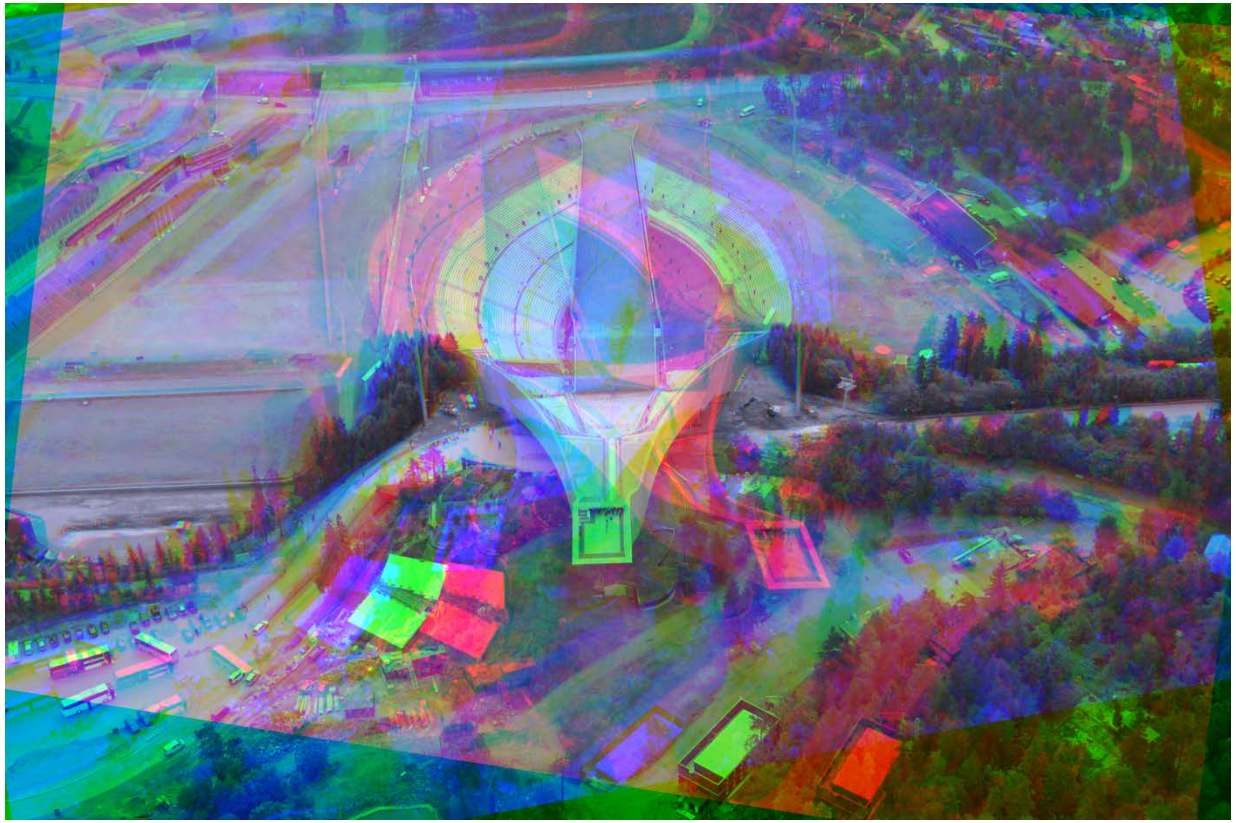


Plane sweep with ground normal compared to fronto-parallel

Ground normal



Fronto-parallel

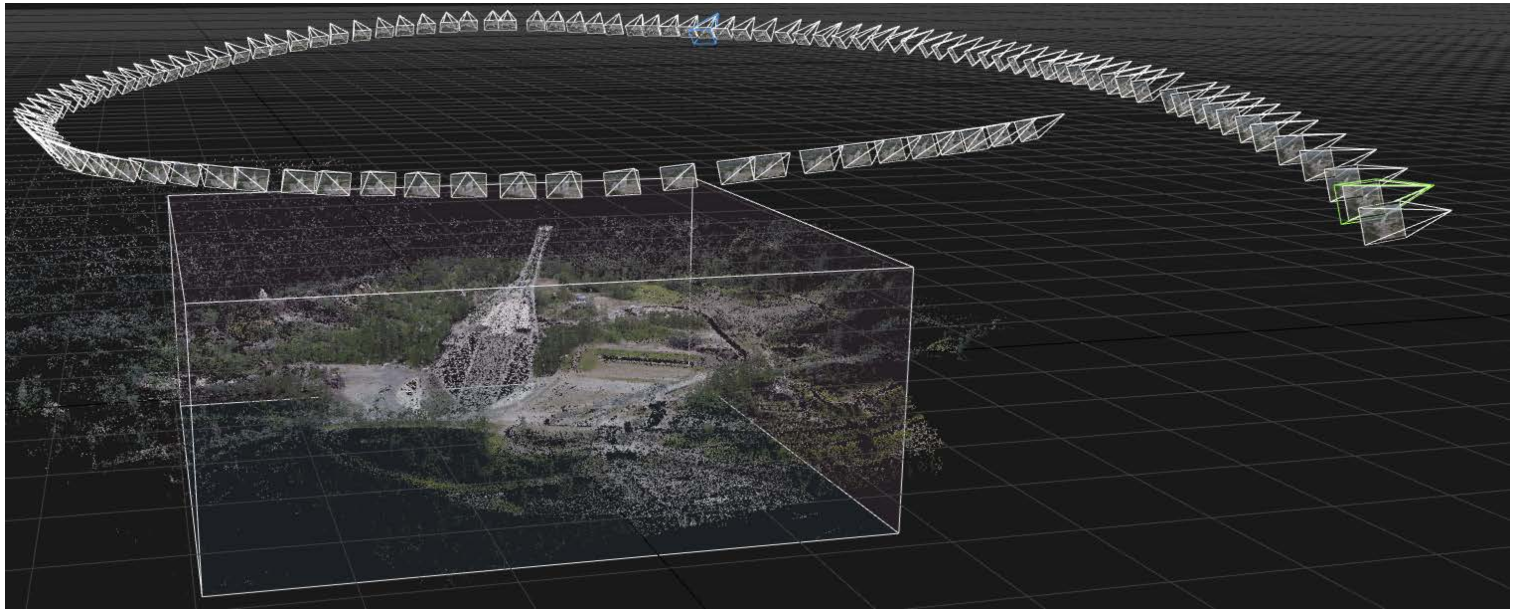


Use surface normal to avoid distortions and get better matching!

From plane sweep depths to 3D surface

1. Choose a set of reference views
   – Can be virtual views!

2. Compute the point clouds for each reference view



From plane sweep depths to 3D surface

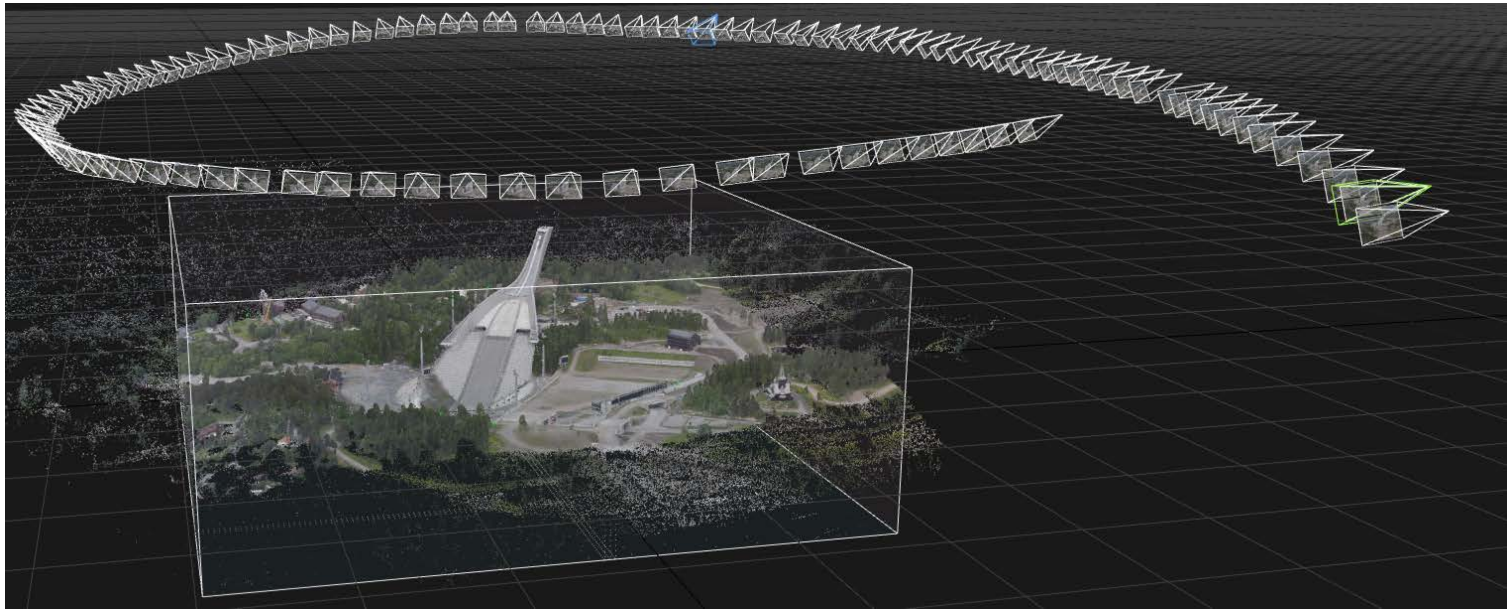
1. Choose a set of reference views
   – Can be virtual views!

2. Compute the point clouds for each reference view

3. Combine the point clouds in a common coordinate system

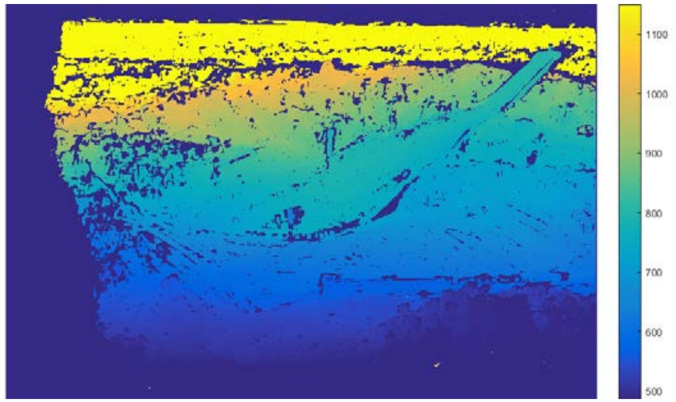
4. Filter out bad points
   – Remove point with little support from other point clouds

5. Perform surface reconstruction



Approaches to multiple-view stereo

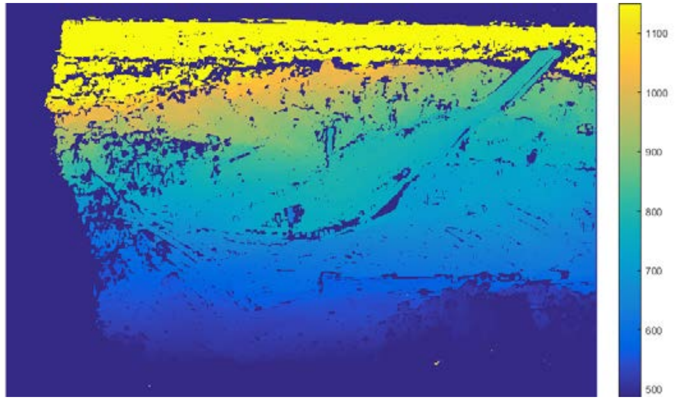
- Plane sweep depth maps
  - Robust and adaptable multiple-view stereo matching
  - Real-time applications
  - Fusion of point clouds from different reference views
  - Sampling of scene depends on the reference views

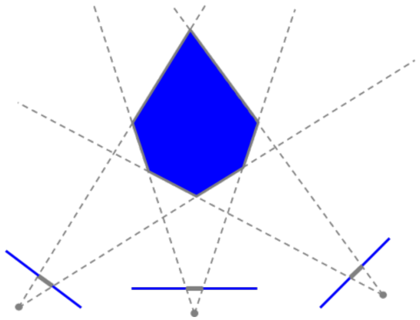


Approaches to multiple-view stereo

- Plane sweep depth maps
  - Robust and adaptable multiple-view stereo matching
  - Real-time applications
  - Fusion of point clouds from different reference views
  - Sampling of scene depends on the reference views

- Volumetric stereo
  - View-independent representation



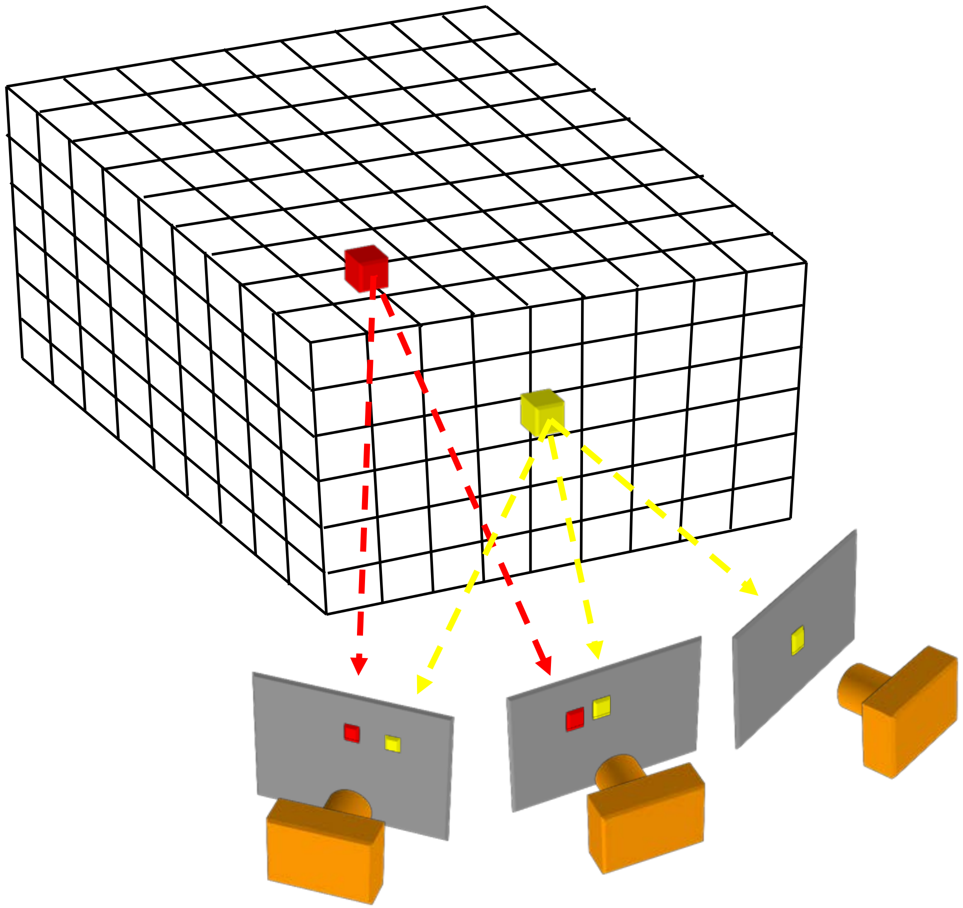


Volumetric stereo

- Volume is discretized as a 3D grid of *voxels*

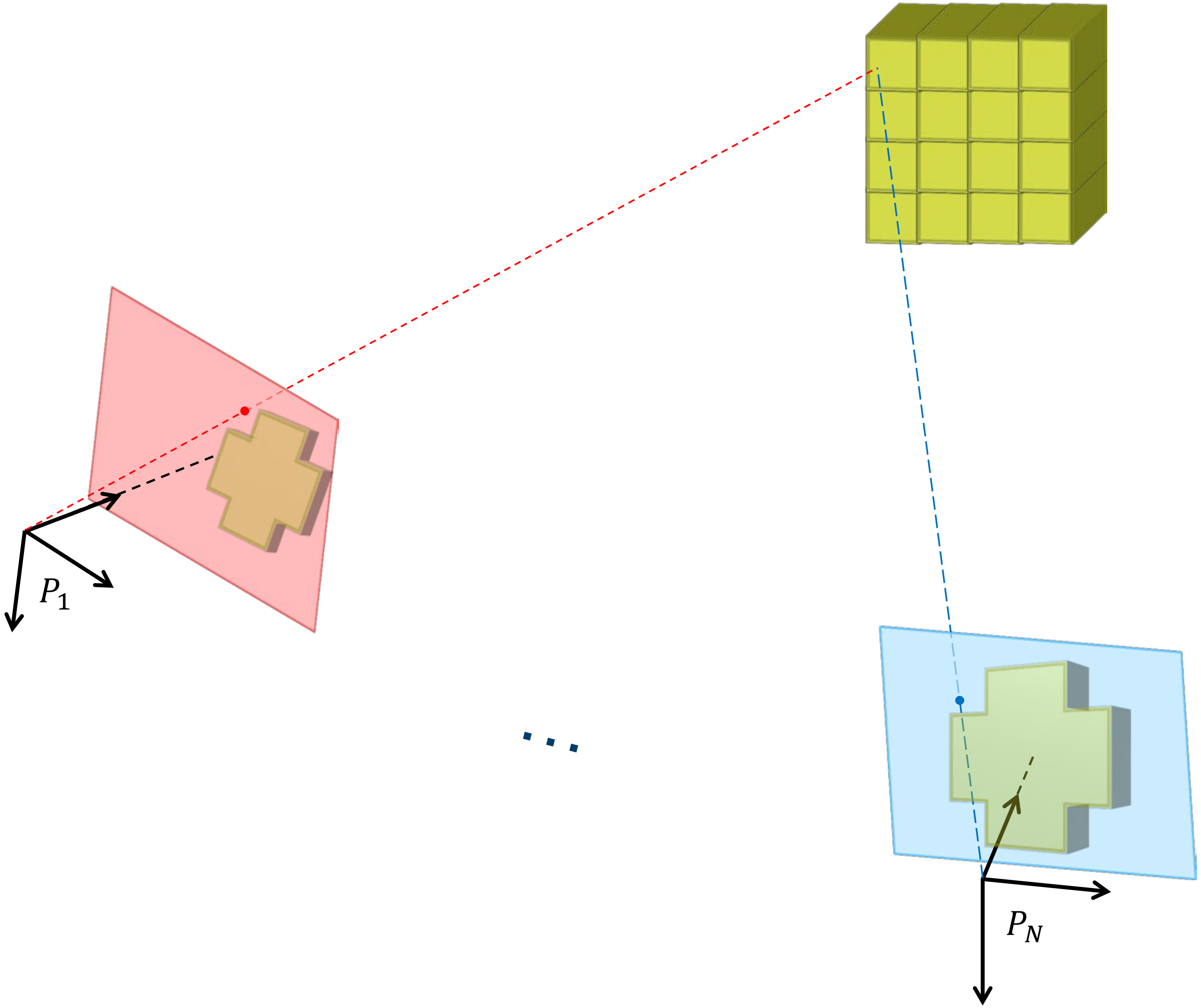
- Goal:
  - Assign RGB values to voxels *photo-consistent* with images

- A *photo-consistent* scene exactly reproduces your input images from the same viewpoints



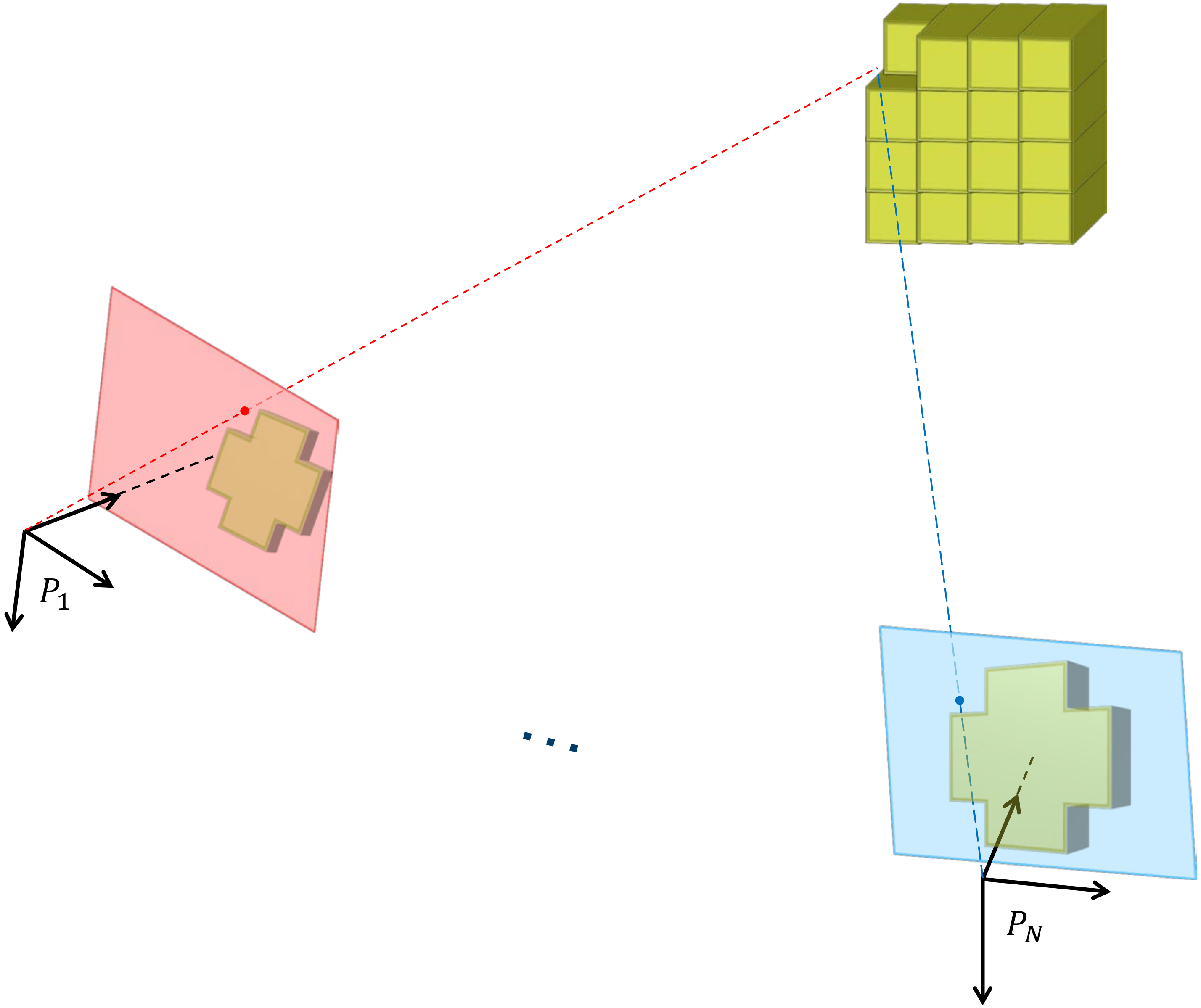
Space Carving

1. Choose a voxel
   on the current surface
2. Project to visible input images



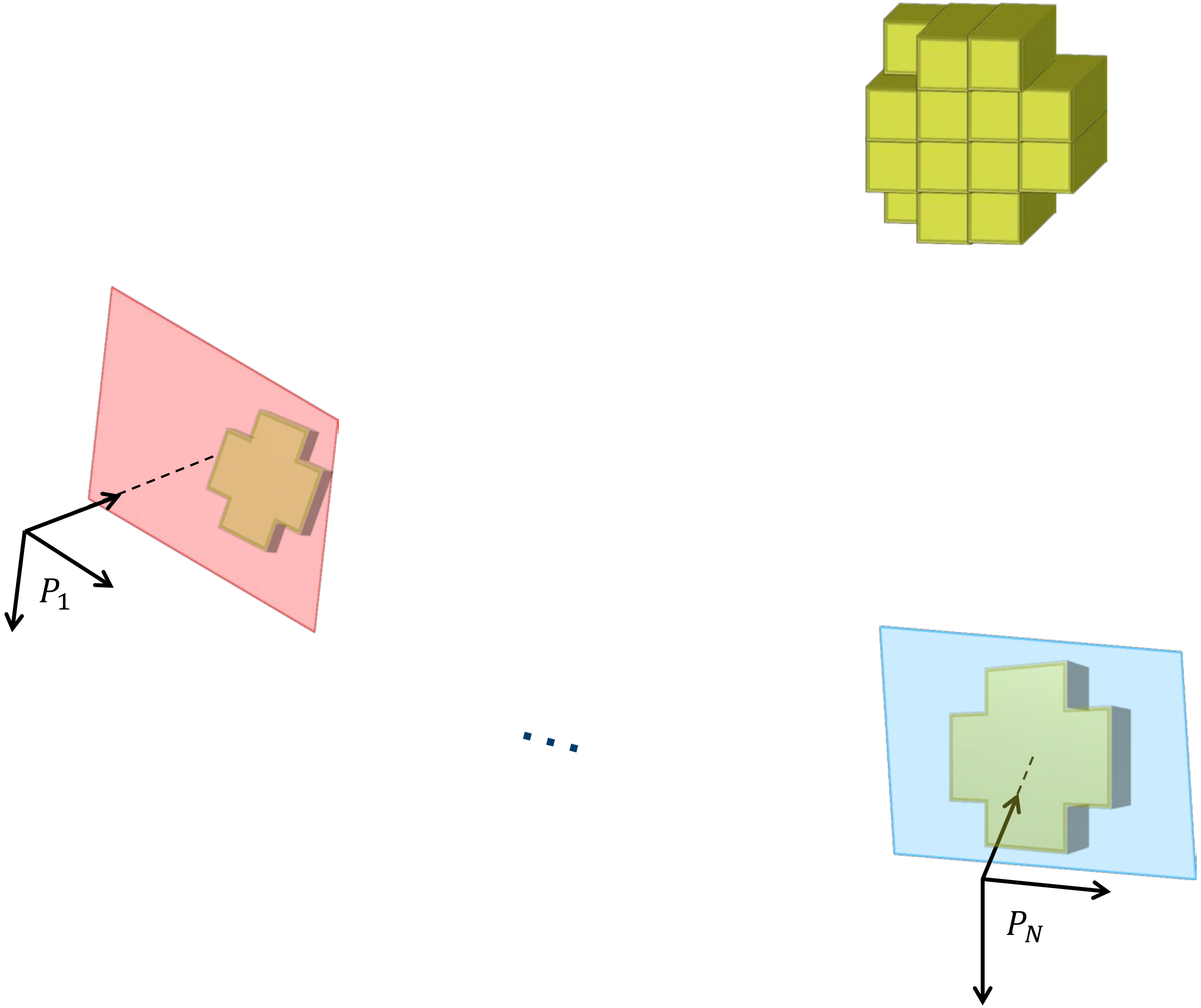
Space Carving

1. Choose a voxel
   on the current surface
2. Project to visible input images
3. Carve if not photo-consistent



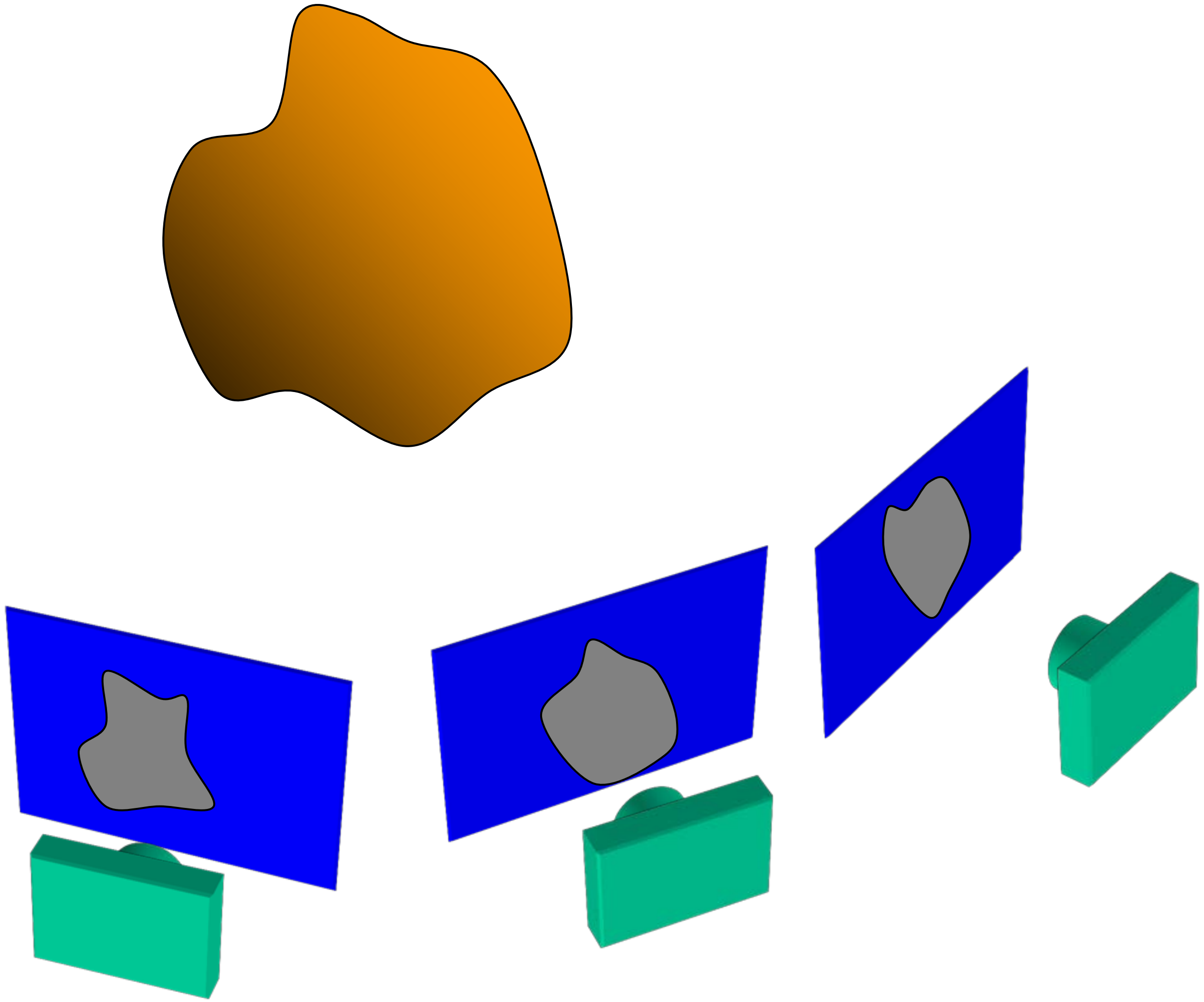
Space Carving

1. Choose a voxel
   on the current surface
2. Project to visible input images
3. Carve if not photo-consistent
4. Repeat until convergence



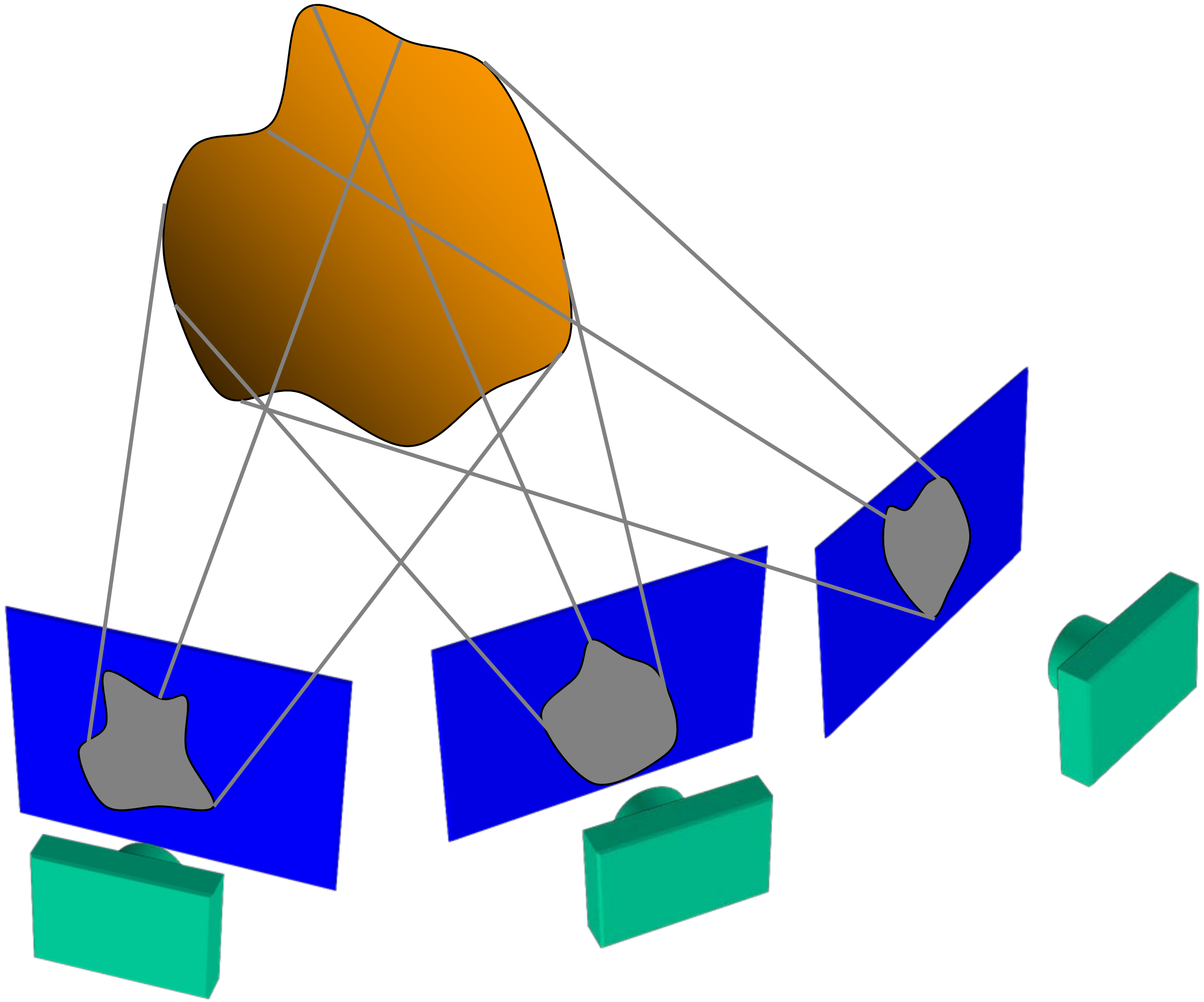
Visual Hull

1. Segment out object from background



Visual Hull

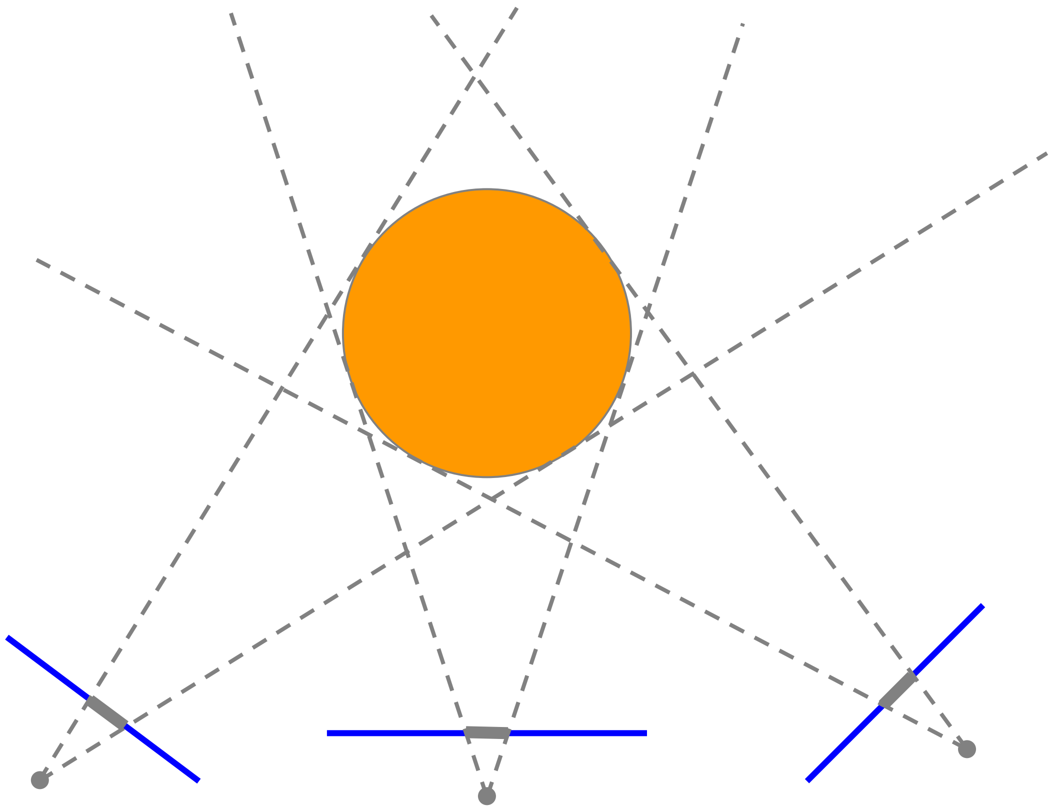
1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes



Visual Hull – Volume intersection

1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes

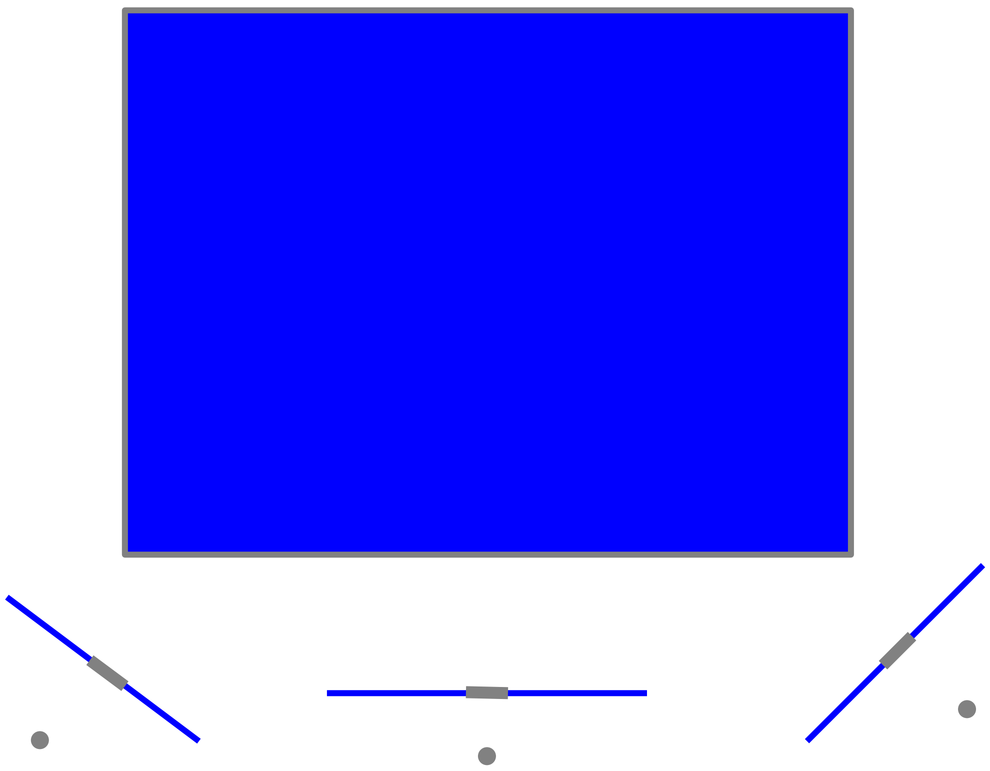
From above



Visual Hull – Volume intersection

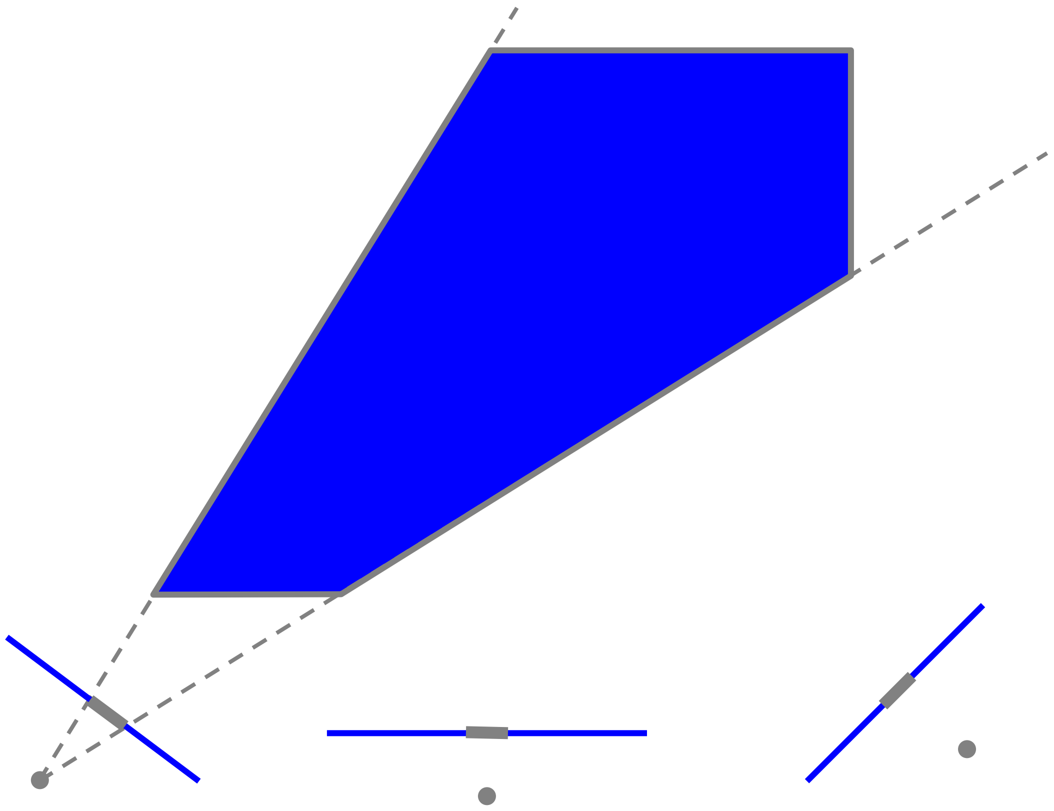
1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes

Initial volume



Visual Hull – Volume intersection

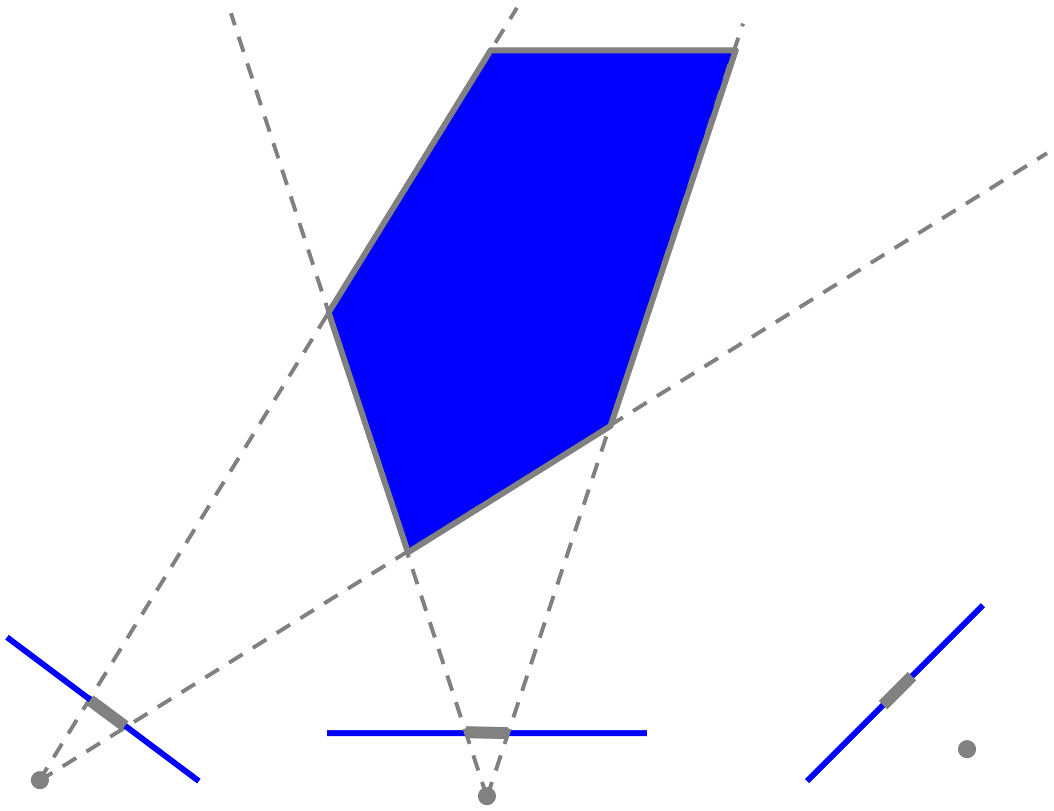
1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes



Visual Hull – Volume intersection

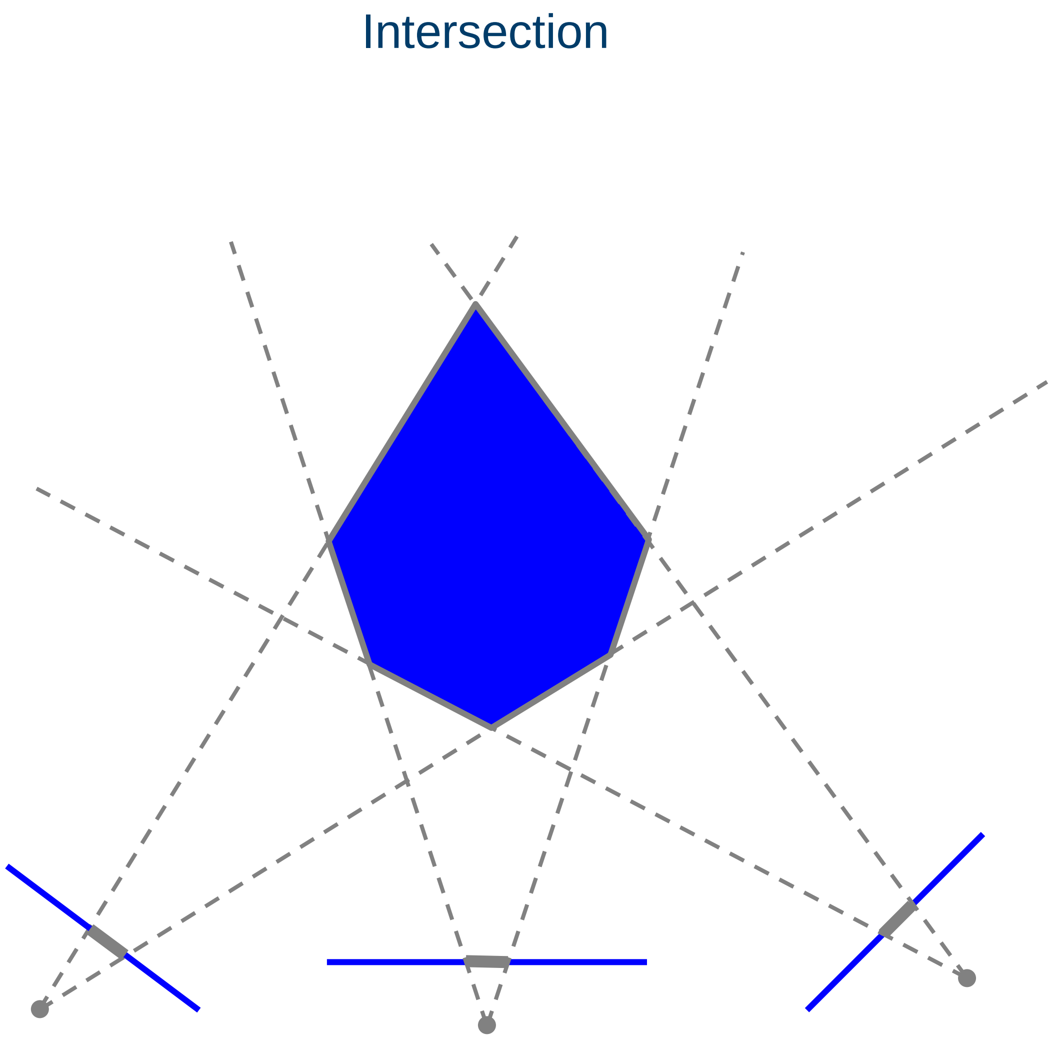
1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes

Intersection



Visual Hull – Volume intersection

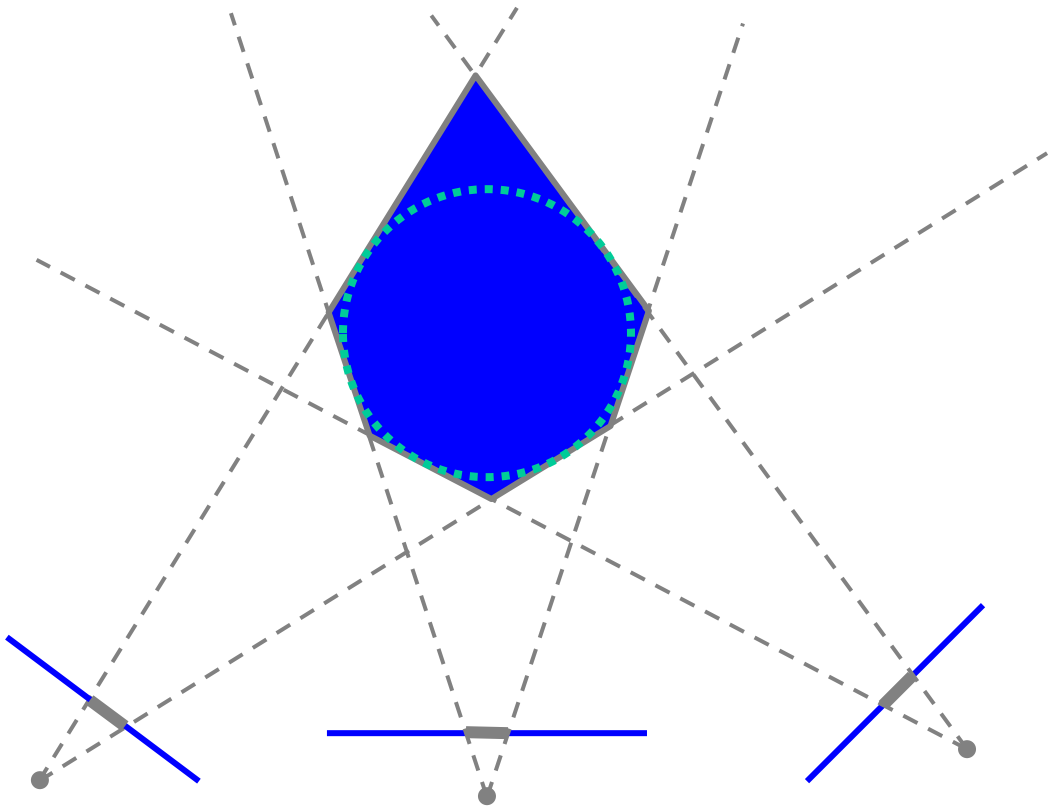
1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes



Visual Hull – Volume intersection

Visual hull
Reconstruction contains the structure

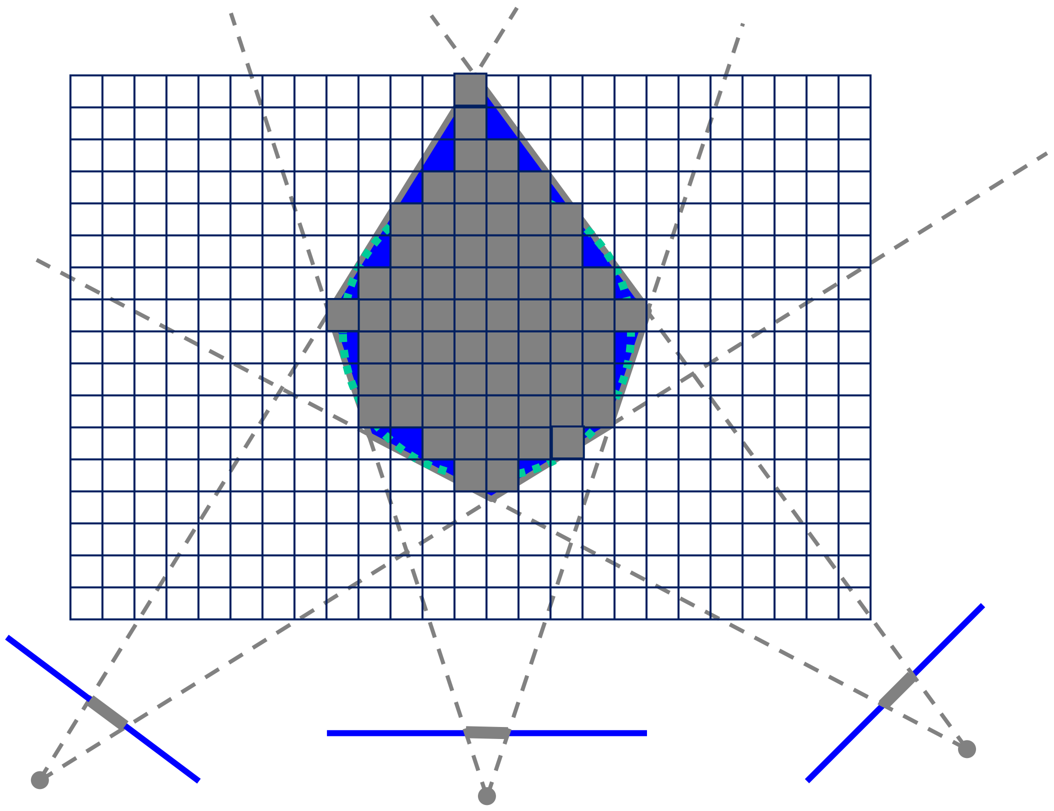
1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes



UNIK4690

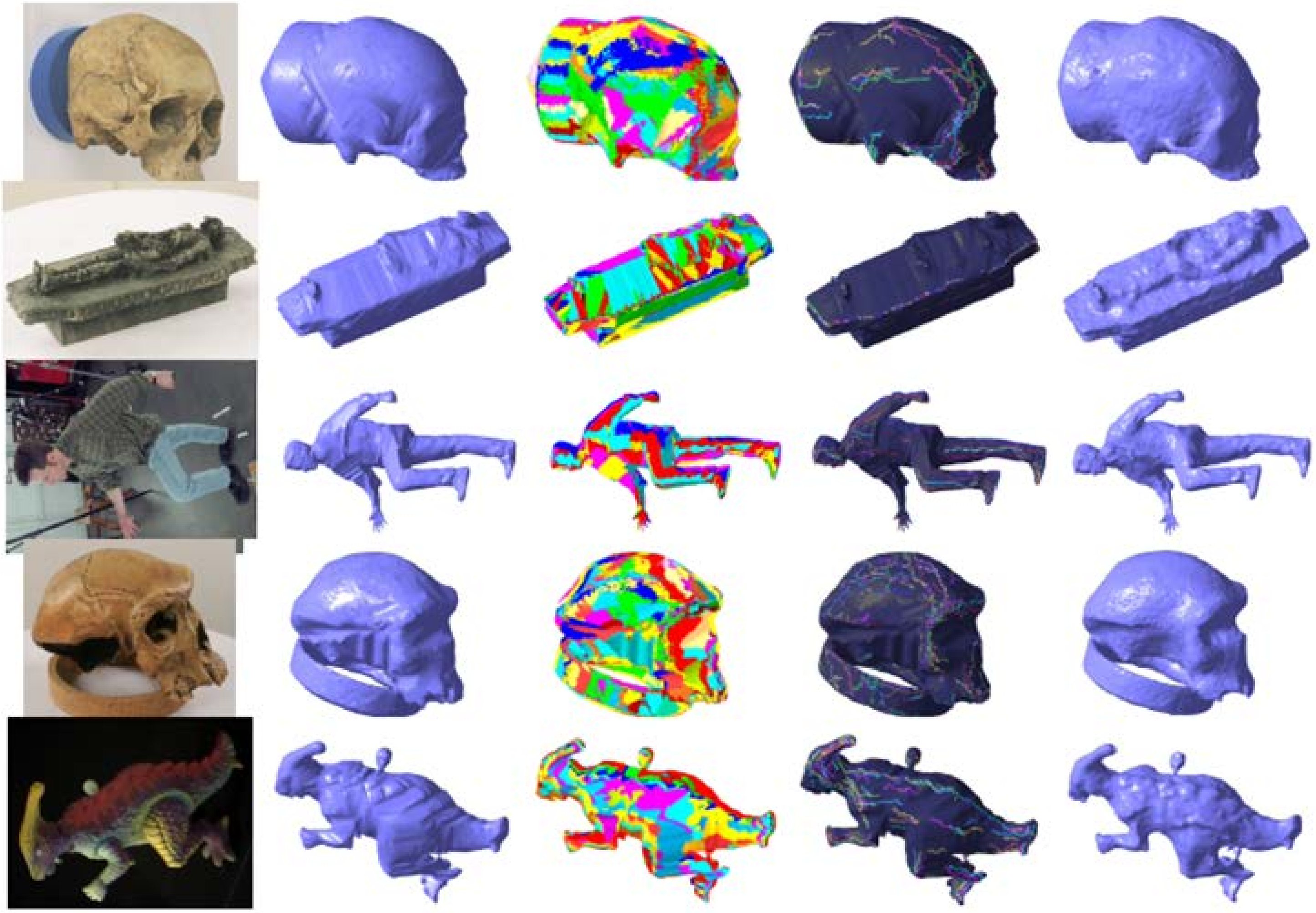
Visual Hull – Volume intersection

1. Segment out object from background
2. Backproject each silhoutte
3. Intersect backprojected volumes



Carved visual hulls

1. Compute visual hull
2. Find rims and constrain them to be fixed
3. Carve the visual hull to optimize photo-consistency

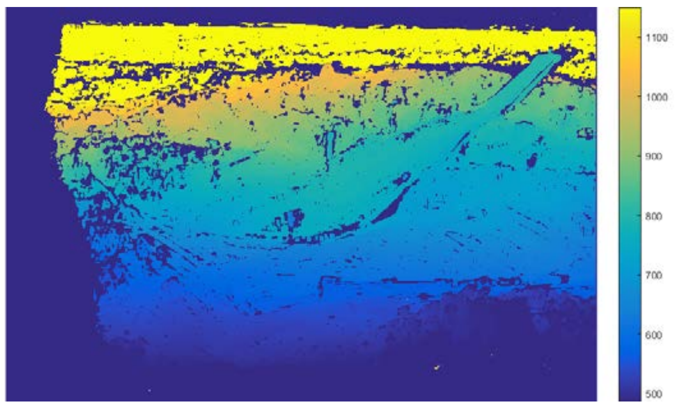


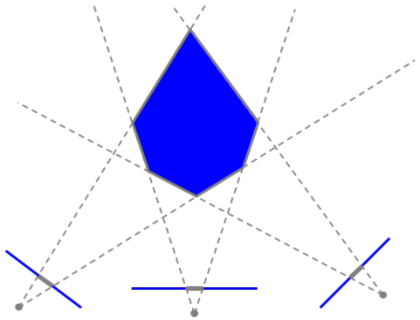
Yasutaka Furukawa and Jean Ponce, Carved Visual Hulls for Image-Based Modeling, ECCV 2006.

Approaches to multiple-view stereo

- Plane sweep depth maps
  - Robust and adaptable multiple-view stereo matching
  - Real-time applications
  - Fusion of point clouds from different reference views
  - Sampling of scene depends on the reference views

- Volumetric stereo
  - View-independent representation
  - Need silhouette extraction
  - Accuracy depends on the density of the grid
  - High computational and memory costs



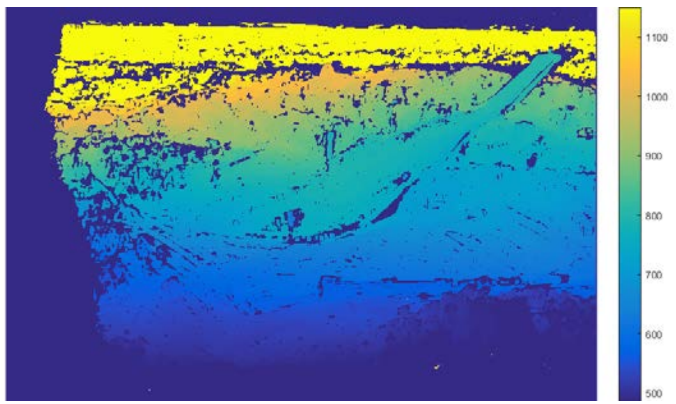


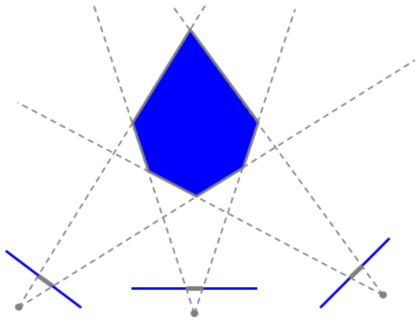
Approaches to multiple-view stereo

- Plane sweep depth maps
  - Robust and adaptable multiple-view stereo matching
  - Real-time applications
  - Fusion of point clouds from different reference views
  - Sampling of scene depends on the reference views

- Volumetric stereo
  - View-independent representation
  - Need silhouette extraction
  - Accuracy depends on the density of the grid
  - High computational and memory costs

- Surface expansion
  - Sparse to dense feature correspondences





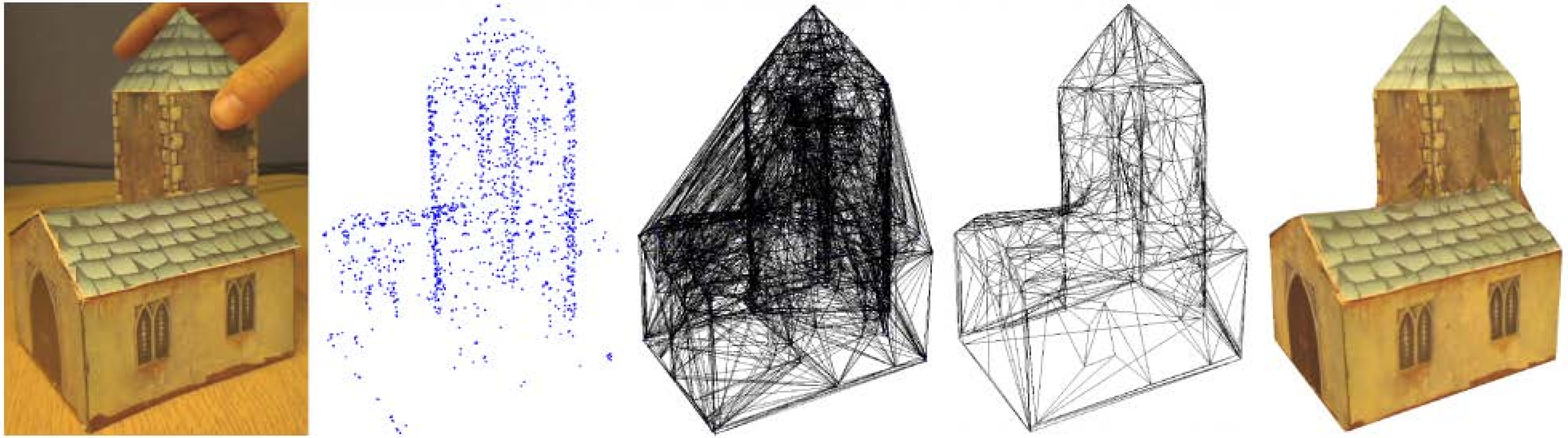
Surface Expansion

1. Extract features and acquire a sparse set of initial matches
2. Iteratively expand matches to nearby locations
3. Use visibility constraints to filter out false matches
4. Perform surface reconstruction

Yasutaka Furukawa and Jean Ponce, Accurate, Dense, and Robust Multi-View Stereopsis, CVPR 2007.

Surface reconstruction from point clouds

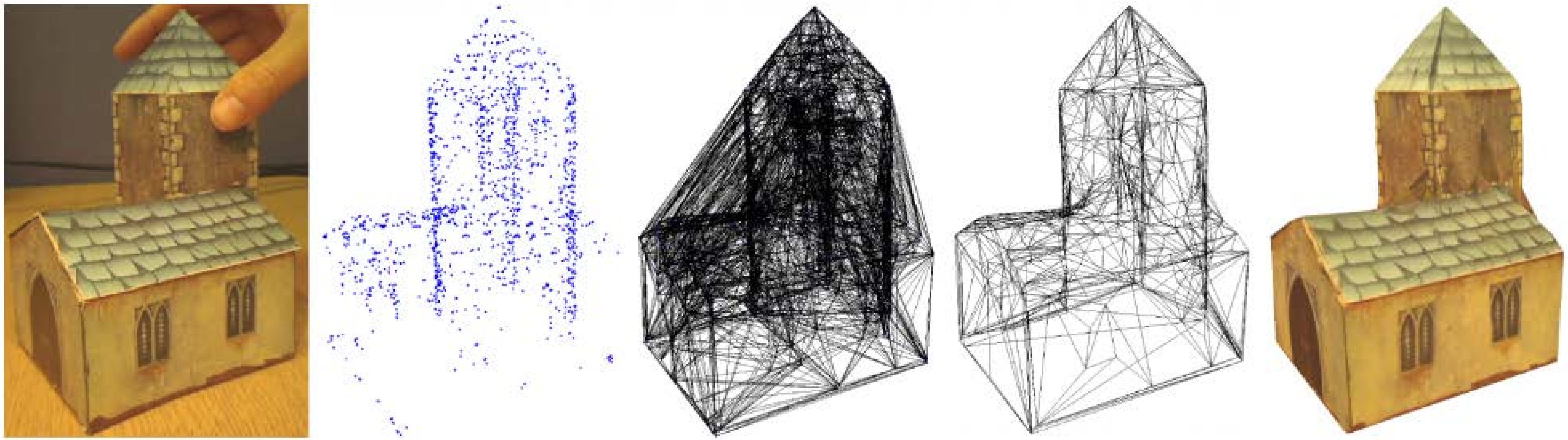
- Exploit the known imaging geometry
  - Remove surfaces that we see through



Qi Pan, Gerhard Reitmayr, Tom Drummond, ProFORMA: Probabilistic Feature-based On-line Rapid Model Acquisition, BMVC 2009.

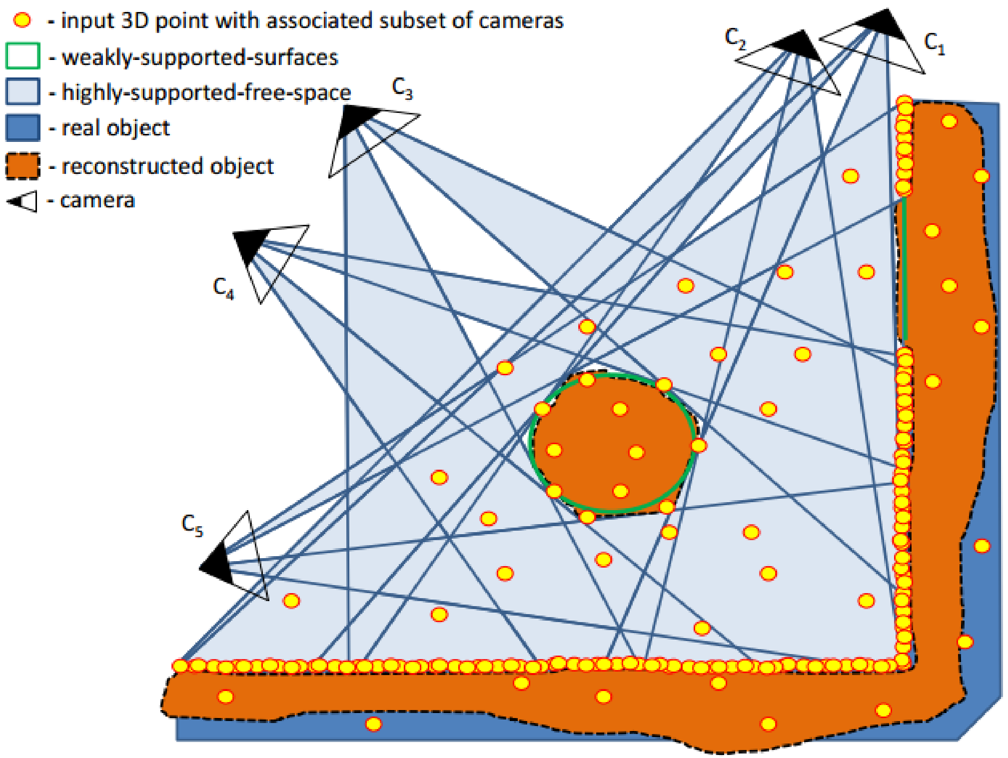
Surface reconstruction from point clouds

- Exploit the known imaging geometry
  - Remove surfaces that we see through
  - A kind of visual hull and space carving for point clouds



Qi Pan, Gerhard Reitmayr, Tom Drummond, ProFORMA: Probabilistic Feature-based On-line Rapid Model Acquisition, BMVC 2009.

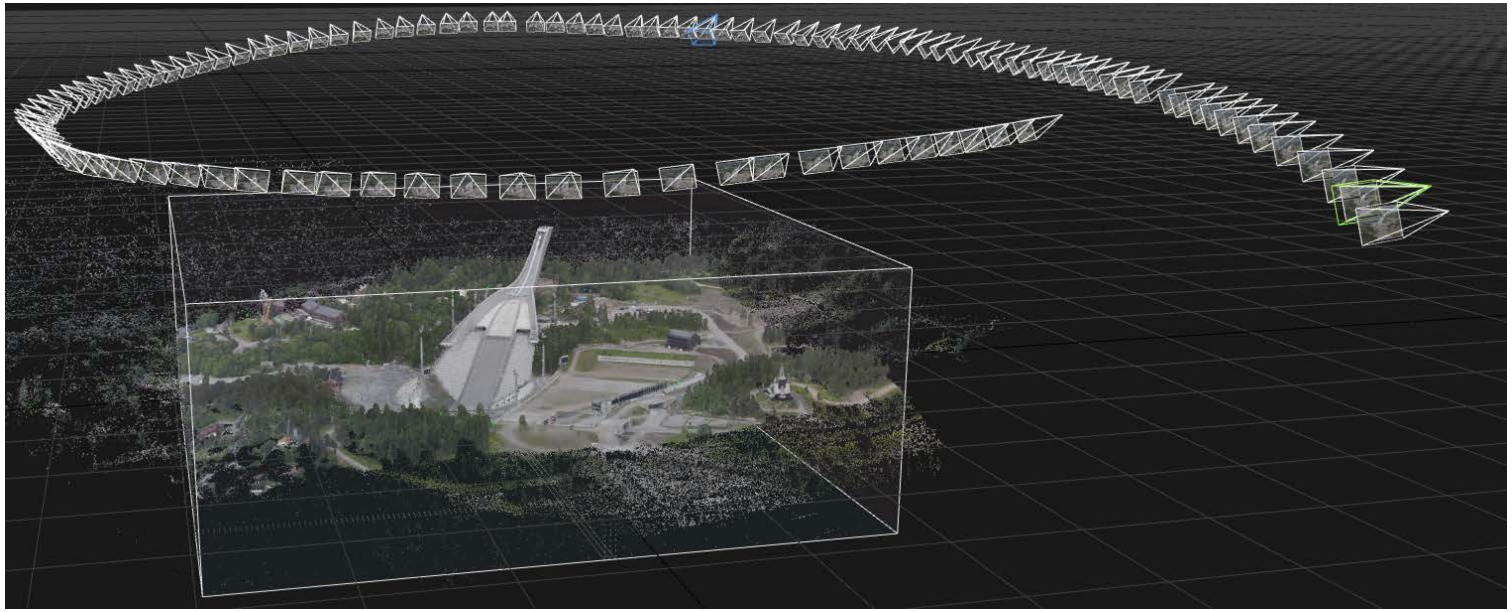
Surface reconstruction from point clouds



Michal Jancosek and Tomas Pajdla, Multi-View Reconstruction Preserving Weakly-Supported Surfaces, CVPR 2011.


Capturing Reality

Summary



- Multi-view stereo
  - Plane-sweep
  - Volumetric stereo
  - Surface expansion

- Surface reconstruction

- Additional reading
  - Szeliski 11.1.2, 11.6

- Multi-view stereo software
  - Patch-based Multi-view Stereo (PVMS)