Automatic Photo-to-Terrain Alignment for the Annotation of Mountain Pictures

Lionel Baboud*, Martin Čadík*,
Elmar Eisemann#, Hans-Peter Seidel*

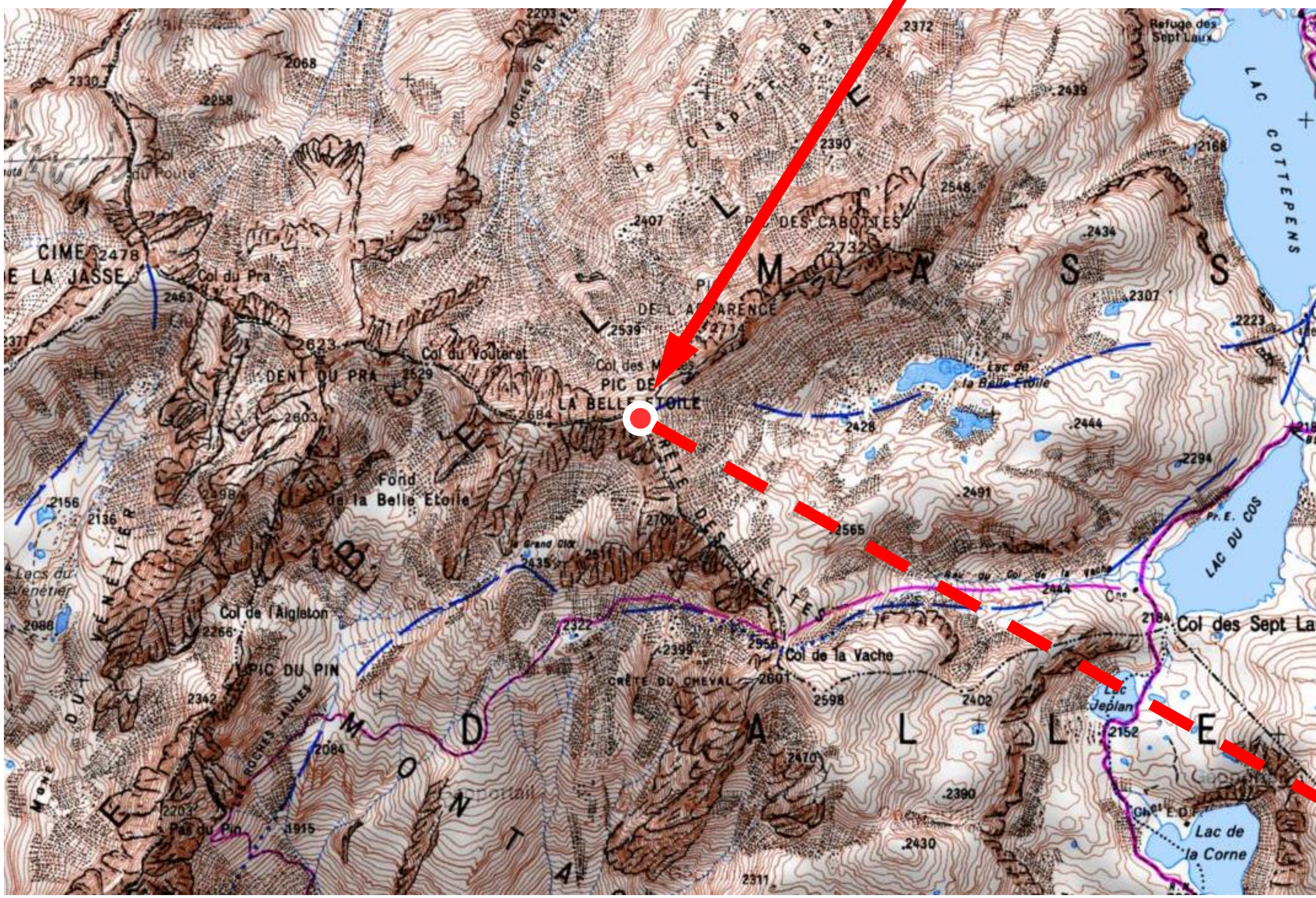
*Max-Planck Institute Informatik
#Telecom ParisTech/CNRS-LTCI

Motivation

Motivation

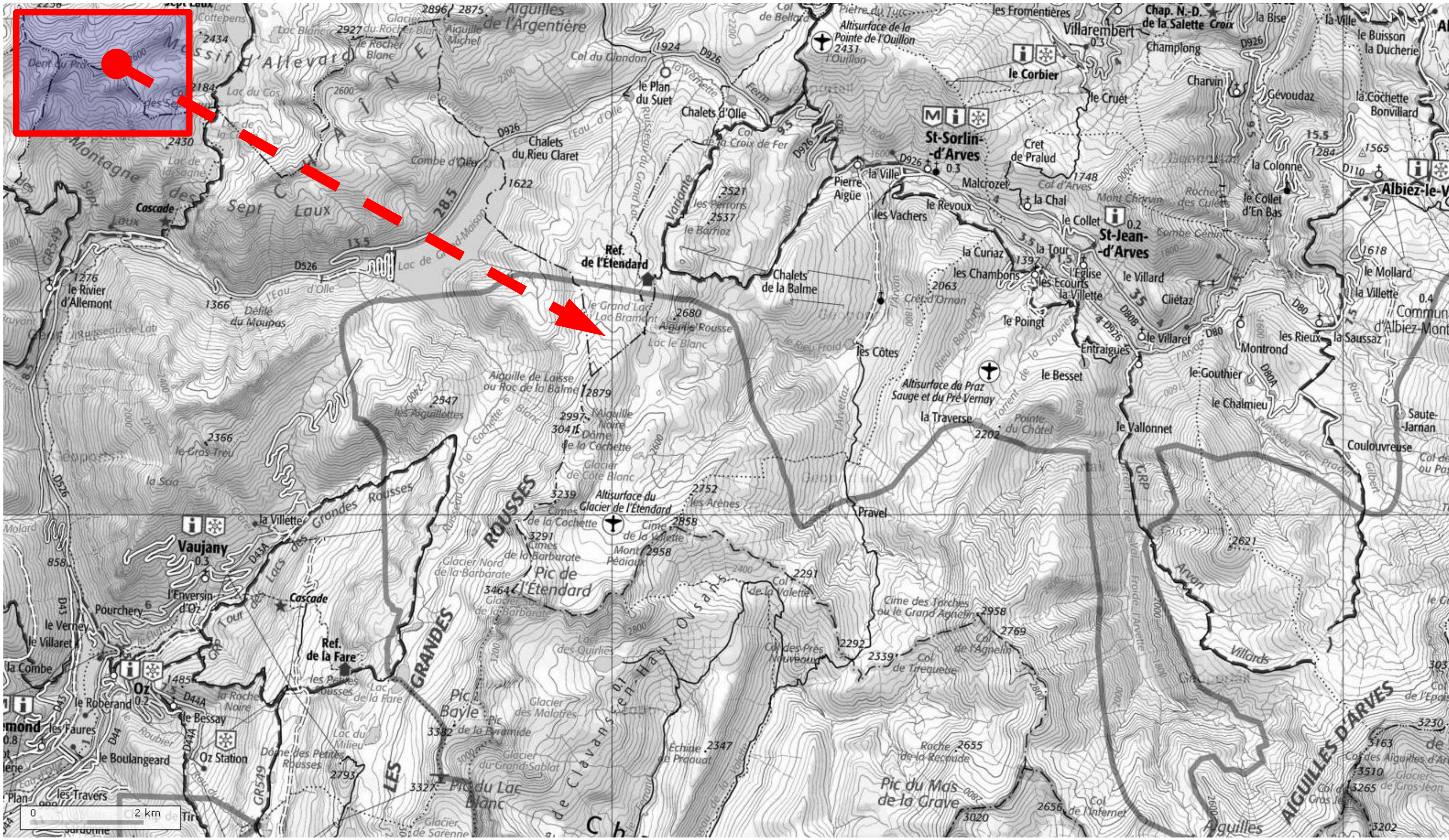
Using topographic maps

photograph's viewpoint

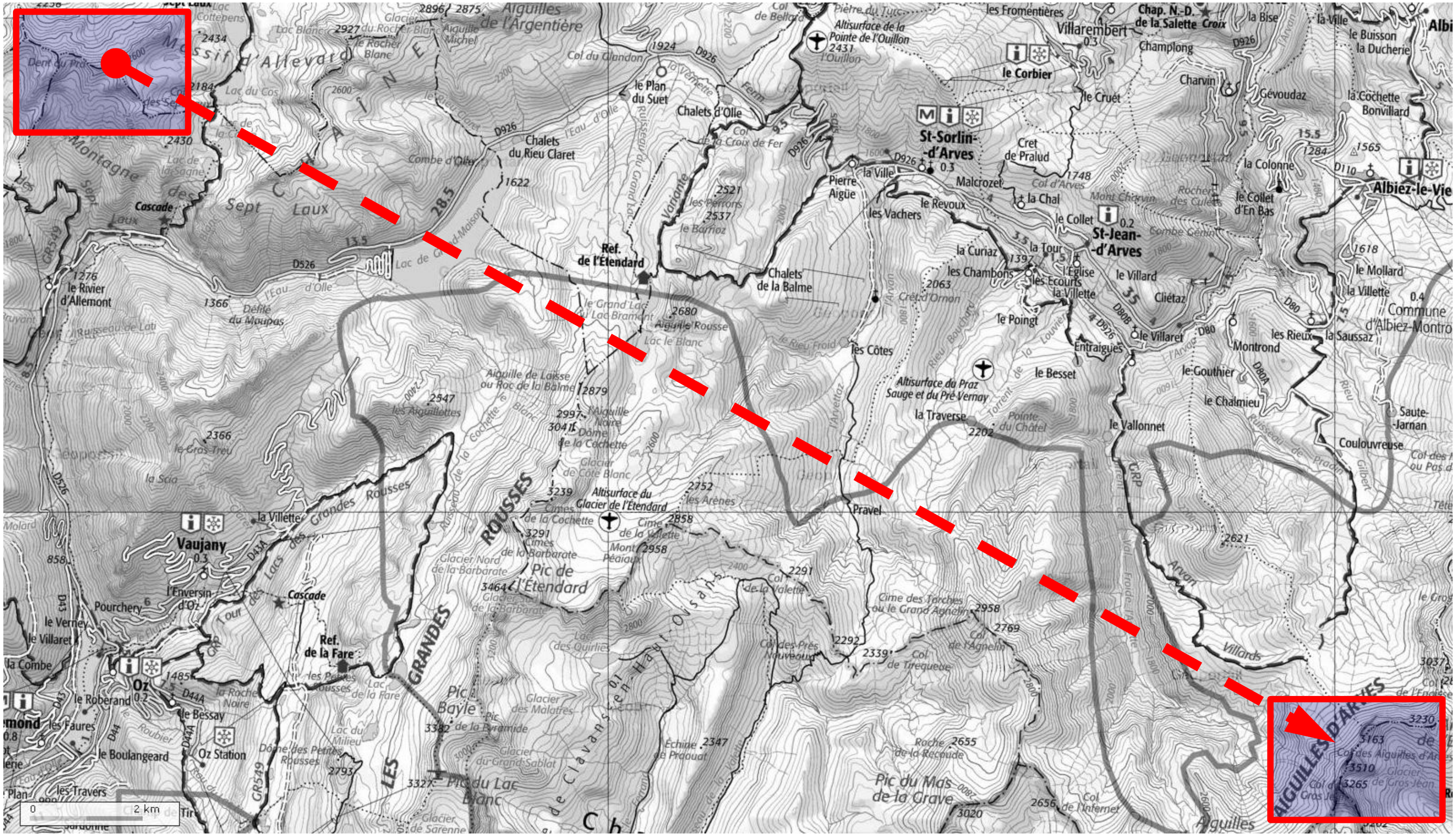


viewing direction

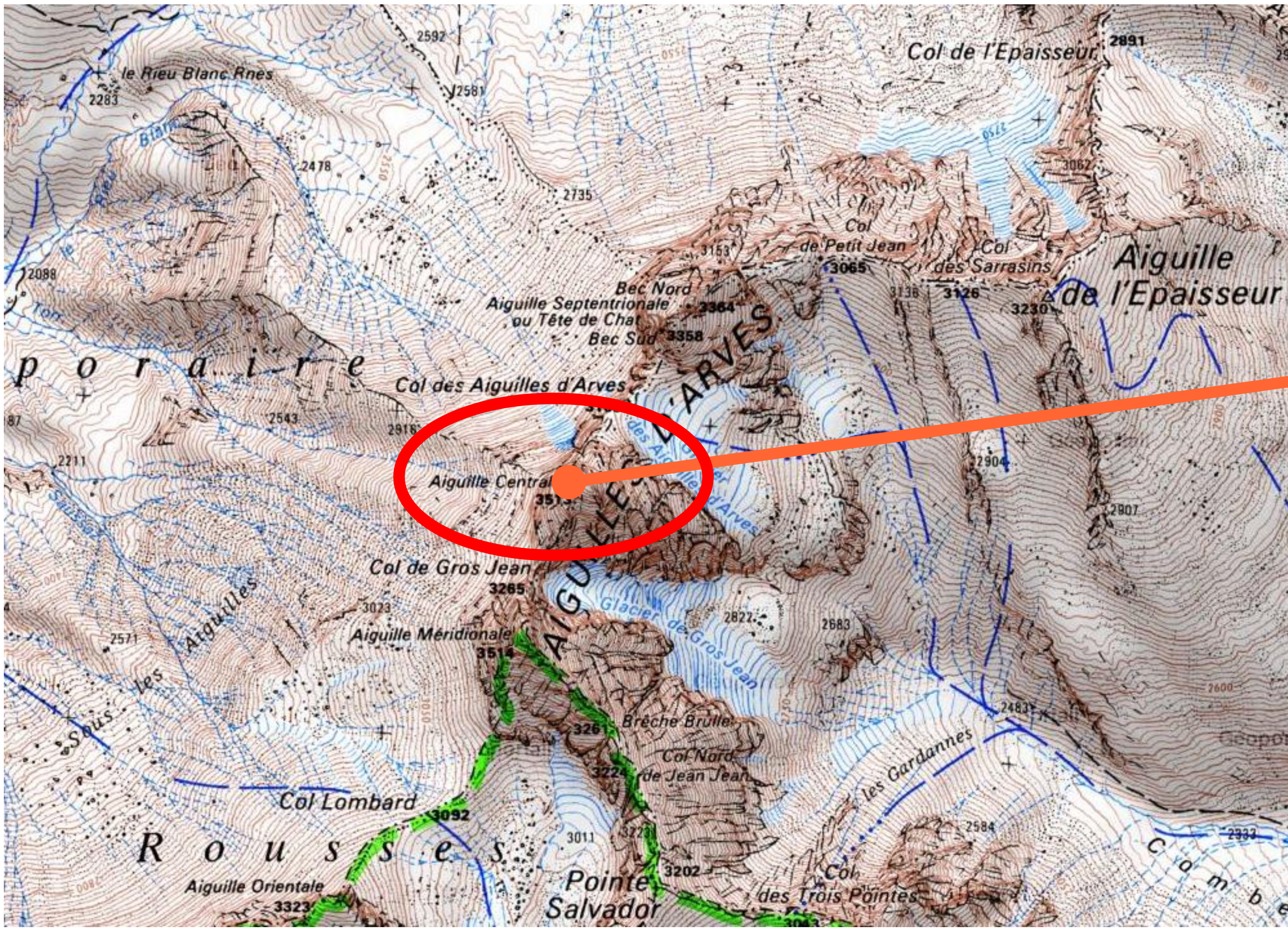
Using topographic maps



Using topographic maps



Using topographic maps



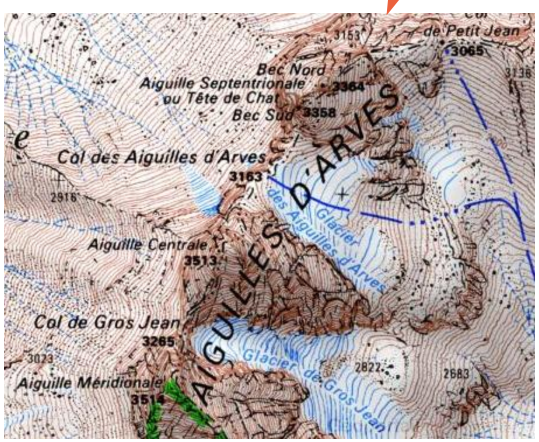
Aiguilles d'Arves
(Aiguille Centrale)
3513 meters

Using topographic maps

- Tedious task
- Difficulties
  1. Quantity of data to scan
  2. Topographical representation $\neq$ visual aspect

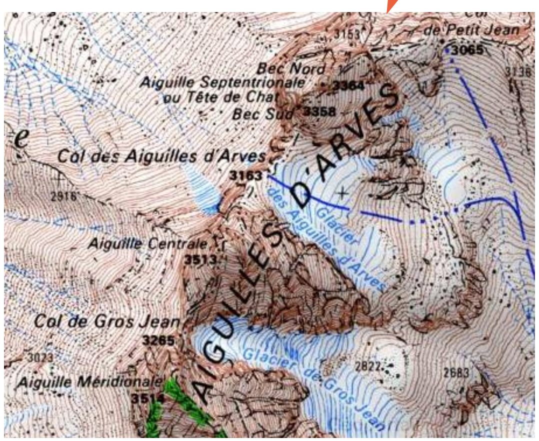
Using topographic maps

- Tedious task
- Difficulties
  1. Quantity of data to scan
  2. Topographical representation ≠ visual aspect



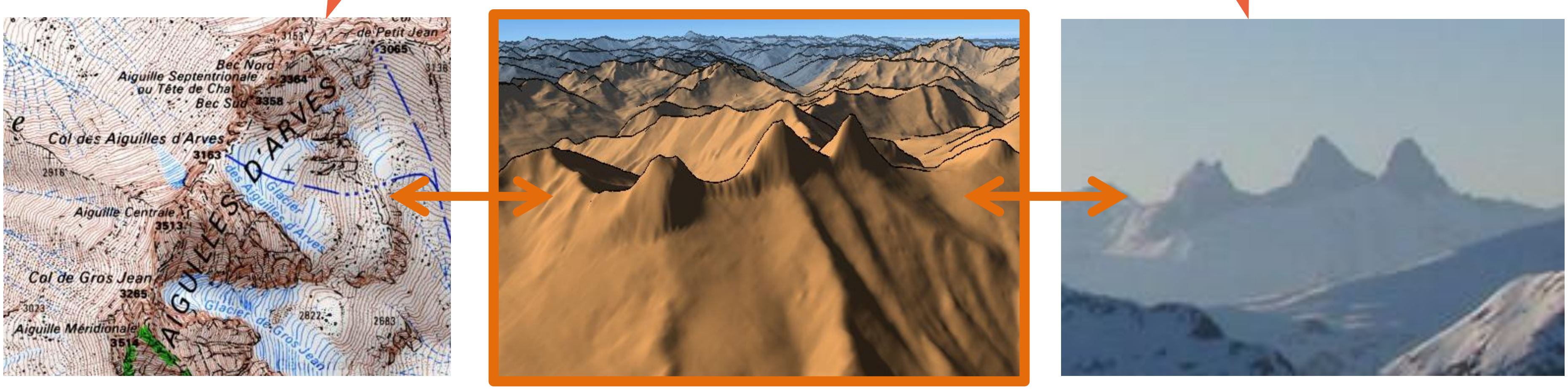
Using topographic maps

- Tedious task
- Difficulties
  1. Quantity of data to scan
  2. Topographical representation ≠ visual aspect



Using topographic maps

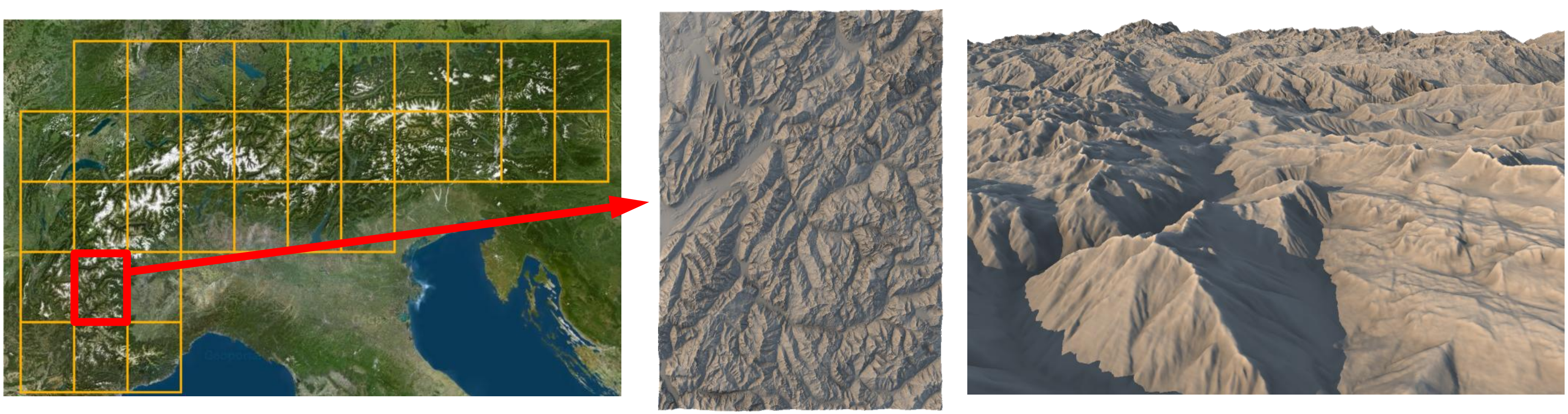
- Tedious task
- Difficulties
  1. Quantity of data to scan
  2. Topographical representation ≠ visual aspect



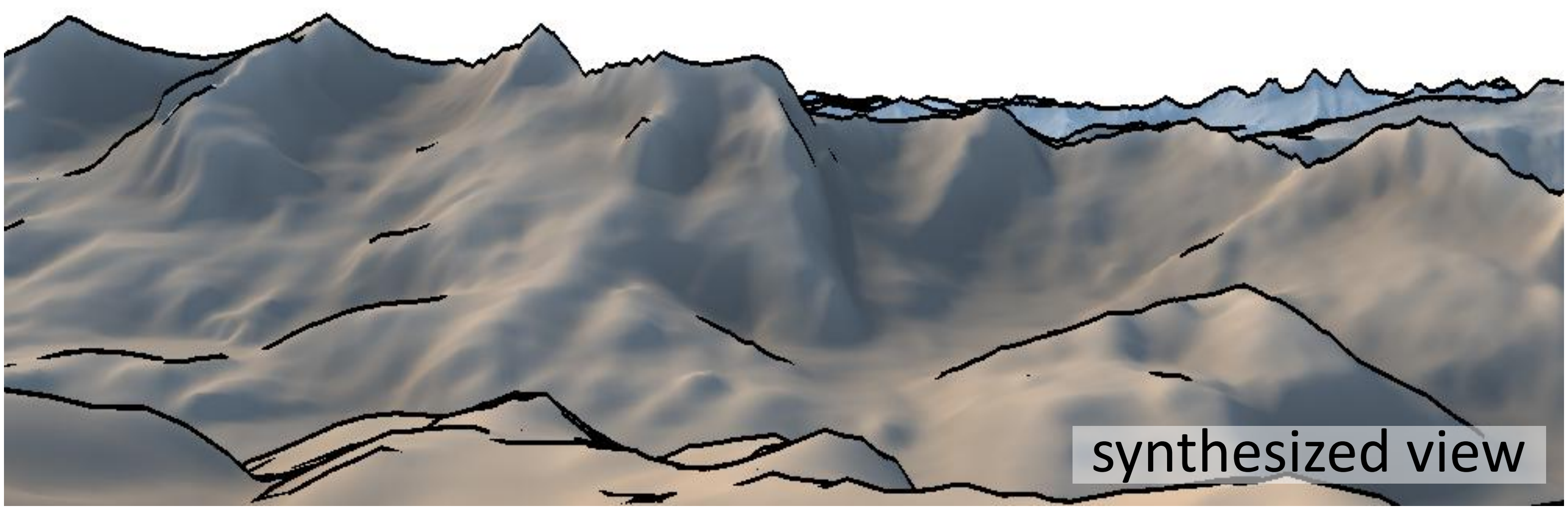
Identifying peaks on a 3D map is way easier

Available elevation data

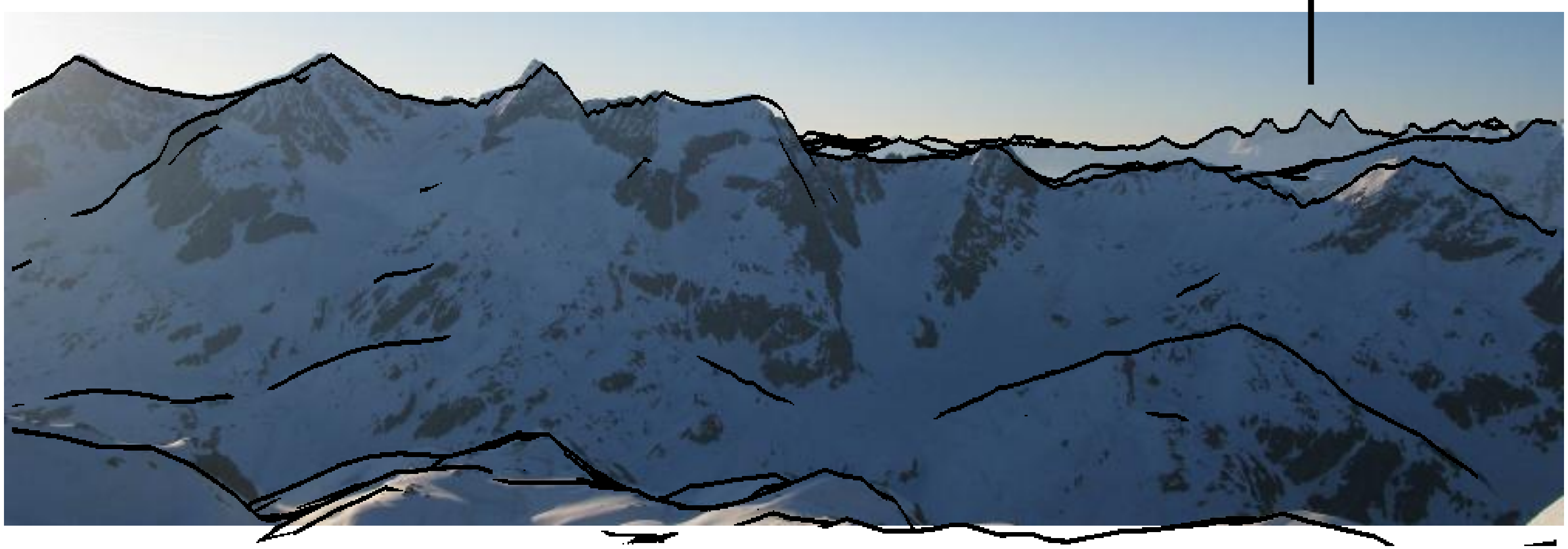
- High resolution elevation maps
  - Alps : SRTM data (NASA), ~25m resolution
  - Rockies (Colorado, USA) : USGS, ~8m resolution



Available elevation data



Available elevation data



Matching is pretty accurate,
now can we compute it automatically?

Problem statement

- Problem = camera pose estimation
- Camera parameters
  - Intrinsic (FOV, etc.)
  - Extrinsic (position, orientation)

Existing approches

- Photogrammetric features

  SIFT **[Lowe 04]**, pano stitching **[Szeliski 06]**,
  photo-tourism **[Snavely et al. 06]**, etc.

➔ **Problematic for outdoor, highly varying environments:**

Existing approches

- Photogrammetric features

  SIFT **[Lowe 04]**, pano stitching **[Szeliski 06]**, photo-tourism **[Snavely et al. 06]**, etc.

- Specific features:

  horizon line curve **[Woo et al. 07]**, peaks **[Mukunori et al. 97]**

Existing approches

- Photogrammetric features

  SIFT **[Lowe 04]**, pano stitching **[Szeliski 06]**, photo-tourism **[Snavely et al. 06]**, etc.

- Specific features:

  horizon line curve **[Woo et al. 07]**, peaks **[Mukunori et al. 97]**

- Manual registration

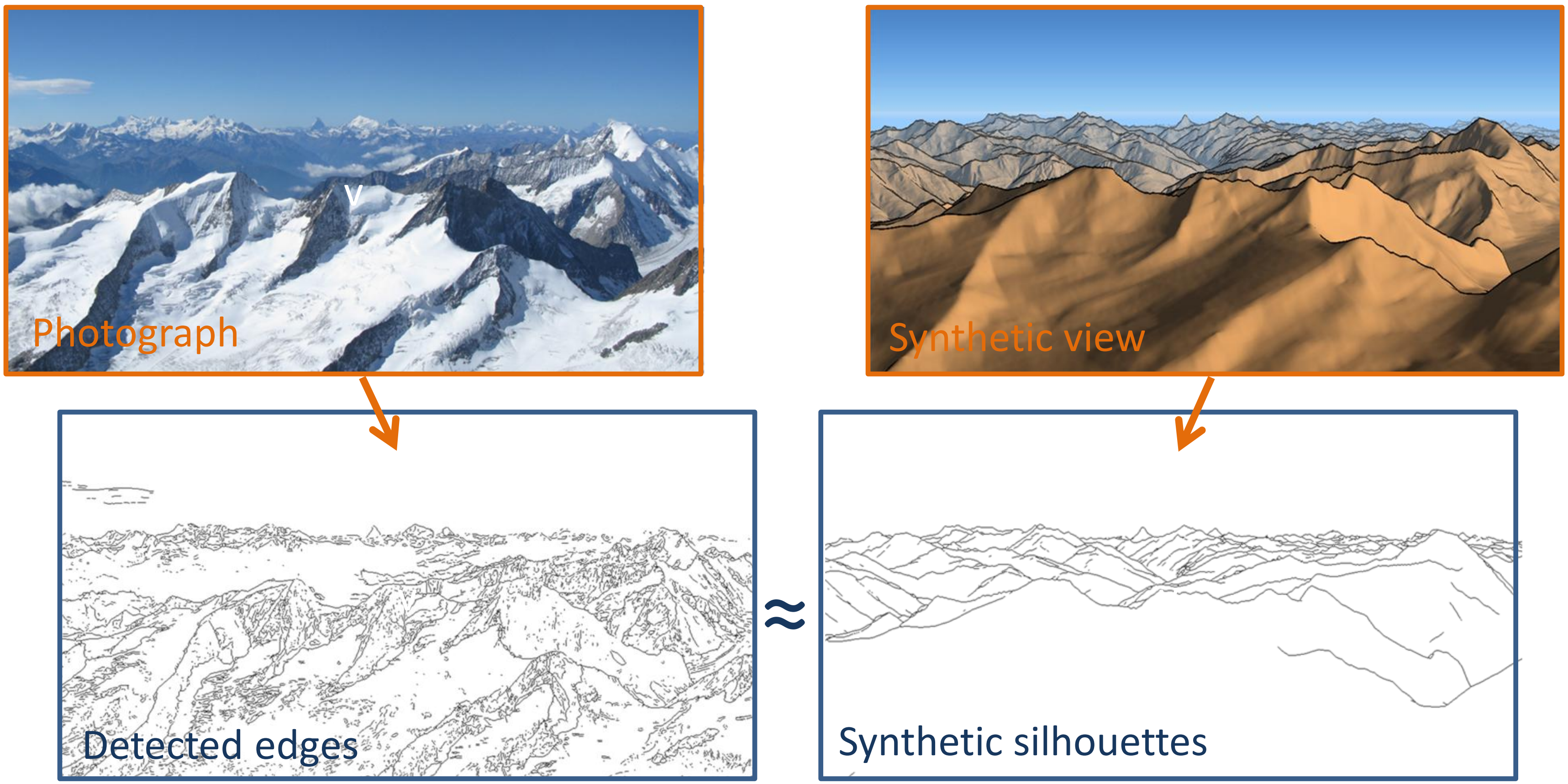
  [http://flpsed.org/gipfel.html]

Which features to rely on?

- Visual variations in mountain scenes
  - Season (snow, grass, trees, sheep, etc.)
  - Lighting (sun position, shadows, etc.)
  - Weather (clouds, atmospheric scattering, etc.)

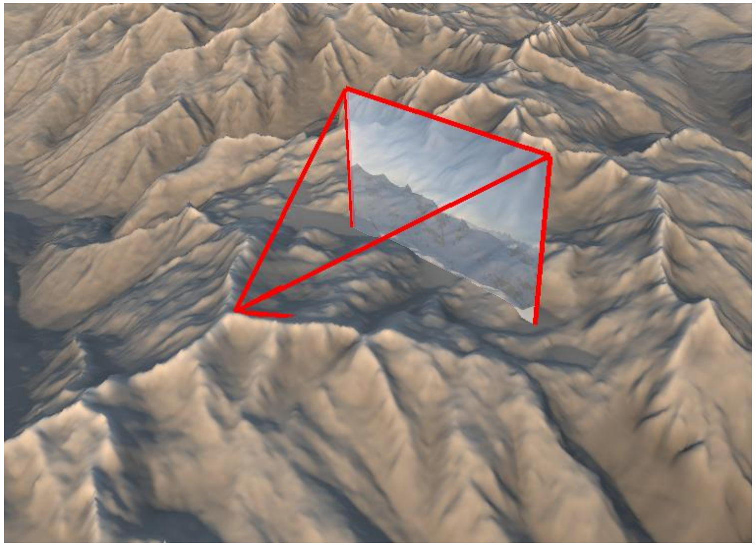
Which features to rely on?

- Robust features: silhouette edges



Assumptions

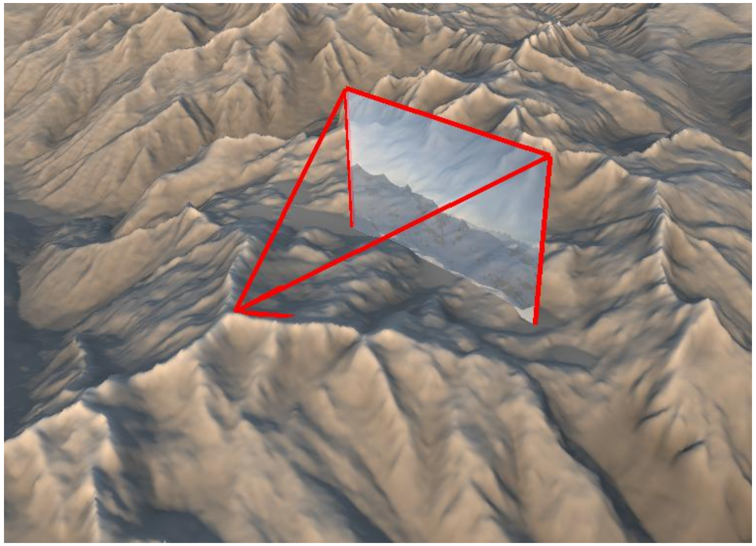
- Camera parameters
  - Intrinsic (FOV, etc.)
  - Position
  - Orientation



Assumptions

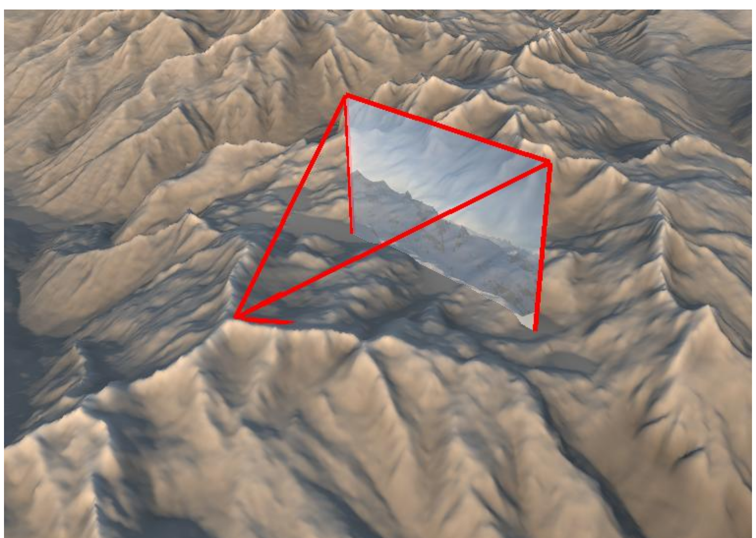
- Camera parameters
  - Intrinsic (FOV, etc.)
  - Position
  - Orientation

Read in attached data (EXIF tags) or camera database

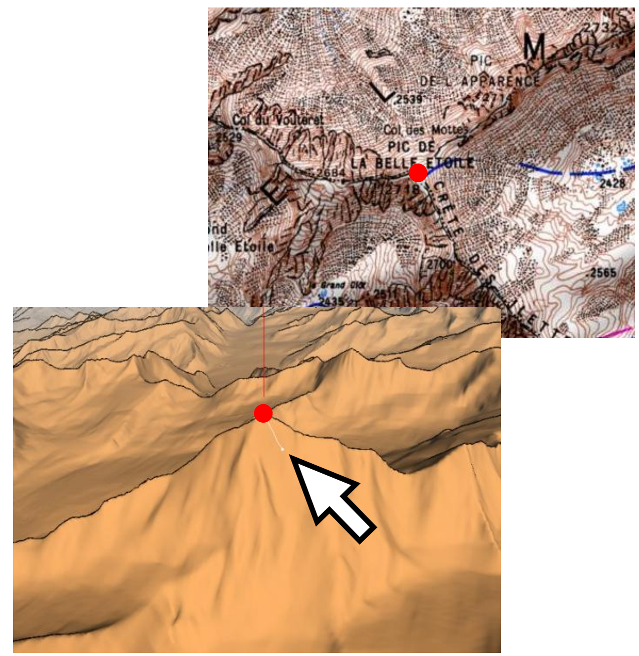


Assumptions

- Camera parameters
  - Intrinsic (FOV, etc.)
  - Position
  - Orientation

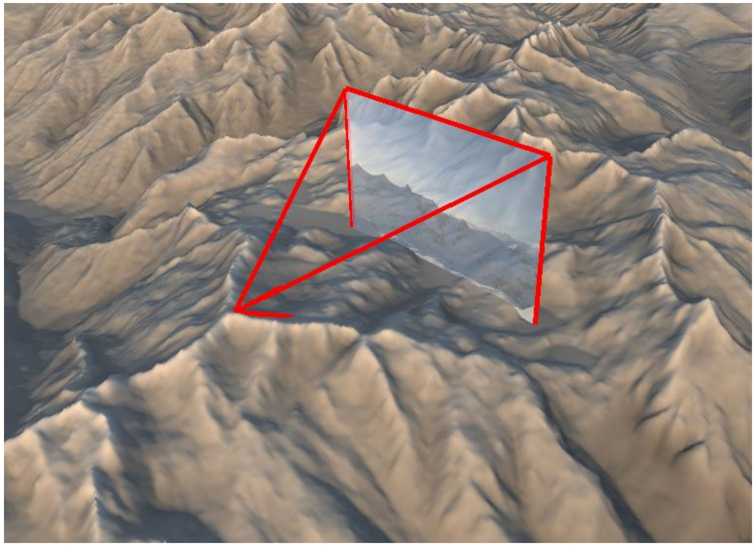


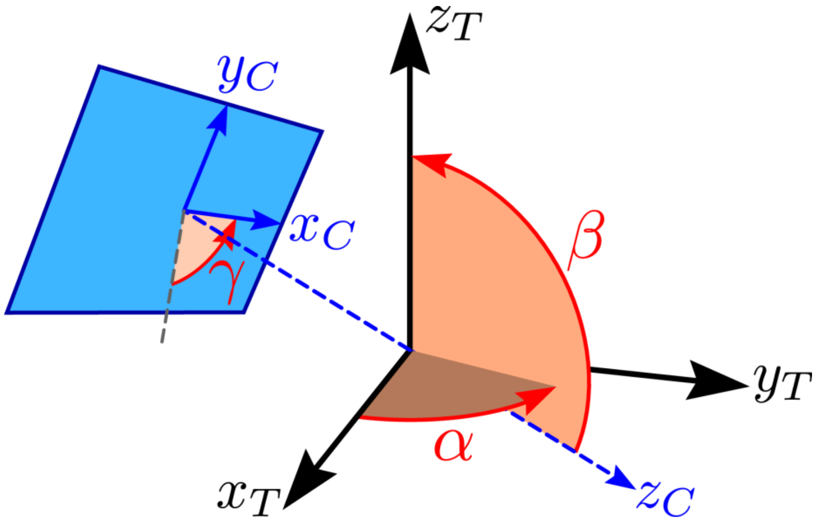
Good estimation
- GPS coordinates
- User input



Assumptions

- Camera parameters
  - Intrinsic (FOV, etc.)
  - Position
  - Orientation

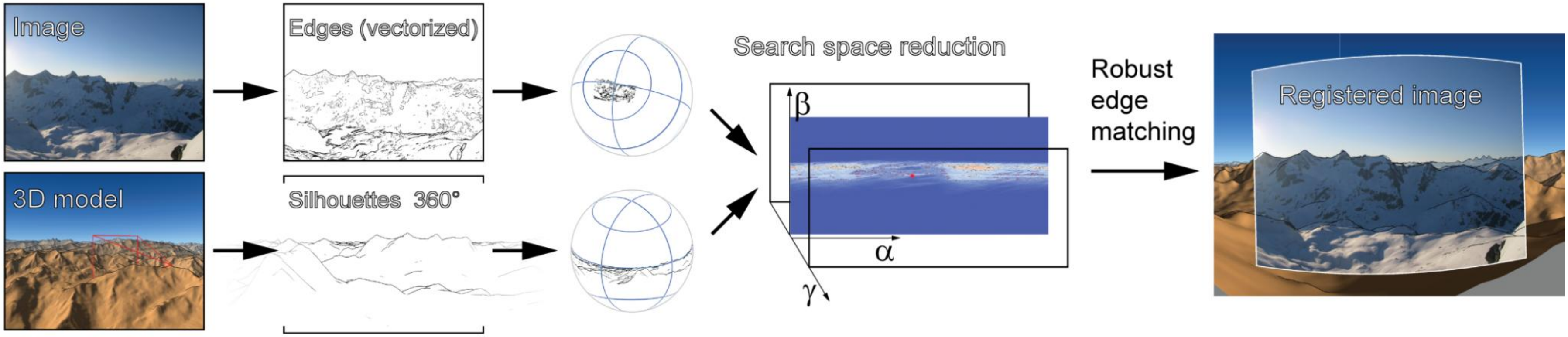




➔ **3 degrees of freedom to determine**

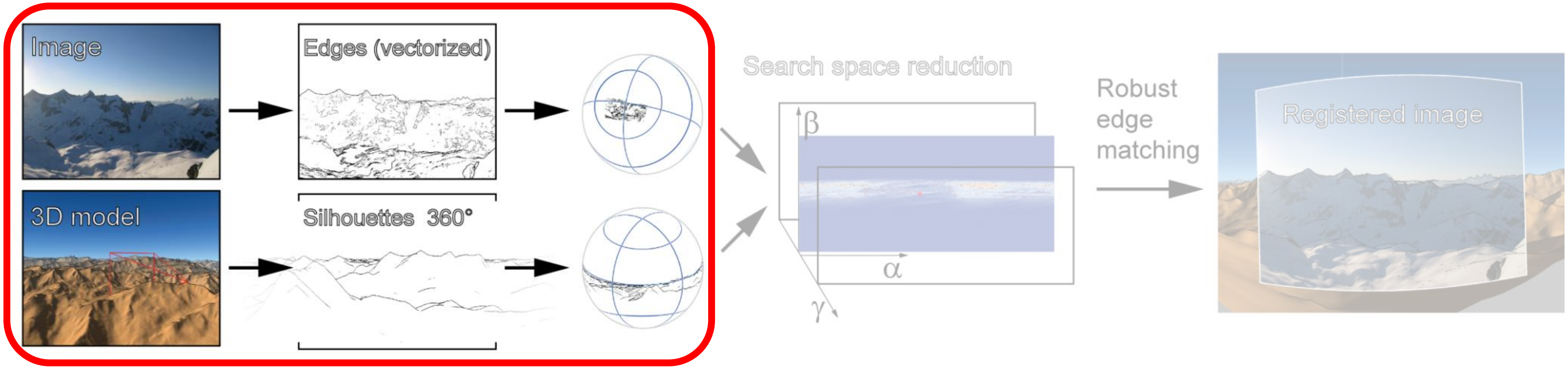
Algorithm

1. Inputs generation
   - Edge detection
   - Panorama synthesis
2. Matching
   - Search space reduction
   - Robust matching
3. Post-processing: annotation



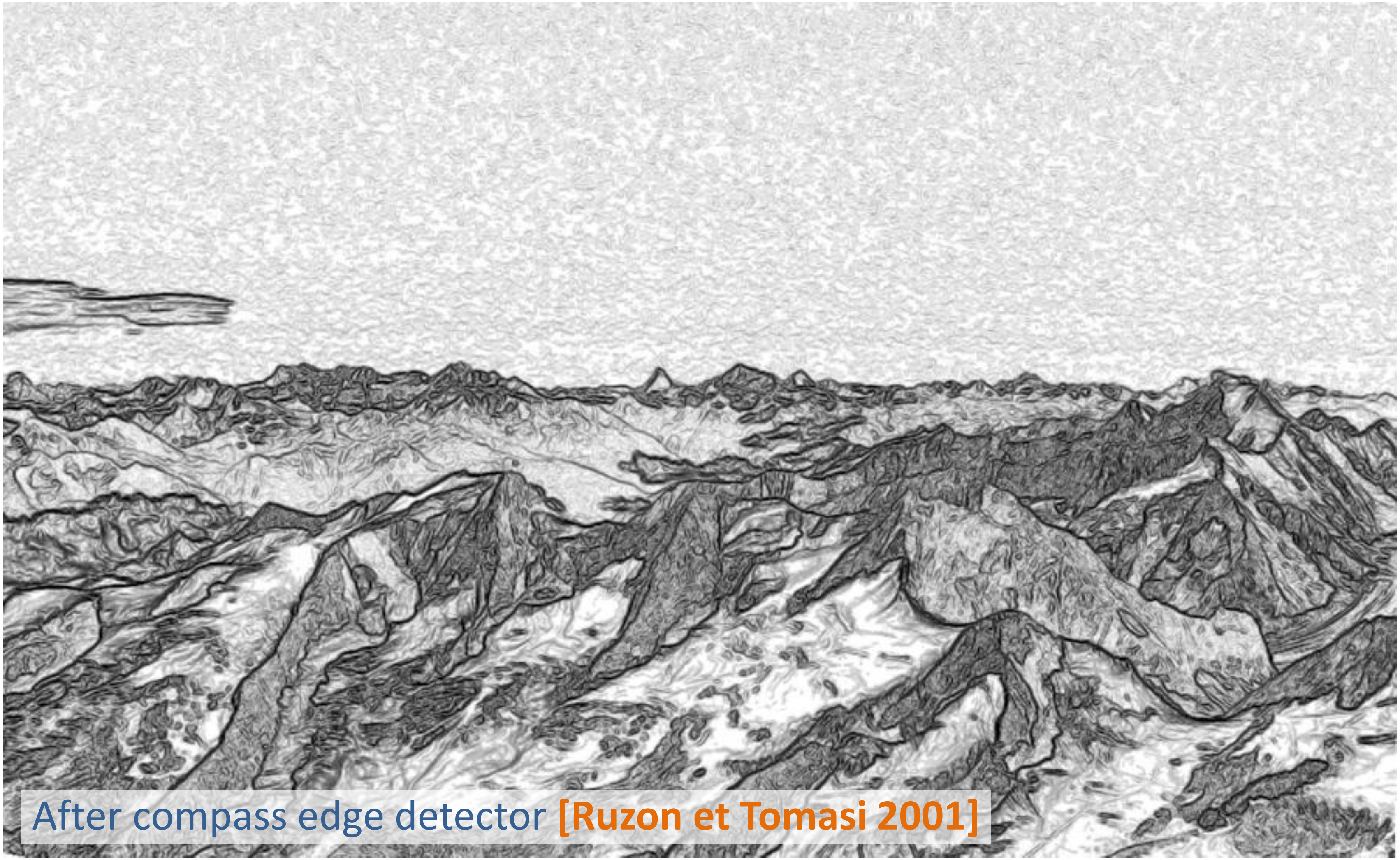
Algorithm

1. Inputs generation
   - Edge detection
   - Panorama synthesis
2. Matching
   - Search space reduction
   - Robust matching
3. Post-processing: annotation



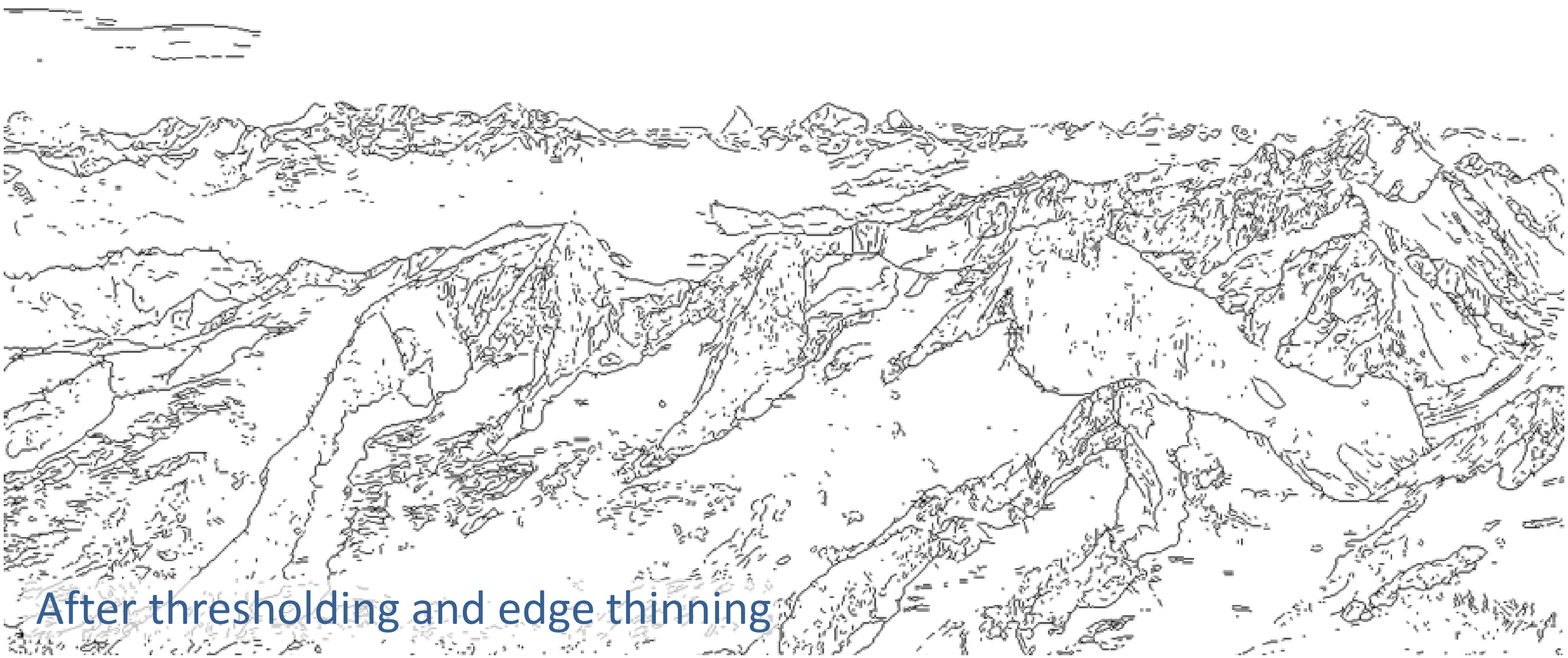
Edge detection

Edge detection



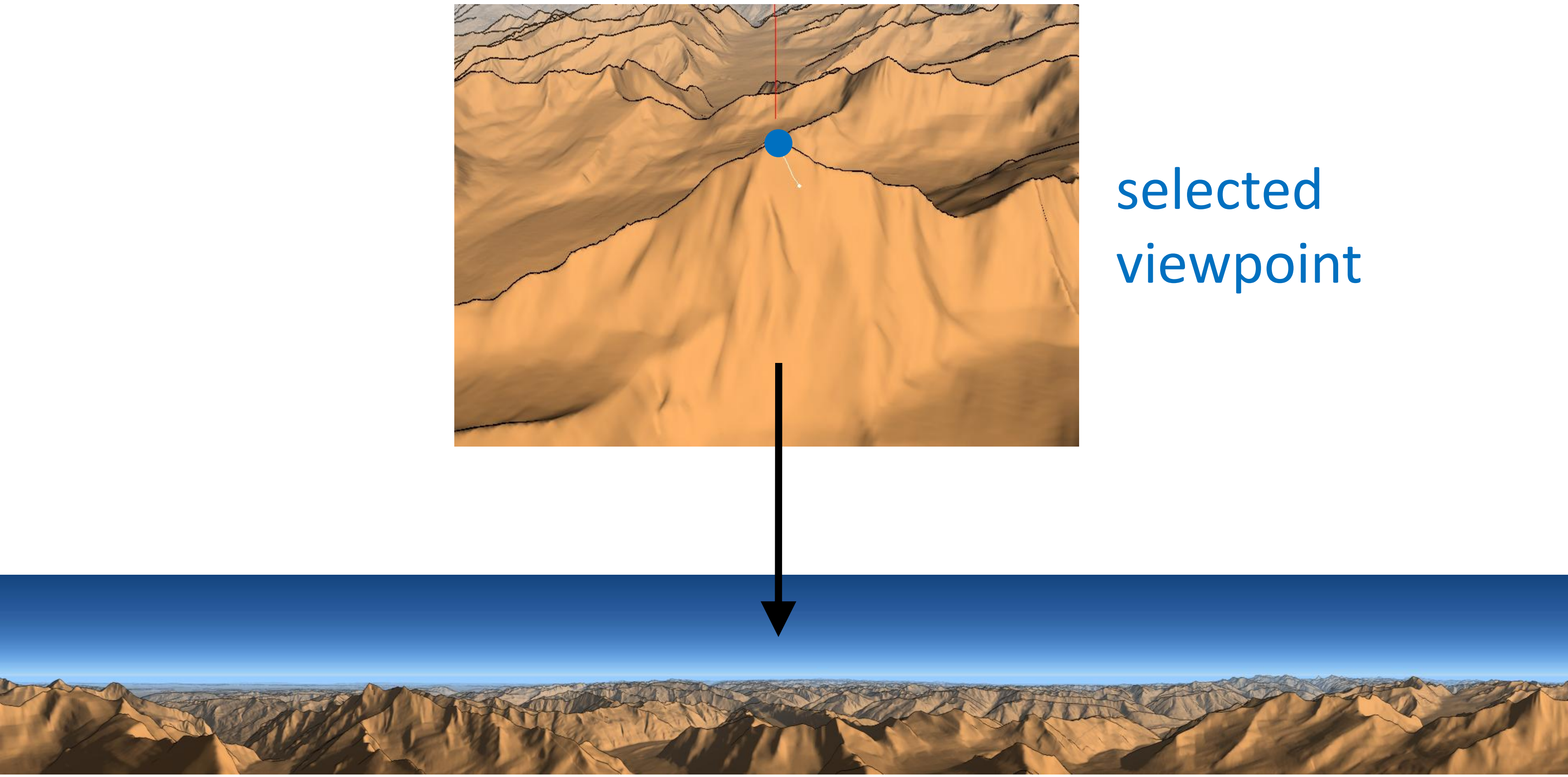
After compass edge detector [Ruzon et Tomasi 2001]

Edge detection



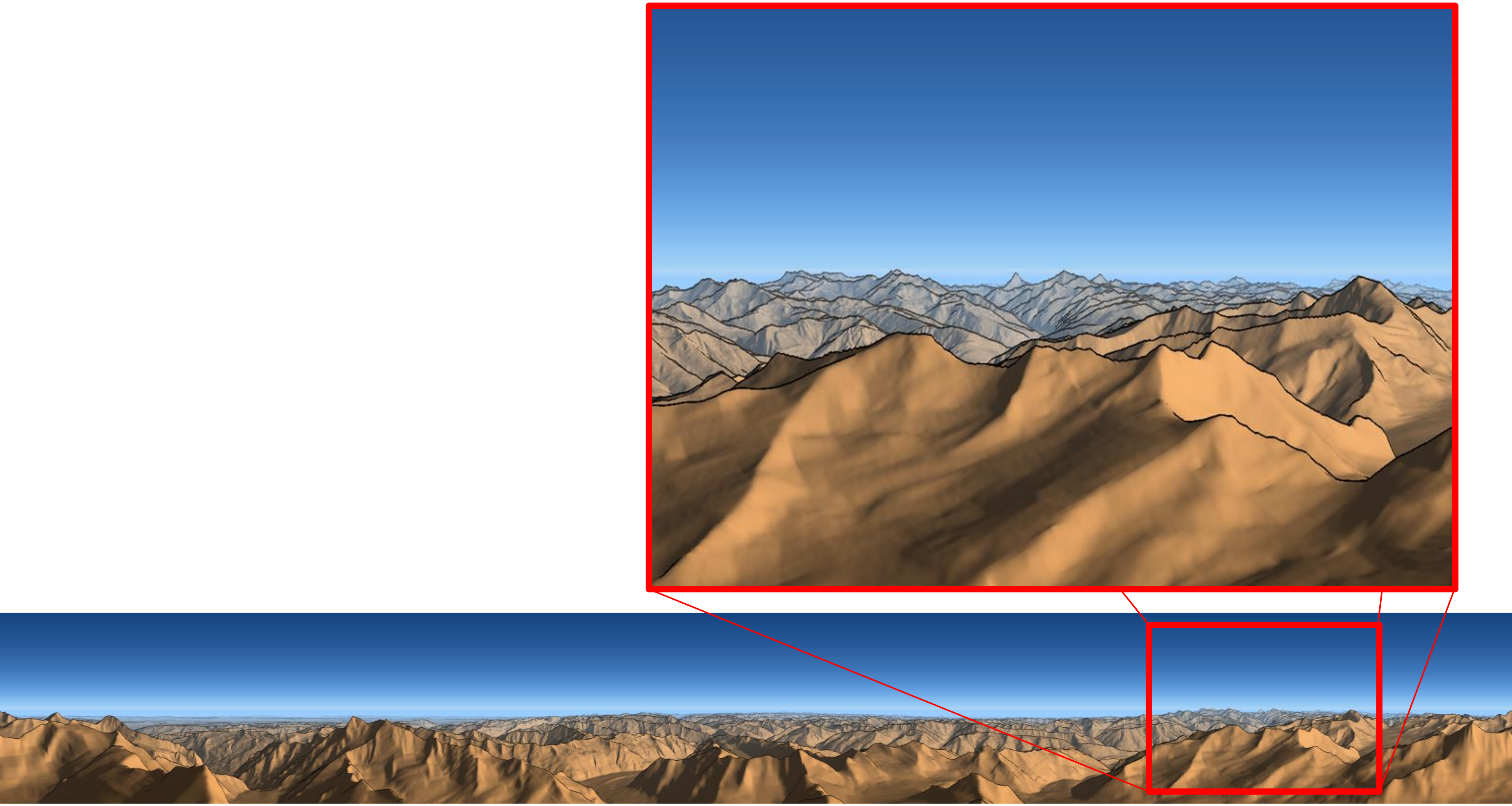
After thresholding and edge thinning

Panorama synthesis



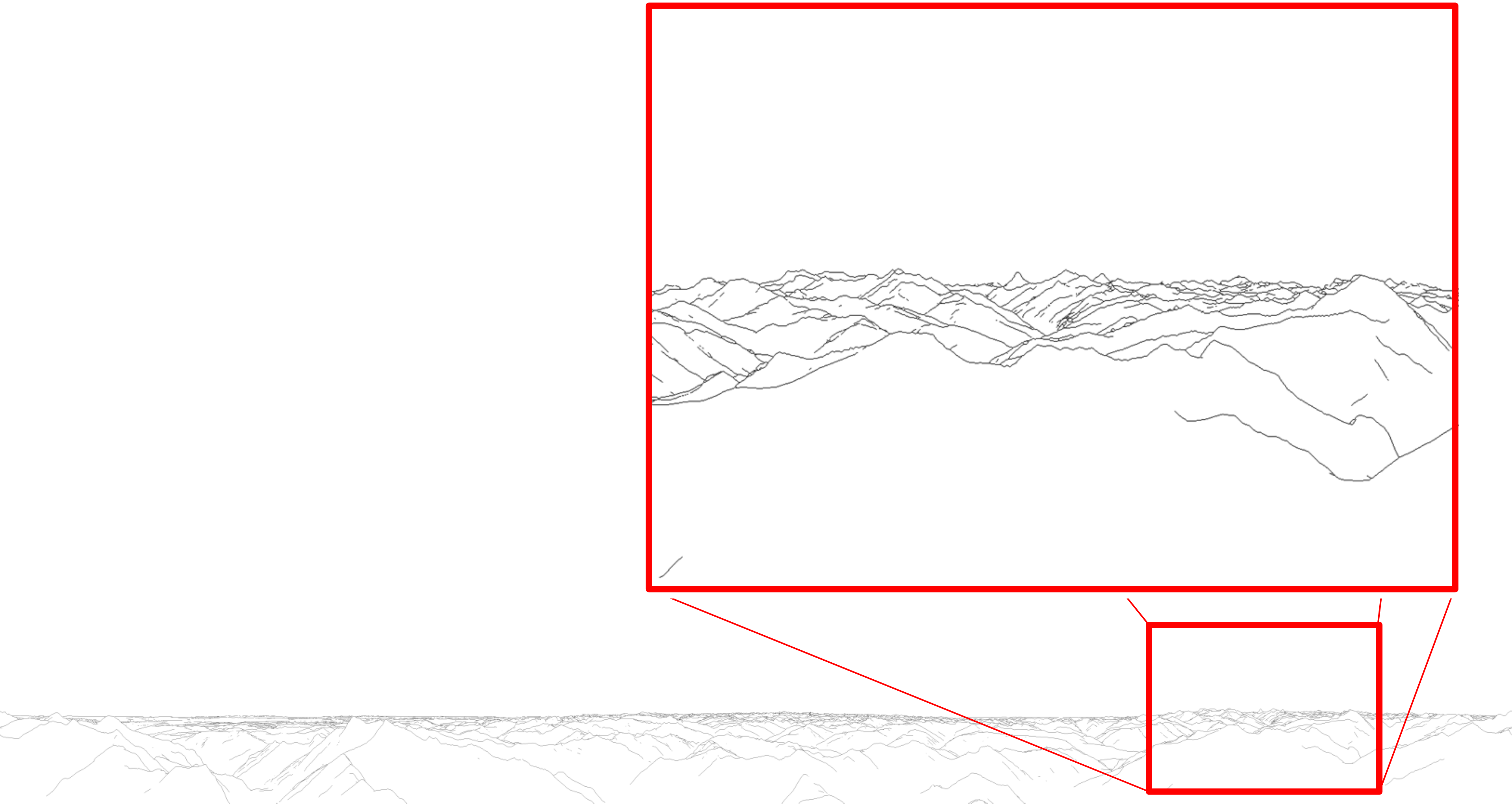
360 synthetic panorama

Panorama synthesis



360 synthetic panorama

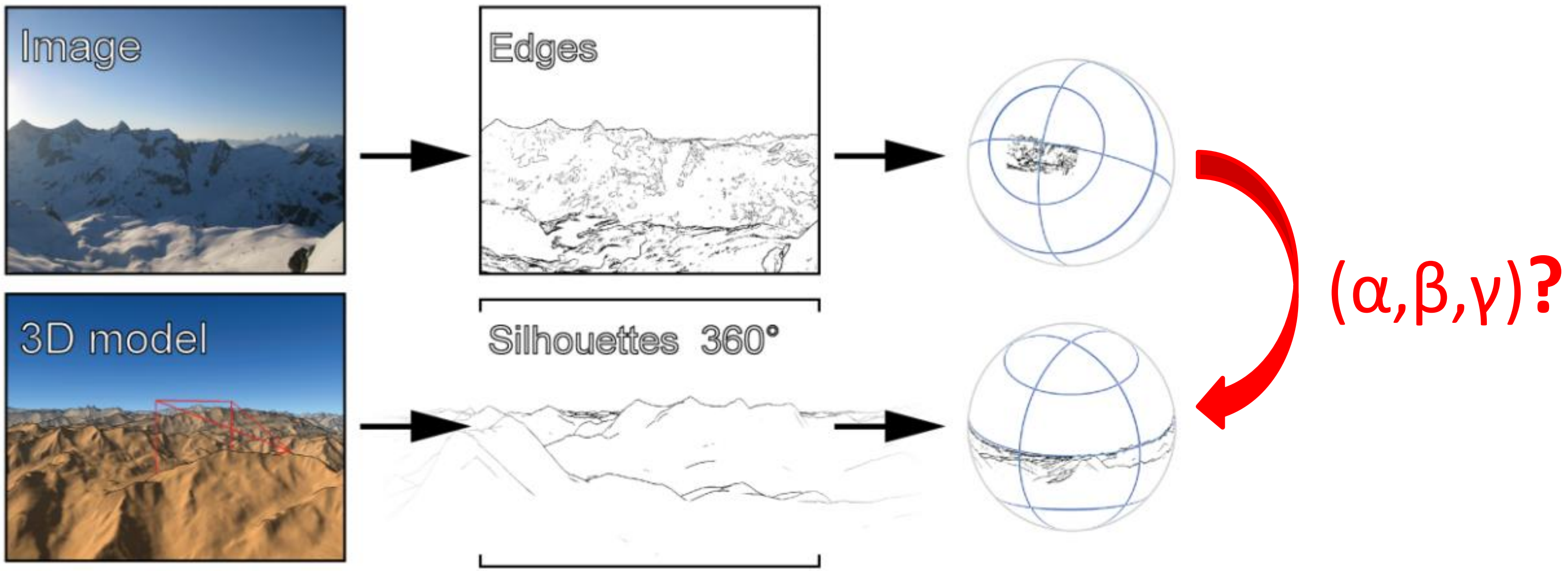
Panorama synthesis



360 synthetic panorama

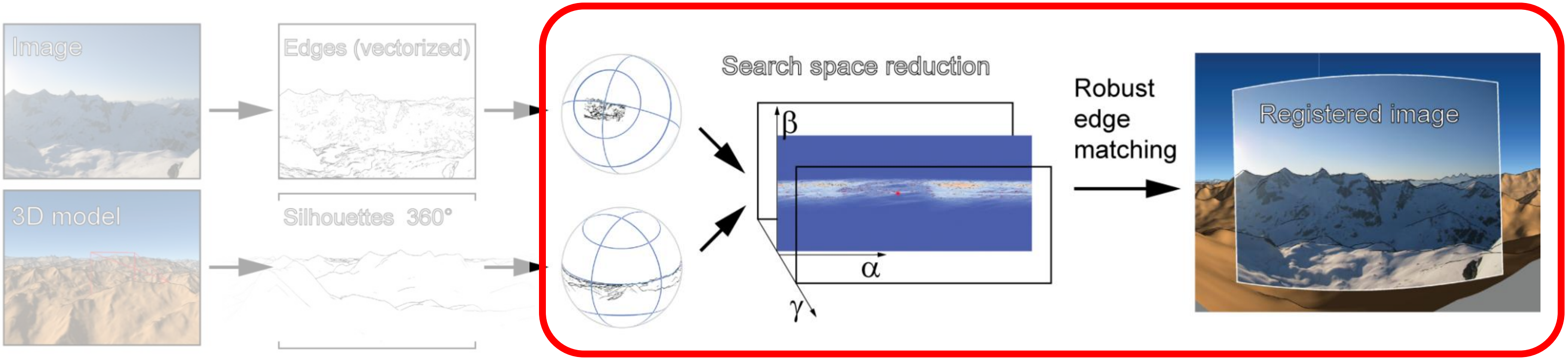
Spherical edge maps

- Spherical images
  - Unifies projection for photo / panorama



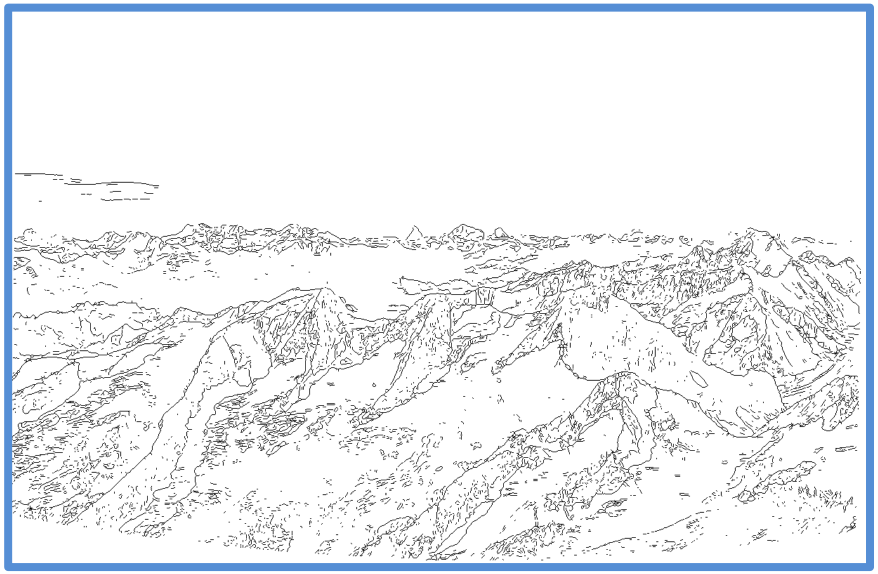
Algorithm

1. Inputs generation
   – Edge detection
   – Panorama synthesis
2. Matching
   – Search space reduction
   – Robust matching
3. Post-processing: annotation

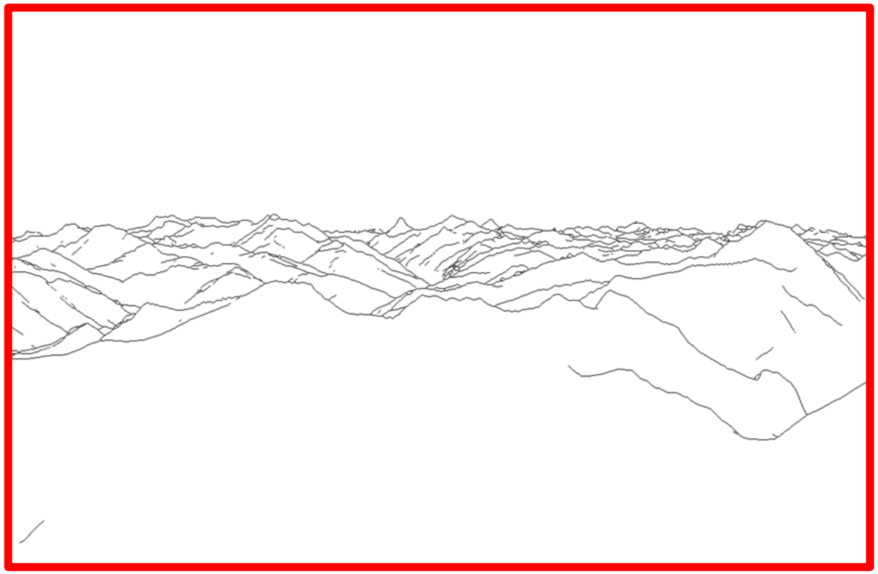


Matching

- Matching silhouette maps needs special care



≈

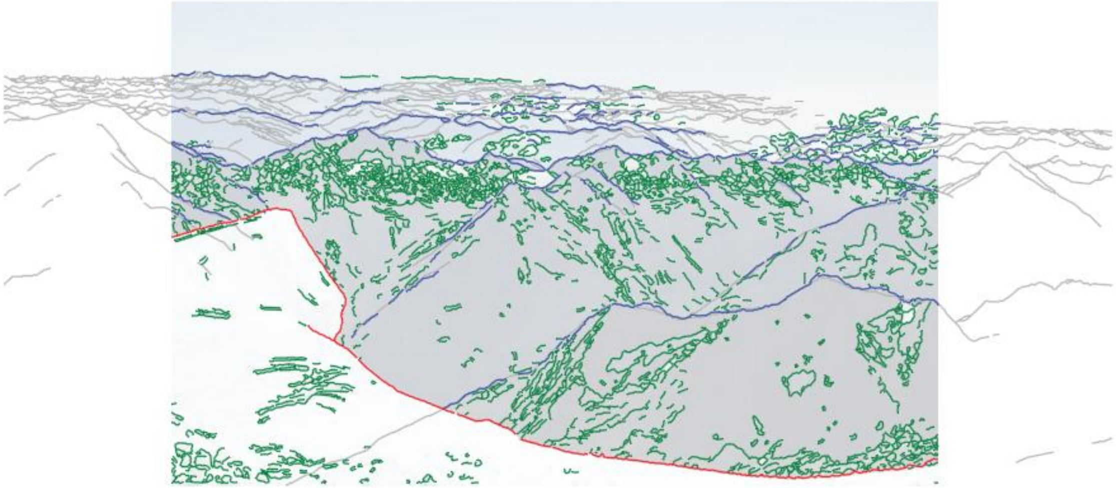


detected edges (input photograph)

theoretically matching synthetic edges

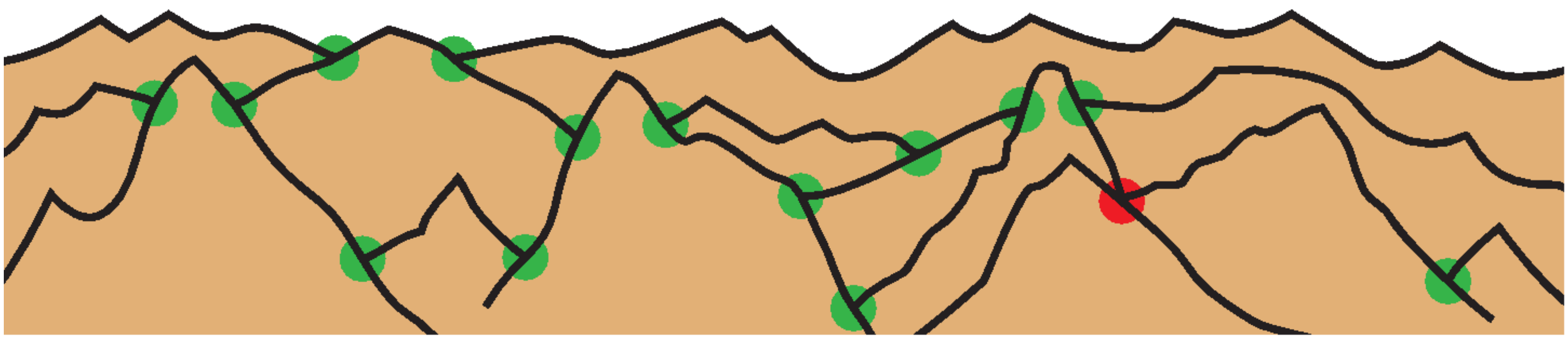
Silhouette Map Matching

- Inaccuracies in detected edges
  - Missing silhouettes
  - Non-silhouette edges, noise
  - Silhouettes but not encoded in the terrain model



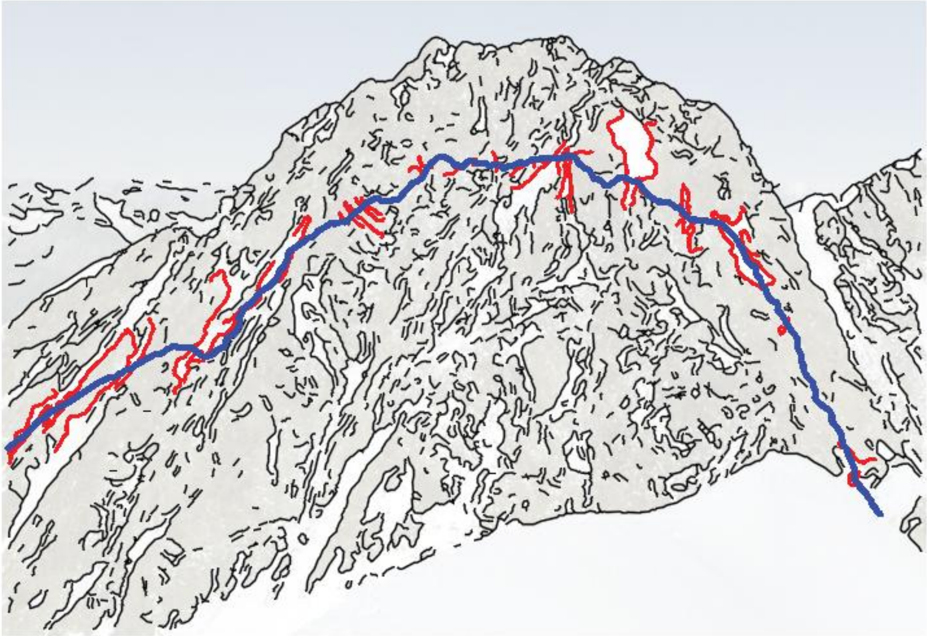
Silhouette Map Matching

- Topological properties of edge maps
  - Silhouette edges always meet in T-junctions
    (crossings are singularities)
  - Non-silhouette edges seldom cross silhouettes
    (e.g. border of forests, snow, grass, etc.)

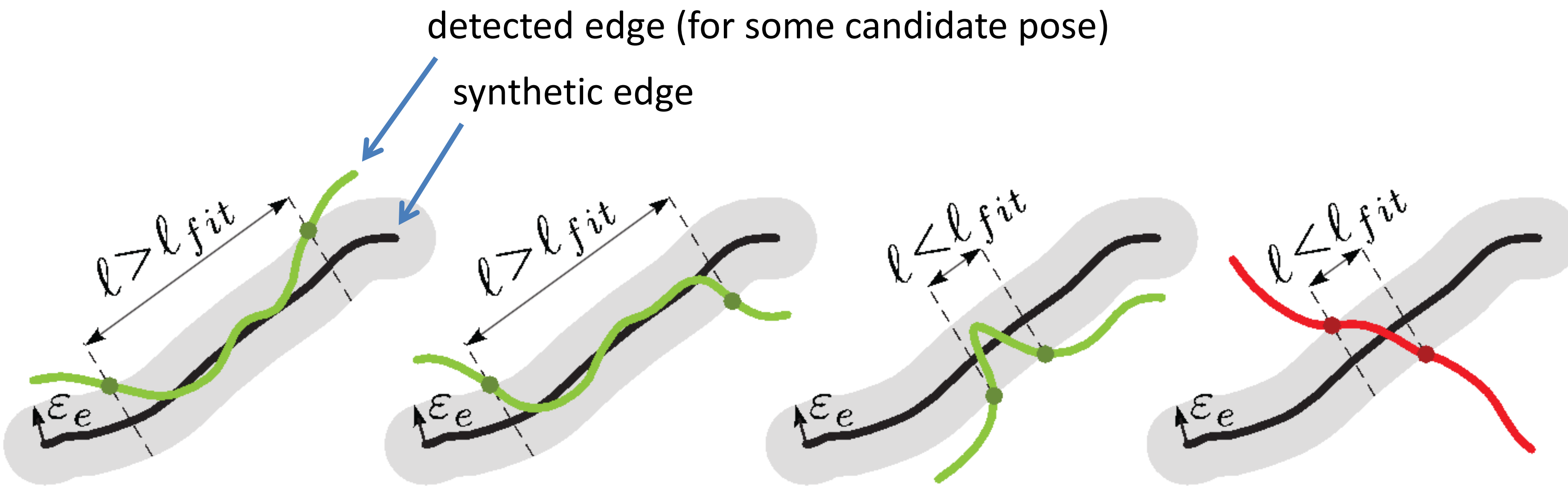


Silhouette Map Matching

- Non-silhouette edges also provide information
  → Use them for matching



Robust Matching Metric



ε_e – tolerance

ℓ_{fit} – following/crossing threshold

Robust Matching Metric

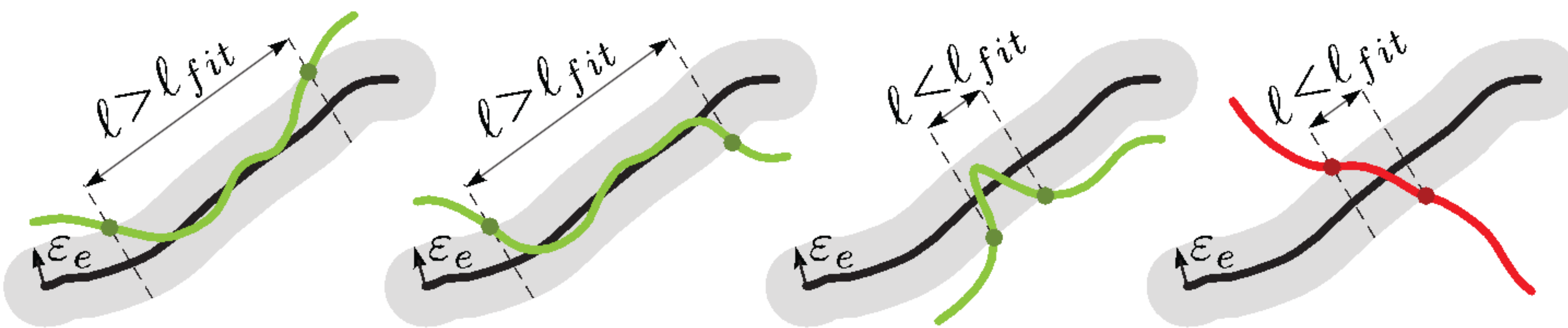
- Compute matching likelihood *E*:

foreach *edge within the* ε_e*-neighboorhood*

if ($\ell \geqslant \ell_{fit}$ **or** *exits on same side*)

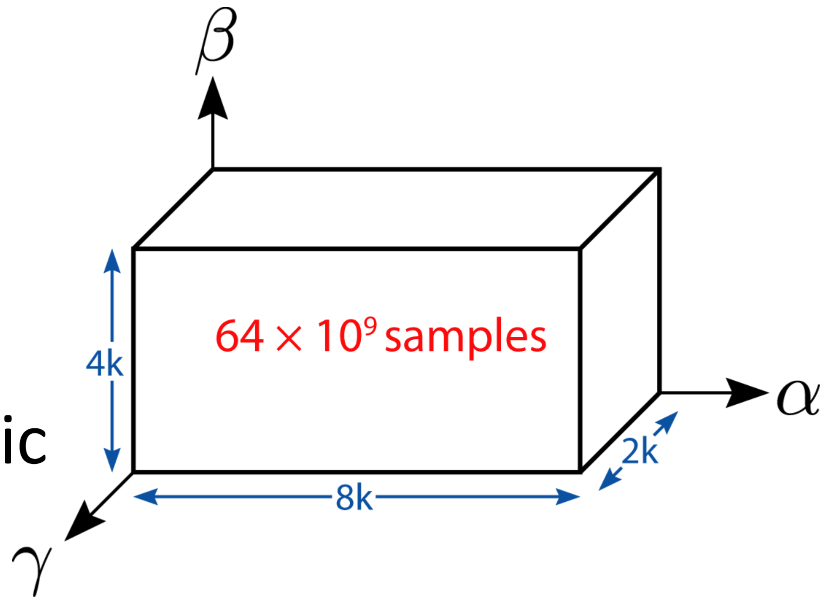
then *E* **+=** $\ell^{a_{fit}}$

else *E* **-=** c_{cross}



Silhouette Map Matching

- Naive implementation:
  - Sample SO(3) densely
  - For each (α,β,γ) sample, evaluate matching metric



➔ Robust, but prohibitively costly (≈ 8h)

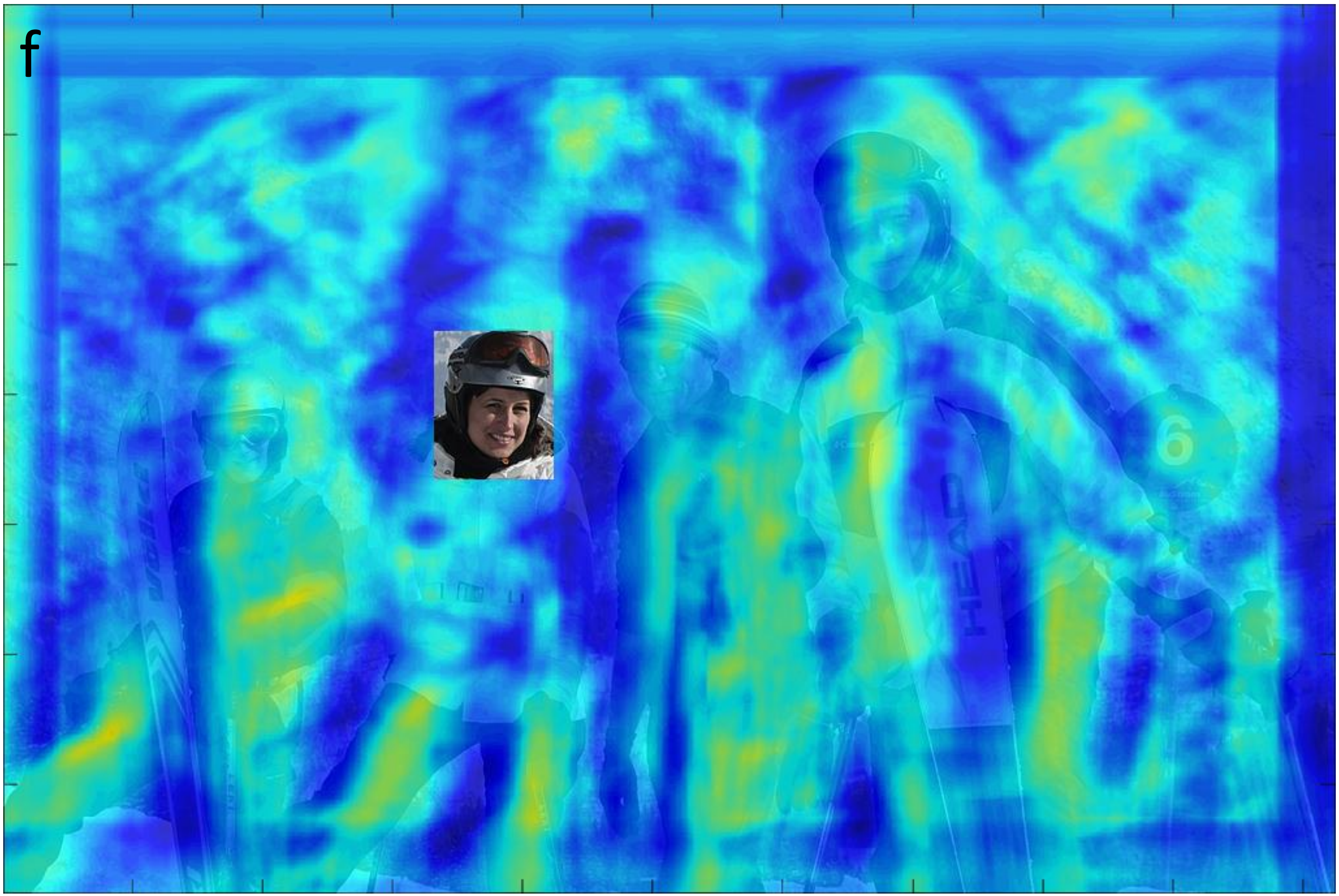
We need prior search-space reduction

➔ Spherical cross-correlation:
8 hours ⟶ 1 minute

Search Space Reduction - CC

- 2D Cross-Correlation (or sliding dot product)
  - $f \star p(x,y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(a,b)\overline{p(x-a,y-b)}dadb$
- Fast calculation using FFT
  - $\mathcal{F}\{f \star p\} = \mathcal{F}\{f\}\overline{\mathcal{F}\{p\}}$
  - $O(n^2 \log n)$
  - Template matching

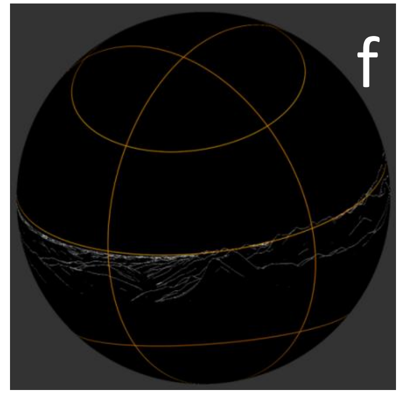
Cross-Correlation: principle in 2D

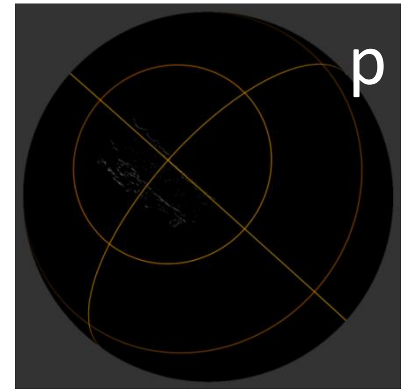


p

Spherical Cross-Correlation

- Spherical images:





- Spherical Cross-Correlation

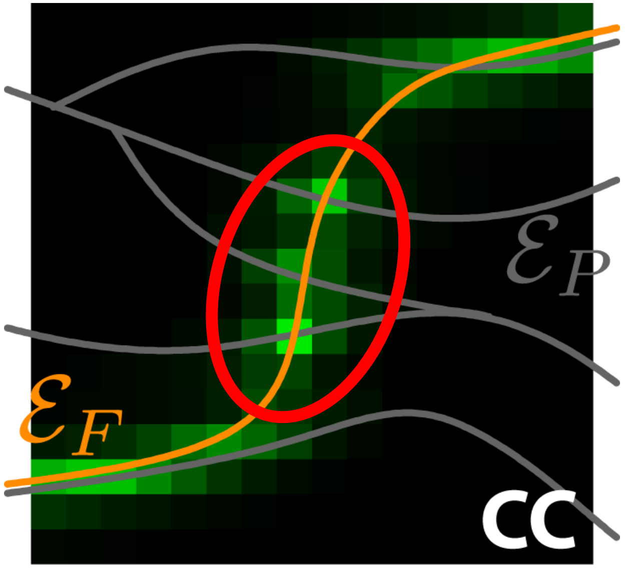
$$\forall g \in SO(3), \quad f \star p\ (g) = \int_{S^2} f(\omega)\overline{p(g^{-1}\omega)}d\omega$$

$$g = (\alpha, \beta, \gamma)$$

- Efficient computation on SO(3)
  – spherical harmonics and FFT **[Kostelec & Rockmore 2008]**
  – $O(n^3 \log n)$

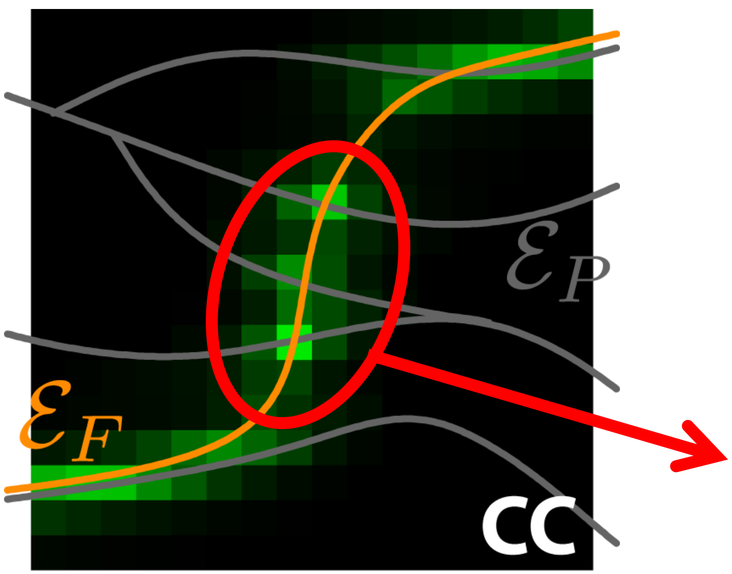
Spherical Cross-Correlation

- Pure cross-correlation
  - Maximizes edges overlap
  - Disregards orientations



Spherical Cross-Correlation

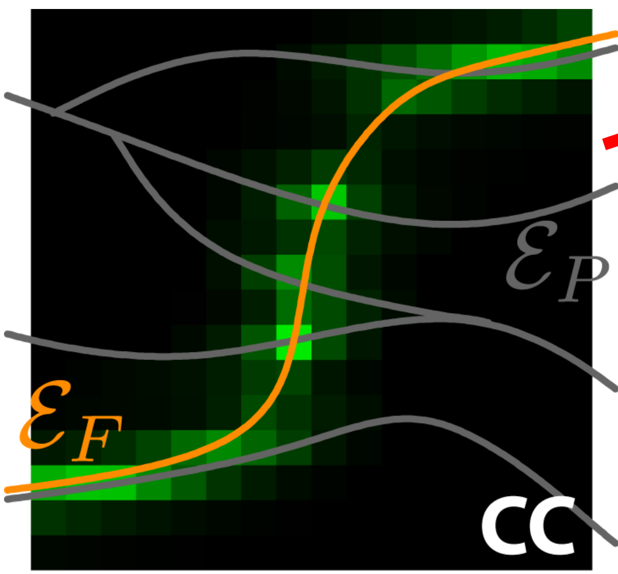
- Pure cross-correlation
  - Maximizes edges overlap
  - Disregards orientations

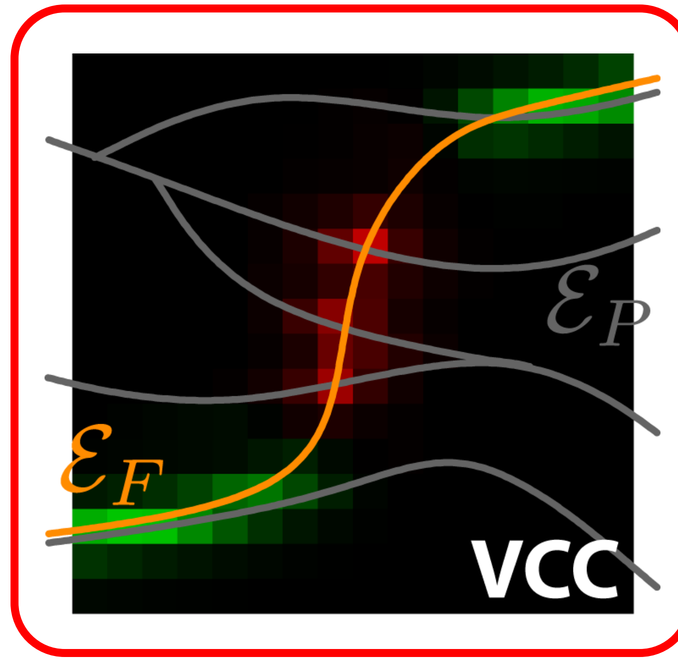


Crossings should penalize the cross-correlation score

Spherical Cross-Correlation

- Pure cross-correlation
  - Maximizes edges overlap
  - Disregards orientations





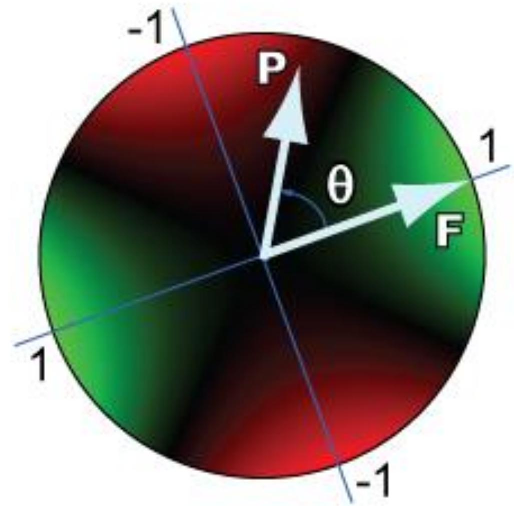
Vector-field Cross-Correlation
(edge-map = 2D vector field)

Vector-field Cross-Correlation

- Angular similarity operator

$$\mathcal{M}(\mathbf{f},\mathbf{p}) = \rho_f^2 \rho_p^2 \cos 2(\theta_f - \theta_p)$$

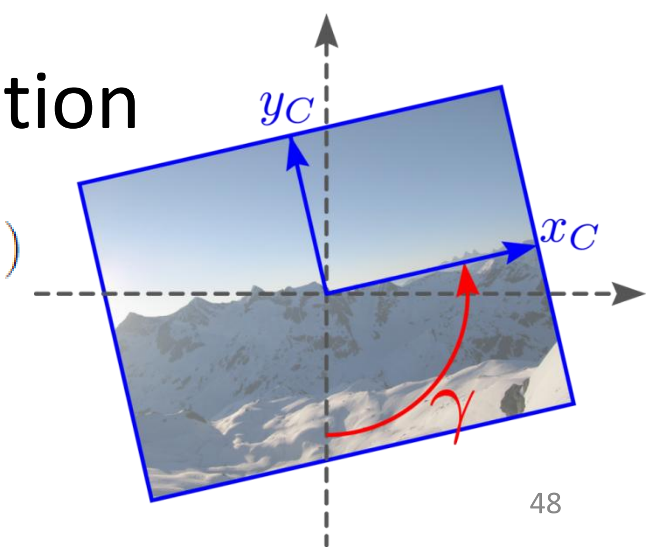
  - positive for parallel vectors
  - negative for orthogonal vectors
  - zero if one vector is zero



- Compensation for photo’s rotation

$$\mathcal{M}_g(\mathbf{f},\mathbf{p}) = \rho_f^2 \rho_p^2 \cos 2(\theta_f - (\theta_p + \gamma + \frac{\pi}{2}))$$

$$g = (\alpha, \beta, \gamma)$$



Vector-field Cross-Correlation

$$\mathrm{VCC}(\mathbf{f},\mathbf{p})(g) = \int_{S^2} \mathcal{M}_g(\mathbf{f}(\omega),\mathbf{p}(g^{-1}\omega))d\omega$$

- Reformulation: 2D vectors $\longrightarrow$ complex numbers

$$\left.\begin{aligned}\hat{f} &= \rho_f e^{i\theta_f}\\ \hat{p} &= \rho_p e^{i\theta_p}\end{aligned}\right\} \longrightarrow \begin{aligned}\mathcal{M}(\mathbf{f},\mathbf{p}) &= \rho_f^2\rho_p^2\cos 2(\theta_f-\theta_p)\\ &= Re\left\{\hat{f}^2\overline{\hat{p}^2}\right\}\end{aligned}$$

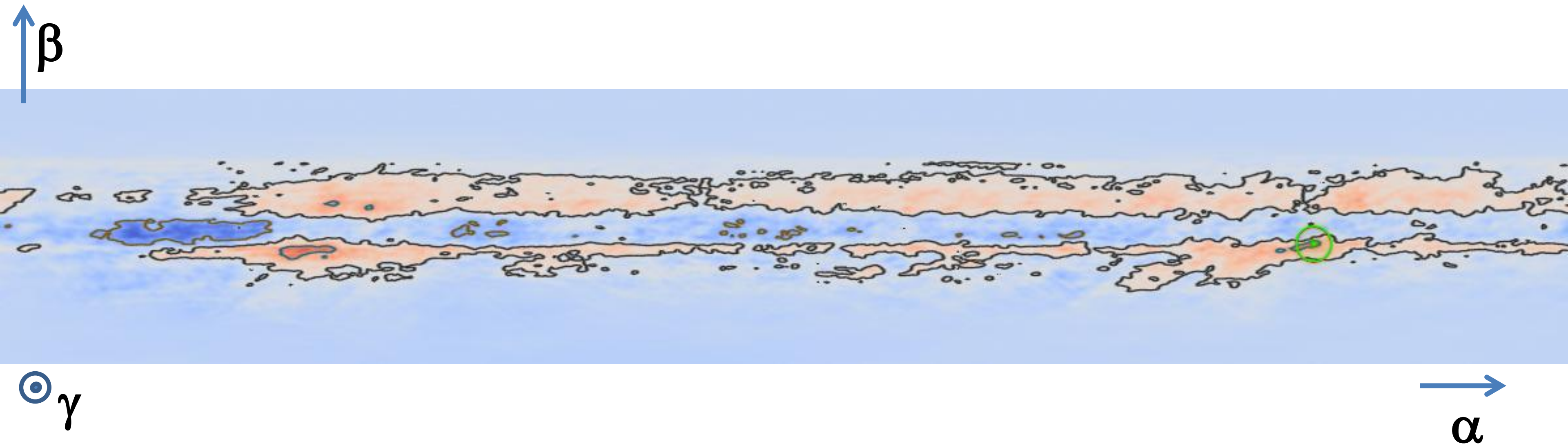
$$\begin{aligned}\mathcal{M}_g(\mathbf{f},\mathbf{p}) &= Re\left\{\hat{f}^2\overline{\left(e^{i(\gamma+\frac{\pi}{2})}\hat{p}\right)^2}\right\}\\ &= -Re\left\{e^{-i2\gamma}\hat{f}^2\overline{\hat{p}^2}\right\}\end{aligned}$$

$\Rightarrow$ $\mathrm{VCC}(\mathbf{f},\mathbf{p})(g) = -Re\left\{e^{-i2\gamma}\,\hat{f}^2\star\hat{p}^2(g)\right\}$

weighted SCC $\longrightarrow$ $O(n^3\log n)$

Vector-field Cross-Correlation

- Sampling: 512^3
- Thresholding:
  - **g**: reduce search space to 0.05% of the highest values



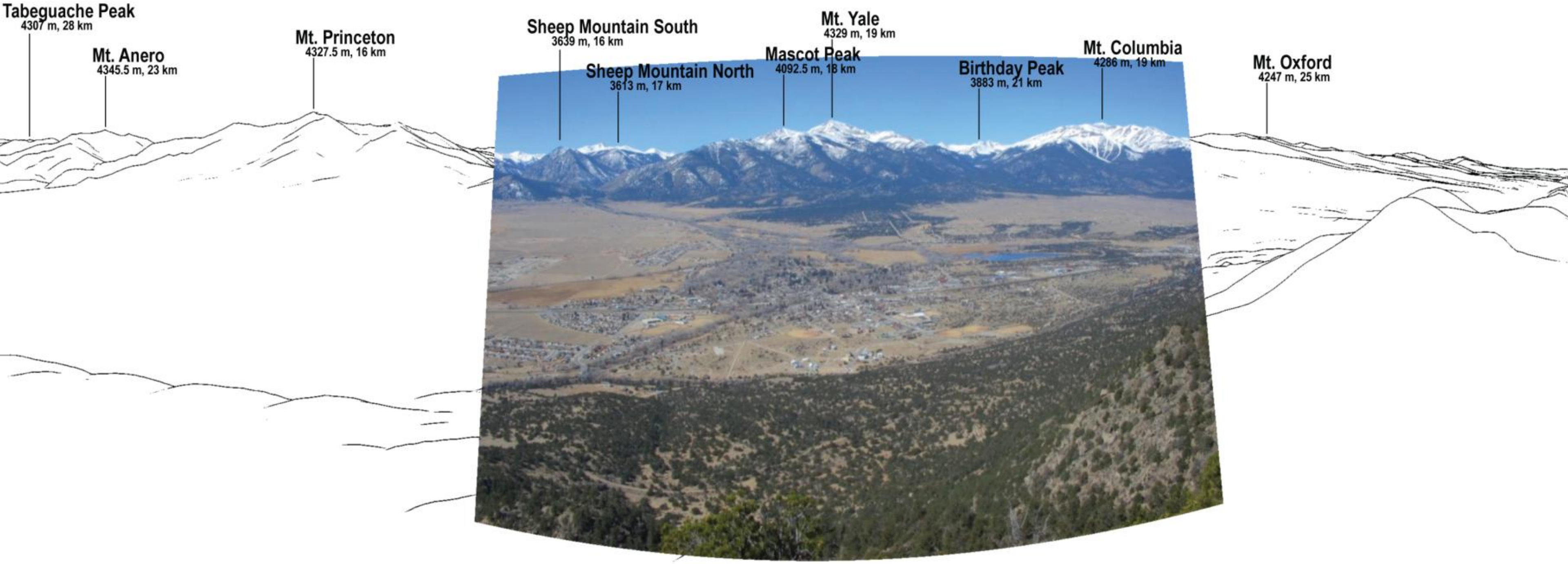
Results: performance

- 28 testing photos + 2 testing videos
  - VCC: maximal at ground truth for 25%
  - With matching metric: **86%**, accuracy within 0.2°
- Requires ~**2min** on current hardware:
  - Compass: ~1min
  - VCC (using SOFT lib): ~40s
  - Matching metric (GPU implementation): ~20s

Results

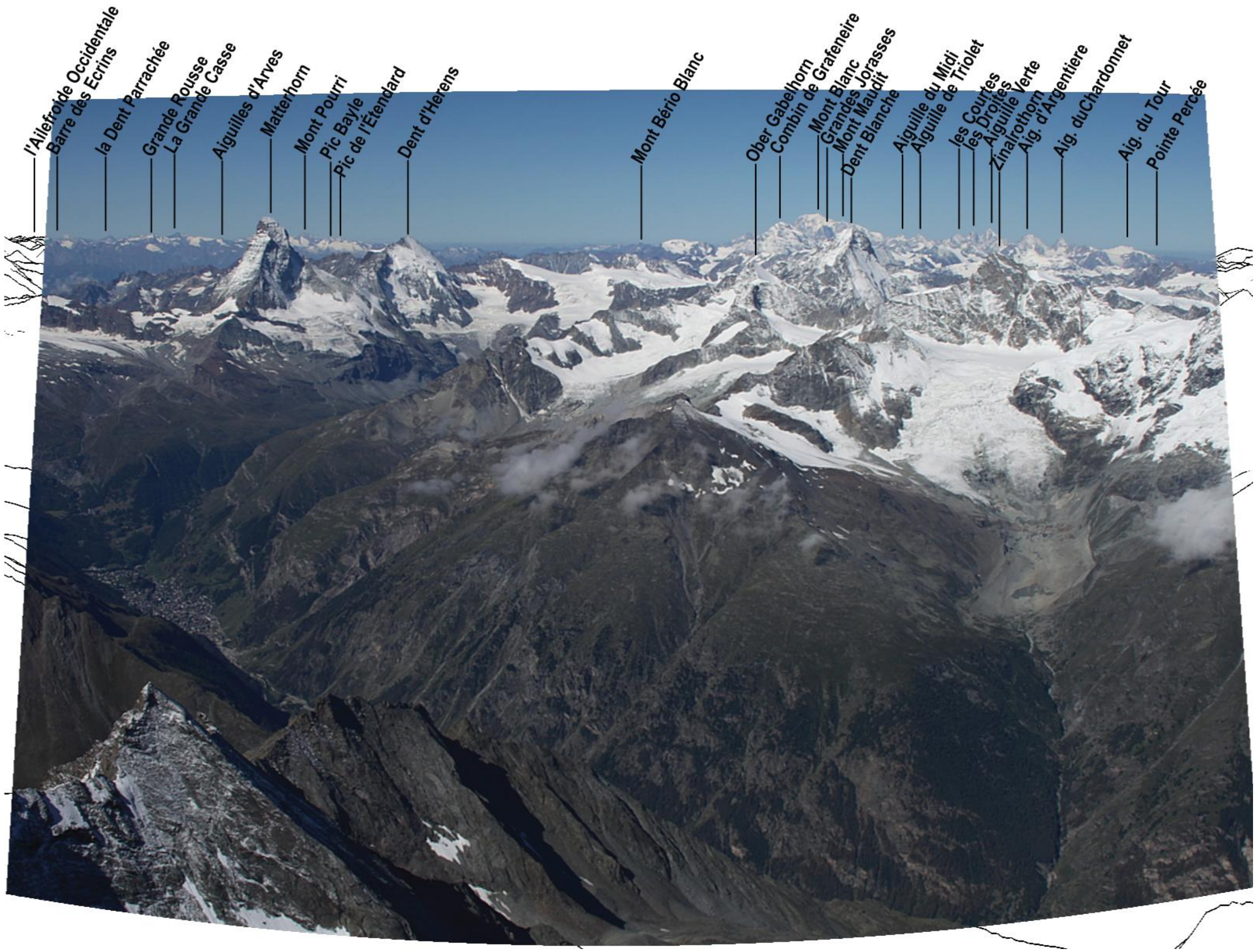
Mount Shavano
4337 m, 25 km
Tabeguache Peak
4307 m, 25 km
Mt. White
4078.5 m, 20 km
Mt. Anero
4345.5 m, 20 km
Mt. Princeton
4327.5 m, 16 km
Sheep Mountain South
3639 m, 17 km
Sheep Mountain North
3613 m, 18 km
Mascot Peak
4093 m, 20 km
Mt. Yale
4329 m, 21 km

Results

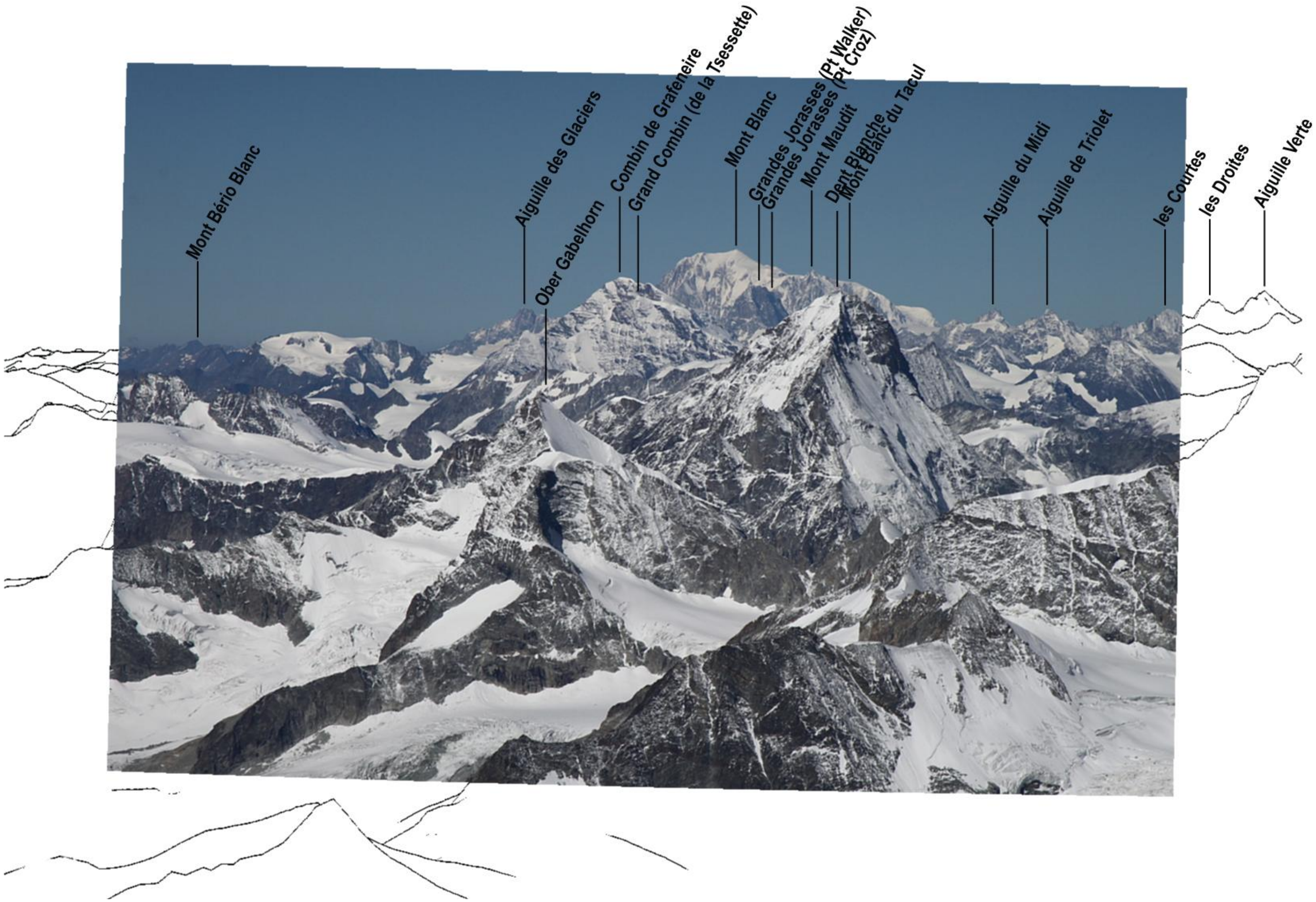


Tabeguache Peak
4307 m, 28 km
Mt. Anero
4345.5 m, 23 km
Mt. Princeton
4327.5 m, 16 km
Sheep Mountain South
3639 m, 16 km
Sheep Mountain North
3613 m, 17 km
Mt. Yale
4329 m, 19 km
Mascot Peak
4092.5 m, 18 km
Birthday Peak
3883 m, 21 km
Mt. Columbia
4286 m, 19 km
Mt. Oxford
4247 m, 25 km

Results

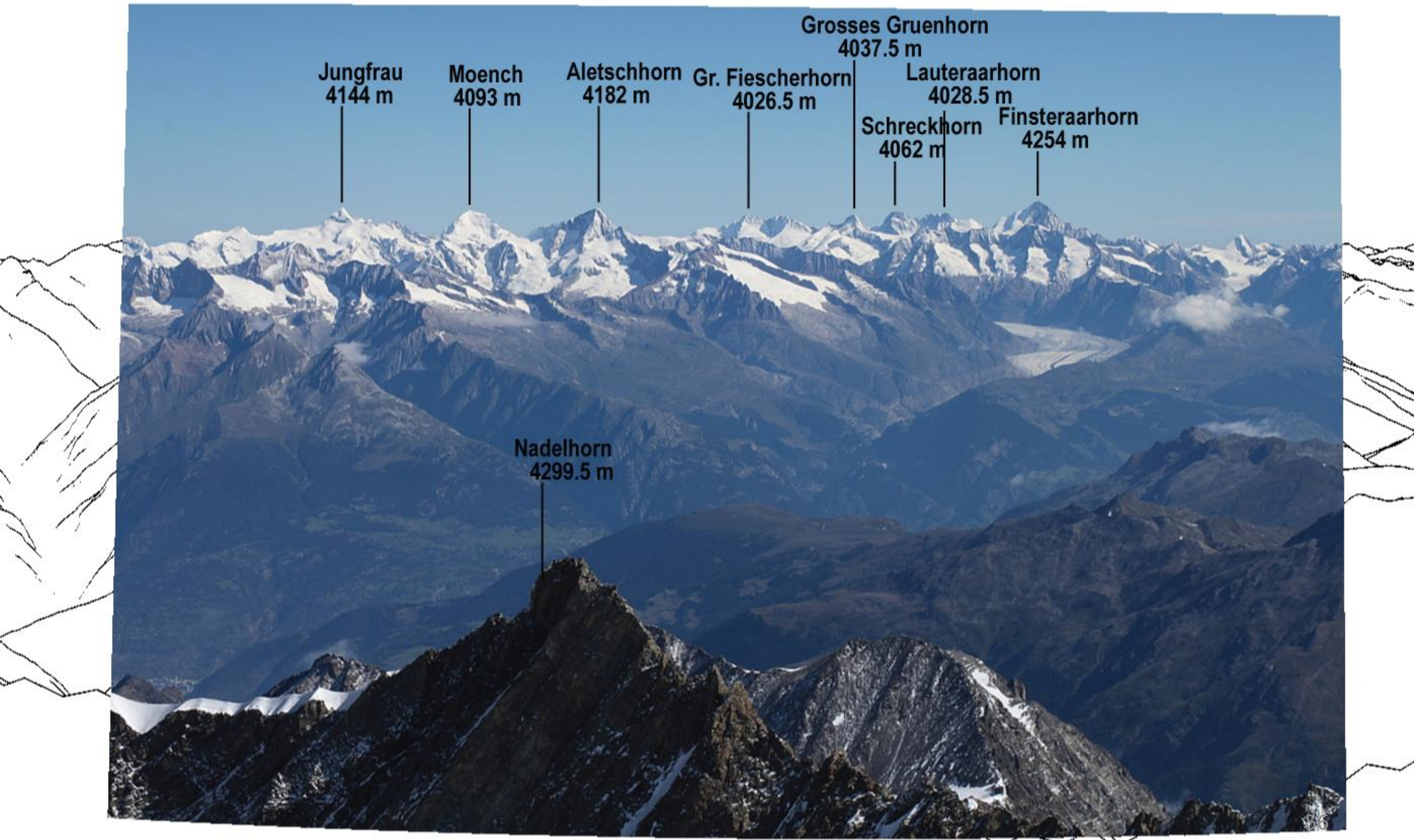

l'Ailefroide Occidentale
Barre des Ecrins
la Dent Parrachée
Grande Rousse
La Grande Casse
Aiguilles d'Arves
Matterhorn
Mont Pourri
Pic Bayle
Pic de l'Étendard
Dent d'Herens
Mont Bério Blanc
Ober Gabelhorn
Combin de Grafeneire
Mont Blanc
Grandes Jorasses
Mont Maudit
Dent Blanche
Aiguille du Midi
Aiguille de Triolet
les Courtes
les Droites
Aiguille Verte
Zinalrothorn
Aig. d'Argentiere
Aig. duChardonnet
Aig. du Tour
Pointe Percée

Results

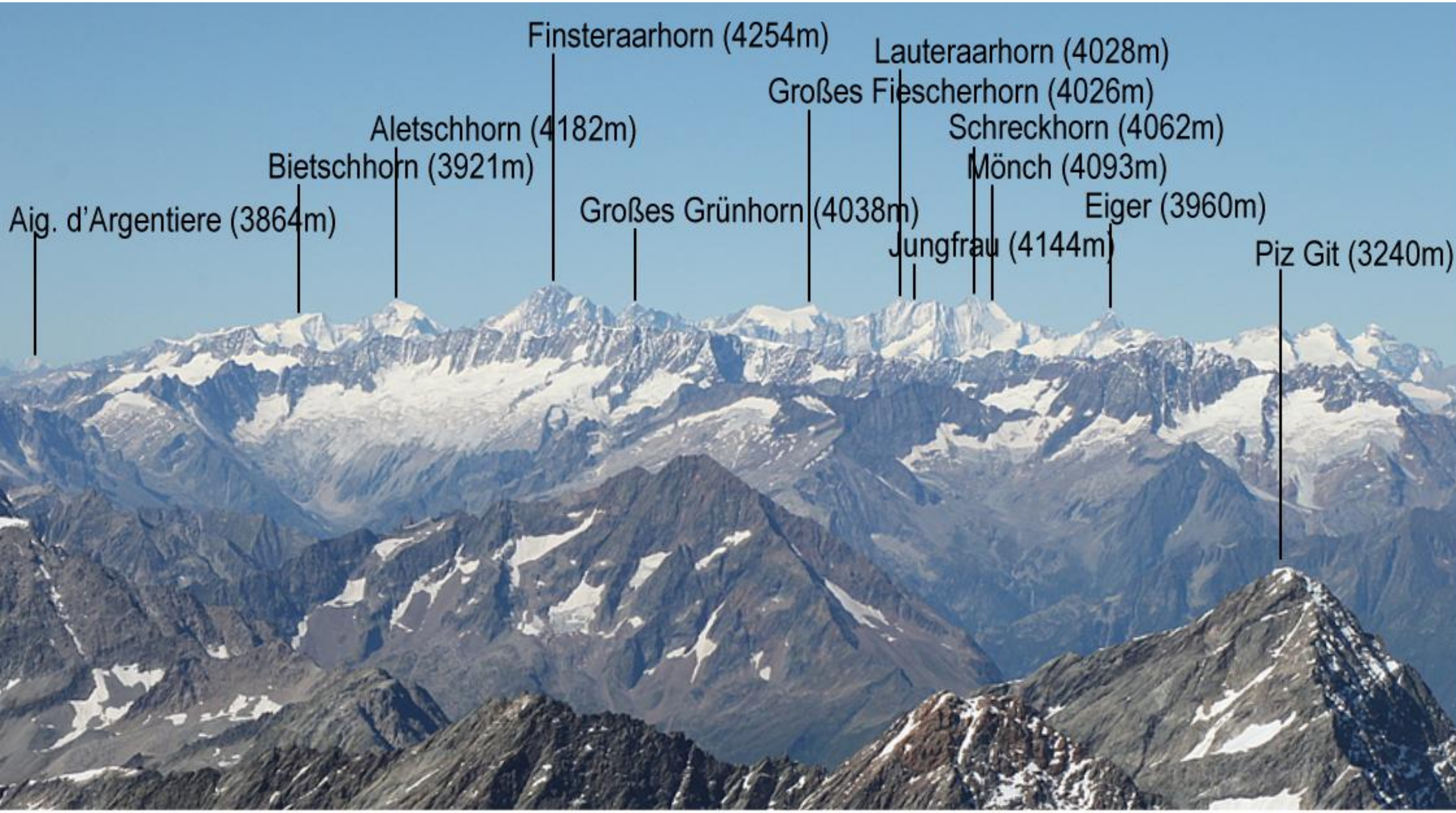


Mont Bério Blanc
Aiguille des Glaciers
Ober Gabelhorn
Combin de Grafeneire
Grand Combin (de la Tsessette)
Mont Blanc
Grandes Jorasses (Pt Walker)
Grandes Jorasses (Pt Croz)
Mont Maudit
Dent Blanche
Mont Blanc du Tacul
Aiguille du Midi
Aiguille de Triolet
les Courtes
les Droites
Aiguille Verte

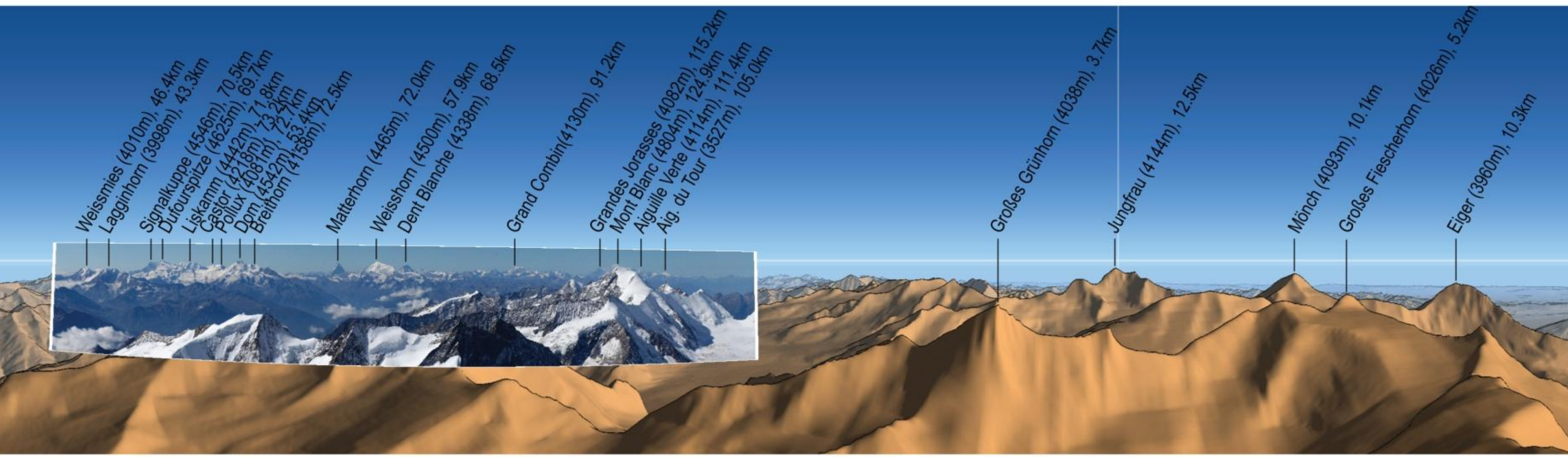
Results


Jungfrau
4144 m
Moench
4093 m
Aletschhorn
4182 m
Gr. Fiescherhorn
4026.5 m
Grosses Gruenhorn
4037.5 m
Lauteraarhorn
4028.5 m
Schreckhorn
4062 m
Finsteraarhorn
4254 m
Nadelhorn
4299.5 m

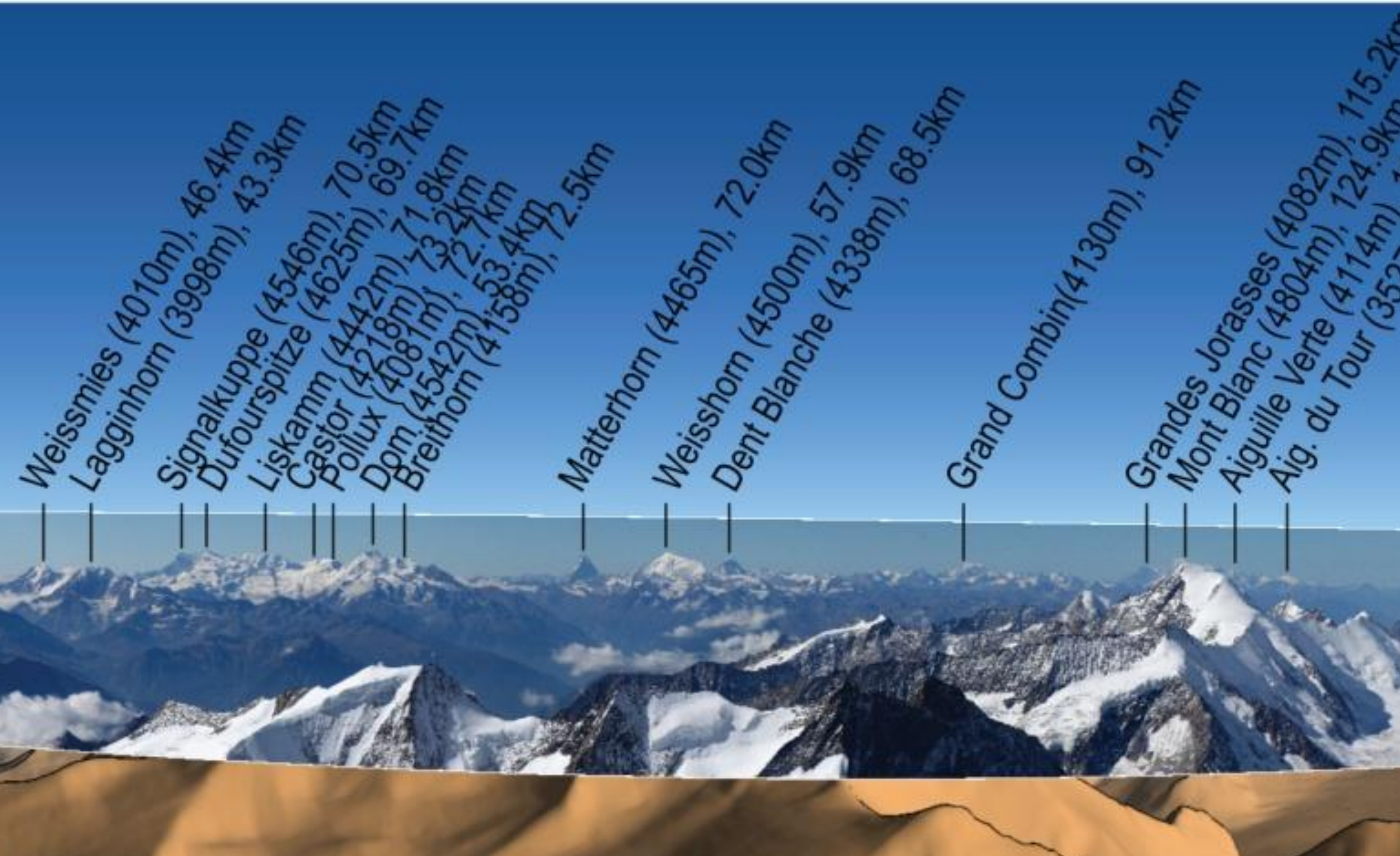
Results


Finsteraarhorn (4254m)
Lauteraarhorn (4028m)
Großes Fiescherhorn (4026m)
Aletschhorn (4182m)
Schreckhorn (4062m)
Bietschhorn (3921m)
Mönch (4093m)
Großes Grünhorn (4038m)
Eiger (3960m)
Aig. d'Argentiere (3864m)
Jungfrau (4144m)
Piz Git (3240m)

Results


Weissmies (4010m), 46.4km
Lagginhorn (3998m), 43.3km
Signalkuppe (4546m), 70.5km
Dufourspitze (4625m), 69.7km
Liskamm (4442m), 71.8km
Castor (4218m)
Pollux (4081m), 72.7km
Dom (4542m), 53.4km
Breithorn (4158m), 72.5km
Matterhorn (4465m), 72.0km
Weisshorn (4500m), 57.9km
Dent Blanche (4338m), 68.5km
Grand Combin(4130m), 91.2km
Grandes Jorasses (4082m), 115.2km
Mont Blanc (4804m), 124.9km
Aiguille Verte (4114m), 111.4km
Aig. du Tour (3527m), 105.0km
Großes Grünhorn (4038m), 3.7km
Jungfrau (4144m), 12.5km
Mönch (4093m), 10.1km
Großes Fiescherhorn (4026m), 5.2km
Eiger (3960m), 10.3km

Results



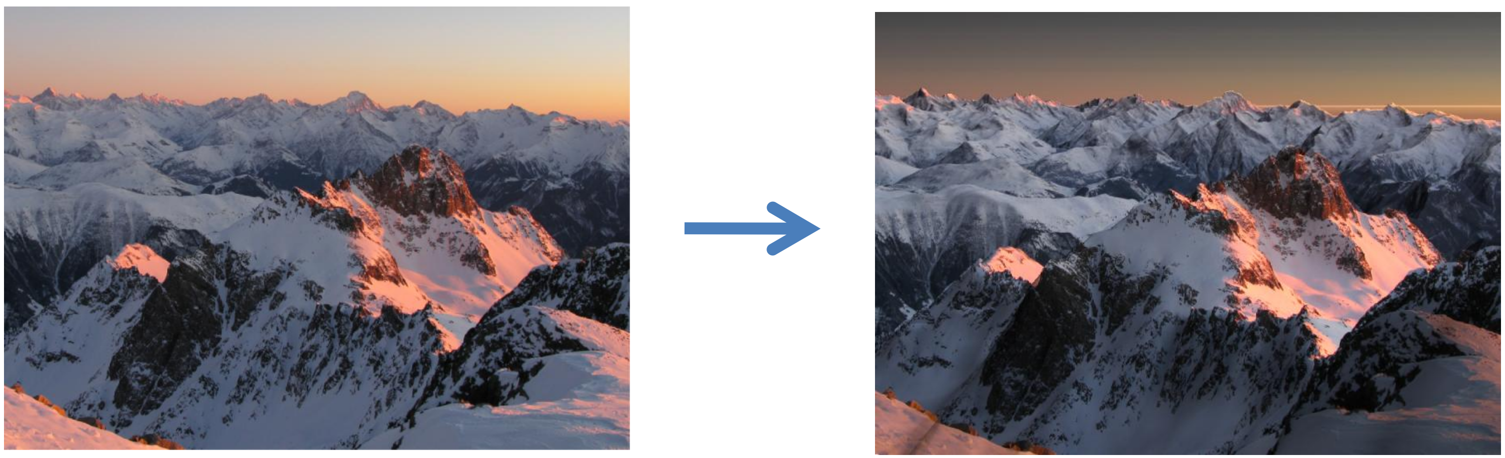
Weissmies (4010m), 46.4km
Lagginhorn (3998m), 43.3km
Signalkuppe (4546m), 70.5km
Dufourspitze (4625m), 69.7km
Liskamm (4442m), 71.8km
Castor (4218m), 73.2km
Pollux (4081m), 72.7km
Dom (4542m), 53.4km
Breithorn (4158m), 72.5km
Matterhorn (4465m), 72.0km
Weisshorn (4500m), 57.9km
Dent Blanche (4338m), 68.5km
Grand Combin(4130m), 91.2km
Grandes Jorasses (4082m), 115.2km
Mont Blanc (4804m), 124.9km
Aiguille Verte (4114m),
Aig. du Tour

Results

- Works for videos as well
- VIDEO

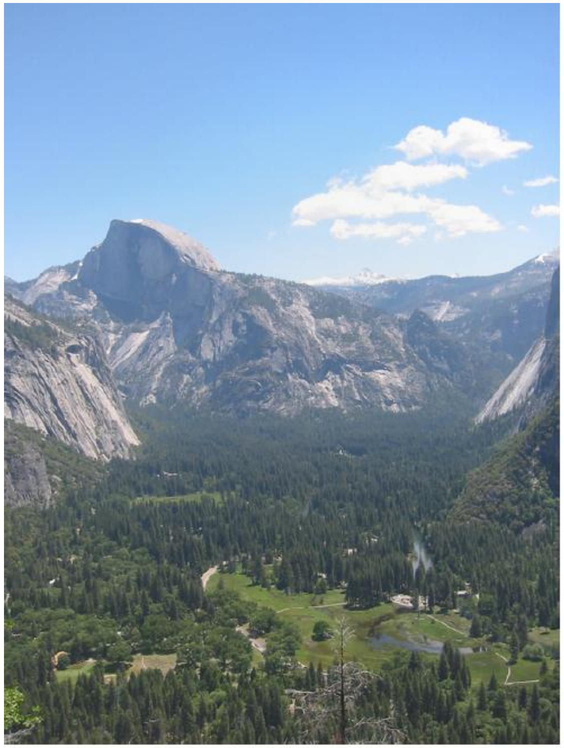
Other applications

- Advanced image enhancement
  - Contrast enhancement/dehazing, etc.



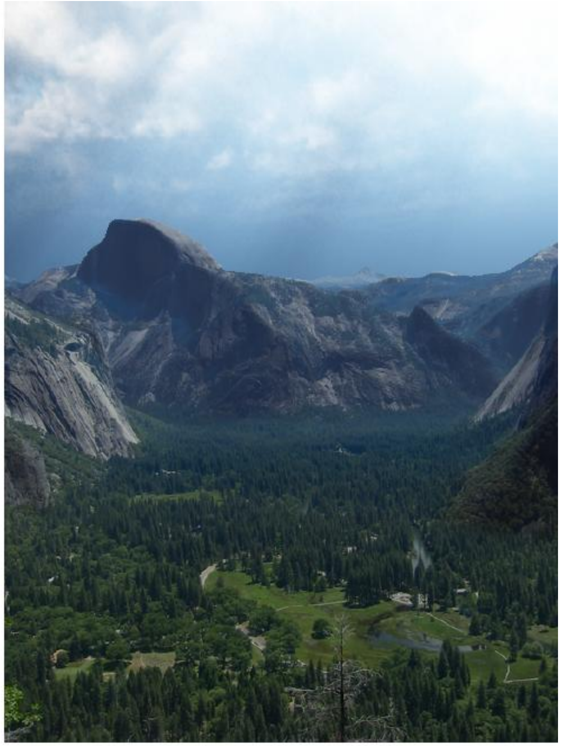
  - 3D objects integration (depth information)

Other applications



Input

Relighted



Relighted

[Kopf et al. 2008]

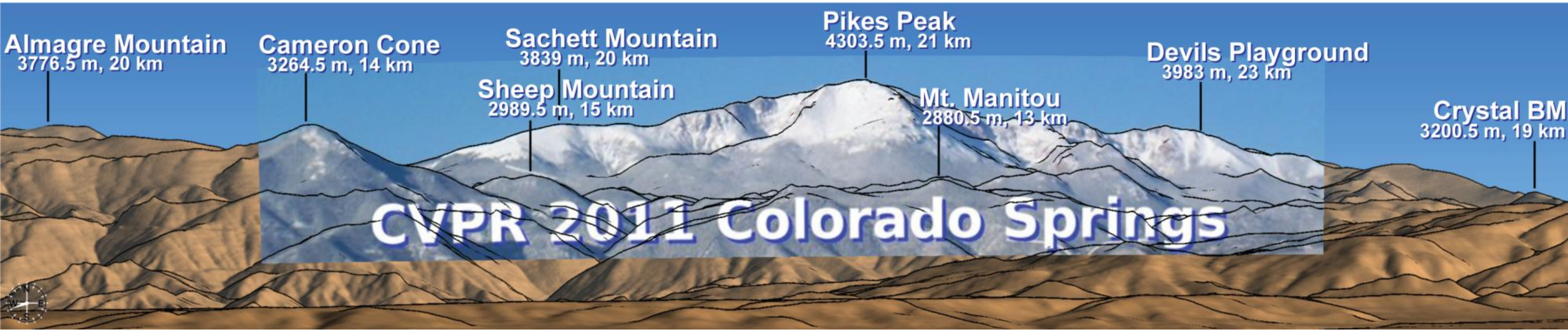
Conclusions and Future Work

- Mountain photo-to-3D model registration technique
  - Robust silhouette-map matching metric
  - Fast space reduction using SCC
  - Many applications (image/video annotation, augmented reality, model-based image enhancement, etc.)

- Future Work
  - Edge detection: other cues (e.g. aerial perspective)
  - Optimization for viewpoint position and FOV
  - Matching reliability prediction
  - Other possible applications of VCC

One last example...

Thanks!



More resources:

http://www.mpi-inf.mpg.de/resources/photo-to-terrain/

References

- **[Snavely et al. 06]** N. Snavely, S. M. Seitz, and R. Szeliski. Photo tourism: Exploring photo collections in 3d. In *SIGGRAPH Conference Proceedings*, pages 835–846, New York, NY, USA, 2006.
- **[Szeliski 06]** R. Szeliski. Image alignment and stitching: a tutorial. *Found.Trends. Comput. Graph. Vis.*, 2(1):1–104, 2006.
- **[Lowe 04]** D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60:91–110, 2004.
- **[Herdtweck and Wallraven 10]** C. Herdtweck and C. Wallraven. Horizon estimation: perceptual and computational experiments. In *APGV '10: Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization*, pages 49–56, USA, 2010.
- **[Woo et al. 07]** J. Woo, K. Son, T. Li, G. S. Kim, and I.-S. Kweon. Vision based uav navigation in mountain area. In *MVA*, pages 236–239, 2007.
- **[Mukunori et al. 97]** P. C. N. Jr, M. Mukunoki, M. Minoh, and K. Ikeda. Estimating camera position and orientation from geographical map and mountain image. In *38th Research Meeting of the Pattern Sensing Group, Society of Instrument and Control Engineers*, pages 9–16, 1997.