



**OPPA European Social Fund
Prague & EU: We invest in your future.**

► Choleski Decomposition for B. A.

The most expensive computation in B. A. is solving the normal eqs:

$$\text{find } \mathbf{d}_s \text{ such that } - \sum_{r=1}^k \mathbf{L}_r^\top \nu_r(\theta^s) = \left(\sum_{r=1}^k \mathbf{L}_r^\top \mathbf{L}_r + \lambda \text{diag } \mathbf{L}_r^\top \mathbf{L}_r \right) \mathbf{d}_s$$

This is a linear set of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, where

- \mathbf{A} is very large approx. $3 \cdot 10^4 \times 3 \cdot 10^4$ for a small problem of 10000 points and 5 cameras
- \mathbf{A} is sparse and symmetric, \mathbf{A}^{-1} is dense direct matrix inversion is prohibitive

Choleski: Every symmetric positive definite matrix \mathbf{A} can be decomposed to $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$, where \mathbf{L} is lower triangular. If \mathbf{A} is sparse then \mathbf{L} is sparse, too.

1. decompose $\mathbf{A} = \mathbf{L}\mathbf{L}^\top$ transforms the problem to solving $\underbrace{\mathbf{L}\mathbf{L}^\top}_{\mathbf{c}} \mathbf{x} = \mathbf{b}$

2. solve for \mathbf{x} in two passes:

$$\mathbf{L}\mathbf{c} = \mathbf{b} \quad \mathbf{c}_i := \mathbf{L}_{ii}^{-1} \left(\mathbf{b}_i - \sum_{j < i} \mathbf{L}_{ij} \mathbf{c}_j \right) \quad \text{forward substitution, } i = 1, \dots, q$$

$$\mathbf{L}^\top \mathbf{x} = \mathbf{c} \quad \mathbf{x}_i := \mathbf{L}_{ii}^{-1} \left(\mathbf{c}_i - \sum_{j > i} \mathbf{L}_{ji} \mathbf{x}_j \right) \quad \text{back-substitution}$$

- Choleski decomposition is fast (does not touch zero blocks)
non-zero elements are $9p + 121c + 66pc \approx 3.4 \cdot 10^6$; ca. $250\times$ fewer than all elements
- it can be computed on single elements or on entire blocks
- use profile Choleski for sparse \mathbf{A} and diagonal pivoting for semi-definite \mathbf{A} [Triggs et al. 1999]
- λ controls the definiteness

Profile Choleski Decomposition is Simple

```
function L = pchol(A)
%
% PCHOL profile Choleski factorization,
%   L = PCHOL(A) returns lower-triangular sparse L such that A = L*L'
%   for sparse square symmetric positive definite matrix A,
%   especially useful for arrowhead sparse matrices.

[p,q] = size(A);
if p ~= q, error 'Matrix must be square'; end

L = sparse(q,q);
F = ones(q,1);
for i=1:q
    F(i) = find(A(i,:),1); % 1st non-zero on row i; we are building F gradually
    for j = F(i):i-1
        k = max(F(i),F(j));
        a = A(i,j) - L(i,k:(j-1))*L(j,k:(j-1))';
        L(i,j) = a/L(j,j);
    end
    a = A(i,i) - sum(full(L(i,F(i):(i-1))).^2);
    if a < 0, error 'Matrix must be positive definite'; end
    L(i,i) = sqrt(a);
end
end
```

► Gauge Freedom

1. The external frame is not fixed: See Projective Reconstruction Theorem, Slide 124

$$\underline{\mathbf{m}}_i \simeq \mathbf{P}_j \underline{\mathbf{X}}_i = \mathbf{P}_j \mathbf{H}^{-1} \mathbf{H} \underline{\mathbf{X}}_i = \mathbf{P}'_j \underline{\mathbf{X}}'_i$$

2. Some representations are not minimal, e.g.
- \mathbf{P} is 12 numbers for 11 parameters
 - we may represent \mathbf{P} in decomposed form $\mathbf{K}, \mathbf{R}, \mathbf{t}$
 - but \mathbf{R} is 9 numbers representing the 3 parameters of rotation

As a result

- there is no unique solution
- matrix $\sum_r \mathbf{L}_r^\top \mathbf{L}_r$ is singular

Solutions

- fixing the external frame (e.g. a selected camera frame) explicitly or by constraints
- imposing constraints on projective entities
 - cameras, e.g. $\mathbf{P}_{3,4} = 1$ this excludes affine cameras
 - points, e.g. $\|\underline{\mathbf{X}}_i\|^2 = 1$ this way we can represent points at infinity
- using minimal representations
 - points in their Euclidean representation $\underline{\mathbf{X}}_i$ but finite points may be an unrealistic model
 - rotation matrix can be represented by Cayley transform see next

► Minimal Representations for Rotation

- \mathbf{o} – rotation axis, $\|\mathbf{o}\| = 1$, φ – rotation angle
- **wanted**: simple mapping to/from rotation matrices

1. Rodrigues' representation

$$\mathbf{R} = \mathbf{I} + \sin \varphi [\mathbf{o}]_{\times} + (1 - \cos \varphi) [\mathbf{o}]_{\times}^2$$
$$\sin \varphi [\mathbf{o}]_{\times} = \frac{1}{2}(\mathbf{R} - \mathbf{R}^{\top}), \quad \cos \varphi = \frac{1}{2}(\text{tr } \mathbf{R} - 1)$$

- hiding φ in the vector \mathbf{o} as in $[\sin \varphi \mathbf{o}]_{\times}$ is not so easy
- Cayley tried:

2. Cayley's representation; let $\mathbf{a} = \mathbf{o} \tan \frac{\varphi}{2}$, then

$$\mathbf{R} = (\mathbf{I} + [\mathbf{a}]_{\times})(\mathbf{I} - [\mathbf{a}]_{\times})^{-1}$$
$$[\mathbf{a}]_{\times} = (\mathbf{R} + \mathbf{I})^{-1}(\mathbf{R} - \mathbf{I})$$
$$\mathbf{a}_1 \circ \mathbf{a}_2 = \frac{\mathbf{a}_1 + \mathbf{a}_2 - \mathbf{a}_1 \times \mathbf{a}_2}{1 - \mathbf{a}_1^{\top} \mathbf{a}_2} \quad \text{composition of rotations } \mathbf{R} = \mathbf{R}_1 \mathbf{R}_2$$

- no trigonometric functions
- cannot represent rotation by 180°
- explicit composition formula

3. exponential map $\mathbf{R} = \exp[\varphi \mathbf{o}]_{\times}$, inverse by Rodrigues' formula

Minimal Representations for Other Entities

1. with the help of rotation we can minimally represent

- **fundamental matrix**

$$\mathbf{F} = \mathbf{U}\mathbf{D}\mathbf{V}^\top, \quad \mathbf{D} = \text{diag}(d, 1, 0), \quad \mathbf{U}, \mathbf{V} \text{ are rotations}, \quad 3 + 1 + 3 = 7 \text{ DOF}$$

- **essential matrix**

$$\mathbf{E} = [-\mathbf{t}]_\times \mathbf{R}, \quad \mathbf{R} \text{ is rotation}, \quad \|\mathbf{b}\| = 1, \quad 3 + 2 = 5 \text{ DOF}$$

- **camera**

$$\mathbf{P} = \mathbf{K} [\mathbf{R} \quad \mathbf{t}], \quad 5 + 3 + 3 = 11 \text{ DOF}$$

2. homography can be represented via exponential map

$$\exp \mathbf{A} = \sum_{k=0}^{\infty} \frac{1}{k!} \mathbf{A}^k \quad \text{note: } \mathbf{A}^0 = \mathbf{I}$$

some properties

$$\exp \mathbf{0} = \mathbf{I}, \quad \exp(-\mathbf{A}) = (\exp \mathbf{A})^{-1}, \quad \exp(\mathbf{A} + \mathbf{B}) \neq \exp(\mathbf{A}) \exp(\mathbf{B})$$

$\exp(\mathbf{A}^\top) = (\exp \mathbf{A})^\top$ hence if \mathbf{A} antisymmetric then $\exp \mathbf{A}$ orthogonal

$$(\exp(\mathbf{A}))^\top = \exp(\mathbf{A}^\top) = \exp(-\mathbf{A}) = (\exp(\mathbf{A}))^{-1}$$

$\det \exp \mathbf{A} = \exp(\text{tr } \mathbf{A})$ a key to homography representation:

$$\mathbf{H} = \exp \mathbf{Z} \text{ such that } \text{tr } \mathbf{Z} = 0, \text{ eg. } \mathbf{Z} = \begin{bmatrix} z_{11} & z_{12} & z_{13} \\ z_{21} & z_{22} & z_{23} \\ z_{31} & z_{32} & -(z_{11} + z_{22}) \end{bmatrix}, \quad 8 \text{ DOF}$$

► Implementing Simple Constraints

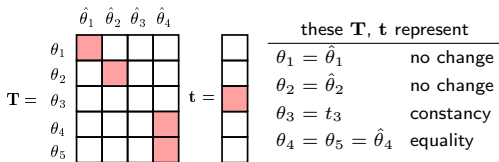
What for?

1. fixing external frame $\rightarrow \theta_i = \theta_i^0$ 'trivial gauge'
2. representing additional knowledge $\rightarrow \theta_i = \theta_j$ e.g. cameras share calibration matrix \mathbf{K}

We introduce reduced parameters $\hat{\theta}$:

$$\theta = \mathbf{T} \hat{\theta} + \mathbf{t}, \quad \mathbf{T} \in \mathbb{R}^{p, \hat{p}}, \quad \hat{p} \leq p$$

Then \mathbf{L}_r in LM changes to $\mathbf{L}_r \mathbf{T}$ and everything else stays the same



- \mathbf{T} deletes columns of \mathbf{L}_r that correspond to fixed parameters **it reduces the problem size**
- consistent initialisation: $\theta^0 = \mathbf{T} \hat{\theta}^0 + \mathbf{t}$
or filter the initialization by pseudoinverse $\theta^0 \mapsto \mathbf{T}^\dagger \theta^0$
- we need not compute derivatives for θ_j that correspond to all-zero rows \mathbf{T}_j
fixed params
- constraining projective entities \rightarrow minimal representations
- more complex constraints tend to make normal equations dense
- implementing constraints is safer than explicit renaming of the parameters, gives a flexibility to experiment
- other methods are much more involved, see [Triggs et al. 1999]
- **BA resource:** <http://www.ics.forth.gr/~lourakis/sba/>

Stereovision

- 4 Introduction
- 5 Epipolar Rectification
- 6 Binocular Disparity and Matching Table
- 7 Image Likelihood
- 8 Maximum Likelihood Matching
- 9 Uniqueness and Ordering as Occlusion Models
- 30 Three-Label Dynamic Programming Algorithm

mostly covered by

Šára, R. How To Teach Stereoscopic Vision. Proc. ELMAR 2010 [referenced as \[SP\]](#)

additional references



C. Geyer and K. Daniilidis. Conformal rectification of omnidirectional stereo pairs. In *Proc Computer Vision and Pattern Recognition Workshop*, p. 73, 2003.



J. Gluckman and S. K. Nayar. Rectifying transformations that minimize resampling effects. In *Proc IEEE CS Conf on Computer Vision and Pattern Recognition*, vol. 1:111–117. 2001.



M. Pollefeys, R. Koch, and L. V. Gool. A simple and efficient rectification method for general motion. In *Proc Int Conf on Computer Vision*, vol. 1:496–501, 1999.

What Are The Relative Distances?



- monocular vision already gives a rough 3D sketch because we understand the scene

What Are The Relative Distances?



Centrum för teknikstudier at Malmö Högskola, Sweden

- we have no help from image interpretation here
- this is how difficult is low-level stereo we will attempt to solve

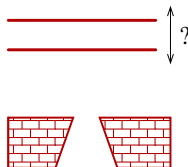
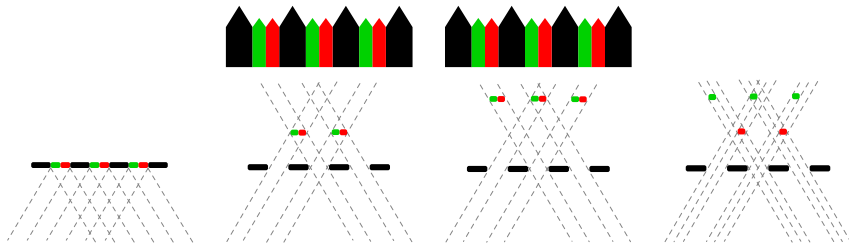
What Are The Relative Distances? (Why?)



- a combination of lack of texture and occlusion → ambiguous interpretation

Repetition: How Many Scenes Correspond to a Stereopair?

Consider the fence and the fortress worlds ...



- lack of texture is a limiting case of repetition

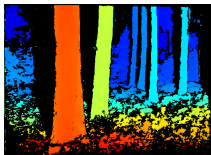
How Difficult Is Stereo?



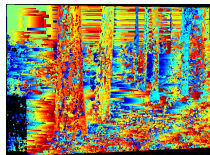
- when we do not recognize the scene and cannot use high-level constraints the problem seems difficult (right, less so in the center)
- most stereo matching algorithms do not require scene understanding prior to matching
- the success of a model-free stereo matching algorithm is unlikely:



left image



disparity map



disparity map from WTA

WTA Matching:

- for every left-image pixel find the most similar right-image pixel along the corresponding epipolar line [Marroquin 83]

Why Model-Free Stereo Fails?

- lack of an occlusion model
- lack of a continuity model

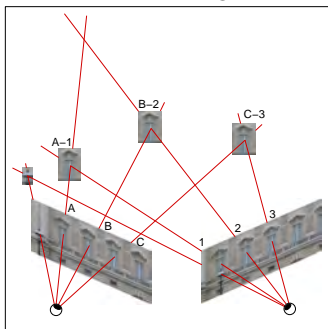
⇒ structural ambiguity



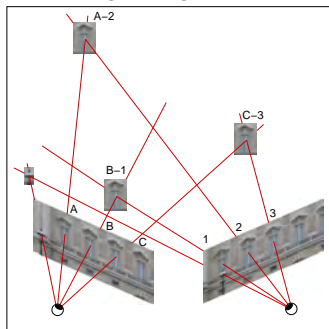
left image



right image

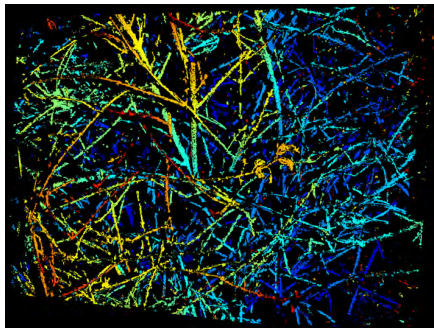


interpretation 1



interpretation 2

But What Kind of Continuity Model Applies Here?



- continuity alone is not a sufficient model
- occlusion model is more primal
- but occlusion model alone is insufficient, since it does not solve structural ambiguity

A Summary of Our Observations and an Outlook

- simple matching algorithms do not work
- decisions on matches are not independent due to occlusions
 - occlusion constraint works along epipolars only
- occlusion model alone is insufficient
 - does not resolve the structural ambiguity
- a continuity model can resolve structural ambiguity
 - but continuity is piecewise due to object boundaries
- in sufficiently complex scenes the only possibility is that stereopsis uses scene interpretation (or another-modality measurement)

Outlook:

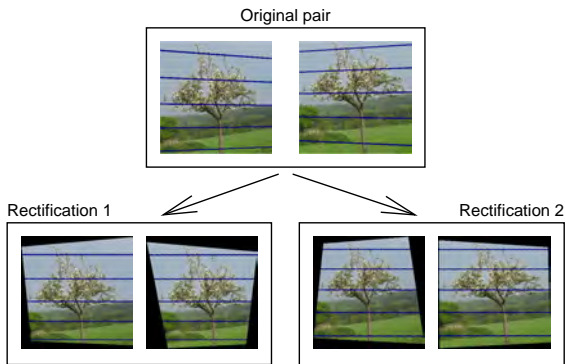
1. represent the occlusion constraint:
 - epipolar rectification
 - disparity
 - uniqueness as an occlusion constraint
2. represent piecewise continuity
 - ordering as a weak continuity model
3. use a consistent framework
 - looking for the most probable solution (MAP)

► Epipolar Rectification

Problem: Given fundamental matrix \mathbf{F} or camera matrices $\mathbf{P}_1, \mathbf{P}_2$, transform images so that epipolar lines become horizontal with the same row coordinate. The result is a standard stereo pair for easier correspondence search

Procedure:

1. find a pair of rectification homographies \mathbf{H}_1 and \mathbf{H}_2 .
2. warp images using \mathbf{H}_1 and \mathbf{H}_2 and modify fundamental matrix $\mathbf{F} \mapsto \mathbf{H}_2^{-T} \mathbf{F} \mathbf{H}_1^{-1}$ or cameras $\mathbf{P}_1 \mapsto \mathbf{H}_1 \mathbf{P}_1, \mathbf{P}_2 \mapsto \mathbf{H}_2 \mathbf{P}_2$.



- there is a 9-parameter family of rectification homographies for binocular rectification, see next

Rectification Example

Four cameras in general position



cam 1



cam 2



cam 3



cam 4



Rectified pairs



pair 1 - 2



pair 2 - 4



pair 1 - 4

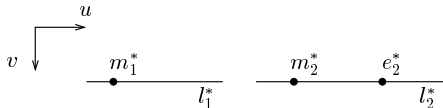


► Rectification Homographies

Cameras ($\mathbf{P}_1, \mathbf{P}_2$) are rectified by a homography pair ($\mathbf{H}_1, \mathbf{H}_2$):

$$\mathbf{P}_i^* = \mathbf{H}_i \mathbf{P}_i = \mathbf{H}_i \mathbf{K}_i \mathbf{R}_i \begin{bmatrix} \mathbf{I} & -\mathbf{C}_i \end{bmatrix}, \quad i = 1, 2$$

rectified entities: \mathbf{F}^* , l_2^* , l_1^* , etc:



corresponding epipolar lines must be:

1. parallel to image rows \Rightarrow epipoles become $e_1^* = e_2^* = (1, 0, 0)$
2. equivalent $l_2^* = l_1^* \Rightarrow l_2^* \simeq l_1^* \simeq \underline{e}_1^* \times \underline{m}_1 = [\underline{e}_1^*]_{\times} \underline{m}_1 = \mathbf{F}^* \underline{m}_1$

both conditions together give the rectified fundamental matrix

$$\mathbf{F}^* \simeq \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}$$

A two-step rectification procedure

1. Find some pair of primitive rectification homographies $\hat{\mathbf{H}}_1, \hat{\mathbf{H}}_2$
2. Upgrade them to a pair of optimal rectification homographies from the class preserving \mathbf{F}^* .

► Geometric Interpretation of Linear Rectification

What pair of cameras is compatible with \mathbf{F}^* ?

- we know that $\mathbf{F} = (\mathbf{Q}_1 \mathbf{Q}_2^{-1})^\top [\mathbf{e}_1]_\times$
- we choose $\mathbf{Q}_1^* = \mathbf{K}_1^*$, $\mathbf{Q}_2^* = \mathbf{K}_2^* \mathbf{R}^*$; then

Slide 77

$$(\mathbf{Q}_1^* \mathbf{Q}_2^{*-1})^\top [\mathbf{e}_1^*]_\times = (\mathbf{K}_1^* \mathbf{R}^{*\top} \mathbf{K}_2^{*-1})^\top \mathbf{F}^*$$

- we look for \mathbf{R}^* , \mathbf{K}_1^* , \mathbf{K}_2^* compatible with

$$(\mathbf{K}_1^* \mathbf{R}^{*\top} \mathbf{K}_2^{*-1})^\top \mathbf{F}^* = \lambda \mathbf{F}^*, \quad \mathbf{R}^* \mathbf{R}^{*\top} = \mathbf{I}, \quad \mathbf{K}_1^*, \mathbf{K}_2^* \text{ upper triangular}$$

- we also want \mathbf{b}^* from $\mathbf{e}_1^* \simeq \mathbf{P}_1^* \mathbf{C}_2^* = \mathbf{K}_1^* \mathbf{b}^*$ \mathbf{b}^* in cam. 1 frame
- result:

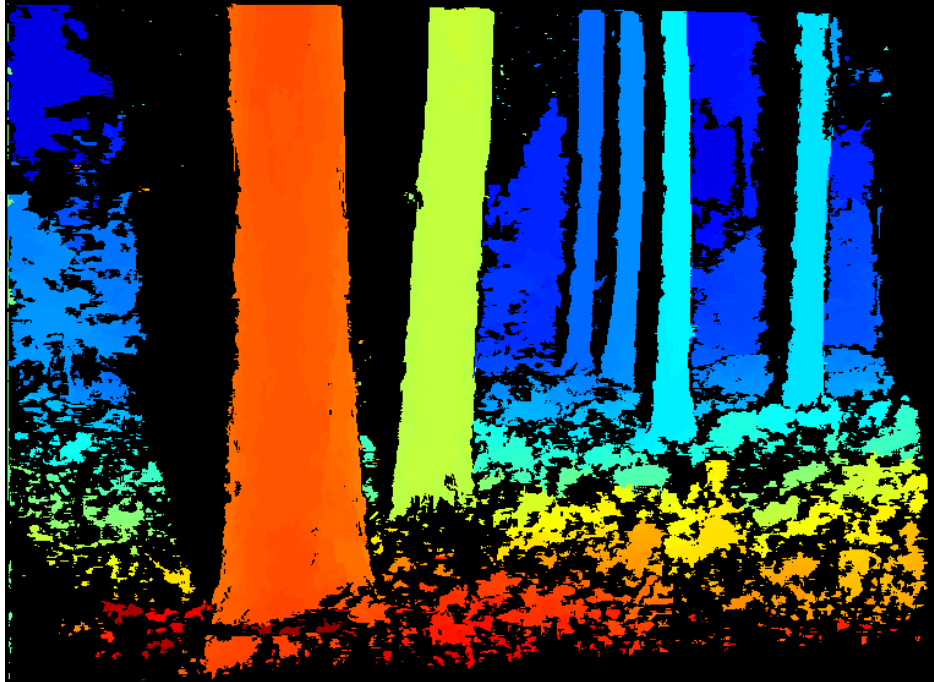
$$\mathbf{R}^* = \mathbf{I}, \quad \mathbf{b}^* = \begin{bmatrix} b \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{K}_1^* = \begin{bmatrix} k_{11} & k_{12} & k_{13} \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{K}_2^* = \begin{bmatrix} k_{21} & k_{22} & k_{23} \\ 0 & f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (29)$$

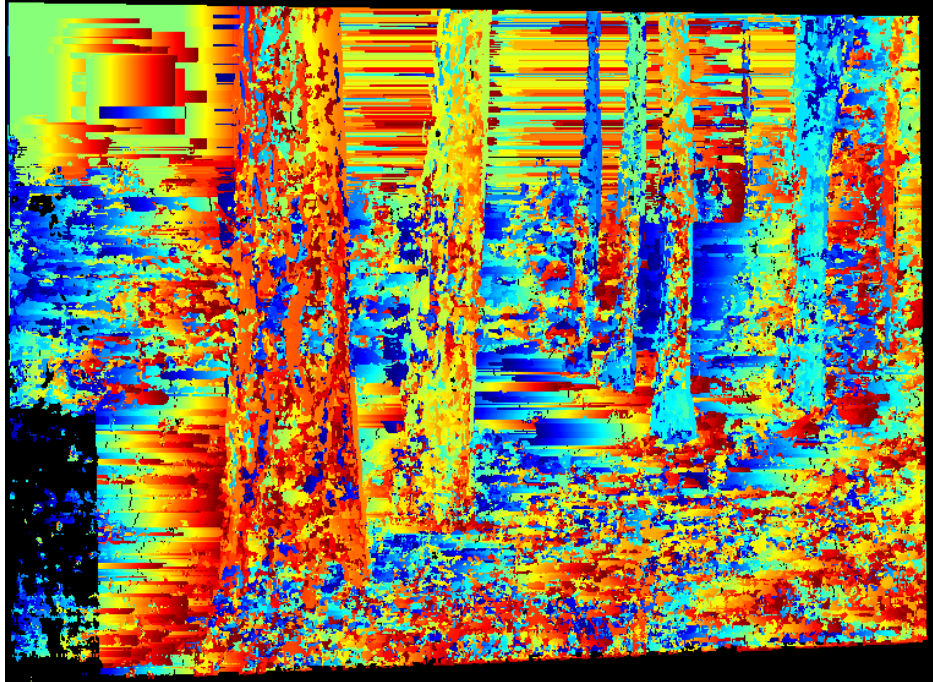
- rectified cameras are in canonical position with respect to each other
not rotated, canonical baseline
- rectified calibration matrices can differ in the first row only
- when $\mathbf{K}_1^* = \mathbf{K}_2^*$ then the rectified pair is called the standard stereo pair and the homographies standard rectification homographies

Thank You

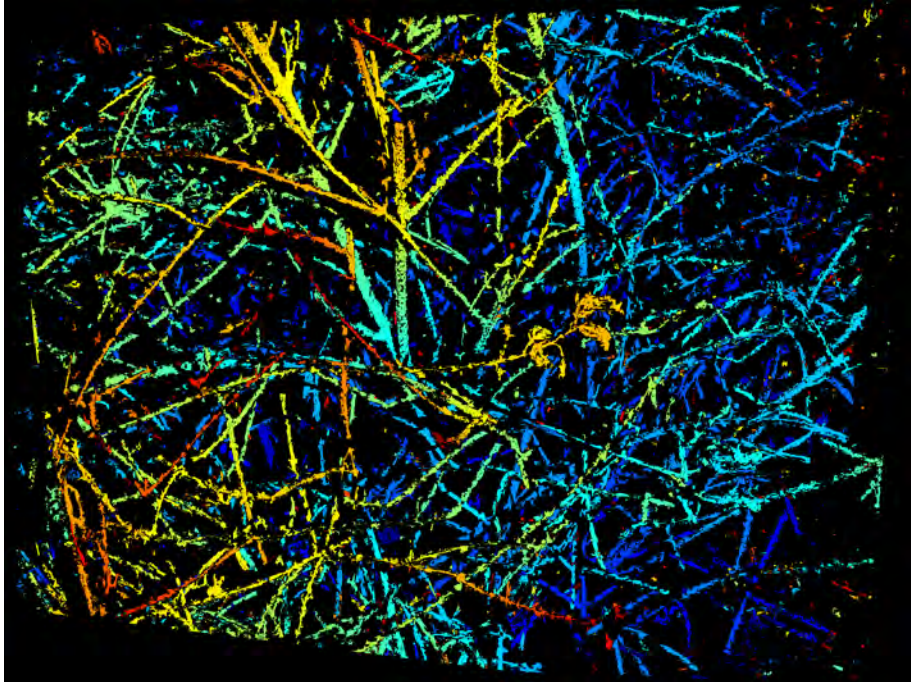














**OPPA European Social Fund
Prague & EU: We invest in your future.**
