University of Ljubljana

Faculty of Computer and Information Science

Computer Vision Laboratory

Mosaic-Based Panoramic Depth Imaging with a Single Standard Camera

Peter Peer, Franc Solina

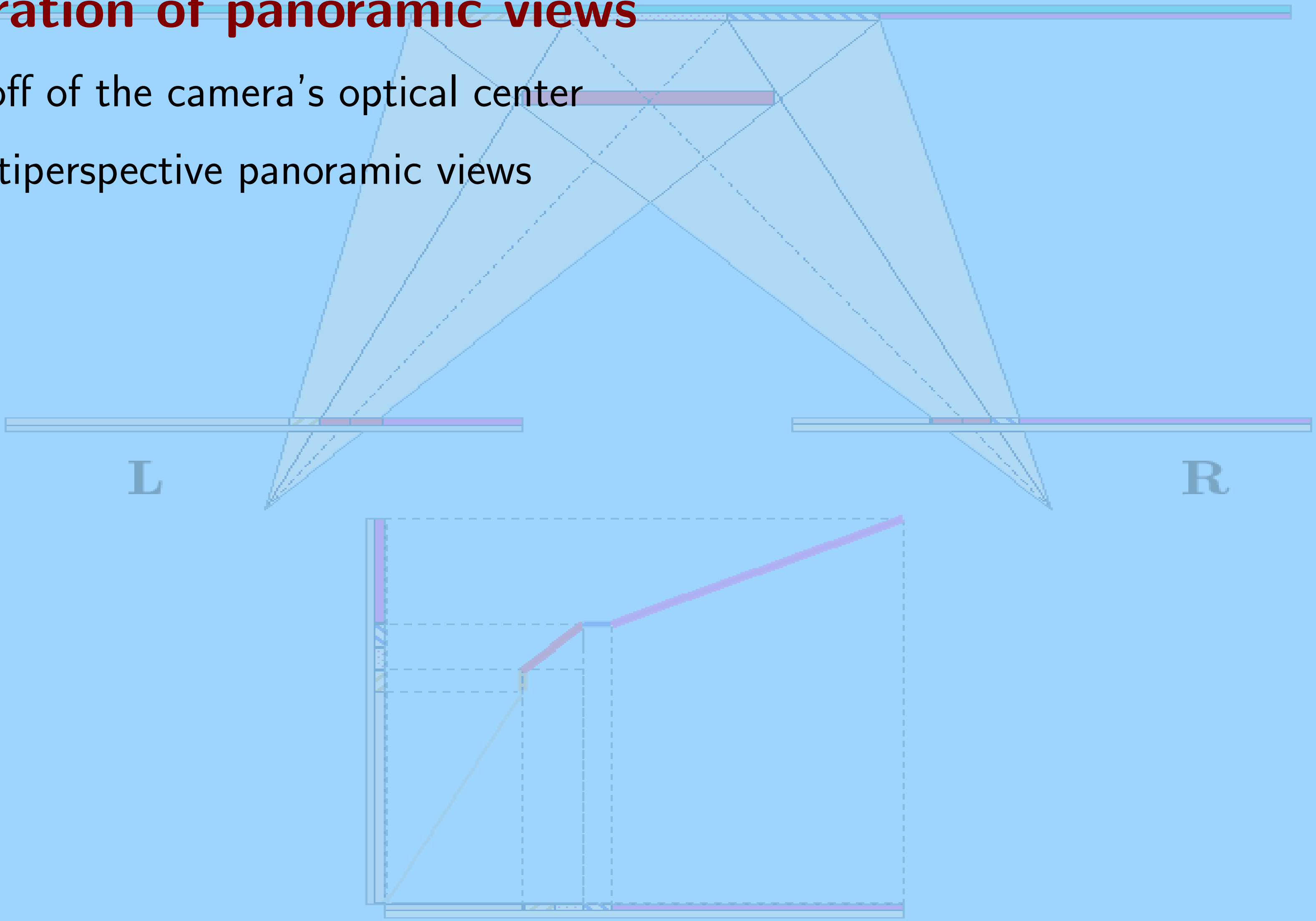
IEEE Workshop on Stereo and Multi-Baseline Vision, 2001 (CVPR)

Overview

1. Generation of panoramic views
2. Geometry of the system
3. Epipolar geometry
4. Stereo reconstruction
5. **Analysis**
6. Summary

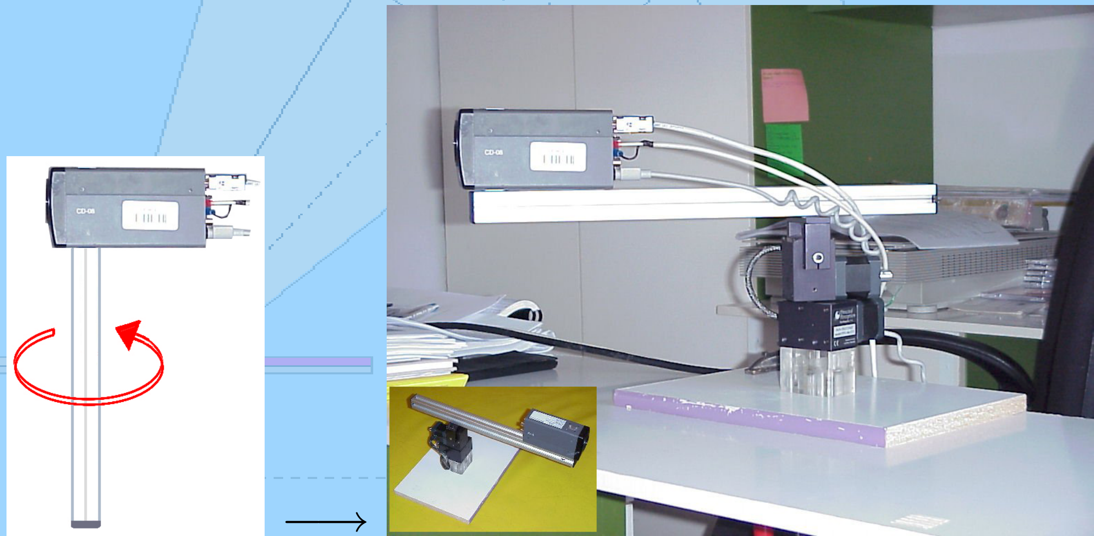
Generation of panoramic views

- Setoff of the camera's optical center
- Multiperspective panoramic views



Generation of panoramic views

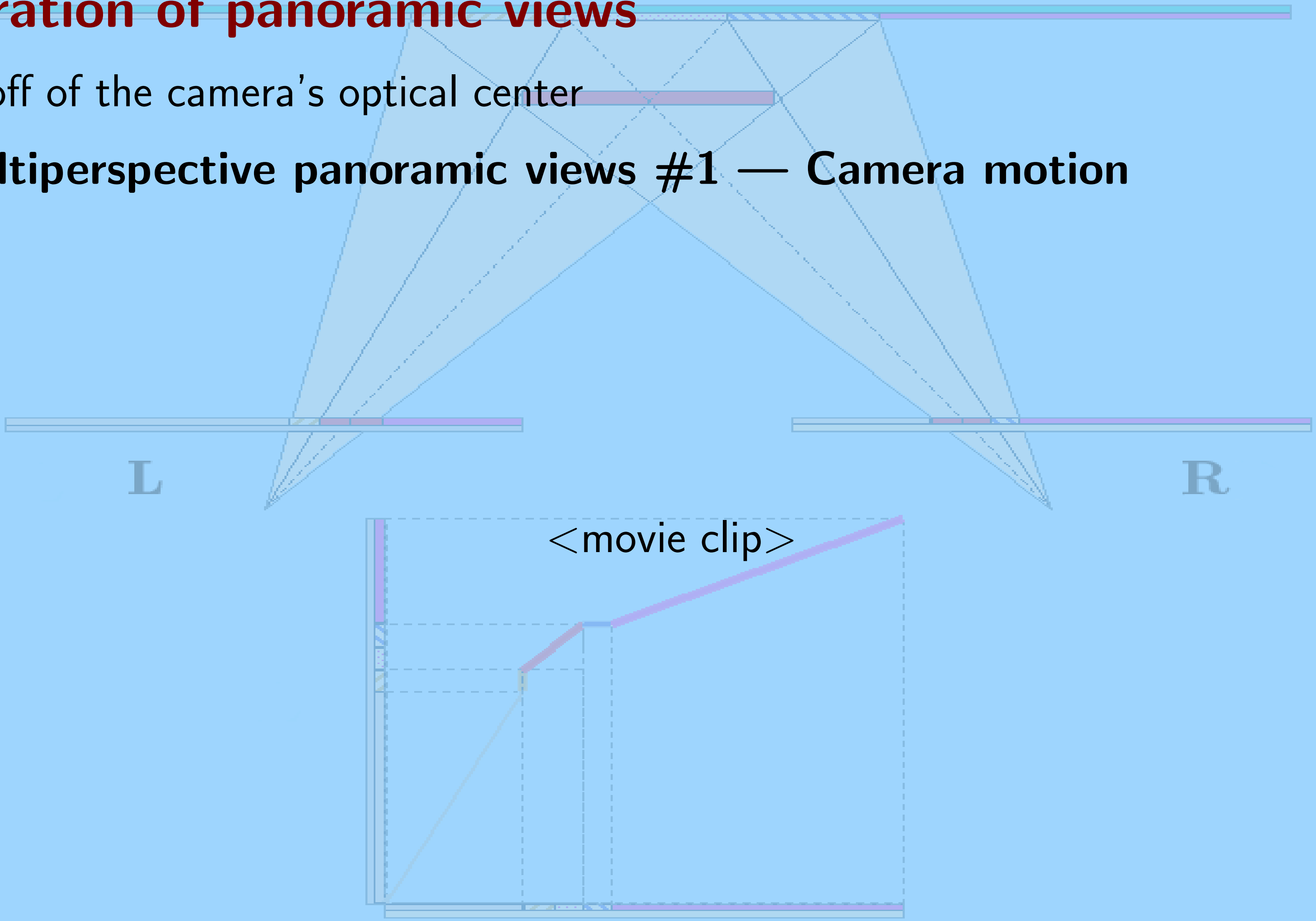
- **Setoff of the camera's optical center**



- Multiperspective panoramic views

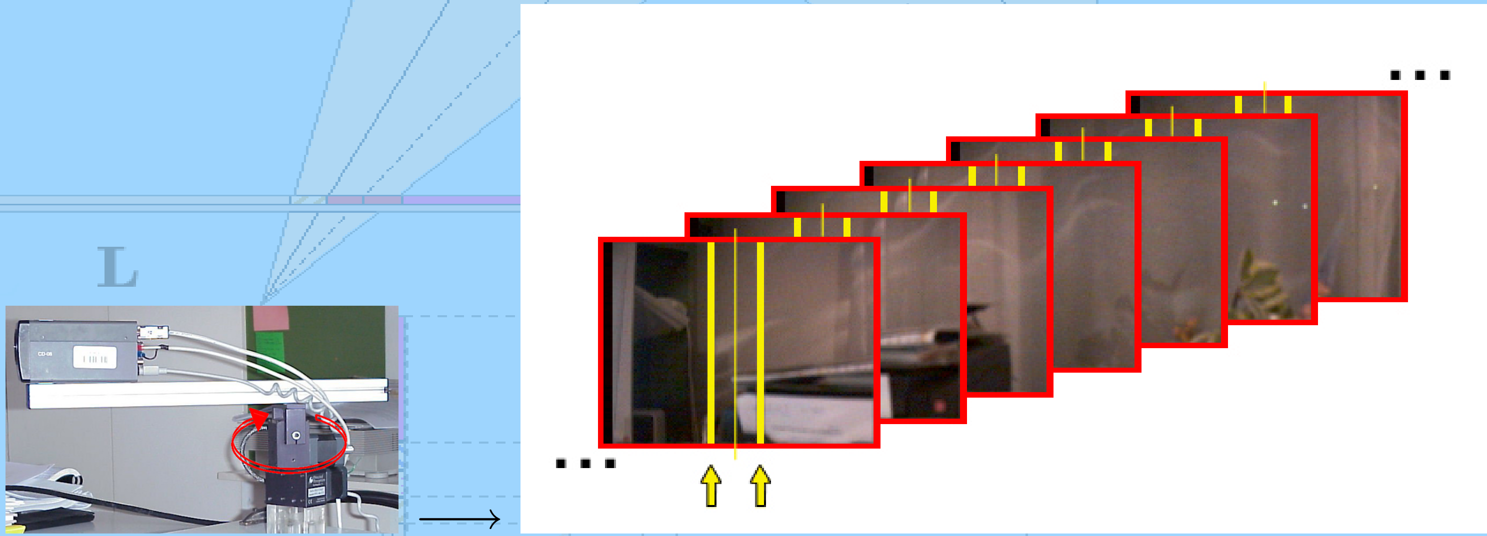
Generation of panoramic views

- Setoff of the camera's optical center
- **Multiperspective panoramic views #1 — Camera motion**



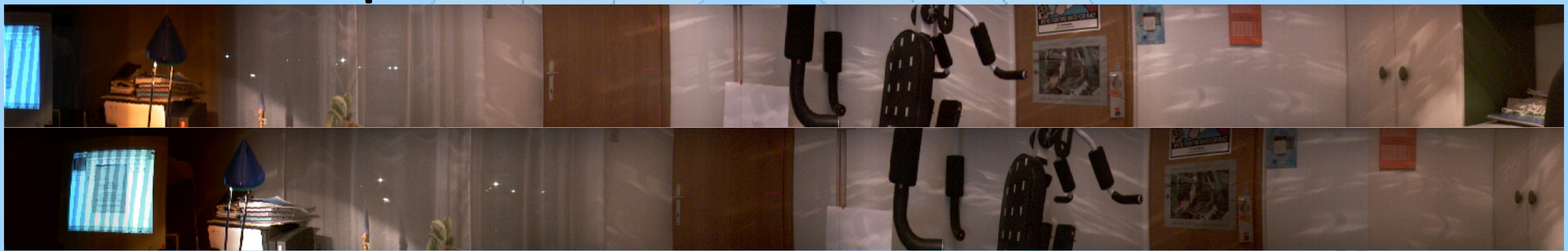
Generation of panoramic views

- Setoff of the camera's optical center
- **Multiperspective panoramic views #2 — Generation**

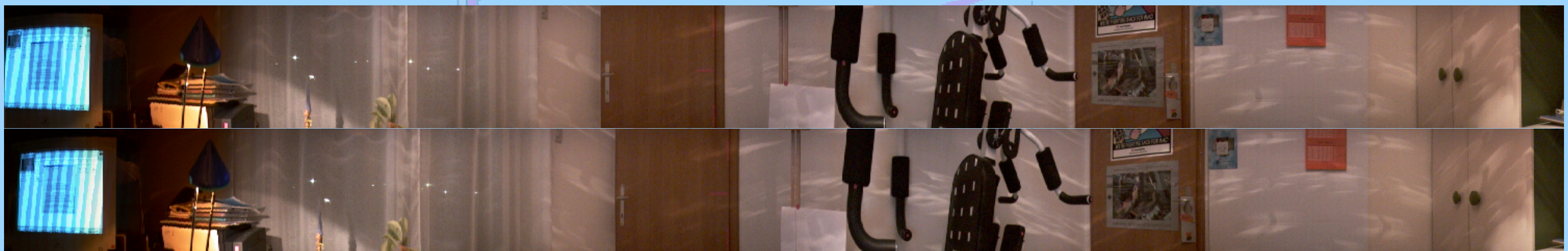


Generation of panoramic views

- Setoff of the camera's optical center
- **Multiperspective panoramic views #3 — Result of the acquisition**



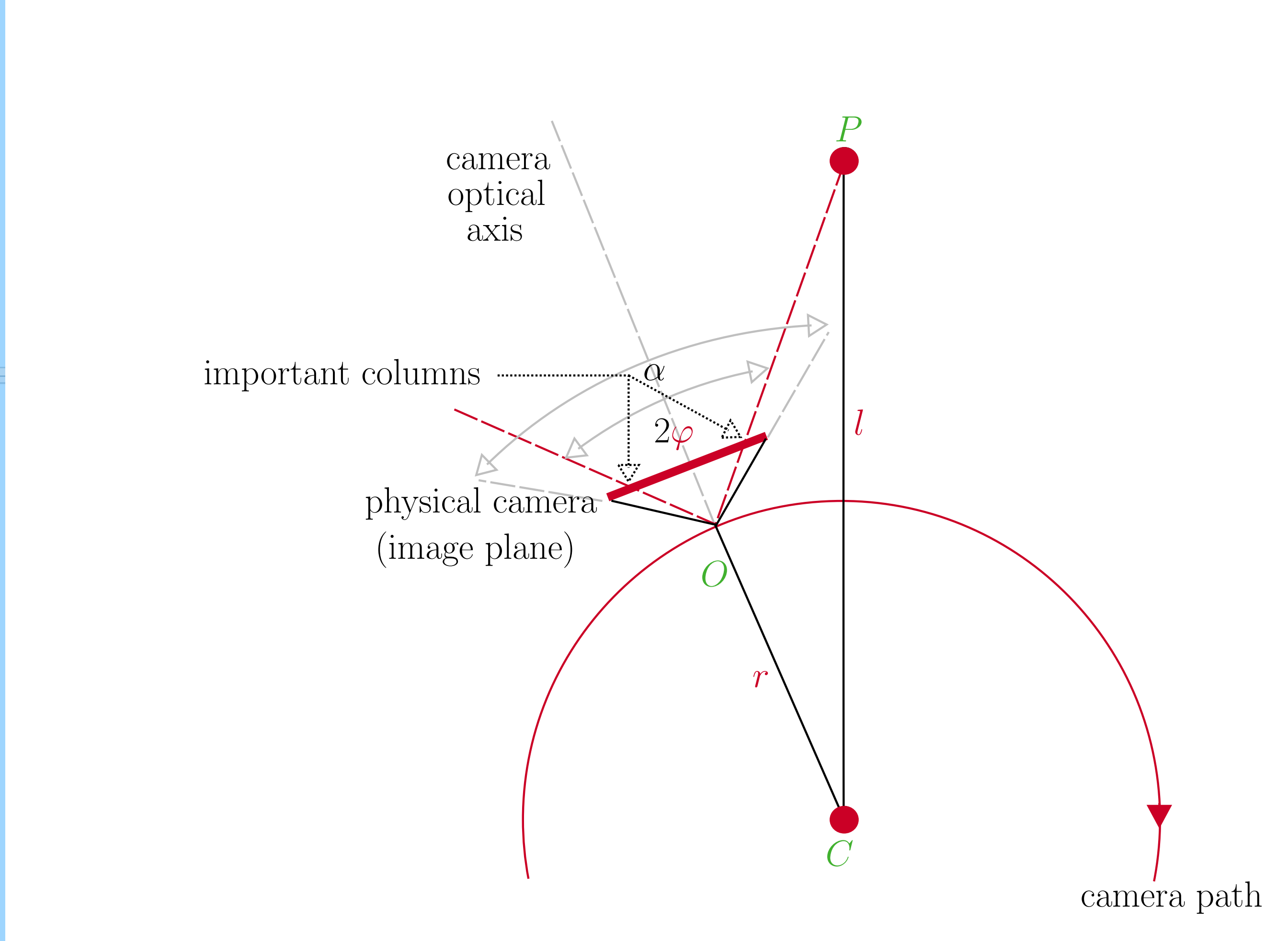
2φ=29.9625°



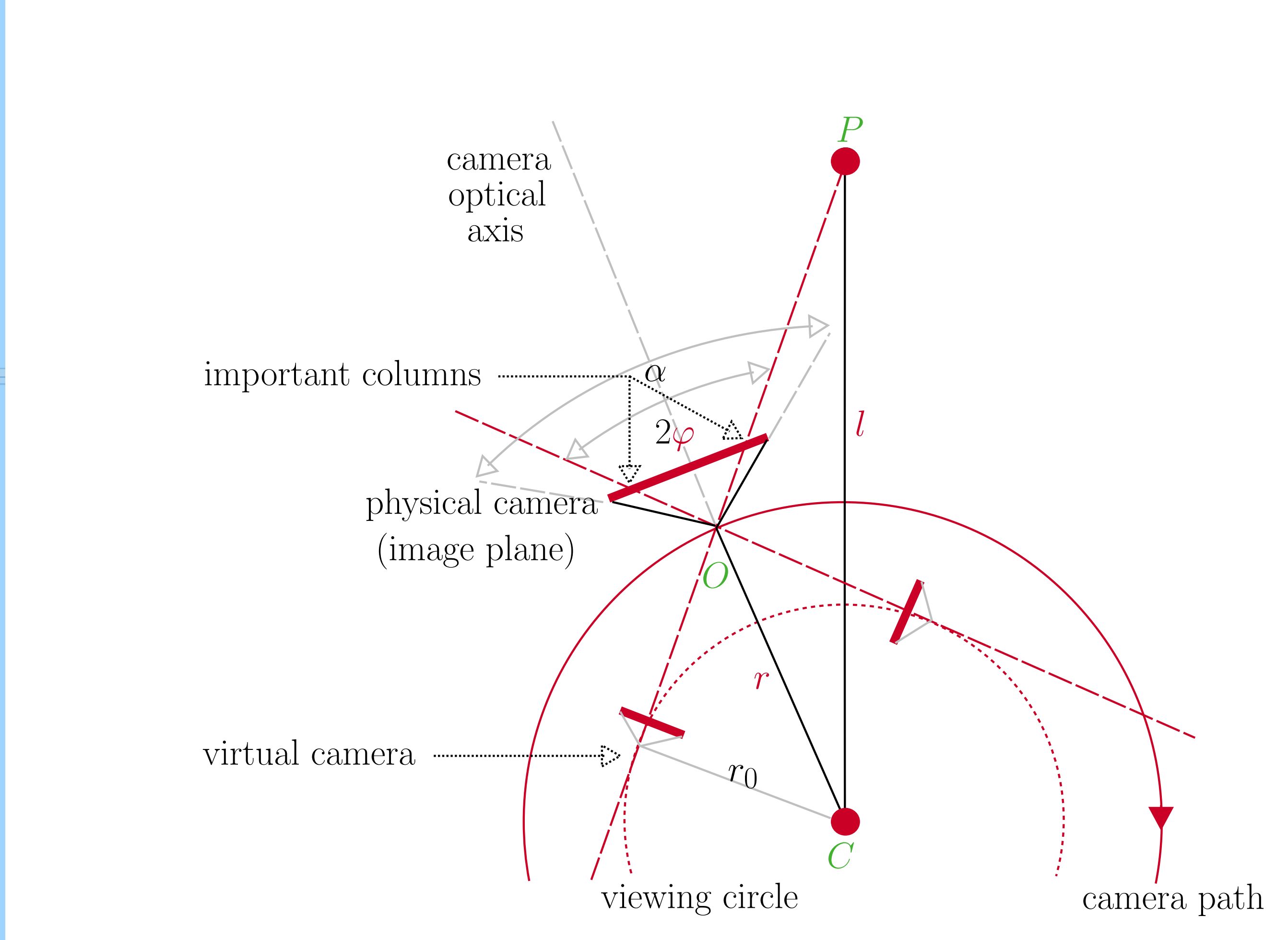
2φ=3.6125°

...and the stereo reconstruction is possible!

- **Multiperspective panoramic views #4 — Geometry of the system**

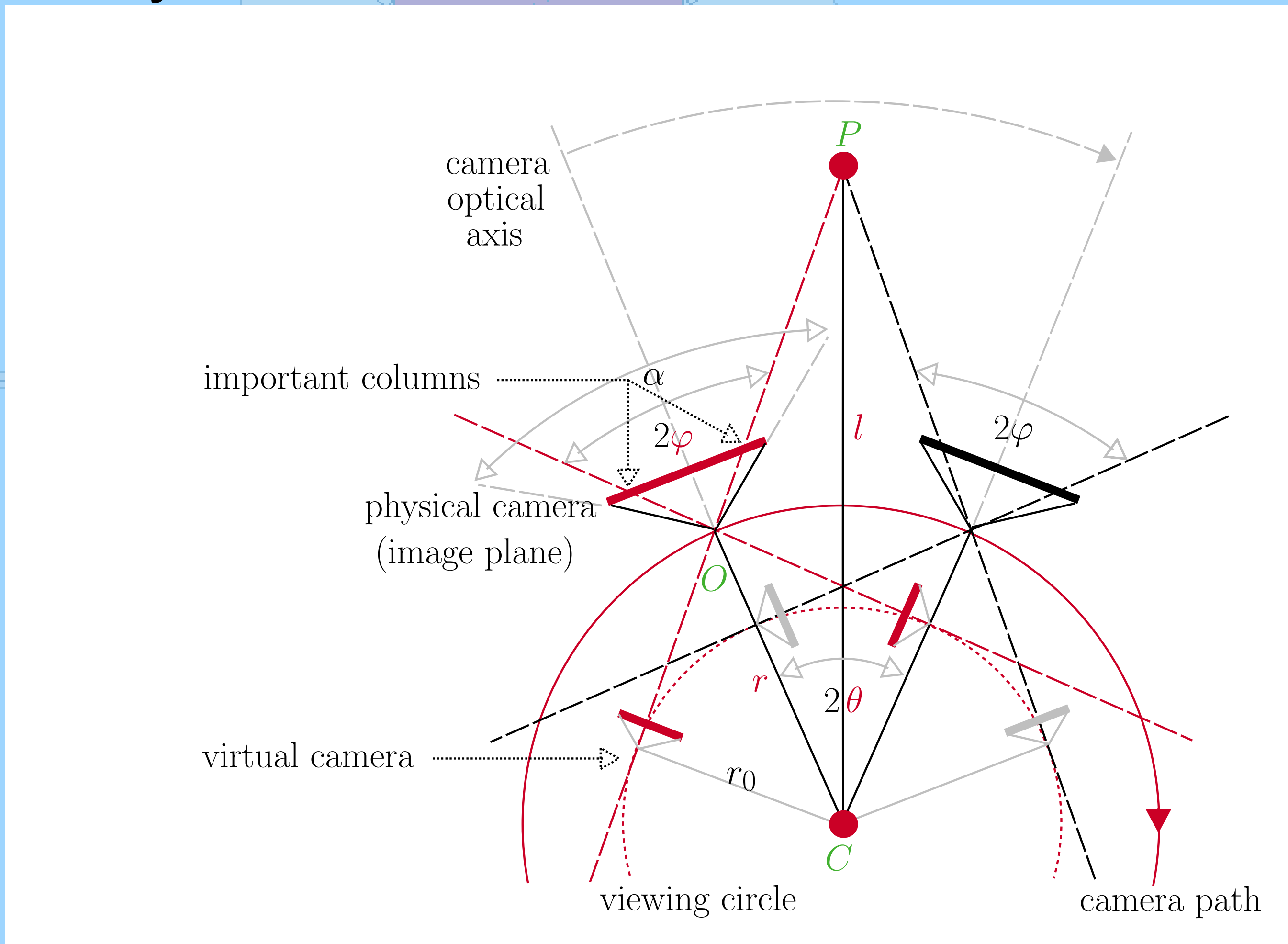


- **Multiperspective panoramic views #4 — Geometry of the system**



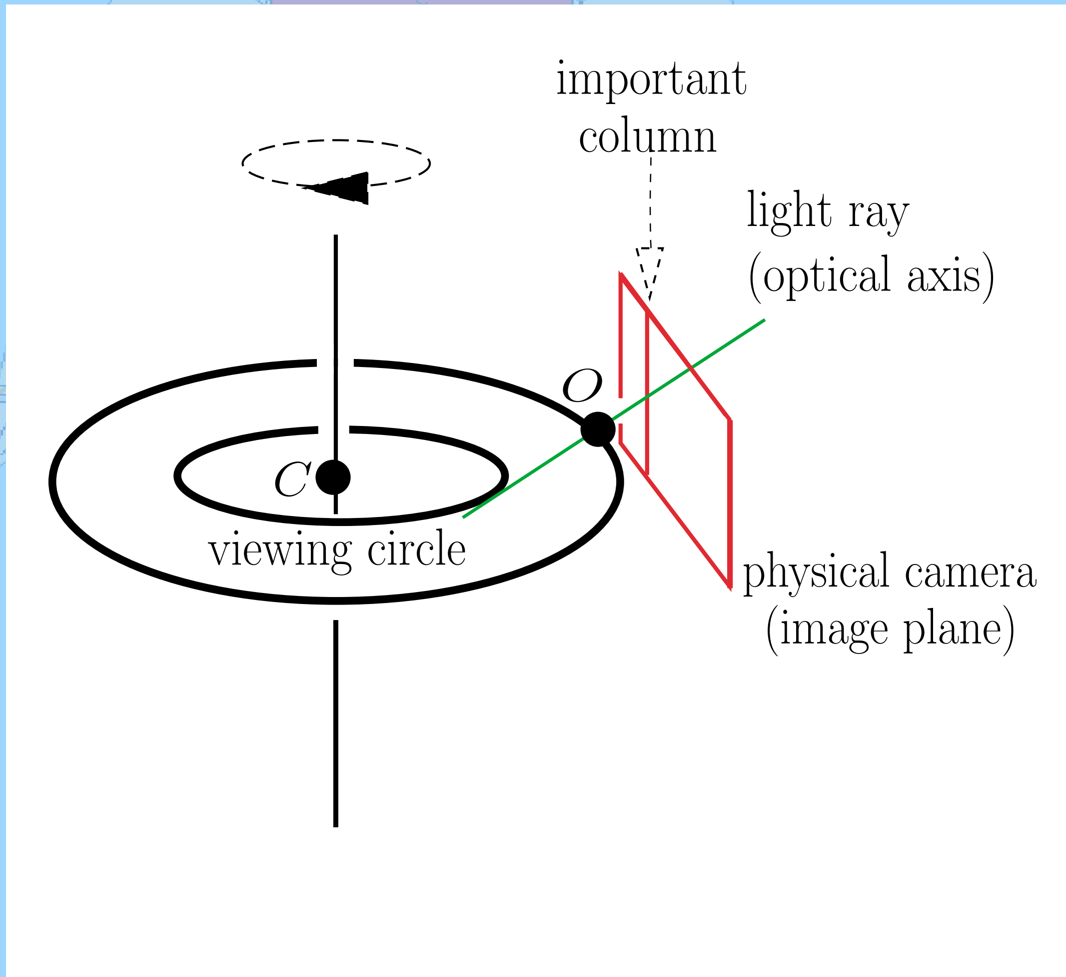
Non-central system

- **Multiperspective panoramic views #4 — Geometry of the system**

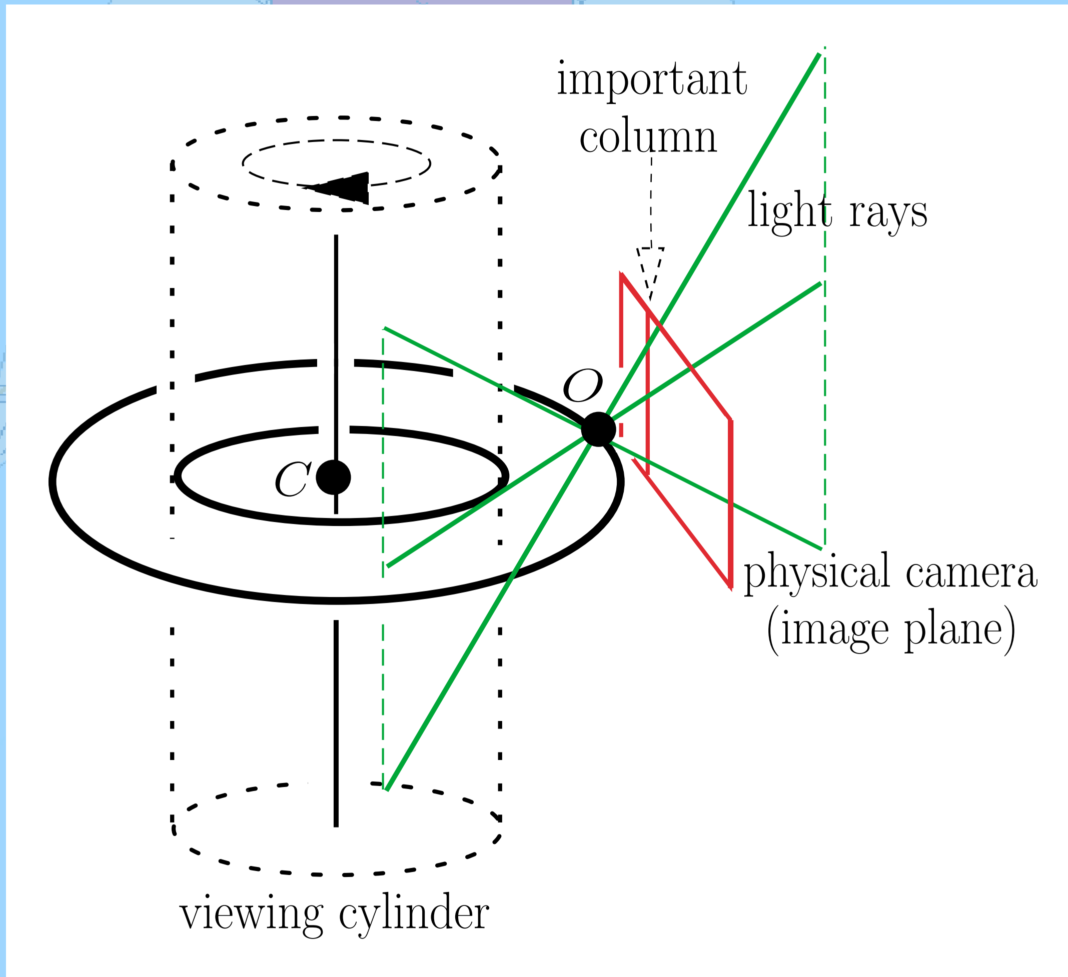


Non-central system

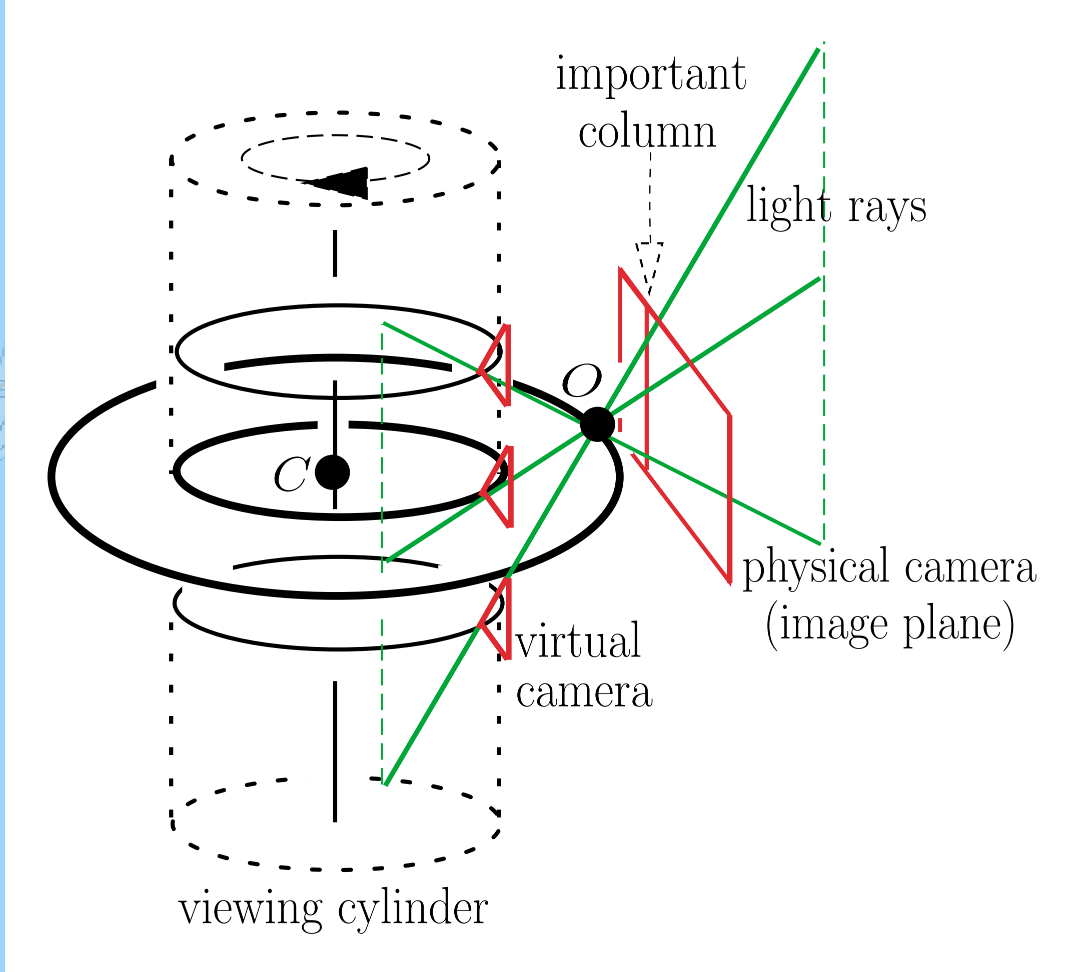
- **Multiperspective panoramic views #5 — Viewing cylinder**



- **Multiperspective panoramic views #5 — Viewing cylinder**

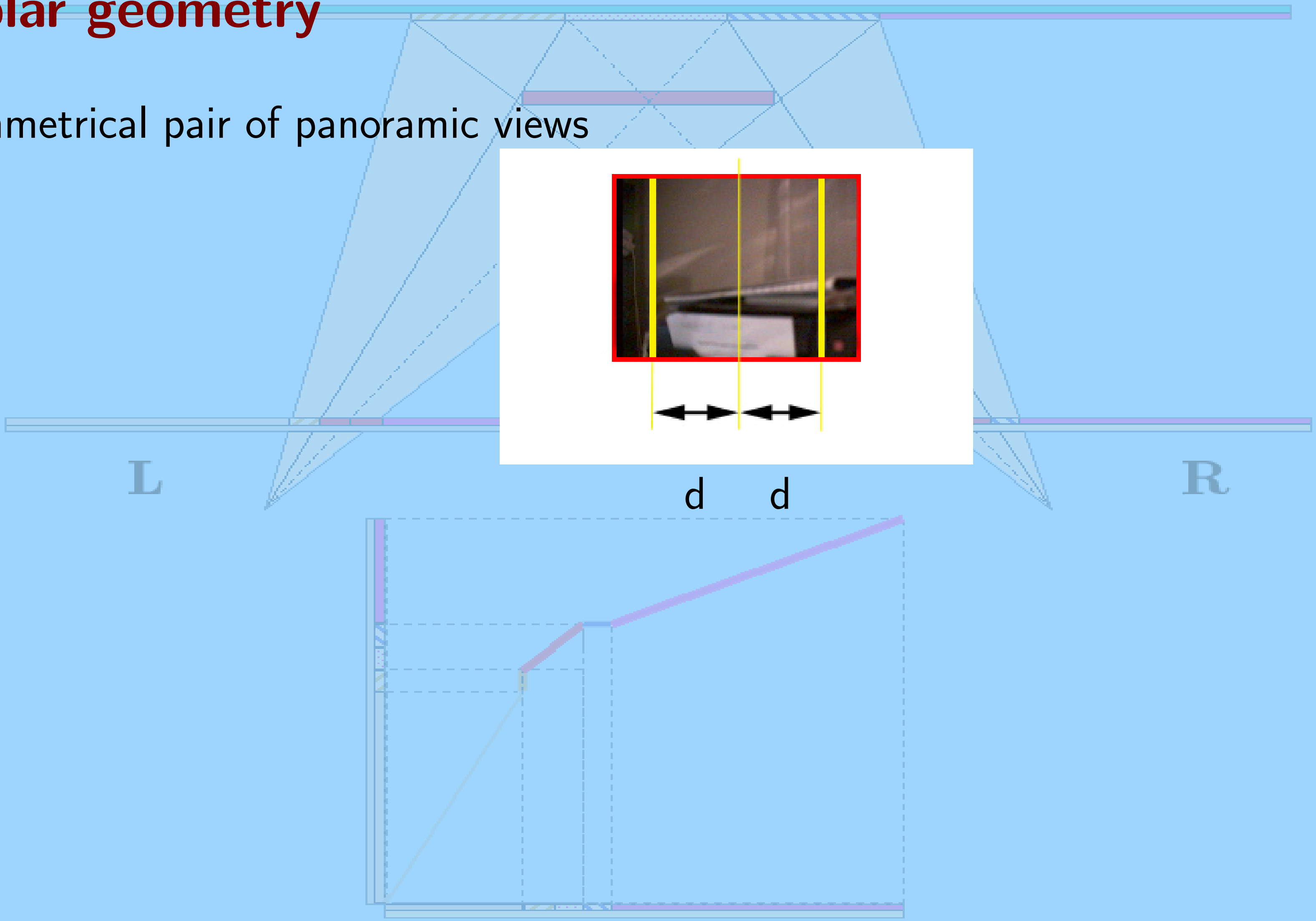


- **Multiperspective panoramic views #5 — Viewing cylinder**



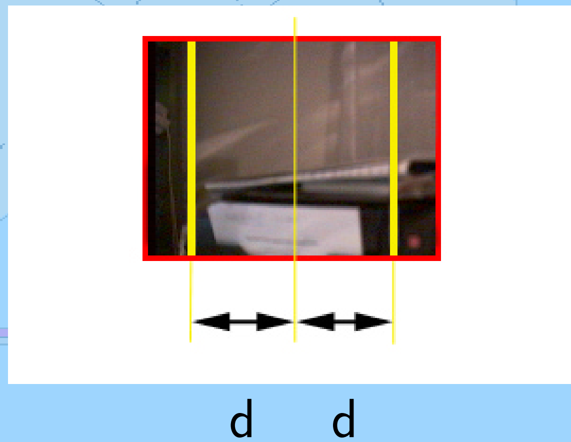
Epipolar geometry

Symmetrical pair of panoramic views



Epipolar geometry

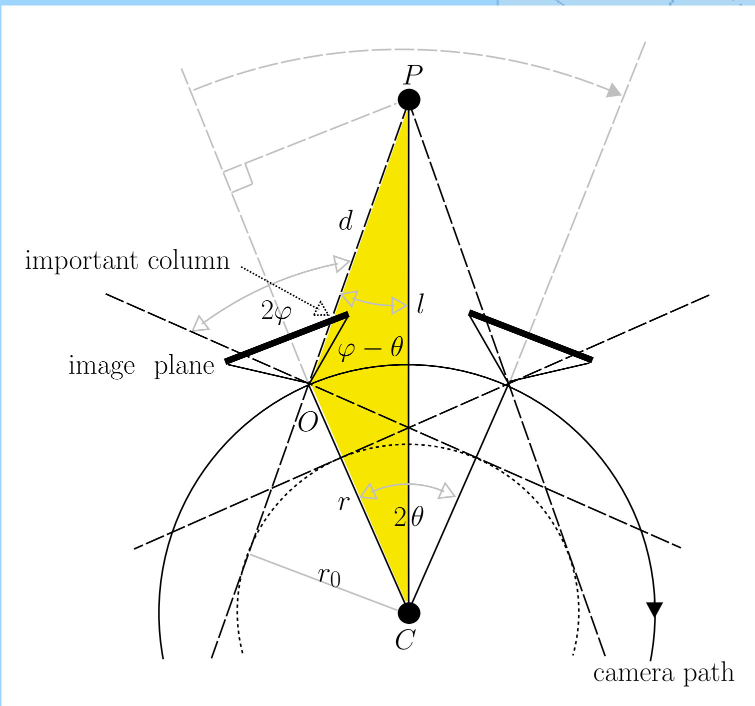
Symmetrical pair of panoramic views



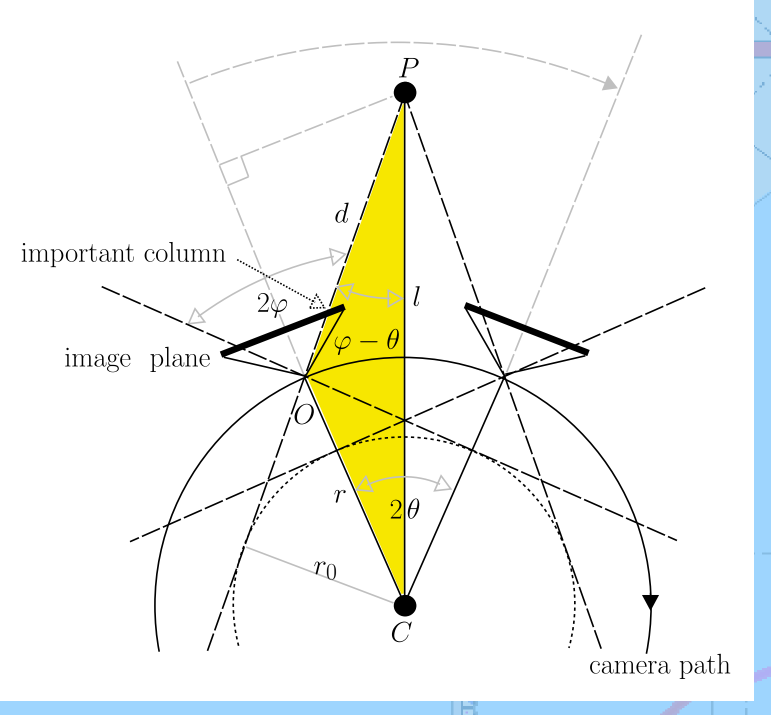
STATEMENT: In this case the epipolar lines are rows of panoramic views.

[Huang]

Stereo reconstruction

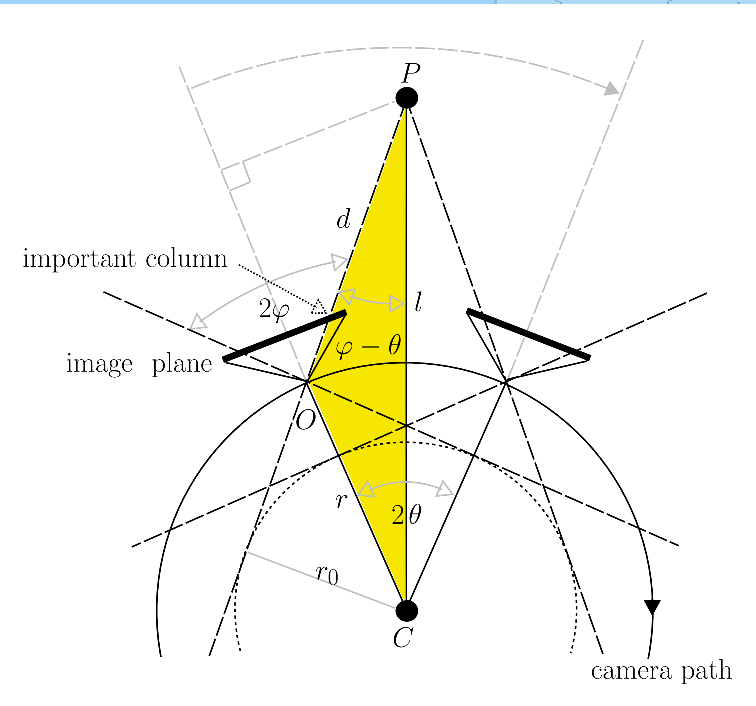


Stereo reconstruction



$$\frac{r}{\sin(\varphi-\theta)} = \frac{d}{\sin\theta} = \frac{l}{\sin(180^{o}-\varphi)}$$

Stereo reconstruction

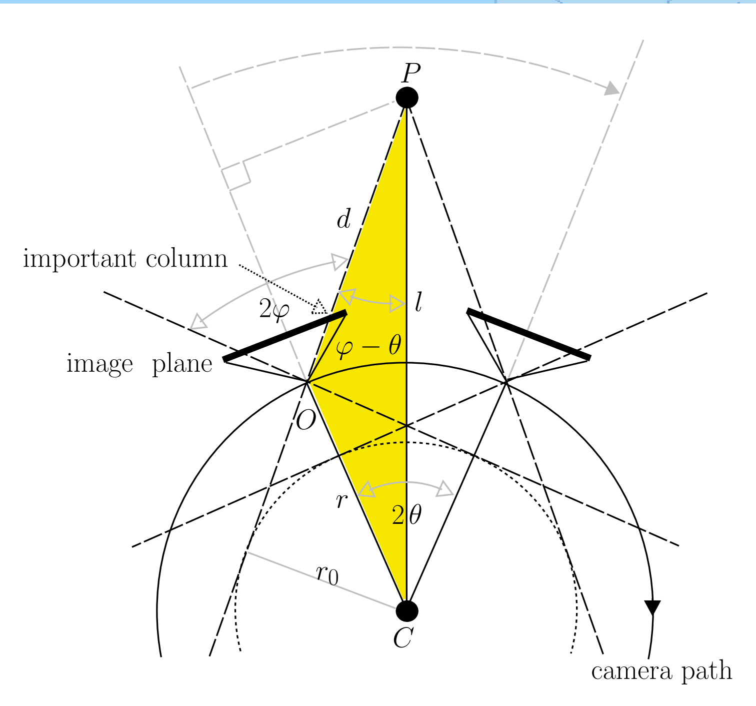


$$\frac{r}{\sin(\varphi-\theta)} = \frac{d}{\sin\theta} = \frac{l}{\sin(180^{\mathrm{o}}-\varphi)}$$

$$l = \frac{r\cdot\sin(180^{\mathrm{o}}-\varphi)}{\sin(\varphi-\theta)} = \frac{r\cdot\sin\varphi}{\sin(\varphi-\theta)}$$

r, φ and θ

Stereo reconstruction



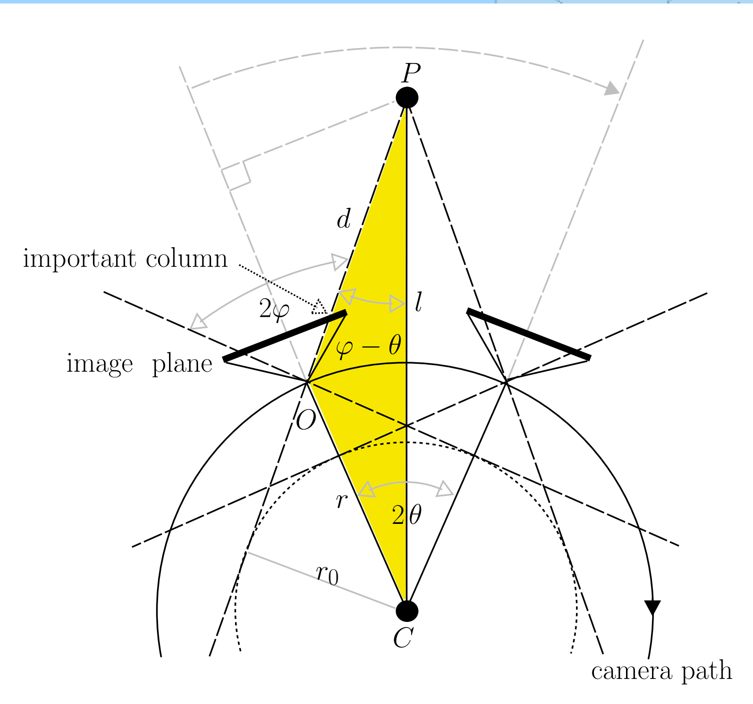
$$\frac{r}{\sin(\varphi-\theta)} = \frac{d}{\sin\theta} = \frac{l}{\sin(180^{\circ}-\varphi)}$$

$$\theta = dx \cdot \frac{\theta_0}{2}$$

$$l = \frac{r \cdot \sin(180^{\circ}-\varphi)}{\sin(\varphi-\theta)} = \frac{r \cdot \sin\varphi}{\sin(\varphi-\theta)}$$

r, φ and θ

Stereo reconstruction



$$\frac{r}{\sin(\varphi-\theta)} = \frac{d}{\sin\theta} = \frac{l}{\sin(180^{\circ}-\varphi)}$$

$$\theta = dx \cdot \frac{\theta_0}{2}$$

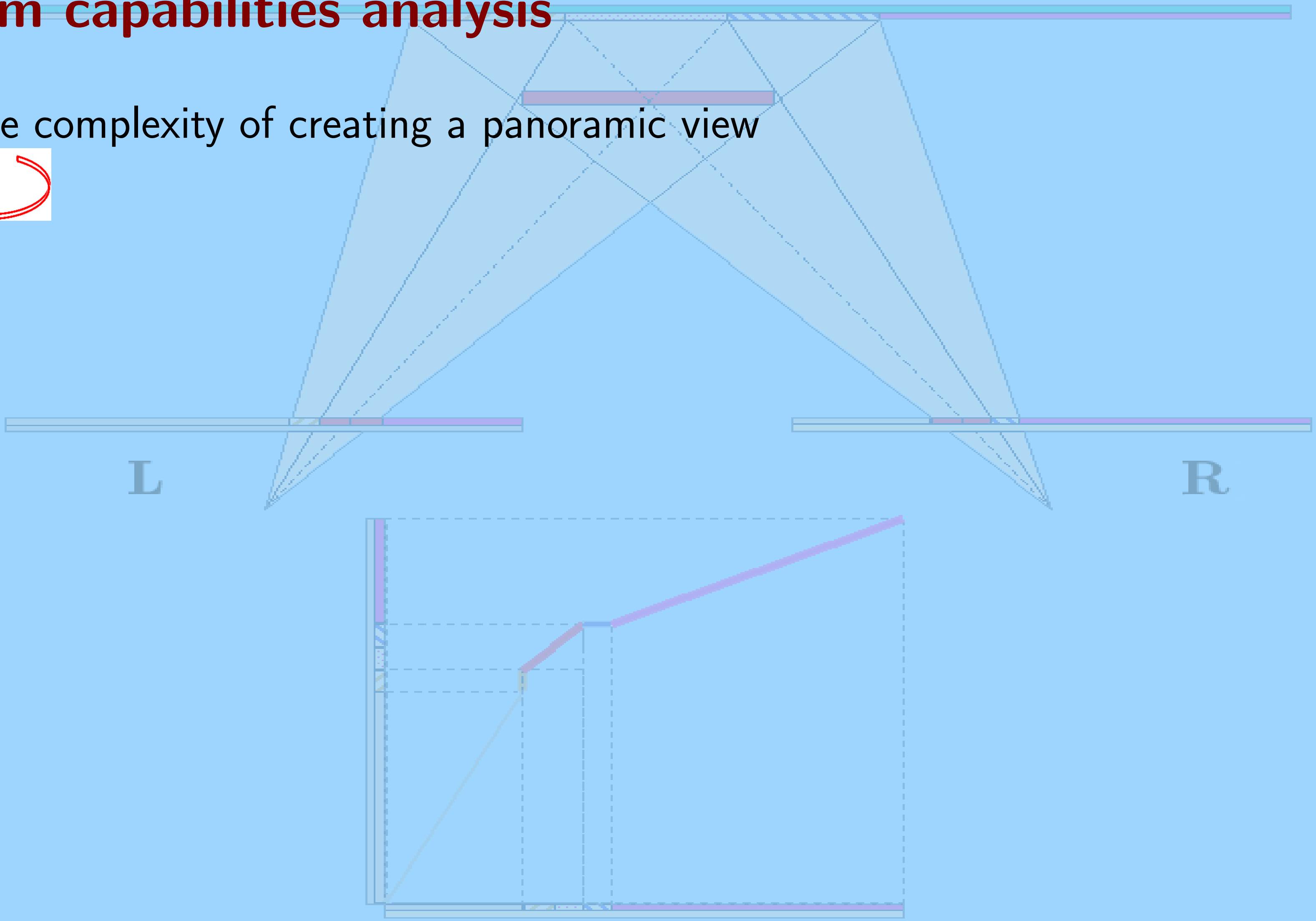
$$l = \frac{r\cdot\sin(180^{\circ}-\varphi)}{\sin(\varphi-\theta)} = \frac{r\cdot\sin\varphi}{\sin(\varphi-\theta)}$$

r, φ and θ

- Normalized correlation
- Back-correlation

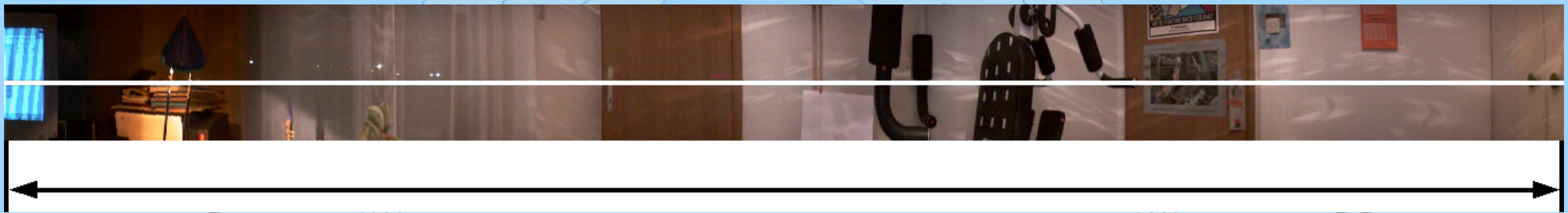
System capabilities analysis

- Time complexity of creating a panoramic view



System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line

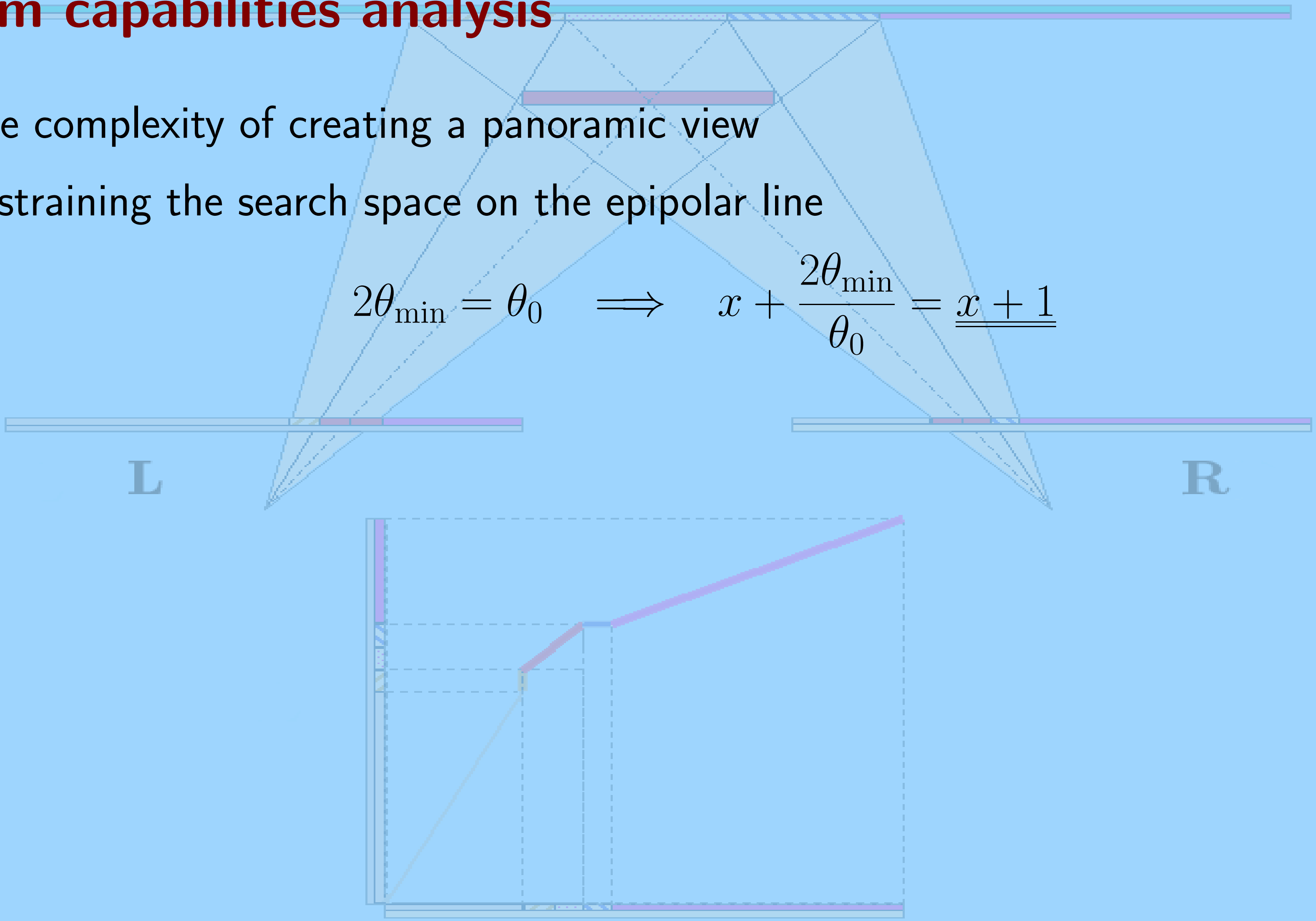


unconstrained length of the epipolar line: 1501 pixels

System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line

$$2\theta_{\min} = \theta_0 \quad \Longrightarrow \quad x + \frac{2\theta_{\min}}{\theta_0} = \underline{\underline{x+1}}$$

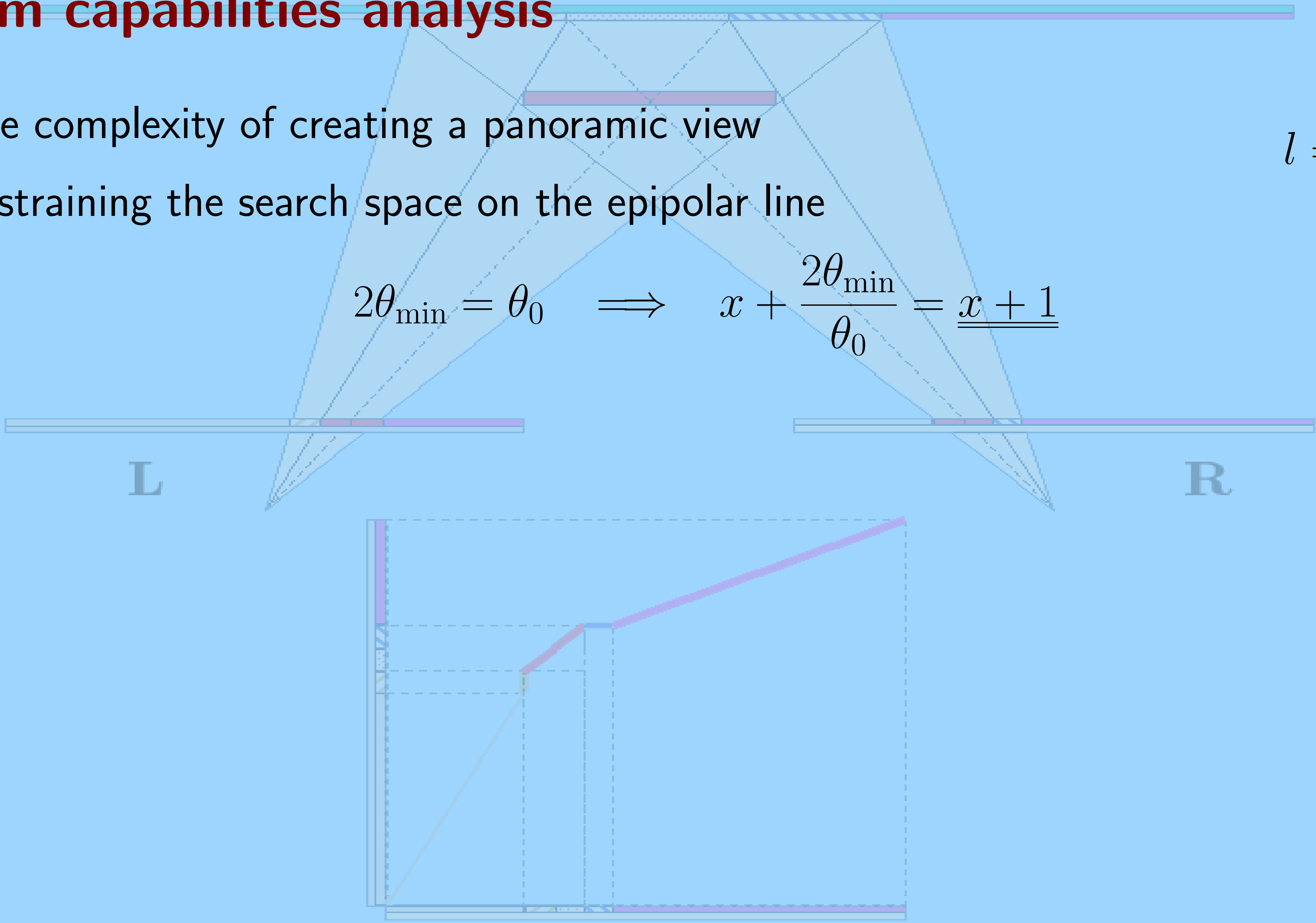


System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line

$$l = \frac{r \cdot \sin\varphi}{\sin(\varphi - \theta)}$$

$$2\theta_{\min} = \theta_0 \quad\Longrightarrow\quad x + \frac{2\theta_{\min}}{\theta_0} = \underline{\underline{x+1}}$$



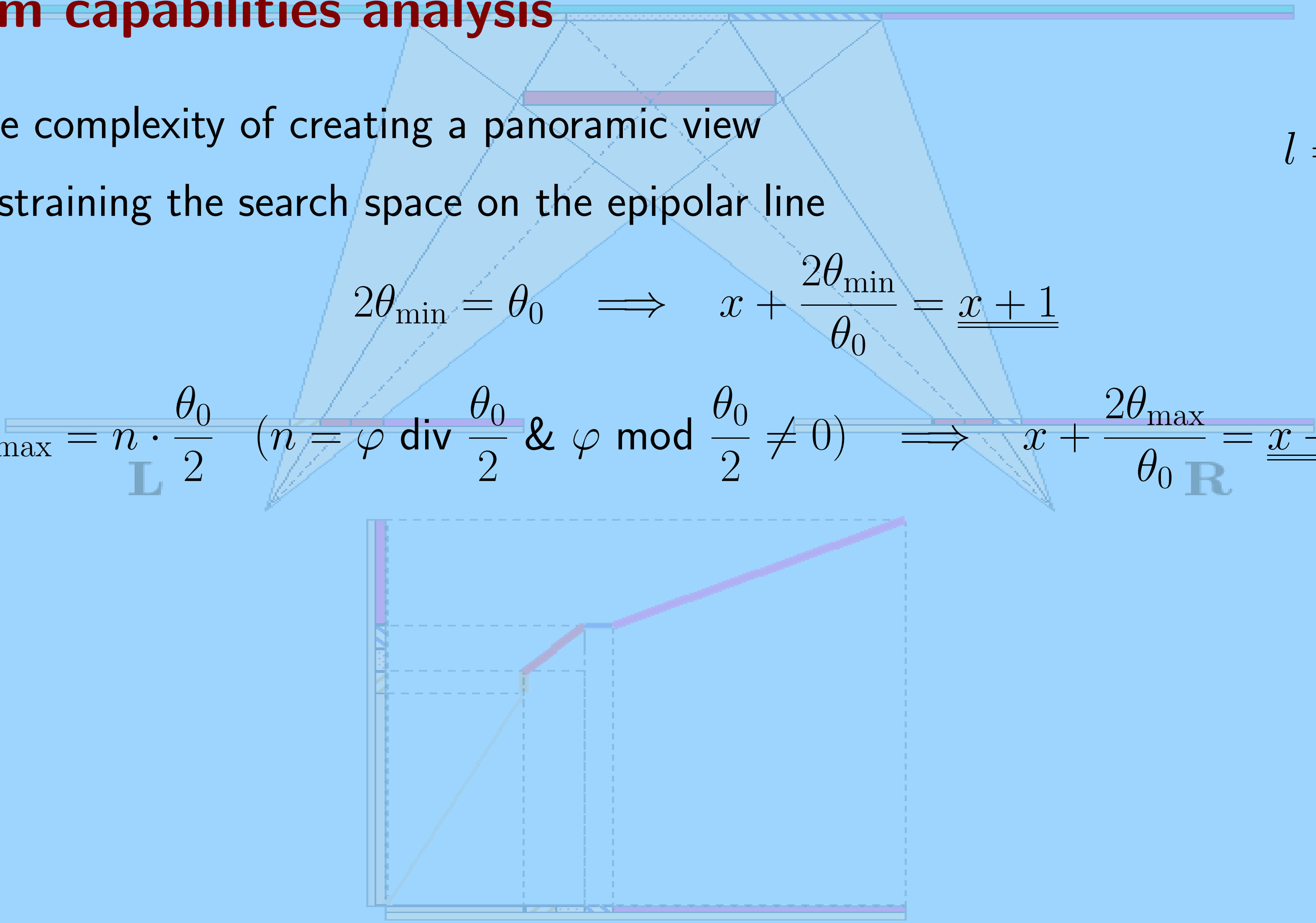
System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line

$$l = \frac{r \cdot \sin\varphi}{\sin(\varphi - \theta)}$$

$$2\theta_{\min} = \theta_0 \quad \Longrightarrow \quad x + \frac{2\theta_{\min}}{\theta_0} = \underline{\underline{x+1}}$$

$$\theta_{\max} = n \cdot \frac{\theta_0}{2} \quad \left(n = \varphi \ \mathsf{div}\ \frac{\theta_0}{2} \ \&\ \varphi \ \mathsf{mod}\ \frac{\theta_0}{2} \neq 0\right) \quad \Longrightarrow \quad x + \frac{2\theta_{\max}}{\theta_0} = \underline{\underline{x+n}}$$



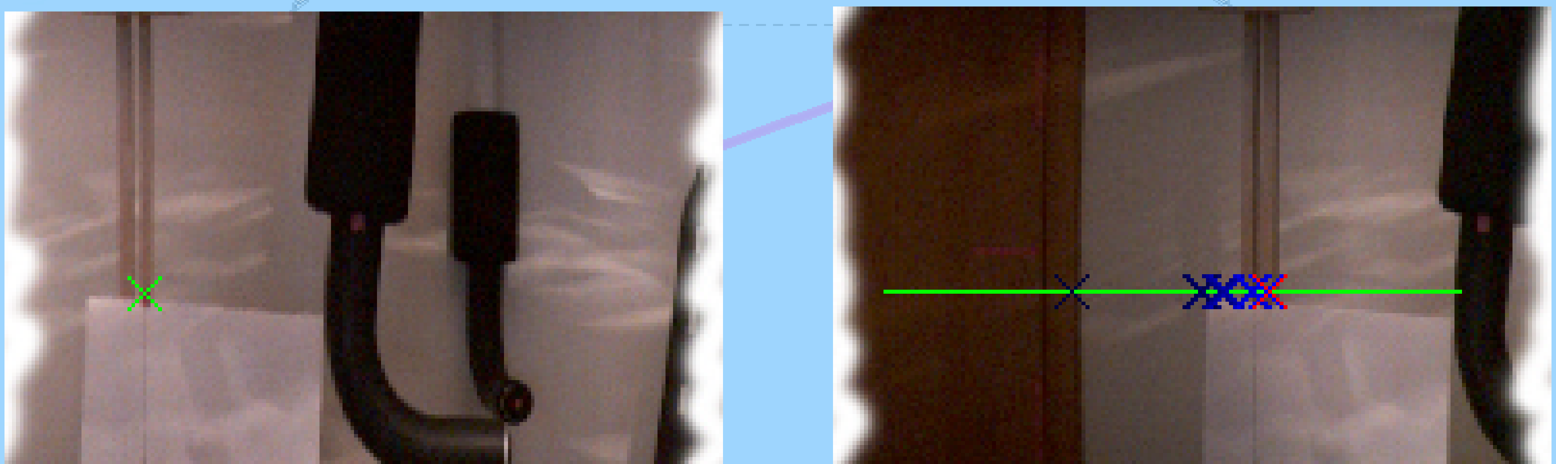
System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line

$$l = \frac{r \cdot \sin \varphi}{\sin(\varphi - \theta)}$$

$$2\theta_{\min} = \theta_0 \quad \Longrightarrow \quad x + \frac{2\theta_{\min}}{\theta_0} = \underline{\underline{x+1}}$$

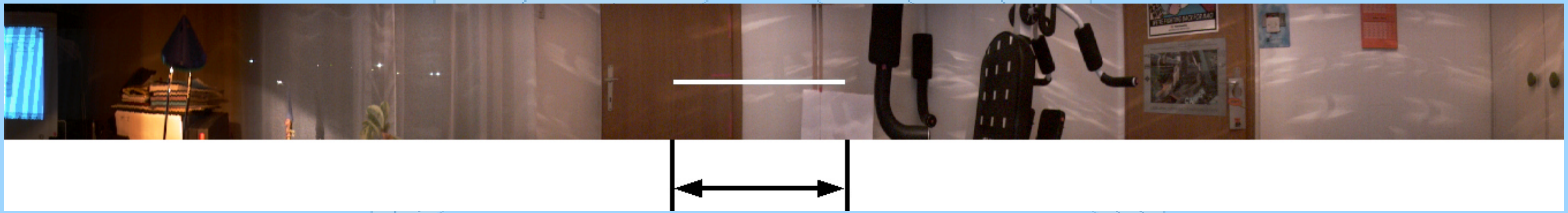
$$\theta_{\max} = n \cdot \frac{\theta_0}{2} \quad \left(n = \varphi \ \mathsf{div}\ \frac{\theta_0}{2} \ \&\ \varphi \ \mathsf{mod}\ \frac{\theta_0}{2} \neq 0\right) \quad \Longrightarrow \quad x + \frac{2\theta_{\max}}{\theta_0} = \underline{\underline{x+n}}$$



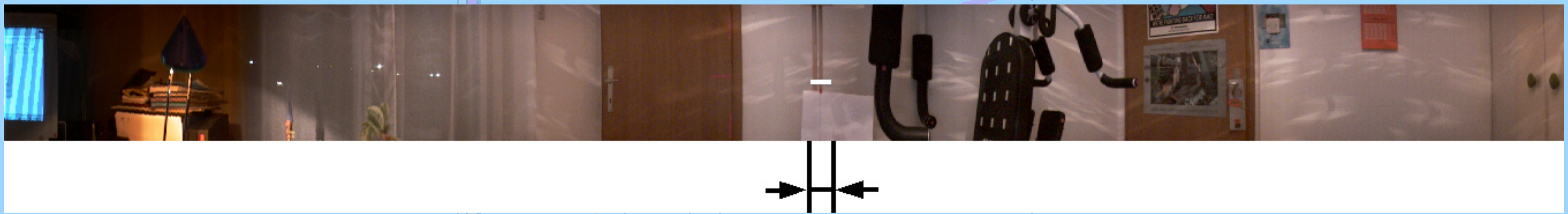
1501 pixels $\Longrightarrow n = \mathbf{149}$ pixels, $2\varphi = 29,9625^{\circ}$ / $n = 18$ pixels, $2\varphi = 3,6125^{\circ}$

System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line



a) constrained length of the epipolar line: **149** pixels, $2\varphi = 29.9625^{\circ}$

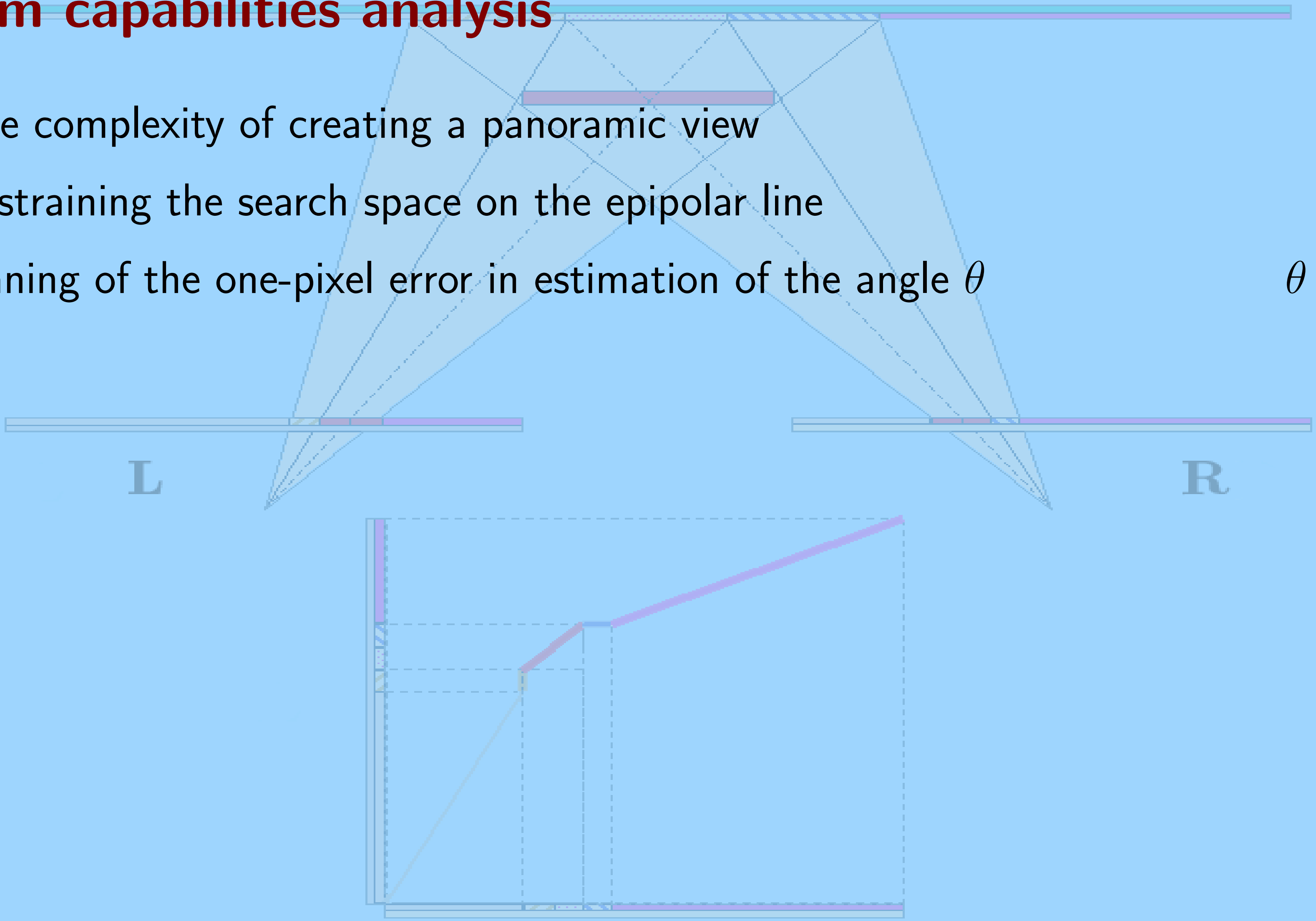


b) constrained length of the epipolar line: **18** pixels, $2\varphi = 3.6125^{\circ}$

System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ

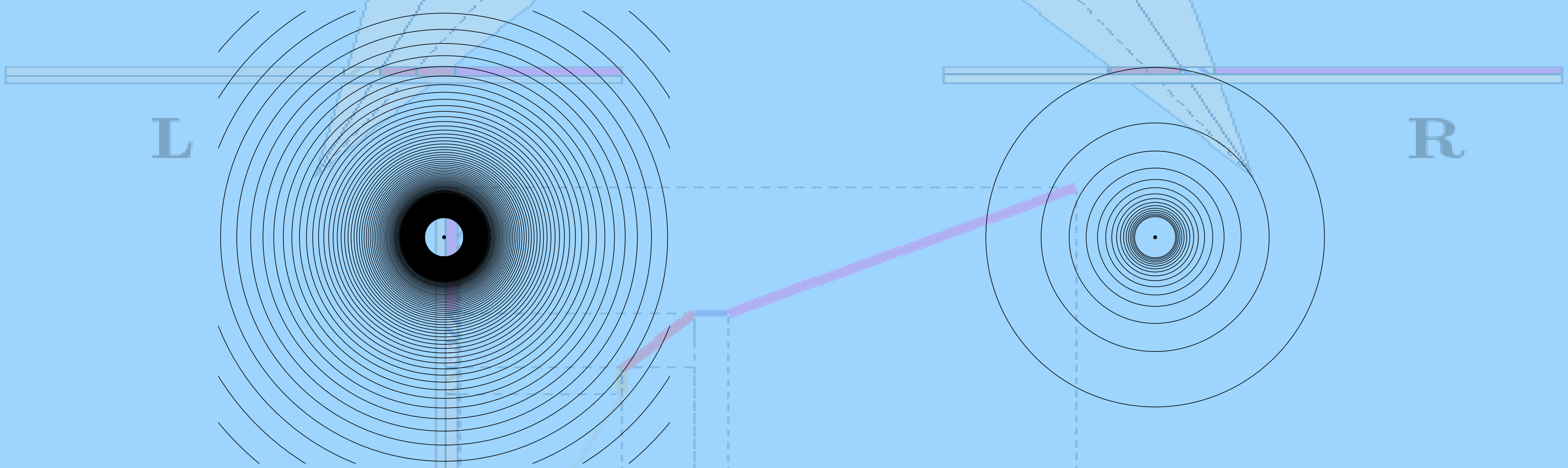
$$\theta = \underline{\underline{dx}} \cdot \frac{\theta_0}{2}$$



System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ

$$\theta = \underline{\underline{dx}} \cdot \frac{\theta_0}{2}$$



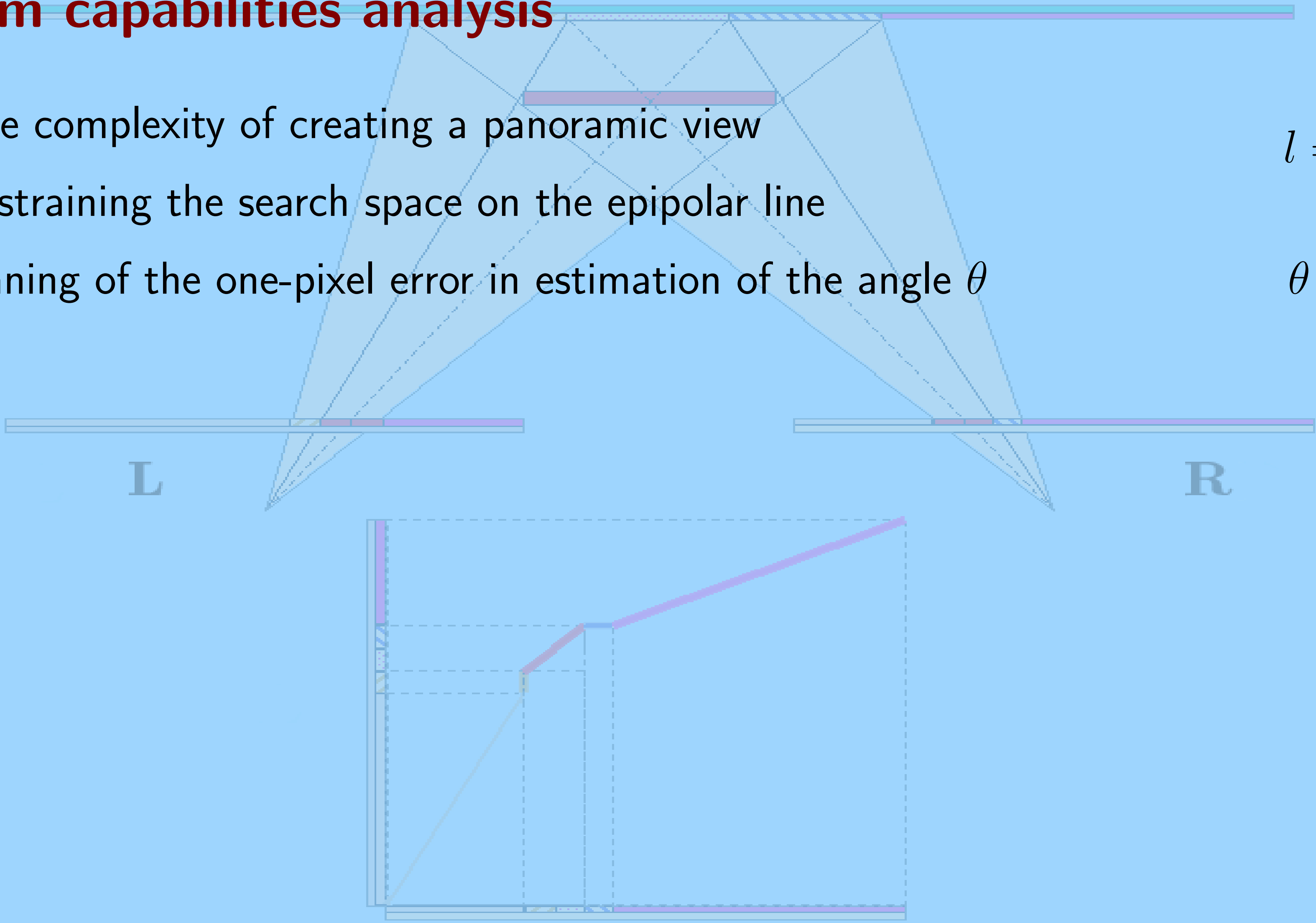
$2\varphi = 29,9625^{\mathrm{o}} \quad \Longrightarrow \quad$ **149** depths $\qquad 2\varphi = 3,6125^{\mathrm{o}} \quad \Longrightarrow \quad$ **18** depths

System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ

$$l = \frac{r \cdot \sin\varphi}{\sin(\varphi - \theta)}$$

$$\theta = \underline{\underline{dx}} \cdot \frac{\theta_0}{2}$$

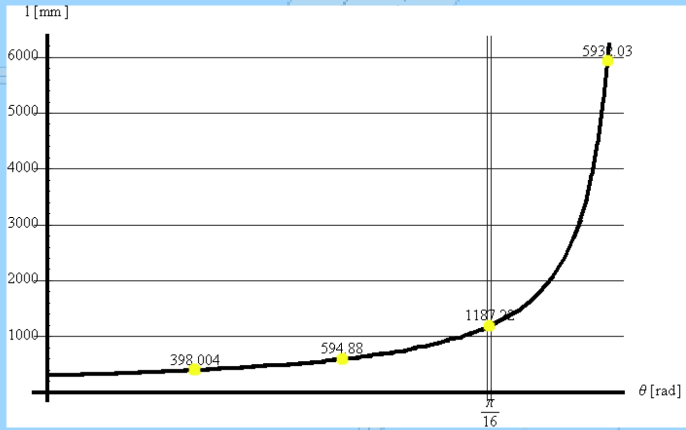


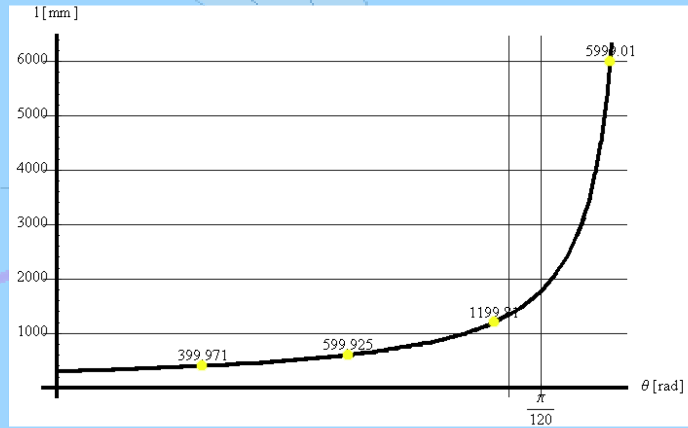
System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ

$$l = \frac{r \cdot \sin\varphi}{\sin(\varphi - \theta)}$$

$$\theta = \underline{\underline{dx}} \cdot \frac{\theta_0}{2}$$





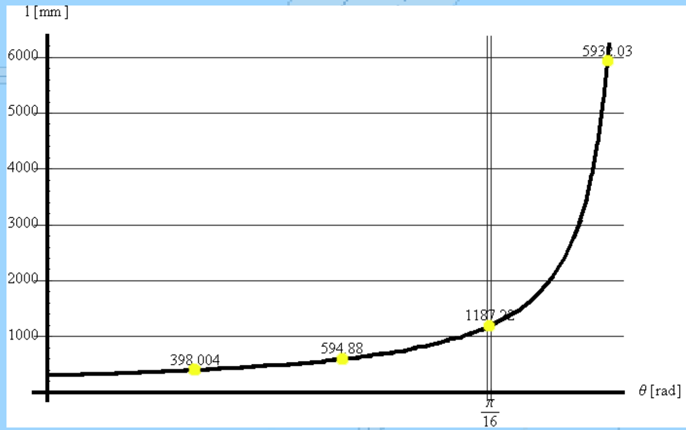
$2\varphi = 29,9625^{\circ} \implies$ **149** depths $\quad 2\varphi = 3,6125^{\circ} \implies$ **18** depths

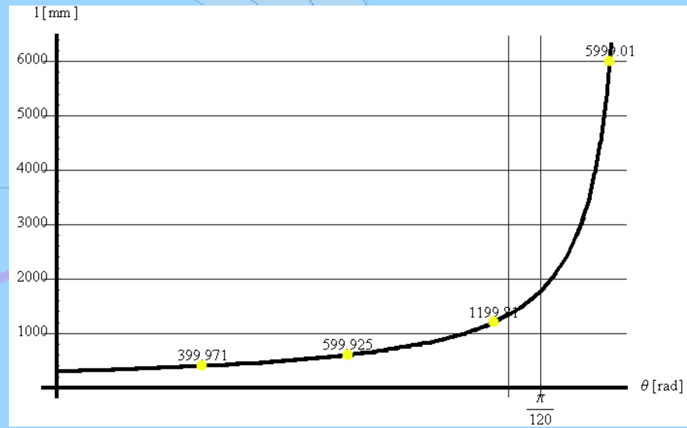
System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ

$$l = \frac{r \cdot \sin\varphi}{\sin(\varphi - \theta)}$$

$$\theta = \underline{\underline{dx}} \cdot \frac{\theta_0}{2}$$



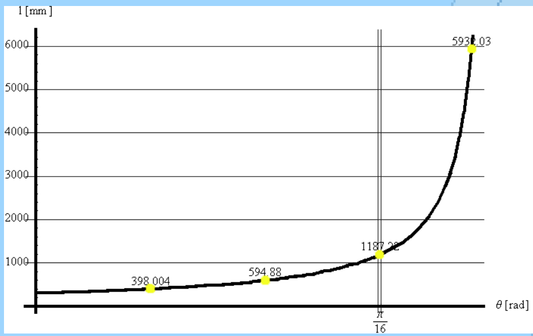


$2\varphi = 29,9625^{\circ} \implies$ **149** depths $\quad 2\varphi = 3,6125^{\circ} \implies$ **18** depths

reconstruction speed $\propto$ 1 / reconstruction accuracy

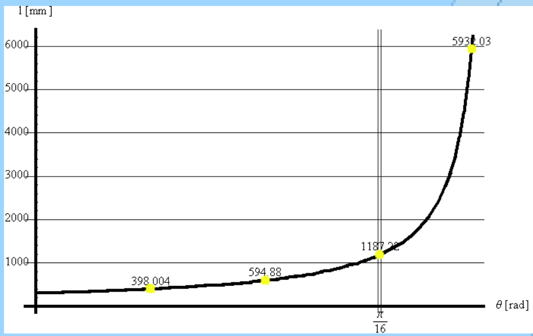
System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ
- Definition of the maximal reliable depth value



System capabilities analysis

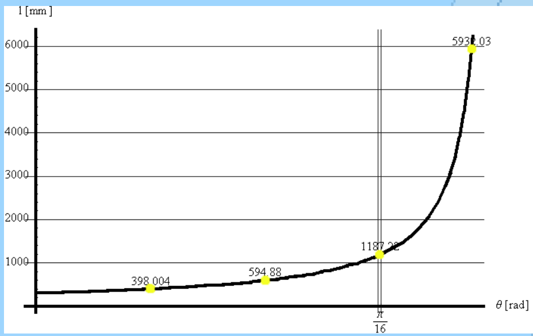
- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ
- Definition of the maximal reliable depth value



	$2\varphi = 29.9625^{\circ}$	$2\varphi = 3.6125^{\circ}$
$\Delta l_{\min}$ [mm]	2	19
$\Delta l_{\max}$ [mm]	**30172**	**81587**

System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ
- Definition of the maximal reliable depth value

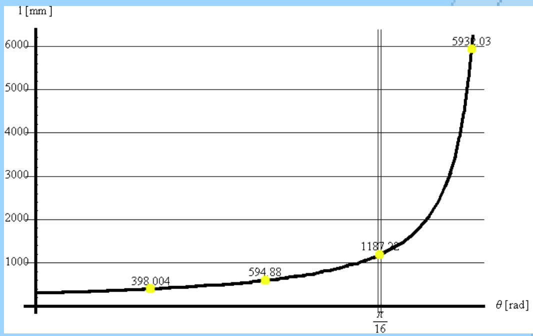


	$2\varphi = 29.9625^o$	$2\varphi = 3.6125^o$
$\Delta l_{\min}$ [mm]	2	19
$\Delta l_{\max}$ [mm]	**30172**	**81587**

define the upper boundary of allowed error size (Δl)

System capabilities analysis

- Time complexity of creating a panoramic view
- Constraining the search space on the epipolar line
- Meaning of the one-pixel error in estimation of the angle θ
- Definition of the maximal reliable depth value



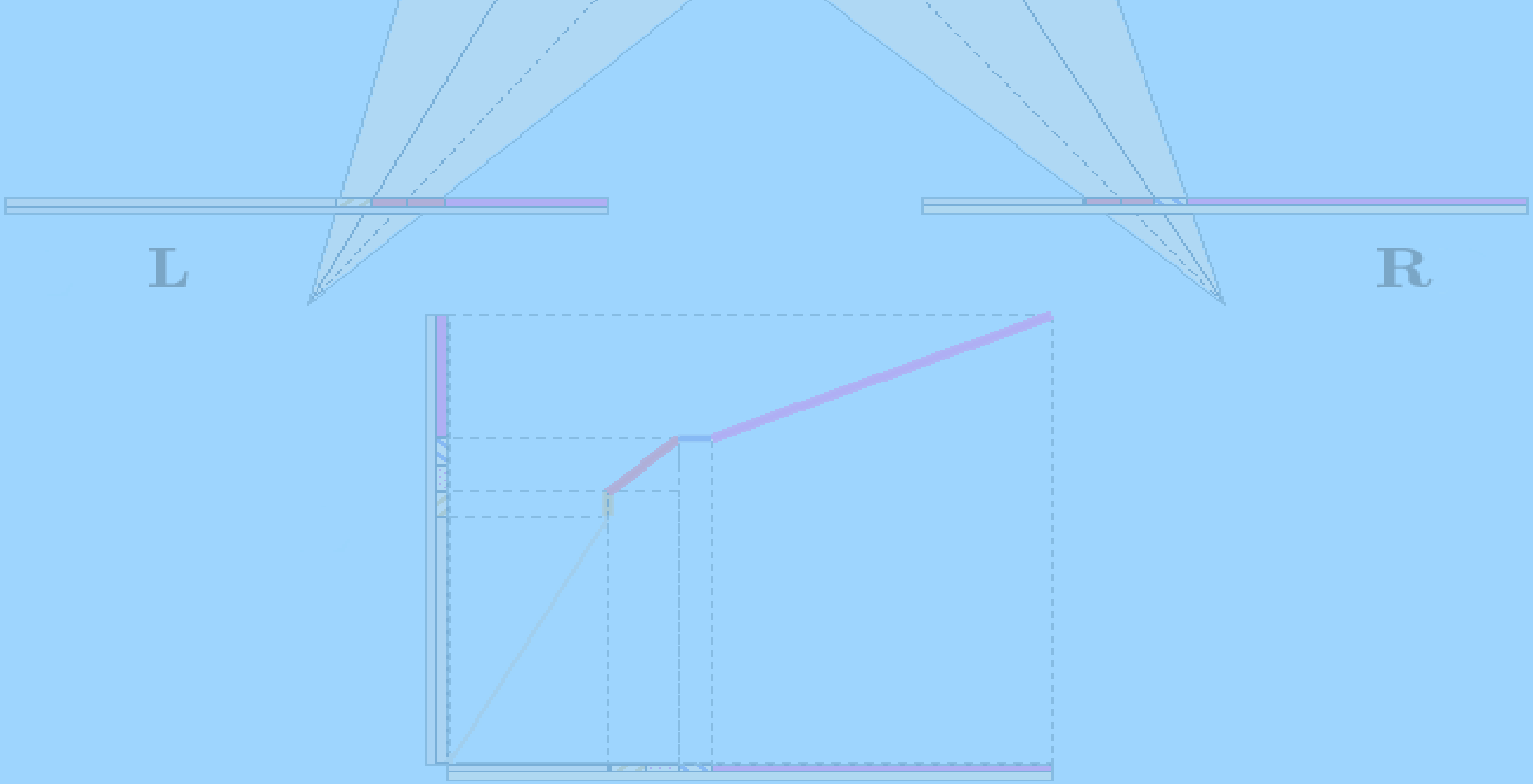
	$2\varphi = 29.9625^{\circ}$	$2\varphi = 3.6125^{\circ}$
$\Delta l_{\min}$ [mm]	2	19
$\Delta l_{\max}$ [mm]	**30172**	**81587**

define the upper boundary of allowed error size (Δl)

$\Longrightarrow$ **reconstruction of small spaces**

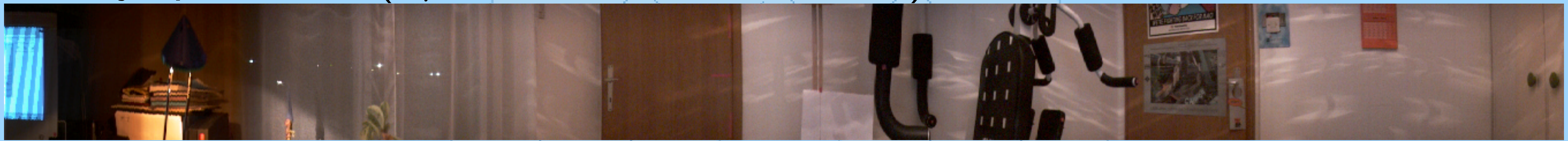
Experimental results #1 — Depth images

a) Left eye panorama ($2\varphi = 29.9625^{\mathrm{o}}$, $r = 30$ cm)

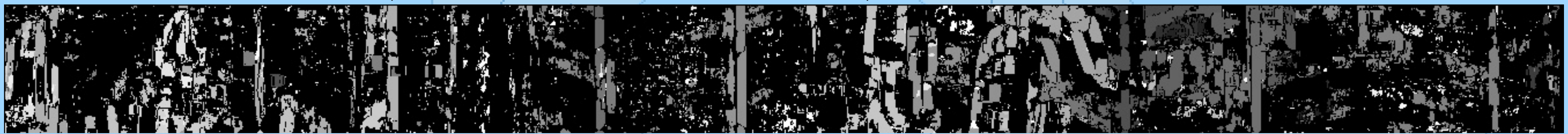


Experimental results #1 — Depth images

a) Left eye panorama ($2\varphi = 29.9625^o$, $r = 30$ cm)

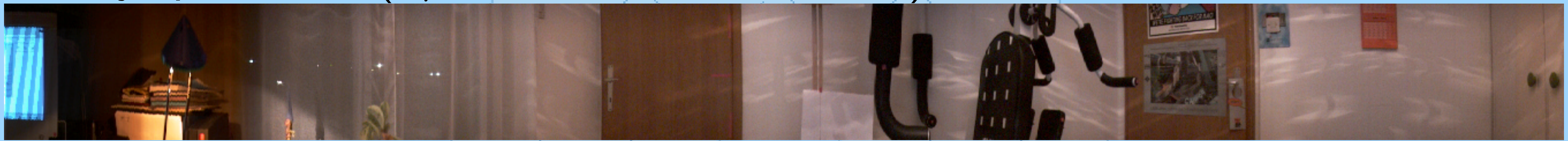


b) Dense depth image / using back-correlation / $t \doteq 6,7$ hours

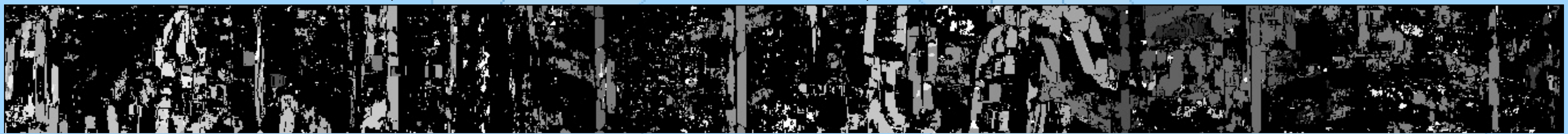


Experimental results #1 — Depth images

a) Left eye panorama ($2\varphi = 29.9625^{\circ}$, $r = 30$ cm)



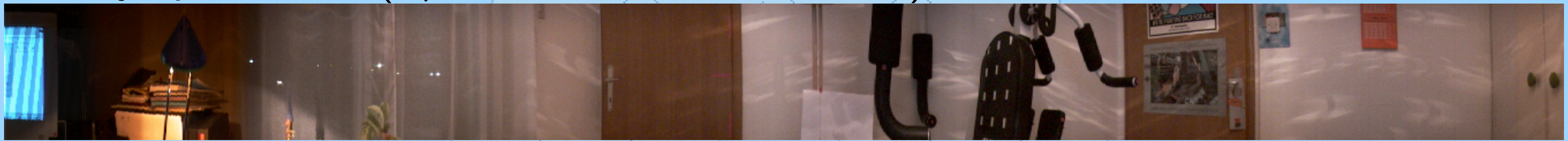
b) Dense depth image / using back-correlation / $t \doteq 6,7$ hours



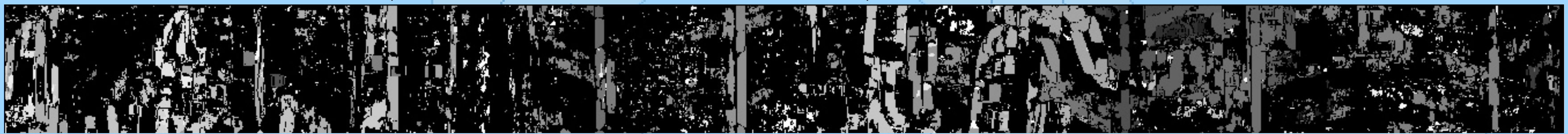
c) Information about the confidence in the estimated depth

Experimental results #1 — Depth images

a) Left eye panorama ($2\varphi = 29.9625^{\circ}$, $r = 30$ cm)



b) Dense depth image / using back-correlation / $t \doteq 6,7$ hours

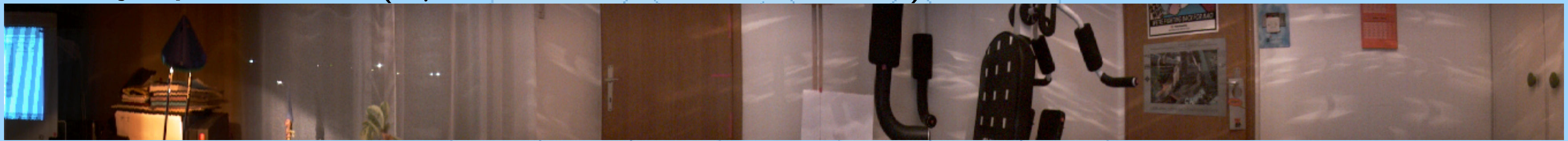


c) Information about the confidence in the estimated depth

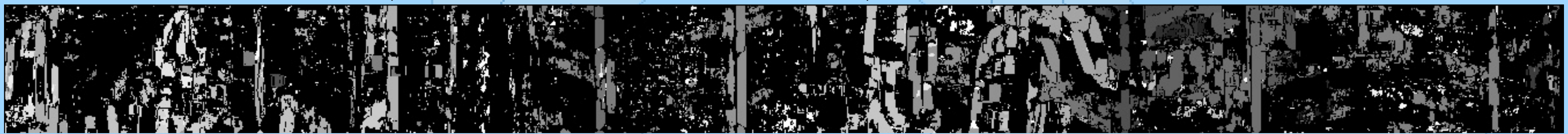
d) Dense depth image after the weighting / no back-correlation / $t \doteq 3,35$ h

Experimental results #1 — Depth images

a) Left eye panorama ($2\varphi = 29.9625^{\circ}$, $r = 30$ cm)



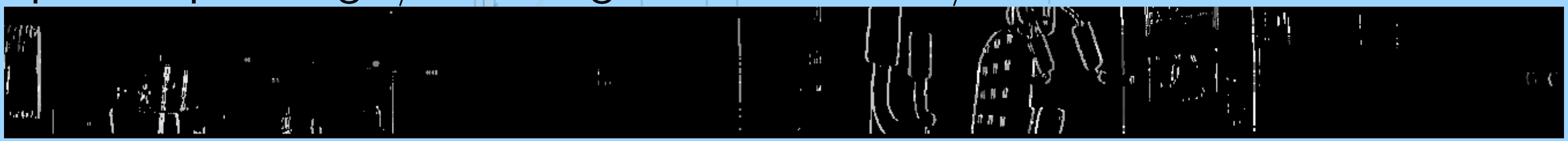
b) Dense depth image / using back-correlation / $t \doteq 6,7$ hours



c) Information about the confidence in the estimated depth

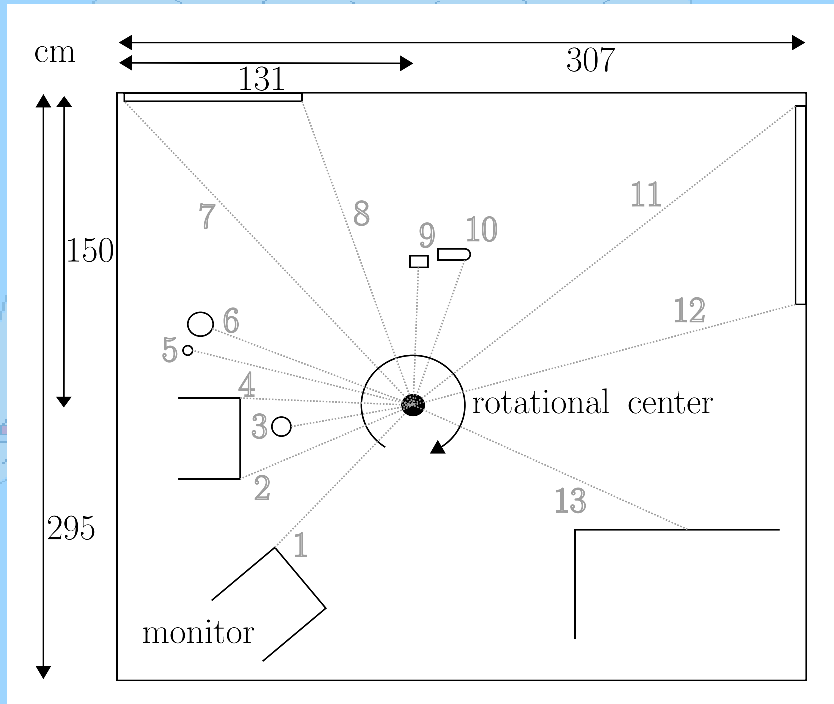
d) Dense depth image after the weighting / no back-correlation / $t \doteq 3,35$ h

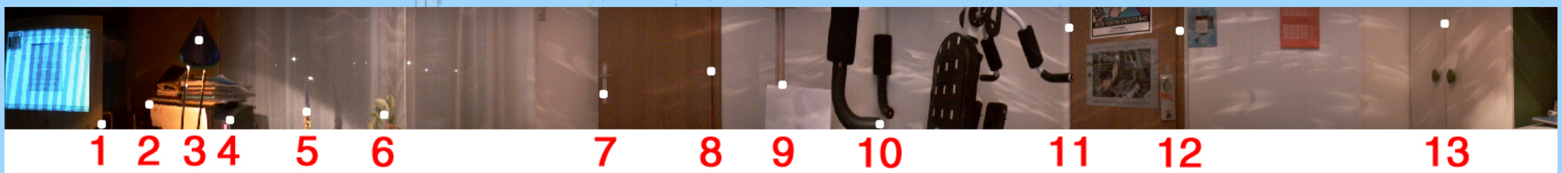
e) Sparse depth image / not using back-correlation / $t = \mathbf{38}$ **sec.**



[Together with the Sobel filter]

Experimental results #2 — Evaluation

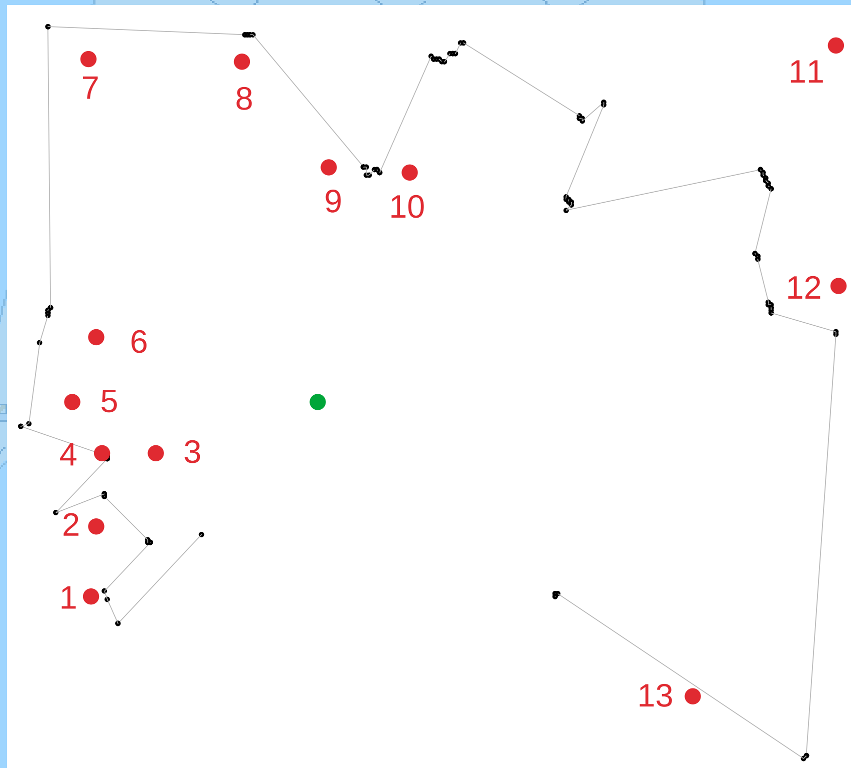


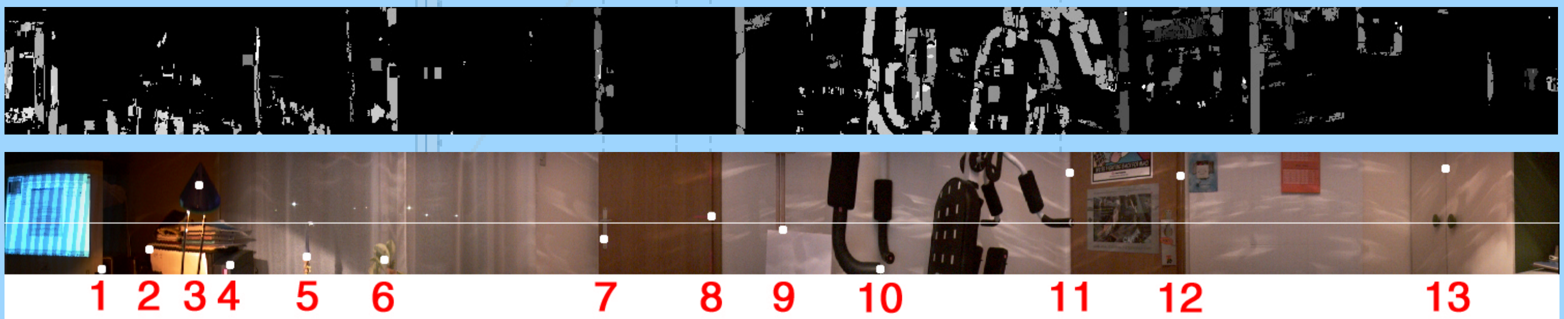


ground-plans of the scene $\Longrightarrow$

Experimental results #2 — Evaluation

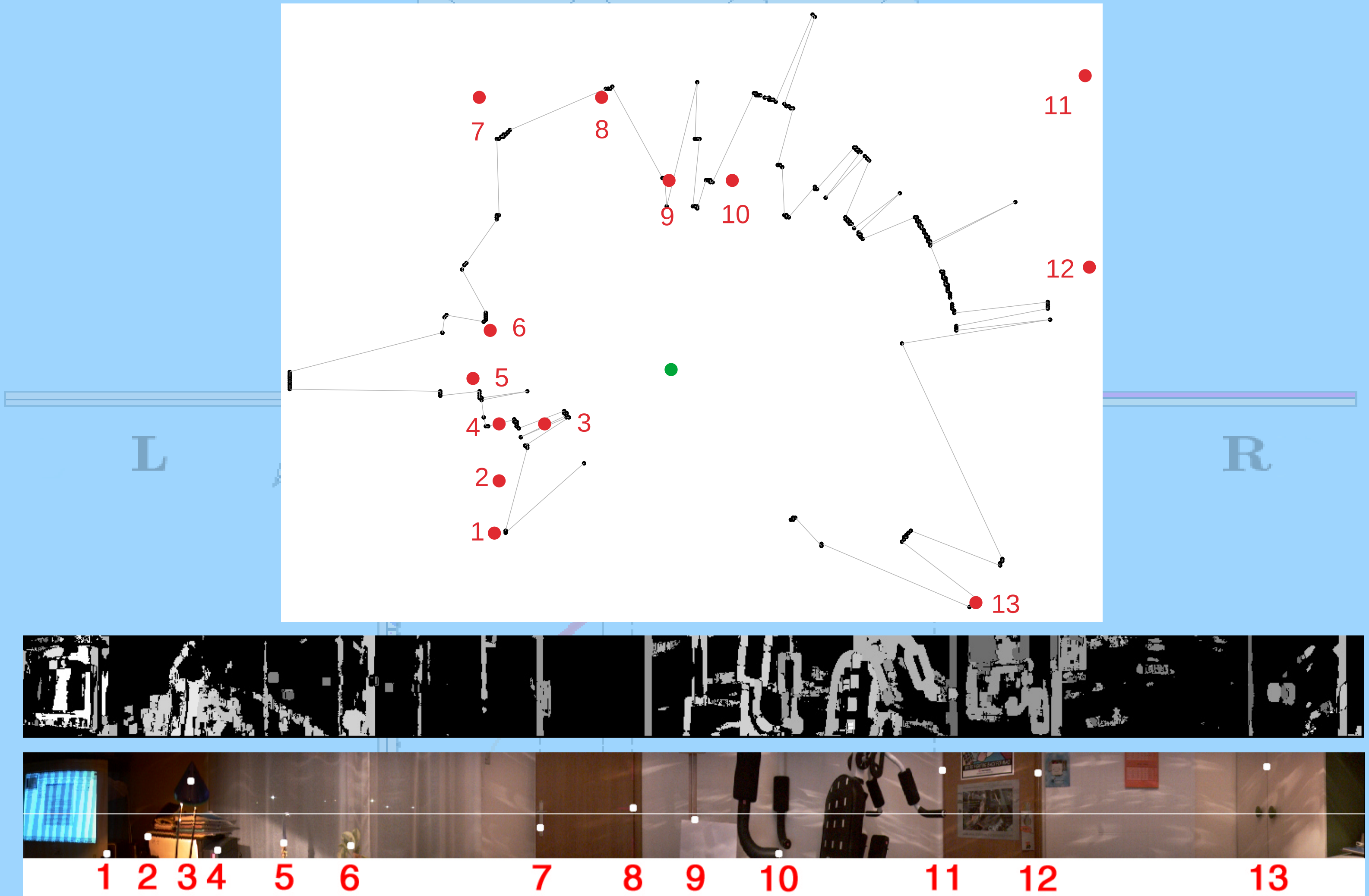
$2\varphi = 29.9625^{o}$





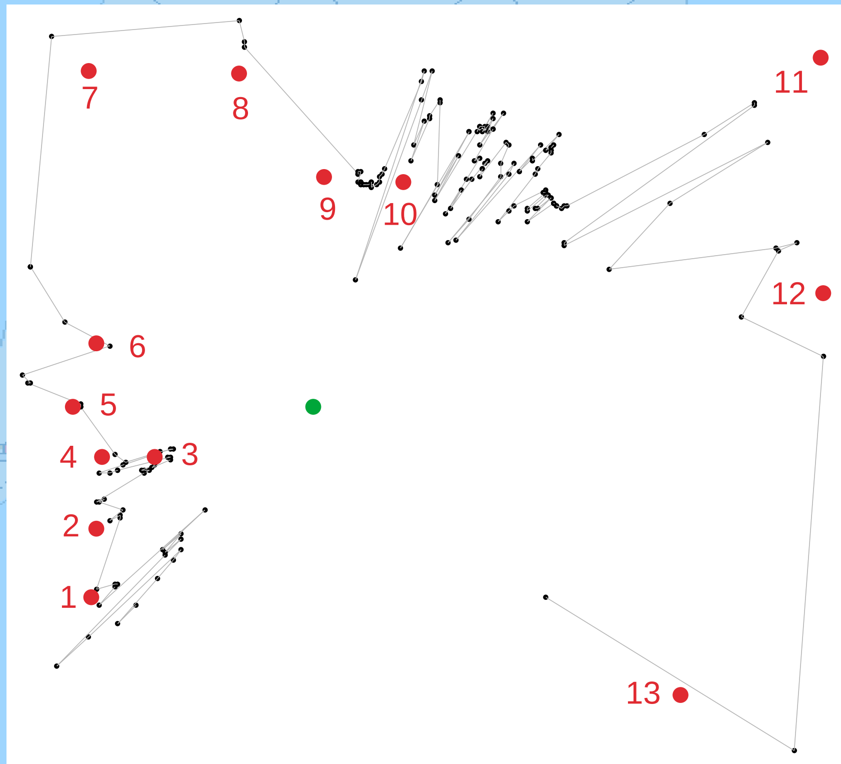
Experimental results #2 — Evaluation

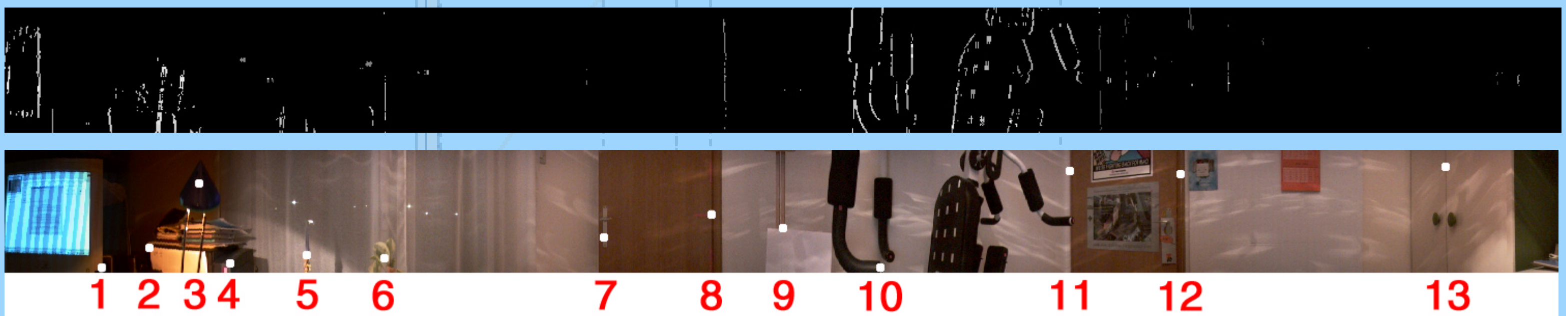
$2\varphi = 3.6125^{\mathrm{o}}$



Experimental results #2 — Evaluation

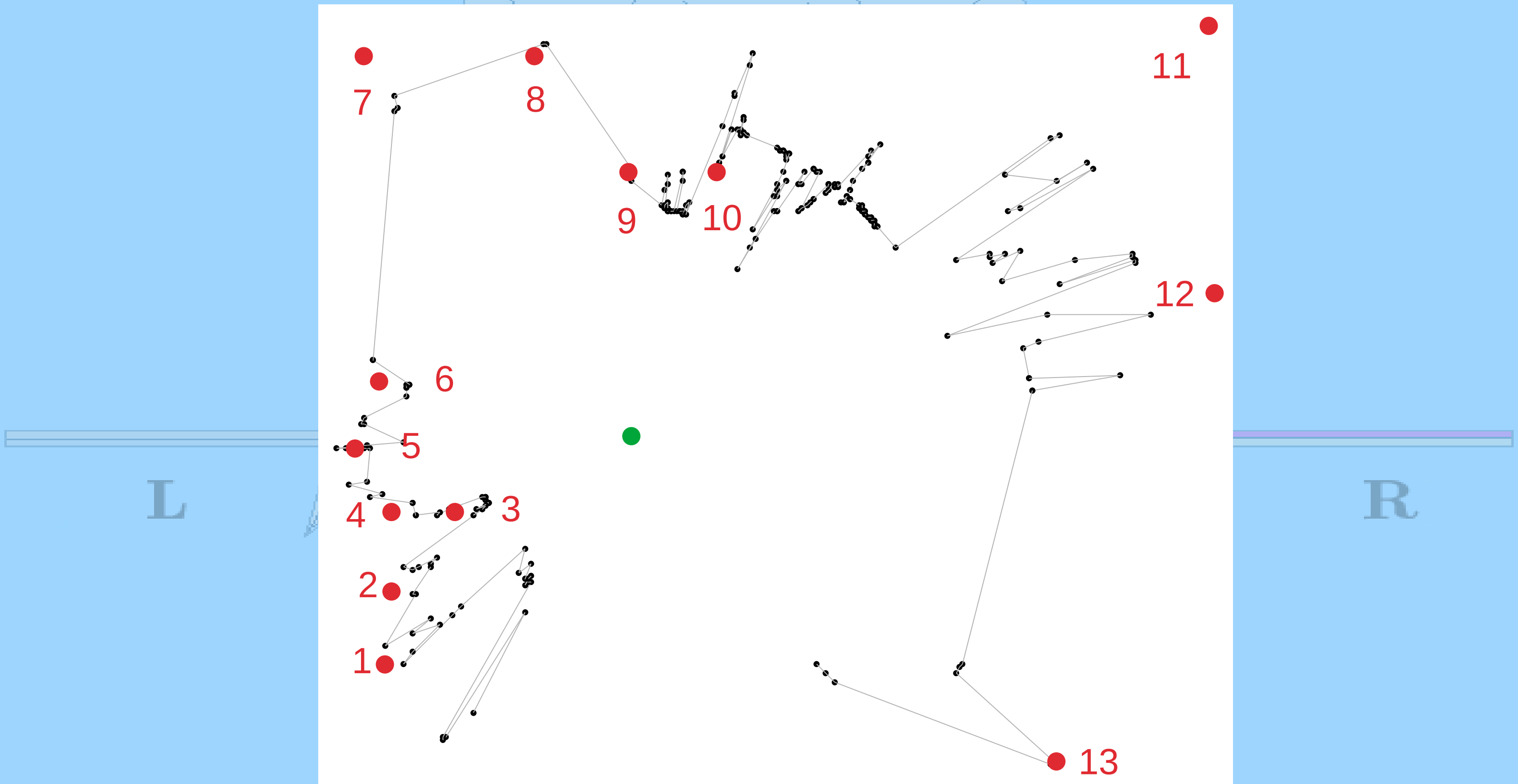
$2\varphi = 29.9625^{\text{o}}$

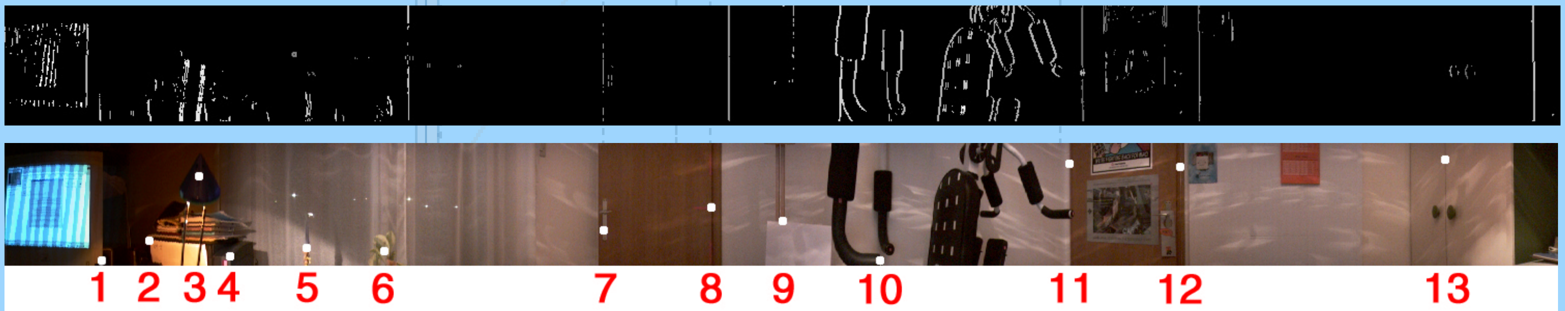




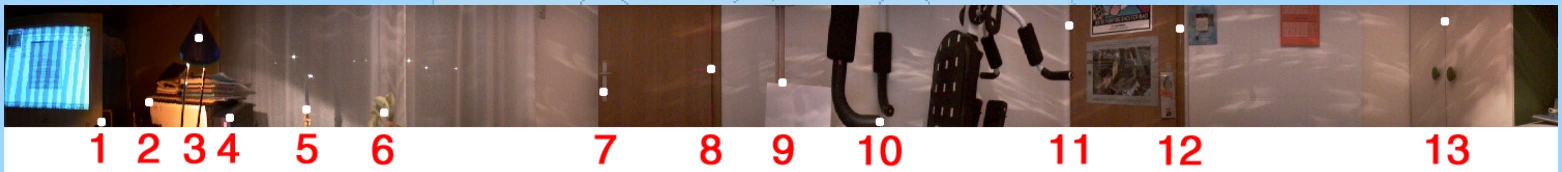
Experimental results #2 — Evaluation

$2\varphi = 3.6125^{\circ}$





Experimental results #3 — The error function on the manually measured points



feature marked in the fig. with	actual distance d [cm]	estimated depth l [cm] for $2\varphi =$		difference $l - d$ [cm (% of d)] for $2\varphi =$	
		3.6125^o	29.9625^o	3.6125^o	29.9625^o
1	111.5	89.4	109	-22.1 (-19.8%)	-2.5 (-2.2%)
2	95.5	76.7	89.3	-18.8 (-19.6%)	-6.2 (-6.5%)
3	64	53.8	59.6	-10.2 (-15.9%)	-4.4 (-6.9%)
4	83.5	76.7	78.3	-6.8 (-8.1%)	-5.2 (-6.2%)
5	92	89.4	89.3	-2.6 (-2.8%)	-2.7 (-2.9%)
6	86.5	76.7	82.7	-9.8 (-11.3%)	-3.8 (-4.4%)
7	153	133.4	159.8	-19.6 (-12.8%)	6.8 (4.4%)
8	130.5	133.4	135.5	2.9 (2.2%)	5 (3.8%)
9	88	76.7	87.6	-11.3(-12.8%)	-0.4 (-0.5%)
10	92	89.4	89.3	-2.6 (-2.8%)	-2.7 (-2.9%)
11	234.5	176.9	213.5	-57.6 (-24.6%)	-21 (-9%)
12	198	176.9	179.1	-21.1 (-10.7%)	-18.9 (-9.5%)
13	177	176.9	186.7	-0.1 (-0.1%)	9.7 (5.5%)

Summary

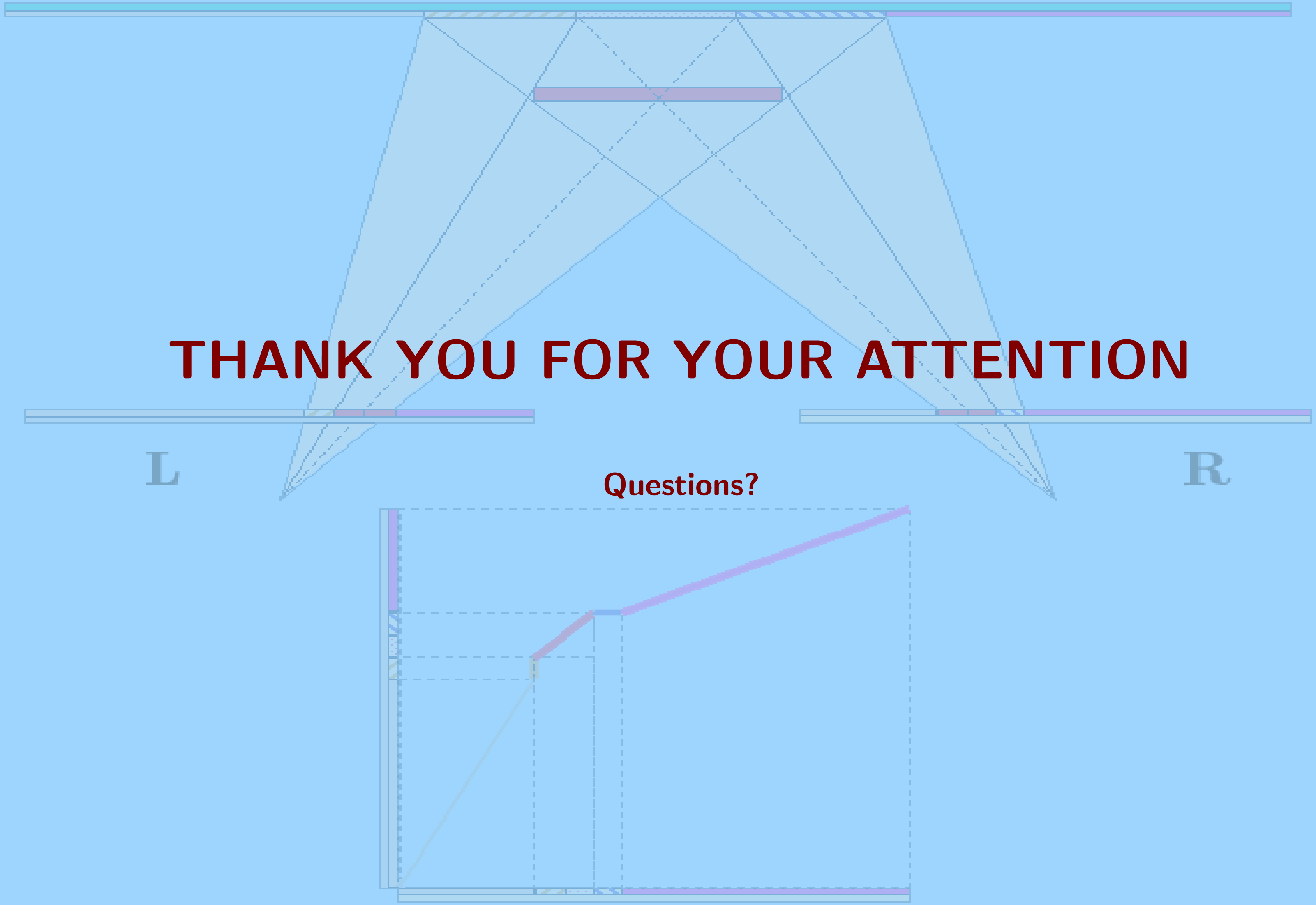
- We are doing a stereo reconstruction with only one standard camera.
- Problem: real time execution of the mosaicing process.
- Epipolar lines are image rows.
- The search space on the epipolar line can be effectively constraint.
- The confidence in the estimated depth is changing with the slope of the curve l (depth).
- The reconstruction time of dense depth images is not acceptable (in most cases).
- The results are very promising but they can get better.
- **Reconstruction of small spaces** — constraint: α

Future work

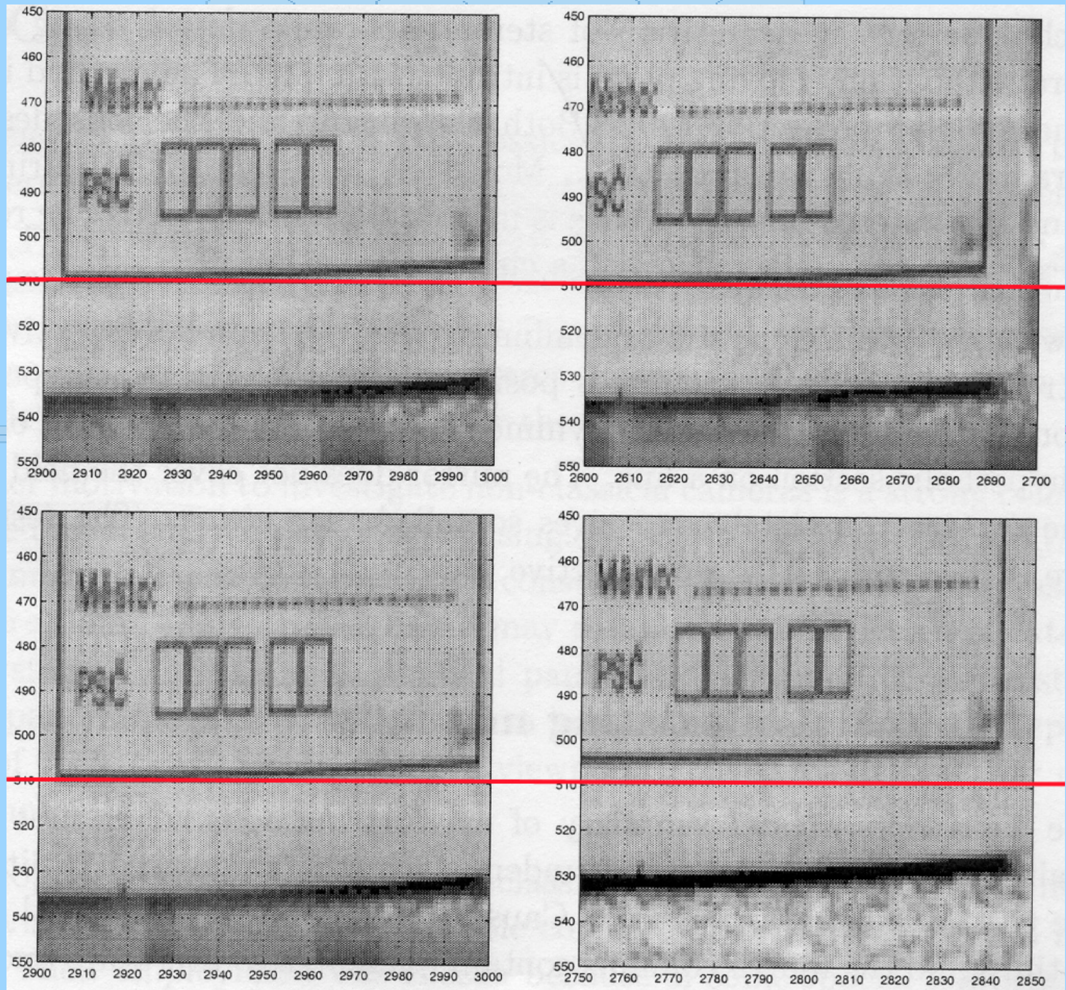
- Solve mentioned problems &
  - identify the size of the systematical error,
  - address the precision of vertical reconstruction,
  - use calibration procedure,
  - use sub-pixel accuracy procedure and
  - enlarge the vertical field of view.
- **The goal:** mobile robot application!

THANK YOU FOR YOUR ATTENTION

Questions?



Appendix #1 — Vertical motion parallax effect



[Huang]

Appendix #2 — Using wider stripes