



Konzistentní rekonstrukce scény
z velkého množství snímků

Daniel Martinec

Vedoucí diplomové práce: Ing. Tomáš Pajdla

1

2

3

4

5

6

7

8

9

10

11

Co je rekonstrukce



m p

- Máme obrázky nějakého objektu (z fotek, filmu).



Neznáme polohy kamer.

- Chceme najít 3D model scény, **polohy kamer**.



1

2

3

4

5

6

7

8

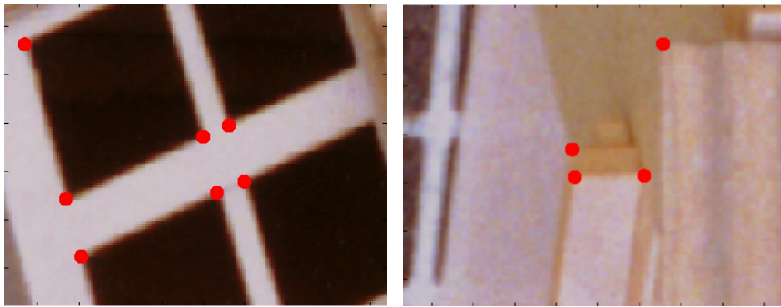
9

10

11

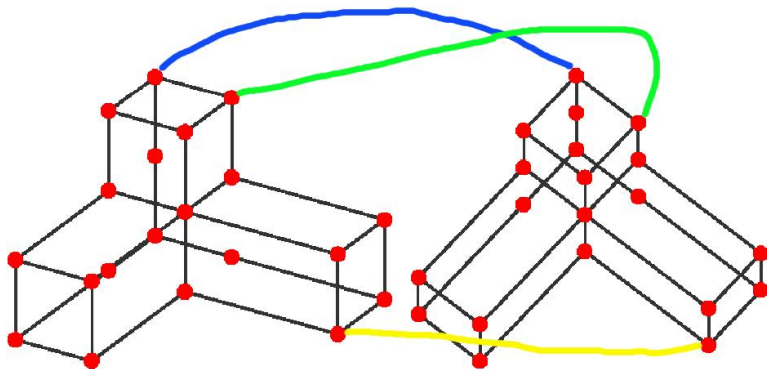


1. V obrázcích se určí význačné body.



Např. rohy.

2. Přiřadí se korespondence těchto bodů.



3. Nalezne se 3D model scény.



1

2

3

4

5

6

7

8

9

10

11

Problém – chybí data



m p

1

2

3

4

... protože v obrázcích nejsou vidět všechny body scény.



5

6

7

8

9

10

11

Formulace úlohy



m p

$$\underbrace{\lambda_j^{(i)}}_{\text{projektivní hloubka}} \quad \underbrace{\mathbf{q}_j^{(i)}}_{\text{bod v obrázku}} = \underbrace{P^{(i)}}_{\text{kamera}} \quad \underbrace{\mathbf{x}_j}_{\text{bod v prostoru}}$$

\cap \cap \cap \cap

\mathcal{R} $\mathcal{R}^{3 \times 1}$ $\mathcal{R}^{3 \times 4}$ $\mathcal{R}^{4 \times 1}$

$$\underbrace{\begin{bmatrix} \lambda_1^{(1)} \mathbf{q}_1^{(1)} & \lambda_2^{(1)} \mathbf{q}_2^{(1)} & \lambda_3^{(1)} \mathbf{q}_3^{(1)} & \dots & \lambda_n^{(1)} \mathbf{q}_n^{(1)} \\ \lambda_1^{(2)} \mathbf{q}_1^{(2)} & \lambda_2^{(2)} \mathbf{q}_2^{(2)} & \lambda_3^{(2)} \mathbf{q}_3^{(2)} & \dots & \lambda_n^{(2)} \mathbf{q}_n^{(2)} \\ \lambda_1^{(3)} \mathbf{q}_1^{(3)} & \lambda_2^{(3)} \mathbf{q}_2^{(3)} & \lambda_3^{(3)} \mathbf{q}_3^{(3)} & \dots & \lambda_n^{(3)} \mathbf{q}_n^{(3)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \lambda_1^{(m)} \mathbf{q}_1^{(m)} & \lambda_2^{(m)} \mathbf{q}_2^{(m)} & \lambda_3^{(m)} \mathbf{q}_3^{(m)} & \dots & \lambda_n^{(m)} \mathbf{q}_n^{(m)} \end{bmatrix}}_Q = \underbrace{\begin{bmatrix} P^{(1)} \\ P^{(2)} \\ P^{(3)} \\ \vdots \\ P^{(m)} \end{bmatrix}}_P \underbrace{[\mathbf{x}_1 \mathbf{x}_2 \mathbf{x}_3 \dots \mathbf{x}_n]}_X$$

- Cíl: rekonstrukce scény s perspektivní kamerou a s **chybějícími daty**.

1

2

3

4

5

6

7

8

9

10

11



- $\lambda_j^{(i)}$ ← epipolární geometrie, (trifokální, kvadrifokální tenzory)
- Nechybí data
 - Q má hodnot (zhruba) 4
 - * $Q = P_{3m \times 4} X_{4 \times n}$
 - Faktorizace pomocí svd (singular value decomposition)
- Chybí data
 - Algoritmus existuje, ale pouze pro ortografickou kameru, tj. $\lambda_j^{(i)} = 1 \forall_{i,j}$.
 1. V Q se vybírají čtveřice sloupců, které na bázi Q kladou omezující podmínku.
 - * Definována i pro neúplné sloupce
 - * „Báze Q leží v lineárním obalu čtveřice sloupců“
 - * Lin. prostor generovaný čtveřicí → komplement → komplement báze Q
 2. Báze Q → doplnění
 3. Faktorizace

1

2

3

4

5

6

7

8

9

10

11



1

2

3

- 1.(a) Spočítá se co nejvíce projektivních hloubek $\lambda_j^{(i)}$ \longrightarrow kompletní podmatice Q
- (b) Tato podmatice se doplní algoritmem pro ortografickou kameru.

4

Kroky (a) a (b) se opakují, dokud není celá Q kompletní.

5

2. Faktorizace kompletní Q .

6

- Minimalizace počtu iterací \longrightarrow přesnější výsledky

7

1. Výběr podmnožiny obrázků \longrightarrow doplnění nejvíce dat.
2. Zobecnění algoritmu doplňování pro perspektivní kameru.

8

9

10

11



Doplňování pro perspektivní kameru

- Podmínka pro použitelnost čtveřice sloupců
 - Využití bodů s neznámou projektivní hloubkou
- Důkazy
 1. Správnost podmínky
 2. Nové možnosti – např.
 - K rekonstrukci scény se 2 obrázky stačí znát 5 bodů a 3 společné projektivní hloubky.
 - Je možné rekonstruovat více než předchozími metodami.
 3. Omezení
 - Méně než 2 projektivní hloubky v obrázku nestačí.

1

2

3

4

5

6

7

8

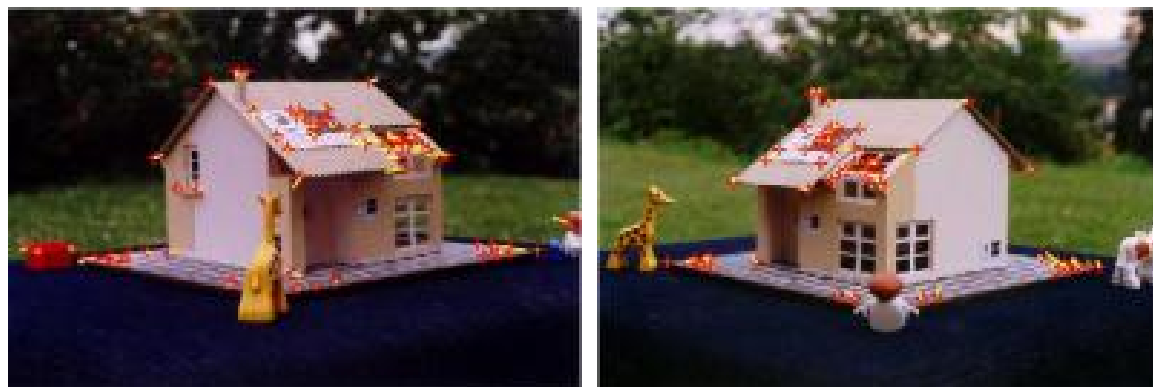
9

10

11



Scéna s domkem — 10 obrázků v rozlišení 2003 × 2952 pixelů



Chyby reprojekce v pixelech:

| Obrázek | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---------|------|------|------|------|------|------|------|------|------|------|
| Součet | 230 | 321 | 292 | 274 | 225 | 174 | 231 | 278 | 202 | 264 |
| Maximum | 11.9 | 10.8 | 15.3 | 14.3 | 9.8 | 14.6 | 15.3 | 15.2 | 8.5 | 15.6 |
| Medián | 1.13 | 0.93 | 0.76 | 1.05 | 0.00 | 0.00 | 1.00 | 1.28 | 0.83 | 0.95 |
| Průměr | 1.37 | 1.91 | 1.74 | 1.63 | 1.34 | 1.04 | 1.38 | 1.66 | 1.21 | 1.58 |

Celkový součet chyby reprojekce: 2495 pixelů

Průměrná chyba reprojekce: 2.52 pixelů na bod

1

2

3

4

5

6

7

8

9

10

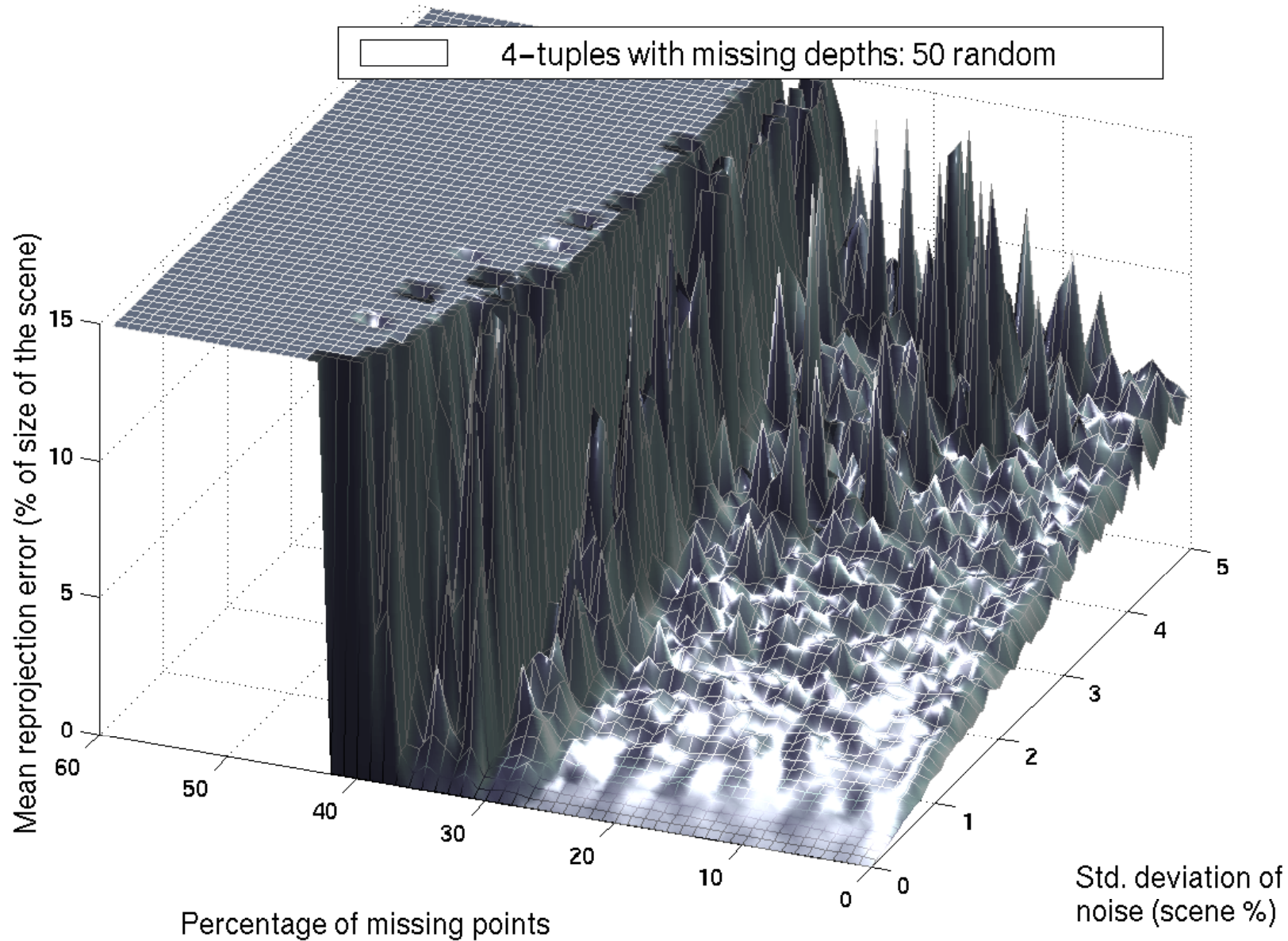
11

Experimenty s umělými scénami



m p

Error of reconstruction by Iri-Jac for missing depths algorithm
(5 images, normal noise, 5 trials)



1

2

3

4

5

6

7

8

9

10

11



- Nový algoritmus, jeho implementace a otestování.
 - Využívá co nejvíce informace o scéně a co nejspravedlivěji.
- Způsob využití bodů s neznámou projektivní hloubkou.
 - Dokázána správnost, možnosti a omezení této metody.
- Integrace algoritmu do systému RecX na katedře kybernetiky FELK CVUT.
 - Srovnatelné výsledky s algoritmy v ReXu.
- Závěr
 - Otázka: Je nový algoritmus kandidátem na obecný algoritmus rekonstrukce scény?
 - Nápady na vylepšení, hledání algoritmu na výpočet velmi málo projektivních hloubek (2-3).

1

2

3

4

5

6

7

8

9

10

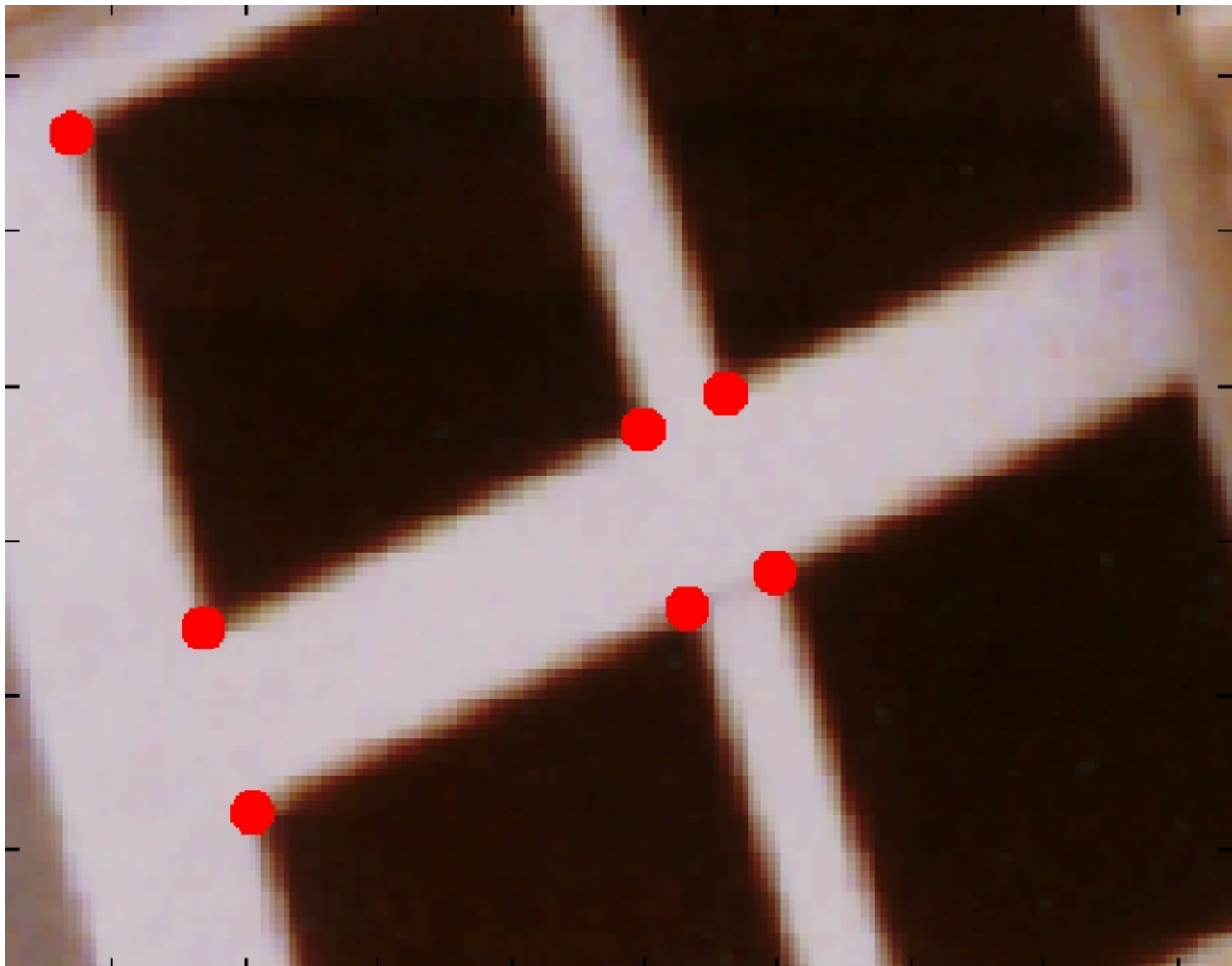
11

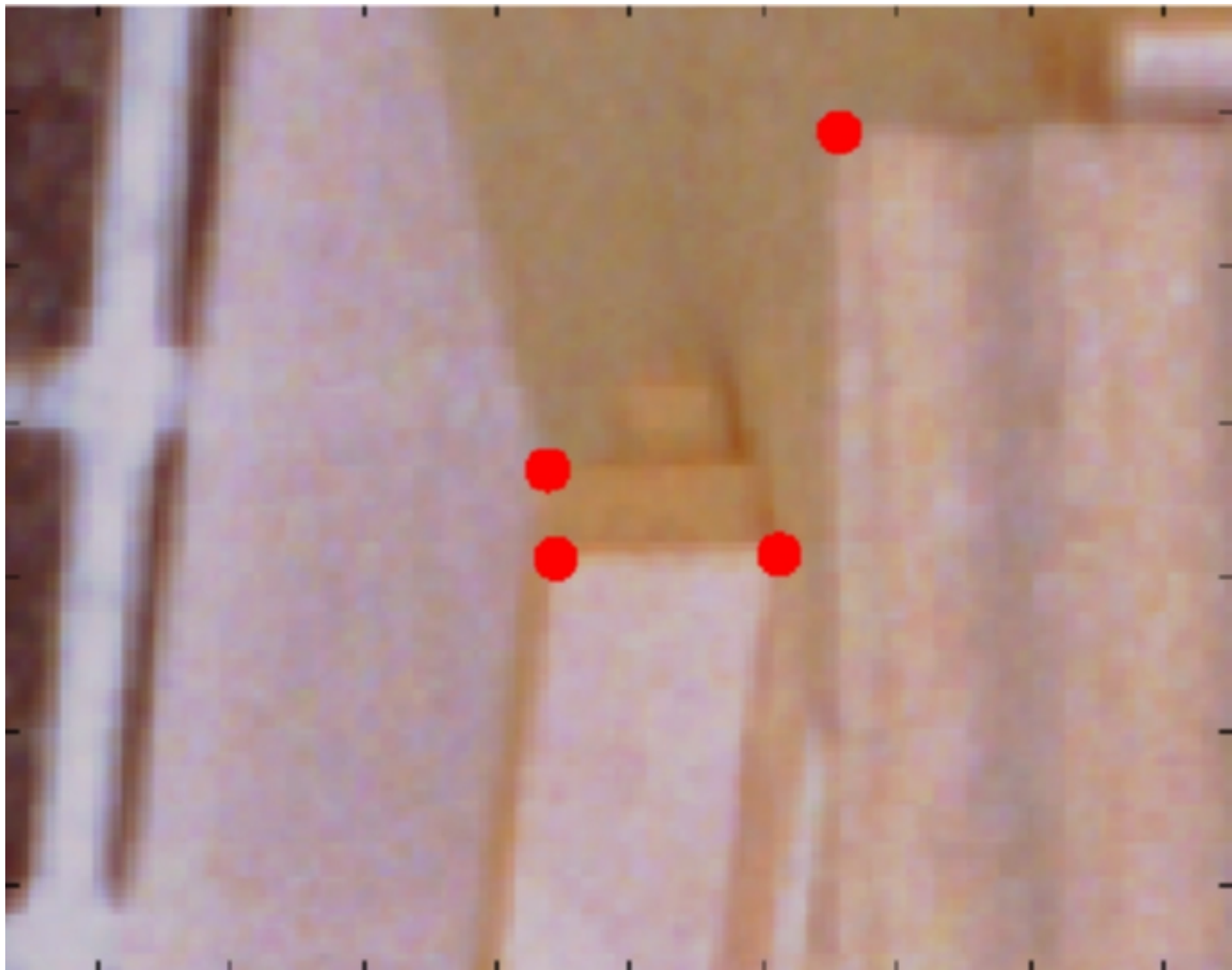


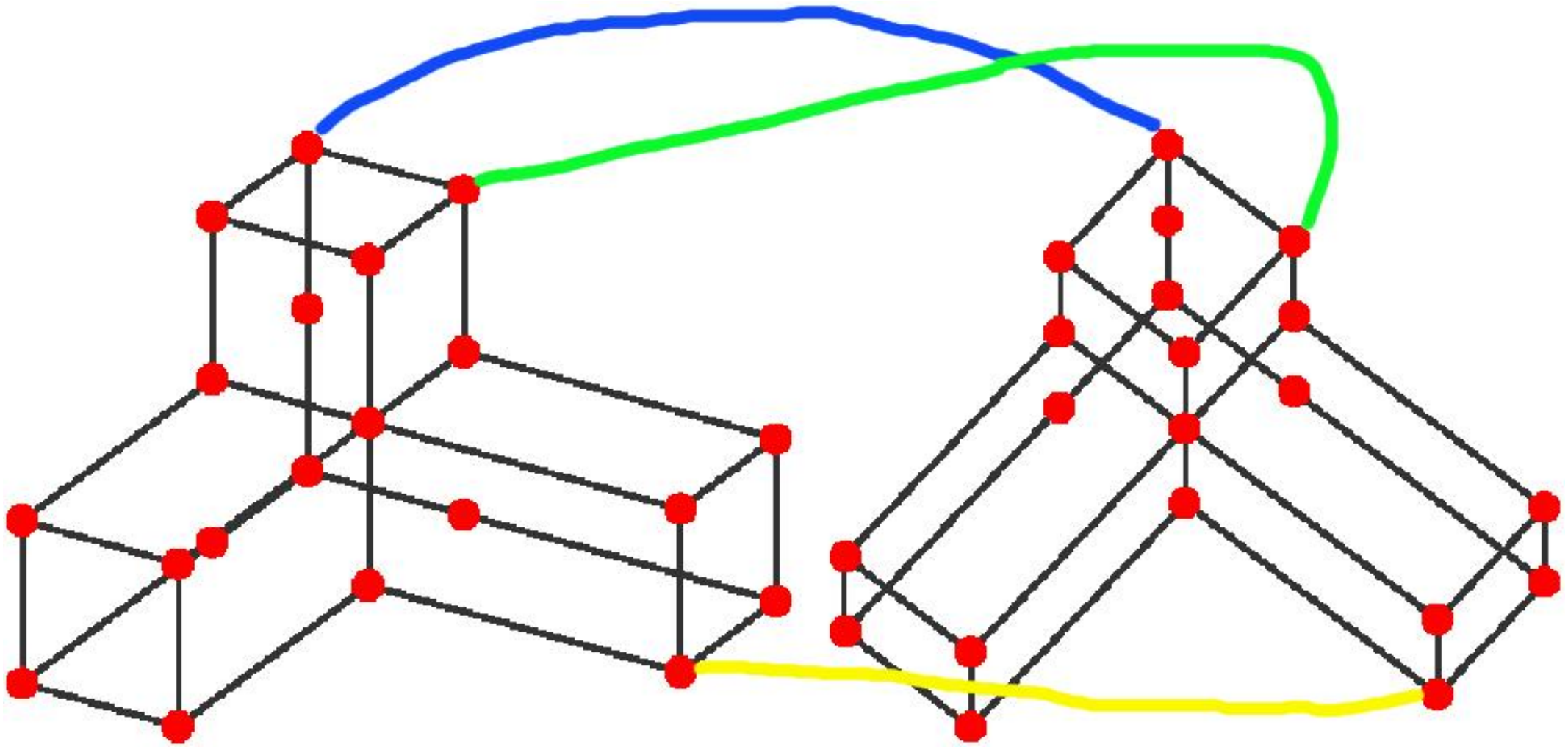






















Error of reconstruction by Iri-Jac for missing depths algorithm
(5 images, normal noise, 5 trials)

