The Complete Rank Transform: A Tool for Accurate and Morphologically Invariant Matching of Structures

Oliver Demetz, David Hafner, and Joachim Weickert

Mathematical Image Analysis Group
Saarland University
Saarbrücken, Germany

Partly funded by the German Research Foundation (DFG)

Illumination changes in reality

source: KITTI benchmark

Illumination changes in reality

source: KITTI benchmark

Illumination changes in reality

source: KITTI benchmark

- Important for optic flow
- Assumption that intensity of objects stays constant violated
- In this work: only assume invariance under monotonically increasing greyvalue rescalings (eg. additive, multiplicative)

$\rightarrow$ develop morphologically invariant descriptor

Outline

Rank Transform (RT) (Zabih and Woodfill 1994)

- Idea:

  How many pixels are smaller than me?

- Invariant under monotonically increasing transformations

4	14	83
88	25	4
3	15	65

Intensity values

	5	

Rank

Rank Transform (RT) (Zabih and Woodfill 1994)

- Idea:

  How many pixels are smaller than me?

- Invariant under monotonically increasing transformations

4	14	83
88	25	4
3	15	65

Intensity values

	5	

Rank

Rank Transform (RT) (Zabih and Woodfill 1994)

- Idea:

  How many pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Resulting signature:

$$s_{\mathrm{RT}} = 5$$

4	14	83
88	25	4
3	15	65

Intensity values

	5	

Rank

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CT}} = (1, \qquad\qquad\qquad)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CT}} = (1, 1, \qquad\qquad)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CT}} = (1, 1, 0, \qquad\qquad)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CT}} = (1, 1, 0, 0, 1, \qquad)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CT}} = (1, 1, 0, 0, 1, 1, 1, 0)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Census Transform (CT) (Zabih and Woodfill 1994)

- Idea:

  Which pixels are smaller than me?

- Invariant under monotonically increasing transformations
- Census signatures carry spatial information
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CT}} = (1, 1, 0, 0, 1, 1, 1, 0)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	1	0
0		1
1	1	0

Census

Complete Rank Transform (CRT)

- Capture the rank of every pixel in the local patch
- Vector-valued signature in every pixel
- Invariant under monotonically increasing transformations
- Carries as much local image information as possible

4	14	83
88	25	4
3	15	65

Intensity values

1	3	7
8	5	1
0	4	6

Census

Complete Rank Transform (CRT)

- Capture the rank of every pixel in the local patch
- Vector-valued signature in every pixel
- Invariant under monotonically increasing transformations
- Carries as much local image information as possible
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CRT}} = (1, 3, 7, \qquad\qquad)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	3	7
8	5	1
0	4	6

Census

Complete Rank Transform (CRT)

- Capture the rank of every pixel in the local patch
- Vector-valued signature in every pixel
- Invariant under monotonically increasing transformations
- Carries as much local image information as possible
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CRT}} = (1, 3, 7, 8, 5, 1, \qquad)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	3	7
8	5	1
0	4	6

Census

Complete Rank Transform (CRT)

- Capture the rank of every pixel in the local patch
- Vector-valued signature in every pixel
- Invariant under monotonically increasing transformations
- Carries as much local image information as possible
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CRT}} = (1, 3, 7, 8, 5, 1, 0, 4, 6)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	3	7
8	5	1
0	4	6

Census

Complete Rank Transform (CRT)

- Capture the rank of every pixel in the local patch
- Vector-valued signature in every pixel
- Invariant under monotonically increasing transformations
- Carries as much local image information as possible
- Resulting signature:

$$\boldsymbol{s}_{\mathrm{CRT}} = (1, 3, 7, 8, 5, 1, 0, 4, 6)^{\top}$$

4	14	83
88	25	4
3	15	65

Intensity values

1	3	7
8	5	1
0	4	6

Census

Outline

- Introduction
- Complete Rank Transform
- **Variational Optic Flow Model**
- Experiments
- Conclusions

Variational Optic Flow Model

- Modify model of Brox et al. (2004)
- Minimise functional for optic flow field $(u, v)^\top : \Omega \to \mathbb{R}^2$

$$E(u, v) = \int_\Omega (\, D + \alpha \, R \,) \, \mathrm{d}x \, \mathrm{d}y$$

- Data term D models constancy of Complete Rank signatures of corresponding positions

$$\left\| \boldsymbol{s}_{\mathrm{CRT}}(x + u, y + v, t + 1) - \boldsymbol{s}_{\mathrm{CRT}}(x, y, t) \right\|^2$$

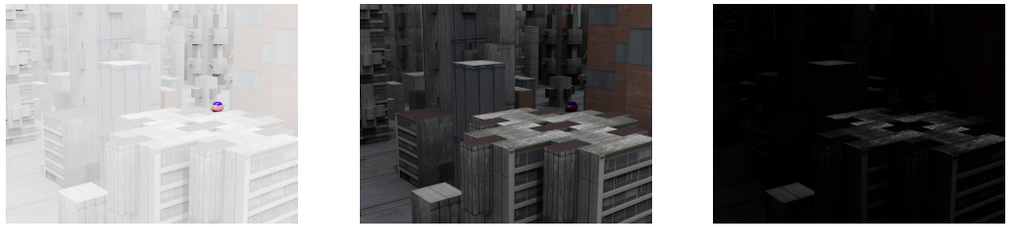
- Regularisation term R leads to piecewise smooth solution
- Both terms equipped with robust estimator functions
- Typical warping-based coarse-to-fine minimisation (Brox et al. 2004)

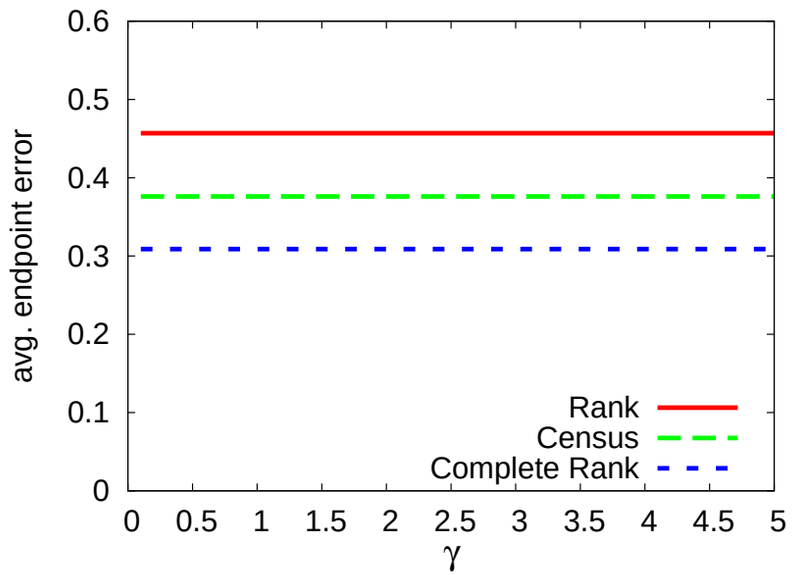
Outline

Experiments - γ Changes

- Adjust second frame with an exponential function $f_\gamma = f_{\text{max}} \cdot \left(\frac{f}{f_{\text{max}}}\right)^\gamma$

$\gamma = 0.1$ $\gamma = 1$ $\gamma = 3$





→ Unconditional morphological invariance
→ Outperform Rank and Census

Experiments - KITTI Vision Benchmark

- GCPR Special Session Imagery (all scenes with lighting changes)

KITTI image sequence:	#11	#15	#44	#74	average
Zimmer et al. 2011	37.3	32.3	23.2	62.9	38.9
Bruhn/Weickert 2005	33.9	47.7	32.4	71.4	46.7
Census Transform	36.5	28.6	28.5	63.8	39.4
Complete Rank Transform	**29.8**	**22.8**	**22.6**	**61.5**	**34.2**

[%] bad pixel error measure bp3

→ CRT outperforms established methods in difficult lighting conditions

- Whole KITTI benchmark (195 testing image sequences)
  - Our method ranks 10th of 33
  - Top 4 methods use stereo information
  - Census-based method by Ranftl et al. (2012) ranks 14th

→ CRT carries enough image information for real-world applications

Experiments - Middlebury Benchmark

- Middlebury Training Set

	rw	dimetr.	grove2	grove3	hydr.	urban2	urban3	yos	avg
RT	.111	.092	.191	.764	.191	.457	1.03	.211	.381
CT	.102	.090	.169	.646	**.147**	.378	.819	.169	.316
CRT	**.100**	**.076**	**.154**	**.585**	.158	**.324**	**.529**	**.150**	**.260**

[pixel] average endpoint error

→ CRT clearly preferable over RT and CT

- Middlebury Benchmark (July 31st 2013)

Method	Average rank
Zimmer et al. (2009)	40.5
ours	45.8
Brox et al. (2004)	52.1

→ even without illumination changes, CRT gives acceptable quality

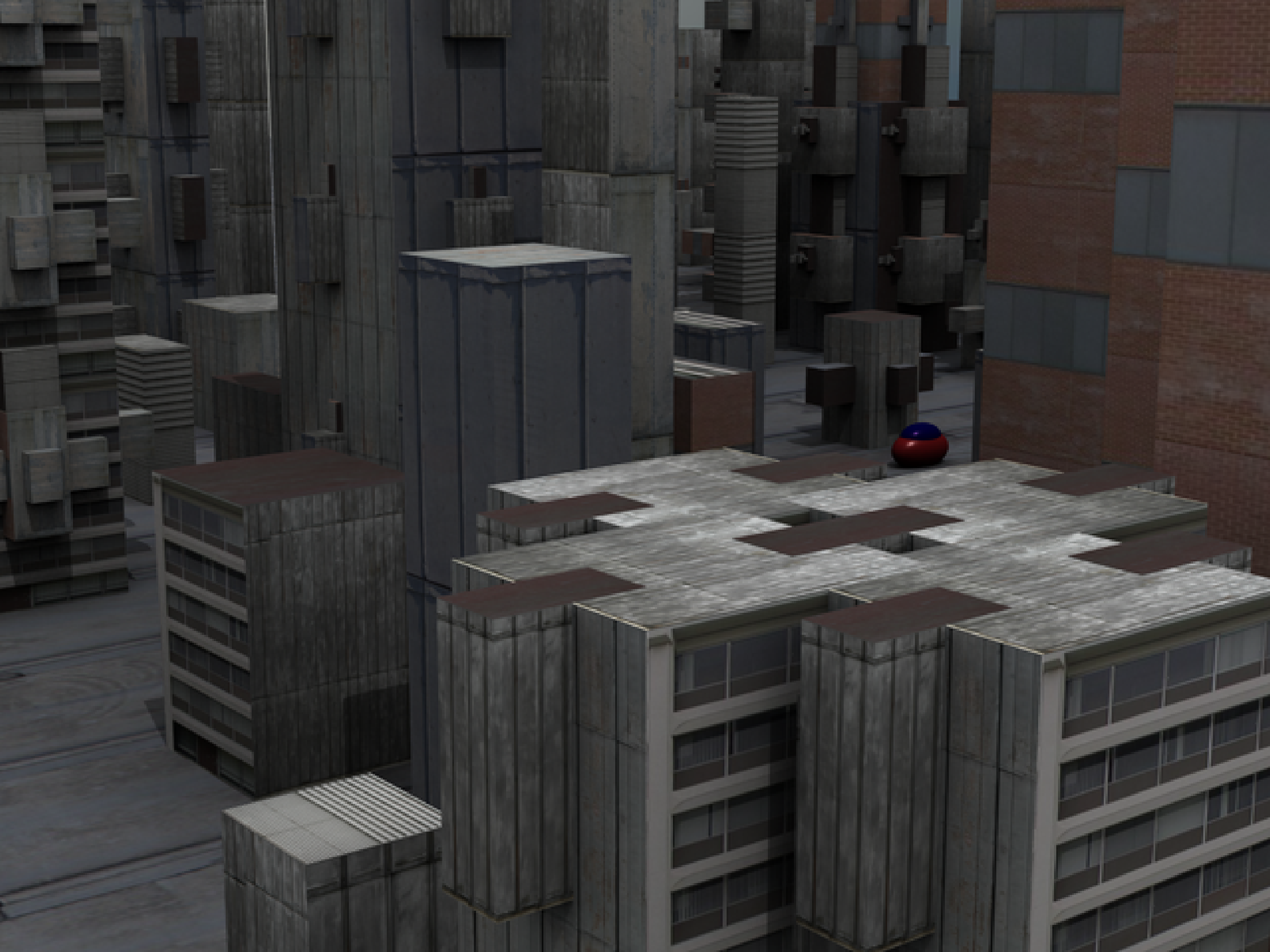
Conclusions

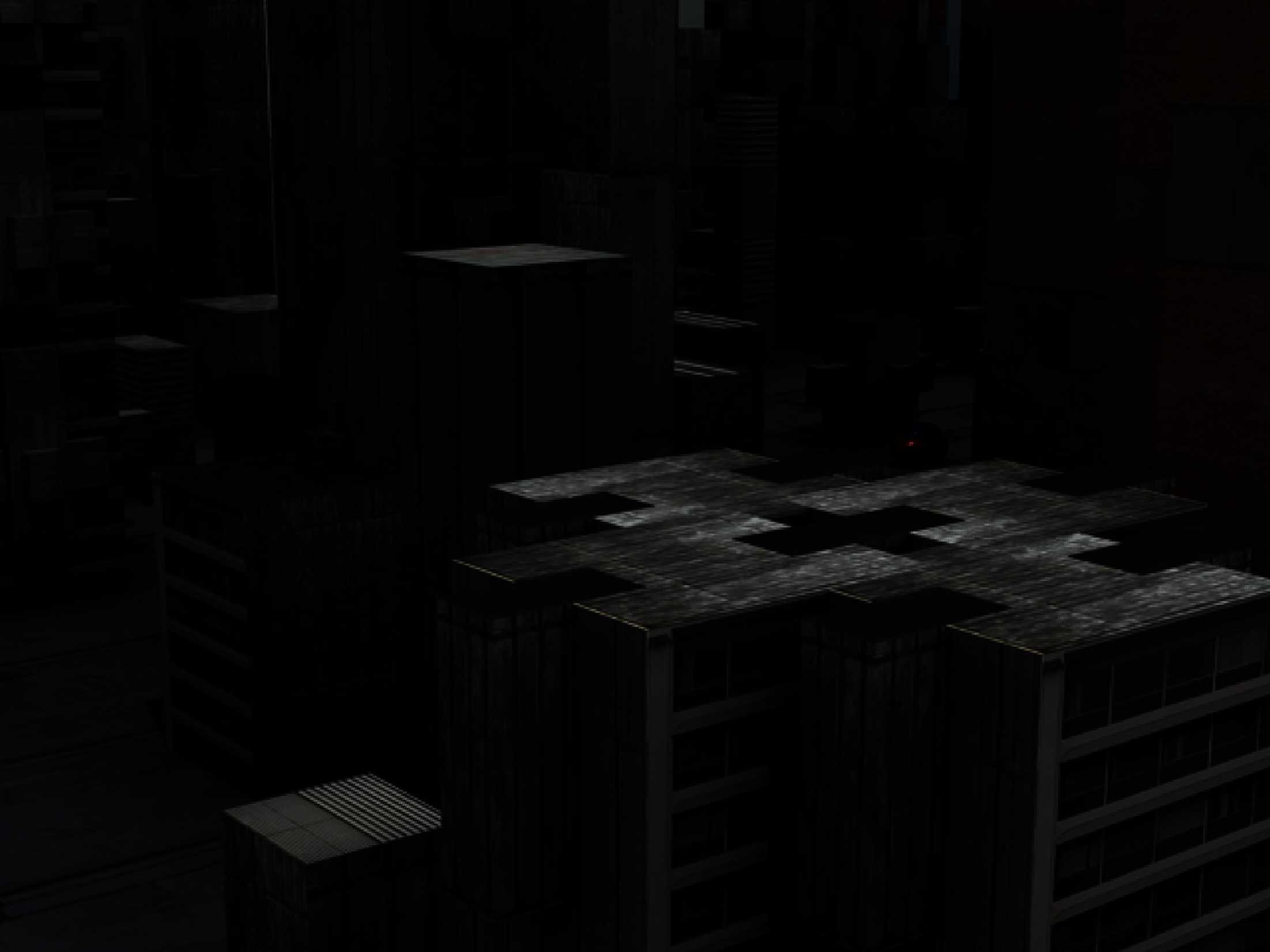
- Morphological invariance handles illumination changes
- Problem: Invariance discards information
- Our solution: CRT that carries as much local image information as possible
- Intentionally kept variational optic flow model simple
- Our proposed CRT clearly preferable over Census and Rank transforms

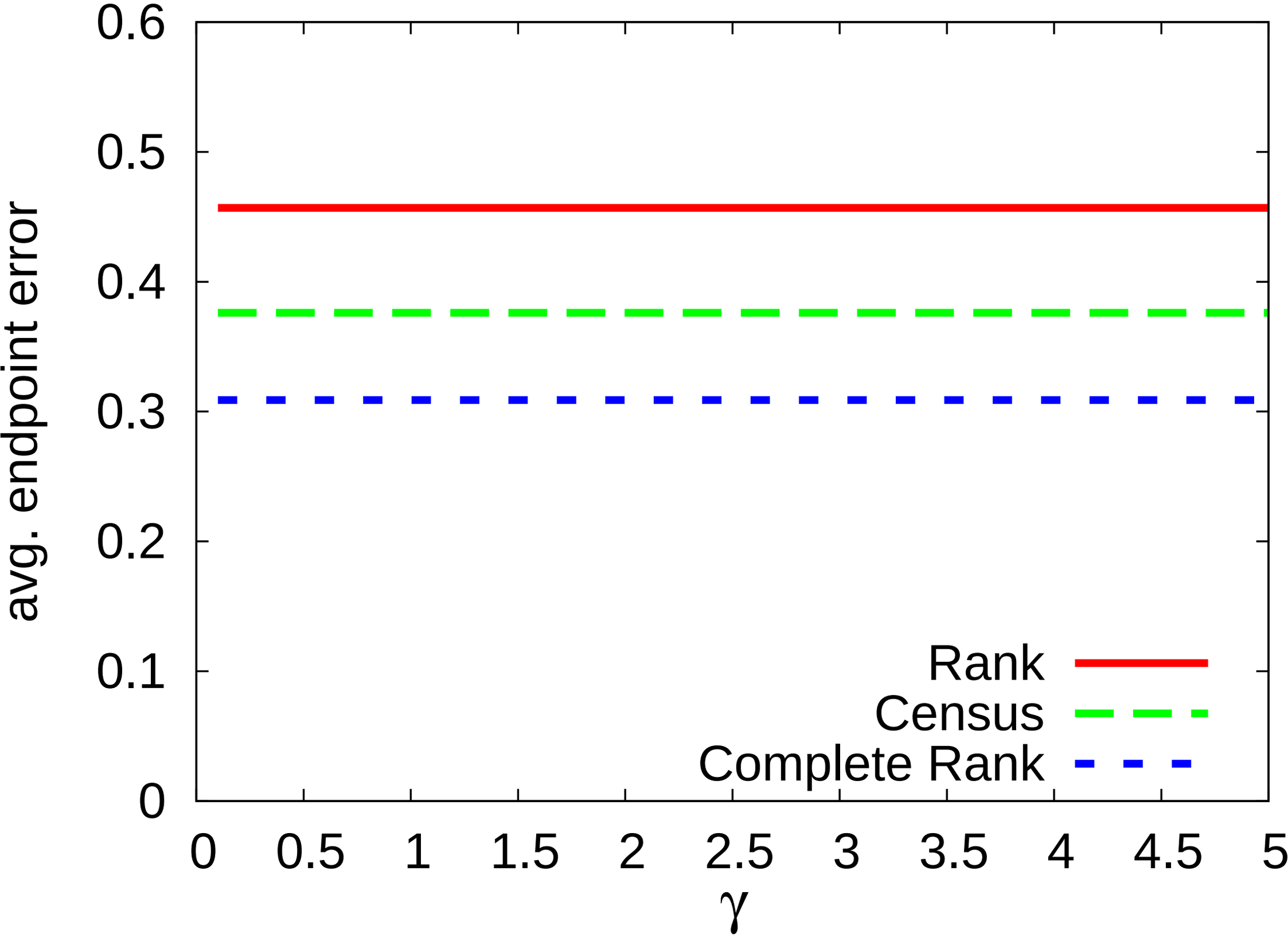
Conclusions

- Morphological invariance handles illumination changes
- Problem: Invariance discards information
- Our solution: CRT that carries as much local image information as possible
- Intentionally kept variational optic flow model simple
- Our proposed CRT clearly preferable over Census and Rank transforms

Thank You !






0.6
0.5
0.4
0.3
0.2
0.1
0
avg. endpoint error
0
0.5
1
1.5
2
2.5
3
3.5
4
4.5
5
γ
Rank
Census
Complete Rank