3D Ordinal Constraint in Spatial Configuration for Robust Scene Recognition

Ching Lik TEO* , Shimiao LI^, Loong-Fah CHEONG , Ju SUN
*DSO National Laboratories; National University of Singapore

Introduction

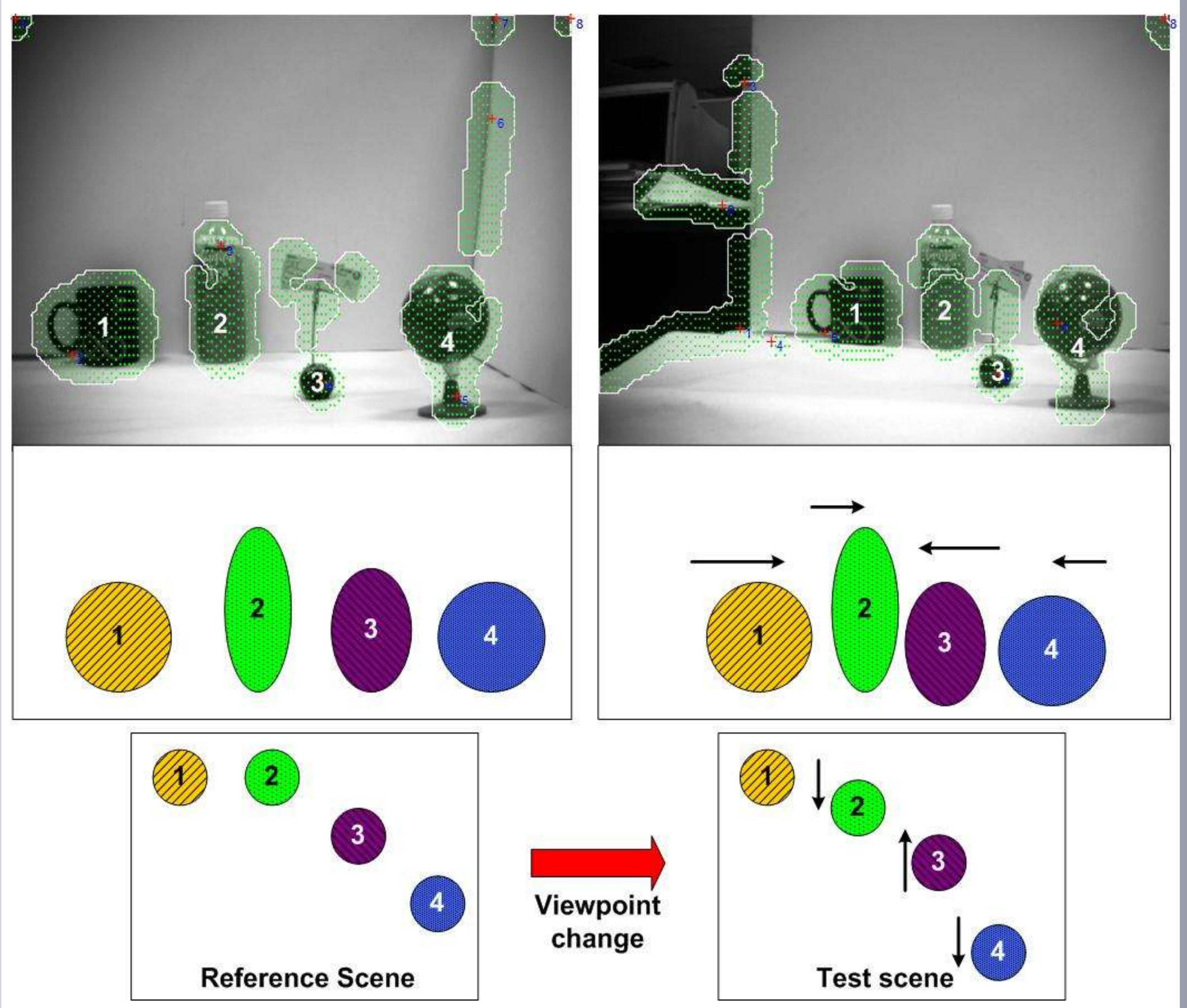
❑ Scene recognition has important uses in robotic navigation and image retrieval.

❑ A scene recognition strategy that integrates the appearance based local SURF features and the geometry based *3D ordinal constraint* is proposed.

❑ The performance is evaluated over four indoor and outdoor databases.

3D Ordinal Constraint in Spatial Configuration

❑ Landmark Ranks

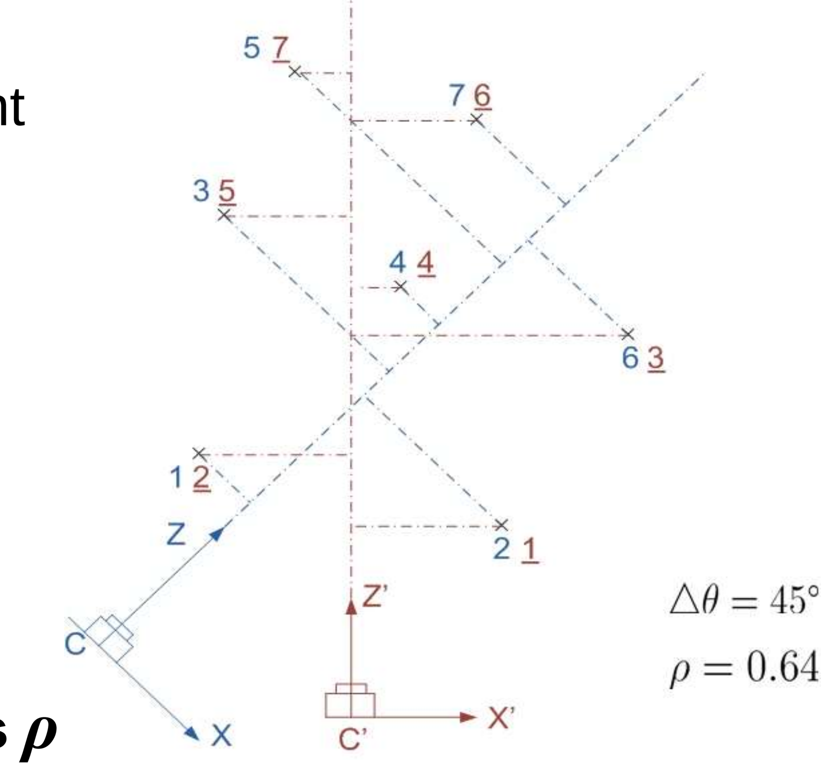


Landmark ranks (in x and z dimensions) under slight viewpoint change.

- Although the absolute position changes, their **ranks** remain invariant.

❑ Rank Correlation Coefficient (RCC)

- As viewpoint changes, some ranks may be perturbed.
- Similarity is captured by the RCC measures – **Spearman's ρ** or **Kendall's τ**

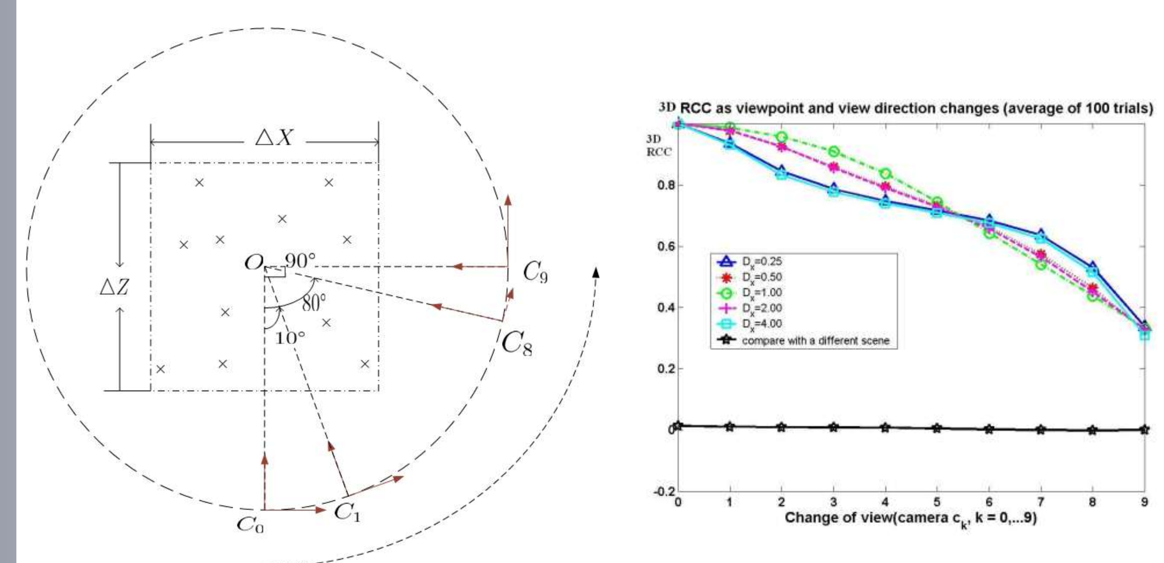


RCC(ρ) in depth dimension.

❑ 3D Rank Correlation Coefficient (3D RCC)

- 3D RCC (ρ_{3D}) is measured in three dimensions – ρ_x, ρ_y, ρ_z.
- Perturbations of rank in different dimensions occur under different types of viewpoint changes and scene configurations.
- RCC in different dimensions compensates one another.
- 3D RCC is defined as:

$$\rho_{3D} = w_Z\rho_Z + w_x\rho_x + w_y\rho_y \qquad (w_Z + w_x + w_y = 1)$$

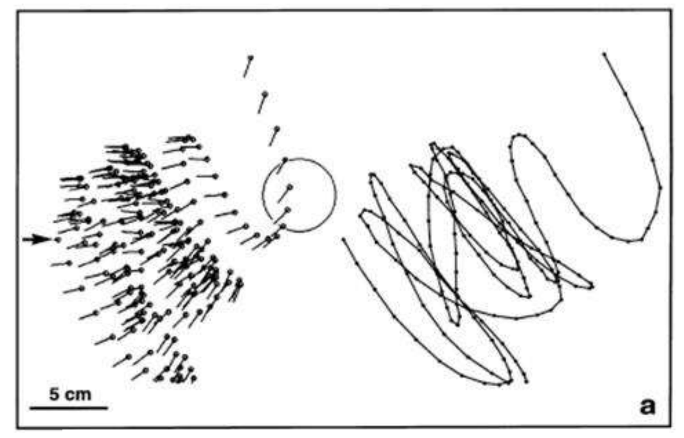


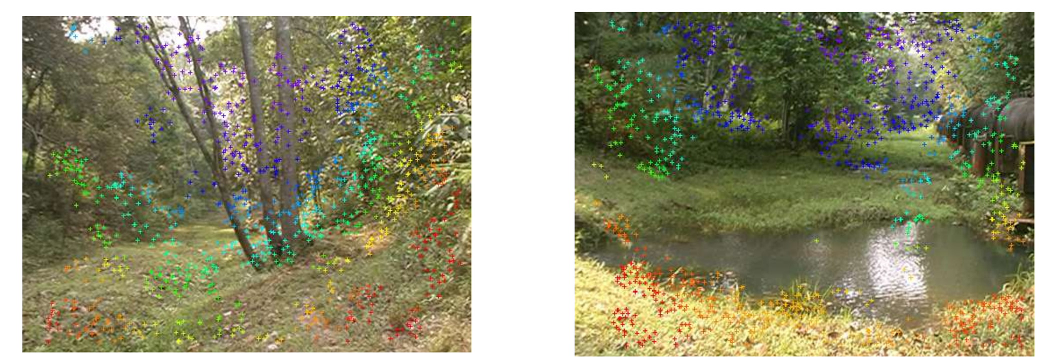
Effect on 3D RCC from simulated viewpoint changes.

Robust Ordinal Depth Acquisition via TBL Motion

❑ Biomimic Turn-Back-and-Look (TBL) motion

- TBL motion provides robust ordinal depth recovery.

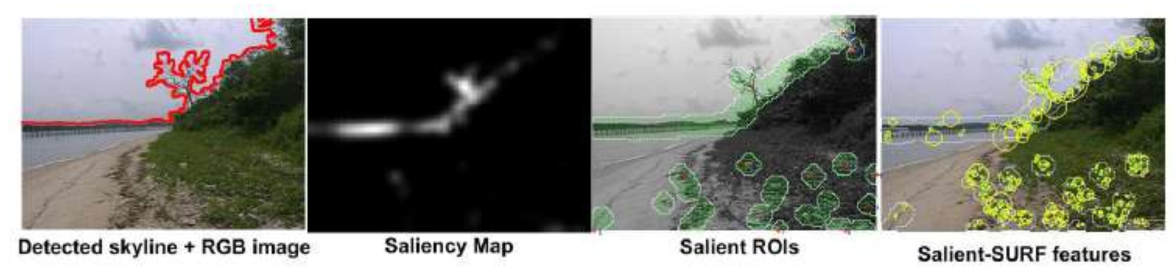




Estimated depth from TBL – red to purple (nearest to furthest)

Scene Recognition Algorithm

❑ The Scene Matrix, S_m



$$s_m(c) = [locs_{N\times 5}\ des_{N\times 65}\ Z_{TBL}]_{N\times 71}$$

- c – c[th] colour space in HSV

❑ Measuring Scene Similarity

- Global Scene Correlation Coefficient, G_c:

$$G_c(s_m^r, s_m^t) = (N_{match}/N_{tot}) \times R$$

N_{match} *Number of matches between the reference and the test.*

N_{tot} *Total number of features in the test*

$R = \frac{\overline{\rho_{3D}+\tau_{3D}}}{2}$ *3D RCC between the reference and the test scenes*

❑ Recognition Decision Module

- A test scene is compared with a database of N_{ref} reference scenes
- The Match Statistics Matrix, m_s summarises the results:

$$m_s = [N_{match}/N_{tot}\ R\ G_c]_{N_{ref}\times 3}$$

The candidate match G_{max}, is the reference scene that yields the largest G_c in m_s. The decision process is as follows:

if $G_{max} < t_d$, *reject the test image;*
elseif $G_{max}/G_{2ndmax} < t_a$, *reject the test image;*
else *accept the test image.*
where t_d and t_a are preset thresholds; G_{2ndmax} is the 2nd largest G_c in m_s.

Experimental Results

Table 1. The four databases. $(N_{ref}, N_{pos}, N_{neg})$ refers to the number of reference scenes, positive test scenes, negative test scenes respectively.

Database	$(N_{ref}, N_{pos}, N_{neg})$	Type
IND	(18, 25, 21)	Indoor
UBIN	(20, 63, 69)	Outdoor coastal
SBWR	(15, 15, 16)	Outdoor enclosed
NS	(20, 41, 52)	Outdoor varied

Table 2. Recognition results of the Proposed SRS(SURF + 3D Ordinal Constraint) (%)

Database	(t_a, t_d)	P_{acc}	P_{rej}	$P_{overall}$
IND	(1.1, 0.03)	84.00	90.48	87.24
UBIN	(1.3, 0.02)	84.13	91.30	87.72
SBWR	(1.2, 0.02)	93.30	100.00	96.65
NS	(1.2, 0.01)	92.68	92.31	92.49

Table 3. Recognition results of the Simple SRS(SURF Only) (%)

Database	(t_a, t_d)	P_{acc}	P_{rej}	$P_{overall}$
IND	(1.2, 0.03)	76.00	71.43	73.71
UBIN	(1.2, 0.01)	69.84	68.12	68.98
SBWR	(1.2, 0.02)	80.00	81.25	80.63
NS	(1.5, 0.02)	70.73	98.08	84.40

Table 4. Recognition results of the Epipolar SRS(SURF + RANSAC based Epipolar Constraint) (%)

Database	(t_a, t_d)	P_{acc}	P_{rej}	$P_{overall}$
IND	(1.2, 0.01)	75.20	71.43	73.32
UBIN	(1.2, 0.01)	57.78	68.12	62.95
SBWR	(1.2, 0.01)	80.00	81.25	80.63
NS	(1.2, 0.01)	80.00	80.77	80.38

Various challenging test and reference scenes in the four databases.

Contact: *tchingli@dso.org.sg; ^shimiao@nus.edu.sg