NUS
National University of Singapore

Facilitating and Exploring Planar Homogeneous Texture for Indoor Scene Understanding

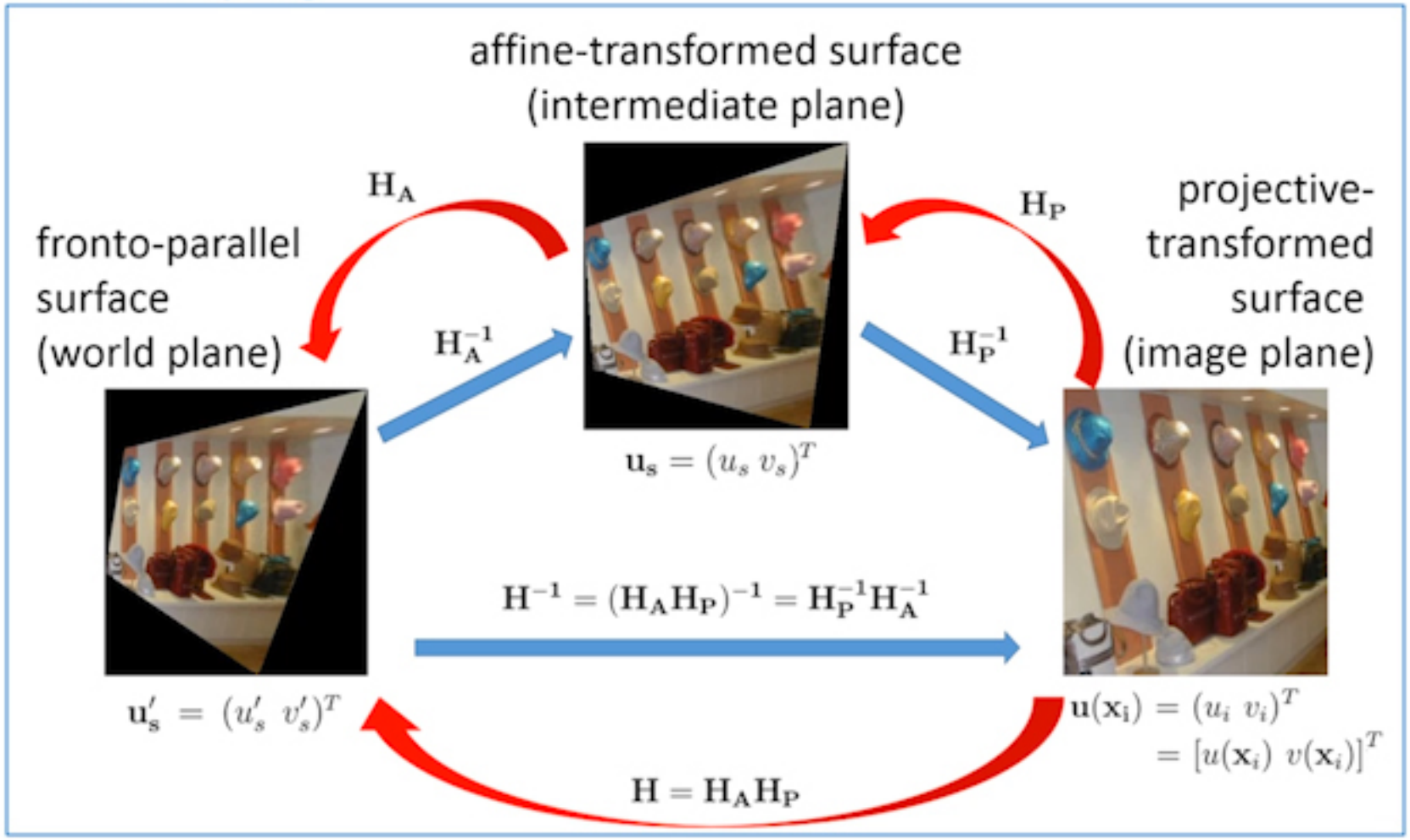
Shahzor Ahmad and Loong-Fah Cheong

National University of Singapore

Motivation & Challenges

Recovering Affine-Ambiguous Homography

- Minimize reprojection error in estimated image frequency to recover h_7, h_8



Estimating Optimal Image Frequency

$$E(f) = \sum_{p \in \mathcal{P}} D_p(f_p) + \sum_{\{p,q\} \in \mathcal{N}} V_{p,q}(f_p, f_q)$$

Graph Cuts Optimization (GCO)

Gabor filter center radial frequency

$$f_p = (\Omega_p, \theta_p) \in \mathcal{L}$$

radial orientation

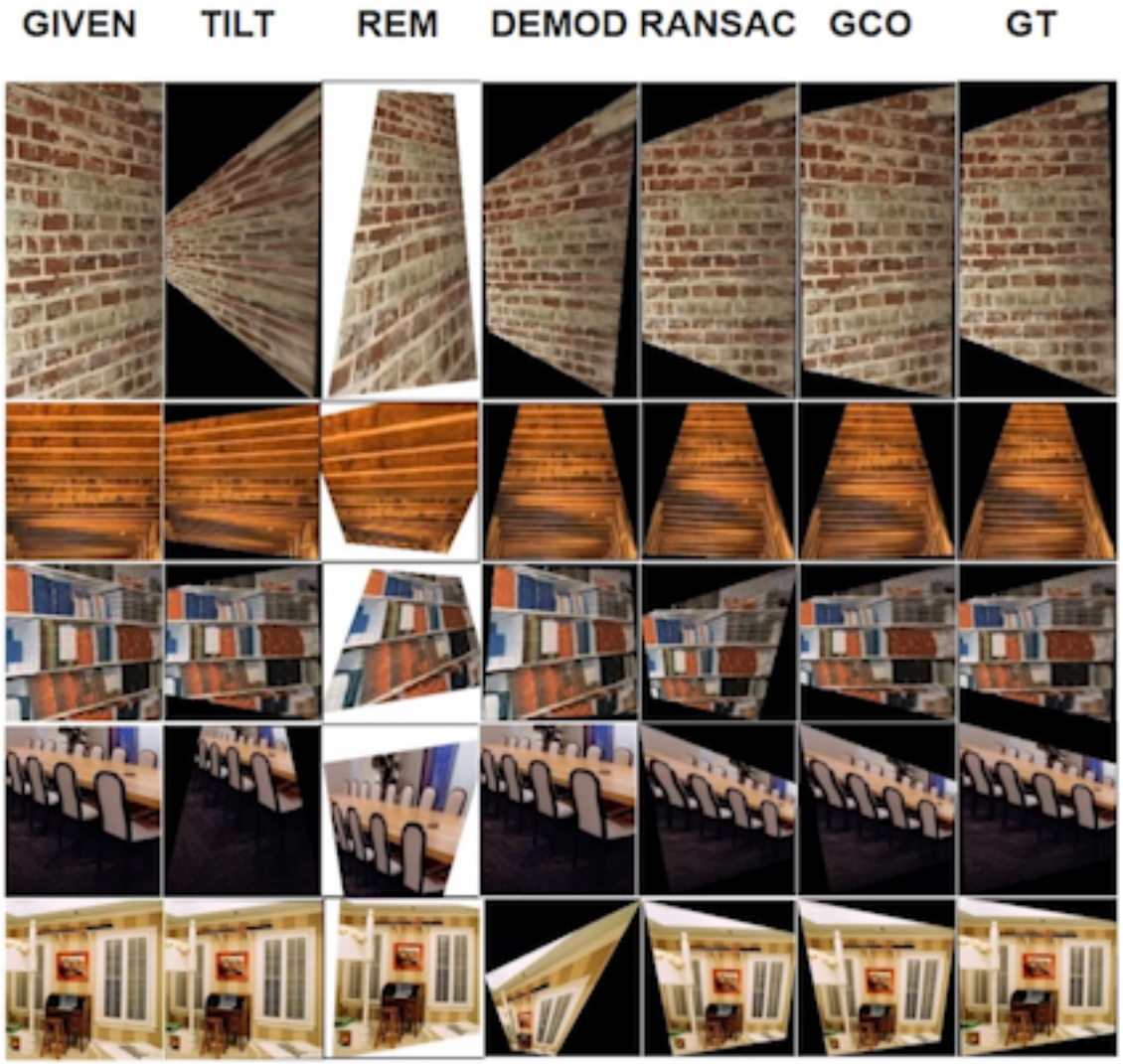
Unary Term:
favours Gabor maximizing response at a pixel
(DEMOD, Super & Bovik, 1995)

$$D_p(f_p) = \frac{\alpha}{A(f_p; p)}$$

Pairwise Term:
smooths radial frequency and orientation

$$V(f_p, f_q) = \beta(\Omega_p - \Omega_q)^2 + \gamma\{(sin\theta_p - sin\theta_q)^2 + (cos\theta_p - cos\theta_q)^2\}$$

Affine Rectification (1/2)



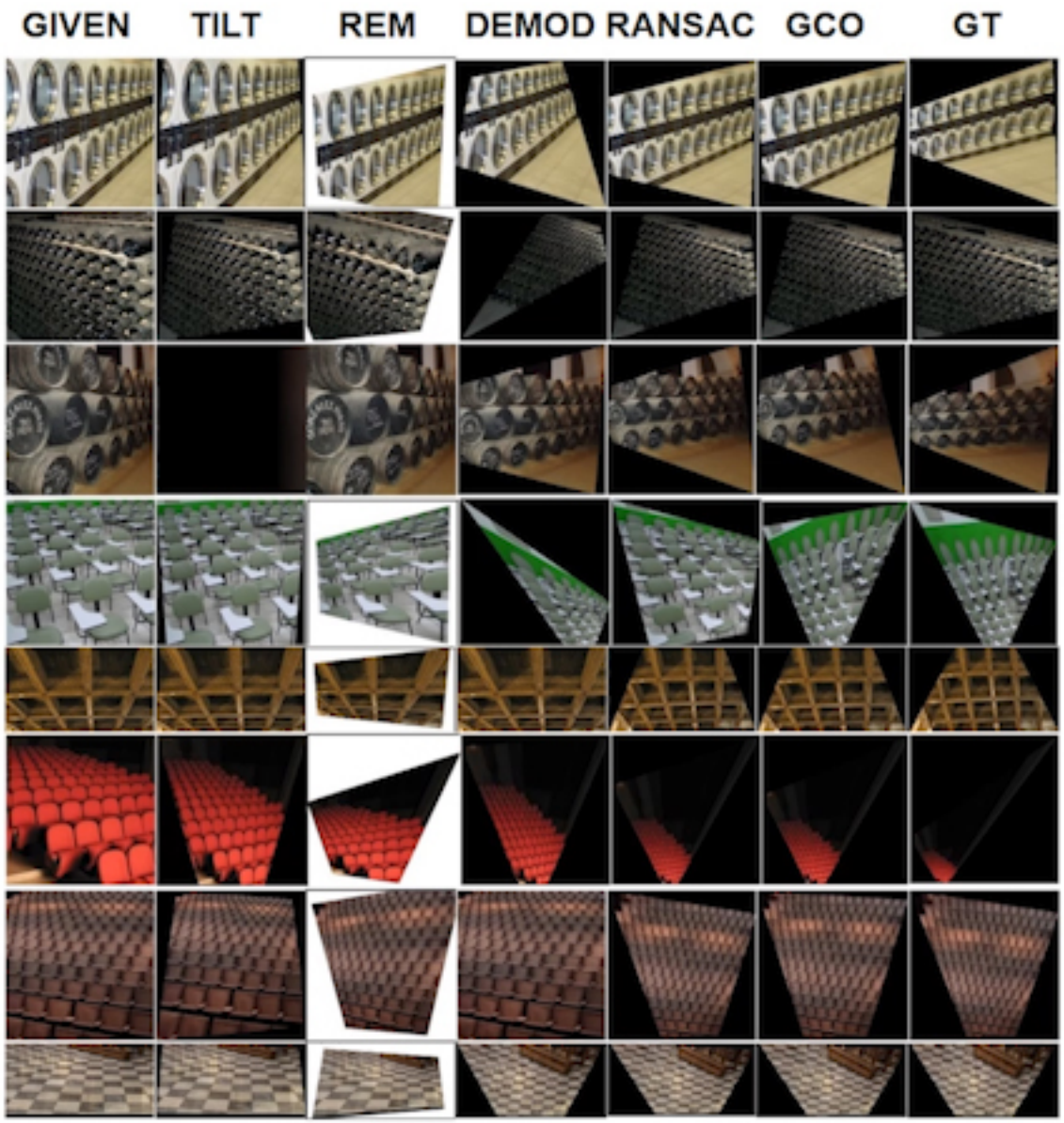
Texture with limited spatial support and/or outliers / clutter

TILT: Transform-Invariant Low-Rank Texture (Zhang et. al, 2010)

REM: Repetition Maximization (Aiger et. al. 2012)

DEMOD: Demodulation (Super and Bovik 1995)

Affine Rectification (2/2)



non low-rank texture

significant perspective distortion

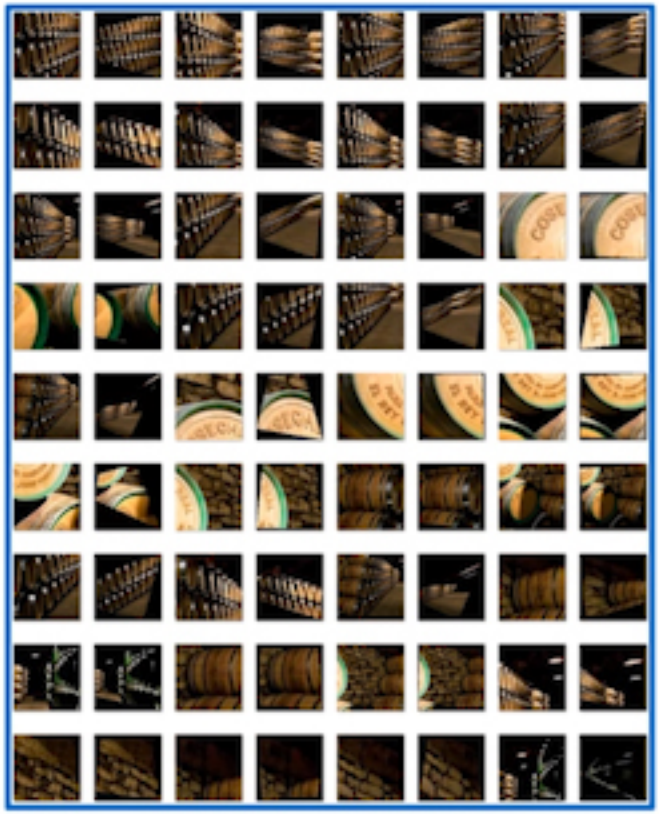
photometric severities

Detection in the Wild

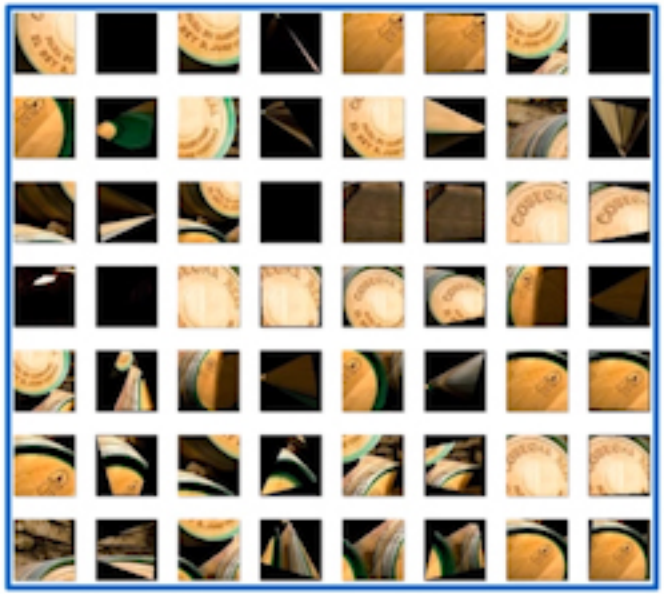
wine_cellar

large perspective distortion, illumination variation

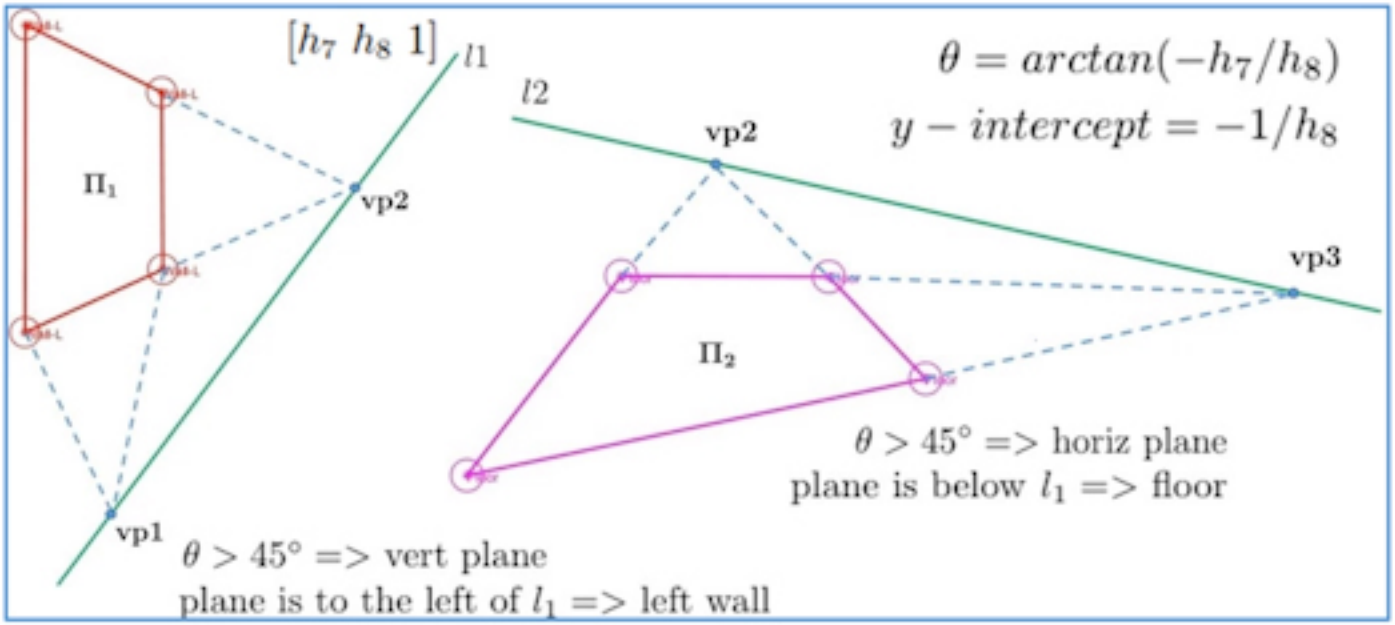
proposed (GCO)



TILT

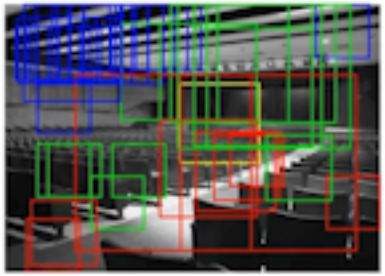


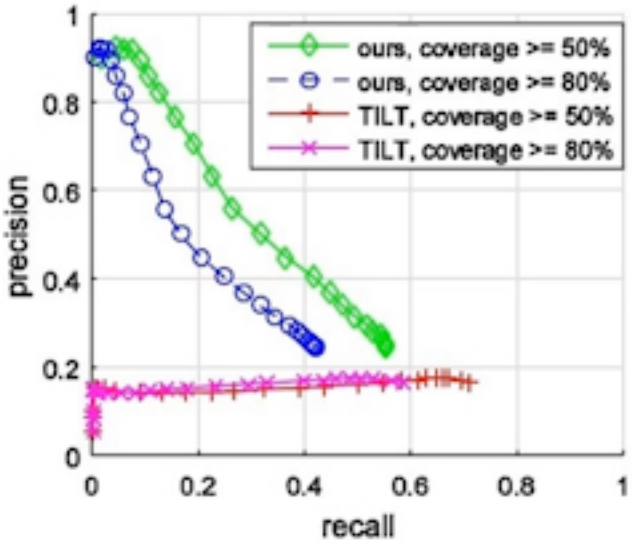
Geometric Class Assignment



annotated

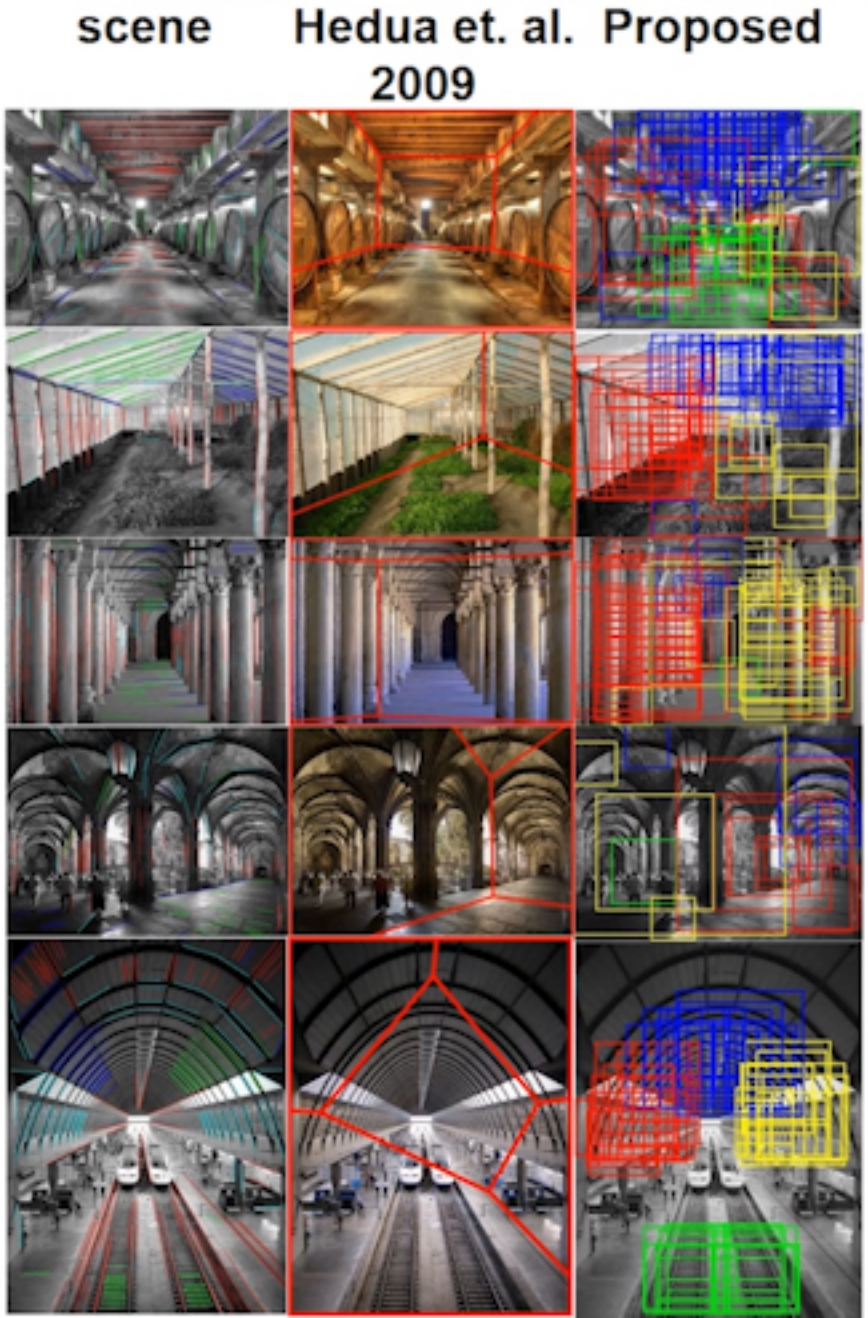
automatic





Our AP = 0.53 vs.
TILT's AP = 0.15

Scene Geometric Layout Estimation



incorrect face localization

>3 principal directions

no straight lines in a principal direction

forked layout

non-Manhattan scene structure

Indoor Scene Recognition (1/2)

Rectification mitigates in-class variation, hence can improve classification performance

Indoor Scene Recognition (2/2)

MIT Indoor 67 (Quattoni & Torralba 2009)

Representation	% Accuracy
LBP_u2(16,2)	37.10%
LBP_u2_Rect(16,2)	**40.84%**
LBP_u2 + LBP_u2_Rect	**41.28%**
CEN	46.44%
CEN_Rect	46.30%
CEN + CEN_Rect	**50.22%**
SIFT	59.14%
SIFT_Rect	57.98%
SIFT + SIFT_Rect	**61.00%**
HOG	57.69%
HOG_Rect	56.65%
HOG + HOG_Rect	**60.42%**
CEN + SIFT + HOG	61.66%
SIFT_Rect + HOG_Rect	60.88%
CEN + SIFT + HOG + SIFT_Rect + HOG_Rect	**64.54%**

Rectification produces features that are discriminative and complementary to regular, non-rectified features.

No learning for feature extraction!

NUS
National University
of Singapore

Thank You!