

A Comparative Study for Single Image Blind Deblurring

Wei-Sheng Lai¹, Jia-Bin Huang², Zhe Hu¹, Narendra Ahuja², Ming-Hsuan Yang¹

¹University of California, Merced ²University of Illinois, Urbana-Champaign

Code and Datasets available at : http://bit.ly/deblur_study

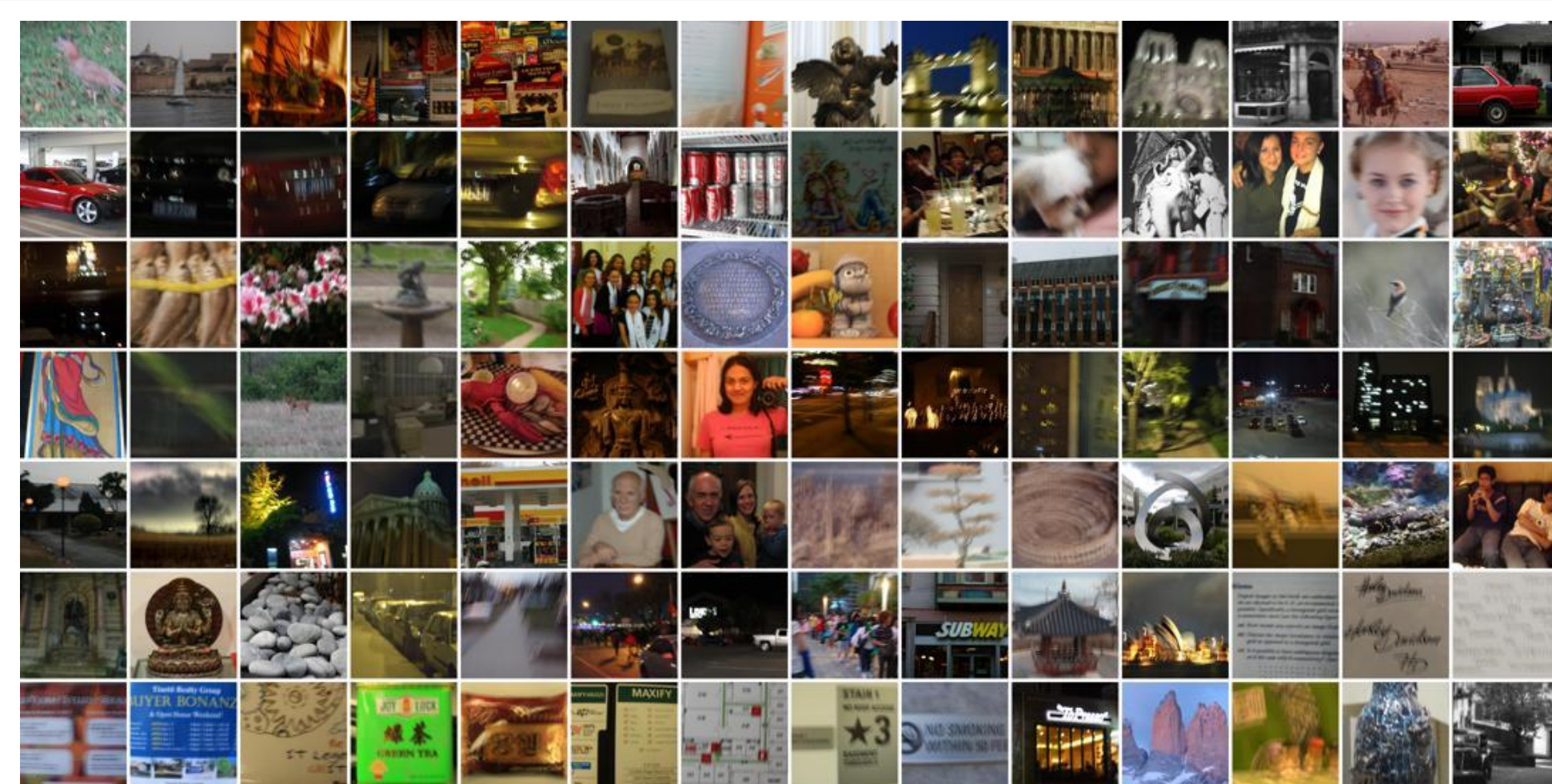
Goal

Performance evaluation of single image blind deblurring algorithms on **real-world blurred images**

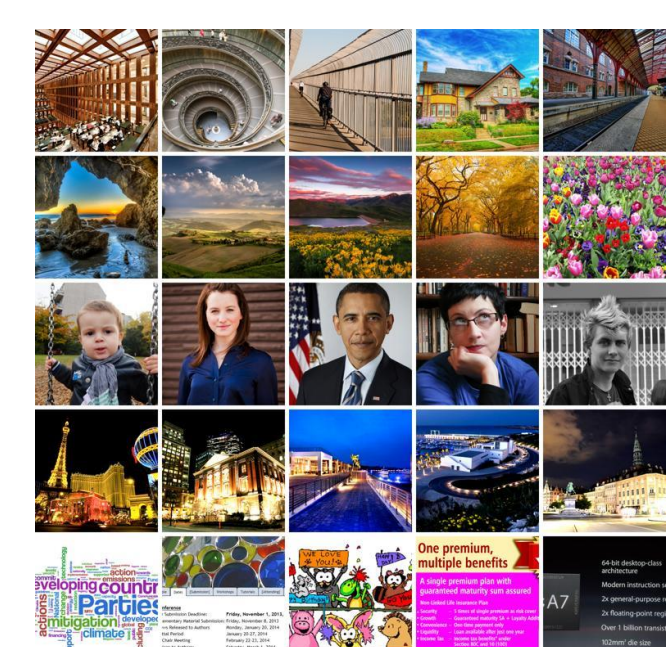
Contributions

1. A dataset of real blurred images
2. Perceptual evaluation of 13 state-of-the-art algorithms on both real and synthetic datasets
3. Performance comparisons of full-reference and no-reference image quality metrics.

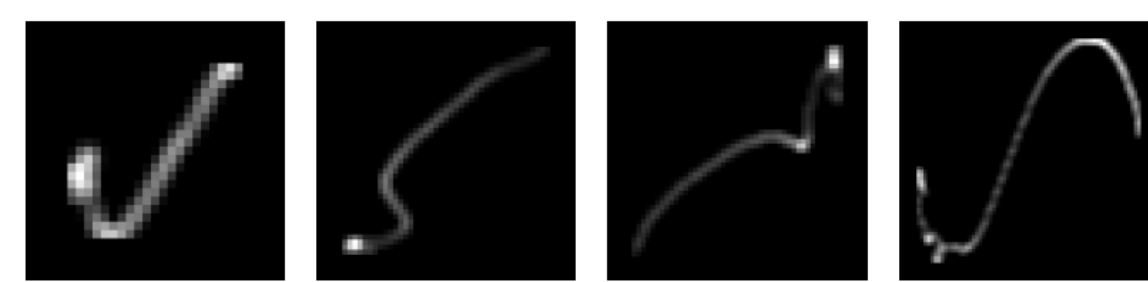
Datasets



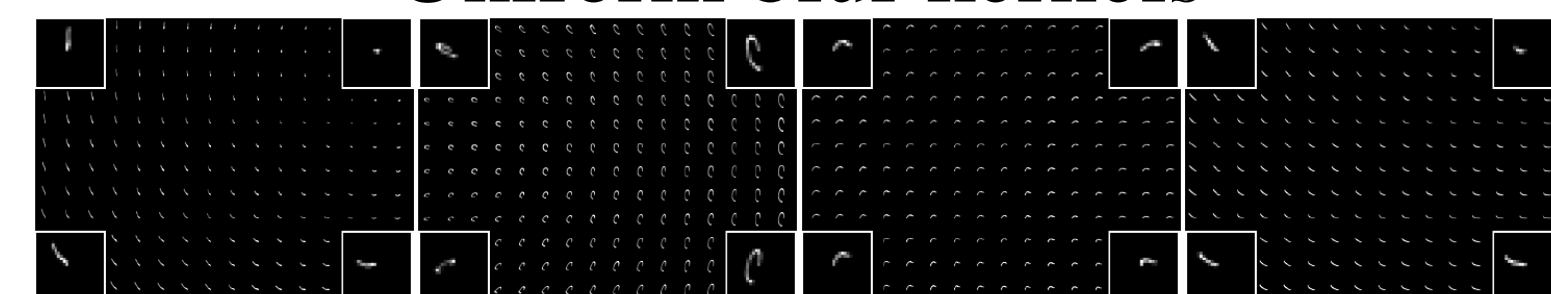
Real dataset



Synthetic dataset



Uniform blur kernels



Non-uniform blur kernels

Analyzing Users' Responses

Bradley-Terry model:

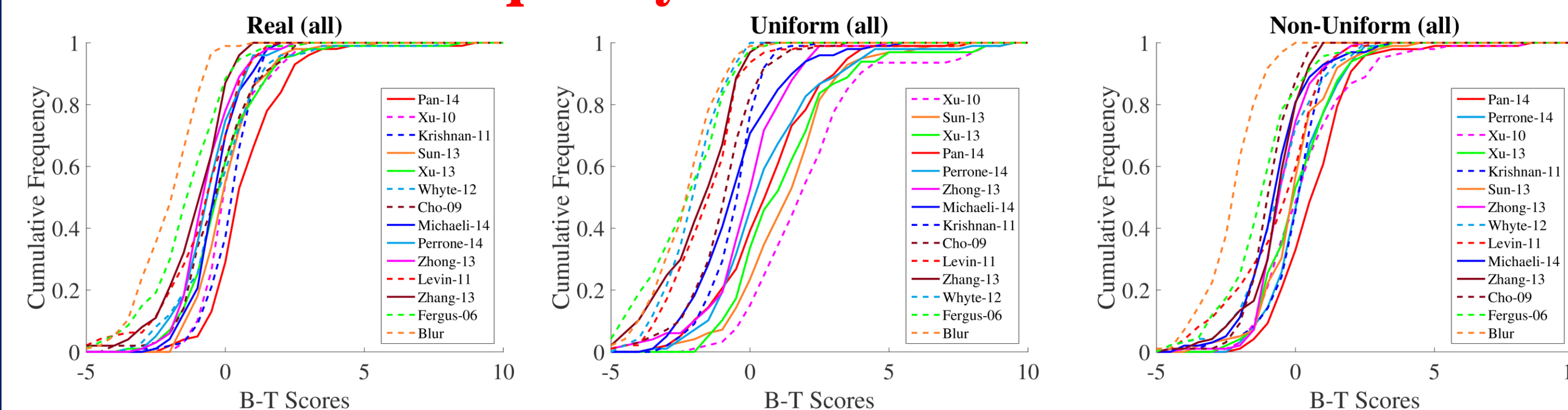
pairwise comparisons (user votes) → real-value B-T scores.

- c_{ij} : #times users choose method i over method j
- s_i : the B-T score of method i
- Probability of choosing method i over method j :

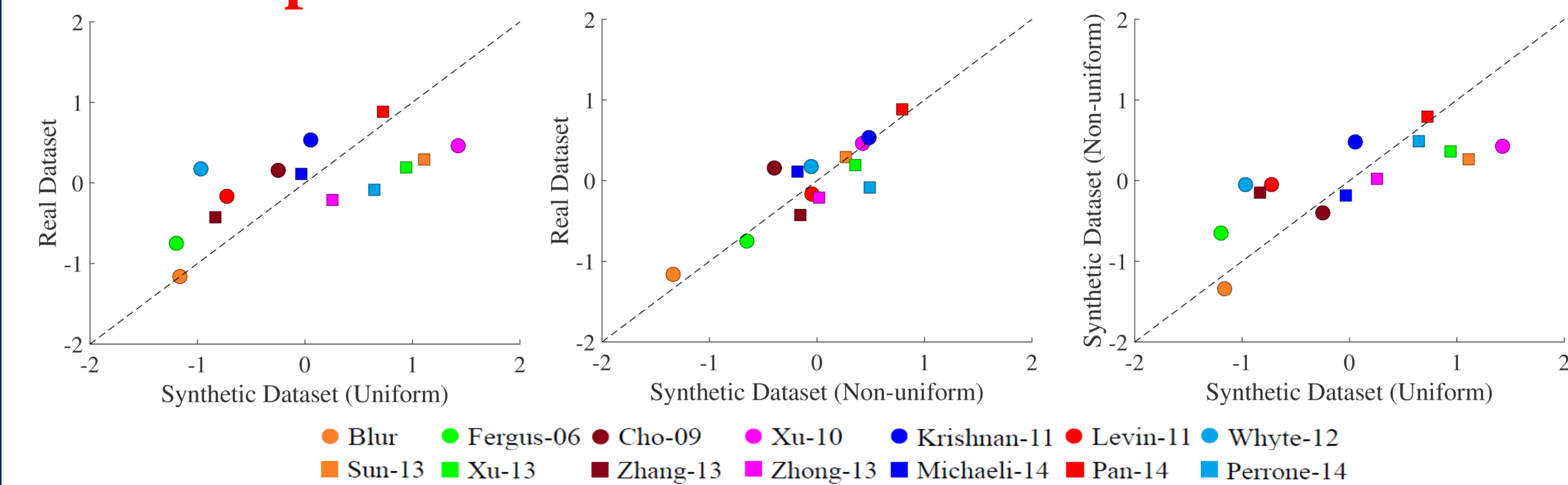
$$p_{ij} = \frac{\exp(s_i)}{\exp(s_i) + \exp(s_j)}$$

- Likelihood function: $L = \prod_{i=1}^M \prod_{j=1, i \neq j}^M p_{ij}^{c_{ij}}$
- Solve B-T scores s_i by maximizing the log-likelihood

Cumulative frequency of B-T scores:



Scatter plots of B-T scores between datasets:



Significance Test

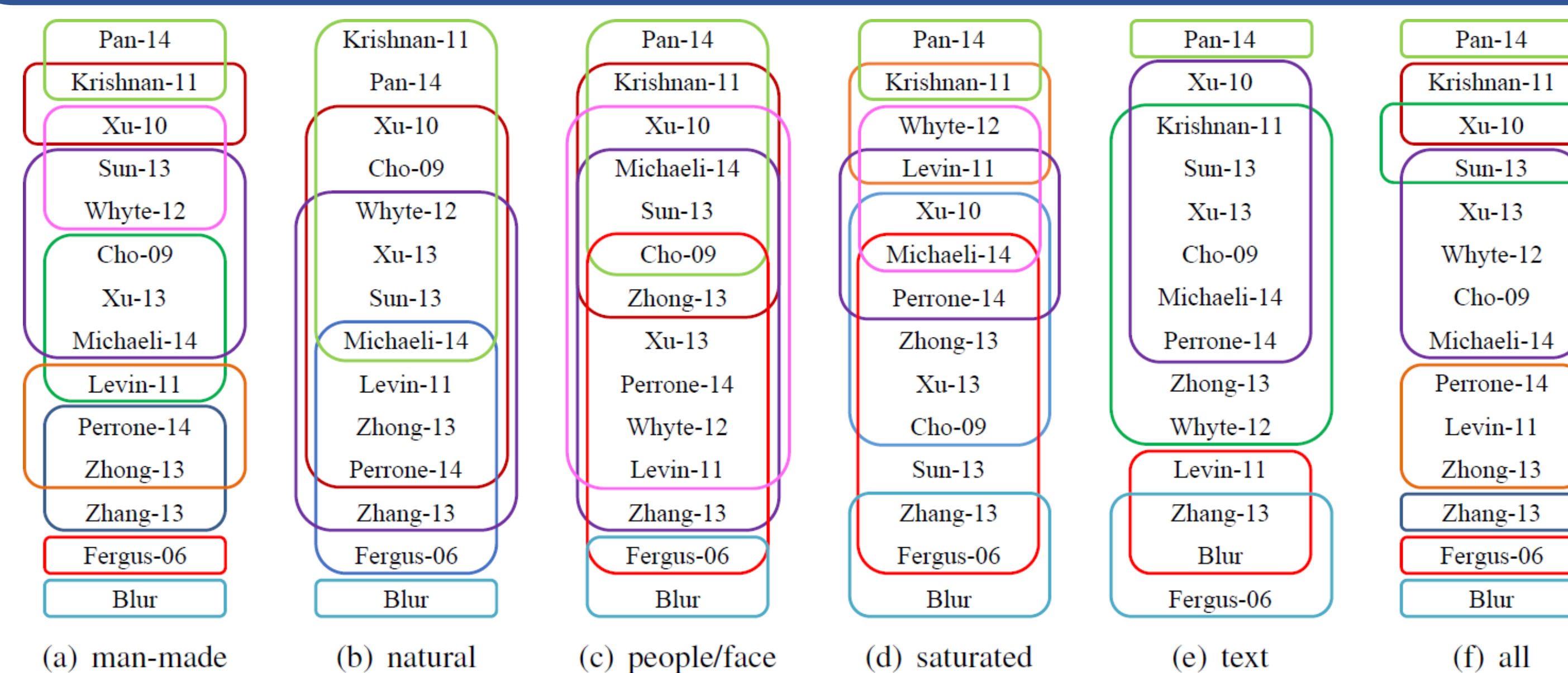
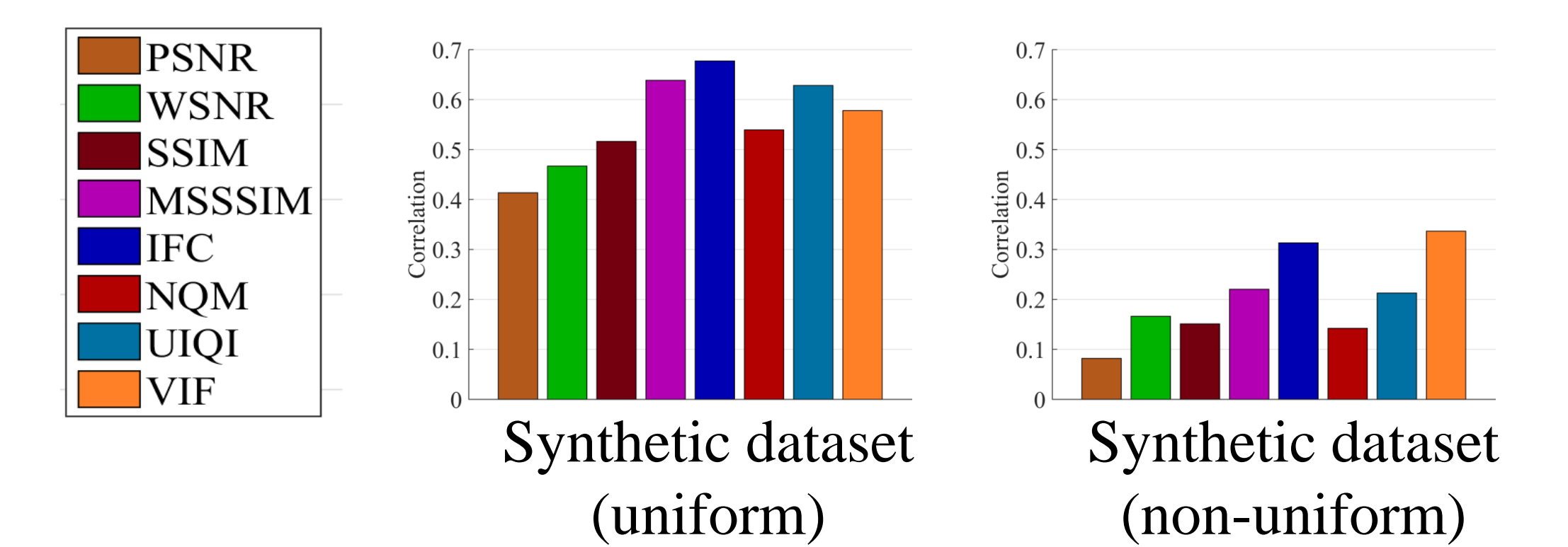


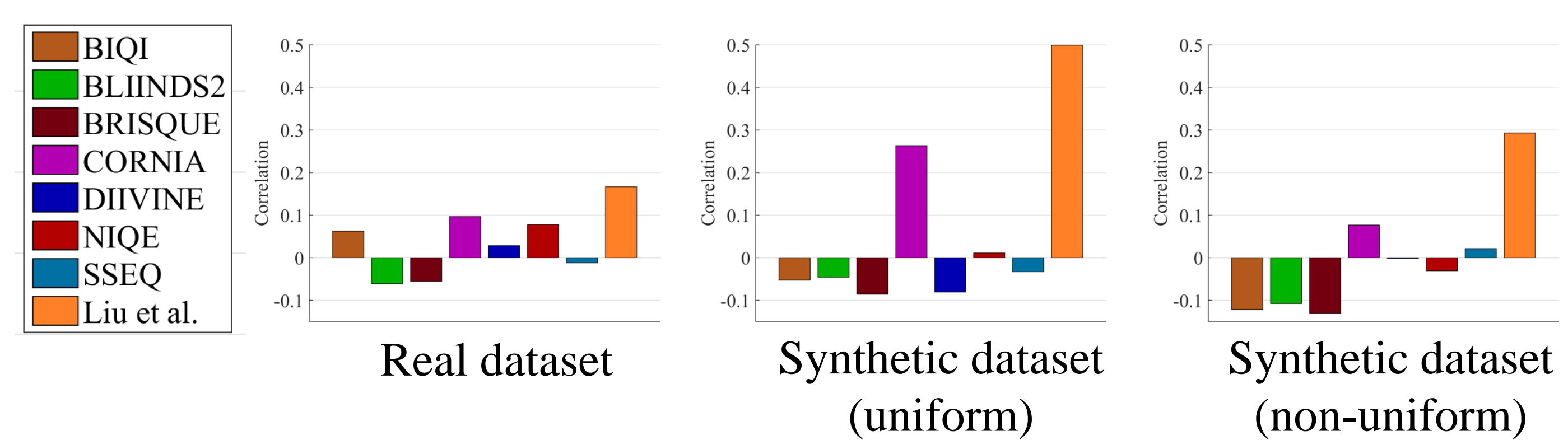
Image Quality Metrics

Spearman's rank correlation coefficients (y-axis): correlation between image quality metrics and human subject scores (B-T scores)

- Full-reference Metrics:



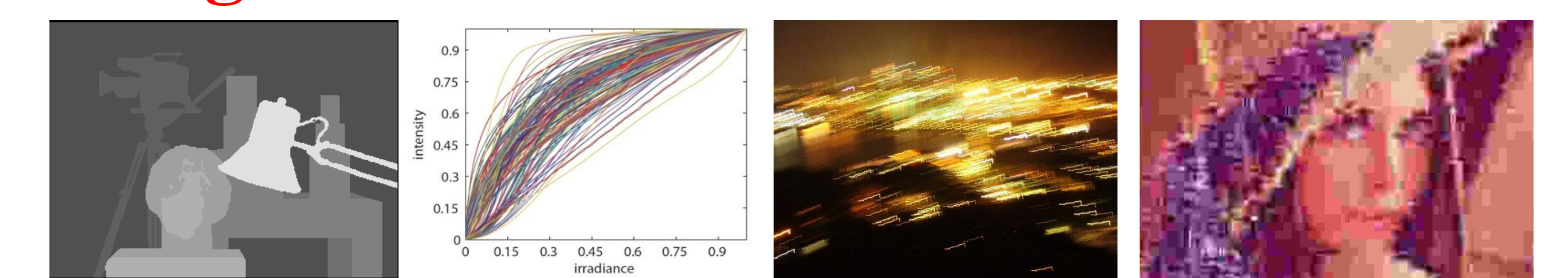
- No-reference Metrics:



Key Observations

1. **Image priors**: sparse gradient and intensity priors are more robust than edge prediction methods on handling real images.

2. **Image formulation**:



Depth CRF Saturation Compression

3. **Datasets**: performance on synthetic datasets does not reflect the performance on real-world blurred images.

4. **Image quality metrics**: IFC/VIF > PSNR/SSIM; all no-reference metrics do not correlate with human perception well.

Dataset	Levin et al. CVPR 2009	Sun et al. ICCP 2013	Kohler et al. ECCV 2012	Ours (real)	Ours (synthetic)
Synthetic/Real	Synthetic	Synthetic	Real	Real	Synthetic
Blur Model	Uniform	Uniform	Non-Uniform	Unknown	Both
Latent Images	4	80	4	100	25
Kernels / Trajectories	8	8	12	100	8
Blurred Images	32	640	48	100	200
Depth Variation	No	No	No	Yes	No
Evaluation	PSNR/SSIM	PSNR/SSIM	PSNR	User-Study	User-Study