Optimal Structured Light à la Carte

Computer Science
UNIVERSITY OF TORONTO

Parsa Mirdehghan Wenzheng Chen Kyros Kutulakos

More information can be found at:
http://www.dgp.toronto.edu/optimalsl

2018 CVPR SALT LAKE CITY • JUNE 18-22

Key Objective: Maximally Accurate & Robust Structured Light

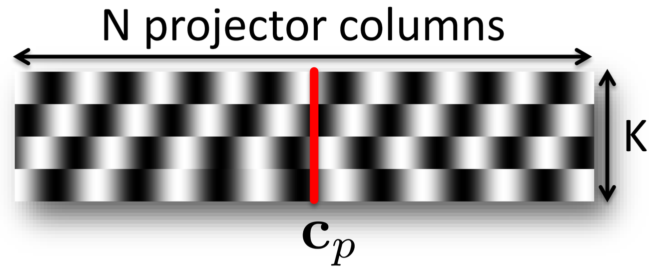
We optimize 3D reconstruction performance of any given projector-camera system

Our emphasis is on the most challenging cases of structured-light triangulation:

- fast acquisition (⇔ very small # of projection patterns)
- low-power projectors (⇔ low-SNR images)

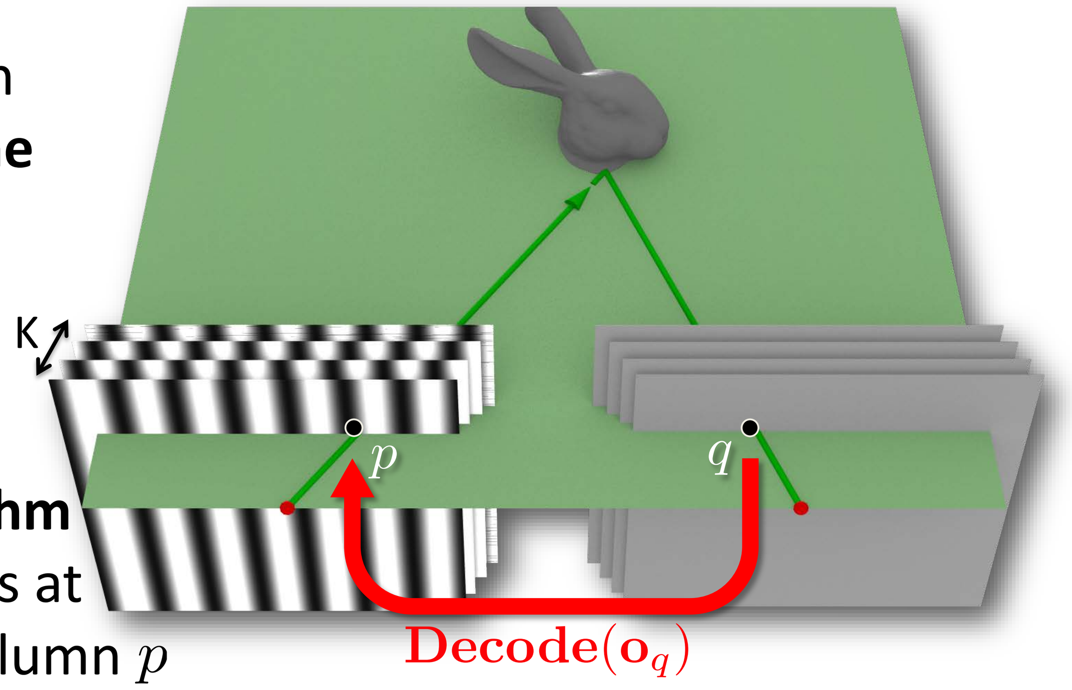
K-Pattern Structured Light: Traditional Approach

1. Adopt an **abstract encoding scheme** that assigns a K-dimensional vector $\mathbf{c}_p$ to each projector column p (e.g., shifted cosines, MPS codes [1], gray codes, XOR codes [3])



2. This defines a sequence of K projection patterns that are used **regardless of the projector-camera system's specs or the scene's geometry & appearance**

3. An **encoding-specific decoding algorithm** then maps the vector of K observations at camera pixel q to its corresponding column p



Our Approach: Optimal Structured Light Framework

1. Structured light formulated as a **maximum-likelihood (ML) classification task** (pixel label ⇔ corresponding projector column)

2. We derive an **encoding-independent decoding algorithm** that is **near-optimal** & **trivial to implement**

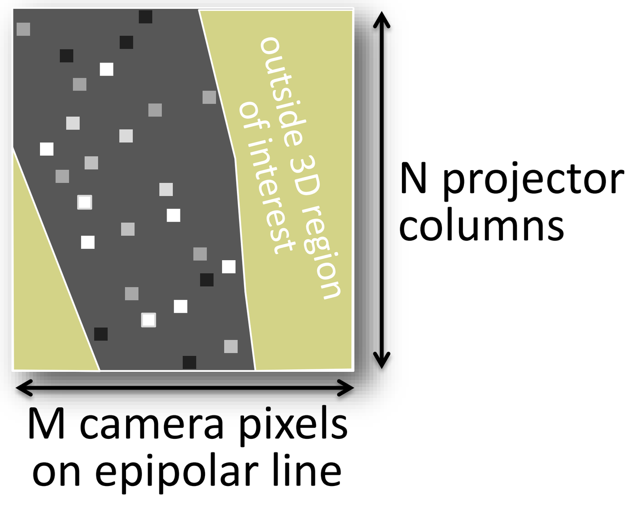
$$\text{Decode}(\underbrace{\mathbf{C}}_{\text{encoding scheme}}, \underbrace{\mathbf{o}_q}_{\text{observation vector}}) = \arg\max_{1\le p\le N} \text{Prob}(\mathbf{o}_q \mid \mathbf{c}_p)$$

3. We derive an efficient algorithm to **generate column encoding schemes on the fly, optimized for the imaging conditions at hand**

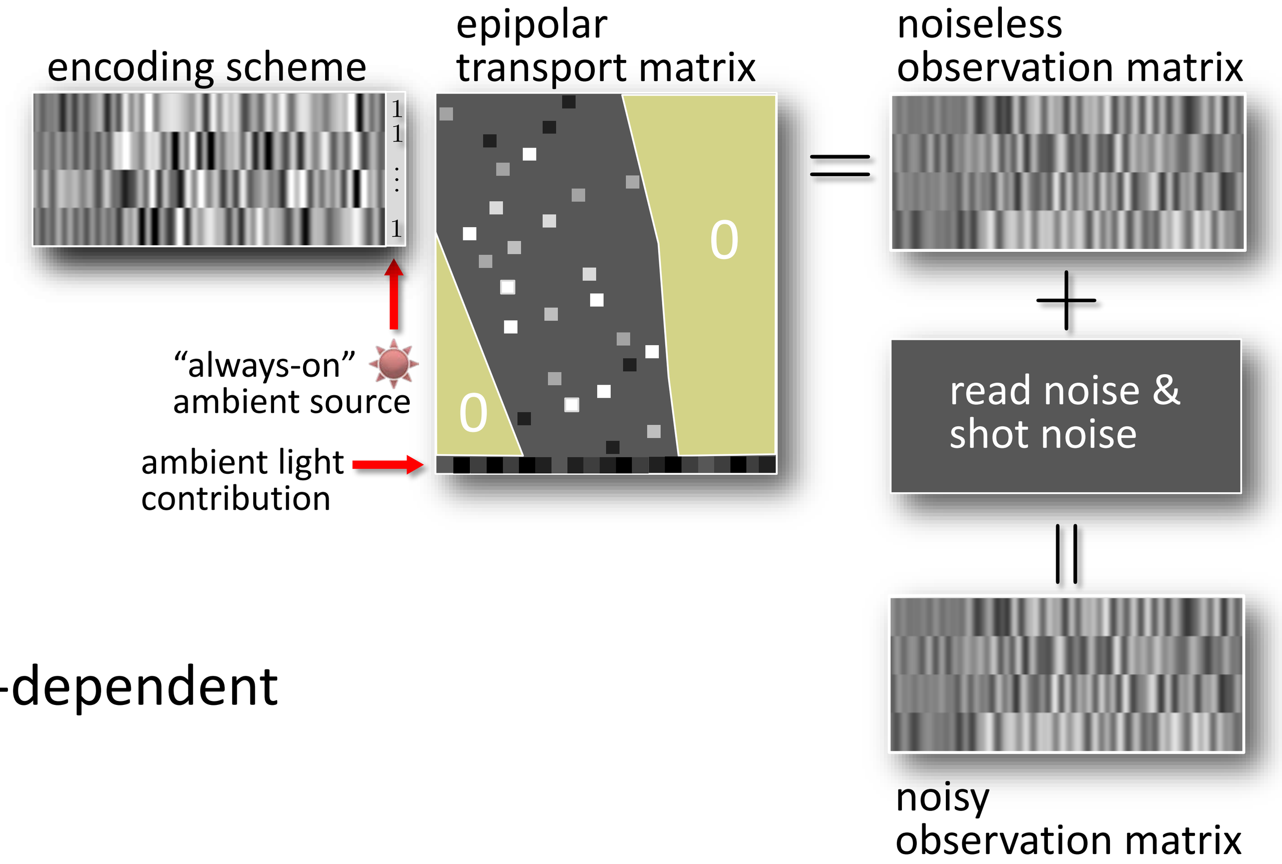
$$\underbrace{\mathbf{C}^*_\epsilon}_{\text{optimal encoding}} = \arg\min_{\mathbf{C}} \underbrace{\mathbb{E}}_{\text{expectation over plausible scenes}}[\underbrace{\text{Error}}_{\text{\# of misclassified pixels in a given scene}}(\mathbf{C}, \underbrace{\epsilon}_{\text{error tolerance: min disparity error for labeling pixel as misclassified}})]$$

Epipolar-Only Generative Image Formation Model

1. randomly assign valid correspondence & albedo to pixels on epipolar line



2. incorporate encoding scheme, ambient light & camera noise



3. (optional) incorporate depth-dependent projector defocus

Contribution #1: A Simple Near-Optimal Generic Decoder

The ZNCC decoder: Choose the projector column whose vector $\mathbf{c}_p$ maximizes its zero-mean normalized cross-correlation (ZNCC) with the observation vector, where:

$$\text{ZNCC}(\mathbf{o}_q, \mathbf{c}_p) = \underbrace{\text{NCC}}_{\text{normalized cross-correlation}}(\mathbf{o}_q - \textbf{mean}(\mathbf{o}_q), \mathbf{c}_p - \textbf{mean}(\mathbf{c}_p))$$

Proposition. The ZNCC decoder is **optimal in the limit**, *i.e.* it yields the ML solution when (a) noise is additive Gaussian, (b) its standard deviation σ is sufficiently small and (c) the variance v of variances of vectors $\mathbf{c}_1$, $\mathbf{c}_2$, ... , $\mathbf{c}_N$ is sufficiently small .

$$\lim_{\substack{\sigma\to 0\\ v\to 0}} \left(\text{Decode}(\mathbf{C}, \mathbf{o}_q)\right) = \arg\max_{1\le p\le N} \text{ZNCC}(\mathbf{o}_q, \mathbf{c}_p)$$

Contribution #2: The Optimal Column Encoding Loss Function

Proposition. For a sufficiently large softmax temperature τ, the number of misclassified pixels in a given scene can be approximated as:

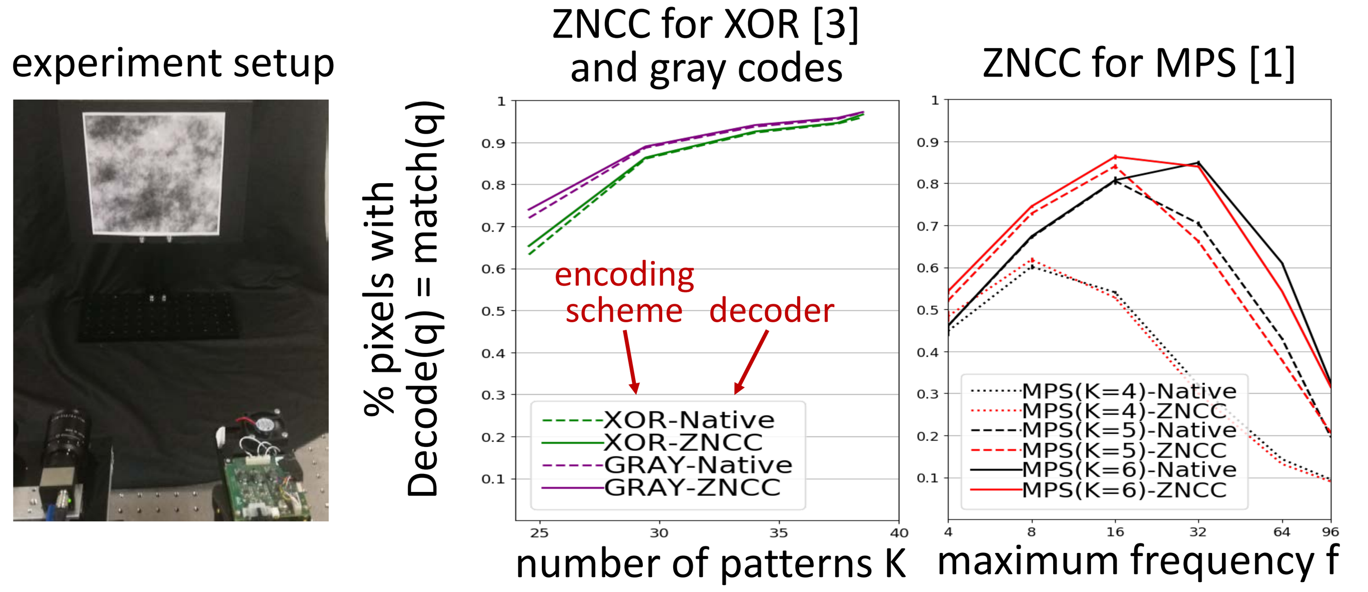
$$\text{Error}(\mathbf{C}, \epsilon) \approx M - \sum_q \text{softmax}_\tau(\text{ZNCC}(\mathbf{o}_q, \underbrace{\mathbf{c}_{\text{match}(q)}}_{\text{pixel q true corresponding projector column}}))$$

Loss function defined over S fair samples (of noise, albedo, ambient light, and correspondences):

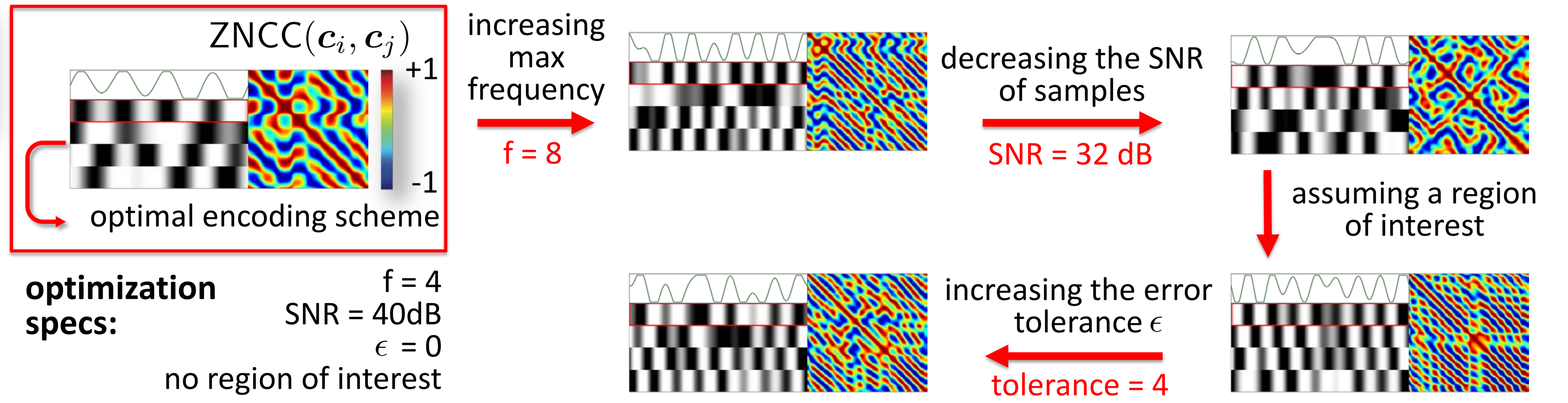
$$\mathbb{E}[\text{Error}(\mathbf{C}, \epsilon)] \approx \left(\frac{1}{S}\right) \sum_{S \text{ samples}} \text{Error}(\mathbf{C}, \epsilon)$$

Decoding Comparison: ZNCC vs. Encoding-Specific Decoding Algorithms

The decoding comparison suggests that there is potentially **no accuracy advantage** to be gained by designing encoding-specific decoding algorithm



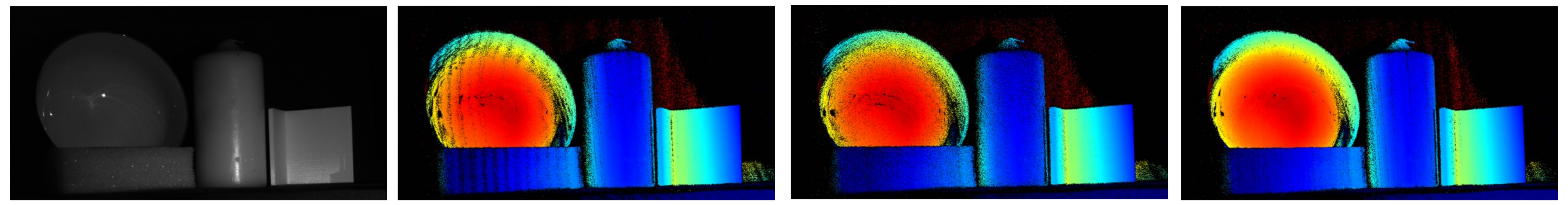
Optimal Column Encodings: Computed in < 3 mins on Standard Laptop



Depth Map Comparison: Optimal Encoding vs. State-of-the-Art Encoding

Optimal encoding optimized for the region of interest in front of camera (K=6)

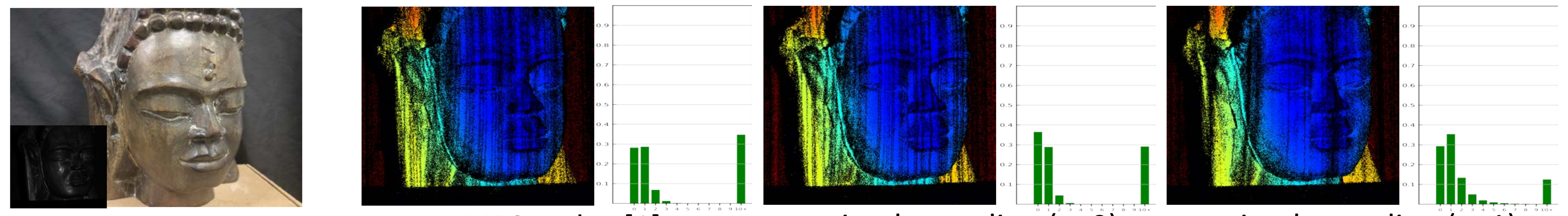
Depth maps showing the pixels whose disparity error $\le$ 1 pixel



scene MPS codes [1] EPS codes [2] optimal encoding

Optimal encoding optimized for low-SNR conditions and an error tolerance (K=4)

Depth maps showing the pixels whose disparity error $\le$1 pixel, and the distribution of disparity error



scene MPS codes [1] optimal encoding (ϵ=0) optimal encoding (ϵ=1)

References

[1] Gupta & Nayar, "Micro Phase Shifting", CVPR'12
[2] Moreno et.al, "Embedded Phase Shifting: Robust Phase Shifting with Embedded Signals", CVPR'15
[3] Gupta et.al, "A Practical Approach to 3D Scanning in the Presence of Interreflections, Subsurface Scattering and Defocus", IJCV'13
[4] Kingma et.al, "Adam: A Method for Stochastic Optimization", ICLR'15