

コンピュータで見る

西野 恒

Ko Nishino



京都大学

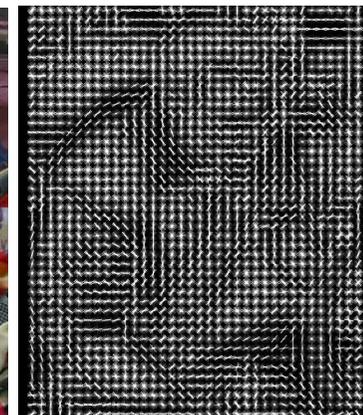
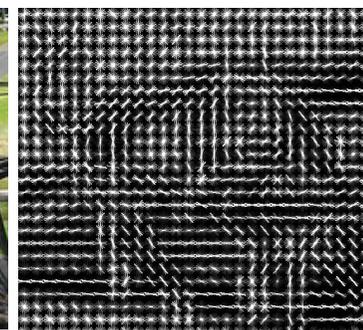
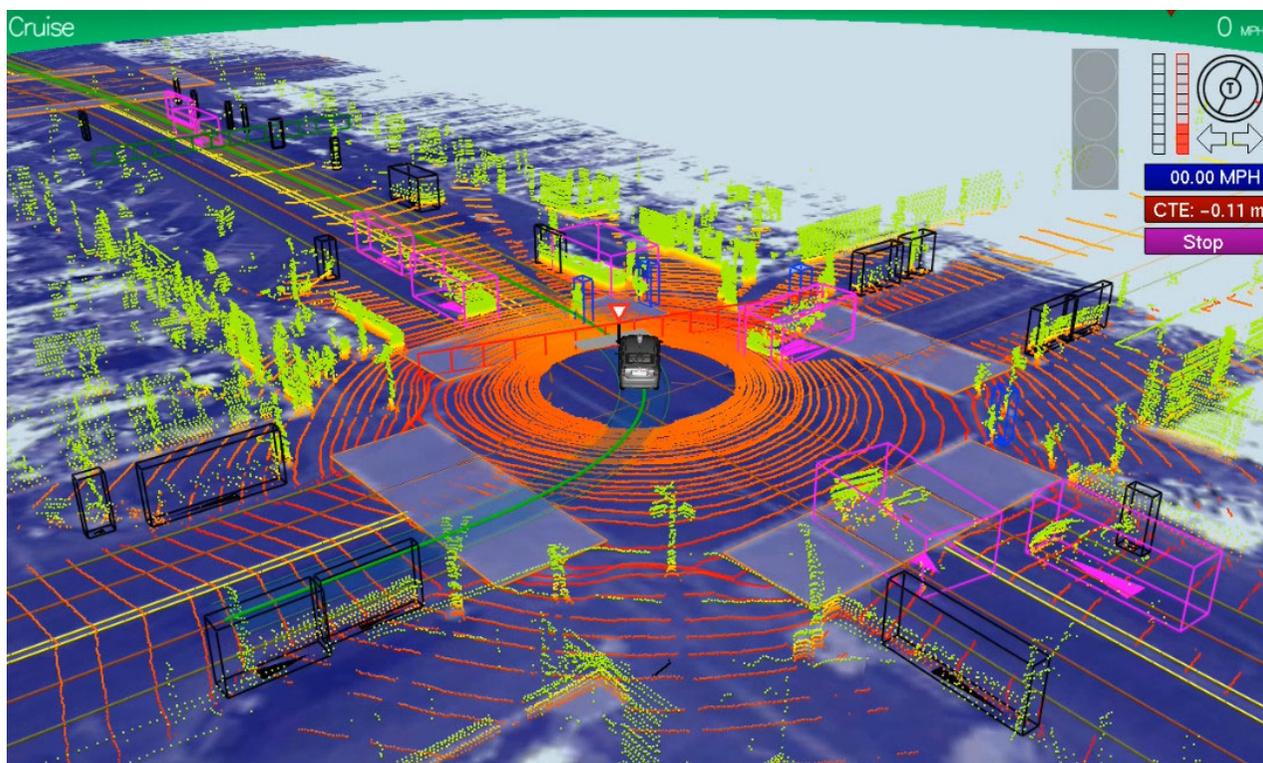
KYOTO UNIVERSITY

コンピュータビジョンの春



コンピュータビジョンは完成？

- コンピュータが見ている世界は我々の世界とは異なる

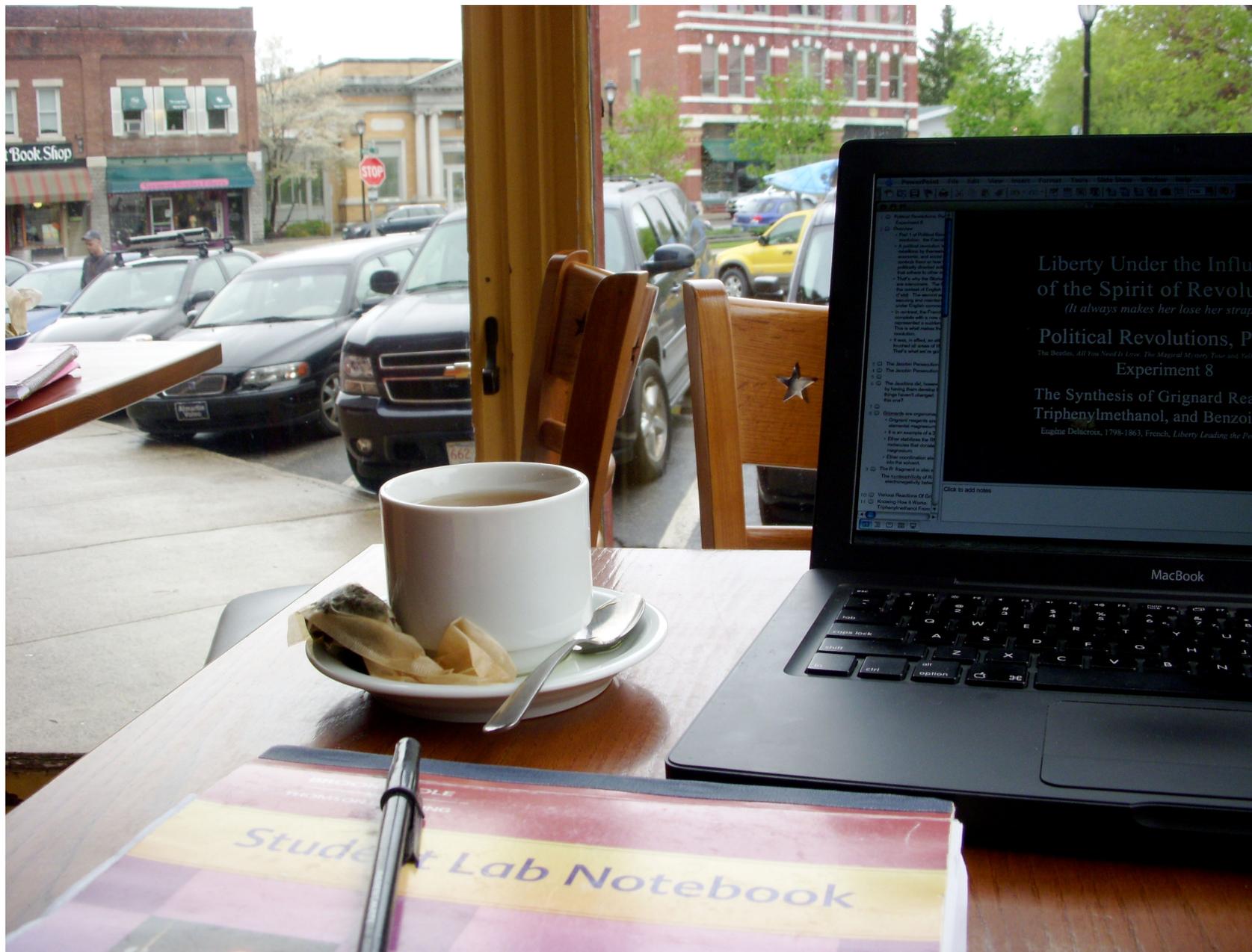


コンピュータビジョンの根本仮定

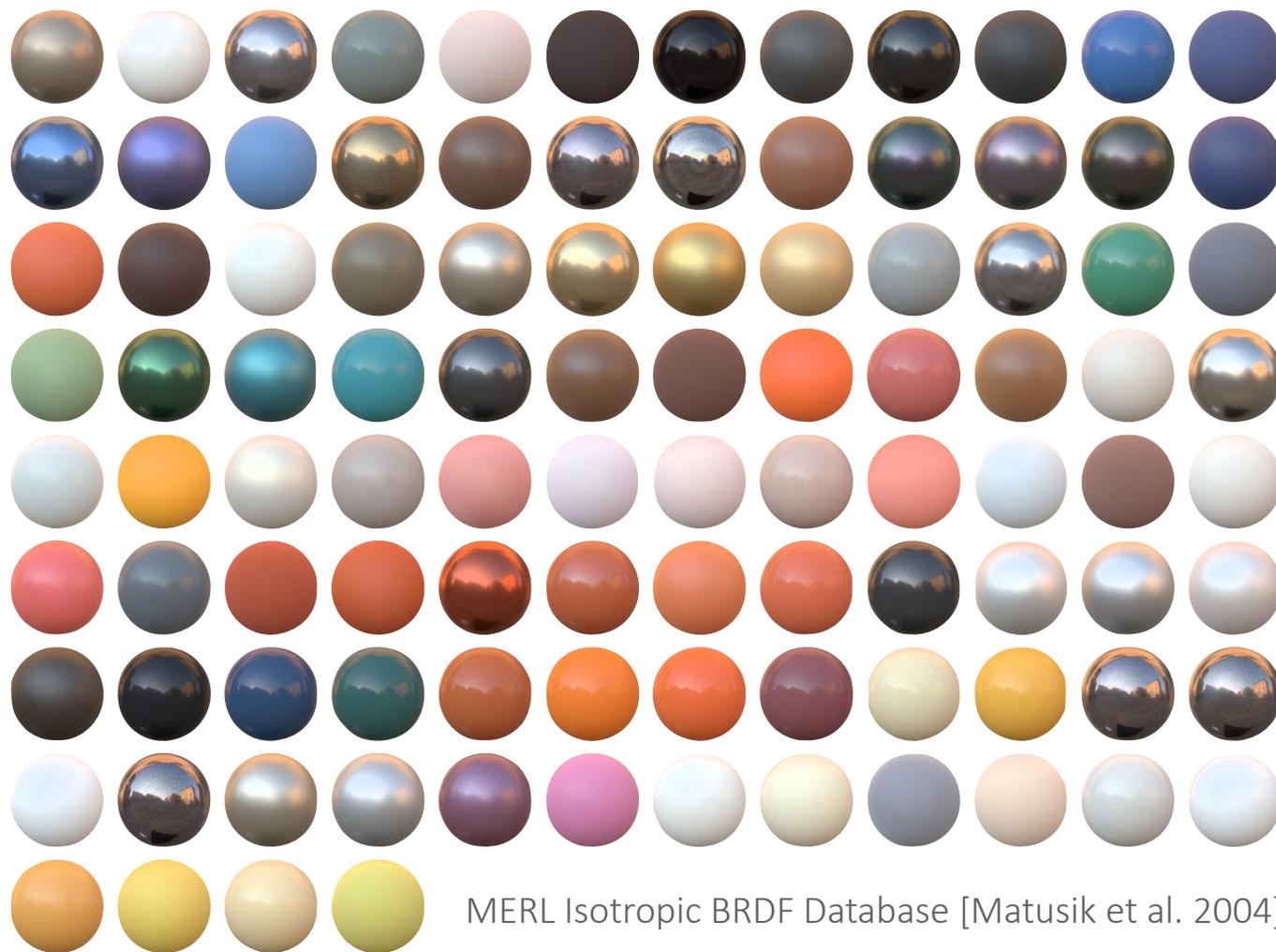
・石と紙の世界



Paper City by Maciek Janicki



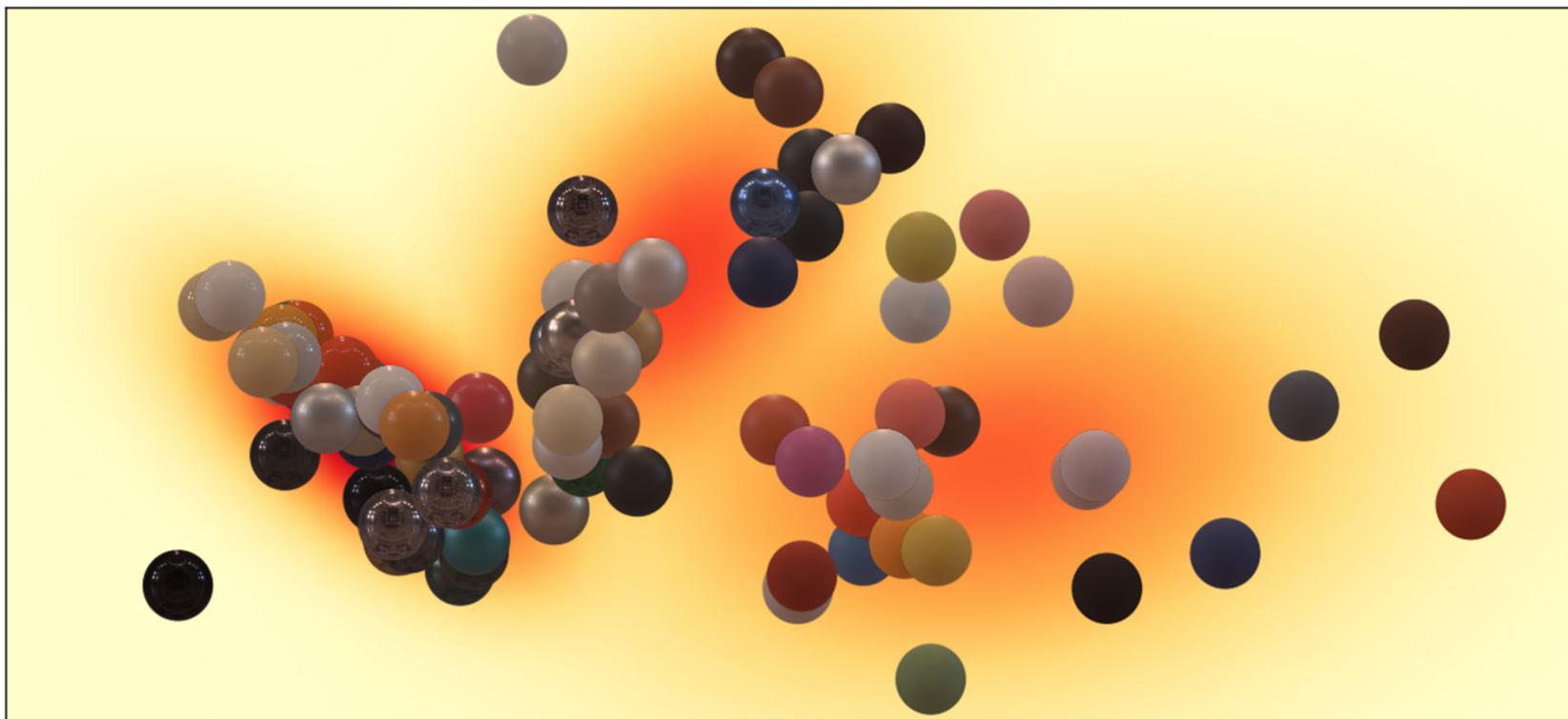
実世界の反射特性



MERL Isotropic BRDF Database [Matusik et al. 2004]

反射特性の事前分布

$$\arg \max_{\mathbf{R}, \mathbf{L}, \mathbf{N}} p(\mathbf{R}, \mathbf{L}, \mathbf{N} | \mathbf{I}) \propto p(\mathbf{I} | \mathbf{R}, \mathbf{L}, \mathbf{N}) p(\mathbf{R}) p(\mathbf{L}) p(\mathbf{N})$$



反射特性と自然光源環境の同時復元

[ECCV 12-1, PAMI 15-1]



入力画像

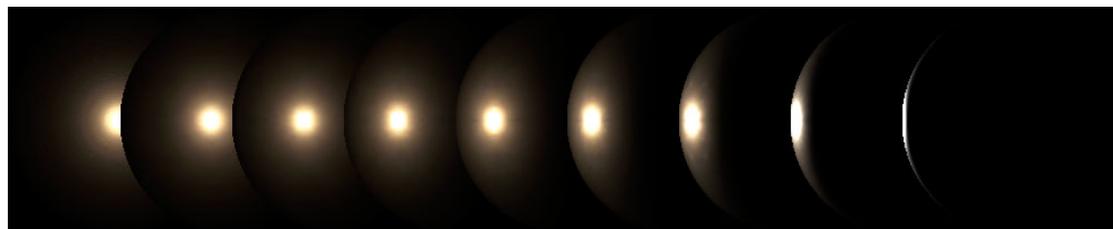


復元画像

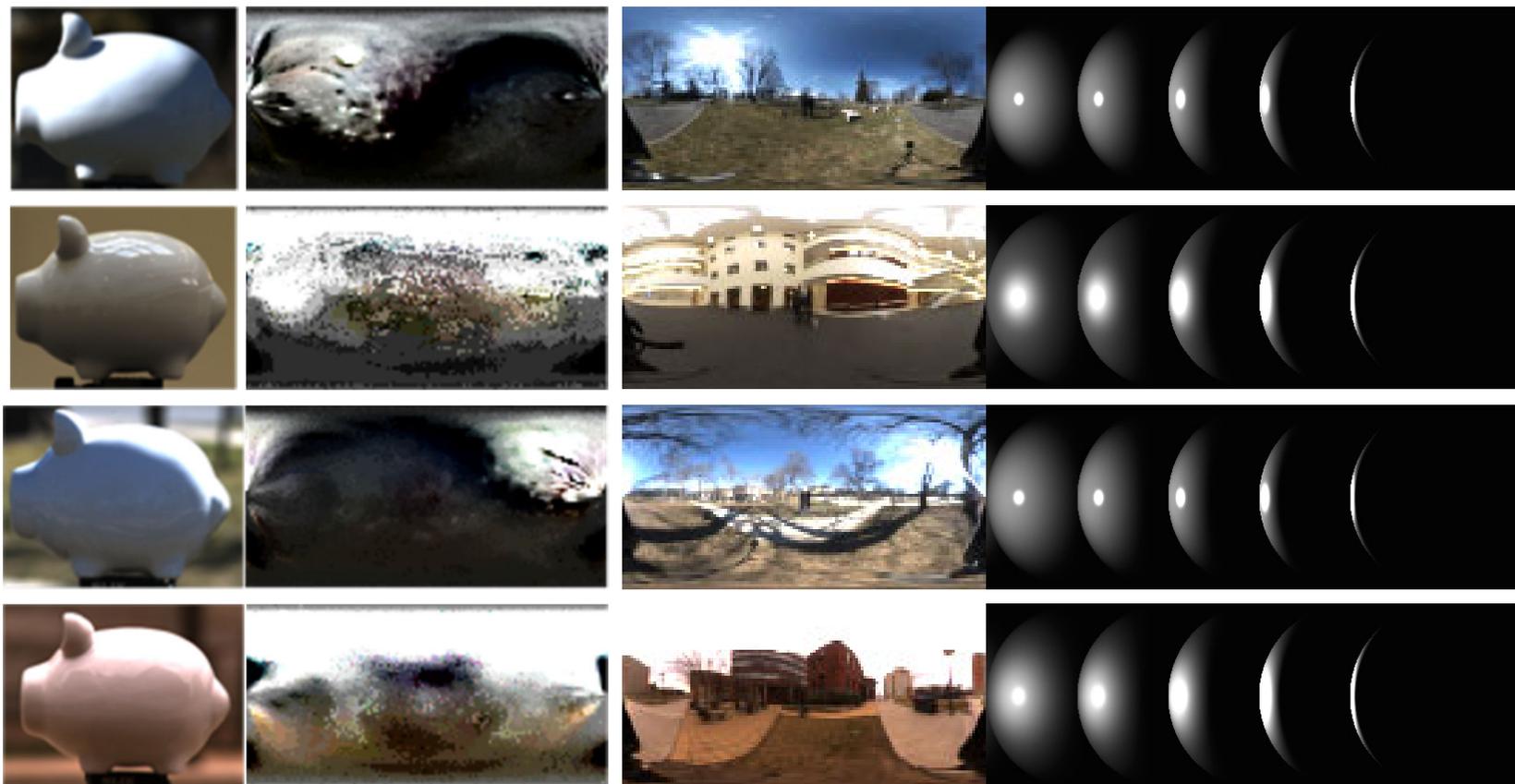
復元結果



真値



反射特性と自然光源環境の同時復元



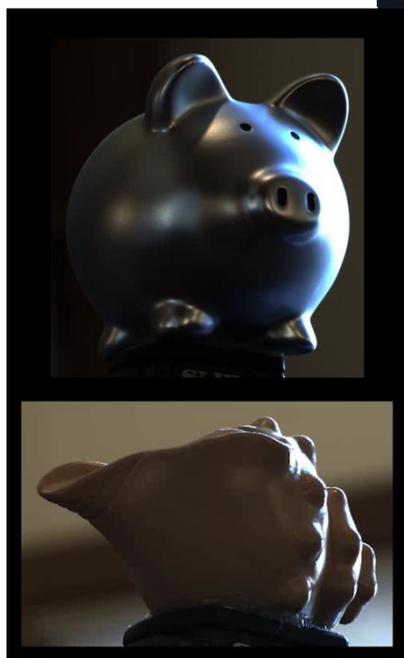
入力画像

光源環境
復元結果

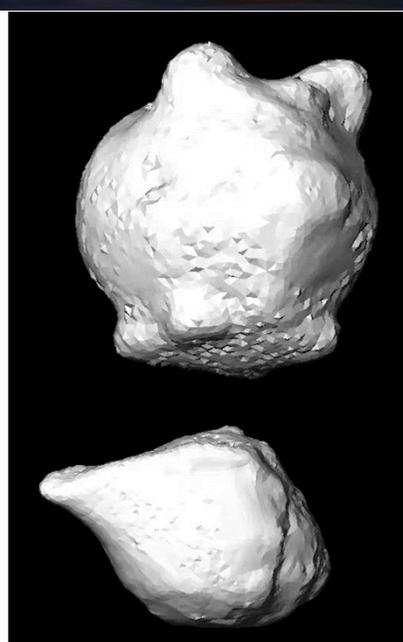
光源環境
真値

反射特性
復元結果

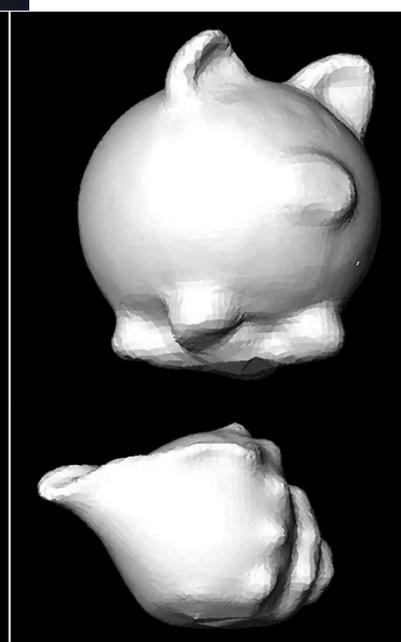
幾何形状と反射特性の同時復元



入力画像列



復元結果

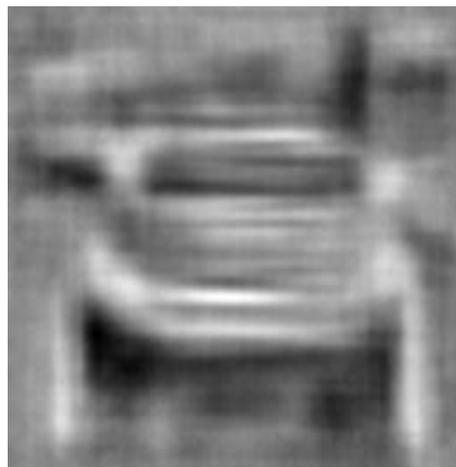
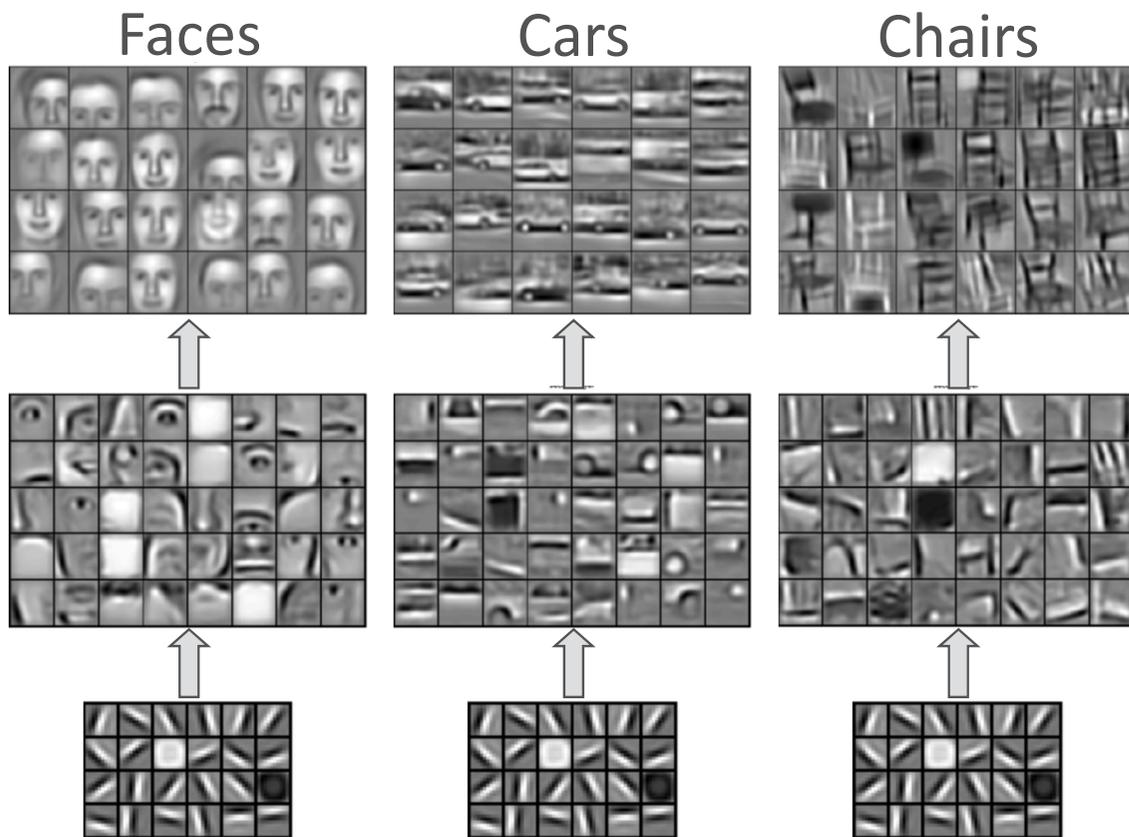


真値

[CVPR 14, PAMI 15-2]

コンピュータビジョンの根本仮定

- 多様性を排除した世界





素材に基づく画像理解



素材認識の難しさ



[Flickr Material Database, Sharan et al. 09]

- 大きなクラス内分散
- 特徴量が不明

材質属性

- 人間の素材認識のように，材質を認識することにより素材を認識できないか？
 - 材質は素材そのものより見えに発現しやすい
 - Fabric is fuzzy or woven but not hard or metallic

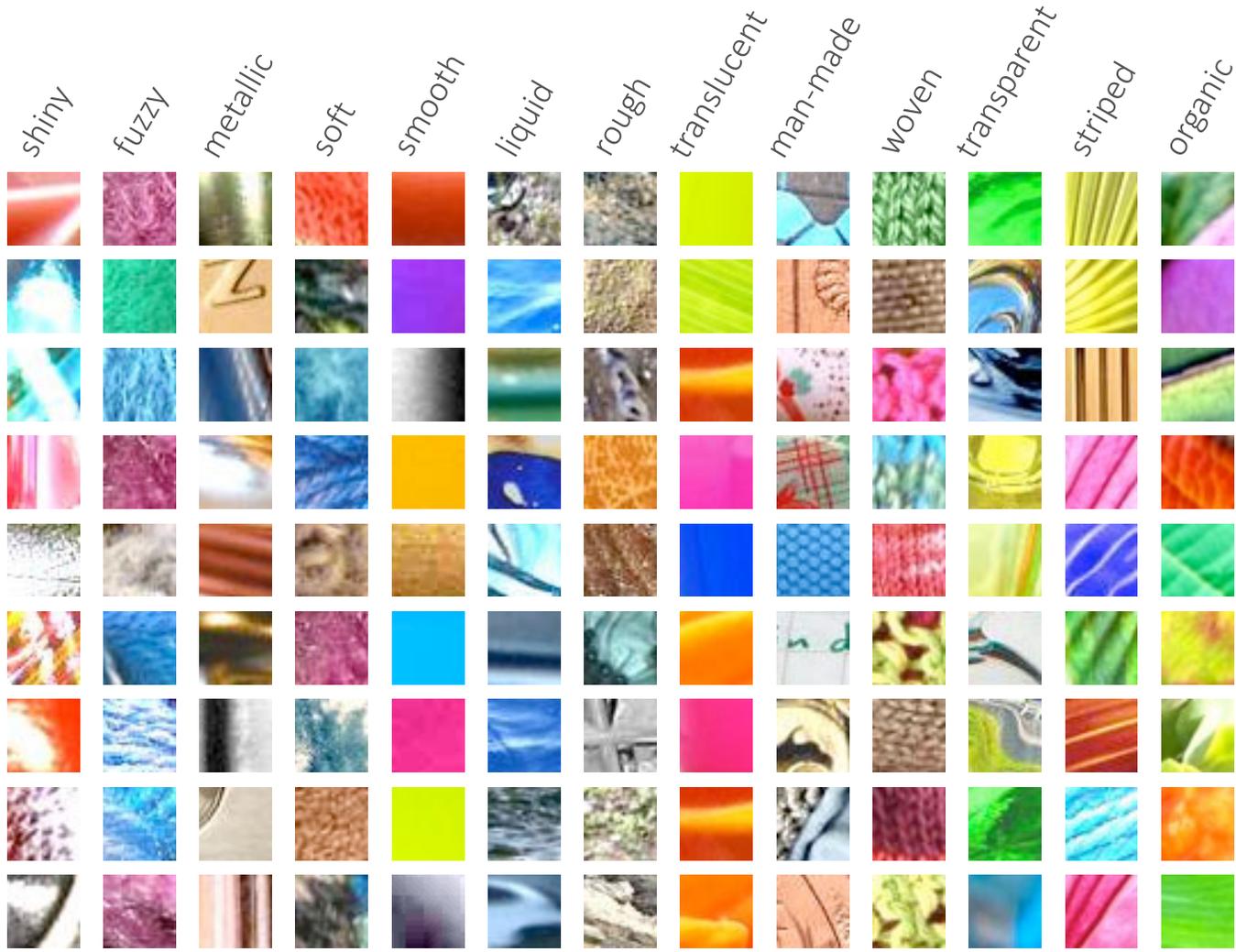


Fuzzy

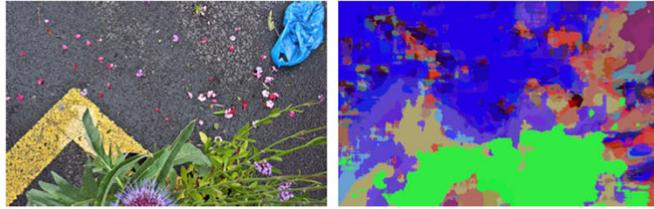


Organic

材質認識



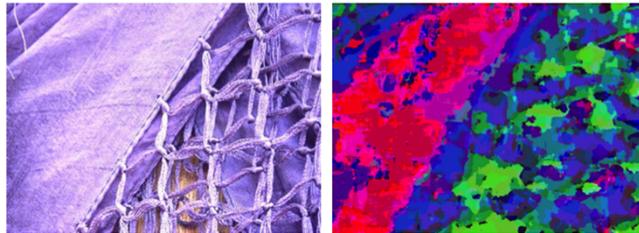
材質認識



■ Manmade ■ Organic ■ Rough



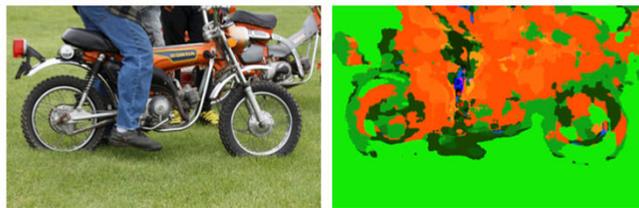
■ Shiny ■ Organic ■ Smooth



■ Smooth ■ Striped ■ Soft



■ Fuzzy ■ Organic ■ Smooth

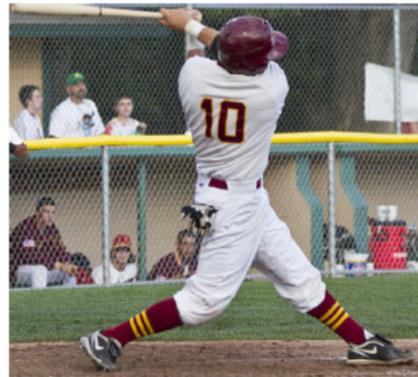
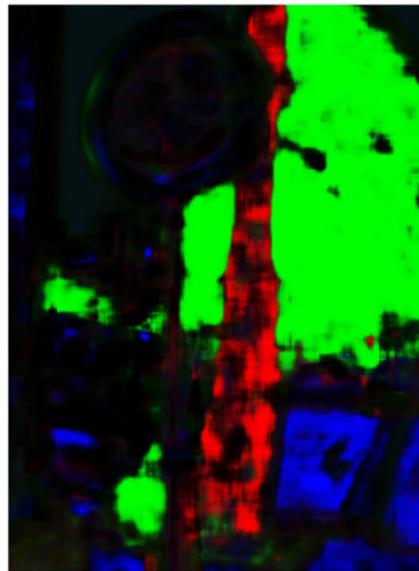


■ Metallic ■ Organic ■ Smooth

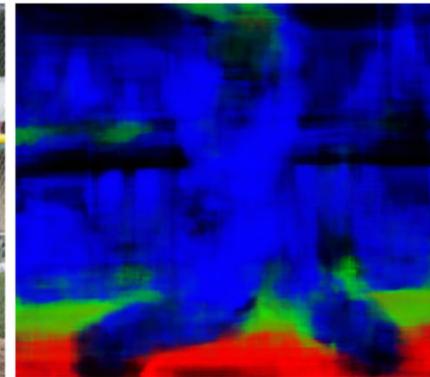
密な素材認識



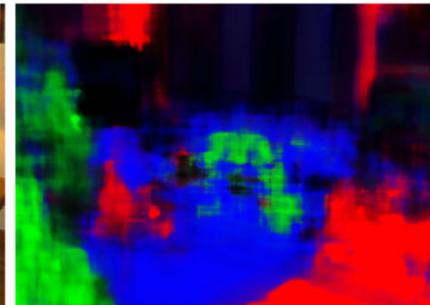
■ Wood ■ Foliage ■ Asphalt



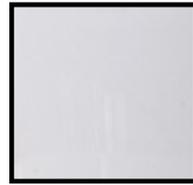
■ Soil ■ Foliage ■ Fabric



■ Wood ■ Foliage ■ Fabric



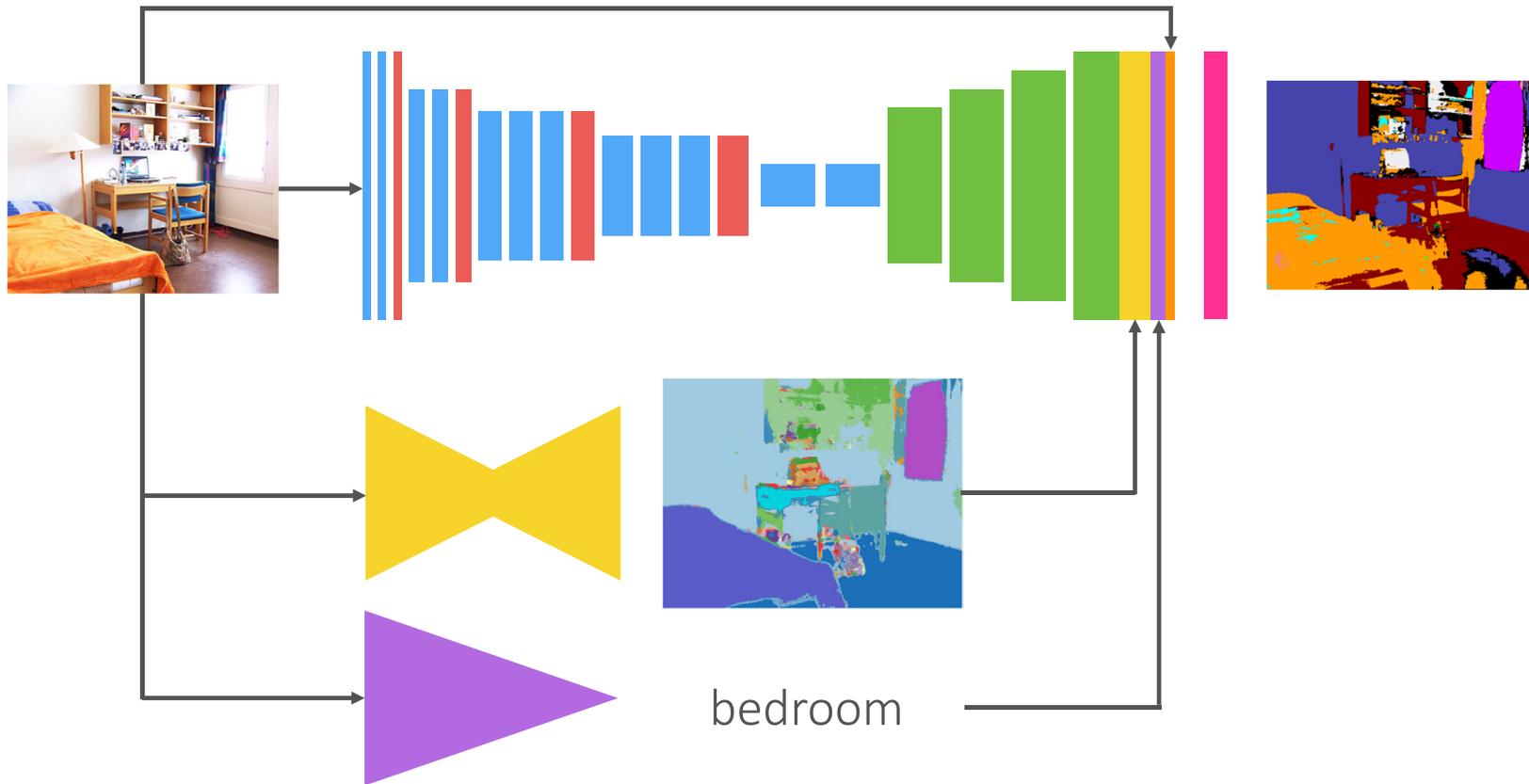
局所視覚情報からの素材認識



局所視覚情報と大局的コンテキスト



コンテキストを用いた素材認識



コンテキストを用いた素材認識



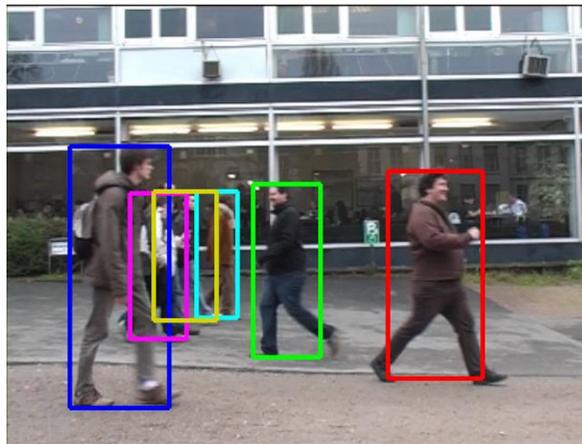
コンテキストを用いた素材認識



- | | | | | | | | |
|---|---|--|--|---|---|---|---|
|  Asphalt |  Ceramic |  Concrete |  Fabric |  Foliage |  Food |  Glass |  Metal |
|  Paper |  Plaster |  Plastic |  Rubber |  Soil |  Stone |  Water |  Wood |

コンピュータビジョンの根本仮定

- まばらで秩序だった世界



Annotated Actions: (gray if not active)

- Holding a dish
- Holding a blanket
- Sitting on sofa/couch
- Snuggling with a blanket
- Sitting in a bed
- Drinking from a cup/glass/bottle
- Holding a cup/glass/bottle of something

Annotated Objects:

- Bed, Blanket, Cup, Dish, Sofa/couch



Script:

A person is drinking coffee while snuggling cozily in a blanket.



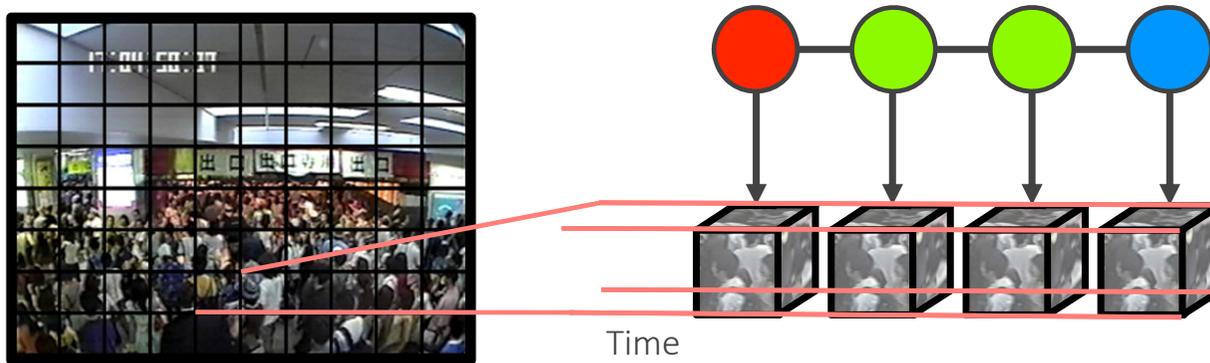
人の流れ



人流のモデル化

・局所的な動き場の時間変遷のモデル化

- 局所的オプティカルフローをvon Mises-Fisher分布で表現
- その平均方向と分散を生成する隠れマルコフモデル



Observation (local motion pattern) $O = \{\mu, \kappa\}$ von Mises-Fisher

Emission density

$$p(O|z) = p(\mu|z)p(\kappa|z)$$

von Mises-Fisher Gamma

異常検出

- 人流モデルからの統計的外れ

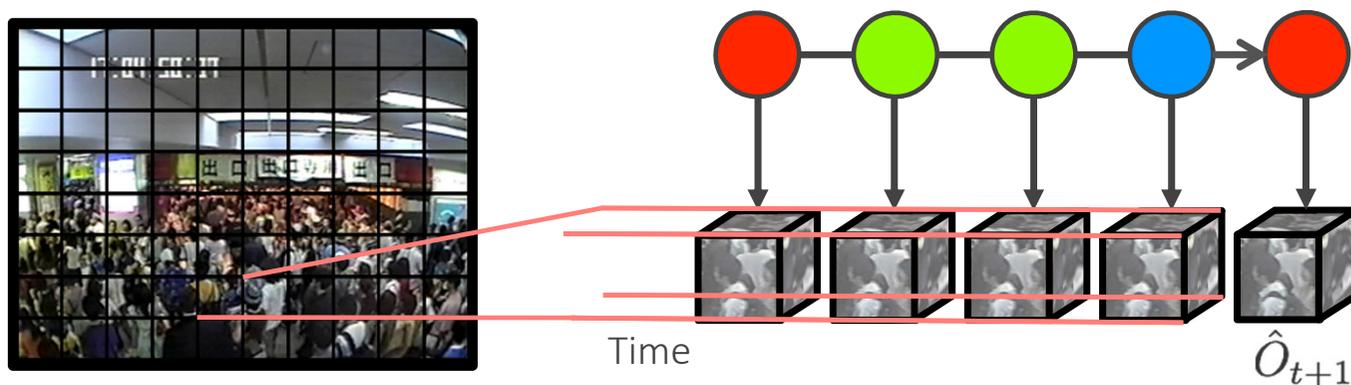


■ true positive ■ false positive ■ false negative



[CVPR09, ECCV12]

局所的動き場の予測



$$\hat{O}_{t+1} = \mathbb{E}[p(O_{t+1}|O_1, \dots, O_t)] = \sum_{z \in Z} \gamma_{t+1}(z) \mathbb{E}[p(O|z)]$$

\swarrow
 $\alpha^T \mathbf{A}$

- 隠れ状態に対する局所動き場の周辺分布
 - 予測された局所動き場は von Mises-Fisher distribution
 - 直接サンプリングできる方向統計分布

人流を用いたベイジアントラッキング

$$\frac{p(\mathbf{x}_t|\mathbf{z}_{1:t})}{\text{posterior}} \propto \frac{p(\mathbf{z}_t|\mathbf{x}_t)}{\text{likelihood}} \int \frac{p(\mathbf{x}_t|\mathbf{x}_{t-1})}{\text{state transition}} \frac{p(\mathbf{x}_{t-1}|\mathbf{z}_{1:t-1})}{\text{prior post.}} d\mathbf{x}_{t-1}$$

人流に基づく動的プライア

(時空間コンテキスト)

Directions

■ Right

■ Left

■ Up



雑踏おけるトラッキング



[CVPR10, ECCV12, PAMI12]

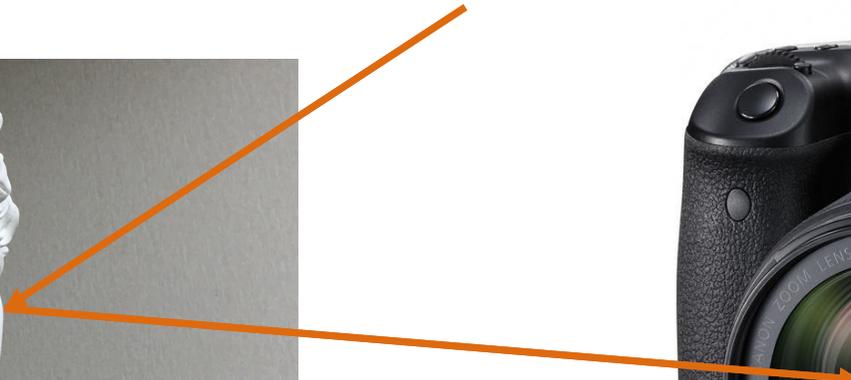
雑踏おけるトラッキング

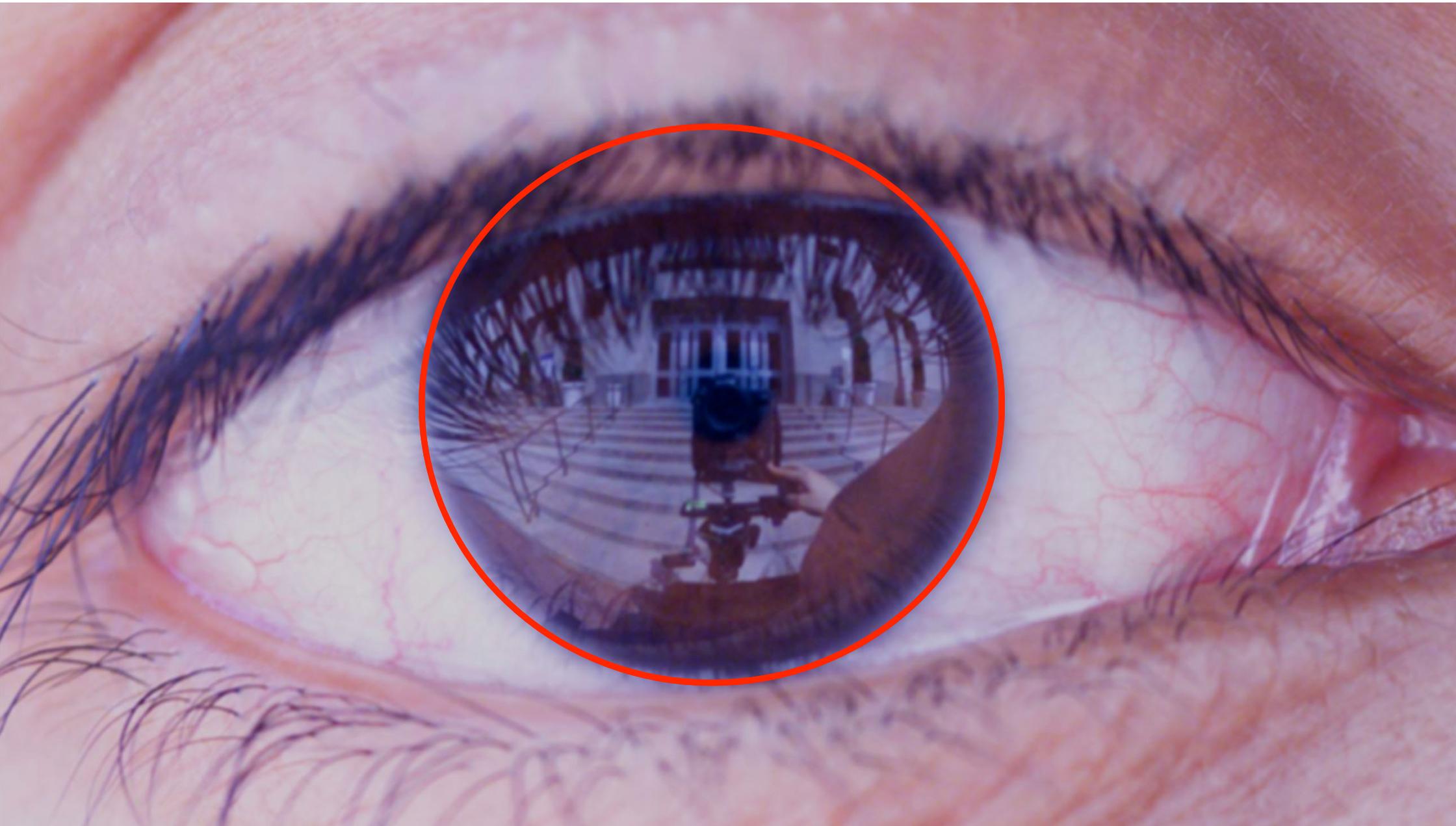


[CVPR10, ECCV12, PAMI12]

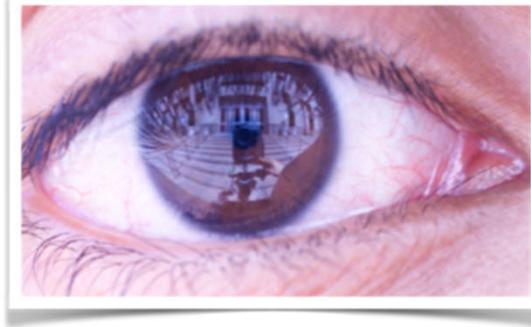
コンピュータビジョンの根本仮定

- 真空中で剛体の世界

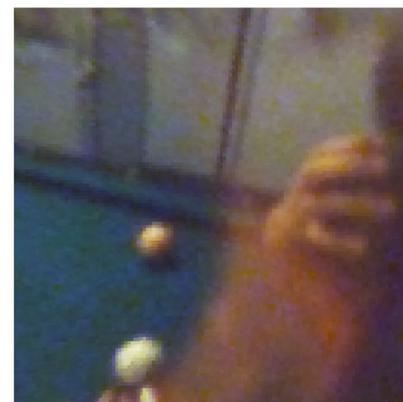
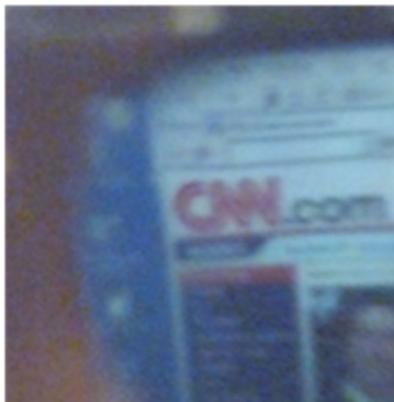
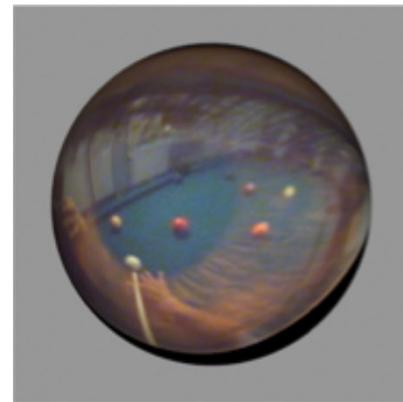
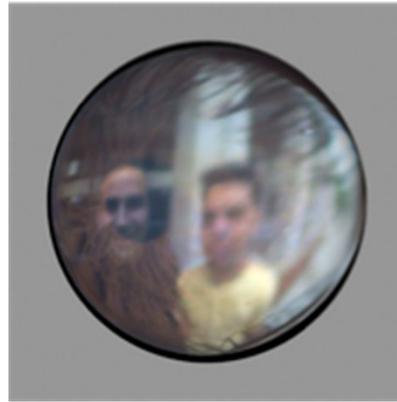
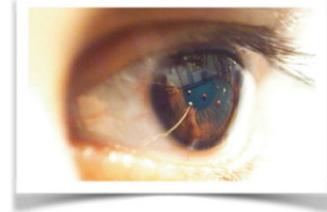
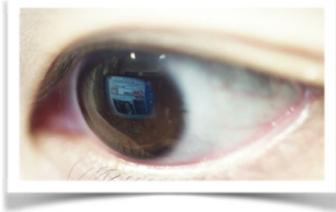




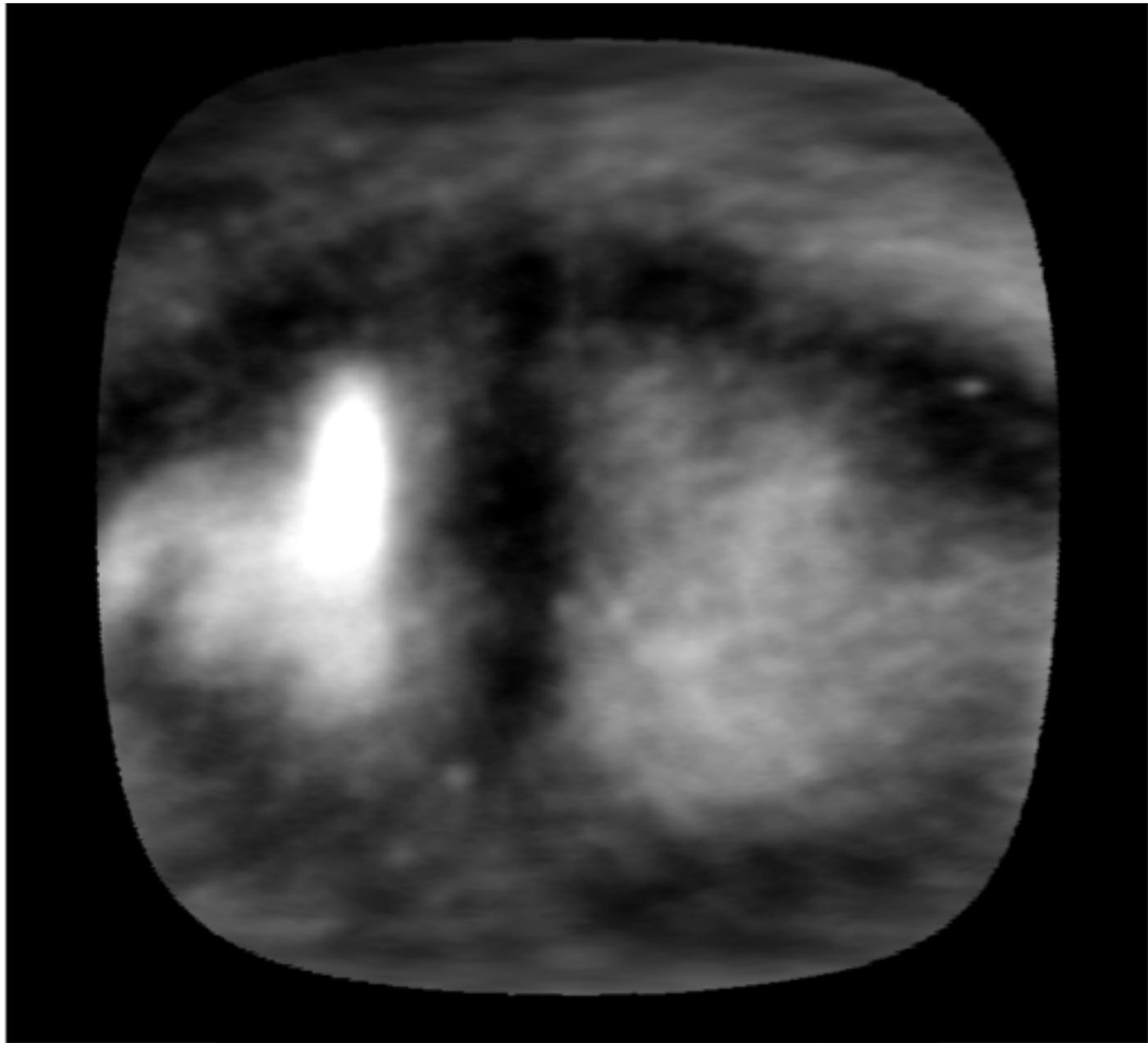
眼の中の世界



[CVPR04, IJCV 06]



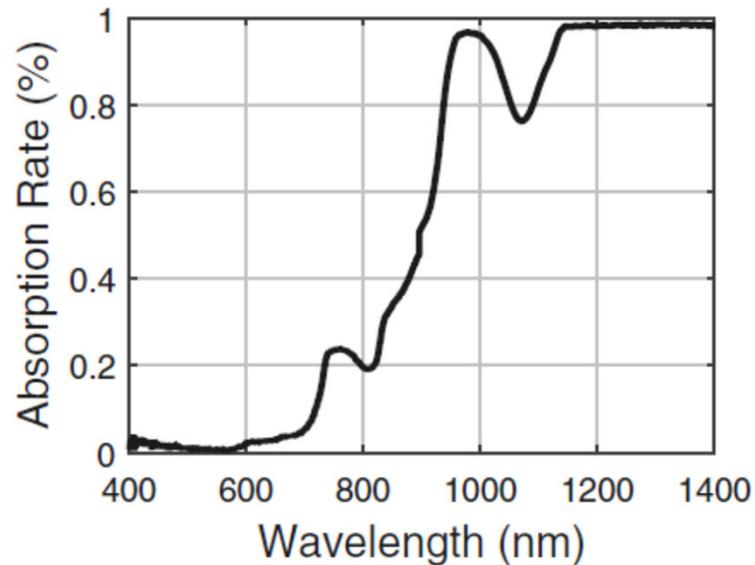
Watching a Bus





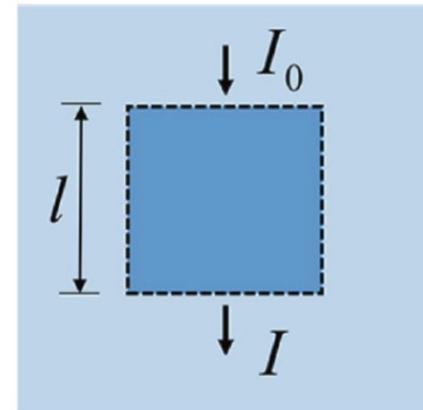
水における光の吸収

近赤外光の水による吸収係数

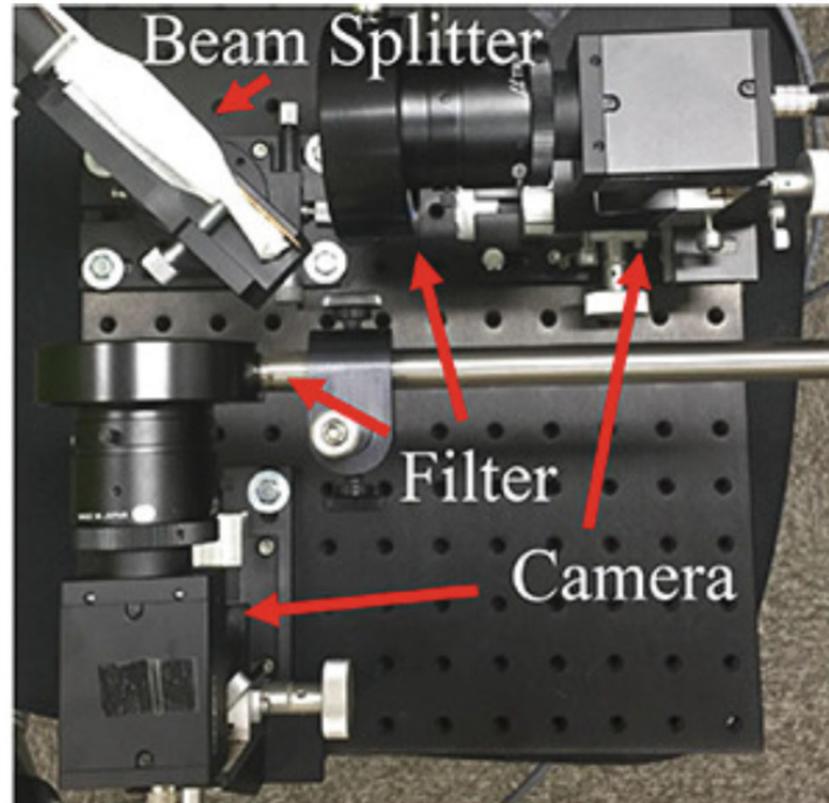


Beer-Lambert Law

$$I = I_0 \exp[-\alpha(\lambda)l]$$



2波長近赤外光イメージング



$$l \approx \frac{1}{2(\alpha(\lambda_2) - \alpha(\lambda_1))} \ln \frac{I(\lambda_1)}{I(\lambda_2)}$$

Shape from Water

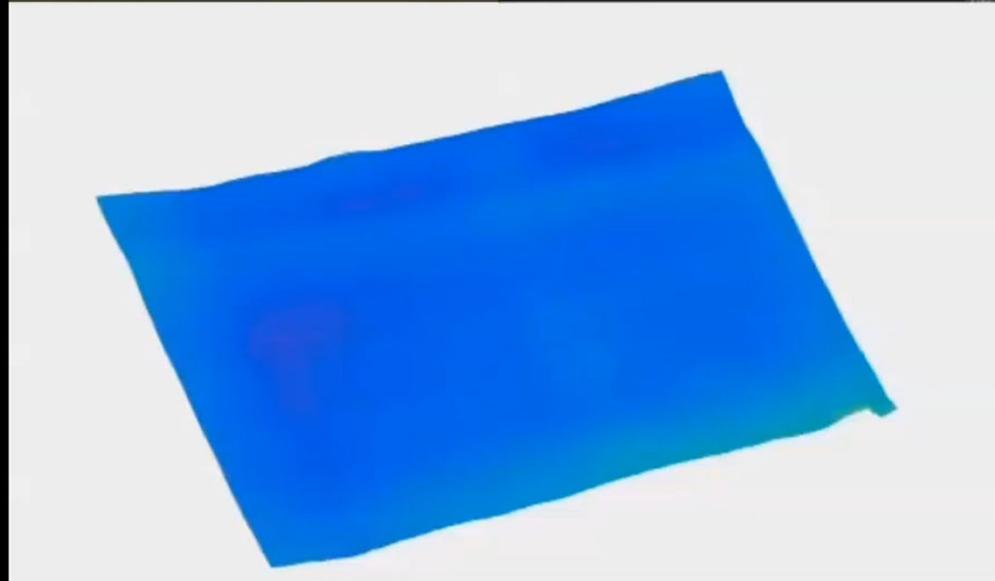
RGB Camera



NIR Camera
(905nm)



3D shape



可変リングライトイメージング

Variable Ring Light:

Capturing Transient Subsurface Scattering with An Ordinary Camera

Ko Nishino^{1,3}, Art Subpa-asa², Yuta Asano², Mihoko Shimano³, Imari Sato³

¹ Kyoto University

² Tokyo Institute of Technology

³ National Institute of Informatics

見るから視るへ

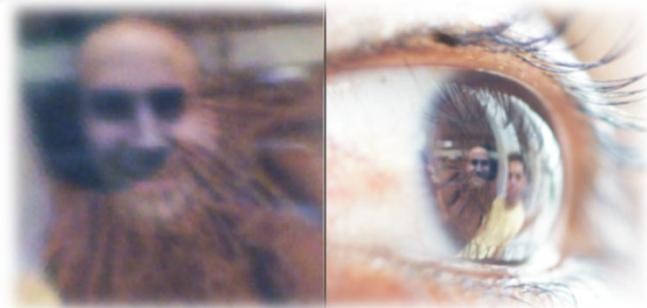


見るから視るへ

人を視る



物を視る



より良く見る

人に寄り添った 社会基盤としてのコンピュータビジョン

