Light Field Vision for Transparent Object Categorization and Segmentation
光场视觉在透明物体分类和分割中的应用

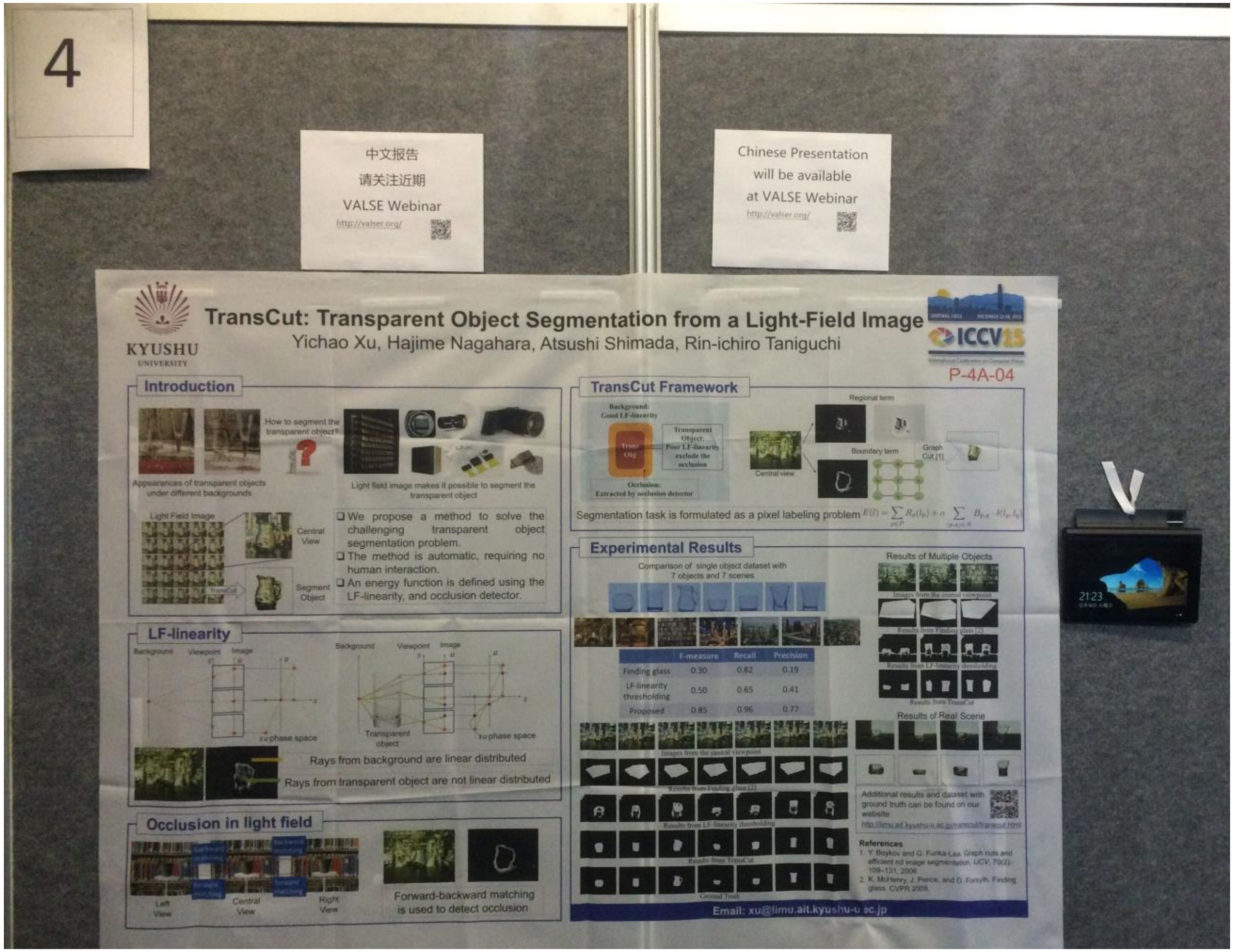
Yichao Xu
徐轶超

xuyichao.cn

KYUSHU UNIVERSITY

Jan. 6, 2016

Just a reminder – Last day P4A-04



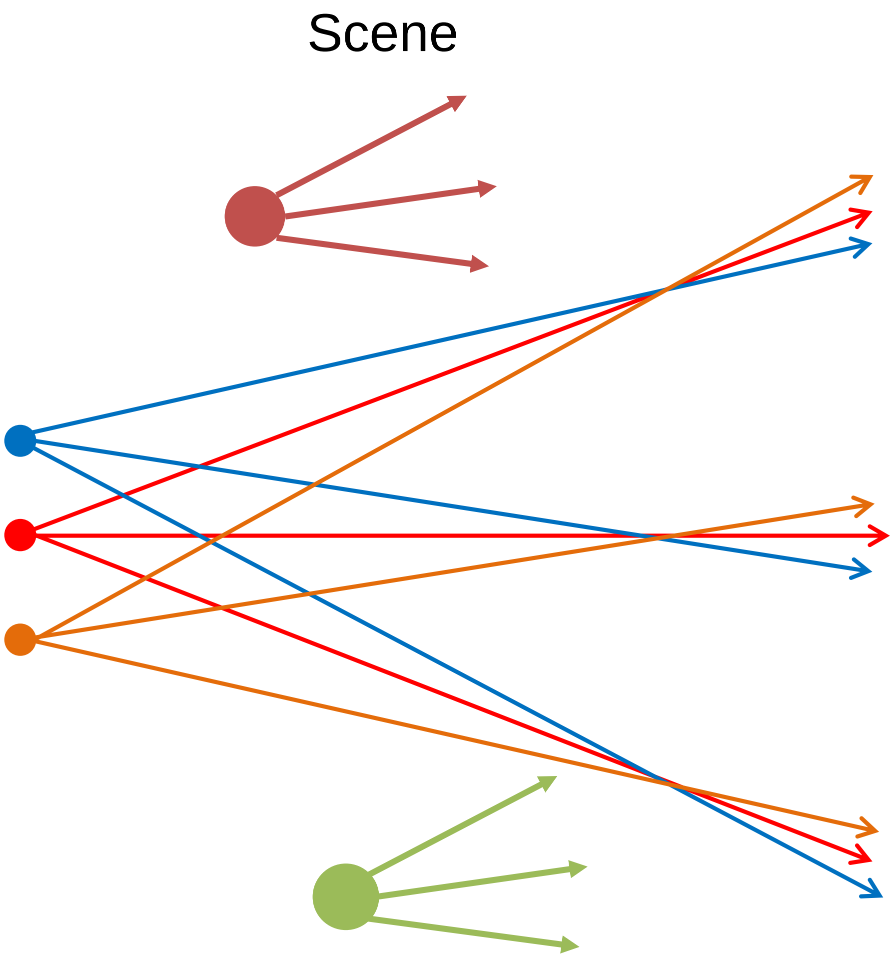
About me

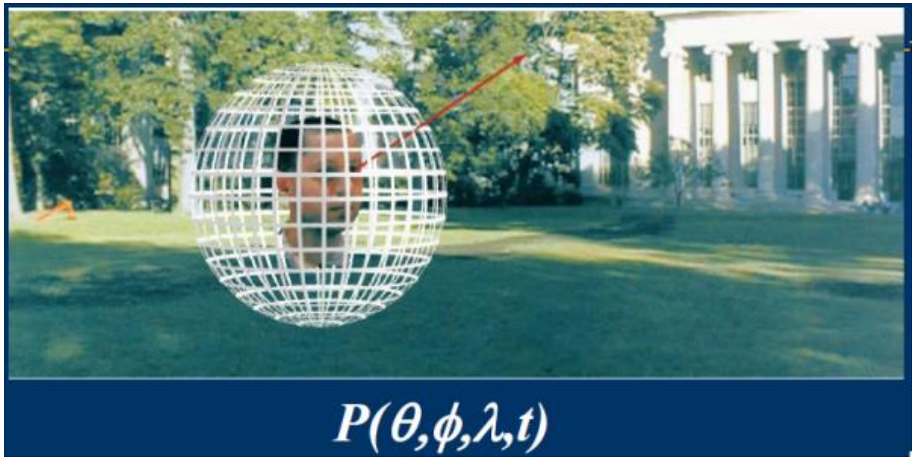
- A: Hometown in Zhejiang
  - Jiaxing
- B: Undergraduate in Beijing
  - BESTI
- C: Master 1 in Anhui
  - USTC, Hefei
- D: Master 2-3 in Shanghai
  - SINAP, CAS
- E: PhD in Fukuoka, Japan
  - Kyushu University

Outline

- Introduction of Light Field Vision
- Transcat: Transparent Object Categorization
- Transcut: Transparent Object Segmentation

Light field



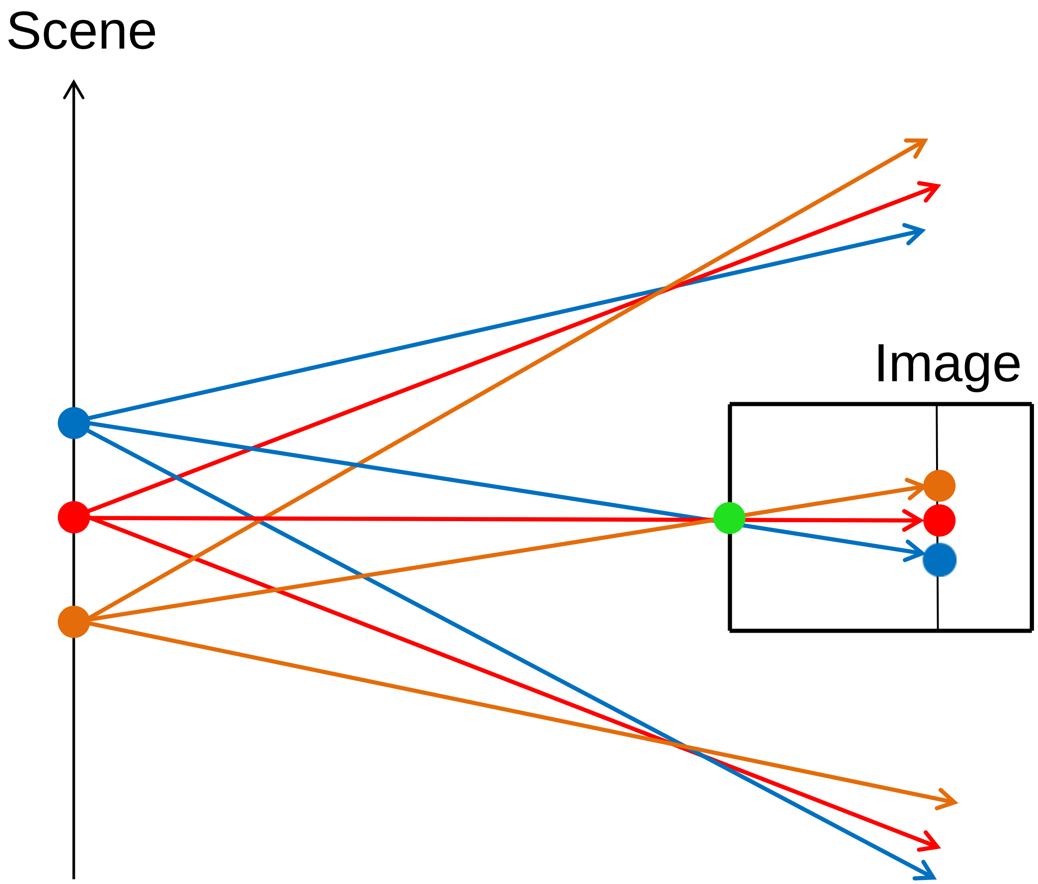


Light field describes all the light rays in the space

Sensors for visual perception

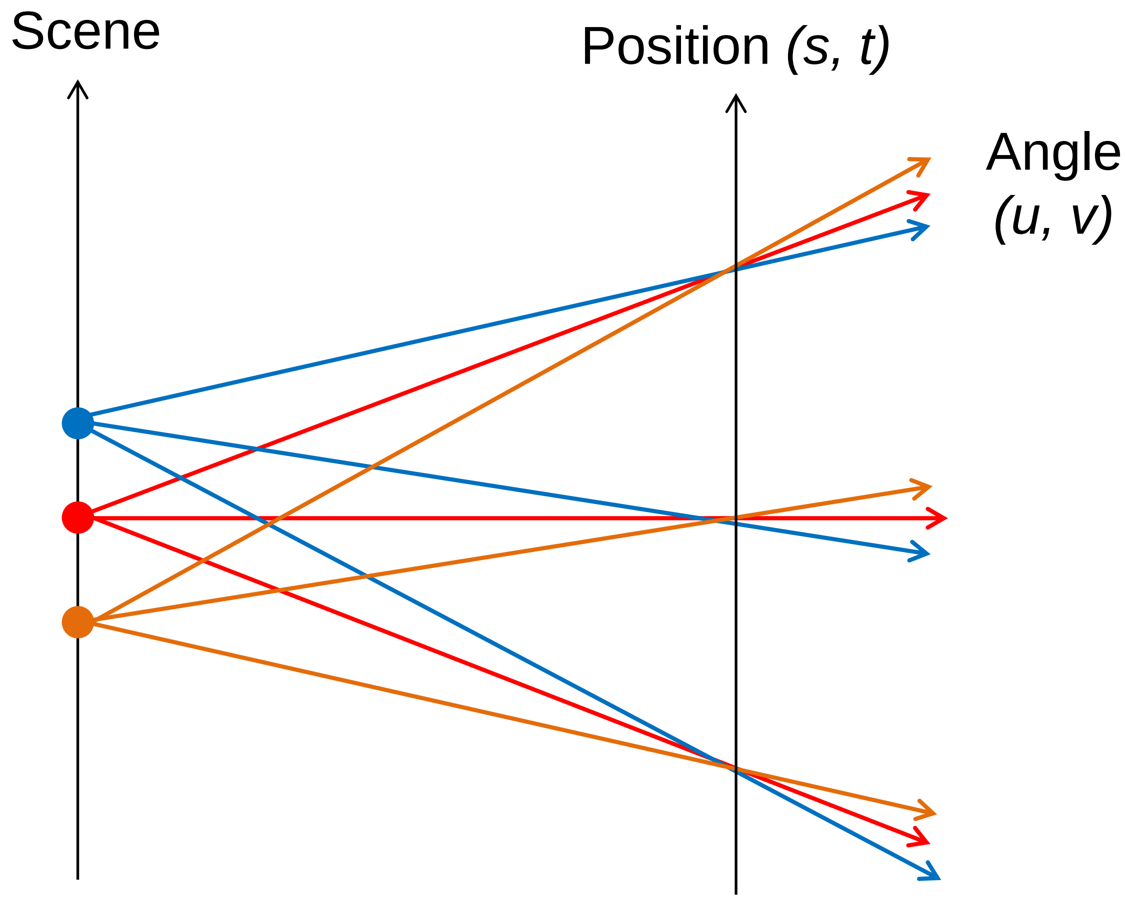
Cameras with CCD and CMOS sensors

Regular camera sensing



Only a few light rays can be captured

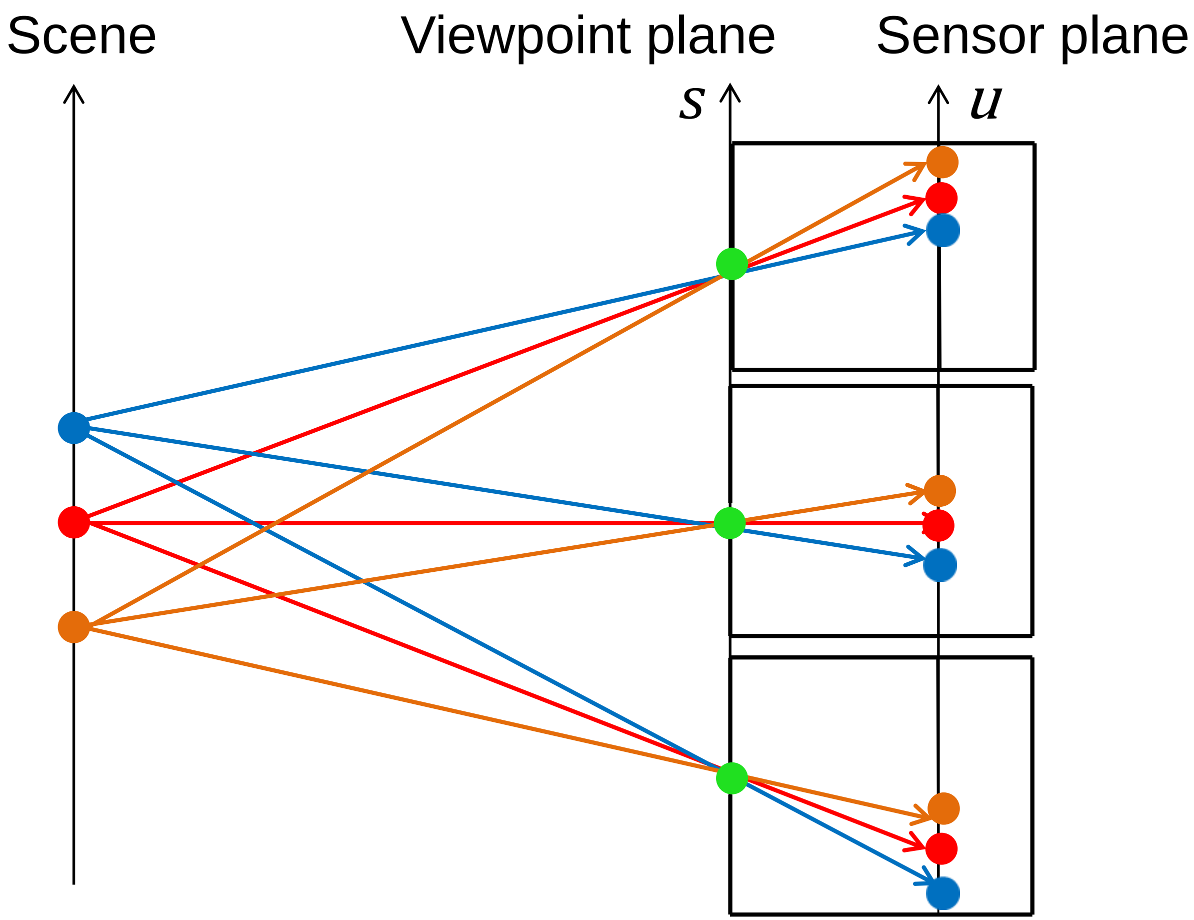
Light field parameterization



4D light field

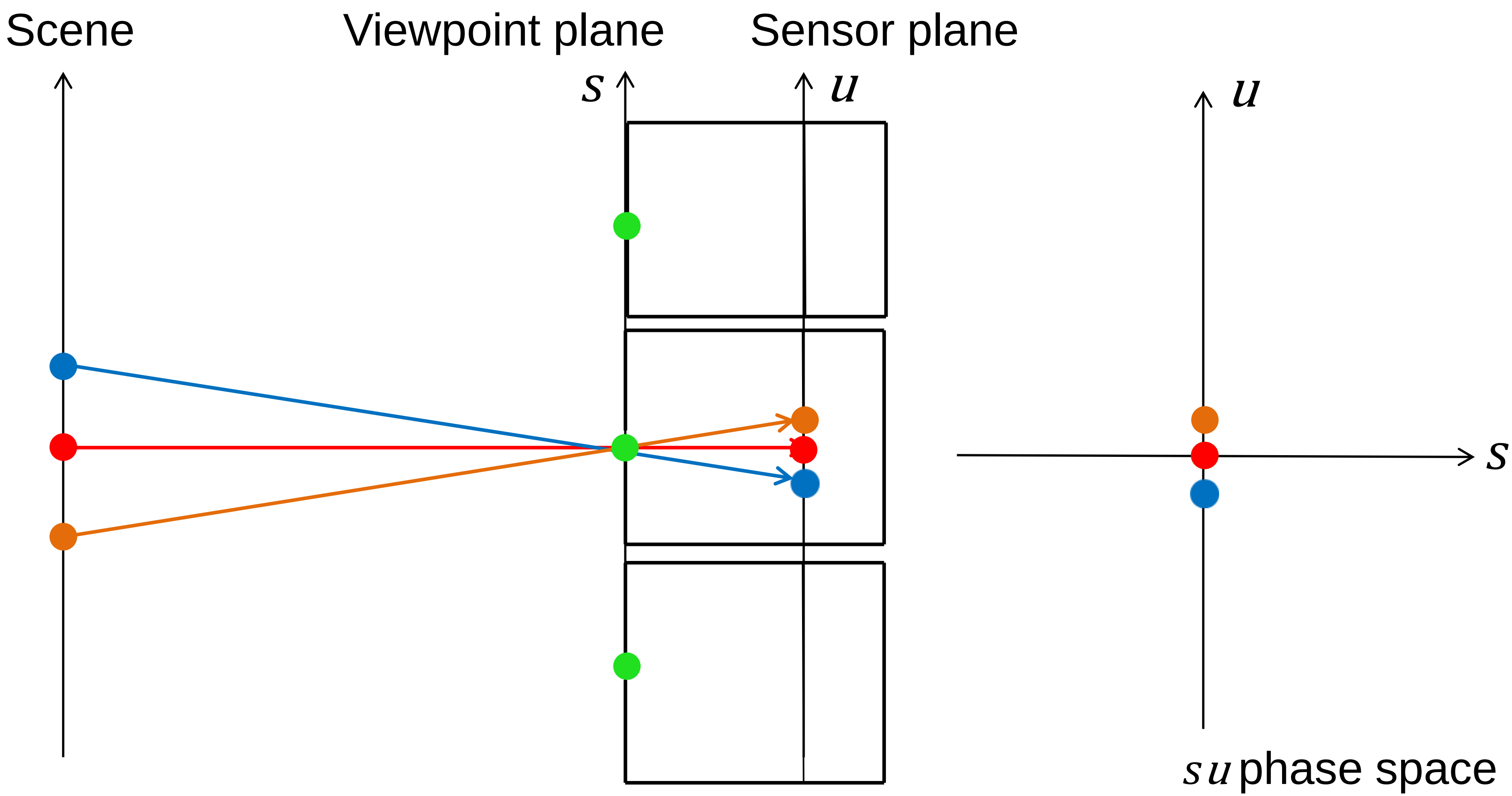
Each light ray can be represented by $L(s, t, u, v)$

Light field sensing



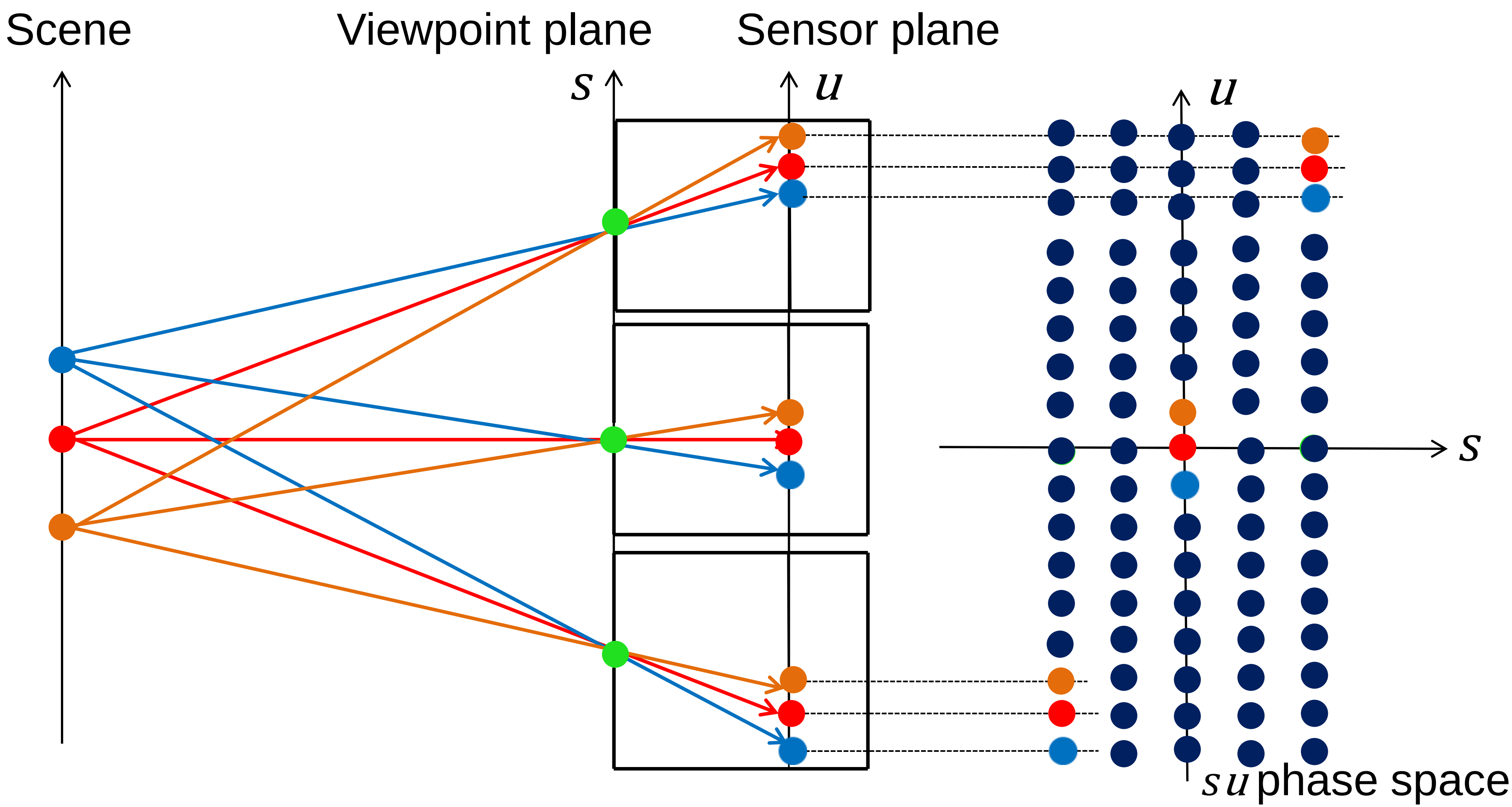
Light field camera can capture richer information

Light field sampling in phase space



Regular camera can only sample sub light field space

Light field sampling in phase space



Light field camera can capture richer information

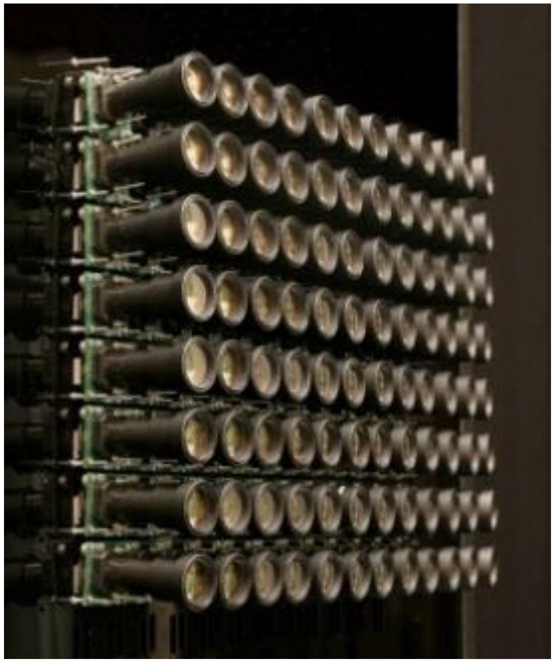
Computational Photography

Multi-focus

Multi-view

Light Field is widely used for Image-based Rendering

Light field cameras

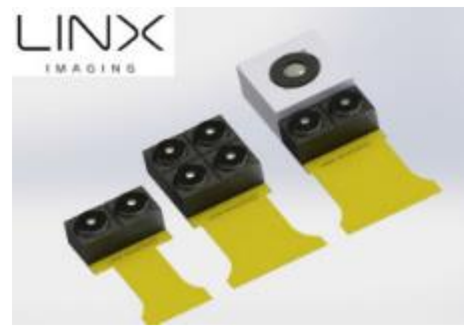


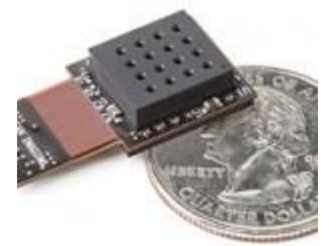
Stanford

Raytrix

Lytro

Viewplus





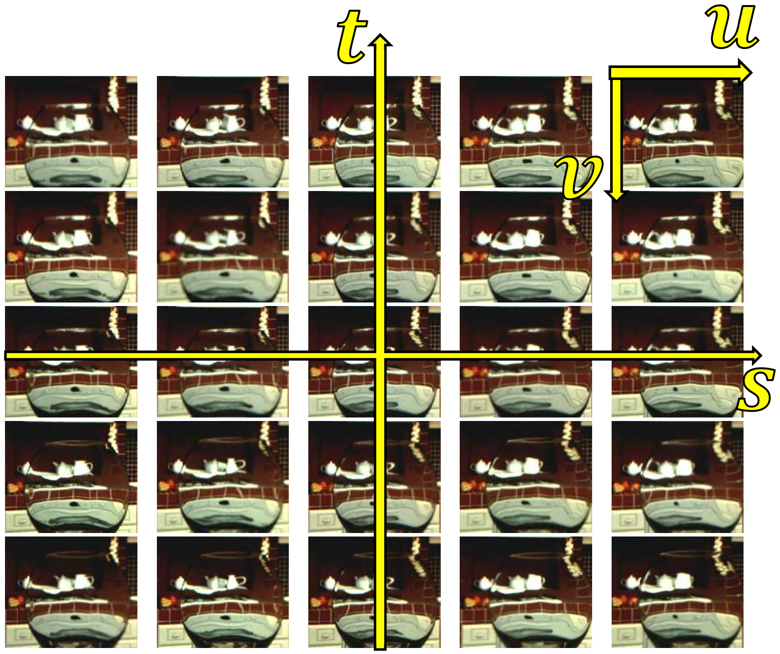
PiCam

Simultaneously record positional and angular information of ray

Obtain rich information with single-shot

Light field vision

Capture



To solve computer vision problems

Computer vision makes our life better

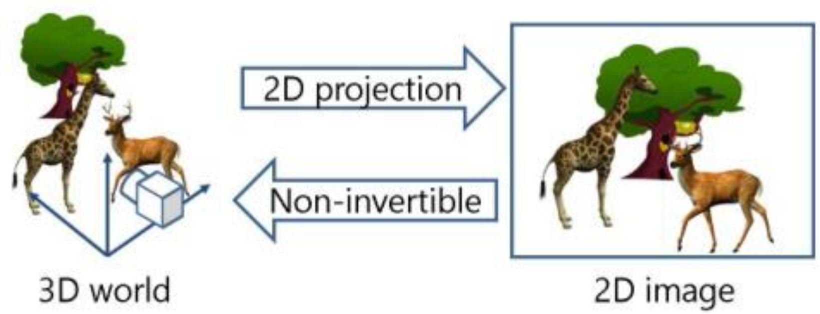
Help us know more

Free our hands

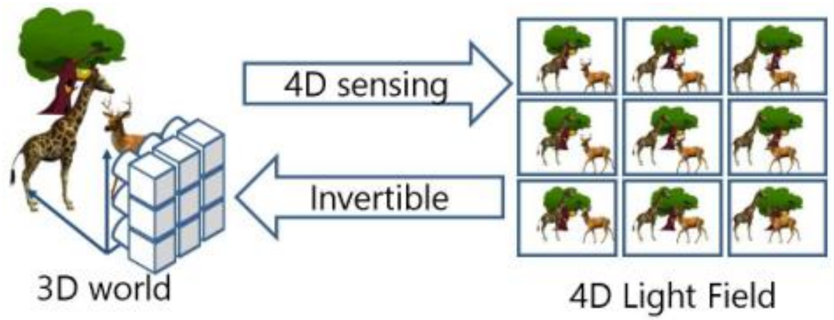
Visual recognition makes it possible

Visual recognition is important in these applications

Advantage of light field vision



Regular Computer vision



Light Field Vision

Redundant information makes it easier to understand the 3D world

Light field vision applications

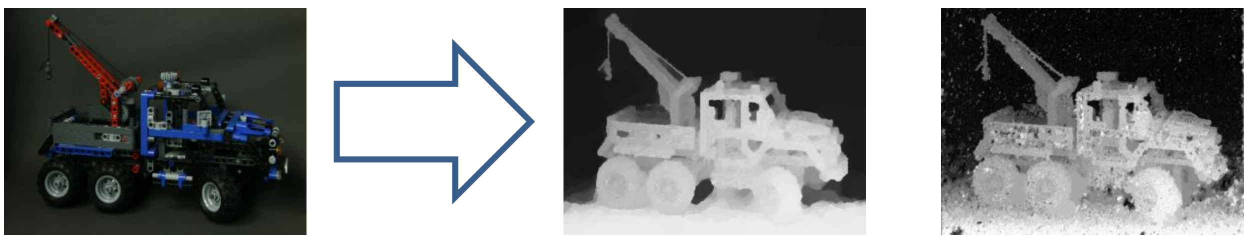
- Surveillance - Accurately detect desired foreground

LF method Conventional

[A.Shimada et al., IPSJ CVA 2013]

- Depth estimation - Accurate and consistent

LF method Conventional



[S. Wanner et al., PAMI2014]

- Salience detection - Accurate in challenge scenes

LF method Conventional GT

[N. Li et al., CVPR2014]

Light Field Vision Application
-- transparent object recognition

Transparent object recognition

Which type is the object?
Where is the object?

Challenge of the target object

Appearance of transparent objects drastically varies with background

Transparent object causes distortion

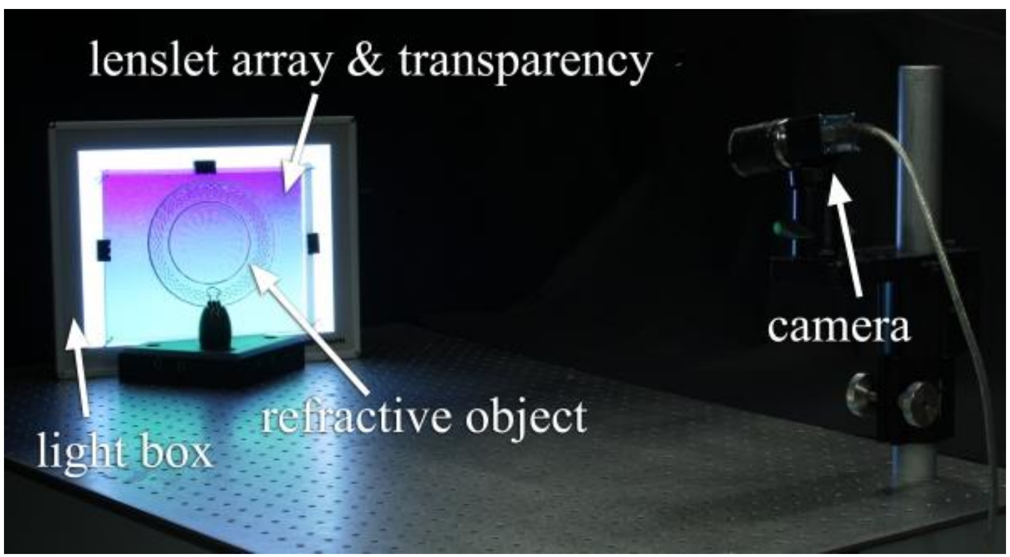
Different objects produce different image of the same scene

Regular computer vision methods cannot understand whether the image is distorted or not without prior knowledge

Know light field from background

Transparent object

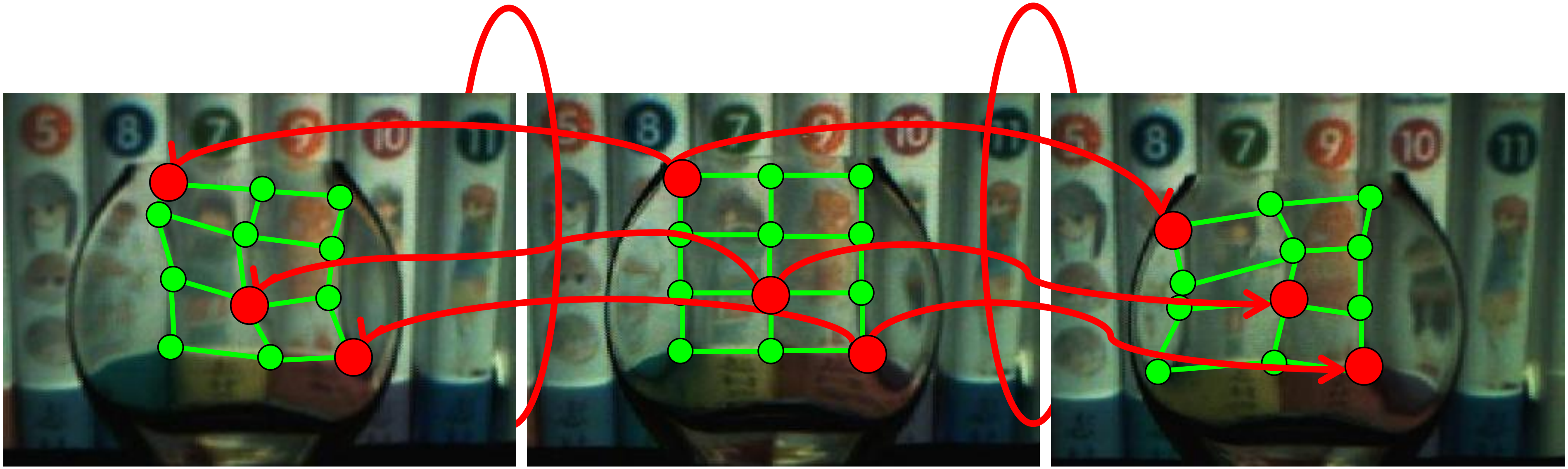
[Ben-Ezra and Nayar, ICCV2003]
Known motion,
Manually tagged feature



[G. Wetzstein et al, ICCV2011]
Known background

Features from Light Field for Transparent Object Recognition

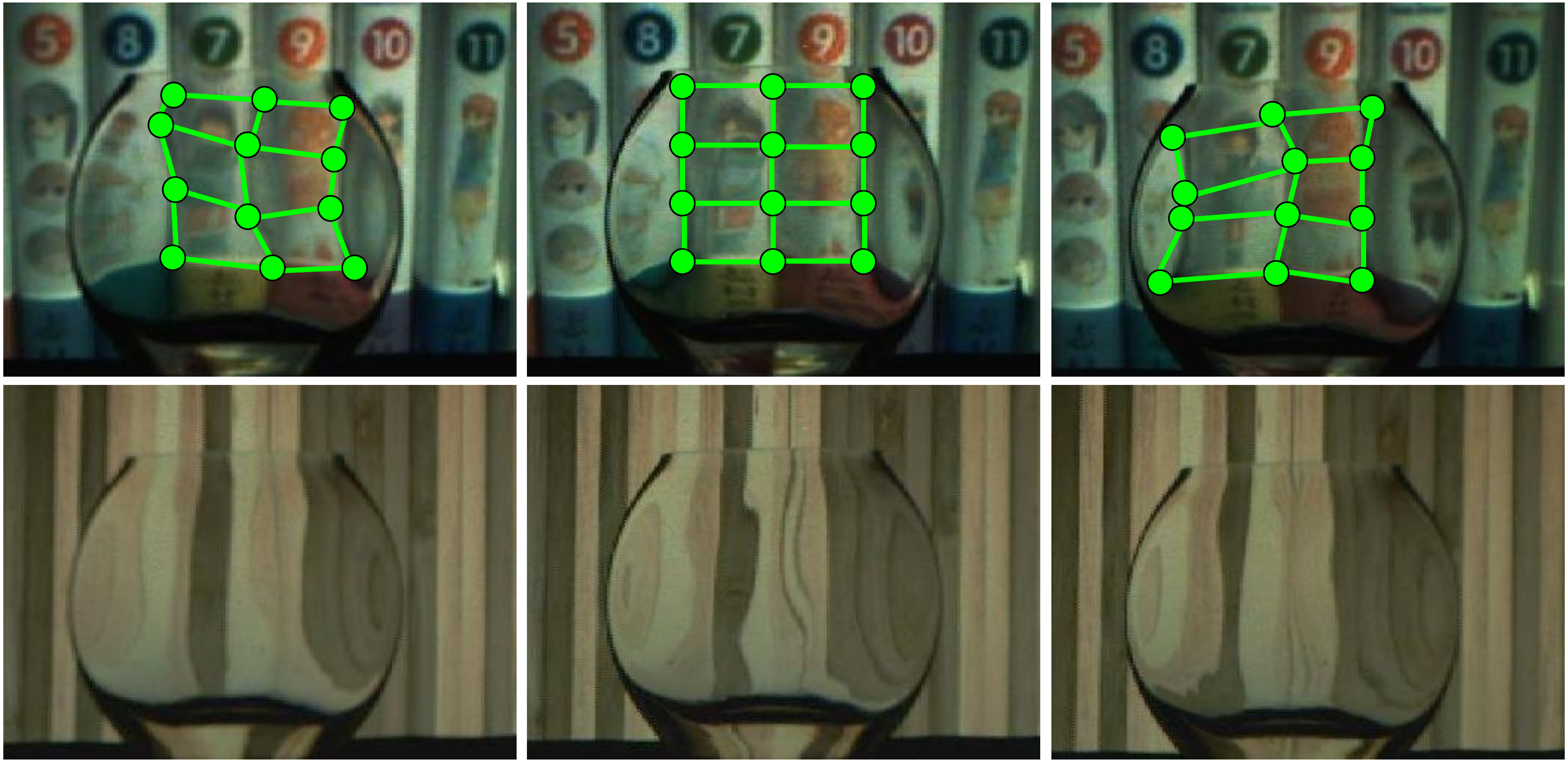
Distortion modeled by light field vision



Background distortion changes with viewpoint

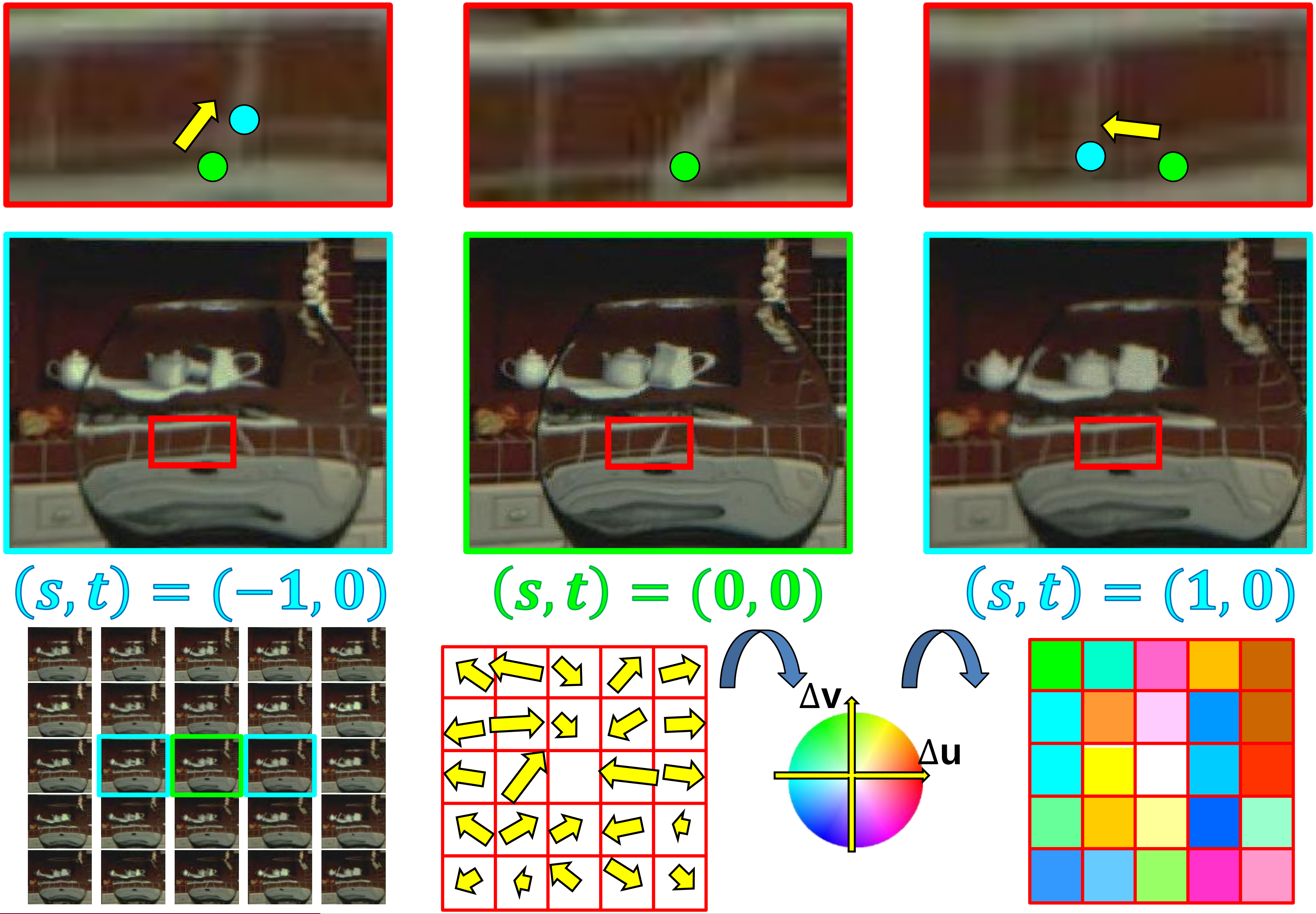
Background distortion is modeled as the correspondences between the viewpoints

Background invariant distortion

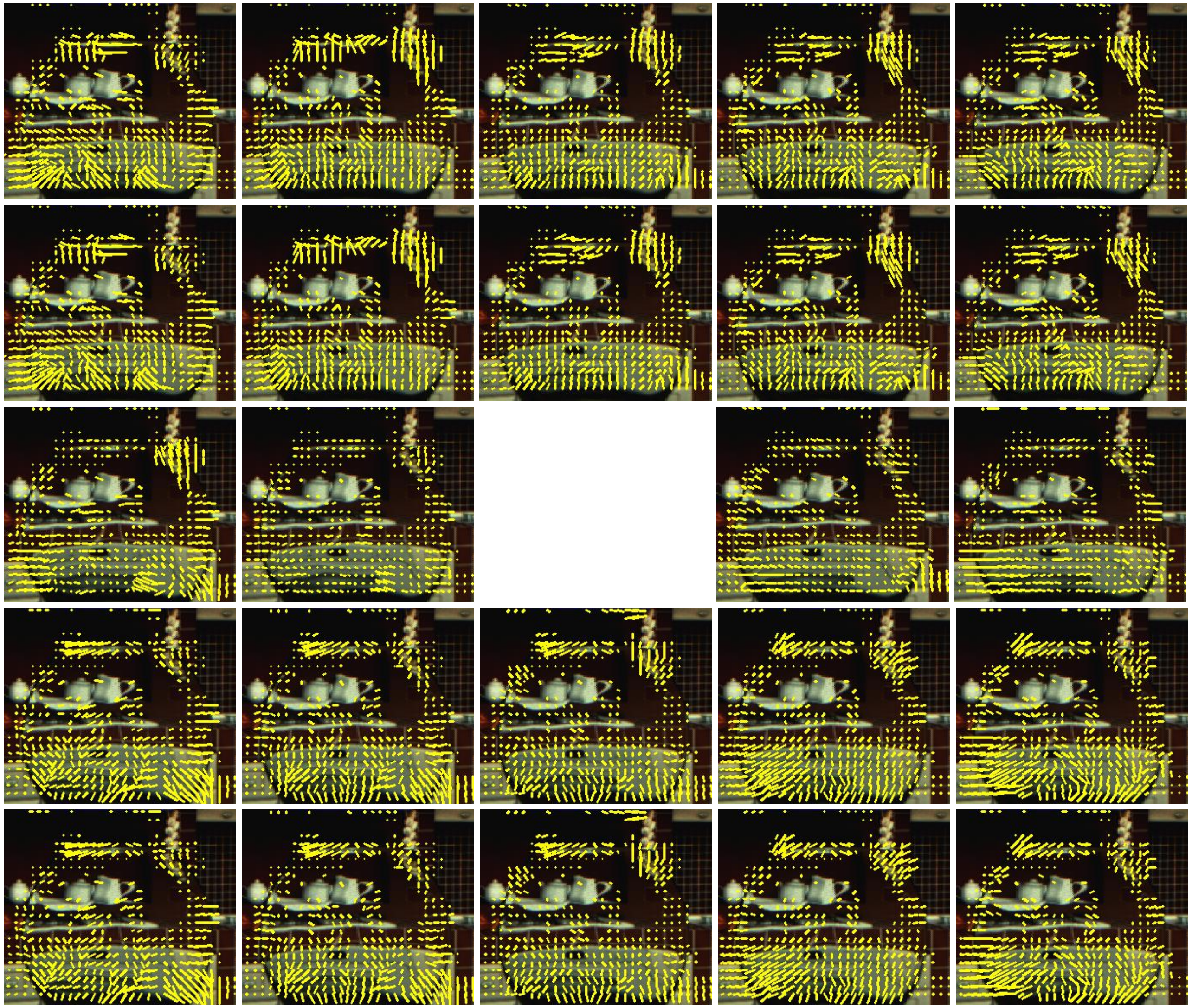


Modeled distortion is independent of background textures

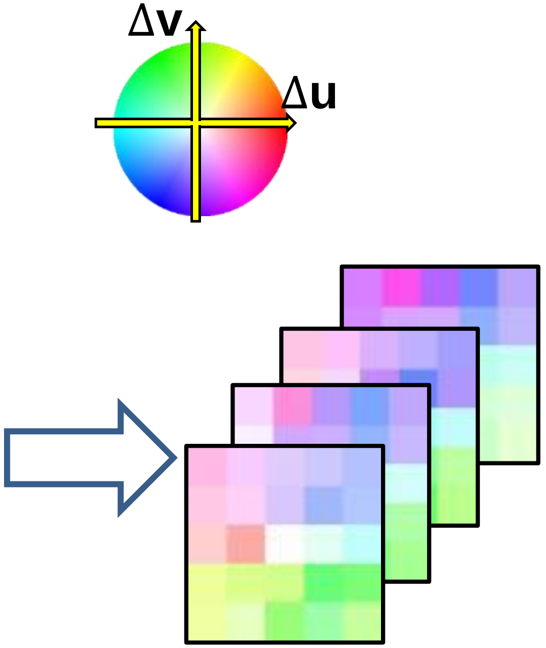
Light Field Distortion (LFD) feature



LFD feature visualization

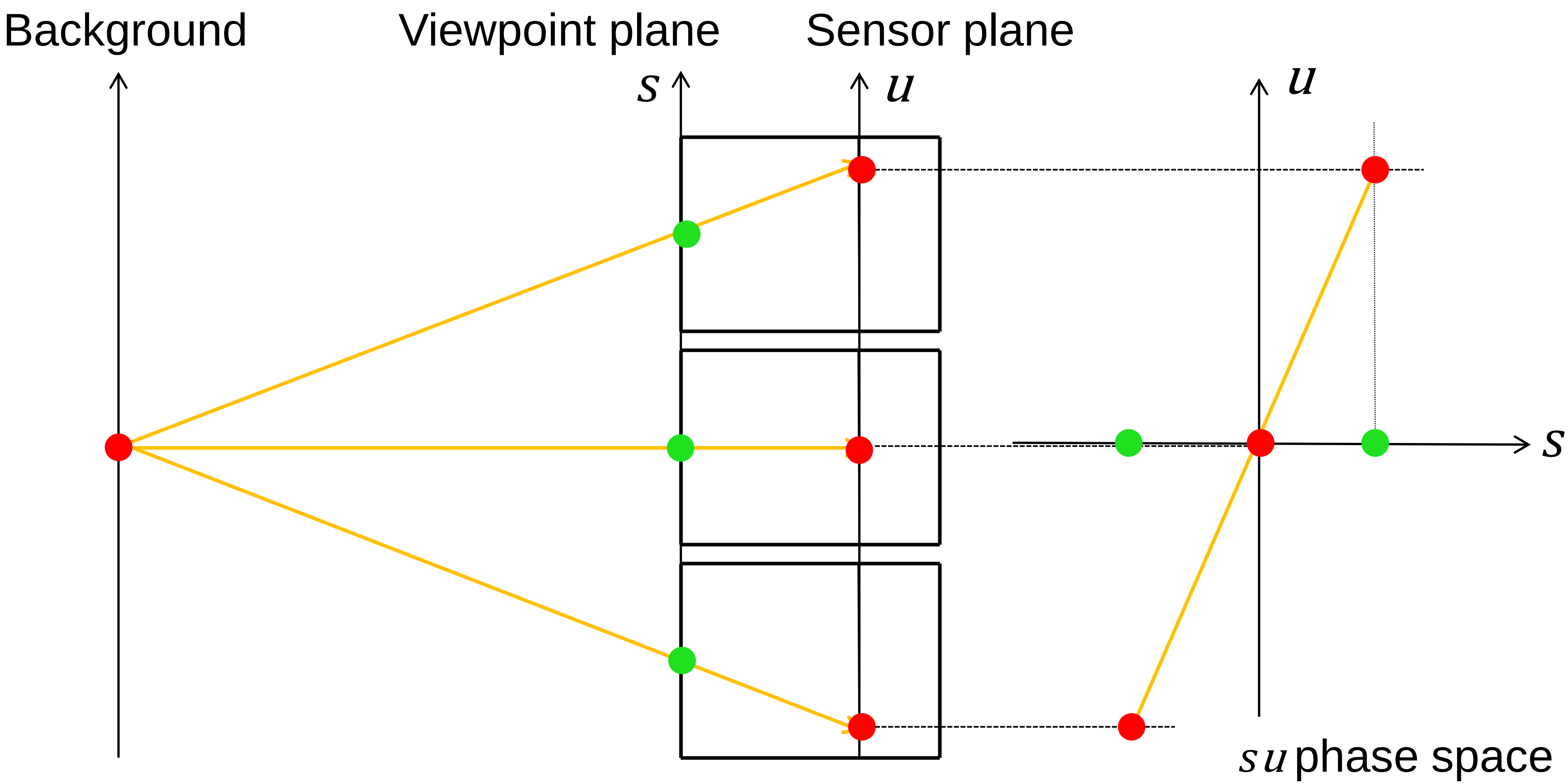


2D vectors on different viewpoints



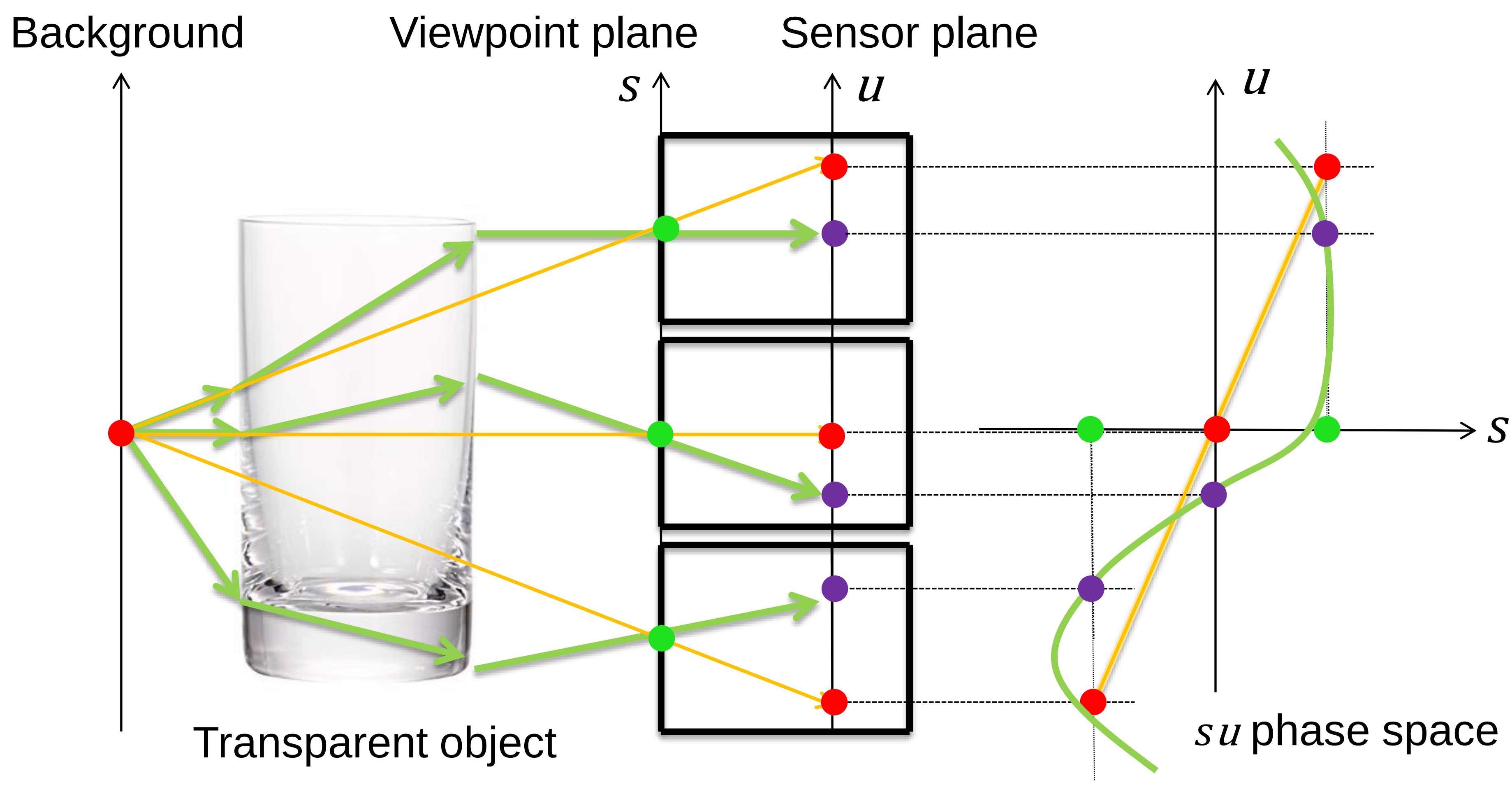
24x2D feature vector for each pixel

Light Field Linearity (LF-linearity)



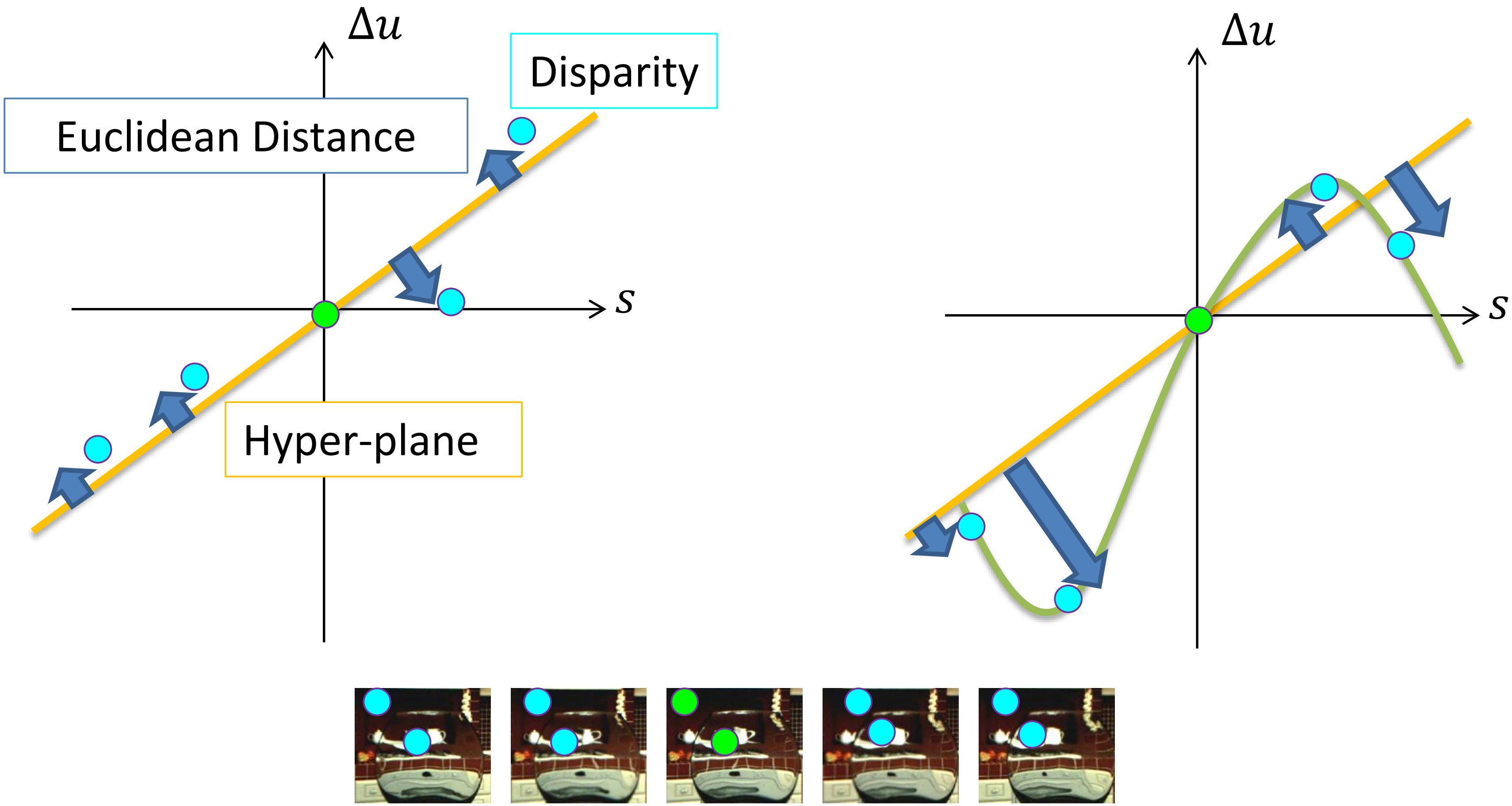
Rays from background are linear distributed

Light Field Linearity (LF-linearity)



Rays from transparent object are not linear distributed

Extract LF-linearity

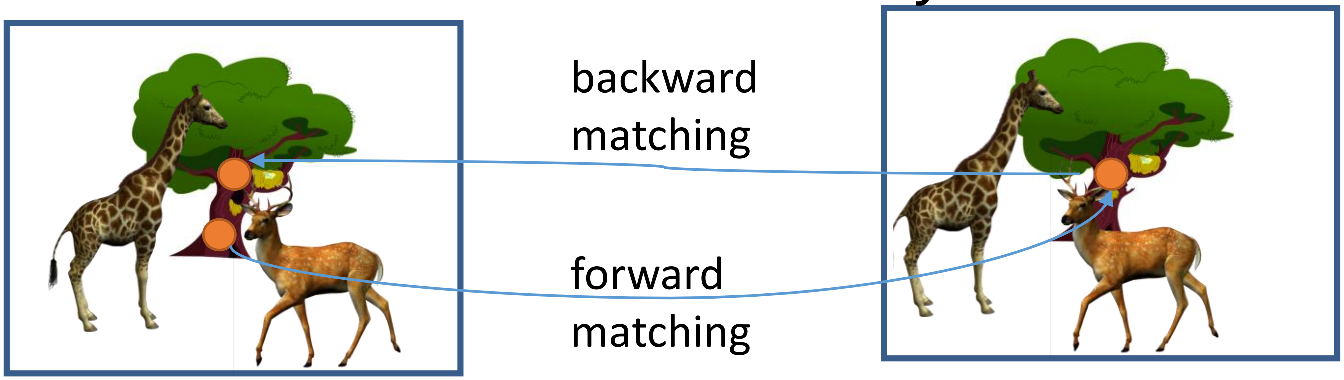


LF-linearity visualization

Central view | LF-linearity

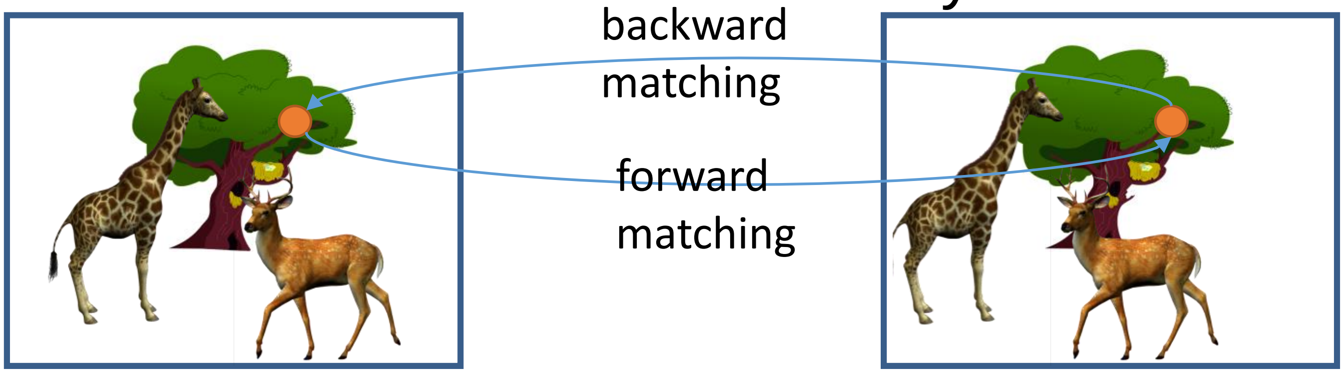
Light Field Consistency (LF-consistency)

Poor consistency



$view(0,0)$ $view(s,t)$

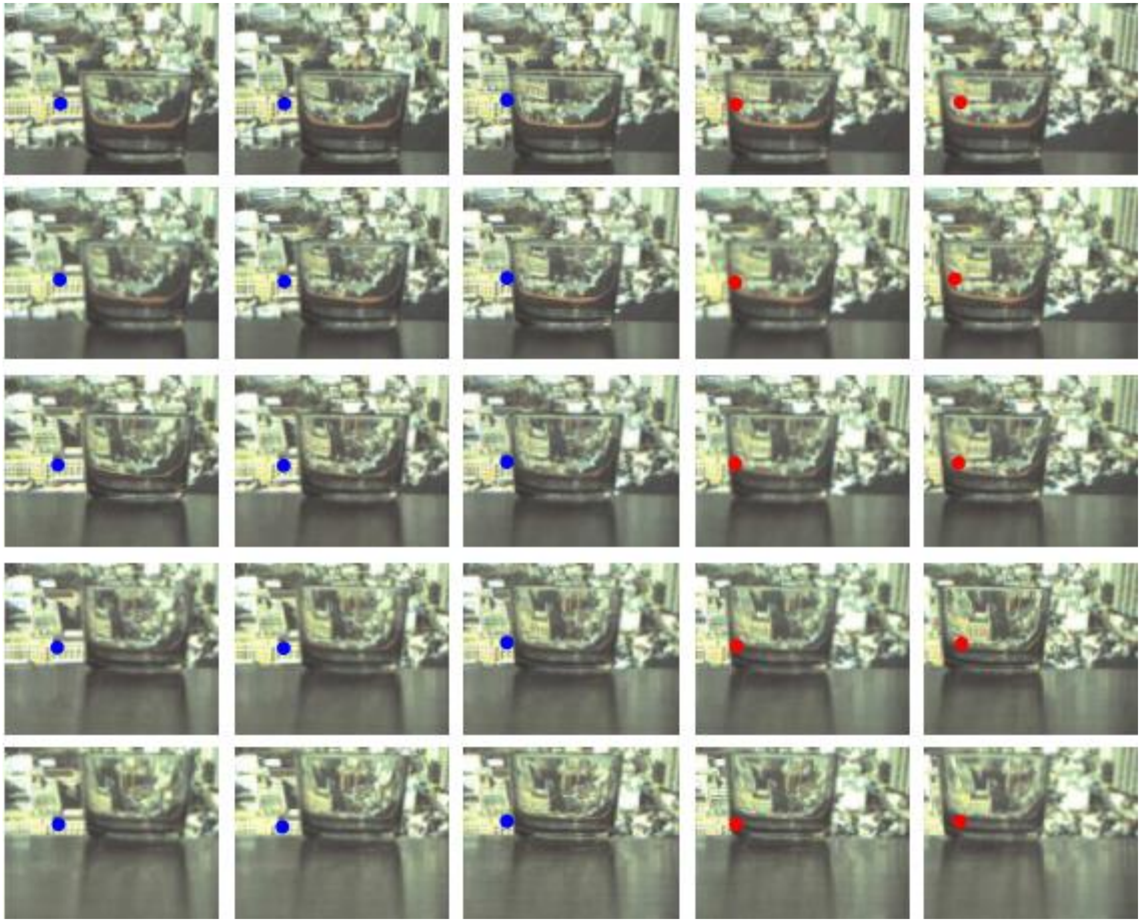
Good consistency



$view(0,0)$ $view(s,t)$

LF-consistency is used for detecting the depth discontinuity

Occlusion in light field



Occlusion detector

0	0	0	0.1	0.1
0	0	0	0.1	0.1
0	0	0	0.1	0.1
0	0	0	0.1	0.1
0	0	0	0.1	0.1

Occlusion is caused by depth discontinuity

Occlusion detectors

0	0	0	0.1	0.1
0	0	0	0.1	0.1
0	0	0	0.1	0.1
0	0	0	0.1	0.1
0	0	0	0.1	0.1

(a) $\theta = 0$

0	0.1	0.1	0.1	0.1
0	0	0.1	0.1	0.1
0	0	0	0.1	0.1
0	0	0	0	0.1
0	0	0	0	0.1

(b) $\theta = 45$

0.1	0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1	0.1
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0

(c) $\theta = 90$

0.1	0.1	0.1	0.1	0
0.1	0.1	0.1	0	0
0.1	0.1	0	0	0
0.1	0	0	0	0
0	0	0	0	0

(d) $\theta = 135$

0.1	0.1	0	0	0
0.1	0.1	0	0	0
0.1	0.1	0	0	0
0.1	0.1	0	0	0
0.1	0.1	0	0	0

(e) $\theta = 180$

0	0	0	0	0
0.1	0	0	0	0
0.1	0.1	0	0	0
0.1	0.1	0.1	0	0
0.1	0.1	0.1	0.1	0

(f) $\theta = 225$

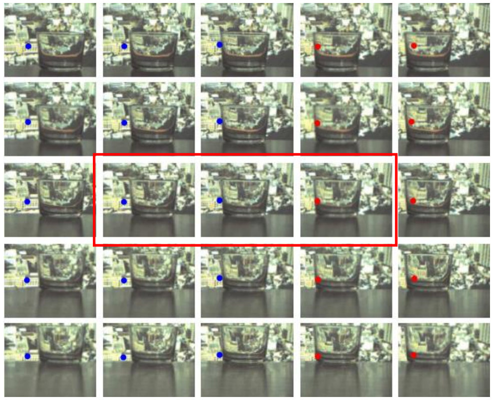
0	0	0	0	0
0	0	0	0	0
0	0	0	0	0
0.1	0.1	0.1	0.1	0.1
0.1	0.1	0.1	0.1	0.1

(g) $\theta = 270$

0	0	0	0	0
0	0	0	0	0.1
0	0	0	0.1	0.1
0	0	0.1	0.1	0.1
0	0.1	0.1	0.1	0.1

(h) $\theta = 315$

Detect occlusion point



0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1
0	0	0	1	1

X

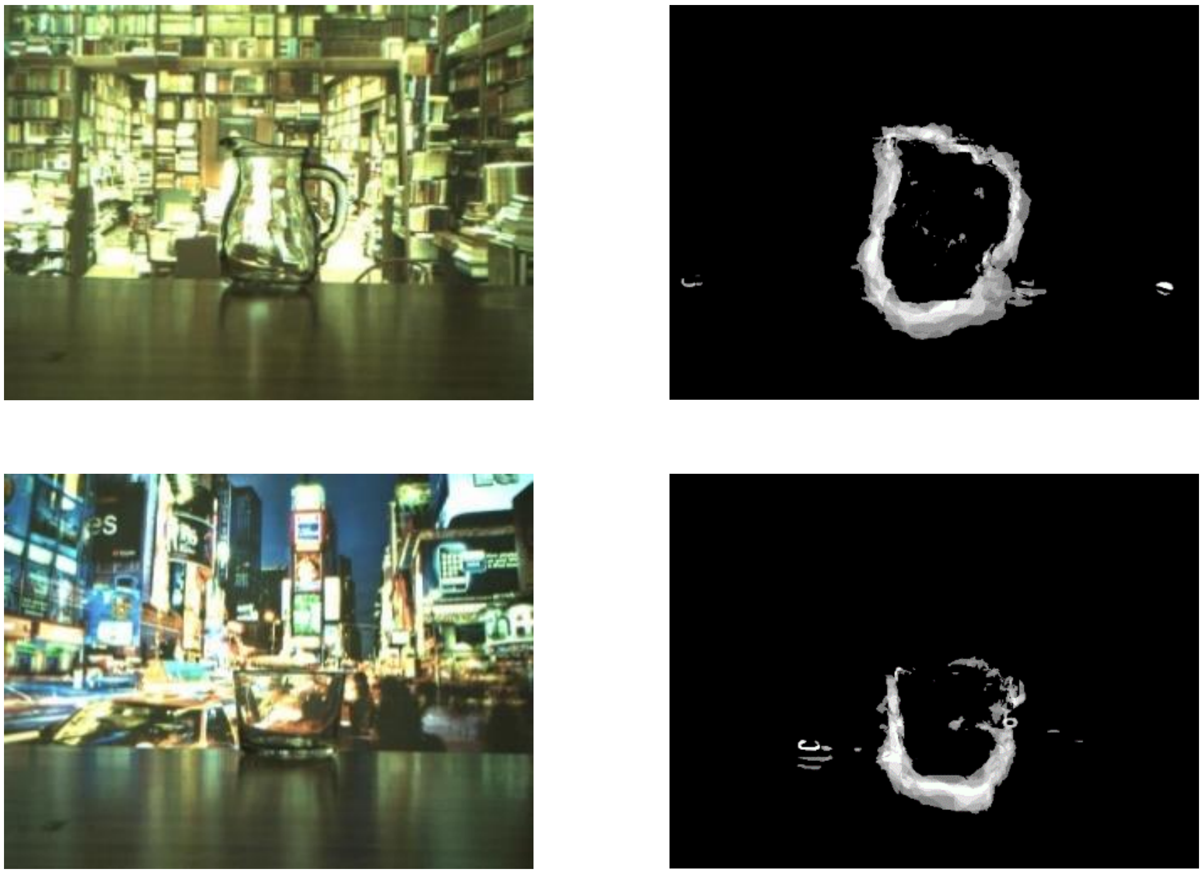
0	0	0	0.01	0.01
0	0	0	0.01	0.1
0	0	0	0.1	0.1
0	0	0.01	0.1	0.1
0	0.01	0.01	0.1	0.1

= 1 0.7

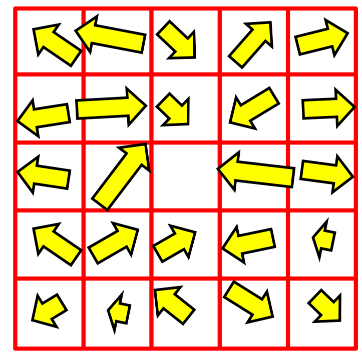
The detected occlusion point is from $\theta = 0$

Detected occlusion visualization

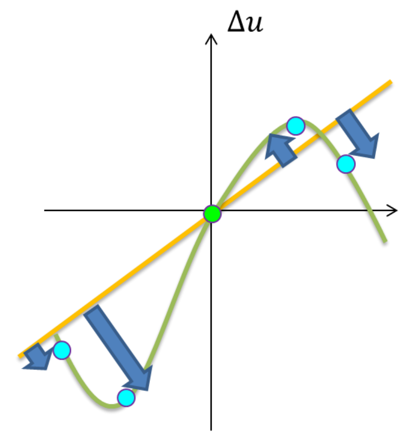
Central view | Occlusion response



Feature and descriptor



- LFD Feature (光场扭曲特征)
  - 2x24 Dimensional vector
  - Describe the distortion pattern



- LF-linearity（光场线性度）
  - A metric to describe how much is the distortion
- Occlusion detector （遮挡检测）
  - Describe the probability of a point to be in the occlusion
  - Occlusion in which direction

0	0	0	0	0
0	0	0	0	0.1
0	0	0	0.1	0.1
0	0	0.1	0.1	0.1
0	0.1	0.1	0.1	0.1

Outline

- Introduction of Light Field Vision
- Transcat: Transparent Object Categorization
- Transcut: Transparent Object Segmentation

TransCat: Transparent Object Categorization

Which object ?

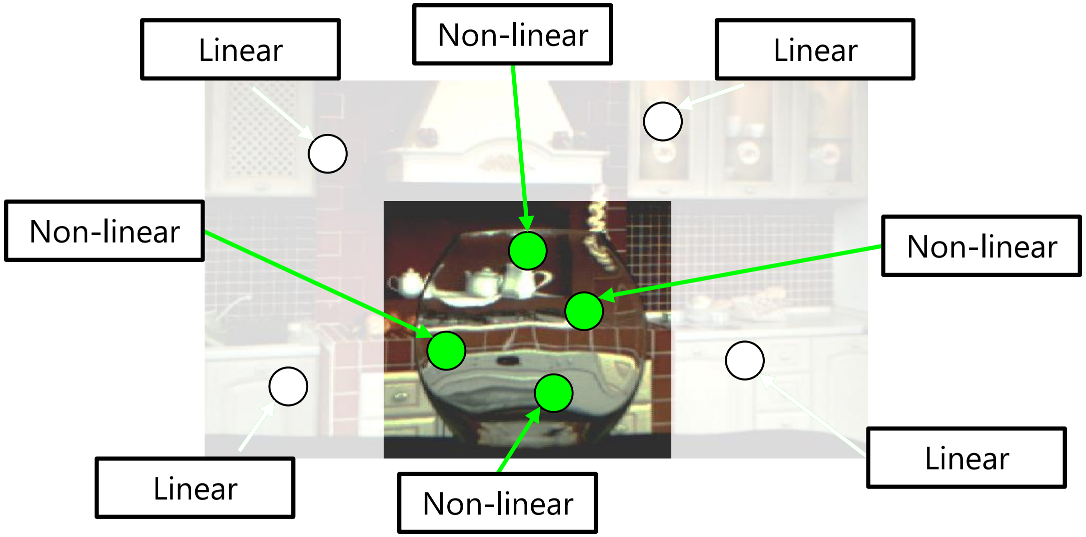
Training pipeline

Filtering the background

Extracting the LFD feature

Training based on Bag of features

Estimate background by LF-linearity



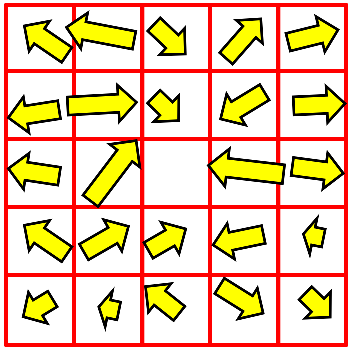
Training pipeline

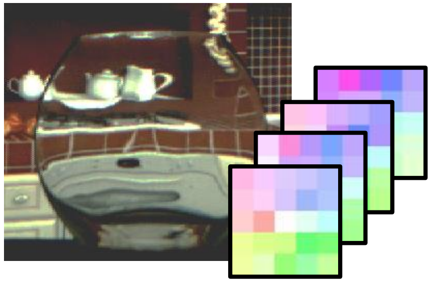
Filtering the background

Extracting the LFD feature

Training based on Bag of features

Estimate relative differences by optical flow



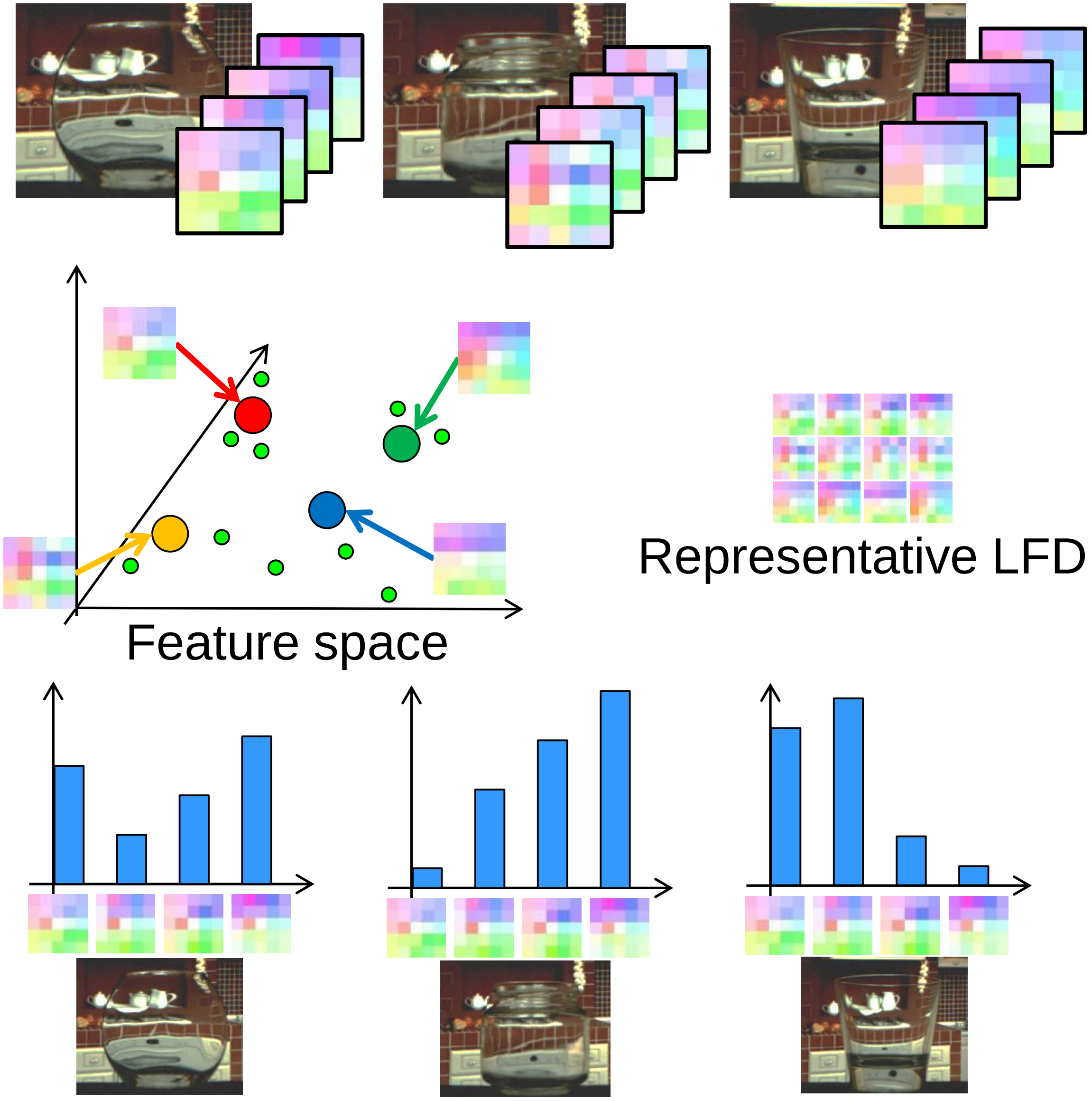


Training pipeline

Extracting the LFD feature

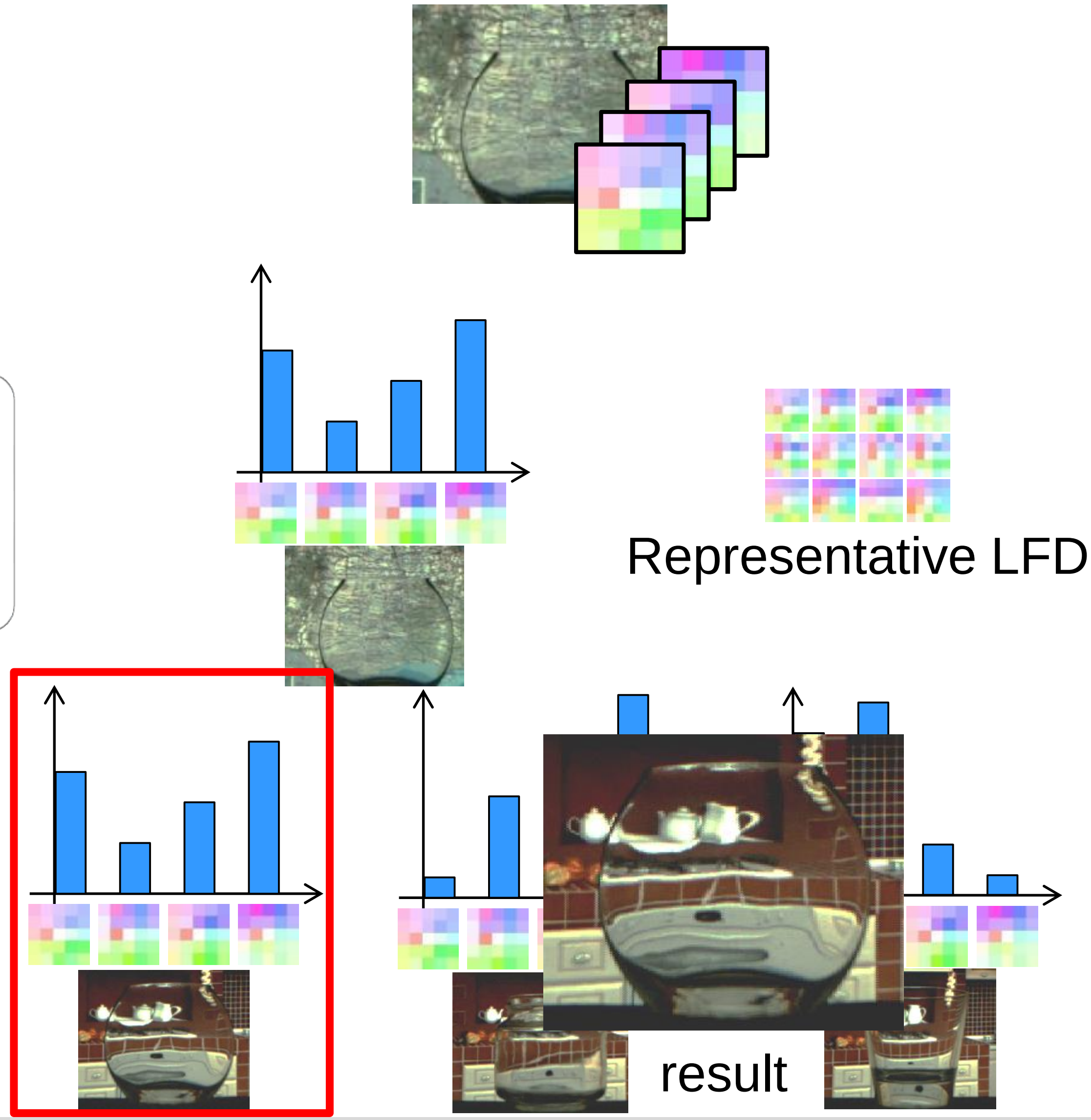
Filtering the background by LF-linearity

Training based on Bag of features



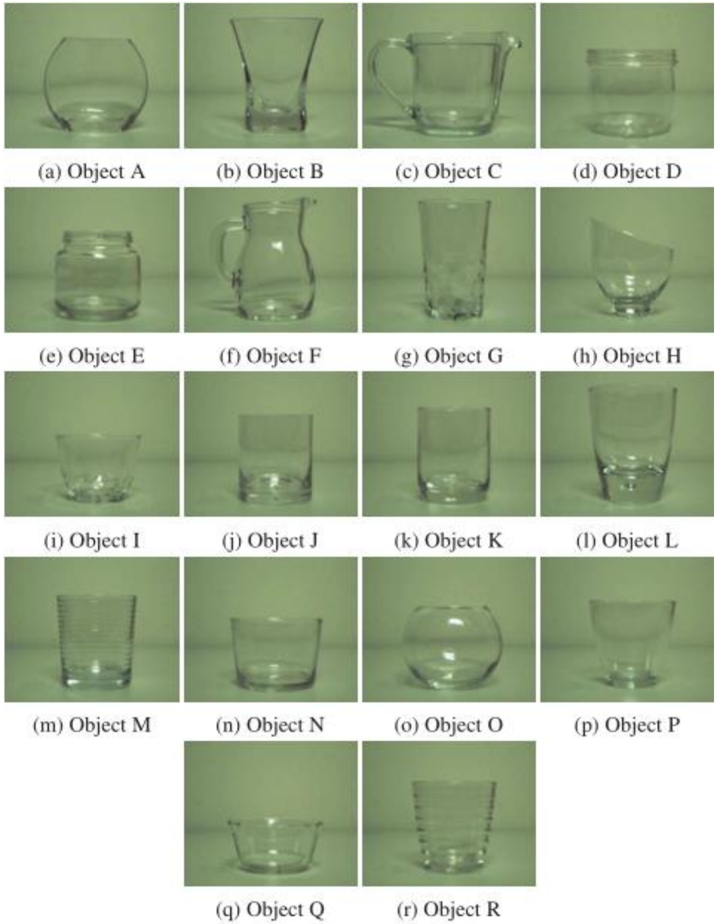
Testing for transparent objects categorization

Categorization based on Bag of features



Experimental setting

18 objects

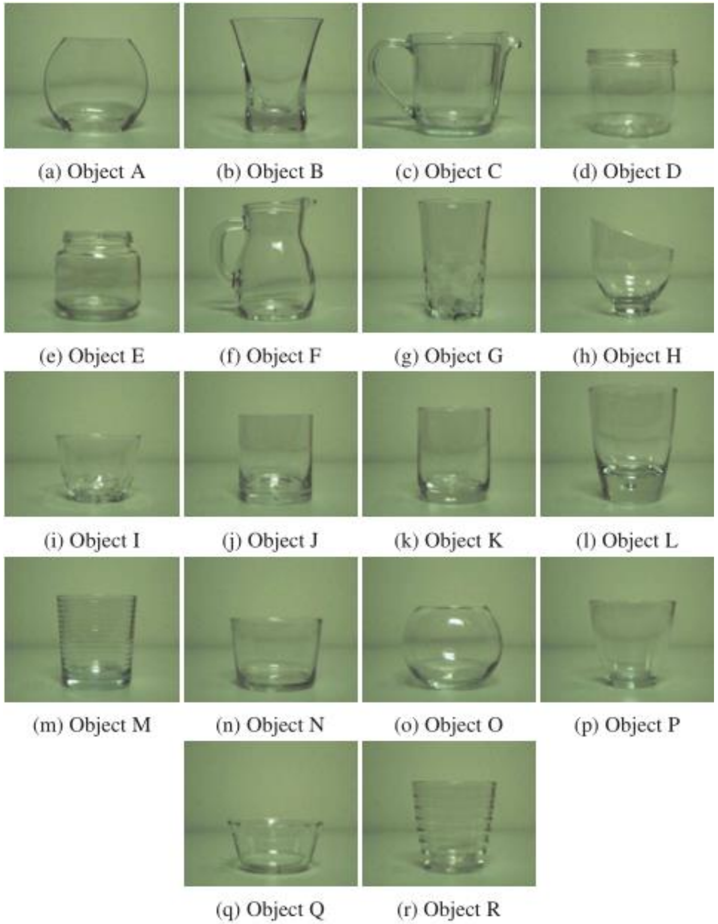


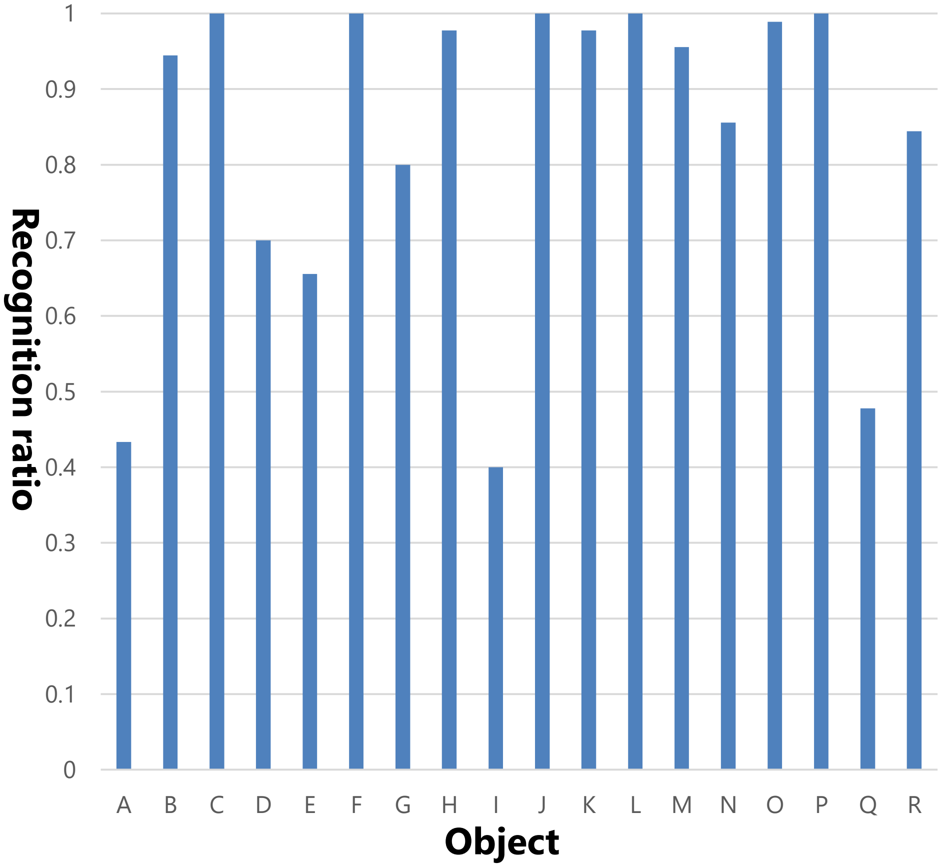
10 backgrounds

Background scenes can be dynamic!

Categorization result

Evaluation by leave-one-out cross-validation

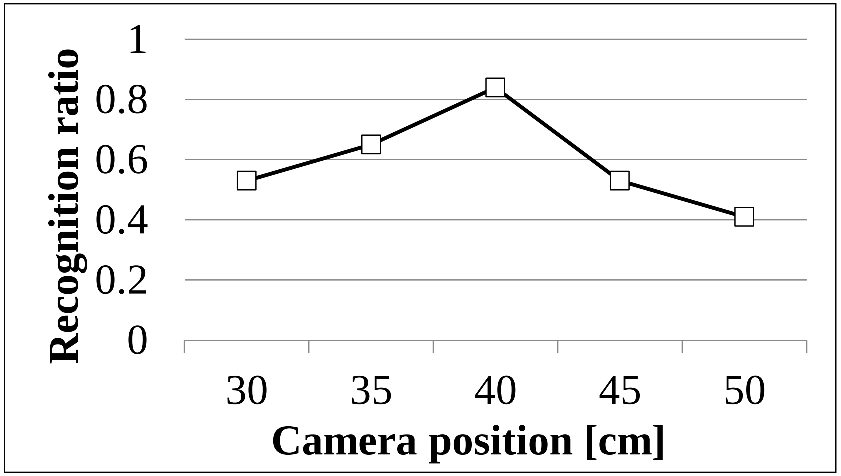


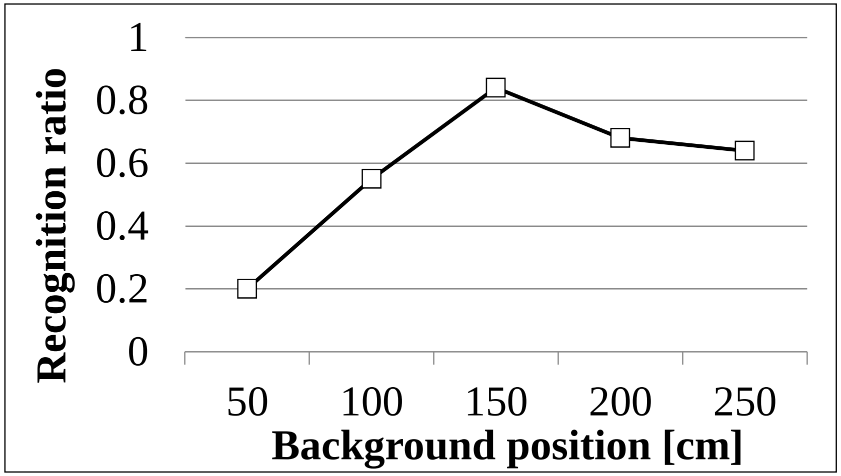


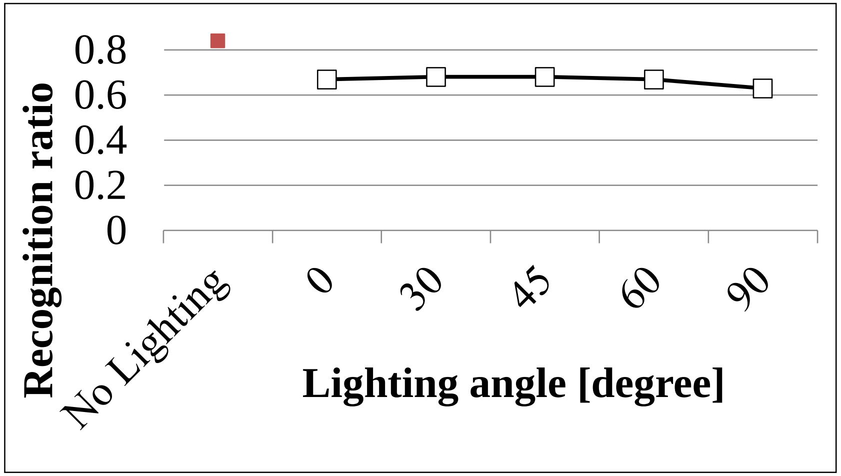
Average categorization accuracy: 84%

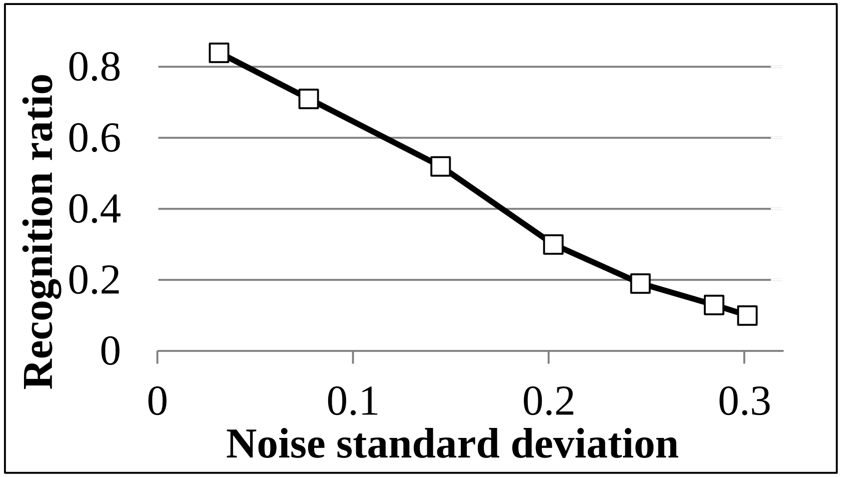
Analysis

- Applicable conditions



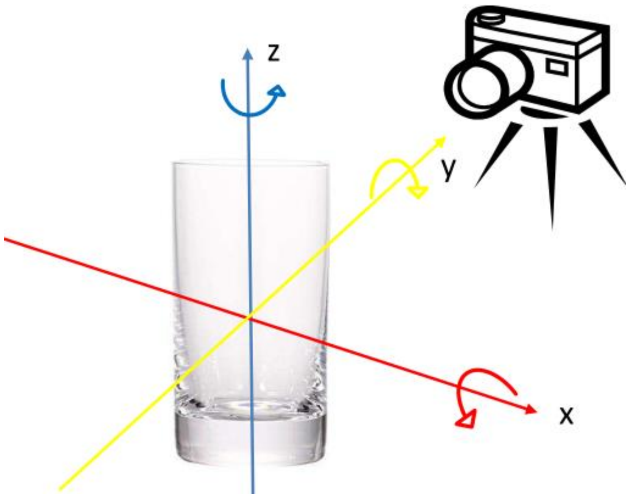


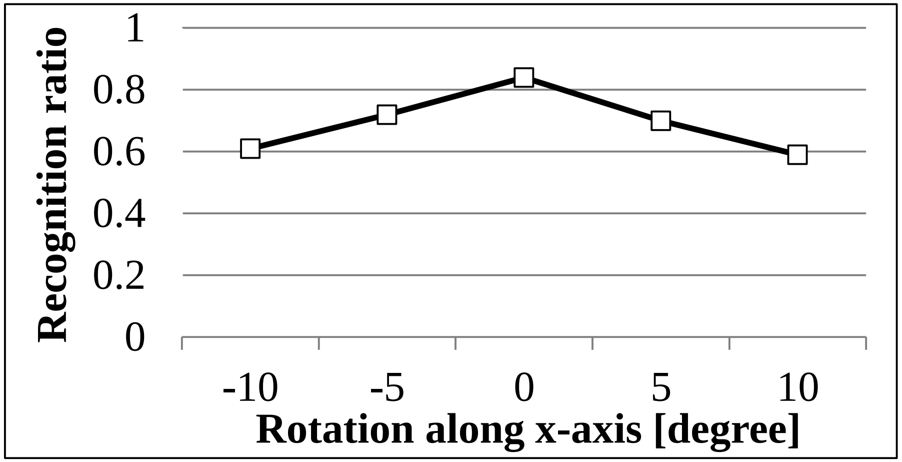


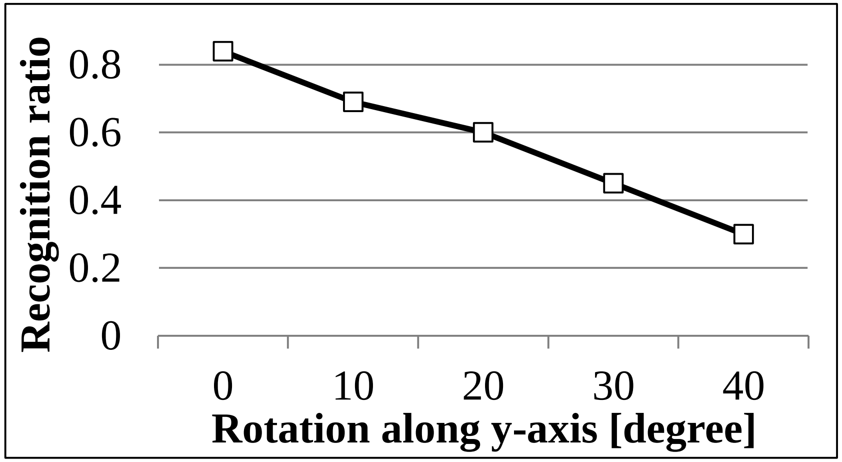


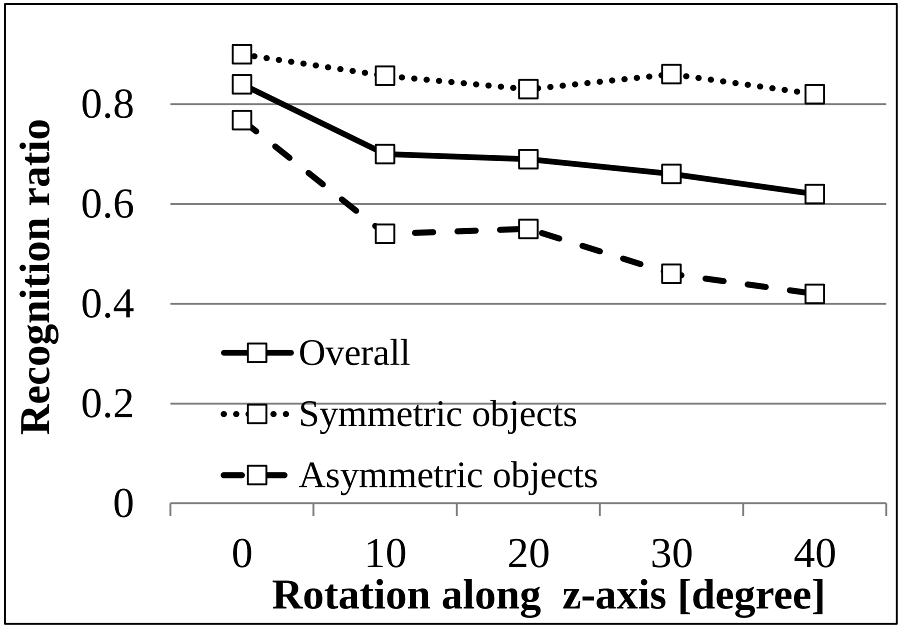
Analysis

- Applicable conditions

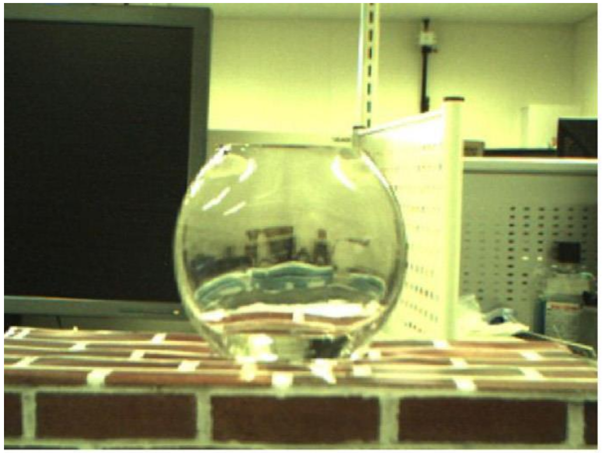








Results for real scene



(a) Indoor

(b) Outdoor

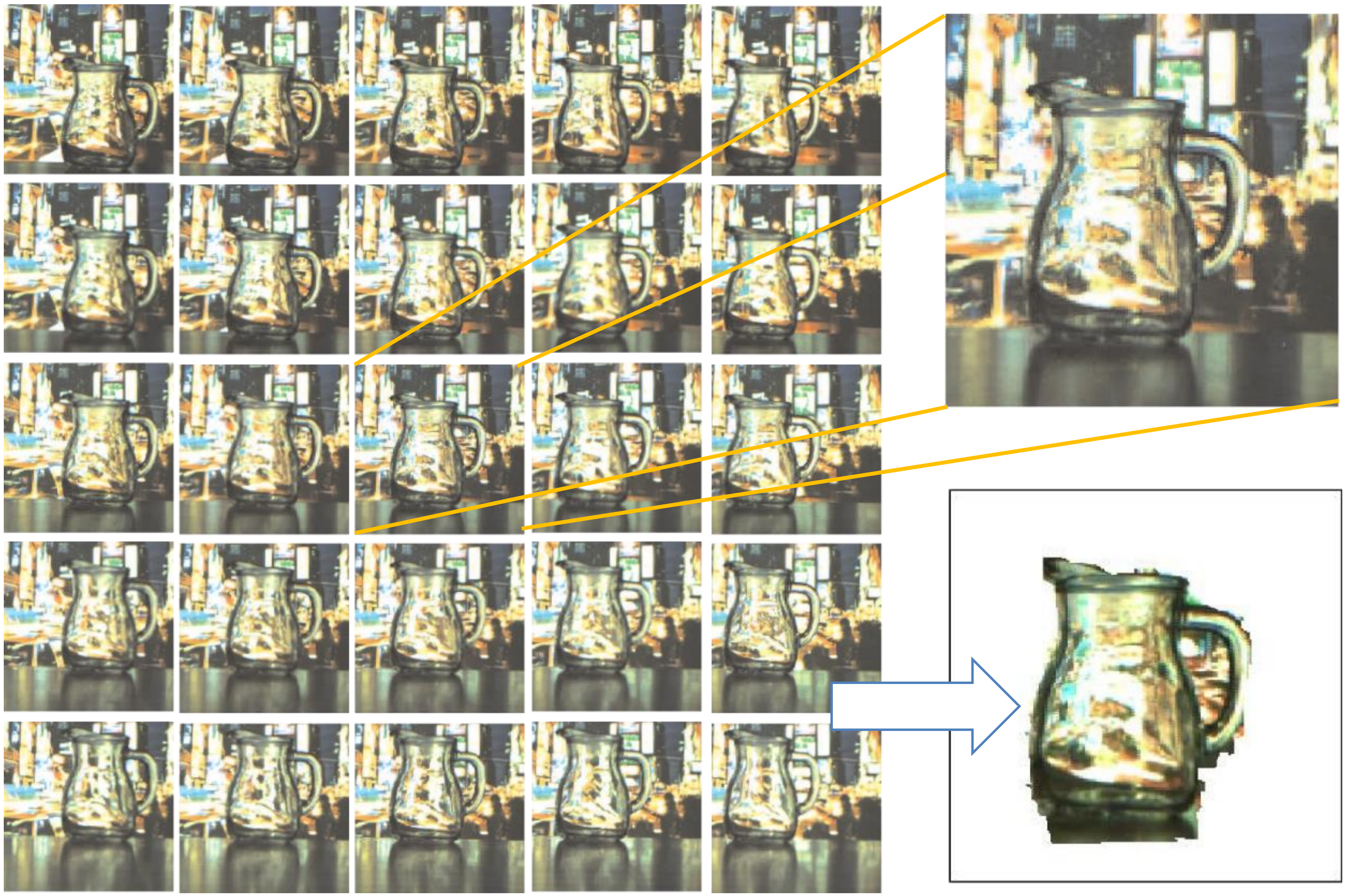
Recognition ratios for real experiment.

	6 objects	10 objects	15 objects	18 objects
Proposed LFD feature	0.766	0.678	0.587	0.533
Standard SIFT	0.160	0.108	0.075	0.063

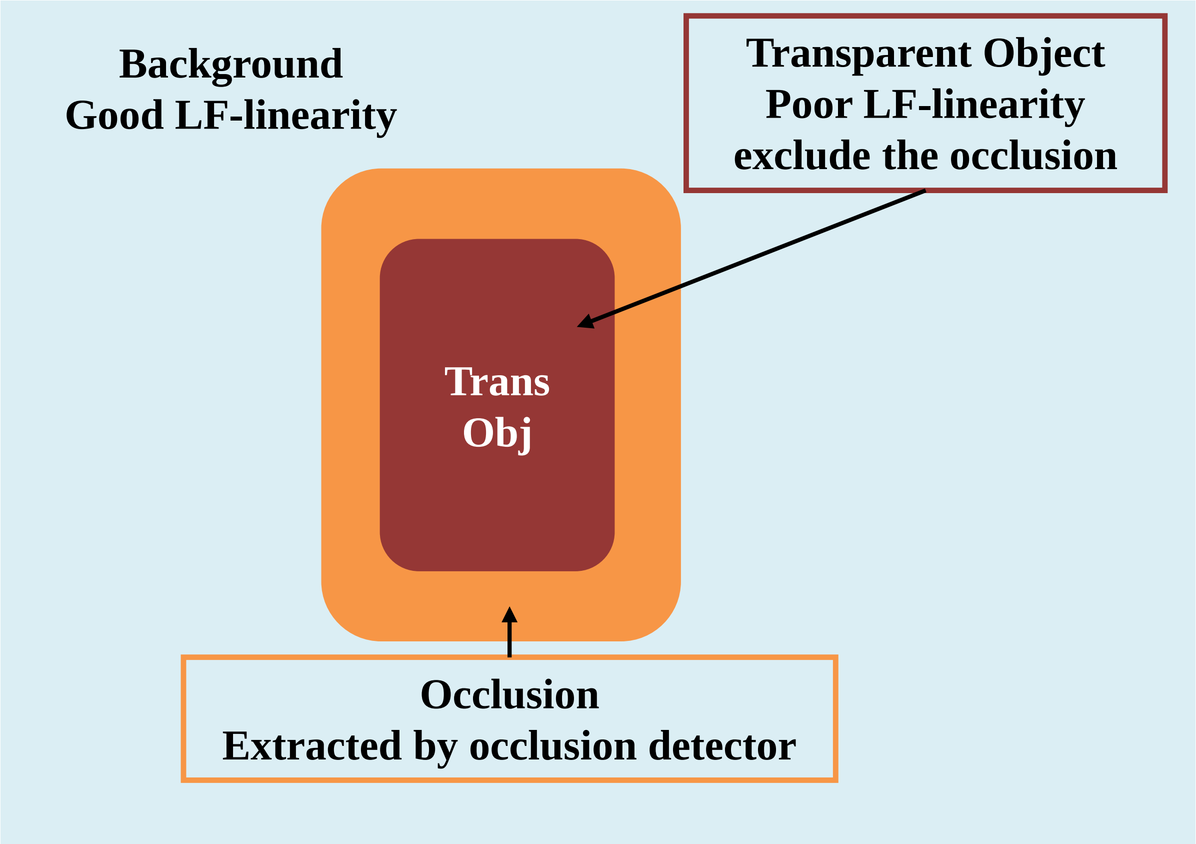
Outline

- Introduction of Light Field Vision
- Transcat: Transparent Object Categorization
- Transcut: Transparent Object Segmentation

Transcut: Transparent Object Segmentation



Properties of different components



$$E(l) = \sum_{p \in P} R_p(l_p) + \alpha \sum_{(p,q) \in N} B_{p,q} \cdot \delta(l_p, l_q)$$

Transparent object segmentation formulated as labeling problem

Regional term

Central view of input light field image

Background penalty

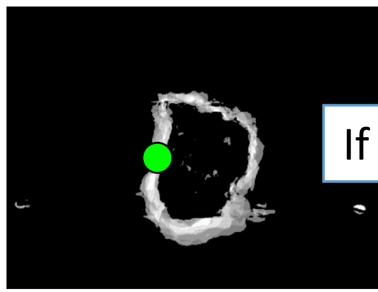
large penalty assigns to pixels that have poor LF-linearity exclude the occlusion area

Foreground penalty

large penalty assigns to pixels with poor LF-linearity in the occlusion or pixels with good LF-linearity

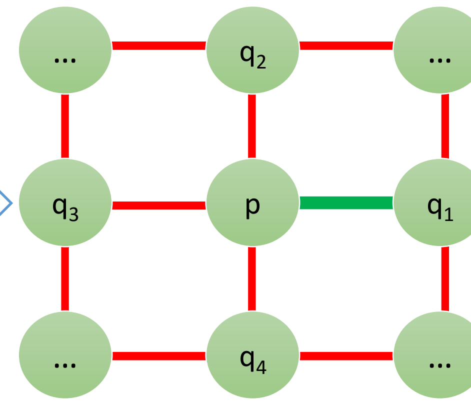
Boundary term

Central view of input light field image

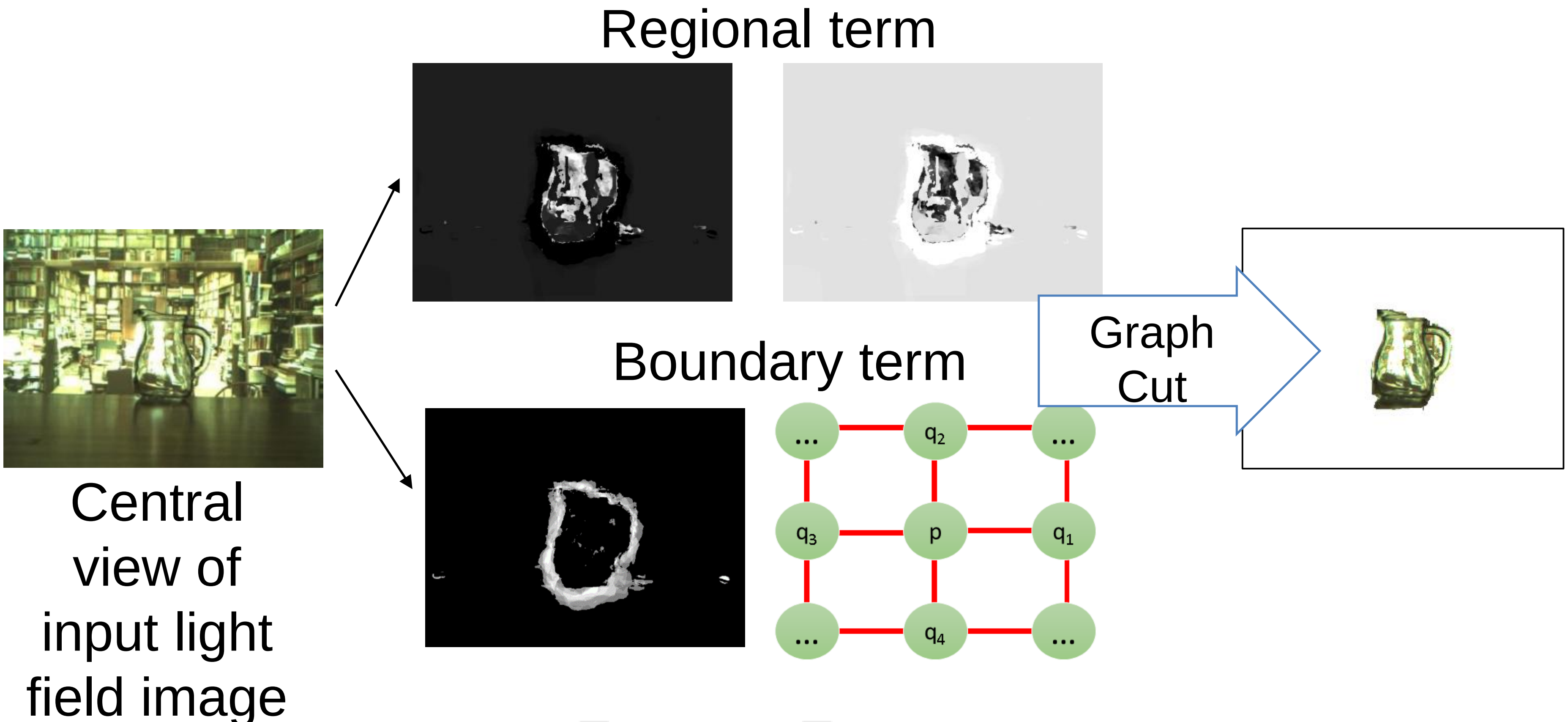


Detected occlusion point $\tilde{O}_p$

If $\tilde{O}_p$ is from $\theta = 0$,



Energy minimization



$$E(l) = \sum_{p \in P} R_p(l_p) + \alpha \sum_{(p,q) \in N} B_{p,q} \cdot \delta(l_p, l_q)$$

Experiments


Object 1


Object 2


Object 3


Object 4


Object 5


Object 6


Object 7


Scene 1


Scene 2


Scene 3


Scene 4


Scene 5

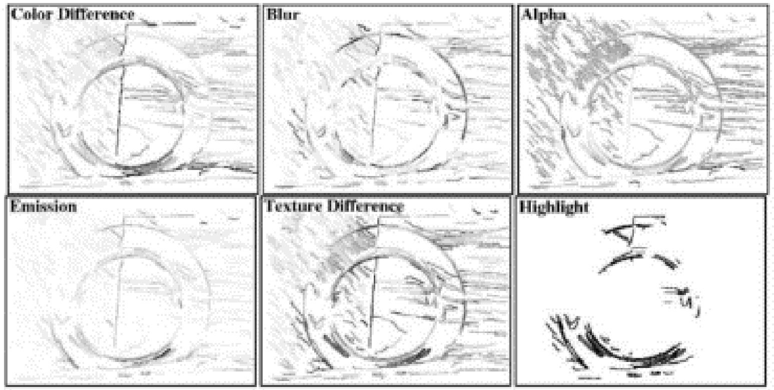

Scene 6


Scene 7

Background scenes can be dynamic!

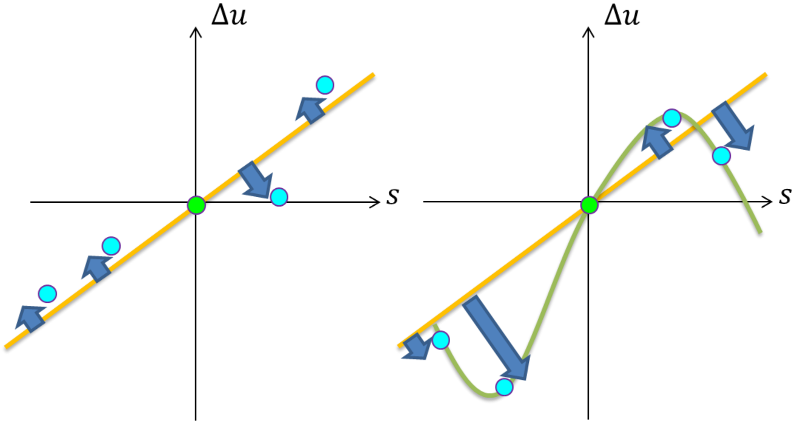
Comparison with related work

6 features from appearance



Finding glass
McHenry et al., CVPR2005

Single feature from LF

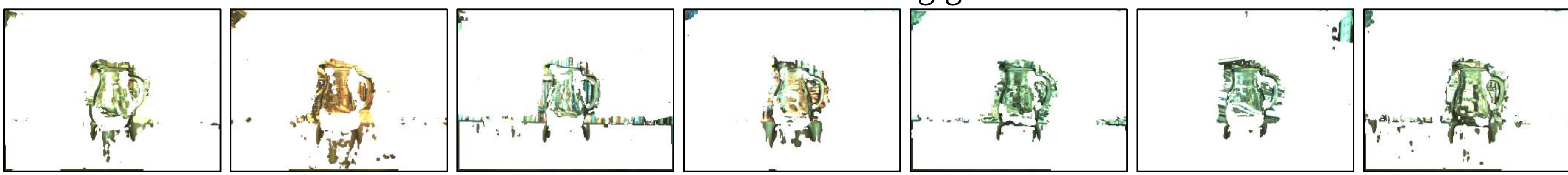


LF-linearity thresholding

Visual comparison

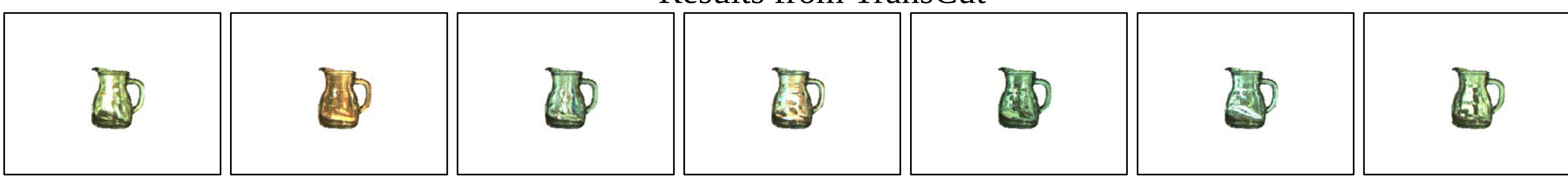
Images from the central viewpoint

Results from Finding glass



Results from LF-linearity thresholding

Results from TransCut



Ground Truth

Quantitative comparison

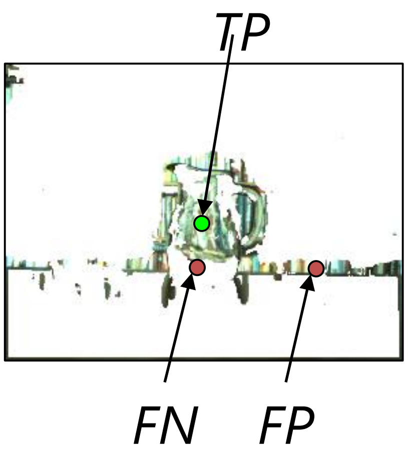
	F-measure	Recall	Precision
Finding glass	0.30	0.82	0.19
LF-linearity thresholding	0.50	0.65	0.41
Proposed	0.85	0.96	0.77

	F-measure	Recall	Precision
Th=1	0.19	0.84	0.11
Th=2	0.34	0.78	0.22
Th=3	0.42	0.76	0.29
Th=4	0.45	0.73	0.33
Th=5	0.48	0.70	0.37
Th=6	0.49	0.67	0.39
Th=7	0.50	0.65	0.41
Th=8	0.50	0.63	0.43
Th=9	0.50	0.61	0.44

$$F = \frac{2 * Precision * Recall}{Precision + Recall}$$

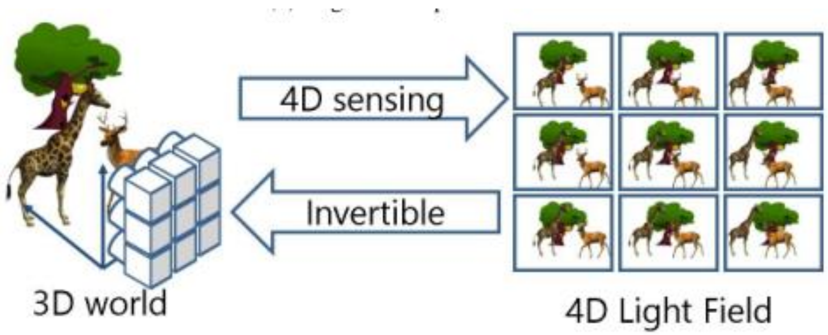
$$Recall = TP/(TP+FN)$$

$$Precision = TP/(TP+FP)$$

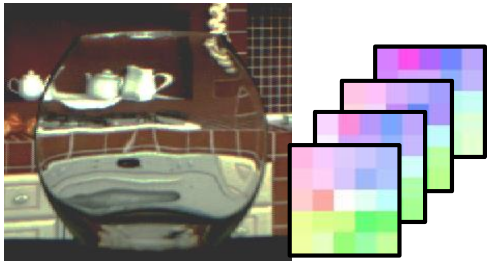


Summary

- Light field vision can get more information for solving vision problems
  - Full space sampling vs. sub-space sampling



- Transparent object categorization

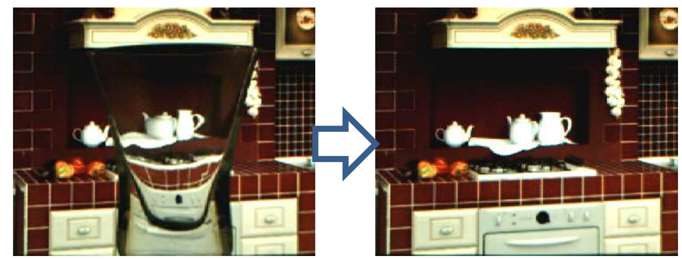


- Transparent object segmentation

Open issues

- Develop robust feature descriptors
  - Distance invariant
  - Rotation invariant
- Apply to other objects
  - Specular objects

- Recover the undistorted background



- Reconstruct 3D shape of transparent objects
  - Natural scenes

Publications

- **Y. Xu**, K. Maeno, H. Nagahara, and R. Taniguchi, “Mobile camera array calibration for light field acquisition,” in International Conference on Quality Control by Artificial Vision (QCAV), pp. 283–290, 2013.
- **Y. Xu**, K. Maeno, H. Nagahara, and R. Taniguchi, “Camera array calibration for light field acquisition,” Frontiers of Computer Science, 2015, 9(5), pp. 691-702.
- **Y. Xu**, K. Maeno, H. Nagahara, A. Shimada, and R. Taniguchi, “Light field distortion feature for transparent object classification,” Computer Vision and Image Understanding, Vol. 139, pp. 122-135, 2015.
- **Y. Xu**, H. Nagahara, A. Shimada, and R. Taniguchi， "TransCut: Transparent Object Segmentation from a Light-Field Image", ICCV 2015, Santiago, Chile

Thank you!