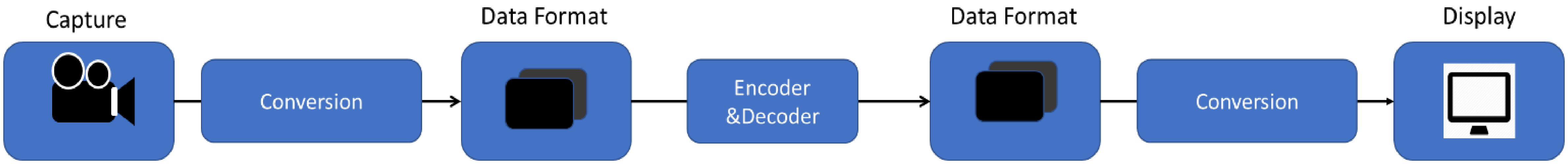
How to achieve dense light field video compression?

Mehrdad Teratani (Associate Professor, Nagoya University)
Xin Jin (Graduate School at Shenzhen, Tsinghua University)



N18445: Activity Report on Dense Light Fields

Dense light field

Muti-cam (multiview)

2004-2008

NAGOYA UNIVERSITY

2012 @ Osaka

NICT

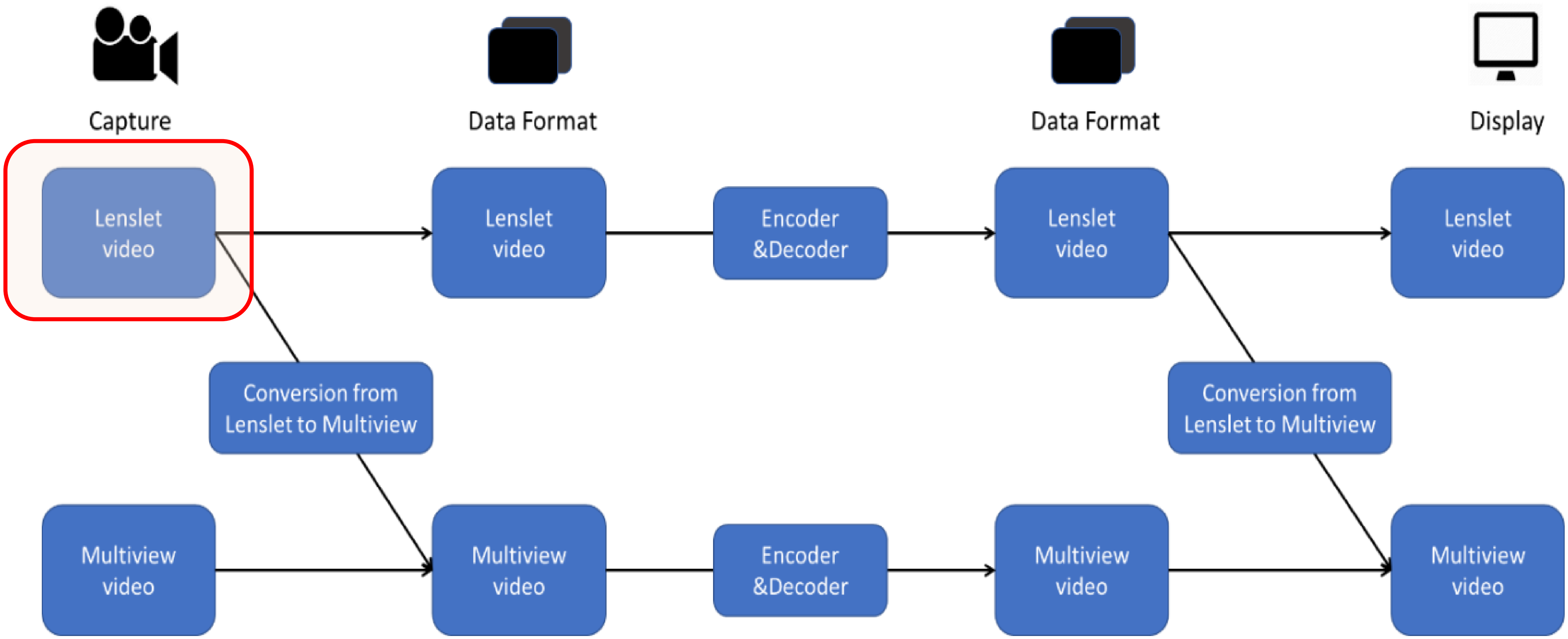
Plenoptic (lenslet)

清華大學

Tsinghua University

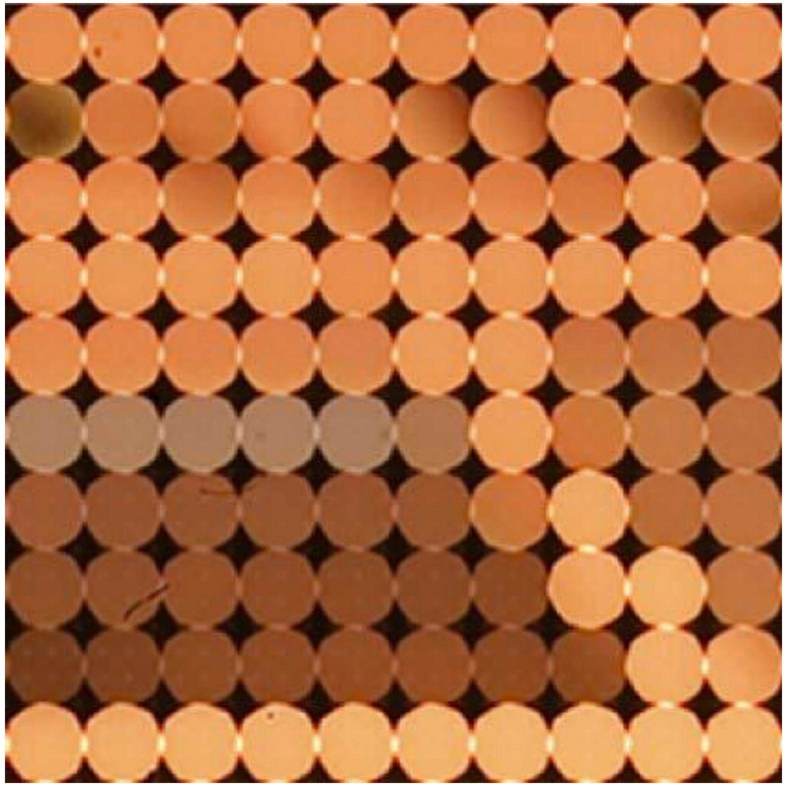
NHK

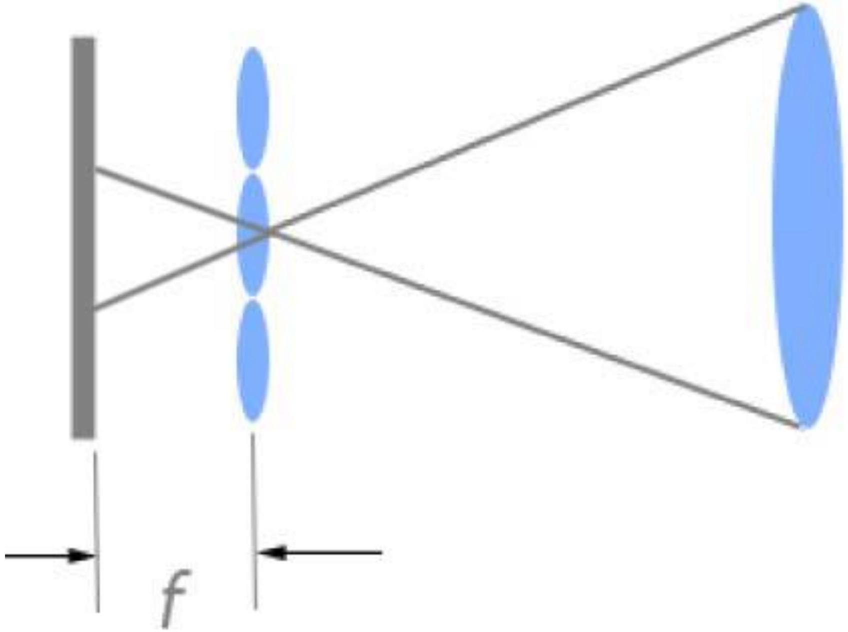
End-to-end system for dense light field



Plenoptic Camera type 1.0

Plenoptic 1.0 (e.g., Lytro)





“Lytro,” https://www.lytro.jp/.

- Spatial resolution = number of microlens.
- Completely defocused relative to main lens image.

GSST Plenoptic 1.0 Lenslet Data

Tsinghua University

M44684: "Toys"

M44684: "Trees"

Lenslet Video Data
Resolution: 8656×6075 Color: 24 bits, PNG Frame rate: 30 fps

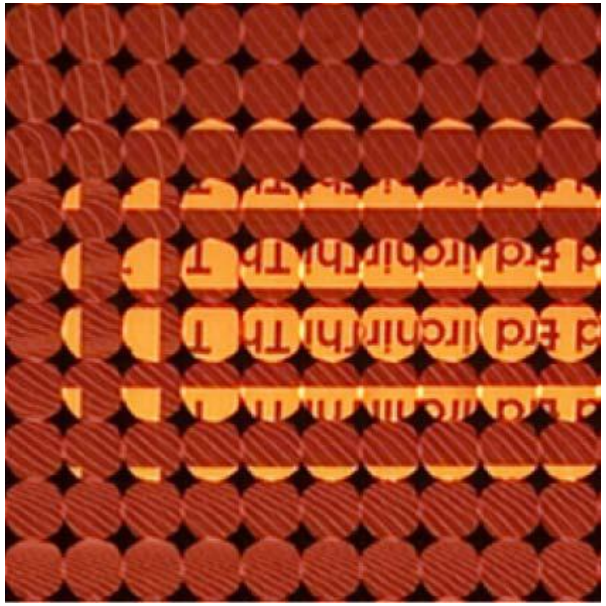
M46258: "Teapots"

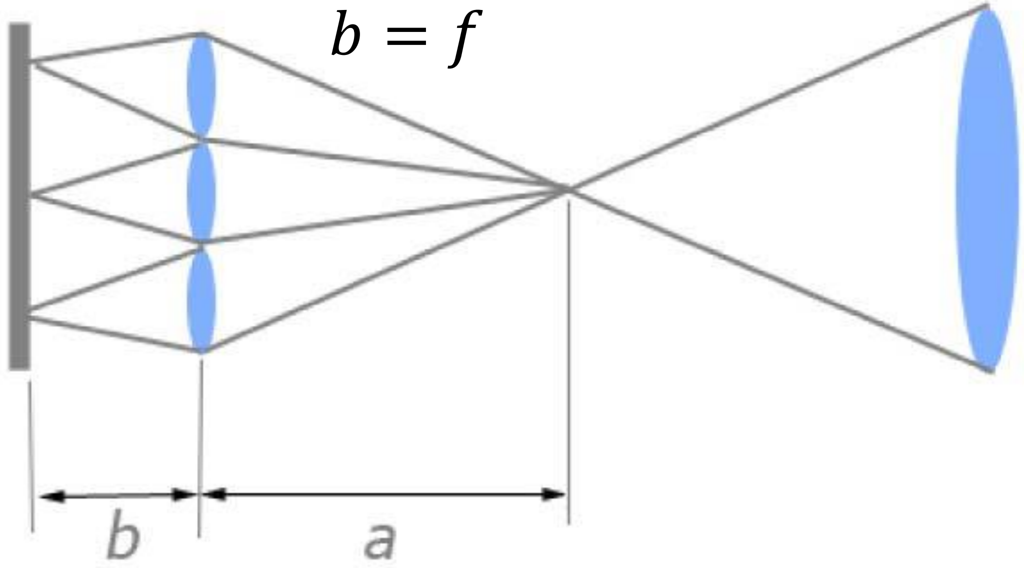
M46258: "Mini-garden"

Lenslet Video Data
Resolution: 8654×6074 Color: 24 bits, PNG Frame rate: 30 fps

Plenoptic Camera 2.0

Single-Focused Plenoptic Camera





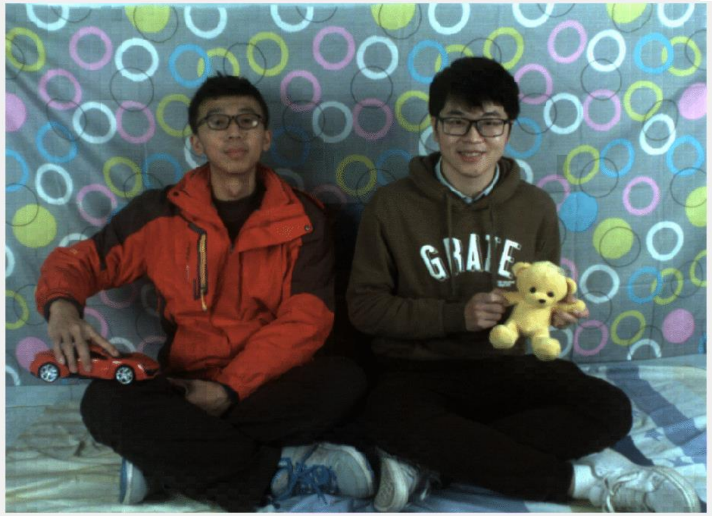
Designed and made by Tsinghua University

- resolution independent of microlenses
- spatial-angular resolution = free tradeoff point.
- Exactly focused on the main lens image.

GSST Plenoptic 2.0 Lenslet data

清華大學 Tsinghua University

M46259: “Boys”



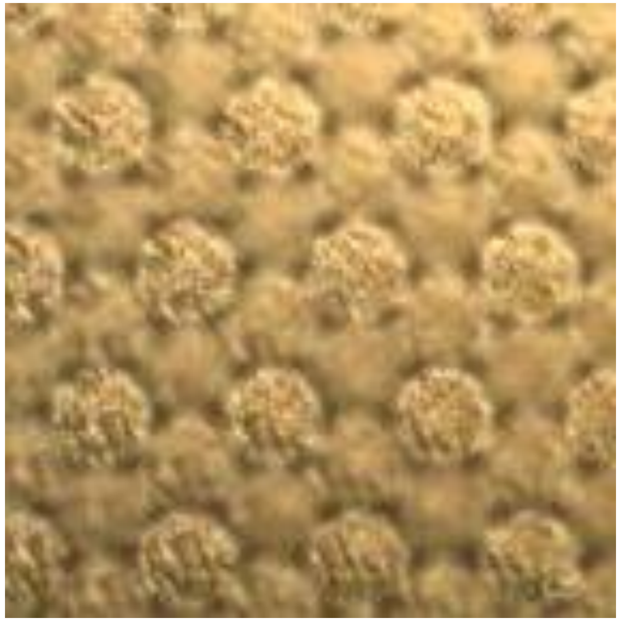
M46259: “Experimenting”

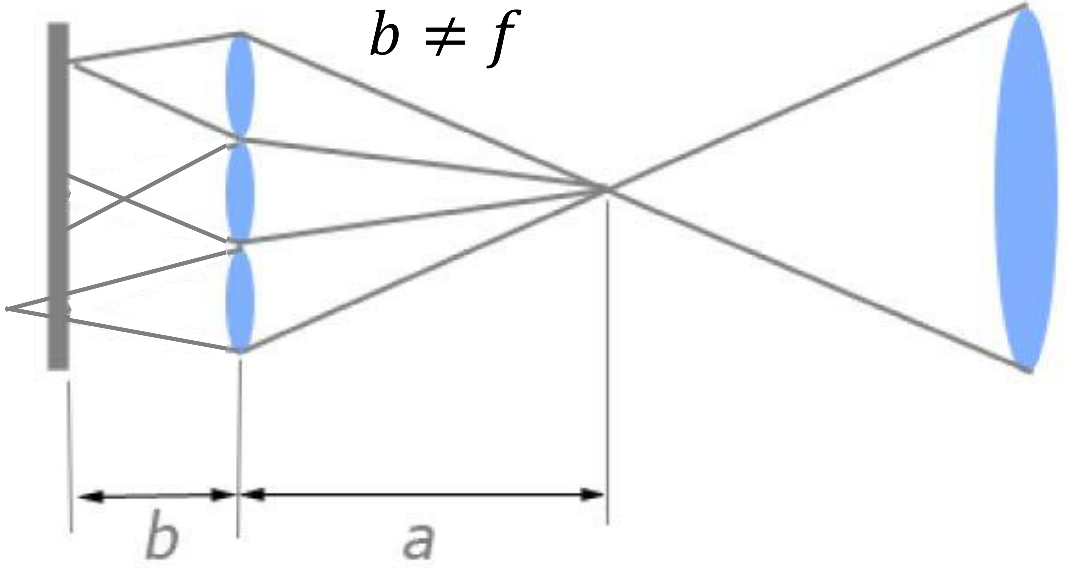
M49007: “cars”

Colored Lenslet Video Data	**Multiview Video Data**
Resolution: 4088×3068 Color: 24 bits, BMP Frame rate: 30 fps Number of frames: 300	Resolution of each view: 926×672 Views: 5×5 Number of frames: 300

Plenoptic Camera 2.0

Multi-Focused Plenoptic Camera





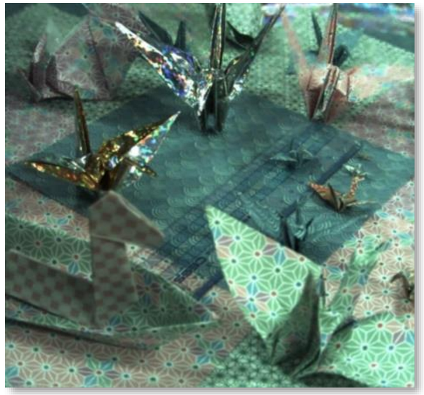
"Raytrix," https://www.raytrix.de/.

- Flexible (3 kind of ML)
- resolution independent of microlenses
- spatial-angular resolution = free tradeoff point.

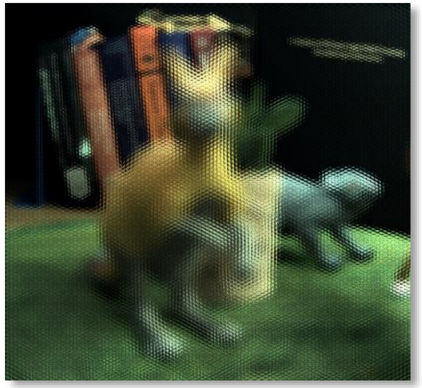
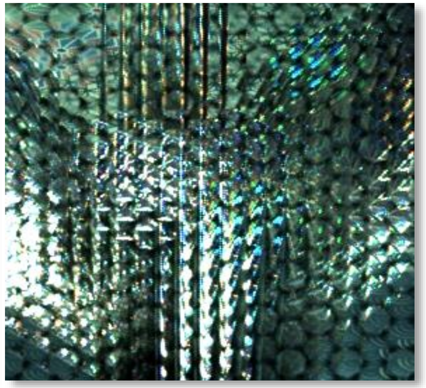
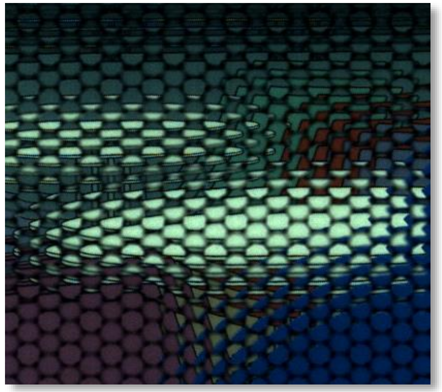
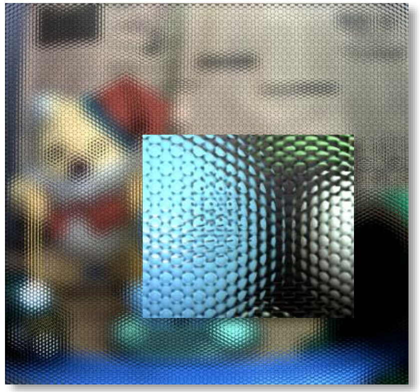
Nagoya University Lenslet data Plenoptic 2.0

NAGOYA UNIVERSITY

Center viewpoint

Lenslet

NagoyaFujita
Fixed Camera
Horizontal view
M47642, M49670

NagoyaOrigami
Fixed Camera
Top view
M47642, M49670

NagoyaDataLeading
Camera on turn table
Horizontal view
M47642, M49670

Tunnel_Train_2
Fixed Camera
Horizontal view
M41787

Colored Lenslet Video
Resolution: **2048x2048 pixels**
Color: 24 bits PNG, and YUV420
Frame rate: 30 fps
Number of frames: 300-400
Camera parameters (SDK output)

Raytrix R5-C- GigE-F2.4 (color)
Main Lens: LMVZ166HC (Kowa)

INRIA Lenslet data Plenoptic 2.0

Colored Lenslet Video
Resolution: **3840 ×2160 pixels**
Color: 24 bits, PNG
Frame rate: 30 fps
Number of frames: 300
Camera parameters (SDK output)

Center viewpoint

Lenslet

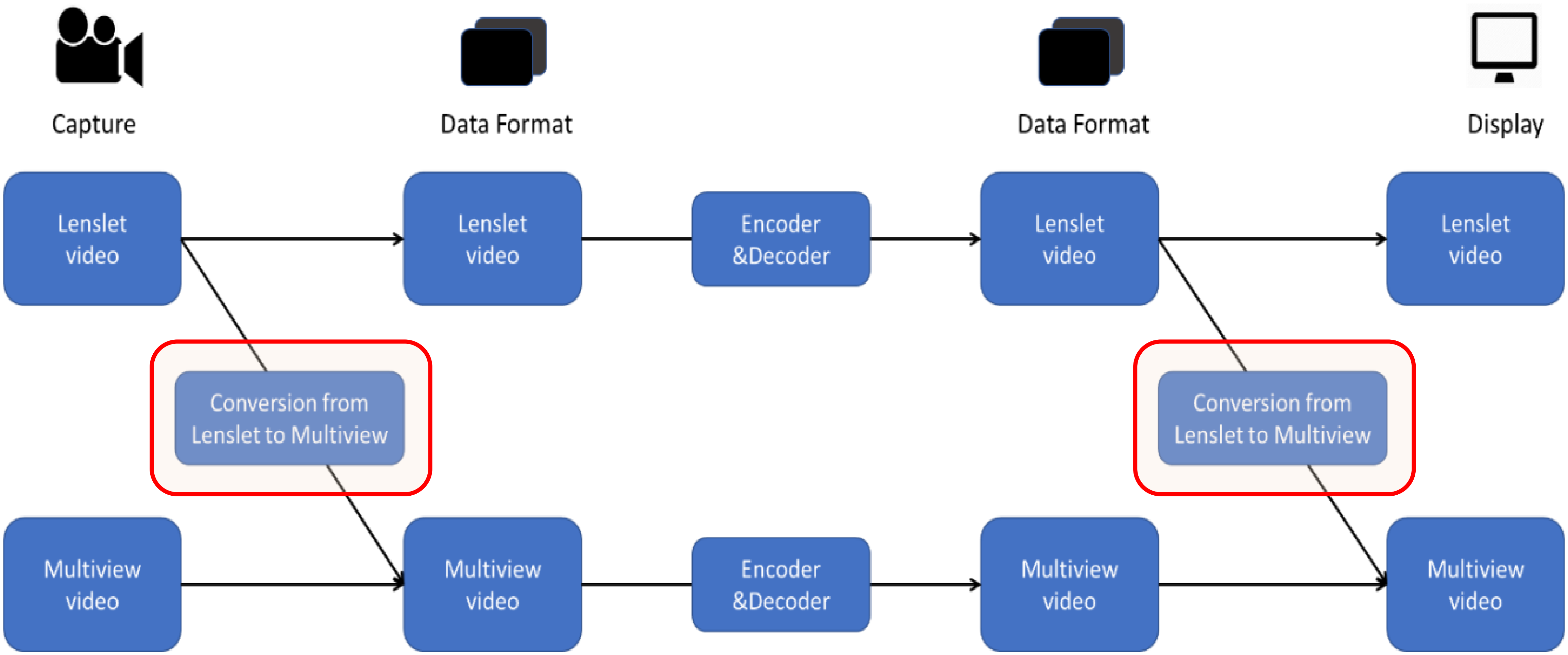
Boxer-IrishMan-Gladiator   Chess-Pieces   ChessPieces-MovingCamera

m42468

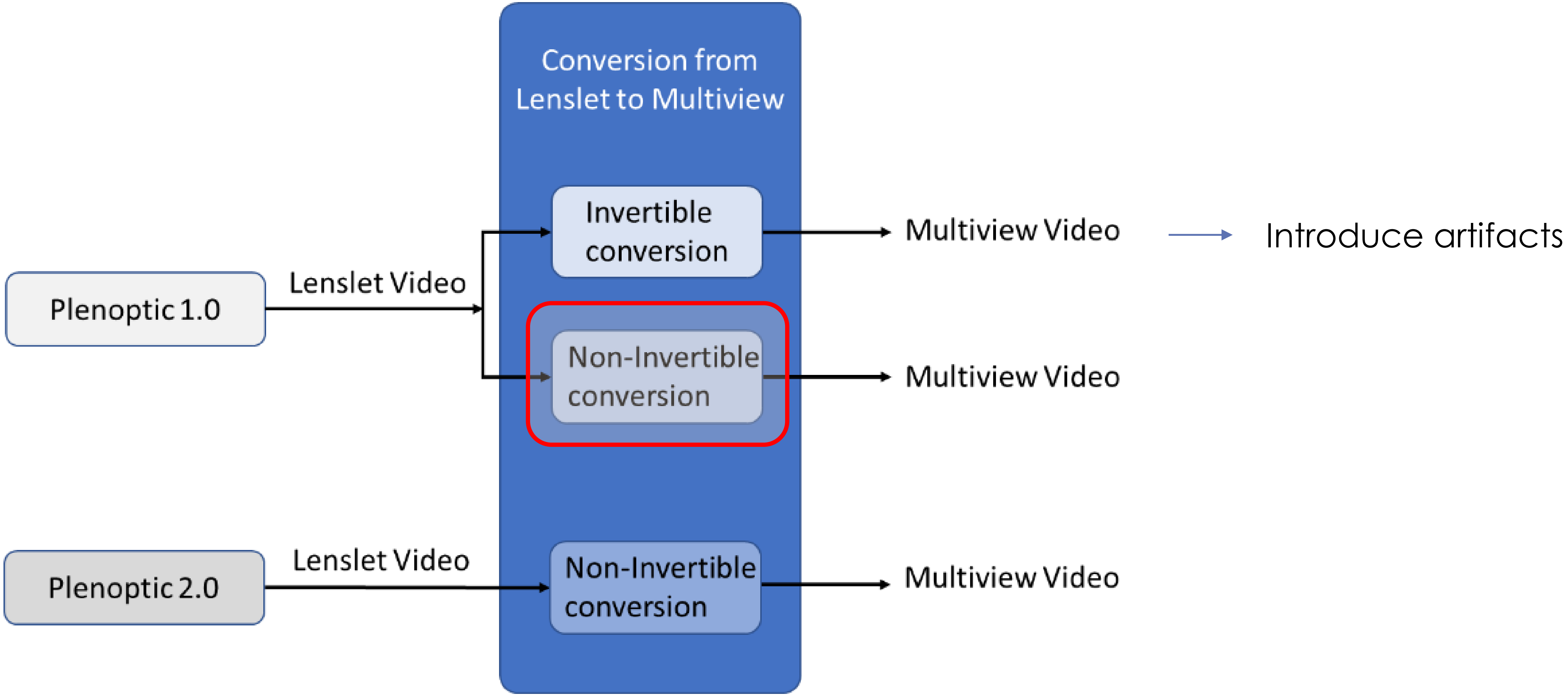
Raytrix R8

“Raytrix,” https://www.raytrix.de/.

End-to-end system for dense light field

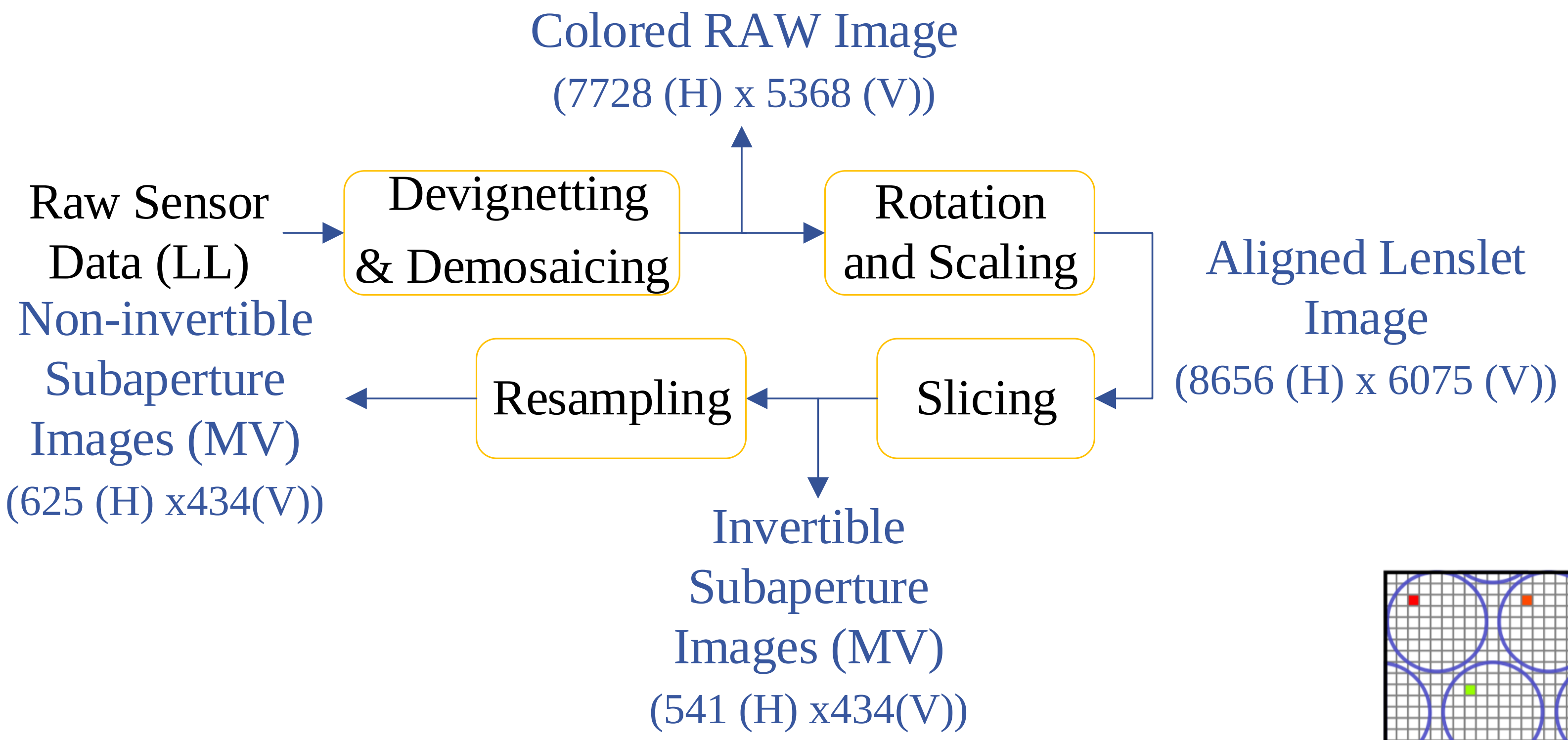


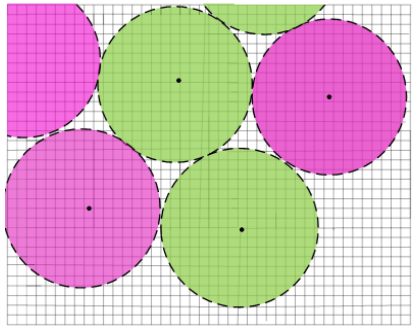
Conversion from Lenslet to Multiview



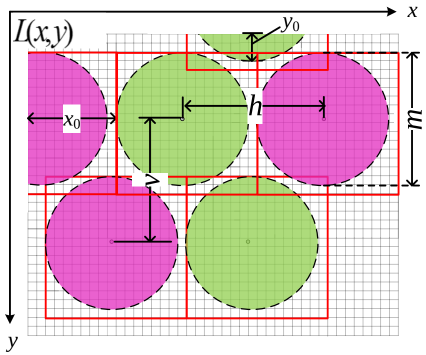
Plenoptic 1.0 Data Conversion Tool

Tsinghua University

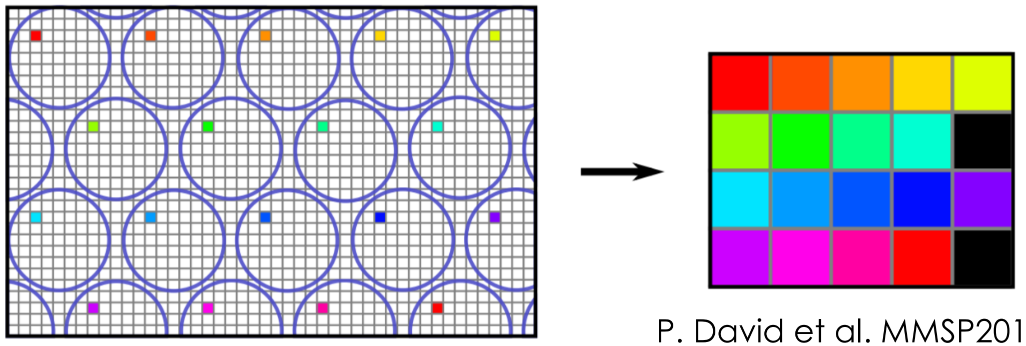




Color Raw image

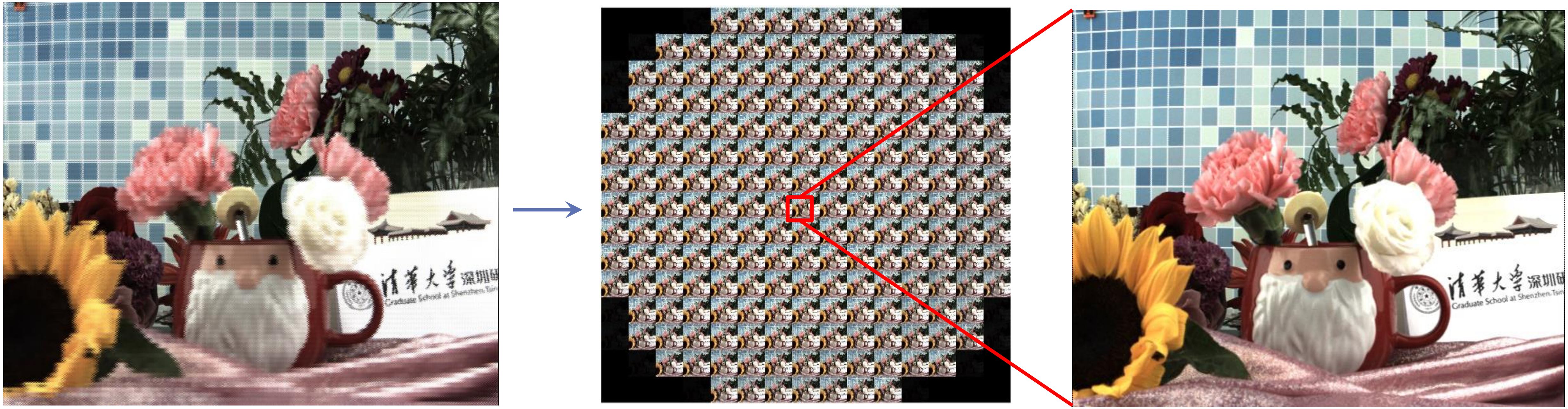


Color Lenslet image

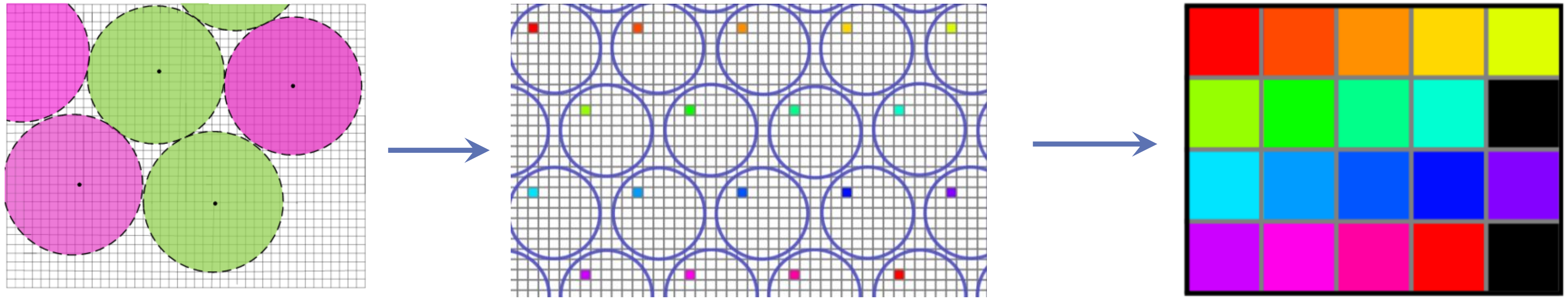


P. David et al. MMSP2017

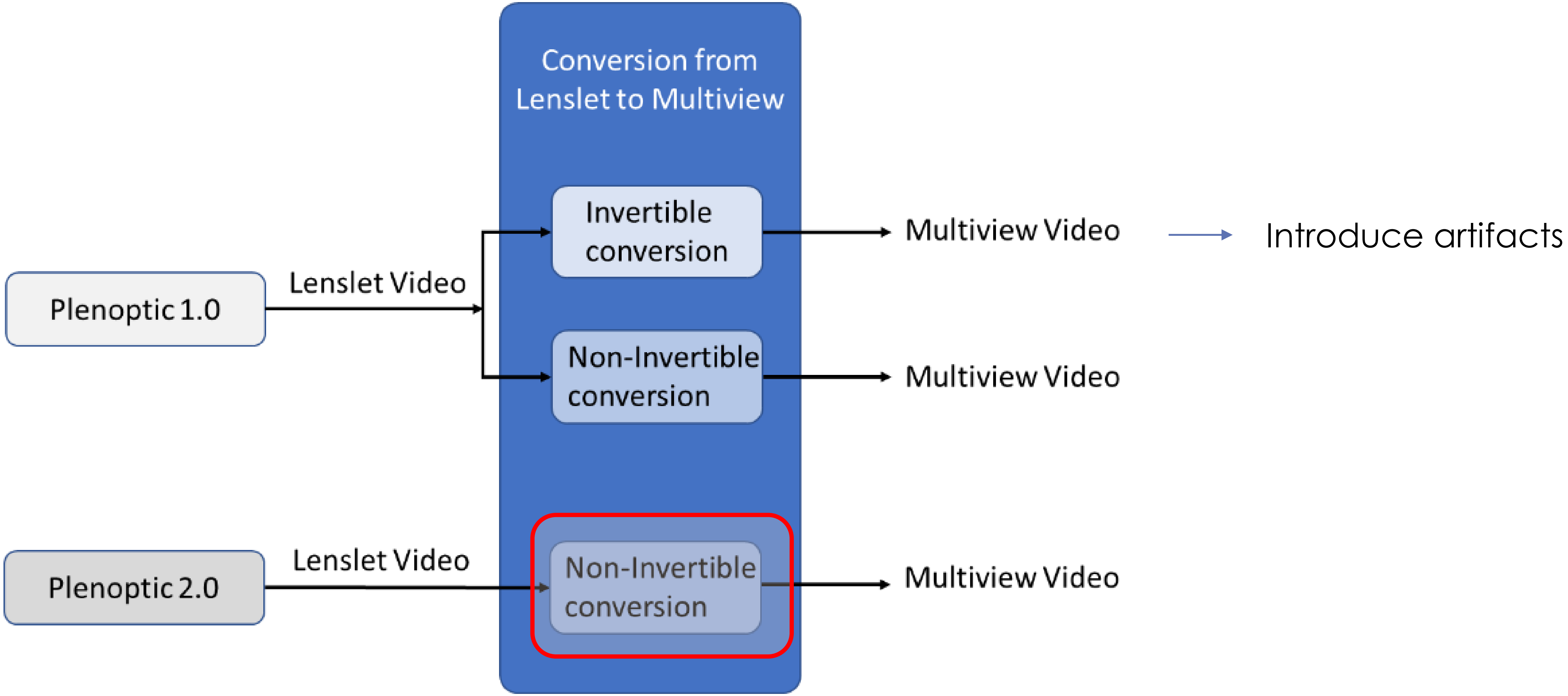
Plenoptic 1.0 Data Conversion Tool



M44684: New Test Sequences "Toys'' and "Trees" Captured by a Light Field Camera @MPEG,Macao

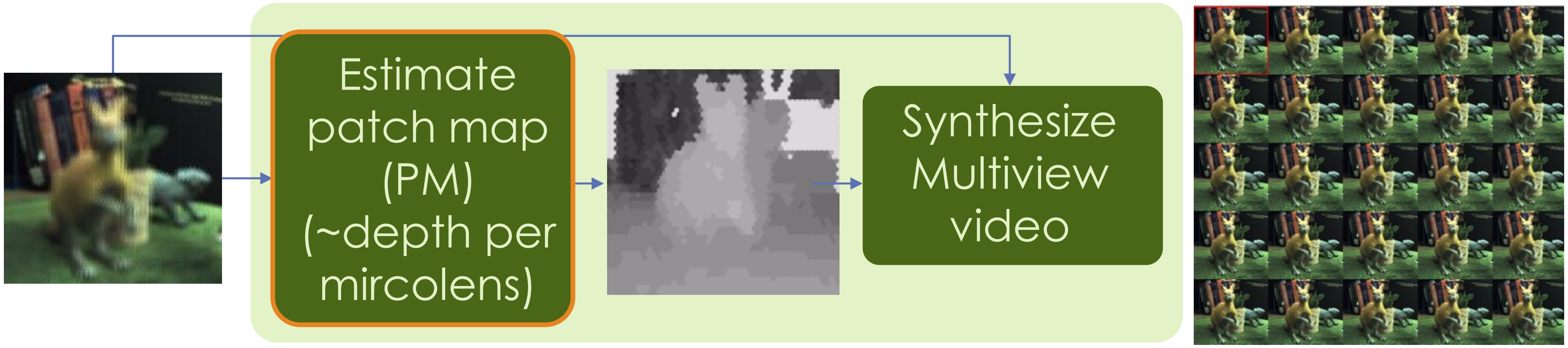


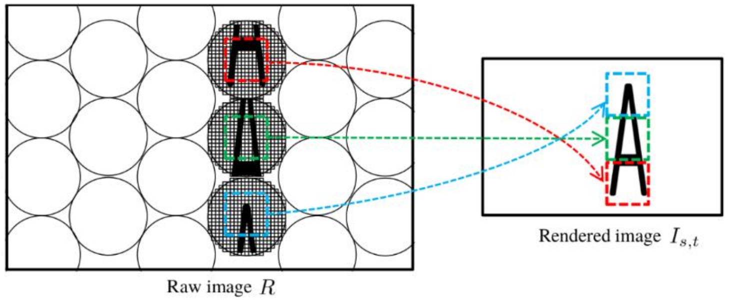
Conversion from Lenslet to Multiview

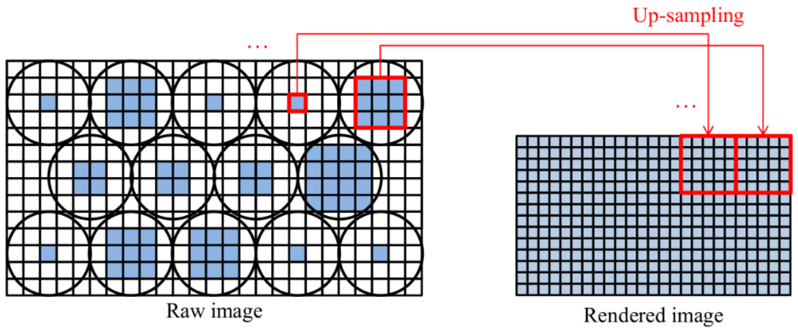


Reference Lenslet content Convertor

NAGOYA UNIVERSITY







$$K(i,j) = \arg\min E_k(i,j)$$

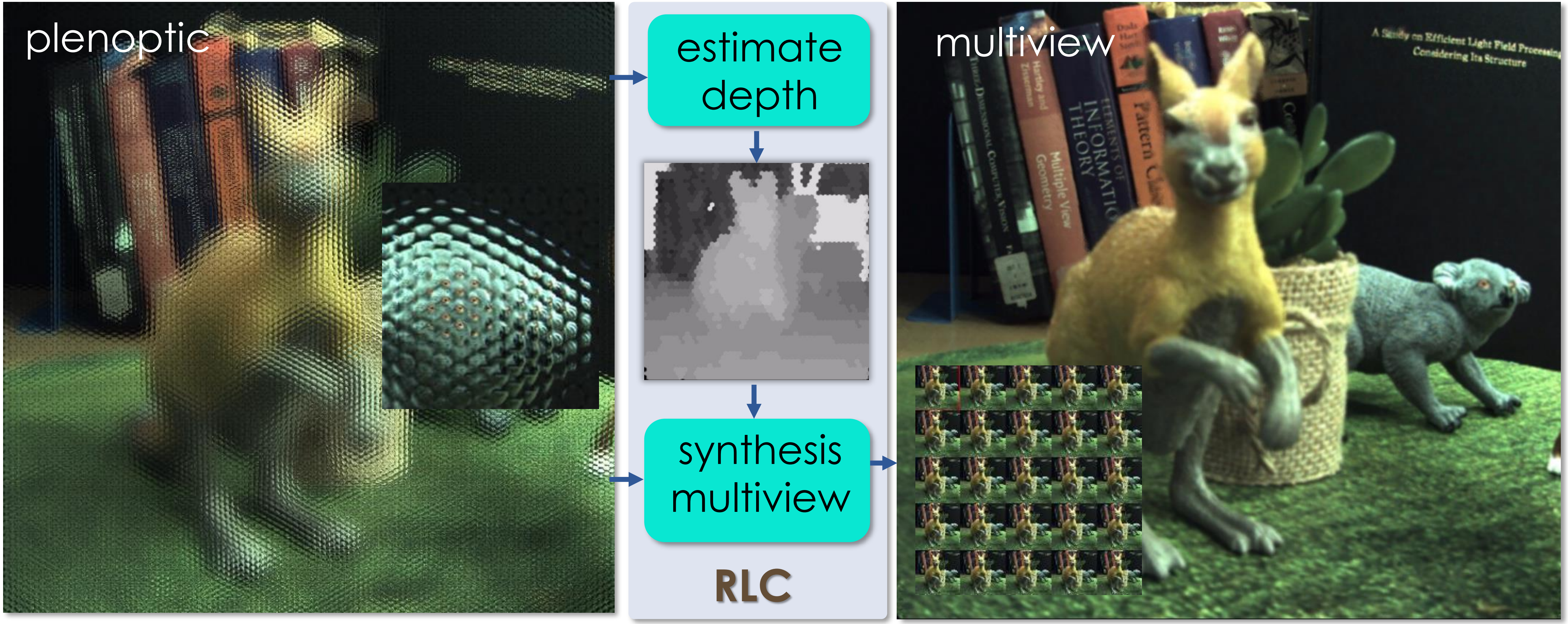
$$E'_k(i,j) = \sum_{s \in \mathcal{S}} \sum_{t \in \mathcal{T}} \sum_{(u,v) \in b(i,j)} \nabla^2 I_k^{(s,t)}(u,v)$$

K(i,j) is the suitable patch size for the *(I,j)*-th microlens, *b(i,j)* contains the pixel indices of patch border at the *(i,j)*-th microlens.
(s,t) and *(u,v)* are view and pixel coordinates.

Reference Lenslet content Convertor (RLC)

MPEG (2017 - 2019), IC3D2018

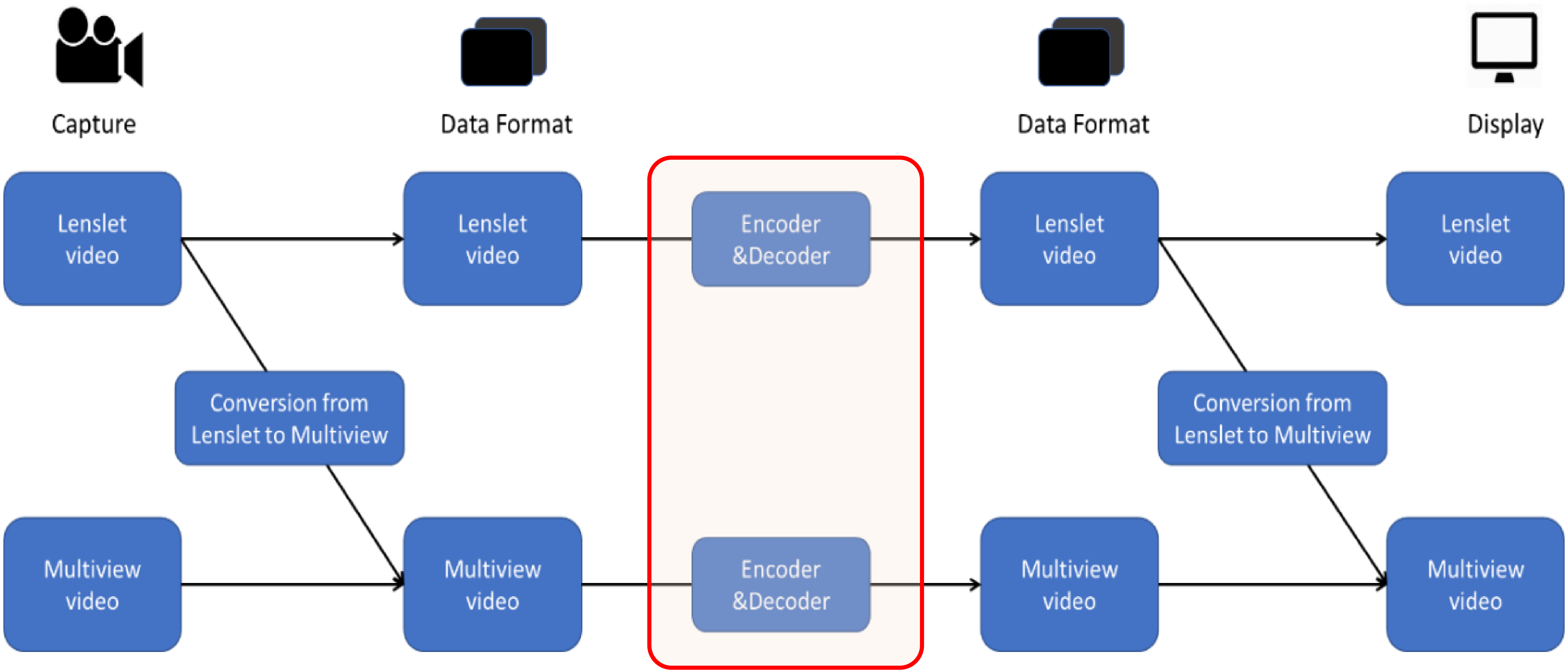
RLC: Plenoptic 2.0 to Multview



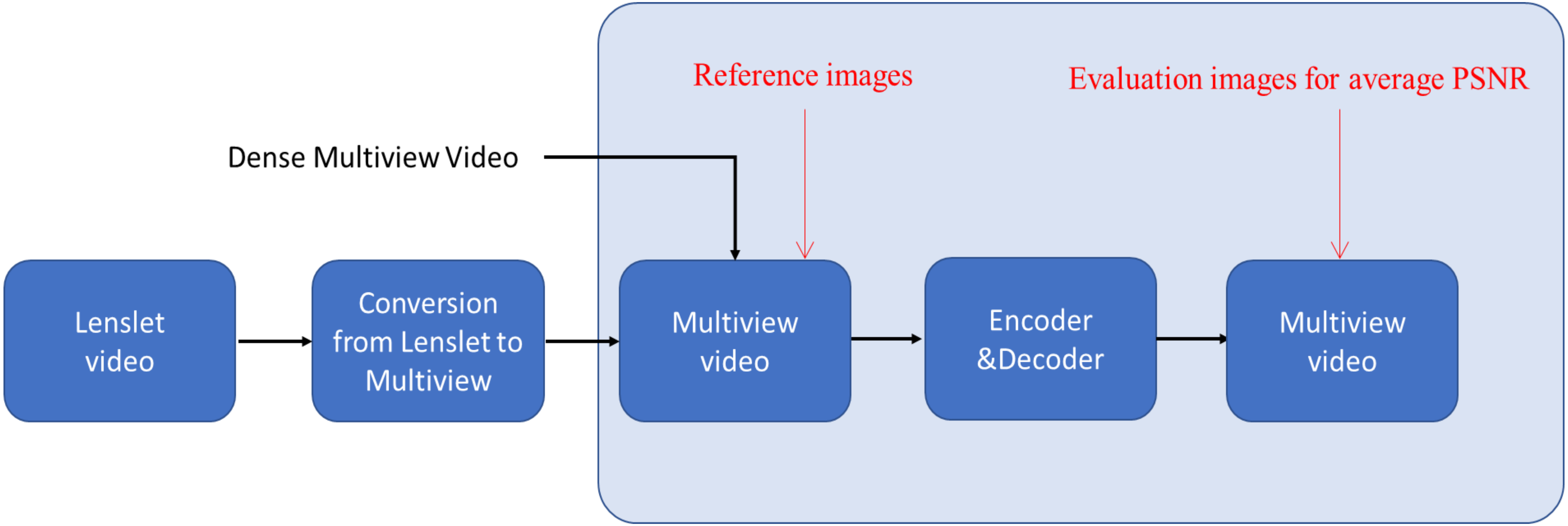
Reference Lenslet content Convertor (RLC)

MPEG (2017 - 2019), IC3D2018

End-to-end system for dense light field

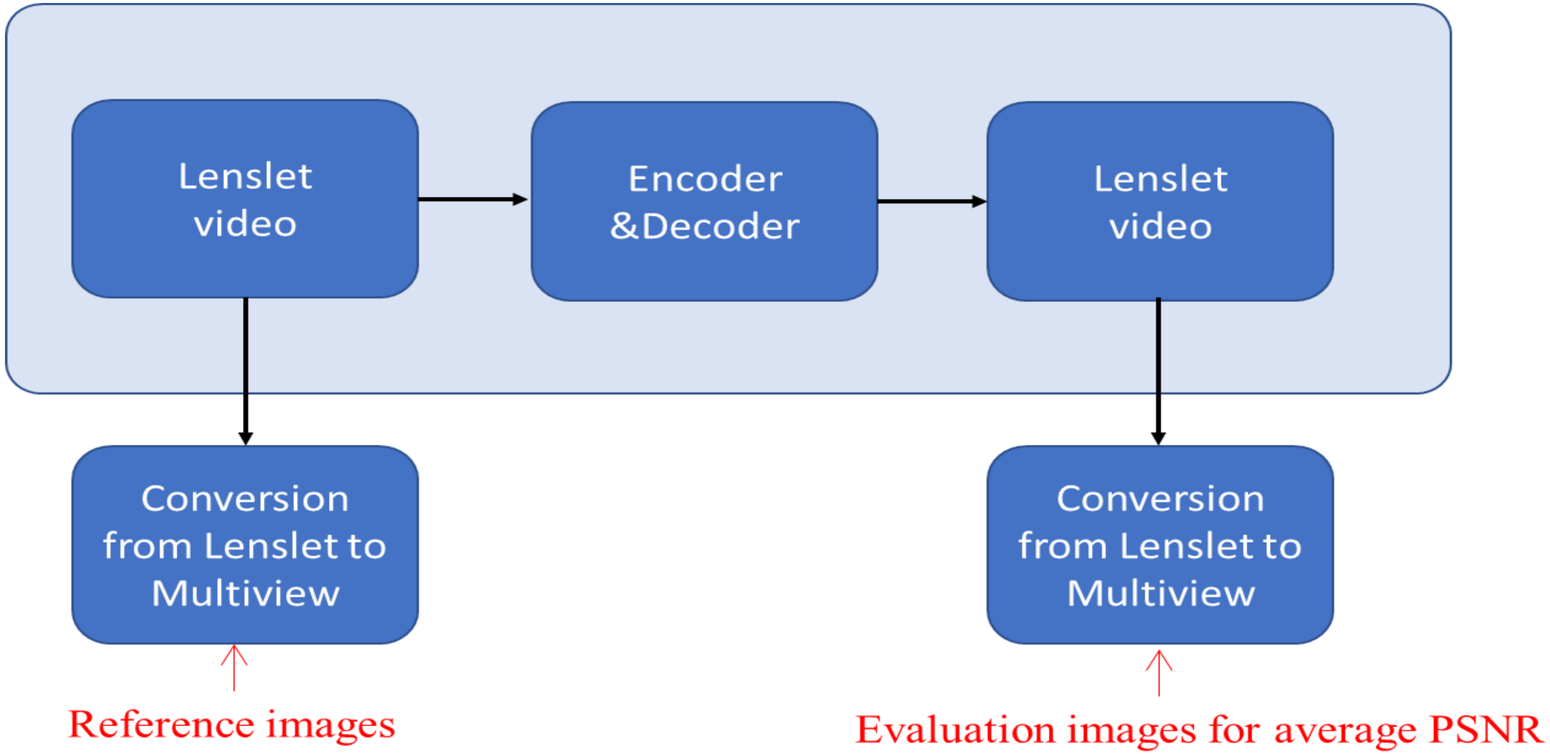


EE_MV: Multiview>Compression>Multiview



Plenoptic 1.0	Plenotic 2.0
HTM16.2	HM-16.9_SCM_8.0

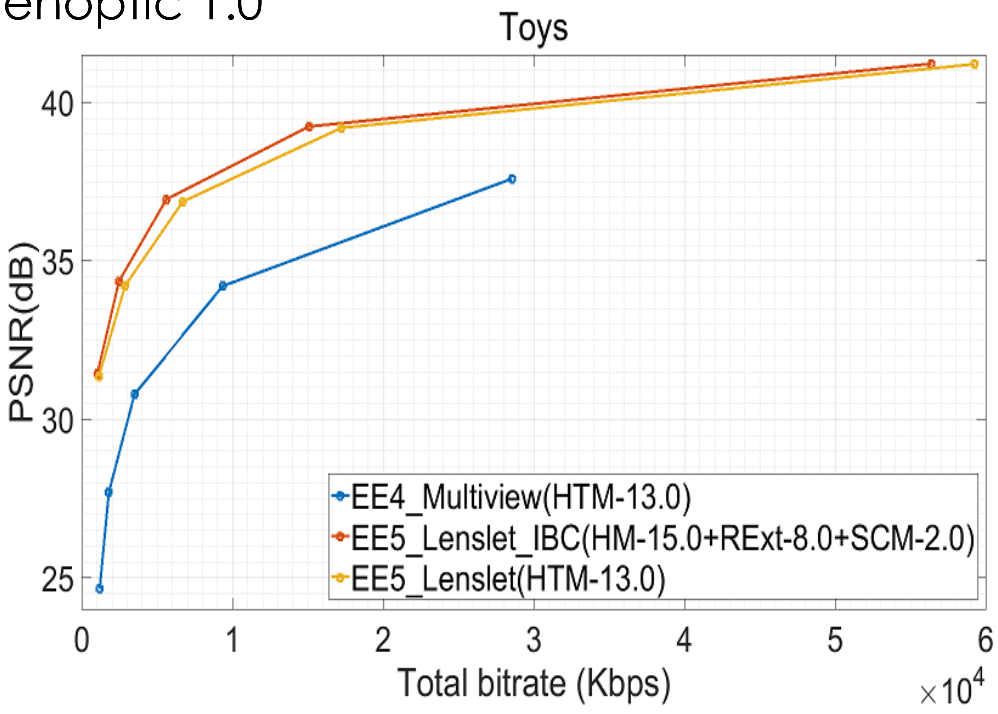
EE_LL: Lenslet>Compression>Lenslet



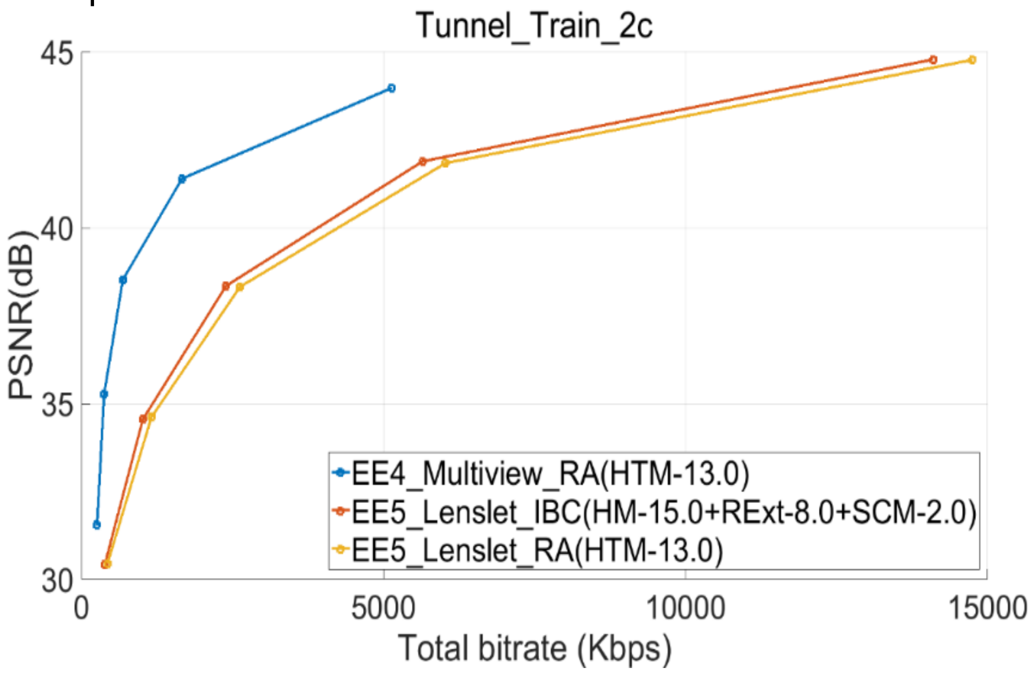
Plenoptic 1.0	Plenotic 2.0
HTM16.2	HM-16.9_SCM_8.0

Compression of plenoptic 1.0 vs plenoptic 2.0

Plenoptic 1.0



Plenoptic 2.0



Plenoptic 1.0	Plenotic 2.0
EE_LL > EE_MV	EE_MV > EE_LL

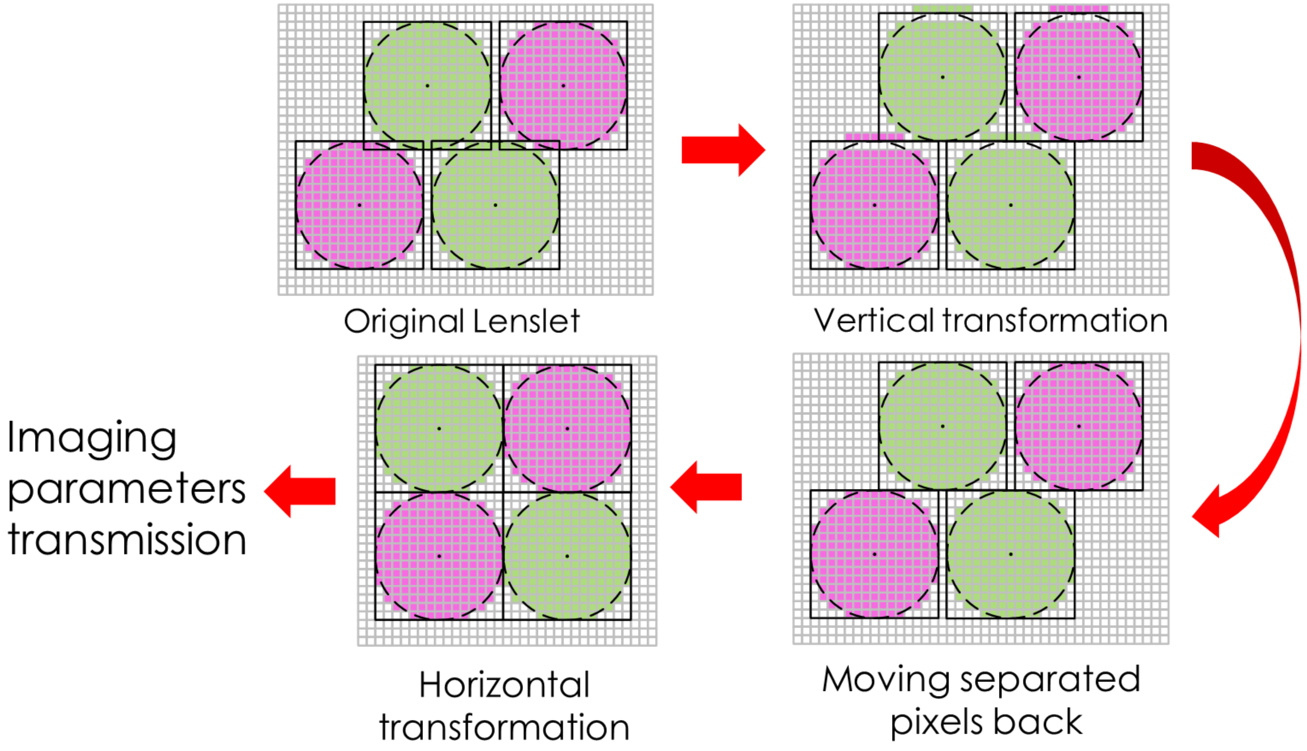
Tsinghua University

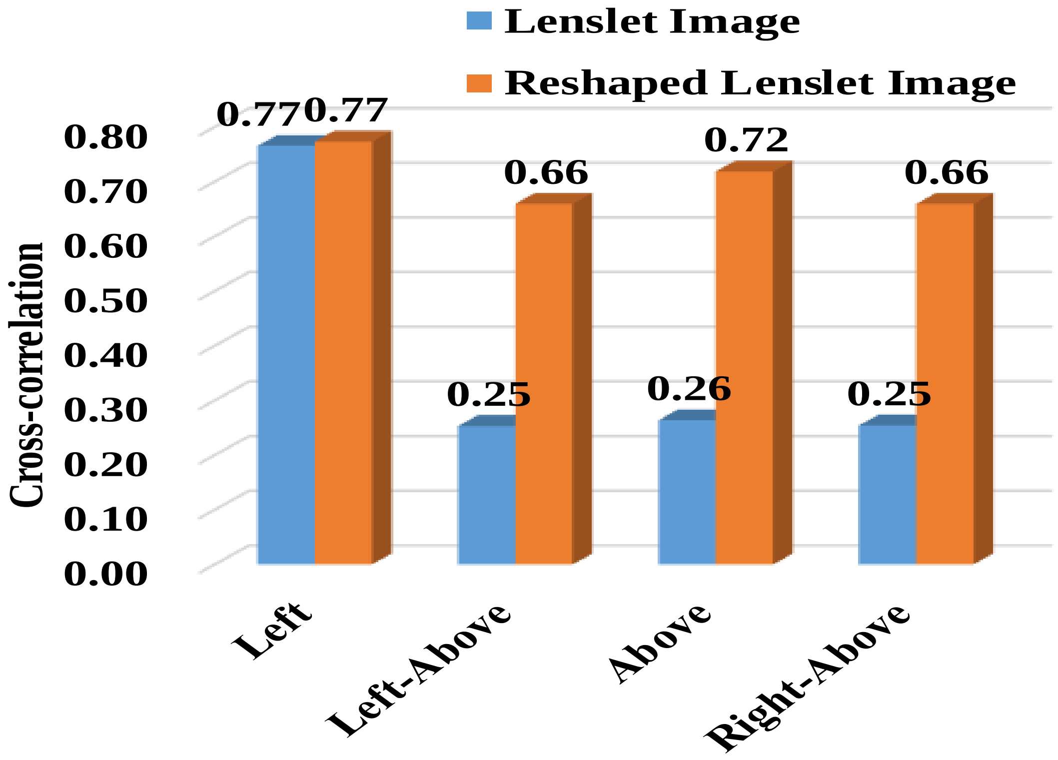
NAGOYA UNIVERSITY

New coding tools for EE_LL

- M44685: Imaging Reshaping (IR)
- M46261: Boundary matching based prediction

Imaging Reshaping (IR)





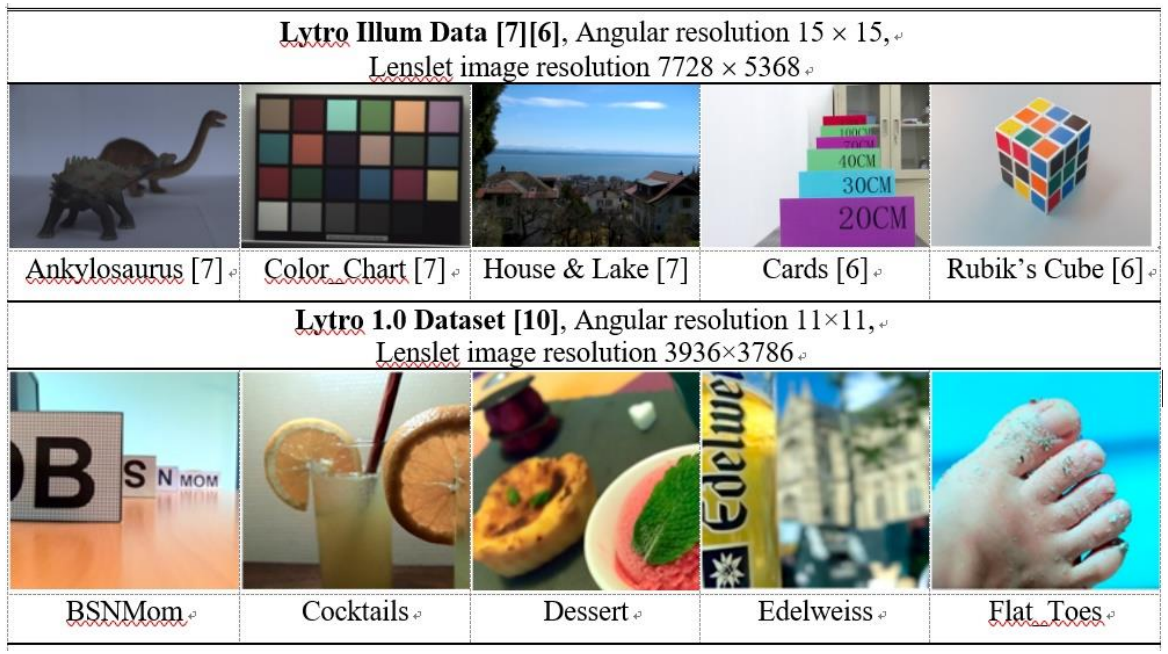
X. Jin, H. Han, and Q. Dai. "Image Reshaping for Efficient Compression of Plenoptic Content." IEEE Journal of Selected Topics in Signal Processing, 11(7): 1173-1186, 2017

Imaging Reshaping

Tsinghua University

Test conditions:

- Reference Software: HM-16.9SCM8.0
- Profile: HEVC Format Range Extension(RExt)
- All Intra Main
- Input Color Format: YUV4:4:4
- QP: 26, 32, 38, 44
- Evaluation: light field performance BD-bitrate

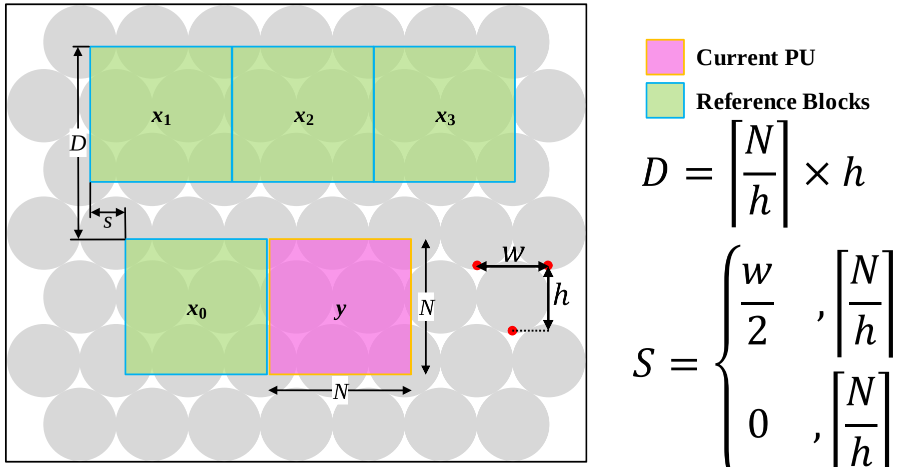


Datasets	Image Name	IR vs. HEVC
Lytro Illum	Ankylosaurus	-30.0%
	Color_Chart	-15.6%
	House & Lake	-34.0%
	Cards	-8.6%
	Rubik's Cube	-13.9%
Lytro 1.0	BSNMom	-23.1%
	Cocktails	-19.8%
	Dessert	-13.6%
	Edelweiss	-4.1%
	Flat_Toes	-20.6%
Average		-18.3%

Boundary matching based prediction

high correlations among the neighboring macropixels → collocated blocks are used to predict the current PU

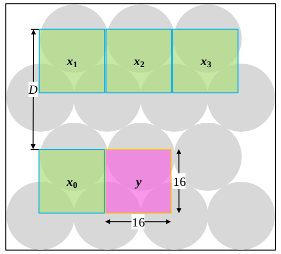
PU size is 32×32 or 64×64



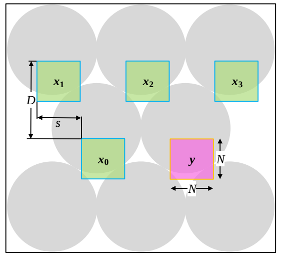
$$D = \left\lceil \frac{N}{h} \right\rceil \times h$$

$$S = \begin{cases} \frac{w}{2} & , \left\lceil \frac{N}{h} \right\rceil \in O \\ 0 & , \left\lceil \frac{N}{h} \right\rceil \in E \end{cases}$$

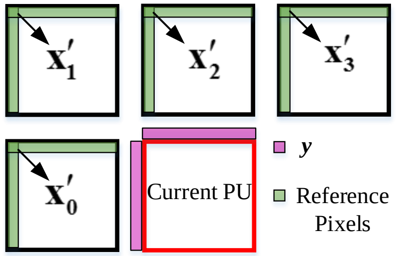
PU size is 16×16



PU size is 4×4 or 8×8

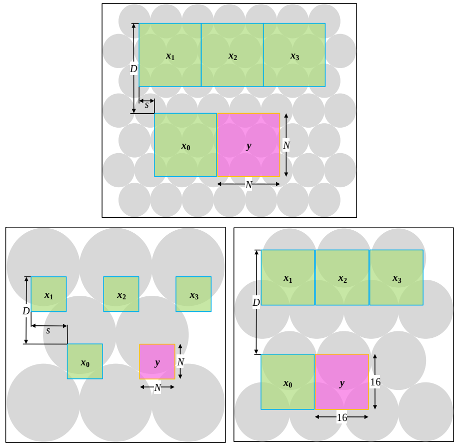


Weights determination



$$\begin{aligned} & minimize\ \|Xw - y\|_2^2 \\ & subject\ to\ \mathbf{1}^T \boldsymbol{w} = \mathbf{1} \\ & \qquad\qquad\quad \boldsymbol{w} \geq \mathbf{0} \end{aligned}$$

Weighted prediction



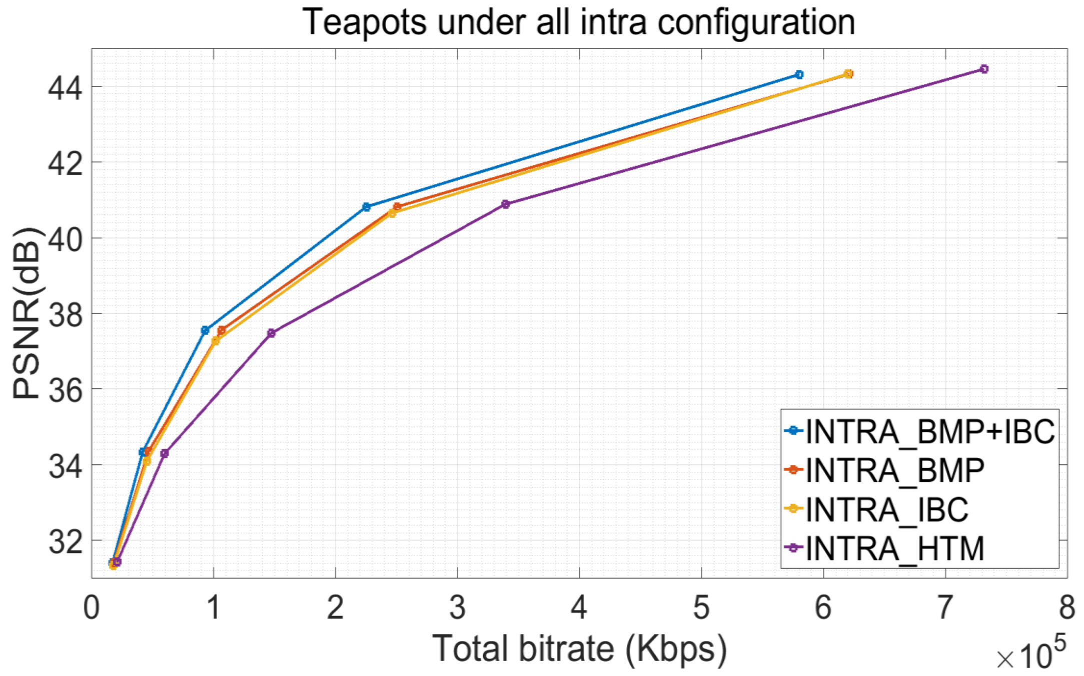
$$y' = w_0 x_0 + w_1 x_1 + w_2 x_2 + w_3 x_3$$

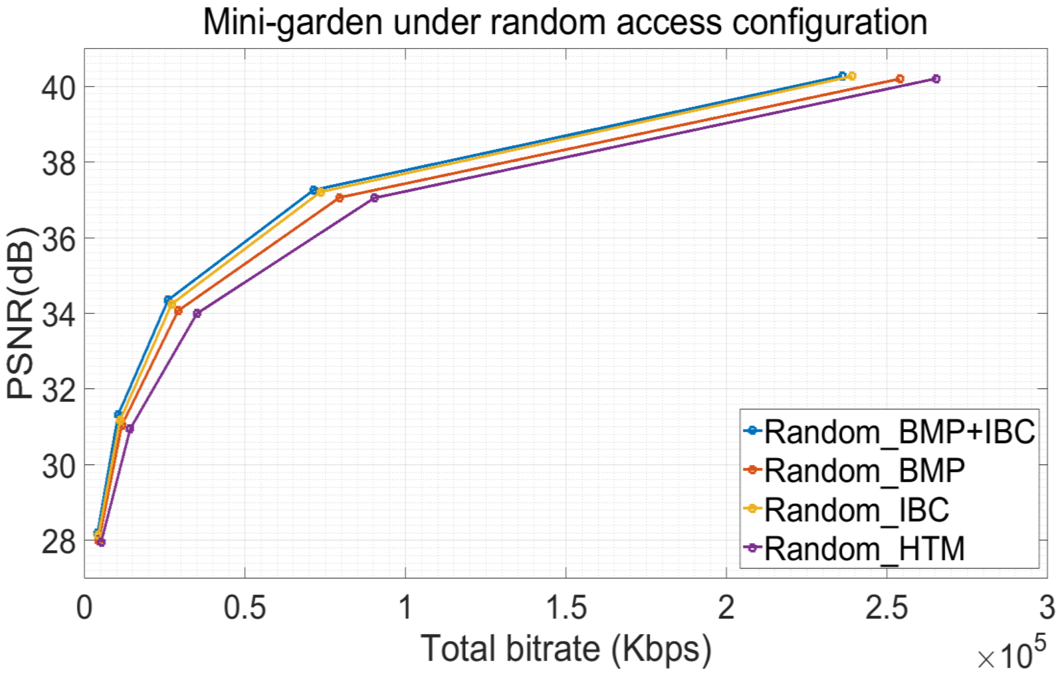
M46261: Boundary matching based prediction

Tsinghua University

Test Conditions

- Test sequences: Teapots, Mini-garden
- Total frames: 60 frames
- Resolution: 8656(H) × 6074(V)
- Anchor: HTM-HM mode
- Compared methods: IBC, BMP, IBC+BMP
- Configuration: all intra, random access
- QP: 24, 30, 36, 42 and 48
- Evaluation: RD curve and BD-Bitrate





Compression Performance on Plenoptic 1.0

Tsinghua University 清華大學

BD-Bitrate result for QP 30, 36, 42 and 48 under the all Intra configuration

	IBC vs. HTM	BMP vs. HTM	IBC+BMP vs. HTM	IBC+BMP vs. IBC	Codec version
Toys	-16.44%	-21.17%	-23.33%	-8.20%	HM15.0+RExt -8.0+SCM-2.0
Trees	-3.69%	-4.14%	-5.36%	-1.74%	
Avg.	**-10.07%**	**-12.66%**	**-14.35%**	**-4.97%**	
Teapots	-22.28%	-24.78%	-32.55%	-13.21%	HM-16.9_SCM_8.0
Mini-garden	-30.30%	-31.98%	-38.30%	-11.54%	
Avg.	-26.29%	-28.38%	-35.43%	-12.38%	
Avg. (All)	**-18.18%**	**-20.52%**	**-24.89%**	**-8.67%**	

BD-Bitrate result for QP 30, 36, 42 and 48under the random access configuration

	IBC vs. HTM	BMP vs. HTM	IBC+BMP vs. HTM	IBC+BMP vs. IBC	Codec version
Toys	-16.62%	-19.25%	-22.36%	-6.84%	HM-15.0+RExt-8.0+SCM-2.0
Trees	-35.58%	-1.50%	-36.03%	-0.72%	
Avg.	**-26.10%**	**-10.38%**	**-29.20%**	**-3.78%**	
Teapots	-21.76%	-11.93%	-27.16%	-6.78%	HM-16.9_SCM_8.0
Mini-garden	-26.52%	-18.20%	-32.26%	-7.72%	
Avg.	**-24.14%**	**-15.07%**	**-29.71%**	**-7.25%**	
Avg. (All)	**-25.12%**	**-12.72%**	**-29.45%**	**-5.52%**	

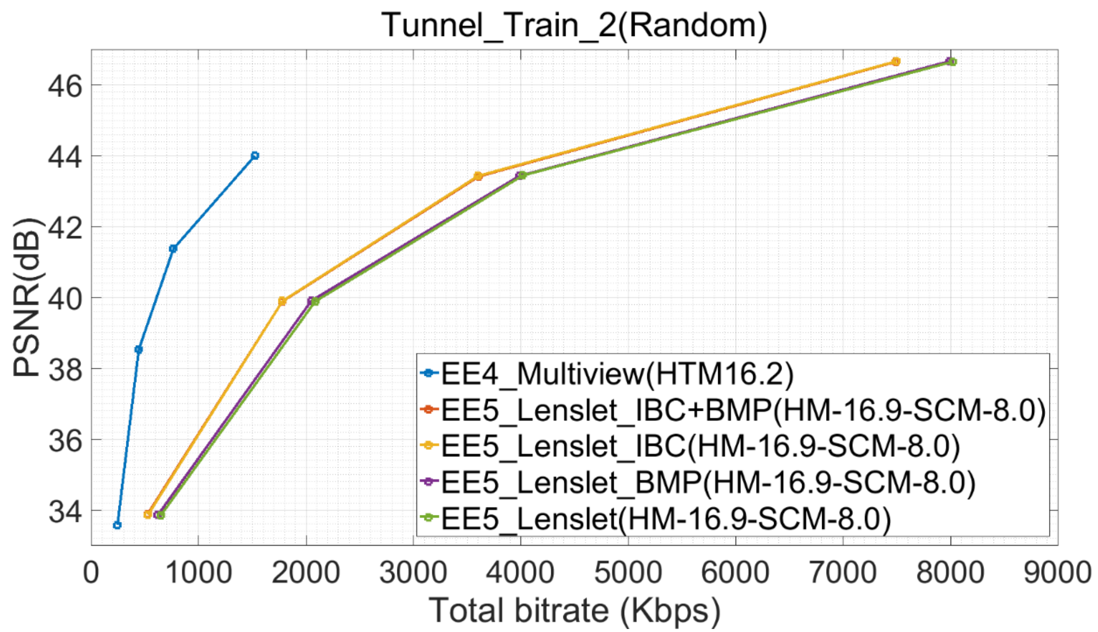
Compression Performance on Plenoptic 2.0

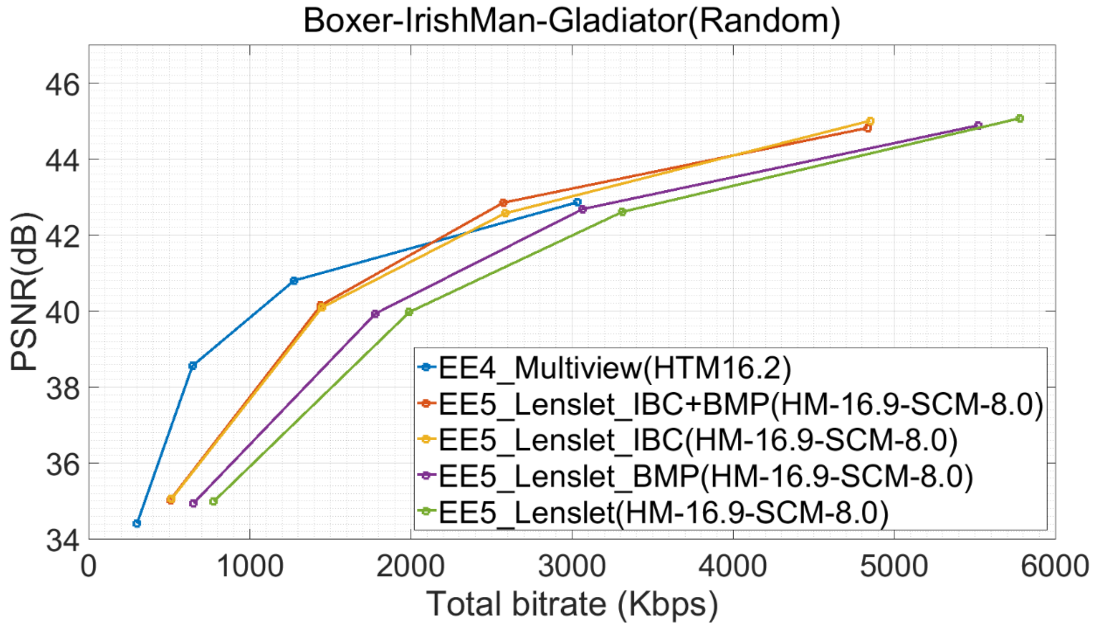
Tsinghua University

Test Conditions:

- Test sequences: Tunnel_Train_2, Chess-Pieces, Boxer-IrishMan-Gladiator and ChessPieces-MovingCamera.
- Total frames: 100 frames
- Resolution: R5:2048(H) × 2048(V),R8:3840 (H) × 2160(V),

- Anchor: HTM-HM mode
- Compared methods: IBC, BMP, IBC+BMP
- Configuration: all intra, random access
- QP: 28, 33, 38 and 46
- Evaluation: BD-Bitrate and RD curve





Compression Performance on Plenoptic 2.0

Tsinghua University

BD-Bitrate result for QP 28,33,38 and 46 under all intra configuration

	IBC vs. HTM	BMP vs. HTM	IBC+BMP vs. HTM	IBC+BMP vs. IBC	Codec version
Tunnel_Train_2	-16.15%	-3.61%	-16.54%	-0.52%	HM-16.9+SCM-8.0
Chess-Pieces	-42.99%	-20.77%	-43.42%	-0.85%	
Boxer-IrishMan-Gladiator	-37.84%	-15.95%	-38.55%	-1.17%	
ChessPieces-MovingCamera	-24.81%	-22.83%	-29.57%	-6.54%	
Avg. (All)	**-30.45%**	**-15.79%**	**-32.02%**	**-2.27%**	

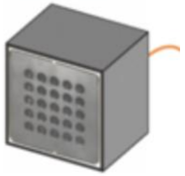
BD-Bitrate result for QP 28,33,38 and 46 under random access configuration

	IBC vs. HTM	BMP vs. HTM	IBC+BMP vs. HTM	IBC+BMP vs. IBC	Codec version
Tunnel_Train_2	-13.66%	-2.01%	-13.63%	0.03%	HM-16.9+SCM-8.0
Chess-Pieces	-30.22%	-13.89%	-30.67%	-0.63%	
Boxer-IrishMan-Gladiator	-27.45%	-10.02%	-29.24%	-2.16%	
ChessPieces-MovingCamera	-17.29%	-13.60%	-19.22%	-2.41%	
Avg. (All)	**-22.16%**	**-9.88%**	**-23.19%**	**-1.29%**	

Light Field Capture to Display System

NAGOYA UNIVERSITY

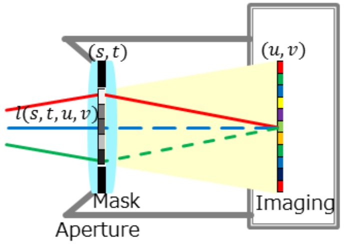
ViewPLUS
ProFUSION 25

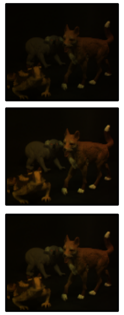


Multi-view camera

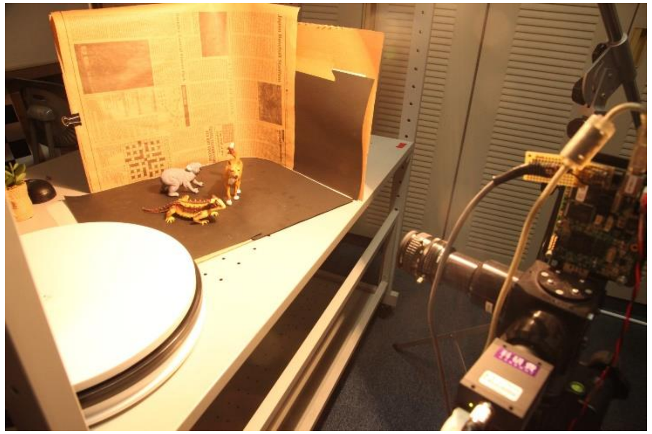
Acquired images

Coded Aperture Camera

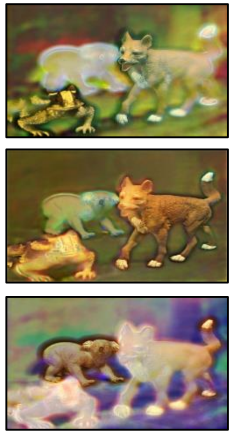


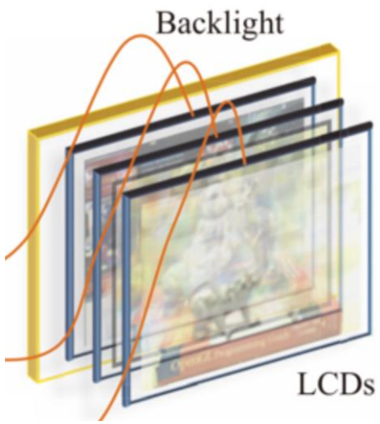


Acquired images



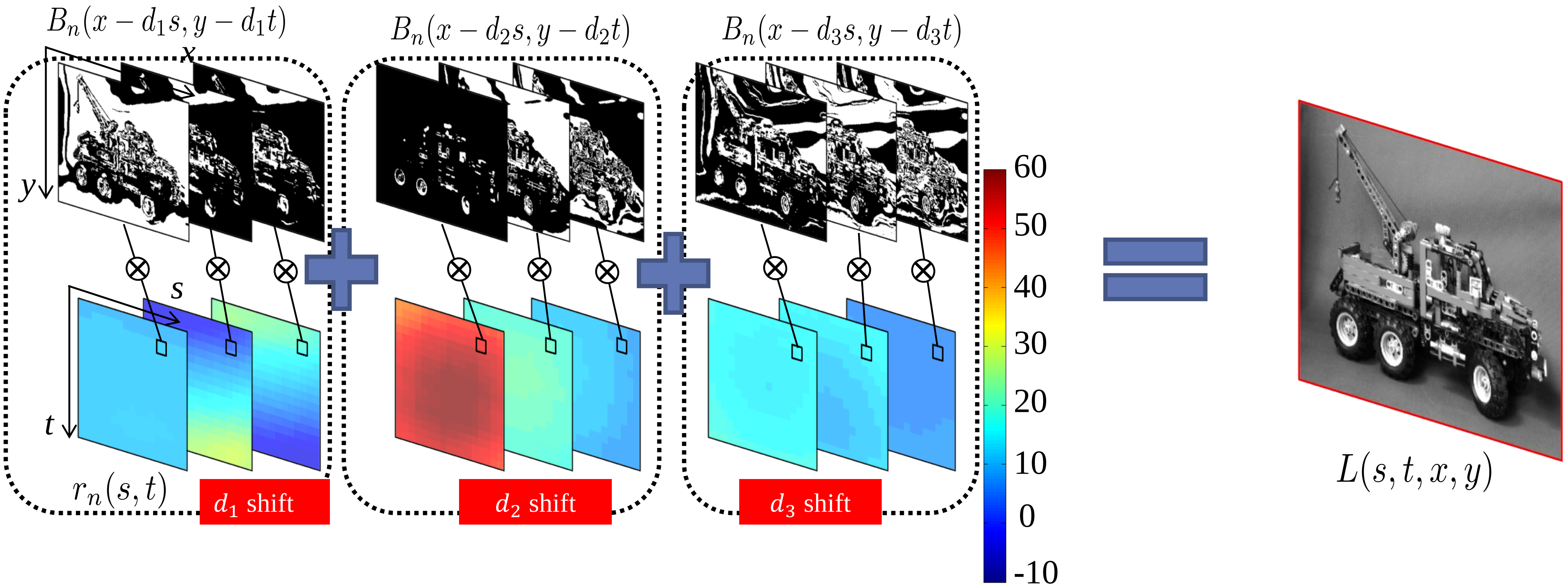
Layered patterns



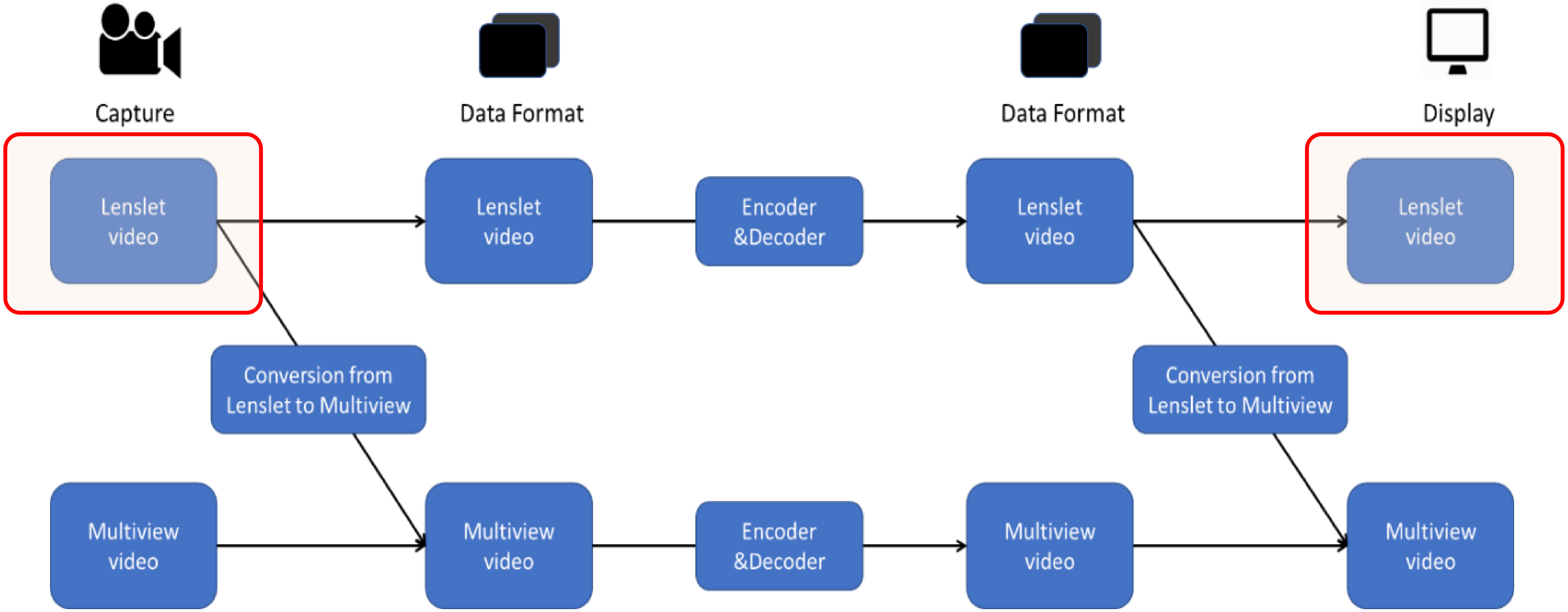


Layer type 3D display

Light Field Representation based on Weighted Sum of Binary Patterns



End-to-end system for dense light field



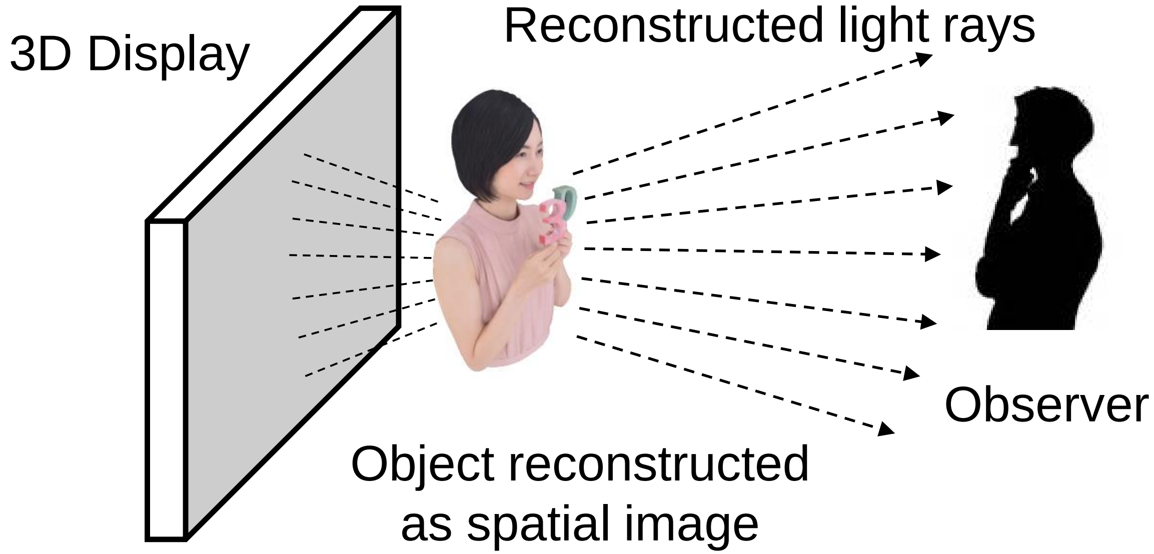
3D TV Based on Spatial Imaging

NHK Science & Technology Research Laboratories

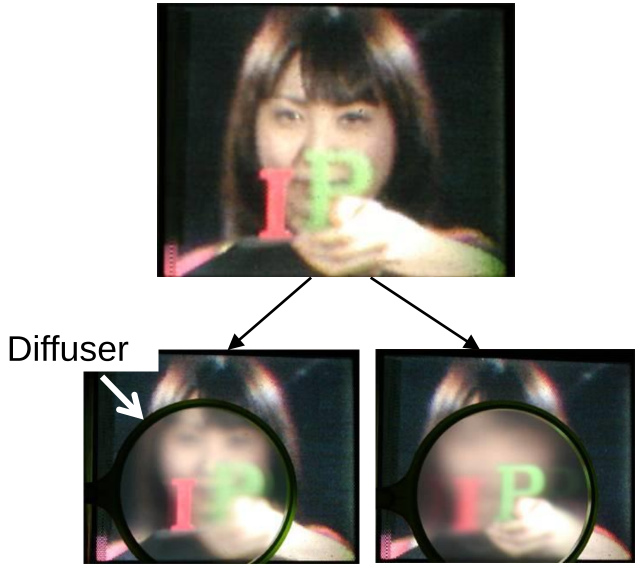
Requirements for 3D TV

- No special glasses
- Full parallax
- Natural 3D image

Spatial Imaging



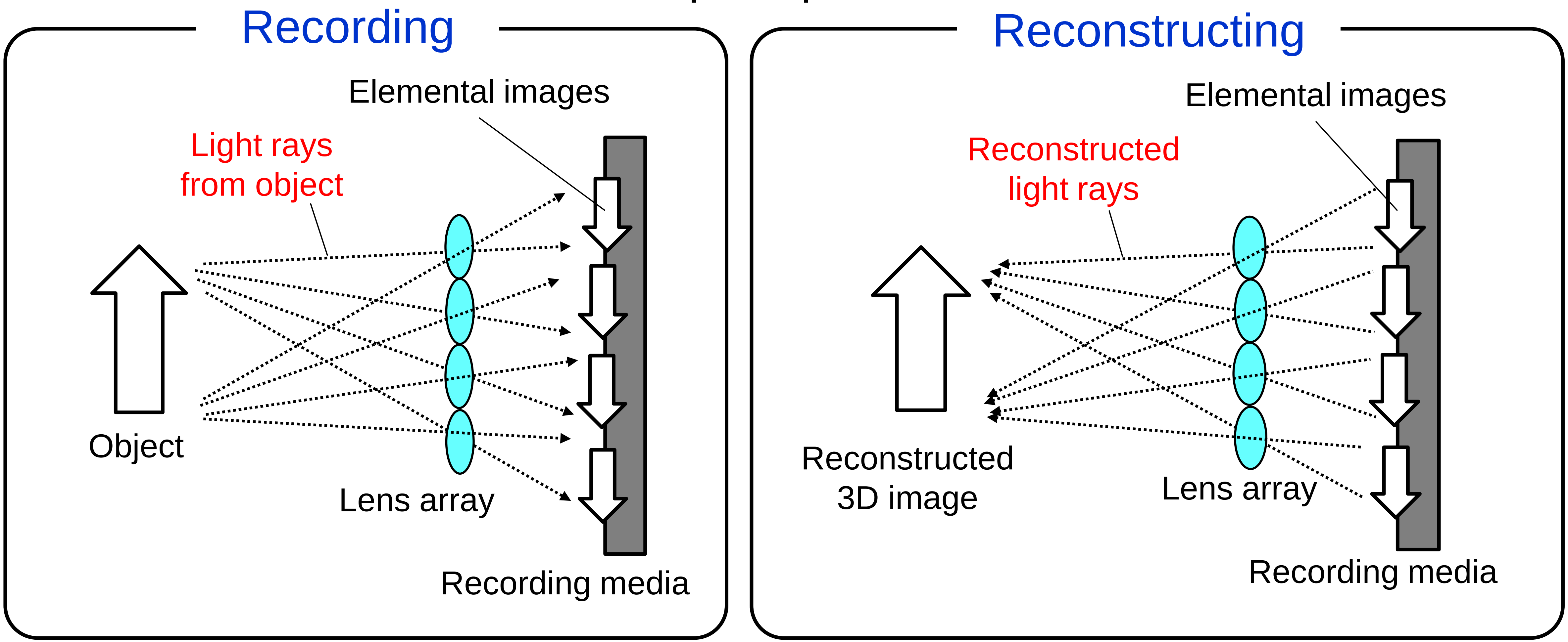
3D image based on Integral Photography



Distance from lens array to 3D image

Integral photography

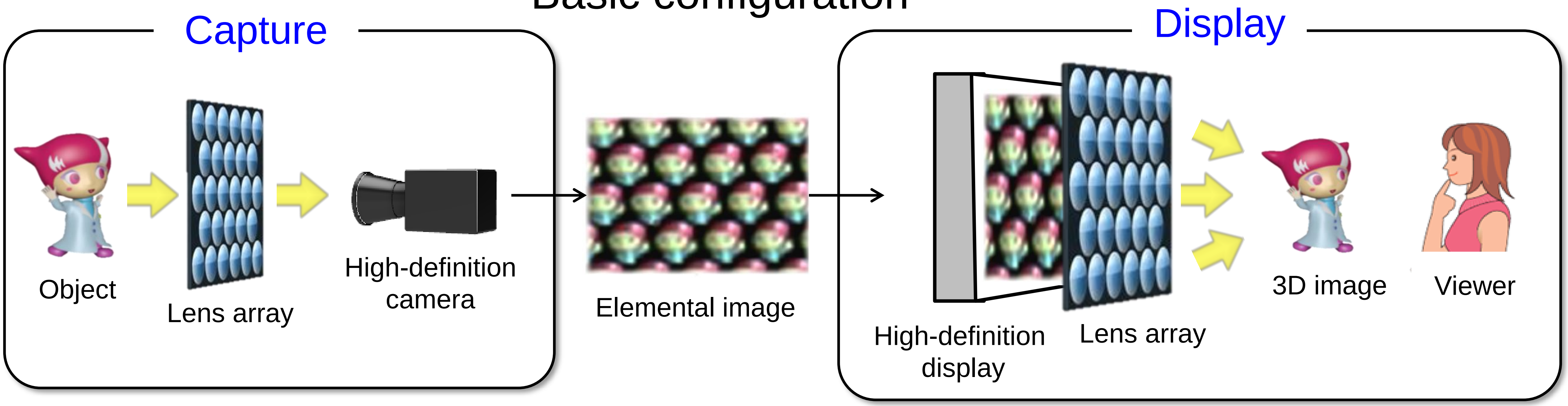
Basic principle



M. G. Lippmann (1908)

Integral 3D TV

Basic configuration



Features Real-time capture and display of moving 3D images

Real objects (not computer graphics) are captured and displayed

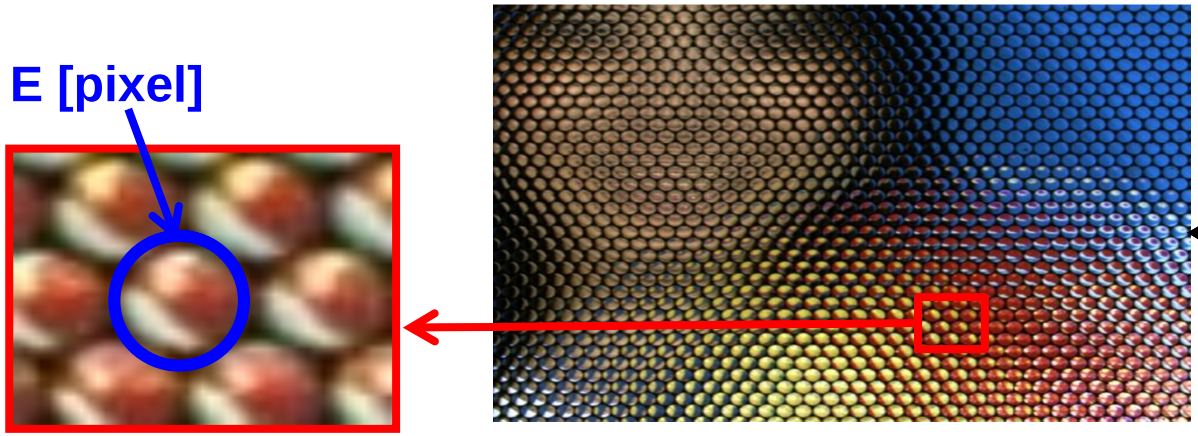
Full-parallax images

Problem Integral 3D system requires huge number of pixels

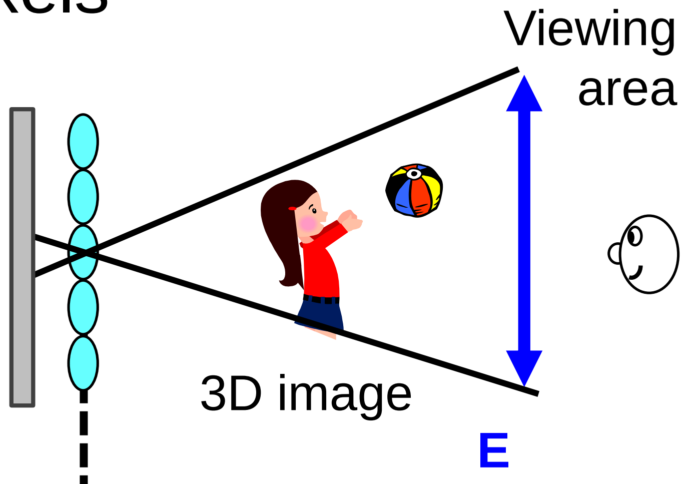
Science & Technology Research Laboratories

Problem with Integral 3D system

- Integral 3D system requires huge number of pixels

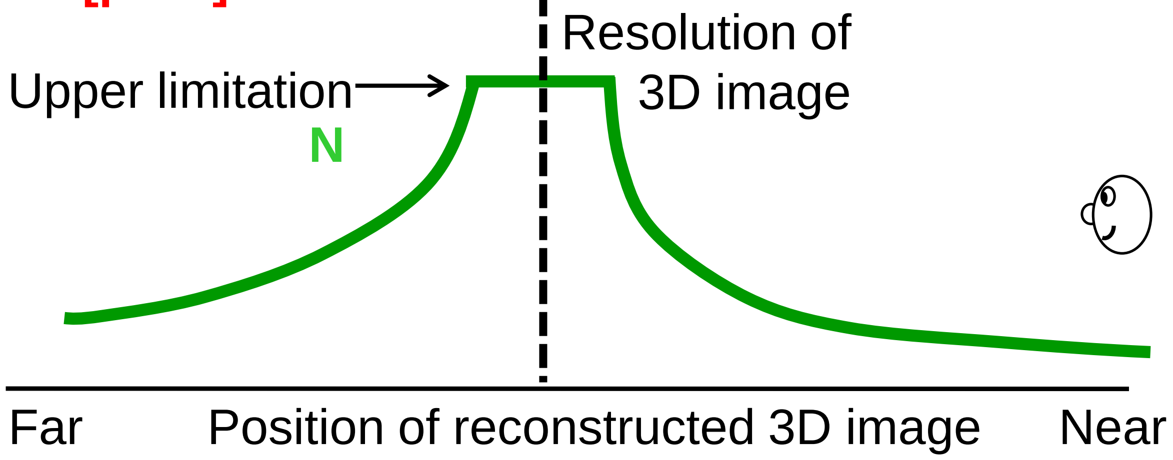


Elemental images M [pixel]



M=N X E

- M: Total pixels in elemental images
- E : Number of pixels in one elemental image
- N : Number of elemental images (lenses)



Concluding Remark

- Lenslet data
  - integral display, 3D modeling, refocusing , multiview rendering.
  - Multimedia, medical applications

- Compression efficiency of lenslet data
  - Utilize the structure of lenslet data for inter/intra predictions
  - Novel image transform and entropy coding methods
  - Utilize machine learning tools

- New compression method for lenslet
  - leads to improvement over the existing standards.

Acknowledgement

清華大學

Tsinghua University

NAGOYA UNIVERSITY

NHK

ULB

INRIA

NICT

National Institute of Information and Communications Technology