Multiview Depth Image Enhancement for Free-viewpoint Television

Pravin Kumar Rana, Zhanyu Ma, Jalil Taghia, and Markus Flierl

Opponents: Nasser Mohammadiha, Haopeng Li and, Jalil Taghia

Internal seminar, Communication Theory, KTH

June 27, 2013

Motivation & Background

Conventional Television

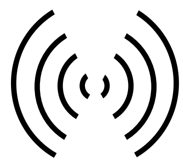
User

Display

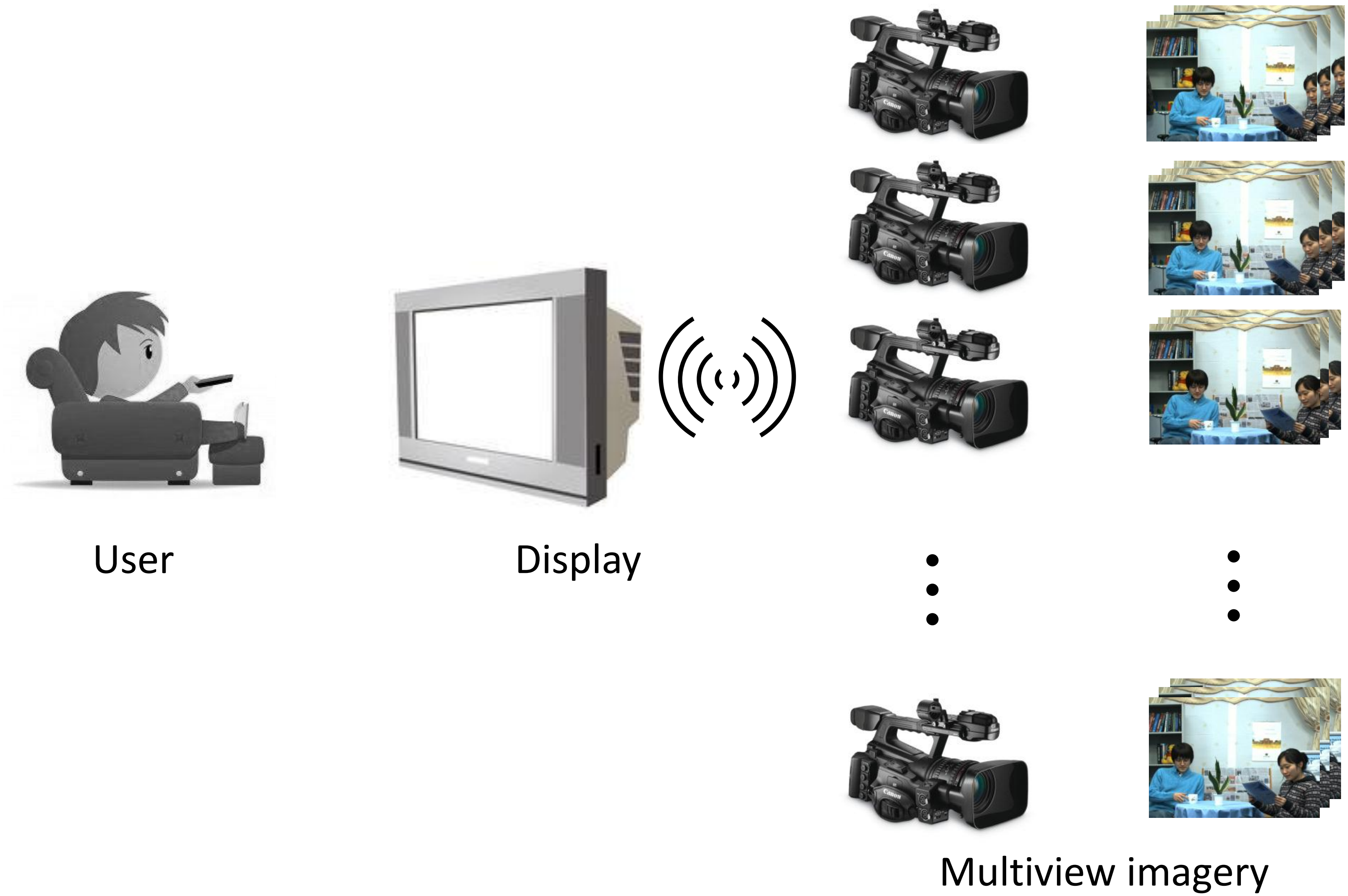
Conventional Television

User

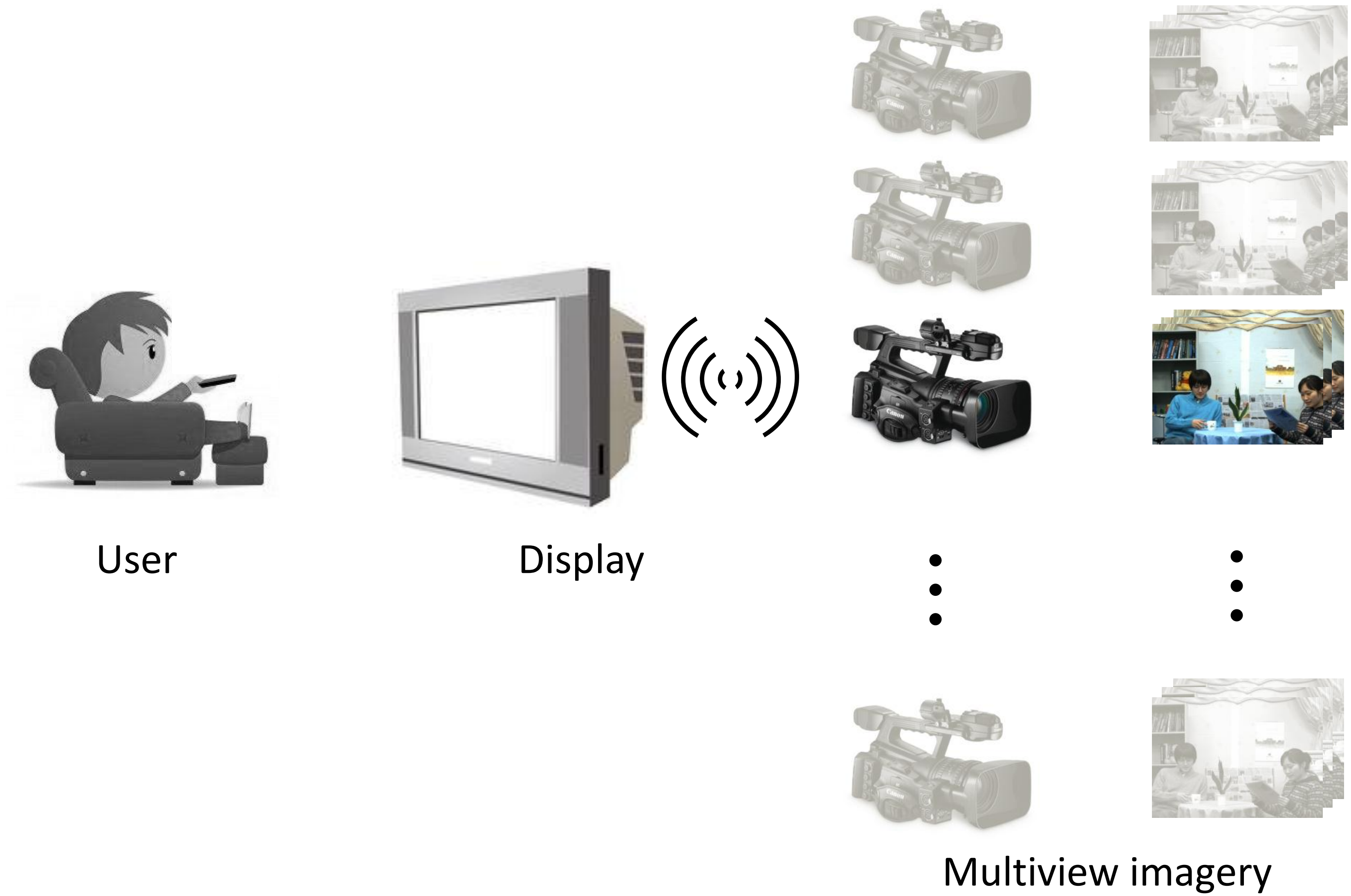
Display



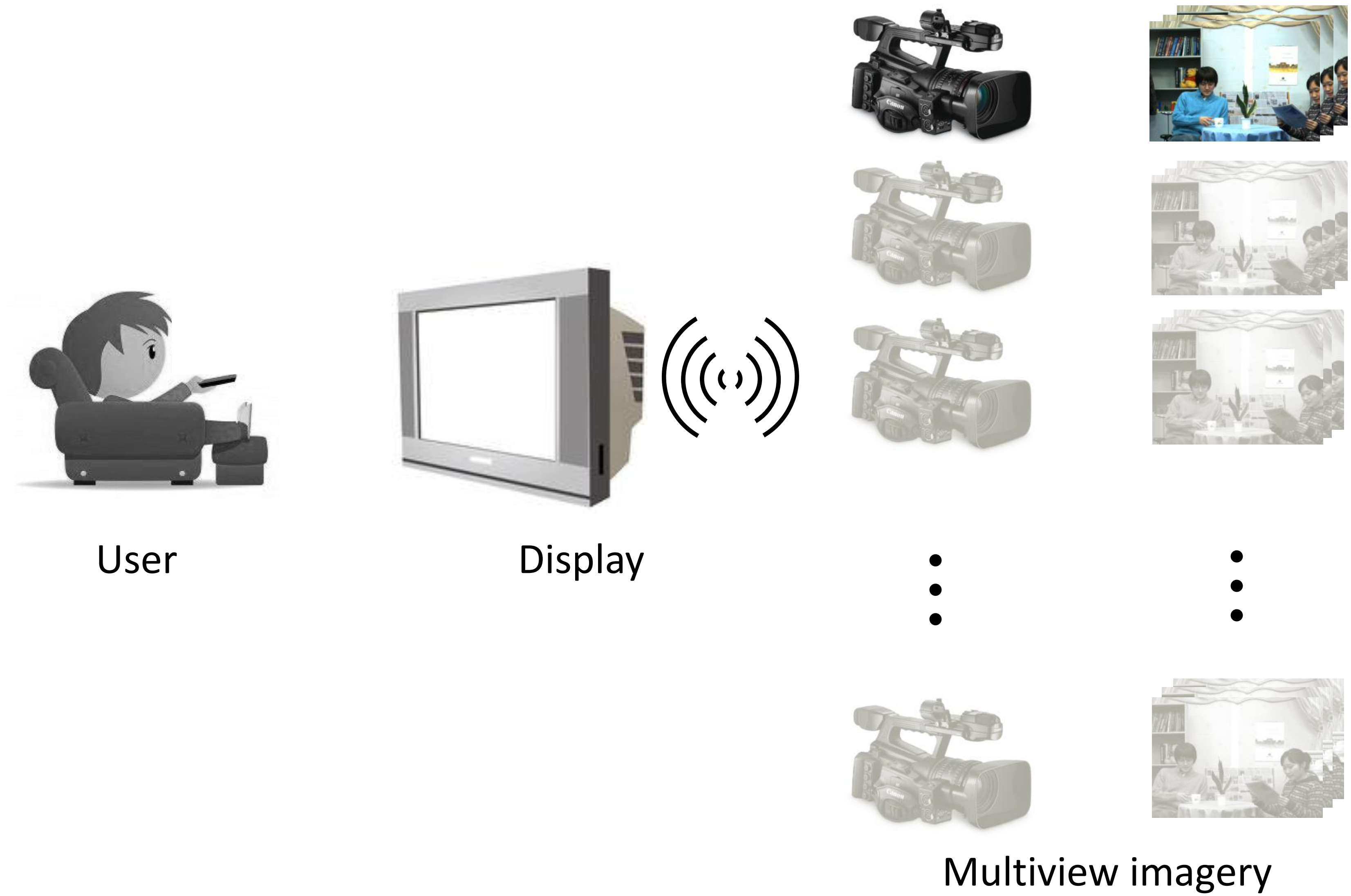
Free-viewpoint Television



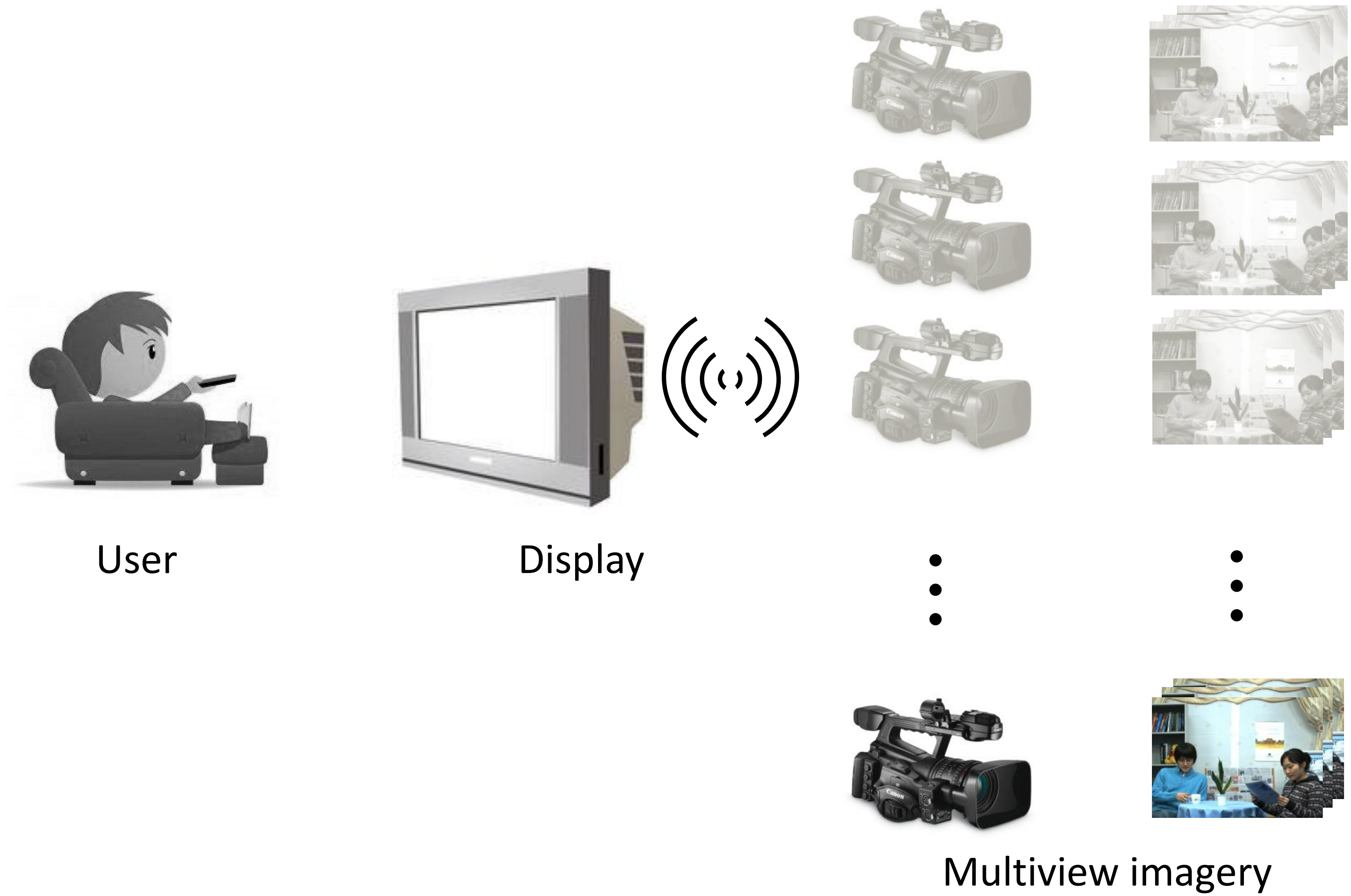
Free-viewpoint Television



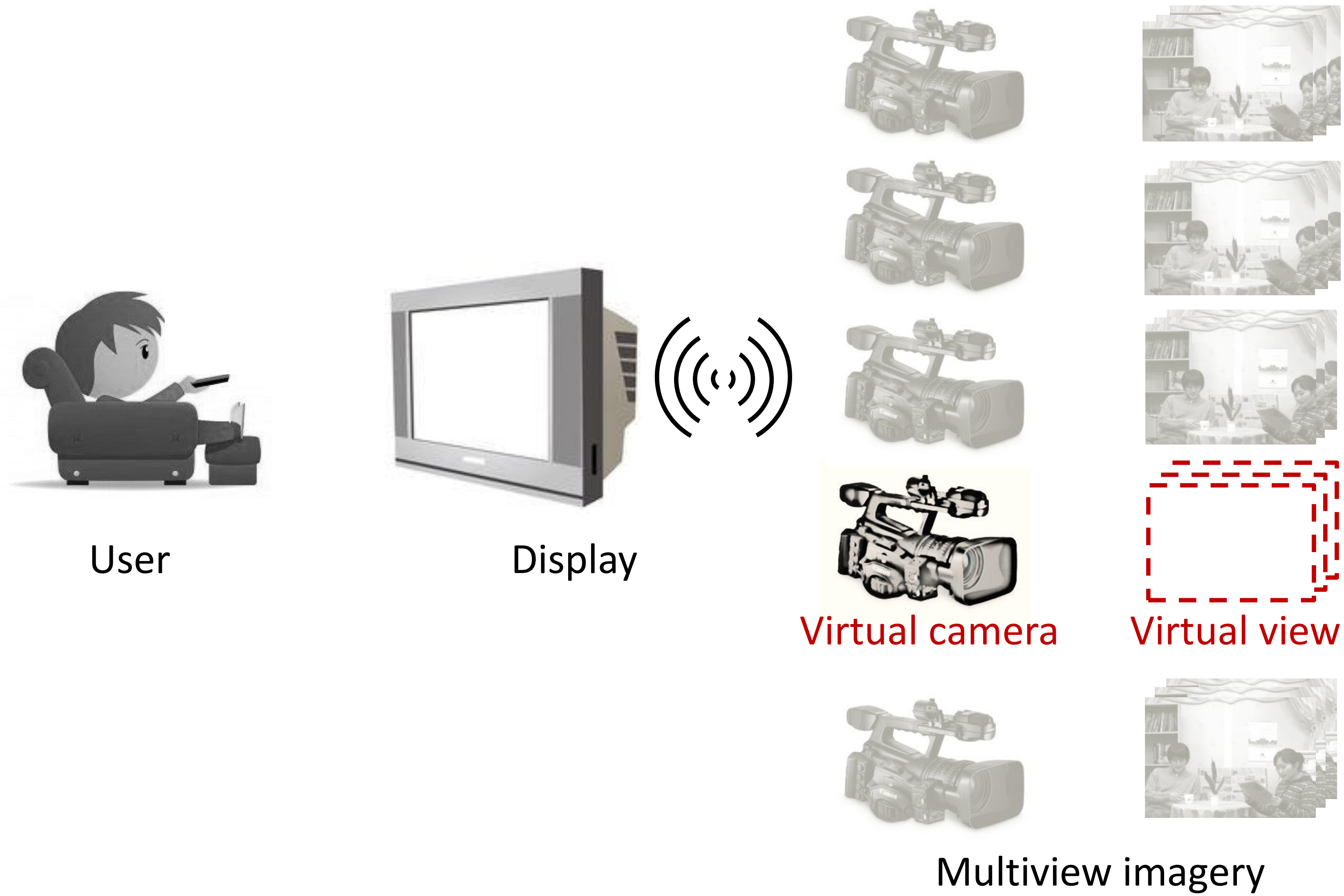
Free-viewpoint Television



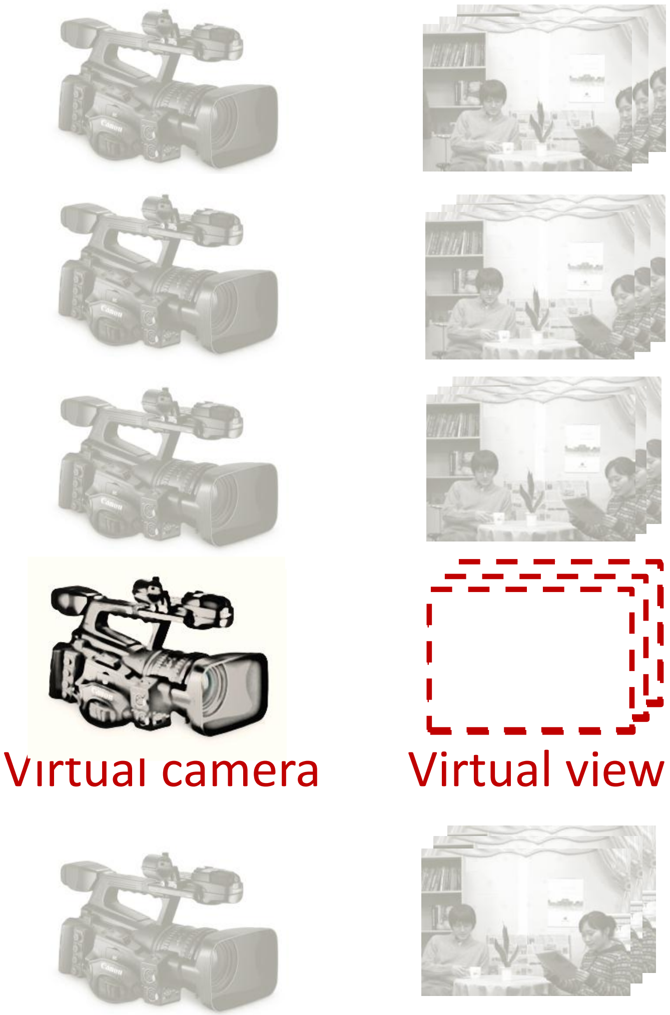
Free-viewpoint Television



Free-viewpoint Television



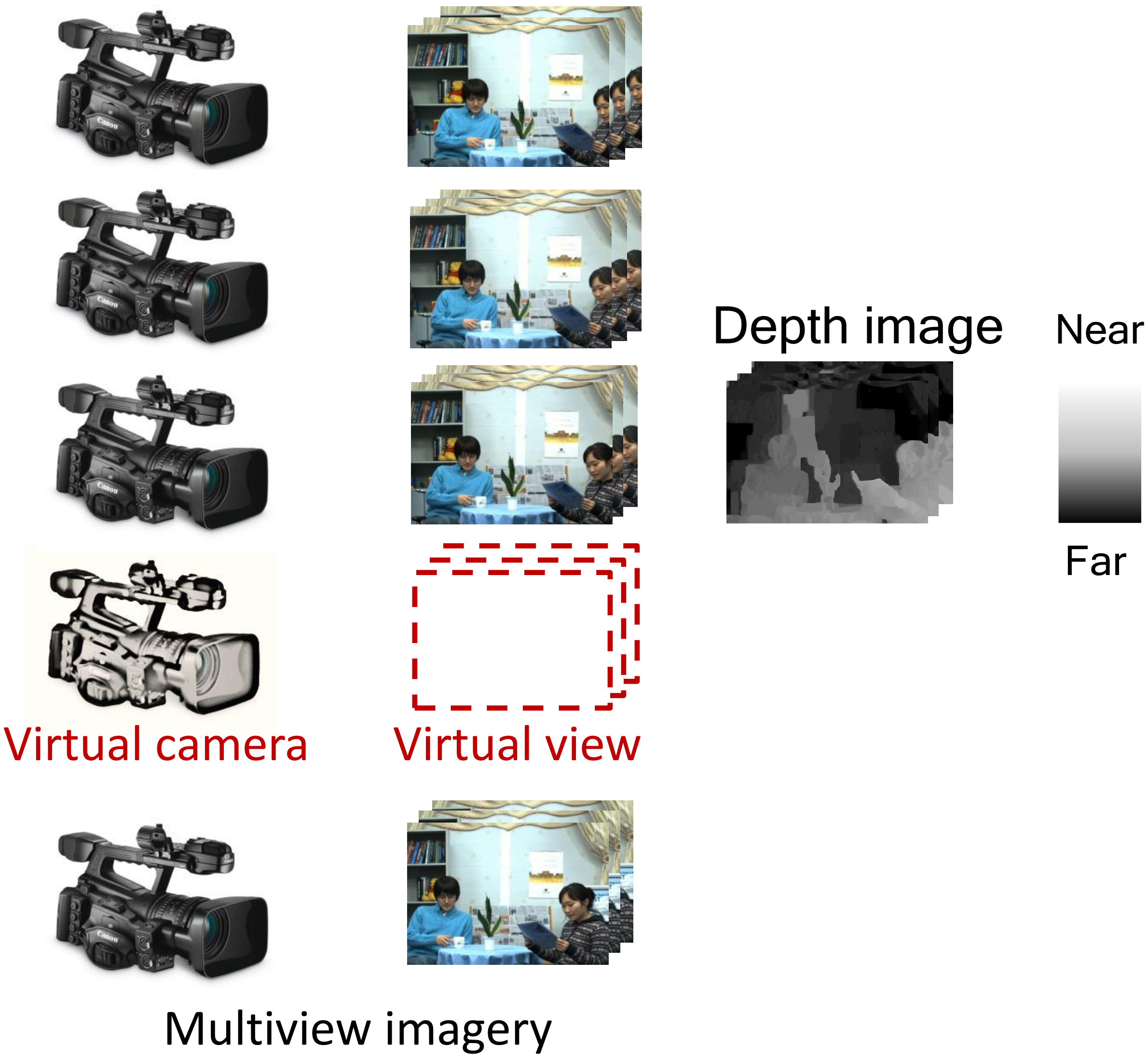
Virtual View Synthesis



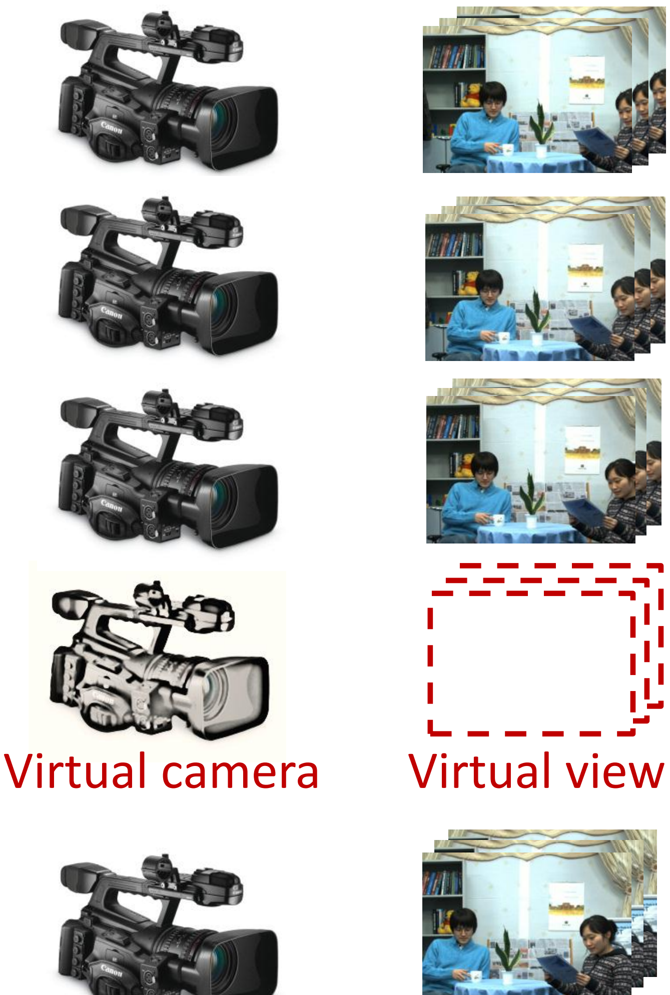
Multiview imagery

Depth Image Based Rendering

Depth Image Based Rendering



Depth Image Based Rendering



Multiview imagery

- Depth pixels represent shortest distance between object points and the camera plane
- To be estimated from multiview imagery

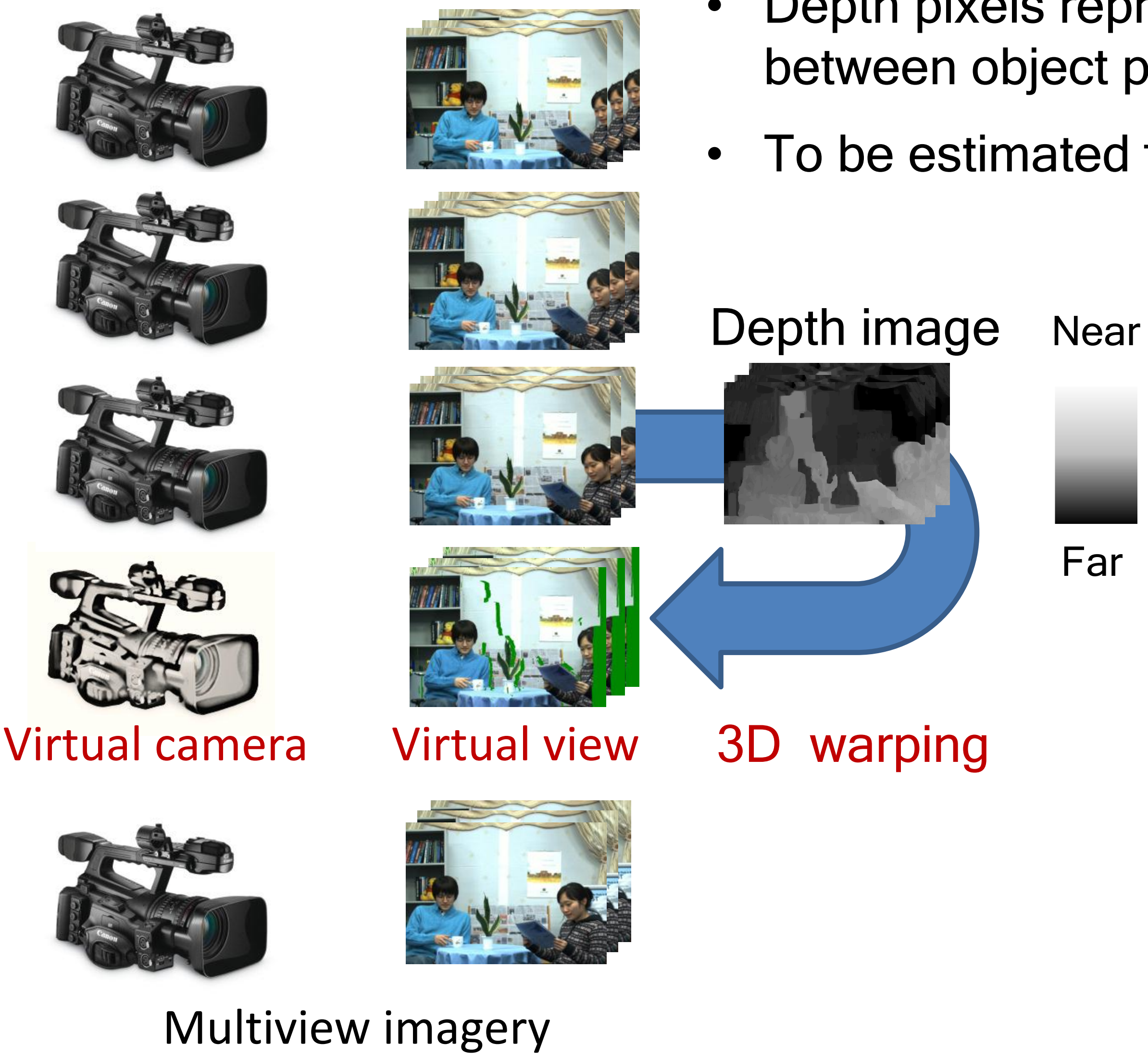
Depth image

Near

Far

Depth Image Based Rendering

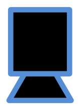
- Depth pixels represent shortest distance between object points and the camera plane
- To be estimated from multiview imagery



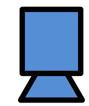
Depth Image Based Rendering

3D Warping

Physical camera
parameters



Virtual camera
parameters

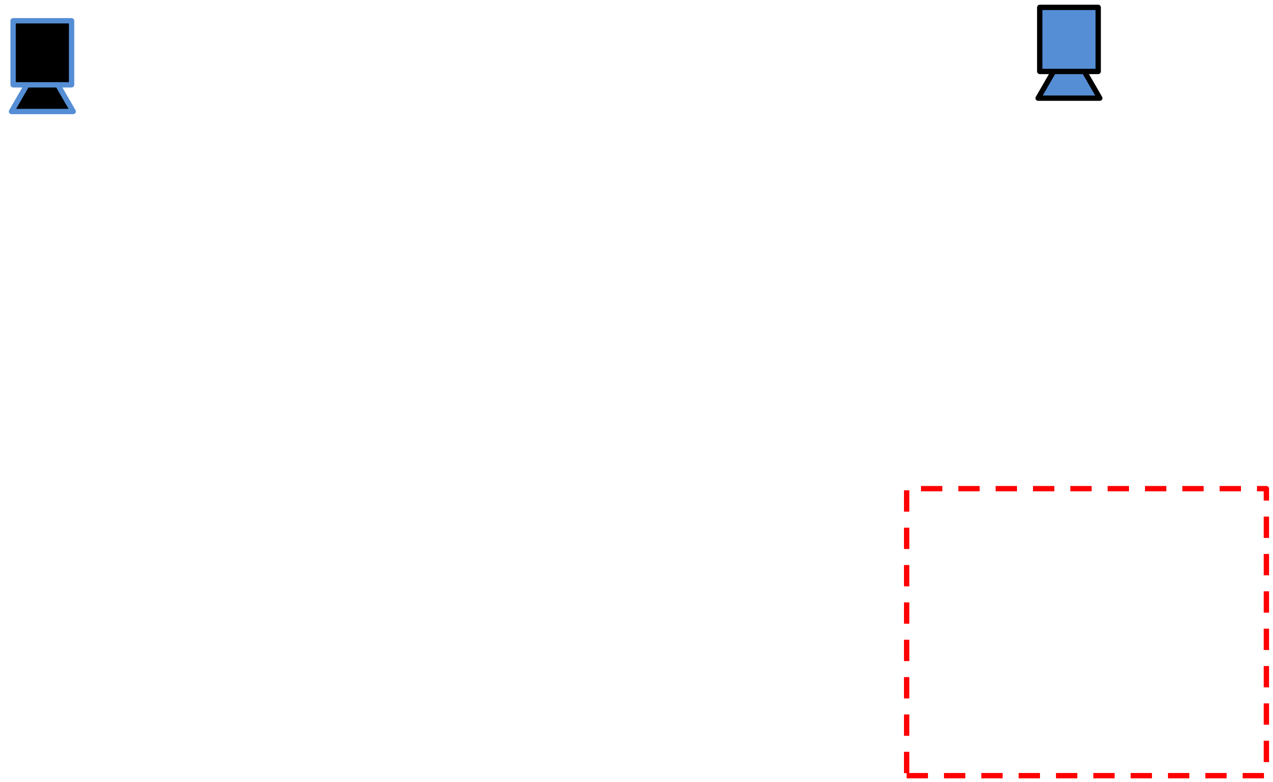


Depth Image Based Rendering

3D Warping

Physical camera parameters

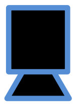
Virtual camera parameters



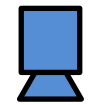
Depth Image Based Rendering

3D Warping

Physical camera parameters

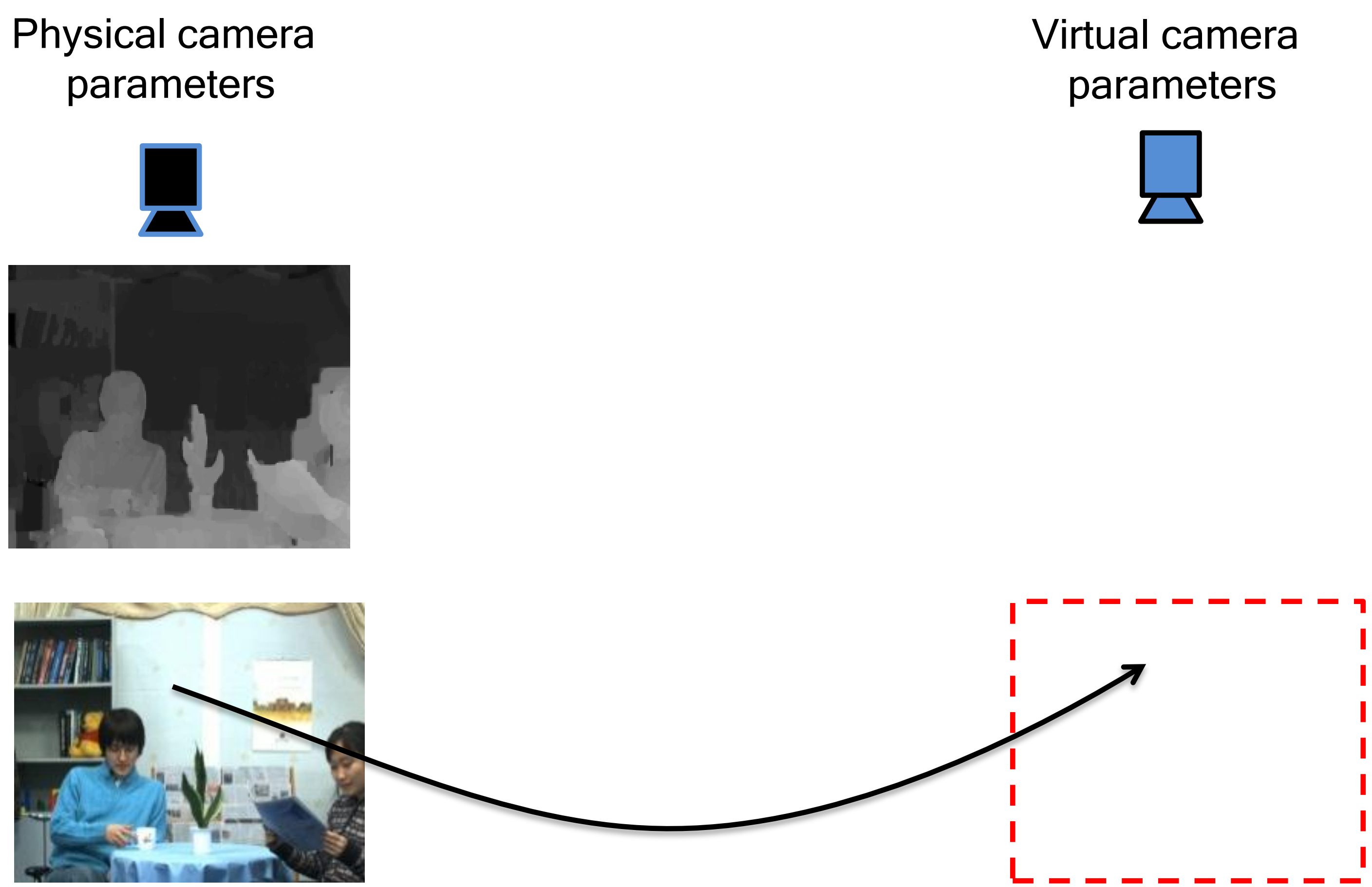


Virtual camera parameters



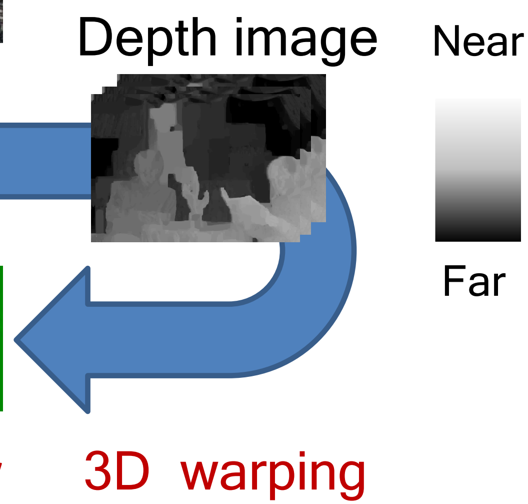
Depth Image Based Rendering

3D Warping



Depth Image Based Rendering

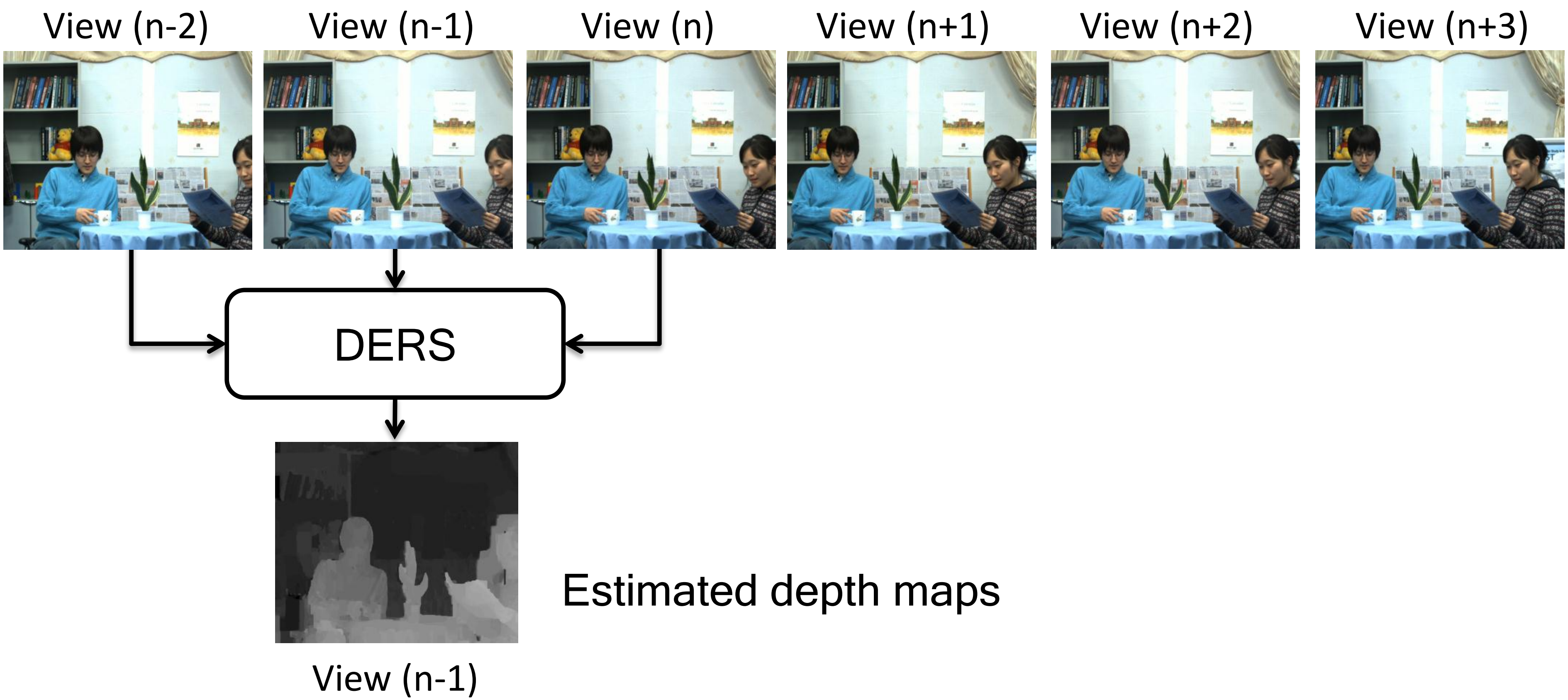
- Depth pixels represent shortest distance between object points and the camera plane
- To be estimated from multiview imagery



[1] C. Fehn: Depth-image-based rendering DIBR, compression, and transmission for a new approach on 3D-TV, SPIE, 2004.

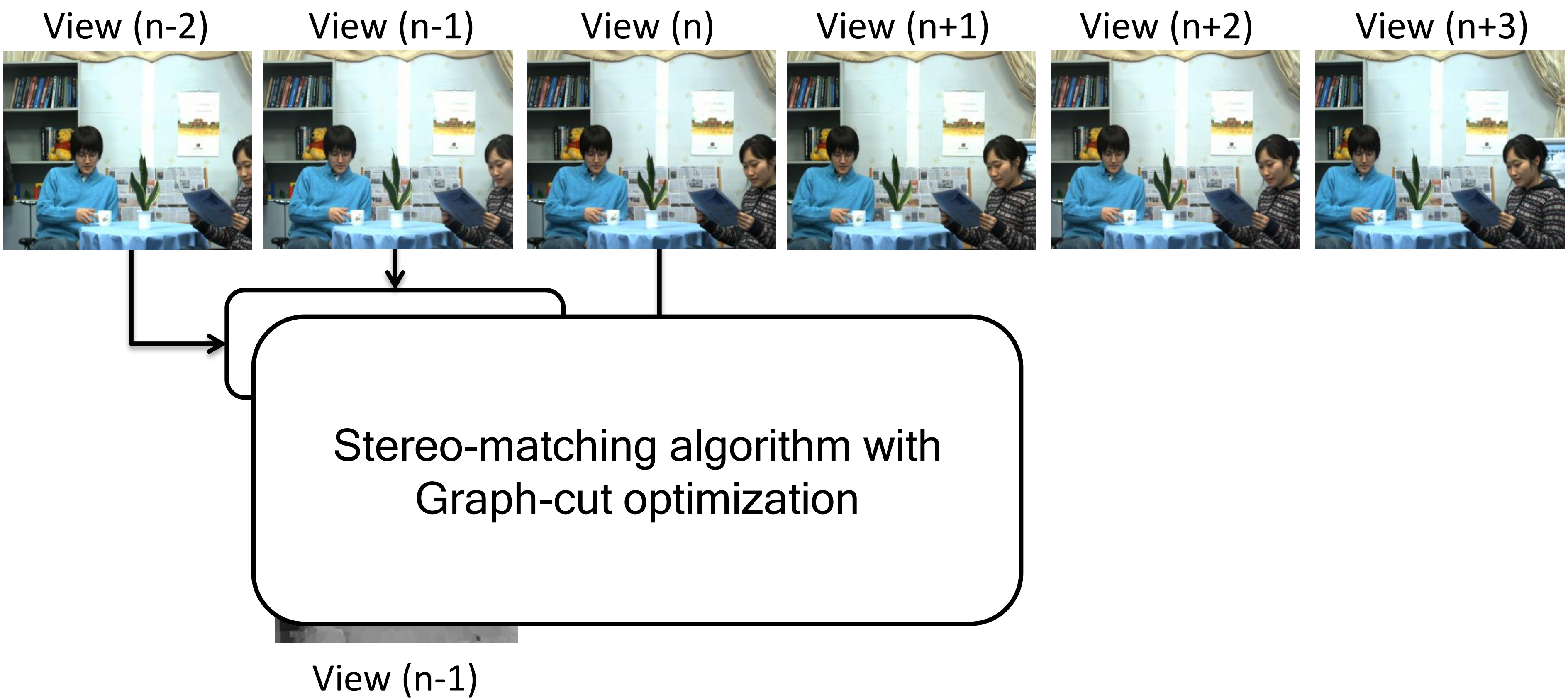
Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)



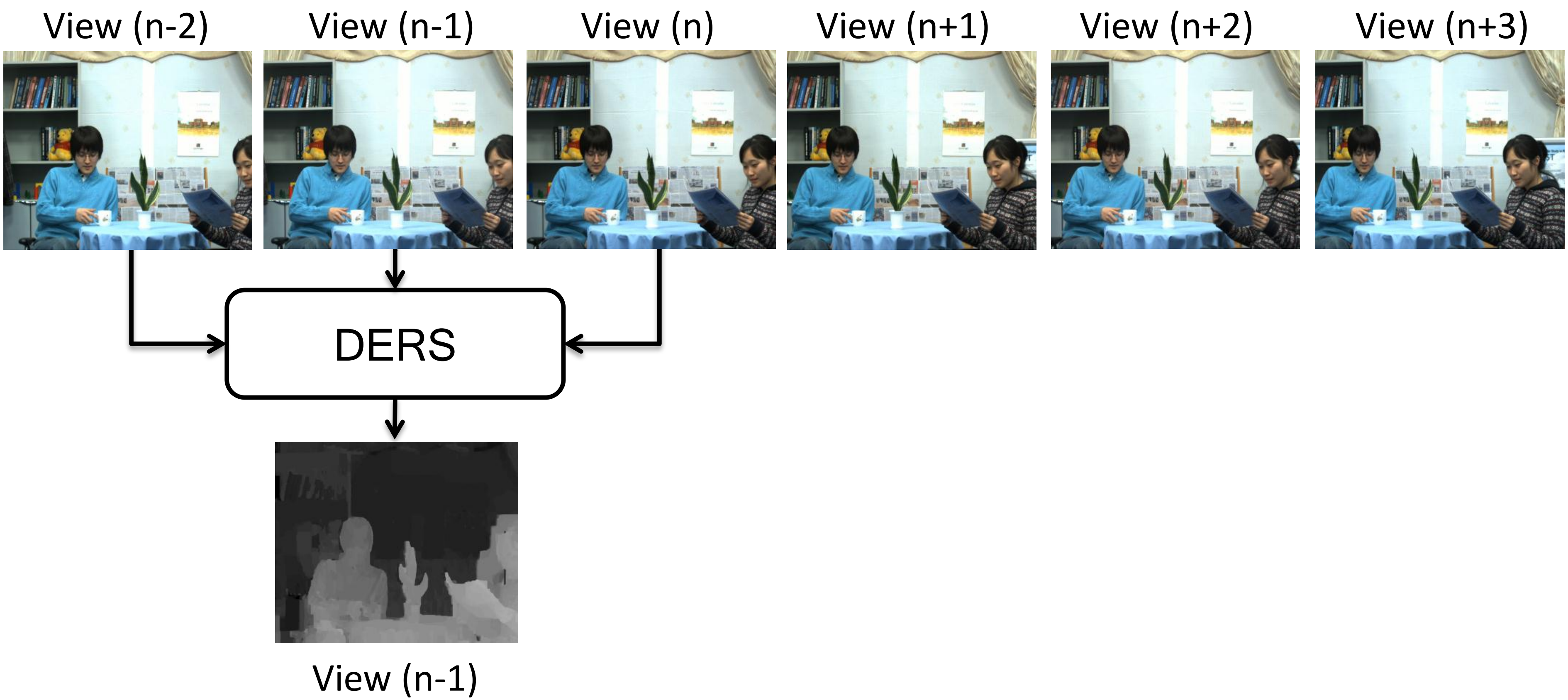
Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)



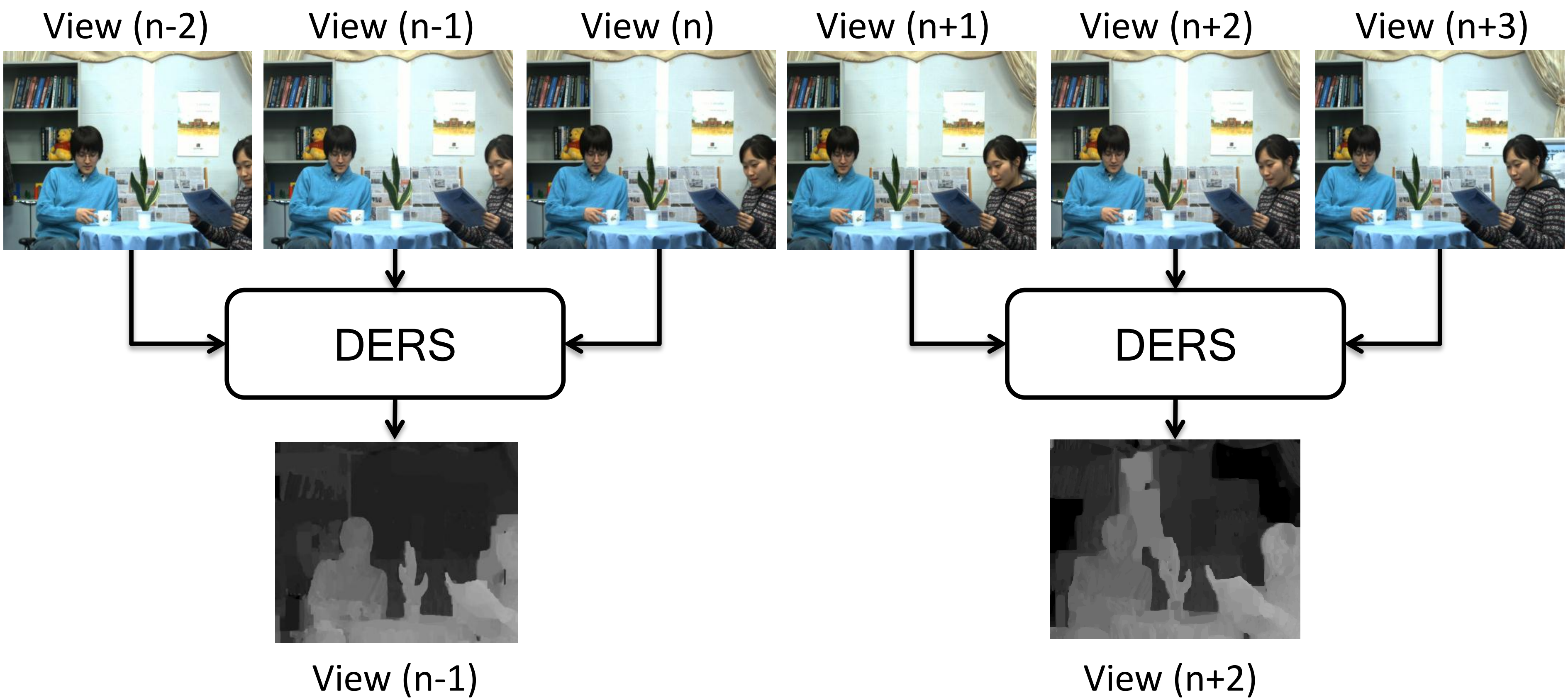
Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)



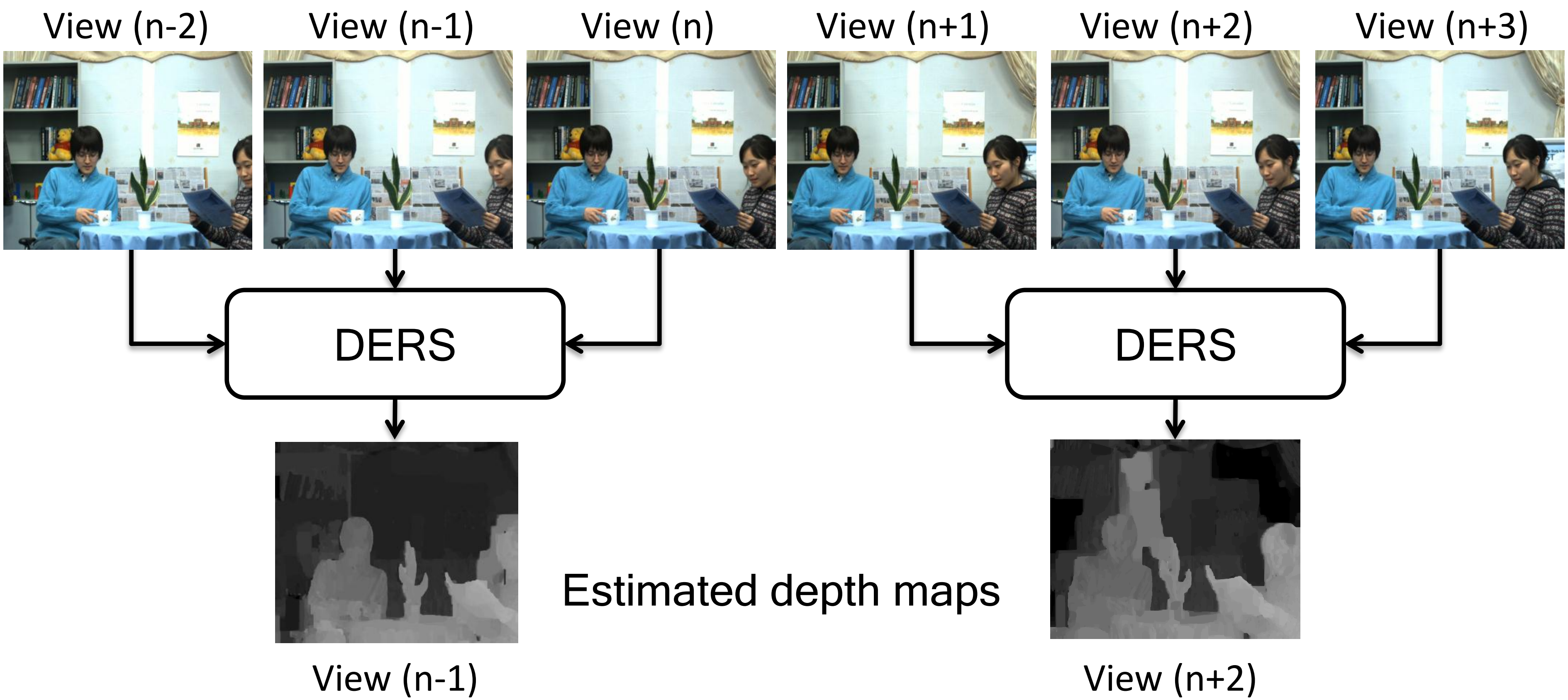
Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)



Depth Image Estimation

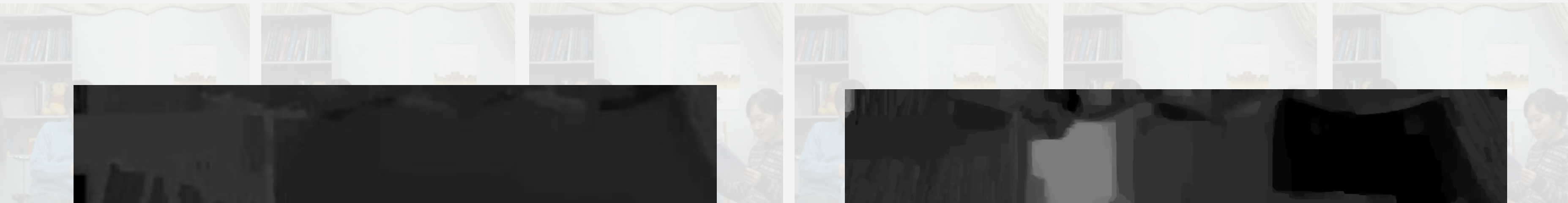
MPEG Depth Estimation Reference Software (DERS)



Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)

View (n-2) View (n-1) View (n) View (n+1) View (n+2) View (n+3)



View (n-1)

View (n+2)

Note: we assume a 1D-parallel camera arrangement

Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)

Note: we assume a 1D-parallel camera arrangement

Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)

Note: we assume a 1D-parallel camera arrangement

Depth Image Estimation

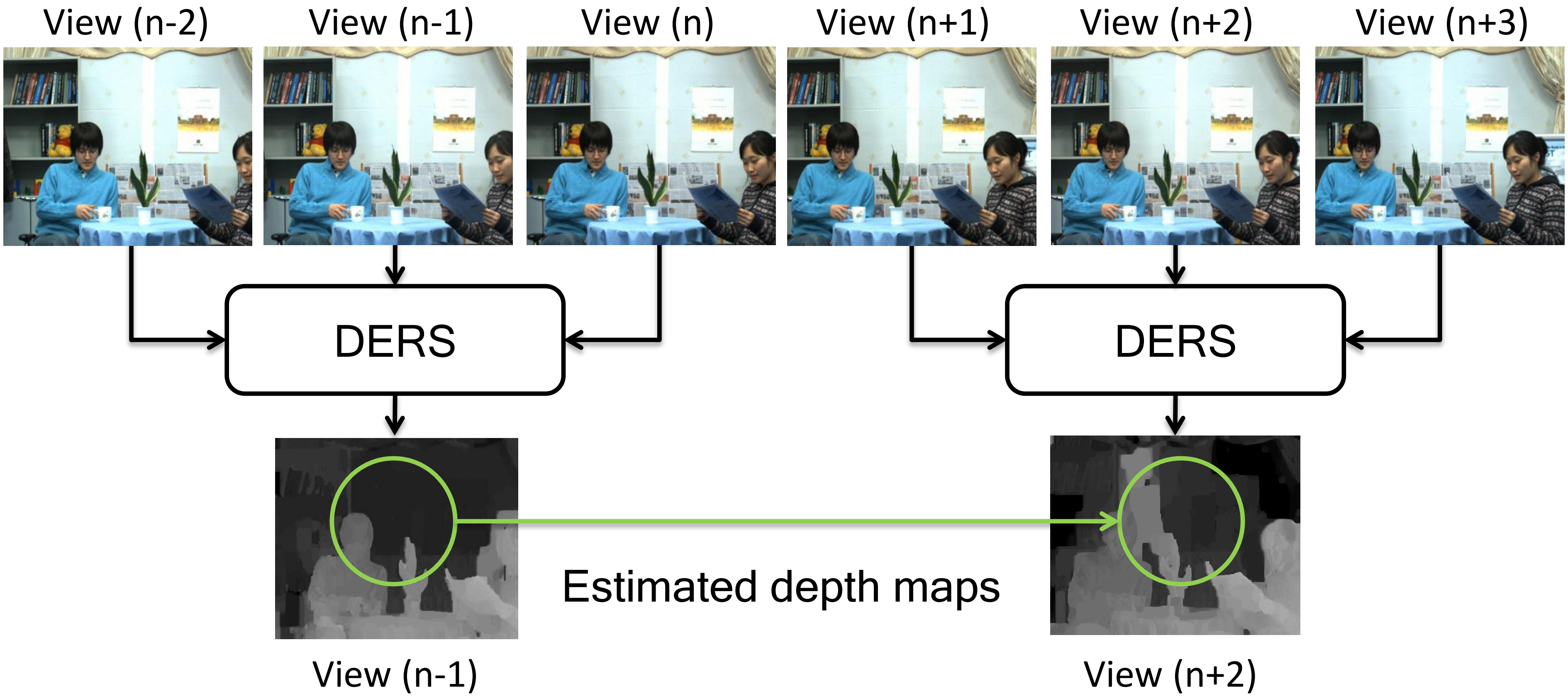
MPEG Depth Estimation Reference Software (DERS)

Problem: Inter-view depth inconsistency

Note: we assume a 1D-parallel camera arrangement

Depth Image Estimation

MPEG Depth Estimation Reference Software (DERS)

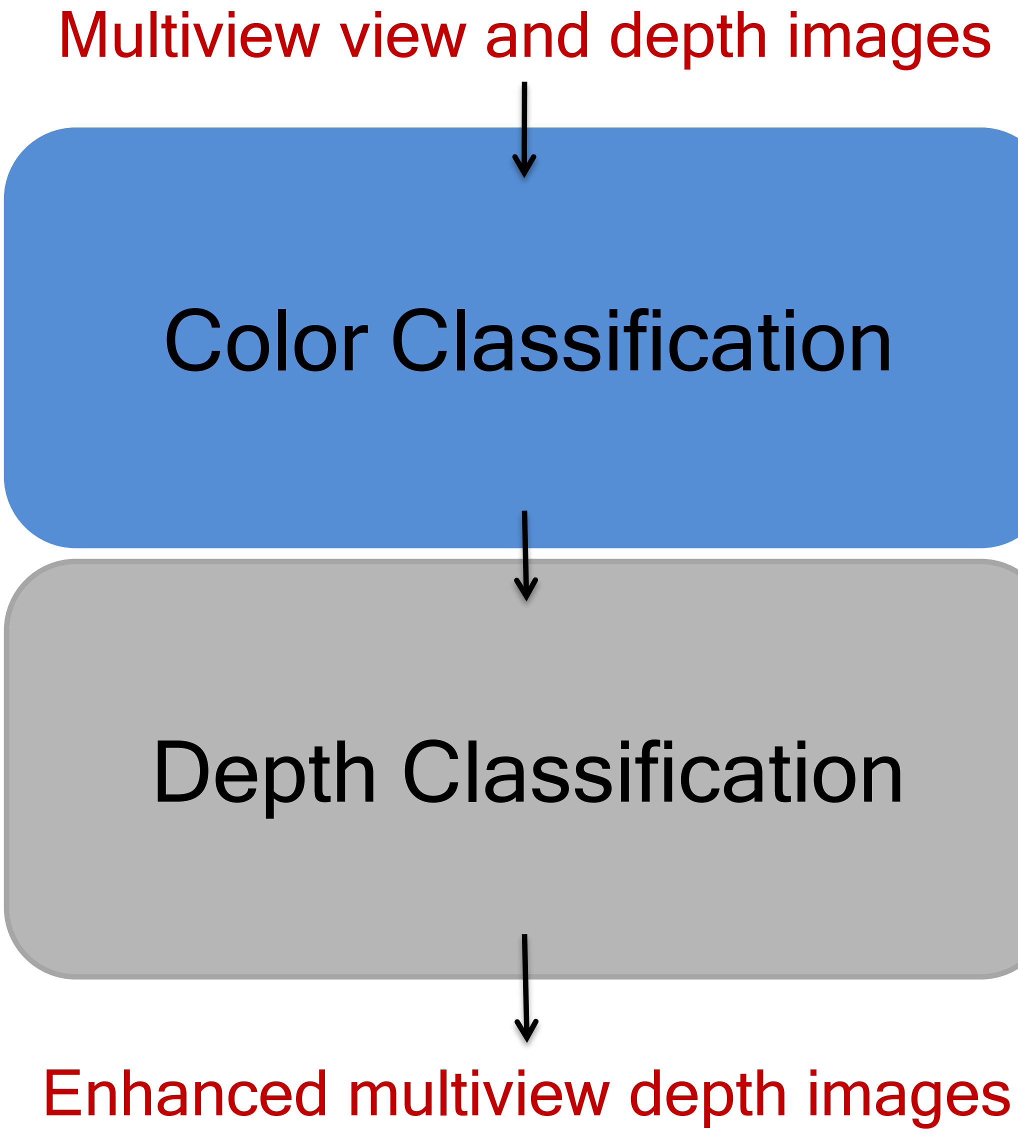


Problem: Inter-view depth inconsistency

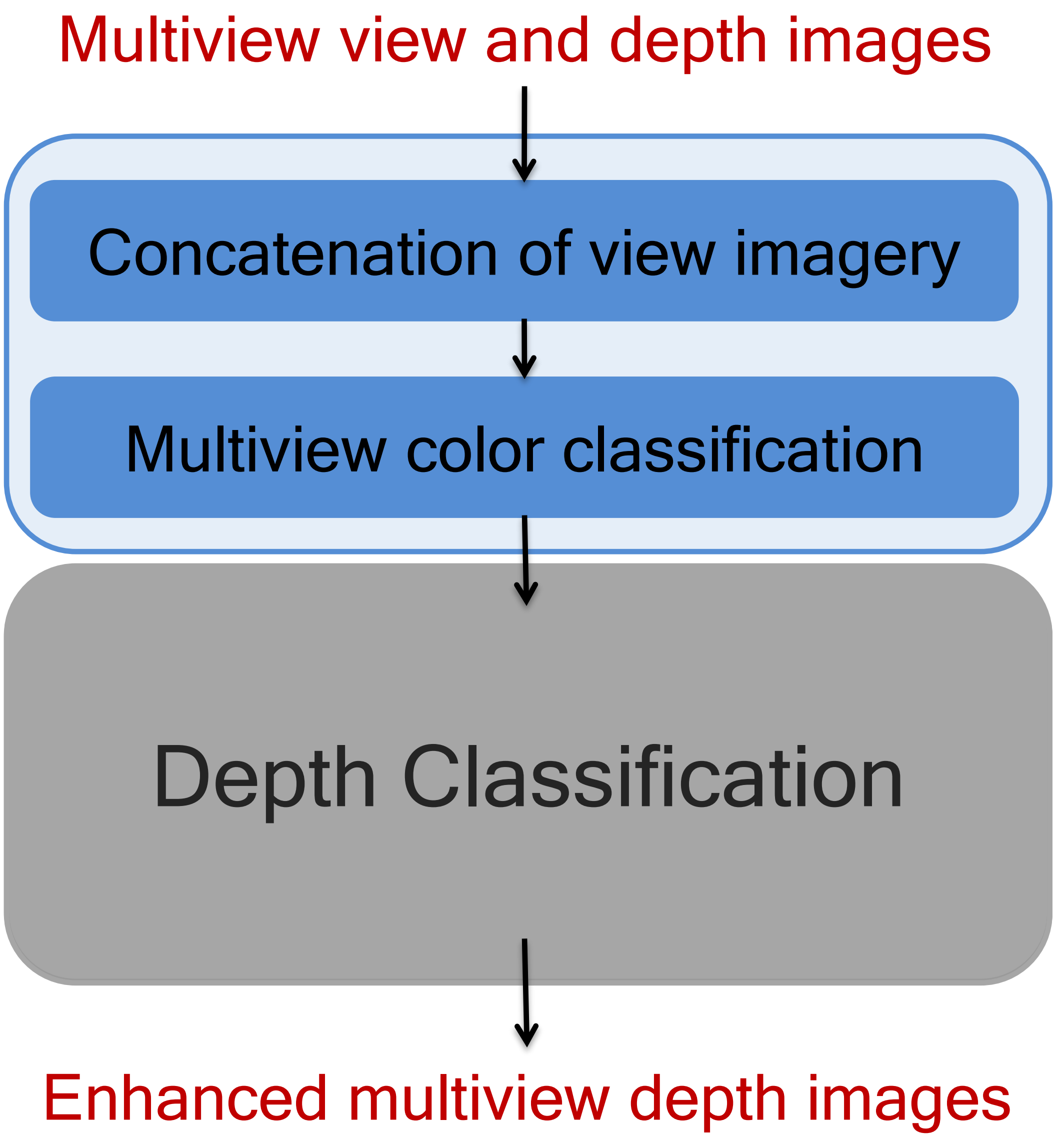
Note: we assume a 1D-parallel camera arrangement

Depth Enhancement Framework

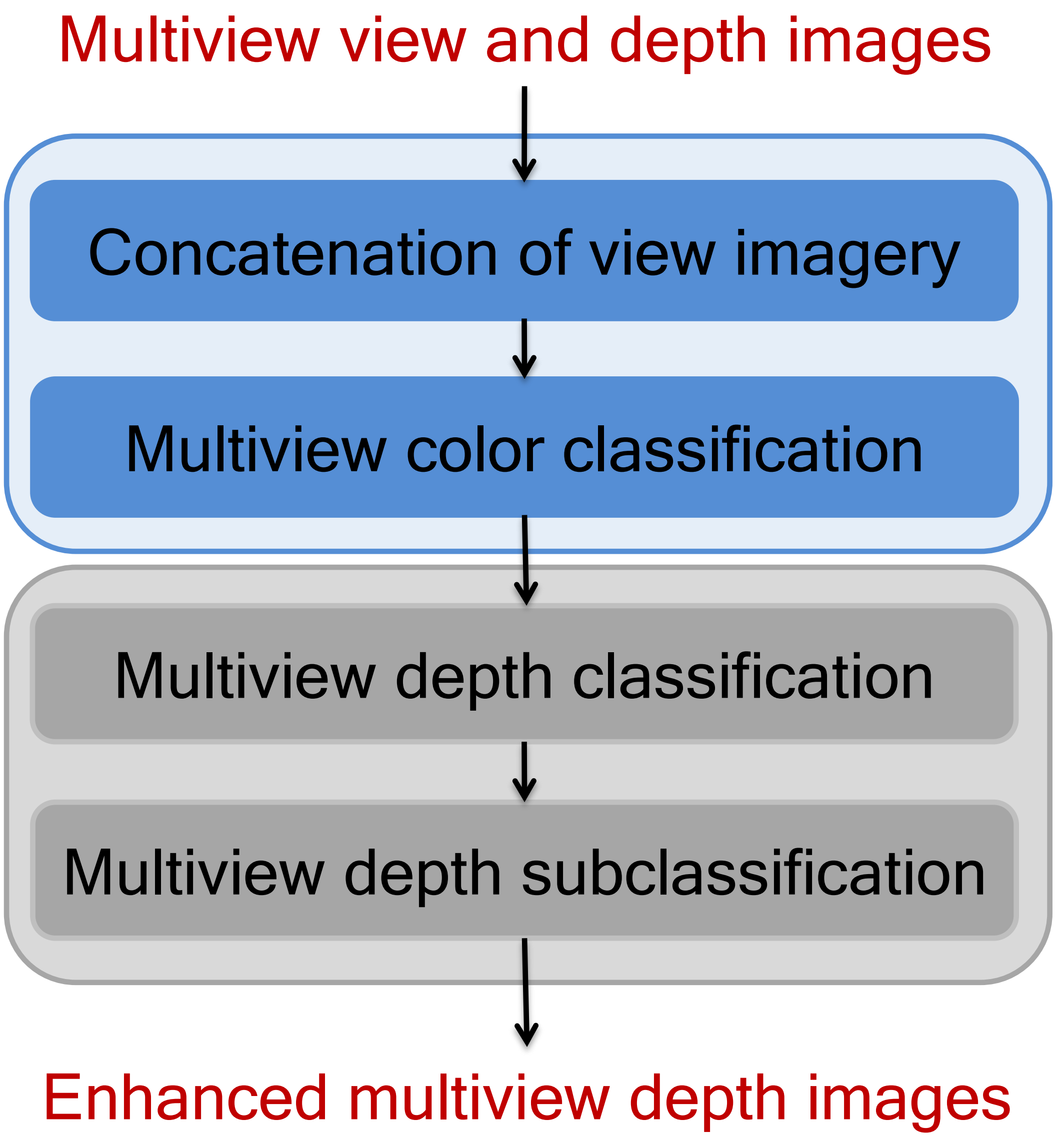
Overview of Depth Enhancement Framework



Overview of Depth Enhancement Framework



Overview of Depth Enhancement Framework



Concatenation of View Imagery

• • • View (n-1) View (n) View (n+1) View (n+2) • • •

View (n-1)

View (n)

View (n+1)

View (n+2)

Concatenation of View Imagery

View (n-1)

View (n)

View (n+1)

View (n+2)

• • • • • •

View (n-1) View (n) View (n+1) View (n+2)

• • • • • •

Multiview Color Classification

View (n) View (n+1)

...

...

RGB Color space

Multiview Color Classification

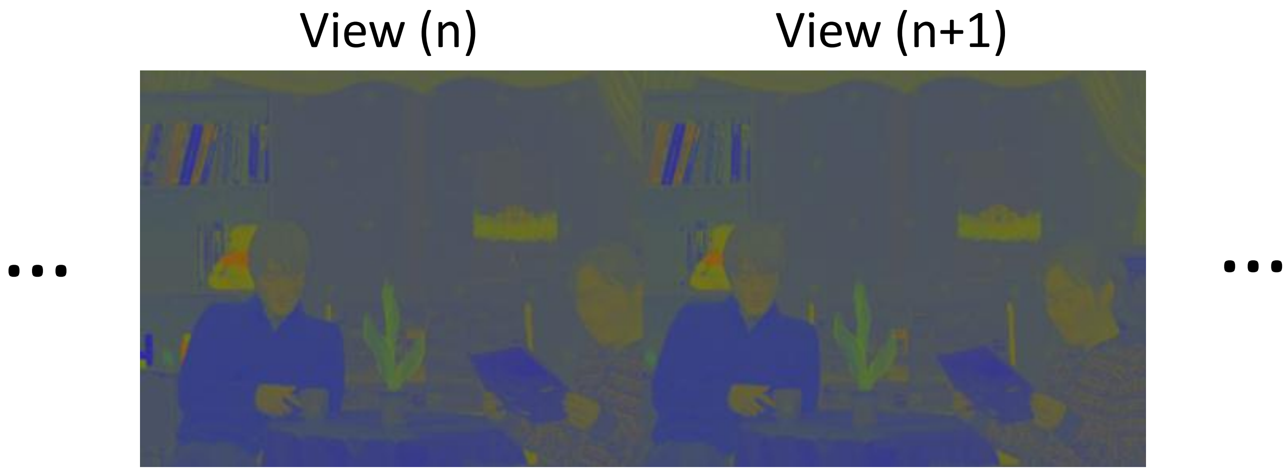
View (n) View (n+1)

...

...

xyz Chromatic color space

Multiview Color Classification



xyz Chromatic color space

- Insensitive to the absolut luminance
- A pixel is described by a vector of three chromaticity coefficients $[x\ y\ z]^T$, where

$$x+y+z = 1$$

Multiview Color Classification

Why variational Bayes inference (VI)?

- The goal of classification is to partition an image into regions each of which has a reasonably homogeneous visual appearance
- Usually, clustering algorithm suffers from one major drawbacks that the number of clusters has to be known
- Bayesian approaches automatically and optimally select the number of clusters
- Use of variational inference (VI) framework for Bayesian approaches gives an analytically tractable solution

[2] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

Multiview Color Classification

Why Dirichlet mixture models (DMM) ?

- The pixel vector in the chromaticity space has
  - nonnegative elements
  - bounded by the interval [0,1]
  - sum to one

- Assume that these pixel vectors are Dirichlet distributed

- DMM with variational inference is used to capture all underlying color clusters in multiview imagery

- It reduces complexity

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

[4] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon: Bayesian estimation of Dirichlet mixture model with variational inference, submitted, 2013.

Multiview Color Classification

Newspaper

Balloons

Kendo

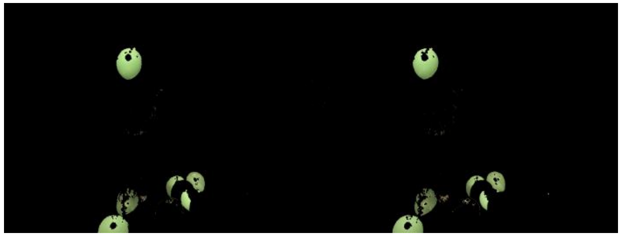
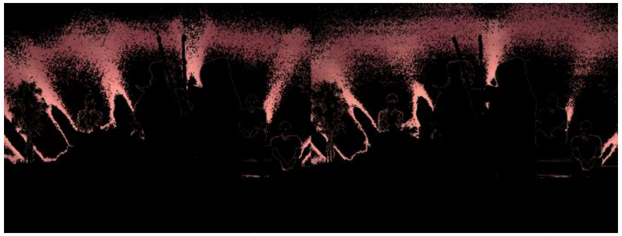
Input multiview data

Multiview Color Classification

Newspaper | Balloons | Kendo



Input multiview data

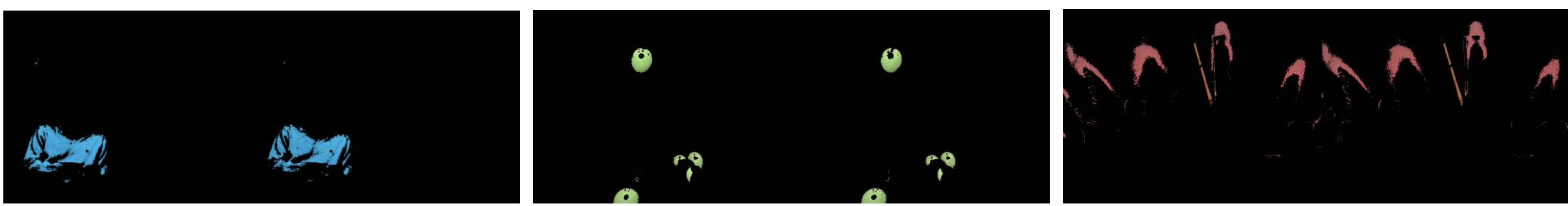
Using VI-DMM

Multiview Color Classification

Newspaper | Balloons | Kendo

Input multiview data

Using VI-DMM



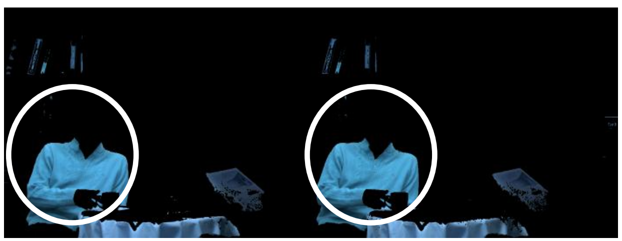
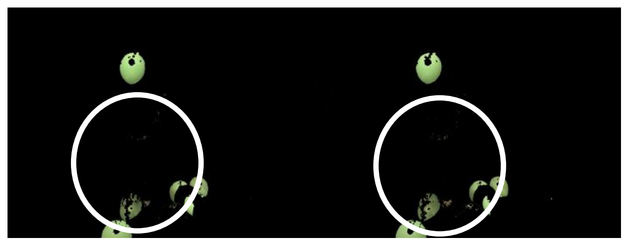
Using VI-GMM

Multiview Color Classification

Newspaper | Balloons | Kendo



Input multiview data


Using VI-DMM



Using VI-GMM

Multiview Depth Classification

Exploiting the per-pixel association between color and depth

View image   Depth image

Multiview Depth Classification

Exploiting the per-pixel association between color and depth

View image　　Depth image

Concatenated view imagery　　Concatenated depth imagery

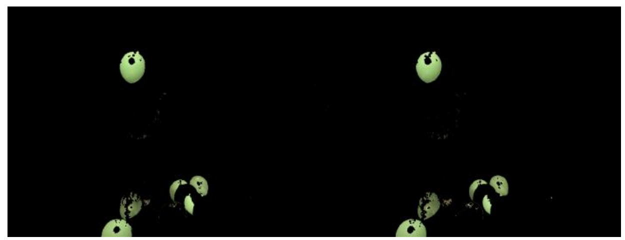
Multiview Depth Classification

Newspaper

Balloons

Kendo

Input multiview data



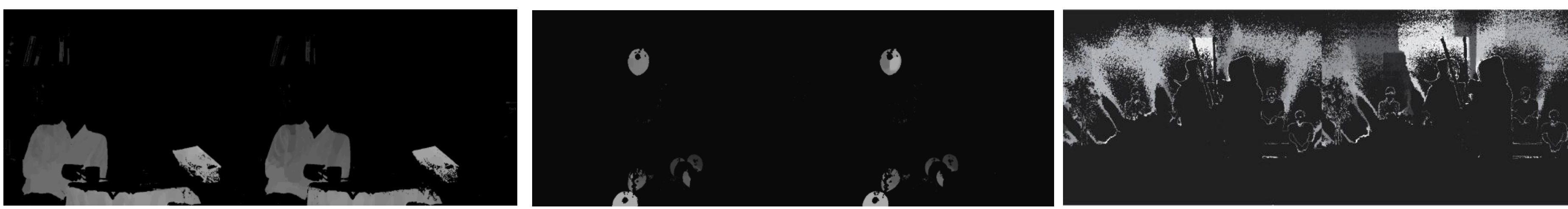
Using VI-DMM

Multiview Depth Classification

Newspaper | Balloons | Kendo

Input multiview data

Using VI-DMM

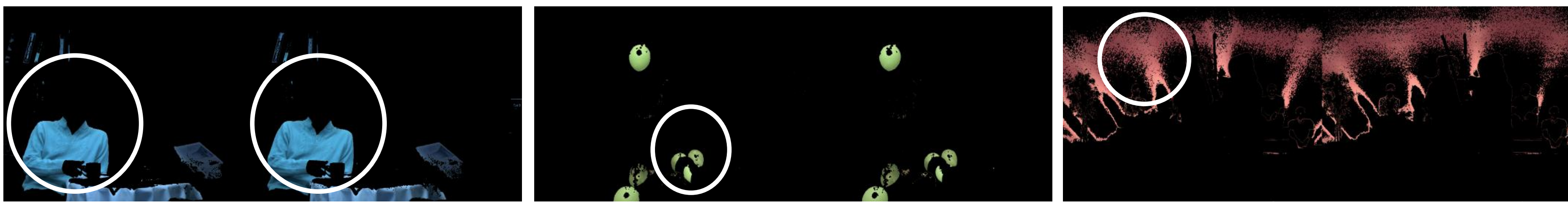


Depth clusters

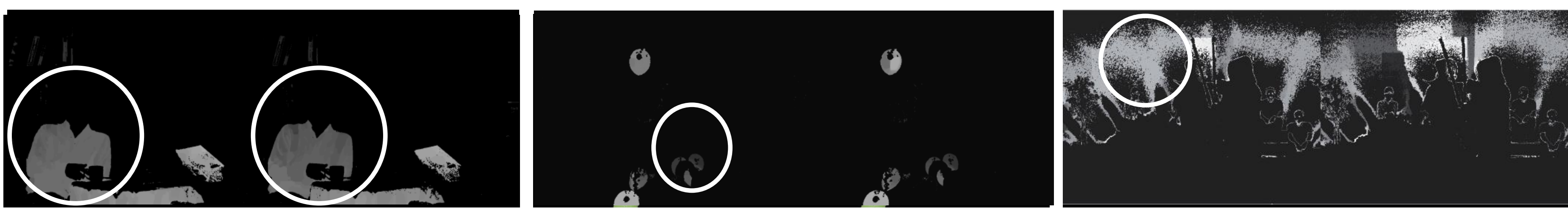
Multiview Depth Classification

Newspaper | Balloons | Kendo

Input multiview data



Using VI-DMM



Depth clusters

Multiview Depth Subclassification

Difference between color and depth clusters

- Members of color cluster have similar colors pixels
- Members of depth cluster may have different depth values

- Why?
  - due to foreground and background depth difference
  - due to inter-view inconsistency

Multiview Depth Subclassification

Mean shift Clustering

- A nonparametric clustering technique
- Knowledge of the number of clusters not required
- Assigns the mean to depth pixels irrespective of the originating viewpoints
- Generative model based approaches imply higher computational complexity

[5] D. Comaniciu and P. Meer: Mean shift: A robust approach toward feature space analysis, IEEE Trans. Pattern Anal. Mach. Intell., 2002.

Experimental Results

Experimental Results

MPEG 3DTV multiview data set

Newspaper
(1024 X 768)

Lovebird1
(1024 X 768)

Kendo
(1024 X 768)

Balloons
(1024 X 768)

Poznan street
(1920 X 1088)

Experimental Results

Multiview data set	Initial number of mixture components	Active number of mixture components (after convergence)	
		VI-GMM	VI-DMM
Lovebird1	100	31	24
Kendo	100	34	15

Depth Image-Based Rendering

MPEG View Synthesis Reference Software (VSRS) 3.5

Enhanced depth map

Left

Enhanced depth map

Right

Depth Image-Based Rendering

MPEG View Synthesis Reference Software (VSRS) 3.5

Enhanced depth map

Left

Reference view

Enhanced depth map

Right

Reference view

Depth Image-Based Rendering

MPEG View Synthesis Reference Software (VSRS) 3.5

Left

Enhanced depth map

Reference view

3D warping

Warped view

Right

Enhanced depth map

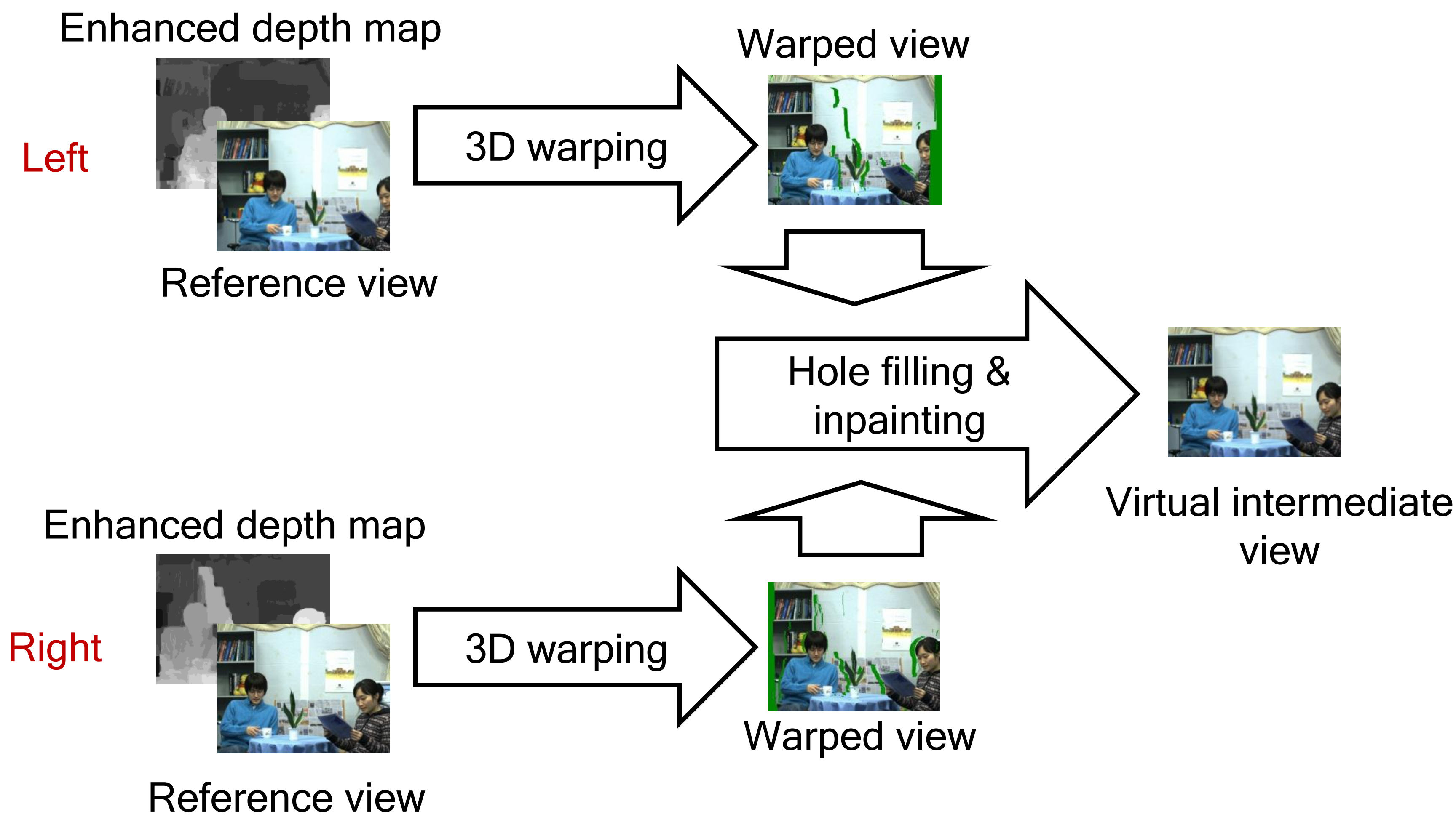
Reference view

3D warping

Warped view

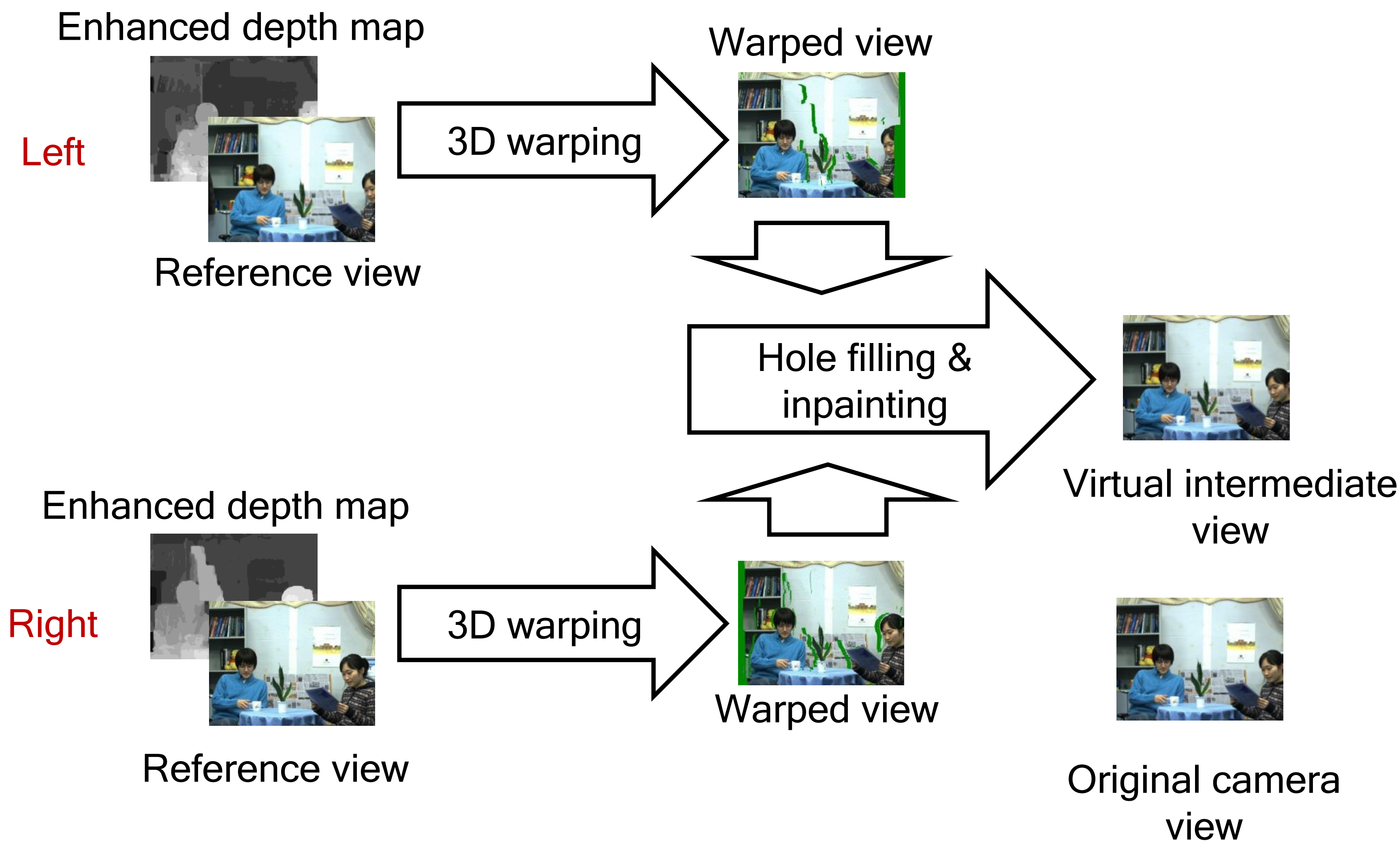
Depth Image-Based Rendering

MPEG View Synthesis Reference Software (VSRS) 3.5



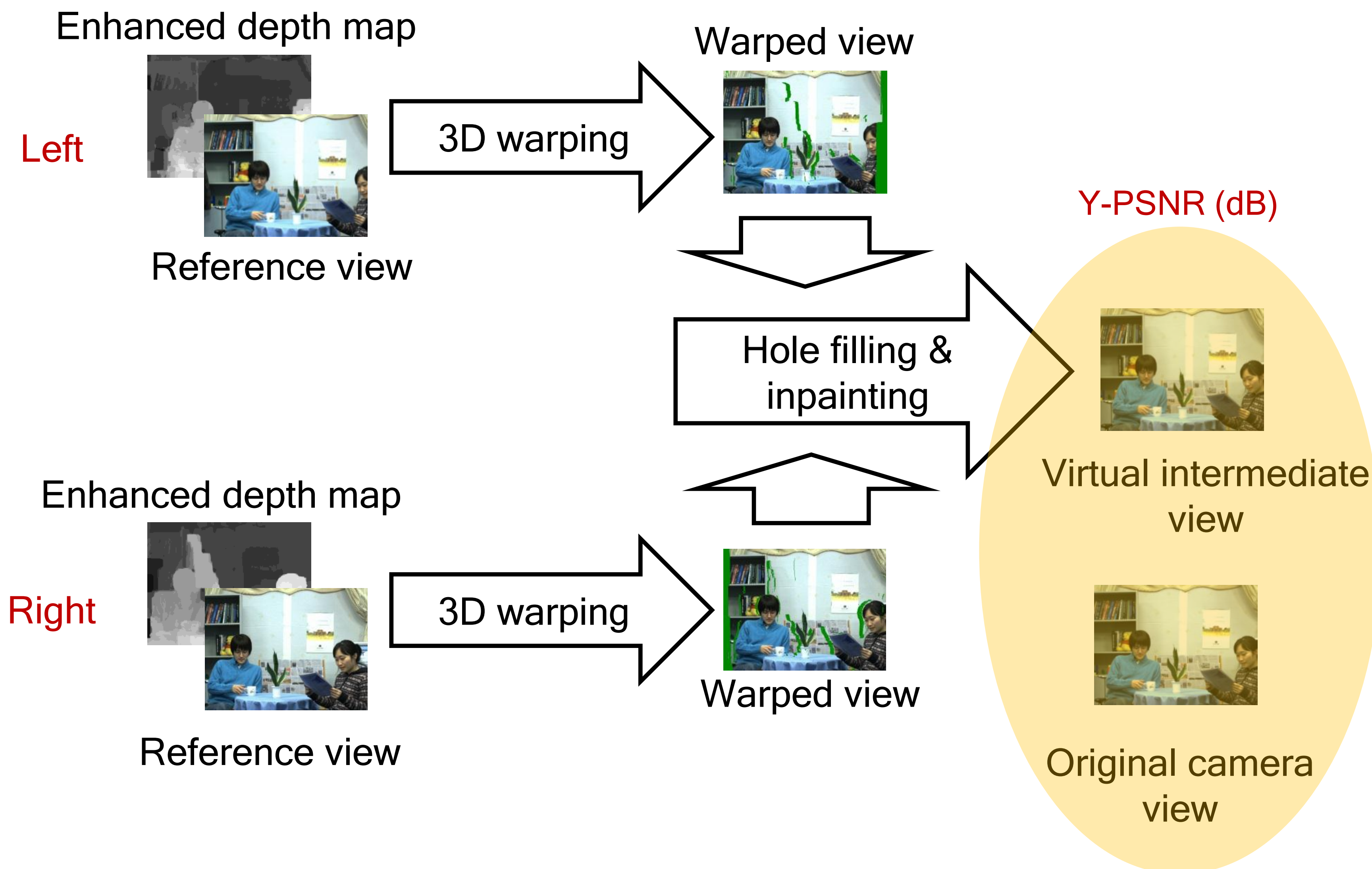
Depth Image-Based Rendering

MPEG View Synthesis Reference Software (VSRS) 3.5



Depth Image-Based Rendering

MPEG View Synthesis Reference Software (VSRS) 3.5



Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VBIGMM + K-Means depth maps [3]	With VBIDMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VBIGMM + K-Means depth maps [3]	With VBIDMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VBIGMM + K-Means depth maps [3]	With VBIDMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VBIGMM + K-Means depth maps [3]	With VBIDMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VBIGMM + K-Means depth maps [3]	With VBIDMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VBIGMM + K-Means depth maps [3]	With VBIDMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VBIGMM + K-Means depth maps [3]	With VBIDMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Objective Results

Test sequence	Input view pair	Virtual view	Y-PSNR [dB]		
			With MPEG depth maps	With VI-GMM + K-Means depth maps [3]	With VI-DMM + Mean-shift depth maps
Newspaper	(4, 6)	5	32.00	32.10	32.11
Kendo	(3, 5)	4	36.54	36.72	39.35
Lovebird1	(6, 8)	7	28.50	28.68	29.04
Balloons	(3, 5)	4	35.69	35.93	36.02
Poznan Street	(3, 5)	4	35.56	35.58	35.72

- K-means sub-clustering
  - Number of cluster : 12

[3] P. K. Rana, J. Taghia, and M. Flierl: A Variational Bayesian Inference Framework for Multiview Depth Image Enhancement, IEEE ISM, 2012.

Subjective Results

Test sequence: Kendo

With MPEG depth map

With VI-DMM+Mean-shift depth map

Subjective Results

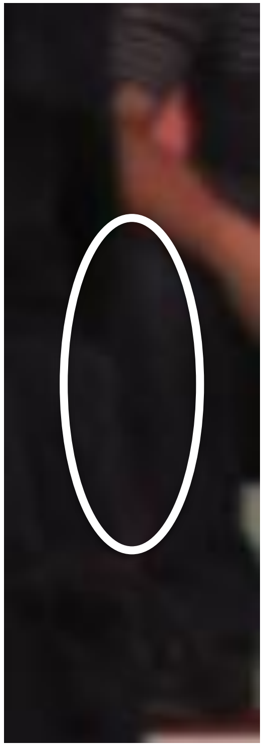
Test sequence: Kendo

With MPEG depth map

With VI-DMM+Mean-shift depth map

Subjective Results

Test sequence: Kendo



Original

With VI-DMM + Mean-Shift depth maps

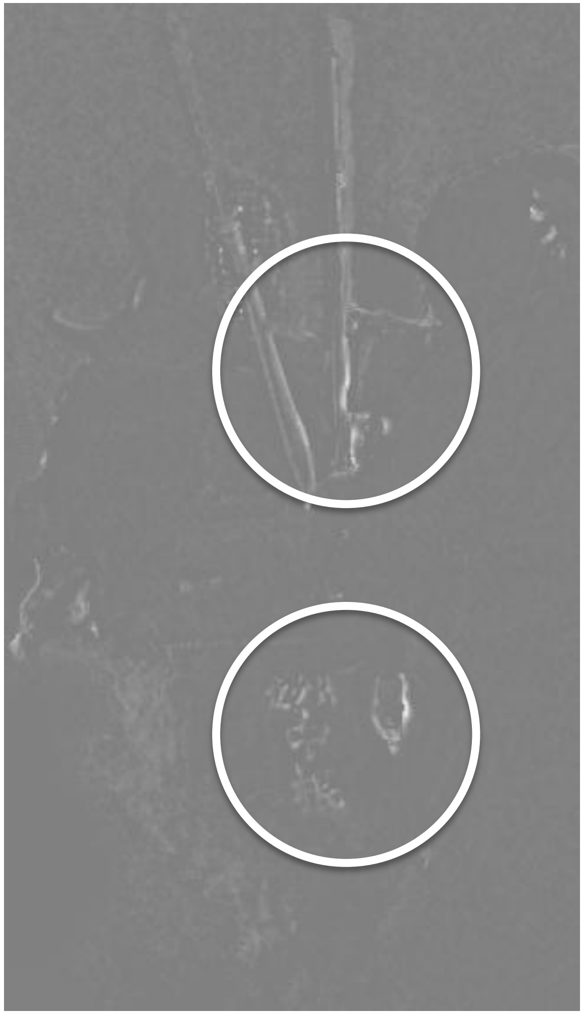
With VI-GMM + K-Means depth maps

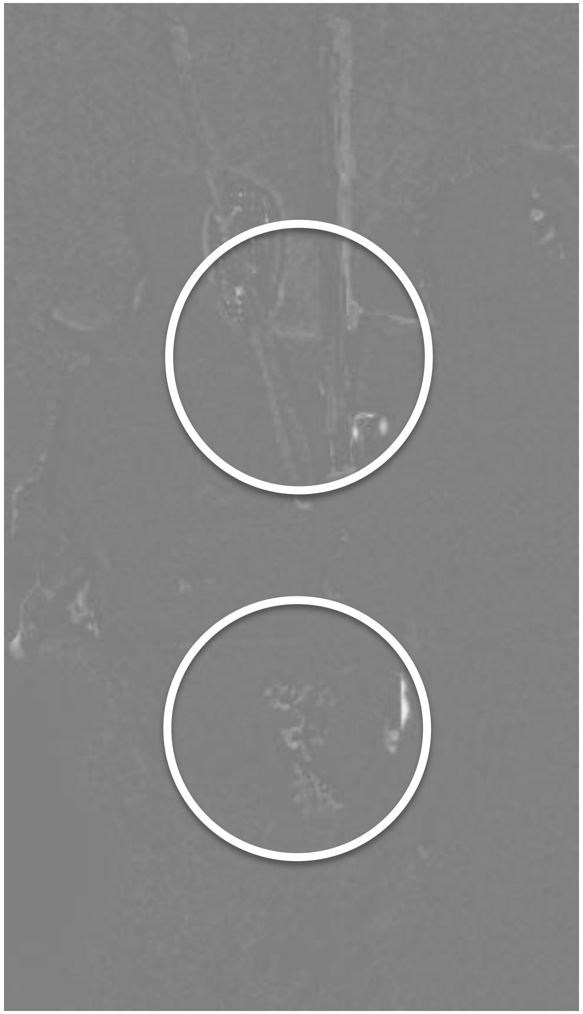
With MPEG depth maps

Subjective Results

Test sequence: Kendo

With VI-GMM +K-Means depth maps





With VI-DMM +Mean-Shift depth maps

Subjective Results

Test sequence: Lovebird 1

With MPEG depth map

With VI-DMM+Mean-shift depth map

Subjective Results

Test sequence: Lovebird 1

With MPEG depth map

With VI-DMM+Mean-shift depth map

Subjective Results

Test sequence: Lovebird 1

Original

With VI-DMM
+Mean-Shift
depth maps

With VI-GMM
+K-Means
depth maps

With MPEG
depth maps

Subjective Results

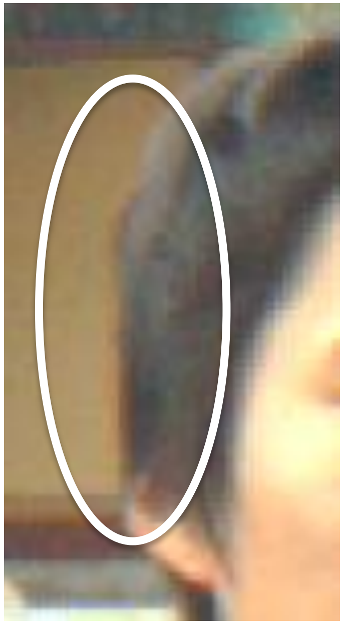
Test sequence: Lovebird 1

With MPEG depth map

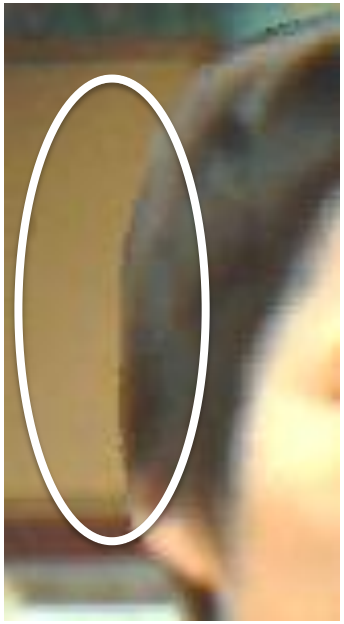
With VI-DMM+Mean-shift depth map

Subjective Results

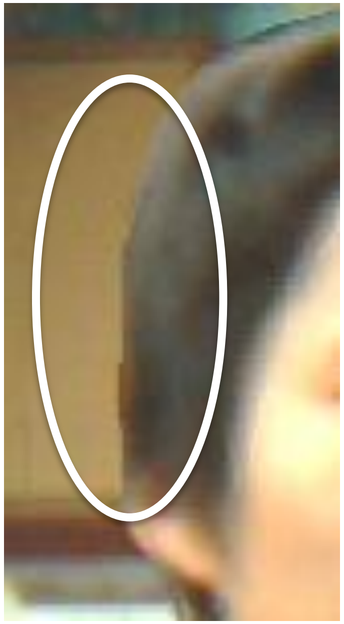
Test sequence: Lovebird 1



Original



With VI-DMM +Mean-Shift depth maps



With VI-GMM +K-Means depth maps

With MPEG depth maps

Conclusions

- The inter-view depth consistency and hence, the free-viewpoint experience improve
- The per-pixel association between depth and color is exploited by classification
- Depth subclassification improves depth maps and hence, view rendering quality
- Both objective and subjective results improve

[6] P. K. Rana, Z. Ma, J. Taghia, and M. Flierl: Multiview Depth Map Enhancement by Variational Bayes Inference Estimation of Dirichlet Mixture Models, IEEE ICASSP, 2013.

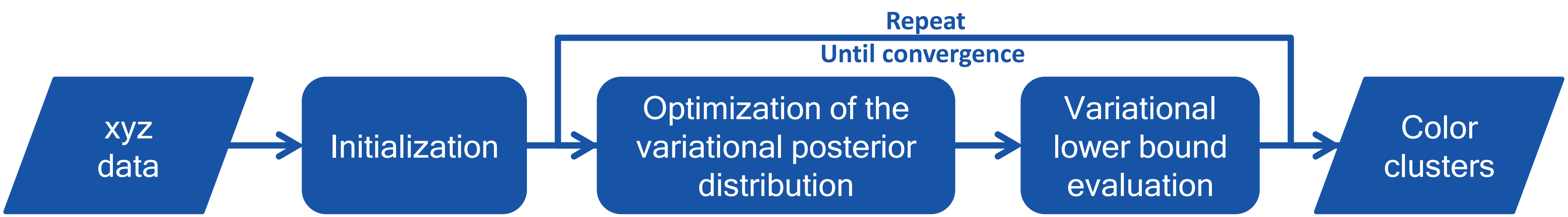
Future Directions

- A fully probabilistic multiview depth image enhancement
  - With improved computational efficiency
  - With improved depth subclassification

Thank You

Multiview color classification

Dirichlet mixture model with variational Bayes inference



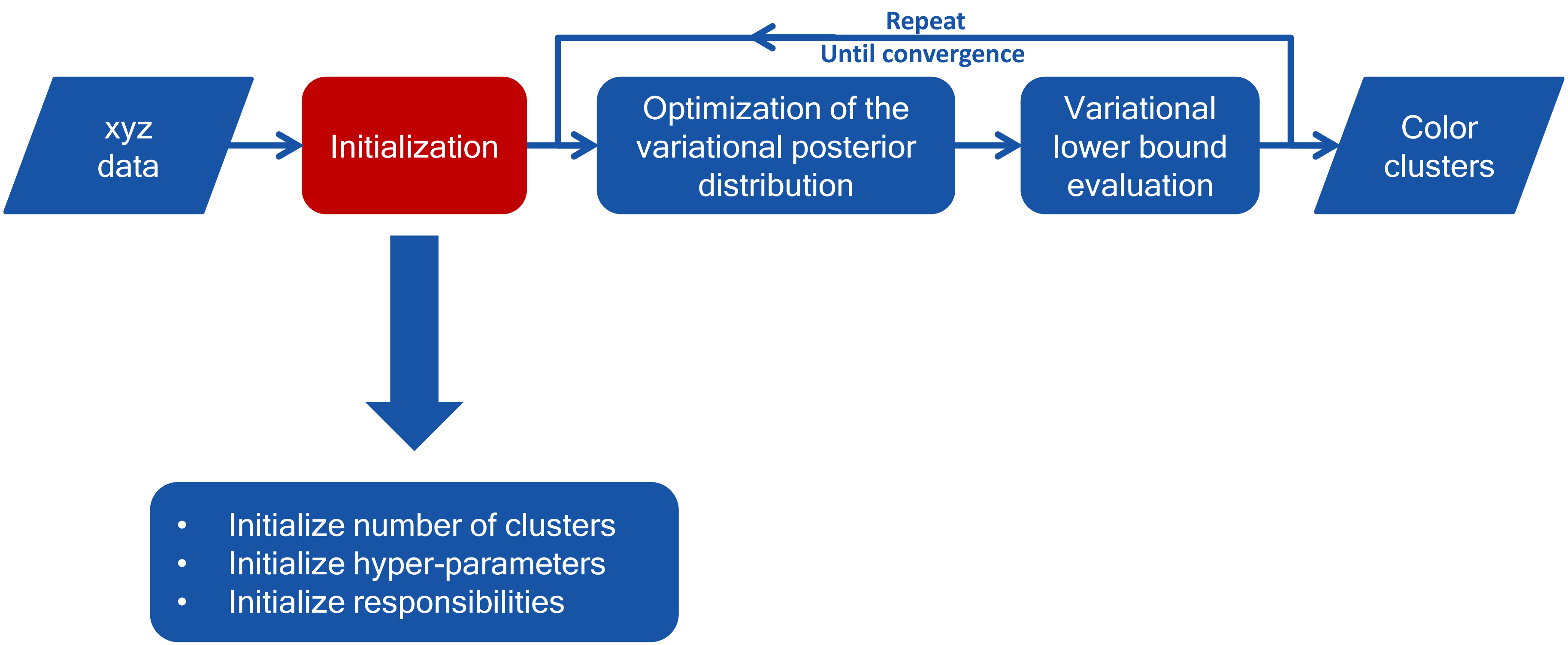
[1] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

[2] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of Dirichlet mixture model with variational inference," IEEE Trans. PAMI, submitted, 2013.

Y

Multiview color classification

Dirichlet mixture model with variational Bayes inference



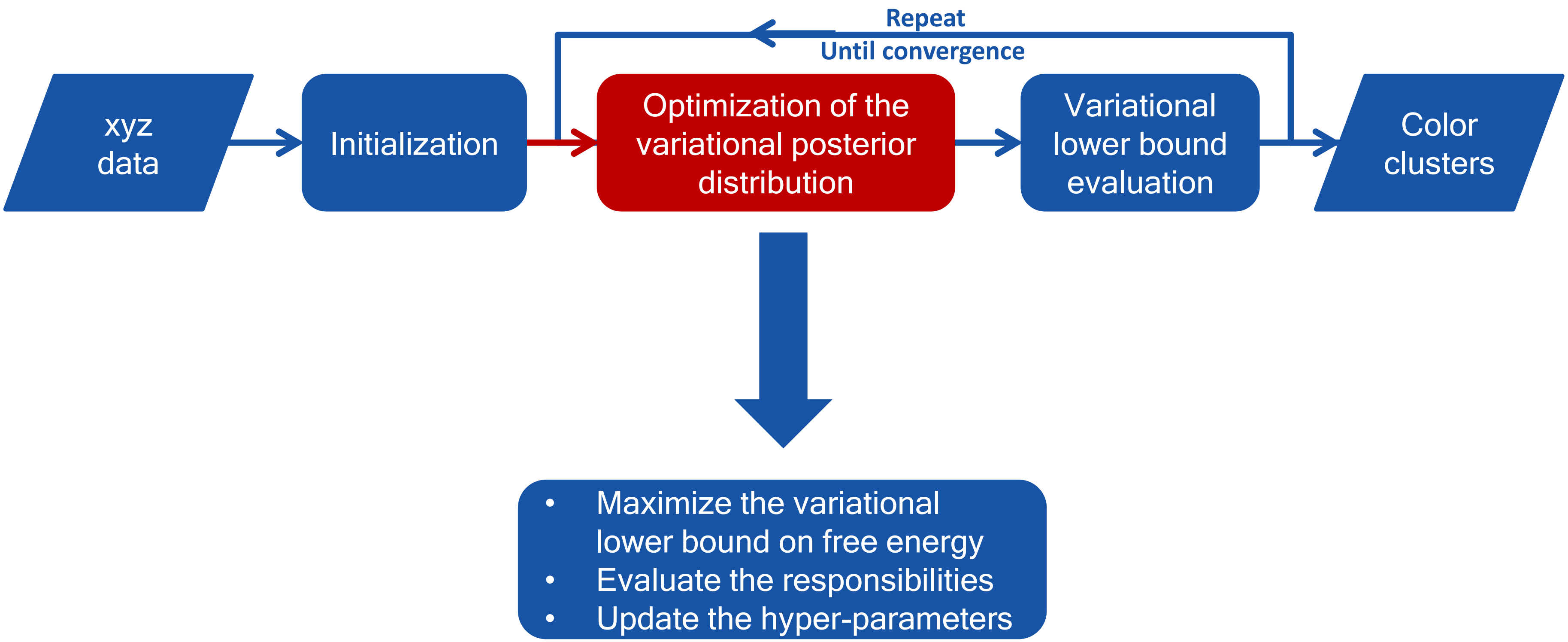
[1] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

[2] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of Dirichlet mixture model with variational inference," IEEE Trans. PAMI, submitted, 2013.

Y

Multiview color classification

Dirichlet mixture model with variational Bayes inference



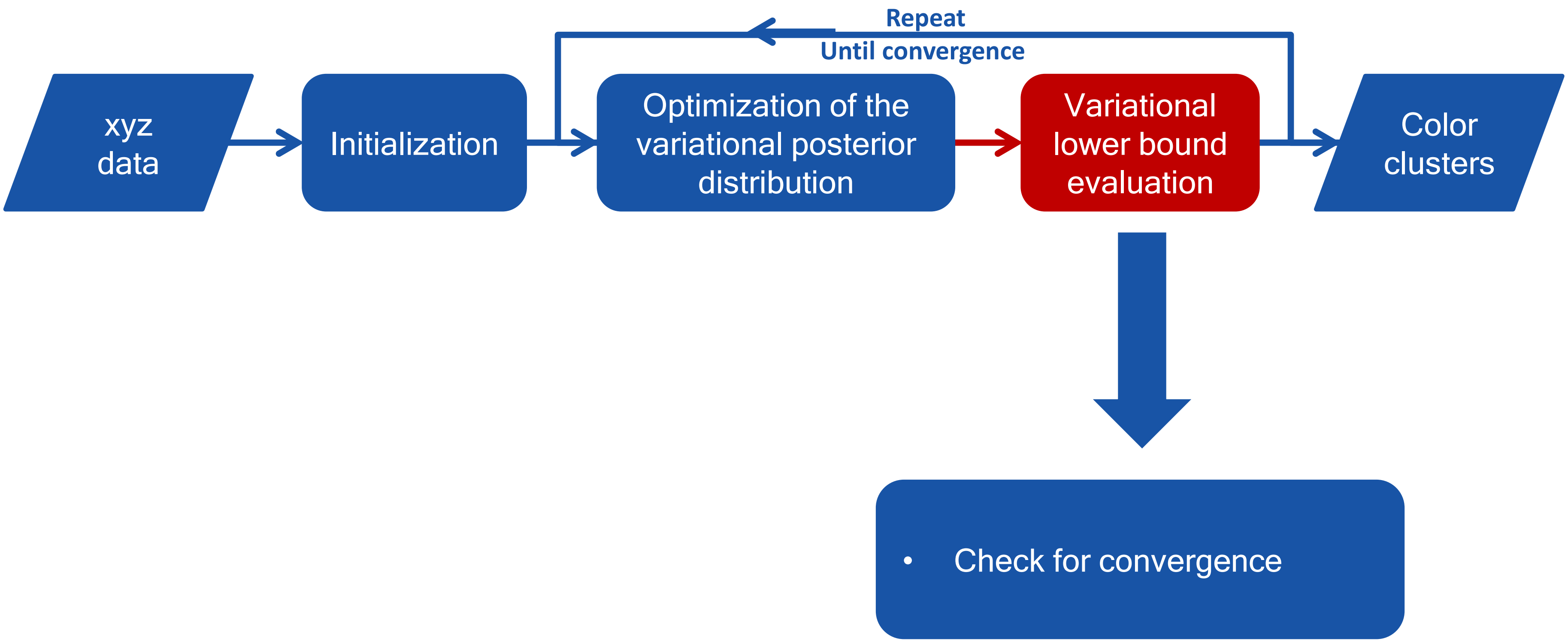
[1] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

[2] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of Dirichlet mixture model with variational inference," IEEE Trans. PAMI, submitted, 2013.

Y

Multiview color classification

Dirichlet mixture model with variational Bayes inference

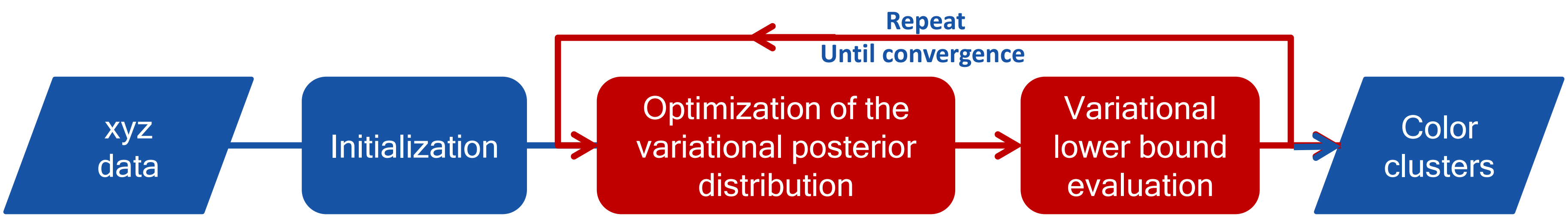


[1] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.
[2] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of Dirichlet mixture model with variational inference," IEEE Trans. PAMI, submitted, 2013.

Y

Multiview color classification

Dirichlet mixture model with variational Bayes inference



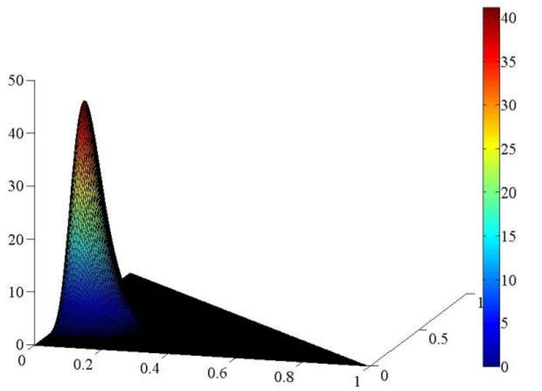
[1] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

[2] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of Dirichlet mixture model with variational inference," IEEE Trans. PAMI, submitted, 2013.

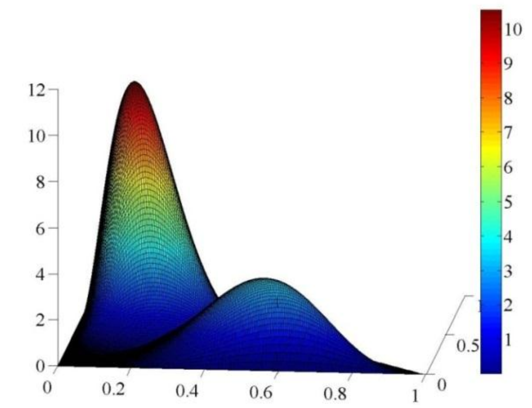
Y

Dirichlet Plot Details

- For probability density function of Dirichlet distribution $\alpha = [2\ 10\ 15]$
- For probability density function of Dirichlet mixture model parameters $\alpha_1 = [6\ 2\ 4]$ and $\alpha_2 = [3\ 8\ 5]$ with mixture weights $\pi_1 = 0.3$ and $\pi_2 = 0.7$, respectively.



Probability density function of Dirichlet distribution



Probability density function of Dirichlet mixture model

Subjective Results

Test sequence: Newspaper

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective Results

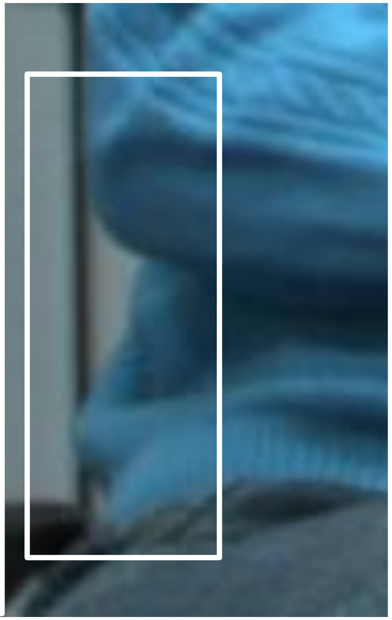
Test sequence: Newspaper

With MPEG depth map

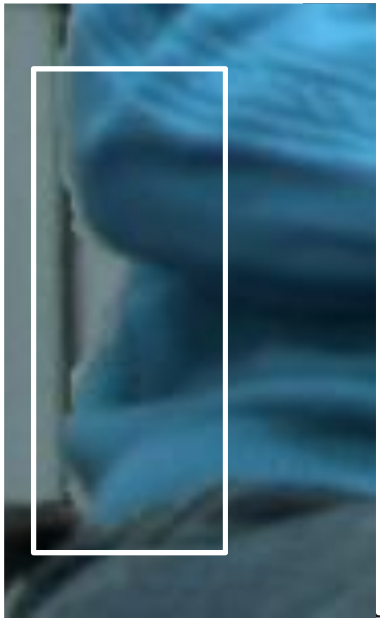
With VBDMM Mean-shift depth map

Subjective Results

Test sequence: Newspaper



Original
depth maps



With MPEG
depth maps

With VBGMM
K-Means
depth maps

With VBDMM
Mean-Shift
depth maps

X

Subjective Results

Test sequence: Newspaper

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective Results

Test sequence: Newspaper

Original
depth maps

With MPEG
depth maps

With VBGMM
K-Means
depth maps

With VBDMM
Mean-Shift
depth maps

Subjective Results

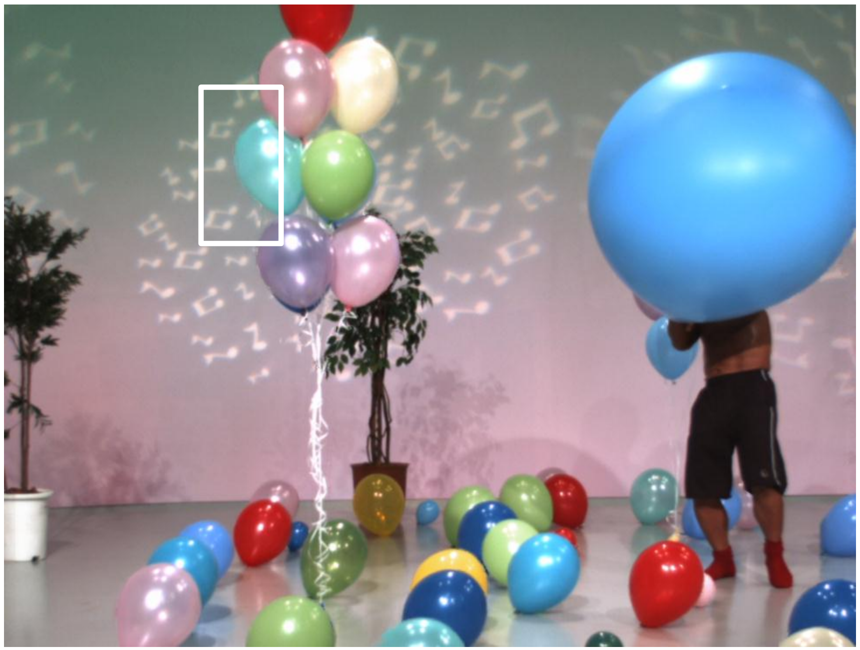
Test sequence: Balloons

With MPEG depth map

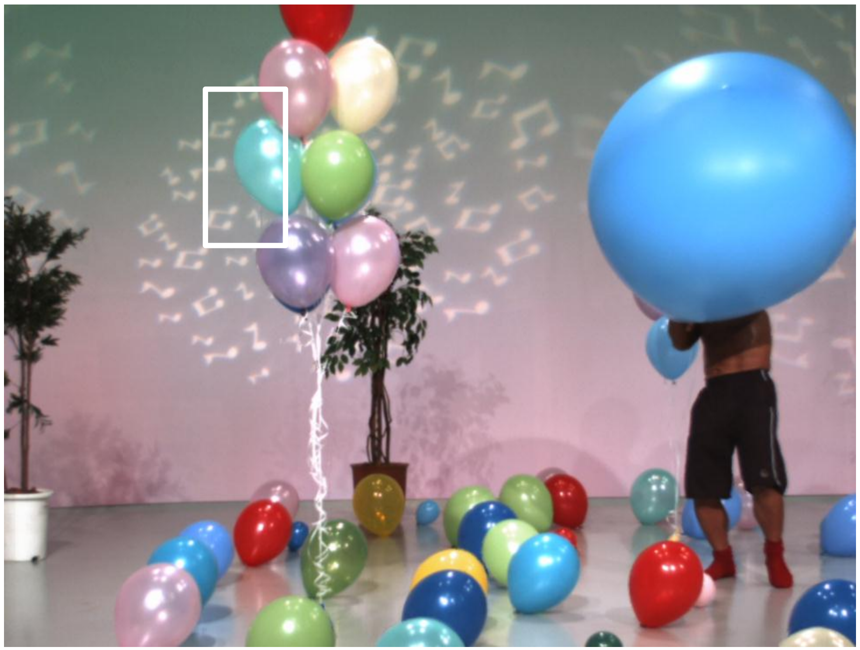
With VBDMM Mean-shift depth map

Subjective Results

Test sequence: Balloons



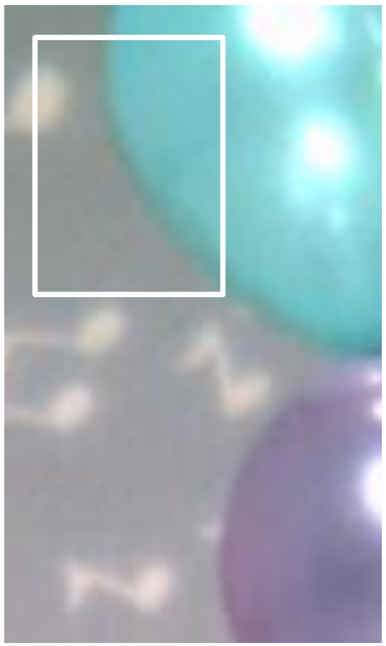
With MPEG depth map



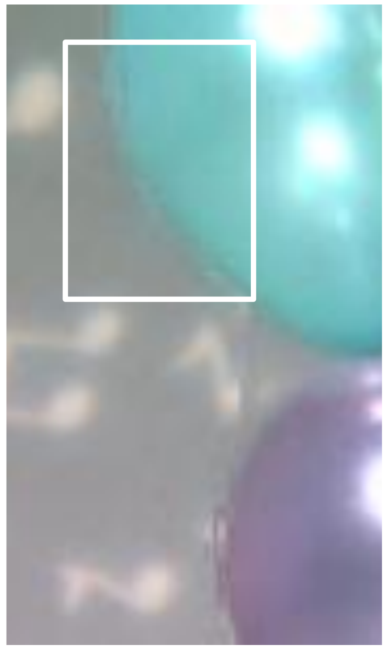
With VBDMM Mean-shift depth map

Subjective Results

Test sequence: Balloons

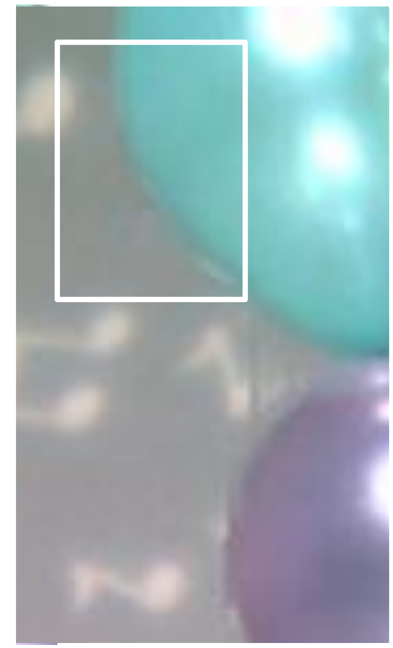


Original
depth maps



With MPEG
depth maps

With VBGMM
K-Means
depth maps



With VBDMM
Mean-Shift
depth maps

Subjective Results

Test sequence: Poznan Street

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective Results

Test sequence: Poznan Street

Original depth maps

With MPEG depth maps

With VBGMM K-Means depth maps

With VBDMM Mean-Shift depth maps

Subjective Results

Test sequence: Poznan Street

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective Results

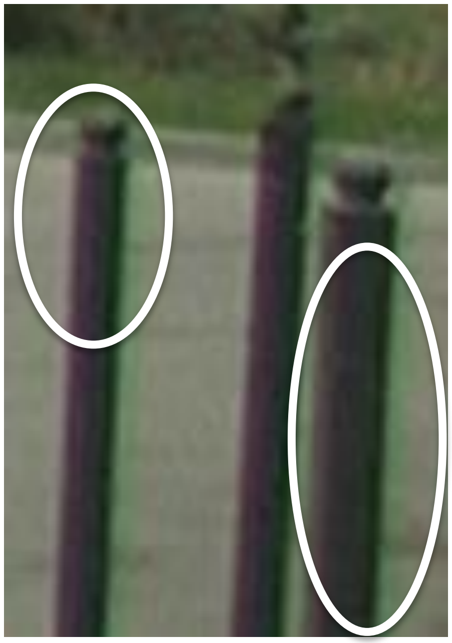
Test sequence: Poznan Street

With MPEG depth map

With VBDMM Mean-shift depth map

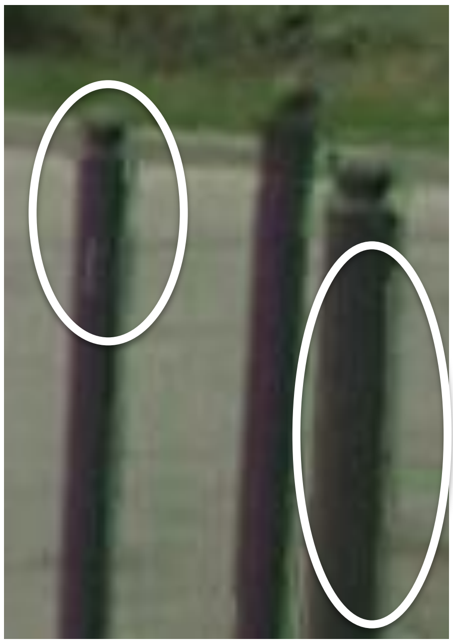
Subjective Results

Test sequence: Poznan Street

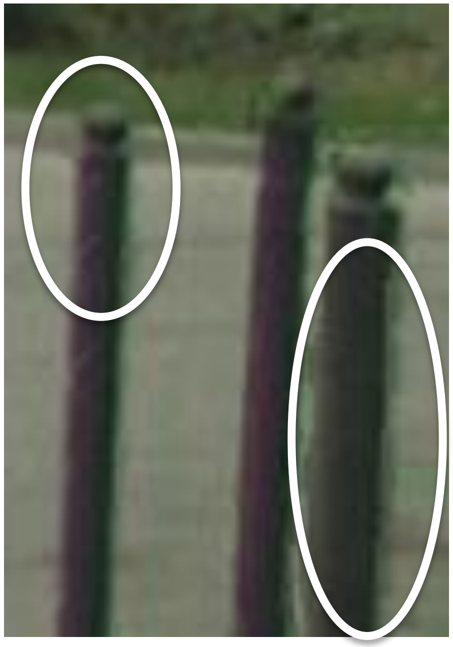


Original

With MPEG
depth maps



With VBGMM
K-Means
depth maps



With VBDMM
Mean-Shift
depth maps