Multiview Depth Map Enhancement by Variational Bayes Inference Estimation of Dirichlet Mixture Models

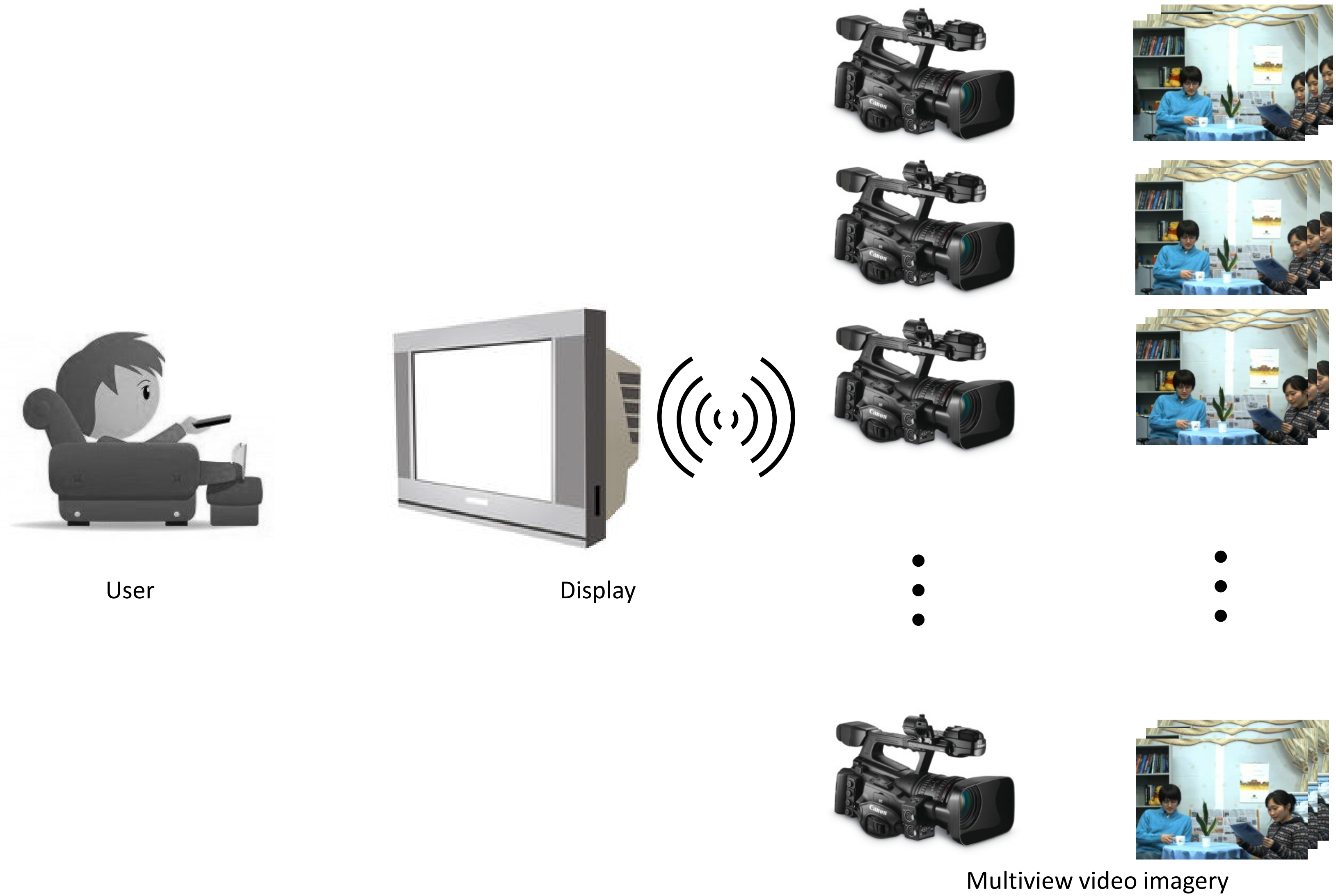
Pravin Kumar Rana, Zhanyu Ma, Jalil Taghia, and Markus Flierl

School of Electrical Engineering
KTH Royal Institute of Technology
Stockholm, Sweden

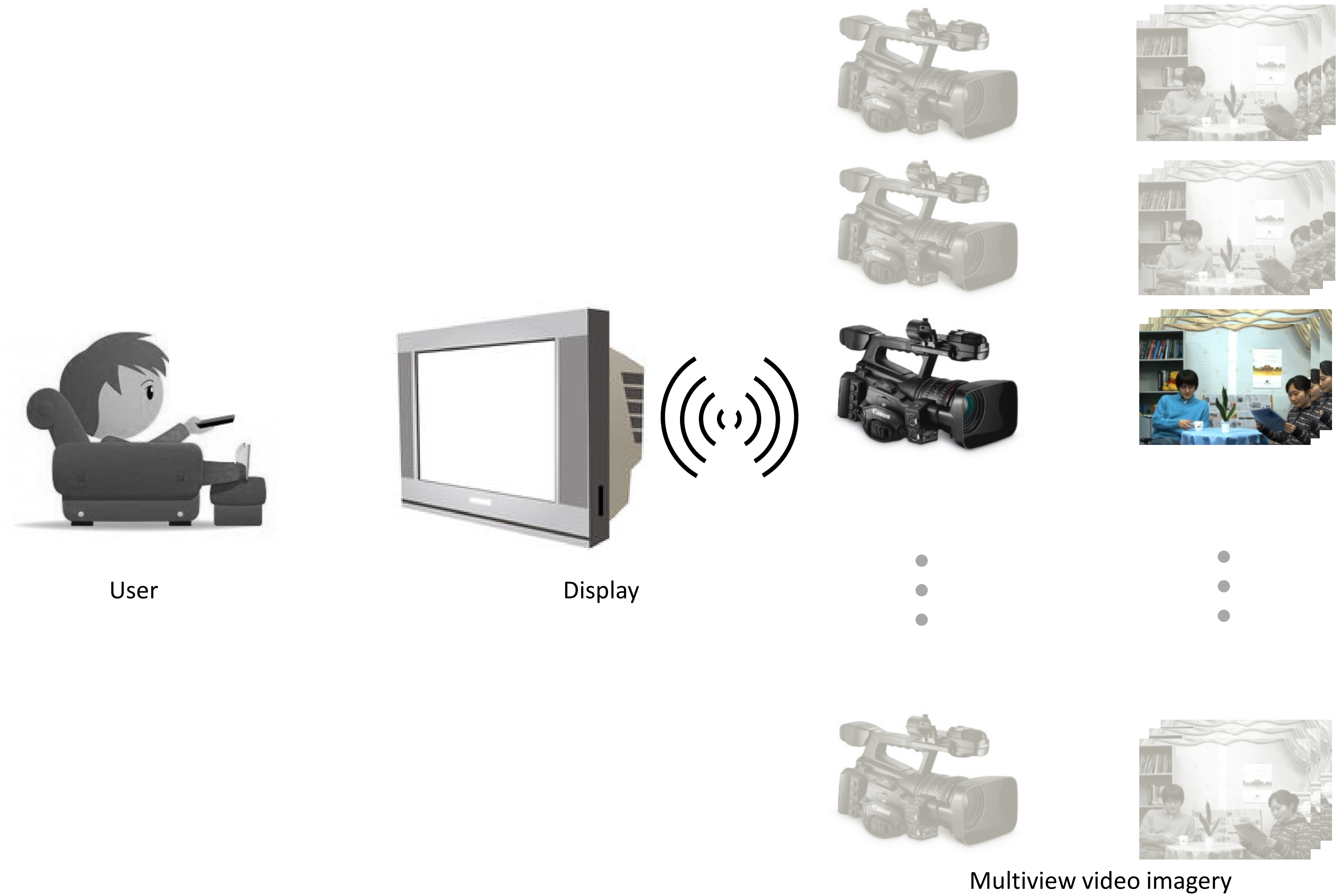
May 14, 2013

Background and motivation

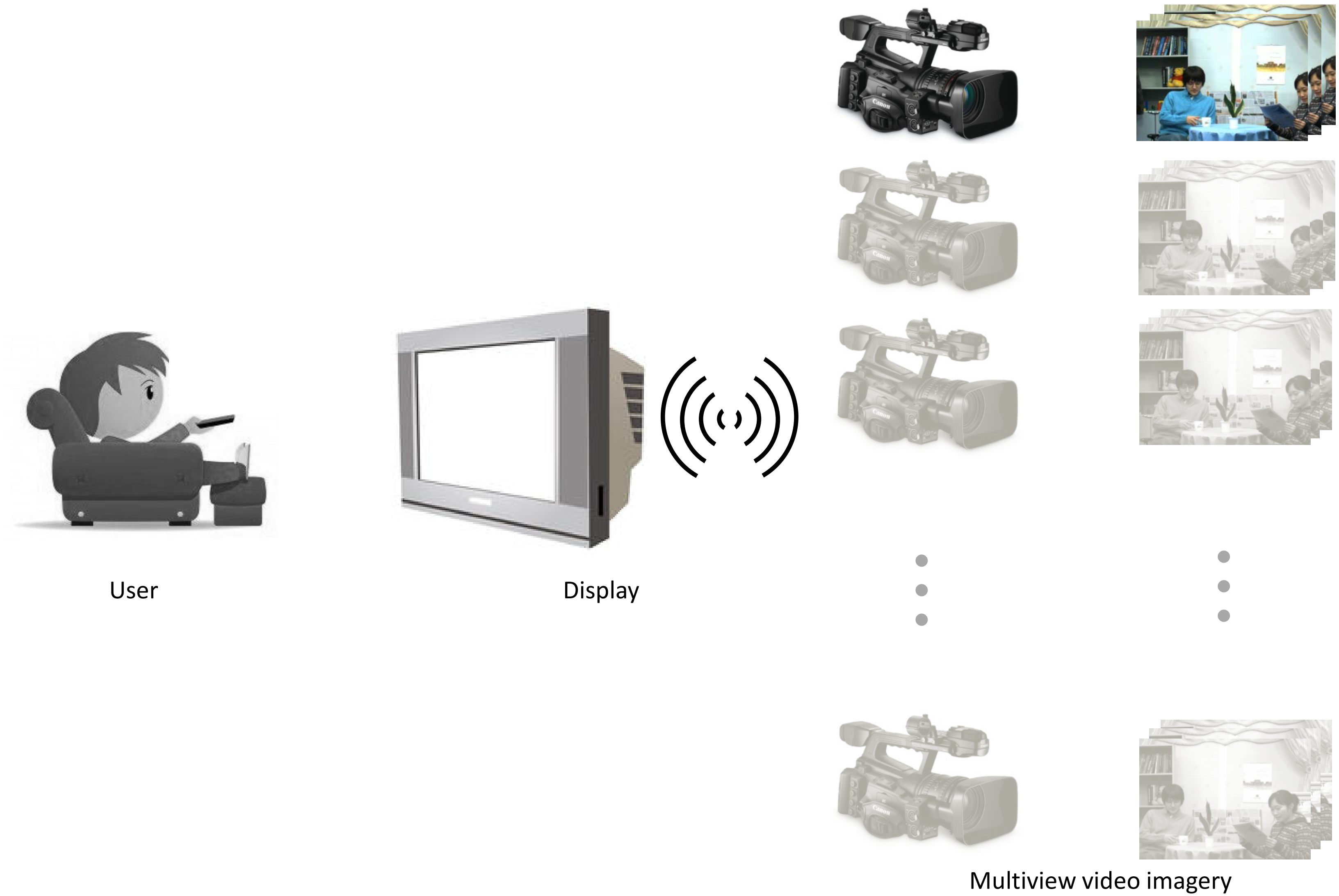
Free-viewpoint television



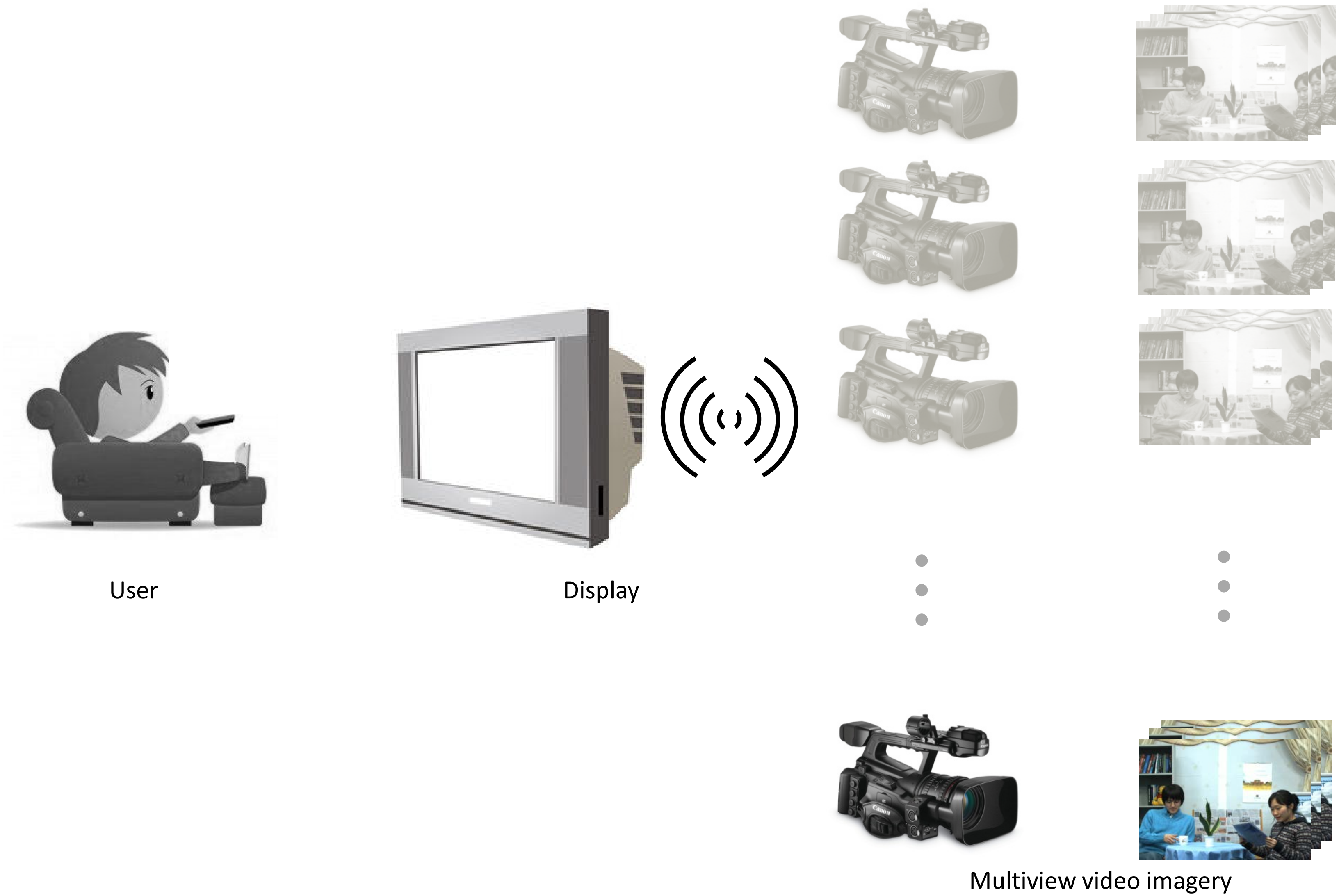
Free-viewpoint television



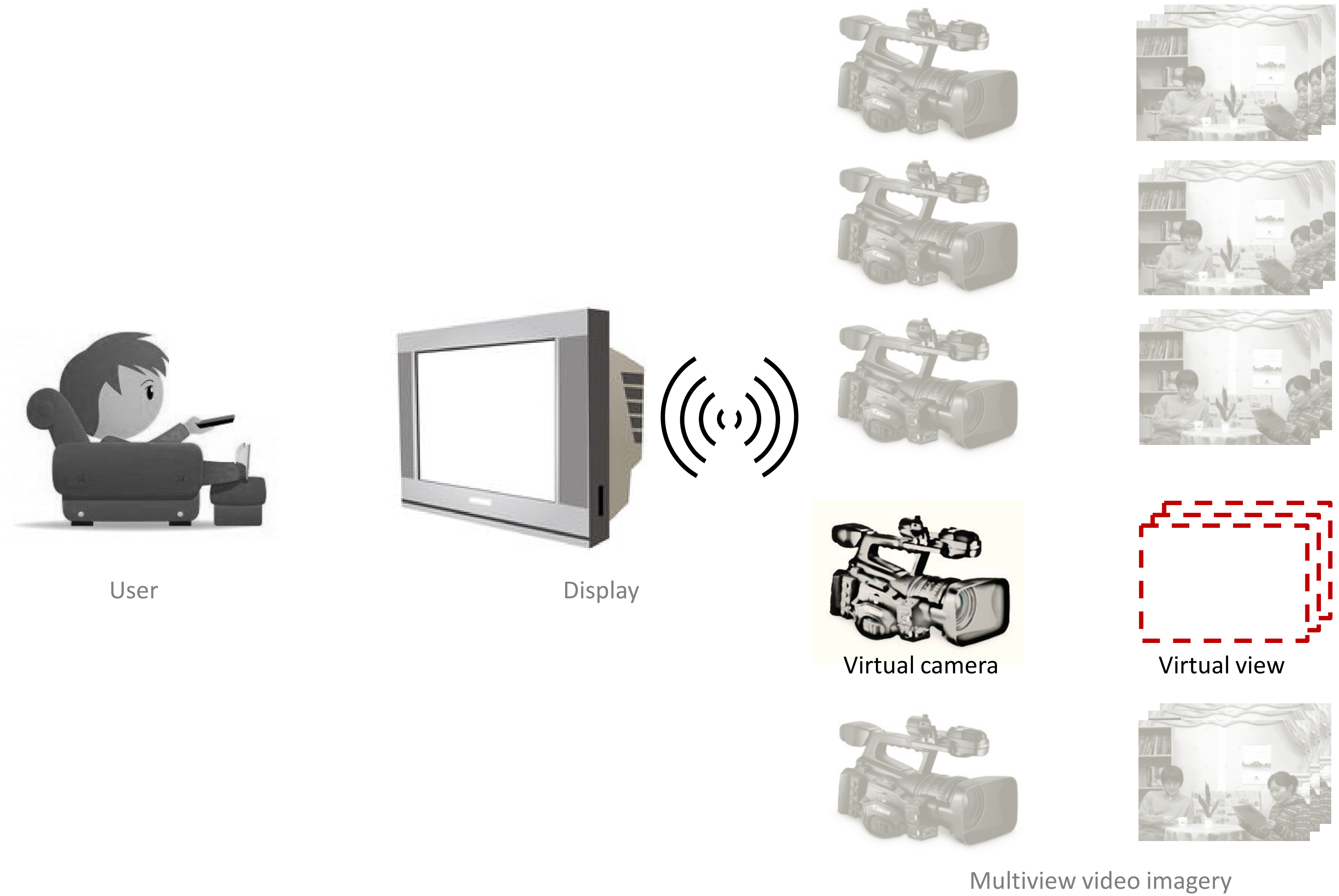
Free-viewpoint television



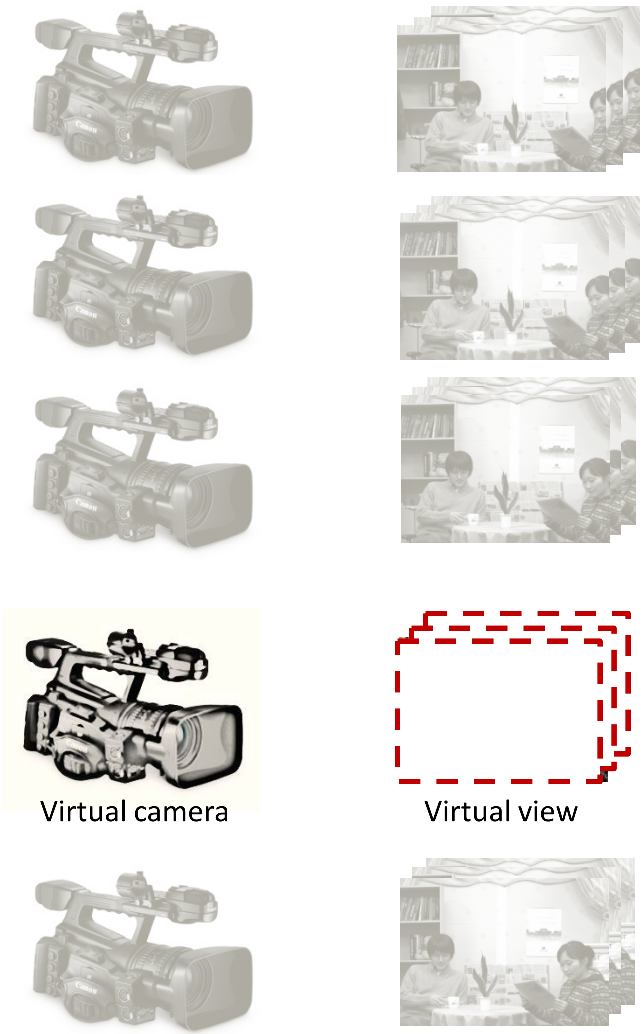
Free-viewpoint television



Free-viewpoint television

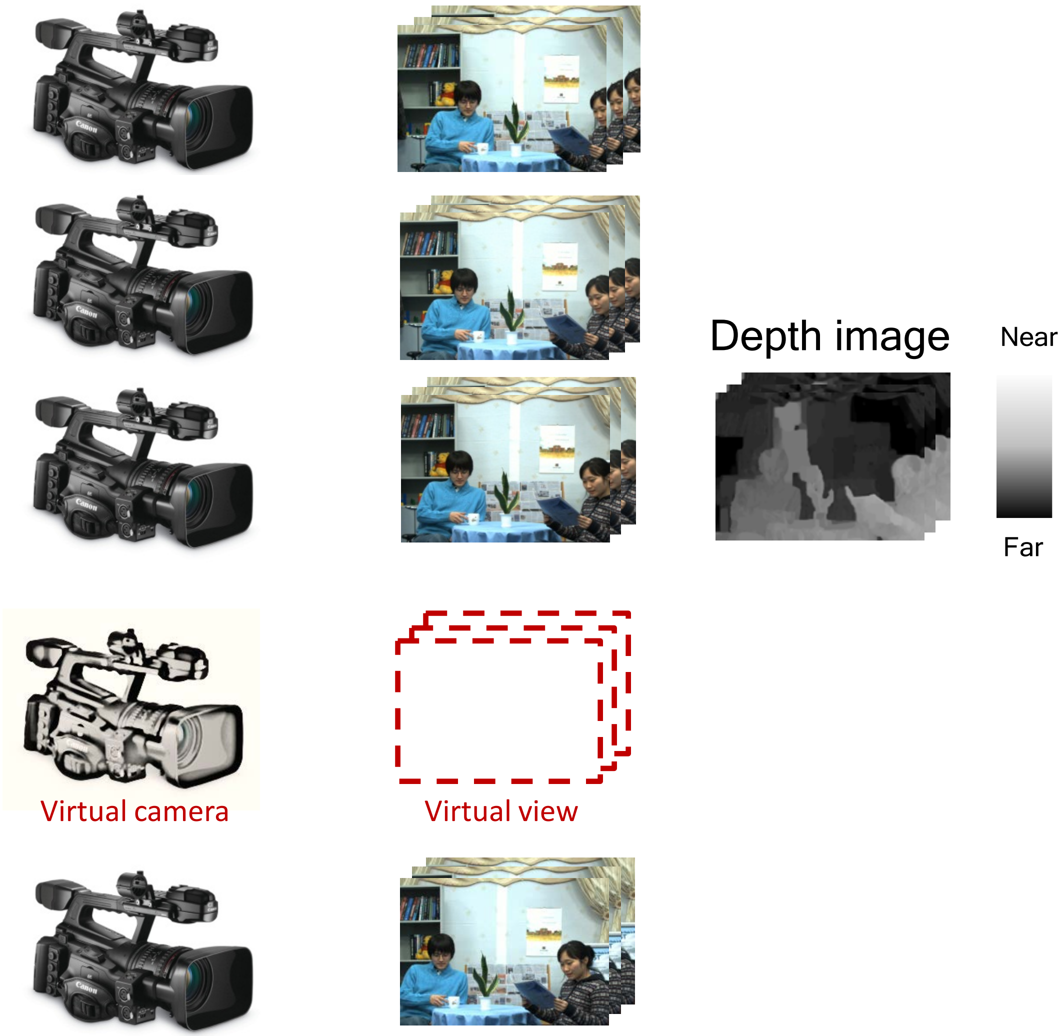


Depth image based rendering



Multiview video imagery

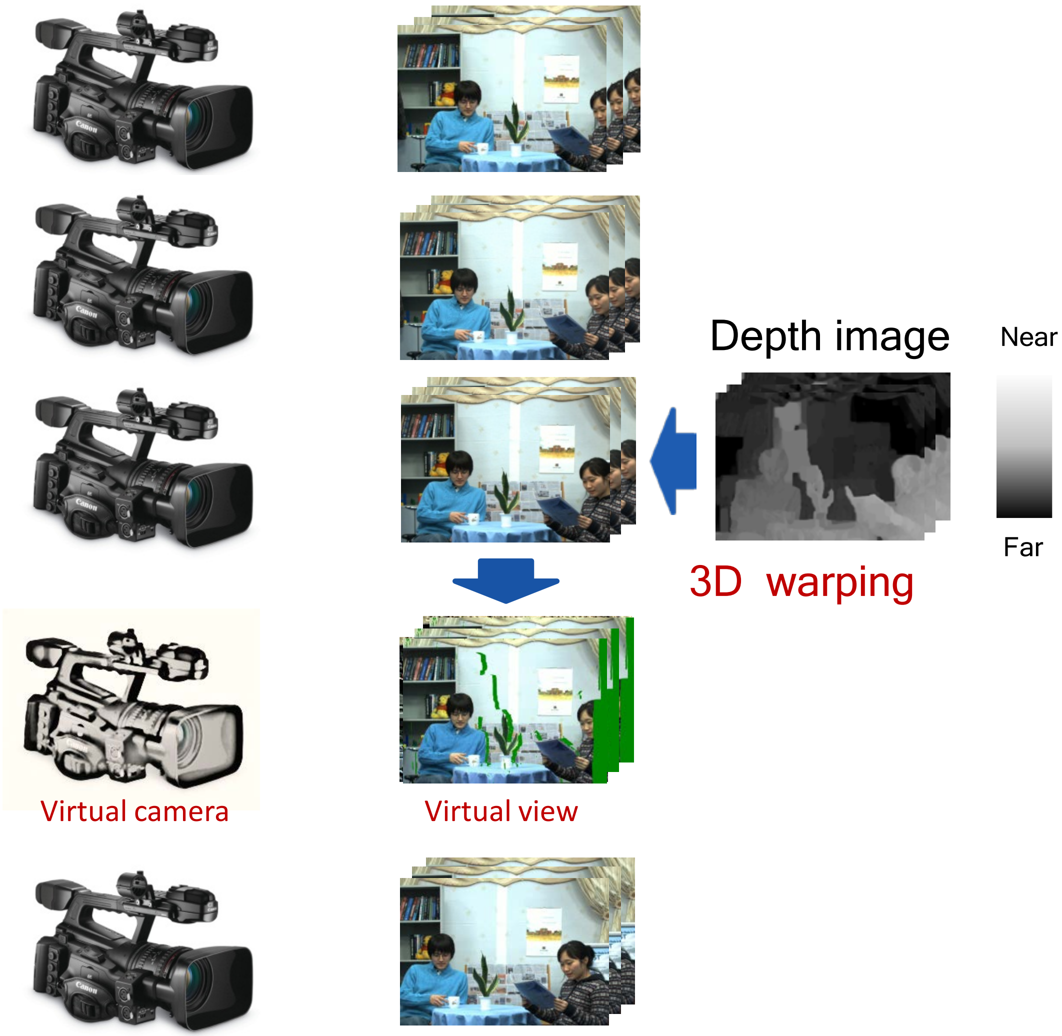
Depth image based rendering



Multiview video imagery

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

Depth image based rendering

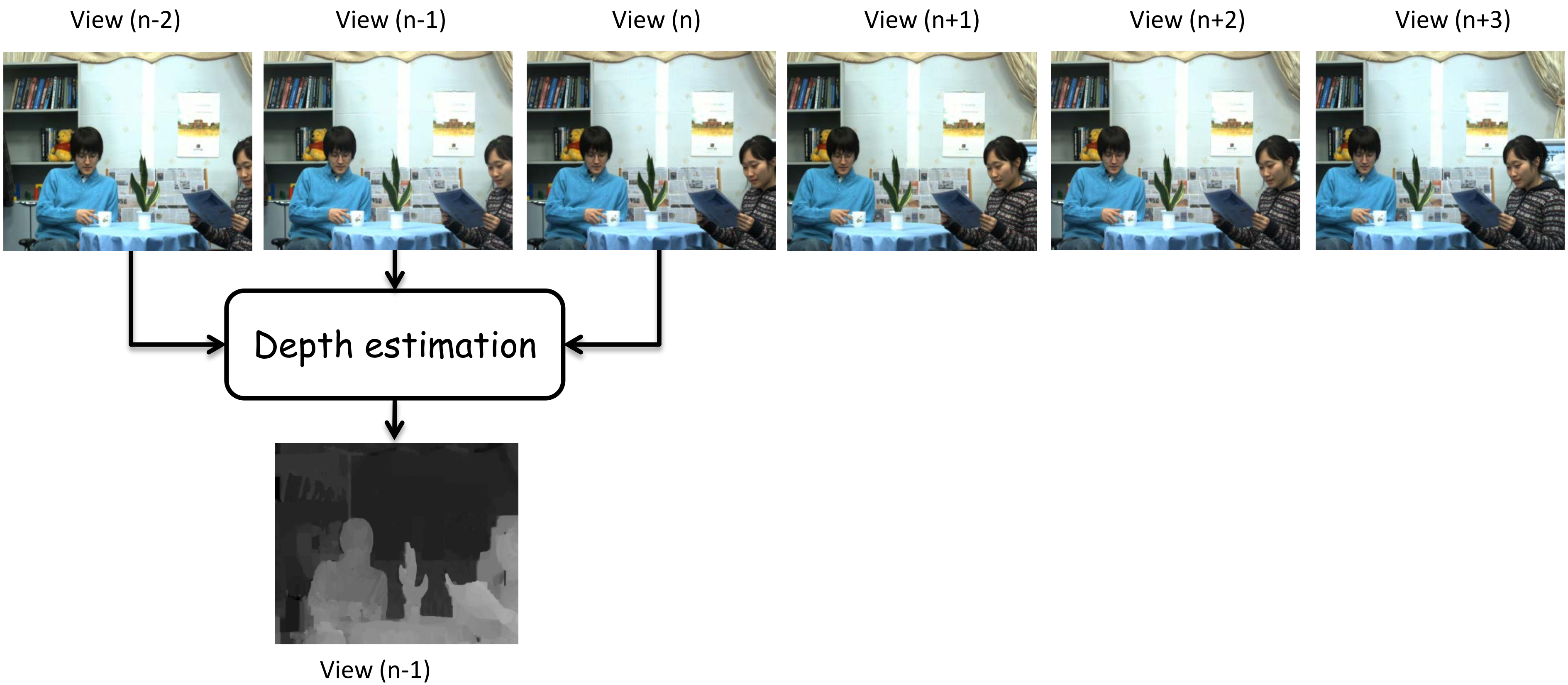


Multiview video imagery

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

Depth estimation

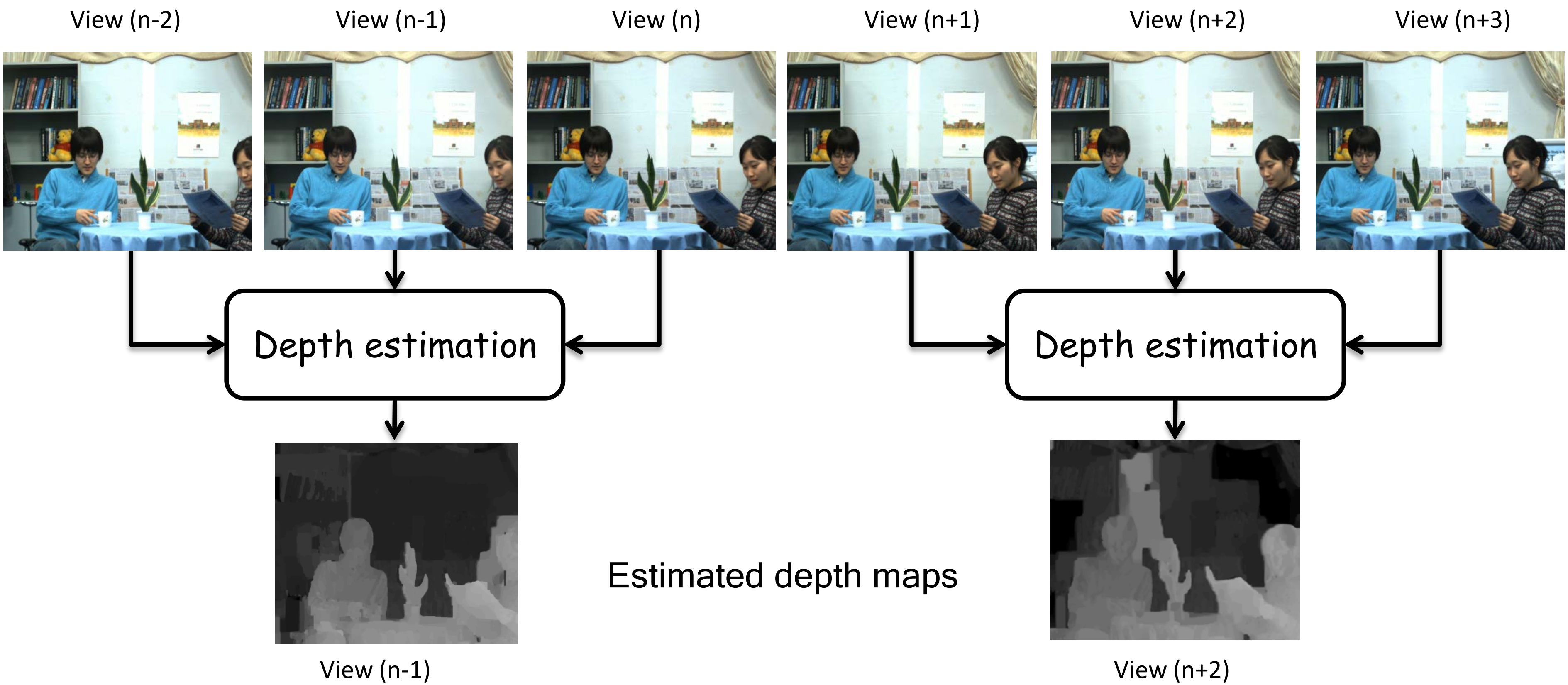
MPEG Depth Estimation Reference Software



Note: we assume a 1D-parallel camera arrangement

Depth estimation

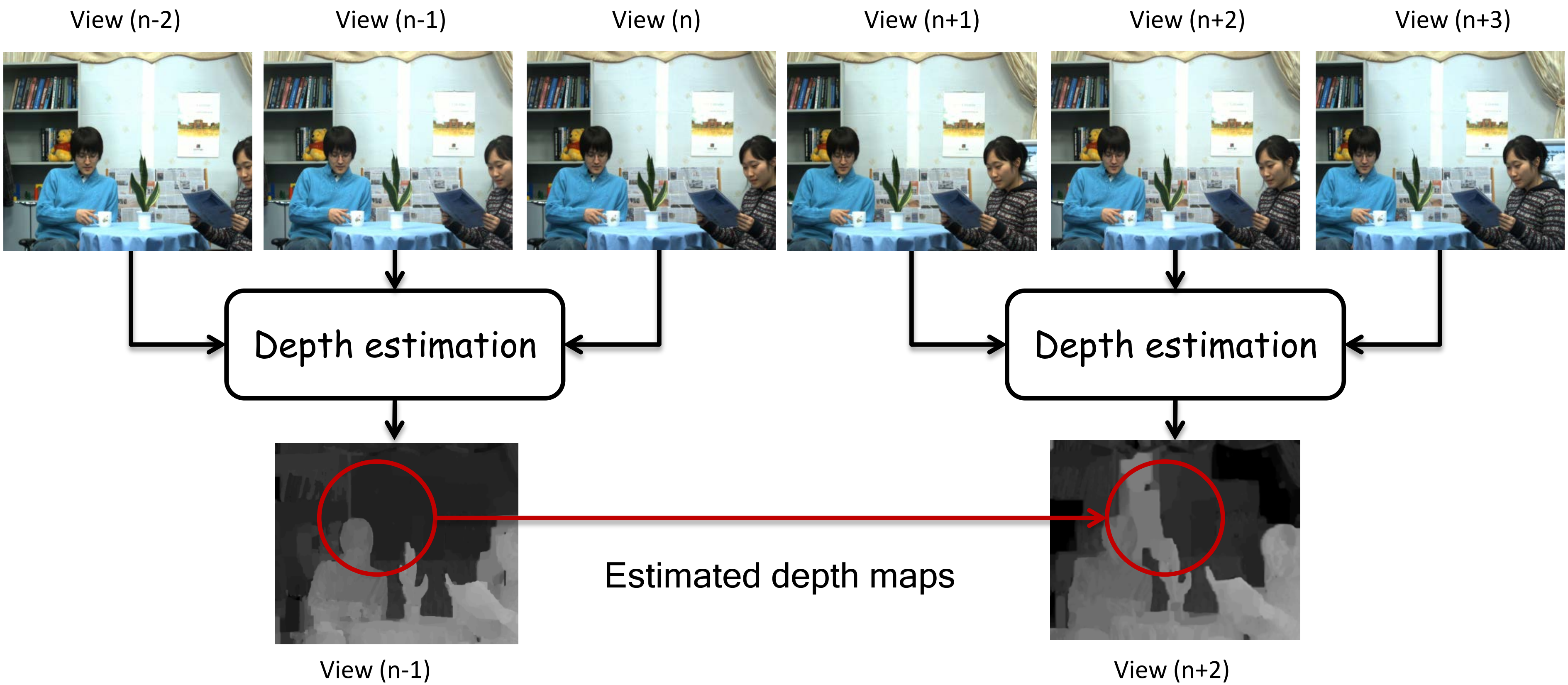
MPEG Depth Estimation Reference Software



Note: we assume a 1D-parallel camera arrangement

Depth estimation

MPEG Depth Estimation Reference Software



Problem: Inter-view depth inconsistency

Note: we assume a 1D-parallel camera arrangement

Improved depth enhancement framework

Prior work on depth enhancement

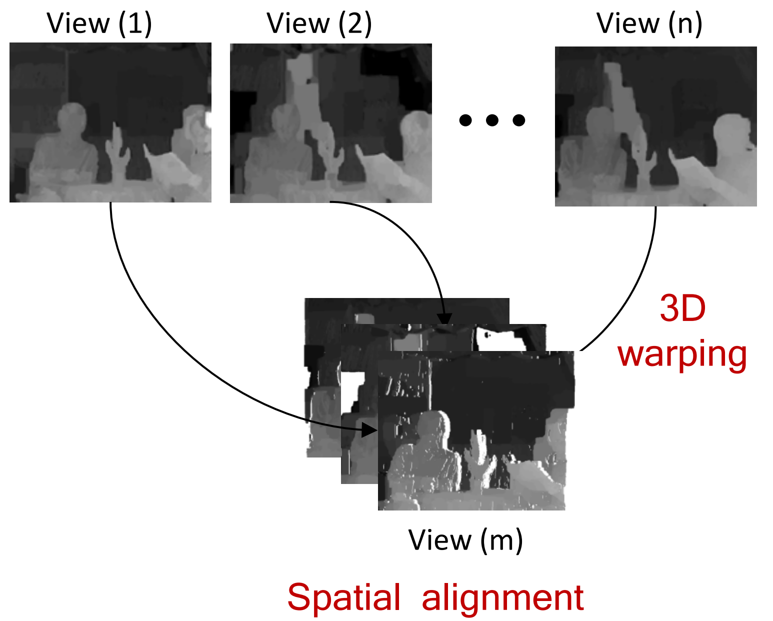
1. Existing methods warp depth images from multiple viewpoints to a common viewpoint for spatial alignment ([1], [2])

2. Warping errors due to the discrete values in depth maps affects enhancement algorithms negatively

3. Our approach ([3]):

   – Exploiting per-pixel associations between depth and color from various viewpoints

   – Use variational Bayes inference to classify color clusters in multiview imagery



[1] P. K. Rana and M. Flierl, "Depth consistency testing for improved view interpolation," IEEE Int. Workshop MMSP, 2010.

[2] E. Ekmekcioglu, V. Velisavljevic, and S. Worrall, "Content adaptive enhancement of multi-view depth maps for free viewpoint video," IEEE J. Sel. Topics Signal Process., 2011.

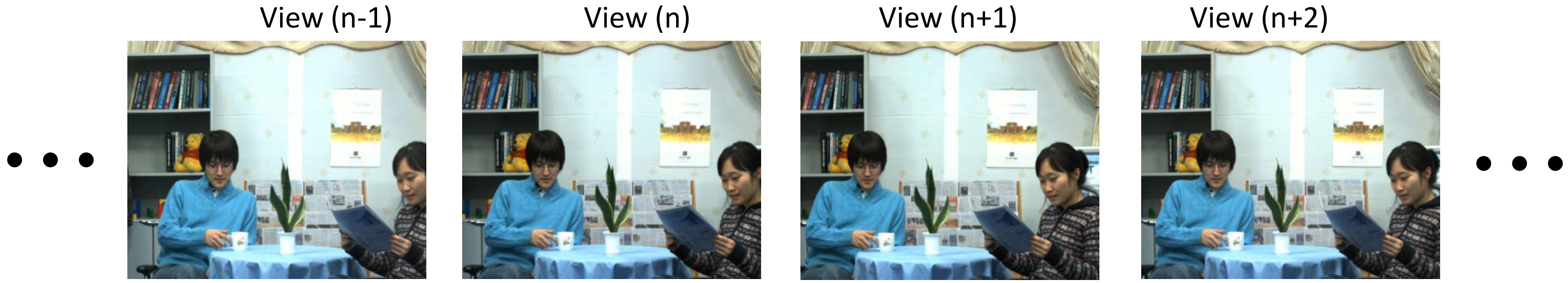
[3] P. K. Rana, J. Taghia, and M. Flierl, "A variational Bayesian inference framework for multiview depth image enhancement," IEEE Int. Symp. Multimedia (ISM), 2012

Overview of depth enhancement framework

- Concatenation of view imagery
- Multiview color classification
- Multiview depth classification
- Depth image enhancement

Concatenation of view imagery

- The captured MVV imagery of the scene has inherent inter-view similarity
- To have a unique model for the captured natural scene,
  - The MVV inter-view similarity is exploited by concatenating views from multiple viewpoints



Concatenation of view imagery

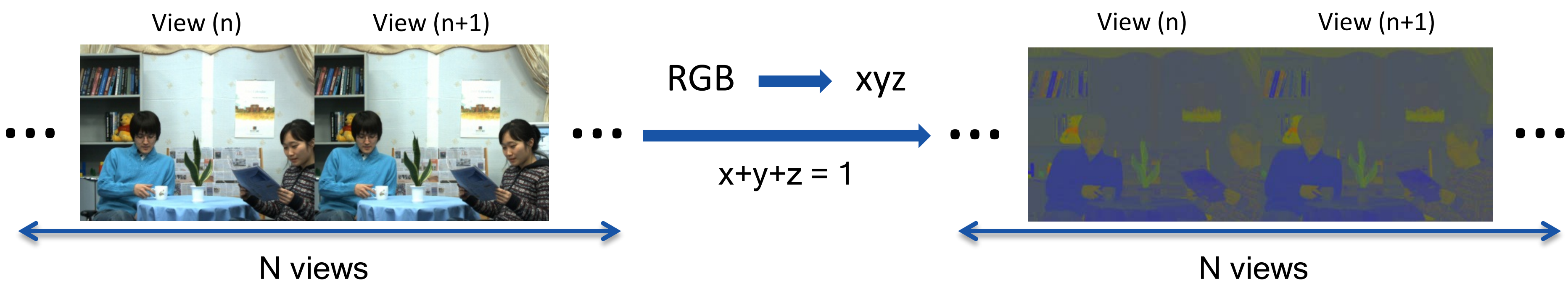
- The captured MVV imagery of the scene has inherent inter-view similarity
- To have a unique model for the captured natural scene,
  - The MVV inter-view similarity is exploited by concatenating views from multiple viewpoints

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

- Use the chromatic color representation to make the procedure insensitive to the absolut luminance
- The chromaticity of a pixel is described by a vector of three chromaticity coefficients $[x\ y\ z]^T$



Multiview color classification

- The goal of classification is to partition an image into regions each of which has a reasonably homogeneous visual appearance

- Usually, classification algorithm, such as expectation-maximization for Gaussian mixtures, suffers from two main drawbacks:
  - model over-fitting and
  - the number of clusters has to be known, (similar to the K-means algorithm)

- With variational Bayes inference (VBI) [4] because
  - no model over-fitting,
  - the number of clusters is treated as a random variable

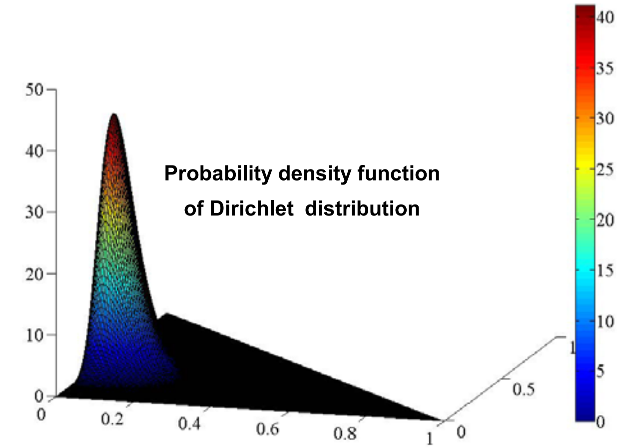
[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

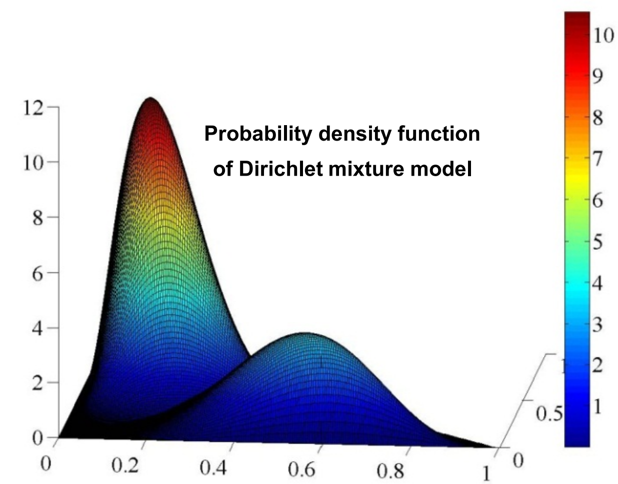
Multiview color classification

Dirichlet mixture model with variational Bayes inference (VBI)

- The vector of image pixels has nonnegative elements and is bounded
  - it can be efficiently modeled by utilizing non-Gaussian distributions [5]

- Based on the pixel vector's properties,
  - assume that the pixel vectors of each cluster are Dirichlet distributed





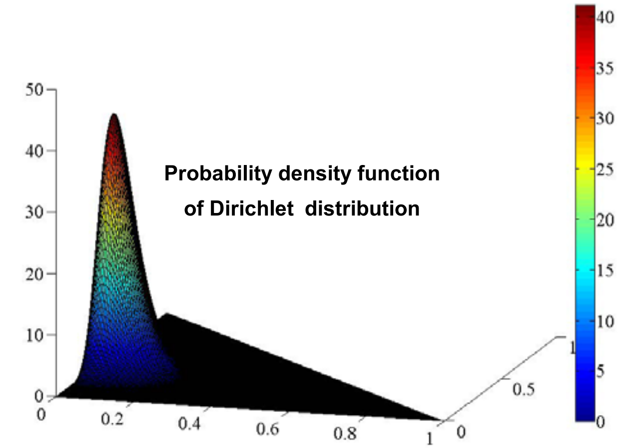
[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

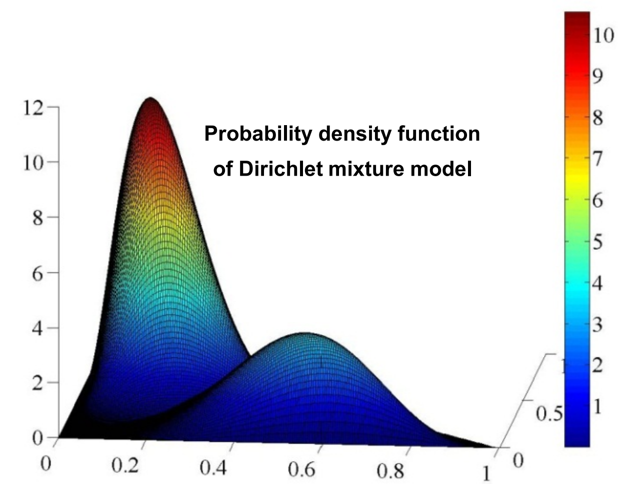
[5] Z. Ma, P. K. Rana, J. Taghia, M. Flierl, and A. Leijon, "Bayesian estimation of Dirichlet mixture model with variational inference," IEEE Trans. PAMI, submitted, 2013.

Multiview color classification

Dirichlet mixture model with variational Bayes inference (VBI)

- The vector of image pixels has nonnegative elements and is bounded
  - it can be efficiently modeled by utilizing non-Gaussian distributions [5]
- Based on the pixel vector's properties,
  - assume that the pixel vectors of each cluster are Dirichlet distributed
- Use Dirichlet mixture model (DMM) with VBI to capture the all underlying clusters in multiview imagery [5]
- It reduces the model complexity when compare to Gaussian mixture model with VBI





[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.
[5] 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.

Multiview color classification

Newspaper

Balloons

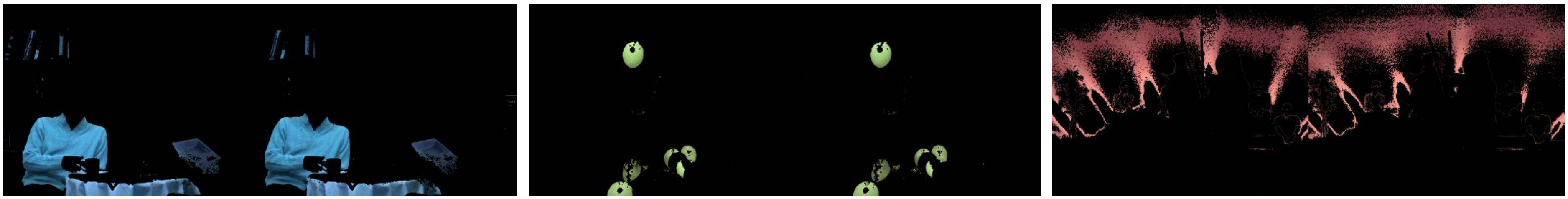
Kendo

Input multiview data

Multiview color classification

Newspaper | Balloons | Kendo

Input multiview data

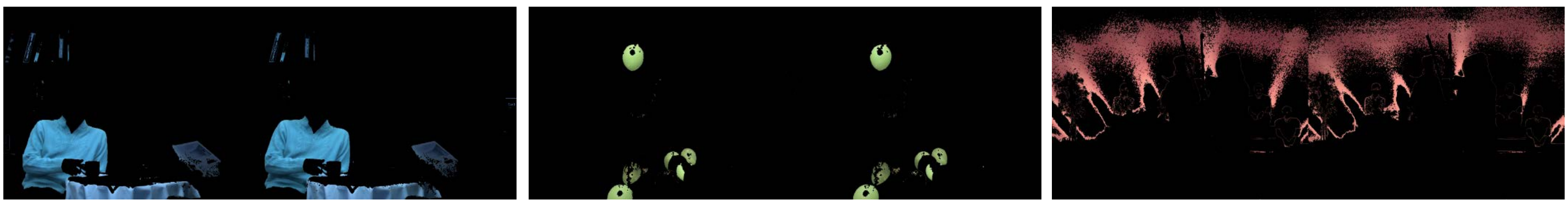


Using Dirichlet mixture model with variational Bayes inference in xyz space

Multiview color classification

Newspaper | Balloons | Kendo

Input multiview data



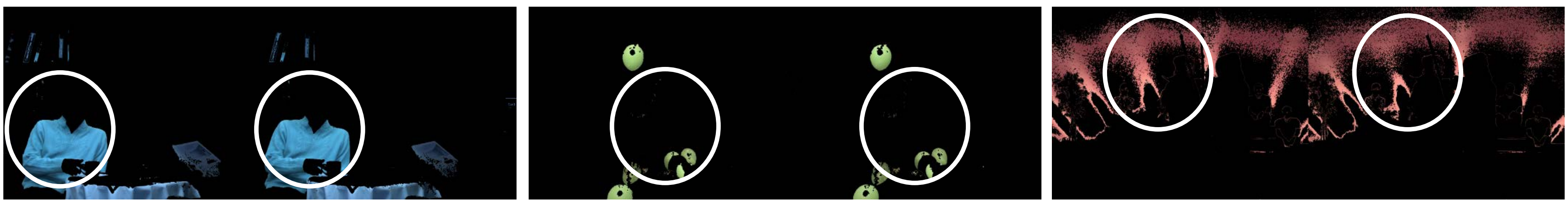
Using Dirichlet mixture model with variational Bayes inference in xyz space

Using Gaussian mixture model with variational Bayes inference in RGB space

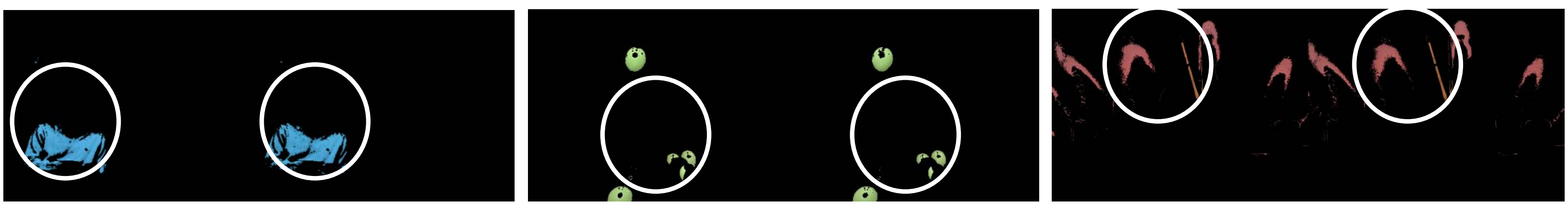
Multiview color classification

Newspaper | Balloons | Kendo

Input multiview data



Using Dirichlet mixture model with variational Bayes inference in xyz space



Using Gaussian mixture model with variational Bayes inference in RGB space

Multiview depth classification

Exploiting the per-pixel association between color and depth

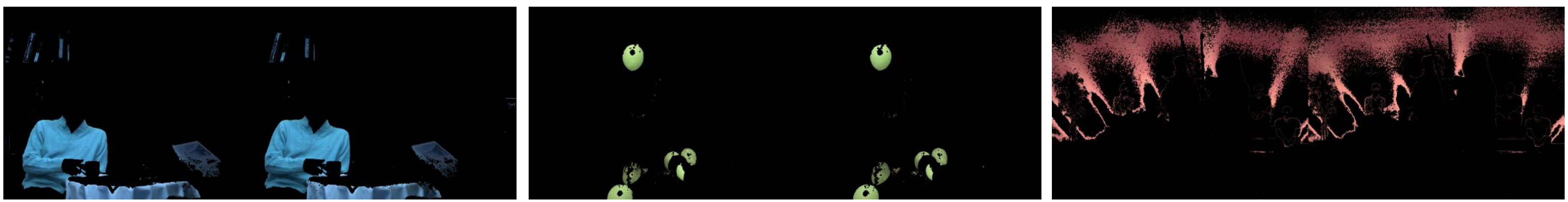
View image Depth image

Concatenated view imagery Concatenated depth imagery

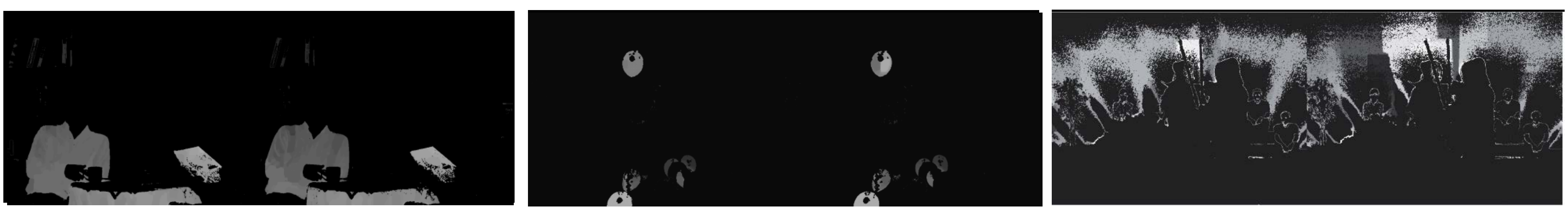
Multiview depth classification

Newspaper Balloons Kendo

Input multiview data



Using Dirichlet mixture model with variational Bayes inference in xyz space



Multiview depth enhancement

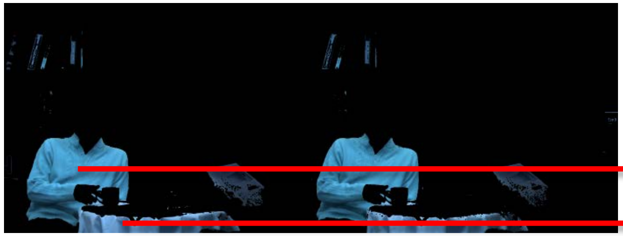
Difference between color and depth clusters

Members have similar colors pixels

Members may have different depth values

Multiview depth enhancement

Difference between color and depth clusters

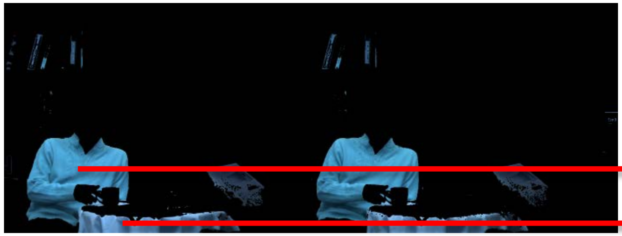


Members have similar colors pixels

Members may have different depth values

Multiview depth enhancement

Difference between color and depth clusters



Members have similar colors pixels

Members may have different depth values

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

Multiview depth enhancement

Difference between color and depth clusters

Members have similar colors pixels

Members may have different depth values

- Why?
  - due to foreground and background depth difference
  - due to inter-view inconsistency
- Our approach: Means-shift sub-clustering [5]
  - a nonparametric clustering technique
  - does not require prior knowledge of the number of clusters
  - does not constrain the shape of the clusters
  - assigns the mean to depth pixels irrespective of the originating viewpoints
  - Bayesian 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 setup

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)

Model complexity

- By measuring the model complexity in terms of the number of free parameters:
  - VBI-DMM requires a smaller model complexity than the VBI-GMM

Input Vector	Initial number of mixture components	Number of free parameters	
		VBI-GMM	VBI-DMM
D	I	I(2D + 1) −1	I(D + 2) −1

- In experiment,
  - Initial number of mixture components = 100

Multiview Data Set	Active number of mixture components	
	VBI-GMM	VBI-DMM
Lovebird1	31	24
Kendo	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

Left

Enhanced depth map

Reference view

Right

Enhanced depth map

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

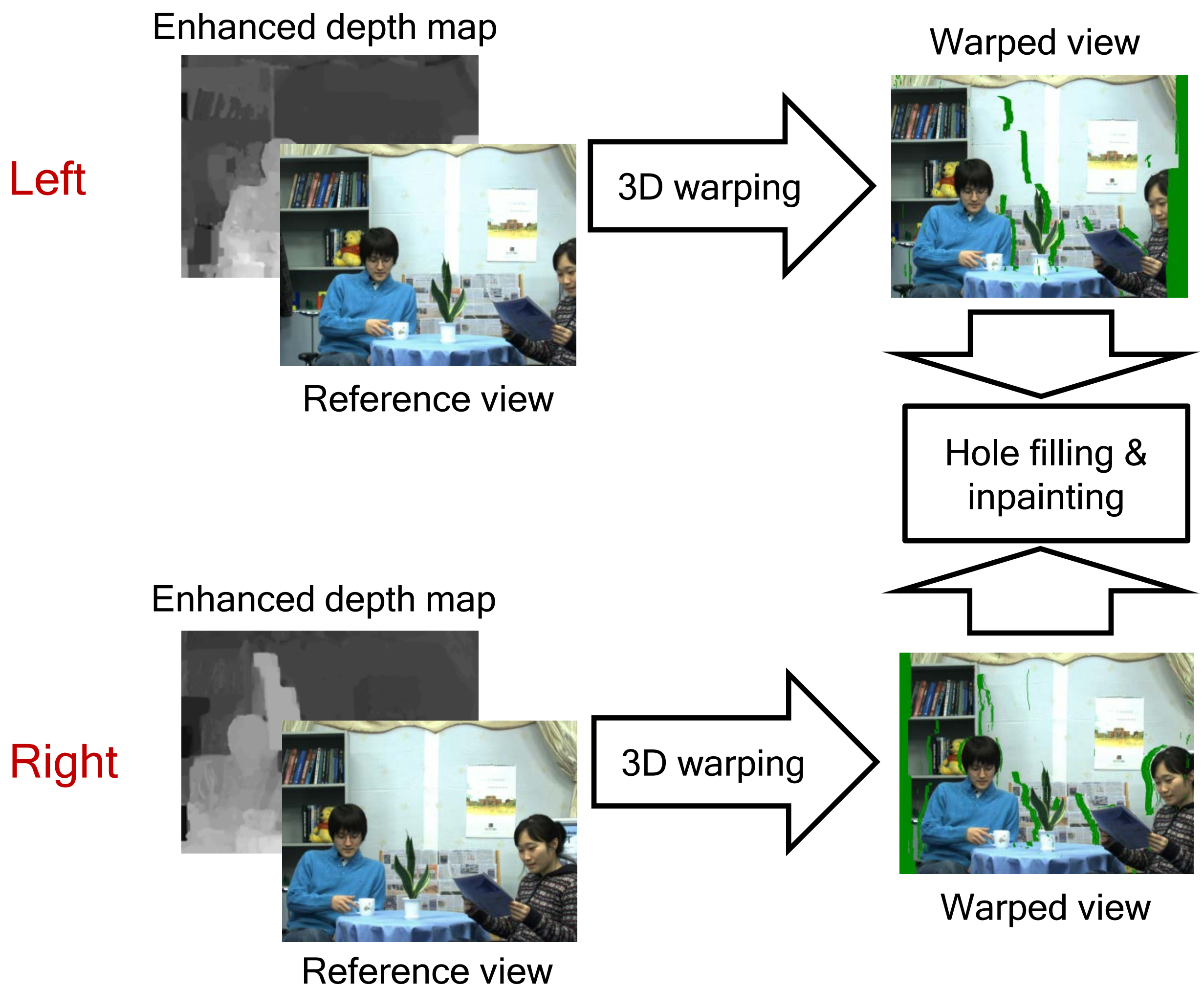
Reference view

3D warping

Warped view

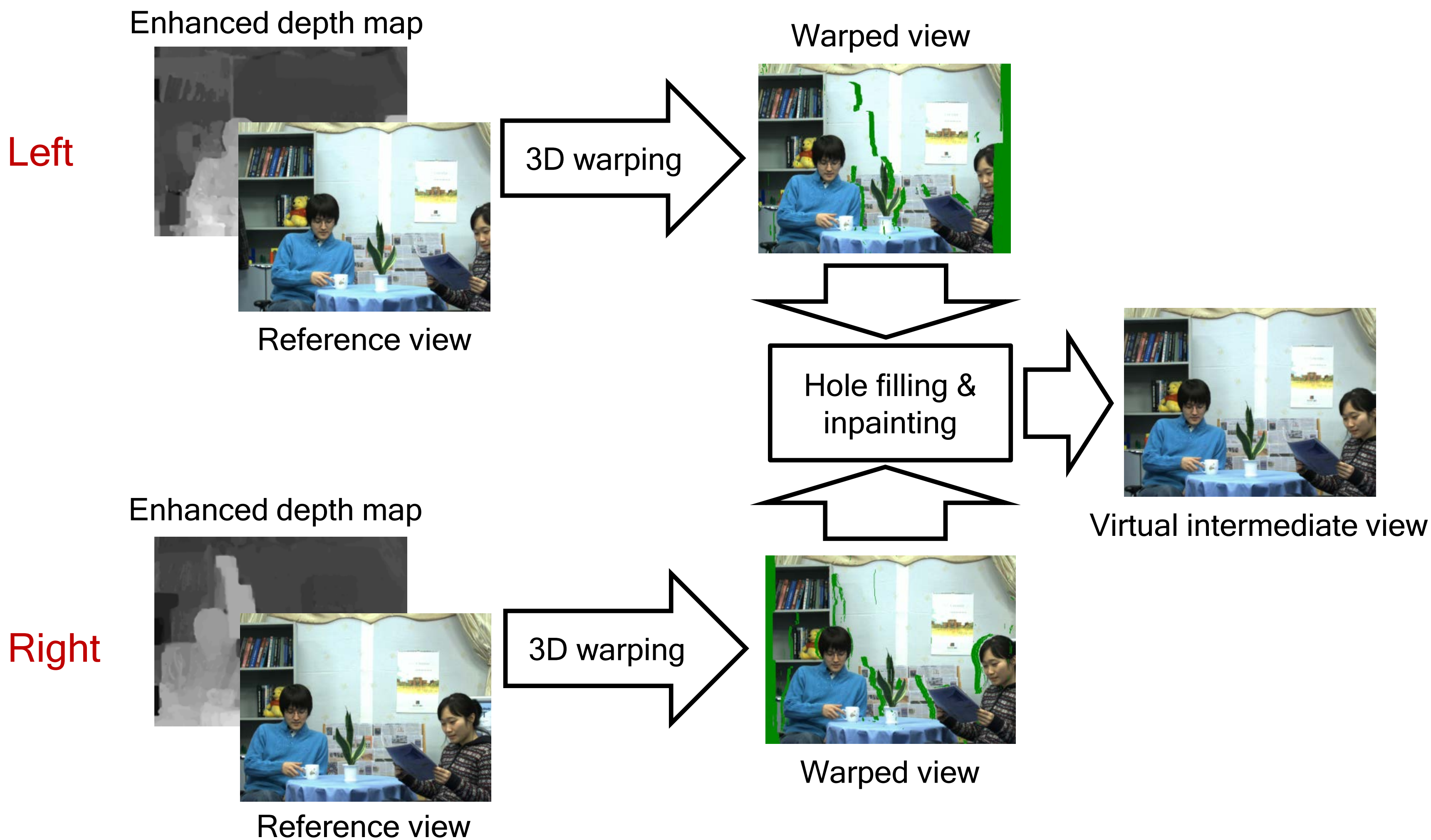
Depth image based rendering

MPEG View Synthesis Reference Software (VSRS) 3.5



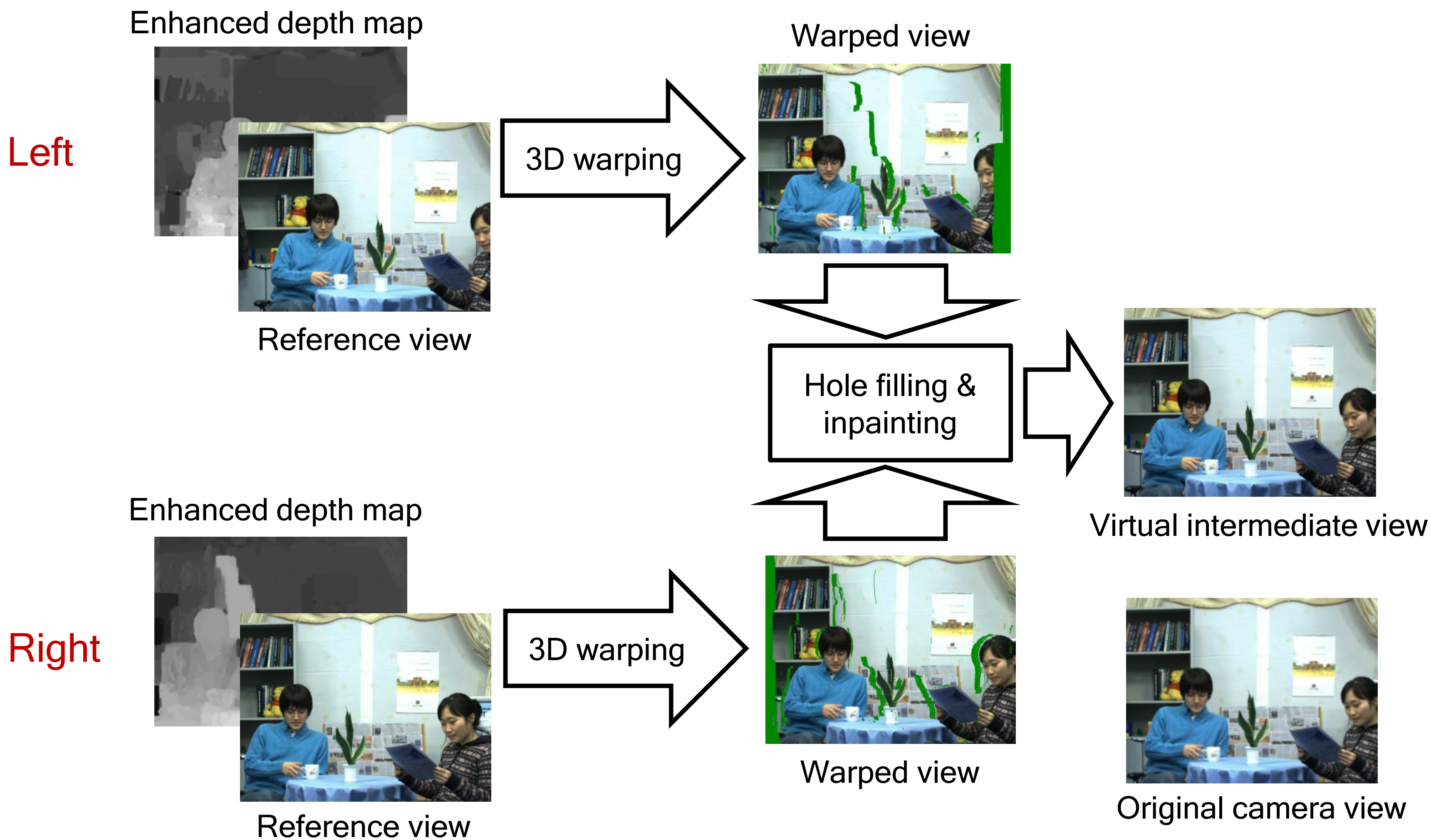
Depth image based rendering

MPEG View Synthesis Reference Software (VSRS) 3.5



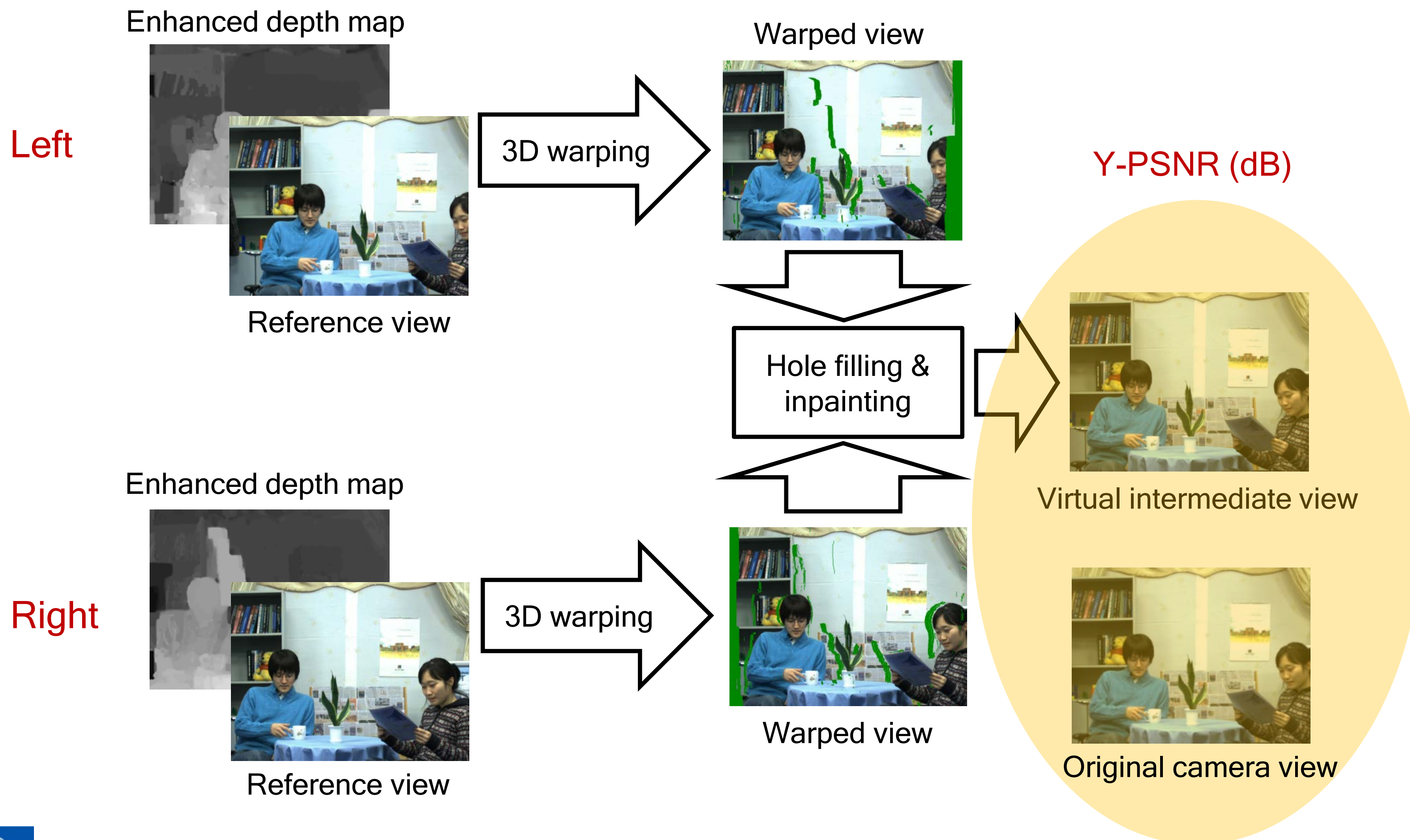
Depth image based rendering

MPEG View Synthesis Reference Software (VSRS) 3.5



Depth image based rendering

MPEG View Synthesis Reference Software (VSRS) 3.5



Objective results

Test sequence	Input views	Virtual view	MPEG VSRS 3.5 [dB]		
			MPEG depth maps	VBGMM K-Means depth maps	VBDMM Mean-shift depth maps
Newspaper	4,6	5	32.00	32.10	32.11
Kendo	3,5	5	36.54	36.72	39.35
Poznan Street	3,5	4	35.56	35.58	35.72
Lovebird1	6,8	7	28.50	28.68	29.04
Balloons	3,5	4	35.69	35.93	36.02

- Color classification
  - Initial number of mixture components = 100
- K-means sub-clustering
  - Number of cluster : 12

Objective results

Test sequence	Input views	Virtual view	MPEG VSRS 3.5 [dB]		
			MPEG depth maps	VBGMM K-Means depth maps	VBDMM Mean-shift depth maps
Newspaper	4,6	5	32.00	32.10	32.11
Kendo	3,5	5	36.54	36.72	39.35
Poznan Street	3,5	4	35.56	35.58	35.72
Lovebird1	6,8	7	28.50	28.68	29.04
Balloons	3,5	4	35.69	35.93	36.02

- Color classification
  - Initial number of mixture components = 100
- K-means sub-clustering
  - Number of cluster : 12

Objective results

Test sequence	Input views	Virtual view	MPEG VSRS 3.5 [dB]		
			MPEG depth maps	VBGMM K-Means depth maps	VBDMM Mean-shift depth maps
Newspaper	4,6	5	32.00	32.10	32.11
Kendo	3,5	5	36.54	36.72	39.35
Poznan Street	3,5	4	35.56	35.58	35.72
Lovebird1	6,8	7	28.50	28.68	29.04
Balloons	3,5	4	35.69	35.93	36.02

- Color classification
  - Initial number of mixture components = 100
- K-means sub-clustering
  - Number of cluster : 12

Objective results

Test sequence	Input views	Virtual view	MPEG VSRS 3.5 [dB]		
			MPEG depth maps	VBGMM K-Means depth maps	VBDMM Mean-shift depth maps
Newspaper	4,6	5	32.00	32.10	32.11
Kendo	3,5	5	36.54	36.72	39.35
Poznan Street	3,5	4	35.56	35.58	35.72
Lovebird1	6,8	7	28.50	28.68	29.04
Balloons	3,5	4	35.69	35.93	36.02

- Color classification
  - Initial number of mixture components = 100
- K-means sub-clustering
  - Number of cluster : 12

Subjective results

Test sequence: Kendo

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective results

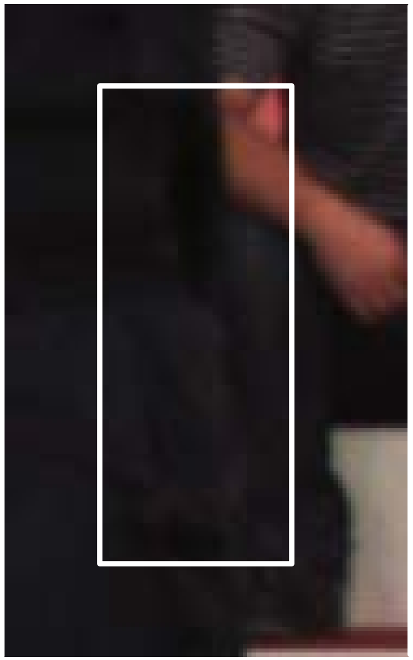
Test sequence: Kendo

With MPEG depth map

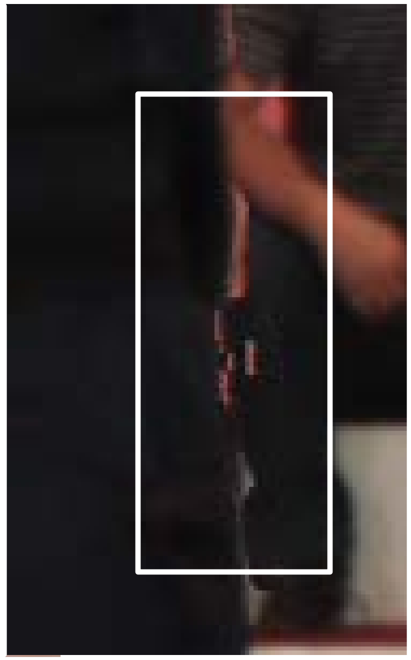
With VBDMM Mean-shift depth map

Subjective results

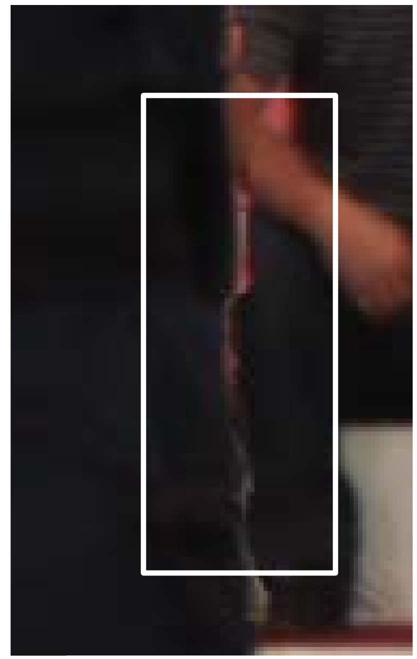
Test sequence: Kendo



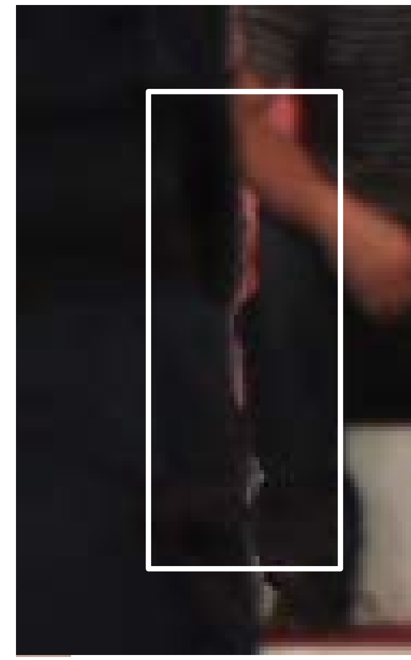
Original
depth maps



With MPEG
depth maps



With VBGMM
K-Means
depth maps



With VBDMM
Mean-Shift
depth maps

Subjective results

Test sequence: Lovebird 1

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective results

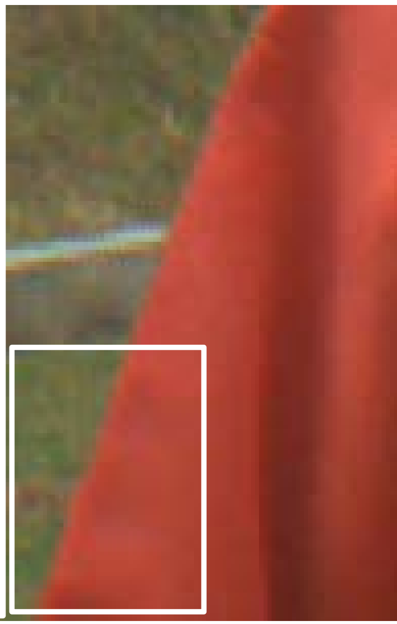
Test sequence: Lovebird 1


With MPEG depth map

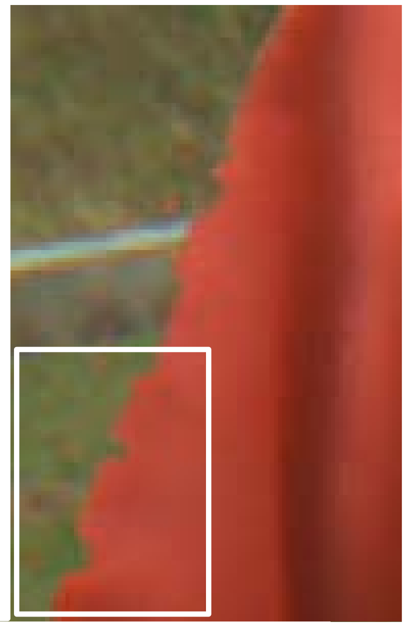

With VBDMM Mean-shift depth map

Subjective results

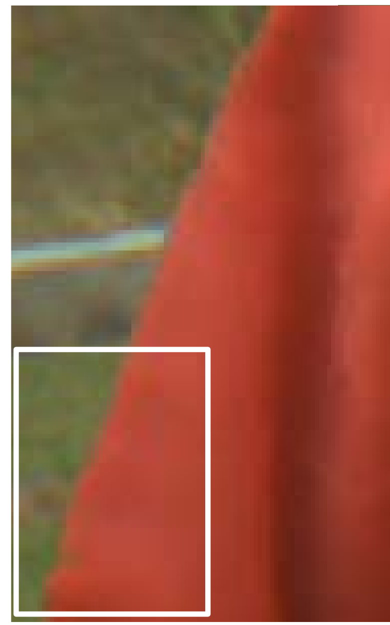
Test sequence: Lovebird 1



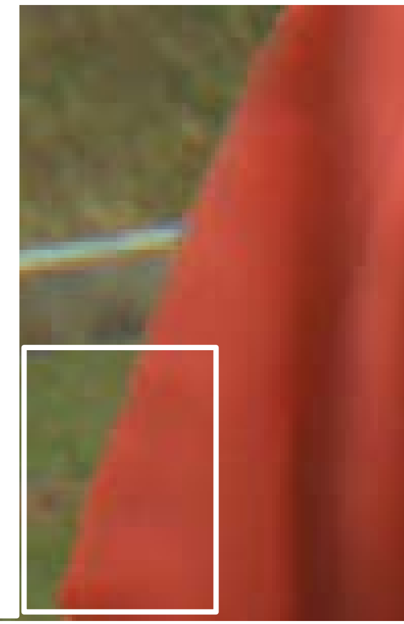
Original
depth maps



With MPEG
depth maps



With VBGMM
K-Means
depth maps



With VBDMM
Mean-Shift
depth maps

Subjective results

Test sequence: Lovebird 1

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective results

Test sequence: Lovebird 1

Original depth maps | With MPEG depth maps | With VBGMM K-Means depth maps | With VBDMM Mean-Shift depth maps

Subjective results

Test sequence: Newspaper

With MPEG depth map

With VBDMM Mean-shift depth map

Subjective results

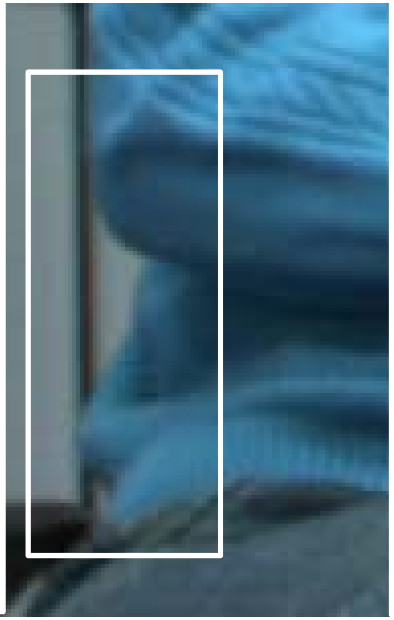
Test sequence: Newspaper

With MPEG depth map

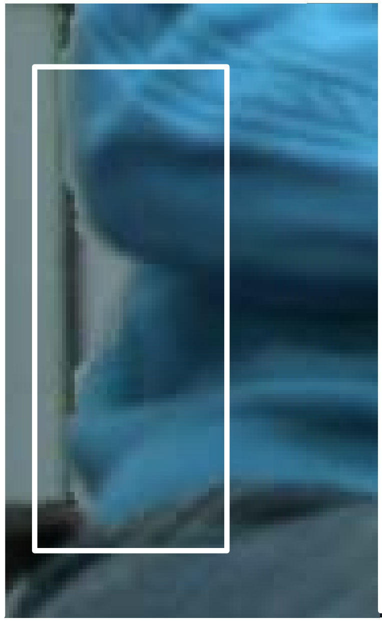
With VBDMM Mean-shift depth map

Subjective results

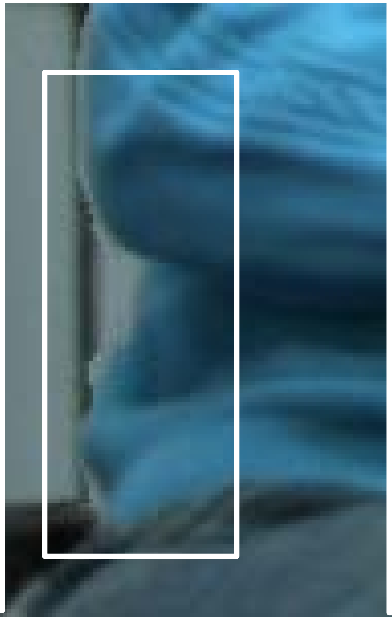
Test sequence: Newspaper



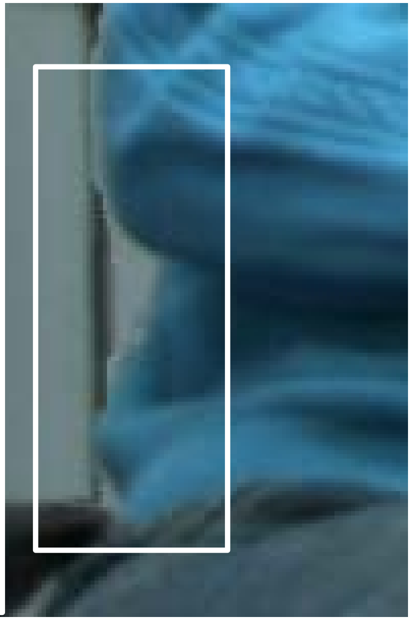
Original
depth maps



With MPEG
depth maps



With VBGMM
K-Means
depth maps



With VBDMM
Mean-Shift
depth maps

Subjective results

Test sequence: Newspaper

With MPEG depth map

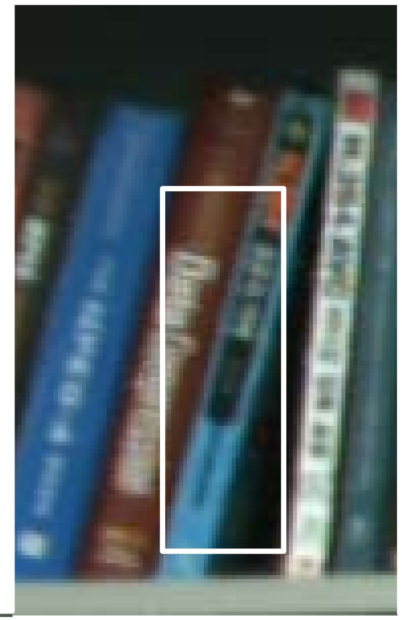
With VBDMM Mean-shift depth map

Subjective results

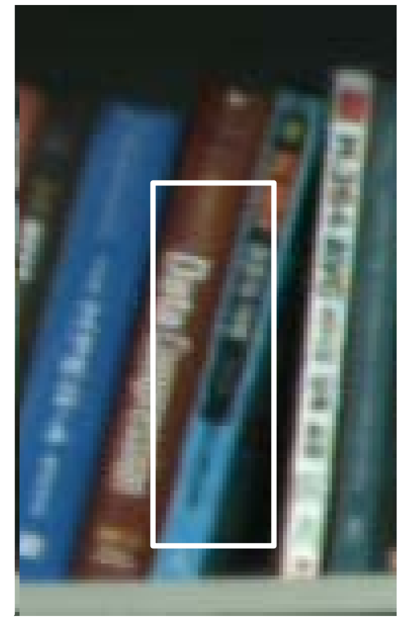
Test sequence: Newspaper

Original depth maps

With MPEG depth maps



With VBGMM K-Means depth maps



With VBDMM Mean-Shift depth maps

Subjective results

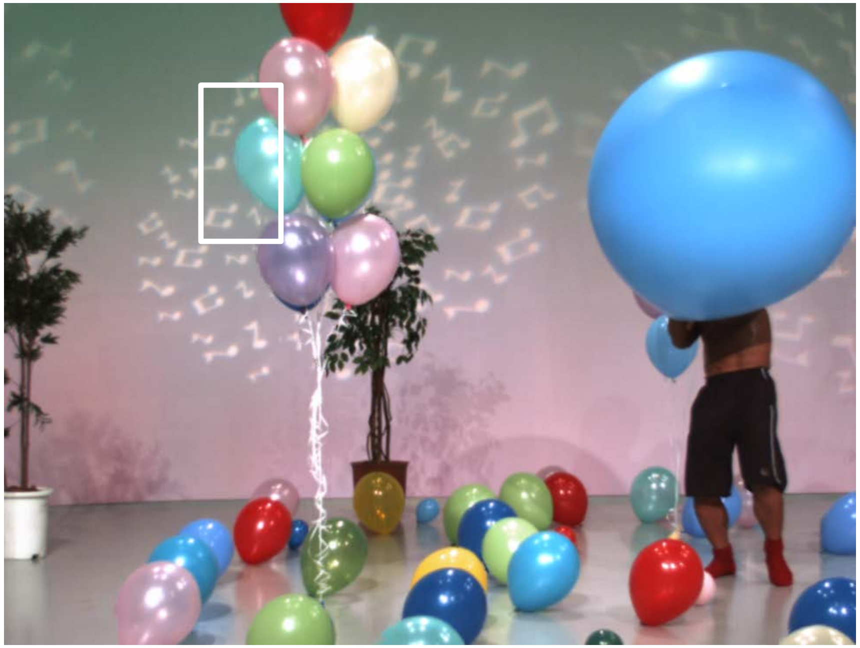
Test sequence: Balloons

With MPEG depth map

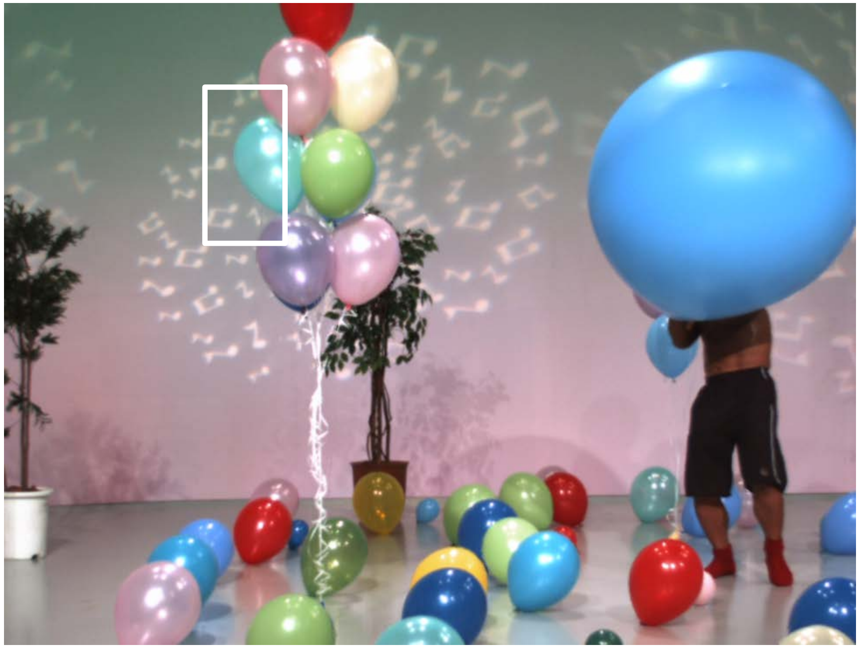
With VBDMM Mean-shift depth map

Subjective results

Test sequence: Balloons



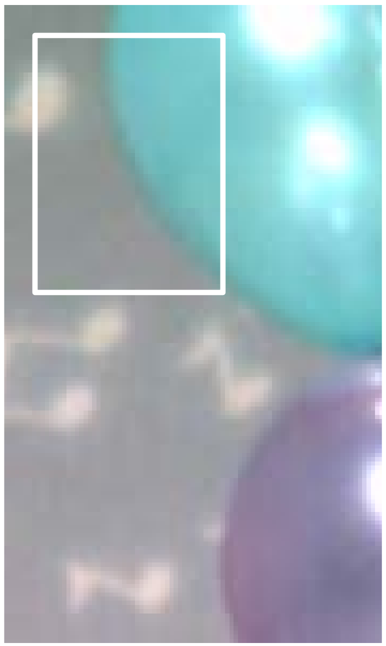
With MPEG depth map



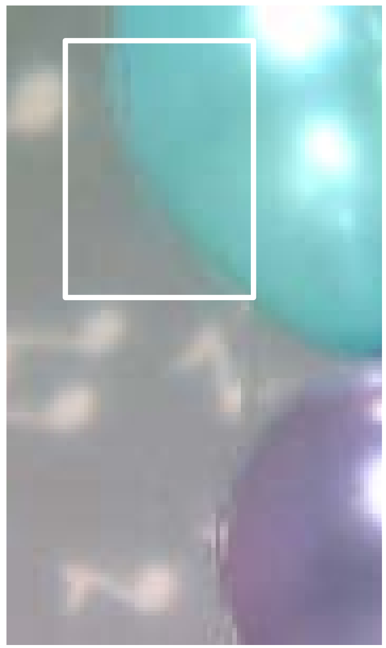
With VBDMM Mean-shift depth map

Subjective results

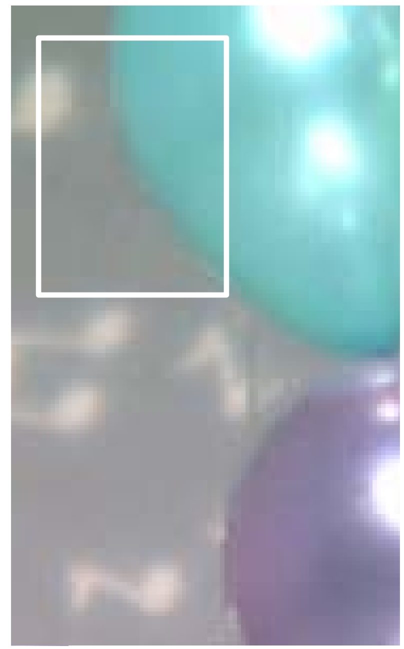
Test sequence: Balloons



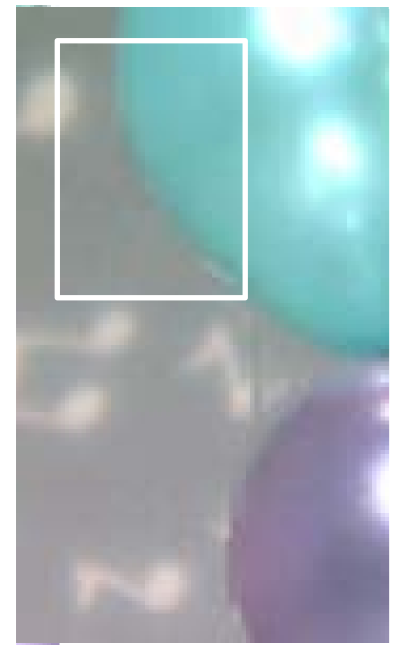
Original
depth maps



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

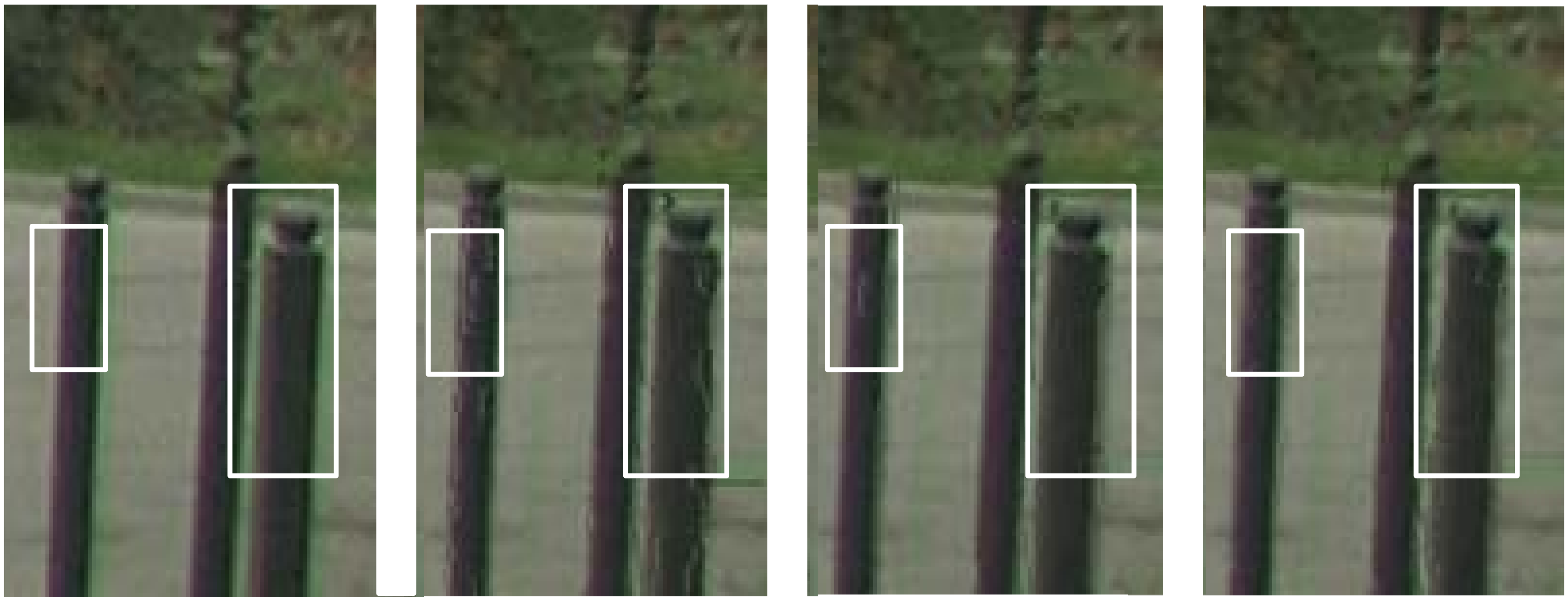
Test sequence: Poznan Street

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

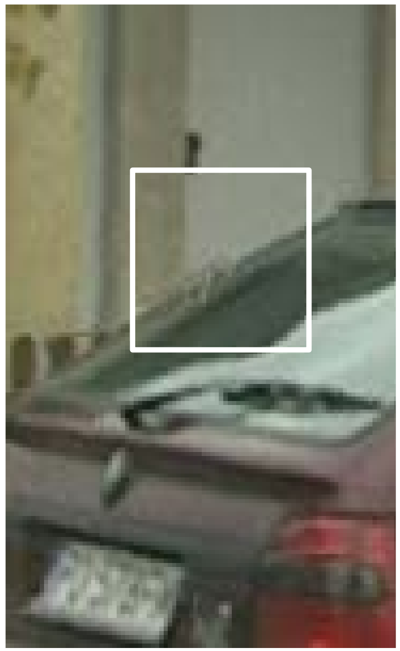
With MPEG depth map

With VBDMM Mean-shift depth map

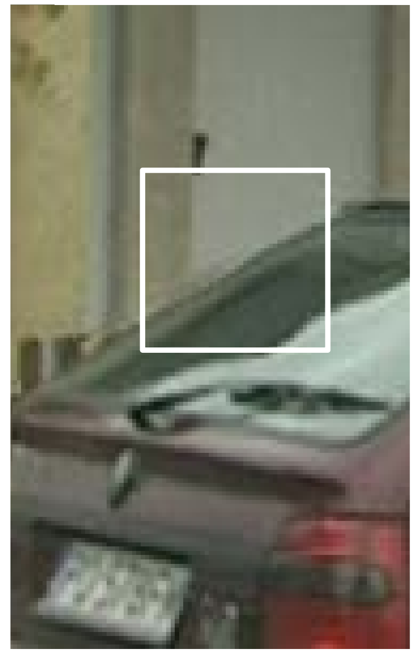
Subjective results

Test sequence: Poznan Street

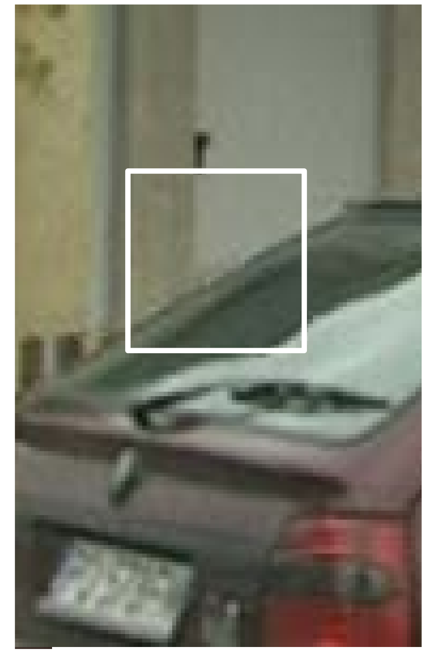
Original
depth maps



With MPEG
depth maps



With VBGMM
K-Means
depth maps



With VBDMM
Mean-Shift
depth maps

Conclusions

- We improved the inter-view depth consistency and hence, enhanced the visual experience of free-viewpoint television

- For that, we exploited the per-pixel association between depth and color by classification

- Color classification in xyz space is accomplished by variational Bayesian inference. Then, color classes are used for depth classification

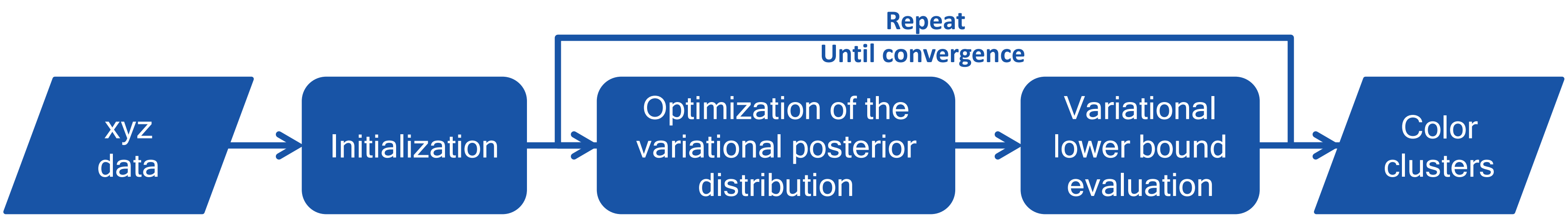
- Depth sub-clustering with Mean-shift improves the depth maps and hence view rendering quality

- Effectiveness of our approach is demonstrated by both objective and subjective results

Thank you

Multiview color classification

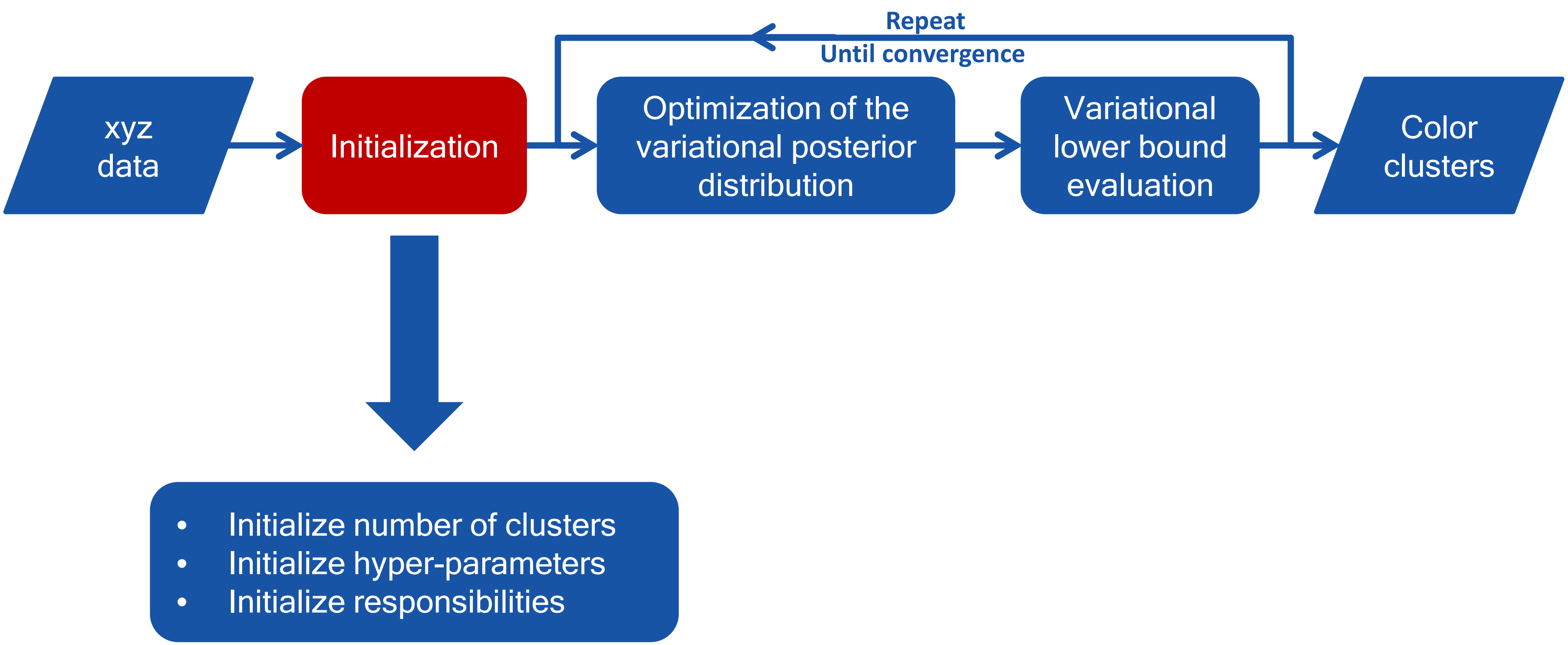
Dirichlet mixture model with variational Bayes inference



[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.
[5] 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.

Multiview color classification

Dirichlet mixture model with variational Bayes inference

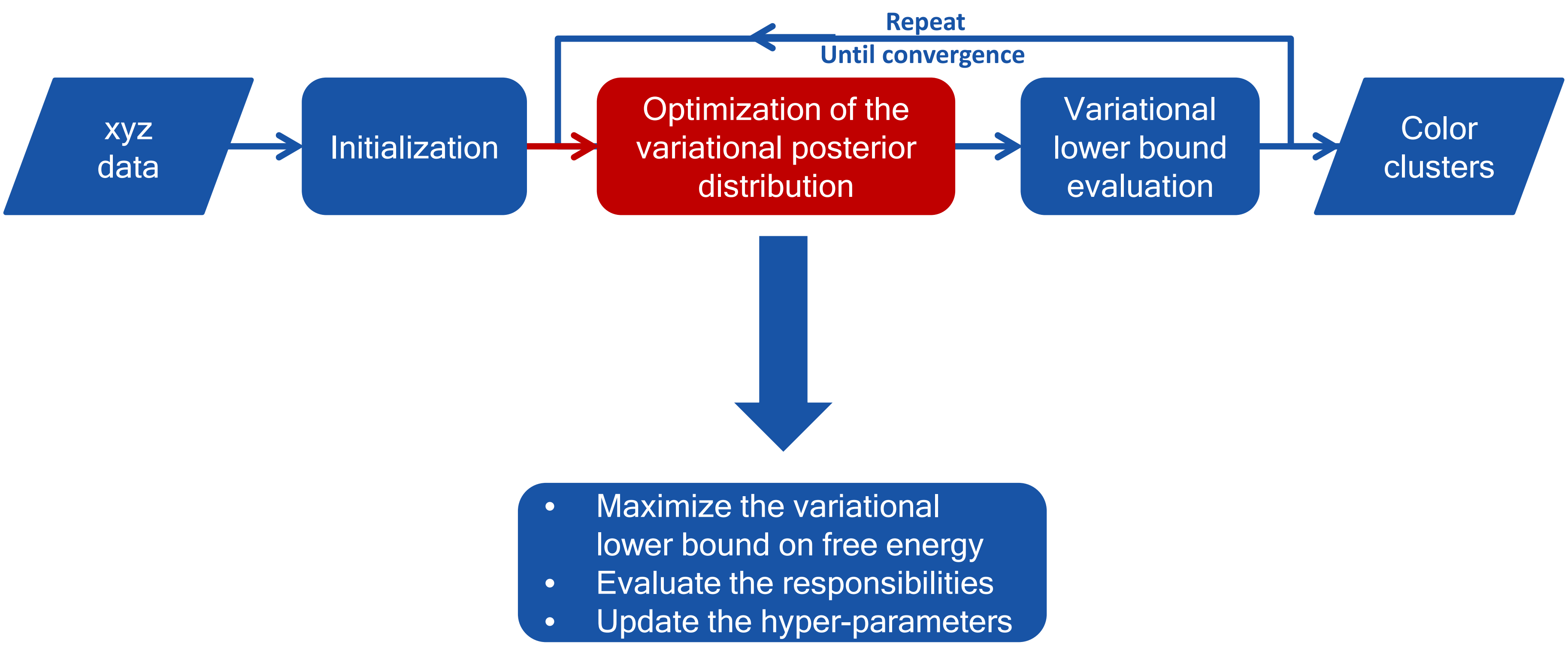


[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.

[5] 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.

Multiview color classification

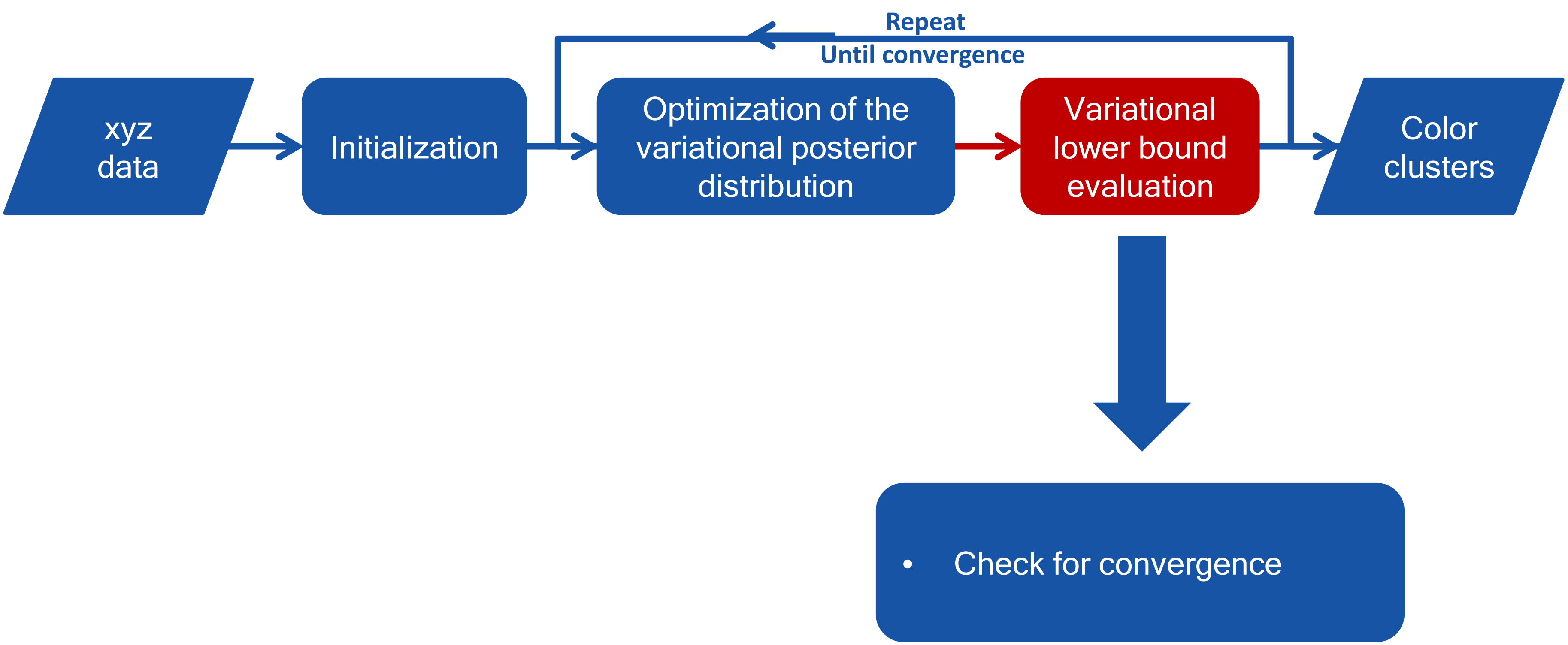
Dirichlet mixture model with variational Bayes inference



[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.
[5] 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.

Multiview color classification

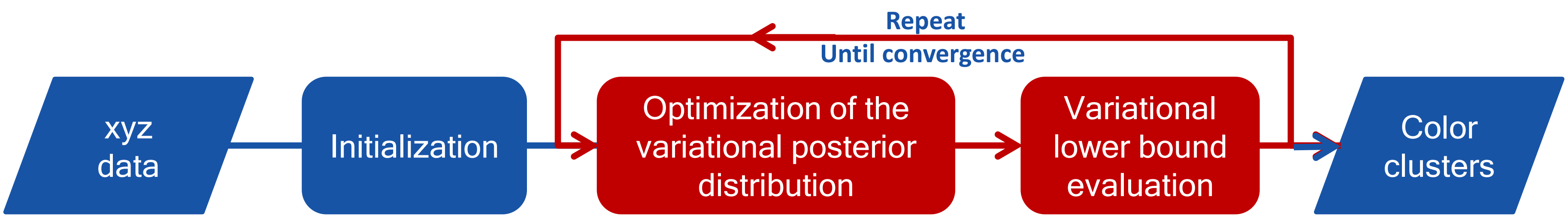
Dirichlet mixture model with variational Bayes inference



[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.
[5] 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.

Multiview color classification

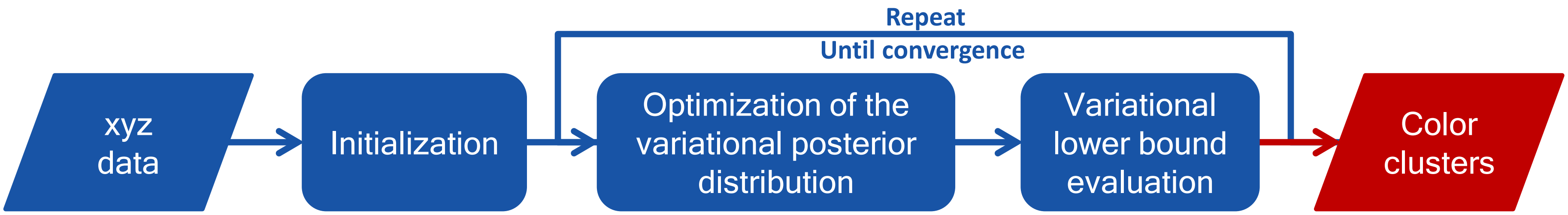
Dirichlet mixture model with variational Bayes inference



[4] C. M. Bishop, Pattern Recognition and Machine Learning, 1st ed. New York: Springer, 2006.
[5] 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.

Multiview color classification

Dirichlet mixture model with variational Bayes inference



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

[5] 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.

Future directions

- Improve color classification performance
- Improve computational efficiency of color classification
- Improve depth sub-clustering
- Improve temporal depth consistency

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.