Wide Angle Virtual View Synthesis Using Two-by-Two Matrix Kinect V2

Huan-Rui Chang and Hsueh-Ming Hang

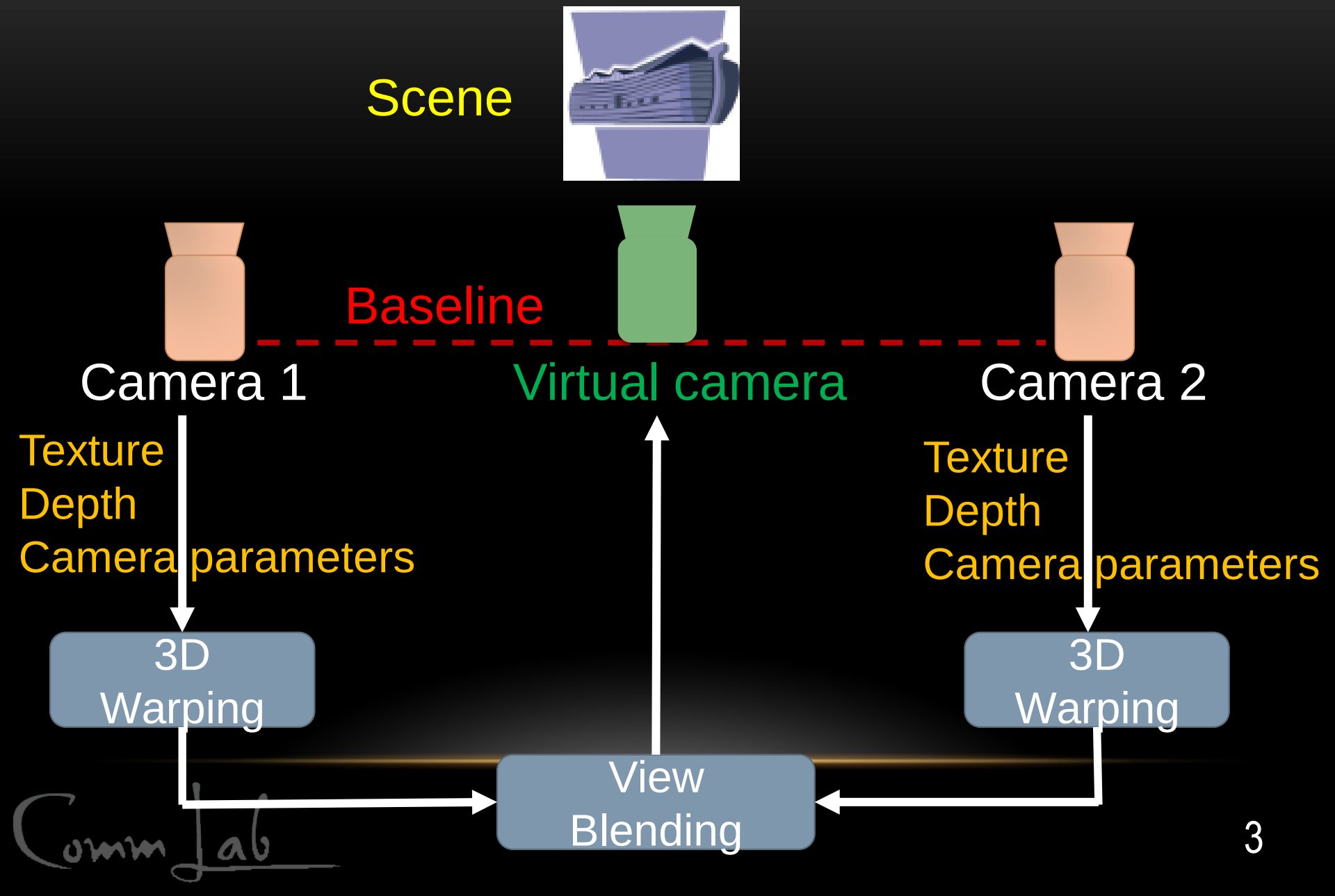
Electronics Inst., Electrical and Computer Engr College,
National Chiao Tung University, Hsinchu, Taiwan
Email: jackychang@commlab.tw, hmhang@mail.nctu.edu.tw

CommLab

Outline

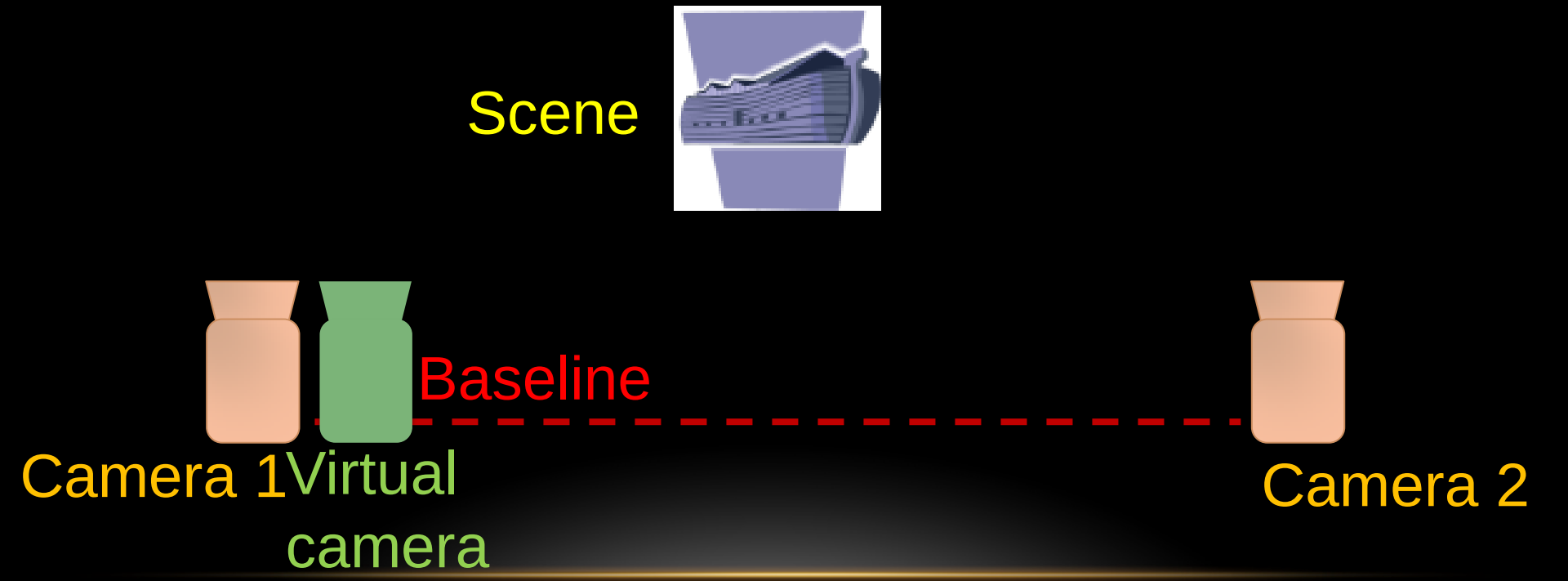
- Introduction
- Preprocessing Before Virtual View Blending
- The Proposed Algorithms
- Experimental Results
- Conclusions
- Future Work

Introduction: View Synthesis Flow Chart



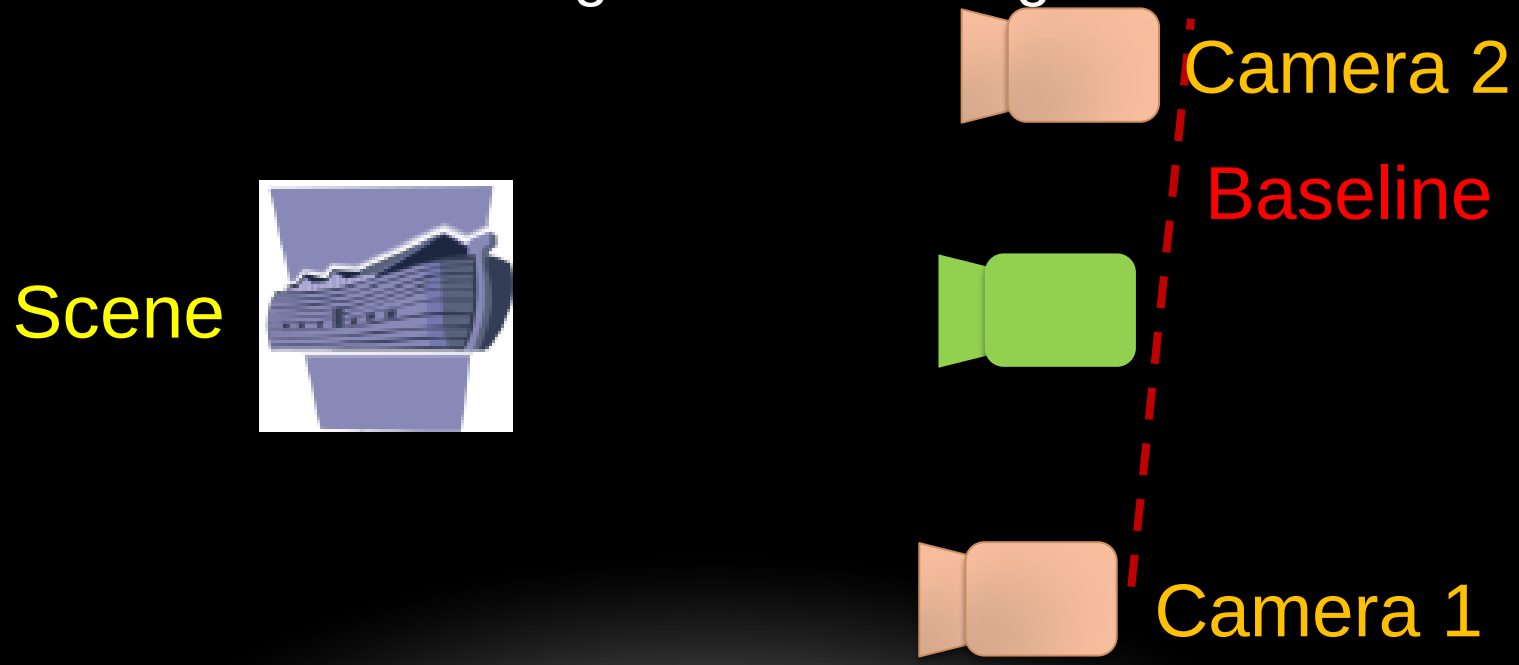
Introduction: Traditional view synthesis

- Traditional virtual view synthesis mainly focuses on how to synthesize virtual view lie in the baseline.



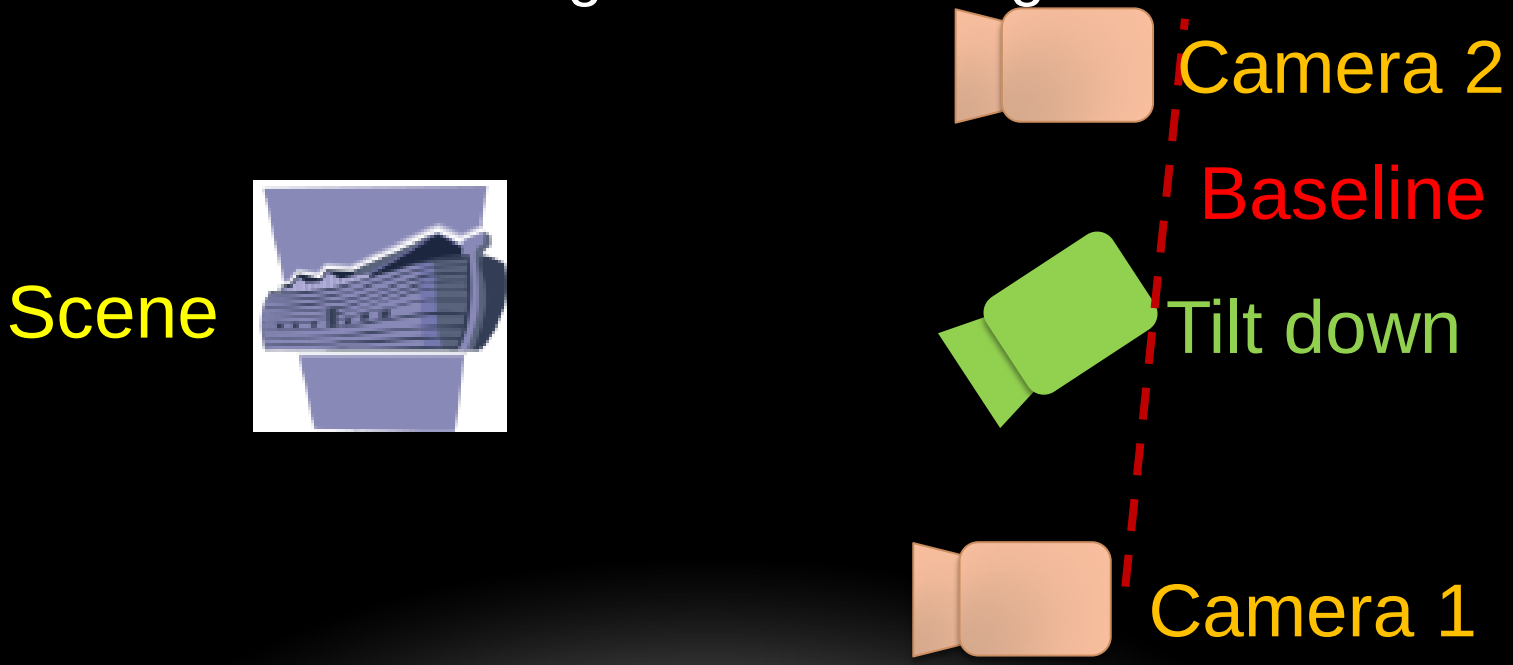
Introduction: Wide angle view synthesis

- Wide-angle synthesis (with zoom in/out or tilt up/down effect) often induces many artifacts like small cracks and large occlusion regions.

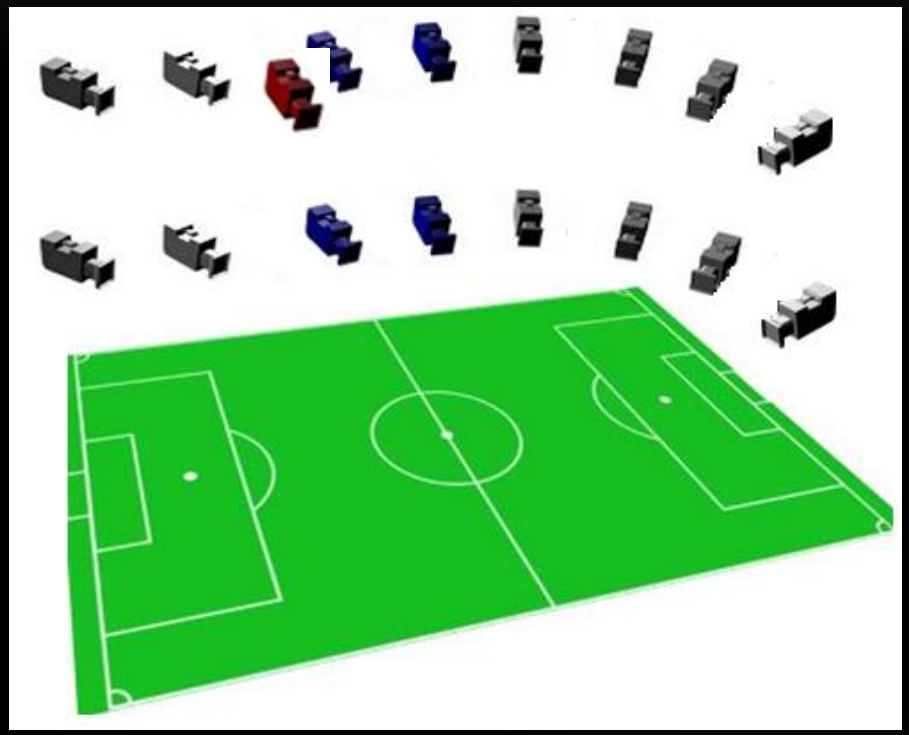


Introduction: Wide angle view synthesis

- Wide-angle synthesis (with zoom in/out or tilt up/down effect) often induces many artifacts like small cracks and large occlusion regions.



Wide view synthesis -- Applications



: Virtual viewpoint

: Reference viewpoint

Picture reference:
https://www.google.com.tw/search?espv=2&biw=1242&bih=585&tbm=isch&sa=1&q=free+viewpoint+video&oq=free+viewpoint+&gs_l=img.1.0.0i19k1.9193.9193.0.9896.1.1.0.0.0.0.217.217.2.1.0....0...1c.1.64.img..0.1.216.RZMyn15Gg2o#imgrc=9b54AWZnDB7mxM%3A

Contributions

- Implementing wide-angle view synthesis for real cases.
- Solve time synchronization between four Kinect sensors.
- Synthesized results robust against texture-depth misalignment and texture-texture misalignment.

Scene

Four Kinect v2 in two-by-two matrix arrangement

Proposed: Synchronizing all the four RGB-D cameras

Depth refinement | Depth refinement | Depth refinement | Depth refinement

3D warping | 3D warping | 3D warping | 3D warping

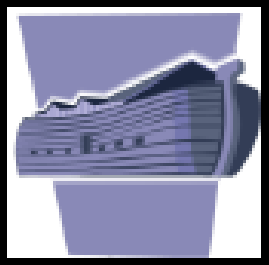
Proposed: Multi-view blending

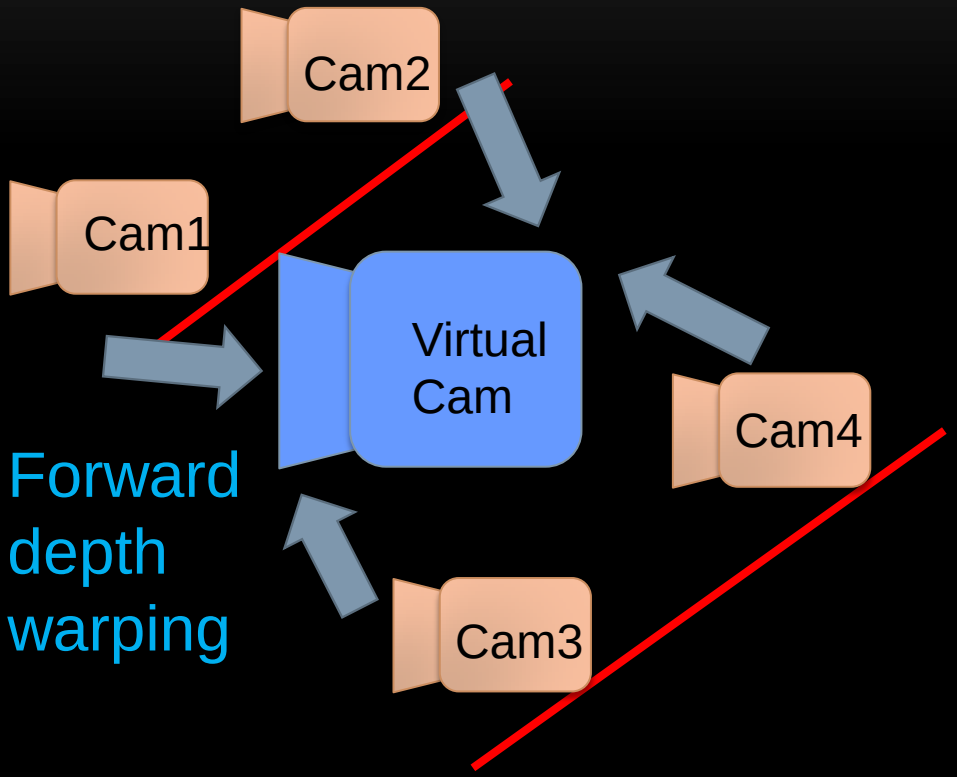
Outline

- Introduction
- Preprocessing Before Virtual View Blending
  - Depth Refinement
  - 3D Warping
- The Proposed Algorithms
- Experimental Results
- Conclusions
- Future Work

3D Warping

Scene





$$\mathbf{p}_v = \frac{1}{z_v}\mathbf{A}_v\left(\mathbf{R}_v\mathbf{R}_r^{-1}\left(z_r\mathbf{A}_r^{-1}\mathbf{p}_r - \mathbf{t}_r\right) + \mathbf{t}_v\right)$$

3D Waring: Original Four Views

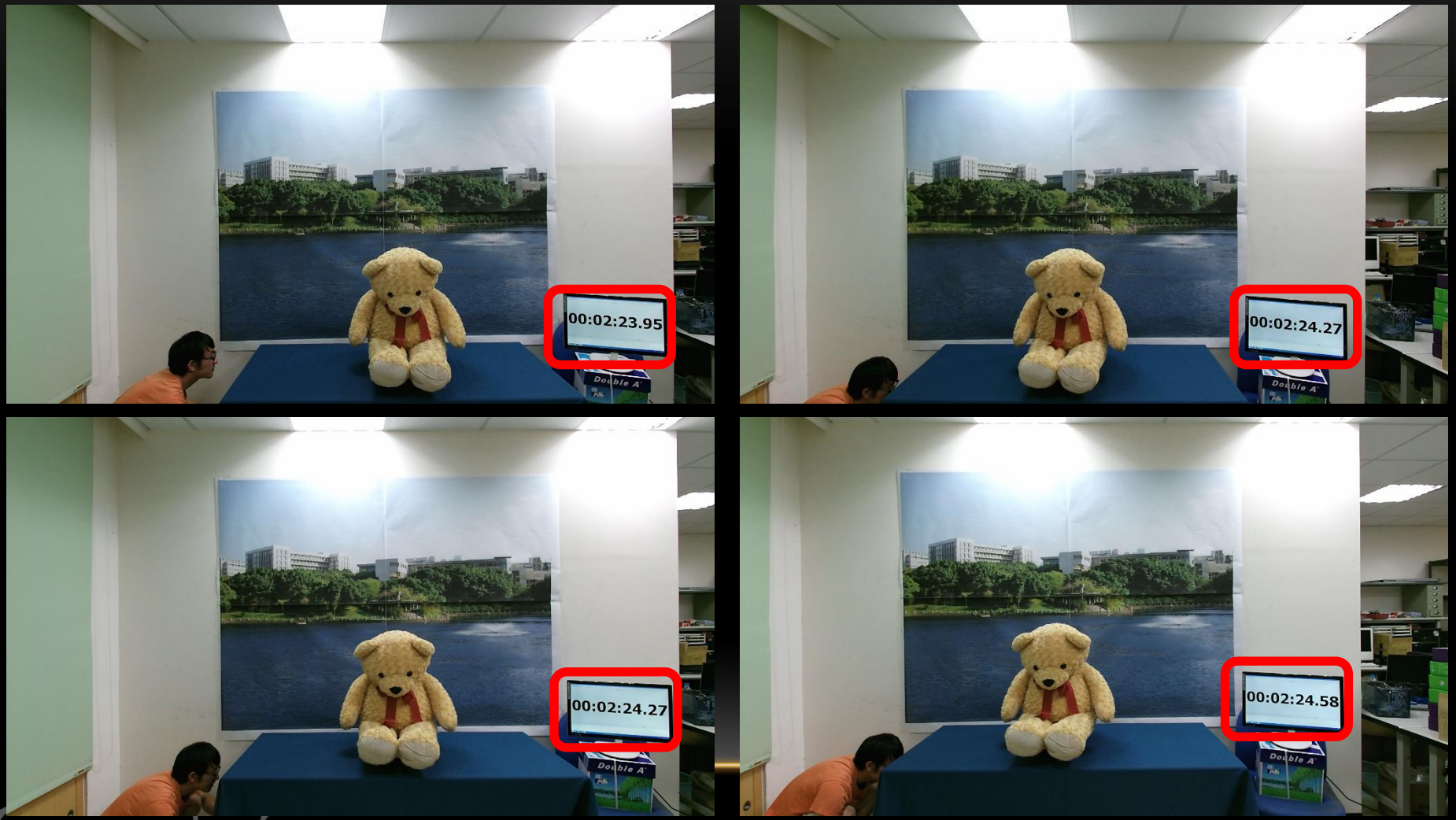
3D Waring: **After** 3D Warping

- Virtual camera: tilt down 5 degrees, and zoom in 100mm.

Outline

- Introduction
- Preprocessing Before Virtual View Blending
- The Proposed Algorithms
  - Synchronizing All the Four Kinect V2 Sensors
  - Multi-view Blending
- Experimental Results
- Conclusions

Synchronizing All Four Kinect V2 Sensors: Before



Max captured time skew=0.12seconds

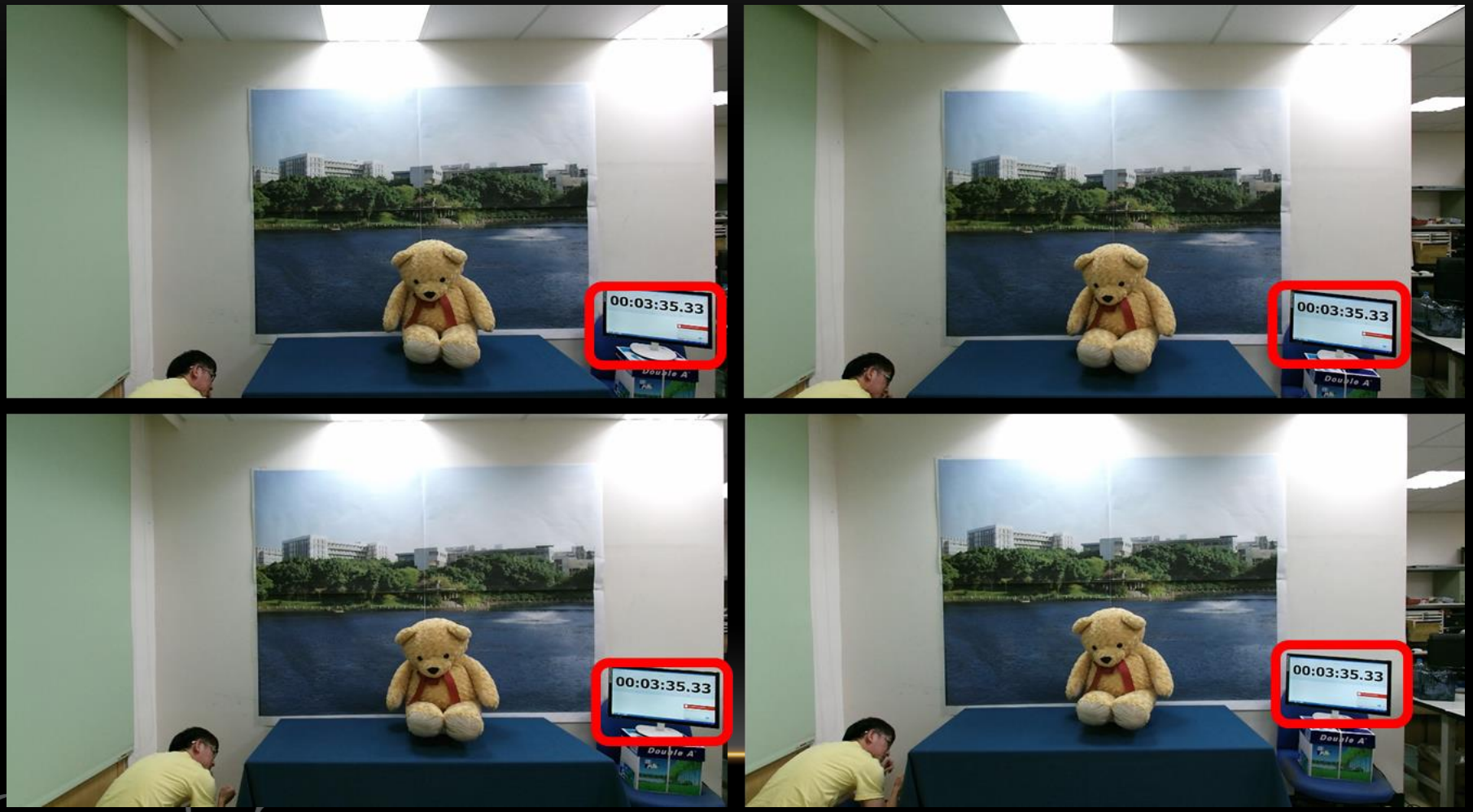
Max captured time skew 0.81s

Synchronizing All Four Kinect V2 Sensors: **Solutions**

- Use NTPserver.exe software to calibrate the system time of each PC.
- All Kinects start recording at the same time.
- All Kinects capture image & depth at 0ms.

- Advantage: Achieve synchronization.
- Disadvantage: Frame rate only 1fps.

Synchronizing All the Four Kinect V2 Sensors: **Results**



CommLab

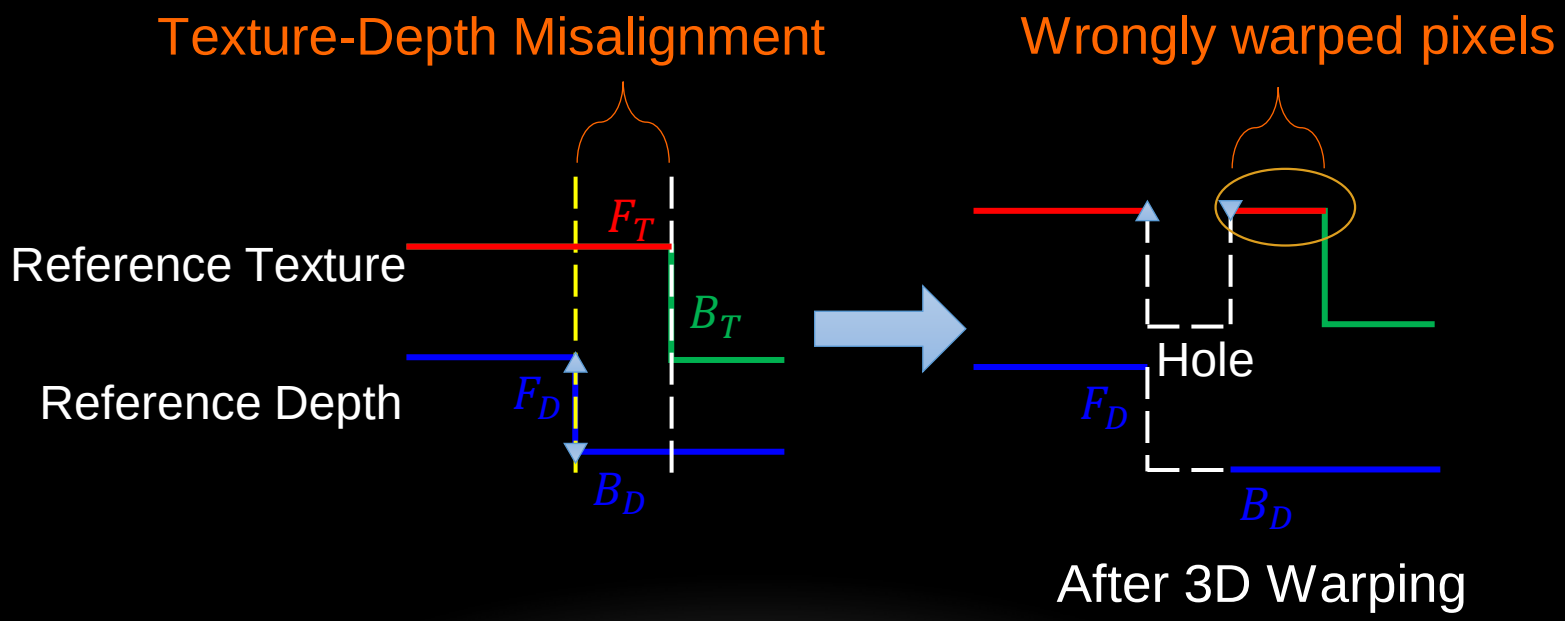
Max captured time skew 0.00s

Outline

- Introduction
- Preprocessing Before Virtual View Blending
- The Proposed Algorithms
  - Synchronizing All the Four Kinect V2 Sensors
  - Multi-view Blending
- Experimental Results
- Conclusions

1st Problem of Warped Virtual Views

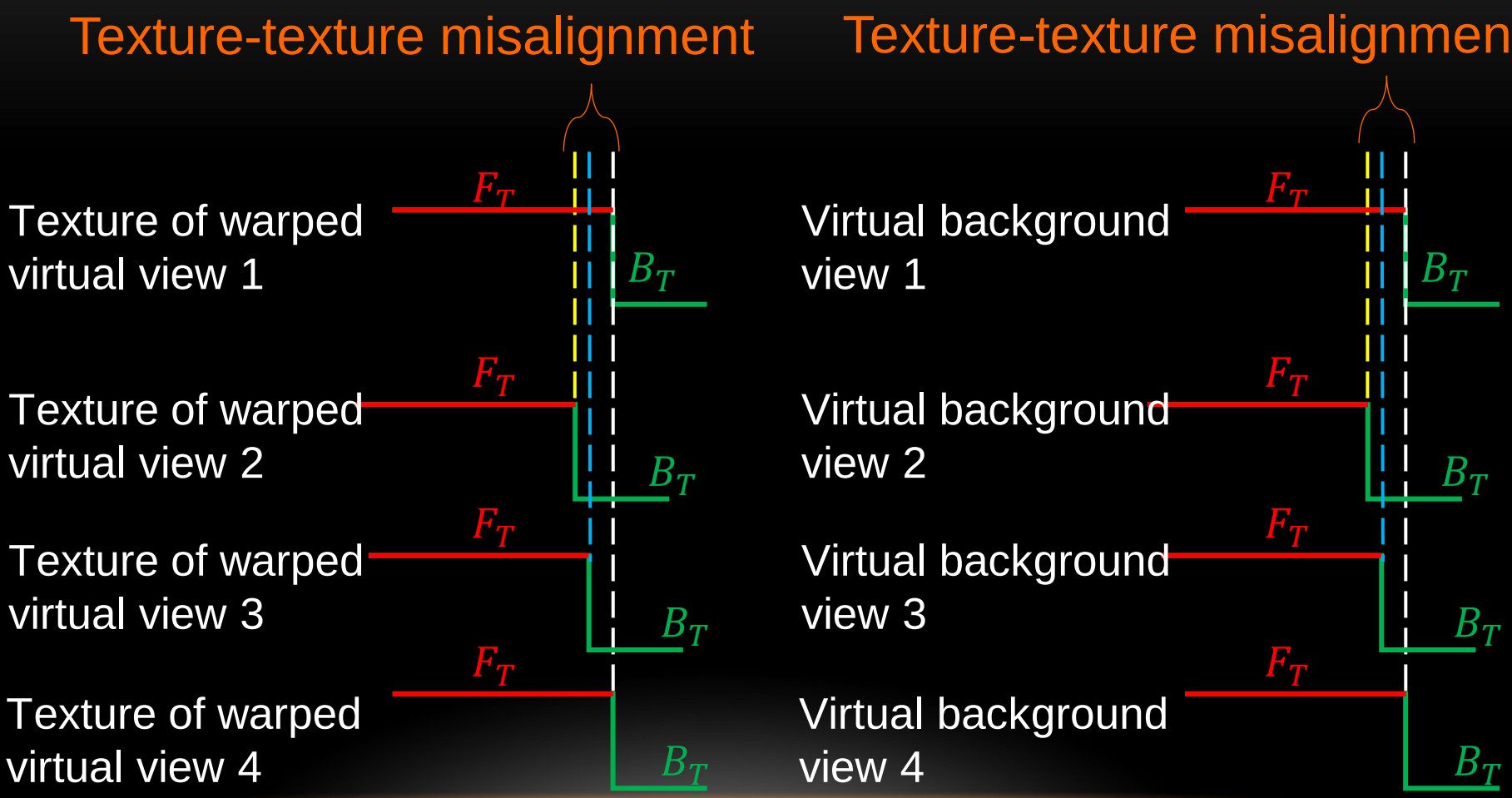
- Misalignment between reference color and depth leads to wrongly warped pixels.



Misalignment between Color & Depth

- Four warped virtual views

2nd Problem of Warped Virtual Views



Texture-texture Misalignment

Proposed Multi-view Blending Algorithm

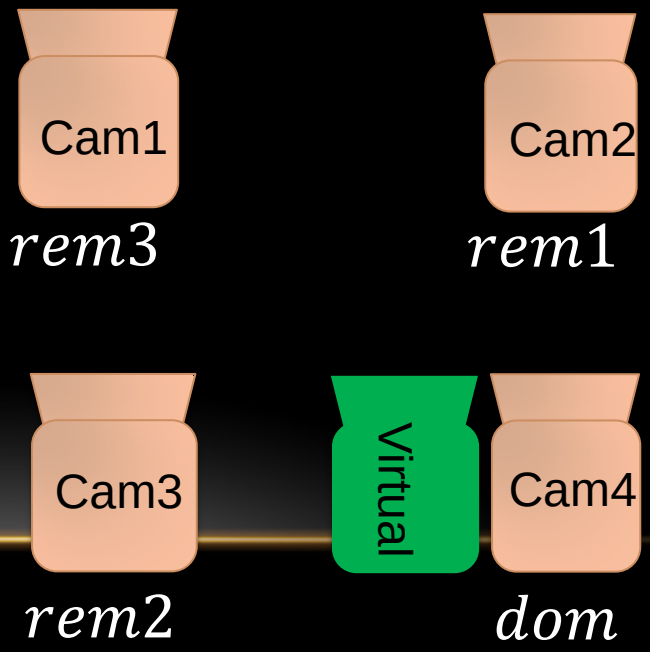
- Purpose: robust against
  1. texture-depth misalignment
  2. texture-texture misalignment

  of the four warped virtual views.

- **Part 1**: render pixels not near depth edge.
- **Part 2**: render pixels near depth edge.
- **Part 3**: render pixels by choosing the best matched pixels from four virtual views.

Proposed Multi-view Blending Algorithm

- Dominant virtual view (denote as *dom*): the warped virtual view whose reference camera position is closest to virtual camera.
- The remaining warped virtual views are denote as rem1, rem2, and rem3.

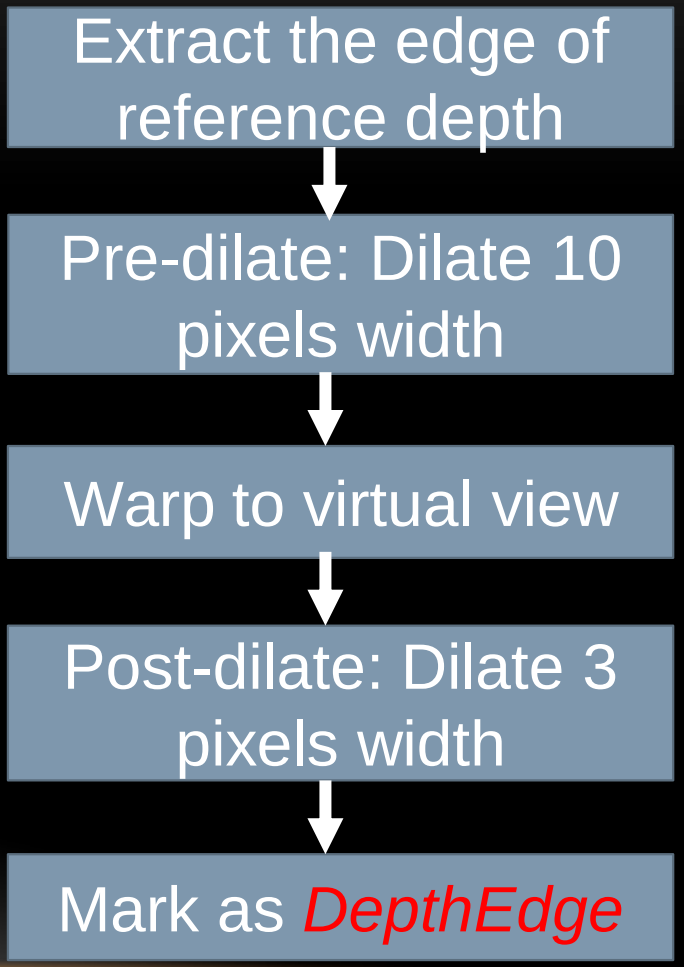


Extract *DepthEdge*

- The depth around object boundary are often unreliable

  → texture-depth misalignment

- Extract the pixels around depth edge as *DepthEdge.*

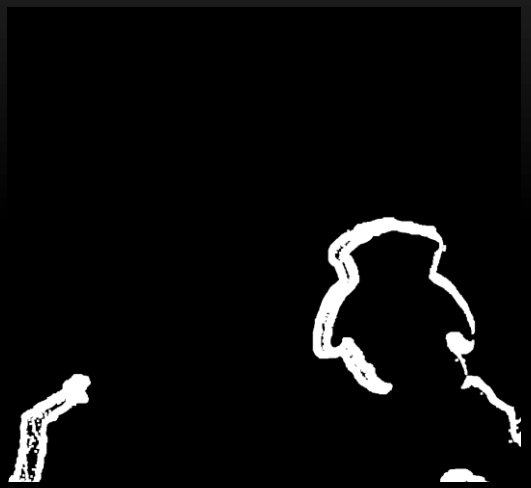


Illustrations of *DepthEdge*

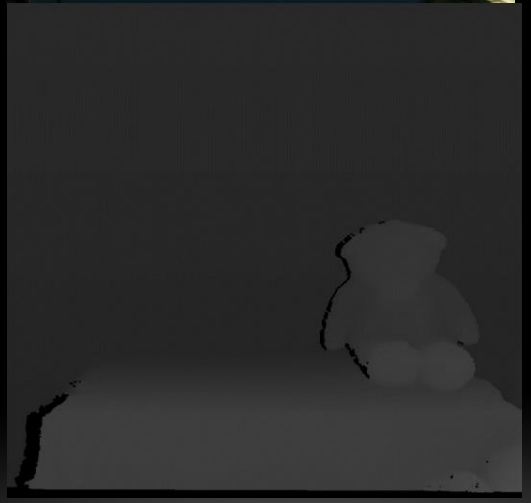
Reference Color

Warped virtual color

DepthEdge



Reference Depth

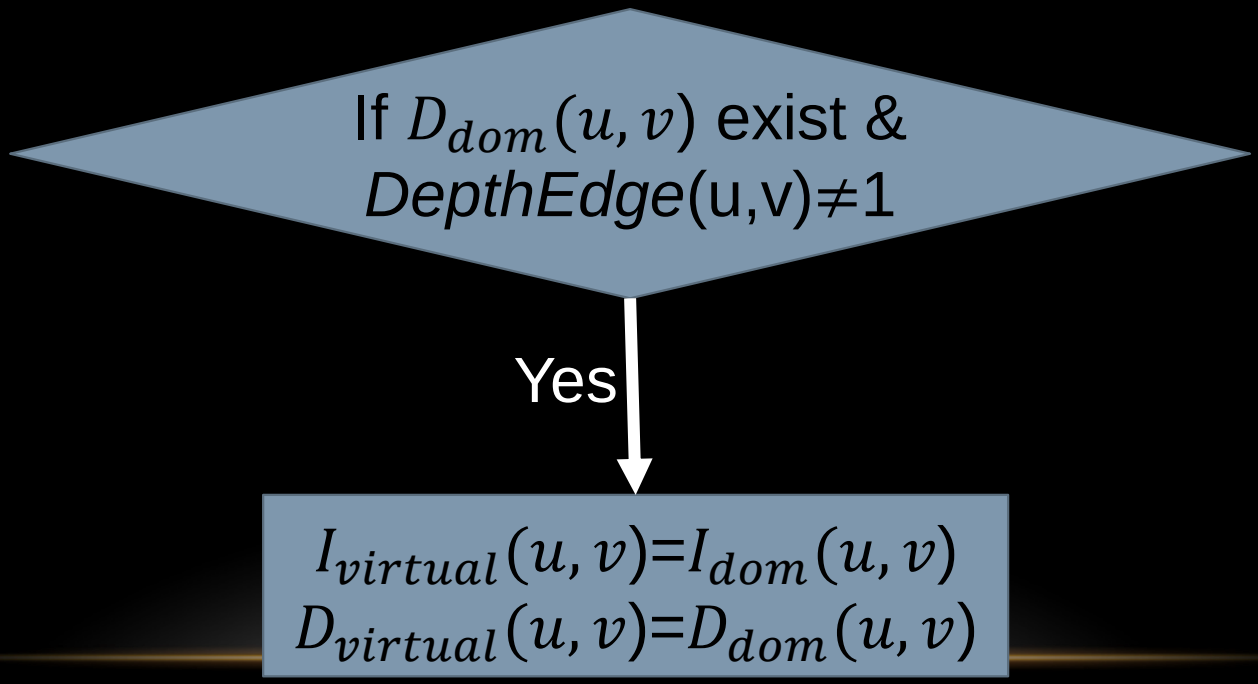


Warped virtual depth

Color of *DepthEdge*

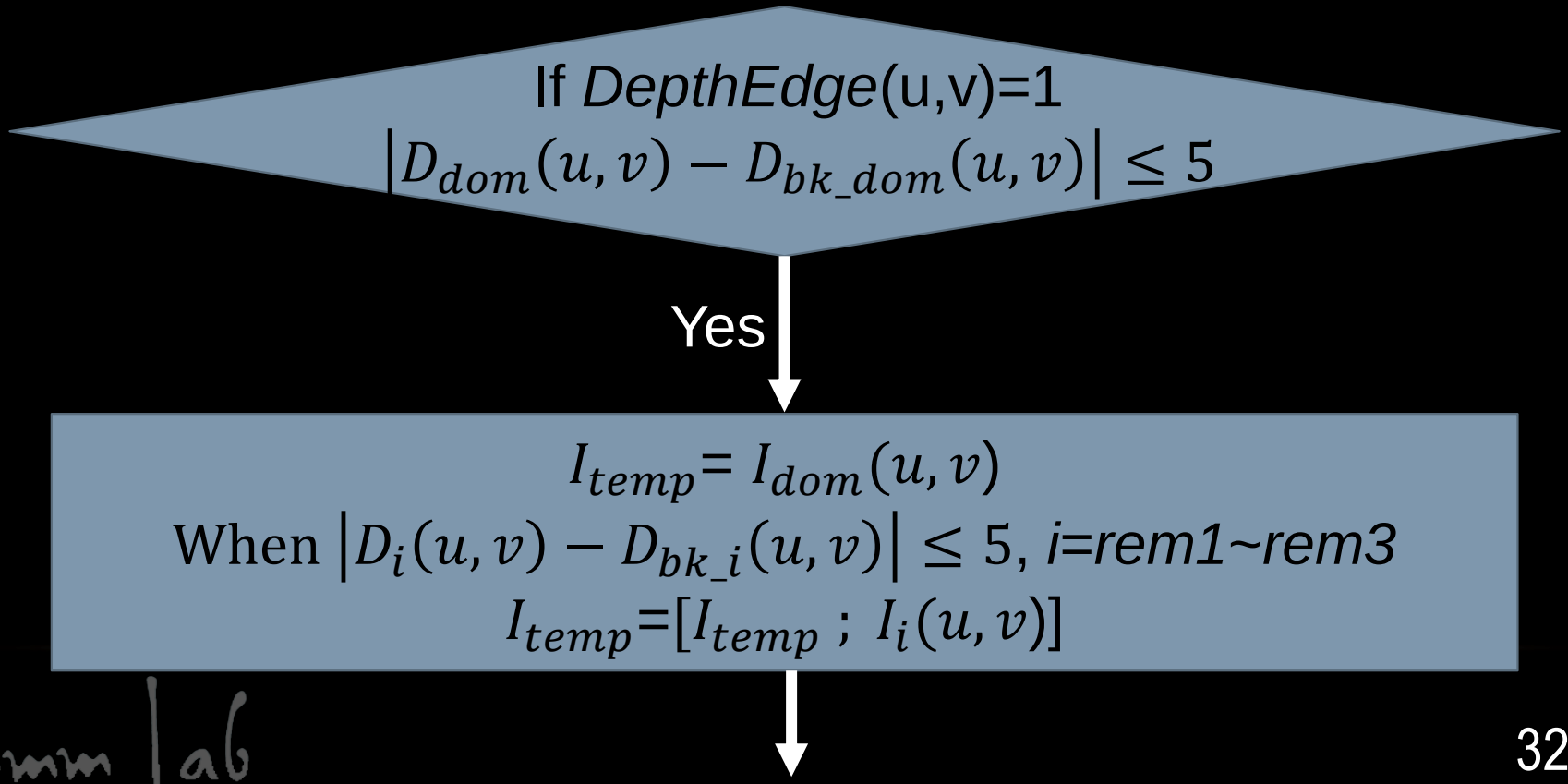
Proposed Multi-view Blending Algorithm: Part 1

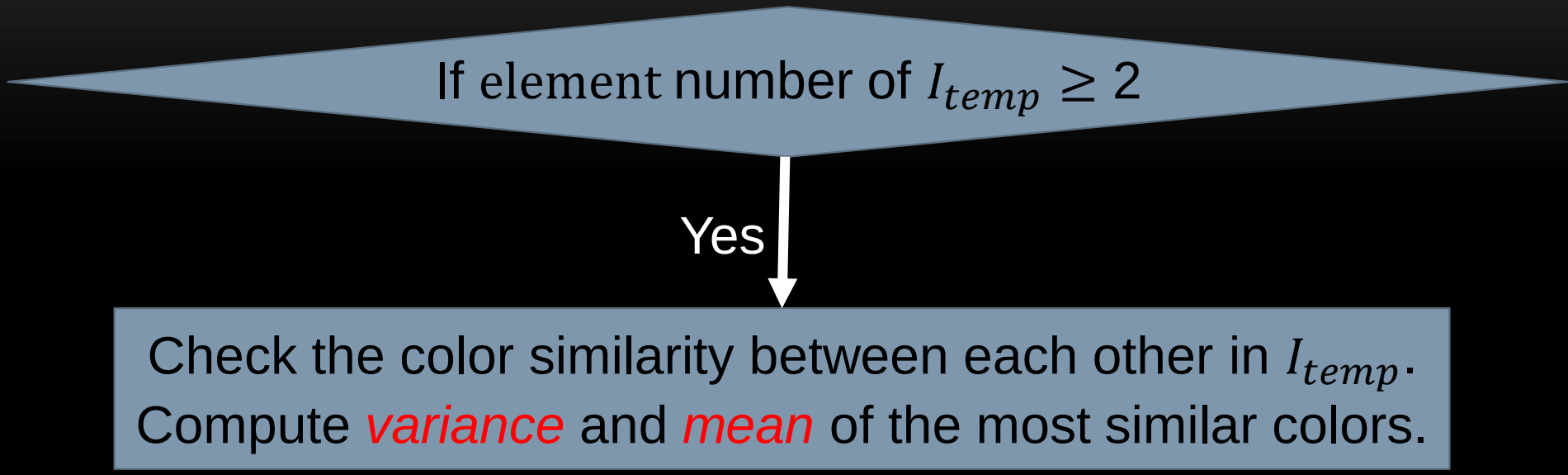
- Use the color and depth pixel of the dominant virtual view to render the color pixel.

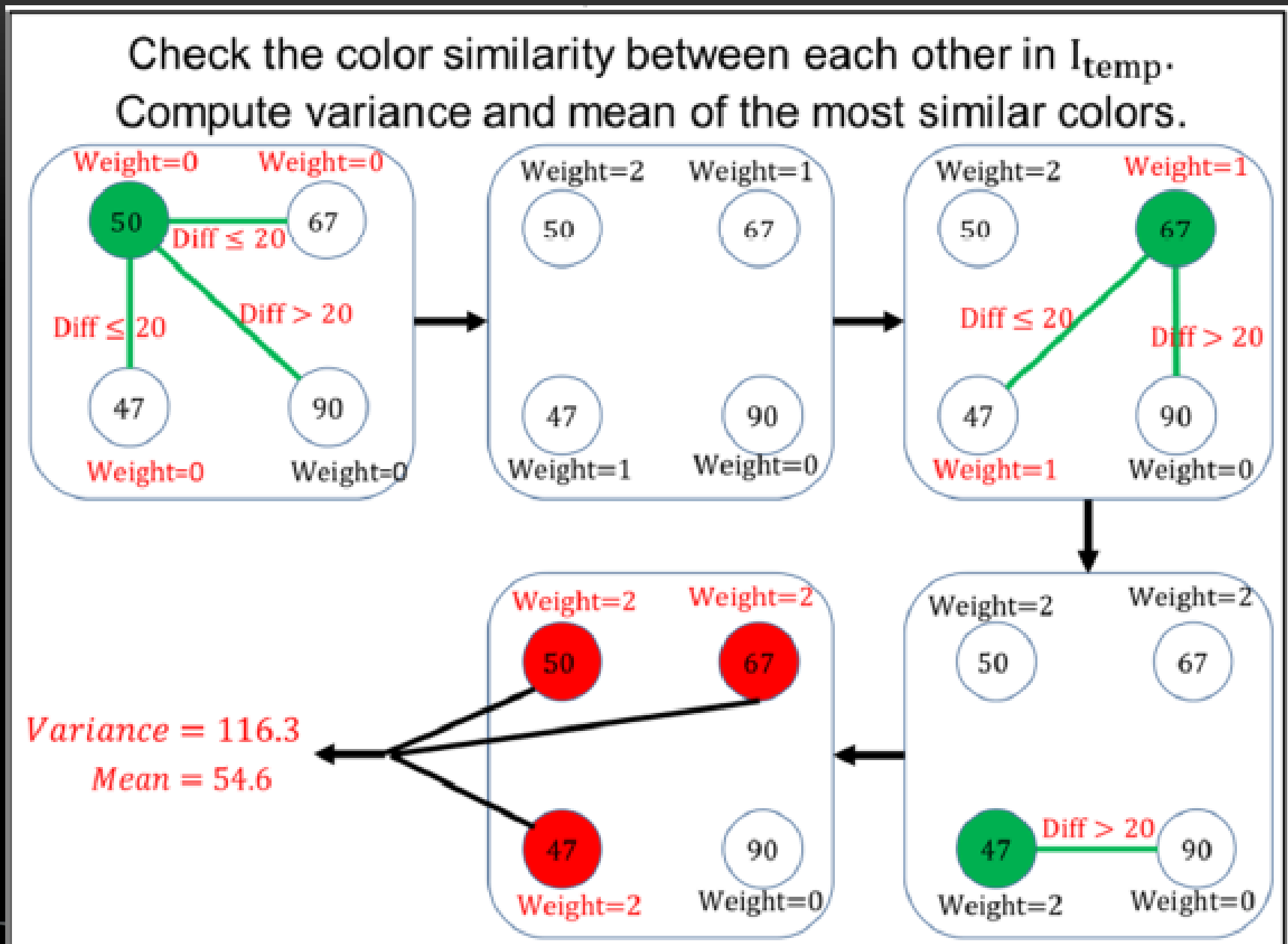


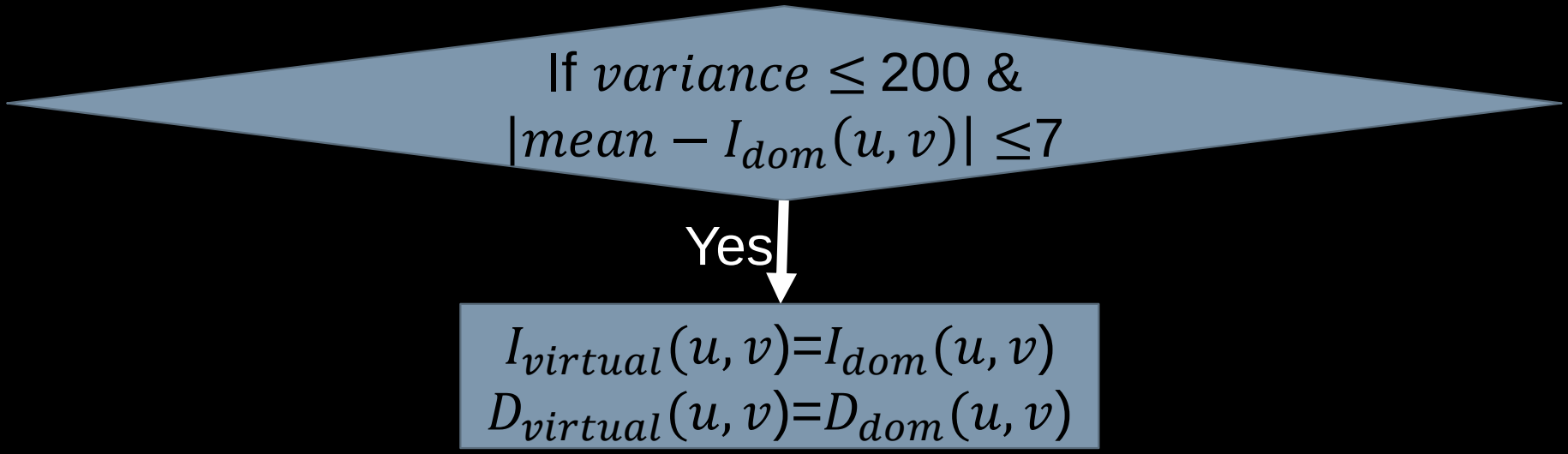
Proposed Multi-view Blending Algorithm: Part 2

- Remove the noise due to texture-depth misalignment and texture-texture misalignment by choosing the correct color pixel from the other warped virtual views.




If element number of $I_{temp} \geq 2$
Yes
Check the color similarity between each other in I_{temp}.
Compute *variance* and *mean* of the most similar colors.

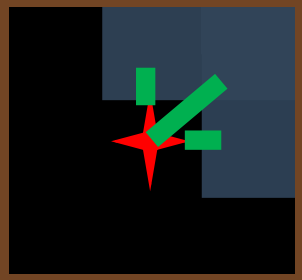

Check the color similarity between each other in I_{temp}.
Compute variance and mean of the most similar colors.
Weight=0
Weight=0
50
67
Diff ≤ 20
Diff ≤ 20
Diff > 20
47
90
Weight=0
Weight=0
Weight=2
Weight=1
50
67
47
90
Weight=1
Weight=0
Weight=2
Weight=1
50
67
Diff ≤ 20
Diff > 20
47
90
Weight=1
Weight=0
Weight=2
Weight=2
50
67
47
90
Weight=2
Weight=0
Weight=2
Weight=2
50
67
47
Diff > 20
90
Weight=2
Weight=0
$Variance = 116.3$
$Mean = 54.6$


If $variance \leq 200$ &
$|mean - I_{dom}(u,v)| \leq 7$
Yes
$I_{virtual}(u,v) = I_{dom}(u,v)$
$D_{virtual}(u,v) = D_{dom}(u,v)$

If $variance \leq 200$ &
$|mean - I_{dom}(u,v)| \geq 7$

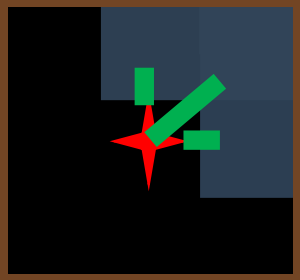
Yes

color of virtual view 1



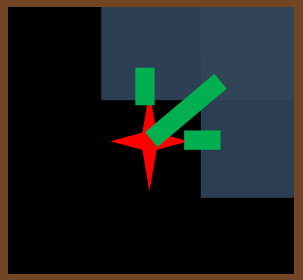
SSD1

color of virtual view 2



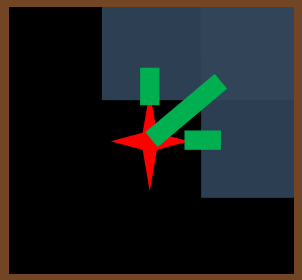
SSD2

color of virtual view 3



SSD3

color of virtual view 4



SSD4

Use the warped color pixel (having the minimum sum of squared errors) to be the synthesized color pixel.

Proposed Multi-view Blending Algorithm: Part 3

- Choosing the best matched color pixels from the four warped virtual views to render the remaining color pixels.

If $D_{virtual}(u, v)$=hole

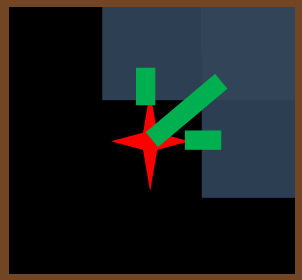
Yes

color of virtual view 1

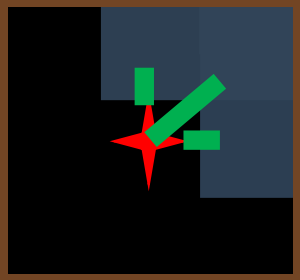
color of virtual view 2

color of virtual view 3

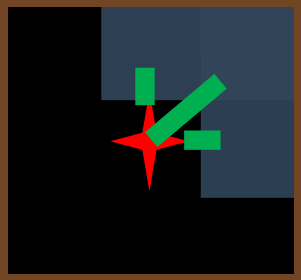
color of virtual view 4



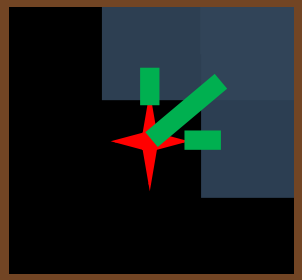
SSD1



SSD2



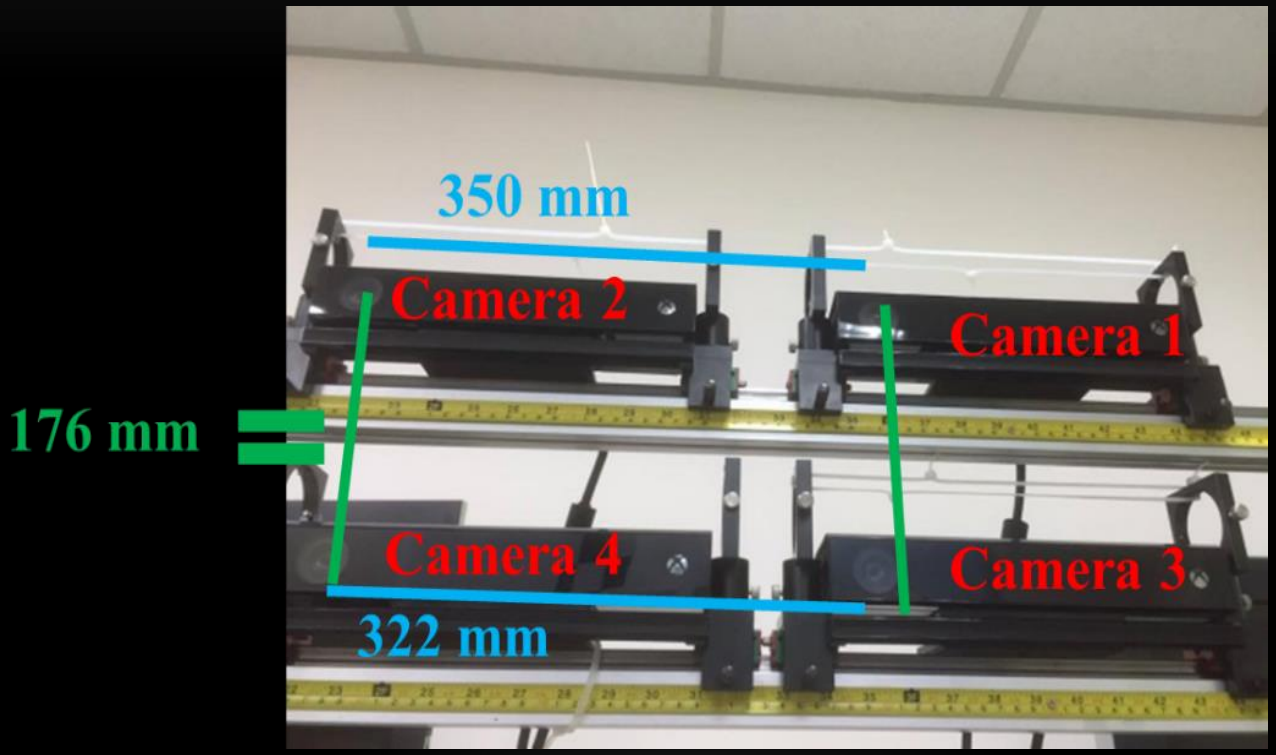
SSD3



SSD4

Use the warped color pixel (having the minimum sum of squared errors) to be the synthesized color pixel.

Experimental Setting

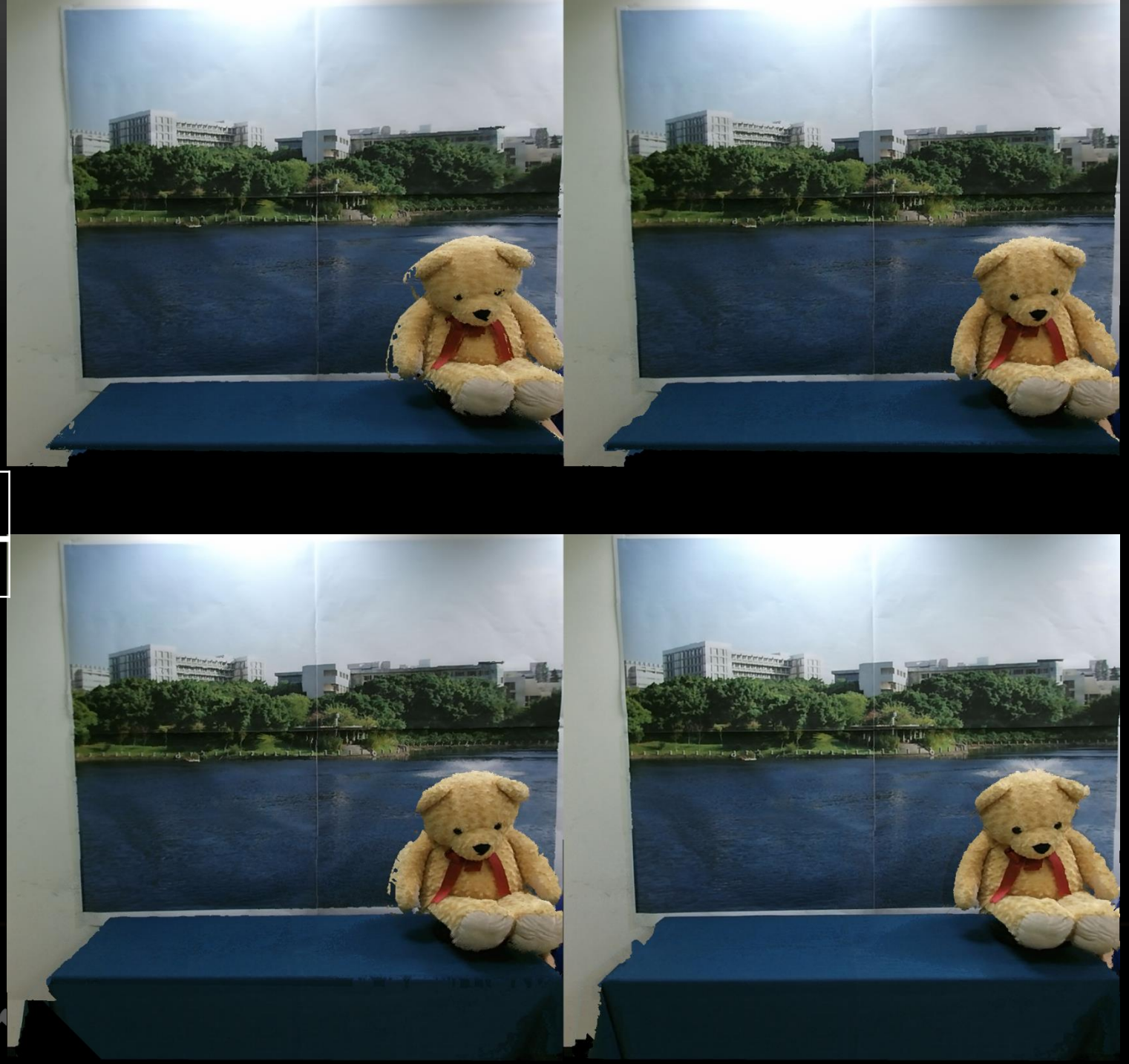


Experimental Results Comparison

- Comparing with the adaptive blending (Ltc method).
- Comparing the results when input camera number varies (from two to four).

T.-C. Lee, C. –L. Chien, and H.-M. Hang, "Virtual view synthesis using quality refiement," in *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, July 2016.

Case 1: Tilt down 5 degrees, and zoom in 100mm



Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Case 2: Zoom in 400mm

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4


Ltc
2
3
4

Ltc	2
3	4


Ltc
2
3
4

Case 3: Zoom out 400mm

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

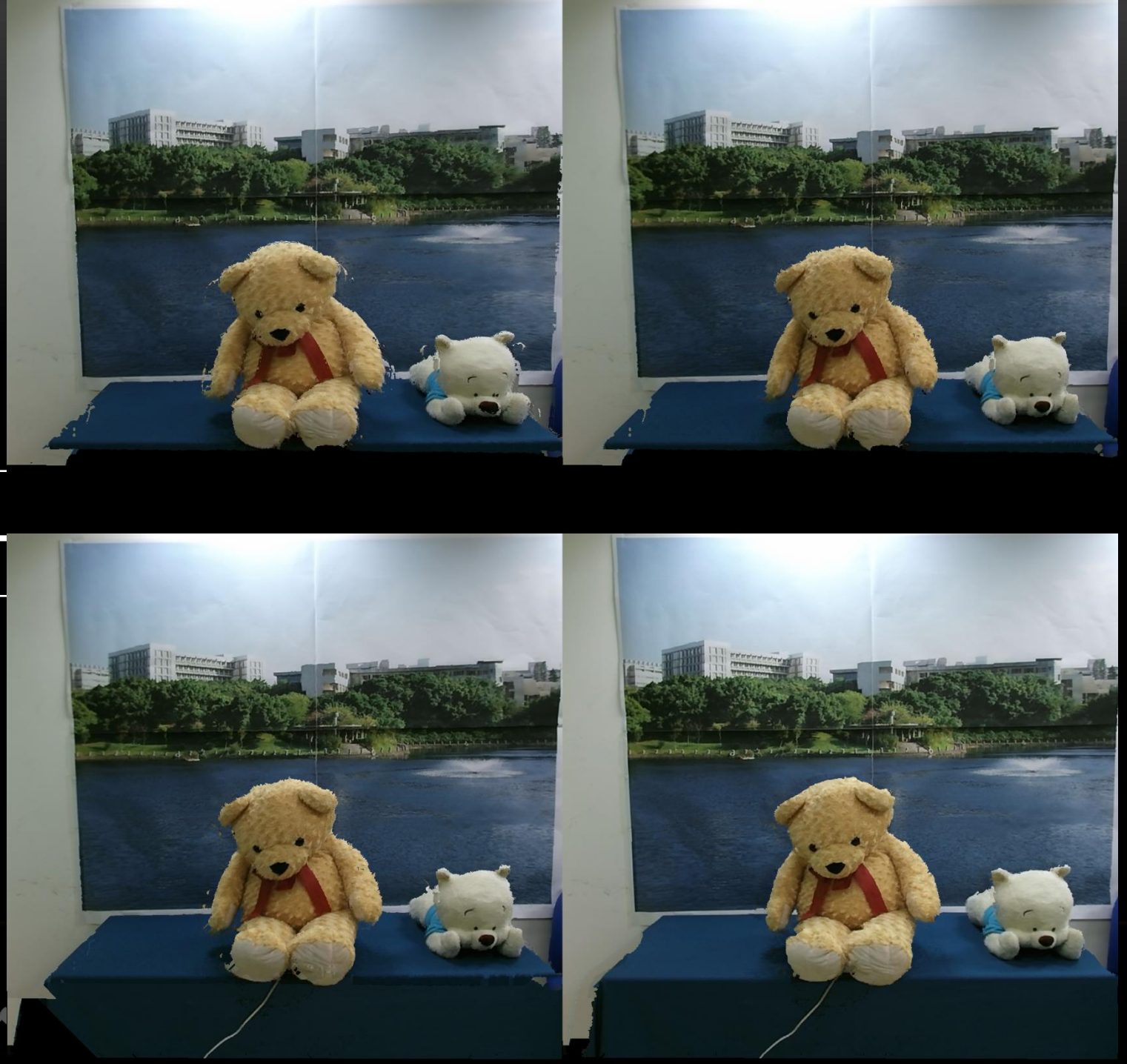
Ltc	2
3	4

Case 4: Tilt down 5 degrees, and zoom in 100mm

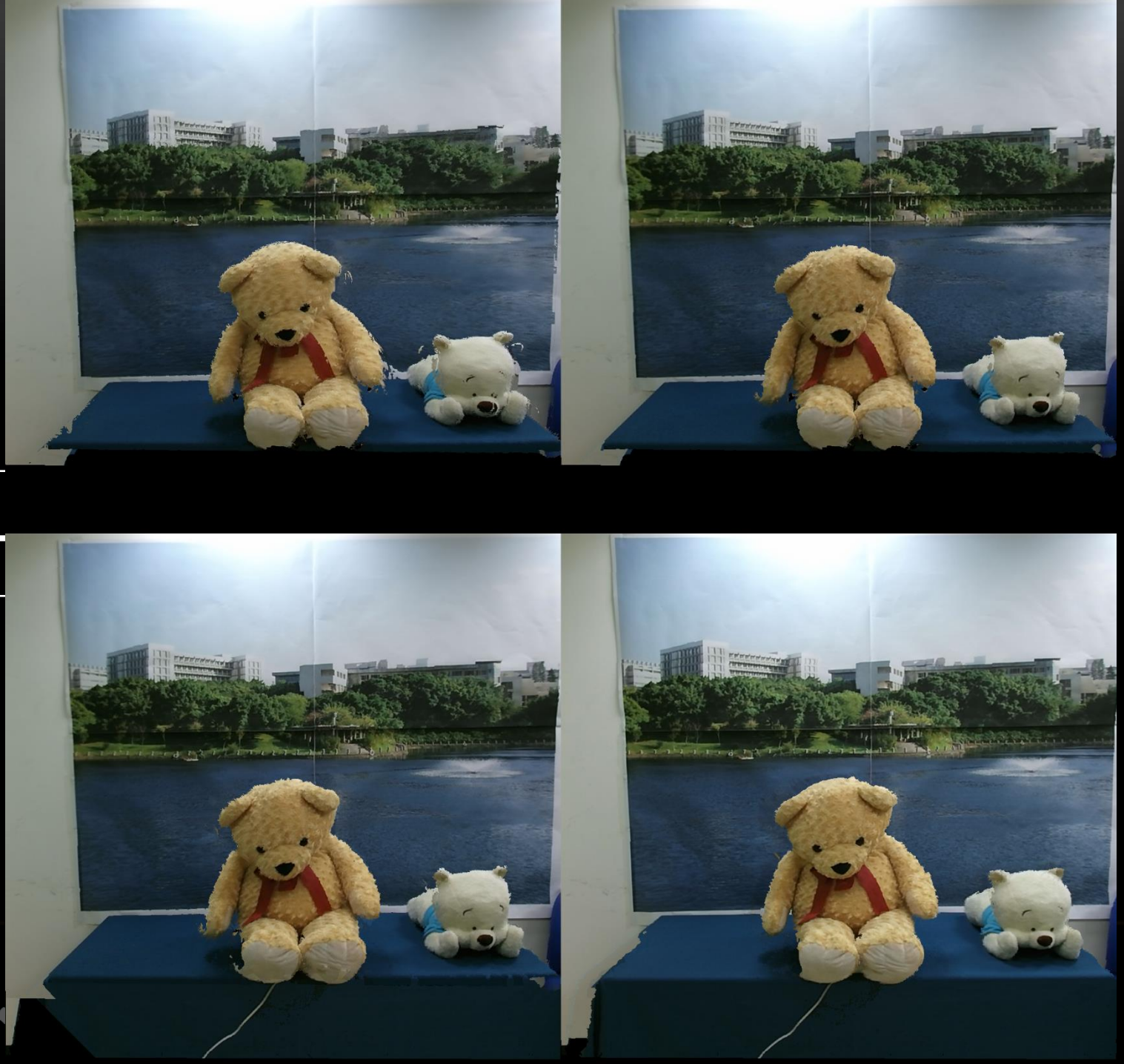
Ltc	2
3	4

Ltc	2
3	4

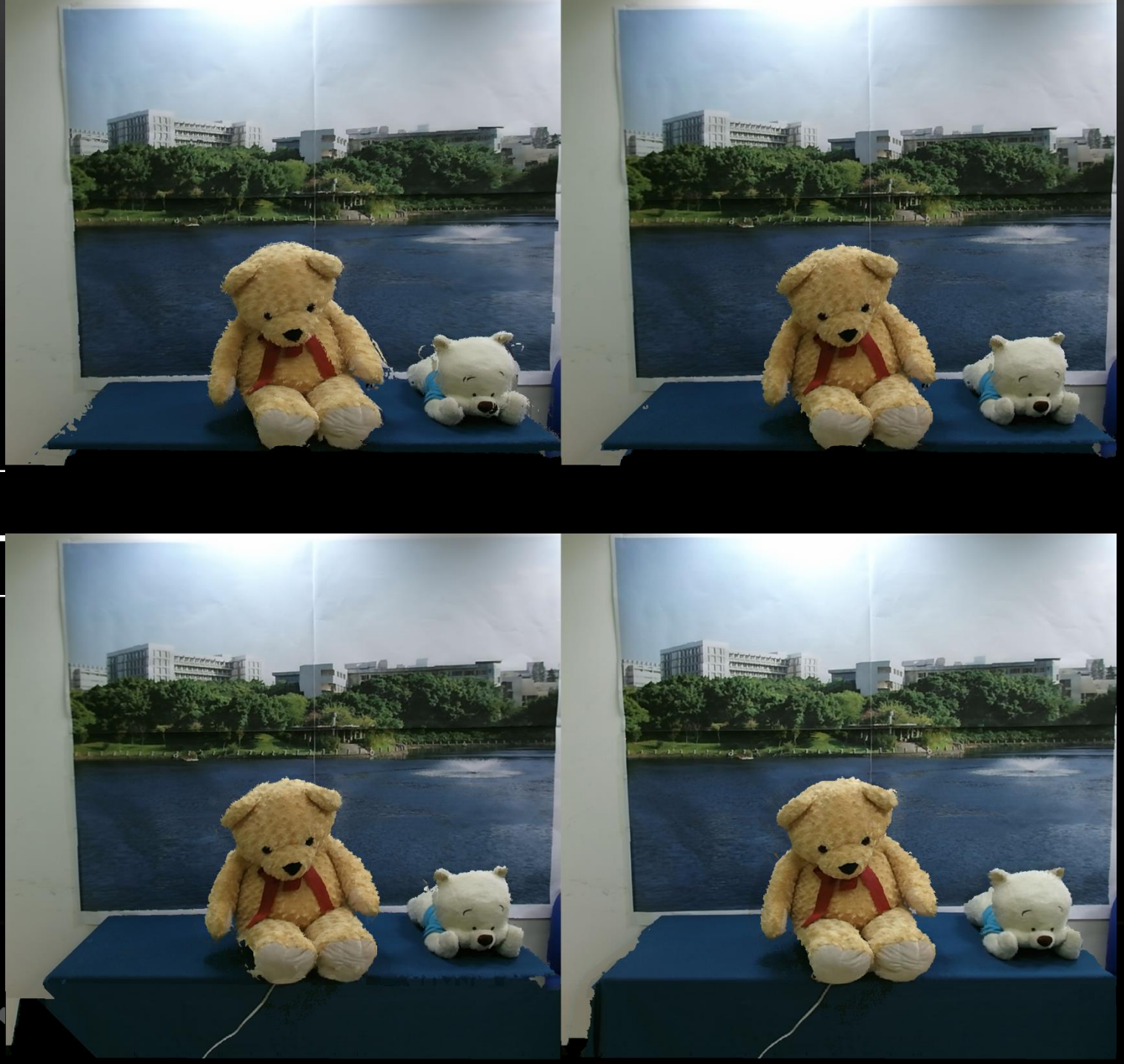
Ltc	2
3	4



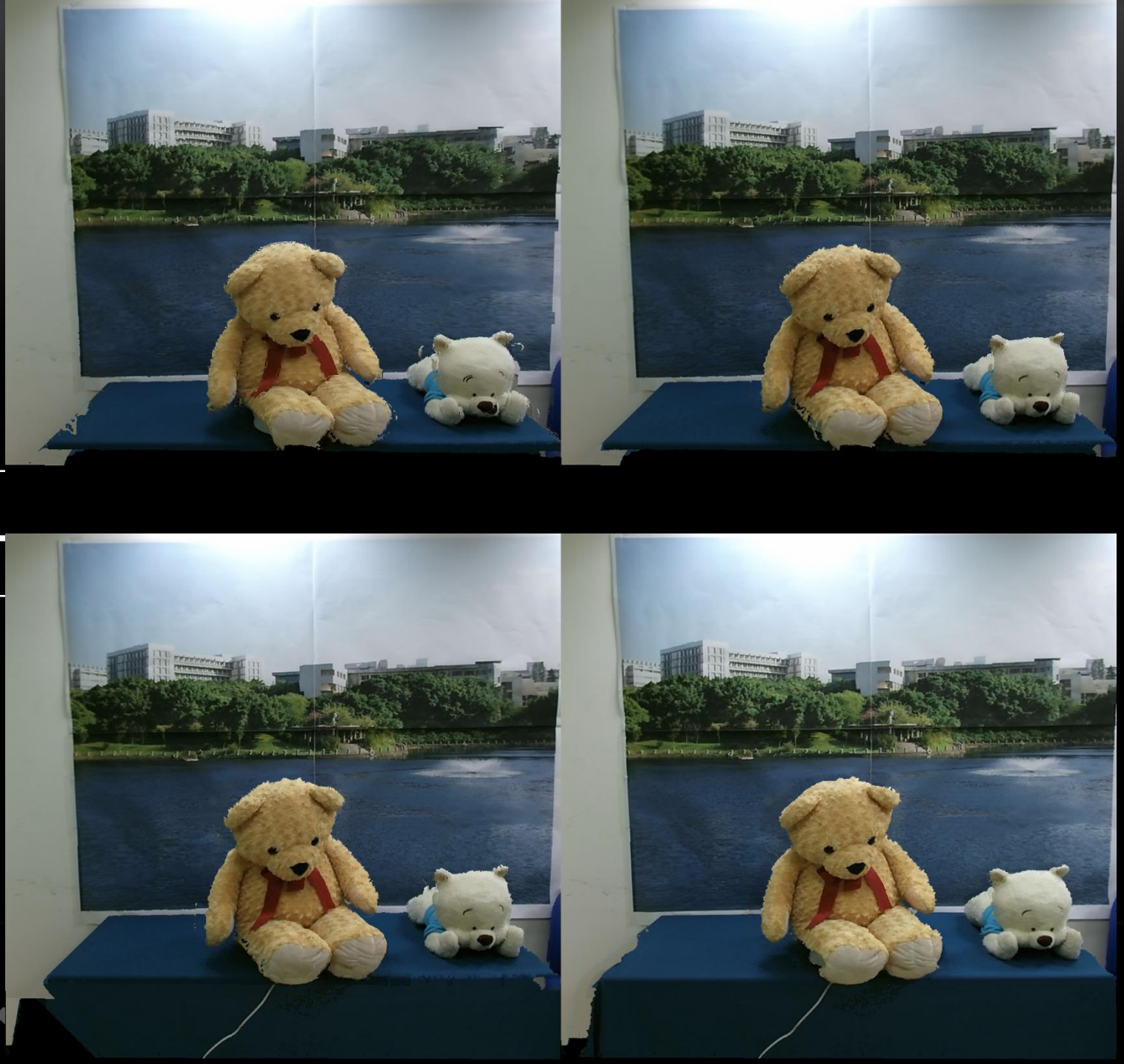
Ltc	2
3	4



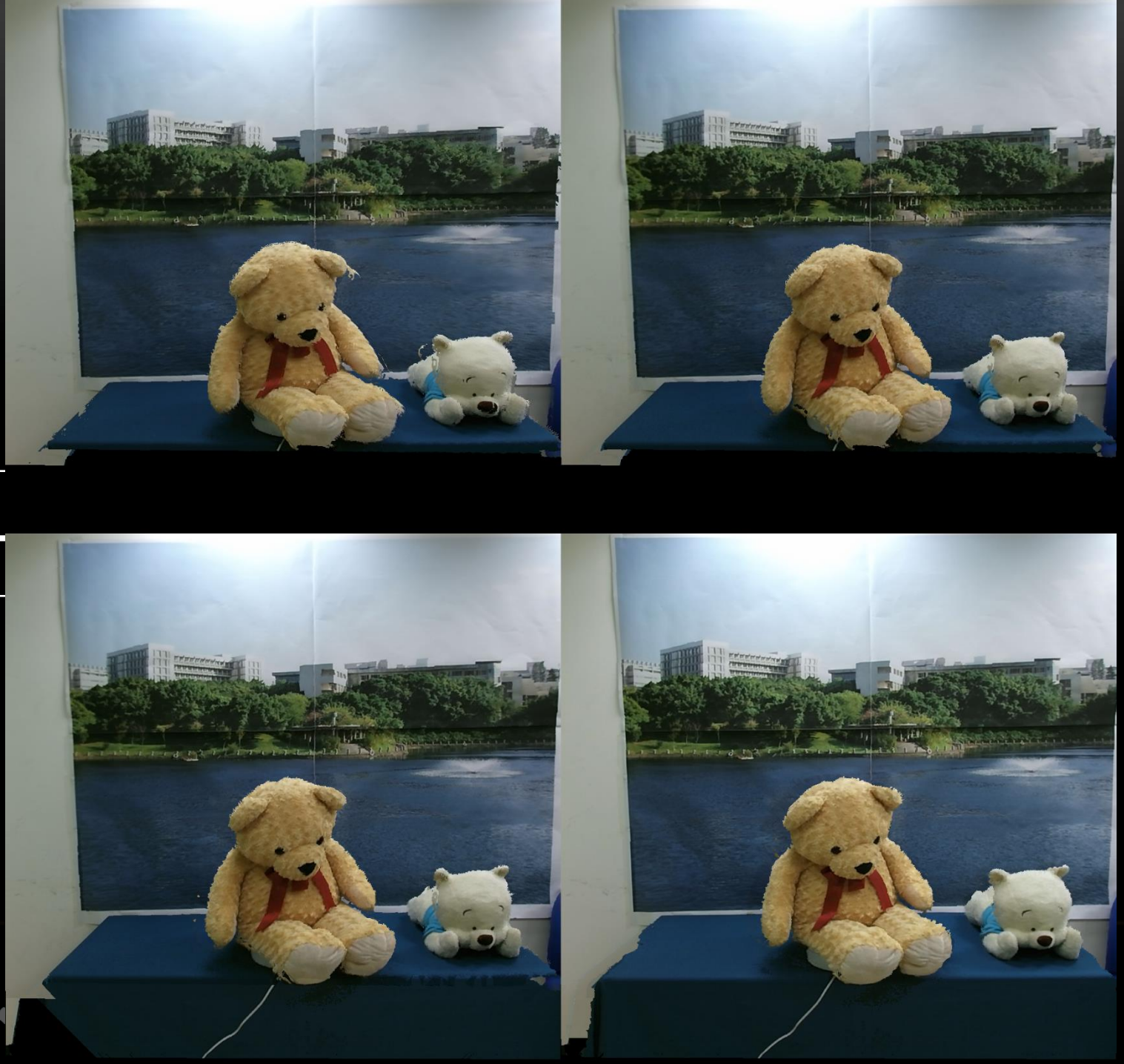
Ltc	2
3	4



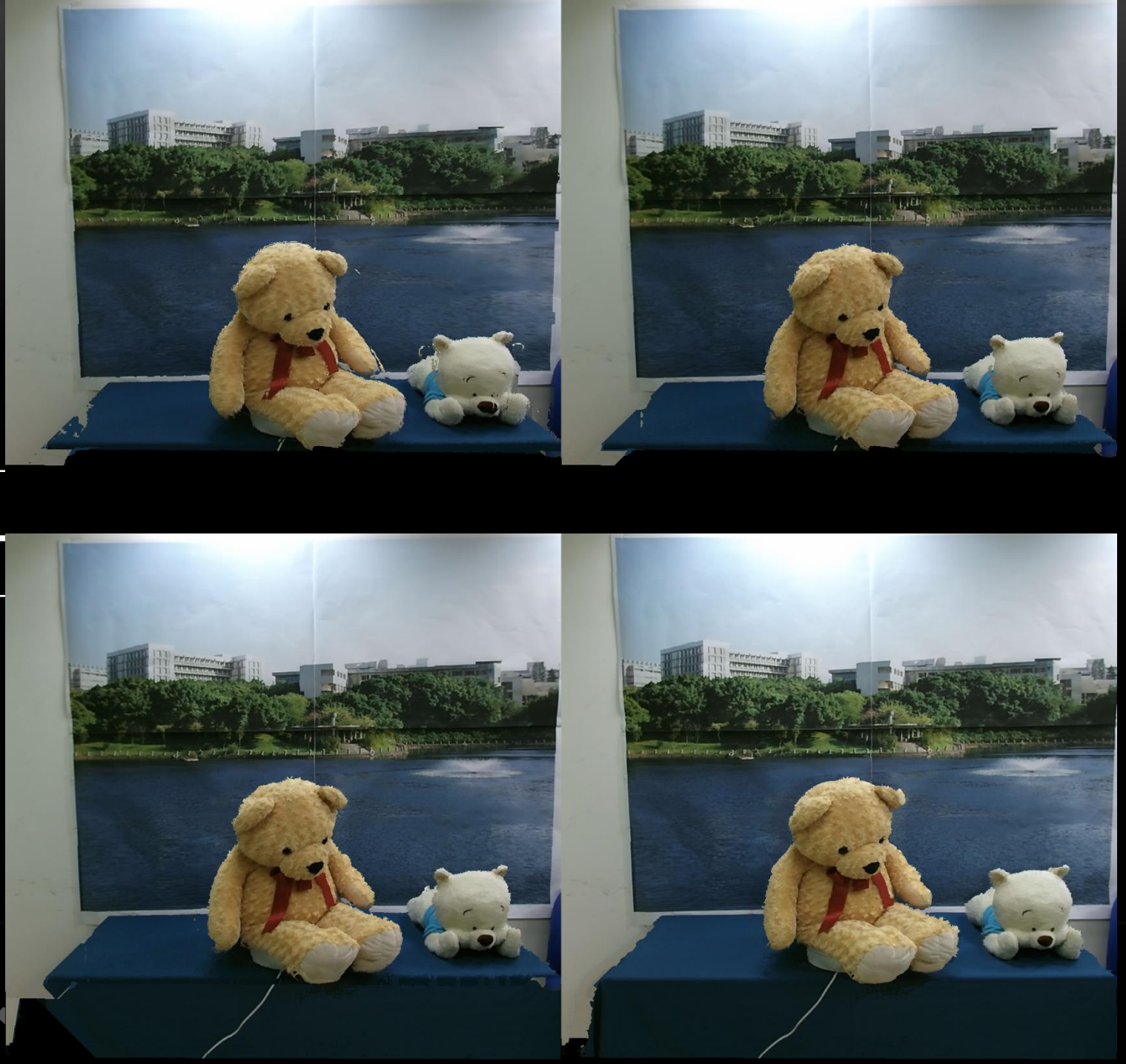
Ltc	2
3	4



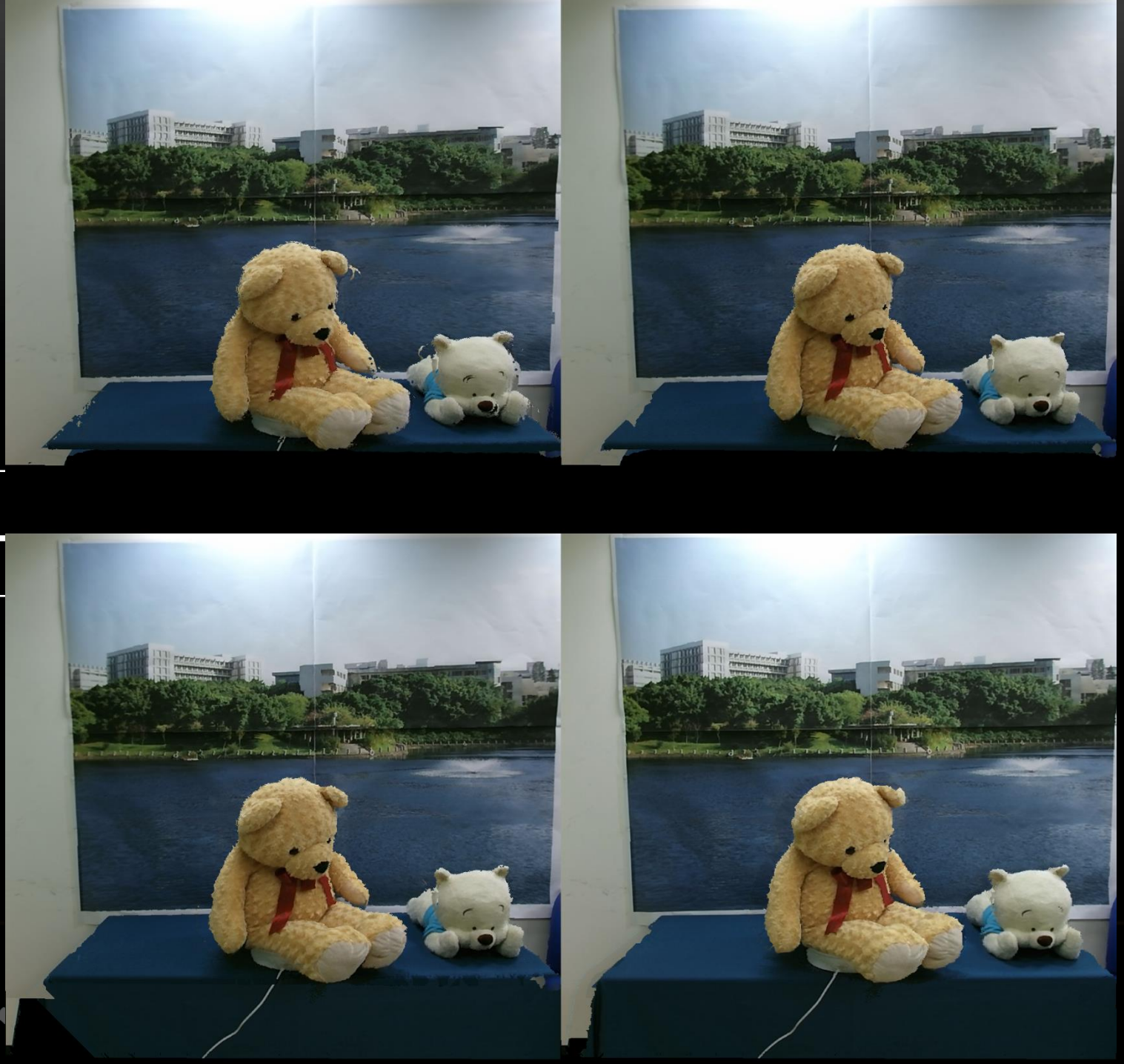
Ltc	2
3	4



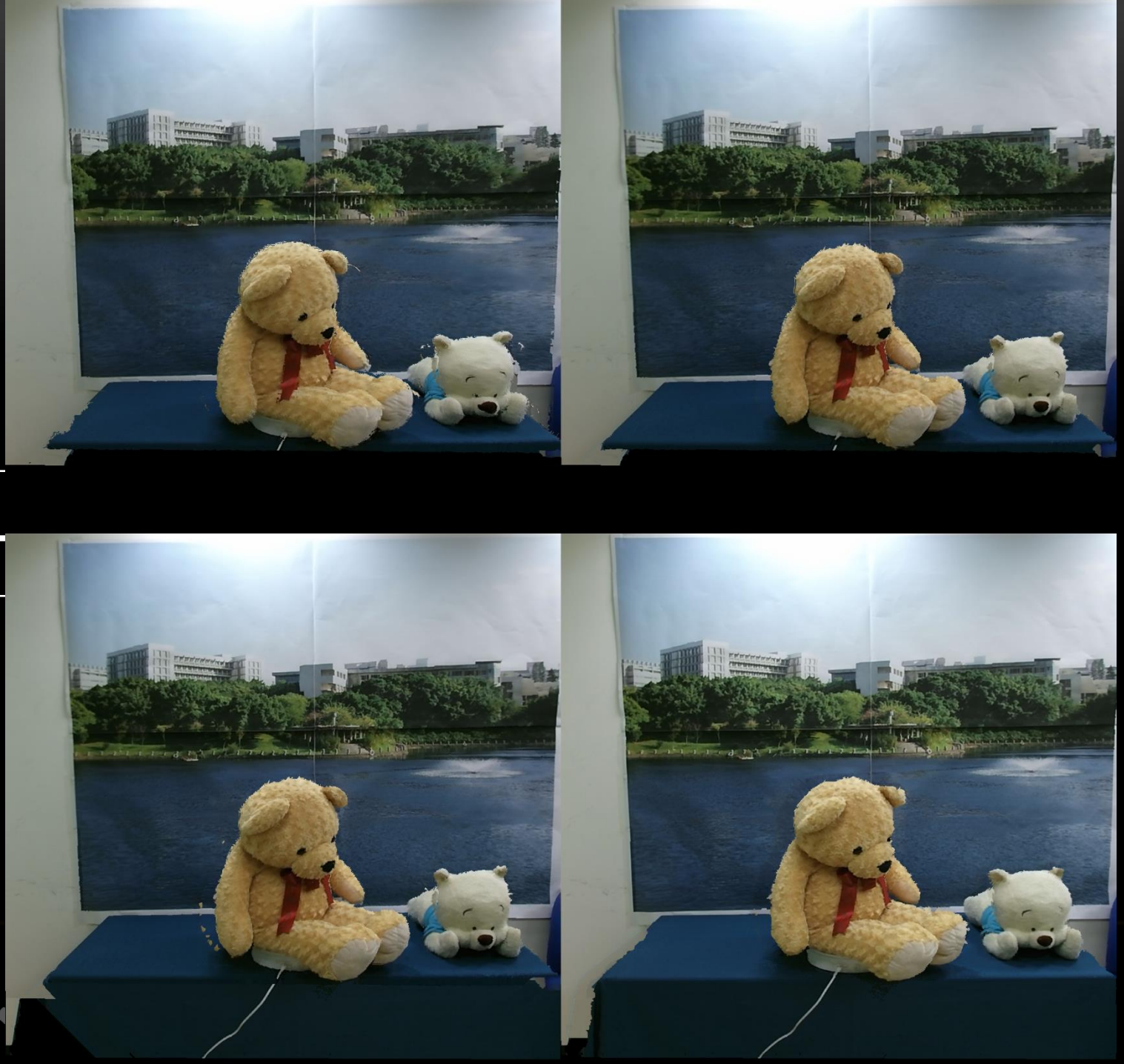
Ltc	2
3	4



Ltc	2
3	4



Ltc	2
3	4



Ltc	2
3	4

Case 5: Zoom in 400mm

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

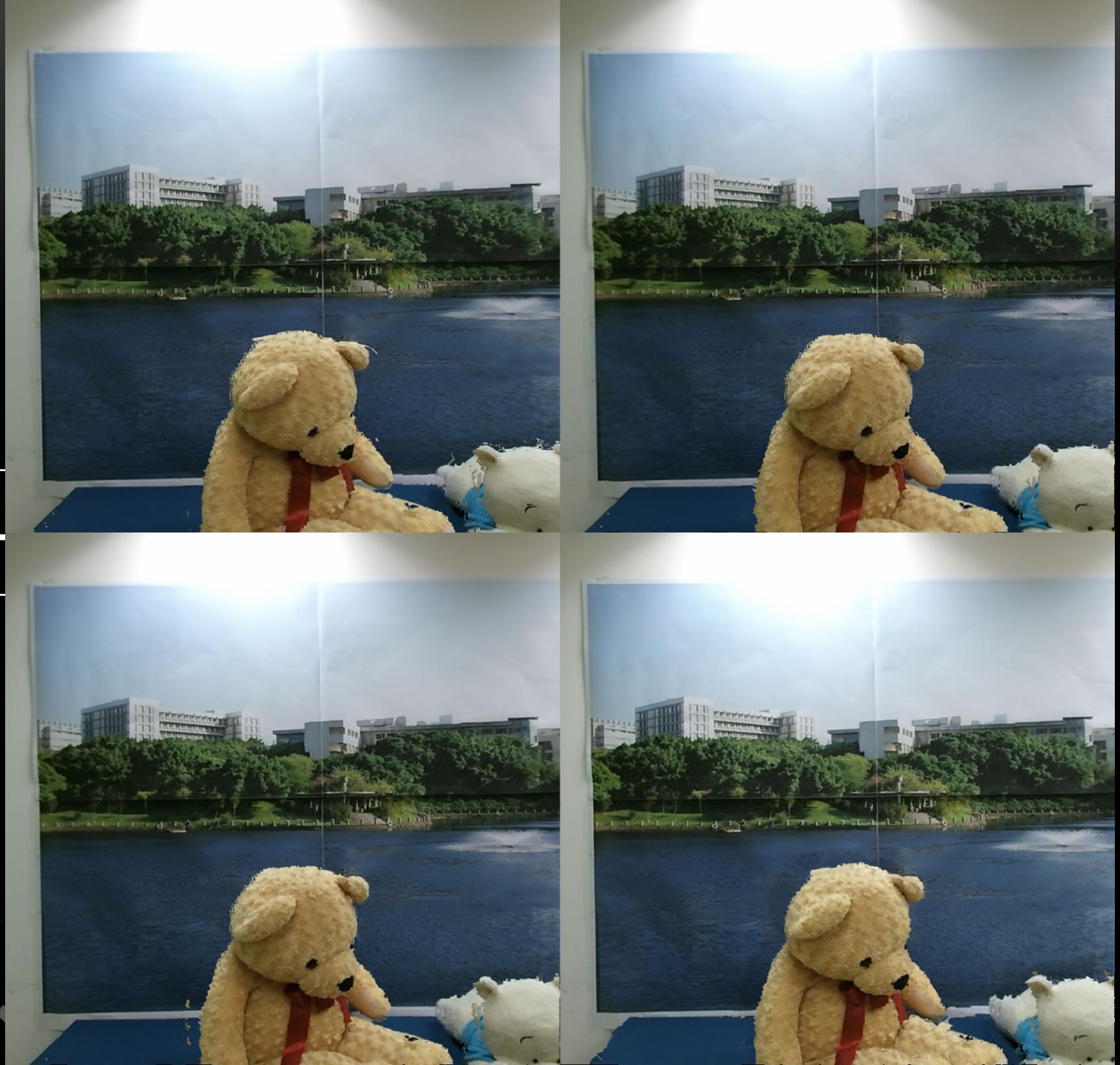
Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4



Ltc	2
3	4

Case 6: Zoom out 400mm

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Ltc	2
3	4

Conclusions

- We solved the synchronization problem between multiple Kinects.
- We proposed multi-view blending algorithm to reduce artifacts due to texture-depth misalignment and texture-texture misalignment.
- The proposed multi-view blending algorithm achieves good subjective synthesized image quality in real world cases.

Future Work

- Using the temporal information of the four warped virtual views to enhance the temporal consistency is worth to try.

References

- C. Zhu, Y. Zhao, L. Yu, and M. Tanimoto, 3D-TVSystem With Depth-Image-Based Rendering: Architectures, Techniques and Challenges, 2013 :Springer
- C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3DTV," in *Electronic Imaging* 2004, pp. 93-104.
- Yang J, Guo F, Wang H, Ding Z "A multi-view image rectification algorithm for matrix camera arrangement", Artificial Intelligence Research, Vol. 3, No. 1, pp18~29, November 2013.
- T.-C. Lee, C. –L. Chien, and H.-M. Hang, "Virtual view synthesis using quality refiement," in *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON)*, July 2016.
- Computer system time calibration software. Available at http://www.stdtime.gov.tw/chinese/EXE/NTPClock.exe

Thank you for your attention!