Collaborative View Synthesis for Interactive Multi-view Video Streaming

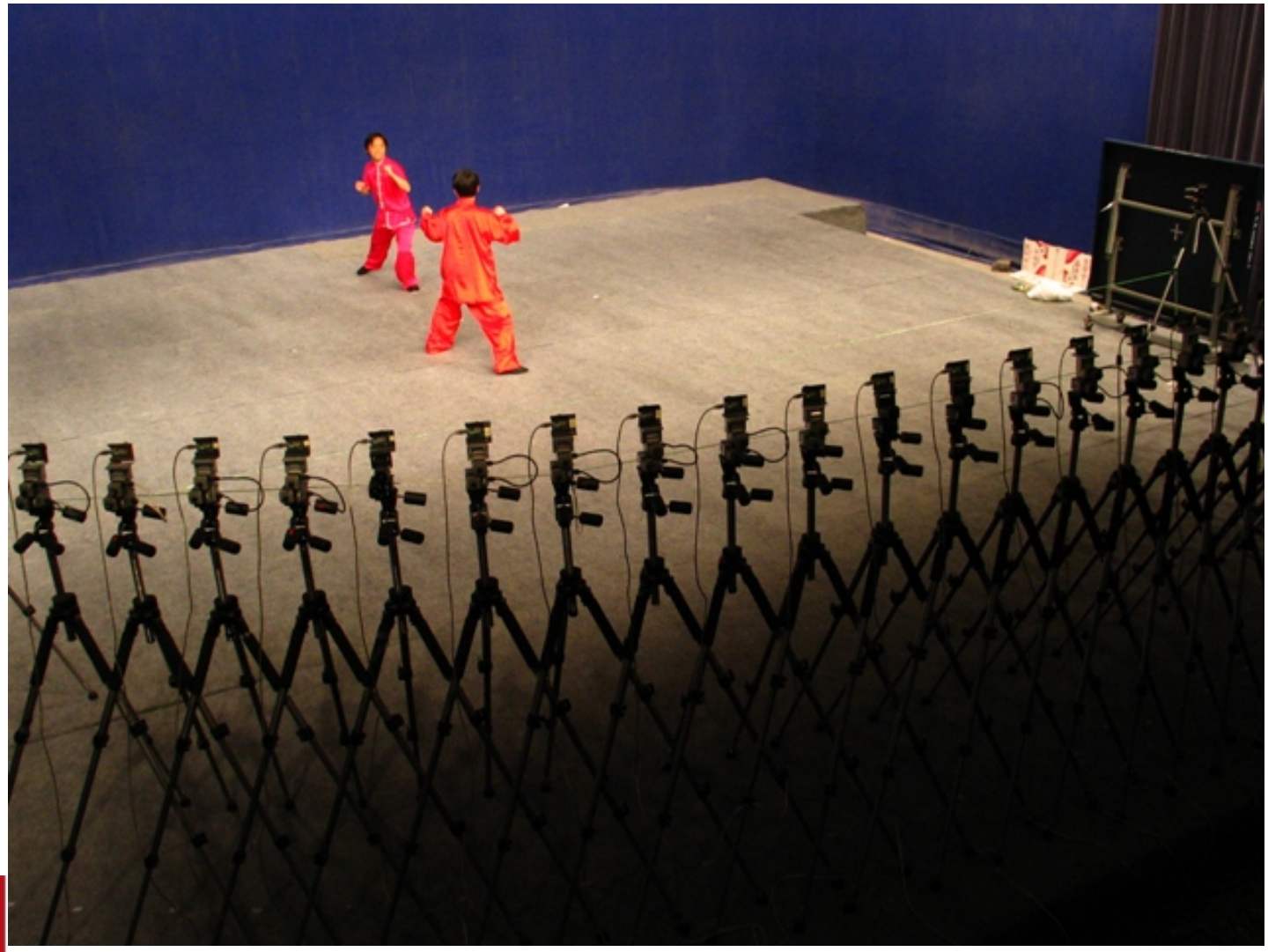
Fei Chen, Jiangchuan Liu, Yuan Zhao, Edith Cheuk-Han Ngai

SFU SIMON FRASER UNIVERSITY
THINKING OF THE WORLD

Outline

- Background
- System Description
- View Synthesis Collaboration Strategy
- View Synthesis Algorithm
- Evaluation
- Conclusion

Background



Background

Character

Multiview video enables users to enjoy the video from different viewpoints.

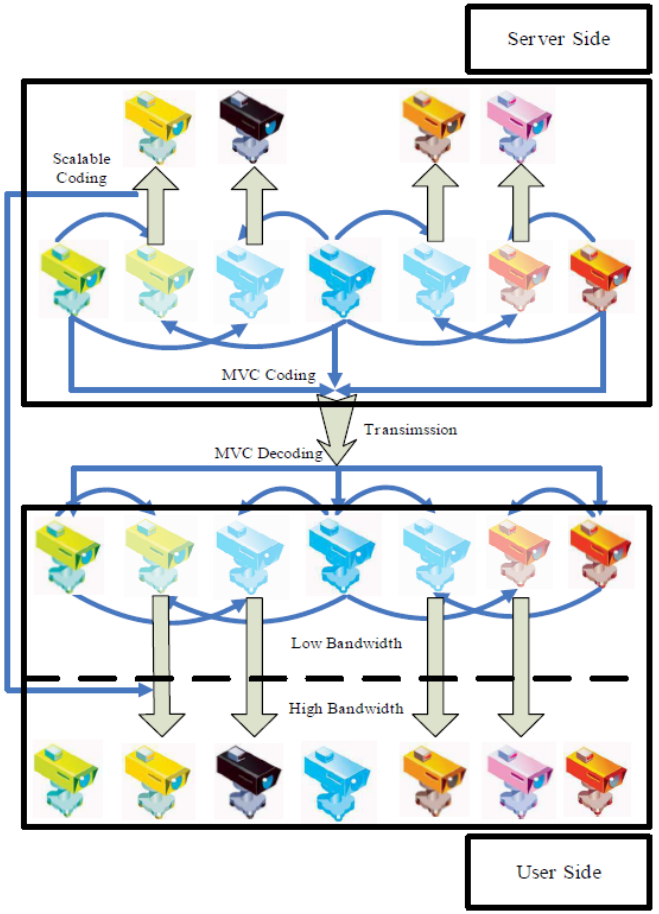
Requirements for multiview streaming

- **Rendering Quality:** to reduce disparity of interview and smooth the view sweeping process.
- **Efficiency:** to guarantee the availability of interactive application
- **Bandwidth Scalability:** adaptive to available bandwidth of users

System Description

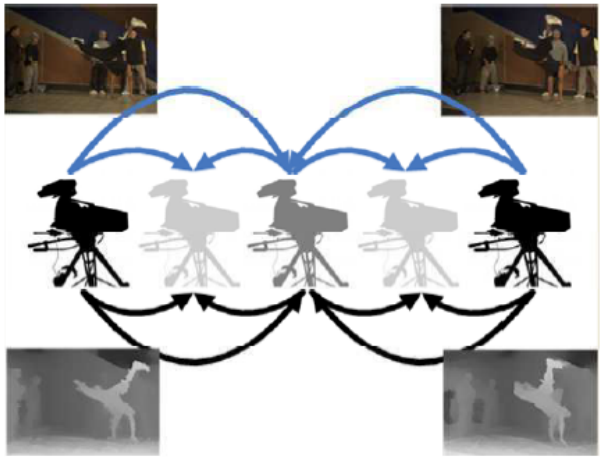
Methods

- Multi view video coding
- Scalable video coding
- View Synthesis Collaboration Strategy

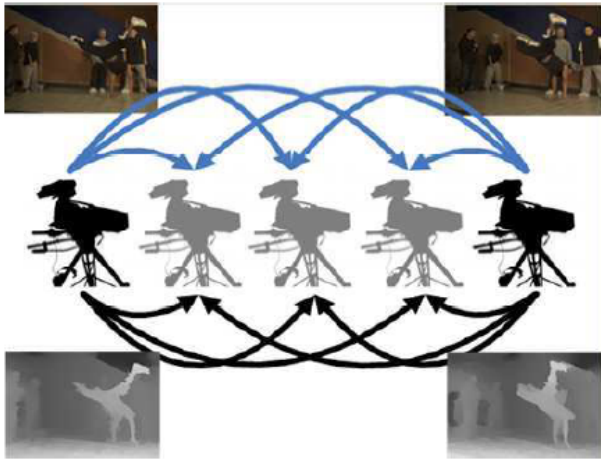


Multiview Streaming Structure

View Synthesis Collaboration



Middle synthesis

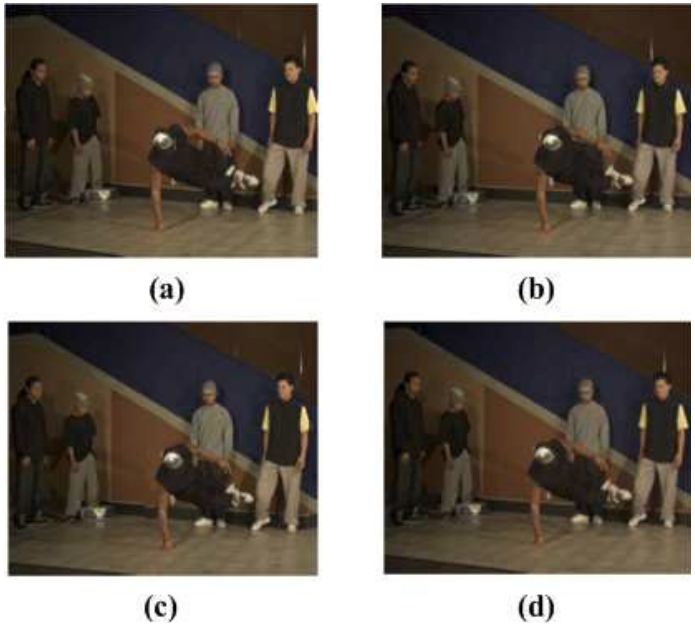


Shift synthesis

Different strategies to generate visual views

View Synthesis Algorithm

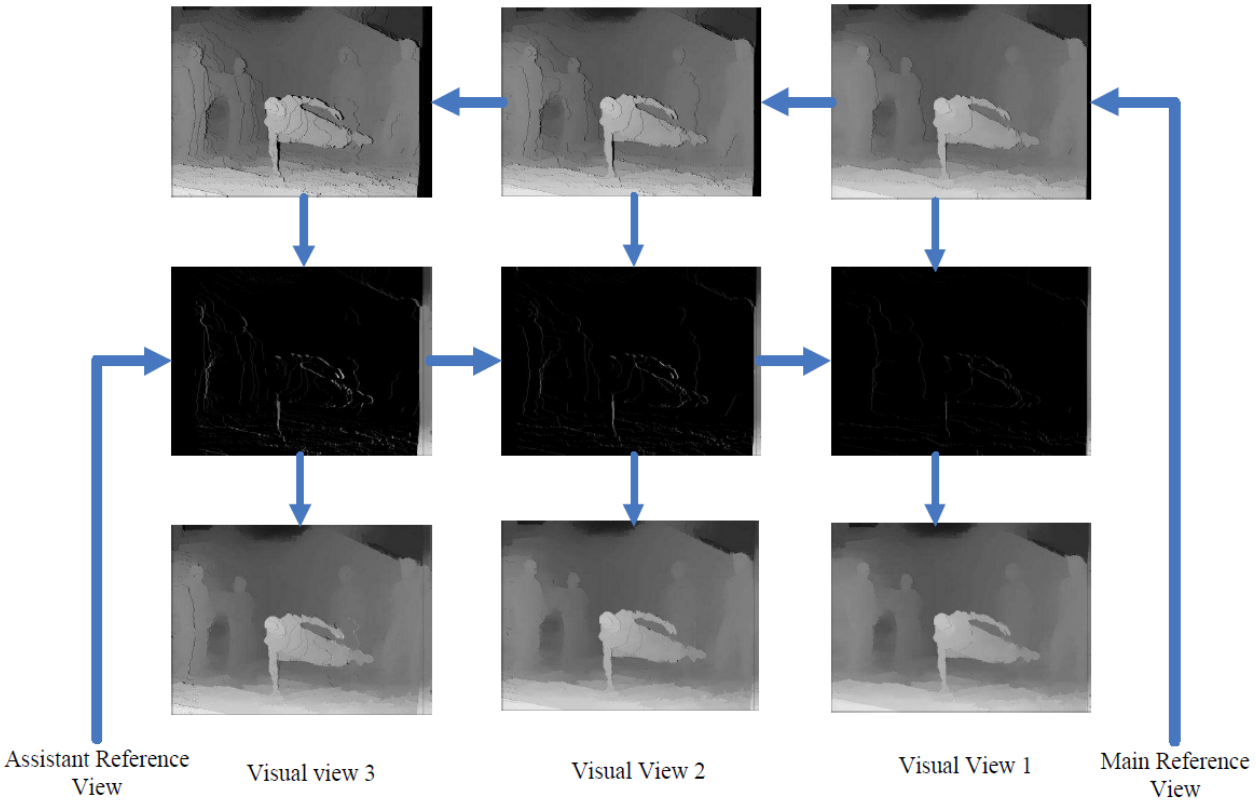
- DIBR: Depth image based rendering



(a)Left Reference View (b)Right Reference View (c)Synthesized View (d)Original View

View Synthesis Algorithm

- S-DIBR: Shift depth image based rendering



S-DIBR

- The shift value from the main reference view to the visual view is:

$$S_0 = \frac{fdn}{\lambda N}, \quad where\ n = 1, 2, \ldots \frac{N}{2}$$

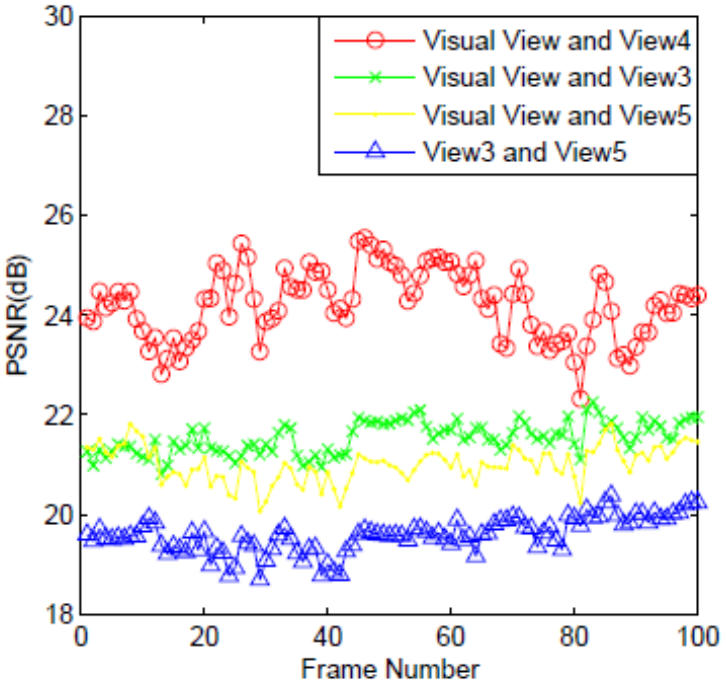
- The shift value from the visual view to the assistant reference view is:

$$S_1 = \frac{fd}{\lambda}(1 - \frac{n}{N}), \quad where\ n = 1, 2, \ldots \frac{N}{2}$$

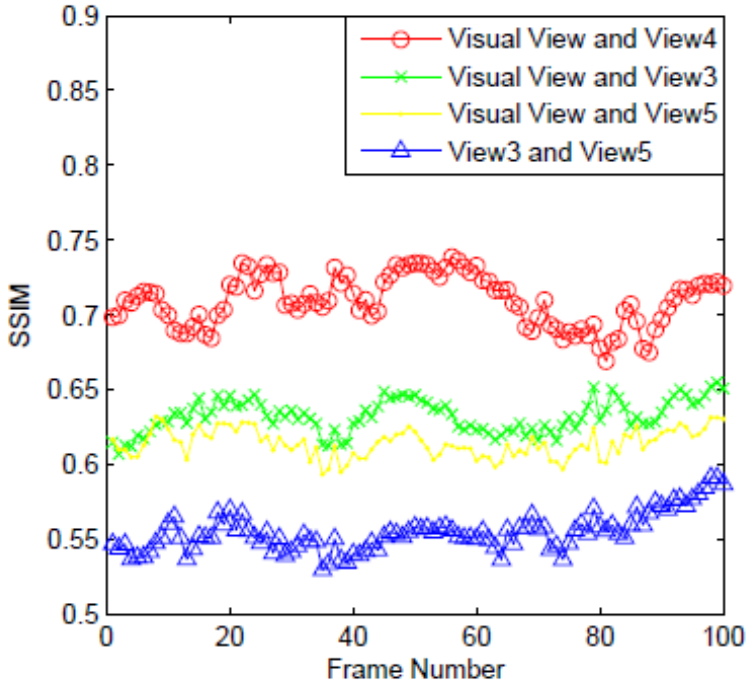
- Therefore we have following relationship:

$$\begin{cases} P_i \subset P_{i+1}, & P = P^A \\ P_{i+1} \subset P_i, & P = P^M \end{cases}$$

View Synthesis Analysis



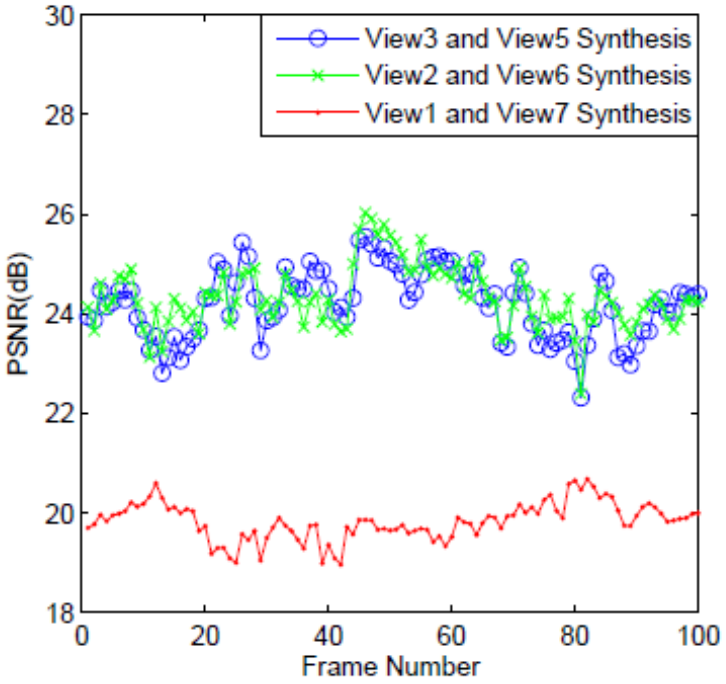
(a) PSNR



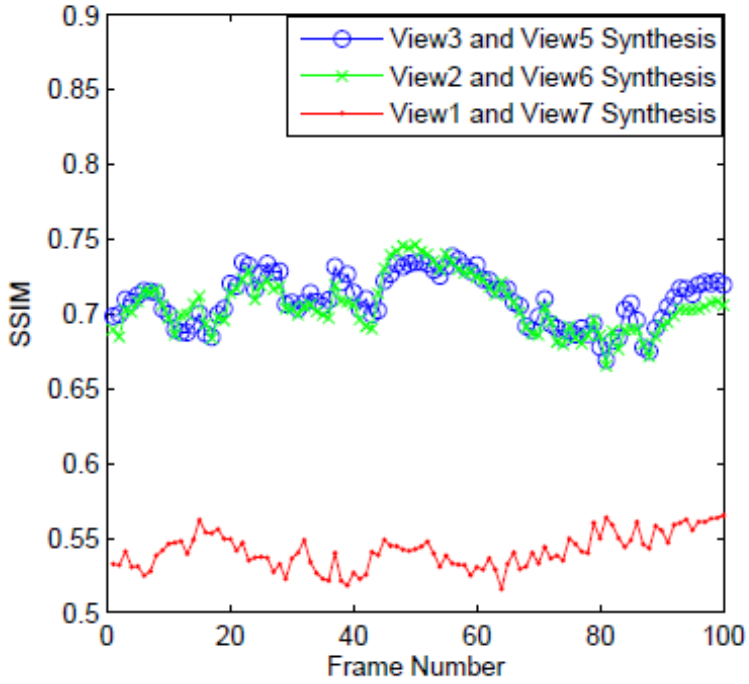
(b) SSIM

Different view comparison

View Synthesis Analysis



(a) PSNR



(b) SSIM

Synthesized with different reference views

Rendering Quality

Visual View	W-DIBR	S-DIBR
V(4)	26.5dB/0.73	25.9dB/0.71
V(4,5)	25.6dB/0.72	25.8dB/0.7
V(3,4,5)	22dB/0.63	24.7dB/0.68
V(2,3,4,5,6)	20.2dB/0.62	20dB/0.6

Rendering quality comparison between W-DIBR(warping DIBR) and S-DIBR(shift DIBR)

- Similar rendering quality for 1 visual view
- S-DIBR keeps the performance stable as the number of visual views increases

Efficiency

- The computation latency:

$$T = N(T_p + T_m + T_t + T_h) = NT_p + \delta$$

- The time cost in pixel projection for S-DIBR:

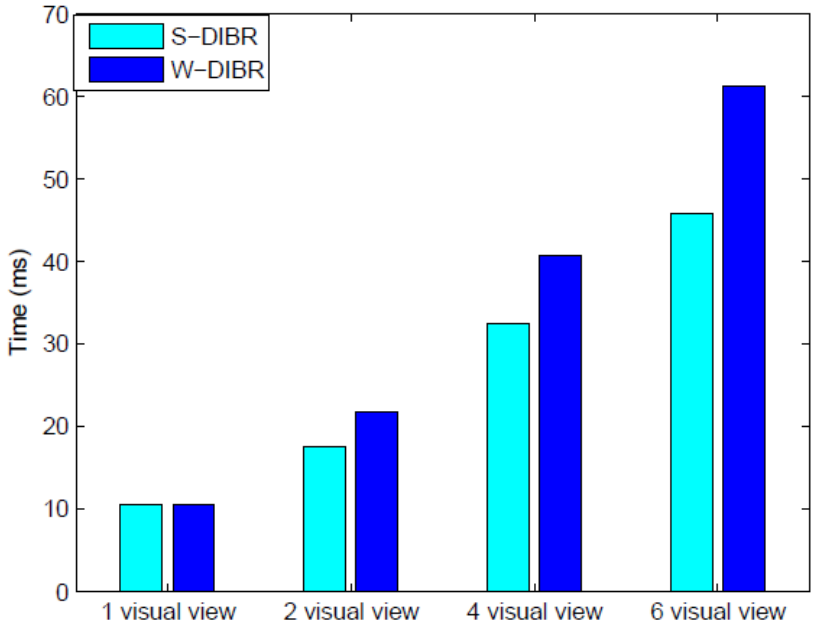
$$\begin{aligned} T_p(S) &= \sum_{j=1}^{N} T_{j,p} \\ &= 2K[(P_1^M + P_1^A) + (P_2^M + P_2^A)...(P_{N/2}^M + P_{N/2}^A)] \\ &= 2KP[\sum_{j=1}^{N/2}\prod_{i=1}^{j}\alpha_i^M + \sum_{j=1}^{N/2}\prod_{i=1}^{j}\alpha_i^A] \end{aligned}$$

- And we have the computation latency reduction:

$$\varphi = \frac{T(D) - T(S)}{T} = 1 - \frac{\sum_{j=1}^{N/2}(\prod_{i=1}^{j}\alpha_i^M + \prod_{i=1}^{j}\alpha_i^A)}{N}$$

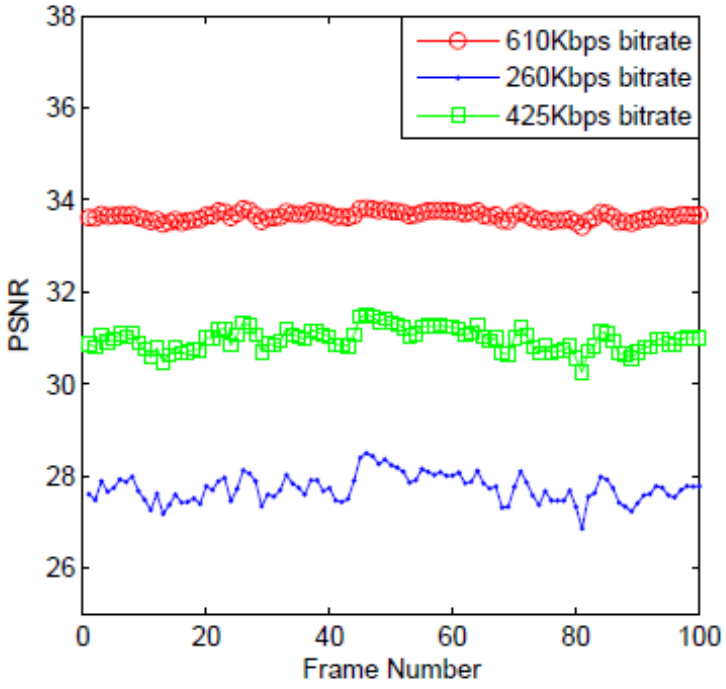
Efficiency

- The computation latency reduction aggregates as the number of visual views increases

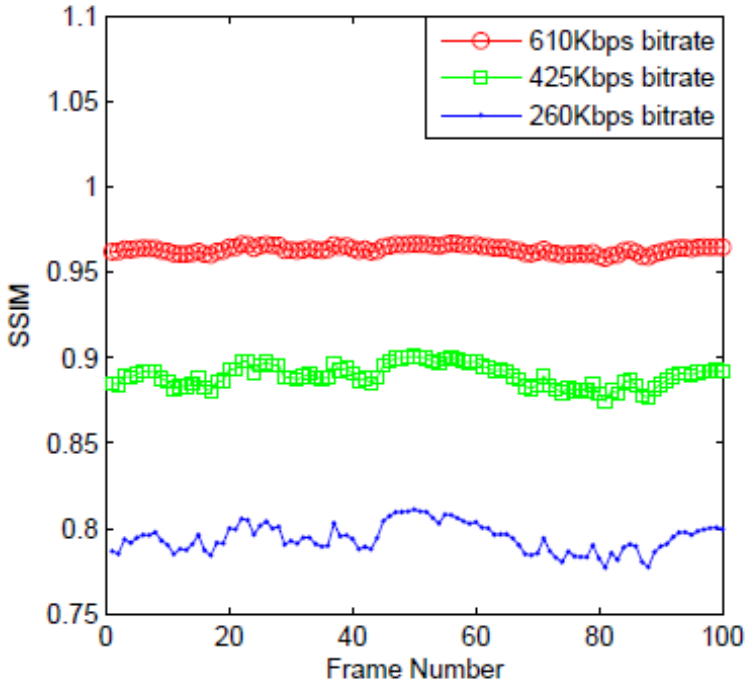


computation latency comparison

Bandwidth Scalability



(a) PSNR



(b) SSIM

scalable rendering with different bitrates

Conclusion

- A collaborative view synthesis strategy for multiview streaming system

- S-DIBR algorithm with rendering quality, efficiency and bandwidth scalability

Thanks

Q & A