Multi-Camera Acquisition and Placement Strategy for Displaying High-Resolution Images for Telepresence Systems

ABM Tariqul Islam[†1], Stephan Ohl[‡1] and Oliver Staadt[§1]

[1]Visual Computing Lab, Institute of Computer Science, University of Rostock, Germany

Abstract

Life-size high-resolution telepresence systems, used for remote collaboration, face the problem of transmitting huge data from multiple viewpoints. We present different strategies focusing on efficient camera selection and acquisition method to discard part of image data for transmission as a preprocess to classical video compression schemes. At the the receiver site, part of the omitted data can be restored by means of super-resolution methods.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Virtual Reality, H.4.3 [Information Systems Applications]: Communications Applications—Computer conferencing, teleconferencing, and videoconferencing

1. Introduction

Now-a-days, to cope with the ever-growing demand of the communication world, the use of remote collaboration has increased at a large scale. The invention of larger displays have added another exciting dimension to this collaboration environment. Although, the integration of multiple cameras in telepresence systems has made it possible to transmit and view the entire scene of the collaboration space, because of the limited network bandwidth [WOS12], there are still challenges to overcome to display life-size images on very large high-resolution display systems [NSS*06]. High-resolution source image transmission is also not encouraged due to bandwidth limitation [WOS12].

We have seen a number of works related to efficient bandwidth consumption for telepresence system such as [LWG05], [WOS12] and recently, we have seen model-driven telepresence system such as [LKB10]. Besides, there also exist various approaches as in [YWMH08] for gaining super-resolution (SR) images. We put our focus on developing a model-driven and bandwidth-efficient smart camera placement and acquisition model which captures motion data at lower resolution and enhance them, at the receiver site, for obtaining higher resolution image using SR meth-

ods. Our work is in the progressive stage: camera acquisition and setup model is being developed and tested and different parts of the method are being refined and optimized.

2. Problem analysis

Our problem domain consists of two sites – local and remote, a display screen on both sites, virtual space between the sites and cameras, placed at certain positions, to capture the communication space. In our setup, see Figure 1, the local-site user is looking at an image of the remote-site user on the display wall, whose images are captured by remote-site cameras.

The recorded resolution of the remote user depends on her position from the camera. The needed resolution, on the other hand, depends on the viewpoint defined by the local user's position. A back-channel could be introduced to send the local viewpoint to the remote site. However, this introduces extra latency. Our goal is to avoid sending data that contribute only little to the reconstructed viewpoint.

3. Our strategy

We propose several strategies for efficient camera acquisition and placement which support low-resolution image stream capture (i.e., bandwidth consumption is reduced); at a later stage, an SR method is applied for enhancement.

[†] tariqul.islam@uni-rostock.de

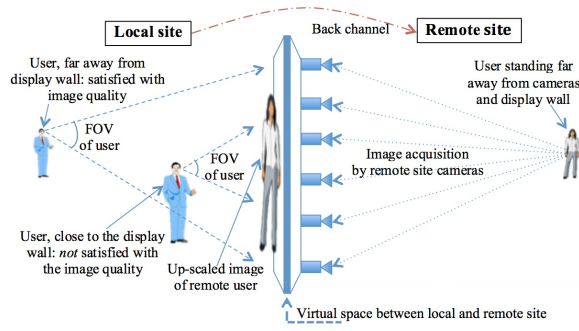[‡] stephan.ohl@uni-rostock.de

[§] oliver.staadt@uni-rostock.de

Figure 1: *Illustration of display resolution considering different positions of the local user.*

3.1. Camera selection based on surface angle

In this strategy, we select those cameras which make minimum angle, ψ in Figure 2, with the surface of the user's image patches. We use tracking position and camera geometry to calculate ψ. We transmit the entire scene covered in the FOVs of the selected cameras and combine them for achieving a SR image. This strategy is applicable when there is *no* back channel between the local and remote sites to transmit the viewing angle of the local user; otherwise, cameras can be selected by estimating the view angle using the virtual camera strategy from [WOS12]. The camera weight function can be expressed as Eq.1:

$$w(C_i) = \sum_{p \in P} \frac{1}{1 + |\psi(p, C_i)|} \tag{1}$$

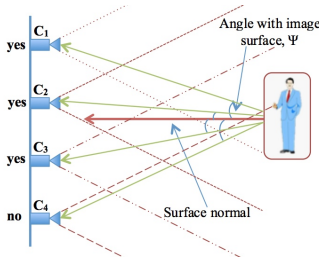with $w(C_i)$ = camera weight for camera C_i and $\psi(p, C_i)$ = angle between camera and surface normal of patch p.

Figure 2: *Camera selection based on camera angle with image surface normal.*

3.2. Framerate adaptation

We explore the possibilities to run different cameras of the camera array at different framerates. This could be done in an unstructured data-driven fashion or with a predefined interleaving pattern, Figure 3(a). At the local site, the image

based rendering algorithms need to be modified to account for missing data. If a back channel is present, it can be used to steer the process by temporal downsampling of less important camera data only. Framerate adaptation can be implemented as a preprocess to standard compression schemes.
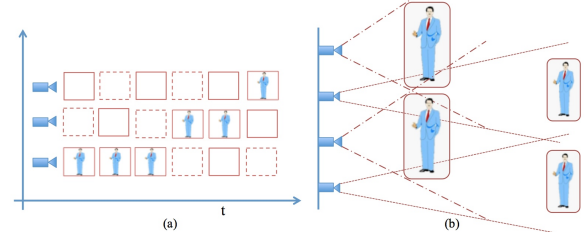

Figure 3: *(a) Framerate adaptation (b) Camera array with different focal lengths and user position.*

3.3. Optimizing camera resolution

To record a user in front of a very high-resolution display at different distances with the same resolution, a mixed focal length camera array can be used, Figure 3(b). In case of a user close to the display, a wide angle lens must be used; whereas, for a distant user, a telephoto lens would be appropriate. We investigate different regular mixed focal length camera array configurations and their implications on image based rendering algorithms. Especially, we will evaluate the effect on known blending strategies.

4. Conclusion

We have proposed a number of strategies for efficient camera selection and acquisition method to display life-size SR image, for teleconference systems, mantaining low-bandwidth usage. We plan to combine some of these strategies, such as combination of *framerate adaptation* and *optimizing camera resolution* strategies, for the collaboration system and evaluate its performance.

References

[LKB10] LIEN J.-M., KURILLO G., BAJCSY R.: Multi-camera tele-immersion system with real-time model driven data compression. *The Visual Computer 26*, 1 (2010), 3–15. 1

[LWG05] LAMBORAY E., WURMLIN S., GROSS M.: Data streaming in telepresence environments. *IEEE Tran. on Visualization and Comp. Graphics 11*, 6 (Nov. 2005), 637–648. 1

[NSS*06] NI T., SCHMIDT G. S., STAADT O. G., LIVINGSTON M. A., BALL R., MAY R.: A survey of large high-resolution display technologies, techniques, and applications. In *Virtual Reality Conference* (2006), vol. 25-29, pp. 223–236. 1

[WOS12] WILLERT M., OHL S., STAADT O. G.: Reducing bandwidth consumption in parallel networked telepresence environments. In *VRCAI* (December 2012), pp. 247–254. 1, 2

[YWMH08] YANG J., WRIGHT J., MA Y., HUANG T.: Image super-resolution as sparse representation of raw image patches. In *CVPR* (June 2008), pp. 1–8. 1