6
Results

We tested our rendering framework with frames from the breakdancing and ballet scenes provided by the Interactive Visual Media Group of Microsoft Research [38]. Each dataset consists of a sequence of 100 frames of a dynamic scene. Their system captured the dynamic scene with 8 high-resolution (1024x768) video cameras arranged along an 1D arc, spanning about 30° from one end to another. Each frame of the sequences has two images: one BMP color image of the real scene and one BMP image for the associated depth map, which was estimated with their specific segmentation-based stereo algorithm.

Along with the frames for the dynamic scenes, they provide calibration data for each of the 8 cameras used for capture, consisting of the calibration and rotation matrices. Since our framework uses OpenGL for rendering, these matrices are converted for the OpenGL analogous matrices, namely the projection matrix and the modelview matrix, respectively.

Even though the provided depth maps in the datasets cannot be considered ground-truth, their quality is acceptable for testing the effectiveness of our proposed rendering method. Figure 6.1 contains examples of one frame of the used scenes.

PUC-Rio - Certificação Digital Nº 0711269/CA

6.1
Rendering quality

In the first test devised to evaluate rendering quality, we used a single frame from the ballet and breakdancers sequences as input and set the viewpoint evenly spaced between two cameras. Figure 6.2 shows the result of this test for frame 48 in ballet sequence, using depth images from cameras 4 and 5 as input, while Figure 6.3 depicts the result for frame 88 in breakdancers sequence, using cameras 2 and 3 as input. It is noticeable how each view consistently completes the missing parts of its counterpart's occlusion areas, generating no visible artifacts.

The second test consisted in setting the virtual camera's position coincidently with a input camera's position, to check the amount of artifacts introduced by the proposed method. Figures 6.4 and 6.5 show that although

Figure 6.1: Sample data from Ballet (left) and Breakdancers(right) sequences: color images on top, and corresponding depth images below [38].

PUC-Rio - Certificação Digital Nº 0711269/CA

Figure 6.2: Synthesized images for one frame of ballet sequence. Left and middle columns respectively correspond to cameras' 5 and 4 warping results, and right column is the final result. First row: occlusion areas not identified and rubber sheets appear. Second row: we apply the proposed labeling approach.

Figure 6.3: Synthesized images for one frame of breakdancers sequence. Left and middle columns respectively correspond to cameras' 3 and 2 warping results, and right column is the final result. First row: occlusion areas not identified and rubber sheets appear. Second row: we apply the proposed labeling approach.

Figure 6.4: Virtual camera positioned coincidently with camera's 5 position in ballet sequence, and cameras 5 and 6 used as reference cameras. Left column: original image. Middle column: synthetic image. Right column: differences between real and synthetic images. Negated for ease of printing: equal pixels in white.

Figure 6.5: Virtual camera positioned coincidently with camera's 5 position in ballet sequence, and cameras 5 and 6 used as reference cameras. Left column: original image. Middle column: synthetic image. Right column: differences between real and synthetic images. Negated for ease of printing: equal pixels in white.

some differences exist between the real and the synthetic image, they are barely noticeable to the human eye. However, those differences happen mainly at objects boundaries, which can become a problem if the method is used to videos: cracks may appear during frames transitions [38].

We also tested the proposed method for videos. We built up videos for the sequences of images provided by Zitnick et al [38]: for each input camera, two videos were generated, one for color and one for depth, keeping the acquisition frame-rate of 15 FPS. Those videos then worked as input for our method, and some results can be seen in the supporting website [24].

We can see that the proposed method performs reasonably well for videos, creating smooth cameras interpolation in real-time, despite the appearance of some artifacts between consecutive frames. This behavior is expected, since we do not deal with temporal coherence during rendering.

PUC-Rio - Certificação Digital Nº 0711269/CA

Figure 6.6: Close-up of rendering result for frame 48 in ballet sequence. Left column: original photo from camera 7. Middle column: estimated depth map. Right column: visible seams below dancer caused by wrong depth estimates.

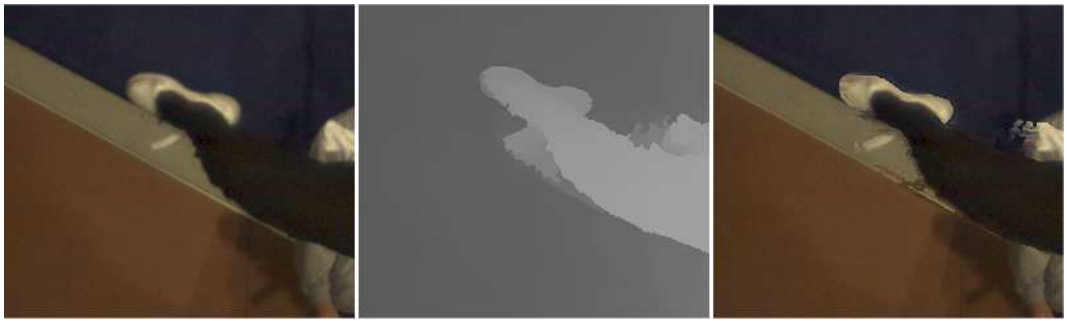

Figure 6.7: Close-up of rendering result for frame 88 in breakdancers sequence. Left column: original photo from camera 7. Middle column: estimated depth map. Right column: visible seams close to dancer's foot caused by wrong depth estimates.

PUC-Rio - Certificação Digital Nº 0711269/CA

6.2 Limitations

Errors in depth maps estimates may cause artifacts, as can be seen in Figures 6.6 and 6.7.

In the first case of Figure 6.6, depth Z for the shorts of the dancer equals the background's depth Z. Since the occlusion areas identified by the proposed algorithm is based on the erroneous depth map, they do not coincide with the real object's boundaries, and seams become noticeable.

Figure 6.7 depicts the same problem with wrong depth estimates for the breakdancers sequence, which also results in visible seams.

Another limitation of the proposed algorithm is the 'shadowing' effect aforementioned in Section 4.3, which happens on frontiers between occlusion areas and blended regions. Figures 6.8(a) and 6.8(b) are examples of this issue, that can be solved by previous color calibration of cameras. Finally, to achieve real-time performance, the proposed method does not handle the matting problem directly. Although dealing with the matting problem effectively, Zitnick et al [38] calculates mattes during a offline step. Our simplification causes a speed-up compared to their approach, avoiding the necessity of a

6.8(a): Shadowing effect for ballet sequence.

6.8(b): Shadowing effect for breakdancers sequence.

Figure 6.8: Close-ups of rendering artifacts (shadowing) for frame 48 in ballet sequence and frame 88 in breakdancers sequence.

6.9(a): Matting absence causes artifacts for ballet sequence.

6.9(b): Matting absence causes artifacts for breakdancers sequence.

Figure 6.9: Close-ups of rendering artifacts (due to matting absence) for frame 48 in ballet sequence and frame 88 in breakdancers sequence.

pre-processing step, but also creates some minor artifacts as shown in Figures 6.9(a) and 6.9(b).

6.3 Time-performance analysis

It is desirable for any IBR method that its time-performance depend solely or mainly on the input images' resolution. In other words, the complexity of the captured scene must not interfere severely on the time needed for view synthesis.

To verify this characteristic in our method, we used a testing set consisting of five sets of depth images. Images' dimensions in each set used were: 320x240, 640x480, 1024x768, 1600x1200, 1920x1440. The computer used was a workstation with a Intel®Core 2 Quad 2.4GHz CPU, with 2 GB RAM, and a NVidia®GeForce 9800 GTX graphics card. The graph in Figure 6.10 depicts the results of the test in that machine. As we can see in Figure 6.10, the performance is linear in the number of images' total number of pixels as desired. Besides, we can notice that the method can be effectively applied for rendering

PUC-Rio - Certificação Digital Nº 0711269/CA

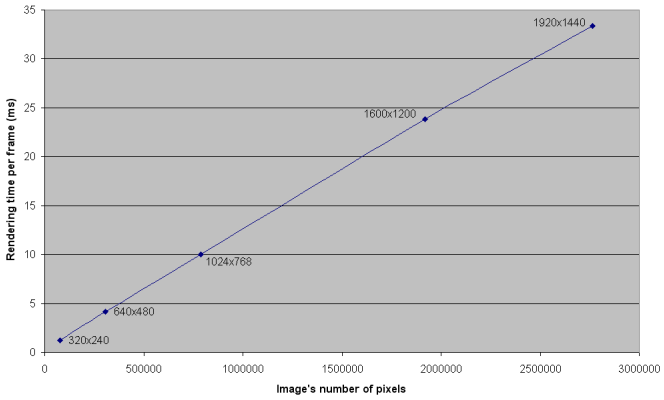

Figure 6.10: Render time X input images' number of pixels.

high-definition (HD) images at interactive rates: a 30 FPS rate was achieved using the mentioned workstation.

6.4 Summary

In this chapter we showed analyses of our method. We could verify that the use of our algorithm generates little noticeable artifacts in rendered images, but also that it has some limitations.

Besides, we could analyze the general time-performance expected for our system, and concluded that our system can be applied for rendering of HD images at interactive rates.

PUC-Rio - Certificação Digital Nº 0711269/CA