ALL-CLEAR IMAGE BASED SYNTHESIS USING CLARITY DEGREE

Xun Cao, Su Xue, Qionghai Dai(Senior Member, IEEE)
cao-x06@mails.tsinghua.edu.cn
Department of Automation, Tsinghua University, Beijing 10084, P.R. China

ABSTRACT

Image based rendering (IBR) usually produces severe artifacts, typically blur and ghost, for objects not located on the focal plane. To obtain an all clear result, previous works either endeavor to recover scene depth or rest on iterations. These methods are difficult and time-consuming, not suitable for real-time IBR applications. In this paper, we propose a novel all-clear synthesis scheme bypassing tedious depth estimation. Our algorithm directly measures the Clarity Degree of local regions in pre-rendered images and then optimal blocks are combined. This measurement is motivated by the observation that mean change energies (MCE) are well consistent with the human vision system's feeling of "clarity". Furthermore, an efficient pseudo-depth filtering is proposed to alleviate the block effect during combination. Our algorithm gives outstanding performances on both synthetic and real world data. It is also proved to be fast enough for real-time implementation.

Index Terms— Image synthesis, image de-blurring

1. INTRODUCTION

Image Based Rendering (IBR) offers a novel alternative to conventional model based rendering when handling complex scenes of difficult geometry, so IBR is suitable for building practical 3D display system. Rendering a high-quality virtual viewpoint image rapidly is critical for 3D experience of users in real-time IBR applications.

With plenoptic sampling theory [1], the scene objects located on the virtual focal plane (VFP) of rendering will be clearly synthesized. Conversely, if the real depth of objects does not match the VFP, those objects will be rendered out of focus with unpleasing visual artifacts such as blur and ghost. Then, synthesizing virtual image with objects at varying depth being all clear, i.e. "all-clear image", is an important goal pursued by researchers.

Nearly all previous works, following the plenoptic sampling theory, sought for an "all-in-focus" rendering to solve this problem. In other words, they focused on recovering the accurate depth distribution of scenes, so different objects can be clearly rendered with varying local focal plane settings, that is, "put into focus". Computing a geometry model [2, 3] totally satisfies the requirements, however, it remains a conventional computer vision difficulty. Other works assigned an accurate depth layer to each part of the image [4, 5], an anti-aliased rendering result is then obtained based on correct depth calculation. Although avoiding model construction, computing depth layers is also difficult and time-consuming. Other hardware-aided methods achieve good result [6], it is not widely feasible for different IBR systems

In this paper, we propose a novel image based synthesis scheme to fast generate an all-clear virtual image, which is very suitable for real-time IBR system application. It is motivated by the observation that "all-in-focus" (or accurate depth) is a sufficient condition of "all-clear", but not essential. Since our ultimate purpose is just to obtain visually pleasing images, the proposed method directly examines the clarity of synthesis results based on special measurement, bypassing difficult "in-focus" rendering or depth recovery. The work in [7] avoids depth estimation too, however, the iterative synthesis is slow to implement.

In our proposed algorithm, the visual quality of a region in result images is formulized by introducing a quantitative index, clarity degree (CD). Considering the nature of human vision system (HVS) that people are most sensitive to regions with intensive changes while degradation of smooth region are usually unnoticeable, we define the clarity degree in local regions by the Mean Change Energy (MCE) criterion, which is easy to calculate and proved consistent with subjective feelings. Then, with this index, we propose a block-based synthesis algorithm which mosaics high quality local regions extracted from different single-depth rendered images to composite a final result.

Different from "all-in-focus" rendering, this approach selects regions with good visual quality rather than those strictly rendered at accurate depth. Thus, in regions whose visual quality is sensitive to focal plane, the correct depth can be easily recovered by our clarity measurement. Meanwhile, other depth-insensitive parts are also good-looking even the depth information is not accurate. In practice, region partition is done in a block-based way for a faster rendering. In order to get a better result, an easy and effective de-blocking post processing is employed. Experiments of both synthesized and real scenes verify the effect of our algorithm, which is realized on-the-fly.

*This work is supported by The National Science Fund for Distinguished Young Scholars (No.60525111) and NSFC for Key Program (No.60432030).

The remainder of this paper is organized as follows. In Section 2, the concept of clarity degree and block combination method is detailed. Section 3 describes the complete real-time synthesis pipeline using light field rendering. We present experimental results in Section 4 and close with some concluding remarks in Section 5.

2. BLOCK COMBINATION BASED ON CLARITY DEGREE (CD)

2.1. Clarity Degree Measurement

As discussed above, quantitative definition of clarity degree is a key problem in later clear block selection from multiple single-depth rendered images. Determined by IBR algorithm's principles, e. g. light field rendering (LFR) used in our system, major factors affecting clarity and sharpness of single-depth rendered images are the existences of blur and ghost (double image), see Figure 1. Thus, it is required that the clarity degree has strong ability to distinguish the clear parts of an image from the blurred and ghosted regions, especially those regions with intensive changes which are much more noticeable for HVS

Fig.1. Artifacts in a single-depth rendered image. Right-top: the ghost or double image; right-down: blur

Given a region *R(x, y)*, we use the magnitude of its derivatives to reflect the changes, namely $|\nabla R(x,y)|$. To reduce interference of noise, a threshold T_{ch} adaptive to local regions is set to remove indistinctive changes. It is simply defined as $|\nabla R(x,y)| + \beta(\max|\nabla R(x,y)| - \min|\nabla R(x,y)|)$, where β can be preset or set interactively (β varies with the noise level). With this threshold, all pixels in *R(x, y)* can be classified into two categories: if $|\nabla R(x,y)| > T_{ch}$, pixel (x, y) is called a "change point" where prominent changes happen; otherwise, if $|\nabla R(x,y)| \le T_{ch}$, (x, y) is called a "smooth point". All change points compose a set *ChSet*.

By summing the magnitudes of derivatives at all change points, we obtain the Total Change Energy (TCE) which reflects the global extent of major changes in this region. Generally, a sharp rendered result has a larger value of Total Change Energy than a blurred one. However, ghosted effects, different from blur-effects, produce double edges phenomenon usually, which increases the Total Change Energy. So, TCE is not efficient enough to tell all kinds of artifacts apart.

Due to the rendering algorithm, the picture of virtual view is synthesized from a set of pre-acquired images which are multiplied by a weigh and then summed up. So, although two ghost edges coexist, they both have weak energies. Then, we solve this misjudging problem by averaging the Total Change Energy onto all change points. We introduce Mean Change Energy (MSE) to describe this:

$$MCE = \frac{\sum_{(x,y)\in ChSet} |\nabla R(x,y)|}{\| ChSet \|} \quad (1)$$

where ||*ChSet*|| indicates the number of elements in *ChSet*. Then we define the Clarity Degree in a region by this MCE value. The higher MCE is, the clearer this region is. This measurement is very effective at selecting clear regions out of both blur and ghost ones, because blurred blocks are distinguished through cumulating prominent change energies with threshold and ghosted blocks are excluded through energy averaging. Figure 2 gives an example.

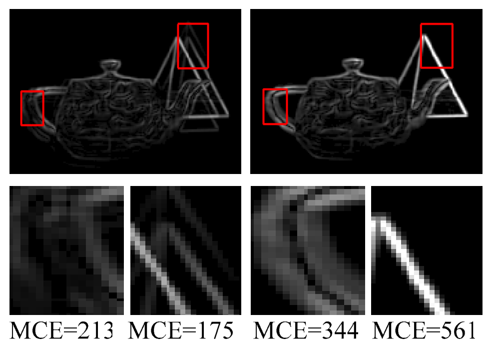


Fig.2. Clarity Degree Measurement by MCE. Top-left: the derivative of a single-depth rendered image; Top-right: the derivative of an all-clear image; The second row: local regions and their MCE values. The higher MCE reflects better clarity.

2.2. Block Combination

With Clarity Degree measurement, synthesizing an all-clear image amounts to selecting clear regions from multiple pre-rendered images at different focal plane to compose a whole. For efficiency, a block-wise partition is employed. For each block in result image, the block with the largest MCE among all pre-rendered ones is selected to fill in.

This combination procedure is easy and efficient without high level information. Because of its definition, clarity degree is mainly effective to blocks full of edges so that they are always recovered by those rendered at correct depth. Although clarity degree measurement is less effective to homogeneous areas, these blocks are not sensitive to the fluctuation of depth selection and still exhibit a moderate appearance, because the degradation in flat regions is less noticeable to HVS. An easy and effective post-processing to alleviate block effects in combination will be detailed in the next section.

3. ALL-CLEAR SYNTHESIS PIPELINE

In this section we will depict the whole pipeline to implement real time all-clear synthesis using the proposed algorithm in practical IBR system.

Step 1. Pre-rendering

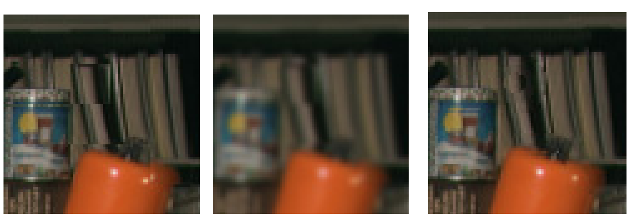
With the input sample images, a series of single-depth rendered images are first created, which are all partly clear with artifacts somewhere. Typical image-based rendering techniques are employed, such as light field rendering [8].

Step 2. Block combination based on Clarity Degree

First partition all pre-rendered images into blocks. In each of these blocks, a Clarity Degree is calculated using equation (1). The blocks located at the same position from different pre-rendered images compose a block set. Among each block set, the block with the maximum CD value is selected to fill in the result image as an optimal one. Record the depth value of the optimal block and all the block-wise depth values form a pseudo-depth map, which is not geometrically accurate but very useful for post-processing.

Step 3.Pseudo-depth map filtering

To alleviate block effects in combination, a smooth filtering is adopted in the pseudo-depth map rather than in the result image. Then we replace pixels whose depths are modified with those pixels in pre-rendered images at new depth. It is more effective to reduce the discontinuity in block depths than in block pixels, because smoothing images causes serious blur while smoothing the depth of rendering maintains visual quality very well. Figure 3 demonstrates the effect of this method.



(a) (b) (c)

Fig. 3. De-block post-processing. (a) Results with obvious block effects. (b) Filtering directly on pixels. Serious blur degrades the image. (c) Filtering on pseudo depth map. Block effects are alleviated with little loss of clarity.

4. EXPERIMENTAL RESULTS

In the experiments, we use dynamical light field rendering [9] as the IBR algorithm. Both synthetic and real world light filed data are used to evaluate our algorithm.

The light field data of two synthetic scenes (see Figure 4 (a)~(g)) are both produced by 3DMAX 7. One of the real world data is the University of Tsukuba “head and lamp” data set (Figure 4 (h)~(j)) ; The other real data, “Akko and Kayo”, is published by Tanimoto Laboratory, Nagoya University, see Figure 4 (k)~(m).

The first synthetic scene is a challenging test because the grid texture is very sensitive to artifacts, which are illustrated in all the single-depth rendered images, see Figure 4 (a)~(c). Our algorithm demonstrates a very good performance as seen in Figure 4 (d). The teapot scene verifies the ability of our Clarity Degree Measurement to distinguish both blur and ghost.

Another two more complex real world scenes containing more details and textures offer bigger challenges. The proposed algorithm cooperating with pseudo-depth filtering is still verified very powerful to create all-clear and sharp results.

Furthermore, the algorithm can be implemented very quickly in our experiments. Given pre-rendered images, the preconditions in almost all “all in-focus” algorithms, the time to synthesize one image is less than 0.11s by our MATLAB program on CPU Pentium2.4GHZ PC. Considering the fact that our light field pre-rendering is done at the speed of 30fps, the whole pipeline can be totally implemented real time with C++ codes and optimization.

Due to the absence of accurate depth information, the quality of the synthesized images is still imperfect. However, considering the tradeoff between quality and computation time, we've got acceptable rendering quality on-the-fly.

5. CONCLUSIONS AND FUTURE WORK

In this paper, we propose an efficient and fast all-clear image based synthesis scheme. By introducing Clarity Degree measurement, we can directly discover local regions of high visual quality and then results can be obtained by an easy combination. Due to the consistence with HVS, our Clarity Degree measurement is very powerful to handle all kinds of complex situations. Cooperating with a novel pseudo-depth filtering to reduce block effects, our algorithm gives visual pleasing results even on very challenging data set. The low calculation complexity guarantees that our algorithm is suitable for real time IBR applications. To apply this scheme in practical 3D systems is our future work.

6. REFERENCES

[1] J.-X. Chai, S.-C. Chan, H.-Y. Shum, X. Tong, “Plenoptic sampling”, Proc. ACM SIGGRAPH 2000 (2000) 307–318.

[2] J. Shade, S. Gortler, L.-W. He, R. Szeliski, “Layered depth images”, Proc. ACM SIGGRAPH 1998 (1998) 231–242.

[3] L. McMillan and G.Bishop, "Plenoplic modeling: An image-based rendering system." in Proc. SICGRAPH'95, 1995, pp.39-46.

[4] Yin Li, Xin Tong, Chi-Keung Tang, Heung-Yeung Shum, "Rendering Driven Depth Reconstruction. "In Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on, Volume: 4 , 6-10 April 2003. Pages:IV - 780-3 vol.4

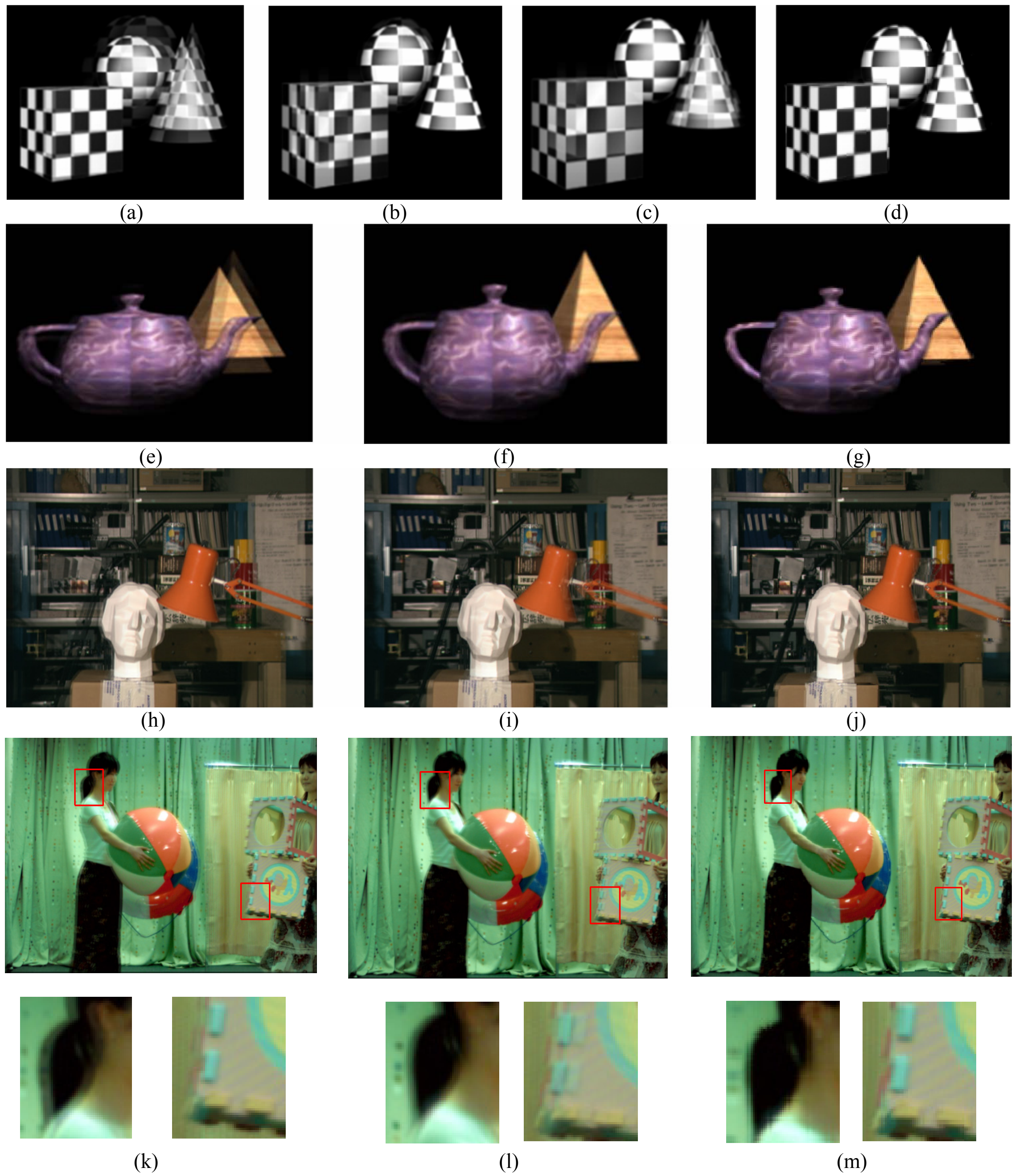


Fig. 4. (a)~(d): Synthetic scene of objects with grid textures. (e)~(g): Synthetic teapot scene. (h)~(j): Real world scene, head and lamp. (k)~(m): Real world scene, Akko and Kayo. In each of the four scenes, from left to right are: the result rendered with const depth at nearer objects; the result rendered with const depth at farther objects; the all-clear images are in the last column. Scene 1 is an exception where the result rendered with depth at every object is displayed. Good visual performances are demonstrated in our all-clear synthesized images.

[5] Keita Takahashi, Akira Kubota, and Takeshi Naemura: "Focus Measurement and All in-Focus Image Synthesis for Light Field Rendering", Systems and Computers in Japan, Vol. 37, No. 1, pp. 1--12. Jan. 2006.

[6] T. Naemura, J. Tago, H. Harashima, "Real-time video based modeling and rendering of 3D scenes", IEEE Comput.Graph. Appl. 22 (2) (2002) 66–73.

[7] A. Kubota, K. Takahashi, K. Aizawa, T. Chen, "All-focused light field rendering," Eurographics Symposium on Rendering (EGSR2004), June 21-23, 2004.

[8] M. Levoy and P. Hanrahan, "Light field rendering," in Proc.SIGGRAPH'96.1996, pp. 31-42

[9] A. Isaksen, L. McMillan, and S. Gortler, "Dynamically reparameterized light fields," in Computer Graphics (SIGGRAPH 2000) Proceedings.2000, pp. 297-306, ACM SIGGRAPH.