Soft 3D Reconstruction for View Synthesis

Primary Paper

ACM Transactions on Graphics (TOG) - Proceedings of ACM SIGGRAPH Asia Dec 2017

With the improved technology in the field of photography new methods of video visualization is being synthesized by the means of virtual-reality, 3D-movies. The major issue which needs to be addressed while constructing an object with such deep properties is that we should not lose the quality, continuity and robustness of the picture captured. For not losing the above-mentioned parameters of the image an algorithm has been proposed in this paper for synthesizing views by utilizing a soft 3D reconstruction.

This Algorithm proposed is anew Image based Rendering framework that retains depth uncertainty from a local 3D reconstruction of the scene to final color synthesis which is better than the Machine Learning approach which takes numerous training examples of images to sometime give low-resolution results and blurry images, furthermore the Machine learning approach is trained to produce final rendered images rather than consistent 3D constructions which is not fully understood until now as to how to use this method for using it to visualize or produce real time image data or to produce other effects such as mixing computer generated content with photographed content. Where as on the other hand our proposed Algorithm "Soft 3D reconstruction for view synthesis" we can produce better results with more efficiency than the existing method blending with wide range of inputs like views from plenoptic cameras, camera array videos and wide baseline image captures.

So the proposed algorithm has 6 steps in it where some of the steps follow the basic idea of getting to know the depth in the input picture and then synthesizing the volume using fast local stereo method, the steps are as follows, first, we compute the depth maps for all input views using a small set of N neighbors and a fast local stereo method, for each input view we reconstruct the scene geometry consequently we perform soft view synthesis and occlusion aware depth estimation.

Taking the new Samsung galaxy smart phone's camera features like wide selfie and virtual shot we can understand the importance of view synthesis because if you take a picture taken using the wide shot specialty then you observe that when you move the phone to-fro or from left to right the picture also moves, which is possible by different algorithms as mentioned in the first paragraph but the proposed algorithm in the paper will do it more efficiently challenging other algorithms.

one can also see that the proposed algorithm will work with wide-range of data-sets it may be structured, un-structured that is it might be the data which is stored in the form of tables or has a particular structure and the data which now a days called as the big data which is not just text data it can be any kind of data like geographical data, extra terrestrial data, DNA data and many more. Furthermore, it not only works with different types of data but also with different types of shots or captures like narrow and wide-baseline captures.

Analyzing the following points, one can strengthen the paragraph above by firstly, comparing the results on interpolation of lytro images and showing SSIM scores in a table format by analyzing those one can get the whole picture of how efficient is our algorithm working. When compared to other algorithms or other methods our algorithm which is proposed in the paper produces consistently better results for thin objects and around occlusions. Secondly, the proposed method handles wider baseline camera arrays with lot of complex content. Third, if you take unstructured wide baseline dataset the method produces results with less artifacts. The final point is that the method performs at producing ligt-field videos. As it utilizes depth maps from many views it is robust to temporal discontinuities in any given depth map, with the result being temporarily easy going even though the algorithm operating on individual frames. This is on the other hand global depth regularization systems that will vibrate largely over frames.

There will be limitations for all most all methods proposed in any field, where as the computer graphics field where image rendering is a complex task the methods which are proposed for them will most probably have limitations like how the algorithm in the paper has. The Limitation is that as our algorithm is view based we reconstruct geometry directly in projective view volumes which are synthesized in the earlier steps, this depth precision limits the amount of free view-point movement away from the source views which can be called as the midpoint or the stationary point. Moreover, this algorithm will not perform spill suppression on the front edges which will result in small color bleeding of the background color onto the foreground.

The limitations of any algorithm can be minimized by small improvements in the whole process and better ideas, in this algorithm to make it more efficient , spill suppression should be performed with precaution . the authors hope that they can use 3D reconstruction which is a vital part of the algorithm to do this work as well since it provides a good idea of the background color. The algorithm can also be made in such a way that it will support in GPU's as now it only work's in CPU's

PlenoPatch: Patch-Based Plenoptic Image Maniplulation

Secondary Paper

IEEE Transactions on Visualization and Computer Graphics (Volume: 23, Issue: 5, May 1 2017)

The relation between the first paper and the second paper is that plenoptic cameras are used to take the image so that soft 3D-reconstruction will be easy to synthesize the view as the plenoptic camera takes a 4D shot of the image through which you can make certain edits like reshuffling, re-centering. Which is not possible if you use traditional cameras and as the traditional 2d cameras doesn't take many parametric captures of the image because of which you cant easily perform soft 3D reconstruction.

With the advancement of photography in the world there is this new fascinating invention named light field or plenoptic cameras which we already mentioned in the description of the paper above. This invention has a great potential as it uses micro-lens to capture the photographic context to take in 4D light into it's memory, which is a great advantage when you are synthesizing views for various purposes like enhancing the image or making a motion image or may be constructing a soft 3d enhanced image with deepness analysis. Even though the research in these plenoptic image capturing camera is not so abundant there is an increase in demand for this cameras as it can be used to shift the view point and refocusing which is greatly used in now a days smart phones. This helps in lots of choices in editing tools which are impossible if you take 2D capturing cameras.

What one should understand in this paper is that without the creation of plenoptic cameras 3D movies and other various fascinating features like wide selfie virtual reality is not possible because just by taking a 2D image one cant perform the tasks like retargeting, refocusing , reshuffling and etc. taking a 3D capture which can be inferred as overlapped images one over another.

Our system is designed for consumer light field data, We therefore take advantage of the fact that the scene content of such data can be approximately treated as over lapped layers, a common approximation in consumer lenselet-based cameras. There is this feature with this camera is that when a photo is captured it will at the beginning recover the depth layers for the central view image, which is then presented to the user for specifying an editing task using rough scribbles, providing a natural user interface that is similar to common interfaces in existing photo editing tools.

Light field capture and lighting has been advanced from taking the picture using large camera arrays for light capturing to capturing using microscopes. New inventions which are developed using compact light field cameras property such as picam and lytro. Plenoptic pictures are being accessed by consumers to compile the raw data taken by those devices. Scientists studied the light field data in the 4D fourier domain, which lead to more amazing works on the usage of spectrum domain processes to analyze the dimensionality gap and renovating the complete light field from the local stack based on the Lambertian assumption.

Completion and reshuffling the data is to remove an object in the inout light field, we mark the pixels in all affected layers in the user specified region as hole pixels and complete each layer using the layered synthesis proceed described. After all affected layers are completed, the entire light field is modifying using the propagation method described / one completion result very effective and interesting in the paper.

Changing depth is the method which can be used to shift the depth of an object in a light field. This is performed in 2 steps, first we remove the user-specified object from its original layer using hole filling. Then the object and its layer masks are copied onto the new depth layer specified by the user as an initialization, followed by a similar patch-based synthesis process to blend it into the target depth layer.

In practical experiments the lytro camera was used to get the light field data. The limitations of the camera are, if the captured image has ground-plane or other non-frontal surface with a smoothly varying depth, the method will not work well. Since it does not take other neighboring layers in to the account representation scenes with contiguous dominant structures which span across multiple layers might cause artifacts in applications like parallax widening. So in the future to overcome this problems there might be a lot of research in considering additional geometric features and inter-layer correlations in the layered scene representation.

Conclusion:

The two papers have 2 different concepts which are related in a way that: using the plenoptic camera in the second paper which is an interactive system for light field editing using recent advances in patch-based image synthesis methods, it represents the light field as an image set with hexagonal layout which make it very flexible for soft 3D reconstruction for view synthesis and consequently for many applications of the process.