論 文 の 内 容 の 要 旨

論文題目　Omnidirectional Image Synthesis in Outdoor Environment by Spatiotemporal View Rendering
（時空間レンダリングによる屋外環境の全方位画像補完 ）

氏　　名　モラレス　エスピノサ　カルロス　エマヌエル

Image synthesis consists of restoring, compositing or generating image frames so that the resulting outputs are coherent and visually pleasant. It is one of the most significant objectives in computer vision, computer graphics and image processing since its applications range from video completion and inpainting to scene re-targeting and image-based rendering.

To synthesize images coherently, many works in the literature have tried to solve the geometric, photometric and temporal consistencies between the real-world input images and the synthesized ones. However, challenges increase specially in outdoor scenes, which not only contain objects with unconstrained motion and complex geometries but also suffer from variant-and-uncontrollable illumination. In addition, conventional perspective images widely used in most prior works pose a geometric restriction on the image synthesis task as the scene cannot be completely seen due to the limited field of view.

Because of these limitations, this thesis proposes a novel framework for omnidirectional image synthesis in outdoor environments. This work leverages omnidirectional images, also known as 360 degree images or panoramas, of outdoor scenes to synthesize them by a novel approach called spatiotemporal view rendering.

The spatiotemporal view rendering methodology assesses geometric cues of a scene by analyzing motion and color variation of pixels through space and time from input panorama sequences. The rationale is straightforward; any point in the 3D world can be modeled by three parameters: position, color and time. Highlighting either the position, color or time from the spatiotemporal view rendering solves the geometric, photometric and temporal consistencies, thus allowing plausible 1) omnidirectional geometric completion, 2) photometric rendering or 3) temporal synthesis.

To begin with, the image synthesis task of omnidirectional geometric completion aims to remove holes, undesired regions from input panoramas, and consistently fill in them.

This task is handled directly analyzing the pixel motion in a sequence of panoramic frames. A spatiotemporal volume is built by consecutively stacking every input panoramic frame in time-wise order. The spatiotemporal view rendering models optimal voxel (pixel in 3D) trajectories along the spatiotemporal volume. Minimizing the color variation of a voxel trough time, optimal trajectories can be modeled to only depend of the distance (depth) from the voxel to the camera. Estimating the voxel's depth, the spatiotemporal view rendering can fill in the holes by propagation of known voxels.

The second image synthesis task aims to solve the photometric issue in outdoor environments. In outdoor scenes, the color of objects fades and blends to the environmental color. This phenomenon, called aerial perspective effect, is caused by the scattering of light due to particles of air and haze in the atmosphere. The aerial perspective becomes more noticeable the hazier the atmospheric condition and the farther the object is from the observer. The omnidirectional photometric rendering is achieved using one parameter called atmospheric turbidity to model the aerial perspective effect. The spatiotemporal view rendering employs the proposed aerial perspective model to attenuate the target voxel's color and blend it to the natural environmental light.

Finally, the objective of the omnidirectional temporal synthesis is to generate novel views in between panoramas captured at different camera locations. To this end, a spatiotemporal volume is built stacking the input panoramas at the beginning and ending of the volume, leaving the in-between space empty. The voxel motion in the spatiotemporal volume is simplified from 3D to 2D by accurately rectifying the input panoramas. Then a dense pixel correspondence between the rectified panoramas is estimated. The spatiotemporal view rendering reconstructs the spatiotemporal volume using the rectified panoramas and the computed dense correspondence. Holes in the spatiotemporal volume that were not filled by the view rendering are inpainted by geometric completion. As a result, the novel views are lifted from the reconstructed spatiotemporal volume.

This thesis proposes a framework to solve the omnidirectional geometric completion, photometric rendering and temporal synthesis needed for image synthesis. Evaluations in challenging outdoor scenes demonstrate that the framework outperforms state-of-the-art image synthesis techniques. Experiments also show that spatiotemporal view rendering can successfully be applied to a variety of applications, including video completion, scene re-targeting for image composition, photometric rendering for Mixed Reality, as well as image interpolation to generate content for Virtual Reality.