論文の内容の要旨

論文題目　Video Completion via Spatio-Temporally Consistent Motion Inpainting
(時空間整合性を利用したモーションインペインティング法によるビデオの修正)

氏　名　ローハス メナンゾロ デラクルーズ

Recent advances in the field of transportation, navigation, and virtual reality, lead to the emergence of on-vehicle cameras. Different types of cameras are mounted either on the outside, on top, on the dashboard of cars. These cameras are used to take pictures or videos of the surrounding and are used in several applications.

Virtual tours use these videos to create a virtual environment where in a person with a monitor can view and navigate. Driving simulations use such videos to create a realistic view of a certain street. Another application is the 3D reconstruction where several point of views of a certain place is necessary. Digital archiving uses these videos to document historical places.

Several issues arise from the use of street videos. The most common issue is the presence of pedestrians especially when the faces are clear enough to be recognized. Another issue is the presence of artifacts such as dead or corrupt pixels that are inherent to the camera. Some are adherent water or smudges and occluding objects on the lens of the camera or on the windshield of a car. All of these artifacts degrade the quality of the video and are problematic when used in applications that require clear view of the surrounding.

In this thesis, we address these issues by completely deleting the information (color, brightness, etc.) of the unwanted parts of the video and redrawing them using the desired values. This redrawing process is called video completion. Video completion is the process of recovering missing parts of videos either by interpolation or duplication of the known parts. The goal of video completion is a visually pleasing output video that is both spatially and temporally consistent. Spatial consistency requires objects to maintain their geometry while temporal consistency requires parts of the same object to move in the same manner.

Numerous methods have been formulated to solve the video inpainting problem. Some work directly extended image inpainting methods to videos. With the addition of a third dimension (time), existing methods result in poorly inpainted sequence, especially when both background and holes are moving. One particular approach is the use of optical flow to propagate pixels with known colors toward the hole. This approach is straight-forward, as long as the optical flow is available in the immediately succeeding of preceding frame. However in most practical cases, such as removing pedestrians in street videos, the holes extend several frames which make immediate copying of colors using motion information insufficient. To solve this problem, one approach is to also estimate the motion inside the hole.

We present an iterative optimization approach that uses optical flow to complete the color frames. Our objective is to find the optical flow and the color of the hole such that the resulting video satisfies the requirement of the visually pleasing video.

One of our contributions is the simultaneous optical flow estimation and inpainting. By incorporating a spatially varying mask function in the data term of the optimization function, we were able to estimate the motion inside the hole by basing it to the motion at its boundary. We use two smoothness constraints to achieve coherent motion among pixels. One is the spatial smoothness constraint that is enforced by using a total variation minimization among neighboring motions. The use of spatial smoothness constraint on the motion allows us to keep the motion of the hole and the boundary in check as well as the total motion inside the hole itself.

The second constraint is the trajectory smoothness measure. We compute the optical flow among three frames and we get a forward and backward flow which we relate to each other via trajectory similarity measure. The trajectory measure or prior is calculated as the ratio of the forward and backward optical flow which is estimated by averaging many point values under the assumption that the motion is purely translational or the rotational motion is very low.

Another contribution of this work is the iterative optimization framework that allows us to use simultaneously compute the optical flow and propagate the color from known parts of the video into the hole. In order to do this, we modify the mask function used in data term the optical flow estimation function by inferring the distance of the frames of the source pixel and the hole. The source pixel is the known color where the optical flow inside the hole points. This distance increases as the frame of the source pixel moves farther away from the frame of the current inpainted hole. The distance is used to regulate the mask function. In our method, we propose a negative exponential value which decreases in value as the distance increases. Using this modified mask function allows us to use the intensity or color values of the inpainted pixels inside the hole as the optical flow estimation process runs. The error inside the hole is also regulated such that the total error inside is less than the total error outside. Our results show an increased quality of the completed video when the mask function is modified in this way.

Finally, we propose a refinement method on the inpainted motion. By estimating the boundary of objects (or motion) inside the hole, we are able to refine the boundary of the motion and therefore increase the accuracy of the completed video. We estimate the edges of objects inside the hole by using the motion outside the hole and computing an affine transformation. We first estimate the location of the general area of the source pixels and then compute the edges in those pixels. Then, the edge is transferred to the hole via affine transformation. After estimating the edges, we then impose an edge preserving optical flow refinement by using a weighted non-local smoothness regularizer.

The generalized framework in this thesis allows for the improvements in computing time by using faster optical flow estimation techniques and graph-based color propagation methods. The framework is designed to accommodate such changes in the methodology and thus covers a wide variety of applications.