Saliency-based Feature Selection Strategy in Stereoscopic Panoramic Video Generation

Haoyu Wang,1, Daniel J. Sandin2 and Dan Schonfeld3

1,3Electrical and Computer Engineering Department, University of Illinois at Chicago, 2Computer Science Department, University of Illinois at Chicago

Purpose

In this paper, we present one saliency-based feature selection and tracking strategy in the feature-based stereoscopic panoramic video generation system. Many existing stereoscopic video composition approaches aim at producing high-quality panoramas from multiple input cameras [2], [3], [11], [14], [15]; however, most of them directly operate image alignment on those originally detected features without any refinement or optimization. The standard global feature extraction threshold always fails to guarantee stitching correctness of all human interested regions. Thus, based on the originally commonly identified feature set, we incorporate the saliency map into the distribution of control points to remove the redundancy in texture-rich regions and ensure the adequacy of selected features in visual sensitive regions. The experiments show that our method can improve the stitching quality of visual important region without impairment to the human less-interested regions in the generated stereoscopic panoramic video.

Outline

In this paper, the proposed general disparity control strategy is established based on the construction of a commonly identified feature set. Then, we combine the disparity map, gradient map and saliency map into one energy map that indicates the visual importance of each pixel in the image. Given the total number of control points we intends to sample, we select those best-matched commonly-identified features in each grid. In the feature tracking process, we also do the local feature update based on the change of energy in each grid.

Commonly Identified Feature Generation

The stitching of binocular views based on those features from the same edge, corner, or object will maintain the structure consistency between the left and right views. For simplicity, we only consider the general task of stitching two pairs of input rectified stereoscopic images I1, I2, I3, I4 and I5. We define four randomly chosen feature descriptors (e.g., SIFT or SURF) from the four images as d1, d2, d3 and d4 respectively. Each descriptor contains one vector d1, d2, d3 and d4 in multiple directions, and two scales dx and dy for the center position point. The score to evaluate the correspondence between them is defined as follows:

\[ R(d_1, d_2, d_3, d_4) = \sum_{i,j=1}^{n} |d_i - d_j| + \alpha |d_i - dy, y_j| + |d_i - dx, x_j| \]

where \( \alpha \) is the focal length and b is the baseline.

Thus, the construction of the commonly identified feature set could be decomposed into multiple optimization problems for each extracted feature descriptor:

Local Feature Update Strategy

Step1: Grid classification to determine which grid need feature update

Step2: Run KLT tracking in adjacent camera views simultaneously

Step3: Filter the tracking result with proposed depth-related constraints

Step4: Fill those missing control point slot with retrieved position

Fig. 1: Constructed commonly-identified feature set. The control points only connected by red line in vertical or horizontal direction will be rejected. Only the control points connected by yellow line in both of vertical or horizontal direction will be selected for Commonly-identified feature set.

Fig. 2: Images from left to the Right are: (a) Original images; (b) Depth map; (c) Gradient map; (d) Saliency map; (e) Energy map; (f) Grid map; (g) Grid map with assigned visual weight.

Fig. 3: Matched control point pair before and after feature selection.

Fig. 4: Images from Left to the Right are: (a) Grid map between consecutive frames (b) Grid-level energy change map (c) Indicator map.

Fig. 5: Update result comparison between different strategies: (a) original detected features; (b) purely detected; (c) purely tracking; (d) proposed strategy.

Fig. 6: Comparison of stereoscopic panorama stitching result

Fig. 7: Comparison of stereoscopic panoramic video stitching result

Table 1: Numerical Comparison of panorama case

Table 2: Numerical Comparison of video case

References

In this paper, we presented a feature selection and tracking strategy that optimizes the distribution of control points in panoramic video generation system according to the saliency change.

Conclusion

This work was based on or partly inspired by the “Stereoscopic Vision toward 3D Image Processing” funded by UIC 1502040G-01.

Any opinion, finding and recommendation expressed in this publication are those of the authors and do not necessarily reflect the views of the funders.

Acknowledgement

This paper is based on work supported in part by the National Science Foundation award CNS-1145663.

References


