Automating the Construction of Dynamic and Multi-Resolution 360° Panorama for Natural Scenes with Moving Objects

Zhigang Zhu[1,2], Guangyou Xu[1], Heng Luo[1], Qiang Wang[1]
[1]Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China
[2]Computer Vision Lab, Computer Science Department, University of Massachusetts, Amherst, MA 01003
(email: zhuzhg@cs.umass.edu)

Abstract

A new approach is presented to automatically build a dynamic and multi-resolution 360° panorama (DMP) from image sequences taken by a hand-held camera. A multi-resolution representation is built for the more interesting areas by means of camera zooming. The dynamic objects in the scene can be detected and represented separately. The DMP construction method is fast, robust and automatic, achieving 1 Hz in a 266 MHz PC.

1. Introduction

A panoramic representation [1-3] of image sequences has a wide application scope, including virtual reality, interactive 2D/3D video, content-based video compression and manipulation, and full-view video surveillance. For virtual reality, it has advantages of simplicity for rendering, photographic quality realism, and 3D illusion experienced by users. For video analysis and coding, it is superior to existing coding approaches in that it is a content-based representation with a very low bit-rate. In the sense of advanced human-computer interaction (HCI), these two categories will merge into a more general approach of interactive video, which adds the flexibility of synthesizing images with interactivity, selectivity, and enhanced field of view and resolution.

2. Motivations

We aim at the generation of realistic 2D/3D panoramas from video sequences with more general motion of a hand-held video camera. The construction of layered 3D panorama from a vibrating translating camera has been reported in [3]. In this work a new approach is proposed to automatically build a dynamic and multi-resolution 360° panorama (DMP), with good image quality, from a video sequence taken by a hand-held camera undergoing 3D rotation, zooming, and small translations. An experimental system has been built and can be easily used by the non-expert.

3. Algorithms

An inter-frame motion model is derived and a pyramid-based motion detection algorithm is used. Although a simple (and stable) rigid motion model is used to estimate inter-frame motion parameters, our mosaicing methodology enables fine mosaicing results.

Theoretically it would be elegant if one construct a cylindrical panorama after the focal length and the three rotation angles have been decomposed. However, experimental analysis has shown that this decomposition is very sensitive to image noise and the accuracy of the motion parameters. Since the motion in our experimental environment is not a pure rotation, which make this difficult problem even harder, we adopt an alternative approach when panning of the camera is the dominant motion and the pan covers more than 360° around the viewpoint. The algorithm consists of motion parameter estimation and refinement as well as automatic head-tail stitching of the full-view panorama.

As the mosaic is being constructed, the difference image between successive frames is analyzed. In order to rapidly detect and separate the dynamic and deformable objects from the scene, both motion and shape information is utilized in our active contour method. The dynamic sub-images of objects are coded separately.

In order to enhance the visual realism, we use a multi-resolution representation for each "interesting" portion of the panorama. The representation is constructed by actually zooming the camera when the more interesting regions of the scene are viewed. A automatic registration between two zoomed frames is achieved in a manner similar to that for the panned frames, but the step followed is to select the representative frames as the components of a multi-resolution representation (instead of mosaicing the frames).

4. Conclusion

The construction of the DMP is fast, robust and automatic. The processing rate is about 1 Hz for 384×288 color images using a Pentium II/ 266 MHz PC. A factor of 2 speed-up can be expected by algorithm optimization and MMX utilization.

References

[1]. S. K. Nayar, Omnidirectional video camera. *Prof. DARPA Image Understanding Workshop*, May 1997:235-241

[2]. M. Irani, P. Anandan, S. Hsu, Mosaic based representation of video sequence and their applications, *IEEE Proc ICCV'95*, pp605-611.

[3]. Z. Zhu , G. Xu, X. Lin, Constructing 3D natural scene from video sequences with vibrated motions, *VRAIS'*98:105–112.