Image and Vision Computing 37 (2015) 20–30

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Stabilization of panoramic videos from mobile multi-camera platforms ☆

Ameer Hamza [a], Rehan Hafiz [a,*], Muhammad M. Khan [a], Yongju Cho [b], Jihun Cha [b]

[a] VISpro Lab, School of Electrical Engineering and Computer Science (SEECS), National University of Sciences and Technology, (NUST), Islamabad, Pakistan
[b] Electronics & Telecommunications Research Institute, South Korea

A R T I C L E I N F O

Article history:
Received 12 July 2013
Received in revised form 7 February 2015
Accepted 9 February 2015
Available online 30 March 2015

Keywords:
Panoramic Videos
Video Stabilization
Video Stitching
Mobile multi-camera platforms

A B S T R A C T

Wide field of view panoramic videos have recently become popular due to the availability of high resolution displays. These panoramic videos are generated by stitching video frames captured from a panoramic video acquisition system, typically comprising of multiple video cameras arranged on a static or mobile platform. A mobile panoramic video acquisition system may suffer from global mechanical vibrations as well as independent inter-camera vibrations resulting in a jittery panoramic video. While existing stabilization schemes generally tackle single-camera vibrations, they do not account for these inter-camera vibrations. In this paper, we propose a video stabilization technique for multi-camera panoramic videos under the consideration that independent jitter may be exhibited by content of each camera. The proposed method comprises of three steps; the first step removes the global jitter in the video by estimating collective motion and subsequently removing the high frequency component from it. The second step removes the independent i.e. local jitter of each camera by estimating motion of each camera content separately. Pixels that are located in the overlapping regions of panoramic video are contributed by neighboring cameras, therefore, the estimated camera motion for these pixels is weighted using the blend masks generated by the stitching process. The final step applies local geometric warping to the stitched frames and removes any residual jitter induced due to parallax. Experimental results prove that proposed scheme performs better than existing panoramic stabilization schemes.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Multimedia technology has seen a rapid evolution in recent years in terms of both quality and quantity of the information delivered through multimedia displays. Wide Field of View (FOV) panoramic images and videos provide ultra-high definition content that when displayed on large high definition displays, provide an immersive experience to the viewers. Panoramic videos are generated either by using fish-eye lens or by stitching together synchronized video frames coming from multiple cameras arranged on a rig [1,2]. The use of stitching-based panorama generation and display systems once limited to geological surveys and surveillance applications [3] has already been extended for entertainment purposes [4–6]. With decreasing cost of commodity cameras, this trend is expected to be more ubiquitous [7] in the entertainment industry, closely followed by household consumer market in the future. Thus, instead of static panorama acquisition systems [3] that arrange closely coupled sensors in a dedicated closed unit [2], affordable panoramic acquisition systems are emerging that employ commodity cameras mounted on a platform. When set on a mobile platform, such a panoramic video acquisition system may suffer from mechanical vibrations that are global to the camera rig or independent to a particular constituent camera. The global vibrations appear as global jitter in the panoramic video frames while the inter-camera vibrations result in jitter in the spatial region of panoramic video contributed by the particular affected cameras. This results in an unpleasant experience for viewers. This is illustrated conceptually in Fig. 1 which shows a single frame of a panoramic video ('*Man video*') that was formed by stitching frames acquired using three cameras. Please refer to the *accompanying video* to watch this panorama sequence. The tail of the arrows in Fig. 1 represents the location of a few salient features tracked using Kanade–Lucas–Tomasi (KLT) feature tracker [8,9]. The yellow, red and green vectors represent the direction of motion of these features at a particular instance in time and are color coded to signify the independent direction of motion for the contents captured from camera 1, camera 2 and camera 3 respectively. Fig. 1 illustrates an example where camera 3 is experiencing a different vibration and hence its motion vectors (green) appear to have a different direction as compared to camera 1 and camera 2. In the panoramic video (Man video), this effect appears as jitter in the region of panoramic video frames that is contributed by camera 3. Since this jitter appears in a sub-section of panoramic frame, we term it as **sub-frame jitter**. The 'Man video' demonstrates this effect for an actual setup and thus supports our motivation for the requirement of a stabilizing scheme particularly aimed at panoramic videos.

The problem of single camera video stabilization has been extensively researched and has reached a certain level of maturity [10–14]. However, for stabilization of panoramic videos captured from multi-

☆ This paper has been recommended for acceptance by Michael Goesele.

* Corresponding author. Tel.: +92 51 90852107; fax: +92 51 8317363.
E-mail address: rehanhafiz@gmail.com (R. Hafiz).

http://dx.doi.org/10.1016/j.imavis.2015.02.002
0262-8856/© 2015 Elsevier B.V. All rights reserved.

Fig. 1. Stitched frame from a three camera panoramic video. The base of the arrows corresponds to a few salient features in the scene. The yellow, red and green lines illustrate the motion of the tracked features extracted from camera 1, camera 2 and camera 3 respectively at a particular time instance. The figure illustrates that content captured from different cameras may exhibit different motion. This motion is because of the independent motion of each camera and not because of moving objects.

camera platforms, little work has been reported. Furthermore, the existing schemes for such systems [15,16] do not account for the inter-camera vibrations experienced in mobile platforms. In this paper, we treat stabilization of panoramic multi-camera videos as a distinct problem from that of a single camera video. We achieve this by classifying the effects of these vibrations as global, sub-frame and local jitter, and by proposing a method to deal with each one of these in a systematic fashion. In summary, global stabilization is achieved by estimating 2D motion models using the tracked feature trajectories over the complete panoramic frame. Sub-frame stabilization is achieved by making use of the information available in the blend masks that are generated by the stitching application [1]. Blend masks are the intensity weights that are used to blend together the images acquired from multiple cameras to generate a seamless panorama. Finally, the local stabilization tackles the residual jitter that might appear in parts of the panoramic scene in the video due to differences in the scene depth. To the best of our knowledge, this is the first scheme that acknowledges that panoramic videos need to be stabilized temporally as a sequence of stitched frames, as well as spatially to handle the sub-frame jitter.

This paper is organized as follows. In Section 2, we provide a brief account of predominant approaches for video stabilization. In Section 3, we provide the necessary background for panoramic video stitching process followed by a description of the proposed stabilization scheme. In Section 4, we discuss the comparison of the results of the proposed scheme with that of two recent stabilization schemes [14,15] for a number of videos. Finally in Section 5, we present our conclusion.

2. Literature review

Single camera video stabilization problem has been thoroughly addressed in previous works over the last decade. Most video stabilization methods comprise of three main steps: Estimating a 2D motion model between subsequent video frames, computing a smooth motion model that removes the unpleasant jitter and finally applying stabilizing geometric transforms to the video frames [11,12,14,17–29]. These methods differ in their approach of 2D motion model estimation which mostly relies on tracking the position of a certain number of features among adjacent video frames or correlating similar regions in subsequent video frames to estimate the motion vectors. Kanade–Lucas–Tomasi tracker (KLT) is one such popular method for tracking salient features in a video sequence [8,9,30]. The information from the tracking is then used to estimate a 2D motion model for the video. Such 2D motion models depend on the behavior of majority content in the scene and hence fail to accurately model scenes with large moving objects or significant parallax. For most single camera videos, this limitation does not pose a serious problem but for wide angle panoramic videos, there are almost always significant variations in the scene in terms of scene depth and moving content. Another class of stabilization algorithms are based on constructing a 3D [10] model of the scene and camera motion using Structure-From-Motion (SFM) method [31]. Such methods then perform video reconstruction on an estimated low frequency 3D camera path. A relatively recently proposed subspace video stabilization approach [13] tracks feature point trajectories and restricts them to a low-dimensional subspace. Feature trajectories corresponding to moving objects are discarded and corresponding smooth trajectories are computed for the remaining feature trajectories. Instead of estimating global geometric transforms as in the previous schemes, Liu et al. use content preserving warping [32]. Both original and smooth feature trajectories are used as handles in order to cater for the local nature of parallax induced jitter. Liu et al. demonstrated greater robustness than SFM-based schemes while handling parallax more effectively than 2D video stabilization.

Certain schemes for stabilization of multi-camera camera systems have also been proposed. In [15] Haq et al. propose an efficient stabilization system for panoramic video acquisition. It extracts the vibration parameters from the output of one of the cameras and extends it to the output of other cameras based on the assumption that all the cameras have the same optical center and the cameras are mutually static. However, as evident from Fig. 1 this may not be true for a number of cases. In [16] Kamali et al. discuss a scheme for stabilization of spherical videos captured by the Ladybug omnidirectional image acquisition system [2]. It makes use of SFM for modeling spherical scene content. Parameters of camera motion model are then smoothed separately and a new smooth camera path is computed. Stabilized frames are rendered by using a technique similar to content preserving warps [32]. This method however, treats each panoramic frame of the video as a single video frame and assumes no individual motion exhibited by each camera. For a large panoramic video acquisition platform (Fig. 2), enforcing such a constraint may not be feasible. Nevertheless, for a wide angle, high resolution panoramic video, computing a 3D model shall involve complex modeling of multiple cameras for SFM and shall be a difficult task to accomplish. To summarize, neither the single camera stabilization schemes nor the ones that are aimed at panoramic videos particularly account for the effects caused due to inter-camera vibrations.

3. Proposed methodology

In this section we present a brief overview of panoramic video generation followed by a detailed description of the proposed video stabilization technique.

3.1. Panoramic video generation

A panoramic video acquisition system consists of a number of cameras mounted on a platform in such a way that the collective viewing angle of the platform is larger than that of a single camera. These cameras are placed and directed in such a way that the view of each camera partially overlaps with the view of at least one other camera. To minimize parallax, an attempt is made to align the optical centers of the cameras either by placing them carefully in a circular array or by using a conical reflecting mirror rig [6]. Furthermore, the overlapping region may also be minimized to limit the visual effects of parallax to a

Download English Version:

https://daneshyari.com/en/article/526853

Download Persian Version:

https://daneshyari.com/article/526853

Daneshyari.com