Engineering Applications of Artificial Intelligence 54 (2016) 39–48

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

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

Improved global motion estimation via motion vector clustering for video stabilization

Bo-Hao Chen [a,*], Andrey Kopylov [b], Shih-Chia Huang [c,*], Oleg Seredin [b], Roman Karpov [b], Sy-Yen Kuo [d], K. Robert Lai [a], Tan-Hsu Tan [e], Munkhjargal Gochoo [f], Damdinsuren Bayanduuren [f], Cihun-Siyong Gong [g,i,*], Patrick C.K. Hung [h]

[a] *Department of Computer Science and Engineering, Yuan Ze University, Taoyuan, Taiwan*
[b] *Institute of Applied Mathematics and Computer Science, Tula State University, Tula, Russia*
[c] *Department of Electronic Engineering, National Taipei University of Technology, Taipei, Taiwan*
[d] *Department of Electrical Engineering, National Taiwan University, Taipei, Taiwan*
[e] *Department of Electrical Engineering, National Taipei University of Technology, Taipei, Taiwan*
[f] *School of Information and Communications Technology, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia*
[g] *Department of Electrical Engineering, Chang Gung University, Taoyuan, Taiwan*
[h] *Faculty of Business and Information Technology, University of Ontario Institute of Technology, Oshawa, Ontario, Canada*
[i] *Chang Gung Memorial Hospital, Linkou, Taiwan*

ARTICLE INFO

Article history:
Received 29 September 2015
Received in revised form
7 May 2016
Accepted 9 May 2016
Available online 24 May 2016

Keywords:
Video stabilization
Motion vector clustering
Shortest spanning path

ABSTRACT

Video stabilization technique is often used in handheld multimedia devices, whereas the difficulties in the accurate extraction aspect of global motion vectors restrict its development. This paper proposes a novel video stabilization approach that is based on the shortest spanning path clustering algorithm for effective and reliable estimation of the global motion vectors. As demonstrated in our experimental results, the proposed approach achieves superior stabilized effectiveness compared with the other state-of-the-art approaches based on both qualitative and quantitative measurements.

1. Introduction

Several types of handheld multimedia devices have been dramatically developed for the last several decades, such as mobile phones, tablets, and so on (Leu et al., 2012; Kherallah et al., 2009). These devices allow observer to acquire videos from anywhere. However, the video acquisition through the handheld multimedia devices usually suffers from annoying perturbations (e.g. unexpected image motion Pandian et al., 2013) caused by the observer's hand shaking.

In response, video stabilization techniques have played an essential role in handheld multimedia devices. The task of video stabilization techniques is to compensate the unwanted image motion and eliminate these annoying perturbations from video streams (Niskanen et al., 2006; Qu et al., 2013). Numerous video stabilization techniques have been proposed by which to improve video quality in the devices. In general, they attain video stabilization through conjunctive use of camera motion estimation, motion filtering, and motion compensation (Puglisi and Battiato, 2011). In particular, the camera motion estimation is the first essential process in the development of video stabilization techniques by which to provide stable effects of both motion filtering and motion compensation. Specifically, the video stabilization techniques are built on a key observation that the affine transform of the frames is caused by camera motion. According to this observation, the stable frames can be acquired by inverting the global affine transformation. Hence, these techniques can be divided into two major categories according to their capability to estimate the camera motion, that are intensity-based approaches (Puglisi and Battiato, 2011; Kwon et al., 2006; Chang et al., 2004; Ko et al., 1999) and feature-based approaches (Battiato et al., 2007; Yang et al., 2009; Litvin et al., 2003; Shen et al., 2009; Kang et al., 2012; Pinto and Anurenjan, 2011; Feng et al., 2013).

Video stabilization techniques belonging to the intensity-based approaches category directly employ the image textures as motion vectors in each frame of a video to estimate the global affine transform and then reconstruct the stable frames. For instance, Puglisi and Battiato (2011) employed a block-matching algorithm to collect numbers of motion vectors while combining with a voting strategy for detection of global motion vector from different spatial locations of a frame. In addition, the video

* Corresponding authors.
E-mail addresses: bhchen@saturn.yzu.edu.tw (B.-H. Chen), schuang@ntut.edu.tw (S.-C. Huang), alexgong@mail.cgu.edu.tw (C.-S. Gong).

http://dx.doi.org/10.1016/j.engappai.2016.05.004
0952-1976/

Fig. 1. FREAK descriptors from frame 361 and frame 362 in video sequence "Statue". The lines across the symbol ○ and symbol + represent the connections of matched features. The solid and dotted black circles represent the incorrect-matched and inconsistent feature points, respectively.

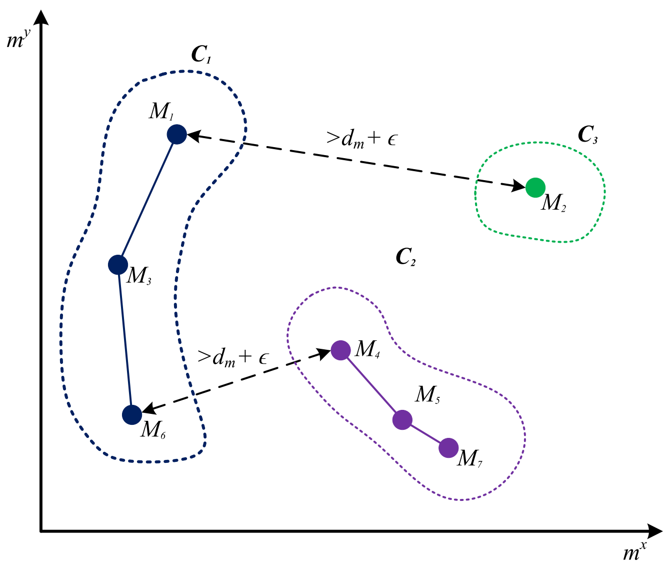

Fig. 2. Illustration of production process for each cluster.

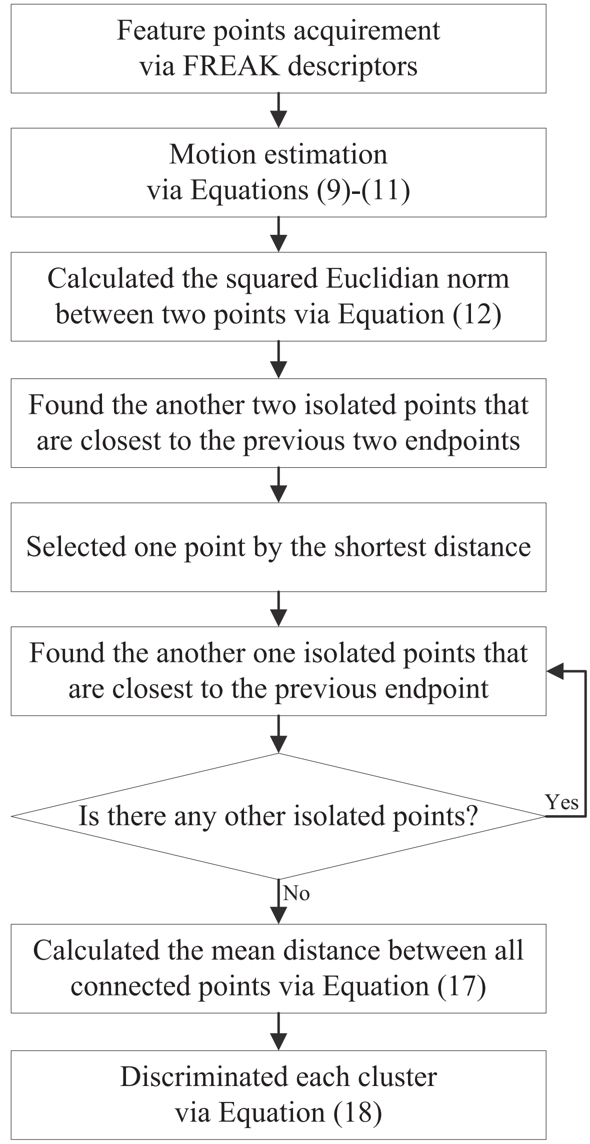

Fig. 3. Flowchart of proposed shortest spanning path based motion vector clustering.

stabilization technique of Kwon et al. (2006) was proposed, in which several motion vectors were estimated first by using the phase correlation-based motion estimation via four rectangular edge sub-images, after which the Kalman filter was used to extract the global motion vectors from those motion vectors. Chang et al. (2004) calculated the optical flows as global motion vectors based on brightness constancy assumption between adjacent frames and the camera motion was then estimated by fitting the optical flow field to a global affine motion model for stabilizing videos. Moreover, the video stabilization technique of Ko et al. (1999) obtained global motion vectors by using gray-coded bit-plane matching from those motion vectors realized by exploiting binary Boolean functions.

On the other hand, feature-based video stabilization approaches locate a sparse set of reliable features in adjacent frames for camera motion estimation. These features can be obtained from Scale Invariant Feature Transform (SIFT) (Lowe, 1999), Speeded Up Robust Features (SURF) (Bay et al., 2008), Kanade–Lucas–Tomasi feature tracker (KLT) (Shi and Tomasi, 1994), Fast Retina Key-point (FREAK) descriptors (Alahi et al., 2012), and so on. Hence, the global motion vectors can be estimated from these features by which to remove the unwanted image motion and further stabilize the videos. In the consideration of computational burden, the most widely adopted method for video stabilization is based on feature-based approaches (Matsushita et al., 2006). Hence, several methods based on the feature-based strategies have been popularly developed and implemented by many video stabilization applications.

The SIFT features are extracted by the method of Battiato et al. (2007); the motion vector integration technique is then utilized to filter the motion vectors produced from these SIFT features as global motion vectors for estimating camera movements. Moreover, Yang et al. (2009) also employed SIFT features to attain the motion vectors, whereupon the global camera motion was estimated by using the particle filters between successive frames. The method of Litvin et al. (2003) minimized the *p*-norm cost function to find the global motion vectors from the extracted feature points. In addition, Shen et al. (2009) proposed a feature-based video stabilization method that extracts features by using both the principal component analysis (PCA) and SIFT, after which the global motion vectors were detected by exploiting the RANdom SAmples Consensus (RANSAC) technique among the motion