



Digital video stabilization based on multilayer gray projection

Fangyu Hu^{a,*}, Jie Ma^a, Liang Shen^a, Huajun Du^b

^a School of Automation, Huazhong University of Science and Technology, Wuhan 430074, China

^b Beijing Aerospace Automatic Control Institute, National Key Laboratory of Science and Technology on Aerospace Intelligent Control, Beijing 100854, China



ARTICLE INFO

Keywords:

Video stabilization
Motion estimation
Differential gray projection
Ring-projection
Circular projection

ABSTRACT

Video stabilization is an important video optimization technology which is aimed at removing unconscious shakiness from videos. In this paper, we present a multilayer gray projection method to estimate image motion. While classical gray projection works only on the videos which just have translation. To achieve this, a multilayer gray projection algorithm is proposed to estimate translation, rotation and scaling motion between target and reference images. First, differential gray projection is applied to the test and target images to estimate relative image translation, the best projection results can help us find corresponding projection center of ring-projection and circular projection in the next two steps. We then calculate the scaling and rotation between test and target images by doing ring-projection and circular projection. The performance of our method is tested by robust experiments, estimation accuracy analysis and experiments over several videos.

1. Introduction

Video stabilization has been widely used in military and civil field since it was proposed in the 1980s. In the process of video information acquisition, image sequence will become unsteady due to the irregular shakiness of camera. This kind of unsteadiness will lead to the fatigue of observer. And it may reduce the precision greatly if the video is used for measuring or observation. Electronic video stabilization technology can reduce and sometimes even eliminate the unsteadiness of videos caused by the shakiness of camera. It will increase the quality of videos, which can decrease the fatigue of observer and increase the accuracy of some application.

In general, digital video stabilization can be decomposed into the following three steps: motion estimation, motion smoothing and motion compensation.

As the first step of digital video stabilization, motion estimation is used to estimate the relative image motion among adjacent images in a video. This kind of motion is also called global motion and the accuracy of global motion estimation is a very important index in the evaluation of digital video stabilization.

In the past few decades, many categories were proposed to accomplish motion estimation. They can be divided into two kinds of approaches: feature-based approaches and intensity-based approaches. Feature-based approaches usually calculate transformation between adjacent frames with distinct features like points [1,2], edges [3] or interested regions [4–6] in an image. These features can be captured

with Scale Invariant Feature Transform (SIFT) [7], Speeded Up Feature Transform (SURF) [8], Kanade–Lucas–Tomasi feature tracker (KLT) [9], Fast Retina Key-point descriptor (FREAK) [10], Binary Robust invariant scalable keypoints (BRISK) [11] and so on.

In 1988, Harris and Stephens proposed Harris feature [12] which has the advantage of high feature detection rate and strong noise resistance. But Harris features are not scale invariant, which can lead to great error if Harris feature detection is used in motion estimation when there exists scaling between adjacent frames. The scaling problem in motion estimation was not solved until the SIFT [13] feature was proposed. The SIFT feature descriptors are invariant to scaling, orientation, illumination changes, and partially invariant to affine distortion [7]. Battiato et al. [14] presented a video stabilization algorithm which accomplishes video stabilization by extracting and tracking SIFT features between adjacent frames in shaky videos. Although the motion estimation based on SIFT feature descriptor has solved the problem of scale change. The 128 dimensional feature description vectors occupy large memory space. The generation and matching of features are both time consuming, which leads to the fact that the electronic image stabilization algorithm based on SIFT feature cannot achieve the real-time requirement. Then an electronic video stabilization algorithm based on SURF [8] feature matching is proposed. The SURF feature descriptor is partly inspired by the SIFT feature descriptor. The standard version of SURF is several times faster than SIFT and more robust against different image transformations than SIFT. Binoy et al. [5] proposed a novel method for robust video stabilization which uses

* Corresponding author.

E-mail address: husthoofy@foxmail.com (F. Hu).

SURF as stable feature points to be tracked between frames for global motion estimation. Although the SURF feature matching is 3 times faster than the SIFT feature matching [5]. It is still difficult for the SURF feature matching to meet the real-time requirements of electronic image stabilization. In 2006, E Rosten and T Drummond developed a new corner detection method called Features from accelerated segment test (FAST) [15]. The greatest advantage of the FAST corner detector is its computational efficiency. In 2010, Calonder et al. proposed to use binary strings as an efficient feature point descriptor called BRIEF [16], which made it possible to achieve real-time electronic image stabilization. After that, Rublee et al. proposed a very fast binary descriptor called ORB [17], which is based on the FAST keypoint detector and the visual descriptor BRIEF and proved to be rotation invariant and resistant to noise. It is aimed at providing a fast and efficient alternative to SIFT. In recent years, algorithms aiming at optimize motion smoothing are proposed and proved to perform well in motion smoothing [18–21].

Feature-based methods are generally more precise than intensity-based methods. But there still exist some shortcomings of feature-based algorithms. Feature-based methods are more prone to local effects. The precision of feature-based algorithms are high relied on the quality of the videos. Poor textural conditions like noise, blur, low intensity and existence of non-texture regions may lead to estimation failure as a consequence of limited number of distinct or reliable features [22]. Intensity-based approaches do not rely on features like points or edges. These methods usually estimate global motion by taking the gray level of whole image either in blocks or in full-area into consideration. They are better alternatives when there exists conditions which can reduce the quality of image like noise, blur and so on.

Intensity-based approaches can be divided into Represent Point Matching (RPM) [23], Block Matching (BM) [24], Gray Projection (GP) [25], Gray Encoded Bit Plane Matching (GEBPM) [26]. Among all intensity-based approaches, the most frequently used methods are BM and GP for they are easier to implement on computer and they tend to be less affected by local image variations and give better performance under challenging scene capture conditions.

BM methods choose several blocks in reference frame and assume that all pixels in the same block have the same motion. A block matching algorithm involves dividing the target frame into several blocks and comparing each of the blocks with a corresponding block in its adjacent neighbors in the reference frame. A vector is created that models the movement of a block. This movement, calculated for all the blocks comprising a frame, constitutes the motion estimated in a frame. Accuracy and real-time of block matching techniques is affected by the size of blocks, the search range and the search strategy. Xu et al. [27] proposed a circular blocking method to estimate the global scaling, rotational and translational motion parameters between adjacent frames in shaky videos. The circular block matching has solve the problem that rectangular block motion search could fail when rotation motion exceeds a certain small range. Battiato et al. [28] have address a robust blocking matching method in which motion vectors are filtered by pre-filtering module through some simple rejection rules based on the goodness of matching to eliminate local motion vectors. G Puglisi et al. [29] proposed a fast and accurate block based local motion estimator together with a global motion estimating algorithm based on voting.

GP methods estimate motion between adjacent frames by cumulating pixel intensities along a fixed angular direction in both frames. Instead of working on parts of the frames, GP methods often work on the whole image, which makes GP methods less affected by local scene variations and perform better under poor textual conditions. AJ Crawford et al. [30] presented a new definition of the relationship between integral projections and motion in an image pair. The resulting new multi resolution gradient based approach is used to estimate dominant motion in image sequences degraded by random shake. Li Shuang et al. [31] proposed a novel video stabilization algorithm which combines gray-scale projection and representative point matching in motion estimation. The algorithm has been proved to be more accurate

than traditional gray scale projection algorithm by experiments. M Veldandi et al. [32] estimate translation, rotation and scale between adjacent frames with improved projection method, experiment results demonstrate the robustness of the algorithm. D Shukla et al. [33] proposed a new video stabilization method based on differential-radon projection which can estimate rotation, translation and scaling. The translation and scaling in horizontal and vertical direction can be estimated accurately.

In this paper, the transformation between images is described by similarity transformation. To improve the accuracy, speed and robustness of motion estimation, we proposed a multilayer gray projection method to estimate translation, rotation and scaling between adjacent frames. In the first step, we use adopt gray projection to estimate translation between adjacent frames. It has been proved that differential gray projection has a better precision than ordinary gray projection on translation estimation for the reason that it has reduced the influence of textureless regions like sky and water areas [33]. So we choose to estimate the translation parameter between two frames with differential gray projection and it turns out to improve the precision effectively. In the second step, we estimate the scaling parameter via an improved ring-projection method. After the first step, we can find the point in target frame which is corresponded to the reference frame center. Take these two points as the centers of ring-projection. Then we can estimate the scaling parameter by analyzing the relevance between two projection curves obtained by ring-projection. In the last step, we estimate the rotation parameter via a new circular projection method. After that a similarity transformation based on the estimated motion parameters is generated for inter-frame stabilization.

In the subsequent part of the paper, a brief introduction to ring-projection, circular projection and the method of how we estimate translation, scaling and rotation parameters is presented in Section 2. A real-time video stabilization framework and implementation steps are discussed in Section 3. In Section 4, we presented several experiments to verify the robustness, accuracy and speed of our estimation method. The conclusion is presented in Section 5.

1.1. Key contribution

In this paper, we proposed a motion estimation method based on multilayer gray projection. The proposed algorithm combined traditional gray projection, ring-projection and circular projection to estimate translation, scaling and rotation with high speed and accuracy. Highlights can be concluded as follows:

1. An improved ring-projection algorithm is introduced to estimate scaling between adjacent frames by calculating the variance of their ring-projection curves. The proposed ring-projection method is proved to have high speed, accuracy and robustness with experiment.
2. A novel circular projection method is proposed to estimate rotation between adjacent frames by calculating their circular projection curves' error function of rotation. It also has been proved to have high speed, accuracy and robustness with experiment.

2. Proposed motion estimation method

In the proposed projection based motion stabilization technique, the required motion parameters i.e. translation, rotation and scaling are extracted using differential gray projection, ring-projection and circular projection. This section presents the parameter extraction algorithms for the three motions sequentially.

2.1. Translation estimation with differential gray projection

In order to achieve better accuracy for ring-projection and circular projection, we first align the center of rotation for adjacent frames by

Download English Version:

<https://daneshyari.com/en/article/6941403>

Download Persian Version:

<https://daneshyari.com/article/6941403>

[Daneshyari.com](https://daneshyari.com)