Signal Processing: Image Communication 47 (2016) 115–130

Contents lists available at ScienceDirect

Signal Processing: *Image Communication*

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

Digital image stabilization using similarity transformation over constrained Differential-Radon warping vectors

Deepika Shukla [a,*], Rajib Kumar Jha [b], Aparajita Ojha [a]

[a] PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, Madhya Pradesh, India
[b] Electrical Engineering Department, Indian Institute of Technology Patna, Bihar, India

ARTICLE INFO

Article history:
Received 9 December 2015
Received in revised form
29 April 2016
Accepted 27 May 2016
Available online 2 June 2016

Keywords:
Motion estimation
Integral projection
Affine transformation
Dynamic time warping
Image stabilization
Video stabilization

ABSTRACT

In this paper an image-stabilizer based on novel Differential-Radon warping for combined camera motions, i.e., translation, rotation and zoom estimation has been presented. Motion estimation using intensity integrals in classical Radon transform (RT) does not consider projection shape for its correlation and leads to erroneous results under local intensity variations. As a solution, *Differential-Radon* (DRadon) matrix utilizing inherent curve variance as matching feature is proposed. The scheme provides peaked correlation curves, which gives better motion accuracy in comparison to the classical RT. DRadon projections of the test and target images computed at an intermediate angular-slice resolution of 0.1 degree are correlated to find relative image tilt, and the best matched projections are warped using Sakoe–Chiba constrained dynamic time warping. The estimated warping vectors are then processed using a new local-optima based *interest vectors* selection scheme to remove various one-to-many-matched warping singularities. A linear-motion model is applied over the extracted interest warping vectors and their least-square solution is used as inter-frame translation and zoom motions. Finally, a composite affine motion stabilizer framework is presented for shaky videos taken under different scene-capture conditions. Comparative stability performance evaluation using peak signal-to-noise ratio and mean structural similarity index is presented for various real-world sequences.

1. Introduction

Unsteady camera platforms induce annoying visible jerkiness in the captured image-sequences, which are very prominent in various real-world applications, such as, surveillance, law-enforcement, military and hand-held camera shooting. This diverse range of video acquisition demands efficient stabilization techniques to provide pleasant and steady-view of the recorded footages for further scene analysis. Any image stabilization system aims to estimate inter-frame camera motions and remove the undesired shakiness efficiently while retaining the intentional camera movements. Online counter-motion adjustment of the camera imaging plane, or the offline processing of captured footage are generally applied to compensate the induced unsteadiness effects. Recent cameras are being designed with in-built motion sensors and dedicated processing circuitry [1,2] to provide online image stabilization by efficiently removing the undesired high-frequency system vibrations. Sensor based techniques [3–6] provide real-time stabilization but the high cost and large space requirements limit their usage in low-cost camera design. Digital techniques [7–30] prove to be economically efficient for video stabilization but they mostly work offline, and process the video after it has been recorded. Fast digital techniques suitable for real-time stabilization generally use frame buffers for required offline processing over stored frames.

In the literature, various existing digital stabilization techniques are broadly categorized as Feature based, and image (or area) based approaches. Features based techniques [7–12] work on distinct image-feature like points [7,8], edges [9] or interest regions [10–12] extraction, which are further matched to find correct transformation between reference and target images. Feature based techniques provide affine motion stabilization, but their reliability highly depends on the quality of captured images. Under challenging scene capture conditions, poor textural detailing like noise, blur, low intensity, and homogeneous regions limit the number of extracted distinctive features, and most of the feature based techniques fail due to their erroneous matching [22]. Image based techniques [13–30] do not rely on selective feature points or regions, rather consider whole image either in blocks or in full for relative motion estimation. These intensity based techniques are observed to be less probable to stabilization failure and provide better stabilization solution under challenging scene capture

* Corresponding author.
E-mail address: deepika.shukla@iiitdmj.ac.in (D. Shukla).

http://dx.doi.org/10.1016/j.image.2016.05.016
0923-5965/

conditions. Block matching techniques [13–17] work on the principle of finding best corresponding match for each selected test block, confined within the dedicated search space of target image. The block displacement required to find the best target match is considered as local motion of the respective test block, and median of the estimated local vectors computed over entire image is termed as global or desired image motion. These block matching techniques are mostly limited to translation estimation domain and suffer from large processing time requirement. In literature, optical flow [18] using relative pixel velocity instead of block-correlation is suggested for affine motion estimation. However, local textural variations in the captured scene may lead to inaccurate intensity flow producing mis-warped pixels as their corresponding distorted local patches in the stabilized target image.

Image projection based techniques [19–30] instead of working on local areas, use global intensity integrals as frame representative feature. Projections are less affected by local image variations and give better performance under challenging scene capture conditions. In literature, motion estimation using projection correlation [19,20] and warping [21–24] is mostly investigated for translational domain only. Instead of working over entire image as whole, block projections have also been reported as time efficient alternative to block-matching techniques [25,26]. Extension of projection techniques to rotation motion estimation is recently suggested using Radon transform [27,28] and multiple rotated images [29] for the required angular projection estimation. Best matched angular projections between consecutive images give the relative tilt estimate, and are further processed for translation and zoom motion estimation [29,30]. Rotating an image at different angles for multiple angular-projections overburdens the processing, and thus, in this work, Radon transform is chosen as an efficient tool for multiple-projection extraction.

A new discretized differential-Radon (DRadon) scheme, based on derivate-projections is proposed which by inheriting angular projection shape instead of classical intensity-integrals provides improved curve matching. In real-world applications, the platform unsteadiness induces oscillatory frame deflections within a small range of motions. Inspired from this, in the proposed DRadon intermediate angular projections are computed at a resolution of 0.1 degree under a confined range of $[-10°, 10°]$. DRadon-slice correlation is used for relative image rotation estimation, and the best matched reference and target DRadon angular slices are then matched using Sakoe–Chiba constrained warping [31,32]. One-to-many miswarping singularities present in the extracted warping vectors are removed by selecting some key vectors termed as *interest warping vectors* [33]. Least square solution applied over the linearly modeled interest vectors gives relative image translation and zoom. These estimated rotation, translation and zoom motion parameters are next combined to get the inter-frame affine motion matrix. Finally, an affine video stabilizer framework based on combined DRadon warping and linear least square solution, is developed and the stabilization performance is evaluated over various categories of real-world videos.

Remaining part of the paper is organized as follows. A brief background to Radon transform, its discrete version applicable to image analysis with properties related to the translation, rotation and scale estimation are presented in Section 2. This section also discusses the key highlights on Sakoe–Chiba constrained band algorithm for DRadon curve warping. Proposed DRadon based angular motion algorithm and the interest warping vector extraction with their usage in translation and zoom estimation is presented in Section 3. A composite affine motion stabilizer framework combining the three estimated parameters is presented in Section 4. An algorithm for the proposed composite motion estimation and frame stabilization is also presented in this Section. Experimental results with motion estimation accuracy and stabilization performance analysis are given in Section 5. Finally, a concluding remark on the proposed stabilizer framework is presented in Section 6.

1.1. Key contribution

In this paper a new constrained differential-Radon estimation incorporating projection shape as the matching feature is proposed for improved motion accuracy. In the literature, multi-projections of Radon transform [27,28] or rotated-images [29] are matched for inter-frame rotation motion estimation. These approaches estimate the consecutive projections at unit degree angular resolution, which in turn restrict the angular motion accuracy at its crude-level. However, in image stabilization application the inter-frame tilt may lie in even more smaller range and as per human visual perceptibility of relative tilt, 0.1° is observed to be a noticeable change in the image content. Inspired from this, in proposed DRadon approach the consecutive projections are estimated at 0.1 degree of incremental angular change. Veldandi et al. [29] have suggested application of linear least square solution of projection warping vectors for combined translational and zoom motion estimation, but presence of one-to-many mismatched vectors affects motion accuracy severely. In the proposed scheme, this problem has been overcome by selecting only local optima-matched interest warping vectors out of the complete set of DRadon slice warping vectors and the linear least square analysis is applied over these interest vectors. Highlights of this research contribution can be summarized as follows.

A new differential-Radon (DRadon) inheriting the projection shape as key matching feature is proposed for improved and peaked correlation curve.

For improved rotational motion accuracy, consecutive angular projections for the constrained DRadon are estimated at a resolution of 0.1 degree instead of unit integers.

Interest warping vectors based on local optima matching are extracted to eliminate the mismatching singularities in the warped DRadon-slices.

Least square solution of the linear motion model applied over interest warping vectors results in improved translation and zoom motion accuracy within the extended range as $[-25, 25]$ and $[0.85, 1.15]$ respectively.

2. Background

The proposed image stabilization technique uses Differential-Radon (DRadon) of test and target images to find their best correlated angular position as the desired rotation motion. Best matched DRadon projections are further warped using constrained dynamic time warping and the estimated warping vectors are processed for relative translational and zoom motion estimation. This section presents a brief introduction of continuous and discrete Radon transform with its useful properties related to inter-frame motions. A background to Sakoe–Chiba constrained warping is also discussed for self-integrity of the proposed DRadon-warping approach.

2.1. Radon transform

Radon transform has its wide applications in field of medical, optics and seismic data acquisition and processing. Radon transform is widely recognized in tomography image reconstruction [34], where it relates a 2D function $f(x,y)$ to the collection of intensity integrals of that function computed at different angles. The continuous form of Radon transform can be mathematically expressed as