Computer Vision and Image Understanding 000 (2016) 1–12

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

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

A mutual local-ternary-pattern based method for aligning differently exposed images

Shiqian Wu[a,*], Lingxian Yang[b], Wangming Xu[b], Jinghong Zheng[c], Zhengguo Li[c], Zhijun Fang[d]

[a] School of Machinery and Automation, Wuhan University of Science and Technology, 430081, China
[b] School of Information Science and Engineering, Wuhan University of Science and Technology, 430081, China
[c] Department of Signal Processing, Institute for Infocomm Research, 1 Fusionopolis Way, 21-01 Connexis, 138632, Singapore
[d] School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, 333 Long Teng Rd, Shanghai, 201620, China

ARTICLE INFO

Article history:
Received 17 November 2015
Revised 1 May 2016
Accepted 29 July 2016
Available online xxx

Keywords:
Image registration
Local ternary pattern
Multi-exposed images
Image mapping function
Histogram-based matching
Coarse-to-fine technique
HDR imaging

ABSTRACT

Saturation and large intensity variations occurred in multi-exposed images offer great challenges to align these images. In this paper, a mutual local-ternary-pattern (MLTP) is proposed to represent differently exposed images for image registration. Different from the classical local ternary pattern (LTP) and its variants, the proposed MLTP has two salient properties: (1) The ternary pattern of one image is not only determined by itself, but also relied on its counterpart; (2) The MLTP is grayscale-adaptive. It is analyzed that the proposed MLTP is a good representation to preserve consistency of differently exposed images. Based on the MLTP-coded images, an efficient linear model derived from Taylor expansion is presented to estimate motion parameters. To improve accuracy and efficiency, image rotation is initially detected by the histogram-based matching, and coarse-to-fine technique is implemented to cope with possibly large movement. Extensive experiments carried out on a variety of synthesized and real multi-exposed images demonstrate that the proposed method is robust to 10 exposure values (EV), which is superior to other methods and current commercial HDR tools.

1. Introduction

In the current world where storage, bandwidth, processing power and sophisticated devices have provided an excellent convenient infrastructure for people to capture, utilize and consume images, there has been an increasing demand for high quality in the manipulated images as compared to what human eyes can see. As a result, super resolution, extended depth of field, high dynamic range (HDR) imaging and other image enhancements by photographic multi-shot techniques have been used in many applications (Juergen and Rainer, 2009). As the input of these technologies is a sequence of images, image registration is crucial to obtain high-quality synthesized output. Currently, combining differently exposed low dynamic range (LDR) images of the same scene has been the most popular approach to generate an HDR image. This paper focuses on aligning multi-exposed images captured by hand-held devices for HDR imaging, which is still a challenging issue in image registration (Oldridge et al., 2011).

1.1. Alignment of multi-exposed images – the state of the arts

Image registration or alignment has been a fundamental problem in image processing and computer vision. There are a large number of techniques proposed for a variety of applications, for example, spatial variations, intensity variations, sensor variations and so on. The papers (Oldridge et al., 2011; Szeliski, 2006) provide a comprehensive survey on image registration. Generally, the registration methods can be classified into two categories (Szeliski, 2006): pixel-based (area-based) methods and feature-based methods. Pixel-based methods find motion parameters through minimizing pixel-to-pixel dissimilarities. Accordingly, such methods are dependent on image intensities. Feature-based methods, on the other hand, first extract distinctive features from each image, e.g., point features (Lowe, 2004), line features (Maintz et al., 1996) etc., then match and warp the features to derive parametric transformations. As feature-based methods do not work directly with image intensities, it is frequently used when illumination (intensity) changes.

So far, feature-based techniques have been adopted to align a set of differently exposed images. In Eden et al. (2006); Tomaszewska and Mantiuk (2007), SIFT (Scale-Invariant Feature

* Corresponding author. Fax: +862768862283.
E-mail address: shiqian.wu@wust.edu.cn (S. Wu).

http://dx.doi.org/10.1016/j.cviu.2016.07.010
1077-3142/

Transform) method was employed to detect feature points (keypoints) in multi-exposed images, and the RANSAC method was used to find best pairs of keypoints and derive motion parameters. In Gevrekci and Gunturk (2007), Gevrekci and Gunturk proposed to detect corners as feature points. To alleviate the influence of intensity variations on extracting feature points, both the algorithms (Gevrekci and Gunturk, 2007; Tomaszewska and Mantiuk, 2007) work in contrast domain, which is further revealed in Gevrekci and Gunturk (2009) that the repeatability rate of the detected features can be improved by about 25%. Meanwhile, pixel-based methods were employed for this task, in which the key idea is to cope with intensity variations. In Cerman and Hlavac (2006); Kang et al. (2003), the multi-exposed images are modeled via camera response function (CRF) proposed by Debevec and Malik (1997). Then, the differently exposed images can be converted to identical exposure via the CRF. Another technique is to perform joint geometric and photometric registration (Aguiar, 2006; Bartoli, 2008; Candocia, 2003; 2005; Hossain and Gunturk, 2011; Luong et al., 2010; Zimmer et al., 2011). In Aguiar (2006); Luong et al. (2010), the intensities of multi-exposed images are modeled as a linear relationship. A two-step iterative algorithm in Aguiar (2006) and total least square in Luong et al. (2010) were proposed to solve geometric registration. In Bartoli (2008), Bartoli modeled photometric mapping as gain and offset terms, and estimated the motion field as well as the photometric mapping parameters. Based on gradient constant assumption, gradient information was used for alignment in Zimmer et al. (2011). In Hossain and Gunturk (2011), two differently exposed images were normalized by the intensity mapping function. In Candocia (2003), Candocia proposed to model multi-exposed images by a nonlinear model which was introduced in Mann (2000), spatial and tonal registrations were then simultaneously performed by Levenberg-Marquardt optimization. The key drawback of this method is that the optimization is very slow as there are many parameters, i.e., $9(q-1)$ parameters, where q is the number of images, to be estimated simultaneously. Moreover, optimization in high dimension is difficult to guarantee global solution. To mitigate the computation burden, an improved solution using piecewise linear comparametric model was proposed in Candocia (2005). A hybrid scheme employing a pixel-based method as well as a feature-based method was proposed in Tico and Pulli (2010), and a two-stage method comprising of image normalization and local-binary-pattern representation was developed in Wu et al. (2014) for registering multi-exposed images. Recently, patch-based methods were prevalent (Hu et al., 2012; 2013; Ramirez et al., 2013; Sen et al., 2012; Zheng and Li, 2015; Zheng et al., 2013) in HDR imaging. The key advantage of these approaches lies in simultaneously dealing with camera movement and object movement to composite ghost-free HDR images. To this end, dense correspondences have to been detected, for example, by PatchMatch method (HaCohen et al., 2011). These patch-based methods are time-consuming, because the computation of dense correspondences and the following joint optimization on image reconstruction is very expensive. Moreover, the quality of HDR images heavily relies on the choice of the reference image because the saturated pixels in the reference image yield big problem in patch matching (Hu et al., 2013; Zheng and Li, 2015; Zheng et al., 2013).

1.2. Issues in current HDR techniques and challenges in exposure-robust alignment

The standard and widely-used method for HDR imaging is to combine differently exposed LDR images of the same scene. The motivation behind this technique is that different exposures capture different dynamic range characteristics of the scene. However, real applications of this technology suffer from two problems (Srikantha and Sidibé, 2012): (1) Misalignment: global motion from hand-held camera results in misaligned images that cause the combined HDR image to look blurry; (2) Ghosting: ghost artifact appears due to dynamic scenes. To well cope with the two problems and generate high-quality HDR images, the choice of reference image is very important (Zheng et al., 2013). To our best knowledge, most of the HDR methods select the middle-exposed image, which is viewed as the best exposed image, as an initial reference, and subsequently processing (aligning and detecting moving objects) the consecutive images. After that, the processed images are served as new references to deal with their consecutive images until all images are done. The key advantage of the progressive method is to alleviate the difference of the underlying images resulting from different exposures. However, the drawback of the progressive solution is the error propagation. For example, two images with large EV interval have large registration error. This implies that the performances of the existing paradigms are limited by the number of images, which affects the quality of final HDR images. Hence, it is desirable to use a fixed reference and develop an exposure-robust method to align the multi-exposed images.

It is highlighted that, two images with large exposure value (EV) interval have significant intensity variations as shown in Fig. 1. Such images pose great challenges for image registration:

(1) A feature (e.g., an indoor feature) detected in one image (the long-exposed image) may not occur in another one (the short-exposed image) as shown in Fig. 1, which results in difficulty to use feature-based methods;
(2) Each image contains severely under/over-exposed regions, where information is loss and few features are detected;
(3) The intensities or intensity gradients are not linearly related with exposure times as supposed in Aguiar (2006); Bartoli (2008); Luong et al. (2010); Zimmer et al. (2011) due to nonlinear CRF property, which yields difficulty to normalize these images for area-based registration;
(4) A specific intensity in one image may map to multiple intensities in the other images, and vice versa. For example, two saturated pixels $Z_1(u)=255$, $Z_1(v)=255$ in a long-exposed image may become $Z_2(u)=255$, $Z_2(v)=240$ in its corresponding short-exposed image. Or one grayscale, say 150 in a short-exposed image, may map to grayscales of 190 or 193 because of noise or quantization error.

Due to the aforementioned issues, two images with large EV interval are usually not correlated. We used Photoshop CS5 to align two images with more than 7EV interval. Results show that Photoshop recognizes them as different scenes and displays “the images intended for alignment should overlap by approximate 40%”.

1.3. Contributions of this paper

To address the issue, i.e., the middle-exposed image is supposed to be the best quality image and selected as reference, then progressive method is employed for image registration, which results in error propagation, the aim of this paper is to provide a robust method so that the alignment performance does not rely on the particular reference image. In other words, even when there are large over/under exposed areas in the selected reference image, and the underlying images have large EV interval, the proposed scheme can still yield high-performance alignment results. The contributions of this paper are summarized as follows:

(1) The invariant representation of multi-exposed images is analyzed, and it is indicated that exposures change the intensities, but keep the relative order of intensities. However, saturation yield inconsistent binary features of multi-exposed images.