J. Vis. Commun. Image R. 36 (2016) 11–27

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Evaluation of color descriptors for projector-camera systems ☆

Aleksandr Setkov *, Michèle Gouiffès, Christian Jacquemin

LIMSI, CNRS, Univ. Paris-Sud, Université Paris-Saclay, Bât 508, Campus Universitaire, F-91405 Orsay, France

ARTICLE INFO

Article history:
Received 28 April 2015
Accepted 6 January 2016
Available online 14 January 2016

Keywords:
Feature descriptors
Feature matching
Color invariance
Geometry compensation
Smart projection
Projector-camera systems
Projection-based augmented reality
Spatial augmented reality

ABSTRACT

Spatial Augmented Reality applications generally use projector-camera systems to control the visual projection appearance by comparing the initial projected and the acquired images. To obtain an accurate geometric compensation, a non-intrusive feature-point matching approach can be exploited which must handle complex photometric distortions due to the spectral devices responses, complex illumination and the mixing of the projected image with the projection surface.

This paper first discusses the invariance properties of existing color descriptors in that application for non-intrusive geometric compensation. Their performance is evaluated using the framework of Setkov et al. (2013) extended by adding the several new test cases: modeled synthetic projections, real-world projections under various illuminants on one and two planar surfaces. Our experimental results show two main conclusions: (1) classical color vision models are hardly suitable to model the distortions in a projector-camera system, and (2) the LHE-based descriptor (Local Histogram Equalization) is the most reliable to compensate real-projections.

1. Introduction

Smart projections generally combine a projector device, and one or several cameras and other proprioceptive devices so as to adapt the projection with respect to the scene conditions such as illumination and geometry. Their main application is Projection-based Augmented Reality (PAR). Cameras acquire corrupted projections and the system compares them with the reference images to estimate and to compensate for the geometric and photometric distortions. The whole process can be split in two parts. The first one is offline-calibration that aims at estimating the relative geometrical correspondence between each projected and captured pixels [2–4]. The second one is online estimation and compensation of the distortions.

According to Bimber et al. [5], the problems that affect the final projection, can be roughly classified as photometric and geometric distortions. The ultimate purpose of all compensation algorithms is to bring the final projection to a such representation that it can be perceived nearly identical to the source image. The photometric compensation is out of the paper scope and, therefore is not described.

Geometric projection correction methods form two groups, *active* and *passive methods*. The methods of the first group project additional patterns that facilitate recovering surface structure and, consequently, compensating geometric distortions. They, for the most part, make use of structured light techniques [6–11]. Although they can address both the static and the dynamic scenarios [9–11,8], they have some drawbacks. Firstly, initial projected content has to be modified to project the light patterns. This can be inappropriate for some applications which require high fidelity of rendering, for example medical PAR applications. Secondly, depending on the coding strategy, a projected pattern either interleaves the image data sequence or is embedded in the original images which can interfere with the human perception. The former requires precise camera-projector synchronization [12], whereas the latter makes it extremely difficult to imperceptibly embed a pattern into the stream using standard commercial cameras [9]. Moreover it puts some constraints on the projected colors and affects the projected frame-rate which can be a restrictive limitation for applications such as teleconferences or video content projections.

More recently, the inversion of light transport matrices [13–15] has gained influence and allows to compensate for radiometric distortion and small superficial roughness, at a high price in terms of memory and time resources.

Another approach is introduced by Zollmann et al. [7] that combines both active and passive reconstruction techniques. The first one, a slow structured light projection, is used as a preliminary step to calibrate the devices off-line. Next, a lighter optical flow method tries to estimate geometric distortions in real-time. If it

---

☆ This paper has been recommended for acceptance by Yehoshua Zeevi.

* Corresponding author. Fax: +33 1 6985 8088.
E-mail address: aleksandr.setkov@limsi.fr (A. Setkov).

http://dx.doi.org/10.1016/j.jvcir.2016.01.006

fails, the structured light method is triggered for recovering a correct geometrical compensation.

Passive methods are only based on the capture and the analysis of the projected image itself. Local features are extracted and matched between the captured and the reference images. Such methods do not require any modification of projected data, however they highly depend on source and target images properties such as the amount of textured patterns. In this case it presents a chicken and egg situation in which geometric and colorimetric compensations are mutually dependent. In other words, geometric compensation generally requires precise correspondences between images but these correspondences are difficult to obtain without prior color compensation. Conversely, in the state of the art, most color compensation methods assume that the geometric distortions have been compensated for.

Since in the standard approach for feature matching, keypoints are extracted only from the grayscale channel, the results may suffer from color distortions present in images. To address this problem, color-invariant descriptors have been recently introduced [16,17]. They are constructed independently from several color channels of the image and then merged into one long descriptor. Preliminarily, some color transforms can be applied to the image in order to produce an invariant representation on the basis of RGB channels.

However, most color invariants have been designed for standard computer vision applications, for which the changing parameters are mainly the camera sensitivity and the illumination spectrum. The color invariance is more difficult to define in the context of video-projector systems, for which the distortions are due simultaneously to the projection and acquisition devices and to the inhomogeneous color and geometry of the scene.

This paper explores the use of local color-invariant descriptors for planar geometric adaptation. The advantage of our approach lies in the fact that geometric compensation can be performed independently or even without photometric compensation. It is achieved when color invariance is defined with respect to a precise model of the physical phenomena involved in the acquisition process.

Therefore, in this work we extend our previous framework [1] which compared several color descriptors with the newly presented Local Histogram Equalization (LHE) method for homography compensation. Besides projective transformations coupled with synthetic color distortions, more scenarios are addressed in this work. We use three test image sets including (1) synthetic images with added Gaussian noise and simulated projections obtained through normal mapping and blended with color backgrounds, (2) real-world projections on a single planar surface under three different illuminations, and (3) projections on two planar surfaces to address a more complex double-homography compensation problem. Test synthetic data was chosen so as to simulate real distortions occurred in the images acquired by a projector-camera system. To that end, we consider two color change models used in computer vision applications.

In the evaluation we consider several state-of-the-art color descriptors. For each type of test set used in the experiments our study shows one or several descriptors that produce the best matching results. By analyzing the results we infer a set of recommendations about which color descriptor is more appropriate for feature-point based homography compensation depending on the types of color distortions in the acquired projection.

The paper is organized as follows. In Section 2 the current state of the art approaches and methods are introduced. Section 3 describes the photometric models that are considered in this work. Then, Section 4 provides a discussion on existing color descriptors, their properties and relevance to the use for geometric compensation. In Section 5 we discuss LHE and analyze its characteristics when used in a projector-camera system. The experimental setup and the chosen evaluation metrics are explained in Section 6. The experimental results obtained in the paper are shown in Section 7. Finally, Sections 8 and 9 are devoted to the discussion and the conclusions, respectively.

2. Related work

Let us now review the state-of-the-art in geometric compensation and color descriptors.

2.1. Geometric compensation

Geometric distortions mostly occur due to non-planar surfaces or planar surfaces that are not an homothetic transformation of the projection surfaces in the scene which make projected images appear geometrically warped to an observer. Some works assume that a prior knowledge of the scene geometry is available [18–20]. This allows applying mapping from the estimated set of points to the expected structure.

Bimber et al. [21] address the problem of unknown dynamic surfaces similarly. The core of their approach relies on off-line geometry measurements for a discrete number of perspectives. During runtime, the correct image warping is obtained by interpolating the measured samples based on a sequence of observer's current perspectives.

As it was described in the introduction, in this paper we are interested in passive, feature matching-based, techniques for geometric compensation. Therefore, some methods appeared in the literature are reviewed below.

Passive methods. The methods of this group [22–24] rely on the content of the projected image itself to estimate the projection transformation. It extracts the most distinctive regions (blobs, lines, corners, SIFT-like points) on the assumption that they do not change drastically under geometric distortions. Yang and Welch [22] developed a method that estimates the deformation iteratively starting from an initial rough estimate. Yamanaka et al. [23] propose a method that performs geometric adaptation to non-planar screens. After preliminary off-line calibration, it estimates the projection surface geometry by matching epipolar lines and by fitting B-spline surfaces.

Unfortunately, none of the described methods can deal with dynamic surfaces in real-time, otherwise, only slight displacements are possible. Another limitation comes from the fact that these methods significantly depend on color consistency between source and target images and hence matching results can suffer when photometric distortions impact the projected images.

Drouin et al. [24] have recently proposed a new approach for passive feature matching in projector-camera systems which is based on spatio-temporal matching of active pixels extracted through background subtraction algorithm. In total, from 15 to 30 frames are required to obtain a full compensation.

We based our work on passive geometric adaptation because it has a number of advantages compared with the active methods. It does not require a projection of several images along with the main content which in turn is kept unmodified. Through this approach, it is easier to achieve real-time performance that most PAR applications require. The focus of our work is to ensure robust feature matching and accurate homography estimation of various projected and captured images under different types of color distortions due to illumination conditions, background, and camera settings. It is achieved by means of local color invariant descriptors. This approach is described in the next subsection.

2.2. Color invariant descriptors

To match a reference image with its projections, in this paper we exploit the local feature matching approach [25,26]. It has been