Image and Vision Computing 32 (2014) 606–615

Contents lists available at ScienceDirect

Image and Vision Computing

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

Accelerometer-based correction of skewed horizon and keystone distortion in digital photography ☆

Enrico Calore [a], Iuri Frosio [a,b,]*

[a] Computer Science Department, University of Milan, Via Comelico 39/41, 20135 Milan, Italy
[b] NVIDIA Research, 2701 San Tomas Expressway, 95050 Santa Clara, CA, USA

ARTICLE INFO

Article history:
Received 19 December 2012
Received in revised form 25 February 2014
Accepted 26 June 2014
Available online 3 July 2014

Keywords:
Accelerometer
Perspective correction
Keystone
Horizon
Digital photography
Sensor fusion
Noise propagation

ABSTRACT

Improper camera orientation produces convergent vertical lines (*keystone distortion*) and skewed horizon lines (*horizon distortion*) in digital pictures; an a-posteriori processing is then necessary to obtain appealing pictures. We show here that, after accurate calibration, the camera on-board accelerometer can be used to automatically generate an alternative perspective view from a virtual camera, leading to images with residual *keystone* and *horizon distortions* that are essentially imperceptible at visual inspection. Furthermore, we describe the uncertainty on the position of each pixel in the corrected image with respect to the accelerometer noise. Experimental results show a similar accuracy for a smartphone and for a digital reflex camera. The method can find application in customer imaging devices as well as in the computer vision field, especially when reference vertical and horizontal features are not easily detectable in the image.

1. Introduction

Photographic images may be affected by perspective distortions due to the attitude of the camera with respect to the depicted object. For instance, when the picture of a building is taken from the ground, the entire facade rarely fits into the frame without tilting the camera: the vertical lines appear then convergent in the acquired picture (*keystone distortion*, Fig. 1c). Similarly, a rolled camera leads to a skewed horizon line (*horizon distortion*, Fig. 1b).

Because of the poor appeal of images affected by perspective distortion (Fig. 1d), camera and lens manufacturers produce *tilt-shift* (TS) lenses [1,2], whose optical elements are manually shifted and tilted until a satisfying perspective is achieved. However, because of the cost and difficulty to use them, TS lenses are employed only by professional photographers in particular contexts (e.g. architectural photography). A simpler and cheaper solution is a-posteriori processing through widely diffused softwares like Photoshop [3], Gimp [4] or Hugin [5]. These apply a perspective transformation to the picture, whose parameters can be reliably estimated only if reference vertical and horizontal features are present in the picture and manually localized with high accuracy. More refined approaches automatically extract vertical and horizontal lines in man-made environments and have been exploited for texture map acquisition and metric measurements [6], camera calibration and 3D reconstruction from a single view [7], robotic navigation [8], perspective correction for documents [9] or digital photography [10]. The accuracy depends in any case on the accuracy on the estimated vanishing points [7]. In [11] the uncertainty on the automatically extracted vanishing points is minimized through the definition of a proper cost function, but the method cannot work if straight lines are not present, as in a portrait or a mountain view.

Integration of inertial sensors and vision systems has been explored to overcome this issue. For instance, after combined calibration of camera and inertial sensors [12–14], the gravity force is used as a reference to improve robot navigation [8,15] or for automotive applications [16]. In [17], accelerometer data and local invariant image features are used to estimate the camera pose and insert virtual objects into the scene, whereas in [18] the camera pose is estimated through the accelerometer to facilitate panorama imaging from multiple views. Accelerometers are used in [19] to estimate the camera pose and vibrations and consequently correct perspective and blur. A method to fully correct the perspective distortions using an accelerometer as a tilt sensor and the camera autofocus system as a distance measurement system is also described in [20].

Despite the large literature, on-board perspective correction is nowadays not present in digital cameras yet. This fact can be explained considering that accelerometer noise and imprecise calibration [17,18,21,22], as well as imprecise estimate of the orientation of the accelerometer with respect to the imaging sensor [8,12,14,15,23], may bias the

☆ This paper has been recommended for acceptance by Enrique Dunn, Ph.D.

* Corresponding author at: NVIDIA Research, 2701 San Tomas Expressway, 95050 Santa Clara, CA, USA.
E-mail address: ifrosio@nvidia.com (I. Frosio).

http://dx.doi.org/10.1016/j.imavis.2014.06.008
0262-8856/

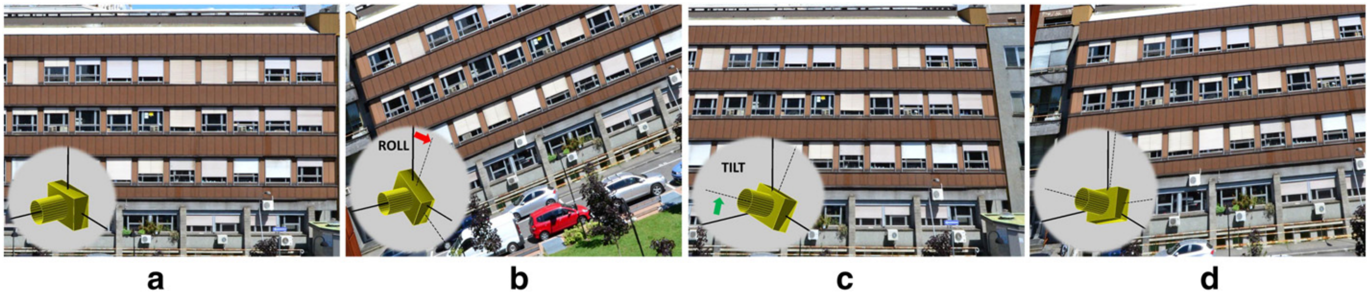

Fig. 1. Pictures of the same building facade acquired with: (a) camera correctly aligned; (b) rolled camera; (c) tilted camera; (d) rolled and tilted camera. Each panel also shows a pictorial representation of the camera orientation in the gravity reference frame.

estimated camera attitude. Here we show that an accurate system calibration makes automatic perspective correction feasible with high accuracy; furthermore, we show how to estimate the uncertainty of the position of each pixel in the corrected picture. The method was tested on a smartphone and on a digital reflex camera equipped with a zoom lens, with or without auto-focus enabled, leading in all cases to images with residual *keystone* and *horizon distortion* which are practically imperceptible at visual inspection.

2. Method

The method is based on estimating the pose of the camera from the output of the on-board accelerometer. The *accelerometer calibration* procedure in [22] is first applied; a set of images of a checkerboard is then taken from different points of view, measuring at the same time the accelerometer output. These data are used for traditional camera calibration [24] and to compute the rotation matrix $\mathbf{R}^{\mathbf{Cam}}_{\mathbf{Acc}}$ describing the orientation of the accelerometer with respect to the camera sensor plane (*camera-accelerometer calibration*, see Fig. 2), through an approach similar to the one described in [13]. After *accelerometer* and *camera-accelerometer calibration*, the camera attitude can be estimated for every acquired picture and an alternative perspective view can be generated. If zoom lenses are used, the proposed method still applies, using interpolation to estimate the camera internal parameters at every zoom focal length.

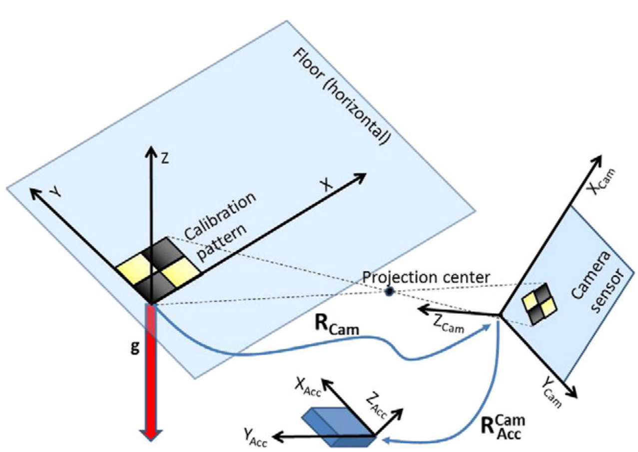

Fig. 2. Calibration setup. The current axes of the camera are obtained rotating the absolute axes by $\mathbf{R}_{\mathbf{Cam}}$. A further rotation by $\mathbf{R}^{\mathbf{Cam}}_{\mathbf{Acc}}$ in the camera reference frame gives the accelerometer axes. The horizontal calibration pattern defines the absolute reference system, where the gravity force is aligned with the -*Z* axis.

2.1. Accelerometer calibration

Let us define $\mathbf{V}_i = [V_{x,i}V_{y,i}V_{z,i}]^T$ the accelerometer voltage output and $\mathbf{g^{Acc}}_i = \left[g_{x,i}g_{y,i}g_{z,i}\right]^T$ the corresponding acceleration vector measured by the accelerometer. These are related by [21,22]:

$$\mathbf{g^{Acc}}_i = \mathbf{S}^{-1}\left(\frac{\mathbf{V}_i}{V_{CC}-\mathbf{O}}\right), \tag{1}$$

where V_{CC} is the voltage power supply, $\mathbf{O} = [O_xO_yO_z]^T$ is the offset vector and $\mathbf{S} = \{S_{ij}\}_{i,j\ =\ x,y,z}$ the sensitivity matrix. **S** and **O** are estimated through the method in [22], which automatically selects the best model (scalar, diagonal or symmetric matrix) for **S** based on the noise level and the number of the calibration data. To this aim, the set $\left\{\mathbf{g^{Acc}}_i\right\}_{i=1..M}$ is acquired with the accelerometer in *M* random orientations. Since in stationary condition the norm of $\mathbf{g^{Acc}}_i$ should equal the gravity acceleration, *g*, **S** and **O** are estimated minimizing the cost function *E*:

$$E = \sum_{i=1}^{M}\left[\left\|\mathbf{S}^{-1}\left(\frac{\mathbf{V}_i}{V_{CC}-\mathbf{O}}\right)\right\|^2 - g^2\right]^2 \tag{2}$$

through the Levemberg–Marquardt algorithm; as initial guess for **S** and **O** the nominal parameters reported in the sensor datasheet are used.

Each component of $\mathbf{g^{Acc}}_i$ is approximately affected by zero mean Gaussian noise with standard deviation $\sigma_g = \sigma_V s/V_{CC}$, where σ_V is the noise standard deviation on each channel of the accelerometer and *s* is the sensor average sensitivity [22]; noise standard deviation can be reduced to $\sigma_g/\sqrt{N_S}$ by averaging N_S samples acquired in stationary conditions.

2.2. Camera-accelerometer calibration

The aim of the *camera-accelerometer calibration* is the computation of the camera internal parameters [25] and of the rotation matrix $\mathbf{R}^{\mathbf{Cam}}_{\mathbf{Acc}}$, which describes the orientation of the accelerometer with respect to the camera sensor plane (Fig. 2).

Camera calibration is performed through the Matlab Toolbox [24]: a set of *N* images of a checkerboard, resting on the horizontal plane ($Z = 0$), is acquired from different positions. Then the Toolbox is used to estimate the camera focals, f_x and f_y, the camera principal point, $[c_x\ c_y]^T$ and a unique distortion parameter. The adoption of this simplified distortion model (also adopted in [26,27]) is justified considering that most of the cameras include a lens model in their firmware which automatically corrects the majority of distortions [1,2]. The use of a unique distortion parameter is furthermore needed for the interpolation scheme adopted with zoom lenses (see Section 2.5).