J. Vis. Commun. Image R. 46 (2017) 165–175

Contents lists available at ScienceDirect

J. Vis. Commun. Image R.

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

Distortion adaptive Sobel filters for the gradient estimation of wide angle images ☆

Antonino Furnari [a,*], Giovanni Maria Farinella [a], Arcangelo Ranieri Bruna [b], Sebastiano Battiato [a]

[a] University of Catania - Department of Mathematics and Computer Science, Italy
[b] STMicroelectronics, Advanced System Technology - Computer Vision, Italy

ARTICLE INFO

Article history:
Received 11 August 2016
Revised 6 March 2017
Accepted 22 March 2017
Available online 24 March 2017

Keywords:
Gradient estimation
Adaptive filters
Radial distortion
Wide angle images

ABSTRACT

We introduce a set of distortion adaptive Sobel filters for the direct estimation of geometrically correct gradients of wide angle images. The definition of the filters is based on Sobel's rationale and accounts for the geometric transformation undergone by wide angle images due to the presence of radial distortion. Moreover, we show that a local normalization of the filters magnitude is essential to achieve state-of-the-art results. To perform the experimental analysis, we propose an evaluation pipeline and a benchmark dataset of images belonging to different scene categories. Experiments on both, synthetic and real images, show that our approach outperforms the current state-of-the-art in both gradient estimation and keypoint matching for images characterized by large amounts of radial distortion. The collected dataset and the MATLAB code of the proposed method can be downloaded at our web page http://iplab.dmi.unict.it/DASF/.

1. Introduction ad related work

Wide angle cameras allow to gather a compact representation of the scene characterized by a wide Field Of View (FOV) [1] (see Fig. 1). Due to their convenient ability to acquire a large portion of the scene by means of a single device, wide angle cameras have gained popularity in many application fields including surveillance, automotive, robotics and first person vision [2–7]. The large Field Of View (FOV) is achieved by performing a non-uniform spatial sampling of the incoming light, which introduces radial distortion. The result is a non-Euclidean representation of the environment such that straight lines in the scene are not mapped to straight lines in the image [1,8]. Since many applications assume that the input images are the result of the perspective projection of the scene to a finite plane, geometrical considerations are often needed when processing wide angle images [3,8,9]. A number of methods have been proposed in the literature to establish a mapping between the distorted wide angle images and their ideal purely perspective counterparts [10–13]. When such a mapping is known and invertible, the most straightforward way to deal with wide angle images consists in explicitly compensating the radial distortion through a rectification process [14]. However, as pointed out in [9,15], this process is expensive in terms of computational time and memory and introduces blurred areas which can reduce the quality of the signal due to the required interpolation. To avoid such problems, researchers have investigated methods to perform feature extraction directly on the wide angle images: the authors of [16,17] studied how to compute the scale space of omnidirectional images; in [9,18,19] the Scale Invariant Feature Transform (SIFT) pipeline [20] is modified in order to be used directly on wide angle images; in [21] SIFT features are derived from wide angle images mapping them to a sphere; in [22] a direct approach to detect people using omnidirectional cameras is proposed; in [23] an algorithm to extract straight edges from distorted images is presented; in [24] gradient-based local descriptors are adapted to be computed directly on distorted images; in [25] a non-parametric method to define a pixel neighbourhood in catadioptric images is introduced.

In this paper we focus on the geometrically-correct estimation of the gradients of distorted images. The image gradient is a fundamental element in image analysis and vision and it is commonly used in a number of applications including image enhancement and edge extraction [26,27], object, scene and key-point representation [20,28–31], and gradient-domain-based image processing [32–35]. As it is shown in Fig. 1, a conventional computation of the gradients directly on wide angle images is deceived by the presence of radial distortion, while in practice many applications

☆ This paper has been recommended for acceptance by M.T. Sun.

* Corresponding author.
E-mail addresses: furnari@dmi.unict.it (A. Furnari), gfarinella@dmi.unict.it (G.M. Farinella), arcangelo.bruna@st.com (A.R. Bruna), battiato@dmi.unict.it (S. Battiato).

http://dx.doi.org/10.1016/j.jvcir.2017.03.019
1047-3203/

(a) Rectilinear Image (b) Wide Angle Image

Fig. 1. Gradient estimation of not-distorted and wide angle images. (a) A not-distorted image along with the gradient directions (solid red arrows) of some sample edges. (b) Wide angle counterpart of (a) along with the gradient directions (solid red arrows) of some sample edges deceived by the radial distortion. The ideal gradient directions are reported as dashed blue arrows. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

require a result similar to the ideal gradients depicted in Fig. 1(b) [9,22]. Some methods for estimating the gradients of wide angle images without performing the rectification already exist: in [9,23] the gradient estimated in the distorted domain with standard Sobel filters is corrected using an adaptive Jacobian correction matrix derived from the differential chain rule; in [22] the gradients of catadioptric images are estimated using an operator defined according to the geometry of the catadioptric mirror.

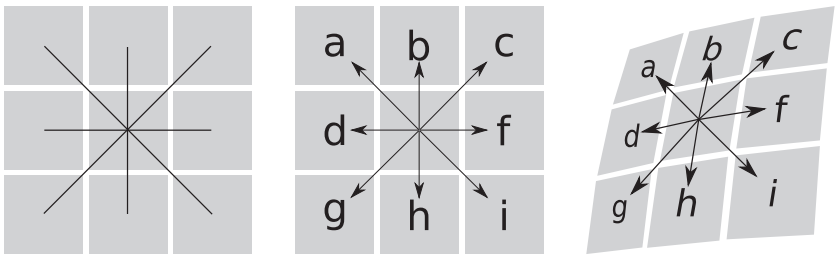
We derive a family of adaptive kernels for the geometrically-correct estimation of the image gradients of wide angle images. The proposed kernels aim to be invariant to radial distortion and hence they are designed to be beneficial for a number of gradient-based applications (such as key-point matching, object and people detection [20,29]), when they are deployed to wide angle camera systems. The derivation of our method is obtained by generalizing the standard Sobel operator to the case of non-Euclidean surfaces in order to take into account the geometrical transformation affecting the image. The derived filters adapt their shape according to the location on which they are computed in order to unevenly weighting the contributes of the estimated directional derivatives to compensate for distortion. The only requirement to compute the proposed filters is that the distortion function is known and invertible. The distortion function can be obtained by calibration when the camera is known, as usually happens in surveillance, automotive and robotics. We extend our previous work [36] introducing a normalization in the formulation of the proposed filters. We show that the new formulation improves the performances of the proposed gradient estimation method on the task of SIFT key-point matching. Moreover, the mathematical derivation is presented in more details, and the previously acquired benchmark dataset is augmented with scene-related attributes. We release the library used for the experiments along with the benchmark dataset to allow the reproducibility of the research.

The main contributions of this work are the following: (1) the proposed distortion adaptive Sobel filters are derived; (2) a novel benchmark dataset of high resolution images covering different scene categories is provided; (3) we define an evaluation protocol to assess the performances of the compared methods on the task of gradient estimation and key-point matching using artificially distorted images.

The reminder of the paper is organized as follows: in Section 2 we introduce the proposed filters; Section 3 defines the evaluation protocol; Section 4 discusses the results. Finally Section 5 concludes the paper providing insights for further research.

2. Distortion adaptive Sobel filters

The Sobel operator was originally proposed by Irwin Sobel in 1968 to estimate the gradient of a digital image [26,37,38]. Sobel proposed to estimate the gradient of the image at a given point by performing the vector summation of the simple central gradient estimates along the 4 main directions in a 3×3 neighbourhood (see Fig. 2(a)). According to Sobel, each of the 4 simple central gradient estimates can be expressed as a vector sum of a pair of orthogonal vectors. Each vector is a directional derivative estimate multiplied by a unit vector specifying the derivative's direction. The value of the directional derivative estimate for a pair of antipodal pixels in a 3×3 neighbourhood is defined as:



(a) (b) (c)

Fig. 2. A diagram of Sobel's rationale. (a) The 4 main directions in a 3×3 neighbourhood of a given point. (b) The 8 simple directional derivative estimates along with the appropriate unit vectors. (c) An example of distorted neighbourhood along with its directional derivative estimates.

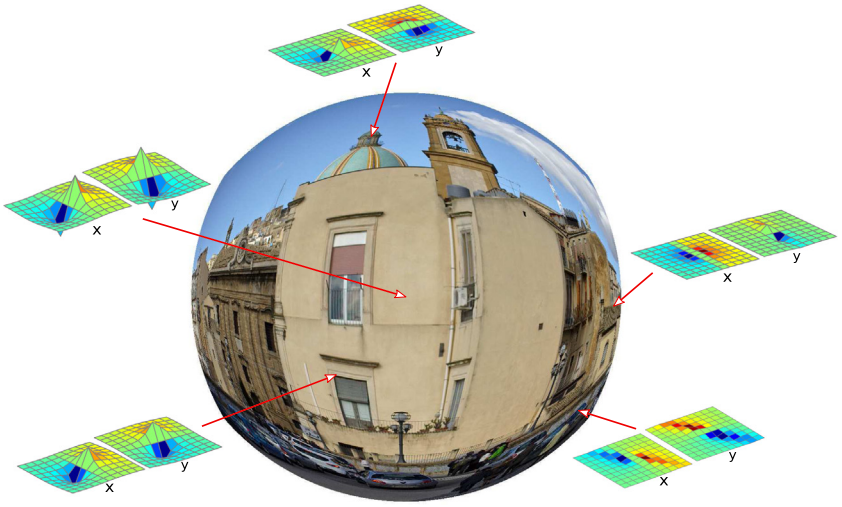


Fig. 3. Some graphical examples of the proposed kernels related to specific positions in the image on which they are computed. Kernel shape adapts to compensate the distortion characterizing a particular image location.

$$\frac{\text{density difference}}{\text{distance to neighbour}}. \tag{1}$$

The direction associated to the derivative estimate is given by the unit vector to the appropriate neighbour. Fig. 2(b) shows a schema of the directional derivative estimates. The reader is referred to [38] for a review of Sobel's rationale. The gradient estimation defined by Sobel can be formulated as the average of the eight oriented derivative vectors as follows:

$$\nabla I(x,y) = \frac{1}{8}\sum_{(s,t)\in S}\left(\frac{I^{s,t}_{x,y} - I^{-s,-t}_{x,y}}{\delta^{s,t}_{x,y}}\cdot\frac{(s,t)}{\sqrt{s^2+t^2}}\right) \tag{2}$$

where I is the considered image, $S = \{(s,t) : -1 \leqslant s,t \leqslant 1\}$, $I^{s,t}_{x,y} = I(x+s,y+t)$, $\delta^{s,t}_{x,y} = \delta((x+s,y+t),(x-s,y-t)),(s,t)$ denotes the vector of components s,t and magnitude $\sqrt{s^2+t^2}$ and δ is a given metric (e.g., Euclidean).

We start from Eq. (2) to build the distortion adaptive Sobel filters. Considering the symmetry of Eq. (2) with respect to the sign of s and t, it is convenient to define:

$$\{S_1,S_2\} \text{ partition of } S\setminus\{(0,0)\} \text{ s.t. } p\in S_1 \iff -p\in S_2. \tag{3}$$

Given the definition in Eqs. (3) and (2) can be rewritten in the following form:

$$\begin{aligned}\nabla I(x,y) = &\frac{1}{8}\sum_{(s,t)\in S_1}\left[\frac{I^{s,t}_{x,y} - I^{-s,-t}_{x,y}}{\delta^{s,t}_{x,y}}\cdot\frac{(s,t)}{\sqrt{s^2+t^2}}\right]\\ &+\frac{1}{8}\sum_{(s,t)\in S_2}\left[\frac{I^{s,t}_{x,y} - I^{-s,-t}_{x,y}}{\delta^{s,t}_{x,y}}\cdot\frac{(s,t)}{\sqrt{s^2+t^2}}\right].\end{aligned} \tag{4}$$

Given the definition in Eq. (3), it is possible to substitute the set S_2 with S_1 changing the signs of s and t in the second summation:

$$\begin{aligned}\nabla I(x,y) = &\frac{1}{8}\sum_{(s,t)\in S_1}\left[\frac{I^{s,t}_{x,y} - I^{-s,-t}_{x,y}}{\delta^{s,t}_{x,y}}\cdot\frac{(s,t)}{\sqrt{s^2+t^2}}\right]\\ &+\frac{1}{8}\sum_{(s,t)\in S_1}\left[\frac{I^{-s,-t}_{x,y} - I^{s,t}_{x,y}}{\delta^{-s,-t}_{x,y}}\cdot\frac{(-s,-t)}{\sqrt{s^2+t^2}}\right].\end{aligned} \tag{5}$$

Grouping the terms related to I and considering that δ is a metric (and hence $\delta((x_1,y_1),(x_2,y_2)) = \delta((x_2,y_2),(x_1,y_1)),\ \forall\ (x_1,y_1),(x_2,y_2)\in\Re^2$), we obtain Eq. (6):

$$\begin{aligned}\nabla I(x,y) &= \frac{1}{8}\sum_{(s,t)\in S_1}\left[\frac{I^{s,t}_{x,y} - I^{-s,-t}_{x,y}}{\delta^{s,t}_{x,y}}\left(\frac{(s,t)}{\sqrt{s^2+t^2}} - \frac{(-s,-t)}{\sqrt{s^2+t^2}}\right)\right]\\ &= \frac{1}{8}\sum_{(s,t)\in S_1}\left[\frac{I^{s,t}_{x,y}}{\delta^{s,t}_{x,y}}\left(\frac{(s,t)}{\sqrt{s^2+t^2}} - \frac{(-s,-t)}{\sqrt{s^2+t^2}}\right)\right] +\\ &\quad+\frac{1}{8}\sum_{(s,t)\in S_1}\left[\frac{I^{-s,-t}_{x,y}}{\delta^{s,t}_{x,y}}\left(\frac{(-s,-t)}{\sqrt{s^2+t^2}} - \frac{(s,t)}{\sqrt{s^2+t^2}}\right)\right].\end{aligned} \tag{6}$$

Leveraging the definition in Eq. (3), we obtain:

$$\begin{aligned}\nabla I(x,y) = &\frac{1}{8}\sum_{(s,t)\in S_1}\left[\frac{I^{s,t}_{x,y}}{\delta^{s,t}_{x,y}}\left(\frac{(s,t)}{\sqrt{s^2+t^2}} - \frac{(-s,-t)}{\sqrt{s^2+t^2}}\right)\right]\\ &+\frac{1}{8}\sum_{(s,t)\in S_2}\left[\frac{I^{s,t}_{x,y}}{\delta^{-s,-t}_{x,y}}\left(\frac{(s,t)}{\sqrt{s^2+t^2}} - \frac{(-s,-t)}{\sqrt{s^2+t^2}}\right)\right].\end{aligned} \tag{7}$$

Considering again the symmetric property of δ and considering the definition in Eq. (3), the Eq. (7) can be finally written as:

$$\nabla I(x,y) = \frac{1}{8}\sum_{(s,t)\in S}\left[\frac{I^{s,t}_{x,y}}{\delta^{s,t}_{x,y}}\left(\frac{(s,t)}{\sqrt{s^2+t^2}} + \frac{(s,t)}{\sqrt{s^2+t^2}}\right)\right]. \tag{8}$$

Let $h_1(x,y,s,t)$ and $h_2(x,y,s,t)$ be defined as follows:

$$h_1(x,y,s,t) = \begin{cases} 0 & \text{if } (s,t) = (0,0)\\ \frac{1}{4}\cdot\frac{1}{\delta^{s,t}_{x,y}}\cdot\frac{s}{\sqrt{(s^2+t^2)}} & \text{otherwise}\end{cases} \tag{9}$$

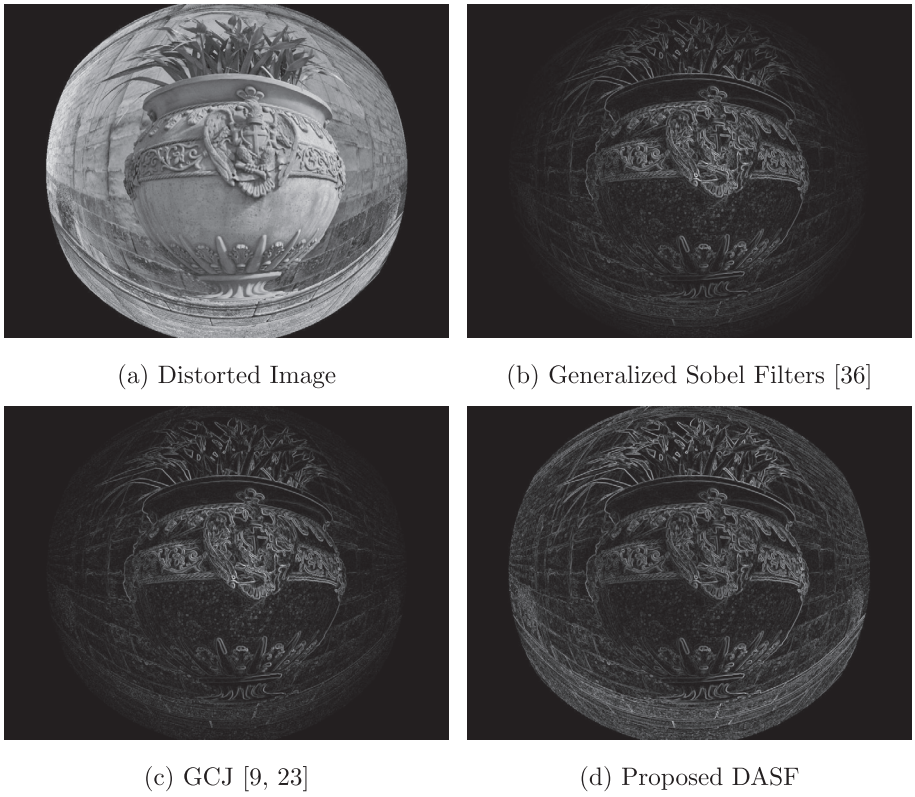
$$h_2(x,y,s,t) = \begin{cases} 0 & \text{if } (s,t) = (0,0)\\ \frac{1}{4}\cdot\frac{1}{\delta^{s,t}_{x,y}}\cdot\frac{t}{\sqrt{(s^2+t^2)}} & \text{otherwise.}\end{cases} \tag{10}$$

Considering the equations above, the gradient estimation can be expressed as follows:

$$\nabla_x I(x,y) = \sum_{(s,t)\in S} I(x+s,y+t)\cdot h_1(x,y,s,t) \tag{11}$$

$$\nabla_y I(x,y) = \sum_{(s,t)\in S} I(x+s,y+t)\cdot h_2(x,y,s,t). \tag{12}$$

Note that if the image has a planar geometry, the neighbourhood is not-distorted, similarly to what is shown in Fig. 2(b) and δ is naturally chosen as the Euclidean distance. In this case, the filters defined in Eqs. (9) and (10) are independent from the location of point (x,y) on which they are applied and Eqs. (11) and (12) are



(a) Distorted Image (b) Generalized Sobel Filters [36]

(c) GCJ [9, 23] (d) Proposed DASF

Fig. 4. An example distorted image (a) along with the magnitudes of gradients estimated using (b) the Generalized Sobel Filters (GSF) defined in Eqs. (9) and (10), (c) the Gradient Correction Jacobian (GCJ) method proposed in [9,23], and (d) the proposed Distortion Adaptive Sobel Filters (DASF).

equivalent to standard convolutions with Sobel filters (up to a factor of 16 in the estimation of the gradient magnitudes as discussed below). If the neighbourhood is not Euclidean (i.e., it is distorted) as in the case of wide angle images (see Fig. 2(c)), the δ function should be chosen according to the underlying geometrical model. In particular, if the distortion function $f : \Re^2 \rightarrow \Re^2$ which maps the not-distorted point of coordinates (u, v) to the distorted point (x, y) is known and invertible, the easiest way to choose a geometrically correct distance metric is to compose the Euclidean distance d with the inverse distortion function f^{-1} as follows:

$$\delta((x_1,y_1),(x_2,y_2)) = d(f^{-1}(x_1,y_1),f^{-1}(x_2,y_2)), \forall (x_1,y_1) \in \Re^2, (x_2,y_2) \in \Re^2. \quad (13)$$

Since f^{-1} is an inverse function, it is also bijective and hence the function δ defined above is a metric (the demonstration of this assertion can be found in A). The exploitation of Eq. (13) corresponds to the projection of the coordinates of the neighbourhood points into the Euclidean space prior to computing the distances in the classic way. In this general case, the terms related to δ in Eqs. (9) and (10) depend on the considered point (x, y) and hence the kernels are adaptive. Fig. 3 shows some graphical examples of the proposed distortion adaptive Sobel filters by considering the position in the image to which they are applied. As it can be noted, the proposed formulation yields kernels which adapt their shape in order to compensate for the radial distortion intrinsic to the different locations of the image. In this case, the computation defined in Eqs. (11) and (12) is not strictly a convolution since the signals defined in Eqs. (9) and (10) also depend on variables s and t. We refer to this computation as "adaptive convolution" of image I with the adaptive kernels h_1 and h_2 as it is intended in [9]. It should be noted that standard Sobel filters overestimate the gradient magnitudes by a scaling factor of 16 [38]. Therefore, a totally compatible formulation of the proposed filters can be achieved by multiplying Eq. (9) and Eq. (10) by 16. As in [36], the just derived filters will be referred to as Generalized Sobel Filters (GSF), in the rest of the paper.

2.1. Local normalization

Given the non-spatially uniform sampling of the incoming light operated by wide angle sensors, wide angle images are intrinsically multi-scale, which implies that the corresponding distance between neighbouring pixels in real world coordinates increases with the distance from the centre of the image. Considering that our filters are derived using the corresponding distances between neighbouring pixels by means of Eq. (13), filters computed in the peripheral areas of the image tend to have smaller coefficients in absolute values. As a result, the magnitudes of the estimated gradient decay in the peripheral areas of the image as it is depicted in Fig. 4(a)–(c). It should be noted that such problem is not specific to our method, but is common to other direct gradient estimation techniques (i.e., the Gradient Correction Jacobian (GCJ) method proposed in [9,23]) as shown in Fig. 4(c). To overcome this limit, we propose to locally normalize the derived GSF filters by the sum of the distances (computed according to the metric defined in Eq. (13)) between the antipodal pairs in the 3×3 neighbourhood:

$$\overline{h}_1(x,y,s,t) = \begin{cases} 0 & \text{if } (s,t) = (0,0) \\ \frac{1}{4} \cdot \frac{1}{\Delta_{x,y}} \cdot \frac{1}{\delta^{s,t}_{x,y}} \frac{s}{\sqrt{(s^2+t^2)}} & \text{otherwise} \end{cases} \quad (14)$$

$$\overline{h}_2(x,y,s,t) = \begin{cases} 0 & \text{if } (s,t) = (0,0) \\ \frac{1}{4} \cdot \frac{1}{\Delta_{x,y}} \cdot \frac{1}{\delta^{s,t}_{x,y}} \frac{t}{\sqrt{(s^2+t^2)}} & \text{otherwise} \end{cases} \quad (15)$$

where

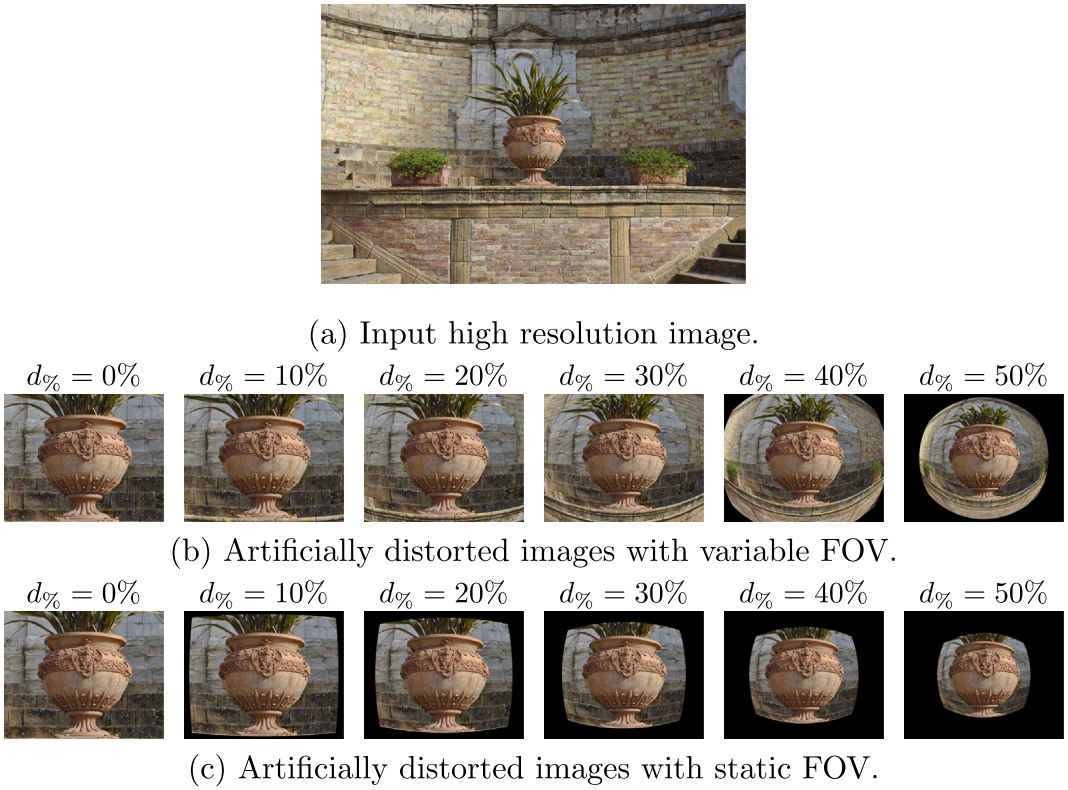


Fig. 5. A comparison of two different ways of artificially adding radial distortion to a reference image. (a) An input high resolution image. (b) Mapping the input high resolution image (a) to several low resolution images simulating a variable Field Of View. (c) Mapping an input low resolution image (the leftmost in the depicted series) to several low resolution images retaining a static Field Of View.

$$\Delta_{x,y} = \sum_{(s,t)\in S} \frac{1}{\delta^{s,t}_{x,y}}. \tag{16}$$

Fig. 4(d) shows the magnitudes of the gradients of the image depicted in Fig. 4(a) as computed using the filters defined in Eqs. (14) and (15). It should be noted that computing the gradients using the proposed locally normalized filters, allows to recover a huge quantity of details in the peripheral areas of the wide angle image. In the rest of the paper, the normalized filters defined in Eqs. (14) and (15) will be referred to as Distortion Adaptive Sobel Filters (DASF).

3. Evaluation protocol

The evaluation is carried on high resolution not-distorted images; radial distortion is artificially added as done in other studies [9,22,24,39]. This experimental approach allows to control the exact amount of distortion characterizing the target images. Moreover, the source (not-distorted) images can be used to compute the reference gradients which serve as a ground truth for the evaluations. In previous works, the radial distortion was added considering a static FOV [9,39] as it is shown in Fig. 5(c). It should be noted that this approach would leave the borders of the distorted image uncovered as the distortion rate increases. We propose to obtain a variable FOV by mapping high resolution reference images (Fig. 5(a)) to low resolution target images as shown in (Fig. 5(b)). In general, the artificial distortion is performed according to the following schema. Let I be the source not-distorted image and let $f : \Re^2 \rightarrow \Re^2$ be the distortion function mapping the not-distorted point $\mathbf{u}$ to its distorted counterpart $\mathbf{x}$. The distorted image is computed by mapping the not-distorted points of the reference image to the distorted locations and interpolating as needed:

$$\hat{I}(x,y) = I(f^{-1}(x,y)). \tag{17}$$

We perform three experiments to assess the performances of the proposed method. The first one aims at measuring the error related to the compared methods when the gradient is estimated, independently from any specific application. The second on experiment aims at assessing the impact of the proposed method on real-world applications. In particular, considering the importance of local feature description and matching [9,20,24,39], we assess the performances of the considered methods on the task of matching densely sampled SIFT features which have been computed using the estimated gradients. The third one is similar to the second one but it is performed on real imagery acquired using different fisheye cameras.

3.1. Experiment 1: Gradient estimation error

The image gradients are usually exploited separating the magnitudes from the orientations. The orientations carry important information about the distribution of edges in the scene, while the magnitudes give insights on the importance of each orientation. For this reason, many algorithms rely on the weighted histograms of gradient orientations [20,29]. We define an error measure by considering the average distance between the local populations (i.e., weighted histograms) of the gradient orientations in the reference image I and in its distorted counterpart $\hat{I}$.

Let $\mathcal{S}$ denote the Sobel operator, $\mathcal{G}$ the gradient estimator under analysis and let the distorted image $\hat{I}$ be divided into n non overlapping regular tiles of size $k \times k$ covering the entire surface: $\{\hat{T}_i\}_{1\leqslant i\leqslant n}$. For each tile $\hat{T}_i$ in the distorted image $\hat{I}$, we consider the related not-distorted tile T_i in the reference image I, which contains the not-distorted counterparts of all the points in $\hat{T}_i$. The error

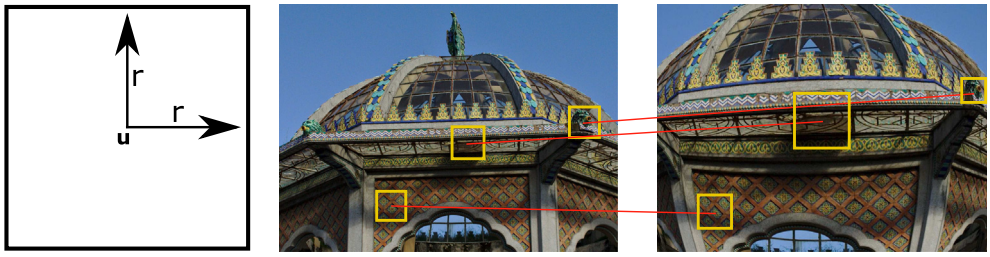


Fig. 6. (a) A schema of a support region and (b) an example of projecting support regions to the distorted space.

related to the gradient estimator $\mathcal{G}$ given the image pair $(\hat{I}, I)$ is defined as:

$$\epsilon(\mathcal{G}, \hat{I}, I) = \frac{1}{n}\sum_{i=1}^{n}\rho(\mathcal{H}(\mathcal{G}\hat{T}_i), \mathcal{H}(\mathcal{S}T_i)) \tag{18}$$

where $\mathcal{H}(\mathcal{G}\hat{T}_i)$ and $\mathcal{H}(\mathcal{S}T_i)$ denote the weighted histograms of the estimated and reference gradient orientations (i.e. Sobel on the reference image) of tiles $\hat{T}_i$ and T_i, and ρ is the metric based on the Bhattacharyya coefficient as defined in [40]:

$$\rho(H_1, H_2) = \sqrt{1 - \sum_{u=1}^{m}\sqrt{H_1^u \cdot H_2^u}}. \tag{19}$$

In Eq. (19), m is the number of bins of histograms H_1 and H_2 and H^u denotes the $u-th$ component of H. In our experiments, the following parameters have been used: (1) each image is divided into tiles of size 24×24 pixels and (2) the histograms have 18 bins evenly spacing the interval $[-180^\circ, 180^\circ]$.

3.2. Experiment 2: Local feature matching

The aim of this second experiment is to assess the performances of the considered methods on the task of local feature matching. To this end, we consider the popular gradient-based SIFT descriptor [20]. The proposed experiment follows the methodology described in [24], which is briefly summarized in the following. Given a reference-target image pair $(I, \hat{I})$, square support regions are considered at multiple scales on the reference image I. In particular, we sample support regions of radii: 32, 64, 128 and 256 pixels at a regular step of 50 pixels. We consider a support region as an entity $\mathcal{R}(\mathbf{u}, r)$ composed by two elements: a centre $\mathbf{u}$ and a radius r (see Fig. 6(a)). Each support region $\mathcal{R}(\mathbf{u}, r)$ is mapped to a corresponding region $\hat{\mathcal{R}}(\mathbf{x}, \hat{r})$ in the distorted space using Eq. (23) to map the not-distorted point $\mathbf{u}$ to its distorted counterpart $\mathbf{x}$. The radius of the distorted support region $\hat{r}$ is computed using the division model as detailed in [9]:

$$\hat{r} = \frac{2r}{1 + \sqrt{1 - 4\xi r^2}}. \tag{20}$$

Fig. 6(b) shows an example of such projection. All the projected regions not entirely contained in the distorted image $\hat{I}$ or which projected radius is under 16 pixels are discarded along with their not-distorted counterparts.

Standard SIFT descriptors $\mathcal{D}$ are computed on the reference support regions $\mathcal{R}$ using the reference gradient estimated with the Sobel operator. Distortion Adaptive SIFT Descriptors $\hat{\mathcal{D}}$ are computed from the projected support regions $\hat{\mathcal{R}}$ using the generic estimator $\mathcal{G}$ as detailed in [24]. Matches between the reference $\mathcal{D}$ and target $\hat{\mathcal{D}}$ descriptors are computed using the nearest neighbour criterion, i.e., descriptor $d \in \mathcal{D}$ is matched to its nearest neighbour $\hat{d} \in \hat{\mathcal{D}}$. Given the known correspondences between the reference and target descriptors, the matching ability score is measured as follows:

$$\textit{matching ability score} = \frac{\#\textit{correct matches}}{\#\textit{matches}}. \tag{21}$$

3.3. Experiment 3: Local feature matching in real imagery

Experiments 1 and 2 are performed in controlled settings simulating the effect of radial distortion on high resolution images. To validate our analysis, we also perform experiments on images acquired using real wide-angle cameras. Let I_1 and I_2 be two real images acquired using a calibrated wide-angle camera, and let I_1^u and I_2^u be their ideal, undistorted counterparts. Since the camera is calibrated, the mapping function f relating each image I_i^u to I_i is assumed to be known (and invertible). We assume that both images represent a planar object acquired from different points of view and that the homography H_{12} relating the two views of the same planar surface is known. Since H_{12} allows to project points of I_1^u to I_2^u, the following mapping can be used to project points of I_1 to I_2:

$$\Psi_{12}(x, y) = f(h_{12}(f^{-1}(x, y))), \tag{22}$$

where h_{12} projects point (x, y) using homography H_{12}.

Given images I_1 and I_2 and the mapping Ψ_{12}, we assess the influence of the proposed gradient estimation method when used to compute and match SIFT descriptors on the two images. Since images I_1 and I_2 are in general related by rotation, scale and viewpoint changes, we first extract radial distortion invariant SIFT keypoints from each image using the method proposed by Lourenço et al. [9]. SIFT descriptors are hence computed for each keypoint employing the gradients obtained using the given estimation method $\mathcal{G}$ under analysis. Descriptors are finally matched using the standard method proposed by Lowe [20].

Ground truth correspondences between keypoints K_1 and K_2 respectively extracted from image I_1 and I_2 are obtained as follows: keypoints K_1 and K_2 are first projected to the reference undistorted space of I_2^u, obtaining the sets K_1^u and K_2^u; each keypoint belonging to K_1^u is assigned to the closest keypoint in K_2^u; associated keypoints which are more than 10 pixels apart, are not considered.

The matching ability score related to the considered image pair I_1 and I_2 is finally computed by counting the fraction of correct matches among the matched descriptors.

3.4. Modelling the radial distortion

To model the radial distortion we rely on the Division Model [9,13], which has proven successful in modelling many real dioptric cameras [41]. The Division Model defines the mapping between the not-distorted point $\mathbf{u} \in \Re^2$ and its distorted counterpart $\mathbf{x} \in \Re^2$ as follows:

Fig. 7. Some randomly chosen images from the proposed Dataset A, for each considered scene-based tag.

Table 1
Numbers of images in Dataset A containing a specific scene-tag.

Scene	Indoor	Outdoor	Natural	Handmade	Urban	Car	Pedestrian	Street
Count	13	87	44	93	51	49	19	50

$$\mathbf{x} = f(\mathbf{u}) = \frac{2\mathbf{u}}{1 + \sqrt{1 - 4\xi||\mathbf{u}||^2}}. \tag{23}$$

All the coordinates are referred to the centre of the distortion, which is conveniently set to the centre of the image in our experiments. Unfortunately, the effects of parameter ξ depend on the dimensions of the image, which makes its interpretation not intuitive. Therefore, we propose to quantify the distortion undergone by an image as the percentage:

$$d_{\%} = 1 - \frac{\hat{r}_M}{r_M} \tag{24}$$

where r_M is the distance from the centre to the corner of the image and $\hat{r}_M$ represents its distorted counterpart. The mathematical relationship between r_M and $\hat{r}_M$, given the distortion parameter ξ, can be derived from Eq. (23):

$$r_M = \frac{\hat{r}_M}{1 + \xi\hat{r}_M^2}. \tag{25}$$

Given Eqs. (24) and (25), the distortion parameter ξ can be computed from the chosen percentage of distortion $d_{\%}$ using the following formula:

$$\xi = -\frac{d_{\%}}{[r_M(1 - d_{\%})]^2}. \tag{26}$$

3.5. Input data

For our experiments, we consider two datasets. The first one is referred to as Dataset A and contains high resolution images to which radial distortion is artificially added to perform experiments in controlled settings. The second one is referred to as Dataset B and comprises real imagery acquired using different fisheye cameras.

Dataset A. We collected a dataset of 100 high resolution images belonging to the following scenes categories: indoor, outdoor, natural, handmade, urban, car, pedestrian, street. The considered categories are relevant to the main application domains where the image gradients are employed [20,29], and consistent with the scene categorization proposed by Torralba and Oliva [42]. Each image is provided with one or more tags related to the above specified scene categories. All images have been acquired using a Canon 650D camera mounting a Canon EF-24 mm lens and have resolution equal to 5204×3472 pixels. Fig. 7 shows some examples of the input images used for the evaluations, whereas Table 1 reports some statistics about the scene-related tags present in the dataset. The dataset can be downloaded at our web page http://iplab.dmi.unict.it/DASF/.

Dataset B. To perform evaluations using real wide-angle images, we also consider the benchmark dataset introduced in [9]. This dataset comprises three image series acquired with fisheye lenses exhibiting different amounts of radial distortion. Each series is provided with calibration images and division model parameters. Fig. 8 shows some samples from the considered dataset. For each image series, we report the related distortion rates computed according to our model (i.e., using Eq. (24)): 13%, 19% and 54%. Each series contains 13 images related by different transformations including viewpoint change, rotation and scale. Images belonging to the same series represent a scene containing the same planar object, acquired from different positions. All image pairs within a series are provided with an homography relating the common planar object. All available 78 image pairs included in each series are considered in our experiments.

4. Results

We compare the performances of the proposed filters (Eqs. (14) and (15)) with the Gradient Correction Jacobian (GCJ) method defined in [9,23], the Generalized Sobel Filters (GSF) proposed in

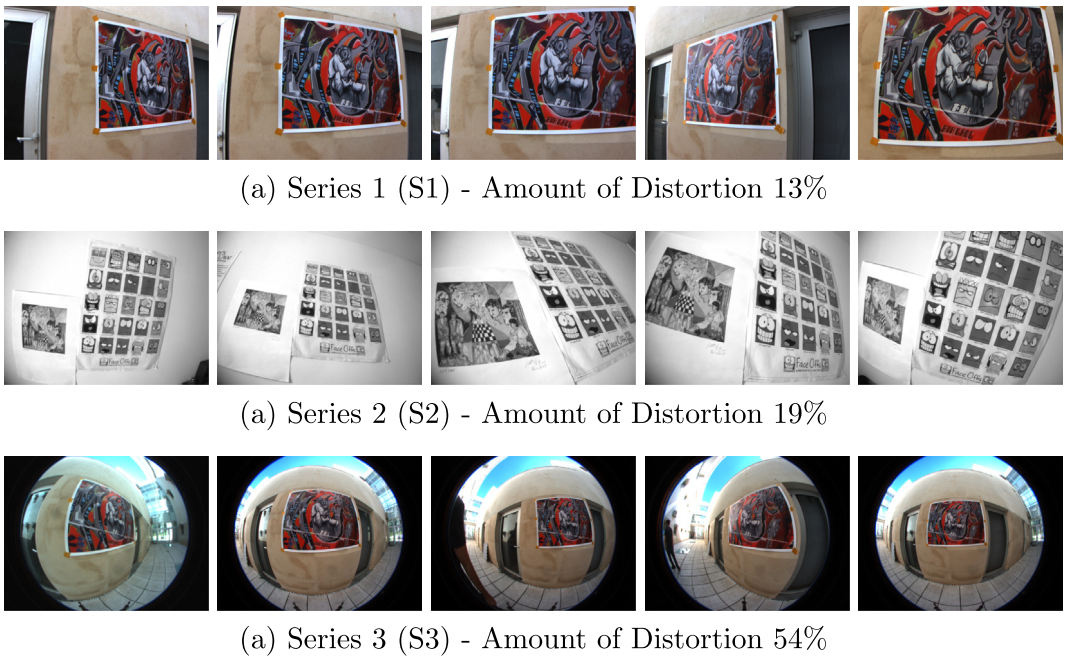


Fig. 8. Some sample images from the three real image series in Dataset B [9].

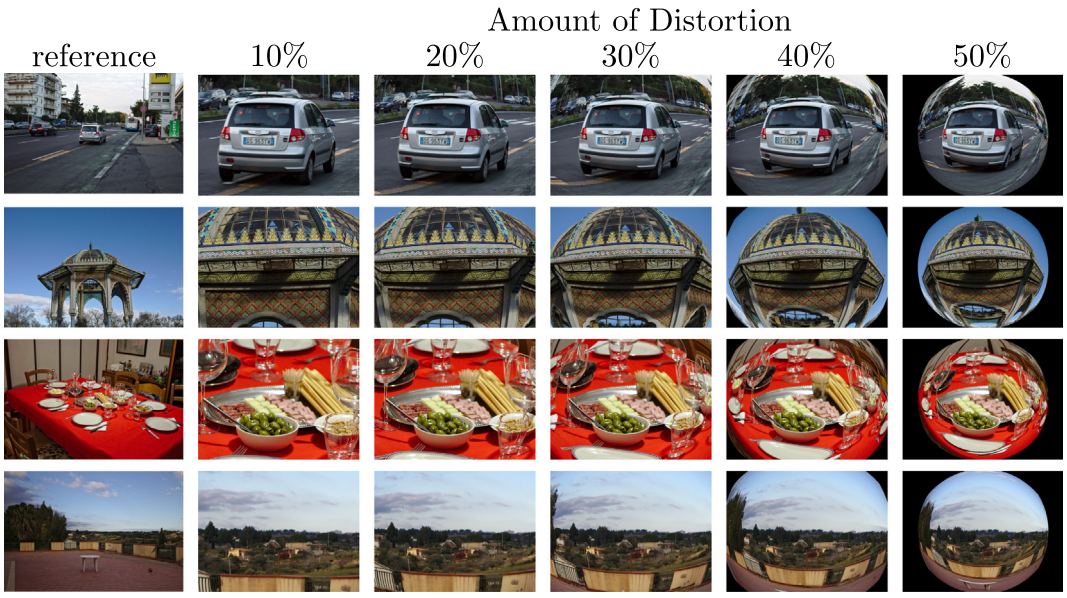


Fig. 9. Some examples of the artificially distorted images used in experiments 1 and 2. The first column reports the full-resolution reference images used to generate the low resolution distorted images reported in the other columns. All image series are derived from Dataset A.

[36] and reported in Eqs. (9) and (10), and two baseline methods consisting in: (1) ignoring the distortion (the gradients are computed using the classic Sobel operator in the distorted space and the result is not corrected); (2) explicitly removing the distortion (i.e., performing rectification) before estimating the gradient using the Sobel operator. In particular, we perform three experiments.

Experiments 1 and 2 are performed on Dataset A. To assess the performances of the considered methods with respect to different degrees of distortion, we processed every image of Dataset A adding varying amounts of distortion ranging from a minimum of 10% to a maximum of 50%. The reference gradients are computed from the reference images using the standard Sobel filters. Fig. 9 shows some examples of the images used in Experiments 1 and 2. Experiment 3 is performed on Dataset B, which comprises real imagery acquired using three different wide angle cameras.

4.1. Experiment 1

Fig. 10 shows the results related to Experiment 1, i.e., the mean error committed by the considered methods with respect to different amounts of radial distortion. Each curve is obtained averaging the error scores related to the 100 images in dataset A and computed using Eq. (18). The legend of Fig. 10 reports in parenthesis the average value of each curve, which should reflect the average performances of the methods with respect to all the considered amounts of distortion. The rectification method allows to achieve

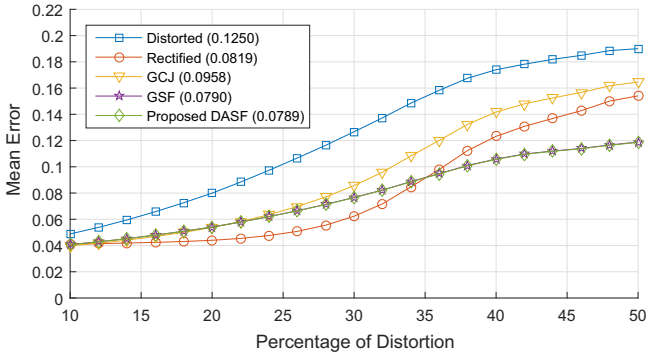


Fig. 10. Results related to Experiment 1. The mean error curves for different gradient estimators on the considered dataset at varying of the percentage of distortion. The average value for each curve is reported in parenthesis in the legend.

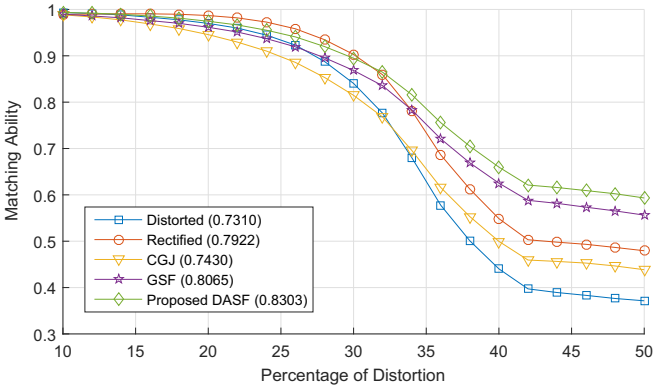


Fig. 11. Results related to Experiment 2. The matching ability for different gradient estimators on the considered dataset at varying of the percentage of distortion.

good results for low distortion rates (where the lost information can still be "guessed" by the rectification process), whereas the error gets higher as the distortion rate increases. The proposed filters perform better than the competitors for distortion rates over the 35%. Table 2 summarizes the average errors related to the subset of images characterized by a specific scene-based tag. The reported results suggest that the proposed filters, the GSF and the GCJ methods offer significant improvements over the distorted gradients. In particular, the proposed method is the best performing when all scene categories are considered (top row of Table 2) and is always among the two best performing methods for each of the scene categories. As it appears clear from Fig. 10 and Table 2, the proposed filters closely match the performances of the GSF method in this experiment. This is due to the fact that, considering regular tiles as small as 24×24 pixels, the effect of the local normalization introduced in Section 2.1 is negligible with respect to the normalization operated by the computation of the weighted histogram of gradient orientations.

4.2. Experiment 2

Fig. 11 shows the results related to Experiment 2, i.e., the matching ability achieved by the considered methods with respect to different amounts of radial distortion. Each curve is obtained averaging the matching ability scores related to the 100 images in Dataset A. As in Fig. 10, the legend of Fig. 11 reports the average value of each curve which should reflect the average performances of the methods with respect to all the considered amounts of distortion. The proposed method retains the highest matching ability for amounts of distortion over the 30%, while it performs comparably to the rectification methods for distortion rates below the 30%. The improvement over the non-locally-normalized GSF filters is in this case evident. Similarly to Tables 2 and 3 reports the average matching abilities related to images characterized by specific image tags. The proposed method is always the best performing (highest score), except in the case of indoor images, where,

Table 2
Results related to Experiment 1. Average errors for different methods and scene types. In each row, the two smallest values are underlined, while the minimum is reported in **bold** letters.

Scene	Distorted	Rectified	GCJ	GSF	Proposed
All	0.1250	0.0819	0.0958	0.0790	**0.0789**
Indoor	0.1172	0.0885	0.1049	**0.0826**	**0.0826**
Outdoor	0.1262	0.0809	0.0945	0.0785	**0.0784**
Natural	0.1129	0.0816	0.0980	**0.0736**	**0.0736**
Handmade	0.1280	0.0819	0.0952	0.0801	**0.0800**
Urban	0.1358	**0.0796**	0.0910	0.0816	0.0813
Car	0.1366	**0.0807**	0.0919	0.0827	0.0824
Pedestrian	0.1417	**0.0748**	0.0846	0.0807	0.0805
Street	0.1357	**0.0807**	0.0921	0.0823	0.0821

Table 3
Results related to Experiment 2. Average matching ability scores for different methods and scene types. In each row, the two largest values are underlined, while the maximum is reported in **bold** letters.

Scene	Distorted	Rectified	GCJ	GSF	Proposed
All	0.7310	0.7922	0.7430	0.8065	**0.8303**
Indoor	0.6394	**0.7434**	0.6685	0.7091	0.7366
Outdoor	0.7447	0.7995	0.7541	0.8210	**0.8444**
Natural	0.6958	0.7594	0.6885	0.7600	**0.7911**
Handmade	0.7472	0.8027	0.7601	0.8236	**0.8452**
Urban	0.7776	0.8260	0.8010	0.8644	**0.8813**
Car	0.7779	0.8248	0.7990	0.8601	**0.8768**
Pedestrian	0.7890	0.8403	0.8252	0.8834	**0.8993**
Street	0.7780	0.8235	0.7974	0.8591	**0.8758**

Table 4
Results related to Experiment 3. Average matching ability for the 78 image pair included in each series of Dataset B. In each row, the two largest values are underlined, while the maximum is reported in **bold** letters.

Series	Distorted	Rectified	CGJ	GSF	Proposed
S1 (13%)	0.6995	**0.7022**	0.6953	0.6907	0.6991
S2 (19%)	0.2646	**0.2746**	0.2542	0.2572	0.2585
S3 (54%)	0.3547	0.3566	0.3430	0.2938	**0.3730**

probably due to the regularity of straight edges, the rectification performs slightly better.

4.3. Experiment 3

Table 4 reports the results related to the Experiment 3, performed on real images. For each series in Dataset B, the table reports the average matching ability for all 78 image pairs. The reported results agree with the previous observations. Specifically, when the amount of distortion is low (e.g., S1 - 13% and S2 - 19% in Table 4), rectification allows to obtain the best results, while all other methods exhibit similar performances. When the amount of radial distortion is larger (e.g., S3 - 54%), the proposed gradients allow better results than all competitors, including rectification. Interestingly, the results highlight the importance of the local normalization introduced in the proposed filters with respect to GSF, especially in the case of images characterized by large amounts of radial distortion.

4.4. Computational times

Differently than standard Sobel filters, the proposed ones adapt their shape according to the location on which they are computed. In order to make computation tractable, adaptive filters can be precomputed when the camera is calibrated. However, since adaptive filters are not in general symmetrical with respect to the horizontal or vertical direction, they are not separable and hence they are less computationally efficient than standard Sobel filters.

Even if the proposed filters are less computationally efficient than standard Sobel filters, their use is still advantageous with respect to the standard pipeline involving image rectification. This is shown in Fig. 12, where we report average computational times for the different gradient estimation pipelines considered in this paper. Results are reported for five different resolutions ranging from qVGA (320×240) to full HD (1920×1080). Standard convolutions with Sobel filters are performed using the MATLAB function *imgradientxy*. Image rectification and interpolation are performed using the standard MATLAB function *interp2*. Convolution with the proposed filters and correction using the CGJ method are performed using optimized C implementations. All times are measured on a laptop equipped with a 2.00 GHz i7 processor and averaged over 100 different runs.

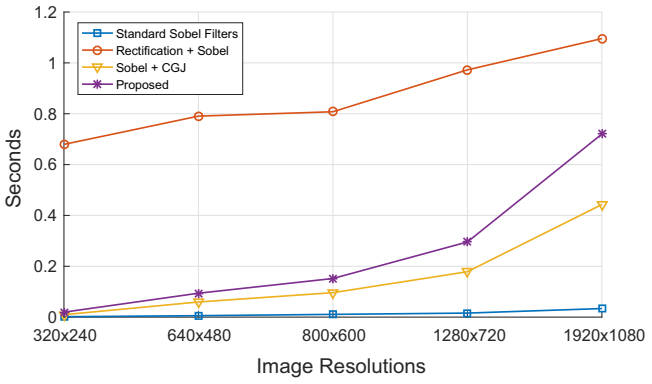


Fig. 12. Per-image average computational times of the considered approaches when computing the gradients of images of different resolutions. Times are averaged over 100 runs.

While the computational time needed by the proposed method increases faster than with other methods, using the proposed filters is still computationally advantageous over performing rectification. Moreover, note that, even in the case of full HD images, the average time required to estimate gradients using the proposed method is less than 1 s.

4.5. Discussion

It should be noted that, despite the loss of performance of direct gradient estimation techniques for distortion rates under 30–35%, it is still advantageous to use such techniques in order to avoid the rectification process, which is computationally expensive and not always feasible [9,15]. This is also confirmed by the computational analysis reported in Section 4.4, which shows how direct gradient estimation techniques are more computationally efficient than performing rectification for all distortion rates. Moreover, the majority of wide angle cameras used in real-world applications, are designed to cover 180° diagonal or vertical FOV corresponding to high distortion rates, as it is discussed in [41].

5. Conclusion and future works

We have proposed a family of distortion adaptive Sobel filters which can be used to correctly estimate the gradient of distorted images. The proposed filters are independent from the adopted distortion model and only require the distortion function f to be known and invertible. We have assessed the performances introducing a dataset of high resolution images belonging to different scene categories and defining an evaluation protocol which measures the error between estimated and reference gradients, and allows to compare the performances on the task of local feature matching. The experiments show that our method outperforms the competitor methods. Future works could be devoted to the integration of the proposed method in popular applications relying on the image gradients such as object and people detection [20,29]. Moreover, additional efforts could be devoted to derive approximations of the proposed filters to obtain separability and constancy with respect to the image radius similarly to what done in [9]. Dataset and MATLAB code can be downloaded from the following page: http://iplab.dmi.unict.it/DASF/.

Acknowledgment

This work has been performed in the project PANORAMA, co-funded by grants from Belgium, Italy, France, the Netherlands, the United Kingdom, and the ENIAC Joint Undertaking.

Appendix A. Demonstration of the metric properties of Eq. (13)

Let be $f : \Re^2 \rightarrow \Re^2$ an invertible function and let $f^{-1} : \Re^2 \rightarrow \Re^2$ be its inverse. Then the function defined as

$$\overline{\delta}((x_1,y_1),(x_2,y_2)) = d(f^{-1}(x_1,y_1),f^{-1}(x_2,y_2)) \quad (13)$$

where $d:\Re^2 \rightarrow \Re^2$ is the Euclidean metric, is a metric.

Demonstration: All the metric properties are demonstrated noting that f^{-1} is a bijective function and leveraging the metric properties of the Euclidean distance:

- Non-negativity: $\forall\, \mathbf{x},\mathbf{y} \in \Re^2\ \exists \mathbf{u},\mathbf{v} \in \Re^2 : \delta(\mathbf{x},\mathbf{y}) = d(\mathbf{u},\mathbf{v}) \geqslant 0$. Then:

$$\forall\, \mathbf{x},\mathbf{y} \in \Re^2,\ \delta(\mathbf{x},\mathbf{y}) \geqslant 0. \quad (A.1)$$

- Coincidence: $\forall\, \mathbf{x},\mathbf{y} \in \Re^2\ \exists\ \mathbf{u},\mathbf{v} \in \Re^2 : \delta(\mathbf{x},\mathbf{y}) = d(\mathbf{u},\mathbf{v})$, moreover, since f^{-1} is bijective, $\mathbf{u}=\mathbf{v} \Longleftrightarrow \mathbf{x}=\mathbf{y}$. Then:

$$\forall\, \mathbf{x},\mathbf{y} \in \Re^2, \delta(\mathbf{x},\mathbf{y}) = 0 \Longleftrightarrow \mathbf{x}=\mathbf{y}. \quad (A.2)$$

- Symmetry: $\forall\, \mathbf{x},\mathbf{y} \in \Re^2\ \exists\ \mathbf{u},\mathbf{v} \in \Re^2 : \delta(\mathbf{x},\mathbf{y}) = d(\mathbf{u},\mathbf{v})$, moreover $\delta(\mathbf{x},\mathbf{y}) = d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u}) = \delta(\mathbf{y},\mathbf{x})$. Then:

$$\forall\, \mathbf{x},\mathbf{y} \in \Re^2, \delta(\mathbf{x},\mathbf{y}) = \delta(\mathbf{y},\mathbf{x}). \quad (A.3)$$

- Triangle inequality: $\forall\, \mathbf{x},\mathbf{y},\mathbf{z} \in \Re^2\ \ \exists\ \mathbf{u},\mathbf{v},\mathbf{p} \in \Re^2 : \mathbf{u}=f^{-1}(\mathbf{x})$, $\mathbf{v}=f^{-1}(\mathbf{y})$, $\mathbf{p}=f^{-1}(\mathbf{z})$. Then, given the triangle inequality for the Euclidean distance $d(\mathbf{u},\mathbf{p}) \leqslant d(\mathbf{u},\mathbf{v}) + d(\mathbf{v},\mathbf{p})$, it follows that:

$$\forall\, \mathbf{x},\mathbf{y},\mathbf{z} \in \Re^2,\ \delta(\mathbf{x},\mathbf{y}) \leqslant \delta(\mathbf{x},\mathbf{z}) + \delta(\mathbf{y},\mathbf{z}). \quad (A.4)$$

References

[1] K. Miyamoto, Fish eye lens, J. Opt. Soc. Am. (1964) 2–3.
[2] L. Puig, J.J. Guerrero, Omnidirectional Vision Systems, Springer, 2013.
[3] C. Hughes, M. Glavin, E. Jones, P. Denny, Wide-angle camera technology for automotive applications: a review, IET Intell. Transport Syst. 3 (1) (2009) 19–31.
[4] S. Battiato, G.M. Farinella, A. Furnari, G. Puglisi, A. Snijders, J. Spiekstra, Vehicle tracking based on customized template matching, International Workshop on Ultra Wide Context and Content Aware Imaging in conjunction with Computer Vision Theory and Applications (VISAPP), vol. 2, 2014, pp. 755–760.
[5] S. Battiato, G.M. Farinella, A. Furnari, G. Puglisi, A. Snijders, J. Spiekstra, An integrated system for vehicle tracking and classification, Exp. Syst. Appl. 42 (21) (2015) 7263–7275.
[6] A. Furnari, G.M. Farinella, S. Battiato, Recognizing personal contexts from egocentric images, in: Workshop on Assistive Computer Vision and Robotics (ACVR) in conjunction with the IEEE International Conference on Computer Vision, 2015, pp. 1–9.
[7] A. Furnari, G.M. Farinella, S. Battiato, Recognizing personal locations from egocentric videos, IEEE Trans. Hum.-Mach. Syst. 47 (1) (2017) 6–18.
[8] M.M. Fleck, Perspective Projection: The Wrong Imaging Model, Department of Computer Science, University of Iowa, 1995.
[9] M. Lourenço, J.P. Barreto, F. Vasconcelos, sRD-SIFT: keypoint detection and matching in images with radial distortion, IEEE Trans. Robot. 28 (3) (2012) 752–760.
[10] J. Kannala, S. Brandt, A generic camera calibration method for fish-eye lenses, International Conference of Pattern Recognition, vol. 1, 2004, pp. 10–13.
[11] D. Scaramuzza, A. Martinelli, R. Siegwart, A toolbox for easily calibrating omnidirectional cameras, in: International Conference on Intelligent Robots and Systems, 2006, pp. 5695–5701.
[12] F. Devernay, O. Faugeras, Straight lines have to be straight, Mach. Vis. Appl. 13 (1) (2001) 14–24.
[13] A.W. Fitzgibbon, Simultaneous linear estimation of multiple view geometry and lens distortion, Computer Vision and Pattern Recognition, vol. 1, 2001, pp. 125–132.
[14] C. Hughes, P. Denny, E. Jones, M. Glavin, Accuracy of fish-eye lens models, Appl. Opt. 49 (17) (2010) 3338–3347.
[15] A. Furnari, G.M. Farinella, A.R. Bruna, S. Battiato, Affine covariant features for fisheye distortion local modeling, IEEE Trans. Image Process. 26 (2) (2017) 696–710.
[16] L. Puig, J.J. Guerrero, Scale space for central catadioptric systems: towards a generic camera feature extractor, in: IEEE International Conference on International Conference on Computer Vision, 2011, pp. 1599–1606.
[17] I. Bogdanova, X. Bresson, J. Thiran, P. Vandergheynst, Scale space analysis and active contours for omnidirectional images, IEEE Trans. Image Process. 16 (7) (2007) 1888–1901.
[18] Z. Arican, P. Frossard, OmniSIFT: scale invariant features in omnidirectional images, in: IEEE International Conference on Image Processing, 2010, pp. 3505–3508.
[19] J. Cruz-Mota, I. Bogdanova, B. Paquier, M. Bierlaire, J. Thiran, Scale invariant feature transform on the sphere: theory and applications, Int. J. Comput. Vis. 98 (2) (2011) 217–241.
[20] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110.
[21] P. Hansen, P. Corke, W. Boles, K. Daniilidis, Scale-invariant features on the sphere, in: IEEE International Conference on Computer Vision, 2007, pp. 1–8.
[22] I. Cinaroglu, Y. Bastanlar, A direct approach for human detection with catadioptric omnidirectional cameras, in: Signal Processing and Communications Applications Conference, 2014, pp. 2275–2279.
[23] M.S. Islam, L.J. Kitchen, Straight-edge extraction in distorted images using gradient correction, in: Digital Image Computing: Techniques and Applications, 2009, pp. 505–512.
[24] A. Furnari, G.M. Farinella, A.R. Bruna, S. Battiato, Distortion adaptive descriptors: Extending gradient-based descriptors to wide angle images, in: Image Analysis and Processing (ICIAP), Lecture Notes in Computer Science, vol. 9280, Springer, 2015, pp. 205–215.
[25] S. Ainouz, O. Morel, D. Fofi, S. Mosaddegh, A. Bensrhair, Adaptive processing of catadioptric images using polarization imaging: towards a pola-catadioptric model, Opt. Eng. 52 (3) (2013) 37001.
[26] P.E. Duda, O. Richard, Hart, Pattern Classification and Scene Analysis, John Wiley and Sons, New York, 1973, pp. 271–272.
[27] J. Canny, A computational approach to edge detection, Pattern Anal. Mach. Intell. 8 (6) (1986) 679–698.
[28] J. Domke, Y. Aloimonos, Deformation and viewpoint invariant color histograms, in: The British Machine Vision Conference, 2006, pp. 509–518.
[29] N. Dalal, B. Triggs, Histograms of oriented gradients for human detection, Computer Vision and Pattern Recognition, vol. 1, 2005, pp. 886–893.
[30] G.M. Farinella, S. Battiato, Scene classification in compressed and constrained domain, IET Comput. Vis. 5 (5) (2011) 320–334.
[31] S. Battiato, G.M. Farinella, G. Gallo, D. Ravì, Exploiting textons distributions on spatial hierarchy for scene classification, EURASIP J. Image Video Process. 2010 (1) (2010) 919367.
[32] P. Pérez, M. Gangnet, A. Blake, Poisson image editing, ACM Trans. Graph. 22 (3) (2003) 313–318.
[33] P. Bhat, C.L. Zitnick, M. Cohen, B. Curless, Gradientshop: a gradient-domain optimization framework for image and video filtering, ACM Trans. Graph. 29 (2) (2010).
[34] S. Battiato, G.M. Farinella, G. Puglisi, D. Ravì, Saliency based selection of gradient vector flow paths for content aware image resizing, IEEE Trans. Image Process. 23 (5) (2014) 2081–2095.
[35] G. Messina, S. Battiato, M. Mancuso, A. Buemi, Improving image resolution by adaptive back-projection correction techniques, IEEE Trans. Consum. Electron. 48 (3) (2002) 409–416.
[36] A. Furnari, G.M. Farinella, A.R. Bruna, S. Battiato, Generalized sobel filters for gradient estimation of distorted images, in: IEEE International Conference on Image Processing, 2015, pp. 3250–3254.
[37] K.K. Pingle, Visual perception by a computer, Autom. Interp. Class. Images (1969) 277–284.
[38] P.E. Danielsson, O. Seger, Generalized and separable Sobel operators, Mach. Vis. Three-dimens. Scenes (1990) 347–379.
[39] A. Furnari, G.M. Farinella, G. Puglisi, A.R. Bruna, S. Battiato, Affine region detectors on the fisheye domain, in: IEEE International Conference on Image Processing, 2014, pp. 5681–5685.
[40] D. Comaniciu, V. Ramesh, P. Meer, Kernel-based object tracking, IEEE Trans. Pattern Anal. Mach. Intell. 25 (5) (2003) 564–577.
[41] C. Hughes, E. Jones, M. Glavin, and P. Denny, Validation of polynomial-based equidistance fish-eye models, in: Signals and Systems Conference, 2009.
[42] A. Torralba, A. Oliva, Statistics of natural image categories, Network: Comput. Neural Syst. 14 (3) (2003) 391–412.