Estimating the 3D Direction of a Translating Camera From a Single Motion-Blurred Image

Giacomo Boracchi

Politecnico di Milano, Dipartimento di Elettronica Informazione Via Ponzio, 34/5 20133 Milano, Italy

Abstract

We propose an algorithm for estimating the 3D motion direction of a camera that undergoes a pure translation. This algorithm exploits a single blurred image, and recovers the 3D translation direction thanks to an accurate analysis of the motion blur, which is characterized by rectilinear smears whose directions and extents typically vary throughout the image. The core of our algorithm is the estimation of the direction of these smears within small image regions that are automatically selected according to the image content. The algorithm has been succesfully tested on camera images and extensively validated with different amount of noise in the images.

Key words: Camera Motion Estimation, Motion Blur, Spatially Variant Blur, Blurred Image Analysis, Blur Estimation, Radial Blur
PACS:

1 Introduction

Motion-blurred images embody information about the motion that took place during the exposure. Nevertheless, it is a challenging problem to estimate the camera 3D-motion given a single motion-blurred image whose content is unknown. We present a novel algorithm for estimating the 3D direction of a translating camera by analyzing a single blurred image, acquired during the camera translation. The core idea of our algorithm is to exploit the blur as a cue for estimating the camera ego-motion. Our motivations are similar to those in [19], where an algorithmic gyroscope based on the analysis of a rotationally blurred image is presented. In case of translation, the blur becomes a crucial information for estimating the camera ego-motion as translational motion cannot be sensed by accelerometers, whereas other motions (such as the shake and the rotation) can be handled exploiting measurements from these sensors.

As a matter of fact, images are usually motion blurred when the light is not enough (this often occurs in a dim indoor environment) thus, the estimation of the blur Point-Spread Function (PSF) has been widely studied in the last decades [7,20]. Most of the PSF estimation algorithms assumes spatially-invariant blur: recently Fergus *et al.* [12] proposed an algorithm for enhancing photographs corrupted by camera shake that estimates the PSF exploiting a prior on the distribution of gradient magnitudes in natural images. Levin [21] considers images where PSFs have rectilinear supports and constant direction, and segments these images into regions where the PSFs have the same extent. Jia [17] uses transparency maps for estimating the PSF, while the algorithms in [27,33,37] exploit a short-exposure noisy image, paired with a long-exposure blurred image. All the aforementioned works, like most of the PSF estimation algorithms, focus on the image restoration, although PSF estimation has been addressed for other purposes like the measurement of targets speed [23,24], or planar scene distance [25].

Some of the algorithms that consider spatially-variant blur focus on the estimation of the depth map of a static scene. These algorithms typically exploit either an image sequence [11,16] acquired controlling the camera settings, or a single image [22] acquired with a coded aperture camera. Nagy *et al.* proposed an algorithm [29] for restoring astronomical images corrupted by spatially variant blur. Sorel *et. al* [32] restore a sequence of blurred images, acquired during a camera motion along an arbitrary curve parallel to the image plane, without any rotations.

In general, the blur caused by camera motion is spatially variant: for example the rotational [4,19,30], the radial [5,34], and the angular [18] blur are spatially variant. Also the camera shake results in spatially-variant blur when the depicted scene is not planar [32]. Among the related works, the work of Rekleitis [31] is the closest to the proposed algorithm, as it addresses the optical flow computation from a single spatially-variant blurred image. This task can be considered as equivalent to estimating direction and extent of a spatially-varying PSF. In [31] the blurred image is divided into a tessellation of blocks, and in each block the PSFs parameters are estimated in Fourier domain. However, this tessellation is fixed and, since the PSF parameters are estimated in Fourier domain, the block size has to be significantly larger than the PSF extent. On the contrary in the proposed algorithm the regions where the PSF is estimated are adaptively selected depending on the image content. More importantly, in [31] the camera 3D ego-motion is not recovered, while this is the main goal of our work.

The paper structure reflects the design of the proposed algorithm and is as follows. Initially, we derive the blurred image formation model by analyzing the effect of a 3D translation of a perspective camera during the acquisition. Thus we show that the resulting blur is spatially variant, characterized by rectilinear smears (Section 2). The directions of these smears are determined by the coordinates (on the image plane) of the epipole, i.e. the vanishing point of the camera motion direction, from now on e. The proposed algorithm estimates the smear directions at some automatically selected image regions using two sorts of blur estimators, which are described in Section 3. The coordinates of e are determined from these local estimates using a voting procedure (Section 4). We can then compute the viewing ray through e (we assume the camera is calibrated) and thus, the direction of the 3D translation of the camera. Section 5 presents the experimental validation on a set of camera images, and the performance evaluation when the amount of noise in the image increases.

2 Problem Formulation

Figure 1**a** illustrates the considered image capture scenario. When the shutter opens, the camera viewpoint is in O, the origin of the canonical 3D reference frame $\mathcal{R}$ [1]. During the exposure interval $[0, T]$ the scene is static, while the camera translates at constant speed, until it reaches F when the shutter closes. Our goal is to estimate, by only analyzing the resulting blurred image I, the 3D direction of $\overrightarrow{OF}$. We assume that the camera intrinsic calibration matrix K is known, so that the 3D direction of $\overrightarrow{OF}$ in $\mathcal{R}$ is the direction of $K^{-1}e$, the viewing ray through the epipole e. The core of our algorithm thus consists in the estimation of the coordinates of e, i.e. the vanishing point of the direction of the camera translation direction. In what follows we indicate the camera translation with $\overline{OF}$, as a single blurred image does not allow us to determine whether the translation was $\overrightarrow{OF}$ or $\overrightarrow{FO}$.

2.1 Blurred Image Model

We assume that the camera sensor has linear response, and we represent a motion-blurred image I as the integration of an infinite number of still images I_t, each one captured with the camera viewpoint in a different position in the space. The equation that describes the blurring process, i.e. the blurred image formation, is

$$I(x) = \int_0^T I_t(x)dt + \eta(x), \quad x = (x_{|1}, x_{|2}) \in \mathcal{X}\,, \tag{1}$$

where, x is a pixel location on the 2D image grid $\mathcal{X} \subseteq \mathbb{Z}$, $x_{|1}$ and $x_{|2}$ indicate the projection of x on the axes of the image coordinate system, $I_t : \mathcal{X} \to \mathbb{R}$ represents the light intensity that reaches each pixel at time t, $[0, T]$ is the exposure interval, and η is Gaussian white noise $\eta(x) \sim N(0, \sigma_\eta^2)$.

Whenever there are no occlusions due to camera translation, the blurring process can be also described as applying a blur operator $\mathcal{K}$ on the original (and unknown)

[1] The z axis of $\mathcal{R}$ is aligned with the camera principal axis, and the other two axes are aligned with the image coordinate system.

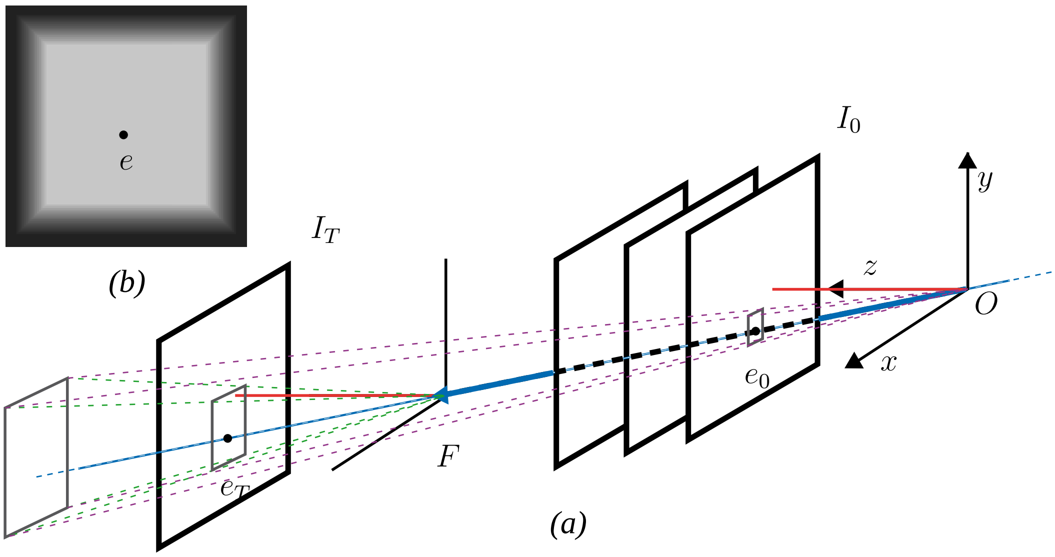

Fig. 1. (a) Camera translation during the exposure. When the shutter opens, the origin O of the $3D$ axis is in the camera viewpoint, with the z axis orthogonal to the image plane. When the shutter closes, the camera viewpoint reaches F; the resulting blurred image is shown in (b).

image I_0. In this case I_0 ideally corresponds, up to a scalar factor, to the image of the same scene, captured from the same camera still in O. Equation (1) becomes

$$I(x) = \mathcal{K}\Big(I_0\Big)(x) + \eta(x)\,. \tag{2}$$

We assume $\mathcal{K}$ is a linear operator, that can be written as [1]

$$\mathcal{K}\Big(I_0\Big)(x) = \int_{\mathcal{X}} k(x,s) I_0(s) ds\,, \tag{3}$$

where $k(x_i, \bullet)$ corresponds to the PSF at pixel x_i. In case of pure camera translation, $k(x_i, \bullet)$ is

$$k(x_i, \bullet) = R_{\theta_i}\Big(\frac{\chi_{[-l_i/2, l_i/2]}}{l_i}\Big)(\bullet - x_i)\,, \tag{4}$$

where $\chi_{[-l_i/2, l_i/2]}$ is the characteristic function of the segment $\{-l_i/2 < x_{|1} < l_i/2,\ x_{|2} = 0\}$ and R_{θ_i} is the rotation of θ_i degrees around the first image axis. Furthermore, the PSF direction θ_i is

$$\tan(\theta_i) = \frac{x_{i|2} - e_{|2}}{x_{i|1} - e_{|1}}, \quad \text{being } e = (e_{|1}, e_{|2})\,. \tag{5}$$

The proof that the PSFs vary as described in Equations (4) and (5) can be easily derived from Equation (1) by means of epipolar geometry[15]. In fact, any couple of images I_{t_1} and I_{t_2}, $t_1, t_2 \in [0, T]$ forms a stereo pair, and thus the corresponding points in these images are related by the epipolar constraints. Let I_0 and I_T be the two images acquired at the initial and at the final camera position, and let the points e_0 and e_T be the images of the line through $\overline{OF}$ in I_0 and I_T, respectively. The points e_0 and e_T correspond to the epipoles of the pair I_0 and I_T. Since the camera undergoes a pure translation, e_0 and e_T have the same coordinates in I_0 and I_T, and thus in the resulting image I they collapse in the epipole e, as illustrated in Figure 1**b**. It follows that all pairs of corresponding points (one in I_0 and the other in I_T) are collinear with e. Therefore, the support of the PSF at x_i is a straight line segment having direction θ_i as shown in Equation (5).

Note that the PSF extents l_i are instead determined by the position of X_i, the scene point that is imaged on x_i. Except few particular cases, e.g. those considered in Section 2.2 or in [5,34], it is not possible to provide a similar description for the PSF extents. In what follows we refer to PSFs of this kind as rectilinear PSFs.
In our model the camera translation is the only cause of blur: defocus, lens aberrations, camera shake, and other blurring factors are not considered.

2.2 Examples of Blur Produced by Camera Translation

The most frequently considered situation is when e belongs to the infinite line of the image plane, and the captured scene is planar and parallel to the image plane. Then all the PSFs have the same direction and extent, and the blur becomes spatially invariant, as shown in Figure 2**a**.

Fig. 2. Images acquired during camera translation. Spatially invariant blur (**a**): the scene is planar, parallel to image plane and $e \to \infty$. Spatially variant blur (**b**): e belongs to the image plane and the scene is planar, the PSF extents are given by Equation (6). Spatially variant blur (**c**): $e \to \infty$ but the scene is not planar, note that PSF directions are constant. Spatially variant blur (**d**), the scene is not planar and e lies on the image plane.

When the camera translation $\overline{OF}$ has a component orthogonal to the image plane, e becomes a point on the image plane (i.e. not at infinity), possibly out of the image grid $\mathcal{X}$. When the scene is planar and parallel to the image plane, the PSF at pixel x_i has extent l_i proportional to $\overline{x_i e}$, the distance between x_i and e. Being $|\overline{OF}|$ the length of the camera translation, and d the distance between the camera viewpoint and the scene plane, we obtain

$$l_i = \overline{x_i e} \frac{|\overline{OF}|}{d + |\overline{OF}|} \,. \tag{6}$$

In this case we refer to radial blur [5,34]: an example of radial-blurred image is shown in Figure 2**b**.
Typically, only the PSFs directions are determined by e, while the extent at x_i

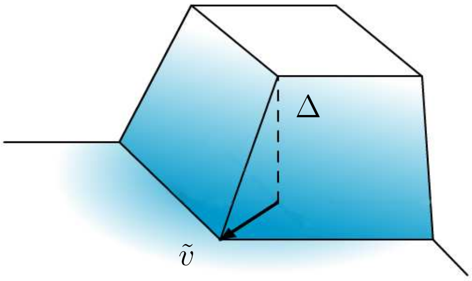

Fig. 3. Image model within a blurred corner region: the displacement vector $\tilde{v}$, and the difference between the corner and the background Δ.

depends on the position of X_i, the corresponding scene point. Thus, even when $e \to \infty$, if the scene is not planar, the blur is spatially variant as shown in Figures 2**c** and 2**d**.

3 Local Blur Analysis

We treat the blur as locally spatially-invariant [1], i.e. we assume that $\forall x_i \in \mathcal{X}$, exists a neighborhood $U_i \subset \mathcal{X}$, of x_i, and a PSF v_i, such that

$$\mathcal{K}\Big(I_0\Big)(x) \approx \int_{\mathcal{X}} v_i(x-s) I_0(s) ds \quad \forall x \in U_i \,. \tag{7}$$

Furthermore, we assume that v_i is a rectilinear PSF having direction θ_i and extent l_i. These assumptions allow us to use methods meant for parametric and spatially invariant blur for estimating the PSF within U_i.
The coordinates of e, and thus the camera 3D translation direction, are obtained by estimating the PSF directions within some automatically selected image regions. For this purpose we exploit two different PSF direction estimation methods, and a procedure for determining the method to be used in each region.

3.1 Local Estimation of PSF Direction

Fourier domain methods, which are widely used for PSF parameters estimation, do not perform adequately on small image regions as they assume periodic signals. Thus we adopt two methods that work in the image domain. Within regions containing an image corner we estimate the PSF direction using the method proposed in [3]. This method estimates the corner displacement vector $\tilde{v}$, which represents the PSF parameters within a image region that contains a corner and that has been blurred with a rectilinear PSF. The method analyzes the image gradient within the

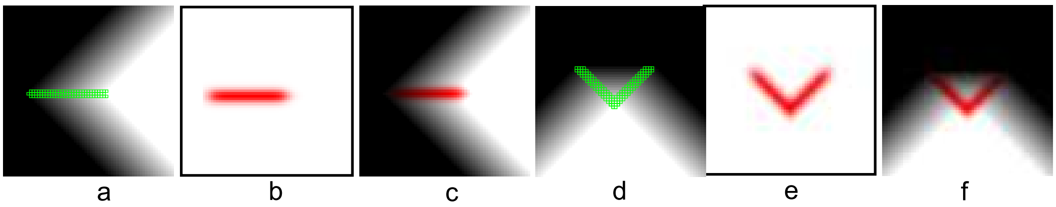

Fig. 4. At corner smears the Harris measure is larger than on blurred edges; examples for two kind of blurred corner regions. Pixels on the corner smears have been manually highlighted (**a** and **d**), the Harris measure in red (**b** and **e**), Harris measure is displayed over the blurred image (**c** and **f**).

blurred-edges and estimates Δ, the intensity gap between the corner and the background, as shown in Figure 3.

Within regions U_i where no blurred edges are present and the image is not flat, the PSF direction is estimated with the method proposed in [36]: the PSF direction θ_i, is given by the direction of the directional derivative filter d_θ having minimum ℓ^1-norm response, i.e.

$$\theta_i = \arg \min_{\theta \in [0, 2\pi]} \Big(\sum_{x \in U_i} |(d_\theta \circledast I)(x)| \Big), \tag{8}$$

where $\circledast$ denotes the 2D convolution.

The regions U_i are selected around some particular pixels: the salient points. Among all the salient points we identify the blurred corners by using the procedure described in the next section.

3.2 Salient Points

We take as salient points the local maxima of the Harris measure [14], which have been used in several feature detection algorithms [26,28]. At pixels having large Harris measure, the Hessian matrix of the sum of square differences has two large eigenvalues and vice versa [14]. Therefore the image in a patch of these pixels is significantly different from any neighboring patch.

Within a blurred corner region, the Harris measure is larger on the corner smears than on the blurred edges. As illustrated in Figure 4, the corner smear is a set of pixels between two blurred edges. Near the corner smears, the image is different w.r.t. any neighboring patch, and thus the Harris matrix has two nonzero eigenvalues. On the contrary, into the blurred edges the Harris measure is negative or zero, as the derivative along the edge direction is zero and one eigenvalue is zero. It follows that, provided that in the original image the corner is binary like those of Figure 4**a**, each corner smear presents at least a salient point. This could be at any pixel on the corner smear; however this is enough for initializing the adaptive corner-region

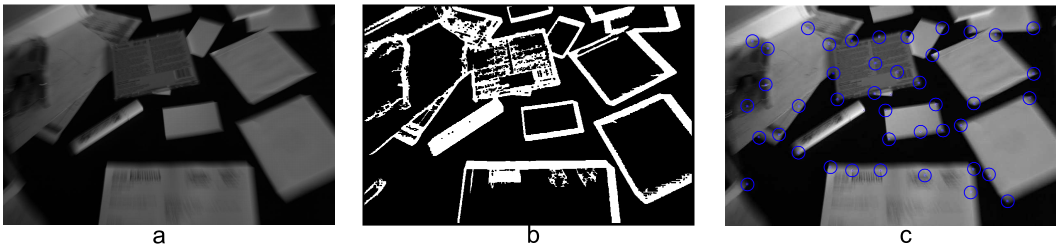

Fig. 5. Blurred corner detection: **a** a blurred image, **b** the mask Γ used to identify blurred edges, **c** blurred-corner candidates.

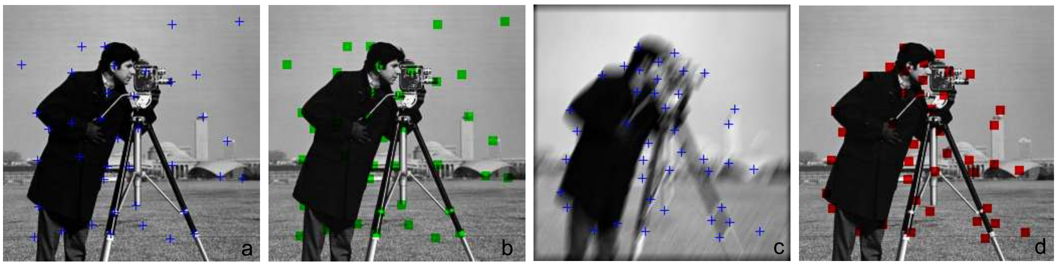

Fig. 6. Salient points in blurred images belong to areas having large Harris measure in the original image: **a** salient points in a test image and, **c** salient points in f_j, a blurred image having PSF direction $60°$ and extent 20 pixels. The average Harris measure of the original image within the red squares **d**, divided by the average Harris measure in the whole original image is $m_j = 1.98$. Near the salient point of the original image, the green squares of **b**, this ratio becomes 2.75.

selection procedure described in Section 3.3.

We consider as blurred-corner candidates those salient points belonging to blurred edges. The blurred edges are identified by the mask Γ,

$$\Gamma = \{x \; s.t. \; ||\nabla I(x)|| > T\}, \tag{9}$$

where T is a threshold parameter, tuned on the minimum acceptable slope for a blurred edge. We also post-process Γ with ordinary morphological operators [13] both to remove isolated points, small areas, thin lines and to widen larger areas. Let the blurred-corners candidates (the salient points in Γ) be $\{x_i\}, \; i = 1, ..., m$: around these points we run the adaptive corner-region selection procedure of Section 3.3, and we discard those that have too small corner area. The remaining $\{x_i\}, \; i = 1, ..., M \;\; M \leq m$, represent the blurred corners and the PSF directions $\{\theta_i\}, \; i = 1, ..., M$ are estimated using [3] within the adaptively selected corner region. Figure 5**a** shows a camera image used as running example, Figure 5**b** shows the corresponding Γ, and Figure 5**c** the blurred corner candidates. Around the salient points that have not been identified as blurred corners (because they do not belong to Γ or because they do not have a corner region large enough) we select circular regions U_i. Let $\{x_i\}, \; i = M, ..., N$ be the remaining salient points having Harris measure over a fixed threshold: we estimate the PSF

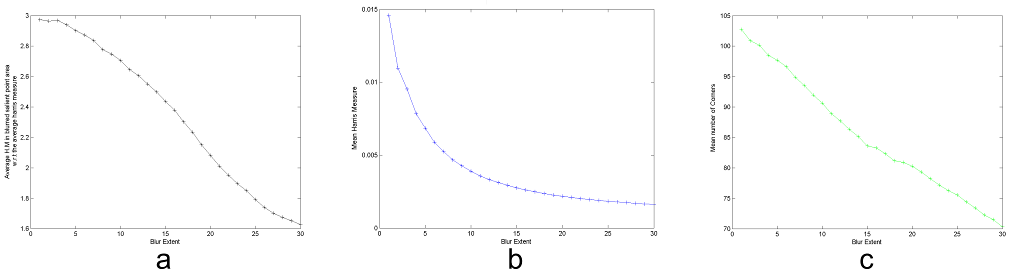

Fig. 7. Salient Points on blurred images: **a** the values of m_j as a function of the PSF extent, **b** the mean Harris measure as a function of blur extent, **c** the mean number of salient point as a function of blur extent.

directions $\{\theta_i\},\ i = M, ..., N$ using Equation (8) within the corresponding circular neighborhoods $U_i,\ i = M, ..., N$.

Equation (8) returns reliable estimates within regions where the original image I_0 has similar, and non zero, ℓ^1-norm response to any directional derivative filter. The original image I_0 typically shows significant variations along any direction within regions centered in a salient point. Therefore Equation (8) takes as the PSF direction, the direction presenting less variations within these regions in the blurred image I. However it is not guaranteed that a salient point of the blurred image I corresponds to a salient point of I_0, the original image.
The following experiment shows that salient points in blurred images typically belong to areas where the original image has a large Harris measure. We consider a dataset of 12 common (256×256 and 512×512) grayscale test images, rescaled in $[0, 1]$, and two sets $\Theta = \{0, 10, \cdots, 170, 180\}$ and $L = \{1, 2, \cdots, 29, 30\}$ for the PSF directions and extents, respectively. We synthetically blur each test image with a convolution against each PSF generated from the parameter pairs in $\Theta \times L$, obtaining $\{f_j\}\ j = 1, ..., 12 \cdot \#(\Theta \times L)$ blurred images. The Harris measure is thresholded against $\tau = 0.0005$ in order to remove low-relevance salient points; we then extract the salient points in each blurred image f_j and we crop, from the corresponding original image, a square of 10 pixels side centered in the salient point. We compute m_j as the average of the Harris measure of the original image within these squares, divided by the average Harris measure of the whole original image, as described in Figure 6. Figure 7**a** shows m_j averaged between images having the same PSF extent: the Harris measure within these squares decreases as the blur extent increases but, even in heavily blurred images, the salient points belong to areas where the original image still presents significant variations. In fact, the average of m_j in the original image (PSF extent 1 pixel) is about 3, and it remains 1.6 in images blurred with a 30 pixels extent PSF. Figures 7**b** and 7**c** show how the mean of the nonzero Harris measure and the number of salient point decrease with the blur extent. Note that m_j is correctly averaged among images having PSF with different direction, as the Harris measure is rotational invariant [14].
This experiment consides spatially invariant blur but, since the Harris measure is a local measure, a similar result holds for the spatially variant blur (although in this

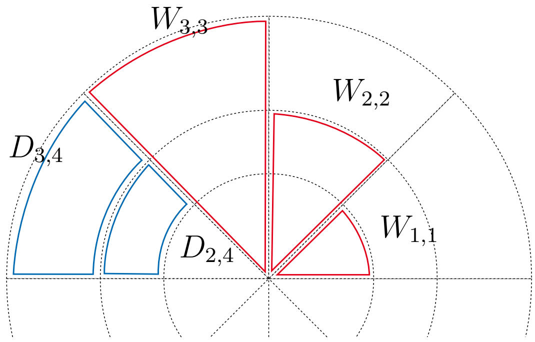

Fig. 8. Shapes of the supports of wedge masks $W_{j,\alpha}$ (red solid line) and their differences $D_{j,\alpha}$ (blue dotted line).

case the salient points tend to concentrate in the least blurred areas).

3.3 Adaptive Corner Region Selection

The corner region is constructed exploiting the fact that a convolution against a rectilinear PSF produces areas (the blurred edges) where the gradient is constant. The gradient uniformity is analyzed using wedge-shaped binary masks $W_{j,\alpha}$, where $j \in \{j_0, ..., J\}$ represents the wedge size and $\alpha \in \{2i\pi/A\}_{i=0,...,A}$ the wedge direction. As illustrated in Figure 8, all the wedges have the vertex in the blurred corner, and wedges along the same direction are nested, i.e. $W_{j,\alpha} \subset W_{j+1,\alpha}, \forall \alpha, \forall j = 1, ..., J-1$. Roughly speaking, we compare the average gradient in each wedge $W_{i,\alpha}$, with the average gradient in the area in between the next nested wedge $D_{j+1,\alpha} = W_{j+1,\alpha} - W_{j,\alpha}$, in order to prevent discontinuities in the region. The corner region is built repeating the following iterative procedure along each direction.

step1 Compute $w_{j,\alpha} = \sum_{x \in W_{j,\alpha}} \nabla I(x)/\#W_{j,\alpha}$, the average gradient in $W_{j,\alpha}$; where $\#W_{j,\alpha}$ denotes the number of elements in $W_{j,\alpha}$.
step2 Compute $d_{j+1,\alpha} = \sum_{x \in D_{j+1,\alpha}} \nabla I(x)/\#D_{j+1,\alpha}$, the average of ∇I on $D_{j+1,\alpha}$.
step3 If ($|w_j - d_{j+1}| > M_1\sigma_\eta$) or ($|d_{j+1}| < M_2\sigma_\eta$), take $\bar{j}_\alpha = j$ and repeat **step1** from $\alpha + 1, j = 3$.
step4 if $j \neq J$, repeat **step1** from $j+1, \alpha$; otherwise $\bar{j}_\alpha = J$, and repeat **step1** from $\alpha + 1, j = 3$.

After having considered all directions in A, we take $U_i = \bigcup_{\alpha \in A} W_{\bar{j}_\alpha,\alpha}$, and we mask U_i with Γ.

This procedure builds regions that are star-shaped w.r.t. the salient point, although it may happen that the blurred corner region is not star-shaped (see the blurred corner in Figure 4**d**). To provide non star-shaped regions, we repeat this procedure in 9 neighboring pixels of each blurred corner, and then we take as U_i those pixels that have been selected at least twice. Note that the procedure does not consider

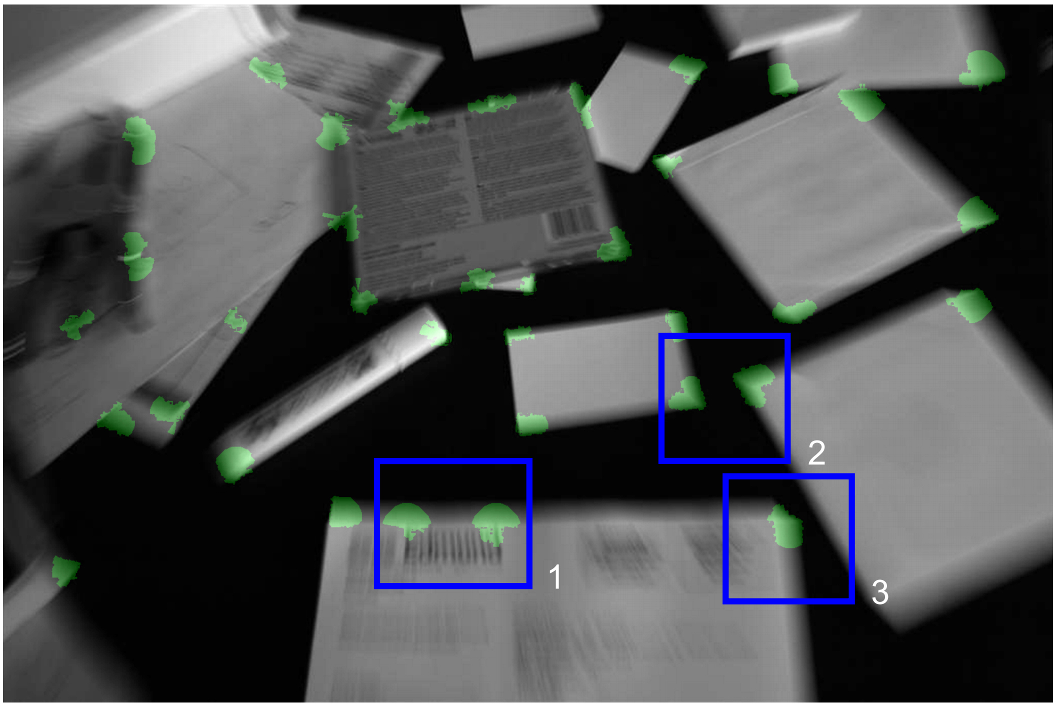

Fig. 9. Selected corner regions where the estimated corner displacement vector has acceptable length.

the smallest wedges in each direction, since on the corner smears the gradient is discontinuous. The parameters $M_1, M_2 > 0$ are fixed, while σ_η is estimated using [9].

3.4 Remarks

The corner-region selection procedure allows us to separately estimate the PSF direction in those areas containing two neighboring corners, see Figure 9 box **2**. Moreover, the selected regions exclude both possible details and blurred edges that do not belong to the corner, since these ones could bias the PSF direction estimation (Figure 9 box **3**). However, textured areas where the average gradient is constant may be erroneously selected (e,g. Figure 9 box **1**). Sometimes spots near a blurred edge are erroneously taken as a salient points, and the blurred edge is identified as corner region. Often, in this case, the estimated PSF has unacceptable extent (e.g. larger than the maximum region size), and thus these estimates are not considered for estimating the epipole. Figure 9 shows only corner regions where the estimated displacement vector has acceptable length. Finally, although the procedure is repeated to construct non star-shaped regions, corner regions like that one in box **3** of Figure 9 are sometimes partially missed.

4 Estimation of the Epipole

The coordinates of e are determined by fusing all the local estimates. From the experimental evidence it emerges that, because of noise or image details that influence the PSF direction estimates, this fusing needs to be done by using a robust procedure. In [19] this issue is addressed by the RANSAC algorithm, while here we solve the equivalent point-fitting problem using an Hough approach. In this case the parameter space, i.e. the space containing all the admissible solutions (the location of e), is the whole image plane, including the line at infinity. For each data $(\theta_i, x_i),\ i = 1, ..., N$, we assign one vote in the parameter space to each parameter (i.e. point of the parameter space) belonging to the straight line having direction θ_i and passing through x_i. After having considered all the data, the parameter that received the largest number of votes is taken as e.

This approach allows us to design an ad-hoc weight function ℓ, which can be used instead of straight line for considering errors in the estimation of both the PSF direction and the corner location. The idea is to assign a full vote to parameters on the aforementioned straight line and a lower vote to nearby parameters. As in [4], the weight function ℓ_{x_i,θ_i} associated to the i-th data is obtained rotating of θ_i degrees and centering in x_i the function

$$\ell(x_{|1}, x_{|2}) = exp\Big[-\Big(\frac{x_{|2}}{(1 + h|x_{|1}|)k}\Big)^2\Big]. \tag{10}$$

The two parameters $h, k > 0$ determine the vote spread from the exact solutions and the localization error, respectively.

As typical in Hough approaches, the votes are assigned in a discrete and finite parameter space, which in our experiments is a grid three times larger than the image. We used $k = \sigma_\eta$ for the estimates coming from Equation (8), and $k = k_0 + \sigma_\eta$ for the estimates at corners, with $k_0 > 0$ a fixed tuning parameter. Figure 10 shows the parameter space in the area of the image grid; the estimated epipole is illustrated in Figure 11.

5 Experiments

The proposed algorithm has been tested on images acquired with a Canon EOS 400D 10-Mpixel camera in two different scenarios. All these images have been acquired with small aperture to reduce the out-of-focus blur, and have been converted in grayscale (in the range [0-255]) and downsampled to half-size, before being processed. In the first scenario, two triplets of images (shown in Figure 12) have been acquired in a controlled environment, mounting the calibrated camera on a 2 DOF planar robot, developed at the AIRLab of Politecnico di Milano [2].

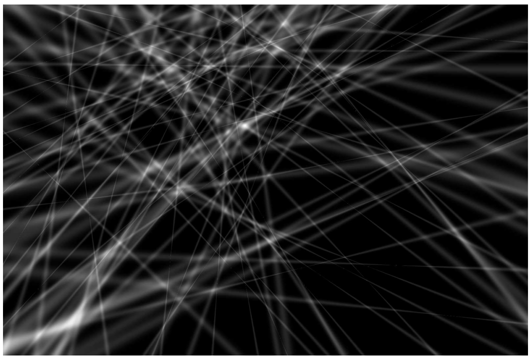

Fig. 10. The region of the parameter space corresponding to the image grid, and the votes corresponding to all the estimates shown in Figure 5.

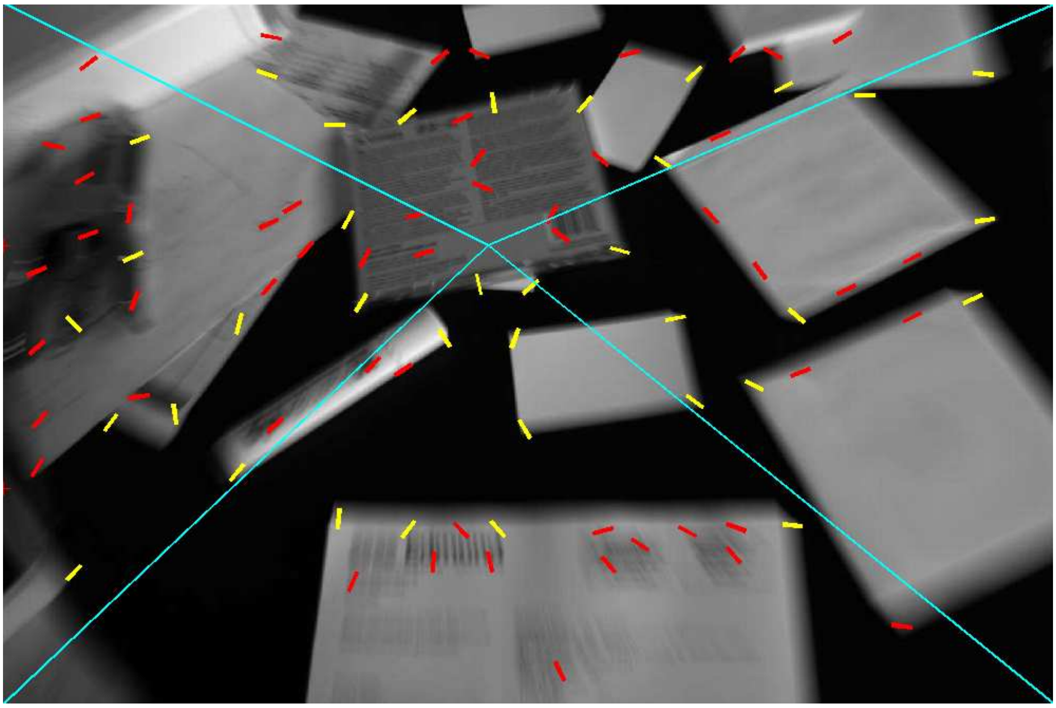

Fig. 11. The estimated epipole corresponds to the intersection of the cyan lines. Fixed-length segments indicate PSF direction estimates: in yellow using [3], in red using Equation (8).

The robot motors are controlled with a proportional integral derivative controller that ensures pure translation at controlled speed, and thus allows us to compare the estimated 3D direction with the ground truth. The ground truth of the 3D motion direction has been computed from each image by appropriately placing a square marker, so that its edges perfectly line up with the camera motion direction; then, we manually select these edges (blue dashed lines in Figure 12) and we compute the viewing ray through their intersection (blue solid lines Figure 12). The estimation error is measured as the amplitude of the angle between the estimated 3D translation direction and the ground truth. The first images of each triplet (Figure 12**a,d**) have been acquired with translation speed v_0, the second ones (Figure 12**b,e**)

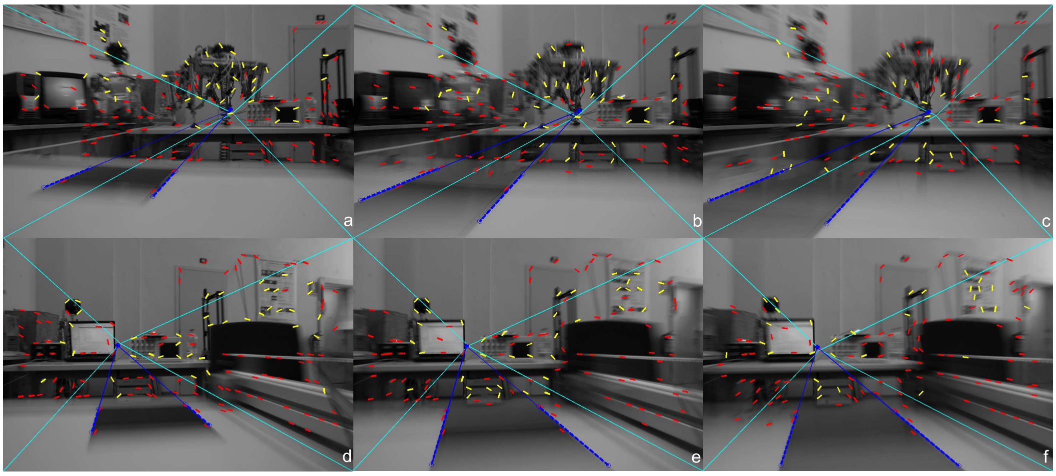

Fig. 12. Experiments on images acquired in an automatically controlled environment. The camera tranlation speed doubles in each column. The blue dashed lines represent the manually selected borders of the square marker, the blue solid ones the fitted lines, and their intersection is the vanishing point of the ground truth. The cyan lines connect the estimated epipole e with the image borders, and the 30 pixels segments are the PSF direction estimates: yellow for estimates at blurred corners [3], and red using Equation (8). The angular errors are: **a** 0.56 ,**b** 0.84 ,**c** 1.15**d** 0.45 ,**e** 0.21 ,**f** 0.52.

with $2v_0$ and the third ones (Figure 12**c,f**) with $4v_0$. The 3D translation direction, all the camera camera settings, as well as the opening position of the camera shutter, were the same in all these images [2], while the depicted scene and location of e has been changed between the two image triplets. In each image of Figure 12 (like in all the other images of Figures 14 - 19), the cyan lines connect the estimated epipole e with the image borders, while the 30 pixels segments within the image show the PSF direction estimates: yellow for estimates at blurred corners (obtained using method [3]) and red for estimates given by Equation (8).

We assume these images as noise free and, to test the algorithm on noisy images, we add a noise term η with standard deviation values $\sigma_\eta = 0, 1, \cdots, 5$. Figure 13 shows how the angular error varies with the noise standard deviation; the results have been averaged over 20 different noise realizations.

In the second scenario we acquired images by translating the camera both manually (Figures 5 and 14), and on a wheeled device (Figures 15 - 18). Even if in this case the ground truth was not available, the cyan lines and the enlarged sections show that the blur directions are correctly represented by the line passing through the estimated epipole. Figures 15- 18**b** shows the selected regions around the detected blurred corners. A squared box of 100×100 pixels around some blurred corner has been zoomed to prove the effectiveness of the adaptive region selection procedure. Figure 19 presents some results in presence of noise with $\sigma_\eta = 5$.

[2] We controlled the camera shutter with a mechanical trigger so that the acquisition started in a fixed position, where the robot speed was stabilized.

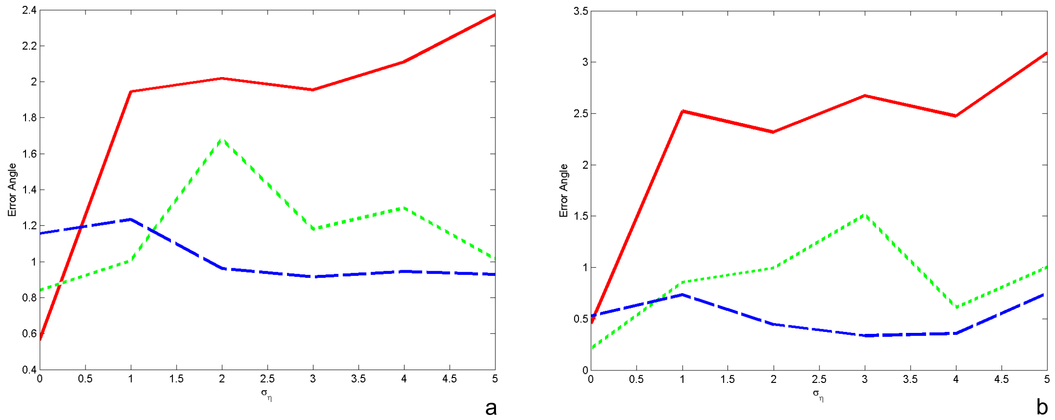

Fig. 13. Angular errors (in $[0, 90]$) as a function of σ_η: **a** triplet of the first row of Figure 12, **b** triplet of the second row. The plots report the average over 20 different noise realizations. The solid red lines correspond to images acquired with translation speed v_0, the dotted green lines with speed $2v_0$, and the dashed blue lines with speed $4v_0$.

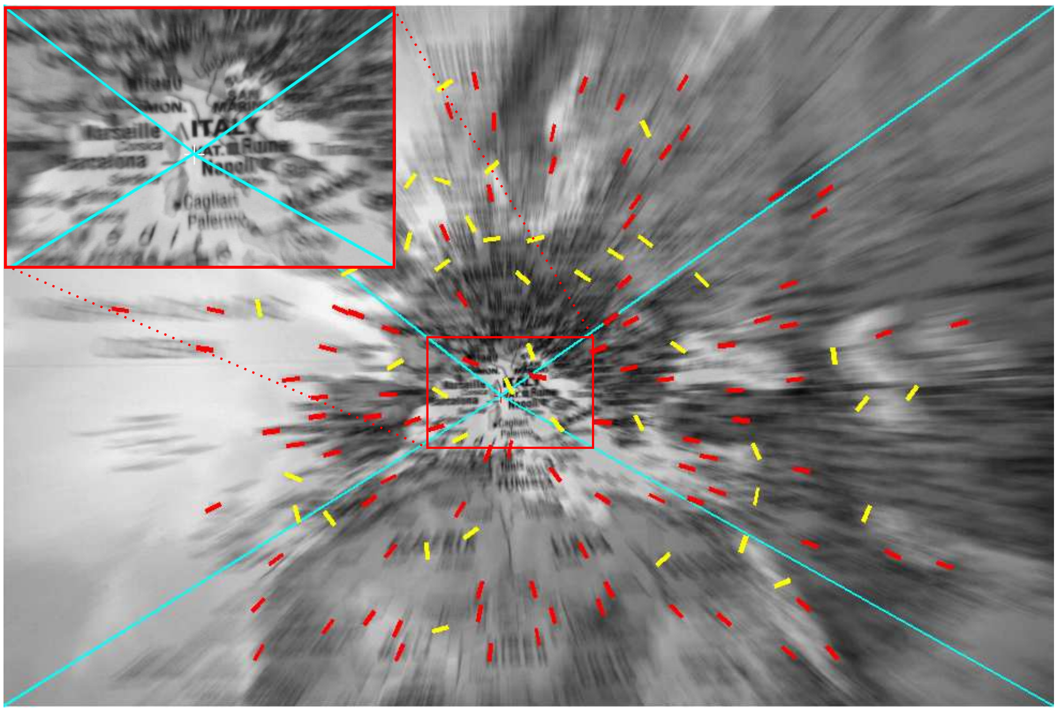

Fig. 14. A radial blurred image: although there are no corners and this produces several outliers, the blur is correctly interpreted as illustrated in the detail. The enlarged section shows that the estimated epipole is correctly in a low-blur area.

The wedge masks used for corner selection have maximum size of 50 pixels, the circular neighborhoods used for estimating PSF direction using Equation (8) have 25-pixels radius. Directional derivatives have been computed using 7-tap filters [10], and the camera has been calibrated by using the Matlab toolbox [6]. In order to exclude low-significance features when the noise increases, we fixed a maximum number of local estimates: we sort the salient point depending on their Harris measure, and we take only the 40 most significant blurred corners and the 100 most significant salient points where PSF direction is estimated as in Equation (8).

5.1 Discussion

While it is clear that the noise reduces the PSF direction estimation accuracy, there is no straightforward relation between the noise amount and the epipole localization error. Plots of Figure 13 and images of Figure 19 show that the noise is not the only factor determining the algorithm accuracy, and that also the following ones have to be considered:

- *Blur*: the plots of Figure 13 show that the algorithm can better cope with noise in heavily blurred images. In fact, when the scene and the camera settings (and thus the exposure) are fixed, we obtain better results at higher camera tranlsation speeds. The plots of Figure 13 and a comparison of the results in Figures 15 and 19**d** show that low-blur areas (which are determined by both the camera displacement and the scene depth) make the algorithm more sensitive to noise.
- *Salient points*: the number of salient points and the way how these are distributed in the image may significantly influence the epipole estimation. Salient points spread in all part of the image usually provide better results. The good performance on noisy images shown in Figures 16 and 19**c** and in plots of Figure 13**a**, is justified by the salient point distribution.
- *Scene content*: when a salient point is located near an edge or a line, Equation (8) typically returns the edge or the line direction. This may produce outliers (as in Figures 15 and 17) or correct estimates, in case the line is pointing to the epipole (like the robot structure on the right side of the images of the second row in Figure 12). Note that it is not unusual that blurred images contain lines or edges that are pointing to the epipole, as these lines are blurred along themselves, and thus may survives even heavy blur. We also experienced that the noise increases the number of salient points by typically introducing new salient points near edges or lines. These two factors motivate the high noise robustness in the images in the second row of Figures 12, as well the corrupted estimates in the images of Figure 19**e,f**.

The considered PSF estimation methods are robust to additive white Gaussian noise: the results of Equation (8) are not influenced by AWGN, and the corner method provided satisfying performance in presence of AWGN [3]. However, it may happen that also PSF directions correctly estimated at blurred corners may be far from pointing to e, because of occlusions and shadows. In fact the blur produced by occlusions during the translation does not satisfy the epipolar constraints, while shadows may be erroneously interpreted as blurred corner edges. An example of blur produced by occlusion is shown in Figure 18 (box **1**), while shadows that have been considered as blurred corner regions are shown in Figure 17**b**. Note that

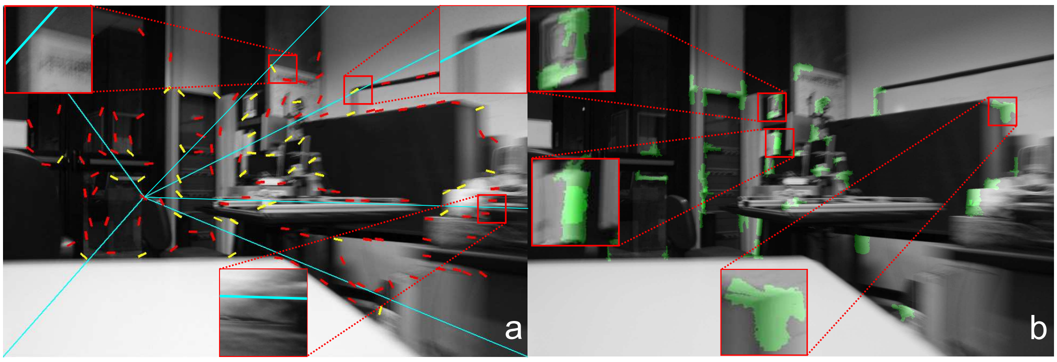

Fig. 15. The PSFs in the left part of the image have short extents as the scene contains far away objects: the PSF directions are not reliably estimated in this area. The corner region selection procedure allows us to estimate separately the PSF direction in two nearby corners, as shown in two enlarged sections.

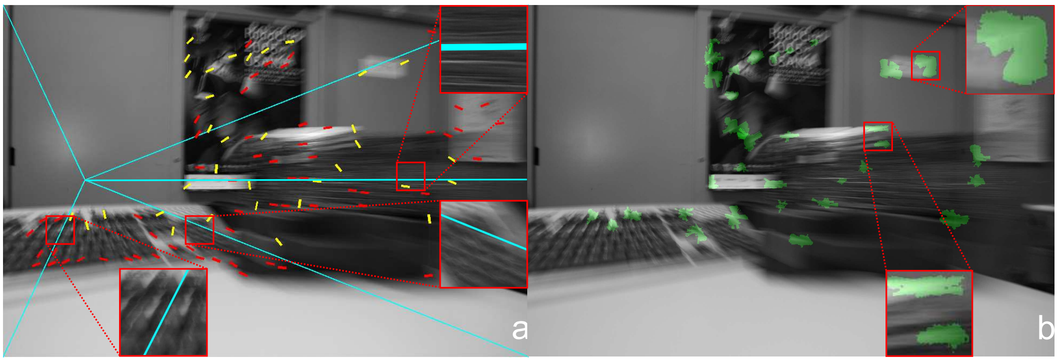

Fig. 16. This image contains two highly textured objects on the foreground that are heavily blurred, and that allow a reliable PSF direction estimation. The epipole estimates in this image are accurate even in presence of noise (see Figure 19**c**) as the salient points are spread enough.

in all our experiments the PSF directions have been estimated on small regions, where the blur can be reasonably approximated as spatially invariant: each corner region highlighted in Figures 15**b** - 18**b** is included in a 100×100 pixels square. Figure 14 shows the effectiveness of the voting algorithm: even if there are no corners and all the 40 estimates coming from [3] are erroneous, the voting procedure is able to cope with such outliers with 100 inliers.

The overall computation time depends on the number of regions where PSF direction is estimated: this is the computationally heaviest part. The corner-region selection procedure is based on local averages and comparisons, therefore its computation cost is linear w.r.t. the number of pixel in the region, like the PSF estimation at corners. Finally the directional derivatives of Equation (8) are computed by using separable filters [10], however the ℓ^1-norm minimization can be sped up with a multiscale implementation.

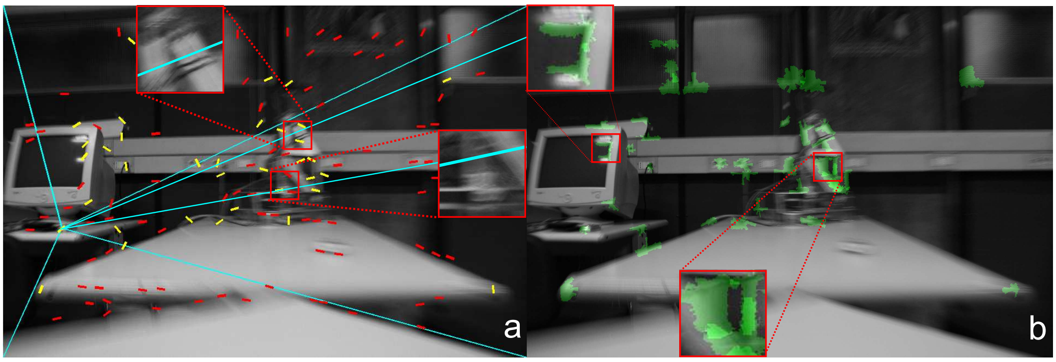

Fig. 17. Several PSF directions estimates have been influenced by the horizontal edges in the central part of the image, resulting in a less accurate estimates of the epipole. The leftmost enlargement of **b** shows how shadows could be erroneusly considered as blurred edges. Note that when noise is introduced, the number of salient points along these lines increases (compare with Figure 19**f**).

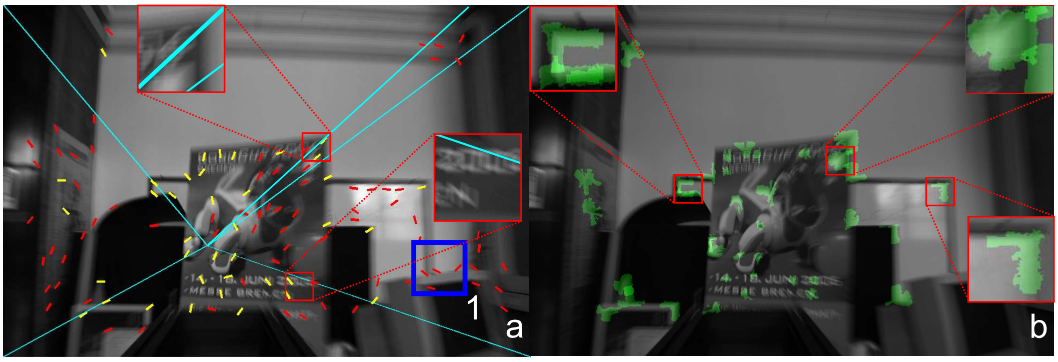

Fig. 18. An example of blur produced by occlusion (box **1**).

6 Conclusions

In a blurred image acquired from a translating camera, the PSF directions and extents are varying through the image pixels according to the camera 3D motion and the scene depth. In this work we devise an image formation model for these images, and we present a single-image algorithm for estimating the camera 3D translation direction, assuming that the camera is calibrated. The algorithm relies on the estimation of the PSF direction within small image regions that are automatically selected according to the image content.

The algorithm can be included in robot vision systems based on frame analysis since these systems (e.g. [8]) have often to handle blurred frames at reduced lightning conditions in indoor environments. Instead of discarding blurred frames, where it is not possible to match map features, the blur can be thus exploited for estimating the camera ego-motion. Moreover, the blur carries information about the *intra*-frame motion, i.e. the motion that the camera undergoes during the acquisition, which in long exposure frames is more meaningful than the *inter*-frame

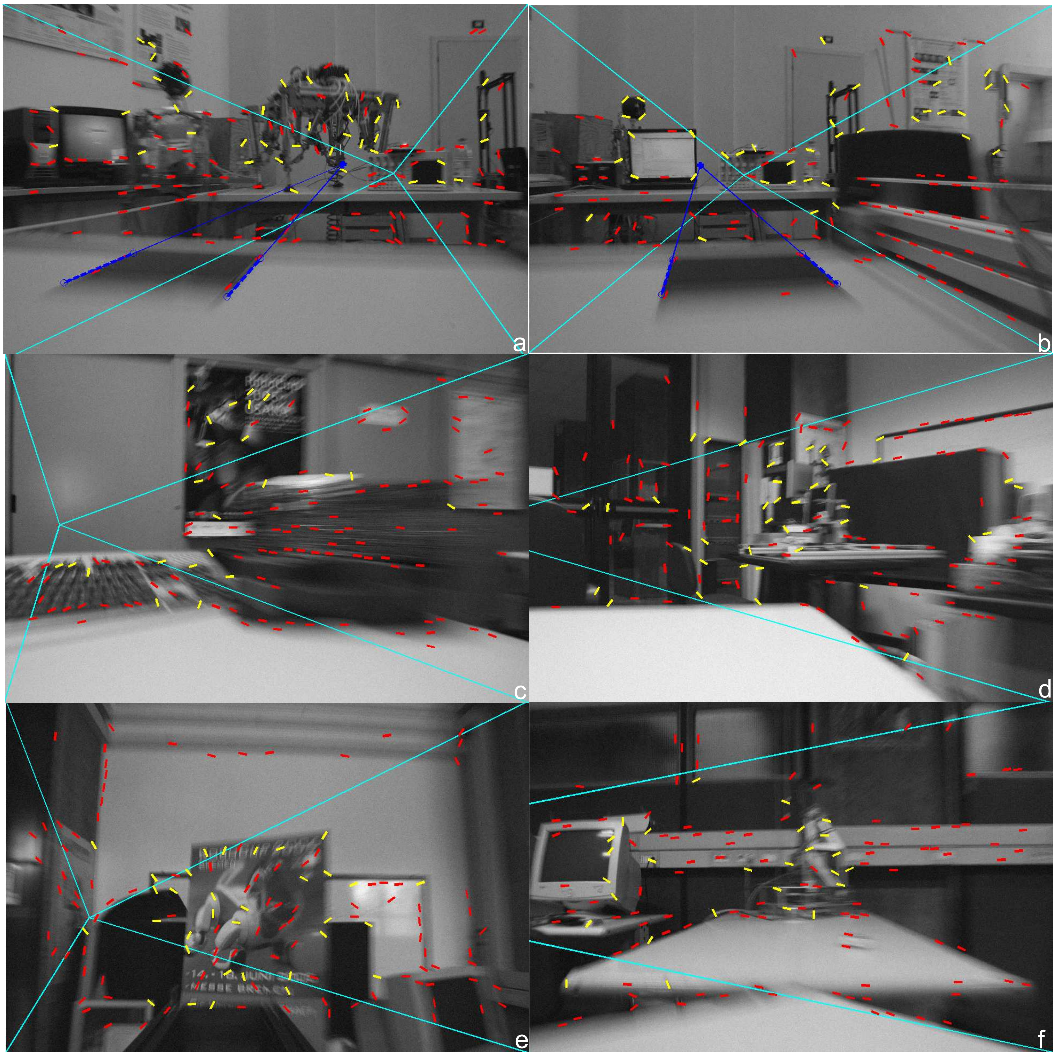

Fig. 19. Algorithm performance in presence of noise ($\sigma_\eta = 5$). In general, the noise makes the PSF direction estimation less accurate, however this impacts differently on the depicted images: **a,c** the algorithm results are close to those on the respective noise-free images as there are several salient point distributed in different image areas; **b** the image structures help the epipole estimation, several salient points have been detected on the robot structure which is parallel to camera translation; **d** the low-blur area in the left part of the image does not allow accurate estimates of e when the image is noisy; **e,f** because of noise several salient points are determined near some image structures that are not pointing at the epipole.

motion, which is typically estimated from a video sequence. The algorithm can be also used to initialize blind-deblurring methods that consider spatially variant blur such as [35], and for estimating the radial blur center [5,34] given a single radial-blurred image.

Here we considered pure camera translation, as the estimation of the 3D motion direction does not require to know the local motion orientation. Future works concern the extension of this approach to analyze the blur produced by other rigid motions. We are also studying robust methods for estimating the PSF within small image regions, in order to extract additional information from a single blurred image. For example, the PSF extents in the considered case allow us to compute the scene depth map, exploiting the relation between the PSF extent in each pixel and the depth of the corresponding scene point. Reliable estimates of the PSF extents allow also to estimate all the radial blur parameters [5,34], and thus to implement a blind radial-deblurring algorithm.

7 Acknowledgments

The author would like to thank Professor Vincenzo Caglioti for the helpful suggestions, and Eng. Simone Tognetti for the image acquisition. Moreover, the author would like to express a particular acknowledgment both to the reviewers and the editor for their efforts in steering this paper towards this much improved form.

References

[1] M. Bertero, P. Boccacci, Introduction to Inverse Problems in Imaging, Institute of Physics Publishing, 1998.

[2] A. Bonarini, L. Mainardi, M. Matteucci, S. Tognetti, R. Colombo, Stress recognition in a robotic rehabilitation task, in: Proc. of ”Robotic Helpers: User Interaction, Interfaces and Companions in Assistive and Therapy Robotics”, a Workshop at ACM/IEEE HRI 2008, vol. 1, University of Hertfordshire, Amsterdam, the Netherlands, 2008.

[3] G. Boracchi, V. Caglioti, Corner displacement from motion blur, in: Inproceedings of ICIAP 2007,Modena, 2007.

[4] G. Boracchi, V. Caglioti, A. Danese, Estimating camera rotation parameters from a blurred image, in: Inproceedings of VISAPP 2008 Conference, Funchal Madeira, 2008.

[5] G. Boracchi, A. Foi, V. Katkovnik, K. Egiazarian, Deblurring noisy radial-blurred images: spatially-adaptive filtering approach, in: Electronic Imaging, Science and Technology, 2731 January 2008, San Jose, California, USA, 2008.

[6] J.-Y. Bouguet, Camera calibration toolbox for Matlab.
URL `http://www.vision.caltech.edu/bouguetj/calib_doc/`

[7] T. Cannon, Blind deconvolution of spatially invariant image blurs with phase, ASSP 24 (1) (1975) 58–63.

[8] A. Davison, Real-time simultaneous localisation and mapping with a single camera (2003).
URL `citeseer.ist.psu.edu/article/davison03realtime.html`

[9] D. L. Donoho, I. M. Johnstone, Ideal spatial adaptation by wavelet shrinkage, Biometrika 81 (3) (1994) 425–455.
URL citeseer.ist.psu.edu/donoho93ideal.html

[10] H. Farid, E. Simoncelli, Differentiation of discrete multi-dimensional signals, IEEE Transactions on Image Processing 13 (4) (2004) 496–508.
URL www.cs.dartmouth.edu/farid/publications/ip04.html

[11] P. Favaro, Shape from focus and defocus: Convexity, quasiconvexity and defocus-invariant textures, in: Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on, 2007.

[12] R. Fergus, B. Singh, A. Hertzmann, S. T. Roweis, W. T. Freeman, Removing camera shake from a single photograph, ACM Trans. Graph. 25 (3) (2006) 787–794.

[13] R. C. Gonzalez, R. E. Woods, Digital Image Processing, Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 2001.

[14] C. Harris, M. Stephens, A combined corner and edge detector, in: Proceedings of the 4th Alvey Vision Conference,, 1988.

[15] R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd ed., Cambridge University Press, ISBN: 0521540518, 2004.

[16] S. W. Hasinoff, K. N. Kutulakos, Confocal stereo, in: Proc. of European Conference on Computer Vision (ECCV), 2006.

[17] J. Jia, Single image motion deblurring using transparency, in: Inproceedings of CVPR 2007, Minneapolis, 2007.

[18] I. Klapp, Y. Yitzhaky, Angular motion point spread function model considering aberrations and defocus effects, Journal of the Optical Society of America A 23 (2006) 1856–1864.

[19] G. Klein, T. Drummond, A single-frame visual gyroscope, in: Proc. British Machine Vision Conference (BMVC'05), vol. 2, BMVA, Oxford, 2005.

[20] D. Kundur, D. Hatzinakos, Blind image deconvolution, SPMag 13 (3) (1996) 43–64.

[21] A. Levin, Blind motion deblurring using image statistics, in: B. Schölkopf, J. Platt, T. Hoffman (eds.), Advances in Neural Information Processing Systems 19, MIT Press, Cambridge, MA, 2007.
URL
http://people.csail.mit.edu/alevin/papers/levin-deblurring-nips06.pdf

[22] A. Levin, R. Fergus, F. Durand, W. Freeman, Image and depth from a conventional camera with a coded aperture, ACM Transactions on Graphics, SIGGRAPH 2007 Conference Proceedings, San Diego, CA.

[23] H.-Y. Lin, Vehicle speed detection and identification from a single motion blurred image, wacv-motion 01 (2005) 461–467.

[24] H.-Y. Lin, C.-H. Chang, Automatic speed measurements of spherical objects using an off-the-shelf digital camera, in: Mechatronics, 2005. ICM '05. IEEE International Conference on, 2005.
URL http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1529229

[25] H.-Y. Lin, C.-H. Chang, Depth recovery from motion blurred images, icpr 1 (2006) 135–138.

[26] D. Lowe, Distinctive Image Features from Scale-Invariant Keypoints, International Journal of Computer Vision 60 (2) (2004) 91–110.

[27] M. V. Marius Tico, Estimation of motion blur point spread function from differently exposed image frames, in: Proceedings of Eusipco 2006, 4-6 September 2006 Florence, Italy, 2006.

[28] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, L. V. Gool, A comparison of affine region detectors, International Journal of Computer Vision 65 (1/2) (2005) 43–72.
URL http://lear.inrialpes.fr/pubs/2005/MTSZMSKG05

[29] J. G. Nagy, D. P. O'Leary, Restoring images degraded by spatially variant blur, SIAM Journal on Scientific Computing 19 (4) (1998) 1063–1082.
URL citeseer.ist.psu.edu/article/nagy96restoring.html

[30] J. J. Qi Shan, Wei Xiong, Rotational motion deblurring of a rigid object from a single image, in: International Conference on Computer Vision (ICCV), Rio de Janeiro, Brazil, 2007.

[31] I. M. Rekleitis, Steerable filters and cepstral analysis for optical flow calculation from a single blurred image, in: Vision Interface, Toronto, 1996.
URL http://www.cim.mcgill.ca/ yiannis/Publications/index.html

[32] M. Sorel, J. Flusser, Space-variant restoration of images degraded by camera motion blur, IEEE Transactions on Image Processing 17 (2) (2008) 105–116.
URL http://dx.doi.org/10.1109/TIP.2007.912928

[33] M. Tico, M. Vehvilainen, Bayesian estimation of motion blur point spread function from differently exposed image frames, in: In Proceedings of 14th European Signal Processing Conference (EUSIPCO), Florence, Italy, September 4-8, 2006, 2006.

[34] C. B. Webster, S. Reeves, Radial deblurring with ffts, in: Proceedings of ICIP International Conference of Image Processing, Sant Antonio, USA, 2007.

[35] M. Welk, D. Theis, J. Weickert, Variational deblurring of images with uncertain and spatially variant blurs., in: DAGM-Symposium, 2005.
URL http://www.mia.uni-saarland.de/Publications/welk-dagm05.pdf

[36] Y. Yitzhaky, N. S. Kopeika, Identification of blur parameters from motion blurred images, Graph. Models Image Process. 59 (5) (1997) 310–320.

[37] L. Yuan, J. Sun, L. Quan, H.-Y. Shum, Image deblurring with blurred/noisy image pairs, ACM Trans. Graph. 26 (3) (2007) 1.