GENERALIZED SOBEL FILTERS FOR GRADIENT ESTIMATION OF DISTORTED IMAGES

Antonino Furnari[⋆]*, Giovanni M. Farinella*[⋆]*, Arcangelo R. Bruna*[†]*, Sebastiano Battiato*[⋆]

[⋆]Department of Mathematics and Computer Science - University of Catania
[†]Advanced System Technology - Computer Vision, STMicroelectronics
{*furnari, gfarinella, battiato*}*@dmi.unict.it, arcangelo.bruna@st.com*

ABSTRACT

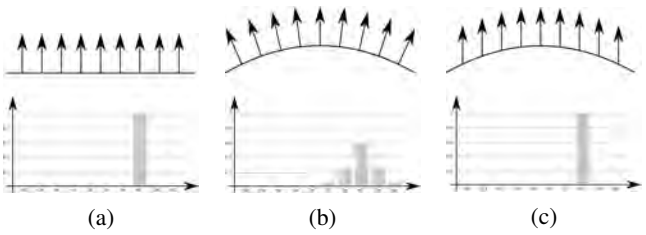
In this paper we tackle the problem of correctly estimating the gradient of distorted images. The proper estimation of the gradient in the presence of distortion is of great interest due to the large number of applications relying on wide angle cameras (e.g., in surveillance, automotive, robotics). To this aim we propose the Generalized Sobel Filters (GSF), a family of adaptive Sobel filters able to correctly estimate the gradient of distorted images. To assess the performances of the proposed method, we acquired a benchmark dataset of high resolution images belonging to different categories which are relevant to application domains where the gradient estimation is usually employed. We build an objective evaluation pipeline and perform experiments which show that our method outperforms the state-of-the-art.

Index Terms— gradient estimation, Generalized Sobel Filters (GSF), radial distortion, omnidirectional vision, gradient-based descriptors

1. INTRODUCTION

The gradient of an image is a low level feature fundamental in image processing and computer vision. It plays the role of a first-order discrete derivative and represents the directional change of the intensity (or colour) in an image. Due to its properties the gradient is employed in a number of applications including image sharpening, enhancement and edge extraction [1, 2], object, scene and key-point representations [3, 4, 5, 6], as well as gradient domain based image processing [7, 8, 9, 10]. The meaning of the estimated gradient is very related to the geometry of the acquired image [1], indeed: i) the gradient vector in a given location always points in the direction of the greatest rate of change of the image intensity; ii) its orientation is always orthogonal to the edge in that image point; iii) its magnitude is proportional to the abruptness of the change of intensity along its direction. Therefore, if the geometry of the image is modified by a distortion, estimating the gradient with classic techniques can lead to erroneous results as briefly shown in Figure 1. As it can be noted, ignoring the effect of the distortion can lead to a misleading representation even for simple descriptors like the histograms of oriented gradients (Figure 1(b)), whereas the majority of the applications (e.g., key-point matching, object detection) would require a result similar to the one shown in Figure 1 (c) (e.g., to guarantee invariance to distortion).

In this paper we tackle the problem of estimating the gradient of distorted images correctly. Even if our analysis would apply in the general settings in which the distortion function is known and invertible, we concentrate on the case in which the distortion is due to the use of a wide angle camera. This case is indeed of great interest due to the number of application domains related to the use of wide angle cameras which would benefit from the correct computation of the image gradient. Wide angle cameras allow to obtain a Field Of View (FOV) as large as the one characterizing the human visual system [12]. This makes these sensors very convenient in many application contexts such as surveillance, tracking, automotive and robotics [13, 14, 15]. A wide angle sensor can be obtained with different designs grouped in two main categories: catadioptric [13, 16, 17] and dioptric [18, 19] systems. Under given conditions (i.e., for central systems [20, 21]) the images acquired by such sensors exhibit a radially symmetric distortion which can be modelled as an invertible function. If such mapping between the distorted space and the rectilinear one is known, the image could be rectified in order to estimate the gradient without the influence of distortion. Unfortunately the rectification process is computationally expansive and the undistorted images contain blurred parts that may affect the feature extraction process [11]. To overcome these problems, it would be ideal to estimate the gradient directly on the distorted image taking into account the geometric deformation affecting the acquired image. Approaches towards feature extraction directly on the distorted domain already exist [11, 22, 23, 24, 25]. In particular researchers already explored the possibility of estimating the gradient without performing the rectification. In [11, 26] the gradient is first estimated in the distorted domain and then corrected using an adaptive Jacobian correction matrix derived through the differential chain rule. In [23] the gradient of catadioptric images is estimated using an operator derived from the geometry of the catadioptric mirror and the gradient orientations are



(a) (b) (c)

Fig. 1. Gradient estimation of a simple feature (i.e., an edge) deceived by distortion. Top row: simple features and the estimated gradient vectors. Bottom row: histograms of gradient orientations. (a) Rectilinear feature. (b) Distorted feature (e.g., acquired with a wide angle sensor). (c) The gradient vectors as they should be computed taking into account the distortion [11].

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.

then explicitly corrected to match the geometry of the original scene.

Unlike the aforementioned previous works, we build on the rationale behind the well-known Sobel operator [27] to obtain the Generalized Sobel Filters (GSF), a family of adaptive filters which, taking into account the image distortion, estimate geometrically correct gradient vectors directly on distorted images. The only requirement for the proposed GSF to be computed is that the distortion function is known and invertible, which makes them useful for both dioptric and catadioptric designs, as well as other possible ones. The main contributions of this work can be summarized as following: 1) we introduce the GSF; 2) we build a benchmark dataset of high resolution images covering different image categories relevant to the main application domains of the image gradient which can be used for evaluation purposes; 3) we propose an objective evaluation pipeline to assess the performances of the gradient estimation methods on distorted images. The reminder of the paper is organized as follows: in Section 2 we introduce the Generalized Sobel Filters; in Section 3 we discuss the evaluation pipeline, the distortion model, the error measure and the benchmark dataset; in Section 4 we discuss the results and finally we conclude the paper in Section 5.

2. GENERALIZED SOBEL FILTERS

The Sobel operator was originally developed by Irwin Sobel in 1968 [27]. Sobel suggests to estimate the gradient vector at a given point by the vector summation of the central gradient estimates along the 4 main directions in a 3×3 neighbourhood (see Figure 2 (a)). Accordingly, each of the 4 simple central gradient estimates is a vector sum of a pair of orthogonal vectors, where each orthogonal vector is a directional derivative estimate multiplied by a unit vector specifying the direction of the derivative. The magnitude of the simple directional derivative $\mathbf{g}$ for an antipodal pair of neighbouring pixels is defined by Sobel as:

$$|\mathbf{g}| = \frac{\text{density difference}}{\text{distance to neighbour}} \tag{1}$$

The direction of vector $\mathbf{g}$ is given by the unit vector to the appropriate neighbour. Figure 2 (b) shows a schema of the directional derivative estimates. The gradient can be formulated as the average of the eight simple directional derivative vectors:

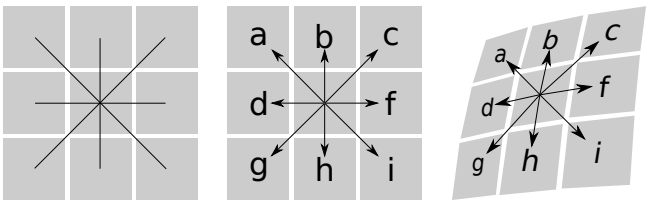
$$\nabla I(x,y) = \frac{1}{8} \cdot \sum_{\text{-}1 \leq s,t \leq 1} \left(\frac{I^{s,t}_{x,y} - I^{\text{-}s,\text{-}t}_{x,y}}{\delta^{s,t}_{x,y}} \cdot \frac{(s,t)}{\sqrt{s^2+t^2}} \right) \tag{2}$$

where I is the considered image, $I^{s,t}_{x,y} = I(x+s, y+t)$, $\delta^{s,t}_{x,y} = \delta\left[(x+s, y+t), (x-s, y-t)\right]$, (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) encoding the geometry of the distortion. Equation (2) can be reduced to the following form after algebraic manipulation:

$$\nabla I(x,y) = \sum_{\text{-}1 \leq s,t \leq 1} I^{s,t}_{x,y} \cdot \mathbf{h}(x,y,s,t) \tag{3}$$

where

$$\mathbf{h}(x,y,s,t) = \frac{1}{4} \cdot \frac{1}{\delta^{s,t}_{x,y}} \frac{(s,t)}{\sqrt{s^2+t^2}}. \tag{4}$$



(a) (b) (c)

Fig. 2. (a) The 4 main directions in a 3×3 neighbourhood. (b) The 8 simple directional derivative estimates. (c) An example of distorted (non-Euclidean) neighbourhood.

Separating the vectorial quantity $\mathbf{h}$ in terms of its x and y components, the proposed Generalized Sobel Filters are defined as:

$$h_1(x,y,s,t) = \frac{1}{4} \cdot \frac{1}{\delta^{s,t}_{x,y}} \frac{s}{\sqrt{(s^2+t^2)}} \tag{5}$$

$$h_2(x,y,s,t) = \frac{1}{4} \cdot \frac{1}{\delta^{s,t}_{x,y}} \frac{t}{\sqrt{(s^2+t^2)}}) \tag{6}$$

where we consider $h_1(x,y,s,t) = h_2(x,y,s,t) = 0$ for $(s,t) = (0,0)$. The original Sobel filters overestimate the gradient magnitudes by a factor of 16 [27], while our formulation does not account for such an approximation. Therefore a formulation of the GSF totally compatible with the original Sobel filters can be achieved multiplying each filter by a factor of 16. It should be noted that the above defined h_1 and h_2 are adaptive kernels which can be used to estimate the x and y components of the gradient vectors directly on the distorted image.

The function δ plays the role of the distance function in the geometrical domain of the input image. If no distortion is present in the image, the neighbourhood is rectilinear (as in Figure 2 (b)) and δ is chosen as the Euclidean distance. In the presence of distortion the neighbourhood cannot be considered rectilinear (Figure 2 (c)) and δ shall be chosen in order to encode such geometry. In some cases (e.g., when the distortion can be modelled using the Riemannian geometry as in [25, 28]) this function is known and can be directly used. However, if the mapping function $(x,y) = f(u,v)$ between the undistorted point of coordinates (u,v) and the distorted point (x,y) is known and invertible, the metric δ can be naturally chosen as the following composition:

$$\delta\left[(x_1,y_1),(x_2,y_2)\right] = d\left[f^{-1}(x_1,y_1), f^{-1}(x_2,y_2)\right], \quad \forall (x_1,y_1),(x_2,y_2) \in \Re^2 \tag{7}$$

where f^{-1} denotes the inverse of f and d is the Euclidean distance. The exploitation of (7) 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 (5) and (6) depend on the considered point (x,y) and hence the convolutions required for the estimation of the gradient are adaptive. In the absence of distortion, the function δ is naturally chosen as the Euclidean distance d. In this case the filters defined in (5) and (6) lose their dependence on x and y and the required convolutions are equivalent to standard Sobel convolutions (up to a factor of 16 in the estimation of the magnitudes as pointed out above). Figure 3 shows some sample GSF computed considering a radial distortion.

	upper left corner			center			upper right corner		
x	-0.3326	0	0.8951	-1.0000	0	1.0000	-0.8951	0	0.3326
	-0.7843	0	0.7843	-2.0000	0	2.0000	-0.7843	0	0.7843
	-0.8951	0	0.3326	-1.0000	0	1.0000	-0.3326	0	0.8951
y	-0.3326	-1.0279	-0.8951	-1.0000	-2.0000	-1.0000	-0.8951	-1.0279	-0.3326
	0	0	0	0	0	0	0	0	0
	0.8951	1.0279	0.3326	1.0000	2.0000	1.0000	0.3326	1.0279	0.8951

Fig. 3. Some examples of the generated GSF in the case of radial distortion. The filters are related to the centre of the image, to the upper left and upper right corners.

3. EVALUATIONS

For the evaluations we have used rectilinear images (see Section 3.3) to which the radial distortion has been artificially added, as done in [11, 22, 23]. Working in this settings, rather than using real wide angle images, is convenient as it allows to control the amount of distortion present in the image. Moreover, the source rectilinear image can be used directly to compute the reference gradient which would serve as a ground truth for evaluation purposes. Given the rectilinear image I, we denote its distorted version with $\hat{I}$, while ∇ and $\mathcal{G}$ denote the classic Sobel and the generic gradient estimators respectively.

3.1. Radial Distortion Modelling

We model radial distortion using the Division Model [11, 29] which defines the relationship between the rectilinear point[1] $\mathbf{u} \in \Re^2$ and its distorted counterpart $\mathbf{x} \in \Re^2$ as:

$$\mathbf{x} = f(\mathbf{u}) = \frac{2\mathbf{u}}{1 + \sqrt{1 - 4\xi||\mathbf{u}||^2}} \tag{8}$$

where parameter ξ allows to select the degree of distortion in the image. Unfortunately, the effects of the parameter ξ depend on the dimensions of the image, which makes its interpretation not intuitive. Therefore, we quantify the distortion undergone by an image as the percentage (expressed in decimal values):

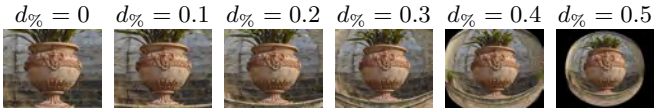
$$d_{\%} = 1 - \frac{\hat{r}_M}{r_M} \tag{9}$$

where r_M represents the maximum radius in the original image (i.e., the distance from the centre to the corner) and $\hat{r}_M$ represents its distorted counterpart. Given the relationship between r_M and $\hat{r}_M$ [11]:

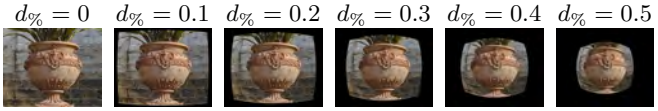
$$r_M = \frac{\hat{r}_M}{1 + \xi \hat{r}_M^2} \tag{10}$$

[1]The coordinates are referred to the centre of the distortion, which is conveniently set to the centre of the image in our experiments.

(a) Input high resolution image.



(b) Artificially added distortion with variable FOV.



(c) Artificially added distortion with static FOV.

Fig. 4. Images with artificially added distortion [22].

the parameter ξ can be chosen to match the given percentage of distortion $d_{\%}$ using the following formula:

$$\xi = -\frac{d_{\%}}{[r_M(1 - d_{\%})]^2}. \tag{11}$$

Once the mapping (8) is known, the distorted image can be computed by mapping the undistorted points to the distorted domain and interpolating when needed:

$$\hat{I}(x, y) = I(f^{-1}(x, y)). \tag{12}$$

In previous works the radial distortion has been added considering a static FOV [11, 22] as it is shown in Figure 4(c). As it is possible to notice, this lets the borders of the distorted image uncovered as the distortion rate increases. We obtain a variable FOV (Figure 4(b)) by mapping high resolution reference images (Figure 4(a)) to low resolution ones as suggested in [22].

3.2. Error Measure

Gradient vectors are usually exploited separating the magnitudes from the orientations. The orientations carry important information about the distribution of the edges in the scene, while the magnitudes suggest which orientations are likely to be more representative than others. For this reason, the weighted histograms of gradient orientations are usually employed [4, 5]. A weighted histogram is defined as a regular one, except that the contribution of a given orientation is weighted by its magnitude. As specified below, we define our error measure by considering the local populations (i.e., weighted histograms) of gradient orientations in both the reference image I and the distorted one $\hat{I}$.

Let $\mathcal{G}$ be 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: $\{\hat{T}_i\}_{1 \leq i \leq n}$. For each tile $\hat{T}_i$ in the distorted image $\hat{I}$, we consider the related undistorted tile T_i in the reference image I, which contains the undistorted counterparts of all the points in $\hat{T}_i$. It should be noted that the number of points in $\hat{T}_i$ can be smaller than

Fig. 5. The considered dataset.

the one of T_i due to the loss of information given by the radial distortion. We define the error committed by the gradient estimator $\mathcal{G}$ given the image pair $(\hat{I}, I)$ as:

$$\epsilon(\mathcal{G}, \hat{I}, I) = \frac{1}{n} \sum_{i=1}^{n} \rho(\mathcal{H}(\mathcal{G}\hat{T}_i), \mathcal{H}(\nabla T_i)) \qquad (13)$$

where $\mathcal{H}(\mathcal{G}\hat{T}_i)$ and $\mathcal{H}(\nabla T_i)$ denote the weighted histograms of the estimated and reference gradient orientations of $\hat{T}_i$ and T_i, and ρ is the metric based on the Bhattacharyya coefficient as defined in [30]:

$$\rho(H_1, H_2) = \sqrt{1 - \sum_{u=1}^{m} \sqrt{H_1^u \cdot H_2^u}} \qquad (14)$$

where m is the number of bins of histograms H_1 and H_2 and H^u denotes the $u-th$ component of H. It should be noted that the Bhattacharyya distance ρ is a measure of dissimilarity normalized in the range $[0, 1]$.

3.3. Dataset

We collected a dataset of 100 high resolution images. All images have been acquired using a Canon 650D camera mounting a Canon EF-24mm lens and have resolution of 5204×3472 pixels. The images depict scenes taken from the following categories: indoor, outdoor, natural, handmade, urban scene, car, pedestrian, street. These categories have been chosen considering the main application domains of the gradient [4, 5] as well as the scene categorization proposed by Torralba & Oliva [31]. Figure 5 shows the thumbnails of the images composing the dataset. The dataset can be downloaded at the url *http://iplab.dmi.unict.it/icip2015/dataset.zip*.

4. EXPERIMENTAL SETTINGS AND RESULTS

We assessed the performances of the proposed GSF against the Gradient Correction Jacobian (GCJ) method introduced in [11, 26]. As baseline for comparisons we also considered two simple approaches in the comparison: 1) ignoring the distortion (the gradients are computed using the Sobel operator in the distorted space and the results are not corrected); 2) explicit removal of the distortion (i.e., rectification) prior to estimating the gradient using the Sobel operator. The reference gradient is computed from the rectilinear reference images using the standard Sobel filters [27]. The high resolution input images are projected to distorted images of standard resolution 1024×768 pixels. The estimation of the local histograms of gradient orientations is performed using the following parameters: 1) each image is divided into tiles of dimensions 24×24 pixels and 2) 18 bins evenly spacing the interval $[-180°, 180°]$ are considered. To assess the performances of the considered method with respect to different degrees of distortion, we processed every image in the dataset in order to add varying percentages of distortion ranging from a minimum of 10% to a maximum of 50%.

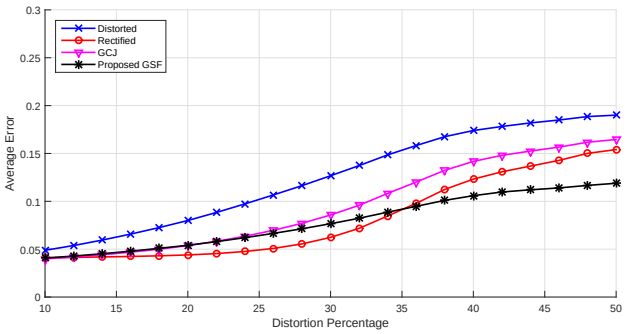


Fig. 6. The mean error for different gradient estimators on the considered dataset at varying of the distortion percentage.

Figure 6 shows the average error for the considered methods with respect to the different amounts of radial distortion. Each curve is obtained averaging the error scores related to the 100 images in the dataset. Both the GSF and GCJ methods return a error smaller than the one of the distorted gradients for all distortion rates. The rectification method allows good results for low distortion rates (where the lost information can be still "guessed" by the rectification process) whereas the error gets higher as the distortion rate increases. The proposed GSF perform better than the other competitors for distortion rates over the 35%. However, it should be noted that for distortion rates under the 35% it is still advantageous to use the proposed GSF since the rectification process is computationally expensive and not always feasible [11, 22].

5. CONCLUSION

We have proposed the Generalized Sobel Filters (GSF), a family of Sobel filters which can be used to correctly estimate the gradient of distorted images. The GSF are independent from the used model of distortion and require only the distortion function f to be known and invertible. To assess the performances of the proposed adaptive filters, we built a dataset of high resolution images belonging to different scene categories and proposed an evaluation pipeline which measures the error of the estimated gradients. The experiments show that our method outperforms the state-of-the-art. Future works could be devoted to assessing the improvement in the performances of popular descriptors such as SIFT and HOG, when the GSF are used to compute the image gradients.

6. REFERENCES

[1] R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, Pearson/Prentice Hall, 2008.

[2] J. Canny, "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.

[3] J. Domke and Y. Aloimonos, "Deformation and viewpoint invariant color histograms.," in *The British Machine Vision Conference*, 2006, pp. 509–518.

[4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[5] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition*, 2005, vol. 1, pp. 886–893.

[6] G. M. Farinella and S. Battiato, "Scene classification in compressed and constrained domain," *IET Computer Vision*, vol. 5, no. 5, pp. 320–334, 2011.

[7] P. Pérez, M. Gangnet, and A. Blake, "Poisson image editing," in *ACM Transactions on Graphics*, 2003, vol. 22, pp. 313–318.

[8] P. Bhat, C. L. Zitnick, M. Cohen, and B. Curless, "Gradientshop: A gradient-domain optimization framework for image and video filtering," *ACM Transactions on Graphics*, vol. 29, no. 2, 2010.

[9] S. Battiato, G. M. Farinella, G. Puglisi, and D. Ravì, "Saliency based selection of gradient vector flow paths for content aware image resizing," *IEEE Transactions on Image Processing*, vol. 23, no. 5, pp. 2081–2095, 2014.

[10] G. Messina, S. Battiato, M. Mancuso, and A. Buemi, "Improving image resolution by adaptive back–projection correction techniques," *IEEE Transactions on Consumer Electronics*, vol. 48, no. 3, pp. 409 – 416, 2002.

[11] M. Lourenço, J. P. Barreto, and F. Vasconcelos, "sRD-SIFT: keypoint detection and matching in images with radial distortion," *IEEE Transactions on Robotics*, vol. 28, no. 3, pp. 752–760, 2012.

[12] M. M. Fleck, "Perspective projection: the wrong imaging model," *Department of Computer Science, University of Iowa*, 1995.

[13] L. Puig and J. J. Guerrero, *Omnidirectional Vision Systems*, Springer, 2013.

[14] C. Hughes, M. Glavin, E. Jones, and P. Denny, "Wide-angle camera technology for automotive applications: a review," *IET Intelligent Transport Systems*, vol. 3, no. 1, pp. 19–31, 2009.

[15] S. Battiato, G. M. Farinella, A. Furnari, G. Puglisi, A. Snijders, and J. Spiekstra, "Vehicle tracking based on customized template matching," in *VISAPP International Workshop on Ultra Wide Context and Content Aware Imaging*, 2014.

[16] S. Baker and S. K. Nayar, "A theory of catadioptric image formation," in *International Conference on Computer Vision*, 1998, pp. 35–42.

[17] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *European Conference on Computer Vision*, 2000, pp. 445–461.

[18] K. Miyamoto, "Fish eye lens," *Journal of the Optical Society of America*, pp. 2–3, 1964.

[19] C. Hughes, P. Denny, E. Jones, and M. Glavin, "Accuracy of fish-eye lens models," *Applied Optics*, vol. 49, no. 17, pp. 3338–47, 2010.

[20] J. P. Barreto, "A unifying geometric representation for central projection systems," *Computer Vision and Image Understanding*, vol. 103, no. 3, pp. 208–217, 2006.

[21] D. Scaramuzza, A. Martinelli, and R. Siegwart, "A toolbox for easily calibrating omnidirectional cameras," in *International Conference on Intelligent Robots and Systems*, 2006, pp. 5695–5701.

[22] A. Furnari, G. M. Farinella, G. Puglisi, A. R. Bruna, and S. Battiato, "Affine region detectors on the fisheye domain," in *IEEE International Conference on Image Processing*, 2014.

[23] I. Cinaroglu and Y. Bastanlar, "A direct approach for human detection with catadioptric omnidirectional cameras," in *Signal Processing and Communications Applications Conference*, 2014, pp. 2275–2279.

[24] J. Cruz-Mota, I. Bogdanova, B. Paquier, M. Bierlaire, and J. Thiran, "Scale invariant feature transform on the sphere: theory and applications," *International Journal of Computer Vision*, vol. 98, no. 2, pp. 217–241, 2011.

[25] Z. Arican and P. Frossard, "OmniSIFT: Scale invariant features in omnidirectional images," *IEEE International Conference on Image Processing*, pp. 3505–3508, 2010.

[26] M.S. Islam and L.J. Kitchen, "Straight-edge extraction in distorted images using gradient correction," in *Digital Image Computing: Techniques and Applications*, 2009.

[27] P.E. Danielsson and O. Seger, "Generalized and separable Sobel operators," *Machine Vision for Three-dimensional Scenes*, pp. 347–379, 1990.

[28] I. Bogdanova, X. Bresson, J. Thiran, and P. Vandergheynst, "Scale space analysis and active contours for omnidirectional images.," *IEEE Transactions on Image Processing*, vol. 16, no. 7, pp. 1888–901, 2007.

[29] A. W. Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *Computer Vision and Pattern Recognition*, 2001, vol. 1.

[30] D. Comaniciu, V. Ramesh, and P. Meer, "Kernel-based object tracking," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 5, pp. 564–577, 2003.

[31] A. Torralba and A. Oliva, "Statistics of natural image categories," *Network: Computation in Neural Systems*, vol. 14, no. 3, pp. 391–412, 2003.