

Fuzzy Edge Detection for Omnidirectional Images

Florence Jacquey, Frédéric Comby, Olivier Strauss

▶ **To cite this version:**

Florence Jacquey, Frédéric Comby, Olivier Strauss. Fuzzy Edge Detection for Omnidirectional Images. Fuzzy Sets and Systems, 2008, 159 (15), pp.1991-2010. 10.1016/j.fss.2008.02.022 . lirmm-00299544

HAL Id: lirmm-00299544

https://hal-lirmm.ccsd.cnrs.fr/lirmm-00299544

Submitted on 17 Jul 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fuzzy Edge Detection for Omnidirectional Images

Florence JACQUEY, Frédéric COMBY and Olivier STRAUSS

LIRMM
Laboratoire d'Informatique, de Robotique et de Microélectronique de Montpellier,
161, rue Ada, 34392 Montpellier Cedex 5 France

Abstract

The use of omnidirectional vision has increased during these past years. It provides a very large field of view. Nevertheless, omnidirectional images contain significant radial distortions and conventional image processing is not adapted to these specific images. This paper presents an edge detector adapted to the image geometry. Fuzzy sets will be used to take into account all imprecisions introduced by the sampling process. The Prewitt filter applied to omnidirectional image will be studied to illustrate this paper.

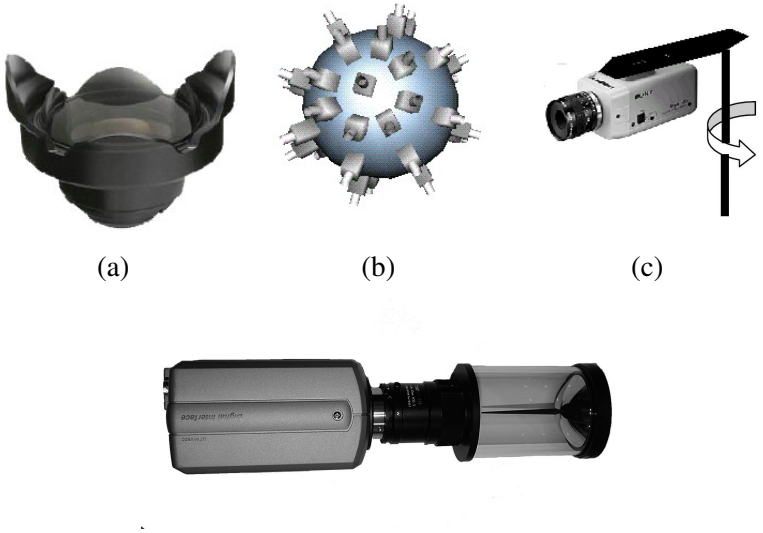
Key words: Omnidirectional images, edge detection, Choquet integral, fuzzy partitioning

1 Introduction

Classical cameras have a very limited field of view (from 30 up to 60 degrees). However, many applications such as robotics, video surveillance, conference, virtual reality representations, etc. need a large field of view. There are many ways to enhance this field of view and obtain an omnidirectional sensor. Historically the first solution was to use a fisheye lens (Figure 1 (a)). The classical optics of the camera is replaced by a very short focal length lens [15]. The field of view is larger than 180 degree but these lens are quite complex to model and expensive. Mutiple-camera devices (Figure 1 (b)) can also be used. In [25], six or more cameras are arranged so that they sample different parts of the visual sphere. This geometric configuration presents some advantages compared to classical small field of view cameras. The rotating camera system (Figure 1 (c)) consists in rotating a classical camera around a line perpendicular to the optical axis [16]. This system provides high resolution panoramic images, but because of the calibration difficulties, the

Email address: {Jacquey,Comby,Strauss}@lirmm.fr (Florence JACQUEY, Frédéric COMBY and Olivier STRAUSS).

reconstruction of a panoramic image from many perspective image is not easy. The scene must be static and real-time applications are quite impossible. The catadioptric system (Figure 1 (d)) is a solution based on a reflecting surface as convex mirror [1]. The imaging system combines a catadioptric system and a camera. In this case the camera perceives its environment through its reflection on a revolution mirror. The camera is omnidirectional and presents a full 360 degree field of view. However, resulting images present a complex geometry which make tough the image processing. It is also possible to combine those previous methods [17]. This paper will focus on omnidirectional sensors using mirrors also known as catadioptric sensors.



(a) (b) (c)

(d)

Fig. 1. (a) Fisheye lens (b) Multi-camera devices (c) Rotating camera (d) Catadioptric camera

Baker and Nayar [4] have shown that there are two kinds of catadioptric systems depending on whether they satisfy the single view point (SVP) constraint or not. SVP sensors are those equipped with parabolic, hyperbolic, elliptic and planar mirrors. However, as planar mirrors actually do not increase the field of view, they are quite useless for omnidirectional vision. Single viewpoint sensors are very useful because they allow the generation of geometrically correct perspective images from one omnidirectional image. The SVP constraint expresses the fact that each pixel in the image measures the irradiance of the light passing through the viewpoint in one specific direction. If the catadioptric system geometry is known, this direction can be computed for each pixel. Thus, the gray level value measured by each pixel can be projected onto a plane at any distance from the viewpoint to form a planar perspective image. These perspective images are easier for human being as they only present small distortions.

Moreover processing omnidirectional images is slightly different from processing

perspective images. With classical perspective images, most of the filtering techniques use kernels. These kernels are called summative as the integral over their definition domain equals 1. The aim of the kernel is to define a weighted neighborhood of each sampled location. This neighborhood provides an interplay between continuous and discrete. It also aims at using a hypothesized ergodicity to reduce the effect of random noise by providing regularization in the interpolation process. This use of neighborhood suppose that the observed scene is located on a camera fronto-parallel plane. Therefore, whatever the kernel position on the image may be, its effect on the image will remain the same. This means the operator is invariant with respect to translations. Omnidirectional cameras cannot satisfy this condition. Let us consider a regular grid pattern composed of white lines on a black plane. Figure 2(a) illustrates this pattern captured by a classical camera (2(b)). The lines thickness remains the same all over the image. An edge detector able to find the two edges of a line at the top of the image will also be able to find it at its bottom. Let us now consider an omnidirectional camera observing the same pattern. Figure 2(c) shows the omnidirectional camera disposition with respect to the grid pattern. Figure 2(d) represents the grid pattern captured by an omnidirectional camera. It is easy to notice that the lines thickness diminishes as one gets closer to the image center. Therefore, a filter detecting two boundaries at the image periphery will surely not be able to do so at the image center. The aim of this paper is to adapt classical filters using a kernel for omnidirectional images. This implies that the kernel size is dependent of its position on the image. Thus its effect remains the same all over the image. This article will focus on a specific edge detector: the Prewitt filter.

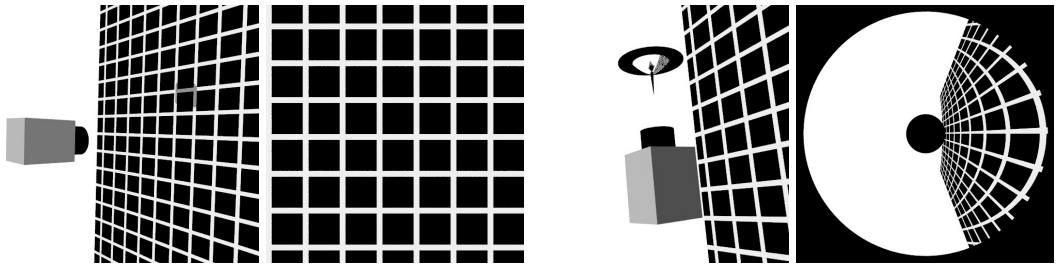
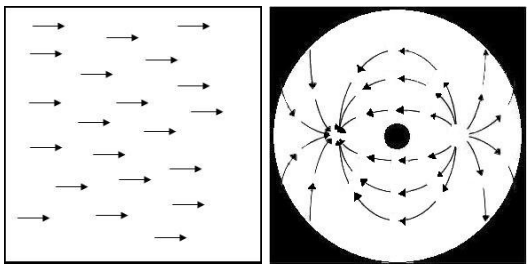


Fig. 2. (a) position of the grid with a classical camera, (b) A regular grid on a plane projected on a perspective image, (c) position of the grid with respect to the camera's, (d) omnidirectional image of the grid.

Although omnidirectional cameras allow to increase the field of view, some problems arise. Anamorphosis in omnidirectional images introduces complexity in image processing and interpretation: for example, optical flow computation. Optical flow is computed from images spatio-temporal derivatives in order to estimate the apparent motion in a digital image sequence. Using appropriate motion models, the pixels apparent motion can be related to the camera motion. Concerning omnidirectional images, a simple camera translation implies a complex apparent motion

[3]. Indeed, a camera translation does not generate an apparent translation of all the pixels in the image (Figure 3).



(a) (b)

Fig. 3. Pixels motion for a classical camera translation (a) and for an omnidirectional camera (b).

Specific treatments to process omnidirectional images have been proposed in the literature. They try to improve the performances of classical treatments. For example, in the article [18], an edge detection composed by two detectors with a 90° rotation is proposed. Nevertheless, this filter is not really adapted to omnidirectional images because the effects of the filter remain the same all over the image without taking into account the image distortion. In [20], Demonceaux presents an approach using wavelet to compute the image gradient. Even if this filter gives better results than classical filters, in visual terms, it is also not adapted to the image geometry. So, there is a real need in, directly, taking into account the image geometry in the definition of image processing algorithms.

2 Geometrical Considerations

2.1 Catadioptric sensor model

As shown in introduction, panoramic images present some distortions. The anamorphosis depends on the shape of the mirror. However, Geyer et Daniilidis have introduced a unifying theory for all central catadioptric sensors in [5]. They proved that the anamorphosis provided by a central panoramic projection is isomorphic to a projective mapping onto a virtual sphere as illustrated on Figure 4. The first step is a central projection to the sphere and the second step, a central projection from a point B (on the sphere's axis) to the retina. The position of B depends on the kind of mirror used in the catadioptric sensor. A parabolic mirror is represented by a point B lying on the north pole of the sphere. For an elliptic or hyperbolic mirror, B lies between the north pole and the sphere's center. A huge advantage of this modelization is that it involves only two parameters (φ and ξ) to map a 3D visible point $M(X, Y, Z)$ with its projection $m(x, y)$ on the panoramic image. This

projection is described by equations (1) and (2), where φ represent the distance between the omnidirectional image and the sphere center, ξ the distance OB and $R^2 = X^2 + Y^2 + Z^2$. Therefore, using this model allows an easy transposition of our algorithm to any kind of SVP catadioptric omnidirectional sensor.

$$x = \frac{(\varphi + \xi)X}{\xi R - Z} \tag{1}$$

$$y = \frac{(\varphi + \xi)Y}{\xi R - Z} \tag{2}$$

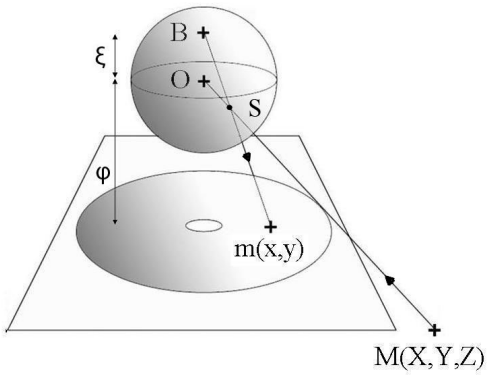


Fig. 4. The unifying model for describing all central catadioptric projections

2.2 New workspace for kernels ?

Applying filters based on convolution kernels with image implies that the kernel effects are invariant by translation. This is true for classical images with classical operators. But as soon as distortions due to projection are involved, this statement does not stand. Then, the aim is to find a space where the kernel can be defined so that its effect on the panoramic image is invariant to translations. Some solutions have been found in the literature.

Concerning echographic images, which are not really omnidirectional images, but still present distortions due to the image acquisition process, Herlin and Ayache try to work directly on the ultrasound images in [12]. They acquired ultrasound scan lines in polar coordinates which creates an important anisotropy in spatial resolution. They geometrically transforms the data from a polar representation to take account for the varying resolution of the data. The proposed method consists in computing a scan-conversion with a low-pass filtering of the cartesian image applied directly to the available polar data.

In [6], Daniilidis, Makadia and Bullow propose to project the omnidirectional image onto a virtual sphere and define a gradient operator on it. However, the gradient operator definition is quite complex and needs to be expressed in spherical coordinates. Moreover, the regular sampling on the panoramic image is changed into a non-regular sampling onto the sphere. The non-regular imprecision brought by this projection is hard to consider.

Bigot, Kachi, Durand et Mouaddib [27] have proposed to develop operators directly on the virtual sphere. A new smoothing method for spherical images is described. The authors have introduced a suitable Wiener filter and have used the tikhonov method for this these images. This method is compared with the most used classical spherical kernels on real and synthetical spherical images. The signal to noise ratio prove the effectiveness of these filters for denoising operation. The treatments can be carried out in real time and the filters are invariant by rotation. The results are thus independent of the position on the sphere.

Tosic, Bogdanova, Frossard and Vandergheynst [29] have presented a new local motion estimation algorithm for omnidirectional images. This algorithm is used to compute the correlation between two spherical images of a scene, taken from arbitrary viewpoints, with the objective to reduce the encoding rate of these images. Their algorithm provides a quite efficient image prediction: the prediction error is almost exclusively composed of high frequency noise. In this article, the authors underline the importance to avoid potential discrepancies induced while unfolding omnidirectional images to implement a classical motion estimation on images.

The image geometry plays a central role in the formation of omnidirectional images and must be carefully taken into account while performing such simple tasks as smoothing or edge detection. To cope with geometrical constraints, Bogdanova, Bresson, Thiran et Vandergheynst [28] have derived new energy functionals and partial differential equations for segmenting images obtained from catadioptric cameras. They have shown that it is possible to find a robust implementation using classical finite difference schemes.

In [30], Demonceaux and Vasseur have used the markov random fields on catadioptric images. Because the neghborhood topology is modified by the mirror, they propose to define a new neighborhood for markov random fields by using the equivalence theorem developed for central catadioptric sensors. The importance of this adaptation is proved for segmentation, image restoration and motion detection.

Another projective space is suggested in [10]: a virtual cylinder surrounding the mirror. Like the sphere, it allows an omnidirectional representation of the observed scene. This view is still distorted but is consistent with the way we are used to see images. Moreover, the cylinder can be approximated as the collection of central strips of projective images rotating around the mirror axis (Figure 5). The main advantage of the cylindrical collector is that when unfolded it can be sampled with a

regular Cartesian grid. When the kernel size remains small, the unfolded cylindrical image part to be convolved with the kernel is quasi equivalent to the one extracted from a perspective image.

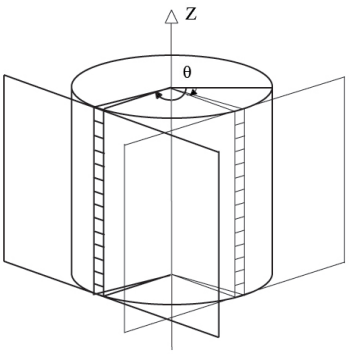


Fig. 5. Quasi equivalence between the projection of a pixel onto the cylinder and onto a perspective image

Another approach consists in reconstructing all perspective images from the omnidirectional image and in processing it. But there are a huge amount of possible reconstructed perspective images which considerably increases the process complexity. Moreover, the resulting images are interpolated and arbitrarily depend of the chosen sampling.

The reconstruction of the omnidirectional image on another support like the cylinder or the sphere, involves also interpolations and smoothing that modify the image information. These errors will be propagated through all the image processing thus increasing their effects. Our approach tries to preserve the original informations until the last processing to minimize the impact of these imprecisions in the processing step.

In this paper, the projective space selected was the cylinder. The main reason is the small difference between treatments on the cylinder and those on computed perspective images. The approximation brought by the cylinder, with respect to the perspective plane, is almost negligible. For example, with a 3×3 kernel, the localization error on the projected kernel is about 5.10^{-5} pixels in the worst case. Moreover, the hypothesis in classical perspective images treatment is that the distance between the perspective camera and the scene is constant. Thus, treatments respect the invariance by translation. It expresses the fact that visual informations will be identically processed whatever their spatial localization may be. We tried to reproduce the same reasoning with catadioptrical cameras. A single view point catadioptric camera produces a view equivalent to the one acquired by a rotating camera. Each position of the rotating camera will suppose that the 3D scene is in a specific plane. The collection of all these planes can be approximated by a cylinder. This is why we have chosen the cylinder as projective space. The image is not mapped on the cylinder; it is only a projective plane, where classical operators can

be easily approximated. With such a process the sensor geometry is integrated in the image processing.

3 Image Processing

3.1 Generalities

As presented in the introduction, the present work document will focus on the Prewitt kernels. First of all the Prewitt filter is a simple bounded convolution mask. As kernels need to be projected onto the cylindrical support, our approach has been limited to bounded kernels. Prewitt filter was just one among the others. In fact, our approach can be applied to all bounded convolution masks.

The Prewitt edge detector is based on a basic criterion: images should contain sharp intensity transition and low noise. When using Prewitt edge detection, the image is convolved with a set of convolution kernels. Each convolution kernel is sensitive to edges in a different orientation (horizontal, vertical and oblique) and provides an estimation of the component of the brightness gradient at every pixel. For each pixel, the local edge gradient magnitude is estimated with the maximum response of all kernels at this pixel location.

In our article, we have just presented an example of the Prewitt filter in the vertical direction, illustrated on Figure 6. The whole set of kernels is produced by taking this kernel and rotating its coefficients circularly.

1	0	-1
1	0	-1
1	0	-1

Fig. 6. 3×3 Prewitt mask of a vertical edge detector

The gray level $F_{i,j}$ of the filtered pixel, in a given direction, at (i, j) coordinates, is given by formula (3)

$$F_{i,j} = \sum_{u=-1}^{1} \sum_{v=-1}^{1} C_{u,v} I_{i+u,j+v} \tag{3}$$

where $C_{u,v}$ is the coefficient value of the Prewitt mask at (u, v) coordinates ($C_{0,0}$ stands for the central element of the mask) and $I_{i+u,j+v}$ is the gray level of the pixel located in $(i+u, j+v)$ on the image. This processing uses a neighborhood of pixels to compute the filtered pixel gray level. For perspective images, the kernel shape is

adapted to the image sampling: a Cartesian kernel convolved with a Cartesian partition of the same resolution. However for omnidirectional images, the resolution depends of the position on the image. Moreover the topology of neighborhood is modified by the projection on the mirror. Any classical kernel operator is therefore not adapted to omnidirectional images.

3.2 Filter Adaptation

The first intuitive solution to adapt the Prewitt edge detector is to project the omnidirectional image on the cylinder, then to process the unfolded image and finally, to back-project it on the omnidirectional plane. However, projections and back-projections involve interpolations thus introducing errors in the processing. As pointed out in [6], it is crucial that data are kept in their original space. Thus, if data are unchanged the operator needs to be modified. Instead of projecting pixels values on another support, the kernel will be geometrically defined and projected on the omnidirectional image.

The proposed algorithm is structured as follows:

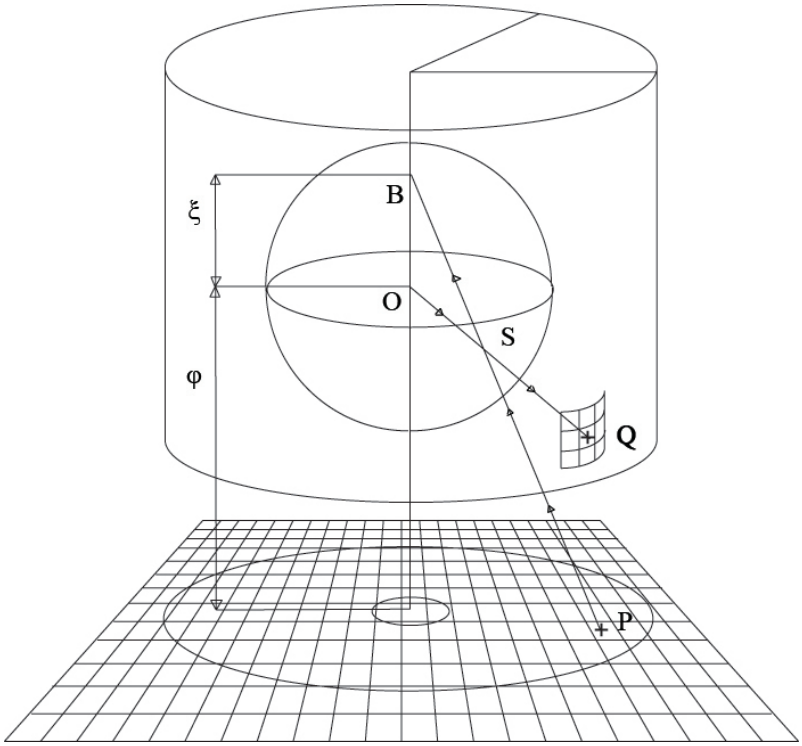
Algorithm 1 Kernel definition and computation

for *each pixel in omnidirectional image* **do**
  Project the pixel center (P) on the cylinder (Q).
  Definition of a Prewitt kernel around Q.
  Back-projection of the kernel on the omnidirectional image.
  Computation of the filtered pixel gray level.
end for

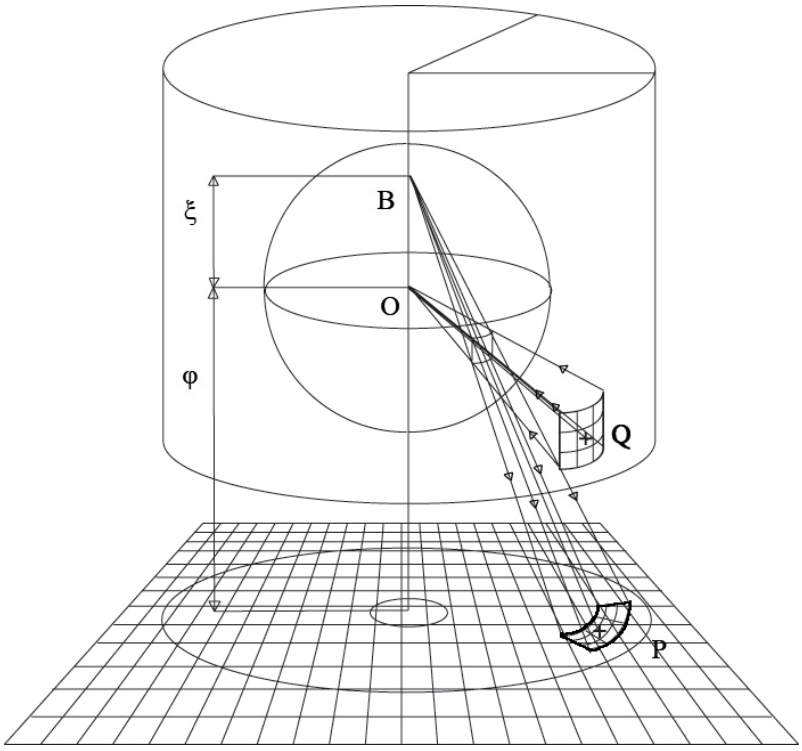
Figure 7 (a) illustrates the projection of a pixel center on the cylinder. It also shows the definition of a Prewitt kernel around this projection. Then, thanks to the extension principle [7], the kernel is back-projected on the panoramic image (Figure 7 (b)) and the gray level value of the filtered pixel is computed with respect to the projected kernel. This comes down to change the kernel size depending on its position on the image.

One problem arises: how to choose the kernel on the cylinder? The solution presented here is to define a kernel centered on the projection of the considered pixel. Its size is arbitrarily determined depending both on its capability to highlight the sought after edges and on the amount of filtering. There is no real rule to set the sampling on the cylinder. However, to remain suitable with the sensor, the cylinder sampling step is chosen proportionally to the omnidirectional image external perimeter (expressed in pixels). Indeed, this perimeter corresponds to the best resolution on the image.

Now another issue is raised: how to compute the gray level associated to the pro-



(a)



(b)

Fig. 7. (a) Center pixel projection onto the cylinder and regular mask reconstruction. (b) Mask projection onto the omnidirectional image

jected kernel? As it is shown on Figure 8 (a) each part of the kernel covers many pixels of the panoramic image. The value of $I_{i+u,j+v}$ cannot be computed as in the equation (3). The gray level value associated to each kernel element must be evaluated to estimate the filtered pixel value. Two solutions are proposed in sections 3.3 and 3.4 to compute the gray levels of the projected kernel elements.

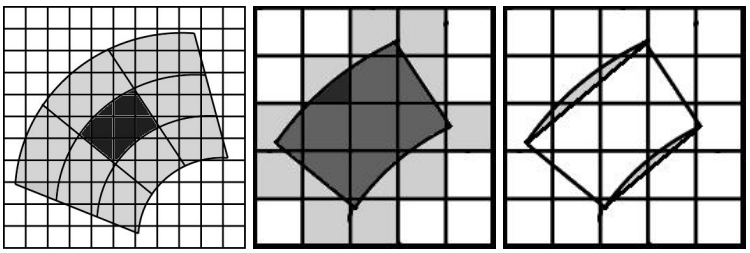
3.3 Proportional Approach

The aim is to transfer the information of the intersected pixels toward a projected element of the kernel. A first solution, inspired from belief transfer, is briefly presented here: the pignistic transfer [14]. Each pixel and mask element is considered as a 2D interval defined by its support and with a uniform membership function. The gray level value of a kernel element is a weighted sum of all intersected pixels gray levels. The weight is proportional to the area of intersection between the pixel and the projected kernel (dark gray shaded area in Figure 8 (b)). This weight is normalized by the area of the projected kernel.

In equation (3) $I_{i+u,j+v}$ has to be replaced by (4)

$$I_{i+u,j+v} \Leftrightarrow \sum_{\Omega(k,l)} \frac{|M_{u,v} \cap P_{k,l}|}{|M_{u,v}|} I_{k,l} \tag{4}$$

where $M_{u,v}$ is the 2-D interval associated to the projection of (u, v) kernel cell on the omnidirectional image, $P_{k,l}$ is the 2-D interval associated to the pixel $p_{k,l}$, $\Omega_{k,l}$ is the set of 2-D intervals $P_{k,l}$ intersected by $M_{u,v}$ (light gray shaded area on Figure 8 (a)) and $I_{k,l}$ is the gray level of the $p_{k,l}$ pixel.



(a) (b) (c)

Fig. 8. (a) The projected mask covers many pixels of the omnidirectional image. (b) The gray level of a projected pixel is proportional to the intersected pixels surface (c) To make easier the intersection area compilation, the mask is approximated by a quadrilateral

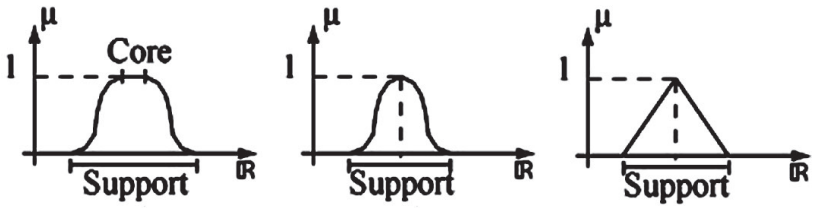
In order to simplify the intersections area computation and to improve the computation time, an approximation of the projected kernel is used (Figure 8 (c)). The surface is approximated by a trapeze. With this approximation, the kernel elements

are independent and do not overlap each other. The projected elements constitute a covered partition of the omnidirectional image.

3.4 Fuzzy Approach

3.4.1 Fuzzy concepts applied to signal processing

The data imprecision is generally modeled by crisp intervals. Fuzzy intervals generalize the notion of crisp intervals. They could be defined by the support and core of their membership function (Figure 9(a)). Fuzzy quantities whose core is reduced to a single value are called fuzzy numbers (Figure 9(b)). In [9] the authors have shown that the triangular fuzzy numbers generalize all mono-modal and symmetric fuzzy numbers (Figure 9(c)). Then, this representation will be used for the remaining of this document to model imprecision.



(a) (b) (c)

Fig. 9. Membership functions of fuzzy quantities: (a) fuzzy interval, (b) fuzzy number, (c) triangular fuzzy number

The possibility theory introduced by Zadeh [8] allows to compare two imprecise data. This kind of data will bold written for the remainder of the document. Dubois and Prade advocate in [7] to use a possibility measure $\Pi(\mathbf{A}, \mathbf{B})$ to estimate the interaction of two imprecise data $\mathbf{A}$ and $\mathbf{B}$. $\Pi(\mathbf{A}, \mathbf{B})$ is defined by formula (5) where Ω is the real axis, μ_A (resp. μ_B) is the membership function of the fuzzy number $\mathbf{A}$ (resp. $\mathbf{B}$).

$$\Pi(\mathbf{A}, \mathbf{B}) = \mathbf{sup}_{\mathbf{\Omega}}(\mathbf{min}(\mu_{\mathbf{A}}(\mathbf{x}), \mu_{\mathbf{B}}(\mathbf{x}))) \tag{5}$$

When computerized signal processing is needed, the continuous signal to process is discretized. This results from the acquisition process. Then, the continuous signal is supposed to be known on discret intervals defined by the sampling step. The data have imprecise locations on these intervals. To represent these imprecisions, we have chosen to replace the classical binary intervals by triangular fuzzy numbers. These numbers overlap each other and allow to represent in a better way the localisation imprecision. Their support equals two times of the sampling step and their mode coincide with the given interval center. A triangular fuzzy number represent

the confidence in the interval value. All these triangular fuzzy number constitute a fuzzy partition of the space $(H_i)_{i\in\{1,...,N\}}$ depicted on Figure (10). Using a kernel for signal processing and triangular fuzzy numbers can ensure a kind of interpolation between discrete and continuous. The aim of this modeling is to transfer the discrete knowledge of the signal to any weighted neighborhood W out of the partition. The value associated to any fuzzy number $\mathbf{W}$ can then be estimated.

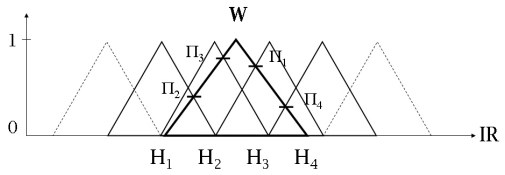


Fig. 10. Interactions between a fuzzy number $\mathbf{W}$ and the fuzzy numbers $\mathbf{H_k}$ of the axis of real partition

This transfer is ensured by using a Choquet integral [11] which allows to give a superior and inferior estimation of $\mathbf{W}$ value. These boundaries are given by equations:

$$Cv_{upper} = \sum_{n=1}^{N} I_{(n)}[v(A_n) - v(A_{n+1})] \tag{6}$$

$$Cv_{lower} = -\sum_{n=1}^{N} -I_{(n)}[v(A_n) - v(A_{n+1})] \tag{7}$$

where I_i is the value of the i^{th} cell $\mathbf{H_i}$, (.) indicates a permutation such that in (6) $I_{(1)} \leq I_{(2)} \leq ... \leq I_{(N)}$ and in (7) $-I_{(1)} \leq -I_{(2)} \leq ... \leq -I_{(N)}$. The $A_n = (\mathbf{H_n}, ..., \mathbf{H_N})$ are binary coalitions of cells whose values are superior or equal to I_n. The biggest confidence that can be given to the coalition A_n is $v(A_n)$ defined in (8).

$$v(A_n) = Sup_{i=n..N}\Pi(\mathbf{H_i}, \mathbf{W}) \tag{8}$$

To characterize this interval its middle: $C_v = (Cv_{upper} + Cv_{lower})/2$ is frequently used, but this choice remains arbitrary.

3.4.2 Application to image processing

The image acquisition process introduces an imprecision on the pixel gray level localization. This imprecision is due to the spatial sampling. It can be modeled with fuzzy sets. Each pixel of the image is considered as a bi-dimensional imprecise quantity. Indeed, for the pixel $P_{i,j}$ located at (i, j) on the image, its gray level localization is unknown within the 2D interval $[i - \Delta_i, i + \Delta_i] \times [j - \Delta_j, j + \Delta_j]$,

where Δ_i (resp. Δ_j) is half the width of the horizontal (resp. vertical) sampling. The 1-D triangular fuzzy numbers are changed in 2-D pyramidal fuzzy numbers (Figure 11(a)) considering the t-norm min (cartesian product of two 1-D triangular fuzzy numbers). All the pixels constitute a strong fuzzy partition of the image.

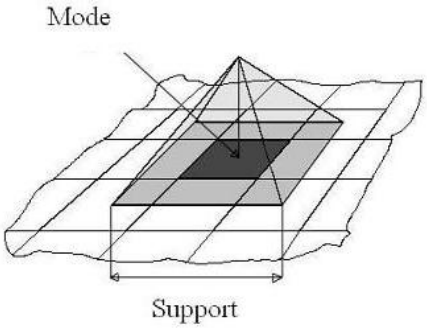


Fig. 11. Fuzzy pixel

As presented in 3.1, each omnidirectional pixel center (core of the 2-D fuzzy pixel) is projected onto the cylinder. A Prewitt kernel is defined around this projection and back-projected on the image. The Prewitt kernel presented on Figure 6 is also a sampled quantity. Therefore, there is an imprecision on the kernel values localizations. Each kernel element can be considered as a fuzzy 2-D number. Figure 12 illustrates the Prewitt kernel on the cylinder. To simplify the graphics, only its fuzzy central element is represented.

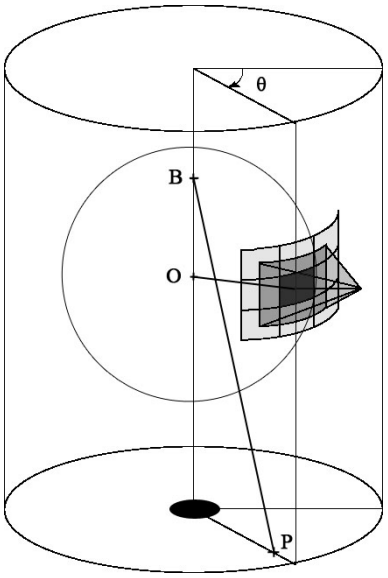


Fig. 12. The mask on the cylinder with the central fuzzy element.

The fuzzy numbers associated to the Prewitt kernel are back-projected on the omnidirectional image with respect to the mirror and camera parameters (Figure 13).

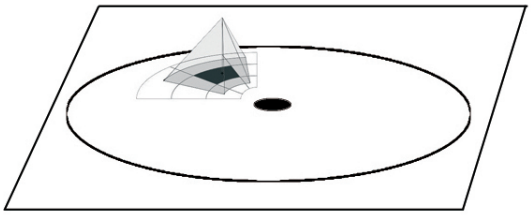


Fig. 13. The projected mask on the omnidirectional image with the central fuzzy element.

As in section 3.3 the projection of each fuzzy kernel element covers many omnidirectional image fuzzy pixels. This projection constitutes a pixel neighborhood. How to determine this neighborhood gray level? The solution presented in this article is to use the Choquet integral to give an upper and lower boundary of this gray level. Equations (6) and (7) are extended in 2D to be used in image processing. Equations remain the same but I_i stands for the gray level of the fuzzy pixel $\mathbf{P_i}$, (.) indicates a permutation such that $I_1 \leq I_2 \leq I_N$. The $A_n = \mathbf{P_n}, ..., \mathbf{P_N}$ are binary 2D coalitions of pixels whose gray levels are superior or equal to I_n. The biggest confidence that can be given to the coalition A_n is $v(A_n)$ defined in (9).

$$v(A_n) = Sup_{i=n..N} \Pi(\mathbf{P_i}, \mathbf{M_{u,v}}) \tag{9}$$

where $\mathbf{M_{u,v}}$ is the projection of a fuzzy Prewitt kernel element, $(\mathbf{P_i})_{\mathbf{i=1..N}}$ is the set of fuzzy pixels intersecting $\mathbf{M_{u,v}}$. The values of $\Pi(\mathbf{P_i}, \mathbf{M_{u,v}})$ are the intersection heights of $\mathbf{P_i}$ and $\mathbf{M_{u,v}}$ as illustrated in figure 14. The computation of $\Pi(\mathbf{P_i}, \mathbf{M_{u,v}})$ is made through a dichotomy approach.

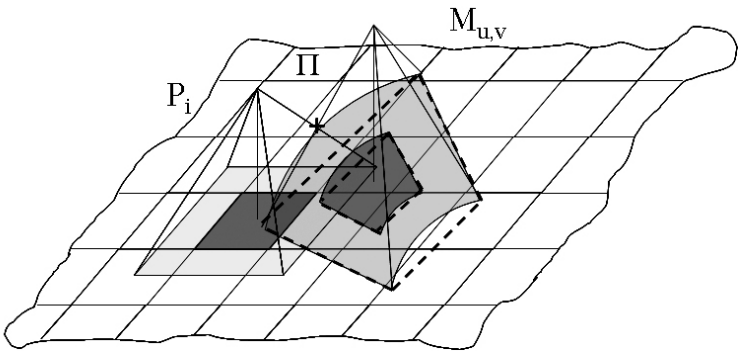


Fig. 14. Computation of $\Pi(P_i, \mathbf{M_{u,v}})$ as intersection height of $\mathbf{P_i}$ et $\mathbf{M_{u,v}}$. The dotted lines represent the approximation of the mask element.

Once the gray level of each kernel element is computed, the computation of the filtered pixel gray level still remains to be done. For kernel element the Choquet integral give a superior and inferior value. Two solutions are suggested.

- The first one consists in directly applying equation (3) with the kernel elements gray level values. The obtained images correspond to the superior and

the inferior estimations of the filtered image. Because of the very low difference between these images, the resulting image is the mean between the two obtained images.

- The second one consists in considering that the fuzzy kernel elements overlap each other. Indeed the kernel is also a discretized quantity, then the localization of kernel values is also imprecise. Therefore the Choquet integral on each fuzzy kernel element allows to give an upper and lower boundary of its gray level. Thus, the resulting filtered pixel gray level will also be characterized by an interval. The kernel gray level computation is relatively simple because the possibility representing the interaction of two contiguous kernel cells is always equal to 0.5 (Figure 15).

$$\Pi(M_{u,v}; M_{u',v'}) = 0.5 \tag{10}$$

with

$$\begin{cases} u' = u \pm k \\ v' = v \pm k' \\ k, k' \in \{-1, 0, 1\}. \end{cases}$$

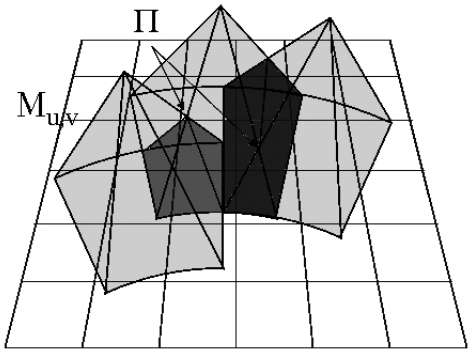


Fig. 15. Interactions between three overlapping fuzzy elements of the kernel.

4 Experiments

This section presents some results obtained with artificial and real images. For synthetic images, all parameters are controlled and for real images the sensor has been calibrated. In both cases an hyperbolic mirror was used. Five approaches are compared in this paper. The first one convolves a combination of the classical Prewitt kernels in all directions (horizontal, vertical and two oblique) with the omnidirectional image. The second one unfolds the image onto a cylinder and applies

the classical Prewitt kernel. Then, the filtered image is back-projected on the omnidirectional plane for comparison. The third one uses the proportional approach presented in this paper. The last approaches, present two possible extensions using fuzzy reasoning. These two variations depend if the Prewitt kernel is considered as an imprecise quantity or not. Indeed, the first one assumes a Prewitt kernel made of precise values and the second one considers each element of the kernel as an imprecise quantity. It uses a Choquet integral to deal with the overlapping of the these elements and to compute their values. Our filter is also compared with a well known optimal edge detector: the Deriche filter [26]. The Deriche filter is an infinite impulse response filter. Its adaptation to omnidirectional images is not easy. It is quite impossible to project it with our algorithm. Nevertheless, it is interesting to compare this optimal filter with our approach to know the performances an the limits of our edge detector.

In the subsection 4.1, artificial omnidirectional images are used to indicate the performances of the edge detector. We quantitatively evaluate the noise sensibility of our filter and the good localization of the detected edges. Real omnidirectional images are presented in subsection 4.2 to illustrate the different edge detectors responses in real conditions of use.

4.1 Quantitative evaluation of edge detector on synthetic images

The comparison between the different edge detector operators is difficult. The results can differ with the interactions between two edges, the luminance or the noise sensibility. Some performance evaluation methods have been proposed in the literature. In this article, two measures are used: the first given by Fram and Deutsch in [22,23], is based on the results analysis and the second, introduced by Canny in [24] based on the filter response analysis.

In this section, artificial images are used because of the total control of various parameters (camera, mirror, etc.). Indeed, calibration parameters such as phi and xi (or those of the camera) have a very crucial role in the projecting scheme. Of course, even if these two values are badly estimated, the edge detector will still work as an edge detector. However, detected edges will not be coherent with the real 3D scene. Thus, it is very important to use an efficient calibration algorithm to avoid mistakes in edge detection.

The benchmark used to test our algorithm is composed of four generated images containing intensities steps. Each image has a different intensity step value : 3,6, 12 and 20 gray levels (Figure16). The intensity of the darkest zone is 147 and remains invariant.

Two transitions are studied: a vertical transition and an oblique transition to 60° from the vertical. As images contain various kind of noise coming from acquisition

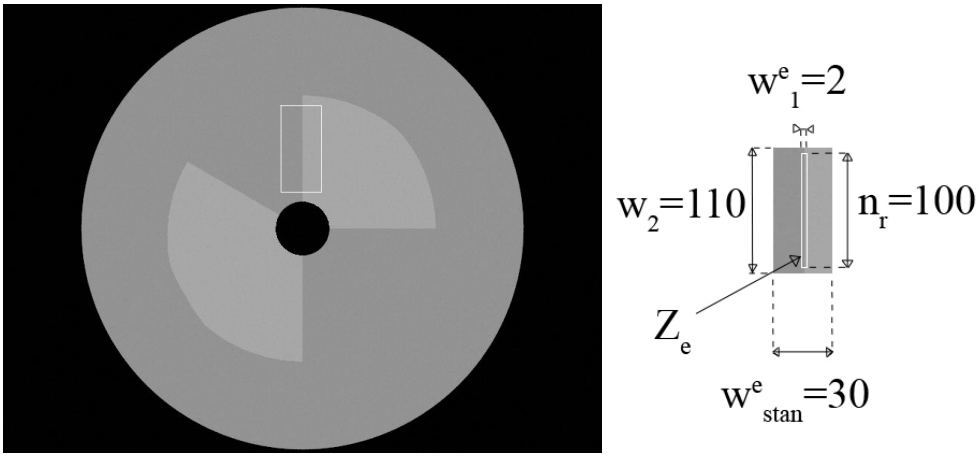


Fig. 16. (a) Noised artificial image with a 20 gray levels step. The considered region is delimited by a white rectangle. (b) Zoom on the edge region. The Fram and Deutsch parameters are detailed on this figure.

process or the 3D scene itself, any edge detection algorithm has to be robust against it. To model this behavior, test have been carried out by adding a gaussian noise of null average and standard deviation equal to two gray levels to the image.

To obtained significant edges, a thresholding step is necessary. This step remove the undesirable edges due to noise and preserve significant edges in the image. The results depend on the threshold value. The best threshold values provide the best ratio between desirable and undesirable edges. Nevertheless, these values are not easy to compute. We have to compromise between the over-detection and the under-detection. In this article we have chosen to extract the most complete edge, even if we detect some noise.

First of all, the edge detector performances can be evaluated using two parameters P_1 and P_2 defined by Fram and Deutsch. The first parameter value P_1 characterizes the freeness from noise of the edge detector output. The second parameter value P_2 measures the distribution of the output over the length of the edge. Both parameters responses are 1 for ideal performance and 0 for random output.

$$P_1 = \frac{n^e_{sig}}{n^e_{sig} + (n^e_{noise} + n^0)\frac{n_{in}}{f n_{tot}}} \tag{11}$$

where $n^e_{sig} = \frac{n^e - n^e_{noise}}{1 - \frac{n^e_{noise}}{n_{in}}}$ and $n^e_{noise} = n^0 \frac{n_{in}}{n_{out}}$

$$P_2 = \frac{\frac{n_r}{w_2} - \left\{1 - \left[1 - \frac{n^e_{noise}}{n_{in}}\right]^{w^e_1}\right\}}{\left[1 - \frac{n^e_{noise}}{n_{in}}\right]^{w^e_1}} \tag{12}$$

Let Z_e be the considered edge zone, n_{tot} the total number of points flagged as edge

points by the edge detector, n_{in} the part of n_{tot} inside the edge region Z_e and n_{out} the part outside the edge region ($n_{tot} = n_{in} + n_{out}$). n^0 and n^e are respectively the numbers of points flagged as edge points outside and inside the edge region after thresholding. w_1^e is the columns number of the edge region (2 columns). n_r is the row number of the edge region and w_2 is the total rows number of the considered zone. Finally, parameter f is used to normalize the edge detector output such as the points proportion between edge region and non-edge region is conserved: $f = \frac{w_1^e}{w_{stan}^e}$, where w_{stan}^e is the columns number (30 columns in our case).

Approaches	Vertical transition		Oblique transition	
	P1	P2	P1	P2
Deriche	0.9638	0.9199	0.9114	0.9231
Prewitt	0.8765	0.9091	0.7992	0.8991
Anamorphoses	1.0000	0.9231	0.7636	0.9181
Proportional	1.0000	0.9231	0.7919	0.9194
Fuzzy without interaction	0.9904	0.7821	0.9194	0.9155
Fuzzy with interactions	1.0000	0.9674	0.9213	0.9168

Table 1
P1 and P2 parameters values for oblique and vertical 20 gray level intensity step.

Table 1 gives some values for P_1 and P_2 obtained with an amplitude step of 20 gray levels.
With a vertical transition, fuzzy, proportional and anamorphosis approaches give the best results with respect to P_1 and P_2 parameters. Some noise points are detected as edges by the classical Prewitt filter, which penalizes the performances of the edge detection. The Deriche filter detects a double edge on the omnidirectional image, which reduces its P_1 score.
With an oblique transition, the performances are slightly different. The Deriche filter and the fuzzy approaches provides the best results and the edge is correctly detected. The classical Prewitt filter still highlights a lot of noise points as edge points and has the worst results for the P_2 parameter. These tests seem to indicate that taking into account the interactions between the kernel elements slightly improves the performances of the edge detector.

To test the robustness of the edge detection with respect to noise level, the artificial image depicted in Figure 16 has been corrupted by noise. Indeed, a Gaussian noise with an increasing variance σ (Figure 17) has been added to the image. The operators performances are quantified by the parameter P_1 which measures the sensitivity of the detector in presence of noise.

Only three algorithms are presented for this test: Deriche, classical Prewitt and Fuzzy. Figure 17 shows that the classical Prewitt filter is the most sensible to noise and can not correctly detect the edge when $\sigma > 10$. The Deriche approach detect an

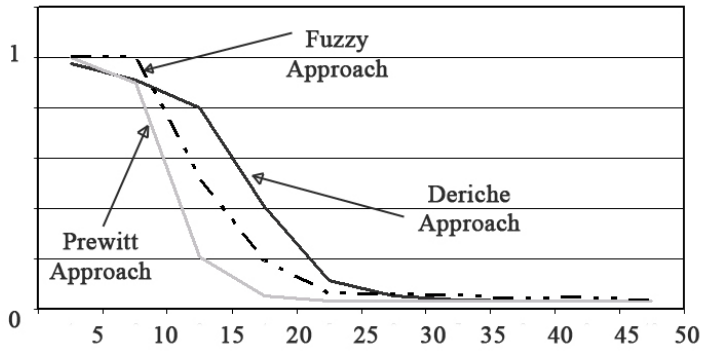


Fig. 17. Comparative P1 values on Gaussian noised image

optimal edges with $\sigma < 15$. The results decrease until $\sigma < 20$. After, edge and noise are confused. With our method, the results are better than with the Prewitt approach but remains less efficient than Deriche one. This could be explained by the fact that the Deriche operator is developed using optimal criteria. However, such detected edges are not consistent with the real 3-D scene because the Deriche operator does not consider the anamorphosis brought by the mirror.

To evaluate edge detector performances, Canny has defined three quantitative criteria:

- Good detection: They should be a low probability of failing to mark real edge points and low probability of falsely marking non-edge points. Since both probabilities are monotonically decreasing functions of the output signal-to noise ratio, this criterion corresponds to minimizing signal-to-noise ratio.
- Good localization: The points marked as edge points by the operator should be as close as possible to the center of the true edge.
- Only one response to a single edge: This is implicitly captured in the first criterion since when there are two responses to the same edge, one of them must be considered false.

These criteria stand for a reference in image processing field. To evaluate the presented methods performances, Canny's criteria will be used in the follow-up.

A method to quantify the operator performances, illustrated on figure 18, is to search the correlation between the detected and the theoretical edge. This method evaluate the rightness of the parameters and the good localization of the detected pixels. The results highly depend on the selected pixels after the thresholding procedure. Then, for each method, the best threshold has been chosen (the one providing the less noise points). Thus, the performances are less dependent of the thresholding procedure. The line minimizing the variance (deviation square sum) between the abscissa x_i of the pixels in Z_e and the theoretical line is computed.

$y = ax + b$ represent the theoretical edge with $a = tan60 = 1.7321$ and $b =$

$u_0 - (v_0 * a) = -119.6152$. If N is the detected pixels number, the parameters c and d from the regression line $x = cy + d$ are given by:

$$c = \frac{n \sum_i (x_i y_i) - \sum_i x_i \sum_i y_i}{N \sum(y^2) - \sum(y)^2} \tag{13}$$

$$d = \frac{\sum_i x_i - m \sum_i y_i}{N} \tag{14}$$

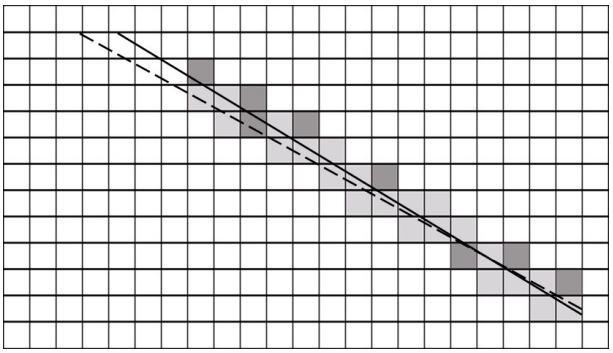


Fig. 18. Linear regression line. The gray area shows the pixels belonging to Z_e. The light gray pixels are detected by the operator and the dark gray pixels are not. The theoretical edge, represented by a dotted line, is computed with all pixels in Z_e. The linear regression line, represented by a dark line, with the pixels detected by the edge detector.

The error percentage on the theoretical edge is given by $|\frac{a-c}{a}|$. Figure 19 presents the error percentage on the theoretical edge obtained with different filters.

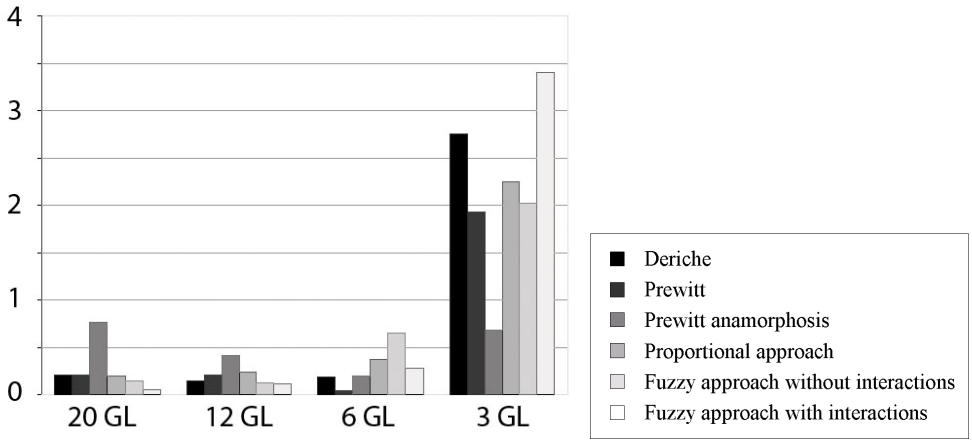


Fig. 19. Percentage of error on the slope for an amplitude step of 20, 12, 6 and 3 gray levels with five approaches

With a low amplitude step (inferior to 5 gray level), no algorithm correctly detects the edge. Indeed the edge is totally lost within the noise. Moreover, all approaches using a projection of the kernel introduce a kind of smoothing. This smoothing has catastrophic effects when signal to noise ratio is low. When the amplitude step is

bigger than 10 gray level, the error percentage decreases. The different approaches provide good results. With an increasing gray level value, the fuzzy approach gives the best results. The anamorphosis method gives the worst because of noise introduced by the retro-projection.

In our approach, all the projected masks can be pre-computed. In this case our computation time is equivalent to computation time obtained with a classical image processing.

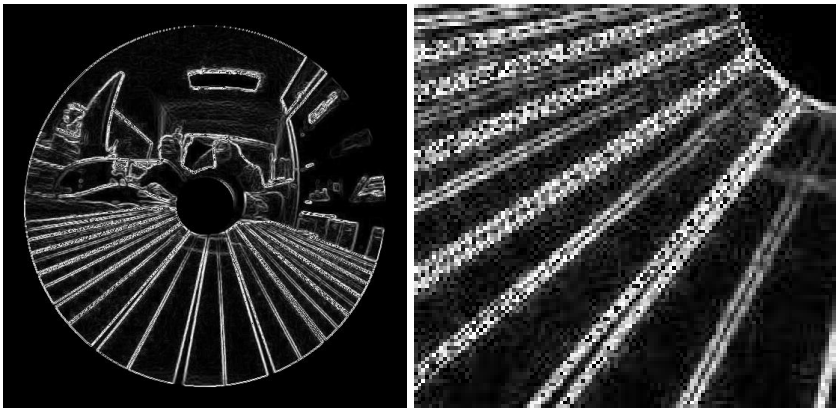
4.2 Real Images Experiments

Figure 20 presents one image used for real image testing. This image includes a grid made of black parallel lines. Half of the lines are thin, the other half are thick. Due to distortions brought by the projection onto the mirror, the black lines seems to get thinner as one gets closer to the image center.



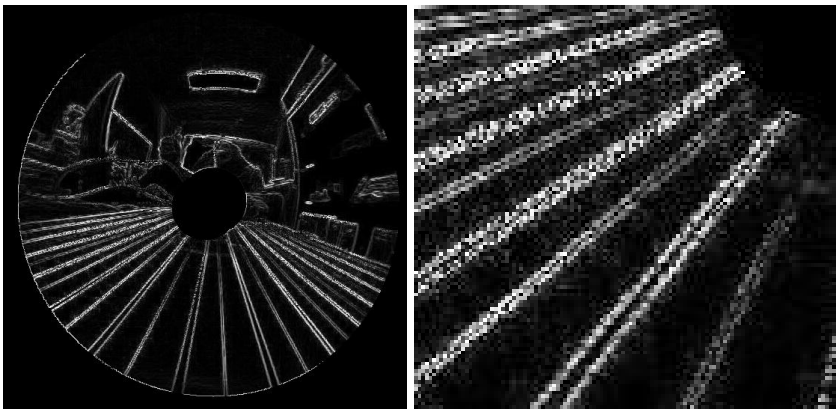
Fig. 20. Original image

We have illustrated a detail of the edges detected with different methods.
The classical Prewitt (Figure 21) edge detector is not adapted to omnidirectional images. The kernel provides a double edge for the thin lines. But concerning the thick lines, it highlights two boundaries at the image periphery that merge at the image center. Thus, the edge detector effects are not the same depending on its position on the image.
The Deriche edge detector (Figure 26) presents very good results. The double edges are correctly detected everywhere on the omnidirectional image. Nevertheless, the size of each line edge is the same from the periphery to the image center. The De-



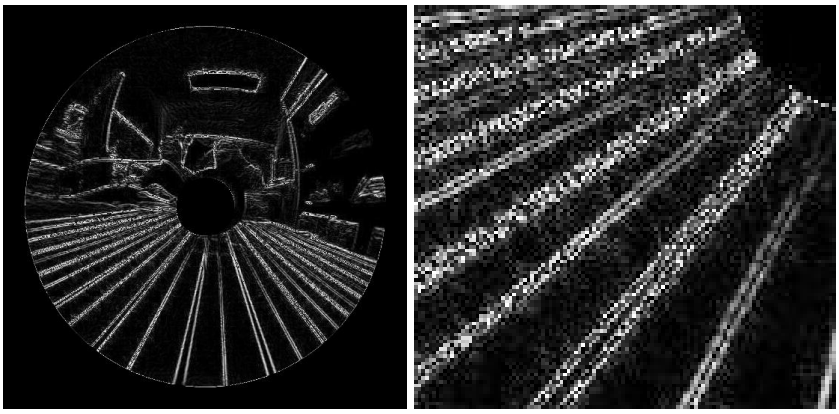
(a) (b)

Fig. 21. (a) Classical Prewitt filter on the omnidirectional image and (b) Zoom.



(a) (b)

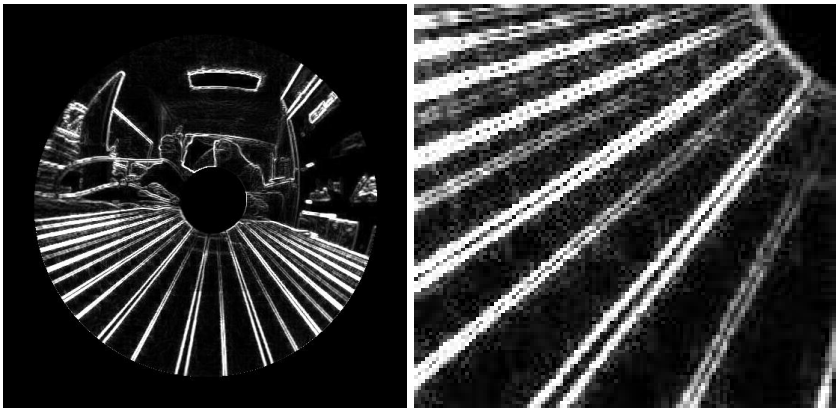
Fig. 22. (a) Classical Prewitt filter applied on the desanamorphosed image and (b) Zoom.



(a) (b)

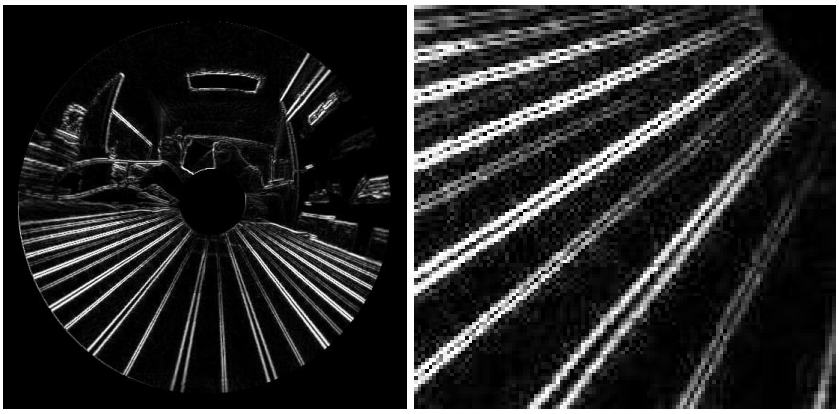
Fig. 23. (a) Prewitt filtering using the proportional approach presented in section 3.3 and (b) Zoom.

riche operator is not adapted the the reality of the 3-D scene captured by the camera. The projective approach (Figure 22) presents better results. The thin lines are high-



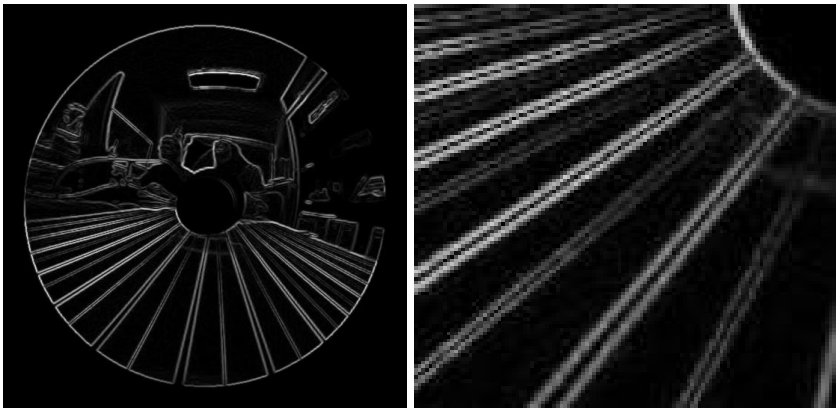
(a) (b)

Fig. 24. (a) Prewitt filtering using the fuzzy approach presented in section 3.4 without considering the interactions between the fuzzy kernel elements of the mask and (b) Zoom.



(a) (b)

Fig. 25. (a) Prewitt filtering using the fuzzy approach presented in section 3.4 and (b) Zoom



(a) (b)

Fig. 26. (a) Deriche filter applied on the omnidirectional image and (b) Zoom

lighted as in the classical approach. However, the thick lines two boundaries are detected from the periphery to the center of the image. Due to projection and

back-projection interpolation effects, the edges are a little noisy. The proportional approach presented in section 3.3, provides slightly better results than with the projective one (Figure 23). Results are less noisy, especially near the image center, because only kernels are projected. Indeed, no interpolation effects are visible. However, it is noticeable that the edge width varies depending on its location on the image. This is due to the imprecision of the Prewitt kernel cells. Indeed, the projection of a kernel cell to the outer image boundary covers a lot of pixels, whereas it covers only few pixel near the image center.
The fuzzy approach (Figure 25) presents the best results. The edges are correctly identified on the omnidirectional image. The two edges of the thick lines are well highlighted and remain separated from the periphery to the image center. As with proportional approach, the thickness of the edge depends on its position on the image because of the kernel imprecision.

Even if the optimal gradient operators (such as Canny-Deriche) provide a good edge detection on omnidirectional image, their behavior remains the same all over the image as if it was a perspective one. They discard the resolution variations. Conversely, our approach keeps the same behavior on the cylindrical projective space. This space, locally close to a perspective image, allows to detect edges in a coherent manner with respect to the real scene. Indeed, the edge thickness depends on the radial position on the omnidirectional image.

5 Conclusions

This article addresses the problem of image processing for omnidirectional images. Classical operators using kernels such as Prewitt edge detector are designed for scenes belonging to a camera fronto-parallel plane. The kernels are usually regularly sampled using a step equivalent to the acquisition device sampling step. Thus, their effects on the image are considered invariant to translation. Omnidirectional images present strong radial resolution variations. Therefore, classical operators are not adapted as their effect will depend on their position on the omnidirectional image. A solution using the sensor geometry to define the operators has been presented. The filtering operator is defined on a regular space: the surrounding cylinder, and back-projected on the omnidirectional image. This method avoids projecting the image onto another space, thus introducing errors or approximations on data. Moreover, fuzzy quantities have been used to deal with localization imprecisions brought by the image sampling effects. Tests carried out have shown that this method allows a coherent edge detection on omnidirectional image. The capacity to distinguish two close edges is conserved whatever their positions on the image may be. Some improvements are already planned. For example, the approximation of the projected kernel needs to be improved. In this case, a compromise has to be reached between the precision and computation complexity. Another interesting question concerns the quantitative criteria which evaluate the performance of

the edge detector. They are not adapted to the omnidirectional image geometry. An adaptation of Fram and Deutsch (or other ones) quantitative criteria for omnidirectionnal image is another real challenge and research is being carried out. The proposed approach can be extended to all discrete summative kernel such as Sobel, Laplacian, gaussian filter, smoothing filter, median filter, etc.

References

[1] E. Hecht, A. Zajac, *Optics*, Third edition, New York, 1994.

[2] M. Fiala, A. Basu, Feature extraction and calibration for stereo reconstruction using non-SVP optics in a panoramic stereo-vision sensor, *Proceedings of the Third Workshop on Omnidirectional Vision, Washington DC, USA*, 79–86, 2002.

[3] T. Svoboda, T. Pajdla, V. Hlavac, Motion estimation using central panoramic cameras, *In IEEE International Conference on Intelligent Vehicles*, 1998.

[4] S. Baker, S. K. Nayar, A Theory of Single-Viewpoint Catadioptric Image Formation, *International Journal of Computer Vision*, Kluwer Academic Publishers, Hingham MA, USA, Volume 35, Number 2, Pages 175–196, 1999.

[5] C. Geyer, K. Daniilidis, Catadioptric projective geometry, *International Journal of Computer Vision*, Volume 43, Pages 223–243, 2001.

[6] K. Daniilidis, A. Makadia, T. Bulow, Image Processing in Catadioptric Planes: Spatiotemporal Derivatives and Optical Flow Computation, *Proceedings of the Third Workshop on Omnidirectional Vision*, Washington DC, USA, Pages 3–10, 2002.

[7] D. Dubois, H. Prade, Plenum Press, *Possibility Theory An Approach to Computerized Processing of Uncertainty*, 1988.

[8] L. A. Zadeh, Knowledge Representation in Fuzzy Logic, *An Introduction to Fuzzy Logic Applications in Intelligent Systems*, R. R. Yager, L. A. Zadeh editors, Kluwer publisher, Boston, Pages 1–25, 1992.

[9] D. Dubois, L. Foulloy, G. Mauris, H. Prade, Probability-Possibility Transformations, Triangular Fuzzy Sets, and Probabilistic Inequalities, *Reliable Computing*, Volume 10, Number 4, Pages 273–297, 2004

[10] O. Strauss, F. Comby, Fuzzy Morphology for Omnidirectional Images, *Proceedings on the International Conference on Image Processing*, Genova Italy, September 11-14, Volume 2, Pages 141–144, 2005.

[11] J.L. Marichal, An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria, *IEEE Transactions on Fuzzy Systems*, Volume 8, Number = 6, Pages 800–807, 2000

[12] I. L. Herlin, N. Ayache, A New Methodology to Analyse Time Sequences of Ultrasound Images, *Rapports de Recherche No 1390*, Rocquencourt, 1991

[13] I. Bloch, H. Maître, Fuzzy Mathematical Morphology., *Annals of Mathematics and Artificial Intelligence*, Volume 10, Number = 1 - 2, Pages 55–84, 1993.

[14] P. Smets, *Uncertainty in Artificial Intelligence 5*, Pages 29–39, M. Henrion, R. D. Schachter, L. N. Kanal, J. F. Lemmer editors, Elsevier Science Publishers, North-Holland Amsterdam, 1990

[15] H. Bakstein, T. Pajdla, Panoramic mosaicing with a 180° field of view lens, *Proceedings of the IEEE workshop on omnidirectional vision*, Copenhagen Denmark, June 2, Pages 60–67, 2002.

[16] S. Peleg, M. Ban-Ezra, Stereo panorama with a single camera, *Proceedings of the IEEE conference on computer vision and pattern recognition*, June, Pages 395–401, 1999.

[17] S. K. Nayar, A. Karmarkar, 360*360 mosaics, *Proceedings of the IEEE conference on computer vision and pattern recognition*, Hilton head, South Carolina, June, Volume 2, Pages 388–395, 2000.

[18] N. Vlassis, Y. Motomura, I. Hara, H. Asoh, T. Matsui, Edge-based features from omnidirectional images for robot localization, *Proc. IEEE Int. Conf. on Robotics and Automation*, Seoul,Korea, 2001.

[19] O. Shakernia, R. Vidal, S. Sastry, Omnidirectional egomotion estimation from back-projection flow, *OMNIVIS 2003* 2003.

[20] C. Demonceaux, D. Kachi-Akkouche, Optical flow estimation in omnidirectional images using wavelet approach, *Conference on Computer Vision and Pattern Recognition Workshop*, Volume 7, Page 76, 2003.

[21] S. Baker, S. K. Nayar, Single Viewpoint Catadioptric Cameras, *Panoramic Vision: Sensors, Theory, Applications*, Ryad Benosman and Sing Bing Kang editors, Springer-Verlag publisher, 2001.

[22] Deutsch, E. S., J. R. Fram, A quantitative study of the Orientational Bias of some Edge Detector Schemes, *IEEE Transactions on Computers*, March, 1978.

[23] Fram, J.R., E.S. Deutsch, On the quantitative evaluation of edge detection schemes and their comparisions with human performance, *IEEE Transaction on Computers*, Volume C-24, Number 6, Pages 616–627, 1975.

[24] J.F. Canny, A computational approach to edge detection, *IEEE Trans Pattern Analysis and Machine Intelligence*, Volule 8, Number 6, Page 679–698, November, 1986.

[25] P. Baker, C. Fermüller, Y. Aloimonos, R. Pless, A spherical eye from multiple cameras (makes better model of the world), *Proceedings of the conference on computer vision and pattern recognition*, Kauaii, USA, Volume 1, Pages 576–583, December, 2001.

[26] R. Deriche, Fast Algorithms for Low-Level Vision, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Volume 1, Pages 78–88, Number 12, January, 1990.

[27] S. Bigot, D. Kachi, S. Durand, E. Mouaddib, Spherical Image Denoising and Its Application to Omnidirectional Imaging, *2nd International Conference on Computer Vision Theory and Applications*, Barcelona, Spain, Pages 8–11, March, 2007.

[28] I. Bogdanova, X. Bresson, J. Thiran, P. Vandergheynst, Scale-space analysis and active contours for omnidirectional images, *IEEE Transactions on Image Processing*, 2007.

[29] I. Tosic, I. Bogdanova, P. Frossard, P. Vandergheynst, Multiresolution Motion Estimation for Omnidirectional Images, *Proceedings of EUSIPCO*, 2005.

[30] C. Demonceaux, P. Vasseur, Markov Random Fields for Catadioptric Image, *Processing. Pattern Recognition Letters*, Elsevier, Volume 27, Pages 1957–1967, December, 2006.