

## Research Article

# Curvilinear Image Regions Detection: Applications to Mobile Robotics

**J. M. Perez-Lorenzo,<sup>1</sup> A. Bandera,<sup>2</sup> R. Marfil,<sup>2</sup> and R. Vázquez-Martín<sup>3</sup>**

<sup>1</sup>Departamento de Ingeniería Telecomunicación, Universidad de Jaén, 23700-Linares, Spain

<sup>2</sup>Grupo ISIS, Departamento de Tecnología Electrónica, Universidad de Málaga, 29071 Málaga, Spain

<sup>3</sup>Centro Andaluz Innovación y Tecnología de Información y Comunicaciones (CITIC), 29590-Málaga, Spain

Correspondence should be addressed to J. M. Perez-Lorenzo, jmperez@ujaen.es

Received 15 June 2010; Revised 25 August 2010; Accepted 19 October 2010

Academic Editor: Yannis Kopsinis

Copyright © 2011 J. M. Perez-Lorenzo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a novel approach for visual features detection, which is based on the presence of objects whose shape can be modelled using cylinders or generalized cylinders. These specific structures are commonly found on indoor and outdoor scenarios, and their image representations, the so-called curvilinear regions, automatically deform with changing viewpoint as to keep on covering identical physical parts of a scene. The method is based on Marr's visual theory that proposes that visual objects can be decomposed in generalized cylinders. Also, part of the method can be compared to the behavior of AOS neurons, placed in the caudal intraparietal sulcus, that respond when an elongated object is visualized. Our detector reliably finds the same curvilinear regions under different viewing conditions. Evaluation results are given to demonstrate the performance of the approach and its ability to be applied for visual features detection in a mobile robot navigation framework.

## 1. Introduction

An autonomous robot must be capable of managing real situations in dynamic and complex environments and of interacting with people and/or other robots. Among other components, navigation is a fundamental capacity that requires the integration of different modules. Self-localization and environment mapping are two that are essential, as they are needed at different levels from low-level control to higher-level strategic decision making or navigation supervision. In order to solve the localization and mapping problems, it is necessary to describe the robot's environment. One popular choice is to use a feature-based map. These maps represent the environment by a set of features or landmarks, such as lines, points, or circles. Localization in these maps is performed by extracting features from the sensor data and associating them with the features in the map. Recognizable landmarks are then essential since they will be used as reference marks to identify locations in the world [1]. There are some advantages to using these maps; their representation is more compact and accurate,

is better suited to representing dynamic objects, and is closer to people's perception than other approaches, such as grid maps. Besides, these approaches allow one to use multiple models to describe the measurement process from different sensors and parts of the environment and to represent the uncertainty for the map and the robot pose. Feature maps are particularly appropriate for structured environments, but it is difficult to model unstructured ones.

Feature-based navigation approaches mainly differ in the method employed to represent the belief of the mobile robot about its current pose or to find and track a safe path to a goal. Furthermore, they can be differentiated according to the type of sensor information that they use. Although range and bearing sensors have been successfully used for feature extraction in mobile robot applications, vision systems constitute an interesting alternative. These systems are passive and of high resolution, and they provide a huge amount of information (color, texture, or shape) which allows disambiguating landmarks for subsequent data association purposes [2]. In this framework, this paper describes a vision-based approach for features detection.

The proposed method exploits a particular object structure, that which can be modelled using cylinders or generalized cylinders. The choice of this type of structure is based on Marr's visual theory, being the biological inspiration of this work. Marr suggests that visual objects can be divided into generalized cylinders that are combined to form objects [3] and therefore can be used to identify and recognize them. The biological inspiration for the choice of the visual features will allow to use regions with a high semantic significance for future applications. The detection of these cylinders in the image could be considered the main disadvantage of the approach, as it depends on the presence of these specific structured landmarks in the scene. However, as it has been aforementioned, feature-based maps allow one to employ different models to describe the perceived environment. In a mobile robot navigation system, this detector could be combined with other visual feature detectors to increase the number of extracted landmarks.

The image representation of each detected landmark will be a curvilinear region. A curvilinear structure represents a line or curve with some width [4]. Thus, curvilinear features differ from conventional line or curve features, which are typically obtained based on edges, and they contain more information. Besides, they constitute semilocal structural items which are more robust to intensity, color, or pose changes than local interest points or intensity-based interest operators [5]. In our approach, curvilinear regions are chosen from the image partition generated by the modified version of the Bounded Irregular Pyramid (BIP) proposed by Vázquez-Martín et al. [2]. The uBIP is a hierarchical structure, whose data structure merges the regular and irregular data structures to obtain a scale-space representation of the input image. In this hierarchy, each level is a reduced version of the level below, and it is constituted by nodes which are joined to irregular-shaped regions at the input image [2, 6]. This decimation process allows one to obtain the set of nodes whose associated regions define a domain-independent segmentation of the image pixels into regions. These nodes are located at different scales, and their associated regions are perceptually homogeneous according to certain criteria. The shape of these regions is adapted to real items of the scene. Hence, continuous geometric changes of the view point transform pixels from a single connected region to a new single one (see Vázquez-Martín et al. [2] for further details). This is one of the main differences with respect to other curvilinear detectors, which typically employ the first steps of Steger's curvilinear detector algorithm [7]. From the obtained image partition, the proposed approach chooses the set of curvilinear regions applying a set of geometric and chromatic conditions. Experimental results show that this detector could be used to build sparse maps, where landmarks are perceptually distinguished and, besides, they usually have an underlying semantic significance.

The paper is organized as follows. After discussing related work to the proposed detector and the biological inspiration in Section 2, the approach for the acquisition of curvilinear visual landmarks is described in Section 3. Section 4 deals with some obtained experimental results.

Finally, the paper concludes along with discussions and future work in Section 5.

## 2. Related Work and Biological Inspiration

Feature-based vision systems for mobile robot localization and navigation identify each scenario or environment pose with a set of features and their spatial distribution. These landmarks must own some invariant and stable properties in order to be detected with high repeatability in images taken from arbitrary viewpoints. Visual feature detectors can be classified into intensity-based detectors and structure-based detectors. The advantage of intensity-based detectors is that no model of landmarks has to be specified to the vision system a priori. Besides, these approaches can generate dense occupancy maps, but comprising of landmarks which do not necessarily present an underlying semantic significance. The disadvantage of such systems is scalability. Thus, these systems are usually implemented in environments, where the number of detected landmarks is relatively small [8]. Moving the robots to a larger environment requires the management and recognition of a much larger number of landmarks. An excessively huge number of landmarks can provoke that the reliability and repeatability of visual features can not always be guaranteed, appearing outliers in feature matching which can lead to unreliable data association [9]. This problem has been addressed by grouping local interest points together and using these groups as landmarks [9] or by imposing a fixed number of landmarks (e.g., the iterative SIFT [10]). Other solution has been suggested by model-based visual landmark detectors. These detectors are employed to build sparse maps using landmarks that have an underlying semantic significance. Thus, image edges [11] or planar quadrangles [12, 13] can be employed to match images. Environment-specific features like walls or doors are used by Horswill [14]. To deal with outdoor environments, Asmar et al. [8] propose a tree detection approach.

In our work, although curvilinear regions are based on the existence of generalized cylinders, a 3D framework is not built. In fact, this work is more similar to the intensity-based models because it works with the projections of the features. These projections are transformed in a way that can be interpreted as an extended intensity-based approach according to Tarr and Bulthof [15]. Given a set of images taken from different viewpoints, the process of finding the projections on each image of real 3D surface patches must deal with the problem that image regions associated to the projections change covariantly with the class of transformation induced by the viewpoint change. When the viewpoint change can be approximated by an affine transformation, approaches which solve this problem are called affine region detectors. The detection of regions which change covariantly with affine transformations was described in detail by Mikolajczyk et al. [16]. In this work, the authors provide a review of affine covariant region detectors and compare their performance on a set of test images under varying imaging conditions. The requirement for these detectors is that they must provide regions whose

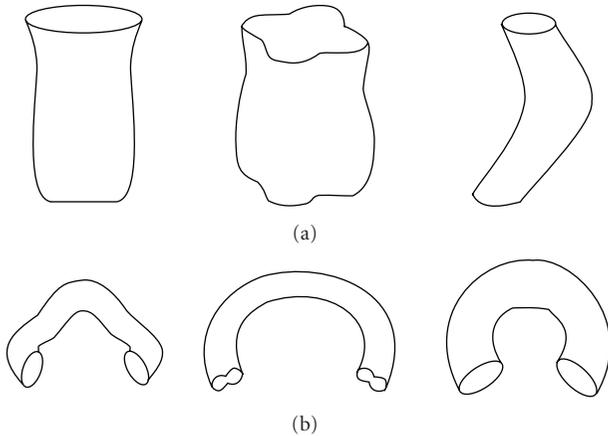


FIGURE 1: Several examples of generalized cylinders. (a) Straight homogeneous generalized cylinders. (b) Planar, right constant cross section generalized cylinders.

shapes depend on the underlying image features, so that they correspond to the projections of the same 3D surface patch on the different images. Although the boundaries of these covariant regions do not have to be associated to changes in image features such as color or texture, some of the approaches described in Mikolajczyk et al. [16] look for these abrupt changes. Thus, the intensity extrema-based region detector (IBR) [17] starts from intensity extrema and studies an intensity-based function along rays emanating from this extrema to define a region of arbitrary shape. The region is delineated by the image points defined over these rays, where the intensity suddenly increases or decreases. A maximally stable extremal region (MSER) [18] is a connected component of an appropriately thresholded image, where all internal pixels have either higher or lower intensity than all the pixels on its outer boundary. Among these extremal regions, the “maximally stable” ones are those corresponding to thresholds, where the relative area change as a function of relative change of threshold is at a local minimum. The Principal Curvature-based regions (PCBRs) detector [5] extracts stable regions within the multiscale principal curvature image. Contrary to other intensity-based detectors, PCBR overcomes local intensity variations within regions by focusing on region boundaries rather than the appearance of region interiors. This paper employs an approach for affine region detection which extends the idea of looking for abrupt changes. However, to detect these boundaries, we use a hierarchical clustering scheme, which groups neighbor image pixels into blobs of homogeneous color. The hierarchical clustering algorithm represents the input image at different levels with decreasing resolution. This hierarchy constitutes a scale-space representation, where salient regions could be detected at different scales. The geometry of salient regions is analyzed to look for those regions which can be the projection of a generalized cylinder.

Classical visual theories like Biederman’s propose that objects can be divided in their constituent components, and so Recognition By Components (RBCs) can be achieved [19].

Marr’s theory suggests that these components are generalized cylinders [3]. Our work is based on the presence of these semantical generalized cylinders in every image, and we consider their semantic significance the main advantage over other invariant region detectors. In [20], one of the first studies with this kind of visual objects is presented. A generalized cylinder can be defined as an arbitrary planar shape, called a *cross section*, swept along an arbitrary 3D curve, called an *axis*. Shafer in [21] uses terms as SHGCs (Straight Homogeneous Generalized Cylinders) to describe generalized cylinders with a straight axis and cross sections of a fixed shape, but of varying size; CGCs (Constant cross section Generalized Cylinders), with a fixed cross section and PRCGCs (Planar, Right Constant cross section Generalized Cylinders), where the axis is planar and the cross sections are orthogonal to the axis (Figure 1). PRCGCs provide an important modeling tool for mostly “snake-like” objects that cannot be modeled by SHGCs [22]. Ulupinar and Nevatia in [22] analyze the properties of these types of cylinders and derive the types of symmetries that the limb boundaries and cross sections of these objects produce on the image plane. Nevatia assures that, combining the detection of this types of cylinders, it is possible to cover a large fraction of the surfaces found in man-made environments. Nevatia’s work is based on the analysis of the extracted edges, the straightness and parallelism properties of lines, and the symmetry of curves. Our work is also based on the analysis of symmetries. However, our detection is more tolerant because it allows several deformations and errors in the boundaries, so we think it is better for working with real images. Nevatia needs a perfect or ideal extraction of the borders of the object and also at least one cross section of the cylinder must be drawn in the image, and this is not always possible when working with real images. Moreover, they need to build the 3D model in the detection process, so the method is slow. In [23], a better method is described to compute the desired descriptions of complex objects from a single image. This is achieved by exploiting projective properties of a class of generalized cylinders and of possible joints between them. However, also at least one cross section of the cylinder must be visible in the borders image. Also a 3D model of the parts of the objects is built, and this is not needed in our approach. Furthermore, the method may not handle two-dimensional compound objects [23].

Other authors as Rao and Medioni have also studied some subclasses of PRCGCs like a torus [24], which can be modeled as a PRCGC with circular cross section and circular axis. They use the property that limb edges of a torus produce parallel symmetry under orthographic projection. They also investigate for which classes of generalized cylinders the contour is symmetric about the projection of the 3D axis. They conclude that the contour of a solid of revolution is symmetric about the projection of its axis for any view. Medioni in [25, 26] uses a contour representation with a symmetry axis. As assured in [26], it is well known that the projections of 3-D mirror symmetric coplanar curves form a skew symmetry in the image plane. The method is based also on a previous detection of edges of the image, which are modeled with B-splines. The method works with these

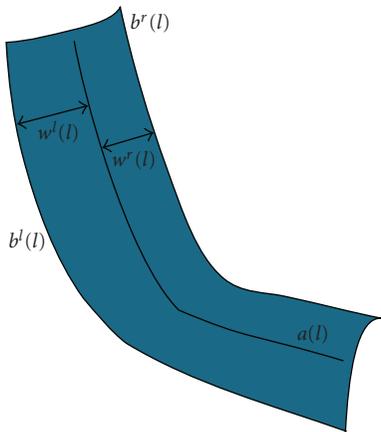


FIGURE 2: Definition of curvilinear region for a continuous case.  $b^l(l)$  and  $b^r(l)$  are the left and right borders, respectively;  $a(l)$  is the medial axis; and  $w^l(l)$  and  $w^r(l)$  are the left and right widths, respectively.

B-splines instead of the image borders, and the symmetry axes are estimated. These axes allow to divide the object in a hierarchical way. The method needs to analyze the object from a local and a global point of view. This can lead to conflicts that should be resulted in a higher level of knowledge. Also, no much experiments with real images are shown, and problems can appear with contours with very low curvature.

To overcome some limitations when working with real images, Nevatia in [27] uses projections of generalized cones (GCs) [20] known also as *ribbons*. For a ribbon, the axis is an arbitrary 2D curve, and the cross sections are simply line segments. As Nevatia assures, 2D ribbons may be viewed as projections of 3D GCs, and given only a single intensity image, it is easier to compute ribbons which may serve as a step toward inferring 3D volume descriptions [23]. Based on the analysis of the medial axis and symmetries, Nevatia makes a scene segmentation detecting the ribbons on the image. Later, these ribbons are grouped into superribbons, objects, and superobjects. Our method has got similarities with this work but also differs from this in several aspects. First, Nevatia works with the extracted edges from the image, and our method is based on a colour segmentation. Also, for estimating the medial axis, Nevatia needs to work with the image rotated in  $m$  equally spaced directions. Our method extracts the medial axis using a distance transform, and there is no need to rotate the image to detect the skeletons to get reliable results. Another difference is that the application proposed in our paper is oriented to detect landmarks for localization, so there is no need to form entire objects or the superribbons described in Nevatia's work. The curvilinear regions that we propose have strong similarities with ribbons; however, the computation of the curvilinear properties that we propose is a novel approach. Our method allows some deformations in the symmetries, in order to obtain a better detection with real images.

Nowadays, there are other methods that detect curvilinear regions although in the literature these methods are more focused on other applications of artificial vision, such as detection of roads in aerial images and blood vessels in biomedical applications. We can find an example in [28], where Liu et al. propose a wide line detector based on isotropic responses via circular masks. The method can work quite well with some types of images. However, the size of the width that a line must have in order to be detected depends on the size of the chosen mask, so those lines whose width is above an established threshold can not be detected. So, the method requires the maximum width of the lines to be estimated before the detection. Liu for his experiments develops a method that detects those lines with a length very high related to its width. We believe that to consider only those types of curvilinear regions is insufficient for our application because it establishes a very restrictive condition. Liu's application is focused on the detection of quite elongated objects such as blood vessels and palmprint images. We think our detection algorithm is able to find a more diversity of regions. In Liu's method, the size of the mask could be increased, but this could lead to an artificial joint of independent regions. Other drawback is the speed of the algorithm since it involves large mask comparison operations for every image pixel. Li et al. in [29] presents a faster method. It consists of two components: a small object detector and a centerline detector. The method uses a fast kernel-based density estimation which is termed as local weighted features (LWFs). The work is focused on the detection of objects whose widths are less than 10 pixels. They eliminate the wide borders created by large objects, so we think that it would not detect many objects that could be of interest for our application. This method has been proved for the detection of roads and rivers in aerial images, but not with the goal to extract natural landmarks of a scene for localization in robotics. Indeed, the authors assert that they still have to apply it to other image processing tasks.

Also, a method that detects elongated objects has got by itself a strong biological inspiration. It has been clearly shown in the visual neuroscience literature that there is a type of neurons, called AOS (axis orientation selective) and placed in the caudal intraparietal sulcus (CIP) area, that responds when an elongated object is visualized [30, 31]. This response is suggested to be used for grasping objects and some recognition tasks. In [32], a method for emulating the AOS neurons behaviour is implemented for a robotic system, and the authors use a short descriptor for an object recognition purpose. However, their algorithm is simpler than the proposed in our work, and we think our work can be used for a wider variety of elongated objects.

### 3. Curvilinear Regions Detection

**3.1. Definition.** We have represented in Figure 2 a region delimited by left and right boundaries,  $b^r(l)$  and  $b^l(l)$ , in a continuous case. A curvilinear region could be represented by a parameter vector  $\{a(l), w^l(l), w^r(l)\}_{l=0\dots L}$ , where  $L$  is the length of the region,  $a(l)$  is a vector defining the axis between

the right and left borders, and  $w^r(l)$  and  $w^l(l)$  are the widths of the curvilinear region. Rao and Medioni in [24] proved that the axis of some right generalized cones, like solids of revolution, projects to the axes of right ribbons, so the axis  $a(l)$  can provide an important information about the region. In order to decide if the set of pixels is a curvilinear region or not, we define four properties, called *curvilinear properties*, that must be satisfied by every curvilinear region.

- (i) *Symmetry*: there must be a geometric similarity around the region axis.
- (ii) *Elongation*: the ratio between the average width and the total length must be less than a predefined threshold.
- (iii) *Parallelism*: the left and right borders must be locally parallel.
- (iv) *Homogeneity*: the colour of the region should be homogeneous.

The first three properties are geometrical properties whereas the last one imposes a colour restriction. This constraint eases the detection stage, and it can be achieved by a colour segmentation of the original image.

The algorithm for detecting curvilinear regions can be divided into several steps, depicted in Figure 3. Firstly, the original image is segmented into homogeneous colour regions using a pyramid algorithm based on the Bounded Irregular Pyramid (BIP) [6]. By this way, the obtained regions comply with the *homogeneity* property. Next, a skeleton extraction based on a distance transform is achieved inside the regions, and the geometric properties of every region are checked looking for those regions with the *symmetry*, *elongation*, and *parallelism* properties. The approach obtains the set of curvilinear regions from this analysis. Image segmentation and the computing of the properties are described in Sections 3.2 and 3.3.

**3.2. Image Segmentation.** The segmentation step employs a colour distance to group the image pixels into a set of blobs, whose spatial distribution is physically representative of the image content. To accomplish this grouping process, the contents of the input image can be described using multiple representations with decreasing resolution. Pyramids are hierarchical structures which have been widely used to represent the perceptual organization of the image by a tree of regions, ordered by inclusion [33]. In this hierarchy, each level is a graph which is at least defined by a set of nodes, which represent regions, connected by a set of arcs, which represent region adjacency relationships. The efficiency of a pyramid to represent the information is strongly influenced by two features: the graph selected to encode the information within each pyramid level and the decimation scheme used to build one graph from the graph below [33]. The choice of a graph encoding determines the information that may be encoded explicitly at each level of the pyramid. On the other hand, the reduction or decimation scheme used to build the pyramid determines the

dynamic of the pyramid (height, preservation of details...). Depending on these two features, pyramids have been classified as regular and irregular ones. Regular approaches have a rigid structure, where the decimation process is fixed. This rigid structure allows to build and process them with a low computational cost. However, this inflexibility can also provoke three main problems: nonconnectivity of the obtained regions, impossibility to represent elongated objects, and shift variance [33]. Irregular pyramids solve these problems using a structure which dynamically adapts to the image layout. However, they require a computational time which is usually higher than the one required by regular pyramids. In order to combine the advantages of regular and irregular pyramids, the Bounded Irregular Pyramid (BIP) was proposed by [34]. The BIP arose as a mixture of regular and irregular structures, whose goal is to obtain accurate results at a low computational cost.

The BIP approximates or even outperforms previously proposed hierarchical segmentation schemes, yet it can be computed much faster [33]. However, it is highly affected by the shift variance problem, that is, it provides an image segmentation which varies when this image is shifted slightly. In this paper, we use the uBIP, an extension of the original BIP structure, which improves the mixture of the regular and irregular decimation processes [2]. The uBIP uses a union-find algorithm to merge the nodes resulting of the regular and irregular decimation processes, allowing that a node of the structure (regular or irregular) can be linked with any type of nodes from its same level.

As it is described in [6], the pixels of the input image can be considered as the nodes of the graph  $G_0$ . Then, the segmentation stage divides the image into regions of uniform color using the uBIP. Contrary to the BIP, this decimation algorithm only runs two consecutive steps to obtain the set of nodes  $N_{l+1}$ . The first process generates the set of regular nodes of  $G_{l+1}$  from the regular nodes at  $G_l$ , meanwhile the second one determines the set of irregular nodes at level  $l+1$ . In this proposal, this second process conducts a union-find decimation algorithm which is simultaneously conducted over the whole set of regular and irregular nodes of  $G_l$  which do not present a parent in the level  $l+1$ .

Let  $G_l = (N_l, E_l)$  be a graph, where  $N_l$  stands for the set of regular and irregular nodes and  $E_l$  for the set of intralevel arcs. Let  $\varepsilon_l^{xy}$  be equal to 1 if  $(x, y) \in E_l$  and equal to 0 otherwise. Let  $\xi_x$  be the neighborhood of the node  $x$  defined as  $\{y \in N_l : \varepsilon_l^{xy}\}$ . It can be noted that a given node  $x$  is not a member of its neighborhood, which can be composed by regular and irregular nodes. Each node  $x$  has associated a  $v_x$  value. Besides, each regular node has associated a boolean value  $h_x$ : the homogeneity [6]. At the base level of the hierarchy,  $G_0$ , all nodes are regular, and they have  $h_x$  equal to 1. Only regular nodes which have  $h_x$  equal to 1 are considered to be part of the regular structure. Regular nodes with a homogeneity value equal to 0 are not considered for further processing. The proposed decimation process transforms the graph  $G_l$  in  $G_{l+1}$ , such that the reduction factor is greater than 1. In our case, we focus on dividing the image into a set of homogeneous blobs. This aim is achieved using the pairwise comparison of neighboring nodes. Then, a pairwise

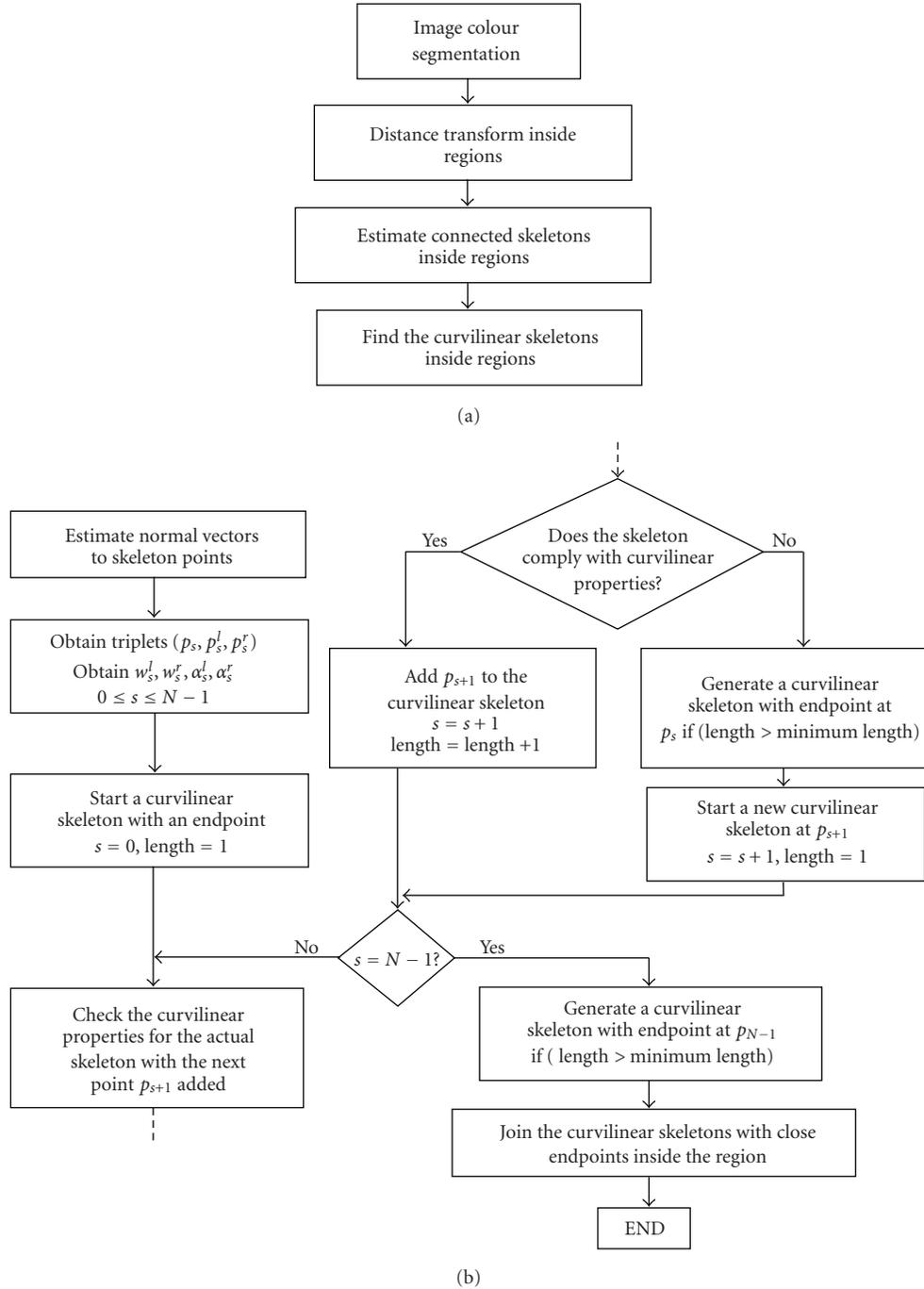


FIGURE 3: (a) An overview of the method. (b) Algorithm to detect the skeletons with curvilinear properties inside a region. Note that several curvilinear skeletons could be detected inside an homogeneous colour region.

comparison function,  $g(v_{x_1}, v_{x_2})$ , is defined. This function is true if the  $v_{x_1}$  and  $v_{x_2}$  values associated to the  $x_1$  and  $x_2$  nodes are similar according to some criteria and false otherwise. The decimation process consists of the following steps

- (1) *Regular decimation process.* The  $h_x$  value of a regular node  $x$  at level  $l+1$  is set to 1 if the four regular nodes immediately underneath  $\{y_i\}$  are similar according to

some criteria, and their  $h_{\{y_i\}}$  values are equal to 1. That is,  $h_x$  is set to 1 if

$$\left\{ \bigwedge_{\forall y_j, y_k \in \{y_i\}} g(v_{y_j}, v_{y_k}) \right\} \wedge \left\{ \bigwedge_{y_j \in \{y_i\}} h_{y_j} \right\}. \quad (1)$$

Besides, at this step, interlevel arcs among regular nodes at levels  $l$  and  $l+1$  are established. If  $x$  is

a homogeneous regular node at level  $l + 1$  ( $h_x = 1$ ), then the set of four nodes immediately underneath  $\{y_i\}$  are linked to  $x$ .

- (2) *Irregular decimation process.* Each irregular or regular node  $x \in N_l$  without parent at level  $l + 1$  chooses the closest neighbor  $y$  according to the  $v_x$  value. Besides, this node  $y$  must be similar to  $x$ . That is, the node  $y$  must satisfy

$$\left\{ \left\| v_x - v_y \right\| = \min(\|v_x - v_z\| : z \in \xi_x) \right\} \wedge \left\{ g(v_x, v_y) \right\}. \quad (2)$$

If this condition is not satisfied by any node, then a new node  $x'$  is generated at level  $l + 1$ . This node will be the parent node of  $x$ . Besides, it will constitute a root node, and its receptive field at base level will be a homogeneous set of pixels according to the specific criteria. On the other hand, if  $y$  exists and it has a parent  $z$  at level  $l + 1$ , then  $x$  is also linked to  $z$ . If  $y$  exists but it does not have a parent at level  $l + 1$ , a new irregular node  $z'$  is generated at level  $l + 1$ . In this case, the nodes  $x$  and  $y$  are linked to  $z'$ .

This process is sequentially performed and, when it finishes, each node of  $G_l$  is linked to its parent node in  $G_{l+1}$ . That is, a partition of  $N_l$  is defined. It must be noted that this process constitutes an implementation of the union-find strategy. The union-find uses tree structures to represent sets. A find operation looks for the parent of a node at level  $l$ . If two nodes at level  $l$  are similar, then a union operation will be performed by setting one of the two nodes to be the parent of both ones at level  $l + 1$ .

- (3) *Definition of intralevel arcs.* The set of edges  $E_{l+1}$  is obtained by defining the neighborhood relationships between the nodes  $N_{l+1}$ . Two nodes at level  $l + 1$  are neighbors if their reduction windows are connected at level  $l$ .

The structure hierarchy stops growing when it is no longer possible to link together any more nodes because they are not similar. The set of nodes which are not linked to any node at upper levels defines a partition of the input image (see [6] for further details).

This decimation process avoids the shift variance problem associated to the BIP. To demonstrate this issue, we compared the proposed modification with the original BIP and with the main irregular structures present in the literature in a colour-based segmentation framework. Obtained results were shown in [2]. They demonstrate that the decimation process of the uBIP avoids the shift variance problem at a cost of a very reduced increase of the computational costs.

**3.3. Medial Axis Extraction and Skeleton Classification.** The skeleton of the region is used for the analysis of the geometric properties that define if a region is curvilinear or not. The skeleton is defined as a subset of pixels that preserve the topological information of the region and it must approx-

imate the medial axis. In this work, a distance transform approach is used for each colour-segmented region, where a distance skeleton is a subset of grid points, such that every point represents the centre of a maximal disc contained in the given component. For estimating the distance transform of a region, we use the algorithm based on the  $d8$ -distance described in [35], which can approximate the distance transform inside the region in only two steps, so it has got a low computational cost. Those pixels which present a local maximum in the distance transform belong to the distance skeleton, and by choosing them we can obtain a skeleton for each region. These distance skeletons are generally not connected, so we postprocess them with morphological operations to obtain connected and smooth skeletons that are used to estimate further geometric properties.

We define the skeleton extracted from the distance transform of a region as the set of connected pixels  $p_s = (i_s, j_s)$ ,  $0 \leq s \leq N - 1$ , being  $N$  the number of pixels of the skeleton. The algorithm tries to join as many pixels as possible to form a curvilinear skeleton. The algorithm starts in an endpoint of the skeleton and it looks for adding the connected pixels checking if the set of pixels satisfies the *symmetry*, *elongation*, and *parallelism* properties. In case that any property is not fulfilled, the curvilinear skeleton is finished and a new curvilinear region will begin with the next positive evaluation. When all the pixels have been evaluated inside a region, the curvilinear skeletons with close endpoints are linked. Those parts of the objects with a skeleton evaluated as a curvilinear skeleton are considered curvilinear regions.

Next, the algorithms in order to check *symmetry*, *elongation*, and *parallelism* properties are presented.

- (i) *Symmetry.* The algorithm tries to evaluate how much asymmetrical the region is. To achieve this, first the normal vector is calculated for each pixel  $p_s$  in the skeleton, and the cross-points between the normal and the left and right borders of the region are estimated. If we define  $p_s^l$  and  $p_s^r$  as these cross-points, then we obtain the triplets  $(p_s, p_s^l, p_s^r)$ ,  $0 \leq s \leq N - 1$  (Figure 4(a)). Then  $w_s^l$  and  $w_s^r$  are calculated, being the Euclidean distance between pixels  $p_s$  and  $p_s^l$  and the Euclidean distance between pixels  $p_s$  and  $p_s^r$ , respectively, (Figure 4(a)). Then, the algorithm estimates

$$\Delta w_s = \left| w_s^l - w_s^r \right|, \quad 0 \leq s \leq N - 1, \quad (3)$$

so, for every point in the skeleton, the algorithm estimates the absolute difference between the widths on the left and right sides of the skeleton. For an object ideally symmetrical around the axis,  $w_s^l = w_s^r$ .

Next, if we consider  $\Delta w_s$  as a signal along the skeleton points, its energy can be estimated as follows:

$$E_{\Delta w}(N) = \frac{1}{N} \sum_{s=0}^{N-1} \left( \Delta w_s - \overline{\Delta w} \right)^2, \quad (4)$$

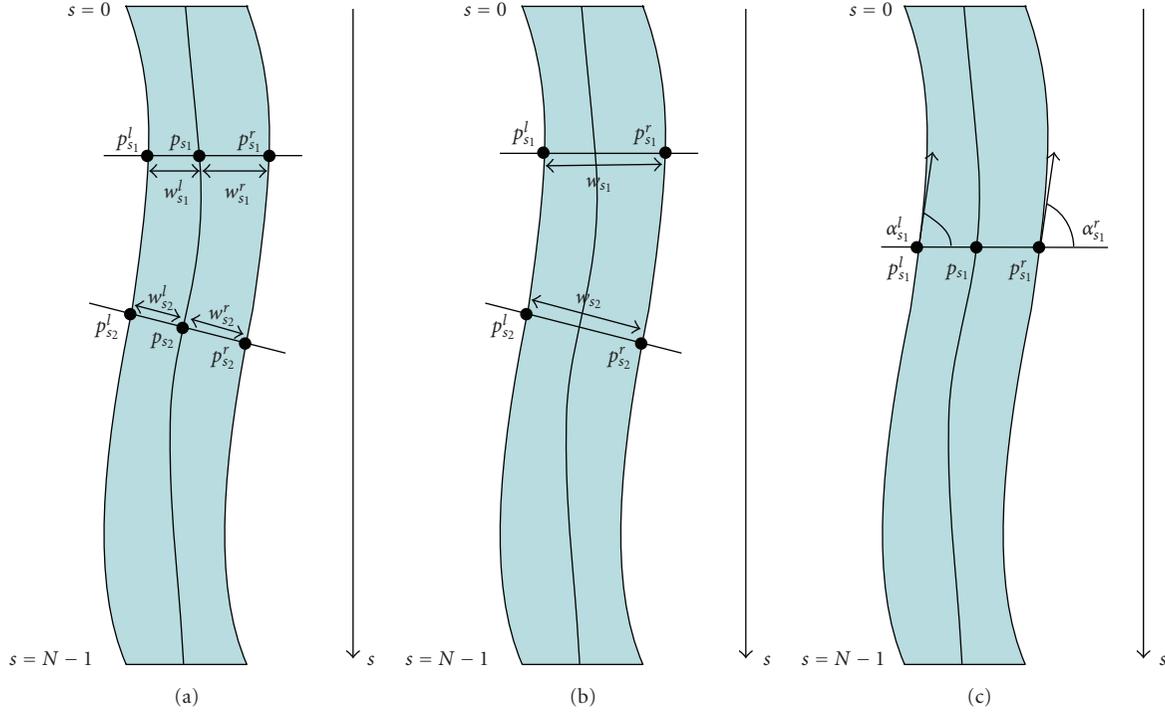


FIGURE 4: Definitions for the estimation of curvilinear properties. (a) Normal vector is estimated for every pixel  $p_s$  of the skeleton. Then, the cross points between the normal vectors and the left and right borders define the  $p_s^l$  and  $p_s^r$  pixels, respectively. The widths  $w_s^l$  and  $w_s^r$  are calculated as the distance between the cross points and the pixel  $p_s$ , where the normal vector is estimated. (b) For the elongation property, the method estimates  $w_s$  as the distance between  $p_s^l$  and  $p_s^r$ . (c) For the parallelism property, the tangential vectors at pixels  $p_s^l$  and  $p_s^r$  are estimated. Then  $\alpha_s^l$  and  $\alpha_s^r$  are defined as the angles between those vectors and the normal vector at  $p_s$ .

being

$$\overline{\Delta w}(N) = \frac{1}{N} \sum_{s=0}^{N-1} \Delta w_s. \quad (5)$$

Equation (4) grows with the asymmetries of the region, so the algorithm uses it as an estimation of how much asymmetrical the region is. Then the maximum value of energy allowed for a curvilinear region is defined as

$$U_E(N) = U \left( 1 - e^{-(N^2/2\sigma_U^2)} \right), \quad (6)$$

being  $U$  and  $\sigma_U$  parameters of the method.

By last, given a skeleton with  $N$  points, the symmetrical property is defined as

$$U_E(N) - E_{\Delta w}(N) \geq 0. \quad (7)$$

It must be noted that  $U_E(N)$  grows with the number of points of the evaluated skeleton, with the goal to detect as longer regions as possible (Figure 5).

- (ii) *Elongation*. For the elongation property, the algorithm tries to evaluate the relation between the length of the skeleton and the average width. Given a position  $s$  in the skeleton, the width  $w_s$  of the region

is estimated as the Euclidean distance between pixels  $p_s^l$  and  $p_s^r$  (Figure 4(b))

$$w_s = w_s^l + w_s^r, \quad 0 \leq s \leq N-1 \quad (8)$$

Then, we can estimate

$$\overline{w}(N) = \frac{1}{N} \sum_{s=0}^{N-1} w_s. \quad (9)$$

The elongation property is defined as

$$L_{\max} - U_w \cdot \overline{w}(N) \geq 0, \quad (10)$$

where  $L_{\max}$  is the maximum length that the region could have with all the connected pixels of the skeleton, and  $U_w$  is a parameter of the method, that must be greater than 1.0 in order to detect elongated regions.

- (iii) *Parallelism*. In a similar way to the previous properties, the algorithm tries to evaluate the degree of local parallelism between the left and right borders of the region. For that purpose, the algorithm in a first step estimates the tangential vectors on the borders at pixels  $p_s^l$  and  $p_s^r$  (Figure 4(c)). Then, the angles between those vectors and the normal vectors to the borders are estimated. Given a position  $s$ , if we define

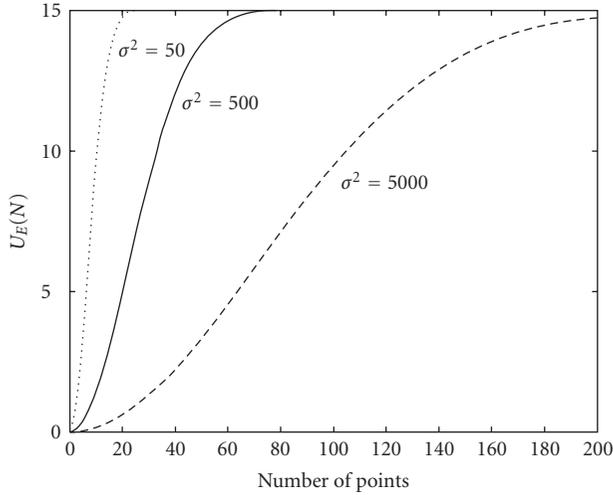


FIGURE 5: Maximum value of energy allowed for the curvilinear regions for different values of  $\sigma_U$  and  $U = 15$ .  $N$  represents the number of points of the skeleton. The method can benefit the detection of longer regions using this function instead of a fixed threshold for the maximum value allowed for  $E_{\Delta w}$ .

those angles as  $\alpha_s^l$  and  $\alpha_s^r$  for the left and right borders, respectively (Figure 4(c)), then the method estimates the difference between them

$$\Delta\alpha_s = \left| \alpha_s^l - \alpha_s^r \right|, \quad 0 \leq s \leq N - 1. \quad (11)$$

Next, we define

$$\overline{\Delta\alpha}(N) = \frac{1}{N} \sum_{s=0}^{N-1} \Delta\alpha_s, \quad (12)$$

being a term that increases with the differences in orientation between left and right borders.

Then, the parallelism property is defined as follows:

$$U_\alpha - \overline{\Delta\alpha}(N) \geq 0, \quad (13)$$

being  $U_\alpha$  a parameter of the method.

As we have mentioned above, the curvilinear regions are those sets of points with skeletons that comply with (7), (10), and (13). The method tries to form as longer regions as possible. If we define  $N_s$  as the number of points of the skeleton in a colour-segmented region, then the length of a curvilinear region is  $N_c = \max\{N_i\}$ , being  $0 < N_i \leq N_s$  and

$$\begin{aligned} U_E(N_i) - E_{\Delta w}(N_i) &\geq 0, \\ L_{\max} - U_w \cdot \overline{w}(N_i) &\geq 0, \\ U_\alpha - \overline{\Delta\alpha}(N_i) &\geq 0. \end{aligned} \quad (14)$$

In Figure 3(b), an overview of the algorithm to detect the curvilinear skeletons is depicted. The method demands that the regions must have a minimum length defined by the application user. For other purposes, this minimum length could be also defined as a percentage of the image size.

## 4. Experimental Results

Several tests have been performed to evaluate the stability of the detected regions. Different toolboxes and protocols have been also used to conduct these tests. We also provide in this Section the environment mapping framework, where the proposed approach is currently applied (Section 4.1) and an estimation of the parameters that the approach employs (Section 4.2).

Figures 6, 7, and 8 present some experiments and the obtained results. In Figure 6, a detection test with simple images is shown. The original images, segmentation stages, and the results of the curvilinear properties evaluation are represented. The pixels that belong to the curvilinear regions have been drawn in green in the last row of the figure. The experiment shows that the different cylinders in the original images are detected as curvilinear regions. In Figure 7, a detection test with a complex scene is depicted. The scene is captured from two different viewpoints, and the experiments show that several objects of interest are detected in both images. In Figure 8, a test of detecting the same object under different viewing conditions is presented. The object is seen from different points of view and is always evaluated as a curvilinear region in a positive way.

A standard dataset (<http://www.robots.ox.ac.uk/~vgg/research/affine>) has been also used to carry out other experiments. Images and Matlab code to carry out the tests have been downloaded. The dataset is composed by eight different scenes, where images have been captured under changing conditions (viewpoint changes, scaling, image blur, jpeg compression, or illumination changes). The datasets can be grouped into structured and textured scenes. Figure 9(a) shows example images from several scenarios. The structured group contains homogeneous regions which present distinctive boundaries, and the textured one contains repeated patterns or textured surfaces. Clearly, the proposed approach could be useful to deal with structured scenarios, where curvilinear regions can appear. Its behaviour in textured scenes has been tested obtaining poor results, as it was expected, with the exception of *Trees* scenario, where some stable regions have been detected. The aim of this test is to measure how many of the obtained regions are found in images under different transformations, relative to the lowest total number of regions detected (where only the part of the image that is visible in both images is taken into account). The ground truth is provided by mapping the regions detected on the images in a set to the reference image of this set using homographies. Then, the measure of repeatability is the relative amount of overlap between regions detected in the reference image and in the other image. In the described framework, the overlap is computed by assuming that output regions of the evaluated detector are represented by ellipses. So the detected regions have been enclosed into ellipses to employ the tests tool. These ellipses have the same first and second moments that the corresponding detected regions. In Figure 10, the results for three of the scenarios are depicted. The proposed detector is compared to other region detectors: the *maximally stable extremal region* detector (MSER) [18], the *intensity extrema-based region* detector (IBR) [17],

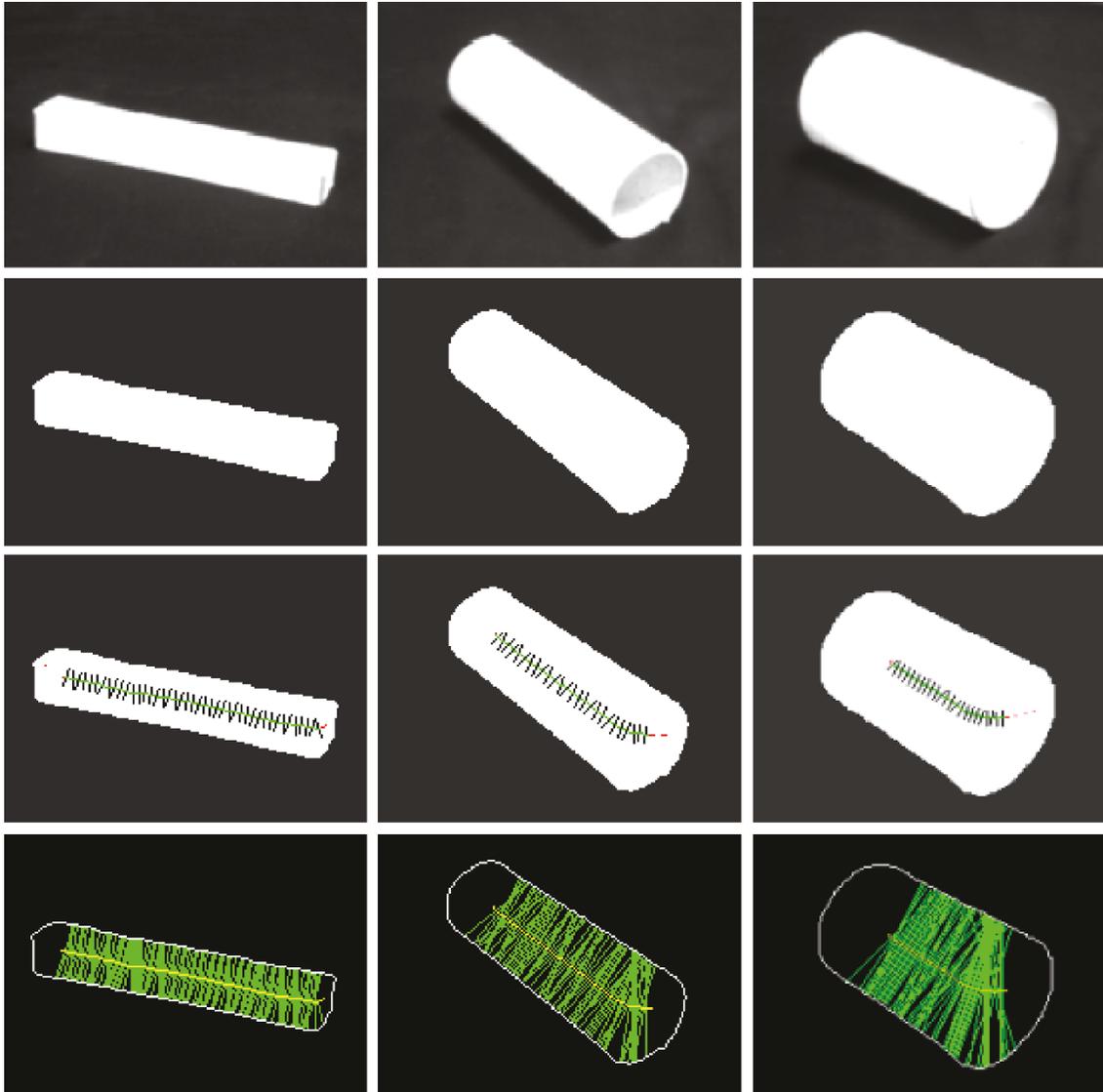


FIGURE 6: Detection test with simple images. The first row contains the original images. In the second row, the results of a segmentation stage are depicted. The third row contains the results of the curvilinear properties evaluation algorithm. The points of the skeletons inside the region of interest have been drawn; the green pixels represent the parts of the skeleton with a positive evaluation in the curvilinear properties, and the red ones those parts with a negative evaluation. Also, the estimated normal vectors to the skeleton points have been drawn in black. Finally, in the fourth row, the skeletons of the detected curvilinear regions are depicted in yellow, and in green the normal vectors that have been prolonged to see the crosspoints with the borders.

and the salient region detector in [2]. The results show that the proposed detector ranks similar to the rest of the approaches, and when the transformation is not severe, it may have a good performance. Also, the proposed approach generates a reduced number of regions compared with MSER and IBR detectors (Figure 9(b)). As it was commented in the introduction section, this is the main disadvantage of the approach. In any case, the repeatability score of detected regions makes it suitable for mobile robots global localization purposes.

The computational cost depends strongly on the size and complexity of the image. For example, in the *Graffiti* images, which are complex images of  $800 \times 640$  pixels, the overall

method needs about 1100 milliseconds in a 2.4 GHz Pentium PC. In average, the segmentation and skeleton stages take the 80% of this time, and the 20% corresponds with the algorithm of detecting the curvilinear properties. But for simpler images of  $320 \times 240$  pixels like the captured by the robot architecture, the overall method needs less than 100 milliseconds. If we consider that for a mobile robot application the visual localization can take place from time to time, the method could be suitable for extracting natural landmarks.

*4.1. Testing the Approach in an Environment Mapping Framework.* The visual landmark detector has been included into

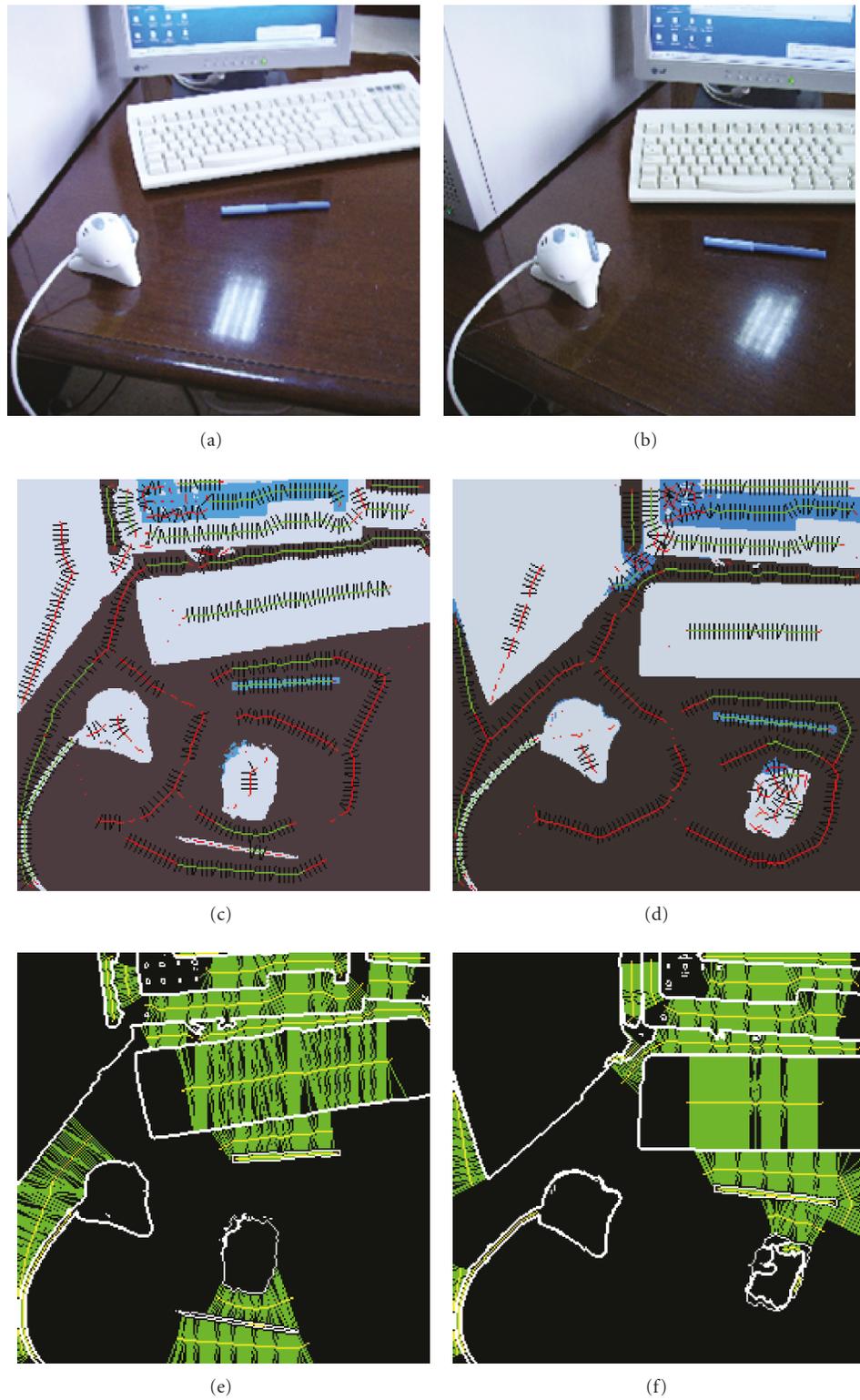


FIGURE 7: Detection test with a complex image seen from different viewpoints. Several interesting objects as the ball pen, keyboard, and webcam cable have been detected in both images. (a)-(b) Original images; (c)-(d) segmented images with the points of the skeletons and the normal vectors. Green points mean skeletons with curvilinear properties and red points skeletons without curvilinear properties. (e)-(f) Final results are depicted (region colours have not been represented for a clearer representation). The curvilinear skeletons are drawn with the normal vectors prolonged to see the crosspoints with the borders.

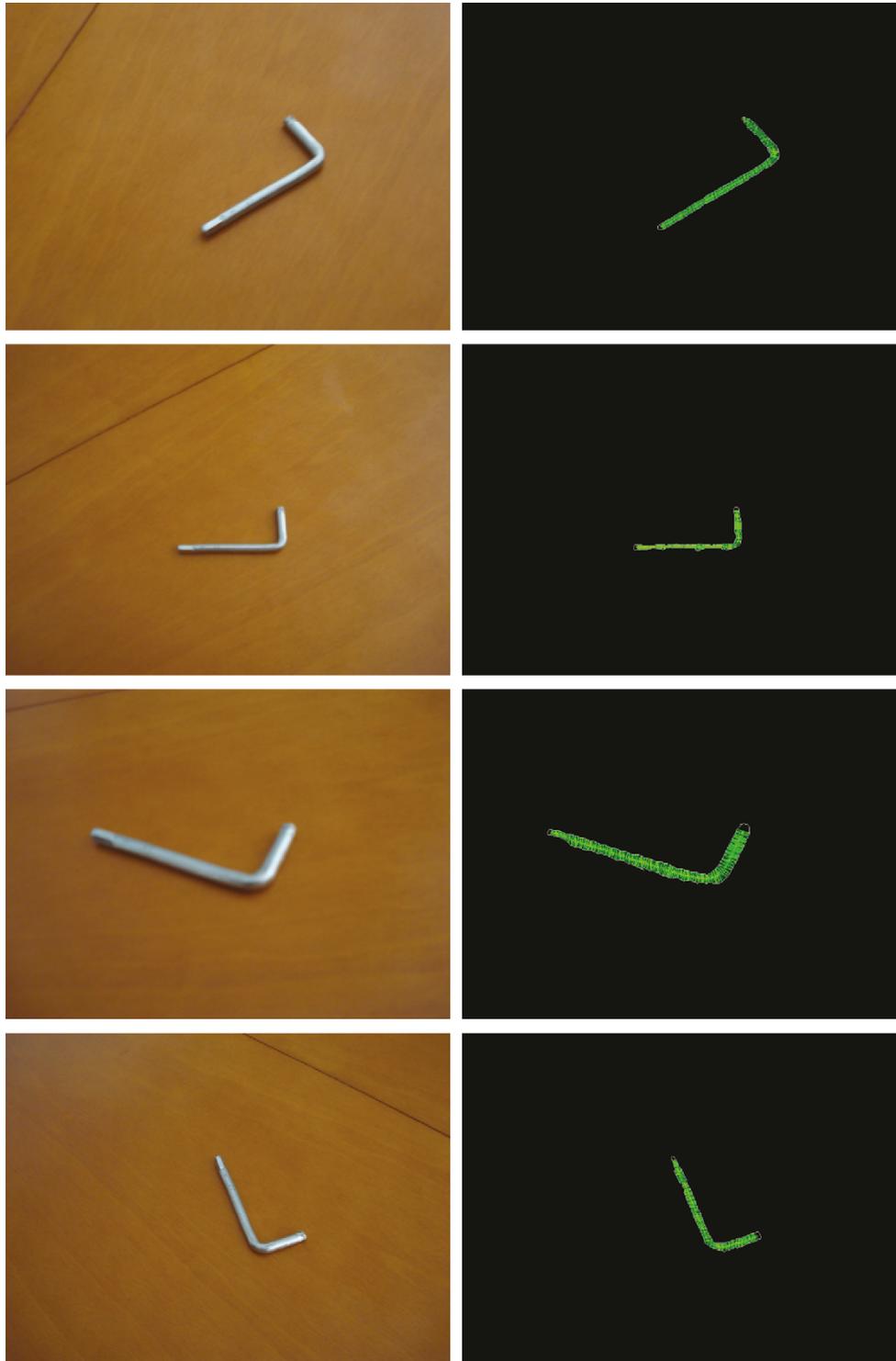
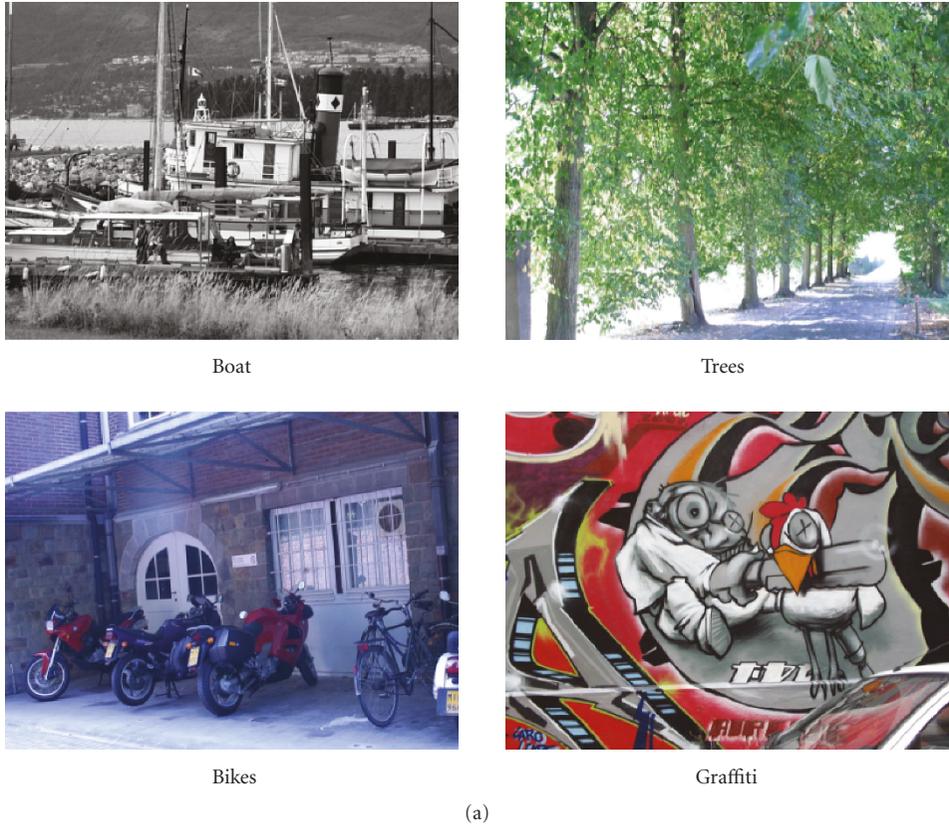


FIGURE 8: Detection test of the same object under different viewing conditions. In the first column, the original images are depicted. In the second column, the curvilinear regions are drawn. The object has been always detected as a curvilinear region.

the software architecture of an ActiveMedia Pioneer 2AT robot. The robot is equipped with an STH-MDCS stereoscopic camera from Videre Design, which is mounted at the front and top of the robot at a constant orientation, looking forward. This stereo camera consists of two 1.3 megapixels,

progressive scan CMOS imagers mounted in a rigid body. Images were restricted to  $320 \times 240$  pixels. The viewpoint invariance of the detector was checked by driving the robot through different environments while capturing real-life stereo images. Figure 11 shows several frames of a typical



(a)

Detector	Number of regions graffiti image
IBR [16]	679
MSER [17]	533
Salient region [2]	147
Proposed	252

(b)

FIGURE 9: (a) Image examples of the standard data set. (b) Comparison with other region detectors with the number of regions detected in the GRAFFITI image.

sequence, where the results of the detected landmarks have been superimposed on the original images.

**4.2. Estimation of Parameters.** The proposed method requires choosing values for a set of parameters. These parameters are the color threshold employed at the segmentation stage and the set of parameters employed by the curvilinear regions detection stage.

In order to choose a color threshold value  $T_{seg}$  which can remain unaltered for the experiments shown in this paper, we have tested several values into a color-based segmentation framework. To evaluate the obtained results, two empirical methods have been employed: the shift-variance (SV) proposed by [36] and the Q function [33]. Shift variance means that the image segmentation provided by a pyramid-based

approach varies when the base of the pyramid is shifted slightly. This is an undesirable effect, so the shift variance (SV) can be taken as a measurement of an algorithm quality. On the other hand, the Q function takes into account that segmented regions must be uniform and homogeneous; that the interior of these regions must be simple, without too many small holes; and that adjacent regions must present significantly different values for uniform characteristics. From the conducted tests, the best choice for the color threshold was  $T_{seg} = 20.0$ .

On the contrary, the set of parameters used at the curvilinear regions detection stage is user specific, that is, they depend on the current application. In the tests of repeatability, the parameters have been adjusted to obtain a stable set of regions.  $(U, \sigma_U, U_w, U_\alpha)$  has been set to  $(5, \sqrt{500}, 2.0, 20$  degrees) and the minimum length to 15 pixels.

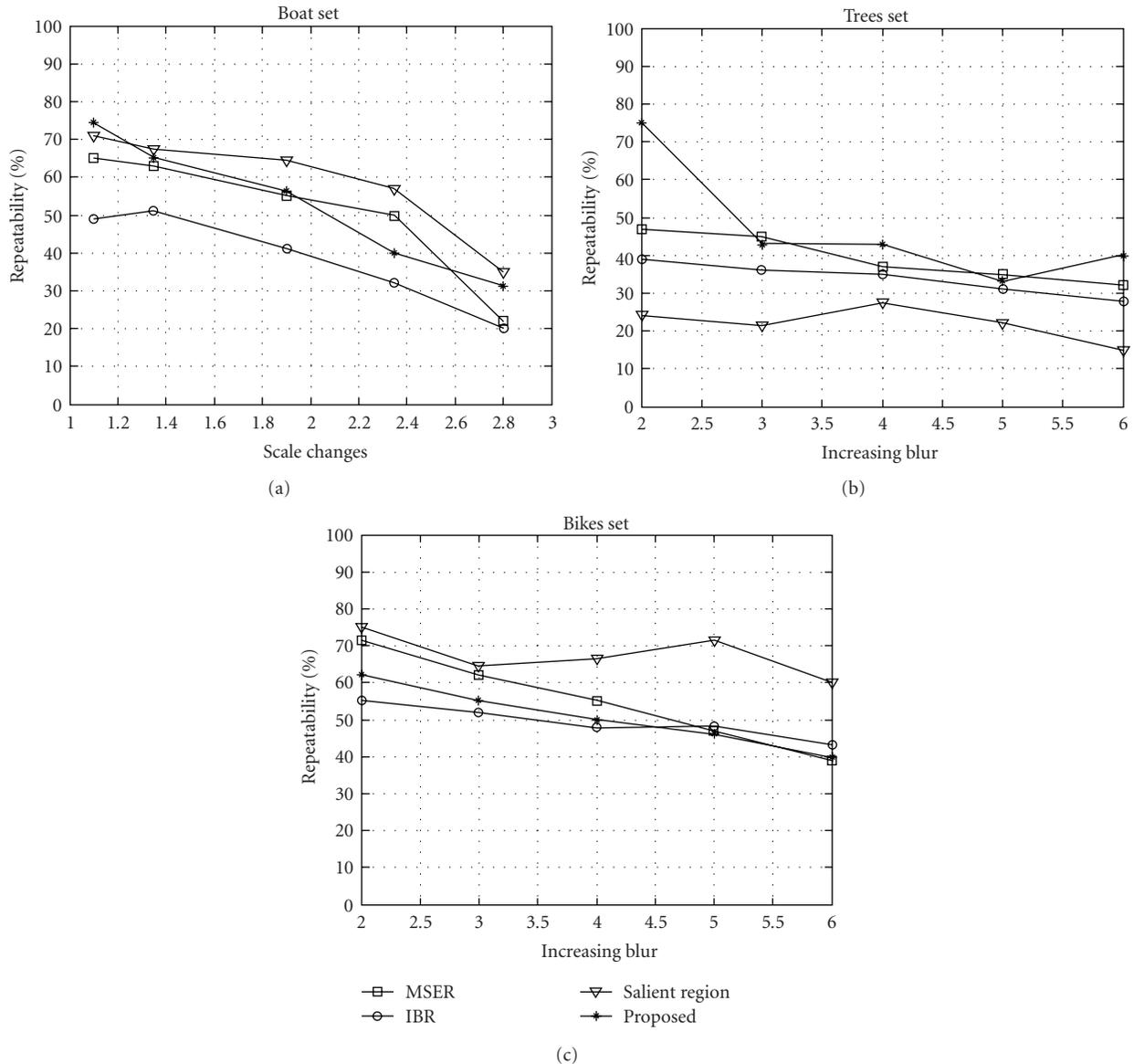


FIGURE 10: Repeatability scores for BOAT, TREES, and BIKES sequences (see Figure 9 (a)).

Other values could be better to detect curvilinear features if the approach is applied in a different framework (roadmap detection, human silhouette delineation ...).

## 5. Conclusions and Future Work

This paper describes a novel visual landmark detection system. The detection is based on the presence of the scene of objects or parts of objects whose shapes can be modelled by cylinders or generalized cylinders, inspired on Marr's visual theory. Thus, the algorithm looks for curvilinear regions, that is, the planar representation of these objects. This biological inspiration allows to detect regions with a higher semantic significance than other approaches. The search of the regions is addressed from a previous segmentation of the acquired image. This process differs from the typical strategy

employed by other approaches, which initially accomplish a borders detection stage. In our tests in real image sequences, the segmentation stage has been able to provide stable partitions at a reduced computational cost (see [33] or [2] for further details). The image regions which satisfy several previously established geometrical properties constitute the set of curvilinear regions. The performance of the detection scheme has been tested and compared to other region detectors.

Our approach can detect stable regions with a high semantical significance. The approach depends on the presence of these specific structures in the scene and, in a mobile robot navigation framework, it could be combined with other visual feature detectors to increase the number of perceived features. Future work will be focused on this integration procedure and the mapping of structured

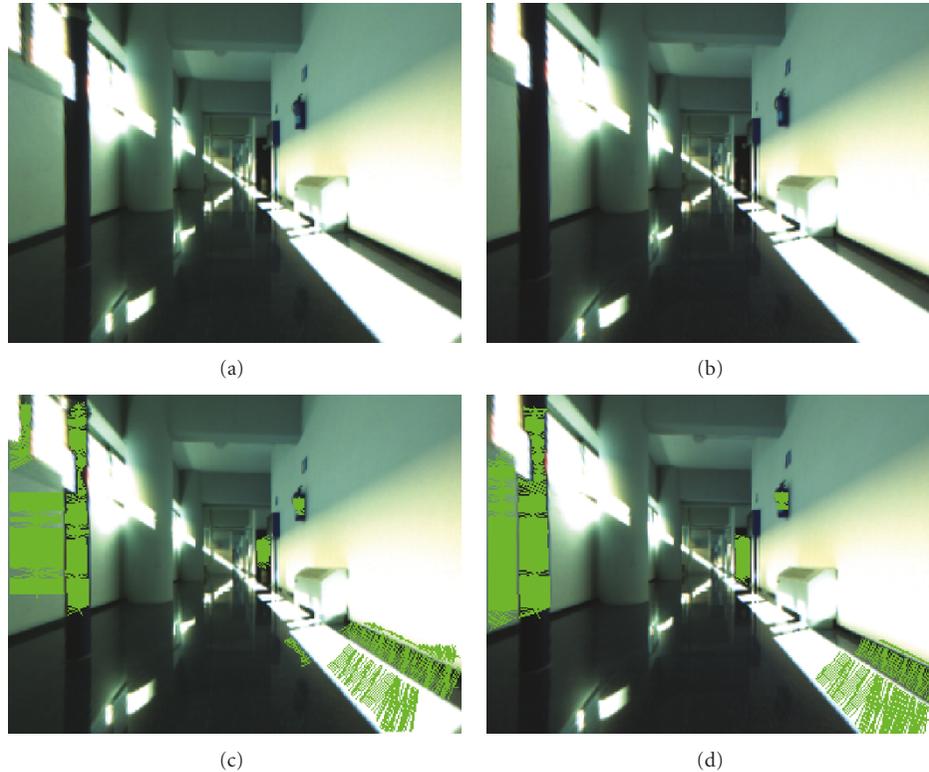


FIGURE 11: Consecutively acquired images by the Pioneer robot, with size  $320 \times 240$  pixels. (a)-(b) Original images. (c)-(d) Original images with the results of the detection superimposed. Several landmarks are detected in both of the images.

environments. Also, a region descriptor will be needed to achieve a scene matching. Although we have done some previous work in [37], we must develop a better descriptor in our future work for the curvilinear regions, in order to achieve a reliable matching for localization purposes.

## Acknowledgments

This work has been partially granted by the Spanish Government and FEDER funds Project no. TIN2008-06196 and by the Junta de Andalucía Project no. P07-TIC-03106.

## References

- [1] P. Núñez, R. Vázquez-Martín, J. C. del Toro, A. Bandera, and F. Sandoval, "Natural landmark extraction for mobile robot navigation based on an adaptive curvature estimation," *Robotics and Autonomous Systems*, vol. 56, no. 3, pp. 247–264, 2008.
- [2] R. Vázquez-Martín, R. Marfil, P. Núñez, A. Bandera, and F. Sandoval, "A novel approach for salient image regions detection and description," *Pattern Recognition Letters*, vol. 30, no. 16, pp. 1464–1476, 2009.
- [3] D. Marr and H. K. Nishihara, "Representation and recognition of thespatial organization of three-dimensional shapes," *Proceedings of the Royal Society of London B*, vol. 200, pp. 269–294, 1978.
- [4] J. H. Jang and K. S. Hong, "Detection of curvilinear structures and reconstruction of their regions in gray-scale images," *Pattern Recognition*, vol. 35, no. 4, pp. 807–824, 2002.
- [5] H. Deng, W. Zhang, E. Mortensen, T. Dietterich, and L. Shapiro, "Principal curvature-based region detector for object recognition," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '07)*, June 2007.
- [6] R. Marfil, L. Molina-Tanco, A. Bandera, and F. Sandoval, "The construction of bounded irregular pyramids with a union-find decimation process," in *Proceedings of the 6th IAPR-TC-15 International Workshop on Graph-Based Representations in Pattern Recognition (GbrPR '07)*, vol. 4538 of *Lecture Notes in Computer Science*, pp. 307–318, 2007.
- [7] G. Steger, "An unbiased detector of curvilinear structures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 20, no. 2, pp. 113–125, 1998.
- [8] D. C. Asmar, J. S. Zelek, and S. M. Abdallah, "Tree trunks as landmarks for outdoor vision SLAM," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2006, pp. 196–203, 2006.
- [9] S. Ahn, M. Choi, J. Choi, and W. K. Chung, "Data association using visual object recognition for EKF-SLAM in home environment," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS '06)*, pp. 2588–2594, October 2006.
- [10] H. Tamimi, H. Andreasson, A. Treptow, T. Duckett, and A. Zell, "Localization of mobile robots with omnidirectional vision using Particle Filter and iterative SIFT," *Robotics and Autonomous Systems*, vol. 54, no. 9, pp. 758–765, 2006.

- [11] J. Folkesson, P. Jensfelt, and H. Christensen, "Graphical SLAM using vision and the measurement subspace," in *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 325–330, 2005.
- [12] J. B. Hayet, F. Lerasle, and M. Devy, "Visual landmarks detection and recognition for mobile robot navigation," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 313–318, June 2003.
- [13] R. Vázquez-Martín, J. C. Del Toro, A. Bandera, and F. Sandoval, "Data- and model-driven attention mechanism for autonomous visual landmark acquisition," in *Proceedings of the IEEE International Conference on Robotics and Automation*, pp. 3372–3377, April 2005.
- [14] I. Horswill, "Polly: a vision-based artificial agent," in *Proceedings of the 11th National Conference on Artificial Intelligence*, pp. 824–829, July 1993.
- [15] M. J. Tarr and H. H. Bulthof, "Image-based object recognition in man, monkey and machine," *Cognition*, vol. 67, no. 1-2, pp. 1–20, 1998.
- [16] K. Mikolajczyk, T. Tuytelaars, C. Schmid et al., "A comparison of affine region detectors," *International Journal of Computer Vision*, vol. 65, no. 1-2, pp. 43–72, 2005.
- [17] T. Tuytelaars and L. Van Gool, "Matching widely separated views based on affine invariant regions," *International Journal of Computer Vision*, vol. 59, no. 1, pp. 61–85, 2004.
- [18] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," in *Proceedings of the British Machine Vision*, pp. 384–393, 2002.
- [19] I. Biederman, "Recognition-by-Components: a theory of human image understanding," *Psychological Review*, vol. 94, no. 2, pp. 115–147, 1987.
- [20] T. O. Binford, "Visual perception by computer," in *Proceedings of the IEEE Conference on Systems and Controls*, 1971.
- [21] S. A. Shafer, T. Kanade, and J. Kender, "Gradient space under orthography and perspective," *Computer Vision, Graphics and Image Processing*, vol. 24, no. 2, pp. 182–199, 1983.
- [22] F. Ulupinar and R. Nevatia, "Shape from contour: straight homogeneous generalized cylinders and constant cross section generalized cylinders," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 120–135, 1995.
- [23] M. Zerroug and R. Nevatia, "Part-based 3D descriptions of complex objects from a single image," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, pp. 835–848, 1999.
- [24] K. Rao and G. Medioni, "Generalized cones: useful geometric properties," *Computer Vision, Graphics and Image Processing*, vol. 10, no. 3, pp. 185–208, 1992.
- [25] H. Rom and G. Medioni, "Hierarchical decomposition and axial shape description," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 10, pp. 973–981, 1993.
- [26] P. Saint-Marc, H. Rom, and G. Medioni, "B-spline contour representation and symmetry detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1191–1197, 1993.
- [27] K. Rao and R. Nevatia, "Describing and segmenting scenes from imperfect and incomplete data," *CVGIP: Image Understanding*, vol. 57, no. 1, pp. 1–23, 1993.
- [28] L. Liu, D. Zhang, and J. You, "Detecting wide lines using isotropic nonlinear filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 6, pp. 1584–1595, 2007.
- [29] S. X. Li, H. X. Chang, and C. F. Zhu, "Fast curvilinear structure extraction and delineation using density estimation," *Computer Vision and Image Understanding*, vol. 113, no. 6, pp. 763–775, 2009.
- [30] H. Sakata, K. I. Tsutsui, and M. Taira, "Toward an understanding of the neural processing for 3D shape perception," *Neuropsychologia*, vol. 43, no. 2, pp. 151–161, 2005.
- [31] E. Chinellato and A. P. Del Pobil, "The neuroscience of vision-based grasping: a functional review for computational modeling and bio-inspired robotics," *Journal of Integrative Neuroscience*, vol. 8, no. 2, pp. 223–254, 2009.
- [32] E. Chinellato and A. P. del Pobil, "Neural coding in the dorsal visual stream," in *Proceedings of the 10th International Conference on Simulation of Adaptive Behavior. From Animals to Animats (SAB '08)*, vol. 2040 of *Lecture Notes in Computer Science*, pp. 230–239, 2008.
- [33] R. Marfil, L. Molina-Tanco, A. Bandera, J. A. Rodríguez, and F. Sandoval, "Pyramid segmentation algorithms revisited," *Pattern Recognition*, vol. 39, no. 8, pp. 1430–1451, 2006.
- [34] R. Marfil, J. A. Rodríguez, A. Bandera, and F. Sandoval, "Bounded irregular pyramid: a new structure for colour image segmentation," *Pattern Recognition*, vol. 37, no. 3, pp. 623–626, 2004.
- [35] G. Klette, "A comparative discussion of distance transformation and simple deformations in digital image processing," *Machine Graphics and Vision*, vol. 12, no. 2, pp. 235–256, 2003.
- [36] D. Prewer and L. Kitchen, "Soft image segmentation by weighted linked pyramid," *Pattern Recognition Letters*, vol. 22, no. 2, pp. 123–132, 2001.
- [37] J. M. Pérez-Lorenzo, S. G. Galán, A. Bandera, R. Vázquez-Martín, and R. Marfil, "Classifying a new descriptor based on Marr's visual theory," in *Proceedings of the 3rd International Work-Conference on the Interplay between Natural and Artificial Computation (IWINAC '09)*, vol. 5601 of *Lecture Notes in Computer Science*, pp. 205–212, 2009.