Qualitative Models of Shape, Size, Orientation and Distance Applied to the Description of Images Containing 2D Objects

Zoe Falomir and M. Teresa Escrig

Universitat Jaume I, Engineering and Computer Science Department
E-12071 Castellón, Spain
{zfalomir, escrigm}@icc.uji.es

Abstract

In this paper, we present an approach to qualitative description of images containing non-overlapping 2D objects. This approach consists of extracting the properties of the main points of the objects in an image and applying qualitative models to them. Qualitative models of shape and size describe visual features of the objects in the image, while qualitative models of orientation and distance describe spatial features of those objects. Finally, a string which describes the image in qualitative terms is obtained. This string can be used to compute the relationships between the objects in the image and also to compare images by obtaining a degree of similarity between them.

Introduction

Service robots need a system of visual perception similar to that of human beings in order to interact with people and to navigate efficiently through real environments dealing with problems such as unpredictable obstacles or changing targets. Qualitative techniques for representing information and reasoning can help robots to interpret their environment: to describe a new place and to recognize a known one only by using the main landmarks/features of that environment.

Human beings use language to describe images. How we do it is studied by psycho-linguistic researchers (Landau & Jackendoff 1993). In general, nouns are used to refer to objects, adjectives to express properties of these objects and prepositions to express relations between them. These nouns, adjectives and prepositions are qualitative labels which extract knowledge from images and which can be used to communicate and compare image content.

Because of the numerical properties of digital images, most of the image treatment in computer vision has been carried out by applying mathematical models and other quantitative techniques to detect objects in an image. Our aim is to use some of these techniques in order to obtain the main points of the objects contained inside an image (such as edges and vertices) and then to apply qualitative models to them, so that we could obtain the main visual and spatial information of these objects.

The approach presented in this paper applies qualitative models of shape (Museros & Escrig 2004), size, distance (Escrig & Toledo 2001) and orientation (Hernández 1991; Freksa 1992) to image description. These models represent visual (shape and size) and spatial (orientation and distance) information about 2D objects in an image. The visual information obtained from an image describes *which* objects are placed in an environment, and the spatial information obtained describes *where* all the objects are located with respect to (wrt) each other and wrt the observer. Contextualizing our approach to robot navigation, a mobile robot with a camera could qualitatively describe its environment by describing the images taken by the camera and, by interconnecting all the images, it could localize itself wrt all the objects contained in any of the images, that is, in any place of the environment.

In the literature, related works which describe images qualitatively have appeared. In (Museros & Escrig 2004), an approach for the qualitative description of the shape of a 2D object contained in an image is applied to assemble tile mosaics. In (Qayyum & Cohn 2007; Bañuelos 2000), qualitative description of images are used for image retrieval in data bases. In (Qayyum & Cohn 2007), landscape images are divided by a grid for its description so that semantic categories (grass, water, etc.) are identified and qualitative relations of compared size, time and topology relations are used to describe the image. In (Bañuelos 2000), images composed by squares, triangles and circles are described qualitatively by using the approach in (Chang *et al.* 1989) which uses projections to find the spatial relations between the objects in an image. In (Socher 1997), a verbal description of an image is provided to a robotic manipulator system, so that it can identify and pick an object. These objects are described qualitatively by predefined categories of type, colour, size, shape and spatial relations. In (Lovett *et al.* 2006), a qualitative description for sketch image recognition is proposed, which defines lines, arcs and ellipses as basic elements and also considers relative position, relative length and relative orientation of pairs of edges.

We believe that all the works described above provide evidence for the effectiveness of using qualitative information to describe and compare images. However, to our knowledge, none of them is intended to be applied to the description of the robot environment and to the enhancement of robot navigation. This is our main aim. Although we have tackled the problem by considering images containing 2D objects, in the near future, we will describe real images taken from the robot environment where doors, corners and other 3D objects will appear.

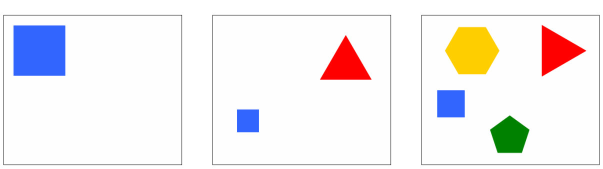
Outlining our Approach for Qualitative Image Description

In order to describe qualitatively images composed of 2D objects, our approach applies qualitative models of shape, size, orientation and distance.

If the considered image contains just one object, only its shape, its size with respect to (wrt) the image, and its orientation wrt the centre of the image can be described. For example, a possible qualitative description of the image in Figure 1a could be "an image containing a medium-sized blue square situated left-front wrt the centre of the image".

If the considered image contains two objects, the relative size of each object wrt each other and the orientation of an object wrt the other objects can be described. For example, a possible qualitative description of the image in Figure 1b could be "an image containing a blue small square and a red medium-sized triangle (…) the triangle is situated right-front wrt the centre of the image and right-front wrt the square".

Finally, if the considered image contains more than two objects, relations of compared distance and relative orientation between objects can be described. For example, a possible qualitative description of the image in Figure 1c could be "an image containing a triangle, an hexagon, a square and a pentagon (…) the square is closer to the pentagon than to the triangle (…) the pentagon is located left-middle wrt the reference system defined from the centroid of the triangle to the centroid of the hexagon, while the square is located left-front wrt the same reference system (…)".



(a) (b) (c)

Figure 1. Images composed by 2D objects.

Table 1 shows the qualitative models that our approach uses to describe images which contain an object, two objects and more than two objects.

Table 1. Qualitative models our approach uses to describe images composed by 2D objects.

Qualitative Model of	Number of Objects 1	2	> 2
Shape	x	x	x
Fixed Orientation	x	x	x
Compared Size (Object wrt Image)	x	x	x
Compared Size (Object wrt Object)	-	x	x
Relative Orientation	-	-	x
Compared Distance	-	-	x

The following sections describe the qualitative models of shape, size, orientation and distance used by our approach. Then, the structure of the string for the qualitative description of any image is presented. Finally, our conclusions and future work are explained.

Qualitative Model for Shape Description

Our approach uses Museros and Escrig's qualitative model for shape description (Museros and Escrig 2004) to obtain a description of any 2D object in an image. This model has been extended to identify regularity and convexity of the objects. Next section summarizes Museros and Escrig's model and, in the following section, our extension to this model is explained.

Museros and Escrig's Approach

Museros and Escrig's approach extracts the boundary of each object in an image by applying Canny's edge detector (Canny 1986). Then the slope between the pixels that compose that boundary is compared in order to obtain the main points of each object: vertices and points of curvature. These points are described qualitatively by this approach, which also obtains the colour and the centroid of each object.

The **vertices** of each object are described by a set of three elements $<A_j, L_j, C_j>$ where:

$$A_j \in \{right, acute, obtuse\},$$
$$L_j \in \{smaller, equal, bigger\} \text{ and}$$
$$C_j \in \{convex, concave\}$$

- A_j or the qualitative amplitude of the angle j is calculated by obtaining the circumference that includes the previous and following vertices of vertex j (j-1 and j+1) (Figure 2). If vertex j is included in that circumference, then the angle is *right*. If vertex j is external to the circumference, then the angle is *acute*. Finally, if vertex j is included in the circle that this circumference defines, then the angle is *obtuse*.
- L_j or the relative length of the two edges related to vertex j is obtained by comparing the Euclidean distance between two segments: the segment defined from vertex j-1 to vertex j and the segment defined from vertex j to vertex j+1. If the first distance obtained is smaller/equal/bigger than the second one, the relative length between the two edges in vertex j is *smaller/equal/bigger*, respectively.

- C_j or the <u>convexity of the angle</u> defined by the edges related to vertex j is calculated by obtaining the segment from the previous vertex (j-1) to the following vertex (j+1). If vertex j is on the left of that segment, then the angle is *convex*. If vertex j is on the right of that oriented segment, then the angle is *concave*.

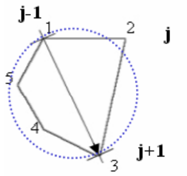


Figure 2. Characterization of a vertex in Museros and Escrig's approach.

The **points of curvature** of each object are characterized by a set of three elements <*curve, TC_j, C_j*> where:

$TC_j \in$ *{acute, semicircular, plane}* and
$C_j \in$ *{convex, concave}*

- TC_j or the <u>type of curvature</u> in point j is obtained by comparing the size of the segment defined by the previous point of curvature j-1 and the centre of the curve (*da* in Figure 3) with the size of the segment defined by the point of curvature j and the centre of the curve (*db* in Figure 3). If *da* is smaller than *db*, the type of curvature in j is *acute*; if *da* is equal to *db*, the type of curvature in j is *semicircular*; and finally, if *da* is bigger than *db*, the type of curvature in j is *plane*.

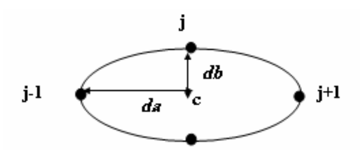


Figure 3. Characterization of a curve in Museros and Escrig's approach.

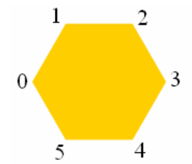
- C_j or the <u>convexity of the point of curvature</u> j is calculated by obtaining the segment from the previous vertex (j-1) to the following vertex (j+1). If vertex j is on the left of that segment, then the angle is *convex*. If vertex j is on the right of that segment, then the angle is *concave*.

Thus, the complete description of a 2D object is defined as a set of qualitative tags as:

[Type, Colour, [A_1,C_1,L_1]|[*curve*,TC_1,C_1], …,
[A_n,C_n,L_n]|[*curve*,TC_n,C_n]]

where *n* is the number of vertices and points of curvature of the object, *Type* belongs to the set {*without-curves, with-curves*}, *Colour* describes the RGB colour of the object by a triple [*R,G,B*] which stands for the Red, Green and Blue coordinates and $A_1,\dots,A_n$, $C_1,\dots,C_n$, $L_1,\dots,L_n$ and $TC_1,\dots, TC_n$, which have been previously explained.

Finally, as an example, Figure 4 shows the qualitative description of the hexagon in Figure 1c provided by Museros and Escrig's approach. The first vertex detected is that with the smaller coordinate x (considering that the origin of coordinates in computer vision is the upper-left corner of the image), while the first vertex described is that defined by the three first vertices obtained.



QualitativeShapeDescription(S)=
[*without-curves*, [255, 206, 0],
[
1: [*obtuse*, *equal*, *convex*],
2: [*obtuse*, *equal*, *convex*],
3: [*obtuse*, *equal*, *convex*],
4: [*obtuse*, *equal*, *convex*],
5: [*obtuse*, *equal*, *convex*],
0: [*obtuse*, *equal*, *convex*],
],
].

Figure 4. Qualitative description of the hexagon in Figure 1c.

Characterizing the Shape of 2D Objects

In Museros and Escrig's approach, a qualitative tag is included in order to distinguish if the object has curves or not (*with-curves*, *without-curves*), so that the comparison process can be accelerated. However, a more accurate characterization of the objects (according to geometry principles) can be defined by using the qualitative features described for each vertex.

The characterization defined for our approach consists on: (1) giving a name to the object that could represent it geometrically, (2) describing the regularity of its edges and (3) defining the convexity of the whole object.

Therefore, **objects without curves** can be characterized by a set of three elements:

[Name, Regularity, Convexity]

where,

Name $\in$ {*triangle, quadrilateral, pentagon, hexagon, heptagon, octagon, …, polygon*},
Regularity $\in$ {*regular, irregular*} and
Convexity $\in$ {*convex, concave*}

- **Name** is the name given to the object depending on its number of edges (or vertices qualitatively described) and it can take values from *triangle* to *polygon*;
- **Regularity** indicates if the object have *equal* angles and *equal* edges (so it is *regular*), or not (so it is *irregular*);
- **Convexity** indicates if the object has a *concave* angle (so it is *concave*) or not (so it is *convex*).

However, for triangular and quadrilateral objects a more accurate characterization can be made.

Triangular objects can be characterized as *right*, *obtuse* or *acute triangles* according to the kind of angles they have, and as *equilateral*, *isosceles* or *scalene triangles* according to the relation of length between its edges. Therefore, the element **Name** for a triangle is made up by three elements:

triangle–Kind_of_angles–Sides_relation

where,

Kind_of_angles ∈ {*right, obtuse, acute*}
Edges_Relation ∈ {*equilateral, isosceles, scalene*}

- **Kind_of_angles** indicates if the triangle has got a right angle (so it is *right*), an obtuse angle (so it is *obtuse*), or if all its angles are acute (so it is *acute*); and
- **Edges_relation** shows, if the edges of the triangle are all equal (so it is *equilateral*), or two equal (so it is *isosceles*), or none equal (so it is *scalene*).

Quadrilateral objects can be also characterized more accurately as *square, rectangle* or *rhombus* depending on the compared length between its edges and on its kind of angles. Therefore, the element **Name** for a quadrilateral is made up by two elements:

quadrilateral–Type_quadrilateral

where,

Type_quadrilateral ∈ {*square, rectangle, rhombus*}

- **Type_of_quadrilateral** specifies if the quadrilateral is a *square* (if all their angles are right and their edges equal), a *rectangle* (if all their angles are right and their opposite edges are equal), or a *rhombus* (if all their edges are equal and their opposite angles are equal).

On the other side, **objects with curves** can be also characterized by a set of three elements:

[Name, Regularity, Convexity]

where,

Name ∈ {*circle*, *ellipse*, *polycurve*, *mix-shape*}
Regularity ∈ {*regular, irregular*}
Convexity ∈ {*convex, concave*}

- **Name** is the name given to the object depending on its properties: *mix-shape* (if the shape of the object is made up by curves and straight edges), *polycurve* (if the shape of the object is made up only by curves), *circle* (if the shape of the object is a polycurve with only four relevant points, two of them defined as semicircular points of curvature) and *ellipse* (if the shape of the object is a polycurve with only four relevant points, two of them defined as points of curvature with the same type of curvature but different from *semicircular*, that is, both *plane* or *acute*).
- **Regularity** regarding to curves is not defined by our approach from the point of view of geometry. We consider 2D objects with circular or elliptical shapes as *regular* and the rest of objects with curvaceous shapes as *irregular*.
- **Convexity** of objects with curvaceous shapes is defined in the same way as for objects containing only straight edges: if an object has a *concave* vertex or point of curvature, that object is defined as *concave*; otherwise it is defined as *convex*.

Finally, according to this characterization, the triangle in Figure 1b will be characterized as *[triangle-acute-equilateral, regular, convex]*.

Qualitative Model of Compared Size

In order to describe images which contain objects with the same features but with different size, a qualitative model of Compared Size (CS) has been developed.

Represented Information. In order to represent the size of the objects in the image a Compared Size Reference System (CSRS) with two levels of granularity has been defined. The reference system with coarse level of granularity represents the size of an object A wrt the size of the image (A wrt Image), while the reference system with fine level of granularity represents the size of an object A wrt the size of another object B (A wrt B). Two reference systems are used in order to distinguish situations where the object compared can be larger than the other object (object wrt object comparison) or not (object wrt image comparison).

The CSRS has three components, CSRS = {US, LAB_{csrs}, INT_{csrs}}, where US refers to the relation obtained after comparing the area of an object wrt the area of another object or wrt the area of the image; LAB_{csrs} refers to the set of qualitative labels which represent compared size; and INT_{csrs} refers to the intervals associated to each compared size label of LAB_{csrs}, which will describe the size of the object in terms of US and which depends on the application.

The CSRS with coarse level of granularity is used to obtain the size of each object wrt the size of the image and it is defined as:

$CSRS_{LAB1}$ = {*small (s)*, *medium (m)*, *large (l)*}

$CSRS_{INT1}$ = {]0, 1/8 us[, [1/8 us, 1/4 us[, [1/4 us, 1us[}.

The CSRS with fine level of granularity is used to obtain the size of each object (A) wrt the size of another object (B) and it is defined as:

$CSRS_{LAB2}$ = {*smaller_than_half (sh)*, *half (h)*, *larger_than_half (lh)*, *equal (e)*, *smaller_than_double (sd)*, *double (d)*, *larger_than_double (ld)*}

$CSRS_{INT2}$ = {]0, 1/2 us[, [1/2 us, 1/2 us],]1/2 us, 1 us[, [1 us, 1 us],]1 us, 2 us[, [2 us, 2 us],]2 us, ∝[}.

Application to qualitative description of images. First, our approach for qualitative description of images calculates: (1) the area of the image (n pixels of height x m pixels of width) and (2) the area of each object by using the cross-product of its vertices, as described in (Goldman 1991). After that, our approach compares the area of each object wrt the area of the image and obtains a qualitative tag of compared size from the reference system of compared size at a coarse level of granularity. Finally, our approach compares the area of each object wrt the area of the other objects in the image and obtains a qualitative tag of compared size from the reference system of compared size at a fine level of granularity.

As an example, the square in Figure 5b will be characterized as: *[2, quadrilateral-square,..., small ..., [0, larger_than_half],[1,smaller_than_half],[3,larger_than_half]],...],* which means "a small square (Object 2) which is larger than half the triangle (Object 0), smaller than half the hexagon (Object 1) and larger than half the pentagon (Object 3)".

Qualitative Models of Orientation

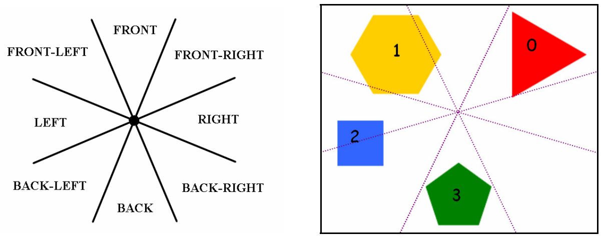
In order to provide a description of the location of the objects wrt the centre of the image, our approach uses Hernandez's qualitative model of Fixed Orientation (FO) (Hernández 1991) and fixes the area named as *front* of the defined reference system to the upper edge of the image.

Moreover, to obtain relative orientations between the objects of an image, our approach applies Freksa's model of Relative Orientation (RO) (Freksa 1992).

Qualitative model of fixed orientation

Represented Information. In order to represent the orientation of an object A wrt the image (A wrt Image) or to represent the orientation of an object A wrt the orientation of another object B (A wrt B), Hernandez's Fixed Orientation Reference System (FORS) is used. This FORS divides space into eight regions (Figure 5a):

FORS $_{LAB}$ = {*front (f)*, *back (b)*, *left (l), right (r)*, *left-front (lf)*, *right-front (rf)*, *left-back (lb)*, *right-back (rb)*}



(a) (b)

Figure 5. (a) Representation of the FORS; (b) FORS applied to image in Figure 1(c).

Application to qualitative description of images. First, our approach for qualitative description of images situates the FORS in the centre of the image in order to obtain a global orientation of all the objects inside that image. Secondly, our approach locates the centre of the FORS on the centroid of each object in order to obtain the location of all the other objects in the image wrt the current one.

The orientation of an object is determined by the union of all the orientation labels obtained for each of the vertices/points of curvature of the object.

Our approach calculates the orientation of a vertex wrt a FORS by obtaining the slope between this vertex and the centre of the FORS and then comparing this slope with the slope of each of the straight lines which define the regions of orientation in the FORS (Figure 5a).

Finally, as an example, the orientation description for the hexagon (Object 1) in Figure 5b is the following string: *[1, hexagon, ..., [front, front_left, left], [[0, left], [2, front, front_right], [3, front, front_left]], ...]* which means "an hexagon (Object 1) located: front/front-left/left wrt the centre of the image, left wrt the triangle (Object 0), front/front-right wrt the square (Object 2) and front/front-left wrt the pentagon (Object 3).

Qualitative model of relative orientation

Represented information. Freksa's model divides the space into 15 qualitative regions by means of a Reference System (RS). This RS is formed by an oriented line determined by two reference points *a* and *b*. From *a* to *b* a line is defined, which determines the left/right dichotomy; another line is defined perpendicular to *b*, which defines the first front/back dichotomy; and another line is defined perpendicular to *a* to establish the second front/back dichotomy. The information which can be represented by this model is the qualitative orientation of a point *c* wrt the RS formed by the points *a* and *b*, that is, *c wrt ab* (Figure 6). This model defines the following regions of orientation:

RORS $_{LAB}$ = { *left-front (lf)*, *straight-front (sf), right-front (rf), left (l), identical-front (idf), right (r), left-middle (lm), straight-middle (sm), right-middle (rm), identical-back-left (ibl), identical-back (ib), identical-back-right (ibr), back-left (bf), straight-back (sb), back-right (br)* }

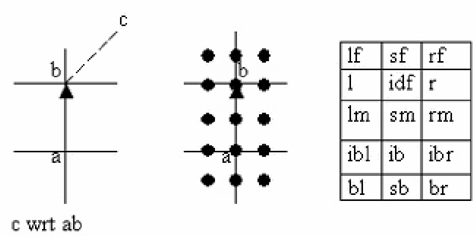


Figure 6. Freksa's model and its iconical representation: RS(a,b) and the 15 qualitative tags of orientation located according to RS(a,b).

Application to qualitative description of images. Our approach establishes a reference system (RS) between all pairs of objects in the image. The points *a* and *b* of the RS are the centroids of the selected objects. All the vertices and points of curvature of the objects that do not compose the RS are located with respect to the corresponding RS. The orientation of an object wrt each RS is the union of all the orientation labels obtained by each vertex or point of curvature of the object.

Our approach calculates the orientation of a vertex wrt a RORS by projecting each vertex to two lines of the RS: first to the line defined by the points *a* and *b* and then to the line perpendicular to *a* or to the line perpendicular to *b*, depending on which are closer to the vertex. If any of the vertices is included in any of these lines, no projection is needed. After having the vertices projected into the RS, we study the orientation of the RS by observing the tendency of the coordinates x and y of the points *a* and *b,* that is if they increase, decrease or remain constant from *a* to *b*. The tendency of the coordinates inside the lines perpendicular to the points *a* and *b* is related to the orientation of the RS. Therefore, if we compare the tendency of the coordinates of a vertex with the tendency of each line of the RS, we could locate the vertex to the *left*/*right* of the line *a*-*b* and to the *front/back* of the line perpendicular to *a* or to the *front/back* of the line perpendicular to *b*, depending on the situation of the vertex. By combining these locations and taking into account if the original vertex was included into a line of the RS or was projected into it, the final orientation of the vertex wrt the RORS is obtained.

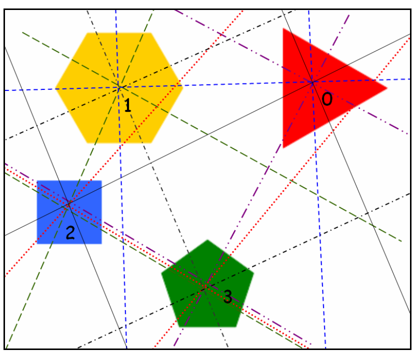


Figure 7. Reference systems obtained for the image in Figure 1c.

In Figure 7, all possible reference systems obtained for the image in Figure 1c are shown. Therefore, according to Figure 7, the orientation of the square (Object 2) is described as the following string:

[2, square, …, [[[0, 1], left_front], [[0, 3], right_middle, right_front], [[1, 0], back_right], [[1, 3], right_middle], [[3, 0], left_middle, back_left], [[3, 1], left_middle]], …]

which means that "the square is located: left-front wrt the RS from the triangle (Object 0) to the hexagon (Object 1), right-middle/right-front wrt the RS from the triangle (Object 0) to the pentagon (Object 3), back-right wrt the RS from the hexagon (Object 1) to the triangle (Object 0), right-middle wrt the RS from the hexagon (Object 1) to the pentagon (Object 3), left-middle/back-left wrt the RS from the pentagon (Object 3) to the triangle (Object 0) and left-middle wrt the RS from pentagon (Object 3) to the hexagon (Object 1)".

Qualitative Model of Compared Distance

In order to describe the distance between the objects in the image, our approach applies Escrig and Toledo's Compared Distance (CD) model (Escrig & Toledo 2001).

Represented Information. In order to compare the distance from the current object or point of view (PV) to another object (A) and the distance from that PV to another object (B), the Compared Distance Reference System (CDRS) defined by Escrig and Toledo is used. This CDRS has two components, CDRS = {RP, LAB}, where RP refers to the Reference Points between which the distances to be compared are calculated; and LAB refers to the labels which represent compared distances. The Referent Points (RP) in CDRS are: the point of view (PV), the first end point (A) to which the first distance is obtained and the second end point (B) to which the second distance is obtained. Compared distance labels are:

$CDRS_{LAB}$ = {*closer_than (ct), nearby (nb), further_than (ft)*}

Application to qualitative description of images. In our approach, the Euclidean distance between the centroid of the current object and the centroid of the other objects in the image is obtained. Then, those distances are compared among them by using the CDRS. Our reference points in the CDRS are: the centroid of the current object or the PV, the centroid of another object (A) to which the first Euclidean distance is calculated, and the centroid of the other object (B) to which the second Euclidean distance is calculated.

As an example, the compared distance for the triangle in Figure 7 is described as the string: *[0, triangle-acute-equilateral, …, [[1,2, closer_than], [2,1, further_than], [1,3, closer_than], [3,1, further_than], [2,3, further_than], [3,2, closer_than]],* which means that "the triangle (Object 0) is: closer to the hexagon (Object 1) than to the square (Object 2), and viceversa, further from the square than from the hexagon; closer to the hexagon (Object 1) than to the pentagon (Object 3), and viceversa, further from the pentagon than to the hexagon; and finally, further from the square (Object 2) than from the pentagon (Object 3), and vicerversa, closer to the pentagon than to the square".

Final String for the Qualitative Description of an Image

Finally, an application that provides the qualitative description of an image containing two-dimensional objects has been implemented.

In general, the structure of the string provided by the application, which describes any image composed by K two-dimensional objects, is defined as a set of qualitative tags as:

[[QVisual, QSpatial]$_1$, …, [QVisual, QSpatial]$_K$]

For each object in the image, its qualitative visual characteristics (QVisual), and its qualitative spatial characteristics inside the image (QSpatial) are described.

The qualitative visual characteristics of the objects (QVisual) consist of the identifier of the object, its name, its regularity, its convexity, its colour in RGB coordinates, the type of object (if it has curves or not), its size wrt the image, the qualitative description of its shape, and a list of compared sizes wrt the other objects:

QVisual = [ObjId, Name, Regularity, Convexity, Colour, Type, Size, QShapeDesc, ListSize$_{wrt}$Obj]

where,

ObjId = N ∈ [0, ∝]
Name ∈ {*triangle, quadrilateral,…, polygon, circle, ellipse, polycurve, mix-shape*}
Regularity ∈ {*regular*, *irregular*}
Convexity ∈ {*convex, concave*}
Colour = [R, G, B] / R, G, B = N ∈ [0, 255]
Type ∈ {*without-curves*, *with-curves*}
Size = {*small*, *medium*, *large*}
QShapeDesc = [[PointQDesc$_1$, …, PointQDesc$_{NmP}$]]
PointQDesc = [Angle|*curve*, Lenght|TypeCurvature, Convexity]

where,
Angle ∈ {*acute, right, obtuse*}
Length ∈ {*smaller, equal, bigger*}
TypeCurvature ∈ {*acute, semicircular, plane*}
Convexity ∈ {*concave*, *convex*}

ListSize$_{wrt}$Obj = [[ObjId, CompSize]$_1$,…, [ObjId, CompSize]$_{K-1}$],

where,
CompSize = {*smaller_than_half*, *half*, *larger_than_half*, *equal*, *smaller_than_double*, *double*, *larger_than_double*}

Finally, the qualitative spatial situation of each object inside the image (QSpatial) consists of (1) the fixed orientation of the current object wrt the centre of the image and a list of fixed orientations of the current object wrt the other objects; (2) a list of relative orientations of the current object wrt all the reference systems in the image; and (3) a list of qualitative compared distances obtained by comparing the distance from the current object to all pairs of other objects in the image.

QSpatial = [Orient$_{wrt}$Image, LOrien$_{wrt}$Obj, LOrien$_{wrt}$RS, LCompQDistance]

where,

Orient$_{wrt}$Image = {*front*, *back*, *left, right*, *front-left*, *front-right*, *back-left*, *back-right, centre*}
LOrien$_{wrt}$Obj =[[ObjId, LOrien]$_1$,…,[ObjId, LOrien]$_{K-1}$]
LOrien = [Orien$_1$, …, Orien$_{no}$]
Orien = {*front*, *back*, *left, right*, *front-left*, *front-right*, *back-left*, *back-right*}
LOrien$_{wrt}$RS = [Orien$_{wrt}$RS$_1$, … Orien$_{wrt}$RS$_m$]
Orien$_{wrt}$RS = [ObjId$_A$, ObjId$_B$, [RelOrien$_1$,..., RelOrien$_{nr}$]]
RelOrien = {*left_front, straight_front, right_front, left, identical_front, right, left_middle, straight_middle, right_middle, identical_back_left, identical_back, identical_back_right, left_back, straight_back, right_back*}
LCompQDistance = [CQDistance$_1$,…, CQDistance$_m$]
CQDistance = [ObjId$_A$, ObjId$_B$, Qdistance]
Qdistance = {*closer_than*, *nearby*, *further_than*}

As a result of the application which implements our approach, the qualitative description of the image shown by Figure 8 is presented. The string obtained describes the four objects contained in the image in the order presented in Figure 8b. Objects containing vertices with smaller coordinates x and y are described first, considering that, in traditional computer vision, the origin of coordinates (x = 0 and y = 0) is located on the upper-left corner of the image. Figure 8b also shows the location of the vertices detected by our approach.

Finally, the qualitative description obtained by our application for the image in Figure 8 is the following one:

[

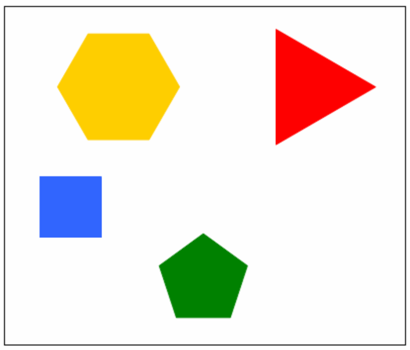
[0, triangle-acute-equilateral, regular, convex, [255,0,0], without_curves, small,
[[acute,equal,convex],[acute,equal,convex],[acute,equal,convex]],
[[1, larger_than_half],[2, smaller_than_double],[3, smaller_than_double]],
[front, right, front_right],
[[1, right], [2, front_right, right], [3, front, front_right]],
[[[1, 2], back_left], [[1, 3], left_middle], [[2, 1], right_front], [[2, 3], left_middle, left_front], [[3, 1], right_middle], [[3, 2], right_middle, back_right]],
[[1, 2, closer_than],[1, 3, closer_than],[2, 1, further_than],[2, 3, further_than],[3, 1, further_than],[3, 2, closer_than]]
],

[1, hexagon, regular, convex, [255, 206, 0], without_curves, small,
[[obtuse,equal,convex],[obtuse,equal,convex],[obtuse,equal,convex],[obtuse,equal,convex],[obtuse,equal,convex],[obtuse,equal,convex]],
[[0, smaller_than_double],[2, larger_than_double],[3, smaller_than_double]],
[front_left, front, left],
[[0, left], [2, front, front_right], [3, front, front_left]],
[[[0, 2], right_middle], [[0, 3], right_middle], [[2, 0], left_middle], [[2, 3], back_left, left_middle], [[3, 0], left_middle], [[3, 2], right_front, right_middle]],
[[0, 2, further_than],[0, 3, closer_than],[2, 0, closer_than],[2, 3, closer_than],[3, 0, further_than],[3, 2, further_than]]
],

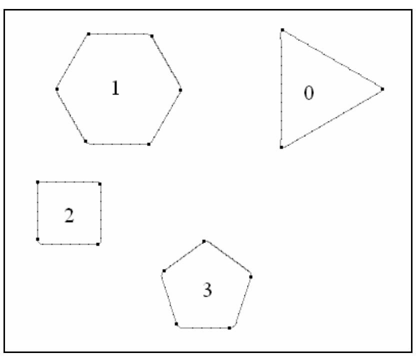
[2, quadrilateral-square, regular, convex, [49, 101, 255], without_curves, small,
[[right,equal,convex],[right,equal,convex],[right,equal,convex],[right, equal,convex]],
[[0, larger_than_half],[1, smaller_than_half],[3, larger_than_half]],
[left, back_left],
[[0, left, back_left], [1, back_left, back], [3, front_left, left]],

[[[0, 1], left_front], [[0, 3], right_middle, right_front], [[1, 0], back_right], [[1, 3], right_middle], [[3, 0], left_middle, back_left], [[3, 1], left_middle]],
[[0, 1, further_than],[0, 3, further_than],[1, 0, closer_than],[1, 3, closer_than],[3, 0, closer_than],[3, 1, further_than]]
],

[3, pentagon, irregular, convex, [0,129,0], without_curves, small,
[[obtuse,smaller,convex],[obtuse,equal,convex],[obtuse,bigger,convex],[obtuse,smaller,convex],[obtuse,equal,convex]],
[[0, larger_than_half],[1, larger_than_half],[2, smaller_than_double]],
[back],
[[0, back_left, back], [1, back_right, back], [2, right, back_right]],
[[[0, 1], left_middle], [[0, 2], left_middle], [[1, 0], right_middle], [[1, 2], left_middle, left_front], [[2, 0], right_middle], [[2, 1], right_middle, back_right]],
[[0, 1, further_than],[0, 2, further_than],[1, 0, closer_than],[1, 2, further_than],[2, 0, closer_than],[2, 1, closer_than]]
].



(a)



(b)

Figure 8. (a) Original image which has been processed by our application; (b) output image after the processes of segmentation and location of the relevant points of the objects. The numbers of the objects have been added previously to the image in order to arrange the qualitative description obtained.

As it can be observed from the qualitative description of the image in Figure 8, Museros and Escrig's method for detecting the vertices of the objects sometimes can be inaccurate, this is the reason why some regular objects are sometimes described as irregular by our approach.

Conclusion and Future Work

This paper has presented an approach for describing images containing 2D objects by applying qualitative models of shape, compared size, fixed and relative orientation and compared distance. Qualitative models of shape and size describe visual features of the objects in the image, while qualitative models of orientation and distance describe spatial features of those objects. As the final result, our approach obtains a string of qualitative labels which can be easily (1) used for computing relationships between the objects in the image and (2) compared to another string that describes another image in order to obtain a degree of similarity between both images.

While the obtained results are encouraging, much research remains in order to apply our approach to real images taken from a camera of a mobile robot. As for future work, we intend to (1) improve the accuracy of the detection process of the vertices of the objects in the image; (2) use the final string obtained by our approach in order to compare images and calculate a degree of visual and/or spatial similarity between them; (3) extend our approach to include topology relations between the objects in the image; and finally (4) apply our approach to qualitatively describe visual landmarks (such as corners or doors) for robot map building and navigation.

Acknowledgements

This work has been partially supported by CICYT and Generalitat Valenciana under grant numbers TIN 2006-14939 and BFPI06/219, respectively. We would like to thank the referees for their comments which helped improve this paper.

References

Bañuelos, J. I.: Algoritmo de indexado en bases de datos de imágenes. Master thesis in Maestría Electrónica supervised by Leopoldo Altamirano Robles. Instituto Nacional de Astrofísica, Óptica y Electrónica. México (2000).

Canny, J. F.; A computational approach to edge detection. IEEE Trans. Pattern Anal. Machine Intell., vol. PAMI-8, pp. 679--697 (1986).

Chang, S-K., Jungert, E., Li, Y.: Representation and retrieval of symbolic pictures using generalized 2D strings. In: SPIE Conference on Visual Communications and Image Processing, pp. 1360--1372. Philadelphia, USA (1989).

Escrig M. T., Toledo F.: Reasoning with Compared Distances at Different Levels of Granularity. In: 9th Conferencia de la Asociación Española para la Inteligencia Artificial (ECAI), Spain (2001).

Freksa, C.: Using Orientation Information for Qualitative Spatial Reasoning. In: Frank, A.U., Campari, I., Formentini, U. (eds.), Theories and Methods of Spatio-Temporal Reasoning in Geographic Space, Proc. Int. Conf. GIS- From Space to Territory. LNCS, vol. 639, pp. 162--178 (1992).

Goldman R. N., Area of Planar Polygons and Volume of Polyhedra. In Arvo J. Ed., Graphics Gems II, The Graphics Gems series, ISBN 0-12-064481-9, pp. 170 -- 171 (1991).

Hernández, D.: Relative Representation of Spatial Knowledge: The 2-D Case. In: Mark, D.M., Frank, A.U. (eds), Cognitive and Linguistic Aspects of Geographic Space, pp. 373--385, Kluwer, Academic Publishers, Dordrecht (1991).

Landau, B., Jackendoff, R.: 'What' and 'Where' in spatial language and spatial cognition. Behavioral and Brain Sciences, 16 (2), 217--265 (1993).

Lovett A., Dehghani M., Forbus K.: Efficient Learning of Qualitative Descriptions for Sketch Recognition. In: 20th International Workshop on Qualitative Reasoning. Hanover, USA, 2006.

Museros L., Escrig M. T.: A Qualitative Theory for Shape Representation and Matching for Design. In: 16th European Conference on Artificial Intelligence (ECAI), pp. 858--862. IOS Press. ISSN 0922-6389 (2004).

Qayyum, Z. U., Cohn, A. G.: Image retrieval through qualitative representations over semantic features. In: 18th British Machine Vision Conference (BMVC2007), pp. 610--619. Warwick, UK (2007).

Socher, G.: Qualitative Scene Descriptions from Images for Integrated Speech and Image Understanding. Dissertationen zur Kunstlichen Intelligenz (DISKI 170). Infix-Verlag, Sankt Augustin (1997).