Belief networks for the perceptual grouping of 3-D surfaces

Liscano, Ramiro; Elgazzar, Shadia; Wong, A.K.C.

Belief Networks for the Perceptual Grouping of 3-D Surfaces

R. Liscano
Ramiro.Liscano@nrc.ca
Institute for Information Technology
National Research Council
Ottawa, Ont. K1A 0R6
CANADA

S. Elgazzar
elgazzar@iit.nrc.ca
Institute for Information Technology
National Research Council
Ottawa, Ont. K1A 0R6
CANADA

A. K. C. Wong
akcwong@watnow.uwaterloo.ca
Dept. of Systems Design
University of Waterloo
Waterloo, Ont. N2K 3G1
CANADA

Abstract

This article introduces an approach, based on Bayesian Networks, for the grouping of 3-D surfaces extracted from data obtained by a laser ranging sensor. A methodology based on the decomposition of an object into its sub-parts is used for specifying the structure of the network. Conditional probabilities are computed using a set of compatibility functions that measure the are a measure in the quality of fit of the data to a model that the features may have come from. These compatibility functions are akin to measures that are used for the perceptual organization of features in the computer vision domain except that they have been developed for 3-D range data. An approach is presented for the mapping of the compatibility functions to conditional probabilities that are required by the Bayesian network. An example of a Bayesian network is presented that models the detection of corners and continuity among planar surfaces and uses both range and intensity values as features sets. The Bayesian network is used to compute a belief value in the formation of corners and continuity among the surfaces which in turn can be used to decide if surfaces should be joined. Results and analysis are presented for an actual set of intensity and range images taken of a typical indoor scene of a robotics laboratory.

1 Introduction

Research in the domain of modeling using 3-D data has primarily focussed on the extraction of 3-D surfaces and volumetric primitives for the purpose of either object recognition or creating more precise models from 3-D sensory data of machined parts. These type of objects can easily be carried and placed in a controlled environment and scanned using a high resolution active sensor. For the modeling of large indoor environments it is necessary to bring the sensor to the environment,

changing the characteristics of the sensed data dramatically. The result of this is that nearly all scans taken in these environments consist of missing data and occluded objects. Recent attempts in the modeling of larger environments [1, 2, 3, 4] from 3-D sensory data have demonstrated it to be a difficult task, due mainly to the amount of missing, sparse, and obscured data. This requires algorithms to be developed that can manage several sources of evidence for determining how surfaces should be grouped and should also maintain belief values in the formation of these groupings.

Research in the modeling of indoor environments has primarily focussed on the incremental synthesis of sensor views and/or position estimation of the sensor [5, 6, 7, 8]. For these systems to become viable tools for Computer Aided Design (CAD) it is necessary to develop approaches that can extract geometric surfaces from the data and hypothesize the formation of more composite features from the surfaces, i.e junctions among the surfaces. In most circumstances a set of heuristics have had to be declared to decide what and how the surfaces should be joined. Unfortunately these heuristics tend to be embedded into the algorithms and no formal approaches are used for maintaining a belief value in the grouping process. This leads to systems that are not easily extendable and the user has no ability to sense of the quality in the grouped surfaces.

This article presents an approach to the grouping of surfaces based on the formalisms developed for Bayesian networks. A Bayesian network offers a cohesive approach to the specification of relationships among feature sets as well as a methodology for the computation of belief values in the existence of a formation among the features. Bayesian networks have been used in the domain of computer vision for object recognition [9, 10, 11, 12], multi-agent vision systems [13], road scene recognition [14], tracking of dynamic objects in images [15], and perceptual grouping for 2-D images [16]. A common thread among those systems in the use of constraints among the features to help either recognize or track objects in a scene. These constraints are a mathematical expression of perceptual organization and the Bayesian network enforces a structure that defines intermediate formations among those features. The Bayesian network also enforces a computational model for belief values in the existence of those formations. This is similar in concept to the system developed by Levitt et al. [9] in that the belief values of the nodes in the Bayesian Network represent a probability of the existence of formations among the features. This is a well sound model and the challenge still remains of determining approaches for constructing the network and determining the conditional probabilities.

2 Constructing the Bayesian Network

This particular implementation of a Bayesian Network is based on the modeling of composite geometrical formations from other fundamental features computed from the sensory data. A composite geometrical formation is created from the aggregation and combination of other geometrical formations until finally it is composed of individual components that can be detected by the sensor.

The constraints that are applied at each level to achieve a geometrical formation can be mapped to a conditional probability for the Bayesian network. The Bayesian Network allows the encoding of the expected geometrical formations that the sensor may detect and infer a belief value in the existence of that formation.

2.1 The Network's Structure

A Bayesian Network is a directed acyclic graph with the nodes representing a hypothesis of the existence of a proposition and the arcs signifying the causal relationship from one proposition to the next. This results in a simple model of the universe but one that can be useful for perceptual grouping. Deciding on the structure and variables of a Bayesian network is a mutual exercise. The approach proposed in this article for guiding the design of the network's structure is the theory of hierarchical decomposition by parts that has been previously used for the modeling of objects [17, 18]. Decomposition by parts is a generic concept that cannot be used directly but as a guide the design of the network. It was also primarily a concept developed for the decomposition of objects for the representation in Computer Aided Design (CAD) that do not take into account the features and components that can actually be detected by a sensor. Figure 1 illustrates a cube decomposed into its components that directly maps to the Bayesian network shown to the right of the decomposition hierarchy.

The process of part decomposition can lead to many solutions and must be guided by the type of sensory data and any intermediate groupings that can be extracted from that data in gathering evidence for the formation of the object. For example, in figure 1 the cube can be decomposed into *Surfaces* and *Vertices* that are formations arising from the detection of *3-D Points* and *Edges*. If it was not possible to detect *Edges* then this structure would collapse to simply the left hand branch modeling the formation of a *Cube* from *Surfaces* and *3-D Points*. The decomposition procedure must also consider the type of compatibility function that will determine how well the data fits the model represented by a node. Each directed edge or set of directed edges can be considered as an application of a constraint to the grouping procedure and therefore implies another compatibility function applied to the data. Later in section 2.2 a procedure will be outlined on how to convert compatibility functions to conditional probabilities, in deciding the structure it is simply a matter of knowing conceptually how this grouping might be performed. For example the decomposition of *Surfaces* to *3-D Points*, in figure 1, can be performed by relaxing an imposed constraint on fitting a set of points to a planar surface. The compatibility function would correspond to a measure of fit of the 3-D points to a planar surface.

Each node in the network is a hypothesized formation representing a boolean variable that reflects the existence of that particular formation from the sensory data. It is possible to create a node that represents several formations [12] but this leads to fairly complicated compatibility functions and a network that is difficult to extend when it is discovered that one particular formation within the node is dependent on another composite formation.

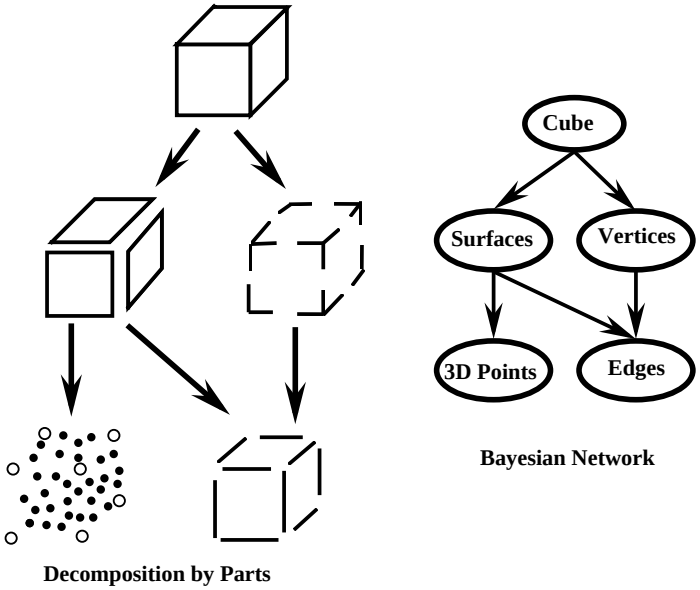

Figure 1: Decomposition by parts of a cube.

The leaf nodes in the network tend to correspond to readily detectable 2-D or 3-D features, for example, edges, 3-D points, intensity, and/or color. It is at this point that any hard evidence will be entered into the network. All other intermediate nodes correspond to hypothesized formations and groupings among the feature sets. For environment modeling the purpose is to detect particular general surface formations that could be used for extracting overall shape of a room.

The steps for creating the structure of the network are the following:

- Define a set of *Boolean* variables that represent the final composite formations that are desired to be hypothesized as well as the variables of the features that are readily detected by the sensor.
- Using the decomposition by parts methodology, decompose the composite formations into sub-formations keeping in mind the following points:
  - Each sub-formations should correspond to the relaxation of a geometric constraint (these will translate directly to compatibility functions).
  - Eventually these sub-formations will connect to the leaf nodes that correspond to the variables representing the features that are readily detected by the sensor.

- Map the links from a parent composite formation to its child sub-formation with a directed edge. These are causal links representing the cause effect relation between the parent grouping to the child sub-grouping.

2.2 The Conditional Probabilities

Conditional probabilities are derived from compatibility functions that measure how well a set of features match a particular perceptual grouping. This perceptual grouping is representative of a particular geometric constraint among the set of features that imposed on the child formation leads to the parent formation. They are not direct measures of how well the features represent a particular formation represented by the node but are simply one measure among several that lead to that desired formation. This is an important point, in that the compatibility function is not representative of the formation of a complicated object but primarily consists of simple perceptual grouping measures leading towards that formation. So for example in terms of the *Cube* formation depicted in figure 1 the compatibility functions measure the planar quality of a surface, proximity among edges, proximity among surfaces, angular relationships among edges, and angular relationships among surfaces. Later in section 3 an actual example is presented for the detection of corners and continuity among planar surfaces. Compatibility functions then are measures that conform closer to those fundamental sets defined originally by the Gestalt laws of organization, i.e. symmetry, proximity, continuity, ... The difference is that they can be developed for different types of sensory data besides the original 2-D human perception world envisioned in the Gestalt laws. With this in mind the following definition for compatibility functions is presented.

A **compatibility function** is a function ($f_{PG}(\mathbf{A})$) based on a geometric constraint applied to a set of attributes $\mathbf{A}$ of a sensory feature that measure how well these features match the perceptual grouping PG.

From this point on, compatibility functions will be referred solely as f_{PG} without the argument $\mathbf{A}$.

In order to place compatibility functions on a same level of reference a mapping function is used to map the results from a compatibility function to a certainty value that exhibits the following properties,

- It is bounded between the interval $(0, 1)$.
- It is a decreasing monotonic function, so that a compatibility value equal to 0 signifies a certainty value of 1. This is equivalent to a perfect match between the features and the perceptual grouping.

There are several of these type of mappings and in this particular implementation we propose the declining S-Curve function, shown in figure 2 and represented mathematically as,

$$S(f_{PG}) = \begin{bmatrix} 1 & \rightarrow & f_{PG} = 0 \\ 1 - 2(f_{PG}/\gamma)^2 & \rightarrow & 0 \leq \beta \\ 2((f_{PG} - \gamma)/\gamma)^2 & \rightarrow & \beta \leq \gamma \\ 0 & \rightarrow & f_{PG} \geq \gamma \end{bmatrix} \quad (1)$$

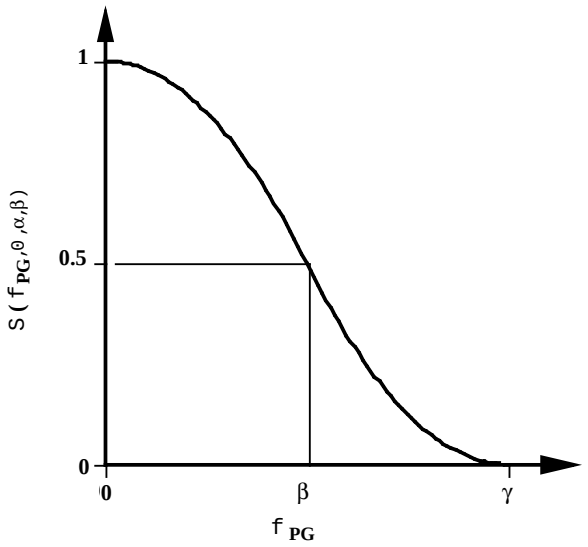

Figure 2: A declining S-curve.

This mapping offers the following advantages to the user:

- It maps a physical measure to a subjective confidence value and therefore separates these 2 conceptually different environments. The user can operate with physical forms of measurements and then apply a subjective measure that represents how well that physical measure comes close to a particular perceptual grouping;
- The S-Curve function allows the user to define the resolution of a match of the geometric features with the perceived formation. It is possible to use the same compatibility functions but different S-Curves for 2 different geometrical formations;

Compatibility functions and constraint functions have the drawback that they are used mostly as a means of determining how well the evidence supports a hypothesized formation. Therefore they cannot be used as a mechanism for determining a-priori probabilities that are required for

causal networks. For example, each node represents a Boolean variable whose states are defined as $A_i = (a_i, \neg a_i)$, where a_i and $\neg a_i$ respectively signify the existence and non-existence of the formation A_i. The compatibility functions can be used directly for computing the conditional probabilities when the parent's states are **TRUE** but no such mapping exists for when they are **FALSE**. This does not completely prevent the use of Bayesian networks but requires that the directional edges be reversed so as to represent consequential knowledge, i.e. from evidence to hypothesis. For example, the cube decomposition represented in figure 1 must now be inverted as shown in figure 3.

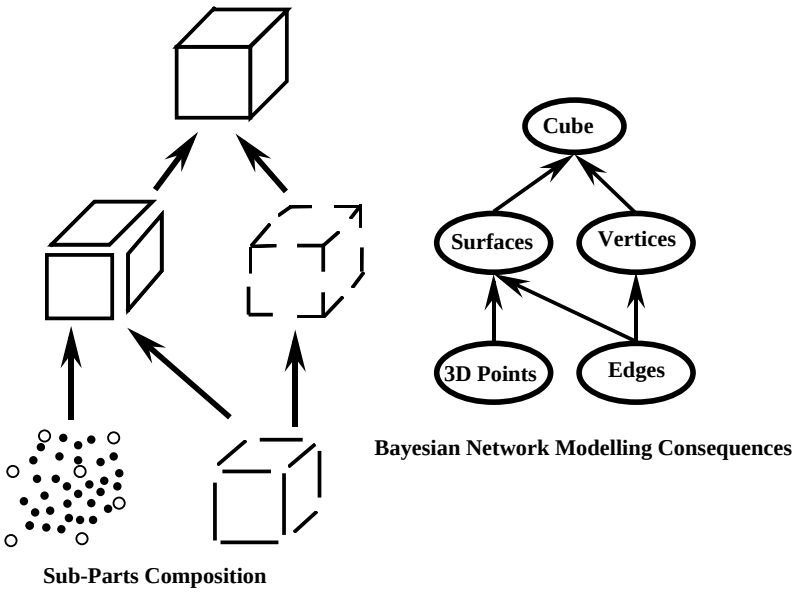

Figure 3: Formation of cube as a consequence of the evidence.

In this direction of computation the network resembles the Perceptual Inference Network (PIN) developed by Sarkar and Boyer [16] with the significant difference that the structure of the network is developed using causal theory and the perceptual grouping constraints are associated with the edges of the network. The advantages of using causality over consequences are the removal of multiple instantiations of similar nodes that can occur when the structure of the network is built using consequences. For example, the existence of a set of parallel lines may be caused by multiple formations and is reflected in a causal network by a node representing parallel lines and multiple parent formations that have caused the parallel formation. When the network is designed to reflect consequences it is fairly easy to introduce another node representing a parallel formation anywhere along the network. This of course is a design methodology than a limitation imposed by the actual

network itself.

Let $pa(A_i)$ be a function that returns the parent nodes of A_i. It is only necessary then to specify the conditional probabilities for the cases when the states of $pa(A_i)$ are **TRUE**. For single parent nodes the S-curve mappings can be mapped directly into conditional probabilities resulting in the following conditional probabilities for serial connections,

$$\begin{array}{rcl} P(a_i|a_j) & = & \mathcal{S}(f_{PG}) \\ P(\neg a_i|a_j) & = & 1 - \mathcal{S}(f_{PG}) \end{array} \tag{2}$$

where $pa(A_i) = A_j$.

For nodes with multiple parents, i.e. convergent nodes, the approach must account for the edges to the other parent nodes. There are 2 possible situations that occur; either the formation is a consequence of similar features being grouped or evidence has to be combined from the grouping of dissimilar feature sets. A simple example of this is portrayed in figure 3, where edges can be used to support the formation of vertices while both 3-D points and edges are used to support the formation of surfaces.

If the composite formation represents the grouping of similar features then a compatibility function can be developed that measures the quality of fit of the features to a model of the composite formation. This is typically what has been developed in the computer vision domain for the grouping of edges. For these type of formations the conditional probabilities can be computed using equation2.

For the situations where a composite formation can come about from several different types of feature sets separate compatibility functions are required for each different set of features and the results of these must be combined. For these cases, each directed edge, e_{A_j,A_i}, from a parent node A_j to a child node A_i has associated with it a compatibility function and a corresponding S-Curve mapping. With this in mind the following calculation for computing the conditional probabilities is proposed,

$$\begin{array}{rcl} P(a_i|pa(A_i)) & = & \displaystyle\prod_{A_j \epsilon pa(A_i)} \mathcal{S}(f_{PG}(e_{A_j,A_i})) \\ P(\neg a_i|pa(A_i)) & = & 1 - \displaystyle\prod_{A_j \epsilon pa(A_i)} \mathcal{S}(f_{PG}(e_{A_j,A_i})) \end{array} \tag{3}$$

These 2 conditions can occur together which can be managed by considering one compatibility function for similar feature sets and combining the compatibility functions of dissimilar feature sets using equation 3.

For the situations when any of the states of $pa(A_i)$ are **FALSE** then the following conditionals apply,

$$\begin{array}{rcl} P(a_i|pa(A_i)) & = & 0 \\ P(\neg a_i|pa(A_i)) & = & 1 \end{array} \tag{4}$$

3 Detecting Corners and Continuity

This section will present an example of a Bayesian network for the detection of corners and continuity among planar surfaces. This network is depicted in figure 4 and models the grouping of surfaces and their respective edges into corners and continuous surfaces.

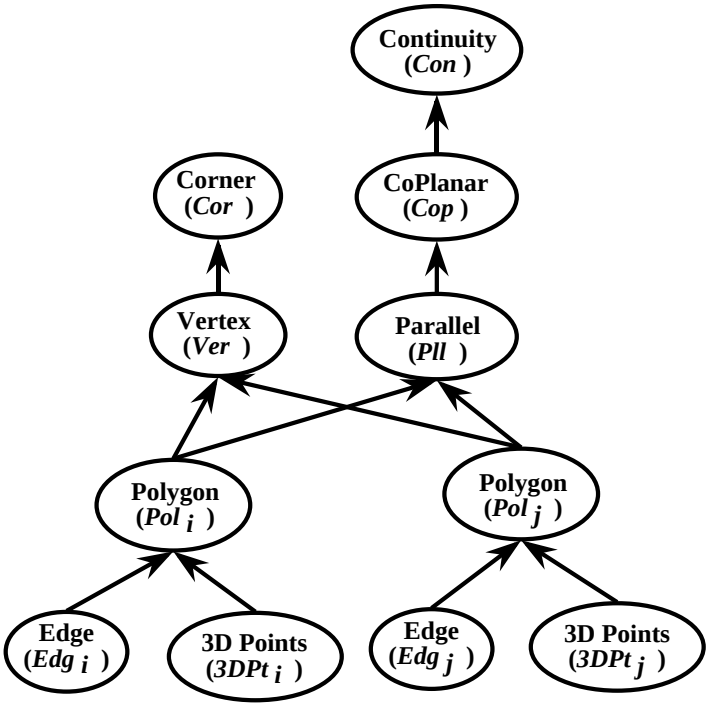

Figure 4: A Bayesian network for the detection of corners and continuity among planar surfaces.

This particular example is a Bayesian network depicting the computation of belief values in the direction from evidence to hypotheses, or in the direction of consequences. Figure 4 is a grouping model depicting the existence of corners (**Cor**) and continuous surfaces (**Con**) from the existence of coplanar surfaces (**Cop**), parallel surfaces (**Pll**), vertices (**Ver**) , polygons (**Pol**), and a set of 3-D points (**3DPt**) and their respective edges (**Edg**). The 3-D points and edges are the only instantiated nodes in the network since they represent actual sensory data, the other nodes are hypothesized formations that contain belief values computed from the instantiated evidence.

This particular structure decomposes the formation of corners and continuous surfaces by relaxing one particular constraint among the grouping for each level in the network as one traverses the network from top to bottom. This is a decomposition by parts operation with intermediate nodes

that depict the relaxation of one single constraint at a time. Figure 4 also notes which perceptual grouping compatibility functions are being applied as one moves up the network. In this network one can see the grouping of similar features; for example, the formation of parallel surfaces among 2 polygons using one compatibility function (f_{prll}), and the grouping of dissimilar features, like the formation of polygons from a set of 3-D points and their corresponding boundary that uses 2 compatibility functions (f_{plan} and f_{edge}).

Also similar compatibility functions are used for the formation of vertices (**Ver**) as well as that for parallel surfaces (**Pll**). These 2 formations are contrary to each other and a different S-curve function is used to actually compute the conditionals.

3.1 Compatibility Functions

A very brief listing of the compatibility functions are presented with appropriate references to any detail information. Some of these functions, in particular proximity among surfaces, are fairly unique and require more explanation than can be performed in this article.

3.1.1 Planar Surfaces (f_{plan})

A compatibility function for determining the quality of fit for a set of points to a planar surface is to take the mean of the squared distance of the points from the surface. The result is the following compatibility function for planar surfaces from 3-D points,

$$f_{plan}(d_j) = \frac{1}{N}\sum_j d_j^2, \tag{5}$$

where d_j is the distance of point j from the surface and N are the number of points that define the surface.

3.1.2 Parallel (f_{prll}) and Coplanar (f_{copl})

A measure of parallelism between two planar surfaces can be derived by evaluating the length of the vector computed from the cross product of the normals to the surfaces. This leads to the following geometrical compatibility function,

$$f_{prll}(S_i, S_j) = |\mathrm{N}_{S_i} \times \mathrm{N}_{S_j}|, \tag{6}$$

where N_{S_i} and N_{S_j} are the normals corresponding to the planar surfaces S_i and S_j.

Coplanar surfaces are a restricted case of parallel surfaces with the added constraint that the angle between the normals of the surfaces and the line joining the center points of the two surfaces is approximately 90°. This leads to the following equation,

$$f_{copl}(S_i, S_j) =< \mathrm{N}_{S_i}, C_{S_j} - C_{S_i} >, \tag{7}$$

where C_{S_j} and C_{S_i} correspond to the location of the centers of the planar surfaces S_j and S_i respectively and the $< a, b >$ operator is a dot product operation between vectors a and b. This is a fairly loose constraint that actually measures how well the centroids among 2 surfaces are aligned that along with equation 6 does lead to coplanar surfaces.

3.1.3 Proximity (f_{prox})

Defining a proximity compatibility function for 3-D surfaces is challenging, because the definition of proximity is not unique. The proposed approach defines a compatibility function among the surfaces by measuring the distance between the boundary of two surfaces and is summarized in the following equation,

$$f_{prox}(S_i, S_j) = \frac{Area(Gap(S_i, S_j))}{Area(S_i + S_j)}, \tag{8}$$

where $Gap(S_i, S_j)$ computes a polygon with a boundary common with the surfaces S_i and S_j based on their respective boundaries. The common boundary is determined by approximating the surface by a polygon and projecting the vertices of the polygon away from the surface until they intersect with another adjacent surface's boundary. This operation can be performed in the 2-D image plane representative of an ordered set of 3-D points. This ordered set of 3-D points is a common representation for scanning type of range sensors. The common boundaries between the 2 surfaces are used to define a 3-D surface composed of triangles of whose surface area can be approximated by adding up the surface area of the triangles. This procedure is rather challenging and the approach offers a unique method for determining adjacent surfaces [19].

3.1.4 Surface Boundary (f_{edge})

This compatibility function measures how well the edge of a planar surface defines a polygon representing that surface. The compatibility function involves a ratio of the area defined by the polygon to the area of the surface itself and is computed using the following equation,

$$f_{edge}(S_i, P_{S_i}) = \frac{Area(S_i) - Area(P_{S_i})}{Area(S_i)}, \tag{9}$$

where $Area(S_i)$ computes the surface area based on the points defining the surface and $Area(P_{S_i})$ computes the surface area based the polygon, P_{S_i} associated with the surface A_i. These two values in most situations are not equivalent since the edges extracted for the surface are straight while the original surface will consist of irregularly shaped edges. The estimation of the surface by a polygon is based on determining high curvature points along the surfaces boundary [19]

4 Experimental Results

The Bayesian network and compatibility functions described in section 3 were applied to a set of planar surfaces extracted from range points acquired using a compact laser camera called BIRIS [20]. The output of the camera consists of one acquisition of 256 range and intensity values along a projected plane of light. To acquire more data than that of a single acquisition the BIRIS sensor was mounted onto a pan and tilt unit. The ability to tilt the sensor is crucial in being able to acquire more data, since the field of view of the sensor is fairly limited. When the Biris Laser Scanner is panned at a constant speed for a fixed tilt angle the result is a rectangular image of dimension 256 x N_{aq}, where N_{aq} is the number of acquisitions taken during the panning sequence. The characteristics of this particular version of the sensor are shown in table 1.

Laser power	24 mW, (two 12 mW laser projectors)
wavelength	680 nm (visible red)
Field of view	19 deg.
Focal Length	20 mm
Accuracy	3 mm @1 m, 14 mm @ 2 m, 42 mm @ 3 m
Range	0.5 m - 5.0 m

Table 1: Specifications for the BIRIS range sensor.

Figure 5 shows the extracted planar surfaces from range data taken of a room with a typical layout approximately the shape of that shown in figure 6. The individual segments in the intensity image are shown using separate gray values to represent each segment. The advantage of maintaining the points in an image is that one can take advantage of the already established scanning ordering inherent in the sensor.

The Bayesian network depicted in figure 4 was applied to each set of adjacent surfaces in figure 5. In this example there were 25 total surface groupings of which the computed belief values are presented in table 2. Table 2 also lists a small comment for each grouping that mentions reasons why the belief values for particular surface formations are low. Low values for both corner and continuity do not necessarily signify bad results they can reflect situations where the edge shared by the 2 adjacent surfaces is in fact a jump edge and the surfaces cannot form either a corner or continuous surface.

The results are consistent in that the formation of a corner is counter to the formation of continuous surfaces. The belief values reflect this in the following manner. When the belief of the formation of a corner is high the belief in the formation of a continuous surface is low, and the converse is true. The exception are cases where both values are close to 0.0 which in some cases are characteristic of surfaces with "Jump" edges and in other cases low certainty values caused by

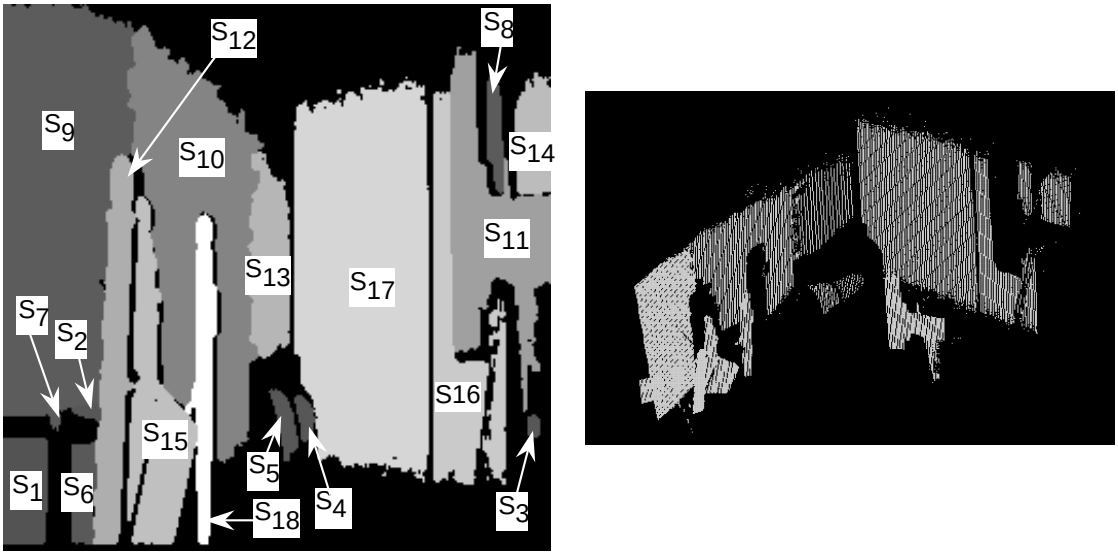

Figure 5: Surfaces extracted from the range data; intensity image (a) and isometric view (b).

high compatibility functions. Using the Bayesian network it is rather simple to infer why these low values in both corners and continuity exists, it is simply a matter of searching through the graph looking for occurrences in low compatibility values.

Figure 7 depicts the results in table 2 as an image where the surface pairs that have belief values (0.7 or over) of being continuous or corners have been displayed using the same intensity value. Surfaces that had low belief values in either being planar or had bad edge representations are depicted in a dark color with only their boundaries showing.

These results suggest that another surface grouping, known as a "Jump Surface", should perhaps be introduced to the model and will be considered for future networks. These type of surfaces are also presented in figure 7. Discontinuities, like "Jump Surfaces", are generally not considered a perceptual grouping formation but certainly do add knowledge that can be used for interpreting a scene.

Some results are not so positive but do reflect the reality of the situation. It would have been desirable to have surfaces S_{16} and S_{17} to result in a high belief value of being a continuous surface. This was not so, the reason was in the low value that surface S_{16} was a planar surface. At least this low certainty value in the surface's planar quality is reflected throughout other formations where surface S_{16} is part of and a low belief value is computed that can hint to a further investigation

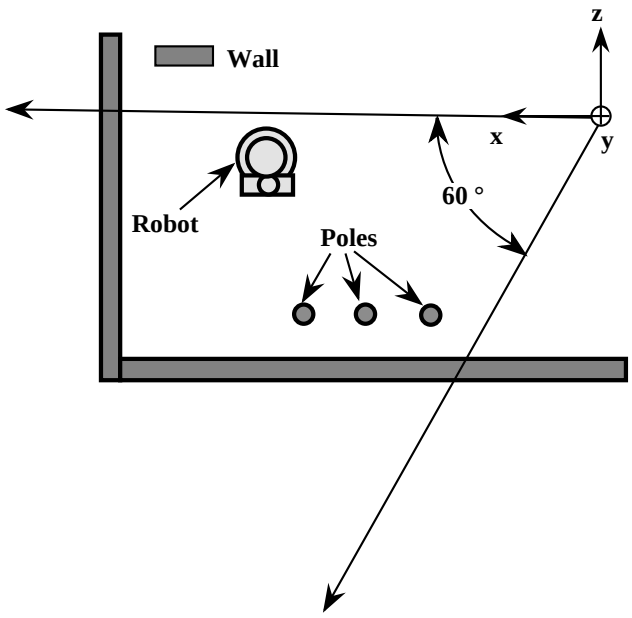

Figure 6: A schematic of the top view of the robot lab showing sensor position.

into the reason behind the small value.

It should be notes that the results obtained are biased by the parameters used in the S-curve mapping. Table 3 lists the values used for the arguments β and γ for S-curve mapping.

These values were subjectively selected but the selection follows a certain reasoning. The proximity value of 0.25 represents the desire to consider surfaces with gaps smaller than 1/4 the sum of the areas of the 2 neighbouring surfaces as proximal, anything beyond 0.5 is not proximal. This same argument can be applied to the edge compatibility value since it is also a ratio of areas. Surfaces are considered parallel if the surface normals are within 10 rads with respect to each other, and they are not parallel if beyond 20 rads. Coplanarity is similar to parallel surfaces and share the same arguments. The parameter values for the quality of the surface being considered as a plane was determined by computing the average in the variance of the data points of a typical set of points gathered from the BIRIS sensor. Using the actual variances in the data as a guide for determining the planarity arguments for β and γ is important since the sensor itself is not an accurate sensor at long ranges. If an arbitrary low absolute value is used the result would be small belief values in the existence of all planar surfaces. These current settings show that the robot and pole surfaces are reasonable planar surfaces which is not the case but because they are closer to the sensor have smaller variances in the range then the background walls.

S_i	S_j	Cor	Con	Comment	S_i	S_j	Cor	Con	Comment
1	6	0.32	0.0	$S(f_{prox}$ low	10	12	0.32	0.0	**OK** Jump Edge
1	9	0.87	0.0	**OK**	10	13	0.01	0.72	**OK**
4	5	0.0	0.0	$S(f_{plan}(S_4)) \approx 0.0$	10	15	0.0	0.0	**OK** Jump Edge
4	17	0.0	0.0	$S(f_{plan}(S_4)) \approx 0.0$	10	17	0.9	0.0	**OK**
5	10	0.0	0.0	$S(f_{plan}(S_5)) \approx 0.0$	10	18	0.0	0.0	**OK** Jump Edge
5	13	0.0	0.0	$S(f_{plan}(S_5)) \approx 0.0$	11	14	0.0	0.0	**OK** Jump Edge
5	17	0.0	0.0	$S(f_{plan}(S_5)) \approx 0.0$	11	16	0.0	0.0	**OK** Jump Edge
6	9	0.63	0.0	$S(f_{edge}(S_6))$ low	12	15	0.0	0.0	**OK** Jump Edge
6	12	0.0	0.0	**OK** Jump Edge	13	17	0.77	0.0	**OK**
8	11	0.0	0.0	$S(f_{plan}(S_8)) \approx 0.0$	15	18	0.0	0.0	**OK** Jump Edge
8	14	0.0	0.0	$S(f_{plan}(S_8)) \approx 0.0$	16	17	0.02	0.27	$S(f_{plan}(S_{16})) \approx 0.0$
9	10	0.87	0.0	**OK**	17	18	0.0	0.0	**OK** Jump Edge
9	12	0.22	0.0	**OK** Jump Edge					

Table 2: Belief values for the formation of corners and continuity among the surfaces in figure 5 (a).

Compatibility	β	γ	Units
proximity	0.25	0.50	unitless
coplanarity	10	20	Rads
parallel	10	20	Rads
planarity	0.578	1.156	cm
edges	0.25	0.50	unitless

Table 3: Values for the arguments β and γ for particular compatibility functions.

5 Conclusions

This article presented an approach in the use of Bayesian networks for the grouping of features and the detection of higher order complex structures. An example was presented of a Bayesian Network for the grouping of 3-D surfaces into either corners or continuous planar surfaces. An approach for the specification of the Bayesian Network was presented that used the concept of part decomposition for determining the structure of the network and perceptual grouping compatibility functions that can be used to determine the conditional probabilities required by the network. It was shown that the design of the network can be done by considering causal relationships from

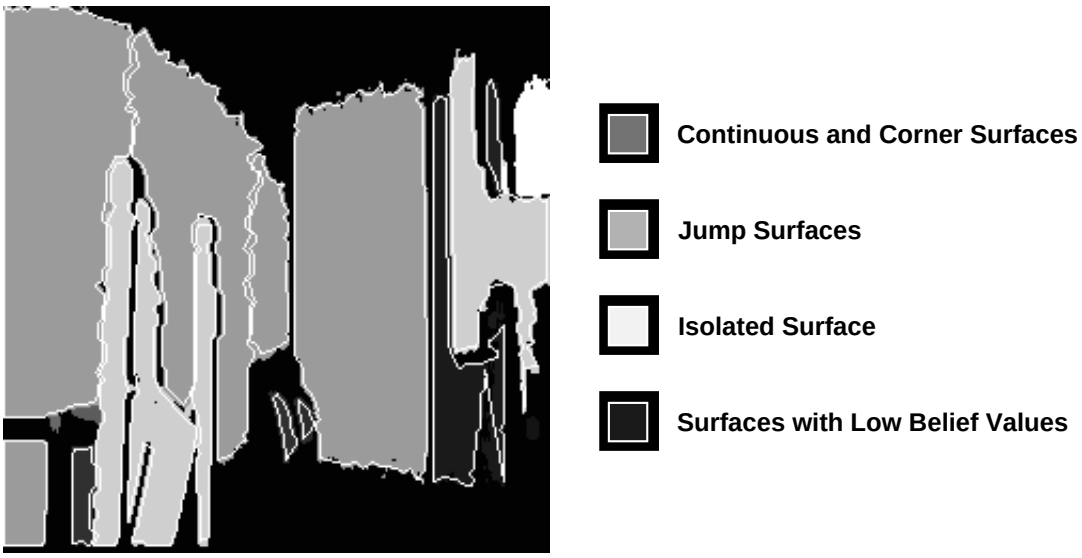

Figure 7: Image of the surfaces depicting those that are continuous, corners, or jump surfaces.

the model to its sub-parts but the actual implementation and computations must be performed in a bottom up fashion from the sub-parts to the model. This is imposed by the inability to specify compatibility functions for the conditions when the parent nodes are **FALSE**. A network was presented that grouped surface data into corners and continuous surfaces that was tested on 3-D sensory data captured from a typical indoor environment.

The grouping results appear to be consistent but what is more beneficial is the declarative nature of procedure implied by the use of a Bayesian network. This facilitates the search for reasons why belief values may be low. The procedural knowledge required to perform the grouping is also well encapsulated as compatibility functions, facilitating the use and development of the network.

To reduce the inherent computational complexity associated with grouping operations the grouping process was only applied to adjacent surfaces. The next challenge is to determine an approach for comparing surfaces that are not directly adjacent without having to compare all the surfaces. Once this issue is resolved it will be possible to apply the same Bayesian network procedure to the detection of occluded surfaces. Clearly there are many other examples that need to be investigated like, 3 surface corners, jump surfaces, and textured surfaces like brick walls. Development has been occurring into a Bayesian Attributed Hypergraph (BAHG) that combines the benefits of Bayesian networks and attributed hypergraphs. This representation will facilitate the computation of the compatibility functions by integrating the relationships and attributes with the creation of the Bayesian network. Currently multiple separate networks are created as the

pair of surfaces are compared. The BAHG combines this into one network and therefore facilitates multiple surface groupings.

References

[1] F. Nashashibi and M. Devy, "3-D incremental modeling and robot localization in a structured environment using a laser range finder," in *Proceedings of the 1993 IEEE International Conference on Robotics and Automation*, (Atlanta, GA), pp. 20–27, IEEE Press, May 1993.

[2] V. Sequeira, J. G. M. Goncalves, and M. I. Ribeiro, "3-D scene modelling from multiple views," in *Proceedings of the SPIE: Videometrics IV*, vol. 2598, (Philadephia, Pa.), pp. 114–127, October 25-26 1995.

[3] J. Hoshino, T. Uemura, and I. Masuda, "Region-based reconstruction of an indoor scene using an intergration of active and passive sensing techniques," in *IEEE Third International Conference on Computer Vision*, (Osaka, Japan), pp. 568–572, IEEE Press, December 4-7 1990.

[4] R. Fayek, *3-D Surface modeling using hierarchical topographic triangular meshes.* PhD thesis, University of Waterloo, Waterloo, Ontario, Canada, 1996.

[5] Z. Zhang and O. D. Faugeras, *3-D Dynamic Scene Analysis: A Stereo Based Approach.* Springer Series in Information Sciences, Berlin: Springer-Verlag, 1992.

[6] N. Ayache, *Artificial vision for mobile robots: stereo vision and multisensory perception.* Cambridge, MA: MIT Press, 1991.

[7] R. Liscano, R. Fayek, A. Manz, E. Stuck, and J.-Y. Tigli, "Using a blackboard to integrate multiple activities and achieve strategic reasoning for mobile-robot navigation," *IEEE Expert*, vol. 10, no. 2, pp. 24–36, 1995.

[8] P. Boulanger and F. Blais, "Range image segmentation, free space determination, and position estimate for a mobile vehicle," in *SPIE Proceedings, Mobile Robots VII*, vol. 1831, (Boston, MA), pp. 444–455, SPIE Press, November 18-20 1992.

[9] T. Levitt, T. Binford, and G. Ettinger, "Utility-based control for computer vision," in *Uncertainty in Aritificial Intelligence 4* (R. D. Schachter, T. S. Levitt, L. N. Kanal, and J. F. Lammer, eds.), pp. 407–421, Amsterdam: Elsevier Science, 1990.

[10] R. Munk-Fairwood, "Recognition of geometric primitives using logic program and probabilistic network reasoning methods," in *Proceedings of SPIE: Applications in AI X; Machine Vision and Robotics*, vol. 1708, (Orlando, FL), pp. 589–600, SPIE Press, 1992.

[11] S.-H. Wang and Q.-G. Gao, "High-level decisions in belief networks based 3-D object recognition," in *IEEE Int. Conf. on Systems, Man, and Cybernetics*, vol. 3, (Vancover, Canada), pp. 2742–2747, IEEE Press, October 1995.

[12] J. Liang, F. C. Jensen, and H. I. Christensen, "A framework for generic object recognition with bayesian networks," in *Intelligent Industrial Automation, IIA'96*, (Reading, UK), pp. C9–15, March 26-28 1996.

[13] F. V. Jensen, H. I. Christensen, and J. Nielsen, "Bayesian methods for interpretation and control in multi-agent vision systems," in *Proceedings of SPIE: Applications in Artificial Intelligence X: Machine Vision and Robotics*, vol. 1708, (Orlando, FL), pp. 536–548, 1992.

[14] G. Foresti, V. Murino, C. S. Regazzoni, and G. Vernazza, "A distributed approach to 3-D read scene recognition," *IEEE Trans. on Vehicular Technology*, vol. 43, pp. 389–406, May 1994.

[15] S. Gong, "Bayesian net for functional integration of visual surveillance system," in *PRICAI94: 3rd Pacific Rim Int. Conf. on Artificial Intelligence*, vol. 2, (Beijing, China), pp. 850–856, August 15-18 1994.

[16] S. Sarkar and K. L. Boyer, "Integration, inference, and management of spatial information using bayesian networks: perceptual organization," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 256–274, Mar. 1993.

[17] T. S. Levitt, "Model-based probabilistic inference in hierarchical hypothesis spaces," in *Uncertainty in Artificial Intelligence* (L. Kanal and J. Lemmer, eds.), pp. 347–356, Amsterdam: Elsevier Science, 1986.

[18] I. Biederman, "Human image understanding," in *7th Scandinavian Conference on Image Analysis*, (Aalborg, Denmark), 1991. late paper.

[19] R. Liscano, S. Elgazzar, and A. K. C. Wong, "A proximity compatibility function among 3-D surfaces for environment modelling," in *5th IASTED International Conference on Robotics and Manufacturing*, (Cancun, Mexico), pp. 161–166, IASTED, May 19-21 1997.

[20] S. Elgazzar, R. Liscano, F. Blais, and A. Miles, "3-D data acquisition for indoor environment modeling using a compact active range sensor," in *IEEE Instrumentation, Measurement and Technology Conference (IMTC '97),*, (Otawa, Ontario), pp. 586–592, IEEE Press, May 19-21 1997.