Robin Burk
CS551

How Can You Tell That's a Door?

1.0 Introduction and Motivation

Object recognition is a task that people learn as small children, but which presents significant challenges for robotics. For some purposes, objects can be recognized by robots through detection of color transitions (outlines) presumed to be unique in a given environment, as with the spherical ball used in robot soccer competitions. Where outline alone is not determinative, but where the robot is only expected to function within a specific environment, techniques such as artificial neural nets can be used to identify environmental objects with high reliability.

A harder task, however – but one which, if successful, would significantly enhance the flexibility and usefulness of autonomously mobile robotic equipment – is the cognitive robotics challenge of object recognition based on probabilistic inferences from a range of attributes commonly used by humans for the same purpose.

Success at this task would confer two benefits. First, it would reduce the training required for object recognition software, allowing probability updates based on future experiences in new environments and removing the stringent requirement that a sufficiently large and representative training data set be collected prior to robot use. Second, and more significantly, a robot which can identify indoor objects based on attributes commonly used by humans can potentially also be given navigation instructions based on such objects and their attributes – i.e. in terms which come naturally to people.

The purpose of this project was to prototype the creation of a Bayesian belief net (BBN) which might be used to deduce the probable identity of an object (door, wall segment or hallway opening). Belief nets capture a graph of causal relationships along with conditional probabilities for the joint occurrence of node values. Once an initial BBN has been established, it can be used in multiple ways. First, on the assumption that the probabilities hold true for new instances, the likelihood of an object identity can be inferred based on a set of observed attribute values. Second, BBN probabilities can be updated using artificial intelligence techniques when new data are collected.

Reasoning based on belief nets is attractive in many ways. BBNs allow complex levels of causality to be represented in a single model which is easily comprehended and validated by humans. Bayesian reasoning has its disadvantages, however, including the fact that full computation of complex nets is intractable (NP complete). Thus to be useful Bayesian models must contain the smallest set of nodes that suffice to make sufficiently accurate predictions for the purpose at hand.

This paper describes an initial effort at constructing a simple Bayesian belief net suitable for distinguishing doors, wall segments and hallway openings in an idealized indoor environment. The objects were chosen to support broader research that builds on the work of Haas and Shimizu in natural language processing for robot navigation commands. Their work occurred in a simulated environment where objects are known. This project begins an effort to move their NLP approach into a real robot which must recognize objects in a variable physical environment.

2.0 The Data

A total of 89 object observations were collected. Of these, 44 were for doors, 32 for wall segments and 13 for hallway openings. The venues for these observations included 3 residences and 2 public institutions (university buildings). Although not necessarily representative of the full set of all indoor environments, it is likely that these offer a suitable basis for initial Bayesian inferences, especially in light of the regularizing role of building codes and custom for some dimensions of interest such as minimum door width and ceiling height.

Attributes measured included height and width of the object, the height of any wall above the object, the width and depth of any frame around the object, color and reflectivity of the surface and the depth of space beyond the object, if any. (Only closed doors were used for this initial effort.) Table 1 summarizes the raw observations.

Table 1. The Primary Distributions (measurements in inches):

	Height	width	hgt above	h/w ratio
For all observations:				
mean	89	56	11	2.02
median	84	36	16	2.24
mode	96	30	0	2.60
stdev	14	36	11	0.71
Doors:				
mean	77	32	21	2.47
median	78	30	20	2.55
mode	78	30	16	2.60
stdev	4	5	6	0.25
Walls:				
mean	96	90	0	1.31
median	96	96	0	1.00
mode	96	120	0	0.80
stdev	0	35	0	0.70
Openings:				
mean	104	61	8	1.83
median	104	41	0	1.95
mode	120	36	0	2.89
stdev	14	37	11	0.68

Robin Burk
CS551

In an effort to find useful categories for a Bayesian belief net, the discrete measurements were binned into three ranges for height, width and height/width ratio, with each range consisting of the mean plus/minus one standard deviation (medium range) and the lower and upper tails of the sample distribution. Frames and depth of space were binned into binary categories (present, absent). Table 2 gives the probability of each resulting object/attribute value pair.

Table 2. Joint probabilities

	low hgt	med hgt	high hgt
door	0.1124	0.3820	0.0000
wall	0.0000	0.1124	0.2472
opening	0.0000	0.1011	0.0449

	low h/w	med h/w	high h/w
door	0.0000	0.4382	0.0562
wall	0.2022	0.1124	0.0449
opening	0.0225	0.1236	0.0000

	low width	med width	high width
door	0.0000	0.4944	0.0000
wall	0.0000	0.1798	0.1798
opening	0.0000	0.1011	0.0449

	frame	no frame
door	0.4944	0.0000
wall	0.0000	0.3596
opening	0.0000	0.1460

	low above	med above	high above
door	0.0000	0.3258	0.1685
wall	0.3596	0.0000	0.0000
opening	0.1011	0.0449	0.0000

	depth	no depth
door	0.0000	0.4934
wall	0.0000	0.3596
opening	0.1461	0.0000

Normalizing the joint probabilities by P(object) gives the likelihoods P(attribute value | object) shown in table 3. As these likelihoods show, several attribute value categories contribute little to identifying objects correctly. Since the computational time needed to reason using a Bayesian belief net is exponential in the number of nodes * attribute categories, collapsing these categories is valuable when we can do so without reducing predictive power.

Therefore, I chose the likelihood breakouts shown in Table 4. as the basis of my belief net. Width was collapsed into 'high width' and 'not high width'. Height/width was not sufficiently predictive on its own to warrant replacing both height and width, nor does it add predictive power when they are present. Therefore it was dropped from the final categories.

Robin Burk
CS551

Table 3. Likelihoods

	low hgt	med hgt	high hgt
door	0.2273	0.7727	0.0000
wall	0.0000	0.3125	0.6875
opening	0.0000	0.6923	0.3077

	low h/w	med h/w	high h/w
door	0.0000	0.8864	0.1136
wall	0.5625	0.3125	0.1250
opening	0.1538	0.8462	0.0000

	low width	med width	high width
door	0.0000	1.0000	0.0000
wall	0.0000	0.5000	0.5000
opening	0.0000	0.6923	0.3077

	frame	no frame
door	1.0000	0.0000
wall	0.0000	1.0000
opening	0.0000	1.0000

	low above	med above	high above
door	0.0000	0.6591	0.3409
wall	1.0000	0.0000	0.0000
opening	0.6923	0.3077	0.0000

	depth	no depth
door	0.0000	1.0000
wall	0.0000	1.0000
opening	1.0000	0.0000

Table 4. Modified categories.

	low hgt	med hgt	high hgt
door	0.2273	0.7727	0.0000
wall	0.0000	0.3125	0.6875
opening	0.0000	0.6923	0.3077

	high width	not high width
door	0.0000	1.0000
wall	0.5000	0.5000
opening	0.3077	0.6923

	frame	no frame
door	1.0000	0.0000
wall	0.0000	1.0000
opening	0.0000	1.0000

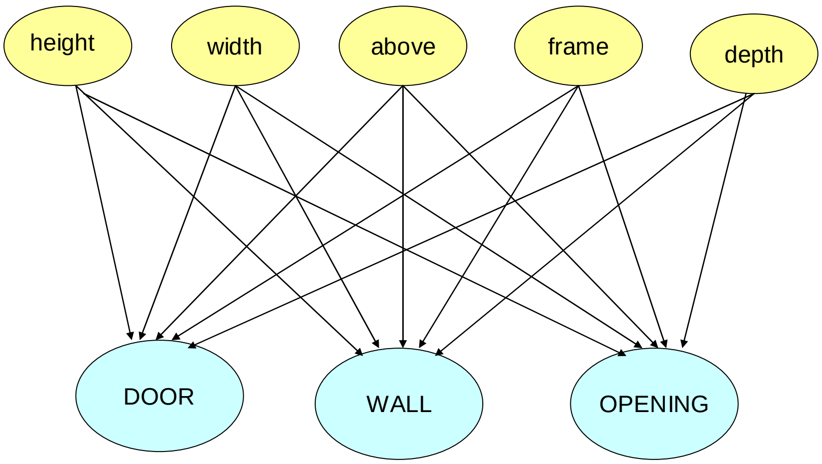
	low above	not low above
door	0.0000	1.000
wall	1.0000	0.0000
opening	0.6923	0.3077

	depth	no depth
door	0.0000	1.0000
wall	0.0000	1.0000
opening	1.0000	0.0000

Robin Burk
CS551

3.0 Constructing and Using the Belief Net

From these likelihood tables (i.e. conditional probabilities) we can now construct a belief net. We will assume the conditional independence of the attributes from one another, resulting in the following graphical structure:



Belief Net for Indoor Object Recognition

These presumed causalities, supplemented by the conditional probabilities listed in Table 4, enable reasoning about the likely identity of an object, given its attribute values.

For instance, we can compute the probability that an object with no frame (~F), a low expanse of wall above it (LA), a not-high width(~HW) and a medium height(MH) is a hallway opening, given everything else we know about indoor environments, as follows:

$$P(O \mid {\sim}F, LA, {\sim}HW, MH) = P(O) * P({\sim}F, LA, {\sim}HW, MH \mid O) \ / \ P({\sim}F, LA, {\sim}HW, MH)$$

The prior probability that an object is a hallway opening is $P(O) = 13/89 = .1461$ since we have no reason to assign priors on any other basis than their frequency in our sample set.

The likelihood of $P({\sim}F, LA, {\sim}HW, MH \mid O)$ is the product of the individual conditional probabilities $P({\sim}F|O)$, etc.:

$$P({\sim}F, LA, {\sim}HW, MH \mid O) = 1 * .6923 * .6923 * .6923 = .3318$$

And the joint probability of $P({\sim}F, LA, {\sim}HW, MH)$ can be obtained from Table 2 by marginalizing on each attribute and multiplying the resulting marginal probabilities:

$$P({\sim}F, LA, {\sim}HW, MH) = .5056 * .4607 * .8989 * .5955 = .1247$$

Therefore, the posterior probability that our observed object is a hallway opening is

$$P(O \mid \sim F, LA, \sim HW, MH) = .1461 * .3318 / .1247 = .3887$$

Note that in this case we did not have an observed value for depth. Given the sample distribution, knowing the value for the depth "beyond" the object collapses the calculation considerably. The same is true with regard to presence or absence of a frame when evaluating the probability that the object is a door. More broadly, however, this example illustrates the robustness of Bayesian reasoning in the face of incomplete data. Since we have no evidence that directly addresses the value for depth, we marginalize over it and are able to calculate the desired posterior probability that the object is a hallway opening.

4.0 Conclusions and Next Steps

This project provided an opportunity for me to investigate what attributes that might be of use for probabilistically inferring the presence of common objects in an indoor environment. A few observations about the results:

First, despite the similarity of widths, heights etc. of many members in each object class, it was useful to bin attribute values into a small number of ranges. Ideally, the nodes in a belief net have attributes with present/absent, true/false values, as this reduces computations significantly. In the case of the net constructed here, one attribute (height) has been given 3 possible values. One next step is to recast the existing data by collapsing low and medium heights into a single category and using the resulting belief net to predict synthetic instances drawn by nested sampling from the data. Another step might be to collect additional data from a much a broader range of buildings.

Second, although color information was collected, it was not integrated into the belief net. Changes in color as collected by digital cameras are a primary means by which intelligent software embedded in robots can detect object edges. However, the qualitative data collected for this effort is inadequate to address color as a factor for software object recognition. Characterizing colors in terms of camera inputs is a non-trivial (and robot-specific) process which will require extensive data collection and analysis in its own right.