Paper Review

"Towards a Cognitive System That Can Recognize Spatial Regions Based on Context"

Ayush Varshney (10183)
February 26, 2013

Synopsis

The paper describes a cognitive system that can recognize spatial regions based on the context they are presented to the system. To understand the problem take this example:
Command: Find the position for the refree in a stadium
The solution depends on the game being played in the stadium which in itself can be determined by the relative position of players, ball, etc. Such areas are referred to as Context Dependent Spatial Regions CDSRs in this paper. The cognitive system described in this paper can learn CDSRs by using semantic labels, Qualitative Spatial Representations (QSRs) and analogy.

Introduction

The paper describes, in specific the problem of making a bot find the front of the classroom based on the configuration of objects such as tables and chairs present in the room (*'the context')*. The work is based on *Dora*, a mobile cognitive robot with a pre-existing multi-layered spatial model [1]. *Dora* uses vision to recognize pre-trained 3D object models and prepares a metric map which basically a set of lines in 2D global coordinate frame.

The strengths of 8 spatial relations between that object (detected by *Dora*) and each of the objects adjacent to it (adjacency being determined using a *Voronoi Diagram [2]*) is calculated. For this, the spatial relationships (s) between adjacent objects is calculated as:

$$s = \begin{cases} \left(1 - \frac{\theta}{90}\right) * \left(1 - \frac{d}{d_{max}}\right) & \text{if} \quad \theta \leq 90^\circ \\ 0 & \text{otherwise} \end{cases}$$

Where

d: the distance between the landmark's centroid and the adjacent object's location
ɵ: the inner angle between the direction vector of the relation and the vector from the origin (the landmark's centroid) to the neighbour's location.
This helps in QSR Extraction.

Methods Used

- **Representing CDSRs**

  Anchor Points are used to represent CDSRs [3]. These are symbolic descriptions which link a conceptual entity to a perceived entity. The perceived objects are the room and the ones recognized by *Dora*. Anchor points are created from the entities perceived using Unary Functions like *YMinXFewestFn*. Example of CDSR representation:

```
(regionType CDSR9 FrontRegion)
(boundarySegment CDSR9
    (YMaxXFewestFn Room3)
    (YMinXFewestFn Room3))
(boundarySegment CDSR9
    (YMinXFewestFn Room3)
    (YMinXFewestFn Group1))
```

- **Using Analogy**

  To save the user from the pain of training for all the contexts the system uses analogy. For example if the Bot is trained where the front of a classroom after perceiving a class and objects in it, it is capable of recognizing the front – of any other class of the same structure. Structure-Mapping Engine [4] is employed to perform analogical matching in the system.

Example System Run

- **Setup**

The approach was simulated on six class rooms and two simulated studio apartments. The base and target representations were created by making *Dora* wander around two different classrooms and a map created. QSRs are generated using map and object data.

- **Evaluation [5]**

Transferred	Manually Encoded	Entire Room
$\bar{p}$=.47 σ=.37, $\bar{r}$=.46 σ=.38	$\bar{p}$=.71 σ=.30, $\bar{r}$=.67 σ=.25	$\bar{p}$=.17 σ=.11, $\bar{r}$=.98 σ=.05

Table 1: Overall Performance Compared Against Target Regions Defined by Three Users

Region	Transferred	Manually Encoded	Entire Room
Front	$\bar{p}$=.32 σ=.33, $\bar{r}$=.49 σ=.41	$\bar{p}$=.60 σ=.29, $\bar{r}$=.83 σ=.19	$\bar{p}$=.16 σ=.10, $\bar{r}$=1 σ=0
Back	$\bar{p}$=.44 σ=.37, $\bar{r}$=.56 σ=.41	$\bar{p}$=.66 σ=.25, $\bar{r}$=.84 σ=.17	$\bar{p}$=.11 σ=.06, $\bar{r}$=.99 σ=.03
Front Rows	$\bar{p}$=.76 σ=.27, $\bar{r}$=.28 σ=.21	$\bar{p}$=.83 σ=.31, $\bar{r}$=.50 σ=.11	$\bar{p}$=.22 σ=.08, $\bar{r}$=1 σ=0
Back Rows	$\bar{p}$=.72 σ=.30, $\bar{r}$=.42 σ=.26	$\bar{p}$=.80 σ=.29, $\bar{r}$=.43 σ=.26	$\bar{p}$=.19 σ=.06, $\bar{r}$=1 σ=0
Kitchen	$\bar{p}$=.60 σ=.05, $\bar{r}$=.59 σ=.34	$\bar{p}$=.78 σ=.20, $\bar{r}$=.71 σ=.13	$\bar{p}$=.16 σ=.02, $\bar{r}$=.92 σ=.13
Office	$\bar{p}$=.00 σ=.00, $\bar{r}$=.00 σ=.00	$\bar{p}$=.78 σ=.29, $\bar{r}$=.55 σ=.20	$\bar{p}$=.08 σ=.03, $\bar{r}$=.94 σ=.06
Living Room	$\bar{p}$=.40 σ=.39, $\bar{r}$=.01 σ=.01	$\bar{p}$=.63 σ=.34, $\bar{r}$=.54 σ=.13	$\bar{p}$=.35 σ=.22, $\bar{r}$=.96 σ=.06

Table 2: Performance by Region Type

These results support the hypothesis that anchor points can provide a symbolic representation on top of sensor data for context-dependent spatial regions, and, when combined with qualitative spatial relations, they facilitate learning from a single example through analogical transfer

References

[1] M.H.C. Gretton; M. Gobelbecker. "Dora, a Robot Exploiting Probabilistic Knowledge under Uncertain Sensing for Efficient Object Search"

[2] Book: Spatial Tessellations – Concepts and Applications of Voronoi Diagrams: Second Ed. Wiley ISBN 0-471-98635-6 Page: 441-446

[3] Klenk, M.; Hawes, N.; Lockwood, K.; Horn, G.S.; Kelleher, J.D.; "Using Anchor Points to Define and Transfer Spatial Regions Based on Context"

[4] Falkenhainer, B.; Forbus K.D.; Gentner, D.; "The Structure-Mapping Engine: Algorithm and Examples"

[5] Klenk, M.; Hawes, N.; Lockwood, K.; Horn, G.S.; Kelleher, J.D.; "Towards a Cognitive System That can Recognize Spatial Regions Based on Context"