Using Visual Information for Spatial Advice-giving in a Primate-inspired Autonomous Robot

Ronald Arkin, Lakshmi Velayudhan, and Michael Pettinati

School of Interactive Computing

Georgia Tech, Atlanta, GA U.S.A. 30332

Email: arkin@gatech.edu

Abstract: This paper presents an approach to apply mental rotations in the context of robot navigation in support of spatial advice-giving. It describes the motivation in terms of cognitive processing, the underlying architecture, and robotic results.

Keywords: robot navigation, mental rotations, spatial cognition

I. Introduction

Robot navigation is a nearly-solved problem as a number of techniques exist for moving a mobile robot from one location to another: SLAM (Thrun et al. 05) and behavior-based control (Arkin 98) to name a few. Some methods are biologically plausible, others not at all.

A question is what role the cognitive asset of mental rotations present in many primates (spatial transformations of abstract representations of objects from one frame of reference to another) can play in robot navigation. Mental rotations refer to the ability to manipulate 2D/3D cognitive object representations. We address how mental rotations can enhance/supplement existing robot navigation. We focus on spatial advice-giving and being able to reconcile multiple frames of reference. This advice may be given through non-verbal graphical communication (maps) (Shepard and Hurvitz 84; Aretz and Wickens 92), non-verbal communication such as gestures, and direct verbal communication.

We explore the use of visual depth information as a basis for extracting abstract representations of the world, comparing these representations with a goal state of similar format and then iteratively providing a control signal to a robot to allow it to move in a direction consistent with achieving that goal.

While we are not particularly interested in reproducing the specific methods by which primates conduct mental rotations, we believe that this system serves a useful purpose to the animal. We posit that mental rotations play a yet-to-be verified role in animal navigation, particularly with respect to the communication of information between conspecifics. The goal is to complement existing robot navigation methods drawing upon models of mental rotation to reconcile differing frames of references. This allows spatial advice to be translated into visual cues prior to a journey. Objectives include creating models by which primates cognitively manipulate spatial information; developing perceptual techniques using depth imagery; and integrating these ideas on robots.

Region segmentation or object recognition of current location and projected goal images are derived from a Kinect sensor. A bootstrap phase initializes object/key segments that correspond to the desired target in both images, after which a feedforward visualization at the aligned orientation (inspired by visual-analog mental rotation models) is performed iteratively on incoming sensor data, guiding the robot. Spatial advice is given during the bootstrap phase regarding destination/waypoints in terms of objects and/or scene features.

Results are presented on a robot using depth imagery. Iterative trajectory refinement provides a complement to existing navigational methods through the reconciliation of differing frames of reference generated from the current vantage-point, the anticipated goal, and spatial advice from the user regarding objects encountered along the way. Robot navigation based on a process inspired by mental rotations in primates has been achieved and demonstrated experimentally.

II. Related Work

We previously reported on the significance of mental rotations in biology (Arkin 2012;Arkin et al 12). Mental rotation is observed in numerous animals (mostly primates): humans (Shepard and Cooper 82;Khooshabeh and Hegarty 10), monkeys (Kohler et al 05), baboons (Hopkins et al 93), and sea lions (Mauch and Dehnhardt 97). Controversy exists regarding the underlying representations used in mental rotation: they are posited to be visual analogues (Khooshabeh and Hegarty 10) or propositional models (Pylyshyn 73), with little resolution on the matter to date.

Our research could perhaps provide the foundation for advice-giving in robotic navigation. If an agent didn't "know" its goal view a priori, something/someone external to the agent could provide a high-level goal description, relating its relative position and orientation to known objects in the scene. The agent could then use these objects with the biologically-inspired algorithm below. For example, if an agent had to fetch something from a cabinet in an office, it might have a map to the office. If the agent knew the appearance of the cabinet, upon arriving at the office it could identify it and navigate relative to the cabinet using our algorithm.

Research has shown that although children and nonhuman primates are robust navigators, they do not use maps or precise distances to navigate to a location (Menzel and Menzel 07;Lourenco and Huttenlocher 07). They instead use mental transformations to overcome changes in perspective and to make assessments about the direction they must move to reach a certain location. Research by Kosslyn et al. (2006) affords evidence for mental images in humans through neuroimaging. Hutcheson and Wedell (2012) discuss how human use qualitative, abstract representations to navigate as opposed to more precise distances or explicit directions; (Menzel et al. 02) observed a similar tactic with nonhuman primates. A Bonobo had to travel from a start location to a designated goal; it did not take a rigid trajectory but varied its path. Starting position did not affect its ability to successfully navigate, implying it possesses the ability to mentally encode the spatial layout of the area and mentally manipulate this encoding to localize and travel.

Mental rotation must be differentiated from perspective-taking spatial abilities, and visual servoing/homing techniques. Hegarty and Waller (2004) show a distinction between cases where an observer is mentally manipulating a scene/scene object to view it differently and cases where the observer is mentally viewing the scene from a different viewpoint. They state humans may use both skills, but most people have a strong preference for one or another. Arkin (2012) observes visual servoing is distinct from the navigational approach described here, as our method is more deliberative. It derives direction using abstract representations of the scene rather than working with image features directly. The three-dimensional, structured scene representations in our algorithm permit correspondence between elements of like form in different scenes (objects or object parts) as opposed to corresponding image features. In advice-giving, full object recognition and semantic labeling is required.

This evidence makes the case that primates maintain abstract mental images and manipulate these representations. The approach below has an agent rotating a "mental" visual representation of an object to inform its motion. We use the visual analog approach to mental rotations. This approach is substantiated by the time dependency of mental rotation (Takano et al. 03) and evidence that good rotators rely on spatial configuration rather than visual information, sometimes segmenting complex objects into parts acted upon individually (Khooshabeh and Hegarty 10). Analogously, our approach focuses on mentally transforming the key object instead of acting upon the entire scene itself.

Some primates exhibit another process called rotational invariance, which unlike mental rotation is time-independent. In some primates, these two processes coexist (Kohler et al. 03). It is posited that as hominids retreated from their arboreal environment, evolving an upright gait where the vertical reference plane gains importance, a dominance of mental rotation over rotational invariance arose. We hypothesize that advice-giving could offer a plausible explanation for the dominance of mental rotation as well. We strive to understand the usefulness of advice-giving in conjunction with mental rotation in navigation. Advice-giving could possibly explain this dominance and we attempt to understand its usefulness in conjunction with mental rotation in navigation.

Two primary spatial transformation strategies commonly employed by humans exist, namely spatial visualization (mental rotation) and spatial orientation (perspective-taking). Though correlated, mental rotation and perspective-taking are dissociated processes (Hegarty and Waller 04;Kozhevnikov et al 06) where the former strategy is preferred when a greater than 90 degree rotation is required (Zacks et al 99, Rigal 96). Studies provide evidence for the existence of perspective-taking capabilities in chimpanzees (Hegarty and Waller 04; Kozhevnikov et al 06). This may indicate a co-evolutionary relationship between mental rotation and perspective-taking. The advice-giving role in these spatial transformation strategies is elucidated by Keyser et al. (2000). The definition of a mutual knowledge base that demarcates the perspectives of two individuals allows disambiguation of the advice specified by one to another. Advice-giving partly plays a role in the disambiguation process by "filtering out" objects that don't fit the advice's specifications. Mental rotation may be employed to identify the object an individual is referring to, by using the information conveyed through advice. Trafton et al. (2005) implement perspective-taking on a robot, illustrating this form of disambiguation. Our navigation algorithm also draws from the concept of mutual knowledge and advice disambiguation. Rigal (1996) offers evidence that asserts the correlation between advice-giving and mental rotation. Gradual cognitive development in children that eventually allows them to accomplish their task may be derived from numerous social interactions they encounter, requiring them to perform some mental rotation to identify objects or scenes. These interactions could take the form of advice-giving.

III. Advice-Giving, Mental Rotation and Robot Navigation

Components of advice-giving in navigation (Psathas 90) are:

1) A named destination – The final goal, for instance an object.
2) Operations describing movement – E.g., Move straight, keep left
3) Operations performed in relation to reference points – E.g., Turn left at the intersection

Other characteristics of advice-giving involve sequential execution of advice and state-of-being verbs.

Our previous results are reported involving representations derived from depth segmentation and object recognition including multi-waypoint scenarios (Arkin et al 12;Arkin et al 2013;Velayudhan and Arkin 15;Pettinati and Arkin 14). The algorithm for single waypoint segmentation-based navigation based on mental rotations appears in Figure 1. Figure 2 illustrates the visual segmentation process with results summarized in Table 1.

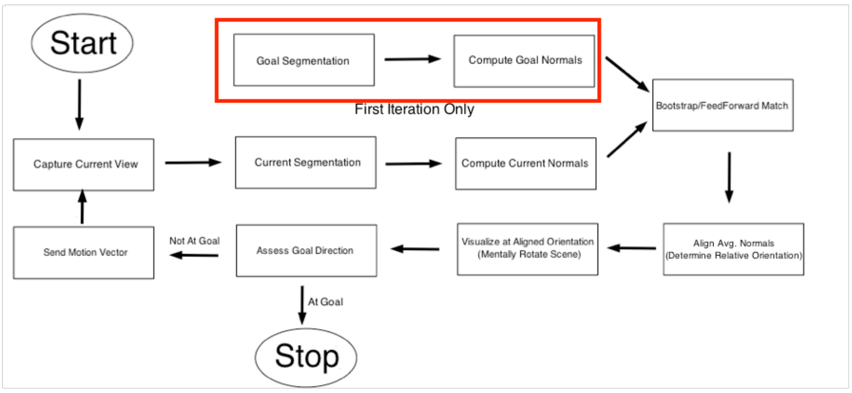

Figure 1: High-level algorithm using mental rotation for navigation

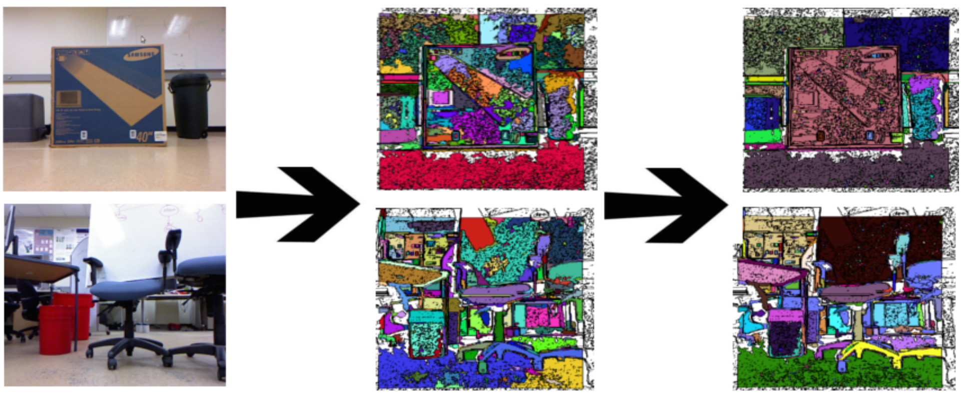

Figure 2: Visual Segmentation (scenes 1-2)

First Scene

Location	Success	Direction Updated on Average	Avg. Rotational Offset from Goal Orientation	Avg. Depth from Goal	Avg. Horizontal Displacement from Goal	Avg. Dist. from Goal
Location 1	70%	5.8	7°±4.74°	13.08cm± 8.25cm	16.49cm± 8.25cm	21.96cm± 8.25cm
Location 2	70%	7.4	1.38°±7.48°	12.21cm± 10.09cm	-6.63cm± 15.44cm	22.65cm± 4.47cm
Location 3	90%	6.3	6.8°±4.66°	8.65cm± 10.83cm	-4.18cm± 7.83cm	13.48cm± 9.44cm

Second Scene

Location	Success	Direction Updated on Average	Avg. Rotational Offset from Goal Orientation	Avg. Depth from Goal	Avg. Horizontal Displacement from Goal	Avg. Dist. from Goal
Location 1	80%	3.7	-1.5°±6.09°	4.34cm± 11.74cm	34.19cm± 10.19cm	36.6cm± 9.47cm
Location 2	90%	6.4	-1.4°±5.41°	16.89cm± 14.03cm	-30.47cm± 12.45cm	37.73cm± 11.93cm
Location 3	100%	5.3	-0.4°±4.65°	13.26cm± 10.03cm	9.7cm± 15.48cm	21.09cm± 12.86cm

Table 1: Segmentation Results

Initial work restricted this algorithm to one scene at a time. If the agent was to travel more realistic distances, it must use waypoints. The modification here is the agent stops to see if this is its final goal upon reaching a waypoint. If it is, it stops. If not, it begins again using this new subgoal location. Currently the user matches the segments at one waypoint to the "key" segments at the next (bootstrapping).

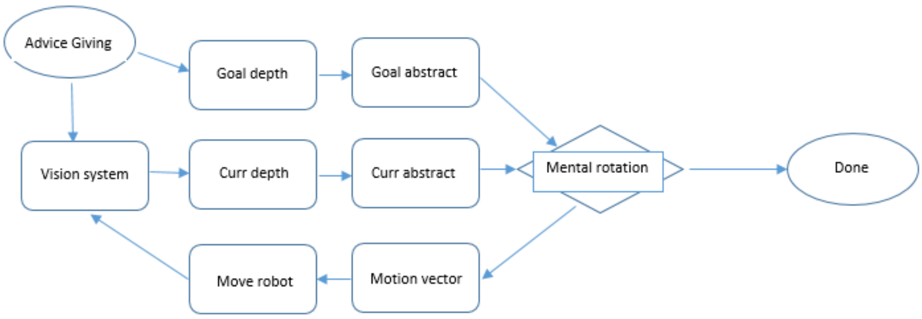

Fig. 3 – Object version with advice-giving

Our final approach (Fig. 3) directly incorporates advice-giving using object recognition. The modified algorithm is divided into two steps. The first involves calculation of the rotation matrix as a result of the mental transformations (mental rotation) on the video captures made during navigation. The second involves the computation of a bounding box that helps track the object location, improving the accuracy of object recognition by constraining the environment. Details appear in (Velayudhan and Arkin 15).

Fig 5: Scene with Multiple Objects

Location	Success %	Avg. angular offset from goal position	Avg. displacement from goal	Avg. horizontal displacement
Location 1 (avg. over 10/14 trials)	71.4 %	+ 8.3 ^{o}	+ 10.05 cm - 30.15 cm	+ 7.67 cm
Location 2 (avg. over 10/15 trials)	66.7 %	+ 2.36 ^{o} - 20.87 ^{o}	+ 2.95 cm - 28 cm	+ 9.2 cm - 5.3 cm
Location 3 (avg. over 10/15 trials)	66.7 %	+ 10.55 ^{o} - 4.9 ^{o}	- 33.79 cm	+ 11.78 cm - 20.6 cm

Table 2: Multiple objects results

This navigation algorithm gives insight into the correlation between advice-giving and mental rotation albeit subjected to certain limitations associated with the object recognition component. As shown by the results, in situations where object recognition has reasonable performance (some incorrect matches can be compensated for using bounding box approach), successful navigation is achieved for 100% of the cases. Even though higher accuracy can be obtained by improving the object recognition, existing results validate how advice-giving and mental rotation fit together in a navigational scenario, the primary goal of this project.

IV. Summary

We show how a process inspired by mental rotation, present in higher order primates, can be incorporated into autonomous navigation. The algorithm allows the robot to navigate in an informed way toward a goal location without explicit planning. This work explores possible correlation between advice-giving and mental rotation. It mimics mental rotation, using a series of transformations on observed scenes to achieve a navigational goal. The information contained within advice helps filter ambiguities in the initial capture (mutual knowledge between advisor and robot) made at the starting position (Keysar et al. 00). The results show that advice-giving could play a role in guiding navigation in scenarios where maps or external aids may be absent. Additionally, it indicates how advice-giving and mental rotation fit together in robot navigation.

Acknowledgements

This research is supported by ONR grant #00014-11-1-0593. We thank Prof. Frank Dellaert and Natesh Srinivasan for their contributions.

References

Aretz, A, Wickens, C, (1992) The Mental Rotation of Map Displays. Human Performance, 5(4):303-328

Arkin R (1998) Behavior-based Robotics. MIT Press, Cambridge MA

Arkin, R (2012) The Role of Mental Rotations in Primate-inspired Robot Navigation. Cognitive Processing, 13(1):83-87

Arkin, R, Dellaert, F, Devassy, J (2012) Primate-inspired mental rotations: Implications for robot control. Proc. ROBIO

Arkin, R, Dellaert, F, Srinivasan, N, Kerwin, R (2013), Primate-inspired Vehicle Navigation using Optic flow and Mental Rotations. Proc. SPIE DSS 2013

Hegarty, M, Waller, D (2004) A dissociation between mental rotation and perspective-taking spatial abilities. Intelligence, 32(2):175-191

Hopkins W, Fagot J, and Vauclair J (1993) Mirror-Image Matching and Mental Rotation Problem Solving by Baboons (Papio papio): Unilateral Input Enhances Performance. Journal of Experimental Psychology:General 122(1):61-72

Hutcheson, A, Wedell, D (2012) From maps to navigation: The role of cues in Finding locations in a virtual environment. Memory and cognition, 40(6):946-957

Keysar, Boaz, et al (2000) Taking perspective in conversation: The role of mutual knowledge in comprehension. Psychological Science 11.1(2000):32-38

Khooshabeh P, Hegarty M (2010) Representations of Shape during Mental Rotations. Spring Symposium, AAAI, Menlo Park.

Kohler C, Hoffman K, et al (2005) Mental Rotation and Rotational Invariance in the Rhesus Monkey (Macaca mulatta). Brain Behav. Evol 66:258-166

Kosslyn, S, Thompson, W, Ganis, G. (2006) The case for mental imagery. Oxford

Kozhevnikov, Maria, et al. (2006) Perspective taking vs. mental rotation transformations and how they predict spatial navigation performance. Applied Cognitive Psychology 20.3(2006):397-417

Lourenco, S, Huttenlocher, J. (2007) Using geometry to specify location: implications for spatial coding in children and nonhuman animals, Psycholog. Res. 71(3):252-264

Mauck B, Dehnhardt G (1997) Mental Rotations in a California Sea-Lion (Zalophus Californianus). Experimental Biology 200:1309-1316

Menzel E, Menzel, C (2007) Do Primates Plan Routes? Simple Detour Problems Reconsidered

Menzel, C, Savage-Rumbaugh, E, Menzel E (2002) Bonobo (Pan paniscus) spatial memory and communication in a 20-hectare forest. Primatology, 23(3):601-619

Psathas, G (1990) Direction-giving in Interaction. Réseaux 8.1:183-198

Pettinati, M, Arkin R (2014) A Primate-inspired Autonomous Navigation Algorithm Using Cognitive Mechanism of Mental Rotation. Proc. SIMPAR 2014:388-399

Pylyshyn Z (1973) What the mind’s eye tells the mind’s brain: a critique of mental imagery. Psychological Bulletin 80:1–24

Rigal, R (1996) Right-left orientation, mental rotation, and perspective-taking: When can children imagine what people see from their own viewpoint?. Perceptual and motor skills 83.3: 831-842

Shepard R, Hurwitz S (1984) Upward Direction, Mental Rotation, and Discrimination of Left and Right Turns in Maps. Cognition 18:161-193

Shepard R, Cooper L (1982) Mental images and their transformations. MIT Press, Cambridge, MA

Takano, Y, Okubo, M (2003). "Mental rotation.", Encyc. CogSci

Thrun S, Burgard W, Fox D (2005) Probabilistic Robotics, MIT Press, Cambridge MA

Trafton, J. Gregory et al (2005) Enabling effective human-robot interaction using perspective-taking in robots. IEEE Transactions SMC A:35.4:460-470

Velayudhan, L, Arkin, R, Primate-inspired Autonomous Navigation Using Mental Rotation and Advice-Giving. In submission 2015

Zacks, J et al. (1999) Imagined transformations of bodies: an fMRI investigation. Neuropsychologia 37.9:1029-1040