Grounding Action in Visuo-Haptic Space using Experience Networks

Arren Glover[1], Ruth Schulz[2], Gordon Wyeth[1] and Janet Wiles[2]
School of Engineering Systems
Queensland University of Technology[1]
School of Information Technology and Electrical Engineering
The University of Queensland[2]
{aj.glover, gordon.wyeth}@qut.edu.au
{ruth, wiles}@itee.uq.edu.au

Abstract

Traditional approaches to the use of machine learning algorithms do not provide a method to learn multiple tasks in one-shot on an embodied robot. It is proposed that grounding actions within the sensory space leads to the development of action-state relationships which can be re-used despite a change in task. A novel approach called an Experience Network is developed and assessed on a real-world robot required to perform three separate tasks. After grounded representations were developed in the initial task, only minimal further learning was required to perform the second and third task.

1 Introduction

The interactive, companionable, robotic worker of the future requires the ability to operate in a range of environments performing many different tasks. The skill-set required often needs to be quite specific to the task-at-hand and the environment, though it is impractical to program environment recognition and movement behaviours for all conceivable situations directly into the robot. Therefore there is a need for the robot to learn the world around it and the consequences of its actions for successful interaction with the entities encountered.

Machine learning algorithms, traditionally used for outcome prediction and data mining, have been adopted by the robotics community as an approach to solve the key problems encountered. Various manifestations of neural networks, Bayesian networks with inference, and error minimisation have been used to solve low-level control of manipulators [Lewis, 1996; Nguyen-Tuong*, et al.*, 2008], obstacle avoidance [Glasius*, et al.*, 1995; Samejima and Omori, 1999], navigation [Koenig and Simmons, 1998; Thrun, 1998] and visual recognition [Fei-Fei*, et al.*, 2007] through both supervised and unsupervised approaches. In keeping with the algorithms original design, the focus is generally on optimising action towards achieving a goal, or to estimate an unknown target function given a feedback signal [Thrun, 1996].

Although it has been shown that behaviours can be learned to solve specific problems using machine learning techniques, an adaptive robot should be able to switch problem goal or context and complete a novel task without having to completely relearn and re-optimise towards solving this new problem. Many reinforcement algorithms [Sutton and Barto, 1998], such as neural networks and Q-learning, store only the optimised outcomes, or a policy, used towards achieving a goal and learning occurs due to the update equations alone. Disregarding the storage and organisation of the data that was used to develop the policy results in development of a policy that has only been learned with the initial goal in mind; it will not bootstrap the learning of a secondary task. These types of learning paradigms do not fit with the nature of the problems encountered when developing an embodied, adaptive robot.

The term *grounding* is most often referred to in the field of artificial intelligence for linking computational symbols to the physical environment [Harnad, 1990]. However the idea of grounding can be extended to that of linking any set of internalised representations to a second set. In general terms, a robot which has learned the outcomes of its actions within the world has grounded its action representations within its sensory representations. Given a goal in sensory space, an action policy can be extracted from the grounded action representations. The 'learning' is achieved through the grounding of actions within sensory information, rather than performing a policy optimisation.

A system that combines motor actions with sensory data in order to learn actions for interacting with the environment performs sensorimotor coordination [Lungarella*, et al.*, 2003]. The principles of sensorimotor coordination have been used to learn self-motion [Bongard*, et al.*, 2006; Lungarella and Berthouze, 2003], manipulation [Metta and Fitzpatrick, 2003], and for forming environment representations [Fitzpatrick and Metta, 2003; Modayil and Kuipers, 2008; Scheier and Lambrinos, 1996]. This paper proposes that a system developing sensorimotor coordination through the principles of grounding will provide the flexible action-outcome learning required for an adaptive robot interacting with the world. Grounding actions within the sensory system instigates a 'knowledge' development rather than only modelling a single solution function. Grounding the relationship between action and sensory state gives the advantage of allowing the agent to re-use the information to complete a novel task.

This paper proposes the Experience Network (EN) as a mechanism to perform sensorimotor grounding. The EN develops representations of visuo-haptic experiences (snapshots of the sensory system at any given time) that are linked through the actions performed by the agent. The relationship between change in sensory state and performed actions can be utilised to learn a task by conceptually navigating back through past experiences in the network.

The EN is evaluated on three typical robotic challenges: learning to navigate to interact with an object, learning to navigate to a particular physical relationship with an object and performing obstacle avoidance. The EN is assessed on a real-world robot platform using a colour segmented object feature. By grounding the actions within the sensory system no further re-learning is required to perform the second and third task after the first task has been achieved.

The development of an EN is described in section 2, the application to a real robotic platform and studies are described in section 3, with results being presented in section 4. A discussion on grounding with respect to the performed studies and future work is presented in section 5 and 6 respectively.

2 Experience Networks

An EN links a set of sensory experiences through the agent's motor actions that are executed to change experiences from one to another. Sensorimotor coordination is achieved as the agent begins to ground the relationship between action and sensation and use the learned relationship to direct future action. The main components of the EN are the experiences and the action links between them.

2.1 Experiences and Actions

The nodes of the EN are the experiences, or snapshots, of the perceived state of the world at a given point in time. Each experience is a combination of visual perception and haptic sensing; what the world looks and feels like. An experience node, e_i is defined:

$$e_i = \langle p, h, n \rangle$$

where p is the perceptual sense, h is the haptic and n is the recognition count.

While sensation is stored in the nodes of the network, action is stored in the links. A link is formed between two experiences when the experience changes from one to the other. The link stores the motor or actuator commands that were performed during the change of experience and the time taken to do so. The link keeps a record of how the agent traversed from one state to the next and forms the motor half of sensorimotor coordination. A link, l_i, is defined as follows:

$$l_i = \langle m, t, n \rangle$$

where m are the motor commands, t is the time taken and n is the traversal count.

2.2 Network Construction

An EN is constructed as the agent senses and acts in the world and grounds more experience and action relationships over time. A naive approach would result in a linear chain of experiences over time; however the same experience can occur multiple times in the agent's lifetime. A recognised past-experience can be revisited forming loops of experiences rather than a single chain. The benefit of closing the loop is that when the same, or similar, experience is recognised the consequences or actions can be recalled and used to direct future action. Given two experiences e_i and e_j the, the probability, $P(e_i=e_j)$, that they are the same is calculated via prior assumption of the given state and by assuming a Gaussian distribution for each element of an experience, as follows:

$$P(e_i = e_j) = P(e_j, t \mid e_i, t-1) \prod_{g=1}^{n} \exp^{(g_i - g_j)^2 / (-v_g)} \qquad (1)$$

where t is the current time, g is an element of an experience, and v_g is the variance of element g.

A new experience is added to the network when the probability of being at a previous experience is less than the experience similarity threshold (E_t). The two experiences e_i and e_j are deemed recognised when $P(e_i=e_j)$ is above the threshold. Every time an experience is recognised and 'revisited' each element of the representative experience (e_g) can be updated to include the newly experienced sensory information.

$$e_g = e_g + (e_{g\,new} - e_g) / n \qquad (2)$$

In a predictable world the actions from a given experience often lead to the same resulting experiences. When two consecutive experiences are visited more than once there will be multiple links between these experiences. Links li and lj can be compared in a similar fashion to experiences using the probability comparison measure $P(l_i=l_j)$.

$$P(l_i = l_j) = \prod_{k=1}^{n} \exp^{(k_i - k_j)^2 / (-v_k)} \qquad (3)$$

where k is an element of a link and v_k is the variance of element k.

A new link will be formed when no current link's probability is above the link similarity threshold L_t. A link that is above the similarity threshold can consolidate each element (l_k) as follows:

$$l_k = l_k + (l_{k_{new}} - l_k) / n \qquad (4)$$

The network is constructed as the agent interacts with the world. For each new sensory experience that occurs, the probability of it being an already known experience is calculated, given all the current experience nodes in the EN. New nodes are added when no current experience is sufficiently similar. New links are created when an action leads from one experience to another which has never been performed before, regardless of when the experience was created. The result over time is a connection graph of sensory experience nodes linked together by the actions that allow them to change from one to another.

The EN grounds actions directly in the sensorimotor space of the agent as the nodes and links are derived from sensorimotor interaction with the world.

2.4 Network Navigation

A central aim for this study is to direct action using grounded sensorimotor representations. The EN grounds representations in such a way that navigation through the network provides a method for using the grounded representations to predict consequences and decide a course of action.

EN navigation initially requires two things: a current experience or position in the network and a goal experience. The current experience can be set as the most likely experience according to Equation 1. Goal experiences can be set according to any applied criteria or constraints on sensory data.

Given the current and goal experiences in the network, an action path can be found through the network to achieve the desired state. The path found will result in a chain of actions that can be performed to change the sensory state of the agent. The path is selected based on a time minimisation algorithm using the time element of each link and the probability of a link achieving a goal. An iterative algorithm is used to calculate the time to goal for each node. The update equation, for node T_i, is similar to that of a single horizon Markov decision process update.

$$T_i = \min_{a \in A}\left(\sum_{j=0}^{N} P_{ij}(a)(l_{ij_t} + T_j)\right) \tag{5}$$

for each action a in all available actions A, where T_j is the time at node j, l_{ij} is the link from node i to node j, and P_{ij} is the probability of action a leading to node j. The $P_{ij}(a)$ is calculated as:

$$P_{ij}(a) = P(j \mid a, i, L_i) = \frac{P(a \mid l_{ij})P(l_{ij})}{\sum_{k \in L} P(a \mid l_{ik})P(l_{ik})} \tag{6}$$

where L_i is the set of links from node i, $P(l_{ij})$ is calculated from the number of times a link is traversed and $P(a|l_{ij})$ is calculated using Equation 3.

The next action to perform, $a*$ is then chosen to minimise time to goal:

$$a* = \min_{a \in A}[T_i(a)] \tag{7}$$

where $T_i(a)$ is the time to goal of experience i given action a is chosen.

It is possible that the EN will not have a well grounded set of nodes and actions, due to incomplete network construction, sensory noise or random environmental disturbances. The probability that the goal will be reached from experience e_i given action $a*$ can be calculated and used to guide behaviours.

$$P(goal \mid e_i) = \sum_{j=0}^{N} P_{ij}(a*)P(goal \mid e_j) \tag{8}$$

where N is the number of links from node i.

A low probability of achieving a goal can mean that the goal is very unlikely to occur from the current state or the network is not well grounded. For example, when the goal state is unknown, or has not previously been experienced, the probability of reaching it is zero. Further exploration of the sensory space is necessary to more concretely ground representations.

3 Experimental Set-up

The EN was demonstrated in its use to ground egocentric motion in the visual and haptic senses of a real-world robot. The grounded representations were learnt and used simultaneously over three goal directed tasks. Importantly, the later tasks show the re-use of knowledge gained in earlier tasks.

3.1 Robot Platform

The experiments were conducted on a Pioneer 3-DX from MobileRobots (see Figure 1). The Pioneer is driven by a 2-wheel differential drive with wheel encoders and was equipped with a two degree of freedom 'gripper' with limited haptic sensors. The gripper was able to detect when an object is positioned within the paddles using infrared break-beams between the paddles. A Logitech Pro 9000 webcam was positioned to provide 320x240 pixel colour images facing forward, including the space between gripper paddles. A Hokuyo URG-04LX laser range finder (LRF) was used to avoid walls but provided no input to the network. All processing was performed on-board using a 2 GHz Pentium M processor.

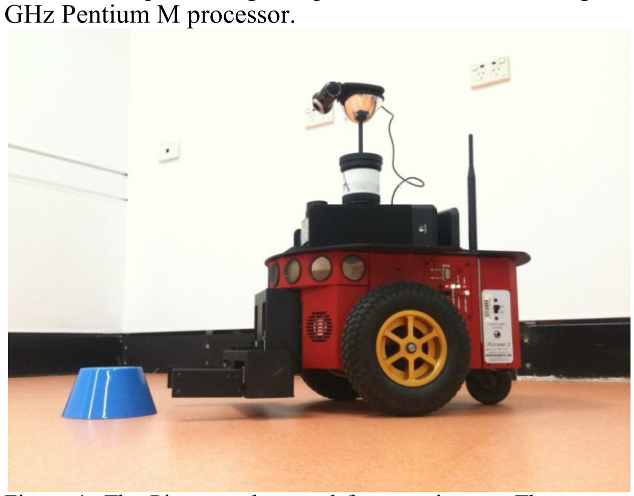


Figure 1: The Pioneer robot used for experiments. The camera was placed to give as large a field of view as possible while still viewing the contents of the gripper. The object used was a distinct blue plastic marker.

3.2 Applying the EN

3.2.1 Experience

The experiences were generated from a combination of visual and haptic sensory information. The camera image was processed to extract patches of a given colour hue as defined in Table 4. The visual half of the experience consisted of the presence, and the x and y coordinates, of the largest detected colour patch. The haptic half of the experience consisted of the value of the infrared break-beam. The sensory experience of the Pioneer is summarised in Table 1.

Table 1: Summary of EN experiences used on the Pioneer

Experience	
Feature Detection	$d_i \in \{true, false\}$
Feature Position	$p_i = \langle x_i, y_i \rangle \in \mathbb{R}$
Break-beam State	$h_i \in \{true, false\}$

The range of experiences available to the Pioneer consisted of any combination of the three elements categorised as per Equation 1 using the parameters in Table 4. The 'undetected feature' ($d_i = false$) is a unique state as it often has multiple links to it and defines all experiences outside the range of sensor limitations.

3.2.2 Link

The link is the action sequence taken between experiences. The Pioneer actions consist of the wheel motion as measured by the wheel encoders. The motion is processed and split into translational and rotational velocity. The time taken for each link is measured in number of frames assuming a constant 10 Hz update rate. The self-motion link of the Pioneer is summarised in Table 2.

Table 2: Summary of EN links used on the Pioneer

Link	
Self Translational Velocity	$T_i \in \mathbb{R}$
Self Rotational Velocity	$R_i \in \mathbb{R}$
Time	$t_i \in \mathbb{R}$

3.2.3 Pioneer Behaviour

The Pioneer selects actions according to two different behaviour patterns. The first behaviour is exploration, in which actions are performed solely to gain more experiences and sensorimotor relationships. The second behaviour is experience following, in which a path to the given goal state has been well grounded and can be followed to achieve it. Grounding still occurs while performing experience following.

The selection of behaviour is based upon the calculation of $P(goal|e_i)$ as per Equation 8 and performed until a new experience is reached, or the experience timeout $E_{timeout}$ occurs (see Table 4). The behaviour is chosen randomly at a ratio equal to $P(goal|e_i)$ where e_i is the current experience. When a well grounded path to the goal is known experience following will always be performed. When the path to the goal is less well grounded, the chance of performing exploration is increased.

Initially the robot has no grounded experiences and exploration of the sensory space will always be performed until a goal is reached. Explorative behaviour is chosen from a set of pre-defined actions shown in Table 3 and performed until a new experience is reached. The action is selected randomly based on the number of times each action has previously been performed from the given state, with less performed actions given preference.

Table 3: Standard behaviours to perform when exploring the sensory space

Action	Translational Velocity	Rotational Velocity
Forward	10mm/sec	0
Backward	-10mm/sec	0
Left	0	5deg/sec
Right	0	-5deg/sec

When experience following, the best link to follow towards a goal is chosen as per Equation 7 and the action a^* associated with the link is performed. as defined in Table 4.

Table 4: Summary of parameters used in the EN and Pioneer behaviours

Parameter	Value
Experience Positional Variance	σ_p = 2000 pixels^2
Link Translation Velocity Variance	σ_t = 750 (mm/s)^2
Link Rotational Velocity Variance	σ_r = 1.5 (deg/sec)^2
Feature Detection Hue	F_{hue} = 214
Minimum Feature Size	F_{min} = 500 pixels
Experience Prior	$P(e_i,t|e_it\text{-}1) = 0.9$
Experience Similarity Threshold	$E_t = 0.5$
Link Similarity Threshold	$L_t = 0.5$
Experience Timeout	E_{timeout} = 5 s

3.3 Overview of Studies

3.3.1 Study 1: Learning to Achieve a Goal

The first study evaluated the effectiveness of the EN to ground motion in the visual and haptic sensory space of the pioneer. The EN was evaluated on its ability to guide the agent towards a goal state of $h_i = true$, that is, moving such that the object feature entered the gripper and triggered the break-beam detector (see Figure 2).

The experiment was conducted in a flat, empty space using a blue marker as the visual feature. The object was placed randomly within the visual field and the EN directed actions based on the behaviours described in section 3.2.3. When the object exited the visual field the Pioneer paused and the object was placed randomly back within the field of view. When the Pioneer achieved the goal state, the Pioneer again paused and the object was placed in a random position. The procedure was performed until the Pioneer achieved the goal state 20 times.

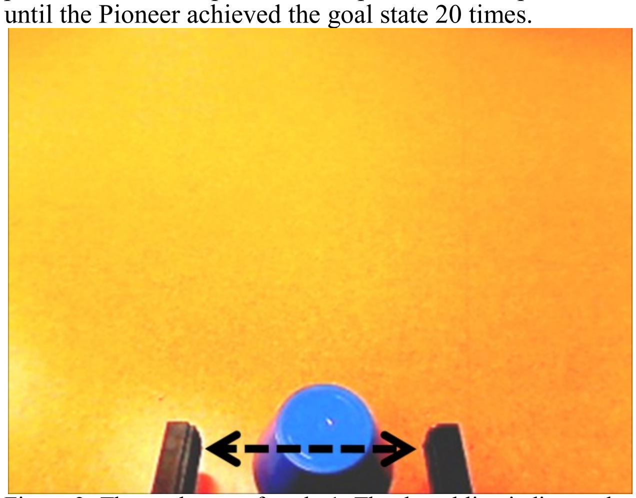


Figure 2: The goal state of study 1. The dotted line indicates the approximate location of the break-beam sensor. The blue object triggered the sensor when located between the gripper paddles.

3.3.2 Study 2: Learning a Second Goal

The second study evaluated the effectiveness of an EN which was grounded when attempting the goal in study 1 to achieve a new goal. The experience-action relationships grounded in study 1 were not explicit to the task of achieving the desired state, but rather in the predictable nature of the environment in which the agent was embodied. A robot with grounded representations should be as effective at achieving any goal state within the sensory space, not only the initial goal, once sufficient grounding has occurred.

The new goal was defined as any experience with the visual property p_i = <50±15, 80±15>. In visual space the goal corresponded to a 30 × 30 pixel box, 80 pixels down and 50 pixels across from the top left corner of the input image (see Figure 3). Using the EN grounding in study one, a further 20 runs of achieving the new goal state was performed using the same methodology as in study 1. An empty EN was also trained in achieving secondary goal only, in order to provide a baseline performance comparison.

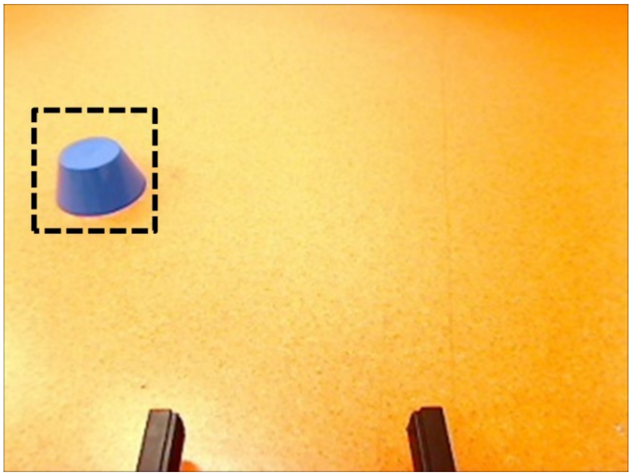


Figure 3: The goal state of study 2. The dotted square indicates the approximate area in which any experience is designated a goal. The goal state is a 30 × 30 pixel box centred over pixel (50, 80).

3.3.3 Study 3: Obstacle Avoidance

The third study evaluated the effectiveness of using the EN to perform obstacle avoidance. Avoiding obstacles involved the EN achieving the 'undetected' state (d_i = *false*). The potential number of goal experiences available to the EN increased as any experience near the 'edge' of the sensory space could lead to the undetected state. A successful avoidance would direct action to remove the obstacle from vision via the closest visual edge and not an edge on the opposite side of the image.

The Pioneer was placed in a 2.5m by 2.5m arena sectioned off by large obstacles (see Figure 4) which could be detected, and avoided, using the LRF. The object (undetectable by LRF) was placed in the centre of the arena which the Pioneer was given the goal of also avoiding.

A third 'wander' behaviour was added to the robot to perform action when not in view of the object. When in the undetected state the robot continually drove forward while turning left and right at random intervals. A command from the LRF or EN to turn also influenced the turning direction to provide smooth transition between behaviours.

The EN grounded in study 1 and study 2 was used as the pre-grounded network, again having never learnt the task prior to performing it. The EN was compared against wander-only baseline behaviour to determine effectiveness of the EN. Three trials of five minutes were performed both with and without the use of the EN.

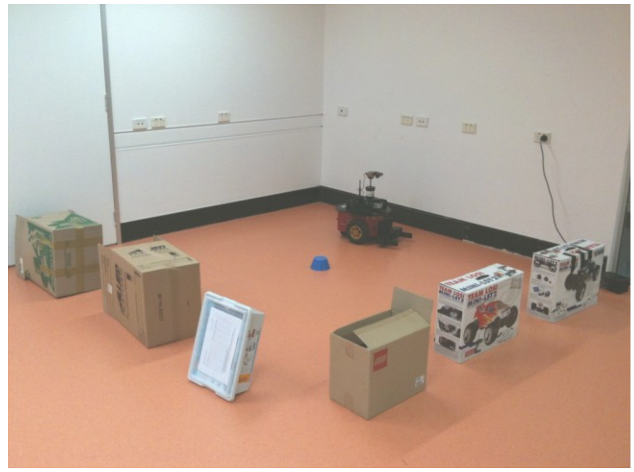


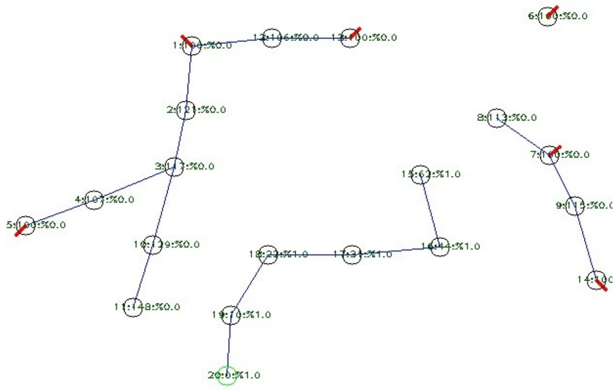
Figure 4: The arena for study 3. A 2.5m x 2.5m area was cordoned off by large obstacles that could be detected by the laser range finder. The Pioneer was set to roam inside the area while avoiding the blue object.

4 Results

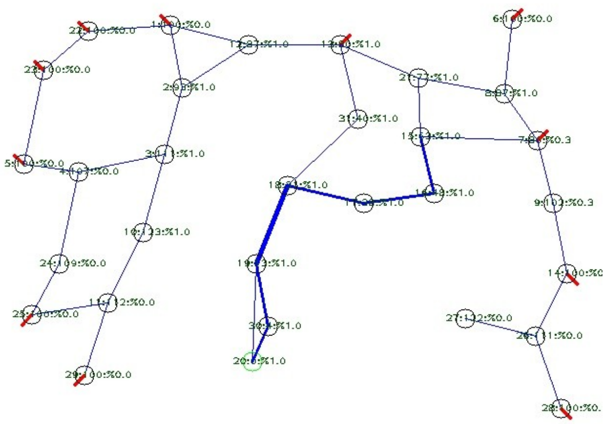
4.1 Study 1: Learning to Achieve a Goal

The initially empty EN began by performing exploration-based actions resulting in experience-action relationships developing across the sensory space. As more of the sensory space was explored the exploration behaviour eventually led to achieving the goal state. The amount of time exploring the sensory space was reduced as more experiences were available as 'entry points' to an action path to the goal, and less sensory space became unknown (see Figure 5). Subsequent runs achieved the goal state more quickly as less of the sensory space was required to be grounded (see Figure 6). To complement the reduction in exploration, the amount of time spent experience following increased. The last five indicate a well grounded EN as the amount of time required for exploration reduced to very low levels, and the majority of time was spent following grounded action-experience relationships.

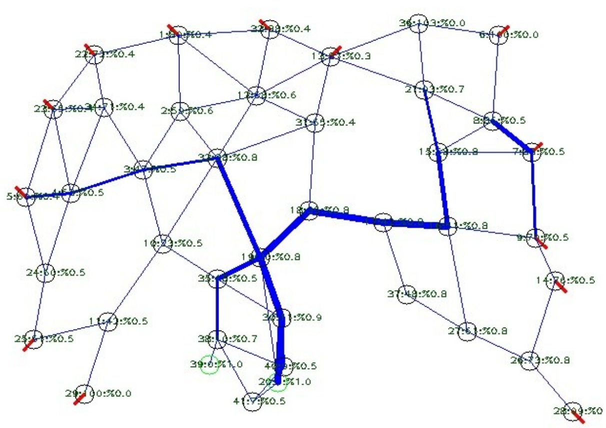
As a measure of control the EN was compared to performing only random exploration movements to achieve the desired goal. Using results averaged over three trials the random movement took approximately 147 seconds, whereas the EN, after grounding, took approximately 10 seconds.



(a)



(b)



(c)

Figure 5: The development of the EN for achieving the goal of haptically detecting the object (h_i = *true*). The ENs are displayed by placing each experience at the (x,y) coordinate of the visual feature (p_i). Links leading to an undetected experience *(d_i=false)* are displayed as short red links pointing away from the centre and are featured on the outside experiences. Each node is labelled with the node id, the time to the goal node and the probability of achieving the goal. The weight of each link signifies the traversal count n. The ENs after 1 successful run (a), after 5 successful runs (b) and after 20 successful runs (c) are shown.

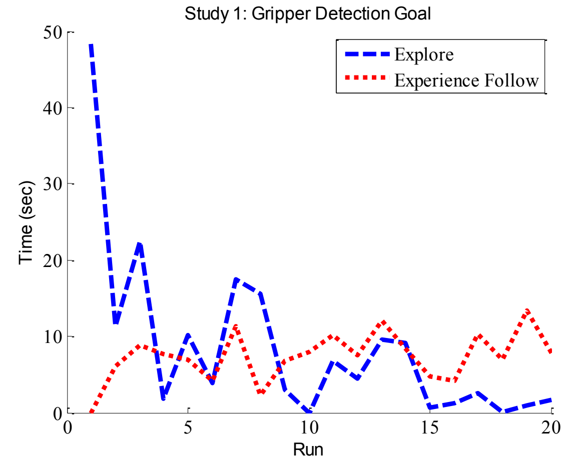


Figure 6: The time to achieve a goal state in study 1. The total time was split into the amount of time exploring and the amount of time following a grounded action path to the goal. As further runs were completed and more sensorimotor relationships were grounded the amount of time having to explore decreased and the amount of time grounded experiences were followed increased.

4.2 Study 2: Learning a Second Goal

The EN pre-grounded on the first task was able to achieve the secondary goal in similar times as it took to achieve the goal of study 1 (see Figure 7). Some further grounding was performed as the number of experiences in increased from 41 to 46 and the number of links increased from 117 to 177, and as indicated by the small amount of time exploring. The learning curve to fulfil the secondary goal was very minimal as the majority of the sensorimotor relationships required were already grounded. The results indicate the effectiveness of grounding to perform tasks within the same sensory space without having to re-learn the new goal.

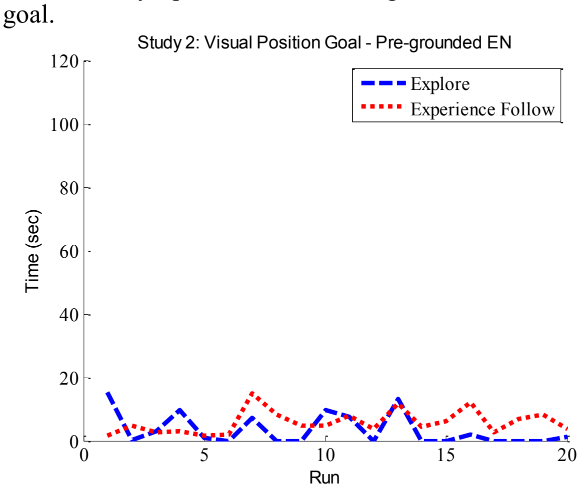


Figure 7: The amount of exploration and experience following to achieve the goal state for the pre-grounded EN in study 2. Although the network was grounded when performing study 1 there is minimal learning curve to achieve the second goal.

The newly grounded network, as expected, required time to ground enough experiences to reliably achieve the goal. The large spike on run 8 was caused as the network grounded the action of continually running into the object and pushing it along the ground, which was performed until the timeout (E_t) occurred.

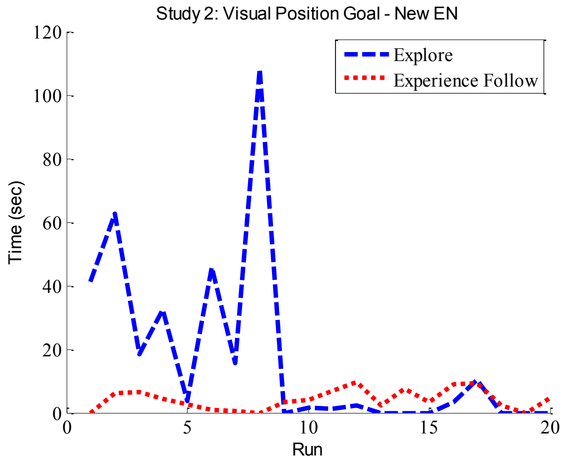


Figure 8: The amount of exploration and experience following to achieve the goal state for the new EN in study 2. The EN requires time to ground the sensorimotor relationships before stable performance is reached.

4.3 Study 3: Obstacle Avoidance

The pre-grounded EN was evaluated on the effectiveness for obstacle avoidance; the opposite of the goal for which it was originally built. Measurements of the total time that the object was visually detected (see Figure 9), and the number of collisions with the object (see Figure 10), were compared to the wander-only behaviour. Results indicate success at performing the task as the amount of time with the object in view was reduced by approximately a half. The EN was quickly able to steer away from the object and choose a new direction to wander. The pre-grounded network did not need to learn the behaviour of colliding with the object to learn what not to do, indicating the usefulness of the grounded sensorimotor relationships.

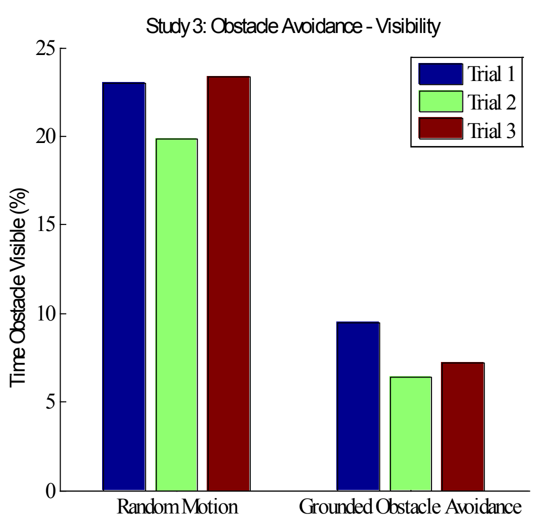


Figure 9: Results of the obstacle avoidance task in terms of amount of time the object was in view of the Pioneer. The avoidance of the EN was compared to the amount of time with random motion. Results for three separate trials are presented and it can be seen that the EN successfully reduces the amount of time the object is visually detected.

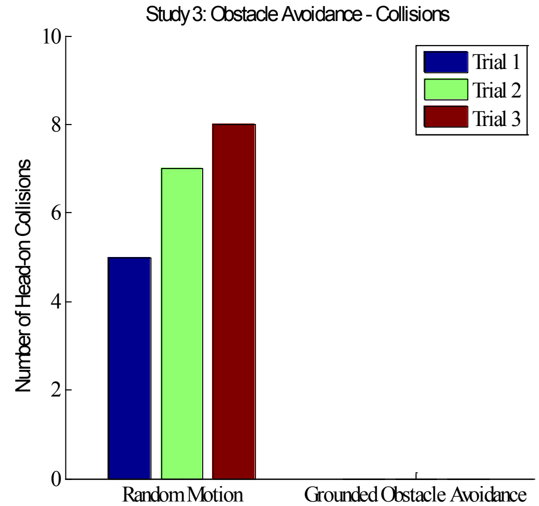


Figure 10: Results of the obstacle avoidance task in terms of the amount of times a front on collision with the visual object occurred. The EN was compared to a random wander within the arena and results for three trials of each are presented. It can be seen that the EN is successful at performing obstacle avoidance given a grounded network.

5 Discussion

An EN was developed to ground the relationship between action and change in sensory state from an egocentric perspective. The network was able to successfully learn the actions required to move an object from any location in the visual space to an experience in which it could be haptically detected. The initially empty EN demonstrated a learning curve during which the amount of time to achieve the goal decreased as more action-state relationships were grounded.

The benefit of grounding over directly supervised learning was demonstrated in the ability of the EN to perform additional tasks without having to relearn optimised actions towards the new goal states. The EN previously grounded when performing study 1 was able to move to a given position relative to an object and also perform obstacle avoidance successfully. Both tasks were able to be completed without repeating the initial learning curve experienced in study 1. Only minimal extra grounding was required, which was bootstrapped from previous grounding without notable performance decrease.

The EN was successful in the second and third task as it had grounded the action-state relationships during study 1, rather than estimating a function that would achieve the same goal. The action-state relationships are a property of the world in which the robot is embodied, not the task itself, and are therefore useful to perform a range of tasks in the sensorimotor space.

An adaptable robot able to interact with different entities in the environments while performing various tasks should be able to subsequently perform a novel task without having to relearn specific actions required to achieve it. Grounding action representations within the sensory representations builds sensorimotor relationships which can be used irrespective of the specific goal. Developing adaptive robots through the principles of grounding has the potential to lead towards robots that can easily adapt and perform novel tasks.

6 Future Work

The EN built in the presented studies provides a foundation for developing grounded representations of sensorimotor coordination. Future work entails increasing the richness of the grounded sensorimotor coordination. In particular issues such as: expanding the range of actions available to allow more complex environment interaction, expanding the range of objects with different affordances which are interacted with, removing the pre-programmed segmentation of the environment, and managing interaction with multiple objects within a cluttered scene, will be investigated. An example problem could be the tidying of a room, which requires a grounded representation of allocentric space (as the location of objects is important for determining whether they have been tidied or not); a grounded representation of different types of object which can be visually recognised; and grounded representations of how different objects should be manipulated.

As more information is used in the sensory input, the environment and the action output, issues arising through the increase in dimensionality will affect performance due to a larger state space and limited computational resources. One of the advantages of sensorimotor coordination on an embodied robot is the ability to reduce dimensionality [Pfeifer and Scheier, 1997]. A system to reduce visual input to a grounded visual-affordance relationship will allow the dimensionality of further systems to be reduced.

To perform more complex tasks the detail and resolution at which the environment can be sensed need to be enhanced. The computer vision community has investigated advanced techniques to represent a scene through identifying interesting or informative points within the image and further probabilistic methods have been developed to manipulate these representations. Employing these techniques will allow for sufficient visual sensory detail. More advanced methods for efficiently storing the complex multi-modal data in experience networks will also need to be investigated.

References

[Bongard, et al., 2006] J Bongard, V Zykov and H Lipson. Resilient machines through continuous self-modeling. Science, 314 (5802):1118, 2006

[Fei-Fei, et al., 2007] L Fei-Fei, R Fergus and P Perona. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. Computer Vision and Image Understanding, 106 (1):59-70, 2007

[Fitzpatrick and Metta, 2003] P Fitzpatrick and G Metta. Grounding vision through experimental manipulation. Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences, 361 (1811):2165, 2003

[Glasius, et al., 1995] R Glasius, A Komoda and S Gielen. Neural network dynamics for path planning and obstacle avoidance. Neural Networks, 8 (1):125-133, 1995

[Harnad, 1990] S Harnad. The symbol grounding problem. Physica D: Nonlinear Phenomena, 42 (1-3):335-346, 1990

[Koenig and Simmons, 1998] S Koenig and R Simmons. Xavier: A robot navigation architecture based on partially observable markov decision process models. Artificial Intelligence Based Mobile Robotics: Case Studies of Successful Robot Systems, 91–122, 1998

[Lewis, 1996] FL Lewis. Neural network control of robot manipulators. IEEE Expert, 11 (3):64-75, 1996

[Lungarella and Berthouze, 2003] M Lungarella and L Berthouze. Learning to bounce: First lessons from a bouncing robot. In 2nd International Symposium on Adaptive Motion in Animals and Machines, 2003

[Lungarella, et al., 2003] M Lungarella, G Metta, R Pfeifer and G Sandini. Developmental robotics: a survey. Connection Science, 15 (4):151-190, 2003

[Metta and Fitzpatrick, 2003] G Metta and P Fitzpatrick. Early integration of vision and manipulation. Adaptive Behavior, 11 (2):109-128, 2003

[Modayil and Kuipers, 2008] J Modayil and B Kuipers. The initial development of object knowledge by a learning robot. Robotics and autonomous systems, 56 (11):879-890, 2008

[Nguyen-Tuong, et al., 2008] D Nguyen-Tuong, M Seeger and J Peters. Local gaussian process regression for real time online model learning and control. Advances in Neural Information Processing Systems, 22 2008

[Pfeifer and Scheier, 1997] R Pfeifer and C Scheier. Sensory-motor coordination: The metaphor and beyond. Robotics and autonomous systems, 20 (2-4):157-178, 1997

[Samejima and Omori, 1999] K Samejima and T Omori. Adaptive internal state space construction method for reinforcement learning of a real-world agent. Neural Networks, 12 (7-8):1143-1155, 1999

[Scheier and Lambrinos, 1996] C Scheier and D Lambrinos. Categorization in a real-world agent using haptic exploration and active perception. Simulation of Adaptive Behaviour, pages 65-74, 1996

[Sutton and Barto, 1998] RS Sutton and AG Barto. Reinforcement learning: An introduction. The MIT press, 1998

[Thrun, 1996] S Thrun. Is learning the n-th thing any easier than learning the first? Advances in Neural Information Processing Systems, 640-646, 1996

[Thrun, 1998] S Thrun. Learning metric-topological maps for indoor mobile robot navigation. Artificial Intelligence, 99 (1):21-71, 1998