
Neodroid Documentation

Release 0.1.1

Christian Heider Nielsen

Nov 14, 2018

Contents

1	Project idea	3
2	Approaches, hypotheses and choice of method	5
3	Relevance relative to the call for proposals	7
4	Architecture	9
5	Abstractions	11
5.1	Actor	11
5.2	Agent	11
5.3	Observer	11
5.4	Reaction	11
5.5	Motion	12
5.6	Motor	12
5.7	Environment State	12
6	References	13
7	About	15

Creating a reality-ready robot brain in virtual reality

Project idea

The idea of Neodroid is to create a reality-ready robot brain in virtual reality. We specifically focus on creating a robot brain capable of humanoid visual-motor ability. Visual-motor ability is the integration between visual perception and motor skills. More specifically, it is the ability to perform constructive tasks integrating both visual perception and motor skills. The motivation behind Neodroid is to enable robots to assist humans in performing such tasks.

Humans have the only known visual-motor function that we truly understand through experience. Even so, we have not been able to reproduce human-level visual-motor function in a robot, despite the super-human non-visual motor skills in robots. Humans learn through a relatively slow process that unfolds over years, from the baby stage, through childhood and into adulthood. In Neodroid, we seek to develop humanoid visual-motor function in an accelerated learning environment, so that a robot can learn in days or weeks what may take a human years to learn.

The perfect accelerated learning environment is virtual reality (VR), where time can be sped up, and robots can move and learn faster and safer than in reality. The vision behind Neodroid is to teach robots how to perform visual-motor tasks, by training their brains in virtual reality. This is similar in philosophy to how Neo in the movie “The Matrix” learnt kung fu over the course of several hours within a simulator – in just a few seconds in the real world. Similarly, in Neodroid we will create a robot brain by training it – using deep learning – in virtual reality, much faster and safer than possible in the real world.

The essence of Neodroid is shown in Figure 1. The ‘baby’ phase is shown to the left in the picture, and involves growing the basic visual-motor ability of the brain in virtual reality, so that it can represent the connections between visual perception and motion. Similar to the baby phase in humans, where the baby brain transitions from a blank brain – tabula rasa – to a brain capable of representing the basic structures of reality. There is no human teacher present in the baby phase – in which environment-assisted learning takes place. Environment-assisted learning implies learning purely by observation of the environment. In Figure 1 this is illustrated by the virtual robot just observing images of the falling strawberries together with environment-provided information on the underlying 3D position and pose of each strawberry. The ‘school’ phase follows, in which a human enters virtual reality and demonstrates the visual-motor associations and motion patterns that define different visual-motor tasks. In the school phase, the brain structure develops further and is capable of representing complex spatio-temporal sequences of visual-motor response patterns. After the school phase, the brain is transferred from the virtual to the real robot, so learnt visual-motor skills can be applied in the real world.

Approaches, hypotheses and choice of method

The research problem in Neodroid is essentially an improved way for humans to teach robots and for robots to learn. The motivation behind this is the often tedious development cycle involved in robotics applications – such as illustrated in Figure 3. Here we have a farmer that wants a strawberry-picking robot. The farmer tells the engineer about the problem, and then the engineer tells the robot – through programming – what the problem is. Some things are usually lost in translation, and this process repeats until the farmer is satisfied. In Neodroid, we are developing technology that can enable a more direct transfer of knowledge from the human to the robot – via virtual reality, as seen in Figure 4.

We are focusing on a narrow set of tasks, involving visual-motor ability. This area of research is an area which enables robot to see and respond with a motion sequence. Existing knowledge in this area is mature when it comes to simple tasks that can be programmatically explained – such as putting a peg in a hole. For simple task sequences, there are robots that can be quite directly taught by a human physically moving the robot [12]. When the tasks become complex, and/or difficult to describe in a program, there is a need for expanded knowledge and this is where we focus – specifically with virtual reality as the medium in which to teach robots. With that in mind, we present the following hypotheses to be investigated in Neodroid:

In terms of methodology, we have chosen deep learning – and focus on the challenge of developing the applicable deep learning architectures. We chose deep learning since it has the necessary tools to handle both image processing (using convolutional neural nets [5,6]) and motion sequences (using recurrent neural networks [13]). This is well suited to our task, in which we want a robot to provide a motion response to a visual stimulus. Virtual reality is the primary medium we will use to provide the example visual stimulus and end-effector motion responses. Such motion responses are essentially the 6 DOF motion of two grippers, without consideration to the pose of the other robot joints or the position in space of the robot. A humanoid robot – Aldebaran Pepper – will be used to demonstrate transfer-learning from virtual reality to the real world. Inverse-kinematics algorithms are needed in order to convert end-effector motion responses to actual motion of the humanoid robot. NTNU and SINTEF already developed such algorithms [14] applicable to a moving humanoid robot.

Given the methodology above, we will investigate our hypotheses through the following objectives: Hypotheses 1) Robots can learn visual-motor ability in virtual reality. 2) Humans can teach robots visual-motor tasks in virtual reality. 3) Visual-motor skills, learnt by a robot in virtual reality, can be transferred to a robot in the real world. Primary objective Develop deep learning architectures for creating a reality-ready robot brain in virtual reality. Secondary objectives Implement a virtual reality environment for training a virtual robot. Develop deep learning architectures for visual-motor tasks. Implement human operation of a robot in virtual reality. Demonstrate robot deep learning in virtual reality. Demonstrate transfer-learning between virtual reality and the real world.

Header 1	Header 2	Header 3
body row 1	column 2	column 3
body row 2	Cells may span columns.	
body row 3	Cells may span rows.	<ul style="list-style-type: none">• Cells• contain• blocks.
body row 4		

Relevance relative to the call for proposals

Neodroid is relevant to the following aspects that FRINATEK scheme promotes: Scientific quality at the forefront of international research Deep learning applied to robot visual-motor tasks is a very nascent research field, with 2015 being the first year in which this has been explored in the simplest of grasping tasks, by Cornell University [1] and later in 2016 at Google [2]. Independently, and at around the same time, we began our early work [15] on this topic – to be published later this year (2016). Hence in this regard, we are in the international forefront already. Beyond this, we have a deep learning paradigm that differs both from that of Cornell and Google. Our approach has greater potential and crosssynergies with the developments within virtual reality (VR), coinciding with the introduction of the first consumer-quality VR technology in 2016. The combination of deep learning and virtual reality is a key novelty of our project and we are the first to apply this to developing robot brains with visual-motor ability.

Boldness in scientific thinking and innovation

The potential for Neodroid is to open doors to a whole new area of research with wide impact for the future. When virtual reality becomes commonplace and of a high quality in a few years, it will enable humans to teach virtual workers and robots what to do – all in virtual reality – and the brains of these virtual workers or robots can be transferred to physical robots anywhere – on- or off-planet. This can revolutionize the way humans teach robots and the way robots assist humans. As a stepping stone, this has the potential to enable greater endeavours for the human race, such as remote construction of factories and resource mining in space, affordable and automated construction of homes on Earth, as well as automated farming and processing of food, personal assistants to humans and manufacturing of robots themselves. Taken further, the potential is the duplication of the visual-motor skills of a human expert anywhere on Earth into many robots anywhere in the Solar System – hence solving the challenge of expensive human labour costs and worker availability currently inhibiting the progress of human evolution. This is the bold potential, 4

and we believe in this light that the Neodroid project is a worthy and realistic first step towards this exciting potential.

Careers for young research talents

Through his early career as research scientist, Dr. John Reidar Mathiassen has had practical implementation and management experience in introducing machine vision and robotics to use in the Norwegian fisheries and aquaculture industries. He is also a visionary and sees the immense potential for future scientific innovation in robotics applications. The combination of practical experience and being a visionary has led him to look for and discover some critical next steps in deep learning, virtual reality and humanoid robotics – steps requiring focus now in order to lay the foundation for the future. Neodroid is a stepping stone for Dr. Mathiassen, and will enable him to develop his career so that he can start building up a research group to develop an exciting new (yet to be named) research field spanning deep learning, virtual reality and robotics.

CHAPTER 4

Architecture

Neodroid is comprised of a collection models designed to be general and simple, to facilitate the user to easily build the mental model of the environment they building. These models include:

5.1 Actor

The concept of the *actor* model is that is that is encapsulates a number of *motor* s as part of one entity. For example as brushless electric motors on a multirotor aircraft, here the actor is the aircraft and the motors in are the brushless electric motors.

5.2 Agent

The *agent* model is an abstraction of some external agent providing *reaction* s to be acted out in the environment. The *agent* abstraction has the responsibility of providing the interface for external agents into the environment, this include passing messages back and forth through a *TCP* connection or likewise.

5.3 Observer

An observer or its 'sensor' synonym fits quite well for thinking about the abstraction at hand. It is responsible for capturing information in the environment to be passed on the agent. It can for example be used for capturing depth images in the environment or tracking a specific variable in the environment.

5.4 Reaction

A reaction is the abstraction of a message passed from an external agent to the *agent* model providing the interface, as a response to the observed state. A reaction includes a boolean variable for resetting the environment and a number of the *motion* s to be acted out the environment by actor and their motors.

5.5 Motion

Motions are abstractions of spendature of energy and where the energy should be spent. These is simply an address on an *actor* and a *motor* with an assigned energy strength, this energy may be positive or negative to indicate for example expansion and contraction of a *motor* or the torque applied to a *motor* clock-wise or anti-clock-wise. Its is the absolute magnitude of the strenght that counts towards total energy spendature. Motions are passed through *reaction* model as messages produced by the agent, this the way the external agent affect the environment.

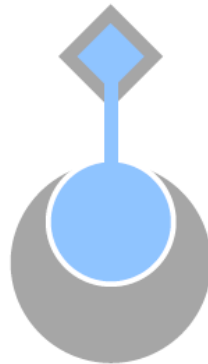
5.6 Motor

The idea of the motor is very general, is can be unary or binary, here binary refers to the sign of the motion/force being applied to it, if a *motor* is unary it can only be affected by a positive force. A motor can apply anything from a spinning motion to a sliding motion, it is easy the realise that both of these can be binary but also unary if chosen. Whereas a rocket motor provide a thrust motion will most likely only be unary.

5.7 Environment State

The *environment_state* encapsulates all information that should be exposed to external agents into one message. This message is comprised of *observer* s with their relevant observational data, *actor* s with their positions and rotations in the environments, how much energy has been spent since the last reset, how many frames/time has passed since last reset and lastly a reward given by some objective function.

and a additional one which is relevant for building the interface for an external agent with only partially observervable state of the environment.



Neodroid

Creating a reality-ready robot brain in virtual reality



CHAPTER 7

About
