



Bio-inspiration for Deployed Autonomous Systems

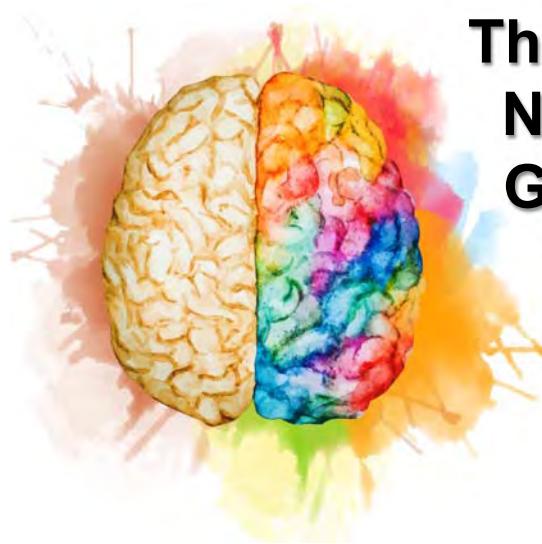
The Good, the Bad, and the Unknowns

Professor Michael Milford | Australian Research Council Future Fellow | Microsoft Research Faculty Fellow | Chief Investigator, Australian Centre for Robotic Vision
michael.milford@qut.edu.au

Overview



Introductions



**The Mystery of
Navigational
Grid Cells in
the Brain**

**Biologically-inspired Mapping
and Navigation & Sensing**



**Translating Research to Industry
Autonomous Vehicles**





michael.milford@qut.edu.au



Twitter: @maththrills



<https://www.youtube.com/milfordrobotics>



<http://www.tinyurl.com/milfordm>



<https://goo.gl/rczsle>



**Today's talk will only cover
a small fraction of our
research and activities**

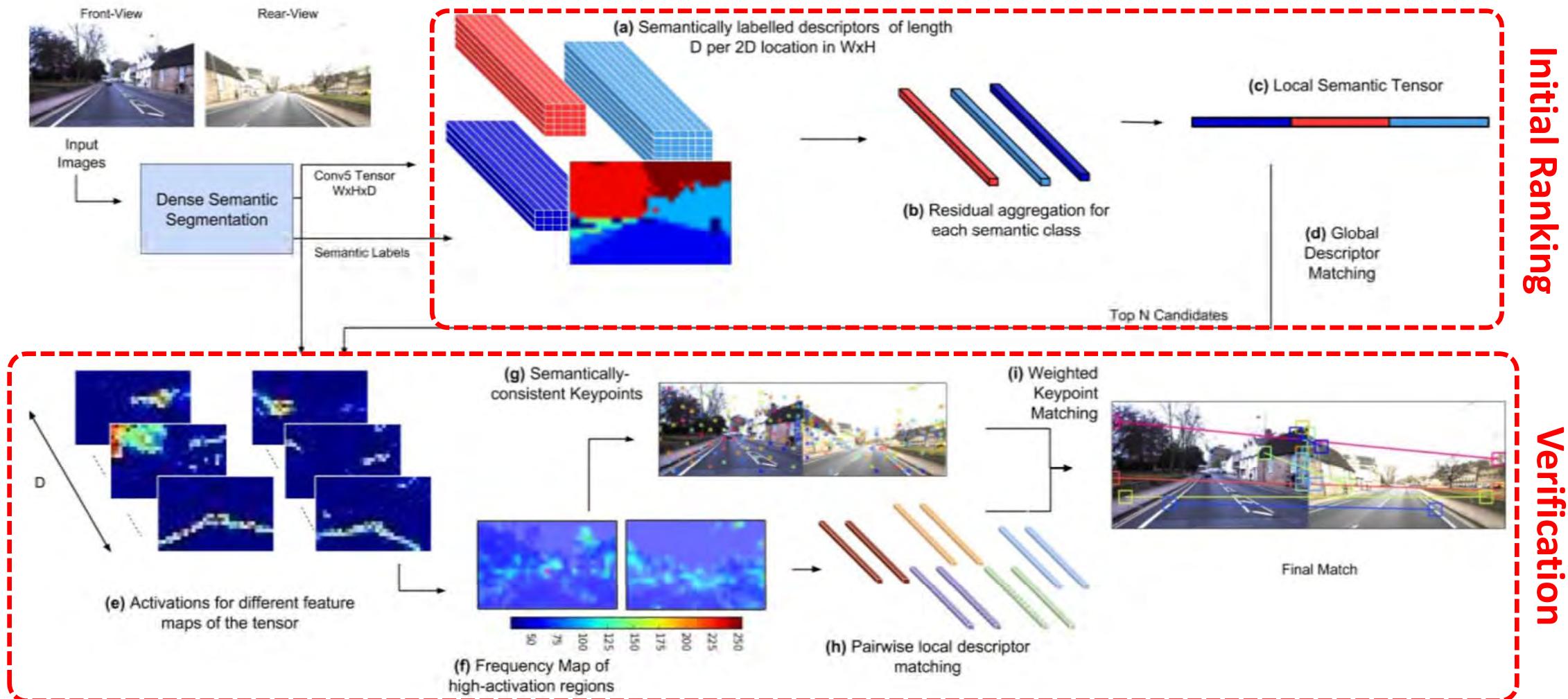
Please reach out to chat:

autonomous vehicles | neuroscience-inspired tech | robotics | computer vision | active navigation | mapping and localization | **machine and deep learning inc. reinforcement learning** | artificial intelligence | startups | STEM (Science Technology Engineering Maths) education | movie & fiction-based edutainment | collaboration

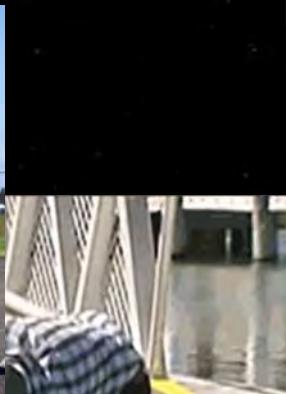
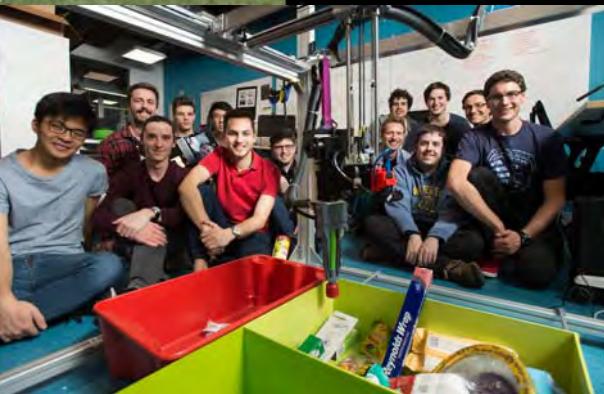
Professor Michael Milford | Australian Research Council Future Fellow | Microsoft Research Faculty Fellow | Chief Investigator, Australian Centre for Robotic Vision

michael.milford@qut.edu.au

LOST? Appearance-Invariant Place Recognition for Opposite Viewpoints using Visual Semantics



Robotics and AI at QUT



Extensive Outreach Engagement Consulting in AVs



Where to go for more information... (high level)

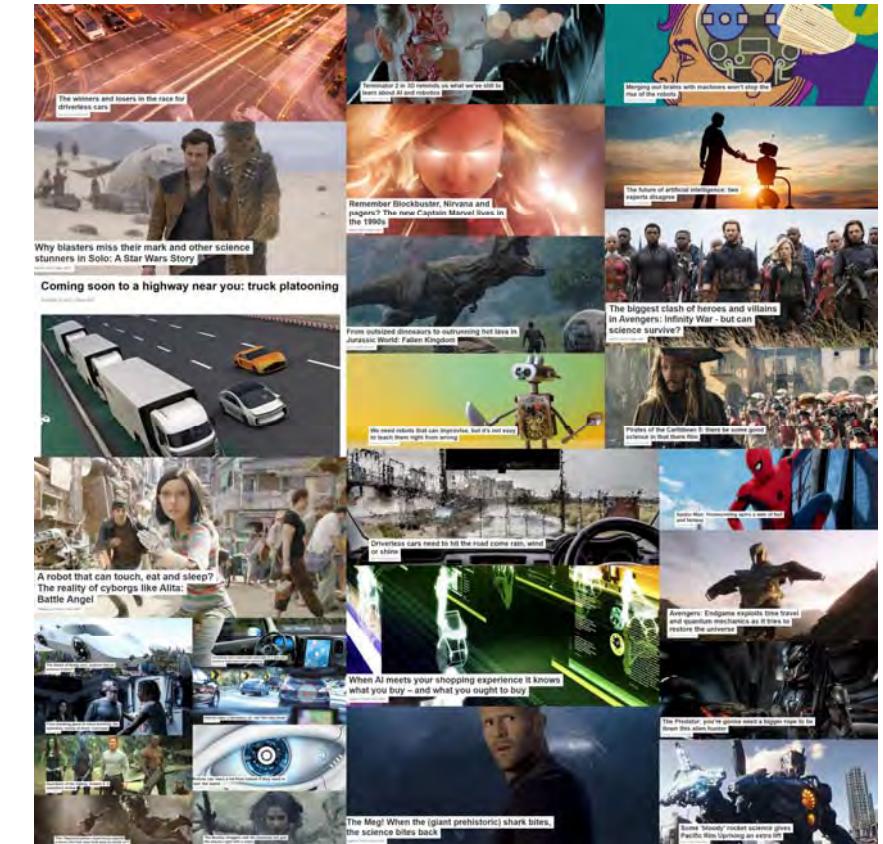
IEEE Spectrum June 2017



Username:
MilfordRobotics

Engineers Australia Create
Magazine Nov 2018

The Conversation and
other media outlets



Publications and Key Survey/Review Papers

Google Scholar: <http://scholar.google.com/citations?user=TDSmCKgAAAAJ>

IEEE TRANSACTIONS IN ROBOTICS VOL. 32, NO. 1 FEBRUARY 2016

Visual Place Recognition: A Survey

Stephanie Lowry, Niko Sünderhauf, Paul Newman, Fellow, IEEE, John J. Leonard, Fellow, IEEE, David Cox, Peter Corke, Fellow, IEEE, and Michael J. Milford, Member, IEEE*

Abstract—Visual place recognition is a challenging problem due to the vast range of ways in which the appearance of real-world places can vary. In recent years, improvements in visual sensing capabilities, an ever-increasing focus on long-term mobile robot autonomy, and the availability of state-of-the-art deep learning techniques have all contributed to significant advances in visual place recognition systems. This paper presents a survey of the visual place recognition research landscape. We start by introducing the concept behind place recognition and the field of place recognition in the animal kingdom, how a “place” is defined in a robotics context, and the major components of a place recognition system. We then discuss the challenges of visual place recognition. Appearance can be a significant factor in visual place recognition. Therefore, we discuss how place recognition systems can implicitly or explicitly account for appearance changes within the visual place recognition, in particular with respect to rapid advances being made in the related fields of deep learning, semantic scene understanding, and video description.

*Index Terms—Visual place recognition, place recognition.

INTRODUCTION

VISUAL place recognition is a well-defined but extremely challenging problem to solve in the general sense; given an image of a place, can a human, animal, or robot decide whether or not this image is of a place it has already seen? Whether referring to humans, animals, computers, or robots, there are some fundamental things a place-recognition system must have:

Manuscript received March 17, 2015; revised September 18, 2015; accepted October 23, 2015. Date of publication November 26, 2015; date of current version February 5, 2016. This paper was recommended by the publications by Australian Centre for Robotic Vision. This work was supported by an ARC Future Fellowships grant (FT140100228) and an ARC Research Project grant (DP140101228). This work was supported by an ARC Future Fellowship (FT140100228) and an ARC Research Project grant (DP140101228).

*Corresponding author. S. Lowry will the Australian Centre for Robotic Vision, School of Electrical Engineering and Computer Science, Queensland University of Technology, 222 Queen Street, St. Lucia, Qld 4000, Australia; e-mail: stephanie.lowry@qut.edu.au. N. Sünderhauf, P. Corke, and M. J. Milford are with the Australian Centre for Robotic Vision, School of Electrical Engineering and Computer Science, Queensland University of Technology, Brisbane, QLD 4000, Australia; e-mail: niko.sunderhauf@qut.edu.au, peter.corke@qut.edu.au, michael.milford@qut.edu.au.

†P. Newman is with the Mobile Robotics Group, Department of Engineering Science, University of Oxford, Oxford OX1 3EL, U.K.; e-mail: paul.newman@eng.ox.ac.uk.

‡J. J. Leonard is with the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: jleonard@csail.mit.edu).

†Corresponding author. D. Cox is with the Department of Mathematics, Colgate University, Hamilton, NY 14846 USA (e-mail: dcox@colgate.edu).

© 2016 IEEE. Personal use of material is permitted. No republication without explicit written permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

Digital Object Identifier 10.1109/TRO.2015.2500823



Fig. 1—Visual place recognition systems must be able to (a) successfully match two images of different places, while also recognizing incorrect matches between pairs of different places.

and much else. First, a place recognition system must have an internal representation—a map—of the environment to compare to the incoming visual data. Second, the place recognition system must remember where it has been or not the current visual input corresponds to a place already seen before, and if so, at which time. Performing visual place recognition can be difficult due to a range of challenges: the appearance of a place can change drastically (see Fig. 1), multiple places in an environment may look very similar, a problem known as perceptual aliasing, and places may not always be revisited from the same viewpoint and position as before.

In robotics, the need for place recognition is relevant given the requirements for long-term autonomy and rapid improvements in visual sensing capabilities and cost. Vision is the primary sensor for many localization and place recognition algorithms [1]–[19]. Place recognition is also a growing research field, as evidenced by citation analyses and a number of dedicated place recognition workshops at recent and upcoming robotics and computer vision conferences including the IEEE International Conference on Robotics and Automation (2014), the International Conference on Computer Vision and Pattern Recognition (2015). The problem of place-based place recognition has also formed a regular component of many other general workshops including the long-running ICRA Workshop on Long-Term Autonomy (2011–2014).

Our aim in writing this survey article is to provide a comprehensive review of the current state of place recognition research that is relevant both in robotics and other fields of research including computer vision and neuroscience. The focus for such surveys is generally to highlight the state-of-the-art and the research future; for example, the almost universal usage of deep learning techniques in state-of-the-art recognition systems in computer vision, and the 2014 Nobel Prize in Physiology or Medicine award to Edvard Moser, May-Britt Moser, and John O’Keefe, who discovered the key representations of place in

The limits and potentials of deep learning for robotics

Niko Sünderhauf¹, Oliver Brock², Walter Scheirer³, Raia Hadsell⁴, Dieter Fox⁵, Jürgen Leitner⁶, Ben Upcroft⁷, Pieter Abbeel⁸, Wolfram Burgard⁹, Michael Milford¹⁰ and Peter Corke¹¹

¹Australian Centre for Robotic Vision, Queensland University of Technology (QUT), Brisbane, Australia
²Robotics and Biology Laboratory, Technische Universität Berlin, Berlin, Germany
³Department of Computer Science and Engineering, University of Notre Dame, IN, USA
⁴University of California, Berkeley, CA, USA
⁵Pi Lab, University of Washington, WA, USA
⁶University of Oxford, Oxford, UK
⁷UCL Brain and Behaviour Department of Electrical Engineering and Computer Sciences, CA, USA
⁸Department of Computer Science, University of Freiburg, Germany
⁹Robotics Institute, University of Queensland, St. Lucia, QLD 4072, Australia
¹⁰Corresponding author: Niko Sünderhauf, Queensland University of Technology (QUT), 2 George Street, Brisbane 4000 QLD, Australia
¹¹E-mail: peter.corke@qut.edu.au

Abstract—The application of deep learning in robotics leads to very specific problems and research questions that are typically not addressed by the computer vision and machine learning communities. In this paper we discuss a number of robotics-specific learning, reasoning, and embodiment challenges for deep learning. We explain the need for better evaluation metrics, highlight the importance and unique challenges for deep robotic learning in simulation, and explore the spectrum between purely data-driven and model-driven approaches. We hope this paper provides a motivating overview of important research directions to overcome the current limitations, and helps to fulfill the promising potentials of deep learning in robotics.

Keywords
Robotics, deep learning, machine learning, IJRR, Vision

1. Introduction

A robot is an inherently active agent that interacts with the real world, and often operates in uncontrolled or detrimental conditions. Robots have to perceive, decide, plan, and execute actions, all based on incomplete and uncertain knowledge. Mistakes can lead to potentially catastrophic results that will not only endanger the safety of the robot’s mission, but can even put human lives at risk, e.g., if the robot fails to detect a child in its path.

The application of deep learning in robotics therefore motivates research questions that differ from those typically addressed in computer vision. How much trust can we put in the predictions of a deep learning system when misclassifications can have catastrophic consequences? How can we estimate the uncertainty in a deep network’s prediction and how can we fix the prediction with prior knowledge and other sources of probabilistic information? How well does deep learning perform in realistic unstructured operating scenarios where objects of unknown class and appearance are regularly encountered?

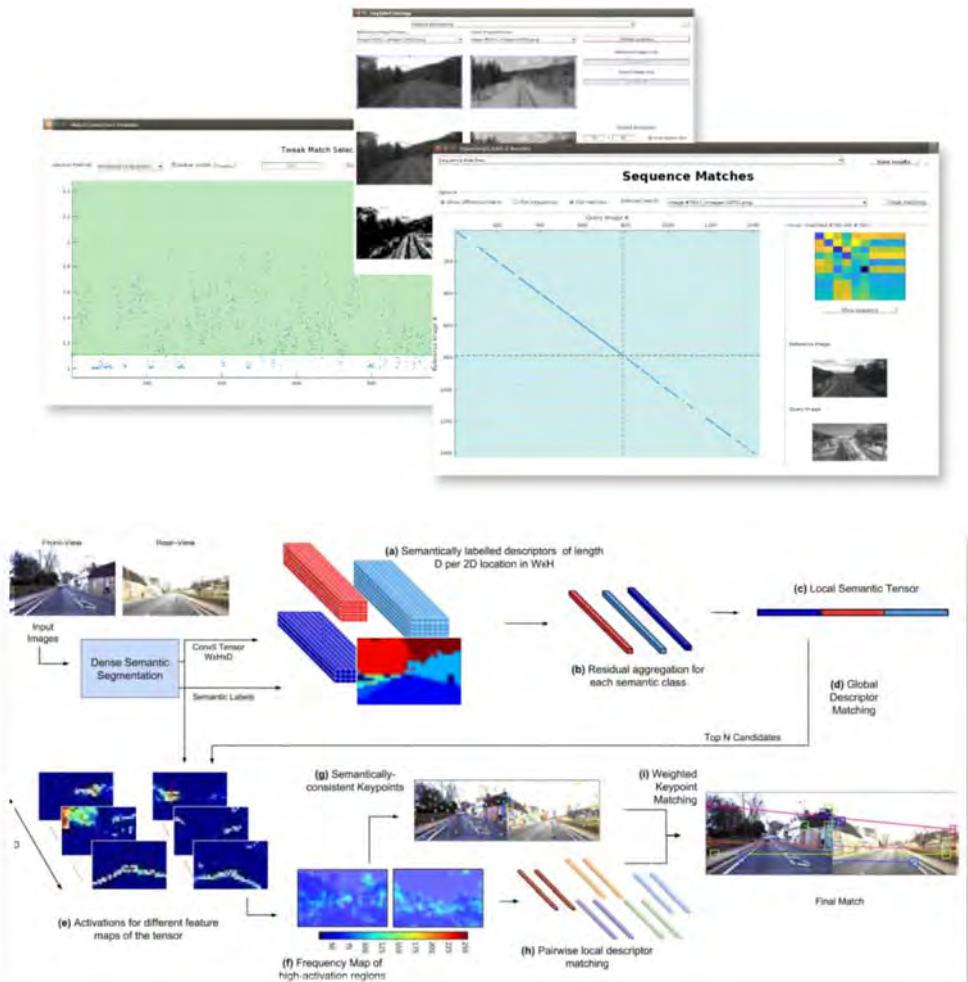
If we want to use data-driven learning approaches to generate motor commands for robots to move and act in the world, we are faced with additional challenging questions. How can we use data from high-quality training data to do we rely on data solely collected in relatively real-world scenarios, or do we require data augmentation through simulation? How can we ensure the learned policies transfer well

Stephanie Lowry, Niko Sünderhauf, Paul Newman, John J. Leonard, David Cox, Peter Corke, and Michael J. Milford,
“Visual Place Recognition: A Survey”, in *IEEE Transactions on Robotics and Automation*, 32 (1), 2016

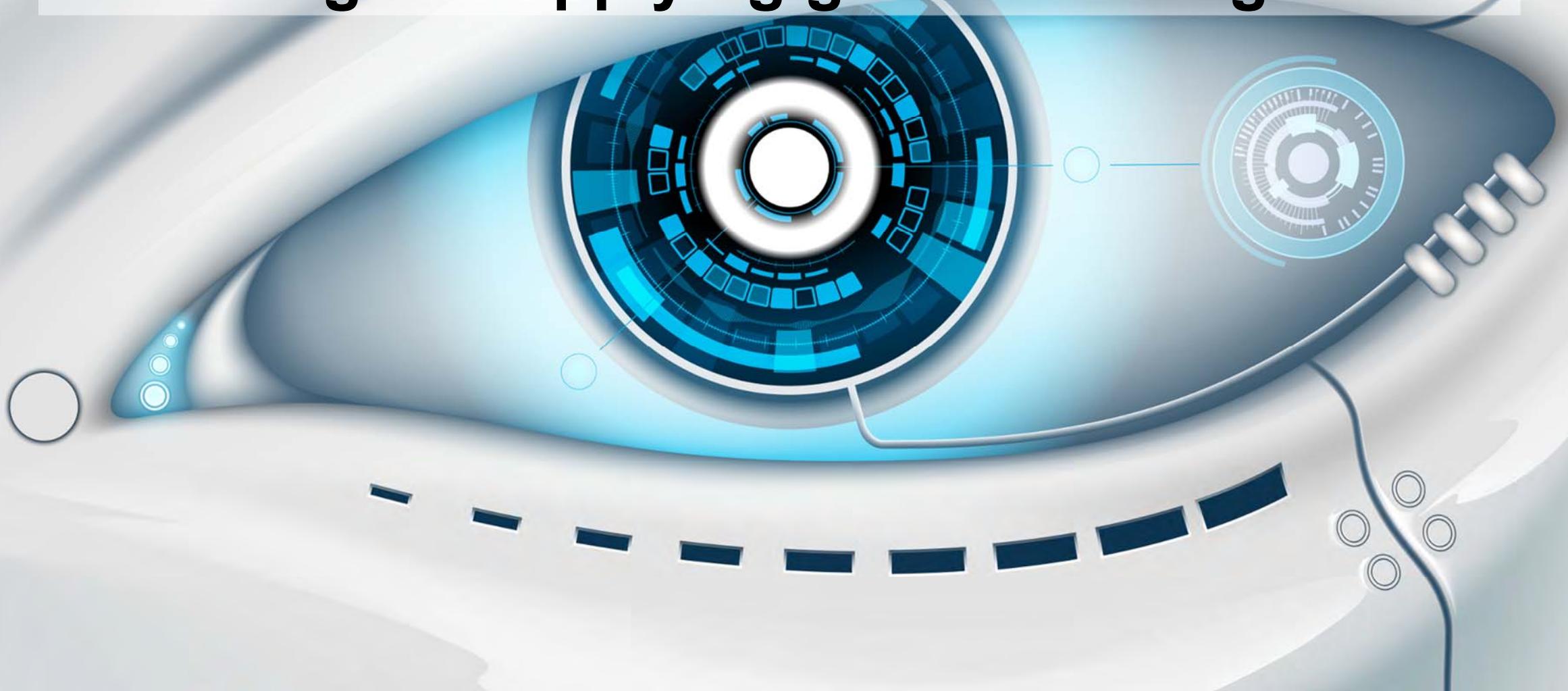
Niko Sünderhauf, Oliver Brock, Walter Scheirer, Raia Hadsell, Dieter Fox, Jürgen Leitner, Ben Upcroft, Pieter Abbeel, Wolfram Burgard, Michael Milford and Peter Corke, **“The limits and potentials of deep learning for robotics”,** in *International Journal of Robotics Research*, 37 (4-5), 2018

Open Source Code and Datasets

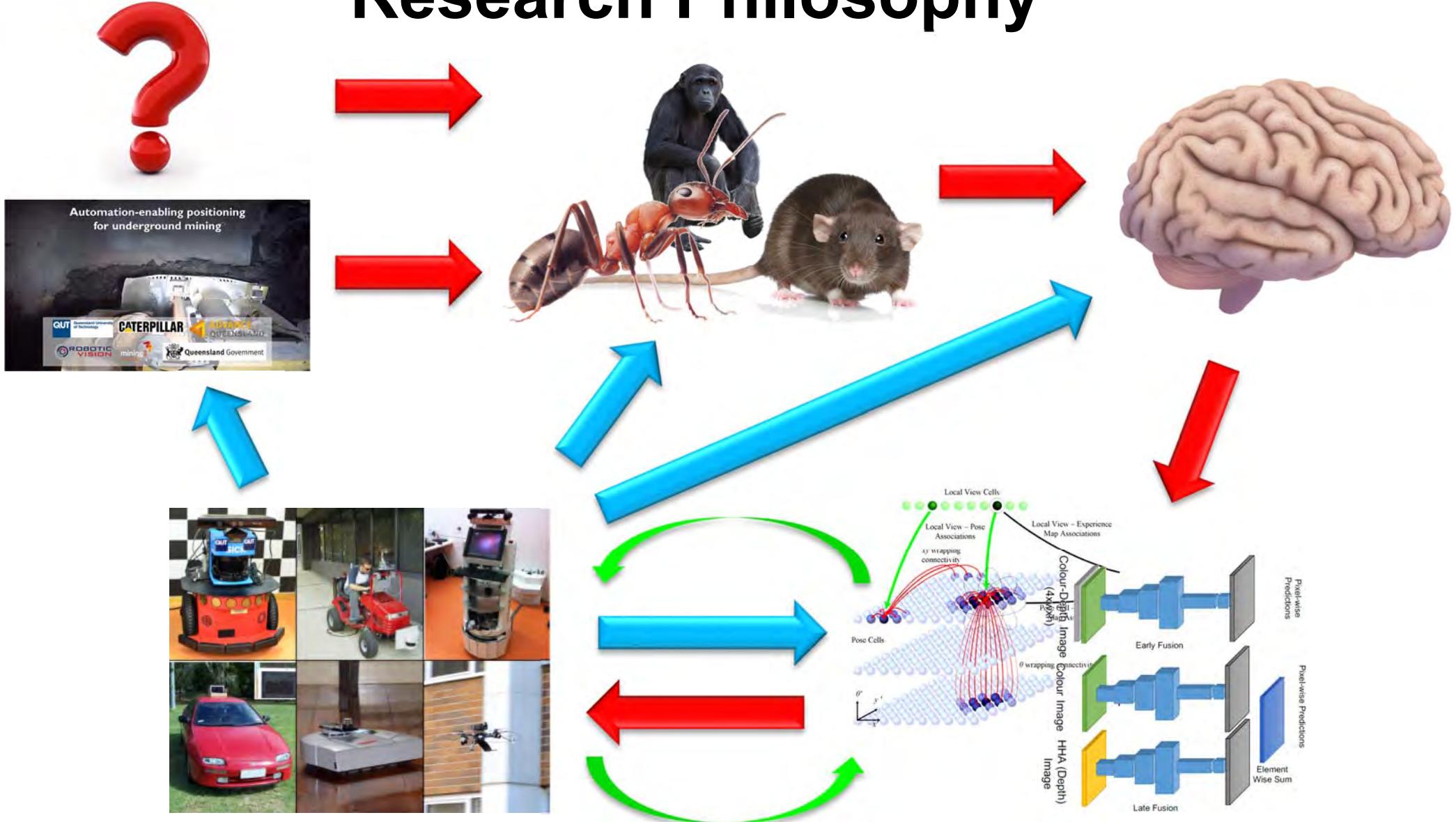
- **OpenSeqSLAM 2.0:** <http://seqslam.com/>
- **OpenSeqSLAM:**
<https://openslam.org/openseqslam.html>
- **OpenRatSLAM:**
<https://code.google.com/p/ratslam/wiki/RatSLAMROS>
- **OpenFABMAP** (also in OpenCV):
<https://github.com/arrenglover/openfabmap>
- **Learning to Navigate at Scale:** rl-navigation.github.io/deployable
- **Local Semantic Tensors:**
<https://github.com/oravus/lostX>
- **Multi-Process Fusion:**
<https://github.com/StephenHausler/Multi-Process-Fusion>
- **Look No Deeper: Recognizing Places from Opposing Viewpoints:**
<https://github.com/oravus/seq2single>



“Understanding spatial and perceptual intelligence as a gateway to understanding, creating and applying general intelligence”



Research Philosophy



RatSLAM*: rat-inspired mapping and navigation

Key contributors

Michael
Milford



Gordon
Wyeth



Janet
Wiles



David
Prasser



David
Ball



Brett
Browning

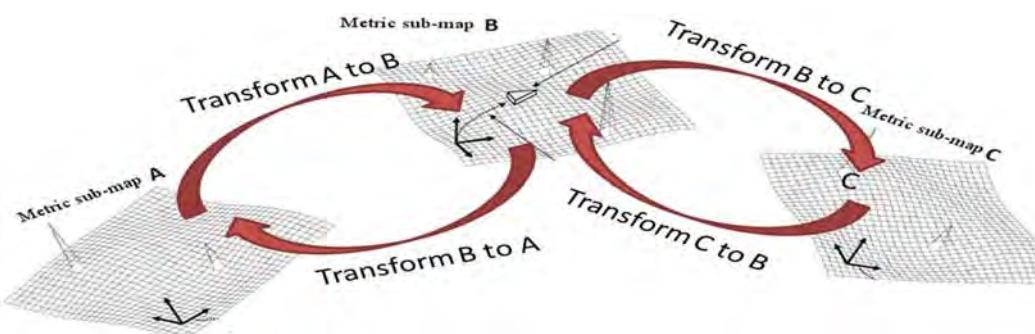
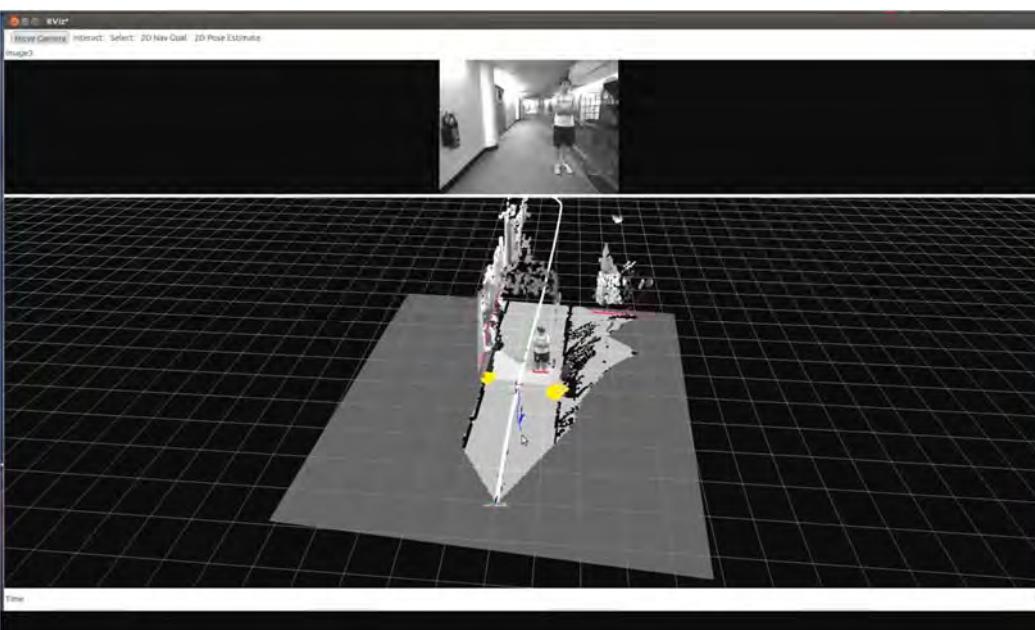


... plus many others over the
past 20 years



***SLAM = Simultaneous Localisation And Mapping**

Spatial Mapping: Robotics versus Nature



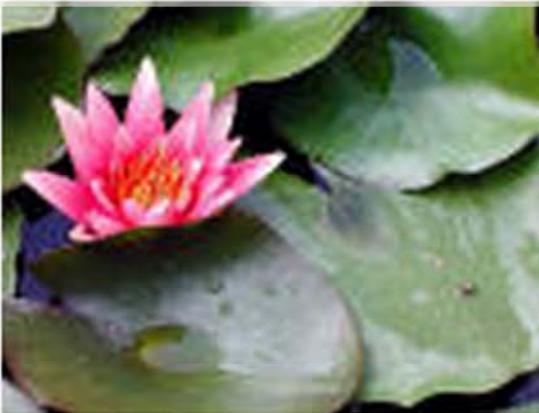
Rats as inspiration



A. P. Buckle and R. H. Smith, Rodent pests and their control. CABI, 2015.

A Jacobson, Z Chen, M Milford , "Autonomous Multisensor Calibration and Closed-loop Fusion for SLAM", *Journal of Field Robotics* 32 (1), 85-122, 2015

Well-characterized Sensing & Perception



Human vision



Normally-pigmented rats have blurry dichromatic vision with a little color



Albino rats may see a very blurry, light-dazzled world

<http://www.ratbehavior.org/RatVision.htm>



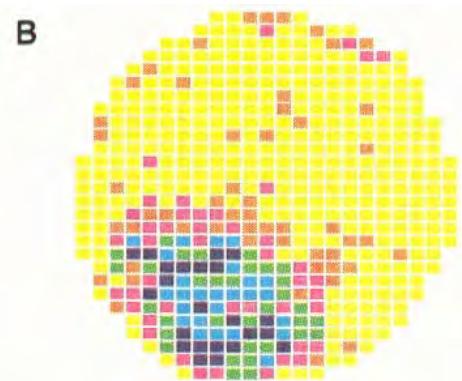
A LED is triggered when the firing rate of the place cell is above 10 Hz.



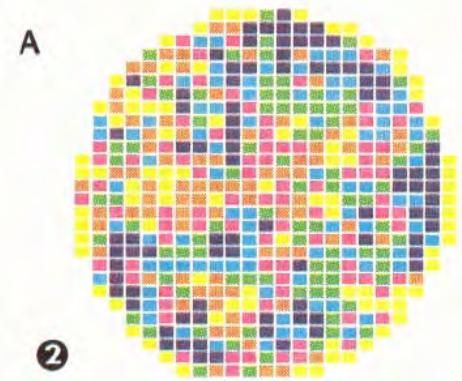
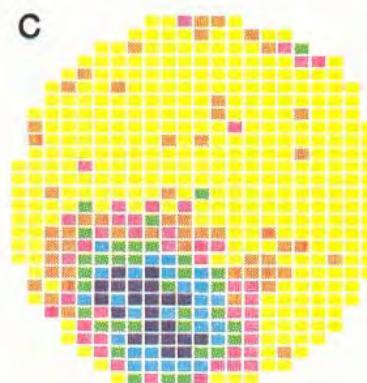
Source: www.gtec.at

Well-characterized Neural Navigation Systems

Place and Head-Direction Cells

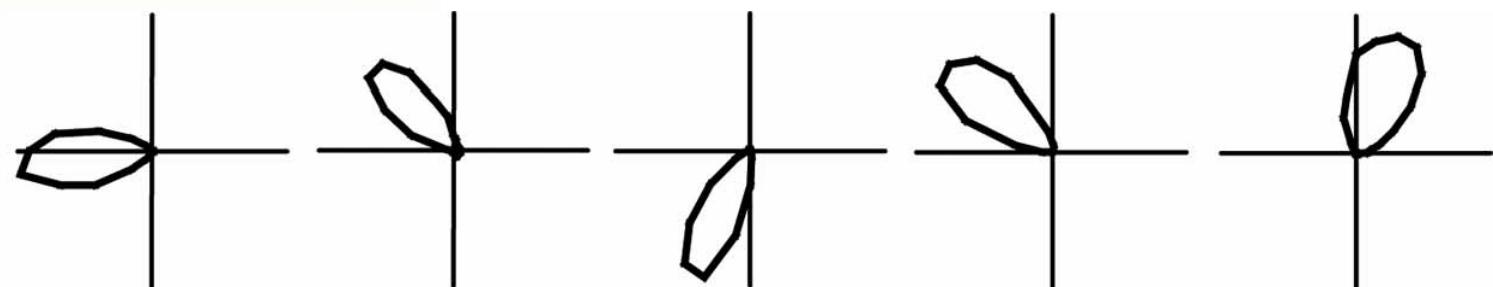


(Muller, R. et al., 1987)



②

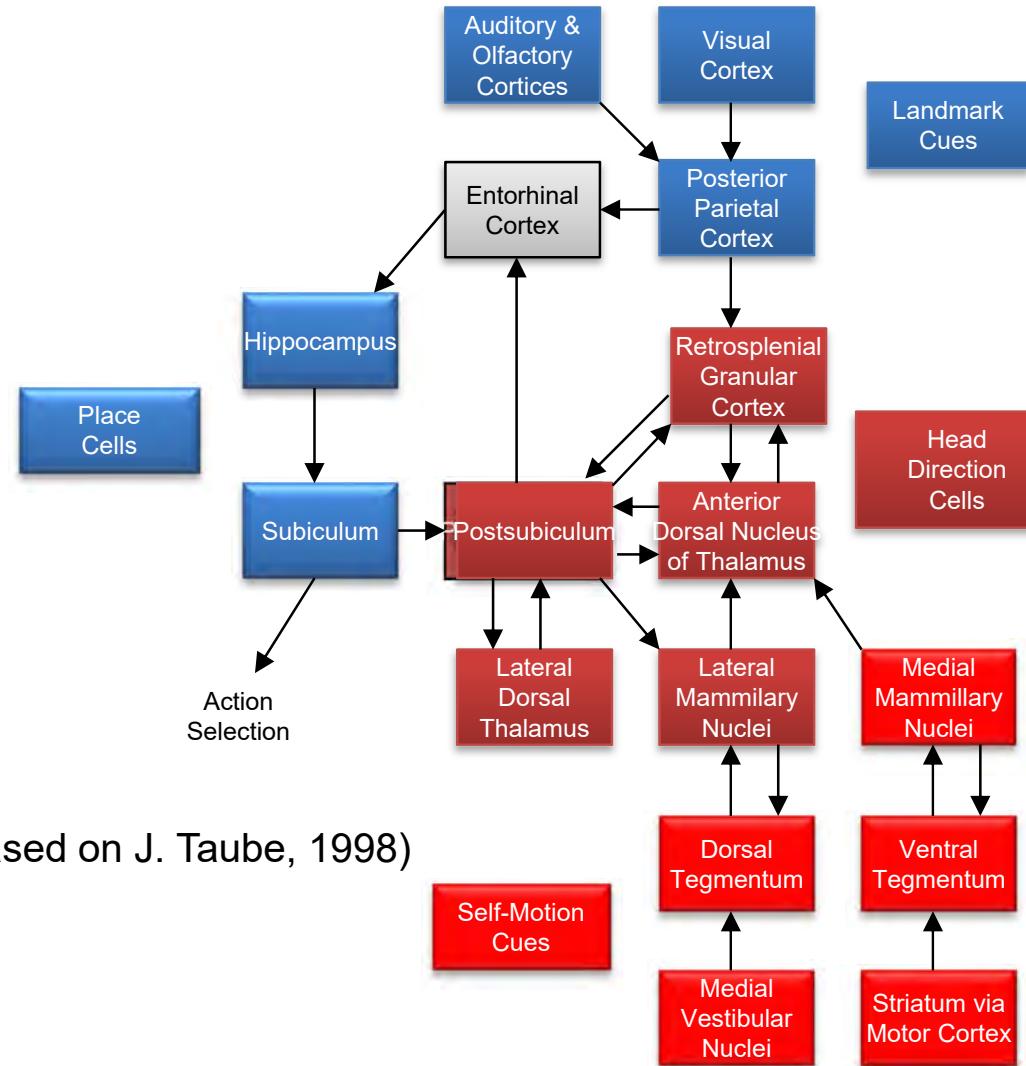
Cells that are perfect
for encoding a robot's
location (1971) and
orientation (1984)



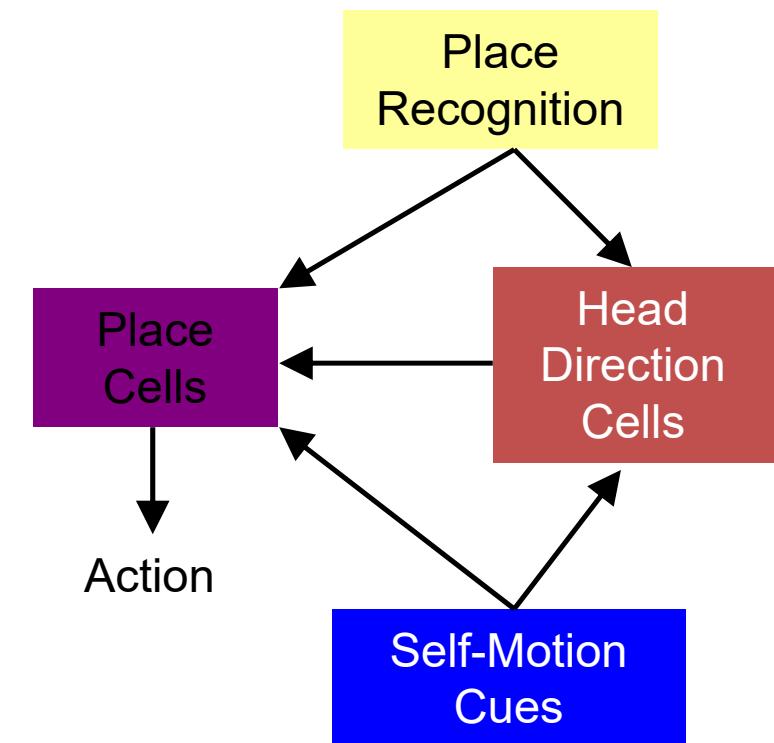
(Yoganarasimha, Yu and Knierim, 2006)

Modelling the Neural System?

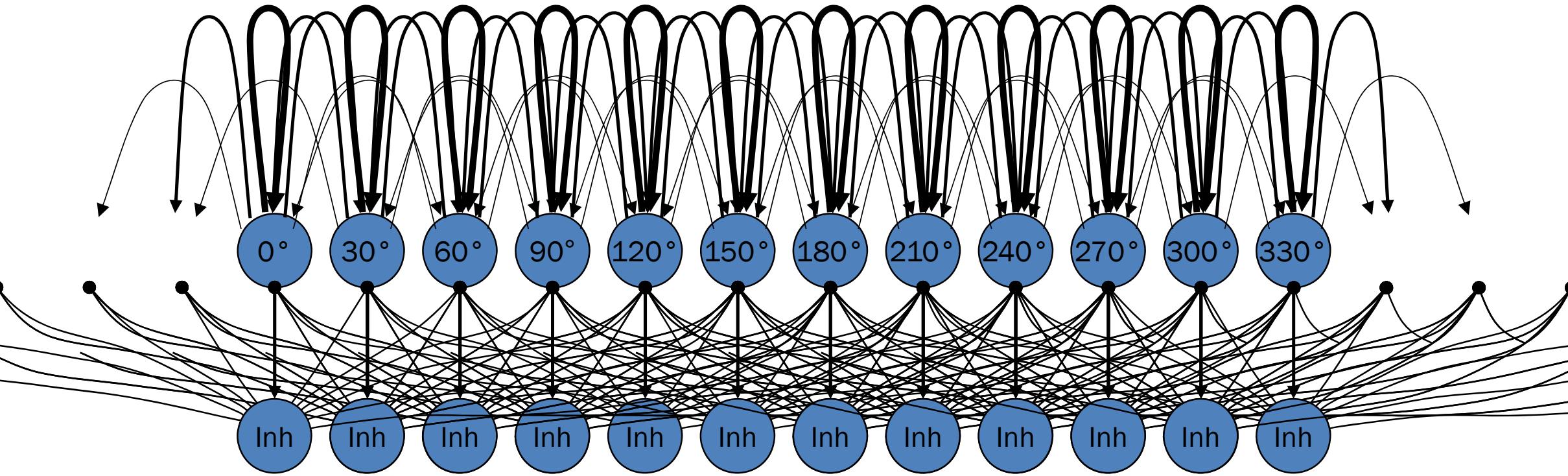
Neuroscientist System Overview



Roboticist Abstraction RatSLAM Mark 1



Modelling with Continuous Attractor Networks (CAN)



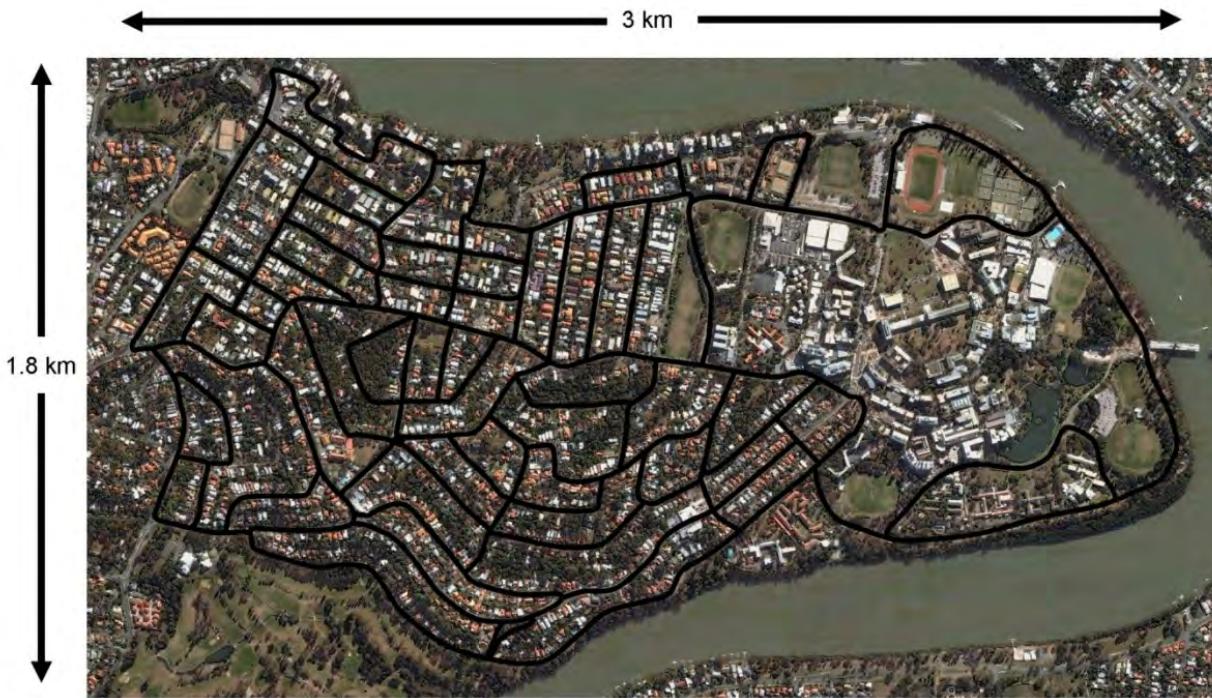
The background of the slide is a detailed, colorful illustration of a suburban neighborhood. It features a grid of streets with dashed white lines, numerous houses represented by grey rectangles with red roofs, and various commercial buildings shown as larger structures with multiple windows. Green areas include lawns, small trees, and a few larger parks with green grass and trees. A few cars are scattered across the roads.

Key Mapping and Navigation Demonstrations

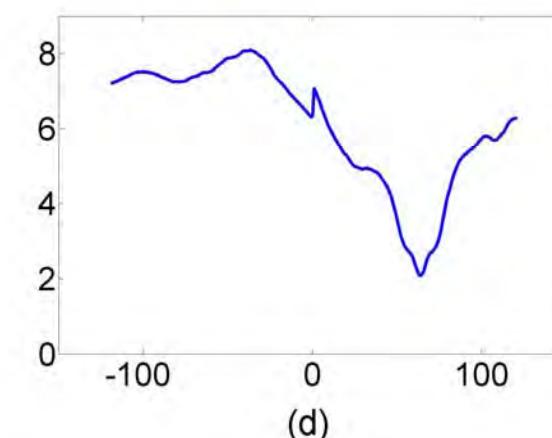
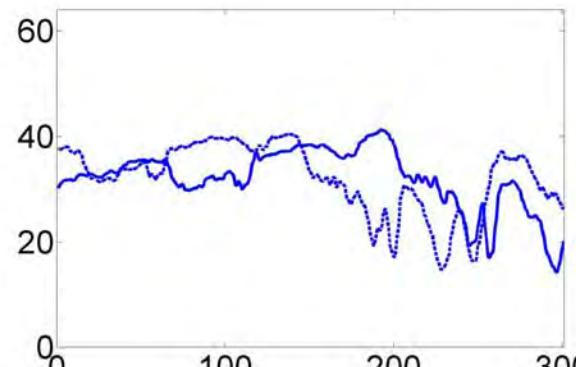
- 1) Mapping a Suburb**
- 2) Persistent Mapping and Navigation**

Mapping a Suburb

- Used vision for local view and odometry.
- Vision from built-in camera of a Mac iBook mounted on experimenter's car.
- Mapped 66 km over just under 2 hours.



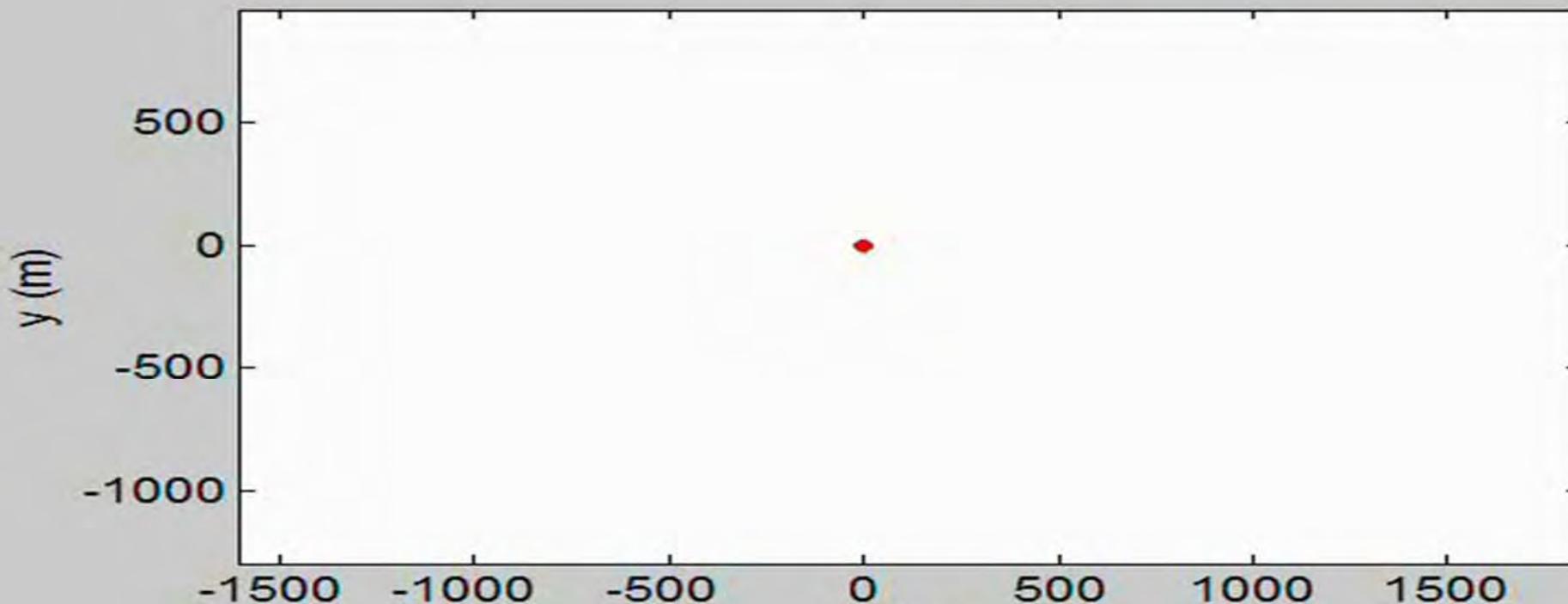
Visual Odometry and Place Recognition

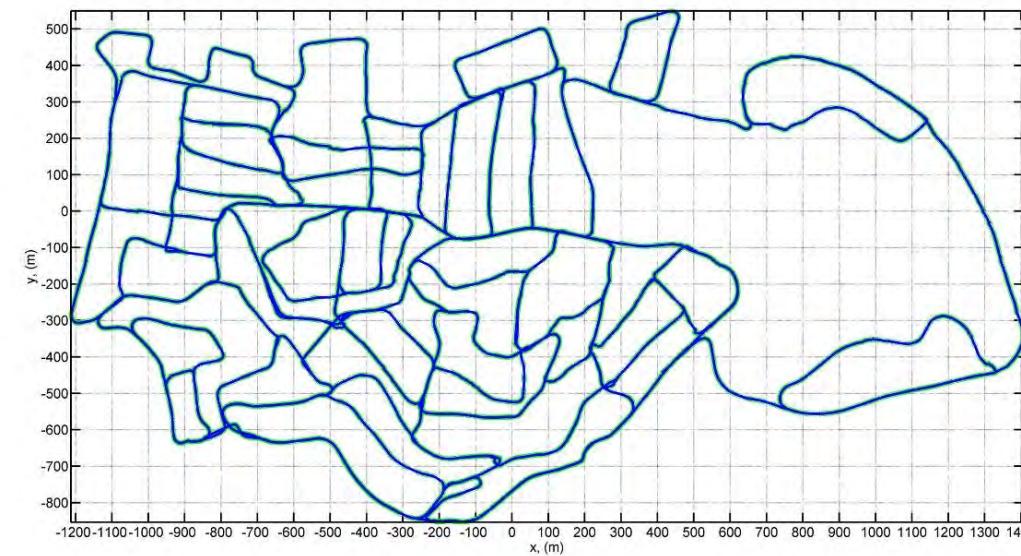
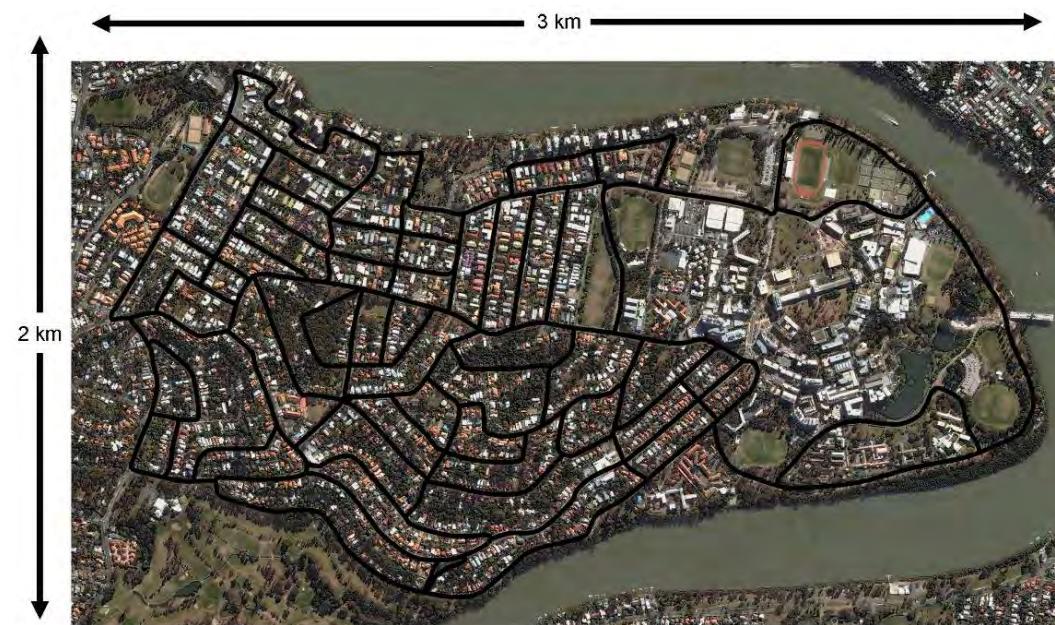


- Forward speed estimated from change in scanline intensity profile between current profile and rotated previous profile.
- Template matching based on profiles with rotation accounted.

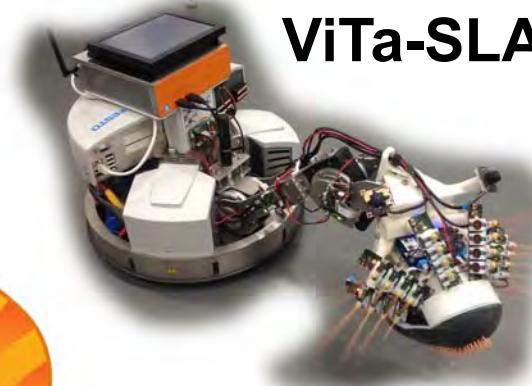
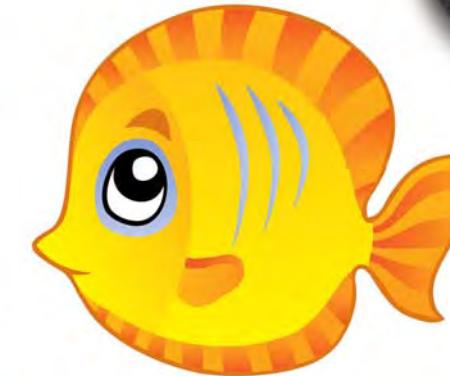
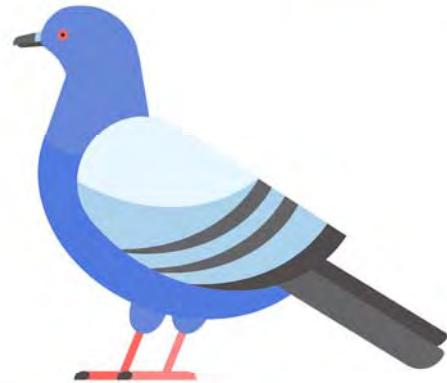
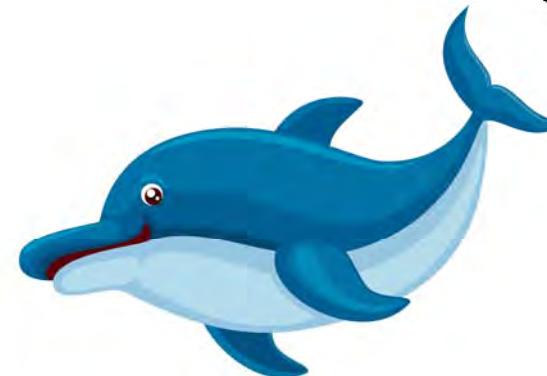
Mapping an Entire Suburb

Time: 1.0 s

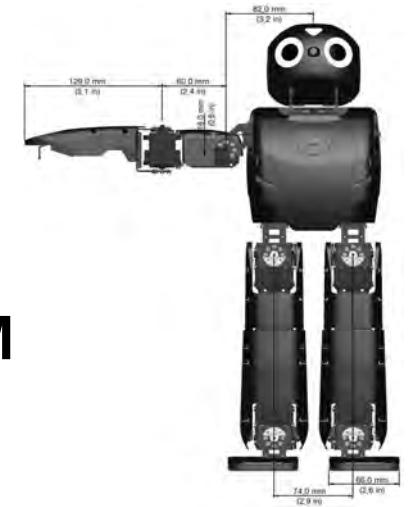




CatSLAM* (Maddern et al) and many others...

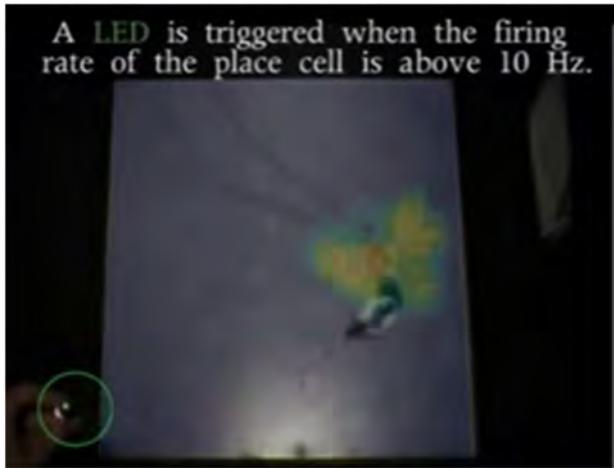


ViTa-SLAM



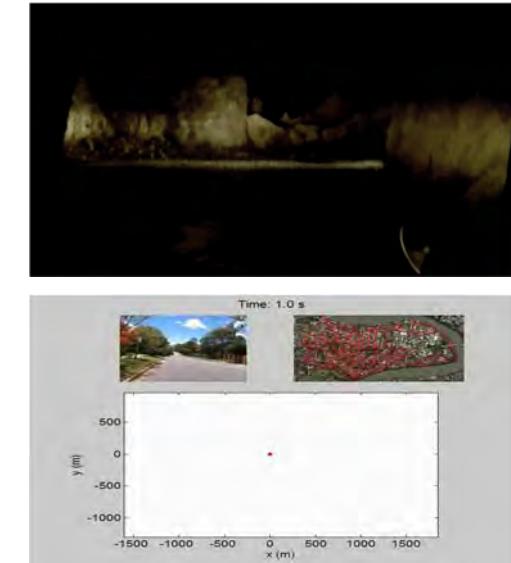
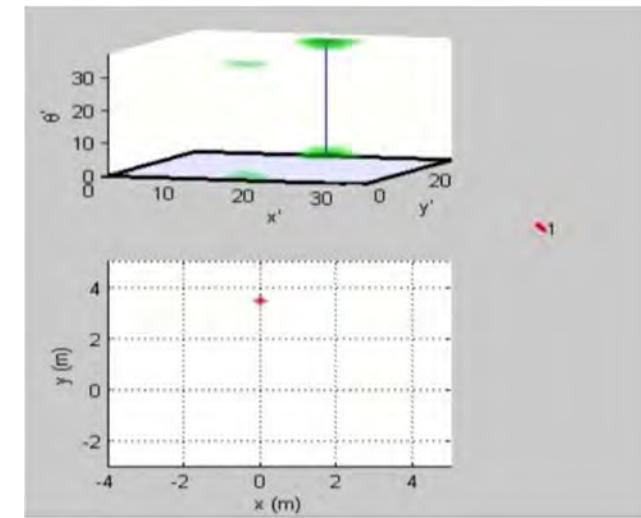
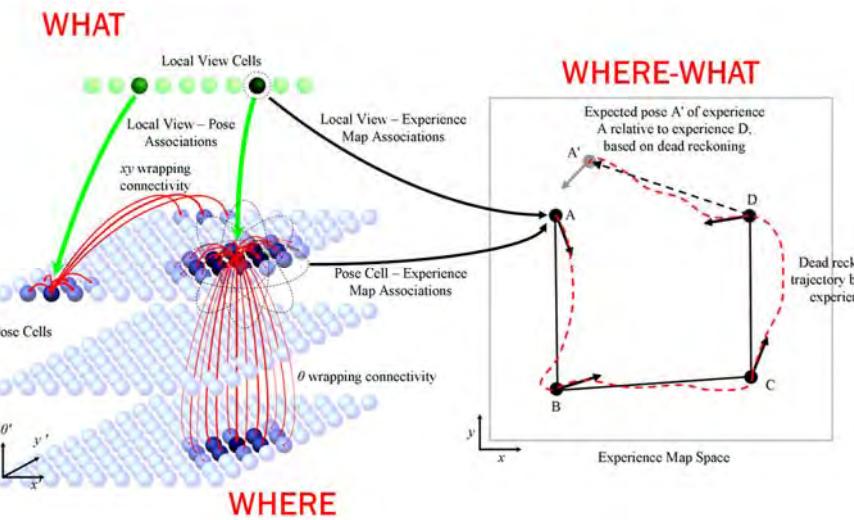
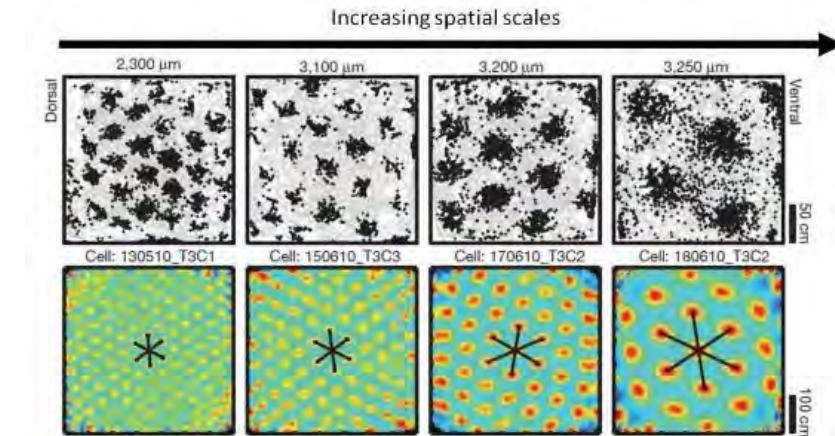
*CatSLAM is
not
biologically
inspired

RatSLAM: Over a decade from neuroscience to deployment



Attractor Connections

0° 30° 60° 90° 120° 150° 180° 210° 240° 270° 300° 330°

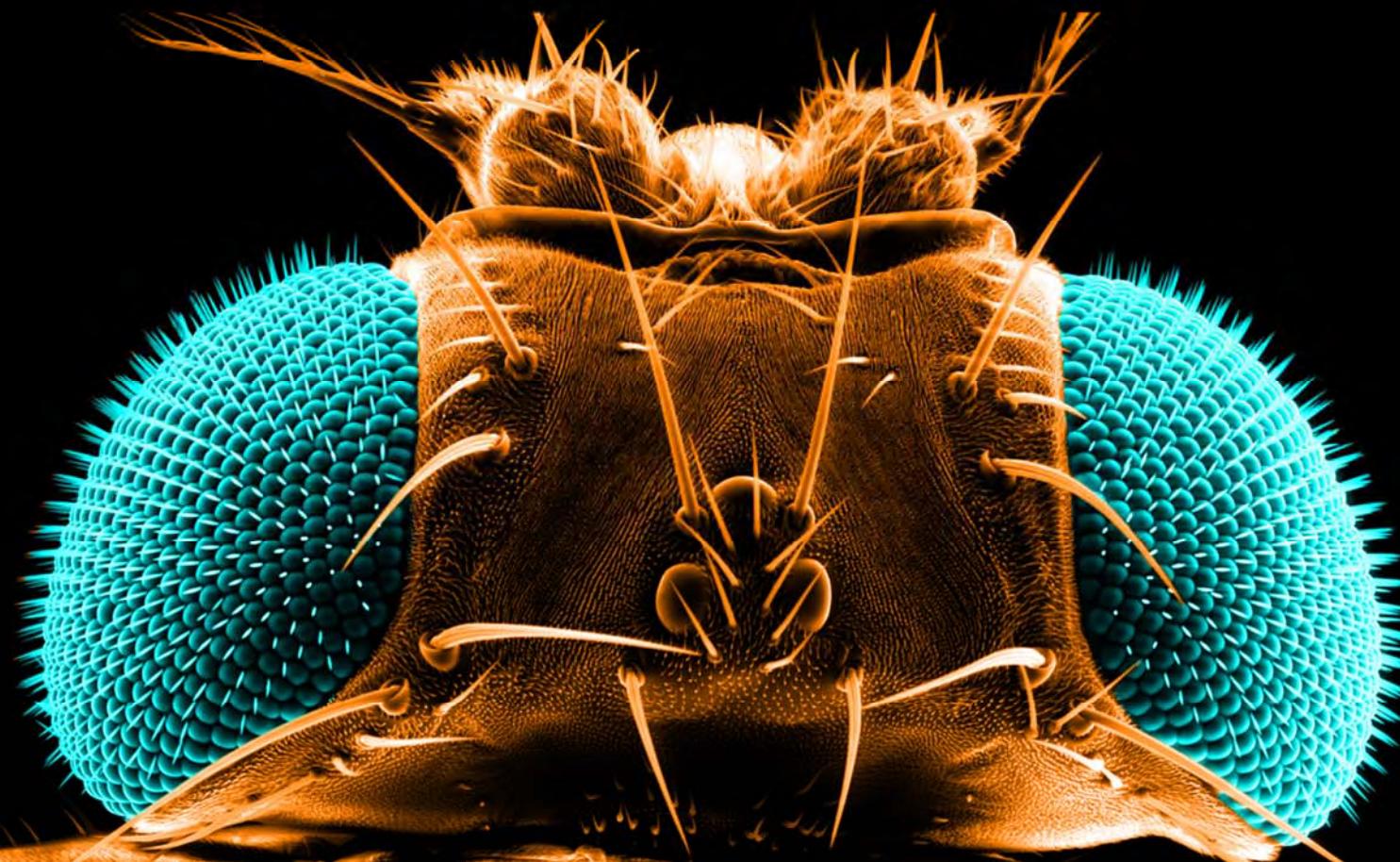


Source: www.gtec.at

MJ Milford, Robot navigation from nature, *Springer Tracts in Advanced Robotics* 41

[2] H. Stensola, T. Stensola, T. Solstad, K. Froland, M. Moser, and E. Moser, "The entorhinal grid map is discretized," *Nature*, vol. 492, pp. 72-78, 2012

Recent & Ongoing Bio-inspired Research



NeuroSLAM: A Brain inspired 6-DOF SLAM System for 3D Environments

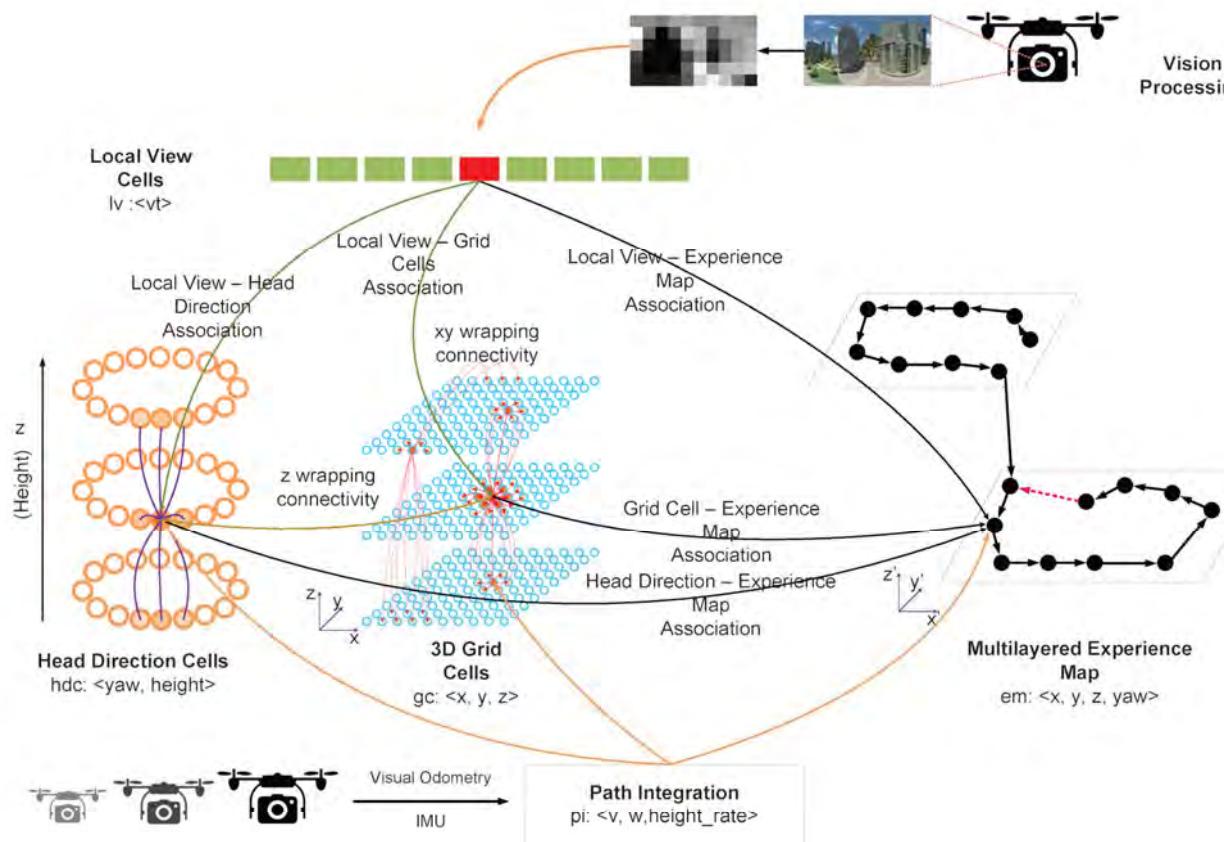
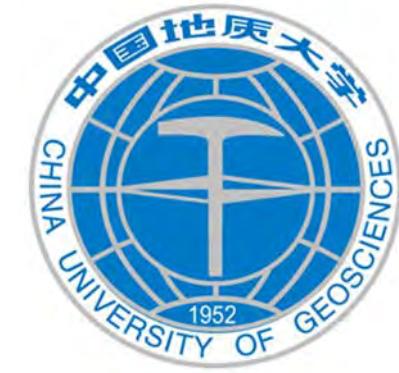
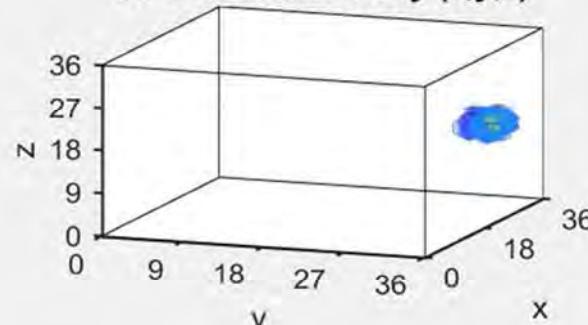
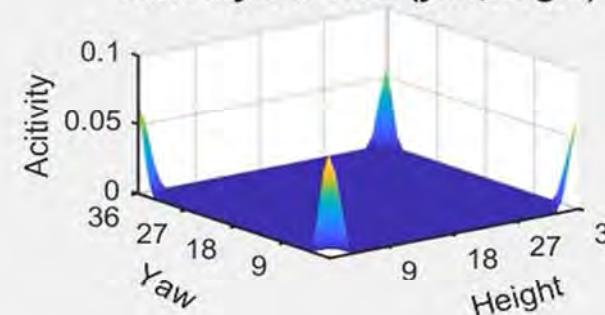
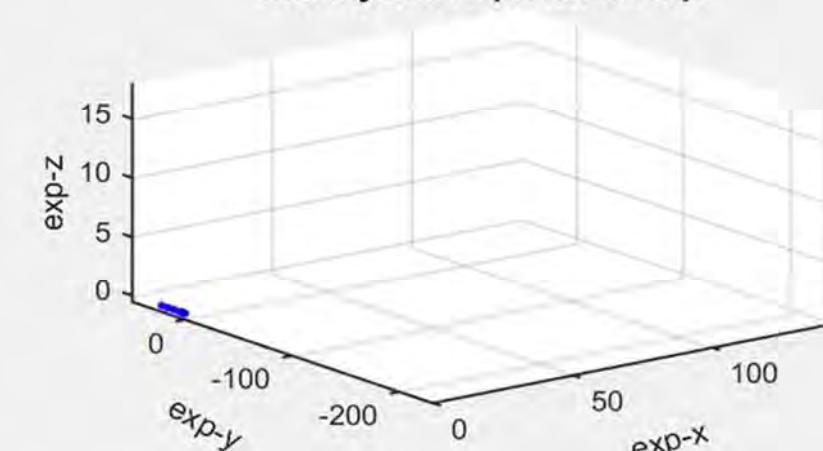
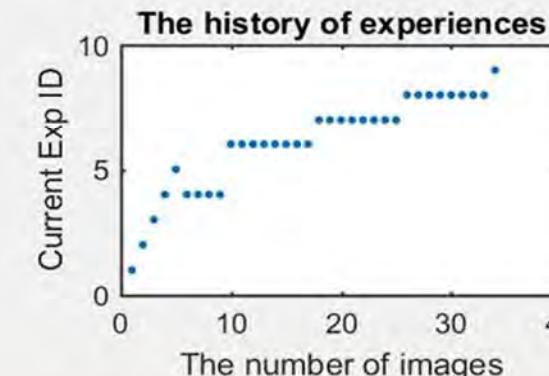
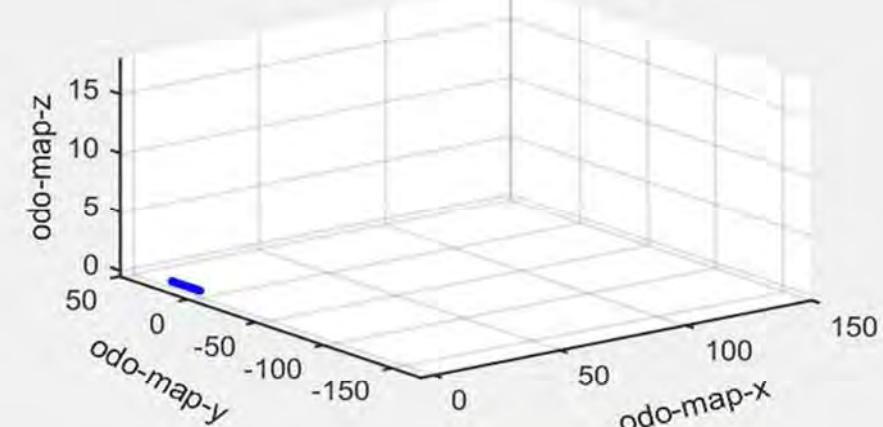
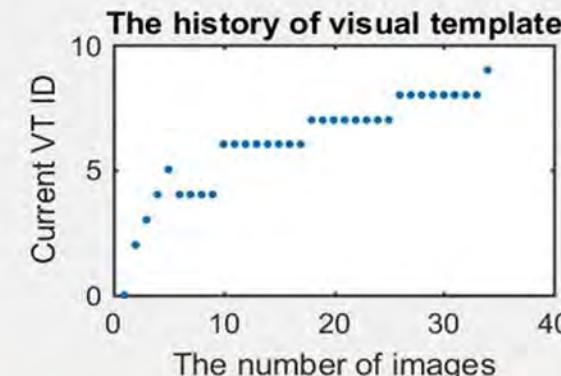


Fig. 1 NeuroSLAM architecture. The system consists of conjunctive pose cells combining the 3D grid cells and multilayered head direction cells, the multilayered experience map and vision modules. The conjunctive pose cell network performs path integration based on the local view cues and self-motion information. Local view cells encode distinct scenes in 3D environment. The self-motion information including translational velocity, altitude velocity and rotational velocity is estimated based on a lightweight 3D visual odometry system. The output from three components of the conjunctive pose cells, local view cells and 3D visual odometry drives the creation of a multilayered experience map, a hybrid spatial representation with topological, metric 3D graphical map of the 3D environment.

Fangwen Yu



This work was supported by the National Key Research and Development Program of China (No. 2016YFB0502200), the Fundamental Research Funds for National University, China University of Geosciences (Wuhan) (No. 1610491T08) and the Hubei Soft Science Research Program (No. LZX2014010).

3D Grid Cell Activity (x,y,z)**Multi-layered HDC (yaw,height)****Multilayered Experience Map****Raw image for vo and vt****Current yaw (decoded from HDC)****Multilayered Odometry Map****The history of visual template**

Winner of a Innovation Grand Prize at the 2019 International Collegiate Competition for Brain-inspired Computing run by Tsinghua University



Bio-inspired Sensing

Event Cameras



Towards Visual SLAM with Event-based Cameras

Michael Milford¹, Hanme Kim², Stefan Leutenegger² and Andrew Davison²

¹Australian Centre for Robotic Vision, Queensland University of Technology

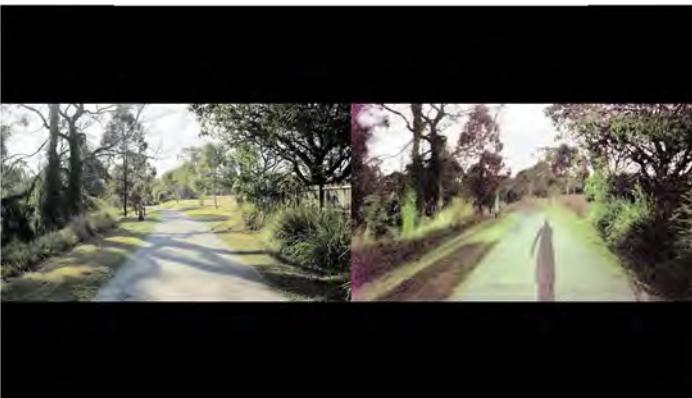
²Department of Computing, Imperial College London

Corresponding author: michael.milford@qut.edu.au

In "The Problem of Mobile Sensors: Setting future goals and indicators of progress for SLAM" Workshop at Robotics and Science Systems 2015



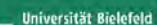
Low light cameras



James
Mount



UV-sensitive cameras

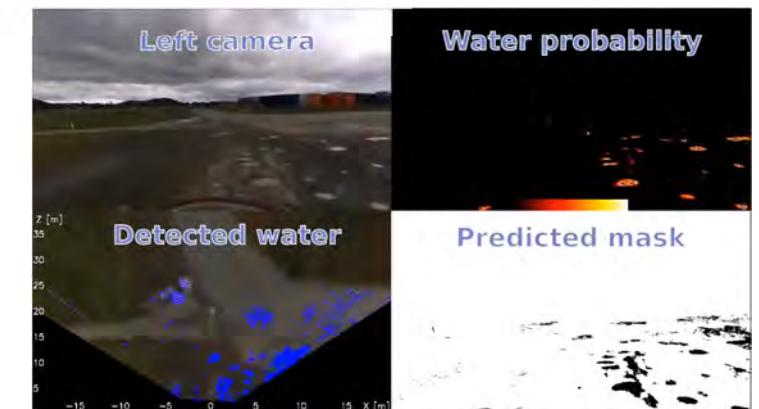


Tom Stone

Stereo Polarized Cameras



Australian
National
University
Chuong Nguyen

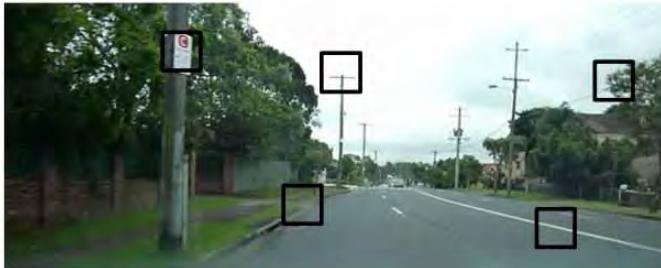




How Bio-inspired Research Can Spur Breakthroughs

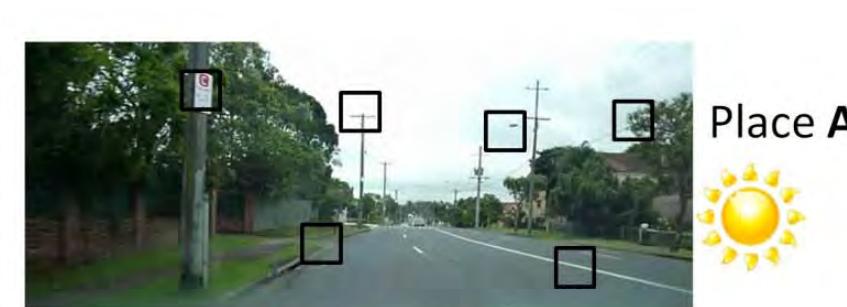
The Core Challenge

Place A



✓ ↕ Same place, low similarity

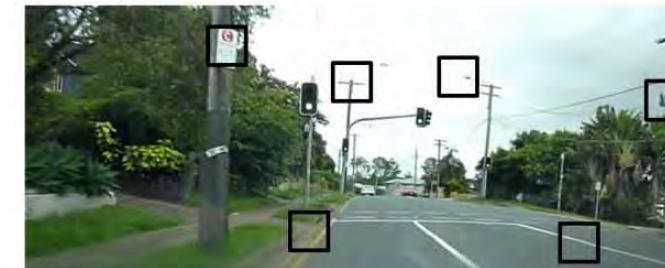
Place A'



Place A



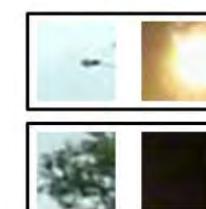
✗ ↕ Difference place, high similarity



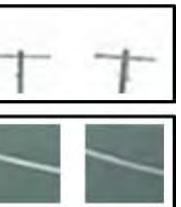
Place B



A-A'



A-B



One of our core research foci

IEEE TRANSACTIONS ON ROBOTICS VOL. 32 NO. 1 FEBRUARY 2016

Visual Place Recognition: A Survey

Stephanie Lowry, Niko Sünderhauf, Paul Newman, Fellow, IEEE, John J. Leonard, Fellow, IEEE, David Cox, Peter Corke, Fellow, IEEE, and Michael J. Milford, Member, IEEE

Abstract—Visual place recognition is a challenging problem due to the vast range of ways in which the appearance of real-world places can vary. In recent years, improvements in visual sensing capabilities, the availability of large-scale datasets, the use of deep learning, and the ability to draw on state-of-the-art results in other disciplines—particularly recognition in computer vision and animal navigation in neuroscience—have contributed to significant advances in the field of visual place recognition. This paper provides a survey of the visual place recognition research landscape. We start by introducing the concepts behind place recognition—the role of place representations, the use of local and global features (as defined in a robotics context), and the need for robustness of a place recognition system. Long-term rated operations have revealed that changing appearance can be a significant factor in visual place recognition failure. We discuss how this has been addressed by using implicitly or explicitly account for appearance change within the environment. Finally, we close with a discussion on the future of visual place recognition, in particular with respect to the rapid advances currently being made in the fields of deep learning, semantic scene understanding, and video description.

Index Terms—Visual place recognition, place recognition.

INTRODUCTION

VISUAL place recognition is a well-defined but extremely challenging problem to solve in the general sense, given the image variability and the lack of context available to determine if this image is of a place it has already seen. Whether referring to humans, animals, computers, or robots, there are some fundamental things a place recognition system must have

(Manuscript received March 17, 2015; revised November 18, 2015; accepted October 23, 2015. This work was partially presented at the IEEE/RSJ International Conference on Intelligent Robots and Systems, Atlanta, GA, October 2015. The paper was recommended by Associate Editor R. Chatila and Editor G. D. Fox upon evaluation of the review committee. This work was supported by the Australian Research Project (ARC) Grant DP130100001, the ARC Future Fellowships FT130100133, and the University of Queensland’s Research Fund.

S. Lowry was with the Australian Centre for Robotic Vision, School of Electrical Engineering and Computer Sciences, Queensland University of Technology, Brisbane, QLD 4000, Australia. She is now with the School of Electrical and Electronic Engineering, University of Manchester, Manchester M13 9PL, U.K. M. J. Milford is with the Australian Centre for Robotic Vision, School of Electrical Engineering and Computer Sciences, Queensland University of Technology, Brisbane, QLD 4000, Australia. E-mail: milford@rmit.edu.au; m.j.milford@qut.edu.au

N. Sünderhauf, P. Corke, and M. J. Milford are with the Australian Centre for Robotic Vision, School of Electrical Engineering and Computer Sciences, Queensland University of Technology, Brisbane, QLD 4000, Australia. E-mail: niko.sunderhauf@qut.edu.au; peter.corke@qut.edu.au

P. Newman is with the Mobile Robots Group, Department of Engineering Science, University of Oxford, Oxford OX1 3PJ, U.K. e-mail: pnewman@eng.ox.ac.uk

J. J. Leonard is with the Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A. e-mail: jleonard@mit.edu

D. Cox is with the Department of Molecular and Cellular Biology, the Institute of Engineering and Applied Sciences, and the Center for Brain Sciences, Harvard University, Cambridge, MA 02138, U.S.A. e-mail: dcox@mit.edu

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TRO.2015.2509882.

Copyright © 2016 IEEE. Personal use is permitted. Any other distribution requires prior permission from IEEE. Please refer to the Terms and Conditions at http://www.ieee.org/publications_standards/publications/rights/index.html for more information.



Fig. 1. Visual place recognition requires both (a) to successfully match (c) previously different images while (b) also rejecting incorrect matches between related images pairs of different places.

Stephanie Lowry, Niko Sünderhauf, Paul Newman, John J. Leonard, David Cox, Peter Corke, and Michael J. Milford, “**Visual Place Recognition: A Survey**”, in *IEEE Transactions on Robotics and Automation*, 32 (1), 2016

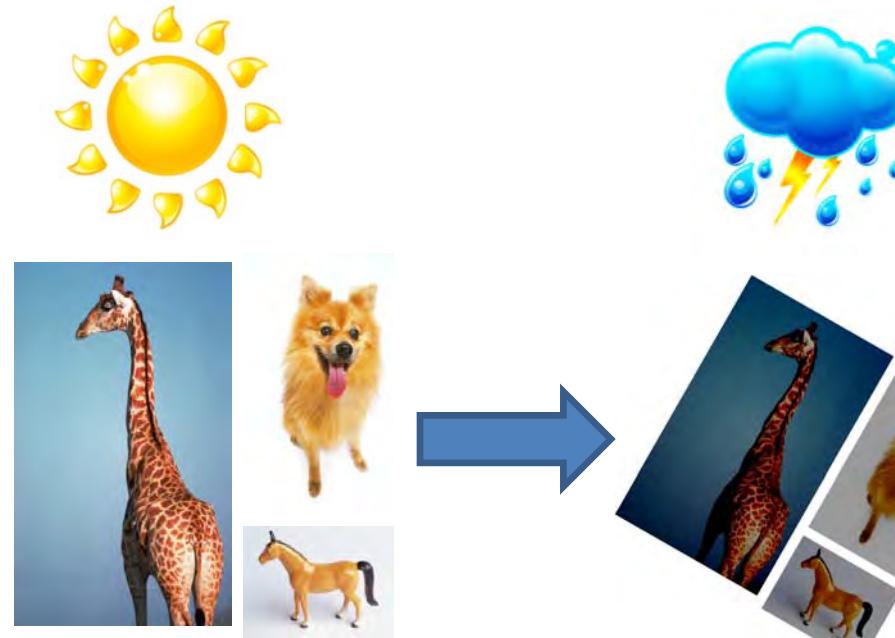
Some papers from 2018-2019

- J Mount, L Dawes, M Milford, “Automatic Coverage Selection for Surface-Based Visual Localization”, *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2019
- S Hausler, A Jacobson, M Milford, “Filter Early, Match Late: Improving Network-Based Visual Place Recognition,” *IEEE International Conference on Robotics and Automation*, 2019
- Sourav Garg, V Babu, Thanuja Dharmasiri, Stephen Hausler, Niko Suenderhauf, Swagat Kumar, Tom Drummond, Michael Milford, “Look no deeper: Recognizing places from opposing viewpoints under varying scene appearance using single-view depth estimation”, *IEEE International Conference on Robotics and Automation*, 2019
- S Hausler, A Jacobson, M Milford, “Multi-Process Fusion: Visual Place Recognition Using Multiple Image Processing Methods”, *IEEE Robotics and Automation Letters* 4 (2), 2019.
- S Garg, N Suenderhauf, M Milford, “Semantic–geometric visual place recognition: a new perspective for reconciling opposing views”, *The International Journal of Robotics Research*, 2019
- J Mao, X Hu, X He, L Zhang, L Wu, MJ Milford, “Learning to Fuse Multiscale Features for Visual Place Recognition”, *IEEE Access* 7, 5723-5735, 2018
- S Garg, N Suenderhauf, M Milford, “Don’t look back: Robustifying place categorization for viewpoint-and condition-invariant place recognition”, *IEEE International Conference on Robotics and Automation*, 2018
- Y Latif, R Garg, M Milford, I Reid, “Addressing challenging place recognition tasks using generative adversarial networks”, *IEEE International Conference on Robotics and Automation*, 2018
- S Garg, N Suenderhauf, M Milford, “Lost? appearance-invariant place recognition for opposite viewpoints using visual semantics”, in *Robotics Science and Systems*, 2018
- L Yu, A Jacobson, M Milford, “Rhythmic representations: Learning periodic patterns for scalable place recognition at a sublinear storage cost”, *IEEE Robotics and Automation Letters* 3 (2), 811-818, 2018

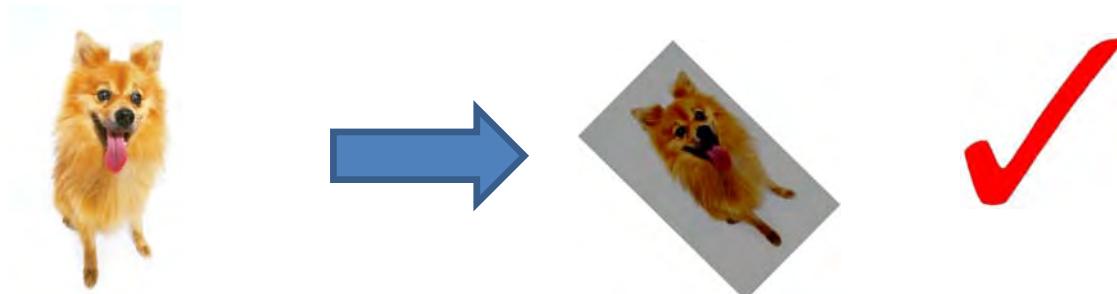


Increasing selectivity & tolerance

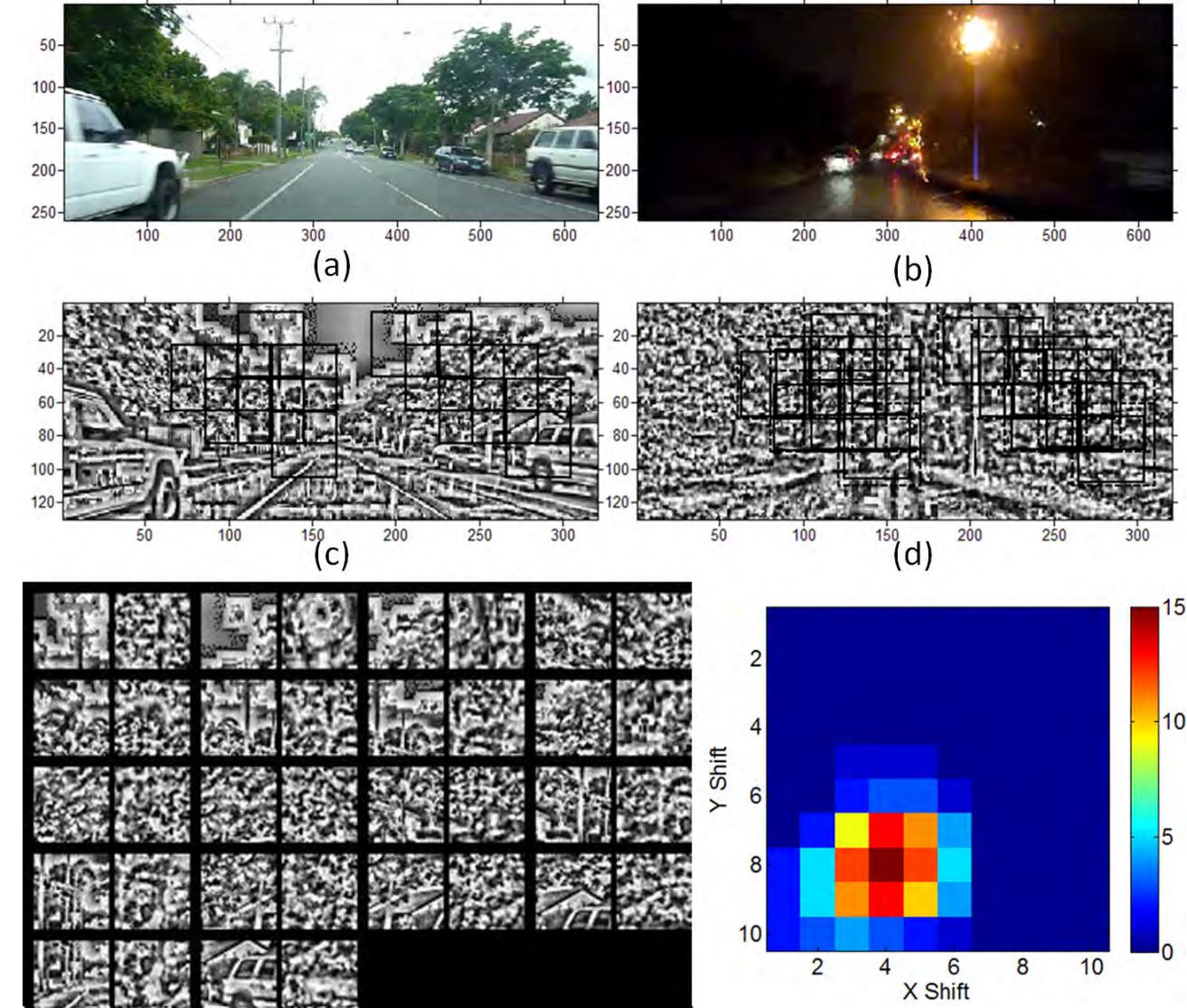
Early stage



Later stage



Correctly confirmed match: true positive

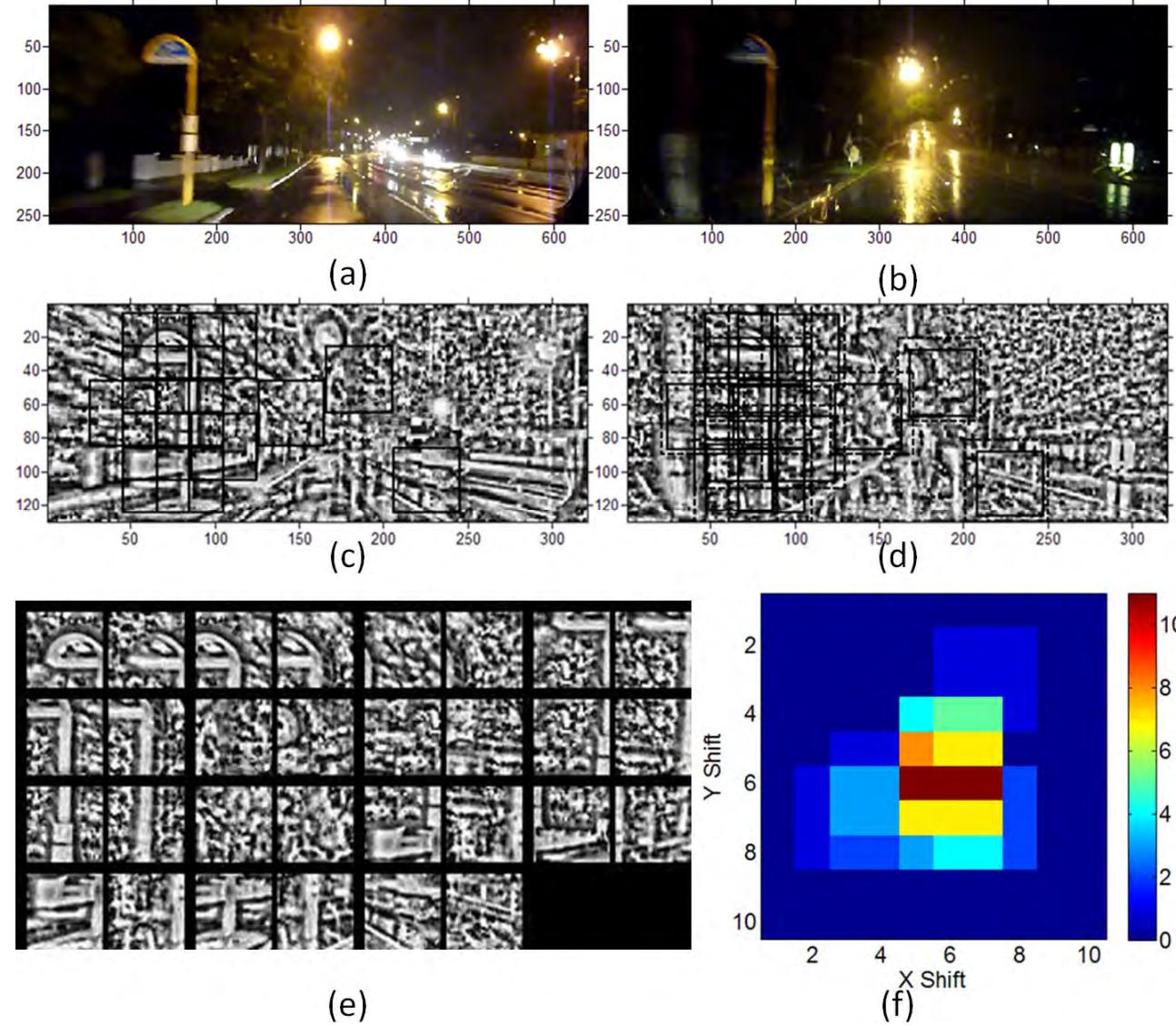


- Michael Milford, Eleonora Vig, Walter J. Scheirer, David D. Cox, "Vision-based Simultaneous Localization and Mapping in Changing Outdoor Environments", in *Journal of Field Robotics*, 31 (5), 2014.
- Michael Milford, Walter J. Scheirer, Eleonora Vig, Arren Glover, Oliver Baumann, Jason Mattingley, David D. Cox, "Condition-Invariant, Top-Down Visual Place Recognition," *IEEE International Conference on Robotics and Automation*, 2014

Correctly rejected match hypothesis: true negative



**Beat the Machine
Game**



- Michael Milford, Eleonora Vig, Walter J. Scheirer, David D. Cox, "Vision-based Simultaneous Localization and Mapping in Changing Outdoor Environments", in *Journal of Field Robotics*, 31 (5), 2014.
- Michael Milford, Walter J. Scheirer, Eleonora Vig, Arren Glover, Oliver Baumann, Jason Mattingley, David D. Cox, "Condition-Invariant, Top-Down Visual Place Recognition," *IEEE International Conference on Robotics and Automation*, 2014



How SeqSLAM Came About

- Experimentation with spike train sequence generation
- Very low resolution images proven by prior RatSLAM bio-inspired work
- Attempting to generate self-sustaining spike trains in software, corresponding to image sequences
- Final SeqSLAM was an algorithmic, non-spiking simplification

| | x_0 | | | | | | | | x_7 |
|---------|---|-----|--|-----|---|-----|---|---|-------|
| Images |  | ... |  | ... |  | ... |  | | |
| Indexes | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Bit 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | |
| Bit 2 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | |
| Bit 3 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | |

BTEL: A Binary Tree Encoding Approach for Visual Localization



Huu Le¹, Tuan Hoang², and Michael Milford³

Contact: Professor Michael Milford, michael.milford@qut.edu.au

¹Chalmers University of Technology, ²Singapore University of Technology and Design, ³QUT



The Nuances of Compression & Storage

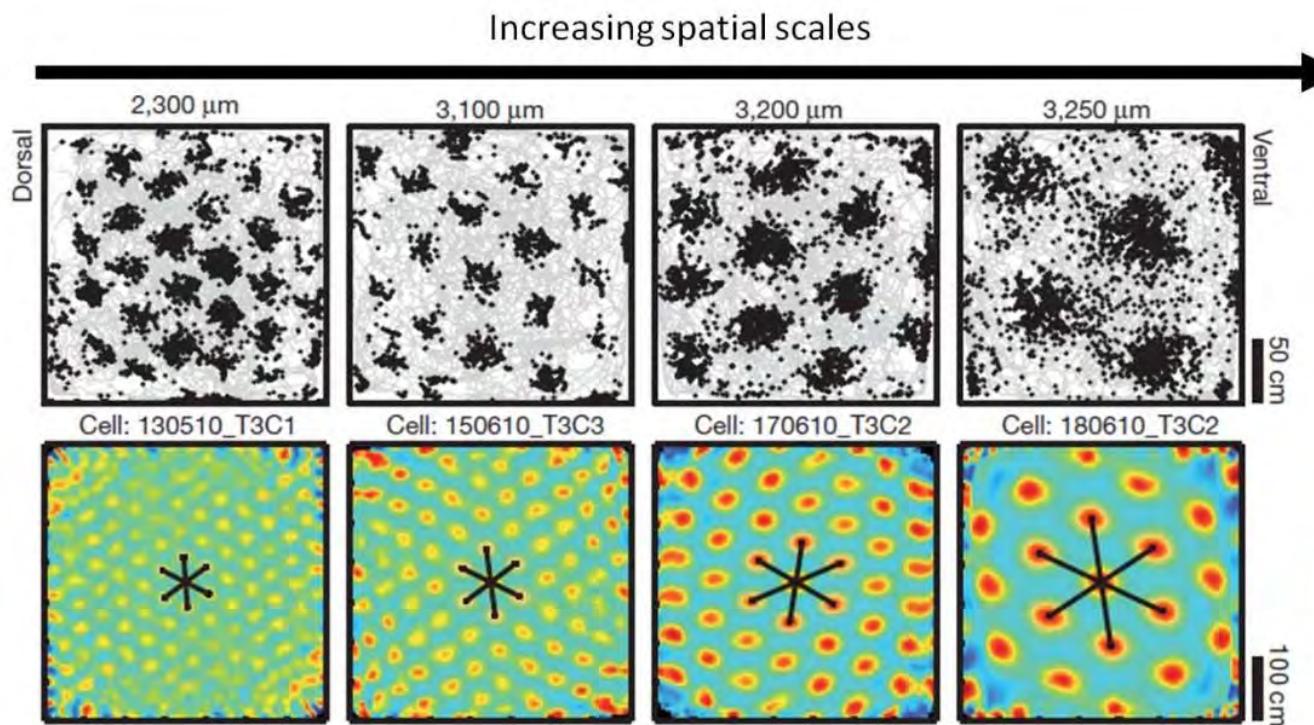
- Early days of robotics: critical factors for deployment & feasibility
- Recent years: move towards focus on maximal recall / accuracy / precision / other performance
- All other things being equal, better compression & storage enables:
 - Cheaper, less bulky / power hungry compute hardware
 - On-board rather than off-board operations
 - Better absolute performance with no growth in compute

Absolute versus Scalability

- Much focus on absolute scalability
- But still at least linear growth
- Can we achieve sub-linear **storage** growth?
- Can we achieve this while maintaining competitive performance?
- Can we achieve sub-linear growth while maintaining compact **absolute** storage requirements?

Grid Cells (2004/2005)

- Multi-scale grid cell mapping. ~5+ scales, $\sqrt{2}$ scaling

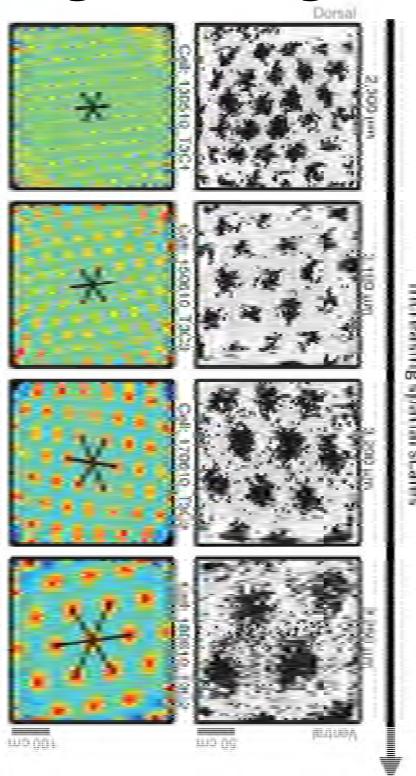


2014 Nobel Prize for
Physiology or Medicine:
Edvard Moser, May-Britt
Moser and John O'Keefe



[2] H. Stensola, T. Stensola, T. Solstad, K. Froland, M. Moser, and E. Moser, "The entorhinal grid map is discretized," *Nature*, vol. 492, pp. 72-78, 2012

A Mystery We've Been Investigating Thoroughly

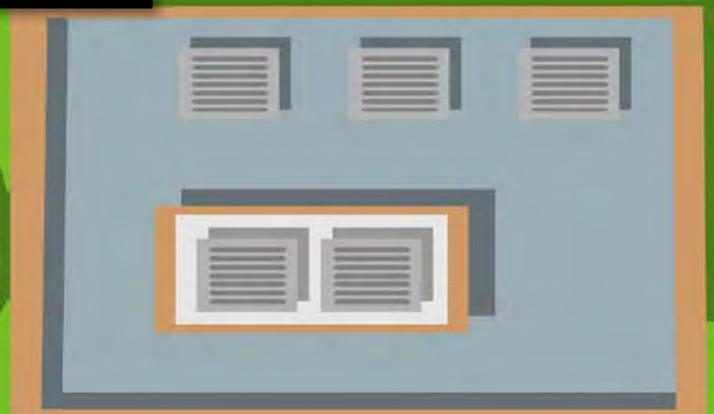
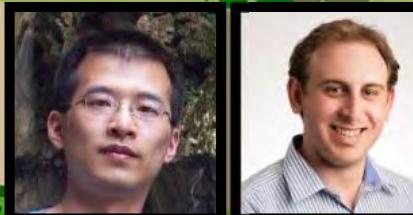


- Why have multiple scales?
- What scale ratio?
- How to set the scales?
- Interaction with all sensing modalities?
- What memory and computational advantages?

H. Stensola, T. Stensola, T. Solstad, K. Froland, M. Moser, and E. Moser, "The entorhinal grid map is discretized," *Nature*, vol. 492, pp. 72-78, 2012

- Huu Le, Tuan Hoang, Michael J Milford, "BTEL: A Binary Tree Encoding Approach for Visual Localization", in *IEEE International Conference on Intelligent Robots and Systems*, 2019
- A Jacobson, Z Chen, M. Milford, *Biological Cybernetics*, 2018
- Litao Yu, Adam Jacobson, Michael J Milford, "Rhythmic Representations: Learning Periodic Patterns for Scalable Place Recognition at a Sub-Linear Storage Cost", in *IEEE Robotics and Automation Letters*, 2018
- Huu Le, Anders Eriksson, Thanh-Toan Do, Michael Milford, "A Binary Optimization Approach for Constrained K-Means Clustering" in *Asian Conference on Computer Vision*, 2018.
- Chen Fan, Zetao Chen, Adam Jacobson, Xiaoping Hu and Michael Milford, "Biologically-inspired Visual Place Recognition with Adaptive Multiple Scales", in press in *Robotics and Autonomous Systems*, 2017.
- Adam Jacobson, Walter Scheirer and Michael Milford, "De ja vu: Scalable Place Recognition Using Mutually Supportive Feature Frequencies", in *IEEE International Conference on Intelligent Robots and Systems*, 2017
- Zetao Chen, Stephanie Lowry, Adam Jacobson, Michael E Hasselmo, Michael Milford, "Bio-inspired homogeneous multi-scale place recognition", in *Neural Networks*, 2015.
- Z Chen, A Jacobson, UM Erdem, ME Hasselmo, M Milford, "Multi-scale bio-inspired place recognition," *IEEE International Conference on Robotics and Automation*, 2014.
- MJ Milford, J Wiles, GF Wyeth, "Solving navigational uncertainty using grid cells on robots", *PLoS Computational Biology* 6 (11), 2010

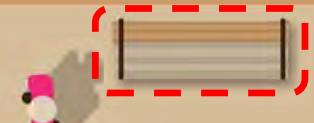
Exploiting a Cyclic World



Feature 1
Frequency F_1



Feature 2
Frequency F_2



Exploiting a Cyclic World

Place ID: 1 2 3 4 5 6 7 8

Feature 1: 0 1 0 1 0 1 0 1

Feature 2: 0 1 2 0 1 2 0 1 2

Memory collision

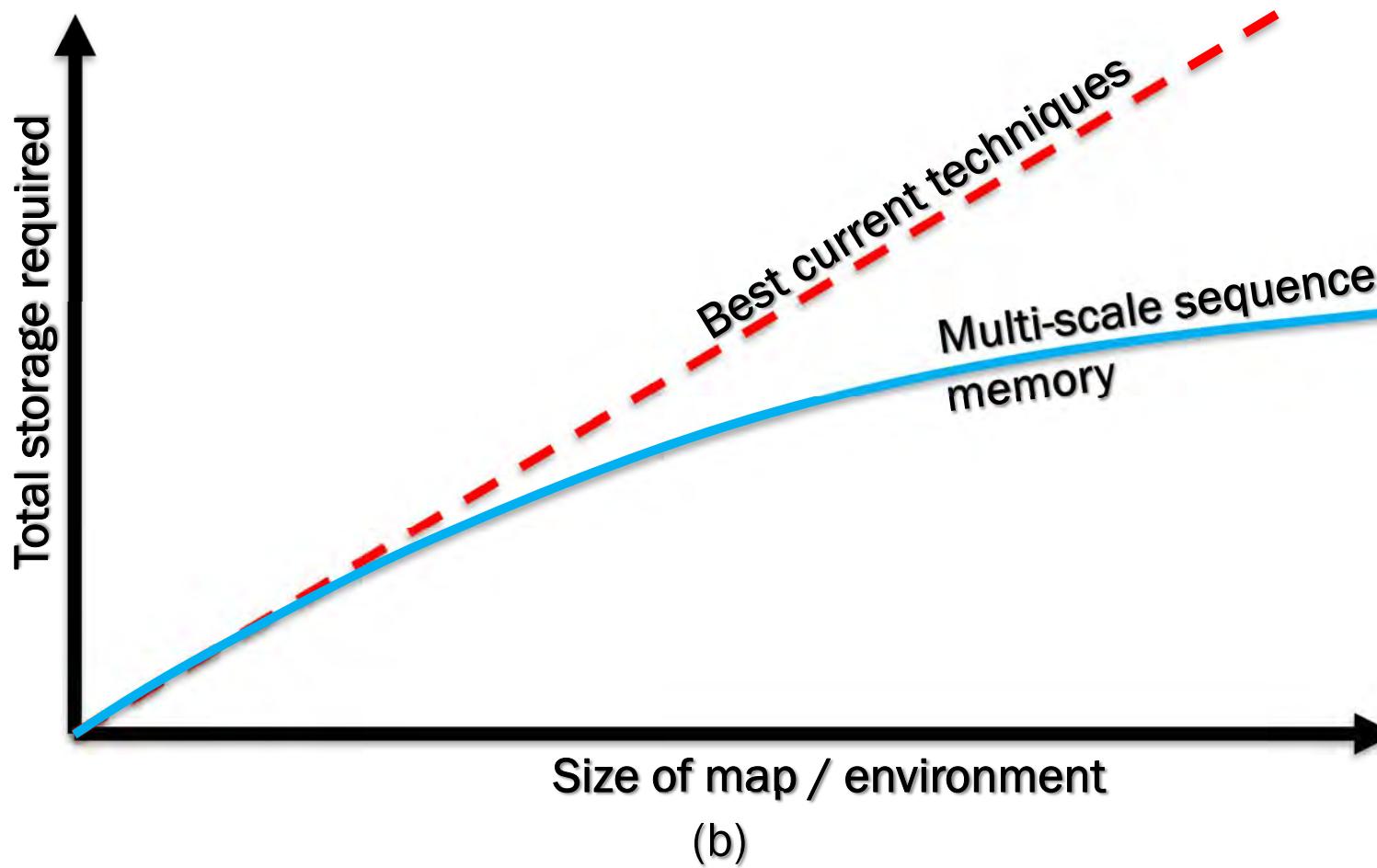
6 places, 5 units of storage

Feature 1: [0 1]

Feature 2: [0 1 2]

- Huu Le, Tuan Hoang, Michael J Milford, "BTEL: A Binary Tree Encoding Approach for Visual Localization", in *IEEE International Conference on Intelligent Robots and Systems*, 2019
- Litao Yu, Adam Jacobson, Michael J Milford, "Rhythmic Representations: Learning Periodic Patterns for Scalable Place Recognition at a Sub-Linear Storage Cost", in *IEEE Robotics and Automation Letters*, 2018
- Huu Le, Anders Eriksson, Thanh-Toan Do, Michael Milford, "A Binary Optimization Approach for Constrained K-Means Clustering" in *Asian Conference on Computer Vision*, 2018.
- Adam Jacobson, Walter Scheirer and Michael Milford, "De ja vu: Scalable Place Recognition Using Mutually Supportive Feature Frequencies", in *IEEE International Conference on Intelligent Robots and Systems*, 2017

Real World Sub-Linear Dataset Compression



- Huu Le, Tuan Hoang, Michael J Milford, "BTEL: A Binary Tree Encoding Approach for Visual Localization", in *IEEE International Conference on Intelligent Robots and Systems*, 2019
- Litao Yu, Adam Jacobson, Michael J Milford, "Rhythmic Representations: Learning Periodic Patterns for Scalable Place Recognition at a Sub-Linear Storage Cost", in *IEEE Robotics and Automation Letters*, 2018
- Huu Le, Anders Eriksson, Thanh-Toan Do, Michael Milford, "A Binary Optimization Approach for Constrained K-Means Clustering" in *Asian Conference on Computer Vision*, 2018.
- Adam Jacobson, Walter Scheirer and Michael Milford, "De ja vu: Scalable Place Recognition Using Mutually Supportive Feature Frequencies", in *IEEE International Conference on Intelligent Robots and Systems*, 2017

| | x_0 | | | | | | | | x_7 |
|---------|-------|-----|---|-----|---|-----|---|---|-------|
| Images | | ... | | ... | | ... | | | |
| Indexes | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | |
| Bit 1 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | |
| Bit 2 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 1 | |
| Bit 3 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | |

BTEL: A Binary Tree Encoding Approach for Visual Localization

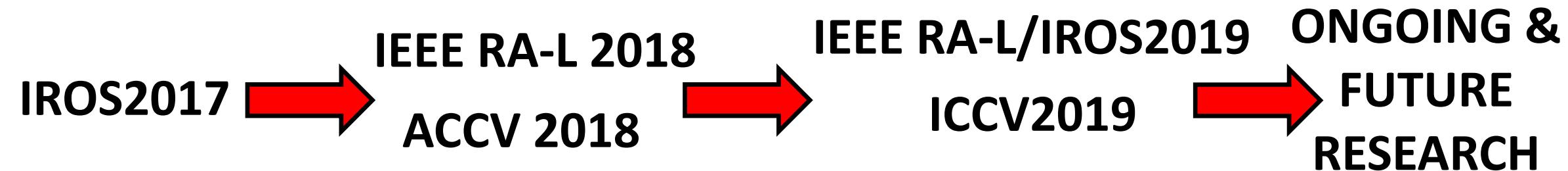
Huu Le¹, Tuan Hoang², and Michael Milford³

Contact: Professor Michael Milford, michael.milford@qut.edu.au

¹Chalmers University of Technology, ²Singapore University of Technology and Design, ³QUT



Future Work



- Co-investigating both **absolute storage compression** and **sub-linear scaling**
- Scaling up to **global-size datasets**



Applications:
Industry and Government Projects

Example Application Areas



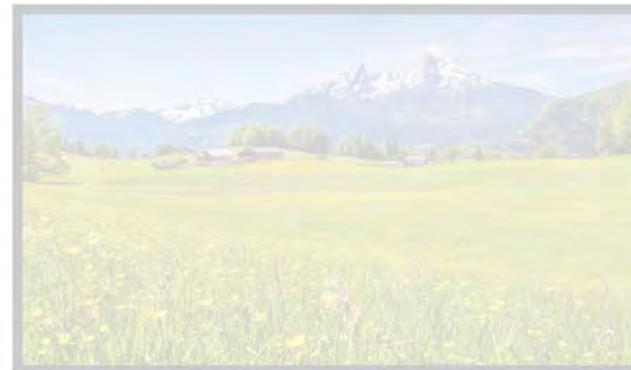
Positioning Systems for Autonomous Mining Vehicles

Robust multimodal
toolpoint positioning



How Automated Vehicles Will Interact With Road Infrastructure Now and in the Future

Robust hazard detection on construction and mining sites



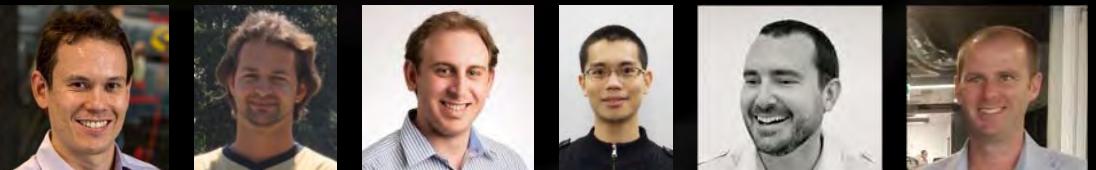
Automating Analysis of Vegetation with Computer Vision: Cover Estimates and Classification



An Infinitely Scalable Learning and Recognition Network



Automation-enabling positioning for underground mining



No drive in the park...

Clear images



Low light

Water



Dust

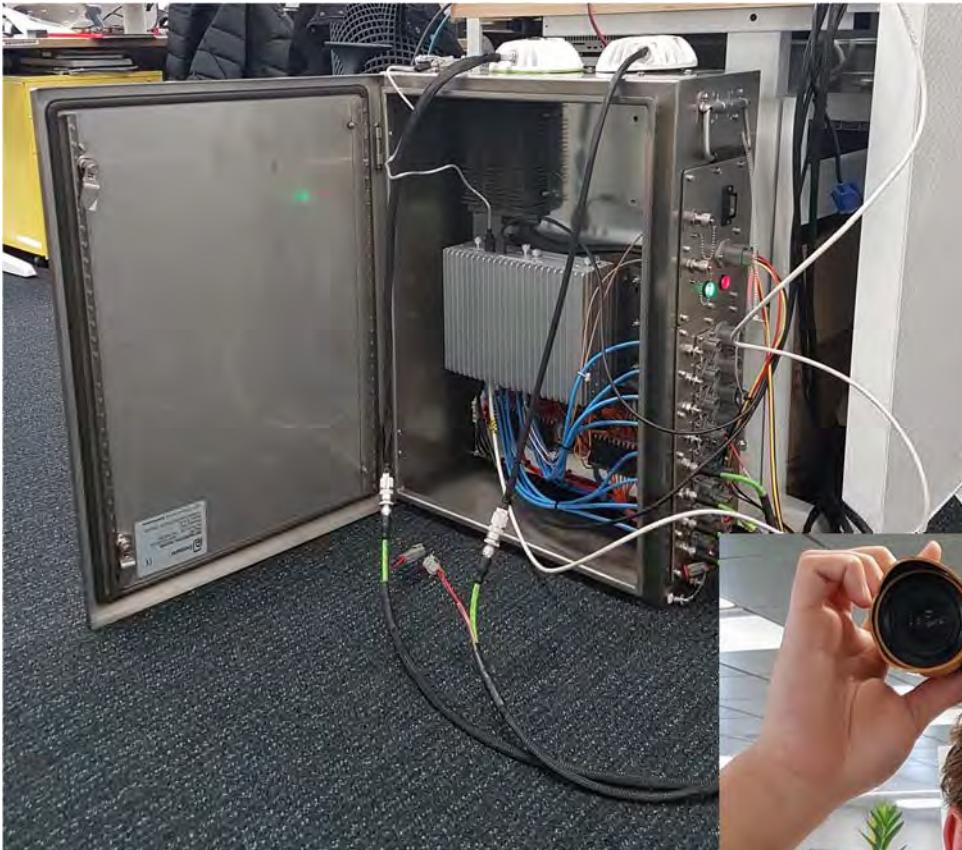


Glare



- Fan Zeng, Adam Jacobson, David Smith, Nigel Boswell, Thierry Peynot, Michael J Milford , “TIMTAM: Tunnel-Image Textually-Accorded Mosaic for Location Refinement of Underground Vehicles with a Single Camera”, in *IEEE/RSJ International Conference on Intelligent Robots*, Macau, China 2019.
- Fan Zeng, Adam Jacobson, David Smith, Nigel Boswell, Thierry Peynot, Michael J Milford , “LookUP: Vision-Only Real-Time Precise Underground Localisation for Autonomous Mining Vehicles”, in *IEEE International Conference on Robotics and Automation*, 2019.
- Adam Jacobson, Fan Zeng, David Smith, Nigel Boswell, Thierry Peynot, Michael J Milford, “Semi-Supervised SLAM: Leveraging Low-Cost Sensors on Underground Autonomous Vehicles for Position Tracking”, in *IEEE/RSJ International Conference on Intelligent Robots*, Madrid, Spain 2018.

The Early Days...



"Octobox"

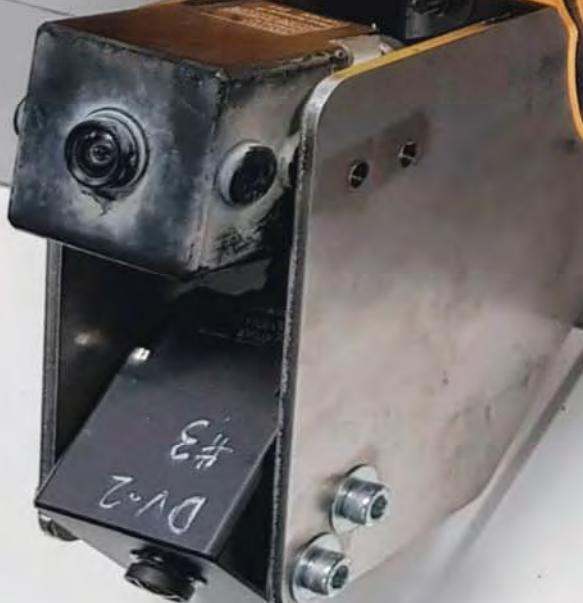


Cameras



Adam

SICK



SICK

162-162

SICK



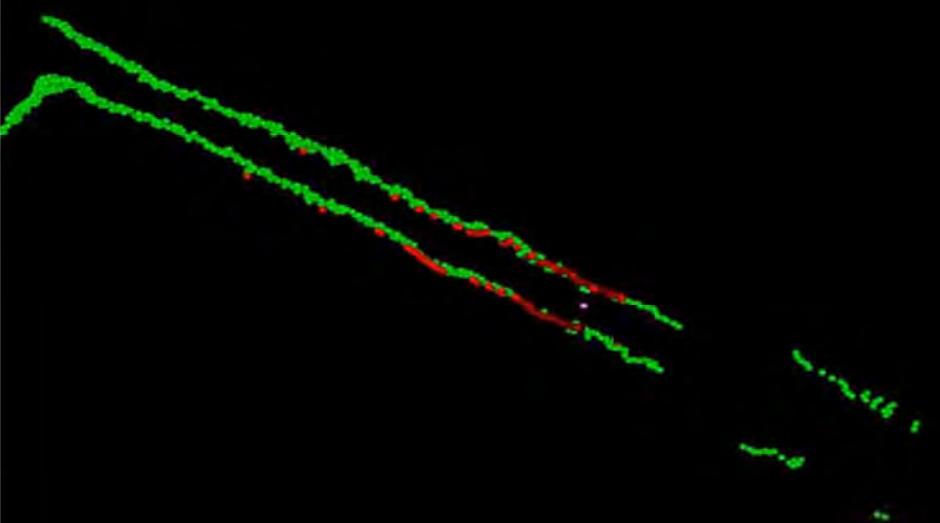


192 168 1 161

192 168 1 161

S10C



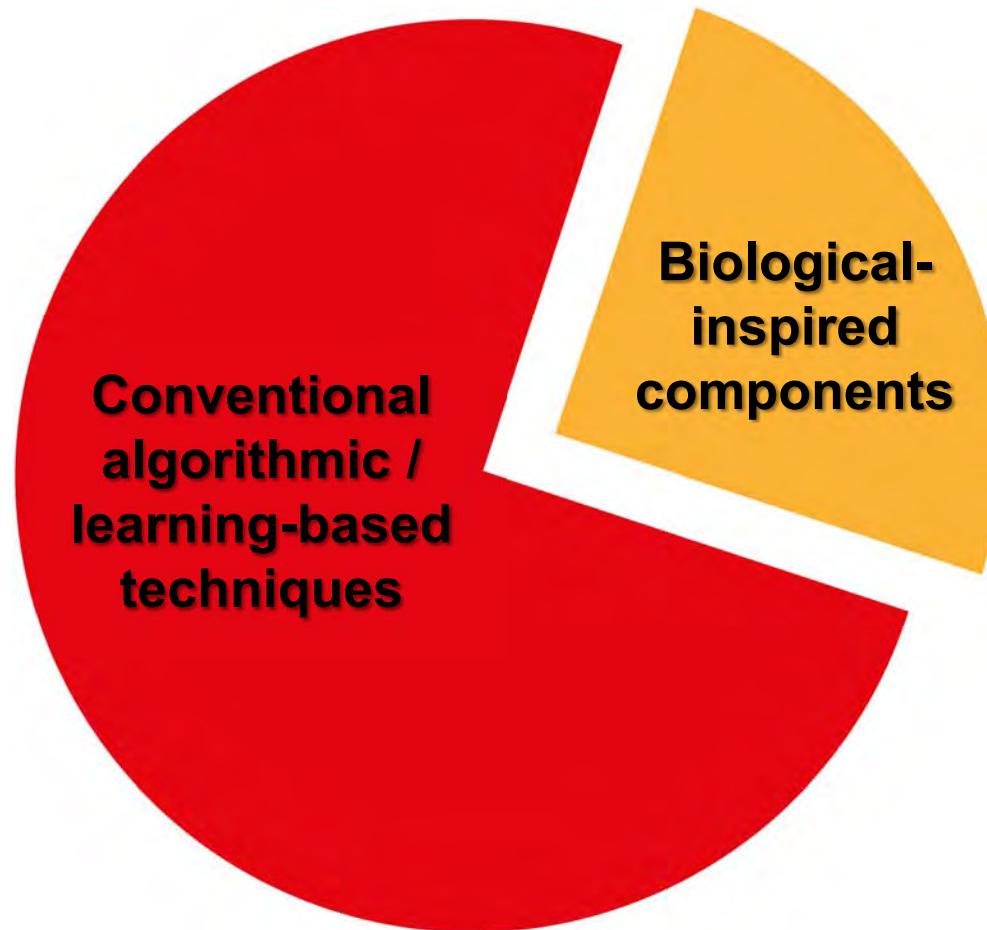


Partially censored for
confidentiality reasons



- Fan Zeng, Adam Jacobson, David Smith, Nigel Boswell, Thierry Peynot, Michael J Milford , “TIMTAM: Tunnel-Image Textually-Accorded Mosaic for Location Refinement of Underground Vehicles with a Single Camera”, in *IEEE/RSJ International Conference on Intelligent Robots*, Macau, China 2019.
- Fan Zeng, Adam Jacobson, David Smith, Nigel Boswell, Thierry Peynot, Michael J Milford , “LookUP: Vision-Only Real-Time Precise Underground Localisation for Autonomous Mining Vehicles”, in *IEEE International Conference on Robotics and Automation*, 2019.
- Adam Jacobson, Fan Zeng, David Smith, Nigel Boswell, Thierry Peynot, Michael J Milford, “Semi-Supervised SLAM: Leveraging Low-Cost Sensors on Underground Autonomous Vehicles for Position Tracking”, in *IEEE/RSJ International Conference on Intelligent Robots*, Madrid, Spain 2018.

Going from biological inspiration to deployment



- Mixture of bio-inspired + conventional.
- Why?:
 - Sensor differences (both limited and opportunistic)
 - AI limitations
 - Embodiment differences
 - Different risk appetites
 - Provability

What Bio-inspiration Made it In

- Image processing techniques partially derived from bio-inspired research
- Short bespoke sequence-matching techniques
- Topological mapping techniques partially derived from bio-inspired research
- Image matching techniques derived from fundamental primate-inspired vision research several years ago



We are hiring!

Current and upcoming roles including PhDs, Postdocs,
Research Engineers, and Academic Roles

We work
here



E-mail: michael.milford@qut.edu.au

QUT

Collaboration Opportunities



Some groups we have published with or held joint grants with



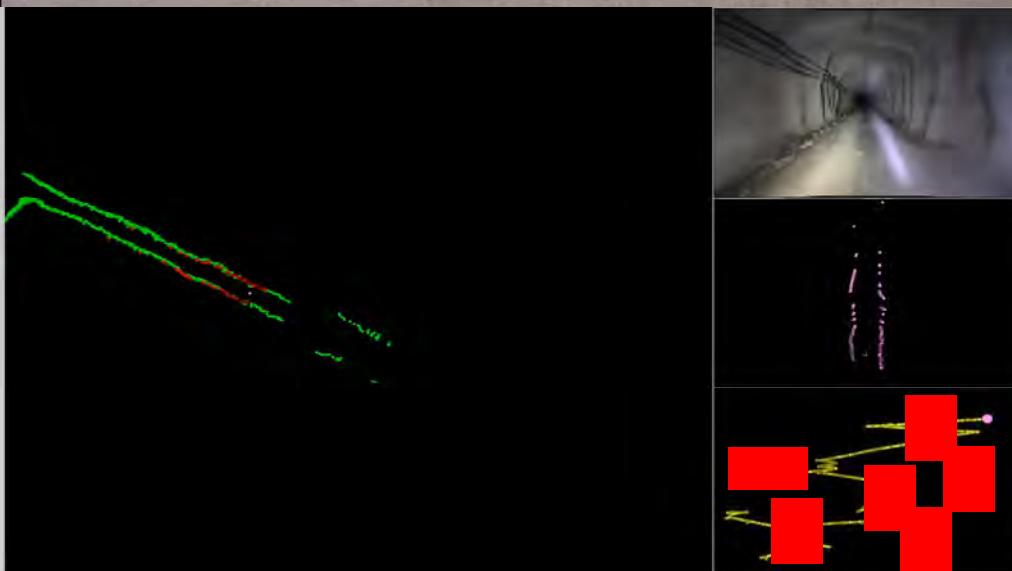
- Access to unique or limited access datasets / sensors / compute
- Non-critical-path but important big picture research problems
- Co-authored publications
- Grants
- Consulting
- Part-time academic roles
- Student & researcher exchanges
- Intern programs
- Co-organization of workshops / conferences etc...

Thank you to our
collaborators,
and our funders,
including:



Australian Government

Australian Research Council



Professor Michael Milford | Australian Research Council Future Fellow | Microsoft Research Faculty Fellow | Chief Investigator, Australian Centre for Robotic Vision
michael.milford@qut.edu.au



michael.milford@qut.edu.au

Twitter: @maththrills

<https://www.youtube.com/milfordrobot>

<http://www.tinyurl.com/milforddm>

<https://goo.gl/rczsle>