ViTa-SLAM: Biologically-Inspired Visuo-Tactile SLAM

Oliver Struckmeier, Kshitij Tiwari, Martin J. Pearson, and Ville Kyrki

***Abstract*— In this work, we propose a novel, bio-inspired multi-sensory SLAM approach called ViTa-SLAM. Compared to other multi-sensory SLAM variants, this approach allows for a seamless multi-sensory information fusion whilst naturally interacting with the environment. The algorithm is empirically evaluated in a simulated setting using a biomimetic robot platform called the WhiskEye. Our results show promising performance enhancements over existing bio-inspired SLAM approaches in terms of loop-closure detection.**

I. Introduction

Of late, there is growing interest in biologically inspired SLAM algorithms. For instance, the rat hippocampus inspired RatSLAM [1] provides a biologically-inspired 3 DOF SLAM architecture which has been shown to work for robots operating under a myriad of environmental conditions. While the vanilla RatSLAM was purely designed as a visual SLAM approach, there are additional works that now explain its usage with other sensory modalities like WiFi [2] and auditory signals [3].

In nature, most mammals like the rats use their whiskers to interact with their environment through contact (Fig. 1). In [4], it was shown that rats rely on whiskers to maintain a cognitive understanding of their surroundings. To this end, this work describes our preliminary findings of extending the RatSLAM algorithm to account for both visual and tactile sensory modalities. The new algorithm will be hereby referred to as ViTa-SLAM. The novel aspect of this algorithm is that it utilizes multiple sensory modalities to estimate the state of the robot which are implicitly fused as opposed to other methods that require explicit weighted fusion like [5].

Fig. 1: Rat interacting with its surroundings while foraging.

II. ViTa-SLAM

ViTa-SLAM comprises of two sensory inputs: *visual* and *tactile*. Below, we elucidate how each of these sensory modalities are processed to obtain the best pose estimate.

This research has received funding from the European Union's Horizon 2020 Framework Programme for Research and Innovation under the Specific Grant Agreement No. 785907 (Human Brain Project SGA2).

Oliver Struckmeier, Kshitij Tiwari and Ville Kyrki are with the Department of Electrical Engineering and Automation, Aalto University, Espoo 02150, Finland {first.last}@aalto.fi

Martin J. Pearson is with the Bristol Robotics Laboratory, Bristol BS16 1QY, U.K (martin.pearson@brl.ac.uk)

A. Perceputal Data Preprocessing

Each visual frame is converted to greyscale (B/W img) and processed to obtain a **local view template** (V) similar to the vanilla RatSLAM [1].

The tactile data contains two kinds of information: *contact points* and *deflection*. *Contact points* $(Cts.)$ refers to the 3D point on the object surface in the world frame where the contact is made. These points are obtained by transforming the whisker contact points from a head centric frame to the world frame and are used to obtain a Point Feature Histogram (PFH) [6].

The *deflection* $(Defl.)$ refers to the amount of bending of the whiskers and is used to obtain the Slope Distribution Array (SDA) [7]. Together, they are referred to as **tactile features** (T). Previous attempts at pure whisker sensor based SLAM can be found in works like the WhiskerRatSLAM [8].

B. Overall System Architecture

The overall system architecture is shown in Fig. 2 wherein the pose cell network (PC) [9] combines the information V, T from both modalities to obtain the best pose estimate. A semi-metric experience map (Exp. Map) [10] is generated to evaluate the model performance.

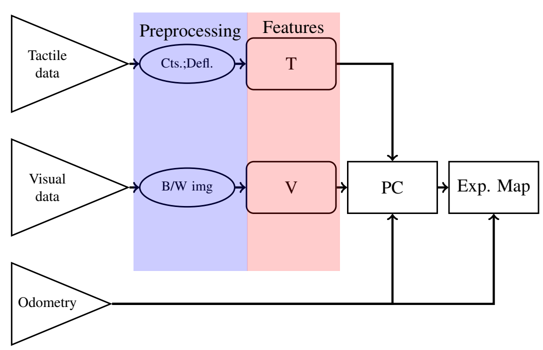

Fig. 2: Overview of Vita-SLAM architecture.

III. Evaluation

In this section, we describe the robot platform and the experimental scenario under which the model performance of ViTa-SLAM was evaluated. Following suit, the results obtained are described.

A. Robot Platform

The robot platform used for empirical evaluation is called the WhiskEye and is shown in Fig. 3. This robot is equipped with 24 whiskers which can be individually controlled and comes with a mobile base. The analog data from the whiskers is sampled at 500Hz. The tactile templates generated from the 3D whisker contact points are published once per whisk cycle.

The whiskers are mounted at the end of a neck which was kept fixed at a desired pose for this work. Aside from these, the platform is also equipped with a static HD (1280×720 pixels) RGB camera.

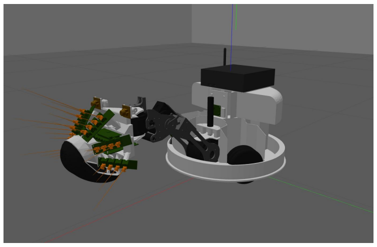

Fig. 3: Simulated robot platform, WhiskEye.

B. Environment

The robot is deployed in a visually sparse scene meaning there are not many diverse visual cues to be processed. Additionally, to break the 1-fold rotational symmetry of the rectangular environment being used, two static landmarks (cylinder and cube) are also placed in the scene.

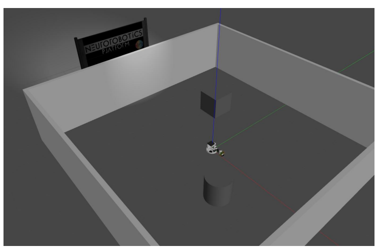

Fig. 4: Environmental setup. 2 landmarks in a visually sparse arena.

C. Behavior

The robot was given a pre-meditated trajectory to perform: It was required to reach landmark 1 (cylinder), revolve around it, then approach landmark 2 (cube), revolve around it and terminate exploration. At all times, the whiskers were being controlled using the Rapid Cessation of Protraction (RCP) protocol [11].

D. Results

In order to evaluate ViTa-SLAM performance, it was directly pitted against vanilla RatSLAM while keeping the platform and environment identical for both settings. As can be seen from Fig. 5, vanilla RatSLAM generated too many novel visual templates which eventually lead to failure of loop-closure detection. The frequent generation of novel templates can be attributed to visual sparsity of the scene. However, with ViTa-SLAM, additional tactile information helps detect loop closures[1] as shown in Fig. 6. There are however challenges with using the passive whisking behavior which leads the whisker array of the robot to collide with landmarks. This induces noise into the tactile data which adversely affects loop-closure detection.

IV. Conclusion and Future Works

In this work, we presented preliminary variant of our novel ViTa-SLAM algorithm which allows for multi-sensory SLAM

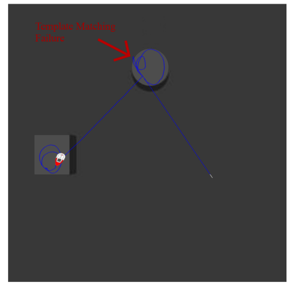

Fig. 5: RatSLAM failure.

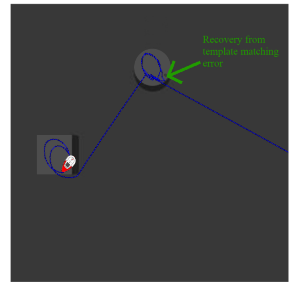

Fig. 6: ViTa-SLAM success.

whilst interacting with the environment through contact. Whilst the state-of-the-art bio-inspired SLAM model called RatSLAM was inspired by the rat hippocampal formations, it was designed purely for non-contact sensing scenarios. Similarly, WhiskerRatSLAM was designed purely for contact-sensing based SLAM. With this work, we have extended the outreach of these bio-inspired SLAM approaches to biomimetic robots bringing us one step closer to transitioning from biologically-inspired to biologically plausible methodologies.

In the future works, we plan on extending our algorithm to higher dimensions to account for the full 6 DOF pose while the algorithm currently can handle upto 3 DOF pose. While this poses significant computational challenges, it is essential to generalizing the applicability of this method. Additionally, we will investigate active sensory switch mechanism to minimize rudimentary sensory data acquisition.

References

[1] M. J. Milford, G. F. Wyeth, and D. Prasser, "Ratslam: a hippocampal model for simultaneous localization and mapping," in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA'04. 2004*, vol. 1, pp. 403–408, IEEE, 2004.

[2] R. Berkvens, A. Jacobson, M. Milford, H. Peremans, and M. Weyn, "Biologically inspired slam using wi-fi," in *2014 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1804–1811, IEEE, 2014.

[3] J. Steckel and H. Peremans, "Batslam: Simultaneous localization and mapping using biomimetic sonar," *PloS one*, vol. 8, no. 1, p. e54076, 2013.

[4] A. Cheung, D. Ball, M. Milford, G. Wyeth, and J. Wiles, "Maintaining a cognitive map in darkness: the need to fuse boundary knowledge with path integration," *PLoS computational biology*, vol. 8, no. 8, p. e1002651, 2012.

[5] M. J. Milford and A. Jacobson, "Brain-inspired sensor fusion for navigating robots," in *Robotics and Automation (ICRA), 2013 IEEE International Conference on*, pp. 2906–2913, IEEE, 2013.

[6] R. B. Rusu, Z. C. Marton, N. Blodow, and M. Beetz, "Learning informative point classes for the acquisition of object model maps," in *Control, Automation, Robotics and Vision, 2008. ICARCV 2008. 10th International Conference on*, pp. 643–650, IEEE, 2008.

[7] D. Kim and R. Möller, "Biomimetic whiskers for shape recognition," *Robotics and Autonomous Systems*, vol. 55, no. 3, pp. 229–243, 2007.

[8] M. Salman and M. J. Pearson, "Whisker-ratslam applied to 6d object identification and spatial localisation," in *Conference on Biomimetic and Biohybrid Systems*, pp. 403–414, Springer, 2018.

[9] D. Ball, S. Heath, J. Wiles, G. Wyeth, P. Corke, and M. Milford, "Openratslam: an open source brain-based slam system," *Autonomous Robots*, vol. 34, no. 3, pp. 149–176, 2013.

[10] M. Milford, D. Prasser, and G. Wyeth, "Experience mapping: producing spatially continuous environment representations using ratslam," in *Proceedings of Australasian Conference on Robotics and Automation 2005*, Australian Robotics and Automation Association Inc, 2005.

[11] R. A. Grant, B. Mitchinson, C. W. Fox, and T. J. Prescott, "Active touch sensing in the rat: anticipatory and regulatory control of whisker movements during surface exploration," *Journal of neurophysiology*, 2009.

[1] Video demonstration available here.