Qingbiao **Li**

15 JJ Thomson Ave, Cambridge, UK, CB3 0FD

+44 (0) 776-304-7574 | ql295@cam.ac.uk | qingbiaoli.github.io | qingbiaoli | Qingbiao Li

Publications

Journal Articles - Robotics & Machine Learning

Lifeng Zhou, Vishnu D Sharma, **Li, Qingbiao**, Amanda Prorok, Alejandro Ribeiro, Vijay Kumar. "Graph Neural Networks for Decentralized Multi-Robot Submodular Action Selection," *IEEE Robotics and Automation Letters* ***(JCR Q2, IF 3.74), Under Review***. 2021. **PDF**

Jiajun Cao∗, **Qingbiao Li**∗, Liping Xu, Rui Yang, Yuejin Dai. "Non-parametric Surrogate Model Method for Low-pressure Turbine Exhaust System," *Journal of Engineering for Gas Turbines and Power (**JCR Q3, IF 1.209**, Under Review)*. 2021

Qingbiao Li, Weizhe Lin, Zhe Liu, Amanda Prorok. "Message-Aware Graph Attention Networks for Large-Scale Multi-Robot Path Planning," *IEEE Robotics and Automation Letters* ***(JCR Q2, IF 3.74)***. 2020. **PDF**

Fernando Gama, **Qingbiao Li**, Ekaterina Tolstaya, Amanda Prorok, Alejandro Ribeiro. "Decentralized Control with Graph Neural Networks," *IEEE Transactions on Signal Processing* ***(JCR Q1, IF 4.931, Under Review)***. 2020. **PDF**

Binyu Wang, Zhe Liu, **Qingbiao Li**, Amanda Prorok. "Mobile Robot Path Planning in Dynamic Environments through Globally Guided Reinforcement Learning," *IEEE Robotics and Automation Letters* ***(JCR Q1, IF 4.931)*** pp. 6932–6939. 2020. **PDF**

Conference Proceedings - Robotics & Machine Learning

Jan Blumenkamp, Steven Morad, Jennifer Gielis, **Qingbiao Li**, Amanda Prorok. "A Framework for Real-World Multi-Robot Systems Running Decentralized GNN-Based Policies," *IEEE International Conference on Robotics and Automation* ***(CCF-B, Qualis-A1), Under Review***, 2021

Amanda Prorok, Jan Blumenkamp, **Qingbiao Li**, Ryan Kortvelesy, Zhe Liu, Ethan Stump. "The Holy Grail of Multi-Robot Planning: Learning to Generate Online-Scalable Solutions from Offline-Optimal Experts," *Conference on Robot Learning*. 2021. **PDF**

Jan Blumenkamp, **Qingbiao Li**, Amanda Prorok. "Evaluating the Sim-to-Real Gap of Graph Neural Network Policies for Multi-Robot Coordination," *IEEE International Conference on Robotics and Automation* ***(CCF-B, Qualis-A1)****, Real World Swarms Workshop*, 2021, **PDF**

Qingbiao Li, Fernando Gama, Alejandro Ribeiro, Amanda Prorok. "Graph Neural Networks for Decentralized Multi-robot Path Planning," *IEEE/RSJ International Conference on Intelligent Robots and Systems* ***(CCF-C, ERA-A, Qualis-A1)***, 2020, **PDF**

Qingbiao Li, Fernando Gama, Alejandro Ribeiro, Amanda Prorok. "Graph Neural Networks for Decentralized Path Planning," *International Conference on Autonomous Agents and MultiAgent Systems* ***(ERA-A, Qualis-A1)***, 2020, **PDF**

Qingbiao Li, Iordanis Chatzinikolaidis, Yiming Yang, Sethu Vijayakumar, Zhibin Li. "Robust Foot Placement Control for Dynamic Walking using Online Parameter Estimation," *IEEE-RAS 17th International Conference on Humanoid Robotics (Humanoids)* ***(ERA-C, Qualis-B2)***, 2017, **PDF**

Journal Articles - Computer Vision

Weizhe Lin, Indigo Orton, **Qingbiao Li**, Gabriela Pavarini, Marwa Mahmoud. "Looking At The Body: Automatic Analysis of Body Gestures and Self-Adaptors in Psychological Distress," *IEEE Transactions on Affective Computing* ***(JCR Q1, IF 10.506)***. Springer, 2020. **PDF**

Qingbiao Li, Jianyu Lin, Neil T Clancy, Daniel S Elson. "Estimation of Tissue Oxygen Saturation from RGB Images and Sparse Hyperspectral Signals based on Conditional Generative Adversarial Network," *International Journal of Computer Assisted Radiology and Surgery* ***(JCR Q3, IF 2.924)***. pp. 987–995. Springer, 2019. **PDF**

Conference Proceedings - Computer Vision

Ruoxi Wang, Dandan Zhang, **Qingbiao Li**, Xiao-Yun Zhou, Benny Lo. "Real-time Surgical Environment Enhancement for Robot-Assisted Minimally Invasive Surgery Based on Super-Resolution," *IEEE International Conference on Robotics and Automation* ***(CCF-B, Qualis-A1)***, 2021, **PDF**

Qingbiao Li, Xiao-Yun Zhou, Jianyu Lin, Jian-Qing Zheng, Neil T Clancy, Daniel S Elson. "Estimation of Tissue Oxygen Saturation from RGB Images based on Pixel-level Image Translation," *The Hamlyn Symposium on Medical Robotics*, 2018, **PDF**

Graph Neural Networks for Decentralized Multi-Robot Submodular Action Selection

Lifeng Zhou,[1⋆] Vishnu D. Sharma,[2⋆] Qingbiao Li,[3] Amanda Prorok,[3] Alejandro Ribeiro,[1] Vijay Kumar[1]

Abstract— **In this paper, we develop a learning-based approach for decentralized submodular maximization. We focus on applications where robots are required to jointly select actions, e.g., motion primitives, to maximize team submodular objectives with local communications only. Such applications are essential for large-scale multi-robot coordination such as multi-robot motion planning for area coverage, environment exploration, and target tracking. But the current decentralized submodular maximization algorithms either require assumptions on the inter-robot communication or lose some suboptimal guarantees. In this work, we propose a general-purpose learning architecture towards submodular maximization at scale, with decentralized communications. Particularly, our learning architecture leverages a graph neural network (GNN) to capture local interactions of the robots and learns decentralized decision-making for the robots. We train the learning model by imitating an expert solution and implement the resulting model for decentralized action selection involving local observations and communications only. We demonstrate the performance of our GNN-based learning approach in a scenario of active target coverage with large networks of robots. The simulation results show our approach nearly matches the coverage performance of the expert algorithm, and yet runs several orders faster with more than 30 robots. The results also exhibit our approach's generalization capability in previously unseen scenarios, e.g., larger environments and larger networks of robots.**

I. Introduction

Submodular maximization problems find a wealth of applications in robotics. Typical examples include environmental monitoring and surveillance [1], [2], target coverage and tracking [3], [4], [5], search and rescue [6], and informative path planning [7]. Such applications ask for teams of robots that act as mobile sensors to jointly plan their actions to optimize submoudlar objective functions. Submodular functions have the diminishing returns property. Examples of such functions include the information-theoretic metrics such as entropy and mutual information [1] and the geometric metrics such as the visibility region [8]. The problems of maximizing submodular functions are generally NP-hard. The most well-known approach for tackling these problems is the greedy algorithm that runs in polynomial time and yields a constant factor approximation guarantee [9], [10].

The greedy algorithm cannot be directly implemented in the scenarios where the robots can only communicate locally due to a limited communication range. To address the issue of local communication, some decentralized versions of the greedy algorithm were designed, where only neighboring information is utilized to choose actions for the robots for optimizing submodular objectives [11], [12], [13], [14]. However, these algorithms either assume a connected communication graph [11], [12] or typically have a worse suboptimality bound than that of the greedy algorithm [10] in the scenarios with limited communication [13], [14].

In this paper, we aim to explore learning-based methods that approach the performance of the greedy algorithm [10]. The graph neural network (GNN) is chosen as the learning method given its nice properties of decentralized communication architecture that captures the neighboring interactions and the transferability that allows for the generalization to previously unseen scenarios [15], [16]. Also, GNN has recently shown success in various multi-robot applications such as formation control [17], [18], path finding [19], and task assignment [20]. To this end, we design a GNN-based learning network that enables robots to communicate and share information with neighbors and selects actions for the robots. We train such a learning network to perform as close as possible to the greedy algorithm by imitating the behavior of the greedy algorithm.

Related Work. Researchers have developed several decentralized/distributed algorithms for tackling submodular maximization problems. For example, building on the local greedy algorithm [10, Section 4], Atanasov et al. designed a decentralized greedy algorithm that achieves 1/2 approximation bound for maximizing submodular objectives [11]. Specifically, the algorithm greedily selects an action for each robot in a sequential order, given all the actions selected so far. A speed-up of the sequential greedy algorithm was developed by Micah and Michael to select exploration actions for mobile robots [21]. However, with limited communication, the robots may not have access to all the previously selected actions. To this end, a few distributed submodular maximization algorithms were devised to execute the sequential greedy algorithm over directed acyclic graphs that may not be connected [13], [14]. The suboptimality bound of these distributed algorithms depends on the property of the communication graph and is typically worse than $1/2$ [13], [14]. Particularly, the worst-case analysis parallels results in [22] where a distributed submodular maximization algorithm is developed for data selection. Other decentralized greedy approaches include the ones that utilize the consensus-based

⋆These authors contributed equally.

[1]L. Zhou, A. Ribeiro, and V. Kumar are with the GRASP Laboratory, University of Pennsylvania, Philadelphia, PA, USA (email: {lfzhou, aribeiro, kumar}@seas.upenn.edu).

[2]V. D. Sharma is with the Department of Computer Science, University of Maryland, College Park, MD, USA (email: vishnuds@umd.edu).

[3]Q. Li and A. Prorok are with the Department of Computer Science and Technology, University of Cambridge, Cambridge, United Kingdom (email: {ql295, asp45}@cam.ac.uk)).

This research was sponsored by the Army Research Lab through ARL DCIST CRA W911NF-17-2-0181.

Non-parametric Surrogate Model Method based on Machine Learning

Jiajun Cao*[a], Qingbiao Li*[b], Liping Xu[a], Rui Yang[c], Yuejin Dai[c]

[a] *Whittle Laboratory, Department of Engineering, University of Cambridge, Cambridge, United Kingdom CB3 0DY*
[b] *Computer Lab, Department of Computer Science and Technology, University of Cambridge, Cambridge, United Kingdom CB3 0FD*
[c] *Shanghai Turbine Works Co., Ltd. Shanghai, China 200240*

Abstract

In this paper, a novel "non-parametric" surrogate model method is introduced. The new method extracts geometric information from the surface mesh of the simulation domain using Graph Neural Networks (GNNs) and predicts the two-dimensional distributions of flow variables (in forms of contour maps) using Convolutional Neural Networks (CNNs). This method can automatically extract relevant geometric information from surface mesh, while existing data-driven surrogate model methods need manual parameterization, which may introduce additional uncertainties. Existing methods can only process geometries defined by their specific parameterization methods because the inputs of existing surrogate models are human-defined geometric parameters, while new methods can process any geometries with the same topology because its input is the surface mesh. This allows users to access more design variations from different sources to create a larger database. In addition, this novel surrogate model method is able to predict the distributions of variables, not only the integrated values of performance. This paper demonstrates this novel surrogate model method with its application on optimization of a low-pressure steam turbine exhaust system (LPES). The new surrogate model uses 10 surface meshes of the LPES as input

Email addresses: jc980@cam.ac.uk (Jiajun Cao*), ql295@cam.ac.uk (Qingbiao Li*), lpx1@cam.ac.uk (Liping Xu), yangrui@shanghai-electric.com (Rui Yang), daiyj3@shanghai-electric.com (Yuejin Dai)

*Jiajun Cao and Qingbiao Li contribute equally to this work.

Message-Aware Graph Attention Networks for Large-Scale Multi-Robot Path Planning

Qingbiao Li[1,*], Weizhe Lin[2,*], Zhe Liu[1], Amanda Prorok[1], *Member, IEEE*

***Abstract*—The domains of transport and logistics are increasingly relying on autonomous mobile robots for the handling and distribution of passengers or resources. At large system scales, finding *decentralized* path planning and coordination solutions is key to efficient system performance. Recently, Graph Neural Networks (GNNs) have become popular due to their ability to learn communication policies in decentralized multi-agent systems. Yet, vanilla GNNs rely on simplistic message aggregation mechanisms that prevent agents from prioritizing important information. To tackle this challenge, in this paper, we extend our previous work that utilizes GNNs in multi-agent path planning by incorporating a novel mechanism to allow for *message-dependent attention*. Our Message-Aware Graph Attention neTwork (MAGAT) is based on a key-query-like mechanism that determines the relative importance of features in the messages received from various neighboring robots. We show that MAGAT is able to achieve a performance close to that of a coupled centralized expert algorithm. Further, ablation studies and comparisons to several benchmark models show that our attention mechanism is very effective across different robot densities and performs stably in different constraints in communication bandwidth. Experiments demonstrate that our model is able to generalize well in previously unseen problem instances, and that it achieves a 47% improvement over the benchmark success rate, even in very large-scale instances that are $\times$100 larger than the training instances.**

I. Introduction

REMARKABLE progress has been achieved for Multi-Robot Path Planing [1]–[4]—a problem that considers the generation of collision-free paths leading robots from their start positions to designated goal positions. In recent years, solutions to this problem have become increasingly important for item retrieval in warehouses [5] and mobility-on-demand services [6].

However, computing a solution that can balance between optimality and real-world efficiency remains a challenge. Current approaches can be mainly classified into *centralized* and *decentralized*. Centralized methods require central units to gather information from all robots and organize the optimal path for each of them, consuming large computational resources. As the system scales, decentralized approaches become increasingly popular, where each robot estimates or communicates others' future trajectories via broadcasting or distance-based communication. Unfortunately, if the communication happens concurrently and equivalently among many neighboring robots, it is likely to cause redundant communication, burden the computational capacity and adversely affect overall team performance. Besides, robust and continuous communication cannot yet be guaranteed due to limited bandwidth, large data volumes, and interference from the surroundings. Additionally, under a fully decentralized framework without any priority of planning, it is very hard to ensure the convergence of the negotiation process [7]. These limitations ultimately affect the optimality of solutions found and the overall resilience of the team to disruptions. Hence, new trends of research focus on ***communication-aware path planning approaches*** by explicitly considering communication efficiency during path generation and path optimization [8], addressing to whom the information is communicated, and at what time [9].

Contributions. In this work, we propose to use a Message Aware Graph Attention neTwork (MAGAT) to extend our previous decentralized framework [4]. Our contributions are summarized as follows:

- We combine a Graph Neural Network (GNN) with a key-query-like attention mechanism to improve the effectiveness of inter-robot communication. We demonstrate the suitability of applying our model on *dynamic communication graphs* by proving its permutation equivariance and time invariance property.
- We investigate the impact of reduced communication bandwidth by reducing the size of the shared features, and then deploy a skip-connection to preserve self-information and maintain model performance.
- We demonstrate the generalizability of our model by training the model on small problem instances and testing it on increasing robot density, varying map size, and much larger problem instances (up to $\times \mathbf{100}$ the number of robots). Our proposed model is shown to be more efficient in learning general knowledge of path planning as it achieves better generalization performance than the baseline systems under various scenarios.

This paper was recommended for publication by Editor-in-Chief Allison Okamura upon evaluation of the Associate Editor and Reviewers' comments. We gratefully acknowledge the support of ARL grant DCIST CRA W911NF-17-2-0181, and a gift through Amazon.com Inc. Z. Liu was supported by the Engineering and Physical Sciences Research Council (grant EP/S015493/1). A. Prorok was supported by the EPSRC (grant EP/S015493/1) and ERC Project 949940 (gAIa). *Corresponding author: Qingbiao Li.*

*Qingbiao Li and Weizhe Lin contributed equally to this work. We also thank Fernando Gama, Binyu Wang and Alex Raymond for their helpful discussions.

[1]Qingbiao Li, Zhe Liu and Amanda Prorok are with Department of Computer Science and Technology, University of Cambridge. (e-mail:{ql295,zl457,asp45}@cam.ac.uk).

[2]Weizhe Lin is with Department of Engineering, University of Cambridge. (e-mail: wl356@cam.ac.uk).

II. Related Work

Learning-based methods have been actively investigated in recent years, and have demonstrated their strengths in designing robot control policies for an increasing number of tasks [10]. These methods have shown their capabilities to offload the online computational burden into an offline learning procedure, which allows agents to act independently based on the learned knowledge, and thus to work in a decentralized manner [3], [11]. Sartoretti et al. [3] have proposed a hybrid learning-based method called PRIMAL for multi-agent path-finding that integrated imitation learning (via

Decentralized Control with Graph Neural Networks

Fernando Gama, Qingbiao Li, Ekaterina Tolstaya, Amanda Prorok, and Alejandro Ribeiro

***Abstract*—Dynamical systems consisting of a set of autonomous agents face the challenge of having to accomplish a global task, relying only on local information. While centralized controllers are readily available, they face limitations in terms of scalability and implementation, as they do not respect the distributed information structure imposed by the network system of agents. Given the difficulties in finding optimal decentralized controllers, we propose a novel framework using graph neural networks (GNNs) to *learn* these controllers. GNNs are well-suited for the task since they are naturally distributed architectures and exhibit good scalability and transferability properties. The problems of flocking and multi-agent path planning are explored to illustrate the potential of GNNs in learning decentralized controllers.**

***Index Terms*—decentralized control, graph neural networks, graph signal processing, flocking, path planning**

I. Introduction

Network dynamical systems are widespread, spanning applications in multiagent robotics [2], [3], smart grids [4], [5], sensor networks [6], wireless communications [7], [8] and traffic control [9]. In all of these situations we encounter teams of autonomous agents that sense their local environment and exchange information with nearby agents; which then proceed to control their individual actions in pursuit of a common global goal [10, Ch. 3]. This somewhat mismatched specification of attaining a *global* goal with *local* sensing and interaction is the defining characteristic of a network dynamical system.

A possible approach to controlling an autonomous team is to collect information at a designated fusion center that decides on actions that the team executes. Such *centralized* controllers for network systems are common – see, e.g., [3], [4], [7], [8] – and their implementation is preferred when feasible. Their advantage is that if communication delays are not significant, the resulting control problem is standard and we can deploy established techniques to design optimal controllers (Sec. II). Their disadvantage is that collecting sensor inputs and disseminating control actions, burdens communication networks. As we scale the number of agents in the team, communication delays ensue. This effectively limits applicability of centralized controllers to small teams.

Scalability to teams with large numbers of agents is more feasible with *decentralized* controllers, in which agents decide on their own actions. This approach engenders scalability by design, but results in optimal controllers that are famously difficult to design [11]. The reason why this happens is that agents that have access to local information, are also agents that have access to different information. This results in the possibility of conflicting actions even for agents that are intent on cooperating (Sec. II-A). If optimal decentralized controllers are unavailable, resorting to heuristics is warranted. In this paper we advocate the use of learned heuristics.

The main contribution of this paper is to develop the use of graph neural networks (GNNs) [12], [13] to learn decentralized controllers (Sec. III). More concretely, we model agent interactions with a communication graph and interpret sensing inputs and control actions as signals supported on the nodes of this graph. We then proceed to use graph convolutional neural networks (GCNNs) [14]–[16] and graph recurrent neural networks (GRNNs) [17], [18] to learn maps from sensor inputs to control outputs. The maps are trained with imitation learning [19]–[23] of expert –ideally but not necessarily optimal– centralized controllers (Sec. III-D).

The key to explain GNNs and their applicability in decentralized control is the notion of a graph filter [24]–[26]. Graph filters are generalizations of Euclidean convolutional filters and are formally defined as polynomials on matrix representations of the graph (Sec. III-A). Given that the sparsity pattern on these matrix representations matches agent connectivity, graph filters can be implemented in a distributed manner. GCNNs rely on graph filter banks composed with *pointwise* nonlinearities (Sec. III-B). Given that pointwise operations are local operations, GCNNs also admit distributed implementations. GRNNs also rely on graph filters and pointwise nonlinearities but differ on the incorporation of a time varying hidden state (Sec. III-C). This makes them more appealing in time varying partially observable problems, but does not affect the admissibility of decentralized implementations.

GNNs admit decentralized implementations by construction and are therefore viable parameterizations to learn decentralized controllers. That they can succeed is a matter of experimental evaluation. Here, we consider flocking (Sec. VI) and path planning (Sec. VII) and deploy GCNNs and GRNNs that are trained to imitate expert centralized controllers. Our first conclusion is that imitation succeeds in both problems over a range of realistic system parameters:

- **(C1) Successful Imitation.** GCNNs and GRNNs are decentralized architectures relying on local information exchanges with neighboring agents. They nevertheless successfully imitate centralized policies that rely on global information.

It is important to remark that in imitation learning we rely on expert centralized controllers for *offline* training of a GNN that we deploy *online* for decentralized control. In this process of transference between systems observed at training to those observed at deployment, networks are likely to change. Asides from the admissibility of decentralized implementa-

The authors are supported by ARL DCIST CRA W911NF-17-2-0181. Q. Li and A. Prorok are also supported by EPSRC grant EP/S015493/1.
F. Gama is with the Elect. Eng. Comput. Sci. Dept., Univ. of California, Berkeley, E. Tolstaya and A. Ribeiro are with the Dept. Elect. Syst. Eng., Univ. of Pennsylvania., and Q. Li and A. Prorok are with the Dept. Comput. Sci. Technol., Cambridge University. Email: fgama@berkeley.edu, {eig,aribeiro}@seas.upenn.edu, and {ql295,asp45}@cam.ac.uk. Preliminary discussions appear in [1], [2].

Mobile Robot Path Planning in Dynamic Environments through Globally Guided Reinforcement Learning

Binyu Wang, Zhe Liu, Qingbiao Li, Amanda Prorok, *Member, IEEE*

***Abstract*—Path planning for mobile robots in large dynamic environments is a challenging problem, as the robots are required to efficiently reach their given goals while simultaneously avoiding potential conflicts with other robots or dynamic objects. In the presence of dynamic obstacles, traditional solutions usually employ re-planning strategies, which re-call a planning algorithm to search for an alternative path whenever the robot encounters a conflict. However, such re-planning strategies often cause unnecessary detours. To address this issue, we propose a learning-based technique that exploits environmental spatio-temporal information. Different from existing learning-based methods, we introduce a *globally guided reinforcement learning* approach (G2RL), which incorporates a novel reward structure that generalizes to arbitrary environments. We apply G2RL to solve the multi-robot path planning problem in a fully distributed reactive manner. We evaluate our method across different map types, obstacle densities, and the number of robots. Experimental results show that G2RL generalizes well, outperforming existing distributed methods, and performing very similarly to fully centralized state-of-the-art benchmarks.**

***Index Terms*—Hierarchical path planning, mobile robots, reinforcement learning, scalability.**

I. Introduction

PATH planning is one of the fundamental problems in robotics. It can be formulated as: given a robot and a description of the environment, plan a conflict-free path between the specified start and goal locations. Traditionally, there are two different versions: off-line planning, which assumes static obstacles and perfectly known environments, and on-line planning, which focuses on dealing with dynamic obstacles and partially known environments [1]. Traditional off-line planning algorithms [2] cannot be directly utilized for solving on-line path planning tasks as they assume that the obstacles are static. One strategy is to plan an initial path and invoke re-planning whenever its execution becomes infeasible [3]. However, re-planning will suffer from time inefficiency (frequent re-planning) and path inefficiency (oscillating movements and detours) due to the absence of motion information of dynamic obstacles. Furthermore, re-planning strategies may fail in the presence of robot deadlocks. Instead of re-planning, some methods include an extra time dimension to avoid potential conflicts [4]. However, this approach increases the number of states to be searched, and additionally requires the knowledge of the future trajectories of dynamic obstacles. If the future movements of obstacles are unknown, one may attempt to model the behavior of dynamic obstacles and predict their paths [5]. Yet, separating the navigation problem into disjoint prediction and planning steps can lead to the 'freezing robot' problem. In that case, the robot will fail to find any feasible action as the predicted paths could mark a large portion of the space as untraversable.

Recently, learning-based approaches have been studied to address on-line planning in dynamic environments [6], [7]. This is popularized by the seminal work [8], which utilized deep neural networks for the function estimation of value-based reinforcement learning (RL). Although RL has demonstrated outstanding performance in many applications, several challenges still impede its contribution to the path planning problem. First of all, when the environment is extremely large, the reward becomes sparse, inducing an increased training effort and making the overall learning process inefficient [9]. Another challenge is the over-fitting issue. The robot is often limited to training environments and shows poor generalizability to unseen environments [10]. Most recent approaches still show difficulties in scaling to arbitrarily large multi-robot systems [7], as the sizes of the robot state, joint action, and joint observation spaces grow exponentially with the number of robots [11]. Thus, the efficiency, generalizability, and scalability of existing RL-based planners can still not fulfill the requirements of many applications.

In order to overcome the above challenges, we develop a hierarchical path-planning algorithm that combines a *global guidance* and a *local RL-based planner*. Concretely, we first utilize a global path planning algorithm (for example, A*) to obtain a globally optimal path, which we refer to as the *global guidance*. During robot motion, the *local RL-based planner* generates robot actions by exploiting surrounding environmental information to avoid conflicts with static and dynamic obstacles, while simultaneously attempting to follow the fixed global guidance. Our main contributions include:

- We present a hierarchical framework that combines global guidance and local RL-based planning to enable end-to-end learning in dynamic environments. The local RL planner exploits both spatial and temporal information within a local area (e.g., a field of view) to avoid potential collisions and unnecessary detours. Introducing global guidance allows the

This work was supported by the Engineering and Physical Sciences Research Council (grant EP/S015493/1), and ARL DCIST CRA W911NF-17-2-0181.

B. Wang is with the Department of Mechanical and Automation Engineering, The Chinese University of Hong Kong, Hong Kong (e-mail: zbwby819@sina.com). Z. Liu, Q. Li and A. Prorok are with the Department of Computer Science and Technology, University of Cambridge, Cambridge CB3 0FD, U.K. (e-mail: {zl457, ql295, asp45}@cam.ac.uk). *Corresponding author: Z. Liu.*

A Framework for Real-World Multi-Robot Systems Running Decentralized GNN-Based Policies

Jan Blumenkamp, Steven Morad, Jennifer Gielis, Qingbiao Li and Amanda Prorok

***Abstract*— Graph Neural Networks (GNNs) are a paradigm-shifting neural architecture to facilitate the learning of complex multi-agent behaviors. Recent work has demonstrated remarkable performance in tasks such as flocking, multi-agent path planning and cooperative coverage. However, the policies derived through GNN-based learning schemes have not yet been deployed to the *real-world* on physical multi-robot systems. In this work, we present the design of a system that allows for fully decentralized execution of GNN-based policies. We provide an open-source framework[1] based on ROS2 and elaborate its details in this paper. We demonstrate our framework on a case-study that requires tight coordination between robots, and present first-of-a-kind results that show successful real-world deployment of GNN-based policies on a decentralized multi-robot system relying on *Adhoc* communication.**

***Index Terms*— Multi-Robot Systems, Graph Neural Network, Robot Learning, Sim-to-Real**

I. Introduction

Researchers have invested significant effort into finding analytical solutions to multi-robot problems, balancing optimality, completeness, and computational efficiency [1], [2], [3], [4]. Recently, data-driven approaches have provided alternatives for some multi-robot problems. Data-driven approaches can find near-optimal solutions to NP-hard problems, which enables fast on-line planning and coordination that is typically required in robotics [5], [6], [7], [8]. Graph Neural Networks (GNNs), in particular, demonstrate remarkable performance and generalize well to large-scale robotic teams for various tasks such as flocking, navigation, and control [9], [6], [10], [11], [12], [13]. In such multi-robot systems, GNNs learn inter-robot communication strategies using latent messages. Individual robots aggregate these messages from their neighbors to overcome inherently local (partial) knowledge and build a more complete understanding of the world they are operating in.

While GNN-based policies are typically trained in *simulation* and in a centralized manner, and therefore assume synchronous communication, resulting policies can be executed either in a centralized or decentralized mode. Evaluating a GNN in the *centralized* mode typically requires execution on a single machine decoupled from the robots that are acting according to the policy [6], [10], [14]. This *(i)* introduces a single point of failure, *(ii)* requires all robots to maintain constant network connectivity, and *(iii)* introduces scalability issues due to computational complexity $\mathcal{O}(N^2)$ where N is the number of robots. In contrast, in the *decentralized* mode, each robot is responsible for making its own decisions. With fully decentralized evaluation, *(i)* there is no single point of failure, resulting in a higher fault tolerance, *(ii)* agents do not need to remain in network range of a router that orchestrates the centralized evaluation, and *(iii)* computation is parallelized across N robots, each with time complexity $\mathcal{O}(N)$, decoupling the time complexity from the number of robots.

All authors are with Department of Computer Science and Technology, University of Cambridge. Email: {jb2270, sm2558, jag233, ql295, asp45}@cam.ac.uk).

[1]https://github.com/proroklab/ros2-gnn

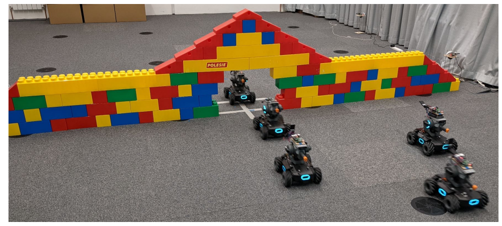


Fig. 1. Five RoboMaster robots need to navigate through a narrow passageway to reconfigure on the other side, as quickly as possible. The robots leverage communication to overcome partial observability of the workspace. To solve this task, we deploy GNN-based policies.

Even though GNNs have an inherently decentralizable mathematical formulation, previous work on GNNs-based multi-robot policies was conducted exclusively in centralized simulations using synchronous communication [10], [11], [9]. For practical reasons, decentralized execution is often unavoidable in the *real-world*, but it is currently unknown whether this contributes to a shift of domains, and how resulting policies are affected. Multi-robot GNNs require inter-robot communication, but real-world wireless communication is noisy, and messages can be lost or delayed, leading to significant performance loss—this is exemplified in prior work that demonstrates the need for appropriate models to overcome these challenges [15], [16], [17], [18]. Further compounding these issues, decentralized policies are typically executed asynchronously, resulting in system states not previously encountered during training.

In this paper, we provide a framework that facilitates the decentralized execution of GNN-based multi-robot policies. We present the results of a suite of real-robot experiments (see Fig. 1) to demonstrate the consequences of this decentralized execution. To that end, we introduce a taxonomy of different evaluation modes and networking configurations. Specifically, we contribute:

1) A ROS2-based software and networking framework for GNNs and other message-passing algorithms to facilitate operation in both simulation and the real-world, and to permit GNN execution in either a centralized or decentralized manner.
2) An ablation study on several forms of execution to quantify performance shifts between centralized execution and three forms of decentralized policy execution, *(i)* offboard (non-local), *(ii)* onboard over routing infrastructure, and *(iii)* onboard with *Adhoc* networking.
3) An open-sourced implementation of our framework[1].

The Holy Grail of Multi-Robot Planning: Learning to Generate Online-Scalable Solutions from Offline-Optimal Experts

Amanda Prorok* **Jan Blumenkamp*** **Qingbiao Li*** **Ryan Kortvelesy*** **Zhe Liu***
Department of Computer Science and Technology
University of Cambridge, UK
{asp45,jb2270,ql295,rk627,zl457}@cam.ac.uk

Ethan Stump*
DEVCOM Army Research Laboratory (ARL), Maryland, USA.
ethan.a.stump2.civ@mail.mil

Abstract: Many multi-robot planning problems are burdened by the curse of dimensionality, which compounds the difficulty of applying solutions to large-scale problem instances. The use of learning-based methods in multi-robot planning holds great promise as it enables us to offload the *online* computational burden of expensive, yet optimal solvers, to an *offline* learning procedure. Simply put, the idea is to train a policy to copy an optimal pattern generated by a small-scale system, and then transfer that policy to much larger systems, in the hope that the learned strategy scales, while maintaining near-optimal performance. Yet, a number of issues impede us from leveraging this idea to its full potential. This blue-sky paper elaborates some of the key challenges that remain.

Keywords: Multi-Robot Planning, Imitation Learning

1 Introduction

Learning-based methods have proven effective at designing robot control policies for an increasing number of tasks [1, 2]. The application of learning-based methods to multi-robot planning has attracted particular attention due to their capability of handling high-dimensional joint state-space representations, by offloading the online computational burden to an offline learning procedure [3, 4]. We argue that these developments point to a fundamental approach that combines ideas around the application of learning to optimization and produce a flexible framework that could tackle many hard but important problems in robotics, including multi-agent path planning [5], area coverage [6, 7], task allocation [8, 9, 10], formation control [11], and target-tracking [12]. In this paper, we motivate this approach and discuss the crucial challenges and research questions.

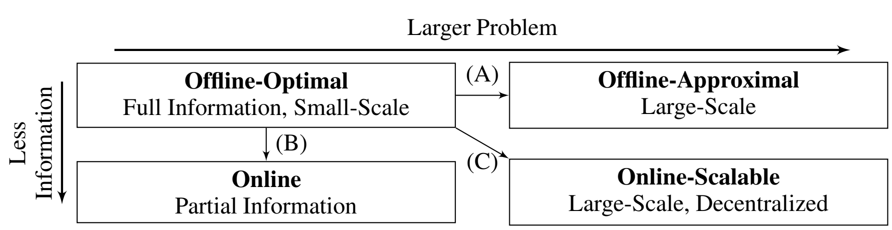


Figure 1: Applications of learning to optimization problems. (A) embodies techniques for learning optimization heuristics; (B) embodies techniques for learning to solve POMDPs; (C) is the emerging topic discussed here, embodying techniques for learning to coordinate large systems in real-world applications

*All authors contributed equally.

Evaluating the Sim-to-Real Gap of Graph Neural Network Policies for Multi-Robot Coordination

Jan Blumenkamp, Qingbiao Li and Amanda Prorok

***Abstract*— Graph Neural Networks (GNNs) are a paradigm-shifting neural architecture that facilitates the learning of complex multi-agent behaviors. Recent work has demonstrated remarkable performance in tasks such as multi-agent path planning and cooperative coverage. However, the policies derived through GNN-based learning schemes have not yet been validated in the real world on physical multi-robot systems. In this paper, we report pioneering experiments that evaluate the performance of GNN-based policies on a multi-robot coordination task that requires a team of five robots to navigate through a narrow passageway. We demonstrate that our policy, which is trained using Reinforcement Learning (RL) in simulation, is capable of being deployed to the real world with a minor drop in performance compared to simulation. These experiments promise to be the foundation for future experiments, which aim to identify sim-to-real challenges in GNNs. To that end, we investigate and characterize the impact of the communication radius and random communication message dropouts to evaluate the robustness of our framework.**

***Index Terms*— Multi-Robot Systems, Robot Learning, Sim-to-Real**

I. Introduction

Efficient and collision-free navigation in multi-robot systems is fundamental to advancing mobility in guiding robots from their origins to designated destinations, for example in industrial plant inspection, item retrieval in warehouses [1], and transportation with self-driving cars in smart cities [2]. Finding an optimal solution to such problems is often not tractable in polynomial time.

Recent work is actively investigating data-driven decentralized methods to approach optimal but costly algorithms by offloading the online computational burden to an offline learning procedure [3], [4], [5]. Particularly, Graph Neural Networks (GNNs) have demonstrated remarkable performance results in simulation and have been shown to generalize well to large-scale robotic teams [6], [5], [7]. However, training neural networks on real robots is a time-consuming and costly process, which yields the need to explore solutions to bridge the gap from simulation into the real world (sim-to-real) using approaches such as zero-shot transfer [8], domain adaptation [9] and domain randomization [10].

Most recent works in the field of sim-to-real have focused on the field of robot vision [11], [12], [13]. Yet, to the best of our knowledge, no work has been conducted on sim-to-real for GNNs. Current GNN models are usually trained with the assumption of synchronous communication, which is hard to ensure in real-world settings. Other challenges include message dropouts, communication delays, and noise. A clear performance metric is essential to investigate the sim-to-real gap, which is difficult in other common multi-agent problems such as flocking. We suggest the following steps to tackle this problem: *(i)* First, a suitable scenario with a clear metric to evaluate the performance of deployed policies in both simulation and in reality has to be designed. *(ii)* Next, the parameters that contribute to the sim-to-real gap, but are not necessarily caused by communication, have to be isolated. *(iii)* After all variables that contribute to the sim-to-real gap but are not part of the communication component are isolated, the impact on changing parameters related to the communication has to be investigated and evaluated. *(iv)* Lastly, methods to minimize the sim-to-real gap can be developed based on these results. In this work, we report the intermediate results after step *(ii)*. In a future work, we intend to extend this work to utilize truly decentralized real-world communication to further investigate the sim-to-real gap for GNNs.

All authors are with Department of Computer Science and Technology, University of Cambridge. (e-mail: {jb2270, ql295, asp45}@cam.ac.uk).

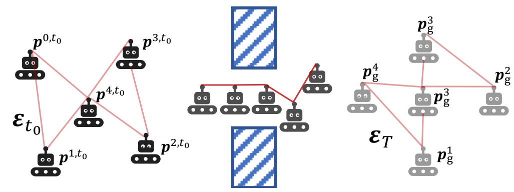


Fig. 1. A robot formation has to reconfigure to move through a narrow passage to the other side of the wall as quickly as possible.

II. Related Work

Within the field of multi-robot navigation, formation control and reconfiguration remain challenging in complex, yet static, environments. For example, robots need to reconfigure the formation to pass through a narrow corridor and rejoin as a team after the passageway. Centralized approaches [14] can be deployed to compute a global and optimal path with a global knowledge of positions, destinations, and environment maps, or by setting up intermediate collision-free configurations for the team of robots. However, this method fails in scaling up to larger teams. *Decentralized* approaches provide an attractive alternative to centralized approaches, because they can reduce the computational overhead, and relax the dependence on centralized units. In prioritized planning [15], [16], each robot is assigned a unique priority, and the algorithm proceeds sequentially from the highest priority robot to the lowest priority one. However, these methods can only obtain sub-optimal solutions.

Graph Neural Networks for Decentralized Multi-Robot Path Planning

Qingbiao Li[1], Fernando Gama[2], Alejandro Ribeiro[2], Amanda Prorok [1]

ABSTRACT

Effective communication is key to successful, decentralized, multi-robot path planning. Yet, it is far from obvious what information is crucial to the task at hand, and how and when it must be shared among robots. To side-step these issues and move beyond hand-crafted heuristics, we propose a combined model that automatically synthesizes local communication and decision-making policies for robots navigating in constrained workspaces. Our architecture is composed of a convolutional neural network (CNN) that extracts adequate features from local observations, and a graph neural network (GNN) that communicates these features among robots. We train the model to imitate an expert algorithm, and use the resulting model online in decentralized planning involving only local communication and local observations. We evaluate our method in simulations by navigating teams of robots to their destinations in 2D cluttered workspaces. We measure the success rates and sum of costs over the planned paths. The results show a performance close to that of our expert algorithm, demonstrating the validity of our approach. In particular, we show our model's capability to generalize to previously unseen cases (involving larger environments and larger robot teams).

KEYWORDS

Multi-Agent Path Finding; Decentralized Planning; Deep Learning; Graph Neural Networks;

1 INTRODUCTION

Efficient and collision-free navigation in multi-robot systems is fundamental to advancing mobility. The problem, generally referred to as Multi-Robot Path Planning (MRPP) or Multi-Agent Path Finding (MAPF), aims at generating collision-free paths leading robots from their origins to designated destinations. Current approaches can be classified as either *coupled* or *decoupled*, depending on the structure of the state space that is searched. While coupled approaches are able to ensure the optimality and completeness of the solution, they involve *centralized* components, and tend to scale poorly with the number of robots [22, 23]. Decoupled approaches, on the other hand, compute trajectories for each robot separately, and re-plan only in case of conflicts [26, 27, 30]. This can significantly reduce the computational complexity of the planning task, but generally produces sub-optimal and incomplete solutions. Balancing optimality and completeness with the complexity of computing a solution, however, is still an open research problem [1, 19].

This work focuses on multi-robot path planning for scenarios where the robots are restricted in observation and communication range, and possess no global reference frame for localization. This naturally arises when considering physical robots equipped with hardware constraints that limit their perception and communication capabilities [12]. These scenarios impose a decentralized structure, where at any given point in time, robots have only partial information of the system state. In this paper, we propose a combined architecture, where we train a convolutional neural network (CNN) [11] that extracts adequate features from local observations, and a graph neural network (GNN) to communicate these features among robots [8] with the ultimate goal of learning a decentralized sequential action policy that yields efficient path plans for all robots. The GNN implementation seamlessly adapts to the partial information structure of the problem, since it is computed in a decentralized manner. We train this architecture to imitate an optimal coupled planner with global information that is available offline at training time. Further, we develop a dataset aggregation method that leverages an online expert to resolve hard cases, thus expediting the learning process. The resulting trained model is used online in an efficient, decentralized manner, involving communication only with nearby robots. Furthermore, we show that the model can be deployed on much larger robot teams than the ones it was trained on.

2 RELATED WORK AND CONTRIBUTION

Related work. Classical approaches to multi-robot path planning can generally be described as either centralized or decentralized. *Centralized* approaches are facilitated by a planning unit that monitors all robots' positions and desired destinations, and returns a coordinated plan of trajectories (or way-points) for all the robots in the system. These plans are communicated to the respective robots, which use them for real-time on-board control of their navigation behavior. Coupled centralized approaches, which consider the joint configuration space of all involved robots, have the advantage of producing optimal and complete plans, yet tend to be computationally very expensive. Indeed, solving for optimality is NP-hard [31], and although significant progress has been made towards alleviating the computational load [6, 21], these approaches still scale poorly in environments with a high number of potential path conflicts.

Decentralized approaches provide an attractive alternative to centralized approaches, firstly, because they reduce the computational overhead, and secondly, because they relax the dependence on centralized units. This body of work considers the generation of collision-free paths for individual robots that cooperate only with immediate neighbors [3, 30], or with no other robots at all [26]. In the latter case, coordination is reduced to the problem of reciprocally avoiding other robots (and obstacles), and can generally be solved without the use of communication. Yet, by taking purely

[1]Qingbiao Li and Amanda Prorok are with the Department of Computer Science and Technology, University of Cambridge, Cambridge, United Kingdom
(Emails: ql295, asp45@cam.ac.uk).
[2]Fernando Gama and Alejandro Ribeiro are with the Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, USA (Emails: fgama,aribeiro@seas.upenn.edu).

Graph Neural Networks for Decentralized Path Planning

Extended Abstract

Qingbiao Li
Dept. of Computer Science and Technology, University of Cambridge
United Kingdom
ql295@cam.ac.uk

Fernando Gama, Alejandro Ribeiro
Dept. of Electrical and Systems Engineering, University of Pennsylvania, USA
{fgama,aribeiro}@seas.upenn.edu

Amanda Prorok
Dept. of Computer Science and Technology, University of Cambridge
United Kingdom
asp45@cam.ac.uk

ABSTRACT

We propose a combined model that automatically synthesizes local communication and decision-making policies for agents navigating in constrained workspaces. Our architecture is composed of a convolutional neural network (CNN) that extracts adequate features from local observations, and a graph neural network (GNN) that communicates these features among agents. We train the model to imitate an expert algorithm, and use the resulting model online in decentralized planning involving only local communication and local observations. We evaluate our method in simulations involving teams of agents in cluttered workspaces. We measure the success rates and sum of costs over the planned paths. The results show a performance close to that of our expert algorithm, demonstrating the validity of our approach. In particular, we show our model's capability to generalize to previously unseen cases (involving larger environments and larger agent teams).

KEYWORDS

Graph neural networks, decentralized multi-agent path planning, inter-robot communication

ACM Reference Format:
Qingbiao Li, Fernando Gama, Alejandro Ribeiro, and Amanda Prorok. 2020. Graph Neural Networks for Decentralized Path Planning. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

Efficient and collision-free navigation in multi-agent systems is fundamental to advancing mobility, where collision-free paths are generated to lead agents from their origins to designated destinations. Current approaches can be classified as either *coupled* or *decoupled*, depending on the structure of the state space that is searched. While coupled approaches are able to ensure the optimality and completeness of the solution, they involve *centralized* components, and tend to scale poorly with the number of agents [7, 8]. Decoupled approaches, on the other hand, compute trajectories for each agent separately, and re-plan only in case of conflicts [10–12]. This can significantly reduce the computational complexity of the planning task, but generally produces sub-optimal and incomplete solutions. The application of learning-based methods to multi-robot motion planning has attracted particular attention due to their capability of handling high-dimensional joint state-space representations, by offloading the online computational burden to an offline learning procedure [1, 3, 4, 9]. Notably, Sartoretti et al. [6] propose a hybrid learning-based method called PRIMAL for multi-agent path-finding that uses both imitation learning and multi-agent reinforcement learning. In contrast to our work, however, the latter approach does not learn explicit inter-agent communication policies.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

2 PROPOSED FRAMEWORK

We formulate the multi-agent path planning problem as a sequential decision-making problem that each agent solves at every time instant t, with the objective of reaching its destination. All components are described as following:

Partial observation processed by CNN: In an environment $(W \times H)$ with static obstacles, each agent has a local field-of-view (FOV) defined by the radius r_{FOV}, beyond which it cannot 'see' anything. The input map $\mathbf{Z}_t^i$ for agent i with size $W_{\text{FOV}} \times H_{\text{FOV}}$ is fed into a CNN that is run internally on each agent. This results in a vector $\tilde{\mathbf{x}}_t^i \in \mathbb{R}^F$ containing F observations, $\tilde{\mathbf{x}}_t^i = \text{CNN}(\mathbf{Z}_t^i)$. These observations can then be communicated to nearby agents.

Communication via GNN: The agents can communicate with each other as determined by the communication network defined by a graph $\mathcal{G}_t = (\mathcal{V}, \mathcal{E}_t, \mathcal{W}_t)$ at time t. Note that $\mathcal{V} = \{v_1, \dots, v_N\}$ is the set of N robots, $\mathcal{E}_t \subseteq \mathcal{V}\times\mathcal{V}$ is the set of edges and $\mathcal{W}_t : \mathcal{E}_t \to \mathbb{R}$ is a function that assigns weights to the edges. Robots v_i and v_j can communicate with each other at time t if $(v_i, v_j) \in \mathcal{E}_t$, i.e. if they are within a communication radius r_{COMM} of each other. The corresponding edge weight $\mathcal{W}_t(v_i, v_j) = w_t^{ij}$ can represent the strength of the communication.

Graph Neural Network: We define a *graph convolution* [2] as linear combination of shifted versions of the signal

$$\mathcal{A}(\mathbf{X}_t; \mathbf{S}_t) = \sum_{k=0}^{K-1} \mathbf{S}_t^k \mathbf{X}_t \mathbf{A}_k \tag{1}$$

where $\{\mathbf{A}_k\}_k$ is a set of $F \times G$ matrices representing the filter coefficients combining different observations. $\mathbf{S}_t^k\mathbf{X}_t = \mathbf{S}_t(\mathbf{S}_t^{k-1}\mathbf{X}_t)$ is computed by means of k communication exchanges with 1-hop neighbors, and is actually computing a summary of the information located at the k-hop neighborhood. The operation $\mathbf{S}_t\mathbf{X}_t$ represents a linear combination of neighboring values of the signal due to the sparsity pattern of $\mathbf{S}_t$.

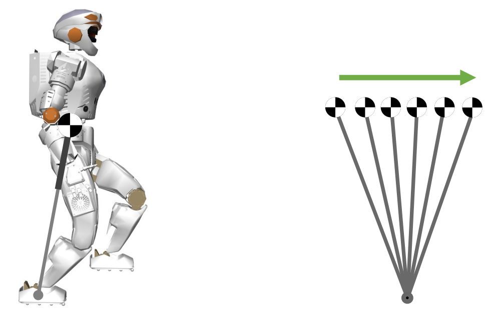
Action Policy: A local MLP (weight-sharing) is trained to predict action $\tilde{\mathbf{u}}_t^i$ taken by robot i, which is computed by a softmax over the probability distribution of motion primitives, includes up, left,

In Proceedings of IEEE International Conference on Humanoid Robots (Humanoids '17), Birmingham, UK (2017).

Robust Foot Placement Control for Dynamic Walking using Online Parameter Estimation

Qingbiao Li, Iordanis Chatzinikolaidis, Yiming Yang, Sethu Vijayakumar and Zhibin Li

***Abstract*— This paper presents an estimation scheme to control foot placement for achieving a desired dynamic walking velocity in presence of sensor and model errors. Inevitable discrepancies, such as sensors' noise, delay, and modelling errors, degrade the performance of model-based control methods or even cause instabilities. To resolve these issues, an on-line parameter estimation approach based on Tikhonov regularisation is formulated using measurement data, which is particularly robust for more accurately approximating the dynamics. The proposed scheme initially uses the foot placement predicted by the linear inverted pendulum model, while the control parameters are being optimised using adequate measurements to represent the real dynamics within and in-between steps; and then, the estimation based control is used to predict the future foot placement accurately in the presence of discrepancies.**



(a) Valkyrie robot (b) COM motion

Fig. 1: Bipedal walking control of the Valkyrie robot using the Linear Inverted Pendulum model (sagittal scenario).

I. Introduction

Humanoid robots, designed with a human morphology, offer advantages of traversing environments that are easily accessible by humans, such as stairs, passageways, rugged terrains, etc. [1] as well as using human-oriented tools [2]. A humanoid robot is a floating-base system of two-legs [3] with morphological adaptation to various surfaces, providing adaptability and maneuverability [4]. They have potentials to be indispensable in emergency and disaster responses, where wheeled robots are limited by the terrain irregularities. In turn, the mechanical complexity of humanoids imposes control challenges compared to wheeled robots.

Many model-based approaches have been studied to address the problem of bipedal locomotion. Kajita et al. [5] proposed the Linear Inverted Pendulum (LIP) model, which regards the robot as a point mass, to generate horizontal motions and keep the Centre of Mass (COM) height constant. Given a target COM motion, the corresponding Zero Moment Point (ZMP) or Centre of Pressure (COP) for achieving it can be analytically computed. LIP model and its extensions have been widely applied in bipedal walking, and its simplified modelling is illustrated in Fig. 1.

However, model-based approaches, e.g. LIP model-based foot placement control, generally have fixed coefficients and parameters manually tuned off-line for controlling legged locomotion [6]. For example, Raibert's control of a one-leg hopping robot has decoupled regulation of hoping height by delivering a fixed vertical thrust during stance, forward speed by foot placement, and an upright posture by exerting a torque around the hip [7]. Therefore, proper tuning of all variables was very crucial and usually relied on experience, which could only be done by experimental trial and error. However, this manual tuning has limitations because parameters might be time-varying or state-dependent. The same problems exist in other model-based approaches, especially when unexpected changes occur [8].

To resolve this, auto-tuning of parameters has been explored given a known control structure [9]. Nakanishi et al. [10] developed a framework to learn bipedal locomotion through movement primitives by locally weighted regression while the frequency of the learnt trajectories is adjusted automatically. You et al. [11] used linear regression based on past measurements for updating the coefficients of an extended formulation based on Raibert's model to achieve accurate velocity tracking. This method improved the system's flexibility to unknown changes, such as a mass offset, and was later extended to bipedal walking and running [12]. However, the convergence rate in You's method is significantly limited, because its formulation has two coefficients coupled with the measured velocity: one is directly for the velocity, and the other is for the velocity error where measured velocity appear as well. Hence, the coupling of these two coefficients resulted in the fluctuation of estimated values.

We propose an online estimation approach derived from the analytic insights of the LIP model, which has a major advantage in comparison with Raibert's linear model, i.e. the decoupling between the current forward velocity and the desired one. As a consequence, we expect a much faster and stable convergence of the coefficients needed for foot placement, thus better adaptation to unexpected changes, e.g. an unknown mass offset. We elaborated on a more robust calculation of the coefficients as well as on the effects of downgraded sensory information.

This paper contributes in the following aspects:

- A rigorous analysis of the propagation of sensor errors in walking control, and the inevitable uncertainties of model-based methods;
- Identification of control parameters derived from the

Looking At The Body: Automatic Analysis of Body Gestures and Self-Adaptors in Psychological Distress

Weizhe Lin, *Member, IEEE,* Indigo Orton, Qingbiao Li, Gabriela Pavarini, and Marwa Mahmoud, *Member, IEEE*

Abstract—Psychological distress is a significant and growing issue in society. Automatic detection, assessment, and analysis of such distress is an active area of research. Compared to modalities such as face, head, and vocal, research investigating the use of the body modality for these tasks is relatively sparse. This is, in part, due to the limited available datasets and difficulty in automatically extracting useful body features. Recent advances in pose estimation and deep learning have enabled new approaches to this modality and domain. To enable this research, we have collected and analyzed a new dataset containing full body videos for short interviews and self-reported distress labels. We propose a novel method to automatically detect self-adaptors and fidgeting, a subset of self-adaptors that has been shown to be correlated with psychological distress. We perform analysis on statistical body gestures and fidgeting features to explore how distress levels affect participants' behaviors. We then propose a multi-modal approach that combines different feature representations using Multi-modal Deep Denoising Auto-Encoders and Improved Fisher Vector Encoding. We demonstrate that our proposed model, combining audio-visual features with automatically detected fidgeting behavioral cues, can successfully predict distress levels in a dataset labeled with self-reported anxiety and depression levels.

Index Terms—Self-adaptors, fidgeting, psychological distress, digital phenotyping, behavioural sensing

1 Introduction

Psychological distress and mental disorders are significant threats to global health [1].[1] According to the World Health Organization (WHO), an estimated 450 million people around the world suffer from neuropsychiatric conditions [3], with depression and anxiety being the most common mental disorders [4]. Despite existing strategies for the treatment of distress, such as depression, it is estimated that nearly two-thirds of people suffering distress have never received help from a health professional [5]. Early detection of distress is consistently noted as a key factor in treatment and positive outcomes. Early detection requires an ongoing assessment to identify distress when it begins. Self-evidently, ongoing assessment at scale is prohibitive when performed manually. As such, automatic detection of signs of psychological distress or specific mental disorders is an active area of research.

Currently, the most effective automated distress detection approaches utilize multi-modal machine learning. These modalities include facial, head, eye, linguistic (textual), vocal, and body.

There are significant challenges to body modality research, particularly within automatic distress detection, including the lack of relevant data, the inability to share much of the data, and the difficulty in gathering such data. Specifically, the combination of full-body data (either sensor-based or video-based) with psychological distress labels is rare. Compounding this rarity is the private and sensitive nature of the data, which means such datasets are rarely shared publicly.

Body expressions, and especially self-adaptors, have been shown to be correlated with human affect, depression and psychological distress [6], [7], [8], [9], [10]. Self-adaptors are self-comforting gestures, including any kind of touching on other parts of the body, either dynamically or statically [11], [12]. Fidgeting, a subset of self-adaptors, is the act of moving about restlessly, playing with one's fingers, hair, or personal objects in a way that is not peripheral or nonessential to ongoing tasks or events [13]. Patients with depression often engage in self-adaptors [14]. Fidgeting has been seen and reported in both anxiety and depression [12]; It is a sign of attention-deficit and hyperactivity disorder, also exhibited by individuals with autism [15]. With manually annotated data, Scherer *et al.* [16] reported a longer average duration of self-adaptors as well as fidgeting for distressed participants.

More recent advances in the state-of-the-art for pose estimation [17] enable accurate pose data on a broader set of datasets and thus open the door for new approaches for body expression analysis and broader incorporation of body features in multi-modal systems.

In this paper, we propose to use a hierarchical model to automatically detect self-adaptors as well as fidgeting, which has been shown to be predictive of psychological distress. We analyzed body gestures and self adaptors in a

- W. Lin is with the Department of Engineering, Cambridge. E-mail: wl356@cam.ac.uk
- I. Orton, Q. Li and M. Mahmoud are with the Department of Computer Science and Technology, Cambridge. Email: {indigo.orton@cl., ql295@, mmam3@}cam.ac.uk
- G. Pavarini is with the Department of Psychiatry, Oxford. Email: gabriela.pavarini@psych.ox.ac.uk

1. This work is an extension of the work in [2], originally published in the proceedings of the IEEE International Conference on Automatic Face and Gesture Recognition (FG) 2020

International Journal of Computer Assisted Radiology and Surgery (2019) 14:987–995
https://doi.org/10.1007/s11548-019-01940-2

ORIGINAL ARTICLE

Estimation of tissue oxygen saturation from RGB images and sparse hyperspectral signals based on conditional generative adversarial network

Qingbiao Li[1,2] · Jianyu Lin[1,3] · Neil T. Clancy[4,5] · Daniel S. Elson[1,2]

Received: 29 January 2019 / Accepted: 7 March 2019 / Published online: 21 March 2019

Abstract

Purpose Intra-operative measurement of tissue oxygen saturation (StO_2) is important in detection of ischaemia, monitoring perfusion and identifying disease. Hyperspectral imaging (HSI) measures the optical reflectance spectrum of the tissue and uses this information to quantify its composition, including StO_2. However, real-time monitoring is difficult due to capture rate and data processing time.

Methods An endoscopic system based on a multi-fibre probe was previously developed to sparsely capture HSI data (sHSI). These were combined with RGB images, via a deep neural network, to generate high-resolution hypercubes and calculate StO_2. To improve accuracy and processing speed, we propose a dual-input conditional generative adversarial network, Dual2StO2, to directly estimate StO_2 by fusing features from both RGB and sHSI.

Results Validation experiments were carried out on in vivo porcine bowel data, where the ground truth StO_2 was generated from the HSI camera. Performance was also compared to our previous super-spectral-resolution network, SSRNet in terms of mean StO_2 prediction accuracy and structural similarity metrics. Dual2StO2 was also tested using simulated probe data with varying fibre number.

Conclusions StO_2 estimation by Dual2StO2 is visually closer to ground truth in general structure and achieves higher prediction accuracy and faster processing speed than SSRNet. Simulations showed that results improved when a greater number of fibres are used in the probe. Future work will include refinement of the network architecture, hardware optimization based on simulation results, and evaluation of the technique in clinical applications beyond StO_2 estimation.

Keywords Intro-operative imaging · Optical imaging · Tissue oxygen saturation · Generative adversarial network

Introduction

Tissue perfusion and oxygenation are important clinical indicators of organ health during minimal access surgery (MAS). Endoscopic hyperspectral imaging (HSI) is a non-invasive optical technique to capture quantitative spectral information with a high spatial resolution based on narrow spectral bands over a virtually continuous spectral range for live tissue diagnostics and monitoring [1]. HSI can be used to estimate oxygen saturation (StO_2) and perfusion, which reflects tissue function and the health of an organ's blood supply. This, in turn, can be applied to various important clinical applications

✉ Qingbiao Li
ql716@ic.ac.uk

Jianyu Lin
xjtuljy@gmail.com

Neil T. Clancy
n.clancy@ucl.ac.uk

Daniel S. Elson
daniel.elson@imperial.ac.uk

1 The Hamlyn Centre for Robotic Surgery, Imperial College London, London, UK

2 Department of Surgery and Cancer, Imperial College London, London, UK

3 Department of Computing, Imperial College London, London, UK

4 Wellcome/EPSRC Centre for Interventional and Surgical Sciences (WEISS), University College London, London, UK

5 Centre for Medical Image Computing (CMIC), Department of Computer Science, University College London, London, UK

Real-time Surgical Environment Enhancement for Robot-Assisted Minimally Invasive Surgery Based on Super-Resolution

Ruoxi Wang[1], Dandan Zhang[1], *Student Member, IEEE*, Qingbiao Li[2], Xiao-Yun Zhou[3], Benny Lo[1], *Senior Member, IEEE*

***Abstract*— In Robot-Assisted Minimally Invasive Surgery (RAMIS), a camera assistant is normally required to control the position and zooming ratio of the laparoscope, following the surgeon's instructions. However, moving the laparoscope frequently may lead to unstable and suboptimal views, while the adjustment of zooming ratio may interrupt the workflow of the surgical operation. To this end, we propose a multi-scale Generative Adversarial Network (GAN)-based video super-resolution method to construct a framework for automatic zooming ratio adjustment. It can provide automatic real-time zooming for high-quality visualization of the Region Of Interest (ROI) during the surgical operation. In the pipeline of the framework, the Kernel Correlation Filter (KCF) tracker is used for tracking the tips of the surgical tools, while the Semi-Global Block Matching (SGBM) based depth estimation and Recurrent Neural Network (RNN)-based context-awareness are developed to determine the upscaling ratio for zooming. The framework is validated with the JIGSAW dataset and Hamlyn Centre Laparoscopic/Endoscopic Video Datasets, with results demonstrating its practicability.**

I. Introduction

In current Robot-Assisted Minimally Invasive Surgery (RAMIS), surgeons have to rely solely from the laparoscopic or endoscopic camera images to carry out the operation. With recent advances in Artificial Intelligence (AI), AI algorithms become increasingly popular in robot-aided MIS surgery to assist surgeons to perform the delicate operations [1].

In current practice, surgeons control the surgical robot based on the visual information obtained from the laparoscope. Generally, a camera assistant is required to control and operate the laparoscope following the surgeon's command to focus and zoom into targeted Region of Interest (ROI) during operation. The camera assistant needs to hold the laparoscope and follow the surgeon's instructions to control the position, distance and zooming ratio of the laparoscope, and even for RAMIS, the surgeon has to switch the manipulator to control the camera. However, the zooming ratio required by the surgeons during surgical operations varies, which depends on the specific operation scenarios. A larger zooming ratio is required for delicate operations such as inspecting the tumors or organs, while a smaller ratio might be sufficient for relatively simple operations such as moving the surgical tools to another target area. Indirect visual control is very inefficient for surgical operations. Besides, simply zooming is not enough to provide the surgeon with a clearer image to assist the operations. Much detailed information, such as the high-frequency fine details, will be lost in the traditional zooming process. In addition, there are only very few laparascopic cameras have optical zooming functions, which can be very costly. Therefore, it is essential to investigate super-resolution methods that could provide the zooming function with high-quality images.

[1]The Hamlyn Centre for Robotic Surgery, Imperial College London, United Kingdom. [2] University of Cambridge, United Kingdom. [3] PAII Inc, MA, USA. Corresponding Author email: d.zhang17@imperial.ac.uk.

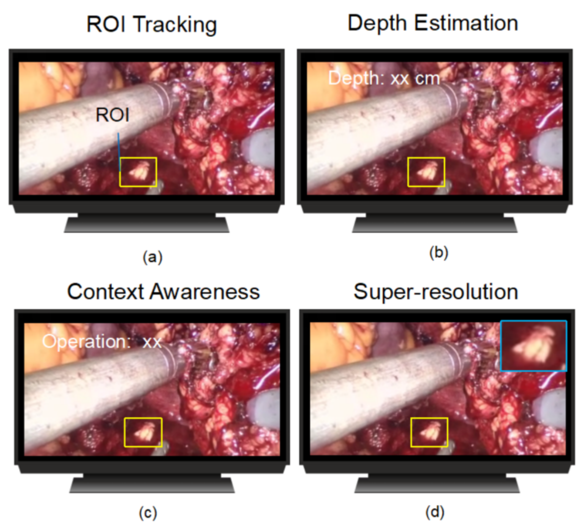


Fig. 1. The pipeline and the typical scenario of the proposed framework. (a) An ROI is initialized and tracked. (b) The depth between the camera and the ROI (xx cm) is estimated. The ROI will not be super-resolved if the estimated depth is less than a pre-defined threshold. (c) The surgeon's next operation will be predicted by RNN-based surgical scene recognition method. The predicted operation will determine the upscaling factor for the super-resolution method. (d) The ROI is super-resolved at a specific upscaling factor determined by the predicted operation and is presented in the top-right corner of the screen or another window.

Related Work. End-to-end deep convolutional neural network-based methods have been actively investigated in recent years to solve the super-resolution problem [2]–[6], of which SRCNN [7] and FSRCNN [8] are the most popular methods.

In order to generate more plausible-looking natural images with high perceptual quality, generative adversarial network (GAN) has attracted huge attention in tackling the super-resolution task [3], [9], [10]. SRGAN [3] uses the SRResNet to construct the generator, and uses the content loss and the adversarial loss as the perceptual loss to replace the Mean-Square-Error (MSE) loss to train the generator. To further improve the visual quality of the high-resolution estimation, ESRGAN [9] has been proposed to mainly focus on constructing the generator by using the Residual-in-Residual Dense Block, and the relativistic discriminator to calculate the adversarial loss.

Estimation of Tissue Oxygen Saturation from RGB Images based on Pixel-level Image Translation

Qing-Biao Li[1,3], Xiao-Yun Zhou[1], Jianyu Lin[1,3],Jian-Qing Zheng[1], Neil T. Clancy[2], Daniel S. Elson[1,3]

[1]*The Hamlyn Centre for Robotic Surgery, Imperial College London, London, UK,*

[2]*Centre for Medical Image Computing, University College London, London, UK*

[3]*Department of Surgery and Cancer, Imperial College London, London, UK*
q.li17@imperial.ac.uk

INTRODUCTION

Oxygenation and perfusion reflect tissue metabolic activity, which can be potentially used to monitor diseases including diagnosis or characterisation of cancer. A minimally-invasive endoscopy technique based on HSI uses narrow spectral bands over a continuous spectral range to capture quantitative spectral information for live tissue diagnostics. StO2 can be estimated from HSI images using the Beer-Lambert law, if oxy- and deoxy- haemoglobin are considered to be the primary absorbers within the visible wavelength range, and using an approximation for scattering [1].

The Imperial College London, Structured Light and Hyperspectral Imaging system (ICL SLHSI) is an optical probe system combining sparse hyperspectral measurements and spectrally-encoded structured lighting, which can accurately estimate StO2 with high spatial resolution [2,3]. However, the requirement for an additional optical probe creates a barrier to widespread adoption. In contrast, RGB cameras are widely used during minimally invasive surgery. Hence, it is worthwhile to develop methodologies for estimating the StO2 from RGB images. Currently, HSI images were first recovered from RGB images via 'super-spectral resolution' and then used to estimate StO2 by linear regression (route 1 in Fig. 1)[3], [4].

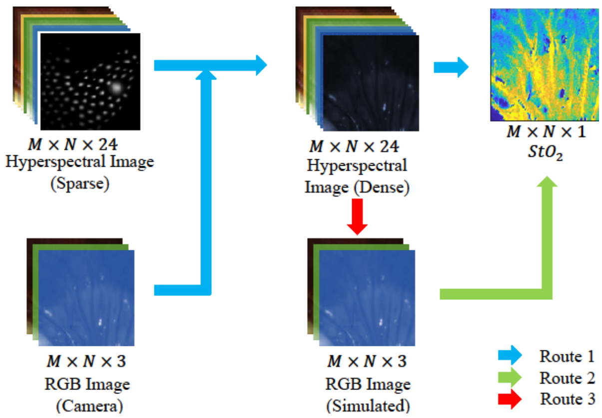


Fig. 1 Processing methods for StO2 estimation.

In this paper, conditional Generative Adversarial Networks (cGAN) [5] are introduced to achieve a pixel-level image translation, estimating the StO2 directly from RGB images without super-spectral resolution (route 2 in Fig. 1). The following sections will describe the network architecture, data collection, and validation setup. The test results on three types of *in vivo* data including rabbit uterus, lamb uterus, and porcine bowel are presented and discussed.

MATERIALS AND METHODS

Simulation of Dataset

Given an HSI image, an RGB image (input – 'x') can be accurately simulated based on the specific RGB camera responses to the corresponding wavelength range (route 3 in Fig. 1). The intensity ratio $\left(\frac{I_\lambda(x)}{I_{0,\lambda}(x)}\right)$ for each pixel (i,j) reflects the attenuation of the incident light spectrum ($I_{0,\lambda}(x)$) due to oxy- $[HbO_2]$ and deoxy-haemoglobin $[Hb]$. Given the molar extinction coefficient of each haemoglobin $\varepsilon(\lambda)$ at specific wavelength λ, the concentration of both haemoglobins can be calculated based on a least squares method [2]:

$$I_\lambda(x) = I_{0,\lambda}(x)e^{-([HbO_2]\varepsilon_{HbO_2}(\lambda)+[Hb]\varepsilon_{Hb}(\lambda)+\alpha)}$$

Where, α is a constant value representing the attenuation caused by scattering and other components within tissue. Thus, the StO2 ground truth can be calculated (image – 'z').

Conditional Generative Adversarial Networks

In the training stage, the RGB images and the StO2 ground truth image are put into a generator whose structure is shown in Fig. 2, outputting an estimated StO2 (synthesized image – 'y'). Concat(x,y) and Concat(x,z) are separately put into a discriminator, outputting the probability of the input to be z. Here Concat() is concatenate, the probability map is a $30 \times 30 \times 1$ map which is useful for pixel-level rather than image-level translation [5]. The discriminator network architecture is shown in Fig. 3.

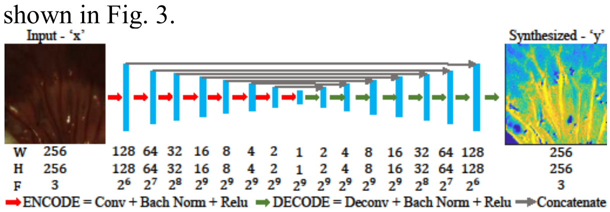


Fig. 2 Generator network architecture (G).