cases

doc_1		doc_2		decision	id
authors	• Arthur Bucker • Luis Figueredo • Sami Haddadin • Ashish Kapoor • Shuang Ma • Sai Vemprala • Rogerio Bonatti	**authors**	• A. Bucker • Luis F. C. Figueredo • S. Haddadin • Ashish Kapoor • Shuang Ma • Sai Vemprala • Rogerio Bonatti	DUPLICATES	144
title	LATTE: LAnguage Trajectory TransformEr	**title**	LaTTe: Language Trajectory TransformEr		
publication_date	2022-08-04 22:43:21+00:00	**publication_date**	2022-08-04 00:00:00		
source	SupportedSources.ARXIV	**source**	SupportedSources.SEMANTIC_SCHOLAR		
journal	None	**journal**	ArXiv		
volume		**volume**	abs/2208.02918		
doi		**doi**	10.48550/arXiv.2208.02918		
urls	• http://arxiv.org/pdf/2208.02918v3 • http://arxiv.org/abs/2208.02918v3 • http://arxiv.org/pdf/2208.02918v3	**urls**	• https://www.semanticscholar.org/paper/97e6b89f8f256289b01b9f31799d957db81f2d4e		
id	id-4103382995225983990	**id**	id-5838064122031208533		
abstract	Natural language is one of the most intuitive ways to express human intent. However, translating instructions and commands towards robotic motion generation and deployment in the real world is far from being an easy task. The challenge of combining a robot's inherent low-level geometric and kinodynamic constraints with a human's high-level semantic instructions traditionally is solved using task-specific solutions with little generalizability between hardware platforms, often with the use of static sets of target actions and commands. This work instead proposes a flexible language-based framework that allows a user to modify generic robotic trajectories. Our method leverages pre-trained language models (BERT and CLIP) to encode the user's intent and target objects directly from a free-form text input and scene images, fuses geometrical features generated by a transformer encoder network, and finally outputs trajectories using a transformer decoder, without the need of priors related to the task or robot information. We significantly extend our own previous work presented in Bucker et al. by expanding the trajectory parametrization space to 3D and velocity as opposed to just XY movements. In addition, we now train the model to use actual images of the objects in the scene for context (as opposed to textual descriptions), and we evaluate the system in a diverse set of scenarios beyond manipulation, such as aerial and legged robots. Our simulated and real-life experiments demonstrate that our transformer model can successfully follow human intent, modifying the shape and speed of trajectories within multiple environments. Codebase available at: https://github.com/arthurfenderbucker/LaTTe-Language-Trajectory-TransformEr.git	**abstract**	â€”Natural language is one of the most intuitive ways to express human intent. However, translating instructions and commands towards robotic motion generation and deployment in the real world is far from being an easy task. The challenge of combining a robotâ€™s inherent low-level geometric and kinodynamic constraints with a humanâ€™s high-level semantic instructions traditionally is solved using task-speciï¬c solutions with little generalizability between hardware platforms, often with the use of static sets of target actions and commands. This work instead proposes a ï¬‚exible language-based framework that allows a user to modify generic robotic trajectories. Our method leverages pre-trained language models (BERT and CLIP) to encode the userâ€™s intent and target objects directly from a free-form text input and scene images, fuses geometrical features generated by a transformer encoder network, and ï¬nally outputs trajectories using a transformer decoder, without the need of priors related to the task or robot information. We signiï¬cantly extend our own previous work presented in [1] by expanding the trajectory parametrization space to 3D and velocity as opposed to just XY movements. In addition, we now train the model to use actual images of the objects in the scene for context (as opposed to textual descriptions), and we evaluate the system in a diverse set of scenarios beyond manipulation, such as aerial and legged robots. Our simulated and real-life experiments demonstrate that our transformer model can successfully follow human intent, modifying the shape and speed of trajectories within multiple environments. Codebase avail- able at: https://github.com/arthurfenderbucker/ LaTTe-Language-Trajectory-TransformEr.git .		
versions		**versions**			