

# Staged Contact-Aware Global Human Motion Forecasting

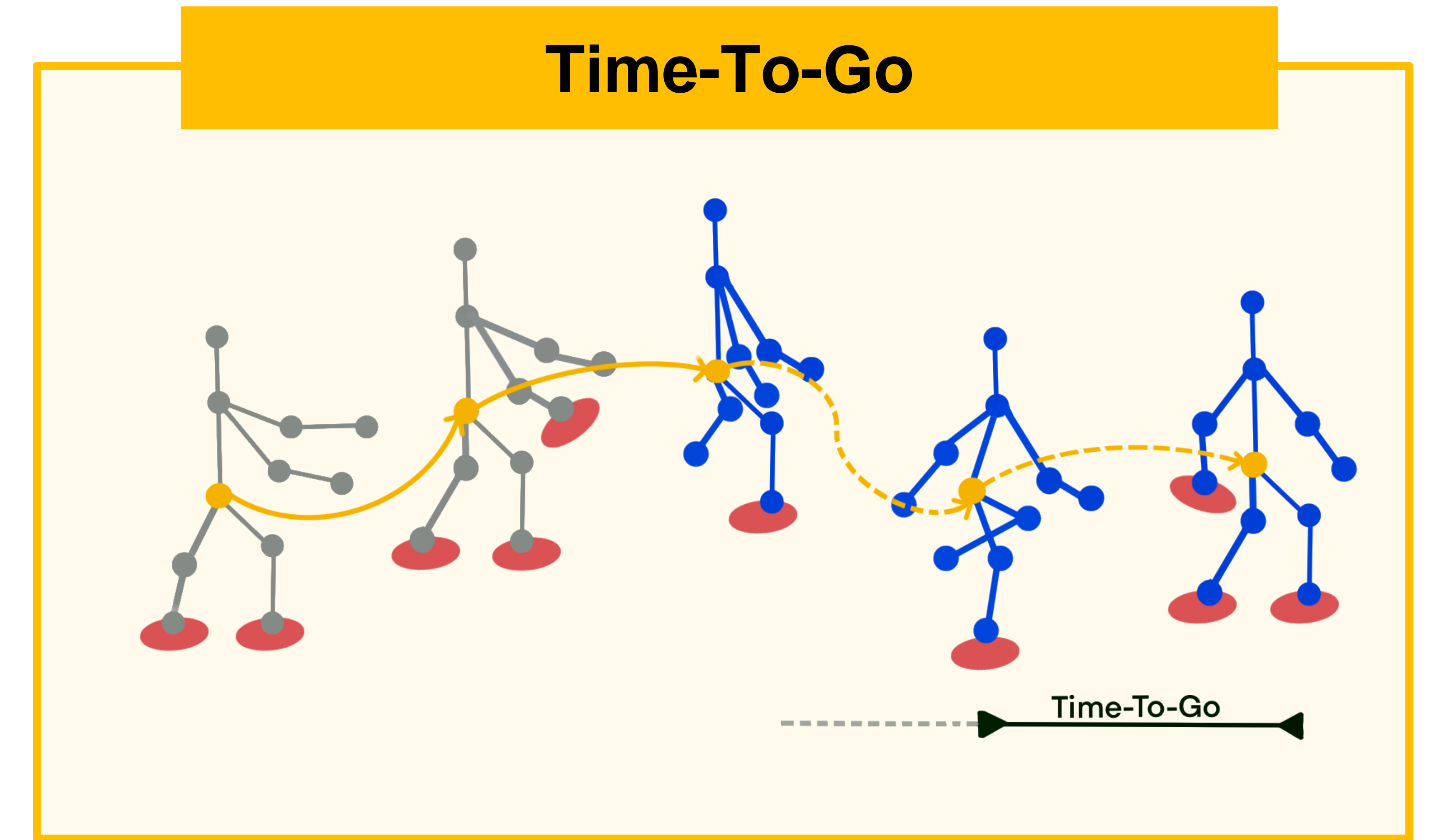
L. Scofano<sup>1</sup>, A. Sampieri<sup>1</sup>, E. Schiele<sup>2</sup>, E. De Matteis<sup>1</sup>, L. Leal-Taixé<sup>2</sup>, F. Galasso<sup>1</sup>  
Sapienza University of Rome<sup>1</sup>, Technical University of Munich<sup>2</sup>

## Human Forecasting in Scene

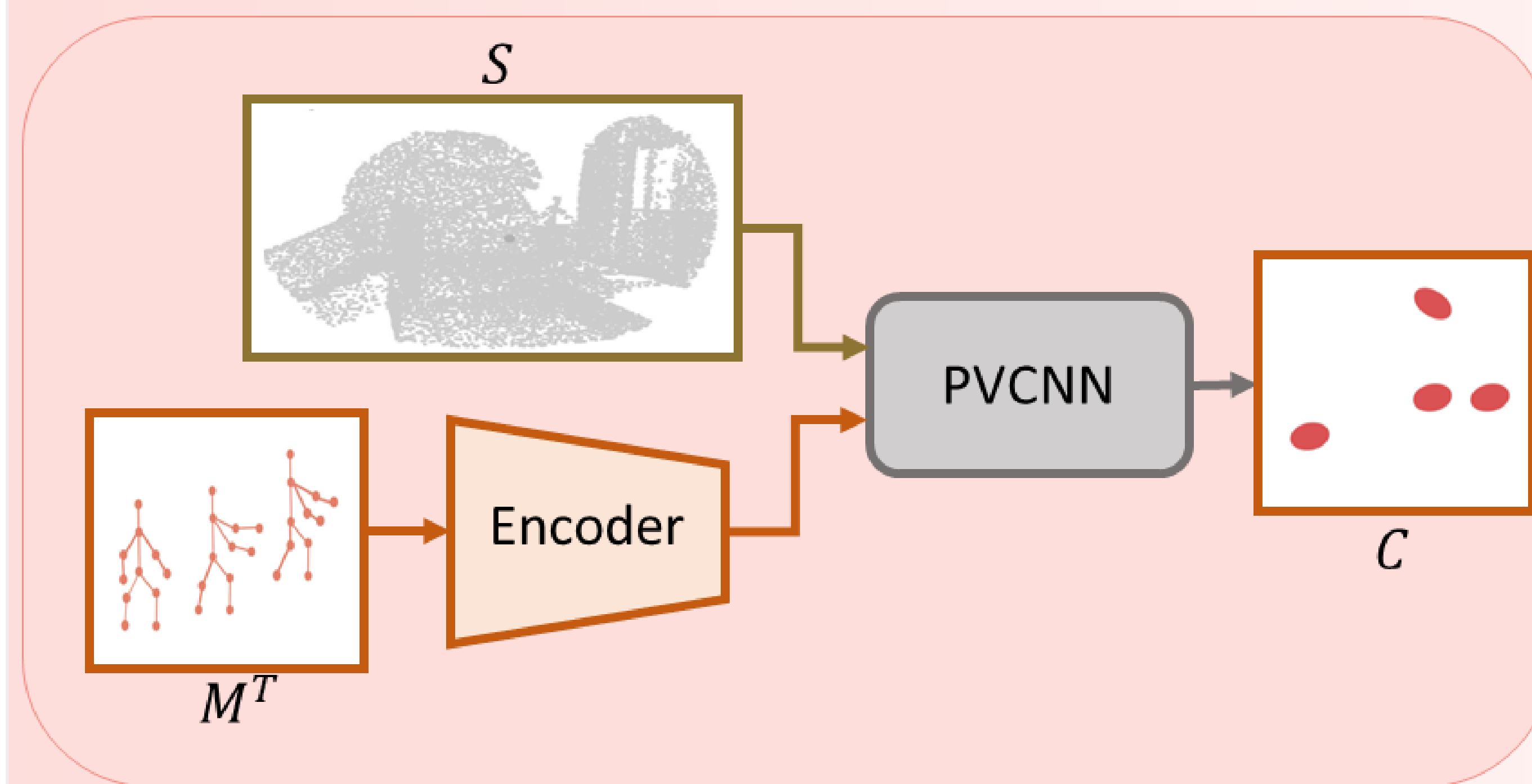
**Task:** Predict the future global motion of a human body in a 3D environment, taking into account the interaction between the human body and the environment.

**Problem:** The previous model performs global pose forecasting the latter as the end-to-end forecasting of future trajectories and poses, in contrast with the coarse-to-fine nature of the task.

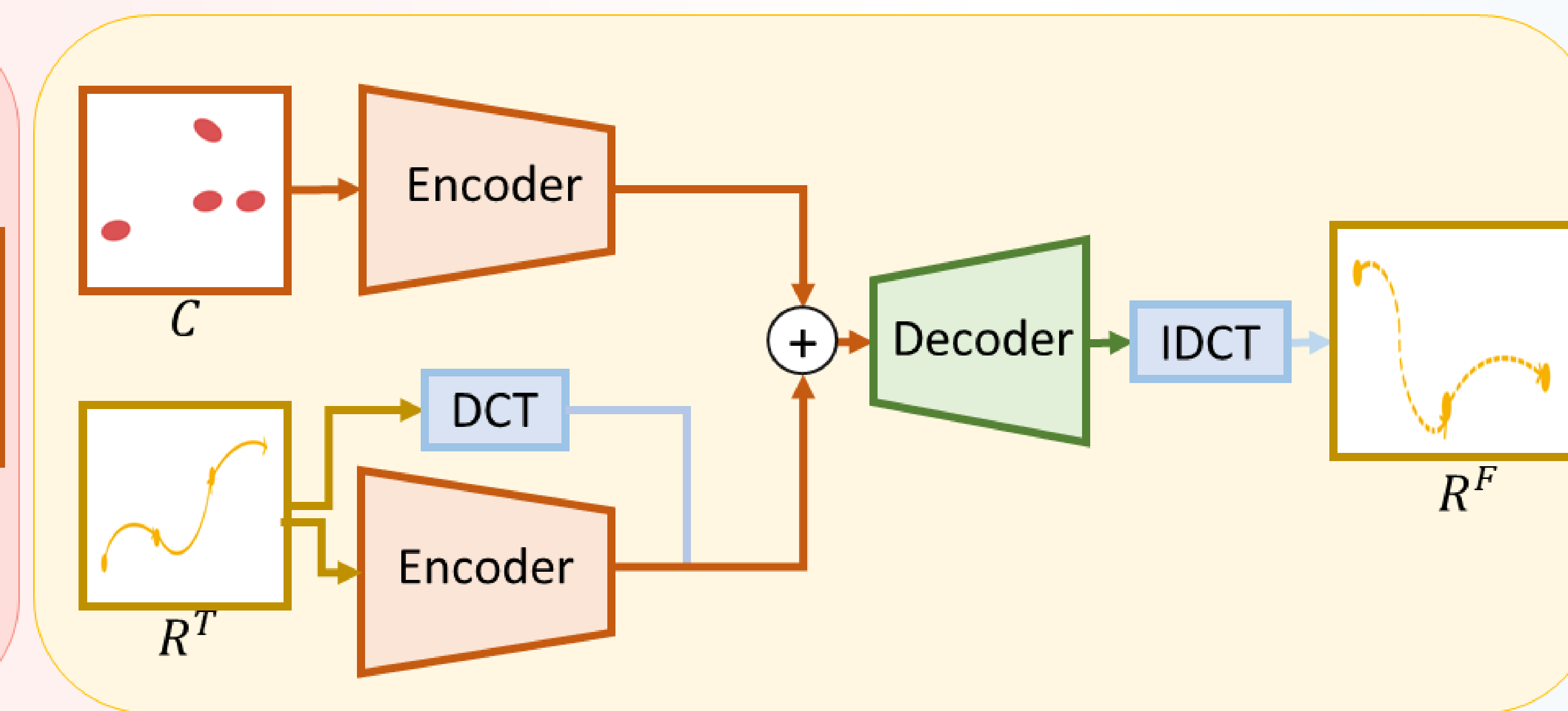
**Motivation:** We decouple global motion forecasting into two subtasks, trajectory forecasting and human pose forecasting



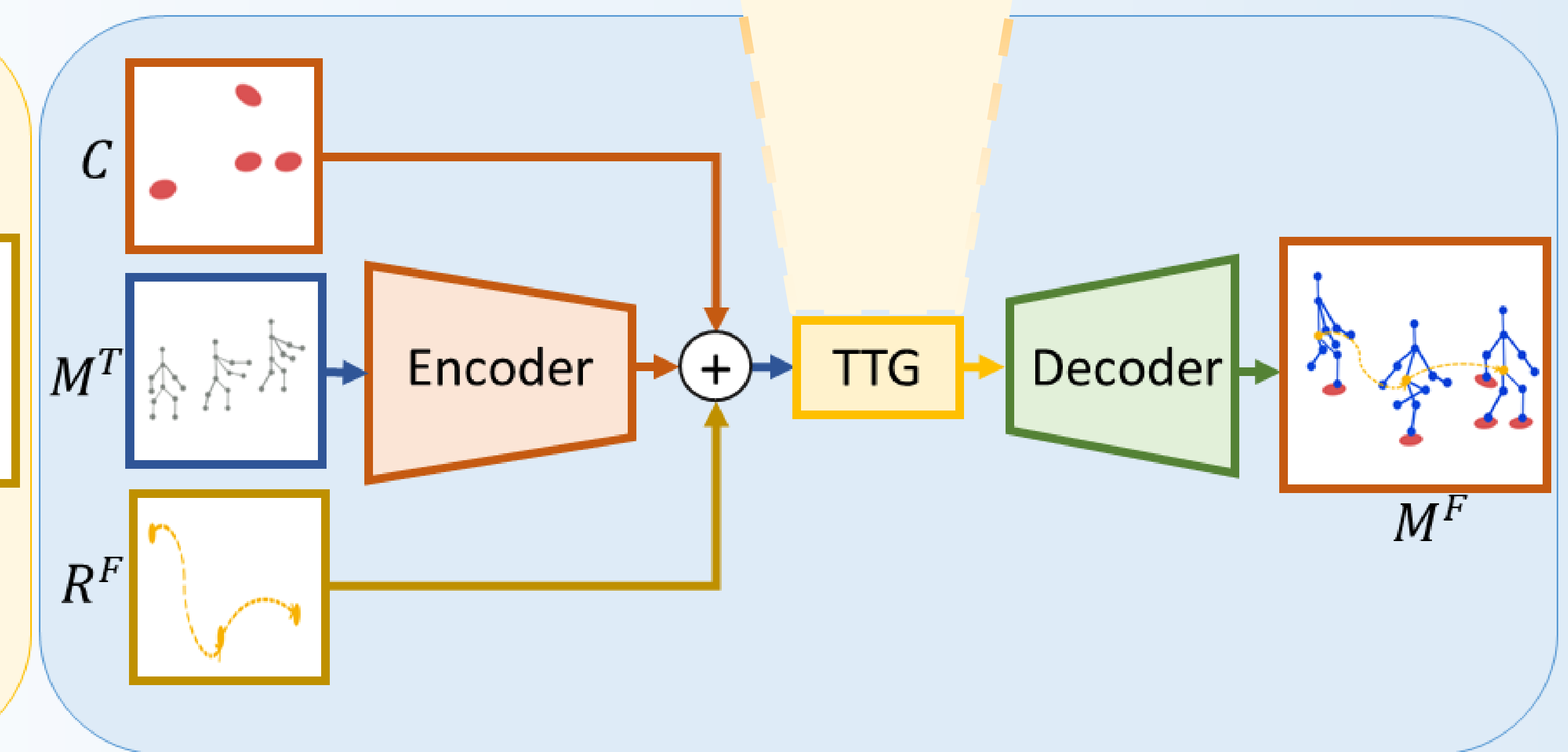
## STAG



Stage 1: Contact Point Estimation



Stage 2: Root Forecasting



Stage 3: Global Pose Forecasting

## Take Home Messages

### STAG:

- addresses global 3D pose prediction via scene, human trajectory, and pose modeling.
- is the first scene-aware model separates trajectory and pose for better coarse-to-fine alignment.
- uses a TTG encoder which provides information about the duration remaining before reaching scene contact and endpoints

## Results

Models	Path Error (mm)					Pose Error (mm)				
	0.5s	1s	1.5s	2s	Mean	0.5s	1s	1.5s	2s	Mean
LTD	67.0	119.3	207.6	375.6	147.4	67.5	93.8	98.9	103.5	80.5
DMGNN	82.7	158.0	227.8	286.9	156.2	47.5	69.1	85.6	95.3	64.9
SLT	<b>45.9</b>	117.0	186.7	267.1	121.8	70.8	181.4	150.2	196.0	112.6
Mao et al.	58.0	103.2	154.9	221.7	108.4	50.8	67.5	75.5	<b>86.9</b>	61.4
<b>STAG</b>	55.4	<b>89.6</b>	<b>127.9</b>	<b>179.3</b>	<b>92.3</b>	<b>48.1</b>	<b>65.3</b>	<b>75.6</b>	88.2	<b>60.3</b>



Code Available!



Project Page