

Saif Punjwani, Bowen Zuo, Micah Grimes, Larry Heck

## Goal

Create a **Gesture Generation Model** that

- Creates realistic **full-body gestures**
- Communicates virtually in real-time
- Embody nuances of **nonverbal cues**
- Engage in **human-gesture dialogue**

## Method

This work introduces a **new approach transforming text into full-body gestures**

- Gestural Transformer Framework
- Allocentric Gesture Dataset
- Text-to-Gesture Sequence Mapping
- Full-Body Gesture Synthesis

Georgia Institute of Technology

## Preliminary Results

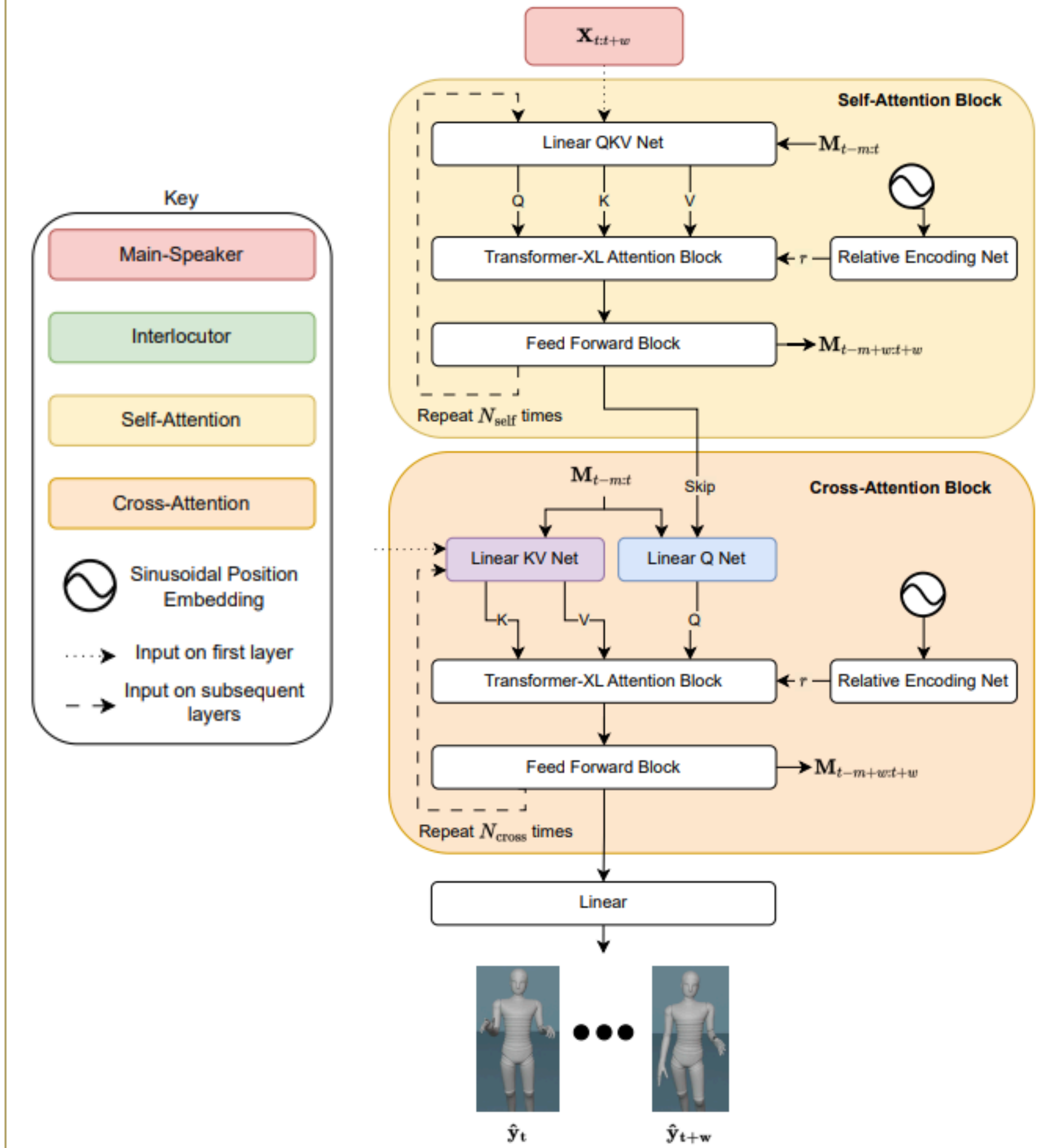
Table 1: Summary of results for human-likeness.

Condition	Median	Mean
NA	71 e [70, 71]	68.4
AM	70 e [69, 70]	67.9
SG	64 e [63, 65]	61.8
SF	63 e [62, 64]	61.2
SJ	59 e [58, 60]	57.1

Table 2: Summary of results for speech appropriateness.

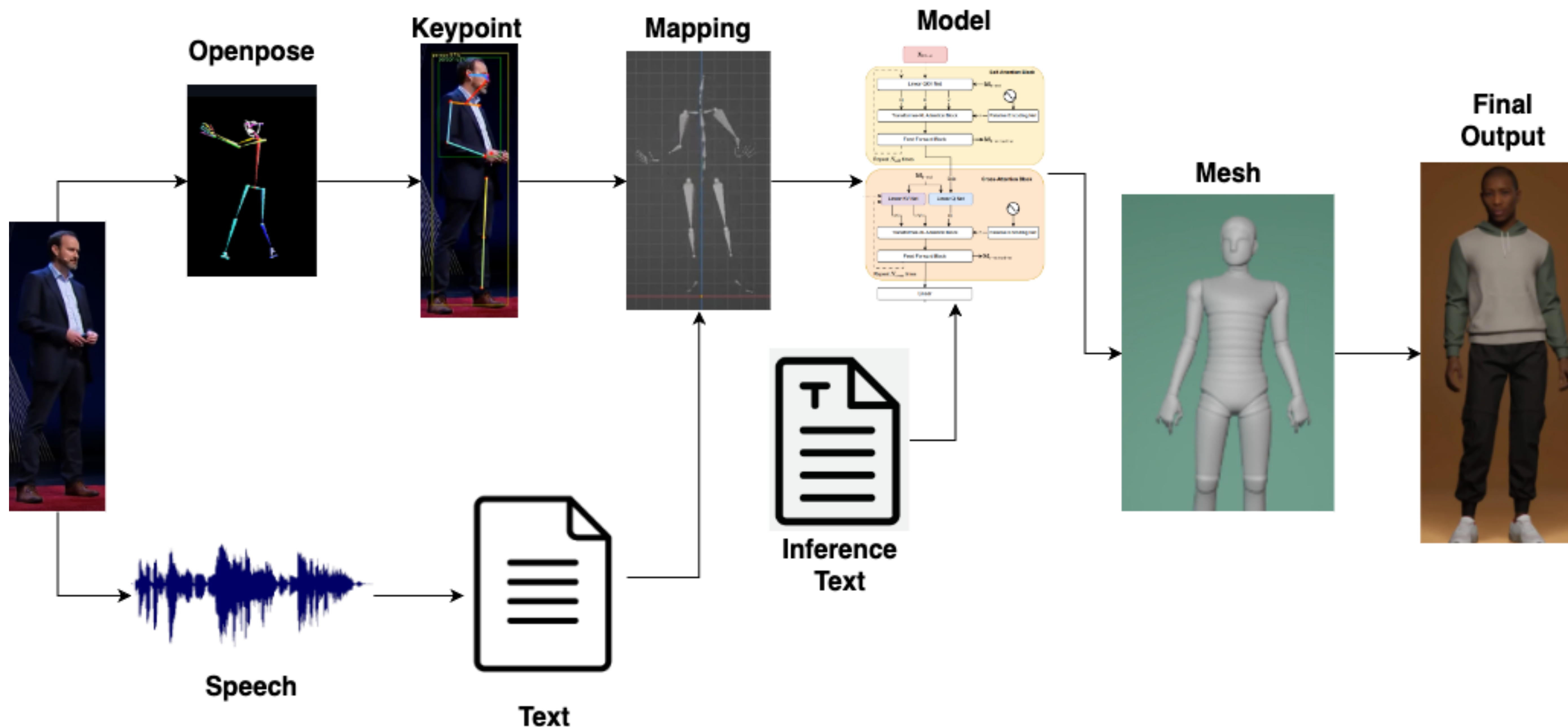
Condition	MAS	Pref.M. (%)
NA	0.81	73.6
AM	0.80	72.2
SG	0.78	70.4
SF	0.76	69.3
SJ	0.74	67.9

## Model Architecture



UEA + AVA Approach

## Gesture Generation Approach



Scan to be in the loop

