

This is the weekly report by Wentao from 2022/10/5 to 2022/11/11

topics

- Gesture Generation
- ACM MM Paper

details

Gesture Generation

主要工作：

- ☒ 文本特征和音频特征的融合
- ☐ 节奏特征的提取和生成

结果：

加入了文本特征之后PCK指标从36.1%提升到了42.3%

(PCK,关键点准确率, 计算预测点和GT之间的距离是否在一个阈值以内)

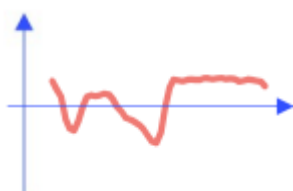
分析原因：

- 1.PCK较低的最主要原因还是数据的问题, 存在一些预测不准确和检测不到的帧
- 2.3维关键点的维度相比2维关键点更难预测, 点的数量更多
- 3.没有学到对应的节奏信息, 一直有动作变化, 没有停顿和间隔

节奏特征提取

我们用一段运动序列的平均偏移量来定义这段运动的节奏特征

$$\mathcal{L}_r = \left\| \widetilde{M}_i^* - (M_i - M_i^m) \right\|$$



M表示某一段运动序列，我们使用一个CNN网络来生成某段给定语音特征（MFCC特征）的运动节奏，得到运动的节奏之后和多模态动作进行合成

后续工作：

待效果达到标准后之后进行主观评估（主要由Q语通专家进行判断，包括和动作和视频相似性，是否满足Q语通语义等）

ACM MM Paper

见proposal