PAJRD20170274---基于CQT的微表情面审方法

郭老师好！该文档中存在两个问题：

1. 该专利中只有视频的相关特征，没有语音的相关特征。是否我之前的表达有误给老师造成了误解？还需要麻烦郭老师再修改一下，将微表情面审方法中的声音特征内容删除。
2. Tsne降维后没有具体的物理意义，不能够将第一维特征记录为声音，第二维度记录为表情。两个维度的特征是用原先4170维的特征向低维空间投影。在此过程中，特征发生了根本性的变化，原始的特征消失了。

**暂定名称**：基于CQT的微表情面审方法

**背景技术**：

传统风控面审流程中，为了核实贷款者信息真实性，会逐一对所提供信息进行询问核查。但这种方法对于贷款者故意隐瞒、欺诈的情形无法有效识别。

**技术方案**：

本发明提供一种**微表情面审情绪识别方法**【暂定发明名称】，该方法包括：

**问题库构建步骤**：构建CQT问题库，该CQT问题库包相关问题、准绳问题、中性问题三种类型的问题，给每种类型的问题分配一个标签，将构建的CQT问题库存储于存储空间；

**问卷生成步骤**：根据预设的第一规则及所述标签从CQT问题库抽取上述三种类型的问题，并按照预设的第二规则对抽取的问题进行排序，生成测试问卷；

**视频切割步骤**：拍摄测试对象回答测试问卷中所有问题的视频，以单个问题为单位对视频进行切割得到每个问题对应的视频片段，并将每个问题的标签分配给该问题对应的视频片段；

**特征提取步骤**：提取每个视频片段中的声音特征向量及表情特征向量，将每个问题对应的视频片段的声音特征向量及表情特征向量视为每个问题的特征向量；

**聚类步骤**：将每种类型的问题对应的特征向量归为一类；

**计算步骤**：计算每种类型问题的中心点特征向量，计算每个相关问题的特征向量与中性问题的中心点特征向量之间的第一距离，计算每个相关问题的特征向量与准绳问题的中心点特征向量之间的第二距离；

**识别步骤**：当第一距离大于第二距离时，判断该测试对象回答该相关问题时隐瞒了真实情绪，当第一距离小于第二距离时，判断该测试对象回答该相关问题时表现出的情绪是真实的。

CQT的全称为comparison question test，其包括三种类型的问题：相关问题、准绳问题、中性问题，定义如下：

中性问题：即与事件本身无关的问题，被测人没有负担，不会引起心理压力和情绪反应；

准绳问题：即能够给被测人造成一定的心理压力，与相关问题相似，但与测试主题无关，明知被测人很可能说谎或会说谎的问题，用以和相关问题进行比较；

相关问题：即和case有关的问题，是测试所要甄别的问题。

CQT的理论基础是，说谎者更害怕与事件（待调查的事件）相关的问题，因而会对相关问题产生更大的心理反应；而诚实者害怕“准绳问题”，因而会对准绳问题产生更大的心理反应。因此，根据被测人对相关问题和准绳问题反应的差异，就可以把说谎和诚实者区分出来。

上述问卷生成步骤中，所述第一规则为：每一份测试问卷至少要有预设数量（例如15个）以上的问题，要保证相关问题、不相关问题、准绳问题三种不同类型的题目各有一定数量（例如每种类型的问题需3个以上），且相关问题及中性问题的数量分别多于准绳问题的数量。所述第二规则为：在测试问卷中，相同类型的问题位置不能相邻。此外，第一题、第二题及最后一题设置为中性问题。“不得相邻”的要求，是为了能明确被测试者在面对不同问题时候的表情变化。比如第三题为准绳问题，用户做出了相关反应，则第四题应当选择相关问题，或者中性问题，以测试与测试对象回答第三题反应的差别。相邻题目设置为不同类别，可以使得用户有较明显的情绪变化。

每种类型的问题分配有一个类型标签，例如可以I代表中性问题，C代表准绳问题，R代表相关问题，以问题为单位对测试对象的回答测试问卷的视频进行切割后，每个问题对应的视频片段被赋予该问题所属类型的标签。

上述特征提取步骤中，提取每个视频片段中的声音特征向量及表情特征向量，包括：

从视频片段中分离出语音及视频；

从分离得到的视频中提取测试对象的面部表情特征（包括面部动作-AU、眼球朝向、头部朝向），统计各面部表情特征在视频中出现的次数及持续的时长，根据统计结果构造表情高阶特征，利用特征筛选算法对所述表情高阶特征进行筛选出最优特征子集；

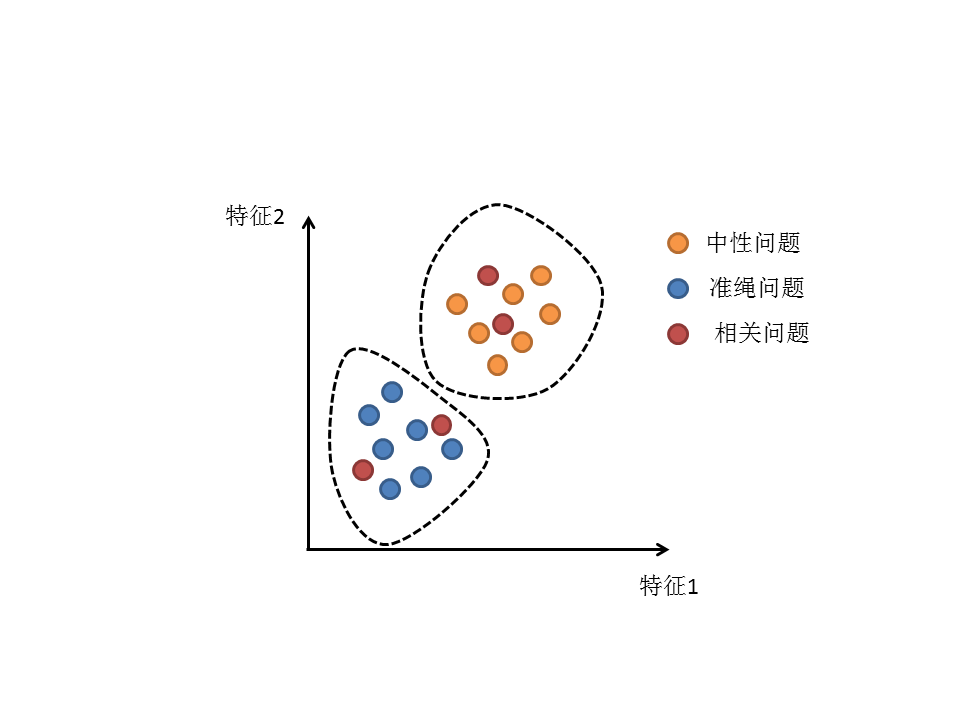
从分离得到的语音中提取低阶音频特征，从音频低阶特征中提取动态回归系数，得到动态音频特征，利用统计函数（如max-最大，min-最小, skewness-峰度， kurtosis-偏度）从所述低阶音频特征及动态音频特征中提取高阶音频特征，利用特征筛选算法从所述高阶音频特征中筛选出最优特征子集（具体方法参提案PAJRD20170275）；

对所述语音、视频的最优特征子集进行降维处理，得到二维空间中的声音特征向量及表情特征向量。

在本实施例中，采用tsne方法将高维（例如4170维）的音频特征及表情特征向低维空间投影降成2维，主要为了可是化展示。记声音特征向量为特征1、表情特征向量为特征2。在降维过程中，特征发生了根本性的变化，原始的特征消失了（虽然新的特征也保持了原特征的一些性质）。因而特征1与特征2并不像原先的高维（例如4170维）特征一样，有具体的物理意义。

通过tsne方法降维成2维，主要为了可是化展示。这里的2个特征是用原先4170维的特征向低维空间投影，所以维数减少了。在此过程中，特征发生了根本性的变化，原始的特征消失了。（虽然新的特征也保持了原特征的一些性质）。特征1与特征2不有具体的物理意义。它们是从原始高维特征中提取出来，能较好表达原始特征的两个特征。记声音特征向量为特征1、表情特征向量为特征2不能这样记录，他们只是坐标的位置，没有含义。

以特征1、特征2为X、Y坐标轴，聚类的示意图如下：



上述计算步骤中，每种类型问题的中心点特征向量，可以利用K-means聚类算法计算得到，也可以计算每种类型问题的特征向量(声音特征向量，表情特征向量)的均值、以均值为该类型问题的中心点特征向量。假设，一共有n1个中性问题、n2个准绳问题、n3个相关问题，计算得到该n1个中性问题的中心点特征向量为（xi,yi），该n2个准绳问题的中心点特征向量为（xc,yc）,则每个相关问题的特征向量（xr,yr）与中性问题、准绳问题的的中心点特征向量之间的欧式距离为：

若d\_ri>d\_rc，则表明测试对象在回答该相关问题时，其呈现出来的情绪波动及紧张程度与回答中性问题时的情绪波动及紧张程度差异较大，与回答准绳问题时的的情绪波动及紧张程度差异较小，则判断测试对象在回答该相关问题时隐瞒了真实情绪，欺诈嫌疑较大；

若d\_ri<d\_rc，则表明测试对象在回答该相关问题时，其呈现出来的情绪波动及紧张程度与回答中性问题时的情绪波动及紧张程度差异较小，与回答准绳问题时的情绪波动及紧张程度差异较大，则判断测试对象回答该相关问题时表现出的情绪是真实的，欺诈可能性较小；

若d\_ri=d\_rc，则表示测试对象回答该相关问题时，其呈现出来的情绪波动及紧张程度与回答中性问题、准绳问题时的情绪波动及紧张程度差异相同，无法判断是否隐瞒了真实情绪，但是d\_ri=d\_rc出现的可能性非常小。