

权 利 要 求 书

- 1、一种声音变换系统，其中，包括：
语音识别模型，至少包括瓶颈层，所述语音识别模型被配置为：将输入的源语音的梅尔倒谱特征，通过所述瓶颈层转换为瓶颈特征；并将所述瓶颈特征通过所述瓶颈层输入至注意力变声网络；
所述注意力变声网络被配置为：将输入的瓶颈特征转换为与目标语音相符的梅尔倒谱特征；
神经网络声码器，所述神经网络声码器被配置为：将与目标语音相符的梅尔倒谱特征转换为仿目标语音并输出；所述仿目标语音为源语音通过转换生成的音频。
- 2、根据权利要求 1 所述的声音变换系统，其中，所述语音识别模型包括五层 DNN 结构，其中第三层或者第四层为所述瓶颈层。
- 3、根据权利要求 1 所述的声音变换系统，其中，所述注意力变声网络包括一层双向 RNN 结构。
- 4、一种声音变换系统的训练方法，应用于权利要求 1-3 任一项所述的声音变换系统，其中，包括：
根据多人语音识别训练语料中的文字与字符编码之间的映射关系，将所述多人语音识别训练语料中的文字转换为编号；
将转换而得的编号和所述多人语音识别训练语料的梅尔倒谱特征输入至所述语音识别模型；
运行反向传播算法；
进行迭代优化至所述语音识别模型收敛，以训练所述语音识别模型。
- 5、根据权利要求 4 所述的声音变换系统的训练方法，其中，所述方法还包括对所述多人语音识别训练语料进行预处理；预处理包括去空白和归一化。
- 6、根据权利要求 4 所述的声音变换系统的训练方法，其中，所述方法还包括：
将目标语音的梅尔倒谱特征转换为瓶颈特征；
将所述目标语音的瓶颈特征输入至注意力变声网络；将所述目标语音的梅尔倒谱特征作为真实值输入至所述注意力变声网络；
使用深度迁移学习的方法训练所述注意力变声网络。
- 7、根据权利要求 6 所述的声音变换系统的训练方法，其中，将目标语音的梅尔倒谱特征转换为瓶颈特征的步骤，由预先训练的所述语音识别模型进行。
- 8、根据权利要求 6 所述的声音变换系统的训练方法，其中，所述方法还包括：
将所述目标语音的梅尔倒谱特征以及所述目标语音输入至神经网络声码器；
使用深度迁移学习的方法训练所述神经网络声码器。
- 9、根据权利要求 6-8 任一项所述的声音变换系统的训练方法，其中，所述方法还包括对所述目标语音进行预处理；预处理包括去空白和归一化。
- 10、根据权利要求 4 所述的声音变换系统的训练方法，其中，所述方法还包括：
参数提取，以得到多人语音识别训练语料的梅尔倒谱特征、目标语音的梅尔倒谱特征以及源语音的梅尔倒谱特征。