

权 利 要 求 书

1. 一种流式端到端语音识别方法，其特征在于，包括：
- 以帧为单位对接收到的语音流进行语音声学特征提取并进行编码；
- 对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行编码输出的激活
- 5 点数量进行预测；
- 根据预测结果确定需要进行解码输出的激活点所在的位置，以便解码器在所述激活点所在的位置进行解码并输出识别结果。
2. 根据权利要求 1 所述的方法，其特征在于，
- 所述分块中包括一帧语音流对应的编码结果；
- 10 所述预测结果包括：当前分块中是否包含需要进行编码输出的激活点；
- 所述根据预测结果确定需要进行解码输出的激活点所在的位置，包括：
- 将包含所述激活点的分块所在的位置确定为所述激活点所在的位置。
3. 根据权利要求 2 所述的方法，其特征在于，还包括：
- 分别确定各帧编码结果的注意力 Attention 系数；所述 Attention 系数用于描述对
- 15 应帧需要进行解码输出的概率；
- 根据所述 Attention 系数对所述预测结果进行验证。
4. 根据权利要求 1 所述的方法，其特征在于，
- 所述分块中包括多帧语音流对应的编码结果；
- 所述方法还包括：
- 20 分别确定各帧编码结果的注意力 Attention 系数；所述 Attention 系数用于描述对
- 应帧需要进行解码输出的概率；
- 根据预测结果确定需要进行解码输出的激活点所在的位置，包括：
- 将同一分块中各帧的 Attention 系数进行比对，并按照大小进行排序；
- 根据分块中包含的激活点数量，将该分块包含的各帧编码结果中 Attention 系数最
- 25 高的对应数量的帧所在的位置，确定为所述激活点所在的位置。
5. 根据权利要求 4 所述的方法，其特征在于，还包括：
- 根据已预测出的激活点的出现频度，对分块的大小进行自适应调节。
6. 根据权利要求 1 至 5 任一项所述的方法，其特征在于，
- 对编码结果进行分块处理，包括：
- 30 将所述编码结果进行缓存；
- 在加入缓存的编码结果帧数达到分块的大小时，将当前缓存的各帧编码结果确定为一个分块。
7. 根据权利要求 6 所述的方法，其特征在于，还包括：
- 在完成对该分块的预测处理后，将该分块的各帧编码结果从缓存中删除。

8. 一种建立预测模型的方法，其特征在于，包括：

获得训练样本集，所述训练样本集中包括多个分块数据以及标注信息，其中，每个分块数据帧包括对语音流的多帧分别进行编码的编码结果，所述标注信息包括每个分块中包括的需要进行解码输出的激活点的数量；

5 将所述训练样本集输入到预测模型中进行模型的训练。

9. 根据权利要求 8 所述的方法，其特征在于，

所述训练样本集中包括同一建模单元对应的多帧语音流被划分到不同的分块中的情况。

10. 一种提供语音识别服务的方法，其特征在于，包括：

10 云服务系统接收到应用系统的调用请求后，接收所述应用系统提供的语音流；

以帧为单位对接收到的语音流进行语音声学特征提取并进行编码；

对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行编码输出的激活点数量进行预测；

15 根据预测结果确定需要进行解码输出的激活点所在的位置，以便解码器在所述激活点所在的位置进行解码获得语音识别结果；

将所述语音识别结果返回给所述应用系统。

11. 一种获得语音识别信息的方法，其特征在于，包括：

应用系统通过调用云服务系统提供的接口，向所述云服务系统提交调用请求以及待识别的语音流，由所述云服务系统以帧为单位对接收到的语音流进行语音声学特征提取
20 并进行编码，对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行编码输出的激活点数量进行预测；根据预测结果确定需要进行解码输出的激活点所在的位置后，通过解码器在所述激活点所在的位置进行解码获得语音识别结果；

接收所述云服务系统返回的语音识别结果。

12. 一种法院自助立案实现方法，其特征在于，包括：

25 自助立案一体机设备接收语音输入的立案请求信息；

以帧为单位对接收到的语音流进行语音声学特征提取并进行编码；

对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行编码输出的激活点数量进行预测；

30 根据预测结果确定需要进行解码输出的激活点所在的位置，以便解码器在所述激活点所在的位置进行解码并确定识别结果；

将所述识别结果录入到关联的立案信息数据库中。

13. 一种终端设备升级方法，其特征在于，包括：

向终端设备提供升级建议信息；

35 接收到终端设备提交的升级请求后，为所述终端设备赋予使用升级后的方式进行流式语音识别的权限，所述升级后的方式进行流式语音识别包括：以帧为单位对接收到的

语音流进行语音声学特征提取并进行编码，对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行编码输出的激活点数量进行预测；根据预测结果确定需要进行解码输出的激活点所在的位置后，通过解码器在所述激活点所在的位置进行解码获得语音识别结果。

- 5 14. 根据权利要求 13 所述的方法，其特征在于，
所述终端设备包括智能音箱设备。
- 15 15. 根据权利要求 13 所述的方法，其特征在于，还包括：
根据所述终端设备提交的降级请求，为所述终端设备关闭所述使用升级后的方式进行流式语音识别的权限。
- 10 16. 一种流式端到端语音识别装置，其特征在于，包括：
编码单元，用于以帧为单位对接收到的语音流进行语音声学特征提取并进行编码；
预测单元，用于对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行
编码输出的激活点数量进行预测；
激活点位置确定单元，用于根据预测结果确定需要进行解码输出的激活点所在的位置，
15 以便解码器在所述激活点所在的位置进行解码并输出识别结果。
17. 一种建立预测模型的装置，其特征在于，包括：
训练样本集获得单元，用于获得训练样本集，所述训练样本集中包括多个分块数据
以及标注信息，其中，每个分块数据帧包括对语音流的多帧分别进行编码的编码结果，
所述标注信息包括每个分块中包括的需要进行解码输出的激活点的数量；
- 20 输入单元，用于将所述训练样本集输入到预测模型中进行模型的训练。
18. 一种提供语音识别服务的装置，其特征在于，应用于云服务系统，包括：
语音流接收单元，用于接收到应用系统的调用请求后，接收所述应用系统提供的语音流；
编码单元，用于以帧为单位对接收到的语音流进行语音声学特征提取并进行编码；
- 25 预测单元，用于对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行
编码输出的激活点数量进行预测；
激活点位置确定单元，用于根据预测结果确定需要进行解码输出的激活点所在的位置，
以便解码器在所述激活点所在的位置进行解码获得语音识别结果；
识别结果返回单元，用于将所述语音识别结果返回给所述应用系统。
- 30 19. 一种获得语音识别信息的装置，其特征在于，应用于应用系统，包括：
提交单元，用于通过调用云服务系统提供的接口，向所述云服务系统提交调用请求
以及待识别的语音流，由所述云服务系统以帧为单位对接收到的语音流进行语音声学特征
提取并进行编码，对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行
编码输出的激活点数量进行预测；根据预测结果确定需要进行解码输出的激活点所在
35 位置后，通过解码器在所述激活点所在的位置进行解码获得语音识别结果；

- 识别结果接收单元，用于接收所述云服务系统返回的语音识别结果。
20. 一种法院自助立案实现装置，其特征在于，应用于自助立案一体机设备，包括：
- 请求接收单元，用于接收语音输入的立案请求信息；
- 编码单元，用于以帧为单位对接收到的语音流进行语音声学特征提取并进行编码；
- 5 预测单元，用于对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行编码输出的激活点数量进行预测；
- 激活点位置确定单元，用于根据预测结果确定需要进行解码输出的激活点所在的位置，以便解码器在所述激活点所在的位置进行解码并确定识别结果；
- 信息录入单元，用于将所述识别结果录入到关联的立案信息数据库中。
- 10 21. 一种终端设备升级装置，其特征在于，包括：
- 升级建议提供单元，用于向终端设备提供升级建议信息；
- 权限授予单元，用于接收到终端设备提交的升级请求后，为所述终端设备赋予使用升级后的方式进行流式语音识别的权限，所述升级后的方式进行流式语音识别包括：以
- 15 帧为单位对接收到的语音流进行语音声学特征提取并进行编码，对已完成编码的帧进行分块处理，并对同一分块中包含的需要进行编码输出的激活点数量进行预测；根据预测结果确定需要进行解码输出的激活点所在的位置后，通过解码器在所述激活点所在的位置进行解码获得语音识别结果。
22. 一种计算机可读存储介质，其上存储有计算机程序，其特征在于，该程序被处理器执行时实现权利要求 1 至 15 任一项所述的方法的步骤。
- 20 23. 一种电子设备，其特征在于，包括：
- 一个或多个处理器；以及
- 与所述一个或多个处理器关联的存储器，所述存储器用于存储程序指令，所述程序指令在被所述一个或多个处理器读取执行时，执行权利要求 1 至 15 任一项所述的方法的步骤。