

权利要求

- 1、一种动态手势识别方法，其特征在于，包括：
对待检测的视频流中的动态手势进行定位，得到动态手势框；
从所述视频流的多帧图像帧中截取与所述动态手势框对应的图像块；
基于所述截取出的图像块生成检测序列；
根据所述检测序列进行动态手势识别。
- 2、根据权利要求1所述的方法，其特征在于，所述对待检测的视频流中的动态手势进行定位，得到动态手势框，包括：
对所述视频流的多帧图像中的至少一帧图像帧进行静态手势定位，得到所述至少一帧图像帧的静态手势框；
根据所述至少一帧图像帧的静态手势框确定所述动态手势框。
- 3、根据权利要求2所述的方法，其特征在于，所述根据所述至少一帧图像帧的静态手势框确定所述动态手势框，包括：
对所述至少一帧图像帧的静态手势框进行放大处理，得到所述动态手势框。
- 4、根据权利要求2或3所述的方法，其特征在于，所述视频流的多帧图像帧中至少一帧图像帧的静态手势框满足：
所述静态手势框位于所述动态手势框内，或者，所述静态手势框与所述动态手势框相同。
- 5、根据权利要求1-4任一项所述的方法，其特征在于，所述根据所述检测序列进行动态手势识别，包括：
确定所述检测序列中多个帧间图像差；
基于所述多个帧间图像差生成图像差序列；
根据所述检测序列和所述图像差序列进行动态手势识别。
- 6、根据权利要求5所述的方法，其特征在于，所述帧间图像差为所述检测序列中两个相邻参考帧之间的图像差。
- 7、根据权利要求5或6所述的方法，其特征在于，根据所述检测序列和所述图像差序列进行动态手势识别，包括：
将所述检测序列输入第一动态手势识别模型，以获得所述第一动态手势识别模型输出的第一动态手势类别预测概率；
将所述图像差序列输入第二动态手势识别模型，以获得所述第二动态手势识别模型输出的第二动态手势类别预测概率；
根据所述第一动态手势类别预测概率和所述第二动态手势类别预测概率，确定动态手势识别结果。
- 8、根据权利要求7所述的方法，其特征在于，所述第一动态手势识别模型为第一神经网络，所述第二动态手势识别模型为第二神经网络，所述第一神经网络和所述第二神经网络的结构相同或不同。
- 9、根据权利要求5-8任一项所述的方法，其特征在于，还包括：
多次截取获得所述检测序列，并多次生成所述图像差序列，以及多次根据所述检测序列以及所述图像差序列进行动态手势识别；
根据每次所述动态手势识别得到的动态手势类型的概率，确定动态手势识别结果。
- 10、根据权利要求7或8项所述的方法，其特征在于，所述根据所述检测序列和所述图像差序列进行动态手势识别之前，还包括：采用以下方法建立所述第一动态手势识别模型：
采集不同动态手势类型的样本视频流；
对所述不同动态手势类型的动态手势框进行标记；
从样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；
以所述动态手势类型作为监督数据，以所述图像序列作为训练数据，训练所述第一动态手势识别模

型。

11、根据权利要求 10 所述的方法，其特征在于，所述以所述动态手势类型作为监督数据，以所述图像序列作为训练数据，训练所述第一动态手势识别模型，包括：

将所述图像序列分为至少一段；

从所述至少一段中抽取预置帧数的图像，堆叠组成图像训练数据；

以所述动态手势类型作为监督数据，以所述图像训练数据训练所述第一动态手势识别模型。

12、根据权利要求 7、8、10、11 任一项所述的方法，其特征在于，所述根据所述检测序列和所述图像差序列进行动态手势识别之前，还包括：采用以下方法建立所述第二动态手势识别模型：

采集不同动态手势类型的样本视频流；

对所述不同动态手势类型的动态手势框进行标记；

从所述样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；

确定所述图像序列中多个帧间图像差；

基于所述多个帧间图像差生成图像差序列；

以所述动态手势类型作为监督数据，以所述图像差序列作为训练数据，训练所述第二动态手势识别模型。

13、根据权利要求 12 所述的方法，其特征在于，所述以所述动态手势类型作为监督数据，以所述图像差序列作为训练数据，训练出所述第二动态手势识别模型，包括：

将所述图像差序列分为至少一段；

从所述至少一段中抽取预置帧数的图像，堆叠组成图像差训练数据；

以所述动态手势类型作为监督数据，以所述图像差训练数据训练所述第二动态手势识别模型。

14、根据权利要求 1-13 任一项所述的方法，其特征在于，所述动态手势的类型包括以下之一或任意组合：挥手手势、点击手势、打手枪手势、抓取手势。

15、一种动态手势识别建模方法，其特征在于，包括：

采集不同动态手势类型的样本视频流；

对所述不同动态手势类型的动态手势框进行标记；

从样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；

以动态手势类型作为监督数据，以所述图像序列作为训练数据，训练第一动态手势识别模型。

16、根据权利要求 15 所述的方法，其特征在于，所述以动态手势类型作为监督数据，以所述图像序列作为训练数据，训练第一动态手势识别模型，包括：

将所述图像序列分为至少一段；

从所述至少一段中抽取预置帧数的图像，堆叠组成图像训练数据；

以所述动态手势类型作为监督数据，以所述图像训练数据训练所述第一动态手势识别模型。

17、根据权利要求 15 或 16 所述的方法，其特征在于，还包括：

采集不同动态手势类型的样本视频流；

对所述不同动态手势类型的动态手势框进行标记；

从所述样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；

确定所述图像序列中多个帧间图像差；

基于确定的多个帧间图像差生成图像差序列；

以所述动态手势类型作为监督数据，以所述图像差序列作为训练数据，训练所述第二动态手势识别模型。

18、根据权利要求 17 所述的方法，其特征在于，所述以所述动态手势类型作为监督数据，以所述图像差序列作为训练数据，训练出所述第二动态手势识别模型，包括：

将所述图像差序列分为至少一段；

从所述至少一段中抽取预置帧数的图像，堆叠组成图像差训练数据；

以所述动态手势类型作为监督数据，以所述图像差训练数据训练所述第二动态手势识别模型。

19、一种动态手势识别装置，其特征在于，包括：

手势定位单元，用于对待检测的视频流中的动态手势进行定位，得到动态手势框；

处理单元，用于从所述视频流的多帧图像帧中截取与所述动态手势框对应的图像块；

检测序列生成单元，用于基于所述截取出的图像块生成检测序列；

手势识别单元，用于根据所述检测序列进行动态手势识别。

20、根据权利要求 19 所述的装置，其特征在于，所述手势定位单元包括：

静态手势框定位单元，用于对所述视频流的多帧图像中的至少一帧图像帧进行静态手势定位，得到所述至少一帧图像帧的静态手势框；

动态手势框确定子单元，用于根据所述至少一帧图像帧的静态手势框确定所述动态手势框。

21、根据权利要求 20 所述的装置，其特征在于，所述动态手势框确定子单元用于：对所述至少一帧图像帧的静态手势框进行放大处理，得到所述动态手势框。

22、根据权利要求 20 或 21 所述的装置，其特征在于，所述视频流的多帧图像帧中至少一帧图像帧的静态手势框满足：所述静态手势框位于所述动态手势框内，或者，所述静态手势框与所述动态手势框相同。

23、根据权利要求 19-22 任一项所述的装置，其特征在于，所述手势识别单元包括：

图像差确定子单元，用于确定所述检测序列中多个帧间图像差；

图像差序列确定子单元，用于基于所述多个帧间图像差生成图像差序列；

动态手势识别子单元，用于根据所述检测序列和所述图像差序列进行动态手势识别。

24、根据权利要求 23 所述的装置，其特征在于，所述帧间图像差为所述检测序列中两个相邻参考帧之间的图像差。

25、根据权利要求 23 或 24 所述的装置，其特征在于，所述动态手势识别子单元用于：将所述检测序列输入第一动态手势识别模型，以获得所述第一动态手势识别模型输出的第一动态手势类别预测概率；将所述图像差序列输入第二动态手势识别模型，以获得所述第二动态手势识别模型输出的第二动态手势类别预测概率；根据所述第一动态手势类别预测概率和所述第二动态手势类别预测概率，确定动态手势识别结果。

26、根据权利要求 25 所述的装置，其特征在于，所述第一动态手势识别模型为第一神经网络，所述第二动态手势识别模型为第二神经网络，所述第一神经网络和所述第二神经网络的结构相同或不同。

27、根据权利要求 23-26 任一项所述的装置，其特征在于，所述手势识别单元还包括：

多次识别控制单元，用于多次截取获得所述检测序列，并多次生成所述图像差序列，以及多次根据所述检测序列以及所述图像差序列进行动态手势识别；

识别结果确定单元，用于根据每次所述动态手势识别得到的动态手势类型的概率，确定动态手势识别结果。

28、根据权利要求 25 或 26 所述的装置，其特征在于，所述手势识别单元还包括：第一动态手势识别模型建立单元；所述第一动态手势识别模型建立单元包括：

样本采集子单元，用于采集不同动态手势类型的样本视频流；

手势框标记子单元，用于对所述不同动态手势类型的动态手势框进行标记；

图像序列构成子单元，用于从样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；

训练子单元，用于以所述动态手势类型作为监督数据，以所述图像序列作为训练数据，训练所述第一动态手势识别模型。

29、根据权利要求 28 所述的装置，其特征在于，所述训练子单元用于：将所述图像序列分为至少

一段；从所述至少一段中抽取预置帧数的图像，堆叠组成图像训练数据；以及，以所述动态手势类型作为监督数据，以所述图像训练数据训练所述第一动态手势识别模型。

30、根据权利要求 25、26、28、29 任一项所述的装置，其特征在于，所述手势识别单元还包括：第二动态手势识别模型建立单元；所述第二动态手势识别模型建立单元包括：

样本采集子单元，用于采集不同动态手势类型的样本视频流；

手势框标记子单元，用于对所述不同动态手势类型的动态手势框进行标记；

图像序列构成子单元，用于从所述样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；

图像差确定子单元，用于确定所述图像序列中多个帧间图像差；

图像差序列确定子单元，用于基于所述多个帧间图像差生成图像差序列；

训练子单元，用于以所述动态手势类型作为监督数据，以所述图像差序列作为训练数据，训练所述第二动态手势识别模型。

31、根据权利要求 30 所述的装置，其特征在于，所述训练子单元用于：将所述图像差序列分为至少一段；从所述至少一段中抽取预置帧数的图像，堆叠组成图像差训练数据；以及，以所述动态手势类型作为监督数据，以所述图像差训练数据训练所述第二动态手势识别模型。

32、根据权利要求 14-31 任一项所述的装置，其特征在于，所述动态手势的类型包括以下之一或任意组合：挥手手势、点击手势、打手枪手势、抓取手势。

33、一种动态手势识别模型建立装置，其特征在于，包括：第一动态手势识别模型建立单元；所述第一动态手势识别模型建立单元包括：

样本采集子单元，用于采集不同动态手势类型的样本视频流；

手势框标记子单元，用于对所述不同动态手势类型的动态手势框进行标记；

图像序列构成子单元，用于从样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；

训练子单元，用于以动态手势类型作为监督数据，以所述图像序列作为训练数据，训练第一动态手势识别模型。

34、根据权利要求 33 所述的装置，其特征在于，所述训练子单元用于：将所述图像序列分为至少一段；从所述至少一段中抽取预置帧数的图像，堆叠组成图像训练数据；以及，以所述动态手势类型作为监督数据，以所述图像训练数据训练所述第一动态手势识别模型。

35、根据权利要求 33 或 34 所述的装置，其特征在于，还包括：第二动态手势识别模型建立单元；所述第二动态手势识别模型建立单元包括：

样本采集子单元，用于采集不同动态手势类型的样本视频流；

手势框标记子单元，用于对所述不同动态手势类型的动态手势框进行标记；

图像序列构成子单元，用于从所述样本视频流的多帧图像帧中截取与动态手势框的标注信息对应的图像块，构成图像序列；

图像差确定子单元，用于确定所述图像序列中多个帧间图像差；

图像差序列确定子单元，用于基于确定的多个帧间图像差生成图像差序列；

训练子单元，用于以所述动态手势类型作为监督数据，以所述图像差序列作为训练数据，训练所述第二动态手势识别模型。

36、根据权利要求 35 所述的装置，其特征在于，所述训练子单元用于：将所述图像差序列分为至少一段；从所述至少一段中抽取预置帧数的图像，堆叠组成图像差训练数据；以及，以所述动态手势类型作为监督数据，以所述图像差训练数据训练所述第二动态手势识别模型。

37、一种手势交互控制方法，其特征在于，

获取视频流；

采用如权利要求 1 至 14 中任一项所述的方法确定所述视频流中的动态手势识别结果；

控制设备执行与所述动态手势识别结果相应的操作。

38、根据权利要求 37 所述的方法，其特征在于，所述控制设备执行与所述动态手势识别结果相应的操作，包括：

根据预先确定的动态手势识别结果与操作指令之间的对应关系，获取与所述动态手势识别结果对应的操作指令；

根据所述操作指令控制设备执行相应操作。

39、根据权利要求 38 所述的方法，其特征在于，所述根据所述操作指令控制设备执行相应操作，包括：

根据所述操作指令对车辆的车窗、车门或车载系统进行控制。

40、根据权利要求 37 所述的方法，其特征在于，所述控制设备执行与所述动态手势识别结果相应的操作，包括：

响应于所述动态手势识别结果为预定义动态动作，控制车辆执行与所述预定义动态动作对应的操作。

41、根据权利要求 40 所述的方法，其特征在于，所述预定义动态动作包括动态手势，所述动态手势包括以下至少之一：单指顺/逆时针旋转、手掌左/右挥动、两指前戳、拇指和小指伸出、手掌朝向下压、手掌朝上向上抬起、手掌向左/右扇风、拇指伸出左/右移动、手掌左/右长滑、掌心向上拳变掌、掌心向上掌变拳、掌心向下掌变拳、掌心向下拳变掌、单指滑动、多指向内捏住、单指双击、单指单击、多指双击、多指单击；

与所述预定义动态动作对应的操作包括以下至少之一：调节音量大/小，歌曲切换，歌曲暂停/继续，接听或启动电话、挂断或拒接电话、空调温度升高或降低、多屏互动、打开天窗、关闭天窗、锁紧门锁、解锁门锁、拖动导航、缩小地图、放大地图。

42、一种手势交互控制装置，其特征在于，所述装置包括：

视频流获取模块，用于获取视频流；

结果获取模块，用于采用如权利要求 19 至 32 中任一项所述的装置确定所述视频流中的动态手势识别结果；

操作执行模块，用于控制设备执行与所述动态手势识别结果相应的操作。

43、根据权利要求 42 所述的装置，其特征在于，所述操作执行模块，包括：

操作指令获取子模块，用于根据预先确定的动态手势识别结果与操作指令之间的对应关系，获取与所述动态手势识别结果对应的操作指令；

操作执行子模块，用于根据所述操作指令控制设备执行相应操作。

44、根据权利要求 43 所述的装置，其特征在于，所述操作执行子模块，用于：

根据所述操作指令对车辆的车窗、车门或车载系统进行控制。

45、根据权利要求 42 所述的装置，其特征在于，所述操作执行模块，还用于：

响应于所述检测结果为预定义动态动作，控制车辆执行与所述预定义动态动作对应的操作。

46、根据权利要求 45 所述的装置，其特征在于，所述预定义动态动作包括动态手势，所述动态手势包括以下至少之一：单指顺/逆时针旋转、手掌左/右挥动、两指前戳、拇指和小指伸出、手掌朝向下压、手掌朝上向上抬起、手掌向左/右扇风、拇指伸出左/右移动、手掌左/右长滑、掌心向上拳变掌、掌心向上掌变拳、掌心向下掌变拳、掌心向下拳变掌、单指滑动、多指向内捏住、单指双击、单指单击、多指双击、多指单击；

与所述预定义动态动作对应的操作包括以下至少之一：调节音量大/小，歌曲切换，歌曲暂停/继续，接听或启动电话、挂断或拒接电话、空调温度升高或降低、多屏互动、打开天窗、关闭天窗、锁紧门锁、解锁门锁、拖动导航、缩小地图、放大地图。

47、一种电子设备，其特征在于，包括处理器，所述处理器包括权利要求 19 至 32 任意一项所述的

动态手势识别装置，或权利要求 33 至 36 任意一项所述的动态手势识别建模装置，或权利要求 42 至 46 任意一项所述的手势交互控制装置。

48、一种电子设备，其特征在于，包括：存储器，用于存储可执行指令；

以及处理器，用于与所述存储器通信以执行所述可执行指令从而完成权利要求 1 至 14 任意一项所述动态手势识别方法或权利要求 15 至 18 任意一项所述的动态手势识别建模方法或权利要求 37 至 41 任意一项所述的手势交互控制方法的操作。

49、一种计算机可读存储介质，用于存储计算机可读取的指令，其特征在于，所述指令被执行时执行权利要求 1 至 14 任意一项所述动态手势识别方法或权利要求 15 至 18 任意一项所述的动态手势识别建模方法或权利要求 37 至 41 任意一项所述的手势交互控制方法的操作。

50、一种计算机程序产品，包括计算机可读代码，其特征在于，当所述计算机可读代码在设备上运行时，所述设备中的处理器执行用于实现权利要求 1 至 14 任意一项所述动态手势识别方法或权利要求 15 至 18 任意一项所述的动态手势识别建模方法或权利要求 37 至 41 任意一项所述的手势交互控制方法的指令。