

行人再识别方法、装置、电子设备和存储介质

相关申请的交叉引用

本申请基于申请号为 201810145717.3、申请日为 2018 年 2 月 12 日的中国专利申请提出，并要求该中国专利申请的优先权，该中国专利申请的全部内容在此以引入方式并入本申请。

技术领域

本申请实施例涉及图像处理技术领域，尤其涉及一种行人再识别方法、装置、电子设备和存储介质。

背景技术

行人再识别是智能视频监控系统的一项关键技术，它旨在通过对给定目标视频和候选视频之间的相似性进行度量，进而在大量候选视频中找出与目标视频中包含同一行人的候选视频。

目前的行人再识别方法主要将一段完整的视频进行编码，利用编码结果对整段目标视频和整段候选视频之间的相似性进行度量，行人再识别的效果差。

发明内容

本申请实施例提供了行人再识别技术方案。

根据本申请实施例的第一方面，提供了一种行人再识别方法，包括：获取包含目标行人的目标视频和至少一个候选视频；对所述目标视频中的每个目标视频片段和至少一个所述候选视频中的每个候选视频片段分别进行编码；根据编码结果确定每个所述目标视频片段和每个所述候选视频片段之间的相似性分值；所述相似性分值用于表征所述目标视频片段与所述候选视频片段中的行人特征的相似程度；根据所述相似性分值对至少一个所述候选视频进行行人再识别。

在一实施例中，所述对所述目标视频中的每个目标视频片段和至少一个所述候选视频中的每个候选视频片段分别进行编码，包括：获取每个所述目标视频片段中的每个目标视频帧的第一目标特征向量和第二目标特征向量以及每个所述目标视频片段的索引特征向量，获取每个所述候选视频片段中的每个候选视频帧的第一候选特征向量和第二候选特征向量；根据

所述索引特征向量、所述第一目标特征向量和所述第一候选特征向量生成注意权重向量；根据所述注意权重向量、所述第二目标特征向量和所述第二候选特征向量获得每个所述目标视频片段的编码结果和每个所述候选视频片段的编码结果。

5 在一实施例中，所述获取每个所述目标视频片段中的每个目标视频帧的第一目标特征向量和第二目标特征向量以及每个所述目标视频片段的索引特征向量，获取每个所述候选视频片段中的每个候选视频帧的第一候选特征向量和第二候选特征向量，包括：分别提取每个所述目标视频帧的图像特征向量和每个所述候选视频帧的图像特征向量；根据每个所述目标视频帧的图像特征向量生成每个所述目标视频帧的第一目标特征向量和第二目标特征向量以及每个所述目标视频片段的索引特征向量，根据每个所述候选视频帧的图像特征向量生成每个所述候选视频帧的第一候选特征向量和第二候选特征向量。

10 在一实施例中，所述根据所述索引特征向量、所述第一目标特征向量和所述第一候选特征向量生成注意权重向量，包括：根据所述索引特征向量和所述第一目标特征向量生成每个所述目标视频帧的目标注意权重向量，根据所述索引特征向量和所述第一候选特征向量生成每个所述候选视频帧的候选注意权重向量。

15 在一实施例中，所述根据所述索引特征向量和所述第一目标特征向量生成每个所述目标视频帧的目标注意权重向量，包括：根据所述索引特征向量、每个所述目标视频帧的所述第一目标特征向量生成每个所述目标视频帧的目标热度图；对所述目标热度图进行归一化处理得到每个所述目标视频帧的目标注意权重向量；和/或，所述根据所述索引特征向量和所述第一候选特征向量生成每个所述候选视频帧的候选注意权重向量，包括：根据所述索引特征向量、每个所述候选视频帧的所述第一候选特征向量生成每个所述候选视频帧的候选热度图；对所述候选热度图进行归一化处理得到每个所述候选视频帧的候选注意权重向量。

20 在一实施例中，所述根据所述注意权重向量、所述第二目标特征向量和所述第二候选特征向量获得每个所述目标视频片段的编码结果和每个所述候选视频片段的编码结果，包括：根据每个所述目标视频帧的目标注意权重向量和第二目标特征向量获得每个所述目标视频片段的编码结果，根据每个所述候选视频帧的候选注意权重向量和第二候选特征向量获得每个所述候选视频片段的编码结果。

25 在一实施例中，所述根据每个所述目标视频帧的目标注意权重向量和第二目标特征向量获得每个所述目标视频片段的编码结果，包括：将每个所述目标视频帧的目标注意权重向量与各自目标视频帧的第二目标特征向量相乘；将每个所述目标视频帧的相乘结果在时间维度相加，得到每个所述目标视频片段的编码结果；和/或，根据每个所述目标视频帧的候选注意

权重向量和第二候选特征向量获得每个所述候选视频片段的编码结果，包括：将每个所述候选视频帧的候选注意权重向量与各自候选视频帧的第二候选特征向量相乘；将每个所述候选视频帧的相乘结果在时间维度相加，得到每个所述候选视频片段的编码结果。

5 在一实施例中，所述根据编码结果确定每个所述目标视频片段和每个所述候选视频片段之间的相似性分值，包括：将每个所述目标视频片段的编码结果与每个所述候选视频片段的编码结果依次进行相减操作；将相减操作的结果在每一个维度上进行平方操作；对平方操作得到的特征向量进行全连接操作得到二维的特征向量；将所述二维的特征向量进行归一化操作，
10 得到每个所述目标视频片段和每个所述候选视频片段之间的相似性分值。

在一实施例中，所述根据所述相似性分值对至少一个所述候选视频进行行人再识别，包括：针对至少一个所述候选视频中的每个所述候选视频片段，将分值最高的预设比例阈值的所述相似性分值相加，作为每个所述
15 候选视频的相似性分值；将每个所述候选视频的相似性分值按照降序进行排列；将排列在前面的一个或者几个所述候选视频确定为与所述目标视频包含同一目标行人的视频。

根据本申请实施例的第二方面，提供了一种行人再识别装置，包括：获取模块，配置为获取包含目标行人的目标视频和至少一个候选视频；编
20 码模块，配置为对所述目标视频中的每个目标视频片段和至少一个所述候选视频中的每个候选视频片段分别进行编码；确定模块，配置为根据编码结果确定每个所述目标视频片段和每个所述候选视频片段之间的相似性分值；所述相似性分值用于表征所述目标视频片段与所述候选视频片段中的行人特征的相似程度；识别模块，配置为根据所述相似性分值对至少一个
25 所述候选视频进行行人再识别。

在一实施例中，所述编码模块，包括：特征向量获取模块，配置为获取每个所述目标视频片段中的每个目标视频帧的第一目标特征向量和第二目标特征向量以及每个所述目标视频片段的索引特征向量，获取每个所述
30 候选视频片段中的每个候选视频帧的第一候选特征向量和第二候选特征向量；权重向量生成模块，配置为根据所述索引特征向量、所述第一目标特征向量和所述第一候选特征向量生成注意权重向量；编码结果获取模块，配置为根据所述注意权重向量、所述第二目标特征向量和所述第二候选特征向量获得每个所述目标视频片段的编码结果和每个所述候选视频片段的编码结果。

35 在一实施例中，所述特征向量获取模块，配置为分别提取每个所述目标视频帧的图像特征向量和每个所述候选视频帧的图像特征向量；根据每个所述目标视频帧的图像特征向量生成每个所述目标视频帧的第一目标特征向量和第二目标特征向量以及每个所述目标视频片段的索引特征向量，

根据每个所述候选视频帧的图像特征向量生成每个所述候选视频帧的第一候选特征向量和第二候选特征向量。

在一实施例中，所述权重向量生成模块，配置为根据所述索引特征向量和所述第一目标特征向量生成每个所述目标视频帧的目标注意权重向量，根据所述索引特征向量和所述第一候选特征向量生成每个所述候选视频帧的候选注意权重向量。

在一实施例中，所述权重向量生成模块，配置为根据所述索引特征向量、每个所述目标视频帧的所述第一目标特征向量生成每个所述目标视频帧的目标热度图；对所述目标热度图进行归一化处理得到每个所述目标视频帧的目标注意权重向量；和/或，根据所述索引特征向量、每个所述候选视频帧的所述第一候选特征向量生成每个所述候选视频帧的候选热度图；对所述候选热度图进行归一化处理得到每个所述候选视频帧的候选注意权重向量。

在一实施例中，所述编码结果获取模块，配置为根据每个所述目标视频帧的目标注意权重向量和第二目标特征向量获得每个所述目标视频片段的编码结果，根据每个所述候选视频帧的候选注意权重向量和第二候选特征向量获得每个所述候选视频片段的编码结果。

在一实施例中，所述编码结果获取模块，配置为将每个所述目标视频帧的目标注意权重向量与各自目标视频帧的第二目标特征向量相乘；将每个所述目标视频帧的相乘结果在时间维度相加，得到每个所述目标视频片段的编码结果；和/或，将每个所述候选视频帧的候选注意权重向量与各自候选视频帧的第二候选特征向量相乘；将每个所述候选视频帧的相乘结果在时间维度相加，得到每个所述候选视频片段的编码结果。

在一实施例中，所述确定模块，配置为将每个所述目标视频片段的编码结果与每个所述候选视频片段的编码结果依次进行相减操作；将相减操作的结果在每一个维度上进行平方操作；对平方操作得到的特征向量进行全连接操作得到二维的特征向量；将所述二维的特征向量进行归一化操作，得到每个所述目标视频片段和每个所述候选视频片段之间的相似性分值。

在一实施例中，所述识别模块，配置为针对至少一个所述候选视频中的每个所述候选视频片段，将分值最高的预设比例阈值的所述相似性分值相加，作为每个所述候选视频的相似性分值；将每个所述候选视频的相似性分值按照降序进行排列；将排列在前面的一个或者几个所述候选视频确定为与所述目标视频包含同一目标行人的视频。

根据本申请实施例的第三方面，提供了一种电子设备，包括：处理器和存储器；所述存储器用于存放至少一个可执行指令，所述可执行指令使所述处理器执行如第一方面所述的行人再识别方法。

根据本申请实施例的第四方面，提供了一种计算机可读存储介质，其上存储有计算机程序，该计算机程序被处理器执行时实现如第一方面所述

的行人再识别方法。

根据本申请实施例的第五方面，提供了一种计算机程序产品，包括：至少一个可执行指令，所述可执行指令被处理器执行时用于实现如第一方面所述的行人再识别方法。

- 5 本申请实施例在进行行人再识别时，获取包含目标行人的目标视频和至少一个候选视频，对目标视频中的每个目标视频片段和至少一个候选视频中的每个候选视频片段分别进行编码，根据编码结果确定每个目标视频片段和每个候选视频片段之间的相似性分值；根据相似性分值对至少一个候选视频进行行人再识别。由于视频片段包含的帧数远远少于整段视频包含的帧数，因此，视频片段中的行人表面信息的变化程度远远小于整段视频中的行人表面信息的变化程度。与对整段目标视频和整段候选视频进行编码相比，对每个目标视频片段和每个候选视频片段进行编码，有效减小了行人表面信息的变化，同时利用了不同视频帧内的行人表面信息的多样性和视频帧与视频帧之间动态相关性，提高了行人表面信息的利用率，提高了编码结果对每个目标视频片段和每个候选视频片段之间的相似性分值计算的准确率，进而可以提高了行人再识别的准确率。
- 10
- 15

附图说明

- 图 1 是根据本申请实施例行人再识别方法的一个实施例的流程示意图；
图 2 是根据本申请实施例行人再识别方法的一个实施例的计算框架示意图；
20 图 3 是根据本申请实施例行人再识别方法的另一个实施例的流程示意图；
图 4 是根据本申请实施例行人再识别方法中的注意编码机制示意图；
图 5 是根据本申请实施例行人再识别装置的一个实施例的结构示意图；
25 图 6 是根据本申请实施例行人再识别装置的另一个实施例的结构示意图；
图 7 是根据本申请实施例电子设备的一个实施例的结构示意图。

具体实施方式

- 下面结合附图（若干附图中相同的标号表示相同的元素）和实施例，
30 对本发明实施例的具体实施方式作进一步详细说明。以下实施例用于说明本发明，但不用来限制本发明的范围。

本领域技术人员可以理解，本发明实施例中的“第一”、“第二”等术语仅用于区别不同步骤、设备或模块等，既不代表任何特定技术含义，也不表示它们之间的必然逻辑顺序。

- 35 参照图 1，示出了根据本申请实施例行人再识别方法的一个实施例的流

程示意图。

本申请实施例的行人再识别方法通过电子设备的处理器调用存储器存储的相关指令执行以下步骤。

步骤 S100、获取包含目标行人的目标视频和至少一个候选视频。

5 本申请实施例中的目标视频可以包含一个或多个目标行人，候选视频中可以包含一个或多个候选行人或者不包含候选行人。本申请实施例中的目标视频和至少一个候选视频可以是来源于视频采集设备的视频图像，还可以来源于其他设备，本申请实施例的目的之一是从至少一个候选视频中

10 得到候选行人与目标行人为同一行人的候选视频。
在一个可选示例中，该步骤 S100 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的获取模块 50 执行。

步骤 S102、对目标视频中的每个目标视频片段和至少一个候选视频中的每个候选视频片段分别进行编码。

15 首先，对目标视频和候选视频进行视频片段切割，生成目标视频中的每个目标视频片段和候选视频中的每个候选视频片段，其中，每个目标视频片段具有固定的时间长度，每个候选视频片段具有固定的时间长度，而且，每个目标视频片段的时间长度与每个候选视频片段的时间长度可以相同也可以不相同。

20 然后，分别对每个目标视频片段和每个候选视频片段进行编码操作，得到每个目标视频片段的编码结果和每个候选视频片段的编码结果。

在一个可选示例中，该步骤 S102 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的编码模块 52 执行。

步骤 S104、根据编码结果确定每个目标视频片段和每个候选视频片段之间的相似性分值。

25 本申请实施例中，每个目标视频片段的编码结果可以认为是每个目标视频片段中的行人特征向量的一种表现形式，每个候选视频片段的编码结果可以认为是每个候选视频片段中的行人特征向量的一种表现形式。或者，编码结果即行人特征向量。若某个目标视频片段与某个候选视频片段之间的行人特征向量相同或者相近，则表示该目标视频片段与该候选视频片段包含同一目标行人的可能性较高，即该目标视频片段与该候选视频片段之间的相似性分值较高；若某个目标视频片段与某个候选视频片段之间的行人特征向量不相同，则表示该目标视频片段与该候选视频片段包含同一目标行人的可能性较低，即该目标视频片段与该候选视频片段之间的相似性

35 在一个可选示例中，该步骤 S104 可以由处理器调用存储器存储的相应指令执行，也可以由被处理器运行的确定模块 54 执行。

步骤 S106、根据相似性分值对至少一个候选视频进行行人再识别。

在得到每个目标视频片段与每个候选视频片段之间的相似性分值之

后, 可以根据相似性分值获得至少一个候选视频的相似性分值。将相似性分值较高的候选视频确定为包含与目标视频中具有同一目标行人的候选视频。

5 在一个可选示例中, 该步骤 S106 可以由处理器调用存储器存储的相应指令执行, 也可以由被处理器运行的识别模块 56 执行。

本申请实施例提出的行人再识别方法可以在如图 2 所示的计算框架下执行。首先, 对视频 (包括目标视频和至少一个候选视频) 进行切割, 生成具有固定长度的视频片段。其中 p 表示目标视频, g 表示至少一个候选视频中的其中一个候选视频, p_n 是目标视频 p 中的一个目标视频片段, g_k 是候选视频 g 中一个候选视频片段。为了衡量目标视频 p 和候选视频 g 中任意两个视频片段的相似性, 利用具有协同注意机制的深度网络。该深度网络以目标视频片段 p_n 和候选视频片段 g_k 作为输入项, 输出项 $m(p_n, g_k)$ 为目标视频片段 p_n 和候选视频片段 g_k 之间的相似性分值。对于目标视频 p 和候选视频 g 中的每两个视频片段 (目标视频片段和候选视频片段), 可以获得若干个视频片段之间的相似性分值。为了对目标视频 p 和候选视频 g 之间的相似性进行有效估计, 可以利用竞争性机制选择相似性较高的部分相似性分值, 通过对这些相似性分值的相加获得对目标视频 p 和候选视频 g 之间的相似性的可靠估计 $c(p, g)$ 。

20 本申请实施例在进行行人再识别时, 获取包含目标行人的目标视频和至少一个候选视频, 对目标视频中的每个目标视频片段和至少一个候选视频中的每个候选视频片段分别进行编码, 根据编码结果确定每个目标视频片段和每个候选视频片段之间的相似性分值; 根据相似性分值对至少一个候选视频进行行人再识别。由于视频片段包含的帧数远远少于整段视频包含的帧数, 因此, 视频片段中的行人表面信息的变化程度远远小于整段视频中的行人表面信息的变化程度。与对整段目标视频和整段候选视频进行编码相比, 对每个目标视频片段和每个候选视频片段进行编码, 有效减小了行人表面信息的变化, 同时利用了不同视频帧内的行人表面信息的多样性和视频帧与视频帧之间动态相关性, 提高了行人表面信息的利用率, 提高了编码结果对每个目标视频片段和每个候选视频片段之间的相似性分值计算的准确率, 进而可以提高了行人再识别的准确率。

参照图 3, 示出了根据本申请实施例行人再识别方法的另一个实施例的流程示意图。

需要说明的是, 本申请各实施例描述的部分均有所侧重, 某实施例未详尽描述的部分可参见本申请其他实施例中的介绍和说明, 不再赘述。

35 步骤 S300、获取包含目标行人的目标视频和至少一个候选视频。

步骤 S302、对目标视频中的每个目标视频片段和至少一个候选视频中的每个候选视频片段分别进行编码。

可选地, 本步骤 S302 可以包括如下步骤:

步骤 S3020、获取每个目标视频片段中的每个目标视频帧的第一目标特征向量和第二目标特征向量以及每个目标视频片段的索引特征向量，获取每个候选视频片段中的每个候选视频帧的第一候选特征向量和第二候选特征向量。

5 一种可选的实施方式中，可以利用神经网络提取每个目标视频帧的图像特征向量和每个候选视频帧的图像特征向量，图像特征向量用于反映视频帧中的图像特征，如行人特征、背景特征等等。针对目标视频帧，根据每个目标视频帧的图像特征向量生成每个目标视频帧的第一目标特征向量和第二目标特征向量以及每个目标视频片段的索引特征向量，索引特征向量
10 量包含了目标视频片段的信息，能够有效分辨有用信息与噪声信息。针对候选视频帧，根据每个候选视频帧的图像特征向量生成每个候选视频帧的第一候选特征向量和第二候选特征向量。具体地，可以根据每一帧特征线性变换生成第一目标特征向量（“键”特征向量）和第一候选特征向量（“键”特征向量），可以根据每一帧特征的另一个线性变换生成第二目标特征向量
15 （“值”特征向量）和第二候选特征向量（“值”特征向量），可以利用长短期记忆（Long Short-Term Memory, LSTM）网络和每个目标视频片段的每个目标视频帧的图像特征向量生成每个目标视频片段的索引特征向量，索引特征向量由目标视频片段生成，作用于目标视频片段自身以及所有的候选视频片段。

20 步骤 S3022、根据索引特征向量、第一目标特征向量和第一候选特征向量生成注意权重向量。

本申请实施例中，第一目标特征向量和第一候选特征向量用于生成注意权重向量。一种可选的实施方式中，针对目标视频帧，可以根据索引特征向量和第一目标特征向量生成每个目标视频帧的目标注意权重向量，可
25 选地，根据索引特征向量、每个目标视频帧的第一目标特征向量生成每个目标视频帧的目标热度图，具体地，根据索引特征向量、每个目标视频帧的第一目标特征向量进行内积操作得到每个目标视频帧的目标热度图；在时间维度上利用 softmax 函数对目标热度图进行归一化处理得到每个目标视频帧的目标注意权重向量。针对候选视频帧，可以根据索引特征向量和
30 第一候选特征向量生成每个候选视频帧的候选注意权重向量，可选地，根据索引特征向量、每个候选视频帧的第一候选特征向量生成每个候选视频帧的候选热度图，具体地，根据索引特征向量、每个候选视频帧的第一候选特征向量进行内积操作得到每个候选视频帧的候选热度图；在时间维度上利用 softmax 函数对候选热度图进行归一化处理得到每个候选视频帧的
35 候选注意权重向量。

注意权重向量用于在编码过程中增强有效的行人特征，是一种具有判别力信息的权重向量，能够减弱噪声信息的影响。

步骤 S3024、根据注意权重向量、第二目标特征向量和第二候选特征向

量获得每个目标视频片段的编码结果和候选视频片段的编码结果。

本申请实施例中，第二目标特征向量用于反映目标视频片段中的每一帧的图像特征，第二候选特征向量用于反映候选视频片段中的每一帧的图像特征。一种可选的实施方式中，针对目标视频帧，根据每个目标视频帧的目标注意权重向量和第二目标特征向量获得每个目标视频片段的编码结果。具体地，将每个目标视频帧的目标注意权重向量与各自目标视频帧的第二目标特征向量相乘；将每个目标视频帧的相乘结果在时间维度相加，得到每个目标视频片段的编码结果。针对候选视频帧，根据每个候选视频帧的候选注意权重向量和第二候选特征向量获得每个候选视频片段的编码结果。可选地，将每个候选视频帧的候选注意权重向量与各自候选视频帧的第二候选特征向量相乘；将每个候选视频帧的相乘结果在时间维度相加，得到每个候选视频片段的编码结果。

本申请实施例的步骤 S302 可以通过注意编码机制实现，即通过对视频片段（目标视频片段和候选视频片段）中不同帧特征的提炼而获得视频片段的编码结果，其过程如图 4 所示。首先对目标视频片段中的每一个目标视频帧和候选视频片段中的每一个候选视频帧提取卷积神经网络特征，根据卷积神经网络特征生成与每一个目标视频帧或每一个候选视频帧对应的“键”特征向量和“值”特征向量，每一个目标视频帧或每一个候选视频帧的“键”特征向量与每个目标视频片段的索引特征向量进行内积操作形成热度图，通过热度图反映目标视频帧或候选视频帧内的每一个特征与全局信息的相关性。将热度图在时间维度上利用 softmax 函数进行归一化操作形成注意权重向量，该注意权重向量与每一视频帧的“值”特征向量在每一个维度上对应相乘，并把不同视频帧获得的结果在时间维度进行相加，进而获得每一个视频片段的编码结果。

步骤 S304、根据编码结果确定每个目标视频片段和每个候选视频片段之间的相似性分值。

一种可选的实施方式中，将每个目标视频片段的编码结果与每个候选视频片段的编码结果依次进行相减操作、平方操作、全连接操作和归一化操作，得到每个目标视频片段和每个候选视频片段之间的相似性分值。具体地，将每个目标视频片段的编码结果与每个候选视频片段的编码结果依次进行相减操作，然后在每一个图像维度上进行平方操作，图像维度包括但不限于：行人图像维度和背景图像维度，其中，行人图像维度包括头部图像维度、上身图像维度、下身图像维度等；背景图像维度包括建筑图像维度、街道图像维度等。平方操作之后所得到的特征向量经过全连接层获得一个二维的特征向量，最后通过 Sigmoid 函数的非线性归一化得到每个目标视频片段和每个候选视频片段之间的相似性分值。

步骤 S306、根据相似性分值对至少一个候选视频进行行人再识别。

一种可选的实施方式中，针对至少一个候选视频中的每个候选视频，

将大于或等于预设阈值的相似性分值或者分值较高的相似性分值（例如，排列在前 20%的相似性分值）相加，作为每个候选视频的相似性分值；将每个候选视频的相似性分值按照降序进行排列；将排列在前面一个或者几个候选视频确定为与目标视频包含同一目标行人的视频。其中，预设阈值可以根据实际情况进行设置，分值较高为相对而言。

5 本申请实施例在进行行人再识别时，获取包含目标行人的目标视频和至少一个候选视频，对目标视频中的每个目标视频片段和至少一个候选视频中的每个候选视频片段分别进行编码，根据编码结果确定每个目标视频片段和每个候选视频片段之间的相似性分值；根据相似性分值对至少一个
10 候选视频进行行人再识别。由于视频片段包含的帧数远远少于整段视频包含的帧数，因此，视频片段中的行人表面信息的变化程度远远小于整段视频中的行人表面信息的变化程度。与对整段目标视频和整段候选视频进行编码相比，对每个目标视频片段和每个候选视频片段进行编码，有效减小了行人表面信息的变化，同时利用了不同视频帧内的行人表面信息的多样
15 性和视频帧与视频帧之间动态相关性，提高了行人表面信息的利用率，提高了编码结果对每个目标视频片段和每个候选视频片段之间的相似性分值计算的准确率，进而可以提高了行人再识别的准确率。

本申请实施例中的候选视频的编码结果是由目标视频片段的索引特征向量与候选视频片段的“键”特征向量而得，在编码过程中，利用目标视频片段的索引特征向量作为指导信息，提高了候选视频的编码结果对确定
20 相似性分值的准确性。利用目标视频片段的索引特征向量估计每一个候选视频帧的注意权重向量，减少候选视频中异常候选视频帧对候选视频片段的编码结果的影响，提升了候选视频中行人再识别的针对性。

本申请实施例将目标视频和候选视频进行片段切割，对目标视频片段和候选视频片段进行编码，当候选视频中的行人在部分候选视频帧中被遮挡时，选择相似性分值较高的候选视频片段作为候选视频的有效候选视频片段，忽略相似性分值较低的候选视频片段。

参照图 5，示出了根据本申请实施例行人再识别装置的一个实施例的结构示意图。

30 本申请实施例提供的行人再识别装置包括：获取模块 50，配置为获取包含目标行人的目标视频和至少一个候选视频；编码模块 52，配置为对目标视频中的每个目标视频片段和至少一个候选视频中的每个候选视频片段分别进行编码；确定模块 54，配置为根据编码结果确定每个目标视频片段和每个候选视频片段之间的相似性分值；相似性分值用于表征目标视频片
35 段与候选视频片段中的行人特征的相似程度；识别模块 56，配置为根据相似性分值对至少一个候选视频进行行人再识别。

本申请实施例的行人再识别装置用于实现上述实施例中相应的行人再识别方法，并具有相应的方法实施例的有益效果，在此不再赘述。

参照图 6, 示出了根据本申请实施例行人再识别装置的另一个实施例的结构示意图。

本申请实施例提供的行人再识别装置包括: 获取模块 60, 配置为获取包含目标行人的目标视频和至少一个候选视频; 编码模块 62, 配置为对目标视频中的每个目标视频片段和至少一个候选视频中的每个候选视频片段分别进行编码; 确定模块 64, 配置为根据编码结果确定每个目标视频片段和每个候选视频片段之间的相似性分值; 相似性分值用于表征目标视频片段与候选视频片段中的行人特征的相似程度; 识别模块 66, 配置为根据相似性分值对至少一个候选视频进行行人再识别。

10 可选地, 编码模块 62 包括: 特征向量获取模块 620, 配置为获取每个目标视频片段中的每个目标视频帧的第一目标特征向量和第二目标特征向量以及每个目标视频片段的索引特征向量, 获取每个候选视频片段中的每个候选视频帧的第一候选特征向量和第二候选特征向量; 权重向量生成模块 622, 配置为根据索引特征向量、第一目标特征向量和第一候选特征向量生成注意权重向量; 编码结果获取模块 624, 配置为根据注意权重向量、第二目标特征向量和第二候选特征向量获得每个目标视频片段的编码结果和每个候选视频片段的编码结果。

20 可选地, 特征向量获取模块 620, 配置为分别提取每个目标视频帧的图像特征向量和每个候选视频帧的图像特征向量; 根据每个目标视频帧的图像特征向量生成每个目标视频帧的第一目标特征向量和第二目标特征向量以及每个目标视频片段的索引特征向量, 根据每个候选视频帧的图像特征向量生成每个候选视频帧的第一候选特征向量和第二候选特征向量。

25 可选地, 权重向量生成模块 622, 配置为根据索引特征向量和第一目标特征向量生成每个目标视频帧的目标注意权重向量, 根据索引特征向量和第一候选特征向量生成每个候选视频帧的候选注意权重向量。

30 可选地, 权重向量生成模块 622, 配置为根据索引特征向量、每个目标视频帧的第一目标特征向量生成每个目标视频帧的目标热度图; 对目标热度图进行归一化处理得到每个目标视频帧的目标注意权重向量; 和/或, 根据索引特征向量、每个候选视频帧的第一候选特征向量生成每个候选视频帧的候选热度图; 对候选热度图进行归一化处理得到每个候选视频帧的候选注意权重向量。

35 可选地, 编码结果获取模块 624, 配置为根据每个目标视频帧的目标注意权重向量和第二目标特征向量获得每个目标视频片段的编码结果, 根据每个候选视频帧的候选注意权重向量和第二候选特征向量获得每个候选视频片段的编码结果。

可选地, 编码结果获取模块 624, 配置为将每个目标视频帧的目标注意权重向量与各自目标视频帧的第二目标特征向量相乘; 将每个目标视频帧的相乘结果在时间维度相加, 得到每个目标视频片段的编码结果; 和/或,

将每个候选视频帧的候选注意权重向量与各自候选视频帧的第二候选特征向量相乘；将每个候选视频帧的相乘结果在时间维度相加，得到每个候选视频片段的编码结果。

5 可选地，确定模块 64，配置为将每个目标视频片段的编码结果与每个候选视频片段的编码结果依次进行相减操作；将相减操作的结果在每一个维度上进行平方操作；对平方操作得到的特征向量进行全连接操作得到二维的特征向量；将二维的特征向量进行归一化操作，得到每个目标视频片段和每个候选视频片段之间的相似性分值。

10 可选地，识别模块 66，配置为针对至少一个候选视频中的每个候选视频片段，将分值最高的预设比例阈值的相似性分值相加，作为每个候选视频的相似性分值；将每个候选视频的相似性分值按照降序进行排列；将排列在前面的一个或者几个候选视频确定为与目标视频包含同一目标行人的视频。

15 本申请实施例的行人再识别装置用于实现上述实施例中相应的行人再识别方法，并具有相应的方法实施例的有益效果，在此不再赘述。

本申请实施例还提供了一种电子设备，例如可以是移动终端、个人计算机（PC）、平板电脑、服务器等。下面参考图 7，其示出了适于用来实现本申请实施例行人再识别装置的电子设备 700 的结构示意图：如图 7 所示，电子设备 700 可以包括存储器和处理器。具体地，电子设备 700 包括一个或多个处理器、通信元件等，所述一个或多个处理器例如：一个或多个中央处理单元（CPU）701，和/或 一个或多个图像处理器（GPU）713 等，处理器可以根据存储在只读存储器（ROM）702 中的可执行指令或者从存储部分 708 加载到随机访问存储器（RAM）703 中的可执行指令而执行各种适当的动作和处理。通信元件包括通信组件 712 和/或通信接口 709。其中，通信组件 712 可包括但不限于网卡，所述网卡可包括但不限于 IB（Infiniband）网卡，通信接口 709 包括诸如局域网（LAN，Local Area Network）卡、调制解调器等网络接口卡的通信接口，通信接口 709 经由诸如因特网的网络执行通信处理。

30 处理器可与只读存储器 702 和/或随机访问存储器 703 中通信以执行可执行指令，通过通信总线 704 与通信组件 712 相连、并经通信组件 712 与其他目标设备通信，从而完成本申请实施例提供的任一项行人再识别方法对应的操作，例如，获取包含目标行人的目标视频和至少一个候选视频；对所述目标视频中的每个目标视频片段和至少一个所述候选视频中的每个候选视频片段分别进行编码；根据编码结果确定每个所述目标视频片段和每个所述候选视频片段之间的相似性分值；所述相似性分值用于表征所述目标视频片段与所述候选视频片段中的行人特征的相似程度；根据所述相似性分值对至少一个所述候选视频进行行人再识别。

此外，在 RAM703 中，还可存储有装置操作所需的各种程序和数据。

CPU701 或 GPU713、ROM702 以及 RAM703 通过通信总线 704 彼此相连。在有 RAM703 的情况下，ROM702 为可选模块。RAM703 存储可执行指令，或在运行时向 ROM702 中写入可执行指令，可执行指令使处理器执行上述通信方法对应的操作。输入/输出 (I/O) 接口 705 也连接至通信总线 704。

- 5 通信组件 712 可以集成设置，也可以设置为具有多个子模块（例如多个 IB 网卡），并在通信总线链接上。

以下部件连接至 I/O 接口 705：包括键盘、鼠标等的输入部分 706；包括诸如阴极射线管 (CRT)、液晶显示器 (LCD) 等以及扬声器等的输出部分 707；包括硬盘等的存储部分 708；以及包括诸如 LAN 卡、调制解调器等
10 等的网络接口卡的通信接口 709。驱动器 710 也根据需要连接至 I/O 接口 705。可拆卸介质 711，诸如磁盘、光盘、磁光盘、半导体存储器等等，根据需要安装在驱动器 710 上，以便于从其上读出的计算机程序根据需要被安装入存储部分 708。

需要说明的，如图 7 所示的架构仅为一种可选实现方式，在具体实践
15 过程中，可根据实际需要对上述图 7 的部件数量和类型进行选择、删减、增加或替换；在不同功能部件设置上，也可采用分离设置或集成设置等实现方式，例如 GPU 和 CPU 可分离设置或者可将 GPU 集成在 CPU 上，通信元件可分离设置，也可集成设置在 CPU 或 GPU 上，等等。这些可替换的实施方式均落入本发明的保护范围。

20 本申请实施例的电子设备可以用于实现上述实施例中相应的行人再识别方法，该电子设备中的各个器件可以用于执行上述方法实施例中的各个步骤，例如，上文中描述的行人再识别方法可以通过电子设备的处理器调用存储器存储的相关指令来实现，为了简洁，在此不再赘述。

根据本申请实施例，上文参考流程图描述的过程可以被实现为计算机
25 程序产品。例如，本申请实施例包括一种计算机程序产品，其包括有形地包含在机器可读介质上的计算机程序，计算机程序包含用于执行流程图所示的方法的程序代码，程序代码可包括对应执行本申请实施例提供的方法步骤对应的指令，例如，获取包含目标行人的目标视频和至少一个候选视频；对所述目标视频中的每个目标视频片段和至少一个所述候选视频中的
30 每个候选视频片段分别进行编码；根据编码结果确定每个所述目标视频片段和每个所述候选视频片段之间的相似性分值；所述相似性分值用于表征所述目标视频片段与所述候选视频片段中的行人特征的相似程度；根据所述相似性分值对至少一个所述候选视频进行行人再识别。在这样的实施例中，该计算机程序可以通过通信元件从网络上被下载和安装，和/或从可拆卸
35 介质 711 被安装。在该计算机程序被处理器执行时，执行本申请实施例的方法中公开的功能。

可能以许多方式来实现本申请实施例的方法和装置、电子设备和存储介质。例如，可通过软件、硬件、固件或者软件、硬件、固件的任何组合

来实现本申请实施例的方法和装置、电子设备和存储介质。用于方法的步骤的上述顺序仅是为了进行说明，本申请实施例的方法的步骤不限于以上具体描述的顺序，除非以其它方式特别说明。此外，在一些实施例中，还可将本申请实施为记录在记录介质中的程序，这些程序包括用于实现根据
5 本申请实施例的方法的机器可读指令。因而，本申请还覆盖存储用于执行根据本申请实施例的方法的程序的记录介质。

本申请实施例的描述是为了示例和描述起见而给出的，而并不是无遗漏的或者将本发明限于所公开的形式，很多修改和变化对于本领域的普通技术人员而言是显然的。选择和描述实施例是为了更好说明本申请的原理
10 和实际应用，并且使本领域的普通技术人员能够理解本申请从而设计适于特定用途的带有各种修改的各种实施例。