



(12)发明专利申请

(10)申请公布号 CN 108805077 A

(43)申请公布日 2018. 11. 13

(21)申请号 201810592758.7

(22)申请日 2018.06.11

(71)申请人 深圳市唯特视科技有限公司

地址 518057 广东省深圳市高新技术产业
园区高新南一道009号中研发园新
产业孵化中心楼610室

(72)发明人 夏春秋

(51)Int.Cl.

G06K 9/00(2006.01)

G06N 3/04(2006.01)

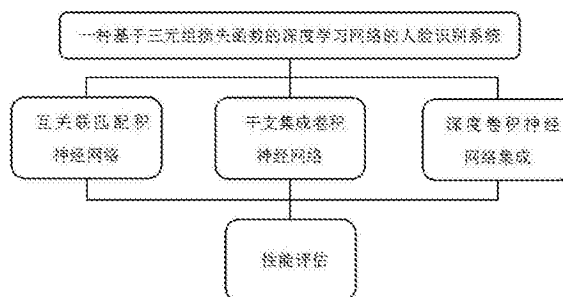
权利要求书2页 说明书4页 附图3页

(54)发明名称

一种基于三元组损失函数的深度学习网络的人脸识别系统

(57)摘要

本发明中提出的一种基于三元组损失函数的深度学习网络的人脸识别系统,其主要内容包括:互关联匹配卷积神经网络、干支集成卷积神经网络、深度卷积神经网络集成和性能评估,其过程为,首先从训练数据集中选择三个样本,包括一个静止ROI(感兴趣区域)、一个与静止ROI相似的正样本和一个与静止ROI不相似的负样本,让这三个样本共同组成了一个三元组;然后将这个三元组输入到深度学习网络中进行训练,训练过程中采用了三元组函数(可拉近相似ROI的距离);最后,相似ROI会形成集聚,从而达到人脸识别的目的。本发明采用了三元组损失函数,相比起传统的人脸识别系统有更高的识别精度,并且运算复杂程度相对较低,运算效率较高。



1. 一种基于三元组损失函数的深度学习网络的人脸识别系统,其特征在于,主要包括互关联匹配卷积神经网络(一);干支集成卷积神经网络(二);深度卷积神经网络集成(三);性能评估(四)。

2. 基于权利要求书1所述的互关联匹配卷积神经网络(一),其特征在于,互关联匹配卷积神经网络(CCM-CNN)采用了矩阵哈达玛乘积,其后是一个全连接层,用于模拟自适应加权互关联技术;采用一种基于三元组优化的方法来学习面部表征判别式,这些面部表征是基于三元组的,包括正样本的和负样本的视频感兴趣区域(ROI)以及对应的静态ROI;为了进一步提高面部模型的鲁棒性,通过生成基于静态和视频非目标个体ROI的合成面部,使CCM-CNN的微调过程包括了多项信息;互关联匹配卷积神经网络主要包括三个部分:特征提取、互关联匹配和三元组损失优化。

3. 基于权利要求书2所述的特征提取,其特征在于,通过特征提取管线实现,用于对同一个对象在不同情况下取得的ROI进行区别性特征的提取;特征提取管线包括三个子网络,分别对应静止、正样本和负样本的面部;每一个子网络都包括9个卷积层,每一个卷积层后是一个空间批标准层、丢失层和线性整流函数层。

4. 基于权利要求书2所述的互关联匹配,其特征在于,互关联匹配主要采用的是一种像素匹配方法,能够高效的对特征映射进行对比,并且估量其匹配相似性;对比过程主要包括三个部分:矩阵乘积、全连接层和Softmax层;此方法采用了特征映射来表示ROI,将这些特征映射相乘从而对三元组的ROI进行自编码,大大降低了对比的复杂性。

5. 基于权利要求书2所述的三元组损失优化,其特征在于,采用一个双向三元组优化函数来高效地训练网络;为了使三元组损失优化与网络兼容,需要在网络中添加额外的特征提取分支;三元组损失可用以下公式表示:

$$\text{Triple Loss} = \frac{1}{L} \sum_{R_i \in B} \sqrt{(1 - S_{t_i p_i})^2 + S_{t_i n_i}^2 + S_{n_i p_i}^2} \quad (1)$$

其中, S_{tp} 、 S_{tn} 和 S_{np} 为互关联匹配中的相似性得分(分别为静止ROI与正样本ROI的对比得分、静止ROI与负样本ROI的对比得分和正样本ROI与负样本ROI的对比得分)。

6. 基于权利要求书1所述的干支集成卷积神经网络(二),其特征在于,干支集成卷积神经网络(TBE-CNN)可用于从整体面部图像和干支网络的面部标记块中提取互补性特征;为了对现实视频数据进行仿真,从静止图像中人工合成(主要采用人工失焦和动态模糊来学习对模糊不敏感的面部表征)模糊训练数据;TBE-CNN包含一个主干网络和多个分支网络,干支网络有一部分公用层,用来植入全局和局部信息,这个方法降低了计算成本并且有效地融合了信息;将干支网络的输出特征示意图串联起来以在全连接层中生成最终的面部表征。

7. 基于权利要求书1所述的深度卷积神经网络集成(三),其特征在于,深度卷积神经网络集成(HaarNet)可高效地学习区别性强的面部表征以满足视频面部识别;HaarNet包含了一个主干网络和三个分支网络,这些网络的设计是用来植入面部特征、姿态特征以及其他区别性特征;此外,为了提升区别能力,HaarNet采用了一种多阶段的训练方法,并且还采用了一个二阶统计标准三元组损失方程从类内和类间变化中获取信息;最终,在一个微调阶段中植入面部ROI的相关性信息,这些信息储存于登入和提升识别精确性的阶段。

8. 基于权利要求书7所述的多阶段的训练方法,其特征在于,主要包括将三元组输入

HaarNet、HaarNet的输出结果进行L₂-标准化表示和进行三元组损失处理,这种训练方法能够高效地优化HaarNet的内部参数。

9. 基于权利要求书1所述的性能评估(四),其特征不在于,使用考克斯人脸数据库(Cox Face DB)对本系统进行评估,Cox Face DB中的面部信息包括了从静止照相机在可控环境下拍摄的高质量面部照片和视频设备在非可控情况下拍摄的低质量面部图像;性能评估主要有两个方面:将静止图片与视频图像进行对比、计算复杂性的评估。

10. 基于权利要求书9所述的计算复杂性的评估,其特征不在于,计算复杂性主要决定于运行流程(匹配静止ROI和视频ROI)的数量、网络参数和层数的数量;计算复杂性决定了人脸识别的效率。

一种基于三元组损失函数的深度学习网络的人脸识别系统

技术领域

[0001] 本发明涉及人脸识别领域,尤其是涉及了一种基于三元组损失函数的深度学习网络的人脸识别系统。

背景技术

[0002] 人脸识别是基于人的脸部特征信息进行身份识别的一种生物识别技术。用摄像机或摄像头采集含有人脸的图像或视频流,并自动在图像中检测和跟踪人脸,进而对检测到的人脸进行脸部的一系列相关处理,通常也叫做人像识别、面部识别。在企业、住宅安全和管理方面,人脸识别技术可用于门禁考勤系统、人脸识别防盗门等;在公安、司法和刑侦方面,可结合人脸识别技术人脸数据库在全球范围内进行罪犯追捕;在电子政务和电子商务的使用方面,由于人脸识别系统采用的是生物特征(而不是传统的字符密码),可以更精确做到当事人在网上的数字身份和真实身份统一,从而大大增加电子商务和电子政务系统的可靠性;此外,人脸识别技术还在航天、电力、边检、教育等领域广泛应用。然而,目前的人脸识别系统运算复杂程度较高,运算效率低,并且识别的精确性也不够高。

[0003] 本发明中提出的一种基于三元组损失函数的深度学习网络的人脸识别系统,先从训练数据集中选择三个样本,包括一个静止ROI(感兴趣区域)、一个与静止ROI相似的正样本和一个与静止ROI不相似的负样本,让这三个样本共同组成了一个三元组;然后将这个三元组输入到深度学习网络中进行训练,训练过程中采用了三元组函数(可拉近相似ROI的距离);最后,相似ROI会形成集聚,从而达到人脸识别的目的。本发明采用了三元组损失函数,相比起传统的人脸识别系统有更高的识别精度,并且运算复杂程度相对较低,运算效率较高。

发明内容

[0004] 针对目前的人脸识别系统运算复杂程度较高,运算效率低,并且识别的精确性也不够高等问题,本发明中提出的一种基于三元组损失函数的深度学习网络的人脸识别系统,先从训练数据集中选择三个样本,包括一个静止ROI、一个与静止ROI相似的正样本和一个与静止ROI不相似的负样本,让这三个样本共同组成了一个三元组;然后将这个三元组输入到深度学习网络中进行训练,训练过程中采用了三元组函数(可拉近相似ROI的距离);最后,相似ROI会形成集聚,从而达到人脸识别的目的。

[0005] 为解决上述问题,本发明提供一种基于三元组损失函数的深度学习网络的人脸识别系统,其主要内容包括:

[0006] (一) 互关联匹配卷积神经网络(CCM-CNN);

[0007] (二) 干支集成卷积神经网络(TBE-CNN);

[0008] (三) 深度卷积神经网络集成(HaarNet);

[0009] (四) 性能评估。

[0010] 其中,所述的互关联匹配卷积神经网络,采用了矩阵哈达玛乘积,其后是一个全连

接层,用于模拟自适应加权互关联技术;采用一种基于三元组优化的方法来学习面部表征判别式,这些面部表征是基于三元组的,包括正样本的和负样本的视频感兴趣区域(ROI)以及对应的静态ROI;为了进一步提高面部模型的鲁棒性,通过生成基于静态和视频非目标个体ROI的合成面部,使CCM-CNN的微调过程包括了多项信息;互关联匹配卷积神经网络主要包括三个部分:特征提取、互关联匹配和三元组损失优化。

[0011] 进一步地,所述的特征提取,通过特征提取管线实现,用于对同一个对象在不同情况下取得的ROI进行区别性特征的提取;特征提取管线包括三个子网络,分别对应静止、正样本和负样本的面部;每一个子网络都包括9个卷积层,每一个卷积层后是一个空间批标准层、丢失层和线性整流函数层。

[0012] 进一步地,所述的互关联匹配,主要采用的是一种像素匹配方法,能够高效的对特征映射进行对比,并且估量其匹配相似性;对比过程主要包括三个部分:矩阵乘积、全连接层和Softmax层;此方法采用了特征映射来表示ROI,将这些特征映射相乘从而对三元组的ROI进行自编码,大大降低了对比的复杂性。

[0013] 进一步地,所述的三元组损失优化,采用一个双向三元组优化函数来高效地训练网络;为了使三元组损失优化与网络兼容,需要在网络中添加额外的特征提取分支;三元组损失可用以下公式表示:

$$[0014] \quad \text{Triple Loss} = \frac{1}{L} \sum_{R_i \in B} \sqrt{(1 - S_{t_i p_i})^2 + S_{t_i n_i}^2 + S_{n_i p_i}^2} \quad (1)$$

[0015] 其中, S_{tp} 、 S_{tn} 和 S_{np} 为互关联匹配中的相似性得分(分别为静止ROI与正样本ROI的对比得分、静止ROI与负样本ROI的对比得分和正样本ROI与负样本ROI的对比得分)。

[0016] 其中,所述的干支集成卷积神经网络,可用于从整体面部图像和干支网络的面部标记块中提取互补性特征;为了对现实视频数据进行仿真,从静止图像中人工合成(主要采用人工失焦和动态模糊来学习对模糊不敏感的面部表征)模糊训练数据;TBE-CNN包含一个主干网络和多个分支网络,干支网络有一部分公用层,用来植入全局和局部信息,这个方法降低了计算成本并且有效地融合了信息;将干支网络的输出特征示意图串联起来以在全连接层中生成最终的面部表征。

[0017] 其中,所述的深度卷积神经网络集成,可高效地学习区别性强的面部表征以满足视频面部识别;HaarNet包含了一个主干网络和三个分支网络,这些网络的设计是用来植入面部特征、姿态特征以及其他区别性特征;此外,为了提升区别能力,HaarNet采用了一种多阶段的训练方法,并且还采用了一个二阶统计标准三元组损失方程从类内和类间变化中获取信息;最终,在一个微调阶段中植入面部ROI的相关性信息,这些信息储存于登入和提升识别精确性的阶段。

[0018] 进一步地,所述的多阶段的训练方法,主要包括将三元组输入HaarNet、HaarNet的输出结果进行 L_2 -标准化表示和进行三元组损失处理,这种训练方法能够高效地优化HaarNet的内部参数。

[0019] 其中,所述的性能评估,使用考克斯人脸数据库(Cox Face DB)对本系统进行评估,Cox Face DB中的面部信息包括了从静止照相机在可控环境下拍摄的高质量面部照片和视频设备在非可控情况下拍摄的低质量面部图像;性能评估主要有两个方面:将静止图片与视频图像进行对比、计算复杂性的评估。

[0020] 进一步地,所述的计算复杂性的评估,计算复杂性主要决定于运行流程(匹配静止ROI和视频ROI)的数量、网络参数和层数的数量;计算复杂性决定了人脸识别的效率。

附图说明

[0021] 图1是本发明一种基于三元组损失函数的深度学习网络的人脸识别系统的系统框架图。

[0022] 图2是本发明一种基于三元组损失函数的深度学习网络的人脸识别系统的网络架构图。

[0023] 图3是本发明一种基于三元组损失函数的深度学习网络的人脸识别系统的训练流程图。

具体实施方式

[0024] 需要说明的是,在不冲突的情况下,本申请中的实施例及实施例中的特征可以相互结合,下面结合附图和具体实施例对本发明作进一步详细说明。

[0025] 图1是本发明一种基于三元组损失函数的深度学习网络的人脸识别系统的系统框架图。主要包括互关联匹配卷积神经网络、干支集成卷积神经网络、深度卷积神经网络集成和性能评估。

[0026] 互关联匹配卷积神经网络,采用了矩阵哈达玛乘积,其后是一个全连接层,用于模拟自适应加权互关联技术;采用一种基于三元组优化的方法来学习面部表征判别式,这些面部表征是基于三元组的,包括正样本的和负样本的视频感兴趣区域(ROI)以及对应的静态ROI;为了进一步提高面部模型的鲁棒性,通过生成基于静态和视频非目标个体ROI的合成面部,使CCM-CNN的微调过程包括了多项信息;互关联匹配卷积神经网络主要包括三个部分:特征提取、互关联匹配和三元组损失优化。

[0027] 干支集成卷积神经网络,可用于从整体面部图像和干支网络的面部标记块中提取互补性特征;为了对现实视频数据进行仿真,从静止图像中人工合成(主要采用人工失焦和动态模糊来学习对模糊不敏感的面部表征)模糊训练数据;TBE-CNN包含一个主干网络和多个分支网络,干支网络有一部分公用层,用来植入全局和局部信息,这个方法降低了计算成本并且有效地融合了信息;将干支网络的输出特征示意图串联起来以在全连接层中生成最终的面部表征。

[0028] 深度卷积神经网络集成,可高效地学习区别性强的面部表征以满足视频面部识别;HaarNet包含了一个主干网络和三个分支网络,这些网络的设计是用来植入面部特征、姿态特征以及其他区别性特征;此外,为了提升区别能力,HaarNet采用了一种多阶段的训练方法,并且还采用了一个二阶统计标准三元组损失方程从类内和类间变化中获取信息;最终,在一个微调阶段中植入面部ROI的相关性信息,这些信息储存于登入和提升识别精确性的阶段。

[0029] 性能评估,使用考克斯人脸数据库(Cox Face DB)对本系统进行评估,Cox Face DB中的面部信息包括了从静止照相机在可控环境下拍摄的高质量面部照片和视频设备在非可控情况下拍摄的低质量面部图像;性能评估主要有两个方面:将静止图片与视频图像进行对比、计算复杂性的评估。

[0030] 其中,计算复杂性的评估,计算复杂性主要决定于运行流程(匹配静止ROI和视频ROI)的数量、网络参数和层数的数量;计算复杂性决定了人脸识别的效率。

[0031] 图2是本发明一种基于三元组损失函数的深度学习网络的人脸识别系统的网络架构图。以互关联匹配卷积神经网络为例,主要包括特征提取、互关联匹配和三元组损失优化三个部分。

[0032] 特征提取,通过特征提取管线实现,用于对同一个对象在不同情况下取得的ROI进行区别性特征的提取;特征提取管线包括三个子网络,分别对应静止、正样本和负样本的面部;每一个子网络都包括9个卷积层,每一个卷积层后是一个空间批标准层、丢失层和线性整流函数层。

[0033] 互关联匹配,主要采用的是一种像素匹配方法,能够高效的对特征映射进行对比,并且估量其匹配相似性;对比过程主要包括三个部分:矩阵乘积、全连接层和Softmax层;此方法采用了特征映射来表示ROI,将这些特征映射相乘从而对三元组的ROI进行自编码,大大降低了对比的复杂性。

[0034] 三元组损失优化,采用一个双向三元组优化函数来高效地训练网络;为了使三元组损失优化与网络兼容,需要在网络中添加额外的特征提取分支;三元组损失可用以下公式表示:

$$[0035] \quad \text{Triple Loss} = \frac{1}{L} \sum_{R_i \in B} \sqrt{(1 - S_{t_i p_i})^2 + S_{t_i n_i}^2 + S_{n_i p_i}^2} \quad (1)$$

[0036] 其中, S_{tp} 、 S_{tn} 和 S_{np} 为互关联匹配中的相似性得分(分别为静止ROI与正样本ROI的对比得分、静止ROI与负样本ROI的对比得分和正样本ROI与负样本ROI的对比得分)。

[0037] 图3是本发明一种基于三元组损失函数的深度学习网络的人脸识别系统的训练流程图。以深度卷积神经网络集成的训练流程为例,此训练方法是一种多阶段的训练方法,主要包括将三元组输入HaarNet、HaarNet的输出结果进行L₂-标准化表示和进行三元组损失处理等过程,这种训练方法能够高效地优化HaarNet的内部参数。

[0038] 对于本领域技术人员,本发明不限制于上述实施例的细节,在不背离本发明的精神和范围的情况下,能够以其他具体形式实现本发明。此外,本领域的技术人员可以对本发明进行各种改动和变型而不脱离本发明的精神和范围,这些改进和变型也应视为本发明的保护范围。因此,所附权利要求意欲解释为包括优选实施例以及落入本发明范围的所有变更和修改。

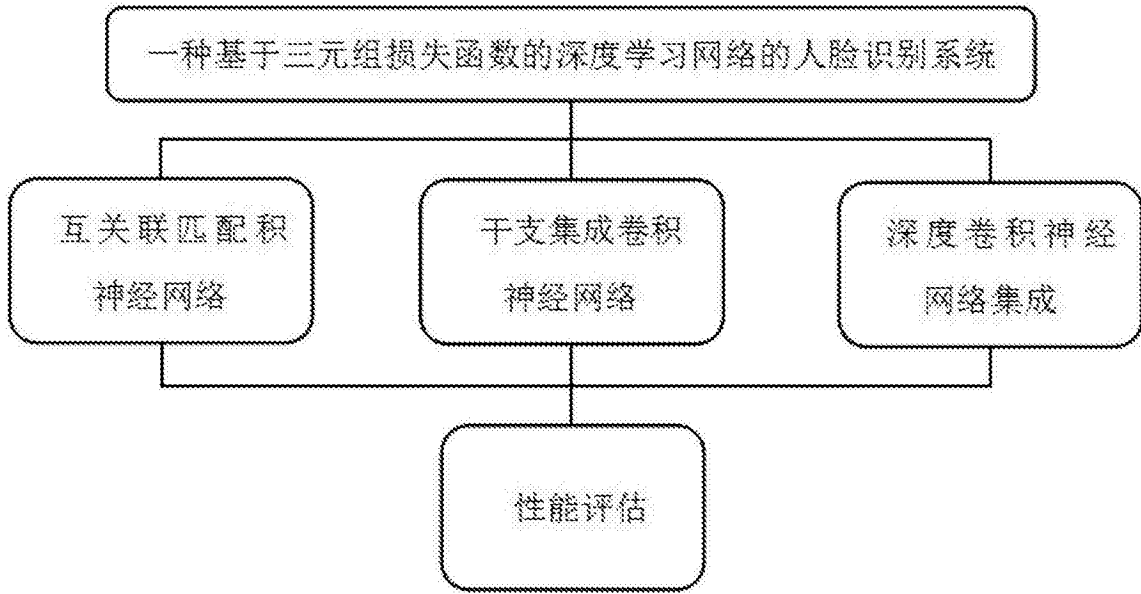


图1

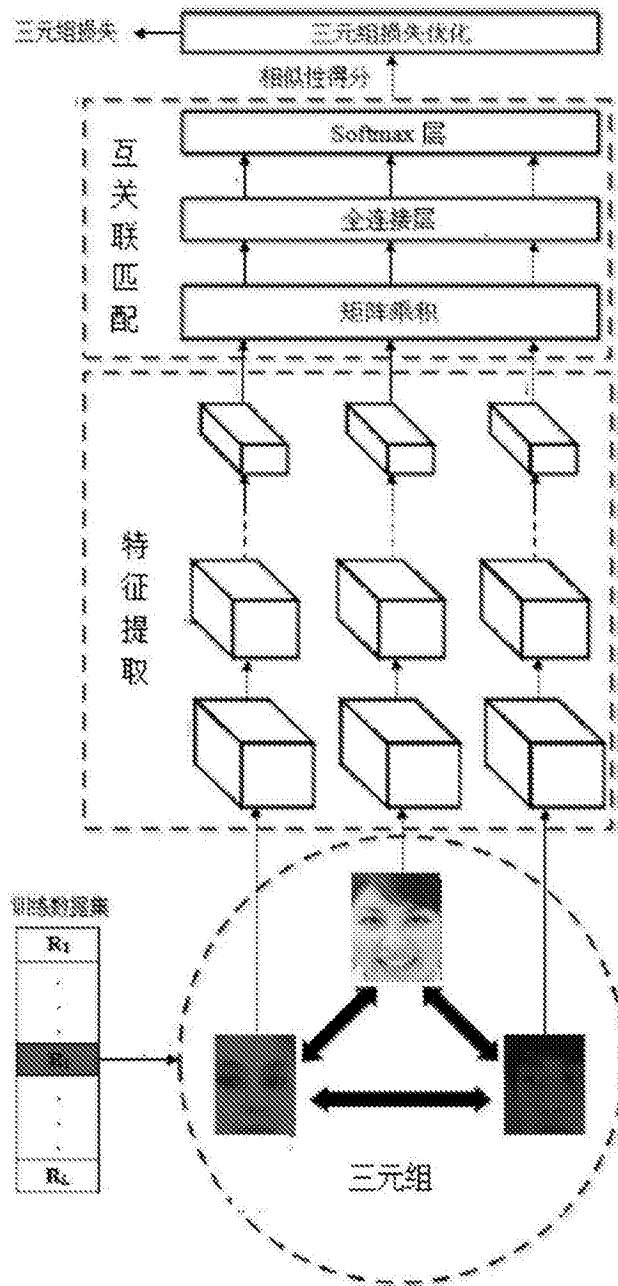


图2

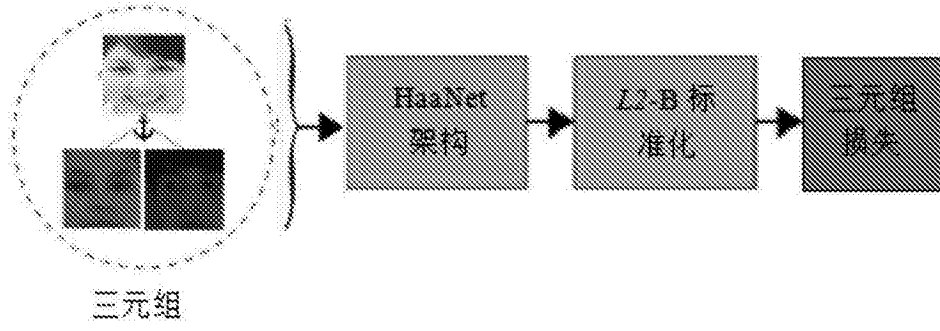


图3