

(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国际局

(43) 国际公布日
2019 年 1 月 10 日 (10.01.2019)



(10) 国际公布号
WO 2019/007004 A1

(51) 国际专利分类号:
G06K 9/00 (2006.01)

(21) 国际申请号: PCT/CN2017/118794

(22) 国际申请日: 2017 年 12 月 27 日 (27.12.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201710536020.4 2017 年 7 月 4 日 (04.07.2017) CN

(71) 申请人: 北京大学深圳研究生院(PEKING UNIVERSITY SHENZHEN GRADUATE SCHOOL) [CN/CN]; 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。

(72) 发明人: 王文敏(WANG, Wenmin); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 张奕豪(ZHANG, Yihao); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 王荣刚(WANG, Ronggang); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 李革(LI, Ge); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 董胜富(DONG, Shengfu); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 王振宇(WANG, Zhenyu); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 李英(LI, Ying); 中国广东省深圳市南山区西丽镇丽水路深圳大学城北大校区, Guangdong 518055 (CN)。 高文(GAO, Wen); 中

(54) Title: IMAGE FEATURE EXTRACTION METHOD FOR PERSON RE-IDENTIFICATION

(54) 发明名称: 用于行人重识别的图像特征提取方法

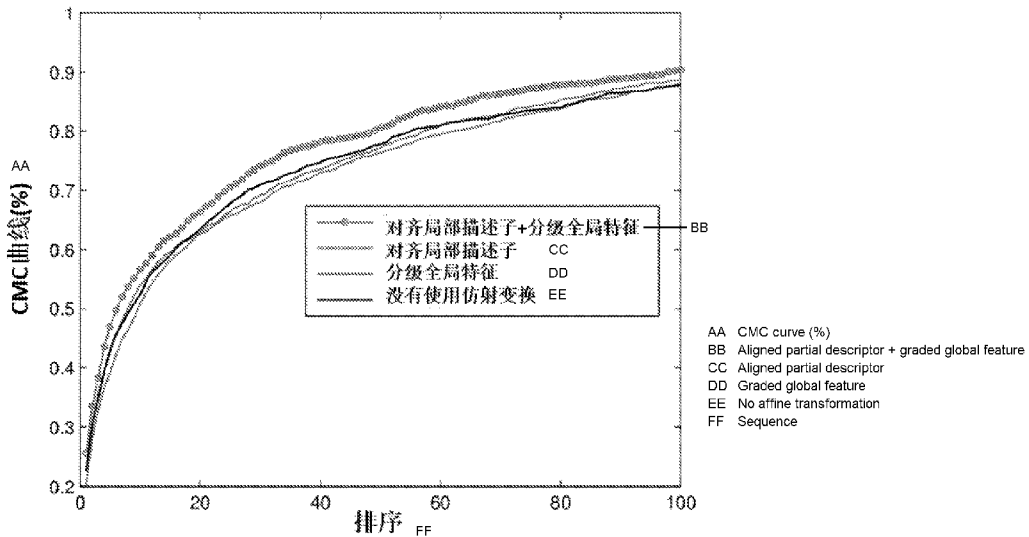


图 6

(57) Abstract: Provided is an image feature extraction method for person re-identification, comprising: performing person re-identification by means of aligned partial descriptor extraction and graded global feature extraction, wherein the aligned partial descriptor extraction is performed by processing an original image by affine transformation and performing a summation pooling operation on image block features of same regions to obtain an aligned partial descriptor, and spatial information between inner blocks of the image for the aligned partial descriptor is reserved; graded global feature extraction is performed by grading a positioned person region block and solving a corresponding feature mean value to obtain a global feature. The invention can resolve the problem of feature misalignment caused by posture changes in person re-identification, and eliminate the effect of unrelated backgrounds on re-identification, thus improving the precision and robustness of person re-identification.

国广东省深圳市南山区西丽镇丽水路深圳大学
城北大校区, Guangdong 518055 (CN)。

(74) 代理人: 北京万象新悦知识产权代理事务
所(普通合伙)(BEIJING WANXIANGXINYUE
INTELLECTUAL PROPERTY OFFICE); 中国北京
市海淀区颐和园路1号北京资源燕园宾馆
六层1625号, Beijing 100080 (CN)。

(81) 指定国(除另有指明, 要求每一种可提供的国家
保护): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG,
BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU,
CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB,
GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS,
JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK,
LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX,
MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,
PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国(除另有指明, 要求每一种可提供的地区
保护): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ,
NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM,
AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG,
CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,
IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT,
RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI,
CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

本国际公布:

— 包括国际检索报告(条约第21条(3))。

(57) 摘要: 一种用于行人重识别的图像特征提取方法, 通过对齐局部描述子提取和分级全局特征提取, 进行行人重识别; 对齐局部描述子提取采用仿射变换对原图像进行处理, 对相同区域的图像块特征进行求和池化操作而得到对齐局部描述子; 对齐局部描述子保留图像内部块与块之间的空间信息; 分级全局特征提取通过对定位的行人区域块进行分级, 求取相应特征均值而得到全局特征。该方法能够解决行人重识别中由于行人姿态变化等带来的特征不对齐问题, 消除无关背景对重识别带来的影响, 由此提高行人重识别的精度和鲁棒性。

用于行人重识别的图像特征提取方法

技术领域

本发明涉及计算机视觉与图像处理技术，尤其涉及一种用于行人重识别的图像鲁棒性特征的提取方法。

背景技术

近年来，随着人们对社会公共安全的要求日渐提升，视频监控系统大量普及，如何利用计算机视觉分析技术，自动化监控及分析视频信息，成为关注的热点。行人重识别是计算机视觉研究中关键的任务。一般来说，给定关于行人的一张图片或者一段视频，行人重识别就是在其它不重合拍摄场景下的图片或者视频中，将同一个人识别出来的过程。尽管相关的研究越来越受到重视，行人重识别的准确率也已经提高了不少，但仍有许多困难需要解决，比如姿态的变化、光照和无关背景的影响等等。

现有传统的行人重识别框架主要由特征提取和度量学习两部分构成。不过，目前大多数方法在提取图像特征的过程中，都忽略了无关背景信息的影响，使得提取好的图像特征带有较多的噪声；其次，为了解决由于拍摄角度不一以及行人姿态的变化问题，大多数方法采取了非精确匹配的策略，使用特定的方式将不同图像块之间的特征进行融合，效果颇为显著，但同时也损失了一部分图像内部的空间信息，使得在分辨一些具有相对位置差异的图像上，能力大大降低。

发明内容

为了克服上述现有技术的不足，本发明提供一种用于行人重识别的图像特征的提取方法，基于图像颜色和纹理特征，在精确匹配的策略下提出新的对齐局部描述子，并设计分级全局特征与之互补，解决由于行人姿态变化等带来的特征不对齐等问题，提高行人重识别的精确度。

本发明的原理是：提出了应用于行人重识别的对齐局部描述子和分级全局特征提取算法。本发明针对现有特征提取方法中非精确匹配策略带来的缺陷，提出基于精确匹配策略下的对齐局部描述子，以及为了消除无关背景的影响，提出了分级全局特征，通过对齐局部描述子与分级全局特征提取，提高行人重识别的性能。对齐局部描述子保留了图像内部块与块之间的空间信息，同时为了解决图像之间的不对齐问题，采用了仿射变换对原图像进行处理，得

到包括原图在内的水平翻转和切变的四张图像，通过叠加，仅对相邻区域的图像块特征进行求和池化操作，得到对齐局部描述子。为了丰富图像特征的表达以及消除背景信息的影响，对原图像中行人的整体轮廓区域及其头部、上半身和下半身分别计算全局特征，最后整合成分级全局特征。实验证明，对齐局部描述子和分级全局特征有助于解决姿态变化问题和弱化背景的影响，保留的图像内部空间信息也有助于行人匹配精度的提高。

对齐局部描述子是基于精确匹配策略所提出来的一种特征，即对于图像内部的块，提取其相应的颜色和纹理特征，保留该图像块的位置信息，不与图像内其它块的特征融合。为了解决行人姿态变化带来两张图像之间的不对齐问题，采用了仿射变换方法对原始图像进行了处理，如水平翻转和水平切变，得到关于同一张图像的四种不同角度的图像扩充。紧接着，在空间上将四张图像叠加在一起，以图像块的特征为基础，对相同图像区域的图像块特征进行求和池化操作，作为该图像块最终的描述子。分级全局特征是基于与局部特征互补的作用而提出来的。传统的全局特征是对整幅图像统计相应的颜色和纹理信息，提取的全局特征往往包含大量的背景信息，而且空间信息紊乱，对于结构差异大但统计信息一致的图像，极易造成误差。分级全局特征不仅考虑到背景的影响，还把属于行人的区域的图像块按照头部、上半身和下半身的人体结构进行分组，并对涵盖重要信息的上半身进一步分组，总共三级划分，增强了全局特征的辨识力。

本发明提供的技术方案是：

一种图像特征提取方法，通过对齐局部描述子提取和分级全局特征提取，进行行人重识别；所述对齐局部描述子提取采用仿射变换对原图像进行处理，仅对相邻区域的图像块特征进行求和池化操作而得到对齐局部描述子；所述对齐局部描述子保留图像内部块与块之间的空间信息；所述分级全局特征提取对定位的行人区域块进行分级，通过求取相应特征均值得到全局特征；包括如下步骤：

1) 对图像进行了预处理，以消除光照对图像的影响；

为了消除光照对图像的影响，本发明采用了多尺度 Retinex 算法对图像进行了预处理；在去除背景信息时，借助了通过元胞自动机进行显著性检测的前景提取方法；关于图像块的表达，颜色信息使用了 nRGB、Lab 和 HSV 三种颜色空间直方图，纹理信息使用了尺度不变局部三值模式 (SILTP) 统计直方图。

2) 提取对齐局部描述子，主要包括以下步骤：

21) 对图像进行仿射变换，得到多张图像(包括原图及变换后的图像)；

22) 分别对上一步得到的每张图像进行图像块分割, 提取特征, 生成相应的特征图;

23) 在空间上叠加所有特征图, 对相邻位置的图像块特征进行求和池化操作, 得到相应局部描述子;

24) 根据局部描述子在图像中所在的位置, 将它们按从左上到右下的顺序连接起来, 得到对齐局部描述子;

3) 提取分级全局特征, 主要分为以下步骤:

31) 使用前景提取方法对图像进行显著性检测, 得到相应的显著图;

32) 将上一步得到的显著图进行对比增强;

33) 定位属于行人区域的图像块;

34) 针对人体的特殊结构, 将属于行人的图像块按照三个不同层级进行归类;

35) 对每一级图像块进行平均池化操作, 得到该级特征;

36) 对三级特征首尾连接起来, 得到分级全局特征;

4) 根据步骤 2) 提取得到的对齐局部描述子和步骤 3) 提取得到的分级全局特征, 使用现有的 XQDA (Cross-view Quadratic Discriminant Analysis, 交叉视觉二次判别分析) 方法作为度量学习方法, 计算图像的相似性, 由此进行行人重识别。

与现有技术相比, 本发明的有益效果是:

本发明提供一种行人重识别方法, 通过对齐局部描述子的提取, 保留了图像局部块之间的空间信息, 并解决由于行人姿态变化等带来的特征不对齐问题; 通过分级全局特征提取, 消除无关背景带来的影响; 通过结合提取的对齐局部描述子和分级全局特征, 提高行人重识别的精度和鲁棒性。具体地, 本发明具有以下优点:

(一) 本发明方法保留不相关图像块的相对位置信息。现有行人重识别方法在特征提取过程中经常忽略, 而本发明通过增加不相关图像块的相对位置信息, 可让后续匹配的准确率更高;

(二) 本发明方法使用仿射变换对输入图片进行处理, 可以一定程度上解决由于行人姿态变化或拍摄角度不一带来的特征不对齐问题;

(三) 本发明方法使用多个层级的行人区域图像块提取全局特征, 可以提高全局特征的鲁棒性。

附图说明

图 1 是对齐局部描述子提取过程示意图；

其中，(a) 为对原图像进行仿射变换（水平剪切和水平翻转），分别得到四张变换后的图像；(b) 为分别对每一张图像进行图像块的特征提取过程；(c) 为将四张特征图在空间上叠加在一起；(d) 为对相邻图像块的特征进行求和池化操作，得到多个局部描述子；(e) 为将每个局部描述子向量相连，得到最终的对齐局部描述子。

图 2 是分级全局特征提取过程示意图；

其中，(a) 为通过显著性检测到显著图；(b) 为对显著图进行对比度增强；(c) 为对属于行人区域图像块进行定位；(d) 为对所有属于行人的图像块特征进行第一级全局特征提取；(e) 为对所有属于行人的图像块分成头部、上半身和下半身三个部分，提取第二级全局特征；(f) 为对所有属于行人上半身的图像块分成三个部分，提取第三级全局特征；

图 3 是本发明实施例在 VIPeR 数据集上与使用 LOMO 特征匹配效果对比。

图 4 是本发明实施例在 VIPeR 数据集上的 CMC 结果。

图 5 是本发明实施例在 QMUL GRID 数据集上与使用 LOMO 特征匹配效果对比。

图 6 是本发明实施例在 QMUL GRID 数据集上的 CMC 结果。

具体实施方式

下面结合附图，通过实施例进一步描述本发明，但不以任何方式限制本发明的范围。

本发明针对现有图像特征提取方法中非精确匹配策略带来的缺陷，提出基于精确匹配策略下的对齐局部描述子和分级全局特征。对齐局部描述子保留了图像内部块与块之间的空间信息，同时为了解决图像之间的不对齐问题，采用了仿射变换对原图像进行处理，得到包括原图在内的水平翻转和切变的四张图像，通过叠加，对相同区域的图像块特征进行求和池化操作，得到对齐局部描述子。为了丰富图像特征的表达，对原图像中行人的整体轮廓区域及其头部、上半身和下半身分别计算全局特征，最后整合成分级全局特征。实验证明，对齐局部描述子和分级全局特征有助于解决姿态变化问题，保留的图像内部空间信息有助于行人匹配精度的提高。

图 1 示意了对齐局部描述子的提取流程，首先是仿射变换，然后是图像块的特征提取，接着对相邻图像块的特征向量进行求和池化操作，最后把不同的局部描述子连接到一起作为图片的对齐局部描述子。具体实施中，提取对齐局部描述子的过程具体包括如下步骤：

1. 把图片统一缩放到 128*48 大小，并使用文献[1]（D. J. Jobson, Z.-U. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. Image Processing, IEEE Transactions on, 6(7):965–976, 1997）记载的多尺度 Retinex 算法消除光照影响，得到预处理后的图片 I。其中，多尺度 Retinex 算法以色感一致性（颜色恒常性）为基础，认为观察者所看到的图像是由物体表面对入射光 L 反射得到的，而反射率 R 是由物体本身决定，不受入射光 L 变化影响。因此，通过去除亮度分量 L 求得反射分量 R，可以达到图像增强的效果。另外，亮度分量可以通过对图像进行高斯卷积求得， σ 是高斯函数的尺度参数。如果使用单一的尺度，取值大时，颜色失真小但细节恢复差；取值小时，细节恢复好但颜色失真大。因此，采用多种尺度参数，进行加权求和，本实验中尺度参数分别为 $\sigma=5$ 和 $\sigma=20$ 。
2. 仿射变换。对预处理后的图片 I 进行水平翻转，得到图片 II，分别对图片 I 和 II 进行水平剪切变换，产生相应图片 III 和 IV，水平剪切变换具体如式 1 所示：

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (\text{式 1})$$

其中， λ 为剪切变换参数，方法中使用 $\lambda=0.2$ 。

3. 分别对上一步得到的四张图片（I、II、III 和 IV）求图像块的特征。

具体采用滑动窗口方法，对窗口内的图像块提取 8*8*8 维的 nRGB 和 HSV 颜色直方图、32 维的 Lab 颜色直方图以及两个阈值为 0.3 的尺度不变局部三值模式（SILTP）统计直方图，两个 SILTP 的比较半径 R 分别为 3 和 5 个像素，其中滑动窗口大小为 8*8，步长为 4*4，每张图像在同一水平方向上通过滑动窗口可以得到 N 个图像块，在同一垂直方向上得到 M 个图像块，提取相应的特征后得到一张特征图，其中有 M*N 个点，每个点对应一个图像块的特征向量，记作 $F_{m,n}^i$ 。

上述滑动窗口方法中，nRGB 是将原来的 RGB 三通道彩色图片中的每一通道的像素值（原来是 0-255）变到指定的 0-8 的范围，8*8*8 是将每一通道都缩放到 0-8 之间，再进一步求相应颜色直方图。比较半径 R 是指目标元素点与离自己 R 个像素远的像素比较大小。

4. 在空间上叠加四个特征向量图，对相邻位置的图像块特征进行求和池化操作，得到相应局部描述子，如式 2 所示：

$$\widetilde{F}_{p,q} = \sum_{i=1}^4 (F_{m,n}^i + F_{m+1,n}^i + F_{m,n+1}^i + F_{m+1,n+1}^i) \quad (\text{式 2})$$

$$\text{其中: } m = 2 \times (p - 1) + 1, n = 2 \times (q - 1) + 1, p \leq \left\lfloor \frac{M}{2} \right\rfloor, q \leq \left\lfloor \frac{N}{2} \right\rfloor$$

其中, $F_{m,n}^i$ 代表了第 i 张特征图中第 m 行第 n 列的图像块特征向量, 将相邻的 16 个图像块特征向量相加, 得到的 $\widetilde{F}_{p,q}$ 表示一个局部描述子, 其中 p 和 q 为正整数, 分别表示为局部描述子所在的位置;

5. 最后, 根据所有局部描述子在特征图的位置, 按从左上到右下的顺序首尾连接成一个特征向量, 作为整幅图片的对齐局部描述子。

图 2 示意了分级全局特征的提取流程, 首先对图片进行显著性检测, 然后对相同水平方向上的元素值进行对比度增强, 接着对图片进行二值化处理, 通过滑动窗口的方法定位行人区域块, 对行人区域块进行分级求取相应特征均值。提取分级全局特征的过程具体包括如下步骤:

1. 把图片统一缩放到 128*48 大小, 并使用文献[1] (D. J. Jobson, Z.-U. Rahman, and G. A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. Image Processing, IEEE Transactions on, 6(7):965-976, 1997) 所述的多尺度 Retinex 算法消除光照影响, 其中尺度参数分别为 $\sigma=5$ 和 $\sigma=20$ 。
2. 使用文献[2] (Y. Qin, H. Lu, Y. Xu, and H. Wang, "Saliency detection via cellular automata," in Computer Vision and Pattern Recognition (CVPR), 2015, pp. 110-119) 记载的方法对预处理后的图片进行前景提取, 得到相应的显著图。该方法首先利用颜色和空间对比度建立一个基于背景的图。然后, 采用一种新的依赖于元胞自动机的传播机制, 通过与邻域的相互作用来发掘相似区域的内在相关性。构造了影响因子矩阵和一致性矩阵, 以平衡各单元下一状态的影响能力。根据更新规则, 同时更新所有单元的显著性值。
3. 采用式 3 对提取得到的显著图作对比增强:

$$p'_{r,i} = \left\lfloor \frac{p_{r,i}}{\max(p_{r,j=1,2,\dots,n}) - \min(p_{r,j=1,2,\dots,n})} \times 255 \right\rfloor \quad (\text{式 3})$$

其中, 分子 $p_{r,i}$ 为第 r 行第 i 列的元素值, 分母为该行元素的极差。得到该行每列元素的 $p'_{r,i}$ 值后, 统一对该行元素进行更新。

4. 使用文献[3] (N. Otsu, "Threshold selection method from gray-level histograms, IEEE transactions on systems man and cybernetics," Systems Man & Cybernetics IEEE Transactions on, vol. 9, no. 1, pp. 62-66, 1979) 中记载的 OTSU 算法自适应选取阈值,

对上一步得到的结果进行二值化处理，得到相应的二值图。其中，OTSU 算法把图像的灰度数按灰度级分成 2 个部分，使得两个部分之间的灰度值差异最大，每个部分之间的灰度差异最小，通过方差的计算来寻找一个合适的灰度级别来划分。所以可以在二值化的时候采用 OTSU 算法来自动选取阈值进行二值化。

5. 采用滑动窗口的方法对二值图进行扫描，其中窗口大小为 10*10，步长为 5*5，对窗口内的元素值求和，其中把大于设定阈值（如 50）的窗口定义为行人区域，其它则为无关区域。
6. 按照提取对齐局部描述子图像块特征的方法，对行人区域的图像块提取颜色和纹理特征。
7. 可将得到的行人区域的图像块划分为 3 个层级。第一级为人的整体（即全部图像块），第二级是按人体的头、上身、下身分，第三级是按人体上身的重要性来细分三块。

本实施例中，层级 1 为所有行人区域图像块；层级 2 按照图像块所在垂直方向上的位置分为头部区、上半身和下关身区域，其中行数小于等于 15 为头部区，行数大于等于 60 为下半身区、中间为上半身区；层级 3 按照上半身图像块的垂直位置再细分为 3 个区，分界条件为：图像块所属位置在 15 到 30 行之间为一区、在 30 到 45 行之间为一区、在 15 到 60 行为一区。对不同层级里的图像块特征向分别作平均池化操作，分别得到 7 个特征向量，然后连接在一起，作为最终的分级全局特征。

8. 最后，通过采取 XQDA 作为度量学习方法，通过式 4 计算两张图像之间的相似性，值越小即说明特征之间距离越小，对应的两张图片越相似：

$$d(x, z) = d_W(x_{F_{AL}}, z_{F_{AL}}) + \beta \times d_W(x_{F_{HG}}, z_{F_{HG}}) \quad (\text{式 } 4)$$

其中， $d(x, z)$ 是指两张图像（ x 和 z ）之间的距离，越小则表明越相似； $d_W(f_1, f_2)$ 是 XQDA 距离函数， $x_{F_{AL}}, x_{F_{HG}}$ 分别为图像 x 的对齐局部描述子和分级全局特征，其中 β 表示全局特征的权重，在这里 $\beta = 1.3$ 。

本发明具体实施中，将本发明方法与现有其他方法进行对比分析，以验证本发明方法的有效性。表 1 和表 2 中，LOMO（Local Maximum Occurrence，局部直方图最大频数特征）在文献（S. Liao, Y. Hu, X. Zhu, and S. Z. Li, Person re-identification by local maximal occurrence representation and metric learning," in Computer Vision and Pattern Recognition (CVPR), 2015, pp. 2197-2206.）中记载。SCNCD（Salient Color Names based Color Descriptor，基于颜色描述子

的显著颜色名)在文献(Y. Yang, Salient color names for person re-identification," in European Conference on Computer Vision (ECCV), 2014, pp. 536-551.)中记载。kBiCov (Covariance descriptor based on bio-inspired features, 基于生物启发特征的协方差描述子)在文献(B. Ma, Y. Su, and F. Jurie, Covariance descriptor based on bio-inspired features for person re-identification and face verification," Image & Vision Computing, vol. 32, no. 6-7, pp. 379-390, 2014.)中记载。LDFV (Local descriptors encoded by fisher vectors, 使用 Fisher 向量编码的局部描述子)在文献(B. Ma, Y. Su, and F. Jurie, Local descriptors encoded by fisher vectors for person re-identification," in European Conference on Computer Vision (ECCV), 2012, pp. 413-422.)中记载。SCSP (Spatially Constrained Similarity function on Polynomial feature map, 基于多项式特征图的部分约束相似函数)在文献(D. Chen, Z. Yuan, B. Chen, and N. Zheng, Similarity learning with spatial constraints for person re-identification," in Computer Vision and Pattern Recognition (CVPR), 2016, pp. 1268-1277.)中记载。DR-KISS (dual-regularized kiss, 双正则化 KISS 算法)在文献(D. Tao, Y. Guo, M. Song, and Y. Li, Person re-identification by dual-regularized kiss metric learning," IEEE Transactions on Image Processing, vol. 25, no. 6, pp. 2726-2738, 2016.)中记载。SSDAL (semi-supervised deep attribute learning, 半监督深度属性学习)在文献(C. Su, S. Zhang, J. Xing, W. Gao, and Q. Tian, Deep attributes driven multi-camera person re-identification," in European Conference on Computer Vision (ECCV), 2016, pp. 475-491.)中记载。NLML (Nonlinear local metric learning, 非线性局部度量学习)在文献(S. Huang, J. Lu, J. Zhou, and A. K. Jain, Nonlinear local metric learning for person re-identification," arXiv preprint arXiv:1511.05169v1, 2015.)中记载。LOMO (Local Maximum Occurrence, 局部直方图最大频数特征)在文献(S. Liao, Y. Hu, X. Zhu, and S. Z. Li, Person re-identification by local maximal occurrence representation and metric learning," in Computer Vision and Pattern Recognition (CVPR), 2015, pp. 2197-2206.)中记载。

图 3 和图 5 分别对比了使用本发明所提出的方法和使用 LOMO 特征在 VIPeR 和 QMUL GRID 数据集上的表现,可以看出,我们的方法都取得了较好的效果。图 4 和图 6 对本发明所提出的方法进行了不同情况下的对比实验,分别为使用了结合对齐局部描述子和分级全局特征的完整方法、只使用了对齐局部描述子作为特征、只使用了分级全局特征作为特征和使用了不做仿射变换的对齐局部描述子和分级全局特征结合的方法,结果展示了局部描述子和全局特征在不同数据集上的不同表现,并且结合的效果达到最优,说明本专利所提出方法的

有效性。

表 1. 本发明方法和现有方法在 VIPeR 数据上的 CMC 结果

方法	Rank=1	Rank=10	Rank=20	参考来源
本发明方法	46.2	85.3	94.4	/
LOMO	40.0	80.5	91.1	CPPR 2015
SCNCD	37.8	81.2	90.4	ECCV 2014
kBiCov	31.1	70.7	82.5	IVC 2014
LDFV	26.5	70.9	84.6	ECCVW 2012

表 2. 本发明方法和现有方法在 QMUL GRID 数据上的 CMC 结果

方法	Rank=1	Rank=5	Rank=10	Rank=20	参考来源
本发明方法	25.6	47.0	56.6	66.3	/
SCSP	24.2	44.6	54.1	65.2	CVPR 2016
DR-KISS	20.6	39.3	51.4	62.6	TIP 2016
SSDAL	22.4	39.2	48.0	58.4	ECCV 2016
NLML	24.5	35.9	43.5	55.3	arXiv 2015
LOMO	16.6	33.8	41.8	52.4	CVPR 2015

表 1 列出了其它一些方法在 VIPeR 数据上的 CMC (Cumulative Match Characteristic, 累计匹配特性) 结果, 与之对比, 使用本发明方法识别性能有较好的提升。表 2 列出了其它一些方法在 QMUL GRID 数据上的 CMC 结果, 与之对比, 本发明方法的准确率最高, 说明了本发明方法的有效性。

需要注意的是, 公布实施例的目的在于帮助进一步理解本发明, 但是本领域的技术人员可以理解: 在不脱离本发明及所附权利要求的精神和范围内, 各种替换和修改都是可能的。因此, 本发明不应局限于实施例所公开的内容, 本发明要求保护的范围以权利要求书界定的范围为准。

权 利 要 求 书

1、一种用于行人重识别的图像特征提取方法，通过对齐局部描述子提取和分级全局特征提取，进行行人重识别；所述对齐局部描述子提取采用仿射变换对原图像进行处理，对相邻区域的图像块特征进行求和池化操作而得到对齐局部描述子；所述对齐局部描述子保留图像内部块与块之间的空间信息；所述分级全局特征提取通过对定位的行人区域块进行分级，求取相应特征均值而得到全局特征；包括如下步骤：

- 1) 对图像进行了预处理，以消除光照对图像的影响；
- 2) 提取对齐局部描述子，包括步骤 21) ~24):
 - 21) 对图像进行仿射变换，得到多张图像；
 - 22) 对仿射变换后产生的所有图像进行图像块分割并提取特征，生成相应的特征向量图；
 - 23) 在空间上叠加所有特征向量图，对相邻位置的图像块特征进行求和池化操作，得到相应局部描述子；
 - 24) 将图像中的所有局部描述子按顺序连接起来，得到对齐局部描述子；
- 3) 提取分级全局特征，包括步骤 31) ~36):
 - 31) 使用前景提取方法对图像进行显著性检测，得到相应的显著图；
 - 32) 将得到的显著图进行对比增强；
 - 33) 定位属于行人区域的图像块；
 - 34) 对行人区域的图像块分层级，得到多级行人区域图像块；
 - 35) 对每一级的行人区域图像块进行平均池化操作，得到该级特征；
 - 36) 对多级特征连接起来，得到分级全局特征；
- 4) 根据步骤 2) 提取得到的对齐局部描述子和步骤 3) 提取得到的分级全局特征，采用度量学习方法计算图像的相似性，由此进行行人重识别。

2、如权利要求 1 所述图像特征提取方法，其特征是，步骤 1) 采用多尺度 Retinex 算法对图像进行预处理。

3、如权利要求 1 所述图像特征提取方法，其特征是，步骤 21) 对图像进行仿射变换，具体是：对预处理后的图片 I 进行水平翻转，得到图片 II；分别对图片 I 和 II 进行水平剪切变换，产生相应图片 III 和 IV；所述水平剪切变换如式 1 所示：

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ \lambda & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (\text{式 } 1)$$

其中， λ 为剪切变换参数。

4、如权利要求 1 所述图像特征提取方法，其特征是，步骤 22) 采用滑动窗口方法求图像块的特征，具体对窗口内的图像块提取 RGB 颜色直方图、HSV 颜色直方图、Lab 颜色直方图、尺度不变局部三值模式 SILTP 统计直方图，得到多个特征向量图，每个特征图中有 $M \times N$ 个点，每个点对应一个图像块的特征向量，记作 $F_{m,n}^i$ 。

5、如权利要求 1 所述图像特征提取方法，其特征是，步骤 23) 具体通过式 2 在空间上叠加四个特征向量图，对相邻位置的图像块特征进行求和池化操作，得到相应局部描述子：

$$\widetilde{F}_{p,q} = \sum_{i=1}^4 (F_{m,n}^i + F_{m+1,n}^i + F_{m,n+1}^i + F_{m+1,n+1}^i) \quad (\text{式 } 2)$$

$$\text{其中: } m = 2 \times (p - 1) + 1, n = 2 \times (q - 1) + 1, p \leq \left\lfloor \frac{M}{2} \right\rfloor, q \leq \left\lfloor \frac{N}{2} \right\rfloor$$

其中， $F_{m,n}^i$ 代表第 i 张特征图中第 m 行第 n 列的图像块特征向量；将相邻的图像块特征向量相加，得到的 $\widetilde{F}_{p,q}$ 表示一个局部描述子，其中 p 和 q 为正整数，分别表示局部描述子所在的位置。

6、如权利要求 1 所述图像特征提取方法，其特征是，步骤 31) 具体采用元胞自动机方法对预处理后的图片进行前景提取，得到相应的显著图。

7、如权利要求 1 所述图像特征提取方法，其特征是，步骤 32) 具体采用式 3 对提取得到的显著图作对比增强：

$$p'_{r,i} = \left\lfloor \frac{p_{r,i}}{\max(p_{r,j=1,2,\dots,n}) - \min(p_{r,j=1,2,\dots,n})} \times 255 \right\rfloor \quad (\text{式 } 3)$$

其中，分子 $p_{r,i}$ 为第 r 行第 i 列的元素值，分母为该行元素的极差；得到该行每列元素的 $p'_{r,i}$ 值后，统一对该行元素进行更新。

8、如权利要求 1 所述图像特征提取方法，其特征是，步骤 33) 定位属于行人区域的图像块，首先采用 OTSU 算法自适应选取阈值，对经对比增强得到的结果图像进行二值化处理，得到相应的二值图；再采用滑动窗口方法对二值图进行扫描，对窗口内的元素值求和；通过设定阈值，将大于设定阈值的窗口定义为行人区域，其它为无关区域。

9、如权利要求 1 所述图像特征提取方法，其特征是，步骤 34) 具体将行人区域的图像块分为三个层级，得到三个层级的行人区域图像块；层级一为所有行人区域图像块；层级二

按照图像块所在垂直方向上的位置分为头部区、上半身和下半身区域；层级三按照上半身图像块的垂直位置再细分为三个区。

10、如权利要求 1 所述图像特征提取方法，其特征是，步骤 4) 采用 XQDA 作为度量学习方法计算图像的相似性，具体通过式 4 计算两张图像之间的相似性：

$$d(x, z) = d_W(x_{F_{AL}}, z_{F_{AL}}) + \beta \times d_W(x_{F_{HG}}, z_{F_{HG}}) \quad (\text{式 4})$$

其中， $d(x, z)$ 是两张图像（ x 和 z ）之间的距离，越小则表明越相似； $d_W(f_1, f_2)$ 是 XQDA 距离函数； $x_{F_{AL}}$ 、 $x_{F_{HG}}$ 分别为图像 x 的对齐局部描述子和分级全局特征； β 表示全局特征的权重。

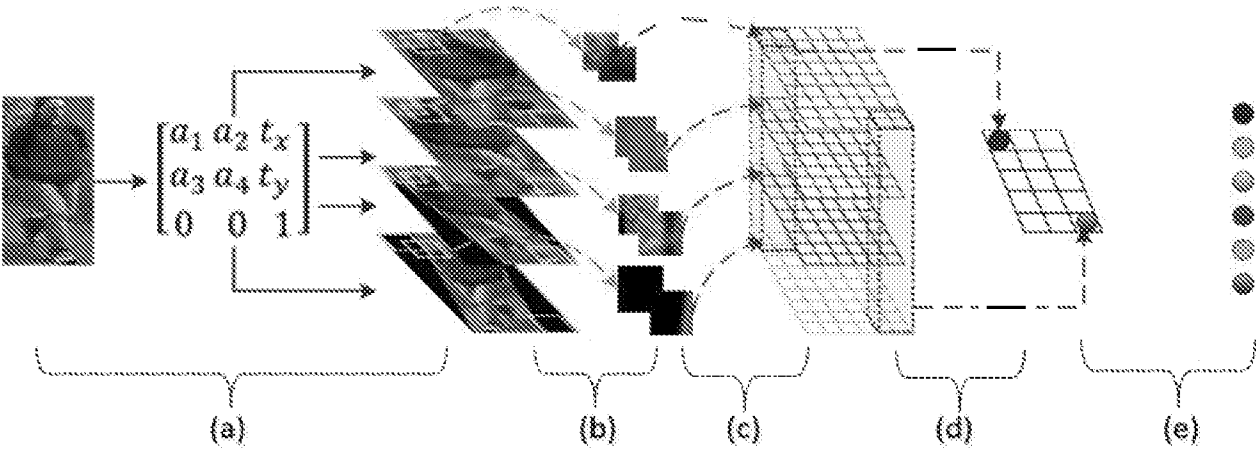


图 1

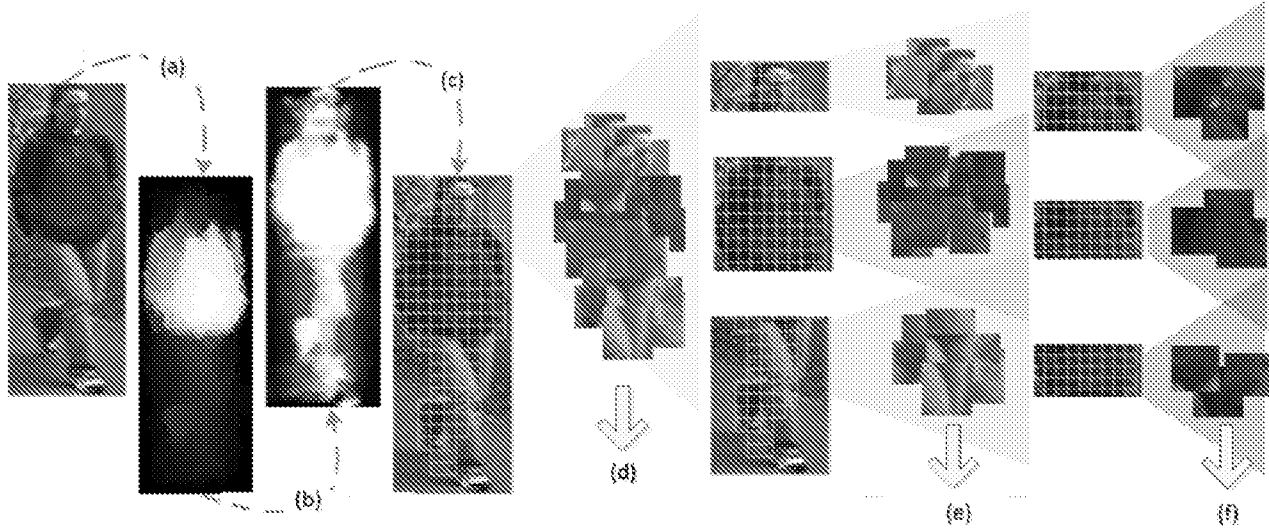


图 2

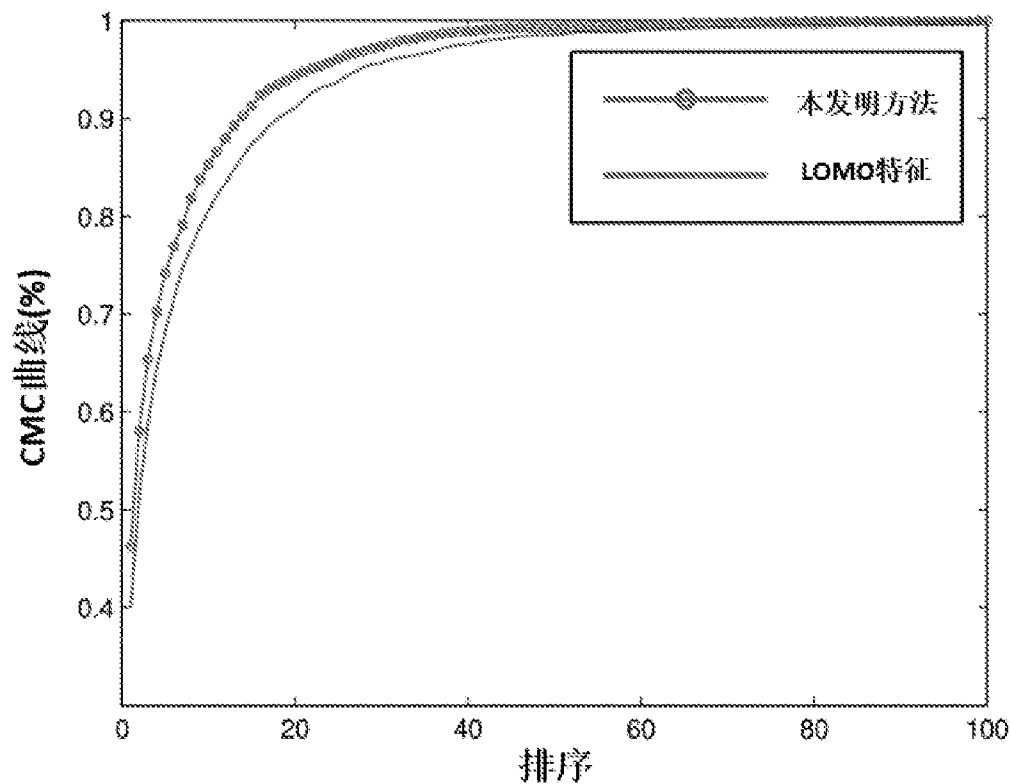


图 3

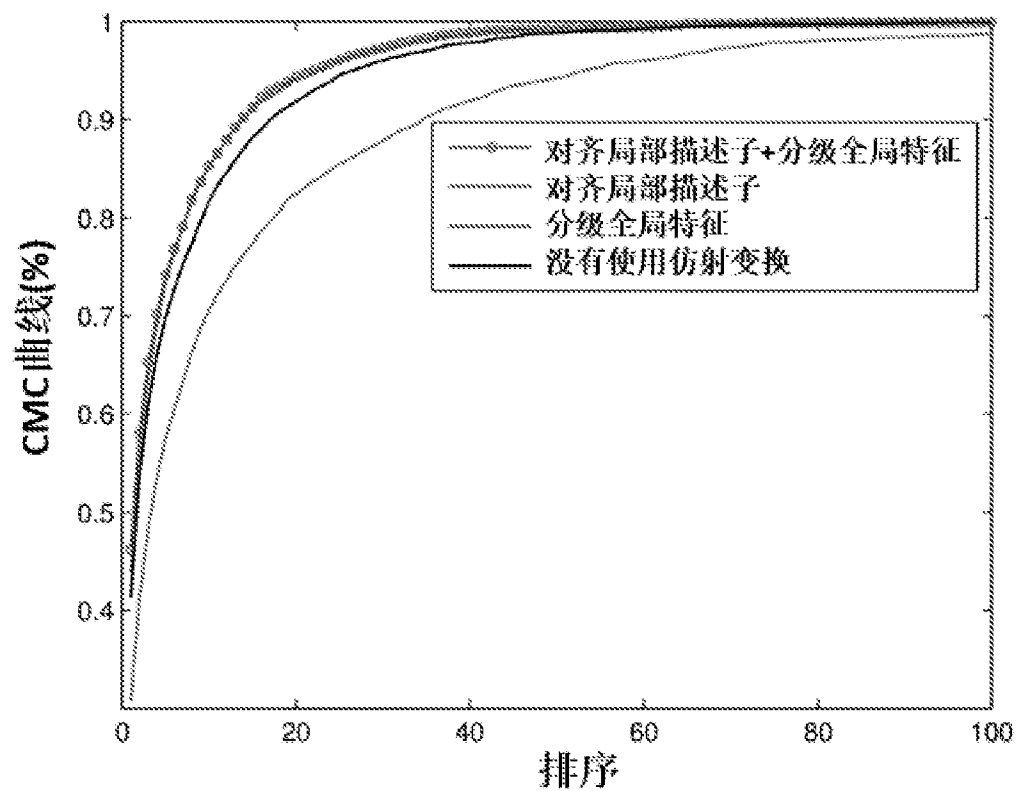


图 4

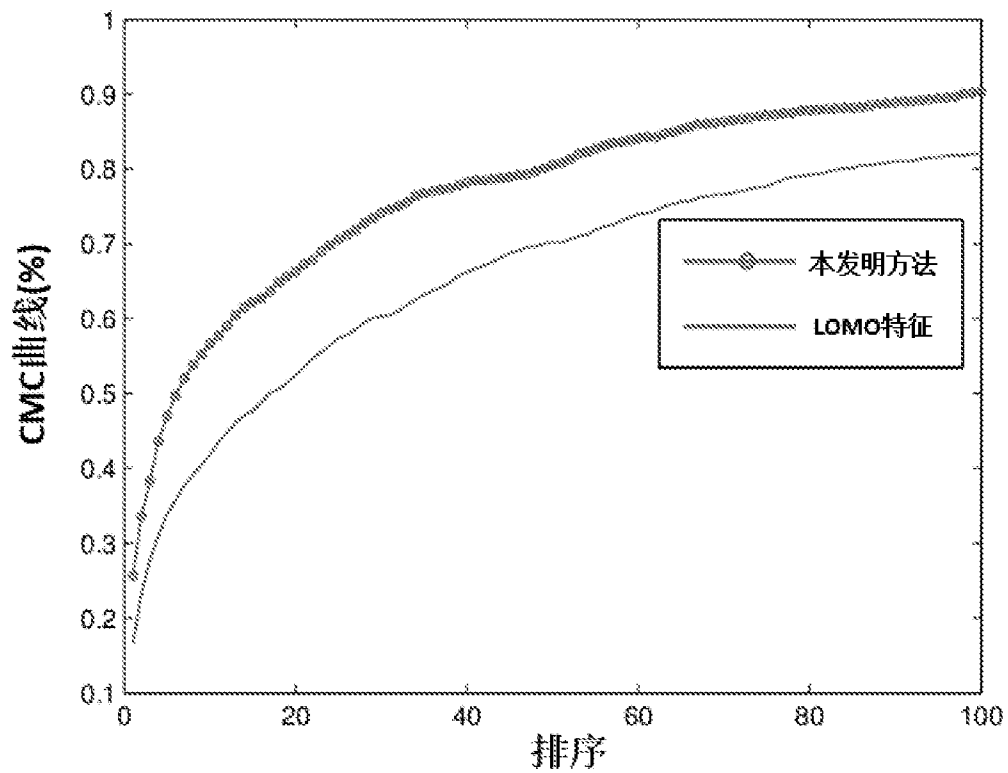


图 5

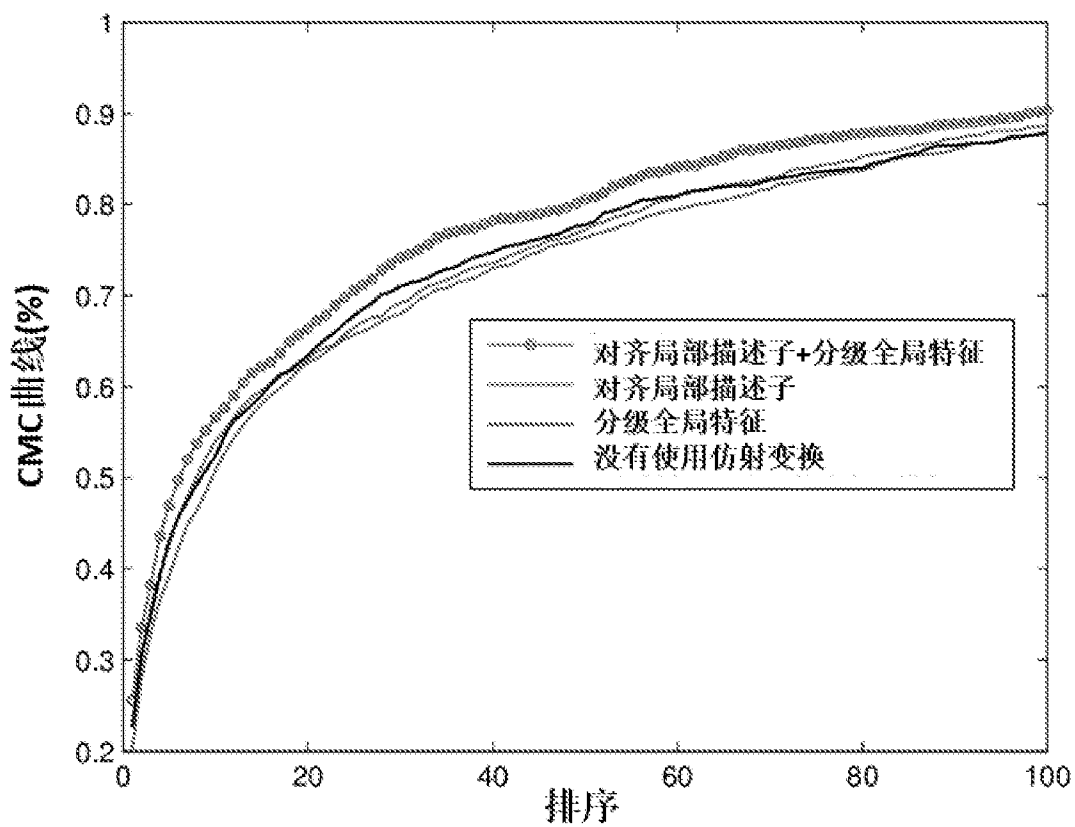


图 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/118794

A. CLASSIFICATION OF SUBJECT MATTER

G06K 9/00 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, IEEE: 行人, 识别, 图像, 特征, 提取, 局部, 描述子, 分级, 光照, 变换, pedestrian, identify, image, feature, extract, global, description, grade, light, transform

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
PX	CN 107316031 A (PEKING UNIVERSITY SHENZHEN GRADUATE SCHOOL) 03 November 2017 (03.11.2017), claims 1-10, and description, paragraphs [0010] and [0062]	1-10
A	CN 102663409 A (XIDIAN UNIVERSITY) 12 September 2012 (12.09.2012), description, paragraphs [0062]-[0095]	1-10
A	CN 105975929 A (PEKING UNIVERSITY SHENZHEN GRADUATE SCHOOL) 28 September 2016 (28.09.2016), entire document	1-10
A	CN 106650806 A (PEKING UNIVERSITY SHENZHEN GRADUATE SCHOOL) 10 May 2017 (10.05.2017), entire document	1-10
A	US 2005117779 A1 (KONICA MINOLTA HOLDINGS, INC.) 02 June 2005 (02.06.2005), entire document	1-10

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
“A” document defining the general state of the art which is not considered to be of particular relevance	“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
“E” earlier application or patent but published on or after the international filing date	“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	“&” document member of the same patent family
“O” document referring to an oral disclosure, use, exhibition or other means	
“P” document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search 22 March 2018	Date of mailing of the international search report 30 March 2018
--	---

Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer JU, Bo Telephone No. (86-10) 53961534
---	--

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.
PCT/CN2017/118794

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 107316031 A	03 November 2017	None	
CN 102663409 A	12 September 2012	None	
CN 105975929 A	28 September 2016	WO 2017190574 A1	09 November 2017
CN 106650806 A	10 May 2017	None	
US 2005117779 A1	02 June 2005	JP 2005157906 A	16 June 2005

国际检索报告

国际申请号

PCT/CN2017/118794

A. 主题的分类 G06K 9/00 (2006.01) i 按照国际专利分类 (IPC) 或者同时按照国家分类和 IPC 两种分类		
B. 检索领域 检索的最低限度文献 (标明分类系统和分类号) G06K 包含在检索领域中的除最低限度文献以外的检索文献 在国际检索时查阅的电子数据库 (数据库的名称, 和使用的检索词 (如使用)) WPI, EPODOC, CNPAT, CNKI, IEEE: 行人, 识别, 图像, 特征, 提取, 局部, 描述子, 分级, 光照, 变换, pedestrian, identify, image, feature, extract, global, description, grade, light, transform		
C. 相关文件		
类 型*	引用文件, 必要时, 指明相关段落	相关的权利要求
PX	CN 107316031 A (北京大学深圳研究生院) 2017年 11月 3日 (2017 - 11 - 03) 权利要求1-10、说明书第[0010]、[0062]段	1-10
A	CN 102663409 A (西安电子科技大学) 2012年 9月 12日 (2012 - 09 - 12) 说明书第[0062]-[0095]段	1-10
A	CN 105975929 A (北京大学深圳研究生院) 2016年 9月 28日 (2016 - 09 - 28) 全文	1-10
A	CN 106650806 A (北京大学深圳研究生院) 2017年 5月 10日 (2017 - 05 - 10) 全文	1-10
A	US 2005117779 A1 (KONICA MINOLTA HOLDINGS, INC.) 2005年 6月 2日 (2005 - 06 - 02) 全文	1-10
☐ 其余文件在C栏的续页中列出。 ☒ 见同族专利附件。		
* 引用文件的具体类型: “A” 认为不特别相关的表示了现有技术一般状态的文件 “E” 在国际申请日的当天或之后公布的在先申请或专利 “L” 可能对优先权要求构成怀疑的文件, 或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件 (如具体说明的) “O” 涉及口头公开、使用、展览或其他方式公开的文件 “P” 公布日先于国际申请日但迟于所要求的优先权日的文件 “T” 在申请日或优先权日之后公布, 与申请不相抵触, 但为了理解发明之理论或原理的在后文件 “X” 特别相关的文件, 单独考虑该文件, 认定要求保护的发明不是新颖的或不具有创造性 “Y” 特别相关的文件, 当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时, 要求保护的发明不具有创造性 “&” 同族专利的文件		
国际检索实际完成的日期		国际检索报告邮寄日期
2018年 3月 22日		2018年 3月 30日
ISA/CN的名称和邮寄地址		受权官员
中华人民共和国国家知识产权局 (ISA/CN) 中国北京市海淀区蓟门桥西土城路6号 100088 传真号 (86-10) 62019451		鞠博 电话号码 (86-10) 53961534

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2017/118794

检索报告引用的专利文件			公布日 (年/月/日)	同族专利			公布日 (年/月/日)
CN	107316031	A	2017年 11月 3日	无			
CN	102663409	A	2012年 9月 12日	无			
CN	105975929	A	2016年 9月 28日	WO	2017190574	A1	2017年 11月 9日
CN	106650806	A	2017年 5月 10日	无			
US	2005117779	A1	2005年 6月 2日	JP	2005157906	A	2005年 6月 16日