

基于图像全局矢量的特征点矢量提取方法

邹 彬, 陆 阳, 钱 钧, 贺峻峰, 李良福, 杨一洲

(西安应用光学研究所, 陕西 西安 710065)

摘 要: 针对传统二进制特征描述子提取算法复杂, 鲁棒性较差等缺点, 提出了一种具有一定全局参考性及鲁棒性的二进制特征矢量; 首先, 对输入的匹配图像求取其 Harris 特征点及其显著图; 然后, 将图像空间网格化, 求取各网格中心, 并按其对应网格的显著度进行排序, 生成全局基准矢量; 最后, 将各个特征点与全局基准矢量进行比较, 生成基于显著度加权的二进制特征矢量. 实验结果表明, 该算法提取的二进制特征在极大减少计算的同时, 还具有良好的特征表征能力.

关键词: 二进制特征; 显著性; 基准矢量; 网格中心

中图分类号: TP391. 41

文献标识码: A

文章编号: 1000-7180(2017)01-0058-04

Binary Image Feature Extraction Method Based on the Image Global Binary Vector

ZOU Bin, LU Yang, QIAN Jun, HE Jun-feng, LI Liang-fu, YANG Yi-zhou

(Xi'an Institute of Applied Optics, Xi'an 710065, China)

Abstract: To overcome the weakness of algorithm complexity and less robust by traditional binary feature extraction algorithm, this paper presents a novel binary feature extraction method with certain global reference and robustness. Firstly, Harris feature points and Saliency map are obtained from the input matching image. Secondly, the image space is partitioned into grid in order to get the grid center, which is sorted according to the saliency degree of the corresponding grid to generate the global reference vector. Finally, the feature points are compared with the global reference vector to generate the binary feature based on the saliency weighted. Experimental results show that the proposed algorithm can greatly reduce the computational burden, and has a good ability of characterization.

Key words: binary feature; saliency; reference vector; grid center

1 引言

近几十年来, 特征点描述子已经成为计算机视觉领域中, 尤其是实现图像配准不可或缺的工具; 同时, 如何在计算能力和存储空间都有限的硬件条件下, 快速提取并合理表征成百上千个特征点, 也一直是众多科研工作者持续研究的热点问题.

在目前众多的图像局部描述子中, 最具代表性的为 SIFT^[1-2]、SURF^[3-4] 等, 其被广泛应用于各类图像系统中. 虽然, 这类传统局部描述子具有较好的效果, 但其特征矢量却付出了高纬度的代价, 尤其在大规模图像库应用中, 特征点数据量呈指数级翻倍

增长, 同时, 由于其特征点间的相似性度量采用欧式距离, 计算复杂度伴随巨大的计算量, 对特征点的邻近搜索提出了严峻挑战; 无论是硬件条件, 还是实时性的要求, 均制约其实际应用.

针对传统特征描述子实际应用中所暴露出来的算法复杂、内存需求高以及计算耗时等问题, 近年来, 许多学者开始着重研究基于二进制表征的局部特征. 其中, 最早最著名的便是由 Calonder 等人提出的 BRIEF^[5] 二进制特征描述子, 该特征描述子充分挖掘特征点的临域信息, 在此范围内随机采样像素点, 并与特征点进行二值化的幅值比较, 以生成特征矢量. 该算法天然继承了二进制运算所具有的内

存需求低和计算速度快等特点.虽然以 BRIEF^[5]为代表的二进制特征描述子提取简单,计算速度快,但是采样点对的选取具有随机性,且容易受到局部噪声干扰,同时,二进制特征具有较弱的区分性和鲁棒性;这些都导致了描述子对图像旋转和尺度变化很敏感,影响其表征能力及匹配精度,制约其进一步的广泛应用^[6].

为了避免现有技术的缺陷,本文提出了一种基于图像全局矢量的二进制特征提取方法,其目的在于提取出具有一定全局参考性及鲁棒性的二进制特征.它通过将图像空间网格化,并基于图像的视觉显著性生成全局基准矢量,然后在此基础上进行二进制特征矢量的提取;在提高特征表征能力的同时,极大避免和减少了特征点对的采集与计算.

2 二进制特征矢量提取算法

以上回顾了传统二进制特征提取算法,下面将介绍本文基于图像全局矢量的特征提取算法.该方法以图像空间网格化为手段,以图像视觉显著性为基础,生成具有全局参考性的基准矢量,并借此提取综合性能更好的二进制特征,主要流程包括 4 个部分.

2.1 Harris 特征点及显著图的概述

应用 Harris 算子提取特征点具有比较稳定的特性,最早由 C. Harris 和 M. J. Stephens 等^[7]学者于 1988 年提出. Harris 算子根据单像素点邻近区域,即某尺寸窗口范围内像素平均量能的变化,而判断该点是否为特征点.由于其具体计算过程主要为导数运算,所以对图像的亮度变化、旋转缩放及噪声等具有较好的鲁棒性.本文采用 Harris 算子提取输入匹配图像的特征点,也便于同其他相关工作进行试验比较,记求取的 Harris 特征点集合为 $P = \{p_i, 1 \leq i \leq L\}$,其中 L 为特征点总数.

而人眼视觉系统是具有主观能动性的智能系统,表现在人们往往能将自身目光的注意力迅速定格在思想意识所感兴趣的目标或区域.这一过程具体包括两个方面,首先是对底层视觉特征的提取和加工,然后是思想控制对该过程的方向引导,两者相互配合共同完成人眼感兴趣区域的选择.而该区域本身能否引起人眼视觉的注意,则主要由其显著性决定.在众多注意力模型中,我们选择最经典的由美国加州大学提出的 Itti 注意力模型^[8],其比较适合处理自然图像,主要思想是对输入图像首先进行多个特征通道和多尺度的金字塔分解,再进行滤波得

到特征图,再对特征图做融合计算得到最终的显著图.记求取输入匹配图像的显著图为 $S(x, y), 1 \leq x \leq W, 1 \leq y \leq H, W \times H$ 为输入图像的空间尺寸, (x, y) 为像素点坐标.

2.2 图像空间网格化

传统二进值提取算法中,需要在特征点周围进行随机像素采样,以组成特征点对序列.这种算法一方面容易受到噪声及图像变换的影响,另一方面采集的是单个特征点的局部区域特征,具有不稳定性,同时计算负担较大.而本文算法

在图像空间上进行标准网格化的划分,各个网格可以看做图像局部特征的重要组成部分^[9],并以各网格中心与特征点组合成点对序列,不但可以避免噪声的影响,而且使得各个特征点具有一组统一的基准比较序列,同时极大减少了计算任务.其具体步骤如下.

① 假设输入匹配图像的空间尺寸为 $W \times H$,将 W 和 H 分量各等分为 η 个区间,则图像空间被划分为互不交叠的网格单元,图像中的每个像素依据其空间位置坐标隶属于一个特定的网格,记为 $\text{Cell}_i = \{I_1, I_2, \dots, I_j, \dots, I_N\}$,即第 i 个网格, I_j 为图像灰度值,其中 $i = 1, \dots, M, j = 1, \dots, N, M = \eta \times \eta$ 代表该图像具有的网格总数, $N = (W/\eta) \times (H/\eta)$ 代表该网格中的像素总数;

② 求取各网格中心:

$$g_i = \frac{1}{N} \sum_{p_j \in \text{Cell}_i} I_j, i = 1, \dots, M, j = 1, \dots, N \quad (1)$$

式中, $I_j \in \text{Cell}_i$ 表示 Cell_i 网格中包含的像素点 $M = \eta \times \eta, N = (W/\eta) \times (H/\eta)$.

2.3 生成图像全局基准矢量

生成图像全局基准矢量的意义在于为所有特征点提供一个统一的比较平台,而不是各个点独立的采用临近区域的点对序列,从而有利于增强同一幅图像在不同情境下特征点之间的区分度.

首先,对各网格中心按其对应网格的显著度进行排序,生成全局基准矢量;记根据 Itti 模型^[8]求取输入匹配图像的显著图为 $S(x, y), 1 \leq x \leq W, 1 \leq y \leq H$,则各网格的显著度为

$$S_i = \sum_{s_j \in \text{Cell}_i} s_j, i = 1, \dots, M, j = 1, \dots, N \quad (2)$$

式中, $s_j \in \text{Cell}_i$ 表示网格 Cell_i 中包含的像素点 I_j 所对应的显著度 $s_j, M = \eta \times \eta, N = (W/\eta) \times (H/\eta)$.

然后,对各网格中心按照其对应网格显著度的大小进行降序排列,得到全局基准矢量为

$$G = \arg \text{desc}\{g_i, 1 \leq i \leq M \mid S_k \geq S_l, k \leq L, 1 \leq l \leq M\} \quad (3)$$

式中, $M = \eta \times \eta$, 这样, 便构成了一个具有旋转不变性的全局基准矢量, 如图 1 所示。

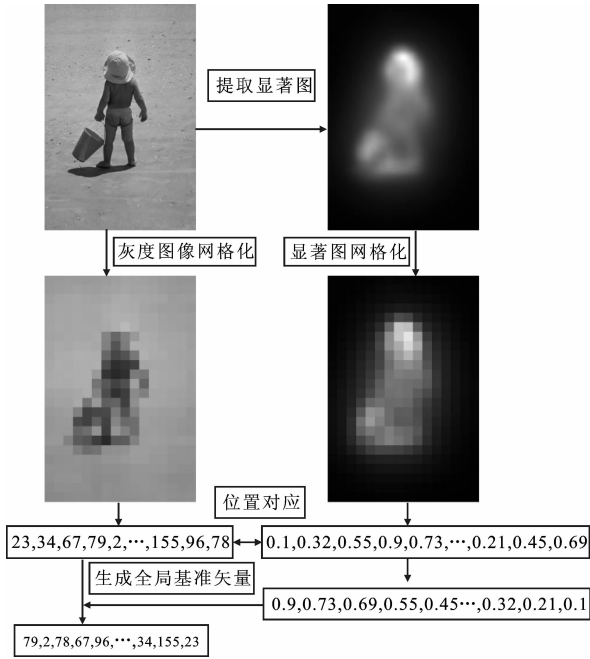


图 1 生成全局基准矢量示意图

2.4 基于显著度加权的特征提取

为了使提取的特征更符合视觉意义, 在进行特征点对的比较时, 没有进行简单的数值比较, 而是基于其各自的显著值进行加权处理, 具体方法如下所示。

记对输入匹配图像求取的 Harris 特征点为 $P = \{p_i, 1 \leq i \leq L\}$, 其中 L 为特征点总数, 记各个特征点的二进制特征集合为 $T = \{T_i, 1 \leq i \leq L\}$, 则 $T_i = (T_i^1, T_i^2, \dots, T_i^L, \dots, T_i^M)$ 由下式生成:

$$T_i^j = \begin{cases} 1, & s_i I_i \geq S_j g_j \\ 0, & s_i I_i < S_j g_j \end{cases}, 1 \leq j \leq M, 1 \leq i \leq L, \quad (4)$$

式中, M 为网格总数, L 为特征点总数, s_i 为第 p_i 个特征点所对应显著度, g_j 和 S_j 分别为第 Cell_j 个网格中心值及其对应的网格显著度。

然后, 在实际应用中, 可以通过对特征点间的二值特征矢量进行海明距离度量, 作为特征点间的相似度距离, 用以进行图像匹配等应用。

3 实验结果

为了验证和评价本文算法, 实验所用的图像数据集为 Mikolajczyk^[10] 等人提出的牛津标准数据集, 含有 8 组 (每组 6 幅) 共 48 幅图像, 包括 6 种不同

变换 (视角改变、亮度变化、图像旋转、尺度变换、图像模糊和图像压缩)。并采用与其相同的评价标准, 即召回率 recall 和 1-precision:

$$\begin{cases} \text{recall} = \frac{m_c}{n_c} \\ 1 - \text{precision} = \frac{m_f}{n} \end{cases} \quad (5)$$

式中, m_f 为实验错误匹配数, m_c 为实验正确匹配数, n_c 为所有正确匹配数, n 为所有匹配数。在实验中将本文算法与 BRIEF^[5]、FREAK^[11] 进行比较, 实验将第 1 类图像中的特征点分别与第 2 类、第 3 类图像中的特征点进行匹配, 空间等分区间 η 的取值满足最终基准矢量为 128 维, 同时每幅图像最多检测 100 个特征点。

对比实验结果如图 2 所示, 本文算法相对于传统算法, 在图像模糊 (bikes) 及旋转和尺度变化 (boat) 情况下, 均获得了较好的性能。

4 结束语

本文通过图像空间网格化, 将各网格中心作为与特征点比较的采样点, 一方面, 为不同的特征点提供了一个统一的基准比较平台, 有利于提升特征向量的表征能力; 另一方面, 对所有特征点采用同一个全局基准矢量进行比较, 不但极大降低了对系统内存的需求, 而且避免了对各个特征点进行随机点对采样所耗费的计算资源与时间; 同时, 求取网格均值作为网格中心, 有效避免了传统方法中采样单个像素点易受噪声影响等不良因素; 并且, 在生成全局基准矢量过程中, 对各网格中心按其对应网格的显著度进行降序排列, 由于视觉显著性不易受到图像旋转、尺度缩放、亮度变化等因素影响, 从而使得基于此提取的特征矢量具有较好的鲁棒性; 此外, 在最后将各个特征点与全局基准矢量进行比较时, 对其进行基于显著度的加权处理, 使特征矢量的提取符合一定的人眼视觉特性, 有效提高了二进制特征的区分度及其描述表征能力。

参考文献:

- [1] Lowe D G. Distinctive image features from scale-invariant key-points[J]. International Journal of Computer Vision, 2004, 60: 91-110.
- [2] 刘辉, 申海龙. 一种基于改进 SIFT 算法的图像配准方法[J]. 微电子学与计算机, 2014, 31(1): 38-42.
- [3] Bay H, Tuytelaars T, Van G. SURF: speeded up robust features[C] // Proceedings of the 9th European Conference on Computer Vision Graz. Berlin: Springer, 2006, 3951: 404-417.

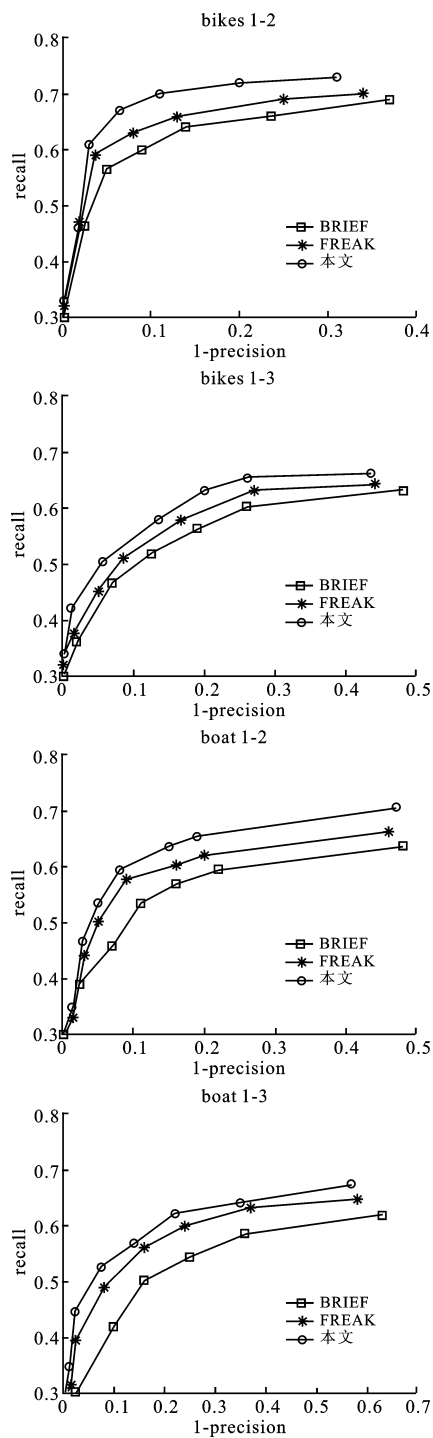


图 2 对比实验结果图

[4] 尚明珠. 一种基于改进 SURF 的图像匹配方法[J]. 微电子学与计算机, 2014, 31(2): 125-128.

[5] Calonder M, Lepetit V, Ozuysal M, et al. BRIEF: computing a local binary descriptor very fast[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34: 1281-1298.

[6] 庄东晔, 张冬明, 张勇东, 等. 随机亮度差量化的二进制特征描述[J]. 中国图象图形学报, 2014, 19(4): 630-636.

[7] Harris C, Stephens M. A combined corner and edge detector[C] // Proceedings of Alvey Vision Conference. Manchester, Britain: BMVA, 1988: 147-151.

[8] Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20 (11): 1254-1259.

[9] Chan K W, Chan K L. Optimisation of multi-level block truncation coding[J]. Signal Processing: Image Communication, 2001, 16: 445-459.

[10] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2005 (27): 1615-1630.

[11] Alahi A, Ortiz R, Vandergheynst P. FREAK: fast retina key-point[C] // Proceedings of the 25th IEEE Conference on Computer Vision and Pattern Recognition. Rhode Island, Providence, USA: IEEE, 2012: 16-21.

作者简介:
邹彬男, (1986-), 工程师. 研究方向为图像处理与识别.
E-mail: zoubean@163. com.