

DOI: 10.11992/tis.201709014

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.tp.20180420.1018.002.html>

改进 SURF 特征的维吾尔文复杂文档图像匹配检索

阿丽亚·巴吐尔¹, 努尔毕亚·亚地卡尔¹, 吾尔尼沙·买买提¹, 阿力木江·艾沙², 库尔班·吾布力¹

(1. 新疆大学 信息科学与工程学院, 新疆 乌鲁木齐, 830046; 2. 新疆大学 网络与信息中心, 新疆 乌鲁木齐, 830046)

摘要: 针对图像局部特征的字袋模型 (Bag-of-Word, BOW) 检索研究中聚类中心的不确定性和计算复杂性问题, 提出一种由不同种类的距离进行相似程度测量的检索和由匹配点数来检索的方法。这种方法首先需要改进文档图像的 SURF 特征, 有效降低特征提取复杂度; 其次, 对 FAST+SURF 特征实现 FLANN 双向匹配与 KD-Tree+BBF 匹配, 在不同变换条件下验证特征鲁棒性; 最后, 基于这两种检索方法对已收集整理好的各类维吾尔文文档图像数据库进行检索。实验结果表明: 基于距离的相似性度量复杂度次于基于匹配数目的检索, 而且两种检索策略都能满足快速、精确查找需求。

关键词: 复杂文档; 维吾尔文档图像; 文档图像分割; 特征提取; SURF 特征; FLANN 双向匹配; KD-Tree+BBF 匹配; 图像检索

中图分类号: TP391.1 **文献标志码:** A **文章编号:** 1673-4785(2019)02-0296-10

中文引用格式: 阿丽亚·巴吐尔, 努尔毕亚·亚地卡尔, 吾尔尼沙·买买提, 等. 改进 SURF 特征的维吾尔文复杂文档图像匹配检索[J]. 智能系统学报, 2019, 14(2): 296-305.

英文引用格式: ALIYA Batur, NURBIYA Yadikar, HORNISA Mamat, et al. Complex Uyghur document image matching and retrieval based on modified SURF feature[J]. CAAI transactions on intelligent systems, 2019, 14(2): 296-305.

Complex Uyghur document image matching and retrieval based on modified SURF feature

ALIYA Batur¹, NURBIYA Yadikar¹, HORNISA Mamat¹, ALIMJAN Aysa², KURBAN Ubul¹

(1. School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China; 2. Network and information center, Xinjiang University, Xinjiang University, Urumqi 830046, China)

Abstract: This study is aimed at the uncertainty and computational complexity of the clustering center in local image features retrieval based on the bag-of-words (BOW) model. A method to retrieve the measure of similarity degree from different kinds of distance and another method that requires using the matching point number as the basis of retrieval are proposed in this paper. In this method, the SURF feature is first modified to effectively reduce feature extraction complexity, and then FLANN (fast library for approximate nearest neighbors) bidirectional matching and KD-Tree + BBF matching are implemented for FAST + SURF features. Feature robustness is verified under different transformation conditions. Finally, all kinds of Uyghur document images that have been classified and sorted based on these two retrieval methods are retrieved. The results of the retrieval experiments indicate that the similarity degree measure retrieval based on distance is inferior to the retrieval based on matching number, and both of these two retrieval strategies can meet the requirements of fast and accurate searching.

Keywords: complex document image; Uyghur document image; document image segmentation; feature extraction; SURF feature; FALNN bidirectional matching; KD-Tree+BBF matching; image retrieval

收稿日期: 2017-09-17. 网络出版日期: 2018-04-24.

基金项目: 国家自然科学基金项目 (61563052, 61163028, 61363064); 新疆大学博士科研启动基金项目 (BS150262), 新疆维吾尔自治区高校科研计划创新团队项目 (XJEDU2017T002).

通信作者: 库尔班·吾布力. E-mail: urbanu@xju.edu.cn.

在当今信息技术高速发展的背景下, 多媒体技术的发展使文档图像在信息的交换中运用越来越频繁。日益增长的需求使文档图像的数量越来越庞大, 这就要求文档图像存储系统能够为用户

提供快速高效的检索服务,为了达成这一目标许多国内外学者进行了卓有成效的研究^[1-2]。王澎等^[3]提出提取图像关键点的 64 维 SURF 特征集,计算每一维上的一阶中心矩与加权绝对中心矩,形成特征向量,将其作为 SVM 分类的依据,对 1 000 幅图片进行分类,获得 86.8% 的正确分类率。赵璐璐等^[4-5],对提取的 64 维 SURF 特征点,基于 FLANN 算法进行双向匹配,对匹配对进行 PROSAC 分析,剔除误匹配对,提高图像匹配精度,有效缩短匹配占用时间。闫丽等^[6],扩展 Haar 响应,生成包含线特征、中心环绕特征、角点特征的 Haar-Like 特征集,提高描述子的区分率,基于欧氏距离比实现遥感图像的快速准确匹配。陈剑虹等^[7]提出改进 SURF 关键点检测,提取图像细节区域的特征点,不经过非极大值抑制有效去除边缘点与低对比点,将优先队列 (BBF) 融入到 KD-Tree 双向匹配中,实现稳定、快速鲁邦特征的精确匹配。罗楠等^[8]对 SURF 特征描述进行改进,用改进 DAISY 描述算子为每个关键点分配主方向,形成维数为 200 的特征向量,并通过最近邻距离比 (NNDR) 匹配目标图像,其最大匹配率达到 95.78%。

王亚文等^[9]基于中心-边缘区域比较模板匹配算法有效跟踪目标区域,提取 SURF 特征,将匹配对数目与模板匹配数目比较自适应调整跟踪窗口,实现目标检测。这些方法分别对特征提取和匹配方式进行不同程度的改进,但是检索时间或检索准确率需要进一步提高。

本文提取改进的 $N \times 64$ 维 SURF 特征,并对其实现 FLANN 双向匹配,依次统计匹配性能参数,并将其作为相似性度量依据,从大规模图像数据库中有效检索输出目标文档图像。此外,为降低检索复杂度,笔者运用基于距离的相似性度量方法,快速查找目标文档图像。

1 FAST+SURF 特征提取

本文对细节信息丰富的维吾尔文复杂文档图像进行研究,不对其进行版面分析,对改进 SURF 特征实现脱离于词袋模型的检索方法,从大规模图像数据库中实现有效检索。本文算法流程如图 1 所示。

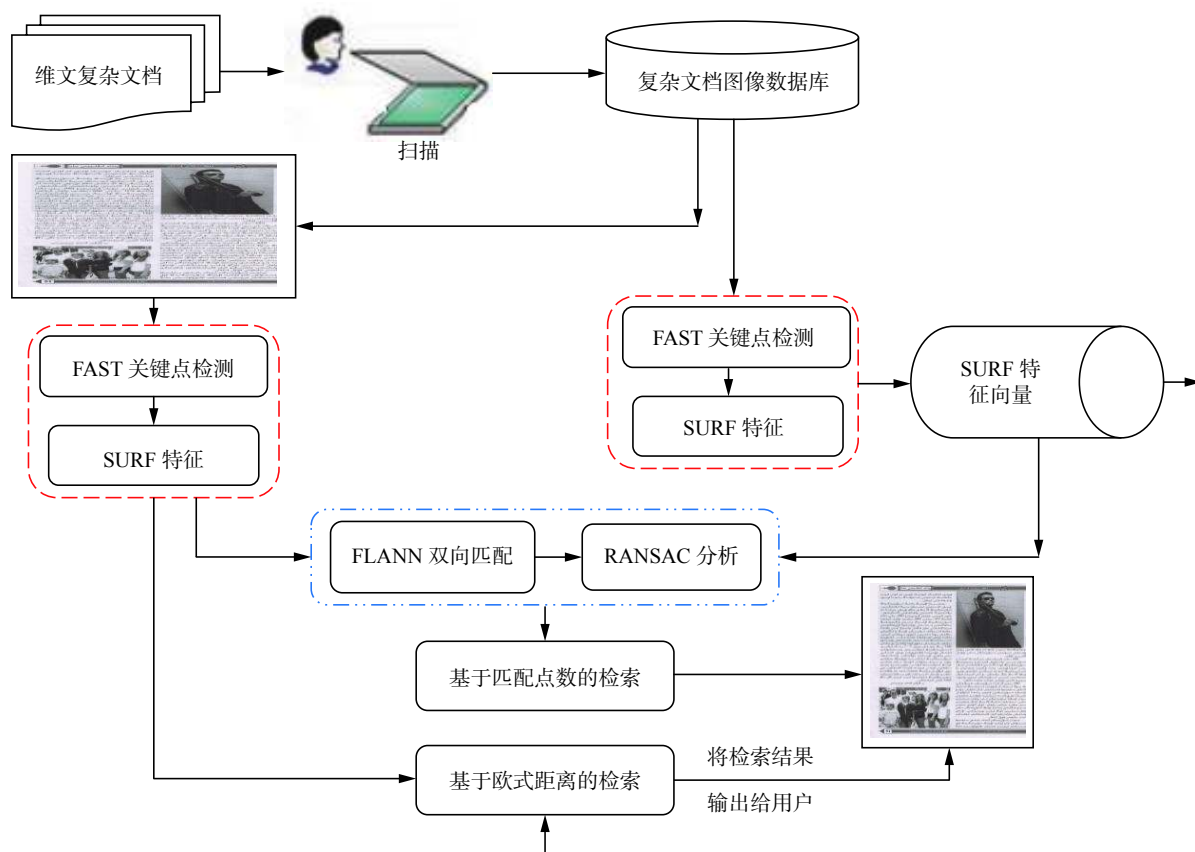


图 1 基于改进 SURF 特征的维吾尔文印刷体复杂文档图像检索流程框图

Fig. 1 The block diagram of Uyghur printed complex document image retrieval based on modified SURF feature

图 1 中的快速鲁棒特征 (SURF) 提取过程与 SIFT 相似, 由关键点检测与特征描述两部分构成。但其维持图像尺寸不变, 以倍数关系改变盒子滤波器的尺寸、尺度, 在积分图像的基础上滤波构建尺度空间, 使得特征检测耗用时间远短于 SIFT。在特征描述中, 统计扇形区域内的 Haar 小波响应值, 确定关键点的主方向, 累加统计划分区域内 X 、 Y 方向的 Haar 值与 Haar 绝对值, 降低计算复杂度。

1.1 原始 SURF 特征向量提取

计算灰度图像在 (x, y) 处的积分值, 并与不同尺度下的高斯函数二阶偏导数做卷积运算。为降低计算复杂度, 用不同尺寸、尺度的盒子滤波器近似替代卷积结果, 形成不同尺寸条件下的 Hessian 矩阵行列式。最后, 计算 Hessian 行列式的迹, 并与相邻尺度下的 26 个像素点进行非极大值抑制 (NMS) 运算, 获取 SURF 关键点的精确位置^[10-11], 其数学表达式如下:

$$I_{\Sigma(x,y)} = \sum_{i=0}^{i \ll X} \sum_{j=0}^{j \ll Y} \text{Gray}(i, j) \quad (1)$$

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{yx}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} = \begin{bmatrix} D_{xx}(x, \sigma) & D_{xy}(x, \sigma) \\ D_{yx}(x, \sigma) & D_{yy}(x, \sigma) \end{bmatrix} \quad (2)$$

式中: $I_{\Sigma(x,y)}$ 为积分图像, $H(x, \sigma)$ 为 Hessian 矩阵行列式, $L_{xx}(x, \sigma)$ 、 $L_{yy}(x, \sigma)$ 分别为 X 方向、 Y 方向二阶偏导数卷积, $L_{xy}(x, \sigma)$ 、 $L_{yx}(x, \sigma)$ 为混合偏导数卷积。 $D_{xx}(x, \sigma)$ 为盒子滤波近似值。对应 SIFT, 构建 4 组尺度空间, 每组含有不同尺寸、尺度下的盒子滤波。其层次表达如图 2 所示。

当盒子滤波器尺寸为 $N \times N$ 时, 其对应的高斯尺度为 $s = \sigma_0 \times N/9$, 其中 $\sigma_0 = 1.2$ 。尺度空间中取每一

点的 Hessian 迹^[12], 其表达式为

$$\text{Det}(H) = D_{xx}D_{yy} - (0.9D_{xy})^2 \quad (3)$$

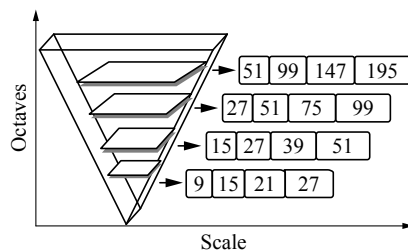


图 2 SURF 尺度空间示意

Fig. 2 Schematic diagram of SURF scale space

在尺度空间被检测出的关键点对文档图像的尺寸、平移和旋转有鲁棒性。给 SURF 局部特征点分配主方向来保证其旋转不变。以 5° 为量级, 在 60° 滑动扇形区域内累加求和每一个像素点在水平、垂直方向上的 Haar 小波响应分量, 求其最大值作为此关键点主方向^[13]。以关键点为中心取 20×20 区域, 将这个区域又按 4×4 的大小划分成 25 个子区域, 然后依次计算每个小区域的像素点的水平方向和垂直方向的 Haar 值以及绝对和值, 最后将提取 64 维 SURF 特征向量。

1.2 改进 SURF 特征提取

原始 SURF 特征相比于 SIFT, 被检测特征点数目少、特征提取占用时间短。但是, 对于图文混排的复杂文档图像, 缩短时间参量不理想。因此, 为实现维文复杂版面图像中关键点的快速检测, 本文提取文档图像的灰度信息和角点。检测角点时用 FAST 算法, 并用 SURF 算子描述, 构成 64 维 FAST+SURF 特征, 有效缩短特征提取时间^[14]。其特征检测算法流程如图 3 所示。

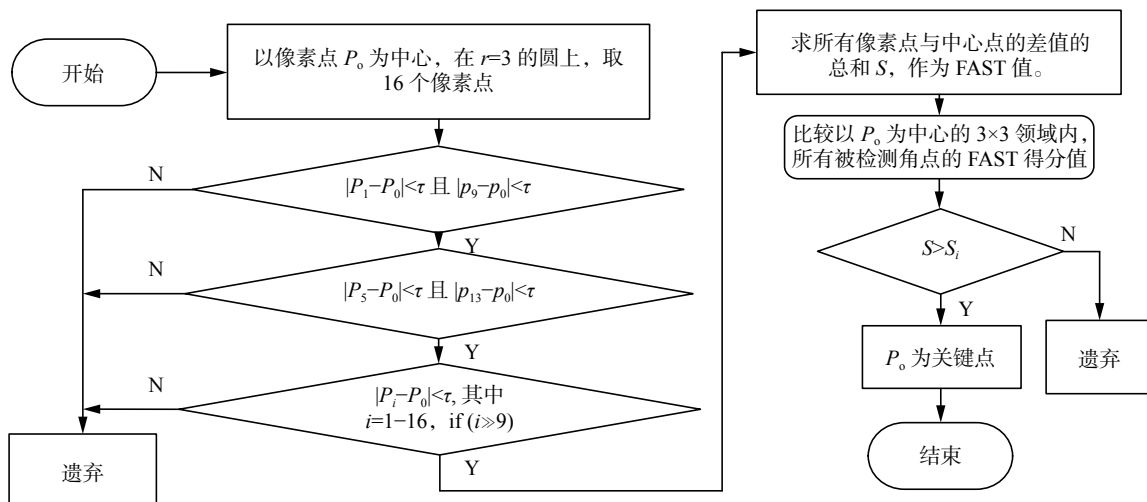


图 3 改进 SURF 特征关键点检测流程框图

Fig. 3 Flow chart of improved SURF feature key point detection.

获取具有平移、缩放不变性的关键点后, 为维持旋转不变性, 给被检测的 FAST 关键点进行 SURF 描述, 统计分子区域内 Haar 小波响应的水平、垂直分量累加值与绝对累加值, 形成 $N \times 64$ 维特征。

2 FAST+SURF 特征匹配分析

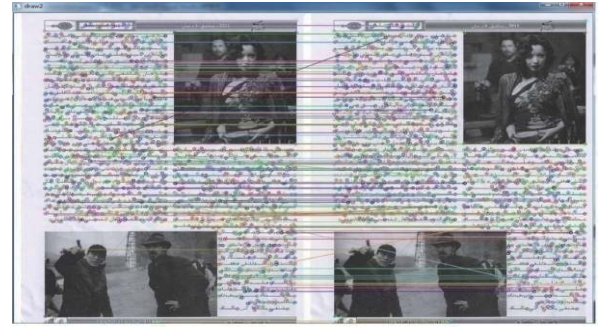
如果采用传统的方法进行特征匹配不仅计算量大, 耗时长, 而且系统的内存占用率较高。维吾尔文复杂文档图像的文字区域中存在较多的黑像素信息, 被检测关键点数目较多。因此, 为有效提高匹配速率, 笔者对不同版面图像实现双向快速近似最近邻 (FLANN) 匹配, 将其结果与 KD-Tree+BBF 匹配结果进行对比分析, 以系统性能作为基本出发点来构建匹配系统, 从而使系统能够对复杂的维吾尔文文档图像进行高效的检索。

2.1 双向 FLANN 匹配

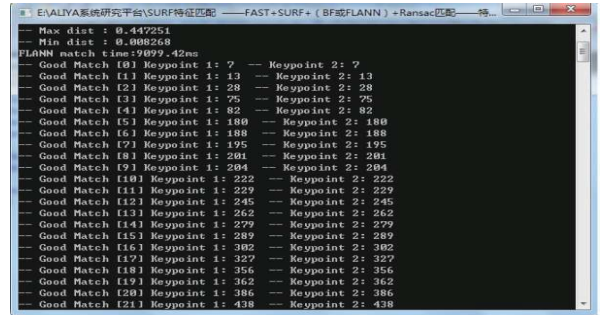
设用户输入查询图像与维吾尔文复杂文档图像库中, 待匹配图像的改进 SURF 特征向量集分别为 $A = [x_1 \ x_2 \ \cdots \ x_N]^T$, $B = [y_1 \ y_2 \ \cdots \ y_N]^T$ 。通过计算 x_i 到特征向量 B 的最大和最小距离, 并设定阈值为 $\gamma = \rho \times \text{Min}_i$ 。如果计算所得距离小于设定阈值, 则记录该点并等待进一步匹配计算, 如果距离大于设定阈值, 则继续进行其他点的匹配。由于 FLANN 算法具有匹配速度快的特点, 所以本文采用 FLANN 算法进行匹配, 并在先后两个方向同时开始匹配, 这样能够快速获得所需匹配对的相似性信息, 从而进行进一步的判断。具体为: 先从 A 到 B 的方向开始匹配, 然后再从 B 到 A 的方向进行匹配, 分别形成位置信息集合 $\{m_j^A, m_k^B\}$ 与 $\{m_k^B, m_j^A\}$ 。这样逐次将 m_j^A 的大小与 m_k^B 进行比较, 如果大小相同则说明 m_k^B 完全匹配。同时在匹配过程中通过 RANSAC 算法求出将要进行匹配的点与影射矩形的距离, 并用此距离与设定的距离阈值进行比较, 这样做的目的是为了能够比较有效地去掉误匹配点对, 从而使个匹配点对能够准确地匹配。原始图像 FALNN 双向匹配结果如图 4 所示。

2.2 KD-Tree+BBF 匹配

KD-Tree 的匹配主要由树形结构的建立与最近邻查找等两个部分组成。KD-Tree 搜索能力与特征向量的维数相关, 维数越大, 搜索能力越差。因此, 本文从改进的 KD-Tree 出发, 将得到的距离结果与预设的阈值相比较, 判断是否为匹配关键点^[15]。本文中对改进 64 维 SURF 特征实现改进 KD-Tree 匹配的过程如图 5 所示。



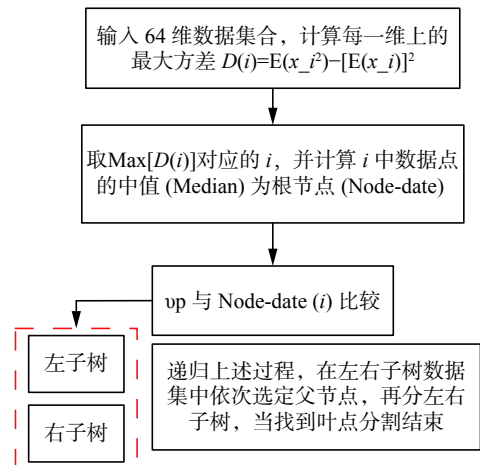
(a) 改进的 SURF 特征匹配效果图



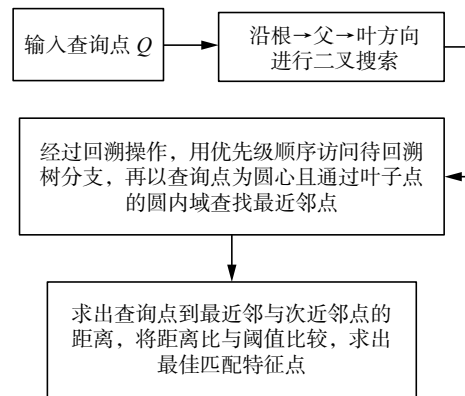
(b) 匹配点关系图

图 4 改进 SURF 特征双向 FLANN 匹配示意

Fig. 4 Schematic diagram of improved SURF features bidirectional FLANN matching



(a) KD-Tree 建立流程^[16]



(b) 特征匹配过程

图 5 改进 KD-Tree 匹配过程描述

Fig. 5 The description of improved KD-Tree matching

匹配系统在不同变换条件下的匹配效率一般是由匹配率 (MR)、正确匹配率 (CMR) 与错误匹配率 (IMR) 来衡量, 它们计算公式分别为

$$MR = \frac{N_c}{N_d} \quad (4)$$

$$CMR = \frac{N_{ac}}{N_c} \quad (5)$$

$$IMR = \frac{N_{ai}}{N_c} \quad (6)$$

式中: N_c 和 N_d 分别为总匹配对数目和被检测的特征点数; N_{ac} 和 N_{ai} 分别指正确匹配对总数和错误匹配对总数。

3 维吾尔文复杂文档图像检索

用户输入查询复杂文档图像, 系统自动获取 64 维改进 SURF 特征向量, 与特征向量库之间基于多种相似性度量算法, 从数据库中查找目标文档图像, 返回用户界面。本文运用两种相似性度量算法实现用户特定目标文档图像的检索, 即基于距离的相似性度量与基于匹配数目的相似性度量。文中用 4 种特征向量间距离相似性度量算法, 其数学表达式为

$$d_{\text{Euclidean}} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (7)$$

$$d_{\text{Manhattan}} = |x_1 - x_2| + |y_1 - y_2| \quad (8)$$

$$d_{\text{Chebyshev}} = \max(|x_1 - x_2|, |y_1 - y_2|) \quad (9)$$

$$d_{\text{Cosine}} = \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + x_2^2} \sqrt{y_1^2 + y_2^2}} \quad (10)$$

对于同一个查询文档图像的改进 SURF 特征向量, 基于上述四种距离分别度量相似性, 实现维吾尔文复杂文档图像的检索, 以检索率为系统性能指标, 评估本系统有效性。在以匹配数目为检索依据的检索系统中, 从匹配数目和正确匹配数目考虑, 统计查询图像与图像库中每幅图像之间的正确匹配数目, 将其按降序排序, 实现文档图像的有效检索。其中越相似复杂文档图像, 则其匹配数目越大。本文计算系统检索率的表达式为

$$R = \frac{N - S}{N - 1} \quad (11)$$

式中: N 是指复杂文档图像数据总的样本数, S 是指系统检索出来的目标文档图像的位置序号。

4 实验结果与分析

4.1 实验数据

收集维吾尔文复杂版面书籍、杂志、公文, 以

不同分辨率 (100 dpi, 200 dpi, 400 dpi) 扫描形成 8 位的 .bmp 格式的维文复杂文档图像, 构造含有 1 000 幅不同分辨率、不同底色信息、不同版面结构的复杂文档图像的数据库。本系统是在 Windows7 环境下用 Visual Studio 2010+openCV-2.4.10 编程开发的。

4.2 实验结果与分析

SURF 特征在检测的时候主要是靠尺度空间层数 (Octaves)、组数 (Intervals)、斑点 (THRES) 阈值的选择。阈值 (Octaves, Intervals, Init-sample, THRES) 选定的不同, 提取的特征点数目也不相同。为验证 FAST+SURF 特征的效率, 在不同的阈值条件下, 对 1 606×2 290 的同一张图像进行实验, 统计关键点数和特征提取时间, 实验结果如表 1 给出。

表 1 在不同阈值条件下同一幅图像检测的参数统计表^[16]
Table 1 The statistical table of detecting parameters using the same image under different threshold conditions^[16]

阈值	阈值参数	关键点数	特征提取时间/s
(4, 4, 2)	0.000 4f	9 290	38.536
	0.004f	2 290	19.248
	0.04f	0	10.195
(4, 3, 2)	0.000 4f	9 290	38.262
	0.004f	2 290	19.378
	0.04f	0	10.29
(3, 3, 2)	0.000 4f	9 177	37.833
	0.004f	2 982	19.326
	0.04f	0	10.221
(3, 4, 2)	0.000 4f	9 177	37.811
	0.004f	2 982	19.275
	0.04f	0	10.134

从表 1 可以得到, 尺度空间阈值的变化对 SURF 特征点的检测尤为重要。层数与层间组数的变化由改变盒子滤波器数目来实现, 其影响表现在对积分图像的处理上, 且特征检测耗用时间上有微小差异。阈值的变化极大影响了最终检索结果。对于每一个候选点, 取其上下邻域空间中的 26 个像素点, 相互比较 Hessian 矩阵行列式值。若候选点 Hessian 行列式值比 26 个像素点行列式值都大或都小且小于阈值, 则此候选点视为关键点。本文考虑多个阈值参数对 SURF 关键点检测的影响, 并为减少时间和计算复杂度, 提出了 FAST+SURF 特征提取算法。为验证该方法的

优越性, 在 (4, 4, 2, 0.000 4f) 阈值下提取相同版面不同尺寸的维吾尔文复杂文档的 SURF 特征, 将其与

不同灰度阈值下的 FAST+SURF 特征进行性能分析, 实验结果如表 2 所示。

表 2 在同一个阈值下同一幅不同尺寸图像中检测的 FAST+SURF 关键点与时间对比表

Table 2 The comparison table of detecting FAST+SURF key points and time using different image sizes in a same threshold

图像尺寸	性能	SIFT	SURF	FAST(50)+surf	FAST(100)+surf	FAST(150)+surf	FAST(200)+surf
803×1 145	关键点数	3 276	3 537	9 767	7 195	4 960	964
	占用时间/s	30.405	15.866	0.021	0.01	0.009	0.009
1 606×2 290	关键点数	10 320	11 516	22 028	9 414	9 299	2 524
	占用时间/s	105.450	51.141	0.031	0.019	0.017	0.015
3 212×4 581	关键点数	27 820	38 115	52 967	17 764	7 839	214
	占用时间/s	250.492	162.491	0.082	0.04	0.035	0.03

经多次实验发现 FAST(100)+SURF 特征数目接近在 (4, 4, 2, 0.000 4f) 阈值下的原始 SURF 特征数目。以检测 FAST 关键点来增加被检测特征点总量为付出, 就可以很大程度上减少在特征点检测上耗费的时间复杂度, 这样不仅能够明显优化匹配效率, 还能够让快速检索维吾尔文复杂文档图像成为可能。从表 2 中能够了解到, 同一幅不同尺寸图像的 SURF 特征点数目随图像尺寸变大而增多。而本文提出的改进 FAST+SURF 特征中, 特征点数目不与图像尺寸成正比关系。FAST 角点检测中, 建立高斯卷积或 Haar 滤波的尺度空间、计算高斯函数导数与积分图像的卷积是完全没有必要的, 我们只需要检测相对周围像素点, 检测都亮或都暗区域, 这样能够明显缩短特征检测时间同时也显著降低计算复杂度。为检测本文

提取特征对旋转、尺度、光照变换的鲁棒性, 进行不相同变换的情况, 选择大小为 1 606×2 290 的维吾尔文复杂文档图像且提取改进 SURF 特征, 同时分别用 FLANN 双向匹配、KD-Tree+BBF 匹配寻找能够精确匹配对总数。基于 FLANN 的双向匹配对阈值的依赖性较强, 其定义为 $\gamma = \max\{\alpha \times \min_dist, \beta\}$ 。为精确定位阈值, 分别在 $\gamma = \beta$ 与 $\gamma = \alpha \times \min_dist$ 下匹配两幅相同版面文档图像, 并得知当 $\gamma = 0.1$ 或 $\gamma = 50 \times \min_dist$ 时匹配性能最可靠。分别在两种阈值下匹配换后的文档图像与原始图像, 统计匹配性能参数得知当阈值为 $\gamma = 0.1$ 时, 系统匹配性能最可靠。在尺寸变换下 FLANN 双向匹配在阈值 $\gamma = 0.1$ 的下匹配与 KD-Tree+BBF 匹配的结果如下表 3 所示。

表 3 维吾尔文档图像尺寸变换下 FAST(100)+SURF 特征点的两种匹配结果^[16]

Table 3 Two kinds of FAST (100)+SURF feature points matching results under Uyghur document image scale transform^[16]

变换参量性能指标	FLANN 双向匹配			KD-Tree+BBF 匹配		
	整幅	1:(1/2)	1:(1/4)	整幅	1:(1/2)	1:(1/4)
关键点总数目	9 414	4 369	2 082	9 414	4 369	2 082
匹配对总数目	1 145	454	200	9 335	4 264	2 017
正确匹配对数目	759	363	172	7 582	4 151	573
正确匹配率/%	66.29	79.96	86	81.22	97.35	28.40
错误匹配对数目	386	91	28	1 753	113	1 444
错误匹配率/%	33.71	20.04	14	18.78	2.65	71.59
内点数目	768	358	59	—	—	—

将原始的文档图像划分为相等的 2 块和 4 块, 随着图像尺寸的减小, 被检测的特征点数目也随之减少, 在匹配时, 匹配到的关键点数目也减少。在特征提取时由于每块剪切图像中文字所占比例的不同, 提取的特征点的位置也不尽相

同。图像中的非文字区域是由图片组成, 特征点的分布较为密集, 干扰点也较多, 关键特征点的选择比较困难, 而文字区域的特征点分布较为均匀, 因此被检测局部关键点数目较多。从表 3 可以得到以下结论, 使用 KD-Tree+BBF 匹配时, 它

的总的匹配对数、正确的匹配对数、错误的匹配对数比使用 FLANN 双向匹配的结果好。实验中,随着图像尺寸的减小,提取到的关键点的总数目也成倍减小,匹配时的正确匹配的数目也成倍减小,相较于 FLANN 双向匹配而言, KD-Tree+BBF 匹配时的正确匹配率要高。但在使用这两种匹配算法时, KD-Tree+BBF 的匹配算法在构建与匹配上需遍历父节点与叶点且回溯,而 FLANN

双向匹配不需要回溯,只需要将近似近邻替代最近邻,若实验图像过大时,每幅图像提取的关键点会变多,而数量庞大的关键点会增加运算的复杂度和运行时间,这也给运行系统的稳定性带来影响,因此当关键点数量庞大时 FLANN 双向匹配的效果较好。为证明 SURF 特征的旋转不变性,对原始图像进行不同角度旋转,并分别用这两种匹配算法进行匹配,统计实验结果如表 4 所示。

表 4 维吾尔文档图像旋转变换条件下 FAST(100)+SURF 特征点的两种匹配结果

Table 4 Two kinds of FAST (100)+SURF feature points matching results under Uyghur document image rotation transform

变换参量性能指标	FLANN 双向匹配					KD-Tree+BBF 匹配				
	原始	+5°	+10°	-5°	-10°	原始	+5°	+10°	-5°	-10°
关键点总数目	9 414	10 373	10 614	9 339	10 423	9 414	10 373	10 614	9 339	10 423
匹配对总数目	1 145	1 213	1 164	1 070	1 254	9 335	6 044	6 691	6 005	6 794
正确匹配对	759	791	757	752	785	7 582	1 802	2 112	2 015	1 845
正确匹配率/%	66.29	65.21	65.03	70.28	62.60	81.22	29.81	31.56	33.56	27.16
错误匹配对	386	422	407	318	469	1 753	4 242	4 579	3 990	4 949
错误匹配率/%	33.71	34.79	34.97	29.72	37.40	18.78	70.19	68.44	66.44	72.84
内点数目	768	12	13	11	11	—	—	—	—	—

逆时针或顺时针旋转印刷体维文复杂文档图像,使图像面积扩大,图像的位置也会发生变化,因此特征点的位置也会改变。由表 4 可以得出, FLANN 双向匹配中,当旋转不同的角度时,图像的整体面积不同,关键点的数也不同。因此匹配时的正确匹配率也在 62.60%~70.28% 变化。可见,匹配对总数变化率与正确/错误匹配率相关。在不进行任何旋转变换时,内点数目为最大值为 768。在 KD-Tree+BBF 匹配中,匹配对数目减少,

从 9 335~6 005,变化范围较大;对于整体结果,匹配结果相对较好的是 KD-Tree+BBF 的匹配率, FLANN 匹配效果较差,但 KD-Tree+BBF 检索工作量大,而 FLANN 检索工作量小。当旋转角度发生变化时, FLANN 双向匹配的结果较好, KD-Tree+BBF 匹配的结果较差。为验证在光照变换条件下的检索效果,对原始文档图像的亮度进行改变,统计不同亮度变换下的两种匹配性能参数,实验结果如表 5 给出。

表 5 维吾尔文档图像光照变换下的 FAST(100)+SURF 特征点的两种匹配结果

Table 5 Two types of FAST (100)+SURF feature points matching results with Uyghur document image illumination transform

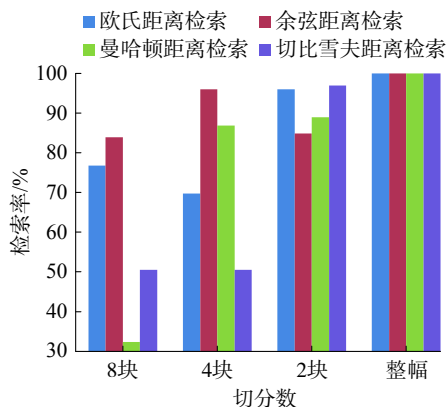
变换参量性能指标	FLANN 双向匹配							KD-Tree+BBF 匹配						
	原始	20	40	60	-20	-40	-60	原始	20	40	60	-20	-40	-60
关键点总数目	9 414	9 882	9 411	9 729	8 963	9 647	30 191	9 414	9 882	9 411	9 729	8 963	9 647	30 191
匹配对总数目	1 145	1 073	1 139	1 055	999	952	2 229	9 335	4 892	4 931	4 374	5 591	5 789	7 048
正确匹配对	759	749	758	705	803	784	1 018	7 582	3 556	1 869	1 638	3 499	3 281	2 652
正确匹配率/%	66.29	69.80	66.55	66.82	80.38	82.35	45.67	81.22	72.69	37.90	37.45	62.58	56.68	37.63
错误匹配对	386	324	381	350	196	168	1 211	1 753	1 336	3 062	2 736	2 092	2 508	4 396
错误匹配率/%	33.71	30.20	33.45	33.18	19.62	17.65	54.33	18.78	27.31	62.10	62.55	37.42	43.32	62.37
内点数目	768	321	91	7	287	216	8	—	—	—	—	—	—	—

表 5 中,在 FLANN 双向匹配中,亮度变亮的变化对匹配的正确率影响较小,在 705~758 波动;而亮度变暗的变化对匹配的正确率影响较大,在

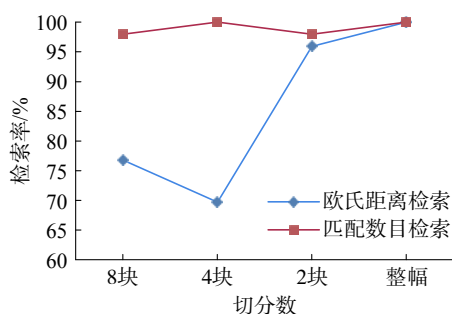
803~1 018,这是受采集的实验样本的影响,亮度过暗时会形成干扰点,因此亮度的变化范围不易过大。在 KD-Tree+BBF 匹配中,当亮度变暗到

-60 时,实验样本整体模糊,样本的背景色也被列入检测范围,对检测到的关键特征点造成干扰,从而对匹配时的匹配数目造成干扰,影响正确匹配率;亮度变暗时提取的特征点数目要多于亮度变亮时的特征点数目,匹配的精度也较差。可知,在亮度变换条件下性能较稳定的是 KD-Tree+BBF 匹配。

由于维吾尔文复杂文档图像库是由不同版面、不同底色、不同分辨率文档图像采集而建立的。在灰度化过程中,不能有效消除文档图像底色,因此底色信息影响关键点检测。此外,将图像几何均匀分割成不同小块与整幅图像匹配时,若图像版面越相似,则匹配率较高;因此从数据库中取不具有底色,相似版面结构的分辨率为 100 dpi,尺寸为 1 606×2 290 的 100 幅维吾尔文复杂文档图像构建小规模图像数据库。从中任选一幅图像,将其几何均匀划分成 2、4、8 块,分别作为查询图像,再以基于匹配距离与匹配数目为依据进行图像检索。由于文档图像尺寸大,被检测特征点数目较多,因此系统检索耗用时间较长,本文不作为系统性能指标。用不同匹配方法所进行的检索实验对比如图 6 所示。



(a) 不同距离相似性度量检索率对比



(b) 基于欧氏距离与匹配数目检索的检错率对比

图 6 改进 SURF 特征实现维吾尔文复杂文档图像检索的性能指标对比

Fig. 6 The performance comparison figure of modified SURF feature based Uyghur complex document image retrieval

将复杂的文档图像分别均匀划分为等面积的 2 块、4 块和 8 块,由于每幅图像的文字分布区域不同,所以分割后的每小块中包含的关键点数目也不同。在本实验中,当被选择查询图像分割成 4 块时,绝大部分为非文字区域,因此提取的特征数目少,并在完整图像库中检索含有本区域的图像时,检索率明显下降。从图 6(a) 可知,不同距离相似性度量算法对检索系统的影响不同。因此在剪切图像研究中,基于匹配距离的检索系统相对性能优于基于匹配数目的检索。当完整文档图像为查询输入时,与目标图像关键点位置一一对应,因此关键点之间距离为最小;当复杂版面图像均匀分割分别作为查询图像时,被检测关键点数目减少,被分割图文面积不均匀,在诸多相似版面中难以找出含有被分割区域的复杂文档图像。从图 6(b) 可知,基于匹配数目的检索性能优于基于距离的检索。由于局部特征的尺寸变换不变性使得对于区域匹配数目相对稳定,因此不同分割条件下检索目标文档图像的检索率都接近 100%。

在实验中,采集的原始图像篇幅较大,特征提取到的特征点数目太多,这对最终匹配点的数目造成很大的影响。因此,为更好地评估系统性能,对采集到的原始维吾尔文复杂文档图像数据库进行修改:1) 对图像进行压缩,将图像压缩成 256×256 尺寸;2) 对图像进行剪切,将图像剪切成 256×256 尺寸。对同一幅文档图像进行两种修改样本实例如图 7 所示。



(a) 压缩图像样本实例

(b) 剪切的图像样本实例

图 7 修改数据库样本实例

Fig. 7 The sample instance of modified database

从图 7 中可以得到,对原始的整幅图像进行压缩,会使图像的清晰度降低。而对原始的整幅图像进行剪切,只是剪切图像的一部分,不会对清晰度造成影响,但篇幅的减少也会使提取的特

征点减少。因此改造后的数据库都会对检索精度有或多或少影响。本文分别从匹配数目、匹配距离这两个方面进行测试,其检索实验结果如表 8、9 所示。

表 8 对剪切文档图像库进行检索实验的统计结果

Table 8 The statistical results of the sheared Uyghur document image retrieval experiment

检索性能指标	检索率/%	总检索时间/s	平均索引时间/s
匹配数目检索	100	1 000	1
欧式距离检索	100	854	0.854
余弦距离检索	100	861	0.861

表 9 对压缩文档图像库进行检索实验的统计结果

Table 9 The statistical results of the compressed Uyghur document image retrieval experiment

检索性能指标	检索率/%	总检索时间/s	平均索引时间/s
匹配数目检索	100	1 636	1.636
欧式距离检索	100	599	0.599
余弦距离检索	100	607	0.607

由表 8 可以得出,由于剪切文档图像是从原始文档图像上分割的一部分,因此检测到的特征点数目也会减少。当输入裁剪图像,在剪切构造的维文复杂文档数据库中,基于 3 种方法检索,其检索率都达到 100%,但检索占用时间不同。匹配数目检索时需要先比较最近邻的特征点之间距离,看距离比值是否在设定的阈值范围内,若是,则相匹配,反之不匹配,这就导致检索时耗时较长。

由于压缩文档只是对原始文档图像的缩小,因此其内容包含整体图像内容。由表 9 可知,在时间损耗上,基于匹配数目的检索系统较多,而基于距离相似性度量的检索时间较少。对于剪切文档和压缩文档这两种数据库,基于匹配数目的检索中提取的特征点数目越多,匹配时的匹配点数目也会增多,则匹配时间也会随之变化。基于距离相似性的检索中,剪切的文档图像篇幅较少,比压缩的文档图像检索用时更短。由于人工剪切采集图像,易受人主观因素的影响;此外原始图像库是由不同分辨率图像构成,采集时分辨率越大则图像获取内容越少,使得图像失去完整性,因此压缩图像检索时间比剪切图像检索时间短。

5 结束语

本文为弥补维吾尔文复杂文档图像在检索领域中的空白,在维吾尔文文档图像检索匹配中运用 SURF 与改进 SURF 特征,使得少数民族文档

图像检索研究进一步推进。文中通过不同阈值条件下检测 SURF 特征,分析多种阈值对其影响。因此,为提高特征阈值适应性及快速检测特征点,将 FAST 角点检测与 SURF 描述相结合,使特征检测时间压缩到 50~1 425 倍。文中笔者对选取的 64 维特征向量运用两种匹配,并在尺寸、旋转、光照变换条件下实验统计匹配率,对两种匹配性能进行对比分析。最后,分别对原始 100 幅文档图像、压缩 1 000 幅图像、剪切 1 000 幅文档图像,基于多种距离度量特征向量间的相似性,检索目标文档图像,将其检索率与基于匹配数目的检索率对比,亮出以匹配数目为检索依据的搜索系统优越性。系统最高检索率都达到 100%。

但是,由于原始文档图像篇幅较大,提取的特征点数目较多,特征点之间的匹配点数目也会增大,因此基于匹配数目的检索系统耗用时间比基于距离的检索系统较长,系统检索所占时间都不太理想。在保证系统较高的检索率的前提下,怎样进一步降低时间开销是下一步研究重点。

参考文献:

- [1] 张敬丽, 张会清, 代汝勇. 基于 MIC-SURF 的快速图像匹配算法[J]. 计算机工程, 2016, 42(1): 210-214.
ZHANG Jingli, ZHANG Huiqing, DAI Ruyong. Fast image matching algorithm based on MIC-SURF[J]. Computer engineering, 2016, 42(1): 210-214.
- [2] ALFANINDYA A, HASHIM N, ESWARAN C. Content based image retrieval and classification using speeded-up robust features (SURF) and grouped bag-of-visual-words (GBoVW)[C]//Proceedings of 2013 International Conference on Technology, Informatics, Management, Engineering, and Environment. Bandung, Indonesia, 2013: 77-82.
- [3] 王澍, 吕学强, 张凯, 等. 基于快速鲁棒特征集合统计特征的图像分类方法[J]. 计算机应用, 2015, 35(1): 224-230.
WANG Shu, LYU Xueqiang, ZHANG Kai, et al. Image classification approach based on statistical features of speed up robust feature set[J]. Journal of computer applications, 2015, 35(1): 224-230.
- [4] 赵璐璐, 耿国华, 李康, 等. 基于 SURF 和快速近似最近邻搜索的图像匹配算法[J]. 计算机应用研究, 2013, 30(3): 921-923.
ZHAO Lulu, GENG Guohua, LI Kang, et al. Images matching algorithm based on SURF and fast approximate nearest neighbor search[J]. Application research of computers, 2013, 30(3): 921-923.
- [5] CHEON S H, EOM I K, HA S W, et al. An enhanced SURF algorithm based on new interest point detection pro-

- cedure and fast computation technique[J]. *Journal of real-time image processing*, 2016.
- [6] 闫利, 陈林. 一种改进的 SURF 及其在遥感影像匹配中的应用[J]. *武汉大学学报(信息科学版)*, 2013, 38(7): 770–773, 804.
- YAN Li, CHEN Lin. A modified SURF descriptor and its application in remote sensing images matching[J]. *Geomatics and information science of Wuhan university*, 2013, 38(7): 770–773, 804.
- [7] 陈剑虹, 韩小珍. 结合 FAST-SURF 和改进 k-d 树最近邻查找的图像配准[J]. *西安理工大学学报*, 2016, 32(2): 213–217, 252.
- CHEN Jianhong, HAN Xiaozhen. Image matching algorithm combining FAST-SURF and improved k-d tree nearest neighbor search[J]. *Journal of Xi'an university of technology*, 2016, 32(2): 213–217, 252.
- [8] 罗楠, 孙权森, 陈强, 等. 结合 SURF 特征点与 DAISY 描述符的图像匹配算法[J]. *计算机科学*, 2014, 41(11): 286–290, 300.
- LUO Nan, SUN Quansen, CHEN Qiang, et al. Image matching algorithm combining SURF feature point and DAISY descriptor[J]. *Computer science*, 2014, 41(11): 286–290, 300.
- [9] 王亚文, 陈鸿昶, 李邵梅, 等. 基于关键点匹配的多策略尺度自适应跟踪算法[J]. *计算机工程与设计*, 2016, 37(1): 247–253.
- WANG Yawen, CHEN Hongchang, LI Shaomei, et al. Multi-strategy scale adaptive tracking algorithm via keypoint matching[J]. *Computer engineering and design*, 2016, 37(1): 247–253.
- [10] 张凤晶, 王志强, 吴迪, 等. 基于 SURF 的图像配准改进算法[J]. *长春理工大学学报(自然科学版)*, 2016, 39(1): 112–115.
- ZHANG Fengjing, WANG Zhiqiang, WU Di, et al. Improved algorithm of image registration based on SURF[J]. *Journal of Changchun university of science and technology (natural science edition)*, 2016, 39(1): 112–115.
- [11] LIU Yanling, MA Sihang. Research on image based on improved SURF feature matching[C]//*Proceedings of 7th International Symposium on Computational Intelligence and Design*. Hangzhou, China, 2014: 581–584.
- [12] EL-GAYAR M M, SOLIMAN H, MEKY N. A comparative study of image low level feature extraction algorithms[J]. *Egyptian informatics journal*, 2013, 14(2): 175–181.
- [13] HUANG Liqin, CHEN Caigan, SHEN Henghua, et al. Adaptive registration algorithm of color images based on SURF[J]. *Measurement*, 2015, 66: 118–124.
- [14] 安维胜, 余让明, 伍玉铃. 基于 FAST 和 SURF 的图像配准算法[J]. *计算机工程*, 2015, 41(10): 232–235, 239.
- AN Weisheng, YU Rangming, WU Yuling. Image registration algorithm based on FAST and SURF[J]. *Computer engineering*, 2015, 41(10): 232–235, 239.
- [15] HUI Dong, YUAN Handian. Research of image matching algorithm based on SURF features[C]//*Proceedings of 2012 International Conference on Computer Science and Information Processing (CSIP)*. Xi'an, China, 2012: 1140–1143.
- [16] 阿丽亚·巴吐尔. 基于局部特征的维吾尔文印刷体复杂文档图像检索研究[D]. 乌鲁木齐: 新疆大学, 2017.
- ALIYA Batur. Research on Uyghur printed complex document image retrieval based on local feature[D]. Urumchi: Xinjiang University, 2017.

作者简介:



阿丽亚·巴吐尔, 女, 1990 年生, 硕士研究生, 主要研究方向为模式识别、图像检索。



努尔毕亚·亚地卡尔, 女, 1970 年生, 讲师, 中国人工智能学会会员, 主要研究方向为图像处理、计算机视觉。



吾尔尼沙·买买提, 女, 1976 年生, 讲师, 中国人工智能学会会员, 主要研究方向为信号与信息处理、模式识别。