

采用快速 SIFT 算法实现目标识别^①

张书真¹ 宋海龙² 向晓燕¹ 赵英男³ (1.吉首大学 物理科学与信息工程学院 湖南 吉首 416000; 2.吉首大学 信息管理与工程学院 湖南 张家界 427000; 3.南京信息工程大学 计算机与软件学院 江苏 南京 210044)

摘要: 在基于原始 SIFT 算法的目标识别中,特征描述符的计算复杂,特征点的匹配时间较长,为此提出一种快速 SIFT 算法。该算法采用同心圆形窗口内的灰度累加值和差分值构建 16 维的简化描述符,并且在目标识别时,按照金字塔结构由粗至精进行特征点匹配。实验表明,在保证目标识别准确率的前提下,快速算法的运算时间比原始 SIFT 算法减少了两个数量级,具有很好的实时性能。

关键词: 目标识别; SIFT 算法; 特征描述符; 尺度空间; 特征点匹配

Fast SIFT Algorithm for Object Recognition

ZHANG Shu-Zhen¹, SONG Hai-Long², XIANG Xiao-Yan¹, ZHAO Ying-Nan³

(1.College of Physics Science and Information Engineering, Jishou University, Jishou 416000, China; 2.College of Information Management and Engineering, Jishou University, Zhangjiajie 427000, China; 3. College of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: In the object recognition based on original SIFT algorithm, the calculation of the feature descriptor is complex and the time of feature point matching is long, so a fast SIFT algorithm is proposed. The fast algorithm takes cumulative gray-scale value and its difference in a concentric circular window to construct a simplified feature descriptor with 16 dimension, and it matches feature points according to a coarse-to-fine pyramid structure when recognising the object. Experiments show that, under the premise of ensuring the recognition accuracy, the computation time of the proposed fast algorithm is reduced two orders of magnitude compared with the original SIFT algorithm, so the proposed algorithm has a good real-time performance.

Keywords: object recognition; SIFT algorithm; feature description; scale space; feature point matching

自动目标识别(AUTOMATIC TARGET RECOGNITION, ART)是模式识别技术的一项重要应用。传统的目标识别技术多是基于目标的全局特征(如灰度分布统计特征、几何形状、矩特征等),这些特征在目标被部分遮挡或发生形变时不再适用,而基于局部特征的目标识别能有效克服上述不足。2004年,Lowé提出一种提取局部特征的SIFT(Scale Invariant Feature Transform)算法^[1],该算法将尺度空间的极

值点作为特征点,并通过计算特征点局部邻域内的梯度方向直方图,来构建特征描述符。研究表明,SIFT算法在目标发生平移、旋转、缩放、亮度变化、部分遮挡、视角变换时都能取得较好的识别效果。然而算法中每个特征点使用了128维的向量进行描述,数据计算量相当大,并且由于特征点很多,使得特征匹配时间过长。针对上述不足,本文将在原始SIFT算法的

^① 基金项目:国家自然科学基金(60573028);湖南省教育厅自然科学基金(07C526)

收稿时间:2009-09-18;收到修改稿时间:2009-10-26

基础上对每个特征点的描述符进行简化,并由尺度空间的结构出发,提出一种由粗至精的快速匹配识别方案,从而有效降低算法运算时间,提高计算机处理的实时性。

1 SIFT算法

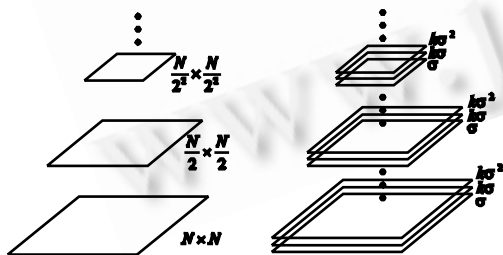
1.1 图像的尺度空间

传统的图像多尺度表达,是通过平滑滤波与亚采样来构建多级金字塔。Lindeberg 将尺度空间表示法用于图像多尺度表达中^[2],它通过在每一级金字塔图像中使用连续平滑滤波,形成多级多层金字塔结构。同时 Koenderink 和 Lindeberg 的研究证明高斯核是实现尺度变换的唯一线性变换核^[3]。一幅图像 $I(x,y)$ 在不同尺度下的尺度空间 $L(x,y,\sigma)$ 可通过图像与高斯核 $G(x,y,\sigma)$ 的卷积来实现,即:

$$L(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \quad (1)$$

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} \exp[-(x^2+y^2)/2\sigma^2] \quad (2)$$

其中 σ 为尺度空间因子,它的值越小表征图像被平滑得越少,相应的尺度也越小。大尺度对应图像的概貌特征,小尺度对应图像的细节特征。图 1 给出了传统高斯金字塔与多级多层高斯金字塔的示意图,如图所示,传统金字塔中第 n 层高斯图像是由第 $n-1$ 层的高斯图像经过高斯平滑和亚采样获得,而采用尺度空间构建的金字塔中,每一级有多层,每层之间的 σ 值相差 k 倍。



(a) 传统高斯金字塔 (b) 多级多层高斯金字塔

图 1 传统金字塔与多级多层金字塔示意图

1.2 SIFT 特征匹配算法

Lowe 提出的 SIFT 算法主要有四个步骤:

1) 尺度空间极点检测。首先将输入图像同高斯核进行卷积,对相邻尺度空间做差分建立 DOG 金字塔,

在 DOG 空间检测极点。极点的检测就是看该点是否均大于或小于它的 26 个邻域点(包括该点相同尺度的 8 个邻域点,以及大于该尺度的 9 个邻域点和小于该尺度的 9 个邻域点)。

$$D(x,y,\sigma) = [G(x,y,k\sigma) - G(x,y,\sigma)] * I(x,y) \\ = L(x,y,k\sigma) - L(x,y,\sigma) \quad (3)$$

在获得的极点后,还必须剔除掉低对比度的极点和不稳定的边缘响应点,以此获得匹配能力和抗噪声能力较强的特征点。精确定位特征点的具体过程可参考文献^[1]。

2) 特征点主方向的确定。为使特征描述符具有旋转不变性,采用梯度直方图来确定特征点的主方向。坐标为特征点的梯度模值与方向为

$$m(x,y) = \sqrt{[L(x+1,y) - L(x-1,y)]^2 + [L(x,y+1) - L(x,y-1)]^2} \quad (4)$$

$$\theta(x,y) = \arctan \left[\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)} \right] \quad (5)$$

对于每个特征点,在特征点为中心的邻域窗口内采样,用直方图统计邻域像素的梯度方向。梯度直方图的范围是 00-3600,每 100 为一柱,共 36 柱,梯度直方图的峰值对应角度代表特征点的主方向。

3) 生成特征描述符。首先将坐标轴旋转 to 特征点主方向,以确保旋转不变性。然后以特征点为中心,选取 16×16 大小的邻域,将此邻域均匀分成 16 个 4×4 的子区域。在每个子区域计算 8 个方向(0°、45°、90°、135°、180°、225°、270°、315°、360°)的梯度累加和,得到梯度直方图,这样对于每个特征点将生成 $16 \times 8 = 128$ 维的特征描述符。

4) 特征匹配。两幅图像的 SIFT 特征描述符都生成后,就利用欧氏距离度量两幅图像特征点的相似性,即当图像 1 中某个特征点 m_1 与图像 2 中欧氏距离最近特征点 m_2 和次近特征点 m_3 ,其距离比值小于某个阈值时,就认为 m_1 与 m_2 相匹配。匹配后还需用 RANSAC 法剔除误匹配对。

2 改进的SIFT算法

实验表明,算法中建立尺度空间和生成特征描述符耗费了大量的时间,同时随着图像尺寸的增大,特征点数量增多,特征匹配时间明显增加,为此本文对原始的 SIFT 算法进行了改进。

2.1 特征描述符的简化

将坐标轴旋转至特征点主方向后，简化特征描述符的具体生成步骤如下：

1) 以特征点为圆心，以 8pixel 为半径作圆形窗口，确定出特征点的邻域范围^[4]，然后分别以 1pixel、2pixel、3pixel、4pixel、5pixel、6pixel、7pixel 为半径，作以特征点为圆心的同心圆，将邻域区域划分为 8 个子区域。

2) 统计各子区域内像素的灰度累加值 f_i ，为保证光照不变性，将灰度累加值归一化，得到

$$\bar{f}_i = \frac{f_i}{\sqrt{\sum_{j=1}^8 f_j}} \quad i=1, 2, \dots, 8 \quad (6)$$

3) 计算各子区域的灰度累加差分 $d\bar{f}_i$ ，利用一阶数值差分可得

$$d\bar{f}_i = \begin{cases} |\bar{f}_i - \bar{f}_{i+1}| & i=1 \\ |2\bar{f}_i - \bar{f}_{i-1} - \bar{f}_{i+1}| & 1 < i < 8 \\ |\bar{f}_i - \bar{f}_{i-1}| & i=8 \end{cases} \quad (7)$$

4) 特征点的描述符由归一化灰度累加值和灰度累加差分构成，即得到一个 16 维的描述符

$$F = (\bar{f}_1, \bar{f}_2, \dots, \bar{f}_8, d\bar{f}_1, d\bar{f}_2, \dots, d\bar{f}_8) \quad (8)$$

通过上述过程建立的特征描述符考虑了邻域内多个子区域的灰度信息与关联，具有平移、尺度、旋转等基本特性。

2.2 匹配算法的改进

在两幅图像间进行基于特征点的目标识别，关键问题是确定特征点匹配对的数目。假设图像间的变换为全仿射变换

$$\begin{bmatrix} x_{21} & x_{22} & \dots & x_{2q} \\ y_{21} & y_{22} & \dots & y_{2q} \\ 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} a_{00} & a_{01} & t_x \\ a_{10} & a_{11} & t_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1q} \\ y_{11} & y_{12} & \dots & y_{1q} \\ 1 & 1 & \dots & 1 \end{bmatrix} \quad (9)$$

其中 $(x_{11}, y_{11}), (x_{12}, y_{12}), \dots, (x_{1q}, y_{1q}) (q \geq 3)$ 为参考图像部分特征点的坐标， $(x_{21}, y_{21}), (x_{22}, y_{22}), \dots, (x_{2q}, y_{2q})$ 为待检测图像部分特征点的坐标。由(9)式可知，变换有 6 个自由度，因此在匹配算法中，若检测到 3 组不共线的精确匹配对，即可识别目标。在实际操作中，为提高识别准确性，可适当增加匹配对个数。

鉴于上述分析，又根据尺度空间结构分析，在目标识别中可考虑匹配过程从金字塔顶级开始，由粗至

精进行匹配，当目标一旦被认定，即可停止匹配过程，其具体步骤如下：

1) 用传统高斯金字塔结构表达待检测图像；

2) 对金字塔顶级图像，进行尺度空间细化表达，并在该尺度内检测 DOG 图像中的极点，精确定位特征点，获得特征主方向及简化特征描述子；

3) 将获得的特征点与参考目标图像的所有特征点进行匹配，去除误匹配对后，若获得的匹配对个数达到给定数目时，即完成识别，若匹配对个数不够，则进入传统金字塔的下一级图像中；

4) 在新一级图像中，重复 2) 3) 步骤，直至获得足够多匹配数目，完成识别。

3 实验结果及分析

实验是在 CPU 为 P4 2.8G，内存为 768MHz 的微机利用 VC++6.0 编程实现的，通过两组实验验证本文的算法。

实验 1: 描述符算法性能的验证。选取 10 幅图像，大小均在 600×450 左右，每幅图都进行光照、尺度、旋转和遮挡变换，分别采用传统 SIFT 描述符和简化 SIFT 描述符进行特征点匹配，运算结果各数据如表 1 所示，其中匹配率^[5]是指描述符匹配成功的特征对数量与两幅图像中重复出现的特征对数量的比值，误匹配率是指错误匹配的特征点数目与描述符计算出的总匹配对数量的比值，表中的数据是先对每幅图像计算后，再统计求得平均值。实验还考查了 500 个特征点的描述符在简化前后的计算时间，结果是简化前计算 500 个特征点的描述符平均耗时 731ms，简化后的平均耗时为 64ms。

表 1 两种算法匹配率与错误匹配率(%)

图像变化	传统 SIFT 描述符		简化 SIFT 描述符	
	匹配率	误匹配率	匹配率	误匹配率
照明增加 20%	95.2	1.6	93.4	1.7
图像放大 1.2 倍	80.6	8.4	81.3	9.5
旋转 15°	52.8	7.2	62.7	7.1
遮挡 10%	78.5	7.8	64.3	8.5

通过实验 1 可看出，简化的描述符虽然在图像被部分遮挡的情况下匹配率有所下降，但在其它变换下

匹配率变化不大, 而简化描述符的计算时间大大低于传统 SIFT 描述符计算时间。



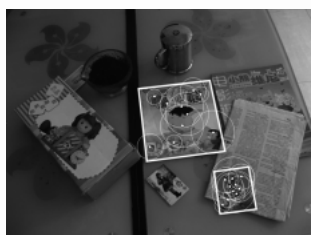
(a) 参考目标 1 (b) 参考目标 2



(c) 待检测图片



(d) 传统方法检测结果



(e) 改进匹配方案检测结果

图 2 目标识别实验

实验 2: 目标识别实验。分别采用传统 SIFT 算法和改进匹配方案的快速 SIFT 算法进行目标识别实验, 快速 SIFT 算法设定匹配对数目为 10。实验结果如图 2 所示, 其中 2(a)、2(b)为参考目标图像, 2(c)为待检测图像, 2(d)为传统 SIFT 算法识别结果, 2(e)为快速 SIFT 算法识别结果。2(d)、2(e)图像中, 用白点表示匹配出的特征点, 白色圆形表示特征点匹配时的尺度大小, 白色矩形框标示出识别得到的目标轮廓。实验

中, 2(a)、2(b)、2(c)的大小分别为 574×657 、 476×564 、 722×538 , 提取到特征点分别为 1374 点、1026 点和 1436。快速 SIFT 算法采用由粗至精的匹配方案, 当匹配对数目达到设定的阈值 10 时, 能准确识别出两个目标。传统 SIFT 算法对匹配对数目没有限制, 最终的匹配对数目 2(a)、2(b)目标分别达到了 48 对和 29 对。实际上, 由于快速 SIFT 算法在匹配成功后即结束了计算, 因此它没有传统 SIFT 算法在小尺度上获取特征点和特征描述符的计算, 也不需要相应小尺度上耗费匹配时间, 因此改进匹配方案的 SIFT 算法在实时性能上改善明显, 这一结果在表 2 中得以反映。

表 2 各种 SIFT 算法的耗时比较(ms)

参考目标	传统算法	算法 1	算法 2	算法 1+算法 2
2(a)	2518	286	735	86
2(b)	2295	214	723	74

目标识别实验还对文中提到的各种 SIFT 算法的耗时进行了比较, 结果如表 2 所示, 其中算法 1 是指仅采用简化描述符的快速 SIFT 算法, 算法 2 指仅采用改进匹配方案的快速 SIFT 算法。

通过比较, 可以看到结合简化描述符和改进匹配方案的快速 SIFT 算法, 其运算时间比传统 SIFT 算法降低了两个数量级, 这非常有利于 SIFT 算法运用到实时处理系统中。

4 结语

本文对 SIFT 特征描述符和特征点匹配方案进行了改进, 提出一种适合目标识别的快速 SIFT 算法。算法以特征点的同心圆形邻域内灰度累加值及其差分作为描述特征, 构建出 16 维的特征描述符, 该描述符对于图像的平移、旋转、缩放等变换具有不变性。在匹配方案中, 采用金字塔结构由粗至精进行特征点匹配, 使得目标识别在较大尺度下能够完成。实验表明, 在大多数变换情况下, 简化的 SIFT 描述符与传统 SIFT 描述符的匹配率相差不大, 而简化描述符的计算时间大大缩短。与此同时, 改进的匹配方案能有效减少了匹配特征点数目, 使目标识别时间大大减少。本文算法的实时性能突出, 但是在对称环境或相似环境下, 算法的特征点匹配率下降明显, 如何在描述符中加入全局元素是下一步要考虑的问题^[6]。另外,

(下转第 186 页)

(上接第 85 页)

目标识别过程中,本算法的匹配对数目设为固定阈值 10,如何能自适应地设置匹配对数目,以达到目标识别的准确性与实时性的统一,也是需要考虑的问题。

参考文献

1 Lowe DG. Distinctive image features from scale invariant key points. *International Journal of Computer Vision*, 2004,60(2):91 - 110.

- 2 Lindberg T, Bretzner L. Real-time scale selection in hybrid multi-scale representation. Griffin L, ed. *Proc. of Scale-space'03*. Berlin, Heidelberg: Springer-Verlag, 2003. 148 - 163.
- 3 Lindberg T. Scale-space theory:a basic tool for analyzing structures at different scales. *Journal of Applied Statistics*, 1994,21(2):224 - 270.
- 4 刘立,彭复员,赵坤,万亚平.采用简化 SIFT 算法实现快速图像匹配. *红外与激光工程*, 2008,37(1):181 - 184.