

SIFT 改进算法在图像配准中的应用

王田甲, 刘国荣

(西安交通大学 电子与信息工程学院, 陕西 西安 710049)

摘 要: 文中对尺度不变特征变换(SIFT)算法进行分析研究, 针对原算法中 128 维的高维描述子提出 60 维方形邻域描述子, 统计邻域梯度信息. 方形邻域描述子较原算法增加了邻域像素统计范围, 增强了关键点的邻域信息; 在配准阶段采用欧氏距离作为度量函数, 用次临近与最邻近之比来对 60 维描述子进行匹配. 通过实验证实, 改进算法的匹配时间是原算法的 30%~60%, 配准精度与原算法相近, 对于复杂图像的配准精度较原算法有所提高, 适用于对实时性要求较高的图像配准场合.

关键词: 尺度不变特征变换算法; 尺度空间; 方形邻域描述子; 特征匹配

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2011)05-0184-05

Improved SIFT Algorithm for Image Matching

WANG Tian-jia, LIU Guo-rong

(School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: The principal of SIFT algorithm is researched in this paper. Due to the descriptor that one feature point needs 128 dimensions, a 60-dimension-square descriptor based on statistic local gradient information is taken forth. Comparing with the original one, the new descriptor expands the scope of neighborhood pixels; the ratio between the first and second closest distance is used to match the 60-dimension descriptors. According to the experiment, the results show that matching time is greatly shortened by 30%~60%; and the new descriptor is competitive with SIFT descriptor in effectiveness. Further more, the new descriptor exhibits good performance in more complicated images. Therefore, it is more suitable in real-time applications.

Key words: SIFT algorithm; scale space; square local descriptor; feature matching

1 引言

图像匹配是计算机视觉和数字图像处理的重要组成部分. 数字图像配准技术在图像拼接、图像融合、图像镶嵌、目标识别与跟踪等领域中都得到了广泛的应用^[1].

图像匹配技术主要分为基于像素方法和基于特征方法^[2]. 基于像素的方法采用计算两幅图像间各点像素灰度的差值, 若差值小于某一阈值则为相同部分. 这种方法简单直观, 但运算量大, 对于光照变化较大, 旋转角度较大的情况下可能会匹配失败. 基于特征的方法首先提取图像的特征, 如: 边缘、轮廓、角点^[3]等, 用相似性度量进行图像间特征点的匹配.

此方法大大改进了基于像素方法出现的种种问题, 成为图像匹配的主要方法.

2004 年 Lowe 基于局部不变量描述符方法, 提出了 SIFT^[4]算法, 有效的解决了由于图像部分遮挡、旋转缩放、仿射变换、尺度变换等引起图像变形等问题. 然而此算法在对特征点进行描述时, 描述子为 128 维向量, 导致在生成描述子以及后续的配准过程中计算量过大. 很多学者提出各类基于 SIFT 算法的改进方法: 主成分分析法(PCA)将描述子降维, 减少计算复杂度, Re 和 Sukthakar 等在归一化梯度场中利用 PCA 方法将描述子降维, 稳健性较高, 但是在精度上有所下降; SURF 方法在计算量上降低, 适用于图片分辨变化较大的情况, 但是在仿射

变换和光照变化下,效果欠佳;Mikeolajczyk 提出了扩展的 SIFT 描述子—GLOH,增强了特征描述子的独特性,但在快速图像匹配的效果上也不尽如人意。

文中针对 SIFT 算法的高维描述子进行降维,同时增强了关键点邻域像素信息,达到了快速和高精度匹配的效果。

2 SIFT 算法原理

SIFT 基本思想^[5]是:首先引入尺度空间概念,建立图像的尺度空间表示,在尺度空间中搜索出图像的局部极值点作为候选关键点,这样定位的关键点具有尺度抗缩放性质,然后精确定位候选点,去除低对比度的关键点和不稳定的边缘响应点,最后将关键点进行特征描述,使之具有独特性。图 1 为 SIFT 算法的示意图。

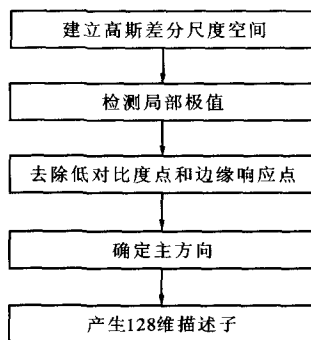


图 1 SIFT 算法示意图

2.1 建立高斯差分尺度空间

尺度空间(scale space)思想是由 Iijima 首先提出,Koenderink 推广,其基本思想是:在视觉信息处理模型中引入一个尺度参数,通过连续变化尺度参数获得不同尺度下的视觉处理信息,然后综合这些信息以深入挖掘图像的本质特征。尺度空间方法将传统的单尺度视觉信息处理技术纳入不断变化的动态分析框架中,因此更容易推导证明了 Gaussian 卷积核是唯一的线性核:

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

将图像高斯卷积得到不同尺度空间中的图像:

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

其中, σ 为尺度空间的空间尺度因子, σ 的大小决定了平滑范围的大小,选择适当的空间尺度因子是建立尺度空间的关键。

为了高效地检测稳定的关键点在尺度空间的位置,引入尺度空间的高斯差分函数的极值与图像的

卷积 $D(x, y, \sigma)$,可以由被常数乘法因子 k 分离的两个邻近尺度的差计算出来:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned}$$

2.2 检测局部极值点

检测 $D(x, y, \sigma)$ 的局部最大值与最小值,每个采样点都要与其当前尺度的 8 个相邻点尺度中上下相邻图像的各 9 个相邻点,共 26 个点进行比较,得到局部的极值,将该点置为候选点。

2.3 去除低对比度点和边缘响应点

针对低对比度的候选点,利用尺度空间函数 $D(x, y, \sigma)$ 的泰勒二次展开式进行最小二乘拟合,通过计算拟合曲面的极值来进一步剔除。高斯差分图像 $D(x, y, \sigma)$ 泰勒展开式为:

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{\partial x^2} x \quad (1)$$

对 D 求导并令其值等于 0,可以解出亚像素精度的偏移量 $\hat{x}$:

$$\hat{x} = -\frac{\partial^2 D^{-1}}{\partial x^2} \times \frac{\partial D}{\partial x} \quad (2)$$

代入式(1)式可得:

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (3)$$

函数 D 在极值点的值 $D(\hat{x})$ 对于排除不稳定的低对比度的极值很有帮助。

针对 DOC 产生的边缘响应点,利用 Hessian 矩

$$\text{阵计算主曲率 } H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

H 的特征值与 D 的主曲率是成比例的。设 α, β 为特征值,且 $\alpha > \beta$,通过矩阵的迹计算特征值的和,通过行列式计算其积:

$$\text{Tr}(H) = D_{xx} + D_{yy} = \alpha + \beta$$

$$\text{Det}(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta$$

设 γ 为最大特征值与小一些的特征值的比率,即 $\alpha = \gamma\beta$,于是有:

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\gamma\beta + \beta)^2}{\gamma\beta^2} = \frac{(\gamma + 1)^2}{\gamma}$$

当两个特征值相等时, $(\gamma + 1)^2/\gamma$ 取得最小值,并随 γ 增大而增大。因此要检验主曲率的比率低于某个阈值 γ 只需要检验式(4):

$$\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(\gamma + 1)^2}{\gamma} \quad (4)$$

2.4 确定主方向

比较稳定的关键点后,利用关键点为中心的邻

域像素梯度信息,确定主方向.首先,用式(5)、(6)计算邻域内各点的梯度值和方向:

$$m(x,y) = \{(L(x+1,y) - L(x,-1,y))^2 + (L(x,y+1) - L(x,y-1))^2\}^{1/2} \quad (5)$$

$$\theta(x,y) = \tan^{-1} \left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)} \right) \quad (6)$$

然后,用直方图统计邻域的像素方向.直方图方向范围是 $0 \sim 360$ 度,LOWE 采用每 10 度为一个格子,共 36 个格子.邻域内各个采样点按照梯度方向 $\theta(x,y)$ 归入所在的格子.直方图的峰值则代表了该关键点处邻域梯度的主方向,即作为该关键点的主方向.

2.5 生成特征描述子

首先,将坐标轴旋转至关键点的主方向,这样就保证了旋转不变性;计算以关键点为中心的 16×16 邻域各点像素梯度值大小和方向,并采用高斯窗加权.然后把特征点邻域范围划分为 4×4 的像素块.在每个 4×4 的像素块中建立梯度方向直方图,此处划分为 8 个区间,并统计 4×4 像素块中每个像素梯度方向属于的区间,这个区间的值为所有梯度方向属于这个区间的像素的梯度幅度大小高斯加权后的累加,用箭头长度表示.因此 4×4 像素块中的每一个像素都可以用一个长度为 8 维的特征描述向量表示.所以对于每一个特征点,一共使用 $4 \times 4 \times 8 = 128$ 维的特征描述向量对其进行描述.

3 SIFT 算法改进

SIFT 算法在高斯差分尺度空间中确定关键点,去除低对比度点和边缘响应点来实现关键点的精确定位,并确定关键点的主方向,达到了尺度、旋转不变性的要求,但是产生的 128 维描述子在匹配过程中耗时过大,影响了算法的效率.

文中针对高维描述子进行改进.原算法中的邻域方向性信息联合的思想能够增强算法的抗噪声能力^[6],同时对于含有定位误差的特征匹配也提供了较好的容错能力.文中也将采用邻域信息构造描述子.改进后的算法如图 2 所示,虚线框为改进部分.

60 维方形邻域描述子的构造方法:以关键点为中心,取临近的 4 邻域为一组,计算各像素的梯度信息 $m(x,y), \theta(x,y)$,统计 8 个方向的梯度值.同样,在第一组向各个方向扩展 2 个像素,此邻域圈为第二组,统计各个像素的 8 方向梯度值.依次类推,在前一组基础上向各个方向扩展 2 个像素,得到第三

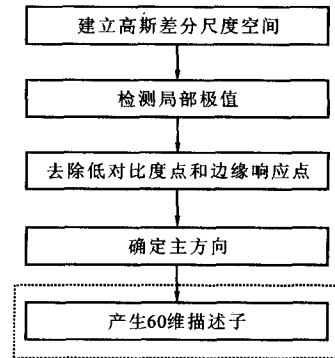


图2 改进后算法示意图

到第八组邻域圈,如图 3 所示.为简化表示,仅取 4 组邻域,得到了 4 组梯度方向值.为了避免关键点发生位移而导致描述子的偏差过大,文中将关键点邻域高斯加权,高斯函数是单值函数,每一邻域像素点权值是随该点与中心点的距离单调增减的,用于减少那些离关键点较远的像素对关键点的影响.

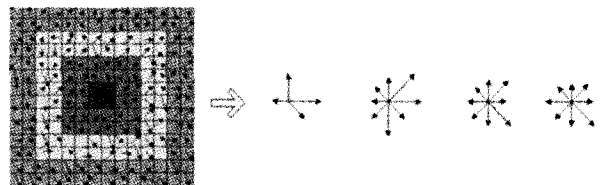


图3 改进后描述子示意图

将第一组的 4 个方向的梯度累加值作为新的描述子向量的第 1 到第 4 个元素,第二组的 8 个方向梯度累加值作为向量的第 5 到第 12 个元素,依次类推,将八组的梯度值排为 $8 \times 7 + 4 = 60$ 元素的向量,此 1×60 的向量即为该关键点的特征描述子.新算法的描述子将特征向量的维数由 128 降到 60 ,同时包含的邻域像素为 30×30 ,大于原算法的 16×16 邻域像素.

4 仿真实验

实验平台为 VISTA 操作系统, CPU 2.00GHz , 内存 2.00GB 的个人电脑,编程环境为 matlab r2010a.采用 Mikolajczyk05^[7]数据集和多种图像进行测试,将测试图进行以下变换:①光照加噪、②旋转压缩、③尺度变换、④模糊变换.匹配策略采用度量向量之间欧氏距离的方法,遍历图中关键的描述子,找到距离差最短的两个点进行匹配.文中采用次临近点(smin)与最临近点(min)的比值(smin/min)来确定更为可靠的匹配点.当此比值大于某一阈值 threshold 时,认为是正确的匹配.

4.1 参数设定

(1)构造高斯差分尺度空间参数关系以及初值

设定:

初始尺度为: $\sigma=1.6$

同频频中各层的迭代关系为: $\sigma=\sigma \times 2^{1/s}$

octave= $\{1,2,3,4\}$, $s=5$

第 octave 组的第一层是上一组的倒数第二层.

尺度空间实质上是对图像进行不同程序的高斯平滑.不同平滑程序的图像对应点的差值表示的是其在平滑过程中的变化程度,越大则其在当前图像平面中梯度幅值越大.生成多个尺度空间目的是为了从多尺度来检测特征点.

(2)去除不明显特征点的对比度阈值为 0.03;去除边缘上的特征点的曲率阈值为 5,当大于 5 时,去除该点.次临近与最邻近之比的阈值 threshold 设为 1.5.

4.2 实验结果

通过实验,改进算法在匹配时间上约为原算法的 30%~60%(见图 4),配准准确率(见图 5)与原算法相近.

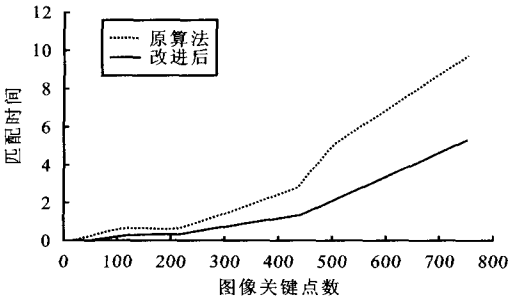


图 4 原算法与改进后算法匹配时间比较

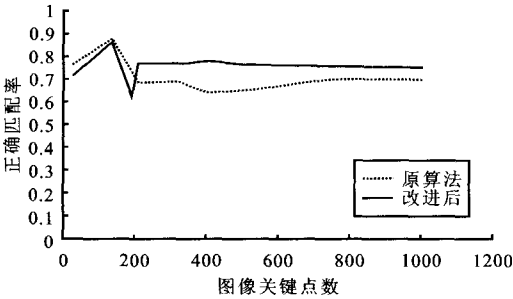


图 5 原算法与改进后算法匹配率的比较

以下是其中三组实验结果:

(1) 采用 Mikolajczyk05 数据集中 Light 组 img1 和 img4,像素值为 900×600 ,本组图片的视角以及像素完全相同,将同一像素位置视为相同点.检测到关键点数分别为 1193,1237.表 1 列举了采用两种算法,成功匹配关键点的个数,正确匹配率以及匹配所用的时间.通过实验 1,可见改进后算法较原算法在匹配正确率相近的情况下,匹配时间大大缩短.

表 1 实验 1 两种算法实验结果的对比

	原算法描述子	改进后描述子
匹配点数	237	195
正确率	79%	78%
匹配时间(秒)	0.81	0.44

(2) 旋转压缩实验

Lena 图像素值为 256×256 将图片旋转 45 度,并做压缩,这样变换图中每一点在原图中都有映射关系,可以精确反映了配准精度.匹配结果如图 6 所示.检测到原图的关键点为 482,变换图的关键点为 135.匹配结果对比见表 2.



(a) 原算法匹配结果



(b) 改进后算法匹配结果

图 6 两种算法匹配结果

表 2 实验 2 两种算法实验结果的对比

	原算法描述子	改进后描述子
匹配点数	101	92
正确率	63%	63%
匹配时间(秒)	0.52	0.27

(3) 采用 Mikolajczyk05 数据集中 Zoom+rotation 组中 img1,预处理压缩后像素值为 400×320 ,匹配检测策略同(2).该图结构较为复杂,特征较多.图中出现了较多的局部相似区域.匹配结果对比见表 3.检测到原图的关键点为 1249,变换图的关键点为 440.

实验 3 表明,方形邻域描述子对关键点描述时扩展了邻域区域,使得描述子更加准确独特,对于有较多的局部相似区域的图像,误匹配率大大降低,从

而提高了正确配准率。

表 3 实验 3 两种算法实验结果的对比

	原算法描述子	改进后描述子
匹配点数	312	242
正确率	64%	78%
匹配时间(秒)	2.87	1.32

5 结束语

改进的 SIFT 算法通过构造 60 维方形邻域描述子,降低了描述子的维数,并且增加了邻域像素统计范围,增强了关键点的邻域信息.通过实验表明,方形邻域描述子匹配的时间是原算法的 30%~60%.当图像特征比较复杂,有多个较小的局部相似区域时,此描述子较原算法有更好的独特性,较原算法减少了误匹配数,提高了配准的精度。

参考文献:

- [1] Lisa Gottesfeld Brown. A survey of image registration techniques[J]. ACM Computing Surveys. 1992, 24(4): 325-376.

- [2] Barbara Zitova, Jan Flusser. Image registration methods: a survey[J]. Image and Vision Computing 2003: 977-100.
- [3] 冯宇平,戴明. 一种基于角点特征的图像拼接融合算法[J]. 微电子与计算机, 2009, 26(7): 21-28.
- [4] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [5] 张伟,刘伟,普杰信. 一种基于 SIFT 和区域选择的图像拼接方法[J]. 微电子与计算机 2010, 27(6): 205-207.
- [6] 张良,王海丽,吴仁彪. 基于改进局部不变特征的兴趣点匹配[J]. 电子与信息学报, 2007(3): 2621-2625.
- [7] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.

作者简介:

王田甲 女, (1984-), 硕士研究生. 研究方向为数字图像处理.

刘国荣 男, (1957-), 副教授. 研究方向为人工智能与数理逻辑.

(上接第 183 页)

5 结束语

文中从 CMOS 图像传感器像素点和双采样电路的工作原理出发,首先分析了提升芯片性能的电压需求,然后提出了电源模块的设计思想,详细介绍了内部电源、Vref1 和 Vref2 电压的产生方法及其关键的技术.最后给出了电源模块的仿真结果,验证了设计的正确性。

参考文献:

- [1] Abbas El Gamal. CMOS Image Sensors[J]. IEEE Circuits&Devices Magazine, 2005.
- [2] 谈新权,何永泰. CMOS 图像传感器及其应用[J]. 激光与红外, 2001, 31(2): 76-77.
- [3] 余有芳. 一种基于电荷泵的 CMOS 图像传感器[J]. 现代电子技术, 2009(18): 185-187.

- [4] 张科,冯全源. 一种带软启动电路的带隙基准电压源的实现[J]. 微电子学与计算机, 2006, 23(12): 179-181.
- [5] 肖海鹏,马卓,张民选,等. 一种低成本高精度 CMOS 基准电压源设计[J]. 微电子学与计算机, 2010, 27(5): 121-124.
- [6] 吕丽峰. TFT-LCD 驱动芯片内置电荷泵研究与设计[D]. 西安: 西北工业大学, 2008.

作者简介:

张学识 男, (1985-), 硕士研究生. 研究方向为混合信号集成电路设计.

魏廷存 男, (1960-), 博士, 教授. 研究方向为混合信号 VLSI 设计和平板显示驱动芯片技术等.

郑 然 男, (1981-), 博士研究生, 讲师. 研究方向为混合信号集成电路设计.