

摘要：验证了一种能够在不同图像之间进行同一个物体相匹配的方法，具有很强的可靠性，称之为SIFT算法（尺度不变特征变换）。SIFT算法能够处理图像间发生的尺度变换、旋转、很大范围内的仿射形变、视角变换、噪声以及光照变换。它的功能十分强大，甚至可以仅仅根据一个简单的物体特征，在一个大型数据库中的许多高品质图像中进行相应目标的寻找与匹配。

关键词：SIFT；特征检测；图像匹配；

中图分类号：TN919.81 文献标识码：A 文章编号：1006-883X (2012) 09-0022-05

收稿日期：2012-05-18

基于 SIFT 算法的目标特征 检测与提取技术研究

阎冲

中北大学信息与通信工程学院 山西太原 030051

一、引言

在现代数字生活中，其很多方面都应用到了目标特征检测和提取的技术，比如摄影测量和遥感、指纹识别、视网膜识别等，甚至在一些更大的领域，图像特征提取技术还起着至关重要的作用，例如能将其应用在国防、军事目标的识别及导弹目标的识别上，而这些关键技术方面的突破正是能够提高我国综合国力的重要象征所在。所以，尺度不变的特征点提取算法及其应用成为了图像处理领域中的一个研究热点，应该加以研究和创新。

SIFT算法其最重要的意义就在于能够处理两幅发生平移，旋转甚至仿射变换的图像，这是很多其他算法所不具备的特点。SIFT特征点是图像的局部特征，其对旋转、尺度缩放、亮度变化保持不变性，对视角变化、仿射变换、噪声也保持一定程度的稳定性。经优化的SIFT匹配算法甚至可以达到实时的要求，还可以很方便的与其他形式的特征向量进行联合。因而近年来，这种尺度不变的特征点提取算法及其应用成为了图像处理领域中

的一个研究热点，并且具有非常重要的意义。

二、SIFT 算法的研究及理论

SIFT 算法其主要思想可以理解为它是利用提取图像局部特征，在尺度空间寻找极值点，提取位置、尺度、旋转不变量来进行图像的匹配。SIFT 算子首先在尺度空间进行特征检测，并确定关键点的位置和关键点所处的尺度，然后使用关键点邻域梯度的主方向作为该点的方向特征，以实现算子对尺度和方向的无关性。^[1]在Mikolajczyk 对包括 Sift 算子在内的十种局部描述子所做的不变性对比实验中，SIFT 及其扩展算法已被证实在同类描述子中具有最强的健壮性。利用 SIFT 方法提取出的特征点对图像尺度和旋转保持不变，对光线变化、噪声、仿射变化都具有鲁棒性。^[2]

1、尺度空间理论

尺度空间理论最早出现于计算机视觉领域，当时其目的是模拟图像数据的多尺度特征。随后 Koendetink 利用扩散方程来描述尺度空间滤波过程，并由此证明高斯

核是实现尺度变换的唯一变换核。Lindeberg、Babaud 等人通过不同的推导进一步证明高斯核是唯一的线性核。

尺度空间理论的主要思想是利用高斯核对原始图像进行尺度变换, 获得图像多尺度下的尺度空间表示序列, 对这些序列进行尺度空间特征提取。对于二维图像 $I(x, y)$ 在不同尺度下的尺度空间 $L(x, y, \sigma)$ 可由图像 $I(x, y)$ 与高斯核 $G(x, y, \sigma)$ 的卷积得到, 公式如下:

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

二维高斯核公式为:

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

其中, σ —尺度因子。

在高斯核中 σ 代表了高斯正态分布的方差, 其值越小则表征该图像被平滑得越小, 其值越大则表征该图像被平滑得越大。^[3]大尺度对应于图像的概貌特征, 小尺度对应于图像的细节特征。因此, 选择合适的尺度因子平滑是建立尺度空间的关键。

2、高斯差分尺度空间 (DoG Scale-Space)

高斯差分算子的计算公式为:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (3)$$

由此看出, 高斯差分尺度空间是利用不同尺度的高斯差分核与图像卷积而生成的, 其目的是为了更有效的在尺度空间检测到稳定的特征点。

DOG 算子计算简单, 是尺度归一化的 LoG 算子的近似。

3、图像金字塔理论

图像金字塔又分为高斯金字塔和高斯差分金字塔 (DOG 金字塔), 检测极值点是在 DOG 金字塔中进行的, 从而提取精确的特征点。

通过拟和三维二次函数以精确确定关键点的位置和尺度 (达到亚像素精度), 同时去除低对比度的关键点和不稳定的边缘响应点 (因为 DoG 算子会产生较强的边缘响应), 以增强匹配稳定性、提高抗噪声能力。

一个定义不好的高斯差分算子的极值在横跨边缘的地方有较大的主曲率, 而在垂直边缘的方向有较小的主曲率。

主曲率可以通过一个 2×2 的 Hessian 矩阵 H 求出:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (4)$$

导数由采样点相邻差估计得到。

D 的主曲率和 H 的特征值成正比, 令 α 为最大特征值, β 为最小的特征值, 则:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta \quad (5)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (6)$$

令 $\alpha = r\beta$, 则:

$$\frac{Tr(H)^2}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(r\beta + \beta)^2}{r\beta^2} = \frac{(r+1)^2}{r} \quad (7)$$

其中, $(r+1)^2/r$ 的值在两个特征值相等的时候最小, 随着 r 的增大而增大。

因此, 为了检测主曲率是否在某域值 r 下, 只需检测:

$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r} \quad (8)$$

取 $r=10^{[5]}$ 。此时, 已经精确的确定了极值点的位置及所处尺度。

最后则需要确定特征点的方向, 利用特征点邻域像素的梯度方向分布特性为每个特征点指定方向参数, 与之后的特征向量种子的方向相结合, 使算子具备旋转不变性。

特征点的梯度模值计算公式为:

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (9)$$

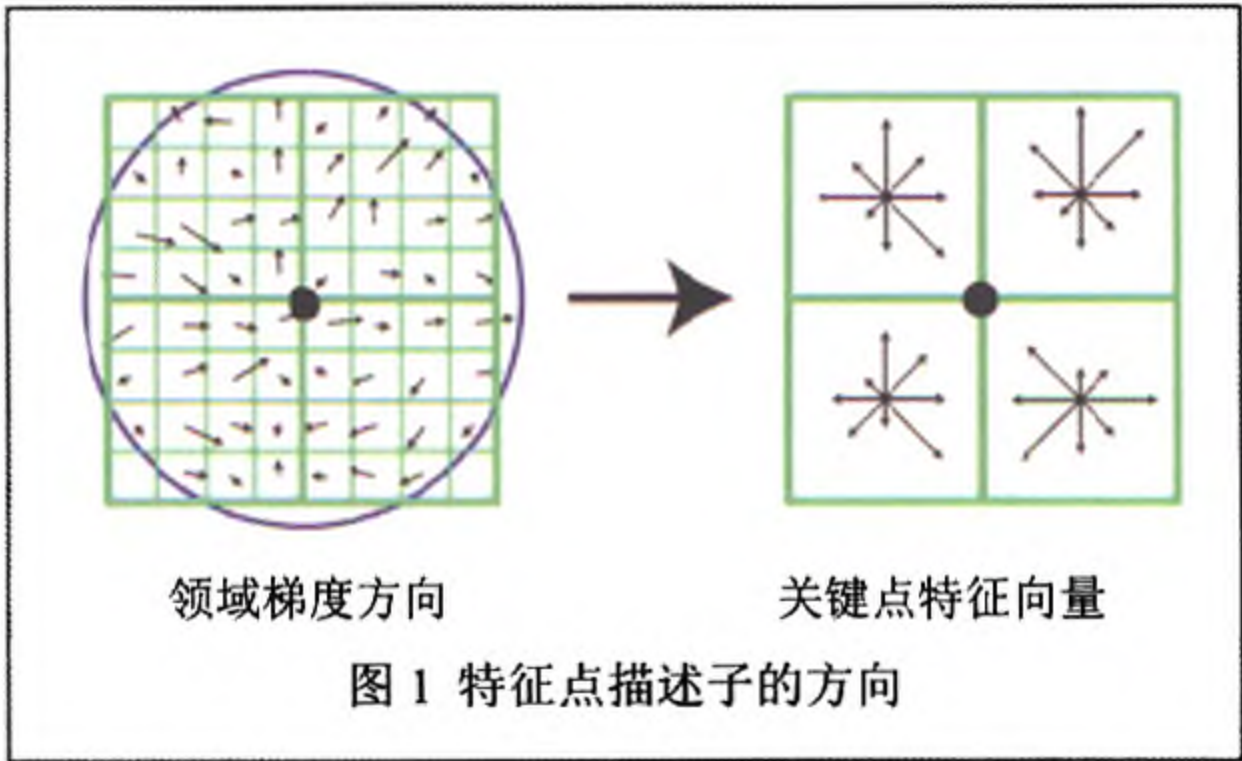
特征点的方向计算公式为:

$$\theta(x, y) = \alpha \tan 2 \cdot \frac{L(X, Y+1) - L(X, Y-1)}{L(x+1, y) - L(x-1, y)} \quad (10)$$

其中, L 所用的尺度为每个关键点各自所在的尺度。

在实际计算时, 我们在以关键点为中心的邻域窗口内采样, 并用直方图统计邻域像素的梯度方向。梯度直方图的范围是 $0 \sim 360^\circ$, 其中每 10° 为一个柱, 总共有 36 个柱。直方图的峰值则代表了该关键点处邻域梯度的主方向, 即作为该关键点的方向。

至此, 图像的特征点已检测完毕, 每个特征点应有的三个信息: 位置、尺度、方向已具备。由此下面便可以开始确定一个 SIFT 特征区域了。



首先将坐标轴 X,Y 任意一个设置为关键点方向, 以确保旋转不变性。然后如图 1, 以特征点为中心, 取其所在的差分尺度空间周围 8×8 的像素点, 注意这里不包括特征点所在的行和列。

接着, 分别求出每个像素点的梯度方向与模值, 图中的箭头方向代表该像素的梯度方向, 箭头长度代表梯度模值, 蓝色的圈代表高斯加权的范围(越靠近关键点的像素梯度方向信息贡献越大)。

最后, 在以坐标轴为分割的 4 个小块, 也就是每 4×4 的一个块上计算 8 个方向的梯度直方图, 得出每个梯度方向的累加值, 便形成了一个种子点。当然, 种子点取的越多, 则其特征点越具有鲁棒性, 程序就越健壮。

实际计算过程中, 为了增强匹配的稳健性, 建议对每个特征点使用 4×4 共 16 个种子点来描述, 每个种子点有 8 个方向向量信息, 这样对于一个特征点就可以产生 $4 \times 4 \times 8$ 共 128 个数据, 最终形成 128 维的 SIFT 特征向量, 所需的图像数据块为 16×16 。此时 SIFT 特征向量已经

去除了尺度变化、旋转、放缩等几何变形因素的影响, 再继续将特征向量的长度归一化, 则可以进一步去除光照变化的影响。

这种邻域方向性信息联合的思想增强了算法抗噪声的能力, 同时对于含有定位误差的特征匹配也提供了较好的容错性。当两幅图像的 SIFT 特征向量即特征描述器生成后, 就可以进行特征向量的匹配了。

三、目标特征提取的实现

这里给出用于实验的原始灰度图像 A、B(图 2、图 3), 图像大小为 256×256 像素。

首先, 通过初始化程序将 RGB 图像(即为 .jpg 格式的图像)转化为灰度图像。这样, 以后文中所提及的图像变量 I 都指的是图像的灰度级。

1、利用计算公式求出图像的尺度空间



图 2 原始灰度图像 A



图 3 原始灰度图像 B

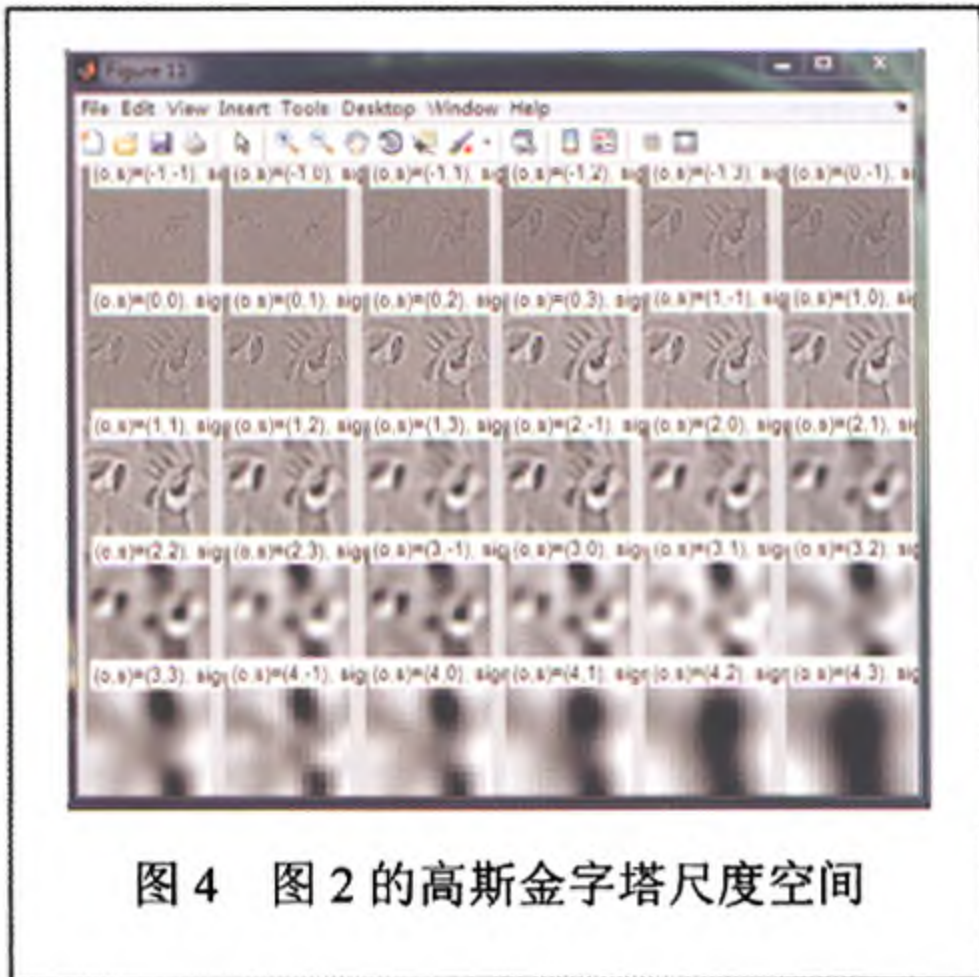


图 4 图 2 的高斯金字塔尺度空间

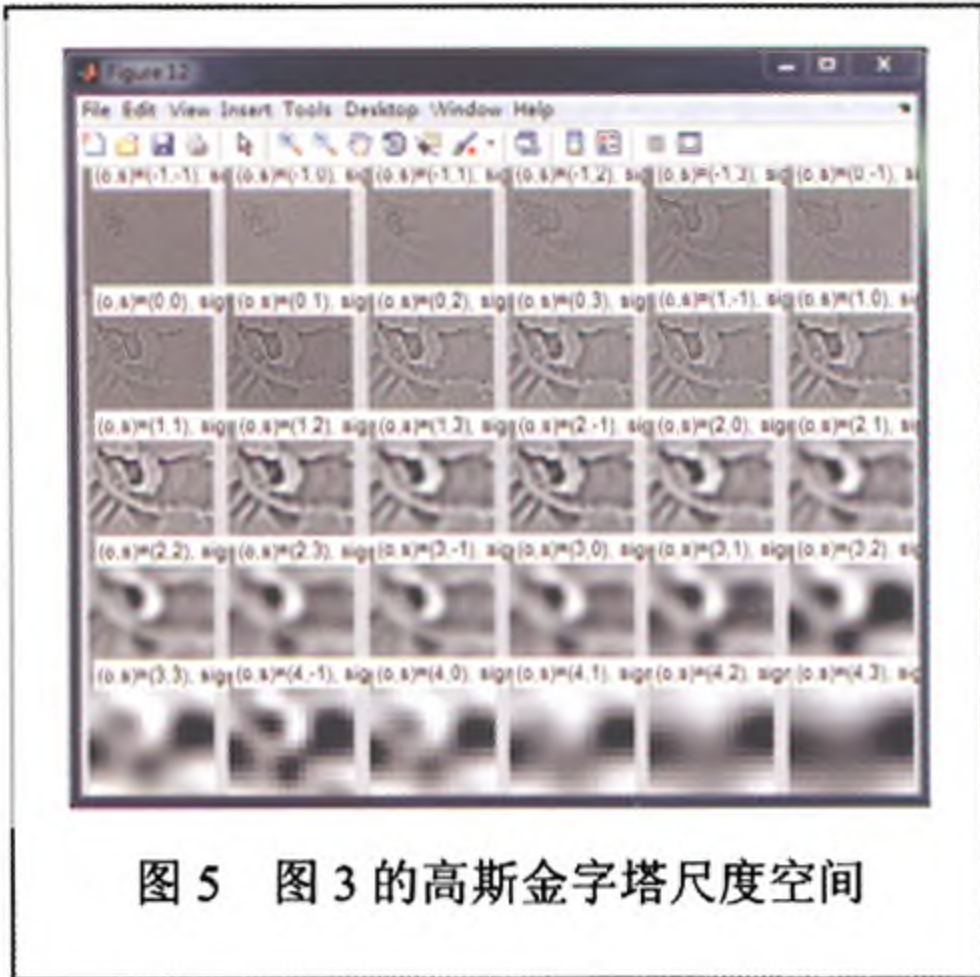


图 5 图 3 的高斯金字塔尺度空间

在 Matlab R2009a 下, 对原始图像进行求尺度空间的操作。读入要处理的原始灰度图像, 令变量 x, y 分别表示图像的行列像素, 写出高斯核, 进行卷积滤波处理。其中 π 为圆周率常量, δ 为尺度因子。^[6]

2、建立图像金字塔

在 Matlab R2009a 下, 通过自定义函数求出金字塔各组各层的尺度空间的操作。令变量 O 为金字塔的组数, S 为金字塔的层数, 通过公式 (1), 令 δ 前的系数 $k = \sqrt{2}$, 逐次求出各层尺度空间, 金字塔第二组第一层的尺度空间为前一组中间层的降采样。

程序运行结果, 图 2 的金字塔各层的尺度空间如图 4 所示。其中 (O, S) 代表第 O 组第 S 层的尺度空间, 将这些尺度空间按金字塔样式排列好就构成了图 2 的高斯金字塔及 DOG 金字塔。

图 3 金字塔各层的尺度空间如图 5 所示。其中 (O, S)

代表第 O 组第 S 层的尺度空间, 将这些尺度空间按金字塔样式排列好就构成了第二幅图像的高斯金字塔及 DOG 金字塔。

3、图像特征点的提取

在 DoG 尺度空间检测极值点。在 Matlab R2009a 下, 编程实现在差分尺度空间中间层次寻找局部极值点, 即对比其周围 26 个像素点的大小, 统计局部极大值和极小值。除去不符合条件的特征 (极值) 点。利用拟和三维二次函数精确确定极值点的位置及尺度, 并去掉不符合条件的极值点, 可以利用编程的技巧等同来解决, 分别将循环的变量始末值设为 2 和 255, 可去除边缘点, 低对比度点可利用将特征点与周围 26 个点中最大值比较的方法处理。

程序结果如图 6、图 7 所示。图中红色箭头为为低对比度或边缘极值点, 应舍去。绿色箭头为精确确定的极

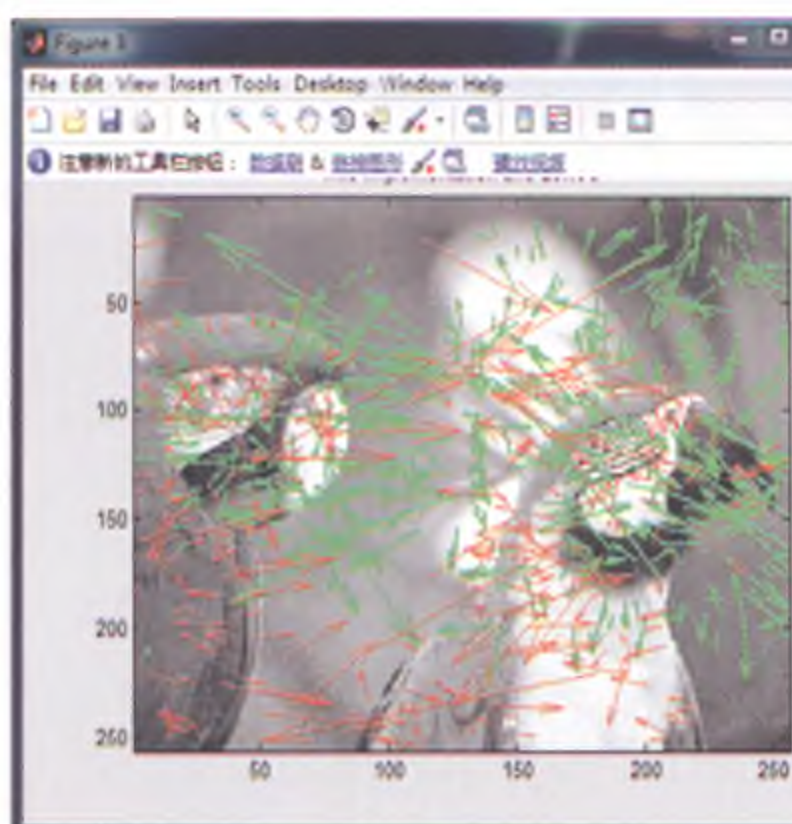


图 6 图像特征点的提取

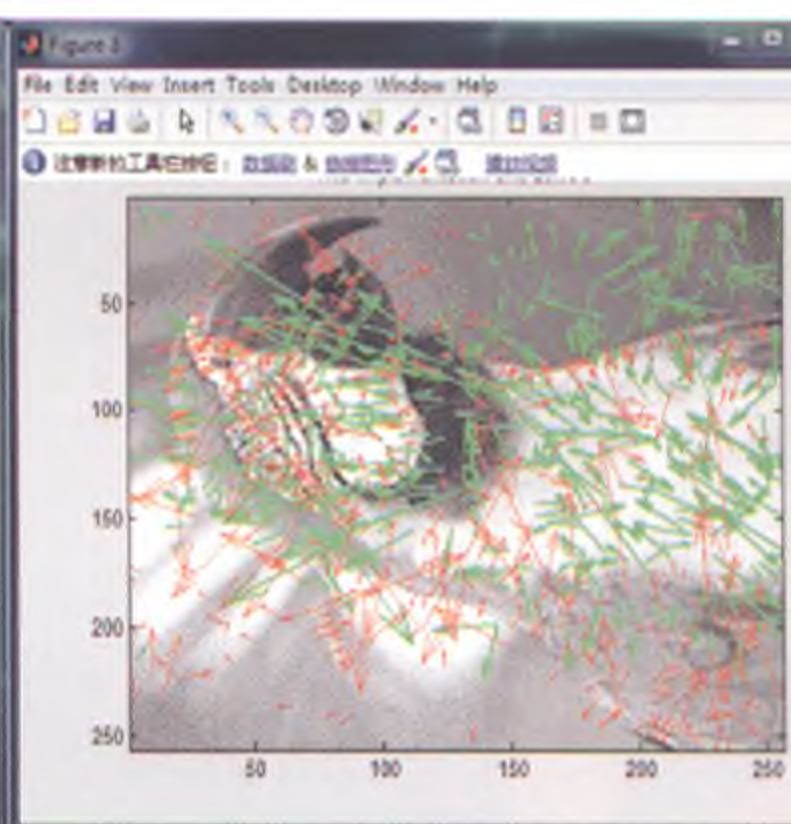


图 7 图像特征点的提取

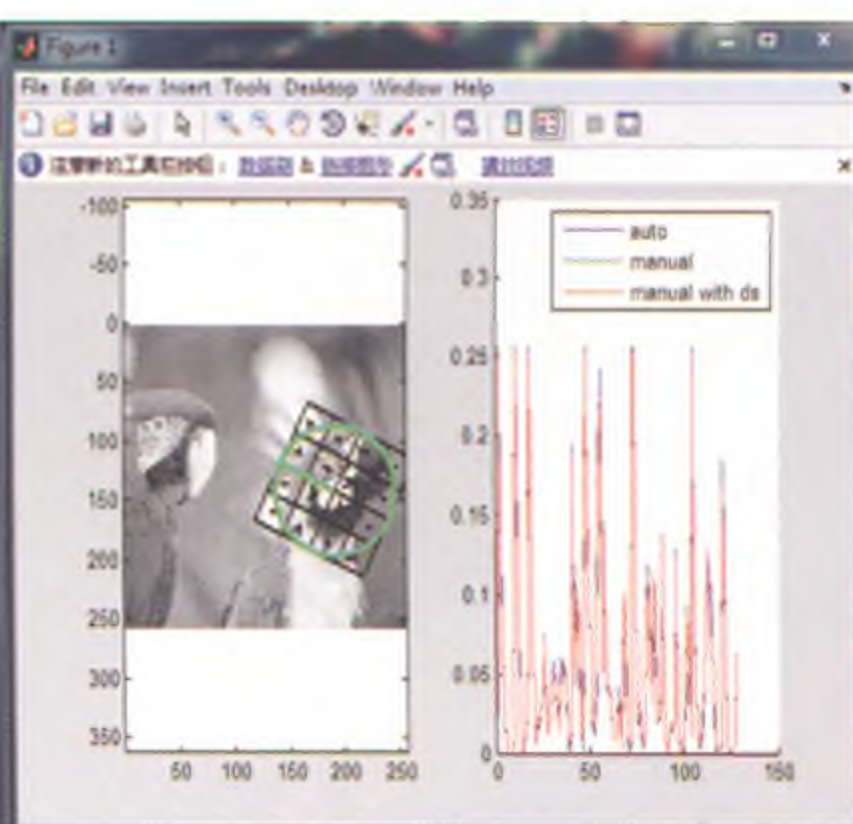


图 8 特征点的主方向

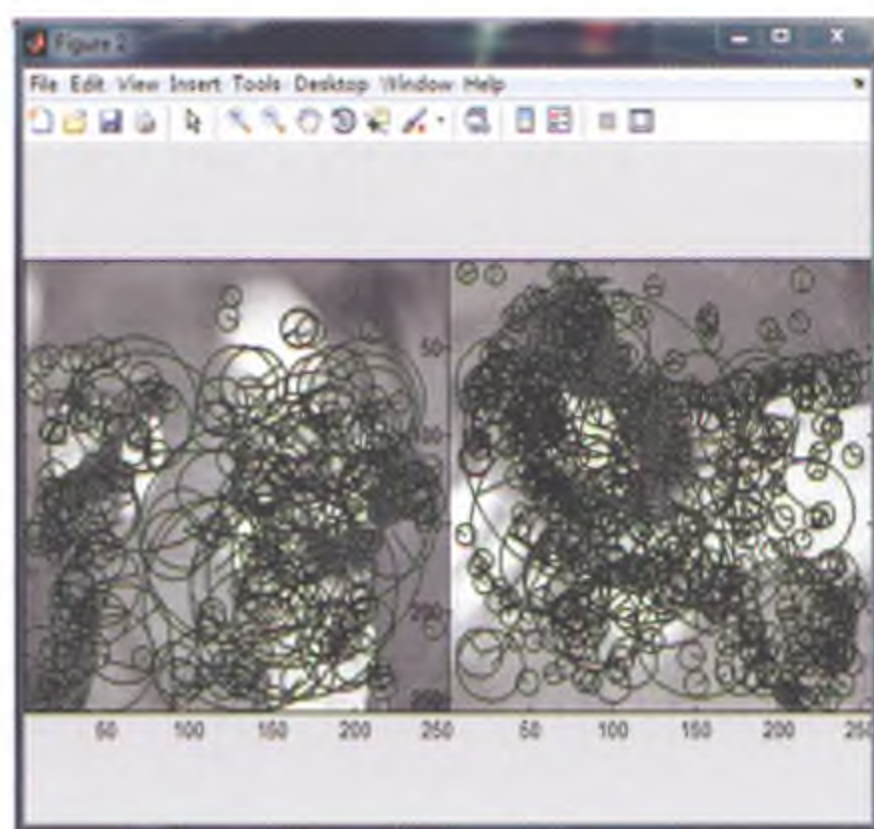


图 9 特征向量的计算结果图

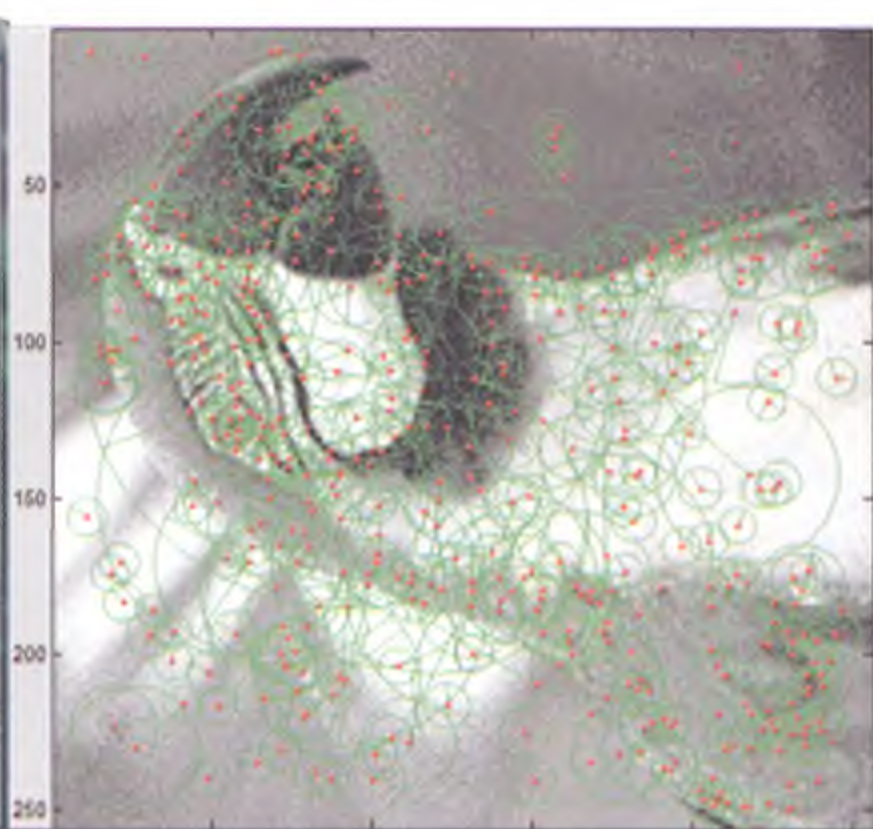


图 10 特征向量的计算结果图

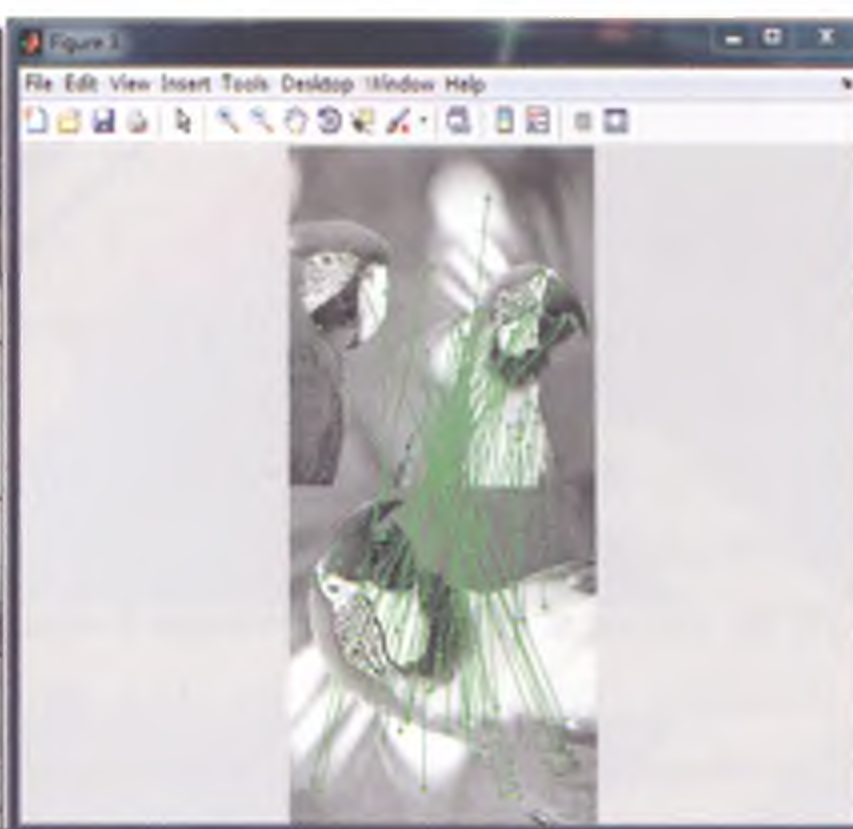


图 11 特征向量的计算结果图

值点。虽然有些绿色点看似是处在边缘,但其实并不是边缘点,只是人眼的视觉不准确而已。

4、计算特征点的主(辅)方向

在以特征点 $a(N)$ 为中心的邻域窗口内采样,并用直方图统计邻域像素的梯度方向。直方图的峰值代表该关键点的邻域梯度方向,即作为该关键点方向。

如图 8 所示,为其中一个特征点的主方向(图中绿线所指方向)。

在极值点的方向确定之后,就已经提取到了位置、尺度、方向三要素完整的可靠性及鲁棒性较强的特征点了。

5、特征向量的计算

下面我们为已经精确确定的特征点来建立其描述子,也就是特征向量。

以特征点 N 为中心,取其所在的差分尺度空间周围 8×8 的像素点,不包括特征点所在的行和列,即 $a(N)-4$ 至 $a(N)+4$, $b(N)-4$ 至 $b(N)+4$ 。利用公式分别求出每个像素点的梯度方向与模值,越靠近关键点的像素梯度方向信息贡献越大。然后以每 4×4 的一个块上计算 8 个方向的梯度直方图,得出每个梯度方向的累加值形成一个种子点。运行结果如图 9、图 10、图 11 所示。

由这些特征向量所描述的一样,图 11 的鹦鹉确实为同一只,当生成了描述子便可以进行图像的匹配了。

四、总结

本论文所做的主要工作是利用 SIFT 算法的设计思路,检测和提取图像的特征点。经优化的 SIFT 匹配算法可以达到实时的要求,还可以很方便的与其他形式的特征向量进行联合。

参考文献

- [1] 王国美,陈孝威. SIFT 特征匹配算法研究[J];盐城工学院学报(自然科学版);2007,20(2):1-5
- [2] 钟力,胡晓峰. 重叠图像拼接算法[J]. 中国图象图形学报,1998,3(5):367-370.
- [3] M. Brown and D. Lowe. Invariant Features from Interest Point Groups[A]. Proceedings of the 13th British Machine Vision Conference[C], Cardiff, 2002:252-262.
- [4] DUCH W, Heural minimal distance methods [A] PIOC of the 3rd Conference on Neural Networks and Their Applications[C]

Kule:[sn],1997:183-188

- [5] L NDEBERG T Feature detection with automatic scale selection [J] International Journal of Computer Vision 1998, 30(2):79-116
- [6] CHO S Y, WONG Jia jun. Probabilistic based recursive model for face recognition [R] Lecture Notes in Computer Science 3641. [S. l.]: Springer Verlag, 2005: 1245 21254.

Research on target feature detecting and extracting technologies based on SIFT algorithm

YAN Chong

(College of information and communication, North University of China, Taiyuan 030051, China)

Abstract: A method is tested and verified for extracting distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene, which is named as Scale Invariant Feature Transform. The features are invariant to image scale and rotation, and are shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise, and change in illumination. The features are highly distinctive, in the sense that a single feature can be correctly matched with high probability against a large database of features from many images.

Keywords: SIFT; feature detection; image matching

作者简介

阎冲, 中北大学信息与通信工程学院, 硕士研究生, 研究方向: 信息与通信系统

通讯地址: 北京市昌平区立汤路 186 号龙德紫金大厦 1 号楼 510 室

邮编: 102218 电子邮箱: yanchongzhi@163.com

读者服务卡编号 004□