

摘要：提出了一种基于哈尔小波分解变换和高斯尺度空间的图像特征点匹配算法。首先利用哈尔小波变换对基础图像进行3层行列分解，然后利用高斯函数卷积核对这些分解图像，进行尺度变换。提出了一个小波高斯金字塔塔林的概念，即对通过小波变换产生的多张不同分辨率的基础图像分别进行高斯尺度变换进而产生一个个独立的高斯金字塔，进而产生独立的高斯差分金字塔林，完成特征点检测。再引进规范化强对比度描述子对特征点进行描述。结果表明：Haar-Gaussian&NICD算法的效果和SIFT算法相当，特征点数量优于SIFT算法，在局部特征匹配方面要更有优势；而且和NICD描述子搭配使用，在运行速度方面要比SIFT算法更快。

关键词：哈尔小波变换；尺度空间；金字塔塔林；特征点检测；特征匹配

中图分类号：TP391.4

文献标识码：A

文章编号：1006-883X(2018)04-0023-06

收稿日期：2018-03-25

一种图像特征点匹配算法 ——Haar-Gaussian&NICD 算法

崔领袖 张奇志 周亚丽

北京信息科技大学 自动化学院，北京 100192

一、引言

随着社会的发展，在图像识别领域小波变换是傅里叶变换之后的另外一个比较有力的工具。原则上讲，传统意义上使用傅里叶分析的地方，都可以用小波分析来替代^[1]。就多分辨率分析而言，小波变换的出现使得图像处理领域又提供了一种新思路，小波框架的产生使得处理数字信号更加的精准。1989年，为了研究图像的纹理特征 Mallat 引入了小波的概念，为后续的基于小波纹理分析的方法提供了扎实的理论基础。小波变换是在不同尺度上研究分析图像纹理细节的一种工具，为更精准的进行图像纹理分类和分析提供了新的思路，在纹理分析中有广阔的发展空间^[2]。

图像的局部特征提取——特征点提取，在实现图像合成、图像校正、识别分类、图像匹配等图像领域常见问题中是及其重要的一个环节，关于如何在一幅图像中精准的定位其稳定的特征点用于后续的相关工作，国内外大量的相关学者和专家都做了深入的研究

工作。

其中有一种较为广为人知的图像特征点提取算法称为Harris角点检测算法^[3]。虽然这种算法对图像平移、旋转、和噪声具有较强的鲁棒性，但是当图像遭遇尺度幅变换之后其效果会很糟糕，因此从一定程度上讲它的使用会受到限制。随着图像技术的发展，越来越多的基于强特征点提取匹配的算法被提出，其中包括SIFT算法^[4]、Harris-Laplacian算法^[5]等。

本文主要讨论一种特征点检测子算法的实现，并与规范化强对比度描述子(Normalized intensity contrast descriptor, NICD)相结合，目的是设计实现一种适用于一定程度的尺度变换、旋转情况下快速提取独特性强特征点的特征检测算法。本文的主要思路是将原始图片进行3层哈尔小波变换，得到不同尺度和不同方向上的图像信息，利用高斯函数进行不同尺度的高斯滤波进而产生哈尔小波高斯金字塔塔林，再利用高斯拉普拉斯函数计算提取相对最稳定的图像强

度特性特征点，在通过计算来修正特征点，最后引进 NICD 描述子对所提特征点进行描述。

二、检测子设计

1、小波理论

小波理论中重要的部分是多分辨率分析即多尺度分析，这也是与我们人的眼睛观察物体时的机理高度相似的。因为我们对图像进行操作所以在这里我们重点讨论二维小波变换。在二维情况下，首先定义一个尺度基函数和平移基函数：

$$\varphi_{j,m,n}(x,y) = 2^{j/2} \varphi(2^j x - m, 2^j y - n) \quad (1)$$

$$\psi_{j,m,n}^i(x,y) = 2^{j/2} \psi^i(2^j x - m, 2^j y - n) \quad i = \{H, V, D\} \quad (2)$$

其中，上标 i —沿列、行、对角方向。

于是，大小为 $M \times N$ 的图像 $f(x,y)$ 的离散小波变换是：

$$W_\varphi(j_0, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \varphi_{j_0,m,n}(x,y) \quad (3)$$

$$\begin{cases} W_\psi^H(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}^H(x,y) \\ W_\psi^V(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}^V(x,y) \\ W_\psi^D(j, m, n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}^D(x,y) \end{cases} \quad (4)$$

其中， j_0 是其中一个开始的任意尺度， $W_\varphi(j_0, m, n)$ 系数为定义 $f(x,y)$ 在尺度 j_0 处的近似， $W_\psi^i(j, m, n)$ 系数对尺度 $j \geq j_0$ 增添了水平、铅垂和斜对角方向的细节。

因为有 $j=0,1,2,\dots, J-1$ 和 $m=n=0,1,2,\dots, 2^j-1$ ，所以通常情况令 $j_0=0$ ，并且选择 $N=M=2^J$ 。图 1 展示了这一过程，在二维情况下我们得到 3 组细节系数——水平系数、铅垂系数和对角线细节系数。

本文中主要应用哈尔小波变换，因为它的基波是最古老的且最简单的正交小波，所以哈尔小波变换是其中最重要的一个环节。哈尔变换可以用如下矩阵形式表示：

$$T = H F H^T \quad (5)$$

其中， F —一个 $N \times N$ 的图像矩阵；

H —一个 $N \times N$ 的哈尔变换矩阵；

T —一个 $N \times N$ 的变换结果。

因为 H 的非对特性，所以在转置是必须要有，其包含哈尔基函数 $h_k(z)$ 。它们定义在连续闭区间 $z \in [0,1]$ ， $k=0,1,2,\dots, N-1$ 上，其中 $N=2^n$ 。

为了得到矩阵 H ，本文在这里定义整数 k ，即 $k=2^p+q-1$ ，其中 $0 \leq p \leq n-1$ 。当 $p=0$ 时， $q=0$ 或 1 ；当 $p \neq 0$ 时， $1 \leq q \leq 2^p$ 。

因此，哈尔基函数为：

$$h_0(z) = h_{00}(z) = \frac{1}{\sqrt{N}}, \quad z \in [0,1] \quad (6)$$

$$h_k(z) = h_{pq}(z) = \frac{1}{\sqrt{N}} \begin{cases} 2^{p/2} & \frac{q-1}{2^p} \leq z < \frac{q-0.5}{2^p} \\ -2^{p/2} & \frac{q-0.5}{2^p} \leq z < \frac{q}{2^p} \\ 0 & \text{其它}, z \in [0,1] \end{cases} \quad (7)$$

$N \times N$ 哈尔变换矩阵的第 i 行包含了元素 $h_i(z)$ ，其中： $z = \frac{0}{N}, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}$ 。

2、Haar-Guassian 金字塔塔林

早在 1962 年，Lijima 提出来了尺度空间的思想，由于 Witkin 和 Koenderink 等人的研究工作，使得尺度空间的思想能够被计算机领域广泛使用^[3]。尺度空间的主要思想是：首先定义一个系数——尺度系数，引入到图像信息处理的模型中，由于尺度系数变化的连续性，会产生很对的尺度描述序列，然后基于轮廓的描述对这些尺度序列进行特征提取，并将提取到的特征数据组合成一种特征向量，进而能够得到图像中相应的边缘信息、角点信息以及不同分辨率上的特征

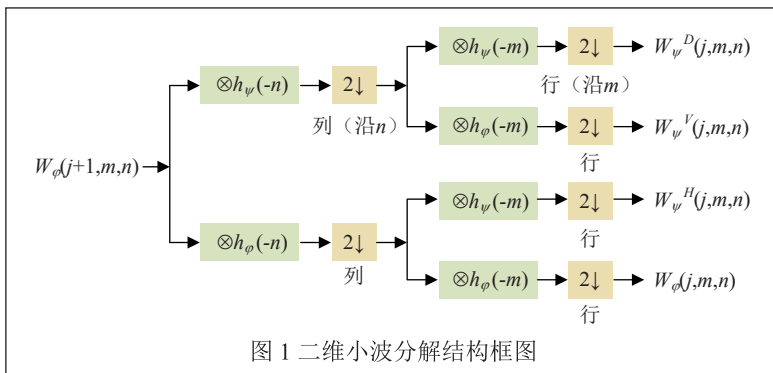


图 1 二维小波分解结构框图

信息等。

尺度空间中是满足视觉不变性的，可以有如下解释：当物体被我们用肉眼观察时，由于背景光强度的不同，图像的光亮被视网膜感知的效果以及对比度是截然不同的，因此分析图像时要求尺度空间算子能够做到不被图像灰度和对比度影响，也就是说在一定程度上满足灰度和对比度不变性。此外，在同一坐标系下，当物体和我们的肉眼观察的发生相对位置变化时，图像位置、大小、角度和形状在我们的肉睛的视网膜上的成像是不同的，所以在进行图像分析时，图像的位置信息、大小信息、角度信息以及仿射变换等信息具备的互不干扰性，也就是满足平移、尺度、欧几里得不变以及仿射不变性。

本文在这里使用一个变换尺度的高斯函数 $G(x, y, \sigma)$ 与原图像 $I(x, y)$ 的卷积和来定义该图像的尺度空间 $L(x, y, \sigma)$ ，即：

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (8)$$

$$\text{其中, } G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\frac{m}{2})^2 + (y-\frac{n}{2})^2}{2\sigma^2}} \quad (9)$$

其中， $m \times n$ 为模板大小， (x, y) 为像素的位置， σ 是正态分布的标准差，作为尺度空间描述子， σ 的值越大，图像越平滑（模糊）， σ 的值越小表示图像被平滑的越少，对应的尺度也就越小。这样在分析图像时，图像的全局特征与相对较大尺度的特征图相对应，细节特征则与相对较小尺度的特征图相对应。

接着本文就要对图像进行哈尔分解。首先对图像进行哈尔变换产生多分辨率序列，并以不同分辨率的哈尔特征图金字塔底层，为了使尺度体现连续性，对每一层的一张图片进行不同参数的高斯模糊，因为每一种分辨率都有三或四种基础特征图，所以多种分辨率的情况下会产生多个金字塔林列，所以在里引进了金字塔塔林的概念。

2002 年，Mikolajczyk 在详细的实验比较中发现尺度归一化的高斯拉普拉斯特征提取函数 $\sigma^2 \nabla^2 G$ 的极值更能产生稳定的特征图像^[5]。与 $\sigma^2 \nabla^2 G$ 非常相似的函数是高斯差分函数 (Difference of Gaussian, 简称 DOG 算子)。 $D(x, y, \sigma)$ 与 $\sigma^2 \nabla^2 G$ 的关系可推导如下：

$$\frac{\partial G}{\partial \sigma} = \sigma^2 \nabla^2 G \quad (10)$$

将微分用差分来代替得到如下：

$$\sigma^2 \nabla^2 G = \frac{\partial G}{\partial \sigma} \approx \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma} \quad (11)$$

$$\text{则有: } G(x, y, k\sigma) - G(x, y, \sigma) \approx (k-1)\sigma^3 \nabla^2 G \quad (12)$$

其中， $k-1$ 为常数，不影响极值点所在地方。

本文使用效率更高的 DOG 算子来替代拉普拉斯算子进行极值点的检测，如下：

$$\begin{aligned} D(x, y, \sigma) &= [G(x, y, k\sigma) - D(x, y, \sigma)] * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (13)$$

为了得到高斯差分图像，计算时将上下两层图像的相减，如图 2 所示，进而进行极值点的检测。

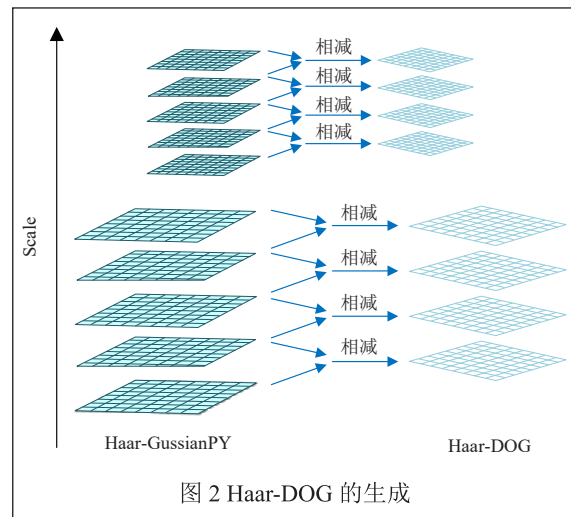


图 2 Haar-DOG 的生成

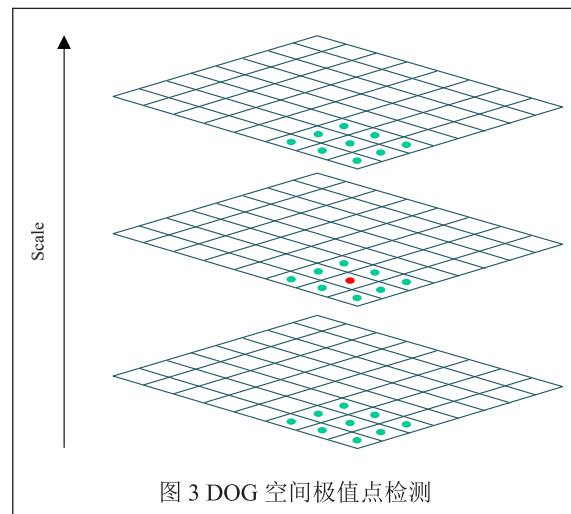


图 3 DOG 空间极值点检测

3、特征点的定位和修正

特征点的定位是通过同组高斯金字塔内各 DOG 相邻之间的比较完成的,按 SIFT 算法的思想将中间点的像素值和周围 26 个点作比较,以保证在尺度和二维图像空间都可以检测到极值点。如图 3 所示。

通过上面步骤检测到的极值点可大致分为两种:强特征性极值点和弱特征性极值点,强特征性极值点的提取是本文的主要目的,而弱特征性极值点的存在往往会对后续的研究工作和应用产生误差。因此本文在引用 SIFT 特征点修正的基础上,首先利用哈尔变换的去噪机理对 DOG 图像进行处理,然后将 DOG 数据进行归一化的处理,并设定合理阈值,最后得到较为整齐的强特征性极值点。

三、特征点描述子选择

在 SIFT 描述子中,采用加权的梯度方向直方图^[6],维度较高大大地影响了计算效率。随后,Ke 等人提出了 PCA-SIFT 描述子降维后进行匹配^[7]。SURF 描述子是在 SIFT 的基础上,计算兴趣点的尺度相关邻域内 X 和 Y 方向的哈尔小波,并选择相对邻域内响应最大的作为主方向^[8]。Huang 等人^[9]提出了对比上下文直方图,利用邻域区域的像素强度之间的线性关系进行特征描述,减少局部特征描述的计算时间,但是他仅仅考虑了常量强照度的变化。

本文采用 NICD 来对本文介绍检测子检测到的特征点进行描述,在将局部邻域内的像素强度规格化后,描述每个子区域中的像素与中心像素的强度对比关系^[10],可得到 48 维或者 64 维的特征描述向量。

四、实验结果与分析

1、实验过程

本文在特征点检测试验中采用重复率^[11]来评价特征点检测子的鲁棒性。首先对原始图像和变换图

像进行特征点的解算,然后求取特征点重复的数量与这两幅图像分别提取的特征点的最小值的比值作为重复率。实际计算中,通过两幅图像提取出的特征点不可能完全的对应,所以,计算时取图像位置小于 2^i (i 为特征点所在不同分辨率的级数, $i=0, 1, 2, \dots$) 个像素。此外,实验中采用 64 维的 NICD 描述子并用特征点检测时的特征点的梯度和偏移量等对描述子向量进行扩展,构成 73 维的描述向量。

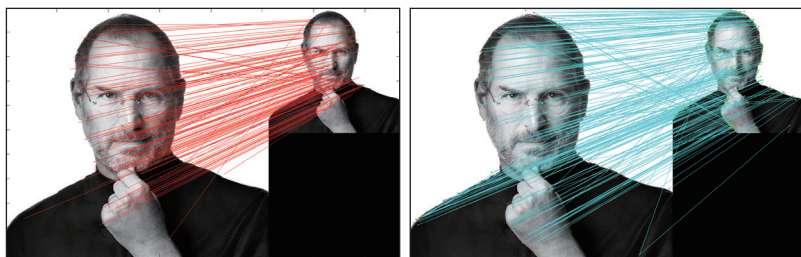
2、评估标准

采用计算余弦相似度的方法来对描述子向量进行匹配,匹配结果用对应特征点连线图来展示。

3、实验结果

(1) 尺度变换匹配实验

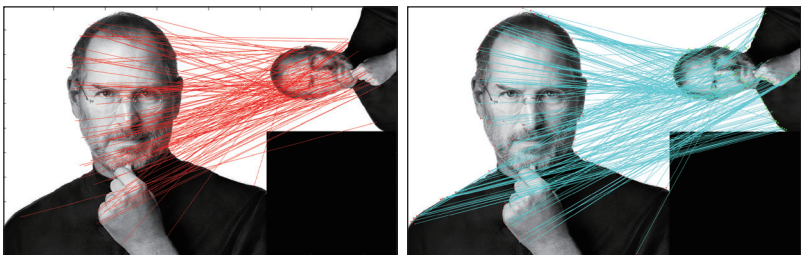
本文使用的原始图片为 $512\text{px} \times 512\text{px}$,变换后图片为 $256\text{px} \times 256\text{px}$ 。图 4 分别展示了 SIFT 和本文所用方法的效果。经过对比可以看出,本文的算子提取的强特征性特征点的数量要优于 SIFT 算法,进而可以更好地应用于图像局部特征的匹配,另外特征描述子的维度只有 SIFT 算子的一半,从匹配速度上来说要优



(a) SIFT 匹配效果图

(b) Haar-Gaussian&NICD

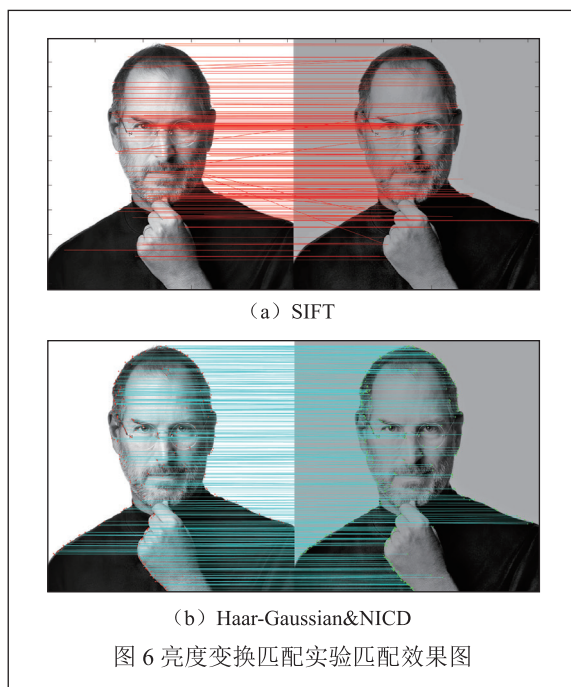
图 4 尺度变换匹配实验匹配效果图



(a) SIFT

(b) Haar-Gaussian&NICD

图 5 尺度和旋转变换匹配实验匹配效果图



于 SIFT，从匹配效果上来说和 SIFT 相当。

(2) 尺度和旋转变换匹配实验

在尺度变换的基础上加上旋转变换，效果如图 5 所示。经过对比可以看出，首先从特征点的数量上来讲要优于 SIFT 算法，在图像局部特征匹配方面的实用性更强，速度上由于维数的降低要优于 SIFT 算法，在这种变换情况下，从匹配效果上来讲要略好于 SIFT 算法。

(3) 亮度变换匹配实验

亮度变换匹配实验匹配效果如图 6 所示。通过对比可以看出，由于 NICD 描述子的构建过程中，去除亮度的相对干扰，对亮度不敏感，所以匹配效果要明显优于 SIFT 算法。

本文分别在图像的尺度变换、旋转尺度变换以及亮度变换的情况下进行了特征点的匹配实验，实验表明本文所提的方法有效，满足图像在以上三种情况变换下所提特征点的强特征性不变，并可以有效的应用于图像的局部特征匹配。

五、总结

特征点的强特征性在尺度、亮度、旋转等变换下

都可以保持不变，继而成为近年来机器视觉和模式识别的研究热点。本文的思路是将哈尔小波变换和高斯函数相结合，构建金字塔塔林的结构作为特征点的提取工具，引入 NICD 快速描述子将局部邻域像素规范化，通过对目标区域中特征点与周围像素点之间的规范化强度差进行计算，并结合特征点检测时的特征点的梯度和偏移量等数据合并组成 73 维的描述特征向量。在实验中分别考虑了尺度、旋转、亮度的变换情况，证明了该算法的有效性。

机器人视觉识别系统都是对局部特征的重要应用，在未来的工作中，我们将把这种特征点检测算法应用到机器人视觉物品识别中，验证它的性能。仿真实验已经证明这种算法的有效性，此外我们还会对该算法进行改进，提出更好的描述子及其搭配用于实时的特征检测和货架商品识别。

参考文献

- [1] 薛智刚, 李巴津. 基于 Haar 小波的图像变换方法的研究[J]. 微计算机信息, 2006, 30: 275-277.
- [2] 孙正. 数字图像处理与识别[M]. 北京: 机械工业出版社, 2014.
- [3] C. G. Harris, M.J.Stephens. A combined corner and edge detector[A]. In: Proc of the 4th Alvey Vision Conference[C]. Manchester:[s. n.], 1988: 147-151.
- [4] David G. Lowe. Distinctive Image Features from Scale Invariant Keypoints[J]. International Journal of Computer Vision 2004, 60 (2): 91-110.
- [5] K. Mikolajczyk, C. Schmid. Indexing based on scale invariant interest points[A]. In: Proceedings of the 8th International Conference on Computer Vision[C]. Vancouver:[s. n.], 2001: 525-531.
- [6] David G. Lowe. Object Recognition from Local Scale-Invariant Features[A]. In: The Proceedings of the Seventh IEEE International Conference on Computer Vision[C]. 1999: 20-27.
- [7] Yan Ke, R. Sukthankar. PCA-SIFT: a more distinctive representation for local image descriptors[A]. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C]. Washington D. C., USA: IEEE, 2004: 506-513.
- [8] Herbert Bay, Tinne Tuytelaars, Luc Van Gool. SURF: speeded up robust features. In: Proceedings of the 9th European Conference on Computer Vision[C]. Graz, Austria: Springer,

2006: 404-417.

[9] C.R.Huang, C. S. Chen, P. C. Chung. Contrast context histogram-an efficient discriminating local descriptor for object recognition and image matching[J]. Pattern Recognition, 2008, 41(10): 3071-3077.

[10] 刘萍萍, 赵宏伟, 臧雪柏, 戴金波. 一种快速局部特征描述算法 [J]. 自动化学报, 2010, 36(1):40-45.

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

[12] A Lanitis, CJ Taylor, TF Cootes. A unified approach to coding and interpreting face images[A]. In: Proceedings of the 5th International Conference on Computer Vision (ICCV' 95)[C]. 1995: 368-373.

An Algorithm for Image Feature Point Matching—Haar-Gaussian&NICD Algorithm

CUI Ling-xiu, ZHANG Qi-zhi, ZHOU Ya-li

(School of Automation, Beijing Information Science and Technology University, Beijing 100192, China)

Abstract: An algorithm based on Haar wavelet transform and Gaussian scale-space is proposed for image feature point matching. First the basic image is carried out three-layer row-column decomposition with Haar wavelet transform, and then scale-space transformation with Gaussian kernel. A conception of Haar wavelet-Gaussian Pyramid forest is proposed, which means the basic images with different resolutions produced by Haar wavelet transform are respectively carried out Gaussian scale-space transformation to generate a series of independent Gaussian pyramids, and then generate independent Gaussian difference pyramid forest to complete detection of feature points. After that the Normalized Intensity Contrast Descriptor(NICD) is introduced to describe the feature points. The experimental results show that the Haar-Gaussian&NICD algorithm has similar performance to SIFT, more feature points than SIFT, which is more advantageous in local feature point matching, and is faster than SIFT in terms of operating speed with the introduction of NICD.

Key words: Haar wavelet transform; scale-space; Gaussian Pyramid forest; feature points detection; feature matching

作者简介

崔领袖: 北京信息科技大学自动化学院, 硕士研究生, 主要就研究方向为模式识别, 图像处理及深度学习。

通讯地址: 北京市海淀区清河小营 12 号北京信息科技大学

邮编: 100192

邮箱: 690914300@qq.com

张奇志: 北京信息科技大学自动化学院, 教授, 博士, 主要研究方向为机器人控制及图像处理。

周亚丽: 北京信息科技大学自动化学院, 教授, 博士, 主要研究方向为机器人控制及信号处理。