

# 基于非监督特征学习的兴趣点检测算法

周来恩 王晓丹

(空军工程大学防空反导学院 西安 710051)

**摘要** 由于兴趣点是图像中的基础、关键特征,因此兴趣点检测是图像配准、图像检索以及图像识别的关键步骤。基于兴趣点对于图像特征响应较为强烈的特性,结合非监督特征学习算法可以自主地从无标签的样本中提取特征的思想,提出了 UFL-ID 兴趣点检测算法。该算法无监督学习了图像的底层特征,对特征进行信息量和各向同性的评价,并利用特征的卷积响应及评价参数寻找图像中的兴趣点。与其他常见的兴趣点检测算法的对比实验表明,该算法具有良好的重复性与抗噪能力。

**关键词** 机器学习,非监督特征学习,自动编码器,兴趣点检测,特征提取

**中图分类号** TP391.41 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2016.9.058

## Unsupervised Feature Learning Based Interest Point Detection Algorithm

ZHOU Lai-en WANG Xiao-dan

(Academy of Anti-missile and Air Defense, Air Force Engineering University, Xi'an 710051, China)

**Abstract** Interest point is of great importance in digital image processing as a kind of critical feature at low level. So the interest point detection is the committed step in image registration, image retrieval and image recognition. In this paper, an unsupervised feature learning based interest point detection (UFL-ID) was presented based on the fact that interest points have stronger feature convolution response than others. The new UFL-based interest point detection algorithm firstly learns low level features in digital images, evaluates the information content and isotropy of learned features, and finally uses features and its evaluation to find interest points. The comparison result demonstrates that using UFL produces great improvements of repeatability and anti-noise property.

**Keywords** Machine learning, Unsupervised feature learning, Auto-encoder, Interest point detection, Feature detection

### 1 引言

兴趣点是像素值在时空邻域变化较大的点,其局部邻域包含了丰富的图像运动信息。兴趣点检测是实现图像匹配、拼接等处理的关键步骤。经过发展,兴趣点检测的方法已有许多成熟模型。

Moravec<sup>[1]</sup>依据以像素为中心的 4 个方向灰度值的变化情况来判断当前像素是否为兴趣点,但是由于其算法只检测了 4 个方向上的灰度变化情况,因此算法是各向异性的。Harris<sup>[2]</sup>为了改进 Moravec 算法的各向异性的问题,通过计算区域窗口中图像任意方向上的灰度自相关值的总和来确定兴趣点。Susan<sup>[3]</sup>则定义了一个圆形模板,通过统计以像素为中心的模板中的核相似区域的面积来判断这个像素是否为兴趣点,由于不存在微分计算,因此算法对噪声较不敏感,但是其几何门限参数的设定严重影响兴趣点提取的质量,且参数的确定多靠人工启发式的方法。DOG 算法<sup>[4]</sup>使用了不同尺度的二维高斯差函数与图像进行卷积,在卷积后的图像中将局部最大值的点标记为兴趣点,由于噪声会影响 DOG 函数

卷积运算,因此会导致局部最大值偏离原兴趣点位置,造成判断误差<sup>[5]</sup>。

结合机器学习,根据利用已有的数据来改善自身算法的性能的特点,学者们提出了许多基于机器学习的兴趣点检测算法。FAST 算法<sup>[6]</sup>利用 ID3 算法建立的决策树判断了以像素为中心、半径为 3 像素的圆形模板的 16 个边缘像素的灰度变化情况,认为这 16 个边缘像素中如果存在连续的 12 个像素的灰度值比中心像素的灰度值大或小,那么中心像素就为兴趣点;FAST 算法利用决策树根据几个边缘像素的灰度变化情况来判断当前点是否为兴趣点,所以在速率和运算代价上表现出众,但是由于检测的像素点数目太少,因此抗噪性能不强<sup>[6]</sup>。NNSIC 算法<sup>[7]</sup>通过建立神经网络来判断输入的点是否为兴趣点,其用于训练的样本中兴趣点度数都为 45°的倍数,虽然神经网络可以模拟高度的非线性过程,并很好地处理分类问题,但是由于输入的训练样本有限,对于样本角度不是 45°的倍数的兴趣点判断错误率较高<sup>[7]</sup>。GNG-C 算法<sup>[8]</sup>通过建立 GNG 神经网络<sup>[9]</sup>来对输入的参数化的曲线进行聚类,之后通过聚类后相邻点之间的角度来判断是否为兴趣点,由

到稿日期:2016-03-01 返修日期:2016-05-24 本文受国家自然科学基金:基于多特征融合和集成学习的多目标识别技术研究(61273275)资助。

周来恩(1992-),男,硕士生,主要研究方向为机器学习、数字图像处理,E-mail:zhou\_laien@126.com;王晓丹(1966-),女,博士,教授,博士生导师,主要研究方向为机器学习、模式识别、数字图像处理。

于需要提取图像的边缘且对边缘进行聚类学习,因此该算法速率较慢<sup>[8]</sup>。

对于现有的基于机器学习的兴趣点检测算法而言,确定兴趣点特征的方法不够理想。在理想状态下,兴趣点检测算法可以自主地确定兴趣点的特征,而不是花费大量时间去启发式地设计特征,且这种手工生成的特征(Hand-Engineered Features)泛化能力不强,对实际的兴趣点检测中存在的噪声、光照变化、传感器差异等干扰的适应能力弱。

非监督特征学习(Unsupervised Feature Learning, UFL)在近期的机器学习领域得到了重要的应用。非监督特征学习的主要目的是从无标签的数据中学习用于表达数据的特征,良好的特征应对输入的数据具有选择性和鲁棒性<sup>[10]</sup>;对于图像而言,好的低层特征不仅可以用于图像去噪、图像分割和图像重建,还可用于构建高层特征<sup>[11]</sup>,而这些高层特征可以完成图像识别、图像分类等复杂任务。针对启发式确定特征造成算法的泛化能力不强的问题,本文结合非监督特征学习算法自主学习图像特征的特性,提出了一种基于非监督特征学习的兴趣点检测算法(Unsupervised Feature Learning based Interest Point Detection, UFL-ID)。图像的特征由非监督特征算法自主提取,对比实验结果表明本文提出的基于非监督特征学习的兴趣点检测算法的泛化能力较好。

## 2 基于非监督特征学习的兴趣点检测算法(UFL-ID)

对于兴趣点而言,由于其像素值在时空邻域变化较大,所以相对于像素较为平滑的区域而言,兴趣点对图像特征的响应会更为强烈。本文基于兴趣点对图像特征响应较为强烈的特性,结合非监督特征学习算法可以自主地从无标签的样本中提取特征的思想,提出了UFL-ID兴趣点检测算法,算法描述如下:

- (1) 利用稀疏自编码器提取图像特征;
- (2) 对提取的图像特征进行信息量和各向同性评价;
- (3) 利用特征的卷积响应及特征评价参数寻找图像中的兴趣点。

### 2.1 提取图像特征

在提取图像特征步骤中,为了解决普通自编码器在提取图像特征时存在的不稳定问题,使用添加了稀疏正则项的自编码器——稀疏自编码器来提取图像的特征。

#### 2.1.1 自编码器

自编码器(Auto-encoder)<sup>[12]</sup>是用于数据降维和特征提取的双层神经网络,第一层是由输入层和中间隐层节点组成的编码层,第二层是中间隐层和输出层节点组成的解码层,编码层输入节点的个数与解码层输出节点的个数相等。对于样本图像,将其构造为一系列 $n \times n$ 的像素块样本 $f(x, y)$ ,之后利用自编码器提取出图像的特征。输入信号 $f(x, y)$ 经过编码器后得到编码信号 $f_{enc}(f(x, y))$ ,编码信号可以看作是输入信号的表示,之后通过解码器将编码信号解码得到解码信号 $f_{dec}(f_{enc}(f(x, y)))$ 。通过不断调整编码器和解码器的参数,使得 $f_{dec}(f_{enc}(f(x, y))) \approx f(x, y)$ ,这样就可以将隐含层 $F_k$ 看作对输入层 $f(x, y)$ 的特征的提取,达到对 $f(x, y)$ 学习特征的目的。自编码网络示意图如图1所示。

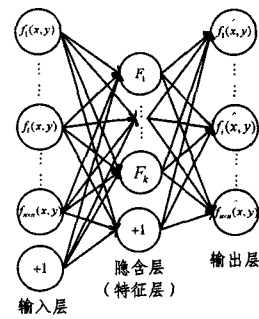


图1 自编码网络示意图

自编码器的隐含层、输出层的激活函数可以表示为:

$$F = S(f(x, y); w, b) \quad (1)$$

$$f(x, y) = S(F; w', b') \quad (2)$$

其中 $S$ 为sigmoid函数:

$$S(x; w, b) = \frac{1}{1 + e^{-w \cdot x + b}} \quad (3)$$

由于sigmoid函数的输出范围为 $[0, 1]$ ,为了使 $h_{w, b}(f(x, y)) \approx f(x, y)$ ,在输入像素块 $f(x, y)$ 时将 $f(x, y)$ 的值规范化到 $[0, 1]$ 之间。

求解自编码器 $h_{w, b}(f(x, y))$ 的参数的方法是依据网络的损失函数使得自编码器参数在解空间中梯度下降。

自编码器 $h_{w, b}(f(x, y))$ 的损失函数为:

$$J(w, b) = \frac{1}{n \times n} \sum_{i=1}^{n \times n} \left( \frac{1}{2} \| h_{w, b}(f_i(x, y)) - f_i(x, y) \|^2 \right) + \frac{\lambda}{2} \sum_{l=1}^{L-1} \sum_{i=1}^{s_l} \sum_{j=1}^{s_{l+1}} (w_{ji}^{(l)})^2 \quad (4)$$

其中, $k$ 表示特征层节点的数目, $L=3$ 表示层的数目, $s_l$ 表示第 $l$ 层的节点数目,其中 $s_1 = n \times n, s_2 = k, s_3 = n \times n$ 。

式(4)中第一项代表输出层与输入层的均方差,表示模型的误差;第二项是权重衰减项,其作用是减少权重的幅度,防止模型出现过拟合;系数 $\lambda$ 表示权重衰减项的权值。

#### 2.1.2 添加稀疏正则项

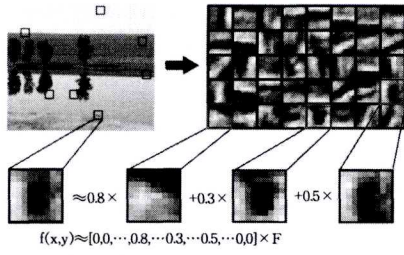
图像构成的复杂性、不确定性以及噪声干扰使得神经网络参数求解问题不稳定,神经网络的参数收敛会不确定或者不稳定。为了解决自编码器参数求解的不稳定问题,可以在求解模型中引入正则化(Regularize)方法。

对于自编码器而言,添加正则化的求解模型为:

$$(\hat{w}, \hat{b}) = \min \{ J(w, b) + \beta \varphi(w, b) \} \quad (5)$$

其中,第二项为正则项, $\beta$ 为正则化参数,用于控制误差项与正则项的平衡; $\varphi(w, b)$ 为理想情况下先验的正则化函数。在添加了正则项后,就可以将不稳定问题转换为稳定。

在求解图像的不稳定问题的模型建立的过程中,正则项的构造包含了图像的先验信息。在正则项的构建方法中,稀疏表示得到了较为广泛的应用<sup>[13]</sup>。稀疏性是指将图像信号表示为一组基的线性组合时只需较少的几个基就可以将信号线性表示出来的特性,其揭示了图像主要几何结构的特征与分布。稀疏性有助于获取数据样本中高相关性<sup>[14]</sup>,对于图像而言,高相关性意味着像素块可以被表示成少量基本元素的叠加,在图像中这些基本元素可能是点、线或者面。稀疏特征在像素块中的表现如图2所示。



图中所示的像素块中只用了 40 个特征的 3 个特征表现

图 2 稀疏特征在像素块中的表现

一种构建稀疏正则项的方法是计算稀疏常数  $\rho$  与第  $j(j \in [1, k])$  个隐藏神经元对于输入样本的平均活跃程度  $\bar{\rho}_j$  的 KL 距离<sup>1)</sup>。其中  $\bar{\rho}_j$  为：

$$\bar{\rho}_j = \frac{1}{n \times n} \sum_{i=1}^{n \times n} S(f_i(x, y); \omega_{ji}^{(1)}, b_{ji}^{(1)}) \quad (6)$$

这样可以得到稀疏正则项的表达式为：

$$\beta \sum_{j=1}^{s_2} KL(\rho \| \bar{\rho}_j) \quad (7)$$

其中  $\beta$  一般取值为 3, KL 为 KL 距离, 其表达式为：

$$KL(\rho \| \bar{\rho}_j) = \rho \log \frac{\rho}{\bar{\rho}_j} + (1 + \rho) \log \left( \frac{1 - \rho}{1 - \bar{\rho}_j} \right) \quad (8)$$

这里通常使  $\rho$  等于一个接近 0 的数(一般为 0.01)。则添加了基于 KL 距离的稀疏正则项的网络损失函数为：

$$\begin{aligned} J_{sparse}(\omega, b) &= J(\omega, b) + \beta \sum_{j=1}^{s_2} KL(\rho \| \bar{\rho}_j) \\ &= \frac{1}{n \times n} \sum_{i=1}^{n \times n} \left( \frac{1}{2} \| h_{\omega, b}(f_i(x, y)) - f_i(x, y) \|^2 \right) + \\ &\quad \frac{\lambda}{2} \sum_{i=1}^{l-1} \sum_{j=1}^{s_1} \sum_{j=1}^{s_2} (\omega_{ji}^{(l)})^2 + \beta \sum_{j=1}^{s_2} \left( \rho \log \frac{\rho}{\bar{\rho}_j} + (1 + \rho) \right. \\ &\quad \left. \log \left( \frac{1 - \rho}{1 - \bar{\rho}_j} \right) \right) \end{aligned} \quad (9)$$

得到神经网络的损失函数之后, 可以通过 BP (Back Propagation) 算法计算损失函数对于网络中权重的偏导数, 之后通过梯度下降法对网络权重进行训练。通过训练, 得到了稀疏自编码器输入层到隐含层(特征层)的参数  $\omega^{(1)}, b^{(1)}$ 。

### 2.1.3 可视化特征

得到训练后稀疏自编码器的参数  $\omega^{(1)}, b^{(1)}$  后, 对于特征层的第  $j$  个节点而言, 使得节点最大激活的输入  $f_j^{\max}(x, y)$  代表了样本在此编码节点的可视化特征。由于求解  $f_j^{\max}(x, y)$  是平凡问题, 为了解非平凡, 必须添加限制条件, 这里限制  $f_j(x, y)$  模长最大值固定为  $l_f$ ：

$$\| f_j(x, y) \| = l_f \quad (10)$$

由于 S 函数为递增函数, 要使得  $S(f_j(x, y); \omega_j^{(1)}, b_j^{(1)})$  最大, 则等价于使得  $\omega_j^{(1)} \cdot f_j(x, y)$  最大, 由向量点乘公式得：

$$\begin{aligned} \omega_j^{(1)} \cdot f_j(x, y) &= \| \omega_j^{(1)} \| \times \| f_j(x, y) \| \times \cos(\theta) \\ &= \| \omega_j^{(1)} \| \times l_f \times \cos(\theta) \end{aligned} \quad (11)$$

其中,  $\theta$  为  $\omega_j^{(1)}$  与  $f_j(x, y)$  的夹角, 当  $\theta = 0$  即  $\omega_j^{(1)}$  向量与  $f_j(x, y)$  向量同向时  $\omega_j^{(1)} \cdot f_j(x, y)$  达到最大, 由此可得编码节点  $j$  的可视化特征为：

$$f_j^{\max}(x, y) = \frac{\omega_j^{(1)}}{\| \omega_j^{(1)} \|} \cdot l_f \quad (12)$$

之后, 将可视化特征  $f_j^{\max}(x, y)$  作为稀疏自编码器从图像中自主提取的特征。

## 2.2 评价特征

稀疏自编码器从图像中自主提取了特征  $f^{\max}(x, y)$  后, 从信息量和各项同性对  $f^{\max}(x, y)$  进行评价。

### 2.2.1 信息量评价

衡量信息量的一个重要指标是熵,  $f^{\max}(x, y)$  的熵  $H$  的计算式为：

$$H_{f^{\max}(x, y)} = - \sum_{t=0}^{m-1} p_{r_t} \log p_{r_t} \quad (13)$$

其中,  $m$  代表图像块  $f^{\max}(x, y)$  中灰度分级的数目,  $p_{r_t}$  代表第  $t$  灰度级出现的概率。信息熵  $H_{f^{\max}(x, y)}$  代表了图像块  $f^{\max}(x, y)$  所含的信息量,  $H_{f^{\max}(x, y)}$  越大说明图像块  $f^{\max}(x, y)$  所含的信息量越多, 通过计算  $H_{f^{\max}(x, y)}$  可以得到提取的每种特征信息量的多少。这里使用归一化的熵  $V_j^H$  作为对编码节点  $j$  的特征的信息量评价：

$$V_j^H = \frac{H_{f_j^{\max}(x, y)}}{\max(H_{f_j^{\max}(x, y)})} \quad (14)$$

### 2.2.2 各向同性评价

各向同性是指图像特征在同一图像中的分量不会因为方向的变换而变化的特性, 即同一特征在不同方向所得到的性能数值相同。为了评价提取特征各向同性, 本文使用了任意方向的自相关值进行评价<sup>[2]</sup>。自相关值的计算公式为：

$$E_{f^{\max}(x, y)} = \sum_{u, v} \omega(u, v) (I(u+x, v+y) - I(u, v))^2 \quad (15)$$

其中,  $I(u, v)$  表示在图像块  $f^{\max}(x, y)$  中滑动的子窗口,  $\omega(u, v)$  是子窗口大小的高斯窗函数,  $I$  代表图像的灰度值, 通过子窗口  $I$  在图像内不停滑动, 计算出图像任意方向的自相关值。  $I(u+x, v+y)$  可以用二阶泰勒函数进行近似：

$$I(u+x, v+y) \approx I(u, v) + I_x(u, v)x + I_y(u, v)y \quad (16)$$

其中,  $I_x$  和  $I_y$  是  $I$  在  $x$  和  $y$  方向上的梯度, 图像的自相关值可以写为：

$$\begin{aligned} E_{f^{\max}(x, y)} &\approx \sum_{u, v} \omega(u, v) (I_x(u, v)x - I_y(u, v)y)^2 \\ &= (x, y) M (x, y)^T \end{aligned} \quad (17)$$

$M$  为二阶矩阵：

$$M = \sum_{u, v} \omega(u, v) \begin{bmatrix} I_x^2(u, v) & I_x(u, v)I_y(u, v) \\ I_x(u, v)I_y(u, v) & I_y^2(u, v) \end{bmatrix} \quad (18)$$

若  $\lambda_1, \lambda_2$  是  $M$  的特征值, 则当  $\lambda_1$  和  $\lambda_2$  十分接近时, 说明图像在不同方向上自相关值的差别较小, 进而说明特征图像各向同性性能较好, 这里使用平方倒数来量化  $\lambda_1, \lambda_2$ ：

$$R^D = \frac{1}{(\lambda_1 - \lambda_2)^2} = \frac{1}{\text{Trace}^2(M) - 4 \times \text{Det}(M)} \quad (19)$$

其中,  $\text{Trace}(M)$  代表矩阵  $M$  的迹,  $\text{Det}(M)$  代表矩阵  $M$  的行列式, 之后将归一化的平方倒数  $V_j^D$  作为对编码节点  $j$  的特征的各向同性评价：

$$V_j^D = \frac{R_j^D}{\max(R^D)} \quad (20)$$

### 2.3 检测兴趣点

现通过稀疏自编码器提取了图像的  $k$  个特征, 并依据信息量和各向同性对特征进行评价, 得到了评价参数  $V$ , 接下来利用兴趣点选取比例  $R_{ip}$ 、特征  $f^{\max}(x, y)$  和特征评价参数  $V$  来检测兴趣点。

由于卷积表示了两个函数所重叠的量, 因此本文使用特征作为卷积掩膜, 对图像进行卷积运算来表示图像中一点在

<sup>1)</sup> Ng, Andrew. CS 224A Lecture Notes. <http://cs224a.stanford.edu/materials.html>

此特征上的分量。对于待检测的整副图像  $IMG(x_{img}, y_{img})$  以及特征  $f^{max}(x, y)$  而言,其对应卷积运算为:

$$C = IMG(x_{img}, y_{img}) * f^{max}(x, y) \\ = \sum_{(m, n) \in \vartheta} IMG(x_{img} - m, y_{img} - n) f^{max}(m, n) \quad (21)$$

其中,  $\vartheta$  为  $f^{max}(x, y)$  掩膜所在的领域。在对图像和每个节点编码的特征图像进行卷积运算后,得到了  $k$  个卷积特征图。之后定义  $p(x_{img}, y_{img})$  点的兴趣点响应函数为:

$$R(p) = \sum_{j=1}^k C_j(x_{img}, y_{img}) V_j^H V_j^D \quad (22)$$

其中,  $V_j^H$  和  $V_j^D$  为第  $j$  个特征的熵评价值和各向同性评价价值。

得到兴趣点响应函数后,将点依据响应值排序,定义响应值最大的前  $|IMG| \times R_{ip}$  个为兴趣点。

UFL-ID 算法描述如下。

#### 算法 1 UFL-ID

输入: 图像 IMG、像素块大小  $n$ 、兴趣点选取比例  $R_{ip}$ 、特征提取数目  $k$   
输出: 图像 IMG 中的兴趣点集合

- Step1 将图像 IMG 分割为一系列  $n \times n$  大小的像素块,将像素块集合记为 patch,并将 patch 中的元素大小归一化到  $[0, 1]$ 。  
Step2 建立稀疏自编码器,输入层和输出层节点个数为  $n \times n$ ,特征层节点个数为  $k$ 。  
Step3 用 patch 训练稀疏自动编码器,得到输入层到特征层的权重  $w^{(1)}$ 。  
Step4 利用  $w^{(1)}$  计算特征  $f^{max}(x, y)$ 。  
Step5 计算特征  $f^{max}(x, y)$  的评价参数  $V$ 。  
Step6 将  $f^{max}(x, y)$  与 IMG 进行卷积运算得到卷积特征图  $C$ 。  
Step7 由  $C$  和  $V$  计算图像中每一个像素的兴趣点响应  $R$ 。  
Step8 对图像中每一个像素进行  $3 \times 3$  像素领域的非极大值抑制后,将像素点依据响应值排序,提取响应最大的前  $|IMG| \times R_{ip}$  个放入集合 INTP 中。  
Step9 返回集合 INTP。

### 3 重复性评价

对于一种兴趣点检测算法,通常对检测出的兴趣点进行重复性评价<sup>[15]</sup>,以此来判断算法的性能。

对于同一空间物体的不同平面图像而言,兴趣点重复性是指用不同图像描述同一空间点时在不同图像中复现的特性,兴趣点重复比例越大,说明有越多的空间点在不同图像中复现,进而说明了兴趣点提取的质量越高。假设现有空间点  $p$ ,  $p$  在图像  $I_i$  和  $I_j$  上的投影为  $p_i$  和  $p_j$ ,那么有:

$$p_j = H_{ji} p_i \quad (23)$$

其中,  $H_{ji}$  是图像  $I_i$  到  $I_j$  的投影变换。

如果  $p_i$  和  $p_j$  都在图像  $I_i$  和  $I_j$  上分别被标定为兴趣点,那么就判定  $p_i$  和  $p_j$  为重复的兴趣点。由于图像  $I_i$  和  $I_j$  对空间点  $p$  的描述存在尺度、3D 角度、灰度等指标的不同,因此存在一定误差,故定义  $p_i$  和  $p_j$  重复:

$$(p_i, p_j) \text{ is repeated} = \begin{cases} \text{true,} & \text{dist}(H_{ji} p_i, p_j) < \epsilon \\ \text{false,} & \text{else} \end{cases} \quad (24)$$

其中,  $\epsilon$  表示容许误差距离,  $\text{dist}(a, b)$  表示点  $a$  与  $b$  的距离。

设在图像  $I_i$  和  $I_j$  上检测出兴趣点集合  $P_i$  和  $P_j$ ,那么点  $p_i \in P_i$  在  $I_j$  上投影为  $D_{ij} = \{d_{ij} | d_{ij} = H_{ji} p_i, p_i \in P_i\}$ ,  $p_j \in P_j$  在  $I_i$  上投影为  $D_{ji} = \{d_{ji} | d_{ji} = H_{ij} p_j, p_j \in P_j\}$ ,则对于图像  $I_i$  和  $I_j$ ,重复的兴趣点集合  $D(i, j; \epsilon)$  为:

$$D(i, j; \epsilon) = \{(d_i, d_j) | \text{dist}(d_i, d_j) < \epsilon, d_i \in D_i, d_j \in D_j\} \quad (25)$$

对于图像  $I_i$  和  $I_j$ ,提取的兴趣点的重复比率  $R_{ij}$  为:

$$R_{ij} = \frac{|D(i, j; \epsilon)|}{\min(|D_i|, |D_j|)} \quad (26)$$

可以使用兴趣点的重复比率对兴趣点检测算法的性能进行评价。

### 4 实验结果

实验讨论了 UFL-ID 算法的特征数  $k$  对兴趣点数目、重复率和速率的影响;同时为了横向比较不同兴趣点算法的性能,将 Harris, SUSAN, FAST12, FAST9, DOG 与本文提出的 UFL-ID 兴趣点检测算法进行了比较。实验中算法所使用的参数如表 1 所列。所有检测算法的非极大值抑制窗口为  $3 \times 3$  像素(除 SUSAN 算法之外)。重复容许的误差距离  $\epsilon$  为 3 像素。

表 1 兴趣点检测算法使用的参数

SUSAN	
模板半径	3.5 像素
核值相似判定阈值	15 像素
Harris	
高斯窗方差	2
DOG	
高斯金字塔层数	3 像素
高斯金字塔尺度	3 像素
检测高斯方差	0.7

图 3 为 UFL-ID 算法在“lean”图像中提取的兴趣点效果图,可以看出,该算法成功标记出了图像中灰度值变化较大的区域。



图 3 UFL-ID 检测兴趣点效果图(特征数目  $k=40$ , 兴趣点数目为 500)

实验使用了“box”和“junk”<sup>[11]</sup>以及牛津大学 vgg 组<sup>[2]</sup>的数据集测试算法的重复性,数据集包含了若干张同一场景的不同角度的照片,这些照片之间较大的扭曲、旋转、形变、模糊、光照变化,可以模拟真实情况下图像的情况。

#### 4.1 评价特征

在重复性测试中,讨论了 UFL-ID 算法的特征提取数目  $k$  对重复率的影响,之后横向对比了不同兴趣点检测算法的重复率和抗噪能力。

##### 4.1.1 特征提取数目对重复性的影响

对于神经网络而言,确定隐含层节点数目十分复杂。输入或输出节点数目、训练样本数目、样本噪声、激活函数的选取等都会对最佳隐含层节点数目产生影响。对于大多数情况而言,在理论上没有可行的方法用来确定最佳隐含层节点数目<sup>[16]</sup>。如果隐含层节点数目过少,会因欠拟合而造成较大的训练误差;如果隐含层节点数目过多,就会因过拟合而造成网

<sup>1)</sup> <http://fastcorner.wordpress.com/2010/09/09/fast-faq.html>

<sup>2)</sup> <http://www.robots.ox.ac.uk/~vgg/research/affine/index.html>

络的泛化性能降低<sup>[17]</sup>。

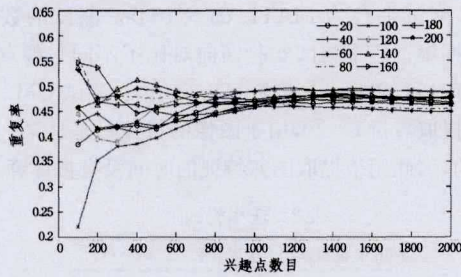


图4 特征提取数目对重复率的影响(测试数据集为 box 数据集)

在稀疏自编码器中,隐含层的数目决定了特征选取的数目,进而对兴趣点提取的质量产生影响。由图4可以看出,从40开始,隐含层数目的增加对兴趣点重复率的影响开始变小,这是由于兴趣点是较低级的特征,对图像的描述是较为底层的,因此只需要少量的基就可以将其线性表示。如图5所示,当特征提取数目为60时,稀疏自编码器提取了很多

彼此较相似的特征,这些特征对计算兴趣点响应值的影响是相近的。

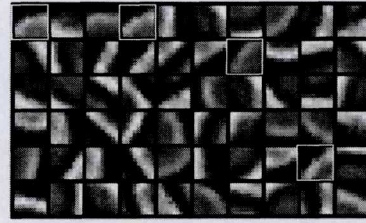


图5 特征提取数目  $k=60$  时稀疏自编码器提取的特征(其中与第一个特征相似的特征用白色框标出)

#### 4.1.2 算法重复性比较

为了更客观地对 UFL-ID 算法的性能做出评价,本文将 Harris, SUSAN, FAST12, FAST9, DOG 与 UFL-ID 算法进行了比较,如图6所示,可以看出各算法的平均重复率在兴趣点提取数目达到500时趋于稳定。

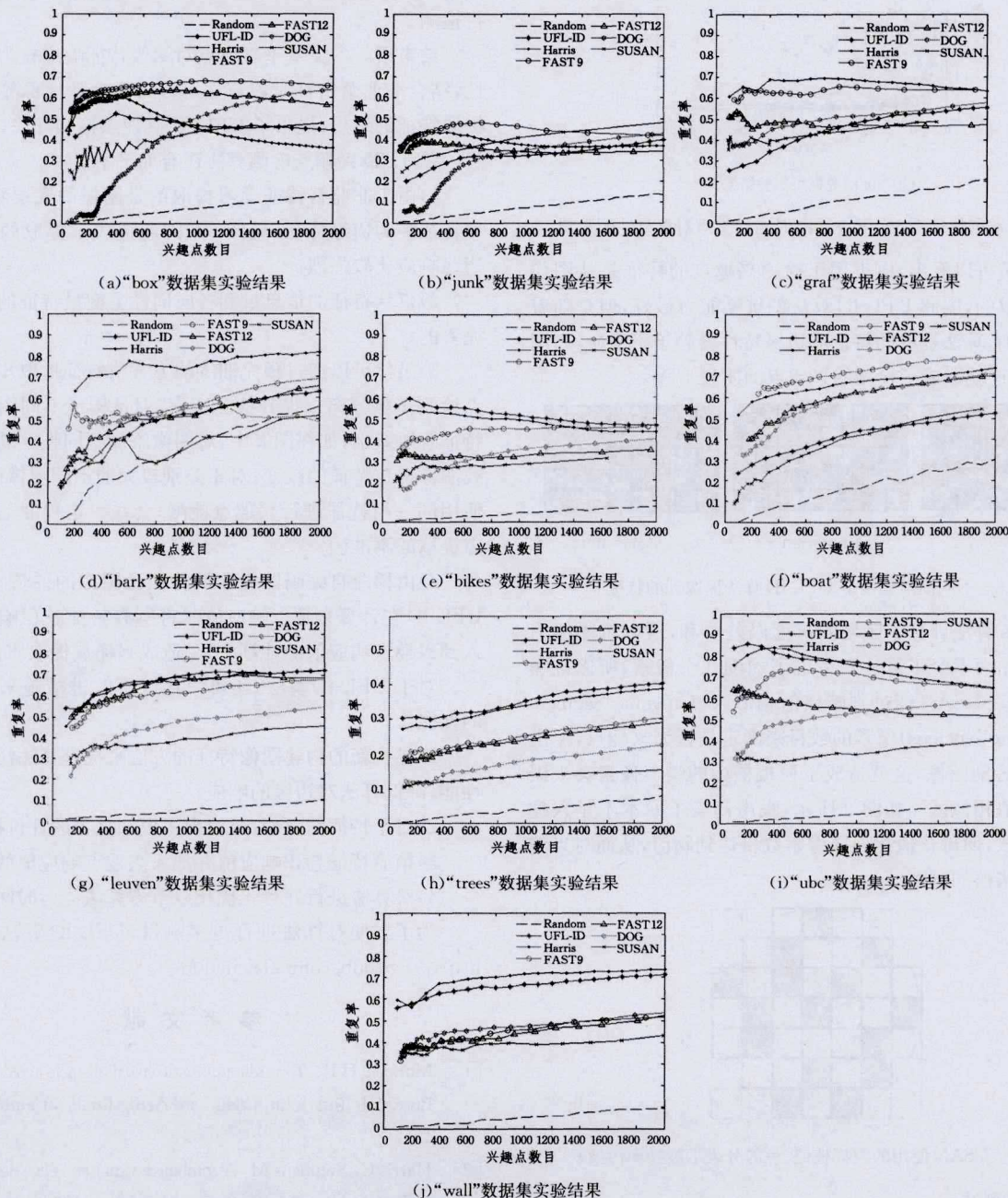
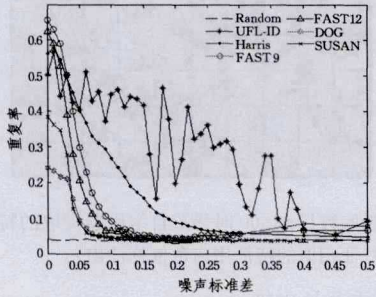


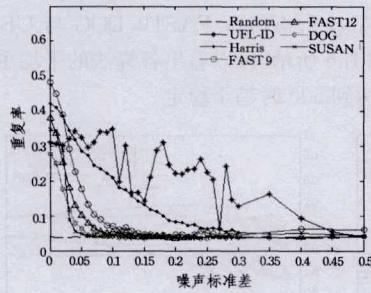
图6 算法重复性随兴趣点提取个数的变化(测试各数据集时,UFL-ID算法的特征提取数目  $k=40$ )

### 4.1.3 图像噪声对重复性的影响

为了测试兴趣点检测算法的抗噪能力,实验还对图像添加高斯噪声,结果如图 7 所示。



(a)“box”数据集实验结果



(b)“junk”数据集实验结果

图 7 兴趣点数量为 500 且  $k=40$  时高斯噪声对重复性的影响

由图 7 可以看出,由于用于检测兴趣点的特征是从图像中自主提取的,因此 UFL-ID 算法的抗噪能力最好,但是随着噪声的增加,算法提取的特征的边缘特性会趋于不明显,如图 8(c)所示,这就造成了检测的兴趣点的质量下降。

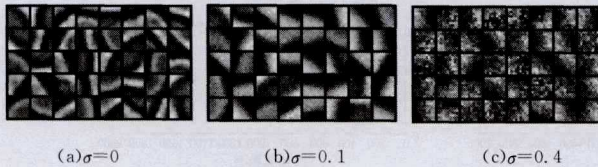


图 8 不同高斯噪声下 UFL-ID 算法提取到的特征

Harris 算法计算了区域的灰度误差总和,所以抗噪能力也较出色。而 FAST 算法只考虑了周围几个像素,所以抗噪能力较差。SUSAN 通过判断核值相似区(Univalve Segment Assimilating Nucleus)像素的数目来确定是否为兴趣点,但由于噪声会污染图像,这就造成了模板核内的噪声像素会被误判属于核值相似区。如图 9 所示,噪声污染了原本不是兴趣点的模板核,使得核值相似区像素数目达到阈值,从而对兴趣点造成了错误的判断。

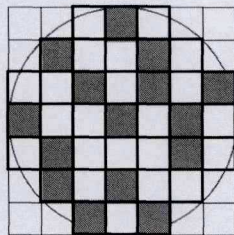


图 9 SUSAN 使用的模板核(灰色部分表示被噪声污染)

### 4.2 速率测试

除了探讨算法的重复率,实验还测试了算法运行所用的

时间,测试计算机 CPU 为 Inter (R) Core (TM) i7-4790 3.6GHz,测试平台为 MATLAB R2014a。测试的数据集为“box”数据集。速率测试实验横向对比了不同兴趣点检测算法的运行时间,兴趣点提取数目为 500。在测试 UFL-ID 算法时,特征提取数目  $k=40$ ,由于图像的特征只是从第一幅图像中提取的,因此用于提取图像特征的时间没有被计算。

表 2 速率测试结果

box 数据集(14 幅图片)	时间(s)
UFL-ID 算法	17.875
Harris 算法	33.940
FAST12 算法	21.317
FAST9 算法	14.227
DOG 算法	32.756
SUSAN 算法	80.688

从表 2 可以看出,FAST 算法的速率最快,本文提出的算法次之,这说明本文的算法可以高效地对图像中的兴趣点进行检测。

**结束语** 本文基于兴趣点对图像特征响应较为强烈的特性,结合非监督特征学习算法可以自主地从无标签的样本中提取特征的思想,提出了 UFL-ID 兴趣点检测算法。

UFL-ID 兴趣点检测算法还有如下不足:

1)由于非监督特征学习提取的是图像的底层特征,而这些特征多为边缘特征(见图 5),因此造成了提取的兴趣点的边缘响应比较强烈;

2)仅从特征的信息量和各向同性去衡量特征的质量是不完善的;

3)虽然可以通过预先训练的方式将特征提取步骤与兴趣点检测步骤分离以提升检测速率,但是构成不同图像的底层特征有所差别,如在图 8 中,对图像添加不同强度的噪声会导致图像稀疏特征的改变;对于差别较大的不同图像而言,如果使用同一组特征进行兴趣点检测,会在一定程度上造成兴趣点提取的不准确;

4)由稀疏自编码器提取的特征为图像的稀疏特征,但是 UFL-ID 在计算兴趣点响应值时将所有特征卷积响应值都加入到兴趣点响应值的计算中,会造成兴趣点提取质量的下降。

对于 UFL-ID 算法,还可以用以下改进措施来提升算法的性能:

1)寻找新的构建图像特征的方法来提升了算法泛化能力和性能,降低算法对边缘的响应;

2)对于特征的评价,可以引入更加完善的评价措施;

3)精简特征卷积响应值并加入兴趣点响应值的计算;

4)对算法进行并行化优化以节省算法运行时间。

为了方便对算法进行更多测试,UFL-ID 算法源码已在 <https://github.com/Zlalien> 开放。

### 参考文献

- [1] Moravec H P. Towards automatic visual obstacle avoidance [C] // Proc. 5th Int. Joint Conf. on Artif. Intell., Cambridge: MA, 1977
- [2] Harris C, Stephens M. A combined corner and edge detector [C] // Proc. 4th Alvey Vision Conf. . Mancheste, 1988; 147-151

(下转封三)

- [3] Pal N R, Pal S K. A review on image segmentation techniques [J]. *Pattern Recognition*, 1993, 26(93): 1277-1294
- [4] Lin Zheng-chun, Wang Zhi-yan, Zhang Yan-qing. Optimal Evolution Algorithm for Image Thresholding[J]. *Journal of Computer-Aided Design & Computer Graphics*, 2010, 22(7): 1201-1206 (in Chinese)  
林正春, 王知衍, 张艳青. 最优进化图像阈值分割法[J]. *计算机辅助设计与图形学学报*, 2010, 22(7): 1201-1206
- [5] Riquelme M T, Barreiro P, Ruiz-Altisent M, et al. Olive classification according to external damage using image analysis[J]. *Journal of Food Engineering*, 2008, 87(3): 371-379
- [6] Rahimi S, Zargham M, Thakre A, et al. A Parallel Fuzzy C-Mean algorithm for Image Segmentation[C]// *Processing Nafips'04*. 2004: 234-237
- [7] Otsu N. A threshold selection method from gray-level histograms[J]. *IEEE Transactions on System Man and Cybernetic*, 1979, 9(1): 62-66
- [8] Li C, Kao C Y, Gore J C, et al. Minimization of region-scalable fitting energy for image segmentation[J]. *IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society*, 2008, 17(10): 1940-1949
- [9] Felzenszwalb P F, Huttenlocher D P. Efficient Graph-Based Image Segmentation[J]. *International Journal of Computer Vision*, 2004, 59(2): 167-181
- [10] Liu T, Xie J B, Yan W, et al. An algorithm for finger-vein segmentation based on modified repeated line tracking[J]. *Imaging Science Journal*, 2013, 61(6): 491-502
- [11] Pham V H, Lee B R. An image segmentation approach for fruit defect detection using k-means clustering and graph-based algorithm[J]. *Vietnam Journal of Computer Science*, 2014, 2(1): 25-33
- [12] Brodi D. Text Line Segmentation With Water Flow Algorithm Based on Power Function[J]. *Journal of Electrical Engineering*, 2015, 66(3): 132-141
- [13] Wang X H, Yi-Gang H E, Zeng Z Z. Optimized Design of the Type-four FIR Filter Based on Neural Networks with Sine Basis Functions[J]. *Journal of Circuits & Systems*, 2003, 8(5): 97-100
- [14] Jia Xu, Sun Fu-ming, Cao Yu-dong, et al. Dorsal Hand Vein Recognition Algorithm Based on Effective Dimensional Feature [J]. *Computer Science*, 2016, 43(1): 315-318 (in Chinese)  
贾旭, 孙福明, 曹玉东, 等. 基于有效维度特征的手背静脉识别算法[J]. *计算机科学*, 2016, 43(1): 315-318
- [15] He J H. A New Iteration Method For Solving Algebraic Equations[J]. *Applied Mathematics & Computation*, 2003, 135(1): 81-84
- [16] Maji P, Roy S. Rough-Fuzzy Clustering and Unsupervised Feature Selection for Wavelet Based MR Image Segmentation[J]. *PLoS One*, 2015, 10(4): e123677
- [17] <http://www4.comp.polyu.edu.hk/~biometrics/MultispectralPalmprint/MSP.htm>
- [18] Hou Zhi-xu, Zhang Jian-xun. A method of Color Image Segmentation Used in Obstacle Recognition[J]. *Journal of Chongqing University of Technology(Natural Science)*, 2016, 30(3): 94-111 (in Chinese)  
侯之旭, 张建勋. 一种彩色图像分割的障碍物识别方法[J]. *重庆理工大学学报(自然科学)*, 2016, 30(3): 94-111

(上接第 294 页)

- [3] Smith S M, Brady J M. Susan—a new approach to low level image processing[J]. *International Journal of Computer Vision*, 1997, 23(1): 45-78
- [4] Zhang X, Wang H, Hong M, et al. Robust image corner detection based on scale evolution difference of planar curves[J]. *Pattern Recognit. Lett.*, 2009, 30(1): 449-455
- [5] Luo Xiao-hui, Li Jian-wei. DOG Model-Based Detector of Corner [J]. *Computer Engineering and Applications*, 2003(11): 87-99 (in Chinese)  
罗晓晖, 李见为. 双高斯差模型用于角点检测研究[J]. *计算机工程与应用*, 2003(11): 87-99
- [6] Rosten E, Porter R, Drummond T. Faster and better, a machine learning approach to corner detection[J]. *IEEE Trans. Pattern. Anal. Mach. Intell.*, 2010, 32(1): 105-119
- [7] Dias P G T, Kassim A A, Srinivasan V. A neural network based corner detection method[C]// *Proceedings of the IEEE International Conference on Neural Networks (ICNN '95)*. IEEE, 1995: 2116-2120
- [8] Sun Wei, Yang Xuan. Image corner detection using topology learning[J]. *Journal of China Universities of Posts and Telecommunications*, 2010, 17(6): 101-105
- [9] Prudent Y, Ennaji A. An Incremental Growing Neural Gas Learns Topologies[C]// *Proceedings of International Joint Conference on Neural Networks*. 2005(2): 1211-1216
- [10] Goodfellow I, Le Q, Saxe A, et al. Measuring invariances in deep networks[J]. *Neural Information Processing Systems*, 2009, 22: 646-654
- [11] Wagner R, Thom M, Schweiger R, et al. Learning Convolutional Neural Networks From Few Samples[C]// *International Joint Conference on Neural Networks (IJCNN)*. Dallas, Texas, USA, 2013: 1-7
- [12] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. *Neurocomputing; Foundations of Research*, 1988, 323(6088): 696-699
- [13] Olshausen B A, Field D J. Sparse coding of sensory inputs[J]. *Current Opinion in Neurobiology*, 2004, 14(4): 481-487
- [14] Mairal J, Bach F, Ponce J, et al. Supervised dictionary learning [C]// *Proceedings of the Advances in Neural Information Processing Systems (NIPS)*. 2009: 1033-1040
- [15] Schmid C, Mohr R, Bauckhage C. Evaluation of interest point detectors[J]. *International Journal of Computer Vision*, 2000, 37(2): 151-172
- [16] Zanchettin C, Ludermir Teresa B, Almeida Leandro M. Hybrid Training Method for MLP: Optimization of Architecture and Training[J]. *IEEE Transactions on Systems Man And Cybernetics Part B-Cybernetics*, 2011, 41(4): 1097-1109
- [17] Sun Jian-ye. Learning algorithm and hidden node selection scheme for local coupled feedforward neural network classifier [J]. *Neurocomputing*, 2012, 79(3): 158-163