

# 改进卷积自编码器的局部特征描述算法

贾琪, 王晓丹, 周来恩, 翟夕阳

JIA Qi, WANG Xiaodan, ZHOU Laien, ZHAI Xiyang

空军工程大学 防空反导学院, 西安 710051

Air and Missile Defense College, Air Force Engineering University, Xi'an 710051, China

JIA Qi, WANG Xiaodan, ZHOU Laien, et al. New local feature description algorithm based on improved convolutional auto-encode. *Computer Engineering and Applications*, 2017, 53(19): 184-191.

**Abstract:** To solve the problem that low-level features extracted by unsupervised learning methods are easily disturbed by image's rotation and scaling as well as difficult to distinguish when used in feature description, a local feature description algorithm is proposed based on improved Convolutional Auto-Encode (CAE-D). Evaluating the convolution kernel's performance by information entropy, a convolution kernel's entropy constraint rule is proposed to improve the distinguish ability of convolution feature description through convolution kernels carrying local information. Traditional SIFT's orientation assignment algorithm is used to assign the main direction of local image before feature description, and the feature-map is down-sampled to enhance rotation-invariance and robustness of the feature description. The results of image matching show that CAE-D is competitive with the performance of KAZE and SIFT descriptor in geometric and photometric deformations and takes 47.14% less time than SIFT.

**Key words:** unsupervised learning; feature description; convolutional auto-encoder; information entropy

**摘 要:** 针对非监督学习方法提取的底层特征用于特征描述时可区分性不强, 对图像旋转、尺度等变换敏感的问题, 提出了一种改进卷积自编码器的局部特征描述算法 (Convolutional Auto-Encoder Descriptor, CAE-D)。CAE-D 算法利用信息熵评价卷积核性能, 提出在 CAE 中添加卷积核信息熵约束规则, 通过均值化卷积核携带的局部特征信息, 提升卷积特征描述的可区分性; 在特征描述前使用传统 SIFT 中主方向分配算法确定局部图像的主方向, 并引入降采样操作, 进一步提升特征描述的旋转不变性及鲁棒性。图像匹配实验结果验证了改进策略的有效性, CAE-D 算法优于当前先进的 KAZE、SIFT, 而运行时间相比 SIFT 缩短了 47.14%。

**关键词:** 非监督学习; 特征描述; 卷积自编码器; 信息熵

**文献标志码:** A **中图分类号:** TP391.4 **doi:** 10.3778/j.issn.1002-8331.1612-0348

## 1 引言

图像特征提取作为计算机视觉中最重要的一环在近年来得到了足够的重视和大量的研究<sup>[1]</sup>。随着深度神经网络和机器学习的发展, 深度卷积神经网络 (Convolutional Neural Networks, CNN)<sup>[2-3]</sup> 等有监督特征学习方法已经取得了良好的效果, 但其性能依赖于标记训练数据的多少与质量好坏, 并且图像局部特征具有的尺度、光照、旋转等不变性、独特性难以手工标记。采用非监督特征学习方法可以自动地从无标记的图像数据中提取隐含信

息, 非常适合处理目前海量的图像、视频数据<sup>[4]</sup>, 以此为基础的方法在图像配准<sup>[5]</sup>、图像识别<sup>[6]</sup>、图像分类<sup>[7]</sup>等领域有着广泛应用。

堆叠的非监督学习网络可以提取图像特征的深层表示与数据内部的隐含信息。文献[8]在卷积受限玻尔兹曼机 (Convolutional RBM, CRBM) 的基础上构建卷积深度信念网络, 使用改进的概率训练函数, 在 MNIST 数据上取得了较高的分类准确率; 文献[9]使用堆叠的卷积自编码器 (Convolutional Auto-Encode, CAE) 作为

**基金项目:** 国家自然科学基金 (No.61273275); 航空科学基金 (No.20151996015)。

**作者简介:** 贾琪 (1993—), 男, 硕士研究生, 研究领域为智能信息处理、计算机视觉, E-mail: jiaqijiaqi1314@126.com; 王晓丹 (1966—), 女, 博士, 教授, 研究领域为智能信息处理、机器学习; 周来恩 (1991—), 男, 硕士研究生, 研究领域为机器学习; 翟夕阳 (1994—), 男, 硕士研究生, 研究领域为智能信息处理。

**收稿日期:** 2016-12-22 **修回日期:** 2017-02-24 **文章编号:** 1002-8331(2017)19-0184-08

CNN网络的预训练方法,取得了比传统CNN网络更好的实验结果;文献[10]组合多层去噪自编码器(Denoising Auto-Encode, DAE),逐层训练网络提取图像深层特征,结合SVM有效实现了图像分类。

然而,图像局部特征属于底层特征,直接利用非监督学习方法进行提取,得到的抽象特征内部差异较大且对图像变换敏感,在实际应用中难以取得较好的效果。为此,文献[11]使用单层非监督学习网络提取图像底层特征,详细分析了不同因素(数据预处理方法、中间隐层节点数量和卷积步长等)对提取底层特征的影响,并在图像分类中尝试不同参数的组合,提升了分类准确率;文献[12]建立了一种基于CAE的图像分类层次模型:特征提取层、卷积层、降采样层、分类层,详细分析了数据白化程度与降采样方法对模型性能的影响,并提出了一种改进的自适应降采样方法;文献[13]利用多尺度卷积核提取图像在不同尺度下的底层特征,并将特征描述与带有方向信息的HOG(Histograms of Oriented Gradients)特征融合,在人脸识别中取得了较高的识别率,但也存在一些不足,采用多尺度的卷积核未能解决卷积核自身对图像变换敏感的问题,融合HOG特征增大了特征描述的维数,提升了算法运行时间;文献[14]提出CKN(Convolutional Kernel Networks)用于图像分类:先使用核函数方法计算像素级特征,再将像素级特征作为输入,利用非监督学习网络提取图像特征,核函数极大地增加了所提取特征的可区分性,同时对多种图像变换具有一定的鲁棒性,但也存在一些不足,核函数运算极大地增加算法运行时间,不适用于对实时性要求高的场合。

针对以上不足,本文提出了一种改进卷积自编码器的局部特征描述算法(CAE-Descriptor, CAE-D),将卷积核与预处理后的局部图像卷积产生的特征响应作为特征描述,称为卷积特征响应,并提出两点改进策略:(1)利用信息熵评价卷积核性能,提出了一种卷积核信息熵约束规则,通过约束卷积核信息熵范围,防止部分卷积核携带过多或不携带局部特征信息,提升卷积特征描述的可区分性。(2)在局部图像预处理中,利用SIFT中性能较好的主方向分配算法给局部图像分配主方向,相比文献[13]直接融合HOG特征,节省算法运行时间的同时增加了准确性。CAE-D算法特征提取时,随机选取若干局部图像块并归一化处理,将图像块输入改进的CAE中训练,得到若干含有底层特征信息的卷积核;特征描述时,首先使用主方向分配算法给待描述局部图像分配主方向,并根据主方向信息旋转局部图像,然后使用卷积核分别卷积局部图像得到特征图,并进行降采样操作,最后,将降采样后的特征图按顺序展开融合后得到卷积特征描述,作为该局部图像的特征描述。图像匹配实验结果表明,CAE-D运行时间相比SIFT更快,其生成的特征描述具有较强的可区分性与一定的鲁棒性,取得

了与基于分布的描述算法相当的匹配效果,对含有结构信息的数据集图像,CAE-D性能更优。

## 2 CAE-D特征提取

### 2.1 自动编码器

自动编码器(Auto-Encode, AE)<sup>[15]</sup>,是用于数据降维和特征提取的神经网络。AE由两层神经网络组成:第一层为编码层,由输入层和中间隐层节点组成;第二层为解码层,由中间隐层和输出层节点组成。编码层输入的节点个数与解码层输出节点的个数相等,目的是通过学习一个恒等函数,能使输入等于输出,从而寻找出原始数据之间隐藏的关联结构。设原始数据为 $x$ 、输出数据为 $\hat{x}$ ,映射函数为 $f_{\text{enc}}$ 和 $f_{\text{dec}}$ ,通过学习不断调整输入层与中间隐层和中间隐层与输出层的连接权重,使得 $f_{\text{dec}}(f_{\text{enc}}(x)) = \hat{x} \approx x$ 。AE网络结构如图1所示。

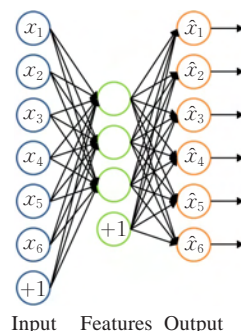


图1 AE网络结构

假设某AE网络由输入层 $x = \{x_1, x_2, \dots, x_m\}$ , 中间隐层 $h = \{h_1, h_2, \dots, h_k\}$ 和输出层 $\hat{x} = \{\hat{x}_1, \hat{x}_2, \dots, \hat{x}_m\}$ 组成。输入层、中间隐层和输出层的节点个数分别为 $m, k, m$ 个。输入层与输出层的节点数目相等,  $\hat{x}$ 代表了中间隐层对输入层数据的重构。相邻两层节点之间的连接权重为 $W^{(1)} = \{w_{11}^{(1)}, w_{12}^{(1)}, \dots, w_{km}^{(1)}\}$ 与 $W^{(2)} = \{w_{11}^{(2)}, w_{12}^{(2)}, \dots, w_{km}^{(2)}\}$ , 其中 $w_{ij}^{(1)}$ 表示输入层节点 $i$ 与中间隐层节点 $j$ 之间的连接权重; 中间隐层和输出层的偏置项分别为 $b^{(1)} = \{b_1^{(1)}, b_2^{(1)}, \dots, b_k^{(1)}\}$ 与 $b^{(2)} = \{b_1^{(2)}, b_2^{(2)}, \dots, b_m^{(2)}\}$ , 其中 $b_j^{(1)}$ 表示中间隐层节点 $j$ 的偏置。中间隐层第 $j$ 个节点的输入值 $h_j$ 定义如公式(1)所示, 经中间隐层后, 输出值 $\hat{x}$ 定义如公式(2)所示。其中,  $a_j$ 为中间隐层输出值, 有 $a_j = f(h_j)$ , 使用Sigmoid激活函数时 $f(h) = \frac{1}{1 + e^{-h}}$ 。

$$h_j = \sum_{i=1}^m w_{ji}^{(1)} x_i + b_j^{(1)} \quad (1)$$

$$\hat{x}_i = f\left(\sum_{j=1}^k w_{ij}^{(2)} a_j + b_i^{(2)}\right) \quad (2)$$

训练AE的目标是学习一个拟合函数,常用的方法是MLP的反向传播算法,通过梯度下降法调整权重参数 $W, b$ 来最小化重构误差。重构误差是指输入数据 $x$

与输出数据  $\hat{x}$  之间的误差  $\Gamma(W, b)$ , 图像数据属于连续数据, 故使用方差误差作为误差评估标准。方法误差如公式(3)所示:

$$\Gamma(W, b) = \frac{1}{2t} \sum_{i=1}^t \|\hat{x}^{(i)} - x^{(i)}\|^2 + \lambda \|W\|^2 \quad (3)$$

在方差误差中, 第二项是权重衰减项, 可以减少模型过拟合, 同时减少权重的值,  $\lambda$  表示权重衰减参数。对于灰度图像, 重构误差通常使用平方差估计。

## 2.2 卷积自动编码器

作为一种非监督学习方法, 卷积自编码器(Convolutional Auto-Encode, CAE)通过卷积核能有效地提取图像局部特征, 其核心思想是在自动编码器(Auto-Encode, AE)<sup>[15]</sup>的基础上引入卷积操作, 可直接处理二维图像, 假设某CAE有  $H$  个卷积核, 第  $k$  个卷积核为  $H_k$ , 偏置为  $b_k, c_k$ , 其余参数与AE相同。对输入数据  $x$ , 第  $k$  个卷积核的输入及解码层的重构数据  $\hat{x}$  分别如公式(4)、(5)所示:

$$h^k = f(x * W^k + b^k) \quad (4)$$

$$\hat{x} = f\left(\sum_{j=1}^H h^j * \overline{W}^j + c^j\right) \quad (5)$$

其中  $\overline{W}^k$  表示第  $k$  个特征图的权矩阵  $W^k$  的转置,  $*$  符号表示二维卷积运算。CAE通过梯度下降法调整权重参数最小化重构误差时, 目标函数梯度值计算方法如公式(6)所示, 其中,  $\delta h$  和  $\delta \hat{x}$  分别表示卷积层和解码层残差, 由于跟本文主要内容无关, 其求解方法在此不再累述。

$$\frac{\partial E(W, b)}{\partial W^k} = x * \delta h^k + \tilde{h}^k * \delta \hat{x} \quad (6)$$

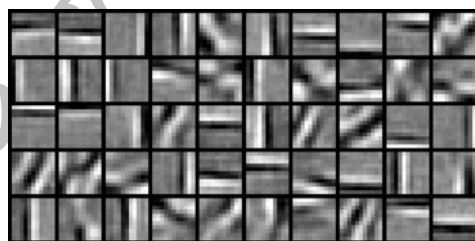
## 2.3 基于信息熵的卷积核评价

从信息论理论得知一幅图像所能表达的信息量可以用信息熵表示, 但存储在计算机上的图像常常被离散化, 为便于对这些离散的图像进行分析处理, 利用图像的相位和振幅分布函数来建立图信息的数学模型<sup>[16]</sup>。设某一灰度图像按灰度值大小划分为  $L$  个等级, 其中第  $i$  ( $i=1, 2, \dots, L$ ) 级灰度出现的概率为  $p_i$ , 则其所带的信息熵  $H$  如公式(7)所示:

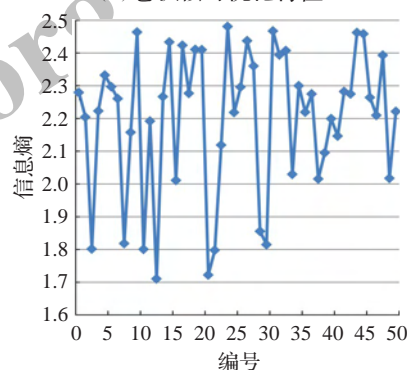
$$H = - \sum_{i=1}^{K-1} p_i \times \ln\left(\frac{1}{p_i}\right) \quad (7)$$

利用基础的CAE提取图像局部特征, 设置灰度值划分  $L$  为20, 卷积核数目  $H$  为50, 卷积核大小为  $11 \times 11$  像素, 迭代次数  $N$  为600, 按照文献[5]中可视化特征的方法, 得到卷积核的可视化特征与信息熵如图2所示。其对应顺序为, 图2(b)编号从小到大依次为图2(a)中从左往右、从上到下的卷积核可视化特征。

通过图2(a)、(b)可知, 信息熵能有效地反映卷积核包含的图像结构信息, 并且当CAE网络接近收敛时, 卷积核的信息熵越大, 包含的结构信息越丰富, 反之则结构信息较少。



(a) 卷积核可视化特征



(b) 各卷积核信息熵取值

图2 CAE训练结果

## 2.4 卷积核信息熵约束规则与重置方法

当卷积核内部的信息熵差异较大时, 训练网络往往存在过拟合<sup>[17]</sup>, 且实验表明, 此时生成的卷积特征描述可区分性较差。针对此不足, 提出在CAE中添加卷积核约束规则, 通过约束规则, 均值化各卷积核信息熵的取值, 防止过拟合的发生的同时提升卷积特征描述的可区分性。其中, 对不符合要求的卷积核, 先进行重置操作, 而后继续迭代直至收敛, 从而获取新的底层特征信息。

约束规则与重置方法的核心思想类似于化学反应中的酸碱中和过程。当卷积核数目一定时, 利用CAE网络提取底层特征, 其含有的底层特征信息量恒定(酸碱平衡时, PH值恒定), 但往往会存在一些卷积核携带过多(酸性)、过少(碱性)的底层信息, 表现为信息熵值过大、过小, 从而减少了卷积特征描述的可区分性。为此, 需要采用重置操作, 重置操作将此类卷积核替换成不携带底层信息的卷积核(移除强酸、碱性物质, 加入弱酸、碱性物质), 而后继续训练CAE网络(搅拌, 促进稳定), 通过反馈调节, 被替换的卷积核将重新携带一定的底层特征信息(重新达到PH值恒定), 但相比之前, 携带的过多、少的底层信息被均值化至其他卷积核中。接下来详细介绍约束规则与重置的实施方法。

卷积核信息熵约束规则以卷积核信息熵取值为评价指标, 方法为: 预先设置中断迭代的次数间隔, 当中断迭代后, 计算当前各卷积核信息熵取值, 判断是否有不符合要的卷积核, 若有则进行重置操作, 反之继续迭代直至下次中断或达到结束条件。约束规则的流程图如图3所示。此外, 综合考虑算法性能与运行时间, 取迭代间隔  $S$  为200, 终止条件  $T$  为600。



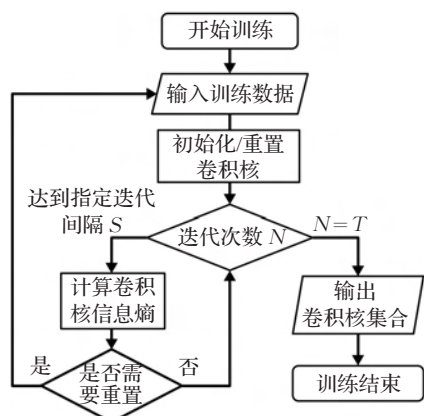


图3 卷积核信息熵约束规则流程图

卷积核重置操作的方法为:使用随机函数产生不携带任何底层特征信息的卷积核,利用该卷积核替换不符合要求的卷积核。其中,随机函数可将卷积核信息熵提升至较高水平,而同一训练数据下,卷积核信息熵总量有限,因此在后续迭代的过程中,重置后的卷积核信息熵将缓慢下降直至收敛,从而预防了过拟合的发生。重置卷积核过程如图4所示。

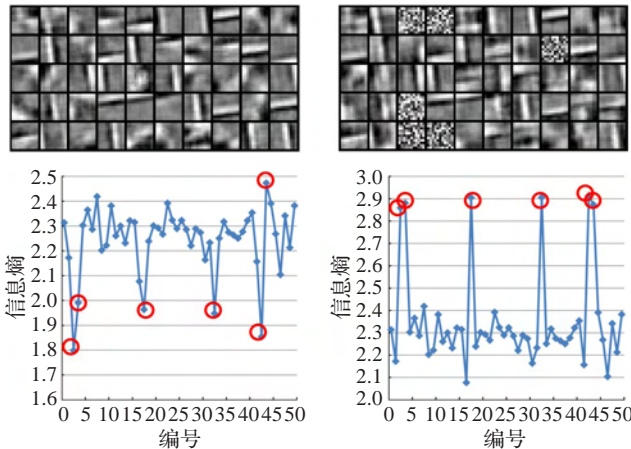
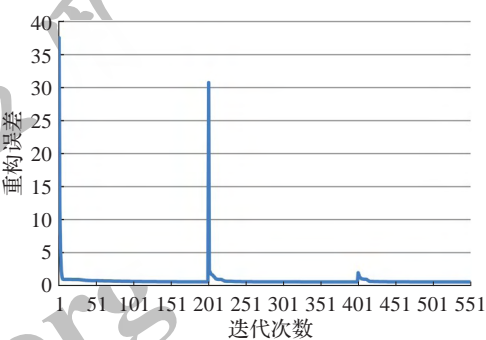


图4 重置卷积核过程(左图为重置前,右图为重置后)

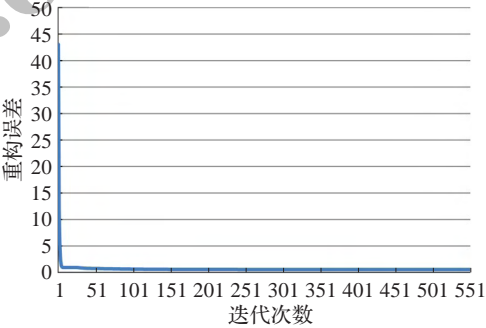
对于卷积核重置标准,文本经过大量实验的统计分析,当卷积核信息熵之间相差小于10%时,卷积特征描述性能较佳。为此,在CAE-D算法中,作为经验参数,卷积核重置阈值设为与平均信息熵取值相差大于10%。

为进一步证明约束规则的有效性性与可行性,分析CAE中添加卷积核信息熵约束规则前后的重构误差变化曲线,当卷积核重置后,CAE网络的重构误差显著上升,但随着迭代的进行,重构误差逐步收敛,表明被重置的卷积核重新获取了部分底层特征信息,同时未出现过拟合,误差曲线如图5所示。

添加约束规则后,得到的卷积核可视化特征与信息熵分布如图6所示,其对应关系与图2中相同。可以看出,相比图2未添加约束规则,添加约束规则后得到的各卷积核信息熵取值差别较小,表明此时各卷积核携带的底层特征具有相同信息量。

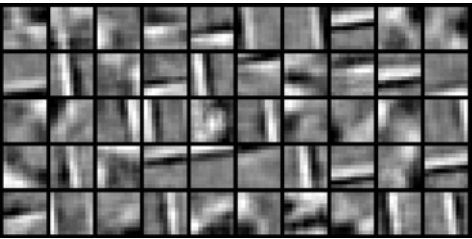


(a)添加卷积核信息熵约束规则

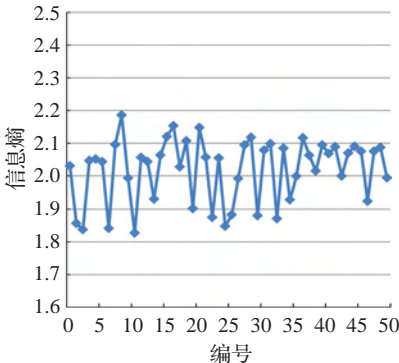


(b)未添加卷积核信息熵约束规则

图5 迭代过程中重构误差曲线



(a)卷积核可视化特征



(b)各卷积核信息熵取值

图6 CAE-D 训练结果

### 3 CAE-D特征描述

#### 3.1 SIFT主方向分配算法

使用SIFT算法<sup>[18]</sup>中特征点方向分配算法可以有效解决图像旋转变换干扰的问题,其步骤为:首先计算局部图像梯度的幅值和幅角,然后利用梯度直方图统计局部梯度信息,最后根据直方图统计结果分配一个或多个主方向。

幅值和幅角的计算方法如公式(8)、(9)所示。其中,  $m(x,y)$  表示图像梯度的幅值,  $\theta(x,y)$  表示图像梯度

的幅角,  $L$  的尺度是关键点所在的尺度。梯度直方图统计与方向分配过程如图 7 所示。此外, 待描述局部图像的预处理方法为: 当主方向分配完成后, 旋转带描述的局部图像至主方向, 供下一步特征描述使用。

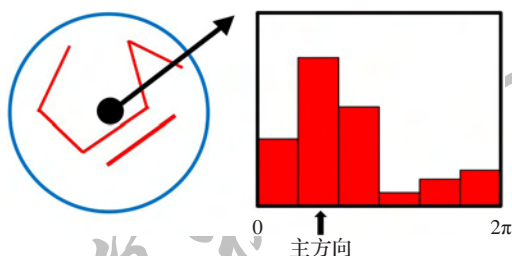


图7 SIFT算法特征点方向分配过程

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (8)$$

$$\theta(x, y) = \arctan\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \quad (9)$$

### 3.2 CAE-D 特征描述层次模型

不同于传统局部图像描述算法从待描述图像中提取特征信息, CAE-D 算法利用改进的 CAE 网络预学习得到的底层特征(卷积核)进行特征描述, 称为卷积特征描述。

为了更清晰地介绍整个算法流程, 提出了 CAE-D 特征描述层次模型, 模型共分为四层: 特征提取层、卷积层、降采样层、特征描述层。CAE-D 层次模型如图 8 所示。

模型中, 特征提取层主要提供具有强可区分性的底层特征, 即卷积核。假设待描述局部图像大小为  $n \times n$  像素, 卷积核大小为  $m \times m$  像素, 模型后三层的实施方法如下: 卷积层使用卷积核卷积局部图像得到特征图, 卷积不考虑图像边界, 并将得到的特征图编号, 特征图反映了待描述局部图像包含底层特征的情况, 其大小为  $(n-m+1) \times (n-m+1)$  像素; 降采样层使用预先设定的大小为  $d \times d$  像素的采样窗口, 对特征图做不重叠最大值降采样, 可以一定程度解决图像尺度变换干扰的问题, 降采样后的特征图大小为  $((n-m+1)/d) \times ((n-m+1)/d)$  像素; 特征描述层主要用于生成卷积特征描述, 先将采样后的二维特征图一维化展开, 而后根据卷积核编号依次写入至特征描述向量中。

CAE-D 算法流程描述如下。

#### 算法 1 CAE-D

输入: 训练图像 IMG、卷积核数目  $H$ 、卷积核大小  $m$ 、局部图像大小  $n$ 、降采样窗口大小  $d$ ;

输出: 局部图像特征描述。

步骤 1 准备训练数据集, 将训练图像 IMG 分割为一系列  $m \times m$  大小的像素块, 合记为  $patch$ , 并将  $patch$  中的像素块大小归一化到  $[0.1, 0.9]$ 。

步骤 2 建立卷积自编码器, 其中输入层和输出层节点个数均为  $m \times m$ , 卷积核个数为  $H$ 。

步骤 3 生成卷积核, 引入卷积核信息熵约束规则与重置方法, 利用  $patch$  训练卷积自编码器, 得到  $H$  个携带不同底层特征的卷积核。

步骤 4 分配特征点主方向, 利用 SIFT 算法中主方向分配算法计算特征点所在局部图像的主方向, 并根据主方向旋转局部图像。

步骤 5 获取特征图, 使用卷积核分别卷积局部图像得到特征图大小为  $(n-m+1) \times (n-m+1)$ , 降采样窗口大小为  $d \times d$ , 降采样后得到的特征图大小为  $((n-m+1)/d) \times ((n-m+1)/d)$ 。

步骤 6 生成特征描述, 将特征图按顺序展开, 融合得到卷积特征描述, 维度为  $H \times ((n-m+1)/d) \times ((n-m+1)/d)$ , 作为该局部图像的特征描述。

## 4 实验结果与分析

### 4.1 实验数据

图像匹配实验采用 Oxford 数据集 ([http://www.robots.ox.ac.uk/~pablo/face\\_3d/](http://www.robots.ox.ac.uk/~pablo/face_3d/)), Oxford 数据集由 8 组像素图片组成, 每组包含 6 张  $800 \times 640$  像素图片, 涵盖了图像视角、尺度与旋转、模糊、照度、JPEG 压缩等变换。此外, 数据集还包含每个场景下的第一张图像与其他所有图像的真实投影变换矩阵。

### 4.2 实验环境与评估准则

系统配置为: Win7 64x 系统, Intel® i7-4790 3.60 GHz 处理器, 8 GB 内存, GT750 Ti 显卡; 像匹配实验的软件环境为: 加载 OPENCV 2.9.1 开源库的 VS 2010 开发环境结合 MATLAB 2014a; 算法时间复杂度分析的软件环境为: MATLAB 2014a。

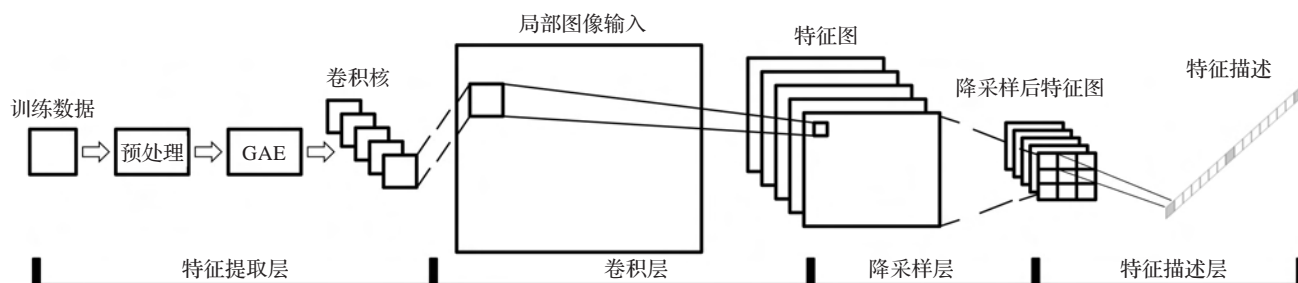


图8 CAE-D层次模型

图像匹配实验中,采用文献[19]的查错率(1-precision)和召回率(recall)评价指标,其计算方法如公式(10)、(11)所示。

错误率 = 1 - 
$$\frac{\text{正确的匹配数目}}{\text{总的匹配数目 (正确 + 错误匹配数目)}} \quad (10)$$

召回率 = 
$$\frac{\text{正确匹配的数目}}{\text{两幅图中一致区域的数目}} \quad (11)$$

选用最近邻距离法(Nearest Neighbor)作为匹配策略,通过数据集中提供的真实投影变换矩阵计算变换后的像素点位置,并规定当两幅图片中的匹配特征点对投影坐标误差小于两个像素时,认为该匹配对为正确匹配,否则为错误匹配。此外,根据各算法中最大、最小匹配误差,动态设置匹配门限做出召回率-查错率曲线图。

4.3 图像匹配实验结果

综合考虑性能与时间复杂度,各算法特征点数目设置为500个。CAE-D参数设置为:卷积核数目H为50,输入图像大小为28×28像素,卷积核大小为11×11像素,降采样窗口大小为6×6像素。此时,融合生成卷积特征向量的维数为((28-11+1)/6)×((28-11+1)/6)×50=450维。

首先,分析特征点主方向分配的预处理方法对算法性能的影响。CAE网络提取的卷积特征属于底层特征,基本不携带方向信息,因此,在实验中对包含旋转、视角变换的图片集,若未采用预处理过程,对局部图像分配主方向,将导致算法功能难以实现,召回率近乎为0。

接下来分析约束规则对算法性能的影响。取视角变换的wall图片集P1-P2(第一张图片与第二张图片匹配,下同)与图像尺度及旋转变换的boat图片集P1-P2进行图像匹配实验,特征点提取统一采用SIFT算法,结果如图9所示。其中CAE-D表示添加卷积核信息熵约束规则,CAE反之,显见,CAE-D算法综合性能更优,在相同查错率时均具有更高的召回率。

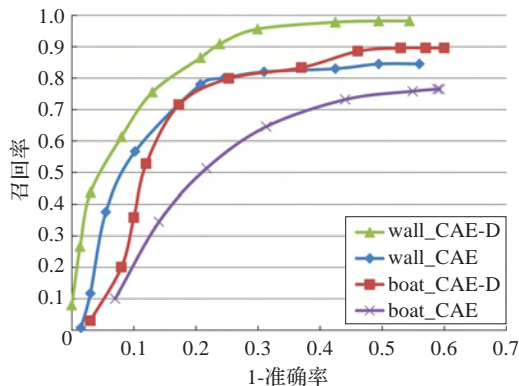


图9 wall,boat 图片集(P1-P2)添加约束规则前后实验结果

为更全面地对CAE-D性能做出评价,选取公认性能较好的SIFT<sup>[18]</sup>、SURF<sup>[20]</sup>、KAZE<sup>[21]</sup>与CAE-D在各图片集中进行对比实验。考虑算法原理不同,特征点位置会一定程度的影响算法性能,故实验中不使用统一的特征点提取算法,但为确保实验的公平性与可靠性,CAE-D算法分别使用各对比算法提取的特征点进行图像匹配

实验,依次表示为(SCAE-D(SIFT+CAE-D)、UCAE-D(SURF+CAE-D)、KCAE-D(KAZE+CAE-D)),实验结果如图10所示。

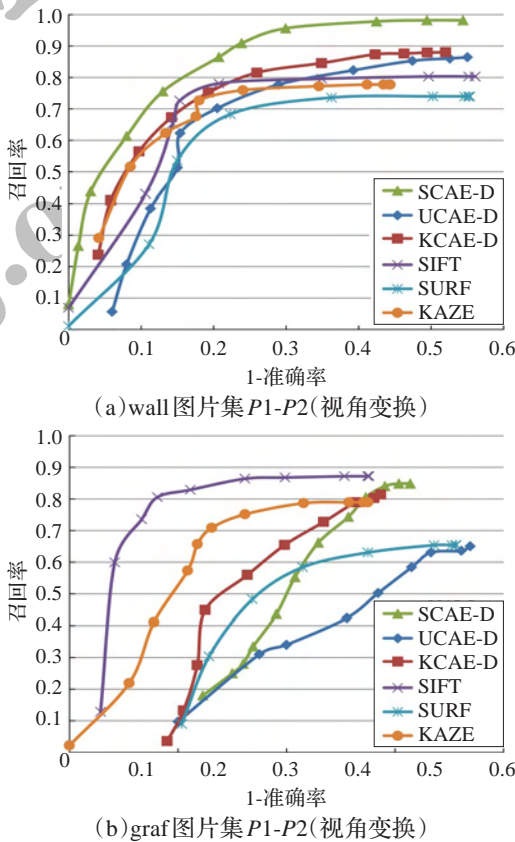
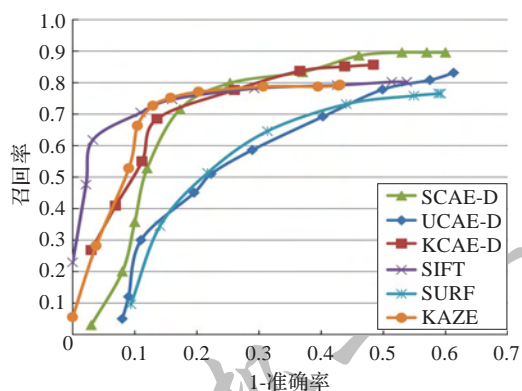


图10 wall,graf 图片集实验结果

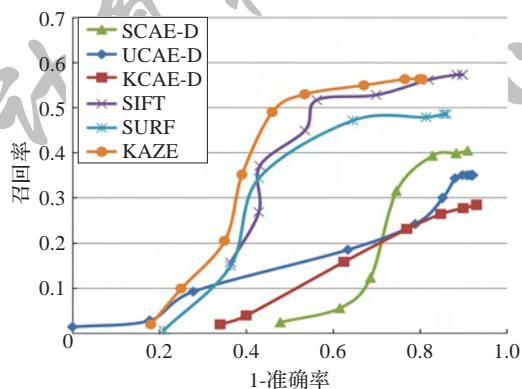
在视角变换实验中,使用数据集中的图片集wall与graf分别测试,其中,wall属于结构图像,graf属于纹理图像,结果如图10所示。可以看出,在结构图像的图片集wall中,表现最好的为SCAE-D,SCAE-D在查错率为33.42%时,召回率达到95.76%,其原因可能是CAE-D算法提取的底层特征属于结构特征,对包含结构信息的图像,可以生成可区分性较高的特征描述;在纹理图像的图片集graf中,CAE-D算法不尽人意,相比几种对比算法性能下降明显,SIFT算法表现最好,其原因可能是纹理图像中的纹理特征属于深层特征,利用底层特征进行描述时可区分性不强。

在尺度与旋转变换实验中,使用数据集中的图片集boat与bark分别测试,结果如图11所示。在图片集boat中,表现最好的为SCAE-D,其原因可能是采用的预处理方法,使CAE-D算法具有一定的旋转不变性,而降采样操作减少了尺度变换的影响;在图片集bark中,CAE-D性能下降显著,其原因可能是岩石图像中结构信息不丰富,CAE-D算法难以得到具有强可区分性的描述,查错率从低阈值开始便较高,而最终的召回率也不理想,而三种对比算法采用基于分布的特征,对特征点周围的细小变化具有较强描述能力,从而在光滑、低纹理图片中取得较好的效果。





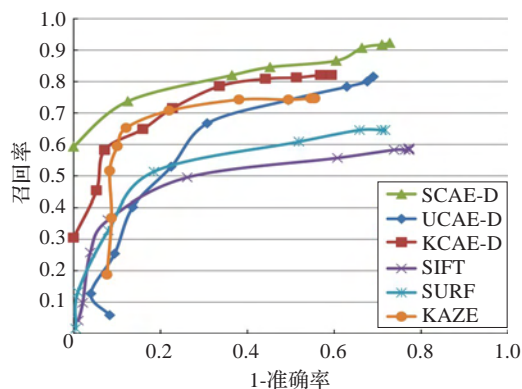
(a)boat 图片集 P1-P2(尺度与旋转变换)



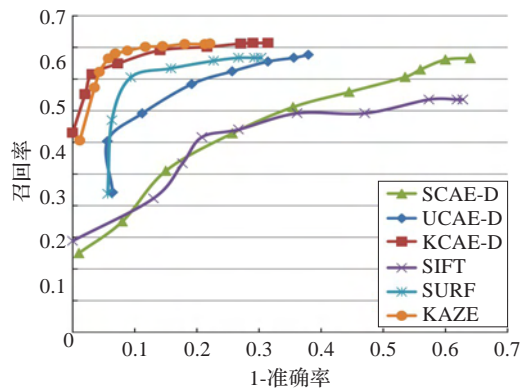
(b)bark 图片集 P1-P3(尺度与旋转变换)

图 11 boat、bark 图片集实验结果

在模糊变换实验中,使用数据集中的图片集trees 与 bikes 分别测试,结果如图 12 所示。可以看出,CAE-D 略



(a)trees 图片集 P1-P2(模糊变换)

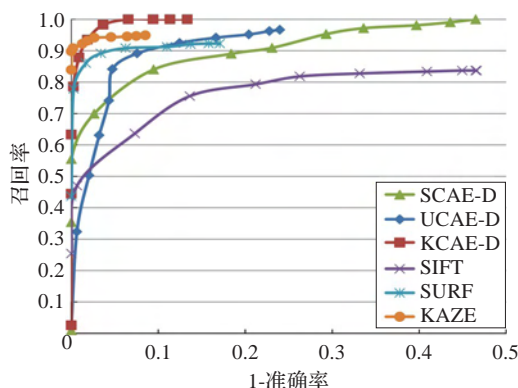


(b)bikes 图片集 P1-P2(模糊变换)

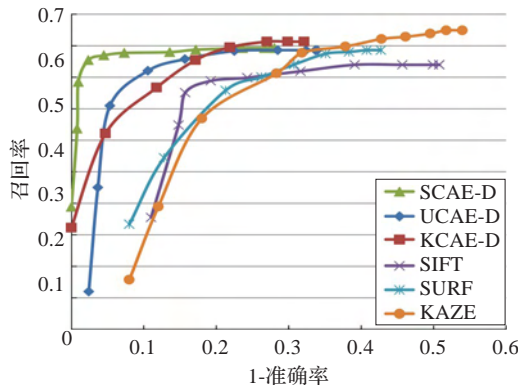
图 12 trees、bikes 图片集实验结果

优于对比算法,图片集trees 中,性能最好的为SCAE-D,其次为KAZE,SCAE-D在查错率为62%时,召回率超过90%,其余算法召回率均在85%以下;图片集bikes 中,性能最好的为KCAE-D与KAZE,两者均取得了较优异的表现。

在 JPEG 压缩变换及照度变换实验中,使用数据集中的图片集 ubc 与 leuven 分别测试,其中, ubc 包含了不同压缩率下的图片, leuven 通过调节相机的光圈获得线性照度的变换,结果如图 13 所示。在图片集 ubc 中,性能最好的为 SCAE-D 算法,其次为 KAZE 算法,表明 CAE-D 算法对图像压缩具有较强的鲁棒性;在图片集 leuven 中,性能前三名依次为 SCAE-D、UCAE-D 算法、KCAE-D 算法,说明 CAE-D 算法对图像照度变换不敏感。



(a)ubc 图片集 P1-P3(JPEG 压缩变换)

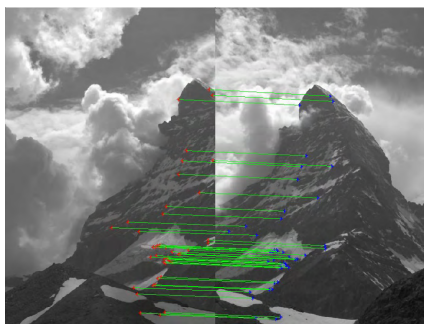


(b)leuven 图片集 P1-P2(照度变换)

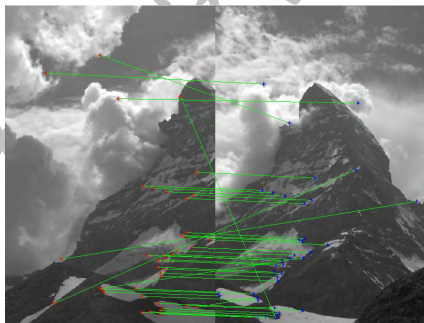
图 13 ubc、leuven 图片集实验结果

#### 4.4 图像匹配实例

图 14 给出了图像匹配结果的实例,匹配采用自然风景的图像,图 14(a)、(b)分别显示了 CAE-D 算法与 SIFT 算法的匹配结果,参数设置同上一节相同,这里只显示前 50 个特征匹配作为比较,通过 RANSAC(Random Sample Consensus)算法估计图像之间的投影变换矩阵并作为几何约束,从而判断误匹配。采用的自然风景图像中细节及结构信息比较丰富,但由于拍摄时间、地点不同,两者之间存在一定程度的视角变换与较大的背景变换,导致图像匹配难度较大。从结果可以看出 CAE-D 算法性能较好,前 50 个特征匹配对均为正确匹配,而 SIFT 算法的表现并不理想,其对图像细小结构变化较为敏感,从而可区分性不强导致查错率较高。



(a)CAE-D 算法:50/50 正确率



(b)SIFT 算法:42/50 正确率

图 14 图像匹配实例

4.5 算法运行时间对比

CAE-D 采用 MATLAB 语言编写,实验中采用的对比算法版本众多,而不同版本之间由于采用了不同的优化策略,运行时间差异较大。为保证客观性,选取 MATLAB 语言编写、不包含交叉编译的 SIFT 算法([https://github.com/tankebuaa/SIFT\\_MATLAB/](https://github.com/tankebuaa/SIFT_MATLAB/))进行对比实验。

实验选取图片集 wall 中第一张图片,利用 SIFT 算法提取 500 个特征点,随机抽取其中 300 个特征点进行特征描述与图像匹配。其中,SIFT、CAE-D 参数设置同匹配实验中保持一致,每个特征点只分配一个主方向。共划分为四个阶段:特征点方向分配、局部特征描述、图像匹配,每个阶段运行 10 次,取平均值为最终结果,结果如表 1 所示。

表 1 SIFT、CAE-D 运行时间 <sub>s</sub>		
算法	SIFT	CAE-D
特征点方向分配	0.939	0.939
局部特征描述	10.002	4.691
图像匹配	0.183	0.250
总运行时间	11.124	5.880

由表可知,CAE-D 算法在局部特征描述阶段具有较快的运行时间,但由于特征描述向量的维数相比 SIFT 高,因此在图像匹配时开销较大。总的看来,CAE-D 相比 SIFT 在运行时间上缩短了 47.14%。

5 总结

本文提出的改进卷积自编码器的局部特征描述算法 CAE-D,将非监督特征学习应用于特征描述中,通过添加改进策略的 CAE 提取图像的底层特征形成卷积

核,再利用卷积核卷积局部图像生成特征描述,具有较强的可区分性。其中,提出的卷积核信息熵约束规则,有效地将卷积核的信息熵差异约束在较小范围,提升了卷积核的描述能力;结合 SIFT 中的主方向分配算法与降采样操作,使生成的卷积特征描述对多种图像变换具有一定的鲁棒性。图像匹配实验结果表明,CAE-D 综合性能达到了当前最优的 KAZE、SIFT 算法,而运行时间相比 SIFT 减少近一半。但算法也存在一些不足与可以进一步改进的方向:

(1)CAE-D 中提取的特征属于底层特征,而底层特征多为图像的结构特征,导致非结构信息图像中生成的特征描述可区分性不强。接下来考虑提取图像的深层特征用于特征描述,合理融合深、浅层特征,进一步提升算法性能。

(2)CAE-D 中描述特征未考虑局部图像不同位置的权重关系,且生成的特征描述维数过高。考虑引入降维方法减少特征向量维数,同时根据信息熵取值赋予卷积核不同权重。

(3)实验中底层特征的提取通过预训练方式完成,但不同类型图像的底层特征有所差别,对差别较大的图像,CAE-D 算法性能不稳定。尝试建立底层特征数据库,根据卷积响应,自适应选取所需卷积核。

下一步,将对以上问题及改进方向进行研究,提出科学、合理的改进措施,进一步提升算法性能。

参考文献:

[1] 施鹏.局部特征描述子算法研究[D].上海:上海交通大学,2008.

[2] Krizhevsky A,Sutskever I,Hinton G E.ImageNet classification with deep convolutional neural networks[J].Advances in Neural Information Processing Systems,2012,25(2):1097-1105.

[3] Sun Y,Wang X,Tang X.Deep learning face representation from predicting 10,000 classes[C]//Proceedings of Conference on Computer Vision and Pattern Recognition,2014:1891-1898.

[4] 敖道敢.无监督特征学习结合神经网络应用于图像识别[D].广州:华南理工大学,2014.

[5] 周来恩,王晓丹.基于非监督特征学习的兴趣点检测算法[J].计算机科学,2016(9),43(9):289-294.

[6] Ranzato M,Huang F J,Boureau Y,et al.Unsupervised learning of invariant feature hierarchies with applications to object recognition[C]//Proceedings of Computer on Vision and Pattern Recognition,2007:1-8.

[7] Romero A,Gatta C,Camps-Valls G.Unsupervised deep feature extraction for remote sensing image classification[J].IEEE Transactions on Geoscience and Remote Sensing,2015,54(3):1-14.

(下转215页)