

基于图像灰度熵的自适应字典学习算法

杜秀丽 左思铭 邱少明

(大连大学通信与网络重点实验室 辽宁 大连 116622) (大连大学信息工程学院 辽宁 大连 116622)

摘要 针对传统图像稀疏表示字典学习算法仅对图像训练学习单一字典,不能很好地对包含不同图像信息的图像块进行最优稀疏表示的问题,将图像灰度熵的思想引入到字典学习算法中,提出基于图像灰度熵的自适应字典学习算法。该算法将图像库作为训练样本,对图像库图像进行分块,计算各子块的灰度熵大小,依据灰度熵大小对子块进行分类,针对不同类别子块,设定不同 K-奇异值分解算法参数,分别进行字典训练,从而得到多个不同的字典。根据灰度熵大小选择训练好的字典对待表示图像子块进行稀疏表示。仿真实验及结果表明,所提算法能够对图像进行较好的稀疏表示,图像的重构效果也得到了明显提升。

关键词 稀疏表示,字典学习,K-奇异值分解,灰度熵

中图分类号 TP391.4 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.05.041

Adaptive Dictionary Learning Algorithm Based on Image Gray Entropy

DU Xiu-li ZUO Si-ming QIU Shao-ming

(Key Laboratory of Communication and Network, Dalian University, Dalian, Liaoning 116622, China)

(College of Information Engineering, Dalian University, Dalian, Liaoning 116622, China)

Abstract Aiming at the problem that the traditional dictionary learning algorithm of image sparse representation only learns a single dictionary for image training, and can not optimally sparsely represent image blocks containing different image information, through introducing the local gray entropy of image into the dictionary learning algorithm, this paper proposed an adaptive dictionary learning algorithm based on image local gray entropy. The proposed algorithm makes use of the image database as training sample. Firstly, the image database is divided into blocks, and the gray entropy of each sub-block is calculated. Then, the sub-blocks are classified according to the size of the gray entropy, and different K-Singular Value Decomposition (K-SVD) parameters are set for different categories of sub-blocks to perform dictionary training respectively, thus obtaining a plurality of different dictionaries. Lastly, a well-trained dictionary is selected for the image sub-blocks to conduct sparse representation according to the size of the gray entropy. Simulation experiment results show that the proposed algorithm can sparsely represent the images better, and the effect of image reconstruction is also improved significantly.

Keywords Sparse representation, Dictionary learning, K-Singular value decomposition, Gray entropy

由 Candes 等^[1]和 Donoho^[2]提出的压缩感知(Compressed Sensing, CS)理论是一种新颖的信号采集理论,该理论突破了“先采集再压缩”的传统采样方式,将采样和压缩相结合,能够使用低于 Nyquist 采样定律所要求的采样率来采样,因此被广泛应用于无线传感器网络、医疗成像^[3]、遥感成像等领域。

压缩感知中,信号在字典下的表示系数越稀疏,重构质量就越高,因此对字典的选择极其重要^[4]。现有的字典构造方法主要有两种:人工构造和字典学习。人工构造方法通过使用事先定义好的数学变换来构造冗余字典,可通过快速傅里叶变换、离散余弦变换等数学变换及少量的参数来计算得到

字典中的原子。人工构造方法虽然相对简单,计算复杂度低,但由于得到的原子的基本形状固定、形态不够丰富,因此该方法不能对图像进行最优的稀疏表示。近年来,人们通过对数据或信号本身进行学习来获得完备字典,这种字典中的原子与训练集中的信号本身相适应,能更好地对原信号进行稀疏表示。早期提出的字典学习算法有主成分分析^[5]、广义主成分分析等。主成分分析由低维的子空间来逼近一组样本;而广义主成分分析用一个子空间来表示单个样本,从而利用多个子空间的联合来建模样本集合。该算法对字典的原子数目进行检测,当子空间的数目和维度过高时,运算复杂度会较大,从而导致计算量增加。随着研究的深入,许多新的字典学

到稿日期:2018-04-08 返修日期:2018-07-27 本文受高速眼图测试关键技术研究基金(L2014495),辽宁“百千万人才工程”培养经费资助。
杜秀丽(1977-),女,博士,教授,CCF 会员,主要研究方向为数字信号处理、通信技术,E-mail:22811623@qq.com(通信作者);左思铭(1993-),男,硕士,主要研究方向为数字信号处理;邱少明(1980-),男,硕士,副教授,主要研究方向为计算机技术与应用。

习理论被提出,包括多成分字典(Multi-component dictionary)^[6]、奇异值分解(Singular value decomposition, SVD)字典^[7]等。SVD在图像的稀疏表示中具有稀疏性、特征保持性、可分性等特点,并被广泛应用于图像处理中^[8]。Engan等^[9]提出了最优方向法(Method of Optimal Directions, MOD)字典学习算法,能够在稀疏表示和字典更新两个阶段快速运行,但无法保证收敛到全局最优解。Mairal等^[10]提出了在线字典学习算法(Online Dictionary Learning, ODL),该算法每次仅从样本集中随机使用一个样本训练字典,根据上次获得的字典,使用最小角度回归(Least-angle Regression, LARS)算法进行系数编码,从而获得表示系数,最后再固定系数并更新字典。Aharon等^[11]提出了K-奇异值分解(K-Singular Value Decomposition, K-SVD)算法,使用SVD算法对K均值聚类算法进行泛化改进。K均值聚类算法认为可以将稀疏表示问题看作是矢量量化问题,但其只能对每个信号使用一个原子近似表示,而K-SVD算法利用SVD方法来更新原子,能够用多个原子的线性组合来表示每个信号,进而训练出与图像相匹配的冗余字典^[12]。目前,使用得最广泛的方法是K-SVD字典学习算法。

传统的分块图像压缩感知对所有的图像块均采用同一字典进行稀疏表示,由于不同的图像块包含的图像信息和特征可能不同,因此不能很好地适应不同类型的图像块^[13]。文献[14]针对单幅图像使用方差对图像块进行分类,对不同类别的图像块分别进行字典训练,从而获得针对此图像各个类别下的图像块最优稀疏表示字典。由于原算法仅针对单幅图像训练字典,当分类过多时每类的数据量较少,这会导致不能训练出足够的冗余字典,且方差仅能评价图像内的灰度差异,体现图像灰度值变化,不能很好地表示图像所包含的信息。对此,本文提出基于灰度熵的图像自适应字典学习算法,通过图像的灰度熵来反映图像的信息丰富程度,同时,采用大数据量的图像库作为训练样本,在图像块稀疏表示过程中根据灰度熵对图像库中的图片进行分块和分类。对不同类的图像块设定不同的训练参数,得到不同大小的稀疏字典,使用训练好的字典对待感知图像进行稀疏表示。随着图像库的不断增多,新图像的稀疏表示性能得到提升。同时,这种方法能够与图像块特征保持很好的一致性,图像块在字典的稀疏表示下能够更加精确,显著提升重构质量。

1 图像稀疏表示和字典学习

1.1 图像稀疏表示

图像信号是一种二维信号,具有结构复杂、数据量大等特点,对图像信号的研究一直是数字信号处理领域的一个重点研究内容。信号的稀疏性是指信号可以使用少数特征向量的线性组合来表示,即稀疏表示,这一特性被广泛应用于信号处理领域,是学者们研究的重点。稀疏表示的基本思想是认为自然图像或者信号可以由一个基函数字典线性叠加表示^[15]。从数学的角度,稀疏表示就是用线性分解来对多维的图像或信号进行表示。图像的稀疏表示模型如下:

$$X = D\theta \quad (1)$$

其中, $X \in R^n$ 表示输入的图像信号, $\theta = [\theta_1, \theta_2, \dots, \theta_M]^T$ 为稀疏表示向量。当 θ 中非零元素的数目为 K 时,称 θ 是 K 稀疏的,或者称图像 X 在字典 D 下是 K 稀疏的。Candes等^[16]通过研究发现,信号越稀疏,信号重建后的精度就越高,信号稀疏表示的目的主要在于寻找一个自适应字典使信号能够最优稀疏表示。因此字典的选择十分重要,它决定了图像反问题的求解质量。目前,构造字典的方法主要有两种:人工构造和字典学习。

人工构造字典均根据原始信号特征调整参数值,进而选取相应的字典原子,生成的字典具有较强的针对性,能够对图像实现有效的稀疏表示。但此类构造方法的效果通用性差,对于特征不明显的图像,可能得不到很好的稀疏表示效果。

近年来,人们开始基于数据或信号本身来学习得到冗余字典,这类字典中的原子能与训练集中的信号本身的特性相符合。与基于解析方法的字典相比,通过学习获得的字典原子的数量更多,形态更丰富,对于图像信号,基于学习得到的字典能够对信号进行更稀疏的表示。

1.2 K-SVD 算法

K-SVD算法是当前字典学习中最常用的算法之一,算法的思想较为简单,首先固定字典 D ,使用匹配追踪(Matching Pursuit, MP)或正交匹配追踪(Orthogonal Matching Pursuit, OMP)算法得到训练信号在该字典下的稀疏系数,然后假定稀疏系数固定,根据稀疏系数更新字典,逐步迭代,直到满足设定好的停止条件。

对于信号 $X \in R^n$,使用 $D \in R^{n \times T}$ 表示包含有 T 个原子的过完备字典, $D = \{d_j\}_{j=1}^T$,其中 d_j 表示字典列。则信号 X 能够通过字典原子线性表示得到,表示系数定义为 $\theta \in R^T$,如式(1)所示。

然而,在使用字典原子对原信号 X 进行线性表示时,通常存在一定的误差,即:

$$X \approx D\theta \quad \text{s. t.} \quad \|X - D\theta\|_2 \leq \epsilon \quad (2)$$

将 N 个信号的集合记为 $Y = \{X_i\}_{i=1}^N$,设其对应的稀疏系数矩阵为 Θ ,在满足给定的稀疏度的条件下找到对于信号 Y 的最优逼近值,即求解如下方程:

$$\min_{D, \Theta} \{ \|Y - D\Theta\|_2 \} \quad \text{s. t.} \quad \forall i, \|\theta_i\|_0 \leq \Gamma \quad (3)$$

其中, Γ 为系数中非零项数目的上限。另一种迭代终止条件是在满足给定的稀疏逼近误差的条件下寻求信号的最优稀疏表示,所需求解的目标方程式如下:

$$\min_{D, \Theta} \{ \|\theta_i\|_0 \} \quad \text{s. t.} \quad \|Y - D\Theta\|_2 \leq \epsilon \quad (4)$$

其中, ϵ 为稀疏逼近误差的上限。同时,也可以同时考虑稀疏性和逼近误差,目标方程式如式(5)所示:

$$\min_{D, \Theta} \{ \|Y - D\Theta\|_2 + \lambda \|\theta_i\|_0 \} \quad (5)$$

对于字典的原子,K-SVD算法是逐列进行更新的。假设固定字典 D 和系数矩阵 Θ ,当更新字典的第 k 列 d_k 时,系数矩阵中对应 d_k 的第 k 行为 θ_k^T ,则式(4)中的惩罚项能够表示为:

$$\begin{aligned}\|Y-D\mathbf{O}\|_2^2 &= \|Y-\sum_{j=1}^k d_j \boldsymbol{\theta}_T^j\|_2^2 \\ &= \|(Y-\sum_{j \neq k} d_j \boldsymbol{\theta}_T^j) - d_k \boldsymbol{\theta}_T^k\|_2^2 \\ &= \|\mathbf{E}_k - d_k \boldsymbol{\theta}_T^k\|_2^2\end{aligned}\quad (6)$$

式(6)中,把 $D\mathbf{O}$ 分解为 k 个秩为1的矩阵的和,假设其中 $k-1$ 项均为固定的,剩下的1项为需要更新的第 k 项。矩阵 \mathbf{E}_k 为去除原子 d_k 的成分后,在所有的 N 个训练信号样本中造成的误差。若此时对 \mathbf{E}_k 使用SVD分解算法来更新获取 d_k 和 $\boldsymbol{\theta}_T^k$,由于在更新 d_k 的过程中并未添加稀疏性约束, $\boldsymbol{\theta}_T^k$ 可能并不稀疏。即 $\boldsymbol{\theta}_T^k$ 与更新 d_k 前 $\boldsymbol{\theta}_T^k$ 的非零元素所处的位置和值不同。可以先对系数中的非零值进行保留,再进行SVD分解,具体步骤如下:定义集合 ω_k 为用到原子 d_k 所有信号的索引构成的集合,即:

$$\omega_k = \{i | 1 \leq i \leq N, \boldsymbol{\theta}_T^k(i) \neq 0\} \quad (7)$$

定义 $\boldsymbol{\Omega}_k$ 为 $N \times |\omega_k|$ 矩阵,它在 $(\omega_k(i), i)$ 处的值为1,其他点为0。 $\boldsymbol{\theta}_R^k = \boldsymbol{\theta}_T^k \boldsymbol{\Omega}_k$ 为向量 $\boldsymbol{\theta}_T^k$ 去除零输入后收缩的结果,向量 $\boldsymbol{\theta}_R^k$ 的长度为 $|\omega_k|$ 。同样地, $\mathbf{Y}_k^R = \mathbf{Y} \boldsymbol{\Omega}_k$ 为使用原子 d_k 的训练信号集合,大小为 $n \times |\omega_k|$ 。定义 $\mathbf{E}_k^R = \mathbf{E}_k \boldsymbol{\Omega}_k$,则:

$$\|\mathbf{E}_k \boldsymbol{\Omega}_k - d_k \boldsymbol{\theta}_T^k \boldsymbol{\Omega}_k\|_2^2 = \|\mathbf{E}_k^R - d_k \boldsymbol{\theta}_R^k\|_2^2 \quad (8)$$

经过上述变换后,再对 \mathbf{E}_k^R 进行SVD分解:

$$\mathbf{E}_k^R = U \Delta V^T \quad (9)$$

d_k 的更新结果 \tilde{d}_k 为 U 的第一列,相应系数向量 $\boldsymbol{\theta}_T^k$ 更新为 V 的第一列乘以 $\Delta(1,1)$ 。即对 \mathbf{E}_k 和 $\boldsymbol{\theta}_T^k$ 进行变换, $\boldsymbol{\theta}_T^k$ 中只保留非零位置的项, \mathbf{E}_k 只保留 d_k 和 $\boldsymbol{\theta}_T^k$ 中非零位置项乘积后的项,从而形成 \mathbf{E}_k^R ,对 \mathbf{E}_k^R 进行奇异值分解,更新 d_k 。

矩阵 $\mathbf{A} \in C^{m \times n}$, \mathbf{A} 的非零奇异值为 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$, $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r$ 是对应于奇异值的左奇异向量, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ 是对应于奇异值的右奇异分量,则可以得到:

$$\mathbf{A} = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T \quad (10)$$

式(10)是对矩阵 \mathbf{A} 的奇异值展开式。对于 $s \leq r$,去除部分较小的奇异值多对应的项,记矩阵 \mathbf{A}_k 为:

$$\mathbf{A}_k = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \dots + \sigma_s \mathbf{u}_s \mathbf{v}_s^T \quad (11)$$

其中, \mathbf{A}_k 是一个秩为 s 的 $m \times n$ 矩阵。 \mathbf{A}_k 是所有秩为 s 的 $m \times n$ 矩阵中,在F范数意义下与矩阵 \mathbf{A} 距离最近的一个矩阵。根据这一性质,字典原子 d_k 与系数向量 \mathbf{X}_T^k 的更新结果的乘积是F范数意义下对 \mathbf{E}_k^R 的最佳逼近。

K-SVD算法由于不需要进行矩阵求逆计算,运算复杂度低,并且在字典更新步骤中对系数矩阵与字典原子进行联合更新,能够提高算法的收敛速度,因此其在实际应用中被广泛使用。

2 基于图像灰度熵的自适应字典学习

图像的纹理复杂度能够反映图像蕴涵的信息量,若图像块的纹理复杂度较高,则该图像块的信息量丰富,包含了大量的图像信息,而一幅图像的灰度熵能够反映图像的信息丰富程度,即图像的纹理复杂度。因此,本文提出采用图像的灰度熵值作为图像纹理复杂度的评价指标,并将其作为图像块的分类指标。

2.1 灰度熵

灰度熵是一种能量空域分布状态不确定性的量度,区别

于根据图像灰度直方图统计所定义的图像熵^[18]。灰度熵同时考虑了直方图的概率作用信息和图像块内灰度级的均匀性,是同一块内像素共同作用的结果。对于 $B \times B$ 大小的图像块,假设像素灰度分别为 m_1, m_2, \dots, m_{n^2} ,则该图像块中的图像灰度熵 h 为:

$$h = - \sum_{j=0}^{n^2} p_j \log_2 p_j / \log_2 B^2 \quad (12)$$

其中, p_j 表示图像上灰度级 j 出现的概率:

$$p_j = \frac{k_j}{\sum_{j=0}^{255} k_j} \quad (13)$$

其中, k_j 表示灰度级为 j 的像素在图像上出现的次数。由式(12)可以看出,图像灰度熵运算是一种非线性的空域滤波器,使用了归一化的处理, $0 \leq h \leq 1$ 。图像的灰度熵能够反映图像块像素灰度的差异程度,同时直接反映了图像块内灰度级的均匀性,能够表示图像的信息丰富程度。图像的灰度熵越大,图像块内能量的空间分布越大,即图像块内像素灰度的差异越大,该图像块所包含的信息量越丰富,图像的纹理复杂度越高。像素灰度级越大,图像分割效果越好,当图像块内所有像素的灰度熵都不同时,图像灰度熵达到最大, $h=1$;反之,图像灰度熵越小,图像内像素灰度差异越均匀,包含的图像信息量越少,纹理复杂度越低。因此,灰度熵是一种很好的对图像块进行分类的评价指标。

2.2 基于图像灰度熵进行自适应字典学习

对于图像的稀疏表示,由于图像数据量大、结构复杂等特性,直接进行二维观测与稀疏表示的难度较大,计算的复杂度与图像信号的大小密切相关。当图像较大时,直接对整幅图像进行稀疏表示及采样重构时计算复杂度太高。因此,现有算法均首先对图像进行分块处理,然后把二维图像块排列成列向量形式,再使用现有算法对其进行观测与稀疏表示^[19]。

在传统的图像分块压缩感知^[20]的稀疏表示字典训练中,所有的图像块均在同一稀疏字典下进行稀疏表示,这对字典原子包含的结构特征的丰富性提出了很高的要求。图像中通常具有多种不同的纹理形态成分,针对不同成分应使用不同的约束^[21]。由于不同的图像块的纹理复杂度不同,图像包含的信息也不同,因此采用相同的稀疏字典不能对蕴含不同信息的图像块进行很好的稀疏表示,不能体现出图像块的差异性,不能很好地提取出图像块所包含的信息。

灰度熵由于能够很好地考虑不同图像块的纹理信息,衡量图像信息的丰富程度,是一种可靠的图像分类阈值指标。因此,本文为提高字典对图像块的适应性,采用基于灰度熵分类图像块的稀疏字典训练方法,基于灰度熵的大小对图像块进行分类,计算各个图像子块的灰度熵数值,按照灰度熵升序的方式排列相应的图像子块,然后将图像子块等分为多个类别,同类图像块使用相同的观测矩阵进行观测。

基于图像块灰度熵的K-SVD字典训练方法的基本框图如图1所示,相关步骤如下:

步骤1 将图像库中的所有图像均分割成若干个 $B \times B$ 大小的子块,共计 S 块。将每个图像块矩阵变为列向量形式。 S 个图像块组成大小为 $B^2 \times S$ 的待训练的数据矩阵 \mathbf{M} 。

步骤2 根据式(12)、式(13)计算矩阵 \mathbf{M} 中每一列的灰

度熵,并按灰度熵从小到大的顺序重新排列 M 中的对应列,从而得到 M_0 。根据灰度熵的大小将 M_0 的列向量分为 C 类,将每连续的 S/C 归为一类,得到 C 个分类后的图像块集合 $M_c(i), i=1,2,\dots,C$ 。

步骤 3 使用 K-SVD 算法对每类图像块 $M_c(i)$ 训练字典,根据当前图像块类别的灰度熵,在字典训练时训练出原子个数不同的多个字典。对于灰度熵较小的图像类,其包含的图像细节较少,图像较平滑,使用含有较少原子的字典就能对其进行有效的稀疏表示。对于灰度熵大的图像类,其包含的图像细节较多,图像纹理度较高,因此需要使用包含更多原子的字典对该类图像块进行稀疏表示。最后保存训练得到的字典。

步骤 4 通过步骤 3 得到每类图像块的训练字典和相应的稀疏表示系数矩阵。字典与相应系数矩阵的乘积为该图像块在该类字典下的表示结果。将各类图像块的字典表示结果按照分类前的索引排列,得到所有图像块的字典表示形式 R_b 。将 R_b 的列向量重新变为图像块的形式,将图像块对应到原图像中相应的位置,从而得到基于图像块灰度熵的 K-SVD 字典表示形式。

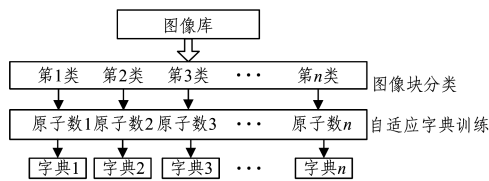


图 1 基于图像灰度熵的自适应字典学习算法框图

Fig. 1 Block diagram of adaptive dictionary learning algorithm based on image gray entropy

3 仿真实验及结果分析

本文使用加利福尼亚理工学院的 101 类图像数据库,即 Caltech101 数据库,Caltech101 总共包含 9 146 幅图像,共有 102 个类别,每一类大约有 40~800 幅图像,大多数类别约有 50 张图像,均为 jpg 格式的图像。从 Caltech101 图像库中选择纹理复杂度较高的 wild-cat 类与纹理复杂度相对较低的飞

机类分别进行仿真实验。两类输入图像分别组成图像库并训练得到字典,能更好地体现出该类图像所包含的图像信息,而对图像进行更好的稀疏表示。分别随机使用两类数据集中 2/3 的图像作为字典训练的样本图像库,再对剩余图像中的随机图像进行稀疏表示和重构。

首先对图像组成的图像库进行分块处理,再分别采用灰度熵和方差两种分类指标来对图像块进行分类,对每类图像块分别训练稀疏表示冗余字典。在对图像库进行分块时,每个图像块的大小为 8×8 ,分别按照灰度熵和方差的大小以从小到大的顺序对其进行排序,并分别将图像库分为 4 类、5 类和 6 类,训练出相应的多个稀疏表示冗余字典。在分为 4 类的情况下,表示单个图像块的字典原子的大小依次为 $64 \times 64, 64 \times 128, 64 \times 192$ 和 64×256 ,多个分类时字典大小依次类推。在字典训练过程中,K-SVD 算法的迭代次数为 10 次,K-SVD 算法中信号稀疏分解算法采用 OMP 算法。在数据观测时,使用采样率为 0.5 的观测矩阵对原信号进行采样。

字典表示图像为 K-SVD 算法训练得到的字典与稀疏表示后得到的最终的系数矩阵的乘积,该结果表明了图像在当前冗余字典下的稀疏逼近情况。表示图像与原图像之间的峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 值越高,图像在该字典下的稀疏逼近效果越好。

本文通过 K-SVD 算法以图像库为样本数据训练多个字典,并对要稀疏表示的图像采用同样的分类指标进行分类,并将分类后的图像与各个类别的字典逐一对应,使用相同类别下的字典对其进行稀疏表示,分别对图像库中的训练集图像与测试集图像进行稀疏表示并重构,计算其平均 PSNR 值。同时,本文将提出的基于灰度熵分类的自适应字典学习算法与传统的 K-SVD 字典学习算法进行对比。首先使用传统的 K-SVD 算法对训练集图像和测试集图像直接进行字典学习,然后再使用所得字典对图像进行稀疏表示,最后对其进行重构并计算测试集与训练集的平均 PSNR 值。为保证对比的有效性,三者使用相同的采样率与重构算法对重构效果进行对比,重构结果如表 1 与表 2 所列,表 1 列出了 wild-cat 类的重构 PSNR 值,表 2 列出了飞机类的重构 PSNR 值。

表 1 针对 wild-cat 类的基于灰度熵与方差分类的字典学习算法与传统 K-SVD 学习算法的重构 PSNR 值

Table 1 Reconstruction PSNR between gray entropy/variance dictionary learning algorithm and traditional K-SVD learning algorithm based on wild-cat category

(单位: dB)

数据集	4 类		5 类		6 类		传统 K-SVD 字典学习算法
	灰度熵	方差	灰度熵	方差	灰度熵	方差	
训练集	25.7384	25.1367	25.7543	25.2891	28.4765	28.2429	26.7002
测试集	25.4489	24.8435	25.6628	25.0181	28.0777	27.9913	26.3613

表 2 针对飞机类的基于灰度熵与方差分类的字典学习算法与传统 K-SVD 学习算法的重构 PSNR 值

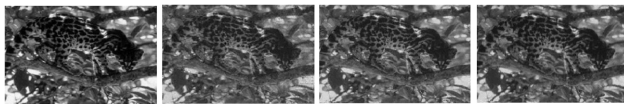
Table 2 Reconstruction PSNR between gray entropy/variance dictionary learning algorithm and traditional K-SVD learning algorithm based on airplane category

(单位: dB)

数据集	4 类		5 类		6 类		传统 K-SVD 字典学习算法
	灰度熵	方差	灰度熵	方差	灰度熵	方差	
训练集	32.3742	32.2851	32.7754	32.6585	43.3998	33.8224	31.8915
测试集	31.5677	31.2388	32.6553	32.5858	33.2336	32.8026	33.1598

由表 1 和表 2 可以看出,当分类数相同时,以灰度熵为分类指标训练出的字典对图像稀疏表示后的重构效果优于以方差为分类指标的重构效果。同时,随着分类数目的增加,图像的重构效果也有所改善。在相同分类数目并同样使用灰度熵或方差作为分类指标的情况下,训练集图像的重构质量高于测试集图像的重构效果。同时,将本文所提算法与传统单幅图像的 K-SVD 字典学习算法进行对比,由于传统的 K-SVD 算法的重构效果只与当前输入图像相关,对测试集和训练集中所有图像的重构 PSNR 值取平均后,由表 2 可以发现当分类数大于或等于 5 时,使用灰度熵作为分类指标的重构效果优于传统的 K-SVD 算法,即能对图像进行更好的稀疏表示。对比不同纹理度的数据集重构效果可以看出,对于纹理度大的数据集,随着分类的增加,基于灰度熵分类指标的重构 PSNR 值的生长趋势大于纹理度较低的数据集。

选取 wild-cat 类中的第 17 幅图像以及飞机图像库中的第 25 幅图像。分别设分类数为 4,5,6,并分别使用方差和灰度熵作为分类指标来训练字典以对图像进行稀疏表示,并对重构后得到的重构图像进行观察。图 2(a)~图 2(d)分别为以灰度熵为分类判决指标,分为 4 类、5 类和 6 类时的 wild-cat 类中的第 17 幅图像的重构图以及原始图。



(a)原图 (b)4 类重构图片 (c)5 类重构图片 (d)6 类重构图片

图 2 第 17 幅图像在灰度熵条件下分类数为 4,5,6 时的重构图像与原图

Fig. 2 Reconstructed images of 17th image under conditions of gray entropy are classified as 4,5 and 6 and original image

图 3(a)~图 3(d)分别为以灰度熵为分类判决指标,以飞机类中的第 25 幅图像为原始图像以及分为 4 类、5 类、6 类时的重构图像。图 4(a)~图 4(c)分别为以方差为分类判决指标,分为 4 类、5 类、6 类时的重构图像。从重构图像中可以发现,随着分类数目的增加,图像重构的 PSNR 值也随之增大。以灰度熵为分类指标的重构质量高于以方差为分类指标的重构质量,重构质量如表 3 所列。



(a)原图 (b)4 类重构图片 (c)5 类重构图片 (d)6 类重构图片

图 3 第 25 幅图像在灰度熵条件下分类数为 4,5,6 时的重构图像与原图

Fig. 3 Reconstructed images of 25th image under conditions of gray entropy are classified as 4,5 and 6 and original image



(a)4 类重构图片 (b)5 类重构图片 (c)6 类重构图片

图 4 在方差条件下分类数为 4,5,6 时的重构图像

Fig. 4 Reconstructed images under condition of variance, and classified as 4,5,6

表 3 灰度熵与方差训练字典对随机图像的重构 PSNR 值

Table 3 Reconstructed PSNR value of random image by gray scale entropy and variance training dictionary (单位:db)

	4 类	5 类	6 类
灰度熵	35.1885	36.0970	36.7106
方差	34.1870	35.6948	36.3463

从飞机图像库中随机抽取 4 幅图像,使用图像本身进行稀疏表示、分类字典训练及重构,用图像库训练得到的字典对图像进行稀疏表示及重构,使用灰度熵作为图像块分类的判决指标,比较图像重构后的 PSNR 值。图像库训练字典与单幅图像训练字典的重构效果情况如表 4 所列。由表 4 可以看出,当单幅图像的分类过多时,重构质量可能会下降。这是因为当分类过多时,单幅图像分类后的每一类数据量可能会较少,冗余字典训练得不够充分,不能很好地对图像进行稀疏表示,导致重构质量下降,而使用图像库的方法则不会产生此问题。

表 4 图像库训练字典与单幅图像训练字典的重构效果比较

Table 4 Comparison of recovery effect of single image training dictionary and images training dictionary (单位:db)

图像序号	单幅图像训练多个字典			图像库训练多个字典		
	4 类	5 类	6 类	4 类	5 类	6 类
100	35.0610	35.9137	15.9838	38.8426	39.1669	39.7168
266	29.6457	33.0838	16.4581	23.8229	35.5428	37.8408
610	25.1743	28.3493	17.4647	38.5893	38.8906	39.4651
666	25.8096	27.6244	17.4302	42.0090	42.8986	42.9615

当图像块的总数一定时,图像块分类数目越多,即分类越细,同一类中的图像块之间的相似性越高,训练出的字典对于图像块的稀疏表示效果就越好。同样地,在分类数目一定时,数据库中的图像数量越多,即每一类中的样本量越大,训练出的字典稀疏表示能力就越高。但对于单幅图像来说,当分类数过多时,可能训练不出合适的冗余字典,从而导致字典对信号的稀疏表示性能下降,图像的重构质量降低。而采用图像库的思想,只要图像库足够大,分类数就可以相应增加,但过大的数据量及分类数会导致运算量的增加,增加功耗。由仿真结果可知,对于纹理复杂度较高与较低的图像,本文所提算法均能提高其图像重构质量,能对图像进行良好的稀疏表示,本文所提算法的思想可行。

结束语 针对图像压缩感知字典学习算法,未考虑不同图像块包含的图像信息不同、仅训练一个固定字典的问题,本文利用灰度熵度量每一类图像块的纹理度,并依据灰度熵大小来对图像块进行分类,针对每一类训练相应的字典,进而对图像进行稀疏表示。仿真结果表明,本文所提算法能够对图像进行较好的稀疏表示,在选择合适的类别时对不同图像都能有效提高 PSNR 指标,图像的重构质量得到了明显提高。

参考文献

- [1] CANDÉS E J, ROMBERG T, TAO T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information[J]. IEEE Transactions on Information Theory, 2006, 52(2): 489-509.

- [2] DONOHO D L. Compressed sensing [J]. *IEEE Transactions on Inform Theory*, 2006, 52(4): 1289-1306.
- [3] SHRIVIDYA G, BHARATHI S H. Application of Compressed Sensing on Magnetic Resonance Imaging: A brief survey [C] // *IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology*. Bangalore, India, 2016: 2037-2041.
- [4] LIAN Q S, SHI B S, CHEN S Z. Research Advances on Dictionary Learning Models, Algorithms and Applications [J]. *Acta Automatica Sinica*, 2015, 41(2): 240-260. (in Chinese)
练秋生, 石保顺, 陈书贞. 字典学习模型、算法及其应用研究进展 [J]. *自动化学报*, 2015, 41(2): 240-260.
- [5] RENE V, YI M, SHANKAR S. Generalized principal component analysis [J]. *IEEE Transactions Pattern Anal Mach Intell*, 2005, 27(12): 1945-1959.
- [6] LIU Z, SONG X N, YU D J, et al. Super-resolution reconstruction algorithm based on multi-component dictionary and sparse representation [J]. *Journal of Nanjing University of Science and Technology*, 2014, 38(1): 1-5. (in Chinese)
刘梓, 宋晓宁, 於东军, 等. 基于多成分字典和稀疏表示的超分辨率重建算法 [J]. *南京理工大学学报*, 2014, 38(1): 1-5.
- [7] YANG S Y, JIN H H, WANG M, et al. Data-Driven Compressive Sampling and Learning Sparse Coding for Hyperspectral Image Classification [J]. *IEEE Geoscience and Remote Sensing Letters*, 2014, 11(2): 479-483.
- [8] LIU X M, LIU Y M. Color image denoising with block K-SVD dictionary learning [J]. *Journal of Nanjing University of Science and Technology*, 2016, 40(5): 607-612. (in Chinese)
刘晓曼, 刘永民. 基于分块 K-SVD 字典学习的彩色图像去噪 [J]. *南京理工大学学报*, 2016, 40(5): 607-612.
- [9] ENGAN K, AASE S O, HUSOY J H. Method of optimal directions for frame design [J/OL]. http://xueshu.baidu.com/usercenter/paper/show?paperid=dc1a54ef946aaa15e2aa439a9d116c9&site=xueshu_se.
- [10] MAIRAL J, BACH F, PONCE J, et al. Online learning for matrix factorization and sparse coding [J]. *Journal of Machine Learning Research*, 2010, 11(1): 19-60.
- [11] AHARON M, ELAD M, BRUCKSTEIN A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. *IEEE Transaction on Signal Processing*, 2006, 54(11): 4311-4322.
- [12] MITTU G P, VIVEK M, JOONKI P. Imaging inverse problem using sparse representation with adaptive dictionary learning [C] // *IEEE International Advance Computing Conference*. 2015.
- [13] ZHANG S Y. Research on Image Compressive Sensing Technology Based on Redundant Dictionary [D]. Jiling: JiLing University, 2016. (in Chinese)
张书扬. 基于冗余字典的图像压缩感知技术研究 [D]. 吉林: 吉林大学, 2016.
- [14] CONG Y L, ZHANG S Y, LIAN Y Y. K-SVD dictionary learning and image reconstruction based on variance of image patches [C] // *8th International Symposium on Computational Intelligence and Design*. 2015: 254-257.
- [15] SUN J D, ZHAO H H. Sparse Representation and Applications in Image Processing [J]. *Infrared Technology*, 2014, 36(7): 533-537. (in Chinese)
孙君顶, 赵慧慧. 图像稀疏表示及其在图像处理中的应用 [J]. *红外技术*, 2014, 36(7): 533-537.
- [16] CANDES E, ROMBERG J, TAO T. Stable signal recovery from incomplete and inaccurate measurements [J]. *Communications on Pure and Applied Mathematics*, 2006, 59(8): 1207-1223.
- [17] ZHANG D S, ZHANG L H. Research on fast dictionary learning algorithm under compressed sensing framework [J]. *Research and exploration in laboratory*, 2015, 34(11): 94-98. (in Chinese)
张得生, 张莉华. 压缩感知框架下快速字典的学习算法 [J]. *实验室研究与探索*, 2015, 34(11): 94-98.
- [18] ZHANG Y L, WANG Y, LU H Z. Block objects detection based on entropy of brightness [J]. *Systems Engineering and Electronic*, 2008, 30(2): 201-204. (in Chinese)
张永亮, 汪洋, 卢焕章. 基于图像灰度熵的团块目标检测方法 [J]. *系统工程与电子技术*, 2008, 30(2): 201-204.
- [19] KHANH Q D, HIUK J S, JEON B. Weighted Overlapped Recovery for Blocking Artefacts Reduction in Block-based Compressive Sensing of Images [J]. *Electronics Letters*, 2015, 51(1): 48-50.
- [20] ZHANG B, LIU Y L. A novel block compressed sensing based on matrix permutation [C] // *Visual Communications and Image Processing*. 2016: 1-4.
- [21] ZHU X, LIU L, JIN P. Morphological component decomposition combined with compressed sensing for image compression [C] // *2016 IEEE International Conference on Information and Automation*. 2016: 1726-1731.