

学校编码: 10384

分类号 _____ 密级 _____

学号: 23320081153321

UDC _____

厦 门 大 学

硕 士 学 位 论 文

基于自适应字典稀疏表示的人脸图像
压缩算法研究

Research on Facial Images Compression
Based on Adaptive Dictionary Sparse Representation

钱坤

指导教师姓名: 丁兴号 副教授

专业名称: 信号与信息处理

论文提交日期: 2011年5月

论文答辩日期: 2011年6月

学位授予日期: 2011年 月

答辩委员会主席: _____

评 阅 人: _____

2011年5月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果,均在文中以适当方式明确标明,并符合法律规范和《厦门大学研究生学术活动规范(试行)》。

另外,该学位论文为()课题(组)的研究成果,获得()课题(组)经费或实验室的资助,在()实验室完成。(请在以上括号内填写课题或课题组负责人或实验室名称,未有此项声明内容的,可以不作特别声明。)

声明人(签名):

年 月 日

厦门大学博硕士学位论文摘要库

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（ ） 1. 经厦门大学保密委员会审查核定的保密学位论文，
于 年 月 日解密，解密后适用上述授权。

（ ） 2. 不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年 月

厦门大学博硕士学位论文摘要库

摘要

图像压缩是图像处理的一个重要环节,是图像存储和传输过程中要解决的重要问题。图像压缩技术已经研究了几十年,取得了很大的成绩,并已制定 JPEG、JPEG2000 等静止图像压缩标准。然而,对于一些使用广泛的特殊图像如人脸图像,由于其图像小且内容丰富,在低比特率条件下压缩重建后的图像质量不能令人满意。为此,需要通过引入新的理论对人脸图像的低比特率压缩进行研究。

图像信息的有效表示是各种图像处理的基础,有效表示是指用甚少的表示系数即可表示图像的主要能量。因此,图像的稀疏表示能力对于压缩性能具有重要的影响。为了解决传统正交变换的稀疏表示能力不足问题,本文研究自适应超完备字典稀疏表示方法,并将其应用于人脸图像的低比特率压缩。具体在以下方面取得了一些进展:

1. 本文提出了一种新的自适应超完备字典的构造算法: **K-PCA** 算法。该算法是一种改进的 **K-SVD** 算法,用于超完备字典的训练学习,以实现图像的有效稀疏表示。与 **K-SVD** 算法相比较,该算法主要在初始字典的获取和原子的更新方面做了改进,以期望相似的样本采用相似的原子表示。相关图像去噪实验证明了该字典训练学习算法的可靠性。进一步,论文将该算法应用于人脸图像的低比特率压缩。相关实验证实,应用本文所提算法构造的超完备字典进行人脸图像的低比特率压缩,其主客观效果都明显优于 JPEG 和 JPEG2000。

2. 为了进一步提高人脸图像的压缩比,本文提出一种基于自适应双字典的人脸图像压缩算法。该算法将基于自适应字典稀疏表示的图像超分辨率重建算法与图像压缩算法有机结合起来,以保证在具有较高重建质量的情况下获得人脸图像的甚低比特率压缩。该算法主要利用对偶的低分辨率和高分辨率图像自适应获得的对偶双字典具有相同的稀疏表示系数这一特点。压缩时,首先对待压缩人脸图像下采样获得低分辨率人脸图像,并在相应的低分辨率字典上稀疏分解图像,再对稀疏分解系数编码以实现压缩;解压缩时,则由稀疏表示系数与对偶的高分辨率字典重建图像。相关实验证实,本文提出的方法能够在保证图像重建质量的情况下,实现甚低比特率条件下(比特率小于 0.125bpp)的人脸图像压缩。

关键词: 人脸压缩; 稀疏表示; 超完备字典; 甚低比特率

厦门大学博硕士论文摘要库

Abstract

Image compression is an important part of image processing, also an important issues to be solved in image storage and transmission. The image compression technology has been studied for decades and has got some achievements, JPEG and JPEG2000 have been developed. However, for some special images such as facial image which is small but has rich content, the quality of reconstructed image is not satisfactory under low bit rate compression condition. Therefore, introducing new theoretical method to low bit rate compression of facial images is very meaningful.

Valid image representation is the basis of image processing, the sparsity of transform coefficients is very important for compression performance. To overcome the insufficiency of commonly used DCT and Wavelet transform, this paper apply the theory of adaptive over-complete sparse representations to the filed of facial image compression. Specifically, progress has been made in the following areas:

Firstly, this paper presents a new adaptive overcomplete dictionary construction algorithm: K-PCA. The algorithm is improved by the K-SVD algorithm, used to train complete dictionary, in order to achieve a valid representation of the image. Compared with the K-SVD, our algorithm do improvement mainly in construction of the initial dictionary and the atom update. By denoising experiments proved the reliability of the dictionary training algorithm. The reliability of the dictionary training algorithm has been proved in image denoising experiments. The algorithm is then applied to facial image compression. The experimental results demonstrate that the proposed method is much better than JPEG and JPEG2000 in both objective performance and visual quality, achieve low bit rate facial image compression.

Secondly, this paper proposes a new facial image compression algorithm based on adaptive dual dictionaries to further increase the compression ratio. With the idea of image super resolution via sparse representation, we can enforce the similarity of sparse representations between the low resolution and high resolution image patch pair with respect to their own dictionaries by jointly training two dictionaries for the low and high resolution image patches. Low resolution facial images are generated by downsampling real facial images, the sparse representation of a low resolution image can be compressed with the low resolution image dictionary. After decompressing, the

sparse representation of a low resolution image can be applied with the high resolution image dictionary to generate a high resolution facial image. The experimental results show that the proposed method can realize facial image compression at very low bit rate (bit rate less than 0.125bpp).

Key words: Facial image compression; sparse representation; overcomplete dictionary; very low bit rate

厦门大学博硕士论文摘要库

目 录

第一章 绪 论	1
1.1 研究背景	1
1.2 技术发展	2
1.3 论文主要内容和结构	4
第二章 图像压缩概述	5
2.1 图像压缩模型	5
2.2 图像压缩分类	6
2.3 静态图像压缩标准	7
2.4 图像压缩效果评价	12
第三章 图像稀疏表示	15
3.1 调和分析图像表示	15
3.1.1 傅立叶变换	15
3.1.2 小波变换	16
3.2 多尺度几何分析图像表示	17
3.2.1 Ridgelet 变换	18
3.2.2 Curvelet 变换	19
3.2.3 Contourlet 变换	19
3.3 基于超完备字典的图像稀疏表示	20
3.3.1 信号超完备稀疏表示理论	21
3.3.2 超完备字典的构建	22
3.4 稀疏表示在图像处理领域的应用	23
第四章 基于自适应超完备稀疏表示的人脸图像压缩方法	25
4.1 信号的稀疏分解算法	25
4.1.1 贪婪算法	25
4.1.2 全局优化方法	27
4.1.3 其他方法	28
4.2 自适应超完备字典的构造	29
4.2.1 K 均值聚类算法	29

4.2.2 K-SVD 算法——K 均值聚类算法的推广	31
4.2.3 K-PCA 算法——K-SVD 算法的改进	34
4.3 基于 K-PCA 字典训练算法的图像去噪及其实验结果分析	36
4.4 基于 K-PCA 字典训练算法的人脸图像压缩及其实验结果分析	39
4.5 本章小结	45
第五章 基于自适应双字典的人脸图像压缩算法	46
5.1 基于自适应双字典的人脸图像压缩	46
5.1.1 超分辨率图像复原方法概述	46
5.1.2 基于稀疏表示的图像超分辨率复原	49
5.1.3 基于自适应双字典的人脸图像压缩	51
5.2 算法实现与实验结果分析	53
5.3 本章小结	59
第六章 总结与展望	61
参考文献	62
致谢	65
附录	66

厦门大学博硕士学位论文摘要

Table of Contents

Chapter 1 Introduction.....	1
1.1 Background of the subject	1
1.2 Technology development	2
1.3 Main work and content arrangements of the thesis.....	4
Chapter 2 Overview on Image Compression	5
2.1 Image compression model.....	5
2.2 Image compression and classification.....	6
2.3 Still image compression standards.....	7
2.4 Evaluation of image compression	12
Chapter 3 Image Sparse Representation	15
3.1 Harmonic analysis for image processing.....	15
3.1.1 Fourier transform	15
3.1.2 Wavelet transform	16
3.2 Image multiscale geometric analysis.....	17
3.2.1 Ridgelet transform	18
3.2.2 Curvelet transform	19
3.2.3 Contourlet transform.....	19
3.3 Image sparse representation based on overcomplete dictionary	20
3.3.1 Overcomplete sparse representation theory	21
3.3.2 Construction of overcomplete dictionary.....	22
3.4 Applications of sparse representation in the field of image processing	23
Chapter 4 Low Bit Rate Facial Image Compression Based on K-PCA Algorithm.....	25
4.1 Sparse representation of signals.....	25
4.1.1 Greedy algorithm	25
4.1.2 Global optimization algorithm	27
4.1.3 Other algorithms	28
4.2 Construction of adaptive overcomplete dictionary	29
4.2.1 K-means clustering algorithm.....	29
4.2.2 K-SVD algorithm.....	31
4.2.3 K-PCA algorithm	34

4.3 Image denoising based on K-PCA algorithm and experiments	36
4.4 Facial images compression based on K-PCA algorithm and experiments..	39
4.5 Brief summary	45
Chapter 5 Facial Images Compression based on Adaptive Two Dictionaries	46
5.1 Facial images compression based on adaptive two dictionaries.....	46
5.1.1 Overview on super resolution image restoration	46
5.1.2 Image super resolution via sparse representation	49
5.1.3 Facial images compression based on adaptive two dictionaries	51
5.2 Realization of the algorithm and analysis of experiment results	53
5.3 Brief summary	59
Chapter 6 Conclusion and Expectation	61
References	62
Acknowledgements	65
Appendix	66

厦门大学博硕士论文摘要库

第一章 绪论

1.1 研究背景

随着科学技术的发展，人们对信息的需求日益增长，每天都有大量的信息用数字进行存储、处理和传输。信息时代带来了“信息爆炸”，使数据量大增，而图像具有直观、形象、易懂和信息量大的特点，是人们获取信息的重要途径，因此，如何有效的对图像进行处理成为存储和传输过程中要解决的一个重要问题。

数字图像处理最早出现于 20 世纪 50 年代，伴随电子计算机的发展，人们开始利用计算机来处理图像信息。常用的图像处理方法有图像增强、复原、编码、压缩等。目前，图像处理技术在许多应用领域受到广泛重视并取得了重大的开拓性成就，如航空航天、生物医学工程、工业检测、机器人视觉、公安司法、军事制导、文化艺术等，使图像处理成为一门引人注目、前景远大的新型学科。

图像信号内容复杂，数据量大，直接对其存储或传输会占用巨大的存储空间和带宽。如一幅大小为 512×512 ，每分量 8bits/pixel 的彩色图像需占据磁盘空间为 768KB，若以 24 帧/秒的速度来传送，那么一张 680MB 容量的 CD-ROM 仅能存储 30 多秒的原始数据，因此，迫切需要一种信息容量少又能满足视觉要求的存储和传输技术，即图像压缩必不可少。图像压缩就是用尽可能少的数据量来表示数字图像，减少数据量的关键在于去除图像数据中的冗余。在数字图像压缩中，可以确定以下三种基本的数据冗余并加以利用，当这三种冗余中的一种或多种得到减少或消除时，就实现了数据压缩。

1. 编码冗余。如果一个图像的灰度级编码，使用了多余实际的编码符号，就称该图像包含了编码冗余。
2. 像素间冗余。图像像素的灰度级是连续变化的，相邻像素的灰度值差别很小，单一像素对于一幅图像的多数视觉贡献是多余的，所以，它的值可以通过与其相邻的像素值为基础进行推测。
3. 心理视觉冗余。在正常视觉处理过程中，各种信息的相对重要程度不同，不重要的信息称为心理视觉冗余。去除心理视觉冗余信息并不会削弱图像的感知质量。

随着技术的不断进步,研究人员提出了多种图像压缩方法,各种图像压缩标准也不断被制定。现有的图像压缩方法大都基于信号非冗余的正交变换,如离散傅立叶变换(DFT),离散余弦变换(DCT),离散小波变换(DWT)等。这些方法将图像从空间域转换到变换域,然后对其进行量化、编码从而达到压缩的目的,有效的减少了图像表示的冗余,取得了良好的压缩效果,因而广泛应用于许多领域。

人脸图像,作为一种特殊的图像,是人类视觉中最常打交道的对象,也是诸如门禁系统、可视电话、会议电视等中主要传输的对象。由于人脸图像出现频繁、数据量巨大,因此研究人脸图像的低比特率压缩,是低带宽条件下的可视电话、会议电话进行信息传输的客观要求。同时,研究对人脸这一特定类图像的压缩算法,可以启发人们对其他类图像的压缩。但目前,还没有针对特定图像的压缩标准,而且传统的图像表示方法不能满足图像在低比特率压缩方面的要求,如离散余弦变换的基函数不具有时间/空间分辨率,仅适合于表示具有周期性特征的信号,小波变换的点状奇异特性使其在表示多维图像时性能明显下降。因此,迫切需要发展一种新的图像压缩方法,实现对人脸图像的低比特率压缩。

本文以静止的人脸图像为研究对象,研究了近些年逐渐兴起的信号稀疏表示(Sparse Representation)理论,提出了新的超完备字典训练算法,并将其应用到人脸图像低比特率压缩领域,同时针对如何进一步提高压缩比的问题,提出了基于自适应双字典的人脸图像压缩算法。

1.2 技术发展

图像压缩编码技术的研究可以追溯到1948年提出的电视信号数字化,至今已有60多年的历史了^[1]。在此期间出现了很多种图像压缩编码方法,特别是到了80年代后期,由于小波变换理论、分形理论、人工神经网络理论、视觉仿真理论等的建立,图像压缩技术得到了前所未有的发展。90年代以后,ITU(国际电信联盟)和ISO(国际标准化组织)制定了一系列图像编码国际标准,其中,最有代表性的静态图像编码标准有JPEG和JPEG2000,使得图像压缩在各个领域得到很大的发展。正如前面所介绍的,传统的图像压缩方法大都基于信号非冗余的正交变换,如JPEG是基于DCT, JPEG2000是基于DWT。JPEG2000相比于JPEG压缩性能更优,主要得益于DWT具有时频局域分析和多尺度特性,能

够较 DCT 更稀疏地表示图像。然而，由于完备正交基函数的单一性，导致其仅对图像中的某些特征能稀疏表示，而自然图像中往往包含有各种不同的特征，要得到图像的稀疏表示仅用一种完备的正交基是远远不够的。

为此，许多学者提出了多尺度几何分析（MGA: Multiscale Geometric Analysis）的方法，主要有 E. J. Candes 和 Donoho 等人提出的 Ridgelet 变换^[2]和 Curvelet 变换^[3]，Donoho 提出的 Wedgelet 变换^[4]和 Edgelet 变换^[5]，M. N. Do 等人提出的 Contourlet 变换^[6]，Francois G. Meyers 等人提出的 Brushlet 变换^[7]，E. Le Pennec 等人提出的 Bandlet 变换^[8]等等。这些变换的提出，都是为了克服小波变换仅对点状奇异性具有稀疏表示能力，而对边缘等一维及其他高维奇异性不能稀疏表示的缺陷。这些变换充分考虑了图像本身所具有的某些几何特征，因此与小波相比能获得更稀疏的图像表示。

虽然上述 MGA 方法能够较小波变换更好地捕获图像中的边缘等奇异特征，但上述方法中的每一种变换仅对图像中的某些特征表示是“有效（即稀疏）”的，而对于其他特征表示并不十分理想。如 Ridgelet 变换，Edgelet 变换等对图像中的线奇异性有效；Curvelet 变换、Contourlet 变换对图像中的曲线和轮廓有效；对于图像中的纹理特征，Brushlet 变换具有较好的表示性能。不同的变换在表示图像特征时，表现出不同的性能。而自然图像中往往包含各种特征，仅仅采用某一种变换很难有效地对图像中的所有特征进行稀疏表示。一种有效的办法是将多种基函数级联构造超完备字典（或称作原子库），在该字典上对图像进行稀疏分解，从而有效捕获图像中的各种特征。但这种方法产生的字典规模巨大，限制了其应用。另外一种办法是依据待研究图像的内容，通过训练、学习的方法自适应获取超完备字典，实现依据图像内容的稀疏表示。总之，为了稀疏表示图像，基于超完备字典的方法是一种趋势。

基于超完备字典的稀疏表示研究始于 20 世纪 90 年代，1993 年 Mallat 等人首次提出应用超完备 Gabor 字典对信号进行稀疏分解的思想，并引入了匹配追踪（Matching Pursuit, MP）算法^[9]。针对稀疏分解过程，已经发展了多种算法，如 MP 算法、OMP 算法、BP 算法、MOF 算法和 BOB 算法^[9-14]等等，目前图像稀疏表示理论已经被成功应用于图像处理的多个方面，如图像压缩^[15,16]、图像去噪^[17]、修复^[18]和识别^[19]等方面。特别是在图像压缩方面，无论是静态图像压缩，还是运动图像压缩，图像稀疏表示都表现出了巨大潜力，在低比特率条件下很多

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库