

关于稀疏编码理论及其应用

庄永文

(厦门大学 福建 厦门 361005)

摘要:稀疏编码的概念源于视神经网络的研究,是对只有一小部分神经元同时处于活跃状态的多维数据的神经网络的表示方法。稀疏编码理论在视神经细胞的响应特性和外部环境刺激的统计特性之间建立一种科学的数量联系,逐渐成为一种有效理解人类神经系统信息加工机制的理论工具,在盲源信号分离、语音信号处理、图像特征提取、自然图像去噪、以及模式识别等方面取得了许多成果,具有重要的实用价值。

关键词:稀疏编码;独立分量分析;图像处理;视觉神经网络

中图分类号:TN919.8

文献标识码:B

文章编号:1004-373X(2008)07-157-04

Sparse Coding Theory and Its Application

ZHUANG Yongwen

(Xiamen University, Xiamen, 361005, China)

Abstract: The concept of sparse coding comes from the study of visual neural network, it is a neural network method for finding a representation of multidimensional data in which each of the components of the representation is only rarely significantly active. Sparse code theory establishes a scientific quantitative link between the information processing mechanisms of visual neurons and the statistics of input visual stimuli, and provides an efficient tool to understand the neural information processing mechanisms. It has been applied in blind source separation, speech signal separation, image feature extraction, natural image denoising and pattern recognition, and it has achieved many results and has important practical value.

Keywords: sparse coding; independent component analysis; image processing; visual neural network

1 引言

图像处理一直是人类视觉研究中一个十分重要的领域,图像处理离不开对生物视觉的了解,如在图像压缩中离散余弦变化系数的量化,就是考虑到人对低频成份的敏感度比高频的强,从而对高频用大的量化步长,低频用较小的量化步长,使人眼看到经过压缩的图像仍有较高的质量。但是,大多数传统的图像处理方法是建立在数字信息处理和概率统计的基础上的,与人的视觉处理还有很大区别。

人类视觉感觉机制的研究表明,人眼视觉系统可看成是一种合理而高效的图像处理系统。对图像来说,相邻像素间的灰度值具有较高的相关性,即一幅图像中的某个像素的灰度值与周围的相邻灰度值通常相差不大,这一特性使得我们可以从部分图像预测整体。

神经生理研究表明,在初级视觉皮层下细胞的感受野具有显著的方向敏感性,单个神经元仅对处于其感受野中

的刺激做出反应,即单个神经元仅对某一频段的信息呈现较强的反应,如特定方向的边缘、线段、条纹等图像特征,其感受野被描述为具有局部性、方向性和带通性的信号编码滤波器。而每个神经元对这些刺激的表达则采用了稀疏编码原则,将图像在边缘、端点、条纹等方面的特性以稀疏编码的方式进行描述。从数学的角度来说,稀疏编码是一种多维数据描述方法,数据经稀疏编码后仅有少数分量处于明显激活状态,大致等价编码后的呈现超高斯分布。

2 稀疏编码的研究历史

(1) 1959 年,David hubel 和 Torsten Wiesel^[1]通过对猫的视觉条纹皮层简单细胞感受野的研究得出这样一个结论:视觉皮层 V1 区神经元的感受野能够对视觉感知信息产生一种“稀疏表示”。

(2) 1961 年,H. B. Barlow^[2]基于 David hubel 和 Torsten Wiesel 提出的结论,提出了“利用感知数据的冗余”进行编码的理论。

(3) 1969 年,D. J. Willshaw 和 O. P. Buneman^[3]提出了基于 Hebbian 学习的局部学习规则的稀疏表示模型。

(4) 1972 年,Barlow^[4]推论出在稀疏性和自然环境的统计特性必然存在某种联系。

收稿日期:2007-09-16

基金项目:国家高技术研究发展计划(863 计划)项目
(2006AA01Z129)

(5) 1987年, Field^[5]提出了主视皮层V1区简单细胞的感受野非常适于学习视网膜成像的图像结构。

(6) 1988年, Michison^[6]明确提出了神经稀疏编码的概念, 然后由牛津大学的 E. T. Roll 等人正式引用。

(7) 1989年, D. J. Field^[7]提出了稀疏分布式编码方法。这种编码方法使响应于任一特殊输入信息的神经细胞数目被减少, 信号的稀疏编码存在于细胞响应分布的四阶矩即峭度中。

(8) 1996年, B. A. Olshausen 和 D. J. Field^[8]在 Nature 发表一篇题为 Emergence of simple cell receptive field properties by learning a sparse code for natural images 的重要论文, 指出自然图像经过稀疏编码后得到的基函数类似于 V1 区简单细胞的感受野的反应特性。

(9) 1997年, B. A. Olshausen 和 D. J. Field^[9]又提出了一种超完备基的稀疏编码算法。

(10) 1997年, Bell 和 Sejnowski^[10]把 ICA 用于自然图像分析, 并且得出一个重要结论: ICA 实际上就是一种特殊的稀疏编码。

本世纪以来, 国外研究人员又提出了许多新的稀疏编码算法, 如 Joshua B. Tenenbaum 和 William T. Freeman^[11]的双线性稀疏编码模型, Bruno A. Olshause^[12]等提出的用小波金字塔结构学习学生图像编码, Aapo Hyvärinen 和 P. O. Hoyer^[13]提出的双层稀疏编码模型, Hoyer^[14]提出一种非负的稀疏编码模型, Donoho^[15]等人提出的经过 l^∞ 范数最小化的通用字典的最佳稀疏表示法等。国内学者也在稀疏编码算法和应用方面做出了一些工作, 如杨谦^[16]利用二维 Gabor 小波函数, 建议了一个基于超定完备基的简单细胞群稀疏编码的计算模型并实现了自然图像的编码。

3 稀疏编码的概念模型及原理

3.1 稀疏编码概念模型

从数学的角度, 稀疏编码目前被假设为是对多维数据进行线性分解的一种表示方法。数学描述如下:

假设输入数据 $X = (x_1, x_2, \dots, x_n)^T$ 为 n 维随机向量, 用 $S = (s_1, s_2, \dots, s_m)^T$ 表示线性转换后的 m 维向量, 假设 $m \leq n$, 那么线性转换矩阵则为 $m \times n$ 维, 记为 M 。则线性转换表示式为:

$$S = MX \quad (1)$$

上式中, M 又称稀疏变换矩阵, 其每一个行向量类似小波变化中的小波基, 线性转换后的稀疏分量 S 满足稀疏分布要求, 并且向量 (s_1, s_2, \dots, s_m) 尽可能地相互独立。

3.2 稀疏编码原理

Olshausen 和 Field 1996 提出了稀疏编码模型^[8], 稀疏编码模型是一个线性叠加模型, 他通过定义稀疏性约束来优化学习得到类似于简单细胞响应特性的基函数。下面从两方面来解释稀疏编码原理:

(1) 对作用于同一刺激的神经元群来说, 稀疏编码指

在作用于同一刺激的神经元群中, 并不是所有的神经元都被激活, 相反, 只有极少数的神经元被激活。如图 1 所示, 大黑斑表示神经元群中被激活的神经元。生物学的研究成果表明, 稀疏编码是一种新陈代谢能量较少的信息处理策略。

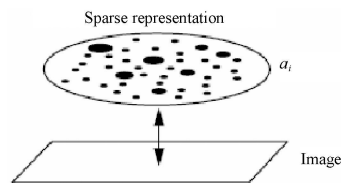


图1 神经元群稀疏编码示意图

(2) 对作用于输入刺激集合的单个神经元来说, 稀疏编码指该神经元响应的分布具有稀疏特性。根据以上所述, 一个神经元群中的神经元在大部分时间都是不激活的, 很自然地, 单个神经元响应的概率分布将在 0 附近有一个峰, 而尾部很扁平, 如图 2 中的实线所示, 正态分布的曲线与之相比要平滑得多。从信息论的角度看, 在具有相同均值和方差的概率分布中, 正态分布具有最大的熵, 而稀疏分布的熵则少得多。

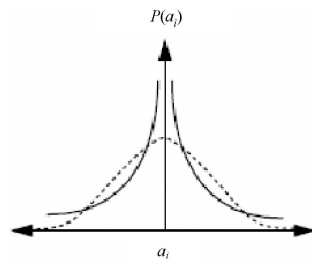


图2 单个神经元稀疏编码示意图

图 2 中实线表示神经元响应的概率分布, 而虚线是具有相同均值和方差的正态分布。

4 独立分量分析

ICA 源于盲源分离问题 (Blind Source Separation, BSS) 的解决。ICA 的起始点基于一个非常简单的假设, 即源信号 S 统计独立, 同时假定源变量为非高斯分布。其基本目标就是寻找一个线性变换, 使得 m 维观测变量 $X = (x_1, x_2, \dots, x_m)^T$ 经过线性变换后, 输出信号 $Y = (y_1, y_2, \dots, y_n)^T$ 之间尽可能保持独立。则 Y 就是源信号 $S = (s_1, s_2, \dots, s_n)^T$ 的近似解, X, Y 和 S 之间的关系如下:

$$X = AS \quad (2)$$

$$Y = WX \quad (3)$$

当 ICA 用于处理自然图像时, 就是一种特殊的稀疏编码方式。

5 稀疏编码的研究现状及存在问题

稀疏编码模型能够表示初级视觉皮层简单细胞外界视觉刺激图像的过程和特征, 但由于哺乳动物的视觉生理

过程很复杂,在用稀疏编码模型模拟初级视觉系统神经元的感受野时,通常做如下假设:

- (1) 输入数据具有稀疏性结构
- (2) 各稀疏系数 S_i 之间是相互对立的
- (3) 稀疏变换是一种线性变换
- (4) 基函数 A 是非奇异的
- (5) 有噪声时,稀疏分量与噪声信号是相互独立的

目前所做的研究都是基于上述假设的稀疏编码模型,研究内容主要有以下几方面:

(1) Olshausen^[9] 等人把超完备基引入稀疏编码模型,利用概率密度估计模型建模 V1 区感受野,并应用于时变的自然图像数据。

(2) Tenenbaum^[13] 等人提出了双线性稀疏编码模型。

(3) Olshausen 和 Simocelli^[17] 从小波变换的角度对自然数据进行神经表示和有效编码。

(4) Hoyer^[14] 提出一种非负的稀疏编码模型。

(5) Donoho^[15] 等人提出的经过 l^∞ 范数最小化的通用字典的最佳稀疏表示法。

(6) 杨谦^[16] 建立了一个基于超定完备基的简单细胞集群稀疏编码的计算模型,并实现了自然图像的稀疏编码。

由于人们目前对主视皮层 V1 区如何对复杂的自然环境进行有效的编码的过程仍然知之甚少,上述稀疏编码模型都是一种启发式的学习方法。稀疏编码理论主要存在以下两个问题:

(1) 理论依据不完善:研究主视皮层 V1 区的稀疏编码首先必须借助感知神经科学的研究成果。另外,必须结合计算技术、统计学理论、估计理论、人工神经网络等基础理论。

(2) 神经元估计模型不精确:稀疏编码算法必须依靠一个好的神经元的计算模型,这个计算模型必须建立在神经生理科学研究的成果上。

6 稀疏编码的应用研究

稀疏编码方法在盲源信号分离^[18]、语音信号处理^[19]、自然图像特征提取^[20]、自然图像去噪^[21]、以及模式识别^[22]等方面取得了许多成果,具有重要的实用价值。

许多研究表明,自然图像在统计上具有冗余特性,用稀疏编码方法可以实现自然图像的特征提取以及消除图像中的高斯噪声,此方法和 ICA 非常相似。标准稀疏编码模型对于描述哺乳动物主视皮层 V1 区的简单细胞活动时仍存在不足之处,一种把多维数据分解成非负稀疏分量的非负稀疏编码算法有效地解决了这个问题。

7 结 语

人的视觉系统是一个强大而复杂的信息处理系统,人

们由客观获得图像信息的主要来源约占视觉信息的 80%,因此模拟视觉系统视皮层神经元输出特性的稀疏编码理论是一个十分重要的领域。目前人们对主视皮层 V1 区仍了解甚少,所以有关 V1 区神经元编码的研究仍处于探索阶段。稀疏编码从最初的概念雏形到理论的形成已有快 50 年的历史,在各个领域取得了重大成果。稀疏编码理论必须结合计算技术、统计学理论、估计理论、人工神经网络等基础理论,又必须建立在神经生理科学研究的成果上,其中还有很多尚未解决或尚未充分解决的问题,这是一个十分值得研究的领域。

参 考 文 献

- [1] Hubel D H, Wiesel T N. Receptive Fields of Single Neurons in the Cat's Striate Cortex[J]. Journal of Physiology, 1959, 148: 547-91.
- [2] Barlow H B. The Coding of Sensory Messages. Current Problems in Animal Behaviors, Chapter XIII, Cambridge University Press, 1961: 331-360.
- [3] Willshaw D J, Buneman O P, Longuet Higgins H C. Nonholographic Associative Memory[J]. Nature, 1969, 222(5): 960-962.
- [4] Barlow H B. Single Units and Sensation: A Neuron Doctrine for Perceptual Psychology[J]. Perception, 1972: 371-379.
- [5] Field D J. Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells[J]. Journal Optical Society, 1987(4): 2 379-2 394.
- [6] Michison G. The Organization of Sequential Memory: Sparse Representations and the Targeting Problem. Organization of Neural Networks, VCH Verlagsgesellschaft, Weinheim, 1988: 347-367.
- [7] Field D J. What the Statistics of Natural Images Tell Us about Visual Coding[J]. Proc. the International Society for Optical Engineering(SPIE), 1989, 1077: 269-276.
- [8] Olshausen B A, Field D J. Emergence of Simple cell Receptive Field Properties by Learning a Sparse Code for Natural Images[J]. Nature, 1996, 381: 607-609.
- [9] Olshausen B A, Field D J. Sparse Coding with an Overcomplete Basis Set: A Strategy Employed by V1 Vision Research, 1997, 37: 3 311-3 325.
- [10] Bell A J, Sejnowski T J. The Independent Components of Natural Scenes are Edge Filters[J]. Vision Research, 1997, 37: 3 327-3 338.
- [11] Joshua B Tenenbaum, William T Freeman. Separating Style and Content with Bilinear Models[J]. Neural Computation, 2000, 12: 1 247-1 283.
- [12] Bruno A Olshausen, Phil Sallee, Michael S Lewicki. Learning Sparse Image Codes Using a Wavelet Pyramid Architecture[J]. Advances in Neural Information Processing Systems, 2001, 13: 887-893.

[13] Aapo Hyvärinen, Hoyer P. O. A Two-layer Sparse Coding Model Learns Simple and Complex cell Receptive Fields and Topography from Natural Images [J]. Vision Research, 2001, 41(18) : 2 413-2 423.

[14] Hoyer P. O. Modeling Receptive Fields with Non-negative Sparse Coding. Computational Neuroscience: Trends in Research, 2003, Elsevier, Amsterdam, 2003.

[15] Donoho D L, Michael Elad. Optimally Sparse Representation in General Dictionaries via l^∞ Minimization[J]. Proceedings of the National Academy of Sciences (PNAS), 2003, 100(5) : 2 197-2 202.

[16] 杨谦, 齐翔林, 汪云九. 视皮层 V1 区简单细胞的稀疏编码策略[J]. 计算物理, 2001, 18(2) : 136-143.

[17] Simoncelli E P. Vision and the Statistics of the Visual Environment[J]. Current Opinion in Neurobiology, 2003, 13(2) : 144-149.

[18] 李清勇, 史忠植. 视觉感知的稀疏编码理论及其应用研究[D]. 北京: 中国科学院研究生院, 2006.

[19] Tuomas Virtanen. Sound Source Separation Using Sparse Coding with Temporal Continuity Objective. Proc. ICMC, 2003.

[20] Oja E, Hyvärinen A, Hoyer P. Image Feature Extraction and Denoising by Sparse Coding[J]. Pattern Analysis & Applications, 1999, 2: 104-110.

[21] Hyvärinen A. Sparse Coding Shrinkage: Denoising of Non-gaussian Data by Maximum Likelihood Estimation[J]. Neural Computation, 1999, 11(7) : 1 739-1 768.

[22] Palmer S E. Vision Science: Photons to phenomenology [M]. London, MIT Press, 1999.

[23] 尚丽, 黄德双. 稀疏编码算法及其应用研究[D]. 合肥: 中国科学技术大学, 2007.

[24] Li Y, Cichocki A, Amari S. Sparse Component Analysis for Blind Source Separation with Less Sensors than Sources. Fourth International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003) Japan, 2003: 89-94.

作者简介 庄永文 男, 硕士研究生。主要研究方向为智能信息处理。

(上接第 156 页)

参 考 文 献

[1] Yang G Z, Huang T S. Human Face Detection in a Complex Background[J]. Pattern Recognition, 1994, 27(1) : 53-63.

[2] Turk M, Pentland A. Eigenfaces for Recognition[J]. Cognitive Neuroscience, 1991, 3(1) : 71-79.

[3] Lin S H. Face Recognition/Detection by Probabilistic Decision-based Neural Network [J]. IEEE Trans. Neural Networks, 1997(8) : 114-132.

[4] Reisfeld D, Wolfson H, Eshurun Y Y. Context-free Attentional Operators: The Generalized Symmetry Transform[J]. Computer Vision, 1995, 14: 119-130.

[5] 彭进业, 俞卞章, 王大凯. 多尺度对称变换及其应用于定位人脸特征点[J]. 电子学报, 2002, 30(3) : 363-366.

[6] Kawato S, Ohya J. Two-step Approach for Real-time Eye Tracking with a New Filtering Technique[A]. IEEE International Conference on Systems, Man & Cybernetics, 2000.

[7] 王忠, 胡步发, 严世榕. 一种改进的对称变换应用于人脸图像眼睛定位[J]. 计算机应用, 2004, 30(11) : 119-121.

[8] Garcia C, Zikos G, Tziritas G. Face Detection in Color Images Using Wavelet Packet Analysis [EB/OL]. <http://www.csd.ucl.ac.uk/tziritas/papers/ICMCS99.pdf>. 2004.

作者简介 陈洁 女, 1982 年出生, 湖北荆门人, 硕士研究生。主要从事图像处理与传输, 人工智能与模式识别等方面的研究。

戴尔可能重回掌上设备市场

DigiTimes 报道, 据台湾业内人士消息, 戴尔可能重新回归掌上设备市场, 并联合富士康开发基于 Windows Mobile 平台的新产品。

去年年初, 戴尔将前摩托罗拉执行副总裁 Ron Garriques 招至麾下, 开始对掌上设备部门进行彻底重组。4 月, 他们的最后一款 PDA 产品 Axim X51 宣布停产, 至此彻底推出掌上设备市场。在经过一年的蛰伏后, 戴尔看来要重新出山了。

据称, 富士康日前也开始从其他台湾手机制造商中寻觅人才, 为的就是准备接受戴尔掌机订单。尽管富士康集团的两家关系企业诚实科技和奇美通讯已经在为其他手机厂商进行智能手机代工。但戴尔希望直接同富士康进行合作, 生产包括 GPS, PDA 和智能手机等产品。

(摘自《驱动之家》)