

一种基于三维自组织映射的立体图像编码方法

覃远年, 徐晓宁

(桂林电子科技大学 信息与通信学院, 广西 桂林 541004)

摘 要: 立体图像的数据量成倍增加, 在图像信息的传输、存储方面遇到很大挑战。自组织映射(SOM)是一种非常有效的聚类算法, 在图像处理、数据可视化等领域得到了广泛的研究与应用。为了提高左图像编码效率提出了一种基于三维频域敏感自组织映射(3DFSSOM)的立体图像编码算法。该算法对左图像进行基于 3DFSSOM 的矢量量化编码; 对矢量量化(VQ)和视差估计(DE)的残差均进行 JPEG 编码; 对视差矢量进行霍夫曼编码。此外, 在 3DFSSOM 算法中引入不等式判据, 来降低失真测度的计算复杂度。通过实验表明: 左图像的编码效率得到了很大改善, 计算复杂度明显降低。

关键词: 立体图像; 自组织映射; 矢量量化; 视差估计

中图分类号: TP183

文献标识码: A

文章编号: 1000-7180(2015)11-0011-04

A Stereo Image Coding Method Based on Three-Dimensional Self-Organizing Map

QIN Yuan-nian, XU Xiao-ning

(College of Information and Telecommunication, Guilin University of Electronic Technology, Guilin, 541004)

Abstract: The data volume of stereo image increases exponentially, so the transmission and storage of image information encounter a great challenge. The self-organizing map is a very efficient clustering algorithm, it is widely researched and applied in image processing, data visualization and other fields. In order to improve the coding efficiency of the left image, this paper presents a stereo image coding algorithm based on the 3D frequency sensitive self-organizing map. Left image is encoded by the vector quantization based on 3DFSSOM in this algorithm. These residuals are encoded by the JPEG, which are obtained by the vector quantization and the disparity estimation. Disparity vector is encoded by the Huffman code. In addition, the proposed algorithm reduces the computational complexity of distortion measure by introducing inequality criterion in 3DFSSOM algorithm. The experiments show that coding efficiency of the left image is improved greatly, the computational complexity of algorithm is also reduced significantly.

Key words: stereo image; self-organizing map; vector quantization; disparity estimation

1 引言

立体图像增加了景物的深度信息, 具有强烈的临场感和真实感, 因此越来越受到人们的追捧。对于立体图像, 目前应用最广泛的是双目视图方案, 利用双目视差原理产生立体感^[1]。与二维平面图像相比, 由于立体图像数据量成倍地增加, 因此, 需要对立体

图像进行压缩编码, 以便能有效地传输和存储。目前主要的立体图像压缩算法有基于块法^[2]、基于小波法^[3]和基于对象法^[4]。在实际应用中, 常采用视差估计补偿算法与国际编码标准 MPEG 相结合的方法对立体图像进行编解码。这种方法对左图像采用基于 JPEG 的编码标准进行编解码, 由左图像和右图像利用视差估计补偿技术得到视差估计误差和视差

收稿日期: 2014-12-27; 修回日期: 2015-02-16

基金项目: 国家自然科学基金地区科学基金项目(61261035)

矢量,对视差估计误差使用类似于 JPEG 标准的方法编码,对视差矢量进行霍夫曼编码^[5].

SOM 算法是基于生物神经网络理论进行训练学习的高效聚类算法,在图像编码^[6]等领域深受学者们的青睐,得到了广泛的应用.在模式匹配过程中,寻找性能最优的码书是 SOM 算法的关键环节.为了提高码书性能,学者们对基本的 SOM 算法进行了大量的改进工作. U. Seiffert 与 B. Michaelis^[7]对算法输入层和竞争层进行了改进,文献^[4]中提出了一种三维 SOM 算法,并对码书邻域结构进行了分析.黎洪松^[8]提出了一种频域敏感自组织映射(FSSOM)算法,通过引入频率敏感函数降低了获胜码矢再次成为响应码矢的可能,使码书中每个码矢都能得到充分利用. Xie Song-zhao 等人^[9]在码书设计过程中,应用不等式判据对基本 SOM 算法进行了改进,结合小波变换和矢量量化对图像进行编码,提高了图像压缩比,降低了算法计算复杂度.

本文提出了一种基于三维频域敏感自组织映射的立体图像编码算法.与传统的使用 JPEG 标准对左图像进行编码方案不同的是,本文采用 3DFSSOM 算法对左图像进行矢量量化编码,由左图和右图经视差估计补偿算法得到视差估计误差,矢量量化误差和视差估计误差均使用离散余弦变换(DCT)加霍夫曼进行编码,对视差矢量进行霍夫曼编码.实验结果表明,本文提出的算法大大缩短了码书训练时间,获得的码书性能更优异,同时左图像的编码效率也得到了很大程度的改善.

2 编解码方案

图 1 中给出了基于 3DFSSOM 算法的立体图像编码框图,相应的解码框图如图 2 所示,算法具体描述如下所示.

步骤 1:对左右图像分块预处理,并使用 3DFSSOM 算法对其进行学习训练,获得性能最优的匹配模式码书.

步骤 2:将左图作为输入矢量,进行基于 3DFSSOM 的矢量量化编码,由左图和矢量量化码矢获得矢量量化误差,并对其进行 DCT 变换、量化和霍夫曼编码.

步骤 3:对编码后的码流进行解码,由解码后的矢量量化误差和矢量量化码矢对左图进行重建,即将解码后得到的矢量量化误差和矢量量化码矢进行求和匹配运算,得到重建左图.

步骤 4:对右图进行分块处理,并将左图当做参考图像利用视差估计补偿技术获得右图的预测图像及视差矢量.

步骤 5:由右图的预测图像和右图进行做差运算得到视差估计误差,对视差估计误差图像分块进行 DCT 变换、量化和霍夫曼编码.

步骤 6:在解码端由重建左图、视差估计误差和视差矢量对右图进行重建,从而实现了立体图像的有效编解码.

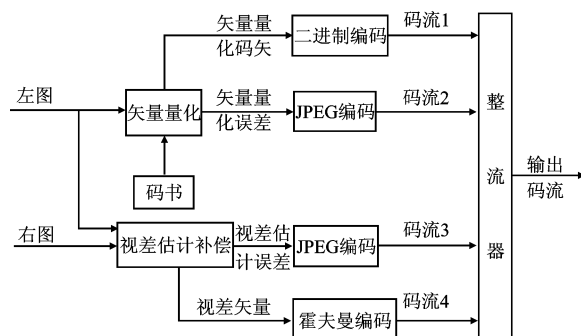


图 1 本文算法编码框图

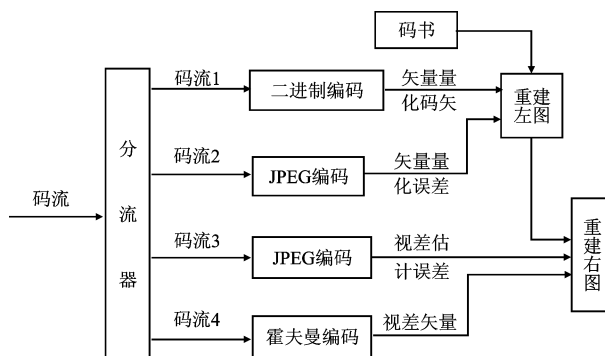


图 2 本文算法解码框图

3 三维 SOM 算法原理及步骤

3.1 三维 SOM 算法原理

如图 3 所示,三维邻域 SOM 网络由二维输入层和三维输出层组成.输出层码书 N 被排列成 $N = a * b * c$ 三维立体结构,合理地对 3 个参数进行分配能有效地提高算法性能.输出层也叫做竞争层,输入层的每个节点与竞争层所有神经元之间通过权值向量来进行联系,若某个神经元在竞争学习过程中获胜了,则该神经元的权值向量就会得到调整,同时它的邻域内神经元的权值向量也会随之调整.神经元拓扑结构图和每个神经元的所有权值在所有输入样本矢量训练结束后得到稳定,图像压缩时所需的码书就由获得的权值矢量构成.

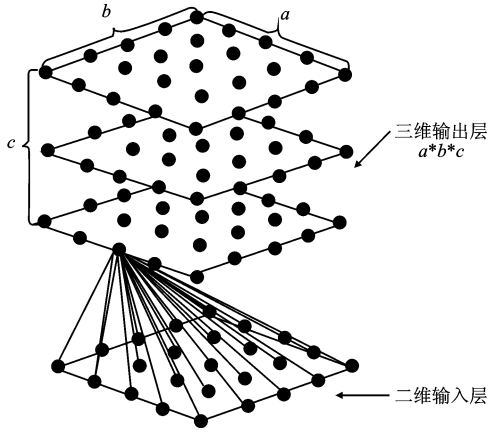


图3 三维 SOM 网络结构

3.2 三维 FSSOM 算法实现步骤

设定网络拓扑结构中输入训练矢量的大小 M , 输出层输出神经元的个数为 N , 输入训练矢量的集合可表示为 $S = \{X_1, X_2, \dots, X_M\}$, 网络输入节点为 k (等于矢量维数), 各输入节点到各输出节点的权值为 $W_{jl}, l = 1, 2, \dots, k, j = 1, 2, \dots, N$, 失真测度为 $d(x_i, w_j)$, 引入频率敏感函数对失真测度进行修正, 以此来降低获胜码矢再次成为响应码矢的可能, 避免“死神经元”的出现, 即

$$d(x_i, w_j) = f(c_j) \sum_{l=1}^k [x_{il} - w_{jl}(t)]^2, 1 \leq j \leq N \quad (1)$$

式中, $f(c_j) = c_j^s$, s 为频率敏感参数, c_j 为响应计数器, 每响应一次 c_j 加 1, t 为迭代次数。

为了降低失真测度的计算复杂度, 使用一个距离不等式判据对失真测度进行修正, 定义一个输入矢量的和值为 $S_{X_i}, i = 1, 2, \dots, M$, 码字 W_j 的和值为 S_{W_j} , 即 $S_{X_i} = \sum_{l=1}^k X_{il}, S_{W_j} = \sum_{l=1}^k W_{jl}$, 容易证明:

$$\sum_{l=1}^k [X_{il} - W_{jl}(t)]^2 \geq (S_{X_i} - S_{W_j})^2 / k \quad (2)$$

由于 $f(c_j) \geq 1$ 所以式(3) 仍然满足, 即:

$$d(x_i, w_j) = f(c_j) \sum_{l=1}^k [X_{il} - W_{jl}(t)]^2 \geq (S_{X_i} - S_{W_j})^2 / k \quad (3)$$

设当前的最小失真为 $d_{\min}$, 并令 $MD = kd_{\min}$, 若

$$(S_{X_i} - S_{W_j})^2 \geq MD \quad (4)$$

则根据式(3) 可得

$$d(x_i, w_j) \geq d_{\min} \quad (5)$$

因此可预先计算码字的和值 S_{W_j} 和 MD 的值, 然后在搜索获胜神经元的过程中判断码字 W_j 的和值 S_{W_j} 是否满足式(4), 若满足则排除相应码字的距

离计算, 降低了算法的计算复杂度。

改进的 3DFSSOM 算法具体实现步骤如下所示。

步骤 1 采用基于方差分类排序法对码书进行初始化 $\{W_{jl}(0) = w_{jl}, l = 1, 2, \dots, k, j = 1, 2, \dots, N\}$ 。

步骤 2 将当前码书中各码矢 $w_{jl}(t)$ 的和值 S_{W_j} 和输入训练矢量 X_i 的和 S_{X_i} 分别计算出来, 同时在每一个神经元中, X_i 要以并行方式输入。

步骤 3 依据式(4) 引入的不等式判据以及式(1) 失真测度的表达式, 搜索具有最小失真的码矢 j^* 作为响应码矢, 即:

$$d(x_i, w_j) = f(c_j) \sum_{l=1}^k [X_{il} - W_{jl}(t)]^2, 1 \leq j \leq N.$$

选择具有最小失真的模式矢量 j^* , 即

$$\tilde{d}_{j^*}(t) = \min d(x_i, w_j), j = 1, 2, \dots, N \quad (6)$$

步骤 4 调整响应码矢 j^* 及 j^* 的拓扑邻域 NE 范围内的码矢

$$w_{jl}(t+1) = \begin{cases} w_{jl}(t) + a(c_j) [x_{il}(t) - w_{jl}(t)], & j \in j^*, NE_{j^*}(t) \\ w_{jl}(t), & \text{其他} \end{cases} \quad (7)$$

式中, $NE_{j^*}(t)$ 为码矢 j^* 的邻域, 邻域函数通常为

$$NE_{j^*}(t) = A_1 + A_2 e^{-t/T_1} \quad (8)$$

$a(c_j)$ 为学习速度, 一般选用单调递减函数:

$$a(c_j) = A_3 e^{-c_j/T_2} \quad (9)$$

步骤 5 返回算法的步骤 2, 直到训练完所有的输入矢量。

4 码书初始化

初始码书的选取方法分离平均法、随机法、分裂法^[10]等。本文采用基于方差排序的码书初始化算法。首先根据方差将码书分为低频和高频两类, 低频码书用于图像平坦区域, 高频用于图像细节区域。然后在高低频码书中分别等间隔抽训练矢量, 抽取的训练矢量的集合即为初始码书。具体步骤如下所示。

步骤 1 将训练图像序列进行分块组成训练矢量集, 计算各训练矢量的方差 $\text{var}(X)$ 。

步骤 2 根据方差对训练矢量集进行排序, 并记录各矢量排序前的原始位置。

步骤 3 设置一个阈值, 根据方差与设定阈值的大小关系把训练集分成低频 X_L 和高频 X_H 两部

分。低频部分为方差低于设定阈值的训练集;高频部分为方差高于设定阈值的训练集。

步骤 4. 计算总的训练矢量中低频和高频子矢量分别所占比例,按照相同比例在高低频部分等间隔抽取训练矢量组成初始码书中的码矢。

5 实验结果

矢量量化过程中码书训练采用立体图像对“plant”和“trees”,大小分别为 480×480 和 472×472 。评价图像质量的客观标准采用峰值信噪比 $PSNR = 10 \lg \frac{255^2}{MSE}$,表达式中 MSE 为均方误差(原图像与重建图像之间)。左图像压缩比 $CR_l = \frac{B_0}{B_a + B_e}$,其中 B_0 为原图像数据比特数, B_a 为矢量量化码矢比特数, B_e 为矢量量化误差编码比特数。右图像压缩比 $CR_r = \frac{B_0}{B_v + B_e}$,其中 B_0 为原图像数据比特数, B_v 为视差矢量编码比特数, B_e 为视差估计误差编码比特数。矢量量化地址码比特数 $B_a = \frac{M \times N}{m \times n} \times \log_2 L$,其中 M 和 N 原始图像像素值, m 和 n 表示所选取的图像块尺寸, L 为码书的尺寸。

与传统的 JPEG 标准结合视差估计算法的编码方案以及文献[11]提出的 2DSOM 矢量量化结合视差估计算法的编码方案相比,本文提出的 3DFSSOM 矢量量化结合视差估计算法的编码方案的改进主要体现在左图像的编码效率上。表 1 和表 2 中给出了“plant”和“trees”立体图像对采用上述三种方案对左图像进行编码的结果。实验结果表明:本文提出的基于 3DFSSOM 矢量量化的编码方案明显提高了左图像的编码效率。对于“plant”左图像,与传统的 JPEG 算法相比,重建左图的质量(PSNR)都有所提高,平均提高了 2.23 dB。与 2DSOM+JPEG 算法相比,重建图像的质量也都有所提高,平均提高了 0.98 dB。对于“trees”左图像,与传统的 JPEG 算法相比,重建图像的质量平均提高了 3.07 dB。与 2DSOM+JPEG 算法相比,重建图像的质量平均提高了 0.86 dB。由表 2 还可以得到,虽然 3DFSSOM 算法较 2DSOM 算法网络结构更复杂,但在引入不等式判据后降低了算法复杂度,编码所需时间大大低于 2DSOM 算法。

6 结束语

本文基于三维自组织映射算法进行矢量量化,

表 1 JPEG 算法与本文算法编码结果比较

图像	JPEG 算法		3DFSSOM+JPEG 算法	
	CR_l	PSNR/dB	CR_l	PSNR/dB
Plant 左图	19.64	33.92	19.86	36.08
	21.89	33.08	21.79	35.57
	24.05	32.33	24.03	34.31
	26.85	31.46	27.26	33.78
trees 左图	19.25	35.83	19.17	38.58
	22.28	34.81	22.47	38.39
	25.32	33.97	25.44	37.53
	28.14	33.24	28.13	35.66

表 2 DSOM+JPEG 算法与本文算法编码结果比较

图像	2DSOM+JPEG 算法			3DFSSOM+JPEG 算法		
	CR_l	PSNR/dB	TIME/s	CR_l	PSNR/dB	TIME/s
Plant 左图	19.57	35.14	247.23	19.86	36.08	173.46
	21.87	34.88	248.39	21.79	35.57	178.65
	23.94	33.36	207.57	24.03	34.31	155.07
	26.67	32.43	205.70	27.26	33.78	153.84
trees 左图	19.35	38.26	304.26	19.17	38.58	209.97
	22.38	37.42	303.44	22.47	38.39	208.30
	25.41	36.18	274.81	25.44	37.53	187.23
	28.08	34.85	248.89	28.13	35.66	177.06

同时结合视差估计补偿算法提出了一种新的立体图像编码算法。通过实验仿真表明本文算法能有效改善左图像的编码效率,同时在码书设计过程中引入了距离不等式判据降低了算法计算复杂度。下一步工作是研究如何将三维自组织映射算法应用到立体视频的运动估计中;研究更高维的自组织映射算法,从而获得性能更优异的矢量量化码书;研究精度更高的视差估计补偿算法以及残差编码算法,从而提高右图像的编码效率。

参考文献:

- [1] Ziegler M, Falkenhagen L, Horst R T, et al. Evolution of stereoscopic and three-dimensional video[J]. Signal Processing: Image Communication, 1998, 14 (1): 173-194.
- [2] Yang Qinyou, Rong Xianwei, Wang Haipeng. Based on false parallax variable block-size stereo image coding method[J]. Advanced Materials Research, 2013(664): 1055-1059.

(下转第 20 页)

- [2] Masuda N. Projection lens shifting mechanism; U. S. Patent 6,547,402[P]. 2003-04-15.
- [3] Li B, Sezan I. Automatic keystone correction for smart projectors with embedded camera[C]// Image Processing, 2004. ICIP04. 2004 International Conference on. Singapore, IEEE, 2004(4): 2829-2832.
- [4] Li Z, Wong K H, Gong Y, et al. An effective method for movable projector keystone correction[J]. Multimedia, IEEE Transactions on, 2011, 13(1): 155-160.
- [5] Wolberg G. Digital image warping[M]. Los Alamitos; IEEE computer society press, 1990.
- [6] Hearn D, Baker M P. Computer graphics[M]. Upper Saddle River; Prentice Hall, 1997.
- [7] Lanman D, Taubin G. Build your own 3D scanner; optical triangulation for beginners[C]// ACM SIGGRAPH ASIA 2009 Courses. Seoul, ACM, 2009: 2.
- [8] Chen P Y, Huang C C, Shiau Y H, et al. A VLSI implementation of barrel distortion correction for wide-angle camera images[J]. Circuits and Systems II: Express Briefs, IEEE Transactions on, 2009, 56(1): 51-55.
- [9] Gribbon K T, Bailey D G. A novel approach to real-time bilinear interpolation[C]// Field-Programmable Technology, 2004 IEEE International Conference on. Singapore, IEEE, 2004: 126-131.
- [10] Barenbrug B, Peters F J, Van Overveld C. Algorithms for division free perspective correct rendering[C]// Proceedings of the ACM Siggraph/Eurographics workshop on Graphics hardware. Switzerland, ACM, 2000: 7-13.
- [11] Kim D, Kim L S. Area-efficient pixel rasterization and texture coordinate interpolation[J]. Computers & Graphics, 2008, 32(6): 669-681.

作者简介:

任力飞 男, (1989-), 硕士研究生, 研究方向为数字集成电路设计和图像处理. E-mail: 1003889037@qq. com.

谢翔 男, (1971-), 副研究员. 研究方向包括集成电路设计、图像处理、生物医学电子等.

李国林 男, (1970-), 副教授. 研究方向包括集成电路设计、互连分析以及 RFIC 设计.

(上接第 14 页)

- [3] Suliman A, Li R. New stereo video coding in redundant wavelet domain[C]// Visual Communications and Image Processing (VCIP). Tainan, 2011: 1-6
- [4] Shiping Zhu, Kamel Belloulata. A novel object based fractal stereo video CODEC[C]// IEEE International Conference on Image Processing, ICIP2005. Magazzini, 2005: 493-497.
- [5] Andreas Grunz, Daniel Memmert, Jurgen Perl. Tactical pattern recognition in soccer games by means of special self-organizing maps[J]. Human Movement Science, 2012, 31(2): 334-343.
- [6] Kathirvalavakumar T, Ponmalar E. Self-organizing map and wavelet based image compression[J]. International Journal of Machine Learning and Cybernetics, 2013, 4(4): 319-326.
- [7] Seiffert U, Michaelis B. Adaptive three-dimensional self-organizing map with two-dimensional input layer[C]// Intelligent Information Systems, Australian and New Zealand Conference. New Zealand, 1996.
- [8] 黎洪松. 一种新的自组织神经网络算法[J]. 北京师范大学学报: 自然科学版, 2005, 41(5): 496-498.
- [9] Xie S, Wang C, Cui C. Vector quantization image coding based on biorthogonal wavelet transform and improved SOFM[C]// Information Science and Technology (ICIST), 2013 International Conference on. China, Yangzhou, IEEE, 2013: 1629-1632.
- [10] Huilin Xiong, Swamy M N S, Ahmad M O. Competitive splitting for codebook initialization[J]. Signal Processing Letters, IEEE, 2004, 11(5): 474-477.
- [11] 李达, 黎洪松. 一种改进的立体图像编码算法[J]. 北京师范大学学报: 自然科学版, 2007, 43(1): 50-52.

作者简介:

覃远年 男, (1971-), 硕士, 高级实验师. 研究方向为信号处理与无线通信系统.

徐晓宁 (通讯作者) 男, (1987-), 硕士研究生. 研究方向为图像视频编解码. E-mail: vipxuxiaoning@126. com.