

【本文信息】祝建英,夏哲雷,殷海兵,等. 基于神经网络的视频编码量化参数选择算法[J]. 电视技术 2012 36(19).

基于神经网络的视频编码量化参数选择算法

祝建英,夏哲雷,殷海兵,华 强

(中国计量学院 信息工程学院,浙江 杭州 310018)

【摘 要】提出了一种基于神经网络的视频量化参数选择,利用已编码帧信息对当前帧的量化参数计算进行调整。神经网络模型选取当前帧的目标比特、当前帧的 MAD 和前一帧编码的 QP 和实际比特数作为输入,输出值经过反归一化及取整处理对视频进行量化。实验在 H.264/AVC 的标准参考软件 JM14.2 中实现,编码时使用该方法在保证码率精确度和平均 PSNR 值基本不变的情况下,减小了 PSNR 的波动。

【关键词】量化参数;神经网络;BP

【中图分类号】TN919.81

【文献标识码】A

Algorithm of Selecting Video Quantization Parameter Based on Neural Network

ZHU Jianying, XIA Zhelei, YIN Haibing, HUA Qiang

(College of Information Engineering, China Jiliang University, Hangzhou 310018, China)

【Abstract】This paper presents QP selection model based on neural network, which uses the information of encoded frame to adjust the QP of current frame. The network picks four features which affect QP most as inputs, then quantizes the video by the output values. Experimental results show less PSNR fluctuation and the same R-D performance, using the proposed model instead of quantization model in JM14.2 reference software.

【Key words】quantization parameter; neural network; BP

视频压缩即在限定速率下得到最优编码。H.264 视频编码标准中,率失真优化(Rate Distortion Optimization, RDO)技术选择一系列编码参数,实现码率和失真之间恰当的折中。码率控制与率失真优化密切相关,是率失真优化技术的典型应用。因此,码率控制作为整个编码器的重要组成部分,是衡量视频编码器性能的重要指标。码率控制算法是将量化参数、实际比特率、编码复杂度之间的平衡关系模型化。其中量化参数(QP)、GoP 结构、目标比特率等参量,都对视频解码质量有影响^[1]。在低码率传输的视频编码中,合理设定每一帧的量化参数,不仅可以节省带宽,利用本就稀少的网络资源,而且可以在有限的条件下解码得到较为优质的图像复原质量。量化参数的设定是一个相对复杂的问题,在 H.264 的参考软件 JM 中就有众多的相关参数。

在码率控制的现有方案中,JVT-G012 作为目前的主流算法,使用 MPEG-4 的二次模型来调节 QP,同时引入基本单元的概念,将码率控制分为 GoP 层、帧层和基本单元层。作为基于经验的率失真模型,该算法虽然可以根据图像复杂度、目标码率等较为准确地预设量化参数,但是输出的可变长编码比特流并不能完全地利用网络带宽并适应固定的带宽环境^[2]。而且该算法对运动剧烈及存在场景切换的视频序列的码率控制不够精确,视频图像的

PSNR 波动较大,一些微小的带宽波都会对视频量化的结果和整个视频的解码质量产生影响;另一方面,为了保持连续帧与帧之间视频质量的平滑,只是引入一个固定的值来限定两个相邻帧之间的 QP 跳动,并没有考虑到已编码帧的信息对当前帧 QP 计算的影响^[3]。

本文主要就帧层的量化参数预测进行了研究,力图找到一种新的解决方案,在可以准确预测量化参数的同时又可以增加编码系统的容错性和自适应能力,得到相对稳定的视频质量。对此,提出了一种基于神经网络的量化模型来快速有效地预测每一帧的帧级 QP,该方法能将已编码信息作为输入之一,充分利用这些信息和当前帧的关系,对当前帧的 QP 进行调整。基本思想是通过对多个序列取样,得到样本数据来训练神经网络,使用训练好的神经网络模拟原来的 QP 预测模型,快速得到量化参数预测值,同时神经网络特有的稳健性能加大系统的自适应能力,达到准确预测编码的目标。

1 神经网络概述

人工神经网络简称神经网络,是一种由许多个简单的处理单元(又称神经元)分布式物理连接组成的复杂网络系统,构成的神经网络具有很强的自适应学习能力和容错性。神经网络在理论和应用研究上的进展使得如今神

基金项目:浙江省自然科学基金项目(ZJNSF Y12F010100)

经网络的种类繁多,功能强大。根据连接方式的不同,现今应用较多的有向感知器网络这样的普通前向网络,也有以回归 BP(Back Propagation)网络为代表的有反馈的前向网络,还有像 Hopfield 这样的相互结合型网络^[4]。其应用广泛性在模式识别、图像处理、控制优化、模拟映射等方面都有体现^[5]。

随着研究的深入,神经网络在视频上应用成果不断。文献[1]提取了4个时间和空间域的特征量,通过神经网络的训练模拟映射码率量化模型,得到的模型对场景变换中I帧的预测更加准确。文献[6]通过计算宏块的平均像素值和平均绝对误差和作为输入,训练BP网络对宏块进行 Intra16x16 和 Intra4x4 模式判断,然后使用 Sobel 边缘检测模式选择算法对 Intra16x16 或 Intra4x4 的子模式进行选择后编码。神经网络在视频压缩、变换编码的性能提升和视频传输控制等方向上也取得了很好的效果。

由神经网络理论可知:多层的前向型网络可以以任意精度逼近非线性连续函数。考虑到不同神经网络的特性,在这里把BP神经网络引入到设计中,以充分利用它的自学习和自适应。BP是一种十分常见的网络模型,其实质是将误差反向传播的学习算法,包含两个过程:信息在网络中的前向传播和误差的反向传播。BP网络强大的学习能力使得它具有很好的函数逼近能力,通过样本的训练样本进行模拟出输入和输出之间的非线性关系。

2 视频量化及方案设计

本文的方案主要是在 H.264 的视频编码标准的基础上提出的,场景切换/预处理、目标码率分配、R-D 模型建立和更新、自适应量化控制和缓冲区管理几个模块共同为QP选择模块服务。帧级及宏块级量化参数QP选择的目标是准确控制输出码流,同时考虑到缓冲区状况,以及图像主观质量^[7]。量化的基本环节是从当前帧减去运动补偿,得到一个残差;将些残差进行 DCT(Discrete Cosine Transform,离散余弦变换)变换,将得到的 DCT 系数进行量化进行熵编码。最后这些量化系数附着着周边信息一起被传送到 NAL(Network Abstraction Layer)层进行传输或储存。

2.1 基于 BP 的 QP 预测

根据 JVT 文档中对于帧级码率控制的描述,帧级QP的计算有以下几步^[8]:首先计算当前帧的目标比特,接着分配剩余比特,用线性模型预测 MAD,最后用二次量化模型计算对应的量化参数。由于神经网络需要先提取数据进行训练,直接采用BP网络与JM软件结合起来的机制,流程如图1所示。

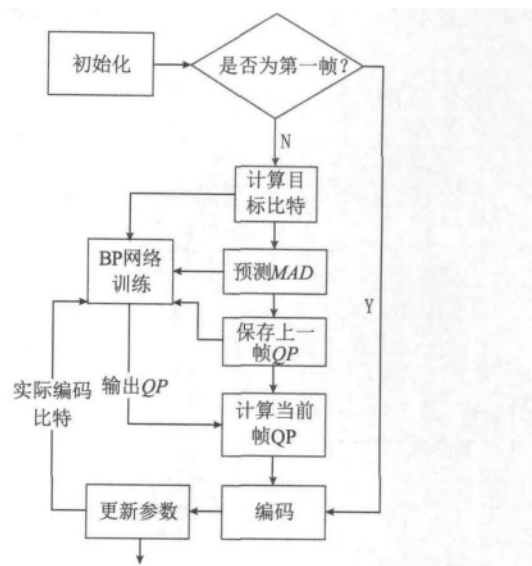


图1 基于BP网络的QP预测流程图

根据量化的主要过程,本文选取了目标比特(TB)、MAD作为特征量。建立BP模型时,结合文献[9]的实验基础,加入前一帧的实际用去的比特数($pre-B$)和量化QP($pre-QP$)来减少输出结果的误差和跳动。网络的结构如图2所示。

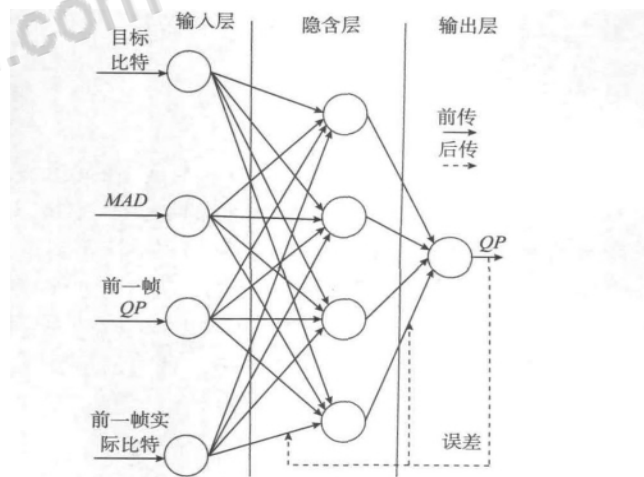


图2 BP网络结构图

本次设计的BP网络有3个输入节点。由经验公式 $m = \sqrt{n+l} + \alpha$ 和 $m = \sqrt{nl} + \alpha$ (其中 m 代表隐含层节点数, n 代表输入层节点数, l 代表输出层节点数, α 为1~10的整数), 隐含层节点数选择在[3,13]区间的整数。整个网络的预测误差由平均绝对误差(Mean-Absolute-Error, MAE)和均方根误差(Root-Mean-Square-Error, RMSE)的大小进行衡量。MAE和RMSE的值越小,代表BP网络的预测越准确。从图3的仿真结果可以看出,在隐含层的节点数为7时,预测的结果最为精确,因此,此网络可表示为BP(4,7,1)。

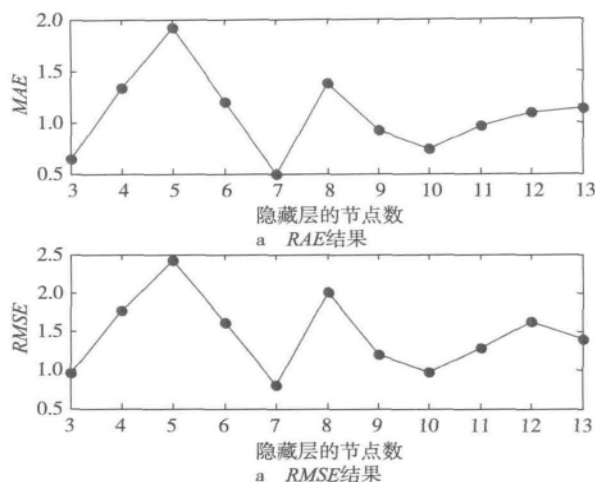


图3 不同节点数的预测误差

2.2 数据处理

第 i 组样本的输入为 $I_i = \{TB_i, MAD_i, pre - B_i, pre - QP_i\}$ 相应整个网络实际输出为 O_i 。节点 j 在第 i 组样本输入时 输出为

$$y_{ij} = f(I_{ij} \cdot w_{ij} + b_i) \quad (1)$$

I_{ij} 是第 i 组输入时 节点 j 的输出, w_{ij} 为上一层节点和下一层节点 j 之间的连接权值, b_i 为节点 i 阈值。 $f(x)$ 为隐含层和输出层的激活函数 即

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

为了使第 i 组样本的期望输出 t_i 与网络的实际输出 O_i 之间的误差函数公式(3) 最小, 训练时 $\varepsilon > 0$ 选取为网络允许最大误差。若 $E < \varepsilon$ 说明网络误差满足要求, 训练结束; 否则在误差梯度下降的方向 重复进行迭代, 调整权值和阈值直到网络满足条件。本文以预测 QP 为目标, 经过实验表明 将网络误差设为 0.001 即可满足要求, 最大训练次数为 500 次。

$$E = \frac{1}{2} \sum (t_i - O_i)^2 \quad (3)$$

为了避免奇异值的出现影响网络的精确度, 在样本输入前要将数据用公式(4) 归一化处理。 $y'(i, j)$ 代表归一化后的数据, $y(i, j)$ 代表原始数据, $y_{\max}(i)$ 与 $y_{\min}(i)$ 分别表示所在行的最大值和最小值。

$$y'(i, j) = \frac{y(i, j) - y_{\min}(i)}{y_{\max}(i) - y_{\min}(i)} \quad (4)$$

最后的数据处理是将 BP 网络独立预测的当前帧 QP 值转换为可供视频编码器使用的量化值。可用公式(5) 对输出进行反归一化 并取整。式中, $x'(i, j)$ 是神经网络的预测值, $x(i, j)$ 表示处理后的 QP 值。 $x_{\max}(i)$ 、 $x_{\min}(i)$ 分别表示样本中目标 QP 的最大值和最小值。 $round$ 进行四

舍五入的取整。

$$x(i, j) = round(x'(i, j) \times (x_{\max}(i) - x_{\min}(i)) + x_{\min}(i)) \quad (5)$$

3 实验仿真结果

结合 H.264 的标准软件, 设计的实验在 JM14.2 中实现并与其比较。实验选取了 Foreman 和 Mobile 两个 cif 序列进行测试。帧频为 30 Hz, 测试带宽为 160 kbit/s, 180 kbit/s, 200 kbit/s, 220 kbit/s, 240 kbit/s 这 5 个点, GoP 结构为 IPPP。

图 4 及图 5 为使用 JM14.2 及本文设计的 BP 网络训练结果, 由于数据篇幅问题只选取 160 kbit/s 和 240 kbit/s 两个点作为代表。从图中可以看到, BP 网络预测的结果与 JM 软件中输出的 QP 曲线基本一致, 而且整个 QP 曲线相对而言起伏比较平稳。

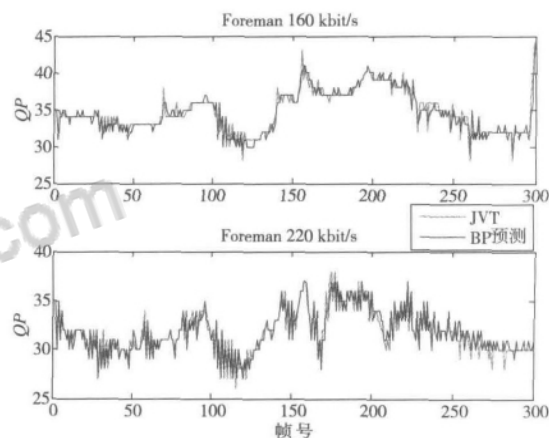


图4 BP 网络对 Foreman 的 QP 预测

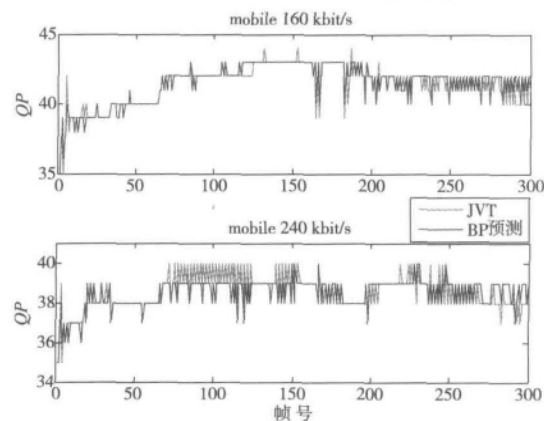


图5 BP 网络对 Mobile 的 QP 预测

将预测得到的 QP 值编码测试序列, 得到了最后编码后序列的峰值信噪比 PSNR(Y) 曲线, 如图 6 及图 7 所示。从图中的 PSNR 比较可以明显看出, 结合 BP 网络来编码的序列, 其 PSNR 曲线更加平滑, 跳变更小, 在部分帧可以观察到在质量平稳的基础上还有 PSNR 的提高。图 8 为两

个测试序列的 R-D 性能曲线,其中本文的方法与 JVT-G012 相比,两条曲线基本一致,由此可说明本文的方法在提高稳定性的同时也保证了视频质量。

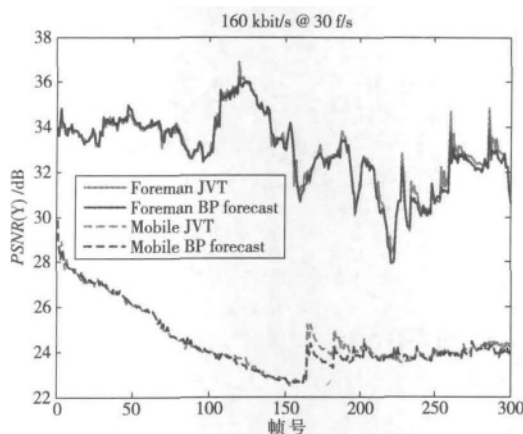


图6 160 kbit/s 下的 PSNR 曲线比较

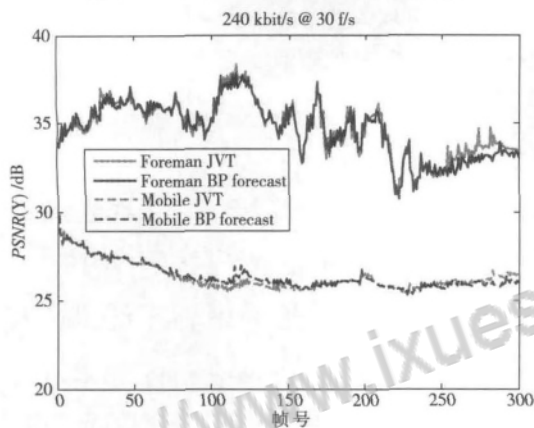


图7 240 kbit/s 下的 PSNR 比较

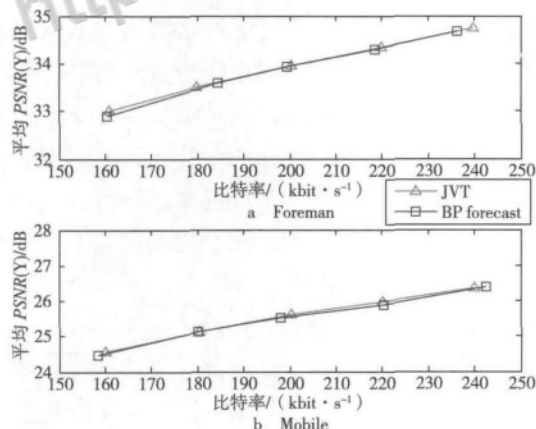


图8 测试序列的 R-D 性能曲线

4 结论

本文提出了一种基于 BP 网络的 H.264 的率失真量化模型。该方案首先使用 BP 神经网络对视频序列进行训练,基于训练的结果预测出测试序列当前帧的帧级量化值,得到每一帧的量化值后对测试序列进行编码。实验结

合表明该方法在保证视频质量几乎不变的情况下,PSNR 的波动,保证了平稳的视频质量。

参考文献:

- [1] CZUNI L, CSASZAR G. Estimating the Optimal Quantization Parameter in H.264 [C]//Proc. 18th International Conference on Pattern Recognition. [S. l.]: IEEE Press, 2006: 330-333.
- [2] LEE H J, CHIANG T, ZHANG Y Q. Scalable rate control for MPEG-4 video [J]. IEEE Trans. Circuits and Systems for Video Technology, 2000, 10(9): 878-894.
- [3] 黄军浩, 毕笃彦, 许悦雷, 等. 基于 H.264 的帧级码率控制和率失真技术研究[J]. 电视技术, 2005, 29(S1): 44-46.
- [4] 飞思科技产品研发中心. 神经网络理论与 Matlab 7 实现[M]. 北京: 电子工业出版社, 2005.
- [5] SUN Xin, LIU Qiongxian, ZHANG Lei. A BP neural network model based on genetic algorithm for comprehensive evaluation [C]//Proc. 2011 Third Pacific-Asia. Conference on Circuits, Communication and system (PACCS) 2011. [S. l.]: IEEE Press, 2011: 1-5.
- [6] 季国飞, 王庆. 一种基于 BP 网络的 H.264 快速帧内模式选择算法[J]. 计算机工程与应用, 2008, 44(13): 100-133.
- [7] 殷海兵. 数字媒体处理技术与应用[M]. 北京: 电子工业出版社, 2011.
- [8] LI Zhengguo, PAN Feng, PANG Keng. Adaptive basic unit layer rate control for JVT: JVT-G012 [R]. [S. l.]: JVT, 2003.
- [9] 李猛. 基于神经网络的视频量化参数预测研究[J]. 辽宁师专学报, 2007, 3(9): 25-28.

责任编辑: 魏雨博

收稿日期: 2012-03-28

国家宽带网络与应用工程技术 研究中心现场验收会召开

9月14日,国家宽带网络与应用工程技术研究中心现场验收评估会议在上海召开。科技部发展计划司蔡文沁副司长和验收评估专家组一行参会并进行现场考察,上海市科委和工程中心依托单位有关领导参加了会议。

工程中心依托上海未来宽带技术及应用工程研究中心有限公司组建。工程中心完成 WDM-PON 原型系统和 IPON 缆桥接入系统研制,建成由宽带中心、上海交大等 5 个节点组成的柔性试验床,建成覆盖 10 万用户的业务试验床,建成覆盖 20 个城市、100 万用户的 NGB 试验床。2012 年 7 月,中心被认为定上海市首批技术创新服务平台——高性能宽带信息网,面向 NGB 行业服务覆盖率超过 90%。

蔡司长在讲话中指出,该中心是一个专注于国家三网融合战略,致力于服务国家下一代广播电视网(NGB 产业)的高新技术企业,要紧紧把握国家战略性新兴产业发展的机遇,力争实现跨越式发展。

论文降重，论文修改，论文代写加微信:18086619247或QQ:516639237

论文免费查重，论文格式一键规范，参考文献规范扫二维码：



[相关推荐：](#)

[一种基于小波和低复杂度矢量量化的视频编码算法](#)

[基于失真-量化模型的视频编码算法优化](#)

[基于通用DSP的视频编码器量化算法研究](#)

[用于HDTV信号压缩的简单递归运动估计技术](#)

[多目标基因算法用于脉冲中子测井数据处理，增强神经网络体的选择](#)

[基于标量量化的多描述视频编码算法](#)

[基于神经网络的视频编码量化参数选择算法](#)

[基于神经网络的视频量化参数预测研究](#)

[基于改进量化策略的XYZ视频编码](#)

[基于自适应门限的视频编码模式快速选择算法](#)