基于单帧图像质量加权的视频质量评价模型

常 青 佟雨兵 张其善 吴今培

(北京航空航天大学 电子信息工程学院，北京 100083) (五邑大学智能技术与系统研究所，江门 529020)

摘 要：利用人眼视觉特性与视频序列时空相关特性，提出了基于视频序列内单帧图像质量加权的视频质量评价模型.其中，单帧图像质量利用峰值信噪比和结构相似性度量作为图像质量的描述参数，采用神经网络(NN，Neural Network)与支持向量机(SVM，Support Vector Machines)建立图像质量评价模型；视频序列质量由序列内单帧图像质量加权衡量，加权因子描述了视频序列内运动及场景变化的剧烈程度.仿真实验结果表明，该模型的输出能有效地反映图像的主观质量.模型预测出的单帧图像质量和视频序列质量的单调性相比 PSNR 分别提高 7.42%和 10.47%，均方根误差相比则提高了 36.06%和 10.48%.

关 键 词：支持向量机；神经网络；视频质量；场景变化

中图分类号：TN 91.21

文献标识码：A **文 章 编 号**：1001-5965(2007)03-0311-04

Video quality assessing model based on single image quality with different weights

Chang Qing Tong Yubing Zhang Qishan

(School of Electronics and Information Engineering, Beijing University of Aeronautics and Astronautics, Beijing 100083, China)

Wu Jinpei

(Institute of Intelligent Technology and System of Wuyi University, Jiangmen 529020, China)

Abstract: Concerning HVS (human visual system) physiological characteristic with the temporal-spatio correlations of video sequence, a video quality assessing model was proposed based on single image quality with different weights. Single image quality was assessed by using NN (neural network) and SVM (support vector machine) with PSNR (peak signal to noise ratio) and SSIM (structure similarity) as two indexes describing image quality. Video quality was assessed by using the quality of each frames in the video sequence with different weights. Those weights described motion and scene changes in the video. The monotonicity of the method for images & video is 7.42% and 10.47% higher than that of PSNR and RMSE (root mean square error) is 35.90% and 10.48% higher than that of PSNR at least. The results from simulation experiments show the model is valid.

Key words: support vector machines; neural networking; video quality; scene changes

一般图像及视频质量客观评价基于人眼视觉系统(HVS，Human Visual System)的生理特征来建模型，如 Watson 的 DCT 矩阵模型[1]、Bradley 的小波域 VDP(Visible Difference Predict)模型等[2].以上模型计算复杂，质量评价流程不够清晰，且以上质量评价模型都是基于一定的假设前提进行的，如一般都认为原始图像质量是完美的且 Weber 定律和 PSF 模型始终成立.而实际图像种类多且差别巨大，当具体实例与假设相差较大时，质量评价的结果会产生较大的偏差.

本文从另一个方面考虑视频质量评价问题，从峰值信噪比(PSNR，Peak Signal to Noise Rati-

收稿日期：2006-04-19

基金项目：国家自然科学基金资助项目(60372018)；航空科学基金资助项目(04F51068)

作者简介：常 青(1962—)，男，天津人，副教授，changq@263.net.

o)这个最为通用的信号质量评价指标入手,引入结构相似度量(SSIM,Structure Similarity),给出了一种利用PSNR和SSIM二者联合构建新图像质量评价模型的方法,该质量评价模型结构简单,易于运算;视频序列的质量由单帧图像质量加权衡量,除了考虑单帧图像质量外,还考虑了视频序列的运动信息及场景变化对人眼感知的影响,这更接近于主观质量评价过程.文末仿真实验的结果证明该模型的输出能有效地反映视频的质量.

1 图像质量描述参数

单帧图像质量评价时采用PSNR和SSIM作为两个图像质量描述参数[3].虽然PSNR不足以全面反映人眼视觉对图像质量的感知,但PSNR是通用的信号质量评价指标,对于多数图像质量评价仍是有效的.因此,本文仍将PSNR作为图像质量评价一个有意义的参考指标.但对图像质量评价模型进行性能测试时,少部分主观质量非常差的图像与客观模型预测值之间的差异大而导致评价模型性能的下降.因此,需要对PSNR产生"野点"(主客观质量评价相差较大的点)的情况进行校正以提高质量评价的准确性.

基于结构失真进行视频质量客观评价的依据是HVS高度适合于提取视觉场景中的结构信息,从而测量结构信息的改变与感知图像质量的变化非常接近[4].图像信号 x 和 y 之间的结构相似性度量值为

$$S(x,y)=[l(x,y)]^{\alpha}\cdot[c(x,y)]^{\beta}\cdot[s(x,y)]^{\gamma} \tag{1}$$

其中,$l(x,y)$,$c(x,y)$和$s(x,y)$分别为亮度、对比度和结构比较函数,三个函数相对独立,$\alpha,\beta,\gamma>0$.实际情况下,$S(x,y)$与$l(x,y)$,$c(x,y)$,$s(x,y)$之间存在非线性的映射关系,式(1)将其固定为三者幂乘积的形式容易产生误差;单独利用SSIM方法进行性能测试时,在主客观质量评价关联图中仍然存在着一些"野点".这些"野点"降低了质量评价的准确性.本文将SSIM值作为描述图像质量的一个参数,结合PSNR对"野点"问题进行考虑.

2 单帧图像质量评价模型

利用PSNR和SSIM作为图像质量描述参数,联合应用神经网络与支持向量机进行图像质量客观评价,具体的操作流程如图1所示.

整个流程分为两个阶段:

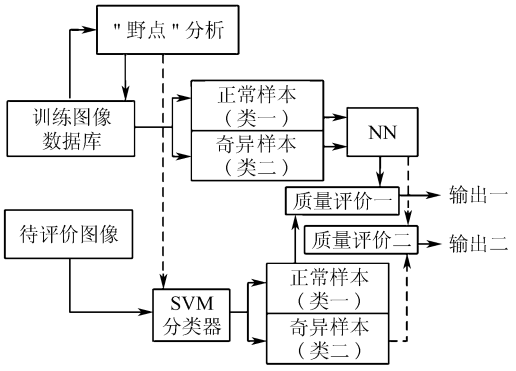


图1 单帧图像质量评价操作流程

①"野点"分析及质量评价映射函数;②"野点"预测"与SVM(Support Vector Machine)分类.

2.1 "野点"分析及质量评价映射函数

定义1 图像质量评价模型的主客观评价关联图中,若某样本其主观评价值与客观评价值相差较大,将其定义为"野点".主客观评价值的偏差可以采用下面两种方式来衡量:

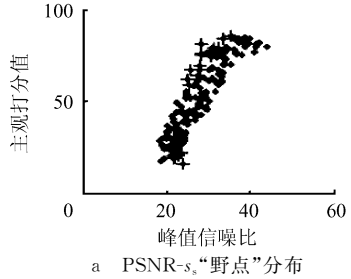
1) 空间距离表示

若$|s_s-e_o|\geqslant\theta_{th}$,将该样本在主客观评价的关联图中定义为"野点";

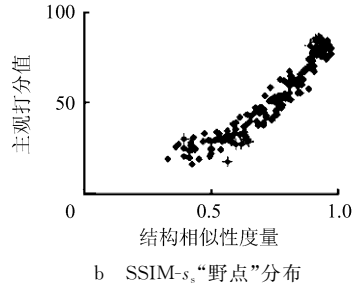
2) 斜率表示

若$(s_s/e_o)\geqslant\theta_{th}$或$(s_s/e_o)\leqslant 1/\theta_{th}$,将该样本在主客观评价的关联图中定义为"野点".

其中,s_s为主观打分值;e_o为客观评价值即为PSNR或SSIM的数值;θ_{th}为门限值,可以通过实验确定.采用斜率表示偏差时,PSNR和SSIM相应的主客观评价关联图中可能的"野点"分布见图2("野点"采用"+"予以标注).



a PSNR-s_s"野点"分布



b SSIM-s_s"野点"分布

图2 "野点"分布图

通过"野点"分析,将训练样本集 S 分为正常

样本集 S_1 和奇异样本集 S_2,利用这两个样本集分别去训练神经网络,得到不同的质量评价映射函数.为简化运算,此处采用 BP(Back Propagation)神经网络,结构如图 3 所示.

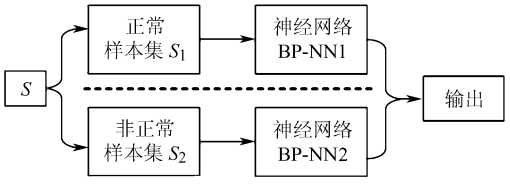


图 3　图像质量评价映射函数框图

将不同样本集的数据作为训练数据,神经网络不断调整网络节点间的权值确定出质量评价映射函数,试验中设定最大训练步数为 8 000,误差 0.05.

2.2　"野点预测"与样本分类

"野点"定义时用到了主观评价值 s_s,但实际图像质量评价时可能没有主观评价值,不能按照上文"野点"分析中的定义来确定"野点",此时利用"野点预测"来确定"野点".

定义 2　处理待评价质量的图像时,在没有主观值的情况下,基于训练样本集"野点"分析的结果,利用 SVM 分类器实现样本归类,即预测出该样本是否为正常样本.这个过程定义为"野点预测".

通过"野点预测"可以对单独采用 PSNR 或 SSIM 可能产生的结果进行估计:估计此时的质量评价值是否会与理想的主观评价值相差大,进而对使用 PSNR 或 SSIM 产生的结果进行调整,减少偏差.

SVM 是一种基于监督学习的分类方法,其强大的分类能力保证了类间距离最大[5].通过获取广义最优分类面 $f(\boldsymbol{x};\boldsymbol{w})=(\boldsymbol{w}\cdot\boldsymbol{x})+b$,使得两类样本的间隔最大,从而完成分类功能.引入松弛变量及对偶原理,得到回归函数

$$f(\boldsymbol{x};\boldsymbol{a})=\sum_{i=1}^{N}(a_i^*-a_i)(\boldsymbol{x}_i\cdot\boldsymbol{x})+b=\sum_{i=1}^{N}(a_i^*-a_i)K(\boldsymbol{x},\boldsymbol{x}_i)+b \tag{2}$$

其中,$K(\cdot,\cdot)=(\boldsymbol{x}_i\cdot\boldsymbol{x})$,称为核函数,满足 mercey 定理;系数 $\boldsymbol{a},\boldsymbol{a}^*$ 为某一个二次规划问题的最优解.式(2)中仅有部分系数非零;非零系数对应的输入量 $\boldsymbol{x}_i(i=0,\cdots,l,l<N)$ 称为支持向量,支持向量位于分类边界或分类面上.

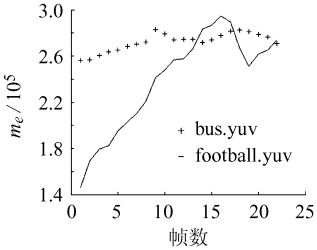
利用"野点"分析的结果,将样本按照 PSNR 和 SSIM 指标归类重排后,作为 SVM 分类器的训练数据,通过训练获得最终的 SVM 分类器模型,该分类器的输入为 PSNR 和 SSIM,输出值 1 表示为正常类,输出值 0 表示利用 PSNR 或 SSIM 进行质量评价可能会产生"野点".

3　视频质量评价

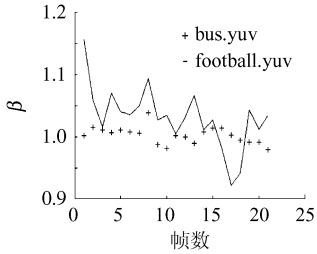
根据实际实验发现,某些明显影响单帧图像质量感知的因素,如块效应,对整个视频序列质量的影响并不突出.人眼对于场景中的运动物体更为关注而可能忽略大部分背景信息.因此,视频质量评价除了利用单帧图像质量外,还必须考虑运动信息对视觉感知的影响.设定视频显示的速度是固定的,图像运动信息主要体现为场景变化的程度.MAD (Mean Absolute Difference) 描述了连续运动的视频内图像在时间上的距离,记作 m_e.定义参数 β 对视频运动信息进行描述

$$\beta_i=m_{e_{i+1}}/m_{e_i} \tag{3}$$

对不同的视频序列,β 值波动越强,说明视频运动信息越丰富,场景变化明显;反之亦然.以 football 和 bus 序列为例,前者有多个物体运动,运动信息丰富,而后者相对简单.其 β 值波动见图 4.



a　MAD 值的变化曲线



b　β 值波动图

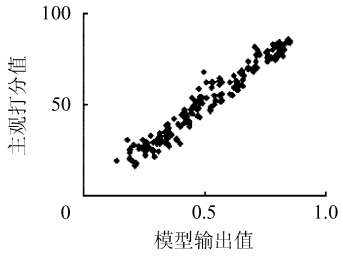
图 4　MAD 和 β 的变化曲线

记单帧图像质量评价为 s_{s_i},视频质量

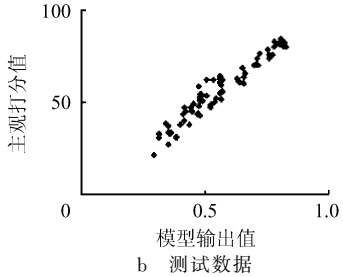
$$s_V=\frac{s_{s_1}+s_{s_2}}{2+\sum_{i=1}^{n-2}\beta_i}+\left(\frac{\beta_i}{2+\sum_{i=1}^{n-2}\beta_i}\right)\times s_{s_{i+2}} \tag{4}$$

4　仿真实验

利用美国 UTexas 大学图像数据库及部分视频序列进行仿真实验[6],结果见图 5.



a 训练数据



b 测试数据

图5 模型输出与主观评价的关联图

由图5可以看出,相比PSNR和SSIM方法,本文图像质量评价模型的主客观质量评价关联图中的"野点"即偏差大的点明显减少.利用100幅图像进行测试,分别给出了"野点"定义中的两种偏差所计算的结果,见表1.相关系数表征了模型预测结果的单调性,离群比表征了模型预测结果的连续有效性[7].综合比较表1中数据可以看出,相比PSNR,模型预测的单调性提高7.42%,模型预测的均方根误差提高36.06%.

表1 单帧图像质量评价模型统计数据

方法	指标		
	相关系数	离群比	均方根误差
PSNR	0.8979	0	0.3053
SSIM	0.9601	0.3947	0.4860
本文模型	0.9645	0	0.1952

视频质量评价时采用70个qcif序列(176×144)和25个cif序列(352×288)对视频质量评价模型进行测试,其相关指标如表2所示.

表2 视频质量评价模型统计数据

方法	指标		
	相关系数	离群比	均方根误差
PSNR	0.6559	0.6000	0.4912
SSIM	0.7075	0.2421	0.3238
本文模型	0.7246	0.0316	0.3864

相比PSNR,本文视频质量评价模型输出的单调性提高10.47%,模型预测结果的均方根误差提高10.48%.试验结果表明质量评价模型的输出能有效地反映视频序列的主观质量.

5 结束语

本文将PSNR和SSIM作为描述图像质量评价的两个参数,联合应用神经网络与支持向量机建立新的图像质量评价模型,图像质量评价模型性能测试中"野点"的数目相对减少,模型性能得以提高.在此基础上给出了基于单帧质量加权的视频序列质量评价模型.基于实际视频序列的实验结果证明该模型输出与视频主观评价值关联性强,能够有效地评价视频质量.本文主要研究了有原始信号作参考时的质量评价方法,而对于没有参考信号的质量评价将是论文进一步研究的方向.

参考文献(References)

[1] Watson A B, Solomon J A. Model of visual contrast gain control and pattern masking [J]. Journal of Optical Society of America, 1997,14(9): 2379-2391

[2] Vanden C J, Branden Lambrecht, Costantini D M, et al. Quality assessment of motion rendition in video coding [J]. IEEE Trans Circuits and Systems for Video Tech, 1999,9(5): 766-782

[3] 佟雨兵,胡薇薇,杨东凯,等.视频图像质量评价综述[J].计算机辅助设计与图形学学报,2006,18(5):1-7
Tong Yubing, Hu Weiwei, Yang Dongkai, et al. Video quality assessment methods overview[J]. Journal of CAD & Computer Graphics, 2006,18(5): 1-7 (in Chinese)

[4] Zhou Wang, Alan C Bovik, Eero P Simoncelli. Structural approached to image quality assessment[M]. 2nd edition, In Handbook of Image and Video Processing. New York: Academic Press, 2005

[5] Vladimir N Vapnik. An overview of statistical learning theory [J]. IEEE Transactions on Neural Networks, 1999,10(5): 988-999

[6] JPEG DATABASE[DB/OL]. 2002-04-15[2005-04-10]. http://live.ece.utexas.edu/index.htm

[7] RRNR-TV Group TEST PLAN[EB/OL]. Draft version 1.7, 2004-06-21[2005-01-10]. http://www.vqeg.org

(责任编辑:娄 嘉)