

文章编号:1004-4213(2011)10-1547-6

基于梯度和运动估计的视频质量评价

刘文娟,李素梅,臧艳军,卫津津

(天津大学 电子信息工程学院,天津 300072)

摘 要:数字视频质量评价在视频压缩、处理以及视频通信领域中有着十分重要的作用.本文在现有视频质量评价方法以及对人类视觉特性研究的基础上,提出了一种基于梯度和运动估计的视频质量评估方法.该方法充分考虑了视频图像的边缘特性、各帧之间的时间相关性以及画面中场景的变化对视频质量的影响等因素.实验结果表明,该方法对基于 H.264 编码的失真视频有更好的评价效果,与人类视觉的主观评价结果一致.

关键词:视频质量评价;基于梯度的结构相似度;运动估计;梯度和运动估计

中图分类号: TN919.82

文献标识码: A

doi: 10.3788/gzxb20114010.1547

0 引言

近年来,视频相关技术的发展极其迅速,视频质量评价也备受关注.数字视频的质量评价方法有主观质量评价和客观质量评价两种方法.主观质量评价是由不同类型的观察者通过大量的主观实验来测评视频质量的好坏,没有任何一种客观评价方法可以完全替代主观评价,但因主观评价方法繁琐、耗资、耗时等缺点,将其应用到实际的视频处理系统中有一定的困难^[1].

客观评价方法相对于主观评价方法具有简便可行的优势.目前,视频质量的客观评价方法已有很多种^[1-10].较常用的是峰值信噪比(Peak Signal Noise Ratio, PSNR)、均方根误差(Root Mean Squared Error, RMSE)^[1-2],这类方法原理简单、易于计算,但是,它们没有考虑到像素间的相关性并忽略了人眼的视觉特性,致使主客观评价结果不一致^[11].Zhou Wang 提出了基于结构相似度(Structure Similarity, SSIM)的图像质量评价方法^[12],这种方法没有考虑到图像的边缘信息,对高度模糊的图像评价效果很差^[13-14];基于 SSIM 方法,文献[13]提出了基于结构边缘信息(Edge Based Structural Similarity, ESSIM)的图像质量评价方法,对模糊失真的图像有一定的改善^[4],但是,上述几种方法没有考虑到视频中帧与帧之间的时间连续性,只能用于图像的质量评价.文献[4]提出一种基于图像结构失

真的视频质量评价方法,文献[10]提出了基于运动信息和结构信息(Motion Information and Structural Similarity, MESSIM)的视频质量评价方法.但是,这两种方法在单帧图像质量评价时都没有考虑到局部失真时的边缘结构信息,且实验证明 MESSIM 方法对高度压缩的视频的质量评价与主观评价结果相差甚远.基于上述原因,本文提出了一种基于梯度和运动估计(Gradient and Motion Estimation, GMES)的视频质量评价方法,该方法不仅考虑了视频图像的结构信息和运动信息等因素对人眼视觉的影响,还考虑了人眼对图像边缘纹理信息的敏感性,并基于 MESSIM 方法做了改进,能够较好地模拟人眼的视觉机制.实验结果表明,本文所提出的方法对因 H.264 编码产生失真的视频的客观评价结果与人眼的主观评价结果一致性较高.

1 视频质量评价的相关算法

1.1 图像结构相似度

人眼视觉系统的主要功能是提取视觉区域内的结构信息.Zhou Wang^[3]等人提出了结构相似度的方法,将失真图像和参考图像分成不重叠的子块,计算失真图像子块和参考图像子块对应位置的结构相似度,建模为三个不同因素的组合:亮度($l(x, y)$)、对比度($c(x, y)$)、结构($s(x, y)$),计算公式为

$$\text{SSIM}(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma \quad (1)$$

式中, α 、 β 、 γ 为三个函数所占的比重.整幅图像的

基金项目:国家自然科学基金(No.61002028)和天津市重点基金(No.09JCZDJC17000)资助

第一作者:刘文娟(1985-),女,硕士研究生,主要研究方向为图像处理和信息处理. Email:liuwenjuan@tju.edu.cn

导师(通讯作者):李素梅(1975-),女,副教授,主要研究方向为立体图像处理、神经网络和机器学习等. Email:tjnklsm@163.com

收稿日期:2011-01-19;修回日期:2011-08-05

评价表达式为

$$\text{MSSIM}(X, Y) = \frac{1}{M} \sum_{j=1}^M \text{SSIM}(x_j, y_j) \quad (2)$$

式中, X, Y 分别代表参考图像和失真图像, x_j, y_j 代表第 j 个图像子块的内容, M 是整幅图像的分块数目. 与 MSE 和 PSNR 相比, SSIM 体现了人眼的一定特征, 具有简单、高效的特点, 然而在评价严重模糊图像时, 结果不理想^[10, 13-14], 并且不适合用于视频质量的评价.

1.2 基于梯度的结构相似度

研究表明, 人眼对图像边缘和纹理信息很敏感, 而 SSIM 的结构比较函数部分只简单地将对应位置的像素进行比较, 并没有考虑图像的边缘和纹理信息. 基于 SSIM 方法, 文献[14]提出了基于梯度的结构相似度 (Gradient Based Structural Similarity, GSSIM). 通过 sobel 算子定义图像中每一像素的梯度信息向量 $V_{ij} = \{dx_{ij}, dy_{ij}\}$, 近似定义图像第 (i, j) 个像素的梯度模值^[15]为

$$\text{AMP}_{ij} = |dx_{ij}| + |dy_{ij}| \quad (3)$$

由式(3)得到参考图像 X 和失真图像 Y 相对应的梯度模值矩阵 X', Y' , 用子块梯度对比度函数 $C_g(x, y)$ 和子块梯度相关系数 $S_g(x, y)$ 代替 SSIM 中的 $c(x, y)$ 和 $s(x, y)$, 则整幅图像的评价表达式为

$$\text{GSSIM}(x, y) = [l(x, y)]^\alpha [C_g(x, y)]^\beta \cdot [S_g(x, y)]^\gamma \quad (4)$$

该文中, $\alpha = \beta = \gamma = 1$, 对于整幅图像的结构相似度的比较为

$$\text{MGSSIM}(X, Y) = \frac{1}{M} \sum_{j=1}^M \text{GSSIM}(x_j, y_j) \quad (5)$$

式中, x_j, y_j 分别表示图像 X, Y 的第 j 个子块, M 是图像中子块的数量.

但是, GSSIM 只适用于图像的质量评价, 将其应用到视频质量评价领域, 需要与视频的运动信息结合起来. 本文正是将 GSSIM 与运动信息结合, 根据人眼对不同区域的敏感度不同而给不同参量赋予不同的权值, 得到新的视频质量评价模型.

1.3 基于运动信息和结构信息的视频质量评价方法

视频质量的评价必须考虑到视频的时间信息或者运动信息, 所以通常来说, 视频失真的评价可以分为三个层面进行: 视频序列单帧图像的局部失真; 视频中某些帧整体失真; 整个视频段失真^[4].

1.3.1 视频序列单帧图像的局部失真

文献[10]对于单帧图像的局部失真采用的是 SSIM, 该帧整幅图像质量的评价为

$$Q_i = \frac{\sum_{j=1}^M \omega_{ij} * \text{SSIM}}{\sum_{j=1}^M \omega_{ij}} \quad (6)$$

式中, Q_i 是视频序列中第 i 帧的图像质量, M 为单帧图像分块的数量, ω_{ij} 为第 i 帧、第 j 块的权重, 且 $\omega_{ij} = E_{ij} \times M_{ij} \times L_{ij}$, E_{ij}, M_{ij}, L_{ij} 分别表示对每一图像子块的结构信息, 运动信息, 亮度信息的调节权重.

1) 结构信息的权值 E_{ij}

用式(3)来计算单帧图像各像素的梯度模值, 由于人眼对于边缘纹理信息较敏感^[16], 边缘部分梯度的模值较大, 因此, 对梯度模值较大的子块赋予较大的权值, 第 i 帧、第 j 块的结构信息权重为^[10]

$$E_{ij} = \begin{cases} 4 & \text{AMP}_{ij} \geq 255 \\ 2 & 85 \leq \text{AMP}_{ij} < 255 \\ 1 & \text{AMP}_{ij} < 85 \end{cases} \quad (7)$$

2) 亮度信息的权值 L_{ij}

研究发现, 人眼在较暗和较亮的背景下, 视觉存在对比度掩盖效应^[4], 根据这一特性, 采用单帧图像子块亮度分量的均值 μ_x 作为局部亮度的估计, 第 i 帧、第 j 块的亮度信息的权重为^[4]

$$L_{ij} = \begin{cases} 0 & \mu_x \leq 40 \\ (\mu_x - 40)/8 & 40 < \mu_x \leq 50 \\ 1 & \mu_x > 50 \end{cases} \quad (8)$$

3) 运动信息的权值 M_{ij}

运动速率不同时, 人眼的视觉特性会存在对视频的模糊失真产生掩盖效应^[4]. 有相同数量的帧图像出现模糊失真时, 人眼的感知随运动速率的不同而不同. 所以需要对不同的运动信息加以不同的权值, 以增强算法的鲁棒性.

首先, 采用基于块的运动估计算法对每个局部抽样区域计算其相对下一帧的运动^[17-19], 得到第 i 帧、第 j 块区域的运动矢量长度 m_{ij} , 并计算运动向量平均值 Ma_i , 然后, 把 Ma_i 分为两种情况, 计算第 i 帧、第 j 块的运动信息加权值 M_{ij} :

① $Ma_i < 16$ 时, 整个视频的背景变化不大, 只有部分对象在运动或变化. 这时, 人眼最关心的是视频中运动和变化的对象, 这时给这部分信息一个大的权值为^[10]

$$M_{ij} = \begin{cases} 1 & m_{ij} \leq 12 \\ 2 - (20 - m_{ij})/8 & 12 < m_{ij} \leq 20 \\ 2 & m_{ij} > 20 \end{cases} \quad (9)$$

② $Ma_i \geq 16$ 时, 视频背景图像运动剧烈, 人眼很难分辨出视频中内容的细节. 我们按照运动剧烈的宏块所占整帧运动宏块的比重划分为两种情况, 如果比重大于 90%, 该帧运动矢量的总长度很大, 人眼更注重相对变化较小的对象, 并对运动变化较小的部分赋予较高权值^[10], 如式(10); 如果比重小于 90%, 则说明大部分物体变化不是非常剧烈, 人眼就

会对高速运动的部分较敏感,并对此部分赋以较高的权值^[10],如式(11).

$$M_{ij} = \begin{cases} 2 & m_{ij} \leq 45 \\ 1 + (55 - m_{ij})/10 & 45 < m_{ij} \leq 55 \\ 1 & m_{ij} > 55 \end{cases} \quad (10)$$

$$M_{ij} = \begin{cases} 1 & m_{ij} \leq 45 \\ 2 - (55 - m_{ij})/8 & 45 < m_{ij} \leq 55 \\ 2 & m_{ij} > 55 \end{cases} \quad (11)$$

1.3.2 视频中某些帧失真和整段视频的失真

对于视频某些帧失真和整段视频的失真采取式(12)处理^[4],得到整段视频的客观质量得分为

$$Q = \frac{\sum_{i=1}^F W_i \times Q_i}{\sum_{i=1}^F W_i} \quad (12)$$

式中, F 表示序列中帧的数量, Q_i 是由式(6)计算得来的第*i*帧视频图像的得分, $Q_i=1$ 表示第*i*帧视频图像没有失真, $Q_i=0$ 表示第*i*帧视频图像已经完全失真. W_i 表示第*i*帧的权值,由于人在观看视频时很容易注意到受损的帧图像,所以对图像质量较差的帧赋予较高的权值,帧图像质量好的赋予较小的权值, W_i 的选取采取原则为

$$W_i = \begin{cases} 1 & 0.8 < Q_i \leq 1.0 \\ 2 & 0.6 < Q_i \leq 0.8 \\ 3 & 0.4 < Q_i \leq 0.6 \\ 4 & 0.2 < Q_i \leq 0.4 \\ 5 & 0.0 < Q_i \leq 0.2 \end{cases} \quad (13)$$

2 基于梯度和运动估计的视频质量评价方法

本文通过 Matlab 仿真,发现文献[10]给出的 MESSIM 对由模糊和压缩引起的视频失真与人眼视觉特性有较好的一致性,但是该方法只考虑了量化步长 $QP \leq 30$ 的视频压缩情况,在视频高度压缩的情况下,人眼看到的压缩视频质量很差,但采用 MESSIM 方法却能得到较高的评分,如表 1. 以量化步长 $QP=38$ 的视频序列 Hehua_QP38.yuv 为例,视频主观平均得分为 2.750 0 分,(主观评价得分标准见表 1),人眼感知效果已经不是很好,而

表 1 视频主观评价得分标准

Score	Video quality classification	Video injury scale
5	Excellent	Imperceptible
4	Good	Perceptible, but not annoying
3	Fair	Slightly annoying
2	Poor	Annoying
1	Bad	Very annoying

MESSIM 得分却达到了较高分数的归一化值 0.869 71. 针对这种情况,考虑到图像边缘纹理信息对人眼视觉的重要性,本文在 MESSIM 的基础上进行了改进.

首先将一帧图像按 16×16 的大小分为 M 块,对单帧图像质量的局部评价采用 GSSIM 方法,于是第 i 帧图像质量的评价得分为

$$Q_i = \frac{\sum_{j=1}^M \omega_{ij} * GSSIM}{\sum_{j=1}^M \omega_{ij}} \quad (14)$$

式中, Q_i 表示视频序列中第*i*帧的图像的质量, ω_{ij} 为第*i*帧、第*j*块的权重,且 $\omega_{ij} = E_{ij} \times M_{ij} \times L_{ij}$, E_{ij} 、 M_{ij} 、 L_{ij} 分别按照上文中 1.3.1 中计算.最后,将式(14)代入(12),得到经 H.264 压缩的视频 GMES 客观评分为

$$Q = \frac{\sum_{i=1}^F W_i \times Q_i}{\sum_{i=1}^F W_i} \quad (15)$$

本文所提算法的实验结构图如图 1.

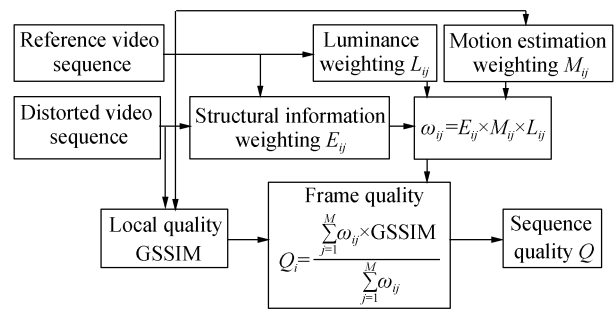


图 1 本文算法结构 Fig. 1 Algorithm structure chart

3 实验数据及实验结果

本文方法是在 Core i5 750@2.67 GHz,内存 4 GB 的平台上,采用 matlab R2010a 版本,对参考视频和失真视频的前 20 帧进行评价.

3.1 实验数据

本文对三个视频序列:hehua.yuv(分辨率为 640×384)、kroger.yuv(分辨率为 640×384)、woshou.yuv(分辨率为 512×384)进行实验,实验数据由天津大学宽带无线通信与立体成像研究所提供.对上述三个视频序列进行 H.264 压缩损失处理(量化参量 QP 分别取为 6、12、18、20、22、24、26、28、30、32、34、36、38、40、42、44、46、48),由每一视频序列得到 18 个压缩视频序列,共得到 54 个压缩失真的视频序列,实验所用数据如图 2.

主观实验分为三组进行,每组让 20 个观察者看 18 对原始视频和失真视频组合,观察者根据被测视频的损伤程度按照双重激励源损伤尺度(Double Stimulus Injure Scale, DSIS)5 级标准^[20]打分(见表 1).

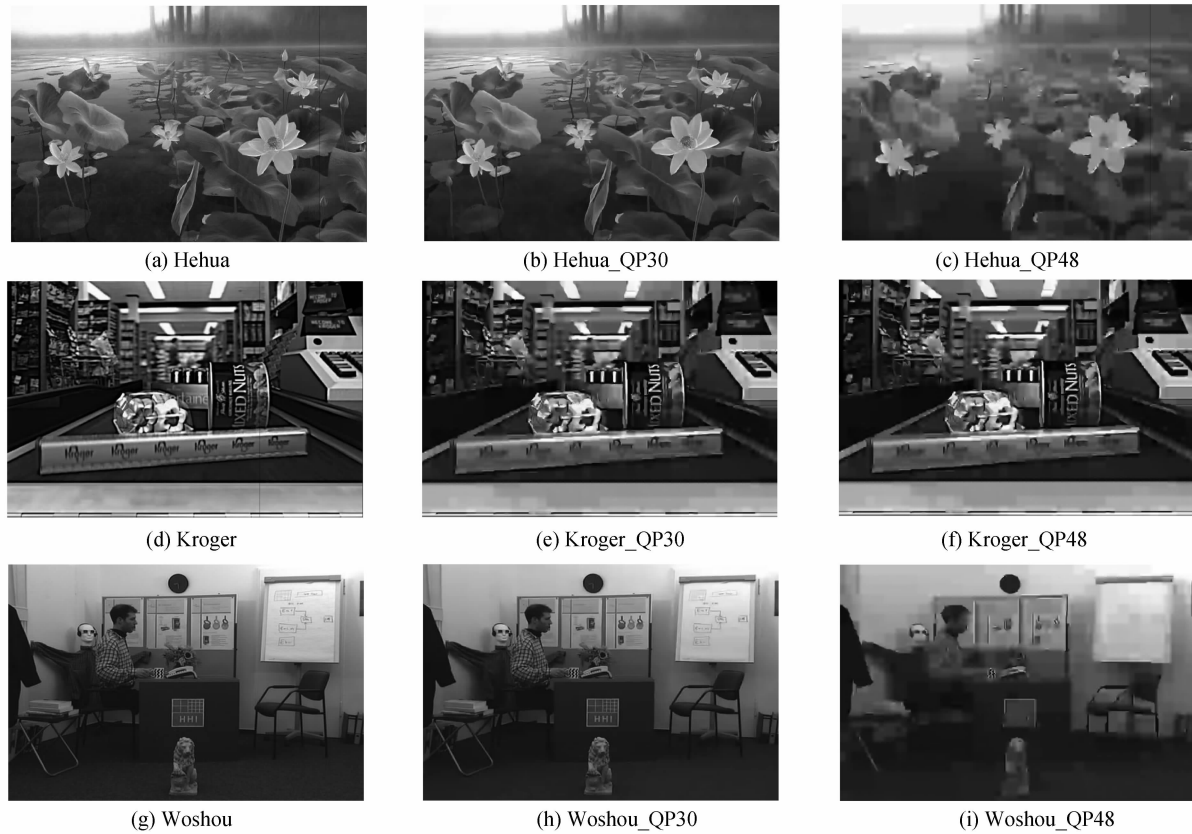


图2 实验原始数据和失真数据的第一帧

Fig.2 The first frame of the original and the distorted data

3.2 实验结果

实验中,首先将每一组视频序列由主观评价得到的分数(sub_s: subjective scores)归一化到 $[0, 1]$ 范围内,得归一化主观评价结果(n_s: normal scores),其中1表示质量最好,0代表质量最差.由MESSIM算法及本文算法得到各视频序列的客观评价结果,并将主客观评价结果进行比较.以视频序列 hehua.yuv 为例,得到 hehua.yuv 视频组各序列的主、客观评价结果,如表2.

由表2看出,当量化步长 $QP < 28$ 的时候,主观评价分数均为5分,得到的图像质量都很好,人眼几乎分辨不出来.当量化步长 $QP > 32$ 时,视频图像会很粗糙,视频失真比较严重.将算法MESSIM和本文所提算法GMES分别对量化步长取 $[6, 48]$ 范围内时所得结果进行比较,MESSIM所得最高分和最低分为 $[0.996\ 85, 0.736\ 85]$,级差仅为0.26,不能充分地体现出视频质量的等级差异,而GMES所得最高分和最低分为 $[0.985\ 76, 0.425\ 97]$,级差达到了约0.56,能够很好体现出视频质量的等级差异.比如,在表2中,当 $QP = 36$ 时,主观平均得分为3分,视频质量为fair,GMES得分为0.655 92符合主观评价,而MESSIM得分却达到了0.891 92的分值,与主观评价的结果有一定的差异;当 $QP = 44$

表2 视频 hehua.yuv 序列实验结果

Table 2 Results of the algorithms for the hehua.yuv sequence

Video sequence	MESSIM	GMES	sub-s	n-s
Hehua_QP6	0.996 85	0.985 76	5	1
Hehua_QP12	0.992 25	0.966 70	5	1
Hehua_QP18	0.981 75	0.919 82	5	1
Hehua_QP20	0.976 61	0.897 69	5	1
Hehua_QP22	0.970 84	0.872 45	5	1
Hehua_QP24	0.963 63	0.843 54	5	1
Hehua_QP26	0.955 70	0.813 73	4.875 00	0.968 75
Hehua_QP28	0.946 63	0.785 59	4.625 00	0.906 25
Hehua_QP30	0.936 36	0.757 19	4.000 0	0.750 00
Hehua_QP32	0.923 63	0.724 22	3.750 0	0.687 50
Hehua_QP34	0.908 89	0.693 06	3.125 0	0.531 25
Hehua_QP36	0.891 92	0.655 92	3.000 0	0.500 00
Hehua_QP38	0.869 71	0.614 22	2.750 0	0.437 50
Hehua_QP40	0.847 23	0.573 87	1.875 0	0.218 75
Hehua_QP42	0.817 10	0.522 80	1.500 0	0.125 00
Hehua_QP44	0.792 38	0.488 83	1.000 0	0
Hehua_QP46	0.767 07	0.458 20	1.000 0	0
Hehua_QP48	0.736 85	0.425 97	1.000 0	0

时,主观平均得分为1分,本文所提算法GMES得分为0.488 83,符合主观评价的质量等级,但是,算法MESSIM得分却达到了0.792 38的分值,与主观评价的等级差别较大,不能很好地反映图像的质

量等级.相比较而言,对于20帧视频图像的处理,MESSIM平均运算时间为68.1s,本文算法平均时间为77.2s,在运算时间相差不多的情况下,本文算法得到较好的评价结果.

为了进一步比较GMES和MESSIM两种算法,对3组视频序列计算由GMES和MESSIM算法所得客观分数O(objective)与归一化主观评价得分S(subjective)的Pearson相关系数 r 式(16)以及均方误差RMSE式(17),得到算法性能比较的结果,如表3.

表3 MESSIM与GMES算法性能比较
Table 3 Comparison of the MESSIM and GMES algorithms

QP	Video sequence	r		RMSE	
		MESSIM	GMES	MESSIM	GMES
6-28	hehua	0.986 85	0.969 62	1.074 63	1.208 37
	Woshou	0.984 97	0.965 21	1.064 25	1.360 41
	Kroger	0.983 54	0.964 38	1.084 67	1.389 72
28-48	Hehua	0.968 98	0.989 15	1.712 58	0.976 10
	Woshou	0.915 55	0.951 06	1.916 22	1.302 89
	Kroger	0.938 04	0.971 64	1.737 02	1.348 27

$$r = \frac{N \sum O_i S_i - \sum O_i \sum S_i}{\sqrt{N \sum O_i^2 - (\sum O_i)^2} \sqrt{N \sum S_i^2 - (\sum S_i)^2}} \quad (16)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (O - S)^2} \quad (17)$$

表3中,Pearson相关系数 r 的值越大,说明算法的客观质量评分O和主观评分S相关性越好, r 的值越小,说明该评价方法的客观质量评分O和主观评分S相关性越差,当 $r=0$ 时,所得客观评价结果与主观评价结果相差很远.而RMSE的值越小,则说明对应算法的预测误差越小.通常,当量化步长 $QP \leq 28$ 时,图像质量会比较好,主观得分基本都在一个等级,由表3可以看出,两种方法的Pearson相关系数 r 均比较大,在0.9643以上,说明此时,两种方法均能较好的与主观评价相一致.但是,当 $28 \leq QP \leq 48$ 时,MESSIM算法的Pearson相关系数 r 均小于0.96898,与主观评价的一致性较差,而本文所提算法GMES的Pearson相关系数 r 最小值为0.95106,最大值亦可达到0.98915,而且GMES的RMSE值比MESSIM都小,说明GMES预测误差也较小.所以对H.264压缩失真的视频质量评价,GMES算法比MESSIM算法更符合人眼的视觉机制.

4 结论

由于人眼对图像的梯度纹理信息较敏感,本文在MESSIM算法的基础上,在视频序列的帧图像中引入了梯度信息,保留了图像像素点间位置的相关

信息,通过计算运动矢量,估算视频场景中物体运动快慢对人眼视觉效果的影响,使得算法更接近于人眼的视觉效果.实验结果表明,本文所提GMES算法比MESSIM算法更符合人眼的视觉机制.

参考文献

- [1] ROHALY A M, CORRIVEAU P, LIBERT J M, *et al.* Video quality experts group: current results and future directions [C]. *SPIE*, 2000, **4067**: 742-753.
- [2] WINKLER S. Digital video quality: vision models and metrics [M]. Switzerland: John Wiley and Sons, 2005: 71-156.
- [3] VRANJES M, RIMAC-DRLJE S, GRGIC K. Locally averaged psnr as a simple objective video quality metric [C]. GRGIC M. 50th International Symposium ELMAR, Zadar, Croatia, 2008: 17-20.
- [4] WANG Z, LU L, BOVIK A C. Video quality assessment based on structural distortion measurement [J]. *Signal Processing: Image Communication*, 2004, **19**(2): 121-132.
- [5] WANG Z, LI Q. Video quality assessment by incorporating a motion perception model [C]. IEEE International Conference on Image Processing, San Antonio, Texas, USA: IEEE Press, 2007, 2: 173-176.
- [6] WANG Z, LI Q. Video quality assessment using a statistical model of human visual speed perception [J]. *Journal of the Optical Society of America A: Optics and Image Science, and Vision*, 2007, **24**(12): B61-B69.
- [7] SESHADRINATHAN K, BOVIK A C. Motion-based perceptual quality assessment of video [C]. *SPIE*, 2009, **7240**.
- [8] ZHANG Wen-jun, WANG Xing-guo, CHEN Guo-bin. A method for objective video quality assessment based on human perceptio [J]. *TV Engineering*, 2004(8): 57-59.
张文俊, 王兴国, 陈国斌. 一种基于视觉感知的视频质量客观评价方法 [J]. *电视技术*, 2004(8): 57-59.
- [9] 叶盛楠. 基于结构相似性的视频/图像质量客观评价 [D]. 北京: 北京工业大学, 2008: 21-31.
- [10] LU Guo-qing, Li Jun-li, CHEN Gang, *et al.* Video quality assessment measurement based on motion information and structural distortion [J]. *Computer Simulation*, 2010, **27**(6): 262-266.
卢国庆, 李均利, 陈刚, 等. 基于运动信息和结构信息的视频质量评价方法 [J]. *计算机仿真*, 2010, **27**(6): 262-266.
- [11] GIROD B. What's wrong with mean-squared error [C]. WATSON A B. Digital images and human vision. MA, USA: MIT Press, 1993: 207-220.
- [12] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity [J]. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600-612.
- [13] CHEN Guan-hao, YANG Chun-ling, PO Lai-man, *et al.* Edge-based structural similarity for image quality assessment [C]. IEEE ICASSP Proceedings, 2006, (2): 933-936.
- [14] CHEN Guan-hao, YANG Chun-ling, SHENG Li-xie. Gradient-based structural similarity for image quality assessment [C]. IEEE International Conference on Image Processing, 2006, 2929-2932.
- [15] ZHANG Jing, SHEN Lan-sun, GAO Jing-jing. Regions of interest detection based on visual attention mechanism [J]. *Acta Photonica Sinica*, 2009, **38**(6): 1561-1565.
张菁, 沈兰荪, 高静静. 基于视觉注意机制的兴趣区检测 [J]. *光子学报*, 2009, **38**(6): 1561-1565.
- [16] DI Hong-wei, LIU Xian-feng. Image fusion quality

- assessment based on structural similarity[J]. *Acta Photonica Sinica*, 2006, **35**(5): 766-771.
- 狄红卫, 刘显峰. 基于图像结构相似度的图像融合质量评价[J]. 光子学报, 2006, **35**(5): 766-771.
- [17] RICHARDSON I E. H. 264 and MPEG-4 video compression video coding for next-generation multimedia[M]. Aberdeen: The Robert Gordon University, 2003. 180-234.
- [18] 张晓星. 基于块匹配的运动估计算法研究与实现[D]. 北京: 北京交通大学, 2008. 7-53.
- [19] LEI Jie, LI Yun-song, ZHOU You-xi, *et al.* An algorithm for multi-spectral image compression based on motion estimation and ROI[J]. *Acta Photonica Sinica*, 2009, **38**(10): 2707-2711.
- 雷杰, 李云松, 周有喜, 等. 基于运动估计和 ROI 编码的干涉多光谱图像压缩[J]. 光子学报, 2009, **38**(10): 2707-2711.
- [20] ITU-R Recommendation BT. 500-11, Methodology for the subjective assessment of the quality of television pictures[S]. Geneva, Switzerland: International Telecommunication Union, 2002.

Video Quality Assessment Based on Gradient and Motion Estimation

LIU Wen-juan, LI Su-mei, ZANG Yan-jun, WEI Jin-jin

(School of Electronics and Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract: Digital video quality assessment plays an important role in video compression, processing and video communications. A method of video quality assessment based on gradient and motion estimation was proposed according to the existing evaluations and human visual system. The new method takes into account the characteristics of the edge information, the temporal correlation between frames, and effects of the scenes' contents changes on the quality of the video. The experiments results show that it is coherent to the human eye's visual effects on the distortion video based on H. 264.

Key words: Video quality assessment; Gradient Based Structural Similarity(GSSIM); Motion estimation; Gradient and Motion Estimation(GMES)