

基于模糊综合评判的客观视频评价模型

纪雯 史浩山

(西北工业大学电子信息学院, 西安 710072)

摘要 为了对视频编码质量进行客观评价,利用模糊综合评判的思想给出了一种客观视频质量评价模型。该模型可同时权衡视频压缩序列的图像清晰度、流畅性、运动抖动性等重要特征,并对其他评价模型和指标具有良好的兼容性。仿真实验结果表明,该模型能够对视频编码序列的多个特征进行较好的综合评价。相比以往评价模型具有实现简洁和可靠度高的优点。

关键词 客观评价 视频编码序列 模糊综合评判

中图分类号: TN919.81 文献标识码: A 文章编号: 1006-8961(2007)11-2006-06

Objective Video Quality Evaluation Model Based on Fuzzy Synthetic Judgment

JI Wen, SHI Hao-shan

(Electronic Engineering Department, Northwestern Polytechnical University, Xi'an 710072)

Abstract Since no well-established methods for objective quality evaluation are available, in this paper, a new efficient and reliable objective video evaluation model is proposed based on fuzzy synthetic judgment. Fuzzy control is combined with the objective evaluation. First, the essential idea of the model is described systematically. Then the fuzzy membership functions of all factors affecting the video quality are determined. And the fact weights are also estimated by proximate approach algorithm. This model highlights a comprehensive evaluation by taking into account various properties of the compressed video including video quality, fluency and motion speed, and has excellent compatibility with other evaluation model or standard. Simulation results demonstrate that the proposed evaluation model outperform some other approaches and achieve more reliability.

Keywords objective evaluation, compressed video sequence, fuzzy theory

1 引言

由于视频压缩算法会导致重构图像的质量降级,所以通常视频编码算法的性能必须根据重构的视频图像质量来评估,然而对视频质量进行评估一直是困扰学术界的难题。由于主观视频质量评价方法耗费成本高,因此在视频通信应用中,对视频编码序列的客观评价及主客观结合评价是目前研究的热点。近年来视频质量专家组(video quality expert group, VQEG)^[1]为此做了大量的工作,但由于几乎不可能建立起一套精确、统一的数学或物理模型,因

此在一定程度上影响了视频质量客观评价方法的准确性。同时由于目前的通信系统对视频传输业务要求越来越高,这就意味着需要一种支持实时视频质量评价,并且还可对视频通信的联合处理提供指导性准则的评价方法,评价模型需简洁且便于硬件实现,以提高视频通信系统的有效性。而现有的众多种评价模型通常是有效性和复杂性的折中,例如多通道HVS(human visual system)评判模型^[2]、基于失真感知的模型^[3]、支持运动特性的分层加权模型^[4]、3维加权信噪比模型^[5]、经验模型^[6]、滤波模型^[7]等。但结合到视频通信应用中这些模型实现起来相对复杂,尤其在传输实时业务时很难在传输

收稿日期:2006-01-11;改回日期:2006-08-14

第一作者简介:纪雯(1976~),女,西北工业大学通信与信息系统专业博士研究生。主要研究方向为无线网络与多媒体通信等。
E-mail: coolman-coolman@163.com

前或传输过程中对视频流做出快速准确的评判。本文将模糊信息控制理论^[8]结合到视频评价体系中,采用模糊综合评判的思想给出一种建立评判标准的快速算法。将影响视频流的图像清晰度、流畅性、运动抖动等重要特征因素采用有限域视频质量因素集的子集来描述,将评价过程转化为模糊综合评判的正问题求解模型,由于因素子集支持视频的多种特性,因此相比以往评价模型,该评判模型具有可靠度高的优点。同时由于视频质量因素集的各向量均是协同考虑通用性和求值复杂度做权衡得出的,因此算法具有复杂度低且实用性强的特点。

2 基于模糊综合评判的客观视频质量评价模型

定义 1 假设有限域视频质量影响因素集 $U = \{u_1, \dots, u_n\}$, 评价集 $V = \{v_1, \dots, v_m\}$, R 是 U 到 V 的一个模糊关系, $R = (r_{ij})_{n \times m}$, 且 $0 < r_{ij} < 1$ 。若有 U 上的任一个模糊子集 A , V 上的模糊子集 B , 它们之间满足 $B = A \circ R$, 则为找到将 U 上的模糊集变为 V 上的模糊集的一个评判。

定理 1 (扩展原理) 给定 $R \in F(U \times V)$, 存在映射 $T_R: F(U) \rightarrow F(V)$, 有 $A \mapsto T_R(A) = A \circ R$ 。则可唯一确定一个从 U 到 V 的模糊评判。

这样寻求一个有关视频质量的评判准则可等效为一个模糊综合评判正问题的求解, 则已知 A 和 R 求 B 的问题。使问题得到简化。该问题描述为几个方面:

(1) 一个已知的评判对象的视频质量因素集 $U = \{u_1, \dots, u_n\}$ 。

(2) 一个已知的评价决策集 $V = \{v_1, \dots, v_m\}$ 。

(3) 一个已知的对各因素的权重分配, 即为 U 上的模糊集 A , 用向量表示为 $A = \{a_1, \dots, a_n\}$ 。其中, a_i 为第 i 个因素 u_i 所具有的权重数。

(4) 获取单因素评判矩阵 $R = (r_{ij})_{n \times m}$, 它由 n 个 V 上的模糊子集, 单因素评判向量组成。其中 $R_i = (r_{i1}, \dots, r_{im})$ 。

(5) 进行 U 到 V 的模糊变换 $A \circ R$, 即合成运算, 得 V 上的模糊子集 B , 即对评判对象得模糊综合评判结果。

2.1 单因素级模型的建立

视频质量影响因素集 U 可依据准则划分为 l 个子集 $U_i, (i = 1, \dots, l)$, 即 $\bigcup_{i=1}^l U_i = U$ 。则对各因素集

按模型方法进行综合评判, 则有评判向量 $R_i = (r_{i1}, \dots, r_{im})$; 相应地有 $R_i = (r'_{i1}, \dots, r'_{im})$; 将各评判结果向量组成评判结果矩阵 $R = (B_1, \dots, B_l)^T = \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{l1} & \dots & b_{lm} \end{pmatrix}$ 。各因素得权重分配为 $A = (a_1, \dots, a_l)$, 其中 $i = 1, \dots, l$ 。则最终评判结果为

$$B = A \circ R = (a_1, a_2, \dots, a_l) \circ \begin{pmatrix} b_{11} & \dots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{l1} & \dots & b_{lm} \end{pmatrix} \\ = \left(\sum_{i=1}^l a_i b_{i1}, \sum_{i=1}^l a_i b_{i2}, \dots, \sum_{i=1}^l a_i b_{im} \right)$$

2.2 多因素级模型的建立

视频质量影响因素集 U 包含 l 个子集 $U_i, (i =$

$1, \dots, l)$, 即 $\bigcup_{i=1}^l U_i = U$ 。对各因素集同样可按模型方法进行综合评判, $B_i = A_i \circ R_i$ 。其中 A_i 为因素子集 U_i 上的权向量, R_i 为对 U_i 的单因素评判矩阵。将各评判结果向量组成合成单因素评判结果矩阵(见 2.1 节)。根据各 U_i 在 U 中所占的地位确定权重分配, 得向量 $A_j = (a_1, \dots, a_l)$ 。做第 2 级模糊变换 $B = A_j \circ R_j$, 得第 2 级评判结果。若 A_j 仍旧只是子集, 则按照逐级评判建立最终评判结果为 $B = A \circ R$, 这样可使得本模型具有良好的兼容性。

2.3 近似试探法求解综合评判的权重分配矩阵

在得出合理的评判结果时, 往往很难确定最佳权重分配矩阵, 但描述为综合评判的逆问题则可大大简化求解, 采用近似试探法^[9]如下:

(1) 设已知评判结果 B 和单因素评判矩阵 R , 求权重分配向量 X , 使满足模糊关系方程 $X \circ R = B$ 。

(2) 可预推测可能得解集 $\{X_1, X_2, \dots, X_k\}$, 通过合成求出 $B_i = X_i \circ R$ 。其中 $i = 1, 2, \dots, k$ 。

(3) 计算贴适度 $n(B_i, B)$, 并用最抖下降法改变步长, 然后用择近原则取最贴近对应的 X'_i 为 X 的近似解: $n(X'_i \circ R, B) = \bigwedge_{j=1}^m n(X'_i \circ R, B)$ 。

(4) 对多因素级权重分配向量 X , 则对上述问题应用递归算法求得。

2.4 支持运动特性及流畅度的客观质量评价模型

由于视频不同于图像, 如果讨论单帧视频质量就转化为讨论图像质量了, 对整体视频序列的意义不大。所以本文选取的压缩视频序列以一个 GOP (group of pictures) 为单位, GOP 的帧数量为 n 帧, 其

中包含 I 帧数量为 1。

定义 2 析取运算: A 或 B , 记 $A \vee B$; 合取运算: A 且 B , 记 $A \wedge B$; 蕴含运算: A 蕴含 B , 记 $A \rightarrow B$; 不相容运算: A 与 B 不相容, 记 $A \wedge B = \emptyset$ 。

定义 3 $A, B, C \dots$ 的可靠性表示 $P(A), P(B), P(C) \dots$ 。

性质 1 $0 \leq P(A) \leq 1; P(\Phi) = 0, P(X) = 1$ 。

性质 2 $P(A \vee B) = P(A) \vee P(B)$;

性质 3 $P(A_1 \wedge A_2 \wedge \dots \wedge A_n) = P(A_2 | A_1) \dots P(A_n | A_1 A_2 \dots A_{n-1})$

性质 4 若 $A \rightarrow B$, 则 $P(B | A) = 1$, 有 $P(A \wedge B) = P(A) \geq P(A \wedge \bar{B})$

定理 2(归纳推理) 若 A 蕴含 B_1 , 蕴含 B_2, \dots 。且 B_1, B_2, \dots 是更容易证实的, 则当 B_1, B_2, \dots, B_n 逐个被证实后, n 越大则 A 的可靠性越大。

证明 当 $P(B_n | A) = 1, P(A \wedge B_1 \wedge \dots \wedge B_n) \neq 0 (n \geq 1)$ 时

根据性质 1, 2, 3, 可得

$P(A) \leq P(A | B_1 \wedge B_2 \wedge \dots \wedge B_n) \leq P(A | B_1 \wedge B_2 \wedge \dots \wedge B_{n-1})$ 。证明完毕。

定理 3(逆向推理) 若 A 与 B 是互不相容的, 则 B 的否定增加了 A 的可靠性。

证明 因为 A 与 B 不相容: $A \wedge B = \Phi, P(A \wedge B) = P(\Phi) = 0$; 又 $P(A \wedge B) = P(A) P(B | A)$, 故有 $P(A) \leq P(A | \bar{B})$ 。证明完毕。

引理 1 当因素集为未知有限域视频质量所有因素集的子集时, 评判结果是最终评判标准的子集。

证明 由析取运算及定理 2, 定理 3, 当因素子集为蕴含时命题成立。证明完毕。

定义 4 有限域视频质量因素集 $U = \{u_1, u_2, u_3, u_4\}$, 其中 u_1 为 MSE, u_2 为帧时间间隔, u_3 为帧的运动特征, u_4 为帧的量化系数。

对得出的各评价向量的说明如下:

$R_1 = \{MSE_1, \dots, MSE_n\}$: MSE 为图像均方差,

$MSE = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [x(i, j) - \hat{x}(i, j)]^2$, 与峰值信噪比 (peak signal to noise ratio, PSNR) 关系为 $PSNR = 10 \log(255^2 / MSE)$ 。这一评价向量符合视频 GOP 内每帧的图像质量评价标准。

$R_2 = \{\Delta t_1, \dots, \Delta t_n\}$ 帧间隔是衡量一段视频流的重要尺度, 通常间隔越大, 其流畅性越差, 尤其当 MPEG 系列和 H. 26X 系列的内核均支持 frame skip 模式, 这样使得其编码帧在时间上有跳跃性, 造成视

觉上的抖动, 在以 PSNR 为视频衡量指标时, 无法描述这一特性, 所以采用帧间隔做评判因素可较好的补充原有评价尺度欠缺时域特征的缺点。

$R_3 = \{motion_1, \dots, motion_n\}$ 通过运动估计取得的图像中每个块的空间偏移量, 由于其偏移量的获得跟运动估计采取的方法及运动矢量的预测方式有较大的关联, 但其运动矢量的特征值向量集通常符合一定共性, 可以作为客观评价运动性指标。将运动特征引入到各帧图像的清晰度评价中, 可以和现有的静止图像评价质量很好地相结合, 充分体现出现视频流的运动性特点, 这也是以往采用 PSNR 准则无法描述的视频流重要特征。

$R_4 = \{\sigma_1, \dots, \sigma_n\}$ 量化参数集合, 通常量化参数越小, 图像越清晰, 之所以增加这个指标做评价向量, 是因为目前视频压缩算法多包含如码率控制、最优运动搜索等技术, 这些技术的很多快速算法都跟量化参数具有一定的函数关系, 因此采用该指标作为评价向量会增强系统可靠性。

对有限域视频质量因素集的分析: u_1, u_2, u_3, u_4 呈现“或”及“蕴含”关系, 依据归纳定理, 得出该评价模型比以往侧重衡量单视频特性的评价模型, 其评价结果更为可靠。

从以上向量特性可以看出, 评价向量值的变化程度与视频的主观质量评价成反比(非反比例)关系, 符合 MSE 评判趋势, 理论上评判结果越接近 0 值时, 表明视频序列性能最佳。鉴于各个因素单位不尽相同但影响趋势一致, 因此也可选择在最终评判时对向量做归一化处理。

对各因素的权重分配向量 $A = (a_1, a_2, a_3, a_4)$ 的说明如下:

当 $A = (1, 0, 0, 0)$ 时, 退化为目前普遍采用的 PSNR 评价体系。其中对评判值做运算 $PSNR = 10 \log(255^2 / MSE)$ 。即得最终的 PSNR 值。

当 $A = (0, 1, 0, 0)$ 或 $(0, 0, 1, 0)$ 时, 退化为目前的视频流畅性或运动性评价标准。

所以目前的视频质量客观 PSNR 评价体系、抖动流畅性评价体系实际为本评价模型的一个子集。

依据 2.3 节, 采用近似试探法求解综合评判的权重分配向量 $A = (0.02, 0.8, 0.1, 0.001)$ 。当然这个权重分配向量也可作为初值, 可对评价体系采用闭环回路控制进行阶段调整, 以达到全局最优权重分配矩阵。这也可作为多因素级评判模型的求解。

本文提出的模型重在表述一种建立评判标准的

方法,以上视频质量因素集的向量均是协同考虑通用性和求值复杂度做权衡得出的,因此该算法实现复杂度低。对于目前客观评判准则,如多通道 HVS 模型^[2]、基于失真感知的模型^[3]、支持运动特性的分层加权模型^[4]、3 维加权信噪比模型^[5]、经验模型^[6]、滤波模型^[7]等均可作为有限域视频质量因素集的子集,依据归纳定理和逆向定理,当这些模型之间为“蕴含”或“或”关系时,综合评判模型的可靠度将增加,依据逆向定理,当这些模型之间不相容时,则在引入时需证明其否定关系才会增加综合评判系统的可靠度。

3 实验结果与分析

为验证本文算法的有效性,以 H. 264 视频编码标准为基础,对 QCIF (quarter common intermediate format) 标准序列做评价测试。其中^[9]:

Foreman 序列是所有视频序列中活动性最强的序列,它包括不同对比度和活动性的头肩像类型,包括杂乱的背景、高强度的噪声和大量前景对象的双向运动。

Carphone 序列有着很多细节的运动背景,在运

行的车上讲话,头部晃动大,有着较多运动的前景对象和不一致的变化背景。

Suzie 序列有着高对比度和适中的噪声,包含在静止和纯文本背景下,在前景位置上人物拿着电话听筒,做着快速的头部移动。

Miss_am 序列活动性稍多,在静止的镜头前活动肩膀和头部。

Claire 序列统一且静止的背景,具有少量的活动,仅有抬头,或动嘴唇及睫毛。

Grandma 序列统一且静止的背景,具有极少量的活动,仅仅是动一下嘴唇或睫毛。

用 3 维直方图表示该模型对上述序列整体评判的结果,从图 1(a)~图 1(f)可以看出其评定结果与主观 DSCQS (double stimulus continuous quality scale) 评定一致,这样可以作为客观指导原则,给出获取最佳的视频传输处理方案所采取的视频序列评判准则。图 4 及图 5 给出针对视频序列的视频质量因素集的权重影响及其结果。

测试视频清晰度及流畅性评价结果时以 foreman 为例进行测试,原始视频帧频为 30Hz,对比帧频为 10Hz (图 2)。从图 3 及图 1(a)的对比可清

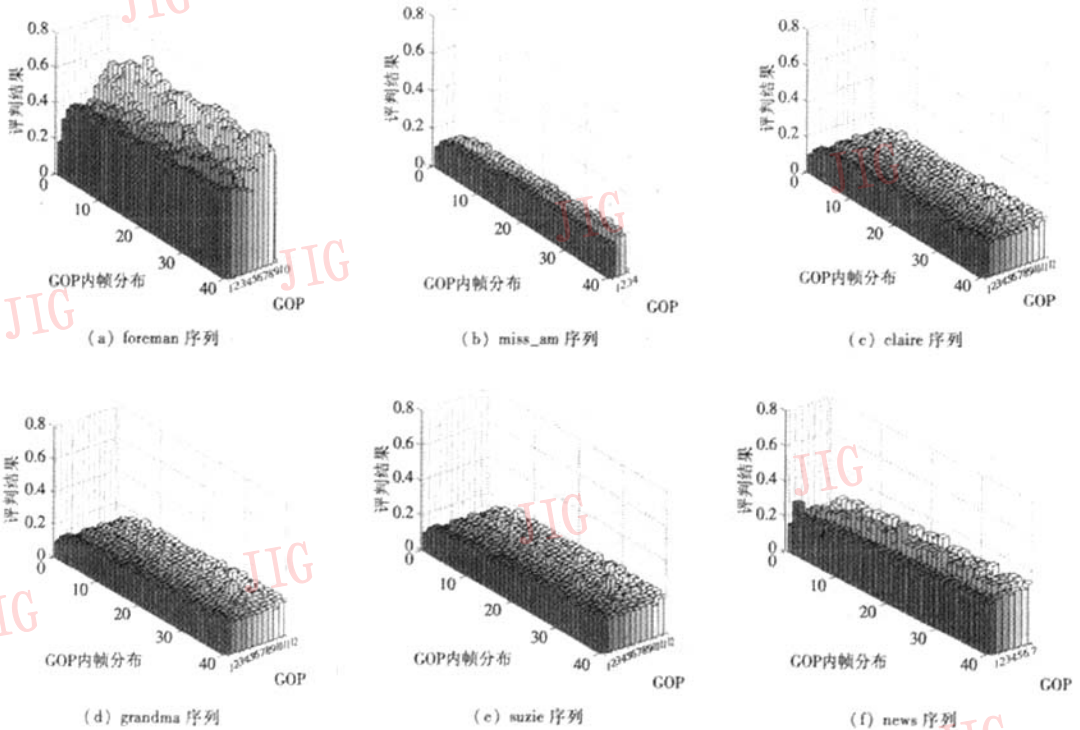


图 1 各个标准视频序列的综合评判结果

Fig. 1 The overall evaluation results of each standard sequence

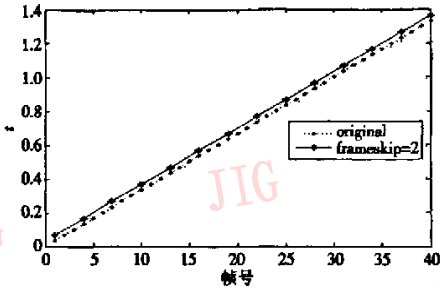


图 2 两段视频序列的采样关系(GOP1 为例)
Fig. 2 Sampling relation between two sequences

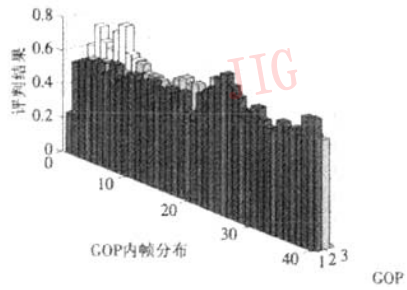
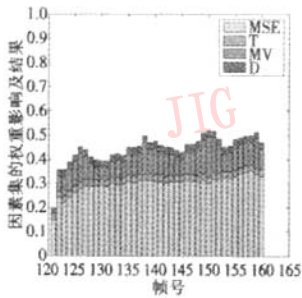
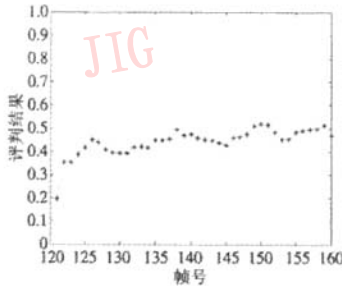


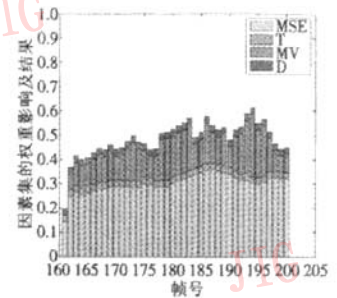
图 3 10Hz foreman 序列的综合评判结果
Fig. 3 Evaluation result of foreman sequence of 10Hz



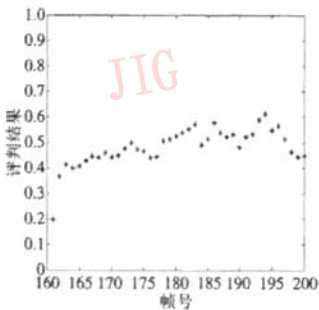
(a) GOP4 时的因素集权重影响及结果



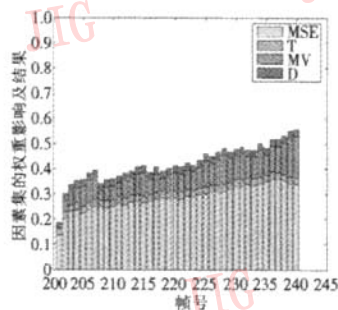
(b) GOP4 的综合评判结果



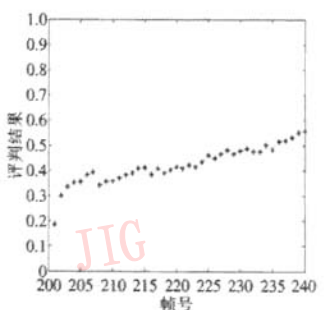
(c) GOP5 时的因素集权重影响及结果



(d) GOP5 的综合评判结果



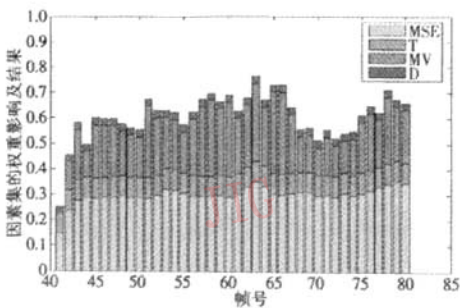
(e) GOP6 时的因素集权重影响及结果



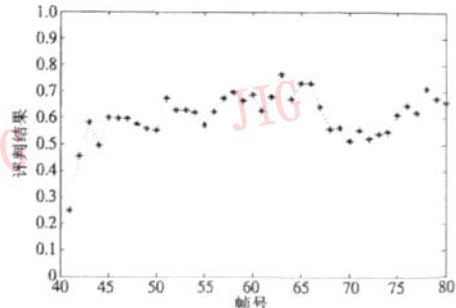
(f) GOP6 的综合评判结果

图 4 30Hz forman 120 ~ 240 帧的评判因素影响显示及测试结果

Fig. 4 120 ~ 240 frames evaluation results and weight factor affects of foreman sequence of 30Hz



(a) 因素集权重影响及结果



(b) 综合评判结果

图 5 10Hz forman 对应 41 ~ 80 帧的评判因素影响显示及测试结果

Fig. 5 Corresponding 41 ~ 80 frames evaluation results and weight factor affects of foreman sequence of 10Hz

楚的显示出本模型对视频流畅性及运动性评判的结果,说明该模型能够很好地反映出视频序列的运动性及相关性特性。图 4 及图 5 给出针对 foreman 运动性最强的片段测试结果,其中 GOP 间隔为 40,由实验结果可以看出,在 MSE 或 PSNR 方面,虽然在 10Hz 时压缩视频的清晰度并不会太大影响,但对 10Hz 视频序列的综合评判比 30Hz 的评判结果差,对比同时间段 foreman 编码序列评价图 5(a)与图 4(a)、图 4(c)、图 4(e),可看出此时因素集权重影响主要体现在 R_2 和 R_3 因子带来的影响上,由于采样间隔加大、帧序列间运动性加剧,导致视频流畅性差,同时视频抖动加剧,因此评判模型反映出 30Hz 更加流畅。这表明了综合评判模型在衡量视频序列综合指标的优越性。

4 结 论

目前的视频质量评价模型主要用于对视频编码器及解码器效果的评价以及其内部的各种算法优化,很多基于 HVS 及其衍生评价算法虽然从结果上接近主观评价,但其复杂度高,而且在实现时需专有的代码模块及算法实现,很难在视频流传输尤其是实时传输时,针对差错弹性、流量控制、视频转码等做出指导性原则。本文提出基于模糊综合评判的模型,结合视频特性并利用编码流已有参数进行快速评判及指导,模型简洁且实现复杂度低,便于硬件实现,同时可将其他模型作为多评判因素集的扩展,使

本模型具有很好的兼容性。实验结果表明,本文提出的综合评判模型能很好的评价视频编码的流畅清晰度、运动特征及流畅性。

参考文献 (References)

- 1 <http://www.its.bldrdoc.gov/vqeg/>.
- 2 Jing Guo, Van Dyke-Lewis. Gabor difference analysis of digital video quality[J]. IEEE Transactions on Broadcasting, 2004, 3(50): 302 ~ 311.
- 3 Ong E, Lin W, Lu Z, et al. Quality evaluation of MPEG-4 and H.261 coded video for mobile multimedia communications[A]. In: Proceedings of 7th International Symposium on Signal Processing and Its Applications[C], Paris, France, 2003: 473 ~ 476.
- 4 Maury M, Hemami S S. Perceived quality metrics for low bit rate compressed video[A]. In: Proceedings of International Conference on Image Processing[C]. Rochester, New York, USA, 2002: III-49 ~ III-52.
- 5 Nannan Wang, Qiang Zhang. Objective quality evaluation of digital video[A]. In: Proceedings of IEEE Asia-Pacific Conference on Circuits and Systems[C], Tianjin, China, 2000: 791 ~ 794.
- 6 Kuhmanch C, Schremmer C. Empirical evaluation of layered video coding schemes[A]. In: Proceedings of International Conference on Image Processing[C], Thessaloniki, Greece, 2001: 1013 ~ 1016.
- 7 Steiger O, Ebrahimi T. Evaluating perceptually prefiltered video[A]. In: Proceedings of IEEE International Conference on Multimedia and Expo[C], Nafplion, Greece, 2005: 1290 ~ 1293.
- 8 Cao Xie-dong. Fuzzy Information Processing and Application[M]. Beijing: Science Press, 2003: 45 ~ 47. [曹晓东. 模糊信息处理及应用[M]. 北京: 科学出版社, 2003: 45 ~ 47.]
- 9 Sadka A H. Compressed Video Communications[M]. New York: Wiley Press, 2002.