

细粒度可伸缩小波视频编码研究

郝燕玲, 王文答

HAO Yan-ling, WANG Wen-da

哈尔滨工程大学 自动化学院, 哈尔滨 150001

College of Automation, Harbin Engineering University, Harbin 150001, China

HAO Yan-ling, WANG Wen-da. Research on fine granularity scalable wavelet video coding. Computer Engineering and Applications, 2010, 46(7): 21-24.

Abstract: For the video application in the uncontrollable network, a new wavelet video coding scheme is proposed. It is suitable for the application of dynamic bit rate transmission and control according to each user's real-time network bandwidth. Its aim is to provide a sustained, clear and fluent experience for video audiences with general optimization in the unstable network. In the wavelet video coding scheme, the W5/3 lift wavelet algorithm is simplified and the hierarchical multi-reference-frame prediction method is introduced into MCTF, the transmission layer stream packing method is adopted according to the data priority, the wavelet embedded block coding method is used based on bit-length search. The test shows that the algorithm proposed in this paper has high compression efficiency and fine scalability in quality and can achieve higher user experience.

Key words: wavelet video coding; dynamic bit rate; media transmission; media control

摘要: 针对不可控网络下的视频应用, 提出一套新的小波视频编码方案, 适用于根据每个用户实时网络状况进行动态码率传输和控制, 目的是为用户在不稳定的带宽下提供持续清晰流畅的全局最优的视频体验。该编码方案简化了 W5/3 提升小波, 在 MCTF 算法中引入层次化多参考帧预测, 并采用新提出的基于优先级的传输层打包方法以及基于块位长的嵌入式零块编码算法。经过测试表明, 提出的算法压缩效率高, 质量可伸缩性好, 能够有效提升用户体验。

关键词: 小波视频编码; 动态码率; 媒体传输; 媒体控制

DOI: 10.3778/j.issn.1002-8331.2010.07.007 **文章编号:** 1002-8331(2010)07-0021-04 **文献标识码:** A **中图分类号:** TN919.81

网络视频业务这几年正以惊人的速度增长。宽带的普及和更多的高清视频内容, 将促使更多的人逐渐习惯通过 Internet 观看视频。另外, 3G 移动网络的兴起也为手机视频业务的推广铺平了道路。

在带宽波动较大的不可控网络环境如 Internet 或移动网络中开展视频业务, 如何提供有效的 QoS(Quality of Service)保障机制, 确保为用户持续提供清晰流畅的视频体验是至关重要的。提出的细粒度可伸缩小波视频编码主要应用于不可控网络中的视频传输, 目的是在带宽波动较大的网络环境下实时跟踪网络状况, 实现精细动态码率调节, 为用户提供持续清晰流畅的视频体验。下面介绍业界在这方面的做法及其优劣势:

(1) 采用多码率文件或多码率媒体源, 由流媒体服务器根据带宽情况在预先编码的多种码率中进行切换。这种方案实现简单, 但需要预先编码生成多码率的文件, 并且无法真正适应复杂多变的网络环境, 码率切换条件难以判断和控制, 码率切换一般跨度较大, 容易导致用户体验急剧变化。

(2) 流媒体服务器采用二次编码的方法, 即在服务器端保存一个较高质量的文件, 在用户请求视频服务时先解码该视频文件, 再根据与用户之间的链路带宽情况实时调整编码参数生成相应码率的媒体流, 实现动态码率传输和控制目的。该方案

能够较快跟踪网络环境的变化, 达到较好的业务体验, 但是流媒体服务器对每用户均需二次编解码, 性能消耗过大, 成本过高, 无法支撑大规模的应用。

提出的细粒度的可伸缩小波视频编码适用于精细的动态码率传输和控制, 可大大降低出现视频播放过程中的停顿、缓冲、中断, 视频画面中的马赛克、拖影等严重影响用户体验的现象发生的概率。经过测试表明, 提出的编码压缩效率高, 质量可伸缩性好。

1 小波视频编码方案

1.1 小波视频编码总体框架

小波视频编码主要有两种框架, 一种是“T+2D”, 即先对图像组(GOP)进行运动补偿时域滤波(MCTF)^[1]形成时间方向上的高低通帧, 接着对每帧进行空域的二维离散小波变换。另一种则是先对 GOP 的每一帧进行空域的二维离散小波变换, 再在小波子带中进行 MCTF。

该方案采用的是“T+2D”的方法, 编码总体框架如图 1 所示。

首先, 视频序列经过 MCTF 降低时间方向上的各帧之间的相关性。这里采用简化的 W5/3 小波杠铃提升^[2]重叠块运动补

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60704018)。

作者简介: 郝燕玲(1944-), 女, 教授, 博导; 王文答(1979-), 男, 博士。

收稿日期: 2009-10-19 修回日期: 2009-12-17

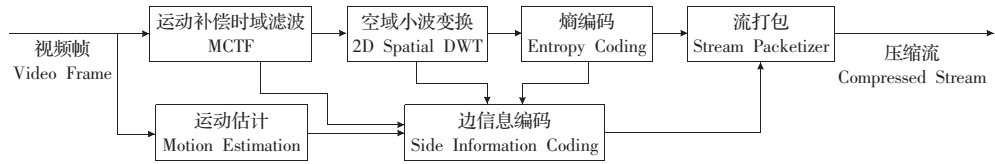


图1 小波视频编码总体框图

偿(OBMC)算法。

小波提升算法的预测环节如公式(1)所示,更新环节如公式(2)所示。

$$H_i = x_{2i+1} + \alpha(x_{2i} + x_{2i+2}) \quad (1)$$

$$L_i = x_{2i} + \beta(H_{i-1} + H_i) \quad (2)$$

当 $\alpha = -1/2, \beta = 1/4$, 即为 W5/3 小波提升算法。

基于在动态码率传输和控制中确保单帧画面清晰度的考虑,提出简化的 W5/3 小波滤波器,即只有预测环节,采用公式(1)进行高通滤波,取消了更新环节,不再通过公式(2)进行低通滤波。

这样虽然对原始帧的编码效率稍有影响,但是却弥补了按照正常小波提升算法进行 MCTF 时若丢弃任何残差帧(高通帧)均会影响所有未丢弃帧的图像质量这一不足。而通过引入多参考帧^[9]则进一步减少简化的 W5/3 提升算法对编码效率的影响。接着,对经过 MCTF 的每一帧在空间上进行二维小波变换以降低空间方向上的相关性。

第三,对经小波变换后的每帧系数采用基于块位长搜索的嵌入零块编码算法^[10]进行编码形成嵌入式压缩码流,内存占用低,运算效率高。

第四,对解码器解码必须的边信息包括参数集、运动矢量、宏块模式及其他控制信息等进行熵编码。

最后,对系数和边信息的压缩数据按优先级进行封装可根据任意目标码率提取而保持该码率下的最优视频效果的压缩流。

1.2 帧结构和可选参考帧集合

该方案在提升算法预测环节突破 MCTF 固定参考帧顺序的模式,允许根据当前帧所处的不同位置按一定的规则和方法选择最优的参考帧。

图2所示的是该方案提出的简化的 W5/3 提升算法预测环节参考帧顺序的一种最典型例子,该例子每 GOP 包含 8 帧(1~8)原始图像,帧 0 为上一个 GOP 的 I 帧,经过 3 级 MCTF 分解得到分布在 3 级分解上的 7 帧高通子带和一帧原始图像(I 帧)。

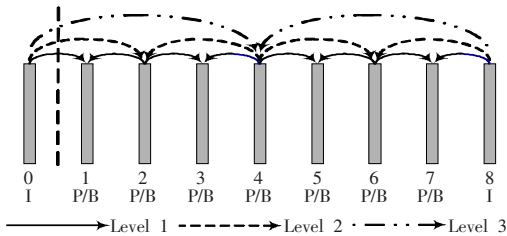


图2 时域滤波帧结构图

第一级分解是在序号 1、3、5、7 位置分别由其相邻的 2 帧图像的预测残差形成的高通子带(帧)。

第二级分解是在第一级分解剩下的 2、4、6、8 中的 2、6 位置分别由其相邻的 0、4、8 中的 2 帧图像的预测残差形成高通子带(帧)。

第三级分解是在第二级分解剩下的 4、8 中的 4 位置分别有其相邻的 0、8 这 2 帧图像的预测残差形成高通子带(帧)。

为了进一步提高预测精度,提出新的多参考帧选择策略。每级分解不再限制仅仅可以由当前帧相邻的 2 帧图像作为参考帧进行预测,而是比当前帧更高级别位置的图像均可以当成参考帧。这样可以在这些参考帧中进行最佳的运动估计。

设某 GOP 包括 N 级 MCTF 分解,则 GOP 的图像帧数为 2^N 。基于时间轴可伸缩原理,在解码器侧的缓存参考帧的最小帧数 $MinRefFrsCnt$ 值是 $N+1$ 。考虑到硬件条件限制,假设参考帧的最大帧数为 $MaxRefFrsCnt=16$,对于不同的应用环境,可通过指定编码参数 $RefFrsCnt$ 来确定缓存参考帧数量(最大可能使用数量)。此时参考帧应符合式(3)。

$$\begin{cases} RefFrsCnt \leq MaxRefFrsCnt = 16 \\ RefFrsCnt \geq MinRefFrsCnt = N + 1 \end{cases} \quad (3)$$

$RefFrsCnt$ 确定之后,就可以按照规则确定各帧预测时的可选的参考帧集合。GOP 原始图像 N 级 MCTF 分解帧间关系构成一棵二叉树,如图3所示的带箭头的黑色实线。从图中可以看出, $RefFrsCnt$ 最小值是 5,除了缓存上一组 GOP 的 I 帧 0 外,每级分解至少需要一帧参考图像。例如对于第 1 帧,最大可能使用的参考帧数量至少是第 0、16、8、4、2 帧。

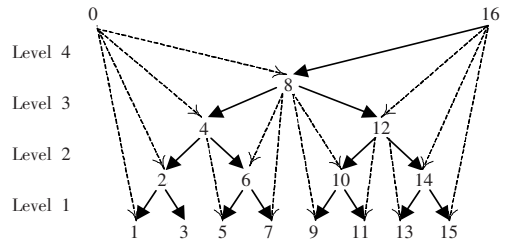


图3 GOP 分解帧间关系构成的二叉树

为了达到最佳的参考帧预测效果,在满足 $RefFrsCnt$ 参考帧数限制条件下,选择参考帧集合时需要的是每个参考帧的集合越大越好,也就是可参考帧数越多越好。基于时间轴上可伸缩要求,低级别的帧在带宽不足情况下可能被丢弃,故处于低级别的帧仅能以高级别的帧作为参考。因此选择参考帧集合时离二叉树根节点近的节点优先级越高。

另外,由于一般情况下距离当前帧越近的参考图像相关性越大,故在选择某节点的参考节点时同一优先级下距离当前节点越近的节点优先选入。举两个例子比较一下,在图3中,当 $RefFrsCnt=7$ 时,第 1 帧的参考帧集合是{0, 16, 8, 4, 2, 12, 6};第 11 帧的参考帧 {0, 16, 8, 4, 12, 10, 14}, 选择帧 14 的原因是第 11 帧距离第 14 帧比第 6 帧近。

1.3 MCTF 运动补偿时域滤波方案

本节主要描述如何在可选的参考帧集合中进行运动估计,并搜索出最佳预测进行 MCTF 运动补偿时域滤波形成残差帧(高通帧)。

在上节中可选的参考帧集合确定之后,下一步问题就是如何针对当前帧的每个宏块在这些可选参考帧中搜索出最佳的宏块划分模式^[11]($16 \times 16, 8 \times 16, 16 \times 8, 8 \times 8, 8 \times 8+, 8 \times 8+$ 是指该子块被进一步分解成 $4 \times 8, 8 \times 4, 4 \times 4$,如图4所示)和最佳预测模式(Skip mode、DirInv mode、Bid mode、Fwd mode、Bwd mode、

如图5)。方法就是搜索宏块划分模式和预测模式找出使公式(4)匹配因子 J 取得最小值的最佳估计。

$$J = SAD(MV, Ref) + \lambda(R(MV) + R(Ref)) \quad (4)$$

式中 MV 是运动矢量, Ref 是指参考帧, SAD 是指当前帧和参考帧中对应的预测块的每个像素的绝对误差和。 λ 是拉格朗日乘子, $R(MV)$ 和 $R(Ref)$ 分别表示编码运动矢量和参考帧序号所需的比特数。由于采用了 Barbell 提升算法, 这里的参考帧经过 Barbell 函数插值滤波后的像素数目是原图像的 4 倍。

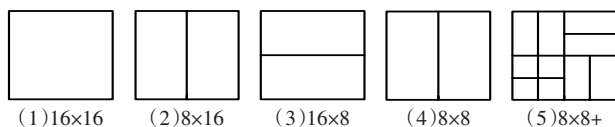


图4 宏块划分模式

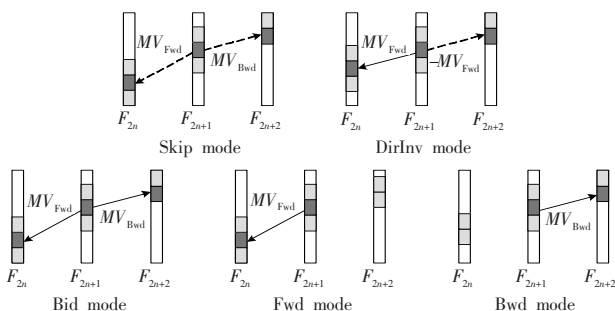


图5 宏块的5种预测模式

为了在保证图像质量的情况下提高搜索效率, 该方案采用类似 H.264 中 UMHexagonS 著名算法^[6]的思想。为了满足 Barbell 提升算法和帧率可伸缩的要求, 提出了如下的搜索流程:

(1) 对于当前帧的每个宏块, 选择 Fwd mode 和 Bwd Mode 分别对可选参考帧集中当前帧之前和之后的两个子集中的参考图像进行 UMHexagonS 搜索, 分别得到当前宏块之前和之后参考帧的最佳匹配块(运动矢量)、宏块划分方法和对应的匹配因子 J 。在前后参考帧子集的搜索过程中均从与当前帧距离最近的参考帧中由近到远的顺序进行搜索。

(2) 根据(1)选中的两帧参考帧采用其他三种预测模式进行搜索和估计, 得到这三种模式下的最佳匹配块(运动矢量)、宏块划分方法和对应的匹配因子 J 。

(3) 选择这五种模式中的匹配因子 J 最小的作为最佳的预测模式。

对当前帧的所有宏块进行以上的运动估计后, 接着就可以根据运动估计的结果对当前帧进行运动补偿。该方案采用的是基于杠铃提升的重叠块运动补偿(OBMC)算法。

1.4 基于优先级的流打包方法

为了便于流媒体服务器进行动态码率传输与控制, 对编码层的压缩数据按图像帧优先级和图像本身数据的优先级排序后进行打包形成网络抽象层(NAL)码流^[7]。这种方案无需对编码层的数据进行解码即可通过简单算法完成几乎任意码率的码流抽取。

在提出的方案中, 打包后的 NAL 码流由一系列 NALU 组成。NALU 的类型包括 I 帧分片、P/B 帧 MV 分片、P/B 帧残差分片、GOP 参数集、图像参数集、帧分隔符、GOP 结束符、流结束符等。NALU 的类型包含在一个字节长度的 NALU 头信息中。

对于类型为 P/B 帧 MV 分片和 P/B 帧残差分片, NALU 头信息中提供一个字段 TemporalLevelMinus1 表示该 P/B 的所处

的时域级别, 值越高越重要, 由该级别以及更高级别的帧构成的视频帧率越低。

如图6所示, 打包后的 NAL 码流根据图像的优先级结合图像显示顺序进行排列。图6中各个图像内部的分片也严格按照嵌入式码流优先级从高到低进行排序。

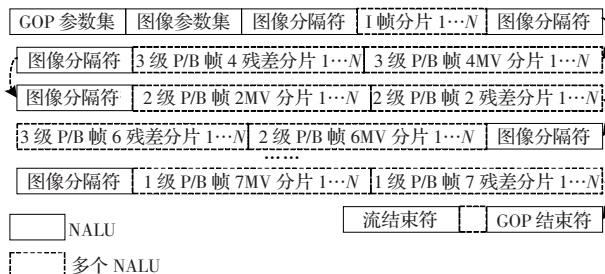


图6 按优先级打包后 NAL 码流

最后, 提出的小波编码方案总结如下: 对于视频序列中的每个 GOP, 根据帧率可伸缩的要求提出多参考帧集合的选择算法, 接着根据预测模式结合 UMHexagonS 进行准确高效的运动估计, 再基于简化 W5/3 小波的杠铃提升重叠块运动补偿(OBMC)算法, 经过 N 级 MCTF 得到 $2^N - 1$ 残差帧, 接着对这 $2^N - 1$ 残差帧和 1 帧关键帧进行空域的 2 维小波变换, 再对经空域 2 维小波变换后的每帧系数采用基于块位长搜索的嵌入零块编码算法进行编码, 最后对图像数据和边信息的熵编码进行打包形成压缩码流。

2 动态码率传输与控制技术

2.1 帧内码率动态微调

小波视频编码方案在空间域采用二维小波提升算法并采用嵌入零块编码算法, 能够支持对从优先级最低的帧的残差分片或 MV 分片进行有选择的丢包或去除包中的部分比特, 快速进行动态码率微调, 进而保证视频在带宽波动的网络环境下流畅播放。而解码器侧通过未丢包的相邻帧进行内插也能够较好地恢复视频质量, 降低对用户主观体验的影响。

2.2 帧率动态调节

由于在 MCTF 编码中采用简化的 W5/3 提升多级分解算法, 省略了更新环节, 不再利用残差帧(高通帧)通过更新环节高优先级的图像帧得到低通帧。因而, 流媒体服务器在实时监测到用户带宽低到一定条件时能够丢弃低优先级的图像数据动态降低帧率而不影响未丢弃帧的图像质量。

另外, 通过提出的 MCTF 层次化的多参考帧的选择算法, 能够支持解码器在接收到低帧率的码流时, 通过未丢弃的运动矢量结合未丢弃的高优先级帧进行内插再提高帧率, 能够减少带宽波动对用户的影响。

3 测试结果分析

对提出的为了方便流媒体服务器进行动态码率传输和控制而设计的小波视频编码算法进行编码性能验证。用标准测试序列 Bus, Foreman, Mobile 和 Football 对算法和 MSRA (Microsoft Research Asia) 小波视频编码算法进行对比测试。选择 GOP 大小为 32, MCTF 级数为 5, 空域小波变换级数为 3, 运动估计的搜索范围为 64, 测试的结果见图 7。从图中可以看出, 提出的编码算法得到的 PSNR 比 MSRA 算法提高 0.1 dB 到 0.4 dB, 在运动激烈的情况下效果更明显。

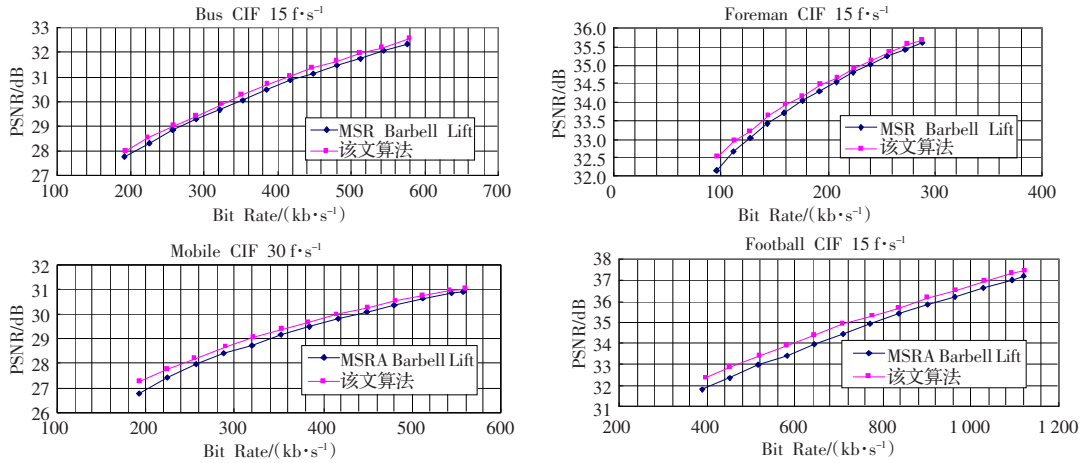


图7 该文编码算法与MSRA算法测试结果对比

4 结论

提出了简化 W5/3 提升小波和多参考帧预测层次化 MCTF 算法,同时采用基于优先级的传输层流打包方法和基于块位长的嵌入式零块编码算法,形成完整的小波视频编码方案。经过对多个标准序列的测试对比,该算法下的 PSNR 比 MSRA 算法提高 0.1 dB 至 0.4 dB,同时具备更好的质量可伸缩性,适用于精细的动态码率传输和控制的媒体应用中。结合传输控制算法,可大大降低视频播放过程中出现停顿、缓冲、中断,视频画面中的马赛克、拖影等严重影响用户体验的现象发生的概率。

参考文献:

- [1] 刘兆广,彭玉华,杨阳.一种基于类 haar 小波的 MCTF 视频编码方案[J].电子学报,2008(2):386-391.
- [2] Xiong Ruiqin,Xu Jizheng,Wu Feng,et al.Barbell-Lifting based 3-D

(上接 20 页)

定理 7 如果 E 是 $V(X)$ 上的模糊熵, $\forall A, B \in V(X)$, 定义

$$d(A, B) = 1 - \frac{E\left(\frac{A \cap B}{2}\right)}{E\left(\frac{A \cup B}{2}\right)}, \text{ 并规定: 当 } E\left(\frac{A \cap B}{2}\right) = 0 \text{ 时, } E\left(\frac{A \cup B}{2}\right) = 1.$$

则 $d(A, B)$ 是 $V(X)$ 上的一个距离测度。

此定理的证明与定理 6 类似,这里不再详述。

5 结束语

通过 Vague 集理论中的一些定义和性质,对 Vague 集的模糊熵、距离测度以及它们之间的关系作了进一步的研究,给出了几个由模糊熵诱导的距离测度公式和几个由距离测度诱导的模糊熵公式,从而说明了 Vague 集的模糊熵和距离测度是可以相互诱导的。这些研究对于讨论 Vague 集间各种度量方法及 Vague 集理论的应用具有重要的意义。

参考文献:

- [1] Zadeh L A.Fuzzy sets[J].Information and Control,1965,8(3):338-353.
- [2] 梁保松,曹殿立.模糊数学及其应用[M].北京:科学技术出版社,2007:65-177.
- [3] Gau W L,Buehrer D J.Vague sets[J].IEEE Transactions on System, Man and Cybernetics,1993,23(2):610-614.
- [4] Deschrijver G,Kerre E E.On the relationship between some extensions of fuzzy set theory[J].Fuzzy Sets and Systems,2003,133(2):

wavelet coding scheme [J].IEEE Transactions on Circuits and Systems for Video Technology,2007,17(9):1256-1269.

- [3] 淑芳,李华.基于 H.264 的多参考帧快速选择算法[J].电子学报,2009,37(1):62-66.
- [4] Wang Wenda,Hao Yanling.A fast wavelet based low memory embedded block coding algorithm[C]//WRI International Conference on Communications and Mobile Computing,6-8 Jan,2009,2:615-619.
- [5] Xiong R,Wu F,Li S,et al.Exploiting temporal correlation with adaptive block-size motion alignment for 3D wavelet coding[J].SPIE,2004,5308:144-155.
- [6] 柴俊华,应骏.自适应搜索长度 UMHexagonS 算法研究[J].中国图象图形学报,2008,13(10):1849-1852.
- [7] Wiegand T,Sullivan G,Bjntegaard G,et al.Overview of the H.264/AVC video coding standard[J].IEEE Trans on CSVT,2003,13(7):560-576.
- [8] Deschrijver G,Kerre E E.On the composition of intuitionistic fuzzy relations[J].Fuzzy Sets and Systems,2003,136(3):333-361.
- [9] Szmidi E,Kacprzyk J.Entropy for intuitionistic fuzzy sets[J].Fuzzy Sets and Systems,2001,118(3):467-477.
- [10] Szmidi E,Kacprzyk J.Distances between intuitionistic fuzzy sets[J].Fuzzy Sets and Systems,2002,114(3):505-518.
- [11] 黄国顺,刘云生.关于 Vague 集的模糊熵[J].计算机工程与应用,2005,41(33):48-50.
- [12] 范平,梁家荣,李天志.Vague 集的新模糊熵[J].计算机工程与应用,2007,43(13):179-181.
- [13] 石玉强,王鸿绪.一种基于 Vague 集间距离的双向近似推理方法[J].小型微型计算机系统,2007,28(4):661-665.
- [14] 石玉强,王鸿绪.Vague 集间距离的新定义[J].琼州大学学报,2006,13(2):34-36.
- [15] 符海东,卢正鼎.基于 Vague 集距离的多评价指标模糊决策方法[J].华中科技大学学报,2003,31(8):77-79.
- [16] 要瑞璞,沈惠璋.基于 Vague 集投影及距离的模糊多指标决策[J].数学的实践与认识,2009,39(2):19-22.
- [17] Liu X C.Entropy,distance measure and similarity measure of fuzzy sets and their relations[J].Fuzzy Sets and Systems,1992,52(6):305-318.
- [18] 李凡,徐章艳,饶勇.Vague 集[J].计算机科学,2000,27(9):12-14.
- [19] Luca A D,Termini S.A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory[J].Information and Control,1972,20(3):301-310.