

一种针对整帧图像丢失的错误隐藏算法

陈建乐 刘济林 陈国斌 薛全

(浙江大学信电系, 杭州 310027)

摘要 在低码率的网络视频通信中,数据包丢失往往导致整帧图像数据的丢失。为了修复丢失的整帧图像数据,提出了一种基于光流场不变性的整帧图像错误隐藏算法。该算法首先使用有理插值对前帧图像的运动矢量场做预处理来得到稠密的运动矢量场;然后采用前向投影技术将前一帧图像内容沿着运动矢量场方向投影到当前图像的对应位置上。由于在图像运动场的突变而形成的重叠和遮挡区域,该算法对其按前景和背景进行区分处理,因而进一步提高了错误隐藏的性能。实验结果表明,和已有算法比较,用该算法修复的图像的峰值信噪比提高了0.18~0.82dB,同时得到了更好的主观图像质量。

关键词 视频通信 错误隐藏 有理插值

中图分类号: TN919.33 **文献标识码**: A **文章编号**: 1006-8961(2007)01-0037-06

An Error Concealment Algorithm for Whole Frame Loss

CHEN Jian-le, LIU Ji-lin, CHEN Guo-bin, XUE Quan

(Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027)

Abstract In low bit-rate packet-based video communications, loss of packet usually results in whole-frame losses. The algorithm dealing with the problem of whole frame loss is proposed in this paper. Rational interpolation is firstly applied to generate accurate dense motion field based on decoded block-based motion field. Then the previous image pixels are projected onto missing frame along the motion vector direction. The identification of overlapped and occluded areas is also used to improve error concealment performance. Experimental results show that the proposed algorithm outperforms other techniques in peak signal-to-noise ratio(PSNR) and provides good visual quality.

Keywords video communication, error concealment, rational interpolation

1 引言

随着视频和网络技术的发展,网络视频在现实生活中得到了越来越多的应用。然而由于目前各种网络,特别是无线网络不能提供完全可靠的传输,因此稳定的网络视频服务仍是当前的热门研究主题。一般情况下,网络视频可分为流媒体视频和实时对话视频两种服务。其中流媒体服务包含视频点播、网上电影院等,在这种情况下,客户端可以使用缓冲区技术来产生较大的延时,以便当出现数据丢失时,客户端有足够的时间要求服务器重传数据。但当网络处于拥塞状态时,由于重传会恶化网络性能,因而

会进一步降低视频服务质量。在实时对话服务中,如可视电话、视频会议等,由于数据的时效性要求很高,因此重传技术带来的时延问题是不能接受的。另外,重传技术会增加传输的带宽要求,这对带宽要求较高的视频传输服务是不合适的。

由于错误隐藏技术不增加传输带宽和信息的时延,因此在有错视频传输中得到广泛关注。目前已有的视频错误隐藏算法^[1~5]由于利用了视频信号在时间和空间上的相关性,在解码端又采用了后处理技术,从而可在一定程度上恢复丢失的图像信息。通常,这些算法都假设一个或者多个连续宏块的数据丢失。虽然这些算法取得了较大的成功,也已在一些传统的视频传输系统中得到应用,但在较低码

收稿日期:2005-07-01; 改回日期:2005-09-02

第一作者简介:陈建乐(1979~),男。2001年于浙江大学信息与电子工程学系获学士学位,现为浙江大学信电系博士生。主要从事视频压缩和传输、视频信号处理等方面的研究方向。E-mail: jlchen@zju.edu.cn

率的网络视频服务中却很少应用。究其原因是,在低码率的情况下,一帧图像的数据量往往小于一个基本网络传输单元,在这种情况下,当网络传输中出现数据包丢失时,则往往对应于一整帧图像内容的丢失。由于大部分已有的假设单个或多个连续宏块丢失的错误隐藏算法不能处理这种情况,因此设计一种能恢复整帧丢失图像的算法,对低码率的网络视频服务具有重要的意义。为此本文提出了一种针对整帧图像丢失的错误隐藏算法。

对于整帧丢失图像的恢复,最简单的方法是使用前面一帧图像内容替代。虽然该方法对基本静止的视频序列能取得较好的错误隐藏效果,但对运动强度较大的视频序列却会产生巨大的失真。为减少运动产生的影响,Belfiore 在假设相邻两帧图像之间的运动场信息相同的基础上,先使用前一帧的运动场来预测当前图像的运动场信息,然后用预测得到的运动场将前一帧图像的内容前向投影到当前图像的对应位置上,最后使用空间域上的中值滤波来得到^[6]未预测到的像素。Baccichet 在最新的视频压缩标准 H. 264/AVC 中实现了实时的整帧图像错误隐藏算法^[7],在该算法中,先假设当前图像的运动矢量在宏块或者子块范围内是常数,然后使用前一帧运动信息预测来得到当前帧各个块的运动矢量,最后根据预测得到的运动矢量找到当前块在前一帧中对应的内容。

在目前的视频压缩标准中,为降低编码比特开销,传输的都是基于块的运动矢量场。由于这种简化的运动场和真实的运动场会存在较大的偏差,因此采用这种运动场预测得到的图像会产生失真。针对这个问题,本文首先使用有理插值函数对解码的运动场进行校正来得到较为准确的稠密运动场,然后使用该运动场通过对前一帧图像进行前向投影来预测当前丢失的图像数据。在对前一帧图像进行前向投影时,必然会出现一些被多次投影到的像素位置(称为重叠区域),还有一些没有被投影到的像素位置(称为遮挡区域),由于这些区域一般出现在图像亮度和运动场的突变处,是最容易产生预测误差的区域,因此合理处理该区域的像素值对提高图像质量意义重大。

2 基于有理插值函数运动场校正

在解码得到的运动矢量场中,一个块内的像素

共用同一个运动矢量值,当一个块包含属于不同物体的像素时,一部分像素使用的运动矢量则会出现偏差,为此,本文算法首先校正部分错误的运动矢量场。由于真实运动场的突变通常出现在图像亮度边界处,因此在图像内容已知的情况下,使用图像亮度值来校正部分错误的运动矢量是可能的。有理函数插值在图像插值和去噪等方面已得到较好的应用,它的特点是能够保持尖锐的图像边缘^[8]。本文算法是先将解码得到的块运动矢量作为块中心位置像素的运动矢量,这样就得到了基本准确的稀疏运动场;然后采用有理函数插值来得到稠密的运动场,其中有理函数的插值系数由被插值点和参考点的亮度差异来决定。

2.1 预处理

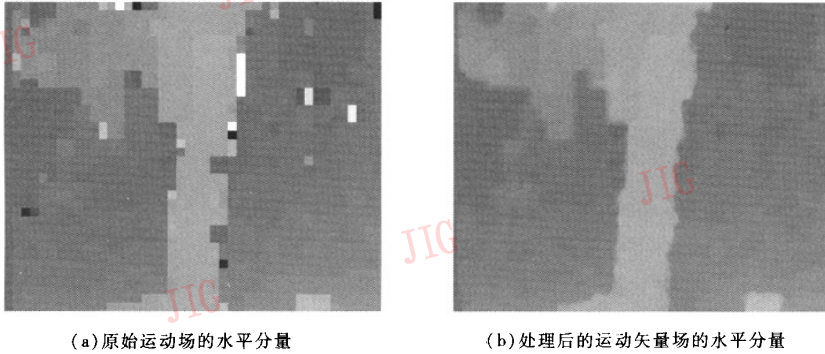
在实际的视频传输系统中,一般解码得到的稀疏运动场中可能会存在错误的运动矢量。图 1(a)给出了 Flower 序列的 H. 264 视频码流解码的运动矢量场的水平分量,从图中可以看出,解码得到的运动场中存在较多的随机错误运动矢量。虽然这些运动矢量在编码时可能可以提高编码效率,然而由于这些错误的运动方向会严重影响采用投影技术的错误隐藏算法的性能,因此在进行有理插值之前,需要更正运动场。从图 1(a)中可以看出,错误块的运动矢量和周围块的运动矢量的相关性非常小。在本文算法中,当一个块的运动矢量和周围所有块的运动矢量的相关性都很小时,则使用周围块运动矢量的中值来替代该块的运动矢量。

2.2 有理插值

假设运动估计采用的块尺寸为 8×8 ,则对大小为 $M \times N$ 的图像,其解码得到的原始稀疏运动场大小为 $(M/8) \times (N/8)$ 。这里使用有理函数,通过多次插值来得到原图像大小的稠密运动场,即每次插值运动场的大小在 X, Y 方向各增大一倍。为降低计算复杂度,这里将 2 维有理函数插值分解成两次 1 维插值实现。图 2 是一次有理函数插值过程的示意图,图中“●”代表已有运动矢量的像素,通过 1 维水平有理函数插值可得到图 2 中标“⊕”的像素的运动矢量,再通过 1 维垂直有理插值来得到图中标“○”的像素的运动矢量。图 2 中 C 点的运动矢量可使用式(1)求得。

$$V_C^{\text{motion}} = (w_A \times V_A^{\text{motion}} + w_B \times V_B^{\text{motion}}) / (w_A + w_B) \quad (1)$$

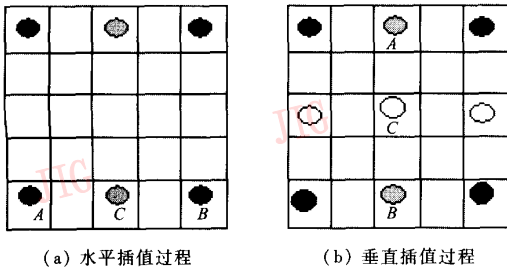
其中, V^{motion} 表示运动矢量, w_A 和 w_B 分别是点 A 和点 B 的加权系数,其可用下式计算:



(a) 原始运动场的水平分量 (b) 处理后的运动矢量场的水平分量

图 1 运动矢量场的水平分量图

Fig. 1 Original motion field and dense motion field generated by proposed algorithm



(a) 水平插值过程 (b) 垂直插值过程

图 2 一维有理插值过程

Fig. 2 Rational filter interpolation procedure

$$w_A = 1 / (k \times (f_A - f_C)^2 + 1) \quad (2)$$

$$w_B = 1 / (k \times (f_B - f_C)^2 + 1)$$

式中, f_A 、 f_B 、 f_C 代表 A、B、C 点的亮度值, k 代表亮度差异的敏感程度, 它的大小决定了有理函数插值的非线性程度, 当 k 为零时, 则有理函数插值变为线性插值。然而随着 k 值增大, 图像噪声对插值效果的影响也越明显。为降低图像噪声以及图像细节纹理对插值效果的影响, 可在进行运动矢量插值前对图像亮度做 3×3 的均值滤波处理。由于加权系数的使用, 使得插值得到的 C 点的运动矢量接近于和它处于同一物体的 A 点或 B 点的值。

2.3 运动场的空间规范化

由于有理函数插值的加权系数由像素点间的亮度差异决定, 这时物体或者背景内部的纹理细节必然会影响到有理函数插值的精度, 从而使运动矢量场内部形成噪声。为降低噪声影响, 可使用 5×5 中值滤波器来对有理函数插值后的运动矢量场进行滤波, 因为中值滤波可以在平滑噪声的同时, 保持尖锐的边缘信息。为降低中值运算复杂度, 这里使用 1

维中值滤波, 即首先进行 5 次 1×5 的水平中值滤波, 再进行一次 5×1 的垂直中值滤波。

图 1 给出了解码的原始运动矢量场和通过预处理、有理函数插值和中值滤波后的稠密运动矢量场的水平分量图, 由该图可以看出, 在原始运动矢量场中, 特别是处于物体边缘区域的错误运动矢量得到了有效地校正, 这就为后续的投影预测操作提供了更准确的运动矢量场, 以降低修复图像的误差。

3 基于光流不变的整帧图像恢复

上一节已得到了前一帧图像校正后的稠密运动场。假设当前丢失图像的运动矢量场和前一帧保持一致, 则根据前一帧的图像内容和运动矢量场就可以预测得到当前图像的内容。本文采用前向投影方法来将前一帧图像的各个像素沿着运动矢量的方向投影到当前图像的相应位置上。运动矢量的精度为半像素, 由于前一帧的每个像素是投影到当前帧的 4 个像素位置上 (如图 3 所示), 因此算法要首先投影得到两倍原图大小的图像内容。在前向投影过程中, 由于大多数像素可投影到当前图像唯一对应的位置, 因此这些位置的像素基本能够得到较好的预测。由于运动矢量场的不连续性, 必然会出现多个像素同时投影到一个位置或者某些位置没有被投影到的现象, 这就是所谓的重叠和遮挡区域。图 4 给出了使用本文校正后的稠密运动矢量场投影以后产生的重叠和遮挡区域, 图中黑色区表示遮挡区域, 白色区为重叠区域。从图中可以看出, 由于遮挡和重叠区域主要处于前景 (树) 和背景 (花草和房子) 的交界处, 因此有效处理该区域的像素, 对提高错误隐

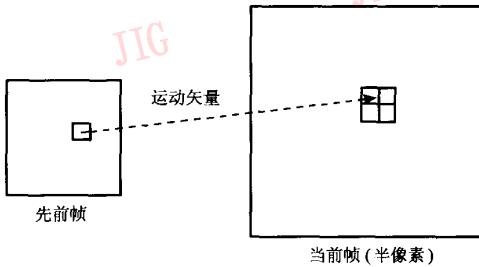


图 3 基于光流不变性的前向投影过程

Fig. 3 Forward projection procedure

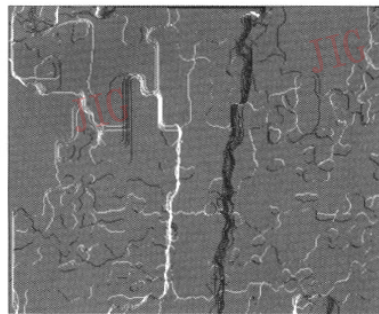
藏的效果有重要的意义。

3.1 重叠区域

由于图像中前景物体和背景的相互移动,使得前一帧图像中的处于前景物体的某些像素和处于背景的某些像素发生重叠,例如图 4 中由树的移动产生的竖条重叠区域。从重叠区域的产生原理可知,该部分的像素值应该由处于前景物体的像素投影得到。一般情况下,由于图像中大部分内容属于背景区域,因



(a) 丢失的原图像



(b) 前向投影后的重叠和遮挡区域

图 4 丢失的原图像和前向投影后的重叠和遮挡区域

Fig. 4 Original missing frame and its map of overlapped and occluded region

3.2 遮挡区域

重叠区域是由前景物体覆盖背景产生的,而遮挡区域的形成则刚好相反,即前一帧被挡住的背景内容由于物体的移动而在当前图像中变成可见区域,例如图 4(b)中树的右边形成的竖条遮挡区域。由于该区域内容在前帧图像中不可见,因此需要利用可见区域的像素对其进行预测。在先前提出的算法中,由于大多使用当前位置周围已有像素的中值或均值作为当前像素值,这种做法的缺点是处于前景物体的像素也被用来预测被覆盖的背景像素,因而不能准确地恢复图像数据。在本文算法中,是优先选取处于背景中的像素作为 7×7 均值滤波器的

此可以将图像运动矢量场的平均值 \bar{V}^{motion} 近似为背景的运动矢量。若将投影到同一位置的前一帧对应像素的运动矢量和平均运动矢量做比较,则矢量差绝对值 $|\Delta V^{motion}|$ 较大的像素被认为处于前景物体上,那么它的值即可作为当前点的像素值。从图 4 中可以发现,在前向投影过程中,运动矢量的偏差还会引起一些伪重叠区域。为确定像素是否处于物体和背景交界处的重叠区域,可采用如下的判断方法。假设 a 点和 b 点投影到同一位置,则使用下式计算:

$$\Delta V_{min}^{motion} = \min(|V_a^{motion} - \bar{V}^{motion}|, |V_b^{motion} - \bar{V}^{motion}|)$$

$$\Delta V_{max}^{motion} = \max(|V_a^{motion} - \bar{V}^{motion}|, |V_b^{motion} - \bar{V}^{motion}|)$$

当 $\Delta V_{min}^{motion} \leq T_1$, 并且 $\Delta V_{max}^{motion} \geq T_2$ 时,这时拥有较大 ΔV^{motion} 值的像素就作为当前点的像素值,否则取两点的平均值作为当前点像素值。由于在实际的投影过程中,由和背景区域有较大相对运动的物体移动形成的重叠区域才是需要特别处理的区域,因此阈值 T_2 应具有较大的值。通过大量的实验统计,取 $T_1 = 3, T_2 = 6$ 时,算法能够进行大尺寸的重叠区域的识别。

输入。当在以当前像素位置为中心的周围的已有像素中,和运动矢量均值 \bar{V}^{motion} 的差的绝对值小于等于某一阈值(该阈值可以取 3.1 节中的 T_1) 的像素超过总像素数量的三分之一时,则使用满足条件的像素值作为均值滤波器的输入;否则使用周围所有已有像素的值作为滤波器的输入。

在得到两倍原图大小的预测图像后,最后一步是使用 2×2 的均值下采样来得到原图像大小的图像内容。

4 实验结果

为了说明本文算法的有效性和稳定性,这里使用

了 3 个具有不同运动强度和空间纹理复杂度的 352×288 格式的视频序列作为测试序列进行了实验。实验时,在最新的视频压缩标准 H. 264 中来实现本文的整帧图像错误隐藏算法。H. 264 编码器生成测试码流时采用的主要编码设置有:全搜索运动检测, 32×32 的运动矢量搜索范围,量化值为 28,帧率为 30 帧/s;考虑算法实现方便,实验中只使用前向预测的编码帧,最小的块模式取 8×8 ,参考帧数目设为 1。实验过程中设置 T_1 的值为 3, T_2 的值为 $2 \times T_1$ 。

本文以恢复图像和原图的峰值信噪比 (peak signal noise ratio, PSNR) 作为评定算法好坏的基本标准,这里不考虑时间上误差积累的问题,因为信噪比越高的图像产生的误差积累越小。本文算法还与文献[6]的算法、文献[7]算法以及最基本的用前一帧图像替代的方法进行了比较,表 1 给出了用这些方法恢复的图像客观质量的比较结果。从表中可以看出,对运动强度较大的 Foreman 和 Flower 序列,前帧图像替代的方法性能非常低,本文算法较用前帧替代的方法可分别得到 3.53dB 和 8.02dB 的增益;对运动强度小的 Akiyo 序列,本文算法也得到了

0.98dB 的增益。相对于文献[6]中的算法,本文算法的增益分别是 0.18dB, 0.49dB, 0.82dB, 由于 Flower 序列图像细节太多影响了有理函数插值算法的效果,因此增益没有 Foreman 序列高。

表 1 各种算法修复图像和原图的峰值信噪比

Tab.1 PSNR comparison of error concealment algorithm

视频序列	峰值信噪比 (dB)			
	前帧替代方法	文献[6]算法	文献[7]算法	本文算法
Foreman	25.26	27.82	27.97	28.79
Flower	16.95	23.63	24.48	24.97
Akiyo	36.56	37.18	37.36	37.54

为了进行主观质量的比较,图 5 给出了原图以及用本文算法和文献[6]算法修复得到的图像效果。从图 5(b)和图 5(c)的比较可以看出,本文算法在图像亮度边缘处可取得更好的修复效果。图 5(d)、5(e)、5(f)分别给出了前一帧图像、文献[6]算法隐藏图像及本文算法隐藏图像和原图的差值图像,由图中可以看出,本文算法可以很好修复由高强度运动带来的变化。

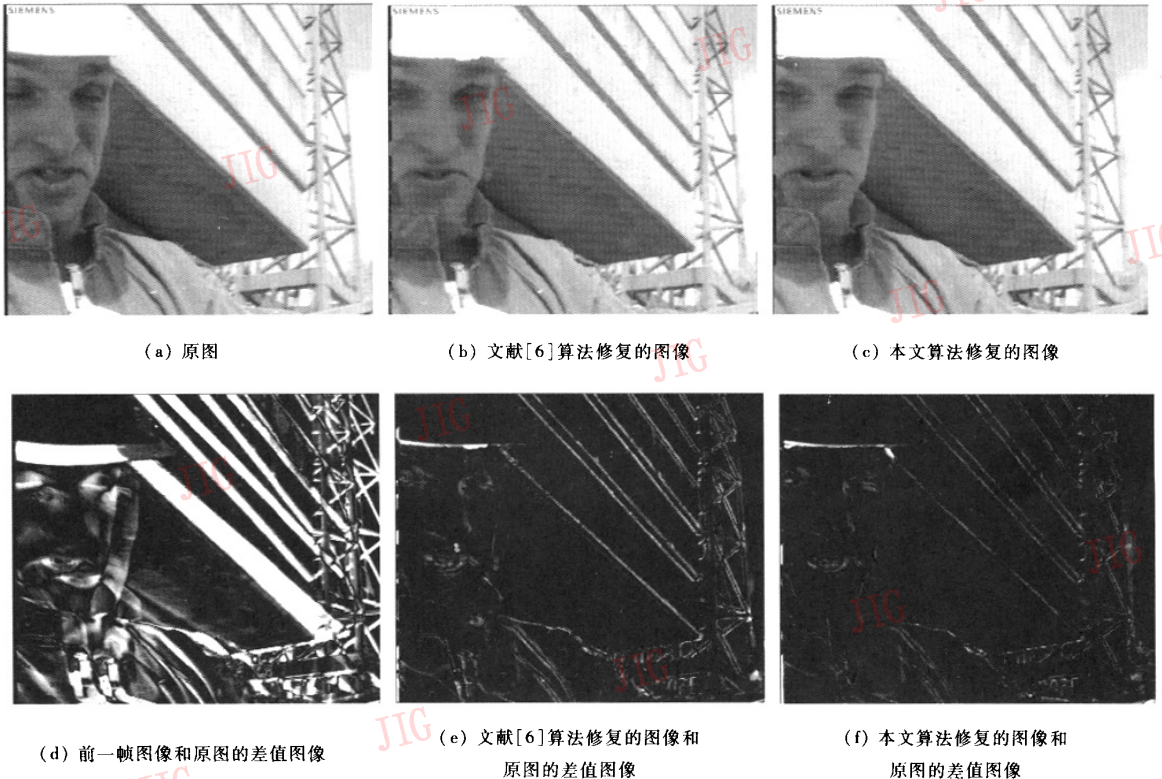


图 5 修复图像的主观图像质量及差值图像的比较

Fig.5 Comparison of visual quality of reconstructed frame

5 算法复杂度

为了在实时的网络解码器中应用,错误隐藏算法必须具有较低的运算复杂度。下面将逐步对本文算法进行复杂度分析。

第 1 步稠密运动场的生成是本文算法的核心部分,其主要包括有理插值和中值滤波两个步骤。为了降低运算复杂度,可将有理函数插值过程的式(1)和式(2)简化来得到如下公式:

$$V_C^{\text{motion}} = (\hat{w}_B \times V_A^{\text{motion}} + \hat{w}_A \times V_B^{\text{motion}}) / (\hat{w}_A + \hat{w}_B)$$

其中, $\hat{w}_A = (f_A - f_C)^2 + 1/k$, $\hat{w}_B = (f_B - f_C)^2 + 1/k$, 这样计算运动矢量的两个分量需要 12 次加法和 10 次乘法。对于中值滤波,文献[9]中的快速算法在实现 1 维滤波时,通过 6 次比较操作就可以完成水平和垂直两个方向的中值滤波。

第 2 步是前向投影过程。对于大部分一般的像素,投影过程只需要通过 2 次加法操作来进行像素的下标计算,再加上赋值操作。对于少部分处于重叠区域的像素,则需要再加数次比较操作,而对于处于遮挡区域的像素需要的运算量最大,其在采用 7×7 均值滤波时,需要 49 次的判断操作和少于 49 次的比较和加法操作。但在通常情况下,处于重叠和遮挡区域的像素只占很小的比例。

算法的最后一步是 2×2 的均值下采样,此步每个像素需要 4 次加法和一次移位操作就可得到最后的像素值。

通过如上分析可见,在通常情况下,为恢复每个像素值,本文算法需要 18 次加法、10 次乘法、6 次比较和 1 次位移操作。由此可以看出,本文算法具有较低的计算复杂度,而随着当前计算机的运算能力的提高,本文算法就可以在实际中得到应用。

6 结论

本文提出了一种用于整帧图像丢失的错误隐藏

算法。该算法首先通过有理函数插值,使运动矢量场的准确性得到有效的提高,而且在前向投影过程中,对重叠和遮挡区域的处理也提高了图像在亮度和运动场突变处的修复效果。实验结果表明,对于具有较高运动强度的视频序列,本文算法可得到 3dB 以上的增益,从而克服了由运动引起的失真。该算法已在 H. 264 解码器中实现,其可以在较低码率的有误视频传输系统中得到应用。

参考文献 (References)

- 1 Wang Yao, Zhu Qin-fan. Error control and concealment for video communication: A review [J]. Proceedings of the IEEE, 1998, 86(5): 974 ~ 991.
- 2 Sun H, Kwok W. Concealment of damaged block transform coded images using projection onto convex set [J]. IEEE Transactions on Image Processing, 1995, 4(4): 470 ~ 477.
- 3 Zeng Wen-jun, Liu Be-de. Geometric-structure-based error concealment with novel applications in block-based low-bit-rate coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 1999, 9(4): 648 ~ 665.
- 4 Li X, Orchard M. Novel Sequential Error-Concealment Techniques Using Orientation Adaptive Interpolation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(10): 857 ~ 864.
- 5 Kung W-Y, Kim Ch-S, Kuo C-C-J. A spatial-domain error concealment method with edge recovery and selective directional interpolation [A]. In: Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing [C], Hong Kong, 2003, 5: 6 ~ 10.
- 6 Belfiore S, Grangetto M, Magli E, et al. Concealment of whole-frame losses for wireless low bit-rate video based on multiframe optical flow estimation [J]. IEEE Transactions on Multimedia, 2005, 7(2): 316 ~ 329.
- 7 Baccichet P, Bagni D, Chimienti A, et al. Frame concealment for H. 264/AVC decoders [J]. IEEE Transactions on Consumer Electronics, 2005, 51(1): 227 ~ 233.
- 8 Ramponi G. The rational filter for image smoothing [J]. IEEE Signal Processing Letters, 1996, 3(3): 63 ~ 65.
- 9 Huang T S, Yang C Y, Tang G Y. A fast two dimensional median filtering algorithm [J]. IEEE Transactions on Acoustics, Speech, Signal Processing, 1979, 27(1): 13 ~ 18.