H.264标准压缩视频的超分辨率重建

黄 哲[1,2], 陈 浩[1], 张 晔[1]

(1. 哈尔滨工业大学 图像与信息技术研究所, 哈尔滨 150001, huangzhe. hit@gmail. com;
2. 北京宇航系统工程研究所,北京 100076)

摘 要: 为利用低分辨率压缩图像序列来重建高分辨率图像序列,提出一种在凸集投影(POCS)方法框架下基于整数DCT域量化噪声模型的针对H.264标准压缩视频的超分辨率重建方法.首先建立压缩视频的降质退化模型,然后根据H.264标准中的整数DCT变换和量化过程建立整数DCT域的量化噪声模型,最后在凸集投影算法的框架下给出了基于整数DCT域量化噪声的超分辨率重建算法.实验表明该算法的超分辨率重建结果的主观质量提高明显,峰值信噪比可达到30 dB,一般迭代5次即可得到良好结果,算法复杂度较低.

关键词: 超分辨率重建; 整数DCT变换; 量化噪声; H.264压缩视频

中图分类号: TN911.73 **文献标志码**: A **文章编号**: 0367-6234(2010)05-0721-06

Super-resolution reconstruction of H.264 compressed video

HUANG Zhe[1,2], CHEN Hao[1], ZHANG Ye[1]

(1. Institute of Image and Information Technology, Harbin Institute of Technology, Harbin 150001, China, huangzhe. hit@gmail. com; 2. Beijing Institute of Astronautical Systems Engineering, Beijing 100076, China)

Abstract: To reconstruct high-resolution (HR) images from a sequence of low-resolution (LR) compressed images, this paper proposes a novel algorithm focused on super-resolution reconstruction of H.264 compressed video, which is based on the integer DCT transform-domain quantization noise. Firstly, models of compressed video and integer DCT transform-domain quantization noise are surveyed. Then the reconstruction algorithm under the POCS theory is proposed. Experimental results demonstrate that this algorithm has a great improvement in subjective visual quality and low computation complexity, in which PSNR can reach 30 dB and iterations are less than 5 times.

Key words: super-resolution reconstruction; integer DCT; quantization noise; H.264 compressed video

图像超分辨率重建是指利用一系列低分辨率图像来恢复原始高分辨率图像的重建过程.基于多帧图像的超分辨率算法最早由Tsai和Huang[1]提出;Stark[2]和Petti[3]等将凸集投影(POCS)理论应用到图像超分辨率重建当中;Schultz和Stevenson[4]建立了基于最大后验概率(MAP)准则的超分辨率重建方法.这些方法在解决传统图像超分辨率重建问题时都取得良好的效果.

因原始视频的数据量十分庞大而难于存储和应用,所以压缩视频成为视频在应用中的主要形式,也形成了多种视频压缩标准,如MPEG和H.26X系列标准等.压缩过程势必引入下采样、模糊和量化噪声等降质过程.为解决这些问题,提高压缩视频的还原质量,针对压缩视频的超分辨率重建逐渐成为研究热点.初期的研究将传统的超分辨率重建方法应用于压缩视频,但因没有考虑压缩过程的降质特点而存在诸多弊端[5-6].后来涌现出很多针对压缩视频的超分辨率重建方法,如Gunturk[7]和Segall[8]等提出在贝叶斯框架下的基于最大后验概率(MAP)准则的重建方法;Altunbasak和Patti等[9-10]利用压缩过程中的量化噪声提出基于凸集投影(POCS)理论的重建方法.这些方法多数是针对以8×8 DCT变换为核心的MPEG等视频压缩标准.如今以4×4整数DCT

收稿日期: 2009-03-24.
基金项目: 黑龙江省自然科学基金资助项目(ZJG04-0701).
作者简介: 黄 哲(1983—),男,硕士研究生;
张 晔(1960—),男,教授,博士生导师.

变换为核心的 H.264 标准已逐渐成为主流,可针对 H.264 标准的超分辨率重建研究却很少.

本文充分考虑 H.264 标准的特点,首先建立视频在压缩过程中的降质模型;其次分析 H.264 标准中整数 DCT 变换和量化的特点,建立整数 DCT 域量化噪声模型;然后提出在凸集投影(POCS)理论框架下的整数 DCT 域压缩视频超分辨率重建算法;最后进行仿真实验,分析实验结果和算法性能.

1　模型的建立

1.1　压缩视频模型

研究如何利用压缩后的低分辨率序列来重建原始高分辨率序列的基础之一,就是建立高分辨率原始视频和低分辨率压缩视频的关系,即压缩视频模型.

本文将所有图像均用一维向量表示. 原始大小为 $pM \times pN$ 的高分辨率图像用 $(p^2MN) \times 1$ 的向量 $\boldsymbol{f}$ 表示,其中 p 为伸缩因子;经过下采样、模糊等降质过程后得到大小为 $M \times N$ 的低分辨率图像用 $(MN) \times 1$ 的向量 $\boldsymbol{g}$ 表示;将 $\boldsymbol{g}$ 压缩后得到大小为 $M \times N$ 的解压缩图像,用 $(MN) \times 1$ 的向量 $\hat{\boldsymbol{g}}$ 表示;$\hat{\boldsymbol{g}}$ 通过超分辨率重建算法得到大小为 $pM \times pN$ 的高分辨率重建图像,用 $(p^2MN) \times 1$ 的向量 $\hat{\boldsymbol{f}}$ 表示. $\hat{\boldsymbol{g}}$ 可由压缩码流获得,所以压缩视频超分辨率重建就是一个利用低分辨率解压缩图像 $\hat{\boldsymbol{g}}$ 来恢复高分辨率重建图像 $\hat{\boldsymbol{f}}$ 的过程.

下面根据视频压缩的过程来建立低分辨率解压缩图像 $\hat{\boldsymbol{g}}$ 和原始高分辨率图像 $\boldsymbol{f}$ 的关系. 对原始高分辨率图像的下采样、模糊等降质过程可以用 $(MN) \times (p^2MN)$ 的矩阵 $\boldsymbol{C}$ 来表示,即 $\boldsymbol{g} = \boldsymbol{Cf}$. 在 H.264 视频压缩标准中,原始视频经过运动补偿、整数 DCT 变换及量化和熵编码后最终得到压缩码流. 因为熵编码为无损编码,并不会造成解压缩图像的降质,所以本文的压缩模型忽略此过程. 设编码算法对 $\boldsymbol{g}$ 的运动估计为 $\hat{\boldsymbol{g}}^{\mathrm{MC}}$,则残差块 $\boldsymbol{R} = \boldsymbol{g} - \hat{\boldsymbol{g}}^{\mathrm{MC}}$. 在整数 DCT 变换和量化过程中,$T$ 表示正向整数 DCT 变换,用 T^* 表示反向整数 DCT 变换;Q 表示量化操作,用 Q^* 表示反量化操作. 可以得到 $\hat{\boldsymbol{g}}$ 和 $\boldsymbol{f}$ 的关系为

$$\hat{\boldsymbol{g}} = T^*(Q^*(Q(T(\boldsymbol{Cf} - \hat{\boldsymbol{g}}^{\mathrm{MC}})))) + \hat{\boldsymbol{g}}^{\mathrm{MC}}. \quad (1)$$

1.2　整数 DCT 变换和量化的噪声模型

与 MPEG 等标准采用 8×8 DCT 变换不同,H.264标准对绝大多数残差码块都采用的是基于 4×4 整数 DCT 变换. 传统的 4×4 DCT 变换为

$$\boldsymbol{Y} = \boldsymbol{AXA}^{\mathrm{T}} = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & -a & a \\ c & -b & b & -c \end{bmatrix} \boldsymbol{X} \begin{bmatrix} a & b & a & c \\ a & c & -a & -b \\ a & -c & -a & b \\ a & -b & a & -c \end{bmatrix}.$$

其中:

$$a = \frac{1}{2},\ b = \sqrt{\frac{1}{2}}\cos\left(\frac{\pi}{8}\right),\ c = \sqrt{\frac{1}{2}}\cos\left(\frac{3\pi}{8}\right).$$

将上述变换表示为

$$\boldsymbol{Y} = (\boldsymbol{C}_{\mathrm{f}}\boldsymbol{X}\boldsymbol{C}_{\mathrm{f}}^{\mathrm{T}}) \otimes \boldsymbol{E}_{\mathrm{f}} = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & d & -d & -1 \\ 1 & -1 & -1 & 1 \\ d & -1 & 1 & -d \end{bmatrix} \boldsymbol{X} \begin{bmatrix} 1 & 1 & 1 & d \\ 1 & d & -1 & -1 \\ 1 & -d & -1 & 1 \\ 1 & -1 & 1 & -d \end{bmatrix} \right) \otimes \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix}.$$

其中 $\boldsymbol{C}_{\mathrm{f}}\boldsymbol{X}\boldsymbol{C}_{\mathrm{f}}^{\mathrm{T}}$ 是二维正向变换核,$\boldsymbol{E}_{\mathrm{f}}$ 是尺度因子矩阵,符号 $\otimes$ 表示两个矩阵中的对应点相乘,$d = c/b \approx 0.41$[11].

H.264 标准为在减少整数 DCT 变换环节中的非整数运算的同时保证变换的正交性,将 b, c 和 d 的值重新分配,就得到了最终的 4×4 整数 DCT 变换,即

$$\boldsymbol{Y} = (\boldsymbol{C}_{\mathrm{f}}\boldsymbol{X}\boldsymbol{C}_{\mathrm{f}}^{\mathrm{T}}) \otimes \boldsymbol{E}_{\mathrm{f}} = \left(\begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix} \cdot \boldsymbol{X} \begin{bmatrix} 1 & 2 & 1 & 1 \\ 1 & 1 & -1 & -2 \\ 1 & -1 & -1 & 2 \\ 1 & -2 & 1 & -1 \end{bmatrix} \right) \otimes \begin{bmatrix} a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \\ a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \end{bmatrix}. \quad (2)$$

其中:$a = 0.5$, $b = \sqrt{2/5}$, $c = 0.5$.

上述整数 DCT 变换就是 H.264 编码器中采用的主要变换方式,是传统 DCT 变换的近似,在保证其正交性的同时拥有和传统 DCT 变换几乎一样的压缩性能,并避免了大量的非整数运算. 其反变换为

$$\boldsymbol{X} = \boldsymbol{C}_i^{\mathrm{T}}(\boldsymbol{Y} \otimes \boldsymbol{E}_i)\boldsymbol{C}_i =$$

$$\begin{bmatrix} 1 & 1 & 1 & 2^{-1} \\ 1 & 2^{-1} & -1 & -1 \\ 1 & -2^{-1} & -1 & 1 \\ 1 & -1 & 1 & -2^{-1} \end{bmatrix} \left(X \otimes \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix} \right) \cdot$$

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2^{-1} & -2^{-1} & -1 \\ 1 & -1 & -1 & 1 \\ 2^{-1} & -1 & 1 & -2^{-1} \end{bmatrix}. \tag{3}$$

因为该变换是正交的,所以有如下关系:

$$T^*(T(\boldsymbol{X})) = \boldsymbol{X}\quad . \tag{4}$$

H.264 标准中的量化部分采用的是尺度量化器,其基本的正向量化和反向量化分别为

$$Z_{ij} = R_{\text{round}}(Y_{ij}/Q_{\text{step}}),$$
$$Y'_{ij} = Z_{ij}Q_{\text{step}}.$$

其中 Y_{ij} 是码块经过整数 DCT 变换后在系数矩阵 Y 中(i,j) 处的系数;Y'_{ij} 是 Z_{ij} 经反向量化后得到的整数DCT系数矩阵 Y' 中(i,j) 处的系数;Q_{step} 为量化步长,由 QP 值决定,R_{round} 函数表示取整操作. 取整操作使量化后的结果丢弃了小数部分,在减少数据量的同时带来的后果是 $Y'_{ij} \neq Y_{ij}$,造成了整数 DCT 系数的失真,即信息的丢失. 这种信息的丢失,可以认为是在原有整数 DCT 系数上面加或者减去一个绝对值小于 $0.5Q_{\text{step}}$ 的噪声,因此,用引入加性量化噪声 N_{ij}^{d}的方式来表示,即:

$$Y_{ij} = Q^*(Q(Y_{ij})) + N_{ij}^{\text{d}}. \tag{5}$$

其中 N_{ij}^{d} 表示由于量化而引入到整数 DCT 域每个系数上的加性量化噪声.

由(4)和(5)得

$$T^*(Q^*(Q(T(\boldsymbol{Cf} - \hat{\boldsymbol{g}}^{\text{MC}})))) = T^*(T(\boldsymbol{Cf} - \hat{\boldsymbol{g}}^{\text{MC}}) + \boldsymbol{n}_Q^{\text{d}}) = \boldsymbol{Cf} - \hat{\boldsymbol{g}}^{\text{MC}} + \boldsymbol{n}_Q. \tag{6}$$

其中 $\boldsymbol{n}_Q^{\text{d}}$表示整数 DCT 域量化噪声的一维向量,$\boldsymbol{n}_Q$ 是表示空域量化噪声的一维向量,且 $\boldsymbol{n}_Q^{\text{d}} = T(\boldsymbol{n}_Q)$.

将(6)带入(1)得

$$\hat{\boldsymbol{g}} = \boldsymbol{Cf} + \boldsymbol{n}_Q.$$

上述推导给出了整数 DCT 域和空域的量化噪声模型,以及原始高分辨图像和解压缩端得到的低分辨率图像的关系. 可以看出压缩端的运动估计并不出现在原始高分辨图像和解压缩低分辨率图像的关系式中,因此本文将整数 DCT 域量化噪声降质模型作为后续超分辨率重建方法的基础.

将前文对码块矩阵的整数 DCT 变换和量化转换为对码块向量的操作. 如果将码块矩阵 $\boldsymbol{X}$ 用 16×1 的一维向量 $\boldsymbol{x}$ 表示,整数 DCT 变换域系数矩阵 $\boldsymbol{Y}$ 用 16×1 的一维向量 $\boldsymbol{y}$ 表示,则(2)和(3)通过矩阵变换可以表示为

$$\boldsymbol{y} = \boldsymbol{T}_{\text{DCT}}\boldsymbol{x},$$
$$\boldsymbol{x} = \boldsymbol{T}_{\text{IDCT}}\boldsymbol{y}.$$

其中 $\boldsymbol{T}_{\text{DCT}}$ 为 16×16 的整数 DCT 变换矩阵,$\boldsymbol{T}_{\text{IDCT}}$ 为 16×16 的反向整数 DCT 变换矩阵.

量化和反量化可以表示为

$$\boldsymbol{z} = R_{\text{round}}(\boldsymbol{y}/Q_{\text{step}}),$$
$$\boldsymbol{y}' = \boldsymbol{z}Q_{\text{step}}.$$

其中 $\boldsymbol{z}$ 为 16×1 的一维向量.

2　基于量化噪声模型的 POCS 超分辨率重建算法

凸集投影算法(POCS)是图像超分辨重建领域的一个经典算法. 凸集指集合中任意两点之间的点仍然在该集合中的闭集合. 假定未知信号 $\boldsymbol{f}$ 是希尔伯特空间中的某个元素,凸集 $\boldsymbol{C} = \{\bigcap_{k=1}^{K} \boldsymbol{C}_k\}$ 是 $\boldsymbol{f}$ 的有效解集,其中凸集 $\boldsymbol{C}_k(k = 1,2,\cdots,K)$ 是由 $\boldsymbol{f}$ 的先验知识确定的 k 个限制凸集,P_k 代表对应于到第 k 个凸集的投影算子,那么凸集投影算法(POCS)给出了从初始值出发得到有效解集中某个有效解的迭代算法,即

$$\boldsymbol{f}^{(i+1)} = P_K P_{K-1} \cdots P_2 P_1 \{\boldsymbol{f}^{(i)}\} = P\boldsymbol{f}^{(i)}.$$

POCS 算法中的初始值理论上可以是任意值,足够致密的有效解集中任意解都是可以接受的重建结果,但初始值和限制集的选择直接决定了算法的收敛速度和重建效果. 下面,本文就根据前面的整数 DCT 域量化噪声模型推导限制集,并给出相应的投影算法来完成超分辨率重建.

2.1　基于量化噪声模型的限制凸集

下面来建立整数 DCT 域量化噪声模型的限制凸集. 因为 R_{round} 函数表示四舍五入的取整操作,可以得出

$$Y_{ij}/Q_{\text{step}} - 0.5 < R_{\text{round}}(Y_{ij}/Q_{\text{step}}) < Y_{ij}/Q_{\text{step}} + 0.5. \tag{7}$$

将 $Y_{ij} = Q_{\text{step}} \times R_{\text{round}}(Y_{ij}/Q_{\text{step}}) + N_{ij}^{\text{d}}$代入上式得

$$-0.5Q_{\text{step}} < N_{ij}^{\text{d}} < 0.5Q_{\text{step}}.$$

根据 $\hat{\boldsymbol{g}} = \boldsymbol{Cf} + \boldsymbol{n}_Q$ 和 $\boldsymbol{n}_Q^{\text{d}} = T(\boldsymbol{n}_Q)$ 可以得到 $\boldsymbol{n}_Q^{\text{d}} = T(\hat{\boldsymbol{g}} - \boldsymbol{Cf})$. $\boldsymbol{n}_Q^{\text{d}}$ 是由 N_{ij}^{d}所构成的一维向量,所以 $\boldsymbol{n}_Q^{\text{d}}$中的每一个元素都应该在区间$(-0.5Q_{\text{step}}, 0.5Q_{\text{step}})$当中.

对于高分辨率原始图像 $\boldsymbol{f}$ 经过下采样后的低分辨率图像 $\boldsymbol{g} = \boldsymbol{Cf}$ 和它的低分辨率解压缩图像 $\hat{\boldsymbol{g}}$ 来说,它们对应的 4×4 的码块都应该满足

$$-0.5Q_{\text{step}} < T_{\text{DCT}}(\boldsymbol{g}(i,j) - \hat{\boldsymbol{g}}(i,j)) < 0.5Q_{\text{step}}.$$

即 $\boldsymbol{g}$ 和 $\hat{\boldsymbol{g}}$ 在对应4×4码块上的差值经过整数DCT

变换后每个变换域系数的值都在区间 $(-0.5Q_{\text{step}}, 0.5Q_{\text{step}})$ 当中. 式中 (i,j) 表示 4×4 码块在图像中的位置.

这样就可以构造一个如下的集合：

$\hat{\boldsymbol{f}}_k \in \{\hat{\boldsymbol{f}}_k: -0.5Q_{\text{step}} < T(\boldsymbol{C}\hat{\boldsymbol{f}}_k - \hat{\boldsymbol{g}}) < 0.5Q_{\text{step}}\}$.

其中 $\hat{\boldsymbol{f}}_k$ 代表高分辨率原始图像，$\boldsymbol{C}\hat{\boldsymbol{f}}_k$ 则代表经下采样后的低分辨率原始图像，$\hat{\boldsymbol{g}}$ 代表低分辨率解压缩图像，T 代表对 $\boldsymbol{C}\hat{\boldsymbol{f}}_k - \hat{\boldsymbol{g}}$ 中的每个 4×4 的码块做整数 DCT 变换. 该集合中的 $\hat{\boldsymbol{f}}_k$ 在经过下采样后与 $\hat{\boldsymbol{g}}$ 的残差图像的每个 4×4 码块，经过整数 DCT 变换之后的系数都在区间 $(-0.5Q_{\text{step}}, 0.5Q_{\text{step}})$ 当中. 这就是本文建立的整数 DCT 变换域量化噪声的限制集.

2.2 量化噪声模型的 POCS 算法

在文献[9-10]，[12]和[13]中提出的针对 MPEG 视频压缩标准中 DCT 变换及量化的凸集投影算法的基础上，利用前文建立的针对 H.264 标准的视频压缩模型和整数 DCT 域量化噪声模型，提出在凸集投影算法框架下的基于整数 DCT 域量化噪声限制集的压缩视频超分辨率重建算法，表达式如下：

$$\hat{\boldsymbol{f}}_{k+1} = P[\hat{\boldsymbol{f}}_k] = \hat{\boldsymbol{f}}_k - \boldsymbol{C}^{\mathrm{T}}\Delta_k. \tag{8}$$

其中：$\hat{\boldsymbol{f}}_k$ 代表第 k 次投影之后的超分辨率重建结果，是一个 $(p^2MN)\times1$ 的向量，$\boldsymbol{C}^{\mathrm{T}}$ 是伸缩矩阵 $\boldsymbol{C}$ 的转置，P 代表投影操作. Δ_k 是一个 $(MN)\times1$ 的空域修正向量，它的每个系数都与低分辨率图像的差值 $[\boldsymbol{C}\hat{\boldsymbol{f}}_k] - \hat{\boldsymbol{g}}$ 中的值一一对应. 将 Δ_k 对应 $[\boldsymbol{C}\hat{\boldsymbol{f}}_k] - \hat{\boldsymbol{g}}$ 中位置为 (m,n) 码块的 16 个系数组成一个 16×1 的一维向量，表示为 $\Delta_k(m,n)$. 这样求解 Δ_k 的问题就转化为求解所有码块对应的 $\Delta_k(m,n)$ 的问题.

令 $D_k(m,n) = T_{\text{DCT}}([\boldsymbol{C}\hat{\boldsymbol{f}}_k](m,n) - \hat{\boldsymbol{g}}(m,n))$，则有

$$\Delta_k^D(m,n;i) = \begin{cases} \boldsymbol{D}_k(m,n;i) - 0.5Q_{\text{step}}, & \boldsymbol{D}_k(m,n;i) > 0.5Q_{\text{step}}; \\ \boldsymbol{D}_k(m,n;i) + 0.5Q_{\text{step}}, & \boldsymbol{D}_k(m,n;i) < -0.5Q_{\text{step}}; \\ \boldsymbol{D}_k(m,n;i), & \text{others}. \end{cases}$$

式中，$\hat{\boldsymbol{g}}(m,n)$ 为一个 16×1 的一维向量，表示解压缩重建图像中位置为 (m,n) 的码块；$[\boldsymbol{C}\hat{\boldsymbol{f}}_k](m,n)$ 为一个 16×1 的一维向量，表示高分辨率图像 $\hat{\boldsymbol{f}}_k$ 经过下采样后得到的低分辨率图像 $\boldsymbol{C}\hat{\boldsymbol{f}}$ 中位置为 (m,n) 的码块；$\boldsymbol{T}_{\text{DCT}}$ 代表整数 DCT 变换矩阵；Q_{step} 为 H.264 标准中的量化步长；$\boldsymbol{D}_k(m,n;i)$ 是 $\boldsymbol{D}_k(m,n)$ 中的第 i 个系数 $(1\leqslant i\leqslant16)$. $\Delta_k^D(m,n)$ 为对应 $\boldsymbol{D}_k(m,n)$ 的 16×1 的整数 DCT 域修正向量，$\Delta_k^D(m,n;i)$ 表示 $\Delta_k^D(m,n)$ 中的第 i 个修正值 $(1\leqslant i\leqslant16)$，$\Delta_k(m,n)$ 是 $\Delta_k^D(m,n)$ 做反向整数 DCT 变换后经过修正得到的，$\boldsymbol{C}_{m,n}^{\mathrm{T}}$ 表示低分辨率图像中位置为 (m,n) 的码块与原始图像之间的伸缩矩阵. $\Delta_k(m,n)$ 的计算公式如下：

$$\Delta_k(m,n) = \frac{\boldsymbol{T}_{\text{IDCT}}(\Delta_k^D(m,n))}{\|\boldsymbol{C}_{m,n}^{\mathrm{T}}\boldsymbol{T}_{\text{IDCT}}\|^2}.$$

当每个码块对应的 $\Delta_k(m,n)$ 都已计算出来后就可得到 Δ_k. 在计算 $\Delta_k^D(m,n)$ 时，如果每一个 $\boldsymbol{D}_k(m,n;i)(1\leqslant i\leqslant16)$ 都在区间 $(-0.5Q_{\text{step}}, 0.5Q_{\text{step}})$ 时，则放弃对这个码块的修正，保持其原始空域的值不变. 通过上面的过程将 Δ_k 计算出后，就可以通过式(8) 迭代计算出 $\hat{\boldsymbol{f}}_{k+1}$. 当所有的 $\boldsymbol{D}_k(m,n;i)$ 都在区间 $(-0.5Q_{\text{step}}, 0.5Q_{\text{step}})$ 时就停止迭代运算，此时的 $\hat{\boldsymbol{f}}_k$ 就是最后的超分辨率重建结果. 迭代运算的起始值 $\hat{\boldsymbol{f}}_0$ 理论上可以设为空间中的任何一个向量，此处将低分辨率解压缩的图像的双线性插值图像作为 $\hat{\boldsymbol{f}}_0$，因为此方法一般得不到收敛的解，所以本算法认为经过若干次迭代之后，若 $\hat{\boldsymbol{f}}_k$ 的质量不再明显变化时就停止迭代，此时的 $\hat{\boldsymbol{f}}_k$ 就是超分辨率重建结果.

3 实验结果及分析

用实验来验证上述整数 DCT 域超分辨率重建方法的性能. 首先，将 CIF (352×288) 大小的 foreman 序列中的一帧作为高分辨率原始图像，对其下采样得到 4 幅 QCIF (176×144) 大小的图像，并构成一个低分辨率序列. 然后用 JM8.6 软件按照 H.264 标准的 baseline 档次，QP 值设为 28，对上述低分辨率图像序列进行压缩，其中第 1 帧作为 I 帧，其他 3 帧为 P 帧，压缩比为 27 倍. 解压缩后得到 4 帧低分辨率序列，对其中的一帧进行双线性插值得到一幅 CIF 大小的图像 f_0. 以 f_0 作为起始点利用上述算法进行迭代得到最终的重建结果，如图 1 所示，细节对比如图 2，对 mobile 序列与上述相同的实验，结果如图 3 所示.

上面的实验中重建结果的峰值信噪比 (R_{PSN}) 相比双线性插值算法有很大的提高，可以达到 30 dB 以上. 同样，重建结果相比双线性插值算法在信噪比(SNR)方面具有明显的优势，提升均在 5 dB 以上. 同样可以看出视觉效果上细节信息更加丰富. 同时本文算法通过较少迭代次数可得到稳定的重建结果，如图 4 所示，两组实验均迭代 5 次即可得到稳定的重建结果. 对 foreman 序列在不同 QP 值下用 JM8.6 软件压缩并用本文算法进行重建，其结果相对于双线性插值算法重建结

果的 R_{PSN} 增量，如图 5 所示. 而 SNR 的增量也有同样的趋势. 因此得出本文算法在 QP 较小的情况下，具有更优的重建效果. 采用分块操作尽可能化简了运算中矩阵和向量的大小，节省了计算时间和所需内存.

（a）原始高分辨率图像　（b）双线性插值图像　（c）本文算法结果

图 1　foreman 序列实验结果

（a）原始细节　（b）插值算法细节　（c）本文算法细节

图 2　细节对比

（a）原始高分辨率图像　（b）双线性插值图像　（c）本文算法结果

图 3　Mobile 序列实验结果

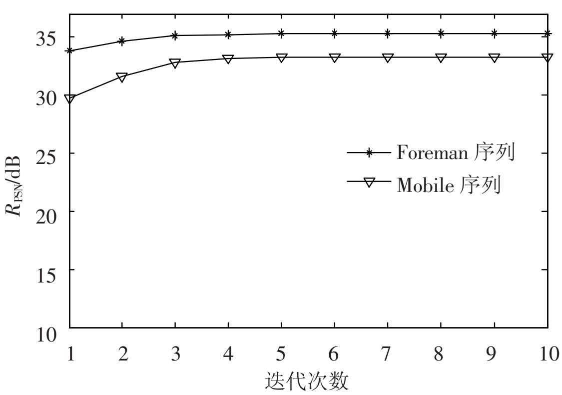

图 4　R_{PSN} 随迭代次数变化曲线

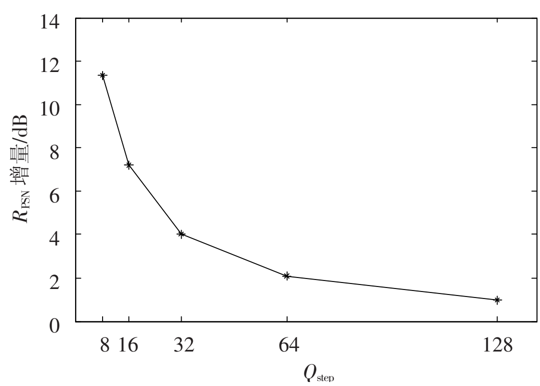

图 5　R_{PSN} 增量与 Q_{step} 的关系

综上所述，本文提出的基于整数 DCT 域量化噪声的压缩视频超分辨率重建方法在峰值信噪比和信噪比方面相比传统的插值算法有较大的提高，而且运算复杂度低，该算法中所使用的大量整数运算使其具有良好的优化前景.

4　结论与展望

本文提出的基于整数 DCT 域量化噪声的压缩视频超分辨率重建方法，针对 H. 264 标准的压缩视频拥有较好的主观重建效果，峰值信噪比较双线性插值等方法有较大提高，收敛速度较快，对计算机资源消耗较小. 但本文的算法仍有很多局限性，例如只考虑量化噪声这一个压缩视频降质因素，并没有考虑由于预测和运动补偿等原因引起的降质过程；算法需要已知降质矩阵 $\boldsymbol{C}$，而通常这个矩阵也具有不确定性. 在未来针对 H. 264 标准的压缩视频超分辨研究当中，可以进一步引入运动补偿等降质模型；同时也可以根据整数 DCT 域量化噪声的统计模型建立在贝叶斯框架下的压缩视频超分辨算法.

参考文献:

[1] TSAI R, HUANG T. Advances in Computer Vision and Image Processing[M]. Greenwich: JAI Press Inc, 1984:317 – 339.

[2] STARK H, OSKOUI P. High-resolution image recovery from image-plane arrays, using convex projections[J]. Journal of the Opt Soc of America, 1989, 6 (11): 1715 – 1726.

[3] PATTI A J, SEZAN M I, TEKALP A M. Super-resolution video reconstruction with arbitrary sampling lattices and nonzero aperture time[J]. IEEE Trans on Image Processing, 1997, 6(8): 1064 – 1076.

[4] SCHULTZ R R, STEVENSON R L. Extraction of high resolution frames from video sequences [J]. IEEE Trans on Image Processing, 1996 (6): 996 – 1001.

[5] SEGALL C A, MOLINA R, KATSAGGELOS A K. High-resolution images from low-resolution compressed video[J]. IEEE Signal Processing Magazine, 2003, 20 (3): 37 – 48.

[6] 徐忠强, 朱秀昌. 压缩视频超分辨率重建技术[J]. 电子与信息学报, 2007, 29(2): 499 – 505.

[7] GUNTURK B K, ANTUNBASAK Y, MERSEREAU R. Bayesian resolution-enhancement framework for transform-coded video[C]//IEEE Int. Conf. Image Processing. Thessaloniki, Greece:[s. n.], 2001: 444.

[8] SEGALL C A, KATSAGGELOS A K, MOLINA R, *et al*. Bayesian resolution enhancement of compressed video[J]. IEEE Trans on Image Processing, 2004, 13 (7): 898 – 911.

[9] PATTI A J, ALTUNBASAK Y. Super-Resolution image estimation for transform coded video with application to MPEG[C]//IEEE International Conference on Image Processing. Kobe, Japan:[s. n.],1999: 179 – 183.

[10] ALTUNBASAK Y, PATTI A J, MERSEREAU R M. Super-resolution still and video reconstruction from MPEG coded video[J]. IEEE Trans. on Circuits and System for Video Technology, 2002, 12(4): 217 – 226.

[11] RICHARDSON IAIN E G. H.264 and MPEG – 4 Video Compression[M]. England: John Wiley and Sons Inc, 2003:189 – 190.

[12] CHAUDHURI S. Super-Resolution Imaging[M]//SEGALL C A, KATSAGGELOS A K, MOLINA R. The International Series in Engineering and Computer Science. Norwell: Kluwer, 2001: 211 – 242.

[13] 胡辛. 压缩图像超分辨重建算法研究[D]. 西安:西安电子科技大学,2006: 37 – 38.

(编辑 张 宏)

(上接第 720 页)

[3] CHEN Y S, LANGFORD W F. The subharmonic bifurcation solution of nonlinear Mathieu's equation and Euler dynamic bulking problems[J]. Acta Mechanica Sinica, 1988, 4(4): 350 – 362.

[4] GOLUBITSKY M, SCHAEFFER D G, STEWART I. Singularities and Groups in Bifurcation Theory[M]. New York: Spring-Verlag, 1988.

[5] 朱照宣. 非线性动力学中的浑沌[J]. 力学进展, 1984, 14(2):129 – 146.

[6] 胡海昌. 多自由度结构固有振动理论[M]. 北京:科学出版社,1987.

[7] WILLIAM B B. An improved computational technique for perturbations of the generalized symmetric linear algebraic eigenvalue problem[J]. International Journal for Numerical Methods in Engineering, 1987, 24(3): 529 – 541.

[8] 陈塑寰, 徐涛, 韩万芝. 线性振动亏损系统的矩阵摄动理论[J]. 力学学报, 1992, 24(6): 747 – 752.

[9] 陈塑寰, 刘中生,赵又群. 振型一阶导数的高精度截尾模态展开法[J]. 力学学报, 1993, 25(4): 427 – 434.

[10] LIU Z S, CHEN S H, ZHAO Y Q. An accurate method for computing eigenvector derivatives for free-free structures[J]. Int. Journal of Computers and Structures, 1994, 52(6): 1135 – 1143.

[11] 韩万芝, 宋大同, 陈塑寰. 计算特征向量摄动量的混合基展开法[J]. 固体力学学报,1995, 16(3): 237 – 243.

[12] 陈塑寰. 结构动态设计的矩阵摄动理论[M]. 北京:科学出版社, 1999.

[13] SEYRANIAN A P, MAILYBAEV A A. Multiparameter Stability Theory with Mechanical Application[M]. Singapore: World Scientific, 2003.

(编辑 张 红)