

文章编号:1004-9037(2013)06-0000-09

# 基于多重假设的视频压缩感知分层重建

常 侃 覃团发 唐振华

(广西大学计算机与电子信息学院, 南宁, 530004)

**摘要:**为了改进视频压缩感知方案的性能,提出了一种基于多重假设的视频压缩感知分层重建方案。该重建方案以图像组为单位进行,首先对关键帧进行空域多重假设预测和残差重建;接下来对图像组中的每个非关键帧分配重建层级,并按照由低至高的顺序逐层重建;每个非关键帧的重建过程逐块进行,需要其参考帧中的时域假设预测集合及当前帧中的空域假设预测集合来对每个图像块做混合多重假设预测,并通过求解全变分最小化问题进行残差重建。实验结果表明,在相同采样率下,与已有的视频压缩感知方案相比,本文方案可以获得质量更高的重建图像。

**关键词:**压缩感知;多重假设;分层重建;全变分最小化

**中图分类号:** TN919.8

**文献标志码:** A

## Multi-hypothesis Based Hierarchical Reconstruction for Compressed Video Sensing

Chang Kan, Qin Tuanfa, Tang Zhenhua

(School of Computer and Electronic Information, Guangxi University, Nanning, 530004, China)

**Abstract:** To improve the performance of Compressed video sensing (CVS), a multi-hypothesis based hierarchical reconstruction method is proposed. In the presented framework, the key frame in Group of picture (GOP) is first predicted by the spatial Multi-hypothesis (MH) method and then reconstructed by the residual reconstruction method. Afterwards, the reconstruction levels are allocated to all the non-key frames in current GOP, after which the reconstruction is processed from the lowest level to the highest one. When reconstructing a non-key frame, block by block reconstruction is carried out. For a target block, both of the temporal data set in reference frames and spatial data set in current frame are used to form a MH prediction, after that the total variation minimization problem is solved to execute residual reconstruction. Experimental results show that compared with the existing CVS methods, our proposed one can get higher quality of reconstructed images at the same sampling rates.

**Key words:** compressed sensing; multi-hypothesis; hierarchical reconstruction; total variation minimization

## 引 言

压缩感知(Compressed sensing, CS)<sup>[1]</sup>是近年在信号处理领域得到广泛重视的新兴理论体系。该理论指出,若某个待采样信号是稀疏的(或者在某个变换域是稀疏的),那么从其有限的(远低于原

始信号维数)不相关测量值中准确恢复出原始信号的可能性极高。CS理论的提出,极大地丰富和发展了信号获取理论,并为其他相关领域的研究提供了新的技术和新的思路<sup>[2-3]</sup>。因为CS的测量过程运算量较低,而重建过程则需要迭代求解复杂度相对较高的最优化问题,所以其理论结构特别适合计算能力和电力受限的各类分布式应用,例如无线多

基金项目:国家自然科学基金(61261023)资助项目;广西自然科学基金(2011GXNSFD018024)资助项目;广西教育厅科研项目(201203YB001)资助项目。

收稿日期:2012-11-19;修订日期:2013-01-20

请加一项 广西自然科学基金  
(2013GXNSFBA019272)

媒体传感器网络、无线视频监控等。

已经有一些学者将 CS 理论应用于视频编码方案。例如, Kang 等人利用了非关键帧与边信息间的相关性进行重建<sup>[4]</sup>, 但是, 时域相关性并未得到充分利用。Do 等人采用已重建帧中的时域相邻块表示当前重建帧中的图像块, 以提高边信息的准确性<sup>[5]</sup>, 并首次采用了残差重建。但是, Do 等人的方法假设了系数的稀疏性, 该假设条件没有普适性。Mun 和 Fowler 设计了以关键帧为参考, 基于运动估/运动补偿 (Motion estimation/ motion compensation, ME/ MC) 的残差重建方案<sup>[6]</sup>; 在文献[6]的基础上, Tramel 和 Fowler 进一步提出了基于多重假设 (Multi-hypothesis, MH) 的残差重建方案<sup>[7]</sup>, 获得了较高质量的重建图像。但是, 在文献[7]的方法中, 没有考虑参考帧距离远近对于重建质量的影响; 在文献[6]的方法中, 也仅考虑了采用相邻关键帧对当前图像组 (Group of pictures, GOP) 中的非关键帧进行双向预测的重建结构, 当 GOP 增大时, 重建质量会有明显下降。文献[8]提出使用变采样率的方式以提升 MH 残差重建方案的性能, 但是其方案仅考虑了一个关键帧与一个非关键帧的简单情况, 仍不能解决大 GOP 下由于关键帧与非关键帧之间相关性下降而导致的重建质量降低的问题。

为了解决文献[6]和文献[7]方案中的缺陷, 本文提出了一种基于 MH 预测的视频压缩感知分层重建方案。在发送端区分关键帧与非关键帧, 关键帧的采样率高于非关键帧的采样率。在接收端, 对关键帧仅利用空域相关性进行 MH 预测和残差重建。对非关键帧, 为其分配相应的重建层级, 按照层级顺序由下至上逐层重建。在每个非关键帧的重建过程中, 高层待重建帧以低层已重建帧为时域参考, 并结合本帧图像的空域相关性进行混合 MH 预测; 对 MH 预测的结果计算测量残差, 并求解全变分 (Total variation, TV) 最小化问题以重建残差信号。

## 1 视频序列的测量

在发送端, 假设视频数据具有  $N$  帧图像, 每帧图像大小为  $W \times H$ 。较为直接的测量方式是把视频数据排列为具有  $W \times H \times N$  行的列向量, 并进行全局测量。但是, 这种测量方法对计算能力和内存的需求比较高, 并且测量延时较大, 并不适合分布式应用。因此, 本文采用运算量较低的逐帧测

量。

为进一步降低对测量端内存和计算能力的需求, 对每帧图像, 采用 Gan 所提出的分块压缩感知 (Block compressed sensing, BCS) 方法<sup>[9]</sup>。具体地, 将每帧分成  $N_B$  个互不重叠的  $B \times B$  大小的图像块, 逐块采用相同的测量矩阵  $\Phi$  进行测量。此时, 对于一帧图像而言, 相当于进行如下操作

$$\mathbf{y} = \Phi \mathbf{x} \quad (1)$$

其中,  $\mathbf{x}$  是  $B^2 \times N_B$  维矩阵, 矩阵的每一列代表按照一维列向量的形式重排的一个图像块;  $\Phi$  是  $m \times B^2$  维测量矩阵, 在本文中采用随机高斯矩阵,  $m/B^2$  是采样率;  $\mathbf{y}$  是  $m \times N_B$  维矩阵, 矩阵中的每一列代表单个图像块对应的测量值。

本文将视频序列划分为若干 GOP, 每个 GOP 中包含一个关键帧和若干非关键帧, 关键帧使用比非关键帧更高的采样率。在接收端, 关键帧仅利用空域相关性进行重建; 非关键帧仅在当前 GOP 内选择时域参考帧, 可综合利用时域和空域相关性进行重建。划分 GOP 的好处在于, 可以有效防止在传输过程中由于误码或丢包所引起的差错传播问题。

## 2 基于多重假设的分层重建

### 2.1 关键帧的空域 MH 预测

关键帧的重建过程不依赖于任何已重建帧。首先, 对关键帧逐块进行独立重建。在本文中, 独立重建图像块等效于求解如下 TV 范数最小化问题

$$x_i = \underset{x_i}{\operatorname{argmin}} x_i \text{TV s. t. } \mathbf{y}_i = \Phi \mathbf{x}_i \quad (2)$$

其中,  $x_i$  代表第  $i$  个图像块的原始数据;  $\mathbf{y}_i$  代表第  $i$  个图像块的测量值。

由于 TV 范数有效表征了图像信号特点, Candès 等人将其引入 CS 理论以重建图像信号。经典的 TV 范数最小化问题求解方法是采用内点法迭代求解 log-barrier 问题, 但是需要较高的复杂度。Li 将 TV 最小化问题转化为 Augmented Lagrangian 问题的形式, 并将该问题拆成两个子问题交替求解<sup>[10]</sup>, 大大简化了求解过程, 该方法被命名为 (TV minimization by augmented lagrangian and alternating direction algorithms, TVAL3)。在本文中, 所有 TV 范数最小化问题均采用 TVAL3 算法进行求解。alternating 无 “-”

在独立重建关键帧之后, 对每个图像块进行空域 MH 预测。基于 Tikhonov 正规化的 MH 预测

中文名: 基于增广拉格朗日和交替方向的全变差最小化算法

由 Tramel 等人首次提出<sup>[7]</sup>,具体地,当前图像块的 MH 预测值为如下问题的解

$$\tilde{\mathbf{x}}_i = \operatorname{argmin}_{\tilde{\mathbf{x}}_i} \|\mathbf{y}_i - \Phi \tilde{\mathbf{x}}_i\|_2^2 \quad (3)$$

由于待预测图像块的测量值  $\mathbf{y}_i$  已知,直接在测量域的预测保证了预测结果的准确性。

在 Tramel 的方法<sup>[7]</sup>中,仅考虑了针对非参考帧的时域 MH 预测方法,未考虑关键帧的 MH 预测方法。本文对关键帧进行空域 MH 预测:在以待预测图像块(假设为第  $i$  个图像块)为中心的搜索窗内,通过逐像素滑动  $B \times B$  大小的掩膜产生空域假设集合,即,以第  $i$  个图像块为中心的搜索窗内的所有  $B \times B$  大小的图像块均视为假设预测块。所有假设预测块的最佳线性组合是空域 MH 预测结果,即

$$\tilde{\mathbf{x}}_i = \mathbf{H}_i^S \omega \quad (4)$$

式中: $\omega$  是最优线性组合系数; $\mathbf{H}_i^S$  是  $B^2 \times K$  维矩阵,代表了第  $i$  个图像块的空域假设预测集合,其每一列是以列向量形式排列的一个假设预测块; $K$  为假设预测块的数量,在空域 MH 预测中,若搜索窗大小为  $\pm \omega$ ,则  $K = (2\omega + 1)^2$ 。

将式(4)代入问题(3)并加入惩罚项,MH 预测问题等效于求解最优线性组合  $\omega$  的问题

$$\omega = \operatorname{argmin}_{\omega} \|\mathbf{y}_i - \Phi \mathbf{H}_i^S \omega\|_2^2 + \lambda \|\Gamma \omega\|_2^2 \quad (5)$$

式中: $\lambda$  是拉格朗日参数; $\Gamma$  是 Tikhonov 矩阵,其为对角矩阵形式

$$\Gamma = \begin{pmatrix} \|\mathbf{y}_i - \Phi \mathbf{h}_1\|_2^2 & \cdots & 0 \\ \cdots & \ddots & \cdots \\ 0 & \cdots & \mathbf{y}_i - \Phi \mathbf{h}_{K/2} \end{pmatrix} \quad (6)$$

其中, $\mathbf{h}_j$  是集合  $\mathbf{H}_i^S$  中的第  $j$  个假设预测块, $j = 1, \dots, K$ 。因为问题(5)的目标函数可微,所以通过对目标函数微分并令微分结果为 0,求出其闭式解如下

$$\omega = ((\Phi \mathbf{H}_i^S)^T (\Phi \mathbf{H}_i^S) + \lambda^2 \Gamma^T \Gamma)^{-1} (\Phi \mathbf{H}_i^S)^T \mathbf{y}_i \quad (7)$$

## 2.2 残差重建方法及性能分析

视频压缩感知的残差重建方法最早由 Do 等人提出<sup>[5]</sup>。其基本思想是待重建信号进行预测,若预测准确,残差信号比原始信号在变换域下更稀疏,因此残差信号的重建误差以极大概率小于原始信号的重建误差。以本文方法为例,假设有待重建图像块的 MH 预测值  $\tilde{\mathbf{x}}_i$ ,对其进行测量

$$\tilde{\mathbf{y}}_i = \Phi \tilde{\mathbf{x}}_i \quad (8)$$

其次,求真实测量值与预测测量值的残差

$$\mathbf{y}_i^r = \mathbf{y}_i - \tilde{\mathbf{y}}_i = \Phi(\mathbf{x}_i - \tilde{\mathbf{x}}_i) = \Phi \mathbf{x}_i^r \quad (9)$$

假设  $\hat{\mathbf{x}}_i^r$  为从  $\mathbf{y}_i^r$  中恢复出来的残差信号,则可

以通过下式获取第  $i$  个图像块的重建值

$$\hat{\mathbf{x}}_i = \hat{\mathbf{x}}_i^r + \tilde{\mathbf{x}}_i \quad (10)$$

根据上述流程,第  $i$  个图像块的重建误差为

$$\mathbf{x}_i - \hat{\mathbf{x}}_{i2} = (\tilde{\mathbf{x}}_i + \mathbf{x}_i^r) - (\tilde{\mathbf{x}}_i + \hat{\mathbf{x}}_i^r)_2 = \mathbf{x}_i^r - \hat{\mathbf{x}}_i^r \quad (11)$$

根据式(11)可知,在残差重建算法中,第  $i$  个图像块的重建误差由预测残差  $\mathbf{x}_i^r$  直接决定。另一方面,若第  $i$  个图像块的预测值足够准确,或者说  $\tilde{\mathbf{x}}_i$  与  $\mathbf{x}_i$  足够接近,则预测残差信号  $\mathbf{x}_i^r$  定会比  $\mathbf{x}_i$  更为稀疏。因此,根据 CS 基本理论,重建  $\mathbf{x}_i^r$  会比重建  $\mathbf{x}_i$  误差更小。

因为 MH 预测的结果满足问题(3),所以第  $i$  个图像块的预测值足够准确,采用残差重建方法预期可获得令人满意的重建质量。本文将残差重建问题也视为 TV 范数最小化问题,即

$$\hat{\mathbf{x}}_i^r = \operatorname{argmin}_{\mathbf{x}_i^r} \|\mathbf{x}_i^r\|_{\text{TV}} \text{ s. t. } \mathbf{y}_i - \Phi \tilde{\mathbf{x}}_i = \Phi \mathbf{x}_i^r \quad (12)$$

公式错误: argmin 中的 x 应在范数内,即 ||x||; TV 是下标; 应空两格才到 s.t.

## 2.3 非关键帧的时域分层重建

在 Tramel 的方法<sup>[7]</sup>中,对非关键帧采用时域 MH 预测。假设对非关键帧中第  $i$  个图像块进行时域 MH 预测,在以关键帧中对应位置处为中心的搜索窗内,通过逐像素滑动  $B \times B$  大小的掩膜产生时域假设集合  $\mathbf{H}_i^T$ 。但是,Tramel 的方法未考虑 GOP 的大小对非关键帧重建质量的影响。事实上,GOP 越大,关键帧与非关键帧之间的距离越远,时域相关性越低,MH 预测效果也越差。另外,为保证 MH 预测准确性,Tramel 的方法给关键帧设置了过高的采样率,这会造成较大的传输负担和速率波动,并不符合实际。

为解决上述问题,借鉴传统视频编码方案中(例如 H. 264/AVC)编码端的层级 B 帧编码结构<sup>[11]</sup>,本文在视频压缩感知系统的重建端采用了时域分层重建,以进一步提升重建图像质量。

在时域分层重建结构中,假设当前重建第  $n$  个 GOP,长度为  $L$ ,需要为所有非关键帧分配相应的时域重建层。该 GOP 共包括  $\log_2 L$  个时域层,第  $k$  个时域层包括  $2^{k-1}$  个非关键帧。在该 GOP 中,将奇数帧分配为第  $\log_2 L$  个时域层;将第  $L/2$  帧分配为第 1 个时域层;其余帧按照金字塔结构分配相应的时域层。在重建 GOP 的过程中,由下层至上层逐层进行重建。

在重建每个非关键帧的过程中,需要选取时域上已重建的前后 2 个最邻近帧作为参考帧以进行 MH 预测,也即

$$\mathbf{H}_i^T = [\mathbf{H}_i^{TF}, \mathbf{H}_i^{TB}] \quad (13)$$

其中,  $\mathbf{H}_i^{TF}$  代表第  $i$  个图像块的前向假设预测集合,  $\mathbf{H}_i^{TB}$  是其反向假设预测集合。将  $\mathbf{H}_i^T$  代替式(4)~(7)中的  $\mathbf{H}_i^S$  即可计算时域 MH 预测结果。需要注意的是,与关键帧的空域 MH 预测情况不同,由于时域分层重建结构中的参考帧已经重建,假设预测集合可直接获得,无须预先独立重建当前非关键帧。

图 1 给出了  $L=8$  时的分层重建结构示例。从图中可以看出,每个非关键帧的参考帧并不限制为关键帧;在金字塔分层重建结构下,待重建帧与其参考帧之间的距离较近,因此时域相关性较强。可以推断出,在分层重建结构下进行 MH 预测比 Tramel 的方法更为准确;综合 2.2 节对残差重建的性能分析可知,时域分层重建所获得的图像重建质量更高。

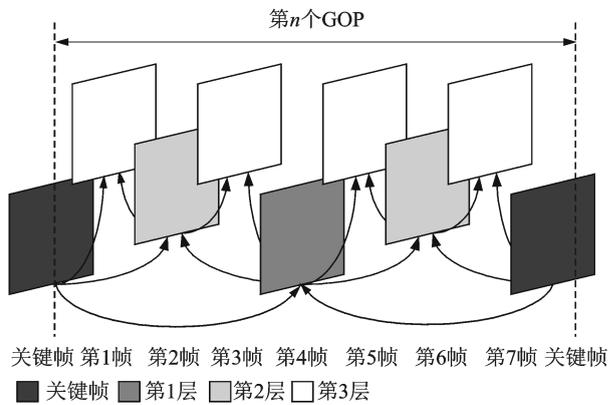


图 1 时域分层重建结构 (GOP=8)

## 2.4 非关键帧的空域时域混合 MH 预测

视频序列通常具备较强的时域相关性和空域相关性。当序列仅包含缓慢、简单的运动时,采用时域 MH 预测可获得令人满意的预测结果;但是,当视频序列包含剧烈、复杂运动时,若仅采用时域 MH 预测,预测效果不能令人满意。为解决此问题,本文进一步提出空域时域混合 MH 预测方法。

空域时域混合 MH 预测仍在时域分层重建结构的基础上逐块进行。用于混合 MH 预测的假设集合包括空域预测集合与前向、后向时域预测集合,即

$$\mathbf{H}_i^M = [\mathbf{H}_i^S, \mathbf{H}_i^{TF}, \mathbf{H}_i^{TB}] \quad (14)$$

混合假设预测集合  $\mathbf{H}_i^M$  的生成过程如图 2 所示。将  $\mathbf{H}_i^M$  代替式(4)~(7)中的  $\mathbf{H}_i^S$  即可计算空域时域混合 MH 预测结果。由于  $\mathbf{H}_i^M$  包含了更多的假设预测块,因此 MH 预测结果也更为精确,对于包含

$\omega$ 是向量

复杂运动及较多纹理细节的视频序列尤为有效。但是,矩阵尺寸的增加也使得最优线性组合系数  $\omega$  的闭式解(7)计算复杂度增大,解码时间变长。此外,由于混合 MH 预测需要当前帧的空域假设集合  $\mathbf{H}_i^S$ , 因此需要采用式(2)预先独立重建当前帧。

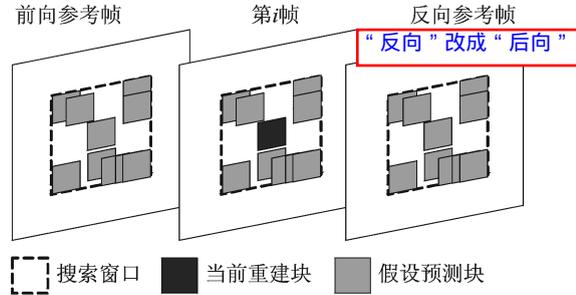


图 2 空域时域混合 MH 预测中假设集合的生成

## 2.5 算法流程

基于 MH 的分层重建以 GOP 为单位进行,单个 GOP 的重建流程如图 3 所示,分步解释如下:

**步骤 1** 重建关键帧。首先按照式(2)逐块独立重建关键帧;接下来,逐块按式(4)进行空域 MH 预测,并通过式(12)做残差重建。

**步骤 2** 按照 2.3 节所述金字塔结构为所有非关键帧分配时域重建层,并设置  $k=1$ ;

**步骤 3** 重建第  $k$  个时域层的所有非关键帧。

**步骤 4** 若所有时域层重建完毕,则当前 GOP 重建完毕;否则设置  $k=k+1$ , 跳转至步骤 3。

步骤4没有和步骤3对齐

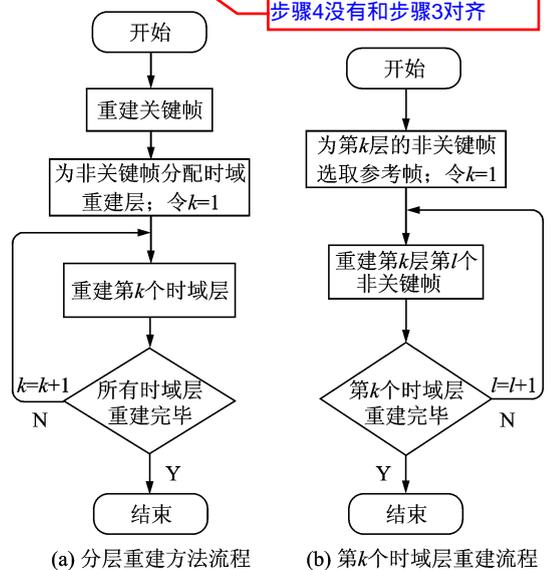


图 3 基于 MH 的分层重建方法

假设当前重建第  $k$  个时域层,重建流程如图 3 (b) 所示,分步骤解释如下

**步骤 1** 按照 2.3 节所述为当前时域层的所

有非关键帧分配参考帧,并设置  $l=1$ 。

**步骤 2** 重建当前时域层中第  $l$  个非关键帧。若采用空域时域混合 MH 预测,首先按照式(2)逐块独立重建非关键帧;接下来,按照 2.4 节所述,逐块进行空域时域混合 MH 预测,并按照式(12)进行残差重建。若仅采用时域 MH 预测,则逐块按照式(13)选取假设预测集合计算 MH 预测值,并按照式(12)进行残差重建。

**步骤 3** 若第  $k$  个时域层重建完毕,则结束本层重建过程;否则设置  $l=l+1$  并跳转至步骤 2。

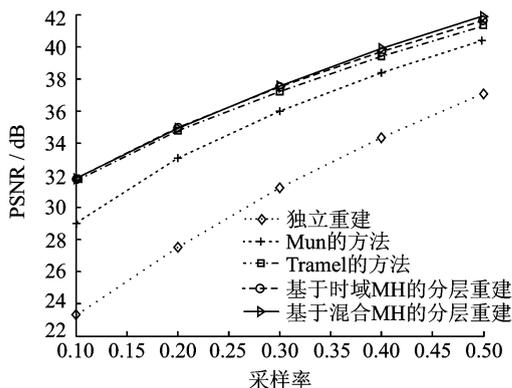
### 3 实验结果及讨论

为测试算法性能,仿真了本文方法,并与 Mun 的方法<sup>[6]</sup>、Tramel 的方法<sup>[7]</sup> 进行比较。在本文方法中,可以设置是否开启非关键帧的混合 MH 预测。当关闭混合 MH 预测时,非关键帧仅进行时域 MH 预测。在 Mun 的方法中,采用 ME/MC 的方式获取当前帧的预测值,仅使用关键帧作为参考帧;在 Tramel 的方法中,采用时域 MH 预测,也仅使用关键帧作为参考帧。公平起见,测试 Mun 的方法和 Tramel 的方法时,关键帧均采用本文方法

重建。此外,Mun 的方法需要迭代重建非关键帧,在本文实验中,迭代次数设置为 3,ME 搜索窗口大小设置为  $\pm 20$ ,搜索精度设置为  $1/4$  像素。

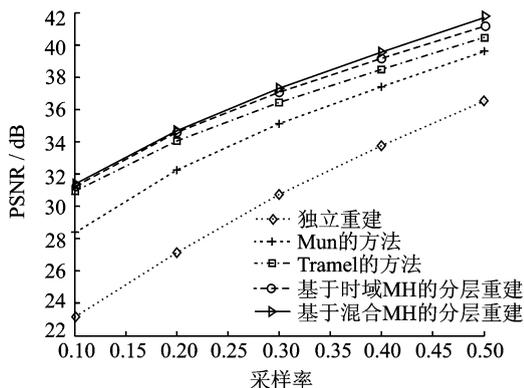
采用的测试序列为 CIF 格式的“foreman”、“coastguard”、“football”和“stefan”,帧率为 30 帧/秒。为了测试不同 GOP 下的算法性能,选取了长度为 8 和 16 的两种 GOP;非关键帧的采样率  $R_1$  设置为 0.1,0.2,0.3,0.4 和 0.5;关键帧的采样率  $R_2 = R_1 + 0.1$ 。分块大小  $B$  设置为 16,用于 MH 预测的搜索窗口  $w$  设置为 15。用于信号重建的 TVAL3 算法中的  $\mu$  选为  $2^{12}$ ,  $\beta$  选为  $2^8$ ,外循环门限设为  $10^{-6}$ ,内循环门限设为  $10^{-3}$ ,最大迭代次数设为 150。在求解 MH 预测最佳系数时,式(5)中的拉格朗日系数  $\lambda$  设为 0.25。实验在 Matlab 平台上进行,计算机中央处理器为 Intel i5-2300,主频为 2.8 GHz,安装内存为 4 GB。

图 4~图 7 分别展示了 4 个序列在 5 个采样率下非关键帧重建图像的平均峰值信噪比(Peak signal to noise ratio, PSNR)比较。作为比较基准,在图中也同时列出了独立重建方法的性能。从图中可知,首先,在相同采样率下,与基于 ME/MC



(a) DOP=8的重建质量比较

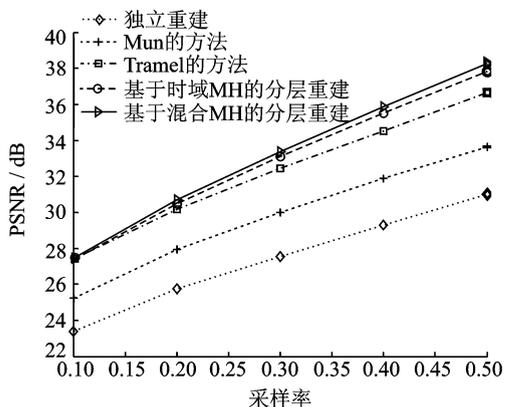
不是DOP是GOP



(b) DOP=16的重建质量比较

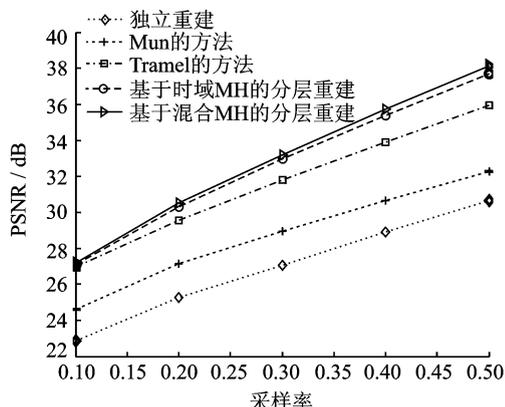
不是DOP是GOP

图 4 “foreman”序列实验结果



(a) DOP=8的重建质量比较

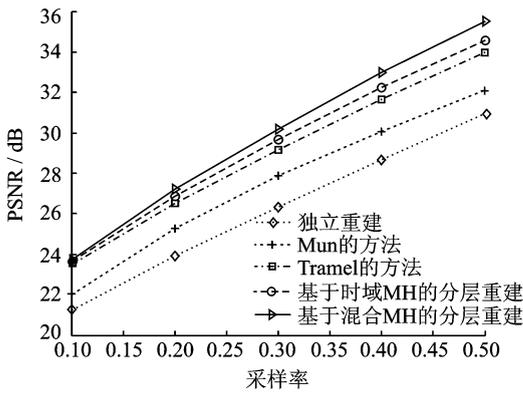
不是DOP是GOP



(b) DOP=16的重建质量比较

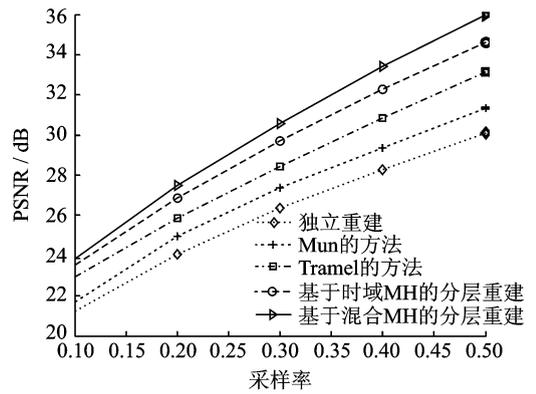
不是DOP是GOP

图 5 “coastguard”序列实验结果



(a) DOP=8的重建质量比较

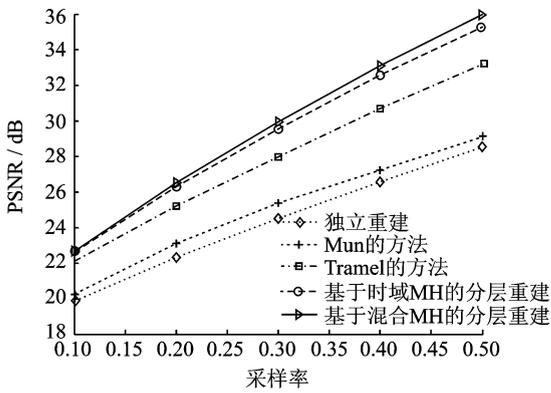
不是DOP是GOP



(b) DOP=16的重建质量比较

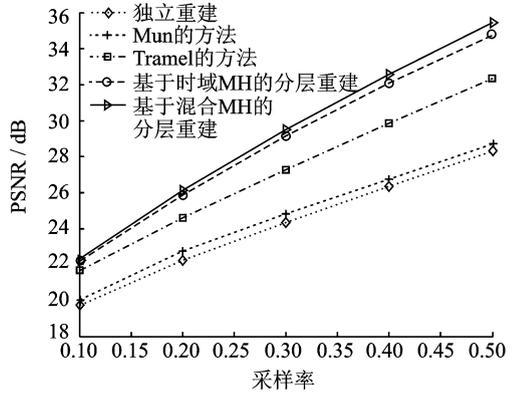
不是DOP是GOP

图6 “football”序列实验结果



(a) DOP=8的重建质量比较

不是DOP是GOP



(b) DOP=16的重建质量比较

不是DOP是GOP

图7 “stefan”序列实验结果

的方法相比,基于MH预测的方法可以获得更精确的预测值,因此残差重建的信号质量更高;其次,由于更合理地利用了时域相关性,本文提出的基于时域MH的分层重建方式可以获得比Tramel的方法、Mun的方法更高的PSNR;另外,若启用本文提出的空域时域混合MH方法以综合利用空域相关性,图像的重建质量还可以进一步提升。

接下来,以Tramel的方法为基准进一步分析图4~图7中本文算法的性能。可以发现,所提出的两种分层重建方法的PSNR增益与序列特性、GOP大小以及采样率相关,具体可总结出如下规律:

(1) GOP越大,分层重建方法的PSNR增益越大。例如“coastguard”序列,在采样率为0.3时,若GOP设置为8,基于时域MH的分层重建可获得0.66 dB的PSNR增益,基于混合MH的分层重建可获得0.82 dB的增益;当GOP扩大为16,PSNR增益分别增加至1.06 dB和1.27 dB。

(2) 序列包含的运动越复杂、剧烈,分层重建方法的PSNR增益越大。例如,对仅包含平缓运动的“foreman”序列,在GOP设置为8时,分层重

建方法的增益不明显,特别是低采样率下,3条率失真曲线几乎重合;对于包含剧烈运动的“football”序列,PSNR增益较大。

(3) 随着采样率的升高,分层重建方法PSNR增益逐渐增大。例如“stefan”序列,当GOP为16时,随着采样率的逐渐升高,基于混合MH的分层重建PSNR增益增大,最高达3.21 dB。

图8展示了“football”序列在GOP为16,非关键帧采样率为0.3下的主观重建质量比较。从图中可观察到,独立重建方法、Mun的方法产生的块效应明显;Tramel的方法产生了较多不规则噪点,图像边缘模糊;与另外3种重建方法相比,本文提出的2种分层重建方法有明显的主观质量提升,边缘和细节信息更为清晰。

此外,为了分析各算法的复杂度,在表1列出了测试序列“football”中,5种算法下非关键帧的平均重建时间。

从表1中可以观察到,在独立重建中,由于不需要对当前帧进行预测,该方法速度最快;为了获取待重建帧的准确预测值,Mun的方法需要在关键帧和待重建帧之间进行ME/MC操作,并对目标

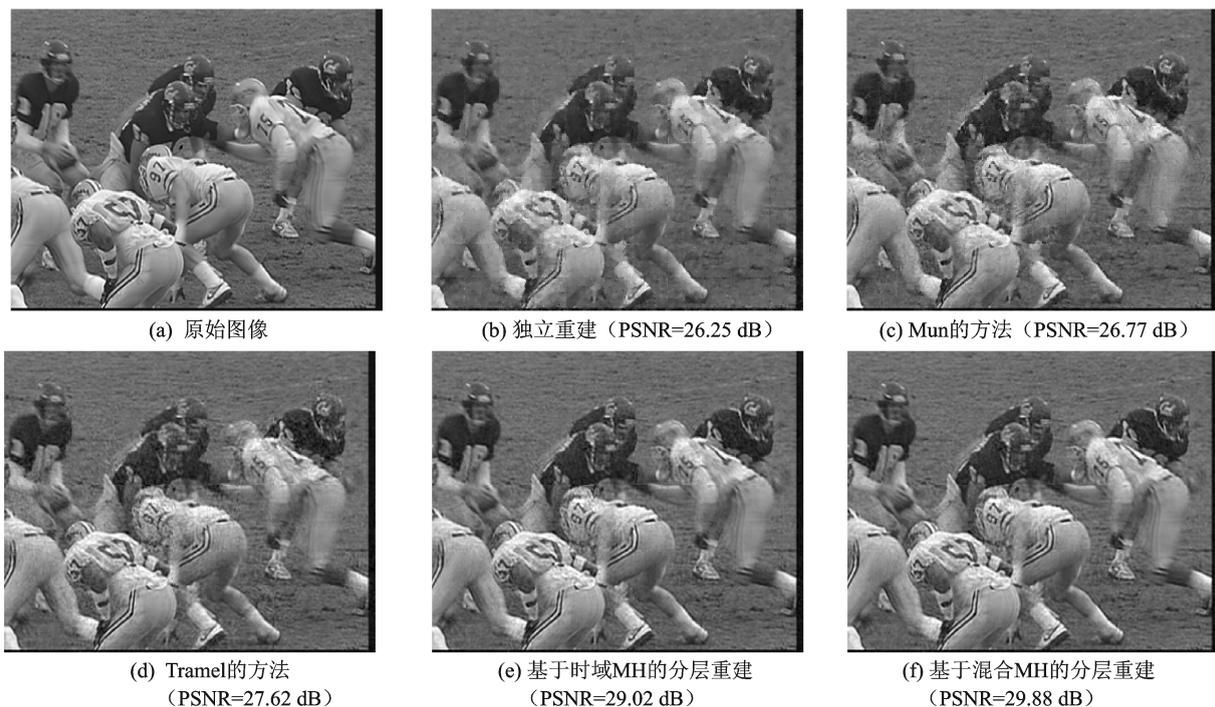


图 8 GOP=16,非关键帧采样率为 0.3 下的“football”序列主观质量比较

帧进行迭代重建,因此复杂度升高;在 Tramel 的方法中,除残差重建外,还需要计算式(7)以获取最优 MH 预测系数组合,所以需要花费较多运算时间;与 Tramel 的方法相比,本文所提出的基于时域 MH 的分层重建并没有增加额外的运算量,故运算时间与 Tramel 的方法基本一致;与时域 MH 方式相比,基于混合 MH 的分层重建方法需要额外重建一次目标帧,并且假设集合中包含更多数据导致了求解最优系数复杂度升高,因此该方法的运算时间最长。因此,本文建议在视频序列变化较小的应用场合中,例如无线视频监控中,关闭混合 MH 方案而仅采用时域 MH 分层重建。此外,还可观察到,随着采样率的升高,5 种重建方法的运算时间都不同程度地增加。主要原因是在高采样率下测量值变多,一方面,用于信号重建的 TVAL3 算法运算时间增加;另一方面,MH 预测最佳系数的求解时间变长。

表 1 “football”序列各重建方法的时间复杂度比较

采样率	独立重建	Mun 的方法	Tramel 的方法	基于时域 MH 的分层重建	基于混合 MH 的分层重建
0.1	3.75	24.22	32.12	32.44	61.17
0.2	4.23	28.13	33.75	33.63	66.36
0.3	5.26	32.72	37.92	38.37	71.65
0.4	6.17	36.45	40.85	41.47	77.54
0.5	6.75	39.67	44.37	45.08	82.13

在实际的编码过程中,量化步骤不可或缺,但

是由量化操作所引入的量化噪声会降低本文算法的性能。首先,量化噪声会影响式(7)中最优系数集合的计算结果,进而降低对待重建帧预测的准确性;其次,在本文算法中,需要采用式(12)进行残差重建,但是,量化噪声会降低重建信号的质量。为了分析量化噪声对本文算法的影响,在本部分的实验中,对测量值进行了均匀量化,测试了 GOP 为 16 下,分别采用 7 byte 量化、6 byte 量化、5 byte 量化的算法性能,并在图 9 展示了 4 个序列的实验结果。从图 9 中可以观察到:

byte 是字节,连续 3 个“byte”应改为“bits”

(1) 量化噪声会较大程度地影响本文算法的性能,量化步长越大,引入的量化噪声越多,算法性能降低也就越多。

(2) 高采样率下算法性能降低的幅度大于低采样率下的情况。主要原因是:高采样率所带来的一部分性能增益被量化噪声所抵消,待重建帧预测结果的准确性降低,因此残差重建算法的优势不能很好地体现。

“byte”应改为“bits”

(3) 相比之下,7 byte 量化引入的量化噪声较少,因此算法性能降低也相对较小。4 个序列在中、低采样率下的 PSNR 降低均在 0.5 dB 以内。综上,推荐在实际编码中采用低采样率与高比特均匀量化的参数组合。

需要注意的是,均匀量化并不是最佳量化方式,可以根据测量方式的特点,采用非线性量化,以

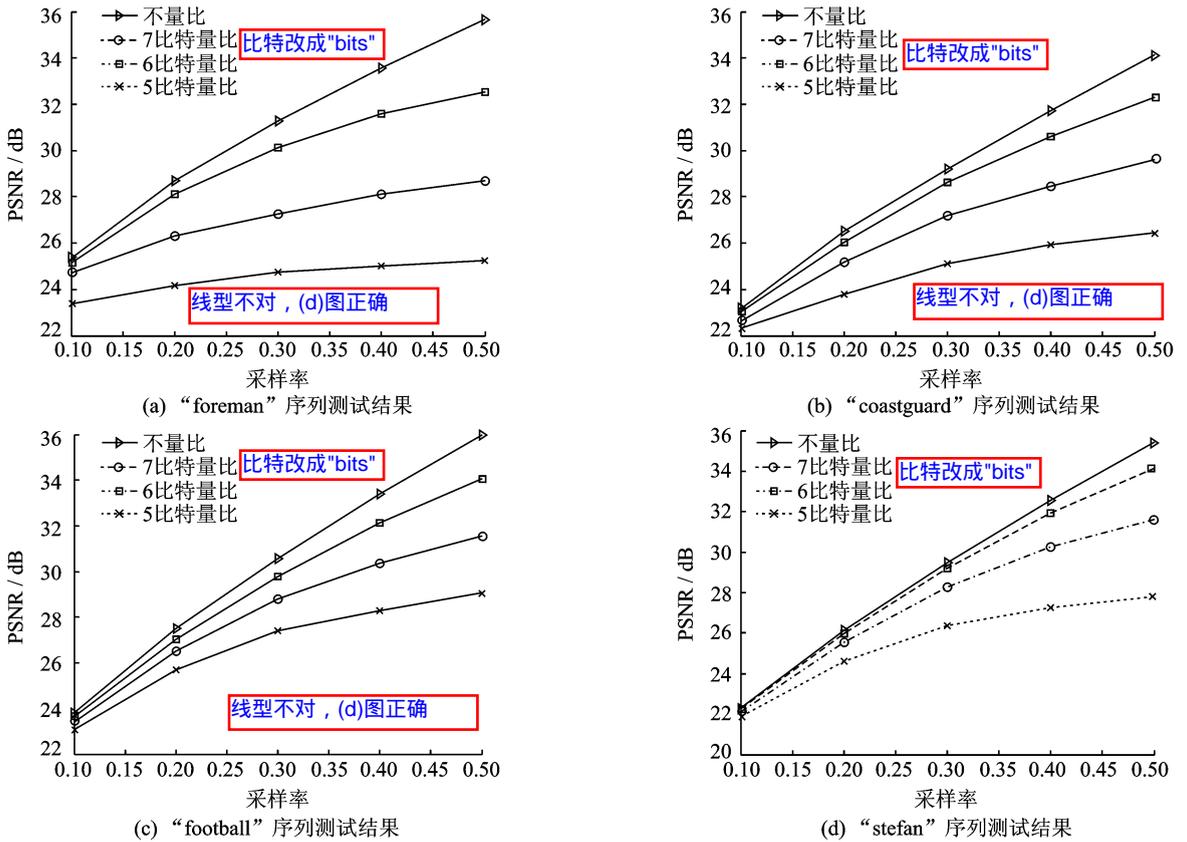


图 9 量化噪声对基于混合 MH 的分层重建算法的影响(GOP=16)

降低引入的量化噪声;此外,研究和运用对量化噪声鲁棒性更强的重建算法,也可以有效提升信号的重建质量。但是,上述改进方法超出了本文的讨论范围,课题组会在后续的工作中做进一步的研究。

## 4 结束语

本文提出了一种基于 MH 的视频压缩感知分层重建方法。对关键帧进行空域 MH 预测和残差重建;根据金字塔结构给非关键帧分配时域重建层,从低层到高层逐层重建;在非关键帧的重建过程中应用了混合 MH 预测和残差重建。实验证明,提出的方法能够有效利用时域空域相关性,从而明显提升重建视频质量,且算法的复杂度增加可接受。进一步地,实验还讨论了量化噪声对本文算法的影响,对采样率与量化比特数的设置给出了建议。未来的工作将主要围绕最佳非线性量化方式展开。

### 参考文献:

[1] Candes E, Roberg J, Tao T. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information [J]. IEEE Trans-

action on Information Theory, 2006, 52(2): 489-509.

- [2] 石光明,刘丹华,高大化,等. 压缩感知理论及其研究进展 [J]. 电子学报, 2009, 37(5):1070-1081.  
Shi Guangming, Liu Danhua, Gao Dahua, et al. Advances in theory and application of compressed sensing [J]. Chinese Journal of Electronics, 2009, 37(5):1070-1081.
- [3] 张弓,杨萌,张劲东,等. 压缩感知在雷达目标探测与识别中的研究进展[J]. 数据采集与处理, 2012, 27(1): 1-12.  
Zhang Gong, Yang Meng, Zhang Jindong, et al. Advances in theory and application of compressed sensing in radar target detection and recognition [J]. Journal of Data Acquisition & Processing, 2012, 27(1): 1-12.
- [4] Kang Li-wei and Lu Chun-shien. Distributed compressive video sensing [C]// Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP). Taipei, Taiwan: IEEE, 2009: 1169-1172.
- [5] Do T, Chen Yi, Nguyen D, et al. Distributed compressed video sensing [C]// Proceedings of IEEE International Conference on Image Processing (ICIP). Cairo, Egypt: IEEE, 2009:1393-1396.

- [6] Mun S and Fowler J. Residual reconstruction for block-based compressed sensing of video [C]// Proceedings of Data Compression Conference (DCC). Snowbird, Utah, USA: IEEE, 2011:183-192.
- [7] Tramel E, Fowler J. Video compressed sensing with multihypothesis [C]// Proceedings of Data Compression Conference (DCC). Snowbird, Utah, USA: IEEE, 2011:193-202.
- [8] 练秋生, 田天, 陈书贞, 等. 基于变采样率的多假设预测分块视频压缩感知 [J]. 电子与信息学报, 2013, 35(1): 203-208.  
Lian Qiusheng, Tian Tian, Chen Shuzhen, et al. Block compressed sensing of video based on variable sampling rates and multihypothesis predictions [J]. Journal of Electronics & Information Technology, 2013, 35(1): 203-208.
- [9] Gan Lu. Block compressed sensing of natural images [C]// Proceedings of the International Conference on Digital Signal Processing. Cardiff, UK: IEEE, 2007:403-406.
- [10] Li Chengbo. An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing [D]. Houston: Department of CAAM, Rice University, 2009.
- [11] Schwarz H, Marpe D, Wiegand T. Hierarchical B pictures [EB/OL]. [http://ftp3.itu.int/av-arch/jvt-site/2005\\_07\\_Poznan/JVT-P014.zip](http://ftp3.itu.int/av-arch/jvt-site/2005_07_Poznan/JVT-P014.zip). 2005-07-19 [2012-11-16].
- 作者简介:**常侃(1983-),男,副教授,博士,研究方向:压缩感知、视频编码, Email: pandack0619@163.com; 覃团发(1966-),男,教授,博士,研究方向:信号处理、多媒体通信、网络编码; 唐振华(1979-),男,副教授,博士,研究方向:视频编码、图像处理。