



星地协同的卫星视频高效压缩方法

肖晶^{1,2} 胡瑞敏^{1,3}

1 武汉大学计算机学院国家多媒体软件工程技术研究中心,湖北 武汉,430072

2 地球空间信息技术协同创新中心,湖北 武汉,430079

3 武汉大学多媒体网络通信工程湖北省重点实验室,湖北 武汉,430072

摘要:新兴的视频卫星数据采集速率与星地实时传输速率间存在巨大差异。传统视频编码技术通过去除单个视频中短时局部冗余以降低待传输数据量,但仍无法满足在现有星地传输带宽下实时数据传输的需求。长程背景冗余由多个遥感视频对同一观测区域重复拍摄所引发,广泛存在于具有一定时间跨度的重复拍摄的卫星遥感视频数据中。针对这一冗余去除问题,首先探讨了影响地貌背景变化的主要因素,提出了长程背景参考字典;其次,提出了基于长程背景参考字典的参考帧生成方法;最后,提出了基于长程背景参考的卫星视频编解码框架。实验表明,通过引入长程背景字典参考,卫星视频压缩后码率较现有方法降低64.38%,有助于推动卫星视频实时监测的发展与应用。

关键词:卫星视频数据;高效视频压缩;长程背景冗余;星地协同

中图分类号:P236

文献标志码:A

视频卫星由于能够有效采集动态遥感数据,且在智慧交通、排污监测、灾害监测等方面具有重大应用潜力,近年来受到遥感领域的广泛关注^[1]。高清视频数据(高空间分辨率和时间分辨率)的获取急速提高数据的采集速率,使其远高于星地间实时传输带宽,阻碍了高码率视频数据的实时下传,制约了卫星视频的实时应用。

现有的视频编码技术可通过对原始采集视频数据的压缩来降低待传输数据的码率,但要实现数据的实时传输,仍需进一步提高数据压缩效率。以吉林一号视频卫星(<http://www.gissky.net/Item/9152.aspx>)为例,视频单帧分辨率高达 $12\,000 \times 5\,000$ 个像素,帧率为15帧/s,每1 s将会产生20 GB的数据。然而,真实的星地实时传输链路的带宽仅为10~20 Mb。即使用现有最先进的高效视频编码标准^[2](high efficiency video coding, HEVC)进行300:1的压缩,待传输数据率与传输带宽仍然存在3~6倍的差距。因此,亟需发展针对卫星视频的高效编码方法。

卫星(轨道高度500~600 km)视频主要以静态背景为主,变化场景或运动目标占比普遍小于

10%。长时间累积观测视频之间存在大量重复区域,而由于地貌变化缓慢,多颗卫星不同时间拍摄的同一区域视频间具有较高相似性,这就产生了一种新的数据冗余,称为长程背景冗余。在长时间范围内对卫星视频数据进行观察可发现,随着卫星视频拍摄数据数量的累积,视频间背景重复的概率逐渐上升,长程背景冗余在数据中的占比也随之增大。如能有效去除卫星视频的长程背景冗余,将有效提升卫星视频的压缩效率。

主流的视频编码技术主要通过宏块的预测来探索视频帧内/帧间的时空相似性^[2-3],可以有效去除单个视频中的局部时空冗余。而现有针对背景不变视频的压缩方法^[4-7]主要通过背景建模去除单个视频因短时背景恒定产生的背景冗余,难以应对多源卫星视频中同一区域背景由于拍摄时间、姿态、高度等引起的图像表观差异。因此,现有的单源视频背景冗余去除方法无法挖掘跨视频的背景相似性,制约了压缩效率的进一步提升。

多源数据压缩理论研究近年来受到国内外的广泛关注,代表性的工作包括基于群体图像的压缩方法^[8-10]。一些研究者主要利用尺度不变特征变

收稿日期:2018-05-08

项目资助:国家自然科学基金(91738302,61671336);中央高校基本科研业务费专项资金(413000048);测绘遥感信息工程国家重点实验室开放研究基金(17E03)。

第一作者:肖晶,博士,副教授,主要从事卫星视频压缩和传输的理论与方法研究。jing@whu.edu.cn

通讯作者:胡瑞敏,博士,教授。hrm@whu.edu.cn

换(scale-invariant feature transform, SIFT)特征来度量不同图像间的宏块相似性。由于SIFT特征具有旋转不变性和光照变化的鲁棒性,因此能够建立起不同图像间的关系,从而实现不同图像间的宏块预测,由此消除多图像间的冗余。最近,多源图像的压缩思想被延伸到多源视频中^[11]。该研究首先在数据集中选取一个基础视频,并通过投影和光照变化等操作对基础视频进行调整,之后通过参考调整后的基础视频来消除待编码视频中的冗余。这些方法以一个数据集为对象,提出了数据集中多源数据压缩的有效方法,但由于在像素层面匹配的宏块在现实世界中并无实际对应关系,导致参考宏块搜索范围大、合成图像存在块效应等问题,不适用于卫星视频中大面积的背景区域的匹配与预测。

在多源数据的压缩中,参考库被提出用于记录所有视频数据中的共性信息,例如,利用2D车辆图像库^[12]或者3D车辆模型库^[13-14]来消除因相似车辆重复而造成的数据冗余。这些工作说明参考库能够有效消除不同视频间的冗余。同时,与数据集的压缩不同,基于参考库的方法记录了

共性信息及其变换方式,比仅在像素域选取参考信息具有更接近真实世界中的相关关系,变换形式更多样,更能提供有效参考的优点。因此本文卫星视频的编码方法也基于参考库的思路,针对卫星视频背景参考库的建立及参考方法来展开。

本文针对卫星视频背景重复的特点,发展了基于长程背景参考字典的视频编码方法,在利用吉林一号真实数据的实验中使用该方法可以平均节省64.38%的码率。同时,本文也讨论了不同的因素对码率节省的影响程度。

1 长程背景参考字典的构建

1.1 长程背景冗余的产生机理

长程背景冗余是由于不同视频对同一地点进行重复拍摄导致这些视频的背景间存在相似性而产生的。如图1所示,两幅影像虽然颜色、角度都有些差异,但可以很明显地看出是同一地区的影像。在大的时间尺度中,长程背景冗余存在以下几个方面的特点。



图1 同一地区不同时间拍摄的两端视频截图

Fig. 1 Appearances of Two Video Clips Shot of the Same Location at Different Time by Two Different Satellites

1)背景结构一致性。由于地貌变化缓慢,特定地区的地貌结构存在一定的时间稳定和持续性。因此,不同卫星视频对该地区反复拍摄时,视频中呈现的背景结构在现实世界中是一致的,称为背景结构的一致性。即使不同视频中,背景图像表观呈现一定的差异,人们也能很容易地判断出这是同一地区。

2)图像表观变化性。由于卫星视频的拍摄环境不同,如自然环境(天气、光照等)和仪器设备环境(卫星平台、镜头等)等,相同地区的地貌在不同视频中呈现出一定的背景图像表观差异。具体来说,包含以下几点。

(1) 投影差异。拍摄地区在视频中的图像

表观是这个地区在卫星相机影像平面上的投影,投影之后的结果受到相机的位置、角度的影响。因此投影差异存在于不同视频的每一帧影像间。

(2) 辐射差异。卫星视频影像的颜色受到拍摄地区环境辐射的影响而呈现一定的变化。这些变化在一定时间内稳定不变,能够通过对同一区域不同图像间的颜色变化建模来实现。

(3) 质量差异。卫星视频影像的质量受很多因素的影响,如传感器本身的光学成像系统、电子信号转换以及平台的运动等。这些因素都可在同一个视频拍摄时间内保持稳定,所以质量的变化可以以视频为单位来进行建模。

1.2 长程背景参考字典的构建

为了建立能够适用于长程背景冗余去除的参考字典,需要遵循以下原则。

1) 覆盖拍摄区域所有位置。长程背景参考字典应能够覆盖卫星轨道上的所有区域。在长程背景参考字典构建的过程中,可以先对重点区域建模,在数据累积的过程中增量添加其他区域参考影像的覆盖范围。参考字典的建立离线完成,并拷贝到卫星上。

2) 便于待压缩卫星视频数据的参考。卫星视频有别于其他通用视频,具有独特的地理空间属性。通过地理位置的索引较像素匹配更有利于视频影像编码时的区域性匹配与参考。因此,长程背景参考字典构建时,需要对参考影像进行地理配准,并在大地坐标系下对参考字典中的影像进行组织。

3) 有利于不同的投影、光照、质量的表达。作为参考影像,生成时应该尽量在历史数据中选取正视影像或经过正射纠正的影像,减小投影差带来的影响。光照条件通过选择白天、曝光条件好、无云干扰的影像来保障。同时,选择高质量、高分辨率的影像,确保在参考帧的生成过程中只有降质、降分辨率的操作。

根据以上 3 点原则,本文在历史数据中选取光照条件好、分辨率高、清晰度高的影像,经过地理配准或正射校正、图像拼接,形成整体的长程背景参考字典。为了提高参考的准确性,考虑构建一个地区多个参考模态的多模态参考字典,但是在本文中对一个地区只使用单张参考图像。

2 参考帧的生成及编码框架

2.1 背景参考帧的生成

预测编码是现有编码算法中去除冗余最有效的环节。首先利用长程背景参考字典中的参考影像生成背景参考帧,为视频帧编码提供高相似性的参考,提高预测效率,去除长程背景冗余。生成背景参考帧主要包含图像匹配、辐射校正和质量校正 3 个步骤。

1) 图像匹配。虽然长程背景参考字典中的参考影像经过严格的地理配准,但受到卫星定位精度及相机镜头畸变等的影响,存在一定的图像偏移与变形,因此通过图像匹配,将参考影像变换到具有与当前待编码视频帧相同的几何结构。需要通过分辨率下采样、同名点匹配、变换预测与图像重采样过程^[15-16] 实现参考影像向当前参考帧的

匹配。

大致的视频拍摄区域由卫星的定位系统获得,并在长程背景参考字典中截取较拍摄范围较大的区域,确保所截取区域能够容纳因定位误差引起的偏差。将所截取的参考影像与待编码视频帧进行同名点匹配,获取两者间的几何变换关系。通常使用 SIFT 特征点匹配^[17] 来寻找同名点。但因为参考影像与待编码视频帧存在分辨率、质量和辐射上的差异,利用 SIFT 特征点进行匹配效果欠佳,无法获取足够的同名点对。为解决该问题,首先将高分辨率的参考影像降分辨率到与当前待编码视频帧相同的分辨率,然后采用针对多传感器遥感影像匹配改进的 SIFT 匹配算法^[18]。该算法采用了基于自相似描述子的特征排序,可对光照变化保持稳定性,有效解决了不同时刻不同传感器拍摄的遥感影像间的辐射和光照差异问题,有助于在参考影像与待编码视频帧间获取有效数量的同名点对。

本文采用透视投影变换来模拟两幅影像之间的关系。根据获取的同名点间的对应关系求解方程组(1),即可得到透视投影变换矩阵

$$\begin{bmatrix} a_0 & a_1 & a_2 \\ a_3 & a_4 & a_5 \\ a_6 & a_7 & 1 \end{bmatrix} \text{ 中的 8 个参数。}$$

$$\begin{cases} x_i^e = \frac{a_0 x_i^r + a_1 y_i^r + a_2}{a_6 x_i^r + a_7 y_i^r + 1} \\ y_i^e = \frac{a_3 x_i^r + a_4 y_i^r + a_5}{a_6 x_i^r + a_7 y_i^r + 1} \end{cases} \quad (1)$$

式中, (x_i^e, y_i^e) 为当前待编码视频帧中的点坐标; (x_i^r, y_i^r) 为下采样后参考图像中同名点的坐标; i 为第 i 组同名点。

获取透视投影参数后,可以通过透视投影变换经过图像重采样生成与待编码视频帧几何一致的参考影像 I_r^e 。

2) 辐射校正。辐射校正主要将参考影像的颜色变换到待编码视频帧的颜色,实现颜色域的一致。本文采用文献[19]中提出的颜色变换模型来实现参考影像的颜色变换。由于视频编码采用 YUV 颜色空间,所以颜色变换也是在 YUV 空间进行。考虑到 YUV 颜色空间与文献[19]方法的 $l\alpha\beta$ 颜色空间相似(第一个颜色通道为亮度,另外两个颜色通道是色度),可直接套用 $l\alpha\beta$ 颜色空间的变换模型来进行参考影像的颜色变换。将 $l\alpha\beta$ 颜色空间替换为 YUV 颜色空间之后的变化模型如下:

$$\begin{bmatrix} Y_t \\ U_t \\ V_t \end{bmatrix} = \begin{bmatrix} \sigma_c^Y & 0 & 0 \\ 0 & \sigma_s^U & 0 \\ 0 & 0 & \sigma_s^V \end{bmatrix} \begin{bmatrix} Y_s \\ U_s \\ V_s \end{bmatrix} - \begin{bmatrix} \bar{Y}_s \\ \bar{U}_s \\ \bar{V}_s \end{bmatrix} + \begin{bmatrix} \bar{Y}_c \\ \bar{U}_c \\ \bar{V}_c \end{bmatrix} \quad (2)$$

式中, $[Y_t \ U_t \ V_t]^T$ 和 $[Y_s \ U_s \ V_s]^T$ 分别是颜色变换之后的参考影像 I_r^c 和颜色变化前仅经过几何变换的参考影像 I_r^g 对应像素的颜色值; $[\bar{Y}_s \ \bar{U}_s \ \bar{V}_s]^T$ 和 $[\sigma_s^Y \ \sigma_s^U \ \sigma_s^V]^T$ 是 I_r^g 影像各颜色通道的均值与方差; $[\bar{Y}_c \ \bar{U}_c \ \bar{V}_c]^T$ 和 $[\sigma_c^Y \ \sigma_c^U \ \sigma_c^V]^T$ 是待编码视频帧各颜色通道的均值与方差。

采用该颜色变换模型后, 几何变换之后的参考影像 I_r^g 根据待编码视频帧的颜色统计特点变换到颜色变换后的参考影像 I_r^c 。

3) 质量校正。在构建长程地理参考字典时, 已经选用了最高清晰度的历史影像数据作为参考影像, 所以质量校正通常是将参考图像模糊化的过程。这里采用最常用的二维高斯模糊函数来对颜色变换后的参考影像 I_r^c 进行质量变换, 生成最终的参考影像 I_r 。

假设质量的变换在整幅影像中是均一的, 采用均值为 0、方差为 σ 的各向同性的高斯模型。实际过程中, 高斯模糊通过将高斯函数转化到一

个 $n \times n$ (本文中选用 5×5)的卷积核对图像进行滤波, 卷积核的参数由方差 σ 与参数到核中心的距离所组成的多项式决定, 由此, 经过高斯模糊后的影像 I_r 的像素值可以由方差 σ 来决定。基于此, 通过在参考影像 I_r^c 和待编码视频帧中随机选取 N 个 5×5 的宏块对, 利用最小二乘方法来最小化以下目标函数, 求取质量变换的模型参数 σ :

$$\operatorname{argmin}_{\sigma} \sum_{k=0}^N (G(\sigma) * B_r^k - p_c^k) \quad (3)$$

式中, B_r^k 是参考影像 I_r^c 中的第 k 个宏块像素值; p_c^k 是待编码视频帧中的第 k 个对应的宏块像素值; $*$ 表示卷积操作; $G(\sigma)$ 是 5×5 的高斯卷积核, 其中各参数受方差控制如下:

$$a_{ij} = \frac{1}{\sum_{ij}} \cdot a'_{ij}, i, j = -2, -1, 0, 1, 2 \quad (4)$$

$$a'_{ij} = e^{-\frac{(i^2 + j^2)}{2\sigma^2}}, i, j = -2, -1, 0, 1, 2 \quad (5)$$

最终的参考影像 I_r 由 $G(\sigma) * I_r^c$ 获得。

2.2 编解码框架

基于长程背景参考的星地协同卫星视频压缩方法流程图如图 2 所示。长程背景参考字典在地面离线生成, 并拷贝到卫星上的编码器。对于新的卫星拍摄视频, 编码端首先生成参考影像, 再编码差异信息, 并和参考影像生成的控制信息一起下传到地面。由于视频每一帧都存在一定的投影变化, 所以初始生成的参考影像在用于后续帧编码

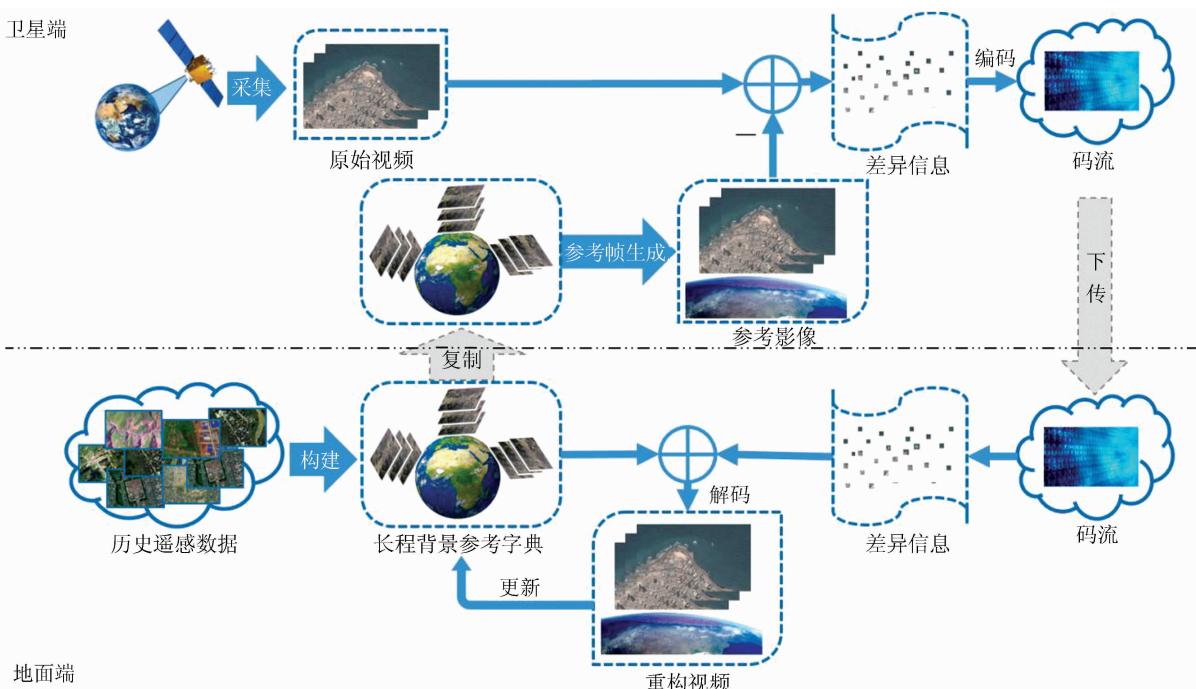


图 2 星地协同的卫星视频压缩流程示意图

Fig. 2 Overall Framework of Satellite-Earth Coordinated Compression of Satellite Video Data

时还要经过更新。地面解码端接收到差异信息后,对其进行解码,并利用参考影像生成的控制信息与解码端的长程背景参考字典结合重新生成参考影像,再与解码的差异信息结合,即可实现视频的解码。具体过程分为如下 4 个步骤。

1) 背景参考影像的生成。传统视频编码时,视频帧被分为 I 帧与 P 帧。 I 帧通常为一个图像组(group of pictures, GOP)的第一帧,只参考自己, P 帧参考前序帧。在基于长程背景参考字典进行编码时,编码端根据 I 帧视频内容采用 § 2.1 的方法生成背景参考影像,为 I 帧提供编码参考。由于生成的背景参考影像并不传输,所以人们需要传输编码端生成背景参考影像的控制信息,用于解码端对背景参考影像的生成。生成控制信息主要包括投影变换矩阵 $\begin{bmatrix} \bar{Y}_c & \bar{U}_c & \bar{V}_c \end{bmatrix}^T$ 和颜色变换参数 $\begin{bmatrix} \sigma_c^Y & \sigma_c^U & \sigma_c^V \end{bmatrix}^T$ 以及质量变换参数 σ 。生成的背景参考影像存储在临时缓冲区中作为 I 帧与后续视频帧编码的参考。

2) 背景参考影像的更新。对于 P 帧来说,辐射与质量的变化都可视为与 I 帧一致,所以 P 帧的参考影像只需要对投影变换做一定的调整。基于前序生成的背景参考影像,利用 § 2.1 中描述的投影变换方法对参考影像进行新的投影变换,并输出新的投影矩阵。更新的参考影像也放入临时缓冲区作为编码参考备选。

3) 影像宏块计算编码与预测模式。对于 I 帧来说,利用生成的背景参考影像作为参考帧进行编码。除了传统的帧内预测编码模式,新增了一种参考背景参考影像的长程预测模式,在编码每一个宏块时,通过率失真优化方法来判定所选的参考模式。由于帧间预测模式的效率通常优于帧内预测模式的效率,所以利用背景参考影像的长程预测模式能够有效降低码率。对于 P 帧来说,更新的背景参考影像作为其参考帧之一,在编码宏块时,在帧内预测、帧间预测与长程预测模式中进行选择。

4) 影像宏块的编码与解码。进行了宏块预测之后,预测残差将通过变换、量化、熵编码进行编码。编码之后,视频帧同样在编码端进行重建,并存储在临时缓冲区作为短时参考帧的备选。

完成视频序列编码之后,背景参考影像的生成参数、预测参数和残差数据组织成码流文件,经过星地间传输链路传输到地面端。然后地面端通过对背景参考影像与视频残差的重建来合成最终的解码视频。

3 实验及结果分析

3.1 编解码框架

本文利用 4 段拍摄于西班牙巴伦西亚的吉林一号视频卫星数据进行数据压缩实验(图 3)。图 3 中分别包含了 3 种场景 4 段视频,两段城区视频,一段农田区域视频和一段海边区域的视频,用来测试本文方法在不同类型地区的效果。所有视频都经过了稳像的预处理,并裁剪到 720P(1 280×720 像素)的分辨率大小。视频帧率为 15 帧/s,每段视频均包含 300 帧。

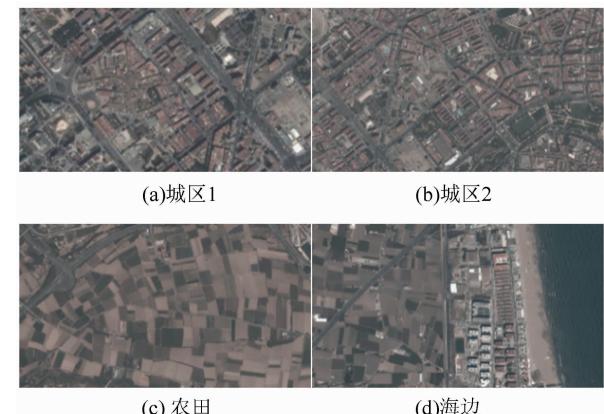


图 3 吉林一号 4 段卫星视频首帧展示图

Fig. 3 The First Frame of Four Satellite Videos
in Satellite Jilin-1

由于暂时无法获取吉林一号卫星的历史数据,但谷歌地球中的影像已经几何校正到大地坐标系下,且较吉林一号卫星的视频帧具有更好的分辨率与质量,满足地理背景参考字典建立的原则,所以本文借用谷歌地球中的影像作为长程背景参考字典中的影像数据。

为了测试本文方法的效果,对比算法选取已经在美国 Skysat 卫星上使用的视频压缩方法 H.264^[20],其编码器为 JM19.0,档次为 baseline profile。 I 帧间隔为 8,编码结构为 IPPP,不打开码率控制开关。其他测试参数均为默认参数。视频压缩测试在 22、27、32、37 这 4 个量化参数(quantization parameter, QP)中进行。采用 Bjontegaard 评价机制^[21],通过对比 BD-PSNR(Bjontegaard delta peak signal-to-noise rate)和 BD-Rate(Bjontegaard delta rate)来评价所提方法的压缩效率。

3.2 实验结果与分析

图 4 是测试视频城区 1 的参考影像生成过程。图 4(a)是待编码视频帧,图 4(b)是从谷歌地

图中下载的原始遥感影像,具有很大的图像覆盖区域。根据卫星视频定位结果,本视频所处位置在遥感影像中红框所对应的部分。图4(c)是经过下采样和几何匹配之后得到的参考影像,图4(d)是经过颜色和质量变换后生成的参考影像。可以看到,生成的参考影像与图4(a)中的待编码视频帧具有很高的相似性,即能够提供效率较高的参考。

本文方法在吉林一号4段视频上的测试效果与H.264标准的对比结果见表1和图5。可以看出,在相同质量情况下,本文算法可以平均降低64.38%的码率,而在相同的码率情况下,本文算法的质量可以平均提高5.44 dB。由此可以证明,地面信息确实可以有效提高卫星视频的压缩效率。

同时,对比4段视频的压缩效率,可以看到,码率提升最大的在农田区域。该区域基本没有建筑物,所以在卫星视频拍摄过程中的视角变化基本不会引起视频帧间的投影差变化,而在利用参考影像进行预测时能够获得较高的预测效率。而海边区域视频序列的码率提升最小,可能是因为水的波浪变化对预测造成了影响,引入了不必要的差异信息。

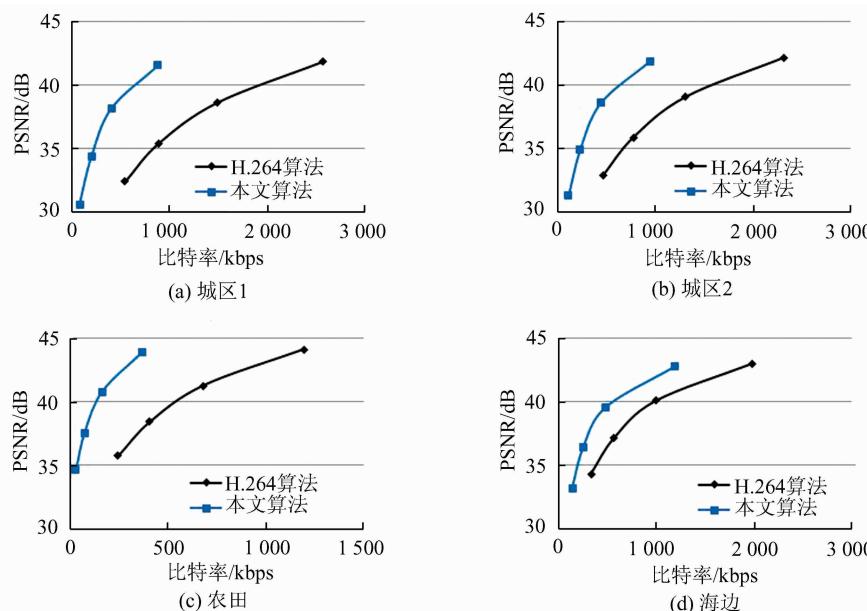


图5 本文算法与H.264的码率-质量曲线结果对比

Fig. 5 RD Curves Comparison of H.264 and the Proposed Method

本文方法带来额外的计算复杂度来源于背景参考帧的生成,包括图像匹配、辐射校正和质量校正3个过程,其他编码过程的复杂度与H.264一致。经测试,由于H.264的编码复杂度较高,编码一帧需要45 s左右的时间,而本文方法用于背景参考帧生成的额外计算量需要5.52 s,增加了

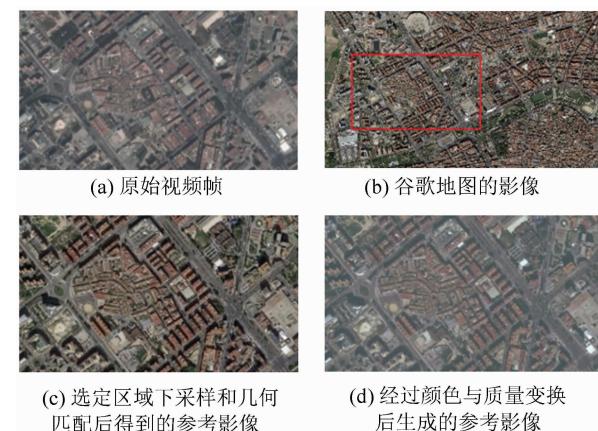


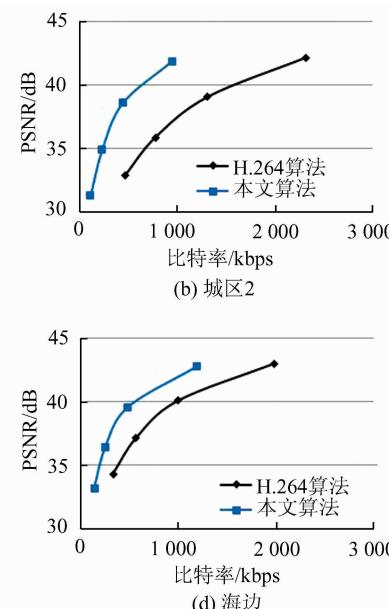
图4 参考影像,视频帧来自测试视频城区1

Fig. 4 Reference Images, Sample Frame Came from Satellite Video Clips Urban-1

表1 本文压缩算法与H.264压缩算法的对比数据

Tab. 1 The Overall BD-PSNR and BD-Rate of the Proposed Method and H.264

视频序列	BD-PSNR / dB	BD-Rate / %
城区 1	6.87	-71.42
城区 2	5.59	-63.82
农田	6.39	-76.26
海边	2.89	-46.02
平均值	5.44	-64.38



12.3%的运算复杂度。

4 结语

本文提出了一种基于长程背景参考字典的星地协同卫星视频高效压缩方法,有助于推动视频

卫星在军事与民用中的实时应用。其关键思想在于利用卫星视频对地面区域观测的重复性引发的视频间的长程背景冗余,通过从采集的历史数据中建立长程背景参考字典,生成长程背景参考,而去除长程背景冗余。与现有的针对单个视频独立编码的技术相比,本文方法将地面拍摄对象信息作为先验知识,并进行合理组织,使其成为可以被星上编码器所利用的信息。利用地面信息协助星上数据编码的星地协同数据压缩方法可有效避免已有地面信息占用传输带宽,而仅下传有效的差异信息。实验证明,本文方法可以较现有星上的视频编码方法平均降低64.38%的码率。

在后续工作中,将进一步研究长程背景字典的构建方法,如多模态字典等。同时,如何针对不同地物类型建立更准确的辐射与质量变换模型也是下一步的研究方向。

参 考 文 献

- [1] Li Deren, Yu Hanruo, Li Xi. The Spatial-Temporal Pattern Analysis of City Development in Countries Along the Belt and Road Initiative Based on Night-Time Light Data [J]. *Geomatics and Information Science of Wuhan University*, 2017, 42(6): 711-720(李德仁, 余涵若, 李熙. 基于夜光遥感影像的“一带一路”沿线国家城市发展时空格局分析[J]. 武汉大学学报·信息科学版, 2017, 42(6): 711-720)
- [2] Sullivan G J, Ohm J, Han W, et al. Overview of the High Efficiency Video Coding Standard [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2012, 22(12): 1 649-1 668
- [3] Wiegand T, Sullivan G J, Bjontegaard G, et al. Overview of the H. 264/AVC Video Coding Standard [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2003, 13(7): 560-576
- [4] Yin L, Hu R, Chen S, et al. Block-Based Global and Multiple-Reference Scheme for Surveillance Video Coding [C]. Pacific Rim Conference on Multimedia, Gwangju, Korea, 2015
- [5] Zhang X, Huang T, Tian Y, et al. Background-Modeling-Based Adaptive Prediction for Surveillance Video Coding [J]. *IEEE Transactions on Image Processing*, 2014, 23(2): 769-784
- [6] Zhang X, Tian Y, Huang T, et al. Optimizing the Hierarchical Prediction and Coding in HEVC for Surveillance and Conference Videos with Background Modeling [J]. *IEEE Transactions on Image Processing*, 2014, 23(10): 4 511-4 526
- [7] Huang Tiejun, Zhang Xianguo, Tian Yonghong, et al. IEEE 1857 Standard for High Efficiency Surveillance Video Compression and Recognition [J]. *Electronic Engineering and Products World*, 2013, 20(7): 22-26(黄铁军, 张贤国, 田永鸿, 等. 支持监控视频高效压缩与识别的 IEEE 1857 标准[J]. 电子产品世界, 2013, 20(7): 22-26)
- [8] Yue H, Sun X, Yang J, et al. Cloud-Based Image Coding for Mobile Devices—Toward Thousands to One Compression [J]. *IEEE Transactions on Multimedia*, 2013, 15(4): 845-857
- [9] Shi Z, Sun X, Wu F. Feature-Based Image Set Compression [C]. IEEE International Conference on Multimedia and Expo, California, USA, 2013
- [10] Wu H, Sun X, Yang J, et al. Lossless Compression of JPEG Coded Photo Collections [J]. *IEEE Transactions on Image Processing*, 2016, 25(6): 2 684-2 696
- [11] Wang H, Tian T, Ma M, et al. Joint Compression of Near-Duplicate Videos [J]. *IEEE Transactions on Multimedia*, 2017, 19(5): 908-920
- [12] Ma C, Liu D, Peng X, et al. Surveillance Video Coding with Vehicle Library [C]. IEEE International Conference on Image Processing, Beijing, China, 2017
- [13] Xiao J, Liao L, Hu J, et al. Exploiting Global Redundancy in Big Surveillance Video Data for Efficient Coding [J]. *Cluster Computing*, 2015, 18(2): 531-540
- [14] Xiao J, Hu R, Liao L, et al. Knowledge-Based Coding of Objects for Multi-source Surveillance Video Data [J]. *IEEE Transactions on Multimedia*, 2016, 18(9): 1 691-1 706
- [15] Wang Feng, You Hongjian, Fu Xingyu. Auto-adaptive Well-Distributed Scale-Invariant Feature for SAR Images Registration [J]. *Geomatics and Information Science of Wuhan University*, 2015, 40(2): 159-163(王峰, 尤红建, 傅兴玉. 应用于SAR 图像配准的自适应 SIFT 特征均匀分布算法[J]. 武汉大学学报·信息科学版, 2015, 40(2): 159-163)
- [16] Xiao Xiongwu, Li Deren, Guo Bingxuan, et al. A Robust and Rapid Viewpoint-Invariant Matching Method for Oblique Images [J]. *Geomatics and Information Science of Wuhan University*, 2016, 41(9): 1 151-1 159(肖雄武, 李德仁, 郭丙轩, 等. 一种具有视点不变性的倾斜影像快速匹配方法[J]. 武汉大学学报·信息科学版, 2016, 41(9): 1 151-1 159)
- [17] Lowe D G. Distinctive Image Features from Scale-Invariant Keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110
- [18] Sedaghat A, Ebadi H. Distinctive Order Based Self-

- Similarity Descriptor for Multi-sensor Remote Sensing Image Matching[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 108: 62-71
- [19] Reinhard E, Ashikhmin M, Gooch B, et al. Color Transfer Between Image [J]. *IEEE Computer Graphics & Applications*, 2002, 21(5):34-41
- [20] Liu Tao. Development of Video Satellite [J]. *Space International*, 2014, 429: 50-56(刘韬. 国外视频卫星发展研究[J]. 国际太空, 2014, 429: 50-56)
- [21] Bjontegaard G. Calculation of Average PSNR Difference Between RD-curves [C]. Document VCEG-M33, ITU-T VCEG Meeting, Austin, TX, USA, 2001
- [20] Liu Tao. Development of Video Satellite [J]. *Space*

Satellite-Earth Coordinated High Efficiency Compression of Satellite Video Data

XIAO Jing^{1,2} HU Ruimin^{1,3}

1 National Engineering Research Center for Multimedia Software, School of Computer Science, Wuhan University, Wuhan 430072, China

2 Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China

3 Hubei Key Laboratory of Multimedia and Network Communication Engineering, Wuhan University, Wuhan 430072, China

Abstract: Video satellites have great potential for military and civil use due to their capabilities in the dynamic remote surveillance. However, the video data collection rate is far higher than the bandwidth between satellite and earth. Video coding is the technology to compression the video data to a low bitrate for transmission, by eliminating the local spatial/temporal redundancies, but there is still a gap between the bitrate of the video data and the satellite-earth bandwidth. Long-term background redundancy is a new type of redundancy existing in the remote satellite video data, caused by the repeated recording of the same place. This type of redundancy becomes significant as the video data of the same place increases. In this paper, we first discuss the factors causing the change of the background, and propose the long-term background reference library. After that, we propose the method for the background reference generation and the coding framework for satellite video data. Experimental results show that 64.38% reduction on video data bitrate can be achieved by using the proposed method, compared to the H.264 video coding standard. The proposed method will boost the applications of video satellite data in the surveillance field.

Key words: satellite video data; video compression; long-term background redundancy; satellite-earth coordination

First author: XIAO Jing, PhD, associate professor, specializes in video compression and processing. E-mail: jing@whu.edu.cn

Corresponding author: HU Ruimin, PhD, professor. E-mail: hrm@whu.edu.cn

Foundation support: The National Natural Science Foundation of China, Nos. 91738302, 61671336; the Fundamental Research Funds for the Central Universities, No. 413000048; Open Research Fund of State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, No. 17E03.