

虚实融合监控视频压缩方法的分析与评价

解愉嘉^{1,2,3}, 刘学军^{1,2,3}

(1. 南京师范大学地理科学学院, 江苏 南京 210023)

(2. 虚拟地理环境教育部重点实验室, 江苏 南京 210023)

(3. 江苏省地理信息资源开发与利用协同创新中心, 江苏 南京 210023)

[摘要] 在阐述虚实融合视频压缩基本思想的基础上, 对虚实融合视频压缩技术流程进行描述. 针对虚实融合视频压缩方法提出评价因子, 在此基础上构建压缩效果评价模型. 实验结果表明, 时间尺度因子对压缩效率的影响受运动前景目标出现情况的制约: 对于没有前景运动目标出现的段落, 单位时长的视频压缩效率高于累积时长; 对于前景运动目标出现的段落, 累积时长的视频压缩效率高于单位时长. 压缩层级因子的变化对压缩效率的影响较显著: 对象投影式相比于图像投影式, 所需使用数据量可压缩至小于后者的 1%; 前景抽象式相较于前后景独立投影式与前景投影式, 所需使用数据量可压缩至小于后者的 1%.

[关键词] 虚拟场景, 虚实融合, 监控视频, 视频压缩

[中图分类号] P208 **[文献标志码]** A **[文章编号]** 1001-4616(2016)03-0130-05

Analysis and Evaluation on the Virtual-Real Fusion Surveillance Video Compression Method

Xie Yujia^{1,2,3}, Liu Xuejun^{1,2,3}

(1. School of Geography Science, Nanjing Normal University, Nanjing 210023, China)

(2. MOE Key Laboratory of Virtual Geographic Environment, Nanjing 210023, China)

(3. Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China)

Abstract: On the basis of describing the idea of the virtual-real surveillance video compression method, this paper describes the technique process of the compression method, proposes the relative evaluation factors, and constructs the evaluation model. The experimental results have shown that, the impact from the time scale factor is conditioned by the appearance of the prospect of moving target. When there is no moving target, the compression rate of unit-time length is higher than the compression rate of cumulative-time length; when there are some moving targets, the compression rate of cumulative-time length is higher than the compression rate of unit-time length. The impact from the compression level factor is remarkable. Comparing to the compression rate of image projection fusion style, the compression rate of object projection style is less than 1%. Comparing to the compression rate of fore-ground independent and prospect projection styles, the compression rate of prospect abstraction projection style is less than 1%.

Key words: virtual scene, virtual reality fusion, surveillance video, video compression

当前, 监控相机被广泛应用于工业控制、交通、安防监控等领域. 监控相机拍摄获得的监控视频拥有巨大的数据量, 但其中有价值的信息是稀疏存在的, 其余部分多为冗余信息. 为实现有效的监控视频浏览分析, 需要对视频进行压缩处理.

当前的监控视频压缩的方法主要分为两类, 一类是图像空间视频压缩, 另一类是虚实融合视频压缩. 图像空间视频压缩通过对视频进行帧预测实现, 其目的在于以压缩数据重构视频. 虚实融合视频压缩通

收稿日期: 2015-12-12.

基金项目: 公安部科技强警基础工作专项资助项目(2014GABJC009)、国家科技支撑计划基金项目(2012BAH35B02)、江苏省优势学科建设项目(164320H116).

通讯联系人: 刘学军, 教授, 博士生导师, 研究方向: DEM 不确定性分析、影视 GIS 等. E-mail: liuxuejun@njnu.edu.cn

过提取视频前后景相关信息实现,其目的在于将视频信息与虚拟场景进行融合表达。当前对于图像空间视频压缩方法的研究很多,但对于虚实融合视频压缩方法的研究较少,尤其对不同约束条件作用下虚实融合视频压缩效果的比较研究十分匮乏。针对上述问题,本文首先描述了虚实融合视频压缩的基本思想流程,分析了压缩数据的层级关系,通过选取评价因子,构建压缩效果评价模型,最后通过实验对不同评价因子作用下虚实融合视频压缩效果进行了分析。

1 虚实融合视频压缩原理

根据相机视域的地理空间位置将视频信息展现于虚拟场景模型中是一种增强虚拟式的表达^[1]。这种表达方法可有效将视频信息与地理空间信息相关联,提供更好的视频-场景信息感知框架^[2-5],并根据视频映射方式的不同分为图像投影式和对象投影式两大类。图像投影式将视频影像依相机参数直接展现于虚拟场景中^[6]。对象投影式则通过智能视频分析,对视频前后景分别处理并进行地理空间投影^[7]。对象投影式的具体方法有前后景独立投影式^[8]、前景投影式^[9,10]、前景抽象式^[11]三种。其中,前后景独立投影式将前景运动目标子图投影至虚拟场景中对应的时空位置,同时将视频背景作为纹理贴图投影至场景模型表面;前景投影式仅将前景运动目标子图展现于场景中对应的时空位置;前景抽象式通过识别前景运动目标类型并抽象为语义图标,展现于地理场景中对应的时空位置。对象投影式融合方法所需数据组合方式构成了如图1所示的压缩层级。

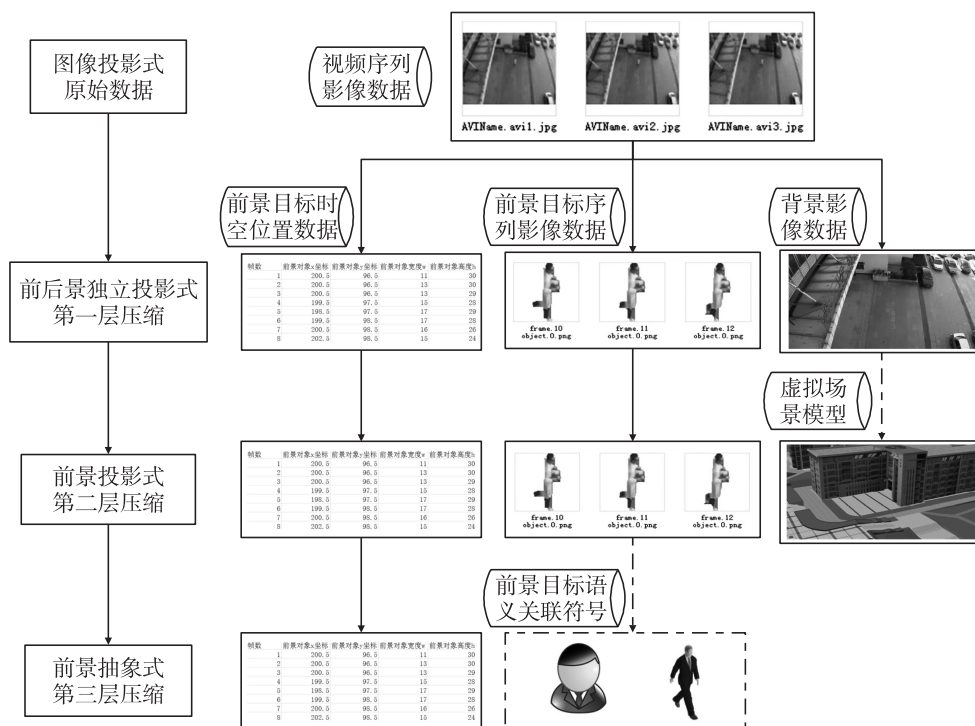


图1 视频压缩数据层级关系示意图

Fig.1 The relationship between the hierarchies of video compressed data

2 压缩效果评价模型

不同的虚实融合方法使用不同类型的视频压缩数据,所需取用的压缩数据量也随之变化。本节基于前文所述的视频压缩技术流程和压缩数据的层级关系,通过分析虚实融合视频压缩过程中的各种影响因素,提出视频压缩效果评价因子并构建相应的评价模型,实现视频压缩效果评价。

2.1 评价因子

本节提出时间尺度因子、压缩层级因子、压缩数据类型这3种压缩效果评价因子,对进行比较的压缩数据相关属性进行界定,为构建视频压缩效果评价模型提供相应的参数。

时间尺度因子:在监控视频中,前景运动目标的时序个数、目标像素数、运动强度等存在随机性,不同时间段内的视频数据压缩效果区别较大,所以需要原始视频时长范围进行界定.评价因子界定规则如下:以参数 t 表示处理视频的时长范围, $\forall t \in T$, $T = \{uni, tol\}$. uni 指视频单位时长,依据预设的背景模型更新参数取值; tol 指视频累计时长.

压缩层级因子:通过构建压缩数据层级关系,可区分相应的虚实融合模式.评价因子界定规则如下:以参数 l 表示数据压缩的层级, $\forall l \in L$, $L = \{1, 2, 3\}$. 数字表达对应的压缩层级.

压缩数据类型:通过引入压缩数据类型评价因子,可对压缩数据进行类型标记.评价因子界定规则如下:以参数 o 表示压缩数据类型, $\forall o \in O$, $O = \{ori, bm, fm, fl\}$, 其中 ori 指原始视频序列影像数据, bm 指背景影像数据, fm 指运动目标子图影像数据, fl 指前景运动目标时空位置数据.

2.2 评价模型

为了对虚实融合视频压缩的数据量缩减效果进行分析,本文构建如下评价模型:

$$F = K_{t,l}(t \in \{uni, tol\}, l \in \{1, 2, 3\}), \quad (1)$$

$K_{t,l}$ 表示时长范围 t 、压缩层级 l 的压缩比倒数(压缩存储的数据量/原始数据量),表达式如下:

$$K_{t,1} = (M_{t,bm} + M_{t,fm} + M_{t,fl}) / M_{t,ori}, \quad (2)$$

$$K_{t,2} = (M_{t,fm} + M_{t,fl}) / M_{t,ori}, \quad (3)$$

$$K_{t,3} = M_{t,fl} / M_{t,ori}, \quad (4)$$

其中, $M_{t,o}$ 表示时长范围 t , 压缩数据类型 o 的某一类数据量.

从式(2)-(4)可看出,不同压缩层级所对应的虚实融合方法不一样:第一层压缩取视频背景影像数据、前景运动目标子图影像数据、前景运动目标时空位置数据之和与原始视频数据进行比较;第二层压缩取前景运动目标子图影像数据、前景运动目标时空位置数据之和与原始视频数据进行比较;第三层压缩仅取前景运动目标时空位置数据之和与原始视频数据进行比较.

3 实验分析与讨论

使用上文所述的压缩效果评价模型,编程进行定量对比实验.实验的硬件参数如下:CPU: Intel(R) Core(TM) i5-2400 CPU @ 3.10 Hz; 内存: 2.99 GB; 编程环境: Visual Studio 2010; 编程语言: C#; 使用开源库: EmguCV. 实验数据为一段时长约 155 s, 帧率 25 fps, 共 3 875 帧的模拟监控视频,在本实验中每 20 帧对视频背景模型进行一次更新.

3.1 实验结果

视频压缩所获压缩数据的累计时长数据量 $M_{tol,o}$ 增长状况如图 2 所示,单位时长的数据量 $M_{uni,o}$ 变化状况如图 3 所示;总时长压缩比倒数 $K_{tol,l}$ 变化情况如图 4 所示,单位时长压缩比倒数 $K_{uni,l}$ 变化情况如图 5 所示.

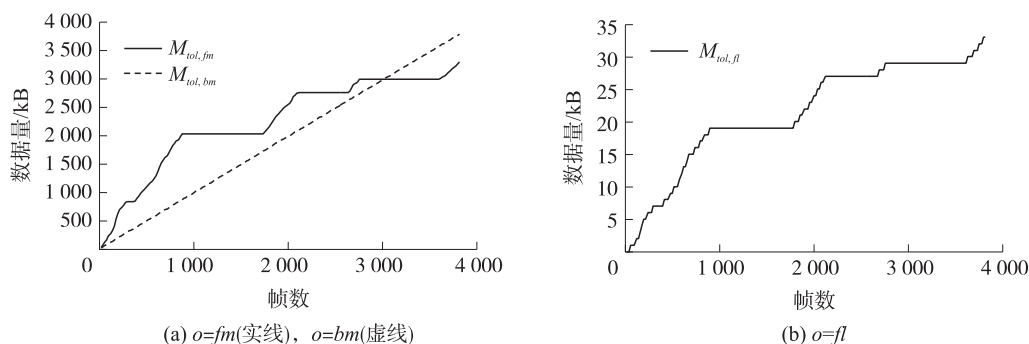
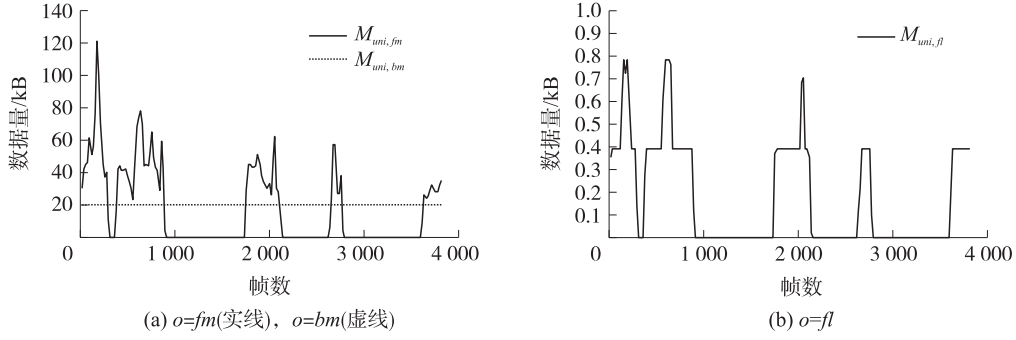
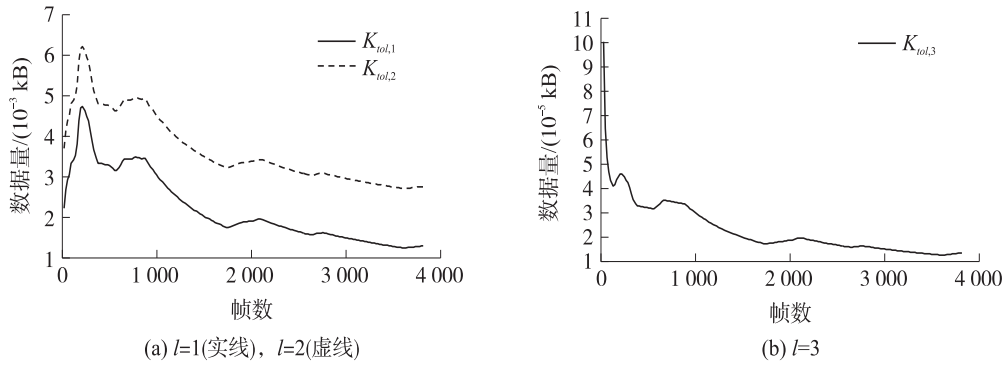
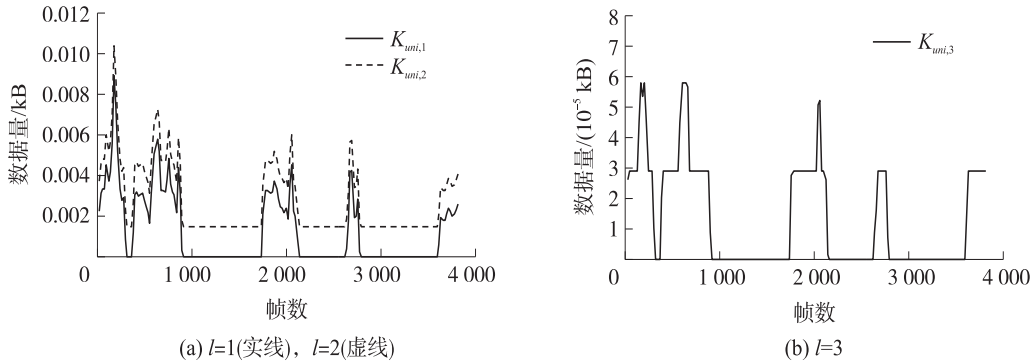


图2 $t=tol$ 时, o 取不同值, $M_{tol,o}$ 的变化情况

Fig.2 $t=tol$ and o in different value, the changes of $M_{tol,o}$

图3 $t=uni$ 时, o 取不同值, $M_{uni,o}$ 的变化情况Fig.3 $t=uni$ and o in different value, the changes of $M_{uni,o}$ 图4 $t=tol$ 时, l 取不同值, $K_{tol,l}$ 的变化情况Fig.4 $t=tol$ and l in different value, the changes of $K_{tol,l}$ 图5 $t=uni$ 时, l 取不同值, $K_{uni,l}$ 的变化情况Fig.5 $t=uni$ and l in different value, the changes of $K_{uni,l}$

从图 2 可看出, $t=tol$ 时, $M_{tol,fm}$ 与 $M_{tol,fl}$ 随前景运动目标的出现与消失而增长或停滞, 而 $M_{tol,bm}$ 增长稳定. 从图 3 可看出, $t=uni$ 时, 在没有前景运动目标出现时, $M_{uni,fm} = M_{uni,fl} = 0$; 在出现前景运动目标时, $M_{uni,fm}$ 与 $M_{uni,fl}$ 出现了较为显著的峰值. 综合上述实验结果, 可见 $M_{t,fl}$ 相比于 $M_{t,fm}$ 和 $M_{t,bm}$ 均存在量级的差别. 从图 4 可看出, $t=tol$ 时, $K_{tol,l}$ 随视频帧数的增长而逐渐趋于平稳. 从图 5 可看出, $t=uni$ 时, $K_{uni,l}$ 的大小变化与视频中前景运动目标个数有关.

3.2 实验分析

将实验结果图 2(a)与图 2(b)、图 3(a)与图 3(b)分别进行对比可发现: 不论时间尺度因子 t 如何取值, 只要满足对 $\forall M_{t,o}$, $M_{t,o} \neq 0$ 成立, 则 $M_{t,fl} < 0.01M_{t,fm}$ 与 $M_{t,fl} < 0.01M_{t,bm}$ 总是成立的. 这一情况说明虚实融合视频压缩方法所获得的不同类型压缩数据存在数据量量级的差别.

将实验结果图 4(a)与图 4(b)、图 5(a)与图 5(b)分别进行对比可发现: 对 $\forall t$ ($t \in T$, $T = \{uni, tol\}$), 当处理的视频帧数相同时, $0.01 > K_{t,1} > K_{t,2} > K_{t,3}$ 总是成立的, 这一情况说明压缩层级因子的变化对压缩

效率的影响较大:对象投影式相比于图像投影式,所需使用视频量可压缩至小于后者的 1%; $K_{i,1}>K_{i,2}>(K_{i,3}\times 100)$ 总是成立的,这一情况说明前景抽象式相较于前后景独立投影式与前景投影式,所需使用视频量可压缩至小于后者的 1%。综合以上分析得出如下结论:不同对象投影式虚实融合方法对视频压缩数据的需求存在量级差别,具体而言,前景抽象式的视频传输存储能力高于前后景独立投影式与前景投影式。

将实验结果图 4(a)与图 5(a)、图 4(b)与图 5(b)分别进行对比可发现:对 $\forall l(l \in L, L=\{1,2,3\})$,当处理的视频帧数相同,且当前视频帧没有前景运动目标出现时, $K_{uni,2}=K_{uni,3}=0$ 且 $K_{iol,l} \geq K_{uni,l}$ 总是成立的;而当前视频帧有前景运动目标出现时, $K_{iol,l} \leq K_{uni,l}$ 总是成立的。从以上分析得出如下结论:在虚实融合视频压缩方法中,时间尺度因子对压缩效率的影响受运动前景目标的出现情况制约:对于没有前景运动目标出现的段落,单位时长的视频压缩效率高于累积时长;对于前景运动目标出现的段落,累积时长的视频压缩效率高于单位时长。

4 结语

本文通过分析压缩数据层级结构,选取相应的评价因子,构建压缩效果评价模型,实现了对虚实融合视频压缩方法的分析与评价。实验结果表明该方法能有效缩减视频信息的数据量,且压缩效率受视频时长范围、虚实融合表达方式等因素影响。考虑到视频数据的复杂性,如较多前景运动目标、光照变化、遮挡等不利于运动目标检测跟踪的情况,对虚实融合视频压缩方法的进一步研究需要通过拥有上述特征的监控视频进行实验分析实现。

[参考文献]

- [1] MILOSAVLJEVIĆ A, RANČIĆ D, DIMITRIJEVIĆ A, et al. Integration of GIS and video surveillance[J]. International journal of geographical information science, 2016, 30(10): 1-19.
- [2] WANG Y, KRUM D M, COELHO E M, et al. Contextualized videos: combining videos with environment models to support situational understanding[J]. IEEE transactions on visualization and computer graphics, 2007, 13(6): 1 568-1 575.
- [3] de HAAN G, SCHEUER J, de VRIES R, et al. Egocentric navigation for video surveillance in 3d virtual environments[C]// 3D User Interfaces 2009, Lafayette, Louisiana, USA. USA: IEEE, 2009: 103-110.
- [4] CHEN K W, LEE P J, HUNG Y P. Egocentric view transition for video monitoring in a distributed camera network[C]// The 16th International Conference on Multimedia Modeling, Chongqing, China. Germany: Springer Berlin Heidelberg, 2011: 171-181.
- [5] WANG Y, BOWMAN D A. Effects of navigation design on contextualized video interfaces[C]// 3D User Interfaces 2011, Singapore. USA: IEEE, 2011: 27-34.
- [6] LEWIS P, FOTHERINGHAM S, WINSTANLEY A. Spatial video and GIS[J]. International journal of geographical information science, 2011, 25(5): 697-716.
- [7] WANG X. Intelligent multi-camera video surveillance: a review[J]. Pattern recognition letters, 2013, 34(1): 3-19.
- [8] YANG Y, CHANG M C, TU P, et al. Seeing as it happens: real time 3D video event visualization[C]// International Conference on Image Processing 2015, Quebec City, Canada. USA: IEEE, 2015: 2 875-2 879.
- [9] BAKLOUTI M, CHAMFRAULT M, BOUFARGUINE M, et al. Virtu4D: a dynamic audio-video virtual representation for surveillance systems[C]// 3rd International Conference on Signals, Circuits and Systems, Medenine, Tunisia. USA: IEEE, 2009: 1-6.
- [10] GIRGENSOHN A, KIMBER D, VAUGHAN J, et al. DOTS: support for effective video surveillance[C]// Proceedings of the 15th International Conference on Multimedia, Augsburg, Germany. Germany: ACM Multimedia, 2007: 423-432.
- [11] TAKEHARA T, NAKASHIMA Y, NITTA N, et al. Digital diorama: sensing-based real-world visualization[C]// International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, Dortmund, Germany. Germany: Springer Berlin Heidelberg, 2010: 663-672.

[责任编辑:丁 蓉]