http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2019.0300

面向 Web 科学可视化的临近空间数据视频化方法

谭剑[1,2], 王圣华[3,*], 郭长顺[2]

(1. 北京邮电大学 网络系统与网络文化北京市重点实验室, 北京 100012;
2. 中国科学院遥感与数字地球研究所, 北京 100094; 3. 北京信息科技大学, 北京 100090)

摘 要: 临近空间在科学、经济和军事等领域具有重要价值,高效便捷地进行临近空间数据 Web 科学可视化和信息提取是临近空间数据应用不断深入和扩展的必然需求,但是,已有的临近空间数据类型繁杂和数据量巨大,成为 Web 传输和实时可视化的关键瓶颈。从数据组织方法角度展开了研究,根据临近空间数据的特点,结合视频压缩原理,提出了将临近空间数据进行分解、插值和色彩空间变换处理形成图像后,选取适宜视频压缩编码方法将图像编码为视频。实验结果表明:所提方法可以实现临近空间数据在 Web 环境下的高效组织,在保障可视化质量的前提下通过降低数据网络传输量实现了临近空间数据的实时可视化需求。研究结果可以直接解决临近空间数据 Web 实时可视化问题,为临近空间科学研究和知识发现提供可视化理论和技术支持,并为类似海量数据 Web 科学可视化提供方法借鉴。

关 键 词: Web 科学可视化; 临近空间数据; 组织方法; 视频压缩; 流传输; 粒子映射
中图分类号: V221+.3; TB553
文献标志码: A **文章编号**: 1001-5965(2020)04-0712-12

临近空间是指海拔高度在 20~100 km 的空间,处于航空与航天的结合部,具有独特的优势和战略价值,有近空间、亚轨道、空天过渡区等别称,包括了平流层、中间层和电离层的部分区域[1-2]。临近空间数据的主要元素包括温度、密度、风场、压力、臭氧含量、电子密度、钠离子密度、流星通量和气辉辐射等[3]。临近空间是临近空间飞行器的飞行区以及各种航天器的通过区,作为从天空到太空的过渡区域,在科学、经济和军事等领域都具有重要的价值[4-6]。

一方面,通过遥感探测和模式模拟等方式已经获取了多源、多维、海量的临近空间数据,如电离层电子密度、大气风场和大气温度等。这些海量的多学科数据资源存在着隐藏的丰富的价值信息,蕴含了尚未发现的科学知识[7]。另一方面,科学计算可视化是将科学计算或观测结果的抽象数据转换为图形与图像在屏幕上进行显示并进行交互处理的理论、方法和技术,它被认为是研究复杂、大规模以及多维数据的有效工具[8-9]。传统的 Web 科学可视化方法将临近空间数据通常表示为数值数据表格或者二维图像,不足以形象直观地表现出临近空间数据的时空分布和变化特征,导致人们从这些结果中获取和感知大量时空信息的能力相对较弱。在这种情况下,需要使用三维图像等视觉产品中表示信息的方法来帮助研究人员理解信息,从而更好地进行科学研究[10]。

随着 HTML5 和 WebGL 技术的出现,以及一些基础数字地球平台的 Web 化,使得将临近空间

收稿日期: 2019-06-17; **录用日期**: 2019-09-22; **网络出版时间**: 2019-10-28 11:40
网络出版地址: kns.cnki.net/kcms/detail/11.2625.V.20191028.0933.001.html
基金项目: 北京市社会科学基金(18YTC030); 北京市属高校高水平教师队伍建设支持计划青年拔尖人才培育计划项目(CIT&TCD201704061); 中国科学院先导 A 项目(XDA19000800)
* **通信作者**. E-mail: 54315266@qq.com

引用格式: *谭剑, 王圣华, 郭长顺. 面向 Web 科学可视化的临近空间数据视频化方法[J]. 北京航空航天大学学报, 2020, 46(4): 712-723. TAN J, WANG S H, GUO C S. Video formatting method of near-space data for Web scientific visualization[J]. Journal of Beijing University of Aeronautics and Astronautics, 2020, 46(4): 712-723 (in Chinese).*

http://bhxb.buaa.edu.cn

数据通过一定的方式集成到 Web 数字地球平台中提供高效高质和方便统一的科学可视化已成为可能。此外,基于 Web 的数字地球平台与临近空间数据圈层体系一致,具有使用方便,可视化效果好等优点,在 Web 数字地球平台上进行临近空间数据的可视化便于科研工作者和公众通过网络及时访问和浏览形象的可视化成果,这在一定程度上能够促进临近空间科学规律的发现和科学知识的广泛传播。

但是,已有的临近空间数据类型繁杂和数据量巨大,成为 Web 传输和实时可视化的关键瓶颈。本文正是以临近空间数据 Web 科学可视化为目标,以开源 Web 数字地球引擎 Cesium 为平台,利用最新的 HTML5 和 WebGL 技术,根据临近空间数据的信息表达原理和结构化特点,研究利用数据编码压缩中的视频压缩编码的方式对数据进行组织。并通过这一新的数据组织方式实验实现了典型临近空间标量场数据和矢量场数据的时空多维动态可视化。

1　临近空间数据属性与 Web 科学可视化需求分析

临近空间数据是临近空间科学的产物,这类数据通常包含多种属性,主要包括:数据类型、数据来源、数据网格类型、数据格式及数据量等。

1.1　数据类型

数据类型是对领域数据的一种划分方式,在临近空间中根据数据所表达的空间特征可以将数据类型划分为标量场数据和矢量场数据。

标量数据是指可以用一个不依赖于坐标系的数字表征其性质的量,例如温度、密度等,该类数据只有大小没有方向[11]。

矢量数据是指需要用不依赖于坐标系的数字及方向表征其性质的量,例如速度、涡度等,该类数据不仅有大小而且具有方向特征。

张量是以矢量为基础扩展得到,其中 m 阶张量由 3^m 个标量组合而成,每个标量中有 n 个元素,因此二阶张量可以用一个 3×3 的矩阵 $\boldsymbol{X}_{ij}$ 来表示。

$$\boldsymbol{X}_{ij}=\begin{bmatrix}x_{11} & x_{12} & x_{13}\\ x_{21} & x_{22} & x_{23}\\ x_{31} & x_{32} & x_{33}\end{bmatrix}$$

当 $m=0$ 时,张量是一个标量;当 $n=1$ 时,张量是一个矢量。或者可以认为,张量是标量或者矢量的组合。

按照标量数据和矢量数据的定义,临近空间数据中的大气温度、电离层电子密度等是标量场数据,其只有大小没有方向特征;临近空间数据中的大气风场是矢量场数据,其不仅有大小,而且具有经向风和纬向风等方向特征。显然不同类型临近空间数据的科学可视化形式不同,本文将在后续研究中进行深入探讨。

1.2　数据来源

临近空间数据主要来源于原位、遥感探测和大气模式模拟。

遥感探测分为地基遥感探测、空基探测和卫星遥感探测。地基遥感探测主要利用雷达系统进行探测,该方法可以探测临近空间大气风场(纬向风和经向风)、大气温度和电离层电子密度等数据。空基探测主要利用火箭、探空气球等进行探测,其可以获取不同高度以及不同时间的临近空间大气温度和大气风场(纬向风和经向风)等数据。卫星遥感探测主要是利用卫星进行探测,其能够提供全球范围或者近于全球范围内的临近空间大气温度、大气风场(纬向风和经向风)和电离层电子密度等数据[2,12]。

除遥感探测以外,模式模拟也是临近空间数据获取的主要来源。目前被广泛使用的大气模式主要有标准大气模式、参考大气模式和数值模式 3 类,这些大气模式可以提供不同条件下的大气温度、大气风场和电子密度等数据[3,13]。不同来源的临近空间数据,理论上只要它们代表相同的临近空间信息,在 Web 科学可视化方面的需求就是一致的。

1.3　数据网格类型

原始的临近空间探测数据由于采样方式、设备误差和采样环境动态变化等原因,几乎不可能保持均匀的空间和时间采样。在实际使用中,研究者通常根据原始探测数据、误差校正方法和临近空间各类型数据模式融合生产不同数据网格上的产品数据,这也是本文后续数据处理的主要数据源(仍有少部分原始数据)。

在临近空间数据中,不同的数据网格类型就会产生不同的数据结构,其在可视化中起到了关键的作用。通过遥感探测和大气模式得到的临近空间数据通常都是离散的,因此要进行临近空间数据的可视化需要着重研究数据网格是如何构造的。

根据数据之间的连接关系,可以将临近空间中的离散数据分为 3 类。第 1 类是结构化网格数据,该类数据在逻辑上能够组织成三维数组。第

http://bhxb.buaa.edu.cn

2 类是非结构化网格数据,该类数据拓扑结构复杂,在逻辑上无规则可循[14]。第 3 类是混合网格数据,该类数据是通过对结构化网格数据以及非结构化网格数据组合得到。一般在实际应用中用到的临近空间数据大多为结构化网格数据,该类数据主要包括以下几种类型:

1)均匀网格结构化数据。如图 1(a)所示,该类数据可以由 3 维矩阵表示,在 x、y、z 三个方向上每个网格单元的长度相等。

2)规则网格结构化数据。如图 1(b)所示,该类数据是由长方体的体元构成,不同行、不同列的六面体长宽高可能不相同,但在同一坐标轴方向上长度是相同的。

3)矩形网格结构化数据。如图 1(c)所示,该类数据中的每个体元都是规则的长方体,在 x、y、z 三个方向上网格点间距可以相等也可以不相等。

4)不规则网格结构化数据。如图 1(d)所示,该类数据虽然也能够用三维数组来表示,但其在空间位置的分布上没有任何规律。

规则数据场是由均匀网格或规则网格构成的结构化数据。经过插值处理后的临近空间大气温度、电离电子总量、大气风场等数据是规则网格结构化数据,这些数据都为规则数据场。

显然,网格空间结构的密度决定了 Web 科学可视化的质量和数据传输要求,网格空间结构的规则情况决定了 Web 科学可视化插值显示的难度,也影响数据传输要求。

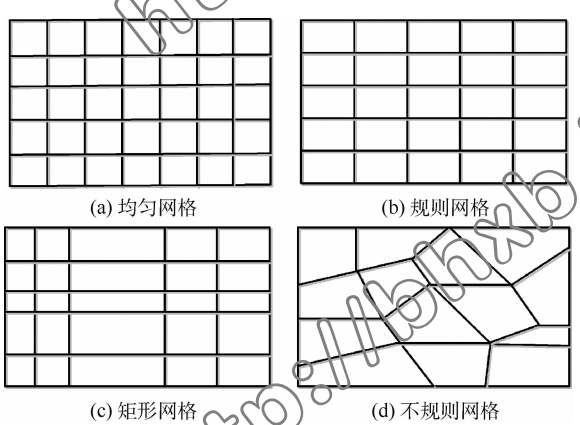


图 1 不同类型的结构化网格

Fig. 1 Different types of structured grids

1.4 数据格式和数据量

数据格式是对临近空间数据结构的封装,一方面由于临近空间长时序数据量非常大,另一方面也由于临近空间各种类数据的空间特征和网格结构均有不同,数据格式的基本要求包括可定制元数据和数据压缩。常见的临近空间数据格式包括 NetCDF、DAT 和 HDF 等,最典型的两种临近空间数据载体格式 NetCDF、DAT 的文件的数据特征如下所示。

1)NetCDF 格式文件

NetCDF 是一种网络通用数据格式,最初是用来在空间与地球科学团体中存储与交换数据。由于其具有灵活性,能够传输海量的面向阵列的数据,因此其被广泛用来存储大气科学及海洋等领域的数据,例如温度、风速、电离层电子密度和压力等。NetCDF 数据的特点主要包括以下几个方面[15]:

① 自描述性。NetCDF 格式文件包含自身的描述信息,如数据获取的时间。

② 高可用性。能够高效的对 NetCDF 数据进行访问,可以直接对需要的数据进行读取,而不必读取整个数据集。

③ 平台无关性。NetCDF 格式文件可以在不同网络平台间进行传输和共享。可以利用多种语言对 NetCDF 格式文件进行读取,例如 MATLAB、Java、Fortran、IDL 等。

NetCDF 格式文件中的数据以数组形式存储。例如:可以用一维数组来存储某个位置处随时间变化的临近空间大气温度,用二维数组来存储某个区域内在某个时刻的临近空间大气温度,用一系列二维数据来存储三维(3D)数据(如某个区域内随时间变化的临近空间大气温度)或四维(4D)数据(如某个区域范围内随时间和高度变化的临近空间大气温度)。图 2 形象地表示出了 NetCDF 格式文件如何以二维数组来实现对 4D 数据的储存。

NetCDF 格式文件主要由维度(dimensions)、变量(variables)和属性(attributes)三部分组成,各个组成部分的特点主要为:

① 维度。用来确定 NetCDF 格式文件的变量维度信息,如经度、纬度、高度等。

② 变量。表示相同类型的多维数组,变量可以为温度、电子密度、风速等。

③ 属性。用来对变量或 NetCDF 格式文件进行描述的信息,如变量的单位等。

2)DAT 格式文件

DAT 不是一种标准文件格式,没有标准化的数据属性结构信息,通常采用 ASCII 或二进制编码格式。在读取该文件中的数据信息时,需要事先对文件中的内容及数据特点有充分的了解。而在临近空间数据存储中,一般作为 NetCDF 格式的补充格式进行存储,其数据结构符合原始 NetCDF 的定义。

http://bhxb.buaa.edu.cn

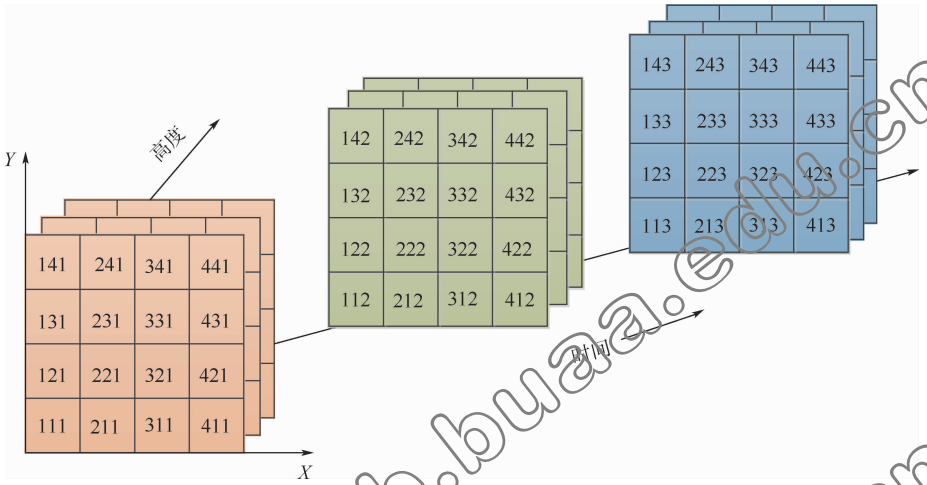


图 2　四维数据(某个区域内随时间和高度变化的数据)

Fig. 2　Four-dimensional data (variation of data with time and height within an area)

3)数据量分析

以典型的临近空间大气温度数据为例,其常见数据格式为 NetCDF,因此大气温度数据集的总数据量随三维空间网格边长和时间分辨率的变化可以由图 3 表示。

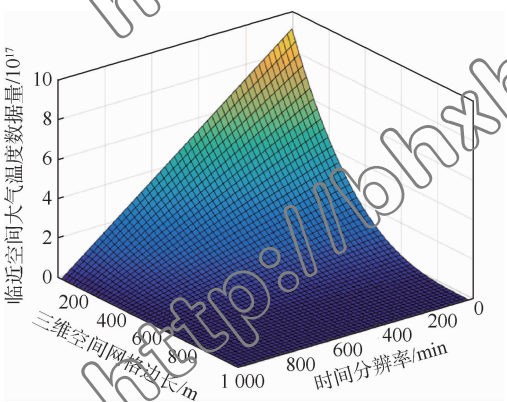


图 3　大气温度数据量随三维空间网格边长及时间分辨率变化

Fig. 3　Variation of atmospheric temperature data amount with length of three-dimensional space grid and time resolution

1.5　临近空间数据 Web 科学可视化的数据组织需求

总体来说,临近空间数据的数据量与空间网格密度和时间分辨率有直接关系,尤其对于三维采样或模式计算来说,随着采样密度的增加(采样网格边长的减小),数据量会呈几何级数增加,而且对更高时间分辨率临近空间数据的需求,比如由年周期变为日周期甚至小时周期,数据量也会呈几十甚至上百倍的增加,如图 3 所示。而不断提高临近空间数据时空分辨率也是中国航空航天事业蓬勃发展的现实需求。

通过以上对临近空间数据特点的分析可以看出,现有的临近空间数据,一方面具有数据量大、数据异构的特点;另一方面,其数据存储方式只是为了长久保持数据,而非为了进行 Web 科学可视化,表现为时序数据耦合性强、数据文件大小随空间网格密度和时间分辨率呈几何级数增长。

但是,稳定的临近空间数据 Web 科学可视化直接受到网络传输速率的限制,而且在单位时间内传输的数据,如果不能形成有效的可视化结果,用户就无法进行实时的交互和浏览。

因此,有必要根据已有的临近空间数据结构进行面向 Web 科学可视化的临近空间数据组织方法改造或改进。

2　基于视频压缩编码的临近空间数据组织方法

在网络环境中进行海量数据传输和可视化方面,不同领域的空间数据压缩技术有很多,主要包括基于有损压缩的 JPEG 图像编码、基于无损压缩的 PNG 图像编码、GZip 压缩编码等[10,16-18]。

比较经典的是传统的视频数据,因为 Web 科学可视化的需要,衍化出多种流媒体数据格式,如 MP4、WebM 等[19-21],其核心思想是利用视频图像时序上的前后的相关性,只增量传输时序帧之间的差异,并且根据视频图像上大量存在的同色图斑,将图形数据进一步压缩以实现在低带宽或不稳定带宽环境下的高效传输和可视化。

同样的,临近空间数据反映了临近空间中的电子密度、温度、风场等连续空间现象的时序变化,时序数据之间具有显著相关性,数据空间网格中存在着大量邻接雷同数据,尤其在网格密度较

http://bhxb.buaa.edu.cn

高时尤为明显。所以根据临近空间数据和传统视频数据在原理上的相似性,采用流媒体视频压缩编码方式进行临近空间数据的重组,将是解决临近空间数据 Web 科学可视化具有潜力的方法。

因此,本文根据典型临近空间标量场数据(临近空间大气温度、电离层电子总量等)和矢量场数据(临近空间大气风场)的特点,研究利用视频压缩编码的方式对数据进行组织,以实现临近空间数据在网络环境中的高效传输。

2.1 典型临近空间数据的可视化结构

四维临近空间大气温度数据为例,本文利用多维数据分解算法[22]将大气温度数据先按照时间维度划分为一系列空间结构相一致的三维体数据,然后将每一时相的数据按高度分层,处理为一系列结构相一致的二维网格数据。如图 4 所示,每个网格点包含某个空间位置的大气温度信息。

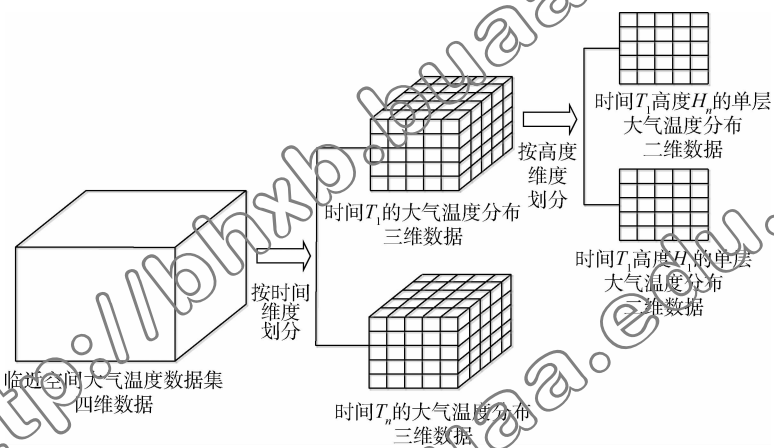


图 4 四维临近空间大气温度数据转换为一系列二维网格示意图

Fig. 4 Schematic diagram of four-dimensional near-space atmospheric temperature data being converted to a series of two-dimensional grids

2.2 临近空间数据规则化插值

由于本文中获取的少数原始临近空间数据不是规则经纬度网格点,因此需要对这些原始数据进行插值处理,以得到规则的网格点数据。本文主要采用双线性插值算法对临近空间数据在经纬度方向上进行插值处理以获得规则的网格点数据,该方法可以在较大程度上确保数据精度[2,5,23-24]。双线性插值算法的原理如图 5 所示。

图 5 中:A_{11}、A_{12}、A_{21}、A_{22}为 4 个已知的原始网格点,D_1、D_2 为中间待求网格点,W 为最终要得到的网格点。(x_i, y_i)、(x_i, y_{i+1})、(x_{i+1}, y_i)、(x_{i+1}, y_{i+1})分别是 A_{11}、A_{12}、A_{21}、A_{22} 4 个网格点的坐标信息,$f(A_{11})$、$f(A_{12})$、$f(A_{21})$、$f(A_{22})$分别是 A_{11}、A_{12}、A_{21}、A_{22} 4 个网格点的属性值,$f(D_1)$和 $f(D_2)$分别是由单线性插值得到的 D_1 和 D_2 网格点的属性值,$f(W)$为 W 网格点的属性值。双线性插值算法分为 3 步:

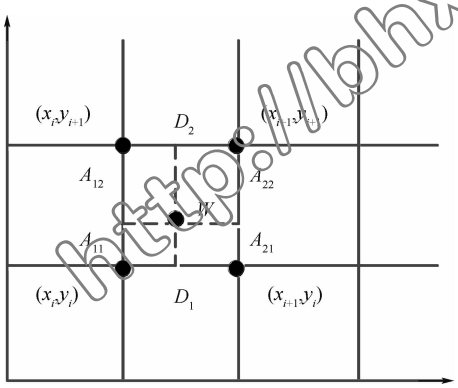


图 5 双线性插值算法原理示意图

Fig. 5 Schematic diagram of bilinear interpolation algorithm principle

1)求得 D_1 点的属性值。

$$f(D_1) = \frac{f(A_{21}) \cdot (x - x_i) + f(A_{11}) \cdot (x_{i+1} - x)}{x_{i+1} - x_i}$$

2)求得 D_2 点的属性值。

$$f(D_2) = \frac{f(A_{22}) \cdot (x - x_i) + f(A_{12}) \cdot (x_{i+1} - x)}{x_{i+1} - x_i}$$

3)求得 W 点的属性值。

$$f(W) = \frac{f(D_2) \cdot (y - y_i) + f(D_1) \cdot (y_{i+1} - y)}{y_{i+1} - y_i}$$

下面以临近空间大气温度、电离层电子总量等典型临近空间标量场数据以及临近空间大气风场典型矢量场数据为例对插值后的数据特点进行简要说明。

如图 6 所示,插值处理后临近空间时序大气温度数据特点为

1)具有经度、纬度、高度、时间四维特征。

2)在高度方向上,数据是按照等距离间隔垂直分层的。

http://bhxb.buaa.edu.cn

3）经纬度方向上,不同高度层上的网格点数据是规则的二维网格数据,即经度方向的间隔和纬度方向的间隔均相等。

4）数据属性值为临近空间大气温度值。

如图 7 所示,插值处理后临近空间时序二维电离层电子总量数据的特点为

1）具有经度、纬度、时间三维特征。

2）经纬度方向上,网格点数据是规则的二维网格数据,即经度方向的间隔和纬度方向的间隔均相等。

3）数据属性值为临近空间电离层电子总量。

如图 8 所示,插值处理后临近空间时序三维大气风场数据特点为

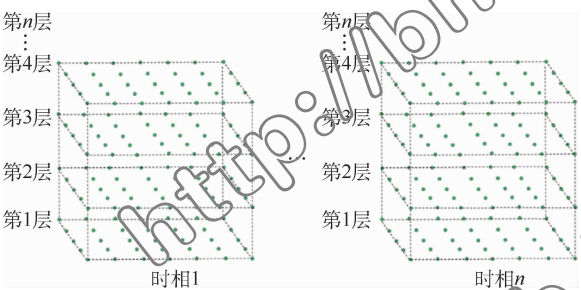


图 6　临近空间时序大气温度数据特点示意图

Fig. 6　Schematic diagram of temporary data characteristics of atmospheric temperature in near-space

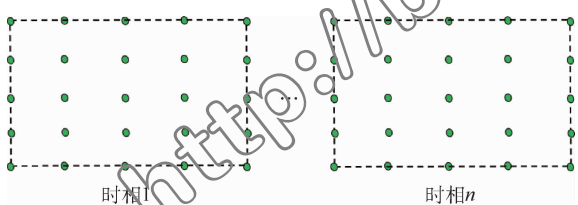


图 7　临近空间时序电离层电子总量数据特点示意图

Fig. 7　Schematic diagram of temporary data characteristics of ionization-layer electron total amount in near-space

1）具有经度、纬度、高度、时间四维特征。

2）在高度方向上,大气风场数据是按照等距离间隔垂直分层的。

3）经纬度方向上,不同高度层上的风场网格点数据是规则的二维网格数据,即经度方向的间隔和纬度方向的间隔均相等。

4）风场速度矢量包括经向风(u)和纬向风(v)两个分量。

2.3　临近空间 Web 科学可视化色彩空间变换

以临近空间大气温度数据为例,当原始大气温度数据被处理为规则二维网格数据后,需要将数据从标量空间转换到 RGB 空间,以便后期将数据编码为视频。本文针对每一高度层中的网格,取每个网格 4 个顶点数据的平均值作为该网格的大气温度值。然后,将每个网格的大气温度值通过式(1)归一化到[0,1]之间,以便于转换到 RGB 色彩空间:

$$D_{(i,j)} = (GE_{(i,j)} - GE_{min})/(GE_{max} - GE_{min}) \quad (1)$$

式中:$GE_{(i,j)}$为每个网格所代表的大气温度值;GE_{max}和GE_{min}分别为整个大气温度数据集的最大值和最小值;$D_{(i,j)}$为原始标量值。

将原始标量值 $D_{(i,j)}$乘以 255 后赋予 RGB 色彩空间的 R 色彩分量,G 色彩分量和 B 色彩分量都赋予 0,因此转换后的 RGB 色彩分量可以用式(2)表示:

$$(R,G,B) = (D_{(i,j)} \times 255,0,0) \quad (2)$$

借助 MATLAB 软件,设计程序将不同高度层的 R、G、B 色彩分量合成为 RGB 图像。图 9 为色彩分量合成后不同高度层的 RGB 图像。

最后,将同一时相、不同高度层的 RGB 图像合成为一张代表该时相临近空间大气温度数据的多层图像,如图 10 所示。

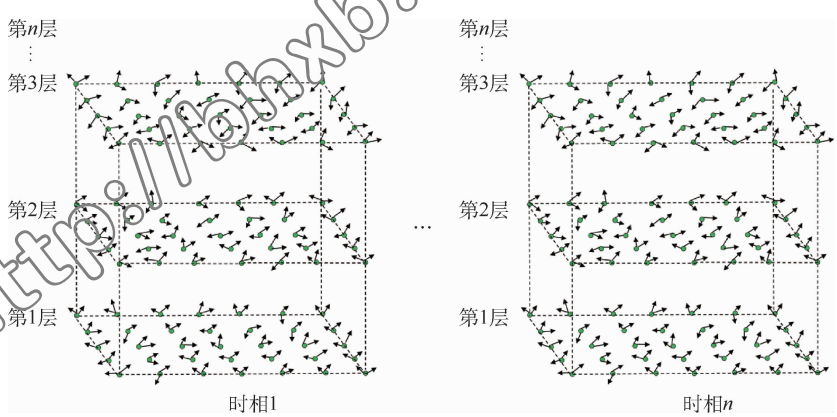
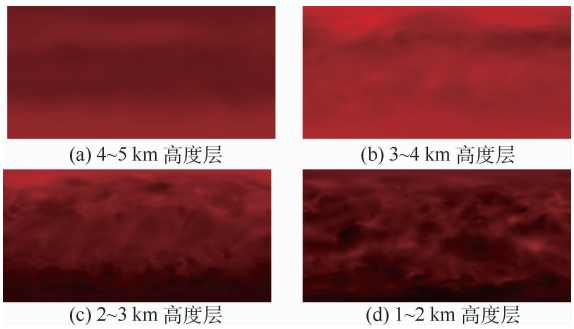


图 8　临近空间时序大气风场数据特点示意图

Fig. 8　Schematic diagram of temporary data characteristics of atmospheric wind field in near-space

http://bhxb.buaa.edu.cn



(a) 4~5 km 高度层 (b) 3~4 km 高度层

(c) 2~3 km 高度层 (d) 1~2 km 高度层

图 9 代表不同高度层临近空间大气温度的 RGB 图像

Fig. 9 RGB image representing atmospheric temperature of near-space at different height levels



图 10 代表某时相临近空间大气温度的多层 RGB 图像

Fig. 10 Multi-layer RGB image representing atmospheric temperature of near-space in a certain time phase

2.4 基于视频编码的临近空间数据压缩

当不同时相的临近空间大气温度原始数据被处理为 RGB 图像后,利用视频编码方法将一系列 RGB 图像压缩为视频数据以实现数据在网络环境中的高效传输和多时相可视化。视频文件的编解码时间和视频质量受到多种因素的影响,其中,最大的影响因素是视频编解码器和色度子采样模型[25]。目前在 Web 环境下应用比较广泛的视频编解码器主要有 H.264、VP8、VP9 等[21]。在相同的参数设置下,与 H.264 编解码器相比,VP8 提供更快的解码速度,但它编码得到的视频质量略低于 H.264;VP9 与 H.264 编解码器编码得到的视频质量相当,也具有相似的解码速度,但 VP9 编解码器可以支持更多的色度子采样模型[22]。不同编解码器的性能比较如表 1 所示。

在视频编码过程中,还应选择色度子采样模型,将在 RGB 色彩空间中编码的大气温度数据转换到由亮度(Y)和色度(UV)组成的YUV色彩空间。在 YUV 模型中,全分辨率 YUV4:4:4 色度子采样模型相比其他 YUV 模型(如 YUV 4:2:0 和 YUV 4:2:2)等保存了更多的数据信息,使用该模型能得到更好的视频质量,如表 2 所示。

因此,综合考虑视频质量和解码效率两方面,本文选择的视频编解码器为 VP9,色度子采样模型为 YUV4:4:4。

表 1 不同视频编解码器性能比较

Table 1 Comparison of performance of different video codecs

视频编解码器	解码速度	编码质量	是否支持 YUV4:4:4
H.264	良	优	否
VP8	优	良	否
VP9	良	优	是

表 2 不同色度子采样模型信息保存量比较

Table 2 Comparison of information storage amount of different color sub-sampling models

色度子采样模型	保存信息量
YUV4:2:0	少
YUV4:2:2	中
YUV4:4:4	多

3 实验条件与数据基础

1)硬件环境:处理器 Intel Core i7-6700HQ,内存 16 GB,硬盘 500 GB,显卡 Nvidia GeForce GTX 960M,显存 4 GB。

2)软件环境:Win10 64 位操作系统,数字地球引擎 Cesium,HTML、CSS、JavaScript、GLSL 等编程语言,开发环境 Sublime。测试所用客户端为 Google Chrome(72.0.3682.86)浏览器。用于视频编码的程序为 FFmpeg,其包含用于 VP9 视频编码的库 libvpx-vp9。

3)数据基础

本文中用到的临近空间数据是由中国科学院国家空间科学中心提供,该数据是由遥感探测数据和大气模式模拟得到的数据经过融合而成,主要包括大气温度、电离层电子总量等典型标量场数据以及大气风场等典型矢量场数据,其具体内容如表 3 所示。

表 3 典型临近空间数据

Table 3 Typical near-space data

数据属性	大气温度	电离层电子总量	大气风场
数据类型	四维标量	三维标量	四维矢量
数据格式	NetCDF	DAT	NetCDF
经度范围/(°)	-180~179.5	70~140	-180~179.5
纬度范围/(°)	-90~90	15~55	-90~90
高度范围/km	20~100	60~100	20~100
起始时刻	2016 年 12 月 21 日 2 时	2018 年 1 月 1 日 1 时	2016 年 12 月 21 日 2 时
结束时刻	2016 年 12 月 30 日 24 时	2018 年 7 月 23 日 24 时	2016 年 12 月 30 日 24 时
经度间隔/(°)	2.5	1	2.5
纬度间隔/(°)		1	
高度间隔/km	1		1
时间间隔/h	2	0.25	2
数据量/GB	1.2	0.8	1.5

http://bhxb.buaa.edu.cn

4　数据预处理

本文中用到的典型临近空间标量场数据大气温度数据以及典型临近空间矢量场数据大气风场数据在纬度方向上不是等间隔的,因此本文利用双线性插值方法对这两种数据进行插值处理使其变为规则数据,插值后数据内容如表 4 所示。

值得注意的是,插值后数据量大幅增加主要来自无效的图像化数据。比如没有利用的 GB 和 Alpha 通道,由于视频图像格式的定义而产生的占位。这一增加是各向同值的,在各类视频压缩算法中,都会完全压缩,不会对视频编码后的数据量产生很大影响。

此外,由 1.4 节分析可知,本研究中临近空间数据预处理时间整体与原始数据量、原始网格不规则程度、原始数据时间、空间缺损程度成正比。应具体数据具体分析,故实验用数据的预处理时间参考价值不大,不在此列出。

当典型临近空间数据被插值处理为规则网格数据后,利用色彩空间变换变换方法将其处理为 RGB 图像,然后利用 VP9 视频编解码器以及色度子采样模型 YUV4:4:4 将 RGB 图像处理为视频数据,为保障数据质量 VP9 编码按照最大冗余即图像质量不损失压缩比进行设置,以作为后期方法验证的无损数据源。

表 4　插值处理后典型临近空间数据

Table 4　Typical near-space data after interpolation

数据属性	大气温度	电离层电子总量	大气风场
数据类型	四维标量	三维标量	四维矢量
数据格式	NetCDF	DAT	NetCDF
经度范围/(°)	-180 ~ 180	70 ~ 140	-180 ~ 180
纬度范围/(°)	-90 ~ 90	15 ~ 55	-90 ~ 90
高度范围/km	20 ~ 100	60 ~ 100	20 ~ 100
起始时刻	2016 年 12 月 21 日 2 时	2018 年 1 月 1 日 1 时	2016 年 12 月 21 日 2 时
结束时刻	2016 年 12 月 30 日 24 时	2018 年 7 月 23 日 24 时	2016 年 12 月 30 日 24 时
经度间隔/(°)	1	1	1
纬度间隔/(°)	1	1	1
高度间隔/km	1	1	1
时间间隔/h	2	0.25	2
数据量/GB	2.68	0.8	3.6

5　基于视频压缩编码的临近空间数据传输效率对比

为了验证视频流在数据传输效率上的优势,在局域网环境下,对基于视频压缩编码方法得到的数据与 NetCDF 格式的数据、DAT 格式的数据以及用 ZIP 封包的分包压缩数据进行了相同可视化质量下数据传输量大小的对比实验。相同可视化质量具体定义为单帧图像分辨率 1 080 P,并且四类数据可视化的实际空间分辨率一致、比例尺一致。此外,为避免传输层协议干扰,统一采用 HTTP 协议进行传输,为避免缓存效率和传输丢包率波动,取连续 25 帧数据传输量的平均值。实验结果如图 11 所示。

从图 11 可以看出,在相同可视化质量、相同网络带宽下,针对临近空间大气温度、电离层电子总量、大气风场等数据,NetCDF 格式需要传输的数据量最大,二进制化的 DAT 格式需要传输的数据量较 NetCDF 格式的数据量缩减接近 50%,对 DAT 格式进行的分块 ZIP 压缩能有效去除冗余空间,缩减率接近 80%,视频数据需要传输的数据量最小,减小幅度超过 80%,因此将 NetCDF 格式及 DAT 格式的数据通过视频编码方法压缩为视频数据可以大大降低需要传输的数据量。即在同等可视化质量、同等网络带宽下,经过视频压缩编码得到的数据在网络环境中传输的时间最短,能够满足可视化时的实时数据需求。

在前端可视化算法方面,采用了基于粒子追踪的多层次移动流线法[26-27]表达全球各区域临近空间大气风场的高度层变化特征。采用了基于 GPU 加速的临近空间标量数据光线投射方法[11,28]对临近空间大气温度、电离层电子总量等典型临近空间标量场数据的标量场数据的实时动态进行了可视化渲染。

图 12 为基于粒子追踪的临近空间大气风场数据多层次移动流线法可视化,数据时刻为 2016 年12 月 21 日 2 时,这种方法能够有效表达出全球各区域临近空间大气风场的高度层变化特征,对临近空间大气风场高度层可视化分析具有很好的应用价值,但是对原始数据密度和时序

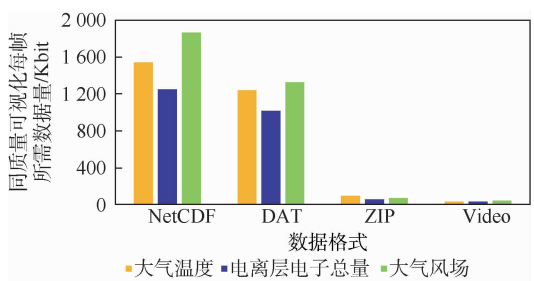


图 11　同质量可视化不同数据格式下的数据传输量对比

Fig. 11　Comparison of data transmission volume under different data formats with the same quality visualization

http://bhxb.buaa.edu.cn

变化要求高。从图 12 中可以看出不同高度层大气风场的变化特征,本文的视频压缩编码方法使得临近空间数据可以支持数字地球的多比例尺缩放和多角度实时交互可视化。

图 13 为利用 2016 年 12 月 22 日 10 时的临近空间大气温度数据得到的不同高度层(从地面起每 50 m 为 1 层)的光线投射体绘制结果。图 14(a)~(d)分别为利用 2016 年 12 月 22 日 16 时、2016 年 12 月 24 日 16 时、2016 年 12 月 26 日16 时、2016 年 12 月 28 日 16 时的临近空间大气温度数据得到的高度层为 50~60 层的光线投射体绘制结果。

值得指出的是,ZIP 方式与视频压缩编码方法的数据传输量比较接近,而且本文方法的临近空间数据预处理时间较长,类似 ZIP 等压缩编码过程。但是视频压缩编码格式不同于 ZIP 等压缩编码,VP9 视频解码是内嵌于 HTML5 标准中的[21]。在传输到 Web 客户端时,在例如 Chrome 等浏览器中这一解压操作已经优化为系统底层机器代码,并且解压图像直接进入显卡显存[20],并不存在 ZIP 等临时文件解压、文件移动和加载过程。如图 15 所示,对客户端接收数据后到图像进入显存阶段(每帧可视化所需数据客户端加载时间)的统计,相比 ZIP 格式,VP9 视频解码加载对可视化的迟滞可以忽略不计,在解压可视化方面具有天然的效率优势。

实验结果表明,将 NetCDF 格式及 DAT 格式的数据通过视频压缩编码方法处理为视频数据可以大大降低需要传输的数据量。即在同等可视化质量、同等网络带宽下,经过视频压缩编码得到的数据在网络环境中传输的时间最短,能够满足 Web 科学可视化的实时数据需求。

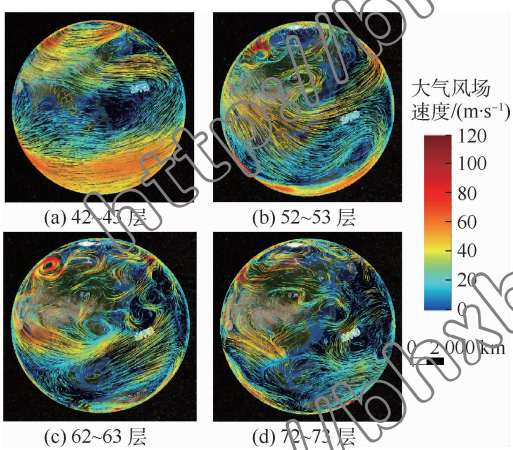


图 12 不同高度层临近空间大气风场绘制结果(流线数量:128×128)

Fig. 12 Rendering results of atmospheric wind field in near-space at different height levels (number of streamlines: 128×128)

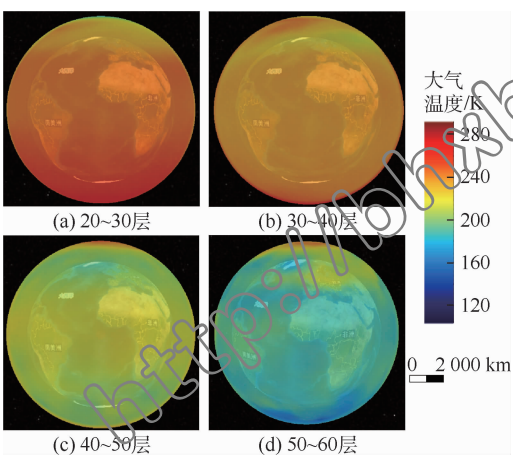


图 13 同一时相不同高度层大气温度光线投射体绘制结果

Fig. 13 Atmospheric temperature ray projection volume rendering results at different height levels in the same time phase

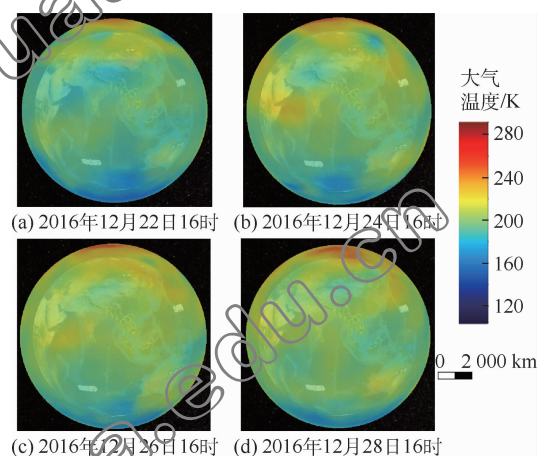


图 14 同一高度层不同时相大气温度光线投射体绘制结果

Fig. 14 Atmospheric temperature ray projection volume rendering results in different time phases at the same height level

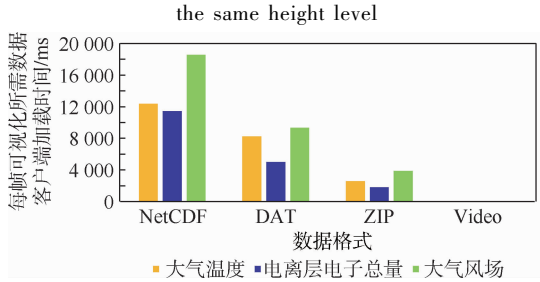


图 15 每帧可视化所需数据客户端加载时间对比

Fig. 15 Comparison of visualization data loading time for each frame in client

6 结 论

1)本文针对临近空间数据的特点和数据在

网络环境中传输面临的问题,研究了数据编码压缩中的视频压缩编码方法,在将数据进行分解、插值、色彩空间变换处理为图像后,选取了合适的视频压缩编码方法将图像编码为视频,实现了临近空间数据的有效组织。在保证相同可视化质量的前提下,经过视频压缩编码后需要传输的数据量和原始 NetCDF 等格式的数据量相比大大减小,总体减小幅度超过 80%,因此所提基于视频压缩编码的临近空间数据组织方法可以大大降低数据的网络传输量,能够实现 Web 环境中数据的实时可视化需求。

2)所提方法的临近空间数据预处理后的视频编码数据修改困难。但是这一缺点在 Web 科学可视化中影响较小,一方面,临近空间数据是由专业机构采集或处理发布,一般科研用户和行业用户既不需要也没有对应的处理工具在 Web 界面中对其进行修改。

3)本文利用基于视频压缩编码的临近空间数据组织方法实现了典型临近空间标量场数据(大气温度等)和典型临近空间矢量场数据(大气风场)的高效组织和实时可视化,取得了令人满意的效果。

临近空间数据中还有许多其他格式和类型的数据。因此在未来的研究工作中,还需要对其他临近空间数据的特点进行研究,以建立全面的临近空间数据增量式数据组织方法。

参考文献(References)

[1] 肖存英,胡雄,杨钧烽,等.临近空间 38°N 大气密度特性及建模技术[J].北京航空航天大学学报,2017,43(9):1757-1765.
XIAO C Y,HU X,YANG J F,et al. Atmospheric density characteristics and modeling techniques of near space 38°N[J]. Journal of Beijing University of Aeronautics and Astronautics,2017,43(9):1757-1765(in Chinese).

[2] 杨钧烽.中纬度临近空间大气风场变化特性研究[D].北京:中国科学院国家空间科学中心,2016.
YANG J F. Study on the variation characteristics of atmospheric wind field in the mid-latitude adjacent space[D]. Beijing:National Space Science Center,Chinese Academy of Sciences,2016(in Chinese).

[3] 刘涛.临近空间大气风场统计预报方法研究[D].北京:中国科学院国家空间科学中心,2017.
LIU T. Research on statistical forecast method for atmospheric wind field in adjacent space [D]. Beijing:National Space Science Center,Chinese Academy of Sciences,2017(in Chinese).

[4] 邓潘.地基激光雷达对临近空间大气的探测研究[D].合肥:中国科学技术大学,2018.
DENG P. Ground-based lidar detection of adjacent space atmosphere [D]. Hefei:University of Science and Technology of China,2018(in Chinese).

[5] 杨钧烽,肖存英,胡雄,等.临近空间风切变特性及其对飞行器的影响[J].北京航空航天大学学报,2019,45(1):57-65.
YANG J F,XIAO C Y,HU X,et al. Wind shear characteristics of adjacent space and its influence on aircraft[J]. Journal of Beijing University of Aeronautics and Astronautics,2019,45(1):57-65(in Chinese).

[6] 郭文杰.532 nm 瑞利多普勒激光雷达若干关键技术和大气重力波观测研究[D].北京:中国科学院国家空间科学中心,2016.
GUO W J. Several key technologies of 532 nm Rayleigh Doppler lidar and atmospheric gravity wave observation[D]. Beijing:National Space Science Center,Chinese Academy of Sciences,2016(in Chinese).

[7] GOODCHILD M F,GUO H D,ANNONI A,et al. Next-generation digital earth[J]. Proceedings of the National Academy of Sciences of the United States of America,2008,109(28):11088-11094.

[8] LACROUTE P H,LEVOY M. Fast volume rendering using a shear-warp factorization of the viewing transformation[M]. New York:ACM,1994:451-458.

[9] LEVOY M. Display of surfaces from volume data[J]. IEEE Computer Graphics and Applications,1988,8(3):29-37.

[10] SUN X J,SHEN S H,LEPTOUKH G G,et al. Development of a Web-based visualization platform for climate research using Google Earth[J]. Computers & Geosciences,2012,47:160-168.

[11] 吴玲达,于荣欢,瞿师.大规模三维标量场并行可视化技术综述[J].系统仿真学报,2012,24(1):12-16.
WU L D,YU R H,QU S. Review of large-scale 3D scalar field parallel visualization technology[J]. Journal of System Simulation,2012,24(1):12-16(in Chinese).

[12] 谢衍新,吴小成,胡雄,等.临近空间全球温度场三维变分同化[J].红外与激光工程,2017,46(8):55-60.
XIE Y X,WU X C,HU X,et al. Three-dimensional variational assimilation of global temperature field in adjacent space[J]. Infrared and Laser Engineering,2017,46(8):55-60(in Chinese).

[13] 冯健,甄卫民,刘钝.电离层环境对地空信息系统的影响效应[J].装备环境工程,2017,14(7):7-11.
FENG J,ZHAI W M,LIU D. Effect of ionospheric environment on ground-air information system[J]. Equipment Environmental Engineering,2017,14(7):7-11(in Chinese).

[14] 王文涛.非结构化网格分片线性矢量场高效可视化技术研究[D].长沙:国防科学技术大学,2016.
WANG W T. Research on efficient visualization technology of unstructured grid segmentation linear vector field[D]. Changsha:National University of Defense Technology,2016(in Chinese).

[15] 曹柱.基于 GPU 的海洋水体环境多维可视化技术研究[D].青岛:中国海洋大学,2015.
CAO Z. GPU-based multi-dimensional visualization technology of marine water environment [D]. Qingdao:Ocean University of China,2015(in Chinese).

http://bhxb.buaa.edu.cn

[16] SUN M, LI J, YANG C W, et al. A Web-based geovisual analytical system for climate studies[J]. Future Internet, 2012, 4(4): 1069-1085.

[17] SARTHOU A, MAS S, JACQUIN M, et al. EarthScape, a multipurpose interactive 3D globe viewer for hybrid data visualization and analysis[J]. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2015, 40(3): 487-493.

[18] LI W W, SONG M M, ZHOU B, et al. Performance improvement techniques for geospatial Web services in a cyberinfrastructure environment-A case study with a disaster management portal[J]. Computers, Environment and Urban Systems, 2015, 54: 314-325.

[19] 张凌燕. 流媒体服务的 OoE 计算方法研究[D]. 北京: 北京邮电大学, 2018.
ZHANG L Y. Research on OoE calculation method of streaming media service[D]. Beijing: Beijing University of Posts and Telecommunications, 2018(in Chinese).

[20] 张志明, 柯卫. 基于 HTML5 的视频通信云服务应用技术研究[J]. 电信科学, 2012, 28(10): 31-37.
ZHANG Z M, KE W. Research on application technology of video communication cloud service based on HTML5[J]. Telecommunications Science, 2012, 28(10): 31-37(in Chinese).

[21] 顾春莲. HTML5 中的音频及视频元素对互联网的影响[J]. 河北省科学院学报, 2011, 28(3): 106-108.
GU C L. The influence of audio and video elements in HTML5 on the Internet[J]. Journal of the Hebei Academy of Sciences, 2011, 28(3): 106-108(in Chinese).

[22] LI W W, WANG S Z. Polarglobe: A Web-wide virtual globe system for visualizing multidimensional, time-varying, big climate data[J]. International Journal of Geographical Information Science, 2017, 31(8): 1562-1582.

[23] 刘鹤丹. 科学计算可视化中三维等值曲面构建方法研究[D]. 沈阳: 东北大学, 2015.
LIU H D. Research on construction method of 3D equivalent surface in scientific computing visualization[D]. Shenyang: Northeastern University, 2015(in Chinese).

[24] 支星, 徐海明. 3 种再分析资料的高空温度与中国探空温度资料的对比: 年平均特征[J]. 大气科学学报, 2013, 36(1): 77-87.
ZHI X, XU H M. Comparison of high altitude temperature and Chinese sounding temperature data of three kinds of reanalysis data: Annual average characteristics[J]. Chinese Journal of Atmospheric Sciences, 2013, 36(1): 77-87(in Chinese).

[25] POYNTON C. Chroma subsampling notation[EB/OL]. San Francisco: Morgan Kaufmann, 2003 (2018-01-24)[2019-06-01]. http://poynton.ca/PDFs/Chroma_subsampling_notation.pdf.

[26] 汪小荣, 陈毅松, 汪国平. 数字海洋系统的向量场可视化[J]. 计算机辅助设计与图形学学报, 2016, 28(12): 2114-2119.
WANG S R, CHEN Y S, WANG G P. Vector field visualization of digital ocean system[J]. Journal of Computer-Aided Design & Computer Graphics, 2016, 28(12): 2114-2119(in Chinese).

[27] 李骞, 范茵, 王吉奎. 基于粒子追踪的风场可视化方法[J]. 解放军理工大学学报(自然科学版), 2005, 6(1): 89-91.
LI Q, FAN Y, WANG J K. Visualization method of wind field based on particle tracking[J]. Journal of PLA University of Science and Technology(Natural Science Edition), 2005, 6(1): 89-91(in Chinese).

[28] 陈静, 邹成, 黄吴蒙, 等. 面向虚拟地球的三维气象场可视化方法[J]. 武汉大学学报(信息科学版), 2016, 41(12): 1563-1569.
CHEN J, ZOU C, HUANG W M, et al. Visualization of 3D meteorological field for virtual earth[J]. Journal of Wuhan University(Information Science Edition), 2016, 41(12): 1563-1569(in Chinese).

作者简介:

谭剑 男, 博士, 副教授。主要研究方向: 虚拟仿真和智能设计。

王圣华 女, 博士, 副教授。主要研究方向: 数字媒体和文化遗产活化。

郭长顺 男, 硕士。主要研究方向: 三维空间信息系统。

http://bhxb.buaa.edu.cn

Video formatting method of near-space data for Web scientific visualization

TAN Jian[1,2], WANG Shenghua[3,*], GUO Changshun[2]

(1. Beijing Key Laboratory of Network System and Network Culture, Beijing University of Posts and Telecommunications, Beijing 100012, China; 2. Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China; 3. Beijing University of Science and Technology Information, Beijing 100090, China)

Abstract: As a transitional region from sky to space, near-space has important value in the fields of science, economy and military. With the development of remote sensing technology and the continuous development of mode simulation, the massive near-space data of higher spatial and temporal resolution is increasing. On the other hand, efficient and convenient Web scientific visualization and information extraction for near-space data are inevitable requirements for the continuous deepening and expansion of near-space data applications. However, the existing data types are cumbersome and the amount of data is huge in near-space, which becomes the key bottleneck of Web transmission and real-time visualization. This paper focuses on the visualization of near-space data on the Web digital globe, and studies it from the point of data organization methods. According to the characteristics of near-space data, combined with the principle of video compression, we propose to decompose, interpolate and color space transform the near-space data to form an image, and then select the appropriate video compression coding method to encode the image into video. Experiment results show that this method can be implemented efficiently. The video organization of spatial data in the Web environment realizes the real-time visualization of near-space data by reducing the amount of data network transmission while ensuring the quality of visualization. The research results in this paper can directly solve the real-time visualization problem of near-space data on Web, provide visual theory and technical support for the near-space science research and knowledge discovery, and provide a reference for the Web scientific visualization on similar massive data.

Keywords: Web scientific visualization; near-space data; organization method; video compression; stream transmission; particle mapping

http://bhxb.buaa.edu.cn

Received: 2019-06-17; **Accepted**: 2019-09-22; **Published online**: 2019-10-28 11:40

URL: kns.cnki.net/kcms/detail/11.2625.V.20191028.0933.001.html

Foundation items: Beijing Social Science Foundation Project (18YTC030); Beijing High-level Teacher Team Building Support Project for Young Talents in Local Universities (CIT&TCD201704061); Chinese Academy of Sciences Pilot A Project (XDA19000800)

* **Corresponding author**. E-mail: 54315266@qq.com