

doi: 10.6046/gtzyyg.2016.02.04

引用格式: 祝佳. Landsat8 卫星遥感数据预处理方法[J]. 国土资源遥感, 2016, 28(2): 21-27. (Zhu J. Analysis of Landsat8 satellite remote sensing data preprocessing[J]. Remote Sensing for Land and Resources, 2016, 28(2): 21-27.)

Landsat8 卫星遥感数据预处理方法

祝 佳

(中国西南电子技术研究所, 成都 610036)

摘要: Landsat 系列卫星是由美国航空航天局和美国地质调查局共同管理的资源遥感系列卫星, 40 多 a 来为地球遥感探测活动提供了大量清晰而稳定的图像数据。卫星遥感数据预处理是获取优质遥感基础图像的第一步, 对后续各级卫星遥感产品的质量有着很重要的影响。针对 Landsat8 卫星原始数据, 对卫星下传所采用的空间数据传输协议和数据传输格式进行了详细的解析, 分析了原始数据从解同步、数据帧解析、任务数据包解析、图像数据获取直到生成 0 级图像产品的步骤; 特别针对存在无损数据压缩的陆地成像仪 (operational land imager, OLI) 数据, 讨论了基于空间数据系统咨询委员会 (consultative committee for space data systems, CCSDS) 相关标准进行无损数据解压缩处理的方法和过程。经数据预处理得到的 Landsat8 卫星 0 级图像产品, 可为 Landsat8 卫星数据应用提供优质的基础图像。

关键词: Landsat8; 遥感卫星; 数据预处理; 数据解压缩

中图法分类号: TP 751.1 **文献标志码:** A **文章编号:** 1001-070X(2016)02-0021-07

0 引言

2013 年 2 月, 美国航空航天局 (national aeronautics and space administration, NASA) 在美国加利福尼亚成功发射了 Landsat 家族的第 8 颗陆地卫星——Landsat8。Landsat8 的发射是“陆地卫星数据连续性任务” (Landsat data continuity mission, LDCM) 的一个重要组成部分, 该任务是纪念 Landsat 系列卫星发射 40 a 而特别制订的。Landsat8 卫星搭载有陆地成像仪 (operational land imager, OLI) 和热红外传感器 (thermal infrared sensor, TIRS)。OLI 被动感应地表反射的太阳辐射和散发的热辐射, 有 9 个波段, 覆盖了从可见光到红外谱段的不同波长范围。与 Landsat7 卫星搭载的传感器 ETM⁺ 相比^[1], OLI 增加了 1 个 C/A 波段 (0.433 ~ 0.453 μm) 和 1 个短波红外波段 (Band 9, 1.360 ~ 1.390 μm)。C/A 波段主要用于海岸带观测; 短波红外波段能反映水汽强吸收特征, 可用于卷云检测。TIRS 是有史以来最先进、性能最好的热红外传感器, 能收集地球热量流失, 可用于了解所观测地带的水分消耗、特别是干旱地区的水分消耗^[2-3]。

本文将针对 Landsat8 卫星的原始数据, 首先对原始数据的解同步、解包等解析步骤进行分析和讨

论; 在提取出任务数据 (mission data) 之后, 特别针对 OLI 数据进行无损数据解压缩处理的方法推演; 最终完成 Landsat8 卫星 0 级图像产品的生成, 为 Landsat8 卫星数据应用提供优质的基础图像。

1 原始数据解析

原始数据解析是获取影像信息的第一步。其过程是利用 Landsat8 卫星下传的原始未加工 (raw) 数据, 依次进行同步、帧解析、数据包解析和任务数据拼接, 最终提取出影像数据和辅助参数、生成 0 级图像产品的过程。

充分了解空间数据传输协议和传输格式是进行原始数据解析的关键。空间数据传输协议是卫星下传数据的传输控制协议, 原始数据同步、数据帧和数据包的解析、影像数据和辅助参数的提取, 均需依据该协议进行; 而数据传输格式描述的是任务数据在传输时所采用的组合方式, 对整景图像的数据提取和对应辅助、校正参数的确定有着重要作用, 直接关系到任务数据的完整获取和准确校正。对于 Landsat8 卫星系统, 其空间数据传输协议和传输格式都是参照空间数据系统咨询委员会 (consultative committee for space data systems, CCSDS) 的遥控遥测相关建议标准^[4-6]制定的。

1.1 数据帧解析

Landsat8 卫星数据以位 (bit) 流的方式,首先通过虚拟通道连接到控制子层,传递到协议函数;随后数据流进入虚拟通道访问子层,通过添加头尾来控制端到端的数据流,从而形成虚拟通道数据单元 (virtual channel data unit, VCDU);接着在进入物理通道层之前添加一个帧同步头,形成通道访问数据单元 (channel access data unit, CADU);最后由物理通道层完成数据流的调制和传输 (图 1)。

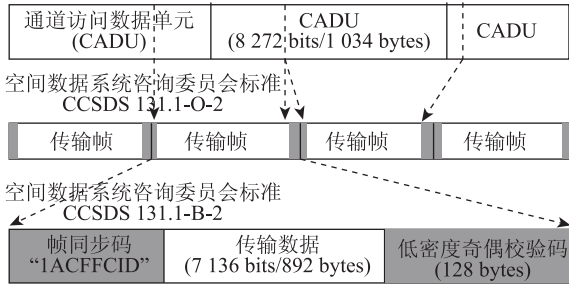


图 1 传输帧与通道访问数据单元关系图

Fig.1 Relation schema between transfer frame and CADU

当地面站接收卫星传输来的原始遥感数据时,需要依据上述的传输控制协议对原始数据位流逐层进行解析,以获取任务图像数据和相应的辅助参考数据。

遥感数据通过数据传输帧进行传输,如图 1 所示,传输帧尺寸为 1 024 字节 (bytes),包括 4 字节的帧同步码 (0x1ACFFC1D)、892 字节的传输数据和 128 字节的低密度奇偶校验码 (low density parity check code, LDPC) 编码信息。通过剥离帧同步码和 LDPC 编码信息,即可从下行数据传输帧中提取出所传输的数据信息;将这些单个 892 字节长的传输数据段拼接起来,又可以组成 1 个新的数据流,即通道访问数据单元 (CADU),亦称为数据帧,每帧尺寸为 1 034 字节,其中包括了 4 字节的帧同步码 (0x352EF853)、6 字节的数据帧主导头和 1 024 字节的数据域。其帧结构如图 2 所示。

帧同步码 0x352EF853 4字节		数据域1 024字节	
数据帧 主导头 6字节	多路协议 数据单元头 M_PDU Header 2字节	M_PDU数据包区域1 022字节	
		第k个包 的后段	第k+1个 数据包
		...	第m个 数据包
			第m+1个 包前段

图 2 通道访问数据单元结构图

Fig.2 Structural chart of CADU

通过帧同步码可同步提取出单个的 CADU 数据帧,从解析数据帧的主导头 (transfer frame primary header) 开始来解析该数据帧。主导头尺寸为 6 字节 (图 3)。

数据帧主导头6字节					
主信道标识		虚拟信道标识		标志域	
版本号	航天器标识	虚拟信道标识	虚拟信道帧计数	回放标志	虚拟信道计数用法标志
2字节		3字节		1字节	

图 3 数据帧主导头结构图

Fig.3 Structural chart of transfer frame primary header

主导头包含以下内容:

1) 版本号。第 0 位和第 1 位被用作数据帧的版本号,设为“01”。

2) 航天器标识。第 2 位到第 9 位用作航天器标识,Landsat8 是“11111010”。

3) 虚拟信道标识 (virtual channel identifier, VCID)。第 10 位到第 15 位为虚拟信道标识,用来标明数据产生的通道。对于 Landsat8,标识“000000”表示来自 OLI 的实时数据;标识“000101”表示来自 TIRS 的实时数据。

4) 虚拟信道帧计数。第 16 到 39 位用作虚拟信道帧计数,计数从 0 开始,每个数据帧计数加 1,直到计数值达到 16 777 215 后重置为 0;设备重启时计数也重置为 0。

5) 回放标志。第 40 位用作回放标志位,“0”表示实时数据,“1”表示回放数据。

6) 虚拟信道帧计数用法标志。第 41 位用作标识是否使用了虚拟信道循环帧计数方式,对于 Landsat8,并未使用循环计数,因而该位设为“0”。

7) 保留空间。第 42 到 43 位为保留位,设为“00”。

8) 虚拟信道帧计数循环标志。第 44 到 47 位用于标识循环帧计数,由于 Landsat8 未使用,因而全部设为“0”。

数据流中紧跟着数据帧主导头的是 1 024 字节长的数据域,其中包含了多个多路协议数据单元 (multi_protocol data unit, M_PDU) 数据包。由于数据包区域长度仅为 1 022 字节,因而数据包区域内的数据包并非都是完整的,可能包含了上一个数据包中未传输完的 1 个数据包的后半段,也可能包含了无法完整传输的 1 个数据包的前半段。

为确保从数据包区域中准确提取出完整的数据包信息,在数据域的头 2 个字节设置了多路协议数据单元头 (M_PDU header),其中的前 5 位作为保留位,全设为“0”;后 11 位为首包头指针 (first header pointer),表示本数据帧内所包含的第一个完整 M_PDU 数据包的包头指针 (模 2 048),其指针值表示第一个数据包的偏移量,即从 M_PDU 数据包区域的起始位到第一个完整 M_PDU 数据包之间相隔的

字节数。此外,特地用 0x7FF 表示本数据帧内没有数据包的包头。

1.2 M_PDU 数据包解析

通过首包头指针定位到完整的 M_PDU 数据包后,就需要对数据包进行相关解析。M_PDU 数据包由 5 个主要部分组成:

1) 数据包主导头 (packet primary header)。长度为 6 字节,其中包含了本数据包的相关标识信息(图 4),对准确解析数据包中的遥感数据有着重要的作用。

M_PDU数据包主导头6字节						
版本号	数据包标识			数据包顺序控制		数据包长度
	数据包类型	副导头标识	应用标识 APID	包顺序标识	数据包序号	
	2字节			2字节		2字节

图 4 数据包主导头结构图

Fig. 4 Structural chart of packet primary header

数据包主导头包含以下内容:

①版本号。“000”。

②数据包类型。“0”。

③副导头标识。“1”表示数据包,带有副导头;“0”表示空闲包,没有副导头。

④应用标识(application process identification, APID)。用 11 位数据表示本数据包所携带数据的类型,是解析遥感数据的关键参数。部分常用 APID 的描述见表 1,其中 APID 为 256—268 表示进行压缩处理后的 OLI 成像仪各波段数据,768—780 表示未压缩的 OLI 成像仪各波段数据。

⑤包顺序标识。“11”。

⑥数据包序号。用 14 位数据连续记录数据包的序号,计数从 0 开始,根据 APID 每个有效数据包计数加 1(空闲数据包不计数),直到计数值达到 16 383后重置为 0,设备重启时计数也重置为 0。

表 1 应用标识描述

Tab. 1 Description of APID

APID	描述	APID	描述	APID	描述
2	OLI 数据行头	264	OLI PAN2_EVEN 波段数据(压缩)	775	OLI PAN2_ODD 波段数据
3	OLI 数据 CRC 校验	265	OLI SWIR2 波段数据(压缩)	776	OLI PAN2_EVEN 波段数据
4	OLI 数据帧头	266	OLI SWIR1 波段数据(压缩)	777	OLI SWIR2 波段数据
5	辅助数据	267	OLI CIRRUS 波段数据(压缩)	778	OLI SWIR1 波段数据
256	OLI PAN1_ODD 波段数据(压缩)	268	OLI DARK 波段数据(压缩)	779	OLI CIRRUS 波段数据
257	OLI PAN1_EVEN 波段数据(压缩)	768	OLI PAN1_ODD 波段数据	780	OLI DARK 波段数据
258	OLI BLUE 波段数据(压缩)	769	OLI PAN1_EVEN 波段数据	1026	TIRS 数据行头
259	OLI C/A 波段数据(压缩)	770	OLI BLUE 波段数据	1027	TIRS 数据 CRC 校验
260	OLI NIR 波段数据(压缩)	771	OLI C/A 波段数据	1792	TIRS 10.8um 波段数据
261	OLI RED 波段数据(压缩)	772	OLI NIR 波段数据	1793	TIRS DARK 波段数据
262	OLI GREEN 波段数据(压缩)	773	OLI RED 波段数据	1794	TIRS 12um 波段数据
263	OLI PAN2_ODD 波段数据(压缩)	774	OLI GREEN 波段数据	2047	空闲数据

⑦数据包长度。2 字节,其数值为 L,单位为字节,即

$$L = (\text{数据包副导头} + \text{偏移位} + \text{任务数据头} + \text{任务数据域长度}) - 1 = \text{任务数据域长度} + 15. \quad (1)$$

任务数据域长度一般为变长,通过该参数确定本数据包的准确长度,以便对任务数据进行解析提取。

2) 数据包副导头(packet secondary header)。长度为 8 字节,主要包含一些数据包的辅助信息。对于 Landsat8 来说,大部分为默认值或保留位,同时也包含了数据域长度和数据源 ID 等参数,用于辅助确认数据包的相关参数信息。

3) 偏移位(offset)。偏移位长度为 4 字节。

4) 任务数据头(mission data header)。长度为 4 字节,用于描述随后的任务数据信息。主要包含 2 个参数:①第 5—15 位为应用标识 APID,此处的 APID 和数据包主导头中的 APID 应保持一致,用于确认任务数据类型;②第 16—31 位为任务数据域长度,与数据包主导头中的数据包长度值 L 相差 15

字节,用于确认任务数据的长度。

5) 任务数据域(mission data)。每个数据包的任務数据域长度不定,由任务数据头中的数据域长度参数确定。任务数据域中的数据就是所需要提取的遥感图像数据信息。

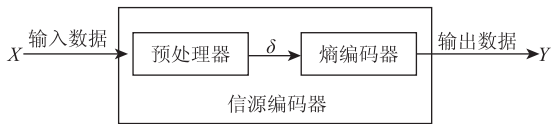
2 图像数据解压缩

Landsat8 的遥感数据主要包括 OLI 成像仪数据、TIRS 传感器数据和相关辅助数据。其中,对于 OLI 成像仪数据,Landsat8 在上传数据时部分进行了数据无损压缩处理,故对这部分数据在提取遥感图像数据时需要进行相应的解压缩处理。由于 Landsat8 只对 OLI 成像仪数据进行了压缩处理,因

而数据解压缩处理只针对 OLI 成像仪数据。

2.1 无损数据压缩

Landsat8 采用 CCSDS 121.0 - B 无损数据压缩建议标准对部分 OLI 成像仪数据进行压缩^[5]。在 CCSDS 121.0 - B - 2 无损数据压缩建议标准中,将 RICE 自适应编码算法作为进行无损数据压缩的核心算法^[4]。RICE 算法的核心思想是使用一系列的变长编码对样本数据进行相应的压缩处理,主要包括预处理器和熵编码器 2 个组成部分,其框架结构如图 5 所示。



(X 为原始数据输入值; δ 为预处理器输出值; Y 为压缩数据输出值)

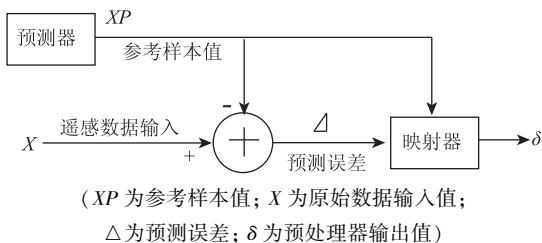
图 5 无损数据压缩结构框图

Fig. 5 Structure diagram of lossless data compression

图 5 中,预处理器的功能在于分析待处理数据样本之间的相关性,使用便于进行熵编码的数学符号来描述待处理的数据样本;而熵编码器则是对经预处理后的数据样本,根据其特点自适应地在一系列变长编码中选取最为合适的编码对样本进行压缩处理,最后在压缩数据前端附上对应的压缩编码标识符。

2.1.1 预处理器

预处理器的主要作用是对待压缩的原始数据进行分析处理,将其变换为有利于熵编码器进行高效率压缩的数据样本形式。由于所进行的数据压缩处理是无损的,因而预处理的过程也应该是可逆的过程。预处理器一般包括预测和映射 2 个步骤,其结构如图 6 所示。



(XP 为参考样本值; X 为原始数据输入值; Δ 为预测误差; δ 为预处理器输出值)

图 6 预处理器结构框图

Fig. 6 Structure diagram of preprocessor

图 6 中,预测器用于减少数据样本间的相关性,预测时要用到参考样本,以此为基础生成差值序列;在 Landsat8 的 OLI 成像仪数据中,每 1 024 行就有 1 行未进行压缩处理的图像数据作为参考样本。映射器主要完成在不改变预测误差统计特性的情况下对预测误差值进行映射变换,映射后得到的 n 位非负整数(与输入数据的 bit 数相同)应便于后续的熵

编码器进行有效处理。

Landsat8 所使用的预测编码映射变换公式为:

若 $XP < 2\ 048$,

$$\lfloor 2(XP - X) - 1 \rfloor \quad 0 \leq X < XP$$

$$\text{则 } \delta = \begin{cases} \lfloor 2(X - XP) \rfloor & XP \leq X \leq 2XP \\ \lfloor X \rfloor & 2XP < X \leq 4\ 095 \end{cases}; \quad (2)$$

若 $XP \geq 2\ 048$,

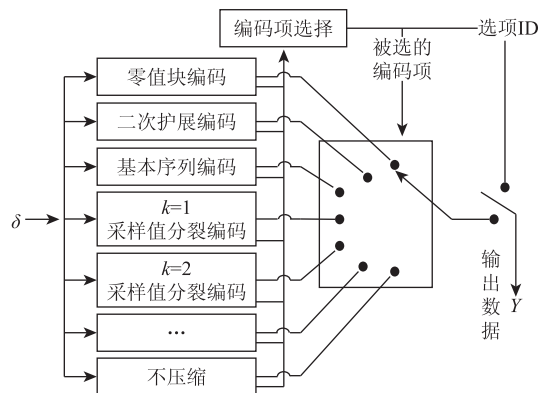
$$\lfloor 4\ 095 - X \rfloor \quad 0 \leq X < (2XP - 4\ 095)$$

$$\text{则 } \delta = \begin{cases} \lfloor 2(XP - X) - 1 \rfloor & (2XP - 4\ 095) \leq X < XP \\ \lfloor 2(X - XP) \rfloor & XP \leq X \leq 4\ 095 \end{cases}. \quad (3)$$

式中: δ 为预处理器输出值; X 为原始数据输入值; XP 为参考样本值; 原始数据和参考样本均为 12 bit 的图像像素值,取值范围为 0 ~ 4 095。

2.1.2 熵编码器

熵编码器模块是 RICE 算法的核心模块,包括 1 组变长编码器集合,其中包含 4 类编码模式:基本序列编码、采样值分裂编码、低熵值编码和无压缩编码。当输入经预处理的数据样本块时,熵编码器会根据数据样本的特点,自适应地选取 1 类最为有效的编码模式对其进行压缩编码;同时在压缩编码数据的前端会附上对应的压缩编码标识符,以便于编码模式的确认和解压缩的处理;最后将编码标识符和压缩编码一起输出。由于在进行熵编码时,每个数据样本可以根据其自身特点自适应地选取编码效率最高的编码模式,因而这种编码算法能有效地适应遥感图像信源统计特性的变化,可大大提高压缩率。熵编码器的结构框图见图 7。



(k 为采样值分裂编码选项参数)

图 7 熵编码器结构框图

Fig. 7 Structure diagram of adaptive entropy coder

熵编码器包括:

1) 基本序列(fundamental sequence, FS)编码。一种简单变长码,“1”为码字结束的信号,用前面“0”的个数代表所编码的信息。其本质是对出现频率越高的数据,分配越短的码字;而对出现频率越低的数据,则分配越长的码字。详细说明见表 2。

表 2 FS 编码规则
Tab. 2 Rule of FS coding

预处理后的数据 δ	基本序列码
0	1
1	01
2	001
...	...
2^{n-1}	000...001 (2^{n-1} 个 0)

2) 采样值分裂编码。为熵编码器中使用频率最多的一类编码模式。其中,分裂编码的选项参数用 k 来表示,对参数为 k 的采样值分裂编码的处理方法是:对 1 个含有 J 个数据样本的待编码块分裂出 k 个低位数据,余下的高位数据将分别依次进行基本序列编码;随后在已完成编码的高位数据流后依次附上分裂出去的低位数据。特别地, $k=0$ 是基本序列编码。RICE 算法中采样值分裂编码的数据压缩率大约在 1 bit 左右。

3) 低熵值编码。RICE 算法^[4]提供了 2 种低熵值编码选项,即二次扩展编码和零值块编码,主要是针对待处理的数据样本值非常小的情况。对于这类数据值很小的样本,采用低熵值编码模式可以获得很好的压缩效果,每个样本点的压缩比可低于 1.5 bits。其中,二次扩展编码先将数据样本块中相邻的 2 个数据合并成数据对,使用一定的方式将数据对转换为新的值;然后对该新值进行基本序列编码,这样一组码字就描述了 2 个数据样本值,大大提高了压缩效率。而零值块编码是针对 1 个或多个连续的待处理数据样本块均为 0 值的情况,对这些连续的待处理 0 值块的个数进行基本序列编码,个数的上限为 64;当连续 0 值块的个数大于 64 时,则分段重新进行个数的编码。

4) 无压缩编码。若上述 3 类编码模式都无法对待处理的数据样本块进行有效的压缩编码时,则不进行任何压缩处理,即采用无压缩编码。处理方式是在经过预处理的数据样本块前端直接附上无压缩编码的标识符后进行输出。

2.2 数据解压缩

解压缩是数据压缩编码的逆过程,与无损数据压缩流程对应。解压缩流程包括熵译码器和后处理器 2 个部分。在熵译码器中需要确定其各参数的配置,以确保熵译码器的工作模式与熵编码器相同。配置参数包括每个样本的位数、数据块的大小和所采用的编码方式。对于 Landsat8 的 OLI 数据,数据样本为 12 bits,任务数据域中的每个压缩数据包(coded data set, CDS)包含了压缩选项标识(option ID)和 16 个数据样本的压缩编码值(表 3)。

表 3 压缩选项对照
Tab. 3 Selected code option identification key

压缩选项标识	压缩编码值
0 值块编码	00000
二次扩展编码	00001
基本序列编码	0001
采样值分裂编码	
$k=1$	0010
$k=2$	0011
$k=3$	0100
$k=4$	0101
$k=5$	0110
$k=6$	0111
$k=7$	1000
$k=8$	1001
$k=9$	1010
$k=10$	1011
$k=11$	1100
$k=12$	1101
无压缩编码	1111

CDS 的结构如图 8 所示。

采样值K分裂编码

Option ID	16个样本值高(12-k)位的FS编码序列	16个样本值低k位的分裂值序列
-----------	-----------------------	-----------------

未压缩的编码

Option ID	未编码的16个样本值序列
-----------	--------------

零值块编码

Option ID	连续零值块数量的FS编码
-----------	--------------

二次扩展编码

Option ID	8个二次扩展样本值的FS编码序列
-----------	------------------

图 8 CDS 结构图

Fig. 8 Structure diagram of CDS

通过解析数据包开头的 4 位或 5 位码字(见表 3),确定所使用的压缩编码方式;然后针对不同的编码方式,解析后面的编码序列。熵编码是可逆的,其译码过程也不复杂。经过熵译码器译码后,数据被送往后处理器。后处理器是预处理器的逆过程,完成与预测和变换相反的功能。对于 Landsat8 数据,其后处理映射变换公式为:

若 $XP < 2\ 048$,

$$\begin{cases} \lfloor XP - (\delta' + 1)/2 \rfloor & 0 \leq \delta' \leq 2XP \text{ 且 } \delta' \text{ 为奇数} \\ X' = \begin{cases} XP + \delta'/2 & 0 \leq \delta' \leq 2XP \text{ 且 } \delta' \text{ 为偶数}; \\ \delta' & 2XP < \delta' \leq 4\ 095 \end{cases} \end{cases} \quad (4)$$

若 $XP \geq 2\ 048$,

$$\begin{cases} \lfloor XP - \delta' + 1 \rfloor / 2 & 0 \leq \delta' \leq (8\ 190 - 2XP) \\ & \text{且 } \delta' \text{ 为奇数} \\ X' = \begin{cases} XP + \delta'/2 & 0 \leq \delta' \leq (8\ 190 - 2XP) \\ \lfloor 4\ 095 - \delta' \rfloor & (8\ 190 - 2XP) < \delta' \leq 4\ 095 \end{cases} \end{cases} \quad (5)$$

式中: δ' 为经过译码后的输入值; X' 为解压缩后的

输出值; XP 为参考样本值; 输入值、输出值和参考样本值均为 12 位的图像像元值, 取值范围为 0 ~ 4 095。经过后处理器的映射后, 就能获得原始的遥感图像数据信息。

3 图像产品生成

在 Landsat8 图像中, 每个 M_PDU 数据包包含 1 行

OLI数据行														
行头	PAN1 ODD	PAN1 EVEN	BLUE	C/A	NIR	RED	GREEN	PAN2 ODD	PAN2 EVEN	SWIR2	SWIR1	CIRRUS	DARK	CRC
TIRS数据行														
行头	10.8 um Row				Dark Row				12 um Row				CRC	

图 9 OLI 和 TIRS 图像数据行结构图

Fig. 9 Structure diagram of image line for OLI and TIRS

根据上述的任务数据协议和数据压缩协议等, 预处理系统主要由以下几个部分组成:

1) 数据获取子系统。负责从下传设备中接收并获取原始的卫星数据文件, 并将其存储到原始数据库中, 以供后续的数据处理使用。

2) 原始数据解析子系统。从原始数据库中读取卫星下传的原始未加工 (raw) 数据进行解析, 根据空间数据传输协议对原始数据进行同步、数据帧解析等处理。

3) 任务数据解析子系统。根据数据传输协议和传输格式对任务数据包进行解析, 提取出任务影像数据帧; 对使用了无损压缩的数据进行解压缩处理, 恢复其原始图像数据, 并提取出对应的辅助参数和校正参数。

4) 图像数据处理子系统。生成各波段图像数据, 提供图像快视浏览。

遥感卫星数据预处理系统输入的是 Landsat8 卫星的原始数据文件, 其内容是二进制的数流; 系统的最终输出是 Landsat8 的 0 级卫星图像信息。软件系统集成开发环境是在 Windows XP SP3 操作系统下的 Microsoft Visual Studio 2010, 软件操作界面采用窗口交互方式。

4 结论

本文通过分析 Landsat8 卫星遥感数据预处理的方法, 梳理出 Landsat8 卫星图像从原始下传数据通过解同步、解包、图像数据获取、数据解压缩直到 0 级卫星图像产品生成的处理流程, 得到如下结论:

1) 按照空间数据系统咨询委员会标准建议的空间数据传输协议和传输格式, 可以逐层对原始数据进行解析, 经过同步、数据帧解析、M_PDU 数

据包解析, 能够从任务数据域中提取出图像数据和辅助参数。
2) OLI 成像仪的部分数据下传时经过了数据压缩, 由于采用的是 CCSDS 无损数据压缩标准, 通过逆向的解压缩处理可以获取原始图像数据。
3) 经过数据预处理过程能够得到 Landsat8 卫星的 0 级图像产品, 可为后续各级卫星遥感产品生成和遥感图像分析提供优质基础。

参考文献 (References):

[1] 王树根. Landsat 系统回顾与展望[J]. 测绘信息与工程, 1998 (1): 1-6.
Wang S G. Review and prospect of the Landsat system[J]. Journal of Geomatics, 1998 (1): 1-6.

[2] 张玉君. Landsat8 简介[J]. 国土资源遥感, 2013, 25 (1): 176-177. doi:10. 6046/gtzyyg. 2013. 01. 30.
Zhang Y J. Introduction to the Landsat8[J]. Remote Sensing for Land and Resources, 2013, 25 (1): 176-177. doi: 10. 6046/gtzyyg. 2013. 01. 30.

[3] 初庆伟, 张洪群, 吴业炜, 等. Landsat-8 卫星数据应用探讨[J]. 遥感信息, 2013, 28 (4): 110-114.
Chu Q W, Zhang H Q, Wu Y W, et al. Application research of Landsat-8[J]. Remote Sensing Information, 2013, 28 (4): 110-114.

[4] CCSDS. CCSDS 121. 0-B-2, Lossless Data Compression, Blue Book, Recommended Standard, Issue 2[S]. Washington DC, USA: CCSDS, 2012.

[5] USGS. Ldcm - Isd - 010, Landsat Data Continuity Mission (LDCM) Space to International Cooperator Interface Specification Document, Version 1. 0[S]. USGS: 2011.

[6] USGS. LDCM - DFCB - 004, Landsat Data Continuity Mission (LDCM) Level 1 (L1) Data Format Control Book (DFCB), Version 6. 0[S]. USGS: 2012.

[7] 唐海蓉. Landsat7 ETM+ 数据处理技术研究[D]. 北京: 中国科学院研究生院 (电子学研究所), 2003.
Tang H R. Research on Landsat7 ETM+ Data Processing[D]. Bei-

- jing: Institute of Electronics, Chinese Academy of Sciences, 2003.
- [8] 初庆伟. Landsat-8 卫星数据预处理关键技术研究[D]. 北京: 中国科学院大学, 2013.
- Chu Q W. The Research of Key Technology of Landsat-8 Data Preprocessing[D]. Beijing: University of Chinese Academy of Sciences, 2013.
- [9] 宋鹏, 张晓林. 关于 CCSDS 无损数据压缩建议的初步研究[J]. 遥测遥控, 2001, 22(1): 8-13.
- Song P, Zhang X L. A Preliminary study of CCSDS lossless data compression proposals[J]. Journal of Telemetry, Tracking, and Command, 2001, 22(1): 8-13.
- [10] 李维忠, 贾小铁, 徐志立. 用于空间数据系统的无损数据压缩技术[J]. 计算机科学, 2003, 33(11): 122-124.
- Li W Z, Jia X T, Xu Z L. Lossless data compression for the space data system[J]. Computer Science, 2003, 33(11): 122-124.
- [11] 张兆亮. 基于 CCSDS 算法的星载图像压缩系统的 FPGA 实现[D]. 西安: 西安电子科技大学, 2010.
- Zhang Z L. Implementation of CCSDS Satellite Image Compression System with FPGAs[D]. Xi'an: Xidian University, 2010.
- [12] 朱云华, 王凤阳, 刘泳, 等. CCSDS 无损数据压缩算法的实现与应用研究[J]. 中国空间科学技术, 2008, 28(4): 40-46.
- Zhu Y H, Wang F Y, Liu Y, et al. Implementation and application study of CCSDS lossless data compression algorithm[J]. Chinese Space Science and Technology, 2008, 28(4): 40-46.

Analysis of Landsat8 satellite remote sensing data preprocessing

ZHU Jia

(Southwest China Institute of Electronic Technology, Chengdu 610036, China)

Abstract: The Landsat series satellites are the remote sensing resource series satellites, which are jointly managed by National Aeronautics and Space Administration and United States Geological Survey. Large quantities of high-resolution and stable image data provided by the Landsat series satellites have created good opportunities for the earth remote sensing exploration activities in the past forty years. Satellite remote sensing data preprocessing is the first step for obtaining remote sensing image, and has an important impact on the quality of the satellite remote sensing product. Aimed at tackling the Landsat8 raw data, the authors dealt in detail with the space data transmission protocol and data transmission format for Landsat8 data downlink. The preprocessing steps for raw data were analyzed, which included synchronization, transfer frame analyzing, unpack, mission data extracting, etc. In addition, the procedure of 0-level image product acquisition was described. Specifically, based on CCSDS (consultative committee for space data systems) recommended standard, the authors also discussed the method and technological process of lossless data decompression for Operational Land Imager (OLI) compressed data. The Landsat8 Level 0 data product obtained by the data preprocessing can provide high-quality basic images for the application of Landsat8 satellite remote sensing data.

Keywords: Landsat8; remote sensing satellite; data preprocessing; data decompression

作者简介: 祝佳(1984-), 男, 博士, 工程师, 主要从事卫星应用地面系统技术方面的研究。Email: zhujiaustc@163.com。

(责任编辑: 刘心季)