基于混合高斯模型的日冕物质抛射探测方法

曾丹丹^①, 白先勇^②, 强振平^③, 李强^①, 季凯帆^{①*}

① 昆明理工大学信息工程与自动化学院, 昆明 650500;

② 中国科学院空间科学与应用研究中心, 北京 100190;

③ 西南林业大学计算机与信息学院, 昆明 650224

* 联系人, E-mail: jkf@cnlab.net

2015-07-01 收稿, 2015-07-29 修回, 2015-07-31 接受, 2015-11-20 网络版发表

国家重点基础研究发展计划(2011CB811401)、空间科学先导专项(XDA04061002)和国家自然科学基金(11303011, 11263004, 11463003)资助

摘要 提出了一种基于混合高斯模型的日冕物质抛射(CME)探测方法. 基本思想是利用自适应的混合高斯模型建立较为稳定的日心极坐标下的日冕图像的动态背景, 从而探测作为前景变化的CME. 采用SOHO卫星上的大视角分光日冕仪(LASCO)观测的2组日冕序列图像作为研究对象, 研究的内容主要包括日冕序列图像的预处理、CME的探测、自适应混合高斯背景差分法与其他多种CME探测方法的对比3个方面. 实验结果表明, 自适应混合高斯背景差分法探测CME是可行的, 它能探测到CDAW手动目录列出的全部CME, 还能探测到CDAW探测不到的强度弱和张角小的CME, 而且探测数量也多于CACTus和SEEDS探测算法.

关键词 CDAW, 混合高斯模型, 日冕物质抛射, 探测, 背景差分

日冕物质抛射(coronal mass ejection, CME)是指在几分钟到几个小时的短时间里, 太阳日冕物质被抛射到行星际空间中的一种大尺度太阳爆发现象, 是太阳系内规模最大, 程度最剧烈的能量释放过程. 日冕物质抛射可以通过日冕仪看到, 迄今为止国外已经有许多空间和地基的日冕观测设备, 空间设备如OSO-7, Skylab, SMM, SOHO, STEREO等卫星上的日冕仪, 地基设备如美国高山天文台的MK3, MK4, k-Cor等.

尽管国外空间和地基日冕仪的数量很多, 遗憾的是至今国内仍然没有拥有自主知识产权的日冕仪观测设备. 无论是开展CME的科学研究还是开展空间环境预报业务, 都依赖于国外的数据. 近年, 随着我国太阳、空间物理以及空间环境预报业务的快速发展, 国内研制日冕仪的呼声也越来越高, 并取得了重大进展^[1,2]. 目前国内的几个空间太阳计划, 如深空

太阳天文台(DSO)^[3]、夸父计划(KUAFU)^[4]、太阳极轨望远镜探测计划(SPORT)^[5]、先进空基太阳天文台(ASO-S)^[6]等都将日冕仪列入其重要载荷之一. 2013年我国首个地基日冕观测站——丽江日冕仪国际合作站^[7]已经建成, 拥有我国自主知识产权的地基日冕仪也正在研制中^[8]. 基于国内已建成或即将建成的日冕仪数据开展CME科学研究或空间环境预报业务之前, 首先要从观测数据中检测CME事件, 所以CME探测识别方法的研究也十分迫切, 这正是本文的出发点. 国内外已有的日冕仪中, 1995年底发射的SOHO(solar and heliospheric observatory)卫星搭载的大视角分光日冕仪(the large angle spectrometric coronagraph, LASCO^[9])以前所未有的时间、空间的分辨率和稳定性观测到了近2个太阳活动周的CME爆发事件. 基于其数据开展的CME探测方法研究也很多, 所以本文将使用LASCO的数据开展CME探测方法研究, 为日

引用格式: 曾丹丹, 白先勇, 强振平, 等. 基于混合高斯模型的日冕物质抛射探测方法. 科学通报, 2016, 61: 1255–1264

Zeng D D, Bai X Y, Qiang Z P, et al. Detection for coronal mass ejection based on Gaussian Mixture Models (in Chinese). Chin Sci Bull, 2016, 61: 1255–1264, doi: 10.1360/N972015-00725

后基于我国日冕仪的数据探测CME储备方法.

目前探测CME主要有2类方法^[10]: (1) 通过人工处理检测, 当前2个主要的人工目录为: 美国戈达德航天飞行中心的the Coordinated Data Analysis Workshop Data Center (CDAW, http://cdaw.gsfc.nasa.gov/CME_list/)目录和美国海军实验室the Naval Research Lab (NRL, <http://lasco-www.nrl.navy.mil/index.php?p=content/cmelist>)目录. 这些目录都是通过观测者对每日的日冕图像进行编辑处理, 然后人工提取CME. (2) 基于计算机图像处理的CME自动探测方法. 第1个基于计算机自动检测CME的方法^[11,12]在2002年被首次提出, 开发了软件CACTus (Computer Aided CME Tracking Software Package, <http://sidc.oma.be/cactus/catalog.php>), 这个方法是利用基于极坐标下的霍夫变换技术对观测到的日冕图像进行CME检测; 2005年, Boursier等人^[13-15]提出了ARTEMIS方法, 这种方法通过将LASCO C2日冕图像转换到Synoptic Maps综合图中, 利用垂直条纹来检测CME; Brueckner等人^[16]在2006年使用形态学和变化阈值的方法在直角坐标系下探测CME; 2008年, Olmedo等人^[17,18]提出了太阳爆发事件探测系统(the solar eruptive event detection system, SEEDS, <http://spaceweather.gmu.edu/seeds/lasco.php>), 这个系统通过极坐标下的图像分割技术完成了CME的探测. 近几年来, 又有一些新的探测CME的方法出现, 如水平集和竞争区域的方法^[19]; 基于多尺度特性的方法^[20,21]; 基于纹理的图像处理技术^[22]、自动多尺度过滤技术^[23]; 频域处理^[24,25]; Savitzky-Golay滤波器^[26], 以及3D重建等^[27,28]. 上述方法很多都是计算相邻2帧图像之间的差, 即帧间差分, 或者减去一个平均的背景图像, 即背景差分法来进行初始的CME探测.

事实上, CME的检测过程可以看成在一个正常的日冕图像序列中检测突变(或者看成是异常事件). 或者说是在一个存在变化, 但相对稳定的冕流动态背景上检测一个快速变化的前景. 因此, 有效的背景表示和背景自适应更新是检测CME的关键技术.

Grimson等人^[29]在1999年提出的自适应混合高斯模型是一个非常有效的复杂背景估计方法; 2001年, Kaewtrakulpong^[30]利用最大期望法对模型参数进行更新, 从而进行动态背景估计; Power等人^[31]改进了算法, 不但提高了算法的速度, 而且能准确地测得背景像素均值. 近年来混合高斯模型在视频动态目标检测中得到了广泛的应用.

本文将混合高斯背景模型引入到CME图像的自动探测中, 称为自适应混合高斯背景差分法. 通过对LASCO C2观测图像的处理表明: 这一方法能探测到CDAW手动目录列出的所有CME, 且比基于帧间差分的SEEDS方法和基于霍夫变换的CACTus的方法在探测数量上具有显著的优势.

1 数据来源

实验数据来源于SOHO搭载的LASCO C2观测的日冕序列图像. 计算的LASCO C2的观测范围是 $2.3 \sim 6.5 R_{\odot}$ ($R_{\odot}$ 代表太阳半径). 使用了2组不同分辨率的实验数据集: 数据集I为1998/05/07 00:00 UT到1998/05/09 20:29 UT的C2 LEVEL 0.5图像序列, 像元分辨率是24.2 arc sec/pixel, 时间分辨率是24~28 min, 视场为 $3.4^{\circ} \times 3.4^{\circ}$; 数据集II为2013/11/14 00:00 UT到2013/11/14 23:48 UT的C2 LEVEL 0.5的图像, 像元分辨率是11.90 arc sec/pixel, 时间分辨率是12 min, 视场为 $3.4^{\circ} \times 3.4^{\circ}$. LEVEL 0.5的数据把C2观测的图像顶部旋转到太阳北极, 并保存为fits文件, 没有做过暗场、平场、渐晕等校正, 也没有去除仪器的杂散光. 此外, C2 level 0.5的数据中除太阳K冕外, 也包含F冕的信息, 一些经过C2视场的恒星、行星和彗星也包含其中. 太阳中心位置、太阳大小、曝光时间等信息可以从fits头文件中得到. 图1(a)和(b)分别显示了数据集II中在2013/11/14 00:00 UT时刻稳定日冕的图像和在2013/11/14 02:48 UT时刻有CME爆发的图像.

2 混合高斯模型方法

每个像素点采用 M 个(一般为3~5个)高斯分布来模拟背景. 其中每个高斯分布的概率密度公式如下:

图1 LASCO观测数据. (a) 稳定日冕图像; (b) 有CME的图像, 箭头指向CME

Figure 1 Observation data from the LASCO. (a) An image of the stable corona; (b) an image of a CME, which is pointed by white arrow

$$\delta(I_t, u_{j,t}, \sigma_{j,t}^2) = \frac{e^{-\frac{1}{2}(\sigma_{j,t})^{-2}(I_t - u_{j,t})^2}}{(2\pi)^{\frac{1}{2}} \cdot \sigma_{j,t}}, \quad (1)$$

其中, $\delta(I_t, u_{j,t}, \sigma_{j,t}^2)$ 表示像素点的概率密度, I_t 表示 t 时刻该点的像素灰度值, $u_{j,t}$, $\sigma_{j,t}^2$ 分别代表在 t 时刻第 j 个单高斯模型的均值和方差。

每个像素点的 M 个高斯分布的总概率为

$$P(I_t) = \sum_{j=1}^M w_{j,t} \cdot \delta(I_t, u_{j,t}, \sigma_{j,t}^2), \quad (2)$$

其中, $w_{j,t}$ 代表在 t 时刻第 j 个单高斯分布占总分布的比例, 即权值。可以看出, 混合高斯模型的核心是估计每个像素的 M 个高斯分布的均值、方差和权值。

(i) 模型初始化。混合高斯模型建模之前要对参数进行初始化, 第1个高斯模型的权值设为 w_0 , 其他都设为较小的权值(本文实验中取0.002), 要保证初始化的权值和为1, 方差都初始为 V_0 , 均值第1个高斯模型取第1幅图像的灰度值, 其他为0。

(ii) 模型参数优化。采用最大期望(expectation maximum, EM)算法^[32,33]来估计混合高斯模型的均值、方差和权值参数。反复迭代EM算法中的求期望和极大化2个步骤对混合高斯分布的参数进行极大似然估计, 直到其收敛为止。

EM算法描述如下, N 个样本 X 的对数似然函数表示如下:

$$l(X | \phi) = \prod_{t=1}^N \log \prod_{j=1}^M w_{j,t} \cdot \delta(I_t, u_{j,t}, \sigma_{j,t}^2), \quad (3)$$

其中

$$\phi = (\theta_1, \dots, \theta_M)^T, \quad \theta_j = (w_{j,t}, u_{j,t}, \sigma_{j,t}^2), \quad \sum_{j=1}^M w_{j,t} = 1. \quad (4)$$

用EM算法对高斯模型的参数估计分为求期望和极大化2步, 求期望的公式如下:

$$B_{j,t} = E(w_{j,t} | I_t; \phi) = \frac{w_{j,t} \cdot \delta_j(I_t; \phi)}{\sum_{j=1}^M \delta_j(I_t; \phi)}, \quad (5)$$

接着是参数的最大化, 包括权值、均值、方差3个的估计。

$$\text{权值估计: } w'_{j,t} = \sum_{i=1}^N B_{j,t}. \quad (6)$$

$$\text{均值估计: } u'_{j,t} = \frac{\sum_{i=1}^N B_{j,t} \cdot I_t}{N}. \quad (7)$$

$$\text{方差估计: } \sigma_{j,t}^{2'} = \frac{\sum_{i=1}^N B_{j,t} \cdot (I_t - u_{j,t})^2}{\sum_{j=1}^N B_{j,t}}. \quad (8)$$

不断重复上述的2个步骤, 反复更新上述3个值,

直到参数的变化不明显, 如: $|u_{j,t} - u'_{j,t}| < \varepsilon$, 本文取 ε 为0.0001。

(iii) 模型更新。由于晃流的准稳态变化^[34], 需要适时对其进行更新, 以适应背景的动态变化。在 i 时刻新的一帧图像出现时, 首先判断该位置的像素观测值是否与前一时刻同一位置的高斯分布模型相匹配, 本文匹配公式如下:

$$|I_t - u_{j,t}| < 2.5 \cdot \sigma_{j,t-1}. \quad (9)$$

满足(9)式则说明观测值与模型中某个分布匹配, 若观测值与模型中多个分布匹配, 则根据优先级的值选取最优匹配(优先级的值最大)进行更新, 根据分布的性质确定优先级: $p_{j,t-1} = w_{j,t-1} / \sigma_{j,t-1}$, 更新参数 $\beta = \alpha \cdot \delta(I_t, u_{j,t}, \sigma_{j,t}^2)$ (其中 α , β 都是参数学习率, 表示背景的更新速度, $0 \leq \alpha < 1$, 不同的实验数据值不同), 最优匹配分布的均值和方差更新如下:

$$u'_{j,t} = (1 - \beta) \cdot u'_{j,t-1} + \beta \cdot I_t, \quad (10)$$

$$\sigma_{j,t}^{2'} = (1 - \beta) \cdot \sigma_{j,t-1}^{2'} + \beta \cdot (I_t - u'_{j,t})^2. \quad (11)$$

如果相应的前一时刻同一位置的模型中没有与观测值匹配的高斯分布, 则将优先级最低的分布的参数重新赋值:

$$w'_{j,t} = w_0, \quad u'_{j,t} = I_t, \quad \sigma_{j,t}^{2'} = V_0. \quad (12)$$

权值的更新如下:

$$w_{j,t} = (1 - \alpha) \cdot w_{j,t-1} + \alpha \cdot \text{match}_{j,t}, \quad (13)$$

其中 $\text{match}_{j,t}$ 为匹配标识, 当有高斯分布与观测值匹配时 $\text{match}_{j,t} = 1$, 否则 $\text{match}_{j,t} = 0$ 。

(iv) 模型学习率改进。由于Kaewtrakulpong等人^[30]提出的算法在背景更新时反复迭代, 收敛较慢。并且全局学习率没有充分利用到每个像素的匹配信息, 对运动目标物质的检测效果不佳。本文结合每个像素的匹配和初始化的信息, 对每个像素定义如下:

$$\alpha = 1 / \text{Smatch}_{k,t}, \quad \text{Smatch}_{k,t} = \text{Smatch}_{k,t-1} + \sum_{j=1}^M \text{match}_{j,t}, \quad (14)$$

其中 $\sum_{j=1}^M \text{match}_{j,t}$ 代表在 t 时刻像素点匹配的高斯分布总个数。 $\text{Smatch}_{k,t}$ 代表到 t 时刻为止像素点 k 被匹配的总次数, 它的值越大说明此像素点是背景的可能性越大, 学习更新率 α 也就越小, 如果其值越小说明背景还不稳定, 学习率更新率大, 要使背景尽快更新。

(v) 运动物质的确定。针对每个像素点, 从 M 个高斯分布中选取 b 个分布 ($1 \leq b \leq M$) 作为背景分布,

使得 $\arg \min_b \sum_{j=1}^b w_{j,t} > T$, 其中 $T(0 \leq T \leq 1)$ 表示背景阈值, 本文取0.8, 若当前处理像素点与所选 b 个分布中的任意一个高斯分布匹配则该像素点被判断为背景, 否则该像素点被判断为前景, 即运动物质CME.

3 日冕物质抛射探测

使用混合高斯模型探测CME之前, 先对level 0.5的原始图像进行预处理. 预处理包括去除噪声、图像标准化和将图像从笛卡尔坐标转换到极坐标3个步骤. 采用 3×3 模板的中值滤波来抑制图像中的噪声, 这个过程基本可以去除视场中的恒星、行星、彗星等信息, 也可以去除图像的一些随机噪声, 然后将图像强度标准化到0~255之间, 并根据观测FITS头中提供的太阳中心位置, 将太阳中心平移到图像中心, 从而对齐整个图像序列. 考虑到太阳圆对称结构以及冕

流和CME的运动方向基本是沿着日面径向, 所以将图像从直角坐标系转换到极坐标. 在2组观测数据集中, 通过计算得知C2的挡板半径约为2.3倍太阳半径, 故只研究2.3倍太阳半径之外的日冕变化, 在极坐标转换的时候, 忽略了小于2.3倍太阳半径的部分. 图2是图1(a)和(b)中大于2.3倍太阳半径之外的日冕转换到极坐标的图像. 该转换是将笛卡尔坐标的图像从原点位置(从南极开始, 正下方对应极坐标的 0° 角)开始逆时针旋转对应到极坐标上, 极坐标的横坐标对应角度(角度范围是 $0^\circ \sim 360^\circ$), 纵坐标对应半径(半径范围为 $2.3 \sim 6.5 R_\odot$).

使用自适应混合高斯背景差分法建立背景模型, 首先要确定高斯分布的个数 M . 本文统计了2组数据中多个位置的背景像素的变化, 发现观测数据集 I 和 II 中大部分的模拟背景像素分别符合5和4个高斯分布的叠加, 因此采用5和4个高斯分布作为数据集

图2 对应的图1转换到极坐标的图. (a) 图1(a)中大于2.3倍太阳半径的日冕极坐标图; (b) 由图1(b)中大于2.3倍太阳半径的日冕极坐标图, 矩形标记的是CME

Figure 2 Convert to the polar coordinate in Figure 1. (a) The polar coordinate image of Fig. 1(a) that greater than 2.3 times of the Sun radius; (b) the polar coordinate image of Fig. 1(b) that greater than 2.3 times of the Sun radius. CME is marked by a rectangle

图3 (网络版彩色)2组数据不同位置的高斯拟合. (a) 观测数据集I背景位置(213, 145)的高斯拟合; (b) 观测数据集II背景位置(606, 312)的高斯拟合

Figure 3 (Color online) The Gaussian fit of two set of data in different location. (a) The Gaussian fitting of the observation data set I in the pixel position of (213, 145) of the background; (b) the Gaussian fitting of the observation data set II in the pixel position of (606, 312) of the background

表 1 使用不同CME探测方法处理数据集 I 和 II 的结果

Table 1 The processing results of data set I and II by different CME detection methods

		总探测数 (个)	CDAW中的 CME的探测 数(个)	CDAW中 CME探测率 (%)
数据集 I	CDAW	19		
	SEEDS	22	14	73.7
	CACTus	12	9	47.4
	自适应混合高斯 背景差分法	25	19	100
数据集 II	CDAW	4		
	SEEDS	9	2	50
	CACTus	2	2	50
	自适应混合高斯 背景差分法	9	4	100

I 和 II 的模拟高斯的个数. 图3(a)显示了数据集 I 中一个较为稳定的日冕像素的多高斯拟合, 图3(b)为数据集 II 的一个像素点的多高斯拟合. 横坐标表示图像背景灰度值, 纵坐标表示对应的灰度值在整个图像序列中的数量.

建立混合高斯背景模型有不同的初始化方法, 如对 M 个高斯分布采用均值等间隔, 方差和权重相同的方法. 或以连续多帧图像中像素点的灰度平均值作为初始化背景的第1个均值, 其余为0, 方差都相等, 并且以第一权重优先的方法来进行初始化. 经过试验, 本文采用的方法是以无爆发的第1幅图像的灰度值作为第1位高斯分布的均值, 其余分布的均值为0; 所有分布的方差都等于 V_0 ; 第1分布的权重的系数为 w_0 , 其余接近0.

然后用EM算法优化参数, 再通过自适应混合高斯背景差分法建立图像序列中每个时刻图像的动态背景模型. 用观测图像减去背景, 可得到包含CME的前景图像.

对于提取出来的前景图像, 先对其进行中值滤波, 然后采用大津法(OSTU)^[35] 阈值分割的方法将其二值化, 并进行区域分割, 再剔除面积较小的区域, 最终提取CME.

4 CME探测实验结果对比与分析

以LASCO观测的2组不同时间段的C2序列图像作为测试对象, 以CDAW手动列表为基准, 用自适应混合高斯背景差分算法得到的结果分别与SEEDS算法和CACTus算法的处理列表对比, 来验证我们提出

的算法的有效性.

自适应混合高斯背景差分法建模数据集I的初始高斯参数: $M=5$, 权值第1位参数 $w_0=0.992$, 其余取等权0.002, 方差都为 $V_0=36$, 背景分布阈值 $T=0.8$, 均值第1位参数取第1帧图像同位置的像素灰度值, 其余为0. 数据集II的初始化参数: $M=4$, $w_0=0.994$, 其余取等权0.002, 方差都为 $V_0=25$, 背景分布阈值 $T=0.8$, 均值第1位参数同样也采用第1帧图像的灰度值, 其余均为0. 自适应混合高斯背景差分法与其他几种探测算法的对比如表1所示(其中探测率等于其他方法测得在CDAW中同时刻的CME总数与CDAW同时段的总数的比值).

对于数据集I, 表1显示自适应混合高斯背景差分能探测到CDAW列出的全部CME, 而且还能探测到一些CDAW未列入的活动. 在探测总数上, 自适应混合高斯背景差分法除了探测到的活动总数比SEEDS和CACTus都多之外, 对CDAW列表中目视可见的CME的探测率也比SEEDS和CACTus要高. 对于数据集II, 表1显示了自适应混合高斯背景差分方法探测到的CME总数与SEEDS探测列表相同, 多于CDAW列表和CACTus列表. 对CDAW中的CME, 自适应混合高斯背景差分方法均可以成功探测, 探测率远高于SEEDS和CACTus.

图4(a)是数据集I中的第27帧(1998/05/07/14:58 UT), 图4(b)是相应的自适应混合高斯背景差分法在极坐标上测得的CME. 两图中矩形代表的是CDAW和自适应混合高斯背景差分法都能测得的CME, 其角宽度为 78° . 椭圆形表示自适应混合高斯背景差分法能探测到, 但CDAW未探测到的CME, 白色箭头标记其初始时的角宽度仅为 18° . 这2个CME的强度比约为9.4:1. 我们也分析了其他用自适应混合高斯背景差分法新探测到的CME, 它们都表现为强度较低、张角较小或与冕流背景大量重叠的特征.

常用的帧间差分法是以序列图像中相邻2帧做差来获取运动目标. 若2帧之间运动不明显时, 差分法则可能探测不到2帧之间的变化. 如在数据集I中, 基于帧间差分的SEEDS算法未能探测到CDAW列表中的5个CME. 为了进一步比较自适应混合高斯背景差分法与帧间差分法的区别, 分别从2个数据集中各选取2帧典型意义的图像来进行比较. 图5(a)和(b)是数据集I中第22帧(1998/05/07/11:05 UT)和51帧(1998/05/08/06:27 UT)的极坐标原始帧, 图5(c), (d)

图4 (网络版彩色)自适应混合高斯探测. (a) 数据I中第27帧; (b) 自适应混合高斯差分法测CME

Figure 4 (Color online) Adaptive mixture Gaussian probe. (a) The 27th frame in the data set I; (b) the CME measured by adaptive Gaussian mixture background subtraction

和(e), (f)是对应图5(a), (b)的帧间差分法和自适应混合高斯背景差分法得到的二值图。

图5(a)和(b)中各有2个CME, 分别用椭圆和矩形标记, 都记录在CDAW中. 而采用自适应混合高斯背景差分法, 也都可以被成功探测到(图5). 但使用帧间差分法, 其中2个用矩形标记的CME就被漏检了. 帧间差分虽然简便, 但是对于慢速运动的物体则难以

提取完整的目标. 相比而言自适应混合高斯背景差分法自适应强, 能准确地检测到动态背景, 故而能检测运动速度缓慢和强度较弱的前景运动。

图6(a)和(b)是数据集II中第6帧(2013/11/14/01:25 UT)和78帧(2013/11/14/16:48 UT)的极坐标像原始帧, 图6(c)~(f)是对应(a)和(b)的帧间差分法和自适应混合高斯背景差分法得到的二值图. 其反映的情况与图4

图5 (网络版彩色)数据集 I 中帧间差分和自适应混合高斯背景差分法对比

Figure 5 (Color online) The comparison of frame subtraction and adaptive Gaussian mixture background subtraction for data set I

图6 (网络版彩色) 测试数据集II中帧间差分法和自适应混合高斯背景差分法对比

Figure 6 (Color online) The comparison of frame subtraction and adaptive Gaussian mixture background subtraction for data set II

基本是一致的。在图5(c)中, 帧间差分法漏检了矩形框标记的CME, 而在图6(d)中, 2个都漏检了, 但自适应混合高斯背景差分法却能成功检测到。

5 结论

本文提出了一种利用自适应混合高斯背景差分模型建立日冕图像动态背景的方法, 从而探测作为前景变化的CME。实验结果表明, 自适应混合高斯背景差分具有较好的效果, 检测到了实验数据集中CDAW手动目录列出的所有CME, 同时也探测到一些额外的强度相对弱的小张角的目标。虽然传统的帧间差分法简单、易于实现, 但从原理上看, 就只是针对单一、稳定的背景。而混合高斯模型具有良好的自适应性, 因此可以较为准确地描述复杂背景, 尤其是动态变化的背景。利用这一方法识别叠加在变化冕流上的CME具有较为明显的优势。同时, 由于在背景建模的过程中, 采用了对序列图像中的多帧图像信息进行统计的方法, 因此大大抑制了单帧图像噪声的影响, 增强了对CME探测的鲁棒性。地基日冕仪的观测数据中存在变化的天空背景, 信噪比较

弱, 如何有效地扣除天空背景以增强信噪比是地基日冕观测的难点之一。从理论上讲, 这一方法对地面日冕观测数据中天空背景的去除非常有益, 扣除天空背景的数据也可以采用此方法探测CME事件。如何将这一方法应用于这些观测也是我们下一步将开展的研究工作。

混合高斯方法存在的最大问题是各种参数和阈值的估计, 如背景初始均值、方差、背景分布的阈值等。这些参数还未能实现自动测量, 需要通过一些先期实验来人工进行选择。这些经验值的选取对算法会有一定的影响。所以, 进一步的改进算法, 提高经验值的准确性以及减少人为的主观性引起的误差变得十分必要。另外, 混合高斯模型的计算量较大, 耗时较长, 还不适应做观测数据的实时处理。但随着计算机硬件技术的发展, 相信这个问题会有改观。最后, CME的观测特征如初始位置、张角、速度、加速度等参数十分重要, 利用自适应混合高斯背景差分方法探测CME参数和其他方法的区别也是我们关注的问题, 未来计划利用LASCO的数据采用更多的样本详细对比其异同。

致谢 感谢美国海军研究实验室、德国马克斯·普朗克研究所、法国天文实验室以及英国伯明翰大学提供的 SOHO/LASCO 数据。其中 SOHO 是欧空局和美国航天局之间的国际合作项目。

参考文献

- 1 Webb D, Howard T. Coronal mass ejections: Observations. *Liv Rev Solar Phys*, 2012, 9: 1285–1295
- 2 General Armament Department of Electronic Information Infrastructure. *Solar Storms Secret* (in Chinese). Beijing: National Defense Industry Press, 2012 [总装备部电子信息基础部. 太阳风暴揭秘. 北京: 国防工业出版社, 2012]
- 3 Li H. Space solar telescope optics coronagraph (in Chinese). In: *Seventh Annu Conference Meeting Manuals Collect Chin Space Sci Inst*, Dalian, 2009 [黎辉. 空间太阳望远镜光学日冕仪. 见: 中国空间科学学会第七次学术年会会议手册及文集, 大连, 2009]
- 4 Tu C Y, Schwenn R, Donovan E, et al. Space weather explorer—The KuaFu mission. *Adv Space Res*, 2008, 41: 190–209
- 5 Wu J, Sun W Y, Zheng J H, et al. Imaging interplanetary CMEs at radio frequency from solar polar orbit. *Adv Space Res*, 2011, 48: 943–954
- 6 Gan W Q. ASO-S: Advanced space-based solar observatory. 40th COSPAR Sci Assembly, Moscow, 2014
- 7 Liu Y. Chinese first aircraft coronagraph settled in Lijiang, Yunnan (in Chinese). *Amateur Astronomers*, 2014, 1: 60–63 [刘煜. 国内首架日冕仪落户云南丽江. 天文爱好者, 2014, 1: 60–63]
- 8 Bo H Y. Research the key technology of foundation coronagraph stray light suppression (in Chinese). Dissertation of Masteral Degree. Beijing: University of Chinese Academy of Sciences, 2012 [卜和阳. 地基日冕仪杂散光抑制的关键技术研究. 硕士学位论文. 北京: 中国科学院大学, 2012]
- 9 Brueckner G E, Howard R A, Koomen C M, et al. The large angle spectroscopic coronagraph (LASCO). *Solar Phys*, 1995, 162: 357–402
- 10 Yashiro S, Michalek G, Gopalswamy N. A comparison of coronal mass ejections identified by manual and automatic methods. *Ann Geophys*, 2008, 26: 3103–3112
- 11 Robbrecht E, Berghmans D. Automated recognition of coronal mass ejections (CMEs) in near-real-time data. *Astron Astrophys*, 2004, 425: 1097–1106
- 12 Bonte K, Jacobs C, Robbrecht E, et al. Validation of CME detection software (CACTus) by means of simulated data, and analysis of projection effects on CME velocity measurements. *Solar Phys*, 2011, 270: 253–272
- 13 Boursier Y, Llebarria A, Goudail F, et al. Automatic detection of coronal mass ejections on LASCO-C2 synoptic maps. *Solar Phys Space Weather Instrum*, 2005, doi: 10.1117/12.616011
- 14 Boursier Y, Lamy P, Llebarria A, et al. The ARTEMIS catalog of LASCO coronal mass ejections. *Solar Phys*, 2009, 257: 125–147
- 15 Floyd O, Lamy P, Boursier Y, et al. ARTEMIS II: A second-generation catalog of lasco coronal mass ejections including mass and kinetic energy. *Solar Phys*, 2013, 288: 269–289
- 16 Brueckner M, Howard R A, Koomen C M, et al. Automatic detection and classification of coronal mass ejections. *Solar Phys*, 2006, 237: 419–431
- 17 Olmedo O, Zhang J, Wechsler H, et al. Automatic detection and tracking of coronal mass ejections in coronagraph time series. *Solar Phys*, 2008, 248: 485–499
- 18 Olmedo O, Zhang J. Partial torus instability. *Astrophys J*, 2010, 718: 433
- 19 Norberto G, Marta M, Julio J, et al. Detection and tracking of coronal mass ejections based on supervised segmentation and level set. *Pattern Recogn Lett*, 2010, 31: 496–501
- 20 Byrne J P, Gallagher P T, McAteer R T J, et al. The kinematics of coronal mass ejections using multiscale methods. *Astron Astrophys*, 2009, 495: 325–334
- 21 Gallagher P T, Young C A, Byrne J P, et al. Coronal mass ejection detection using wavelets, curvelets and ridgelets: Applications for space weather monitoring. *Adv Space Res*, 2011, 47: 2118–2126
- 22 Goussies N, Stenborg G, Vourlidis A, et al. Tracking of coronal white-light events by texture. *Solar Phys*, 2010, 262: 481–494
- 23 Byrne J P, Gallagher P T, McAteer R T J, et al. Automatic detection and tracking of CMEs II: Multiscale filtering of coronagraph data. 2012, arXiv:1207.6125
- 24 Zeng Z X, Wei Y L, Li J S. A method to automatic detecting coronal mass ejections in coronagraph based on frequency spectrum analysis. In: *Proceedings of the 2012 International Conference of Modern Computer Science and Applications*. New York: Springer, 2013
- 25 Zeng Z X, Liu Y, Wei Y L. Detection for coronal mass ejection based on the mutation spectrum analysis (in Chinese). *Huazhong Univ Sci Tech: Nat Sci Ed*, 2012, (S1): 379–383 [曾昭宪, 刘毅, 魏雅利. 基于频谱突变分析的日冕物质抛射识别方法. 华中科技大学学报: 自然科学版, 2012, (S1): 379–383]

- 26 Byrne J P, Long D M, Gallagher P T, et al. Improved methods for determining the kinematics of coronal mass ejections and coronal waves. *Astron Astrophys*, 2013, 557: A96
- 27 Liewer P C, De Jong E M, Hall J R, et al. Stereoscopic analysis of the 19 May 2007 erupting filament. *Solar Phys*, 2009, 256: 57–72
- 28 Liewer P C, Hall J R, Howard R A, et al. Stereoscopic analysis of STEREO/SECCHI data for CME trajectory Determination. *J Atmos Solar-Terrestrial Phys*, 2011, 73: 1173–1186
- 29 Grimson W E L, Stauffer C, Romano R, et al. Using adaptive tracking to classify and monitor activities in a site. In: *Proceedings/CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. 1998. 22
- 30 Kaewtrakulpong P, Bowden R. An improved adaptive background mixture model for real-time tracking with shadow detection. *Video-Based Surveillance Systems*. New York: Springer, 2002. 135–144
- 31 Schoonees J A, Power P W. Understanding background mixture models for foreground segmentation. In: *Proceedings Image and Vision Computing New Zealand*, 2002. 2003. 267–271
- 32 Setiawan W. Combining of image classification with probabilistic neural network (PNN) approaches based on expectation maximum (EM). *Int J Comput Sci Issues*, 2012, 9: 92–102
- 33 Sun D F, Chen Z G, Liu W J. Based on maximum parameters likelihood estimation in EM algorithm (in Chinese). *J Henan Univ (Nat Sci)*, 2002, (4): 35–41 [孙大飞, 陈志国, 刘文举. 基于 EM 算法的极大似然参数估计探讨. *河南大学学报: 自然科学版*, 2002, (4): 35–41]
- 34 Chen Y. Research coronal dynamics: Streamer, coronal mass ejection and their interaction (in Chinese). *Chin Sci Bull (Chin Ver)*, 2013, 58: 1620–1650 [陈耀. 日冕动力学研究进展: 冕流、日冕物质抛射及其相互作用. *科学通报*, 2013, 58: 1620–1650]
- 35 Wu Y Q, Wu W Y, Pan Z. A two-dimensional histogram oblique for Otsu thresholding fast iterative algorithm (in Chinese). *Graphics*, 2009, 30: 89–96 [吴一全, 吴文怡, 潘喆. 二维直方图区域斜分 Otsu 阈值分割的快速迭代算法. *图学学报*, 2009, 30: 89–96]

Detection for coronal mass ejection based on Gaussian Mixture Models

ZENG DanDan¹, BAI XianYong², QIANG ZhenPing³, LI Qiang¹ & JI KaiFan¹

¹ Faculty of Information Engineering and Automation, Kunming University of Science and Technology, Kunming 650500, China;

² Center for Space Science and Applied Research, Chinese Academy of Sciences, Beijing 100190, China;

³ College of Computer and Information, Southwest Forestry University, Kunming 650224, China

A new automatic detection algorithm extracted solar moving target coronal mass ejection (CME) is proposed in this paper. The CME releases huge quantities of matter and electromagnetic radiation solar-terrestrial space from the Sun. When the ejection is directed toward the Earth, it even effects the life of human being. Automatic detecting the CME for observed image is very useful for studying those solar activities. The CME detection can be considered as detection and tracing moving object in a complicated background. Therefore, according to the gradient characteristic of the background and the demand of real-time processing, we developed a dynamic new background estimation algorithm that based on the Adaptive traditional Gaussian Mixture Model. An expectation-maximization algorithm is applied to improve the initialization of the model, and a learning rate for each pixel in the sequence of image is adaptively for updating the background of coronal sequential images. The CME can be detected as a foreground object after subtracting the background that is estimated by the adaptive Mixture Gaussian Models from the original image extraction. The processing is under the polar coordinate of heliocentric. Two image sequences of coronal observed by the Large Angle Spectroscopic Coronagraph (LASCO) in the SOHO satellite were used. The paper gives the details of the preprocessing of the image sequences of coronal, detecting of the CME and the results comparison between the proposed and other CME automatic detection methods. The detection rate, false alarm rate, and detection of the amount of the CME are discussed. The experimental results show that the method is practicable and effective for detecting the CME. Compared with the manual detection of moving target detection, such as CDAW, the automatic CME detection method is more rapid and powerful. The method not only can detect all of the CMEs listed on the CDAW CME catalog, but also the CME with weaker intensity and smaller angle. Furthermore, it has better performance than the algorithm of the CACTus and the SEEDS. The method with adaptive mixture Gaussian background subtraction has higher detection rate, lower false alarm rate and more effectively for detecting the CME than the SEEDS and CACTus, but it takes more computer time and needs more computational capabilities. Our model has a preferable adaptive in the case of uncertain factors, and correspondence quickly.

CDAW, Gaussian Mixture Models, coronal mass ejection, detection, background subtraction

doi: 10.1360/N972015-00725