

一种基于支持向量机的射电可见度数据自动标注方法*

代慧梅¹ 梅 盈^{1,2} 王 威³ 邓 辉¹ 王 锋^{1,2†}

(1 昆明理工大学云南省计算机技术应用重点实验室 昆明 650505)

(2 中国科学院云南天文台 昆明 650011)

(3 中国科学院国家天文台 北京 100012)

摘要 对中国明安图超宽频谱射电日像仪(Mingantu Ultrawide Spectral Radioheliograph, MUSER)观测所得到的可见度数据进行标注(Flag), 以剔除数据中的异常值是后续成图处理的一个重要工作. 研究中利用支持向量机(Support Vector Machine, SVM)技术建立可信可见度数据标记模型, 进而利用模型对可见度数据测试样本集进行测试标注. 结果表明, 该方法与传统基于统计的方法相比准确率有明显改进, 可较好地判断出故障天线, 对MUSER故障天线引起的可见度数据失真标记正确率可达到86%左右, 且不受太阳爆发活动对数据的影响.

关键词 太阳: 活动, 太阳: 射电辐射, 恒星: 成像, 方法: 数据分析, 技术: 其他

中图分类号: P161; **文献标识码:** A

1 引言

明安图超宽频谱射电日像仪(Mingantu Ultrawide Spectral Radioheliograph, MUSER)是我国自行研制的可以同时以高时间、空间和频率分辨率对太阳进行射电频谱成像的设备^[1]. MUSER项目分为两期完成: 第1期MUSER-I(低频阵)由40面4.5 m口径的抛物面天线及其接收设备组成, 在64个频点上成像; 第2期MUSER-II(高频阵)由60面2 m口径的抛物面天线及其接收设备组成, 在528个频点上成像^[2].

与光学望远镜直接成像不同, 射电望远镜主要是接收天体射电信号的强度、相位及偏振等信息, 进而通过后续处理实现成图. 在观测中, 观测数据通常会受到各种因素(例如电磁干扰等)的影响. 此外, 天线故障、馈源故障、通道差错、系统增益等数据接收设备故障也会导致观测数据的异常. 在观测数据的处理过程中, 对这些异常数据进行判断、标注和剔除是射电观测数据处理的重要工作.

2015-06-21收到原稿, 2015-07-22收到修改稿

*中国科学院-国家自然科学基金委员会天文联合基金重点项目(U1231205)、国家自然科学基金项目(11103005, 11263004)、云南省应用基础基金重点项目(2013FA013, 2013FA032)共同资助

† wf@cnlab.net

在射电研究领域初期, 对于射电天文数据的异常处理一般靠人工进行记录比对. 但随着观测数据的日益庞大, 人工处理已经完全不切实际. 随着对领域研究的逐渐深入, 为了高效地对失真数据进行评估和标注, 各望远镜观测项目根据其阵列方式、天线数量、基线长度等性能参数指标, 均分别建立了匹配其数据特征的数据标记模型. 如在阿塔卡玛毫米/亚毫米波阵列望远镜(Atacama Large Millimeter/submillimeter Array, ALMA)^[3-4]中, 为了标记失真数据, 研究者们采用了CASA(Common Astronomy Software Applications)里的TFCrop、RFlag等方法; 为了实现巨米波射电望远镜(Giant Metrewave Radio Telescope, GMRT)^[5-6]失真观测数据的自动标记, Prasad和Chengalur通过研究分析采用了FLAGCAL^[7].

显而易见, 在MUSER进入试观测以后, 为了实现自动数据处理, 自动剔除由于天线故障引起的失真可见度数据, 提高数据的准确性和可靠性, 是后续MUSER数据网格化、洁化成图处理的重要前导工作. 在前期研究工作中曾利用GMRT里面的VSR(Vector to Scalar Ratio)标记算法及ABC(Antenna/Baseline/Channel)标记算法对MUSER数据标记进行了尝试, 结果令人不满意.

本文在利用模式识别技术, 结合现有开源软件包LIBSVM的基础上, 利用支持向量机技术(Support Vector Machine, SVM)^[8]研究了可见度数据的异常标记方法, 通过已知基线损坏情况构造训练集. 随后的一系列实验结果表明, SVM技术可以较好地实现对MUSER观测中异常数据的自动标记.

2 支持向量机技术

支持向量机^[8-9]是一种监督式学习的方法, 是在统计学习的基础上发展起来的一种新的机器学习方法, 它是建立在统计学习理论的VC维理论和结构风险最小化原则上的, 避免了局部极小点(支持向量机是一种凸二次优化问题, 能够保证极值点是全局最优解), 通常用来进行模式分类以及回归分析. 本文采用SVM对MUSER可见度数据进行评估, 就是用了SVM的分类思想, 即SVM通过最大化决策边界的边缘来找到最优超平面.

简单来说, 假设输入数据为 $\{y_i, \mathbf{x}_i\}$, $i=1, 2, \dots, k$. $y_i \in \{-1, 1\}$ 为类标签, $\mathbf{x}_i \in \mathbf{R}_n$ 为输入向量, k 为样本长度, n 为向量特征的维度. 在训练样本线性可分时, SVM寻找最优超平面问题可简化为求:

$$\begin{aligned} \min \quad & \frac{1}{2} \|\mathbf{w}\|^2 = \min \frac{1}{2} (\mathbf{w}^T \cdot \mathbf{w}), \\ \text{s.t.} \quad & y_i (\mathbf{w} \cdot \mathbf{x}_i) \geq 1, \end{aligned} \quad (1)$$

其中, $\mathbf{w}$ 为权值矢量. 在解决实际问题时, 碰到的大都是线性不可分的情况. 对于这一点, SVM里面的核函数 $K(\mathbf{x}_i, \mathbf{x}_j)$ 会把输入 $\{y_i, \mathbf{x}_i\}$ 从低维空间映射到高维空间. 这时候, SVM寻找最优超平面问题转化为:

$$\begin{aligned} \max \quad & w(a) = \sum_{i=1}^k a_i - \frac{1}{2} \sum_{i=1, j=1}^k a_i a_j y_i y_j K(\mathbf{x}_i, \mathbf{x}_j), \\ \text{s.t.} \quad & \sum_{i=1}^k y_i a_i = 0, \quad 0 \leq a_i \leq C, \end{aligned} \quad (2)$$

其中, $i=1, 2, \dots, k$; a_i 是拉格朗日多项式; C 是常数.

经过一系列计算, 得到最优解 $\mathbf{a}^*$, 最优权重向量 $\mathbf{w}^*$ 和最优偏置向量 $\mathbf{b}^*$. 对于未知分类向量 $\mathbf{x}$, 最终最优分类函数为:

$$f(\mathbf{x}) = \text{sgn}((\mathbf{w}^*, \mathbf{x}) + \mathbf{b}^*) = \text{sgn}\left(\sum_{i=1}^k \mathbf{a}^* y_i K(\mathbf{x}_i, \mathbf{x}) + \mathbf{b}^*\right). \quad (3)$$

例如, 对于线性可分情况的分类问题, 通俗地讲, 就是用一条直线把属于不同类别的样本点分开. 以平面坐标系中的直线方程为例说明, 平面坐标系中直线方程为 $Ax+By+c=0$, 引入向量概念后, 方程可以写成 $\{\mathbf{A}, \mathbf{B}\} \cdot \{x, y\} + c = 0$, 其中 $\{\mathbf{A}, \mathbf{B}\}$ 就是方程的法向量, 再把直线方程进行简化, 得到 $\mathbf{w} \cdot \mathbf{x} + b = 0$ 的形式. 当我们输入 $\{y_i, \mathbf{x}_i\}$, 规定 $\mathbf{x}_i$ 属于第1类时, y_i 为1; $\mathbf{x}_i$ 属于第2类时, y_i 为-1. 对此两类问题, 则直线两侧的样本点分为正类和负类, 进一步用符号函数的方式推断点 $\mathbf{x}$ 所对应的类别的决策函数即为 $y=f(\mathbf{x})=\text{sgn}(\mathbf{w} \cdot \mathbf{x} + b)$, 根据符号函数的定义, y 的取值要么为1, 要么为-1. 此时的分类问题为: 对于任意给定的 $\mathbf{x}$, 根据训练集预测出结果属于正类或是负类. 仅通过训练集信息解不出参数 $\mathbf{w}$ 和 $\mathbf{b}$, 为使 $f(\mathbf{x})$ 对原有样本预测误差最小, 可将问题转化为期望误差最小、经验风险最小, 最后结合统计学习理论又将问题转化为结构风险最小, 选取恰当置信范围, 得到经验风险最小的函数即为最优函数.

现如今支持向量机已被广泛应用到各行各业中^[10-11]. 在天文学领域中也有SVM的身影, 例如利用SVM从分光光度法推断天体物理参数^[12], 利用SVM模型进行自动化短期太阳耀斑爆发的预测^[13]等. 但是在射电天文领域, 利用SVM进行天线阵列故障情况判断以剔除异常数据还没有开展过相应的研究工作.

MUSER观测中可见度数据进行标注这一问题的实质, 可以看成是一个分类问题, 可信数据以及由天线故障造成的失真数据各属一类. 本文重点研究如下3个关键问题:

- (1)SVM方法是否可以用于射电可见度数据的判断?
- (2)所获得的模型是否鲁棒, 是否会受到太阳活动影响?
- (3)小样本训练出来的模型能否适用于较大样本的测试集数据, 满足全天观测的需要?

为加快研究工作, 在研究中直接采用开源的SVM软件库LIBSVM作为底层开发包. LIBSVM是台湾林智仁(Chih-Jen Lin)教授2001年开发的一套支持向量机软件库, 这套库程序小, 运用灵活, 开源且易于扩展, 可以很方便地对数据做分类或回归. 在LIBSVM基础上, 根据MUSER实际需求, 进行了相关修改, 在训练时采用网格遍历的方式选取合适的最优惩罚因子 c 和核函数参数 γ , 利用综合性能最优的径向基函数RBF^[14]对观测数据进行训练建模, 得到SVM库, 以此为随后判断数据可靠性与否进行评估及数据标注校正作铺垫. 实验过程中所用到的样本训练集和待分类的测试样本集都是MUSER阵列实际观测所得到的数据. 基于SVM的MUSER可见度数据评估过程如图1所示.

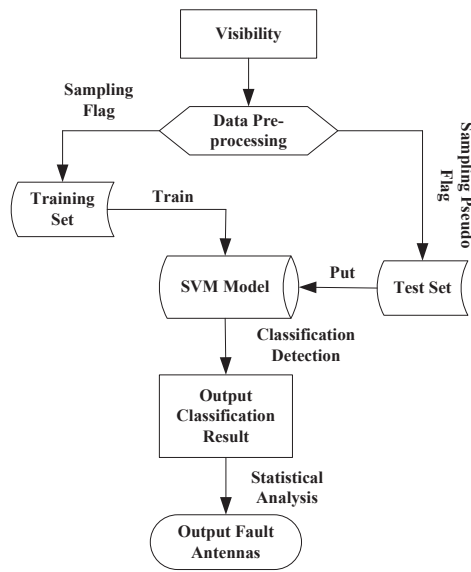


图 1 可见度数据评估标注流程图

Fig. 1 The flowchart of visibility evaluation and flag

3 数据分析与实验

3.1 数据来源

根据科学目标要求, MUSER天线的排列选用了综合性能最优的螺旋阵列. 天线成像视场 0.5° – 7° , 其中最长基线达3 km. 各天线与后端数字接收设备间的信号传输通过光纤完成. 在本文中, 以MUSER-I所获得的观测数据作为研究对象. MUSER-I整体呈现出三臂螺旋结构, 是由中心IA0号天线和A轴编号IA1-IA13、B轴编号IB1-IB13以及C轴编号IC1-IC13的天线共同构成的以中心天线为轴的近似同心圆. 为了方便后续数据计算, 研究中将中心天线编号0, A轴天线编号1-13, B轴天线编号14-26, C轴天线编号27-39.

由综合孔径成像原理, 对于一个由 n 面天线构成的天线阵, 可以构成 $n(n-1)/2$ 架干涉仪, 因此得到 $n(n-1)/2$ 条基线, 也可以在空间频率平面得到 $n(n-1)/2$ 个 (u, v) 空间频率域上的点. 这些点的分布即UV覆盖, 实验中研究的观测可见度数据就是这些UV覆盖上的点. MUSER-I有40面天线, 共有780条基线, 一次数据采集可以获得780个可见度数据.

MUSER每3 ms接收1帧一个极化下的16个通道的可见度数据, 在25 ms内完成两个极化下的64个通道的8个数据帧的采集.

3.2 数据预处理

在数据处理过程中, 根据原始观测数据中的可见度数据计算出振幅和相位信息, 作为判断数据是否异常的特征值.

幅值 A 按如下计算公式计算:

$$A = \sqrt{R^2 + I^2}, \quad (4)$$

相位 φ 的计算公式如下:

$$\varphi = \arctan(I/R), \quad (5)$$

其中, R 和 I 分别代表可见度数据的实部与虚部.

根据现场记录, 研究中选择了2014年11月11日的观测数据进行分析与实验, 人工记录当日存在故障的天线编号分别为: 4、7、10、11、12、13、16、17、18、19、24、25、26、36、38、39. 在已知上述故障天线的情况下, 针对MUSER数据特征, 按照一旦天线故障则其对应的基线都标注为异常的原则, 结合LIBSVM软件包对样本格式的要求, 构建了相应的训练集及测试集, 提取出来的样本数据集示例见表1. 由于样本取的都是天线的互相关值, 最终确定样本集特征维有通道channel, 天线ant1, 天线ant2, 振幅 A , 相位 φ . 其中Label是类别标签, 取值+1表示该组天线ant1和ant2都是正常工作的, 取值-1表示其中至少有1面天线发生故障.

表 1 样本数据集示例
Table 1 Examples of sample set

Label	Index1: channel	Index2: ant1	Index3: ant2	Index4: A	Index5: φ
+1	1: 0	2: 0	3: 1	4: 165628.868558	5: 3.126498
+1	1: 0	2: 0	3: 2	4: 59602.612644	5: -2.325843
+1	1: 0	2: 0	3: 3	4: 41734.774589	5: 0.216889
-1	1: 0	2: 0	3: 4	4: 3090.906016	5: 2.454725
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
-1	1: 0	2: 0	3: 7	4: 2007.880724	5: -0.805473
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

由于样本中幅值波动较大, 为保证程序在运行时收敛速度更快, 在进行训练前, 对准备好的训练集进行了归一化处理, 对于同一训练模型下的测试集, 利用保存的规则进行相同幅度缩放.

3.3 模型建立及测试

3.3.1 样本集训练

研究中构建了爆发太阳模型和宁静太阳模型两组训练样本集. 为使样本分布均匀, 两组模型的训练集都是取连续3 min观测数据的前6组(共18帧 780×18 条记录)拼接而成. 爆发太阳的训练集主要分别选取了2014年11月11日12时20分的太阳爆发数据、2014年11月11日12时21分宁静太阳数据、2014年11月11日12时22分宁静太阳数据前6帧, 经过数据预处理后组成一个18帧的样本训练集train_20+21+22_scale; 宁静太阳训练集train_21+22+23_scale则由2014年11月11日12时21分宁静太阳数据、2014年11月11日12时22分宁静太阳数据、2014年11月11日12时23分宁静太阳数据中前6帧合并而成. 相应地, 分别以现有观测数据2014年11月11日12时20-23、25、26分钟数据选取不同帧数作为模型测试集, 测试分类器性能.

训练数据时, 一般要有优化的参数惩罚系数 c 和径向基系数 γ . c 过大或者过小, 都会

使模型的泛化能力变差, γ 的取值则影响能否在特征空间中找到最优超平面. 现有的对于SVM寻参的过程都是基于经验的, 为较准确找到 c 和 γ , 研究中利用网格化处理训练集, 得到爆发太阳组和宁静太阳组最优 c 均为 8 192, 最优 γ 均为 8.

结合寻找到的 c 和 γ , 采用径向基函数RBF, 对两个样本训练集分别进行训练, 最后得到爆发太阳组和宁静太阳组两个模型: train_20+21+22-18_scale.model、train_21+22+23-18_scale.model.

3.3.2 模型测试

对所建模型进行验证, 主要是用到LIBSVM里面的svm-predict文件, 把测试集以参数形式放到模型中, 可以得到通过模型的预测标注文件, 该文件存放的是真正的预测结果. 考虑到通过原始svm-predict进行的预测输出, 只是有一个单一的关于基线数据的类标签, 并没有其对应的天线信息, 因此很难从中看出天线故障情况. 为了使预测输出更适用于MUSER数据, 便于随后对可见度数据进行评估判别, 在实验中对原始svm-predict进行一些修改, 使输出结果中有相应标记类标签的MUSER阵列天线信息.

为了研究太阳活动对模型的影响, 在爆发太阳组模型中, 测试集都是宁静太阳可见度数据; 在宁静太阳组模型中, 有太阳爆发的测试集test_1220_scale和其余宁静太阳的测试集.

为测试小样本数据训练得到的分类器能否对较大样本测试数据进行评估, 在爆发太阳组模型和宁静太阳组模型中, 样本测试集均选取了10帧数据(共780×10条记录)和100帧数据(共780×100条记录), 这里不一一列举.

3.4 结果分析与讨论

3.4.1 方法可行性验证

两组模型下相应测试集完成预测标注. 爆发太阳组模型train_20+21+22-18_scale.model下, 相应的测试集标注结果见表2、表3.

表 2 爆发太阳模型组10帧样本测试准确率

Table 2 Accuracies of 10 frames of test sample sets in the modeling group for solar eruption

Test Sample	Accuracy/(%)	Correct Prediction		Note
		(Sample Size (780×10))		
test_1223-10_scale	84.68	6605		Quiet
test_1225-10_scale	86.24	6727		Quiet
test_1226-10_scale	85.86	6697		Quiet

表 3 爆发太阳模型组100帧样本测试准确率

Table 3 Accuracies of 100 frames of test sample sets in the modeling group for solar eruption

Test Sample	Accuracy/(%)	Correct Prediction (Sample Size (780×100))	Note
test_1223-100_scale	85.35	66573	Quiet
test_1225-100_scale	85.93	67031	Quiet
test_1226-100_scale	86.09	67150	Quiet

宁静太阳组模型train_21+22+23-18_scale.model下, 相应的测试集标注结果见表4、表5.

表 4 宁静太阳模型组10帧样本测试准确率

Table 4 Accuracies of 10 frames of test sample sets in the modeling group for quiet

Sun			
Test Sample	Accuracy/(%)	Correct Prediction (Sample Size (780×10))	Note
test_1220-10_scale	85.91	6701	Burst
test_1225-10_scale	86.59	6754	Quiet
test_1226-10_scale	85.58	6675	Quiet

表 5 宁静太阳模型组100帧样本测试准确率

Table 5 Accuracies of 100 frames of test sample sets in the modeling group for quiet

Sun			
Test Sample	Accuracy/(%)	Correct Prediction (Sample Size (780×100))	Note
test_1220-100_scale	85.65	66807	Burst
test_1225-100_scale	86.39	67384	Quiet
test_1226-100_scale	85.80	66924	Quiet

分析对比表2和表4, 表3和表5, 可以发现在相同样本容量下, 无论是以太阳爆发数据建立的模型还是以宁静太阳可见度数据建立的模型, 训练集预测标注最终正确率均维持在86%左右; 由表4或表5中的数据, 可以看出宁静太阳模型下, 爆发数据测试集和宁静太阳测试集经过模型后的准确率大致相同, 都维持在86%左右.

对表2和表3、表4和表5分别进行对比, 发现容量为10帧的测试样本和容量为100帧的测试样本在18帧模型下预测出来的结果基本稳定在86%左右.

综合以上数据, 再与前期利用VSR等算法得出的标记准确率0.667相比, 我们可以发现, SVM方法基本能够对MUSER可见度数据进行有效评估, 并且该方法大致不受太阳活动的影响. 对于小样本训练得到的SVM模型, 一定范围内可以对相较之大的样本进行

预测标记.

3.4.2 故障天线判断分析

在对可见度数据标注的基础上, 为进一步给出对故障天线的判断, 研究中采用了对故障数据的统计方法. 通过对每一组测试集预测出来的故障标注结果进行统计, 根据各天线的标注比例反推出故障天线的编号. 表6给出了测试集test_1225-10_scale在爆发太阳模型下的统计结果. 其中, 标注百分比 F 按 $F=S/(2 \times T)$ 计算(S 代表样本测试集中单个天线被标注的次数, T 代表样本测试集中所有天线被标注出的总次数).

表 6 爆发太阳模型下2014年11月11日12点25分数据天线标记结果
Table 6 Flagged antennas at 12:25 on November 11, 2014 in the modeling group for solar eruption

Rate of Flagged (Flagged		Rate of Flagged (Flagged	
Antenna No.	Numbers/(Total Flagged	Antenna No.	Numbers/(Total Flagged
	Numbers $\times 2$))/(%)		Numbers $\times 2$))/(%)
0	2.70	20	2.0
1	1.94	21	1.47
2	2.42	22	1.42
3	2.49	23	1.95
4	2.94	24	3.04
5	2.14	25	3.07
6	2.27	26	3.07
7	2.81	27	2.81
8	2.18	28	1.93
9	1.93	29	1.94
10	3.06	30	2.43
11	3.07	31	1.96
12	3.07	32	2.10
13	3.06	33	2.01
14	2.41	34	2.22
15	2.58	35	2.03
16	3.07	36	2.90
17	3.07	37	2.14
18	3.07	38	3.01
19	3.07	39	3.07

研究中, 也对其他模型进行了同样的统计, 结果和表6类似, 通过分析统计结果中每面天线标注的百分比, 得到一个结果, 无论模型是基于太阳活动数据或是宁静太阳数据建立的, 测试样本集是宁静太阳的或是爆发的, 每个表中都出现的一个共性特征是: 编

号为0、4、7、10、11、12、13、15、16、17、18、19、24、25、26、27、36、38、39的天线被标注的百分比均高于2.50. 而已知故障的天线编号为: 4、7、10、11、12、13、16、17、18、19、24、25、26、36、38、39, 通过与已知故障天线编号进行比对, 可以发现通过该方法建立的模型标注出了绝大多数的故障天线, 可以满足对MUSER数据可靠与否的评估. 除此之外, 通过一系列统计判别, 采用SVM方法还能够找到失真数据对应的故障天线, 便于后期的校正工作. 因此, 实验过程中把坏损标记2.5%作为分类的阈值, 坏损标记高于2.5%认为是故障天线, 坏损标记低于2.5%认定为正常工作天线.

以上统计结果充分说明基于统计的故障天线判断方法是可信的.

4 结束语

本文针对中国太阳射电日像仪数据处理要求, 找到了一种基于支持向量机技术的MUSER可见度数据异常的标注方法, 可以获得约86%的较高的准确率, 同时, 也可以在进一步统计的基础上标注出具体的故障天线. 方法已经应用于MUSER的数据处理中, 并取得了较好的效果. 除此之外, 本方法也为射电观测数据在异常检测方面提供了一种新思路.

参考文献

- [1] Yan Y, Zhang J, Wang W, et al. EM&P, 2009, 104: 97
- [2] 高姣姣, 王锋, 戴伟, 等. 天文研究与技术, 2013, 10: 365
- [3] Brown R L, Wild W, Cunningham C. AdSpR, 2004, 34: 555
- [4] Thompson A R, Moran J M, Swenson JR G W. Interferometry and Synthesis in Radio Astronomy. 2nd Edition. New York: John Wiley & Sons, 2001: 12-23
- [5] Swarup G. IJRSP, 1990, 19: 493
- [6] Begum A, Brogan C L, Karachentsev I D, et al. MNRAS, 2008, 386: 1667
- [7] Prasad J, Chengalur J. ExA, 2012, 33: 157
- [8] 肖建华. 智能模式识别方法. 广州: 华南理工大学出版社, 2006: 113-119
- [9] 雷雨, 赵丹宁. 天文学报, 2014, 55: 216
- [10] 高媛媛, 刘强国. 四川理工学院学报: 自然科学版, 2010, 23: 531
- [11] 曾鸣, 林磊, 程文明. 计算机工程与应用, 2013, 49: 7
- [12] Liu C, Bailer-Jones C A L, Sordo R, et al. MNRAS, 2012, 426: 2463
- [13] Qahwaji R, Colak T. SoPh, 2007, 241: 195
- [14] 党建武, 刘云伍, 王阳萍, 等. 计算机应用, 2011, 31: 1010

A Radio Visibility Data Auto-Flag Method Based on Support Vector Machine

DAI Hui-mei¹ MEI Ying^{1,2} WANG Wei³ DENG Hui¹ WANG Feng^{1,2}

(1 Computer Technology Application Key Lab of Yunnan Province, Kunming University of Science and Technology, Kunming 650505)

(2 Yunnan Observatories, Chinese Academy of Sciences, Kunming 650011)

(3 National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012)

ABSTRACT The Mingantu Ultrawide Spectral Radioheliograph (MUSER) has entered the trial observation stage. After the construction of data acquisition and real-time storage system, it is urgent to automatically flag and eliminate abnormal visibility data so as to improve the image quality. In this paper, according to the observational records, we create a credible visibility set, and further obtain a corresponding model by using support vector machine (SVM) technology. The results show that the SVM is a robust approach to flag the MUSER visibility data, and could reach the accuracy of about 86%. Meanwhile, the approach would not be affected by solar activities such as flare eruptions.

Key words sun: activity, sun: radio radiation, stars: imaging, methods: data analysis, techniques: miscellaneous