

自适应中值滤波处理云雷达数据的软件实现

孙罗娜

(简阳市气象局,四川 简阳 641400)

摘要:因为毫米波测云雷达具有更接近小粒子尺度的短波长,更适用于探测弱云,并且具有高时空分辨率,能够更精确地反映出云的结构,所以比普通的天气雷达更适合监测云的变化。毫米波测云雷达在探测时,由于探测环境、非气象因素以及雷达本身噪声的影响,导致雷达回波经常出现很多散点噪声,为了减小这些散点噪声的影响,需要对数据进行滤波。标准的中值滤波法容易实现且去噪效果明显,但这种方法模糊了数据中的尖锐部分和不连续部分,将影响对大气云层的精细结构的分析。自适应中值滤波方法不仅可以剔除散点噪声,并且能够有效地保留数据的局部细节。使用VC实现中值滤波方法和自适应中值滤波方法,对毫米波测云雷达平面位置显示、距离高度显示和固定指向扫描3种方式下的反射率因子、径向速度和速度谱宽进行预处理,并使用Ka波段的毫米波云雷达实测数据进行验证,最后对两种滤波方法的结果进行对比分析。结果表明:相对于传统的中值滤波方法,自适应中值滤波法的滤波性能更好,不仅可以剔除散点噪声,并且能够有效地保留数据的局部细节。

关键词:气象雷达;毫米波测云雷达;中值滤波

1 毫米波测云雷达简介

云的探测对云物理、人工影响天气、气候变化和航空航天等各个领域都有着非常重要的意义,是气象业务与气象学科研究的重点、热点之一。尽管目前已经发展了很多遥感设备可以对云进行观测,但都无法得到高时空分辨率下的云的水平结构和垂直结构,而毫米波雷达具有更接近小粒子尺度的短波长,更适合用来探测弱云,能够准确、迅速、细致的提供时空连续的云体内部的三维气流结构的演变规律,是探测非降水云和弱降水云的有效工具之一^[1]。

毫米波雷达与天气雷达相比,在气象探测上具备以下优点:(1)探测小粒子的能力较强,如:云,沙尘暴和雾等。(2)具有明显的多普勒效应,多普勒速度的分辨率较好,测速度的精度较高。(3)在小的天线孔径下可以得到窄波束,方向性好,空间分辨力极高^[2]。不仅能够穿透云从而描述云内部的物理结构,并且还能够连续监测云的垂直剖面变化,这弥补了其他气象雷达的不足^[3]。毫米波雷达有3种常用的扫描方式:平面位置扫描(PPI)、距离高度扫描(RHI)和固定指向的时间高度扫描(THI)。

2 自适应中值滤波

自适应中值滤波法通过自适应选择窗口大小减小噪声的密度,自适应选择窗口大小的好处是可以处理空间密度更大的脉冲噪声,当噪声为平滑的非脉冲噪

声时,可以保护更多的图像细节,是剔除探测数据中的散点噪声的有效途径之一。

2.1 单窗口确定噪声

滤波算法中最重要的步骤之一是确定噪声点,如果该数据点的强度值大于某一个阈值则判断为噪声点,否则为信号点^[4]。在毫米波测云雷达的数据中,如果一个数据点的强度值并不接近于其领域的数据点值,通常该点就是噪声点;反之,如果一个数据点的强度值很接近于其领域的数据点值,那么该点为信号点。

首先采用大窗口进行噪声检测,然后再根据数据被噪声影响的程度动态选择窗口大小进行噪声剔除。因为这种自适应变化窗口大小的方法只是对检测出来确定为噪声点的数据进行剔除,所以不会明显增加计算时间。

单窗口确定噪声的基本步骤为:假设数据块中位置为 (i,j) 点的数据值为 $R(i,j)$,以这个点为中心,形成一个窗口大小为 3×3 的检测窗口。首先找出该数据点单元内的最大值 $R(i,j)_{\max}$ 以及最小值 $R(i,j)_{\min}$,将与数据点单元内最大值 $R(i,j)_{\max}$ 和最小值 $R(i,j)_{\min}$ 这两个数据点的值不相等的其他数据点构成另外一个新的数据点单元,然后求出这个新的数据点单元内所有数据点的平均值 $R(i,j)_{\text{average}}$ 。再使用标准方法对数据点是否为噪声点进行判断。

标准方法的定义为:如果该点与平均值 $R(i,j)_{\text{average}}$ 的差值大于该数据单元标准差 σ 的3倍,同时该点的值 $R(i,j)$ 与最大值 $R(i,j)_{\max}$ 或最小值 $R(i,j)_{\min}$ 中的一个最值相等,则判定该点为噪声点,否则为

信号点。

$$R(i,j) - R(i,j)_{\text{average}} \geq 3\sigma$$

$$R(i,j) = R(i,j)_{\min} || R(i,j) = R(i,j)_{\max}$$

2.2 单窗口滤除噪声

通过统计数据块中散点噪声的个数,其结果可以对数据块被噪声影响的程度进行评估,从而动态选择滤波窗口的大小。设参数 r 用于对数据块被噪声影响的程度进行评估,该值等于数据块内噪声点的总数与数据块内所有数据点的总数之比。当参数 r 的值比较小的时候,选择较小尺寸的滤波窗口,以确保在滤除噪声点的同时能够保护数据细节;当参数 r 的值比较大的时候,选择较大尺寸的滤波窗口,可以加强去除噪声的能力。根据这个原则判断数据块被噪声点影响的程度,然后自适应选择不同大小的窗口进行滤波,选择窗口大小的标准为

$$l = \begin{cases} 3, & r < r_1 \\ 5, & r_1 \leq r \leq r_2 \\ 7, & r_2 < r \leq r_3 \\ 9, & r_3 < r \end{cases}$$

当数据块被噪声影响的程度超过所设定的门限值时,则根据数据块被噪声影响的程度自适应选择不同大小的窗口进行滤波。由对雷达探测数据的统计与分析可知: $r_1=0.002$, $r_2=0.01$, $r_3=0.05$ 。通过对雷达回波的数据进行处理,得到数据块内噪声点的个数,根据数据块内噪声点的总数与所有数据点的总数之比 r 对数据被噪声污染的程度进行判断,在实验中选取以上几个门限值自适应选取滤波窗口的大小,最终去除噪声点的效果较好。

当数据点经过噪声检测之后,根据标准方法被判定为噪声点和非噪声点,对噪声点则采用自适应中值滤波对其进行滤除,对非噪声数据点的数据不进行处理。选择窗口内几个特定方向上的数据点进行排序,然后对这几个方向上的数据点进行排序后的各个中值进行加权运算,能够较好地保护数据的细节。

对于识别出的每个噪声点,滤波窗口尺寸大小的值为 $L=(2l+1)$,对位置为 (i,j) 的数据点的值为 $R(i,j)$,为了保护数据的局部细节,所以选择滤波窗口内 $0^\circ, 45^\circ, 90^\circ$ 和 135° 这 4 个特定方向上的数据进行排序处理。

以 $l=1$ 为例的滤波求解过程为:某一滤波方向对应的窗口为需要处理的噪声点,由此可得到 4 个窗口,其各个方向上数据加权得到的最终值为

$$S_0(i,j) = \{R(i,j), n \in [j-l, j+l], n \neq j\}$$

$$S_{45}(i,j) = \{R(m,n), m,n \in [j-l, j+l], (m+n = (i+j)), m \neq i, n \neq j\}$$

$$S_{90}(i,j) = \{R(m,j), m \in [i-l, i+l], m \neq i\}$$

$$S_{135}(i,j) = \{R(m,n), m,n \in [i-l, i+l], \text{abs}(m+n-i-j) = 2l, m \neq i, n \neq j\}$$

对以上 4 个窗口的中值滤波结果加权求和后即可得到噪声点滤波后所对应的值,其值为

$$R(i,j)_{\text{filter}} = \sum_{k=0}^{135} C_k S_k(i,j)$$

式中, k 分别为 $0, 45, 90, 135$; $S_k(i,j)$ 是 4 个特定方向上的数据加权后的最终值; C_k 是加权系数, C_k 的值由 4 个特定方向上的数据进行中值滤波后的结果决定,即:

$$C_k = \frac{S_k(i,j)}{\sum_{k=0}^{135} S_k(i,j)}$$

3 毫米波测云雷达滤波技术实现

首先读取基数据文件并判断文件类型,然后根据不同文件类型读取径向数据并提取文件中的反射率、速度、谱宽,识别背景噪声并剔除,再判断散点噪声并分别使用中值滤波和自适应中值滤波进行滤波,最后对两种滤波方法结果进行对比分析。流程图如图 1 所示。

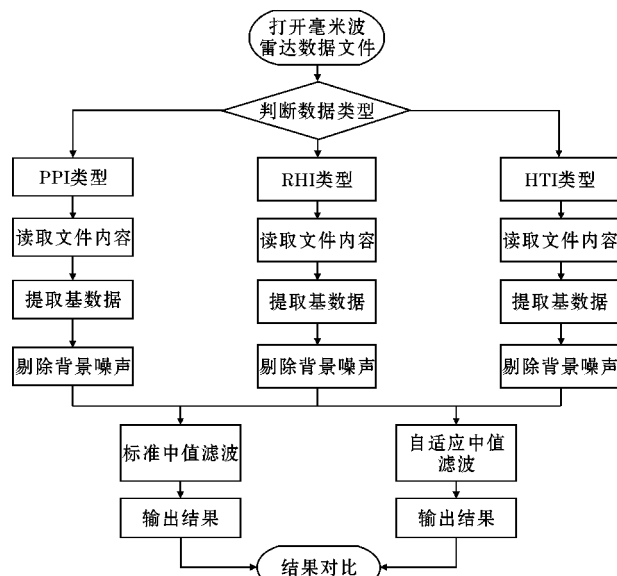


图1 基本流程

文中的数据来自位于成都信息工程大学的地基 Ka 波段磁控管测云雷达,雷达的基本参数如表 1 所示。

表 1 Ka 波段磁控管测云雷达参数

参数	参数值
工作频率	35GHz ± 200MHz
波长	8.6 mm(K 波段)
波束宽度(水平、垂直)	0.2°
脉冲宽度	0.5 μs
脉冲重复率	1500 Hz
距离分辨率	37.5 m
距离库数	1280
盲区	375 m
最大探测距离	48000 m

3.1 毫米波测云雷达的基数据格式及读取

3.1.1 毫米波测云雷达的数据文件类型及判断

平面位置扫描(方位扫描): PPIYYYYMMDDH-HMMSS. CDRPPI。
距离高度扫描(俯仰扫描): RHIYYYYMMDDH-HMMSS. CDRRHI。
固定指向的时间高度扫描: HTIYYYYMMDDH-HMMSS. CDRHTI。

YYYYMMDD 为年月日,HHMMSS 为时分秒,字符的个数表示时间所占的位数。
判断数据文件类型,只需判断出数据文件名称的最后 3 位字母即可。用 C 语言程序实现过程如下:
if(str[length-2] == 'I'&&str[length-3] == 'P'&&str[length-4] == 'P') type=1;
else if(str[length-2] == 'I'&& str[length-3] == 'H'&&str[length-4] == 'R') type=2;
else if(str[length-2] == 'I'&&str[length-3] == 'T'&&str[length-4] == 'H') type=3;
其中, str[] 为文件名, length 为文件名的长度。

3.1.2 毫米波测云雷达的基数据格式及读取

基数据包由基数据头,数据,尾标记 3 个部分组成。基数据头和尾标记的主要作用是存储扫描信息和基数据包校验。反射率、速度、谱宽数据由 DataBuffer 数组存储,数组大小为 MAXRANGBINS * MAXPRONUM,其中 MAXRANGBINS 最大为 1280、MAXPRONUM 最大为 3。
基数据包的结构体定义如下:
typedef struct _CDR_BASEDATA_
{
TagCDRBaseDataHeadHead;
short DataBuff [MAXRANGBINS * MAXPRONUM];
unsigned intEndFlag;

} tagCDRBaseData;

其中, tagCDRBaseDataHead 为基数据头结构体, EndFlag 为文件结束的尾标记,在 DataBuff 中储存一个径向中的反射率、速度、谱宽, BaseData 为基数据头的结构体。
平面位置扫描一圈存储一个文件,文件中为 900 个径向数据。每个基数据包中包含一个径向数据,因此一个 PPI 文件理论上有 900 个基数据包;距离高度扫描储存一个半程扫描的数据;时间高度扫描存储一天的数据,其基数据包的个数由实际工作参数决定。
基数据读取方法如下:

打开文件读取数据后存入指针 p 中,再从 p 中读取每个径向上各个数据库中的值,读取径向数据头标识,并存于 len 中,然后读取每一行的径向个数, p[Radial] 为存有基数据头,反射率、速度、谱宽等的指针,每读取一个径向数据,径向个数进行一次累加。用 C 语言程序实现为:
while(true)
{
fseek(fp, -2 * MoveSize, SEEK_CUR); len = fread(&p[Radial], size, 1, fp);
Radial++;
}

MoveSize 为 int 型数据的大小, fread 表示将文件中的反射率、速度、谱宽数据存入指针 p 中, Radial 表示径向个数。

3.1.3 基数据格式转换

所有反射率、速度、谱宽的数据都存于 p 指针中,数据为 16 位有符号数,通过加权后存储。权值为 0.005。真值=读取值× 0.005,若计算结果大于 100 则为无效值。

3.2 剔除背景噪声

确定背景噪声是滤波算法处理之前的步骤之一,该数据点的强度若大于某一个噪声阈值则为背景噪声点。论文所设门限值为-7800,如果该数据点大于这个门限值则为信号点,否则为噪声点,将数据中小于门限值的数据设置为无效值即可剔除背景噪声。

3.3 固定窗口中值滤波

设定固定窗口大小为 3×3,用径向个数和数据库个数分别除以 3,计算出窗口值,数据不够则补 0,向上取整。使用 for 循环将所有数据分为 3×3 的窗口并进行中值滤波。其步骤为:对窗口内的 9 个数据进行冒泡法排序,如果有效值的个数为奇数,则将排序后的最

中间一个数记为中值,否则将中间两个值的平均值记为中值。

3.4 自适应中值滤波

自适应中值滤波的基本流程如图2所示。

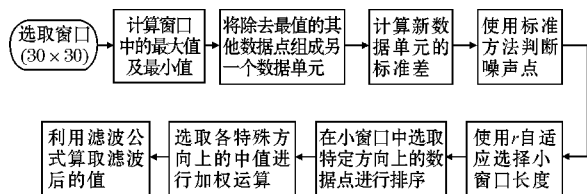


图2 自适应中值滤波法基本流程

根据滤波公式得到自适应中值滤波后的数据:

$$R(i, j)_{filter} = \sum_{k=0}^{135} C_k S_k(i, j)$$

$$C_k = \frac{S_k(i, j)}{\sum_{k=0}^{135} S_k(i, j)}$$

其中, $S_k(i, j)$ 表示加权所得最终值, $k=0, 45, 90, 135$; C_k 为加权系数, 其值由固定窗口的中值滤波对特定方向的数据进行滤波得到, 是某个方向上的中值滤波结果占4个方向上中值滤波结果之和的比重所决定的。

3.5 计算误差参数

完成滤波后, 分别计算标准中值滤波和自适应中值滤波的归一化均方误差 (NMSE)、峰值信噪比 (PSNR)、平均绝对值差 (MAE) 作为误差参数, 并对其结果进行对比。

3个定量参数的定义如下:

$$\sigma_{NMSE} = \frac{\sum_{i=1}^M \sum_{j=1}^M [R(i, j)_{filter} - R(i, j)]^2}{\sum_{i=1}^M \sum_{j=1}^M [R(i, j)]^2}$$

$$R_{PSNR} = 10 \lg \frac{M^2 [R(i, j)_{max}]^2}{\sum_{i=1}^M \sum_{j=1}^M [R(i, j)_{filter} - R(i, j)]^2}$$

$$M_{MAE} = \frac{\sum_{i=1}^M \sum_{j=1}^M |R(i, j)_{filter} - R(i, j)|}{M^2}$$

由公式可知: 峰值信噪比 R_{PSNR} 越大, 平均绝对值差 M_{MAE} 越小, 算法效果越好^[8]。

3.6 滤波结果存入数据文件

对原始数据做了标准中值滤波和自适应中值滤波后, 数据发生了变化, 将滤波后的径向数据中的速度, 谱宽, 反射率3项数据存入文件。使用循环方式存储

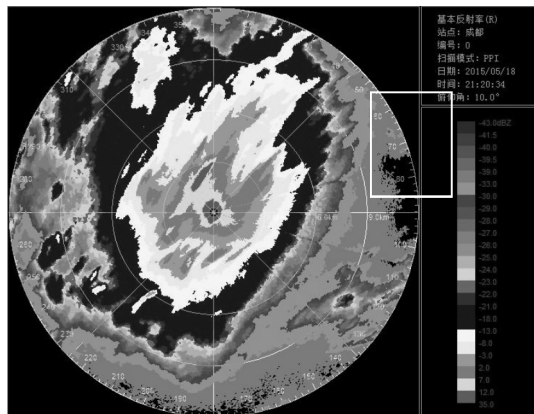
滤波后的数据, 一次存入一个径向内的数据。将最后结果存入指针 p 中。C语言程序实现如下:

```
fwrite(&p[i], p[i].Head.BaseSize, 1, ff);
```

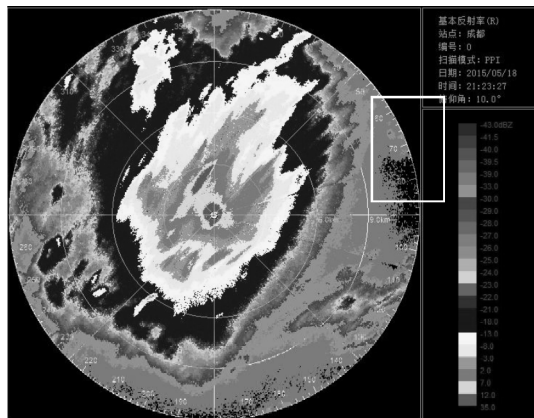
$fwrite$ 表示一次读取一个径向上的数据, 进行数据处理后存入指针 p 中。

4 PPI 实例

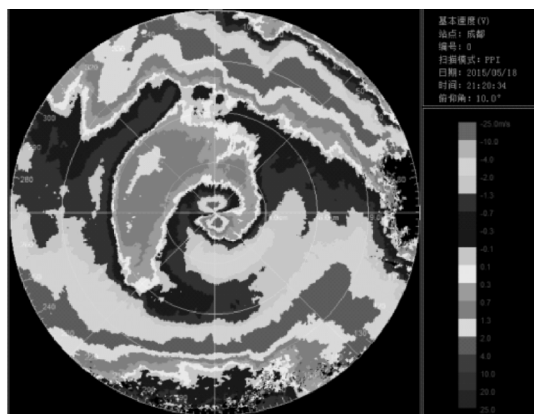
图3是分别根据固定中值滤波和自适应中值滤波后得到的结果。



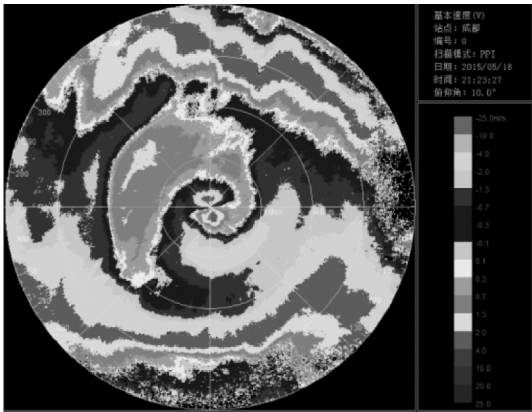
(a) 固定中值滤波后的反射率因子



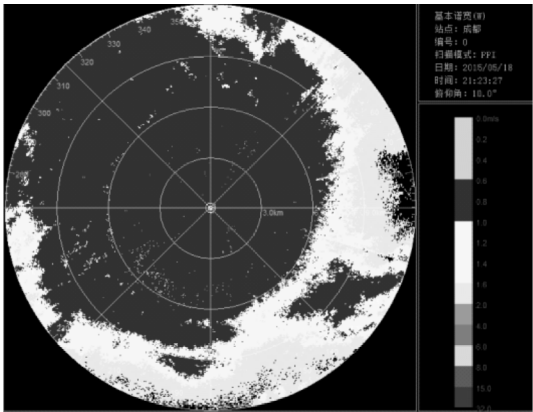
(b) 自适应中值滤波后的反射率因子



(c) 固定中值滤波后的径向速度

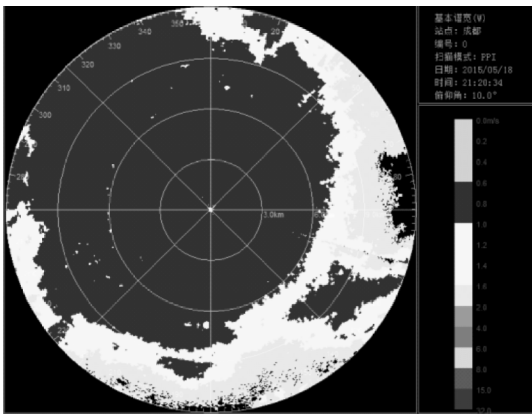


(d) 自适应中值滤波后的速度



(f) 自适应中值滤波后的速度谱宽

图 3 固定中值滤波、自适应滤波后的数据图



(e) 固定中值滤波后的速度谱宽

从图 3 可以直观地看出:常规滤波方法的效果虽然更好一些,却模糊了峰值信息,但是自适应中值滤波可以滤除数据中的噪声点,并保留局部数据。

3 种类型文件进行固定窗口中值滤波和自适应中值滤波后的 3 个定量参量如表 2-表 4 所示。

从以上 3 种扫描方式的定量参量可以看出:自适应中值滤波的归一化均方误差、平均绝对值差比中值滤波更小,峰值信噪比比中值滤波更大,表明自适应中值滤波算法的效果更好。

3 种扫描方式中唯有反射率数据与以上原则不符,尚未找到原因。

表 2 PPI 类型文件参量

	σ_{NMSE}		R_{PSNR}		M_{MAE}	
	固定中值滤波	自适应中值滤波	固定中值滤波	自适应中值滤波	固定中值滤波	自适应中值滤波
反射率	0.009488	0.011983	0.387720	-0.626157	0.905249	1.388781
速度	0.009447	0.005602	-12.436483	-10.167169	0.765252	0.557143
谱宽	0.007501	0.005169	-13.113173	-11.495943	0.687575	0.520178

表 3 RHI 类型文件参量

	σ_{NMSE}		R_{PSNR}		M_{MAE}	
	固定中值滤波	自适应中值滤波	固定中值滤波	自适应中值滤波	固定中值滤波	自适应中值滤波
反射率	0.001908	0.005315	14.580850	10.131861	0.210034	0.515589
速度	0.001603	0.001237	-20.992231	-19.864157	0.176823	0.167439
谱宽	0.001484	0.000991	-6.451509	-4.695626	0.145593	0.100306

表 4 HTI 类型文件参量

	σ_{NMSE}		R_{PSNR}		M_{MAE}	
	固定中值滤波	自适应中值滤波	固定中值滤波	自适应中值滤波	固定中值滤波	自适应中值滤波
反射率	0.007715	0.007781	12.057999	12.021480	0.763413	0.980106
速度	0.005422	0.002460	-25.105461	-21.673954	0.475858	0.250493
谱宽	0.004491	0.001963	-15.681072	-12.087660	0.466069	0.263316

5 结束语

因为毫米波测云雷达具有更接近小粒子尺度的短波长,更适用于探测弱云,具有高时空分辨率,能够更精确地反映云的结构,所以在文中选择毫米波测云雷达监测云的变化。由于探测环境、非气象因素以及雷达本身噪声的影响,导致雷达回波出现散点噪声,为了减小散点噪声的影响,需要对数据进行滤波。标准中值滤波法容易实现且去噪效果明显,但会模糊数据中的尖锐和不连续部分,影响对大气云层的精细结构的分析。自适应中值滤波法不仅可以剔除散点噪声,并且能够保留数据的局部细节。论文中使用 VC 实现了这两种滤波法对毫米波测云雷达 3 种方式下的数据进行处理,并使用实测数据进行验证,最后对两种滤波方法的结果进行对比分析。结果表明:相对于传统的中值滤波方法,自适应中值滤波法的滤波性能更好。

参考文献:

- [1] 王德利,王改利. 自适应中值滤波在云雷达数据预处理的应用[J]. 激光技术,2012,36(2):221-224.
- [2] 仲凌志,刘黎平,葛润生. 毫米波测云雷达的特点及其研究现状与展望[J]. 地球科学发展,2009,26(7):92-93.
- [3] 仲凌志. 毫米波测云雷达系统的定标和探测能力分析及其在反演云微物理参数中的初步研究[J]. 中国气象科学研究. 2009,18(4):2-16.
- [4] X M ZHANG, B Sh XU, Sh Y DONG. Adaptive median filtering for image processing[J]. Journal of Computer-aided Design & Compute Graphice, 2005,17(2):295-299.
- [5] 梁雯,刘松林. 图像中心加权中值滤波的改进与应用[J]. 中国图象图形学,1997,10(9):215-216.
- [6] 柳树,于忠党. 一种改进的开关中值滤波算法[J]. 微计算机应用. 2009,36(9):202-204.
- [7] 邓秀勤,熊勇. 用于图像出路的加权中值滤波算法[J]. 计算机技术与发展,2009,34(3):167-169.
- [8] X K WANG, F LI. Improved adaptive median filtering[J]. Computer Engineering and Applications, 2010,46(3):175-176.