

文章编号:1007-2780(XXXX)XX-0001-07

## 基于离线标定的快速图像拼接算法

于国栋,王春阳,兰孝野,徐鹏宇,张月,刘晓辰,李忠琦  
(中国人民解放军63869部队,吉林白城137001)

**摘要:**针对大场景、大视场图像应用和一般图像拼接算法速度慢的问题,采用了基于离线标定的快速图像拼接算法。在实际应用中多台相机位置固定,采用离线标定计算出图像拼接的单应性矩阵,实时拼接中直接加载该矩阵进行计算,从而省去了大量的特征提取和配准时间。在快速图像对齐的基础上,提出了一种基于YUV色彩空间的拉普拉斯金字塔的快速图像融合算法,并且相对传统拉普拉斯金字塔融合算法有更好的融合效果和更快的速度。实验结果表明,文章算法的图像融合速度是传统拉普拉斯图像融合的3倍。满足应用场景中实时图像拼接速度需求,具有良好的图像拼接效果。

**关键词:**图像拼接;全景监控;特征提取;图像融合

**中图分类号:**TB851+.5; **文献标识码:**A **doi:**10.37188/CJLCD.5-yjyxs2020-0330于国栋(彩)E

### Fast image mosaic algorithm based on offline calibration

YU Guo-dong, WANG Chun-yang, LAN Xiao-ye, XU Peng-yu, ZHANG Yue, LIU Xiao-chen, LI  
Zhong-qi

(63850 Troops of PLA, Baicheng 137001, China)

**Abstract:** Aiming at the problem of large scene, large field of view image applications and the slow speed of traditional image stitching algorithms, a fast image stitching algorithm based on offline calibration is adopted. In practical applications, the positions of multiple cameras are fixed, and offline calibration is used to calculate the homography matrix of image--stitching, which saves a lot of feature extraction and registration time. On the basis of fast image alignment, this paper proposes a fast image fusion algorithm based on the Laplacian pyramid of YUV color space, and it has better fusion effect and faster speed than the traditional Laplacian pyramid fusion algorithm. Experimental results show that The image fusion speed of the article algorithm is three times that of traditional Laplacian image fusion. Meet the real-time image stitching speed requirements in application scenarios, with good image stitching effects.

**Key words:** image stitching; Laplacian pyramid; feature extraction; fast fusion

### 1 引言

对于一些大视场、大场景,特别是近些年比较热门的360度全视场监控等应用中,由于单个

摄像机的视场和功能限制,通常难以够覆盖完整的角度。一般对大视场范围的图像采集或者全景视频监控有两种解决方式,一种是使用广角摄像头;另一种是使用多台固定的摄像机覆盖完整

收稿日期:2020-12-01;修订日期:2021-03-08.

基金项目:军队试验技术研究项目

的监控范围,然后再经过图像拼接获取完整的监控区域<sup>[1]</sup>;但是由于受相机功能和价格的限制,广角摄像机往往难以被采用。为此,需要寻找一种在大视角拍摄情况下,能快速获取高质量画质的图像采集技术。

图像拼接是将在同一时间和地点拍摄的几张具有一定重合区域的照片拼接成一张全景图像的技术,是计算机视觉和图形学领域的研究课题之一。全景图像拼接技术广泛应用于安防监控、战场态势感知、虚拟现实技术和医学图像等多种领域<sup>[1-2]</sup>。传统图像拼接过程一般包含图像预处理、图像配准、图像映射和图像融合等几个步骤<sup>[3]</sup>。

传统的图像拼接算法是基于输入的每一帧图像进行特征提取和计算配准的拼接方法。具体步骤是先提取特征点,然后再进行特征匹配与融合等步骤。但是对于实时的图像拼接场景,传统方法需要消耗大量的空间和时间,因此若是对每一帧图像都进行特征提取和匹配,是无法满足实时拼接的要求。出现鬼影等问题。泊松图像融合算法,虽然融合效果较好,但计算量较大。此外还有基于最佳缝合线和相关的改进的算法,但是都无法满足的实时拼接的要求。传统的图像融合算法中,基于多分辨率的图像融合算法是在RGB三个通道上做拉普拉斯金字塔并实现融合,计算量大。

文章提出了一种基于YUV色彩空间的拉普拉斯金字塔的快速图像融合算法,解决适用于多摄像头合成全景视频拼接难题。系统由多台安放在固定位置的相机组成,具有固定的重合面积,并且使用相同的外部触发,对实时采集到的图像进行拼接。考虑到实际应用场景中相机的角度、位置和重叠的区域固定,所以采用离线标定,在离线标定阶段就计算出相机的映射关系,然后保存到本地。在每次拼接的时候加载映射矩阵,无需再重复进行特征提取和匹配,可以省去拼接过程中大量的前期工作,直接进入图像拼接、融合步骤。

## 2 离线标定

### 2.1 SIFT特征匹配

离线标定的第一阶段是待拼接的图像进

行特征提取和配准。常用的特征提取算法中,基于尺度不变的有SIFT算法<sup>[4]</sup>和SURF等算法<sup>[5-6]</sup>,这些算法各有优缺点。由于在实际应用中使用离线标定方法,所以无需关注在离线标定阶段中特征提取和全局配准所消耗的时间,因此选用效果最好地SIFT算法<sup>[7]</sup>。

传统SIFT特征匹配使用最近邻算法,通过计算特征点之间的欧式距离判断相似度。计算基准图像中某一个特征 $R_i$ 和另外一个图像中所有 $n$ 个特征 $S_1, S_2, \dots, S_n$ 的欧式距离 $d_1, d_2, \dots, d_n$ 后,找出最小值 $d_k$ 和次小值 $d_p$ ( $1 \leq k \leq n, 1 \leq p \leq n, k \neq p$ )。当 $\frac{d_k}{d_p} < T$ 时,则认为 $R_i$ 和 $S_k$ 匹配, $T$ 为阈值。

假设 $R_i$ 和 $S_k$ 是一对配准点的SIFT描述符,SIFT特征向量为128维,即包含128个数据<sup>[8]</sup>,其中 $R_i = (r_{i1}, r_{i2}, \dots, r_{i128})$ , $S_k = (s_{k1}, s_{k2}, \dots, s_{k128})$ ,则 $R_i$ 和 $S_k$ 的欧氏距离计算公式如下:

$$d(R_i, S_k) = \sqrt{\sum_{i=1}^{128} (r_{ij} - s_{kj})^2}, \quad (1)$$

### 2.2 单应性矩阵计算

在摄像机数学模型中,一个三维空间中的点 $P = [X, Y, Z, 1]^T$ 映射到摄像机二维平面上的点 $p = [x, y, 1]^T$ 满足以下关系:

$$p = MP, \quad (2)$$

假设两个摄像机的相机映射矩阵分别为 $M_1$ 和 $M_2$ ,通过这两个相机拍摄的图像中有相同点 $P$ ,而点 $P$ 在两幅图像中的齐次坐标分别为 $p_1$ 和 $p_2$ ,则有下列公式成立:

$$p_1 = M_1^{-1} p_1 = M_2^{-1} p_2, \quad (3)$$

$$p_1 = M_1^{-1} M_2^{-1} p_2 = H p_2, \quad (4)$$

这样,摄像机中相应点的变换就可以用一个 $3 \times 3$ 的单应性矩阵 $H$ 来表示了。两幅图像之间的单应性矩阵 $H$ 可以表示成如下形式:

$$H = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix}, \quad (5)$$

$H$ 为非奇异矩阵,因此对于点 $p_1, p_2$ 用非齐次坐标代替齐次坐标,设 $p_1$ 和 $p_2$ 的非齐次坐标为 $p_1 = [x_1, y_1, 1]^T$ 和 $p_2 = [x_2, y_2, 1]^T$ ,则二维投影变换的关系可表示为:

$$x_1 = \frac{a_{11}x_2 + a_{12}y_2 + a_{13}}{a_{31}x_2 + a_{32}y_2 + 1}, \quad (6)$$

$$y_1 = \frac{a_{21}x_2 + a_{22}y_2 + a_{23}}{a_{31}x_2 + a_{32}y_2 + 1}, \quad (7)$$

由于一对匹配点只能确定2个线性方程,因此理论上需要至少4对匹配点才可以求出投影变换矩阵 $H$ 。对于实际工程中有 $n$ 个点( $n > 4$ )时,最终可利用最小二乘法计算出投影变换矩阵 $M$ 。

### 3 快速图像融合算法

#### 3.1 传统的拉普拉斯金字塔图像融合算法

图像融合是将图像配准之后的待拼接图像在空间上进行叠加。图像融合直接影响拼接后图像的质量。传统的图像融合算法,例如加权平均融合计算量小,可用于快速图像拼接,但融合效果较差,容易出现鬼影等问题<sup>[8-12]</sup>。泊松图像融合算法,虽然融合效果较好,但计算量较大。此外还有基于最佳缝合线和相关的改进的算法,但是都无法满足的实时拼接的要求。传统的图像融合算法中,基于多分率的图像融合算法是在 $RGB$ 三个通道上做拉普拉斯金字塔并实现融合,计算量大。

本文采用一种改进的多分率图像融合方法,提出在 $YUV$ 色彩空间中进行拉普拉斯金字塔图像融合。该方法可以减少大量的计算,并且图像融合效果更佳。

(1)首先构造高斯金字塔,将原始图像作为金字塔的最底层图像 $G_0$ ,使用高斯卷积核 $g$ 对图像进行卷积,然后再下采样得到上一层图像 $G_1$ 。重复进行卷积核下采样操作 $i$ 次即可获取第 $i$ 层图像 $G_i$ ,迭代多次,即可形成图像高斯金字塔。

$$G_{i+1} = DOWN(G_i \otimes g_{3 \times 3}), \quad (8)$$

其中 $g_{3 \times 3}$ 是高斯卷积核。 $DOWN$ 表示下采样计算。

(2)构造拉普拉斯金字塔,在构建高斯金字塔的过程中,持续的卷积核下采样会造成图像高频信息的丢失,所以使用拉普拉斯金字塔来保存图像的高频信息。将上一层的高斯金字塔图像进行上采样后再进行卷积操作,使用该层的高斯金字塔图像减去卷积后的图像即可得到当前层的拉普拉斯金字塔图像

拉普拉斯金字塔中第 $i$ 层图像:

$$L_i = G_i - UP(G_{i+1}) \otimes g_{3 \times 3}, \quad (9)$$

其中: $L_i$ 是拉普拉斯金字塔中第 $i$ 层的图像, $G_i$ 是高斯金字塔中第 $i$ 层的图像, $UP$ 表示上采样计算。

(3)融合后恢复图像,处理的过程需要对每一层金字塔实融合。处理的过程从上到下,最终获得原始图像。

完成每一层的图像融合之后,要根据融合后的拉普拉斯金字塔重构高斯金字塔,并最终得到实际融合后的图像,计算公式如下:

$$G_i = L_i + DOWN(G_{i+1}) \otimes g_{3 \times 3}, \quad (10)$$

#### 3.2 改进的拉普拉斯金字塔图像融合算法

##### (1) 色彩空间转换

$RGB$ 和 $YUV$ 色彩空间的互相转换,公式如下:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.14713 & -0.28886 & 0.436 \\ 0.615 & -0.51499 & -0.10001 \end{pmatrix} \quad (11)$$

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1.113983 \\ 1 & -0.39465 & -0.58060 \\ 1 & 2.03211 & 0 \end{pmatrix} \quad (12)$$

##### (2) 处理 $Y$ 通道

$YUV$ 色彩空间中, $Y$ 通道表示像素点的亮度信息并反映图像中灰度值的动态范围。 $YUV$ 色彩空间将亮度和颜色分开, $Y$ 部分指标是融合后图像的亮度信息, $UV$ 通道表示颜色信息。在图像融合中,高对比度的区域应给更高的权重,所以拉普拉斯金字塔可以很好地保留高对比度的区域。

拉普拉斯图像融合,其实就是权重融合,本文从两个方面考虑权重,对比度因子和颜色深度因子:

对比度因子公式如下:

$$C(i, j)_k = \left| Y_k(i+1, j) + Y_k(i-1, j) + Y_k(i, j+1) + Y_k(i, j-1) - 4Y_k(i, j) \right|, \quad (13)$$

其中: $Y_k(i, j)$ 是第 $k$ 层中坐标为 $(i, j)$ 的 $Y$ 通道值。

颜色深度因子 $E_k(i, j)$ 公式如下:

$$E_k(i, j) = \left| U_k(i, j) \times V_k(i, j) + \epsilon \right|, \quad (14)$$

其中： $U_k(i, j)$ 和 $V_k(i, j)$ ，分别表示第 $i$ 层输入图像的在坐标 $(i, j)$ 处的 $U$ 通道和 $V$ 通道的值， $\epsilon$ 是一个很小的值，避免权重为0。

最终的实际权重为：

$$W_k(i, j) = C_k(i, j) \times E_k(i, j), \quad (15)$$

使用这个权重来完成图像融合，最终可得到融合后的 $Y$ 通道图像。

(3) 处理 $UV$ 通道

$$U(i, j) = \text{Max}_{1 < k \leq N} U_k, \quad (16)$$

$$V(i, j) = \text{Max}_{1 < k \leq N} V_k, \quad (17)$$

根据上述公式获取融合后 $UV$ 通道的值，即相同像素点位置，选择两个中较大值的 $U$ 和 $V$ 。最后将融合后的 $YUV$ 色彩空间图像转换为 $RGB$ 色彩空间的图像即可。

## 4 实验验证

为验证文章方法的有效性，在场区生活区选取一个宽阔场地验证实际拼图效果。文章选用的图像获取设备为华为Mate20，该手机内置镜头为徕卡三摄镜头，采用“广角+超广角+5倍长焦距”的组合。其中，广角镜头1200万像素( $f/1.8$ 光圈，等效焦距27 mm)，超广角1600万像素( $f/2.2$ 光圈，等效焦距17 mm)，长焦镜头800万像素( $f/2.4$ 光圈，等效焦距52 mm)。程序运行环境Visual Studio，算法实现工具OpenCV4.1.0，操作系统是Windows10，CPU配置为Corei7-8750, 2.20GHz。

在拍摄场景前，随机选取两个拍摄位置，相距大约5m，拍摄图像并保存。

图像处理分两个阶段：离线标定阶段和实时拼接阶段。这两个阶段的主要任务和工作如下：

(1) 离线标定阶段：

离线标定阶段主要包括特征提取、特征匹配和单应性矩阵的计算。

离线标定阶段的任务是提取特征点和进行特征匹配，最终根据匹配的结果计算出单应性矩阵。该矩阵在实时拼图时加载，直接用于图像映射，可节省大量的特征提取和配准时间。

(2) 实时拼接阶段：

实时拼接阶段包括加载单应性矩阵计算和

图像融合。



图1 待拼接图像

Fig. 1 Diagram of images to be spliced

图1为待拼接的两张图像，分别取全景图像的左右两部分，经特征匹配和校准后，得到配准的特征点。如图2所示，绿色圆圈表示SIFT特征提取和配准后得到的特征点。使用这些配准的特征点计算得到透视变换矩阵为：

$$\begin{bmatrix} 0.478139 & -0.014861 & 663.436953 \\ -0.348395 & 0.879932 & 100.616037 \\ -0.000480 & -9.830015 & 1 \end{bmatrix} \quad (18)$$



图2 SIFT算法特征配准结果

Fig. 2 Result of SIFT algorithm feature registration

新建两张空图 $L$ 和 $R$ ，高度与原图相等，宽度是原图的两倍，将待拼接的左图拷贝到图 $L$ 中待拼接的右图经单应性矩阵计算后投影到图 $R$ 中，得到对齐后的图像，如图3所示。图4为加权平均融合算法得到的结果，从拼接效果可以看出，使用SIFT特征提取和配准之后的特征点计算得到的单应性矩阵可以使图像获得较好的对齐效果。

虽然使用该单应性矩阵有较好的图像对齐效果，但是也存在一定的缺点，经过矩阵运算对

齐后的图像,其右侧相对于原图缺少了一部分。产生这种情况的主要原因为:右侧的点经过单应性矩阵计算后其坐标超出目标图像矩阵的范围,导致超出范围的这部分被抛弃了。在实际应用中,如果图像侧边不重要的话,可以忽略。但如果侧边图像有必要保留,可通过修改单应性矩阵的方式保留该部分图像信息。

单应性矩阵公式如下:

$$\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \quad (19)$$

其中子矩阵  $\begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \end{bmatrix}$  表示图像的旋转信息,  $\begin{bmatrix} x_{13} & x_{23} \end{bmatrix}^T$  表示图像的平移信息,  $x_{13}$  表示横轴方向上的偏移,  $x_{23}$  表示纵轴方向上的平移。要保留图像的全部信息只需要使用该单应性矩阵计算原图最右侧两点变换后的像素坐标,即可得到图像在横轴的偏移值,使用该偏移值修改原单应性矩阵即可保留图像的全部信息。



图3 图像对齐

Fig. 3 the diagram of image alignment

图像对齐后,分别采用加权平均算法、改进的渐入渐出算法、传统拉普拉斯金字塔融合算法和文章算法,对图像进行融合处理,其效果如图4-7所示。相对于使用加权平均融合算法,无论是传统的拉普拉斯图像融合还是文章算法都有更好的融合效果。加权平均融合算法和改进的渐入渐出算法在一些细微之处无法将两幅图像中的同一个物体融为一体,而是当作两个物体同时保留在图像上,融合效果较差,而拉普拉斯算法确可以识别并做了很好的融合。

使用3个不同大小的融合区域(320\*1 440、430\*1 440、730\*1 440)进行多次融合取处理时长的平均值,比较3种算法的融合速度。如表1所示,加权平均融合算法虽然有较快的处理速度,但是融合效果不佳。改进的渐入渐出图像融合方法相对于加权平均融合算法图像拼接效果更好,但是也会带来更复杂的计算。传统拉普拉斯算法和本文算法处理效果几乎一致,但是处理效率差别较大,文章算法的处理时间明显快于传统拉普拉斯金字塔融合算法,处理时间提高了

近3倍。



图4 加权平均融合算法

Fig. 4 the diagram of weighted average fusion algorithm

可见,相比于传统拉普拉斯金字塔融合算法,文章的方法在不降低处理质量的前提下,处理速度更快,效率更高,针对处理海量高清图像处理时,更能体现文章算法的优越性。

## 5 结论

快速图像拼接中采用离线标定的方法可节



图5 改进渐入渐出图像融合算法

Fig. 5 the diagram of improved gradual in gradual out image



图7 YUV色彩空间拉普拉斯融合

Fig. 7 the diagram of Laplacian fusion in YUV color space



图6 RGB色彩空间拉普拉斯融合

Fig. 6 the diagram of Laplace fusion in RGB color space

表1 文章算法与传统拉普拉斯金字塔融合算法处理速度对比结果

Tab. 1 the results of speed comparison between the two algorithms

图像大小	Unit:s		
	320*1440	430*1440	730*1440
加权平均	0.012	0.016	0.022
改进渐入渐出	0.019	0.023	0.031
传统算法	0.053	0.067	0.103
文章算法	0.016	0.025	0.029

省大量计算时间。在传统拉普拉斯金字塔图像融合基础上,提出基于YUV色彩空间的拉普拉斯金字塔图像融合,将图像亮度和颜色分离,只对Y通道做拉普拉斯金字塔图像融合,相对传统

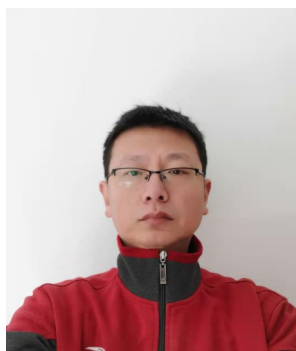
拉普拉斯图像融合速度提升了近3倍,简化实现过程,也同样可以获取较好的图像融合效果。

#### 参考文献:

- [1] 杜往泽,宋执环,闫文博,等. 单摄像头旋转监控下的快速图像拼接[J]. 中国图象图形学报, 2016, 21(2): 245-254.  
DU W Z, SONG Z H, YAN W B, *et al.* Fast image stitching under single camera rotate monitoring [J]. *Journal of Image and Graphics*, 2016, 21(2): 245-254. (in Chinese)
- [2] 谷雨,周阳,任刚,等. 结合最佳缝合线和多分辨率融合的图像拼接[J]. 中国图象图形学报, 2017, 22(6): 842-851.  
GU Y, ZHOU Y, REN G, *et al.* Image stitching by combining optional seam and multi-resolution fusion [J]. *Journal of Image and Graphics*, 2017, 22(6): 842-851. (in Chinese)
- [3] 杨春德,成燕菲. 基于拼接缝自适应消除和全景图矫直的快速图像拼接算法[J]. 计算机应用, 2019, 39(10): 3053-3059.  
YANG C D, CHENG Y F. Fast image mosaic algorithm based on adaptive elimination of stitching seam and panorama alignment [J]. *Journal of Computer Applications*, 2019, 39(10): 3053-3059. (in Chinese)
- [4] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [5] 李倩,江泽涛. 二值化的SIFT特征描述子及图像拼接优化[J]. 中国图象图形学报, 2016, 21(12): 1593-1601.  
LI Q, JIANG Z T. Binary quantized SIFT feature descriptors for optimized image stitching [J]. *Journal of Image*

- and Graphics*, 2016, 21(12): 1593-1601. (in Chinese)
- [6] 雍玫,王阳萍,雷晓妹,等. 基于改进ORB特征检测的全景视频拼接[J]. 计算机应用与软件,2017,34(5): 182-188.  
YONG J, WANG Y P, LEI X M, *et al.* Panoramic video stitching method based on improved orb feature detection [J]. *Computer Applications and Software*, 2017, 34(5): 182-188. (in Chinese)
- [7] 徐鑫,孙韶媛,沙钰杰,等. 一种基于改进RANSAC的红外图像拼接方法[J]. 激光与光电子学进展,2014,51(11):111001.  
XU X, SUN S Y, SHA Y J, *et al.* A method of infrared image mosaic based on improved RANSAC [J]. *Laser & Optoelectronics Progress*, 2014, 51(11): 111001. (in Chinese)
- [8] 宋建辉,闫蓓蕾. 基于SIFT的无人机航拍图像快速拼接技术研究[J]. 计算机应用与软件,2018,35(2):230-234.  
SONG J H, YAN B L. Research on rapid mosaic technology of UAV aerial image based on SIFT [J]. *Computer Applications and Software*, 2018, 35(2): 230-234. (in Chinese)
- [9] 作建宁,郭宝龙,冯宗哲. 一种基于兴趣点匹配的图像拼接方法[J]. 计算机应用,2006,26(3):610-612.  
WU J N, GUO B L, FENG Z Z. An image mosaic method based on interest points matching [J]. *Computer Applications*, 2006, 26(3): 610-612. (in Chinese)
- [10] 罗永涛,王艳,张红民. 结合最佳缝合线和改进渐入渐出法的图像拼接算法[J]. 红外技术,2018,40(4): 382-387.  
LUO Y T, WANG Y, ZHANG H M. Image-stitching algorithm by combining the optimal seam and an improved gradual fusion method [J]. *Infrared Technology*, 2018, 40(4): 382-387. (in Chinese)
- [11] 李冰,李范鸣,谢江荣,等. 基于红外面阵扫描图像的配准和融合[J]. 激光与红外,2017,47(2):246-251.  
LI B, LI F M, XIE J R. Image registration and fusion based on infrared array scanning detector [J]. *Laser & Infrared*, 2017, 47(2): 246-251. (in Chinese)
- [12] 郑悦,程红,孙文邦. 邻域最短距离法寻找最佳拼接缝[J]. 中国图象图形学报,2014,19(2):227-233.  
ZHENG Y, CHENG H, SUN W B. Finding an optimal seam-line through the shortest distance in the neighborhood [J]. *Journal of Image and Graphics*, 2014, 19(2): 227-233. (in Chinese)

#### 作者简介:



于国栋(1982-),男,辽宁大连人,硕士,工程师,2013年于武汉大学获得硕士学位,主要从事弹道终点坐标测试方面的研究。E-mail: daya9527@126.com