**《数字图像处理》课程论文**

**(2020-2021学年第1学期)**

**论文题目：基于APSO自动搜参与肤色的人脸检测算法**

**学生姓名： 黎羿江**

**提交日期： 年 月 日 学生签名：**

|  |  |  |  |
| --- | --- | --- | --- |
| **学 号** |  | **座位编号** |  |
| **学 院** |  | **专业班级** |  |
| **课程名称** |  | **任课教师** |  |
| **教师评语：** | | | |
| **本论文成绩评定：** **分** | | | |

**基于APSO自动搜参与肤色的人脸检测算法**

**学生姓名（宋体，四号字，加粗）**

**摘要：**本文设计实现了一个图片处理系统，实现对照片的人脸检测，彩色图片转换为素描图片，彩色图片油画效果，水彩效果以及点画艺术的转化。传统基于人脸肤色的人脸检测算法在进行预处理后往往基于先验知识对候选区域进行过滤，本文提出基于高斯概率函数的肤色预测，同时基于自适应粒子群算法（APSO）对过滤器进行自适应超参数搜索，优化检测结果。

**关键词：**图像处理 人脸检测 肤色检测 粒子群算法 超参数搜索

# 1、引言

数字图像处理是指将图像信号转换成数字信号并利用计算机对其进行处理的过程。随着人工智能的兴起，以计算机视觉为代表的图像处理蓬勃发展，一系列图像处理软件如photoshop，天天p图等等相继面世。作为图像处理的一大组成部分，人脸检测是人脸识别，跟踪，身份验证等技术的前提，是图像识别的重要研究方向，越来越为人们所重视。Yang在文献中给出人脸检测的定义：对于一张给定的图像，人脸检测确定图像中是否有人脸的存在，如果有则给出人脸在图像中的位置。出于人脸的高度变化性，人脸检测存在一系列的挑战。人脸图像的不同光照，人脸的不同肤色，表情，胡须以及眼睛，头发以及其他物体等等各种遮挡都给人脸检测带来一系列困难。肤色作为人脸的主要特征部分，特征面积大，颜色信息各向同性，可以很好的反应人脸的位置，因此成为人脸检测的重要手段。

本文在此基础上，提出基于高斯的人脸肤色判断以及基于自适应粒子群算法（APSO）的过滤连通域超参数搜索的人脸检测算法，对原有的基于肤色的人脸算法进行优化改进，并且达到一个较优的效果。另外，本文设计并实现了一个图像处理系统与应用，实现了一系列图像转化算法以及上述人脸检测算法。本文第一部分主要介绍提出的基于高斯的人脸肤色判断以及基于自适应粒子群算法（APSO）的过滤连通域超参数搜索的人脸检测算法以及实现，第二部分简要说明本文设计并实现的图像处理系统。

# 基于APSO自动搜参与肤色的人脸检测算法

基于肤色的人脸检测算法通过统计模型判断每个像素点是否是人脸，从而划分出人脸区域，一般流程包括：图像的预处理，像素肤色判断，连通域过滤，在利用统计模型进行像素级肤色判断后，可以得到一个二值图，其中0代表这个像素并非人脸，1代表这像素为人脸，通过连通域搜索算法得到连通域作为人脸候选区域，然后通过一系列指标如空洞面积比，长宽比以及面积密度等等筛选过滤非人脸区域，得到最终的人脸区域后画出人脸边框。本文采用高斯模型对像素进行肤色判断，并且提出一种基于APSO的指标阈值自动搜索算法对连通域过滤的一系列指标阈值进行搜索。算法在给出数据集的训练集上进行高斯模型的训练，并在其验证集合上进行阈值超参搜索。具体流程如下：

加图

## 图像预处理

图像分析中，图像质量的好坏直接影响识别算法的设计与效果的精度，因此在图像分析（特征提取、分割、匹配和识别等）前，需要进行预处理。图像预处理的主要目的是消除图像中无关的信息，恢复有用的真实信息，增强有关信息的可检测性、最大限度地简化数据，从而改进特征提取、图像分割、匹配和识别的可靠性。我们进行了高斯滤波，正则化，图像锐化增强

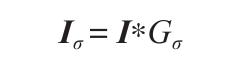
**1.1高斯滤波**

滤波器分为很多种，有方框滤波、均值滤波、高斯滤波等。高斯滤波是一种线性平滑滤波，适用于消除高斯噪声。

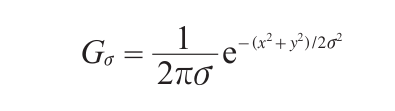
**1.1.1高斯噪声**

首先，噪声在图像当中常表现为一引起较强视觉效果的孤立像素点或像素块。简单来说，噪声的出现会给图像带来干扰，让图像变得不清楚。  
高斯噪声就是它的概率密度函数服从高斯分布（即正态分布）的一类噪声。如果一个噪声，它的幅度分布服从高斯分布，而它的功率谱密度又是均匀分布的，则称它为高斯白噪声。高斯白噪声的二阶矩不相关，一阶矩为常数，是指先后信号在时间上的相关性。高斯滤波器是根据高斯函数的形状来选择权值的线性平滑滤波器  
**1.1.2 高斯函数和高斯核**

理论上，高斯分布在所有定义域上都有非负值，这就需要一个无限大的卷积核。实际上，仅需要取均值周围3倍标准差内的值，以外部份直接去掉即可。



其中 \* 表示卷积操作; Gσ 是标准差为σ 的二维高斯核,定义为:



**1.2图像正则化**

**1.3图像锐化**

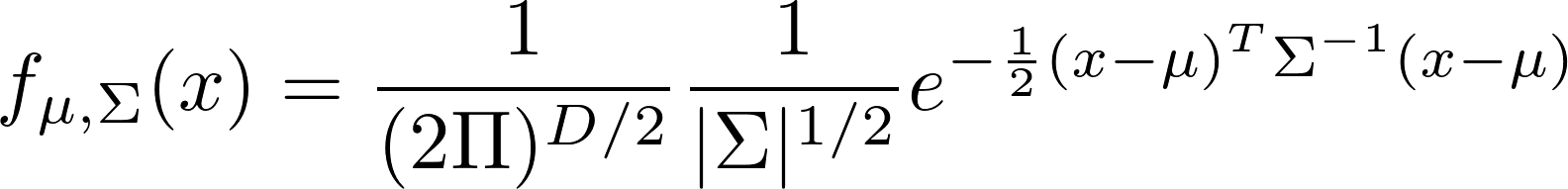
图像锐化是一种突出和加强图像中景物的边缘和轮廓的技术。图像锐化的本质是增加邻域间像素的差值，使图像的突变部分变得更加明显。图像的卷积计算除了可以完成模糊去噪、边缘检测等任务外，还可以实现图像锐化/增强的功能。一般也通过Laplacian滤波加原图权重像素叠锐化空间滤波器用来增强图像的突变信息，图像的细节和边缘信息。平滑滤波器主要是使用邻域的均值（或者中值、积分）来代替模板中心的像素，消弱和邻域间的差别，以达到平滑图像和抑制噪声的目的；模糊图像，称为低通滤波器。锐化滤波器则使用邻域的微分作为算子，增大邻域间像素的差值，使图像的突变部分变的更加明显。锐化的作用是加强图像的边沿和轮廓，通常也成为高通滤波器

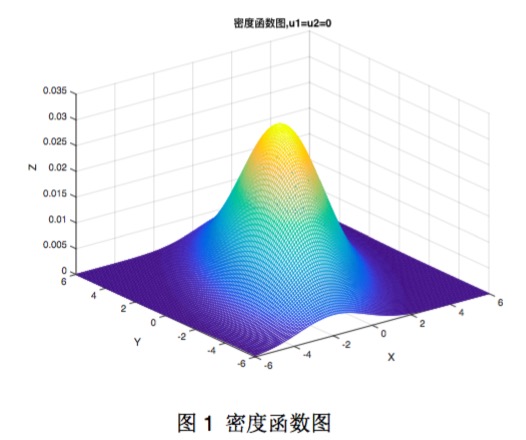
Butterworth通滤波器函数为：



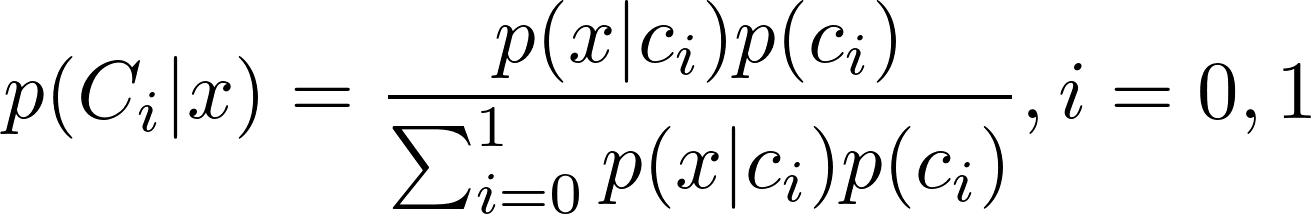
## 基于高斯朴素贝叶斯函数的肤色判断

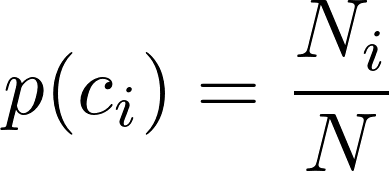
本模型通过高斯概率模型与朴素贝叶斯进行像素级肤色判断，模型读入人脸像素的YCbCr值，输出一个范围在0-1之间的概率值，概率越接近1，此像素为人脸像素的可能性越高，反之，此像素为背景像素的可能性越高。模型对于每一类（背景或人脸）训练一个高斯概率分布：





其中D为特征数，/private/var/folders/g6/0qn_9msd2vl_8dwl99wcfwq80000gn/T/com.kingsoft.wpsoffice.mac/wpsoffice.lrESxcwpsoffice，wpsoffice为多元高斯分布的参数。在本文中D=3。对于这个高斯概率分布，输入一个点的YCbCr，返回这个点属于此类的概率。为了使得高斯概率模型尽可能拟合数据集，我们将训练集合中两个类的点分别训练两个高斯概率模型，通过极大似然估计求得两个高斯概率模型的/private/var/folders/g6/0qn_9msd2vl_8dwl99wcfwq80000gn/T/com.kingsoft.wpsoffice.mac/wpsoffice.lrESxcwpsoffice，wpsoffice，分别为类内点的协方差矩阵与均值。通过朴素贝叶斯确定一个点最终的概率：



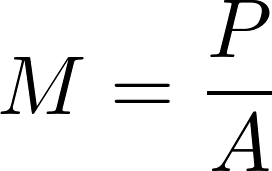


其中wpsoffice为第i个类，x是给点像素点的YCbCr向量，在本文中i为0或1分别代表背景与人脸。wpsoffice由第i个类的高斯概率分布给出，为wpsoffice为第i个类占比。

## 连通域过滤

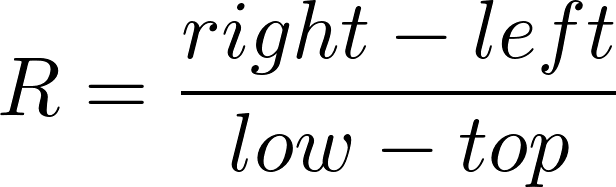
在得到像素级的概率图后，对概率大于0.5的点预测为人脸点，反之背景点，得到二值图。通过DFS或者BFS得到八连通域。对于每一个连通域，设计了一系列指标用于筛选：

* 面积密度



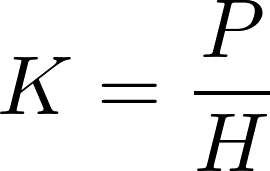
其中P为连通域像素数量，A为连通域外接矩阵面积。考虑区域为矩形时，M趋近于1，为圆形时M约为pi/4，当区域为细长物体时，比例变小，通过设定合理的阈值即可去除细长弯曲与细长倾斜的区域。

* 长宽比



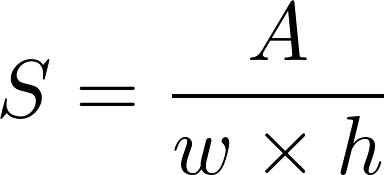
通过计算最小内接矩形的长宽比，可以清楚过于扁平的物体。

* 空洞密度



其中P为连通域的像素数量，H为连通域所占面积（注意区分H与A，H为整个连通域所占面积，包括内部空洞，而A为连通域的内接矩阵）。通过计算像素值与所占面积之比，可以估算连通域中空洞的大小。考虑人脸存在眼睛，嘴巴与鼻子等部位，这些部位可能不会被预测为人脸，因此在二值图中，人脸区域应该存在一定空洞。但是一些其他区域也可能因为光照变化等原因，产生内部空洞，因此需要设立合适阈值进行过滤。

* 面积占比



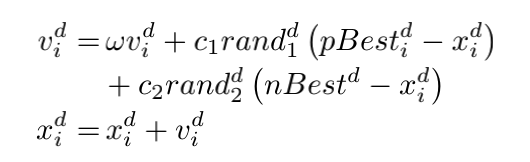
其中A为连通域最小外接矩阵面积，w为图像宽度，h为图像长度。我们考虑人脸在图像中应该占据一定大小的位置，所以过于小的候选连通域不作为人脸进行考虑。

在经过一系列过滤后，得到过滤后的所有人脸连通域，将其最小外接矩形作为人脸边框。

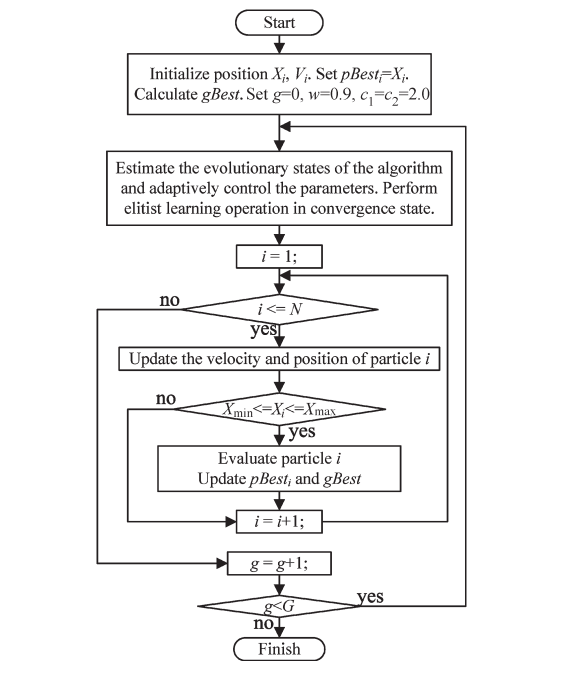
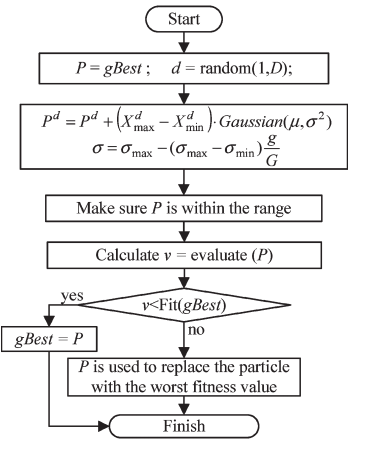
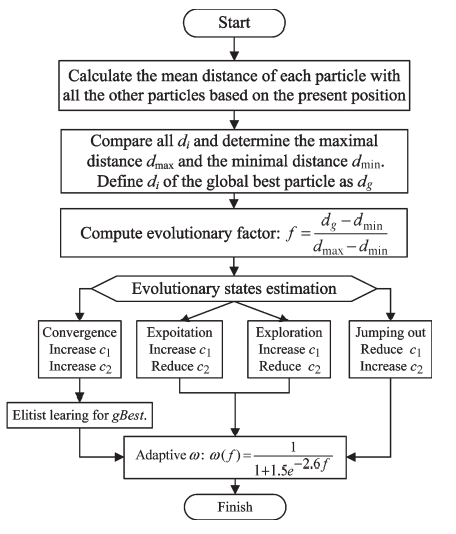
## 基于自适应粒子群算法对连通域过滤超参数的搜索

上述连通域过滤算法中存在一系列指标阈值超参数需要确定，传统方法通过人为经验进行设定，这种方法对不同数据集效果不一，不能自适应调整超参数。因此，在本文中提出基于自适应粒子群算法对指标阈值超参数进行自动搜索，以获得最佳超参数。通过观察鸟群社会行为，发现群体中对信息对社会共享提供演化优势，并以此为基础建立粒子群算法。粒子群算法通过学习群体最优值与个体最优值更新粒子，从而搜索到一个较好到参数。詹于2009年提出的APSO（adaptive particle swarm optimization）通过不断更新粒子群的w，c1与c2，能够自适应搜索最优参数。APSO适用于较小维度的参数（论文中一般在30维左右），同样适用于本问题。因此，本文考虑采用APSO自适应搜索过滤连通域的阈值超参数。

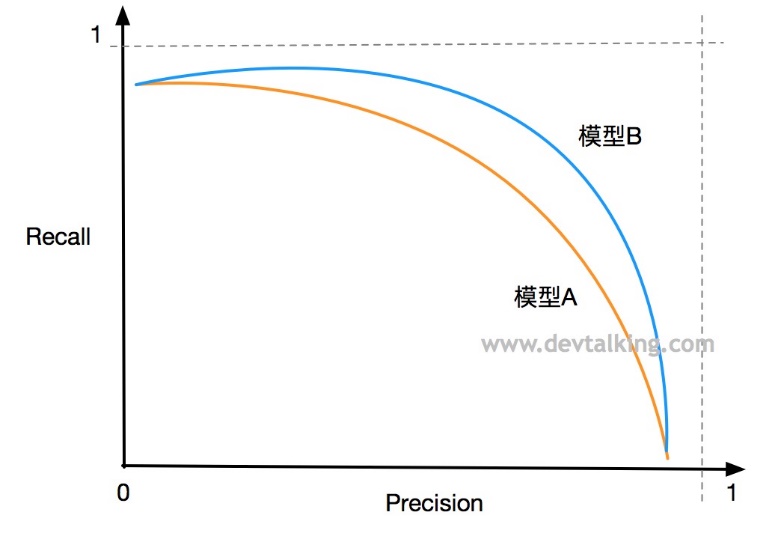
APSO采用常规PSO框架，通过个体最优值以及全局最优值更新粒子。



创新性的是，APSO引入ELS以及ESE模块。ELS模块通过跳出局部最优值引导种群，ESE自适应调整w，c1和c2。一下示意图分别为ESE，ELS模块示意图以及APSO算法流程图。



对于一个给定的粒子群算法，需要为其参数定义评估函数。在本文题中，则是如何评估一组过滤连通域的阈值超参的好坏。由于人脸检测问题本质其实是一个目标检测问题，我们采用目标检测的评估指标mAP（mean average precision）对一组超参数的效果进行评估，由于只有人脸一个类别，实际上mAP退化为AP（average precision）。AP通过计算验证集上所有图像的总的准确率（正确的预测框占所有预测框的比例）与对应召回率（正确预测框占所有标签框的比例），画出P-R曲线并计算曲线下方面积得到。准去率与召回率是一组相互权衡的指标，在不同的准确率下将有不同的召回率。



# 图像处理系统与应用

除了设计与实现基于APSO自动搜参与肤色的人脸检测算法外，我们同时设计并实现了一个图像处理系统与应用，这个应用以微信小程序为前端，以flask框架作为后端，实现了一系列图片转换算法，包括照片的人脸检测，三种彩色图片转换为素描图片算法，彩色图片油画效果算法，水彩效果以及点画艺术的转化算法。

介绍彩图转素描算法的详细过程，并且贴图，微信小程序前端app图以及各种效果图

出于篇幅原因，本文只说明彩色图转素描算法设计。素描，也就是绘制轮廓，本质就是边缘检测。不过本文所述的算法并没有使用常见的边缘检测方法如sobel算子等，而是通过前向差分（forward difference）计算图像梯度（image gradient）来寻找边缘。

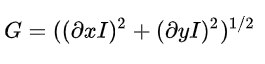
# 3.1前向差分和图像梯度

简单地说，给定一个数列，那么它的前向差分的结果就是数列中每个数的后一个数与当前数字的差。



由于图像是一个矩阵（二维单通道灰度图像），因此图像I的前向差分要分为x轴和y轴方向两种，也可以称之为沿x轴和沿y轴的图像梯度，分别用符号  和  表示。IMG_261 是和原图同样大小的矩阵，每一列都是原图中下一列与当前列的差值，最后一列保持不变。IMG_262也是和原图等大小的矩阵，每一行都是原图中下一行与当前行的差值，最后一行保持不变。

因此，  中，所有垂直边缘会凸现出来，而  中，所有水平边缘会凸现出来。我们只需要将这两个梯度矩阵合并就可以得到整体的轮廓。合并公式如下：



# 3.2图像梯度 vs 边缘检测算子

相比于卷积式边缘检测（即使用算子对图像进行卷积，配合恰当的阈值来判断哪些点是边缘），图像梯度计算出的边缘相对更“嘈杂”，更符合我们绘画时的真实情况。

# 3.3降噪

上文所说的图像梯度G的质量很容易受到噪声影响，而且当轮廓复杂时，这样提取的轮廓会过于破碎、短线条太多。因此他们采用8方向卷积的方式对G进行进一步的滤波降噪。

所谓8方向卷积，就是使用八个不同方向的线段作为卷积核，分别对图像梯度G进行卷积，然后将卷积结果合并，就得到最终的轮廓图像。

# 3.3.1方向卷积核

IMG_269

如图所示，8个卷积核为大小相同的正方形，其边长可以随意设置。论文中将边长设置为G矩阵长或宽的1/30。过卷积核中心做一水平直线（直线上所有像素值为1，直线外所有像素值为0），将该直线以卷积核中心为轴旋转，取每旋转22.5°后的图片为一个卷积核。这样就得到8个不同方向的卷积核，分别用于提取这8个方向的轮廓。

# 3.3.2 原图卷积

在使用卷积核之前，先将其归一化，也就是使得卷积核中所有像素值相加为1。然后用这8个归一化卷积核与图像梯度G进行卷积，并保证卷积后的图像大小与G相同，得到的8个图像就是这8个方向的轮廓特征。

# 3.3.3 卷积结果合并

论文中没有直接将8个卷积结果叠加，其合并过程分两步：

（1）处理卷积结果

我们在卷积时保证了图像大小不变，因此卷积结果与G中的像素可以一一对应。对于每个像素点位(i, j)，看8个卷积结果中该点处的像素值，取最大的值所在的卷积结果，令该卷积结果在(i, j)处的值等于G(i, j)，令另外7个卷积结果在(i, j)处的值为0。

也就是只保留每个点附近最强烈最明显的方向的轮廓特征，以此来达到降噪的目的。

（2）合并卷积结果

将上述8个处理后的卷积结果分别与他们对应方向的归一化卷积核再次进行卷积。最后将这次得到的8个新的卷积结果叠加为一张图片。

再次卷积的目的是进一步提纯轮廓特征，使得降噪更加彻底。从而最大限度的避免细碎边缘的产生。

# 4反色输出

至此，核心运算完成。但是此时图片底色是黑色，因为上述过程导致边缘以外的像素值为0，也就呈现黑色。所以最后要进行反色，使之变为白色背景黑色线条。反色操作非常简单，在uint8格式下将每个像素值替换为255和它的差值即可，这样就实现了黑白翻转。

最后，将所有像素值映射到[0, 1]区间，就得到了最终的线描效果图。我们对一些图片进行测试，结果如下：







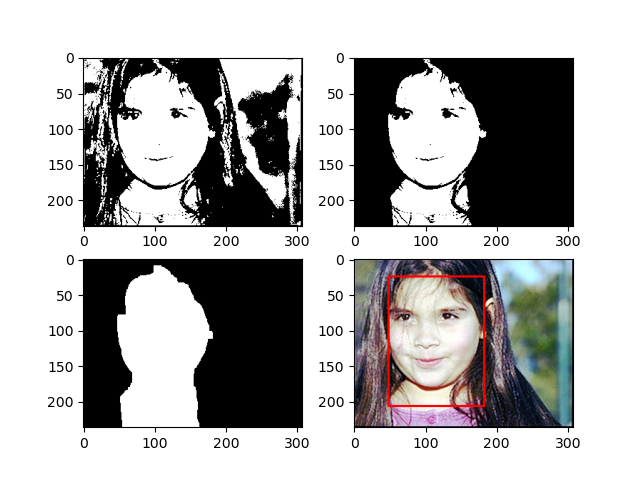
# 4、实验与结论

本文采用helen数据集（<http://www.ifp.illinois.edu/~vuongle2/helen/），helen数据集是一个人脸分割图像数据集，总共包含2330>张图片，每张图片为400\*400大小，每张图片给像素级别的人脸，眼睛，眉毛，鼻子，嘴巴等标签。由于原数据集较大，我们提取其中500张作为测试集合，50张作为验证集，并且在剩余图像中进行检测。

利用基于自适应粒子群算法对连通域过滤超参数的搜索算法，我们得到最优参数为：

M = 0.2957，R = 0.637，K = 0.899， S = 0.163

利用最优参数我们对若干测试集图像进行测试，可以看到大部分图像可以很好的识别出人脸位置，少部分图像会出现多个框或者框偏移问题。



此图为人脸检测算法中间过程图，从左到右，从上到下分别为高斯贝叶斯得到的二值图，经过过滤操作得到的二值图，经过开闭操作得到的二值图以及最后的检测结果。

在具体的检测算法实现过程中，我们发现由于人脸通常与脖子等区域同时出现在图片中，人脸框往往会将脖子等区域等于识别为人脸（由于脖子肤色与人脸相近），因此为了提升准确率，我们人为将识别框的下界上挪，以达到一个更好的效果。

# 5、不足与改进

* 我们的算法采用简单高斯朴素贝叶斯，在大量像素数据点情况下拟合效果较差，后续改进 考虑复杂模型如高斯混合模型，支持向量机或者深度神经网络等等。
* 在实验上，由于时间原因，无法多次重复实验，绘制实验图像以及进行消融性实验。
* 在实现上，粒子群算法每一次迭代，每一个粒子都需要在测试集上进行检测，耗时巨大， 超参寻找上存在性能问题，可以考虑通过提前存储图像中间处理结果以及将算法改为并行算法 对连通域过滤超参数搜索算法进行加速。

# 6、小组成员分工

黎羿江：

在本次课程设计中，我主要参与对基于APSO算法的设计以及实现，对高斯贝叶斯模型进行训练，以及对联通域过滤进行实现。同时也撰写了相应部分的论文。

**参考文献**