# 人脸检测-试验报告

**组长-杨旭俞 成员：屠文翔，吴凤，林海云，陈泽**

**目录**

目录

[人脸检测-试验报告 1](#_Toc324012653)

[1. 算法原理： 1](#_Toc324012654)

[1.1 AdaBoost 1](#_Toc324012655)

[1.2 gentleBoost 2](#_Toc324012656)

[2. 数据准备： 2](#_Toc324012657)

[3. 特征提取： 3](#_Toc324012658)

[4. 试验过程 4](#_Toc324012659)

[5. 试验分析 4](#_Toc324012660)

## 算法原理：

### 1.1 AdaBoost

****算法针对不同的训练集训练同一个弱分类器，然后把这些在不同训练集上得到的分类器集合起来，构成一个强分类器。理论证明，只要每个弱分类器分类能力比随机猜测要好，当其个数趋向于无穷个数时，强分类器的错误率将趋向于零。****算法中不同的训练集是通过调整每个样本对应的权重实现的。

最开始的时候，每个样本对应的权重是相同的，在此样本分布下训练出一个基本分类器h1(x)。对于h1(x)错分的样本，则增加其对应样本的权重；而对于正确分类的样本，则降低其权重。这样可以使得错分的样本突出出来，并得到一个新的样本分布。同时，根据错分的情况赋予h1(x)一个权重，表示该基本分类器的重要程度，错分得越少权重越大。在新的样本分布下，再次对基本分类器进行训练，得到基本分类器h2(x)及其权重。依次类推，经过T次这样的循环，就得到了T个基本分类器，以及T个对应的权重。最后把这T个基本分类器按一定权重累加起来，就得到了最终所期望的强分类器。

****算法的具体描述如下：

假定X表示样本空间，Y表示样本类别标识集合，假设是二值分类问题，这里限定Y={-1,+1}。令为样本训练集，其中。

1. 始化m个样本的权值，假设样本分布为均匀分布：，表示在第t轮迭代中赋给样本的权值。
2. 令T表示迭代的次数。
3. For t=1 to T do

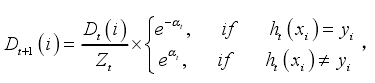
根据样本分布，通过对训练集S进行抽样（有回放）产生训练集St。

在训练集St上训练分类器。

用分类器对原训练集S中的所有样本分类。

得到本轮的分类器：X →Y，并且有误差。

令。

，

其中，Zt是一个正规因子，用来确保ΣiDt+1(i)=1。

end for

1. 最终的输出为：

### 1.2 gentleBoost

## 数据准备：

我们采用ORL数据库：共有400幅人脸图像（40人，每人10幅，大小为112px\*92px）

大多数图像的光照方向和强度差不多。

* 但有少许表情，姿势，伸缩的变化，眼镜对得不太准，尺度差异在10%
* 并不是每个人都有所有这些的变化的图像，即有些人姿势变化大一点，有些还带了眼镜。

如下图所示：

## 特征提取：

使用方法：基于PCA的人脸特征抽取

PCA（主成分分析Pincipal Component analysis）

特征提取是指通过已有的特征的组合（变换）建立一个新的特征子集。PCA的实质是在尽可能低好地代表原始数据的前提下，通过线性变换将高维空间的样本数据投影到低纬空间中。

理论推导

(1)0维的情况

考虑在0维空间（1个点），如何以1个D维向量**X**0（D维中的一个点）来表示n个样本。使得**X**0到这n个样本的距离平方和E(X0)，这种情况下样本均值能够最好地代表n个样本的D维向量 。(从n个样本距离平方和推导)

(1)1维的情况

考虑在1维空间（D维空间中的一条线）:

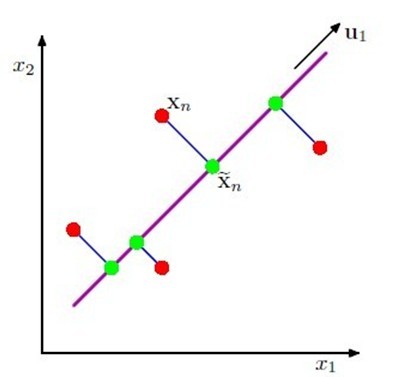
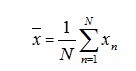


图1 红点代表原始数据点；绿点代表被映射到低维空间后的点；紫线代表映射平面。

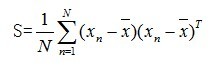
映射到1维空间。我们设低维空间的方向向量为D维单位向量u1,并且具有正交性,即u1Tu1=1。那么每一个样本数据点xn被映射到1维空间后就表示成了u1Txn。我们令原始N个样本数据均值向量为http://hiphotos.baidu.com/chb_seaok/pic/item/d194e81abf24cb2d41a9ad6a.jpg

（1）

那么映射后的数据方差就为：

http://hiphotos.baidu.com/chb_seaok/pic/item/180f98fedf17ef2dd7887d25.jpg（2）

注：此处S为原始数据集的协方差矩阵

（3）

我们所希望的低维空间是能使等式（2）值最大的空间，即方差最大化。那么问题就转化为求解等式（2）的最大值。

因为u1向量是正交向量，所以我们引入拉格朗日乘子法求解等式（2）得最大值。构造条件限制等式：

http://hiphotos.baidu.com/chb_seaok/pic/item/5ef4bede073ded8777c6383e.jpg（4）

由高等数学知识可知，我们要求解关于u1的等式（4）的最大值，只需要令（4）对u1求导令其等于0，得：

http://hiphotos.baidu.com/chb_seaok/pic/item/ce0f6dd0647dcb683bf3cf16.jpg（5）

由线形代数知识可知，http://hiphotos.baidu.com/chb_seaok/pic/item/2cf7f4c02954773db219a8f7.jpg必为协方差矩阵的特征值，而u1为其对应的特征向量。

我们将由1维情况扩展到M>1维情况，协方差矩阵S应该有M个特征特征值：http://hiphotos.baidu.com/chb_seaok/pic/item/c65451787b40069c0bd18712.jpg，其对应的特征向量应为：u1,…,un。

## 试验过程

## 试验分析