

# 图像处理第三次大作业 - 人脸识别

17373489 张佳一

## 1 前言






**人脸识别**是指能够**识别或验证**图像或视频中的**主体的身份**的技术，目的是完成**识别**功能。该术语需要和人脸检测进行区分，人脸检测是在一张图片中把人脸定位出来，完成的是**搜寻**的功能。

自七十年代以来，人脸识别已经成为了计算机视觉和生物识别领域被研究最多的主题之一。基于人工设计的特征和传统机器学习技术的传统方法近来已被使用非常大型的数据集训练的深度学习神经网络取代。

在本次作业中，我对比较流行的几种人脸识别方法进行了较为全面的总结及对比分析，其中既包括比较经典的传统方法：*特征脸方法 (Eigenfaces)*、*线性判别分析法 (Fisher Faces)*、*局部二值模式直方图法*，也有深度学习方法。

## 2 数据集介绍

本次作业使用了助教提供的第一个数据集 **PIE dataset**，包含了68个人在**五种不同姿态下**的共11554张面部图像数据，并以MAT格式的文件保存，如下：

 Pose05_64x64.mat	2017/5/8 13:10	MAT 文件	13,358 KB
 Pose07_64x64.mat	2017/5/8 13:10	MAT 文件	6,531 KB
 Pose09_64x64.mat	2017/5/8 13:10	MAT 文件	6,543 KB
 Pose27_64x64.mat	2017/5/8 13:10	MAT 文件	13,346 KB
 Pose29_64x64.mat	2017/5/8 13:10	MAT 文件	6,543 KB

其中 Pose05\_64x64.mat 的文件信息如下：

```
In [38]: mat_pose05 = loadmat('./dataset/Pose05_64x64.mat')
```

```
In [39]: mat_pose05
```

```
Out[39]: {'__header__': b'MATLAB 5.0 MAT-file Platform: nt, Created on: Mon May 08 13:10:30 2017',
  '__version__': '1.0',
  '__globals__': [],
  'isTest': array([[0.],
  [0.],
  [0.],
  ...,
  [1.],
  [0.],
  [0.]]) ,
  'fea': array([[ 8,  9,  9, ..., 35, 36, 42],
  [ 6,  7,  7, ..., 38, 32, 31],
  [ 8,  8,  6, ..., 26, 25, 23],
  ...,
  [11, 10,  8, ..., 36, 40, 42],
  [ 9, 11,  9, ..., 29, 31, 36],
  [12,  9,  9, ..., 29, 34, 38]] , dtype=uint8),
  'gnd': array([[ 1],
  [ 1],
  [ 1],
  ...,
  [68],
  [68],
  [68]] , dtype=uint8)}
```

```
In [40]: mat_pose05['fea'].shape
```

```
Out[40]: (3332, 4096)
```

由上图可知，每条 `fea` 与 `gnd` 分别为图像数据和标签，`isTest` 为1时，表明为测试图，不参与训练。此外，我们还能看到姿态为**Pose05** 的图片共有3332张，每张图片数据为一个  $64 \times 64 = 4096$  的像素矩阵构成。其余四种姿态的数据信息与之类似，这里不在一一列举。

下面展示的是 `label` 为1的人的在五种姿态下的图像（具体见 `/image/sample` 文件夹下）：



可见，这五张图像分别是测试者的“左侧”、“仰视”、“俯视”、“正视”以及“右侧”角度拍摄的。

对于数据集的处理思路较为简单，不再赘述，代码如下：

```
1 class Dataset():
2     path = '../dataset/PIE dataset/'
3
4     def __init__(self, dataset_name=None):
5         self.ds_name = dataset_name if dataset_name is not None else
        'PoseAll' # e.g. Pose05
6         self.ds_path = '../dataset/PIE dataset/' + dataset_name +
        '_64x64.mat' if dataset_name is not None else None
7
8
9     def load(self):
10        if self.ds_path is None:
```

```

11         print("ERROR, You didn't passed parameter 'dataset_name' when
you built the Dataset(). Maybe you want to replace load() with load_all()
?")
12         self.load_all()
13         return
14
15     mat = loadmat(self.ds_path)
16     self.size = mat['gnd'].shape[0]
17
18     train_data = []
19     train_label = []
20     test_data = []
21     test_label = []
22
23     for i in range(self.size):
24         gray = mat['fea'][i].reshape((64, 64))
25         data = cv2.merge([gray])
26         label = mat['gnd'][i][0]
27
28         if mat['isTest'][i] == 0.0:
29             train_data.append(data)
30             train_label.append(label)
31         else:
32             test_data.append(data)
33             test_label.append(label)
34
35     self.train_data = np.array(train_data)
36     self.train_label = np.array(train_label)
37     self.test_data = np.array(test_data)
38     self.test_label = np.array(test_label)
39
40
41     def load_all(self):
42         self.size = 0
43         train_data = []
44         train_label = []
45         test_data = []
46         test_label = []
47
48         # Pose05
49         mat = loadmat('../dataset/PIE dataset/Pose05_64x64.mat')
50         size = mat['gnd'].shape[0]
51         self.size += size
52         for i in range(size):
53             gray = mat['fea'][i].reshape((64, 64))
54             data = cv2.merge([gray])
55             label = mat['gnd'][i][0]
56
57             if mat['isTest'][i] == 0.0:
58                 train_data.append(data)
59                 train_label.append(label)
60             else:
61                 test_data.append(data)
62                 test_label.append(label)
63
64         # Pose07
65         mat = loadmat('../dataset/PIE dataset/Pose07_64x64.mat')
66         size = mat['gnd'].shape[0]

```

```

67     self.size += size
68     for i in range(size):
69         gray = mat['fea'][i].reshape((64, 64))
70         data = cv2.merge([gray])
71         label = mat['gnd'][i][0]
72
73         if mat['isTest'][i] == 0.0:
74             train_data.append(data)
75             train_label.append(label)
76         else:
77             test_data.append(data)
78             test_label.append(label)
79
80     # Pose09
81     mat = loadmat('../dataset/PIE dataset/Pose09_64x64.mat')
82     size = mat['gnd'].shape[0]
83     self.size += size
84     for i in range(size):
85         gray = mat['fea'][i].reshape((64, 64))
86         data = cv2.merge([gray])
87         label = mat['gnd'][i][0]
88
89         if mat['isTest'][i] == 0.0:
90             train_data.append(data)
91             train_label.append(label)
92         else:
93             test_data.append(data)
94             test_label.append(label)
95
96     # Pose27
97     mat = loadmat('../dataset/PIE dataset/Pose27_64x64.mat')
98     size = mat['gnd'].shape[0]
99     self.size += size
100    for i in range(size):
101        gray = mat['fea'][i].reshape((64, 64))
102        data = cv2.merge([gray])
103        label = mat['gnd'][i][0]
104
105        if mat['isTest'][i] == 0.0:
106            train_data.append(data)
107            train_label.append(label)
108        else:
109            test_data.append(data)
110            test_label.append(label)
111
112    # Pose29
113    mat = loadmat('../dataset/PIE dataset/Pose29_64x64.mat')
114    size = mat['gnd'].shape[0]
115    self.size += size
116    for i in range(size):
117        gray = mat['fea'][i].reshape((64, 64))
118        data = cv2.merge([gray])
119        label = mat['gnd'][i][0]
120
121        if mat['isTest'][i] == 0.0:
122            train_data.append(data)
123            train_label.append(label)
124        else:

```

```

125         test_data.append(data)
126         test_label.append(label)
127
128     self.train_data = np.array(train_data)
129     self.train_label = np.array(train_label)
130     self.test_data = np.array(test_data)
131     self.test_label = np.array(test_label)
132
133
134     def gen_dataset(self):
135         # ensure dataset type as int32
136         trainset = {
137             'data': self.train_data.astype(int),
138             'label': self.train_label.astype(int)
139         }
140         testset = {
141             'data': self.test_data.astype(int),
142             'label': self.test_label.astype(int)
143         }
144         return trainset, testset
145
146
147     def describe(self):
148         # describe dataset and generate sample image to ../image/sample/
149         print('Dataset Name: ' + self.ds_name)
150         print('Dataset Size: ' + str(self.size))
151
152         if self.ds_name != 'PoseAll':
153             img = self.train_data[0]
154             cv2.imshow(self.ds_name, img)
155             cv2.waitKey(0)
156             cv2.destroyWindow(self.ds_name)
157             cv2.imwrite('../image/sample/' + self.ds_name + '_sample.png',
158                 img)

```

在这里要感谢助教对数据集提前做了筛选与处理，相比图像格式的数据，Mat格式不仅占用空间小，也极大的方便了我们读取图像数据，给助教一个大大的赞！

## 3 传统人脸识别方法

人脸识别的第一步，就是要找到一个模型可以用简洁又具有差异性的方式准确反映出每个人脸的特征。在识别人脸时，先将当前人脸采用与前述同样的方式提取特征，再从已有特征集中找出当前特征的最邻近样本，从而得到当前人脸的标签。

本小节基于 OpenCv 实现了 EigenFaces、FisherFaces、LBPH 等三种经典的传统人脸识别算法，并对比评估了这三种方法在五种姿态下的准确率（包括单项评估以及整体评估），最后对实验结果做了简单的分析总结。

### 3.1 Eigen Faces

Eigenfaces 通常也被称为特征脸，是一种从主成分分析（Principal Component Analysis, PCA）中导出的人脸识别和描述技术。

### 3.1.1 基本原理

特征脸方法的主要思路就是将输入的人脸图像看作一个个矩阵，通过在人脸空间中一组正交向量，并选择最重要的正交向量，作为“主成分”来描述原来的人脸空间。为了更好地理解特征脸方法，我们需要先了解PCA的主要过程。

#### PCA的主要过程

在很多应用中，需要对大量数据进行分析计算并寻找其内在的规律，但是数据量巨大造成了问题分析的复杂性，因此我们需要一些合理的方法来减少分析的数据和变量同时尽量不破坏数据之间的关联性。于是这就有了**主成分分析方法**，其核心思想为：

- **数据降维**，减少变量个数，确保变量独立；
- **去除噪声**，发现数据背后的固有模式。

#### PCA的主要过程：

1. **特征中心化**：将每一维的数据（矩阵A）都减去该维的均值，使得变换后（矩阵B）每一维均值为0；
2. 计算变换后矩阵B的**协方差矩阵C**；
3. 计算协方差矩阵C的**特征值和特征向量**；
4. 选取**大的特征值**对应的特征向量作为“**主成分**”，并构成新的数据集；

#### 特征脸方法

特征脸方法就是对原始数据使用PCA方法进行降维，获取其中的主成分信息，从而实现人脸识别的方法。这种方法的核心思想是认为同一类事物必然存在相同特性（主成分），通过将同一目标（人脸图像）的特性寻在出来，就可以用来区分不同的事物了。

#### 特征脸方法的过程：

1. 将训练集中的图像数据拉成一行，然后将这N张人脸图像组合在一起形成一个大矩阵A。
  - 比如本例中人脸图像大小为 64x64，那么矩阵维度就是  $N * 4096$ 。
2. 将N个人脸在对应的维度求平均，得到一个“平均脸”；
3. 将矩阵A中N个图像都减去“平均脸”，得到新矩阵B；
4. 计算B的协方差矩阵；
5. 计算协方差矩阵的特征值和特征向量（特征脸）；
6. 将训练集图像和测试集图像都投影到特征向量空间中，再使用聚类方法（最近邻或k近邻等）得到里测试集中的每个图像最近的图像，进行分类即可。

### 3.1.3 代码实现

1 |

### 3.1.4 实验结果

特征脸识别的具有一定的局限性，它需要保证人脸图像满足：

- 待识别图像中人脸尺寸接近特征脸中人脸的尺寸；
  - 从PCA方法的过程可以看出，特征脸识别的方法是以每张人脸的一个维度（可以看出是矩阵的一列）为单位进行处理的，求得特征向量（特征脸）中包含训练集每个纬度的绝大部分信息。但是若测试集中人脸尺寸不同，那么与特征脸中纬度的也就没法对应起来。
- 待识别人脸图像必须为正面人脸图像。

若不满足这两个条件，则识别错误率将会很高。

## 3.2 Fisher Faces

PCA方法是 EigenFaces 方法的核心，它找到了最大化数据总方差特征的线性组合，但它的缺点在于在操作过程中会损失许多特征信息。因此，如果损失的信息正好是用于分类的关键信息，可能会导致无法完成分类。

FisherFaces 采用线性判别分析（linear discriminant analysis, LDA）实现人脸识别。线性判别分析最早由Fisher在1936年提出，可以用来处理两类的线性判别问题，也被称为“Fisher判别分析法”。

### 3.2.1 基本原理

线性判别分析在对特征降维的同时考虑类别信息。其主要思想是：在低位表示下，相同的类应该紧密地聚集在一起；而不同的类别应该尽可能地分散开，并且它们之间的距离尽可能地远。

简单来说，线性判别分析就是要尽力满足以下两个要求：

- 类别间的差别尽可能大
- 类别内的差别尽可能小

做线性判别分析时，首先将训练样本集投影到一条直线A上，让投影后的点满足：

- 同类间的点尽可能地靠近
- 异类间的点尽可能地原理

做完投影后，将待测样本投影到直线A上，根据投影点的位置判定样本的类别，就完成了识别。

### 3.1.2 代码实现

```
1 |
```

### 3.1.3 实验结果

## 3.3 LBPH

LBPH (Local Binary Pattern Histograms, 局部二值模式直方图) 所使用的模型基于LBP (Local Binary Pattern, 局部二值模式) 算法。LBP最早是被作为一种有效地纹理描述算子提出的, 由于在表述图像局部纹理特征上效果出众而得到广泛应用。

### 3.3.1 基本原理

LBP算法的基本原理是, 将像素点A的值与其最邻近的8个像素点的值逐一比较:

- 如果A的像素值大于其临近点的像素值, 则得到0;
- 如果A的像素值小于其临近点的像素值, 则得到1。

最后, 将像素点A与其周围8个像素点比较所得到的0, 1值连起来, 得到一个8位的二进制序列, 将该二进制序列转换为十进制数作为点A的LBP值。

例如, 将下图的中间点76作为阈值, 对齐8邻域像素进行二值化处理, 结果如右图所示。

128	36	251
48	76	9
11	213	99

1	0	1
0		0
0	1	1

完成二值化以后, 任意指定一个开始位置, 将得到的二值结果进行序列化, 组成一个8位的二进制数。

例如, 从当前像素点的正上方开始, 以顺时针为序得到二进制序列“01011001”。

最后, 将二进制序列“01011001”转换为所对应的十进制数“89”, 作为当前中心点的像素值, 如右图所示。

128	36	251
48	76	9
11	213	99

	89	

对图像逐像素用以上方式进行处理, 就得到了LBP特征图像, 这个特征图像的直方图被称为LBPH, 或称为LBP直方图。

为了得到不同尺度下的纹理结果, 还可以使用圆形邻域, 将计算扩大到任意大小的邻域内, 由于与上述处理方式类似, 这里具体细节不再赘述。

人脸的整体灰度由于受到光纤的影响, 经常会发生变化, 但是人脸各部分之间的相对灰度会基本保持一致。而LBP的主要思想正是以当前点与其邻域像素的相对关系作为处理结果, 这样在图像灰度整体发生变化时, 能够保证从LBP算法中提取出的特征具有不变性。因此LBP在人脸识别中得到了广泛的应用。



### 3.3.2 代码实现

1 |

### 3.3.3 实验结果

## 3.4 实验结果对比与分析

### 3.4.1 算法耗时

### 3.4.2 准确率

## 4 基于卷积神经网络的人脸识别方法

---

卷积神经网络（CNN）是人脸识别方面最常用的一类深度学习方法。深度学习方法的主要优势是可用大量数据来训练，从而学到对训练数据中出现的变化情况稳健的人脸表征。这种方法不需要设计对不同类型的类内差异（比如光照、姿势、面部表情、年龄等）稳健的特定特征，而是可以从训练数据中学到它们。深度学习方法的主要短板是它们需要使用非常大的数据集来训练，而且这些数据集中需要包含足够的变化，从而可以泛化到未曾见过的样本上。

用于人脸识别的 CNN 模型可以使用不同的方法来训练。其中之一是将该问题当作是一个分类问题，训练集中的每个主体都对应一个类别。训练完之后，可以通过去除分类层并将之前层的特征用作人脸表征而将该模型用于识别不存在于训练集中的主体。

对于基于 CNN 的人脸识别方法，影响准确度的因素主要有三个：训练数据、CNN 架构和损失函数。因为在大多数深度学习应用中，都需要大训练集来防止过拟合。一般而言，为分类任务训练的 CNN 的准确度会随每类的样本数量的增长而提升。这是因为当类内差异更多时，CNN 模型能够学习到更稳健的特征。但是，对于人脸识别，我们感兴趣的是提取出能够泛化到训练集中未曾出现过的主体上的特征。因此，用于人脸识别的数据集还需要包含大量主体，这样模型也能学习到更多类间差异。

## 5 实验结果对比与分析

---

## 6 总结

---

特征脸算法对光照十分敏感。

(1) 主成分分析 (PCA) ——Eigenfaces (特征脸) ——函数:  
cv2.face.EigenFaceRecognizer\_create ()

PCA: 低维子空间是使用主元分析找到的, 找具有最大方差的哪个轴。

缺点: 若变化基于外部 (光照), 最大方差轴不一定包括鉴别信息, 不能实行分类。

(2) 线性判别分析 (LDA) ——Fisherfaces (特征脸) ——函数:  
cv2.face.FisherFaceRecognizer\_create()

LDA: 线性鉴别的特定类投影方法, 目标: 实现类内方差最小, 类间方差最大。

(3) 局部二值模式 (LBP) ——Local Binary Patterns Histograms ——函数:  
cv2.face.LBPHFaceRecognizer\_create()

PCA和LDA采用整体方法进行人脸辨别, LBP采用局部特征提取, 除此之外, 还有的局部特征提取方法为:

盖伯小波 (Gabor Waelets) 和离散傅里叶变换 (DCT) 。

## 参考文献

---

- [1] <https://blog.csdn.net/wanghz999/article/details/78817265>
- [2] <http://openbio.sourceforge.net/resources/eigenfaces/eigenfaces-html/facesOptions.html>
- [3] <https://blog.csdn.net/u010006643/article/details/46417127>
- [4] <https://blog.csdn.net/smartempire/article/details/21406005>
- [5] <https://blog.csdn.net/smartempire/article/details/22938315>
- [6] [https://www.cnblogs.com/tianyalu/p/6136970.html?utm\\_source=itdadao&utm\\_medium=referal](https://www.cnblogs.com/tianyalu/p/6136970.html?utm_source=itdadao&utm_medium=referal)
- [7] <https://www.cnblogs.com/little-monkey/p/8118938.html>