

# 会议室场景分析报告(请勿外传)

龙良曲

## 1. 实验目的与方案

### 1.1. 实验目的

实验采用 Sensetime 最新 SDK(R103)，在会议室多种环境条件下采集人脸照片数据，并完成对应维度下人脸识别的 Accuracy 分析和人脸验证的 ROC 分析。

### 1.2. 实验环境

- ✚ **901 会议室：**关门，拉上窗帘，采用四种灯光环境，并且在每种灯光下分别拍摄三个角度的人脸照片（正脸图片，小角度偏转人脸图片和大角度偏转人脸图片）：  
灯光一：关上所有灯，模拟弱光环境  
灯光二：开单侧光，模拟左右非均匀光照  
灯光三：开双侧光，模拟均匀正常光照  
灯光四：关补光灯，开头顶灯，模拟上下非均匀光照
- ✚ **502 会议室：**两个摄像头分别架在不同的位置同时采集，开会议室所有灯，被拍摄人员在会议室走动和坐下，模拟正常会议室开会场景

### 1.3. 实验方案

#### 1.3.1. 识别率分析

- ✚ **rank n 曲线：**每个维度识别前 n 个结果的正确识别率
- ✚ **distractor n 曲线：**制作人脸数据库时，增加不同数量的人脸（distractor）形成不同的数据库，然后每个维度的每个取值做 rank n 曲线（这里只做 n=0 和 9）

#### 1.3.2. ROC 分析

ROC 曲线：在每个维度的不同取值下，两个相同的人脸图片组成正样本对，两个不同的人脸图片组成负样本对。

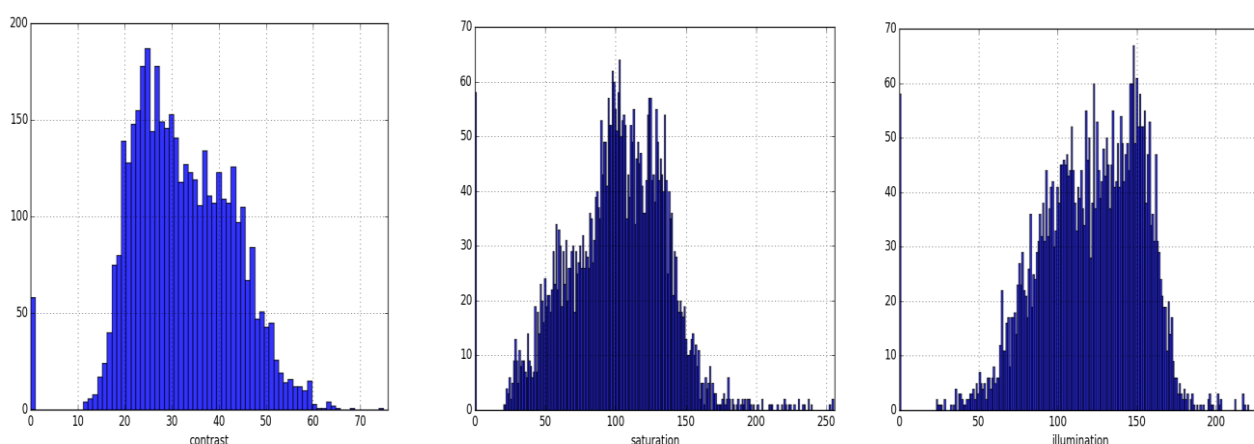
- $T_p$ ：正样本识别为正样本的数目；
- $T_n$ ：负样本识别为负样本的数目；
- $F_n$ ：负样本识别为正样本的数目；
- $F_p$ ：正样本识别为负样本的数目；

$$\text{tpr} = \frac{t_p}{t_p + f_p}$$

$$\text{fpr} = \frac{f_p}{t_n + f_p}$$

## 1.4. 人脸照片统计

本次共采集 80 人（62 男，18 女），总共 4187 张照片（不同维度数据量分布见附录）。



人脸照片对比度，饱和度，亮度分布

由采集到 4187 张人脸照片的对比度，饱和度和亮度的分布可以大致认为对比度（25～45），饱和度和亮度（75～175）范围内是现实场景中常见的正常情况。而人脸维度在眼睛（1 睁开），头发遮挡（1 无遮挡），嘴巴（1 闭合），角度（5 正脸）和光照（2 均匀正常光）是现实场景中常见的正常情况。

		eye		hair		mouth		angle					light			
取值		1	2	1	2	1	2	1~3	4	5	6	7~9	1	2	3	4
照片数		360	24	360	33	360	172	207	286	360	677	43	163	360	86	66
Test样本对数/正样本对数		1500/1521	13/14	1500/1521	120/129	1500/1521	500/524	850/892	1000/1068	1500/1521	6000/6765	90/92	250/266	1500/1521	75/78	55/57
contrast	0~20	0.05	0	0.05	0.03	0.05	0.03	0.12	0.12	0.05	0.07	0.02	0.16	0.05	0.02	0.11
	20~30	0.51	0.46	0.51	0.06	0.51	0.46	0.5	0.52	0.51	0.32	0.19	0.47	0.51	0.12	0.33
	30~40	0.36	0.54	0.36	0.85	0.36	0.41	0.27	0.27	0.36	0.24	0.67	0.27	0.36	0.34	0.27
	40~inf	0.09	0	0.09	0.06	0.09	0.09	0.12	0.1	0.09	0.37	0.11	0.1	0.09	0.52	0.29
saturation	0~25	0	0	0	0	0	0	0	0	0	0	0	0.01	0	0	0
	25~50	0.01	0	0.01	0	0.01	0.01	0.05	0.11	0.01	0.04	0.02	0.06	0.01	0	0.02
	50~75	0.03	0	0.03	0	0.03	0.04	0.08	0.13	0.03	0.18	0	0.12	0.03	0.05	0.03
	75~100	0.07	0	0.07	0.06	0.07	0.08	0.13	0.28	0.07	0.37	0.07	0.34	0.07	0.07	0.2
	100~125	0.34	0.33	0.34	0.36	0.34	0.3	0.22	0.36	0.34	0.34	0.65	0.33	0.34	0.19	0.27
	125~150	0.54	0.63	0.54	0.58	0.54	0.56	0.42	0.11	0.54	0.06	0.26	0.08	0.54	0.17	0.32
	150~175	0.02	0.04	0.02	0	0.02	0.02	0.1	0	0.02	0	0	0.05	0.02	0.31	0.17
illumination	175~inf	0	0	0	0	0	0	0	0	0	0	0	0.01	0	0.21	0
	0~25	0	0	0	0	0	0	0	0	0	0	0	0	0	0.01	0
	25~50	0	0	0	0	0	0	0	0	0	0	0	0.05	0	0	0.05
	50~75	0.01	0	0.01	0.06	0.01	0.02	0.02	0.01	0.01	0.01	0	0.23	0.01	0.05	0.08
	75~100	0.04	0	0.04	0.09	0.04	0.04	0.08	0.14	0.04	0.12	0	0.4	0.04	0.44	0.17
	100~125	0.16	0	0.16	0.18	0.16	0.13	0.24	0.38	0.16	0.22	0.05	0.23	0.16	0.35	0.24
	125~150	0.45	0.67	0.45	0.48	0.45	0.56	0.51	0.26	0.45	0.3	0.12	0.09	0.45	0.14	0.3
	150~175	0.34	0.33	0.34	0.18	0.34	0.24	0.14	0.2	0.34	0.33	0.84	0.01	0.34	0	0.12
	175~inf	0.01	0	0.01	0	0.01	0.01	0	0.01	0.01	0.01	0	0	0.01	0.01	0.05

人脸维度的图片属性分布

从人脸维度来说，在其他人脸维度正常的情况下：

**眼睛**，眼睛睁开对比度分布为 86% 正常，饱和度分布为 96% 正常，亮度分布为 98% 正

常；眼睛闭合对比度分布为 100% 正常，饱和度分布为 100% 正常，亮分布为 100% 正常。即使眼睛闭合的数据比较少，但眼睛睁开和闭合在对比度，饱和度，亮的分布大致一致，故眼睛维度变化时，图片属性维度不会对识别率和 roc 曲线造成影响。同理，头发遮挡维度和嘴巴维度变化时，图片属性维度均不会造成影响。

**角度**，从对比度来说，俯视，小角度，正脸，仰视的弱对比度：正常对比度：强对比度约为 1：8：1，但是大角度时，却有接近 4 成的照片处于强对比度；从饱和度来说，正脸和仰视 95% 以上都是正常饱和度，小角度和大角度的弱饱和度：正常饱和度约为 2：8；从亮度来说，所有角度的亮度 97% 均处于正常。**分析角度与对比度，角度与饱和度之间有一定的相关关系，故在角度维度变化时，对比度和饱和度不能全部取正常情况。**

**光照**，从对比度来说，均匀光照（弱光和正常光），其弱对比度：正常对比度：强对比度约为 1：8：1，而非均匀光照的强对比度所占比例很大；从饱和度来说，均匀弱光的弱饱和度和度相对于其他光照较多，上下非均匀光照的强饱和度相对于其他光照较多；从亮度来说，均匀弱光的弱亮度相对于其他光照较多。**分析有两个原因：一是照片数量相对较少，二是光照与对比度，饱和度和亮度之间存在必然的相关关系，故在光照维度变化时，其照片属性维度不能全部取正常情况。**

从图片属性维度来说，在其他图片属性取值正常的情况下：

		contrast			
取值		0 ~ 20	20 ~ 30	30 ~ 40	40 ~ inf
照片数		131	944	980	884
Test样本对数/正样本对数		400/437	8000/8816	7000/7432	5500/5593
eye	1	0.99	0.96	0.95	0.95
	2	0.01	0.04	0.05	0.05
hair	1	0.99	0.98	0.94	0.91
	2	0.01	0.02	0.06	0.09
mouth	1	0.92	0.79	0.73	0.77
	2	0.08	0.21	0.27	0.23
angle	1~3	0.15	0.14	0.08	0.03
	4	0.31	0.22	0.21	0.22
	5	0.18	0.38	0.36	0.13
	6	0.35	0.24	0.31	0.59
	7~9	0.01	0.02	0.04	0.02
light	1	0.34	0.2	0.15	0.1
	2	0.59	0.68	0.69	0.6
	3	0.02	0.02	0.04	0.2
	4	0.06	0.09	0.12	0.1

对比度的人脸维度分布

**对比度**，眼睛，头发遮挡比例接近于 19：1，嘴巴比例接近于 3：1，此三人脸维度分布比较一致。对于角度来说，正脸和小角度在对比度 20~40 较多，但弱对比度和强对比度时，大角度却更多，说明对比度和角度有一定的相关关系；而光照在对比度的不同取值情况下，分布比较一致，这里不能表现对比度和光照之间的相关关系。

		saturation							
取值		0 ~ 25	25 ~ 50	50 ~ 75	75 ~ 100	100 ~ 125	125 ~ 150	150 ~ 175	175 ~ inf
照片数		3	105	234	499	812	616	88	28
Test样本对数/正样本对数		0/0	250/290	500/567	1900/1937	5000/5142	4000/4632	100/117	8//9
eye	1	1	1	0.95	0.98	0.95	0.94	0.94	0.82
	2	0	0	0.05	0.02	0.05	0.06	0.06	0.18
hair	1	1	0.97	0.92	0.97	0.96	0.92	0.99	0.93
	2	0	0.03	0.08	0.03	0.04	0.08	0.01	0.07
mouth	1	1	0.95	0.86	0.82	0.75	0.68	0.88	0.82
	2	0	0.05	0.14	0.18	0.25	0.32	0.12	0.18
angle	1~3	0.33	0.07	0.03	0.04	0.06	0.13	0.15	0
	4	0	0.28	0.21	0.25	0.21	0.15	0.16	0.18
	5	0.67	0.03	0.09	0.13	0.31	0.49	0.51	0.71
	6	0	0.63	0.67	0.56	0.38	0.19	0.17	0.07
	7~9	0	0	0	0.02	0.05	0.03	0.01	0.04
light	1	0	0.63	0.65	0.33	0.14	0.04	0.07	0.21
	2	1	0.33	0.29	0.53	0.73	0.81	0.28	0
	3	0	0.04	0.02	0.02	0.02	0.06	0.41	0.79
	4	0	0	0.04	0.12	0.11	0.08	0.24	0

饱和度的人脸属性分布

**饱和度**，眼睛，头发遮挡比例接近于 19: 1，嘴巴比例接近于 8: 2，此三个人脸维度分布比较一致。对于角度来说，饱和度在 125 以下大角度的比较多，饱和度在 125 以上正脸较多，说明饱和度和角度有一定的相关关系；而光照在饱和度小于 75 的情况下，均匀弱光照比较多，在饱和度 75~150，均匀正常光比较多，在饱和度 150 以上，却是上下非均匀光比较多，这里说明饱和度和光照有一定的相关关系。

		illumination							
取值		0 ~ 25	25 ~ 50	50 ~ 75	75 ~ 100	100 ~ 125	125 ~ 150	150 ~ 175	175 ~ inf
照片数		0	3	71	319	379	732	585	26
Test样本对数/正样本对数		0/0	0/0	80/87	750/780	1000/1182	5000/5192	5000/5573	10//14
eye	1	0	1	0.96	0.96	0.94	0.95	0.97	0.96
	2	0	0	0.04	0.04	0.06	0.05	0.03	0.04
hair	1	0	0.67	0.86	0.97	0.92	0.94	0.97	1
	2	0	0.33	0.14	0.03	0.08	0.06	0.03	0
mouth	1	0	1	0.79	0.83	0.81	0.68	0.75	0.85
	2	0	0	0.21	0.17	0.19	0.32	0.25	0.15
angle	1~3	0	0	0.04	0.05	0.09	0.12	0.05	0
	4	0	0.67	0.24	0.26	0.27	0.17	0.16	0.19
	5	0	0	0.54	0.37	0.31	0.39	0.24	0.19
	6	0	0.33	0.17	0.31	0.32	0.3	0.48	0.62
	7~9	0	1	0.01	0.01	0.02	0.02	0.08	0
light	1	0	0	0.58	0.51	0.25	0.06	0.02	0.04
	2	0	0	0.1	0.3	0.46	0.8	0.89	0.65
	3	0	0	0.2	0.15	0.12	0.02	0	0.04
	4	0	0	0.13	0.03	0.17	0.12	0.09	0.27

亮度的人脸属性分布

**亮度**，0~50 的照片比较少，不予考虑。眼睛比例接近于 19: 1，头发遮挡比例接近于 9: 1，嘴巴比例接近于 3: 1，此三个人脸维度分布比较一致。对于角度来说，亮度在 50~150，角度分布约为 1: 2: 3: 3: 1，亮度 150 以上，反而是大角度比较多一些。对于光照来说，亮度小于 100 的均匀弱光较多，亮度大于 100 的均匀正常光比较多，说明亮度与光照有必然的相关关系。

**数据分布结论：**眼睛，头发遮挡，嘴巴与对比度，饱和度，亮度之间无明显的相关关系，

并且分布比较一致,故相互之间不会对识别率和 ROC 曲线造成影响;角度,光照与对比度,饱和度,亮度之间有比较明显的相关关系,故在某一维度变化的时候,另一方面维度不能取正常情况。综合上述关系和数据量的综合考虑,在做 rank-n 识别率和 ROC 分析时,人脸维度某一维度变化时,仅控制其他人脸维度正常,图片属性某一维度变化时,也仅控制其他图片属性正常。

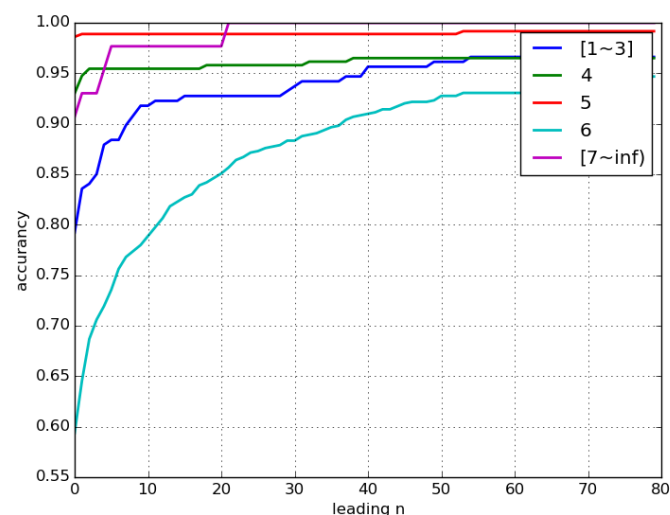
## 2. 测试结果分析

### 2.1. 人脸角度

#### 2.1.1. 编号约定

- 🚦 [1-3]: 人脸俯视一定角度
- 🚦 4: 人脸向左/右侧小角度(约 30 度)
- 🚦 5: 人脸正视
- 🚦 6: 人脸向左/右侧大角度(约 60 度)
- 🚦 [7-8): 人脸仰视一定

#### 2.1.2. Rank n 曲线



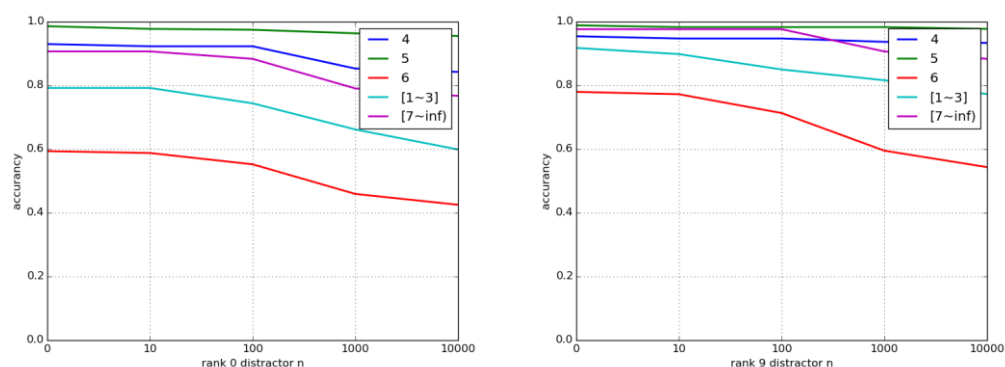
本次采集的人脸数据共 80 类,我们针对每种维度下的识别情况绘制了前 rank n 的曲线如上图。

- 🚦 正面的人脸识别率非常高,在 rank 0 就取得了最高的识别率,并且随着 rank n 的增大并没有显示提升
- 🚦 仰视的人脸照片在前 rank 3 识别率并没有小角度高,但随着 rank n 的增大,识别率显著提升,甚至在 rank 21 左右接近于 1

🚦 俯视的人脸照片识别率不高，大角度的识别率最低

### 2.1.3. Distractor n 曲线

针对每种维度，我们采用了添加 distractor 的方法进行识别测试，一共测试了 distractor 为 0, 10, 100, 1000, 10000 的情况，当 distractor 数据源不足时，我们从其它数据集（如 MegaFace）上抽取了部分照片，并绘制如下曲线。



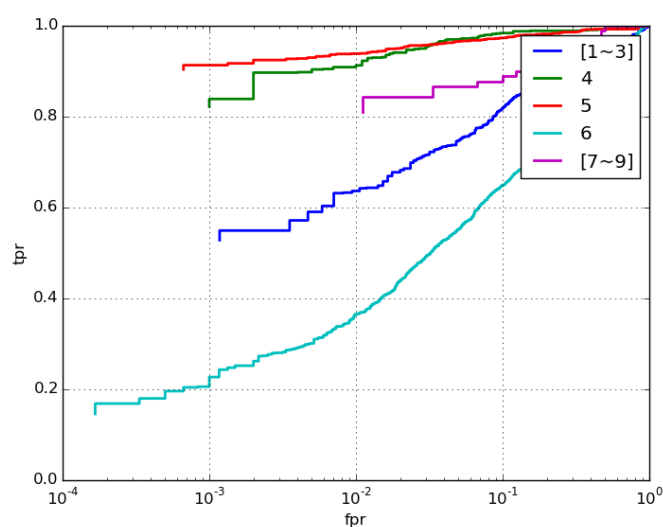
上图中左边为 rank0 下不同的 distractor 的识别率曲线，右边为 rank10 下不同的 distractor 识别率曲线。

🚦 随着 distractor 的增加，正面角度影响并不明显，其它维度明显下降

🚦 Rank 10 的 accuracy 普遍好于 rank 0

### 2.1.4. ROC 曲线

我们从数据集中抽取了一定数量的正样本对与负样本对，根据不同的 threshold 来分类识别结果，并绘制对应的 ROC 网线。



🚦 正面和小角度的 precision 与 sensitivity 均较好，并且差别不大

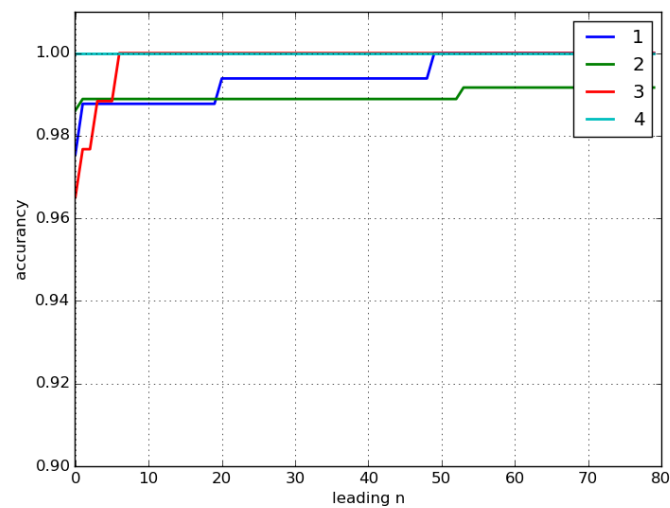
🚦 仰视识别情况要好于俯视，大角度的识别情况最差

## 2.2. 环境灯光

### 2.2.1. 编号约定

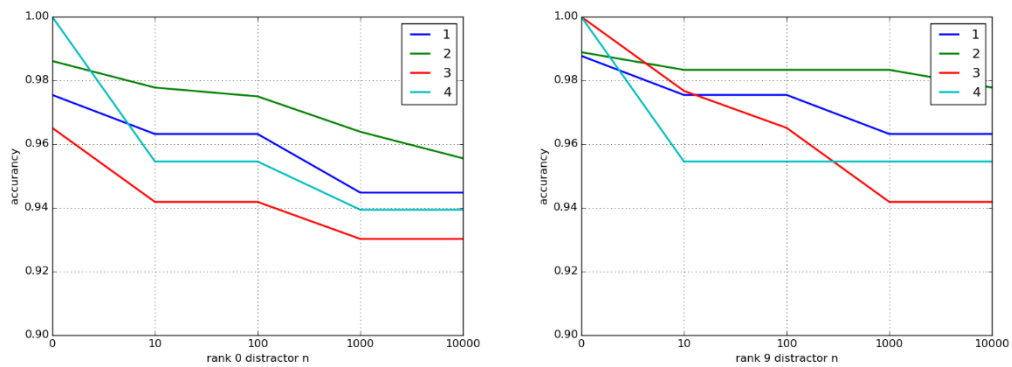
- ✚ 1: 弱光
- ✚ 2: 正常光照
- ✚ 3: 单侧光
- ✚ 4: 顶光

### 2.2.2. Rank n 曲线



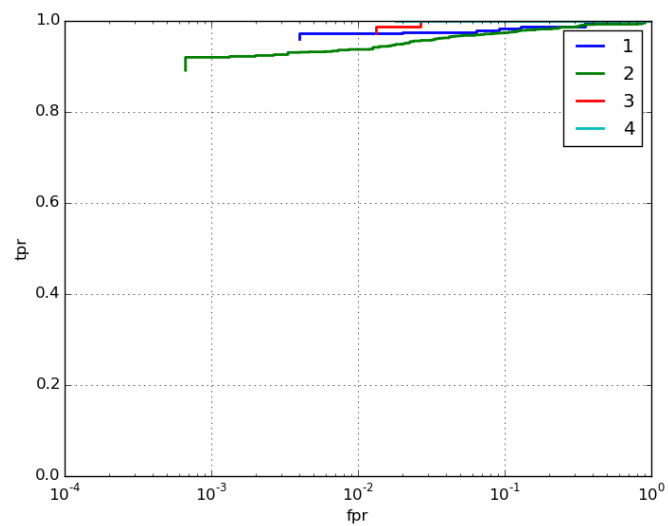
- ✚ 正常光照在 rank0 时识别较高，但弱光在 rank20 左右反超正常光照
- ✚ 顶光在这次测试中 accuracy 达到 1，可能是由于样本数量的问题(样本分布为：163，360，86，66)
- ✚ 单侧光在 rank0 时识别率最低，在 rank5 左右接近于 1

### 2.2.3. Distractor n 曲线



- 正常光照识别率一直保持较高
- 弱光识别情况与正常光基本一致，识别率略低
- 单侧光照识别率较低
- 顶光在 distractor 为 10 之后开始急速下降，说明 distractor=0 时可能包含了大量的噪声

### 2.2.4. ROC 曲线



- 正常光照识别最差（可能是由于数据分布不理想）

## 2.3. 头发遮挡

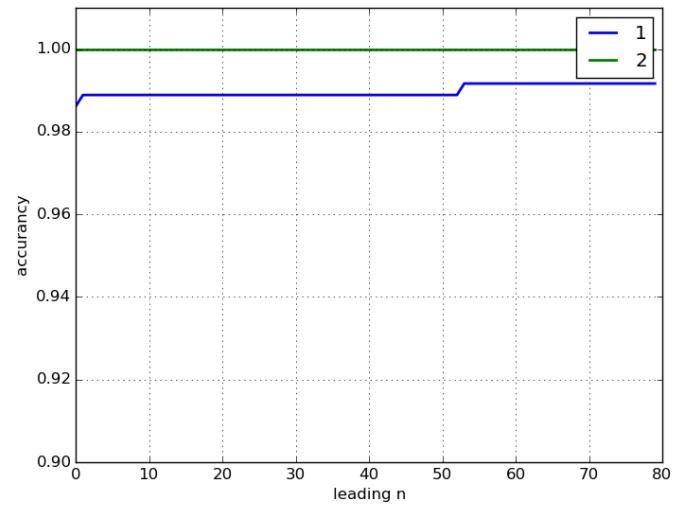
### 2.3.1. 编号约定

- 1: 无遮挡



2: 有遮挡

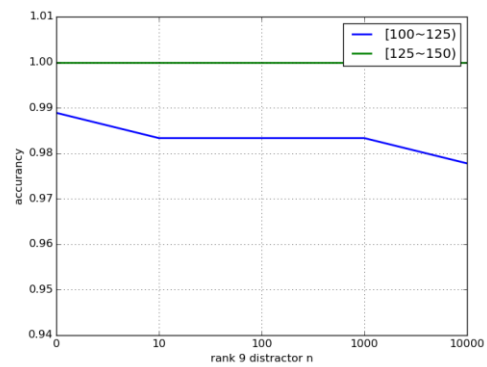
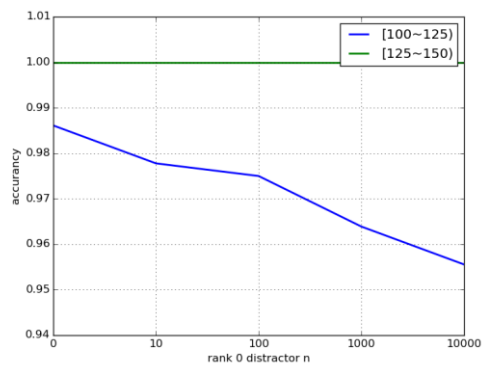
### 2.3.2. Rank n 曲线



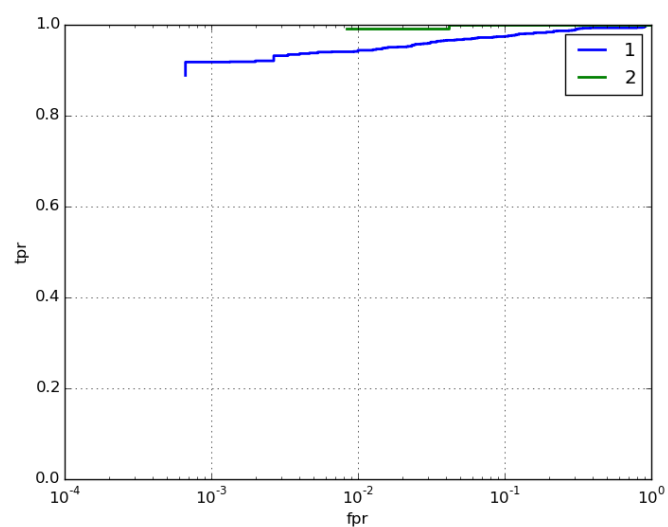
由于其它维度均正常，遮挡对识别率的影响不大

样本分布为(360,33)，因此出现遮挡 accuracy=1 的情况

### 2.3.3. Distractor n 曲线



### 2.3.4. ROC 曲线

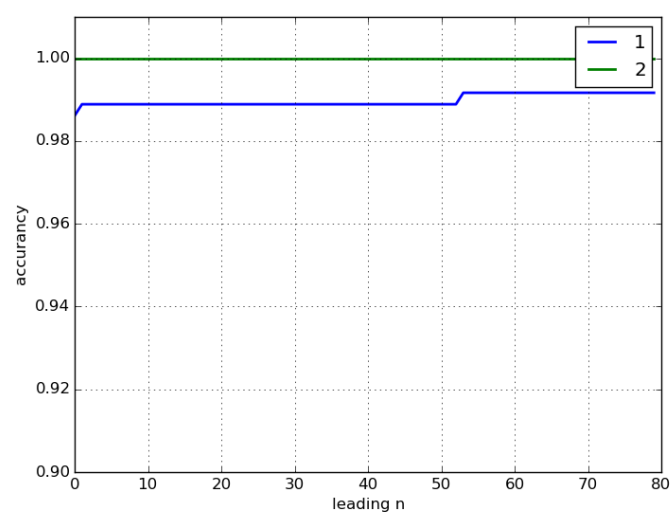


## 2.4.眼睛张开与闭合

### 2.4.1. 编号约定

- 1: 眼睛张开
- 2: 眼睛闭合

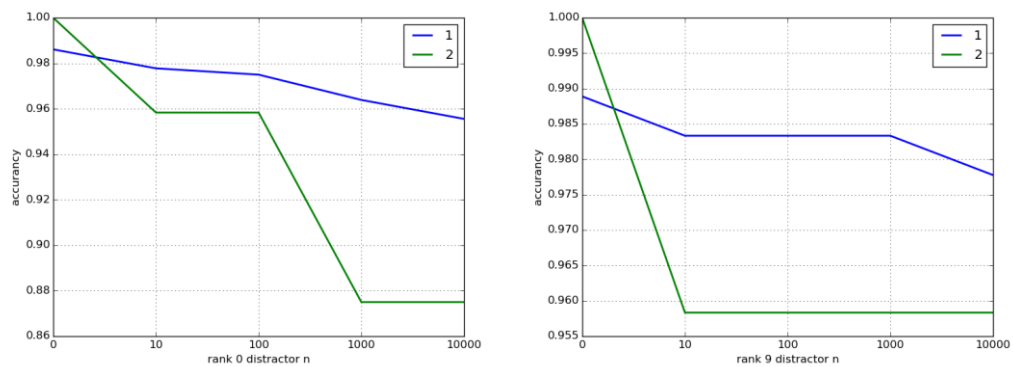
### 2.4.2. Rank n 曲线



在其它维度正常的情况下，眼睛张开与闭合对识别影响不大

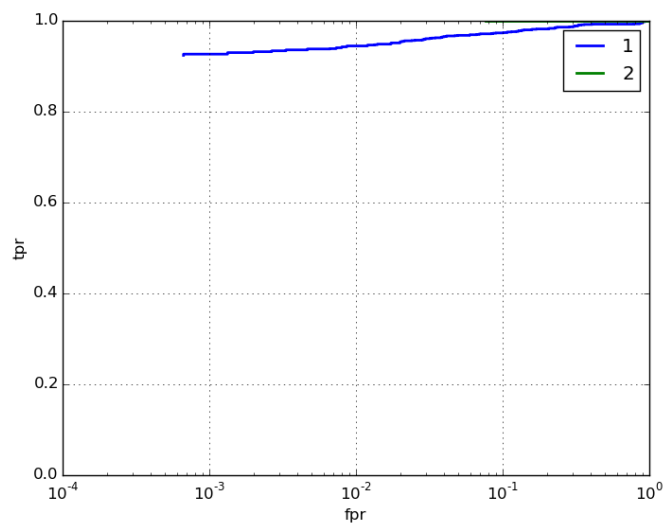
🚦 样本分布为(360,24)，因此出现眼睛闭合时 accuracy=1

### 2.4.3. Distractor n 曲线



🚦 眼睛张开识别情况整体优于眼睛闭合

### 2.4.4. ROC 曲线

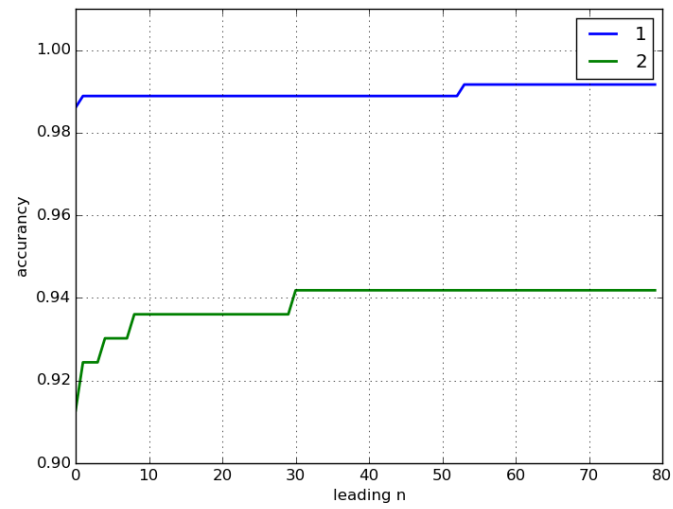


## 2.5. 嘴巴张开与闭合

### 2.5.1. 编号约定

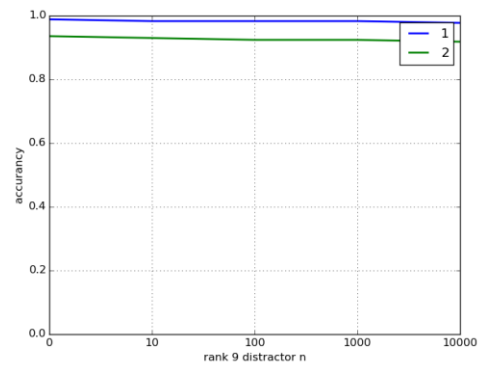
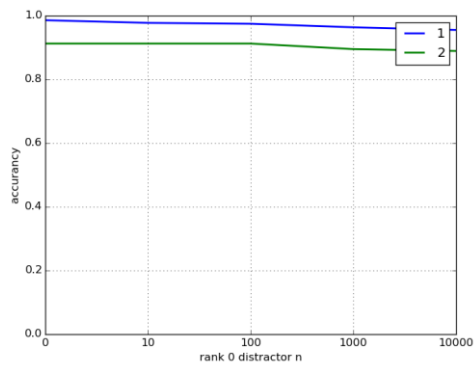
- 🚦 1: 嘴巴闭合
- 🚦 2: 嘴巴张开

### 2.5.2. Rank n 曲线

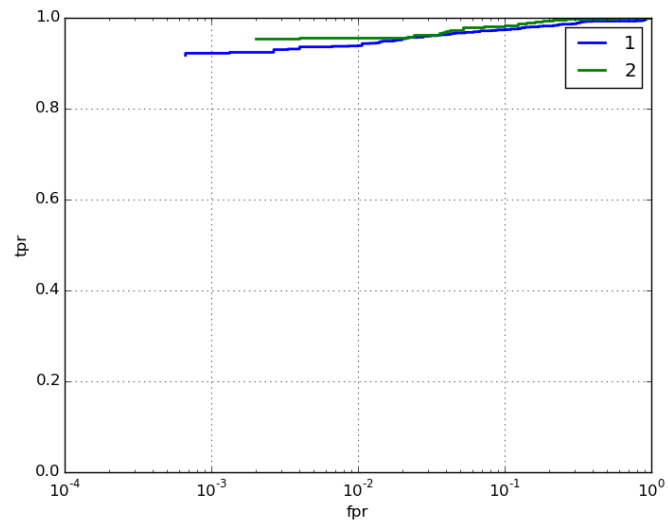


- 相对于遮挡、眼睛维度，嘴巴维度的影响更大一些
- 嘴巴闭合时识别率均优于嘴巴张开
- 数据分布为(360,172)，相对于遮挡、眼睛维度，此维度分析结果更具有可信度

### 2.5.3. Distractor n 曲线



### 2.5.4. ROC 曲线



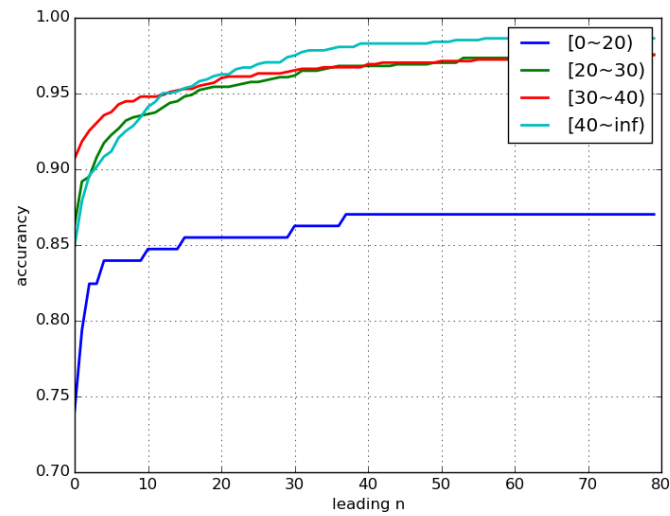
嘴巴闭合识别较差，与 rank  $n$ /distractor  $n$  反而不一致（可能是由于数据分布不理想）

## 2.6. 对比度

### 2.6.1. 编号约定

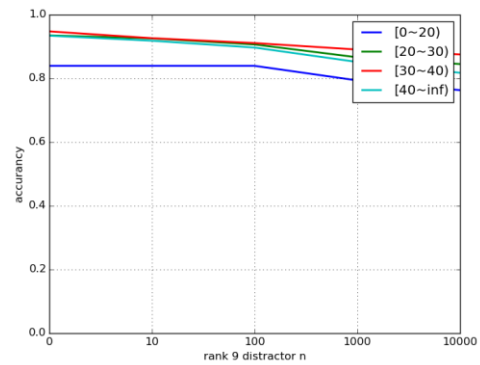
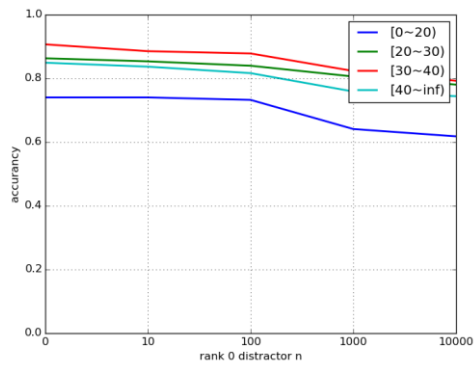
- 0-20: 对比度为 0-20
- 20-30: 对比度为 20-30
- 30-40: 对比度为 30-40
- 40-inf: 对比度为 40 以上

### 2.6.2. Rank n 曲线



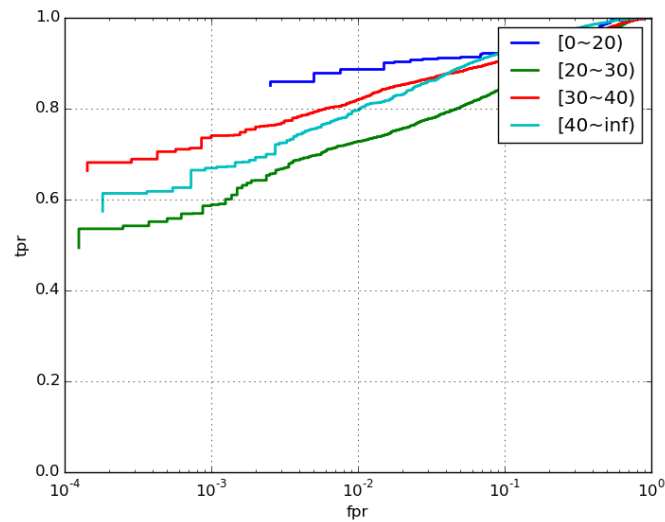
- 对比度在 30-40 时 rank0 取得了最好的 accuracy, 0-20 时最差
- 随着 n 的增加 20-40 之间 accuracy 基本一致, 40 以上的对比度照片 accuracy 增加不少

### 2.6.3. Distractor n 曲线



- 30-40 对比度优于 20-30, 优于 40+, 0-20 最差

## 2.6.4. ROC 曲线



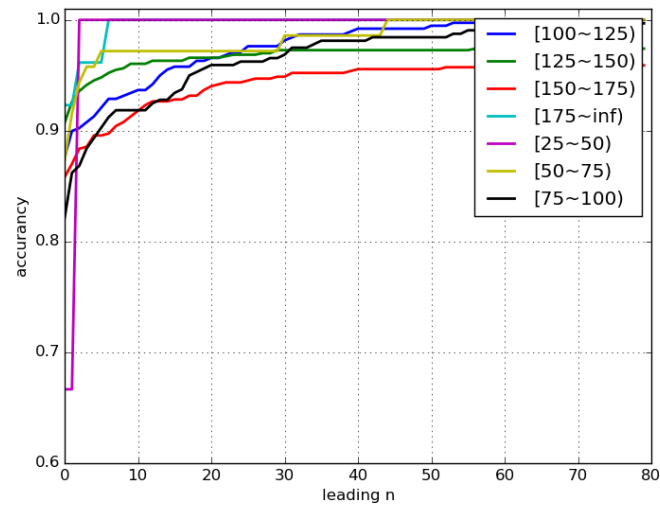
✚ rank n/distractor n 显示 30-40 对比度识别较好，但 ROC 曲线中并没有明显差别

## 2.7.照片亮度

### 2.7.1. 编号约定

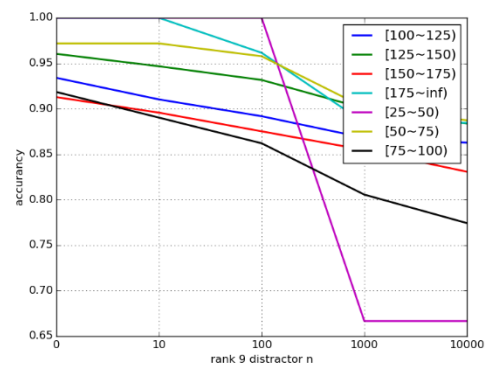
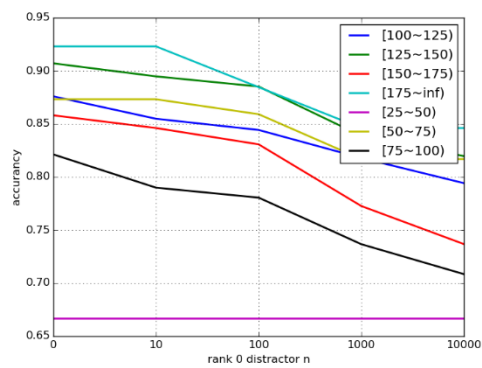
- ✚ 25-50: 亮度为 25-50
- ✚ 50-75: 亮度为 50-75
- ✚ 75-100: 亮度为 75-100
- ✚ 100-125: 亮度为 100-125
- ✚ 125-150: 亮度为 125-150
- ✚ 150-175: 亮度为 150-175
- ✚ 175-inf: 亮度为 175 以上

## 2.7.2. Rank n 曲线



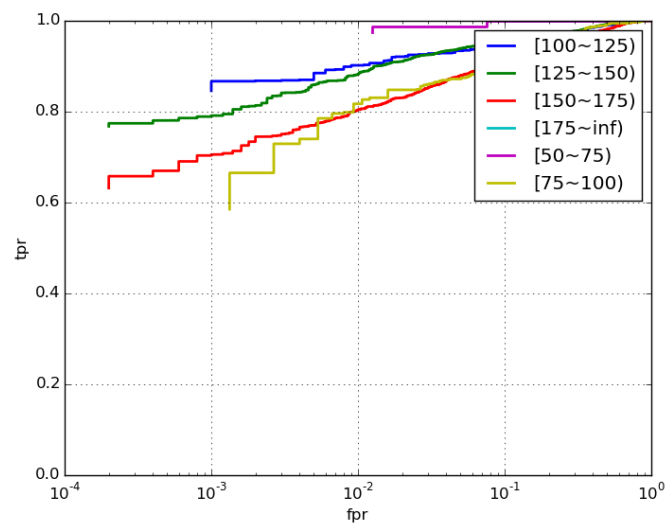
- ✚ 175 以上的亮度在 rank0 时识别情况最好，25-50 最差，但 25-50 在 rank3 时突然 accuracy=1，而 175 以上的亮度在 rank6 时 accuracy=1
- ✚ 亮度整体识别情况：(175+)>(50-75)>(125-150)>(100-125)>(150-175)>(75-100)>(25-50)
- ✚ 25-50 的样本数量为 3，50-75 样本数量为 71，175+ 样本数量为 26

## 2.7.3. Distractor n 曲线





## 2.7.4. ROC 曲线



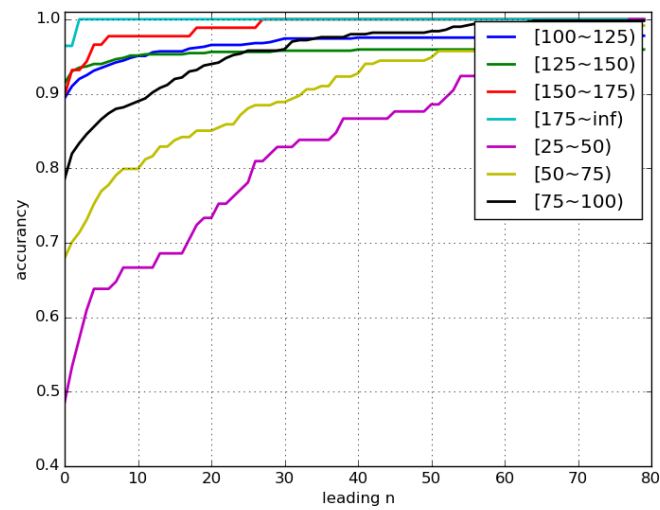
- ✚ 175+与 50-75 的识别非常理想
- ✚ 100-125 与 125-150 之间的识别情况比较接近
- ✚ 75-100 与 150-175 之间识别情况比较接近

## 2.8. 饱和度

### 2.8.1. 编号约定

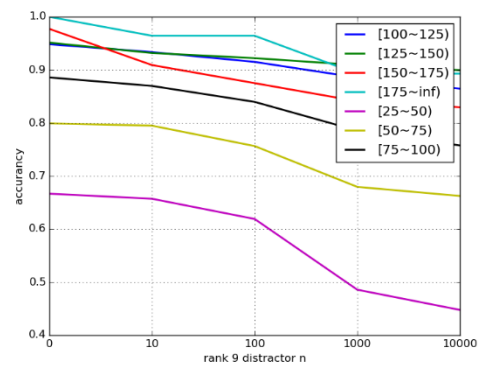
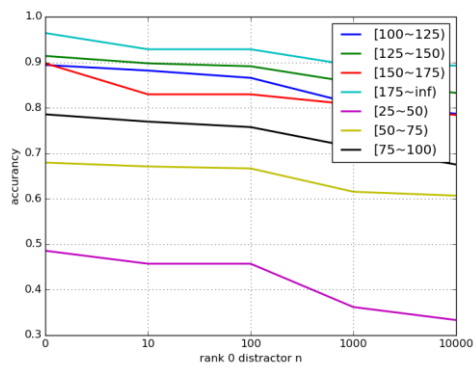
- ✚ 25-50: 饱和度为 25-50
- ✚ 50-75: 饱和度为 50-75
- ✚ 75-100: 饱和度为 75-100
- ✚ 100-125: 饱和度为 100-125
- ✚ 125-150: 饱和度为 125-150
- ✚ 150-175: 饱和度为 150-175
- ✚ 175-inf: 饱和度为 175 以上

## 2.8.2. Rank n 曲线

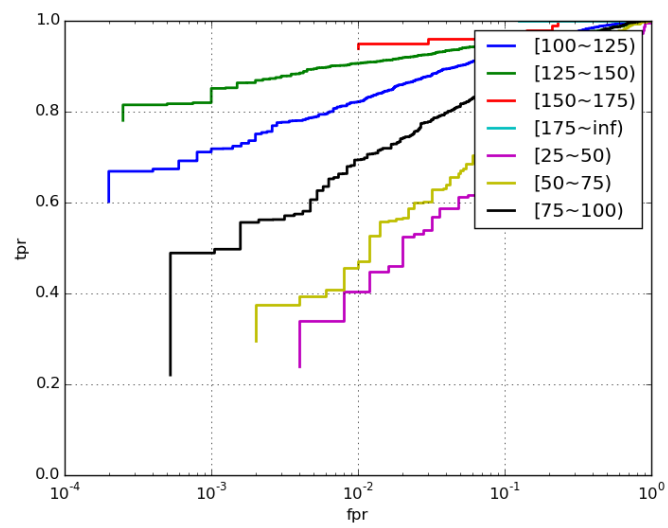


✚ 175 以上饱和度和识别率最高，在 rank10 之前的整体识别情况为：175+>(150-175)>(125-150)>(100-125)>(75-100)>(50-75)>(25-50)，可以看出，饱和度越大，识别率越高

## 2.8.3. Distractor n 曲线



## 2.8.4. ROC 曲线



- 175 以上的样本对为 8: 8, 因此出现了  $tpr=1, fpr=1$  的理想情况
- 饱和度越高, 识别情况越好, 规律非常理想

### 3. 小结

场景因素的分析是一个十分难以精确掌控的过程,为了分析单一因素对算法识别情况的影响,需要采集大量其它因素正常、被分析的维度分布多样化的数据集。数据集的获取是一个十分困难的过程,很难得到理想且数据量丰富的照片资源。

我们尽可能地构建了一个单一维度变化的分析环境,通过参考标准的 MegaFace 数据集分析方法,我们选取了 rank n、distractor n 和 ROC 曲线来分析维度对识别结果的影响。Rank n 和 distractor n 曲线仅作为一个参考曲线,ROC 曲线比较全面反映了算法的识别性能。通过计算 ROC 曲线的面积 AUC 与最佳工作点 Distance 可以有效的衡量算法在数据集上的性能。

分析结果总体来讲与我们的直观感觉基本吻合,例如侧脸大角度时识别率急速下降,正常光照时识别情况最好。但是通过客观的分析,我们得出了一些不那么好理解的结论,如弱光对识别的影响并不大,睁开眼睛也并不会提升多大识别率,这些结论也许不会有多大的意义,但至少可以让我们更理性地认识整个系统。

下面列出了一些不那么直观的结论:

1. 侧脸小角度识别情况非常接近正面,仰视识别率要高于俯视,大角度的识别率最差
2. 弱光对识别影响并不大,顶光对识别是有提升的
3. 头发遮挡、眼睛张开与闭合对识别结果的影响存在,但并不明显
4. 嘴巴闭合时识别情况要更好
5. 对比度在 30-40 时识别情况最好,但影响并不大
6. 亮度在 175+与 50-75 时识别最好
7. 饱和度越大,识别情况越好

另外,我们也发现了三个矛盾的结论,在灯光、嘴巴张开情况与对比度分析中,ROC 曲线的结论反而与前面的结论不一致,有可能是数据集不理想导致。