**武汉市公安局动态人像识别系统**

**测试技术方案**

**一、测试目的**

了解各厂商动态人像监控比对系统的实际性能表现；评估现有人脸监控比对系统产品对公安实战应用需求的满足程度；测试在实验室及现场模拟环境下人脸识别技术及其软（硬）件的性能；根据测试结果，针对性地优化系统实施方案，优化相关设备，改进人脸识别算法，优化软件系统，为后期人脸识别技术手段建设选型提供数据依据。

**二、组织方式**

采用实验室测试和实战测试相结合的方式。实验室测试占80%，实战测试占20%。

**三、测试环境**

（一）实验室测试。测试各厂商人脸识别软件在同一环境下对同一任务的完成度。

1、硬件及软件。组织方提供服务器（联想RD640 2\*E5-2620 6核12线程 2.5G 15M三级缓存/2\*8G内存/3\*600G 15K SAS RAID5/2\*800W电源）1台、台式机（i5-4440，3.10GHZ，4G内存，1TB硬盘，WIN7旗舰版SP1-64位操作系统）1台。参测厂商自行提供参测软件及辅助工具，参测厂商向组织方提供系统配置清单。

2、测试源。组织方提供人脸库、视频样本（任选海康、大华、三星或宇视中一种200万像素摄像机录制的原始AVI视频）及静态测试照片（主要包含质量、角度、环境等不确定因素造成的非标准人脸相片）。测试视频从表1视频样本中四段。视频样本类型见表1，播放1次。

表1 视频样本类型表

|  |  |
| --- | --- |
| 样本序号 | 人员 |
| 视频样本1 | 单人顺序通过 |
| 视频样本2 | 双人并行通过 |
| 视频样本3 | 多人自然通过 |
| 视频样本4 | 其他复杂情况 |

另有1段逆光视频，1段遮挡视频作为附加题目

3、实验室测试环境搭建。

**功能测试**（按细则要求功能实现），重点展示各参测单位的特色功能，注意功能列表提交；**性能测试：**10万人脸库测试4段AVI格式录像，作为评分，另附加2段作为附加测试；15万以上人脸库静态图片测试。电子质检院提供服务器1台和客户端计算机1台。

人脸库标本为10万张标准照片（按需加入若干目标人员标准照片），照片为JPG格式，各厂商按自己算法自行建库。

（二）实战测试。

测试各厂商人脸识别软件在实际环境下对同一任务的完成度。

在武汉火车站一个安全检查门后安装摄像机，每个企业安装，测试两天，有特殊要求可提前提出，期间不定时安插测试人员进入，检测效果。提供测试库实战，整个一套系统可自带安装，也可由检测方提供摄像机等设备（事先提出）。

1、硬件及软件。参测厂商自行提供参测软件及辅助工具，参测厂商向组织方提供系统配置清单。

2、实战环境搭建。实战地点武汉火车站，由参测厂商搭建硬件和软件系统，组织方提供电源和人脸库。

（三）自主实战。参测厂商自行提供参测软件及辅助工具，在武汉火车站，利用公安业务主题库开展人脸捕捉、识别、报警等项目的实战应用。

**四、测试内容**

1、软件符合性测试：

| 序号 | 检验项目 | | 技术（标准）要求 |
| --- | --- | --- | --- |
| 1 | 安装  与卸载 | 系统安装 | 按照安装手册进行安装可以成功。 |
| 系统卸载 | 可以正常卸载。 |
| 2 | 功能 | 功能模块 | 详见软件产品测试功能列表。 |
| 3 | 用户  界面 | 界面输入 | 操作方式符合用户习惯。 |
| 光标可以正常定位在输入区域内。 |
| 输入域可以正常输入，并能进行有效性检查。 |
| 界面文字 | 无乱字符出现。 |
| 界面显示 | 在正常分辨率下界面显示不变形。 |
| 4 | 用户  文档 | 版本一致性 | 用户文档的版本与实际使用的软件版本一致。 |
| 文档描述与程序的一致性 | 用户文档描述和软件实际功能一致。 |
| 用户文档描述 | 可以理解。 |
| 提供了纸介质的说明。 |
| 用户手册的主要功能和关键操作有相应的描述。 |
| 5 | 安全  稳定性 | 软件容错性 | 发生错误时有提示，对数据进行了有效性检查。 |
| 运行稳定性 | 没有发生由于软件错误导致的退出系统和死机现象。 |
| 6 | 易用性 | 易操作性 | 考虑使用的快捷方便，具有快捷工具按钮。 |
| 易理解性及易学性 | 有专业基础的用户对系统易理解、易学习。 |

2、本次测试包括两个主要方面：功能和性能。其中：功能测试占15分，性能测试占75分，主观评价项目由专家组评判，占10分。

（一）功能指标测试。总分15分，各功能具体分值见表2。

表2 功能测试项目表

|  |  |  |
| --- | --- | --- |
| 功能模块 | 功能描述 | 备注 |
| 采集、比对功能（5分，每个小项1分） | 人脸采集、存储、查询。 | 此项共计15分 |
| 人脸分类建库。 |
| 监控自动比对功能。 |
| 自动过滤系统（采集过程中不能满足质量要求的照片自动滤除）。 |
| 支持RTSP协议高清摄像机、视频接入。 |
| 报警管理（5分，每个小项1分） | 报警有提示。 |
| 报警结果单笔、批量导出Excel或Word。 |
| 报警组合查询功能：按逻辑库名称、目标人姓名、目标证件号码、相似度、报警时间组合查询。 |
| 查询应无明显延迟感（不大于1秒）。 |
| 报警记录可以查看与删除功能，支持批量删除。 |
| 巡检比对结果管理（1.5分，每个小项0.5分） | 比对结果审核。 |
| 人像比对结果目标人信息查看及人脸照片放大功能。 |
| 人脸比对信息打印功能。 |
| 用户管理（1.5分，每个小项0.5分） | 可以新建、编辑、删除用户。 |
| 具有锁屏功能。 |
| 界面中有用户、单位或地址显示。 |
| 系统管理（1分，每个小项0.5分） | 参数设置（比对相似度上下限，比对人脸像素上下限）。 |
| 日志管理。 |
| 数据管理（1分，每个小项0.5分） | 人脸库增删改，单笔录入，批量录入。 |
| 设置单笔或批量的布控与撤控。 |
| 特色功能 | 厂商自家特色（报名时提供该特色功能列表）。 | 主观评价 |

（二）性能指标测试。总分75分（实验室测试分数为75\*80%=60分，实战测试分数为75\*20%=15分）。实验室动态人脸视频的识别测试占50分，带偏转角度及图片质量较差的静态照片的识别测试占10分。

1、实验室动态人脸视频测试。目标人员每正确报警一次加m分（m取50与4段基本视频中目标人通过总人次的比值），非目标人员每错误报警一次扣m/2分，目标人员每错误判定一次扣m/2分，计算出得分pi。4段测试视频，测试过程中每段播放1次。此项得分a=∑pi（i=1,2,3,4）。该得分0分下限，50分上限。

得分pi公式如下：

pi=mx-m(y+z)/2

pi------动态人脸视频测试一次的得分

m-----正确报警一次所加分数

x------目标人员正确报警次数

y------非目标人员错误报警次数

z------目标人员错误判定次数

2、实验室静态照片识别测试。取50张目标测试照片（20张正面人员照片，30张侧面人员照片），采用JPG格式，15万以上人脸库。目标人员正确报警一次加0.2分，非目标人员错误报警一次扣0.1分，目标人员错误判定一次扣0.1分。此项0分下限，10分上限。

表3 静态图片性能测试项目表

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 测试项 | 测试项说明 | 备注 |
| 1 | 目标人员照片正确报警次数（次） | 当目标人员测试照片提交比对时，系统正确命中目标人员入库照片即算正确报警。 |  |
| 2 | 非目标人员照片错误报警次数（次） | 当非目标人员测试照片提交比对时，系统只要报警，既为错误报警一次。 |  |
| 3 | 目标人员照片漏报次数（次） | 当目标人员测试照片提交比对时，若系统在整个过程中均未能正确报警，则为目标人员漏报一次。 |  |
| 4 | 目标人员照片错误判定次数（次） | 当目标人员测试照片提交比对时，系统错误判定为其他人员的次数。 |  |
| 5 | 静态图片正确报警率（%） | 目标人员正确报警次数/目标人员测试样本总数×100％。 |  |
| 6 | 静态图片目标人员漏识率（%） | 目标人员漏报次数/目标人员测试样本总数×100％。 |  |
| 7 | 静态图片目标人员误识率（%） | 目标人员错误判定次数/目标人员测试样本总数×100％。 |  |
| 8 | 静态图片非目标人员识别错误率（%） | 非目标人员错误报警次数/目标人员测试样本总数×100％。 |  |

3、实战测试。打分规则与计算公式均与实验室动态人脸视频识别测试相同（实战测试中m取15与目标人（组织方演员个数）通过总人次的比值）。此项0分下限， 15分上限。

表4 动态视频性能测试项目表

|  |  |  |  |
| --- | --- | --- | --- |
| 序号 | 测试项 | 测试项说明 | 备注 |
| 1 | 目标人员正确报警次数（次） | 当目标人员通过一次时，系统报警，不论报警几次，为正确报警一次。 |  |
| 2 | 非目标人员错误报警次数（次） | 当非目标人员通过时，系统只要报警，既为错误报警一次。 |  |
| 3 | 目标人员漏报次数（次） | 当目标人员通过时，若系统在整个过程中均未能正确报警，则为目标人员漏报一次。 |  |
| 4 | 目标人员错误判定次数（次） | 当目标人员通过通道时，系统错误判定为其他人员的次数。 |  |
| 5 | 报警响应时间（仅在实战中测算） | 目标人出现至报警响应的时间长度。 | 主观评价 |
| 6 | 正确报警率（%） | 目标人员正确报警次数/目标人员通过总次数×100％。 |
| 7 | 目标人员漏识率（%） | 目标人员漏报次数/目标人员通过总次数×100％。 |
| 8 | 目标人员误识率（%） | 目标人员错误判定次数/目标人员通过总次数×100％。 |
| 9 | 非目标人员识别错误率（%） | 非目标人员错误报警次数/非目标人员通过次数×100％。 |
| 10 | 入库成功率 | 输入人脸库图片成功率。 |
| 11 | 采集率 | 视频场景中采集人次与通过人次比值。 |
| 12 | 系统可靠性、稳定性 | 实战测试中系统可靠性、稳定性。 |
| 13 | 码流测试 |  |

（三）主观评价项目（总分10分+专家对实战的讲评）。

1、主观评价项目（10分）。由专家具体评分。

2、实战评价。所有测试完成后，参测方可进行48小时完全自主实战测试，主要评估系统稳定性、抗压性等内容。此项作为公安机关参考。

最后总得分K计算如下

K=a+b+c+d+e

a------实验室动态人脸视频测试得分

b------实验室静态照片识别测试得分

c------实战测试得分

d------主观项目评分

e------功能测试得分

**五、测试流程**

1、**公告：**在湖北省电子质检院官网发布测试公告。

2、**报名：**按公告要求，到湖北省电子质检院报名参测，填写委托书并交测试费用。

3、**测试排序：**报名时抓阄排测试顺序，报名结束后按顺序安排测试时间（电话通知参测点位联系人）。

4、**试验室测试：**每个参测单位一天安装部署，一天测试。测试主要内容：**功能测试**（按细则要求功能实现），重点展示各参测单位的特色功能，注意功能列表提交；**性能测试：**10万人脸库测试4段AVI格式录像，作为评分，另附加2段作为附加测试；15万以上人脸库静态图片测试。电子质检院提供服务器1台和客户端计算机1台。

5、**实战测试：**在武汉火车站一个安全检查门后安装摄像机，每个企业安装，测试两天，有特殊要求可提前提出，期间不定时安插测试人员进入，检测效果。提供扒窃库实战，整个一套系统可自带安装，也可由检测方提供摄像机等设备（事先提出）。

6、**自主实战测试：**完全自主实战测试由参测厂商自行决定是否参加，组织方统一安排时间，地点、提供电源和相关主题库。

7、**专家评分：**将测试结果及录像提供专家组，对结果分析评估。

湖北省电子信息产品质量监督检验院

2015年6月23 日