****



**本 科 毕 业 设 计**

院 系 计算机科学与技术系

专 业 计算机科学与技术

题 目基于K匿名及差分隐私的定域功能性影像隐私保护

年 级 2013级 学 号 131220124

学生姓名 肖鹏

指导老师 华景煜 职 称 讲师

论文提交日期 2017.05.22

**南京大学本科生毕业论文（设计、作品）中文摘要**

题目： 基于K匿名及差分隐私的定域功能性影像隐私保护

计算机科学与技术 院系 计算机科学与技术 专业 2013 级本科生姓名： 肖鹏

指导教师（姓名、职称）： 华景煜、讲师

摘要：

近些年来，随着机器学习及数据挖掘研究的深入与进步，大数据产业变得十分火热，无论是学术机构或者是社会部门，都十分关注大数据分析带来的收益与相关产业的发展。就目前的数据交换现状来看，数据的交互过程中存在严重泄露的风险，使得拥有大量数据的部门如刑警部门在需求相关学术机构对其所持有的数据进行分析时会有对数据泄露风险的担忧。

本文的主要内容是针对影像数据（照片，视频资料等），在大数据交互分析中会产生泄漏隐私风险的部分，人脸信息，寻求了一种较为安全的隐私保护构想与发布方式。通过借助差分隐私模型，通过对人脸统计数据添加干扰噪音，从而达到即便攻击者已经掌握了除必要数据以外的辅助数据，且拥有查询统计数据的能力，该攻击者仍然无法从影像资料中推测出具体到个体的特征。同时为了更好的满足数据可用性，保护手段中同时保障了数据发布的K匿名性。为此我们将该算法的实现分成三部分，第一部分利用人脸识别算法，将影像资料中的人脸数据提取到标准尺寸，并按规则划分。预先提取需要处理的资料。第二部分是本保护算法的重点，依据不同规则划分出的人脸单位进行内部面部相似度比较，每个个体提取出前K个与之最相似的人脸。将每个个体与之最相似的K－1个个体进行加权线性叠加，建立新的对应个体。并对统计量进行指数机制处理。第三部分，将新生成的个体与原数据影响进行面部相似度比较，重现提取前K个分析是否达到保护要求，并对发布统计数据的差分隐私的效果进行检验，并将新生成的个体重新覆盖到原始的影像资料中进行发布。

为了验证算法的有效性，收集356个不同的个体数据划分后进行交叉测试。实验结果证明经过标准化提取以及规则划分和加权线性重叠处理后，新生个体与原始个体平均相似度只有30％,在所有数据个体最相似排名中基本都大于1/K\*总体，满足K匿名的要求，统计量经过差分隐私保护后也不会严重破坏数据可用性。同时，整个算法执行过程在单规则分类下，167个个体处理完成及输出比对分析结果需要15min，这些结果说明本文提出的算法是高效可行的。

关键词：影像隐私保护；差分隐私；K匿名；图像识别

**南京大学本科生毕业论文（设计、作品）英文摘要**

THESIS：The Design and Implementation of A CFO-Feature Estimation Algorithm for Wireless Network Devices

DEPARTMENT：Computer Science and Technology

SPECIALIZATION：Computer Science and Technology

UNDERGRADUATE：Peng Xiao

MENTOR：JingYu Hua

ABSTRACT：

In recent years, developments in machine learning and promotions in data mining have boosted the big data industry greatly. As a consequence, attention from both academic institution and social sector has been intensively paid on the benefits of big data analysis and its impacts on related industries. But under current data exchanging circumstance, data interaction process has severe leakage risk, which exaggerates the concern of the data-heavy institutions, such as criminal police organization, especially when they have to seek help from related academic institutions to analyze their data held.

This paper will mainly demonstrate one safer privacy protection method and release mode, which focuses on face information, one main risk of image data (photos, videos, etc.) in big data interaction process. Using the Differential privacy model and adding disturbance noise to statistical data, and the ability to query statistical data, this method could prevent attacker from inferring specific individual actions from the video record even they have had the auxiliary data but no subject features. At the same time, in order to fulfill the data availability, the protection method also ensures K anonymity of data dissemination. We introduced three parts to realize the algorithm. In first part, data will be extracted in advance using the face recognition algorithm to extract the standard size of the face data in video information and distribute them by criteria. Second part is the essential component. Internal face similarity of face unit will be compared according to the different algorithmic rules. Each unit will extract the top K most similar face, and build new unit by Weighted linear superposition of the K-1 data. And use index mechanism modified statistical data. In the final part, face similarity will be compared between new unit and the original data, top K most similar data will be extracted again to analyze their security completeness, the effect of differential privacy on the publication of statistical data is examined and new generated unit will cover in the original video data.

In order to testify the effectiveness of this algorithm, 356 various individual data were collected to implement cross-test. The test result proves that the similarity rate is below 30% between new unit and original ones after standard extraction, compared according to the different algorithmic rules and Weighted linear superposition fulfill the requirements of K anonymity and statistics after differential privacy protection will not seriously undermine the availability of data. Meanwhile, the processing, outputting and comparing time of 167 units in classified by simple algorithmic rule is 15 minutes. Both results prove that the algorithm presented in this paper is efficient and feasible.

KEY WORDS: Image privacy protection，Differential privacy，K anonymity，image recognition

**目录**

[第1章 绪论 1](#_Toc481671756)

[1.1研究背景和意义 1](#_Toc481671757)

[1.2研究现状 2](#_Toc481671758)

[1.3本文研究内容 3](#_Toc481671759)

[1.4论文结构安排 4](#_Toc481671760)

[第2章 相关工作与背景介绍 6](#_Toc481671761)

[2.1灰度直方图 6](#_Toc481671762)

[2.1.1三通道RGB图像及MAT矩阵 6](#_Toc481671763)

[2.1.2灰度直方图转换算法 8](#_Toc481671764)

[2.2直方图对比算法 8](#_Toc481671765)

[2.2.1 Correlation 8](#_Toc481671766)

[2.2.2 Chi-Square 9](#_Toc481671767)

[2.2.3 Intersection 9](#_Toc481671768)

[2.2.4 Bhattacharyya距离 10](#_Toc481671769)

[2.3 图像可调加权线性合成算法 10](#_Toc481671770)

[2.4 隐私模型 11](#_Toc481671771)

[2.4.1 K匿名 11](#_Toc481671772)

[2.4.2 差分隐私 12](#_Toc481671773)

[2.5小结 13](#_Toc481671774)

[第3章 具体实现方案 15](#_Toc481671775)

[3.1方案概览 15](#_Toc481671776)

[3.2数据预处理 16](#_Toc481671777)

[3.2.1按规则划分数据集 16](#_Toc481671778)

[3.2.2人脸识别提取标准化 16](#_Toc481671779)

[3.2.3人脸类型提取统计 18](#_Toc481671780)

[3.3 相似度对比 19](#_Toc481671781)

[3.3.1 人脸图像灰度直方图化 19](#_Toc481671782)

[3.3.2 直方图交叉对比 21](#_Toc481671783)

[3.3.3 提取最相似的前K个主体 22](#_Toc481671784)

[3.4 统计量差分隐私保护与图像加权线性叠加 24](#_Toc481671785)

[3.4.1 统计量差分隐私保护 24](#_Toc481671786)

[3.4.2 图像处理权值选择 25](#_Toc481671787)

[3.4.3 叠加生成新主体 26](#_Toc481671788)

[3.5 生成样本测试 27](#_Toc481671789)

[3.5.1 K匿名检测 28](#_Toc481671790)

[3.5.2 差分隐私测试 30](#_Toc481671791)

[3.6算法实现总体流程图 31](#_Toc481671792)

[3.7数据发布模型流程图 32](#_Toc481671793)

[第4章 实验 34](#_Toc481671794)

[4.1实验环境与数据集 34](#_Toc481671795)

[4.2算法精度 34](#_Toc481671796)

[4.3算法时间效率 36](#_Toc481671797)

[4.4实验结论 37](#_Toc481671798)

[第5章 总结与展望 38](#_Toc481671799)

[5.1本文工作总结 38](#_Toc481671800)

[5.2本文的不足与展望 39](#_Toc481671801)

[参考文献 V](#_Toc481671802)

[致谢 VIII](#_Toc481671803)

第1章 绪论

1.1研究背景和意义

近些年来，随着机器学习及数据挖掘研究的深入与进步，大数据产业变得十分火热，无论是学术机构或者是社会部门，都十分关注大数据分析带来的收益与相关产业的发展，数据持有者与研究机构拥有大量的合作机会。但同时，数据发布所面临的安全风险就值得格外关注。从大数据产业的角度来看，数据拥有庞大的内在价值，同时从数据收集的个体来讲，数据中又同时携带有大量的个体隐私，如搜索纪录，银行卡信息，生理信息。相较于其他而言，文字数据类信息要相对容易进行匿名化处理，但是对于带有重要生理信息如人脸等的视频，照片影像类资料在发布过程中，其需要既要兼顾个人信息匿名化与数据可用性，这就使得数据隐私保护变得较为困难。

举个例子，刑警发现一起小区内盗窃案件，确定不是外来盗贼，刑警部门想要通过摄像头拍下来的视频资料分析小区内住户的日常行为有无异常。假设现在有多家研究机构可以帮助警方来进行分析，但是警方就会十分担忧这些影像资料发布后会带来的住户人脸数据泄露风险。

为了有效解决影像资料在发布时的隐私泄漏问题，人们如今采取了许多措施，比如直接把人像位置扣去，增加对应的编号等。还有给人脸面部打马赛克进行模糊化处理的。然而影像资料和普通的字符串数据资料不同，在分析起端需要进行图像识别处理工作，以上的处理方式会严重的造成数据可用性的下降，给分析工作造成困，另外大部分的处理发布也没有考虑满足差分隐私模型。在种种方案都不尽如人意的情况下，我们考虑结合K匿名及差分隐私的构想，同时兼顾数据可用性的前提下，选择通过在确定工作范围下，依据具体规则（如某特定地点）划分，提取人脸特征，比对后加权线性合成，替换原始人脸图像特征。

为了完成提取，对比，合成新的人脸图像同时兼顾完成差分隐私及K匿名的要求，需要寻求合适的人脸图像特征比对度量单位，还有合成中权值的选择。故接下来会介绍下目前图像特征相似度量以及一些图像隐私工作的研究现状。

1.2研究现状

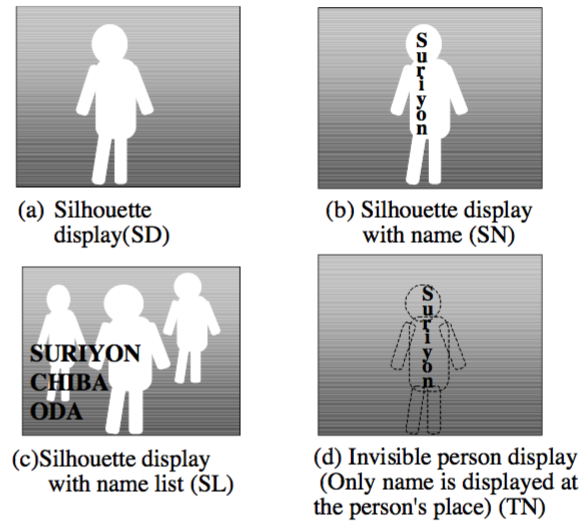
图像相似度比较是人脸识别中一个重要组成部分，也同时广泛应用于大数据分析，机器学习，信息检索，统计学等多个领域。图像相似度对比的算法有很多，应用场景与要求也差异颇大，最常用的是数字图片灰度直方图化后，通过四种直方图对比标准来判定两张图片的相似程度。

对比的要求首先要将图像转化为灰度直方图，灰度直方图是灰度级的函数，描述图像中该灰度级的像素个数（或该灰度级像素出现的频率）：其横坐标是灰度级，纵坐标表示图像中该灰度级出现的个数（频率）。这里要求图像是单通道的灰度图像，对应建立一维直方图，对于彩色图像的例如RGB图像，可以对应不同通道建立多维直方图。之后可以应用Correlation，Chi-Square，Intersection，Bhattacharyya距离来度量相似程度。对于上面的四种对比标准，属于统计学领域范畴，考虑到本次实验的数据标准参考标准照片库，为保证试验的有效性，将对实验数据结果通关以上四种标准混合检验。

图像发布，传输中出现的隐私泄漏问题在近几年社交平台发展，大数据分析技术提升中越发引人关注，也有很多学者提出了自己的保护方式及观点。在一份方案中[3]，探究了在facebook上发布私人照片的隐私泄漏问题，其探究的面较广，不仅限于人脸信息，还有信用卡，身份证等。作者选择通关图像识别技术将敏感部分直接制空后再发布。但是作者并没有探讨这些影像资料在学术研究及大数据挖掘中的价值，其保护方式会对影像数据的可用行严重破坏。如下图

**图1-1 （1）中图片处理样例**

还有一篇论文中考虑到了视频资料的研究价值[2]，给出了逐步递进的四种处理图像的方式，Silhouette display，Silhouette display with name，Silhouette display with name list，Invisible person display (Only name is displayed at the person's place)，其核心思想是依据不同的保护等级对视频资料进行保护，可以说很大程度上保胡了图像的分析可用性，但是没有细致的探讨对差分隐私的考虑，可以说攻击者如果掌握基本的附加信息如身份ID等还是可以很容易的确定具体主体的行为数据。同时，他保护的方式是对整个人体形态进行处理，故对微小的行为信息有一定的干扰。如下图：



**图1-1 （2）中图片处理方案**

另外还有一篇论文中[1]详细探讨了图片隐私保护的分级依据对图片数据可用性，考虑了差分隐私模型定义了完整一套规则去度量不同的图片模糊手段对于图像数据破坏程度。文章中比对了几种基本的图片保护处理方式，加马赛克，直接抹黑，图像聚合，线性叠加等不同方式下度量的高低，依据其试验结果，线性叠加保护的效果远好于普遍认为聚合处理较优的结果。但是文中仅仅考虑了单通道的灰度图像没有考虑三通道情况下的叠加效果，另外也没有探讨在具体环境及应用场景和统计数据的可用性分析。

1.3本文研究内容

本文为了获得统一标准化的处理后的图像，首先对所有图片进行人脸识别并且将提取的人脸标准化提取，按照模拟的具体规则进行归类，然后基于opencv直方图算法将图像转化为灰度直方图，在同规则类下的人脸直方图互相按照直方图比对规则选择前K个最相似的主体。

然后开始进行加权线性偶合，权数的选择依据四种不同的直方图比较规则进行优化，合成后的图片重新规则话后依据同上规则划分后保存。

接着，为了检验合成后图片的K匿名性，将合成后的图像重新按照之前的比较方式进行比较，重新选择前K个最相似的合成后主体，其汇总结果与第一轮进行比对看加入噪音后的新主体的K匿名性是否达到要求。

人脸中存在很多的特征信息，如性别，戴眼镜与否，瞳色，瞳距，人脸类型等，各种信息都包含了主体的个人隐私信息但是对于数据的可用性有很重要的作用，在影像资料经过保护后，这些信息被模糊了，为了保证数据的可用性得到最大的保证，本文借助差分隐私模型中从人脸类型的角度进行保护。在图像库统计脸型分布，在查询时使用指数机制，调整隐私参数，寻找最优的保护效果。

最后，将训练好的参数投入系统重新生成最后的加入噪音后的图像与统计数据共同发布，并统计最终比对结果。

1.4论文结构安排

第一章为绪论，主要介绍了目前影像资料在发布中可能出现隐私泄漏的风险于目前的研究现状。同时简单的对本文研究的关于在考虑差分隐私以及K匿名安全的保护系统进行了描述，并介绍了其核心的图像对比算法标准以及目前前人关于影像隐私保护的研究现状。

第二章简单地介绍了与本文相关的工作背景和概念定义。首先介绍三通道图像的概念及灰度直方图转换算法。然后介绍图像相似度对比的算法基础，并分别对K匿名性，差分隐私概念进行了介绍，为后面算法据其进行参数选择做好准备。

第三章将本文将图像加密算法分成了3部分进行详细描述，图像标准化提取按规则划分、内部相似度排序加权叠加和轮换检测方法调整参数。并分析了相关部分的运行效率，最后给出了算法结构图。

第四章对实验结果进行分析，统计了在各权值下各个数据集的测试结果，总结了在权值变化下得到的验证结果，在目前数据集上该算法的K匿名成功率92.7%以上，处理效率都较高。

第五章总结了本文的主要研究内容，并对其进行总体评价。盘点了本算法的优势与限制，并指出了就不足之处接下来的工作方向，提出了对于算法的加速可能方案。

第2章 相关工作与背景介绍

2.1灰度直方图

直方图英文名称是(HISTOGRAM) [6]，相机上显示的直方图和PHOTOSHOP使用的直方图都是灰度直方图，从图形上说，它是一个二维图，用坐标表示。横坐标表示图象中各个像素点的灰度级。(0到255个级别,一般人眼能够分辨的只有32个级别，人眼对光的强度变化非常敏感，而对颜色的变化就比较弱，目前，流行的视频压缩软件都是应用这一原理，比如RM)它是多种空间域处理技术的基础。直方图操作能够有效用于图像增强；提供有用的图像统计资料，其在软件中易于计算，适用于商用硬件设备。纵坐标为各个灰度级上图象各个像素点出现的次数或概率.各个软件细分程度不同。

1.表征了图像的一维信息。只反映图像中像素不同灰度值出现的次数（或频数）而未反映像素所在位置。

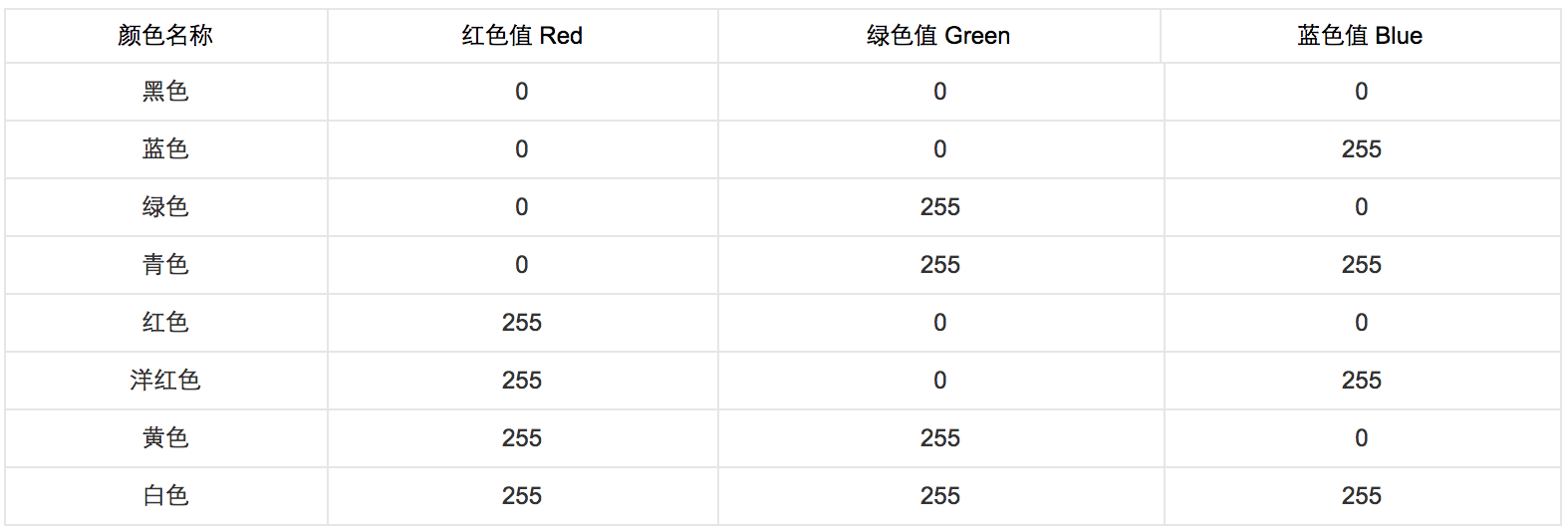
2.与图像之间的关系是多对一的映射关系。一幅图像唯一确定出与之对应的直方图，但不同图像可能有相同的直方图。

3.子图直方图之和为整图的直方图。

2.1.1三通道RGB图像及MAT矩阵

RGB色彩模式是工业界的一种颜色标准[6]，是通过对红(R)、绿(G)、蓝(B)三个颜色通道的变化以及它们相互之间的叠加来得到各式各样的颜色的，RGB即是代表红、绿、蓝三个通道的颜色，这个标准几乎包括了人类视力所能感知的所有颜色，是目前运用最广的颜色系统之一。

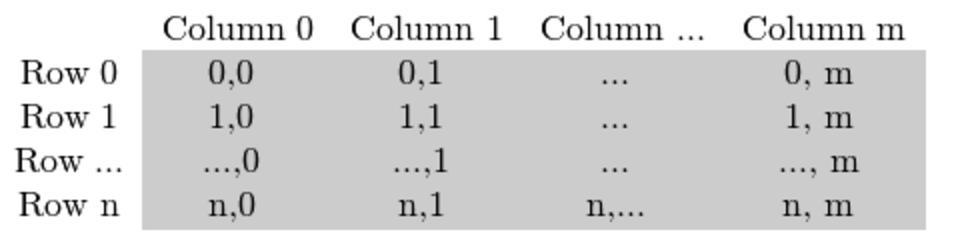
RGB 颜色称为加成色，因为您通过将 R、G 和 B 添加在一起（即所有光线反射回眼睛）可产生白色。加成色用于照明光、电视和计算机显示器。例如，显示器通过红色、绿色和蓝色荧光粉发射光线产生颜色。绝大多数可视光谱都可表示为红、绿、蓝 (RGB) 三色光在不同比例和强度上的混合。这些颜色若发生重叠，则产生青、洋红和黄。



**图2-1 基本颜色三通道数值**

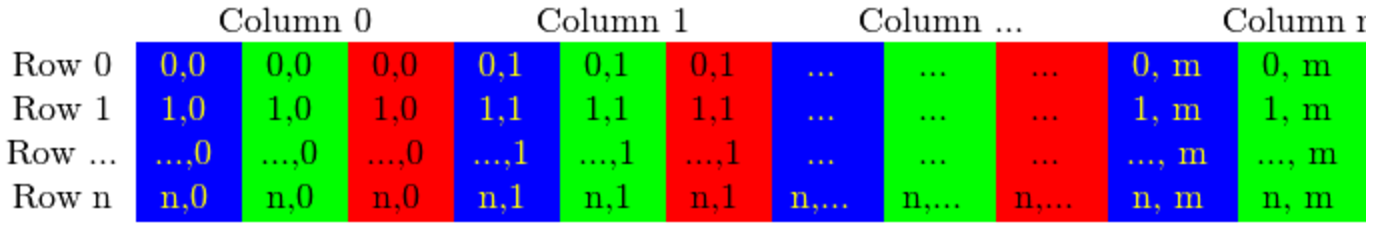
**在**对图片处理时，例如访问每个像素点，修改色彩等通常也是将图片每个像素点分离出三通道数值再进行操作，Opencv中的Mat矩阵给我们提供了方便快捷的操作方式。

如果图像是一幅灰度图像，他就像这样，从左到右，从上到下，依次是矩阵的每一行每一列，这时候矩阵Mat(i,j)的值自然就是当前点的灰度值了。



**图2-2 灰度图像Mat矩阵构成**

而对于一幅彩色图像，由于它的像素通道并不是一个，所以每一列又分为了通道。拿常见的RGB图像来说，就像这样：



**图2-3 RGB图像Mat矩阵构成**

从上图可以看出内存中是如何存放多通道图像，但是实际上在RGB模型中，Mat每一列依次是Blue，Green，Red顺序与名称相反。在本文中，在我们讨论图片相似度对比时往往会优先把RGB图像灰度化，而在讨论加权线性合成时往往是三通道图像。

2.1.2灰度直方图转换算法

灰度直方图[7]是灰度级的函数，描述图像中该灰度级的像素个数（或该灰度级像素出现的频率）：其横坐标是灰度级，纵坐标表示图像中该灰度级出现的个数（频率）。其算法为：

其中，代表某像素个数，代表总像素数，高维直方图可以理解为图像在每个维度上灰度级分布的直方图。常见的是二维直方图。如红-蓝直方图的两个分量分别表示红光图像的灰度值和蓝光图像灰度值的函数。其图像坐标（Dr,Db）处对应在红光图像中具有灰度级Dr同时在蓝光图像中具有灰度级Db的像素个数。这是基于多通道数字图像，二维中对应每个像素统计个变量。本文中在对比图像相似度时是将RGB图像转化为灰度图像，求得其灰度直方图再进行对比的。

2.2直方图对比算法

为了度量多幅图片之间的相互相似程度[8]，我们在将图片灰度直方图化后，需要选择相应的标准进行度量，以下将对统计学中四种比较标准进行介绍。

2.2.1 Correlation

Correlation相关系数[10]，或称线性相关系数、皮氏积矩相关系数（Pearson product-moment correlation coefficient，PPCC）等，是衡量两个随机变量之间线性相关程度的指标。它由卡尔·皮尔森（Karl Pearson）在1880年代提出，现已广泛地应用于科学的各个领域。依据相关现象之间的不同特征，其统计指标的名称有所不同。如将反映两变量间线性相关关系的统计指标称为相关系数（相关系数的平方称为判定系数）；将反映两变量间曲线相关关系的统计指标称为非线性相关系数、非线性判定系数；将反映多元线性相关关系的统计指标称为复相关系数、复判定系数等。

相关系数是衡量观测数据之间相关程度的一个指标，相关关系是一种确定性关系，相关系数是研究变量之间线性相关程度的量，一般情况下，相关系数越大表明相关程度越高，其算法公式如下：

其中

N是直方图中bin的数目。

2.2.2 Chi-Square

Chi-square test方检验[11]是用途非常广的一种假设检验方法，它在分类资料统计推断中的应用，包括：两个率或两个构成比比较的卡方检验；多个率或多个构成比比较的卡方检验以及分类资料的相关分析等。

卡方检验就是统计样本的实际观测值与理论推断值之间的偏离程度，实际观测值与理论推断值之间的偏离程度就决定卡方值的大小，卡方值越大，越不符合；卡方值越小，偏差越小，越趋于符合，若两个值完全相等时，卡方值就为0，表明理论值完全符合，其算法公式如下：

2.2.3 Intersection

Intersection直方图交叉距离[12]是一种较为特殊的集合差异检验，通过取两个类中相同属性值中较小的值累加，其累加之值为度量两个集合直方图差异的数据，值越大，两个集合越相似，其计算方式如下：

2.2.4 Bhattacharyya距离

巴氏距离[13]，它在统计学中用于测量两种离散概率分布的可分离性。在直方图相似度计算时，巴氏距离获得的效果最好，但计算是最为复杂的。设为的巴氏距离计算结果，其值完全匹配为1，完全不匹配则为0。

Bhattacharyya系数（Bhattacharyya Coefficient, 巴氏系数）是对两个统计样本的重叠量的近似计算。巴氏系数可用来对两组样本的相关性进行测量。

计算巴氏系数涉及到对该两个样本的重叠部分进行基本形式的积分。两个样本值的积分被分成指定数目的部分。而每一个样本的每一个部分的成员数被用于下式中：

2.3 图像可调加权线性合成算法

两个图片的融合，首先两个图片大小必须一致，如果不一致怎么办，要使用ROI区域，找到大图片中要和小图片进行融合的尺寸。两个图片的像素值，按照给定的权值进行叠加，就可以将两个图片融合起来。

其算法表达式如下：

在三通道RGB图像下，其算法模型如下：

B = psrc1[j\*channels] \* alpa + psrc2[j\*channels] \* beta + gama;

G = psrc1[j\*channels + 1] \*alpa + psrc2[j\*channels + 1] \*beta + gama;

R = psrc1[j\*channels + 2] \*alpa + psrc2[j\*channels + 2] \*beta + gama;

if (B>255)

B = 255;

if (G>255)

G = 255;

if (R>255)

R = 255;

pdst[j\*channels] = B;

pdst[j\*channels + 1] = G;

pdst[j\*channels + 2] = R;

2.4 隐私模型

随着信息技术的发展,个人隐私的概念发生了很大的变化,人们对隐私信息的保护也越来越重视。另一方面,信息的大量增长,使人们陷入了数据丰富而知识匮乏的困境,数据挖掘技术能够对海量数据进行处理,发现数据中潜在的关联关系。然而,数据挖掘技术的应用增加了隐私泄露的风险,因此数据需要在发布前进行匿名化处理,以提供对个人隐私信息的保护[26]。为了度量某种保护手段所起到的作用，近些年研究出现了很多种隐私保护模型，例如K匿名，差分隐私，l-Diversity等，本文中设计的影像保护算法模型兼顾数据的可用性及匿名性，以下就对涉及到的K匿名及差分隐私模型进行简单叙述。

2.4.1 K匿名

数据表的k-匿名化[9] ( k-anonymization) 是数据发布时保护私有信息的一种重要方法。 k-匿名技术是1998 年由Samarati和Sweeney提出的 ,它要求发布的数据中存在一定数量(至少为k) 的在准标识符上不可区分的记录,使攻击者不能判别出隐私信息所属的具体个体,从而保护了个人隐私, k-匿名通过参数k指定用户可承受的最大信息泄露风险。k-匿名化在一定程度上保护了个人的隐私,但同时会降低数据的可用性。因此, k-匿名化的研究工作主要集中在保护私有信息的同时提高数据的可用性。

假设一条微数据记录对应于一个个体，从广义集合论的角度出发，K-匿名就是要求在一个集合中只能以不大于1/K的概率确定任何一个元素，即要求任何一个元素，在集合中至少存在K-1个和其相同副本元素。个体特征通过相应的属性值来区分。将上述K-匿名概念引申到关系型数据库中，称一个关系表是K匿名化的，当且仅当对于任何一条微数据，在所有准标识属性的投影中，该表中存在至少k-1条记录与之完全相同。

k-匿名的基本思想是通过匿名化（如泛化和隐匿）原始数据集中的某些属性值形成满足一定匿名要求的匿名数据集并可用于数据发布。而在信息发布时实施数据隐私保护要考虑两个方面问题:（1）确保数据发布过程中隐私不泄露；（2）发布的匿名共享数据具有实用性[27]。

本文中描述的影像资料隐私保护算法正是借助k匿名的概念，保证分组发布的用户数据即使攻击者掌握了用户的身份ID信息，也只有小于等于1/K的概率对应到相应的主体信息。

2.4.2 差分隐私

为解决当前信息越来越发达的社会所带来的用户隐私泄露问题，差分隐私模型[28]是一种被广泛认可的严格的隐私保护模型。它通过对数据添加干扰噪声的方式保护所发布数据中潜在的用户隐私信息，从而达到即便攻击者已经掌握了除某一条信息以外的其他信息，该攻击者仍然无法推测出这条信息。因此，这是一种从数据源头彻底切除隐私信息泄露可能性的方法。然而，该模型的最基础来源是基于抽象的数学理论，使得该领域的入门需要较高的门槛。

数据集D 发布时，通过删除标识符属性（例如姓名、ID号等）能够在一定程度上保护个人隐私，但这远远不够，要特别注意到，数据集中还有其他属性，例如：生日、性别、居住地、是否抽烟、是否饮酒等，用这些信息来猜测个人身份，是不是类似数据挖掘中的分类？即将一个没有类别标签的条目识别归类，训练集看成敌手A 从其它地方获得的具有标识符属性的与 D 有属性交集的数据集T，类别标签当然就是某个人。在这方面，美国曾经有几个经典案例，被起诉方通常都是赔了几百万。

在隐私保护方面，k-anonymity 及其扩展模型影响深远且被广泛研究，但随着研究的深入，该系列模型也面临着许多新型攻击的挑战，于是2006年就出现了DP 。DP 在攻击者A 在拥有最大背景知识条件下，仍能抵御各种攻击。

差分隐私有以下几点几本概念[29]：

1. 差分隐私（ϵ-DP）：通常来说[30]，差分隐私机制的实现是通过对数据集添加拉普拉斯噪声或者是指数机制来实现的，其核心构想就是对于两个数据集 ，两者假设只有一定数量的数据差异，此时通过一个满足差分隐私（ϵ-DP）查询算法A，去获取具体的条目，其应满足：
2. 全局函数敏感度[29]：任意函数,其敏感度为：

D1与D2最多相差一条数据记录，函数的敏感度越低，差分隐私机制能保留越多的信息，精确度越高，输出是否时实数可将函数f分为实数型与非实数型，分别对应拉普拉斯机制与指数机制。

1. 拉普拉斯机制[30]：对于实数型函数，可以对应有三个输入，数据库D，函数f，隐私参数e，添加的噪音服从概率密度函数为

的拉普拉斯分布，其中，即噪音的量与数据库无关，只与函数的敏感度与隐私参数有关。

1. 指数机制[31]：对于非实数型函数，以概率p从结果集R中选择r作为输出，其中 为结果r在数据集D上的效用性值，q的敏感度为

其中

本文中通过对影像中人脸类型的统计数据分布进行差分隐私的指数机制处理后再发布。

2.5小结

图像相似度是一个很大的学术问题，关于此问题的学术讨论也曾出不穷，尤其是针对各种复杂的光线环境，角度变换等情况。本文讨论的测试样本均建立与ID照片等，故为了算法的效率选择了直方图比对，同时为了算法的严谨性将四种比较标准交叉检验。在本节对图像，直方图算法，对比标准进行了简单的阐述，为后面在交叉测试中参数的选择进行铺垫。

数据的隐私发布也是一个十分火热的话题，K匿名及差分隐私是目前最适用的数据发布隐私模型，但同时两者的应用效果也略有差异，在不同问题的优缺性也存在差异，本文在设计保护模型中同时兼顾这两种保护模型，故在本节对以上所需概念进行了阐述，为后面算法的扩展及噪音的添加说明打下基础。

第3章 具体实现方案

3.1方案概览

本章将介绍基于K匿名及差分隐私的定域功能性影像隐私保护系统及其具体的实现细节。由于本文的影像隐私保护系统的实现主要分成四部分，所以下面对于具体算法的介绍也分成四部分进行描述。

首先我们对初始的从影像中提取的标准人脸数据集进行处理，我们根据不同的风险要求以及考虑数据可用性，如在本文中的设定是我们在已知某个小区发生了偷窃事件，因此我们需要将监控中出现的人脸资料划分为是本小区居民以及外来者。划分的规则是可以多变的，需要根据不同的资料来源以及需求进行变更，具体的细节会在后面一部分细致阐述。对划分后的数据我们进行标准化裁剪后保存进预处理完成数据集，统计人脸类型的统计分布数据。

对于已经提取完成的数据集，接下来在第二步中我们将其直方图化，接下来使用在绪论部分介绍过的四种直方图比对标准进行数据集内比对排序，优选出最相似的前K个个体，最后系统的出的结果需要将不同的标准交叉使用来提高系统可信度。

做好准备后，我们根据提取出的前K个最相似个体进行线性合成，选择规定好的权值线性合成，合成后的个体重现保存进完成匿名化数据集。对于之前保存的统计量，通关差分隐私进行指数分布处理然后保存心得分布。

最后对匿名化数据集中的数据，重新进行相似度比对，然后对于对比结果分析其K匿名性以及对差分隐私模型的满足能力。训练后的结果反馈进权值选择与噪音控制。

通过以上的简单描述，我们知道本文关于基于K匿名及差分隐私的定域功能性影像隐私保护系统的主要工作有：

初始化数据（划定分类 → 数据标准化）→ 主体直方图化及相似度对比排序 → 加权线性合成，噪音处理 → 新主体安全性验证及反馈参数调整

3.2数据预处理

在本文中需要因为涉及到复杂的隐私泄漏问题还有数据可利用问题，对数据的预先划分是必须的，既可以保证数据的高效可用性也可以预先对差分隐私保护提供支持。同样在本系统中，进行大幅度的图像处理工作，因此对于原始数据的统一标准化是必须的。接下来本节将从这两个方面进行详细的讨论。

3.2.1按规则划分数据集

基于本文的主要背景，我们设定背景为在一个小区内发生了盗窃案件，警方怀疑是内部人员所为，且拥有一个月内小区进出口，楼群间的监控视频，警方想要借助多方研究机构给予支持分析小区中用户，及外来人员的行为习惯。但是想要通过本系统进行影像隐私保护，使得不会在发布影像的过程中出现个人隐私泄漏的问题。基于这样的背景，我们考虑可以把所有在影像中出现的ID划分成两类，一种是属于本小区的居民，另一类是不属于本小区的居民。这样划分有以下好处。首先是对主题群进行了大类的划分，使得在进行了加注噪音后仍然可以较轻松的进行划分，使得不会对之后的影像数据分析造成困难。另外这样的大类划分有效保证了划分群体内部的匿名性与统一性，过细的划分会给保障差分隐私带来极大的困难。

对于使用本模型的其它背景，需要根据具体的需要进行划分，建议在保证K匿名的基础上，每个数据集内的数据要大于等于20\*K个，同时划分时需要考虑到大环境的隐秘性进一步对差分隐私的安全进行保护，例如在本文背景中就不能够划分成每栋楼一个数据集。因为如果这样，旁观攻击者可以针对每栋楼的用户进行分析，某用户进入楼后不久，某住户的灯光开启了，这样就很容易锁定到个体。因此在大类划分中对于这个要求也需要考虑其中。

3.2.2人脸识别提取标准化

对于每一个划分好的数据集，我们首先使用opencv的人脸探测算法（face\_cascade.detectMultiScale）求的每个个体的人脸的位置坐标及轮廓范围，之后按照标准处理图要求以人脸中心为坐标提取出固定大小的图片并保存，这里我们使用的初识主体图像标准为：

Dimensions：144 × 192 pixels

Resolution：300 pixels/inch

Color Space：RGB

而对于我们提取出的照片，我们的图像标准是：

Dimensions：75 × 75 pixels

Resolution：72 pixels/inch

Color Space：RGB

我们将所有提取出的照片按原始的主体名称保存在新的文件夹中，等待之后的处理。整个该部分的算法如下描述：

算法3.1 Mat detectandgive( string path)

输入：个体原始图像的路径

输出：检测出人脸标准化提取后的图片生成的Mat矩阵

1. 将图片数据通过imread方式读入，

if(读入成功)

原始图像灰度化; //为之后的人脸检测做基础

1. face\_cascade.detectMultiScale（）

初始化faces数组保存探测结果

探测图像中出现的人脸，其中心位置的坐标（x,y），其轮廓的大小（weight，length），将每一个人脸的信息保存进faces数组。

//由于现在的标准化图像中只有一个人脸，故实际上数组中只有一个元素

1. for(faces中每个人脸单元)

初始化空Mat矩阵roi\_img//准备保存探测的人脸图像

从原始图像中以：

长：Range(faces[i].y +5,faces[i].y+80)

宽：Range(faces[i].x +5,faces[i].x+80)

截取图像保存进roi\_img中。

return 每一个roi\_img

步骤2中是调用了opencv的人脸检测算法实例，是除了直方图比对及图像合成外最耗时的步骤。

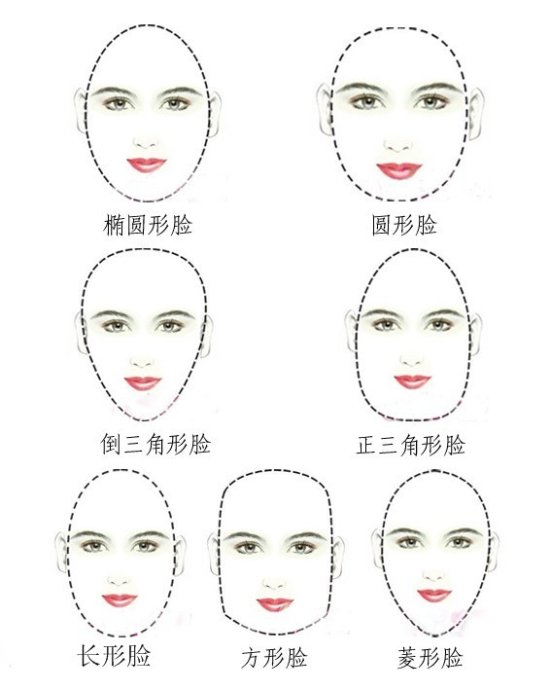
而在步骤3中，对于每个主体实际上在步骤2中已经保存了其面部轮廓的范围，然而为了后续合成的方便，我们将基础范围扩大到提取图像标准。总的来说，这部分的算法整体时间复杂度是O(1)。其生成图像如下图



**图3-1 标准化提取图样例**

3.2.3人脸类型提取统计

本文借助差分隐私模型中从人脸类型的角度进行保护。故我们首先对人脸类型进行划分[32]，类型如下图：



**图3-1 标准化提取图样例**

椭圆形脸，也就是蛋型脸。蛋型脸的额头与颧骨部位基本一样宽，比下颌稍宽一些，脸的宽度约大概是脸的长度的2/3。

长型脸的特征是脸型比较瘦长，额头、颧骨、下颌部位基本齐宽，比较显眼的是长形脸脸宽小于脸长的2/3。

倒三角型脸俗称心形脸，最明显的特征便是额头处最宽，额头往下均匀的过渡到尖下巴。

方型脸的特征是额头、颧骨、下颌的宽度基本相同，感觉四四方方的，是各种脸型中最容易辨认的。

圆型脸的特征和方型脸一样，都是额头、颧骨、下颌的宽度基本一样，最大的区别就是，圆型脸在脸型的转角处显得比较圆润丰满，有时候被称为娃娃脸，不像方形脸那样棱角分明。

正三角形脸也被称为洋梨型脸，这种脸型的特征是额头较窄，下颌最宽，从正面看去，像极了一个正三角形。

菱形脸俗称钻石脸，菱形脸的特征是颧骨部位最宽，额头和下颌逐渐变窄。所以菱形脸看上去会显得比较尖锐。

我们在提取的标准图上依据以上分类规则对数据集进行统计，并将统计后结果保存，准备在后面的通过指数机制处理。

3.3 相似度对比

经过上面的处理之后，此时我们已经得到了要进行比较操作的原始数据集，按照算法需求我们需要对人脸进行对比与选取。本文中选择使用灰度直方图进行比对，因此接下来将会对灰度直方图化以及比对算法进行介绍。

3.3.1 人脸图像灰度直方图化

在第二章中我们已经对灰度直方图化有了简单的介绍，其公式及算法也已经在上文介绍了，这里就不再赘述，在opencv中主要计算直方图的算法是cvCreateHist()。直接对我们设计的图像转换直方图算法进行描述。该部分的具体算法描述如算法3.2所示。

算法3.2 IplImage\* myhist(string Path)

输入：转换图像路径

输出：IplImage图像指针

1. 读入图像并进行灰度化，初始化以下变量：

直方图尺寸，灰度级的范围hist\_height ＝ 256，

创建一维直方图，统计图像在[0 255]像素的均匀分布

然后使用函数cvCreateHist()计算灰度图像的一维直方图

1. 绘制灰度直方图：

创建一张一维直方图的“图”，横坐标为灰度级，纵坐标为像素个数

统计直方图中的最大直方块max\_value

分别将每个直方块的值绘制到图中

for（hist大小）

像素i的概率作为基准为bin\_val，

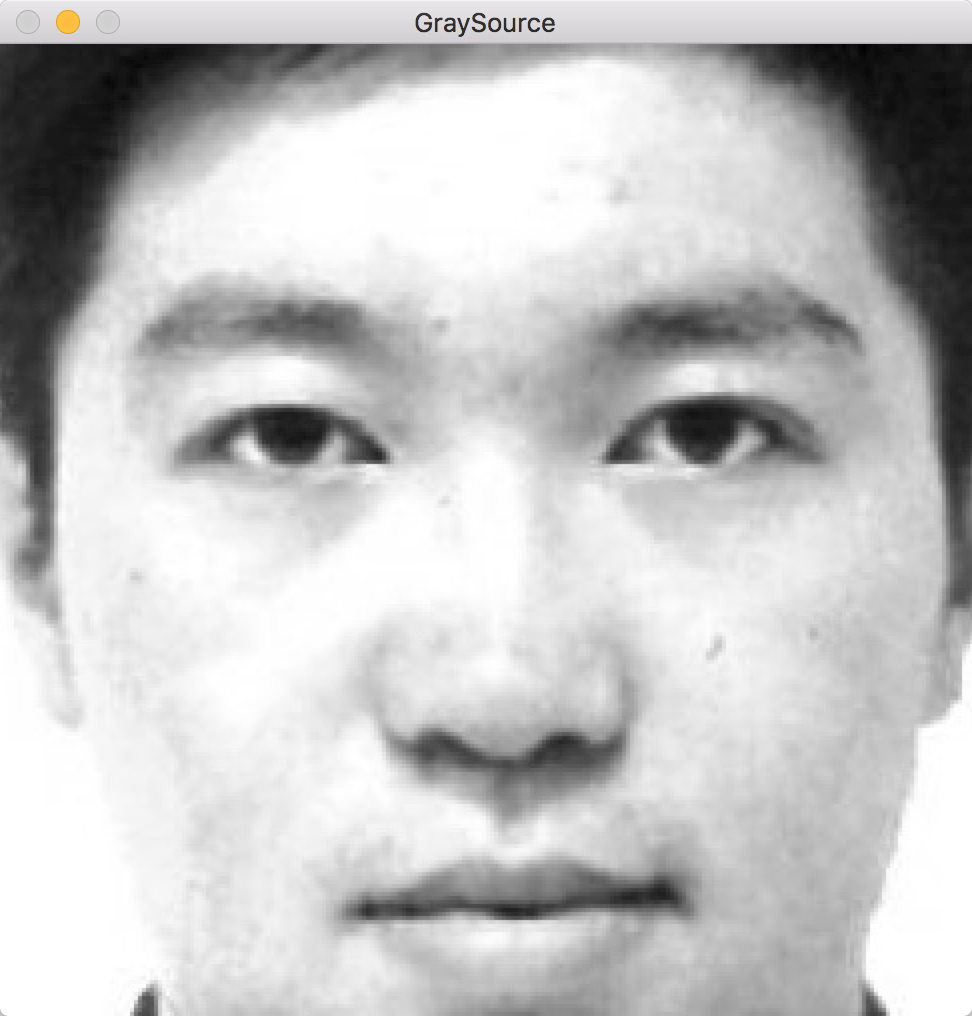
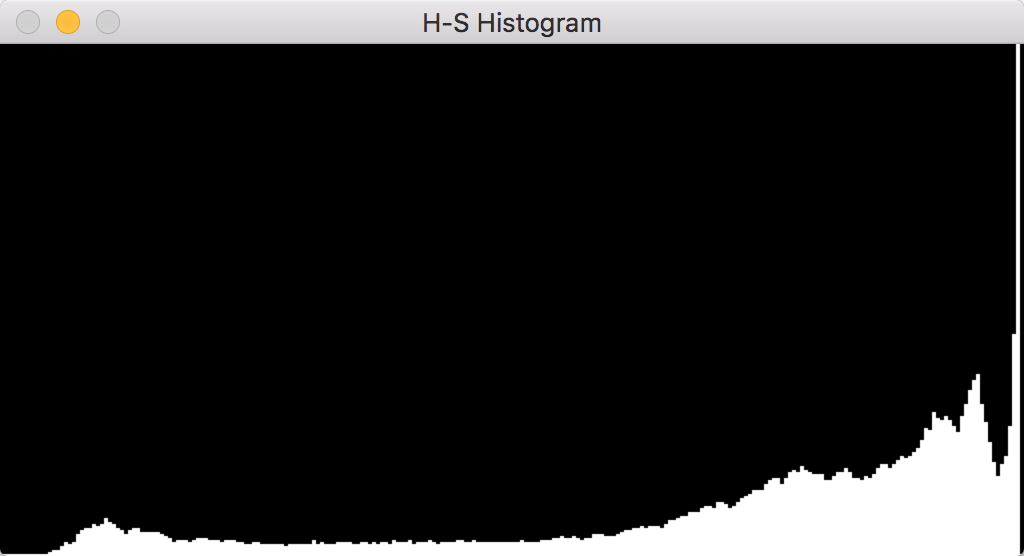
将bin\_val\*hist\_height/max\_value作为要绘制的高度绘制在图中，并输出

1. 返回图的指针

步骤1中需要对图片进行灰度化为了之后的比对我们需要使用灰度直进行比较。步骤2中的输出是为了检查直方图转化是否成功，在样本进行系统测试时不必显示输出。

以上3个步骤进行一次遍历，这部分的算法整体时间复杂度也是O(M)，是线性的，其中M是hist大小。

整个步骤提取和灰度直方图化的样例如下：

**图3-3 灰度转换及灰度直方图化**

3.3.2 直方图交叉对比

在获得每个对象主体转化后的灰度直方图后，我们需要使用直方图对比算法获得每个主体与之相似排序。计算相似度的算法原理已经在第二章进行了介绍，我们在系统中需要进行交叉对比，即使用A算法进行比对后，生成的图像不止要用A算法进行检验，同时还要使用其他三种算法进行交叉比对，我们在本环节就先以Chi-square test方检验为例，并将结果反馈。该部分的具体算法描述如算法3.3所示。

算法3.3 double CompareHist\_CV\_COMP\_CHISQR(IplImage\* image1, IplImage\* image2)

输入：主体图像，客体图像

输出：相似度度量数值

1. 调用算法3.2，获得两幅图像的灰度直方图

主体的灰度直方图为：Histogram1；

客体的灰度直方图为：Histogram2

1. 调用opencv灰度直方图比对算法：

cvCompareHist(Histogram1, Histogram2, CV\_COMP\_CHISQR);

计算结果就是两个图像的相似度度量值result

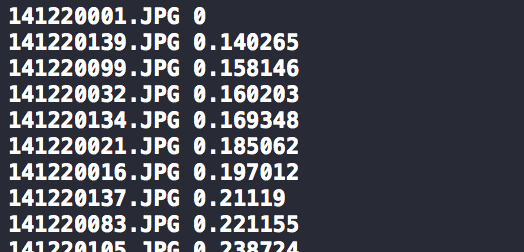
回收垃圾

1. 返回result

步骤1中调用了算法3.2的简化版本，屏蔽了其输出部分。

步骤2中的输出是方检验的数值，由第二章中介绍可知，在数值为0时代表两幅图最相似，数字越大代表差异越大。整个算法内嵌与算法3.4中使用。

以上3个步骤是线性的，这部分的算法整体时间复杂度也是O(1)。其计算结果样例（排序后的）如图3－3，第一列代表主体编号，第二列代表相似度度量值。从图中可以清晰看出自己与自己对比相似度量值为0，代表最相似。



**图3-4 直方图对比结果样例**

3.3.3 提取最相似的前K个主体

如中所述，直接给图片抹黑，加马赛克等暴力的方式并不能兼顾影像资料的可用性及隐私性质。本文考虑兼顾K匿名安全模型及差分隐私模型的要求，同时为了保障新生成个体与其它个体的可区分性，我们选择主体A，并将一个分组中的其它图像与之相比较，并选择前K个与之最相似的准备在后续的算法中加权线性叠加成新的主体。故本节算法3.4如下：

算法3.4 bool get\_10\_top\_path()

输入：主体图像池结构（包含地址等信息）

输出：前K个最相似的图像，相似度量值，地址->文件

1. 打开文体准备写入（覆盖式）
2. for（池中大小，指标i）

｛ 初始化每一项新路径；

初始化新的对比结果池；

for（池中大小，指标j）

｛ 打开i，j指向的文件，将i作为主体，j作为客体；

调用算法3.3获取对比结果，

保存进对比结果池（主体，课题，路径，数值）；

｝

使用冒泡排序将对比结果池排序；

循环将对比结果池中（对比结果，路径，名称）前K个写入文 件

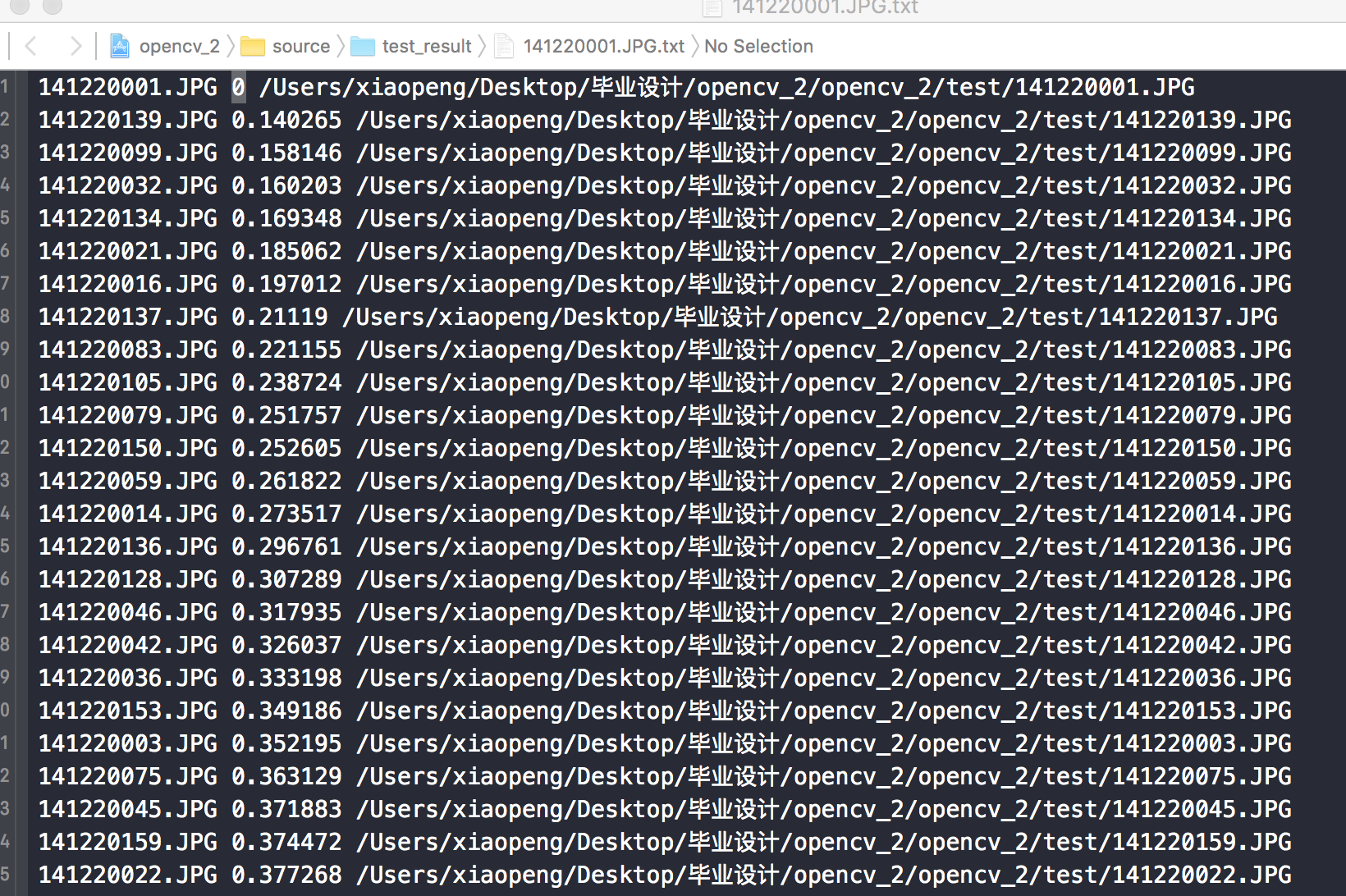
｝

1. 关闭文件，若以上运行无异常，返回true，否则抛出异常返回false

步骤2中调用了算法3.3的简化版本，屏蔽了其输出部分。

步骤2中在对比时使用的比较算法标准是可更换的，在交叉对比调整参数环节需要变更不同的算法标准。这部分主要耗时的是直方图对比算法调用及文件读写。

这部分的算法整体时间复杂度也是O()。其计算结果样例（排序后的）如图3－4，第一列代表主体编号，第二列代表相似度度量值，第三列代表路径。文件名代表此次比较主体编号。目前K取值为25，当然可以根据不同的主体数量及安全需求进行变更。



**图3-5 前K相似提取结果样例**

3.4 统计量差分隐私保护与图像加权线性叠加

在文章[1]中，作者讨论对比了线性叠加与像素点聚类两种方式对原始图像可用性破坏的程度以及对于K－different性质的满足情况，结论是线性叠加相较于像素点聚类拥有更好的运行效率及保护效果，因此考虑到我们的应用场景，我们对于提取的原始主体也通过线性叠加进行保护。但是作者[1]仅仅分析了两个个体的情况，考虑到个体与个体之间需要明显的可辨别性，同时保障K匿名及差分隐私要求，不能仅仅选择简单的均等线性叠加，需要添加一定的干扰噪音，同时对于选择的个体在合成时需要进行权值分配，其实际操作及原因会在下面具体阐述。

加权线性叠加是整个算法系统的核心步骤，是最后生成保护后图像的生成步骤，在重复调整参数后其输出结果就是最终的保护后图像主体。

3.4.1 统计量差分隐私保护

首先我们设定查询函数，即查询者有能力多次获得数据集中主体用户的人脸类型分布。则我们设置查询分布函数：根据输入的数据集返回脸型分布图，以个体的脸型数量总和作为分布单元。根据全局敏感度的定义，分布函数的敏感度为1，即,实用函数返回值就是各脸型的统计个数。

对于的选择需要根据标准要求，即对真实性影像程度标准进行反馈调节，在经过多次反馈调整后，反馈过程将在检测部分以及试验部分细说。本文使用作为标准。算法关键就是2.4.2.4中的公式表达。

故本节算法3.5如下：

算法3.5 bool result\_after\_DP(int \*valure, int sum，double E)

输入：每种脸型的统计个数以及脸型种类数

输出：经过指数机制保护后的各种类占比

1. 新建double result［sum］；

for（sum）//依据指数机制循环计算

｛

result［循环变量i］ ＝ exp（E\*value[循环便利i]／2）；

｝

1. 计算result数组的总和保存到double sum；

for（sum）

｛

result［循环变量i］ ＝ result［循环变量i］／sum；

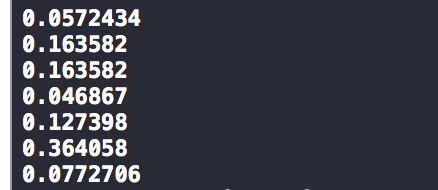
｝

1. 输出result的结果

步骤1中实现了指数机制的功能，以每种类别的个体数量作为可用性函数以及E为作为参数进行计算。

步骤2中将第一步中的运算结果进行标准化，将其概率正比于计算结果，再保存输出即可。

该部分的结果需要经过差分隐私部分检查经过多次反馈调节后确定参数后输出。输出样例如下图3-6：



**图3-6 指数机制后的统计分布**

3.4.2 图像处理权值选择

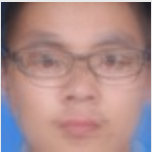
权值选择是图像加权线性叠加的重要一步，是不出现相同合成图像的重要保证。对于在上文中讨论平均线性叠加，我们可以考虑一种情况，我们选择主体A，经过算法3.1,3.2,3.3后我们假设与其最相似的图像是B，C，D..。经过试验在数据集大小在167的情况下，有60%的概率出现对于主体B，其最相似的图像排序会A，C，D..呢么对于保护后图像:

可以看出，如果K值选择较小，且是平均线性叠加，在很大概率上A’与B’的区分度会非常低，这与本系统的目的显然背道而驰，因此我们需要对合成的元素加上权值，使得对任意一个主体X，起最相似的元素为X1，X2..，每一项的权值为k1，k2..，则合成公式为：

对于调整权值，我们使用反馈机制，即首先设定k ＝ k1 ＝ k2 ＝ ..＝ 1/K；合成结束后，通过3.5.1的K匿名检测反馈K匿名满足性来修改k值，最重在本文使用的当前数据集下我们规定以下k值：

个

不同的权值选择合成后会带来明显的差异，以下是两幅在同一主体下不同权值选择情况下获得结果：

**图3-7 k恒等于0.7 图3-8 k为上述标准**

3.4.3 叠加生成新主体

在之前权值等都选择结束后，就开始进行合成，这里我使用了像素点控制进行合成，之前在第二章中已经对RGB的三通道图像进行了简单的阐述，合成算发也在第二章进行了过程化介绍，对于像素点操作需要分不同通道进行执行。

故本节算法3.6如下：

算法3.6 bool get\_color\_new\_picture(string Path)

输入：每个主体前K个相似客体排序文件地址

输出：加权线性合成的图像

1. 打开输入地址文件，读入每一个客体图片并根据算法3.1完成标准化后保存进客体池结构
2. 新建初始化合成后图像容器pdst

for（图像行结构）

｛ for（图像列结构）

｛ pdst.B ＝ 池结构中各个客体的B通道数值乘以各自k相加；// pdst.B是pdst的蓝通道

pdst.G ＝ 池结构中各个客体的G通道数值乘以各自k相加；// pdst.G是pdst的绿通道

dst.R ＝ 池结构中各个客体的R通道数值乘以各自k相加；

// pdst.R是pdst的红通道

if (pdst.B>255)

B = 255;

if (pdst.G>255)

G = 255;

if (pdst.R>255)

R = 255;

｝

｝

1. 将pdst保存到test\_new目录，以原始的主体标识符命名，若以上运行无异常，返回true，否则抛出异常返回false

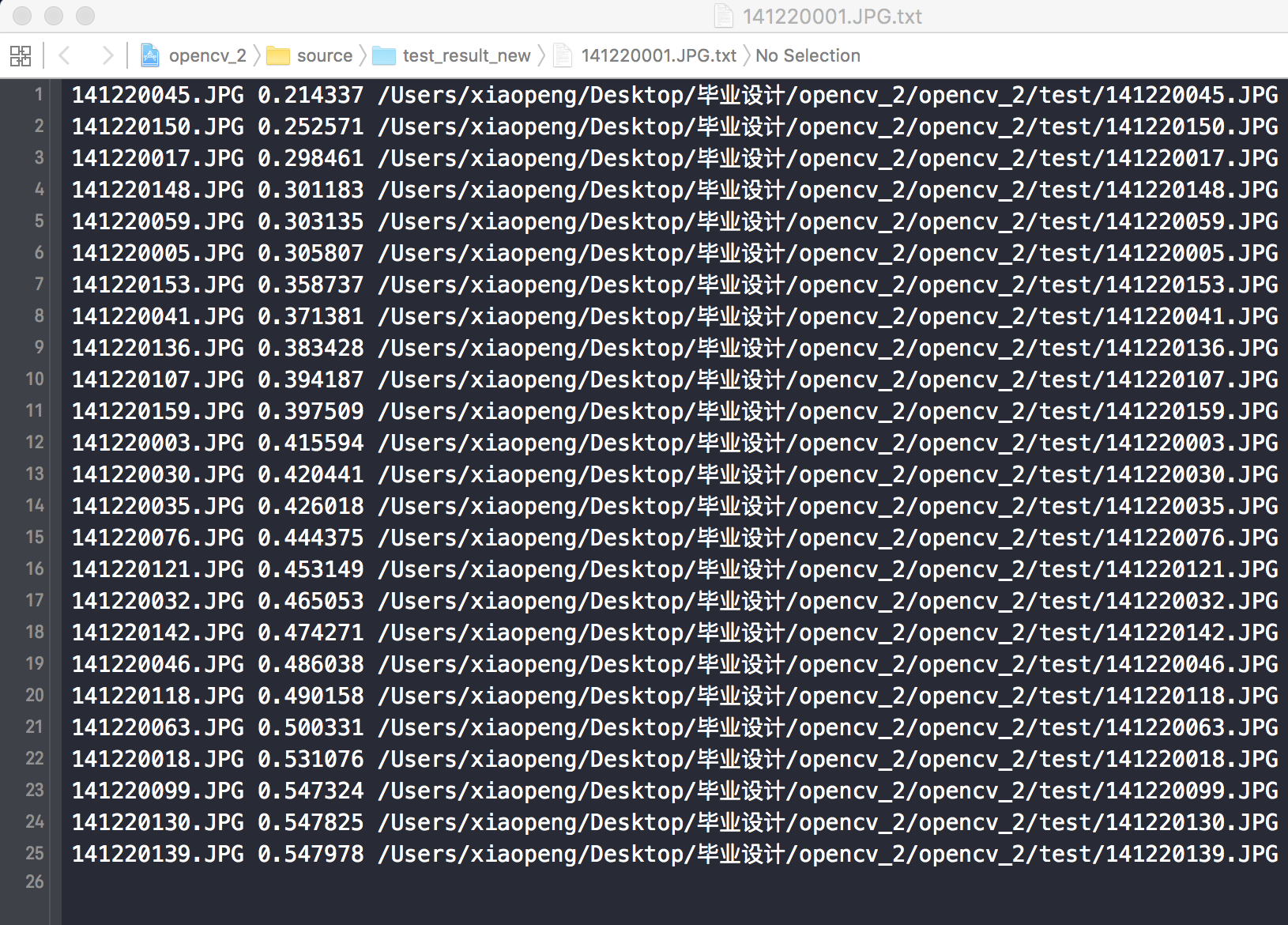
步骤1中调用了算法3.1的简化版本，屏蔽了其输出部分，其算法结果直接保存进池中供后续计算使用。

这部分的算法整体时间复杂度是O(M\*N\*K)。其中M，N代表图像的行像素点数和列像素点数，K为选择的前K个相似数。

该部分生成的的结果以原始的主体标识符命名，在参数调整完毕后投放进测试，其合成样例为图3-8。

3.5 生成样本测试

对于新生成的经过处理的主体数据集，我们再次使用算法3-3，3-4进行比对，只不过这一次比较的对象主体是经过处理的单位，客体是原始的数据集，我们需要通过新的比较数据去度量我们的权值选择及噪音添加是否能够满足我们的预期，其结果样例为图3-9，其中文件名是主体名称，第一列是前K最相似客体排名，第二列是主体对于每个客体的相似度度量数值，第三列是每个客体的保存位置。



**图3-9 保护后主体与原始图像进行对比后结果**

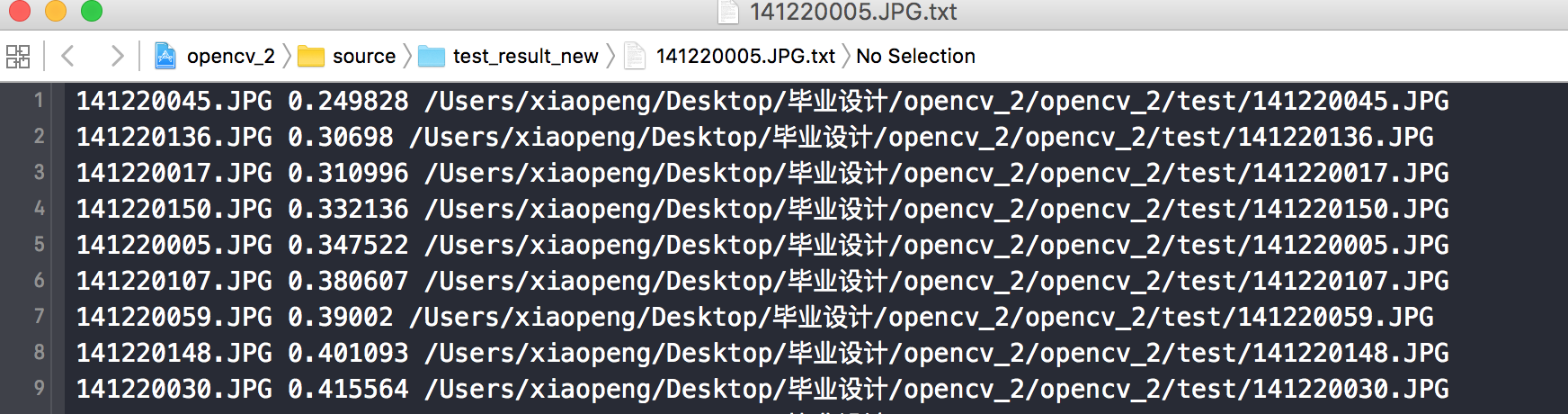
在获得新的对比文件后，我们需要对其结果进行K匿名性及差分隐私检测以对权值及噪音进行调整。其具体方法会在后续小节中叙述。

3.5.1 K匿名检测

为了检测生成的保护后图像是否满足K匿名要求[18]，我们需要做以下规定。基于K匿名的要求，攻击者无法以大于1/K的概率确定对应的主体单位，则我们排除攻击者获取其他情报的可能，单纯就人脸识别角度来看，新生成的主体与原始主体的相似度在所有排名中不能高于：

这里M为一个数据集中所有主体个数，对于我们测试集M为167，K为25，则L为6.68取整后为6，换句话说我们经过3.6生成的新对比文件中如图3-9，对于新主体A，旧主体A不能出现在排序的前6位，我们遍历新的对比数据集如果前6位没有出现则该主体的保护为‘True’，否则为‘False’我们遍历所有的数据集若‘false’占比超过30%则反馈之前权值选择层，降低主体的权值占比。则我们设定K匿名的成功率为：

经过多次优化，就依上文的权值选择，在当前的测试集中，成功率Kp为97.2%，其中一个不满足样例如图3-10:



**图3-10 K匿名检测false样例**

这个是对于新主体005号，可以发现其对比结果中005号排名第5，因此这是一个false的样例。就当前的数据集，通过102次参数调整，当前权值下K匿名的Kp最高。

这部分具体的算法描述如下：

算法3.7 double K\_check(string Path，int comp\_numb)

输入：每个主体前K个相似客体排序文件地址

输出：Kp

1. 打开输入地址文件，将其中文件地址加入池结构

True\_numb = 0;

False\_numb = 0;

1. for（池大小）

｛

读取文件，获得其前comp\_numb条主体标识符；

Find ＝ false；

for（comp\_numb）

｛ 对比条目与文件名代表的主体标识符是否一致；

Find ＝ true；

｝

if（Find）

False\_numb++；

Else

True\_numb ++;

｝

1. 返回True\_numb／（True\_numb＋False\_numb）

该部分算法复杂度为O(M\*comp\_numb)，其中M为主体个数，comp\_numb为L；

3.5.2 差分隐私测试

为了检测与度量经过指数机制保护后的统计数据与真实值之间有没有出现较大误差使得数据可用性变差，需要给出标准，我们设定真实值与保护后的值之间的距离为：

经过多次反复测试，发现为了保证我们对任意类型去掉1个个体时，新的反馈到原始个体数量时结果差异小于1，并且要求单项差异不能达到一倍，因此我们设定，实验表明这个范围内可以满足我们的需求并且经过训练集测试最终确认作为标准，拿测试集来看，表格3-1是计算样例，通过算法3.5获得。

**表3-1 数据集样例**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | 椭圆形脸 | 圆形脸 | 倒三角脸 | 正三角形脸 | 长形脸 | 方形脸 | 菱形脸 |
| 真实值 | *5.98802%* | *18.5629%* | *18.5629%* | *3.59281%* | *15.5689%* | *28.1437%* | *9.58084%* |
| 保护值 | *7.88913%* | *16.4526％* | *16.4526%* | *6.85848%* | *13.8113%* | *28.8032%* | *9.73265%* |

由上表我们计算得到0.0026732,较好的满足的我们要求，并且经过验证，每次尝试在某类脸型中删去1个个体，获得新保护值D与前一次特定类型的旧值相比，反馈到总体上看不会大于1，故满足要求。

这部分具体的算法描述如下：

算法3.8 bool D\_check(double \*R，double \*D,int sum)

输入： 真实值R，保护后值D，种类数量sum

输出：

1. 初始化double L；

for（sum）

｛

L ＋＝ pow（（D［i］－R［i］），2）；

｝

1. 如果0.002 >L > 0.005 z则反悔true，否则返回false

该部分算法复杂度为O(sum)，其中sum为种类个数。

3.6算法实现总体流程图

在经过上面的介绍后，本文就定域性影像隐私保护算法的总体实现步骤就如图3-11所示：

**图3-11 保护算法流程图**

读入所有数据集

Start

按照隐私保护要求划分数据集

调用detectandgive( string path)

进行标准化人脸提取

调用get\_10\_top\_path()算法（其中包括直方图变换算法及比对算法）获取前K个最相似个体

参数选择及调用bool result\_after\_DP(int \*valure, int sum，double E)

调用get\_color\_new\_picture(string Path)算法进行合成

End

调用K\_check(string Path，int comp\_numb)

D\_check(double \*R，double \*D,int sum)检查，结果若满足输出结果完成，否则反馈调节参数

3.7数据发布模型流程图

对于系统算法的部分上文中已经进行了较为详尽的阐述，最后关于本算法的实际应用环境以及数据发布模型流程如图3-12所示：

**图3-12 数据发布模型流程图**

用户获得数据，并根据不同需求再次向数据持有者申请

发布需求方要求发布的统计数据

发布经过保护的影像资料

经过DP，K  
匿名检查后准备发布

需求方预发布数据提出保护需求

系统获取原始数据文件进行运算

第4章 实验

4.1实验环境与数据集

本次实验中的计算环境如下所示：

硬件环境：

CPU: Intel(R) Core(TM) i5-3537U CPU@2.GGHz

内存空间: 8GB

硬盘: 500GB

软件环境：

操作系统：OSX 10.11.3

开发平台：Xcode Opencv

开发语言：C++

本次实验中所使用的数据集由互联网，高校数据库提供，数据分两组，共计

318个，全部符合主体图像标准。以下是具体的数据情况。

**表4-1 数据集情况**

|  |  |  |
| --- | --- | --- |
| （序号）数据组号 | 数据集大小（个） | 男主体／女主体 |
| （1）A组（训练集） | 167 | 139/28 |
| （2）B组（测试集） | 151 | 120/31 |

4.2算法精度

本次实验在评估算法结果时采用了同3.5.1章节的评估要求，新生成的主体与原始主体的相似度在所有排名中不能高于：

我们遍历所有的数据集若原始主体在整体排名小于A的占比超过30%则反馈之前权值选择层，降低主体的权值占比，同时我们设原始主体在整体排名大于A的样例总数为。则我们设定K匿名的成功率为：

对于表4-1中所示的数据集的实验测试结果如表4-2所示（表中的斜率误差衡量的标准值是手动测量，存在少许误差），在A组经过10轮反馈调整后采用本算法的聚类精度结果为97.2%，B组经过系统运行后达92.3%，说明该算法的保护效果良好。每一轮的计算结果也如表4-3所示。

**表4-2 实验结果**

|  |  |  |
| --- | --- | --- |
| （序号）数据组号 | 数据集大小（个） | 值 |
| （1）A组（训练集） | 167 | 97.2%（10轮反馈后） |
| （2）B组（测试集） | 151 | 92.3% |

**表4-3 反馈调整过程**

对于差分隐私要求，本实验采用了3.5.2章节中的评估要求，即要求真实值与保护后的值之间的距离：

我们在测试过程中发现当时，的偏差已经过大，故本实验针对进行分解寻求最优的的取值，我将这个区间进行了10等分进行测试，最后确定时训练集达到要求且达到极小值说明能够满足要求，数据集，及最终结果如表4-4，训练过程如表4-5：

**表4-4 Dp实验结果**

|  |  |  |
| --- | --- | --- |
| （序号）数据组号 | 数据集大小（个） | 值（ |
| （1）A组（训练集） | 167 | 0.0026732 |
| （2）B组（测试集） | 151 | 0.00384079 |

**表4-5 Dp分割训练过程**

4.3算法时间效率

实验中的各个数据集进行测试该算法的所需时间如表4-6所示，每个数据集的大小可以参见表4-1。

**表4-6算法运行时间**

|  |  |  |
| --- | --- | --- |
| 数据集 | A组 | B组 |
| 时间(min) | 12.3 | 11.3 |

本算法的大部分算法都是低于n方的，主要耗时点在于两次各主体间的相似度比较，目前实验阶段算法使用单线程结构，效率没有达到最优，之后对于较大的数据集可以考虑加入如hadoop，spark等大数据的处理方式，分布式的进行比较汇总可以较快的得到算法结果。

4.4实验结论

以上的实验结果表明，通过K匿名测试，我们可以获得一份有保证的保护后影像资料以及获得对应的经过差分隐私保护却不会严重失真的统计数据。这样的保护方式和发布方式可以提供给用户一种可靠的影像数据发布方式以及对于使用者而言最大化保证了数据的可用性。通过实验，我们就训练集获得了极优的参数进行了调整并对应的在测试集上进行了测试，获得了期望的保护效果。

综上验证了文中提出的基于K匿名以及差分隐私的定域功能性影像资料保护算法可以在一个较快的时间内对影像资料进行隐私化处理并准备好发布数据，提供给用户一个安全的，有效的影像数据结果。

第5章 总结与展望

5.1本文工作总结

影相资料中携带的隐私信息种类繁多，在资料保存，传递中泄漏的风险，途径也是防不胜防。在各种隐私信息中又数主体个人信息即人脸信息最令人着重注意，在各种类型的数据交换中，尤其是资料持有者与使用者之间的交换中其泄漏的风险总会变成各种阻碍。本文研究了之前学者就相关问题的讨论与工作结果的优势与缺略，结合K匿名模型及差分隐私模型，既考量隐私信息的安全性同时又在最大程度保护数据的可用性，建立一套带有反馈调节机制的，在确定区域使用的，进行功能性划分的保护，发布系统。并通过试验证明了系统可用性与成功率。

本文的主要工作内容有以下几点：

1、调查并研究了几种实用度最高的几种图像隐私保护系统，由于本文关注的重点是人脸隐私，且初衷在于既保护隐私，又要保证数据的可用性，故不能纯粹使用抹黑或者马赛克，因此参考了K匿名数据，差分隐私模型的保护方法，尽可能在保障数据可用性的情况下保障数据隐私。

2、介绍了多方式交叉比较模型的基本思想和标准，通过不同标准的交叉比较保障了算法成果的真实有效性，并提高了训练后模型的可靠性。

3、关于算法的参数设置问题，本文利用检测反馈模型进行设置，既通过每次结果对预期的符合度反馈调节参数配比，及噪音添加模式来逐步优化算法结果，保障算法的致优性。

4、为了提高算法的时间效率与操作性能，一方面设计了过程式算法结构，一方面在功能过程中省略繁杂的重复的过程，比如在运行时考虑将图像信息结构化统一池结构保存，每次取用时无需重复读写。

5、详细介绍了文中关于基于K匿名及差分隐私的定域功能性影像隐私保护算法的实现，并通过实验验证了该算法的可行性和可靠性。最后通过流程图形式详细展示了整个算法模型的实现流程与数据的处理，发布模型。

5.2本文的不足与展望

本文的工作在取得一定成果的同时依然有许多方面需要得到进一步的改善：

1、文中的算法在进行优化时需要根据反馈调整参数，目前整个系统的参数调整仍然是依靠人工赋予，下一步我们在继续优化系统的过程中希望能够设计新的算法进行动态参数调整，让系统最后直接送出最优结果即可，不再需要大量人工。

2、虽然算法中对比较操作，合成操作的IO操作进行了优化，但是大量时间仍然耗费在其中，尤其当数据集很庞大时，算法复杂度为O（N2），但其中比较操作实际上是可以分布式操作的，下一步我们希望可以借助大数据分析算法hadoop，spark等，在大规模图像处理时可以分布式多线程的完成算法结构。，，

3、在实验验证中发现文中算法虽然对大部分实验数据处理结果较好且能保留大部分信息点，但是对于某类型个体信息如行吧等会有相当大的模糊化，我们考虑在下一步可以将这类要求加入算法第一步的功能性分类中以达到需要的结果。

4、算法一开始的定域功能性数据划分预处理的最优方案获得没有较多自动化处理机制以及人脸划分部分没有进行自动化，在大规模数据集下需要继续进行分类部分的架设，大部分要求还是需要用户在使用中提出具体要求然后自己进行分类，下一步我们希望结合[1]中的De-identification概念及机器学习的思路寻找自动化途径将数据进行初期划分以节省开销。

参考文献

[1] Elaine Newton，Latanya Sweeney，Bradley Malin:Preserving Privacy by De-identifying Facial Images. [C]March 2003 CMU-CS-03-119 School of Computer Science Carnegie Mellon University Pittsburgh, PA 15213-3890

[2]Suriyon TANSURIYAVONG，Shin-ichi HANAKI:Privacy protection by concealing persons in circumstantial video image. [C][Workshop on Perceptive User Interfaces](http://xueshu.baidu.com/usercenter/data/journal?cmd=jump&wd=confuri%3A%28ec1cde5a645e01db%29%20Workshop%20on%20Perceptive%20User%20Interfaces&tn=SE_baiduxueshu_c1gjeupa&ie=utf-8&sc_f_para=sc_hilight%3Dpublish&sort=sc_cited), 2002:1-4

[3] Lam Tran，Deguang Kong, Hongxia Jin，Ji Liu. Privacy-CNH: A Framework to Detect Photo Privacy with Convolutional Neural Network Using Hierarchical Features [C].AAAI 2016: 1317-1323

[4] Jihun Hamm:Minimax Filter: Learning to Preserve Privacy from Inference Attacks. [C]CoRR abs/1610.03577 (2016)

[5] Reza Shokri, Marco Stronati, Vitaly Shmatikov:Membership Inference Attacks against Machine Learning Models. [C]CoRR abs/1610.05820 (2016)

[6] S Süsstrunk，R Buckley，S Swen：Standard RGB Color Spaces；［C］Color & Imaging Conference, 1999, 7:127-134(8)

[7] 刘益新，郭依正：灰度直方图特征提取的Matlab实现［J］；《电脑知识与技术》, 2009, 5(32):9032-9034

[8] C Dwork：Differential Privacy: A Survey of Results［C］；Theory & Applications of Models of Computation, International Conference, Tamc, Xian, China, April, 2008, 4978:1-19

[9] ME Nergiz，C Clifton：Thoughts on k-anonymization［J］《Data & Knowledge Engineering》, 2006, 63(3):622-645

[10] H Schreier，JJ Orteu，MA Sutton：Image Correlation for Shape, Motion and Deformation Measurements［J］；Springer US, 2009:565-600

[11] MD Franzen：Chi-square［J］；Springer New York, 2011:559-559

[12] McMorris，R F：Topics in intersection graph theory［C］；SIAM,, 1999

[13] T Jebara，RI Kondor：Bhattacharyya Expected Likelihood Kernels ［C］；Conference on Computational Learning Theory & Kernel Machines 2003, 2777:57-71

[14] 李月龙，封举富：基于最小扭曲变换的正面人脸图像合成［J］；《计算机辅助设计与图形学学报》, 2011, 23(6):1085-1090

[15] Suman Jana, Arvind Narayanan, Vitaly Shmatikov:A Scanner Darkly: Protecting User Privacy from Perceptual Applications. ［C］IEEE Symposium on Security and Privacy 2013: 349-363

[16] Cynthia Dwork, Kunal Talwar, Abhradeep Thakurta, Li Zhang:Analyze gauss: optimal bounds for privacy-preserving principal component analysis. ［C］STOC 2014: 11-20

[17] Martín Abadi, Andy Chu, Ian J. Goodfellow, H. Brendan McMahan, Ilya Mironov, Kunal Talwar, Li Zhang:Deep Learning with Differential Privacy.［C］ ACM Conference on Computer and Communications Security 2016: 308-318

[18] Yining Wang, Yu-Xiang Wang, Aarti Singh:A Deterministic Analysis of Noisy Sparse Subspace Clustering for Dimensionality-reduced Data.［C］ ICML 2015: 1422-1431

[19] Kevin W. Bowyer:Face recognition technology: security versus privacy. ［C］IEEE Technol. Soc. Mag. 23(1): 9-19 (2004)

[20] Tom A. M. Kevenaar, Geert Jan Schrijen, Michiel van der Veen, Anton H. M. Akkermans, Fei Zuo:Face Recognition with Renewable and Privacy Preserving Binary Templates.［C］ AutoID 2005: 21-26

[21] Ahmad-Reza Sadeghi, Thomas Schneider, Immo Wehrenberg:Efficient Privacy-Preserving Face Recognition.［C］ ICISC 2009: 229-244

[22] Charu C. Aggarwal, Philip S. Yu:A General Survey of Privacy-Preserving Data Mining Models and Algorithms. ［C］Privacy-Preserving Data Mining 2008: 11-52

[23] Luminita Vasiu:Biometric Recognition - Security and Privacy Concerns. ［C］ ICETE (Invited Speakers) 2004: 3

[24] Moo-Ryong Ra, Ramesh Govindan, Antonio Ortega:P3: Toward Privacy-Preserving Photo Sharing.［C］ NSDI 2013: 515-528

[25] Zekeriya Erkin, Martin Franz, Jorge Guajardo, Stefan Katzenbeisser, Inald Lagendijk, Tomas Toft:Privacy-Preserving Face Recognition.［C］Privacy Enhancing Technologies 2009: 235-253

[26] 石香灵：隐私保护模型研究［D］；《重庆大学》, doi： 10.7666/d.y1666001:2009

[27] 任向民：基于K-匿名的隐私保护方法研究［D］；《哈尔滨工程大学》:2012

[28]Ninghui Li, Min Lyu, Dong Su, Weining Yang:Differential Privacy: From Theory to Practice. Synthesis Lectures on Information Security, Privacy, & Trust, Morgan & Claypool Publishers 2016, pp. 1-138

[29] 欧阳佳, 印鉴, 刘少鹏：一种有效的差分隐私事务数据发布策略[C]；中 国计算机学会人工智能会议. 2013.

[30] Dwork C, Mcsherry F, Nissim K： Calibrating Noise to Sensitivity in Private Data Analysis[J]； Proceedings of the Vldb Endowment, 2012, 7(8):637-648.

[31] Mcsherry F, Talwar K. Mechanism Design via Differential Privacy[C]// Foundations of Computer Science, 2007. FOCS '07. IEEE Symposium on. IEEE, 2007:94-103.

[32] 公艳艳, ShuiRende, GongYanyan,等：人脸的类别与特定身份信息提取方式研究[J]； 心理科学, 2008, 31(4):896-900.

致谢

本文研究工作的顺利完成离不开导师华景煜老师和学长孙弘毅对我的悉心指导，同时还有教研室组长仲盛老师的监督。首先，非常感谢华景煜老师帮助我确定毕业设计研究课题，并指导我进行论文编写，提出大量修改意见。其次，非常感谢孙弘毅学长在刚开始进行毕业设计时向我提供大致研究思路，并及时解答我在完成毕业设计过程中遇到的各种问题。

通过本次毕业设计，我得到了许多成长，自己独立完成工作的能力、调研资料的能力、代码编写的能力等等，都得到大大提高。所以非常感谢学校开设毕业设计这样的课题来进一步检验和锻炼我们的研究和创新能力。同时，也要感谢大学四年教导我的各位老师，您们传授的知识是我完成毕业设计的基础，您们渊博的知识和严谨的学术研究作风都给我留下深刻的印象，今后也会继续向您们学习。

另一方面，由于后期自己身体不好，感谢系主任和导师批准我在家完成毕业设计剩下的工作。而在家网络条件受限导致我无法下载相关文献，非常感谢谭琦佳、易璇、朱禹龙、顾颖颖等同学在我需要的时候，及时帮我下载资料，使我的毕业论文的撰写工作得以顺利进行。

最后，再次向所有在大学四年里给予我关心和帮助的人们表示由衷的感谢！