



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2020
04
VOL.25

ISSN1006-8961
CN11-3758/TB



海岛岸线自动分割 P768



第25卷第4期 (总第288期)

2020年4月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿, 均视为同意在本刊网站及CNKI等全文数据库出版, 所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用, 所有内容, 未经许可, 不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 顾行发
编辑出版 《中国图象图形学报》编辑出版委员会
通信地址 北京市海淀区北四环西路19号
邮 编 100190
电子信箱 jig@radi.ac.cn
电 话 010-58887035
网 址 www.cjig.cn

广告发布登记号 京朝工商广登字20170218号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief Gu Xingfa
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
Address No. 19, North 4th Ring Road West, Haidian District, Beijing, P. R. China
Zip code 100190
E-mail jig@radi.ac.cn
Telephone 010-58887035
Website www.cjig.cn

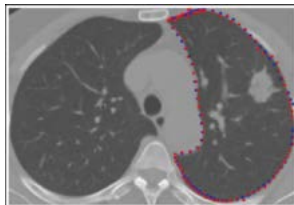
Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ

国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



自适应在线判别外观学习的
分层关联多目标跟踪
(第0708页)



低秩和主动形状模型的肺区
分割(第0759页)



结合曲率滤波的HTM算法去
除遥感影像云雾(第0791页)

综述

深度学习目标检测方法综述

赵永强, 饶元, 董世鹏, 张君毅 0629

图像处理和编码

结合矩阵分解与差分隐私的人脸图像发布

张啸剑, 付聪聪, 孟小峰 0655

图像分析和识别

多尺度深度特征融合的变化检测

樊玮, 周末, 黄睿 0669

面向类内差距表情的深度学习识别

陈亮, 吴攀, 刘韵婷 0679

带后缀印刷体维吾尔文映射关系检索

伊克萨尼·普尔凯提, 阿布都萨拉木·达吾提, 艾斯卡尔·艾木都拉 0688

结合混沌鸟群算法的阴极铜板表面缺陷检测

王卓, 张长胜, 李伟, 钱俊兵, 唐都作, 蔡兵, 常以涛 0697

自适应在线判别外观学习的分层关联多目标跟踪

方岚, 于凤芹 0708

图像理解和计算机视觉

多方向Laplacian能量和与tetrolet变换的图像融合

沈瑜, 陈小鹏, 杨倩 0721

双重对偶生成对抗网络的跨年龄素描—照片转换

吴柳玮, 孙锐, 阚俊松, 高隽 0732

医学图像处理

残差密集相对平均CGAN的脑部图像配准

王丽芳, 张程程, 秦品乐, 蔺素珍, 高媛, 窦杰亮 0745

低秩和主动形状模型的肺区分割

孙申申, 田丹, 吴微, 康雁, 赵宏 0759

遥感图像处理

改进Deeplab网络的遥感影像海岸线分割

王振华, 钟元芾, 何婉雯, 曲念毅, 徐利智, 张文苹, 刘智翔 0768

卷积神经网络提取风云影像土壤湿度

韩颖娟, 周顺武, 王文文, 张学艺, 张承明, 栗佩琪 0779

结合曲率滤波的HTM算法去除遥感影像云雾

杨珂珂, 贾渊, 沈川 0791

ChinaMM 2019会议专栏

利用稳健非负矩阵分解实现无监督高光谱解混

宋晓瑞, 吴玲达, 孟祥利 0801

关键点深度特征驱动人脸表情识别

王善敏, 帅惠, 刘青山 0813

融合区域和边缘特征的水平集水下图像分割

孙杨, 陈哲, 王慧斌, 张振, 沈洁 0824

深度时空能量特征表示下的人体行为识别

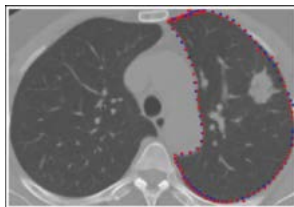
巢新, 侯振杰, 李兴, 梁久祯, 宦娟, 刘浩昱 0836

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Multi-object tracking based on adaptive online discriminative appearance learning and hierarchical association(P0708)



Lung segmentation by active shape model approach based on low rank theory(P0759)



Haze and cloud removal from remote sensing image using HTM algorithm based on curvature filtering(P0791)

Review

Survey on deep learning object detection

Zhao Yongqiang, Rao Yuan, Dong Shipeng, Zhang Junyi 0629

Image Processing and Coding

Private facial image publication through matrix decomposition

Zhang Xiaojian, Fu Congcong, Meng Xiaofeng 0655

Image Analysis and Recognition

Multiscale deep features fusion for change detection

Fan Wei, Zhou Mo, Huang Rui 0669

Depth learning recognition method for intra-class gap expression

Chen Liang, Wu Pan, Liu Yunting 0679

Mapping relationship retrieval of Uyghur-printed suffix word

Eksan Firkat, Abdusalam Dawut, Askar Hamdulla 0688

Cathodic copper plate surface defect detection based on bird swarm algorithm with chaotic theory

Wang Zhuo, Zhang Changsheng, Li Wei, Qian Junbing, Tang Duzuo, Cai Bing, Chang Yitao 0697

Multi-object tracking based on adaptive online discriminative appearance learning and hierarchical association

Fang Lan, Yu Fengqin 0708

Image Understanding and Computer Vision

Image fusion of multidirectional sum modified Laplacian and tetrolet transform

Shen Yu, Chen Xiaopeng, Yang Qian 0721

Double dual generative adversarial networks for cross-age sketch-to-photo translation

Wu Liuwei, Sun Rui, Kan Junsong, Gao Jun 0732

Medical Image Processing

Image registration method with residual dense relativistic average CGAN

Wang Lifang, Zhang Chengcheng, Qin Pinle, Lin Suzhen, Gao Yuan, Dou Jieliang 0745

Lung segmentation by active shape model approach based on low rank theory

Sun Shenshen, Tian Dan, Wu Wei, Kang Yan, Zhao Hong 0759

Remote Sensing Image Processing

Island shoreline segmentation in remote sensing image based on improved Deeplab network

Wang Zhenhua, Zhong Yuanfu, He Wanwen, Qu Nianyi, Xu Lizhi, Zhang Wenping, Liu Zhixiang 0768

Extracting soil moisture from Fengyun satellite images using a convolutional neural network

Han Yingjuan, Zhou Shunwu, Wang Wenwen, Zhang Xueyi, Zhang Chengming, Li Peiqi 0779

Haze and cloud removal from remote sensing image using HTM algorithm based on curvature filtering

Yang Keke, Jia Yuan, Shen Chuan 0791

Column of ChinaMM 2019

Unsupervised hyperspectral unmixing based on robust non-negative matrix factorization

Song Xiaorui, Wu Lingda, Meng Xiangli 0801

Facial expression recognition based on deep facial landmark features

Wang Shanmin, Shuai Hui, Liu Qingshan 0813

Level set method combining region and edge features for segmenting underwater images

Sun Yang, Chen Zhe, Wang Huibin, Zhang Zhen, Shen Jie 0824

Action recognition under depth spatial-temporal energy feature representation

Chao Xin, Hou Zhenjie, Li Xing, Liang Jiuzhen, Huan Juan, Liu Haoyu 0836

中图分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2020)04-0655-14

论文引用格式: Zhang X J, Fu C C and Meng X F. 2020. Private facial image publication through matrix decomposition. Journal of Image and Graphics, 25(04):0655-0668 [张啸剑, 付聪聪, 孟小峰. 2020. 结合矩阵分解与差分隐私的人脸图像发布. 中国图象图形学报, 25(04):0655-0668] [DOI: 10.11834/jig.190308]

结合矩阵分解与差分隐私的人脸图像发布

张啸剑¹, 付聪聪¹, 孟小峰²

1. 河南财经政法大学计算机与信息工程学院, 郑州 450046; 2. 中国人民大学信息学院, 北京 100872

摘要: 目的 人脸图像蕴含着丰富的个人敏感信息, 直接发布可能会造成个人隐私泄露。为了保护人脸图像中的隐私信息, 提出 3 种基于矩阵分解与差分隐私技术相结合的人脸图像发布算法, 即 LRA (low rank-based private facial image release algorithm)、SRA (SVD-based private facial image release algorithm) 和 ESRA (enhanced SVD-based private facial image release algorithm)。方法 为了减少拉普拉斯机制带来的噪音误差, 3 种算法均将人脸图像作为实数域 2 维矩阵, 充分利用矩阵低秩分解与奇异值分解技术压缩图像。在 SRA 和 ESRA 算法中, 如何选择矩阵压缩参数 r 会直接制约由拉普拉斯机制引起的噪音误差以及由矩阵压缩导致的重构误差。SRA 算法利用启发式设置参数 r , 然而 r 值增大导致过大的噪音误差, r 值减小导致过大的重构误差。为了有效均衡这两种误差, ESRA 算法引入一种基于指数机制的挑选参数 r 的方法, 能够在不同的分解矩阵中挑选合理的矩阵尺寸来压缩人脸图像, 然后利用拉普拉斯机制对挑选的矩阵添加相应的噪音, 进而使整个处理过程满足 ϵ -差分隐私。结果 基于 6 种真实人脸图像数据集, 采用支持向量机 (support vector machine, SVM) 分类技术与信息熵验证 6 种算法的正确性。从算法的准确率、召回率、F1-Score, 以及信息熵度量结果显示, 提出的 LRA、SRA 与 ESRA 算法均优于 LAP (Laplace-based facial image protection)、LRM (low-rank mechanism) 以及 MM (matrix mechanism) 算法, 其中 ESRA 算法在 Faces95 数据集上的准确率和 F1-Score 分别是 LRA、LRM 和 MM 算法的 40 倍、20 倍和 1 倍多。相对于其他 5 种算法, ESRA 算法对数据集大的变化相对稳定, 可用性最好。结论 本文算法能够实现满足 ϵ -差分隐私的敏感人脸图像发布, 具有较好的可用性与鲁棒性, 并且为灰度人脸图像的隐私保护提供了新的指导方法与思路, 能有效用于社交平台和医疗系统等领域。**关键词:** 人脸图像; 隐私保护; 差分隐私; 矩阵分解; 低秩分解; 奇异值分解

Private facial image publication through matrix decomposition

Zhang Xiaojian¹, Fu Congcong¹, Meng Xiaofeng²

1. School of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou 450046, China;

2. School of Information, Renmin University of China, Beijing 100872, China

Abstract: **Objective** Facial images are widely used in many applications such as social media, medical systems, and smart transportation systems. Such data, however, are inherently sensitive and private. Individuals' private information may be leaked if their facial images are released directly in the application systems. In social network platforms, attackers can use the facial images of individuals to attack their sensitive information. Many classical privacy-preserving methods, such as

收稿日期: 2019-06-28; 修回日期: 2019-08-30; 预印本日期: 2019-09-06

基金项目: 国家自然科学基金项目 (61502146, 91646203, 91746115, 61772131); 河南省自然科学基金项目 (162300410006); 河南省科技攻关项目 (162102310411); 河南省教育厅高等学校重点科研项目 (16A520002); 河南省高等学校青年骨干教师培养计划项目 (2017GGJS084); 河南财经政法大学青年拔尖人才资助计划项目

Supported by: National Natural Science Foundation of China (61502146, 91646203, 91746115, 61772131); Natural Science Foundation of Henan Province, China (162300410006)

k -anonymous and data encryption, have been proposed to handle the privacy problem in facial images. However, the classical methods always rely on strong background assumptions, which cannot be supported in real-world applications. Differential privacy is the state-of-the-art method used to address the privacy concerns in data publication, which provides rigorous guarantees for the privacy of each user by adding randomized noise in Google Chrome, Apple iOS, and macOS. Therefore, to protect the private information in facial images, this paper proposes three efficient algorithms, namely, low rank-based private facial image release algorithm (LRA), singular value decomposition (SVD)-based private facial image release algorithm (SRA), and enhanced SVD-based private facial image release algorithm (ESRA), which are based on matrix decomposition combined with differential privacy. **Method** The three algorithms employed the real-valued matrix to model facial images in which each cell corresponds to each pixel point of facial images. Based on the real-valued matrix, the neighborhood of some facial images can be defined easily, which are crucial bases to use Laplace mechanism to generate Laplace noise. Then, LRA, SRA, and ESRA rely on low-rank decomposition and SVD to compress facial images. This step aims to reduce the Laplace noise and boost the accuracy of the publication of facial images. The three algorithms use the Laplace mechanism to inject noise into each value of the compressed facial image to ensure differential privacy. Finally, the three algorithms use matrix algebraic operations to reconstruct the noisy facial image. However, in the SRA and ESRA algorithms, two sources of errors are encountered: 1) Laplace error (LE) due to Laplace noise injected and 2) reconstruction error (RE) caused by lossy compression. The two errors are controlled by r parameter, which is the compression factor in the SRA and ESRA algorithms. Setting the compact parameter r constrains the LE and RE. The SRA algorithm sets the parameter in a heuristic manner in which one may fix the value in terms of experiences. However, the choice of r in the SRA algorithm is a problem because a large r leads to excessive LE, while a small r makes the RE extremely large. Furthermore, r cannot be directly set based on the real-valued matrix; otherwise, the choice of r itself violates differential privacy. Based on the preceding observation, the ESRA algorithm is proposed to handle the problem caused by the selection of the parameter r . The main idea of the ESRA algorithm involves two steps: employing exponential mechanism to sample r elements in the decomposition matrix and injecting the Laplace noise into the elements. According to the sequential composition of differential privacy, the two steps in the ESRA algorithm meet ϵ -differential privacy. **Result** On the basis of the SVM classification and information entropy technique, two group experiments were conducted over six real facial image datasets (Yale, ORL, CMU, Yale B, Faces95, and Faces94) to evaluate the quality of the facial images generated from the LRA, SRA, ESRA, LAP (Laplace-based facial image protection), LRM (low-rank mechanism), and MM (matrix mechanism) algorithms using a variety of metrics, including precision, recall, F1 score, and entropy. Our experiments show that the proposed LRA, SRA, and ESRA algorithms outperform LAP, LRM, and MM in terms of the abovementioned six metrics. For example, based on the Faces95 dataset, $\epsilon = 0.1$ and matrix = 200×180 were set to compare the precision of ESRA, LRM, LRA, and LAP. Result show that the precision of ESRA is 40 and 20 times that of LAP, LRA, and LRM. Based on the six datasets, ESRA achieves better accuracy than LRA and SRA. For example, on the Faces94 dataset, the matrix = 200×180 was set and the privacy budget ϵ (i.e., 0.1, 0.5, 0.9, and 1.3) was varied to study the utility of each algorithm. Results show that the utility measures of all algorithms increase when ϵ increases. When ϵ varies from 0.1 to 1.3, ESRA still achieves better precision, recall, F1-score, and entropy than the other algorithms. **Conclusion** Theoretical analysis and extensive experiments were conducted to compare our algorithms with the LAP, LRM, and MM algorithms. Results show that the proposed algorithms could achieve better utility and outperform the existing solutions. In addition, the proposed algorithms provided new ideas and technical support for further research on facial image release with differential privacy in several application systems.

Key words: facial image; privacy protection; differential privacy; matrix decomposition; low-rank decomposition; singular value decomposition (SVD)

0 引言

信息技术的快速发展使得人脸图像的获取与识别变得尤为容易。社交网络服务平台分析用户上传

的手机中的人脸图像时,通过对比多张人脸的相似度,判断是否同一个人;医疗系统通过分析收集的用户人脸图像,能够分析出性别、年龄、表情等多种人脸属性。然而人脸图像中蕴含着个人敏感信息,直接发布或者上传,会造成泄露。如何在保护个人隐

私的前提下,发布人脸图像数据是目前亟待解决的问题。近年来研究者提出了多种隐私保护方法来保护人脸图像,通常使用的方法包括 k -匿名、访问控制、隐私加密等。Newton 等人(2005)和 Gross 等人(2008)利用匿名化技术提出了 k -same方法,对发布的灰度人脸图像进行匿名化操作,使攻击者通过发布的人脸图像重新甄别出用户身份的概率不超过 $1/k$ 。然而,匿名化技术通常依赖于很强的背景知识假设,而相应的假设在现实中并不完全成立,进而使匿名化技术失效。例如 Ilia 等人(2015)指出,将匿名化的人脸图像放到 Facebook 社交平台,攻击者通过该平台额外的身份特征可以推理出匿名人脸图像中个人的社会安全号码,凭借社会安全号码,攻击者可以成功甄别出个人的疾病、住址等敏感信息。Sadeghi 等人(2009)采用访问控制技术对社交平台上人脸图像的访问权限进行限制,达到了隐私保护的目,但是该方法并不是从根本上对人脸图像进行隐私保护,而是在访问设置方面对图像进行保护。如果攻击者具有一定的背景知识和访问权限,则可以获取用户的原始人脸图像。Gross 和 Acquisti (2005)利用同态加密技术对灰度人脸图像进行加密处理,防止通信过程被甄别。然而数据加密技术同样是先对攻击做出相应的假设,再基于这些假设设计相应的加密算法,显然这类加密方法仅能针对某些特定的攻击。

当攻击者具有特定的背景知识时,匿名化与数据加密技术就会在一定程度上泄露用户的敏感人脸图像信息。目前,差分隐私已经成为一种新型隐私保护技术,该技术不关心攻击者具有多少背景知识,通常利用随机噪音实现隐私保护(Dwork 等,2006; Dwork 和 Lei,2009),研究对象已遍布关系数据、空间数据、流数据等多种数据类型,但对人脸图像数据鲜有涉及,主要挑战来自人脸图像的高维度以及如何保持发布精度。本文基于差分隐私提出高效的灰度人脸图像发布方法。主要贡献如下:

1)为了减少人脸图像中噪音的影响,分别基于低秩分解与奇异值分解技术提出3种图像压缩方法,即 LRA (low rank-based private facial image release algorithm)、SRA (singular value decomposition (SVD)-based private facial image release algorithm) 和 ESRA (enhanced SVD-based private facial image release algorithm),针对压缩的图像添加相应的拉普

拉斯噪音。

2)为了减少 SRA 和 ESRA 方法带来的重构误差,提出一种基于指数机制的分解矩阵尺寸选择方法,利用指数机制在不同的分解矩阵中挑选合适的矩阵尺寸进行图像压缩,并且使挑选过程满足差分隐私。

3)理论证明,本文人脸图像发布方法能够满足 ϵ -差分隐私,实验表明在6种真实的人脸库上具有有效性与可用性。

1 相关工作

随着网络的普及,用户数据的隐私保护受到研究者的广泛关注。目前,差分隐私保护模型依靠系统的统计学基础,打破了传统保护模型思路,成为研究热点。直方图发布(Xu 等,2013)、空间数据发布(Qardaji 等,2013)、图数据发布(Zhang 等,2015)、流数据发布(Chen 等,2017)、时序数据发布(Rastogi 和 Nath,2010)和高维数据发布(Su 等,2016)等已开展相应的研究。灰度人脸图像是一种较为复杂的个人数据,通常蕴含敏感信息,具有复杂性与易共享性。目前还没有研究者采用差分隐私技术对其进行保护,若共享平台直接发布,可能对个人隐私造成威胁。利用差分隐私保护技术发布人脸图像是亟待解决的问题。

傅里叶变换与小波变换是人脸图像处理常用的压缩技术,已有研究者采用这两种技术处理1维隐私数据。Rastogi 和 Nath(2010)借助傅里叶变换与差分隐私技术发布1维时序数据,主要思想是对1维时序数据进行傅里叶变换,然后对傅里叶系数添加拉普拉斯噪音;Xiao 等人(2011)利用小波变换与差分隐私技术发布1维序列数据,主要思想与傅里叶变换类似,对小波系数进行噪音处理。然而这两种方法无法直接应用于人脸图像,原因是无法直接采用 Rastogi 和 Nath(2010)以及 Su 等人(2016)的方法度量人脸图像发布的全局敏感性,进而无法计算需要的噪音数量。

实数域矩阵是人脸图像常用的表示形式,已有工作利用矩阵技术批处理查询(Li 等,2015; Yuan 等,2015,2016)与推荐系统(Friedman 等,2016)。Li 等人(2015)利用矩阵机制响应批处理查询,结合批处理矩阵搜索最优的策略矩阵,根据策略矩阵的

敏感性添加拉普拉斯噪音,最后生成满足差分隐私的响应结果。然而 Yuan 等人(2015)指出矩阵机制具有较高的计算复杂性,响应精度通常低于直接在原始数据添加噪音的响应结果。主要原因是要求策略矩阵满足摩尔-彭若斯广义逆,而这一条件使得搜索到的策略矩阵未必是最优解。Yuan 等人(2015)结合 Li 等人(2015)的不足提出了利用矩阵低秩分解机制响应最终的批处理查询,该机制基于 1 阶增广拉格朗日算子求最终的策略矩阵,进而避免策略矩阵必须满足摩尔-彭若斯广义逆约束。随后 Yuan 等人(2016)指出由于问题求解空间的非线性与非凸性,可能导致低秩策略求解过程的非收敛性,进而导致找不到全局最优的策略矩阵。此外, Friedman 等人(2016)在评分矩阵中利用矩阵分解技术,根据损失函数与梯度下降法求解评分矩阵的分解结果,最后对分解结果添加噪音。然而评分矩阵的高度稀疏性与人脸图像矩阵有着本质区别。本文基于矩阵低秩与奇异值分解提出 3 种满足差分隐私的人脸图像发布方法,基本思想是利用矩阵分解将人脸图像中的特征信息提取出来,对其表示的特征信息矩阵添加相应的拉普拉斯噪音,发布的人脸图像具有较高的可用性。

2 定义与问题

2.1 基于近邻矩阵的差分隐私定义

给定一幅人脸图像,将每一个像素点作为一个单元,则人脸图像可以表示成一个 2 维数据矩阵 $A_{m \times n}$,其中 m 表示矩阵 A 的行数, n 表示 A 的列数。在给出差分隐私形式化定义之前,结合 A 给出近邻关系的定义。

定义 1 近邻矩阵。给定一幅尺寸为 $m \times n$ 的人脸图像

$$A_{m \times n} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$A_{m \times n} \in \mathcal{C}^{m \times n}$$

式中, $\mathcal{C}^{m \times n}$ 表示 $m \times n$ 的实内积空间, a_{ij} 表示人脸图像中的某个像素点。用向量将 A 表示成 $A = (a_1, a_2, \dots, a_n)$, 其中 $a_i = (a_{1i}, a_{2i}, \dots, a_{mi})^T$ 。如果 A

与 A' 相差一个向量,则 A 与 A' 互为近邻矩阵 (Dwork 等, 2006)。

根据近邻矩阵 A 与 A' ,重新定义 ε -差分隐私模型。

定义 2 ε -差分隐私。给定近邻矩阵 A 与 A' , 设 M 为一个随机人脸图像发布方法,若 M 在近邻矩阵 A 与 A' 上的任意一个输出结果 I 满足

$$Pr(M(A) = I) \leq \exp(\varepsilon) \times Pr(M(A') = I) \quad (1)$$

则 M 满足 ε -差分隐私。式(1)中, ε 表示发布人脸图像所需要的隐私代价, Pr 为概率。 ε 值的大小直接制约方法 M 的隐私保护程度。

实现差分隐私保护最基本的要求是查询全局敏感性与噪音方法 (Dwork 等, 2006)。查询敏感性是度量某一查询在近邻矩阵上的最大改变量。利用查询敏感性与隐私代价的比率作为噪音尺度参数,借助典型分布(拉普拉斯分布、指数分布等)产生差分隐私所需的噪音,利用噪音即可对数值型与非数值数据进行 ε -差分隐私保护。

2.2 矩阵分解

矩阵分解是一种常用的数字图像压缩方法。本文利用矩阵低秩分解与矩阵奇异值分解来压缩人脸图像。

定义 3 矩阵低秩分解。给定一幅人脸图像,其表示矩阵为 $A_{m \times n}$ 。设 k ($k \leq \min\{m, n\}$) 为其秩,若存在秩为 k 的两个矩阵 $B_{m \times k}$ 与 $L_{k \times n}$,使得

$$A_{m \times n} = B_{m \times k} L_{k \times n} \quad (2)$$

则称其为 $A_{m \times n}$ 的低秩分解。式(2)中, $B_{m \times k}$ 表示 $m \times k$ 的列满秩矩阵, $L_{k \times n}$ 表示 $k \times n$ 的行满秩矩阵。由式(2)可知,对 $A_{m \times n}$ 分解之后的主要特征信息保留在矩阵 $L_{k \times n}$ 中。因此利用噪音机制对矩阵 $L_{k \times n}$ 进行扰动则可以保护 $A_{m \times n}$ 中的敏感信息。

定义 4 矩阵奇异值分解。设 A 为一幅人脸图像的表示矩阵,该矩阵对应的奇异值分解为

$$A = Q_A^{m \times m} D_A^{m \times n} (P_A^{n \times n})^T \quad (3)$$

式中, Q_A 是 $m \times m$ 的正交矩阵, D_A 是 $m \times n$ 的对角阵, P_A^T 是 $n \times n$ 的正交矩阵。当 $m > n$ 时, D_A 是由一个 $n \times n$ 的对角子矩阵与 $0^{(m-n) \times n}$ 组合而成的 $m \times n$ 的对角阵。

2.3 问题描述

给定一幅人脸图像及其相应的表示矩阵 A ,利用矩阵分解技术将 A 拆解为多个矩阵,而多个矩阵

的意义并不相同,在添加噪音时,只需对包含主要信息的矩阵采用拉普拉斯机制扰动,然后利用矩阵乘积运算重新恢复 $\mathbf{A}$,进而得到 $\tilde{\mathbf{A}}$ 。在获得 $\tilde{\mathbf{A}}$ 的过程中存在两种误差,一种是拉普拉斯机制导致的噪音误差 $LE(\tilde{\mathbf{A}})$,另一种是矩阵重新构建过程中的重构误差 $RE(\tilde{\mathbf{A}})$ 。因此,发布人脸图像 $\tilde{\mathbf{A}}$ 的总体误差可以表示为

$$Er(\tilde{\mathbf{A}}) = RE(\tilde{\mathbf{A}}) + LE(\tilde{\mathbf{A}}) \quad (4)$$

在设计人脸图像发布方法时,应尽可能使总体误差 $Er(\tilde{\mathbf{A}})$ 最小。

3 基于矩阵分解的人脸图像处理方法

3.1 LAP 发布方法

LAP(Laplace-based facial image protection)方法是直接利用拉普拉斯噪音(Dwork 等,2006)对 $\mathbf{A}$ 进行噪音扰动,发布扰动后的数据。LAP 是最为直接的基于拉普拉斯机制的方法。

添加噪音的形式化公式可以表示为

$$\tilde{\mathbf{A}} = \mathbf{A} + Lap\left(\frac{\Delta_1 \mathbf{A}}{\varepsilon}\right) \quad (5)$$

式中, $Lap(\cdot)$ 表示对矩阵加噪, $\Delta_1 \mathbf{A} = \max_j \sum_{i=1}^m |a_{ij}|$,即为矩阵 $\mathbf{A}$ 最大的列范数。由式(5)可知,LAP 方法是对 $\mathbf{A}$ 中的每个数值添加拉普拉斯噪音,即 $\tilde{a}_i = a_i + lap(\Delta_1 \mathbf{A} / \varepsilon)$, $Lap(\cdot)$ 表示对行向量加噪。

LAP 发布方法满足 ε -差分隐私(Zhang 等,2015)。然而, $\mathbf{A}$ 为实数域矩阵,通常会造成比较大的 $\Delta_1 \mathbf{A}$,进而导致发布结果的可用性很低。若采用误差平方和的期望值来度量 $LE(\tilde{\mathbf{A}})$,则 LAP 方法产生的误差为

$$\begin{aligned} LE(\tilde{\mathbf{A}}) &= E\left(\sum_{i=1}^n error(\tilde{a}_i)\right) = E\left(\sum_{i=1}^n (\tilde{a}_i - a_i)^2\right) = \\ &E\left(\sum_{i=1}^n \left(a_i - a_i + \sum_{j=1}^m lap\left(\frac{\Delta_1 \mathbf{A}}{\varepsilon}\right)\right)^2\right) = \\ &\sum_{i=1}^n \frac{2m(\Delta_1 \mathbf{A})^2}{\varepsilon^2} = 2mn \left(\frac{\Delta_1 \mathbf{A}}{\varepsilon}\right)^2 = \\ &2mn \left(\frac{\max_j \sum_{i=1}^m |a_{ij}|}{\varepsilon}\right)^2 \end{aligned} \quad (6)$$

式中, E 为期望。由式(6)可知,当图像矩阵的尺寸与 $\Delta_1 \mathbf{A}$ 的数值较大时,产生的总误差就较大。因此,直接利用 LAP 发布尺寸较大的人脸图像,会导

致发布结果的可用性较低。为此,本文基于矩阵分解提出了 LRA、SRA 以及 ESRA 等 3 种发布方法。

3.2 基于矩阵分解的发布方法

3.2.1 基于矩阵低秩分解的发布方法

首先利用矩阵低秩分解对原始人脸图像矩阵进行压缩,具体分为两步:1)采用矩阵分解的方法,将 $\mathbf{A}_{m \times n}$ 拆解为两个低秩矩阵 $\mathbf{B}_{m \times k}$ 与 $\mathbf{L}_{k \times n}$;2)对 $\mathbf{L}_{k \times n}$ 添加拉普拉斯噪音,即 $\mathbf{L}_{k \times n} + lap(\Delta_1 \mathbf{L} / \varepsilon)^k$,进而获得满足差分隐私的人脸图像矩阵 $\tilde{\mathbf{A}}_{m \times n} = \mathbf{B}_{m \times k}(\mathbf{L}_{k \times n} + lap(\Delta_1 \mathbf{L} / \varepsilon)^k)$ 。结合式(6)度量基于低秩矩阵分解产生的拉普拉斯误差 $LE(\tilde{\mathbf{A}})$,即

$$\begin{aligned} LE(\tilde{\mathbf{A}}) &= \sum_{i=1}^n error(\tilde{a}_i) = E\left(\sum_{i=1}^n (\tilde{a}_i - a_i)^2\right) = \\ &E\left(\sum_{i=1}^m \sum_{j=1}^n \left(b_i \left(l_j + lap\left(\frac{\Delta_1 \mathbf{L}}{\varepsilon}\right)\right) - b_i l_j\right)^2\right) = \\ &E\left(\sum_{i=1}^m \sum_{j=1}^n \left(b_i \cdot lap\left(\frac{\Delta_1 \mathbf{L}}{\varepsilon}\right)\right)^2\right) = \\ &\left(\sum_{ij} B_{ij}^2\right) \frac{2(\Delta_1 \mathbf{L})^2}{\varepsilon^2} \end{aligned} \quad (7)$$

式中, $\Delta_1 \mathbf{L} = \max_j \sum_{i=1}^k |l_{ij}|$ 。重写式(7)可得

$$LE(\tilde{\mathbf{A}}) = 2\left(\sum_{ij} B_{ij}^2 \times (\Delta_1 \mathbf{L})^2\right) \frac{1}{\varepsilon^2}$$

由于 $\mathbf{A}_{m \times n} = \mathbf{B}_{m \times k} \mathbf{L}_{k \times n}$ 分解的不唯一性,导致不同的 $LE(\tilde{\mathbf{A}})$ 。若要获得可用的人脸图像,只要获得较小的 $\sum_{ij} B_{ij}^2 \times (\Delta_1 \mathbf{L})^2$ 即可,较小的 $\sum_{ij} B_{ij}^2 \times (\Delta_1 \mathbf{L})^2$ 与 $\mathbf{A}$ 的最优分解 $\mathbf{B}_{m \times k}$ 和 $\mathbf{L}_{k \times n}$ 密切相关,本文结合随机梯度下降方法,将该问题转化为无约束最优化问题,利用正则化矩阵分解迭代地学习出 $\mathbf{B}_{m \times k}$ 与 $\mathbf{L}_{k \times n}$,求解该问题所需的目标函数为

$$\min_{\mathbf{B}, \mathbf{L}} \sum_{a_{ij} \in \mathbf{A}} [(a_{ij} - b_i(\tilde{l}_j))^2 + \tau(\|\mathbf{b}_i\|^2 + \|\tilde{l}_j\|^2)] \quad (8)$$

式中, $(\tilde{l}_j)^T = (l_j)^T + lap(\Delta_1 \mathbf{L} / \varepsilon)^k$, τ 表示正则因子。结合式(8),在正则化损失函数梯度相反的方向上同时更新 $\mathbf{B}_{m \times k}$ 与 $\mathbf{L}_{k \times n}$,得

$$\begin{aligned} b_i &\leftarrow b_i + \gamma(e_{ij} \cdot (l_j + lap(\frac{\Delta_1 \mathbf{L}}{\varepsilon})) - \tau \cdot b_i) \\ l_j &\leftarrow l_j + \gamma(e_{ij} \cdot b_i - \tau \cdot l_j) \end{aligned}$$

式中, e_{ij} 表示每次迭代产生的误差函数, $e_{ij} = a_{ij} - b_i \cdot (l_j)^T$,常数 γ 用来确定最小化误差的速率,通常称为学习速率。

基于上述分析,本文提出 LRA 算法,具体步骤如下:

输入:人脸图像 A ,隐私参数 $\varepsilon, k = \text{rank}(A)$, 随机梯度下降迭代次数 r ,学习速率参数 γ ,正则化因子参数 τ 。

输出:满足 ε -差分隐私的人脸图像 $\tilde{A}$ 。

1) $A_{m \times n} = B_{m \times k} L_{k \times n}$

2) $i = 1 : k$

3) for $a_{ij} \in A$

4) $e_{ij} = a_{ij} - b_i(l_j + \text{lap}(\Delta_1 L / \varepsilon)^k)$

5) $b_i \leftarrow b_i + \gamma(e_{ij} \cdot (l_j + \text{lap}(\Delta_1 L / \varepsilon)^k) - \tau \cdot b_i)$

6) $l_j \leftarrow l_j + \gamma(e_{ij} \cdot b_i - \tau \cdot l_j)$

7) $\tilde{l}_j = l_j + \text{lap}(\Delta_1 L / \varepsilon)^k$

8) return $\tilde{A}_{m \times n} = B_{m \times k} \tilde{L}_{k \times n}$

LRA 算法采用随机梯度下降方法通过迭代得到 $\tilde{A}$ 的最优分解 $B_{m \times k}$ 与 $\tilde{L}_{k \times n}$,进而减小 $LE(\tilde{A})$ 。结合 Yuan 等人(2015)提出的奇异值边界理论与式(8),得到定理 1。

定理 1 假设 $\{\lambda_1, \lambda_2, \dots, \lambda_k\}$ 为矩阵 A 对应的奇异值。若 LRA 采用随机梯度下降方法获得最优分解 $B_{m \times k}$ 与 $\tilde{L}_{k \times n}$,则 $LE(\tilde{A})$ 的上界为 $2 \sum_{i=1}^k \lambda_i^2 / \varepsilon^2$ 。

证明:由于 $A_{m \times n} = B_{m \times k} L_{k \times n}$ 分解的不唯一性,假设 $k = n$,此时 $A_{m \times n} = B_{m \times n}, L_{n \times n} = I$ 。因此, $\Delta_1 L = 1$ 。根据矩阵 A 的 Frobenius 范数定义可知,

$\sum_{ij} B_{ij}^2 \times (\Delta_1 L)^2 = \|A\|_F^2$ 。由于 $\|A\|_F^2 = \sum_{i=1}^k \lambda_i^2$ 成立,由低秩分解的 $LE(\tilde{A}) = 2(\sum_{ij} B_{ij}^2 \times (\Delta_1 L)^2) \frac{1}{\varepsilon^2}$

可知,LRA 的误差上界为 $LE(\tilde{A}) = 2 \sum_{i=1}^k \lambda_i^2 / \varepsilon^2$ 。

将 LRA 算法的误差上界与 LAP 的误差进行比较,得到定理 2。

定理 2 $LE_{\text{LRA}}(\tilde{A}) \leq LE_{\text{LAP}}(\tilde{A})$

证明:设 $LE_{\text{LRA}}(\tilde{A})$ 与 $LE_{\text{LAP}}(\tilde{A})$ 分别为 LRA 和 LAP 算法产生的拉普拉斯误差。

由定理 1 可知, $LE_{\text{LRA}}(\tilde{A})$ 的上界为 $2\|A\|_F^2 / \varepsilon^2$,而 $\|A\|_F^2 = \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$,进而 $LE_{\text{LRA}}(\tilde{A})$ 的上界可以重写为 $2 \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 / \varepsilon^2$ 。

由式(6)可知, $LE_{\text{LRA}}(\tilde{A}) = 2mn \left(\frac{\Delta_1 A}{\varepsilon} \right)^2$,而

$$\Delta_1 A = \max_j \sum_{i=1}^m |a_{ij}|, \text{则}$$

$$LE_{\text{LRA}}(\tilde{A}) = 2mn \left(\max_j \sum_{i=1}^m |a_{ij}| \right)^2 / \varepsilon^2$$

根据

$$\sum_{j=1}^n \left(\max_j \sum_{i=1}^m |a_{ij}| \right)^2 \geq \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$$

可知

$$2mn \left(\max_j \sum_{i=1}^m |a_{ij}| \right)^2 / \varepsilon^2 \geq 2 \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2 / \varepsilon^2$$

进而可以证明 $LE_{\text{LRA}}(\tilde{A}) \leq LE_{\text{LAP}}(\tilde{A})$ 。

3.2.2 基于矩阵奇异值分解的发布方法

利用低秩分解可以获得较好的 $LE(\tilde{A})$ 。然而,由于 LRA 算法使用迭代操作求取最终的最优分解,而迭代操作使得人脸图像的信息损失严重,进而导致较低的可用性。为了获得更低的 $LE(\tilde{A})$,本文采用奇异值分解来压缩人脸图像。 A 的奇异值分解为 $A = Q_A^{m \times m} D_A^{m \times n} (P_A^{n \times n})^T$ 。

根据奇异值分解的性质可知矩阵 $D_A^{m \times n}$ 中的奇异值按照从大到小排序,并且减少特别快,进而导致较少的奇异值之和占据了全部奇异值之和的大部分比例。基于此可以选择 r 个奇异值以及对应的左右奇异向量来近似描述矩阵 A ,即

$$A = Q_A^{m \times m} D_A^{m \times n} (P_A^{n \times n})^T \approx Q_A^{m \times r} D_A^{r \times r} (P_A^{r \times n})^T$$

基于此思想,提出 SRA 方法,具体步骤如下:

输入:人脸图像 A ,隐私参数 ε ,启发式参数 r 。

输出:满足 ε -差分隐私的人脸图像 $\tilde{A}$ 。

1) $A = Q_A^{m \times m} D_A^{m \times n} (P_A^{n \times n})^T$ //对图像进行奇异值分解

2) $Q_A^{m \times r} D_A^{r \times r} (P_A^{r \times n})^T \leftarrow A$ //取 D_A 的前 $r \times r$ 个数值

3) for $i = 1 : r$

4) for $j = 1 : r$

5) if $i = j$ then

6) $\tilde{D}_A^{r \times r}(i, j) \leftarrow D_A^{r \times r}(i, j) + \text{lap}(\Delta_1 D_A^{r \times r} / \varepsilon)$

7) return $\tilde{A} = Q_A^{m \times r} \tilde{D}_A^{r \times r} (P_A^{r \times n})^T$

根据步骤 2)可知,对奇异值分解得到的对角阵 $D_A^{m \times n}$ 启发式地取其前 $r \times r$ 个矩阵数值得到矩阵 $D_A^{r \times r}$,添加拉普拉斯噪声得到对角阵的噪声矩阵 $\tilde{D}_A^{r \times r}$,再将其与 $Q_A^{m \times r}$ 和 $(P_A^{r \times n})^T$ 两个正交矩阵进行

矩阵乘积运算, 重构得到满足 ε -差分隐私的人脸图像 $\tilde{A}$, 但在过程中忽视了 $(m-r) \times (n-r)$ 个对角阵 D_A 的数值, 产生了重构误差 $RE(\tilde{A})$, 本文采用平方和误差度量 $RE(\tilde{A})$, 具体为

$$RE(\tilde{A}) = \sum_{i=r+1}^m |Q_A|^2 + \sum_{i=r+1}^m \sum_{j=r+1}^n |D_A|^2 + \sum_{i=r+1}^n |P_A^T|^2 \quad (9)$$

式中

$$Q_A = (q_1, \dots, q_m), q_i = (q_{i1}, \dots, q_{mi})^T$$

$$P_A^T = (p_1, p_2, \dots, p_n)^T, p_i = (p_{i1}, p_{i2}, \dots, p_{in})$$

根据式(9)可知, 由于对角阵 D_A 忽视了 $(m-r) \times (n-r)$ 个矩阵数值, 进而产生重构误差, 其大小为 $\sum_{i=r+1}^m \sum_{j=r+1}^n |D_A|^2 = \sum_{i=r+1}^m \sum_{j=r+1}^n |D_A(i, j)|^2$ 。 Q_A 与 P_A 分别忽视了 $(m-r)$ 列和 $(n-r)$ 行数值, 二者产生的重构误差为

$$\sum_{i=r+1}^m |Q_A|^2 + \sum_{i=r+1}^n |P_A^T|^2 = \sum_{i=r+1}^m |q_i|^2 + \sum_{i=r+1}^n |p_i|^2$$

根据 SRA 算法的步骤可知, 对角阵前 $r \times r$ 个矩阵数值 $D_A^{r \times r}$ 添加拉普拉斯噪音 (步骤6), 因此 SRA 算法产生的噪音误差上界为

$$\begin{aligned} LE(\tilde{A}) &= E(\tilde{A} - A)^2 = \\ &E\left(Q_A^{m \times r} \left(D_A^{r \times r} + \text{lap}\left(\frac{\Delta D_A^{r \times r}}{\varepsilon}\right)\right) \times \right. \\ &\quad \left. (P_A^{r \times n})^T - Q_A^{m \times r} D_A^{r \times r} (P_A^{r \times n})^T\right)^2 = \\ &E\left(Q_A^{m \times r} \text{lap}\left(\frac{\Delta D_A^{r \times r}}{\varepsilon}\right) (P_A^{r \times n})^T\right)^2 = \\ &2 \sum_{ij} (Q_A^{m \times r})_{ij}^2 (\Delta D_A^{r \times r})^2 \sum_{ij} ((P_A^{r \times n})^T)_{ij}^2 / \varepsilon^2 = \\ &2 \|A\|_F^2 / \varepsilon^2 = 2 \sum_{i=1}^r \lambda_i^2 / \varepsilon^2 \end{aligned} \quad (10)$$

式中, $\Delta_1 D_A^{r \times r} = \max_i \sum_{j=1}^r |D_A^{r \times r}(i, j)|$ 。

根据定理2, 可得定理3。

定理3 $LE_{\text{SRA}}(\tilde{A}) \leq LE_{\text{LRA}}(\tilde{A})$ 。

证明: 根据定理2可知, $LE_{\text{LRA}}(\tilde{A})$ 的误差上界为 $2 \sum_{i=1}^k \lambda_i^2 / \varepsilon^2$, 其中, k ($k \leq \min\{m, n\}$) 为 A 的秩。 $LE_{\text{SRA}}(\tilde{A})$ 的误差上界为 $2 \sum_{i=1}^r \lambda_i^2 / \varepsilon^2$ 。 然而启发式选取的 r 远小于 k , 可知 $2 \sum_{i=1}^k \lambda_i^2 / \varepsilon^2 > 2 \sum_{i=1}^r \lambda_i^2 / \varepsilon^2$ 。

进而可知, $LE_{\text{SRA}}(\tilde{A}) \leq LE_{\text{LRA}}(\tilde{A})$ 成立。

根据式(9)和式(10), 可得定理4。

定理4

$$Er_{\text{SRA}}(\tilde{A}) \leq \sqrt{\sum_{i=r+1}^m |Q_A|^2} + \sqrt{\sum_{i=r+1}^m \sum_{j=r+1}^n |D_A|^2} + \sqrt{\sum_{i=r+1}^n |P_A^T|^2} + \sqrt{2} \sum_{i=1}^r \lambda_i / \varepsilon$$

证明: 用均方根误差度量 $RE(\tilde{A})$, 则

$$RE(\tilde{A}) =$$

$$\sqrt{\sum_{i=r+1}^m |Q_A|^2 + \sum_{i=r+1}^m \sum_{j=r+1}^n |D_A|^2 + \sum_{i=r+1}^n |P_A^T|^2} \leq \sqrt{\sum_{i=r+1}^m |Q_A|^2} + \sqrt{\sum_{i=r+1}^m \sum_{j=r+1}^n |D_A|^2} + \sqrt{\sum_{i=r+1}^n |P_A^T|^2}$$

用均方根误差度量 $LE(\tilde{A})$ 的误差上界, 则

$$LE(\tilde{A}) = \sqrt{2 \sum_{i=1}^r \lambda_i^2 / \varepsilon^2} = \sqrt{2} \sqrt{\sum_{i=1}^r \lambda_i^2} / \varepsilon$$

用均方根误差度量 SRA 算法的总体误差

$Er_{\text{SRA}}(\tilde{A})$, 可得

$$\begin{aligned} Er_{\text{SRA}}(\tilde{A}) &= E \|\tilde{A} - A\|_2 = E(\sqrt{\|\tilde{A} - A\|_2^2}) \leq \\ &\sqrt{E(\|\tilde{A} - A\|_2^2)} = \sqrt{(RE(\tilde{A}))^2 + (LE(\tilde{A}))^2} \leq \\ &RE(\tilde{A}) + LE(\tilde{A}) \leq \\ &\sqrt{\sum_{i=r+1}^m |Q_A|^2} + \sqrt{\sum_{i=r+1}^m \sum_{j=r+1}^n |D_A|^2} + \sqrt{\sum_{i=r+1}^n |P_A^T|^2} + \\ &\quad \frac{\sqrt{2} \sqrt{\sum_{i=1}^r \lambda_i^2}}{\varepsilon} \end{aligned}$$

由定理4可知, 当 $\Delta_1 D_A^{r \times r}$ 与 r 值确定后, 即可计算出 $\tilde{A}$ 的发布误差。 其中, r 值的选择至关重要。 r 值越大, $RE(\tilde{A})$ 越小, 而 $LE(\tilde{A})$ 越大; 相反, r 值越小, $LE(\tilde{A})$ 越小, 而 $RE(\tilde{A})$ 越大。 这两种情况均使发布的人脸图像的可用性较差, 因此需要选择合适的 r 值来均衡这两种误差。 SRA 方法启发式地选择 r 值, 通常无法均衡 $RE(\tilde{A})$ 与 $LE(\tilde{A})$ 。 此外, 如果直接基于原始矩阵 A 选择 r 值, 会破坏 ε -差分隐私, 原因是 r 值的选择是数据相关的。 若要基于 A 选择 r 值, 需要满足 ε -差分隐私。 为此, 本文基于指数机制 (McSherry 和 Talwar, 2007) 提出一种 r 值选择的 ESRA 算法, 具体步骤如下:

输入: 人脸图像 A , 隐私参数 ε , A 的秩 k 。

输出: 满足 ε -差分隐私的人脸图像 $\tilde{A}$ 。

- 1) $A = Q_A^{m \times m} D_A^{m \times n} (P_A^{n \times n})^T$, 对图像进行奇异值分解
- 2) $\forall 1 \leq r \leq k$, 计算每个 r 值的打分函数

$U(\mathbf{A}, r)$, 其中, $U(\mathbf{A}, r)$ 满足

$$U(\mathbf{A}, r) = \sqrt{\sum_{i=r+1}^m |\mathbf{Q}_A|^2} + \sqrt{\sum_{i=r+1}^m \sum_{j=r+1}^n |\mathbf{D}_A|^2} + \sqrt{\sum_{i=r+1}^n |\mathbf{P}_A^T|^2} + \frac{\sqrt{2} \sqrt{\sum_{i=1}^r \lambda_i^2}}{\varepsilon}$$

3) 按概率 $Pr(r) \propto \exp\left(-\frac{\varepsilon U(\mathbf{A}, r)}{2\Delta_1 U}\right)$ 选择一个合适的 r 值;

4) $\mathbf{Q}_A^{m \times r} \mathbf{D}_A^{r \times r} (\mathbf{P}_A^{r \times n})^T \leftarrow \mathbf{A}$, 取 $\mathbf{D}_A$ 的前 $r \times r$ 个数值;

5) for $i = 1 : r$

6) for $j = 1 : r$

7) if $i = j$ then

8) $\tilde{\mathbf{D}}_A^{xr}(i, j) \leftarrow \mathbf{D}_A^{xr}(i, j) + \text{lap}(\Delta_1 \mathbf{D}_A^{xr} / \varepsilon)$

9) return $\tilde{\mathbf{A}} = \mathbf{Q}_A^{m \times r} \tilde{\mathbf{D}}_A^{r \times r} (\mathbf{P}_A^{r \times n})^T$

ESRA 算法首先在 $\mathbf{D}_A^{m \times n}$ 矩阵中计算每个 r 值可能的打分函数值(步骤2)), 然后采用指数机制按概率 $Pr(r)$ 挑选合适的 r 值(步骤3)), 并针对挑选的 r 值, 采用拉普拉斯机制对 $\mathbf{D}_A^{r \times r}$ 矩阵中每个对角数值进行噪音扰动(步骤5)—9)), 最后运用矩阵乘积代数运算, 获得 $\tilde{\mathbf{A}}$ 。

ESRA 算法利用指数机制实现 r 值的挑选, 然而指数机制的关键是如何设计高效的打分函数, 打分函数的优劣直接决定着挑选的 r 值的合理性, 而合理的 r 值能够均衡 $RE(\tilde{\mathbf{A}})$ 与 $LE(\tilde{\mathbf{A}})$ 。本文利用定理4中的总体误差来设计打分函数 $U(\mathbf{A}, r)$, $U(\mathbf{A}, r)$ 越大说明总体误差越小, 具体为

$$U(\mathbf{A}, r) = \sqrt{\sum_{i=s+1}^m |\mathbf{Q}_A|^2} + \sqrt{\sum_{i=s+1}^m \sum_{j=s+1}^n |\mathbf{D}_A|^2} + \sqrt{\sum_{i=s+1}^n |\mathbf{P}_A^T|^2} + \frac{\sqrt{2} \sqrt{\sum_{i=1}^r \lambda_i^2}}{\varepsilon} \quad (11)$$

由式(11)与 ESRA 算法的步骤3)可知, 每个 r 值的抽样概率为

$$Pr(r) \propto \exp \times \left(-\varepsilon \left(\sqrt{\sum_{i=r+1}^m |\mathbf{Q}_A|^2} + \sqrt{\sum_{i=r+1}^m \sum_{j=r+1}^n |\mathbf{D}_A|^2} + \sqrt{\sum_{i=r+1}^n |\mathbf{P}_A^T|^2} + \frac{\sqrt{2} \sqrt{\sum_{i=1}^r \lambda_i^2}}{\varepsilon} \right) / 2\Delta_1 U \right) \quad (12)$$

式中, $\Delta_1 U$ 为打分函数 $U(\mathbf{A}, r)$ 的全局敏感度。为了度量 $\Delta_1 U$, 本文结合式(11)给出定理5。

定理5 $\Delta_1 U \leq \Delta_1 \mathbf{Q}_A + 2\Delta_1 \mathbf{D}_A + \Delta_1 \mathbf{P}_A^T$ 。其中, $\Delta_1 \mathbf{Q}_A$, $\Delta_1 \mathbf{D}_A$, $\Delta_1 \mathbf{P}_A^T$ 分别为矩阵 $\mathbf{Q}_A$, $\mathbf{D}_A$, $\mathbf{P}_A^T$ 的最大列范数。

证明: 由式(11)可知, $\Delta_1 U$ 由两部分组成:

$\sqrt{\sum_{i=r+1}^m |\mathbf{Q}_A|^2} + \sqrt{\sum_{i=r+1}^m \sum_{j=r+1}^n |\mathbf{D}_A|^2} + \sqrt{\sum_{i=r+1}^n |\mathbf{P}_A^T|^2}$ 的敏感度 $\Delta_1 U_1$ 和 $\sqrt{2} \sqrt{\sum_{i=1}^r \lambda_i^2} / \varepsilon$ 的敏感度 $\Delta_1 U_2$, 即 $\Delta_1 U = \Delta_1 U_1 + \Delta_1 U_2$ 。

首先, 满足

$$\Delta_1 U_1 = \Delta_1 \left(\sqrt{\sum_{i=r+1}^m |\mathbf{Q}_A|^2} + \sqrt{\sum_{i=r+1}^m \sum_{j=r+1}^n |\mathbf{D}_A|^2} + \sqrt{\sum_{i=r+1}^n |\mathbf{P}_A^T|^2} \right) \leq$$

$$\Delta_1(\|\mathbf{Q}_A\|_2) + \Delta_1(\|\mathbf{D}_A\|_2) + \Delta_1(\|\mathbf{P}_A^T\|_2)$$

放缩上式可得

$$\Delta_1(\|\mathbf{D}_A\|_2) = \max(\|\mathbf{D}_A\|_2 - \|\mathbf{D}'_A\|_2) \leq \max(\|\mathbf{D}_A - \mathbf{D}'_A\|_2) \leq \max(\|\mathbf{D}_A - \mathbf{D}'_A\|_1) \leq \Delta_1 \mathbf{D}_A$$

同理

$$\Delta_1(\|\mathbf{Q}_A\|_2) \leq \Delta_1 \mathbf{Q}_A$$

$$\Delta_1(\|\mathbf{P}_A^T\|_2) \leq \Delta_1 \mathbf{P}_A^T$$

进而推理可得

$$\Delta_1 U_1 \leq \Delta_1 \mathbf{Q}_A + \Delta_1 \mathbf{D}_A + \Delta_1 \mathbf{P}_A^T \quad (13)$$

由 L_1 范数不等式推理可得

$$\Delta_1 U_2 = \Delta_1 \mathbf{D}_A^{r \times r} \leq \Delta_1 \mathbf{D}_A \quad (14)$$

由式(13)和式(14)推理可得

$$\Delta_1 U \leq \Delta_1 \mathbf{Q}_A + 2\Delta_1 \mathbf{D}_A + \Delta_1 \mathbf{P}_A^T$$

成立。

3.2.3 基于矩阵 $\tilde{\mathbf{D}}_A^{r \times r}$ 的求精处理方法

根据 SRA 算法的步骤6)和 ESRA 算法的步骤8)可知, 通过添加拉普拉斯噪音 $\text{lap}(\Delta_1 \mathbf{D}_A^{r \times r} / \varepsilon)$ 扰动 $\mathbf{D}_A^{r \times r}$ 中的奇异值, 进而得到 $\tilde{\mathbf{D}}_A^{r \times r}$ 。设 $\mathbf{D}_A^{r \times r}$ 中的奇异值为 $\{\lambda_1, \lambda_2, \dots, \lambda_r\}$, 且从大到小排序。扰动后噪音矩阵 $\tilde{\mathbf{D}}_A^{r \times r}$ 中的奇异值为 $\{\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_r\}$ 。然而, 由于添加拉普拉斯噪音的原因, $\{\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_r\}$ 中有些值可能出现乱序, 例如, $\{\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\lambda}_3\} = \{10, 14, 9\}$ 。此时 $\tilde{\mathbf{D}}_A^{r \times r}$ 中的奇异值不是按照从大到小排序, 为了满足奇异值的特性, 需要对其进行后置

求精处理,本文对于此乱序现象采取凸优化方法进行处理。

设 $\{\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_r\}$ 为求精处理后的奇异值序列,则相应的凸优化问题可转换化为

$$\min \sum_{i=1}^r (\bar{\lambda}_i - \bar{\lambda}_i)^2$$

$$\bar{\lambda}_i \geq \bar{\lambda}_{i+1}, 1 \leq i \leq r$$

根据定理6可求出上述目标函数的可行解。

定理6 设

$$I_m = \min_{j \in [m, r]} \max_{i \in [1, j]} \text{mean}[i, j]$$

$$N_m = \max_{i \in [1, m]} \min_{j \in [i, r]} \text{mean}[i, j]$$

则在条件 $\bar{\lambda}_i \geq \bar{\lambda}_{i+1}$ 约束下的可行解为

$$\{\bar{\lambda}_1, \bar{\lambda}_2, \dots, \bar{\lambda}_r\} = \{I_1, \dots, I_r\} = \{U_1, \dots, U_r\}$$

其中, $\text{mean}[i, j] = \sum_{h=i}^j \bar{\lambda}_h / j - i + 1$ (Hay 等, 2010)。

例如, $\{\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3\} = \{10, 14, 9\}$, 后置求精后为 $\{\bar{\lambda}_1, \bar{\lambda}_2, \bar{\lambda}_3\} = \{12, 12, 9\}$ 。

综合定理1—6及LAP、LRA、SRA和ESRA算法,可得定理7。

定理7 LAP、LRA、SRA和ESRA算法均满足 ε -差分隐私。

证明:根据拉普拉斯机制(Dwork 等, 2006), LAP、LRA和SRA算法满足 ε -差分隐私。下面根据指数机制(McSherry和Talwar, 2007)和差分隐私序列组合性质(McSherry, 2010)证明ESRA算法满足 ε -差分隐私。

假设ESRA算法步骤3需要的隐私预算为 ε_1 , 则根据指数机制(McSherry和Talwar, 2007), 选择每个 r 值的概率为

$$Pr(r) = \frac{\exp\left(-\frac{\varepsilon_1 U(A, r)}{2\Delta_1 U}\right)}{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A, r')}{2\Delta_1 U}\right)} \quad (15)$$

给定 A 及其近邻 A' , 对任意 r 值 ($r \in [1, k]$), 根据式(15), 可得

$$\frac{Pr(ESRA(A, r))}{Pr(ESRA(A', r))} = \frac{\exp\left(-\frac{\varepsilon_1 U(A, r)}{2\Delta_1 U}\right)}{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A, r')}{2\Delta_1 U}\right)} = \frac{\exp\left(-\frac{\varepsilon_1 U(A, r)}{2\Delta_1 U}\right)}{\exp\left(-\frac{\varepsilon_1 U(A', r)}{2\Delta_1 U}\right)} = \frac{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A', r')}{2\Delta_1 U}\right)}{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A', r')}{2\Delta_1 U}\right)}$$

$$\left(\frac{\exp\left(-\frac{\varepsilon_1 U(A, r)}{2\Delta_1 U}\right)}{\exp\left(-\frac{\varepsilon_1 U(A', r)}{2\Delta_1 U}\right)} \right) \times \left(\frac{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A', r')}{2\Delta_1 U}\right)}{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A, r')}{2\Delta_1 U}\right)} \right) \leq \exp\left(\frac{\varepsilon_1}{2}\right) \times \left(\frac{\sum_{r' \in [1, k]} \exp\left(\frac{\varepsilon_1}{2}\right) \times \exp\left(-\frac{\varepsilon_1 U(A, r')}{2\Delta_1 U}\right)}{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A, r')}{2\Delta_1 U}\right)} \right) \leq \exp\left(\frac{\varepsilon_1}{2}\right) \times \exp\left(\frac{\varepsilon_1}{2}\right) \times \left(\frac{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A, r')}{2\Delta_1 U}\right)}{\sum_{r' \in [1, k]} \exp\left(-\frac{\varepsilon_1 U(A, r')}{2\Delta_1 U}\right)} \right) = \exp(\varepsilon_1)$$

假设ESRA算法步骤5)—9)需要的隐私预算为 ε_2 , 由拉普拉斯机制(Dwork 等, 2006)可知, 步骤5)—9)满足 ε_2 -差分隐私。

因此,根据差分隐私序列组合性质(McSherry, 2010)可知,ESRA算法满足 ε -差分隐私, 其中, $\varepsilon = \varepsilon_1 + \varepsilon_2$ 。

4 实验与结果分析

为了研究算法的可用性,首先在Lena的灰度人脸图像上进行实验(如图1所示)。



图1 Lena原始图像

Fig. 1 Original image of Lena

图2显示了6种算法的发布结果。从图2可以看出,ESRA算法的效果最好。为了验证算法的有效性与可用性,基于真实人脸库,采用支持向量机(support vector machine, SVM) + 主成分分析(principal component analysis, PCA)方法对上述6种方法进行检验,同时为了度量单幅图像的发布精度,

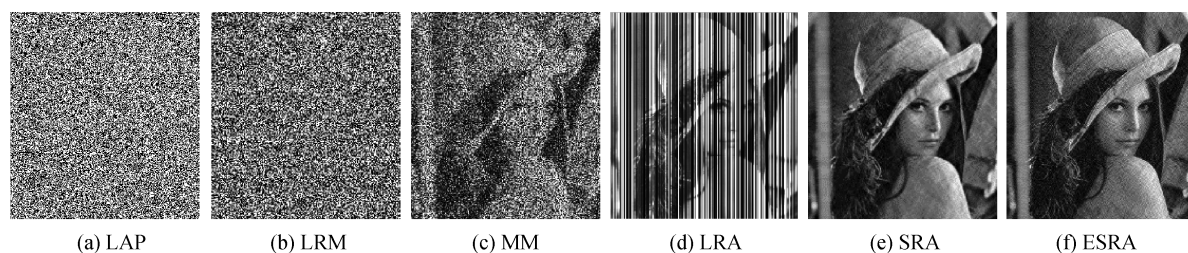


图2 Lena 加噪图像

Fig. 2 Add noise images of Lena ((a) LAP; (b) LRM; (c) MM; (d) LRA; (e) SRA; (f) ESRA)

采用信息熵度量上述 6 种方法处理后的图像信息损失。

4.1 人脸数据库

人脸数据库主要包括 CMU、ORL、Yale、Yale B、Faces95 和 Faces94 人脸库等 6 种数据集,具体细节如表 1 所示。为了采用 SVM + PCA 对上述 6 种数

据集进行有效分类,本文对数据集中每个人拥有的人脸图像数目进行均分,其中前半部分作为训练集,后半部分作为测试集。例如,CMU 中每人有 32 幅图像,前 16 幅图像作为训练集,后 16 幅作为测试集。采用准确率、召回率以及 F1-Score 等 3 种可用性度量标准来衡量上述 6 种算法的可用性。

表 1 6 种数据集信息描述

Table 1 Description of six dataset information

数据集	数据集大小	包含人数	每人拥有图像数	图像大小/像素	像素点数
Yale	150	15	10	231 × 195	45 045
ORL	400	40	10	112 × 92	10 304
CMU	640	20	32	120 × 128	15 360
Yale B	900	30	30	192 × 168	32 256
Faces95	1 440	72	20	200 × 180	36 000
Faces94	3 040	152	20	200 × 180	36 000

4.2 实验结果分析

SVM + PCA 方法首先采用主成分分析法降维并去除数据之间的相关性,其次采用 SVM 训练器对数据集进行训练,最后采用 SVM 产生的分类函数对测试数据集进行分类,并计算分类的正确率(包括准确率、召回率、F1-Score)。 ε 分别取 0.1、0.5、0.9、1.3, $r = 100$, $\gamma = 0.01$, $\tau = 0.1$ 。利用 non-privacy 算法作为参照方法,该方法的实验结果是在原始人脸图像上的识别结果,以此为基准,哪种方法的实验结果越接近 non-privacy 的实验结果,说明哪种方法的效果越好。

4.2.1 基于 ε 变化的可用性对比分析

首先分析隐私预算 ε 变化对上述 6 种方法可用性的影响。实验结果如图 3—图 8 所示。可以看出:1) LRM 算法在 6 种数据集上的可用性最差,因为 LRM 算法采用优化问题寻找矩阵最佳分解时,过

多的迭代操作造成矩阵中的真实值被严重扭曲。2) LAP 算法虽然在原始人脸矩阵中直接添加拉普拉斯噪声,但可用性优于 LRM 算法。3) LRA 算法尽管同样采用迭代操作寻找矩阵的最佳分解,然而在 6 种数据集上,其可用性优于 LRM 和 LAP 算法,因为 LRA 算法基于无约束最优化问题,以较少的迭代次数学习出矩阵的最优分解。例如, $\varepsilon = 0.1$ 时, LRA 算法在 CMU 数据集的准确率是 LAP 算法和 LRM 算法的 1 倍多。4) MM 算法的可用性优于 LRA 算法,因为 MM 算法是直接对原始图像进行噪声处理,不需要经过迭代操作,避免了图像信息的严重损失。5) SRA 算法比 LRA 算法在可用性上有进一步的提高,因为在对图像进行噪声处理之前, SRA 算法采用奇异值分解对图像数据先进行压缩,降低了全局敏感性对噪声大小的影响。例如, $\varepsilon = 0.1$ 时, SRA 算法在 Yale 数据集上的准确率、召回

率和 F1-Score 分别是 LRA 算法的 3 倍多、1 倍多和 2 倍多。6) ESRA 算法在 6 种数据集上的可用性表现最好, 因为 ESRA 算法基于指数机制有效抽取合适的奇异值对原始矩阵进行合理压缩, 有效均衡了拉普拉斯误差与矩阵的重构误差, 提高了人脸图像

的发布精度。例如, $\epsilon = 0.1$ 时, ESRA 算法在数据集 Faces95 上的准确率分别是 LAP、LRM 和 MM 算法的 45 倍、22 倍多和 1 倍多。与其他 5 种算法相比, ESRA 算法对于数据集大小的变化表现相对稳定, 正确性最好。

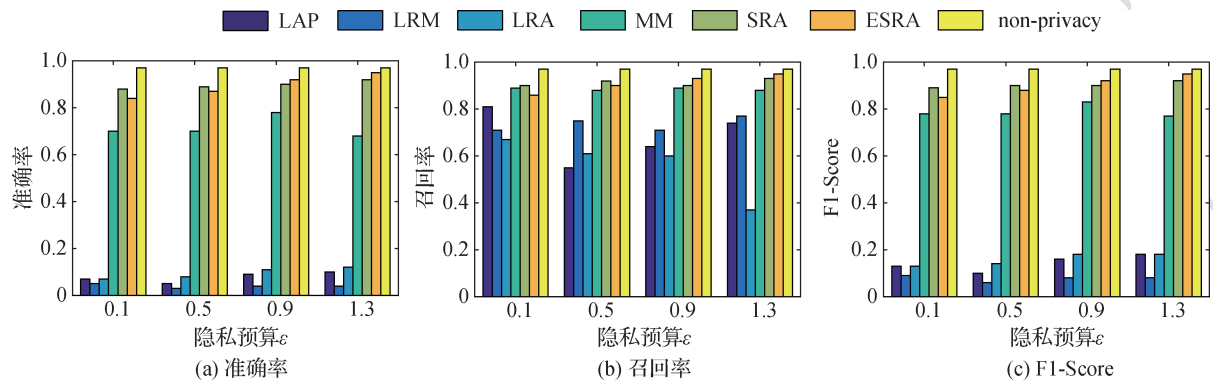


图3 在 CMU 数据集上 ϵ 变化时可用性变化情况

Fig. 3 Usability changes of CMU dataset when vary ϵ ((a) accurate; (b) recall; (c) F1-Score)

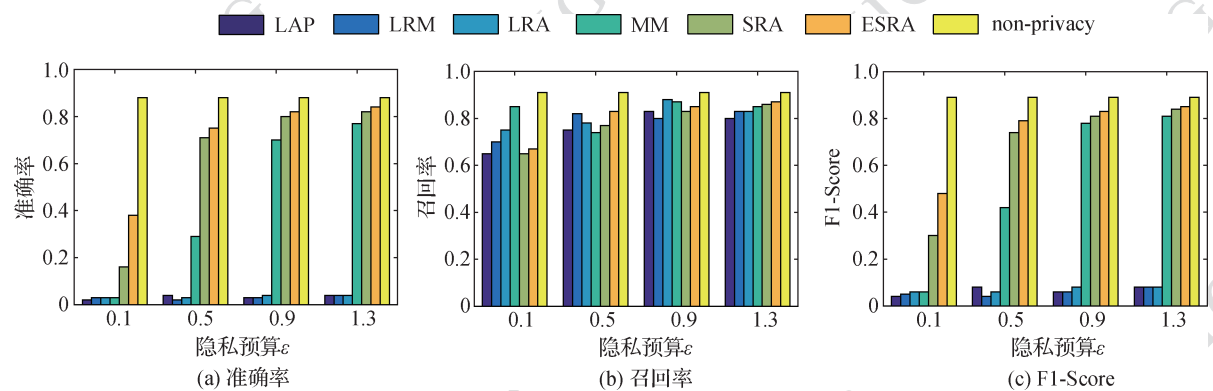


图4 在 ORL 数据集上 ϵ 变化时可用性变化情况

Fig. 4 Usability changes of ORL dataset when vary ϵ ((a) accurate; (b) recall; (c) F1-Score)

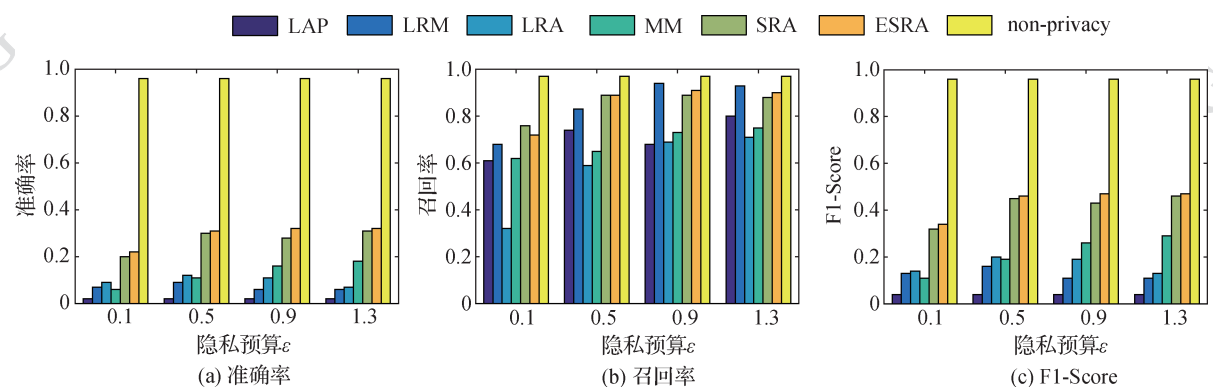
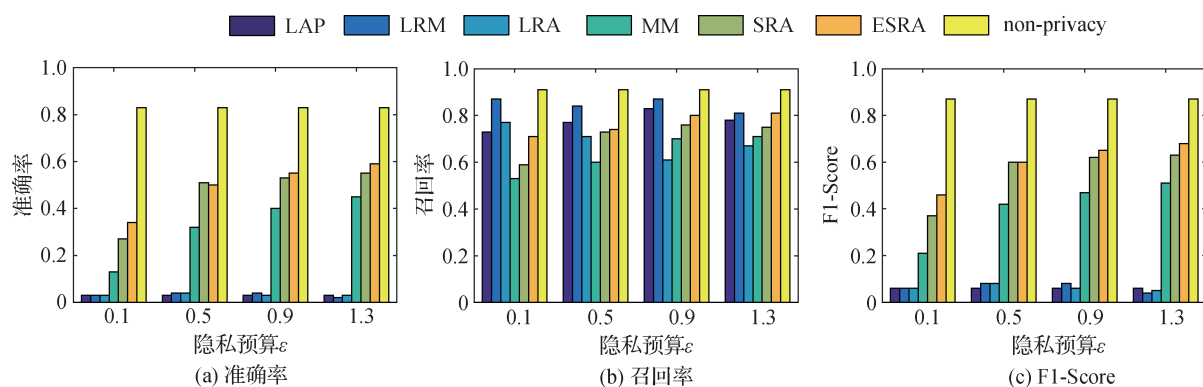
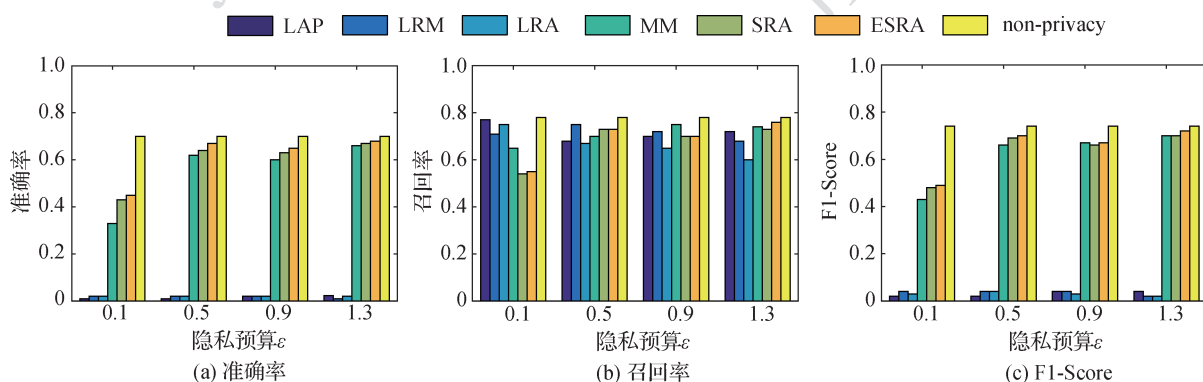
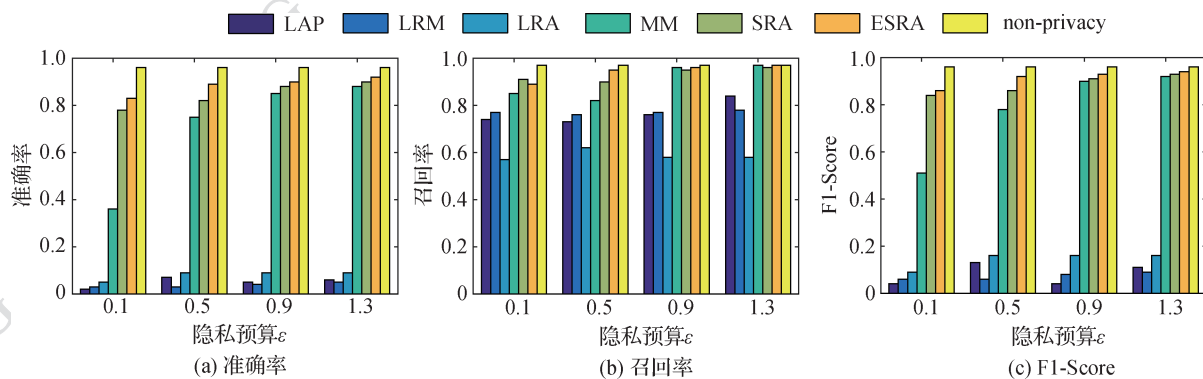


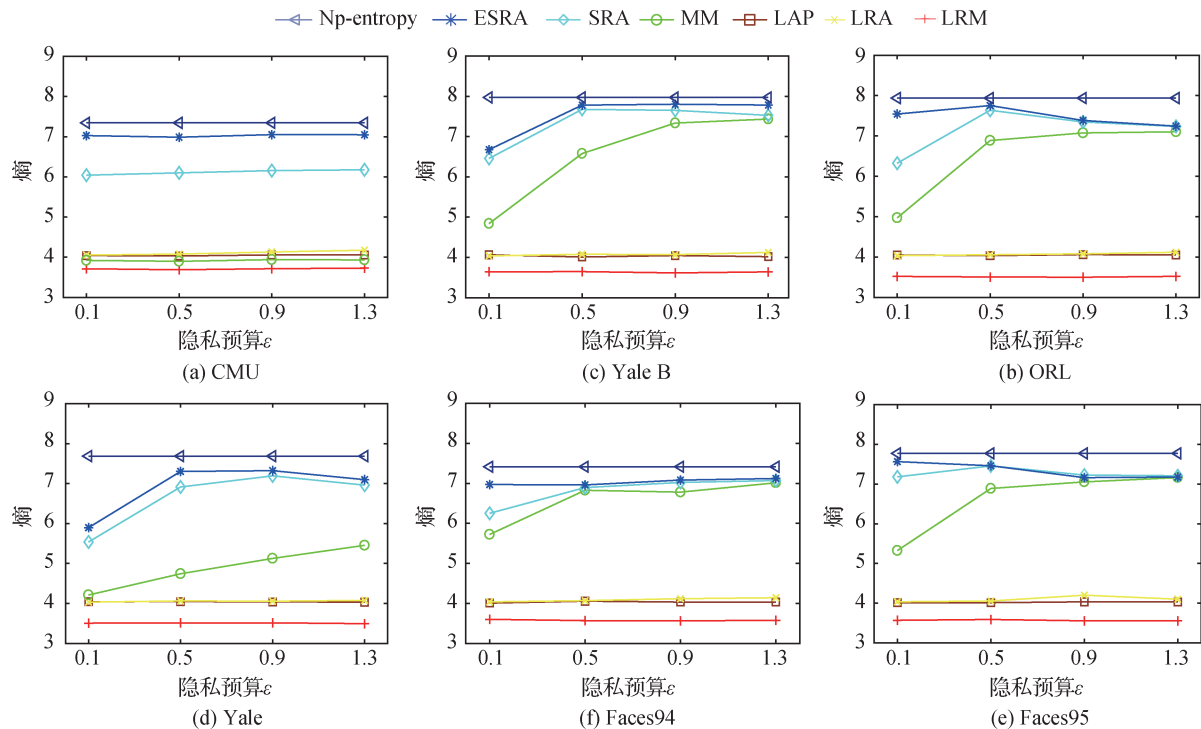
图5 在 Yale 数据集上 ϵ 变化时可用性变化情况

Fig. 5 Usability changes of Yale dataset when vary ϵ ((a) accurate; (b) recall; (c) F1-Score)

图6 在 Yale B 数据集上 ε 变化时可用性变化情况Fig. 6 Usability changes of Yale B dataset when vary ε ((a) accurate; (b) recall; (c) F1-Score)图7 在 Faces95 数据集上 ε 变化时可用性变化情况Fig. 7 Usability changes of Faces95 dataset when vary ε ((a) accurate; (b) recall; (c) F1-Score)图8 在 Faces94 数据集上 ε 变化时可用性变化情况Fig. 8 Usability changes of Faces94 dataset when vary ε ((a) accurate; (b) recall; (c) F1-Score)4.2.2 基于 ε 变化的单幅图像发布精度对比分析

为了有效度量单幅人脸图像蕴含的信息量,采用信息熵 H 作为度量标准, $H = -\sum_{i=0}^{255} p_i \log p_i$, 其中 p_i 表示图像中每个像素点所占的比率。Np-entropy 表示原始图像的信息熵,哪种算法的信息熵越接近 Np-entropy,表示哪种算法发布图像的可用性越高。

在 6 种人脸图像库随机抽取 1 幅人脸图像,分析 ε 变化对单幅图像信息熵的影响,结果如图 9 所示,可以看出,经过 LRM 算法处理过的信息熵最小,原因是 LRM 算法过多的迭代次数导致严重的信息损失; SRA 与 ESRA 算法比较接近 Np-entropy 值,原因是奇异值分解能够充分压缩原始图像,并且能够保留原始图像的主要信息; MM 算法介于 SRA 算法与

图9 6种数据集下 ϵ 变化时单幅图像可用性情况Fig.9 Single image usability changes of six datasets when vary ϵ

((a) CMU; (b) ORL; (c) Yale B; (d) Yale; (e) Faces95; (f) Faces94)

LRA 算法之间,并在 ϵ 大于 0.5 时,逐渐接近 SRA 算法的表现,因为 MM 算法利用矩阵机制处理图像时,无需迭代处理。从 6 种数据集单幅图像的表现来看, SRA 与 ESRA 表现比较好,优于其他 4 种算法。

5 结 论

针对灰度人脸图像发布中的隐私保护问题,首先提出了一种基于矩阵分解的人脸图像发布算法 LRA。LRA 算法将矩阵分解问题当做一种无约束的最优化问题,然后利用正则化矩阵分解,以较少的迭代次数学习出原始矩阵的最优分解。随后分析了 LRA 算法的不足,并提出了两种基于矩阵奇异值分解的人脸图像算法 SRA 与 ESRA。SRA 算法利用启发式设定奇异值的个数,根据设定的奇异值个数压缩原始矩阵。ESRA 算法根据启发式设定奇异值个数的不足,利用指数机制合理地抽取合适的奇异值个数,对原始矩阵进行有效压缩。实验结果表明, LRA、SRA 和 ESRA 算法在达到人脸图像保护效果的同时具有较好的可用性。

今后的研究方向应考虑如下两个方面:1) 如何

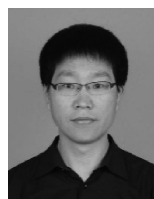
利用树结构(四分树、KD-树等)对人脸图像进行分割,根据分割过程与结果合理地使用噪音机制;2) 由于人脸图像与其图像特征密切相关,如何在满足差分隐私的情况下获取人脸特征,并根据这些隐私人脸特征重构人脸图像,同时应具有较好的分类特性。

参考文献 (References)

- Chen Y, Machanavajjhala A, Hay M and Miklau G. 2017. PeGaSus: data-adaptive differentially private stream processing//Proceedings of 2017 ACM SIGSAC Conference on Computer and Communications Security. Dallas, TX, USA: ACM: 1375-1388 [DOI: 10.1145/3133956.3134102]
- Dwork C. 2006. Differential privacy//Proceedings of the 33rd International Colloquium on Automata, Languages and Programming (ICALP). Venice, Italy: Springer: 1-12 [DOI: 10.1007/11787006]
- Dwork C and Lei J. 2009. Differential privacy and robust statistics//The 41st Annual ACM Symposium on Theory of Computing (STOC). Bethesda, MD, USA: ACM: 371-380 [DOI: 10.1145/1536414.1536466]
- Dwork C, McSherry F, Nissim K and Smith A. 2006. Calibrating noise

- to sensitivity in private data analysis//Proceedings of the 3rd Theory of Cryptography Conference (TCC). New York, USA: Springer: 363-385 [DOI: 10.1007/9783540327.32532]
- Friedman A, Berkovsky S and Kaafer M A. 2016. A differential privacy framework for matrix factorization recommender systems. *User Modeling and User-Adapted Interaction*, 26(5): 425-458 [DOI: 10.1007/s11257-016-9177-7]
- Gross R and Acquisti A. 2005. Information revelation and privacy in on-line social networks//Proceedings of 2005 ACM Workshop on Privacy in the Electronic Society. Alexandria, VA, USA: ACM: 71-80 [DOI: 10.1145/1102199.1102214]
- Gross R, Sweeney L, Cohn J, de la Torre F and Baker S. 2008. Model-based de-identification of facial images//2008 American Medical Informatics Association Annual Symposium. Washington: [s. n.]: 262-273
- Hay M, Rastogi V, Miklau G and Suciu D. 2010. Boosting the accuracy of differentially private histograms through consistency. *Proceedings of the VLDB Endowment*, 3(1/2): 1021-1032 [DOI: 10.14778/1920841.1920970]
- Ilija P, Polakis I, Athanasopoulos E, Maggi F and Ioannidis S. 2015. Face/off: preventing privacy leakage from photos in social networks//Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security. Denver, Colorado, USA: ACM: 781-792 [DOI: 10.1145/2810103.2813603]
- Li C, Miklau G, Hay M, McGregor A and Rastogi V. 2015. The matrix mechanism: optimizing linear counting queries under differential privacy. *The VLDB Journal*, 24(6): 757-781 [DOI: 10.1007/s00778-015-0398-x]
- McSherry F. 2010. Privacy integrated queries: an extensible platform for privacy-preserving data analysis. *Communications of the ACM*, 53(9): 89-97 [DOI: 10.1145/1810891.1810916]
- McSherry F and Talwar K. 2007. Mechanism design via differential privacy//Proceedings of the 48th Annual IEEE Symposium on Foundations of Computer Science. Providence, RI, USA: IEEE: 94-103 [DOI: 10.1109/FOCS.2007.66]
- Newton E M, Sweeney L and Malin B. 2005. Preserving privacy by de-identifying face images. *IEEE Transactions on Knowledge and Data Engineering*, 17(2): 232-243 [DOI: 10.1109/TKDE.2005.32]
- Qardaji W, Yang W N and Li N H. 2013. Differentially private grids for geospatial data//Proceedings of the 29th IEEE International Conference on Data Engineering (ICDE). Brisbane, QLD, Australia: IEEE: 32-33 [DOI: 10.1109/ICDE.2013.6544872]
- Rastogi V and Nath S. 2010. Differentially private aggregation of distributed time-series with transformation and encryption//Proceedings of 2010 ACM SIGMOD International Conference on Management of data. Indianapolis, Indiana, USA: ACM: 735-746 [DOI: 10.1145/1807167.1807247]
- Sadeghi A R, Schneider T and Wehrenberg I. 2009. Efficient privacy-preserving face recognition//Proceedings of the 12th International Conference on Information Security and Cryptology. Seoul, Korea: Springer: 229-244 [DOI: 10.1007/978-3-642-14423-3_16]
- Su S, Tang P, Cheng X, Chen R and Wu Z Q. 2016. Differentially private multi-party high-dimensional data publishing//Proceedings of the 32nd IEEE International Conference on Data Engineering. Helsinki, Finland: IEEE: 205-216 [DOI: 10.1109/ICDE.2016.7498241]
- Xiao X K, Wang G Z and Gehrke J. 2011. Differential privacy via wavelet transforms. *IEEE Transactions on Knowledge and Data Engineering*, 23(8): 1200-1214 [DOI: 10.1109/TKDE.2010.247]
- Xu J, Zhang Z J, Xiao X K, Yang Y, Yu G and Winslett M. 2013. Differentially private histogram publication. *The VLDB Journal*, 22(6): 797-822 [DOI: 10.1007/s00778-013-0309-y]
- Yuan G Z, Yang Y, Zhang Z J and Hao Z F. 2016. Convex optimization for linear query processing under approximate differential privacy//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD). San Francisco, CA, USA: ACM: 2005-2014 [DOI: 10.1145/2939672.2939818]
- Yuan G Z, Zhang Z J, Winslett M, Xiao X K, Yang Y and Hao Z F. 2015. Optimizing batch linear queries under exact and approximate differential privacy. *ACM Transactions on Database Systems*, 40(2): 1-47 [DOI: 10.1145/2699501]
- Zhang J, Cormode G, Procopiuc C M, Srivastava D and Xiao X K. 2015. Private release of graph statistics using ladder functions//Proceedings of 2015 ACM SIGMOD International Conference on Management of Data. Melbourne, Victoria, Australia: ACM: 731-745 [DOI: 10.1145/2723372.2737785]

作者简介



张啸剑, 1980年生, 男, 副教授, 硕士生导师, 主要研究方向为差分隐私、数据库。

E-mail: xjzhang82@ruc.edu.cn



孟小峰, 通信作者, 男, 教授, 博士生导师, 主要研究方向为数据融合与知识融合、面向新型硬件大数据管理、大数据隐私管理、大数据分析、交叉性研究。

E-mail: xfmeng@ruc.edu.cn

付聪聪, 女, 硕士研究生, 主要研究方向为差分隐私、图像处理。E-mail: 1425055812@qq.com