

热点追踪

基于对比关系网络的图像补全

中科院自动化研究所 周晓强 赫然

图像补全任务的难点之一在于对补全结果结构合理性和纹理协调性的要求。在本文中，我们提出了一个两阶段的补全网络来解决此问题。我们首先利用分割重建网络预测残缺图像的完整语义分割图，之后在第二阶段通过图像生成器恢复细粒度的图像细节。在第二阶段的图像生成器中，我们新提出了一种基于图的关系网络，对残缺图像中存在的关系进行建模。此外，为了促使关系网络中的图网络部分学习到更好的节点特征表示，我们提出了对比损失函数，对图网络中的节点特征表示进行优化。相关成果获 ICPR 2020 最佳科学论文奖。

图像补全是指对图像中缺失的区域进行内容补全，要求补全后的图像看起来自然逼真。图像补全技术在图像恢复、图像编辑等领域具有广泛应用。传统图像补全算法^[1,2]使用残缺图片背景区域中的像素或图像块对缺失区域进行补全。当图片残缺区域较大，或待补全内容具有较复杂的语义结构时，这类传统算法的补全结果往往有明显瑕疵。

近年来蓬勃发展的生成对抗网络给图像补全带来了新的解决思路。研究人员利用卷积神经网络和生成对抗网络强大的表征学习和图像重建能力，提出了若干基



图 1 图像补全示意图

于生成对抗网络的图像补全算法^[3,4,5]。然而，如何同时生成结构合理且纹理逼真自然的补全结果，仍是图像补全任务面临的一大挑战。为了解决上述问题，我们提出了基于对比关系网络的图像补全算法。整体框架包含两个阶段：分割重建网络和图像生成网络。补全过程示意图如图 1 所示，整体网络框架图如图 2 所示。

我们首先利用分割重建网络预测残缺图像的完整语义分割图，之后在第二阶段通过图像生成器恢复细粒

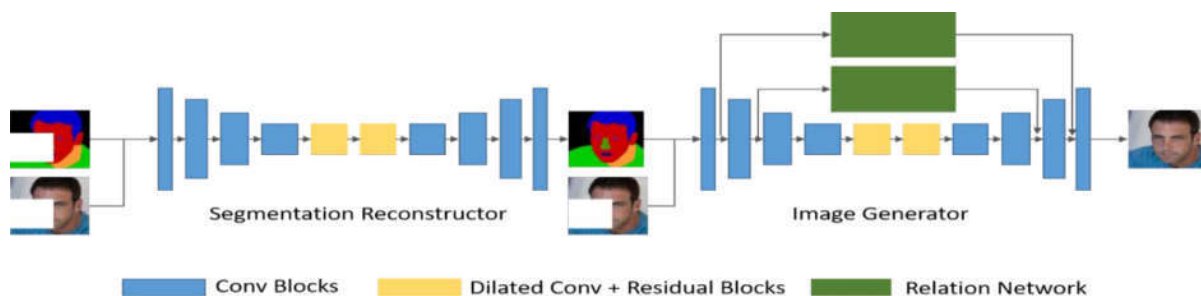


图 2 整体网络框架

度的图像细节。第一阶段中补全的完整语义分割图会对第二阶段的图像细节生成提供指导。

在第二阶段的图像生成器中，我们新提出了一种基于图的关系网络，来对残缺图像中存在的关系进行建模。关系网络的具体结构如图 3 所示。

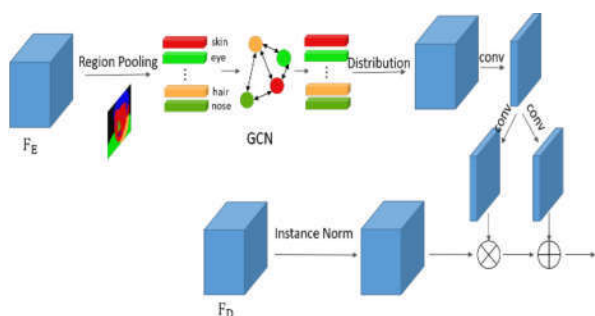


图 3 关系网络

在关系网络中，我们既考虑了同一语义区域内像素之间的内部关系，也考虑了不同语义部分之间的相互关系，以此来提高图像纹理的一致性和协调性。此外，为了促使关系网络中的图网络部分学习到更好的节点特

征表示，我们提出了对比损失函数。对比损失函数通过扩大负样本对之间的特征差异，同时提高正样本对的特征相似度，来对图网络中的节点特征表示进行优化。

为了测试算法的有效性，我们在 CelebA-HQ 和 DeepFashion 数据集上分别进行图像补全实验。定量结果和定性结果对比如表 1 和图 4 所示。与最新相关算法相比，提出的方法不仅在定量实验中取得了更好的性能，同时在定性分析实验中实现了对残缺图像更逼真、自然的补全结果。

表 1 不同算法定量指标对比

		PSNR	SSIM	FID
CelebA-HQ	GMCNN ^[3]	27.35	0.8839	8.81
	EdgeConnect ^[4]	26.60	0.8724	8.38
	GatedConv ^[5]	27.12	0.8814	7.74
	StructureFlow ^[6]	27.48	0.8885	7.21
	Ours	28.61	0.9013	5.94
DeepFashion	GMCNN ^[3]	22.40	0.8074	9.41
	EdgeConnect ^[4]	22.79	0.8162	10.56
	GatedConv ^[5]	23.32	0.8175	8.24
	StructureFlow ^[6]	23.20	0.8166	8.00
	Ours	24.16	0.8197	7.37



图 4 不同算法补全结果对比

责任编辑 崔海楠

参考文献

- [1] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, “Image inpainting,” in Proceedings of the Conference on Computer Graphics and Interactive Techniques, pp. 417–424, 2000.
- [2] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, “Patchmatch: A randomized correspondence algorithm for structural image editing,” ACM Transactions on Graphics, vol. 28, no. 3, pp. 24–33, 2009.
- [3] Y. Wang, X. Tao, X. Qi, X. Shen, and J. Jia, “Image inpainting via generative multi-column convolutional neural networks,” in Advances in Neural Information Processing Systems, pp. 331–340, 2018.
- [4] K. Nazeri, E. Ng, T. Joseph, F. Z. Qureshi, and M. Ebrahimi, “Edgeconnect: Generative image inpainting with adversarial edge learning,” arXiv preprint arXiv:1901.00212, 2019.
- [5] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, “Free-form image inpainting with gated convolution,” in Proceedings of the IEEE International Conference on Computer Vision, pp. 4471–4480, 2019.
- [6] Y. Ren, X. Yu, R. Zhang, T. H. Li, S. Liu, and G. Li, “Structureflow: Image inpainting via structure-aware appearance flow,” in Proceedings of the IEEE International Conference on Computer Vision, pp. 181–190, 2019.



周晓强

中科院自动化研究所博士生，导师为谭铁牛院士。主要研究方向为计算机视觉、图像生成。
Email: xiaoqiang.zhou@cripac.ia.ac.cn



赫然

中科院自动化研究所研究员，博导。主要研究方向为计算机视觉和模式识别。
Email: ran.he@ia.ac.cn

热点追踪

虚拟服装试穿前沿

台湾阳明交通大学电子研究所 陈婕云 帅宏翰 郑文皇

时尚与我们的生活密不可分，例如：妆容以及穿搭，都充分展现个人特质。随着深度学习与计算机视觉的发展，有许多前瞻性的技术能够帮助我们找到适合自己的时尚风格，例如：虚拟上妆、虚拟服装试穿、服装搭配推荐等。本文聚焦于介绍虚拟服装试穿的技术发展，效果展示如图1。试想我们购买新衣服不外乎两种方式：(i) 上街亲自去试穿购买，以及 (ii) 上网看模特的试穿效果购买。上街买衣服可以很直接地判断新衣服适不适合自己，但过程（逛街->挑选到喜欢的衣服->进一步试穿确认是自己喜欢的版型）却非常耗费时间。而近年逐渐流行网购，却存在许多问题，例如：商品的实际色泽与网络商品图不同；衣服穿在模特身上的感受度，和在消费者身上感受度不同；网络上的商品图无法呈现不同姿势角度的试穿效果；衣服的尺寸不合适等。因为这些问题导致网购的退货率极高，这一来一往对于消费者以及店商，都是成本。

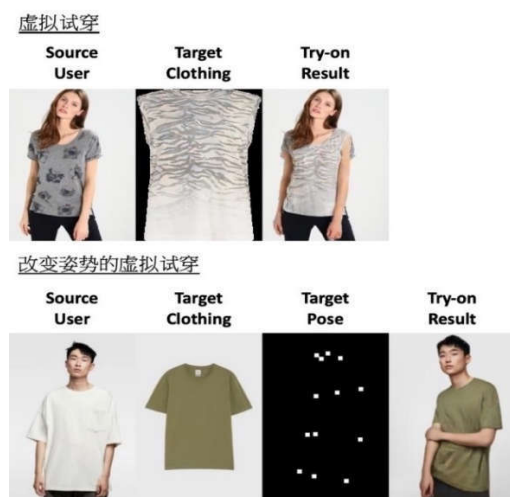


图1 虚拟试穿效果展示

因此我们基于深度学习中生成对抗网络 (Generative Adversarial Network, GAN) 的概念，设计了多姿势虚拟试穿模型^[1]，如图2所示。不仅达到虚拟试穿的效果，更提供多姿势角度的呈现效果。我们基于渐进式学习，将模型分为三个阶段，(i) 姿势引导语义分割转换 (Pose-guided Parsing Translator)，(ii) 语义分割区块上色 (Segmentation Region Coloring)，以及 (iii) 重点区块优化 (Salient Region Refinement)。在阶段 (i)，我们先将要试穿的使用者照片，简化为语义分割图，再基于 ResNet，根据要试穿的衣服掩膜 (mask) 以及目标姿势，来生成变换姿势且试穿后的人体语义分割结果。接着在阶段 (ii)，借由前一个阶段生成的语义分割图，来引导真实色彩的上色区域，帮助模型认识人体的不同部位，例如：脸、头发、手臂、衣服区块等。此阶段使用了条件式生成对抗网络 (cGAN)，特别设计了判别器 (Discriminator) 的目标，并非单纯以生成的图片以及其对应的 groundtruth 作为目标，我们将目标设计为成对的形式：生成的图片 - 真实图片，groundtruth - 真实图片。这样的设计能够帮助生成器不仅生得像 groundtruth，也生得像真实的图片。最后，在阶段 (iii)，我们针对试穿衣服时，消费者最在乎的两个区块：脸以及衣服，分别使用 cGAN 以及 UNet 来进行优化，生成更多的细节，例如：发丝、五官轮廓、衣服皱褶等。优化脸部的 cGAN 架构。与阶段 (ii) 使用的架构相似，我们将阶段 (ii) 的 cGAN 架构去掉全连接层，以保存更多细节，用于优化脸部。此方法^[1]发表于 ACM Multimedia 会议上，并由下方链结观看更多结果。

<https://github.com/fashion-on/FashionOn.github.io>

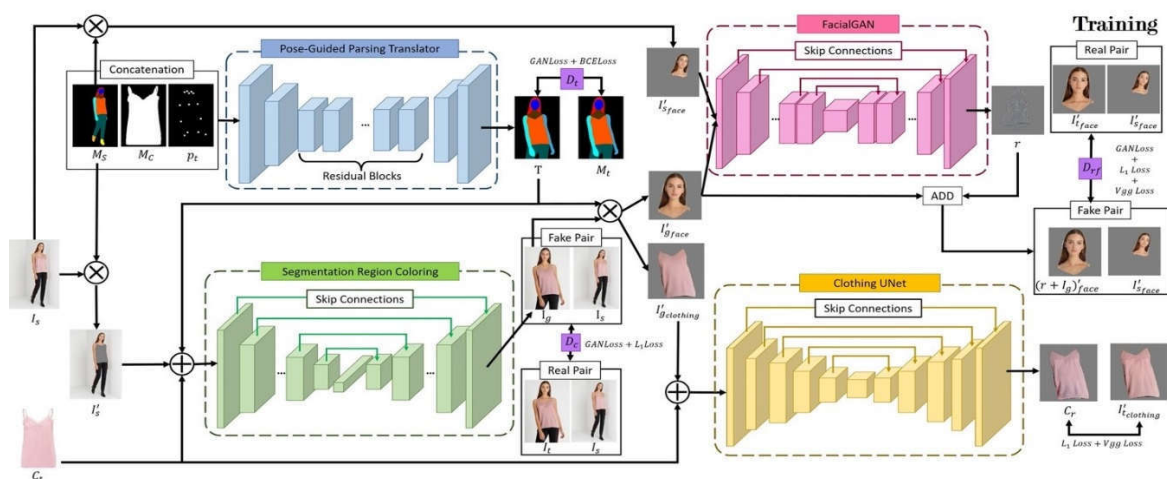


图2 多姿势虚拟试穿架构图

由于衣服多样性非常复杂，例如：有没有衣领、有没有图案、有没有钮扣等诸多细节，我们强化方法^[1]的优化模块，针对有穿着服饰的区域进行更细致的优化，使得细节保存更加拟真，并以此方法获得2020 CVPR Look Into Person Challenge冠军（多姿势虚拟服装试穿任务）。由于改变姿势会使得原本的下半身信息损失，因此针对下半身的裤子/裙子做优化，同时基于注意力机制（attention）改进衣服优化模块。下身优化使用两个编码器（encoder），以及一个解码器（decoder），分别对转换姿势前的下身以及转换姿势后的下身抽取特征，藉由转换姿势前的下身信息，来强化转换姿势过程中丢失的信息，将两者融合后，解码出优化后的下身。上衣使用三个编码器，分别对下列三种衣服图片抽取特征值：(i)目标试穿服饰-代表全局信息、(ii)转换姿势后的衣服区块-为主要的信息来源，以及(iii)前两者与转换姿势前的衣服区块的串接（concatenation）-这边会侦测转换姿势前到转换姿势后的衣服信息间的光流（flow），并基于(ii)的信息，使用注意力机制来将(i)全局的信息以及(iii)姿势间衣服的流动信息，将细节生得更丰富。效果如图3。更详细的方法细节请参考下方公开视频。

<https://www.bilibili.com/video/BV1f54y1B7VJ/>

除此之外，我们发现特定的姿势，能够更好地呈现特定衣服的优点，例如：素色的T-shirt，搭配侧身的姿势，就可以呈现肌肉线条，效果如图4所示。因此我们延伸^[1]，分析衣服特征与姿势间的关联性，基于衣服来生成适合的姿势，并进一步进行虚拟试穿。同时，我们设计一个新的架构来优化衣服，使用两个编码器，分别对

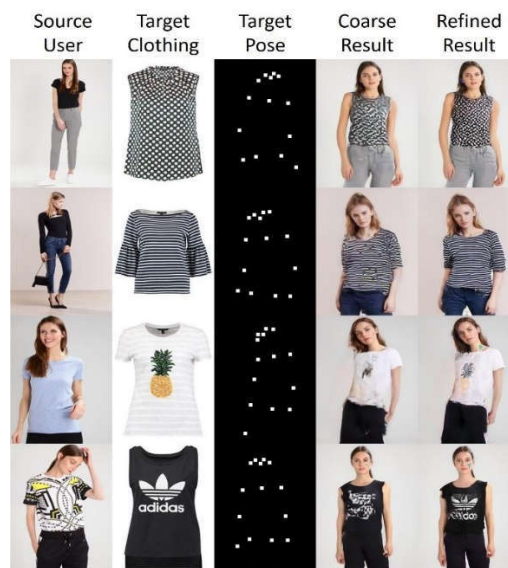


图3 2020 CVPR LIP Challenge 冠军方法试穿效果图

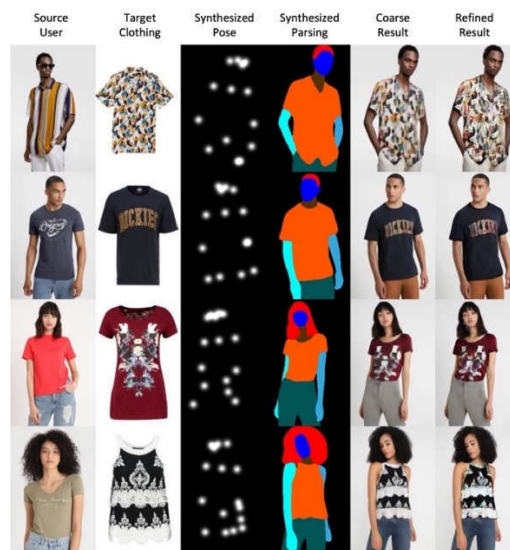


图4 根据衣服生成适当姿势的虚拟试穿效果展示

输入的衣服以及阶段 (ii)生成的衣服区块抽取特征, 再通过解码器, 生成优化后的衣服, 接着使用全局 (global) 以及区域性 (local) 判别器, 不仅判断生得像不像衣服, 更关注细节, 帮助生成细节含量更丰富的衣服。相关方法^[2]已发表至 IEEE Transactions on Neural Networks and Learning Systems (2021)。

针对时尚与计算机视觉的全面性发展, 我们与北京大学王选计算机研究所刘家瑛老师团队合作, 整理了近两百篇相关的论文、对应的数据集以及开源代码, 发表在 ACM Computing Surveys (2021)^[3]。

责任编辑 魏秀参

参考文献

- [1] C.-W. Hsieh, C.-Y. Chen, C.-L. Chou, H.-H. Shuai, J. Liu, and W.-H. Cheng, "FashionOn: Semantic-guided Image-based Virtual Try-on with Detailed Human and Clothing Information," The 27th ACM International Conference on Multimedia (MM 2019). (Oral, acceptance rate=26.5%).
- [2] C.-L. Chou, C.-Y. Chen, C.-W. Hsieh, H.-H. Shuai, J. Liu, and W.-H. Cheng, "Template-Free Try-on Image Synthesis via Semantic-guided Optimization," To appear in IEEE Transactions on Neural Networks and Learning Systems (TNNLS), 2021.
- [3] W.-H. Cheng, S. Song, C.-Y. Chen, S. C. Hidayati, and J. Liu, "Fashion Meets Computer Vision: A Survey," To appear in ACM Computing Surveys (CSUR), 2021.



陈婕云

台湾阳明交通大学电子研究所硕士生。主要从事计算机视觉、虚拟服装试穿相关研究, 获得 2020 CVPR LIP Challenge 冠军(多姿态虚拟服装试穿任务)。
Email: cycchen.ee09g@nctu.edu.tw



帅宏翰

台湾阳明交通大学电机系副教授。主要从事机器学习、多媒体处理、计算机视觉相关研究, 指导学生获得 2020 CVPR LIP Challenge 冠军(多姿态虚拟服装试穿任务)。
Email: hhshuai@nctu.edu.tw



郑文皇

台湾阳明交通大学电子研究所特聘教授。主要从事多媒体信息处理、计算机视觉、人工智能相关研究, 指导学生获得 2020 CVPR LIP Challenge 冠军(多姿态虚拟服装试穿任务)。中国计算机学会杰出会员, 并担任计算机视觉专委会委员。
Email: whcheng@nctu.edu.tw

热点追踪

非接触可撤销掌纹识别研究进展

南昌航空大学 冷璐 南昌大学 闵卫东

一、引言

掌纹 (Palmprint) 识别利用手掌的线、点和纹理等特征鉴别身份, 具有多方面优势, 包括丰富的鉴别信息 (区域面积大)、较少的限制条件 (低分辨率和少量磨损情况下仍然可保持较高的识别精度)、较高的用户接受度、较强的隐蔽性 (不易泄露) 等优势。因此, 掌纹迅速发展成为生物特征家族中的后起之秀^[1]。

1.1. 非接触掌纹识别的意义和技术挑战

接触式掌纹识别时, 用户手掌需要接触采集装置表面, 导致以下问题: 大量用户触摸装置表面, 容易导致疾病传染, 疫情时期更为明显; 容易导致采集设备表面污染, 影响后续样本采集质量; 生物特征可能残留装置表面, 被不法分子获取, 增加信息泄露风险; 部分传统国家或民族禁止异性用户接触同一物体表面; 自由性和灵活性较低, 降低用户舒适度。

因此, 需要非接触式 (Contactless) 掌纹识别技术。然而, 非接触掌纹识别具有背景复杂多变、光照变化明显、手掌姿态和位置难以控制等挑战, 必须设计解决这些严峻挑战的实用化技术方案。

1.2. 生物特征模板保护的迫切需求

已有生物特征识别技术主要关注识别精度, 未充分考虑安全性以及隐私性。生物特征具有不变性, 不能像密码那样, 为了保证安全性而定期灵活更改。生物特征具有唯一性, 若多个数据库保存相同的生物特征模板, 可能导致跨数据库攻击。若某一数据库中的生物特征模板泄露, 则其他存有相同生物特征模板的数据库也会受

到威胁。生物特征还可能含有疾病、基因缺陷等信息, 直接存储会造成严重的隐私泄露风险。

二、研究现状

2.1. 非接触掌纹识别

本部分介绍非接触掌纹识别中的四项技术, 包括辅助拍摄、特征级优选及融合、数据压缩降维、下采样。

(1) 掌纹辅助拍摄技术

完全自动化的识别过程虽然可以提高用户的体验感和舒适度, 但是对算法精度、复杂度等提出了较高要求。在拍摄预览界面绘制辅助图形, 引导用户正确摆放手部位置和姿态, 从而在不明显降低用户舒适度的同时, 降低算法复杂度, 提高抵御各类干扰的鲁棒性以及复杂环境的识别精度^[2]。

(2) 鉴别能量分析及多源空间拓展

通过鉴别能量分析 (Discrimination Power Analysis, DPA) 可以对每维特征进行鉴别性统计和排序, 优选高鉴别性的特征。特征优选不仅可以提高精度, 还降低了模板保存和匹配计算时的数据量, 从而降低了原始特征的隐私泄露。

以左右手的掌纹作为两个实例, 构建双源特征空间, 并在此空间进行动态加权 DPA, 实现多实例掌纹的特征级融合^[3]。由于多源空间比单源空间具有更大的特征搜索范围, 可选择更多鉴别性分量, 提升了识别精度。此技术解决了特征级融合的融合源之间的特征尺寸和类型的不一致问题, 可实现不同维度和异质型特征的特征级融合和特征优选, 具有广泛通用性。

(3) 基于二维双向随机投影的数据压缩降维

随机投影 (Random Projection, RP) 通过随机矩阵直接对原始高维数据进行压缩降维。与传统子空间算法相比, RP 可以完全无需训练, 具有极高的速度。通过构造稀疏 (Sparse) RP、二维 (2D) RP、双向二维 (Two-directional 2D) RP, 可明显降低 RP 的计算量和存储开销^[4]。此技术可广泛应用于包括掌纹识别在内的各类数据压缩和降维场景。

(4) 掌纹编码中的下采样策略

掌纹纹理特征编码可以不需要训练, 存储开销低, 匹配速度快。传统编码直接保留每个尺寸为 4×4 的块内左上角的点。然而每个块内左上角的点不一定能准确表达整个块的信息, 并且对平移错位等干扰比较敏感。为此, 提出极限下采样保留每个块内具有最明显鉴别性特征的点^[5]; 提出民主投票下采样通过块内所有的点参与投票, 共同决策下采样结果^[6]。

2.2. 掌纹特征模板保护

通过特定双因变换, 将原始生物特征转变为扭曲或映射的可撤销形式。若可撤销模板受到威胁, 重新调整控制扭曲或映射的可变因子, 生成新的可撤销生物特征, 保证受保护模板具备替换和重新发布能力。两个因子缺一不可, 和单因认证相比, 具有更高的安全性^[7]。



图1 可撤销生物特征双因认证框架

构造了多样化、可撤销、可更新、可复用的掌纹特征模板, 包括二维掌纹哈希码 (2DPalmHash Code, 2DPHC)^[8]、二维掌纹相位码 (2DPalmPhasor Code, 2DPPC)^[9], 避免了跨数据库的交叉匹配, 保证每个数据库中掌纹模板的安全性和隐私性, 也避免了直接使用原始掌纹特征引发的安全隐患和隐私风险。此框架可推广至其他生物特征模态, 解决推广应用的瓶颈问题。

2.3. 掌纹密码系统

若能够提取可靠稳定的生物特征直接作为密钥, 代替传统密钥并输入现有密码认证系统, 构造生物特征密

码系统 (Biometric Cryptosystem), 可充分发挥生物特征和密码学两种认证方式的双重优势。

然而生物特征比对和密钥比对具有明显差异。生物特征比对具有模糊性, 当注册和认证模板的差异度小于一定阈值, 即认定为合法用户。而传统系统的认证密钥要求绝对精确匹配。

模糊承诺 (Fuzzy Commitment) 和模糊保险箱 (Fuzzy Vault) 是两种典型的生物特征密码系统框架。基于 2DPHC 构造出可撤销掌纹模糊保险箱^[10], 有效抵御多种攻击。鉴于 2DPHC 的水平错位容忍性, 提取免于对齐的特征集合。此外, 将行单生模式扩展为行共生模式, 明显提高了重构精度。

三、未来发展方向

本文论述了非接触可撤销掌纹识别中三项技术的研究进展, 综合考虑了精度、效率、安全、隐私等多方面性能。未来发展方向包括以下几个方面。

(1) 拓展至其他生物特征模态

其他生物特征模态的常见方法通常可以在掌纹识别中使用, 因此掌纹的相关技术框架为其他模态应用提供了重要的参考和借鉴, 理论上可以移植到其他模态。

(2) 完备和严谨的定量化理论

由于生物特征模板保护和生物特征密码系统涉及模式识别、计算机视觉、网络空间安全等多个学科, 需要有机融合并充分借鉴各学科领域的研究方法、思路 and 评价标准, 形成更为完备、严谨的定量化理论。

(3) 远程网络环境的多因子认证

远程网络环境通常面对无处不在的安全威胁以及有限的计算、存储、通信等资源束缚。需要根据具体的安全等级和实用性要求, 设计在精度、效率、安全等多方面综合权衡的多因子 (身份证等物品、密码等知识、掌纹等生物特征) 融合认证算法和远程协议^[11]。

随着生物特征模板保护等研究的不断深入和完善, 生物特征识别技术的应用规模和范围将不断扩大, 为人们提供更为广泛的便捷认证服务。

责任编辑 储璐

参考文献

- [1] Fei L K, Lu G M, Jia W, Teng S H and Zhang D. 2019. Feature extraction methods for palmprint recognition: a survey and evaluation. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 49(2): 346-363 [DOI: 10.1109/TSMC.2018.2795609]
- [2] Leng L, Gao F M, Chen Q and Kim C S. 2018. Palmprint recognition system on mobile devices with double-line-single-point assistance. Personal and Ubiquitous Computing, 22(1): 93-104 [DOI: 10.1007/s00779-017-1105-2]
- [3] Leng L, Li M, Kim C S and Bi X. 2017. Dual-source discrimination power analysis for multi-instance contactless palmprint recognition. Multimedia Tools and Application, 76(1): 333-354 [DOI: 10.1007/s11042-015-3058-7]
- [4] Leng L, Zhang J S, Chen G, Khan M K and Alghathbar K. 2011. Two-directional two-dimensional random projection and its variations for face and palmprint recognition//Proceedings of International Conference Computational Science and Its Applications. Santander, Spain: 458-470 [DOI: 10.1007/978-3-642-21934-4_37]
- [5] Yang Z Y, Leng L and Min W D. 2020. Extreme downsampling and joint feature for coding-based palmprint recognition. IEEE Transactions on Instrumentation & Measurement, 70: 1-12 [DOI: 10.1109/TIM.2020.3038229]
- [6] Leng L, Yang Z Y and Min W D. 2020. Democratic voting downsampling for coding-based palmprint recognition. IET Biometrics, 9(6): 290-296 [DOI: 10.1049/iet-bmt.2020.0106]
- [7] Patel V M, Ratha N K and Chellappa R. 2015. Cancelable biometrics: A review. IEEE Signal Processing Magazine. 32(5): 54-65 [DOI: 10.1109/MSP.2015.2434151]
- [8] Leng L and Zhang J S. 2013. PalmHash Code vs. PalmPhasor Code. Neurocomputing, 108: 1-12 [DOI: 10.1016/j.neucom.2012.08.028]
- [9] 冷璐, 黎明, 张明仁, 金千植. 2015. PalmPhasor 算法性能的理论分析. 软件学报. 26(5): 1237-1250 [DOI: 10.13328/j.cnki.jos.004594]
- [10] Leng L and Teoh A B J. 2015. Alignment-free row-co-occurrence cancelable palmprint Fuzzy Vault. Pattern Recognition, 48(7): 2290-2303 [DOI: 10.1016/j.patcog.2015.01.021]
- [11] Amin R, Islam S K H, Biswas G P, Khan M K, Leng L and Kumar N. 2016. Design of anonymity preserving three-factor authenticated key exchange protocol for wireless sensor network. Computer Networks, 101: 42-62 [DOI: 10.1016/j.comnet.2016.01.006]



冷璐

南昌航空大学副教授。主要研究方向为生物特征识别、生物特征模板保护、计算机视觉。
Email: leng@nchu.edu.cn



闵卫东

南昌大学教授。主要研究方向为生物特征识别、计算机图形图像处理、计算机图形学、云计算分布式系统和智慧城市信息技术。
Email: minweidong@ncu.edu.cn