

不正な認証を防ぐための顔画像の非識別化に関する検討 A study on de-identification of face image for preventing outsiders from authorization

内田秀継* 安部登樹* 山田茂史*
Hidetsugu Uchida Narishige Abe Shigefumi Yamada

キーワード 顔画像、生体認証、セキュリティ、プライバシー、匿名化

あらまし

顔認証は、顔情報から個人を特定する生体認証方式である。近年の深層学習の発展による照合精度の飛躍的な向上を背景に、顔認証は一般的な生体認証として普及が進んでいる。顔認証はその照合精度の高さに加え、利便性の高さ、つまり、一般的なカメラで利用可能、照合のための動作が簡便、といった点も普及の一因となっている。その一方で、顔画像は第三者が容易に入手可能であるため、それを用いた顔認証に対する不正な認証やプライバシーの侵害が問題となる。そこで、顔画像の個人性を加工によって消去することで不正な認証およびプライバシー侵害を避けるための顔画像匿名化技術が検討されている。顔匿名化技術は、顔画像の個人性の識別を困難にすること（非識別化）を目的とした技術であり、非識別化の対象によって、以下の3つに分類できる。

- 1)人の目および顔認証に対して非識別化する[1]、
- 2)人の目のみに対して非識別化する[2]
- 3)顔認証のみに対して非識別化する[3]

1),2)は、どちらも視覚的な個人性を非識別化する技術であるため、匿名化後の画像が人目に触れることによるプライバシーの侵害を防ぐことが出来る。その一方で、画像が持つ視覚的な個人性自体が重要となる場合、例えば、SNSの投稿やビデオ通話など顔画像が自身のものであることが重要である場合などでは適用が難しい。そこで、顔画像が持つ個人性を含めた視覚的な情報を最大に保存しつつ、第三者による不正な認証を防ぐ技術(すなわち、3)の人の目における個人性を保存したまま、顔認証のみに対して非識別化する技術)が注目されている。

視覚的な個人性を保存したまま顔認証での認証を困難化する手法として、adversarial example を利用した方

法が検討されている。Adversarial example を用いた手法では、顔認証用に訓練された深層学習モデルの勾配情報を利用して、識別結果に影響を与えるノイズを画像に付加することで顔認証に対する非識別化を行う。勾配情報はモデル依存性が高いため、非識別化の効果を様々な顔認証モデルに対して汎化させるためには、画像に付加するノイズの強度を上げる必要があり、画像の画質が損なわれる場合がある。そこで、顔画像に含まれる特徴点(目、鼻、口)の配置や形状に摂動を与えることによる非識別化を検討する。特徴点の情報は、様々な顔認証で共通して参照される情報であるため、高い汎化性能が期待できる。また、視覚的な個人性・画質を保存するために、顔画像の認証のしやすさを評価し、それに応じて加工強度を調整する方法についても検討を行った。

提案法の効果を検証するために Labeled Face in the Wild を用いた実験を実施した。実験では、登録画像群に対して、非識別化を行った照合画像をクエリして類似度に基づいた検索を行った場合の検索精度と、非識別化による画像変化の客観評価値(PSNR,SSIM)を評価した。Adversarial example を用いた従来法では、Rank-1 の検索精度が 0.25 (PSNR=29.3,SSIM=0.78)であったのに対して、提案法では、Rank-1 の検索精度が 0.11 (PSNR=31.0,SSIM=0.92)と改善した。

参考文献

- [1] M. Maximov, *et al.*, “CIAGAN: Conditional Identity Anonymization Generative Adversarial Networks”, CVPR2020
- [2] 神津 他, “ステガノグラフィを用いたプライバシ保護顔認証とその安全性評価”, 第 24 回 画像の認識・理解シンポジウム
- [3] X. Yang, *et al.*, “Towards Face Encryption by Generating Adversarial Identity Masks”, CVPR2022

* 富士通株式会社 神奈川県川崎市中原区上小田中 4-1-1, Fujitsu Ltd., 4-1-1 Kamikodanaka, Nakahara-ku, Kawasaki, Kanagawa