

Mob Scene Filter：プライバシー保護のための 容貌変化システムの提案

川瀬 佑司^{1,a)} 白石 英明¹ 吉田 成朗¹ 伊藤 裕之² 斉藤 直樹² 川本 哲也³ 鳴海 拓志⁴
谷川 智洋⁴ 廣瀬 通孝⁴

概要：本研究では、表情・視線情報を損なうことなく映像中の人物の肖像権を保護する、容貌変化システムを提案する。従来では人物の顔をモザイク処理してぼかしているが、人物の表情や視線が読み取れず映像の意図する文脈が損なわれる問題がある。本システムでは、映像中の人物に対して顔の各パーツの形や位置を変えることで他人のように変形する「顔変形フィルタ」と、別人の顔を重畳することで他人のようにみせる「顔重畳フィルタ」のフィルタを用い、表情や視線の情報を残した自然な容貌変化を実現している。

Mob Scene Filter: A proposal for protecting privacy by Changing Facial Appearance

KAWASE YUJI^{1,a)} SHIRAISHI HIDEAKI¹ YOSHIDA SHIGEO¹ ITO HIROYUKI² SAITO NAOKI²
KAWAMOTO TETSUYA³ NARUMI TAKUJI⁴ TANIKAWA TOMOHIRO⁴ HIROSE MICHITAKA⁴

Abstract: Mob Scene Filter (MSF) aims to prevent identification of an individual identity by changing one's facial appearance. Currently, image processing techniques, such as the mosaic and the blur, have been used to prevent identification of a person's identity. These techniques, however, make people difficult to understand the context of the video, because the facial expressions of the characters and the background in the video are masked. So, we propose a system called MSF that changes not only the facial textures but also the arrangement of the facial regions and their size. The system first identifies a face and facial features in a video. Then, the system superimposes another person's face on those identified faces, and transforms the size of facial regions and the arrangement.

1. はじめに

本研究では、表情・視線情報を損なうことなく映像中の人物の肖像権を保護する、容貌変化システムを提案する。映像技術の発達に伴い、放送映像もハイビジョン、さらには 4K、8K と映像の高解像度化が急速に高まっている。一

方で放送映像の高解像度化は、遠くからの撮影で映り込んだ人物の個人特定が可能になるほど高精細化しており、映像に映り込んだ人のプライバシーや肖像権を侵害してしまう危険性がある。このように映像の高精細化に伴い新たなプライバシー問題が生じている。プライバシーを保護し、肖像権侵害を防止するために、撮影者は被写体に対し可能な限り同意を得なければならない。しかし、被写体に対して同意を得ることが困難な場合がある。例えば街中など大勢の通行人が映り込むときは、一人ずつ同意を得ることは非常に困難である。加えて映像の高精細化は、従来の映像では被写体が遠くて個人特定できなかった人物までも、鮮明に映し出してしまい、同意を得ることが困難な被写体のプライバシーを侵害してしまう。

一般的に放送映像などではモザイク処理や黒塗り処理を

¹ 東京大学大学院学際情報学府
Graduate School of Interdisciplinary Information Studies,
The University of Tokyo
² 北海道日興通信株式会社
Nikko Telecommunications Co., LTD.
³ 中京テレビ放送株式会社
Chukyo TV. Broadcasting Co., LTD.
⁴ 東京大学大学院情報理工学研究所
Graduate School of Information Science and Technology,
The University of Tokyo
a) kawase@cyber.t.u-tokyo.ac.jp

施し、個人特定を困難にすることで肖像権の保護を行っている。しかし、同時に被写体の表情を読み取ることも困難にしているため、視聴者が元映像の意図する文脈を誤って解釈してしまう危険性がある。この問題を解決するためには、モザイク処理であればモザイクを薄く、または黒塗り処理であれば範囲を狭くすることが効果的だが、これでは逆に個人特定できる危険性が高まる。すなわち、従来の処理では、肖像権の保護と表情・視線の反映はトレードオフの関係にある。またモザイク処理や黒塗り処理は画像処理を施している部分が非常に明示的であるため、処理部分が目立ち視線が向いてしまうだけでなく、景観や映像コンテンツの品質も損なう。そこで処理部分が非明示的で、肖像権の保護と表情・視線反映の両方を実現する、より進んだ肖像権保護の対策が求められている。

また肖像権は日本国憲法第 13 条幸福追求権や人格権の一環と解釈されており、システムを適用した時に適用者が感じる不快感が強いと、名誉棄損に抵触する恐れがある。つまり、本研究で提案する手法は、肖像権の保護と表情・視線反映の両方を実現することを目指すだけでなく、適用された人に不快感が少ないシステムを構築する必要がある。

そこで本研究は、元映像の表情・視線を反映させて自然に人物の容貌を別人に変化させることにより、処理部分の品質を低下させることなく、群衆の映った場面で個人特定されない手法を提案する。具体的には、肖像権・著作権が発生しない画像（パブリックドメイン画像）として登録された他人の顔を重畳することで、群衆の映った場面（Mob Scene）の匿名化ができると考えた。また、元映像に映り込んだ人物の顔の各パーツの位置の大きさ・位置を変化させることでも、群衆の映った場面の匿名化ができると考えた。

本稿では、本システムを適用したときに個人特定されないものになっているか、処理部分が目立たず自然なものになっているか、システムを適用することによる適用者と鑑賞者の不快感が強くないか、について調査する。

顔画像（図 1）に対して本システムを適用した際の画像処理結果の一例を図 2 に示す。他者の顔を重畳しつつ、目と目の幅を広げる、口を大きく、輪郭を大きく変形する処理を行った。

2. 肖像権侵害と対策

まず本研究における肖像権侵害について定義し、肖像権侵害を防ぐための既存の顔画像処理技術について述べる。

2.1 肖像権侵害

肖像権は、肖像を有する自己の意思に反して肖像を撮影・公表されることを拒絶し、精神的自由を保護する権利を表している [1]。肖像権は、法律による明文の規定はないが、判例で認められている権利であり、判例によって形成されるため、今なお形成されつつある権利である。また、日



図 1: ソース画像

Fig. 1 Source Image.



図 2: MSF 処理後の画像

Fig. 2 Mob scene filtered image.

本国憲法第 13 条幸福追求権や人格権の一環として解釈されている。本研究では、肖像権侵害を「肖像を有する人物の承諾を得ずに映像の公表によって、他者から肖像を有する人物が個人特定され、法律上保護されるべき人格権を侵害されること」と定義する。

2.2 肖像権侵害対策としての顔画像処理

肖像権の対策として、放送映像や週刊誌などにおいてはモザイク処理や、ぼかし処理、黒塗り処理など顔に対して画像処理が適用されることが多い。しかし、前述したように、これらの処理では元人物の表情を読み取ることが難しくなってしまう。

本節では上記とは別の方法で肖像権保護や肖像権保護への利用が可能な顔画像処理手法について事例を挙げる。

D.Bitouk らは、多量に保存されている顔写真データベースを用いて、顔写真を別人の顔に自動で置き換える手法を提案し、静止画において自然な顔の置き換えを可能にした [2]。この手法では顔姿勢推定した画像同士で角度の合うものを選択し、色、ライティング補正したものを重畳に使用していたが、顔の角度が類似したものしか顔を置き換えることができなかった。そこで Y.Lin らは重畳する顔画像の 3D モデルを作成してから、D.Bitouk らの手法を用いることに

よって、顔の角度が異なる画像に対して、顔を置き換えることが可能な手法を提案した [3]。これらの手法は、重畳する顔画像に対して変形処理などは行わず、重畳画像の表情のまま重畳を行う。そのため、元画像・映像の表情は反映されず、元映像の意図する文脈を残せるとは言い難い。

表情を反映しながらも容貌を別人にみせる手法も研究されている。例えば、B.Samarzija らは予め登録しておいた顔画像から表情が近い画像を検索し、顔を置き換えることで個人特定されない手法を提案している [4]。この手法では、予め重畳する人物の様々な表情の顔画像を用意しないと成り立たず、群衆の映った映像に適用するには、相当な手間がかかる。また重畳対象を複数人想定した場合、莫大に画像データが必要になる。そのため群衆の映った場面に適用するには、前処理の手間が非常にかかり、現実的ではない。

一方、前処理や膨大な顔画像データベースを必要としない、顔の形状に合わせて別人の顔を変形して重畳する研究もある。一枚の顔画像を素材とし、顔の形状に合わせて別人の顔を変形して重畳する手法を A.Niswar らが提案している [5]。この手法では少ないリソースで人物の顔を別人に置き換えることを可能にし、表情を反映させることができている。しかし、実際とは異なる形状に変形しており不自然になるという問題がある。

また、Y.Nakashima らは、重畳される顔の特徴点に対応した重み付きマスク画像と、重畳される顔の表情・形状に変形させた重畳画像を用いることで、より自然に表情を反映しつつ別人の顔に置き換える手法を提案している [6]。しかし、表情は反映させられても視線や口の開閉まで反映できない問題がある。

肖像権保護を目的としない顔画像処理の利用法として、中里らはビデオ会議においてクリエイティビティを向上させるために、リアルタイムに他者の容貌を疑似的に変化させる手法を提案した [7]。この手法では動画内の顔を自動検出し顔姿勢を推定し、予め用意した他人の顔画像を動画の顔の姿勢に合わせて重畳している。しかし、Alpha Blend によって顔画像が合成されるため、重畳する人物と重畳元の人物の肌の色が合わない場合は不自然な合成となる。また、瞬きや口の開閉が反映されない問題もある。しかしこの手法は、重畳に使用する画像が 1 枚で良いため、群衆の映った場面では一人ひとり別の処理を施さなければいけない本研究の目的に適している。そこで、この手法に改善を加え本研究で提案する重畳手法を構築する。

他方、以上で述べたような他人の顔画像を重畳する以外にも、顔のパーツを変形させて、別人に見せる手法を考える。吉田らは、リアルタイムに表情を変化させて視覚的にフィードバックすることで、快・不快感情を喚起させる手法を提案した [8]。そして、疑似的に生成した「頬と口角を上げた表情」を視覚的にフィードバックすることによ

って快感情を、「眉をしかめた表情」では不快感情を喚起させることができる可能性を示した。この研究では、Image Deformation[9] と呼ばれる画像処理手法を用いて、体験者の表情画像から自身の表情変化として違和感のない表情画像をリアルタイムに生成することを実現した。本研究では、この表情変化手法を、顔の各パーツの配置や大きさを変化することに応用して、容貌を別人に変化させる方法を構築する。

3. 容貌変化システム：Mob Scene Filter

3.1 容貌変化システムの提案

本研究では他人の顔を重畳するだけでなく、被写体自身の顔パーツの位置や形を変形することで、表情・視線反映と肖像権保護の両立が可能なシステムの構築を目指す。本システム Mob Scene Filter (MSF) は顔変形フィルタ、顔重畳フィルタ、この二つを組み合わせた複合フィルタの 3 つの画像処理フィルタからなる。

従来の顔重畳手法では顔全体を重畳するため、個人特定を防ぐ効果は高いが、表情や視線といった情報を損なう。そこで重畳する顔画像を対象の表情に合わせて変形し重畳する手法が提案されたが、表情は残せるものの視線情報は失い、また個人特定を防ぐ効果も従来と比べると劣ってしまう。そこで、本研究では、表情・目線を残すため、目や口などを除いた顔重畳を行い、さらには顔変形によって顔パーツの布置情報も変化させる。顔重畳、顔変形の両手法を組み合わせることで、人間の顔認知特性に基づいた、匿名性の高い容貌変化システムを構築できるのではないかと考えた。

3.2 顔姿勢推定

MSF 処理を行うためには、まず顔の向きや顔パーツの位置といった顔姿勢の推定が必要となる。顔姿勢推定としてまず、映像から取得した複数人が映った画像から顔の探索を行う。顔領域の検出は、Adaboost と Haar-like 特徴量を利用し、計算が高速かつパターン認識性能の高い処理を実現した Viola-Jones 法を用いた [10]。これにより検出された各顔領域に対して、検出された範囲内にある顔の輪郭や眉、目、鼻、口などの位置情報を取得する。顔の姿勢推定には Face Tracker ライブラリ [11] を使用した。Face Tracker ライブラリは多くのユーザや条件下で使用可能でありロバスト性が高いこと、加えて顔データについて学習済みであり新しい学習を必要としないといった利点があるため使用した。元映像から切り抜いた顔画像に対し、姿勢推定を行った例を図 3 に示す。顔画像の姿勢推定後は、各フィルタで処理の流れが異なる。

3.3 顔変形フィルタ

顔変形フィルタの場合、3.2 の顔姿勢推定で得た位置情報

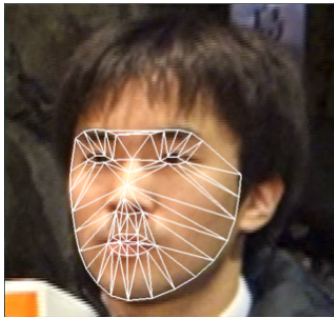


図 3: 顔姿勢推定結果

Fig. 3 Result of face pose estimation.



図 4: 顔変形処理

Fig. 4 Face deformation processing.

を制御点とし、顔の変形を行い、他者の容貌に変化させる。顔画像の変形手法には Schaefer らの rigid MLS method を使用した [9]。

本フィルタでは変形のパラメータとして、眉を大きさ、眉と眉の距離、眉の高さ、目を大きさ、目と目の距離、目の高さ、鼻の横幅、鼻筋の高さ、鼻の高さ、口の大きさ、口の高さ、輪郭の大きさ、の計 12 個のパラメータがある。顔画像に対し、顔変形処理を適用した例を図 4 に示す。ここでは、鼻幅を広く、眉と眉の距離を狭く、口の位置を高く、輪郭が大きくなるようにパラメータを操作し変形を行った。

3.4 顔重畳フィルタ

顔重畳フィルタの場合、3.2 の顔姿勢推定で得た位置情報をもとに、別人の顔画像を変形して重畳することで、他者の容貌に変化させる。顔重畳フィルタでは、まず重畳画像を変形し、それに合わせたマスク画像の作成、そして顔重畳を行う。ここでは、それぞれの処理に分けて説明する。

3.4.1 重畳画像の変形

顔の特徴点の位置情報から、パブリックドメインとして登録した顔画像の中から選択された画像を、重畳される映像中の画像の両目の間の距離に応じてアフィン変換を行い拡大・縮小する。そして、重畳する顔画像を重畳される画像の顔姿勢に応じて、メッシュの点の位置情報を変更し、変形処理を施したのち、重畳される画像に重畳処理を施す。なお、重畳する画像を拡大・縮小するにあたり、両目の間の距離を参照している理由は、重畳において目が最も特徴があり、目がずれると非常に不自然であると判断したため

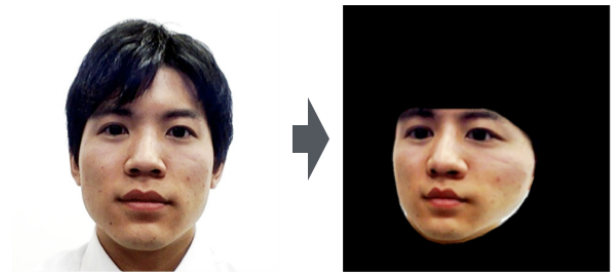


図 5: 重畳画像の変形

Fig. 5 Deformation of the superimposing image.

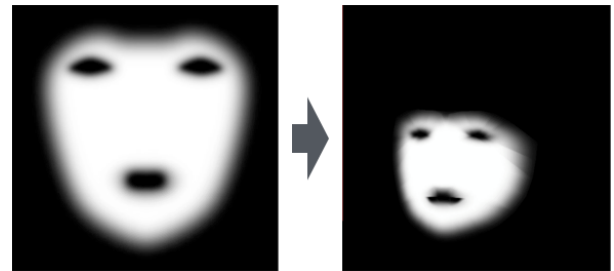


図 6: マスク画像作成

Fig. 6 Generate the mask image.



図 7: 顔重畳処理

Fig. 7 Face superimposed processing.

である。予め用意した重畳する顔画像に対し、変形処理を施した例を図 5 に示す。

3.4.2 マスク画像の作成

重畳処理に使用するマスク画像も 3.4.1 と同様の処理を行い作成する。重畳範囲は、顔重畳の自然さや個人特定率について調査した白石ら [12] を参考に、細かな表情を反映できるように目と口を重畳範囲から除いたマスク画像を使用する。作成したマスク画像とマスク画像の変形例を図 6 に示す。

3.4.3 顔重畳

本手法では、肌の色を適用者に合わせられる Poisson Blend と、重畳の境界を自然にできる Alpha Blend を組み合わせる。まず Poisson Blend により肌の色が自然になるように顔重畳し、顔の輪郭部分と目、口などの重畳に alpha の重みづけをし Alpha Blend を行い、表情・目線を残す処理を加えた。図 7 に顔重畳処理を行った例を示す。

3.5 複合フィルタ（顔変形+顔重畳）

複合フィルタの場合は、顔変形フィルタと顔重畳フィルタの両過程の処理を施す。つまり、顔の姿勢推定で得た位置情報をもとに、顔の変形処理を施し、変形した顔画像の各点の位置情報を元に顔画像を変形して重畳することで、他者の容貌に変化させる。

4. MSF の性能評価実験

容貌変化システムを、従来使用されてきたモザイク処理・黒塗り処理と比較することにより、システムの有効性について評価・考察を行う。また提案した MSF の各フィルタの有効性についても比較して評価・考察を行う。システムの有効性、各フィルタの有効性については、表情反映がされているか、また名誉毀損に抵触するような不快感がないかどうかという点で評価する。

4.1 実験概要

WEB カメラから取得した映像に対して、5つの画像処理フィルタ（モザイクフィルタ、黒塗りフィルタ、顔重畳フィルタ、顔変形フィルタ、複合フィルタ）を施せるシステムを体験してもらう。そして、各フィルタに対して、どの程度不快に感じるか、どの程度表情が読みとれるか、他人に個人特定されないかアンケートで調査した。

4.2 被験者

被験者は 20 代から 50 代の男性 13 名、女性 3 名で平均年齢は 26.0 歳である。

4.3 手続き

被験者は上部に WEB カメラが設置された PC の前に座り、自身の顔画像に画像処理が施された映像を眺める。今回はモザイク・黒塗り・顔重畳・顔変形・複合の 5つのフィルタを使用し、被験者の顔画像に対し画像処理を施した。まず被験者には、被験者の顔が PC ディスプレイ中心に位置するように、Web カメラの角度や椅子の高さを調整してもらった。そして、被験者に WEB カメラの前で顔の向きを変えたり、笑顔・悲しい顔など表情を変化させたり自由に体験させつつ、フィルタを通して変化した自身の顔を眺めさせた。各フィルタに対して、最低でも 1 分間体験してもらい、その後アンケートと自由記述を記入してもらった。

顔重畳フィルタでは予め用意した 10 枚の顔画像から自由に選択して重畳を行ってもらった。顔変形フィルタでは、目の大きさ・口の大きさ・鼻幅の大きさ・輪郭の大きさ・輪郭の横幅・眉の高さ・目の高さ・鼻の高さ・口の高さ・眉の高さ・目と目の間の距離・眉と眉の間の距離・鼻筋の長さの項目の中から、過半数以上の変形項目のパラメータを変更してもらった。複合フィルタでは、顔重畳フィルタと顔変形フィルタ両方の操作を、自由に組み合わせて行っ

表情解読度（平均+標準誤差）

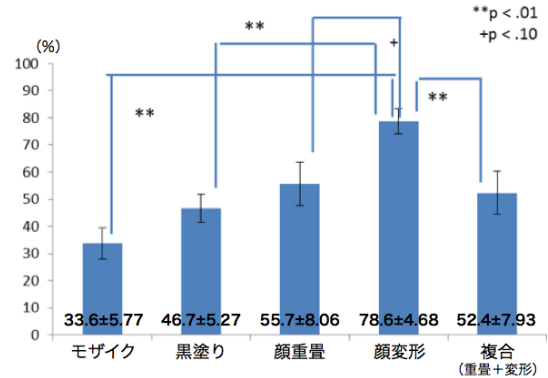


図 8: 画像処理と表情解読度

Fig. 8 Facial expression decoding degree by image processing.

不快感（平均+標準誤差）

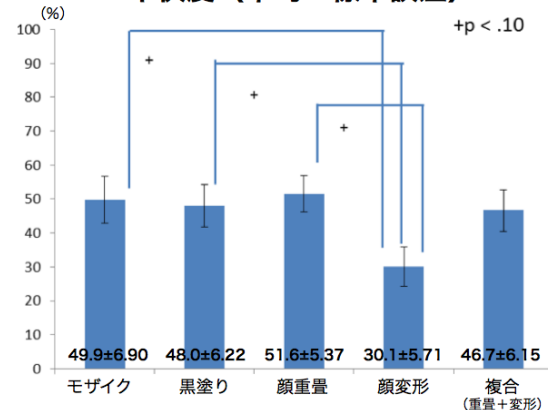


図 9: 画像処理と不快感

Fig. 9 Discomfort degree by image processing.

てもらった。

5つのフィルタを施す順番は被験者ごとにランダムになるように調整した。また表情をあまり変化させていない被験者には、笑ったり、悲しい顔をするように指示した。

4.4 評価手法

被験者に各フィルタを適用した時の表情解読度と不快度の質問項目に対して主観評価尺度（Visual Analogue Scale: VAS）で測定し、各項目の平均値と標準偏差を算出する。その後、各項目の平均値の比較を行うために、被験者内で一要因分散分析と Holm 法による多重比較を行った。また被験者に各フィルタを適用した後に、どの順に他者から見て自身を特定されづらいと感じたかを順位づけしてもらい、順位を測定する。そして平均順位を明らかにする。

4.5 実験結果

各フィルタを適用した時の表情解読度と不快度に対する VAS の結果を、図 8 と図 9 に示す。それぞれの平均+標準誤差を算出した結果、モザイクの場合 33.6 ± 5.77 、黒

表 1: 個人特定されにくい順位

Table 1 Ranking of the difficulty to identify the individual.

フィルタ	中央値
複合	1.75
モザイク	1.88
顔重畳	2.94
顔変形	4.91
黒塗り	4.25

塗りの場合 46.7 ± 5.27 , 顔重畳の場合 55.7 ± 8.06 , 顔変形の場合 78.6 ± 4.68 , 複合で 52.4 ± 7.93 だった. 表情解読度に関して分散分析を行った結果, 有意差がみられた ($F(4,15)=7.20$, $p<.05$). Holm 法による多重比較の結果, モザイクと顔変形に有意差があり ($p<.01$), 黒塗りと顔変形に有意差があり ($p<.01$), 顔重畳と顔変形に有意差があり ($p<.05$), 複合と顔変形に有意差がみられた ($p<.01$).

各フィルタの不快感を VAS で測定し, 平均+標準誤差を算出した結果, モザイクの場合 49.9 ± 6.90 , 黒塗りの場合 48.0 ± 6.22 , 顔重畳の場合 51.6 ± 5.37 , 顔変形の場合 30.1 ± 5.71 , 複合で 46.7 ± 6.15 だった. 不快感に関して分散分析を行った結果, 有意差がみられた ($F(4,15)=2.99$, $p<.05$). Holm 法による多重比較の結果, モザイクと顔変形に有意傾向があり ($p<.10$), 黒塗りと顔変形に有意傾向があり ($p<.10$), 顔重畳と顔変形に有意傾向がみられた ($p<.10$).

他者から見て自身を特定されにくいと感じたかの各フィルタでの順位の中央値を表 1 に示す.

また, Friedman 検定の結果, 有意差 ($p<.01$) がみられた. Bonferroni 法によって有意水準を調整した上で, Wilcoxon の符号付順位検定を一つごとに繰り返し多重比較を行った結果, モザイクと黒塗り間で有意差があり ($p<.01$), モザイクと顔変形間で有意差があり ($p<.05$), 黒塗りと複合間で有意差があり ($p<.05$), 顔重畳と顔変形間で有意差があり ($p<.05$), 顔重畳と複合間で有意差があり ($p<.01$), 顔変形と複合間で有意差が見られた ($p<.01$).

自由記述では, 重畳処理を行う場合, 重畳する顔画像の年齢や性別が一致していないと違和感が高まり, 不快感が強まるという意見が見受けられた. また, 顔重畳がおこなわれる条件では, 口を大きく動かした場合に, 口周辺において違和感のある変形が行われていると感じるといった意見が得られた.

4.6 考察

顔変形フィルタでは他のフィルタと比較し, 不快感の平均値が大きく下がることからフィルタを掛けられる対象に対して, 最も表現の質を低下させない処理であることが示唆された. しかし, 他の MSF 処理に比べ, 他者から見て自身を特定されやすいという欠点があることも示唆された. 一方, 顔重畳フィルタは表情解読度と不快感ともに従来手

法と大きな差が見られず, 単独で使うことが難しいと考えられる. 複合フィルタはモザイク処理と同じ程度に, 他者からみて個人特定されづらいフィルタであることが示唆され, 提案する 3 つの MSF の中で最も肖像権保護に適していることが示唆された. そして,

- 複合フィルタは, 顔変形フィルタ・顔重畳フィルタよりも, 個人特定を防ぐ効果が高いことが示唆されている
- 複合フィルタは, 表情に関して顔変形に及ばないまでも, 顔重畳フィルタと遜色ないほどの表情解読度がある
- 今後アルゴリズムを改良し, より自然に表情の反映が行えれば表情解読度が高まる見込みがある

といった理由から MSF の 3 つのフィルタのなかで, 本研究の目的である表情解読と肖像権保護の両立を最も満たしているフィルタは, 複合フィルタと結論付ける.

5. MSF の群衆映像適用実験

本実験では, 映像中の人物の顔に MSF を適用したときに, 鑑賞者がフィルタ処理された人物の元の顔を特定できるか検証する.

5.1 実験概要

予め撮影した映像に MSF を適用し, 他者から見て映像中のフィルタ処理された人物の個人特定ができるかを検証する. また, 個人特定された原因がフィルタによる影響なのか調べるために, フィルタが施されていることを認識できた割合を測定する.

5.2 被験者

被験者は 20 代の男性 10 名と女性 1 名で平均年齢は 23.2 歳である.

5.3 手続き

被験者はディスプレイの前に座り, PC ディスプレイに映し出される MSF の複合 (重畳+変形) 処理が施された映像を眺め, アンケートに回答する.

まず, 被験者にアンケート用紙に記載された顔画像群を眺めてもらった. このアンケートに記載された人物が映像中に登場するかを判断できるよう, 最低でも 1 分間, 被験者がアンケートの質問を十分回答できるよう満足いくまで眺めるよう指示をした. 図 10 にアンケートで使用した画像の一部を示す. 質問では, アンケートに記載されている人物が映像に映り込んでいるかを判断して, チェックマークをつけてもらった. そして, 記載されている人物が映像に映り込んでいると感じた場合, その人物に何らかの画像処理が施されているか判断してもらう. 映像は 2 回繰り返し再生した. 実験で使った映像のキャプチャを図 11 に



図 10: アンケートに使用した人物例

Fig. 10 Example image used in our questionnaire at subjective evaluation.



図 11: 実際に使用した映像

Fig. 11 The video that was actually used.

示す。

映像は被験者がアンケート用紙を見ながら判断できるよう、29.97 フレーム/秒の映像を半分の 15 フレーム/秒に修正した。なお映像の再生や切り替えなど、画面の操作は実験者側で行った。MSF の重畳処理に使用する画像は予め用意した 10 枚の画像を、映像中の各人物に行い重畳した。変形も映像中の各人物に対して異なるように、変形パラメータがランダムになるように設定した。アンケートに記載した顔画像は 9 人で、映像中に映り込んだ顔画像の中から、正面顔で映り込んだフレーム数の多い上位 10 人の中から選択した 7 人と、映像に映っていないダミー 2 人を使用した。被験者は、この映像鑑賞とアンケート記入の作業を、各映像に行う。

5.4 評価手法

映像を鑑賞しながら、アンケートに記載された人物が映像中に登場したか判断してもらい、その正答率から個人特定率を測定する。フィルタを適用した人物・フィルタを適用していない人物の群に分けて、それぞれの群での個人特定正答率の平均値と標準誤差を算出する。

5.5 実験結果

MSF を適用した人物と、適用していない人物間で個人特定率の平均 + 標準誤差を算出した結果を図 12 に示す。フィルタ無しでは $90.9 \pm 8.71(\%)$ 、フィルタ有では $84.5 \pm 10.9(\%)$ だった。フィルタ間の個人特定率に関して t 検定を行った結果、有意傾向がみられた ($p < .10$)。

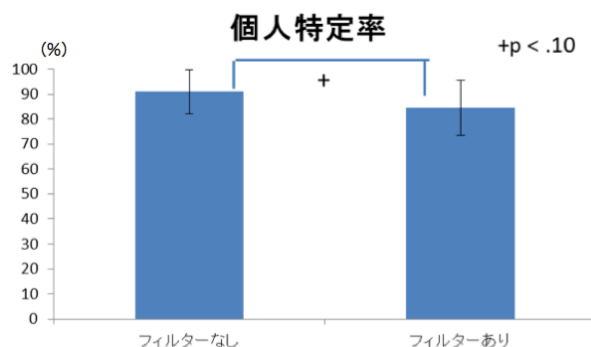


図 12: フィルタ適用と個人特定率

Fig. 12 Personal specific rate by the filter application.

5.6 考察

MSF を適用すると、何も処理を施さないよりも他者から見て個人特定される割合は減少し、有意傾向が見られた。しかしながら、フィルタを適用しても、個人特定率が $84.5 \pm 10.9(\%)$ と、個人特定を妨げる機能がまだ不十分であることが示唆された。

また、顔が画面外に出るときにフィルタが外れ、その瞬間に個人特定できたと被験者に指摘された。現状では、映像中の人物への顔認識が失敗し、フィルタで処理を施した部分が歪んでしまったり、重畳する顔画像が時々ずれるという問題がある。顔認識が上手くいき、フィルタで施した処理部分が歪まず、重畳した顔画像がほとんどずれなかった人物の個人特定率は 45.45 % であった。そのため、顔認識の精度を改善することで、これ以上の個人特定率の減少が期待される。

また、今回の実験では、映像を通常半分の速度で再生し、映像と写真画像をいつでも見比べることができ、さらに 2 回映像を鑑賞していたため、単純に映像を 1 回眺めてもらうよりも遥かに個人特定が行いやすい環境であった。今後は、通常の映像視聴方法から離れないような実験環境を設計し MSF の評価を行う。

6. おわりに

本研究では表情や目線を読み取れることと肖像権保護の両立を実現するための画像処理システム MSF (顔重畳フィルタ・顔変形フィルタ・複合 (重畳 + 変形) フィルタ) を構築した。

MSF を用いて、表情読み取りと肖像権保護の両立ができるか、またシステム適用者に対して不快感が強くないか評価する実験を行った。実験結果から、顔重畳手法の改善が必要ではあるものの、複合フィルタが最適であると結論付けた。また、他者に MSF を適用した時の評価実験を行い、肖像権保護できるかどうか評価実験を行った。映像を鑑賞し、映像中の人物をどのくらい個人特定できるか・フィルタ処理されていると気づいたか検証した。その結果、フィルタ適用者に対する個人特定率が減少したことから肖像権

保護の役割を果たしていることが示されたが、現状では効果が薄かった。これらの実験から、提案手法が肖像権を保護できる可能性が示唆されたが、顔認識の精度を改善する必要があることが分かった。

今回は、肖像権保護のために顔に着目して加工を行い、プライバシー保護できる可能性が示唆されたが、個人特定できないようにする方法は、顔部分に加工を施すだけではない。顔の変形・重量だけでなく、髪型を変形させたり置き換える方法や、服装の色を変えるとといった方法で、容貌を別人に変化させることも考えられる。顔だけでなく他の部位にも加工を施すことで、より良い容貌変化システムを構築できる可能性があると考ええる。

参考文献

- [1] 大家重夫. 肖像権. 太田出版, 2007.
- [2] Dmitri Bitouk, Neeraj Kumar, Samreen Dhillon, Peter Belhumeur, and Shree K Nayar. Face swapping: automatically replacing faces in photographs. *ACM Transactions on Graphics (TOG)*, Vol. 27, No. 3, p. 39, 2008.
- [3] Jens Garstka and Gabriele Peters. View-dependent 3d projection using depth-image-based head tracking. In *8th IEEE International Workshop on Projector/Camera Systems PROCAMS*, pp. 52–57, 2011.
- [4] Branko Samarzija and Slobodan Ribaric. An approach to the de-identification of faces in different poses. In *Information and Communication Technology, Electronics and Microelectronics (MIPRO), 2014 37th International Convention on*, pp. 1246–1251. IEEE, 2014.
- [5] Arthur Niswar, Ee Ping Ong, and Zhiyong Huang. Face replacement in video from a single image. In *SIGGRAPH Asia 2012 Posters*, p. 12. ACM, 2012.
- [6] Yuta Nakashima, Tatsuya Koyama, Naokazu Yokoya, and Noboru Babaguchi. Facial expression preserving privacy protection using image melding. In *Proceedings of IEEE International Conference on Multimedia and Expo (ICME2015)*, 2015.
- [7] Naoto Nakazato, Shigeo Yoshida, Sho Sakurai, Takuji Narumi, Tomohiro Tanikawa, and Michitaka Hirose. Smart face: enhancing creativity during video conferences using real-time facial deformation. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pp. 75–83. ACM, 2014.
- [8] 吉田成朗, 鳴海拓志, 櫻井翔, 谷川智洋, 廣瀬通孝. リアルタイムな表情変形フィードバックによる感情体験の操作. ヒューマンインタフェース学会論文誌, Vol. 17, No. 1, pp. 15–26, 2015.
- [9] Scott Schaefer, Travis McPhail, and Joe Warren. Image deformation using moving least squares. In *ACM Transactions on Graphics (TOG)*, Vol. 25, pp. 533–540. ACM, 2006.
- [10] Michael Jones and Paul Viola. Fast multi-view face detection. *Mitsubishi Electric Research Lab TR-20003-96*, Vol. 3, p. 14, 2003.
- [11] Jason M Saragih, Simon Lucey, and Jeffrey F Cohn. Deformable model fitting by regularized landmark mean-shift. *International Journal of Computer Vision*, Vol. 91, No. 2, pp. 200–215, 2011.
- [12] 白石英明, 吉田成朗, 伊藤裕之. プライバシー保護のための mob scene における自然な顔合成手法の検討. 日本バー

チャルリアリティ学会大会論文集 Proceedings of the Virtual Reality Society of Japan, Annual Conference, 第 19 巻, pp. 472–475. 日本バーチャルリアリティ学会, 2014.