

<b>氏 名</b>	さとう じゅんや 佐藤 惇哉
本籍（国籍）	岩手県
学位の種類	博士(工学)
学位記番号	工博 第275号
学位授与年月日	平成29年3月23日
学位授与の要件	学位規則第5条第1項該当 課程博士
研究科及び専攻	工学研究科デザイン・メディア工学専攻
<b>学位論文 題目</b>	<b>Photographic Environment Independent Multiview Face Detection and Tracking Using Template Generation by Genetic Algorithm</b> (遺伝的アルゴリズムによるテンプレート生成を用いた撮影環境に依存しない多方向顔検出と追跡)
学位審査委員	主査 教授 今野 晃市 副査 准教授 明石 卓也 副査 准教授 藤本 忠博

## 論文内容の要旨

In this thesis, photographic environment independent multiview face detection and tracking is addressed. The photographic environment means a variety of illumination conditions, backgrounds, appearances of objects, and qualities of images in this thesis. The multiview face detection and tracking in these practical environments is a difficult problem. In order to solve this problem, novel methods using template generation and template matching with genetic algorithm (GA) are proposed.

This research mainly focuses on two researches. In the first research, high-speed multiview face detection and tracking in various photographic environments is addressed as a basic research. The proposed method adopts template matching. In this method, a template is created in advance as the first step. Then, a region whose pattern is similar to the template is searched by the sliding window, etc. This is the basic concept of template matching. However, one template is generally able to detect only one pattern. Hence, many patterns of templates are necessary to detect multiview faces. Since generating many templates and calculating all the matching scores are inefficient, a novel method is proposed.

First of all, one 2D face model, whose shape is rectangle, is created. Next, three parameters are defined to generate templates from the model for multiview

faces. By optimizing these parameters using GA, the optimal template can be found efficiently. The GA is a method to find a global optimum within a given search resource. There exist population as search points in a search space and each individual consists of chromosomes. The chromosomes are parameters and they are encoded as bit strings. After the population is evaluated using a fitness function, the chromosomes are updated by iterating genetic operations, such as selection, crossover, and mutation. Since the population evolves to acquire high fitness, a global optimum or its approximate solution can be obtained finally. By introducing GA to the template generation, the optimal templates can be acquired efficiently.

The GA is also applied to target object search instead of sliding window. This method scans the whole target image using a search window with a variety of scales. This is equal to that all the combinations of the coordinate and the scale parameters are checked. Since this approach is inefficient and in-plane rotation is not considered, GA is applied. By optimizing geometric transformation parameters, such as parallel translation, scale, and in-plane rotation, all the parameters are simultaneously optimized and the target object can be localized efficiently.

For experiments, a challenging 60 video dataset was created. Subjects were recorded under various photographic environments. The proposed method is compared to a machine learning-based method and a face tracking method on the dataset. As a result, high accuracy and fast processing of the proposed method are confirmed.

In the second research, multiview face tracking on privacy protected videos is addressed. Recently, many surveillance cameras have been set at public and private spaces. Along with this, many computer vision techniques, such as pedestrian detection, action recognition, and face recognition are also applied. This is necessary for crime deterrence, accident detection, and so on. However, some people may suffer psychological pressure about recording their faces. If reducing the pressure is considered, the recorded videos must be preprocessed to protect the privacy. For this purpose, blur and pixelation can be applied and they are often used in news programs. However, the computer vision techniques cannot be used on the preprocessed videos since the original pixel information are changed. Since my proposed multiview face detection and tracking method in the first research of this thesis is also unable to be applied completely, a new method is proposed.

First of all, a new preprocessing method is proposed to generate a privacy protected image and preserve the pixel value. In this method, a  $n \times n$  filter is set in a target image and a neighborhood number is determined randomly. Then, the

pixel values between the centered pixel of the filter and randomly selected neighborhood are replaced. By applying this filter to the whole image, privacy protected images are generated while the original pixel value is preserved.

Since the color information is preserved, the proposed method uses color histograms as a template to track a multiview face. In this research, Cr and Cb histograms are used because they are robust to illumination changes. The histograms are obtained from a frontal target face. By optimizing geometric transformation parameters using GA and localizing a region whose color histograms are similar to the template in target images, the multiview face tracking can be achieved.

For experiments, a challenging video dataset was recorded under a variety of photographic environments, and the proposed method is compared to related works. As a result, the effectiveness of the proposed method is confirmed.

As described, this thesis contributes photographic environment independent multiview face detection and tracking using by GA.

## 論文審査結果の要旨

本論文は、撮影環境に依存しない多方向顔検出と追跡の研究について述べている。この研究における撮影環境とは、様々な照明環境や背景、物体の見え、画像品質の違いを意味している。これらの異なる撮影環境において、実用的な多方向顔検出と追跡を達成するために、テンプレートマッチングと遺伝的アルゴリズム (genetic algorithm, GA) を組み合わせた手法をベースとして、新たな手法を提案している。具体的には、手動で作成されたひとつの2次元顔モデルからGAで最適なテンプレートを自動で生成し、また、生成されたテンプレートと最もよくマッチングする対象画像中の領域をGAで効率的に探索することで、先行研究よりも高速かつ高精度な多方向顔検出と追跡を達成している。また、この提案された基礎的技術を応用することで、プライバシーが保護された動画における多方向顔追跡手法も提案している。この手法においては、元の画素情報を保存しつつプライバシーが保護可能な画素の置き換えフィルタを提案している。この前処理によって生成された動画で被験者の多方向顔を追跡し続けるために、被験者の正面顔から取得された色ヒストグラムをテンプレートとして取得し、これと最もよくマッチングする領域をGAで効率的に探索することで、従来よりも高精度な多方向顔追跡精度を達成している。

本論文の構成は以下の通りである。

第1章は序論であり、この研究の背景や取り組んでいる内容、提案手法のアプローチといった概要が記述されている。

第2章では、本論文に関わる理論的背景について記述している。はじめに、顔検

出に関するこれまでの先行研究について簡単に説明している。次に、本研究で注目しているテンプレートマッチングの概要について述べ、テンプレートマッチングの問題点について言及している。例えば、多方向顔を検出するためにはそれぞれの顔向きに対応するテンプレートを事前に作成しておき、ひとつずつマッチングする必要がある。しかし、このアプローチは非効率的という問題がある。また、マッチングのためには画像全体を走査する必要があるが、計算コストが高いため、これも非効率的である。これらの問題を解決するために、本研究ではメタヒューリスティックアルゴリズムに注目することを述べている。ここでは例として、GAと粒子群最適化 (PSO)、差分進化について簡単にアルゴリズムを紹介している。その後、これらのアルゴリズムをテンプレートマッチングに適用する具体的な方法について説明し、GAとPSOの性能比較を動画像処理で行っている。

第3章では、第2章で述べられている技術を発展させることで、高速かつ高精度に多方向顔を検出、追跡できる手法について説明している。この手法では、あらかじめ手動で2次元の顔モデルを作成しておき、このモデルから様々な顔向きに対応するテンプレートを生成するためのパラメータを定義している。これらのパラメータをGAで最適化することで、自動で最適なテンプレートを生成する手法を提案している。また、生成されたテンプレートと最もよくマッチングする領域の探索もGAで行うことで、先行研究よりも高速かつ高精度な検出と追跡精度を達成している。

第4章では、第3章で構築した手法の、応用的研究例について述べている。具体的には、プライバシーが保護された動画における多方向顔追跡に取り組んでいる。近年、人間の顔を対象としたビジョンアプリケーションが増えてきているが、顔を撮影されることに精神的苦痛を感じる人が存在する。可能な限りその苦痛を軽減するためには、ブラーリングやピクセル化といった前処理が必要であるが、元の画素情報が改変されてしまうため、処理された画像で物体検出を実現するのは困難な問題である。第3章で構築した手法も適用することはできない。この問題を解決するために、画素情報を保存しつつプライバシー保護画像を生成できる新たな前処理手法を提案している。また、処理された動画で多方向顔を追跡し続けるために、色ヒストグラムを追跡のための特徴量として使用し、テンプレートマッチングとGAの組み合わせで高精度に顔を追跡し続ける手法を提案している。

第5章は結論であり、本論文のまとめが記述されている。

以上、本論文は撮影環境に依存しない高精度な多方向顔検出と追跡手法を新規に提案し、その有効性と有用性を示したものであり、メディア工学分野やコンピュータビジョンの発展に寄与するところが大きい。

よって、本論文は博士（工学）の学位論文として合格と認める。

## 原著論文名（1編を記載）

**Junya Sato, Takuya Akashi: High-speed Multiview Face Localization and Tracking**

with a Minimum Bounding Box Using Genetic Algorithm, IEEJ Transactions on Electrical and Electronic Engineering, Vol.12, No.5, 2017 (予定)