

# 深度カメラを用いた耳介の特徴点抽出法の提案と評価

2017SC026 河合 翼 2018SC002 浅野雅仁

指導教員：石原靖哲

## 1 はじめに

現在世の中では、パスワードや IC チップ内蔵のデバイスを用いる個人認証システムが広く普及している。さらに、マイナンバーカードを IC カードリーダーやスマホのリーダーモードなどの IC チップから読み取ることで本人確認ができる、内閣府が推奨する公的個人認証システムがある。しかし、これらの個人認証システムは盗聴、盗難、推測、貸し借り、コスト、紛失などの恐れがあり、その場合他人でも容易に本人を偽ることが可能である。そこで注目されている認証システムが指紋など固有性の高い人間の身体的特徴をデータ化して本人確認に用いるバイオメトリクス認証、いわゆる生体認証である。

生体認証とは、誰もが持っている特徴である普遍性、本人以外は同じ特徴を持たない唯一性、時間の経過とともに変化しない永続性、この3つの条件を満たす生体的な測定結果を用いて、本人かどうかを自動的に確認する技術である。さらに、生体認証の利点として本人の意識に影響しないこと、管理の依存度が低いなどの運用強度が強い点である。現在、Android 端末や iPhone 端末のロック解除方法には、暗証番号、パスワード、指紋認証、顔認証、虹彩認証や静脈認証などの様々な解除方法がある。その中でも顔認証は、内蔵カメラを利用し、精度・速度ともに十分である。しかし、コロナ禍において日常でマスクを着用することがあたりまえになったいま、顔認証をストレスなく利用することが難しくなっている。

このような社会的背景のもと、本研究では、上記の普遍性、唯一性、永続性の3つの条件を満たす耳介に注目する。南谷らによる耳介を用いたユーザ認証システム [4] では、防犯カメラの画像から耳介認証を行っている。ゆえに、耳の形がきれいに撮れていないかもしれないという前提条件が与えられている。それに対し、我々はスマートフォンのロックを解除するのに耳介認証を用いることを想定する。そうすることで、人々が自分の耳の形をカメラに向けてはっきり見せてくれるという前提条件が置ける。その南谷らが提案する7つの耳介の特徴点の内の1つである対耳輪を深度カメラ (RealSense Depth Camera D435i) で特徴点抽出する。具体的には、OpenCV でビデオキャプチャとして深度カメラを利用する。そして、OpenCV Haar Cascade による右耳のオブジェクト検出を行う。この際、対耳輪の深度が最も浅い位置になるようにフレームを5分の1スケールで行う。深度カメラの Depth Sensor (IR Stereo Camera) から深さを撮影し、その深度データを OpenCV の Python 上で数値化する。深度カメラからの情報を使い、縮小したフレーム内での深度の最小値を得

る。そして深度の最小値に対応する  $x$  座標、 $y$  座標を特定することで対耳輪の特徴点抽出を行う。これらの工程を経て、先行研究の特徴点抽出の手法との差別化と、耳介認証システムの実用化を目標とする。我々は、目標を達成するために深度カメラで撮影した耳介の特徴点を抽出する実験と、あらゆる場面での実用化を想定して Bluetooth イヤホンという障害物を用いた実験を行う。

本研究の成果として、オブジェクト検出で縮小したフレーム内にある深度の最小値とそれに対応する、 $x$  座標、 $y$  座標を Python 上で数値化することで対耳輪の深度、 $x$  座標、 $y$  座標の抽出を15枚中15枚という精度で取得することができた。Bluetooth イヤホンを用いた実験でも15枚中12枚という精度で取得することができた。このように、深度カメラを用いて先行研究の特徴点抽出法とは別角度からアプローチする特徴点抽出法を提案することができた。

## 2 関連研究

### 2.1 耳介認証に関する研究

南谷ら [4] は、人間の耳の形を利用した個人認証 (耳介認証) のシステム構築を試みている。そもそも、耳介認証とは、耳の穴の凹凸の個人差を利用して、個人を特定する生体認証技術である。耳の形は複雑に入り組んでいて、人によって形状が比較的異なる。また、加齢や体格の変化によっての変化が少なく認証精度も高いため、犯罪などの捜査でも利用されている。

南谷らは入り組んだ人の耳介を利用し、犯罪などの捜査で犯人を特定できるように個人認証を行っている。その認証方法は大きく2つに分けて特徴点を抽出し認証している。特徴点の抽出方法としてまず1つ目に、画像内から耳の形状を見つけ出す方法である。2つ目に各々の耳の特徴点を見つけ出して特徴点を抽出している。前半の粗探索では、上方に位置する対耳輪前脚、対耳輪後脚、対耳珠の3点の特徴点を抽出している。髪の毛が上から下に向かって伸びていることから、これらの3点が確認できれば全体が確認できると判断している。後半では7つの特徴点すべてを抽出している (図1)。

犯行現場近くの監視カメラでは、遠距離で犯人を撮影している。そのため、映像でのあらゆる解像度の場合が想定される。さらに、犯人を撮影したときはいつも同じ角度から撮影されることは想定されない。さらに設置している監視カメラの画像の画質は全て同じではなく、画像が良好ではないものがある。この条件下において、ある一定の数値までは認証精度は高く、撮影角度も一定の角度までは認証できている。撮影角度、解像度共に舟状窩の識別ができなくなると認証精度が変わり、適用の限界を示している。舟

状窩とは、耳介の最外周縁である C 字状の輪郭をした耳輪の内側縁の溝のことである。犯人の特徴の一例として、イヤリングをつけた時での認証を試しているが、イヤリングをつけたことによって認証への精度は影響しないことがわかった。

## 2.2 深度カメラに関する研究

深度カメラと RGB カメラを使用し、物体検出を行っている先行研究として島川らの研究 [6] が挙げられる。

島川らの目的は、視覚障害者が自立して生活する際ににおける転落、転倒や衝突による事故を解決し、視覚障害者への歩行を支援することである。

島川らの研究は、深度画像と RGB 画像の特徴を統合することでできた画像に特徴点追跡手法を導入することで階段とその段差を検出する研究である。島川らは、ASUS 社の Xtion PRO LIVE という RGB-Depth カメラをモーションキャプチャデバイスとして RGB 画像と深度画像の両方を補完的に利用している。まず、同時に撮影した RGB 画像と深度画像に対し、エッジフィルタおよびハフ変換を行う。そして、それぞれ得られたエッジ画像を統合し、それを階段の段差候補エッジ画像とする。この段差候補エッジ画像から選定される段差エッジの座標を特徴点としてオプティカルフローを用いて追跡すると RGB 画像上で段差エッジを追跡できるので、エッジフィルタおよびハフ変換の処理を省略し、処理時間を短縮することができる。この手法により、従来の方法と比べ平均処理時間を短縮し、大幅な検出成功率が改善された。しかし、太陽光などによる強い光があたる状況下では深度カメラでの深度距離が正しく測定されないことがわかった。

## 3 提案する手法

### 3.1 使用する機器

本研究にて使用する深度カメラ (RealSense Depth Camera D435i) について説明する。RealSense は Intel 社が提

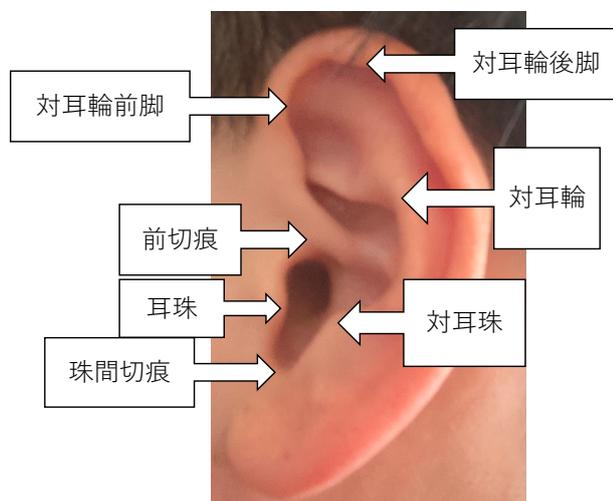


図 1 耳介の特徴点

供する深度カメラである (図 2)。

深度とは、カメラから物体までのことである。Depth Camera D435i では、深度計測が可能なステレオビジョンのカメラで、二つの深度センサ、RGB センサ、IR 投射器を搭載し、電源供給やデータのアウトプットはすべて USB で行うことができる [1]。さらに、Depth Camera D435i は人の目で判断できないほどのわずかな動きの変化を高精度に検出する IMU を追加することで、カメラが動作するあらゆる状況において、深度の認識を調節でき、正確に深度測定をすることができる。本研究ではカメラに耳をはっきりと見せていると前提条件を置いているので正確に深度測定ができる Depth Camera D435i が適していると考えられる。

### 3.2 ライブラリ

OpenCV (Open Source Computer Vision Library) は、画像処理を行う膨大な関数を持つオープンソースのライブラリである。一般的な二次元の画像処理、ヒストグラム処理、テンプレートマッチング、オプティカルフロー、および顔認証など多様なアプリケーションを開発できる関数を持っている。OpenCV を利用すると、数行のコードを記述するだけで簡単な画像処理プログラムを開発でき、Pythonをはじめ、C/C++、Java、MATLAB など様々なプログラミング言語で公開されている。

### 3.3 システムの概要

まず、OpenCV Haar Cascade による耳の検出を行う。OpenCV では、カスケード識別器の学習済みファイルが事前に用意されている。カスケード識別器のファイルの対象物体として、両目や、右目、左目、メガネ、顔、笑顔、全身、上半身、下半身、右耳、左耳など様々な学習済みファイルが存在する [2]。これらの学習済みファイルを GitHub からダウンロードして、深度カメラの RGB カメラから取り込んだ映像をオブジェクト検出する [5, 185–187 ページ] にあるソースリストを元に右耳をオブジェクト検出するプログ



図 2 RealSense Depth Camera D435i

ラムに書き換えた。さらに、対耳輪を対象に特徴点抽出を行うために、対耳輪の深度が最も浅い位置と見込まれる領域になるようフレームの幅と高さを5等分割した。ここで、対耳輪とは、耳介の最外周縁であるC字状の輪郭をした耳輪の内側縁の溝である舟状窩とほぼ平行に走っている隆起のことである。まず、フレーム左上頂点の座標を決めるために幅と高さをそれぞれ5分の1倍する。次にフレーム右下頂点の座標を決めるためにフレームの幅と高さをそれぞれ5分の2倍する(図3)。そのプログラムを以下に示す。

```
if len(rightear) > 0:
    for rect in rightear:
        cv2.rectangle(frame,
                       (rect[0]+rect[2]//5,
                        rect[1]+rect[3]//5),
                       (rect[0]+rect[2]*2//5,
                        rect[1]+rect[3]*2//5),
                       (0,0,255), thickness=1)
```

このフレーム内で取得できる深度の最小値とその深度に対応する  $x$  座標,  $y$  座標を Python 上で数値化するシステムを実装した。そのプログラムを以下に示す。

```
min=100000
minx=0
miny=0
for i in range(rect[0]+rect[2]//5,
                rect[0]+rect[2]*2//5):
    for j in range(rect[1]+rect[3]//5,
                  rect[1]+rect[3]*2//5):
        min > frame1[i,j]:
            min=frame1[i,j]
            minx=i
            miny=j

print(min,minx,miny)
```

さらに、以下のプログラムを書き加えることでフレーム内で取得できる深度の最小値を緑のマークで可視化した。

```
cv2.rectangle(frame,
              (minx, miny),
              (minx+1, miny+1),
              (0, 255, 0), thickness=1)
```

これらの編集を行うことで、縮小されたフレーム内で取得できる深度の最小値と、その深度の最小値に対応する  $x$  座標,  $y$  座標を対耳輪として出力し、対耳輪を対象に特徴点抽出を行う。

#### 4 深度カメラを用いた耳介の特徴点抽出の実験

本実験では、我々が開発した深度カメラを用いた耳介の特徴点抽出システムを使用する。右耳の対耳輪の深度情報

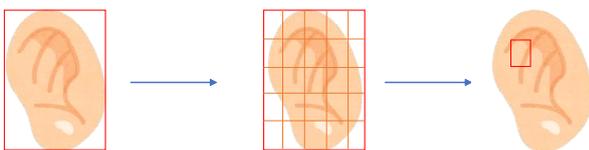


図3 フレームの縮小方法

から  $x$  座標,  $y$  座標を取得することで我々が開発したシステムの評価を行う。実験は、被験者の右側頭部が深度カメラに向くように椅子に座ってもらい行う。被験者の右耳が写るようにカメラの高さを調節する。さらに、このシステムはコロナ禍でマスク着用が義務化されている世の中での実装を目指しているため実験中は被験者はマスクを着用して行ってもらう。評価方法は、プログラム実行中(15秒間)、1秒ごとに1枚スクリーンショットを撮る。この15枚のスクリーンショットのうち何枚が対耳輪を正しく抽出できているか人手で確認する。

我々が開発した耳介の特徴点抽出システムの実験結果を図4に示す。実験を行った結果、RGBカメラでは対耳輪の深度が最も浅い位置になるように赤フレームの表示がされた。さらに、赤フレーム内において、深度カメラから最も近い位置にある対耳輪の位置を緑のマークで表示することができた。Pythonのコンソール上では左から緑のマークが表示されている対耳輪の深度情報、その深度情報に対応する  $x$  座標,  $y$  座標を数値化することができた。対耳輪の深度,  $x$  座標,  $y$  座標を15枚中15枚という精度で取得できた。

作成したシステムでは、被験者の耳の角度によって深度情報が読み取れず緑のマークが赤フレームの左上に表示されてしまうことがあった。この場合、実行結果は0と出力された。さらに、耳を検出している際に手などの耳とは関係ない部分がRGBカメラで表示されるとシステムが固まってしまった。これは、我々がシステムを作成する際に使用したカスケード識別器の学習済みファイルが耳を検出するためにあらゆる人の耳の形を学習したものであるため、耳とは違う学習していないものが表示されることによってシステムが固まってしまうと考えられる。

#### 5 イヤホン着用時における耳介の特徴点抽出の実験

我々が開発した耳介の特徴点抽出システムを実用化させるためには、あらゆる場面において認証できる必要がある。近年では、Bluetooth イヤホンの使用率が高まり、着用し



図4 RGBカメラでの出力結果

ている人を多くの場面で見かける。我々はこの Bluetooth イヤホンに注目した。ユーザが Bluetooth イヤホンを着用していても耳介の特徴点抽出が可能になれば、認証するユーザは着脱して認証するという手間を省くことで、ストレスなく認証することができる。本実験は被験者に Bluetooth イヤホンを着用してもらい前節同様に実験と評価を行う。

実験を行った結果、RGB カメラでは Bluetooth イヤホンの着用していても対耳輪の深度が浅い位置に赤フレームの表示がされた。さらに、Bluetooth イヤホンを着用していない時と同様に赤フレーム内において緑のマークが表示され、Python のコンソール上に左から緑のマークが表示されている対耳輪の深度情報、その深度情報に対応する  $x$  座標、 $y$  座標が数値化されていた。対耳輪の深度、 $x$  座標、 $y$  座標を 15 枚中 12 枚という精度で取得できた。Bluetooth イヤホンを着用していない時と比べると赤フレームの検出が遅いことがわかった。これは、我々がシステムを作成する際に使用したカスケード識別器の学習済みファイルにはさまざまな耳の形や、ピアスやイヤリング着用時の写真を利用して機械学習を行っている [3]。しかし、Bluetooth イヤホンを着用した耳介の写真は利用されていなかった。そのため、我々が障害物として使用した Bluetooth イヤホン着用時における右耳のオブジェクト検出は検出速度が着用していない時と比べて遅くなったと考えられる。さらに、Bluetooth イヤホンを着用していない時と比べて耳の角度が少しでもずれてしまうと赤フレームが検出されなかった。そして深度カメラから最も近い位置にある対耳輪の深度情報が読み取れず緑のマークが赤フレーム内に表示されないことが Bluetooth イヤホンを着用していない時と比べて多くあった。これは、作成したシステムのプログラムによるものだと考える。赤フレーム内において深度が最も浅い位置を表示する緑のマークは、プログラム上、右耳のオブジェクト検出が行えたうえで表示される設計になっている。そのため、深度カメラから最も近い位置にある対耳輪の深度情報を緑のマークで表示されなかったと考えられる。

上記 2 つの課題は残ってしまったがその中でも Bluetooth イヤホンを着用していない時と同様に赤フレーム内において緑のマークが表示され、Python のコンソールに左から緑のマークが表示されている対耳輪の深度情報、その深度情報に対応する  $x$  座標、 $y$  座標を数値化することができた。これは、対耳輪が Bluetooth イヤホン着用時でも右耳のオブジェクト検出および特徴点抽出において比較的影響が少ないからだと考えられる。

## 6 まとめ

### 6.1 本研究の成果

本研究では、対耳輪を対象として深度カメラの RGB カメラをビデオキャプチャとして耳のオブジェクト検出を行うことができた。さらに、オブジェクト検出で縮小したフレーム内にある深度の最小値とそれに対応する、 $x$  座標、 $y$

座標を Python 上で数値化することで対耳輪の深度、 $x$  座標、 $y$  座標の抽出を 15 枚中 15 枚という精度で取得することに成功した。イヤホン着用時における特徴点抽出実験では、近年使用率が高まっている Bluetooth イヤホンを用いて実験と評価を行った。Bluetooth イヤホンを着用していない時と比べて、耳介のオブジェクト検出が遅くなってしまふことと、それに伴って深度、 $x$  座標、 $y$  座標の取得ができない確率は高まった。しかし、対耳輪が Bluetooth イヤホン着用によって直接隠れることはない所以对耳輪の深度、 $x$  座標、 $y$  座標の抽出を 15 枚中 12 枚という精度で取得することに成功した。

### 6.2 今後の課題

我々が開発した対耳輪を対象とした耳介の特徴点抽出システムでは、右耳を検出している際に手などの右耳とは関係のない部分が RGB カメラで表示されるとシステムが固まってしまう。また、Bluetooth イヤホンを着用し耳介の特徴点抽出を行った際には、Bluetooth イヤホンを着用していない時と比べて赤フレームの検出が遅くなってしまふという現象がおきた。この原因としては我々はシステムを作成する際に使用したカスケード識別器によるものだと考えたため、今後の課題としてカスケード識別器の学習として Bluetooth イヤホンを着用している耳の画像をカスケード識別器に学習させることが挙げられる。我々が開発した耳介の特徴点抽出システムは対耳輪のみを対象としているので、今後の課題としては南谷らが提案する 7 つの特徴点のすべてを抽出し、その抽出した特徴点の深度情報から  $x$  座標、 $y$  座標を取得することである。この研究は人の耳の形状を使ったスマートフォンなどにおける個人を識別する耳介認証システムのベースに生かされると考える。

### 参考文献

- [1] IntelRealSenseDepthCameraD435i. <https://www.intelrealsense.com/depth-camera-d435i/>.
- [2] opencv/opencv-contrib. [https://github.com/opencv/opencv\\_contrib/tree/1311b057475664b6af118fd1a405888bad4fd839/modules/face/data/cascades](https://github.com/opencv/opencv_contrib/tree/1311b057475664b6af118fd1a405888bad4fd839/modules/face/data/cascades).
- [3] shivangbansal/haar-cascade-ear-training. <https://github.com/shivangbansal/Haar-Cascade-Ear-Training>.
- [4] 南谷崇成. 耳介による個人認証システムの撮影角度の差に対するロバスト化. Ph.D. dissertation, 埼玉工業大学大学院工学研究科, 2016.
- [5] 北山直洋. Python で始める OpenCV4 プログラミング. 株式会社 カットシステム, 2019.
- [6] 島川 学, 村上 聡一郎, 清田 公保. RGB-Depth 画像を用いた階段検出に関する研究. Technical report, FuzzySystemSymposium, 2015.