

顔認識を利用した リアルタイム視覚効果生成システム

本間春輝

2022年2月9日

目 次

1	はじめに	2
2	背景	2
3	目的	3
4	使用ライブラリ	3
5	物体検出機能の確認	3
6	深度画像取得の確認	5
7	顔認識と深度画像を利用した背景除去の実装	7
8	今後の課題	10
9	おわりに	10

1 はじめに

近年、顔認識を利用した技術や製品を目にすることが多くなってきたが、テレビや映画といった映像の分野における利用はまだ多くない。そこで本研究では、顔認識技術の映像分野での利用を目的とする。

2 背景

近年の映像作品は図1, 図2のような視覚効果を利用したものが多く存在する。しかしそういった編集を人が手作業でする際には、大きな労力と時間を費やす必要がある。そういった人への負担を軽減するために、本研究では図1, 図2のような映像における視覚効果のリアルタイム生成を目指す。



(a) 映画 (アリス・イン・ワンダーランド [1])



(b) CM (サントリー
ウイスキー Q[2])

図 1: 頭部拡大映像の例

3 目的

本研究の目的は2つある。1つめが「顔認識を利用した人の頭部領域の抽出」である。顔認識は OpenCV の物体検出機能を用いて行う。2つめが抽出領域への拡大処理である。抽出を行った領域に拡大処理を施し、もとの画像と合成することで図1のような画像の実現を行う。

4 使用ライブラリ

以下は今回のプログラムを作成するにあたり、使用したライブラリである。

使用ライブラリ

OpenCV

画像処理や画像解析をするためのライブラリ。

OpenKinect

Kinect を Windows、Linux、および Mac で使用できるようにする
無料のオープンソースライブラリ。

5 物体検出機能の確認

カメラを用いて撮影された映像から人間の顔がどのように検出されるのかを確認するために、カメラ映像から顔認識を行い、検出された座標にモザイク処理を施すプログラムを作成した。図2はカメラ映像から顔の検出を行った結果の画像であり、図3は検出された領域にモザイク処理を施した画像である。

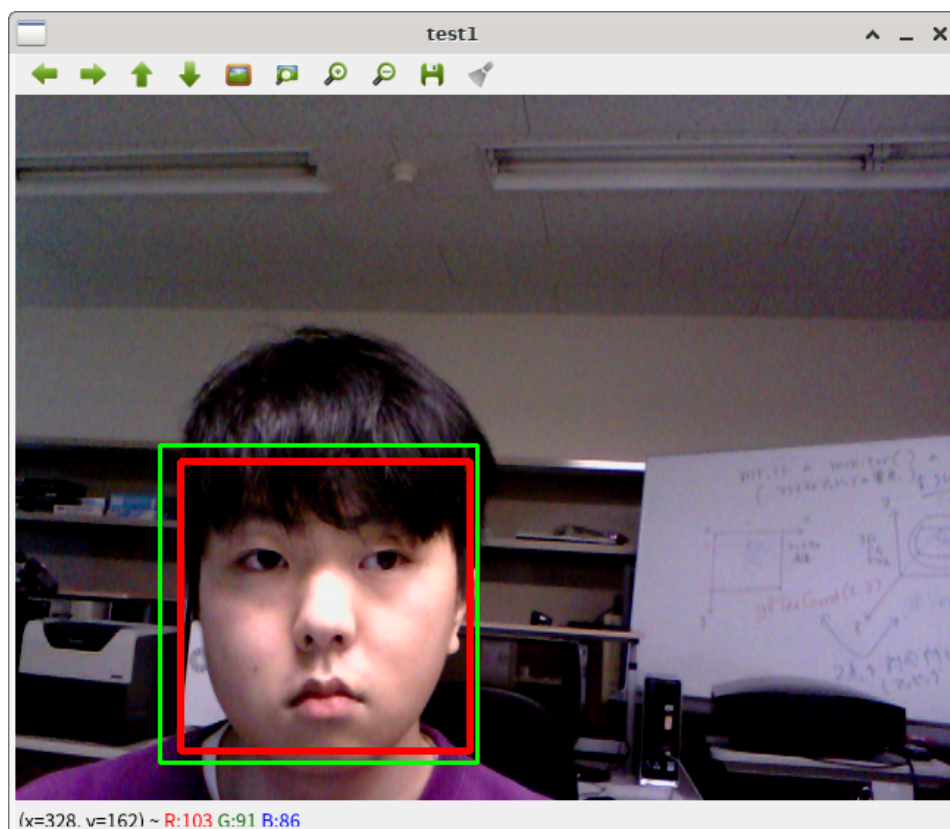


図 2: 顔検出の結果画像

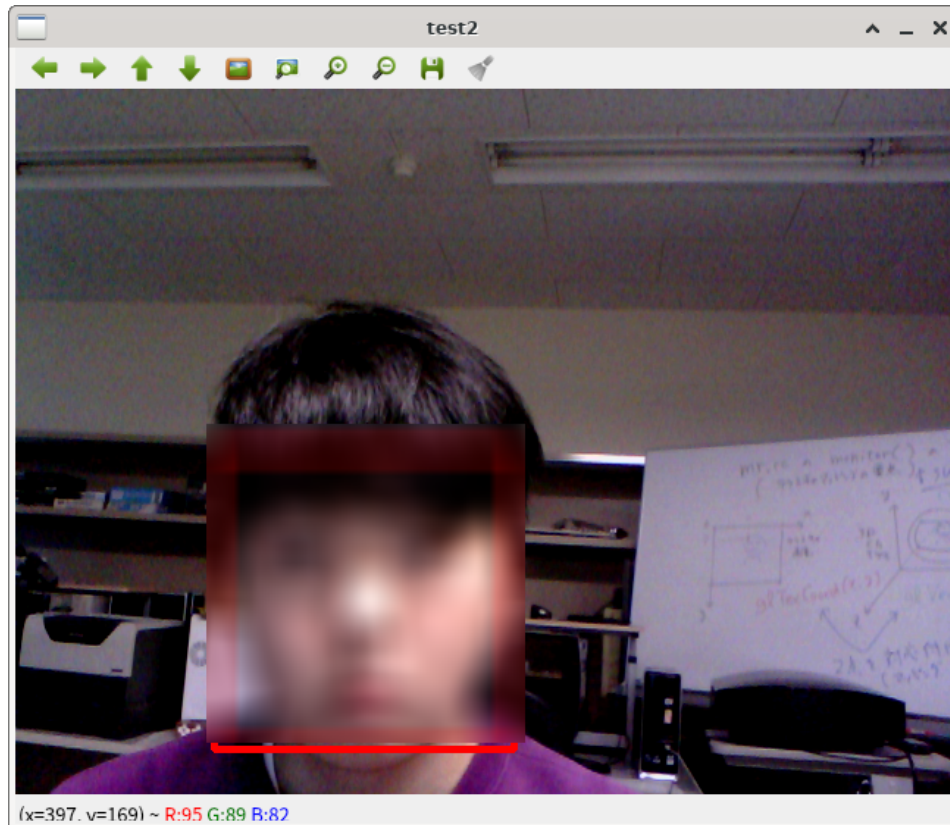


図 3: モザイク処理を行った画像

顔が検出された領域は緑色の枠、モザイク処理を施しているのが赤色の枠である。この画像を見ると、顔認識によって抽出された領域が頭部ではなく顔のみになっていることがわかる。この結果から頭部領域の抽出は顔認識だけでは実現できないことがわかったので、別の機能を利用する必要があることがわかった。

6 深度画像取得の確認

顔認識を行った結果から頭部領域を抽出するために Kinect を用いて取得した深度画像を利用する。深度とは深度センサーによって取得されたセンサーからの距離であり、深度画像は画像全体に写っている物のカメラからの距離を計測したデータである。深度画像を利用するにあたって、取得される情報がどのような情報なのか確認をする必要があるため、深度画像を取得し、距離ごとに色をつけた画像を生成するプログラムを作成した。



図 4: 取得した深度画像

図 4 は実際に深度画像を取得した結果をもとに、深度ごとに色をつけた画像であるこの画像ではカメラに近い部分を赤色、遠い部分を青色で表示しており、黒い部分は検出ができなかった部分である。この画像を見ると、カメラに近い位置にいる人間の体は赤色、背景などのカメラから遠いものは青色に近い色として表示されている。この結果から、深度画像を利用し、顔の座標の深度を基準として深度画像を取得することができれば、画像の背景除去を行えるのではないかと考えた。

7 顔認識と深度画像を利用した背景除去の実装

確認を行った物体検出機能と深度画像を利用し, 顔座標の深度を利用した背景除去を行うプログラムの作成を行った.

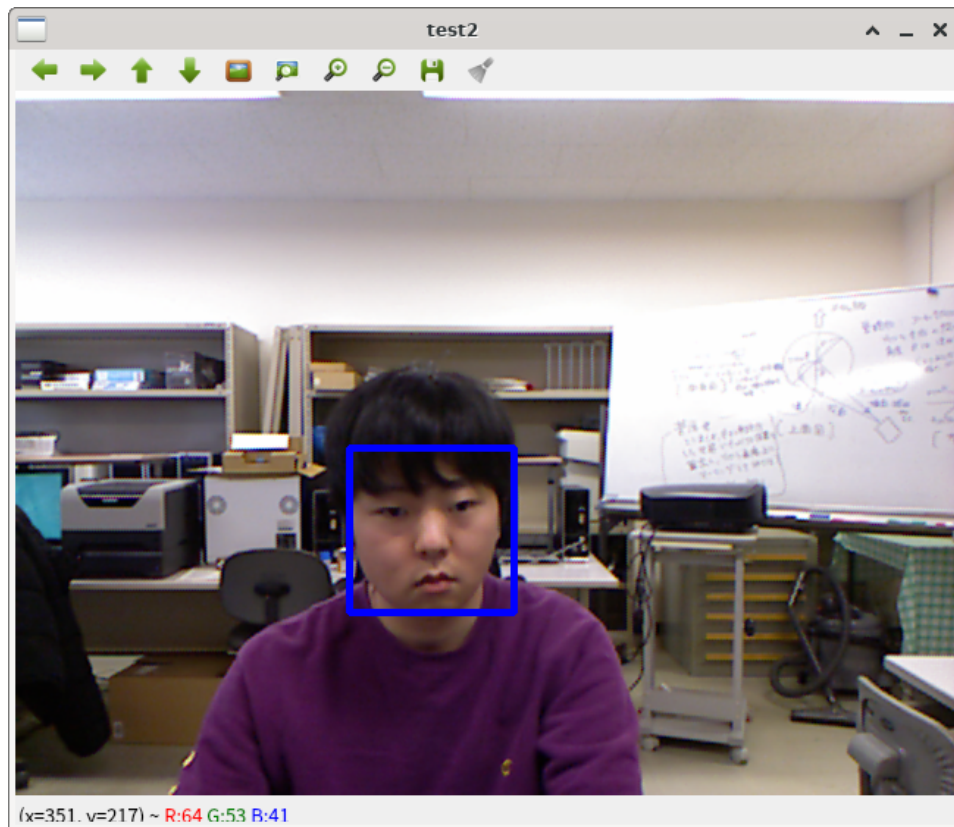


図 5: 取得した深度画像

図 5 は顔認識の結果である. 確認を行ったときと同様に顔を認識し, 四角形枠を表示しているが, このプログラムでは, それに追加で四角形枠の中心座標を別に取得している.

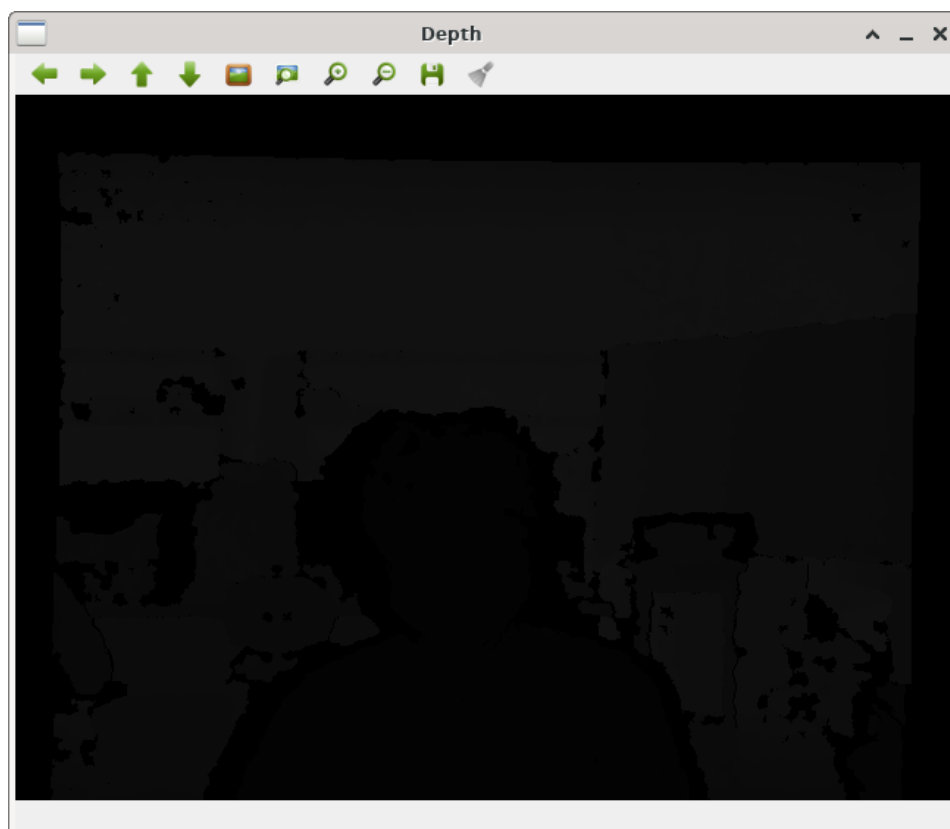


図 6: 取得した深度画像

図 6 は図 5 の画像と同時に取得している深度画像である。この画像によって画像全体の深度と、検出を行った際に取得していた四角形枠の中心座標の深度を取得している。

そして取得した四角形枠の中心座標をもとに画像から認識した人間の顔が存在する深度の算出を行い、算出された深度から人間の体があると考えられる深度の推定を行う。

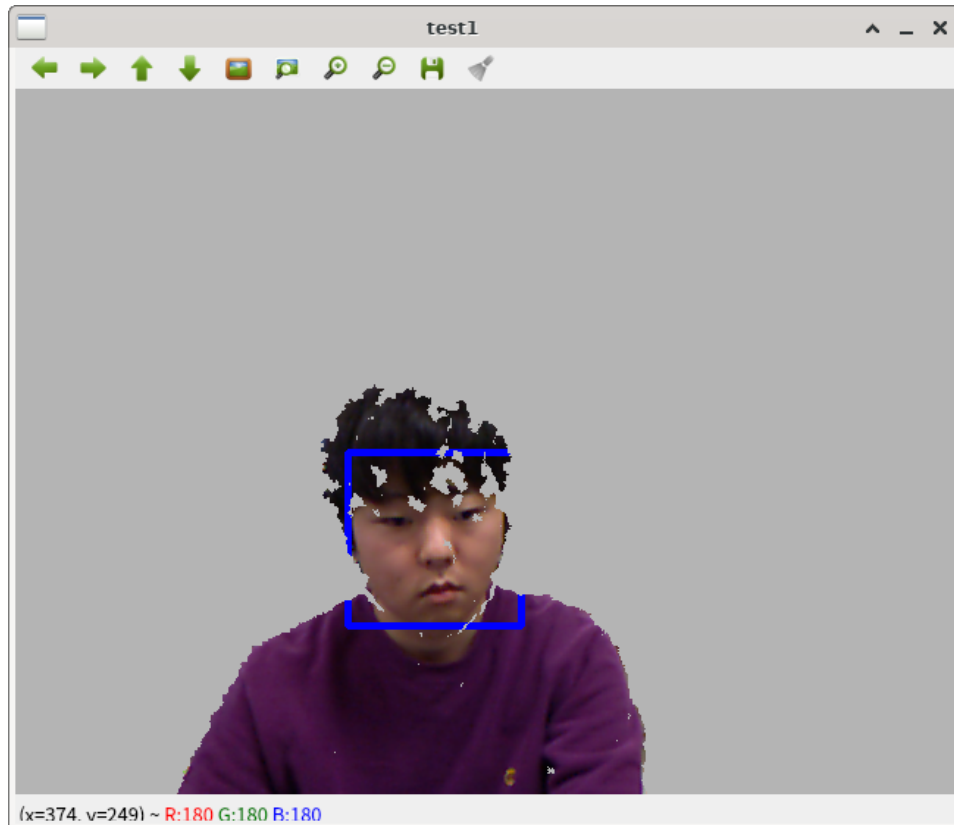


図 7: 背景除去の結果画像

図 7 は推定した深度をもとに背景の除去を行った結果である。背景除去は推定された深度と画像全体の深度を比較し、カメラからの距離が近すぎる部分と遠すぎる部分の座標の画像の除去を画像処理によって行った。その結果として図 7 のような人間の体のみを残した画像の生成を行うことができた。しかし Kinect の性能上、髪の毛などのツヤのある部分や顎などの深度が急激に変化する部分の深度検出の精度が良くないため、画像のような抽出された領域が一部にノイズがかかってしまったような表示になっている。

8 今後の課題

本研究では顔検出と深度センサの併用による背景除去を施すことで、映像から人間のみを抽出することができた。今後の課題として、現在抽出されている領域が頭部のみではなく体全体なので現在抽出されている領域と追加の機能を利用して頭部のみの領域抽出を行う必要があること、Kinect の性能の問題で安定していなかった髪や顎などの境界が曖昧な部分の検出安定化を図るために、画像処理のノイズ除去の利用などを検討する必要があること、抽出された頭部領域の拡大処理や拡大処理後の合成画像作成などが挙げられる。

9 おわりに

OpenCV と深度画像を利用することで、リアルタイムで背景除去を行うことはできたが、抽出領域を頭部のみにすることや、領域への拡大処理などの動作を実装することができなかった。体の領域抽出についても、検出が安定していないため、今後は検出の安定化と抽出領域を頭部のみにする機能の実装を進める必要がある。

参考文献

- [1] ディズニー: “作品情報 | アリス・イン・ワンダーランド -映画- ディズニー”, <https://www.disney.co.jp/movie/alice-time/about.html>, 参照 Oct.2021.
- [2] サントリー: “サントリーウィスキー Q CM デュランデュラン”, <https://www.youtube.com/watch?v=tWxApUkrjVY>, 参照 Oct.2021.