

360° リアルタイム映像配信における 機械学習を用いた非被写体へのモザイク処理

Mosaic Processing for Non-Subjects Using Machine Learning in 360° Real-Time Video Distribution

東京工科大学 メディア学部 メディア学科 菊池研究室
遠藤 直人
指導教員 菊池 司, 研究協力者 Subroto Prasetyo

キーワード: 360° 映像, リアルタイム, 映像配信, 機械学習, モザイク処理

1. 研究背景及び研究目的

監視社会と呼ばれる街頭カメラを用いた対物対人認識を利用した防犯システムが海外では用いられている。

本研究では認識や人物認識を行うだけではなく、普段目にする機会の少ない画角を用いた映像へマスク処理を行い、アーティストのコンサートやテレビ番組の生中継を行う際にどう生かすことができるのかを分析していく。普通のカメラであれば作り手の見せたい部分のみを切り取って映像を作ることができるが、今回用いるのは 360° カメラである。360° 映像は Youtube や Twitter 上などで目にする事が多いが、この映像の画角を用いて、映像業界への新たな手法とその簡略化を行うものである。

2. 360° 映像へのモザイク処理

映像業界は撮影機材や編集ソフトの発展によって 360° という画期的な表現方法を確立してきている。NHK が実際にテレビ映像の放送と同時に 360° 映像の配信を行った際の調査結果レポート「公共放送による 360° 映像の VR 配信の意義 ～2020 年とその先に向けて～[1]」では「臨場感がある」「普段見られない部分が見えた」「受動的でない点が面白かった」などの意見が出ており、テレビ放送という分野は前向きに 360° 映像や画像を活用していくことは十二分にある。

しかしながら、街中でそのカメラを使用する際にその視野の広さから不必要なものが写り込んでしまうことがあるはずである。プライバシー上の問題から顔を撮影されたくない方、著作権的に移

すことのできない衣服や看板などのプリント、さらには撮影スタッフと現代社会において権利的な問題として苦勞している動画制作者は数多くいるはずである。

3. 関連研究

360° 映像に関する研究はごく少数である。映像表現としてのアプローチ方法としては、映像作品としての表現方法の一部であるにとらえて制作物を制作するというのが主である。宮脇ら[2]は 360° 動画をリトルプラネットなどの展開方法別による印象評価を行うことでより良い 360° 動画を制作するための客観的な理論を示している。



[2]宮脇氏の制作した映像作品

4. 実装方法



[3] システム構成

```
1 import numpy as np
2 import cv2
3 face_cascade = cv2.CascadeClassifier('/Users/ryunosuke/anaconda3/share/OpenCV/haarcascades/haarcascade_frontalface_alt2.xml')
4 cap = cv2.VideoCapture(0)
5 frame_count = 0
6 #画像のスケール
7 scale = 1.5
8
9 while (1):
10     ret, frame = cap.read()
11     if ret == True:
12
13         frame_count += 1
14         gray = cv2.cvtColor(frame, cv2.COLOR_BGR2GRAY)
15         faces = face_cascade.detectMultiScale(gray, scale, 5)
16
17         #顔認識のモザイク処理
18         for (x, y, w, h) in faces:
19             frame = cv2.rectangle(frame, (x, y), (x + w, y + h), (255, 0, 0), 3)
20             face = frame[y:y+h, x:x+w]
21             face_blur = cv2.GaussianBlur(face, (23, 23), 0)
22             frame[y:y+h, x:x+w] = cv2.cvtColor(face_blur, cv2.COLOR_BGR2RGB)
23
24         cv2.imshow('Face Detection', frame)
25
26         k = cv2.waitKey(1)
27         if k == ord('q'):
28             cv2.imwrite('frame_count.jpg', frame)
29         if k == ord('n'):
30             break
31
32     else:
33         break
34
35 cap.release()
36 cv2.destroyAllWindows()
```

[4] [3]のシステム図を使用した際のプログラム例

顔認識を元にしてモザイク処理を行っているため、顔認識自体の精度が高くない(OpenCVの場合はカメラに対して顔が向いていないと顔として認識されない)などの問題点により、Google Vision API という高度な画像分析を行うことのできる外部クラウドサービスを利用し追実験を行った。

5. Google Vision API を用いた画像処理

Python をベースとして、[3]のシステム図の OpenCV 部分を Google Vision API を変更した機械学習を用いた顔認識によるモザイク処理の実験を行った。Google Vision API とは、Google Cloud Platform が提供する画像認識サービスのことで、機械学習モデルを利用し対象とする画像を分析することで様々な情報を取得することができる。しかし、インターネットを経由して情報を取得してくるといったクラウドサービスの性質上、

本研究の目的の一つである「リアルタイム」に対してその真価を発揮させにくいという問題点が発生してくるのである。Google Vision API では情報の処理を PC 内のみで行うことができないため、映像自体に遅延が発生してしまう。4 秒おきに 1.5 秒程度の画面のフリーズが確認されており、顔部分の認識できる距離や角度は精度として上がるものの映像自体が定期的にフリーズしてしまうことは大きな問題である。

6. まとめと今後の展望

本研究では 360° コンテンツの更なる利用方法を提案すべく、360° リアルタイム映像配信における機械学習を用いた非対象物への自動検出及びモザイク処理を Python 上で OpenCV を用いた。顔認識の精度は 2~3m 圏内の範囲とあまり高くないもののリアルタイム性のある配信を行うことが可能であるという結果を確認することができた。

360° リアルタイム映像内の非対象物へのモザイク処理が今後の映像業界での 360° 映像配信という分野をより幅広く発展させるためのきっかけとなるよう、今後は顔認識の精度の向上及びライブラリ機能の拡張を視野に研究を進めていきたいと考えている。

7. 参考文献

[1] 山口勝 “公共放送による 360° 映像の VR 配信の意義~2020 年とその先に向けて~”

https://www.nhk.or.jp/bunken/research/domestic/pdf/20171001_4.pdf

[2] 宮藤巧真, 菊池可 “360 度映像における印象に及ぼす要素の研究”

[5] OpenCV documentation

http://opencv.jp/opencv-2.1/cpp/object_detection.html