

顔認識を用いた動画撮影支援のための構図判定モデル

塩山 昌志[†] 鎌原 淳三[†]
[†] 神戸大学大学院海事科学研究科

1. 概要

カメラ機能付きスマートフォンの普及により、静止画のみならず動画もインターネット上で公開されたり中継されたりすることが増えている。撮影者の多くは素人であるため、映像の質が高くないものも多い。原因には専門家であるカメラマンが持っている撮影に関する知識・ノウハウ[1]を非専門家である一般ユーザーが持ち合わせていないということが考えられる。本稿では、従来研究[2]に加えて人物の構図に着目し、映像フレーム内の顔位置が適切になるよう支援するのに必要な構図判定のモデルを提案する。実験で専門家であるプロのカメラマンが撮影した映像から被写体の顔の大きさや顔の中心位置の関係を調べ、構図判定モデルを構築した。このモデルに基づいて別の映像を評価したところ、対象フレーム数の96%以上がモデルの範囲内に収まっており、モデルに従うことが確認できた。

2. プロの映像の顔位置

本研究ではAndroidスマートフォンの顔認識技術を用いるため認識できる人の顔が正面で全部写っていることが条件となる。認識される顔領域は正方形である。今回は映像の被写体が1人か2人のときを対象としそれぞれ1Face, 2Faceとする。TVのトーク番組から抽出した顔面積のヒストグラムから、頻度上位5位までの顔面積区間ごとに顔領域中心のy座標(以下、顔y位置と呼ぶ)の平均を求めたところ、顔面積が大きくなるにつれて顔領域中心のy軸の位置が直線的に下がることがわかった。顔y位置の平均直線を中心直線とする。

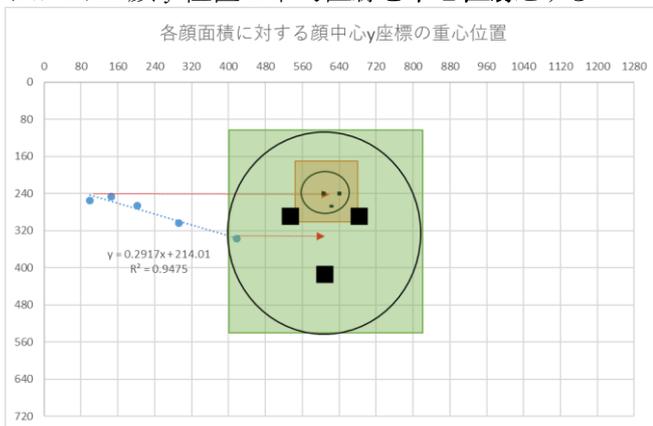


図1: 1Face 頻度上位5位までの顔面積区間ごとの顔領域中心y座標の平均の位置の変化

3. 構図判定モデルの提案

実験より1Face, 2Face共に顔y位置の中心直線からの距離はほぼ正規分布であった。±2σの範囲のデータのみをプロットしたものが図2である(1Face)。

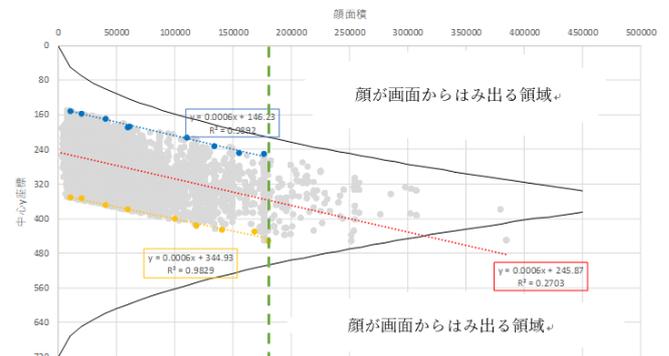


図2: 1Face 中心直線からの距離が±2σの範囲のプロット

ここから顔幅が顔面積の1割から6割までは赤色の中心直線を画面の高さhを用いて式(1)とし、モデルとする。ただしxは顔面積,yは顔y位置,minFは認識できる最小の顔幅の1/2である。

$$y = ax + \left(\frac{h}{3} - a * \min F\right) \quad (1)$$

$$\left(a = \frac{0.5h - 0.33h}{(h+0.6)^2 - (h+0.1)^2}\right)$$

6割以降では、顔が画面からはみ出ないという制約の方が強くなるためy=h/2で一定とする。2Faceの場合は顔幅が1割から4割のところは1Faceと同様の結果が出たが、それ以降のところは2人の顔同士の位置関係などの要素も考慮しなければならないと思われる。

4. 評価とまとめ

実験とは全く別のTV番組から抽出した顔データを用いて検証を行った結果、対象フレームに対して1faceで96.5%, 2Faceは左右それぞれ97.2%, 96.2%のデータがモデルの範囲内に含まれた。これによりモデル式が別の映像でも成立することが確認できた。

モデルに基づいて撮影中の顔面積における適切な顔y位置を撮影者に示すことで、より適切な構図となる撮影支援を行うことができると考える。

参考文献

- [1] 益子広司・内田一夫(2013)映像カメラマンのための構図完全マスター, 玄光社出版
- [2] 堀川(2016)「顔認識を用いた構図判定を行う映像撮影支援機構の開発」, 神戸大学大学院海事科学研究科修士論文