

符号化歪みを有する動画像に対する動体検知

田中哲哉† 中村海斗† 氏三浦康之†

湘南工科大学

1. はじめに

監視カメラは施設や市街地等に設置されており、個人でも監視カメラを所有し自宅などに設置する事も可能である。また、動体検知機能を備えた物も存在する。本研究では、監視カメラで撮影したものをサーバに転送し保管する場合や、カメラ本体の画質や保管領域の都合で圧縮率の高い動画を扱わざるを得ない場合を想定する。

動画を圧縮すると画質が劣化してしまうが、目には見えにくい情報から削減していくことから一部の動体検知法への影響が少ない可能性がある。そこで、フレーム間差分法を対象とし、圧縮した動画像への動体認識精度の検証を行い、精度が向上する条件を発見することを本研究の目的とする。また、以前行った研究から動体検知の閾値を下げてノイズ除去を強くかければ精度が向上する可能性があることが判明したため、今回はその検証を行う。

2. 画像処理

2.1 フレーム間差分法

フレーム間差分法とは、動画像から動体を検知する手法の一つである。動画像はいくつかの静止画像を連続して表示する事で動きを表現している。フレーム間差分法は静止画像の1フレームに着目し、その直前のフレームと比較して差分を抽出する事で動体を検出する手法である。この際映像は二値化される。

どの程度の差までを検出するか決定する値を閾値と呼ぶ。閾値が高いほど動体検出の精度が上がり、閾値が低いと検出精度が落ちる代わりに情報量が増える。

2.2 メディアンフィルタ

メディアンフィルタとは、ノイズを除去する画像処理フィルタの一つである。

注目したピクセルの周りから画素値を探索して整列し、整列したものの中から中央にある値と着目したピクセルの画素値を置き換える。エッ

ジを保持したまま画像のノイズを除去できるため、フレーム間差分法と相性が良い。

3. 動体検知性能の比較

3.1 比較手法

動画像の圧縮率とノイズ除去フィルタを適用することで動体検知精度にどのような変化が起きるか把握するため、定量的な評価を行う必要がある。評価をする際、非圧縮動画像にメディアンフィルタを適用したものと、圧縮率やメディアンフィルタの強度を変えたものを比較し、F値という尺度で測定する。

3.2 F値

F値とは、機械学習において性能評価を行う際に用いられる尺度の一つである。

precision と recall の両方がバランスよく高い箇所を求めた数値をF値と呼ぶ。

• precision

precision とは、F値を求める際に利用される尺度の一つである。プレジジョンが高いほど誤検出が少なくなるが、情報量が少なくなっている傾向がある。適合率とも呼ばれている。

• recall

recall とは、precision と同じくF値を求める際に利用される尺度の一つである。

recall が高いほど情報量が増えるが、誤検出が多くなる傾向がある。precision とはちょうど反対の関係にある。

$$precision = TP / TP + FP \quad (1)$$

$$recall = TP / TP + FN \quad (2)$$

$$F \text{ 値} = \frac{2 \times precision \times recall}{precision + recall} \quad (3)$$

TP 正確に検知されたピクセルの総数

TN 動体のうち、検知されなかったピクセルの総数

FP 非動体のうち、検知されたピクセルの総数

FN 非動体のうち、検知されなかったピクセルの総数

4. 実験

4.1 実験形式

4000kbps に圧縮した動画像を閾値 20 でフレーム間差分法にかけメディアンフィルタを適用し

Inter-frame Differencing Method for Compressed Moving Images

†Tetsuya Tanaka, Kaito Nakamura, Yasuyuki Miura
Shonan Institute of Technology

た動画を正解データとし、検証データにはそれぞれ

2000kbps,1000kbps,700kbps,500kbps,300kbps,200kbps,100kbps に圧縮したものを、閾値を 5, 10, 15, 20, 25 としてフレーム間差分法を使用した動画像を用意し、メディアンフィルタの範囲を 5×5、7×7、9×9 動画像をそれぞれ用意する。

今回の実験では人物が映っている動画像を対象とした。実験に用いた動画像を図 1 に示す。



図 1 実験に使用した動画像

表 2 正解画像よりメディアンフィルタサイズが小さい動画像の閾値別 F 値一覧

メディアン 5×5				
閾値別 F 値				
Bps	閾値 10	閾値 15	閾値 20	閾値 25
2000kbps	0.559	0.758	0.835	0.804
1000kbps	0.561	0.747	0.814	0.782
700kbps	0.555	0.727	0.789	0.755
500kbps	0.55	0.689	0.745	0.717
300kbps	0.55	0.636	0.665	0.637
200kbps	0.533	0.56	0.56	0.531
100kbps	0.372	0.379	0.382	0.371

表 3 正解画像と同じメディアンフィルタサイズで実験した動画像の閾値別 F 値一覧

メディアン 7×7				
閾値別 F 値				
bps	閾値 10	閾値 15	閾値 20	閾値 25
2000kbps	0.615	0.809	0.885	0.807
1000kbps	0.624	0.801	0.856	0.783
700kbps	0.624	0.781	0.828	0.754
500kbps	0.617	0.74	0.781	0.719
300kbps	0.609	0.675	0.692	0.639
200kbps	0.565	0.58	0.573	0.53
100kbps	0.384	0.385	0.384	0.371

表 4 正解画像よりメディアンフィルタサイズが大きい動画像の閾値別 F 値一覧

メディアン 9×9				
閾値別 F 値				
bps	閾値 10	閾値 15	閾値 20	閾値 25
2000kbps	0.615	0.821	0.859	0.757
1000kbps	0.624	0.815	0.837	0.738
700kbps	0.624	0.8	0.815	0.712
500kbps	0.617	0.765	0.774	0.685
300kbps	0.609	0.699	0.694	0.614
200kbps	0.565	0.591	0.572	0.518
100kbps	0.384	0.387	0.381	0.364

4.2 実験結果

行った実験の結果を表 2 から表 4 に示す。

2000kbps に圧縮した動画像のうち最も高い F 値が 0.885 となる一方、300kbps に圧縮した場合に最も高い F 値が 0.699 となった。このことから、大きく劣化した画像においてもある程度正しい検知が可能であることが分かる。また、大きな動体が映っている動画像について、高い圧縮率による圧縮を行った場合メディアンフィルタのサイズを大きくして閾値を低く設定する事により、F 値が改善する事が分かった。

大きな動体が映っている動画像においては、recall に比べて precision が下がりにくいため、閾値を下げて、より多くの動体候補を検出したうえで、サイズの大きなメディアンフィルタにより強めのノイズ除去を行う手法が有効であったものと考えられる。

なお、小さな動体に対しては、この手法が必ずしも有効ではないことも確認されている。

5. まとめ

本実験により、圧縮した動画像に動体検知をかけても動体検知精度が大きく劣化しない場合があることがわかった。また、動体が大きい場合は閾値を下げてノイズ除去フィルタを大きくかけることで精度が上がる事が発見された。

参考文献

[1] Tetsuya Tanaka, Yasuyuki Miura, “Study on Inter-frame Differencing For Compressed Moving Images, IEEE International Conference on Consumer Electronics - Taiwan (IEEE2017 ICCE-TW), 2017.06