

3次元コードにおける動画像符号化用フィルタ処理の検討

坂田 敬悟†

杉浦 彰彦‡

静岡大学 情報学部科学科†

静岡大学 情報学部科学科‡

1. はじめに

近年携帯電話の普及に伴い、その性能も格段に向上している。その中の 1 つに QR コードを読み取る技術がある。既に QR コードのような様々な 2 次元コードの研究が行われている。しかし、2 次元コードには情報量に制限がある。その問題を解決するため、近年 3 次元コードについて研究が行われている。

3 次元コードにはいくつか種類があるが[1][2]、本研究ではコードに時間軸を持たせた 3 次元コードについて研究を行う。まず原動画像に対してフィルタ処理を行う。これを本研究では前処理フィルタと呼ぶ。続いて動画像符号化後にフィルタ処理を行う。このフィルタのことを後処理フィルタと呼ぶ。この 2 つのフィルタの最適な組み合わせを見つけ、認識率の向上を図ることを本研究の目的とする。

2. 原理

2.1 マーカ

3 次元コードを用いるにあたって、本研究ではマーカにキャリブレーションマーカ (図 1) を用いる。このマーカは照明や撮影角度に強く、ロバスト性に優れている。また本研究で用いるマーカはそれ自体が数値情報を持つ[3]。実験はこのマーカを正しく認識できたかを調査する。また実験用動画像は、適当なマーカをランダムに抽出したものをを用いる。

2.2 フィルタ

本研究では積分型のフィルタと微分型のフィルタを用いる。積分フィルタを用いると画像全体が滑らかになる。逆に微分フィルタを用いると、画像中のエッジ部分が強調される。積分フィルタには単純平滑化(sim)、メディアン法(med)を用い、微分フィルタには Sobel オペレータによるエッジ保存平滑化(sob)、テンプレート

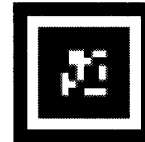


図 1: キャリブレーションマーカ

を用いたエッジ保存平滑化(tem)をそれぞれ用いる。これらのフィルタ処理を施し符号化を行った場合、情報量は“積分フィルタ<微分フィルタ”という関係が成り立つ。本研究では情報量を 1 つの基準とし、この関係性を利用したフィルタ処理についても着目し研究を行う。

2.3 符号化

原動画像を符号化する手法として、本研究では「H.264/AVC」を用いる。この H.264 では、符号化に I, P, B フレームという 3 つのフレームを用いる。順方向予測を用いるフレームを P フレーム、双方向予測を用いるフレームを B フレーム、フレーム間予測を用いないフレームを I フレームという。3 次元コードの認識には、B フレームが認識率に悪影響を及ぼすことが分かっている[4]。したがって H.264 符号化を行う際に、I および P フレームのみを用いて符号化を行う。また、Deblocking filter も 3 次元コード認識に対して悪影響を及ぼすため使用しない。

3. 実験

実験では 3 次元コードを低ビットレートで符号化した後、各フィルタ処理が認識率にどう影響したかを調査する。手順として、はじめに撮影した 3 次元コード (原動画像) に対して前処理フィルタを施す。続いて H.264 で低ビットレート符号化を行う。符号化後のファイルサイズは 46500 キロバイト前後とし、実験動画像全体でファイルサイズの誤差を上下 5%以内に抑える。また、画質統一のため 1 つの動画像間で生成される動画像のファイルサイズの誤差は、上下 2%以内とする。最後に、後処理フィルタを施し、認識率調査を行う。なお今回の実験では、秒間 30 フレームで 10 秒間、サイズ 320×240pixel の 3 次元コードを用いる。

Filtering Approach on Video Encoding by Three Dimension Codes.

†SAKATA Keigo, Shizuoka University. Faculty of Infomatics Department of Computer Science.

‡SUGIURA Akihiko, Shizuoka University. Faculty of Infomatics Department of Computer Science.

3.1 認識率の良い動画像に対する実験

撮影された原動画像をフィルタ処理なしで H.264 符号化を行い、認識率の良い動画像 10 本に対して実験を行う。前提条件として、無圧縮の状態では認識率 100%、符号化後に認識率 90% 以上の結果となるものを採用する。実験の結果、どのフィルタ処理の組み合わせが効果的であったかを調査する。

3.2 認識率の良くない動画像に対する実験

3.1 と同様の実験を別の動画像 10 本に対して行う。この結果と 3.1 の結果を比較することで、元々の認識率に応じてフィルタ処理の組み合わせがどう変わっていくかを調査する。

4. 結果

4.1 認識率の良い動画像に対する実験結果

実験動画像 10 本に対し、フィルタ処理を施し認識率調査を行った。フィルタ組み合わせのパターンごとに平均を取り、まとめた結果を図 2 に示す。図の縦軸は認識率、横軸はフィルタの組み合わせを表している。フィルタ組み合わせは、“non” はフィルタ処理なしを表し、それ以外は各フィルタの頭文字で表す。例えば、“sim-med” は前処理に単純平滑化、後処理にメディアン法を用いたことを表す。最も認識率の良い結果となったのは前処理に単純平滑化、後処理に単純平滑化を用いた場合となった。フィルタ処理なしのパターンと比べると 3% 以上の認識率向上が見られた。また、前処理に Sobel オペレータによるエッジ保存平滑化を用いると認識率が減少する傾向が見られた。

4.2 認識率の良くない動画像に対する実験結果

4.1 と同様に平均でまとめた結果を図 3 に示す。グラフの見方は図 2 と同様である。こちらは 4.1 の結果とは異なり、前処理に単純平滑化、後処理にテンプレートを用いたエッジ保存平滑化を用いた場合が最も良い結果となった。フィルタ処理なしのパターンと比べると 5% 以上の認識率向上が見られた。また 4.1 と同様、前処理に Sobel オペレータによるエッジ保存平滑化を用いると認識率が減少する傾向が見られた。

5. まとめ

4.1 の実験結果から、前処理後処理ともに積分フィルタを用いた場合に認識率の良い結果となった。これは認識率がある程度良い場合、微分してエッジを強調するよりも、2 重に積分することでノイズを抑えた方が認識率改善につながる

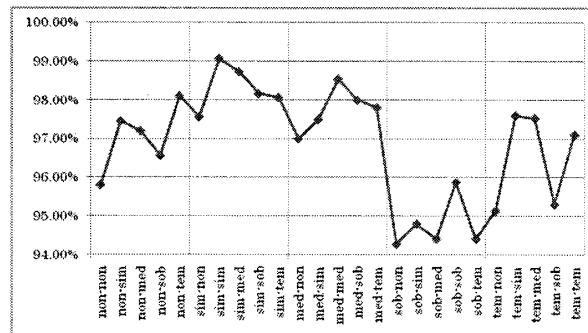


図 2：認識率の良い動画像に対する結果

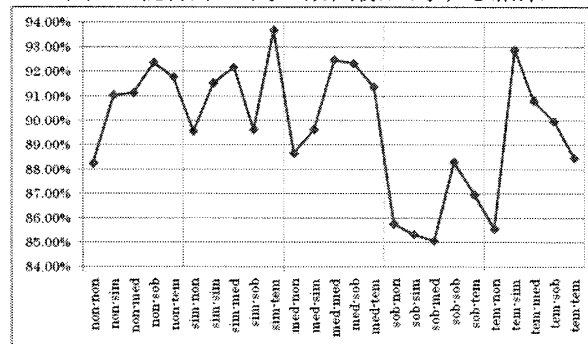


図 3：認識率の良くない動画像に対する結果

ためだと考えられる。逆に 4.2 のような場合には、前処理に積分フィルタ、後処理に微分フィルタを用いた場合に認識率の良い結果が得られた。これは、積分フィルタで符号化後の情報量を下げ、その分符号化時のビットレートを上乗せし、微分フィルタでエッジを強調することで認識率向上につながったためだと考えられる。

6. 今後

今回の実験では比較的簡単なフィルタを用いて実験を行った。今後はさらに効果的なフィルタ選択のため、新しいフィルタを作成し実験を行っていく。また、時間軸に考慮したフィルタリング手法である、3 次元フィルタについても研究を行っていく。

Reference

- [1] 神原誠之, 横矢直和, “現実環境の照明条件と奥行きの実時間測定による仮想物体の陰影表現が可能な拡張実感” 画像の認識・理解シンポジウム講演論文集 (MIRU2004), Vol. II, pp. 247-252, Jul. 2004.
- [2] 宇田隆哉, 伊藤雅仁, 淡谷浩平, 重野寛, 松下温, “三次元パターン認識を用いた携帯型端末向け電子チケットシステム” 情報処理学会研究報告, Vol. 2002, No. 12 (2001-CSEC-016), pp. 133-138, Feb. 2002.
- [3] 久保人志, 杉浦彰彦, “楽譜情報の 3 次元コード化と携帯電話を利用した再生手法” 電気学会論文誌, C Vol. 128, No. 10, pp. 1582-1588, 2008.
- [4] 田中琢也, 杉浦彰彦, “低ビットレート符号化が動画認識に与える影響調査と対策改善” 豊橋技術科学大学修士学位論文 2008 年