

ビデオ高速探索における画像特徴量と音響特徴量の比較*

芳賀鉄平, 杉山 雅英 (会津大学 大学院)

1 まえがき

蓄積された長時間の映像信号, 音響信号やビデオ信号の特徴時系列から指定の信号を高速に探索することを目的として様々な手法が検討されてきた [1]. 我々は高速探索法として DSCC 法 [2] を提案し, 画像情報や音響信号における Active 探索からの改善率などを示した [?]. 本報告では探索性能を F 値を用いて評価し適切な探索条件について述べる.

2 評価実験の枠組

2.1 探索手法

セグメント間の距離を VQ 符号帳により作成した出現確率ベクトル間の l_1 距離で定義し, 探索閾値 θ 以下であれば目的セグメントとする. 本報告では DSCC 法の区分化中心判定法と距離ブルーニング法を用いる. 探索の手順は蓄積ベクトルを区間に区分化し, その代表ベクトル間の距離行列を用いて判定を行い, 判定できない場合に Active 探索を用いて判定を行う.

2.2 F 値

雑音状況下では正解位置を正しく検出できるとは限らない. 例えば正解位置を検出できない場合もあるし, 逆に検出した位置が正解でない場合がある. そこで Recall(式 (1)), Precision(式 (2)), F-measure(式 (3)) によって探索の性能を評価する

$$\text{Recall} = \frac{\text{Number of detected correct sections}}{\text{Number of correct sections}} \quad (1)$$

$$\text{Precision} = \frac{\text{Number of detected correct sections}}{\text{Number of detected sections}} \quad (2)$$

$$F - \text{measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3)$$

2.3 実験条件

評価実験に用いた蓄積信号は 45 分のテレビ番組であり, 探索セグメントは 7.5 秒とした. 画像情報はフレームレート 29.9frames/s, ピクセルサイズ 704×480, MPEG-2 圧縮された映像から各フレームを文献 [3] を参考に 16 分割し RGB 濃度を平均し正規化し 16 次とした. 音響情報の分析は標準化周波数 16kHz, 窓長 32ms, フレーム更新周期 32ms とし LPC ケプストラム (CEP) と MFCC の 2 種類を使用する. CEP は LPC 分析 14 次, ケプストラム分析 16 次を用い, MFCC も 16 次の特徴を用いる. 探索セグメントは時間的に等間隔に選んだ 30 箇所とした. VQ 符号帳は LBG 法で作成し大きさは 32, 64, 128, 256, 512 とした. 評価実験では Pentium4(3.2GHz), メモリ 2GB, OS Vine Linux 3.2 を載せたコンピュータを用いた.

3 評価実験

画像情報, 音響情報を用いて探索性能の比較を行う. 雑音環境下の評価実験では 2 種類の雑音, 即ち MPEG-2 形式ビデオを再生してビデオテープに録画し再度 PC に MPEG-2 として取り込んだ信号, MPEG-2 を MPEG-1 に変換し画像情報, 音響情報のビットレートを半分に削減した信号で実験を行う. 実験結果は両雑音とも同様の傾向を持つので両雑音で実験を行った結果の平均値を示す.

3.1 無雑音信号の探索

表 1 に探索閾値 θ を 0.10 に固定し VQ 符号帳の大きさを变化させた結果を示す. 音響情報は CEP を用いた結果を示すが MFCC を用いても探索性能に大きな変化はない. また, 探索では DSCC 法と Active 探索法を組み合わせ探索を行っており, Active 探索法では距離を計算するたびに VQ 符号帳を用いて量子化をしている. よって, Active 探索で探索する場合に特徴量の次元に比例して探索処理時間が遅くなる [4]. 雑音がない環境での探索では特徴量の次元が小さく距離計算回数の少ない音響情報を用い, サイズ 32 の VQ 符号帳を用いるのが有効である.

表 1: DSCC 法における画像情報と音響情報の比較
距離計算回数, 探索処理時間

VQ 符号帳	画像情報		音響情報	
	計算回数	時間 (ms)	計算回数	時間 (ms)
32	32.7	5.3	42.1	5.0
64	45.5	8.0	36.4	5.0
128	81.7	17.3	40.2	8.3
256	124.4	35.3	44.1	15.0
512	178.4	61.6	50.0	28.6

平均 l_2 歪み区分化法, 区分化閾値 $\varepsilon = 0.00005$

3.2 雑音環境下の探索

3.2.1 雑音環境下での各特徴量の性質

図 1 は蓄積ベクトルと雑音付加ベクトルを時間的に等間隔に 8000 点取りそのベクトル間の出現確率ベクトル間の l_1 距離の平均値である. 結果は両雑音で実験した結果を平均したものである. この値が小さい程, 雑音の影響が小さいことを示している. 音響情報より画像情報の方が距離が小さいこと, VQ 符号帳の大きさが大きくなる程音響情報に対しては距離が大きくなること, また, 音響情報の CEP と MFCC を比較すると MFCC の方が若干距離が小さい. これは MFCC は低域周波数を拡大し, 高域を圧縮しているため雑音に対して頑健である. 評価実験では音響特徴量として雑音に頑健な MFCC を用いる.

3.2.2 F 値による評価

画像情報と音響情報 (MFCC) の各 VQ 符号帳における F 値と探索閾値 θ の関係を図 2, 図 3 に示す. 両特徴

*Comparison between Image and Sound Features in A Quick Video Retrieval Algorithm, by T.Haga, M.Sugiyama (Graduate School, The Univ. of Aizu)

表 2: F 値を最大とする探索閾値とその処理時間
30 セグメント探索の平均値, 処理時間

VQ 符号帳	画像情報				音響情報 (MFCC)			
	探索閾値	F 値	距離計算回数	処理時間 (ms)	探索閾値	F 値	距離計算回数	処理時間 (ms)
32	0.1	0.882	31.5	4.6	0.3	0.914	119.7	9.0
64	0.1	0.922	44.9	7.3	0.4	0.921	195.2	18.3
128	0.1	0.926	80.8	15.8	0.5	0.914	272.8	27.0
256	0.1	0.926	123.8	33.1	0.6	0.921	381.5	51.3
512	0.1	0.954	177.9	61.8	0.8	0.893	586.5	114.0

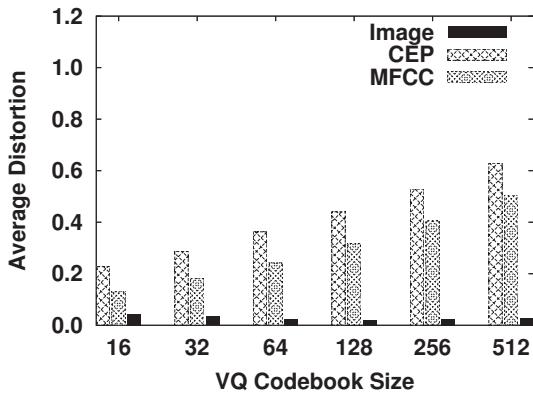


図 1: 蓄積信号と雑音付加信号の平均 l_1 距離

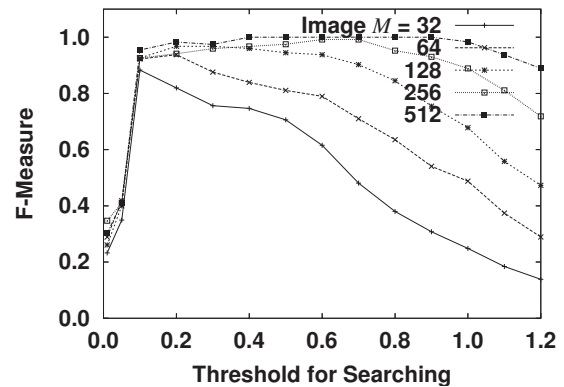


図 2: 音響情報に対する F 値と探索閾値との関係

量ともに探索閾値 θ が小さい場合は目的信号を検出できず F 値が低い, また探索閾値 θ を大きくすると目的信号でない箇所を誤検出してしまうので F 値が低くなる傾向にある. 画像情報と音響情報を比較すると画像情報の方が F 値が高い. これは図 1 に示した通り画像情報の方が蓄積信号と雑音付加信号との距離が小さいからである. 画像情報では VQ 符号帳が大きくなるにつれ F 値も大きくなる傾向にあり, 音響情報では VQ 符号帳のサイズが大きくなるにつれ F 値を最大にする探索閾値 θ の値が大きくなる. 画像情報に対しては探索閾値 θ をほぼ一定に設定できる. 画像情報は探索閾値を 0.10, 音響情報は最大の F 値を与える探索閾値での DSCC 法の結果を表 2 に示す. F 値も高く, 探索時間, 距離計算回数から画像情報が探索に適している.

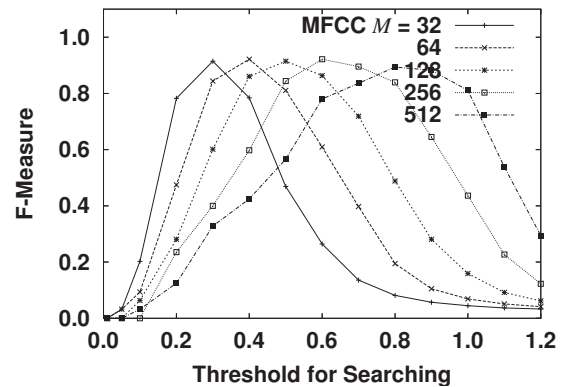


図 3: 音響情報に対する F 値と探索閾値との関係

4 むすび

目的信号に雑音が付加されていない環境では特徴量の次元が低い音響情報を用いて小さいサイズの VQ 符号帳を用いるのが有効である. 雑音環境下での探索では音響情報より画像情報の方が雑音信号との距離が小さく, F 値も高く探索処理時間も短い. 画像情報の中で高い F 値を得るためには処理時間がかかるがサイズ 512 の VQ 符号帳が探索に適している. また, サイズ 64 の VQ 符号帳を用いた場合でも十分 F 値が高く処理速度も速い. 今後の課題として雑音環境下での頑健な探索法の検討がある.

参考文献

- [1] G.A.Smith, et al, "Quick Audio Retrieval Using Active Search", Proc. of ICASSP98, Vol.6 p.3777(1998).
- [2] 杉山, 岡本, "距離空間と出現確率時系列の幾何学的性質に基づくセグメント高速探索法", 情報処理学会論文誌, (2006)
- [3] K.Kashino, et al, "A quick video search method based on local and global feature clustering", Proc. of ICPR, vol.3, pp.894-897 (2004-04).
- [4] 芳賀, 杉山, "ビデオ検索における映像情報と音響情報の比較", 情報処理学会東北支部研究会, 5-3-05 (2006-01).