1P-5

高速セグメント探索法によるビデオ検索 *

– 映像情報と音響情報の比較 –

○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] を提案しその有効性を示した．本報告では長時間のビデオデータからの探索を目的として，探索における VQ 符号帳の大きさの影響及び雑音に対する耐性における映像情報と音響情報の性質の違いを述べる．

2 探索手法と評価実験の枠組

セグメント間の距離を VQ により作成した出現確率ベクトル間の l_1 ノルムで定義し, 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 M1,M3 及び Active 探索を用いる．

- M1：区分化中心判定法
  区間に区分化し代表ベクトル $\boldsymbol{c}_i$ を用いて判定
- M3：距離プルーニング法
  代表ベクトル j, k 間の距離 $d_1(\boldsymbol{b}_j, \boldsymbol{b}_k)$ を用いて判定
- AS：Active 探索法
  M1+M3 で判断できない場合に AS を用いて判定

2.1 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．映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 [4] を参考に 16 分割し RGB 濃度を平均し正規化した．音響情報の分析は標本化周波数 16kHz, 窓長 16ms, フレーム更新周期 32ms, 窓関数はハミング窓, LPC 分析 14 次, ケプルトラム分析 16 次さらに Δ ケプストラム 16 次, Δ^2 ケプストラム 16 次とした．探索セグメントは等間隔に選んだ 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．実験は Pentinum4(3.2GHz), メモリ 2GB, OS Vine Linux3.2 を載せたコンピュータを用いた．

2.2 映像情報と音響情報の違い

図 1 に映像情報と音響情報の探索セグメントと全ての蓄積ベクトルとの出現確率ベクトル間の l_1 距離 ($\sum_{m=1}^{M} |p_m - q_m|$) の平均値とその標準偏差を示す．$l_1$ 距離は 2 以下の性質がある．これらは探索の際の距離計算回数を決める要素となる．映像情報は VQ 符号帳の大きさにかかわらず平均距離が大きく標準偏差も小さい, 音響情報は VQ 符号帳の大きさに依存して平均距離が大きくなることが判る．一般的に VQ 符号帳の大きさが大きくなると平均距離も大きくなるが, これは出現確率ベクトルの要素毎の差を足しあわせて求めるので大きくなる傾向にある．

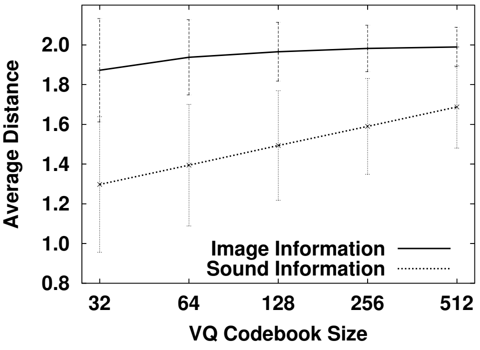


図 1: 映像 · 音響情報の出現確率間の l_1 距離の平均値と標準偏差の比較

3 評価実験

3.1 DSCC 法 (M1+M3)+AS での評価

映像, 音響信号における DSCC 法における性質の違い, Active 探索 (AS) からの距離計算回数, 探索時間の改善率を求める．

3.1.1 DSCC 法の計算量と処理時間

探索閾値 θ を 0.10 に固定し VQ 符号帳の大きさを変化させ映像特徴, 音響特徴を用いた DSCC 法 (M1+M3)+AS の距離計算回数と探索時間を表 1 に示す．各数値は 30 セグメント探索の平均値である．探索閾値 θ が 0.10 のように十分小さい場合は M1+M3 で大部分が処理される．M1+M3 の計算では区分化の際の小球の数が少ない音響情報のほうが距離計算回数が少ない．

表 1: DSCC 法における映像情報と音響情報の比較
探索の平均値 (距離計算, 探索時間: 秒, 区分化数)

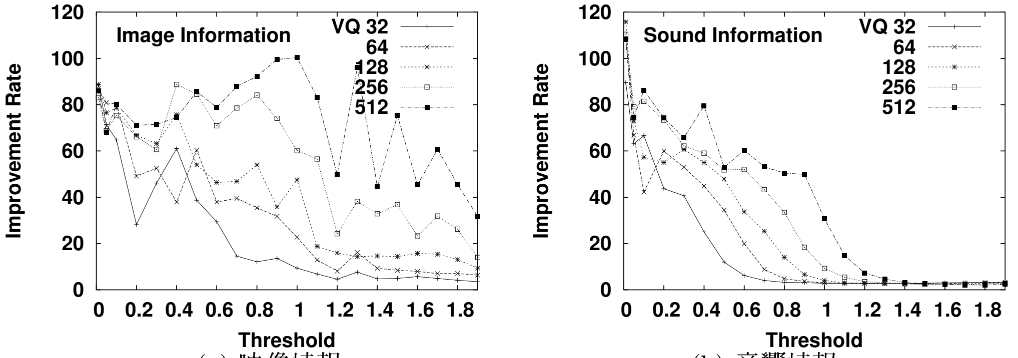
VQ	映像情報			音響情報		
符号帳	計算	時間	小球数	計算	時間	小球数
32	32.7	0.0053	12691	36.9	0.0083	11448
64	45.5	0.0080	12844	42.7	0.0233	11431
128	81.7	0.0173	13169	41.6	0.0320	11252
256	124.4	0.0353	13703	40.5	0.0413	11007
512	178.4	0.0616	14574	44.2	0.0726	10711

平均 l_2 歪み区分化法, 区分化閾値 $\varepsilon = 0.00005$, 探索閾値 $\theta = 0.10$

3.1.2 距離計算, 探索時間の改善率

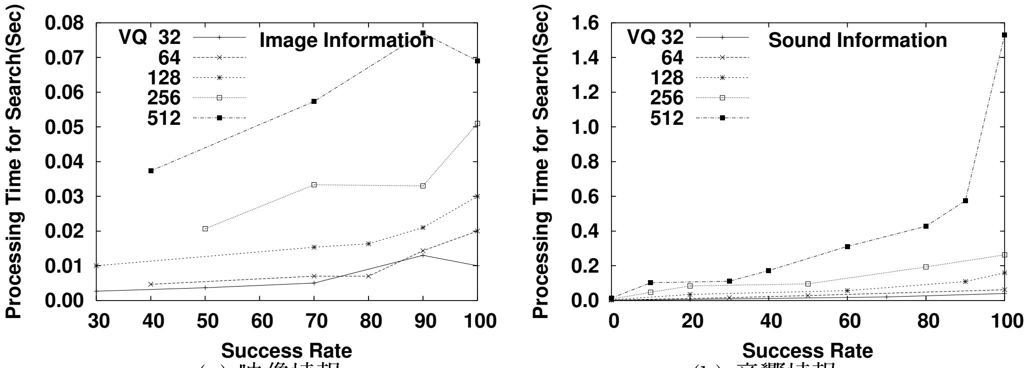
探索時間の改善率 ($\frac{\text{AS}}{\text{DSCC+AS}}$) を図 2 に示す．探索閾値が小さい値では音響情報の方が改善率が高いのは M1+M3 において小球が少いためである．改善率は探索閾値 θ が小さい場合に高く, 探索閾値 θ を大きくするほど低くなった．これは探索閾値が大きくなるにつれて M1+M3 で判定できない箇所が増え処理の重い AS を用いる回数が増えるためである．

*A Quick Video Search Using An Efficeient Segment Searching, by T.Haga, M.Sugiyama (Graduate School, The Univ. of Aizu)



(a) 映像情報　(b) 音響情報

図 2: DSCC 法の探索時間の改善率と探索閾値 θ の関係



(a) 映像情報　(b) 音響情報

図 3: 成功率と探索時間との関係

3.2 雑音付加セグメントに対する探索

3.2.1 雑音に対する耐性

目的信号に雑音を加えての評価を行った．本報告の雑音は MPEG-2 を再生しビデオテープに録画した後, 再度 PC に MPEG-2 として取り込むことで与えた．図 4 は蓄積ベクトルと雑音の付加された 8000 箇所のセグメント間の出現確率ベクトル間の l_1 距離の平均値である．音響より映像の方が距離が小さく雑音の影響が小さく雑音の影響が小さいことがわかる．

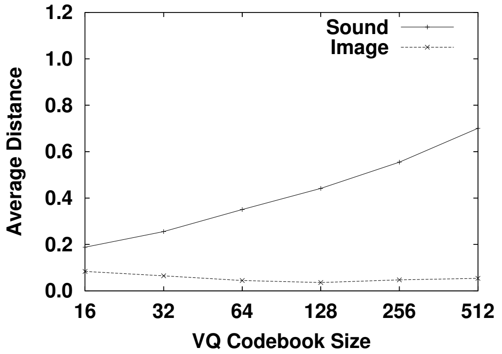


図 4: 雑音付加セグメントと元セグメント間の平均 l_1 距離の比較

3.2.2 探索成功率と処理時間の関係

探索の成功率と探索時間との関係を図 3 に示す．ここで成功率は目的セグメントを検出できた場合を成功とし, わき出しについては考慮していない．探索閾値 θ を $0.01, 0.05, 0.1, 0.2, 0.3, \cdots, 1.9$ と変化させた．図 4 から音響情報に対しては探索成功率を上げるためには VQ 符号帳を小さくするか探索閾値 θ を大きくするの 2 つの方法がある．図 3 の映像情報と音響情報を比較すると映像の方が高い成功率で短時間で探索を行えることが判る．これは映像情報の方が雑音を付加した場合の l_1 距離が小さいので探索を行う場合に探索閾値 θ を音響情報ほど大きくする必要が無いので短い時間で探索が可能であるためである．

4 むすび

探索は映像の方が l_1 距離の平均値が大きく容易である．DSCC 法では小球数が少ない音響の方が改善率が高い．雑音付加セグメント探索では音響より映像での探索の方が低い探索閾値 θ で探索可能, 処理時間も速いので有効である．今後の課題として大規模な蓄積データへの適用, 音響情報の特徴量の検討, 雑音下で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,” 出現確率の幾何学的性質を用いたセグメント高速探索法”, 信学技報, PRMU2005-22, pp.1-6 (2005-06).

[3] 杉山,”Active 探索におけるノルムと類似度との関係”, 信学技報, SP2004-173, pp.53-58 (2005-03).

[4] K.Kashino,et al., ”A quickvideo search method based on local and global feature clustering”, Proc. of ICPR, vol.3, pp.894-897 (Aug. 2004).