

キーワード画像抽出による動画像サマリの作成

1F-10

中島康之 堀 裕修
KDD研究所塩原 敏充*
東京理科大学 工学部 電気工学科

1 はじめに

デジタル動画像圧縮技術や大容量蓄積メディアの普及により、動画像処理技術への関心が高まっている。このうち動画像の検索や認識を目的とした画像処理については、シーンの記述やカット点検出などの検討が進められている[1-6]。

これらの処理においては動画像情報の構造化が重要になる。構造化においては画像の分類が必要になり、要素技術としてシーンの切り替わり点であるカット点の検出や対象画像の検出があげられる。このうち、対象画像の検出はキーワードとなる画像と同一クラスタに属する画像を画像ファイル中から検索するもので、同一クラスタ画像を動画像情報から抽出することにより、キーワード画像を中心とした動画像を再構成することができる。例えばTVニュースは、基本的にアナウンサー画像と資料映像等から構成されており、比較的規則的なパターンを持っているといえる。さらにアナウンサー画像では通常テロップと共に概要説明が行われており[5]、アナウンサー画像とテロップ画像をキーワードとして再構成することにより、付随する音声情報と文字情報（テロップ）からニュース全体の概要把握が可能になる。

本報告では、カット点画像と関連するショットを階層的に用いてテロップが現われるアナウンサー画像を検出することにより動画像サマリを作成したので報告する。

2 シーンの種類について

動画像データについては、例えばMPEG1技術により1/140に圧縮することが可能となったが、そのデータ量は10分間でも約100MBを必要とするため、依然として膨大な情報量となる。このため、動画像処理に必要な情報量の軽減が必要になるが、シーンの切り替わり点を示すカット点は画像シーケンスを大まかに分類する上で有効なツールであり、カット点単位の処理により処理効率を向上することができる。カット点の検出については、いくつかの方法が提案されているが[2,3]、ここではMPEG1圧縮データからのカット検出が可能[3]を用いた。

3 キーワード画像の登録と抽出

特定の画像の抽出についてはこれまでいくつかの方法が提案されている。例えば、色の組み合わせ（RGB）から隣接する領域に含まれる色のペアから対象物の存在を検出する方式[4]や音声の特徴から女性の存在を検出する方式などがある[1]。アナウンサーのように固定的な画像

を対象物として検出する場合は、指定された画像（キーワード画像）と画面全体のレベルで大体一致する画像を検出することにより同一クラスタに属する画像の検出が可能になる。

検出する方式についてはカット点検出方式で用いられている方式をいくつか用いて検討を行った。カット点検出方式は画像が大きく変化する点を検出する方式であるが、逆に用いればほとんど変化しない画像も検出することが可能であると考えられる。そこで、指定された画像と各画像の相関を測定するために、画素変化面積法[2]と色差ヒストグラム相関法[3]で実験を行った。画像はMPEG1で1.2Mbit/sに圧縮された画像からまずカット点検出を行い、カット点画像でキーワードとなる画像を指定し、その画像と各カット点画像を上記の方法を用いて比較し、キーワード画像の検出を行った。画像としては約4分のニュース画像を3種類(News1,2,3)用い、いずれもカット画像の中で最初に現われるアナウンサー画像をキーワードとして登録して、検出実験を行った。図1は画素変化面積法を用いた場合の検出結果で、●は未検出、□は過剰検出を示す。図からもわかるように、どのようなThreshold（画素変化面積S）に対しても未検出あるいは過剰検出が発生している。これに対して、色差ヒストグラム相関を用いた場合、図2のように未検出も過剰検出も発生しないエリア($\rho=0.955-0.98$)が得られている。さらに図3は色差ヒストグラム相関法を改良した方式で、画面を上下に2分し、上下それぞれで得られた相関係数がいずれもある値以上の場合に検出したもので、この場合、アナウンサー画像のように水平方向に比べて上下方向に移動の少ない画像での検出に有効である。図からもわかるように未検出と過剰検出のカーブはなだらかになっており、図2に比べて、Thresholdに対するエラー感度は低く、高い検出精度を保つことが可能となる。

これらの結果から、上下画面の色差ヒストグラム相関値 ρ がいずれも $\rho>TH1$ のときキーワード画像と同一クラスタ画像と判定した。

4 テロップの出現の検出

テロップでは、文字部とその周辺で大きな輝度差が生じている。このため、輝度ヒストグラムをフレーム間で比較して、高輝度部と低輝度部で大きな変化があったフレームをテロップ候補画面として検出する方式が提案されている[6]。しかしながら、実際には縁どりがあまり低輝度でないテロップや、高輝度の物体が現われる画像において誤検出が発生する恐れがある。

文字は多くのエッジ成分から構成されていることから、高周波成分を多く含んでいる。一番簡単な周波数解析として2次のアダマール変換を考えると、隣接画素間の輝度差分値によりエッジ成分の存在を把握することができる。一般に文字部での隣接画素の輝度差分値は非常に大

Construction of audio-video summary through the extraction of key scenes

Y.Nakajima, H.Hori, KDD R&D Labs.

*T.Shiobara, Dep.of Elec.Eng., Science Univ. of Tokyo

きいが文字と背景の関係から、中間的な縁どりもある。そこで、テロップが現われる画面下1/4の領域において、4ライン毎に12ライン分の画素について隣接画素差分値があるThresholdを越える画素数をフレームごとに求めた。図4はテロップを含むアナウンサー画面での画素数の変化を示したものである。●、□、*はそれぞれ隣接画素差分値が32、64、96以上の画素数を示す。またフレームは0.5秒ごとにサンプルされたフレームである。テロップは142フレームから現われているが、差分値が32以上では画素数が大きく変化しているのに対して、64以上では変化量が小さくなり、96以上ではほとんど判別が困難である。実験を行った画像News1、News2いずれも差分値が32以上の場合、テロップ出現時にステップ型の変化が現われることが確認できた。また、アナウンサー画像のテロップは、通常カット点から数秒のちに現われ5秒以上は存在している。従って、カット点から3サンプルフレーム(1.5秒)のNa(隣接画素差分値 ≥ 32 の画素数)の平均値に対して、6サンプルフレーム(3秒)以上連続してTH2以上上昇している場合にテロップの出現と判定した。

5 実験

MPEG1で圧縮されたニュース画像ファイルで、キーワード画像の検出とテロップの出現を検出することにより、動画像サマリの作成実験を行った。まず、カット点検出によりカット点画像を求めた。次にカット点画像で最初に現われるアナウンサーをキーワード画像に指定し、上下分割型の色差ヒストグラム相関により各カット点画像と比較して、同一クラスタ画像の検出を行った。テロップの検出については、同一クラスタ画像のカット点から次のカットまでの画像(ショット)で0.5秒毎にサンプルされたフレームのみを検出の対象とした。テロップが存在するショットをもとにMPEG1圧縮データの編集を行い、動画像サマリを作成した。TH1,TH2については、News1,2,3からTH1=0.93とし、News1,2からTH2=100とした。

表1に約17分のニュース画像News4について、キーワード画像およびテロップ画像検出を行った結果を示す。表からもわかるように実験に用いた画像では色差ヒストグラムと隣接画素差分によりテロップが現われるアナウンサー画像が正しく検出されている。検出時間は、100MIPSのWSで、テロップまでの検出が約7分で完了した。

表1 News4での実験結果

| | 正検出 | 過剰検出 | 未検出 |
|---------|-----|------|-----|
| カット点画像 | 105 | 24 | 8 |
| キーワード画像 | 10 | 0 | 0 |
| テロップ画像 | 6 | 0 | 0 |

6 まとめ

MPEG1で圧縮された動画像ファイルからカット点を検出し、カット点画像群から色差ヒストグラム相関を用いて同一のクラスタに属する画像を抽出し、隣接画素差分により文字の出現を判定し、該当のショットのみを編集することにより同一クラスタシーンを再構成し、ニュース画像において動画像サマリの自動作成を行った。

日頃ご指導頂くKDD研究所、浦野所長、村上次長、お

よび古賀リーダーに感謝致します。

参考文献

- [1] 坂内: "マルチメディアデータベースの展望", 信学技報, IE94-10, pp.1-8, (1994-5).
- [2] 大辻 他: "突出検出フィルタを用いた映像カット点検出法", 信学論,D-II, pp.519-528, (1994-3)
- [3] 中島: "フレーム間輝度差分と色差相関による圧縮動画像データからのカット検出", 1994信学春全大,D-501(1994)
- [4] 上田 他: "動画像解析に基づくビデオ構造の視覚化とその応用", 信学論, D-II, pp.1572-1580, (1994-8)
- [5] 大久保、中川: "AV情報構造化技術とその情報要約への応用", 情処研報, IM15-4, pp.25-32, (1994-3)
- [6] 根本 他: "テロップの認識による資料映像の検索について", 1994信学春全大, D-427,(1994)

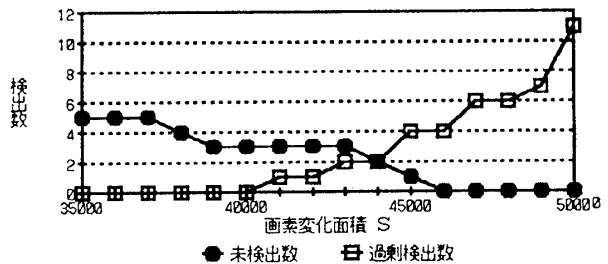


図1 画素変化面積法による同一クラスタシーンの検出

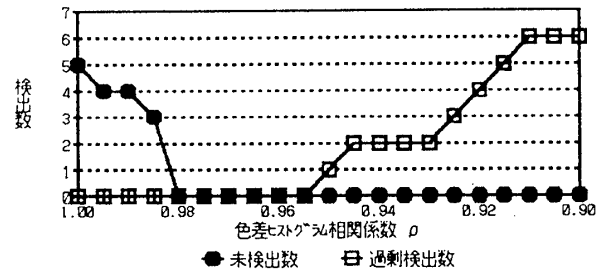


図2 色差相関法による同一クラスタシーンの検出

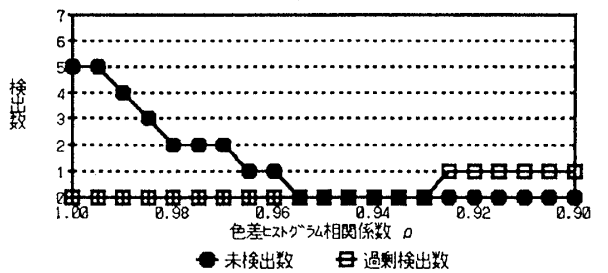


図3 上下画面色差相関法による検出

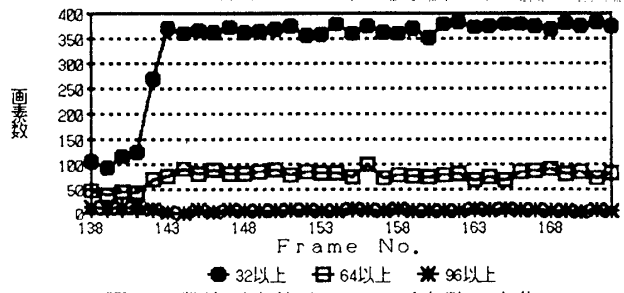


図4 隣接画素差分による画素数の変化