

圧縮動画データからのディゾルブ検出

5N-3

氏原清乃 中島康之 米山暁夫
KDD研究所

1 はじめに

動画画像インデクシングは、電子図書館や映像データベースにおける動画画像の効率的な検索や編集での重要なステップとなっている。カット点は映像の基本的な単位であるシーンの境界点を示すものであり、インデクシングの基本ツールとして、近年活発にその検出方式に関して研究が行われている。特に最近では瞬時にシーンが変わるカット点のみならず、編集効果を用いてシーンが変化するディゾルブなどの特殊カットについての検出方式が報告されている[3-6]。

本稿では、以前報告したDCT DC成分の簡易復号化処理[1]とそのDC画面内アクティビティの変化検出処理により、特殊効果を含むカット点を高精度で検出方法について述べる。

2 従来のディゾルブ検出法

ディゾルブは基本的には2つの異なったシーンが空間的に合成されながら1つのシーンから他のシーンに移行するプロセスであり、例えば輝度の変化をディゾルブに対してモデル化して、ディゾルブを検出する方法が提案されている[3,4]。しかしながら、大きな動きがあるシーンもディゾルブシーンと誤って検出される恐れがあるため、輝度変化の単調性に注目してディゾルブ検出の精度向上を試みたものも報告されている[5]。

また、ディゾルブにおけるシーン合成プロセスに注目してディゾルブ検出を行う方法も提案されている[6]。この場合、ディゾルブ前後に比べ、シーン合成部分で画面内アクティビティが減少する特性を利用している。従って、例えば被写体に微妙な動きがある場合は、輝度が単調に増減して輝度変化による検出では誤検出する可能性が出てくるが、アクティビティを用いた場合はディゾルブと動きではアクティビティの変化特性が異なるため、動きに比較的強い検出が期待できる。そこで本稿では画面内アクティビティに着目し

てディゾルブ検出方式を検討した。

3 ディゾルブ検出法

ディゾルブ検出は、以前報告した圧縮符号化データからのカット点検出[1,2]をベースに方式の検討を行った。このカット点検出ではMPEG符号化データからI、B、P各フレームのDCT DC成分を抽出し[2]、DCT DC画像データからカット点検出を行っている。このため、本ディゾルブ検出についても、各フレームのDC画像データを用いて検出を行った。

まず、ディゾルブ区間においては、各画面内の輝度成分のアクティビティ（画面内偏差）は、下に凸のUカーブ状になる。しかしながら実際の画像においては、動きや局所的な変動によりアクティビティの変化は雑音を伴うことが多く、雑音部分をディゾルブ部分と誤って検出する可能性がある。そこで、このような雑音を除去する目的で、フィルタリング処理を行った。通常ディゾルブでは数十フレーム以上の区間に渡りシーンが変化するため、タップ長の長い時間フィルタリングが局所的な変動を吸収するために適しているといえる。そこで、簡単な時間フィルタリング処理として、現フレームを含むViフレーム間の移動平均処理を用いて雑音除去を行った。

さらに、Uカーブの形状はディゾルブ区間長によっても変化してしまうため、[6]で用いているような単純なピーク検出では比較的平坦なUカーブを検出できない可能性がある。このため、移動平均値の差分値DMVを用いて変化点の検出を行った。図2にディゾルブを含むシーケンスにおける移動平均差分の例を示す。ディゾルブ区間は、図の左側のように下に凸のピークから上に凸のピークに変化する部分として検出することができる。ただし、フェード区間、動きながらのディゾルブ、あるいは画面の形状によって正または負どちらかの単一のピークになることがある。図の右側は比較的平坦なシーンから変化のあるシーンにディゾルブで移行する場合の例である。従って、検出はDMVの絶対値があるしきい値 β を越えた場合ディゾルブ候補とした。

また、ディゾルブ区間では、連続するフレーム間の相関は小さくても、ある程度フレーム間隔を

Detection of dissolve scene change from compressed video

Kiyono UJIHARA, Yasuyuki NAKAJIMA
Akio YONEYAMA
KDD Co.Ltd., R&D Labs.

とったフレーム間相関は大きくなる傾向がある。そこで、Df フレーム離れた2つのフレームのDC画像の色差相関 ρ [2]を求め、先のディゾルブ候補について、 ρ がしきい値 γ 以下の場合ディゾルブと判定した。

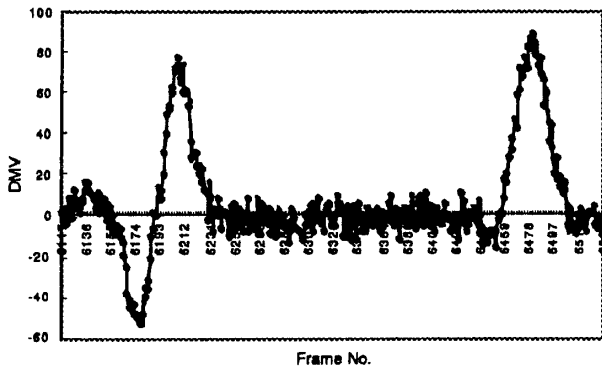


図2 移動平均差分 DMV

4 比較実験

以前報告したカット検出方式[1]と本方式によるディゾルブ検出を加えた場合の統合カット検出の比較実験を行った。なお、統合カット検出ではまずディゾルブ検出を行い、ディゾルブ検出でディゾルブ以外と判定されたフレームについて以前のカット検出を適用して検出した。

画像は、MPEG1で符号化されたニュース、料理番組、バラエティー、CMを含んだおよそ1時間のTV番組で、ワークステーション(約100MIPS)上のHDに蓄積され、そのビットストリームから検出を行なった。使用したパラメータは従来法に関しては、両方式ともに[1]で用いたものを採用し、ディゾルブ検出に関しては、 $Vf=30, Df=45, \beta=20, \gamma=0.6$ とした。

表1に検出結果を示す。ここで、方式1は従来方式で検出を行なったもの、方式2は今回提案した方式を加えて検出を行なったものである。

尚、表中の検出率は以下の計算式を用いている。

$$\text{検出率}[\%] = \frac{[\text{検出カット数(カット)}] - [\text{過剰検出(過)}]}{[\text{検出カット数} - [\text{過剰検出}] + [\text{未検出(未)}]} \times 100$$

5 考察

表1からもわかるように、方式2では一部の番組を除いて、どの画像でも検出率が向上している。これは、方式1で検出できなかった特殊効果を用いたカット検出が方式2で検出可能となっているため、ディゾルブだけでなくワイプに対してもある程度検出可能であることがわかった。ただし、

ディゾルブ検出過程での過剰検出が全過剰検出数の増加をもたらしており、今後その他の未検出カットと合わせて、過剰検出への対策も検討する必要があるといえる。

また、方式1と方式2の処理時間はほぼ同等で、ディゾルブ検出による処理量の増加は非常に小さいと言える。

表1 検出結果

	方式1				方式2			
	カット	過	未	検出率 [%]	カット	過	未	検出率 [%]
News	58	1	2	96.6	63	4	0	100
料理	28	0	1	96.6	31	2	0	100
バラエティ	283	3	5	98.2	285	7	7	97.5
CM	213	7	36	85.1	222	7	28	88.4
Total	582	11	44	92.8	601	20	34	94.5

6 まとめ

本稿では、圧縮動画データからのカット点検出の高精度化について述べた。DCT DC成分の簡易復号化とDC画面内アクティビティの変化を用いることにより、通常のカット点の検出に加えて、ディゾルブなどの特殊効果を用いたカット点の検出を可能にし、全体の検出率の向上を図ることが可能であることを確認した。

日頃御指導頂くKDD研究所、村上所長、および羽鳥SGL、古賀SGLに感謝致します。

参考文献

- [1] 氏原, 中島, 堀, 加納: "簡易復号処理による圧縮動画データからのカット点検出", 情処全大第51回, 6S-9[1995-9]
- [2] 中島: "フレーム間輝度差分と色差相関による圧縮動画データからのカット検出", 信学春季全大, D-501[1994-3]
- [3] 村田, 中村, 大田: "映像におけるカット変わりの自動検出", 情処全大第50回, 6D-7[1995-3]
- [4] 井岡: "動画におけるディゾルブ画面切り替えの検出", 情処全大第51回, 6S-8[1995-9]
- [5] 長坂, 宮武, 谷口: "輝度変化の単調性に着目したディゾルブ検出", 信学春季全大, D-615[1996-3]
- [6] J. Meng, Y. Juan, and S.F. Chang, "Scene change detection in a MPEG compressed video sequence, SPIE Vol.2419, pp.14-25, 1995