

6R-08 ビデオデータ配信時におけるシーン分割法の研究 II

小林健一[†], 飯倉道雄[†], 伊原征治郎[†], 川田直樹^{††}

[†] 日本工業大学工学部

^{††} トランスコスモス株式会社

1 はじめに

本章では、ライブ中継にデジタル・ビデオを使用する場合の、リアルタイム・シーン切り替わり検出法について述べる。現在提案されているシーンの切り替わり検出法には、動きベクトルを利用する方法 [1] や、画像の色強度を用いた発見的的手法による方法 [2] があげられる。これらの方法でシーンの切り替わりを検出する場合は、一般に特徴量抽出のためにデジタル・ビデオを一旦ファイルに格納しておく必要がある。本稿では、リアルタイムにシーンの切り替わりを検出するために、隣り合う 2 枚のフレームの明度の分散分析を行ってシーン切り替わりを検出する方式を提案する。

2 分散分析によるシーン切り替わり検出

1 つのフレームは、1 枚の静止画像である。明度を統計的に扱うために、フレームの画素を横 1 ライン (Line) ごとに使用し、1 ライン分の明度をもとに分散分析を行う。i 番目のフレームの j 番目のラインの k 番目の画素 (Pixel) を p_{ijk} とし、明度は c_{ijk} とする。i のとる範囲は 1 から l 、j のとる範囲は 1 から m 、k のとる範囲は 1 から n とする。 c_{ijk} を利用して明度のダイナミックレンジの平均値を求め、サンプルとする。

1 ラインから q 個おきに画素 p_{ijk} を取り出して明度 c_{ijk} を算出することを $r (= \frac{n}{q})$ 回繰り返す。集合 C_{iju} とする。u の範囲は $u = 1, \dots, q$ であり、1 ラインあたり q 個の集合 C_{iju} を得る。式 (2) により、ダイナミックレンジ d_{iju} を求め、得られた q 個のダイナミックレンジから式 (3) を用いて平均値を求める。

$$C_{iju} = \{c(i, j, u), c(i, j, u + q), \dots, c(i, j, u + (r - 1)q)\} \quad (1)$$

$$u = 1, \dots, q$$

$$d_{iju} = \text{Max}[C_{iju}] - \text{Min}[C_{iju}] \quad (2)$$

$$S_{ij} = \frac{\sum_{u=1}^q d_{iju}}{q} \quad (3)$$

ここで $\text{Max}[C_{iju}]$ は明度の集合 C_{iju} の最大値、 $\text{Min}[C_{iju}]$ は C_{iju} の最小値である。

抽出したサンプル S_{ij} をもとに 1 元配置の分散分析 [4] を行い、統計量 F を算出する。この統計量 F は、自由度 $1, 2m - 2$ の F 分布に従うことが知られている。そこで、隣り合う 2 フレームの明度の平均値について、有意水準 α に対し、確率 P が

$$P(F > F_0) = \alpha \quad (4)$$

となるように F_0 を定め、シーン切り替わりを判定する領域として

$$R = (F_0, \infty) \quad (5)$$

を求め、 F が領域 R に含まれれば、シーンの切り替わりが発生したと判定する。

3 実験と考察

実験に使用したデジタル・ビデオは、毎秒 15 フレームで構成されており、総再生時間は 30 秒である。1 フレームは、 160×120 画素の静止画像である。有意水準 $\alpha = 0.001$ とし、分散分析の結果 F 値が領域 R 内にあればシーンが切り替わったと判定する。このとき $F_0 = F_{238}^1(0.001) = 11.1$ である。

Video Scene Detection for Real-time Video Manipulation II

Kenichi Kobayashi[†], Michio Iikura[†], Seiji Iihara[†], Naoki Kawata^{††}

[†] Nippon Institute of Technology

4-1 Gakuendai, Miyashiro, Minami-Saitama, Saitama 345-0826 Japan

^{††} Trans Cosmos

3-3-3 Akasaka, Minatoku, Tokyo 107-0052 Japan

表 1: シーン切り替わり検出結果

検出法	総検出数 (回)	正当検出 (回)	過剰検出 (回)	未検出 (回)	シーン検出率 (%)	過剰検出率 (%)
発見的手法	20	20	0	0	100.0	0.0
分散分析	28	7	21	13	35.0	105.0
χ^2 検定	128	11	117	9	55.0	585.0

また、シーン切り替わり検出法として、 χ^2 検定をもちいた手法 [3] が提案されている。そこで χ^2 検定によるシーン切り替わり検出の実験を行った。自由度は明度の階調数より 255、有意水準 $\alpha = 0.001$ とし、 χ^2 検定により $\chi^2 > \chi_0^2(0.001) = 330.5177$ のとき、シーンの切り替わりが発生したと判定した。分散分析による検出法と χ^2 検定による検出法との比較

デジタル・ビデオのシーンの切り替わりを発見的手法によって判定した結果を基準とし、分散分析による検出結果と χ^2 検定による検出結果を比較した結果を表 1 に示す。発見的手法ではシーンを人間が判断し、フラッシュやストロボにより画面が点滅した場合や、カメラの前を人物が横切った場合は連続したシーン内のフレームであると判定した。 χ^2 検定による検出法は過剰検出率が非常に高い。これは、画像の一部分の明度が変化した場合でも実現値 χ^2 値が χ^2 分布の 0.1% 点 χ_0^2 を超えてしまうことが原因である。 χ^2 検定による検出法は、発見的手法によるシーン検出と比べて非常にシーン検出数が多く、今回のデジタルビデオでは平均して 4 フレームに 1 回シーン切り替わりを検出している。このことから、シーン切り替わり検出法としては、分散分析による検出法が優れているといえる。

4 シーン切り替わり検出処理時間

デジタル・ビデオのような連続メディア・データは、時系列で連続的に変化するデータである。このようなデータの正当性は、その値の正当性と時間的な正当性も要求される。そこで、処理速度の違う CPU において 1 回の分散分析に要する平均演算時間を求めた (表 2)。

今回計測に用いた最も遅いプログラム実行環境 (Pentium Pro・200MHz) においても、1 回の分散分析に要する平均演算時間は 6msec 程度であり、この処理時間はフレーム読み取り時間間隔 (約 66msec) に対して十分短いことが分かった。また、その標準偏差 (0.167) より、約 6.13msec ($5.63 + 0.167 \times 3$) 以下で 99% 以上のデータが処理されることがわかり、処理時間の安定性が確認された。

表 2: 1 回のシーン切り替わり検出処理時間の平均値および標準偏差

CPU	平均演算時間 (msec)	標準偏差
Pentium III 700MHz	1.59	0.089
Pentium II 450MHz	2.31	0.108
Pentium Pro 200MHz	5.63	0.167

5 おわりに

デジタル・ビデオにおけるリアルタイムでのシーン切り替わり検出のため、分散分析を用いる方法を提案し、評価実験により有効性を確認した。

本検出法を用いて実験を行った結果、検出処理時間は Pentium II 450MHz の CPU を使用した場合、 160×120 画素の画像に対して 2.31[msec] であった。秒間 30 フレームのデジタル・ビデオでは、利用可能時間の 7% で検出が終了する。

今後の課題として、本検出法でデジタル・ビデオをシーンに分割した場合のデータ圧縮率に与える影響についての検討があげられる。

参考文献

- [1] 浅水, 長谷川, 北島: 動画像符号化のためのテーブル参照による動き補償, 電気通信学会論文誌 Vol. j82-D-II No.6(1999).
- [2] 是津, 上原, 田中: 時刻印付オーサリンググラフによるビデオ映像のシーン検索, 情報処理学会論文誌 Vol.39 Number 4(1998).
- [3] 大倉: 画像データの検索に関する研究,
<http://www.tech-lab.pref.saitama.jp/kenkyu/19970708.htm>
- [4] 竹内: 数理統計学, 東洋経済新報社 (1963).