

分散分析によるデジタル・ビデオのリアルタイム・シーン切り替わり検出法

小林 健一[†] 飯倉 道雄[†]
川田 直樹^{††} 伊原 征治郎[†]

デジタル・ビデオを利用してライブ中継などを行う場合、データを蓄積せずに、高速にシーンの切り替わりを検出するリアルタイムでのシーン検出技術が必要である。本稿では、リアルタイムでデジタル・ビデオのシーン切り替わりを検出する方法として、明度のサンプルを分散分析する検出法を提案する。本検出法では、隣り合う2フレーム間で画素の明度をサンプル抽出し分散分析を行い、その結果を利用してシーンの切り替わりを検出する。サンプル抽出法として、明度の平均値を用いる方法と、明度のダイナミックレンジの平均値を用いる方法について評価実験を行った。実験の結果、両方法ともリアルタイムでのシーンの切り替わりの検出に有効であることを確認した。

Real-time Detection of Video Scene Breaks Using Analysis of Variance

KEN-ICHI KOBAYASHI,[†] MICHIO IKURA,[†] NAOKI KAWATA^{††}
and SEIJIRO IHARA[†]

Real-time detection of video scene breaks without storing video data is very important for a live broadcast by digital video. In this paper, we present a method of real-time detection of scene breaks using analysis of variance with samples of luminosity between frames. The method detects scene breaks using the result of analysis of variance with samples based on luminosity of pixels in continuous 2 frames. We have experimented the method with samples of the average luminosity and dynamic range of luminosity, respectively. Both results showed the method can detect scene breaks efficiently.

1. はじめに

近年、マルチメディア情報通信技術の向上により、デジタル・ビデオをライブ中継に利用することが可能となってきた。リアルタイムで生成されるデジタル・ビデオに対してシーン情報を付加することで、対話的な画像編集が可能となる。たとえば防犯カメラなどで、異常検知にシーン情報を利用することが考えられる。

シーン情報を付加するためには、シーンの切り替わりを検出する必要がある。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),2)} や、画像の色強度を用いた発見的手法による方法³⁾ 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、一般にデジタル・ビデオをいったんファイルに格納しておく必要がある。これは、デジタル・ビデオ全体から特徴量を抽出してからシーン

切り替わりの検出を行うためである。このような方法では、リアルタイムにシーン検出を行うことが困難である。

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。本検出法の実験結果は、リアルタイムでのシーン切り替わりの検出に有効であることが明らかになったので報告する。

2. シーン切り替わり検出法

デジタル・ビデオは複数のシーンから構成され、1つのシーンは何枚かの類似したフレームから成る。デジタル・ビデオ・データは、一般にデータ圧縮処理を行った状態でビデオ・クライアントに転送される。連続するフレーム間での差分圧縮が行われるため、フレームの類似の度合いによって、圧縮されたデータの量が変化する。フレーム間でシーンが切り替わると類似度が低くなり、圧縮率がいちじるしく低下するので、フレーム間の類似度判定を行ってシーンの切り替わりを

[†] 日本工業大学工学部

Nippon Institute of Technology

^{††} トランス・コスモス株式会社

Trans Cosmos Inc.

検出する必要がある。

現在、シーン切り替わり検出法として、動きベクトルを利用する方法と、発見的手法による方法が利用されている。

動きベクトルを利用した検出法

連続するフレーム間に表示されている物体に注目し、その物体の動きを検出することにより、連続するフレームが、同じシーン内のものであるか、別のシーンのものであるかを判定する手法である。

直前のフレームに表示されていた物体が現在のフレーム内に発見できなかった場合、シーンの切り替わりと判定する。MPEG⁴⁾では動きベクトルを利用した圧縮技術が採用されており、MPEG形式のデジタル・ビデオでは、動きベクトルを利用してシーン切り替わりを検出した結果をそのまま圧縮処理に利用できる可能性がある。しかしこの手法では、物体を認識しその移動を予測しなければならないため、他の検出法と比較して処理負荷が増大する。また、動きベクトル抽出処理は発見的に行われるため、抽出までにかかる時間が安定しないことも問題である。

発見的手法による検出法

フレーム個々の画像から特徴量を抽出し、デジタル・ビデオ全体での特徴量の変化をもとに、シーンの切り替わりを決定する手法である。特徴量としては、画像の色強度の平均値などが用いられる。

この手法では、最終的に人間の主観的判断によりシーンの切り替わりを決定するのが一般的である。デジタル・ビデオ・データ全体にわたって特徴量を抽出し、その結果を比較する必要があるため、リアルタイムでのシーン切り替わりの検出は困難である。

3. 分散分析によるシーン切り替わり検出

上に述べたように、リアルタイムでシーンの切り替わり検出を行う場合には、処理にかかる時間が短く、フレームごとの処理時間が一定であることが重要である。また、隣り合う2フレームの特徴量のみで検出を行わなければならない。本稿では、隣り合うフレームの明度をサンプルとした分散分析を利用するシーン検出法を提案する。

1つのフレームは、図1に示すように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} とする。 i のとりうる範囲は1から l 、 j のとりうる範囲は1から m 、 k のとりうる範囲は1から n とする。画素は通常赤

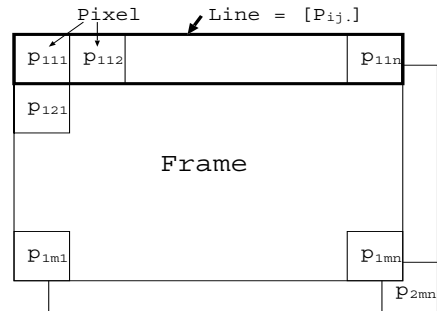


図1 画像データとライン

Fig. 1 Frame data and line data.

(r) 緑 (g) 青 (b) の色強度を表しているので、以下の式(1)により、画素 p_{ijk} の明度 c_{ijk} を算出する⁶⁾。

$$c_{ijk} = 0.3 \times r_{ijk} + 0.59 \times g_{ijk} + 0.11 \times b_{ijk} \quad (1)$$

r_{ijk} , g_{ijk} , b_{ijk} は、画素 p_{ijk} のrgb色強度である。 c_{ijk} を利用してサンプルを抽出し、隣り合う2フレーム間で分散分析を行うことにより、フレーム間の類似度を判定しシーンの切り替わりを検出する。サンプル抽出法として、1ライン分の明度の平均値をサンプルとする方式と、明度のダイナミックレンジの平均値をサンプルとする方式を試みる。

明度の平均値を用いる方法

1ラインごとに明度の平均値を求め、これをサンプルとして分散分析を行う。1つのサンプル S_{ij} は、以下の式(2)によって算出される。

$$S_{ij} = \frac{\sum_{k=1}^n c_{ijk}}{n} \quad (2)$$

ここで S_{ij} はサンプルの値、 n は1ラインの画素数、 c_{ijk} は画素 p_{ijk} の明度である。

明度のダイナミックレンジの平均値を用いる方法

1ラインごとに明度のダイナミックレンジの平均値を求め、これをサンプルとして分散分析を行う。図2に示すように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(5)により、各 C_{iju} の要素内で最も大きい値から最も小さい値を引いた差をダイナミックレンジ d_{iju} とし、得られた q 個のダイナミックレンジから式(6)を用いて平均値を求める。

$$C_{iju} = \{c(i, j, u), c(i, j, u + q), \dots, c(i, j, u + (r - 1)q)\} \quad (3)$$

$$u = 1, \dots, q \quad (4)$$

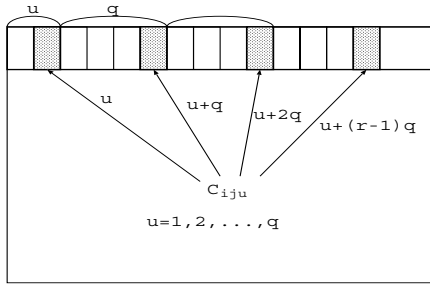


図 2 ダイナミックレンジのサンプリング

Fig. 2 Sampling method of dynamic range.

$$d_{iju} = \text{Max}[C_{iju}] - \text{Min}[C_{iju}] \quad (5)$$

$$S_{ij} = \frac{\sum_{u=1}^q d_{iju}}{q} \quad (6)$$

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.

これらの方法で抽出したサンプル S_{ij} をもとに 1 元配置の分散分析を行う. 統計量 F は以下の式により算出する⁷⁾.

$$F = \frac{V_F}{V_E} \quad (7)$$

$$V_F = \sum_{i=1}^2 m(\bar{S}_i - \bar{S}_{..})^2 \quad (8)$$

$$V_E = \frac{\sum_{i=1}^2 \sum_{j=1}^m (S_{ij} - \bar{S}_i)^2}{2m - 2} \quad (9)$$

$$\bar{S}_{..} = \frac{\sum_{i=1}^2 \sum_{j=1}^m S_{ij}}{2m} \quad (10)$$

$$\bar{S}_i = \frac{\sum_{j=1}^m S_{ij}}{m} \quad (11)$$

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について, 有意水準 α に対し, 確率 P が

$$P(F > F_0) = \alpha \quad (12)$$

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

$$R = (F_0, \infty) \quad (13)$$

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.

4. 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. このデジタル・ビデオの中には, フェイドアウトやカメラのストロボ発光を含むシーンが存在する. 実験は, 明度の平均

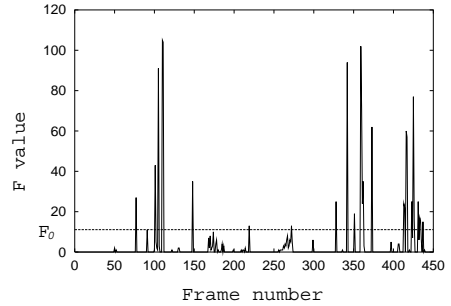


図 3 明度の平均値を用いた場合の実験結果

Fig. 3 Experimental result of average.

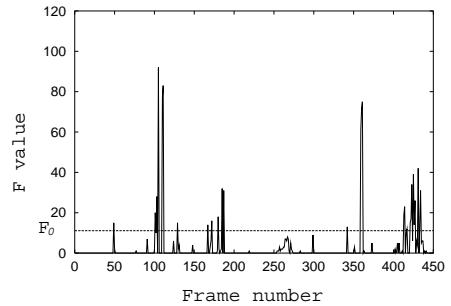


図 4 ダイナミックレンジの平均値を用いた場合の実験結果

Fig. 4 Experimental result of dynamic range.

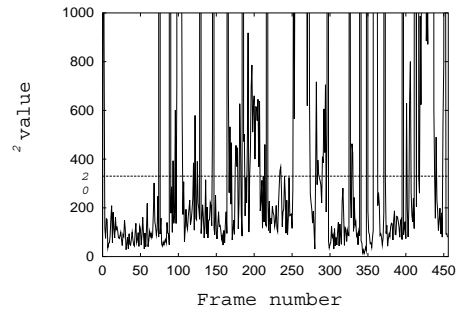


図 5 明度に対して χ^2 検定を用いた場合の実験結果

Fig. 5 Experimental result of χ^2 test.

値をサンプルとした場合と, 明度のダイナミックレンジの平均値をサンプルとした場合について行う. 隣り合う 2 フレームからサンプル $S_{ij} (i = 1, 2; j = 1, 2, \dots, m)$ を求め, 分散分析を行う. 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.

また, シーン切り替わり検出法として, χ^2 検定を用いた手法⁵⁾ が提案されている. これは, 隣り合うフレーム間の明度の分布型を比較することにより, シーンの切り替わりを検出する手法である. そこで同一

表 1 シーン切り替わり検出結果
Table 1 Result of experiments.

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
明度	27	11	16	9	55.0	80.0
ダイナミックレンジ	28	7	21	13	35.0	105.0
χ^2 検定	128	11	117	9	55.0	585.0

のデジタル・ビデオに対し 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 $\chi^2 > \chi_0^2(0.001) = 330.5$ のとき, シーンの切り替わりが発生したと判定した。

4.1 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

明度の平均値をサンプルとした場合の実験結果の例を図 3 に示す。縦軸は分散分析の F 値, 横軸はフレームの番号を表している。統計量 $F > F_0$ となるフレーム間では, シーンの切り替わりが発生したと判定する。類似度の高いフレーム間では, F 値は非常に小さい値をとるのに対し, シーンの切り替わりと判定される箇所では, F 値は極端に高い値をとることが分かる。同様にダイナミックレンジの平均値をサンプルとした場合の測定結果を図 4 に示す。この方式においても, 類似度の高いフレーム間での F 値に比べ, シーンの切り替わりと判定される箇所での F 値が極端に高い。また, 両検出法においてほとんど同じ箇所 F 値が高くなっている。 χ^2 検定を使用したシーンの切り替わりの検出実験の結果を図 5 に示す。 χ^2 検定を使用した場合, シーンの切り替わりと判定される回数が非常に多く, 意味のある単位ごとにシーンの切り替わりを検出できなかった。

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。発見的手法によるシーン切り替わり検出とは, デジタル・ビデオの中で意味的に連続したシーンを人間が判断し, シーンの切り替わりを判定する方法である。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。表 1 において総検出数は, 各手法で検出されたシーン切り替わりの回数である。正当検出数は, 発見的手法によってシーン切り替わりと判定されたフレーム間をシーン切り替わりとして検出した回数である。過剰検出数は, 発見的手法による検出でシーン切り替わりと判定され

なかったフレーム間をシーン切り替わりとして検出した回数である。未検出数は, 発見的手法による検出でシーン切り替わりと判定されたフレーム間をシーン切り替わりとして検出できなかった回数である。

シーン検出率 D_c は次の式 (14) より算出する。

$$D = \frac{d_c}{d_h} \quad (14)$$

ここで d_c は正当検出数, d_h は発見的手法による総検出数である。過剰検出率 D_e は次の式 (15) より算出する。

$$D = \frac{d_e}{d_h} \quad (15)$$

ここで d_e は過剰検出数, d_h は発見的手法による総検出数である。シーン検出率の比較結果より, 明度の平均値および χ^2 検定による検出法が発見的手法によるシーン切り替わり検出結果に近いことが示された。しかし, 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, シーン切り替わりの検出には明度の平均値の分散分析による検出法が優れているといえる。

4.2 分散分析による検出法と統計量 F の変化

分散分析による検出法でシーン切り替わりが検出されたフレームと統計量 F の変化を図 6, 7, 8, 9 に示す。図中, $F_a(f_1, f_2)$ は図 3 での f_1, f_2 番フレーム間の明度の平均値を用いた分散分析の F を示す。 $F_d(f_1, f_2)$ は図 4 での f_1, f_2 番フレーム間の明度のダイナミックレンジの平均値を用いた分散分析の F である。 $F_a(49, 50) = 0.0 < F_0 = 11.1$, $F_d(49, 50) = 15.2 > F_0 = 11.1$ は, シーン切り替わりが発生したフレーム間の統計量 F である。平均値の分散分析の結果では明度が平均化された結果, 同一シーンとして判定されるが, ダイナミックレンジの平均値の分散分析の結果は $F > F_0$ となり, シーンの



49
 $F_a(49, 50) = 0.0 < F_0 = 11.1$
 $F_d(49, 50) = 15.2 > F_0 = 11.1$



50
 $F_a(50, 51) = 2.0 < F_0 = 11.1$
 $F_d(50, 51) = 2.3 < F_0 = 11.1$



51
 図 6 シーン切り替わり
 Fig. 6 Example: Scene breaks.



110
 $F_a(110, 111) = 105.1 > F_0 = 11.1$
 $F_d(110, 111) = 77.5 > F_0 = 11.1$



111
 $F_a(111, 112) = 104.0 > F_0 = 11.1$
 $F_d(111, 112) = 83.8 > F_0 = 11.1$



112
 図 7 カメラのフラッシュ
 Fig. 7 Example: Flash by camera.

切り替わりとして検出できていることが分かる。50, 51 番フレームは同一シーン内の連続するフレームであり、統計量 F は $F_a(50, 51) = 2.0 < F_0 = 11.1$, $F_d(50, 51) = 2.3 < F_0 = 11.1$ で同一シーンであると判定されている。

1 シーン内で明度が大きく変化すると、シーンの切り替わりでないフレーム間で、シーン切り替わりが発生したと判定される。1 シーン内での明度の変化には、主に以下の要因が考えられる。

- フラッシュやストロボの影響
- 被写体そのものの運動・変化
- カメラの移動
- デジタル・ビデオ・データに対する特殊効果

カメラのフラッシュを含むシーンでの統計量 F の変化を図 7 に示す。110, 111 番および 111, 112 番フレーム間では、フラッシュの影響によりフレーム間の明度が大きく変化し $F_a > F_0$, $F_d > F_0$ となり、連続してシーン切り替わりが発生したと判定されている。カメ

ラの移動による明度の変化の例として図 8 にズームの例を示す。秒間フレーム数が多くなると連続するフレーム間での画像の変化が小さくなる。このため $F < F_0$ となり、同一シーンとして判定される可能性が高くなる。秒間フレーム数の少ないデジタル・ビデオでのズームの様子を示すために行った 184 番フレームと 189 番フレームでの分散分析の結果は、明度の平均値の分散分析では $F_a(184, 189) = 47.5 > F_0 = 11.1$ となり、シーンの切り替わりとして判定されている。ダイナミックレンジの平均値を用いた場合は画像の拡大の影響を受けておらず $F_d(184, 189) = 5.7 < F_0 = 11.1$ となり、同一シーンとして判定される。

特殊効果の例として、フェードアウト効果を含むシーンを図 9 に示す。フェードアウトやフェードインは、連続するフレーム間でフレーム全体の明度が平均的に変化する特殊効果である。ダイナミックレンジは明度の最大値と最小値の差をとるため、この効果の影響を受けにくく、フェードアウト、フェードイン



183

$$F_a(183, 184) = 1.3 < F_0 = 11.1$$

$$F_d(183, 184) = 0.0 < F_0 = 11.1$$



184

$$F_a(184, 189) = 47.5 > F_0 = 11.1$$

$$F_d(184, 189) = 5.7 < F_0 = 11.1$$



189

図 8 ズームの例

Fig. 8 Example: zoom.



270

$$F_a(270, 271) = 6.8 < F_0 = 11.1$$

$$F_d(270, 271) = 0.5 < F_0 = 11.1$$



271

$$F_a(271, 276) = 15.1 > F_0 = 11.1$$

$$F_d(271, 276) = 10.1 < F_0 = 11.1$$



276

図 9 フェードアウトの例

Fig. 9 Example: fadeout.

は 1 つのシーンと判定できる可能性が高くなる。図 9 の例においても $F_a(270, 271) = 6.8 < F_0 = 11.1$, $F_d(270, 271) = 0.5 < F_0 = 11.1$ となって同一シーン内の連続するフレームであると判定される。秒間フレーム数の少ないデジタル・ビデオでのフェードアウトの様子を示すために 271 番フレームと 276 番フレームを比較した場合、ダイナミックレンジの平均値の分散分析の結果は $F_d(271, 276) = 10.1 < F_0 = 11.1$ となり同一シーンであると判定された。これに対して、明度の平均値を用いて検出を行う場合は、フェードイン・フェードアウトによる明度の変化の影響を受けやすい。明度の平均値の分散分析の結果は $F_a(271, 276) = 15.1 > F_0 = 11.1$ となりシーンの切り替わりと判定された。

5. 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。デジタル・ビデオにおける隣合う 2 フレームについて、その類似性をリアルタイムに計測するためには、その処理時間は、次の要件を満足する必要がある。

- (1) フレーム読み取り時間間隔に比べて十分短い処理時間
- (2) 同一のフレーム・サイズに対して一定の処理時間
今回実験に用いたビデオ・データ (15 フレーム/秒, 160×120 画素/フレーム, 455 フレーム) について、明度の平均値をサンプルとした分散分析によるシーン切り替わり検出処理時間を計測した。全フレーム (455 枚) を対象にして、すべてのフレーム隣接箇所 (454 箇所) で分散分析を行い、累積演算時間 (454 箇所分) を計測した。この累積演算時間とフレーム隣接箇所数 (454) より、1 回の分散分析に要する平均演算時間を求めた。

分散分析の計算には、フレーム・サイズに対してそ

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
Table 2 Averages and standard deviations of processing time for a scene detection.

プログラム 実行環境	CPU	平均演算時間の 平均値 (msec)	標準偏差
Linux 2.0.36	Pentium III 700 MHz	1.59	0.089
Linux 2.2.14	Pentium II 450 MHz	2.31	0.108
Linux 2.0.36	Pentium Pro 200 MHz	5.63	0.167

の計算量が一定となる方式を用いた。したがって、同一のフレーム・サイズを持つデジタル・ビデオ・データに対しては、本来、同一の処理時間が期待される。しかし、多重処理をしているプログラム実行環境においては、同一の処理時間が保証されない。そこで、前述の計測を複数回(30回)行い、1回の分散分析に要する平均処理時間の平均値および標準偏差を求めた。さらに、演算速度の異なる3種のプログラム実行環境において計測した(表2)。

今回計測に用いた最も遅いプログラム実行環境(PentiumPro・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分かり、処理時間の安定性が確認された。

6. 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。1フレーム内の画素から明度を抽出し、これをもとにサンプルを求め、分散分析を行うことにより、隣り合う2フレーム間のサンプルのみでシーンの切り替わりが検出可能である。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) Tae-Hyun, H. and Dong-Seok, J.: Detection of Video Scene Breaks Using Directional Informa-

- tion in DCT Domain, *Proc. 10th ICIAP* (1998).
- 3) 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, No.4 (1998).
 - 4) MPEGTV: MPEG Pointers and Resources. <http://www.mpeg.org/MPEG/>
 - 5) 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>
 - 6) 太田: 色彩工学, 東京電機大学出版局 (1993).
 - 7) 竹内: 数理統計学, 東洋経済新報社 (1963).

(平成12年6月26日受付)

(平成13年3月9日採録)



小林 健一 (学生会員)

1996年日本工業大学工学部システム工学科卒業。1998年同大学大学院工学研究科(博士前期課程)システム工学専攻修了(工学修士)。現在、同専攻博士後期課程在学中。VODシステムおよび教育支援環境に関する研究に従事。



飯倉 道雄 (正会員)

1969年電気通信大学電気通信学部電波通信学科卒業。1971年同大学大学院電気通信研究科修了。1975年日本工業大学工学部専任講師。1982年同助教授。並列分散処理, 情報教育支援環境, 学習支援環境等に関する研究に従事。教育システム情報学会, 電子情報通信学会, 日本計算機統計学会各会員。



川田 直樹

1998年日本工業大学工学部システム工学科卒業。2000年同大学大学院工学研究科(博士前期課程)システム工学専攻修了(工学修士)。現在トランス・コスモス(株)に勤務。



伊原征治郎

1962年早稲田大学第一理工学部電気工学科卒業。以降(株)富士電機, 通商産業省工業技術院電子技術総合研究所等を経て, 現在, 日本工業大学システム工学科教授。工学博士。電気学会会員。