

## 多変量解析を用いた動画像特徴量抽出方式

真島 恵吾<sup>†</sup> 大竹 剛<sup>†</sup> 合志 清一<sup>†</sup> 小川 一人<sup>†</sup> 室田 逸郎<sup>‡</sup>  
難波 誠一<sup>†</sup> 小山田 公之<sup>†</sup>

<sup>†</sup>NHK 放送技術研究所 〒157-8510 東京都世田谷区砧 1-10-11

<sup>‡</sup>NHK 放送技術局 〒150-8001 東京都渋谷区神南 2-2-1

E-mail: majima.k-fu@nhk.or.jp

あらまし インターネット等により流通するコンテンツの識別を目的とし、動画像特徴量抽出方式を提案する。本方式は、信号処理と多変量解析を用いて動画像コンテンツの特徴量を取得することにより、コンテンツ全体を少ない特徴量データで表現するとともに、一部のデータが欠落している編集コンテンツであっても原コンテンツに対する同一性の検出を高精度に行うことが可能である。放送番組を用いて本方式の評価実験を行った。その結果、各番組の特徴量は相互に良く分散して分布すると共に、編集番組の特徴量は対応するオリジナル番組の特徴量に近接して分布することから、編集され不正に流通するコンテンツの識別を本方式により効果的に行うことが可能であることを明らかにした。

キーワード 動画像特徴量, 信号処理, 多変量解析, 放送番組

## Method for Extracting Characteristic Information of Video using Multivariate Analysis

Keigo Majima<sup>†</sup> Go Ohtake<sup>†</sup> Seiichi Gohshi<sup>†</sup> Kazuto Ogawa<sup>†</sup> Itsuro Murota<sup>‡</sup>  
Seiichi Namba<sup>†</sup> and Kimiyuki Oyamada<sup>†</sup>

<sup>†</sup>NHK Science & Technical Research Laboratories 1-10-11 Kinuta, Setagaya-ku, Tokyo, 157-8510 Japan

<sup>‡</sup>NHK Broadcast Engineering Department 2-2-1 Jinnan, Shibuya-ku, Tokyo, 150-8001 Japan

E-mail: majima.k-fu@nhk.or.jp

**Abstract** A method for extracting characteristic information for identifying video content is proposed in this paper. It extracts characteristic information from video content using signal processing and multivariate analysis. It enables the precise identification of edited content with the original content, representing the whole content with a small amount of characteristic information data. Experiments using broadcasting content give good results: the values of the data for the original content spread enough, whereas those for the edited content are close to the original ones.

**Keyword** characteristic information of video, signal processing, multivariate analysis, broadcasting content

### 1. はじめに

近年のネットワークの高速化や記録媒体の大容量化に伴い、大容量のデジタルコンテンツを公衆通信回線により配信または光ディスクなどの記録媒体により流通させることが容易に行える環境が実現されている。デジタルコンテンツは、配送や蓄積が容易である一方で、権利者や配信事業者(以下、「プロバイダー」と呼ぶ)の許可なくコンテンツを複製、再送信し、または改ざんするなどの不法行為を行うこともまた容易であり、このことがコンテンツ流通の大きな阻害要因となっている。特に、権利者の許可無くコンテンツの複製、再送信を行う不法行為は、権利者に多大な金銭的不利益をもたらすため、不正流通コンテンツを高精

度かつ短時間に識別し、不法利用の抑止を図ることができる技術的手段が求められている。

また、インターネットの普及やデジタル情報機器の高性能化に伴い、動画像など大容量マルチメディアデータの利用が一般化しつつある。このため、インターネット上やストレージ内に存在する膨大なデジタルコンテンツの中から、利用者が求める特定のコンテンツを効率良く検索する手法が求められている。

MPEG-21などの標準化では、不正流通コンテンツの識別を電子透かしや“fingerprint”<sup>1)</sup>により行うことが検討されている<sup>2)</sup>が、画像処理や編集などの攻撃に弱く、適切な解決策が見出せていないのが現状である。

一方、コンテンツの検索等を行う目的で、映像デー

タの輝度や色情報等からその特徴を現すデータを抽出し、これらをコンテンツ間の同一性や類似性の検定に利用する特徴量抽出技術が有効な手段と考えられている。例えば、連続した動画像フレームの輝度や色信号データの時間軌跡を用い、類似性検定を行う特徴量抽出技術が提案されている<sup>3)</sup>。これを編集されたコンテンツの識別に適用した場合、時間軸波形を特徴量データとして直接比較するため、検出精度が低下すると共に、長時間の放送番組などに対しては特徴量のデータ量が膨大となる。

また、映像信号の特徴を記述するための特徴量抽出アルゴリズムが MPEG-7「ビジュアル記述」に規定されている。「ビジュアル記述」は、主として、映像信号の類似検索・フィルタリングのために用いられることが想定されており、このうち色形等の低レベル特徴量記述として、色の空間的な配置を周波数軸上で表現する色配置 (Color Layout) 記述子が定義されている<sup>4)</sup>。本手法は、画像フレームに対して高精度の検索を実現しているが、動画像コンテンツに適用した場合、特徴量のデータ量が多大となり、また、時間的な編集に対する耐性も考慮されていない。

そこで本稿では、主として不正流通コンテンツの識別を目的として、コンテンツが編集され部分的にデータが失われても原コンテンツとの同一性を検出可能とする特徴量抽出方式を提案する。

## 2. 不正流通コンテンツの識別

図1に、動画像の特徴量を用いた不正流通コンテンツの識別方法を示す。プロバイダーにおいて、外部のインターネットや光ディスク等の記録媒体により流通するコンテンツの特徴量を抽出し、プロバイダーが保有する参照コンテンツの特徴量と比較することにより、両者の同一性を検出する。この場合、例えば適当な閾値を設け、両者の特徴量間の差が閾値以下であれば、同一コンテンツと判定するものとする。

## 3. 動画像特徴量抽出方式

本章では、新たに提案する動画像特徴量抽出方式のアルゴリズムについて述べる。

### 3.1. 動画像コンテンツの三次元周波数解析<sup>5)</sup>

動画像コンテンツの三次元周波数解析手法について説明する。本手法は、動画像コンテンツに対し、空間および時間方向に関して周波数解析を行うものである。先ず、図2に示す様に、動画像データとして輝度信号 (Y)、色差信号 (Cb, Cr) を用い、各フレームを水平8ブロック x 垂直8ブロックに分割し、各ブロックにおける全画素データを平均化して8画素 x 8画素の縮小画像を作成する。次に、縮小画像の Y, Cb, Cr

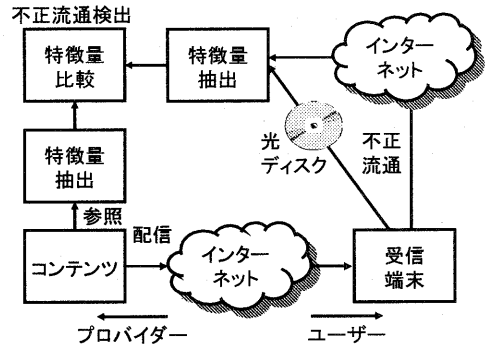


図1 不正流通コンテンツの識別方法

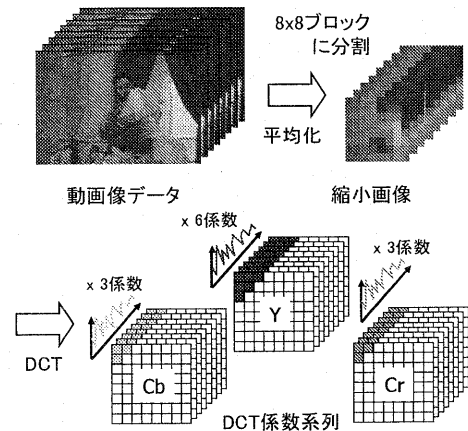


図2 DCT係数の時系列データの作成

の画素データから DCT 係数を算出し、これらをフレーム方向に並べ、DCT 係数の時系列データを作成する。ここでは、前述の MPEG-7「ビジュアル記述」における色配置記述子に準じて、輝度信号については DC から 6 係数、色差信号については各々 DC から 3 係数、合計 12 係数を用いた。図3に、動画像コンテンツの輝度信号における DC 係数の時間変化を示す。DC 係数を 3000 フレームについて時系列にプロットした。縦軸は DC 係数値、横軸は連続するフレーム番号を表す。また、図4は、図3のグラフの一部を時間軸 (フレーム) 方向に拡大して示したものである。図から、隣接するフレームの係数の差は小さく、5秒 (150 フレーム) 程度の時間内では、シーンチェンジなど大きな画面変化が発生する場合以外は係数の変動は少ないことがわかる。そこで、DCT 係数データを一定間隔ごとに選択して配列し、この操作を1フレームずつ順次ずらしながら繰り返すことにより、複数のサンプルデータ系列の組を得る。このような操作を行うことにより、動画像コンテンツ全体の DCT 係数の時間変化を、その長さ

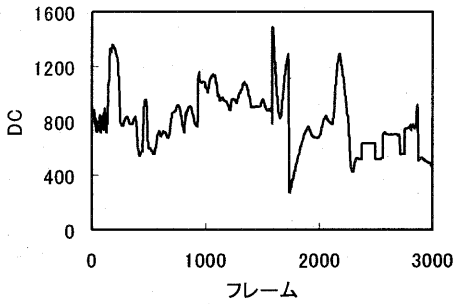


図3 輝度信号における DC 係数の時間変化

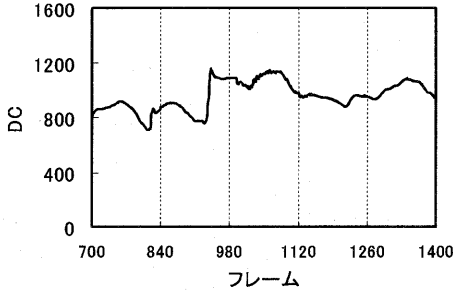


図4 輝度信号における DC 係数の時間変化 (拡大)

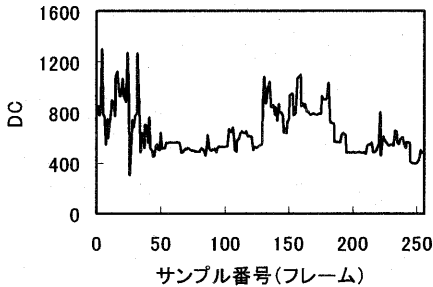


図5 輝度信号における DC 係数サンプルデータ系列

よらず、少ないデータ数により表現することができ、  
 妥当な長さのデータ系列に対して以下の時間軸方向の  
 周波数変換処理を行うことが可能となる。図5に、約  
 10分間(17920フレーム)の動画コンテンツの輝度  
 信号における DC 係数のサンプルデータ系列波形の事  
 例を示す。ここでは、動画コンテンツ全体の DC 係  
 数を70フレーム間隔で選択して配列している。従っ  
 て、このような256サンプルからなる時系列波形デー  
 タが70組得られることになる。

続いて、図6に示すように、輝度信号の DCT 係  
 数  $Y_0$ - $Y_5$ 、および各色差信号の DCT 係数  $Cb_0$ - $Cb_2$ 、  
 $Cr_0$ - $Cr_2$  において、複数組のサンプルデータ系列それ  
 ぞれに対して FFT を行い平均化する。その結果、12  
 の周波数応答波形が得られることになる。図7に、周

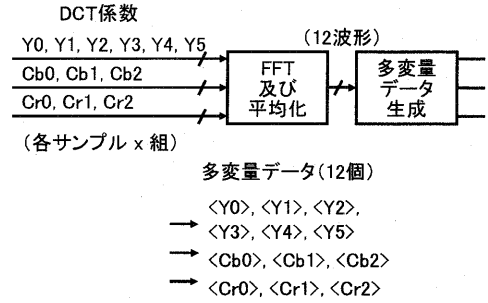


図6 多変量データの生成

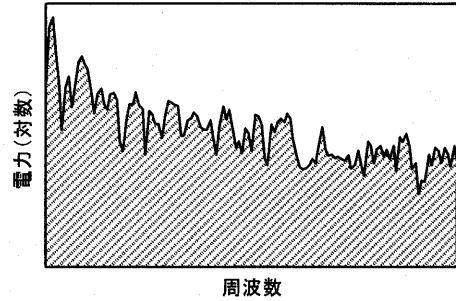


図7 多変量データの生成

波数応答波形の事例を示す。縦軸は対数表現による電  
 力、横軸は右端をナイキスト周波数とする周波数を表  
 す。次に、各 DCT 係数において、対応する周波数応  
 答データの DC からナイキスト周波数までの電力の対数  
 データを総和する、すなわち図7の斜線部の面積を求  
 めることにより、12個の多変量データ  $\langle Y_0 \rangle$ ,  $\langle Y_1 \rangle$ ,  
 $\langle Y_2 \rangle$ ,  $\langle Y_3 \rangle$ ,  $\langle Y_4 \rangle$ ,  $\langle Y_5 \rangle$ ,  $\langle Cb_0 \rangle$ ,  $\langle Cb_1 \rangle$ ,  $\langle Cb_2 \rangle$ ,  
 $\langle Cr_0 \rangle$ ,  $\langle Cr_1 \rangle$ ,  $\langle Cr_2 \rangle$  を生成する。なお、以後、 $\langle \cdot \rangle$   
 の表記は、その成分に対する変量データを意味するも  
 のとする。

### 3.2. 動画コンテンツの多変量解析<sup>6)</sup>

本節では、動画コンテンツから得られた多変量デ  
 ータを用いて、多変量解析<sup>7)</sup>における主成分分析に  
 より特徴量を抽出する手法について説明する。

主成分分析は、複数の変量を線形結合した合成変  
 量の分散を最大とする係数を求め、この時の合成変  
 量を主成分としてデータの分析を行うものである。

まず、動画コンテンツの特徴量となる主成分を計  
 算するために用いる係数(以下、これを「特徴係数」  
 と呼ぶ。)を求める。そのため、複数の基準コンテン  
 ツを用いて、複数組の多変量データを生成し、これら  
 を基準変量データとする。ここでは、基準コンテン  
 ツの数を  $m$ 、各コンテンツの変量データ数を  $n$  とする。  
 (1) 式に、 $i$  番目のコンテンツの多変量データ  $x_{ij}$  を用いた  
 合成変量  $u_i$  を示す。

$$u_i = a_1 x_{i1} + \dots + a_j x_{ij} + \dots + a_n x_{in} \quad (1)$$

但し、 $i, j$  はそれぞれ  $1 < i < m, 1 < j < n$  の整数、  
 係数  $a_1, \dots, a_j, \dots, a_n$  は、  
 $a_1^2 + \dots + a_j^2 + \dots + a_n^2 = 1$  (2)  
 を満たす定数とする。

また、(3) 式に、合成変数  $u_i$  の分散  $Su^2$  を示す。

$$\begin{aligned} Su^2 &= \{1 / (m - 1)\} \{(u_1 - \bar{u})^2 + \dots + (u_i - \bar{u})^2 \\ &+ \dots + (u_m - \bar{u})^2\} \\ &= a_1^2 S_{11} + \dots + a_j^2 S_{jj} + \dots + a_n^2 S_{nn} \\ &+ 2a_1 a_2 S_{12} + \dots + 2a_1 a_n S_{1n} \\ &+ \dots + 2a_j a_{j+1} S_{j(j+1)} + \dots + 2a_j a_n S_{jn} + \dots \\ &+ 2a_{n-1} a_n S_{(n-1)n} \end{aligned} \quad (3)$$

但し、 $\bar{u}$  は、 $u_i$  の平均値、

$S_{11}, \dots, S_{jj}, \dots, S_{nn}$  は、それぞれ変数  $x_{i1}, \dots, x_{ij}, \dots, x_{in}$  の分散、

$S_{12}, \dots, S_{j(j+1)}, \dots, S_{(n-1)n}$  は、それぞれ変数  $x_{i1}$  と  $x_{i2}, \dots, x_{ij}$  と  $x_{i(j+1)}, \dots, x_{i(n-1)}$  と  $x_{in}$  の共分散を表す。

主成分分析は、条件式 (2) の下で、(3) 式に示した合成変数  $u_i$  の分散  $Su^2$  を最大とする係数  $a_1, \dots, a_j, \dots, a_n$  を求めるものである。ラグランジュの定数変化法を用いて、(4) 式の条件式を得る。

$$S \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} = \lambda \begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} \quad (4)$$

但し、

$$S = \begin{pmatrix} S_{11} & \dots & S_{1j} & \dots & S_{1n} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ S_{1j} & \dots & S_{jj} & \dots & S_{jn} \\ \vdots & & \vdots & & \vdots \\ \vdots & & \vdots & & \vdots \\ S_{1n} & \dots & S_{jn} & \dots & S_{nn} \end{pmatrix} \quad (5)$$

また、 $\lambda$  は (4) 式に固有の定数を表す。

(5) 式に示す行列  $S$  は、各変量の分散・共分散行列であり、(4) 式は、係数の組  $(a_1, \dots, a_j, \dots, a_n)$  を固有ベクトル、定数  $\lambda$  を固有値とする行列  $S$  の固有値問題となる。そこで、行列  $S$  の最大の固有値に対応する固有ベクトルを第 1 主成分の特徴係数とし、次に行列  $S$  の 2 番目に大きい固有値に対応する固有ベクトルを第 2 主成分の特徴係数とし、行列  $S$  の 3 番目に大きい固有値に対応する固有ベクトルを第 3 主成分の特徴係数とする。

図 8 に、多変量解析を利用した動画コンテンツの特徴量算出手順を示す。ここでは、 $n = 12$  として、3.1 節で示した多変量データ  $\langle Y0 \rangle, \langle Y1 \rangle, \langle Y2 \rangle, \langle Y3 \rangle, \langle Y4 \rangle, \langle Y5 \rangle, \langle Cb0 \rangle, \langle Cb1 \rangle, \langle Cb2 \rangle, \langle Cr0 \rangle, \langle Cr1 \rangle, \langle Cr2 \rangle$  を用いることとし、予め  $m$  個の基準コンテンツの多変量データから第 1-3 主成分の特徴係数  $(a_1 \langle 1 \rangle, \dots, a_{12} \langle 1 \rangle), (a_1 \langle 2 \rangle, \dots, a_{12} \langle 2 \rangle), (a_1 \langle 3 \rangle, \dots, a_{12} \langle 3 \rangle)$  を算出する。

続いて、特徴量抽出を行う動画コンテンツの 12 個の多変量データから、3 つの特徴係数を用いて (6) - (8) 式に示す第 1 - 3 主成分  $u \langle 1 \rangle, u \langle 2 \rangle, u \langle 3 \rangle$  を計算する。

$$\begin{aligned} u \langle 1 \rangle &= a_1 \langle 1 \rangle \langle Y0 \rangle + \dots + a_6 \langle 1 \rangle \langle Y5 \rangle + a_7 \langle 1 \rangle \langle Cb0 \rangle + \\ &+ a_8 \langle 1 \rangle \langle Cb1 \rangle + a_9 \langle 1 \rangle \langle Cb2 \rangle + a_{10} \langle 1 \rangle \langle Cr0 \rangle + \\ &+ a_{11} \langle 1 \rangle \langle Cr1 \rangle + a_{12} \langle 1 \rangle \langle Cr2 \rangle \end{aligned} \quad (6)$$

$$\begin{aligned} u \langle 2 \rangle &= a_1 \langle 2 \rangle \langle Y0 \rangle + \dots + a_6 \langle 2 \rangle \langle Y5 \rangle + a_7 \langle 2 \rangle \langle Cb0 \rangle + \\ &+ a_8 \langle 2 \rangle \langle Cb1 \rangle + a_9 \langle 2 \rangle \langle Cb2 \rangle + a_{10} \langle 2 \rangle \langle Cr0 \rangle + \\ &+ a_{11} \langle 2 \rangle \langle Cr1 \rangle + a_{12} \langle 2 \rangle \langle Cr2 \rangle \end{aligned} \quad (7)$$

$$\begin{aligned} u \langle 3 \rangle &= a_1 \langle 3 \rangle \langle Y0 \rangle + \dots + a_6 \langle 3 \rangle \langle Y5 \rangle + a_7 \langle 3 \rangle \langle Cb0 \rangle + \\ &+ a_8 \langle 3 \rangle \langle Cb1 \rangle + a_9 \langle 3 \rangle \langle Cb2 \rangle + a_{10} \langle 3 \rangle \langle Cr0 \rangle + \\ &+ a_{11} \langle 3 \rangle \langle Cr1 \rangle + a_{12} \langle 3 \rangle \langle Cr2 \rangle \end{aligned} \quad (8)$$

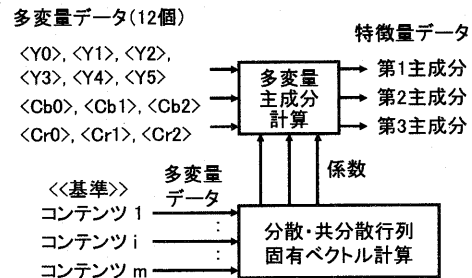


図 8 動画コンテンツの特徴量算出手順

## 4. 放送番組を用いた評価実験

### 4.1. 実験内容

本方式を用いて複数の放送番組の特徴量の分布を求めた。また、各放送番組のデータを部分的に削減することにより編集コンテンツとして特徴量を求め、原コンテンツの特徴量と比較することにより、本方式の識別性能を評価した。

### 4.2. 評価に用いた放送番組

表 1 に、本評価実験に用いた放送番組の一覧を示す。NHK-BS 標準テレビ放送における、ジャンルや内容の異なる約 10 分間の 10 番組を用いて評価を行った。

### 4.3. 実験結果

#### 4.3.1. DCT 係数サンプルデータ系列波形

表 1 に示す各放送番組を用いて、各フレームの縮小画像から DCT 係数の時系列データを求めた。図 9 (a), (b), (c) に、それぞれ番組 1 の輝度信号および色差信号における DC 係数の 1 サンプルデータ系列波形を示す。各番組において、輝度信号は DC から 6 係数、色差信号は各々 DC から 3 係数について同様の波形データを取得した。また、番組全体の 10% のデータを冒頭から削除し、編集コンテンツとして以降の処理を行った。

#### 4.3.2. 周波数応答波形

図 10 (a), (b), (c) に、それぞれ図 9 (a), (b), (c) に対応する編集前後の周波数応答波形を示す。周波数は、サンプリング周波数で正規化した。図から、編集前後の波形は良く一致することを確認した。なお、他の DCT 係数についても、同様の結果が得られている。これらの周波数応答波形データから各放送番組の多変量データを取得した。

#### 4.3.3. 多変量解析を用いた特徴量評価

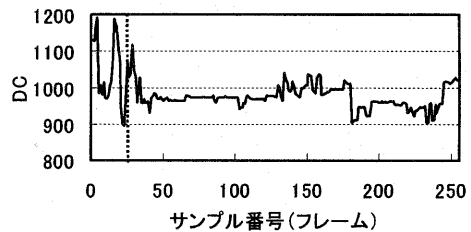
表 1 に示す各放送番組を基準コンテンツとして、多変量解析における主成分分析を行い、特徴係数を求めた。また、これらの特徴係数を用いて各放送番組の多変量データから特徴量を算出した。図 11 (a), (b) に、各放送番組の特徴量の分布を示す。併せて、番組全体の 10% を冒頭から削除して編集した番組の特徴量の分布も同図に示す。特徴量を表す各軸に対応するデータは、主成分分析における第 1-3 主成分を意味し、それぞれ 32 ビットで表現した。両図から、各オリジナル番組の特徴量は相互に良く分散して分布すると共に、編集番組の特徴量は対応するオリジナル番組の特徴量に近接して分布することを確認した。実際、各番組の第 1-3 主成分から成る 3 次元ベクトル間の差の 2 乗ノルムを評価した結果、全ての編集番組について対応するオリジナル番組との距離が他に比べて最も小さい。このことから、適当な閾値を設けることにより、編集番

表 1 評価に用いた放送番組

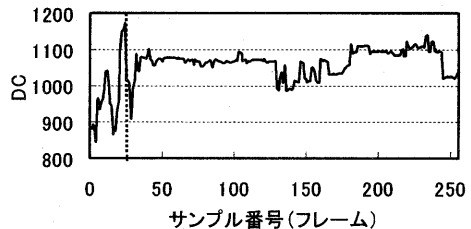
番号	番組タイトル
1	連続テレビ小説(「まんてん」)①
2	連続テレビ小説(「まんてん」)②
3	アーカイブドラマ(「夢千代日記」)
4	BSニュース50①
5	BSニュース50②
6	日本列島ふるさと発
7	BSエアロビック
8	BSおかあさんといっしょ
9	衛星アニメ劇場(「冒険者」)
10	第17回日本ゴールドディスク大賞



(a) Y 信号



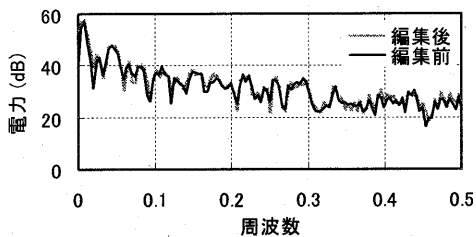
(b) Cb 信号



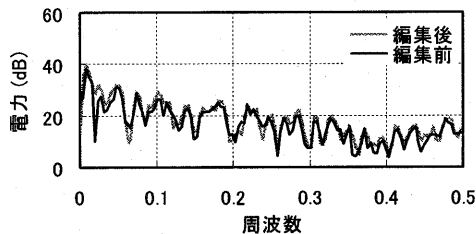
(c) Cr 信号

図 9 DC 係数のサンプルデータ系列波形

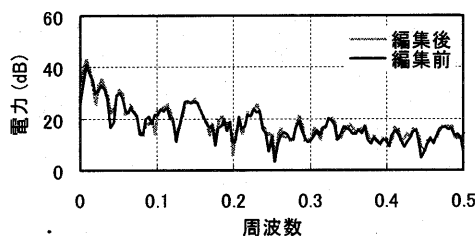
組を他番組と区別して識別することが可能であることが明らかになった。



(a) Y 信号

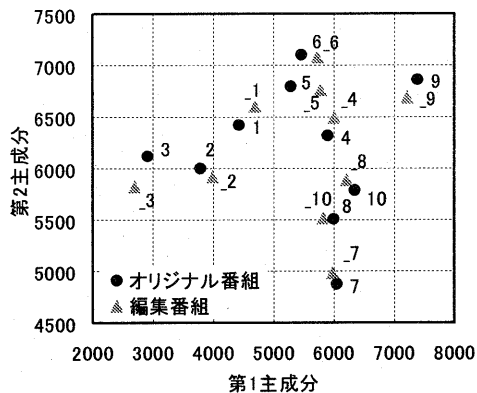


(b) Cb 信号

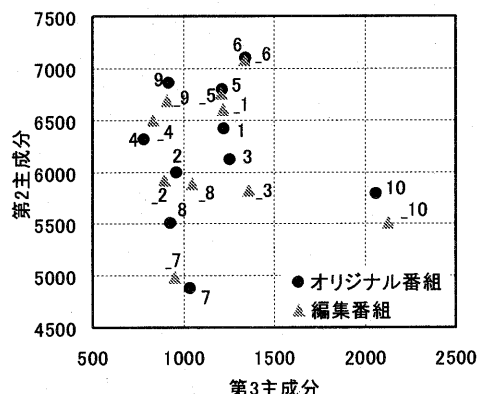


(c) Cr 信号

図 10 周波数応答波形 (DC 係数)



(a) 第 2-1 主成分



(a) 第 2-1 主成分

図 11 多変量解析を用いた番組特徴量の分布

## 5. むすび

広帯域ネットワークにおける不正流通コンテンツの識別を目的とし、新たな動画像特徴量抽出方式の提案を行った。本方式は、3次元周波数解析と多変量解析を用いて動画像コンテンツの特徴量を取得することにより、コンテンツ全体の特徴量を少ないデータ量で表現するとともに、一部のデータが欠落している編集コンテンツであっても原コンテンツに対する同一性の検出を高精度に行うことが可能である。

放送番組を用いて本方式の評価実験を行った。その結果、各番組の特徴量は相互に良く分散して分布すると共に、部分的にデータが失われた編集番組の特徴量は対応するオリジナル番組の特徴量に近接して分布することから、編集され不正に流通するコンテンツの識別を本方式により効果的に行うことが可能であることを明らかにした。

本手法は、インターネット上やストレージ内に存在する膨大な動画像コンテンツから、目的とする特定のコンテンツを検索するシステムにも適用可能である。今後は、さらに多くのコンテンツを用いて本方式の評

価を進めると共に、コンテンツの検索・識別システムや認証システムへの応用について検討する予定である。

## 文 献

- [1] <http://ipsi.fraunhofer.de/delite/Projects/MPEG7/>.
- [2] ISO/IEC JTC1/SC29/WG11 N5229 "Requirements for the Persistent Association of Identification and Description of Digital Items".
- [3] 高橋, 富永, 杉浦, 横井, 寺島, "特徴的な動画像の画紋を用いた高能率動画像検索法", 画像電子学会誌, 29, 6, pp. 818-825, 2000.
- [4] ISO/IEC 15938-3 "Information Technology - Multimedia Content Description Interface - Part 3: Visual, 6.6 Color layout", pp. 42-50, 2002.
- [5] 真島, 大竹, 合志, 小川, 室田, 難波, 小山田, "著作権保護を目的とした動画像特徴量抽出方式の検討", 2003 映情学年次大, no. 15-5, Aug. 2003.
- [6] 大竹, 真島, 合志, 小川, 室田, 難波, 小山田, "多変量解析を用いた動画像特徴量抽出方式", 2003 映情学年次大, no. 15-6, Aug. 2003.
- [7] 柳井, 高根, "多変量解析法 (新版) 現代人の統計", 朝倉書店, 東京, 1985.