

## 教育映像コンテンツのビデオポッドキャストイングと ブラウジング

柳沼 良知・鈴木 一史・清水 康敬

教育コンテンツのデジタル化、データベース化が進むにつれて、ユーザが必要なコンテンツを効率的に選び出すための検索／ブラウジング技術は重要な技術の1つとなってくる。ここでは、特に、映像コンテンツを対象とし、まず、映像配信技術の1つであるビデオポッドキャストイングによる映像コンテンツの配信機能を実装し、それらの映像を構成する画像の色特徴に応じて3次元空間中に画像を配置することによる映像のブラウジング機能を実現した。また、映像コンテンツをブラウジングする場合、映像を構成する画像の色特徴の類似によるだけでなく、同一映像内の画像は近くに表示させるといった表示の仕方が必要となる場合もある。このため、ユーザが色特徴による検索を重視するか、同一映像内の検索を重視するかに応じて、映像コンテンツの検索が効率良く行えるようにそれぞれの画像の表示位置を変え、ブラウジングを行う手法の開発をあわせて行った。

### キーワード

映像コンテンツ、ビデオポッドキャストイング、ブラウジング、検索、データベース

### 1. はじめに

計算機の処理能力の向上や記録媒体の大容量化に伴い、教育の分野においても、教育コンテンツのデジタル化、データベース化が盛んに行われるようになってきている。このような環境では、蓄積された多量の教育コンテンツの中から、ユーザが必要なコンテンツを効率的に選び出すための検索／ブラウジング技術は重要な技術の1つとなってくる。検索／ブラウジング技術に関しては、従来から、映像シーンの分類<sup>[6,8,12]</sup>、検索のための自動キーワード付与<sup>[7,13-16]</sup>等について数多くの研究がされてきた。そこで、本稿では、より実用に近い部分を目指し、まず、映像配信技術の1つであるビデオポッドキャストイングによる映像コンテンツの配信機能を実装し、それらの映像を構成する画像の色特徴に応じて3次元空間中に画像を配置することによる映像のブラウジング機能を実現した。

3次元空間中に画像を配置することによる映像検索システム<sup>[2,3,11]</sup>は従来から存在するが、2次元のライブラリを用いて、遠くの画像は縮小し、画像を重ねて表示することで擬似的に3次元を表現しているものや、表示の際にそのシステムに独自の表示ソフトウェアが必要で、表示できるPC環境が限定されているものが多かった。今

回のシステムでは、3次元空間の表示にVRMLを用いているため、一般的なVRMLのビューアで表示を行うことができ、PC環境を特に限定せずに、完全な3次元空間を実現できる。ビューアの機能により、ユーザは、この3次元空間の中を水平移動、回転等をしながら自分の必要な画像を探し出すことができる。

また、ビデオポッドキャストイングに関しては、もとのビデオポッドキャストイングには、キーワードベースの検索機能しかないため、画像検索との連携が不可欠となるが、ビデオポッドキャストイングと3次元画像検索システムとの組み合わせはない。ビデオポッドキャストイングと映像検索を組み合わせることで、一度見た映像を再び探し出したり、赤い対象物が映っている映像を探すといった、より柔軟な映像の検索を実現することができる。ビデオポッドキャストイングの実装については、2章、VRMLを用いた映像コンテンツのブラウジング手法については、3章で詳しく述べる。

一方、映像コンテンツをブラウジングする場合、同一映像内の検索が頻繁に行われるのであれば、単に色特徴の似た画像を近くに表示するのではなく、同一映像内の画像を近くに表示させた方が、より検索を効率的に行うことができる。このため、ユーザが色特徴による検索を重視するか、同一映像内の検索を重視するかに応じて、映像コンテンツの検索が効率良く行えるようにそれぞれの画像の表示位置を変え、ブラウジングを行う手法の開発をあわせて行った。これにより、すべての人に同じ検

素インタフェースを提供するのではなく、個々のユーザの検索目的に応じた適切な検索インタフェースを提供することができるようになる。同一映像内の画像を近い部分に配置する点、および、その配置をユーザの検索目的に応じて変える点は、従来のシステムにはなかったものである。この、ユーザの検索目的に応じて映像コンテンツの表示を行う手法については、4章で述べる。

## 2. ビデオポッドキャストの実装

近年、iPodをはじめとするハードディスクやフラッシュメモリを用いた携帯音楽プレーヤが広く普及し始めている。また、大学等においても、こういった携帯音楽プレーヤを語学学習等に用いる試みが行われ始めている。こういった携帯音楽プレーヤやPCを対象とし、RSSと呼ばれるメタデータの記述方式を用いて音声データを配信する仕組みがポッドキャストである。最近では、携帯音楽プレーヤに映像再生機能がついたものが出てきており、ポッドキャストと同様の方式で映像を配信するビデオポッドキャストも行われるようになってきた。

ここでは、まず、次章以降述べる映像ブラウジングの対象として利用する、ビデオポッドキャストの実装について述べる。ビデオポッドキャストと映像ブラウジング機能の関係を図1に示す。RSS (Really Simple Syndication) と呼ばれるメタデータを用いて映像の配信を行うのがビデオポッドキャストであり、ユーザは、映像配信用のWebページに付与されたRSSをビデオポッドキャストに対応したソフトウェアで読み込み、必要な映像をダウンロードして利用する。3次元表示による映像検索のページは、映像配信用のWebページと相互にリンクされており、映像検索機能を利用する場合は、ユーザは、検索結果として得られた映像のURLをビデオポッドキャストに対応したソ

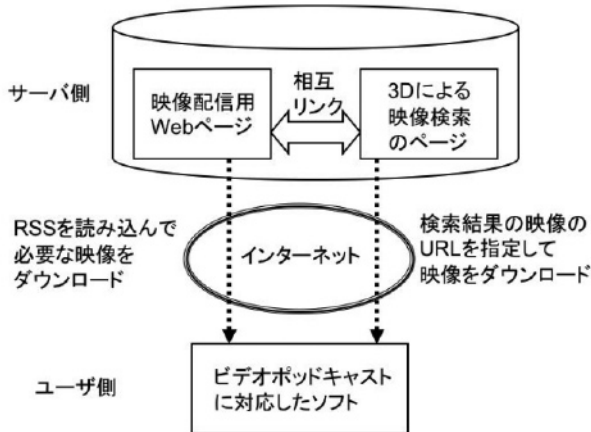


図1 システムの構成

フトウェアで指定して、映像をダウンロードして利用する。

ビデオポッドキャストで配信する映像としては、NIMEの教育支援システム研究公開サイト<sup>[10]</sup>の中の10本の映像を用いた。映像の長さは、それぞれ約1分程度である。図2(a)が映像配信用で作成したWebのトップページである。それぞれのタイトルをクリックすることで、図2(b)のように、それぞれの映像を再生するページへと移動することができる。これらの映像は、iPodで再生できる形式となっている。

ビデオポッドキャストでは、配信する映像の情報をXMLをベースとしたRSSを用いて記述する。RSSは、



(a) トップページ



(b) 各ムービーのページ  
図2 映像配信用Webページ

Webサイトの更新情報などを公開する技術として開発されたものであり、ニュースの配信やBlogの更新情報の配信などで広く用いられてきた。

今回作成したRSSを図3に示す。〈channel〉タグは、TVのチャンネルのようにまとまった番組を表すタグであり、そこから続く〈title〉タグには、チャンネルのタイトル、〈link〉タグには、そのチャンネルに関連するリンク先(ここでは、映像配信のトップページ)、〈description〉タグには、このチャンネルの説明が記述されている。また、〈language〉タグには、記述に用いられている言語、〈copyright〉タグには、著作権者、〈pubDate〉タグには、情報の更新日時が記述されている。実際に配信される映像に関する情報は、〈item〉タグを用いて記述される。〈item〉タグの中には、〈channel〉タグと同様のタグを用いて、配信する映像のタイトル、リンク先(ここでは、それぞれの映像のページ)、説明、更新日時が記述され

```
<?xml version="1.0" encoding="utf-8" ?>
<rss version="2.0">
<channel>
<title>Video Pod Casting of Open.NIME</title>
<link>http://www.nime.ac.jp/~yaginuma/vidocast/</link>
<description>Open.NIMEのVideo Pod Casting</description>
<language>ja-jp</language>
<copyright>NIME</copyright>
<pubDate>Fri, 25 Nov 2005 06:00:00 +0900</pubDate>
<item>
<title>ビデオ観察・分析支援ツールCIA0</title>
<link>http://www.nime.ac.jp/~yaginuma/vidocast/p1.html</link>
<description>ビデオ観察・分析支援ツール「CIA0(Collaborative
Implement for Active Observation)」は、メディア教育開発センター研究開発プロジェクト「国際的通用性の高い先進リソースとその流通方略の研究開発」(平成13~15年度)において開発されました。このソフトウェアには、ビデオ記録の観察、分析、検討を行う際、それらの活動を効率的に進められるように複数のビデオの同期再生、発語記録の同期表示、発語記録からのビデオ場面の検索などの機能が用意されています。カウンセリングや授業の観察学習、行動分析や会話分析など、さまざまな分野での教育や研究に利用できます。その目的は (1)教師が教室で用いる教材として、ビデオ素材の能動的な観察による協調的な知識構築や経験の共有を支援する、(2)研究者のビデオ分析ツールとして、専門家がビデオデータを詳細に分析するのを効率化し、それをプレゼンするのを支援する、(3)将来的には、(1)のような教育を遠隔で行うためのツールとして、インターネットを介して、それぞれがビデオを観察して得られた知見を持ち寄って、議論をすることを支援する。</description>
<pubDate>Fri, 25 Nov 2005 06:00:00 +0900</pubDate>
<enclosure url="http://www.nime.ac.jp/~yaginuma/vidocast/m1.m4v" length="4360715" type="video/x-m4v" />
</item>
<item>
<title>視線ポインタ</title>
<link>http://www.nime.ac.jp/~yaginuma/vidocast/p2.html</link>
<description>手を使わずに、目の動きだけで三次元空間内の仮想物体をポインティングできるインタフェースです。利用者の眼球運動を視線追跡装置で計測し、頭部位置を磁気センサーで測定します。両計測データにより、利用者の仮想空間中の視線方向が導出されます。システムは、およそ以下の仕様で動作します。 ☆遅延: 0.62秒 ☆誤差:1.34度 ☆精度:0.8度 これにより、利用者の自由な動きを許容しながら、三次元仮想空間に存在する仮想物体を瞬時にポインティングすることができます。利用する前に、視線方向をキャリブレーションする作業が必要です。</description>
<pubDate>Thu, 24 Nov 2005 06:00:00 +0900</pubDate>
<enclosure url="http://www.nime.ac.jp/~yaginuma/vidocast/m2.m4v" length="3991299" type="video/x-m4v" />
</item>
...
</channel>
</rss>
```

図3 RSS ファイル

ている。映像、画像、音声といったメディアをRSSで扱う場合は、〈enclosure〉タグを利用し、〈enclosure url="http://www.nime.ac.jp/~yaginuma/vidocast/m1.m4v" length="4360715" type="video/x-m4v"/>のように、関連するデータを記述することができる。

作成されたRSSファイルは、情報を配信するWebサイト上に置かれる。今回作成したRSSファイルは、図2(a)のトップページに書かれているように、http://www.nime.ac.jp/~yaginuma/vidocast/index.rssに置かれている。ユーザは、iTunes<sup>[5]</sup>等のビデオポッドキャストに対応したソフトウェアでこのRSSファイルを読み込むことで映像の取り込み、再生を行うことができる。(iTunesの場合、「詳細設定」メニューから「Podcastを登録」を選び、RSSファイルのURLを登録する。)

図4(a)は、図3のRSSデータをiTunesで読み込んだ画面である。最初の行には、RSSに記述されたチャンネルのタイトル、更新日時、説明が表示されており、2行目以降には、そのチャンネルに含まれる映像のタイトル、更新日時、説明が表示されている。説明については、別



(a) メタデータの読み込み



(b) ムービーの再生

図4 RSSの読み込みと映像再生



ウィンドウで全文を読むことができる。必要と判断した映像は、タイトル脇にある取り込みボタンをクリックすることでPCに取り込むことができる。図4(b)は、取り込んだ映像の再生画面である。これらの映像は映像再生に対応したiPod等へ書き出して再生することもできる。

以上により、携帯映像端末等への映像配信を行うためのビデオポッドキャストの実装を行った。しかし、ビデオポッドキャストでは、キーワードでしか検索を行うことができない。これらの映像を対象とした色特徴による検索機能の実現については、次章で述べる。

### 3. VRMLを用いた映像コンテンツのブラウジング

ビデオポッドキャストにより多量の映像が配信される場合、必要なコンテンツを効率的に選び出すための検索／ブラウジング機能が重要になる。しかし、ビデオポッドキャストのもともとの機能としては、キーワードによる検索機能しかないため、多量の映像コンテンツを扱う場合、色特徴による検索機能の実現は不可欠なものとなる。

ここでは、映像を構成する画像の色特徴に応じて3次元空間中に画像を配置することにより映像のブラウジングを実現した。ユーザは、3次元表示された画像の中から欲しい映像を探し、それをiTunes等のソフトウェアで自分のPC端末にダウンロードし、その後、PCで再生、あるいはiPodなどのプレーヤで視聴する。(iTunesの場合、具体的には、「詳細設定」メニューから「ストリームを開く」を選び、図2(b)のような映像の検索結果のページに記されている映像のURLを入力することで、映像の取り込みを行う。)

今回、画像の3次元表示に用いたのはビデオポッドキャストの実装に用いたそれぞれ約1分の映像10本で、それらの映像から、2秒ごとに取り出した305枚の画像に対して処理を行った。

具体的な処理手順は、以下の通りである。

1) 処理用に画像を一定の大きさの非圧縮のRGBファイルに変換する。また、表示用に画像を一定の大きさのJPEGファイルに変換する。画像の変換には、フリーの画像処理ソフトであるImageMagick<sup>[4]</sup>を利用した。

2) 画像のRGBは、そのまま処理を行ったのでは人間の色の感覚とは必ずしも一致しないため、RGBを人間の感覚により近い色相、彩度、明度に変換する。色相は、赤いか青いかという色合い、彩度は色の鮮やかさ、明度は明るさである。画像の特徴量として、検索キーとして重要と考えられる色相について12段階のヒストグラムを作成し、それぞれの度数を成分とする12次元のベクトルを作成した。図5左の画像から作成した色相のヒストグラムが図5右のヒストグラムであり、横軸は、赤、黄、緑、シアン、青、マゼンタといった色に対応し、



図5 色相のヒストグラム

縦軸がそれぞれの色の度数を表している。(この例では、画像に映っているPCの画面の色がシアンで面積も大きいことから、色相のヒストグラムでシアンに対応する部分の度数が大きくなっている。)この色相のヒストグラムは、映像の場面の変わり目(シーンチェンジ)の検出などで広く用いられている特徴量である。

3) 12次元のベクトルはそのままでは表示できないため、主成分分析を行い、次元を落とす。主成分分析の結果、固有値の値が1以上になるものは3つあったため、その固有値に対応した3軸を選び出し、12次元のベクトルを3次元ベクトルへと変換した。主成分分析には、オープンソースの数値計算ソフトウェアであるOctave<sup>[9]</sup>を用いた。

4) それぞれの画像を3次元ベクトルで表される位置に表示するため、3次元空間を記述する言語であるVRML (Virtual Reality Modeling Language) を用いて、図6のVRMLファイルを作成した。1つのTransform{ }の部分で1枚の画像の表示を行っている。ここでは、各画像をすべての方向から見やすくするため、各画像の3次元座標に対応する場所に立方体を描き、その表面に各画像の貼付けを行った。translationの部分で3次元の座標を指定し、geometry Box{ }の部分で立方体を描き、texture ImageTexture{ }の部分で立方体への画像の貼付けを行っている。また、それぞれの画像をクリックすると、RSSに記述されているリンク先に移動するようにした。リンク先は、urlの部分で指定されている。

作成したVRMLファイルは、VRMLのビューアを用いることで表示することができる。ここでは、VRMLのビューアとしてCortona VRML Client<sup>[1]</sup>を用いた。図7のHTMLを作成することで、Webブラウザ上でVRMLファイルを見ることができる。〈EMBED〉タグのSRCの部分で表示するVRMLファイルを指定し、WIDTH、HEIGHTで表示する大きさの指定を行っている。図8(a)は、VRMLファイルの表示結果である。類似した画像が近くに集まって表示されていることが分かる。ビューアの機能により、ユーザは、この3次元空間の中を水平移動、回転等しながら自分の必要な画像を探し出すことができる。また、それぞれの画像をクリックすると図8(b)のように、リンク先のWebページを開くことができる。

RSSには、情報の更新日時が記述されており、新しい映像のみを探したい、ある映像と同時期に更新された映像を探したい、といった場合など、映像を3次元空間で

```
#VRML V2.0 utf8
Background { skyColor 1 1 1 }
Transform{
translation 1.780872 -6.701775 -1.641594
children Anchor{
url "../videocast/p1.html"
parameter "target=_blank"
children Shape{
appearance Appearance{
texture ImageTexture { url "p1_x1.jpg" }
}
geometry Box{ size 0.3 0.3 0.3}
}
}
}
Transform{
translation -0.875815 0.641291 2.676988
children Anchor{
url "../videocast/p1.html"
parameter "target=_blank"
children Shape{
appearance Appearance{
texture ImageTexture { url "p1_x2.jpg" }
}
geometry Box{ size 0.3 0.3 0.3}
}
}
}
...

```

図6 VRML ファイル

```
<BODY BGCOLOR=white>
<BR>
<EMBED SRC="a21x.wrl"
TYPE="x-world/x-vrml"
WIDTH="900"
HEIGHT="700"
VRML_SPLASHSCREEN="FALSE"
VRML_DASHBOARD="TRUE"
VRML_BACKGROUND_COLOR="#000000"
CONTEXTMENU="TRUE"
>
</BODY>

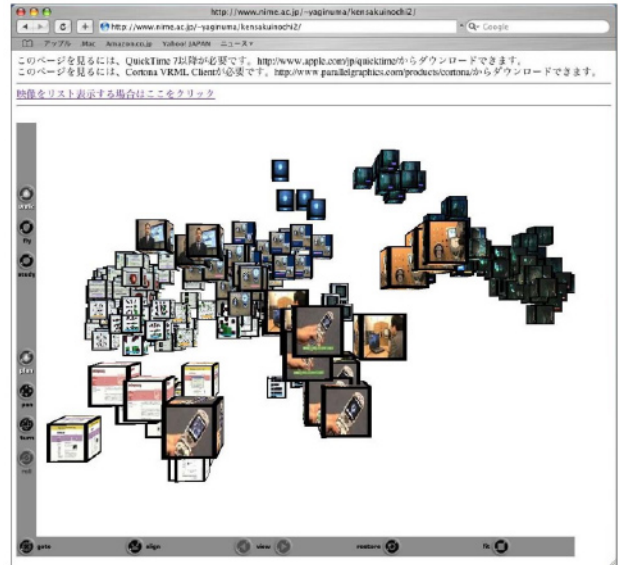
```

図7 HTML ファイル

表示する場合に、時間情報をも使って画像を配置したい場合もある。これは、色特徴の3軸のうち、上位2軸をXY軸とし、時間軸をZ軸とすることで、手前に新しい情報、奥側に古い情報を配置することで実現できる。

具体的な手順は、以下の通りである。

1) 情報が、1日前に更新された場合には、-1、2日前に更新された場合には、-2というように、時間情報



(a) VRML ファイルの表示



(b) リンク先の表示

図8 VRML ファイルの表示

を数値化する。

2) 数値化した時間の平均を求め、その平均をそれぞれの時間から引く。これにより、時間の平均をZ=0にすることができる。

3) 時間情報を表すZ軸と、色特徴の3軸のうち上位2軸を選んだXY軸とでは、スケールが異なる。このため、X軸方向の(最大値-最小値)とY軸方向の(最大値-最小値)を求め、これらの平均とZ軸方向の(最大値-最小値)が一致するように、Z軸を伸縮する。これにより、Z軸方向の広がり、X軸、Y軸方向と同程度になる。

4) このようにして求めた3次元座標をもとに、先に述べた方法と同様に、VRMLファイルを作成する。

色特徴と時間情報を使って、画像を3次元空間に配置した例を図9に示す。手前側(図では時間軸を見やすくするため斜め上の視点から見ているため、左下)に新し

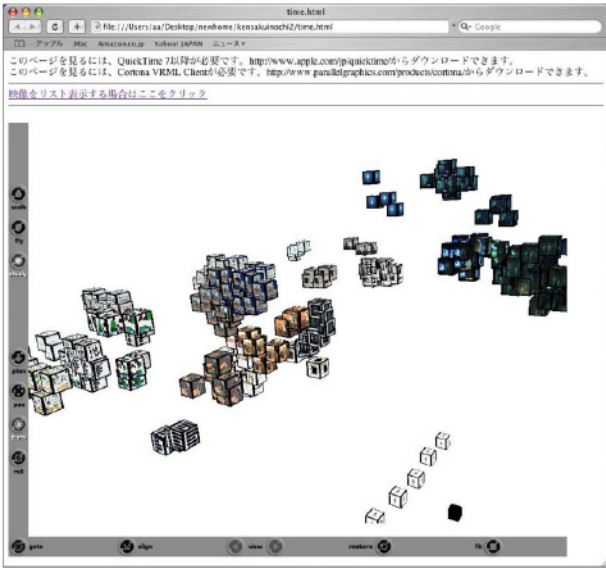


図9 時間情報の表示

い情報、奥側（図では右上）に古い情報が配置されている。また、これと垂直方向は、色特徴によって画像が配置されており、ユーザは、色特徴と時間情報を使って必要な映像を探し出すことができる。探し出した映像は、ビデオポッドキャストに対応したソフトウェアでURLを指定し、ユーザ端末にダウンロードすることで、ローカルで利用することができる。

以上により、ビデオポッドキャストコンテンツの画像特徴による検索機能を実現することができた。これにより、一度見た映像を再び探し出したり、赤い対象物が映っている映像を探すといった、より柔軟な映像の検索を実現することができる。

#### 4. 検索目的に応じた映像コンテンツの表示

3章で述べた映像ブラウジング手法は、色特徴により画像を3次元空間に配置するものであるが、それぞれの画像がどの映像に含まれるものかについては、考慮していない。しかしながら、同一映像内の検索が頻繁に行われるような場合には、同一映像内の画像は近くに表示させた方が、検索を、より効率的に行うことができる。このため、色特徴による検索を重視するか、同一映像内の検索を重視するかに応じて、画像の表示位置を変え、ブラウジングを行う手法の開発を行った。提案手法は、パラメータを変えることにより、色特徴のみにより画像を配置することも、同じ映像に属する画像を同じ場所に配置することも可能であり、また、その中間的な配置を行うことができる。

具体的な提案手法による処理を図10に示す。ここで、映像*i*に含まれる*j*番目の画像の座標を $x_{i,j}$ とする。また、映像*i*に含まれる画像の座標の平均を $m_i$ とする。この時、

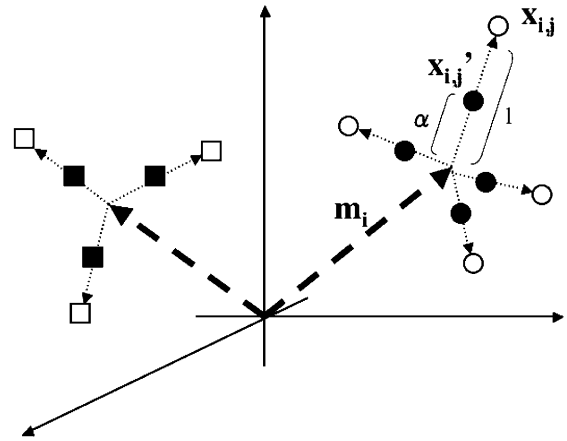


図10 検索目的に応じた映像コンテンツ表示手法

画像のベクトル $x_{i,j}$ は、

$$x_{i,j} = m_i + (x_{i,j} - m_i)$$

と書くことができる。 $(x_{i,j} - m_i)$ は、映像*i*に含まれる画像の座標の平均 $m_i$ から画像の座標 $x_{i,j}$ に向かうベクトルとなる。提案手法では、このベクトルに対して、0から1の間の値をとる変数 $a$ をかけ、拡大縮小することで画像の表示位置を変える。すなわち、画像の変換後の座標を $x'_{i,j}$ とすると、

$$x'_{i,j} = m_i + a \cdot (x_{i,j} - m_i)$$

となる。

$a$ の値が1の時、画像の表示位置はもとの画像特徴のみを使った画像位置と同じになる。また、 $a$ の値を0にした場合、映像*i*に含まれるすべての画像は、それらの座標の平均 $m_i$ の位置に表示される。そして、 $a$ の値をこれらの間に設定することで、画像特徴を優先して画像を配置するか、同じ映像に含まれる画像を近くに配置するか、重みを変えることができ、ユーザの利用目的に応じて画像の配置を変えることができるようになる。

パラメータ $a$ を変えた場合の画像の配置の変化を図11に示す。図11(a)は、 $a=1$ の場合であり、画像特徴のみによって画像が配置されている。この場合の画像の配置は、図8と同じものであるが、同じ映像内の画像に対して外接する直方体もあわせて表示することで、同じ映像に含まれる画像を探しやすいようにしている。図11(e)は、 $a=0$ の場合であり、この場合は、同じ映像に含まれる画像は、それらの平均の位置にすべての画像が表示されている。映像の数が10であるため、10箇所画像が表示されている。実際には、305枚の画像全部を表示しているが、表示位置が重なるため、各映像の最後のNIMEのロゴのみが10枚表示されている。図11(b)、図11(c)、図11(d)は、それぞれ $a=0.75$ 、 $a=0.5$ 、 $a=0.25$ の場合であり、 $a$ の値が変わるにつれて連続的に画像の位置が変化していくことが分かる。

パラメータ $a$ の変化による画像位置の変化の定量的な



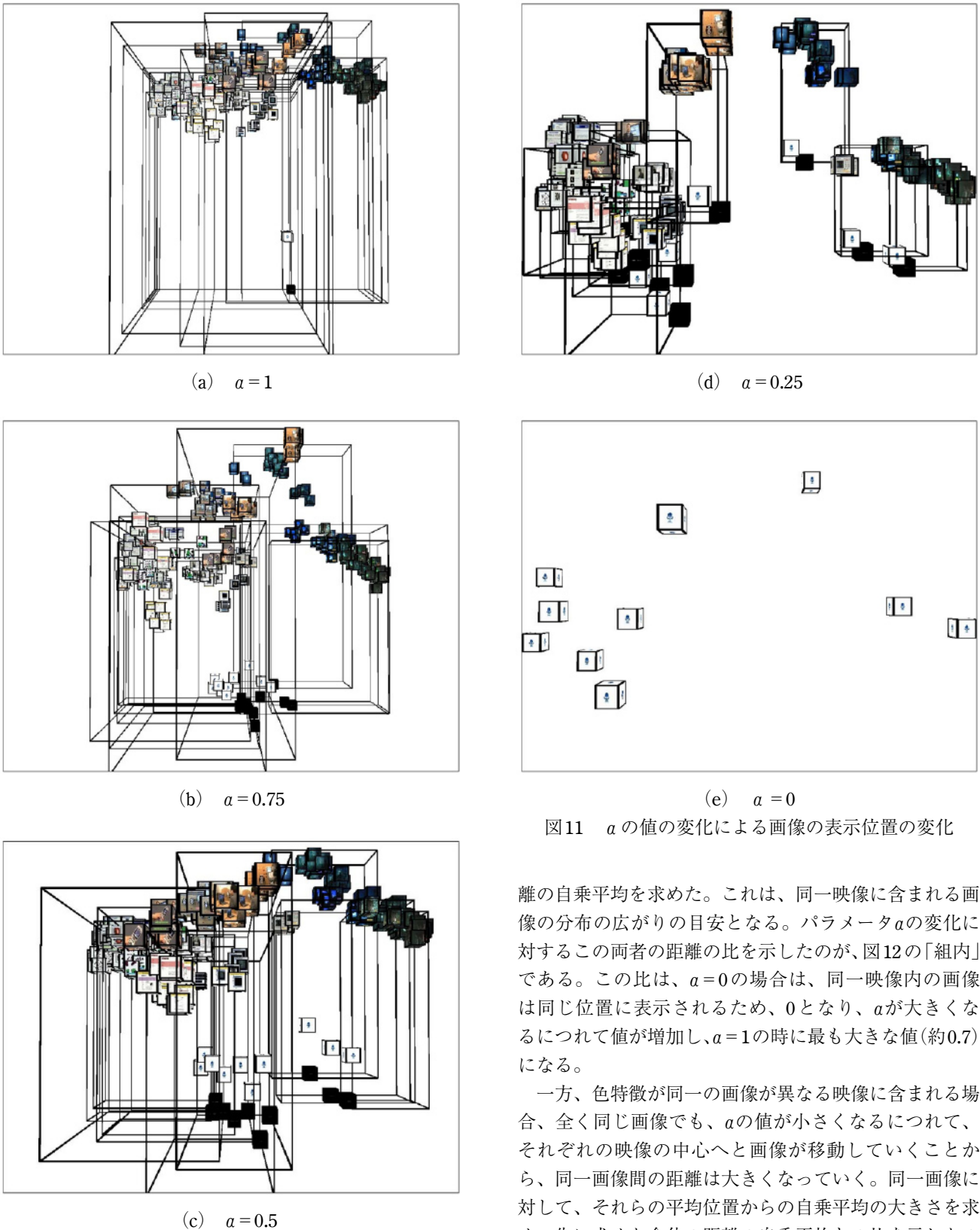


図11  $\alpha$ の値の変化による画像の表示位置の変化

離の自乗平均を求めた。これは、同一映像に含まれる画像の分布の広がりを目安となる。パラメータ $\alpha$ の変化に対するこの両者の距離の比を示したのが、図12の「組内」である。この比は、 $\alpha=0$ の場合、同一映像内の画像は同じ位置に表示されるため、0となり、 $\alpha$ が大きくなるにつれて値が増加し、 $\alpha=1$ の時に最も大きな値(約0.7)になる。

一方、色特徴が同一の画像が異なる映像に含まれる場合、全く同じ画像でも、 $\alpha$ の値が小さくなるにつれて、それぞれの映像の中心へと画像が移動していくことから、同一画像間の距離は大きくなっていく。同一画像に対して、それらの平均位置からの自乗平均の大きさを求め、先に求めた全体の距離の自乗平均との比を示したのが、図12の「同一画像」である。(対象とした画像は、すべての映像の最後に表示されるNIMEのロゴの画像である。)  $\alpha=1$ の場合、同一画像は同じ位置に表示されるため、値が0となっており、 $\alpha$ の値が小さくなればなるほど、ほぼりニアに1に近づいていくことが分かる。例

評価を行うため、まず、すべての画像の平均位置からそれぞれの画像への距離の自乗平均を求めた。この値は、すべての画像の分布の広がりを目安となる。次に、同じ映像に属する画像の平均位置からそれぞれの画像への距

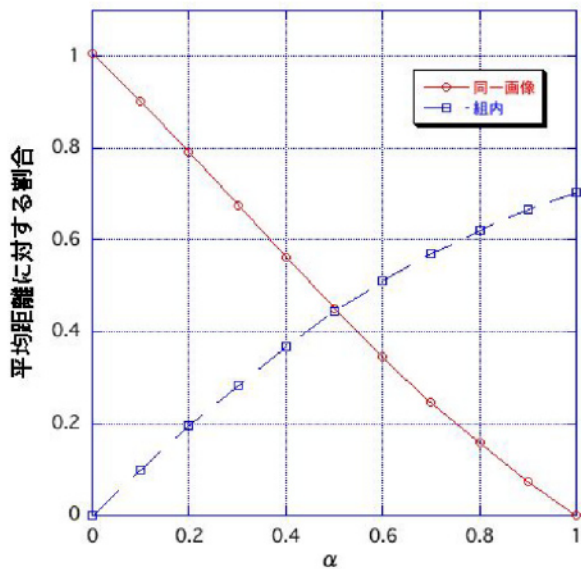


図12 画像の分布の大きさ

例えば、 $\alpha$ の値が0.5の場合、同一画像の自乗平均距離と、同一映像内の画像の自乗平均距離がほぼ等しくなることが分かる。このため、例えば、ユーザが色特徴による検索と同一映像内に含まれる画像の検索を同程度の割合で行う場合には、 $\alpha$ を0.5と設定すれば良い。

以上により、ユーザの検索目的に応じて、画像の表示位置を変え、検索を効率的に行えるようにするブラウジング手法を実現した。これにより、すべての人に同じ検索インタフェースを提供するのではなく、個々のユーザの検索目的に応じた適切な検索インタフェースを提供することができる。

今回の実験では、映像を3次元空間で表示するためVRMLを利用した。これは、3次元空間での配置等の情報をテキストファイルで作成すれば、プラグインを用いることでWebブラウザ上で表示可能であり、また、3次元空間内の移動や回転はプラグイン側で行われるため、こういった操作のプログラムを別途作成する必要がないためである。しかしながら、利用環境によっては、Webブラウザ内での表示ができない場合などもあり、また、一般にプラグインに用意された以上の複雑な処理を行うことが難しい。これを解決するためには、処理的には複雑になるものの、Java環境で3D処理を実現するJava3D等を用いることが考えられる。また、今回は、一度に表示するデータ数が少なく、すべてのデータをVRMLで読み込み、処理を行っているが、データ数が数万、数十万と増えた場合には、サーバ側にデータベースを用意し、必要なデータのみを処理に用いるといったことが必要となる。このような教育コンテンツのデータベースとしてNIMEには、NIME-gladがあり、こういったデータベースと映像検索システムの連携は今後の課題である。NIME-gladのコンテンツは、LOM (Learning

Object Metadata) と呼ばれる学習コンテンツ用のメタデータによりデータが管理されているが、こういったキーワードベースの検索キーと、今回行ったような映像特徴による検索キーの統合も今後の重要な課題となる。

## 5. まとめ

本稿では、ビデオポッドキャストによる映像コンテンツの配信機能を実装し、それらの画像の色特徴に応じて3次元空間中に画像を配置することによる映像のブラウジング機能を実現した。また、ユーザの検索目的に応じて、画像の表示位置を変え、効率的なブラウジングを実現する手法についてもあわせて述べた。

今後は、映像コンテンツの色や形といった画像特徴による検索機能の高度化を更に行うとともに、映像コンテンツのキーワードによる検索機能との連携等について、更に検討を行っていく予定である。

(平成18年1月17日受付)

## 参考文献

- [1] Cortona VRML Client, <http://www.parallelgraphics.com/products/cortona/>
- [2] 大規模マルチメディア検索システム (MultiMedia Meister), <http://www.ntt.co.jp/RD/OFIS/active/2005pdf/hot/ap/01.html>
- [3] 映像中の見たいシーンを眺めて探せる新技術を開発, <http://pr.fujitsu.com/jp/news/2002/08/23.html>
- [4] ImageMagick, <http://www.imagemagick.org/>
- [5] iTunes, <http://www.apple.com/jp/itunes/>
- [6] Lienhart, R., Effelsberg, W., and Jain, R., "Towards a visual grep: A systematic analysis of various methods to compare video sequences", Proc. of SPIE, vol. 3312, pp.271-282, 1998
- [7] 森本容介、室田真男、清水康敬、"教育用動画検索システムと時間情報同期方法の開発"、電子情報通信学会論文誌、Vol. J88-D-I, No. 10, pp.1515-1524, 2005
- [8] 長坂晃朗、宮武孝文、上田博唯、"カットの時系列コーディングに基づく映像シーンの実時間識別法"、電子情報通信学会論文誌、Vol. J79-D-II, No. 4, pp.531-537, 1996
- [9] Octave, <http://www.octave.org/>
- [10] Open. NIME: 教育支援システム研究公開サイト, <http://open.nime.ac.jp/>
- [11] 類似性検索技術を用いた画像マネジメントシステム, <http://www.englink21.com/i-eng/guest2/0c0401.htm>
- [12] 杉山善明、有木康雄、"多重部分空間法に基づくテレビニュース映像の自動分類"、電子情報通信学会論文誌、Vol. J81-D-II, No. 9, pp.2112-2119, 1998
- [13] 谷口行信、南 憲一、佐藤 隆、桑野秀豪、児島治彦、外村佳伸、"SceneCabinet: 映像解析技術を統合した映像インデクシングシステム"、電子情報通信学会論文誌、Vol. J84-D-II, No. 6, pp.1112-1121, 2001
- [14] Wactlar, H.D., Kanade, T., Smith, M.A., and Stevens, S.M., "Intelligent access to digital video: The informedia project",



IEEE Computer, vol. 29, no. 5, pp.46-52, 1996

[15] 柳沼良知、坂内正夫、“DP マッチングを用いたドラマ映像・音声・シナリオ文書の対応付け手法の一提案”、電子情報通信学会論文誌、Vol. J79-D-II、No. 5、pp. 747-755、1996

[16] 柳沼良知、鈴木一史、清水康敬、“RSSを用いた教育用映像／画像のブラウジングの検討”、電子情報通信学会技術研究報告、105(205)、pp.1-6、2005



やぎぬま よしとも  
柳沼 良知

昭63東大・工・計数卒。平2同大大学院修士課程（理・物理）了。同年東大生産技術研究所技官。平5同助手。平12メディア教育開発センター助教授。平13より総合研究大学院大学助教授併任。工博。マルチメディア処理、マルチメディアデータベースなどの研究に従事。



すずき もとふみ  
鈴木 一史

1994米国ユタ州立大学コンピューターサイエンス学部卒（BS）。1997筑波大学大学院理工学研究科修士課程了。2000筑波大学連携大学院（電子技術総合研究所）工学研究科博士課程了（工博）。2000文部科学省大学共同利用機関・メディア教育開発センター研究開発部助手。2004独立行政法人・メディア教育開発センター研究開発部助手。2005同センター助教授。情報処理学会、画像電子学会、日本感性工学会、ACM、IEEE各会員。



しみず やすたか  
清水 康敬

1940年長野県生まれ。東京工業大学工学部電気工学科卒業。東京工業大学助手、助教授、教授、教育工学開発センター長、大学院社会理工学研究科長を経て、2001年3月に定年退職、東京工業大学名誉教授。現在、独立行政法人メディア教育開発センター理事長、国立教育政策研究所・教育研究情報センター長（兼務）。日本教育工学会前会長。工学博士。

## Video Podcasting and Browsing of Educational Movie Contents

Yoshitomo Yaginuma · Motofumi Suzuki · Yasutaka Shimizu

When educational contents are digitized and managed by databases, the retrieval/browsing function which enables users to acquire necessary contents efficiently becomes one of the most important techniques. In this paper, at first, video podcasting of movie contents, which is one of the movie delivery methods, is realized. Then, movie browsing method, which places images in 3D space based on color features, is realized. When movie contents are displayed, it is also important to display them by placing images which belong to the same movie close each other. Therefore, a display method is also developed, which changes locations of images according to whether the user wants to retrieve by color features or wants to retrieve within the movie.

### Keywords

Movie Contents, Video Podcasting, Browsing, Retrieval, Database