2. 生活を見る

2-1 マルチメディアコンテンツの自動解析

——ニュースデータベースの構築に向けて——

有木 康雄[†]

1. ま え が き

映像や音声は、コンテンツ(情報の内容)を表現して人間に伝えるためのメディア(媒介物)である。映像や音声がデジタル化された結果、それはハードディスクに蓄積され、コンピュータによって処理されるようになった。放送や通信は、距離を超えて情報を伝達するためのメディアである。コンテンツを表現している映像や音声デジタル化されると、通信や放送という伝達メディアもデジタル化され、コンピュータと一層強く結びつくようになる。いわゆるC&C(Computer and Communication)技術の発展である。

このC&Cは水道や電気のようなライフラインであり、情報が流れるためのインフラストラクチャである。しかし、いくらC&Cが発達してもその中を流れるコンテンツが豊かでないと、文化的な社会を形成することは難しい。この点から、C&C(Contents)が21世紀のキーワードであるといわれている。

放送がデジタル化され、CS、BS、地上波、CATVを区別しない統合デジタル放送が始まると、放送は巨大なコンテンツ生産空間となる。通信の代表であるインターネットは、FTTHや高速ホームネットワークによって、巨大なサイバースペースとなるであろう。メディアのデジタル化がもたらすものは、コンテンツの大量

生産であり、我々は有限の時間のうちに必要なものを検索するすべを持たなければならない。

この点から、本稿では、マルチメディアコンテンツのデジタル化と、大量のコンテンツの中から必要なものを取り出すコンテンツの自動解析やインデキシング、分類方法について述べる。

2. マルチメディアコンテンツのデジタル化

2.1 放送のデジタル化

21世紀は放送のデジタル化で始まるといっても過言ではないであろう。すでにデジタル放送を開始しているCS放送に加え、BS放送、地上波放送のデジタル化が2000年、2003年に予定されている。CATVのデジタル化は時間の問題であり、インターネットにおいてはデジタル放送がすでに始まっている。放送がデジタル化されると、画質が良くなるとともに、デジタル圧縮技術により従来より放映できるチャンネル数が多くなる。すなわち、放映されるコンテンツ(情報の内容)がますます多くなるのである。すべてを見る時間はなく、自分にとって必要なコンテンツだけを、取捨選択して情報を取らざるを得ないが、このコンテンツの取捨選択に時間をかけるわけにはいかない。この点から、コンテンツの自動解析が必要になってくる。

放送がデジタル化されることによって、放映されたコンテンツは、いつでもハードディスクなどの記憶装置に蓄積され、計算機で処理することができるようになる。したがって、CS、

BS、地上波、CATV、インターネットからのデジタル放送などを区別なく受信できるセットトップボックスや、受信したコンテンツを8時間ほどハードディスクに録画しておくホームサーバが、家庭に設置されるようになるであろう。ホームサーバから必要とするコンテンツを高速に検索したり、番組のダイジェストを再生するには、コンテンツ製作サイドで付与され、放送とともに送られてくるインデックス情報が必要である。これが付与されていない場合には、ホームサーバ上のコンテンツを自動解析するソフトが必要となる。

NHKではISDBという統合デジタル放送と、ISTVという統合サービス型受像機の構想が進められている。これは、見たい番組をホームサーバに録画する「一括自動受信」だけではなく、ニュースや天気予報など最新の情報を、ホームサーバ上で自動更新する「定時ごと自動更新」や、好みの俳優が出ている番組などを自動選択して録画してくれる「インテリジェント選択」などの機能が盛り込まれている。これは、必要なものの情報や自分の好みなどをあらかじめ指定しておいて、それに合うものを選択して配信してくれるプッシュ型サービスである。

2.2 通信のデジタル化

放送が、局から視聴者へという1方向であったのに対して、通信は本来的に双方向である。インターネットに加入していれば、誰でも情報を発信することができるし、情報を得ることができる。ここに、質の高いコンテンツを配信する放送と違い、雑多で分散したコンテンツを持つインターネットの性格が

[†] 龍谷大学 理工学部

“Watching of Living; Automatic Analysis of Multimedia Contents—Toward Construction of News Database—” by Yasuo Ariki (Department of Science and Technology, Ryukoku University, Shiga)

ある。インタネットオーシャンと呼ばれる大海から、自分に必要なものを高速に取ってくるには、検索エンジンが必要である。もし、雑多で分散したコンテンツの中に映像や音響といったデータが含まれていれば、通常検索エンジンで検索することは不可能であり、マルチメディアコンテンツを自動解析して検索できるソフトが必要であろう。

インタネットでは、テキストを表現する HTML、音楽を表現する MIDI、CG を表現する VRML が規格化され、映像や音声を圧縮して解凍するソフトが開発されている。これらの表現やソフトを用いて、通信カラオケやインタネット放送、インタネットライブ、さらに遠隔教育や遠隔医療が可能な状態となっている。すなわち、雑多で分散した大量のコンテンツを持つサイバースペースの誕生である。NTT は、電話局と電柱の間に敷設されている電話線を光ファイバに置き換える「 π システム」の導入を、1998 年 3 月より開始した。電柱から家庭まで、光ファイバを敷設すれば高速通信回線 FTTH (Fiber to the home) が出来上がる。家庭の中では、IEEE 1394 によりホームネットワークがはられ、ホームサーバや AV 機器間を大量の映像や音響信号が伝送されるであろう。まさに、家庭の中も外も、高速通信回線を通して、大量のマルチメディアコンテンツが溢れるのである。

2.3 デジタル化を支える技術

映像や音響信号のデジタル化は MPEG-1, 2, 4, 7 によって規格化が進められている。MPEG-1 では、CD-ROM から映像を読み出すことを目的に、1.5 Mb/s 程度に圧縮される。MPEG-2 は、デジタル放送に使われるもので、4~6 Mb/s 程度に圧縮される。2 GB のハードディスクには、約 1 時間程度の放送を録画することができる。

MPEG-4 は現在、規格化が進められており、オブジェクトベース符号化とも呼ばれている。この符号化では、映像中に映っている人物や物体、テロップや音声などはオブジェクトとして

切り出され、個別に符号化される。さらに、オブジェクトとして、CG や合成音も含めることができ、個別に符号化される。さらに、個々のオブジェクトを組合せて映像を合成するために、シーン記述も符号化されて送られる。視聴者は、ある人物や物体を消去して映像を組み立てたり、視点の位置を移動させて映像を組み立てることができる。つまり、マルチメディア処理として、映像を構成するオブジェクトに触れることができるのである。しかし、人物や物体を切り出す方法は、規格の対象外である。膨大なデータ量を持つ映像に対して、人手でオブジェクトを切り出すことは現実的ではなく、オブジェクトの解析や自動切り出しの方法が必要となるであろう。

MPEG-7 は、これから規格化が行われる符号化方式である。この符号化では、インデックス情報や注釈など、コンテンツに対する情報が付加されることになっている。この付加情報を用いることで、コンテンツの検索や分類が可能となり、ホームサーバから必要なコンテンツを高速に引き出すことができるようになる。MPEG-4 と同様、人手でインデキシングや注釈をつけることは可能ではあるが、コストが高くなる。質の高いサービスではないが、一通り可能であるというレベルを想定するなら、マルチメディアコンテンツの自動索引付けの方法が必要となるであろう。

3. コンテンツの自動解析

前章では、放送や通信のデジタル化により、マルチメディアコンテンツの自動解析が必要になることを述べた。本章では、筆者が研究している「マルチメディアコンテンツの自動解析」に関する研究内容を紹介する。

3.1 コンテンツ解析の目的

テレビを中心とする放送、新聞や書物のような出版、映画やゲームのような娯楽がデジタル化され、通信のインフラであるインタネットと融合することにより、マルチメディアの環境が構築されつつある。このマルチメディア環境は、我々が知的活動を行う情報

空間であるが、この空間には、未整理で構造化されていない分散したデータが溢れている。例えば、電子博物館や美術館、図書館が保有するデータは、一般的に組織化されているが、テレビで放映されているニュースやアニメ、ドラマ、映画のようなビット情報は、検索の対象としにくい未整理で構造化されていないデータである。

我々は知的活動を行うために、この未整理で大量のデータの中から、興味ある情報を十分に、また高速に入手したい。このためにはデータの表現形式にとらわれず、データの持つコンテンツ(情報の内容)を引き出して記述しておく必要がある。これは、図書館の本やビデオショップのビデオテープといった表現単位をバラバラにし、コンテンツ単位で情報検索できるようにすることを意味する。本には、コンテンツをブラウジングし、アクセスするための目次や索引がある。ニュースやアニメ、ドラマ、映画のようなビデオにも、連続的なビット情報の上に目次や索引、更に内容を表現する記述や要約を作成しようというのが、コンテンツ解析の目的である。

時間と労力があれば、コンテンツ製作者が索引や内容記述を作ればよいであろう。しかし、ニュースのような一過性に近いコンテンツに対しては、そういうわけにはいかず、索引付けや内容記述の自動化が望まれる。

3.2 ニュース映像の解析

放送のデジタル化とともに、多チャンネル化を迎え、ニュース番組の放映が更に多くなるものと思われる。このような状況では、多くのニュース番組の中から、ニュースのダイジェストを見たいとか、最も知りたいニュースだけを見たいといった要求が、視聴者に生じる。

ニュース番組のこのようなシステムは、ニュースオンデマンド (News On Demand: NOD) システムと呼ばれ、マルチメディアデータベースの観点から研究が進められている¹⁾。この NOD システムは、文字・映像・音声情報で構成されるニュース記事から、コンテンツにかかわる属性情報を抽出

し、ニュース記事とともにデータベースに格納するシステムである。

ニュースデータベースには、データベース記事だけでなく、記事の特徴づけるキーワード(文字、音声、映像)が属性情報として格納されているので、利用者の要求に応じてニュース記事を検索し、更に新たな番組として再構成して、利用者に提供することができる。

このNODシステムにおいて、ニュース記事から手でキーワードを抽出して索引付けすることは、膨大な量の記事を扱うことから、かなりの労力を費やすことが予想される。そのため、NODシステムでは、ニュース原稿を解析し、キーワードを自動抽出して索引付けする研究が進められている。文字・音声・映像といった利用可能なメディアから、ニュース記事の特徴づけるキーワードを、計算機によって自動的に抽出して索引付けし、記事の内容を分類したり検索したりできるようにすることが、ニュース記事に対するコンテンツ解析である。

3.3 アプローチ

ニュース番組は、政治・経済・スポーツといった内容的に異なる複数の記事から構成されており、各データベース記事は、意味を表現する音声、内容をまとめる文字、情景や状況を伝える映像といったメディアを使って表現されている。

ニュース記事への索引付けと分類では、ニュース番組に含まれている複数の記事が処理対象となるため、ニュース映像を1つ1つの記事に分割する記事切り出し処理が必要である。そこで、まず図1に示すように、放映されたニュース番組をデジタルビデオとして格納しておき、カット検出後、記事のまとまりを検出して、自動的に記事単位に分割する²⁾。

その後、意味を伝えるニュース音声に対してワードスポッティングを行い、抽出したキーワードを基に記事の分類を行う³⁾。同様に、内容をまとめるテロップ中の文字認識を行って形態素解析し、キーワードを抽出する。このキーワードを基に記事の分類を行

図1 ニュース映像の索引付けと分類に対するアプローチ

う⁴⁾。

情景や状況を伝える映像においては、映像に現れる人物の顔領域を自動的に抽出し、追跡・認識する⁵⁾。この結果得られる人物名を索引としてニュース記事に付与し、これをもとに記事

进行分类する。同様にして、同一話者の発話区間を切り出して話者認識し、これを基に記事を切り出すことも可能である⁶⁾。また、スポーツニュースについては、クラスタリングによって代表的な映像を自動抽出しておき、これを

基にスポーツニュースを分類する⁷⁾。

分類された記事に対しては、キーボードを音声またはキーボードで入力することにより、記事検索が行える。また、映像中の興味ある人物やテロップ、あるいはニュース映像中の聞き慣れない単語に対して、対話的に問い合わせることにより情報を検索することができる⁸⁾⁹⁾。

以下には、このアプローチにもとづいて、現在構築しているニュース映像の解析システムについて述べる。

4. ニュース映像の記事切り出し

4.1 DCT 特徴のクラスタ化

ニュース映像は幾つかの記事から構成されており、記事は幾つかのシーンから構成されている。さらにシーンは何枚かのフレームから構成されている。そのためシーンとシーンの変わり目であるカット点の画像(カット点フレーム)は映像全体の概要を示しており、記事を切り出すためには全フレームを処理する必要がなく、カット点フレームのみを処理すれば充分である²⁾。この点から、ニュース映像のカット点検出を先に行い、この結果を基に記事検出を行って処理時間を短縮している。

また、ニュース映像の各フレームのデータ量は多いので、ニュース映像の蓄積を考えるとフレームを圧縮しておくことが望まれる。この圧縮されたフレームに対してカット点検出を行うことができれば、ニュース映像のデータベース化にとって都合が良い。このような考えから、ニュース映像の各フレームに対して、JPEG で採用されている DCT (離散余弦変換) を行い圧縮するとともに、DCT によって得られた直流・交流成分の変化を求めてカット検出を行っている。

カット検出の従来法では、主としてフレームをいくつかのブロックに分割し、隣接するフレーム間で対応するブロックの直流成分を求め、この差が閾値を越えていればカットを検出するといった、局所的な変化に注目していた。このため、カメラフラッシュやラ

ンプの点滅などのように、画像の一部または全体の明るさが変化する場合などで、誤検出を生じるという問題があった。この点に関して、隣接するフレームは類似しているという性質に基づいて、連続するフレーム集合でクラスタを形成する方法を採用している。カットの検出は、形成された2つのクラスタを分割するフレームとして結果的に求めることができる。

4.2 カット検出の方法

ニュース映像は図2(a)に示すように、いくつものフレームから構成されている。各フレームを図2(b)に示すように、 $L \times M$ 個のブロックに分割し、各ブロックにおいて DCT を行うと、低い周波数にパワーが集中するという特徴を持つ。したがって、各ブロックから低周波成分に対応する低次の DCT 成分 K 個を抽出すると、フレーム全体では、 $L \times M \times K$ 個の特徴量が得られる。このようにして求めた特徴量は、多次元空間中の1点として表される。

同じシーン内の隣接したフレームから求めた特徴量は、多次元空間において近いところに存在し、クラスタを形成する。ところが、シーンが変わると、多次元空間中の点はこのクラスタから逸脱しはじめるので、この逸脱をもってカット検出を行うのである。

クラスタからの逸脱はオーバー数によって判定する。オーバー数とは、クラスタを図2(c)のように軸無相関の

正規分布で近似し、その第 i 軸に関する平均 μ_i と標準偏差 σ_i に対して、入力されたフレームの DCT 成分 x_i が

$$|x_i - \mu_i| > 3\sigma_i \quad (1)$$

となる成分の個数である。オーバー数が $(L \times M \times K)/2$ を越えるフレームをアウトフレームと呼び、アウトフレームが連続して P フレームを越えた場合、クラスタから逸脱したと考えてカット検出を行う。

カット検出が行われると、その後1秒間のフレームを使って新しくクラスタを形成する。カット検出が行われなければ、新しい入力フレームによってクラスタの平均と分散を更新する。

4.3 記事切り出しの方法

ニュース番組の構成は、スタジオにおいてニュースキャスタが記事内容を紹介した後、幾つかの現場のシーンに移り、またスタジオシーンに戻るといった繰り返しの構成をとっている。この構成は、スタジオシーンを始点とするループ構造として現れるので、このループ構造を検出することによって、スタジオシーンを同定することができる²⁾。

図3は、検出したカット点フレームを■で表している。類似したシーンは近くに集まるため、ニュース映像は何度もスタジオのシーンに戻りループを形成する。このループの始点となるカット点(ループポイント)を検出する。図3に示すように、たまたま小さいループを形成するような場合もあるが、

図2 クラスタの形成

図3 ループ点とカット点

スタジオのシーンのように長いループを形成することはない。

こうして検出されたループポイントがスタジオシーンとなり、2つのスタジオシーンに挟まれたフレームの時系列を、記事として切り出している。

5. テロップ認識と記事分類

5.1 テロップの切り出しと認識

テロップフレーム中では、文字の色は単一であり、縁どりがはっきりしているという特徴がある。この特徴をもとに、文字列領域を切り出し、次に個々の文字を切り出して認識する。

まず、文字の縁どりがはっきりしていることから、入力フレームに対し

て、1次微分を求めてエッジ画像を作り、横方向に投影してヒストグラムを求める(図4(b))。文字が横1列に並んでいる場合には、他の部分との差は大きく、ヒストグラムの山と谷から安定して文字領域を抽出できる(図4(c))。この結果をさらに縦方向に投影することによって、文字列領域を切り出す(図4(d))。

次に、文字の色が単一であることを利用し、すでに得られた文字列領域に対して判別分析を行い、閾値を決定して2値化¹⁰⁾する。文字の周囲の背景を除去するために、文字列の矩形に接する連結成分を除去する。こうして得られた文字列に対して、縦方向に投影

を行い、文字間を抽出して文字を切り出す。

最後に、切り出された文字に対して大きさを正規化した後、 9×9 画素の小領域ごとに輪郭線の方法を 45° おきの4方向に量子化して、加重方向指数ヒストグラム¹¹⁾を求める。1文字の大きさは 64×63 画素であり、小領域の数は 7×7 となることから、1文字は $196(7 \times 7 \times 4)$ 次元のデータとして観測空間上で表現される。この196次元の特徴空間において、部分空間法¹²⁾により文字を認識する。認識された文字に対して形態素解析を行い、2文字以上の名詞をキーワードとして抽出している。

5.2 記事分類の方法

一般にニュース記事は、いくつかの分野に分類することができる。図5は、「朝日新聞の記事データベース分類表索引¹³⁾」の一部であり、分類には大分類、中分類、小分類と呼ぶ3階層がある。大分類の数は10、中分類の数92、小分類の数737があり、キーワード約12000語が用意されている。したがって、朝日新聞の記事データベースには、最下層の小分類に平均16個ほどのキーワードが用意されている。

そこで、テロップ認識より得られたキーワードをもとに、小分類で用意されている単語とマッチングをとり、その単語がどの大分類(総類、政治、経

(a) 入力画像

(b) エッジ画像とエッジヒストグラム

(c) 文字領域画像

(d) 文字列の矩形領域

図4 テロップ文字の切り出し

図5 記事分類の階層構造

済、労働、文化、科学、社会、事件、スポーツ、国際)に対応するのかを調べることで、記事内容の分類が行える。

具体的には、形態素解析で得られたキーワードとキーワードを含む索引すべてを、分類表索引を用いて検索し、その索引がどの大分類に属しているかをカウントする。そのカウントが一番多かったものを、このキーワードによる大分類結果とする。各キーワードによる大分類結果のANDをとって、最終の分類結果としている。

6. 音声単語による記事分類

6.1 記事分類の流れ

ニュースキャストの音声に対して、単語スポッティングを適用して索引付けし、ニュース記事を分類する。この索引付けと分類システムは、次の2段階で構成されている。

- (1) まず、分類分野とキーワードの関係をもとに、キーワードが分類分野に寄与する確率を求めておく。
- (2) 次に、ニュースキャストの音声に対して、単語スポッティングを行い、キーワード列とキーワードの音響的な存在確率を求める。この抽出されたキーワードがニュース記事に対する索引

となる。

- (3) 最後に、(1)で求めた知識的な確率と、(2)で求めた音響的な確率を掛け合わせて、ニュース記事を分類する。

6.2 記事分類の方法

キーワード列をもとに記事を分類する場合、まず各キーワードと記事の分類分野(トピック)との関係を求めておく必要がある。トピック n に対するキーワード w の寄与率 $P(n|w)$ を(2)式で求める。

$$P(n|w) = \frac{\text{トピック } n \text{ に含まれる単語 } w \text{ の数}}{\text{全トピックに含まれる単語 } w \text{ の総数}} \quad (2)$$

これは、キーワード w を含むトピック n の事後確率に対応している。寄与率を求めるデータベースとしては、テロップと同じく「朝日新聞記事データベース分類表索引」を用いた。これには約12000のキーワードが載っており、キーワード毎に分類付けがなされている。この索引表を用いて、スポッティングで用いるすべてのキーワードについて、事前に(2)式で示される寄与率を求めておく。

ニュースキャストの音声から、キーワード列 $w_1, \dots, w_k$ が抽出されると、トピック n に対する事後確率 $P(n|w_1, \dots, w_k)$ は、(3)式で求める。

$$P(n|w_1, \dots, w_k) = \sum_{i=1, \dots, k} P(w_i) \times P(n|w_i) \quad (3)$$

このようにしてキーワードの音響的な確率 $P(w_i)$ と、キーワードの分類分野に関する知識的な確率 $P(n|w_i)$ を統合する。最後に、トピック n の事後確率 $P(n|w_1, \dots, w_k)$ が最大となるトピックを、その記事のトピックであると判定し分類する。

7. スポーツ映像の分類

スポーツニュースのような映像では、カメラ位置の制約や演出意図などの点から、画像内に映る対象の大きさや位置には制約がある。したがって、特定のスポーツを特徴づける代表的な画像が存在し、これらの画像を使ってシーンを分類することが可能である。

そこで、画像上の物理的なオブジェクトを明確に認識することなく、画像の物理的特徴量である輝度値や空間周波数を利用することによって、シーンの構図に関する全体的な特徴をもとにニュース記事の分類を行う⁷⁾。この手法により、スポーツニュース映像を自動的に分類・蓄積し、利用者の見たいシーンだけを再生することのできるビデオオンデマンドシステムを開発できるであろう。

1つのスポーツは、1つの代表的な画像だけで表すことはできない。そこで、カテゴリーをさらにクラスタに分類する多重部分空間法を用いている¹⁴⁾。

射影行列の計算方法などは部分空間法と同様である。部分空間法との違いは、もとの1つの部分空間を複数の小さな部分空間に分割する点にある。各カテゴリーの学習データにベクトル量子化¹⁵⁾を施して、複数のクラスタを構成し、各クラスタ毎に部分空間を作り多重部分空間を構成する。分類の際には、入力パターンをすべての部分空間に射影し、射影量の最も大きい部分空間を含むカテゴリーに入力パターンを分類している。

8. 顔認識と話者認識

映像処理としては、登場人物を索引付けして検索するために、顔の向きや大きさに関わらず、顔を切り出して追跡し、認識することが必要となる。映像中の人物の顔写真を、あらかじめ180°の方位からとっておくことはできないので、映像中から顔データを収集し、人物の顔モデルを自動構築しなければならない。このためには、映像中から顔を自動抽出して追跡し、すでに知っている人物と認識できれば追跡を継続し、人物名を索引として映像に張り付ける⁵⁾。もし人物が認識できなければ、新たにその人物の顔モデルを自動構築する。この延長上には、動作などのイベントや、ハイライトシーンといったイベント抽出の研究がある。

音声処理としては、発話者を索引付けして検索するために、話者照合の技術を用いて、ニュース音声からアナウンサの発話区間(話者区間)を切り出す処理や、座談会において参加者の発話区間を切り出す処理が必要になる⁶⁾。ニュース音声では、アナウンサの発話区間を自動抽出することにより、ニュース記事の切り出しと要約が可能となる。また、座談会では、参加者の発話区間を自動的に切り出すことにより、

各参加者の意見をデータベース化することが可能となる。

9. むすび

マルチメディアのデジタル化がもたらすものは、コンテンツの大量生産であり、我々は有限の時間のうちに必要なものを検索するすべを持たなければならない。この点から、本稿では、マルチメディアコンテンツのデジタル化と、大量のコンテンツの中から必要なものを取り出すコンテンツの自動解析やインデキシング、分類方法について筆者の研究を交えて述べた。

(1998年10月20日受付)

〔文 献〕

- 1) J. Kamahara, S. Shimojo, A. Sugano, T. Kaneda, H. Miyahara, T. Yahata and S. Nishio: "A News on Demand System with Automatic Program Composition and QoS Control Mechanism", International Journal of Information Technology, 2, 1, pp. 1-22 (1996)
- 2) 有木康雄: "DCT 特徴のクラスタリングに基づくニュース映像のカット検出と記事切り出し", 信学論, J80-D-II, 9, pp. 2421-2427 (1997)
- 3) 有木康雄: "音声単語スポットティングに基づくTVニュース記事の自動分類", 信学論, J82-D-II, 1 (1998)
- 4) 茂木祐治, 有木康雄: "ニュース映像中の文字認識に基づく記事の分類", 信学技報, PRU 95-240, pp. 33-40 (1996)
- 5) 杉山善明, 有木康雄: "部分空間射影による顔領域の追跡と学習", 信学技報, PRMU 97-162, pp. 77-82 (1997)

- 6) 西田昌史, 有木康雄: "自動学習による話者セグメンテーション", 信学技報, SP 97-57, pp. 1-6 (1997)
- 7) 杉山善明, 有木康雄: "多重部分空間法に基づくTVスポーツ・ニュース映像の自動分類", 信学論, J81-D-II, 9, pp. 2112-2119 (1998)
- 8) 石川則之, 有木康雄: "顔画像の指示による情報検索", 画電学研資, 97-04-05, pp. 17-20 (1997)
- 9) 田頭茂明, 有木康雄: "ニュース音声中の語彙反復による情報検索一部分空間射影に基づく話者正規化の応用", 情処学音声言語情報処理研資, pp. 99-104 (1996)
- 10) 大津展之: "判別および最小2乗規準に基づく自動しきい値設定法", 信学論, J63-D, 4, pp. 349-356 (1980)
- 11) 鶴岡信治, 栗田昌徳, 原田智夫, 木村文隆, 三宅康二: "加重方向指数ヒストグラム法による手書き漢字・ひらがな認識", 信学論, J70-D, 7, pp. 1390-1397 (July 1987)
- 12) エルッキ・オヤ著(小川英光, 佐藤誠訳): "パターン認識と部分空間法", 産業図書(1986)
- 13) 朝日新聞社ニューメディア本部: "朝日新聞記事データベース分類表索引", 朝日新聞社(1992)
- 14) 茂木祐治, 有木康雄: "手書き漢字認識のための多重部分空間法について", 平6信学総全大, D-304, p. 312 (1994)
- 15) Y. Linde, A. Buzo, and R. M. Gray: "An Algorithm for Vector Quantizer Design", IEEE Trans. Commun., COM-28, pp. 84-95 (1980)

ありき やすお
有木 康雄

1979年, 京都大学大学院工学研究科情報工学専攻博士課程修了。1980年, 京都大学工学部助手, 1990年, 龍谷大学理工学部電子情報学科助教授, 1992年, 教授となり, 現在に至る。1987年~1990年, エディンバラ大学客員研究員。映像処理・認識, 音声処理・認識に従事。工学博士。