

ニュース映像における人物の分離による背景の場面推定

井手 一郎<sup>†</sup> 浜田 玲子<sup>††</sup> 坂井 修一<sup>††</sup> 田中 英彦<sup>††</sup>

Background scene identification by character masking in news video

Ichiro IDE<sup>†</sup>, Reiko HAMADA<sup>††</sup>, Shuichi SAKAI<sup>††</sup>, and Hidehiko TANAKA<sup>††</sup>

あらまし 映像資源の効率的な再利用や検索への要請が高まるなか、ニュース映像へのテキストによる自動索引付けに関する研究が活発に行われている。筆者らは、映像中の音声やテキストを取捨選択せずにそのまま利用する既存の単純な手法よりも質の高い索引付けを行うために、画像内容との対応をいわゆる4W (Who, Where, When, What) に分類して考慮する機構を提案している。本稿では、その機構の実現のための画像内容解析部において必要となる場面推定手法の提案と、実際のニュース映像に適用した実験結果を紹介する。場面推定を行うにあたり、ニュース映像に人物像が大きく写る傾向が強いことを考慮し、人物領域を分離して背景領域のみの画像特徴量を利用する点が提案手法の特徴である。この結果、人物の存在自体も画像特徴量として加味した場合、人物領域の分離を行わない場合と比較して、7%から8%の場面推定性能の向上が見られ、提案手法の有効性が示された。また、評価のために事前に設定した正解場面別の評価結果から、各画像特徴量の有効性が場面により異なることが示され、適用場面に応じて特徴量を複合的に利用することにより、一段の性能向上の可能性が示唆された。

キーワード 映像データベース, 索引付け, 画像内容解析, 領域分割

1. まえがき

1.1 研究の背景

映像資源の効率的な再利用や検索への要請が高まるなか、米国 Carnegie Mellon 大学における Informedia プロジェクト [11] をはじめとして、映像データベースを構築するための自動索引付けに関する研究が内外で盛んに行われている。しかし、既存の一般的な自動索引付け手法は、映像中の音声やテキストから得られる情報を主に取捨選択せずに利用しており、画像内容との対応を考慮したものは、人物の顔と名前の対応を考慮した手法 [10] などに散見されるに過ぎない。

1.2 画像内容を考慮した索引付け

筆者らはこれらの要請と問題点を考慮して、映像中のテキストから得られる情報を利用しつつも、画像内容との対応を考慮した自動索引付け機構を提案している。索引付け対象としては、将来の再利用や検索とい

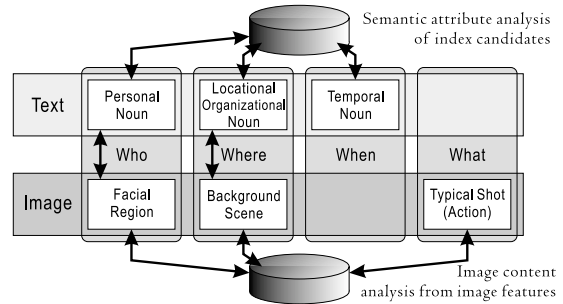


図1 映像内容との対応を考慮した索引付け機構の概要  
Fig.1 Concept of the indexing system considering correspondence with video content.

う点で利用価値が高く、具体的な画像内容に関する資料的価値が高いニュース映像を扱う。

具体的には、図1に示すように、画像及びテキストから得られる情報の各々について、いわゆる4W (「いつ (When)」、「どこで (Where)」、「誰が (Who)」、「何を (What)」) と呼ばれる属性毎の内容解析を行い、各々の対応に基づく索引付けを行うことを目指している。ニュース映像に対して、これらの属性による検索要求を想定することは自然であり、このような限

<sup>†</sup> 国立情報学研究所, 東京都  
National Institute of Informatics, 2-1-2 Hitotsubashi,  
Chiyoda-ku, Tokyo 101-8430, Japan

<sup>††</sup> 東京大学大学院工学系研究科, 東京都  
Graduate School of Engineering, The University of Tokyo,  
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan

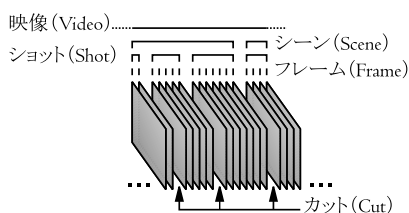


図2 映像の構成と用語の定義  
Fig.2 Hierarchical structure and term definition of video.

定は妥当であると考える。

本稿では、この機構のなかの画像内容解析部について取り上げ、ニュース映像中に人物像が頻出することを考慮し、人物領域分割による内容解析、特に背景領域からの場面推定手法の提案と、内容推定実験による手法の評価結果について報告する。

### 1.3 用語の定義

図2に映像の構成を示す。映像はフレームと呼ばれる静止画像の集合から構成され、画像的に連続なフレーム群をショットと呼ぶ。また、ショット間の不連続点をカットと呼ぶ。図1に示した機構ではショット単位での索引付けを目標としている。また、画像的・内容的に連続するショット群をシーンと呼び、ニュース映像において、内容的なシーンは話題に対応する。

## 2. 人物領域分離による画像内容解析

### 2.1 手法の概要

従来より、領域分割による画像内容の解析は行われてきたが、一般性の維持を考慮して、多少の伸縮はあるにせよ、ブロックへの定型的な分割が主流であった。しかし、ニュース映像とは、主に人間社会に生起する様々な事象に関する情報提供手段であるため、一般に人物像が大きく写ることが多い。そのため、ニュース画像の内容解析を行う際には、人物像の存在を考慮した処理や領域分割が重要となる。また、そのような性質から、人物という特定の事物の存在を前提とした処理を導入しても、一般性が大きく損なわれることはない。このような人物像に注目した領域分割により、従来の定型的なブロック分割による画像内容解析の際に排除しきれなかった、前景に出現する人物像による画像特徴量の変動の影響を排除し、その結果として、より正確な解析の実現を期待する。

人物像の存在、位置、人数などを基準とした、典型的ショット分類によるニュース映像の大雑把な内容推定手法も存在し [2], [6], 図1の機構でも「行為 (What)」

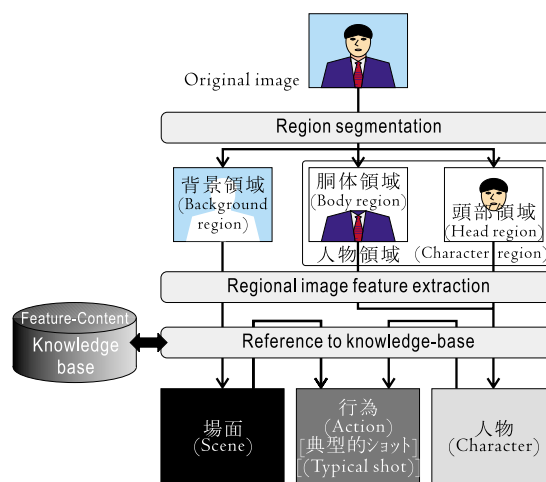


図3 人物領域分割による画像内容解析  
Fig.3 Image contents analysis by character region segmentation.

の推定に利用する。しかし、このような画像内容解析は、あくまでも画像全体としての解析であり、4Wのうちでも特に、「人物 (Who)」と「場面 (Where)」の解析を行うためには、画像中の人物領域 (人物像) と背景領域の各々を分離して解析する必要がある。また、画像全体の内容解析を行う際にも、「人物 (Who)」と「場面 (Where)」を個別に把握することにより、より正確な推定が可能になるとと思われる。

そこで本研究では、図3に示すように、顔領域の大きさや向きなどが一定条件を満たすような人物領域 (顔領域を含む頭部領域と胴体領域からなる) が画像中に存在する場合、人物領域とそれ以外の背景領域とに領域を分割し、領域毎に内容解析を行う。

このうち、「人物 (Who)」に関しては、顔認識に関する研究が盛んに行われているうえに、ニュース画像中の人物の顔領域を抽出し、テキストから得られる情報中の人物名と対応付ける研究も既に行われている [10]。また、「行為 (What)」に関しても、前述のように大雑把な推定手法は存在する [2], [6]。そこで、本稿では「場面 (Where)」に関する画像内容を考慮した索引付けを実現する際に必要となる画像処理部の機能として、人物領域を除去した背景領域の画像特徴に基づく場面推定手法の紹介と、実際のニュース画像を用いた推定実験による評価結果を示す。

### 2.2 人物領域と背景領域の分離

前節で述べたような画像内容の解析を行うためには、人物領域と背景領域を分離する必要がある。

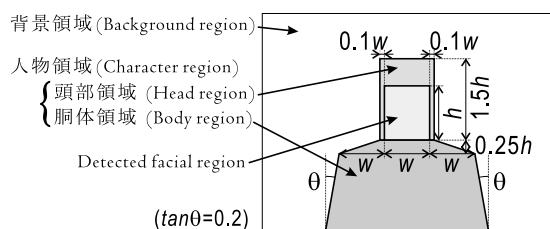


図4 顔領域を基準として人物領域を決定するテンプレート  
Fig. 4 Template for character region estimation from facial region.

そこで、ニュース映像中の人物は、(1) 良好な照明条件下で、(2) 正面から撮影されることが多い、という仮定に基づき、現在の技術でも比較的实现が容易な顔領域抽出を利用し、図4に示すような、顔領域を基準としたテンプレートにより人物領域を決定する簡便な手法を適用してみた。顔領域抽出には、Carnegie Mellon大学において開発されたニューラルネットワークによる訓練を用いたFace Detector [9]を利用した。

しかし、現実的には様々な状況下で撮影される画像において、上記の仮定（特に仮定(2)）は必ずしも満たされず、顔領域が正しく抽出されないことや、正面からの撮影を前提としたテンプレートの適用ができないことがあり、以下の実験では人手により人物領域を切り出したものを用いている。この際、人物領域の輪郭は、本来の輪郭から最大5ピクセル程度の逸脱を許して直線で近似した<sup>(注1)</sup>。

ただし、スタジオにおけるキャストの人物領域の分離に限り、実験に用いた全事例(231件)について顔領域抽出が過不足なく成功し、テンプレートの適用による領域分割も、人手による切り出しと同水準で行えたため、自動的に分離した結果を用いた。

### 3. 背景領域の画像特徴に基づく場面推定

以上のようにして前景に存在する人物領域を分離して得られる背景領域の画像内容、つまりどのような場面であるか、を推定する。そのためには、背景領域の画像特徴量と画像内容(場面)との関係に関する知識を用いる必要がある。一般の画像に対して、このような知識の獲得や記述は著しく困難である。しかし、ニュース映像においては、周期的あるいは集中的に特

(注1): この程度の逸脱を許しても、標準的な位置・大きさの顔領域に対して、最悪でもフレーム全体の1%程度のピクセルにしか影響せず、画像特徴量への影響は誤差の範囲内である。また、以下の実験においては、背景領域の画像特徴しか見ないため、人物領域の若干外側に輪郭をとるようにし、人物像が背景領域にくいこむことを防いだ。

定の事象に関連する話題を取り上げることが多いため、一定量以上の事例が集まれば、頻出場面に関しては、画像特徴量と場面との関係に関する知識ベースを構築することは現実的であると予想される。

本研究では、このような性質に基づく予想のもとに、背景領域の画像特徴量と場面の関係に関する事例に基づく場面推定を行った。このような単純な事例に基づく推定手法は、本稿における実験のように比較的小規模な事例群を想定したものである。従って、より大規模な事例群を対象とする際には、弁別能力、計算量の観点から、他の推定手法の導入が必要となるであろう。

#### 3.1 関連研究：画像特徴に基づく内容推定

画像特徴量と画像内容との関連付けを行う研究としては、初期のものとして、形容詞を中心とした印象語と画像特徴量の対応を心理実験から統計的に求める栗田らによる絵画データベースに関する研究[5]や、木本による感性語を用いた画像検索の試み[4]などがある。しかし、これらの感性工学における試みは、ニュース映像のように具体的事象(主に名詞)を対象とする場合とは問題点やその解決法が異なる。

従来は、詳細な物体のモデルに基づく具体的事物の認識や、対象分野を極めて限定した画像特徴量と画像内容の関係をを用いた内容推定が行われてきたに過ぎず、より一般的な内容推定を行う手法は少ない。そのようなものの一つとして、森らは、百科事典中の画像と説明文から特徴量と内容の関係を獲得する手法[8]を提案している。この手法では、まず一般的な語の共起関係により単語クラスタ空間を形成しておく。次に単語クラスタ空間中の単語間距離に基づき画像に付随する説明文間の類似度を計算し、それを反映させて構築した画像特徴量空間中で、説明文が類似した画像をクラスタ化する。これにより、問い合わせ画像に関連した説明文や語句の検索が可能になるが、百科事典が扱う対象があまりに一般的過ぎるため、良好な検索性能が発揮されない。

一方、孟らは適切にクラスタ化した訓練事例画像との画像特徴量の比較によるスポーツ映像を応用例としたジャンル推定手法[7]を提案している。この手法では、事前に訓練事例を分類して最適なクラスタ形成を行っておき、評価事例と各クラスタとの類似度に基づく分類を行う。手段や目的は類似しているが、画像全体に対する内容推定であり、本稿で提案する手法のように、画像の構成内容までは立ち入っていない。

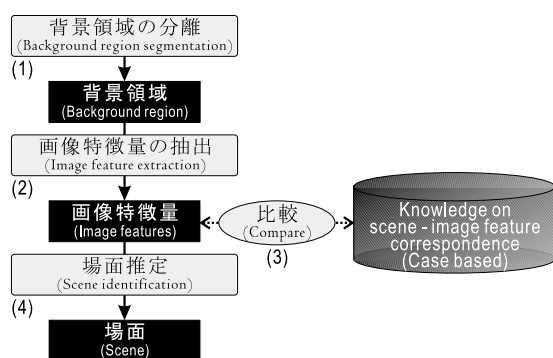


図5 場面推定の手順  
Fig. 5 Scene identification procedure.

### 3.2 場面推定の手順

図5に示す手順により、人物領域を分離した背景領域の画像特徴量に基づく場面推定を行った：

- (1) 背景領域の分離：  
一定の向き、大きさの人物像が存在する画像について、人物領域を切り出す。
- (2) 画像特徴量の抽出：  
切り出した人物領域を分離し、背景領域の画像特徴量を抽出する。人物像が存在しない場合は、画像全体を背景領域とする。
- (3) 類似度比較：  
推定を行う画像と知識ベース中の画像の間で、特徴量に基づき類似度を比較する。
- (4) 場面推定：  
類似度が上位に挙がった事例の場面に関する知識から推定を行う。

## 4. 場面推定実験

### 4.1 実験条件

以下に、場面推定実験を行った際の諸条件について述べる。

#### 4.1.1 用いた画像

実験には、2つの異なる時期に集中的に録画した15分の全国版ニュース映像20本中の一部、817ショットの先頭フレームの画像を用いた。場面の違いによる推定性能の差異を調べるために、いくつかの特定の頻出場面を中心に選び、訓練事例としての利用と評価の際の正解判定のため(のみ)に、事前に目視により正解を定めておいた。具体的には、

- (1) 閣議控室
- (2) 国会議場

表1 設定場面毎の事例数

Table 1 Numbers of images in the pre-classified scenes.

場面	人物あり	人物なし	合計
(1) 閣議控室	22	10	32
(2) 国会議場	31	21	52
(3) 報道会見場	11	6	17
(4) 法廷	6	23	29
(5) スタジオ	231	0	231
その他	124	332	456
合計	425	392	817

表2 画像キャプチャの際の条件

Table 2 Conditions of video capturing.

項目	条件
空間解像度	横 320×縦 240 ピクセル四方
色解像度	24bit (R,G,B 各色 8bit)
時間解像度	15 フレーム毎秒
フレーム圧縮	JPEG
ビデオ圧縮	なし(非圧縮)

- (3) 報道会見場
  - (4) 法廷
- の4種類の国内報道に関係する頻出場面と、
- (5) スタジオ

を設定した。表1に各場面毎の事例数を、人物像の有無に分けて示す。なお、画像中に人物が3人まで前向きに大きく写っている場合に「人物あり」とみなして領域分割の対象とし、それ以外は「人物なし」とした。

また、画像はアナログビデオを用いて録画したものを、ビデオキャプチャボードを用いて最低限のJPEG圧縮によりフレーム単位でデジタル化した。キャプチャの際の詳細な条件を表2に示す。

#### 4.1.2 画像特徴量

様々な画像特徴量が存在するなかで、本実験では、以下の2通りの色に関する統計的特徴量を用いた：

- (1) 色出現頻度分布(ヒストグラム)

画像中のピクセルの色の出現確率の分布。本実験では、RGB空間を線形に64分割した色ブロック毎の確率を求めたため、64次元ベクトルとして表現される。

- (2) 色共起頻度分布(コリログラム)[1]

画像中の一定距離離れたピクセル間の色の組合せの出現確率の分布。本実験では、RGB空間を線形に16分割した色ブロックの組合せ毎に距離1~4の確率を求めたため、(16×16×4=)1,024次元ベクトルとして表現される。

なお、RGB空間の色ブロックへの分割は以下の式に従って行った：

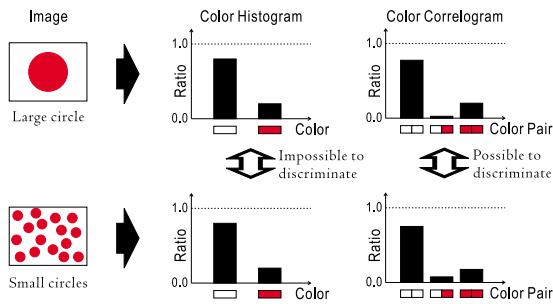


図6 色ヒストグラムと色コリログラムの差異  
Fig.6 Difference between color histogram and correlogram.

$$\text{色ブロック} \equiv \left[ \frac{R + G \times N + B \times N^2}{N^3} \times n \right] \quad (1)$$

ここで、 $0 \leq R, G, B \leq N - 1$ ,  $N = 256$  (各色 8bit),  $n = 16$  (色コリログラムの場合),  $n = 64$  (色ヒストグラムの場合)である。

両者には図6に示すような差異があり、ヒストグラムは画像全体のマクロな色彩の特徴を、コリログラムはミクロな特徴を反映する。これらの特徴量を採用した理由は、目視による場面の推定の際には、色彩的な特徴が大きな手がかりになると思われることと、両特徴量の性質の差異による推定性能への影響を見るためである。

以下の実験では、各々の特徴量を独立して用いて評価を行っているが、より多くの場面の導入や、場面毎の弁別能力の向上のためには、他の特徴量を含めて、統合的に用いる必要がある。

#### 4.1.3 類似度評価尺度

以上の2つの特徴量ベクトルの各々同士を比較する類似度評価尺度として、次式で求められるような、ベクトル  $\vec{F}_1, \vec{F}_2$  間の角度  $\theta$  の余弦 (値域: 0~1) を用いた:

$$\vec{F}_1 \text{ と } \vec{F}_2 \text{ の類似度} \equiv \cos \theta = \frac{\vec{F}_1 \cdot \vec{F}_2}{|\vec{F}_1| |\vec{F}_2|} \quad (2)$$

#### 4.1.4 推定性能評価尺度

場面推定性能の評価にあたり、以下に定義するような正答率を用いる。実際の場面推定の際には、最も類似している場面か、上位  $n$  位中の過半数を占める場면을推定結果とするため、この指標により場面推定性能が示される。

- 正答率 ( $n = 1, 3, 5, 7, 9$ ):

各設定場面に属す全評価事例のうち、知識ベース中の

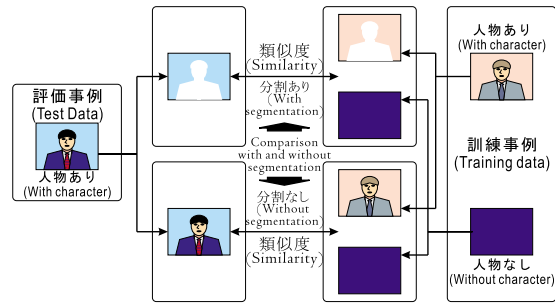


図7 人物領域分割の有無による類似度評価  
Fig.7 Similarity comparison with and without character region segmentation.

事例の類似度上位  $n$  位中に正答が過半数を占めたものの割合。結果的に  $n = 1$  の値は、最も類似している場面が正答であったものの占める割合を示す。

## 4.2 実験手順

基本的には、以上の実験条件のもとに、図5に示した手順に従って実験を行うが、より具体的な手順を以下に記す。

- (1) 背景領域分離:

「人物あり」の事例について、目視により人物領域を切り出す。スタジオのキャストのみ、Face Detector と図4のテンプレートによる切り出し結果を用いる。

- (2) 特徴量抽出:

背景領域の色ヒストグラムと色コリログラムの両者を求める。人物領域が存在しなければ、画像全体を背景領域とみなす。

- (3) 類似度評価:

評価事例と訓練事例 (全事例から評価事例を除いた事例の集合) の類似度の評価を行う。

- (4) 場面推定:

(3) を全事例に対して相互に適用し、設定場面別に正答率を集計する。つまり、知識ベース中の全  $N$  個の事例について、自らを評価事例とし、残りの  $N - 1$  個の事例を訓練事例とみなして推定した正答率を集計する。

なお、図7に示すように、(2) から (4) の処理において、人物領域分割の効果を見るために、「人物あり」の事例に対しても、人物領域の切り出しを行わずに画像全体での比較も併せて行った。

このようにして、以下の3通りの類似度評価条件で2通りの画像特徴量について実験を行った:

- 類似度評価:

- 人物領域分割を行わない場合 (従来手法; 結果のグラフ中で “w/o segmentation” と表記)

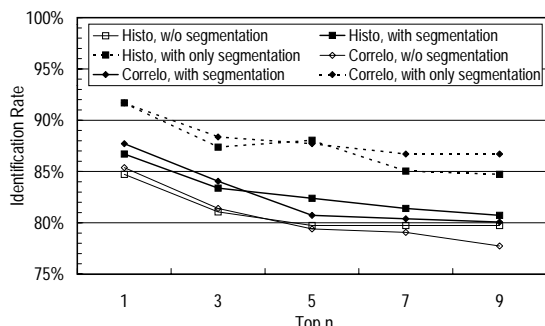


図8 場面推定結果 (総合評価; 正答率)  
Fig. 8 Overall scene identification rate.

— 人物領域分割を行い, 全訓練事例との類似度を評価した場合 (提案手法; 結果のグラフ中で “with segmentation” と表記)

— 人物領域分割を行い, 訓練事例中の人物領域が存在するもののみを評価事例・訓練事例として評価した場合 (提案手法; 結果のグラフ中で “with only segmentation” と表記)

● 画像特徴量:

- 色ヒストグラム
- 色コリログラム

4.3 実験結果と考察

図8は, 評価のために設定した表1中の場面(1)から(5)の推定結果を総合して評価したグラフである。この結果から, 以下のことが言える:

●  $n$  が小さいときに, コリログラムはヒストグラムよりも最大で1%程度良い結果を示し,  $n$  が大きくなるにつれ, 最大で2%程度逆の傾向が見られる。これは, コリログラムは非常に類似した画像との類似度が高くなるものの, わずかな変動に敏感であり, ヒストグラムではその逆の性質があることを示している。これらの性質は, 4.1.2で述べた両者のマイクロ・マクロな性質から説明できる。

● 提案手法である人物領域分離により, 2%から3%の性能向上がみられる。さらに, 訓練事例中の人物領域が元来存在しないものを除いた類似度評価の結果, 7%から8%の性能向上が見られた。後者の結果から, 人物領域の存在そのものも画像特徴量として考慮することが重要であることが示唆された。

次に, 画像特徴量として色ヒストグラムあるいは色コリログラムを用いた場合,  $n$  を1あるいは3に設定して評価した場合における, (1) 閣議控室 (cabinet), (2)

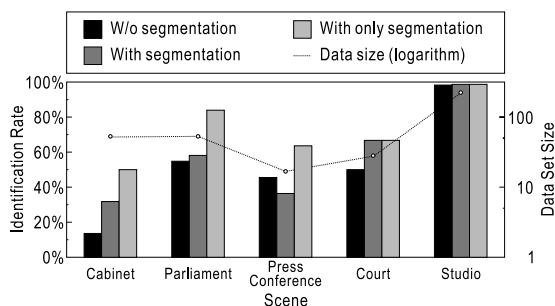


図9 場面推定結果 (正答率 ( $n=1$ ), ヒストグラム)  
Fig. 9 Scene identification rate ( $n=1$ , histogram).

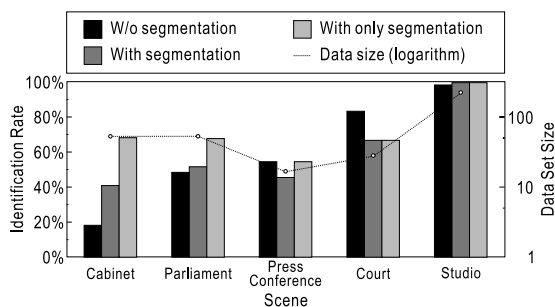


図10 場面推定結果 (正答率 ( $n=1$ ), コリログラム)  
Fig. 10 Scene identification rate ( $n=1$ , correlogram).

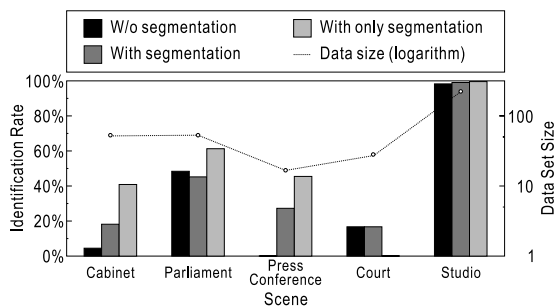


図11 場面推定結果 (正答率 ( $n=3$ ), ヒストグラム)  
Fig. 11 Scene identification rate ( $n=3$ , histogram).

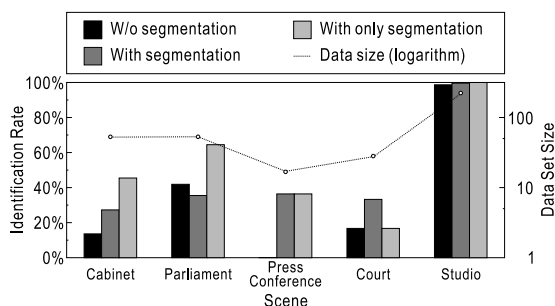


図12 場面推定結果 (正答率 ( $n=3$ ), コリログラム)  
Fig. 12 Scene identification rate ( $n=3$ , correlogram).

国会議場 (parliament), (3) 報道会見場 (press conference), (4) 法廷 (court), (5) スタジオ (studio), の各設定場面別の推定性能を図 9 から図 12 に示す。また, グラフ中には, 各設定場面毎の事例数を線グラフで対数表示してある。これらの結果から以下のことが言える:

- 総合評価で述べたように, 提案手法の有効性が基本的には示された。しかし, 場面によって有効性に差異が見られる。

- 色ヒストグラムと色コリログラムにも場面によって有効性に差異が見られる。

これらから, 人物領域分離及び用いる画像特徴量の有効性は, 対象とする場面により異なることが分かる。ここで, 場面毎の特徴には, 撮影しているカメラに関する物理的制約 (設置できる場所・動かせる角度など) も含み, 頻出場面においても, このような制約や場面固有の様々な画像的特徴により, 画像的な典型度は無視できない範囲で一定の変動を示すと思われる。

- 場面推定性能はおおまかに事例数の対数に比例している。

上位  $n$  件の類似ショットの過半数を占める場面に基づいて推定を行ったため, この傾向は特に  $n = 3$  の場合 (図 11 及び図 12) に顕著である。訓練事例の規模が大きく多様であるほど, より多くの類似事例が存在する確率が高くなるはずであり, 小規模で多様性の低い事例群を用いて, 類似度上位  $n$  件の過半数を正しい場面で占めることは困難である。

以上の実験においては, 人物領域が正しく分離できた場合を想定したが, 実際には輪郭の正しい切り出しに失敗することも多い。失敗の原因は, (1) 顔領域の検出もれ, (2) 顔領域の誤検出, (3) 図 4 のテンプレートの適用の失敗, に大別できる。(1) は顔領域抽出手法の性能に依存するが, (1a) 正面顔でない場合, (1b) 顔の大きさが小さい場合, (1c) 照明が暗い場合, に失敗する傾向がみられる。このうち, (1b) 及び (1c) の場合は背景領域から分離しなくとも影響は少ないと思われるが, (1a) の場合は提案手法による推定が失敗する可能性が高い。次に, (2) の場合に誤って人物領域が切り出されても, 誤検出した顔領域の位置や大きさ次第で背景領域の画像特徴が十分とれれば問題ない。一方, (3) の場合は, 正面向きの姿勢でないときや俯き加減のときに生じるが, 図 4 のテンプレートが正面向きの姿勢であるため, 人物像がとり得るほぼ最大の

領域が切り出されることから, 人物領域に背景領域の一部が含まれてしまうものの, 背景領域の場面推定には大きく影響しないと考えられる。

## 5. むすび

本稿では, ニュース映像の自動索引付け機構の一部として, 背景領域から人物領域を分離して場面を推定する手法を紹介した。提案手法は, ニュース映像において, 特定の場面は頻出するという特徴を利用し, 事前に索引付けしておいた事例との類似度評価により推定を行うものである。

評価実験として, 実際のニュース映像中の 817 のショットに対して, 人物領域の分離が正しく行えたという前提の下にこの手法を適用したところ, 実用可能な水準の推定性能を示し, 領域分割のみでなく, 人物の存在自体も画像特徴量として併せて考慮することが重要であることが示唆された。なお, 実験において, 人物領域の切り出しは主に人手で行ったものの, 特定の良好な条件下では顔領域を基準とした単純なテンプレートの適用による自動化が可能であり, 将来のこの分野における要素技術の進展により, より一般的な自動化も可能になるものと思われる。

一方で, 事前に設定した場面別の評価において, 以下のように今後検討の余地が残される点が洗い出された:

- 設定場面毎に特徴をより良く表す画像特徴量が異なり, 複数の特徴量を重み付きで組み合わせて用いる必要がある。重みは訓練事例群から事前に得ておくことが考えられる。

- 場面推定性能が事例数の対数にほぼ比例する傾向にあることから, 様々な状況に対応し得るに十分な事例を確保するためには, 事例数を指数規模で増やす必要がある。

しかし同時に, 事例数を増やすことにより場面間の弁別能力の低下が懸念されるため, 前項に挙げたような重み付けを用いて様々な画像特徴量を導入する必要がある。

謝辞 Face Detector [9] の使用を快諾して下さった, Henry Rowley 博士に感謝する。

## 文 献

- [1] J. Huang, S. R. Kumar, and R. Zabih, "Image Indexing Using Color Correlograms," Proc. IEEE Conf. on Computer Vision and Pattern Recognition '97, pp.762-768, June 1997.
- [2] 井手一郎, 山本晃司, 浜田玲子, 田中英彦, "ショット分類

- に基づく映像への自動的索引付け手法,” 信学論 (D-II), Vol.J82-D-II, No.10, pp.1543-1551, Oct. 1999.
- [3] I. Ide, R. Hamada, S. Sakai, and H. Tanaka, “Semantic Analysis of Television News Captions Referring to Suffixes,” Proc. Fourth Intl. Workshop on Information Retrieval with Asian Languages (IRAL'99), pp.37-42, Nov. 1999.
- [4] 木本晴夫, “感性語による画像検索とその精度評価,” 情処学論, Vol.40, No.3, pp.886-898, March 1999.
- [5] 栗田多喜夫, 加藤俊一, 福田郁美, 板倉あゆみ, “印象語による絵画データベースの検索,” 情処学論, Vol.33, No.11, pp.1373-1383, Nov. 1992.
- [6] Y. Nakamura, and T. Kanade, “Semantic Analysis for Video Contents Extraction -Spotting by Association in News Video-,” Proc. Fifth Intl. Multimedia Conf. (ACM Multimedia'97), pp.393-402, Nov. 1997.
- [7] 孟 洋, 佐藤真一, 坂内正夫, “事例画像を用いたシーン分類による映像索引付け手法,” 第5回信学知能情報メディアシンポジウム論文集, pp.53-60, Dec. 1999.
- [8] 森 靖英, 高橋裕信, 岡 隆一: “画像と単語の空間配置データベースに基づく画像理解の試み,” 第4回信学知能情報メディアシンポジウム論文集, pp.127-132, Dec. 1998.
- [9] H. D. Rowley, S. Baluja, and T. Kanade, “Neural Network-Based Face Detection,” IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.20, No.1, pp.23-38, Jan. 1998.
- [10] S. Satoh, Y. Nakamura, and T. Kanade, “Name-It: Naming and Detecting Faces in News Videos,” IEEE Multimedia, Vol.6, No.1, pp.22-35, March 1999.
- [11] H. D. Wactler, A. G. Hauptmann, M. G. Christel, R. A. Houghton, and A. M. Olligschlaeger, “Complementary video and audio analysis for broadcast news archives,” Commun. ACM, Vol.43, No.2, pp.42-47, Feb. 2000.
- (平成2000年9月30日受付, 2001年1月31日再受付)

#### 井手 一郎 (正員)

平6東大・工・電子卒。平8同大学院工学系研究科情報工学専攻修士課程了。平12同研究科電気工学専攻博士課程了。博士(工学)。同年より国立情報学研究所助手。自然言語処理, 統合メディア処理に興味を持っている。平7第51回情報処理学会全国大会奨励賞受賞。人工知能学会, 情報処理学会, ACM各会員。

#### 浜田 玲子 (学生員)

平10東大・工・電子情報卒。平12同大学院工学系研究科電気工学専攻修士課程了。修士(工学)。現在同専攻博士課程在学中。日本学術振興会特別研究員。情報処理学会会員。

#### 坂井 修一 (正員)

昭56東大・理・情報科学卒。昭61同大学院工学系研究科情報工学専門課程了。工学博士。同年工業技術院電子技術総合研究所入所。この間平3~4米国マサチューセッツ工科大学招聘研究員, 平5~6RWC超並列アーキテクチャ研究室室長。平8~10筑波大学電子・情報工学系助教授。平10より東京大学大学院工学系研究科助教授。計算機システム一般, 特にアーキテクチャ, 並列処理, スケジューリング問題, マルチメディアなどの研究に従事。平2情報処理学会論文賞, 平3日本IBM科学賞, 平7市村学術賞, 平7ICCD Outstanding Paper Awardなど受賞。情報処理学会, IEEE, ACM各会員。

#### 田中 英彦 (正員)

昭40東大・工・電子工学卒。昭45同大学院工学系研究科博士課程了。工学博士。同年東京大学工学部講師。昭46同助教授。昭62より同教授, 現在同大学院工学系研究科教授。この間昭53~54ニューヨーク市立大学客員教授。計算機アーキテクチャ, 並列処理, 人工知能, メディア処理, 自然言語処理, 分散処理, CAD等の研究に興味を持っている。著書「非ノイマンコンピュータ」, 「情報通信システム」, 共著書「計算機アーキテクチャ」, 「VLSIコンピュータI, II」, 「ソフトウェア指向アーキテクチャ」, “New Generation Computing”編集長。情報処理学会, 人工知能学会, 日本ソフトウェア科学会, IEEE, ACM各会員。