

## 顔の部分領域別動き照合によるニュース映像からの同一場面検出

熊谷はるか<sup>†</sup> 道満 恵介<sup>††,†</sup> 井手 一郎<sup>†</sup> 出口 大輔<sup>†††,†</sup> 村瀬 洋<sup>†</sup>

<sup>†</sup> 名古屋大学大学院情報科学研究科 〒464-8601 愛知県名古屋市千種区不老町

<sup>††</sup> 中京大学工学部 〒470-0393 愛知県豊田市貝津町床立 101

<sup>†††</sup> 名古屋大学情報連携統括本部 〒464-8601 愛知県名古屋市千種区不老町

E-mail: <sup>†</sup>kumagaih@murase.m.is.nagoya-u.ac.jp, <sup>††</sup>kdomain@sist.chukyo-u.ac.jp, <sup>†</sup>ide@is.nagoya-u.ac.jp,  
<sup>†††</sup>ddeguchi@nagoya-u.jp, <sup>†</sup>murase@is.nagoya-u.ac.jp

**あらまし** 本稿では、同一場面、つまり、同一のイベントを異なる視点で撮影した同一場面をニュース映像から検出する手法の改良について報告する。従来、我々は、音声特徴と画像特徴を組み合わせた同一場面検出手法を提案してきた。しかし、音声特徴は話者と音声一致する同一場面しか検出できないため、画像特徴のさらなる精度向上が課題であった。具体的には、従来手法では、異なる顔部品に属する特徴点軌跡が誤って対応付き、誤検出が起こるといった問題があった。そこで顔領域を照合する際に、縦に3分割した部分領域同士を照合する改良方法を提案する。部分領域同士を照合することにより、各部分領域に含まれる顔部品同士を照合することができるため、照合精度の向上が期待できる。実際に放送されたニュース映像で同一場面検出実験を行い、提案手法の有効性を確認した。

**キーワード** 同一場面検出, 画像特徴, 部分領域, 顔部品, ニュース映像

## Scene Duplicate Detection from News Videos by Sub-Region-based Motion Matching of a Face

Haruka KUMAGAI<sup>†</sup>, Keisuke DOMAN<sup>††,†</sup>,

Ichiro IDE<sup>†</sup>, Daisuke DEGUCHI<sup>†††,†</sup>, and Hiroshi MURASE<sup>†</sup>

<sup>†</sup> Graduate School of Information Science, Nagoya University, Japan

<sup>††</sup> School of Engineering, Chukyo University, Japan

<sup>†††</sup> Information and Communications Headquarters, Nagoya University, Japan

E-mail: <sup>†</sup>kumagaih@murase.m.is.nagoya-u.ac.jp, <sup>††</sup>kdomain@sist.chukyo-u.ac.jp, <sup>†</sup>ide@is.nagoya-u.ac.jp,  
<sup>†††</sup>ddeguchi@nagoya-u.jp, <sup>†</sup>murase@is.nagoya-u.ac.jp

**Abstract** In this report, we propose a method to detect scene duplicates from news videos. Scene duplicates are video clips taking the same event from different perspectives. To detect scene duplicates, we have previously proposed a method that combines audio features and image features. However, detectable shots with audio features were restricted. So, the accuracy with image features was needed to be improved. However, in the previous method, there was a problem that feature point trajectories of different facial parts were erroneously matched. In order to deal with this problem, the proposed method divides the face region into three sub-regions, and performs the matching between them. In the way, the same facial parts contained in each partial region are expected to be matched correctly and the accuracy should improve. Scene duplicate detection experiments were conducted using actual news videos. From the results, we confirmed the effectiveness of the proposed method.

**Key words** Scene duplicate detection, image feature, sub-regions, facial parts, news video

### 1. はじめに

近年、記憶装置や通信技術の発達により、放送映像を大量に

蓄積できるようになった。放送映像の中でも、ニュース映像は実世界の出来事を記録しており、資料的観点から価値が高い。ニュース映像を資料として活用するためには、大量に蓄えた



図 1 同一場面の例

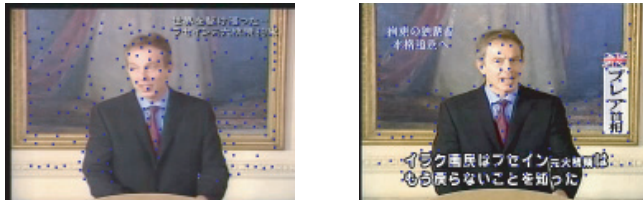


図 2 特徴抽出のための特徴点（青い点）が人物領域に少ない映像 [1]

ニュース映像を内容に基づいて扱う技術が必要となる。本研究では、そのような技術の中でも、「同一場面」をニュース映像から検出する技術に注目する。同一場面とは、図 1 に示すような、ある同一の時点・場所で起きたイベントを撮影した映像のことである。同一場面を利用すれば、異なるニュース番組で放映された、同一のイベントに関する多様な映像を、撮影視点によらず関連付け、様々な視点から視聴することができるようになる [2] [3]。よって、同一場面は、ニュース映像を内容に基づいて扱うための重要な手がかりとなる。

同一場面の検出を目的とした研究として、瀧本らは映像中のフラッシュを利用した検出手法を提案している [4]。しかしこの手法は、フラッシュによる特徴のみを用いるため、フラッシュが使われていない映像は検出できない。Wu らは、映像中の人物の特徴点の不均一性パターンに基づいた検出手法を提案している [1]。この特徴点の不均一性パターンは、人物の動きの緩急を表す特徴である。人物の動きの緩急は、撮影視点の違いに頑健であり、また、同じ撮影対象でも撮影時刻によって変化するため、撮影時刻を区別することができる。しかしこの手法は、図 2 に示すような、背景に文字や絵があるような場合（記者会見映像など）は検出が困難である。このような映像では、背景に特徴点が集まり、人物に特徴点が少ないため、人物の動きの緩急を正しく抽出できないためである。実際のニュース映像にはそのような状況が頻発するため、この手法をそのまま適用することはできない。また人物の動きの緩急は、ある程度の撮影視点の違いに頑健であるが、撮影視点が大きく異なる場合には検出が困難である。したがって、より高精度な検出のためには、撮影視点に影響されにくい特徴の利用が必要である。

以上の問題を解決するため、我々は音声特徴と画像特徴を併用する同一場面検出手法を提案した [5]。画像特徴による照合では、映像の各フレームを顔領域と背景領域に分割し、それぞれの領域同士を独立に照合し、その結果を統合した。顔領域の照合には既存手法と同様に特徴点の不均一性パターン [1] を利用し、背景領域の照合には色ヒストグラムを利用した。実験を行い、撮影視点異なる映像に対する有効性を確認した。しかし、音声特徴は話者と音声一致する同一場面しか検出できない



(a) 提案手法

(b) 従来手法 [5]

図 3 画像特徴による照合の概念

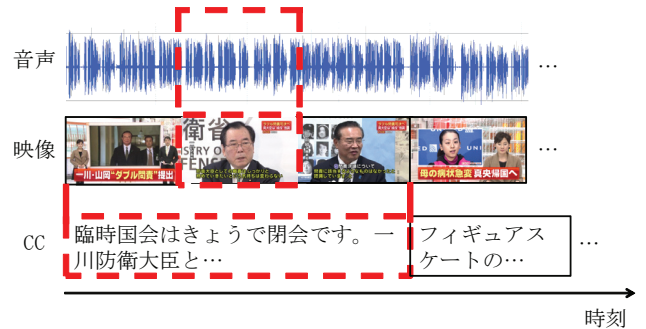


図 4 ニュース映像の一部と提案手法における入力データ（破線部分が入力データを表す）

め、画像特徴のさらなる精度向上が課題であった。具体的には、従来手法による顔領域の照合では、異なる顔部品に属する特徴点軌跡が誤って対応付き、誤検出が起こるといった問題と、マイクなどにより口唇領域が隠され正確に特徴を抽出することができず、誤検出が起こるといった問題があった。

そこで本稿では、図 3(a) に示すように、画像特徴の照合の際に、顔領域を縦に 3 分割し、部分領域間で照合するという同一場面検出手法の改良について報告する。従来手法は、図 3(b) に示すように顔領域全体間で照合した。これに対し、提案手法では部分領域間で照合することにより、各部分領域に含まれる顔部品間で照合することができる。また、照合の際にマイクなどに隠されやすい下部の領域を使用しないことで、より正確な照合を期待する。

以降、2. で提案手法について述べた後、3. で評価実験について述べ、その結果について考察する。最後に 4. で本報告をまとめる。

## 2. 提案手法

提案手法では、ショット、音声、CC (Closed Caption; 文字放送字幕) を 1 組の入力データとし、入力データ間の照合を行う。まず、入力データについて説明する。図 4 は、ニュース映像の一部である。横軸は時刻を、映像の各サムネイル画像はショットを表す。この図に示すように、入力データとして、映像、音声、CC のそれぞれ一部分を用いる。それぞれ破線で囲まれた部分が入力データである。映像は、入力としてショットを使用する。音声は、入力としてショットに対応する区間を使

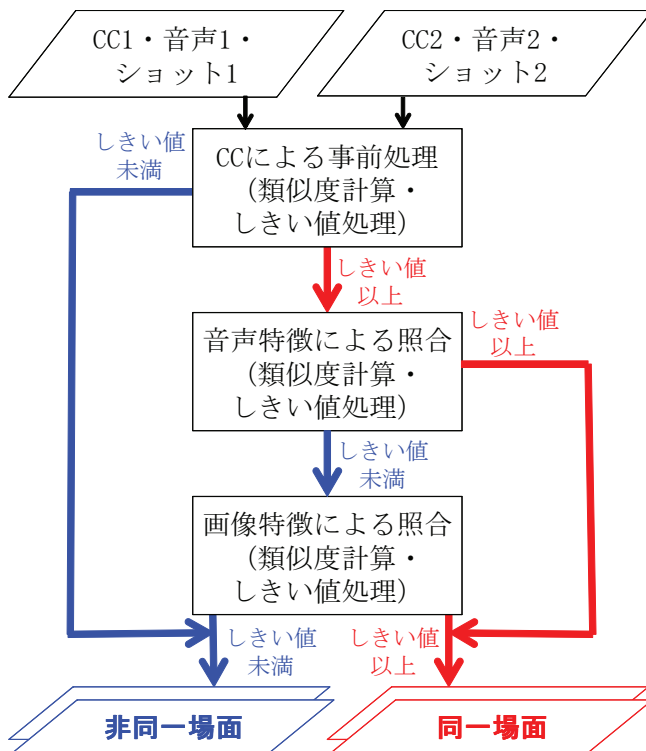


図5 提案手法の処理の流れ

用する。また CC は、入力としてショットが含まれるストーリーに対応するテキストを使用する。ここでストーリーとは、1つのイベントを扱った、ニュース映像の意味的な最小単位である。

次に、提案手法の処理の流れを図5に示す。まず、CCによる事前処理を行う。入力 CC 間の類似度がしきい値未満の場合は、その段階で入力データ対を同一場面でないとして判定し、以降の処理から除外する。次に、音声特徴による照合を行い、類似度がしきい値以上の入力データ対は同一場面と判定し、以降の処理から除外する。最後に、画像特徴による照合を行い、類似度をしきい値処理し、入力データ対が同一場面か否かを判定する。以降、各処理について順に詳述する。

### 2.1 CCによる事前処理

CCによる事前処理は、類似度の計算としきい値処理からなる。類似度の計算では、まず CC を形態素解析し、品詞判別を行う。そして、得られた形態素の品詞情報を用いてキーワードとなる語を抽出する。ここでは、形態素解析に日本語形態素解析システム JUMAN [6] を使用し、キーワードは形式名詞を除く名詞と未定語とした。そして、得られたキーワードとその出現頻度を特徴ベクトルとし、その余弦距離を計算し、類似度とする。

同一場面は同一時刻の映像であるため、その映像を含むストーリーの CC も類似することが期待される。そのため、計算された類似度がしきい値  $\theta_{cc}$  未満の入力データ対は同一場面でないとして判定し、以降の処理から除外する。なお、CCによる事前処理では、ショットに対応する CC だけでなく、ショットが含まれるストーリー全体の CC を用いる。その理由は、ショットに対応する CC には、類似度を安定して算出できないほど長さが短いものが存在すること、テロップが表示される際に CC が提

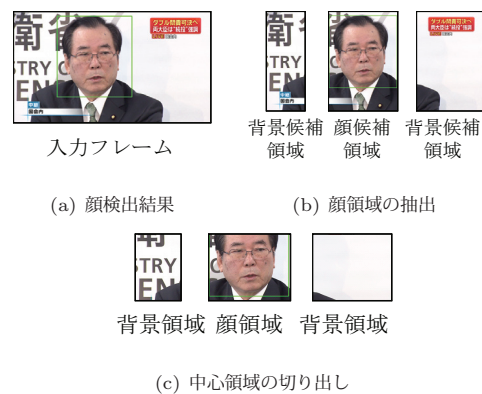


図6 提案手法における領域分割処理の流れ ((a),(b),(c)の順に行う)

供されないことがあるためである。これに対して、ストーリーは1つのイベントを扱った、ニュース映像の意味的な最小単位であるため、ある程度の長さになる。また、ストーリー中にアナウンサーによる発言が存在しないことは稀であり、アナウンサーが発言する際には CC が提供される。そのため、ストーリーに対応する CC を用いると、安定して類似度を計算できると考えた。

### 2.2 音声特徴による照合

音声特徴による照合は、入力音声に対する事前処理、音声間での類似度計算、類似度に対するしきい値処理からなる。事前処理では、まず、雑音を除去するために、入力音声から人の話し声を主に含むと考えられる周波数帯域 0.3k~4kHz [7] を取り出す。次に、処理時間を短縮するために、サンプリングレートを 8kHz に揃える。さらに、音声のボリュームの正規化を行う。最後に、極端に短い音声では同一場面か否かを判別できないため、入力音声対がある時間よりも短い場合、その音声を照合から除外する。

音声間の類似度として、入力音声の片方をサンプリングレートの逆数秒ずつずらしながら計算された音声波形の相互相関を利用する。最終的に、計算された相関の最大値を入力音声対の類似度とする。

映像中の話者と音声一致する同一場面は、撮影時刻が同じであるため、音声ほぼ一致する。そのため、音声特徴によって撮影視点に関わらず同一場面を検出できる。そこで、音声特徴の類似度がしきい値  $\theta_a$  以上の入力データ対は同一場面であると判定し、以降の処理から除外する。

### 2.3 画像特徴による照合

画像特徴による照合は、領域分割、領域別の照合、照合結果の統合、しきい値処理からなる。以降、各処理について説明する。

#### 2.3.1 領域分割

まず、図6(a)に示すように、入力ショットの各フレームから顔を検出する。顔検出には Viola らのオブジェクト検出手法 [8] などを用いることができる。次に、図6(b)に示すように、検出された顔の横幅に基づいて、入力ショットを横に3分割する。中央の領域を顔候補領域とし、左右の領域を背景候補領域とする。さらに図6(c)に示すように、顔候補領域と背景候補領域に分割したショットから、テロップが挿入される可能性が低い



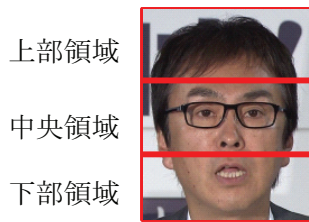


図 7 顔領域の部分領域への分割



図 8 顔領域の下部領域が隠されるショット対とその類似度

中心領域を切り出す。切り出された中心領域の中央部分を顔領域とし、左右を背景領域とする。また、背景領域はショット中で大きく変化することが少ないと仮定し、類似度計算には入力ショットの最初のフレームのみを用いる。

### 2.3.2 領域別の照合

領域別の照合は、人物領域の照合と、背景領域の照合からなる。以降、各処理について説明する。

#### a) 人物領域の照合

顔領域の照合は、 $W_u$  らによる映像中の人物の特徴点の不均一性パターンに基づく同一場面の検出手法 [1] を準用する。本研究では、図 7 に示すように、顔検出結果を縦に 3 分割した部分領域それぞれで顔領域の類似度を算出する。以降、これらの部分領域を上から上部領域、中央領域、下部領域と呼ぶ。人の顔部品の位置関係はおおよそ決まっているため、これらの部分領域間を照合することで、異なる顔部品間で動きの緩急が照合されることを防ぐ。ただし、図 8 に示すように、下部領域はマイクなどで隠されることがあるため、正確に人物の動きの緩急を抽出することが難しい。実際に、顔領域全体でこのショット対の類似度を計算すると 0.83 と高くなっており、誤検出の原因となっていることがわかる。そこで、上部領域と中央領域を使用し、下部領域は使用しない。また、 $W_u$  らの手法では、類似度の計算に特徴点軌跡の一部を用いているが、本手法では、部分領域間の類似度を計算する際、各部分領域で検出された特徴点軌跡をすべて使用する。ここでは、特徴点軌跡の検出を顔領域に限定するため、ほぼすべての特徴点軌跡が有効に働くと考えたためである。最後に、各部分領域間の類似度から顔領域の類似度を次式で計算する。 $S_f$  は入力ショット間の顔領域の類似度を表す。また、 $S_{f_2}$ 、 $S_{f_1}$  はそれぞれ、中央領域の類似度と上部領域の類似度を表す。

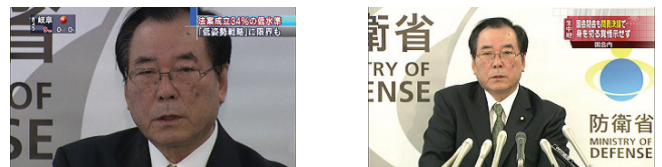


図 9 映像の明るさが異なる同一場面例

表 1 実験に用いた映像 (2012 年 6 月 7 日～2012 年 6 月 13 日)

放送局	番組名	放送時刻
NHK 総合	ニュース	12:00 ~ 12:20
東海テレビ	FNN スピーク	11:30 ~ 11:55
中京テレビ	NNN ストレイトニュース	11:30 ~ 11:55
CBC テレビ	JNN ニュース	11:30 ~ 11:55
メ〜テレ	ANN ニュース	11:45 ~ 12:00

$$S_f = \alpha S_{f_2} + (1 - \alpha) S_{f_1} \quad (1)$$

ここで、 $\alpha$  は 0 から 1 の値をとる重み係数である。

#### b) 背景領域の照合

背景領域の類似度として、色ヒストグラム間の距離を用いる。図 9 に示すように、ニュース映像は、録画状況や各放送局の規定によって同一場面でも映像の明るさが異なる場合がある。この明るさの違いによる検出漏れや誤検出を避けるために、色ヒストグラムには HSV 値の色相 (Hue) と彩度 (Saturation) を用いる。まず、領域分割されたショットの最初のフレームから色相と彩度を求め、2 次元ヒストグラムを計算する。次に、ヒストグラム間の Bhattacharyya 距離を求め、1 から減じることで背景領域の類似度  $S_b$  を計算する。

### 2.3.3 個別照合結果の統合としきい値処理

顔領域の類似度  $S_f$  と背景領域の類似度  $S_b$  から、入力ショット間の類似度  $S$  を次式で計算する。

$$S = \beta S_f + (1 - \beta) S_b \quad (2)$$

ここで、 $\beta$  は 0 から 1 の値をとる重み係数である。

そして、計算された類似度にしきい値処理を適用し、入力データ対が同一場面か否かを判定する。類似度がしきい値以上であった場合は、入力データ対を同一場面と判定し、しきい値未満であった場合は同一場面でないとして判定する。

## 3. 実験・考察

実際に放送されたニュース映像に本手法を適用することで、その有効性を評価した。

### 3.1 実験用映像

本実験では、2012 年 6 月 7 日から 2012 年 6 月 13 日に放送された複数の放送局のニュース映像を使用した。使用した放送局と番組、放送時刻を表 1 に示す。また、ニュース映像のフレームレートは 29.97 [frame/sec] であり、解像度は 1,440 × 1,080 [pixels] であった。これらの映像から、人物を中心に撮影されたショットを人手で計 678 ショット切り出した。全てのショットの組み合わせは  ${}_{678}C_2 = 229,503$  組であり、そのうち同一場面対は 840 組存在した。なお、同一場面の検出精度への影響を排

表 2 実験結果：領域分割の有無による同一場面検出精度の比較

手法	顔領域の 3 分割	AP	最大 F 値
提案	有	0.646	0.648
従来 [5]	無	0.418	0.438

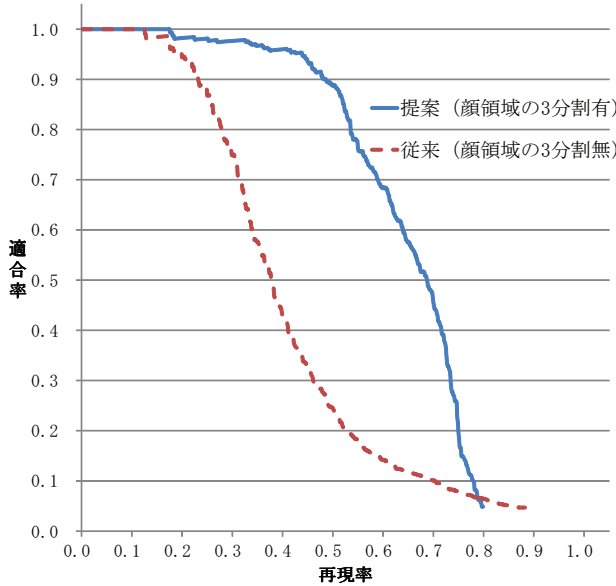


図 10 実験結果：領域分割の有無による同一場面検出実験の再現率—適合率グラフの比較

除するため、262 個のストーリーと顔領域は全て人手で抽出した。

### 3.2 実験方法

提案手法の有効性を評価するために、顔領域の照合の際に顔領域を縦に 3 分割し照合対象領域を制限する処理を行う提案手法と、それを行わない従来手法 [5] の同一場面検出精度を比較した。評価指標として、再現率と適合率、最大 F 値、Average Precision (AP) を使用した。なお、CC による事前処理で用いるしきい値と音声特徴による照合で用いるしきい値は全ての手法で同じ値を用いた。ここで、以降の再現率—適合率グラフは、類似度の順位が上位のものから一つずつ検出していく、それらが同一場面对であったときの適合率と再現率を用いて描いた。

### 3.3 実験結果

実験結果を表 2 に示す。提案手法の AP は 0.646、最大 F 値は 0.648 となり、両者とも提案手法の方が高かった。また、再現率と適合率は図 10 のようになり、大部分において従来手法よりも提案手法の精度が高かった。これらのことから、提案手法の有効性を確認できた。

### 3.4 考察

提案手法の精度が向上した理由は、顔領域を 3 分割したことで、異なる顔部品が照合されることを防げたためであると考えられる。顔領域を 3 分割しなかった場合と、顔領域を 3 分割した各領域のみを使用した場合の結果を表 3 と図 11 に示す。なお、他の特徴による影響を除くために、特徴は人物の動きの緩急のみを使用した。この結果からわかるように、顔領域を 3 分割したいずれの領域を使用した場合も、3 分割しなかった場合

表 3 実験結果：照合領域を変えた同一場面検出精度の比較

照合領域	AP	最大 F 値
顔領域全体 (従来 [5])	0.121	0.201
上部領域	0.328	0.396
中央領域	0.353	0.424
下部領域	0.318	0.390

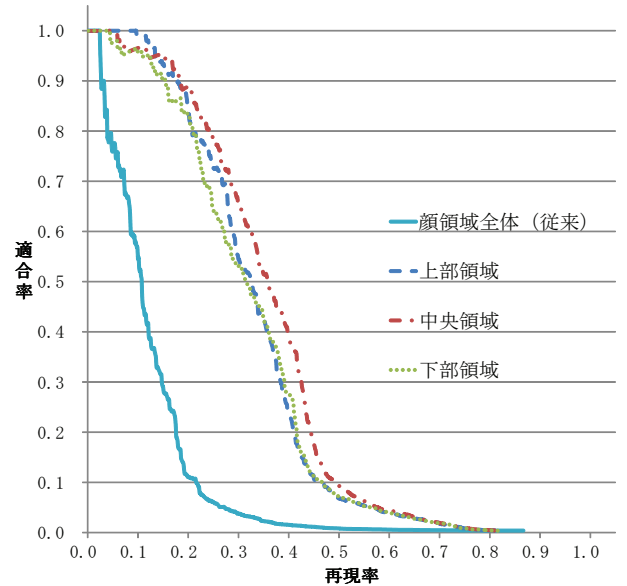


図 11 実験結果：使用領域を変えた同一場面検出実験の再現率—適合率グラフの比較

表 4 実験結果：部分領域の組み合わせを変えた同一場面検出精度の比較

部分領域の組み合わせ	最大 AP	最大 F 値
上部領域と中央領域	0.374	0.438
中央領域と下部領域	0.355	0.427
上部領域と下部領域	0.350	0.419
上部領域と中央領域と下部領域	0.375	0.440

よりも精度が高かった。このことから、顔領域を 3 分割し、部分領域間で照合を行うことの有効性を確認できた。

また、各部分領域の組み合わせにおける最大 AP を表 4 に示す。上部領域と中央領域を組み合わせた場合の最大 AP と最大 F 値は、表 3 の中央領域のみを照合した場合の精度よりも高かった。また、上部領域と中央領域に下部領域を組み合わせた場合の最大 AP と最大 F 値の上昇幅は、上部領域に中央領域を組み合わせた場合の精度の上昇幅に比べ小さかった。このことから、各部分領域の類似度を統合したときに、上部領域と中央領域の組み合わせが最適であることがわかる。下部領域を用いることに効果がないのは、図 12 に示すように、マイクなどで隠れることが多く、人物領域の動きの緩急が正確に抽出できないためだと考えられる。丸い点が特徴点を表し、矩形が部分領域を表す。図から、特徴点の一部がマイクに検出されていることがわかる。このような状況では、マイクの動きの緩急が検出されてしまい、人物領域の動きの緩急が正確に抽出できない。



図 12 下部領域がマイクで隠れるショット例 (丸い点が特徴点を, 矩形が部分領域を表す). 特徴点の一部がマイクに検出されている.



(a) 中央領域の場合: 検出された特徴点が多い  
(b) 上部領域の場合: 検出された特徴点が少ない

図 13 顔を正面から撮影したショットにおける各領域の特徴点検出結果の例 (丸い点が特徴点を, 矩形が領域を表す)



(a) 中央領域の場合: 目や鼻に検出された特徴点が集  
(b) 上部領域の場合: 耳にある程度の特徴点が存在

図 14 顔を横から撮影したショットにおける各領域の特徴点検出結果の例 (丸い点が特徴点を, 矩形が領域を表す)

また, 上部領域の類似度と中央領域の類似度を統合するのが最適であるのは, 各部分領域はそれぞれ安定して照合できる撮影視点が異なるためだと考えられる. これに関して, 中央領域は顔を正面から撮影した同一場面对の類似度が大きく, 上部領域は顔を横から撮影した同一場面对の類似度が高い傾向があった. まず, 顔を正面から撮影したショットについては, 図 13 に示すように中央領域は目や鼻を含むため特徴点検出が容易で, 多くの特徴点検出されている. 一方, 上部領域については顔を横から撮影したショットでは, 中央領域の方が人物の動きの緩急をより正確に抽出でき, その結果安定して照合できたと考えられる. 次に, 顔を横から撮影したショットについては, 図 14 に示すように中央領域は目や鼻に特徴点が集まっている. そのため, 人物が画面の奥を向くような動作をすると, 特徴点の追跡が難しい. 一方, 上部領域は耳にある程度の特徴点検出されており, 横顔において安定して特徴点追跡できる. よって, 顔を横から撮影したショットでは上部領域の方が人物の動きの緩急をより正確に抽出できる. その結果, 安定した照合を行うことができると考えられる. 以上のことから, 撮影視点によって安定して照合を行うことができる部分領域は異なるといえる. さらに, 顔を横から撮影したショットでは, 耳などの顔の側面の顔部品を含む部分領域を使用した方が精度が向上する可能性がある. 今後, 顔検出結果と連動して,

撮影視点の違いを考慮した領域分割の方法を検討する必要がある.

#### 4. むすび

本稿では, ニュース映像における顔の部分領域別動き照合による同一場面検出手法の改良について報告した. また, 提案手法の有効性を確認するために, 実際に放送されたニュース映像からの同一場面検出を行った. 提案手法では, 音声特徴と画像特徴により照合を行った. また, 画像特徴による照合では, 入力ショット対を顔領域と背景領域に分割し, それぞれの領域の照合結果を統合して類似度を計算した. 顔領域の照合では, 顔領域を 3 分割し部分領域間で照合することにより, 特徴点軌跡の誤対応を抑制した. また, マイクなどで隠されやすい下部領域を使用しないことで, より正確な照合を目指した.

提案手法の有効性を確認するために, 実際のニュース映像を用いた同一場面検出実験を行った. 顔領域の照合の際に顔領域を縦に 3 分割し照合対象領域を制限する処理を行う提案手法と, それを行わない従来手法の精度を比較した. その結果, 提案手法の最大 F 値と AP が最も高く, 提案手法の有効性を確認できた. 今後は, 顔向きの変化に頑健な顔検出の結果と連動した, 撮影視点の違いを考慮した領域分割の方法を検討する必要がある. また, 同一場面には意味的な関連があるため, 検出した同一場面を利用してニュース映像間を意味的に関連づけることができる. このような関連づけに関する応用についても, 今後検討する必要がある.

**謝辞** 本研究の一部は科学研究費補助金及び国立情報学研究所との共同研究による.

#### 文 献

- [1] X. Wu, M. Takimoto, and S. Satoh, "Scene duplicate detection based on the pattern of discontinuities in feature point trajectories," Proc. 16th ACM Int. Multimedia Conf., pp.51-60, Oct. 2008.
- [2] N. Katayama, H. Mo, and S. Satoh, "News shot cloud: Ranking TV news shots by cross TV-channel filtering for efficient browsing of large-scale news video archives," Proc. 17th Int. MultiMedia Modeling Conf., pp.284-295, Jan. 2011.
- [3] A. Ogawa, T. Takahashi, I. Ide, and H. Murase, "Cross-lingual retrieval of identical news events by near-duplicate video segment detection," Proc. 17th Int. MultiMedia Modeling Conf., pp.287-296, Jan. 2008.
- [4] 瀧本政雄, 佐藤真一, 坂内正夫, "大容量放送映像アーカイブからの同一フラッシュシーン映像の発見," 信学論 (D), vol.J89-D, no.12, pp.2699-2709, Dec. 2006.
- [5] 熊谷はるか, 道満恵介, 井手一郎, 出口大輔, 村瀬 洋, "音声特徴と画像特徴を利用したニュース映像からの同一場面検出," 信学技報, MVE2013-24, Sept. 2013.
- [6] 京都大学黒橋・河原研究室, "日本形態素解析システム JUMAN version 7.0," Jan. 2012, <http://nlp.ist.i.kyoto-u.ac.jp/index.php?JUMAN> [2014 年 1 月 20 日アクセス].
- [7] NTT 西日本, "高音質電話," <http://flets-w.com/hikaridenwa/service/kouon/> [2014 年 1 月 20 日アクセス].
- [8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proc. 2001 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, pp.511-518, Dec. 2001.