博 士 請 求 論 文

映像への自動索引付けに関する研究

— 統合メディア処理による索引付けとそのニュース映像への適用 —

指導教官 田中 英彦 教授

東京大学大学院 工学系研究科 電気工学専攻

学籍番号：67108 井 手 一 郎

平成 11 年 12 月 17 日 提出

内容梗概

近年の電気通信技術の飛躍的な発展に伴い，放送媒体の多様化や多チャンネル化が進行し，視聴者に提供される映像量は着実に増加している．このような増加を背景とし，コンテンツ不足を補うための再利用や，視聴者による番組選択のための検索を考慮して，映像を整理して蓄積する必要が急速に高まっている．殊に，ニュース番組は，その内容の重要性と利用価値の観点から，索引付けを行う価値が高い．しかし，索引付け作業の大半は，放送台本などをもとに人手に頼っているのが現状であり，付与単位も一般的には話題単位の大雑把なものでしかない．そのため，時々刻々増え続ける情報に対する速報性，そして，話題よりも細かい単位でのきめ細かな検索要求に応えるためにも，ニュース映像への索引付けの自動化が求められている．

このような要請を受け，近年，映像への自動索引付けに関する研究が盛んに行われている．その実現手段として，従来の画像認識による正攻法ではなく，いわゆる統合メディア処理，つまり様々なメディアを統合的に処理する手法によるものが 1990 年代中頃以降特に注目されるようになった．その背景としては，1980 年代以降の計算機による処理能力の飛躍的な向上に伴い，従来は計算量的に非現実的とみなされていた処理が可能になり，複数メディアを統合的に扱う余裕ができてきたことや，各メディア単体での処理に対する手法的限界が見えてきたことが挙げられるだろう．人間による外界の認識がいわゆる五感から総合的に行われることを考えても，複数メディアからの情報を統合的に利用することの利点は大きい．

これらの研究の多くはテキスト中の語の出現や，出現頻度や話題中の相対的な「珍しさ」を基準にした統計的な索引の抽出を行う単純な手法である．他方，テキストに対する全文検索的手法で検索を行い，そもそも索引付けを行わない手法すら存在する．これらの手法は比較的単純かつ高速に実現可能であり，その点においては実用的である．しかし，映像データベースとして重要であるはずの，索引と画像内容との具体的な対応を必ずしも機構的には保証していない．

このように，画像とテキストの統合処理を謳っている索引付け手法の多くは，画像にテキストから得られる情報を単純に付与しているに過ぎない．また，具体的な画像内容と索引との対応を考慮した手法であっても，人物の顔など，特定の事物のみを対象としたものが主流である．

本研究では，このような従来の索引付け手法の問題点をふまえて，日本語テレビニュー

ス映像において，映像中に含まれる画像情報とそれに付随するテキスト情報の双方を統合的に利用した統合メディア処理による索引付けの自動化を目指した．単純ではあるが実用的な従来手法とは異なる手法により，画像内容と具体的に対応した索引をあらかじめ付与しておく意義としては，主に以下の 2 点が挙げられる．

1. 膨大な量の映像に対する検索結果の絞り込みや要約が重要な課題となっているが，最初の検索結果そのものをより精選することができる．

2. 具体的事象に関する画像というよりは，映像中の話題そのものに興味がある一般利用者を対象としたデータベースにおいては，必ずしも具体的事物が検索結果の画像中に存在する必要はない．しかし，映像の再利用や編集を行う製作者を対象とする場合においては，効率的に具体的事象が存在する画像を検索できる．

また，このような映像内容を考慮した索引付けを実現するために必要ないくつかの要素技術，特に，以下の 2 点について重点的に述べる．

1. 索引候補のテキストとして利用する字幕の語義属性解析

2. ニュース映像における頻出場面・典型的構図の多さに着目した事例に基づく画像内容解析

1. の実現にあたり，「人物」，「場所・組織」，「時相」を示し得る接尾辞を収集した辞書を形態素解析済みテキストコーパスから作成し，辞書の性能評価をかねて，実際の字幕の語義解析を行った．その結果，全体的に見て，高再現率・低適合率の結果が得られた．これは，後段の統合処理時に雑音を除去できることを考えると，実用的な結果である．

2. については，まず明示的に分類条件を記述して表現した典型的ショットを複数設定し，実際のニュース映像の分類実験を行ったところ，高適合率・低再現率の結果が得られた．この理由として，顔領域の存在を分類条件に含む典型的ショットに対し，顔領域検出の困難さが悪影響を及ぼしていることが挙げられる．また，画像特徴量と画像内容の対応関係を用いた内容解析の実例として，ニュース映像に頻出する国内報道関連の場面の推定を行った．この際，ニュース映像に人物像が頻出することを考慮して，人物領域を除去した背景領域の画像特徴量を用いた場面推定を行ったところ，除去しなかった場合と比較して正確に行うことができた．

最後に，1. 及び 2. の結果を統合した最終的な索引付けにより，事例数の少なさにより実用的な水準の性能は示されなかったものの，比較的簡単な特徴量のみを利用したにもかかわらず，属性別の具体的事物に関する索引付けに成功した．

以上の結果，従来手法にはない映像内容を考慮した自動索引付け機構実現への道筋を示すとともに，必要となるいくつかの重要な要素技術の提案とその有効性を示した．

目 次

図 目 次

表 目 次

第 1 章

序論

1.1　研究の背景

近年の電気通信技術の飛躍的な発展に伴い，放送媒体の多様化や多チャンネル化が進行し，視聴者に提供される映像量は着実に増加している．

　このような増加を背景とし，コンテンツ不足を補うための再利用や，視聴者による番組選択のための検索を考慮して，映像を整理して蓄積する必要が急速に高まっている．殊に，ニュース番組は，その内容の重要性と利用価値の観点から，索引付けを行う価値が高い．しかし，索引付け作業の大半は，放送台本や取材メモを頼りに人手で行っているのが現状であり，付与単位も一般的には話題単位の大雑把なものでしかない．そのため，時々刻々増え続ける情報に対する速報性，そして，話題よりも細かい単位でのきめ細かな検索要求に応えるためにも，ニュース映像への索引付けの自動化が求められている．

　このような要請を受け，近年，映像への自動索引付けに関する研究が盛んに行われている．その実現手段として，従来の画像認識による正攻法ではなく，いわゆる統合メディア処理，つまり様々なメディアを統合的に処理する手法によるものが 1990 年代中頃以降特に注目されるようになった．その背景としては，1980 年代以降の計算機による処理能力の飛躍的な向上に伴い，従来は計算量的に非現実的とみなされていた処理が可能になり，複数メディアを統合的に扱う余裕ができたことや，各メディア単体での処理に対する手法的限界が見えてきたことが挙げられるだろう．人間による外界の認識がいわゆる五感から総合的に行われることを考えても，複数メディアからの情報を統合的に利用することの利点は大きい．

　これらの研究の多くはテキスト中の語の出現や，出現頻度や話題中の相対的な「珍しさ」を基準にした統計的な索引の抽出を行う単純な手法である．他方，何らかの手段で書き下したテキストに対する全文検索的手法で検索を行い，そもそも索引付けを行わない手法すら存在する．これらの手法は比較的単純に実現可能であり，その点においては実用的である．しかし，映像データベースとして重要であるはずの，索引と画像内容との具体的な対応を必ずしも機構的には保証していない．

1.2 研究の目的と意義

このように，画像とテキストの統合処理を謳っている索引付け手法の多くは，画像にテキストから得られる情報を単純に付与しているに過ぎない．また，具体的な画像内容と索引との対応を考慮した手法であっても，人物の顔など，特定の事物のみを対象としたものが主流である．

本研究では，このような従来の索引付け手法の問題点をふまえて，日本語テレビニュース映像において，映像中に含まれる画像情報とそれに付随するテキスト情報の双方を統合的に利用した統合メディア処理による索引付けの自動化を目指した．単純ではあるが実用的な従来手法とは異なる手法により，画像内容と具体的に対応した索引をあらかじめ付与しておく意義としては，主に以下の 2 点が挙げられる．

1. 膨大な量の映像に対する検索結果の絞り込みや要約が重要な課題となっており，様々な提示手法が検討されているが [102, 144, 152, 166]，最初の検索結果そのものをより精選することができる．

2. 具体的事象に関する画像というよりは，映像中の話題そのものに興味がある一般利用者を対象としたデータベースにおいては，必ずしも具体的事物が検索結果の画像中に存在する必要はない．しかし，映像の再利用や編集を行う製作者を対象とする場合においては，効率的に具体的事象が存在する画像を検索できる．

1.3 本論文の構成

本論文では，まず第 2 章において検索と索引付けを中心に映像データベース及び，索引付けの実現手法となる統合メディア処理に関して，各々基礎的事項や関連研究などについて述べる．

その後，第 3 章において提案手法，すなわち，映像内容を考慮した統合メディア処理による索引付けの概要を紹介し，続く第 4 章及び第 5 章において，その実現上必要となる，索引候補のテキストの語義属性解析及び画像特徴量に基づく画像内容解析について詳述し，実際のニュース映像への適用を通じて評価を行う．第 6 章においては，第 4 章及び第 5 章の処理の結果をうけた，実際の索引付けの結果を示す．

そして最後に，第 7 章において結論と今後の課題をまとめる．

第 2 章

映像データベースと統合メディア処理

本章では，まず一般的な映像の定義と構成について述べたあと，既存の映像検索手法の長所と短所を，問合せとデータ表現の形式で分類して紹介する．更に本研究で前提とする記号による画像への索引付けを，記号の種類に分けて紹介する．

2.1　映像の定義と構成

本節では，映像の定義を行った後に，一般的な構成を画像的観点及び内容的観点から述べる．

2.1.1　映像の定義

本論文では，図 2.1に示すように，画像と同期して放送される音声，文字放送などを全て含んだマルチメディアデータの集合体を映像と呼ぶ．

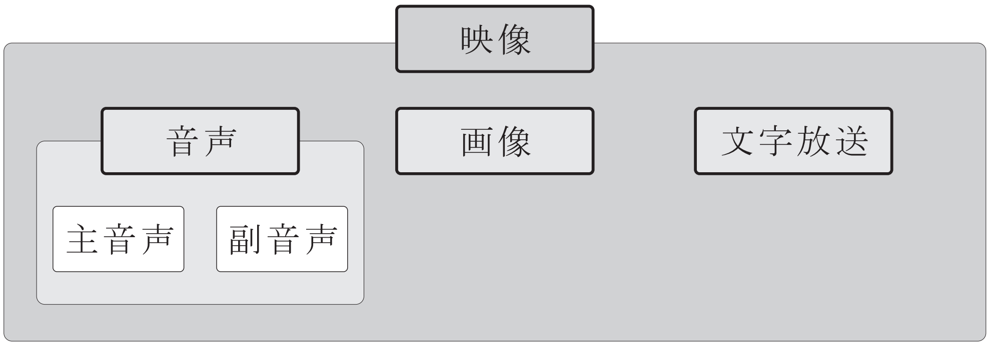


図 2.1: 映像の構成要素

2.1.2　映像の構成

ここでは，映像の一般的な構成を画像的観点及び内容的観点から述べる．

映像の画像的構成

一般に映像は，図 2.2に示すような画像的階層により構成されている．

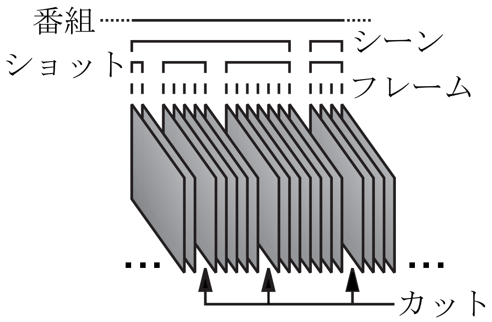


図 2.2: 映像の画像的構成

用語の定義　以下に，図 2.2中に示した用語の定義を行う．

- フレーム（Frame）
  映像を構成する最小単位の静止画像．日本でアナログ地上波放送に採用されている NTSC 放送方式のテレビ映像は，毎秒 30 枚のフレームから構成される．
- ショット（Shot）
  画像的に連続なフレームの集合．通常のテレビ映像中には，編集やカメラの切り替えなどの画像表現上の操作により，不連続点が存在する．このような不連続点をカット（Cut）と呼ぶ．本研究では，ショットを索引付け対象の最小単位とする．
- シーン（Scene）
  ショットの集合．画像的シーンと内容的シーンが存在する．
  - 画像的シーン（Graphical Scene）
    画像的に類似したショットの集合．
  - 内容的シーン（Semantic Scene）
    内容的に類似したショットの集合．通常，一つ以上の画像的シーンを包含する．

映像の内容的構成

映像の内容的構成は，その映像の属するジャンルにより異なった特徴を示す．いわゆるホームビデオ等を用いて個人で撮影した映像など，映像表現上あるいは編集上の自由度が高いものほどそのジャンルに共通した特徴を見出すことは難しい．しかし，専門家による撮影，編集，構成といった人的要因によるもののほか，撮影場所，機材，放送時間による制約なども加わり，多くのテレビ映像において比較的高い共通性が見られる [160]．

殊に本研究で主に扱うニュース映像においては，多くの情報を短時間に簡潔にかつ的確に伝えるという番組の性質上，多岐にわたる話題を扱いながらも，非常に画一的な構成が見られることが知られている [31, 133, 167]．図 2.3に，実際の 15 分間の全国版ニュース番組の内容的構成を分析した結果を示す．内容的シーンは一つ以上の画像的シーンからなり，話題は一つ以上の内容的シーンからなる．また，話題の先頭にはスタジオでキャスタがニュース内容を紹介するキャスタシーンがほぼ確実に存在し，話題境界を検出するための有力な手がかりとなる．

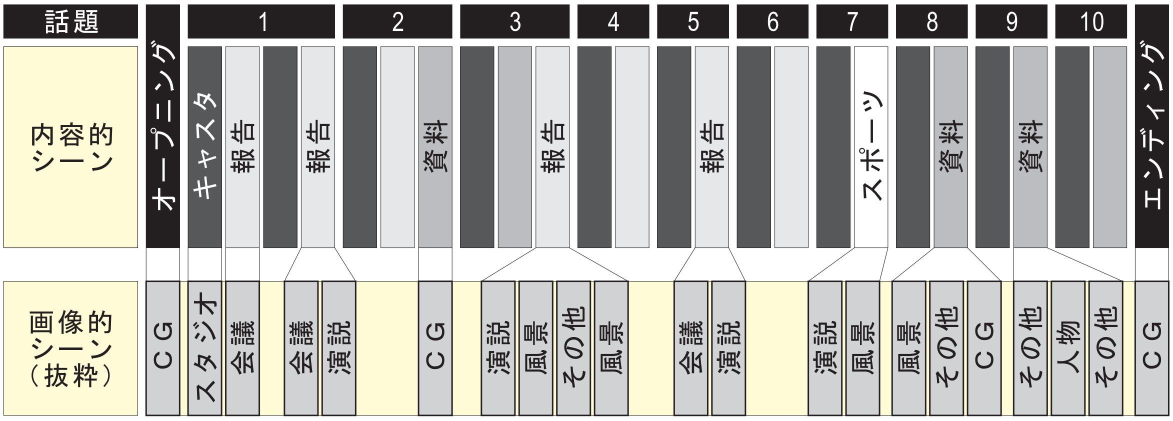


図 2.3: 映像の内容的構成

この例からも明らかなように，内容的構成は画像的特徴と深い関わりをもち，特に人間による人間のための放送である以上，画像中の人物に関する特徴量は内容的構成を把握するうえで非常に重要である．

2.2 映像データベースにおける検索と索引付け

2.2.1 画像検索

本節では，様々な画像検索手法を紹介する．具体的には，図 2.4に示すように，問合せとデータが各々画像であるか，あるいは抽象化された記号であるかに分け，各々による問合せとデータ表現の特徴について述べる．ここで記号とは，画像に対して何らかの抽象化をはかったもので，特徴を表すような数値やテキストによる表現のことである．

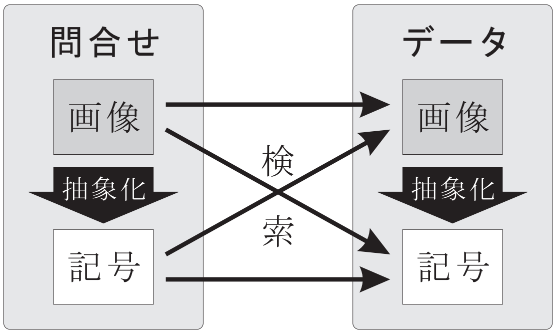


図 2.4: 問合せとデータの形式による検索手法の組合せ

以下に各々の特徴について述べる．

画像による問合せ

- 抽象的記号では表現しにくい問合せに適切
  感性的特徴や，複雑な形状などによる直接の問合せが可能だが，一方で抽象的記号で表現できる場合には入力作業が煩雑になる．
- 検索対象に関する具体的な画像的知識がある場合に有効
  検索対象事物の一部または全体の画像が既に検索者の手元にある場合 [124, 139] や，スケッチなどにより簡単に表現できる場合 [113] に有効である．

記号による問合せ

- 抽象的記号による問合せが可能
  複雑な形状など，具体的な画像的特徴を詳しく指定できない場合に有効である．
- 記号の抽象化の程度により表現の差を吸収した問合せが可能
  たとえばテキストによる問合せの場合，概念辞書や類義語辞書の利用により，問合せの表現の差を吸収可能である．

- 類似画像検索に不適
  検索対象事物の画像があらかじめ検索者の手元にある場合や，記号による表現が困難な場合には不適である．

画像によるデータ表現

- 前処理が不要
  記号による索引が不要なため，そのための前処理が不要である．一方で，データ毎に問合せ画像とのパターンマッチング的照合を行う必要があるため，照合に時間を要する．

- データの一般化が困難
  同一事物に関する個体差や視点の差による違いを吸収して一般化するためには膨大な量の 3 次元モデルをもつ必要があり，非現実的である．

記号によるデータ表現

- 照合時間が高速
  画像のパターンマッチング的照合と較べて，抽象化された記号による照合では，照合する情報量も少なく，また，ハッシュなどを用いた高速照合手法も存在する．一方で，抽象的記号による索引を生成するために，手動，自動を問わず，何らかの前処理が必要となる．

- データの一般化が可能
  記号化の段階で抽象化がはかられているので，同一事物や類似事物を一般化して表現することが可能

2.2.2 記号による画像への索引付け

本研究では，自然言語による具体的事物の検索を主眼におくため，2.2.1で取り上げた問合せ及びデータ形式のうち，記号による問合せ・データ表現を採用する．そこで，画像への索引付けの際に利用する記号の種類により分類して，各々の特徴と該当する既存の索引付け手法の紹介を行う．

画像特徴量による索引付け

ここでは，画像特徴量による索引付けの特徴と適用例について述べる．画像特徴量とは，画像から得られる色調や動きなどの具体的な事物の認識を伴わない特徴量のこととし，高度な認識による抽象化された特徴量は含めない．ただし，色共起頻度分布のように，抽象的ではあるものの，何らかの形で構造情報を保存した単純な特徴量 [116, 117, 141] も提案されている．

特徴

- 自動索引付けが容易
  索引付けに用いる特徴量の種類にもよるが，原始的な特徴量ならば，高速にかつ自動的に抽出可能である．

- 感性的検索との親和性
  感性的表現と画像特徴量とは比較的親和性が高く，相互の関係が判明すれば，感性的検索が実現可能である．

- 柔軟な類似画像検索が容易
  特徴量同士の類似性判定による「画像–画像」検索よりも柔軟な類似画像検索が可能である．

- 具体的事物の表現が困難
  画像特徴量のみで具体的事物を表現するのは困難である．

適用例　以上の特徴から，この種類の索引付けは，具体的事物による検索よりも感性的表現による検索 [126] や類似画像検索 [113, 135] など，より直感的な検索に適している．具体的には，絵画データベース [127]，商標や紋様のデータベース [124] への適用例が見られる．

音声特徴量による索引付け

ここでは，音声特徴量による索引付けの特徴と適用例について述べる．ここで音声特徴量とは，高度な認識処理を伴わずに，音声から得られる周波数スペクトルや声紋などの特徴量のことである．

特徴

- 音声認識を行わずに音声からの情報を利用可能
  音声認識の性能を気にすることなく，音声中に含まれる情報を直接利用する．

- 音声情報に固有の特徴を利用可能
  男声と女声の区別による話者の性別判断や，話者弁別が行える．

- 画像中の事物との対応が保証されない
  音声中の情報と実際の画像中の事物との対応は保証されていない．

適用例　以上の特徴から，この種類の索引付けは，特定のキーワードの音声レベルでの検出や，話者の特定などに適していると言える．具体的には，音声ストリーム中から歓声や音楽など特定の意味をもつ音声を検出して索引付けする手法 [107, 134] や，話者・性別特定による索引付けへの応用，キーワードスポッティングによる映像検索への適用例 [104, 149, 169, 181] が見られる．

テキストによる索引付け

ここでは，テキストによる索引付けの特徴と適用例について述べる．ここでテキストとは，画像認識や音声認識の結果を含め，数値による特徴量ではなく，何らかの「自然言語」で表現された概念のことである．

特徴

- 索引の高度な抽象化が可能
  自然言語という高度に抽象的な形式により表現できる．
- 具体的事物による検索の実現が可能
  具体的事物による問合せが可能であり，非常に実用的である．
- 抽象化時の視点の固定
  抽象化のための情報の取捨選択の際の視点に固定された索引になってしまい[1]，異なる視点からの検索ができない．

適用例　以上の特徴から，この種類の索引付けは，大量の画像から具体的事物を抽出するのに適しており，本手法をはじめ，自動索引付け手法の多くがこれを採用している．具体例については，2.4で紹介する．

[1] 索引付けにおける視点に関するモデルと傾向については，修士論文 [1] を，実際にこのような問題を考慮した索引付け手法については [150] を参照されたい．

2.3 統合メディア処理

本節では，複数のメディア処理を統合的に扱う，統合メディア処理を用いたシステムの具体例を，メディアの組合せ毎に紹介する．

2.3.1 統合メディア処理の定義と目的

統合メディア処理とは，複数の異なるメディアから得られる情報を統合的かつ相補的に用いる処理のことである．人間による外界の認識が，五感，すなわち，視覚，聴覚，嗅覚，味覚，触覚の五種類の感覚から総合的に行われることを考えても，解析対象の状況を理解するためには，単一メディアからの情報だけでなく，複数メディアからの情報を統合的に利用することの利点は大きい．

そこで，従来の単一メディアから得られる情報を処理するのみではなく，複数のメディアからの情報を統合的かつ相補的に処理することにより，解析対象の状況のより良い理解，及び単一メディア処理では困難な問題の回避による効率的な処理を目的とする．

ここでは，メディアとして，現在の計算機で一般的に扱われる以下の 3 通りのものを考え，図 2.5に統合メディア処理の概念図を示す．

- テキスト
- 画像
- 音声・音響

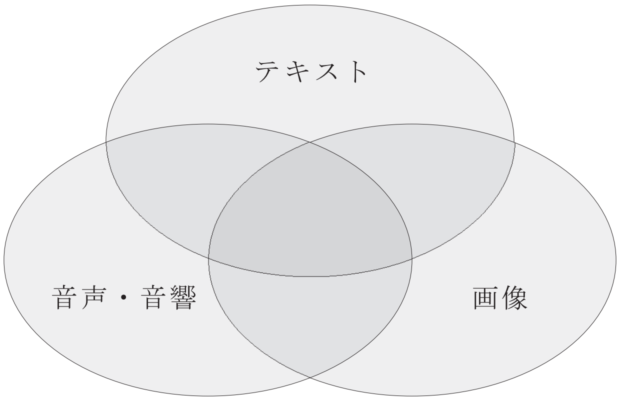


図 2.5: 統合メディア処理の種類

以下では，図中の各メディア処理の交差領域毎に具体的な適用例を紹介する．なお，統合メディア処理による映像への索引付けに関する研究の紹介は，本節では行わずに，2.4で行う．

2.3.2 テキスト–画像の統合処理

テキスト–画像の統合処理は，映像への自動索引付けにしばしば用いられる手法である．実際の自動索引付けへの適用例は，2.4で紹介することとし，ここでは映像以外への適用例を紹介する．

報道写真と説明文の対応付け

PICTION は新聞に掲載されている人物像を含む写真と，写真に付随して画像内容を描写するキャプションを統合的に利用して，写真中の人物像と人物名を関連付ける手法である [168]．具体的には，画像から顔領域候補と大きさを抽出し，キャプションから登場人物候補と相対的位置関係（左右の隣接関係など）に関する情報を抽出し，これら相互の様々な制約を充足させることにより，対応関係を得ている．この手法は統合メディア処理の分野の黎明期に登場したものであり，従来の画像処理のみによる人物認識の計算量的，精度的困難さに対する解決法として提案された．

一方，山田らは，写真ニュースに注目し，写真内容に関連する記事と写真中の人物像との対応付けを行った [183]．両メディア中に共通して存在する人物を対応付ける必要があるため，話題となっている人物の各メディアにおける出現形態に応じた重み付けを経た統合処理を行っている．

図面と解説文の対応付け

図鑑や教科書に掲載されている図面と解説文の対応付けを行う研究がある [143, 145, 178]．これらの手法では，解説図中の図形やテキストの位置関係，連接関係，包含関係などを解説文とともに利用して図面や概念の理解を試みている．しかし，部分的に自動化されていない困難な処理を含むほか，解説図面というやや特殊な画像の理解を目的としたものであるため，一般的な画像内容理解への応用は困難である．

2.3.3 画像–音声・音響の統合処理

映像中の特定の現象を検出するために，画像と音声・音響を統合的に利用する研究例を紹介する．

Chang らは，フットボールの試合中の歓声と画像中の白線やゴールポストの映り方を利用して，どのようなプレーが行われているかを推測している [107]．音響情報と画像情報を効果的に統合処理しているが，一般的な映像に対する内容解析への適用は困難である．

一方，杉山らは，ニュース映像を対象とし，話者認識により検出される同一話者区間と，様々な画像特徴から判定されるショット分類とを統合的に処理し，意味のある単位での再構成を行っている [169]．このような処理は，映像への索引付けを行うための前処理として重要である．

2.3.4 音声・音響–テキストの統合処理

他とは異なり，この組合せに関する研究は，音声認識の分野で，認識結果の修正などを目指して古くから行われている．

統合メディア処理の分野における研究例としては，映像と放送台本の同期をとるために，話者の性別と登場人物の対応をとる研究などが存在する [181, 182]．

2.3.5 テキスト–画像–音声・音響の統合処理

以上に紹介した研究は，主に 2 つのメディアを統合的に利用したものだが，真の統合メディア処理の実現には，テキスト，画像，音声・音響の 3 つのメディアを統合的に利用する必要がある．

このような対応付けを総合的に行ったものとして，柳沼らによるドラマ映像における画像，音声，台本の対応付けに関する研究 [181] がある．画像及び字幕などのテキストと，音声認識したテキストを対応付ける手法は数多く存在するが，このように実際に 3 つのメディアの各々から得られる特徴を利用した統合処理はあまり見られない．

2.4 映像への索引付け自動化における統合メディア処理

本節では，映像の内外に付随するテキストを利用した統合メディア処理による自動索引付けの代表例をいくつか紹介する．

2.4.1 映像中に存在するテキストの利用

ここでは，映像中に存在するテキストを利用した自動索引付け手法を紹介する．

クローズドキャプションの利用

クローズドキャプションを利用した手法は，同サービスの提供が先行している欧米で活発であるが，全文検索的手法によりそもそも索引付けを行わない手法 [105] も多い．ここでは，その中でも特筆すべき 2 つの研究例を紹介する．クローズドキャプションの含む内容は基本的に音声を書き下したものであるが，話者や話題の境界が特殊な表現により示されることが多く，そのような情報を利用した処理も行われている．

Informedia プロジェクト　まず，大規模な映像データベースとして，また映像への自動索引付け手法として，いずれにおいても先駆的なプロジェクトであり，一定の成果をおさめている Informedia プロジェクト内の News-on-Demand システムについて紹介する．

Informedia プロジェクトは，1994 年から米国科学財団（NSF）の支援のもとに Carnegie Mellon 大学（CMU）において行われているオンライン電子映像図書館プロジェクトである．Informedia では，データベースへの映像の蓄積及び索引付けを自動的に行い，ネットワークを介してディジタル化された映像の検索と提供を実現し，初等中等教育機関への提供を通じた利用者調査までを総合的に行っている [121, 173, 174, 202]．対象とする映像としては，当初は，Pennsylvania 州 Pittsburgh 市の地方テレビ局 WQED, Virginia 州 Fairfax 郡公立学校区, 英国のオープンユニバーシティの映像資産から提供を受けた 1,000 時間分の編集前後のドキュメンタリおよび教育用映像を蓄積し，同プロジェクト内の News-on-Demand プロジェクトでは，Cable News Network（CNN）社から提供を受けたニュース映像を蓄積している [175]．

News-on-Demand システムは，このような枠組の中でニュース番組を蓄積する電子映像図書館である．このシステムは放送された映像から，音声認識とクローズドキャプションの利用により得られるテキスト情報による索引付けを行い，検索サービスを提供するものである．

索引付けの際には TF-IDF 法による語の「珍しさ」に基づくキーワード抽出を行っている．ここで，TF-IDF とは，Term Frequency Inverse Document Frequency の略で，一般的なテキストコーパスを基準とした語の「珍しさ」の指標であり，次式のような関数型で定義される．

$$f_{tf-idf}(w) = \frac{f_{tf}(\text{ある文書（区間）中の単語 } w \text{の出現確率})}{f_{df}(\text{一般的な大規模テキストコーパス中の単語 } w \text{の出現確率})} \tag{2.1}$$

このような統計値に基づく評価関数による索引付けは，完全な自動化を実現する点においては有効であるものの，映像内容と索引との対応を保証していない [175]．

Name-It プロジェクト　次に，自動的に画像中の人物像とクローズドキャプションの解析から得られた人物名との対応付けを行う手法として，Name-It システムがある．このシステムは，画像中の人物の顔領域とクローズドキャプション中の人物名との対応付けを行い，画像，テキスト双方からの検索を可能にしている [158, 159]．

画像からの検索としては，特定の人物像の画像による問合せに対して，顔の特徴に基づく類似度判定により，画像中にその人物が存在すれば，対応する人物名を返す．一方，テキストからの検索としては，特定の人物名による問合せに対して，一致する人物名に対応する人物像を含む画像を返す．

このように，複数メディアの対応付けによる双方向性をもつ検索手法として，非常に優れているものの，対象が人物に限定されている点で，一般的なニュース映像の索引付け手法としては，汎用性に欠ける．

音声の利用

音声認識結果のテキストの利用は，基本的にクローズドキャプションの利用と同じ処理を経るが，音声独自の利用方法としては，ワードスポッティングが挙げられる．ワードスポッティングを索引付けに用いる際には，事前にキーワードとなり得る語を登録しておき，それらの出現を検出することになる．

このような研究として，有木らによる音声，字幕，画像を利用した索引付けと分類の研究がある [104]．この研究では，キャスタの発話を認識し，その中からキーワードをワードスポッティングにより抽出したものと，字幕やフリップ中の文字を認識したものから，話題の内容を分類する [138, 149]．

字幕の利用

字幕の出現を索引付けに利用した研究は数多く行われてきたが，必ずしも画像内容と対応しないことや，同じ映像中に出現する字幕でも，出現形態により内容の抽象度が異なるため，即座に索引として利用しにくい，などの問題があった．そこで，字幕の意味解析を行い，選択的に索引付けを行うことが試みられている．

渡辺らは，映像の分類や理解のために必要であるとして，言語的性質や出現位置を利用した字幕の意味解析を行った [180]．

一方，佐藤らは，字幕の文字認識結果の誤りをクローズドキャプション中の語により修正したうえで，両者に共起する語を索引として適切な語と判断して，字幕の選択的な索引付けを行っている [161]．

本研究もこのような試みの一つと位置付けられる．

ショット分類による索引付け

ニュース映像において，映像内容と索引の対応を保証するために，映像を典型的なショットに分類し，各典型的分類に適した索引付けを行う試みとして，ショット分類による索引付けが存在する [4, 142]．

これらの手法は，ニュース画像が多くの典型的特徴をもつショットからなることを前提とし，画像的に分類した典型的ショット分類に対して，共通する特徴に適した属性の索引付けを行うことにより，画像内容と索引との対応を保証している．

しかし，いずれの手法にしても，典型的ショット分類を明示的規則により記述する必要があるため，分類可能な種類には限界がある．本研究でも，第 5 章で述べるように，特定の映像内容解析にはショット分類を採用するが，一般には事例に基づくより一般的な内容解析を行う．

2.4.2 映像外に存在するテキストの利用

ここでは，映像外に存在するテキストを利用した自動索引付け手法を紹介する．

シナリオ文書と映像の対応付け

シナリオ文書は，映像外に存在するテキストとして，一般に入手可能なテキストではないが，製作者側における索引付けを考えると，十分実用的な手法である．具体的には，シナリオから得られる登場人物や動作に関する情報と，映像や音声から得られる情報を DP マッチングを用いて時系列に対応付けて，存在–行動マップを作成する手法である [120, 181, 182]．

しかし，これらの手法は事前に丁寧に作成されたシナリオが存在する必要があり，ドラマやドキュメンタリ番組など，そのようなシナリオに基づき，綿密に編集された映像への適用には適しているものの，ニュース番組などには不適である．

テキスト教材と映像の対応付け

映像外に存在するテキストとして，テキスト教材（副読本）がある．主に教養番組のために刊行されているものであり，その中には料理番組を始めとして，何かを製作する過程を解説したものがある．このようなテキスト教材と映像との対応付けを行う研究が存在する [16]．

このような製作過程は，シナリオのように時系列に直列に扱えるものではなく，順序関係の時間的な不可逆性や並列性に関する制約が存在し，それらを考慮した映像との対応付けが必要となる．

具体的には，テキスト教材中に存在する製作過程を解説した文章の構造解析により，手順の不可逆性や並列性を反映したフローグラフを作成し，各操作要素単位に画像や音声から得られる情報を参照した対応付けを行う．

この手法では，テキスト教材が必要なうえに，手順の解析に対象領域固有の辞書（料理番組ならば素材名や調理操作など）が必要となるため，様々な話題が登場するニュース映像への適用は不適である．

新聞記事とテレビニュース映像の対応付け

一方，直接ニュース映像と連動して製作されるわけではないが，実社会の出来事と連動して間接的に関連した内容のテキストを含む新聞記事との対応付けに関する研究がある [179]．この手法では，ニュース映像中の字幕と新聞記事の各々に共通して出現する名詞の出現位置などを考慮して話題・記事単位の類似度を求め，両者の対応を決定する．このようにして，同期がとれていない 2 つのメディアにおける類似した内容の情報を対応付けることにより，相互に不足している情報を補完し合うことが可能になる．

2.5 まとめ

本章では，まず，映像データベースへの検索を，問合せ及びデータ表現の形式が画像であるか記号であるかにより分類して，各手法の特徴を紹介した．その結果，本研究で扱うニュース映像に対しては，テキストによる問合せ及びデータ表現が適していると判断し，そのためにテキストによる索引付けを行うことにした．

次に，テキストによる索引付けを実現するために本研究で採用する統合メディア処理による様々な対応付けや索引付けの手法に関する研究例を網羅的に紹介した．

第 3 章

映像内容を考慮した統合メディア処理による索引付け

　本章では，提案手法，すなわち，映像内容を考慮した統合メディア処理による索引付けの全体像と各処理部 :(1) 自然言語処理部，(2) 画像処理部，(3) 統合処理部の概要について述べる．

3.1　提案手法の概要

本節では，提案手法の全体像について述べる．

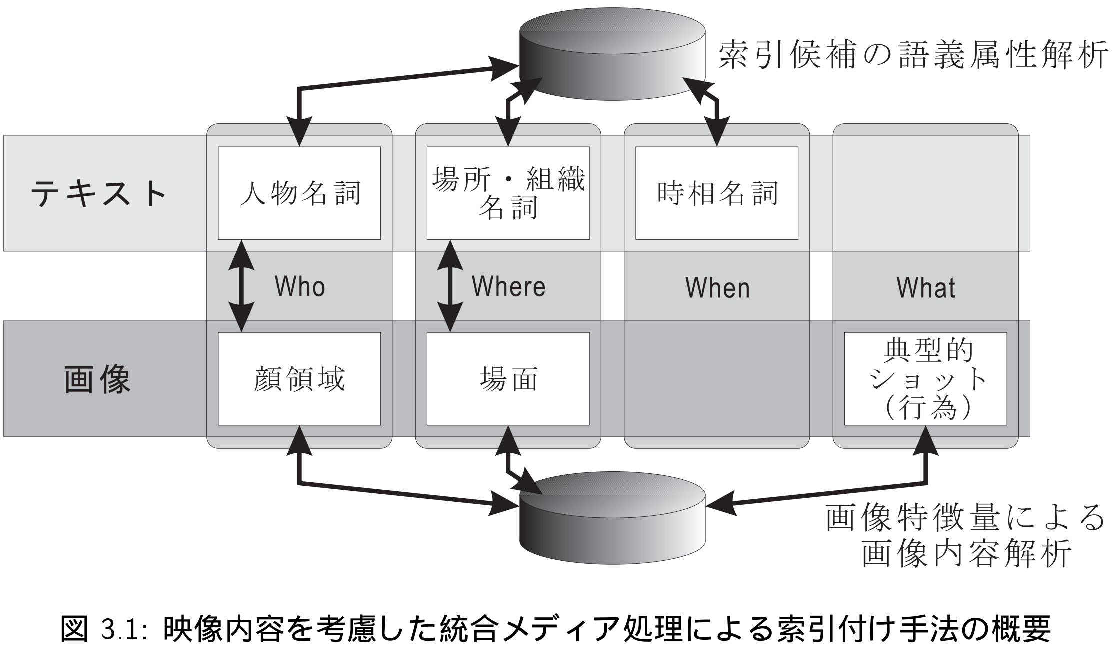


図 3.1: 映像内容を考慮した統合メディア処理による索引付け手法の概要

図 3.1に提案手法による索引付けの概要を示す．提案手法においては，映像に対して 4 つの異なる属性別の索引を付与することを目指した．4 つの属性とはすなわち（1）いつ（when）（2）どこで（where）（3）誰が（who）（4）何を（what）の，いわゆる 4W と呼ばれるものである．ニュース映像に対して，これらの属性に対する検索要求が多いと考えることは自然であり，これらに限定することは妥当であると考えられる．

付随するテキストからこのような属性を考慮して索引を抽出する手法そのものに新規性は少ないが，提案手法では画像内容との対応を考慮しつつ適切な索引を取捨選択する点で，既存手法と一線を画する．

特に（2）及び（3），更には場合によっては（4）に関する対応を考慮する際には，画像特徴と画像内容の関連に関する知識ベースが不可欠となる．この知識ベースの内容は，画像特徴と画像内容の関連に関する知識であれば，明示的に記述された規則であっても，事例に基づき統計的に獲得された関係であってもよい．

このようにして，画像中に実際に存在する事象を属性別に確認した索引付けの実現を目指した．

3.2 自然言語処理部

本節では，提案手法における自然言語処理の概要及び，映像中に存在する様々なテキスト情報の種類と性質について述べ，本研究で字幕を索引のための情報源として採用するに至った理由を述べる．

3.2.1 自然言語処理の概要

提案手法では（1）人物（誰が）（2）場所・組織（どこで）（3）時相（いつ）（4）行為（何を）に関する索引付けを行う．このうち（1）から（3）に関しては，図3.2に示すように，本研究独自の手法で索引候補の語義属性解析を行って分類する．なお（4）に関しては，テキストからの解析は困難であるため，画像処理の結果を利用する．

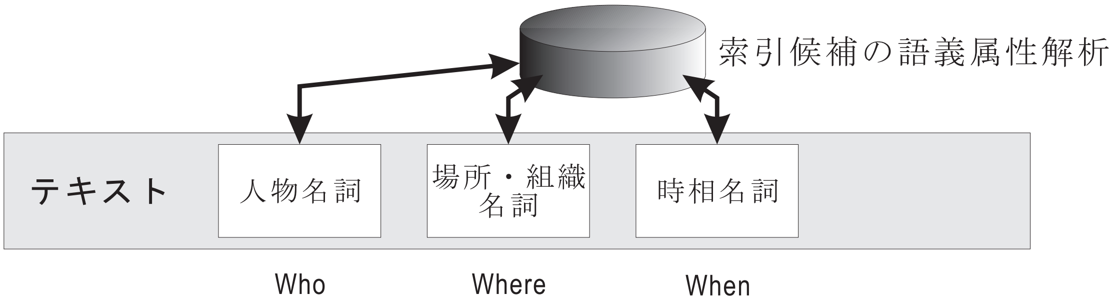


図 3.2: 自然言語処理の概要：索引候補の語義属性解析

3.2.2 映像中のテキスト情報

テレビ映像中に存在するテキスト情報としては表3.1に示すようなものが挙げられる．

表 3.1: テレビ映像中に存在するテキスト情報

メディア	種別
音声（Audio）	主音声（Main Audio）
	副音声（Sub Audio）
キャプション（Caption）	字幕・テロップ（Open Caption）
	クローズドキャプション（Closed Caption）

以下に各々について簡単に述べる．

主音声

基本的には映像中の被写体の発する音声であるが，ドラマにおけるナレーションやニュースにおけるキャスタの原稿朗読のように，必ずしも被写体の発声と一致しないことや，具体的な画像内容と直接関係ないこともある．しかし，映像中のテキスト情報としては最も情報量が多い情報源である．

副音声

音声多重放送の機能を利用し，主音声と並行して放送される音声である．一部のドラマなどにおいて，主に視覚障害者のために画像内容の描写を行っている．そのほか，2ヶ国語放送の際の外国語翻訳音声や，娯楽番組において本放送とは全く無関係な内容の音声が提供されることもある．

字幕・テロップ

画像の上にあらかじめ挿入される文字列．音声の書き下しや翻訳を字幕，他のものをテロップと区別することもあるが [193]，本論文では両者を合わせて字幕と呼ぶ．

（社）日本放送協会（NHK）では，可読性に留意して，横方向の場合は 1 行あたり 15 文字を 2 行まで，縦方向の場合は 1 列あたり 11 文字を 2 列までを，同時に表示する字幕の長さの目安としている [177]．

一般に漢字文化圏のアジア諸国のテレビ映像においては高頻度で字幕が出現するが，欧米では最も字幕の出現頻度が高いと思われるニュース映像においても，最低限必要な人物や場所・組織に関する非常に簡単なものが出現するに過ぎない．これは，表意文字よりも表音文字の方が冗長になりやすく，解像度の点から詳しい情報を表示しきれないこと，表音文字では省略がしにくいこと，などが理由として考えられる．なお，実験に用いた映像中の字幕の出現頻度は，毎分 7.68 件，毎ショット 0.97 件であった[1]．

クローズドキャプション

主に聴覚障害者のために，文字放送の形で提供されるディジタルテキスト[2]．米国においては生放送を含め，大多数の番組で主音声を書き下したものが文字放送により提供されている[3]．しかし，音声認識精度や仮名漢字変換を伴う日本語入力上の問題から，わが国では依

[1] 国立国語研究所の「テレビ放送の語彙調査」[192] によれば，報道系番組における画像中のテキストの出現頻度は毎分 13.6 語である．また，これとは異なる標本において，出現した語のうち字幕類は 72.7%を占めていた．標本が異なるため単純には算出できないが，この統計からも字幕の出現頻度は毎分 10 語程度であることが分かる．

[2] 米国連邦規制 47CFR79.1: Closed captioning of video programming によれば，クローズドキャプションは以下のように定義されている：“The visual display of the audio portion of video programming contained in line 21 of the vertical blanking interval (VBI) ...”

[3] 1996 年改正通信法の施行により，米国連邦通信委員会（FCC）は，一部の例外を認めつつ，2006 年元日を目処に 1998 年元日以降に製作された映像番組の 100%，それ以前のものの 75%を対象にクローズドキャプションの付与を行う義務を放送事業者に課した．現在はこれよりは若干緩い規制のもとに段階的に移行中である．また，1993 年 7 月以降に米国内で販売される 13 インチ以上のすべてのテレビジョン受像機に解読装

然，ドラマやドキュメンタリを中心とした事前入力可能な少数の番組で提供されるのみである[4]．特に，生放送の日本語ニュース映像への付与は困難であり，ようやく 2000 年度に話者を限定した音声認識を用いてキャスタの原稿朗読部のみに対して試行的に始まろうとしている段階である．

これらのうち，字幕と副音声は，画像内容をそのまま反映した内容を多く含み，索引として利用するのに適している [27, 199]．主音声やクローズドキャプションなど，他の情報源はいずれも原理的には索引候補となり得るが，画像内容との対応が明確ではなく，また一般に冗長なテキストであるため，索引となり得る語句や文を抽出するために，今日の自然言語処理技術では依然困難な処理が必要になる．さらに，音声認識技術は進歩を遂げているものの，主・副音声の認識結果を索引付けに利用できる水準には達していない．

そこで，本研究ではニュース映像において重要かつ具体的な画像内容を反映した情報を簡潔に表現したテキスト情報として，字幕を索引候補として用いることにする．

3.2.3 接尾辞に基づく字幕の語義属性解析

3.2.1で述べたような索引候補のテキストの語義属性を明らかにするために，接尾辞に基づく語義属性解析を行う．

表 3.2: 接尾辞による語幹の語義の変化の例

橋本			→	橋本	[名詞-普通名詞]			
橋本	+	首相	→	橋本	[名詞-固有名詞-人名]	+	首相	[名詞-普通名詞]
橋本	+	市	→	橋本	[名詞-固有名詞-地名]	+	市	[名詞-普通名詞]

一般に特定の名詞が普通名詞であるか固有名詞であるか，固有名詞の場合でも人名，地名，組織名のいずれであるかを，前後の文脈を考慮せずにその名詞単独から判断することは人間にとっても著しく困難である．実際には，表 3.2に例示するように，接尾辞により複合名詞や名詞句の語義が決定されると考えられる．この例では「橋本」単独では語義を決定できないが「首相」と結合することにより人名であることが「市」と結合することにより地名であることが，各々明らかになる．

このように，一般に日本語において，接尾辞を解析することにより名詞句全体の語義を決定できる．本研究では，このような性質を仮定したうえで，接尾辞に注目することにより，大部分が複合名詞や名詞句からなる字幕の語義属性解析を行う．

置の内蔵が義務づけられている．

[4]NHK では 1999 年度には，地上波 2 局及び衛星波 1 局において，のべ 62 番組（毎週 49 時間 38 分）に対してクローズドキャプションの提供を行っている [203]．

3.3 画像処理部

本節では，提案手法における画像処理の概要及び，様々な画像特徴量の種類と性質について述べる．

3.3.1 画像処理の概要

提案手法では（1）人物（誰が）（2）場所・組織（どこで）（3）時相（いつ）（4）行為（何を）に関する索引付けを行う．図 3.3に示すように，このうち（1）及び（2）に関しては画像を人物領域と背景領域に分割し，各領域の画像特徴量から内容推定を行い（4）に関しては，典型的ショットの分類により，おおまかな行為内容（演説，集会など）の推定を行う．なお（3）に関しては，画像からの解析は困難であるため，自然言語処理の結果を利用する．

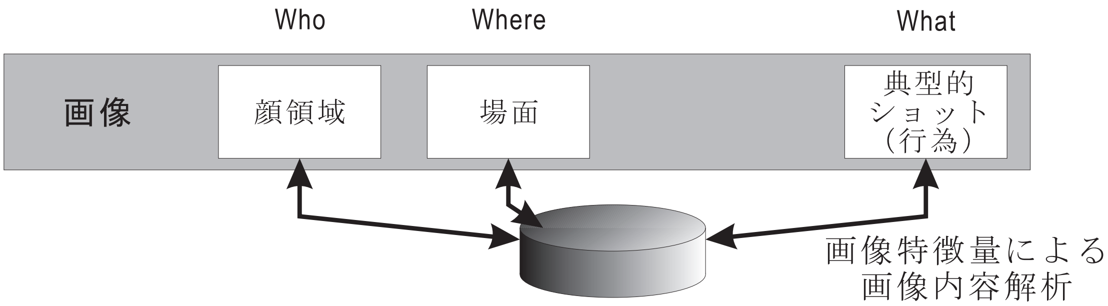


図 3.3: 画像処理の概要：画像特徴量による画像内容の解析

3.3.2 様々な画像特徴量

画像特徴量としては表 3.3に示すようなものが挙げられる．以下に各々について述べる．

色彩に関する特徴量

色彩的特徴は画像内容を把握する際には非常に重要な手掛かりとなる．その中でも代表的な，色出現頻度分布とは，あるピクセルがある特定の色である確率分布であり，色共起頻度分布とは，一定距離離れた 2 ピクセルが特定の色の組である確率分布である．

具体的には RGB や HSI といった様々な表色系における色成分の値に対してこのような統計的処理を施す．

表 3.3: 様々な画像特徴量

種別	特徴量
色彩	色出現頻度分布（Color Histogram） 色共起頻度分布（Color Correlogram）
テクスチャ	エッジ強度（Edge Intensity） 直線成分（Linear Segment）
動き	オプティカルフロー（Optical Flow）
構図	空間周波数解析（Spatial Frequency Analysis） 領域分割（Regional Segmentation）
具体的事物	顔領域（Facial Region） その他

テクスチャに関する特徴量

エッジとは画像中の領域の輪郭部のことであり，明るさが急激に変化する境界線として抽出する．画像特徴量としては，エッジの強度，密度，直線成分の方向などを利用し，形状の複雑さ，画像中の領域の込み入り具合，平行線分の抽出による建物など人工物の検出などに用いられる．

動きに関する特徴量

オプティカルフローとは各ピクセルにおける動きベクトルの向きのことである．オプティカルフローを解析することにより，カメラの動きや画像中の特定領域の動きなどを把握できる．暗い領域や特徴点が取りにくい領域において正しく動きが検出できないこと，多大な計算量を要求することなどから，実用性に若干の問題がある．

構図に関する特徴量

空間周波数解析の代表例として，離散余弦変換（DCT）係数が挙げられる．これにより，画像のおおまかな構図を抽象的な特徴量として抽出できる．また，画像を複数の領域に分割し，各領域から上記の各種特徴量を抽出することによる構図の把握を行う手法もある．

具体的事物に関する特徴量

特定の対象領域に限定した認識を考える際には，目的に応じて具体的な事物のモデルを知識として保有し，認識することがしばしば行われる．このような処理も一種の特徴量抽出

と考えることができる．本研究では，人物像が頻繁に出現するというニュース映像の特徴を考慮し，顔領域の検出を行い，特徴量の一つとする．

3.3.3 画像特徴量に基づく画像内容解析

3.3.1で述べたような画像内容の把握のために，画像特徴量に基づく画像内容解析を行う．

ニュース映像の特徴として，人間の社会的行為を扱った内容が多い点が挙げられる．そのため，画像中に人物像，特に上半身が大きく映っていることが非常に多い．一方，頻出する話題については，それらの人物が登場する背景の場面（場所）が共通であることが多い．

このような特徴を考慮すると，被写体の人物及び背景の領域を分離して各々の画像特徴量を解析することにより，人物及び場面の内容を解析することができると考えられる．頻出する人物や場面に関しては，人物領域や背景領域が類似した画像特徴を示すはずであり，推定のための知識ベースの構築も比較的容易であると考えられる．また，おおまかな行為の推定には，人物の存在や人数を主な分類規則とした典型的ショット分類を行えばよい．

そこで本研究では，画像中に一定以上の大きさの人物が存在する場合は，画像を人物領域と背景領域に分割し，知識ベースを参照することにより，各領域毎の画像特徴量に基づく画像内容の解析を行う．

3.4 統合処理部

本節では，提案手法における自然言語処理結果と画像処理結果との統合処理の概要及び，その際に利用する知識ベースについて述べる．

3.4.1 統合処理の概要

統合処理部では，自然言語処理部で語義属性解析された索引候補を画像内容に応じて対応付ける．つまり，図3.4に示すように（1）人物に関する索引候補は，対応する人物像に対して（2）場所・組織に関する索引候補は，対応する場面に対して索引として付与する．このように属性を考慮した統合処理を行うことにより，画像と具体的に対応した索引付けが実現される．

なお（3）時相及び（4）行為に関する索引付けに関しては，各々単一メディアからの抽出となるため，統合処理部では扱わない．

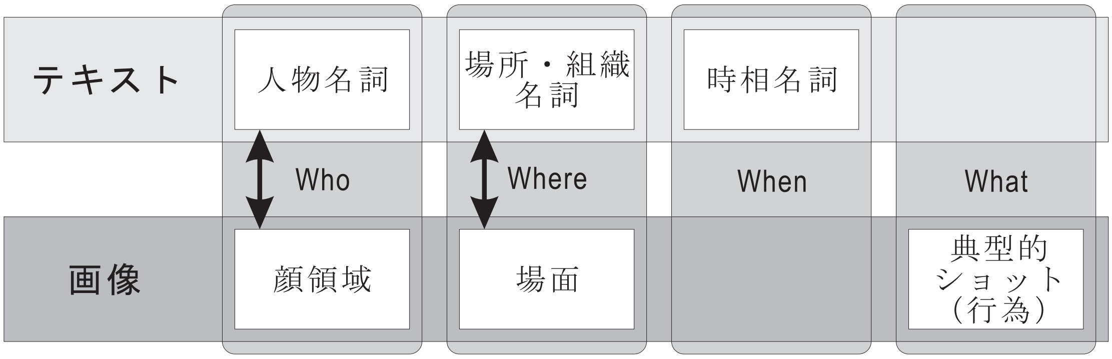


図 3.4: 統合処理の概要：索引候補と画像の属性レベルでの対応を確認した対応付け

3.4.2 画像特徴と画像内容の関連に関する知識ベースの利用による索引付け

このような索引付けを行うにあたり，画像特徴量と画像内容の関連に関する知識ベースを利用する（1）及び（2）に関しては，明示的な記述が困難であるため，ニュース映像中に頻繁に登場する同一人物及び同一場面に関する大量の事例画像を用いて知識ベースを構築し，利用する（4）に関しては，典型的ショット分類により得られるので，最初から明示的に記述された知識（規則）を用いた対応付けがなされていると考えられる．

原始的な画像特徴量を用いてこのような知識を獲得することは，一般に極めて困難だが，ニュース映像のように頻繁に同一人物や同一場所が出現する場合には実現可能と思われる．

更に，このような知識ベースを利用することにより，共起するテキストが存在しない映像についても，大雑把な内容推定を行うことができる．そのため，満足な索引付けを行う

ためには若干情報量が不足している字幕のような情報源を利用する際に有効であると思われる．

3.5 まとめ

本章では，映像内容を考慮した統合メディア処理による索引付け機構の全体像を提案し，自然言語処理，画像処理，統合処理の各処理部の役割について簡単に述べた．各処理の詳細及び評価実験については，続く第 4 章から第 6 章において述べる．

第 4 章

接尾辞に基づく字幕の語義属性解析

本章では，映像内容を考慮した索引付けに必要となる，索引候補のテキストに対する処理として，字幕の語義属性解析について述べる．

まず，字幕の性質について述べ，具体的事物を指示する名詞句が多いことを統計的に示す．次に，本研究で提案する独自の名詞句の語義属性解析手法を提案し，そのために必要となる属性別名詞辞書の作成過程を示す．そして最後に，作成された辞書を用いて，実際のニュース映像中に出現する字幕の語義属性解析を行った結果を示し，評価を行う．

4.1 字幕の語義属性解析

本節では，字幕の言語的性質について述べたあと，それらの性質を考慮して，本研究で独自に提案する字幕，つまり名詞句の語義属性解析手法を紹介する．

テレビニュースの字幕の解析に関する研究としては，言語的解析に加え，表示位置などを利用する手法がある [180]．この手法は本研究で必要とする技術そのものであるものの，字幕の表示位置などを利用する点で，デザインは番組毎に異なり，また同一番組でも定期的に変化するため，汎用性の点から問題があると考えた．そこで，以下ではテキスト情報のみによる字幕の語義属性解析について論じる．

4.1.1 字幕の言語的性質

字幕は他の一般のテキストと比較して，文法的にも内容的にも特殊な性質を有する．そのため，既存の自然言語処理技術をそのまま流用することに限界がある反面，それらを逆に利用した処理も可能である．

以下にそれらの性質を文法的特徴と内容的特徴とに分けて述べる．

文法的性質

今日では日本語に限っても，形態素解析システム JUMAN [205] や茶筅 [207] をはじめ，構文解析システム KN パーザ [204] など，一般文を対象とした高度な言語処理技術が容易に利用可能な形で存在する．しかし，字幕の処理に関しては，以下に列挙するように，通常のテキストとは異なる文法的傾向がみられる [177, 180] ため，通常のテキストを前提として設計されたこれらの技術を必ずしもそのまま利用することはできない．なお本研究では，より簡潔に重要な内容を含むことが多い，非文章的字幕を主に利用する．

- 文章的字幕
  1. 単文
  2. 主語の省略
  3. 体言止め
  4. 特定の助詞（に，か，も，は，を，へ）で終わる
- 非文章的字幕
  5. 文中の助詞の省略
  6. 名詞列からなる名詞句

表 4.1: テレビニュースに出現する字幕の種類と比率

	種別	例	比率	主な文法的性質
(a)	場所・組織	東京大学	28.11%	6
(b)	人物	小渕首相	13.19%	6
(c)	タイトル†	“首相訪米へ”	9.43%	3，4
(d)	時相	昨日	7.00%	6
(e)	発言††	“大きな音がして外を見ると...”	5.52%	2，口語文
(f)	放送技術†††	中継	3.52%	6
(g)	描写	無傷だった先頭車両	1.55%	1，2，3，5，6
	その他		31.67%	

† 主に新しい話題の冒頭に出現するもの．
†† 発言の要約や翻訳を含む．
††† 放送技術上のもの．

内容的性質

一方，字幕の内容的な性質にも一定の傾向が見られる．表 4.1に，370 分間のニュース映像に出現した 2,842 件の字幕を内容的性質毎に分類し，それらの占める割合を示す．この結果から（a）（b）（g）を合わせた 42.86%の字幕が画像中に実際に映っている事物の索引候補となり得ることや（a）（b）（d）（g）を合わせた 49.86%の字幕が，いわゆる 4W と呼ばれる「いつ（when），どこで（where），誰が（who），何を（what）している」の各属性の索引候補となり得ることが分かる．更に（c）も一般的に大雑把な話題単位での索引として利用可能であり，合わせて，前者で 52.29%，後者で 59.29%の字幕が索引の候補となり得る．

出現頻度は，各々毎分 3.29 件，同 3.83 件，あるいは毎ショット 0.42 件，同 0.49 件である．この頻度では本格的な索引付けを行うには不十分であるが，画像処理部で行う場面推定や，最終的な索引付けの際に行う類似画像中の字幕の利用により補うことを考えれば十分である．

4.1.2 名詞句の語義属性解析

4.1.1で述べたような言語的性質を考慮し，主に名詞句からなる（1）場所・組織（2）人物，（3）時相に関する字幕の語義属性解析を行う．

まず，既存の名詞（句）解析手法として，関連研究を紹介し，次に独自の字幕，つまり名詞句の語義属性解析手法を提案する．

一般のテキストにおける名詞の語義属性解析

類似した語義属性解析問題として，第 6，7 回の Message Understanding Conference（MUC）における課題の一つとして定義された，Named Entity task [108, 171] がある．この課題では，提示されたテキストに対して，人物，場所，組織，時相，数値に関する表現を抽出してタグ付けすることを参加者に課している．最終的な評価の結果，再現率，適合率ともに 95%以上の手法がいくつか上位に入選している [172]．

一方，那須川は一般的なテキスト中の固有名詞の語義属性解析手法を提案している [147]．この手法は隣接する文間の文脈を解析することにより，固有名詞が人物の名前であるか，場所・組織の名前であるかを解析する．

これらの課題や手法は，人物，場所，組織に関する表現の抽出及び分類を行うという点においては本研究で必要とする技術に非常に近いが，いずれも固有名詞に解析対象を限定している点で異なる．このような固有名詞の解析と（固有名詞を含む場合を含めた）一般の名詞句の解析とでは，本質的に異なる困難さを孕む問題である．表 3.2で例示したように，基本的に固有名詞単体ではその語義を決定することは不可能であり，那須川の手法や Named Entity task の解法のように，文内の格構造や周囲の文脈から推測することになる．しかし，そのような情報を得にくい字幕に対する適用は困難であり，ここで提案する接尾辞に基づく語義属性解析手法により，単体での解析が可能となる．

接尾辞に基づく名詞句の語義属性解析手法

以上の字幕の性質と既存手法による処理の問題点を考慮し，本研究では字幕に対して単独で語義属性解析を行うことを目指した．表 3.2に示したように，一般に日本語において，接尾辞を解析することにより名詞句全体の語義を決定できる[1]．本研究では，このような性質を仮定したうえで，接尾辞に注目することにより，大部分が複合名詞や名詞句からなる字幕の語義属性解析を行う．

図 4.1にこのような方針による字幕の語義属性解析の例を示す．まず，字幕に形態素解析を施して接尾辞を切り出し，切り出した接尾辞を各語義属性別名詞辞書中の語と比較することにより解析を行う．なお，この例や以下では接尾辞に重点をおいて議論するが，提案手法は「俳優」や「台所」のように，単独で人物や場所を示す普通名詞をも収集，解析対象とする．

[1]この性質はおそらく中国語，朝鮮語をはじめとする他の東アジアの諸言語にも該当すると思われる．

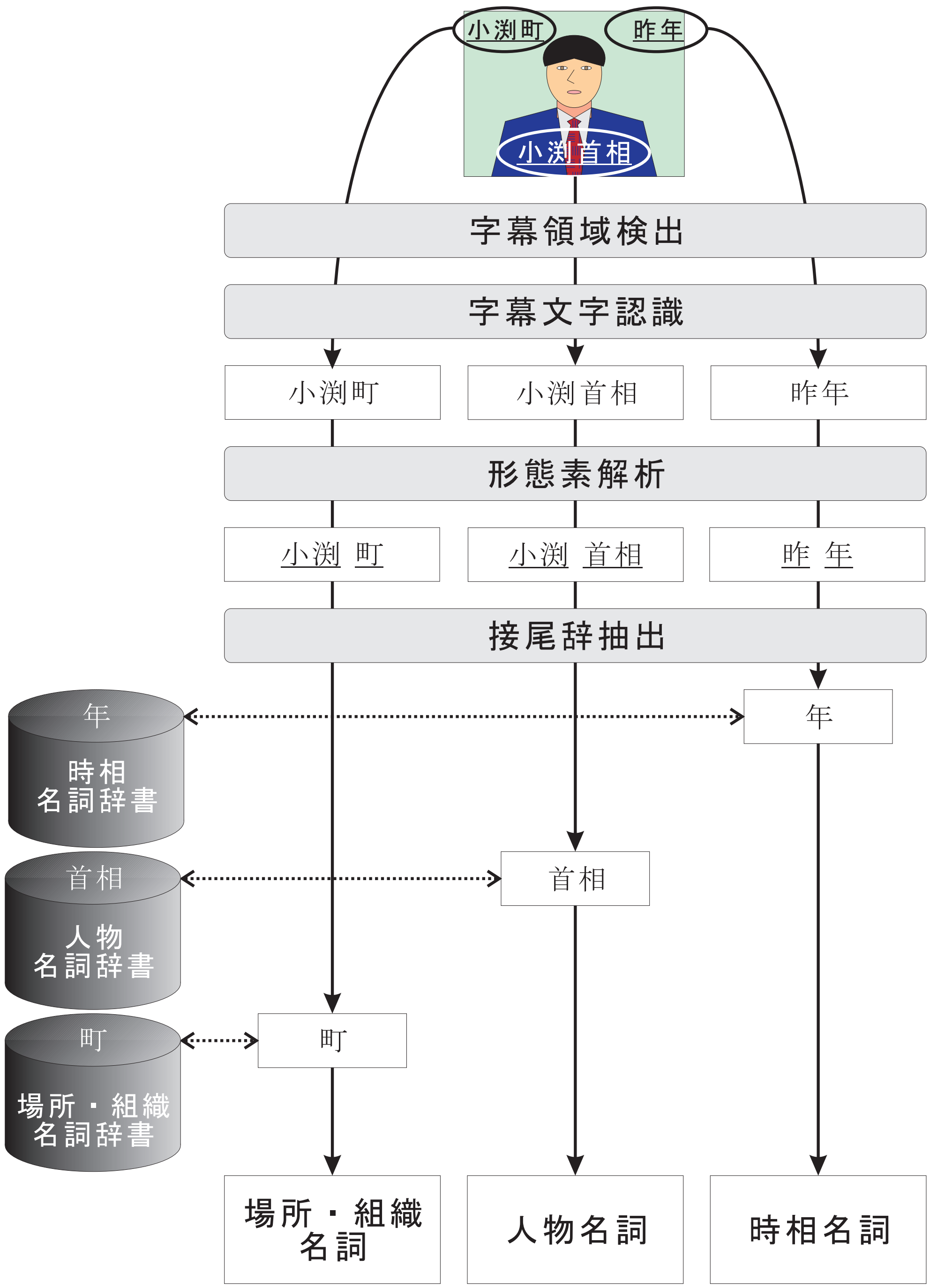


図 4.1: 接尾辞に注目した字幕の語義解析の例

4.2 語義属性別名詞辞書の作成

本節では，4.1で提案した接尾辞に注目した字幕の語義属性解析手法を実現するために必要となる語義属性別名詞辞書の作成について述べる．具体的には（1）人物（2）場所・組織，（3）時相を示す接尾辞を集めた辞書を各々作成する．

前述のように，純粋な接尾辞以外に「俳優」「台所」「今日」のように単独で人物，場所・組織，時相を示す名詞も収集対象とする．なお，固有名詞は，単独での語義属性解析が困難なことや，形態素解析時に解析されることが多いことなどから，収集対象から除外した．

辞書の作成にあたり，表4.2に示す2つのテキストコーパスを利用した．これらのテキストコーパスは人手で形態素解析された結果が本文にタグ付けされており，形態素の区切りや種別を正しいものと考えて利用する．これらを合わせて47,374文から一定の条件のもとに接尾辞の抽出を行った．

表 4.2: 語義属性別名詞辞書の作成に用いたテキストコーパス

コーパス名	収録文数	出典
RWC テキストコーパス（第1版）[210] 中の RWC-DB-TEXT-95-2	27,418 文	1994 年版毎日新聞
京都大学テキストコーパス（第2版）[206]	19,956 文	1995 年版毎日新聞
合計	47,374 文	

以下では，各々の語義分類別名詞辞書の作成手順と収集結果について述べる．

4.2.1 人物名詞の収集

ここでは，人物名詞，すなわち「～博士」のように人物を示す接尾辞や「俳優」のように単独で人物を示す普通名詞の収集手順と結果について述べる．

コーパスからの抽出基準

人物名詞は表4.3に示す基準に従ってコーパスから抽出した．ここでは，接尾辞「ら」及び「たち」が原則として複数の人間（のみ）を表すのに用いられることを利用した．

また，表4.3の最下段にこの基準による抽出例を，表4.4に抽出された語の一部を頻度の高い順に列挙する．

類義語辞書による語彙の拡張

表4.3に示した基準によるコーパスからの抽出により136件の接尾辞が抽出されたものの，汎用的な字幕解析に適用するには語彙が不十分である．そこで，類義語辞書を用いて語彙

表 4.3: コーパスからの人物名詞抽出条件

以下のパターンと語及び形態素が一致する部分から，下線部の語を抽出する．
RWC，京大両コーパスでは形態素体系が異なるため，各々における条件を示す．
抽出例は日本語形態素解析システム JUMAN による出力結果を示す．

語	* →	<u>*</u>	「ら」 「たち」
RWC 形態素	[名詞-*] →	[名詞] [名詞-非自立-*] [名詞-接尾]	[名詞-接尾]
京大 形態素	[名詞-*] [接尾辞-名詞性名詞接尾辞] →	[名詞-普通名詞] [接尾辞-名詞性名詞接尾辞]	[接尾辞-名詞性名詞接尾辞]
抽出例	田中 [名詞-人名] 英彦 [名詞-人名]	<u>教授</u> [名詞-普通名詞]	ら [接尾辞-名詞性名詞接尾辞]

の拡張を行った．拡張にあたっては，類義語辞書の分類段落のうち，抽出された語が含まれるものに含まれる語は全て該当する語義属性を有すると判断した．類義語辞書としては，「分類語彙表」(以下，旧分類語彙表) [208] 及び「『分類語彙表』形式による語彙分類表 (増補版)」(以下，新分類語彙表) [209] を用いた．各々の類義語辞書の規模を表 4.5に比較して示す．各分類段落は平均 6.3 から 6.8 程度の語を含み，この規模の語彙の拡張が期待される．

なお，両類義語辞書による語彙の拡張の際に該当した分類項目及び段落を付録に一覧にして示す．

収集結果

表 4.6に辞書の作成過程で得られた語彙の推移を示す．最終的に，旧分類語彙表による拡張で 860 件，新分類語彙表による拡張で 1,795 件の人物名詞が収集された．

4.2.2 場所・組織名詞の収集

ここでは，場所・組織名詞，すなわち「～駅」のように場所や組織を示す接尾辞や「台所」のように単独で場所や組織を示す普通名詞の収集手順と結果について述べる．

表 4.4: 抽出された人物名詞（上位 73 件）

順位	抽出された語	頻度	順位	抽出された語	頻度	順位	抽出された語	頻度
1	者	75	20	委員長	6	49	代理	2
2	氏	57	27	大使	5	49	代行	2
3	さん	51	27	社長	5	49	隊員	2
4	会長	32	27	幹事	5	49	村長	2
5	長	26	27	課長	5	49	蔵相	2
6	議員	21	31	大統領	4	49	船長	2
7	人	16	31	総書記	4	49	青年	2
7	教授	16	31	選手	4	49	少将	2
9	家	15	31	社員	4	49	女性	2
10	幹部	14	31	市長	4	49	助手	2
11	首相	13	31	軍人	4	49	所長	2
12	被告	10	31	患者	4	49	主	2
13	業者	9	38	労相	3	49	室長	2
14	長官	8	38	両氏	3	49	次官	2
14	相	8	38	役員	3	49	佐	2
14	書記長	8	38	副総裁	3	49	顧問	2
14	客	8	38	部長	3	49	教師	2
14	員	8	38	生	3	49	幹事長	2
19	代表	7	38	主席	3	49	医	2
20	容疑者	6	38	経験者	3	49	委員	2
20	職員	6	38	君	3	49	ランナー	2
20	助教授	6	38	監督	3	49	ダイバー	2
20	局長	6	38	学生	3	49	さま	2
20	関係者	6	49	難民	2		（以下省略）	
20	官	6	49	団長	2			

表 4.5: 語彙拡張に用いた概念辞書の規模

	大分類項目数	小分類項目数	分類段落数	のべ収録語数	（異なり数）
旧分類語彙表	4	822	5,136	36,780	32,548
新分類語彙表	4	842	10,334	87,743	70,858

表 4.6: 収集された人物名詞の語彙の推移
分類項目及び分類段落は分類語彙表収録語の属するものの数を示す

	旧分類語彙表		新分類語彙表	
	コーパスからの抽出	分類語彙表による拡張	コーパスからの抽出	分類語彙表による拡張
分類語彙表収録語数	102	826	117	1,776
(分類段落数)	79	79	125	125
(分類小項目数)	42	42	59	59
分類語彙表非収録語数	34	34	19	19
総語彙数	136	860	136	1,795

コーパスからの抽出基準

場所・組織名詞は表 4.7に示す基準に従ってコーパスから抽出した．ここでは，格助詞「から」,「で」,「に」,「へ」,「より」,「にて」が場所や方向を示すのに用いられることを利用した．

また，表 4.7の最下段にこの基準による抽出例を示した．

人物名詞の削除

表 4.7に示した基準に従った抽出では，表 4.8に例示するように，本来は人物名詞である語が混入してしまう．これは，格助詞「へ」が表 4.7最下段の例では方向を示すのに対し，表 4.8の例では目的格の人物を指すことに示されるように，助詞の用法の多様性によるものである．このようにして混入した人物名詞を除去するために，ここで抽出された語から，4.2.1で抽出された人物名詞を削除する．

表 4.9に人物名詞を削除した後の抽出された語の一部を頻度の高い順に列挙する．

類義語辞書による語彙の拡張

人物名詞の削除後に 696 件の接尾辞が残ったものの，汎用的な字幕解析に適用するには語彙が不十分である．そこで，4.2.1で述べたように，類義語辞書を用いて語彙の拡張を行った．

収集結果

表 4.10に辞書の作成過程で得られた語彙の推移を示す．最終的に，旧分類語彙表による拡張で 4,621 件，新分類語彙表による拡張で 7,908 件の場所・組織名詞が収集された．

表 4.7: コーパスからの場所・組織名詞抽出条件

以下のパターンと語及び形態素が一致する部分から，下線部の語を抽出する．
RWC，京大両コーパスでは形態素体系が異なるため，各々における条件を示す．
抽出例は日本語形態素解析システム JUMAN による出力結果を示す．

語	*	* →	*	「から」 「で」 「に」 「へ」 「より」 「にて」
RWC 形態素	[名詞-固有名詞] [名詞-固有名詞-組織] [名詞-固有名詞-地域-*]	[名詞-*] →	[名詞] [名詞-非自立-*] [名詞-接尾]	[助詞-格助詞]
京大 形態素	[名詞-地名] [名詞-組織名]	[名詞-*] [接尾辞-名詞性名詞接尾辞] →	[名詞-普通名詞] [接尾辞-名詞性名詞接尾辞]	[助詞-格助詞]
抽出例	東京 [名詞-地名]	大学 工学部 [名詞-普通名詞] [名詞-普通名詞]	構内 [名詞-普通名詞]	へ [助詞-格助詞]

表 4.8: 場所・組織名詞抽出に人物名詞が混入する例

抽出例は日本語形態素解析システム JUMAN により形態素解析した結果を示す

抽出例	東京 [名詞-地名]	大学 工学部 [名詞-普通名詞] [名詞-普通名詞]	教授 [名詞-普通名詞]	へ [助詞-格助詞]

表 4.9: 抽出された場所・組織名詞 (上位 100 件)

順位	抽出された語	頻度	順位	抽出された語	頻度	順位	抽出された語	頻度
1	市	163	33	事件	16	66	北部	8
2	内	108	33	五輪	16	66	経済	8
3	側	78	33	空港	16	66	機	8
4	大会	60	38	共和国	14	66	決議	8
5	県	46	38	部	14	66	上空	8
6	市内	43	40	議会	13	66	センター	8
7	駅	39	41	所	12	66	各地	8
7	政府	39	41	内部	12	66	気象台	8
7	戦	39	41	付近	12	77	店	7
10	問題	38	41	地震	12	77	法	7
11	署	36	45	当局	11	77	省	7
12	地方	31	45	遺跡	11	77	以外	7
13	区	30	45	出身	11	77	諸島	7
14	大震災	29	45	港	11	77	国境	7
14	国内	29	45	情勢	11	77	協議	7
16	会談	28	45	基地	11	77	政策	7
16	選	28	45	戦争	11	77	劇場	7
18	地裁	27	45	社	11	77	訪問	7
19	町	25	45	周辺	11	77	大陸	7
20	市場	23	45	前	11	77	筋	7
20	間	23	55	本部	10	77	地検	7
22	地区	22	55	ホール	10	77	各国	7
22	会	22	55	高裁	10	77	大	7
24	館	20	55	会館	10	92	支局	6
25	沖	19	59	ら	9	92	紙	6
25	会議	19	59	局	9	92	支部	6
25	地域	19	59	支店	9	92	線	6
28	村	18	59	南部	9	92	テレビ	6
28	場	18	59	都	9	92	関係	6
30	軍	17	59	圏	9	92	系	6
30	諸国	17	59	勢力	9	92	全体	6
30	語	17	66	島	8	92	選挙	6
33	州	16	66	高原	8		(以下省略)	
33	時代	16	66	政権	8			

表 4.10: 収集された場所・組織名詞の語彙の推移
分類項目及び分類段落は分類語彙表収録語の属するものの数を示す

	新分類語彙表		
	コーパスからの抽出	人物名詞の削除	分類語彙表による拡張
分類語彙表収録語数	581	378	4,460
（分類段落数）	532	486	486
（分類小項目数）	275	264	264
分類語彙表非収録語数	184	161	161
総語彙数	765	539	4,621
	旧分類語彙表		
	コーパスからの抽出	人物名詞の削除	分類語彙表による拡張
分類語彙表収録語数	674	607	7,819
（分類段落数）	764	697	697
（分類小項目数）	318	307	307
分類語彙表非収録語数	91	89	89
総語彙数	765	696	7,908

4.2.3 時相名詞の収集

ここでは，時相名詞，すなわち「～後」のように時相を示す接尾辞や「今日」のように単独で時相を示す普通名詞の収集手順と結果について述べる．

コーパスからの抽出基準

時相名詞は表 4.11に示す基準に従ってコーパスから抽出した．ここでは，格助詞「から」，「に」，「より」が時相を示すのに用いられることを利用した．なお，表 4.11には記していないが，京都大学テキストコーパスに限って［名詞-時相名詞］の形態素を有する全単語も併せて抽出した．

また，表 4.11の最下段にこの基準による抽出例を，表 4.13に抽出された語の一部を頻度の高い順に列挙する．

類義語辞書による語彙の拡張

表 4.11に示した基準によるコーパスからの抽出により 156 件の接尾辞が抽出されたものの，汎用的な字幕解析に適用するには語彙が不十分である．そこで，4.2.1で述べたように，類

義語辞書を用いて語彙の拡張を行った．

表 4.11: コーパスからの時相名詞抽出条件

以下のパターンと語及び形態素が一致する部分から，下線部の語を抽出する．
RWC，京大両コーパスでは形態素体系が異なるため，各々における条件を示す．
抽出例は日本語形態素解析システム JUMAN による出力結果を示す．

語	<u>*</u>	「から」 「に」 「より」
RWC 形態素	[名詞-副詞-*] [名詞-非自立-副詞可能-*]	[助詞-格助詞]
京大形態素	[名詞-副詞的名詞]	[助詞-格助詞]
抽出例	<u>ころ</u> [名詞-副詞的名詞]	から [助詞-格助詞]

収集結果

表 4.12に辞書の作成過程で得られた語彙の推移を示す．最終的に，旧分類語彙表による拡張で 1,215 件，新分類語彙表による拡張で 2,144 件の時相名詞が収集された．

表 4.12: 収集された時相名詞の語彙の推移

分類項目及び分類段落は分類語彙表収録語の属するものの数を示す

	旧分類語彙表		新分類語彙表	
	コーパスからの抽出	分類語彙表による拡張	コーパスからの抽出	分類語彙表による拡張
分類語彙表収録語数	128	1,187	136	2,124
（分類段落数）	114	114	213	213
（分類小項目数）	70	70	104	104
分類語彙表非収録語数	28	28	20	20
総語彙数	156	1,215	156	2,144

表 4.13: 抽出された時相名詞（上位 99 件）

順位	抽出された語	頻度	順位	抽出された語	頻度	順位	抽出された語	頻度
1	ため	465	35	通常	11	66	当初	3
2	時代	145	36	途中	10	66	途上	3
3	中	128	37	新年	9	66	前夜	3
4	間	123	38	方	8	66	生涯	3
5	前	117	38	昨年	8	66	初め	3
6	時	116	38	今日	8	66	時刻	3
7	ところ	79	38	ごと	8	66	月	3
8	うち	65	42	翌日	7	66	一発	3
9	際	54	42	明日	7	66	一昨年	3
10	場合	47	42	以外	7	66	よう	3
11	上	46	45	夜	6	66	さなか	3
12	直後	45	45	昔	6	80	来年	2
13	以来	43	45	春	6	80	当代	2
14	日	40	45	最中	6	80	通り	2
15	後	31	45	今後	6	80	中期	2
16	ほか	30	45	元	6	80	他	2
17	ころ	27	51	本来	5	80	世界一	2
18	戦後	25	51	当日	5	80	初頭	2
19	当時	24	51	このまま	5	80	最初	2
19	時期	24	51	あと	5	80	今度	2
19	まま	24	55	末	4	80	今月	2
22	とき	23	55	全部	4	80	今回	2
23	結果	22	55	前年	4	80	頃	2
24	今年	20	55	前日	4	80	月比	2
24	現在	20	55	初日	4	80	去年	2
26	度	16	55	最近	4	80	夏	2
26	直前	16	55	午前	4	80	一町	2
26	長期	16	55	元年	4	80	一章	2
29	以降	15	55	以前	4	80	一国	2
29	なか	15	55	以上	4	80	ついで	2
29	うえ	15	55	ゆえ	4	80	それぞれ	2
32	所	12	66	毎日	3		（以下省略）	
32	史上	12	66	年	3			
32	たび	12	66	年来	3			

4.3 実装と評価

本節では，4.2で作成した語義属性別名詞辞書を用いた，字幕の語義属性解析手法の実装について述べ，評価実験の結果を示す．

4.3.1 用いる既存技術

形態素解析：日本語形態素解析システム JUMAN

接尾辞による字幕の語義属性解析を行うために必要な形態素解析は，京都大学において開発された日本語形態素解析システム JUMAN [205] を用いた．JUMAN による形態素解析の役割は（1）形態素分割と，分割された形態素の文法的素性を明らかにする（2）狭義の形態素解析に分かれる．このうち，本手法では主に（1）の機能を利用し（2）については，末尾の形態素の品詞が名詞などの体言であることの判定を行うためのみに利用する．

日本語を始めとする東アジア諸語の表記法は，形態素間を空白などの分離記号で明示的に区切られないため，欧米諸語などに対する形態素解析と異なり，形態素分割が重要かつ困難な問題となる [191]．JUMAN における形態素分割は，以下の手順によって行われる [205]．

- ある特定の位置から始まる全ての可能な形態素を辞書引きによって得る
- 辞書引きによって得られた個々の形態素に対して，直前の位置に存在する全ての形態素との連接可能性の検討及びコストの計算を行う
- 連接可能性の検討により連接不可能と判断した形態素間の接続は行わず，その位置で連接可能なもののうち，コストが最小のものに対して，利用者が定義した閾値よりも差が小さいもののみを接続する

なお，4.3.2の実験 4.1 の中で部分的に解析結果を利用する JUMAN の固有名詞及び時相名詞に関する標準辞書中の収録語彙数は，表 4.14の通りである．

表 4.14: 日本語形態素解析システム JUMAN 標準辞書中の固有名詞・時相名詞の語彙数

種　別	固有名詞				時相名詞
	人　名	地　名	組織名	（合計）	
語彙数	20,317	12,714	560	33,591	1,095

類義語辞書：分類語彙表

4.2で作成した辞書は，各々の語義属性について，具体的には以下の 2 つの辞書からなる．

- 該当分類段落番号一覧
- 非収録語一覧

ここで，分類段落番号とは，分類語彙表の特定の分類段落を示す番号であり，非収録語とは，コーパスから抽出した語のうち，分類語彙表に収録されていなかったものである．分類語彙表とは，国立国語研究所において作成された類義語辞書であり，図 4.2に示すような階層構造からなる．

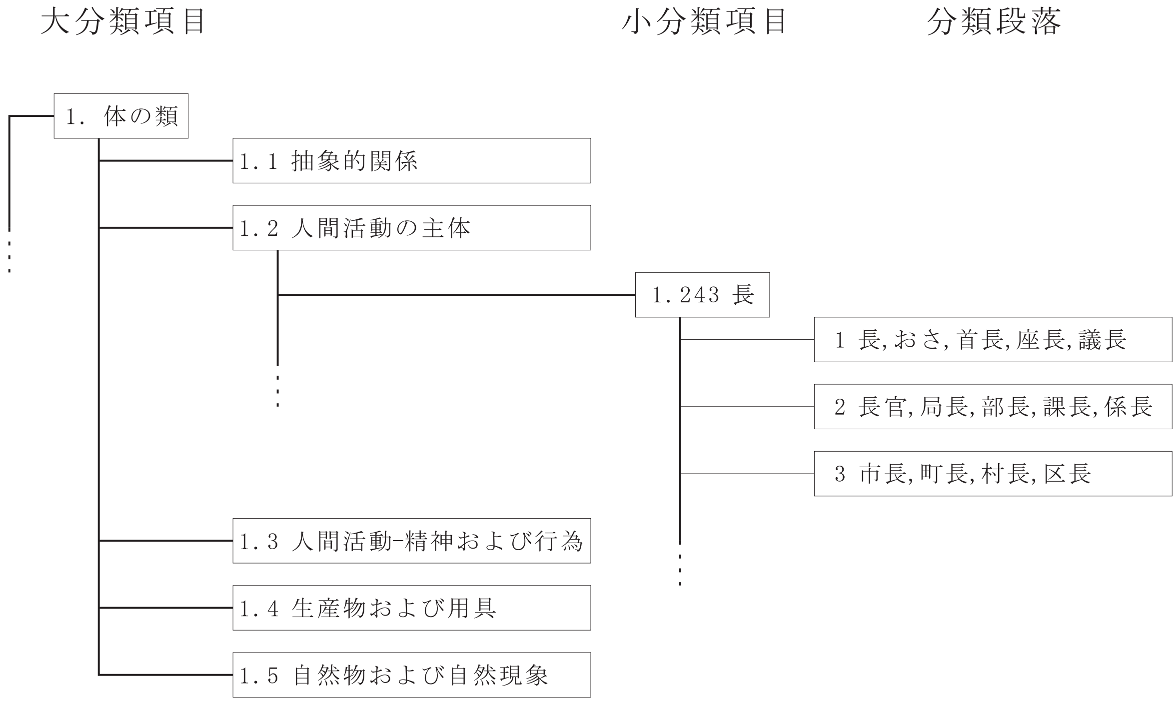


図 4.2: 分類語彙表の階層構造

本研究では，4.2.1で述べたように「分類語彙表」(旧分類語彙表)[208] 及び「『分類語彙表』形式による語彙分類表(増補版)」(新分類語彙表)[209] の 2 つを比較しながら用いた．各々の類義語辞書の規模は表 4.5に比較して示したが，以下に特徴について簡単に述べる．

旧分類語彙表は，「分類語彙表」の名称で 1964 年に国立国語研究所資料集の第 6 冊として刊行された類義語辞書である．語彙は，「現代雑誌九十種の用語用字」中の使用頻度の高い語や「教育基本語彙」から収録している．そのため，元来の収録語彙が新聞やニュース番組で用いられる語彙とは若干異なるうえに，今日の基本的語彙は当時のそれと較べると，

大きく変わっている．そのため，本研究で利用する場合に，語彙が必ずしも十分でないこと，分類項目の粒度が適切でないことが考えられる．

同所では 1981 年以来増補作業を行っており，1996 年に「『分類語彙表』形式による語彙分類表（増補版）」の名称で新分類語彙表が評価用として公開された．なお，新分類語彙表は正式刊行以前の評価版であるため，提供されたものに独自の修正を加えたこと，それでもなお若干の誤謬や曖昧な分類があることに注意されたい．

字幕検出・認識

3.3や第 5 章で述べる画像処理の本流と並行して，字幕領域の検出と文字認識を行う必要がある．高精度の OCR（光学的文字認識）技術は存在するものの，テレビ映像中の字幕は，走査線の本数（NTSC 方式の場合 525 本）が少ないために低解像度であることと，文字の背景が均一でないことから，既存手法での認識は困難である．これらの問題は，特に漢字など複雑な形状の文字を扱う際に顕著であるが，新聞記事などの装飾付き見出しからの文字の分離 [164, 197] や，ニュース画像中の字幕の分離に関する優れた手法 [115, 128, 132, 138, 163] も提案されつつある．このような技術の進歩とあわせて，将来の放送のディジタル化に伴い，字幕に相当する情報が電子的なテキストとして映像に付随して放送される可能性も考慮にいれ，本稿では，字幕認識は実装せずに人手で書き下したものを利用した実験結果を示す．

4.3.2 実験 4.1：ニュース字幕に対する語義属性解析

4.2で作成した語義属性別名詞辞書を用いて，実際のニュース映像中の字幕の解析を行った．

実験条件

字幕解析は，図 4.1に示した手順に従って行った．本研究においては，字幕領域検出と文字認識は目視と人手により行い，形態素解析には日本語形態素解析システム JUMAN を用いた．JUMAN は形態素の境界とその属性または品詞を出力するので，ここでは解析対象の字幕の末尾の形態素を各語義属性別名詞辞書と比較して解析を行った．

一方，JUMAN 自身も固有名詞辞書及び時相名詞辞書を内蔵しており，末尾に人名，地名，組織名の固有名詞が存在する字幕，及び一部の時相名詞に関しては JUMAN の出力結果を解析結果として採用した．

このような方針に従い，以下に示す手順によって語義属性解析を行った．

- JUMAN を用いて形態素解析を施す
- もし，末尾の形態素が固有名詞・時相名詞と解析されたら，JUMAN の解析結果（人名，地名，組織名，時相）を利用
- それ以外の場合は，語義属性別名詞辞書を参照して判断

実験に用いた字幕は表 4.1で分析したものと同じ，370 分間のニュース映像に出現した 2,546 件の字幕である．なお，話題の冒頭のタイトル字幕と，CG による模式図などに現れる字幕については，実験対象から除外した．これは，これらの場合は画像処理により容易に検出できること，これらの字幕は直接画像内容を描写することが少ないためである．

なお，評価の際に用いる正解は第三者により与えられたものを用いた．

4.3.3 実験結果と評価

実験結果

解析結果を表 4.15から表 4.17に示す．また，図 4.3及び図 4.4 に再現率及び適合率に分けて図示した．図表中「本手法単体」に示す数値は，4.2で作成した辞書を利用した本手法単独での解析結果を「JUMAN 単体」に示す数値は JUMAN 単独での解析結果を示す．前述のように，末尾に人名，地名，組織名の固有名詞が存在する字幕，及び一部の時相名詞に関しては JUMAN により解析される「総合」は両処理により解析された総合的結果を示す．

以下のように定義される，適合率と再現率を評価尺度として用いた．各手法の結果に対する適合率は，各手法単独の性能を示し，再現率は，総合的結果の再現率に対する各手法の貢献度を示している．

$$
\begin{aligned}
\text{適合率} &= \frac{\text{正答数}\ (N_c)}{\text{正答数}\ (N_c) + \text{誤答数}\ (N_m)} && (4.1)\\
\text{再現率} &= \frac{\text{正答数}\ (N_c)}{\text{正答数}\ (N_c) + \text{洩れ数}\ (N_o)} && (4.2)
\end{aligned}
$$

評価

実験の結果，いずれの語義属性においても高い再現率が得られた．このことは，画像処理など他の処理からの制約により，特定の語義の索引を求める際に重要な要素である．人物字幕以外における適合率は，本手法単体としても総合的にも全体に低かった．これは，辞書作成時に多くの不適切な語が混入したためと考えられる．

しかしながら，人物字幕の解析については，再現率，適合率のいずれにおいてもほぼ実用的な解析性能が得られた．また，時相字幕についても，本手法を用いなければ極端に低い解析性能しか得られない事も示されている．

誤答と洩れの理由として，主に表 4.18に示すような 5 通りのものが考えられる．原因(a)及び(e)は作成した辞書の性能を直接反映したものである．しかし，原因(a)のうちの 4 割程度はより細かい形態素に分解されていれば正しく解析できたものである．このような場合への対処法として，現在の完全一致による照合ではなく，最長後方一致による照合を採用すれば良いと思われるが，逆に誤答を増やすおそれもある．原因(d)は語義を扱う際に直面する本質的な問題であり，本手法のように字面のみによる解析手法では容易に解

表 4.15: 字幕解析結果（人物）
N_c：正答数，N_m：誤答数，N_o：洩れの数

	旧分類語彙表による語彙拡張					新分類語彙表による語彙拡張				
	N_c	N_m	N_o	適合率	再現率	N_c	N_m	N_o	適合率	再現率
本手法単体	251	23	—	91.61%	67.11%	260	63	—	80.50%	69.52%
JUMAN 単体	48	54	—	47.06%	12.83%	48	54	—	47.06%	12.83%
総合	299	77	75	79.52%	79.95%	308	117	66	72.47%	82.35%

表 4.16: 字幕解析結果（場所・組織）
N_c：正答数，N_m：誤答数，N_o：洩れの数

	旧分類語彙表による語彙拡張					新分類語彙表による語彙拡張				
	N_c	N_m	N_o	適合率	再現率	N_c	N_m	N_o	適合率	再現率
本手法単体	361	618	—	36.87%	45.24%	427	570	—	42.83%	53.51%
JUMAN 単体	279	13	—	95.55%	34.96%	279	13	—	95.55%	34.96%
総合	640	631	158	50.35%	80.20%	706	583	92	54.77%	88.47%

表 4.17: 字幕解析結果（時相）
N_c：正答数，N_m：誤答数，N_o：洩れの数

	旧分類語彙表による語彙拡張					新分類語彙表による語彙拡張				
	N_c	N_m	N_o	適合率	再現率	N_c	N_m	N_o	適合率	再現率
本手法単体	165	101	—	62.03%	82.50%	165	259	—	38.92%	82.50%
JUMAN 単体	22	0	—	100.00%	11.00%	22	0	—	100.00%	11.00%
総合	187	101	13	64.93%	93.50%	187	259	13	41.93%	93.50%

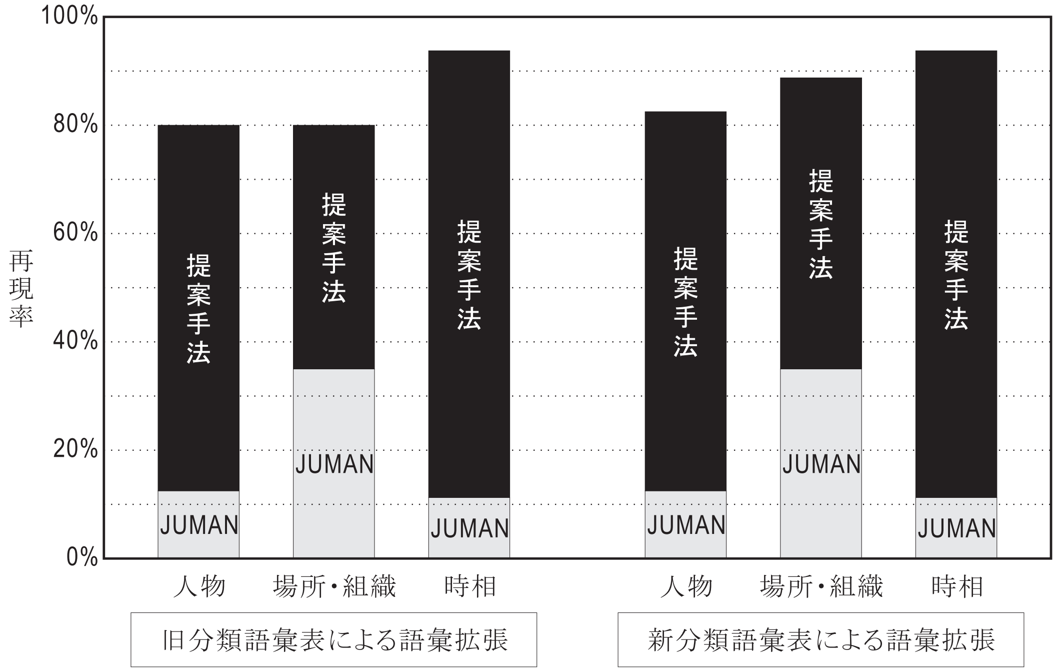


図 4.3: 字幕の語義属性解析性能（再現率）

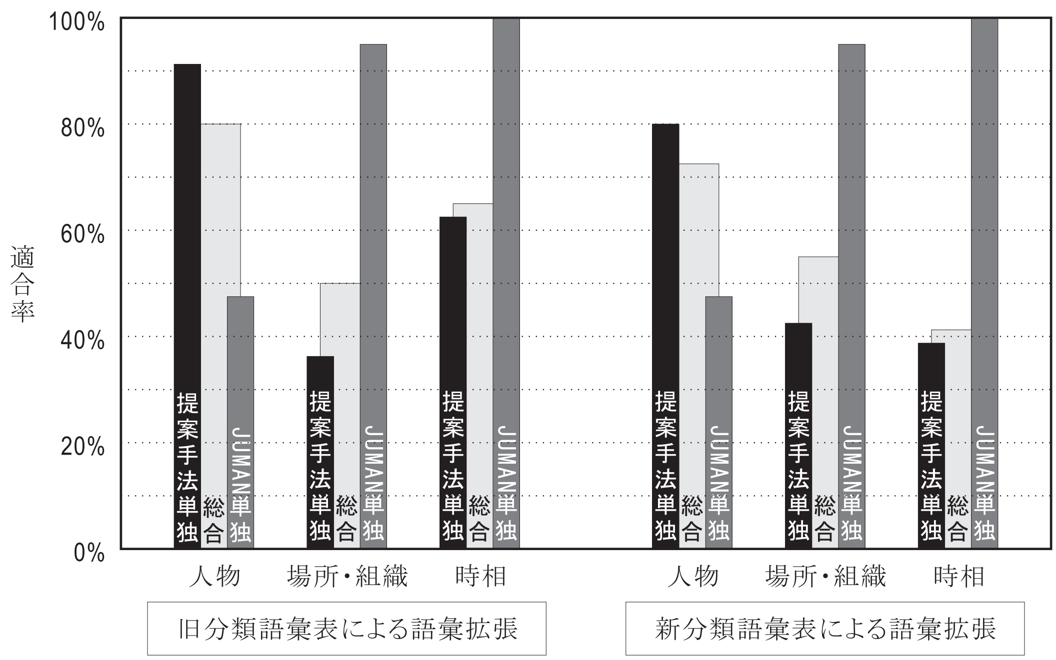


図 4.4: 字幕の語義属性解析性能（適合率）

表 4.18: 字幕の語義属性解析実験の誤答と洩れの原因

	原因	比率
(a)	作成した辞書中の普通名詞の語彙不足	44.0%
(b)	JUMAN 固有名詞辞書・時相辞書の語彙不足	38.8%
(c)	語義の多様性	9.1%
(d)	JUMAN による形態素解析の誤り	5.7%
(e)	作成した辞書への不適切な語の混入	2.4%

決できない．原因（b）と（d）については，JUMAN の語彙及び解析性能の向上をまつしかない．

4.4 まとめ

本章では，字幕の語義属性解析手法と，解析に必要となる語義属性別名詞辞書の作成過程を示した．作成された語義属性辞書の収録語彙数は，人物名詞，場所・組織名詞，時相名詞が旧分類語彙表による語彙拡張の結果，各々，860 語，4,621 語，1,215 語，新分類語彙表による拡張の結果，各々，1,795 語，7,908 語，2,144 語となった．また，これらの辞書を用いて実際のニュース映像中に出現した 2,546 件の字幕に適用して評価実験を行った結果を示した．

第 5 章

画像特徴量に基づく画像内容解析

　本章では，映像内容を考慮した索引付けに必要となる，画像に対する処理として，画像特徴量に基づく画像内容解析手法について述べる．

　まず，本論文中の実験で用いる画像特徴量の性質について述べ，次に，本研究で提案する人物領域の分割による画像内容推定手法を提案する．その後，提案手法の適用例として，実際のニュース映像に対して，ルールベースの典型的ショット分類に基づく行為推定及び，背景領域の画像特徴量解析に基づく事例ベースの場面推定を行った結果を示し，本手法の評価とする．

5.1　画像特徴量

本節では，表 3.3に挙げた様々な画像特徴量のうち，5.3の実験で用いる色彩及び顔に関する特徴量の性質について述べる．

5.1.1　色彩に関する特徴量

ここでは，色出現頻度分布（Color Histogram）と色共起頻度分布（Color Correlogram) の特徴について述べる．

　色出現頻度分布は構造情報をまったく保存しないものの，画像全体の大域的な色調を良く表す特徴量である．一方，色共起頻度分布は近隣の色の共起頻度に基づくため，局所的な色調をやや抽象的な方法で保存する特徴量である．一例として，図 5.1に示すように，着色面積が等しい大円と小円の集合について，前者は弁別できないのに対し，後者では弁別できるという特徴がある．

5.1.2　顔に関する特徴量

顔に関する特徴量は，表 3.3に挙げた様々な画像特徴量のうちの具体的事物の認識に含まれるが，ニュース画像など多くの画像中に顔領域が存在するため，比較的一般性の高い特徴

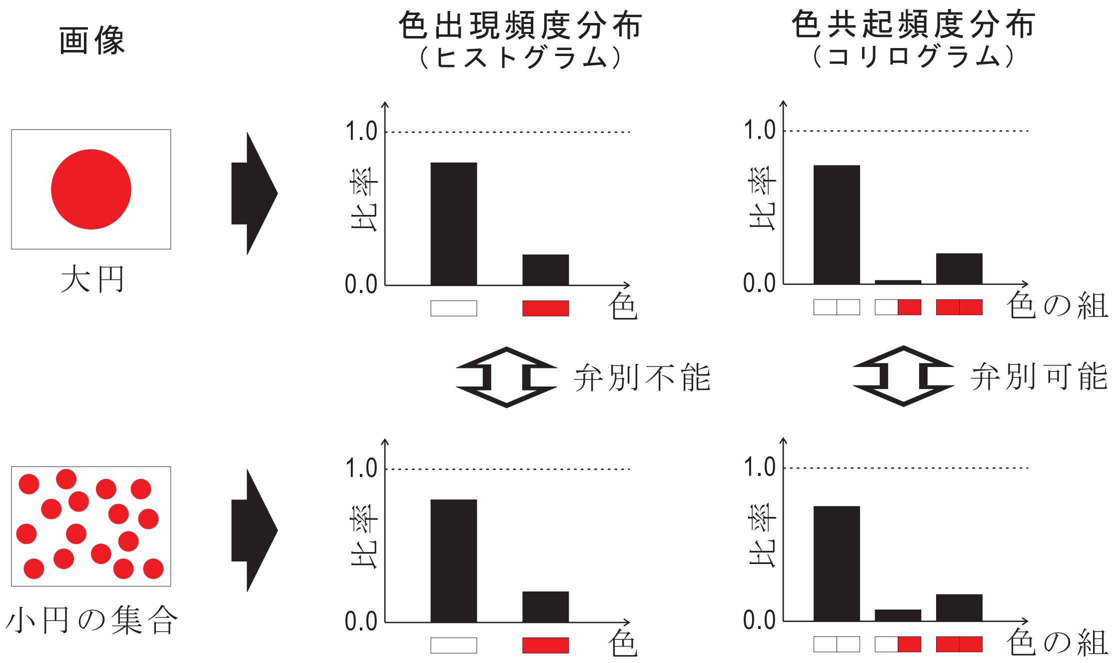


図 5.1: 色頻度分布と色共起頻度分布の弁別能力の差異

量とみなせる．

顔に関する特徴量としては，顔領域の検出による個数，位置，面積などの簡単な特徴から，顔領域内の顔部品の位置及びそれらの相互関係 [185]，固有顔ベクトルによる特徴表現にいたるまで，用途に応じて様々なものが存在する．

このような特徴量を用いた人物識別や同一人物追跡に関する研究は既に多くなされているため [106, 111, 114, 155, 200]，本論文では単純な個数，位置，面積を典型的なショットの分類に用いるにとどめる．

5.2 画像特徴量に基づく画像内容解析

本研究では，画像特徴量から画像内容を解析するために，あらかじめ訓練事例画像から画像特徴量と画像内容との関連に関する知識を獲得しておく．これにより，索引付けを行いたい未知の映像の画像特徴量から画像内容を解析する手がかりとする．本節では，そのような画像特徴と画像内容の関連に関する知識獲得を目指した既存の研究を紹介した後に，本研究で提案する人物領域と背景領域を分離した画像内容解析手法について述べる．

5.2.1 画像特徴量と画像内容の関連に関する知識獲得

画像特徴量と画像内容を描写する索引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 [127] による絵画データベースに関する研究や，木本による感性語を用いた画像検索 [126] などの試みがある．しかし，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．

また，森らによる手法 [137] は，まず語の共起関係により単語クラスタ空間を形成し，次に単語クラスタ空間中の単語間距離に基づき百科事典中の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．しかし，この手法では単語ではなく文との関係を見ているため，獲得された関係を一般的に用いることは難しい．

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を提案している [136]．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，提案手法のように画像の構成内容の推定は行っていない．

この他にも様々な画像分類，内容推定，モデル構築に関する研究があるが [101, 109, 110, 118, 131, 140, 153, 156, 162, 170, 195]，汎用性の高い分類規則，つまりより一般的な画像特徴と画像内容との関係を自動的あるいは統計的に獲得する手法は見当たらない．

5.2.2 領域分割による画像内容の解析

3.3.3で述べたように，ニュース画像中には人物像が大きく映っていることが多いため，人物領域（頭部領域と胴体領域からなる）と背景領域を分離した画像内容解析を行う．図 5.2 に，このような領域分割による画像内容の解析機構を示し，以下に各手順の処理について順に述べる．

人物に着目した領域分割

ショット中の各フレームに対して，人物領域と背景領域の分割を行う．全フレームに対して内容推定を行うのが望ましいが，計算量の点から，ショットを代表する 1 フレームを推定対象とする．ここでは，ショットの冒頭の 1 フレームを代表フレームとして採用した．

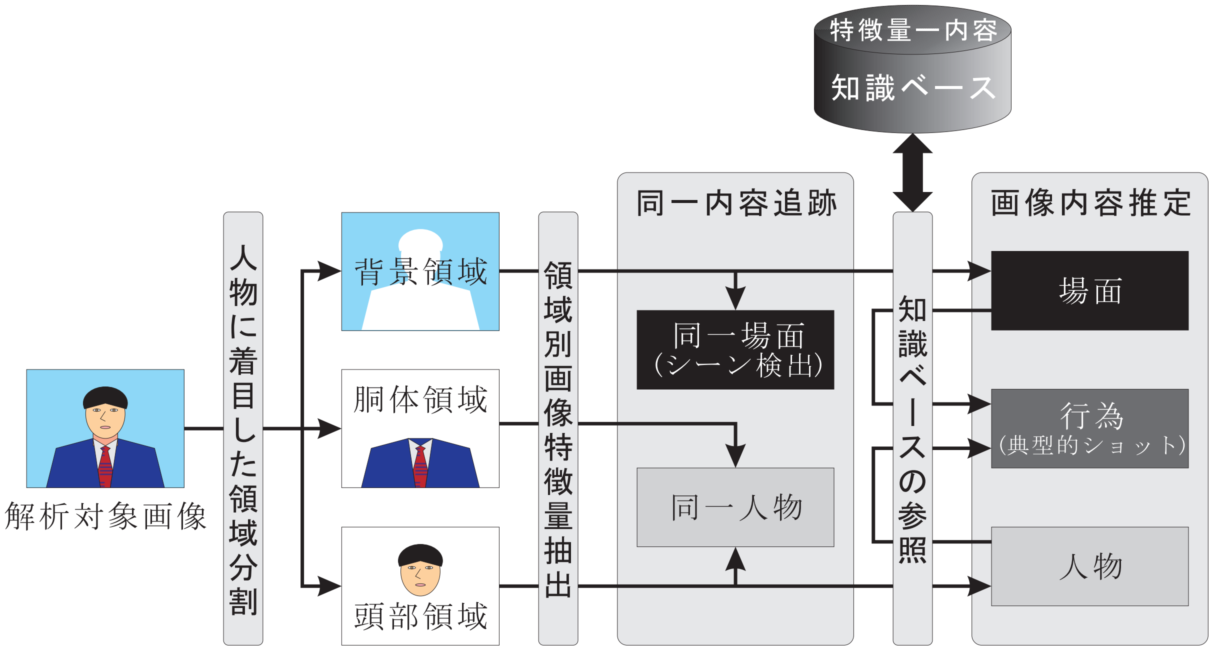


図 5.2: 領域分割による画像内容の解析

ニュース映像の特徴として，人間の社会的行為を扱った内容が多い点が挙げられる．そのため，画像中に人物像，特に上半身が大きく映っていることが非常に多い．このような特徴を考慮して，被写体の人物像と背景の領域を分離して領域別に画像特徴量を抽出し，内容解析を行うことにより，画像中の人物と場面に関する解析を行う．

領域別画像特徴量抽出

次に，分割した領域別に画像特徴量を抽出する．画像特徴量としては，3.3.2で挙げたように様々なものがある．しかし，以下の実験においては，計算量の点から処理が比較的容易であり，かつロバストなものを中心に，具体的事物の認識は行わず，5.1で紹介した色彩と顔に関する特徴量のみを用いた．

同一画像内容追跡

ショットをまたいだ同一画像内容の追跡を行う．このような追跡を行うことにより，特定の人物，場所・組織，行為に関する索引付けを行う時間的範囲を決定する．このような追跡により「存在–行動マップ」[165] のように特定の画像内容の出現区間を把握することが可能になる．

画像中の同一人物の追跡は，他の手法でも行われており [104, 158, 159]，エッジや動きを利用したより高度な手法 [125] も存在するが，ここでは，ニュース映像では良好な照明条件の下に人物の正面から撮影された画像が多いことを利用して，分割した領域毎の特徴量の

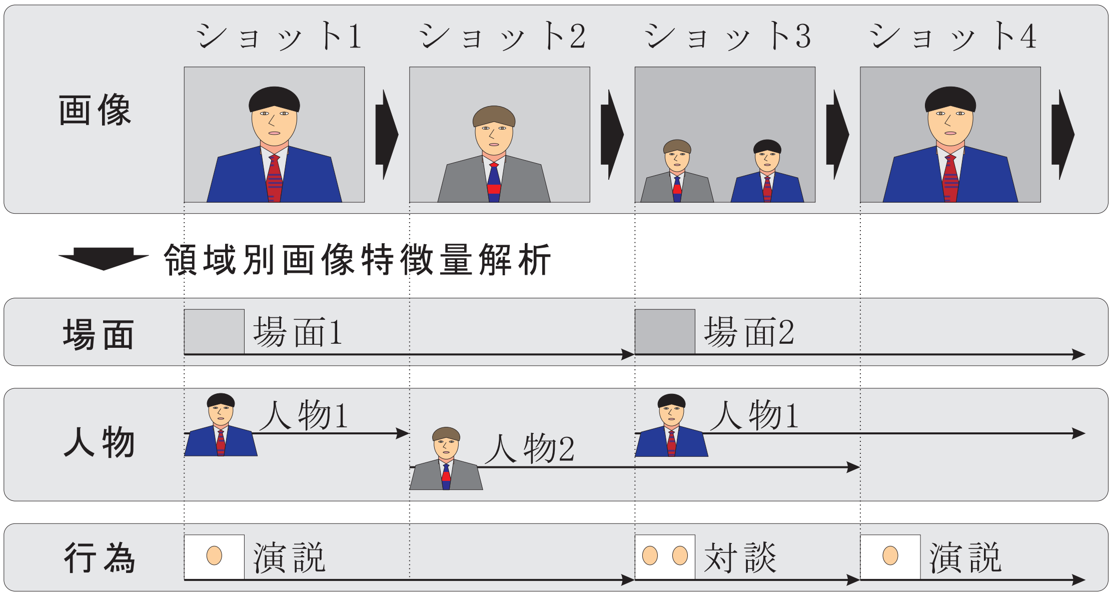


図 5.3: 領域別画像特徴量解析による同一人物・場面追跡

変化により判断する簡易な手法を想定している．図 5.3に同一画像内容追跡の例を示す．

知識ベースの参照による画像内容推定

以上のように領域毎に抽出された特徴量をもとに，表 5.1 に示すように，各領域の画像内容の推定を行う．また，推定された人物と場面に関する画像内容から，典型的ショット分類により，状況や行為に関する画像内容を推定する．以下に，各々の内容推定について述べる．

表 5.1: 領域毎の内容推定

領域	推定内容
頭部領域	人物（Who）
胴体領域	人物（Who）
背景領域	場面（Where）
全領域	状況，行為（What）

人物・場面の内容解析　人物・場面の内容解析は，訓練事例と解析対象事例との画像特徴量の類似度を利用して行う．

特徴量ベクトル同士の類似度を評価する際には，次元や値域の違いなどを考慮して，各特徴量を表現する特徴量ベクトル毎の類似度の重み付き和により総合的な類似度を求めることとする．つまり，2 枚のフレーム F_i, F_jの対応する領域 rにおける特徴量 kの特徴量ベクトルを$\overrightarrow{v_{r,k}(F_i)}, \overrightarrow{v_{r,k}(F_j)}$とし，領域別の総合的な類似度評価関数 $S_r(F_i, F_j)$ を，各特徴量毎の類似度評価関数 $s_r(\vec{x}, \vec{y})$ によって次式のように表す．

$$S_r(F_i, F_j) = \sum_k \alpha_{r,k}(c) s_r(\overrightarrow{v_{r,k}(F_i)}, \overrightarrow{v_{r,k}(F_j)}) \tag{5.1}$$

ここで，$\alpha_{r,k}(c)$ は，類似度評価対象の 2 つ領域 rが同じ属性 c をもつと仮定した場合の特徴量 kに対する重みである．

たとえば，$\alpha_{\text{背景, 色出現頻度}}(\text{スタジオ})$ は，比較対象の 2 つのフレームの背景領域をスタジオであると仮定した際の色出現頻度ベクトル間の類似度に対する重みである．このような重みは，領域の属性毎に各特徴量に対する重要性が異なることを考慮して設定する必要がある．

行為の内容解析：典型的ショット分類 行為の内容解析は，図 5.2に示すように，人物・場面の内容解析の結果を受けて行われる．つまり，人物に関する解析結果（人物の存在，人数など）と場面に関する解析結果（具体的場所，屋内外など）を統合した解析を行う．

これを定式化すると，式 5.1で示した領域別類似度の重み付き和の画像全体の類似度 $S(F_i, F_j)$ として，次式のように表せる．

$$S(F_i, F_j) = \sum_{l \in \text{背景, 頭部, 胴体}} \beta_l(c) S_l(F_i, F_j) \tag{5.2}$$

ここで，$\beta_l(c)$ は，類似度評価対象の 2 つフレームが同じ属性 c をもつと仮定した場合の領域 lに対する重みである．

具体的には，主に登場人物の存在や人数，放送上の表現手段を基準として，以下の 4 通りの典型的なものへのショット分類を行った．これらの分類は，実験に用いた全ショットの 57%を占め，残りはいずれにも分類されなかった．

- 演説ショット（Speech Shot）
  人物が会見を行ったり，レポータが中継地から報告を行っている，画面の中央部に 1 つの顔が比較的大きめに映っているもの
- 集合ショット（Gathering Shot）
  会議や集会など，映像中に複数の人物が存在するもの
- キャスタショット（Anchor Shot）
  スタジオでキャスタが放送原稿を読み上げているもの

- コンピュータグラフィックス（CG）ショット（Computer Graphics Shot）
  CG（コンピュータグラフィックス）で作成されたもので，参考資料的な肖像写真や解説のための説明図など

5.3 実装と評価

本節では，画像的特徴量と映像内容との関係の獲得及び典型的ショット分類の実装について述べ，評価実験の結果を示す．

5.3.1 用いる既存技術

画像のディジタル化

画像のディジタル化のために，ビデオキャプチャボードを利用する．今日では高機能なものが比較的安価に入手できるが，横 320×縦 240 ピクセルの解像度，色標本化レートを 24bit，通常の NTSC 放送方式と同じ 30 フレーム毎秒（fps = frames per second）の条件でディジタル化することを考えると，15 分の番組を保存するのに非圧縮で 5.8GB 程度の記憶媒体が必要になる．この程度の比較的大容量の記憶媒体は存在するものの，依然今日の計算機で気軽に大量に扱えるデータ規模ではない．

そこで実際は，キャプチャの際にハードウェア的，ソフトウェア的に何らかの圧縮を施すことが多い．ところが，フレーム圧縮に一般的に用いられる JPEG 圧縮は通常，HSI 表色系の色相（H, Hue）に強力な空間的不可逆圧縮を施すため，色彩に関する画像特徴を抽出する際に問題が生じる．また，動画像の圧縮に一般的に用いられる MPEG 圧縮によってもブロックノイズが生じるため，同様に様々な画像特徴を抽出する際に問題が生じる．

本研究においては，可能な限り良い特徴量を抽出するために画像圧縮を最低限にとどめ，通常の NTSC 放送方式の半分の 15fps の標本化を行った．また，通常のビデオキャプチャボードではディジタル化の際の JPEG 圧縮は避けられないため，最低限の圧縮率に設定した．そのため，一部の処理を除き，HSI 表色系ではなく，RGB 表色系において処理を行った．その他のディジタル化の際の諸条件を表 5.2に示す．

表 5.2: 画像のディジタル化の際の諸条件

項目	条件
CPU	Intel MMX Pentium II 400MHz
主記憶容量	256MB
ビデオキャプチャボード	MiroVIDEO DC30
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，合計 16,777,216 色）
時間解像度	15 フレーム毎秒
キャプチャ時フレーム圧縮	JPEG
ビデオ圧縮	なし（非圧縮）

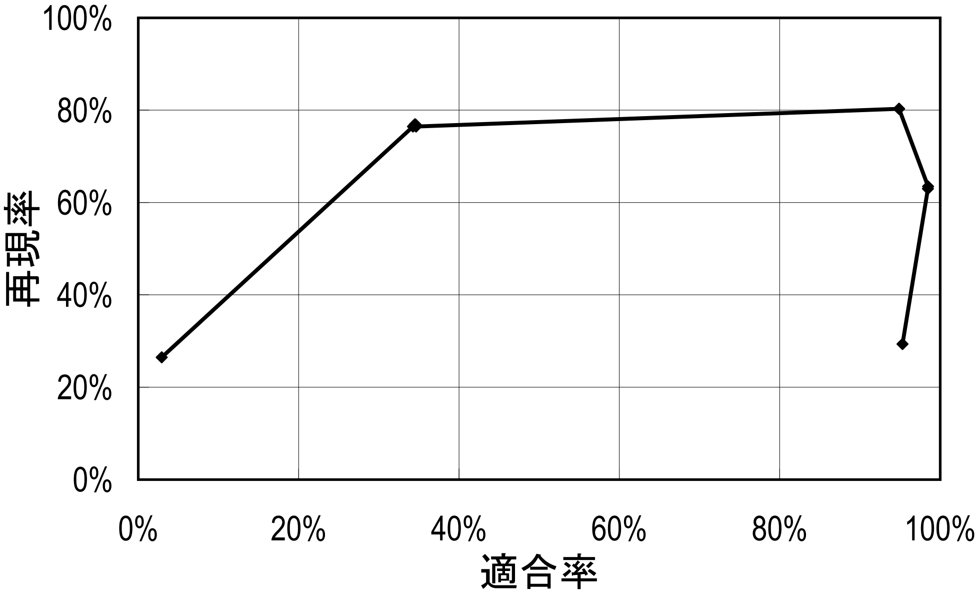


図 5.4: カット検出性能：再現率と適合率の関係

前処理により分割されたショット中の各フレームに対して，人物領域と背景領域の分割を行う．ショット中の全フレームに対して内容推定を行うのが望ましいが，計算量の点から，ショットを代表する 1 フレームを推定対象とする．ここでは，ショットの冒頭の 1 フレームを代表フレームとして採用した．

カット検出

カット検出には，大辻ら [151] や長坂ら [141] により比較検討された種々の手法に始まり，非圧縮画像に対して高精度でカット検出が行える離散余弦変換（DCT）特徴のクラスタリングによる手法 [103] や MPEG 圧縮のデータ形式を利用した圧縮画像に対する手法 [122, 148] まで様々なものが存在する．

最も単純な手法としては，色彩やエッジに関する画像特徴量の変化に注目するものがあり [187]，連続するフレームの画像特徴量が著しく変化する点を検出し，カット点と判断するものである．これらの手法は単純かつ高速なため広く使われていたが，様々な特殊効果を用いたショットの切り替えや，瞬間的に特徴量が大きく変化するフラッシュ光の影響などに対して脆弱である．

一方，DCT 特徴のクラスタリングによる手法は，まず，フレーム全体を分割した各領域に対して DCT を施し，DCT 係数を求める．次に，得られた特徴量で張られる空間にフレーム毎の特徴量ベクトルをプロットしてクラスタを形成させ，一定フレーム連続して分布から逸脱した場合にカットを検出する．

本研究では，この非圧縮画像に対して良い検出精度が報告されている離散余弦変換（DCT）

特徴による手法 [103] を採用する．この手法は，フラッシュ光などによる瞬間的な変化やある種の特殊効果に対する耐性があるとされている．

この手法は，まず，フレーム全体を $x \times y$に分割した各領域に対して DCT を施し，n 個の DCT 係数を求める．次に，得られる合計 xyn 個の特徴量で張られる空間にフレーム毎の特徴量ベクトルをプロットしてクラスタを形成させ，一定フレーム連続して分布から逸脱した場合にカットを検出する．

図 5.4に，30 分間のニュース映像（計 208 カット）に対して，クラスタの大きさを 15 フレーム，逸脱フレーム数を 5 フレームとし，逸脱と判断する閾値を 0.5σから 4.0σの間で変えたときの再現率と適合率との関係を求めた結果を示す．この結果，逸脱と判断する閾値を適切に設定することにより，95%程度の適合率と 80%程度の再現率でカットを検出できることが分かった．

また，ここで得られた最適な閾値 2.0σの条件下で以下の実験に用いた 370 分間のニュース映像（計 1,505 カット）に対して実験を行ったところ，前後 0.5 秒以内の誤差で再現率 84.7%，適合率 66.4%の結果が得られた．

代表フレーム抽出

前処理により分割されたショット中の各フレームに対して，人物領域と背景領域の分割を行う．ショット中の全フレームに対して処理を行うのが望ましいが，計算量の点から，ショットを代表する 1 フレームを推定対象とする．代表ショットを決定するための手法も研究されているが [101, 142]，ここでは簡便のために，ショットの冒頭の 1 フレームを代表フレームとして採用した．

顔領域抽出

顔領域抽出については非常に多くの手法が提案されているが，本研究では以下の手順に示すような比較的簡単な手法で十分であると判断した．

1. 肌色領域抽出
   肌色領域の抽出には，修正 HSI 表色系 [130] を用いた．このうち，I（明度）は暗い領域を排除するためにのみ用いる．実際の様々な人種の顔領域の色を参考に，H（色相）-S_m（修正彩度）空間中の矩形領域を肌色領域と設定した（図 5.5参照）．Hと S_m の値がこの領域に含まれるピクセルは肌色であると判断した後，数ピクセルのクラスタ成長を経て，一定のピクセル数以上のクラスタを顔領域と判断した．

2. テンプレートマッチング
   テンプレートマッチングは，肌色領域として抽出された手，壁，机の表面などを除外するために行う．いくつかの顔領域の Iから複数の解像度の平均顔を作成しておき，抽出された肌色領域の大きさに応じて比較を行う．なお，撮影条件の違いによる影響を避けるために，肌色領域の Iをフレーム全体の Iの値によって正規化しておく．

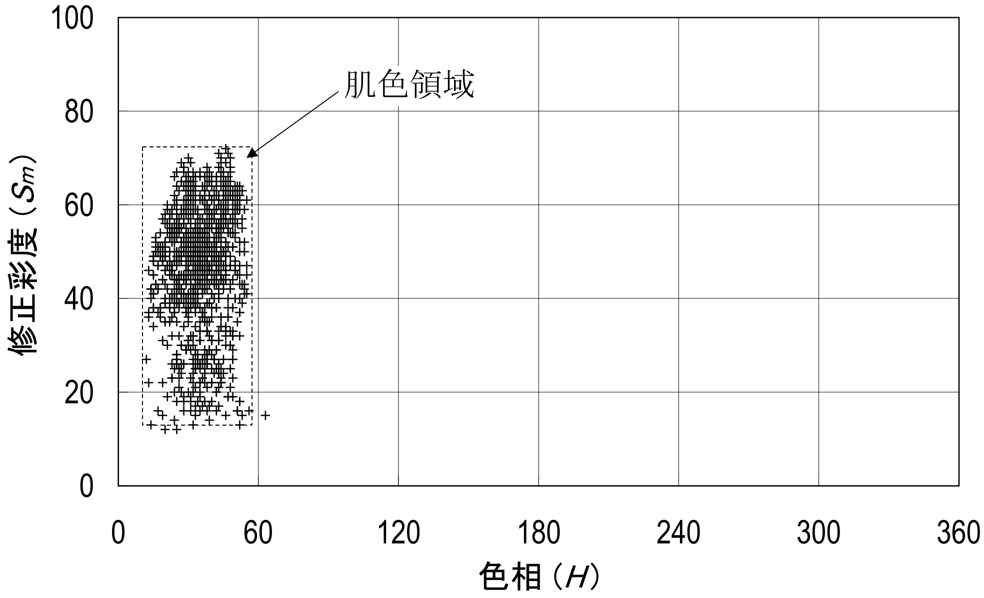


図 5.5: H-S_m空間における肌色領域の分布

一方，試みにニューラルネットワークによる学習を用いたツール：face detector [211] を用いたところ，良好な照明条件下の正面顔という条件を満たすスタジオ内のキャスタの画像 173 件に対して，100%の精度で過不足なく顔領域の抽出に成功した．しかし，他のより一般的な画像中の人物に関しては，必ずしも照明条件が良好でないことや正面顔でなかったために，多くの検出洩れや誤検出が見られた．

5.3.2 実験 5.1：典型的ショット分類

ここでは，75 分間の画像（計 468 ショット）に対して，以下の 4 通りの典型的ショット分類への分類を行った．

1. 演説ショット（Speech Shot）
2. キャスタショット（Anchor Shot）
3. 集合ショット（Gathering Shot）
4. コンピュータグラフィックス（CG）ショット（Computer Graphics Shot）

図 5.6に各ショット分類に属する画像の例を示すとともに，以下に分類規則の詳細について述べる．

演説ショット（Speech Shot）

演説ショットは以下の 2 段階の処理により認識する．

1. 顔領域の検出
   5.3.1に記した手法で顔領域を抽出する．
2. 唇の動きの検出
   資料画像の肖像写真や単に画面の中央に人物が存在するだけのショットを除くために必要な処理である．会見・報告ショットやキャスタショットの場合，映っている顔は良好な照明条件下で撮影された正面顔であることが多く，顔領域が検出されれば，唇の位置を推定することは比較的容易である．
   推定された唇の位置の周囲での動きが顔領域の他の位置よりも大きければ，唇の動きがあるものと判断する．
   なお，本実験でこの処理は，この基準に従って目視により行ったが，将来的には自動化できるものと考えている．

キャスタショット（Anchor Shot）

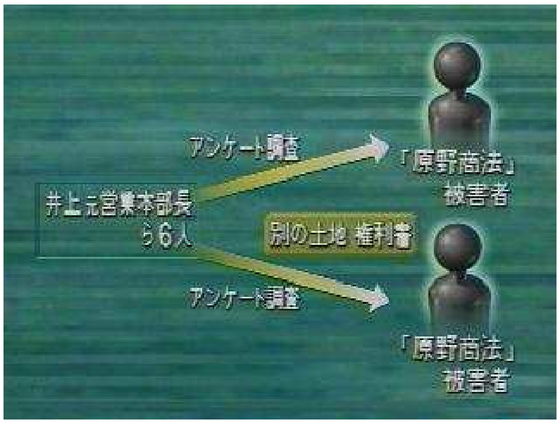
キャスタショットは，スタジオでキャスタが放送原稿を読み上げているショットであり，ニュース映像の内容的構成を把握するうえで重要な手がかりになる．

- 演説ショットからの分離
  キャスタショットの検出は基本的に会見・報告ショットと同様に行うため，これらから分離する必要がある．キャスタショットの特色として，スタジオ内で同じような構図で撮影されるため，相互に画像的に非常に類似した特徴を示し，かつ出現頻度が高いことがあげられる．これらの特色を利用して，会見・報告ショットの検出後，それ

演説ショット

集合ショット

キャスタショット



ＣＧショット

図 5.6: 各ショット分類に属する画像の例

ら相互間の画像的類似度を求め，その結果得られた最大かつ最密なクラスタに含まれるショットをキャスタショットとして分離する．なお，画像的類似度は，色ヒストグラムの差異をχ^2検定して求めた．

- 話題境界の検出
  キャスタショットを検出する最大の目的は，話題の境界を検出することにある．しかし，話題の途中に出現することもあり，単純に出現と境界を同一視することはできない．そこで，話題の境界のキャスタショットの特色である，タイトル字幕の出現を利用する．図 5.7に示すように，タイトル字幕の出現前後で画像全体のエッジ強度が大きく変化する．そこで，エッジ強度に大きな変化があるキャスタショットを話題の先頭と判断する．
  なお，このような特徴は必ずしも欧米のニュース映像では見られず，そのような場合

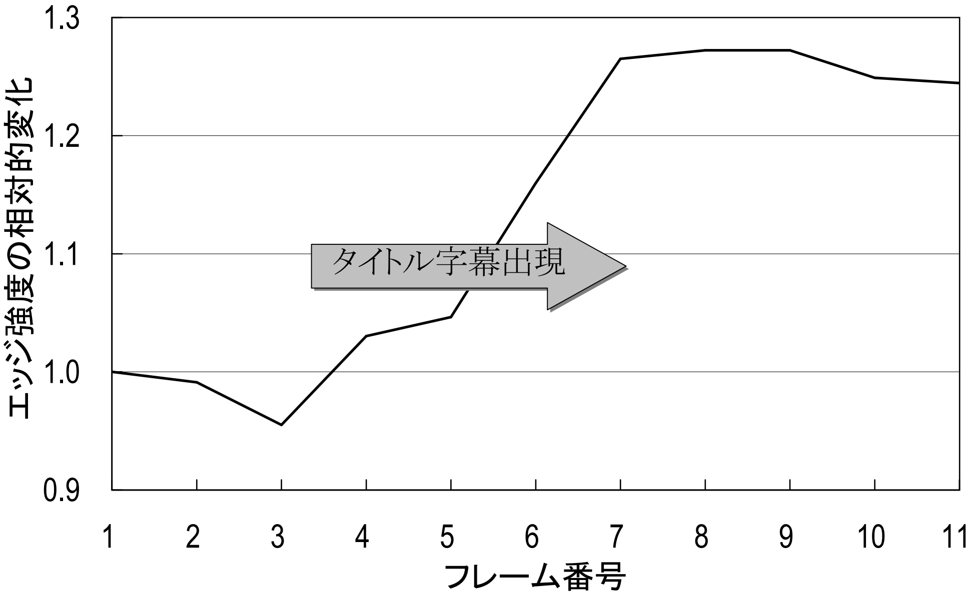


図 5.7: タイトル字幕の出現に伴うエッジ強度の相対的変化

には，話題境界の検出に関する他の手法 [133] を採用する必要がある．

集合ショット（Gathering Shot）

集合ショットは，会議や集会など，映像中に複数の人物が存在するショットである．一定の大きさ以上の複数の顔領域が存在する場合に集合ショットを検出する．

CG ショット（Computer Graphics Shot）

CG ショットは，コンピュータグラフィックス（CG）で作成されたショットであり，参考資料的な肖像写真や解説のための説明図など，他のショット分類や字幕解析に悪影響を及ぼしかねない内容を含む．そこで，索引付け機構から除去するために分類を行う．CG ショットの特色としては，比較的動きが少ない点があげられ，ショット中の静止時間の合計に基づき検出する．本実験では，この時間を 1 秒に設定した．

5.3.3 実験 5.2：背景領域の画像特徴量からの場面推定

ここでは，300 分間の画像の一部（計 817 ショット）に対して（1）閣議控室（2）国会議場（3）報道官会見（4）法廷（5）スタジオの 5 通りの国内報道に関係する場面を設定し，場面推定を行った．各場面毎の事例数を表 5.3に示す．各事例の分類は評価のために事前に目視により行い，各設定場面毎に，人物領域が（1）存在する場合（人物あり）（2）存在しない場合（人物なし）（3）両者を合わせた場合（総合），の 3 通りに分けて評価を行った．

表 5.3: 設定場面毎の事例数

	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
（6）	その他	124	332	456
	合計	425	392	817

図 5.9，5.10に，各設定場面に属する事例の例を示す．

人物に着目した領域分割

一般にニュース映像においては，話題中で主題となっている人物は良好な照明条件下で正面から撮影されていることが多い．5.3.1で紹介したように，輪郭や動きを利用した領域分割手法も存在し [112, 123, 129, 189]，本来は正確な輪郭抽出を行うべきだが，ここではこのような特徴を利用して，顔領域の位置と大きさを基準としたテンプレートにより頭部領域及び胴体領域を決定する．図 5.8に実測に基づいて定義したテンプレートを示す．

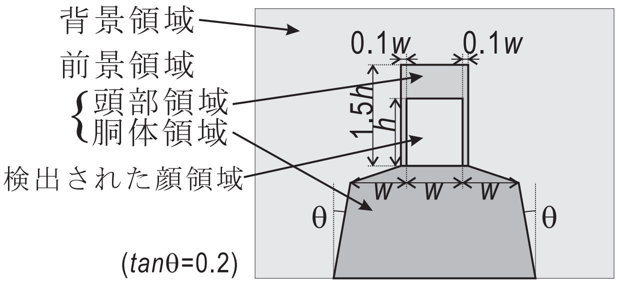


図 5.8: 顔領域を基準とした頭部領域と胴体領域の決定

閣議控室 （人物あり 1）

国会議場 （人物あり）

閣議控室 （人物あり 2）

国会議場 （人物なし 1）

閣議控室 （人物なし）

国会議場 （人物なし 2）

図 5.9: 設定場面に属する事例（1）

報道会見場 （人物あり）

報道会見場 （人物なし）

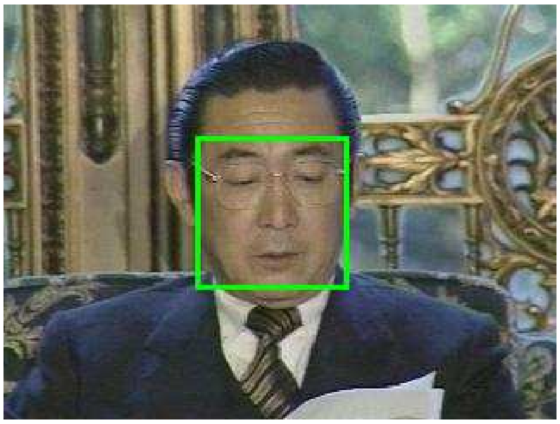
法廷 （人物あり）

法廷 （人物なし）

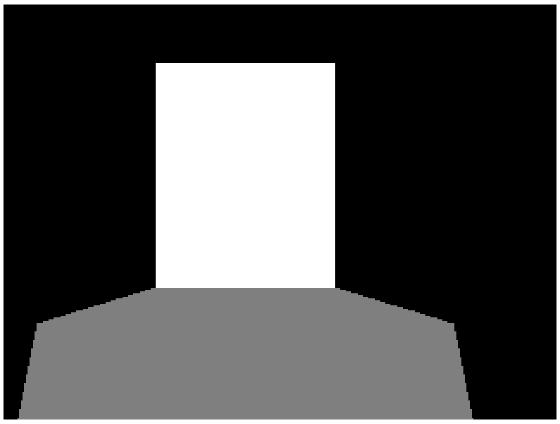
スタジオ （人物あり 1）

スタジオ （人物あり 2）

図 5.10: 設定場面に属する事例（2）



元画像から抽出した顔領域



顔領域を基準に分割した各領域

図 5.11: 顔領域を基準とした頭部・胴体・背景領域決定の例

このようなテンプレートを用いることにより，図 5.11（左）に示すように顔領域が検出された画像から，同図（右）に示すような領域分割が行われる．

ただし，本実験では，場面解析手法単独での評価を行うため，スタジオ内のキャスタ以外は，目視により正しい領域を抽出したものを実験に用いた．

背景領域の特徴量抽出

本実験では，5.3.1で紹介した（1）色出現頻度分布（ヒストグラム）（2）色共起頻度分布（コリログラム）の 2 種類の色彩に関する画像特徴量を各々別に用いて評価を行った．

以下に各々の定義を行う．ここで，特徴量を抽出する対象画像を F，縦横のピクセル数を各々m, n，各ピクセルの色の階調を c_{max}とする．また，F中のピクセル p のとる色を $c(p)$ と表す．

色出現頻度分布（Color Histogram）　色頻度分布 $\{H(c_i) \mid i = 1, 2, ..., c_{max}\}$ とは，画像 F 中において，色 c_iのピクセルが存在する確率であり，次式のように定義され，c_{max}次元で表される．

$$H(c_i) \equiv Pr\{(x, y) \mid p \in F, c(p) = c_i\} = \frac{\mid \{p \mid p \in F, c(p) = c_i\} \mid}{mn}$$

色共起頻度分布（Color Correlogram）　色共起頻度分布 $\{C_d(c_{i,j}) \mid i, j = 1, 2, ..., c_{max}\}$ とは，画像 F中において，距離 d 離れた 2 つのピクセル p_a, p_b間で色 c_i, c_j が共起する確率であり，次式のように定義され，c_{max}^2次元で表される [117]．

$$C_d(c_{i,j}) \equiv Pr\{(p_a, p_b) \mid p_a, p_b \in F, c(p_a) = c_i, c(p_b) = c_j, \|p_a - p_b\| = d\}$$

$$= \frac{|\{(p_a,p_b) \mid p_a,p_b \in F, c(p_a)=c_i, c(p_b)=c_j, \|p_a-p_b\|=d\}|}{H(c_i)N(d)}$$

ここで $N(d)$ は，あるピクセルから距離 d 離れて存在するピクセルの個数である．簡便のために 8 近傍（チェス盤）距離を用いる場合は，$N(d)=8d$ である．

表 5.4に実験に用いた具体的な次元数などを示す．

表 5.4: 実験に用いた画像特徴量

特 徴 量	色 階 調	共起距離	次 元 数
色出現頻度分布（Color Histogram）	64	—	64
色共起頻度分布（Color Correlogram）	16	1 ~ 4	1,024

事例に基づく場面推定

以上のようにして，各ショット毎の背景領域の画像特徴量が得られ，次にこれをもとにして，場面推定を行う．場面推定は，推定を行おうとするデータと知識ベース中の知識との類似度に基づいて行う．具体的には，特徴量ベクトル間の類似度，すなわち，場面推定を行いたい画像の画像特徴量ベクトル$\vec{F}_e$と，訓練事例の画像特徴量$\vec{F}_t$との類似度を求める．類似度としては，次式で定義されるベクトル同士のなす角度θの余弦（$0 \leq \cos\theta \leq 1$）を用いる．

$$\vec{F}_e\text{と}\vec{F}_t\text{の類似度} \equiv \cos\theta = \frac{\vec{F}_e \cdot \vec{F}_t}{|\vec{F}_e||\vec{F}_t|} \tag{5.3}$$

知識として，訓練事例の特徴を統計的処理により抽象化した代表ベクトルを用いることも考えられるが，同一内容の訓練事例が特徴量空間中で必ずしも稠密な分布をするとは限らない．そこで，本実験では，訓練事例を抽象化せずに全て保存しておき，それら一つ一つとの類似度から判定する，事例に基づく推定を行った．

図 5.12に示すように，人物領域が存在する場合は背景領域を，存在しない場合は全領域の画像特徴量を用いた．実験としては，領域分割の効果を検証するために，前者について分割を（1）行わなかった場合（分割なし）と（2）行った場合（分割あり）を比較した．また，分割を行った場合でも，人物の存在の有無を含めた場面の推定という観点から（3）分割を行ったもの同士に限定した類似度評価（分割のみ）も併せて行った．

なお，本来は訓練事例と評価事例を分離した実験を行うべきであるが，多様な評価事例に対して一定規模の場面推定を行うためには，膨大な事例数が必要となる．そのため，本実験では，各事例を評価事例，残りの全事例を訓練事例とみなして，相互に評価した結果をまとめて統計処理した．

図 5.12: 場面推定実験の類似度評価条件

5.3.4 実験結果と評価

実験 5.1：典型的ショット分類

表 5.5及び図 5.13に，典型的ショットへの分類結果を示す．なお，評価に用いた正解は人間が判断した．

表 5.5: ショット分類実験の結果

ショット分類	正検出	誤検出	検出洩れ	再現率	適合率
演説	55	3	12	83%	95%
キャスタ	53	0	0	100%	100%
〃 （先頭）	37	0	6	86%	100%
集合	70	13	62	53%	84%
CG	14	1	0	100%	93%
総合	229	17	80	74%	93%

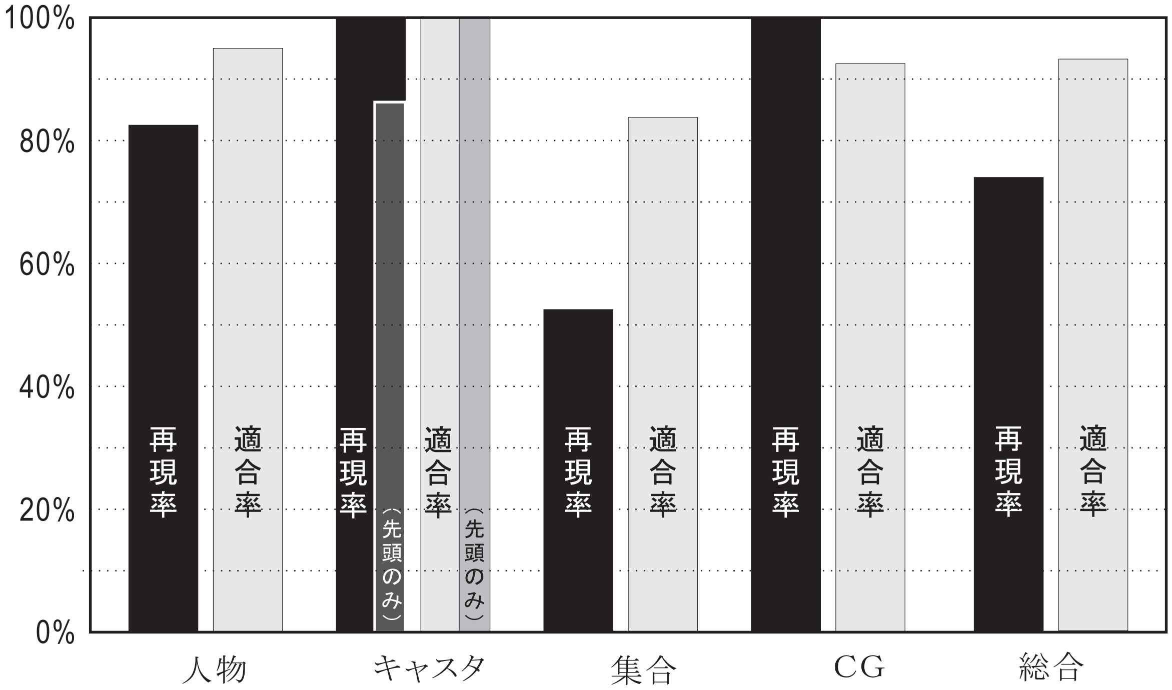


図 5.13: ショット分類性能

誤検出と検出洩れの主な原因は以下の通りであった．

- 演説ショットで，顔領域が正面顔でなかったため唇の動きが検出できなかった．
- 非常に短い話題の先頭のキャスタショットにタイトルが現れなかった．
- 集合ショットで，背後から撮影されていたために顔領域が検出できなかった．
- 静止物体を撮影した映像を CG ショットと誤った．

字幕の語義属性解析と同様に，誤検出による不適切なショットの混入は，対応付けの際に対応する索引候補が存在しない可能性が高く，許容できる．しかし，検出洩れは極力減らす必要があり，今後，顔領域検出を中心とした要素技術の進歩が望まれる．

実験 5.2：背景領域の画像特徴量からの場面推定

結果を以下の 2 つの指標を用いて評価した．

- 類似度上位 n 位正答率（$n = 1, 2, ..., 10$）
  各設定場面に属するべき全事例について，類似度上位 n 位中に正答が含まれている割合．
  上位の正答率が高いほど，設定場面に属する事例間の類似性が強い（特徴量空間中の稠密度が高い）ことを示す．

- 場面推定性能：類似度上位 n 位過半率（$n = 1, 3, 5, 7, 9$）
  各設定場面に属するべき全事例について，類似度上位 n 位中に正答が過半数を占めている割合．結果的に $n = 1$ の際の値は，最も類似している場面が正答であった事例の割合を示す．
  実際の場面推定に際しては，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面を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．

前者の結果を，図 5.16から図 5.41にかけて，後者の結果を，図 5.42から図 5.67にかけて，各々人物領域が（1）存在しない事例（人物なし），（2）存在する事例（人物あり），（3）両事例を合わせたもの（総合）及び，色彩に関する画像特徴量として（a）色ヒストグラム，（b）色コリログラムを用いた場合の組合せ別に示す．

これらの結果に見られる傾向を表 5.6（類似度上位正答率でみた場合），及び表 5.7（場面推定性能でみた場合）に一覧にして示し，以下にまとめる．

- 各場面に対する画像特徴量の適合性
  推定対象となる場面によって適する画像特徴量が異なる．全般に色コリログラム（表中「C」と記載）が色ヒストグラム（表中「H」と記載）よりも優れた性能を示した．しかし，頻出場面でも事例間の変異が比較的大きい，人物領域が存在する国会議場の背景領域や，人物領域が存在しない報道会見場の全領域においては，逆の傾向を示した．これは，このような場面においては，より大雑把な特徴量である色ヒストグラムが雑音耐性において優位であることを示している．

- 人物領域分割の効果
  概ね，人物に着目した領域分割による場面推定の効果が見られたが，分割しない方が良い場合もあった．この原因は主に，人物領域の存在を画像特徴量として考慮していないことによると思われる．
  そこで，人物領域の存在そのものも画像特徴量とみなし，人物領域が存在する評価事例に対して，存在しない訓練事例を解から省いた場合（分割のみ）には，より顕著に領域分割の効果が現れている．

また，各場面について最も良い条件による場面推定性能を表 5.14及び表 5.15に示す．前者では類似度が最上位のものを，後者では類似度上位 3 事例中の過半数（2 事例）を占めるものを場面として推定した．

表 5.6: 実験結果に見られる傾向（類似度上位正答率で見た場合）
色ヒストグラム（H）と色コリログラム（C）を比較．
（○：効果あり，△：判別不能，×：効果なし）

傾向	閣議控室						国会議場						報道会見場						法廷						スタジオ	
	人物 有		人物 無		全体		人物 有		人物 無		全体		人物 有		人物 無		全体		人物 有		人物 無		全体		人物有	
	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C
色特徴量の優位性		C		C		C	H			C	H		△		H		H			C		C		C		C
分割：なし＜あり	○	○	○	○	○	○	△	×	×	○	△	×	×	×	×	×	×	×	○	×	△	○	△	△	○	○
分割：なし＜のみ	○	○	−	−	○	○	○	○	−	−	○	○	○	△	−	−	○	△	○	×	−	−	○	×	○	○

表 5.7: 実験結果に見られる傾向（場面推定性能で見た場合）
色ヒストグラム（H）と色コリログラム（C）を比較．
（○：効果あり，△：判別不能，×：効果なし）

傾向	閣議控室						国会議場						報道会見場						法廷						スタジオ	
	人物 有		人物 無		全体		人物 有		人物 無		全体		人物 有		人物 無		全体		人物 有		人物 無		全体		人物有	
	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C	H	C
色特徴量の優位性		C		C		C	H			C	H		△		H		△			C		C		C		C
分割：なし＜あり	○	○	△	○	○	○	×	△	×	○	×	△	△	△	×	×	△	○	○	×	○	○	○	○	○	○
分割：なし＜のみ	○	○	−	−	○	○	○	○	−	−	○	○	○	○	−	−	○	○	△	×	−	−	△	×	○	○

これらの表に，訓練事例として用いた事例数を併せて対数尺で示した．推定性能は必ずしも実用的水準には達していないが，事例数の対数とほぼ比例していることから，本実験では正しく推定するに足る事例数を供給できていないためと考えられる．事例数を指数規模で増やし，様々な状況に対応し得るだけの量を供給できれば，推定性能が向上するものと期待される．

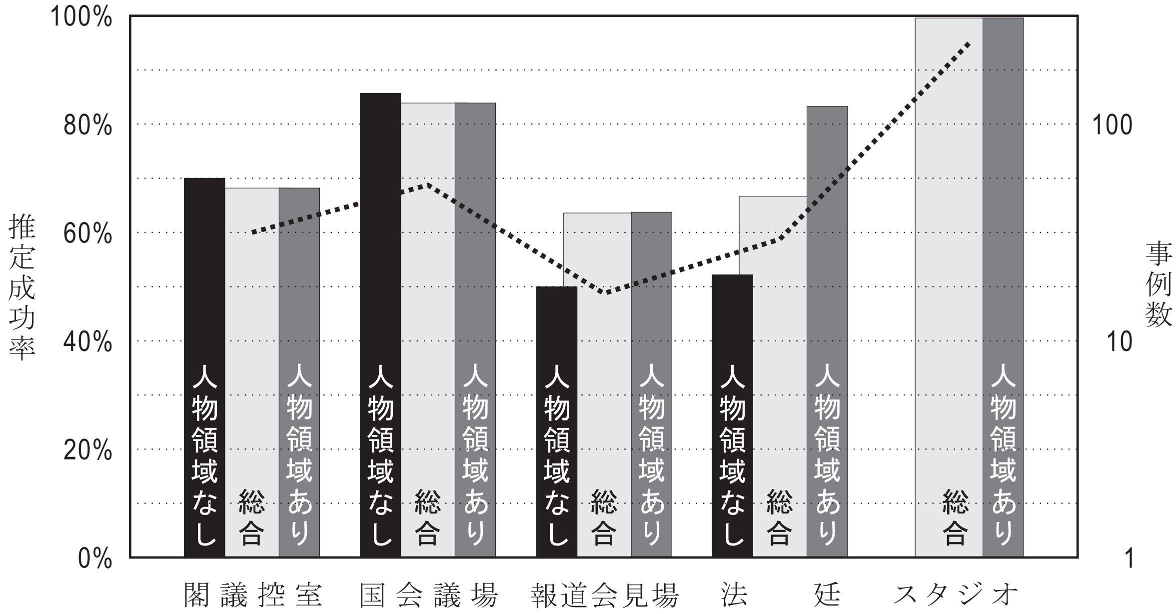


図 5.14: 場面推定性能と事例の規模 (類似度最上位の正答率)
棒グラフが性能を百分率で，線グラフが規模を対数尺で示す．

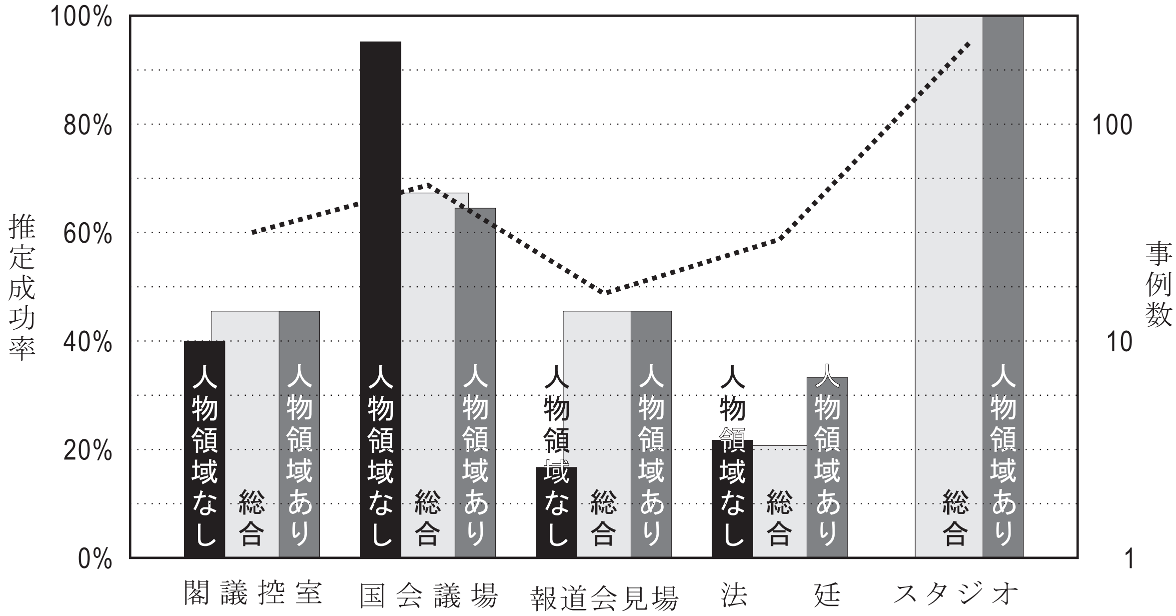


図 5.15: 場面推定性能と事例の規模 (類似度上位 3 事例中過半数の正答率)
棒グラフが性能を百分率で，線グラフが規模を対数尺で示す．

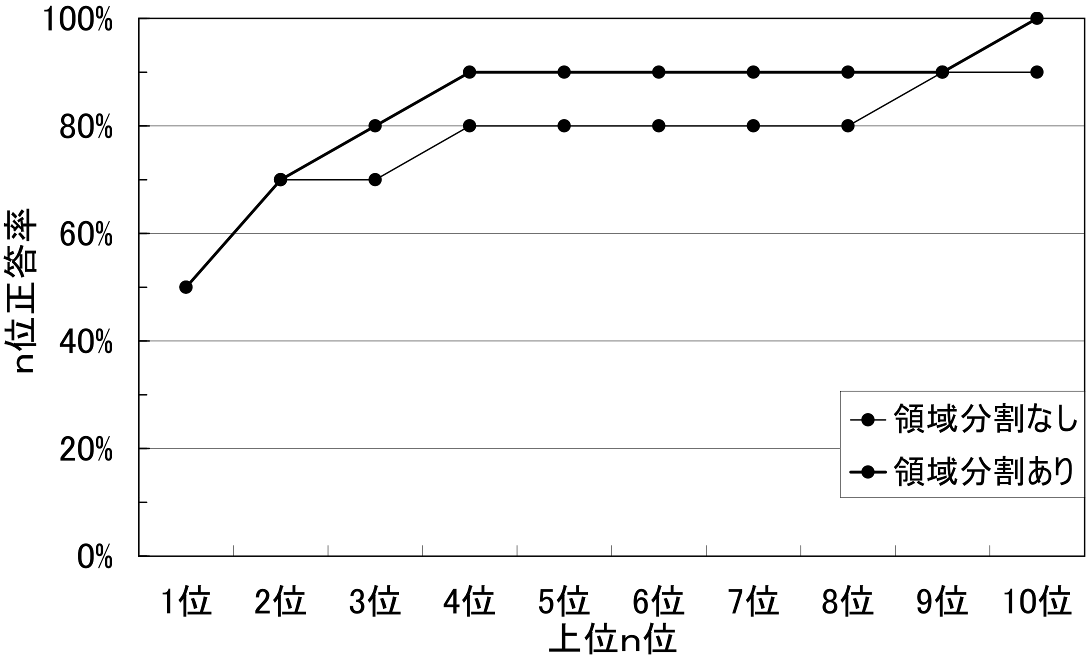


図 5.16: 類似度上位 n 位正答率（闇議控室：人物なし，ヒストグラム）

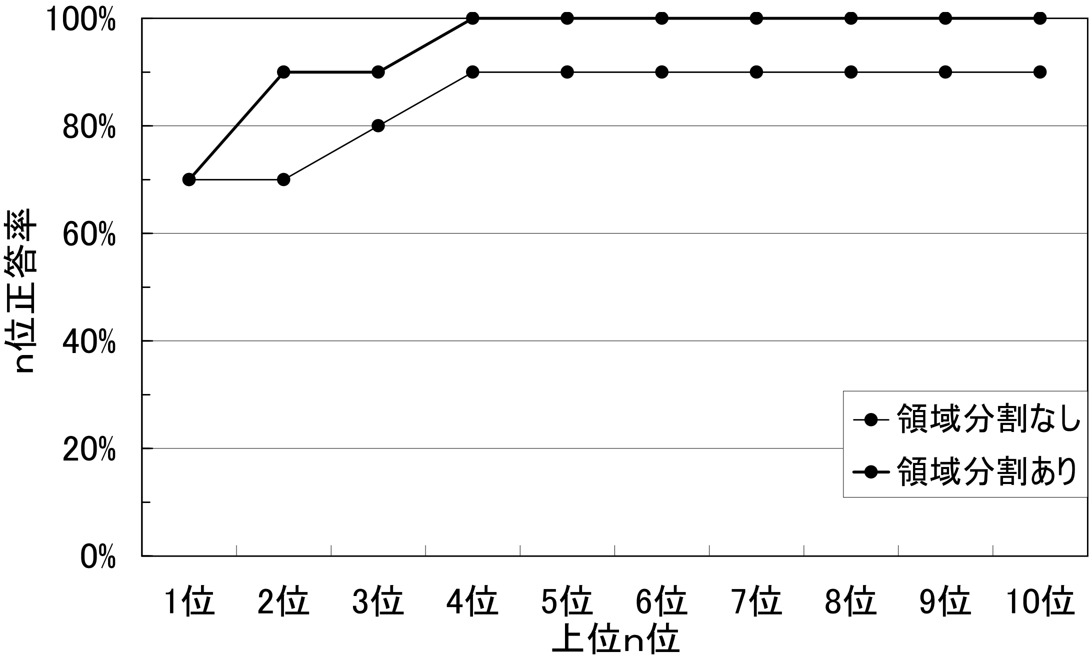


図 5.17: 類似度上位 n 位正答率（闇議控室：人物なし，コリログラム）

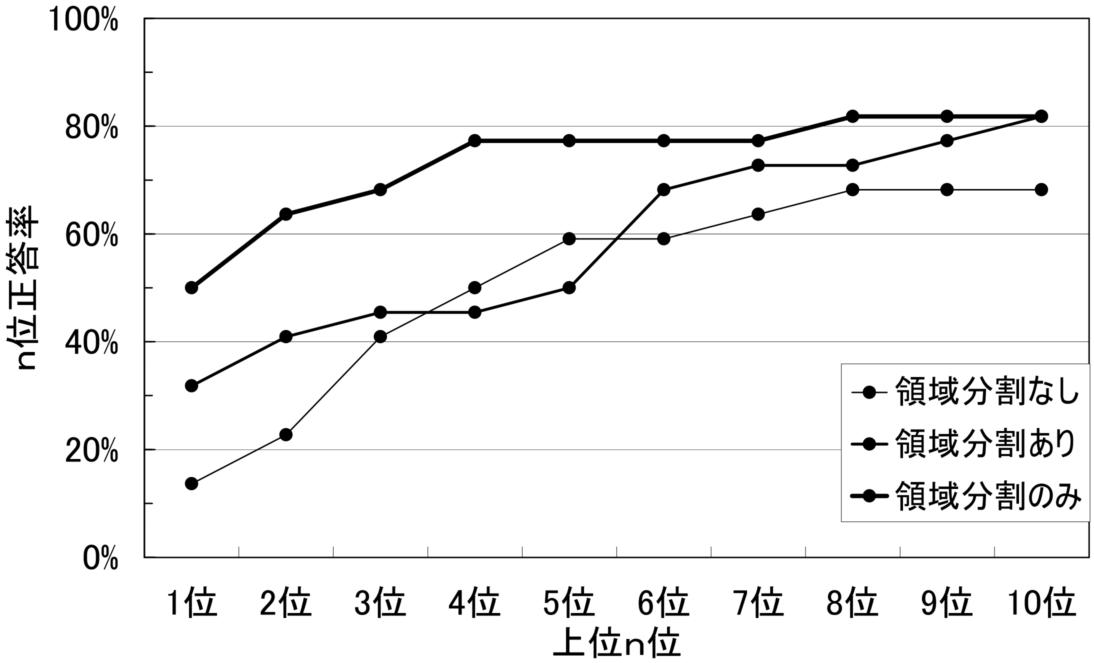


図 5.18: 類似度上位 n 位正答率（閣議控室：人物あり，ヒストグラム）

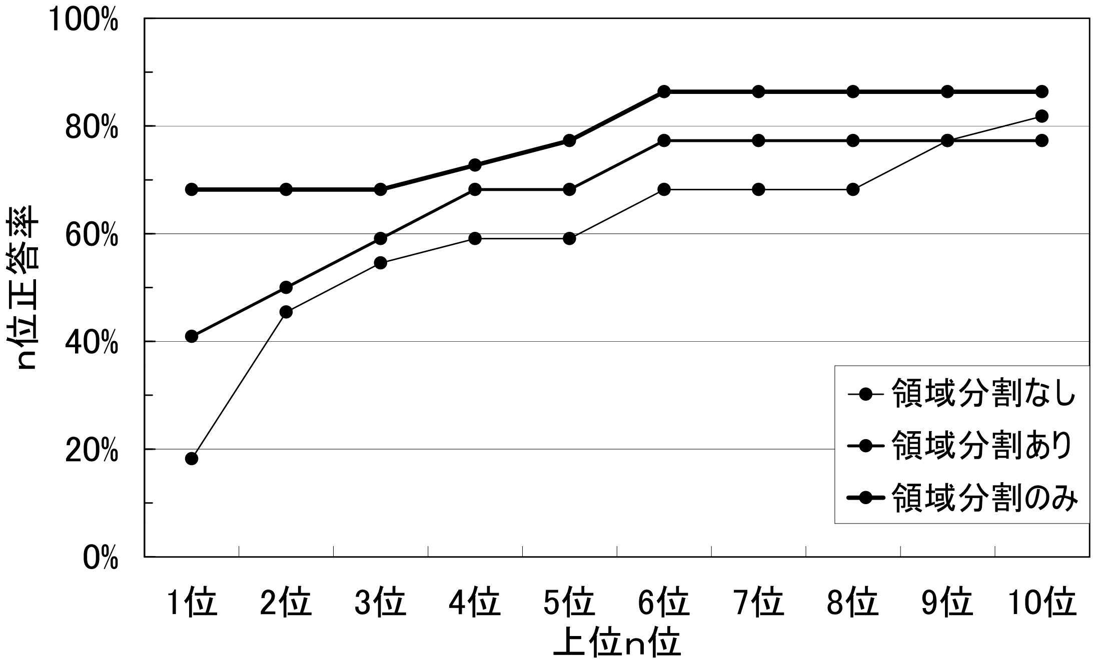


図 5.19: 類似度上位 n 位正答率（閣議控室：人物あり，コリログラム）

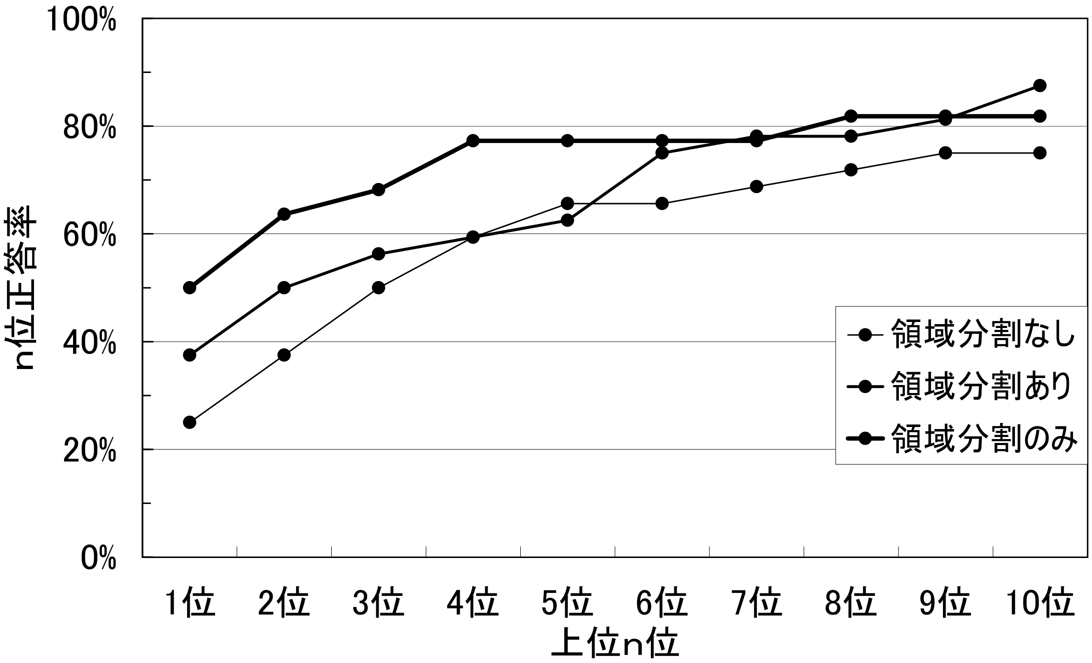


図 5.20: 類似度上位 n 位正答率（閣議控室：総合，ヒストグラム）

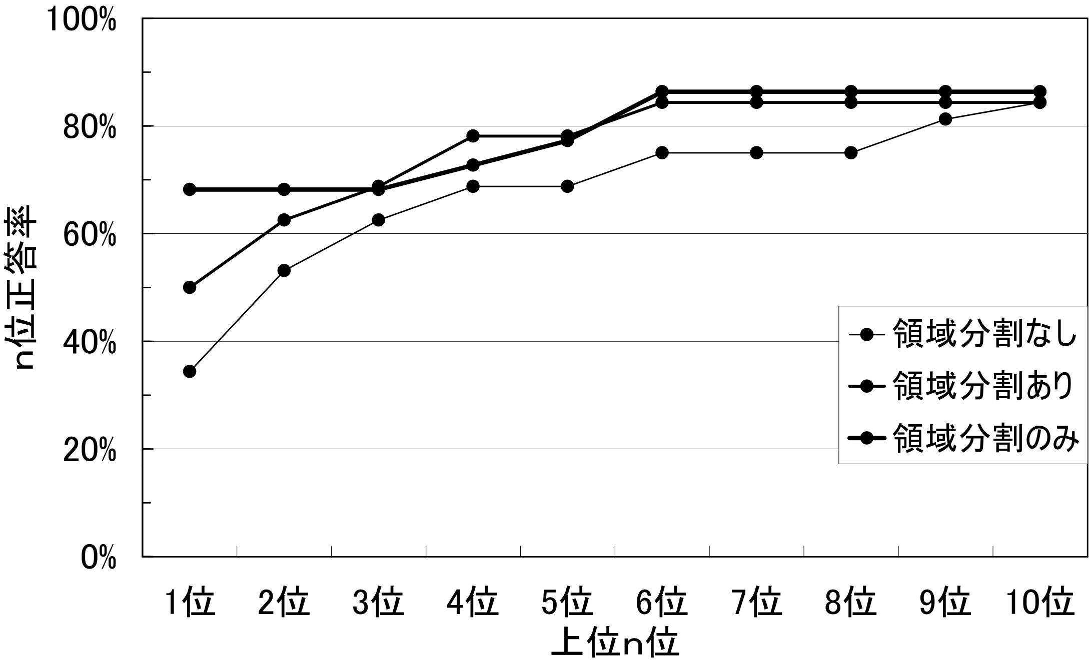


図 5.21: 類似度上位 n 位正答率（閣議控室：総合，コリログラム）

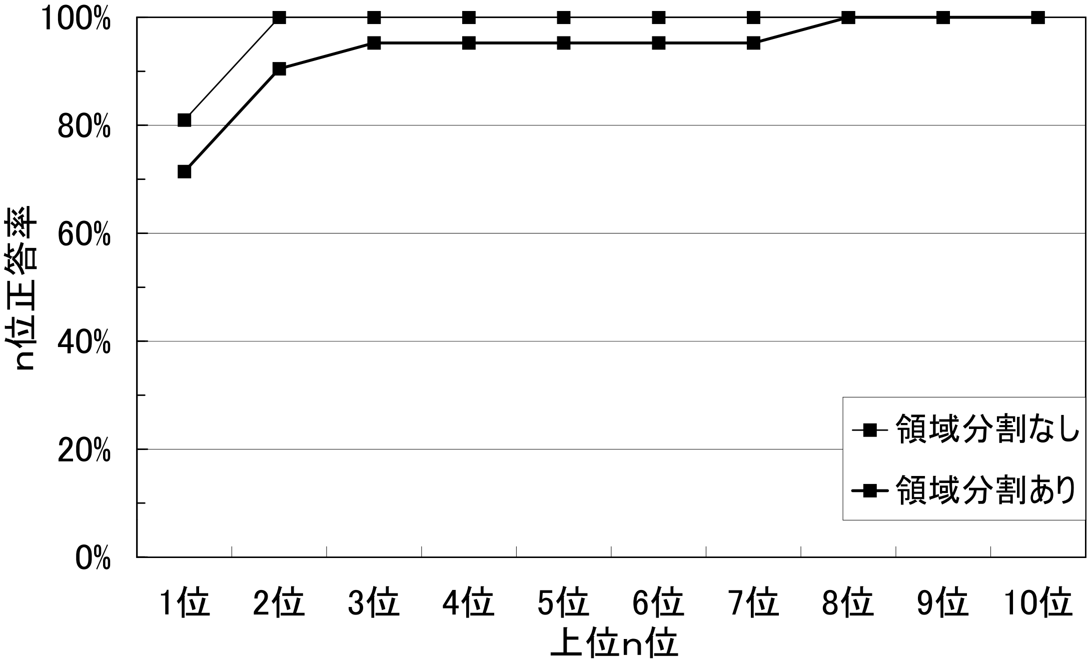


図 5.22: 類似度上位 n 位正答率（国会議場：人物なし，ヒストグラム）

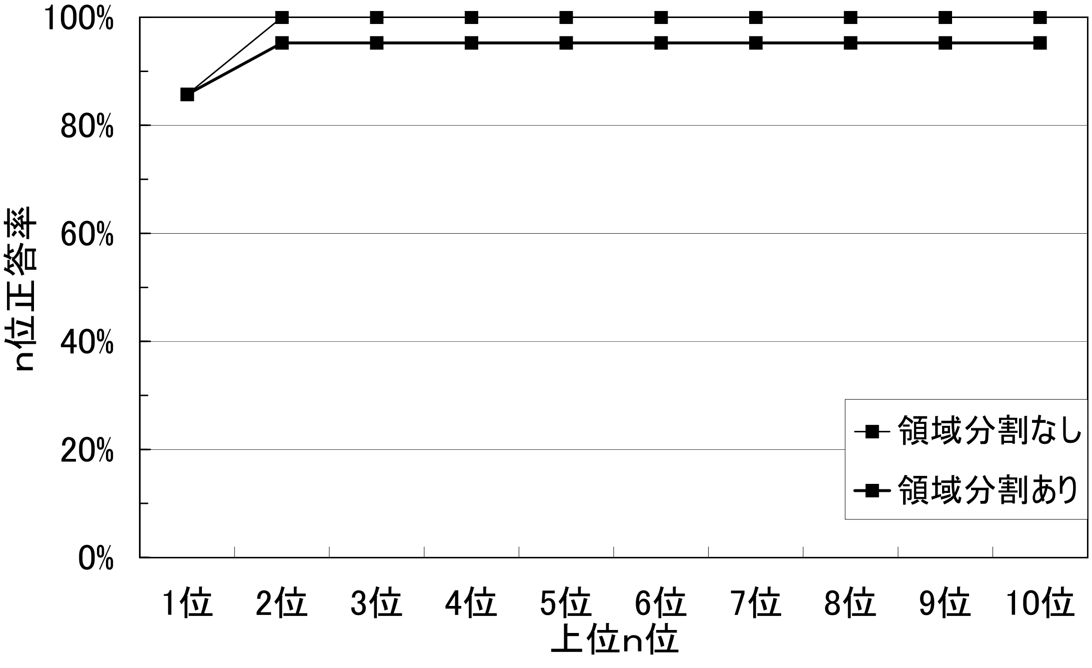


図 5.23: 類似度上位 n 位正答率（国会議場：人物なし，コリログラム）

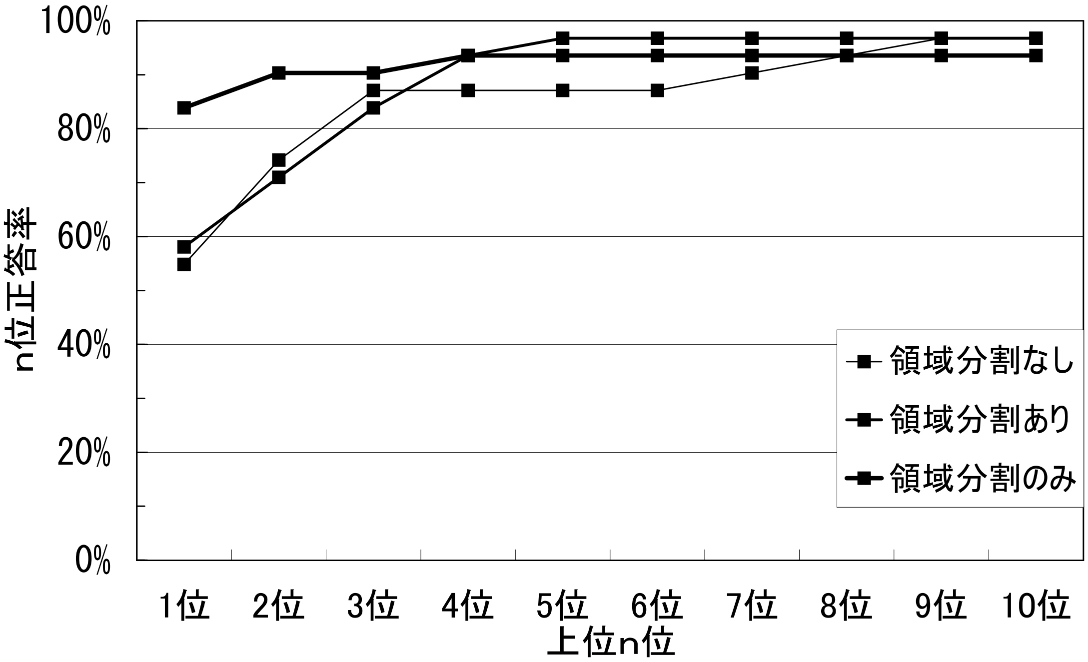


図 5.24: 類似度上位 n 位正答率（国会議場：人物あり，ヒストグラム）

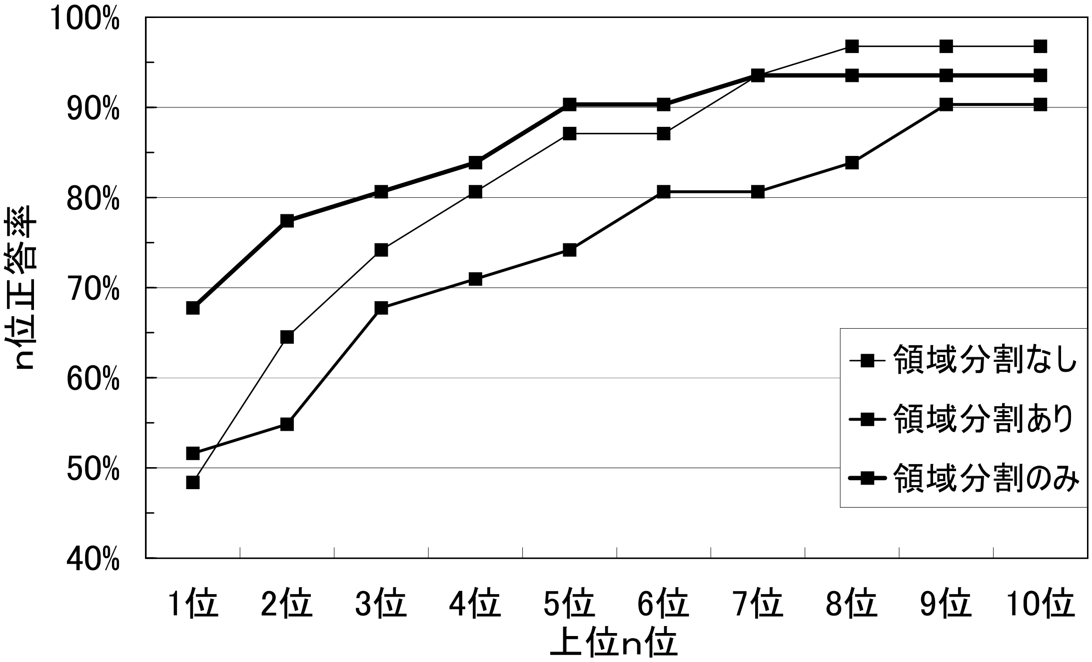


図 5.25: 類似度上位 n 位正答率（国会議場：人物あり，コリログラム）

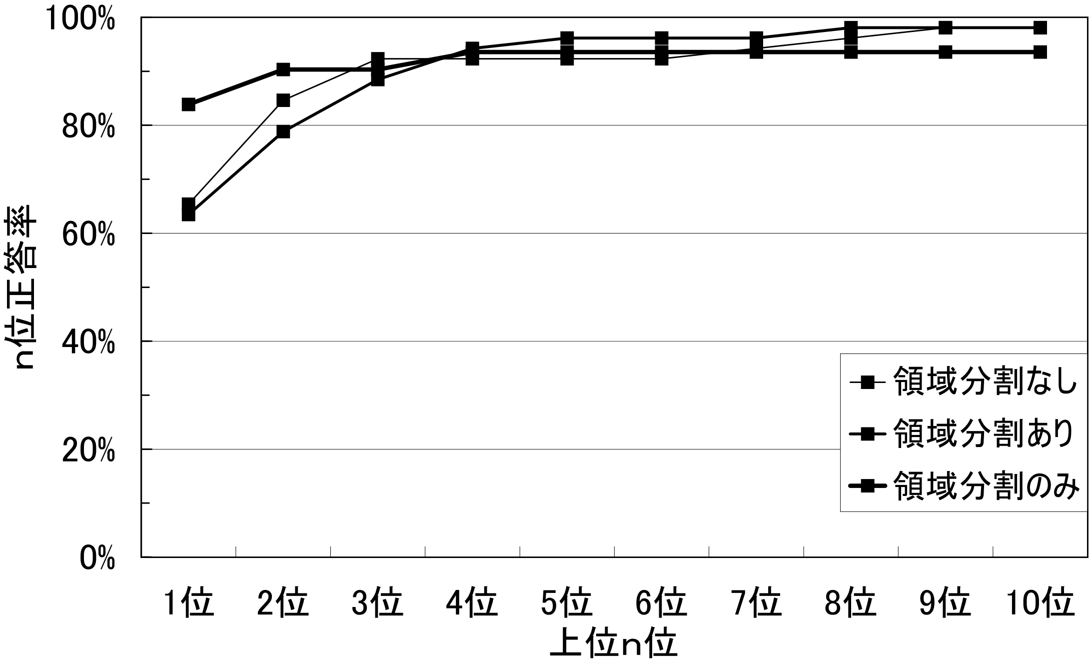


図 5.26: 類似度上位 n 位正答率（国会議場：総合，ヒストグラム）

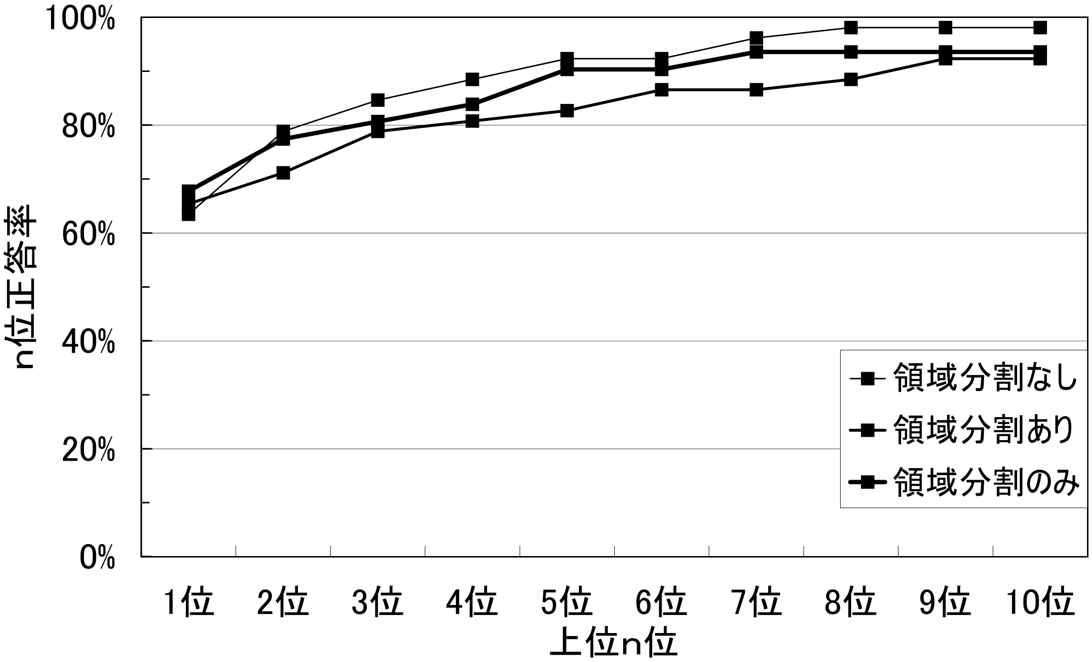


図 5.27: 類似度上位 n 位正答率（国会議場：総合，コリログラム）

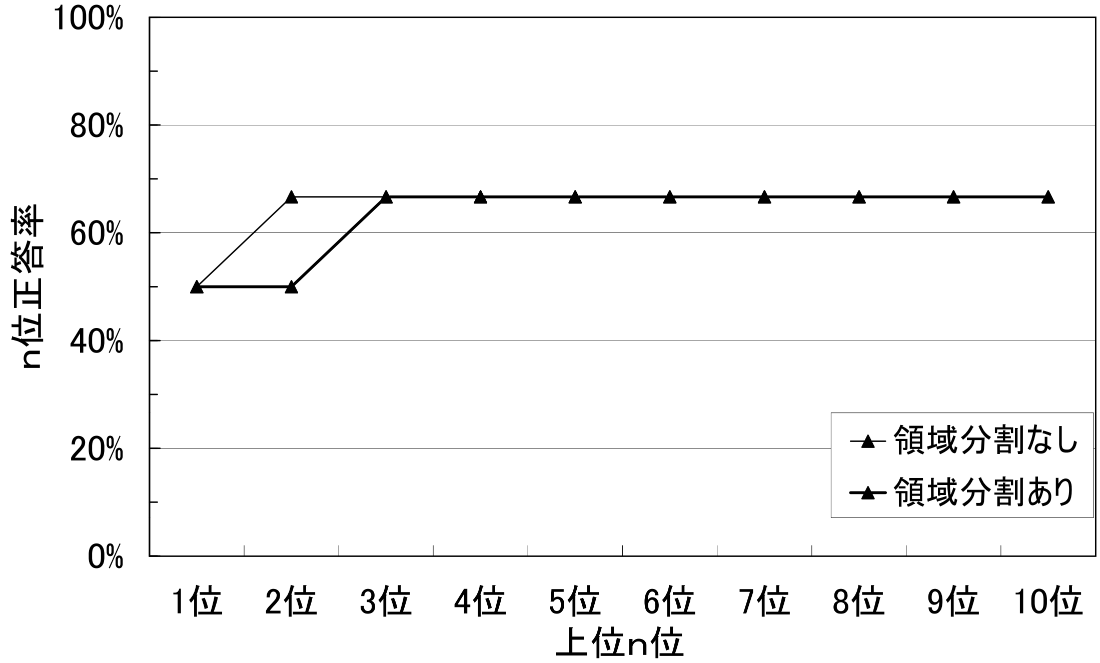


図 5.28: 類似度上位 n 位正答率（報道会見場：人物なし，ヒストグラム）

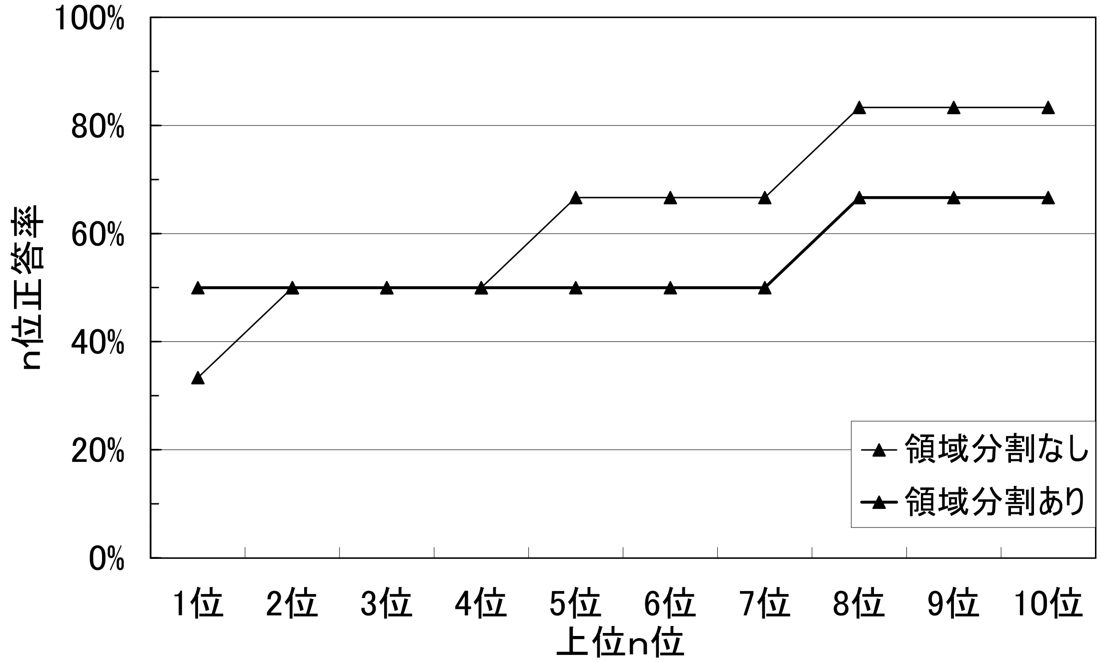


図 5.29: 類似度上位 n 位正答率（報道会見場：人物なし，コリログラム）

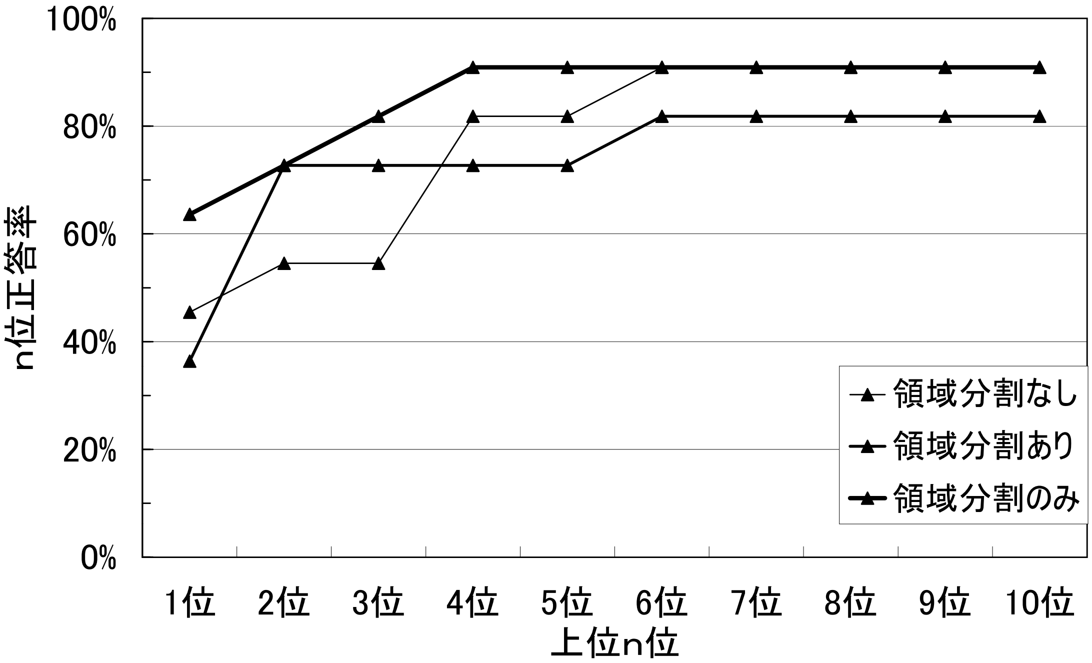


図 5.30: 類似度上位 n 位正答率（報道会見場：人物あり，ヒストグラム）

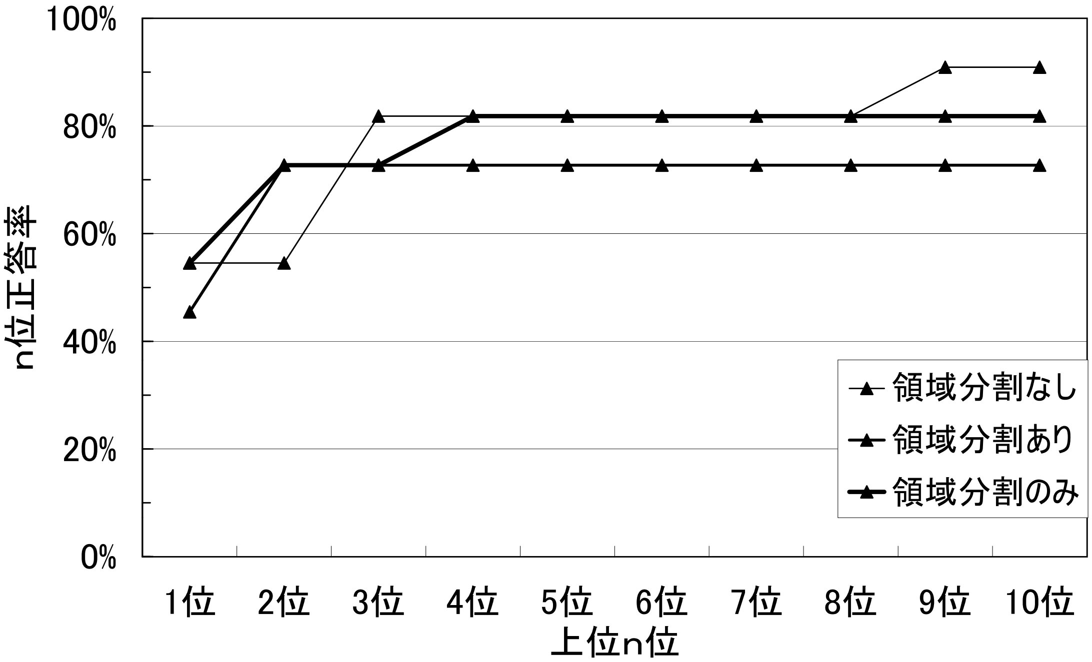


図 5.31: 類似度上位 n 位正答率（報道会見場：人物あり，コリログラム）

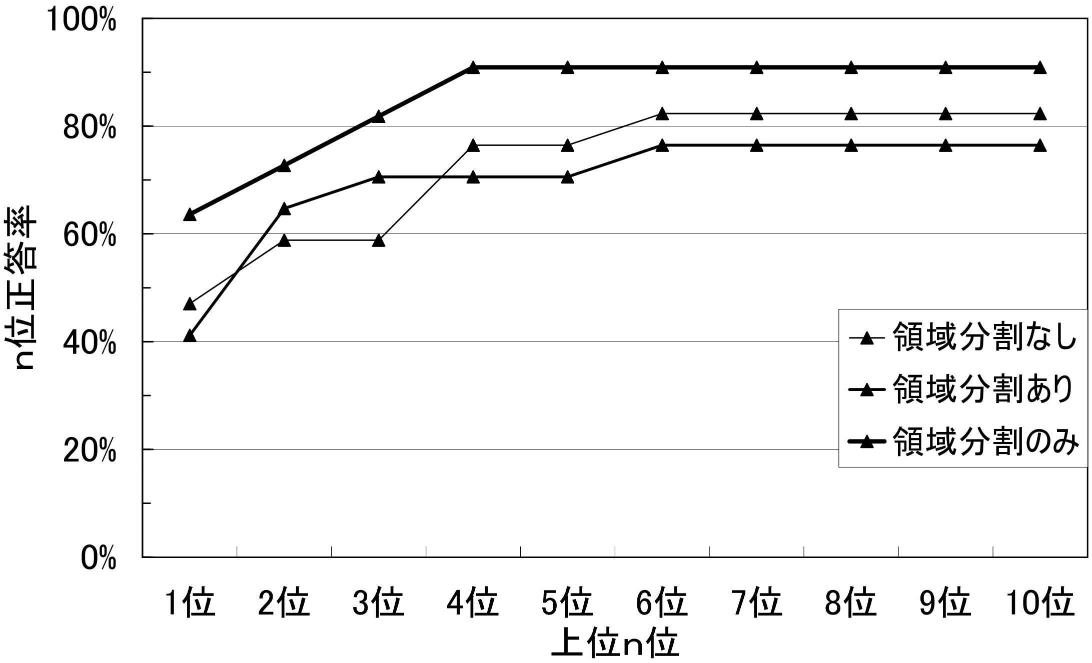


図 5.32: 類似度上位 n 位正答率（報道会見場：総合，ヒストグラム）

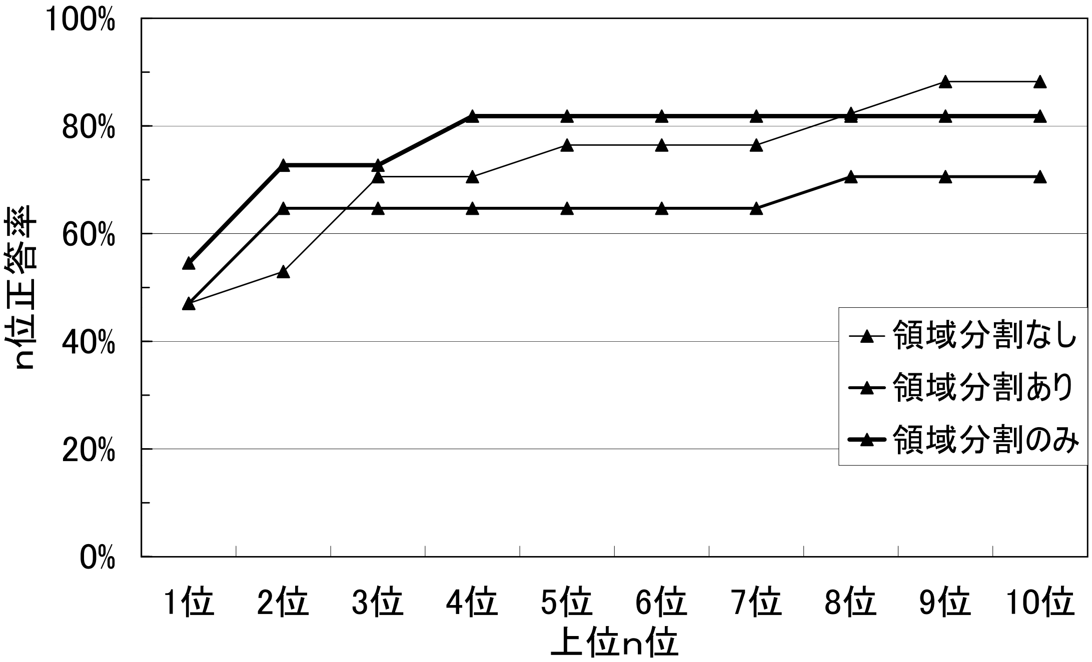


図 5.33: 類似度上位 n 位正答率（報道会見場：総合，コリログラム）

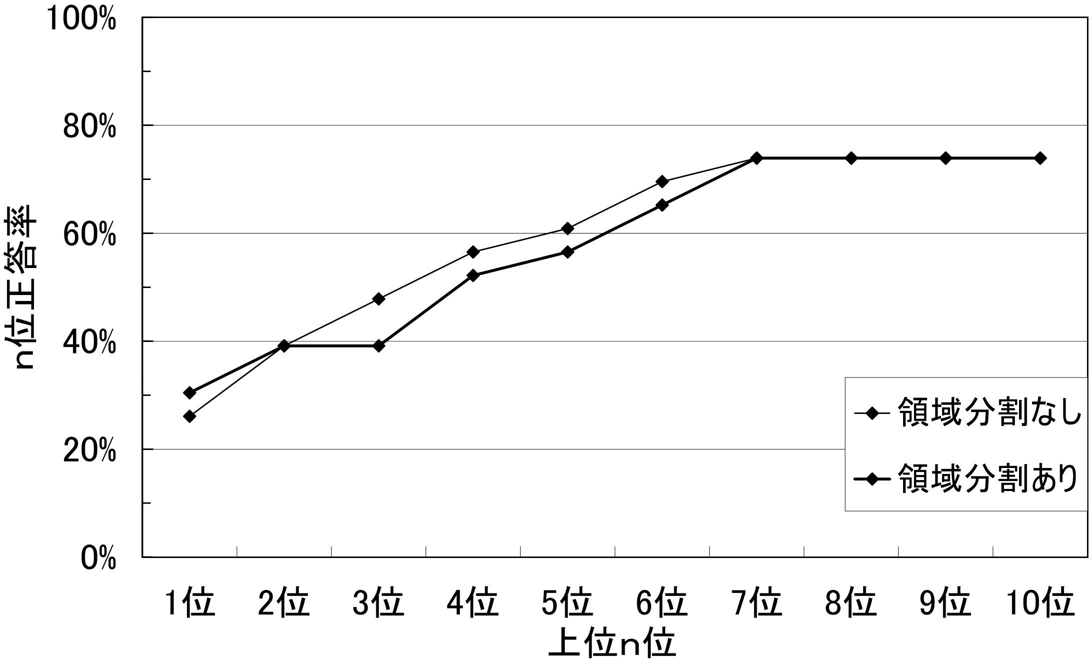


図 5.34: 類似度上位 n 位正答率（法廷：人物なし，ヒストグラム）

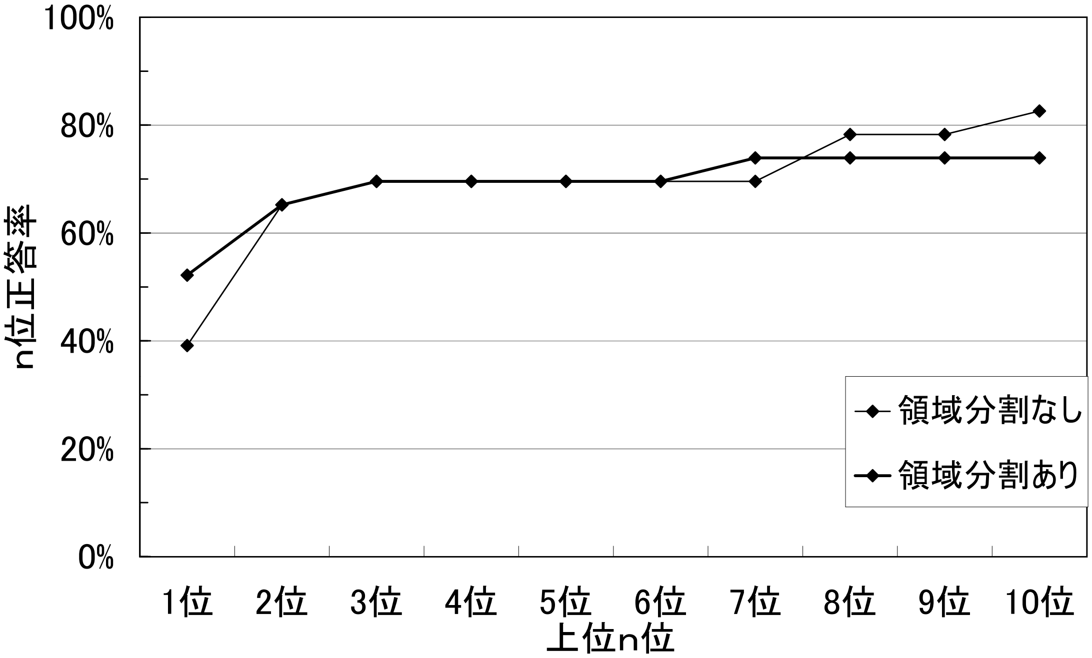


図 5.35: 類似度上位 n 位正答率（法廷：人物なし，コリログラム）

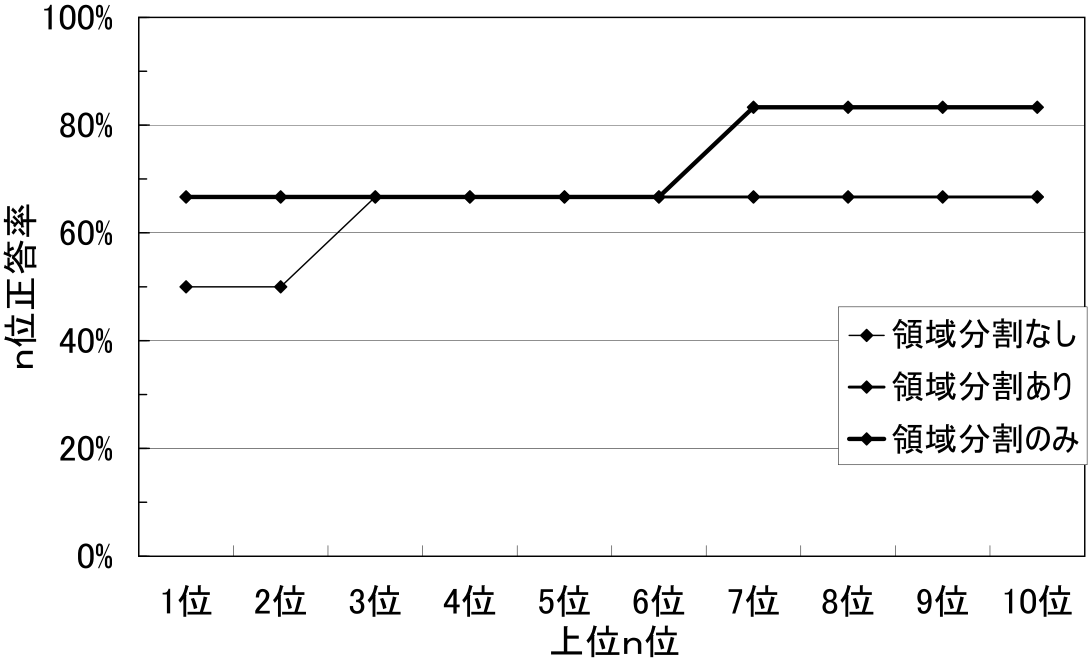


図 5.36: 類似度上位 n 位正答率（法廷：人物あり，ヒストグラム）

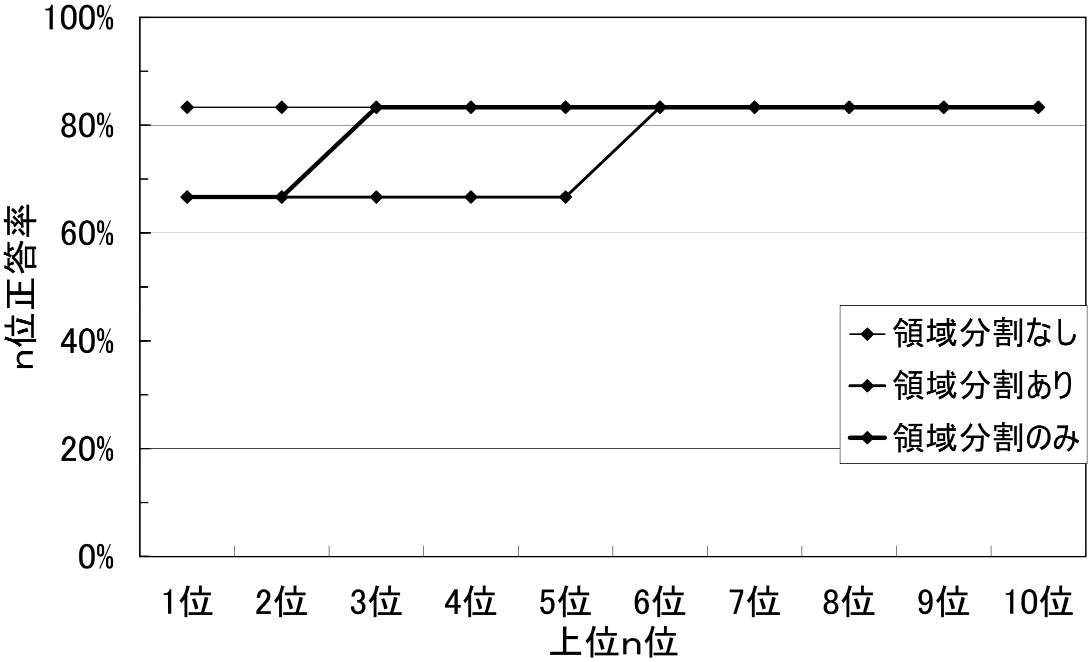


図 5.37: 類似度上位 n 位正答率（法廷：人物あり，コリログラム）

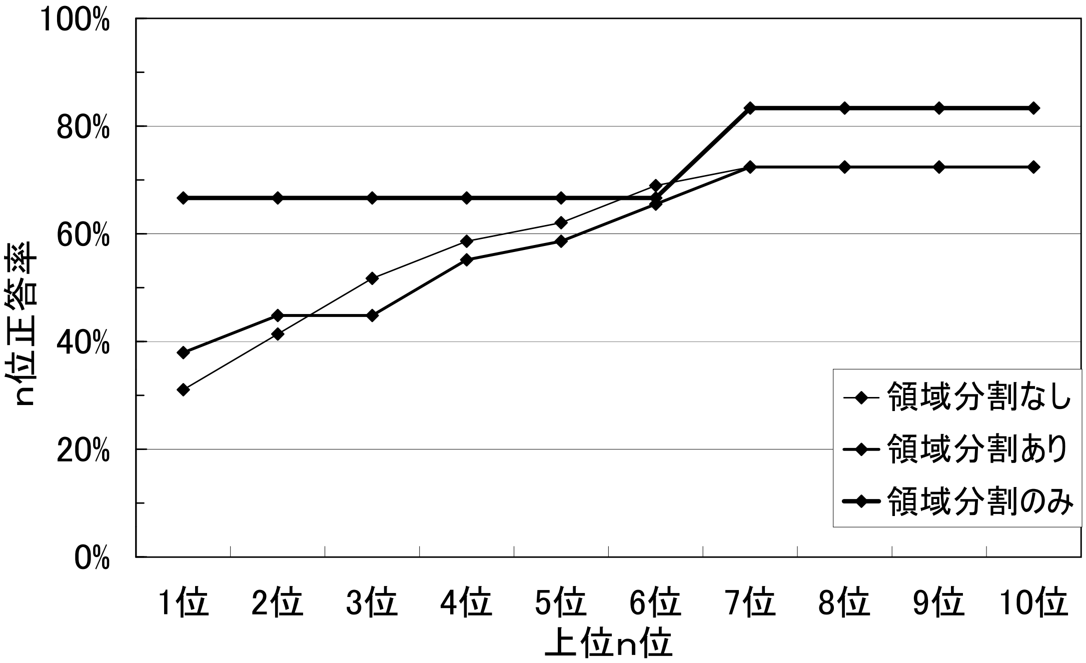


図 5.38: 類似度上位 n 位正答率（法廷：総合，ヒストグラム）

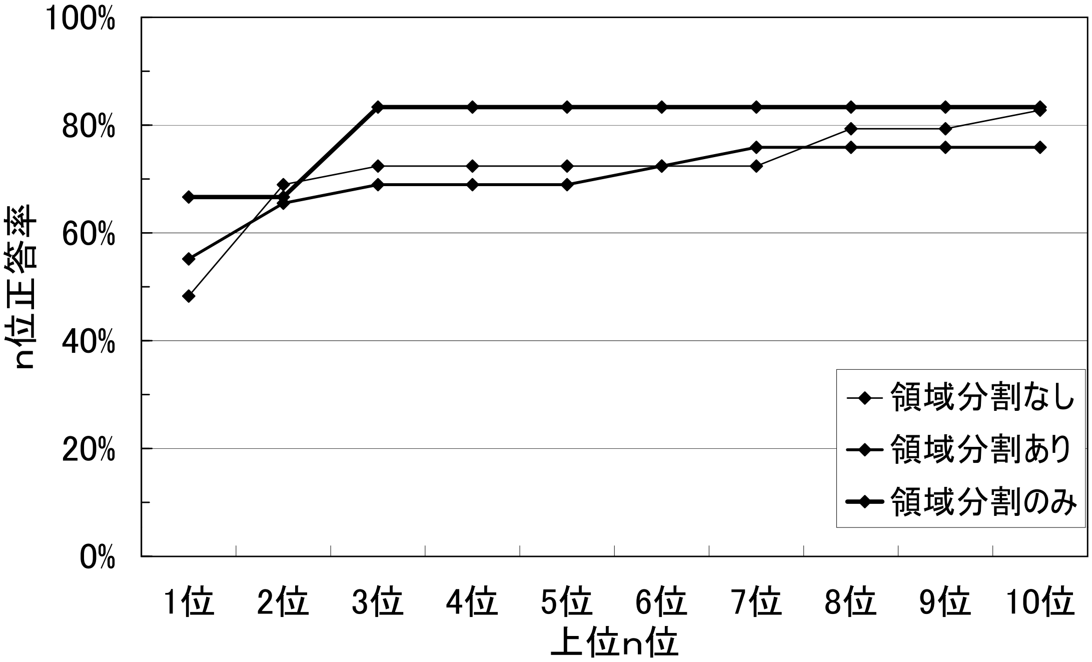


図 5.39: 類似度上位 n 位正答率（法廷：総合，コリログラム）

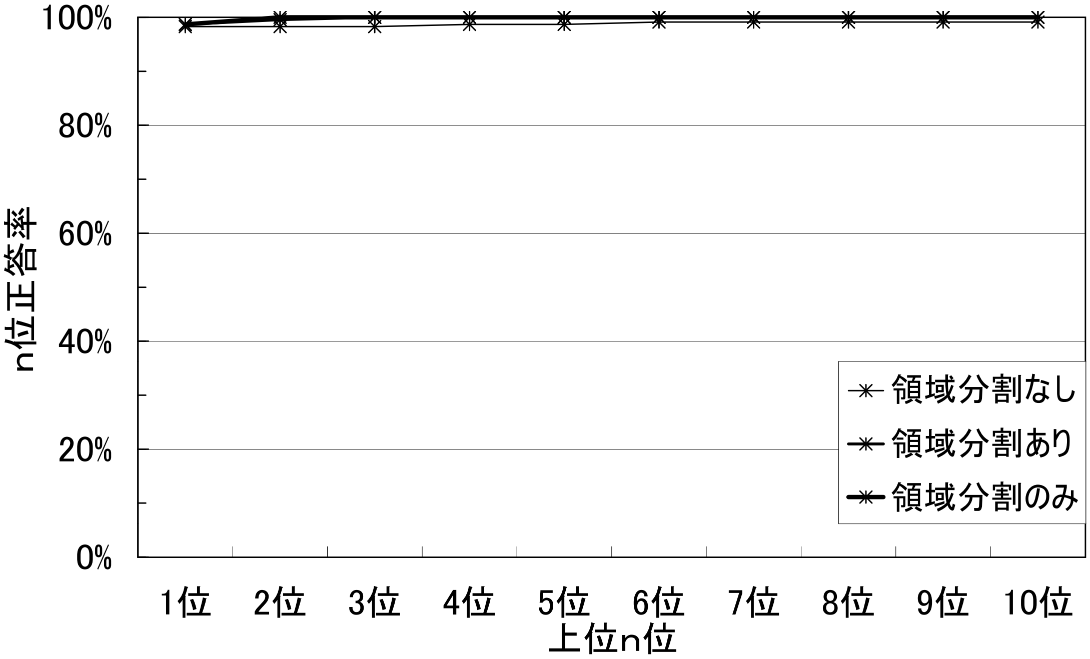


図 5.40: 類似度上位 n 位正答率（スタジオ：人物あり，ヒストグラム）

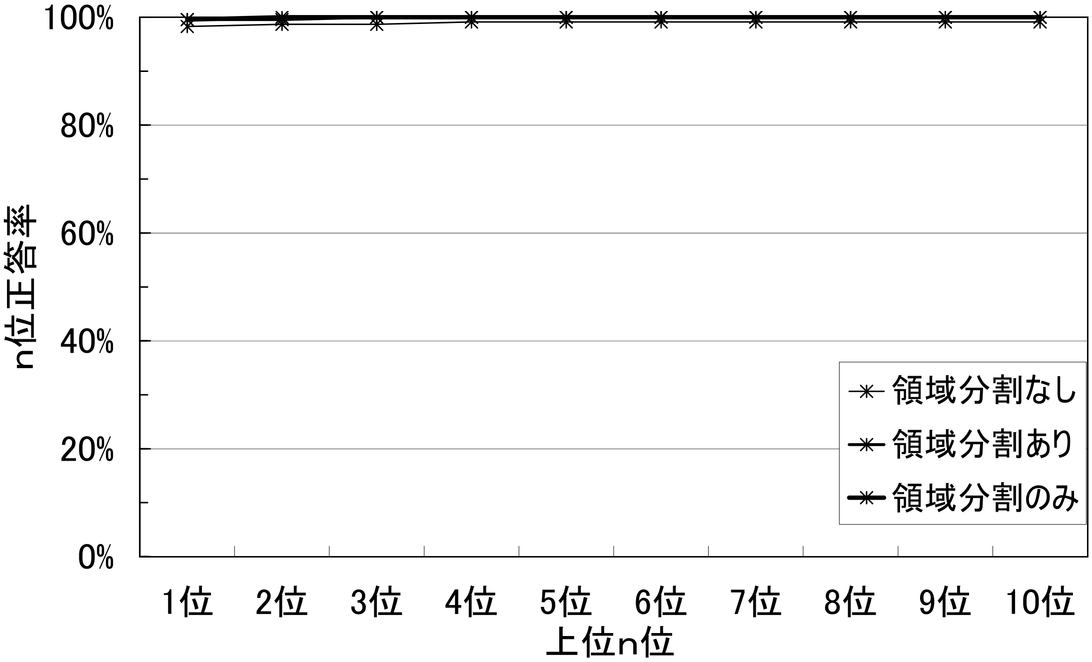


図 5.41: 類似度上位 n 位正答率（スタジオ：人物あり，コリログラム）

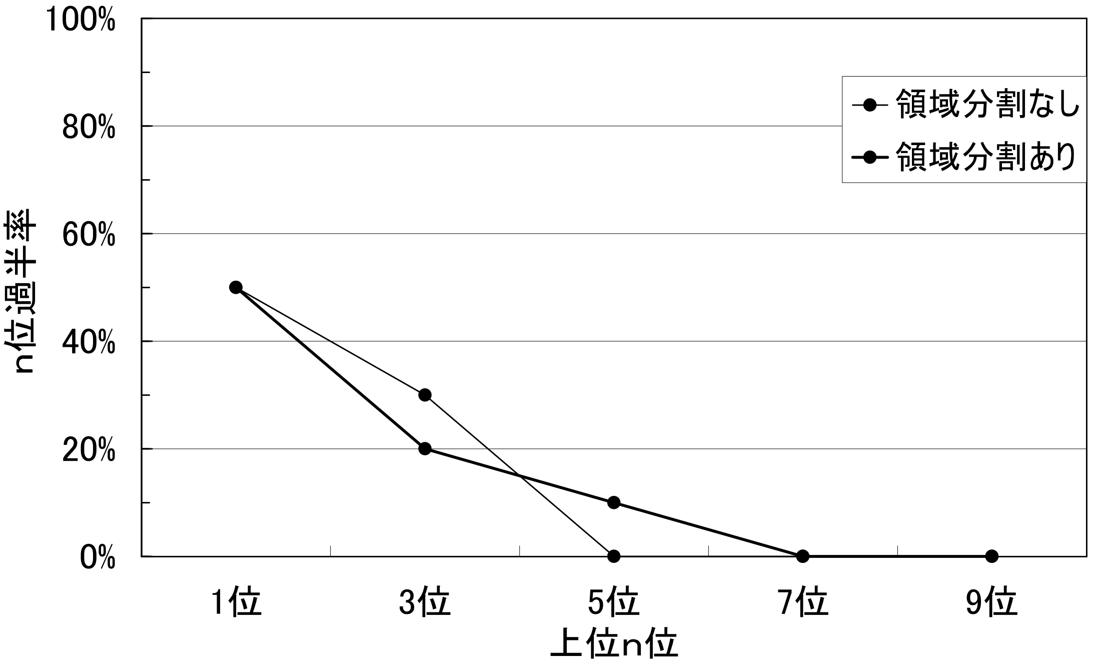


図 5.42: 場面推定性能（閣議控室：人物なし，ヒストグラム）

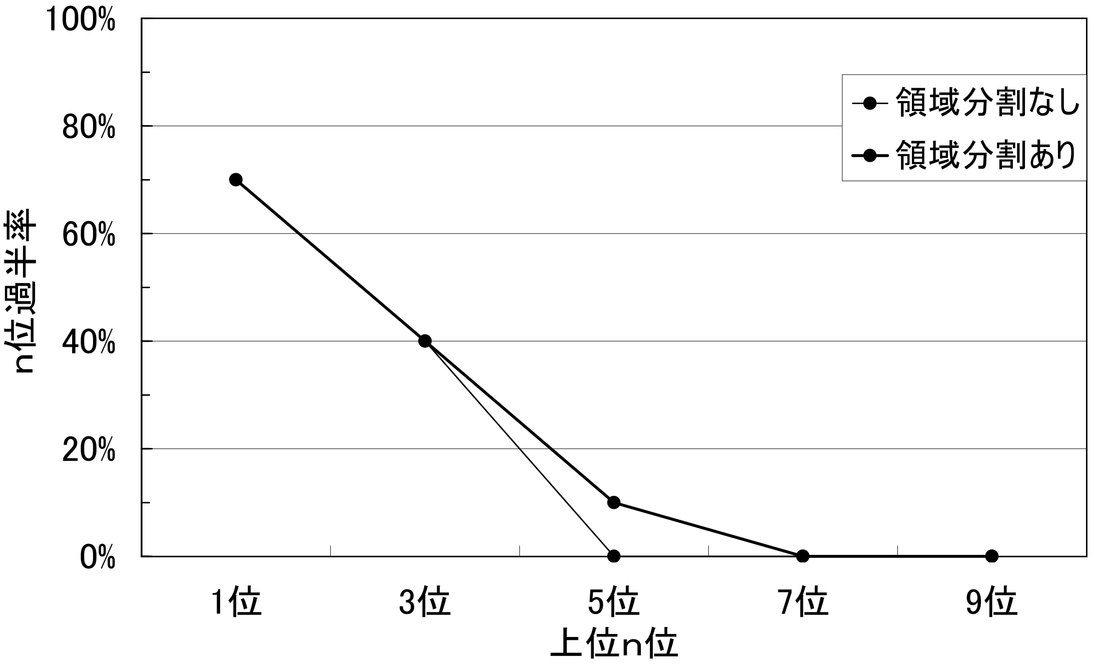


図 5.43: 場面推定性能（閣議控室：人物なし，コリログラム）

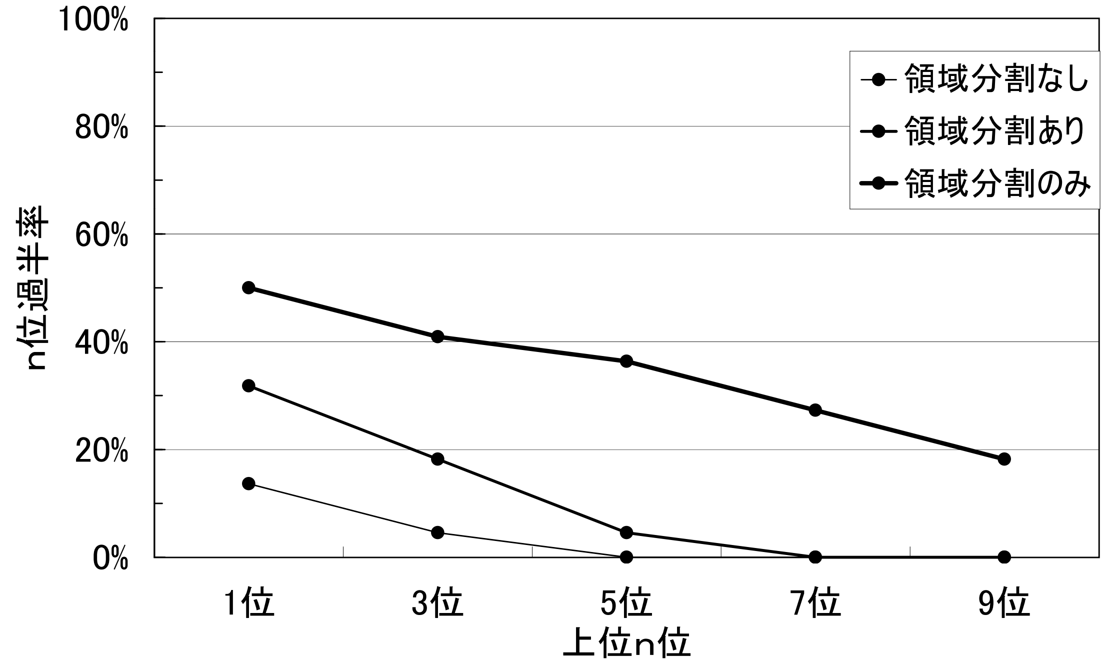


図 5.44: 場面推定性能（閣議控室：人物あり，ヒストグラム）

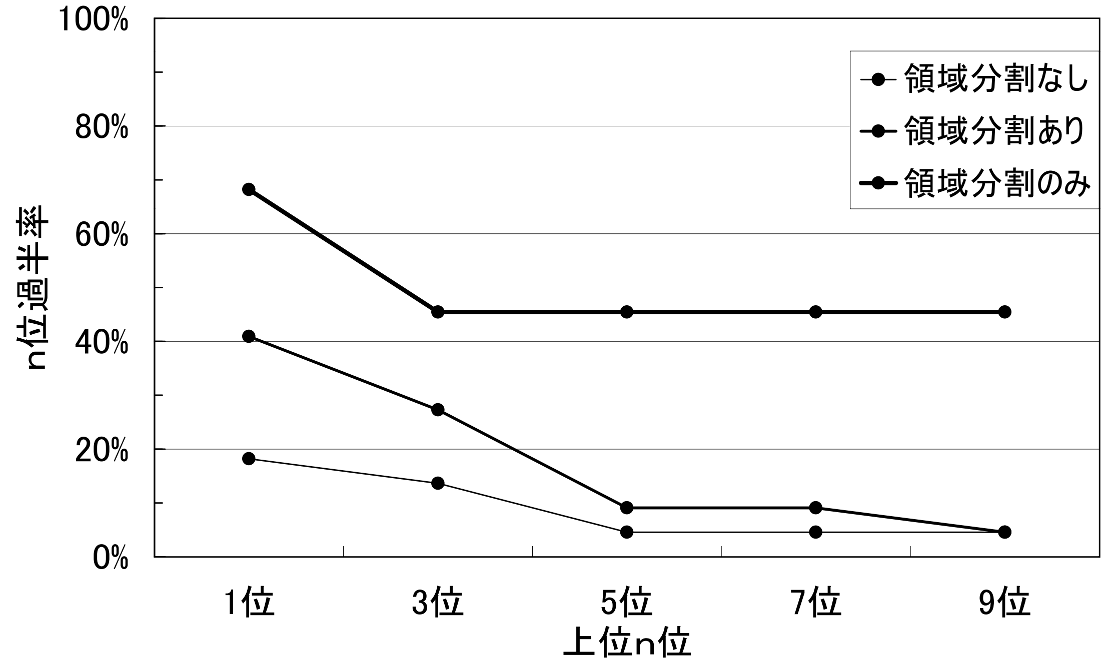


図 5.45: 場面推定性能（閣議控室：人物あり，コリログラム）

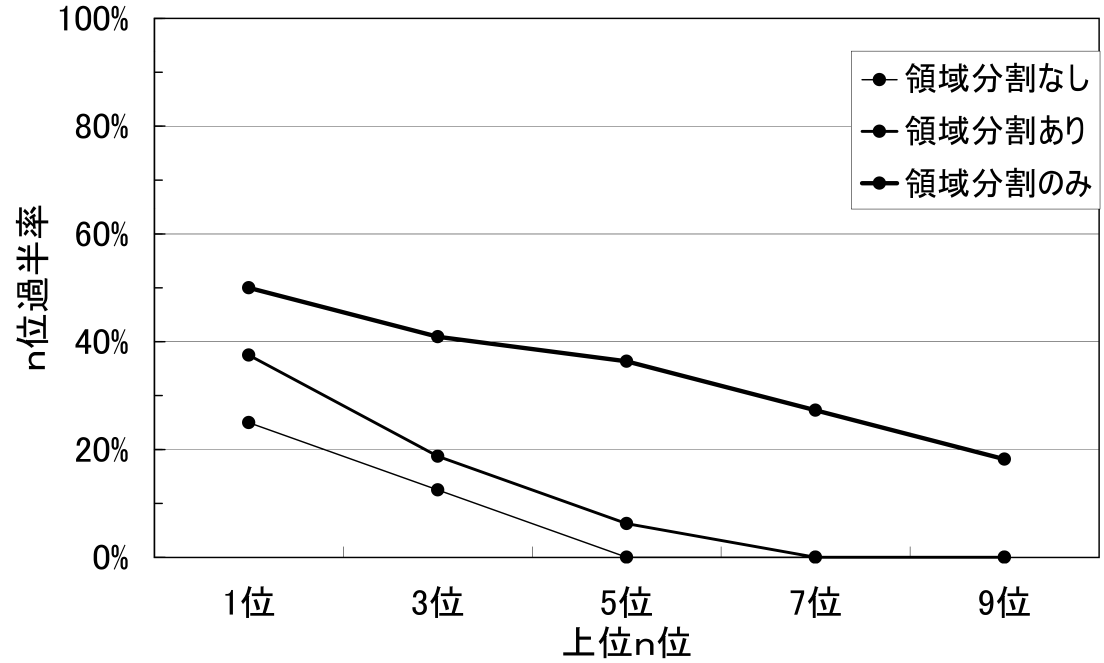


図 5.46: 場面推定性能（閣議控室：総合，ヒストグラム）

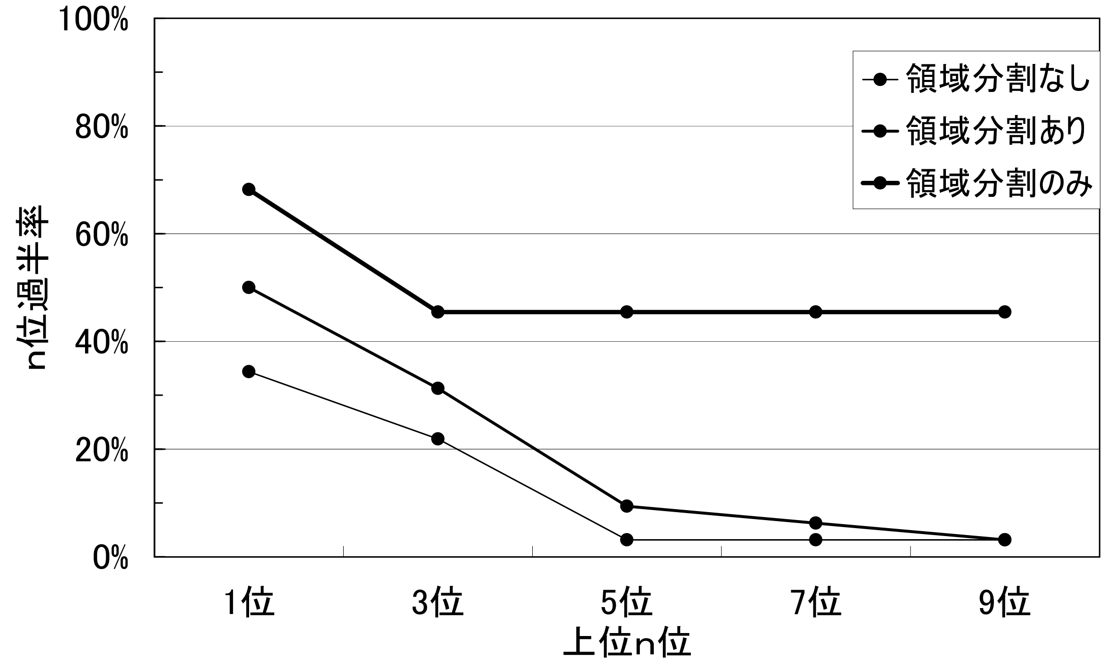


図 5.47: 場面推定性能（閣議控室：総合，コリログラム）

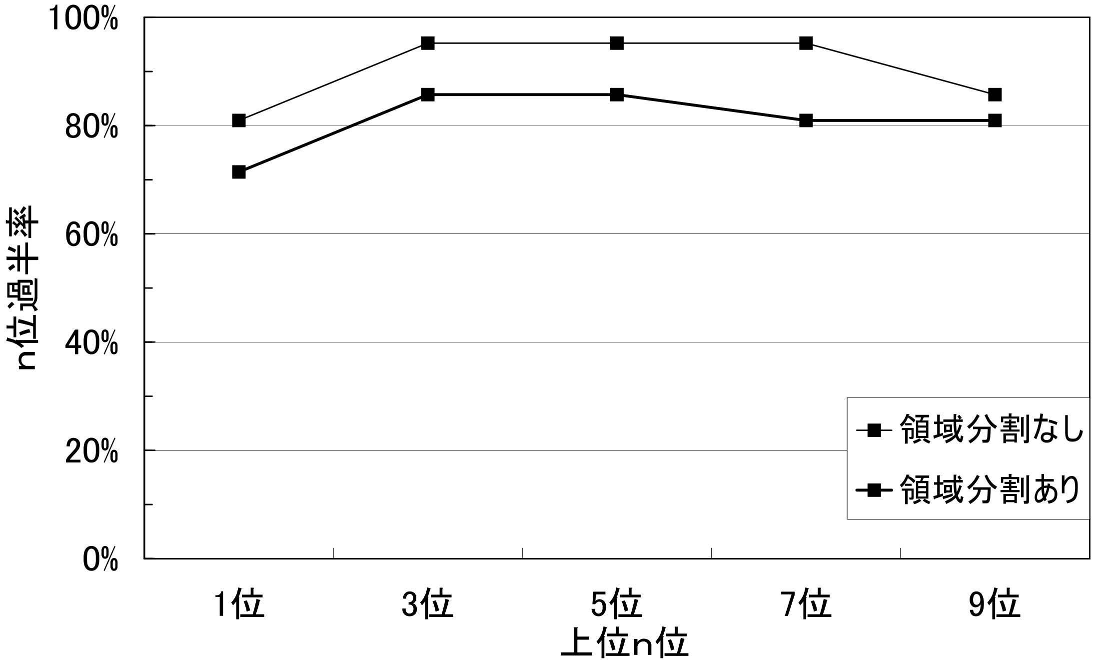


図 5.48: 場面推定性能（国会議場：人物なし，ヒストグラム）

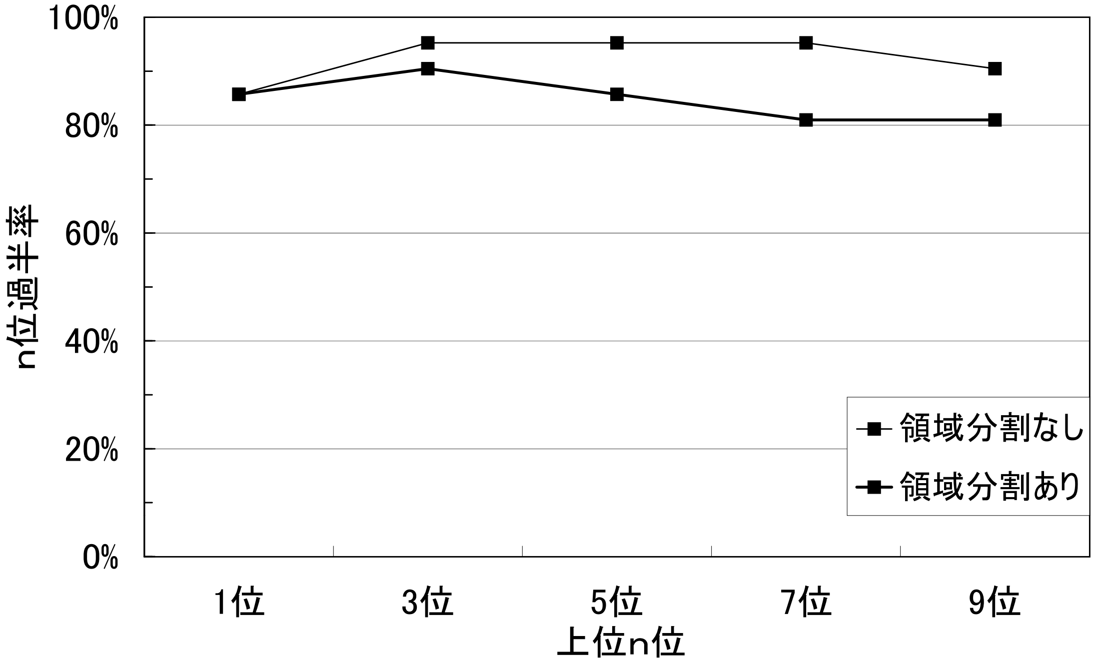


図 5.49: 場面推定性能（国会議場：人物なし，コリログラム）

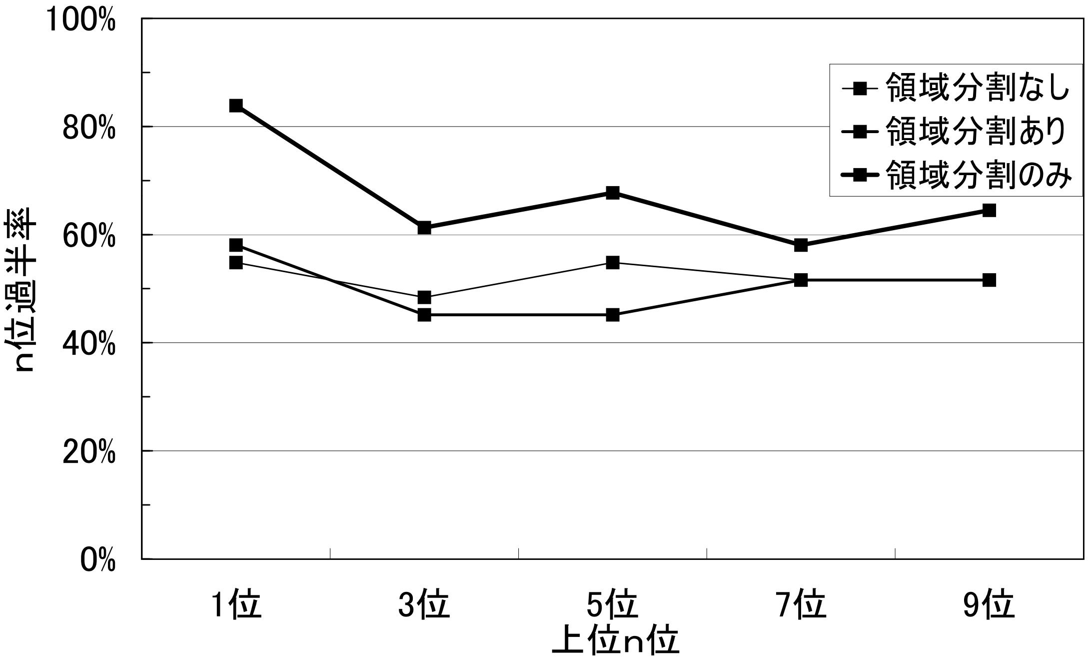


図 5.50: 場面推定性能（国会議場：人物あり，ヒストグラム）

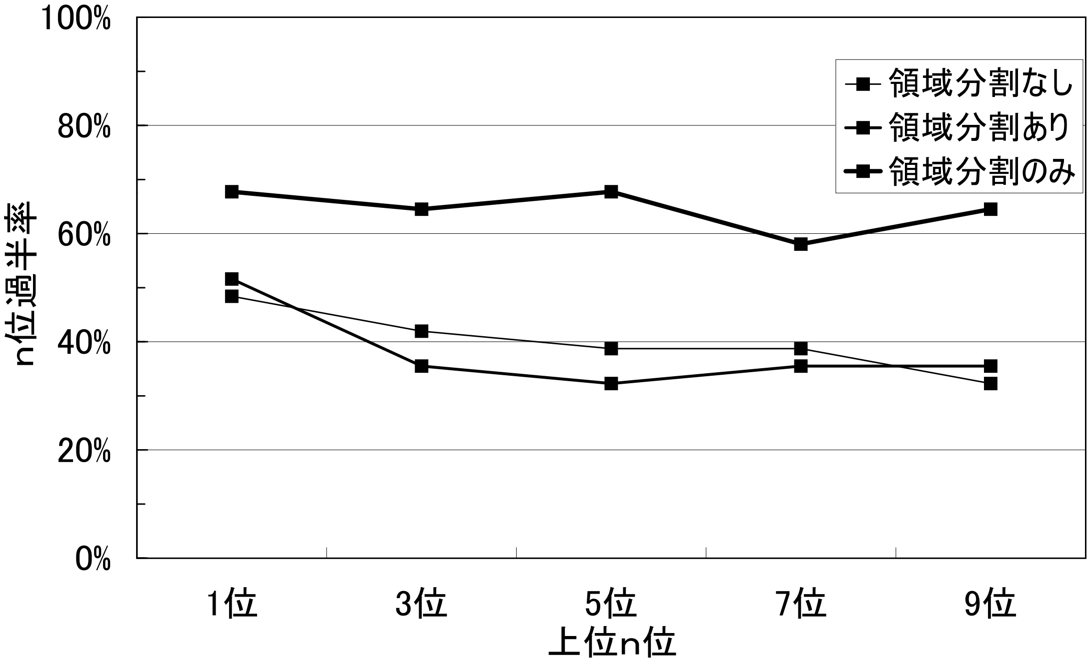


図 5.51: 場面推定性能（国会議場：人物あり，コリログラム）

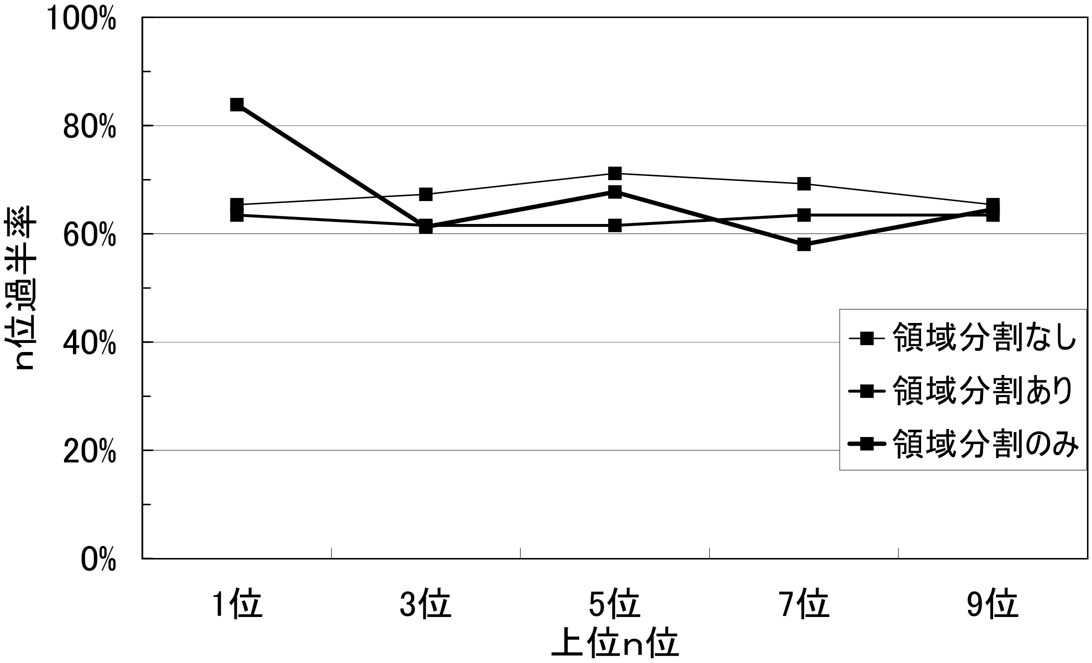


図 5.52: 場面推定性能 (国会議場：総合，ヒストグラム)

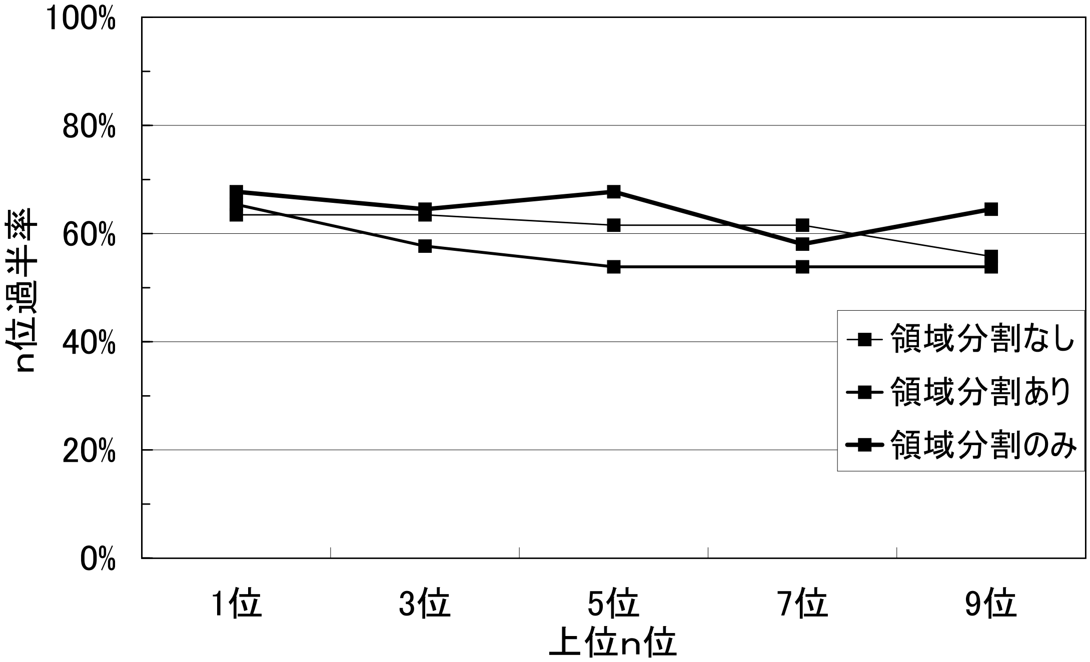


図 5.53: 場面推定性能 (国会議場：総合，コリログラム)

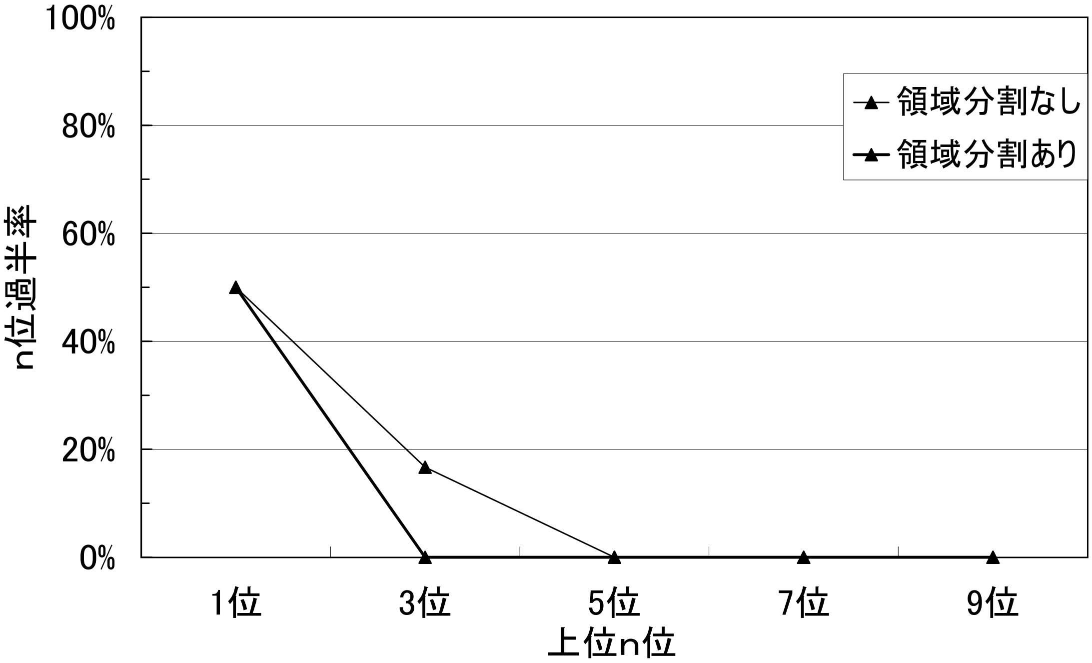


図 5.54: 場面推定性能（報道会見場：人物なし，ヒストグラム）

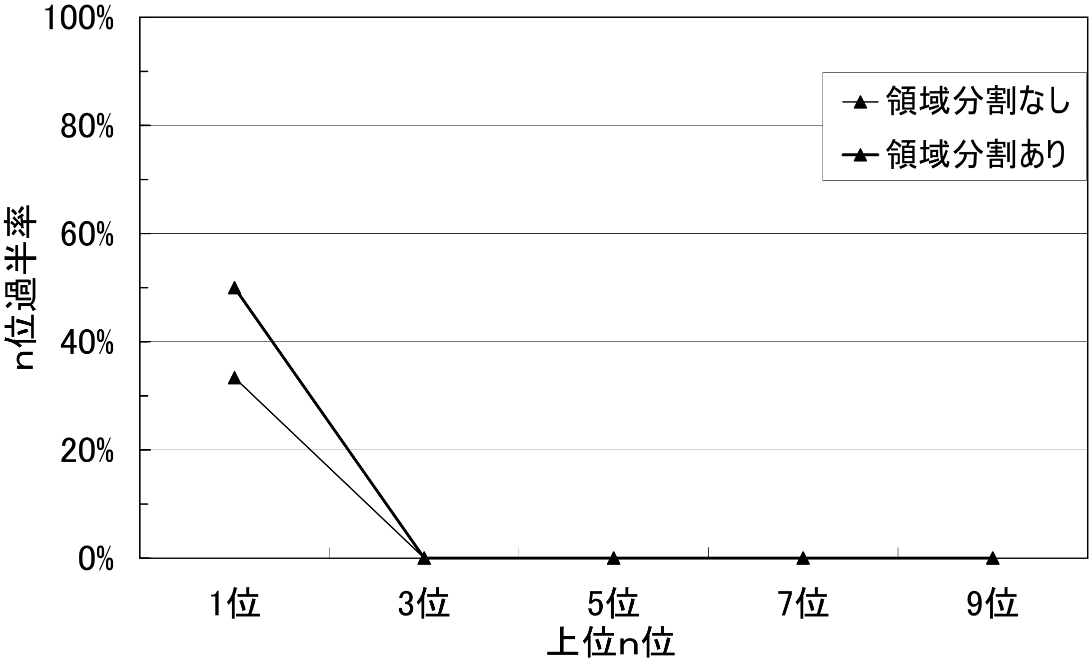


図 5.55: 場面推定性能（報道会見場：人物なし，コリログラム）

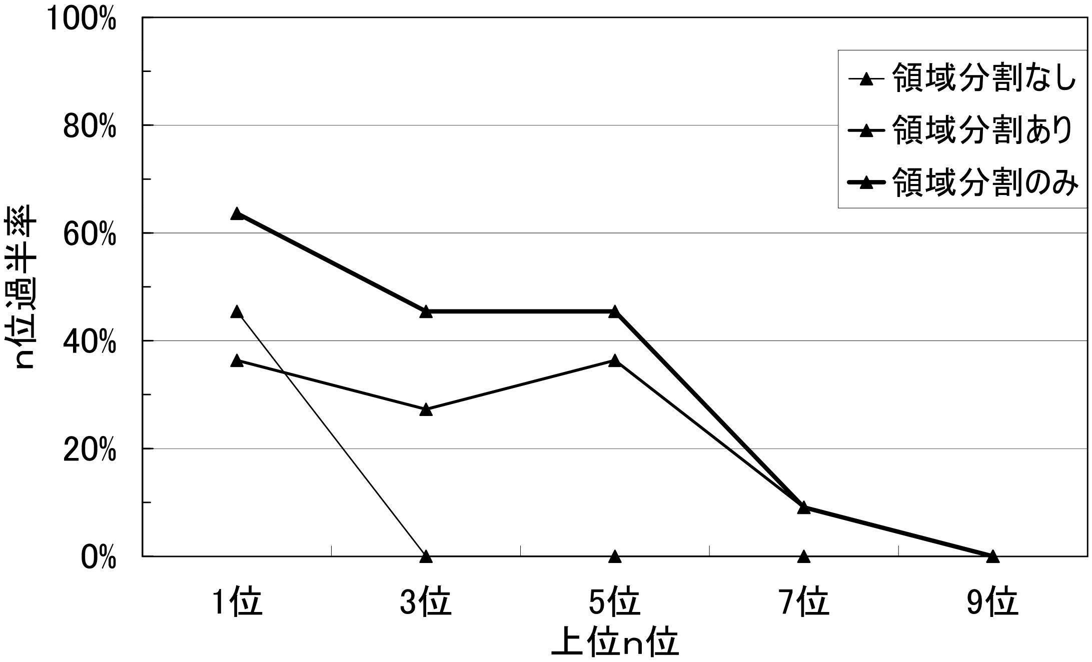


図 5.56: 場面推定性能（報道会見場：人物あり，ヒストグラム）

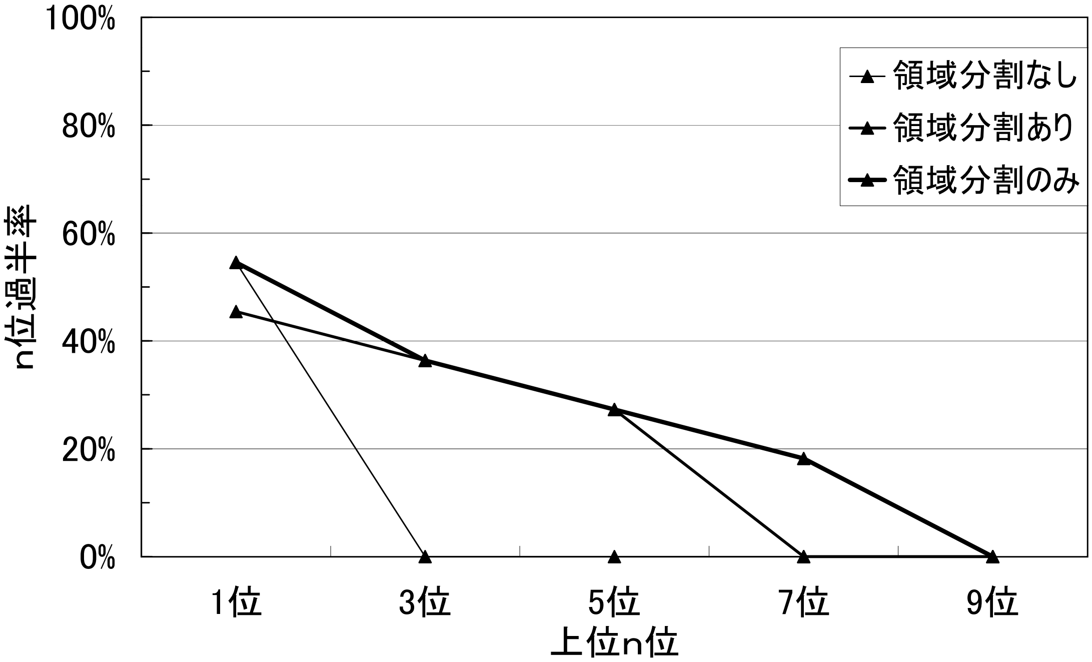


図 5.57: 場面推定性能（報道会見場：人物あり，コリログラム）

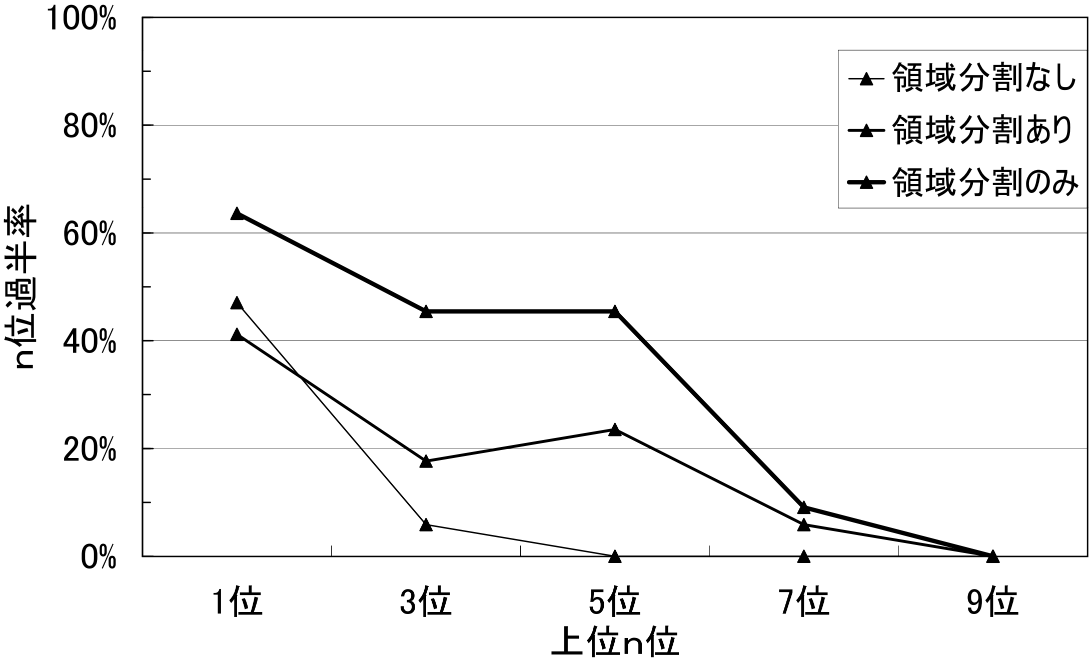


図 5.58: 場面推定性能（報道会見場：総合，ヒストグラム）

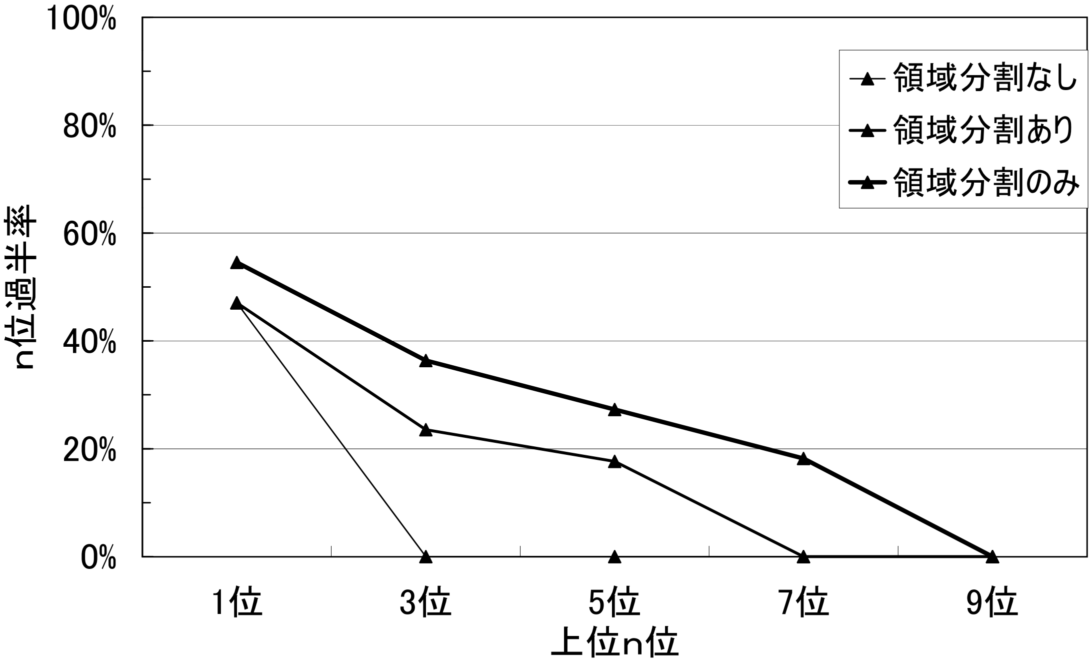


図 5.59: 場面推定性能（報道会見場：総合，コリログラム）

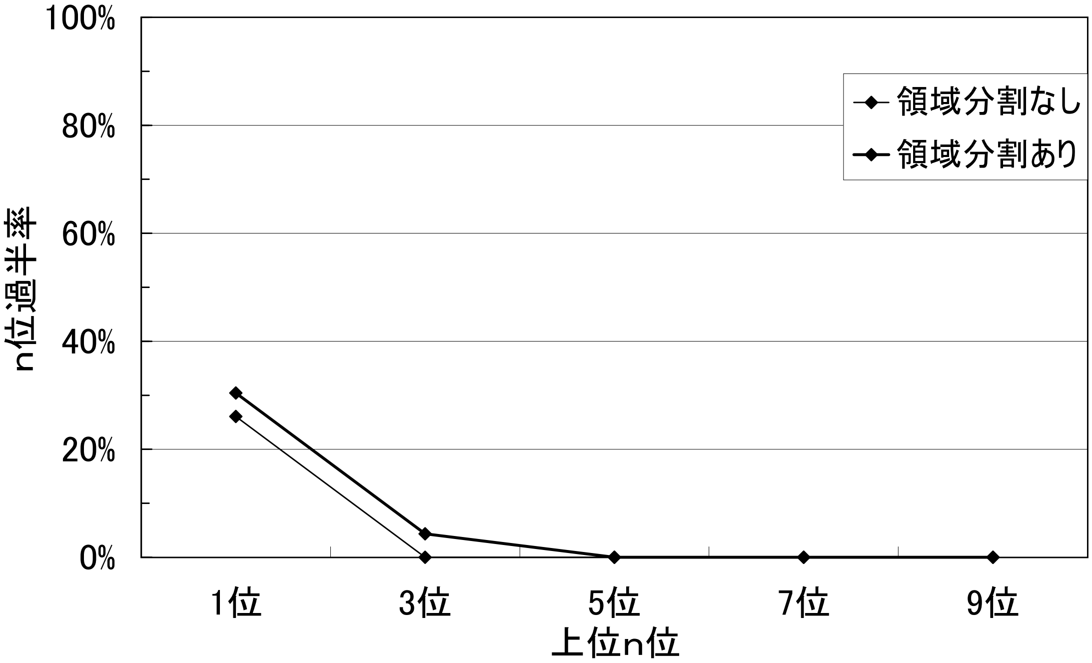


図 5.60: 場面推定性能（法廷：人物なし，ヒストグラム）

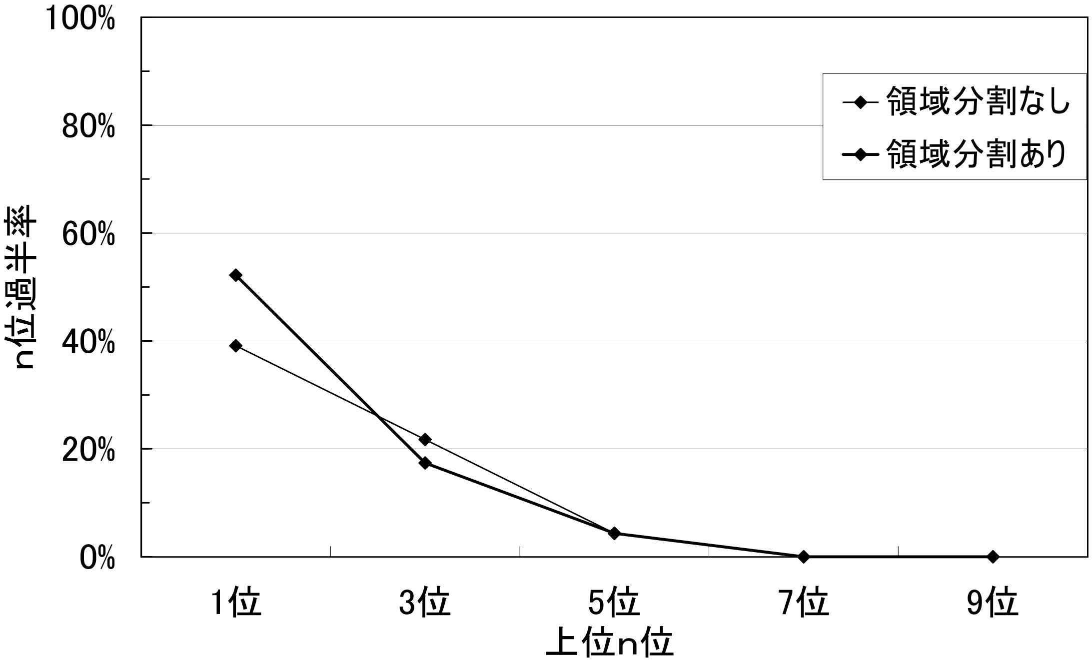


図 5.61: 場面推定性能（法廷：人物なし，コリログラム）

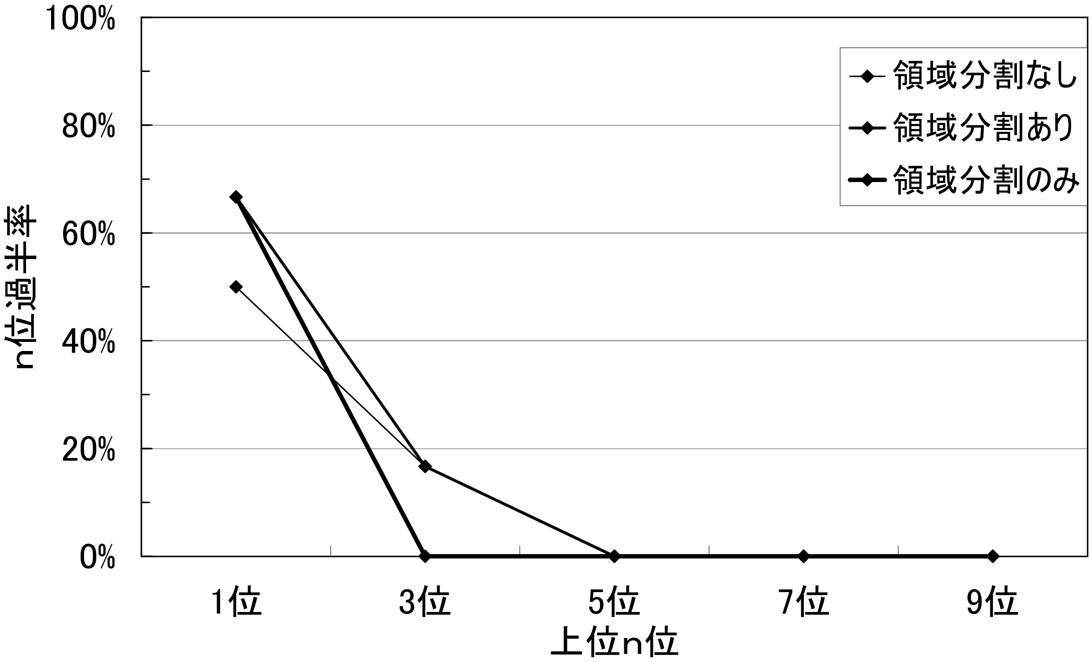


図 5.62: 場面推定性能 (法廷 : 人物あり , ヒストグラム)

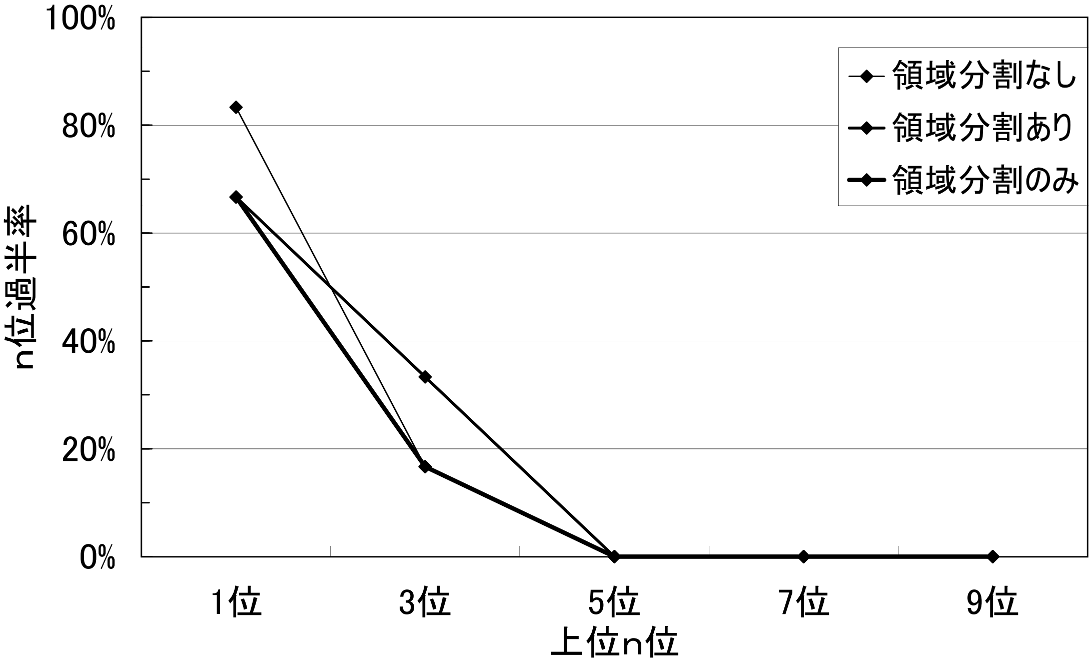


図 5.63: 場面推定性能 (法廷 : 人物あり , コリログラム)

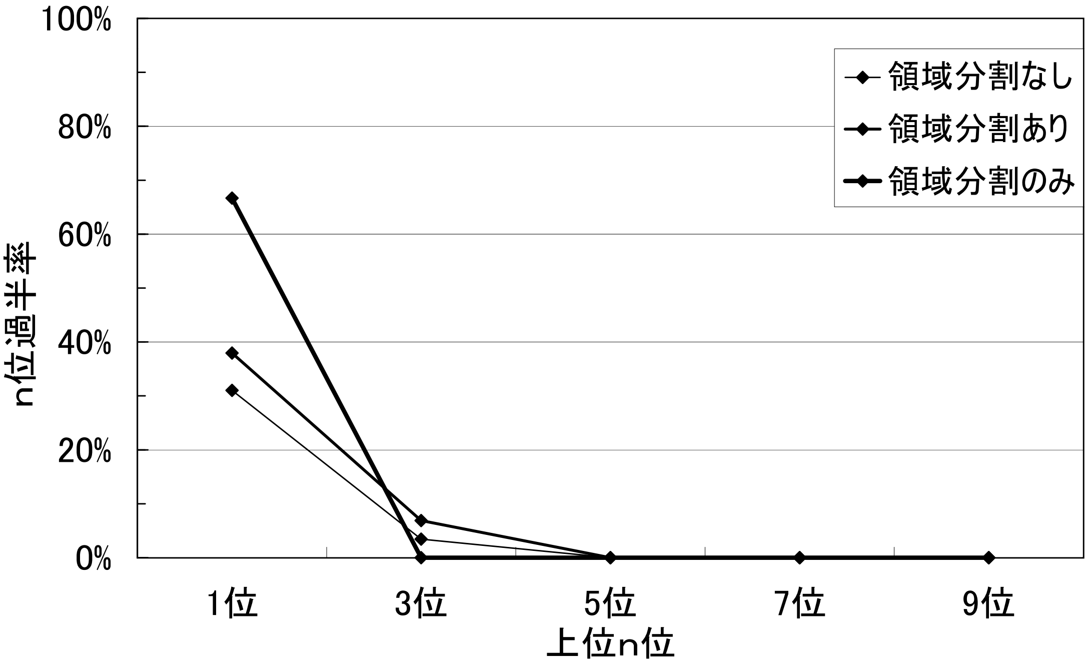


図 5.64: 場面推定性能（法廷：総合，ヒストグラム）

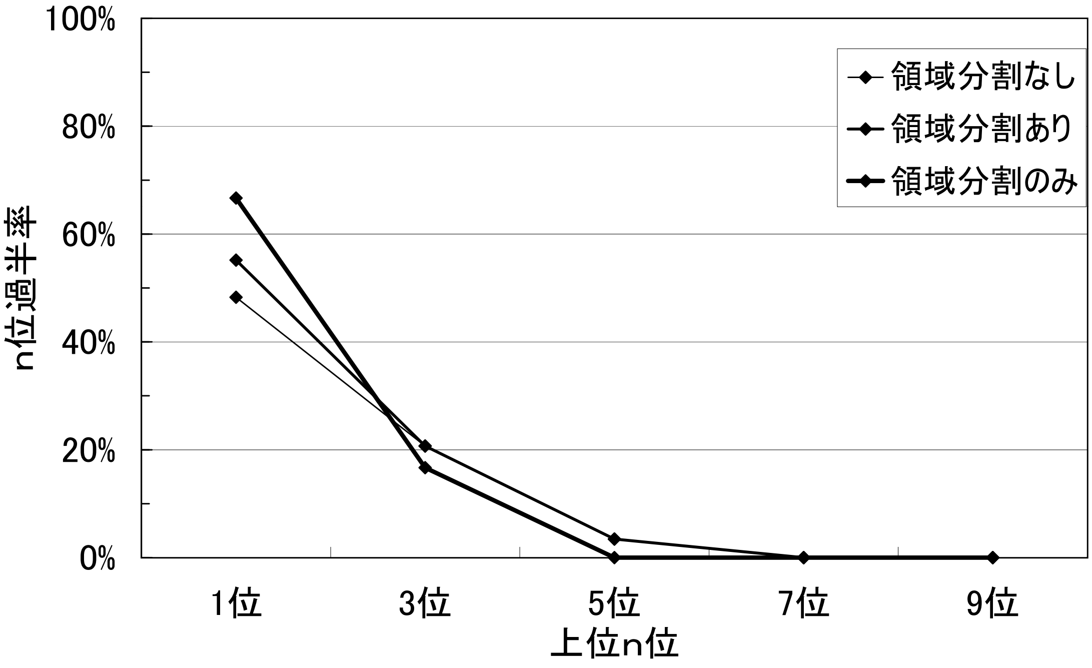


図 5.65: 場面推定性能（法廷：総合，コリログラム）

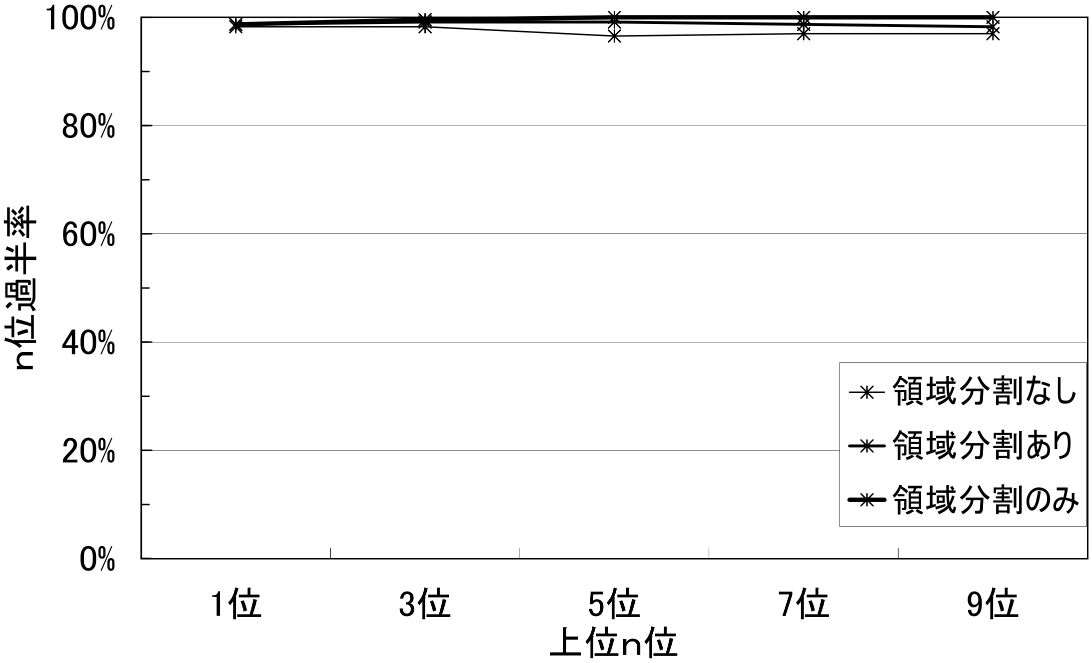


図 5.66: 場面推定性能（スタジオ：人物あり，ヒストグラム）

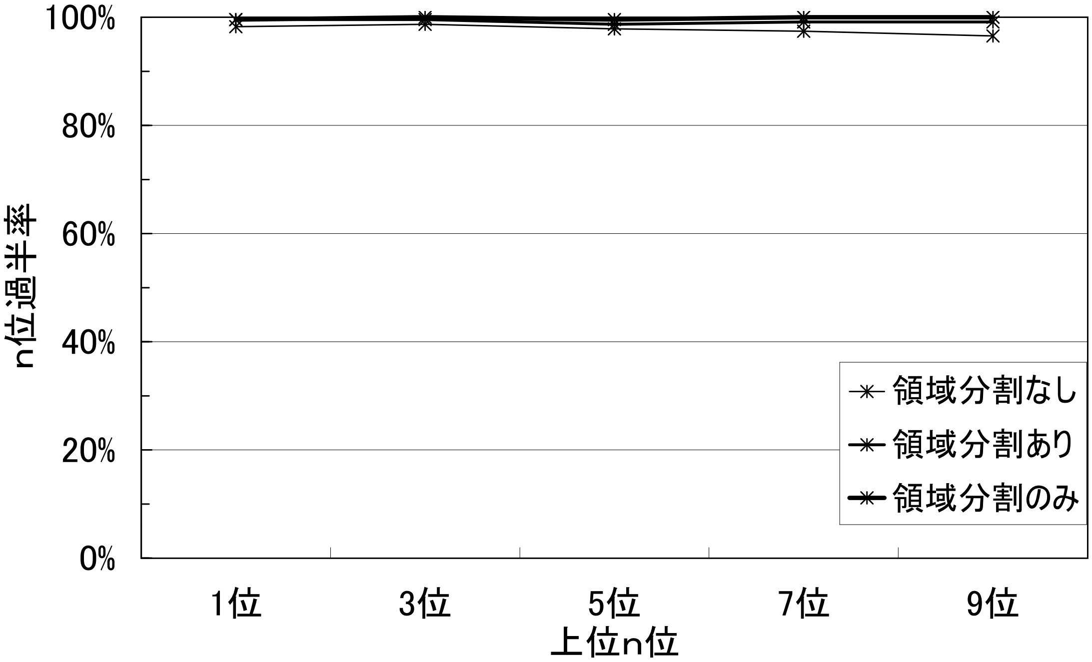


図 5.67: 場面推定性能（スタジオ：人物あり，コリログラム）

5.4 まとめ

本章では，画像特徴量に基づく画像内容解析手法として，典型的ショット分類及び画像特徴量と画像内容の関連に関する事例に基づく解析手法の枠組について述べた．具体的な画像内容解析の例として，ショット分類及び場面推定に関する実験を行った．特に，色彩に関する比較的単純な画像特徴量と顔領域の存在のみを用いて，頻出場面の推定に一定の割合で成功し，推定性能と用意した訓練事例数との関係から，事例数を指数規模で増やせば性能の向上が期待できることを示した．

第 6 章

テキストと画像の属性を考慮した索引付け

本章では，第 5 章及び第 6 章で行った自然言語処理と画像処理の結果をうけた，最終的なテキストと画像の属性を考慮した索引付けについて述べる．

6.1 テキストと画像の属性を考慮した索引付け：実装と評価

本節では，自然言語処理及び画像処理の結果得られる，索引候補及びショットのうち，属性が一致するもの同士の対応付けに関する実験について述べる．以下では，4W のうち，テキストと画像との対応を考慮する（1）Who（人物名詞–人物像）（2）Where（場所・組織名詞–場面）の 2 つの属性に関する索引付けについて述べる．

具体的には，索引付け対象画像との画像的類似度が高いショット中に出現するテキスト情報から，人物及び場所の属性を考慮して選択したものを用いて索引付けを行う．索引付けの詳細については，以下の実験手順のなかで述べる．

6.1.1 実験 6.1：ニュース映像に対する自動索引付け

以下の実験においては，表 5.3に示した合計 817 ショットを事例群として索引付けを行い，映像内容と索引との対応の正誤は目視により人間が判断した．

人物に関する索引付け

図 6.1に示すように，以下の手順に従って，表 5.3に示したもののうち，スタジオを除く人物像が存在する映像，合計 194 ショットを対象として索引付けを行った．

- 5.3.2で述べた手法により，キャスタ（スタジオ）ショットの出現を利用して，話題の境界を検出しておく
- 索引付け対象ショット S_tが属する話題中において，S_tと画像的類似度が高いショット $S_{r_i}(t)$ を選択（本実験では類似度の閾値を$\theta = 0.9$ に設定）

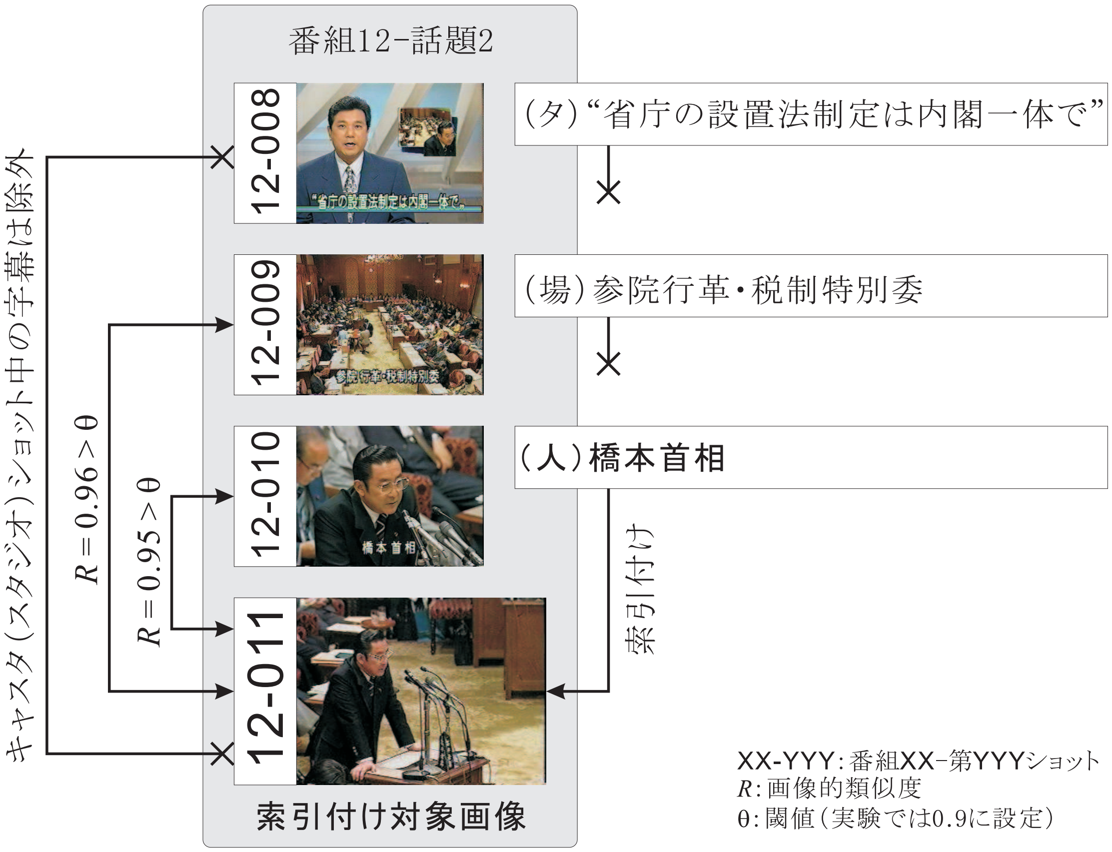


図 6.1: テキストと画像の属性を考慮した索引付け（人物名詞–人物像）

- $S_{r_i}(t)$ 中に出現する字幕のうち，キャスタ（スタジオ）ショットに現れるタイトル以外のもので，語義属性が「人物」であるものを索引候補に選択
- 抽出したショットと S_t との画像的類似度を索引候補の適合度とし，適合度が高い順に索引候補を提示

ショット間の人物の画像的類似度は，本来は人物領域の顔の特徴などに基づいて行うべきであるが，先行研究が多数存在し，同一人物検出は本研究の主題ではないため，実験 5.2 で求めた結果を利用し，色ヒストグラムと色コリログラムの 2 種類の特徴量について，領域分割を行わなかった場合（分割なし）で比較した．

場面に関する索引付け

図 6.2に示すように，以下の手順に従って，表 5.3に示したもののうち，対応の正誤が比較的明確に判断し得る，スタジオ及びその他の場面を除く映像，合計 130 ショットを対象として索引付けを行った．

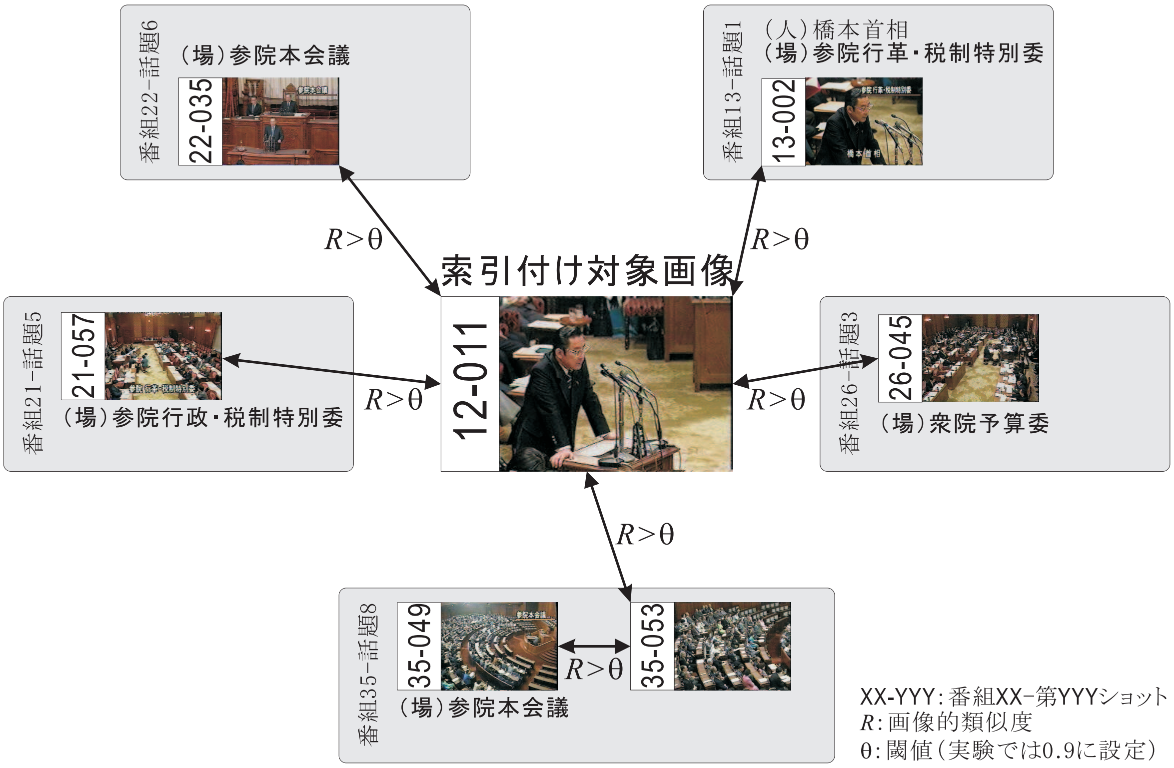


図 6.2: テキストと画像の属性を考慮した索引付け（場所・組織名詞−場面）

- 5.3.2で述べた手法により，キャスタ（スタジオ）ショットの出現を利用して，話題の境界を検出しておく
- 索引付け対象ショット S_tと画像的類似度が高いショット $S_{r_i}(t)$ を選択（本実験では閾値を$\theta = 0.9$ に設定）
- 各 $S_{r_i}(t)$ が属する話題中において，$S_{r_i}(t)$ と画像的類似度が高いショット $S_{r_j}(r_i)$ を選択（本実験では閾値を$\theta = 0.9$ に設定）
- $S_{r_i}(t), S_{r_j}(r_i)$ 中に出現する字幕のうち，キャスタ（スタジオ）ショットに現れるタイトル以外のもので，語義属性が「場所・組織」であるものを索引候補に選択
- 抽出したショットと S_tとの画像的類似度を索引候補の適合度とし，適合度が高い順に索引候補を提示

ショット間の場面の画像的類似度は，実験 5.2 で求めた結果を利用し，色ヒストグラムと色コリログラムの 2 種類の特徴量について，人物像が存在する画像において，領域分割を（1）行わなかった場合（分割なし），（2）行った場合（分割あり）で比較した（図 5.12 参照）．

6.1.2 実験結果と評価

人物に関する索引付け

図 6.3から図 6.6に人物に関する索引付けを行った結果を便宜上設定した場面別に示す．対応付け手法の限界を知るために，字幕の語義属性として，解析結果を用いた結果（図 6.5及び図 6.6）のほかに，正答を用いた結果（図 6.3及び図 6.4）を示す．

図 6.3及び図 6.5に人物に関する索引付けを行ったもののうち，適合度 1 位のものが正答であったショットの割合を示す．この評価は，話題中に適合度が閾値θ以上の人物に関する字幕が現れず，索引候補が 1 件も存在しなかった場合は含んでいない．そこで，索引付け対象の全ショットに対して，適合度 1 位のものが正答であったものの割合として，総合的な索引付けの成功率を図 6.4及び図 6.6に示す．

これらに見られる特徴と，場面毎の字幕の出現形態による原因の考察を以下に列挙する．

- 国会議場や報道会見場でほぼ完全に索引付け成功
  画像中の人物に関する字幕が頻出するため
- 閣議控室では成功率が低い
  人物に関する字幕があまり現れないものの，現れた場合は画像中の人物を的確に示す
- 法廷では成功率とともに正答率も低い
  人物に関する字幕があまり現れないうえ，現れても画像に写らない原告や被告などの人物に関するものが多い

この結果，全体的にみて 8 割程度の高い正答率は得られたものの，成功率は 4 割程度にとどまった．しかし，本実験で省略した同一人物検出を行えば，その精度によっては正答率が下がる危険性があるものの，他のショットから索引候補となる人名を抽出できるため，成功率は高くなるはずである．

場面に関する索引付け

表 6.9から表 6.24に場所に関する索引付けを行った結果を実験 5.2 で便宜上設けた頻出場面別（スタジオを除く）に示す．また，表 6.7及び表 6.8に索引付け対象の全ショットに対する総合的結果を示す．

ここでも，対応付け手法の限界を知るために，字幕の語義属性として，解析結果を用いた結果（各ページ下段）のほかに，正答を用いた結果（同上段）を示す．場面に関する索引付けにおいては，索引候補数がグラフに示した上位 10 件より少ない場合はほぼ存在しなかったため，人物に関する場合に分離して議論した正答率と成功率をほぼ同一視する．また場面の正答としては，必ずしも具体的な討議内容を含んだ会議名などの一致ではなく，類似した場面であると判断されるものを含めた．

これらに見られる特徴と，場面毎の字幕の出現形態や人物の有無による原因の考察を以下に列挙する．

- 人物領域が存在しない場合には比較的高正答率
- 報道会見場では完全に索引付け失敗
  会見場であることを示す字幕が全く現れなかった
- 人物なしの場合が人物ありの場合と比較して高正答率
  人物領域と背景領域を分離しても，大映しになった画像の背景の画像的特徴は，わずかな条件の違いによっても大きく変化する
- 人物領域分割の効果はわずかにみられる程度
  便宜上設定した場面において，人物領域（特に胴体領域）の画像特徴量は比較的共通しており，人物領域分割を行わなくても特徴が得やすく，領域分割の効果が現れにくい

この結果，全体的にみて，最上位で 2 割から 3 割程度，上位 10 位で 5 割から 6 割程度の正答率にとどまった．字幕属性に正答を用いた場合と比較して解析結果を用いた場合に，正答率が 2 割程度低いのは，場所・組織名詞の語義属性解析の低適合率（実験 4.1 参照）が統合処理によっても解消しきれなかったことによる．

しかし，上位 10 位でみると，比較的高い正答率が得られていることから，より画像特徴を抽象的に表現する特徴量の採用や，類似度計算の工夫により，一層の性能向上がはかれるものと思われる．

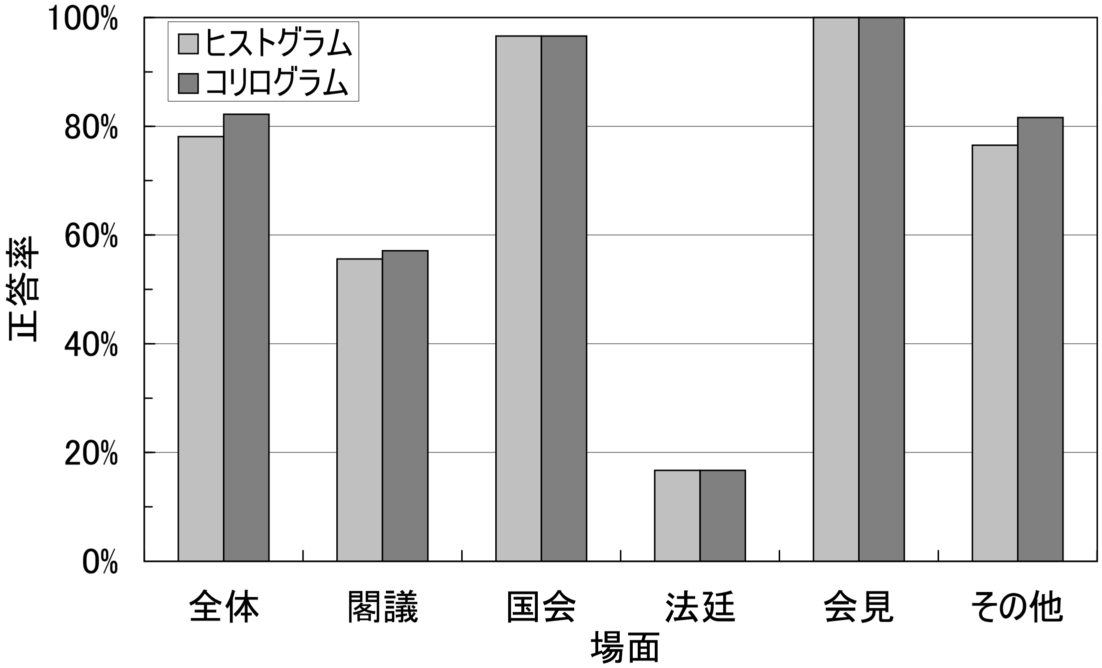


図 6.3: 人物索引付け正答率（1 位率，字幕属性：正答）

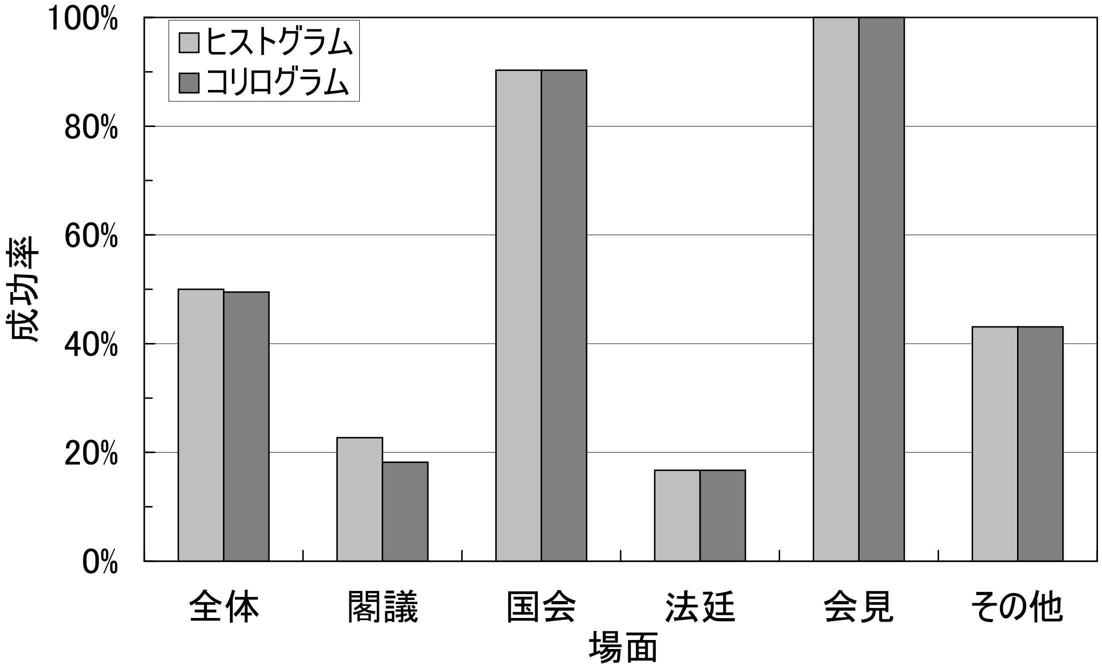


図 6.4: 人物索引付け成功率（1 位率，字幕属性：正答）

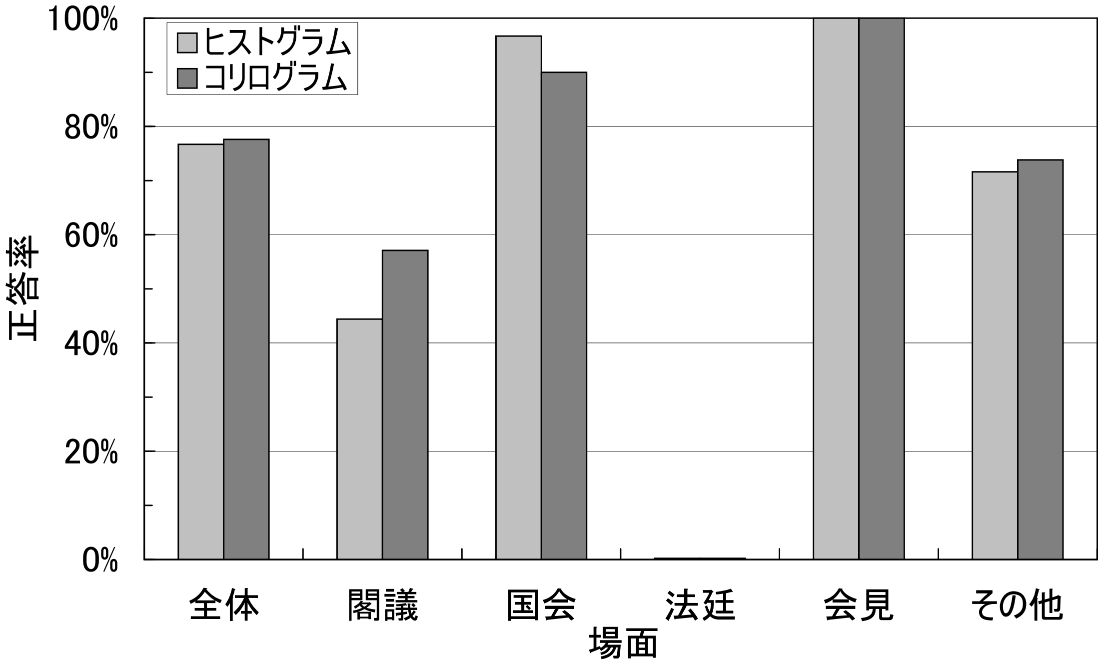


図 6.5: 人物索引付け正答率（1 位率，字幕属性：解析結果）

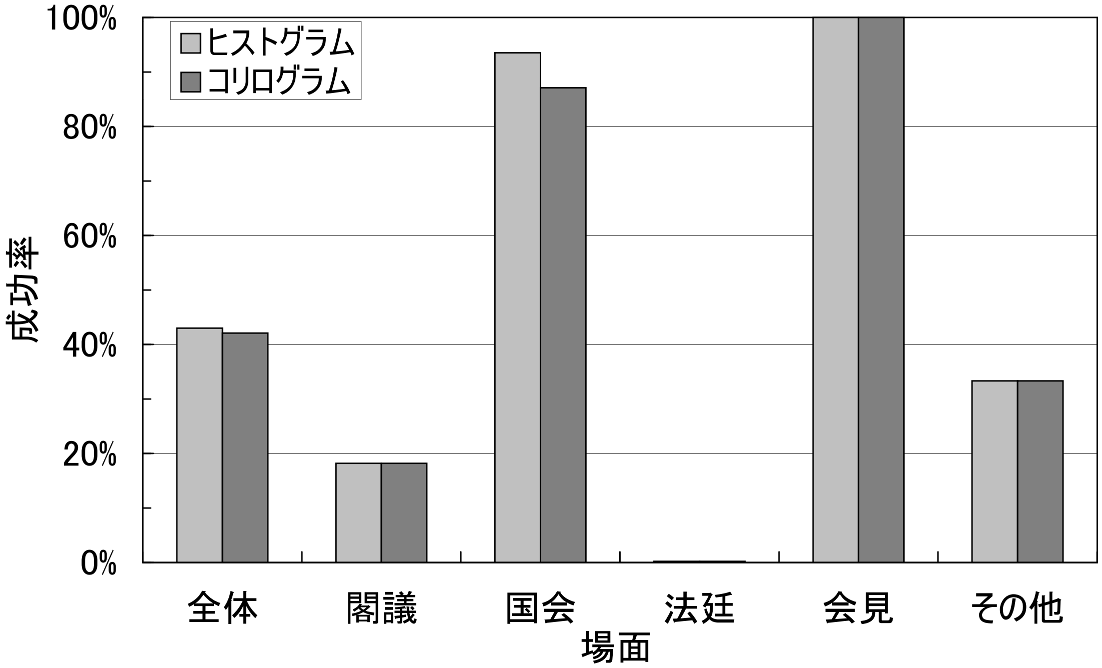


図 6.6: 人物索引付け n 位成功率（1 位率，字幕属性：解析結果）

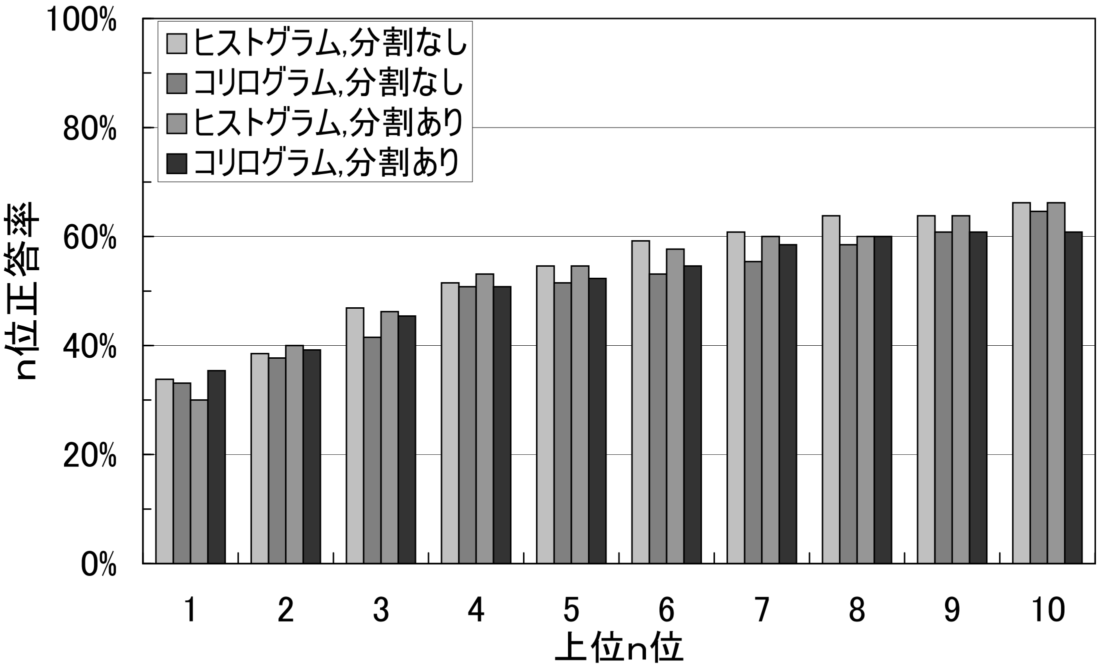


図 6.7: 場面索引付け上位 n 位正答率（全体，字幕属性：正答）

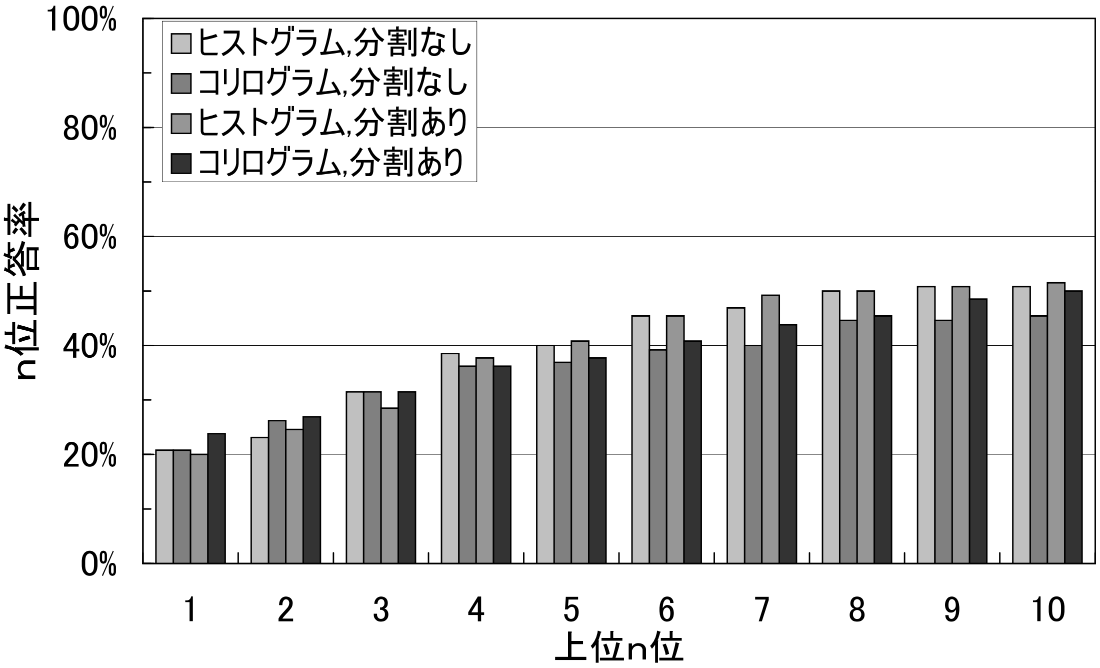


図 6.8: 場面索引付け上位 n 位正答率（全体，字幕属性：解析結果）

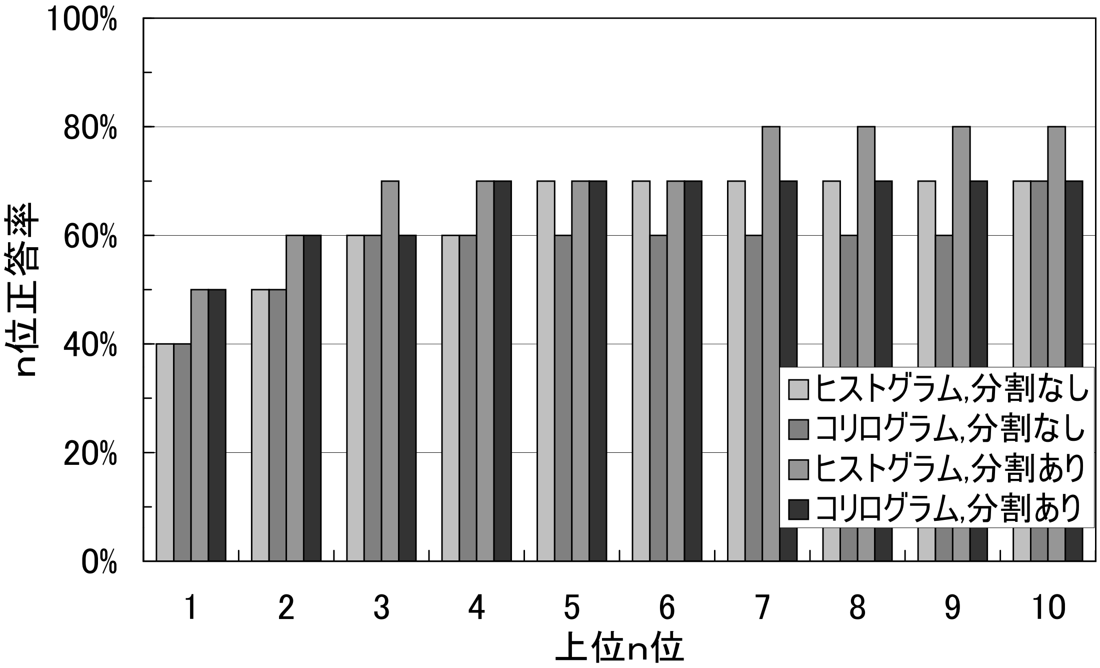


図 6.9: 場面索引付け上位 n 位正答率（閣議控室：人物なし，字幕属性：正答）

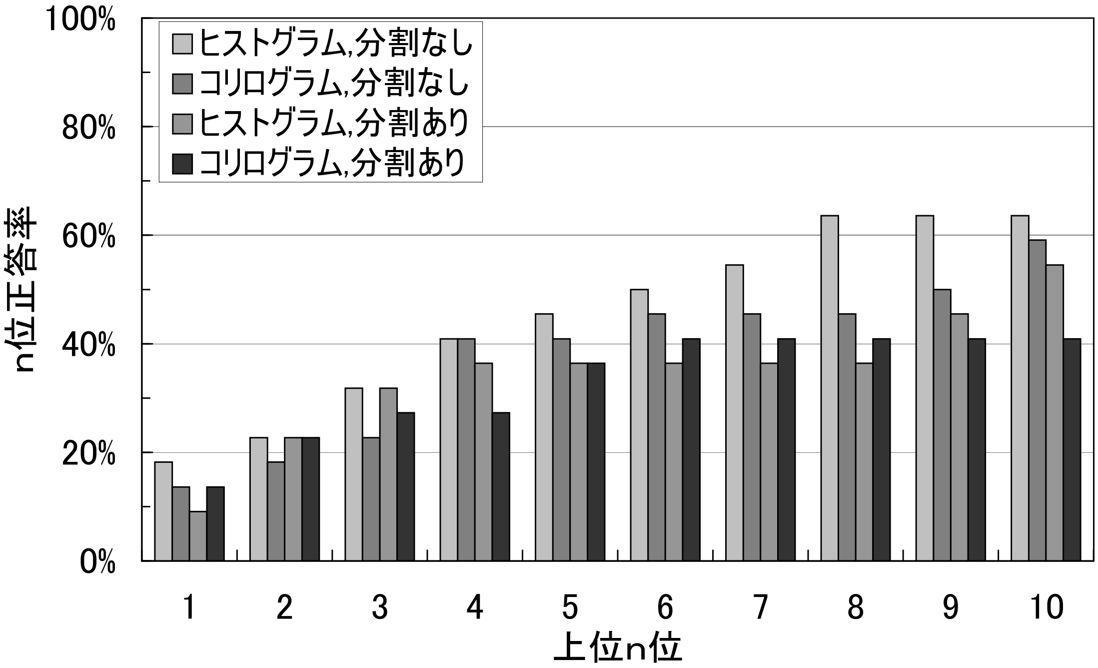


図 6.10: 場面索引付け上位 n 位正答率（閣議控室：人物あり，字幕属性：正答）

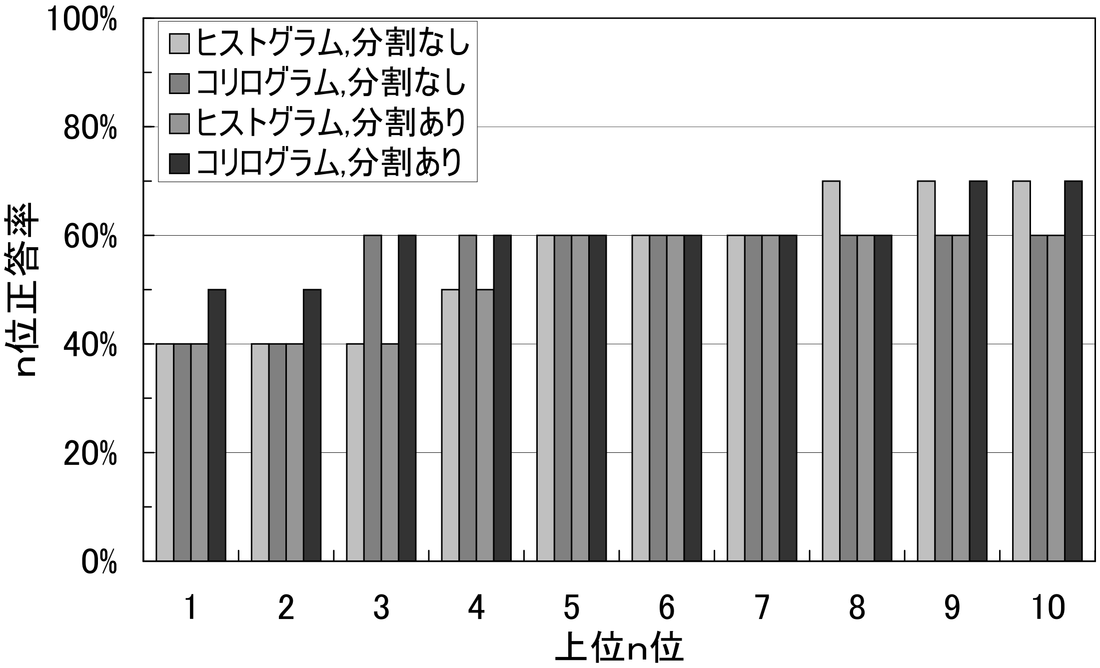


図 6.11: 場面索引付け上位 n 位正答率（閣議控室：人物なし，字幕属性：解析結果）

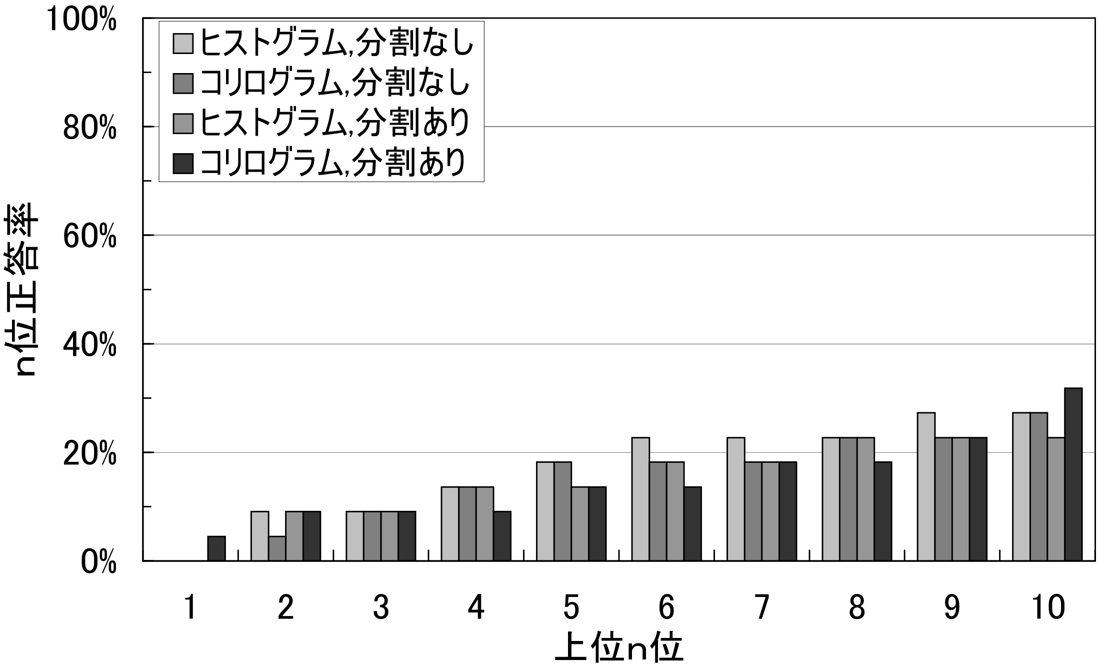


図 6.12: 場面索引付け上位 n 位正答率（閣議控室：人物あり，字幕属性：解析結果）

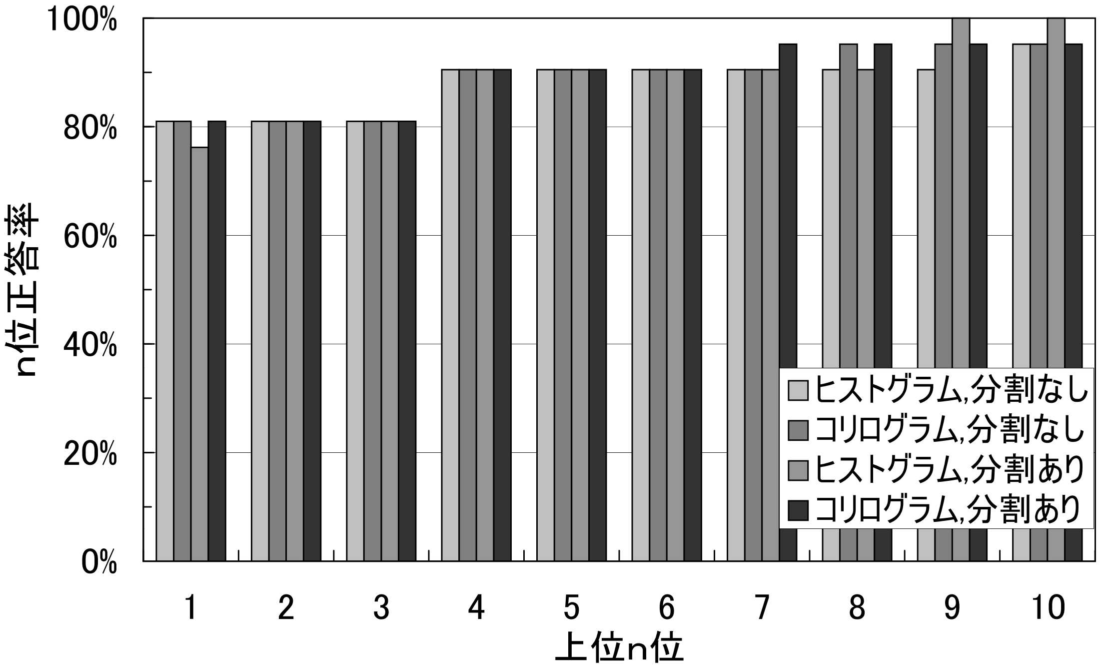


図 6.13: 場面索引付け上位 n 位正答率（国会議場：人物なし，字幕属性：正答）

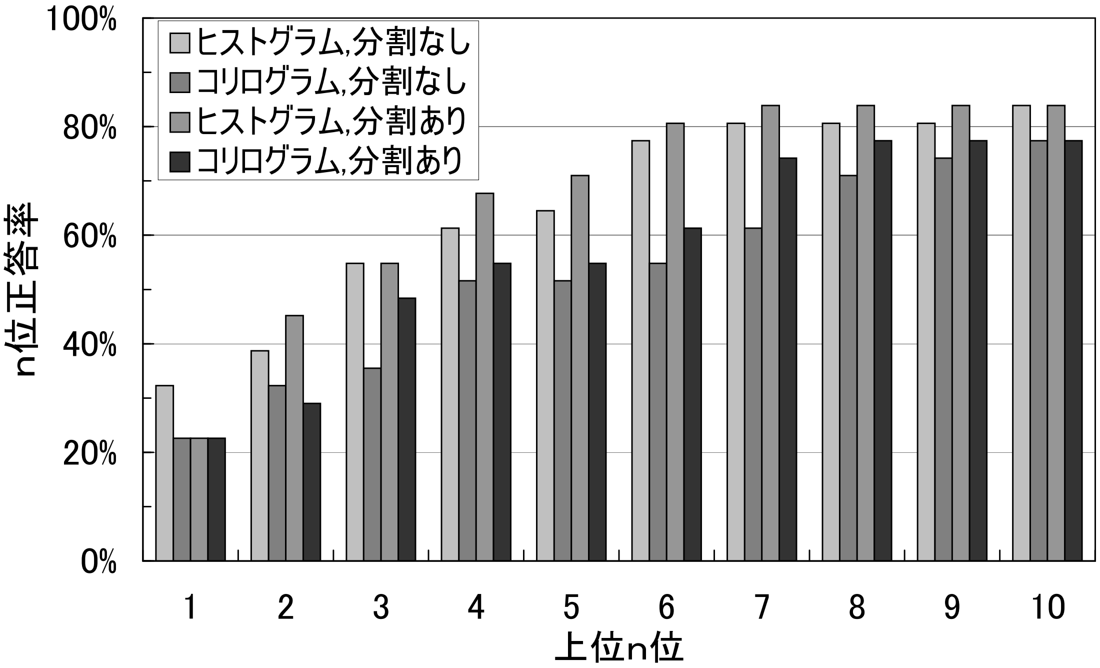


図 6.14: 場面索引付け上位 n 位正答率（国会議場：人物あり，字幕属性：正答）

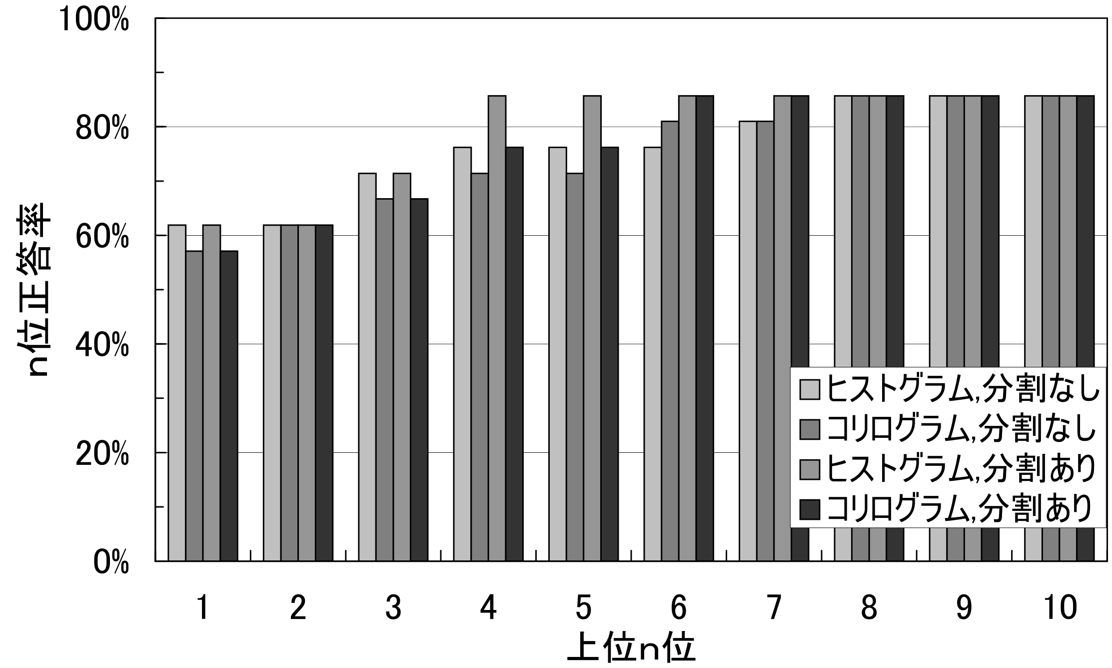


図 6.15: 場面索引付け上位 n 位正答率（国会議場：人物なし，字幕属性：解析結果）

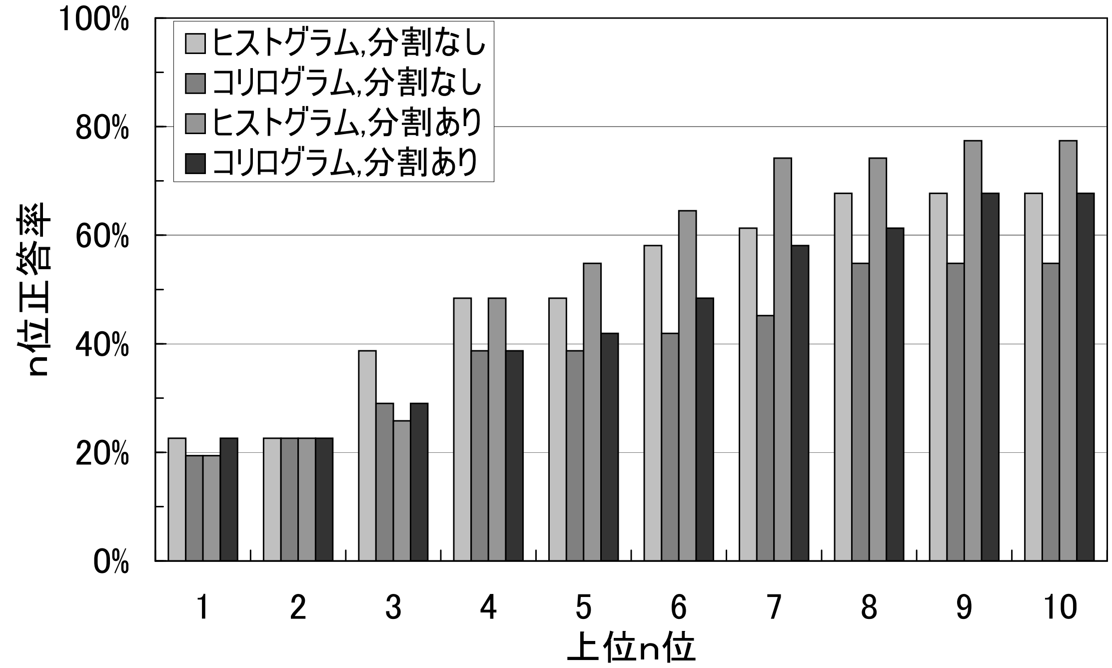


図 6.16: 場面索引付け上位 n 位正答率（国会議場：人物あり，字幕属性：解析結果）

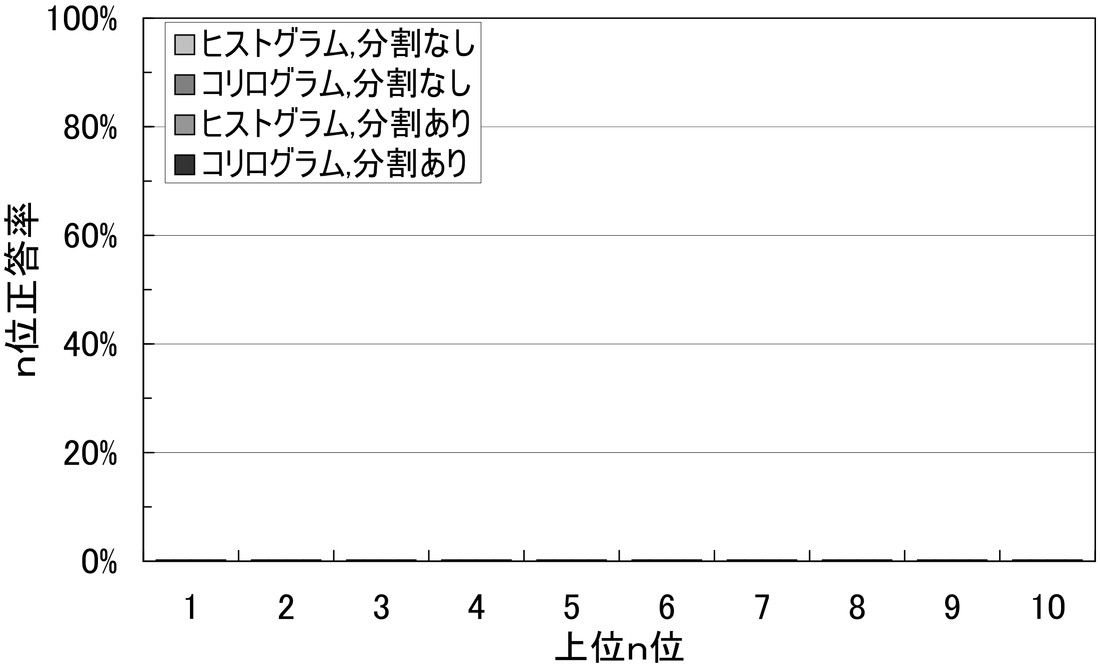


図 6.17: 場面索引付け上位 n 位正答率（報道会見場：人物なし，字幕属性：正答）

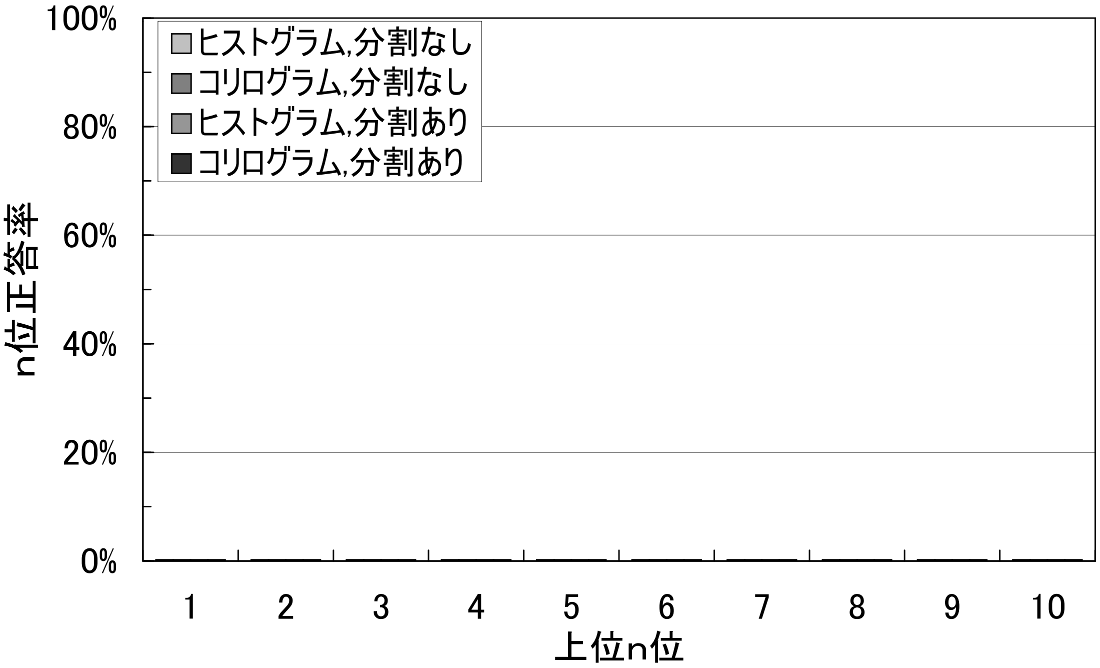


図 6.18: 場面索引付け上位 n 位正答率（報道会見場：人物あり，字幕属性：正答）

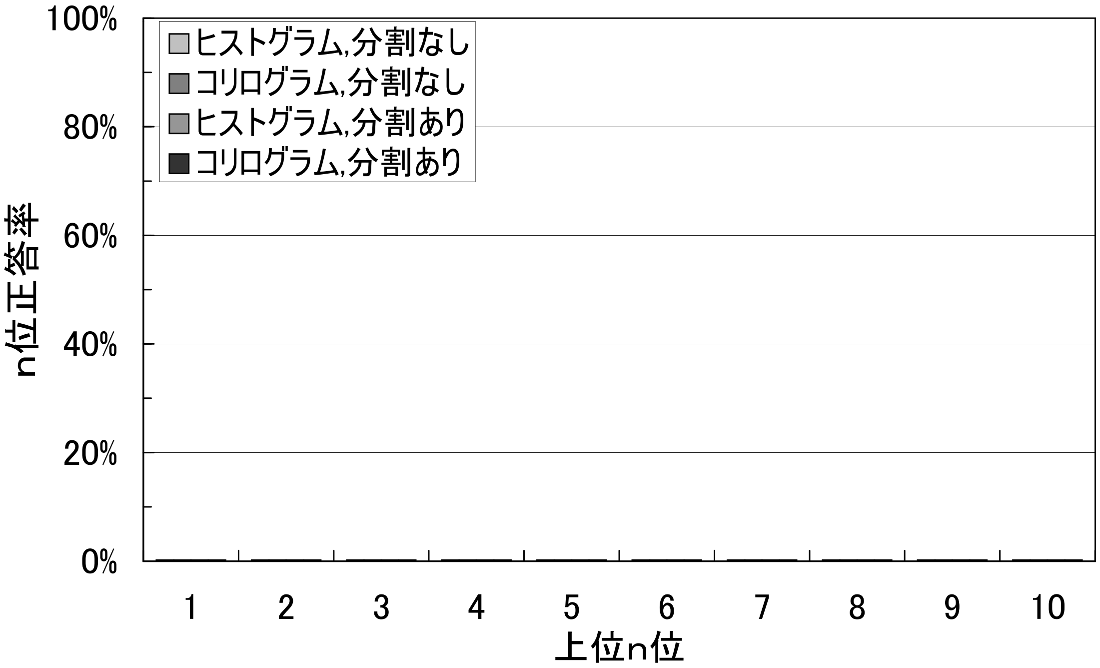


図 6.19: 場面索引付け上位 n 位正答率（報道会見場：人物なし，字幕属性：解析結果）

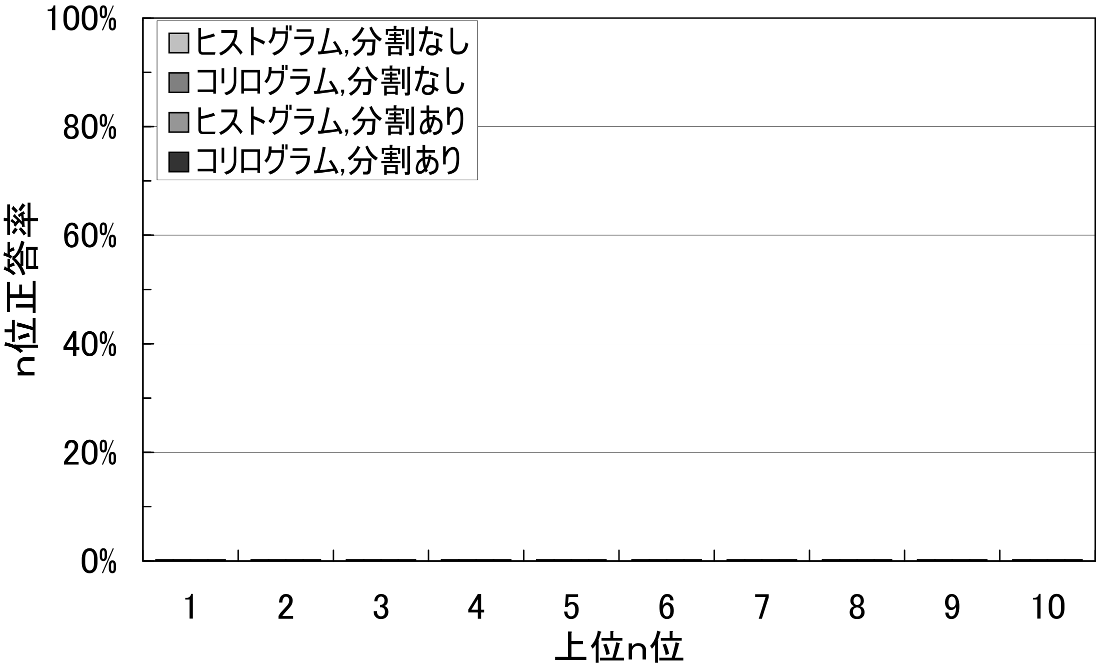


図 6.20: 場面索引付け上位 n 位正答率（報道会見場：人物あり，字幕属性：解析結果）

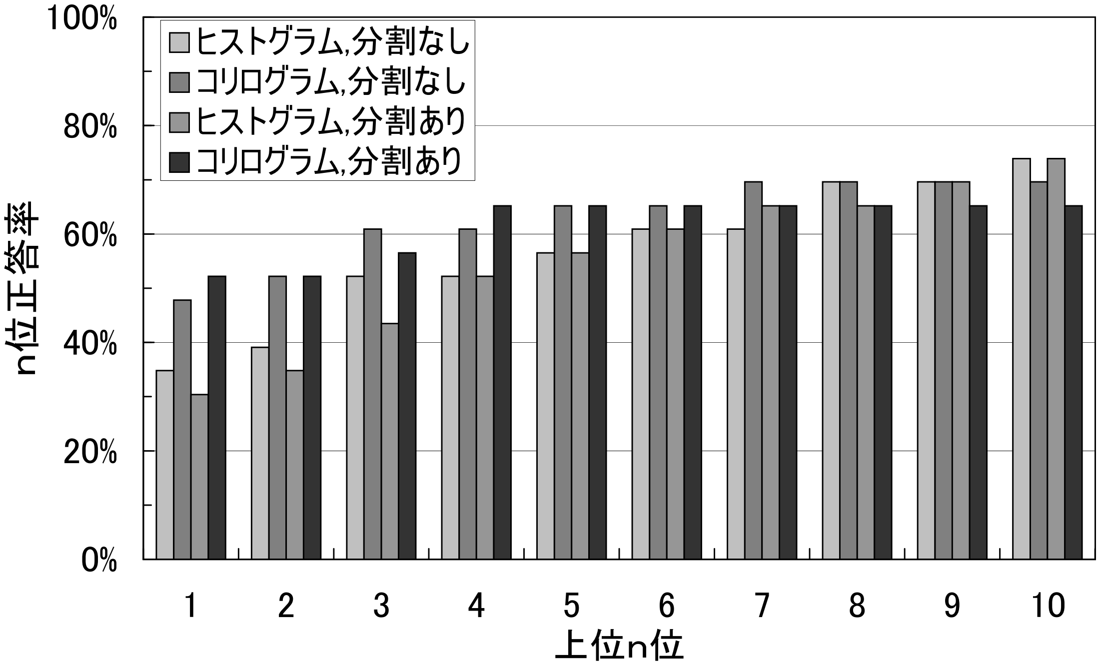


図 6.21: 場面索引付け上位 n 位正答率（法廷：人物なし，字幕属性：正答）

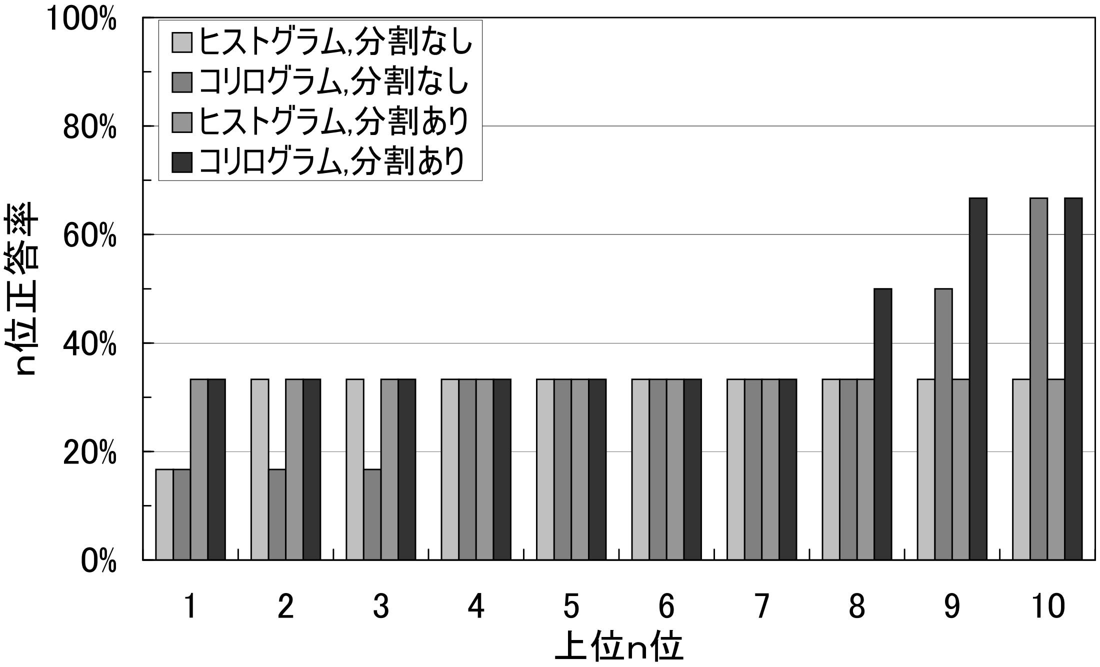


図 6.22: 場面索引付け上位 n 位正答率（法廷：人物あり，字幕属性：正答）

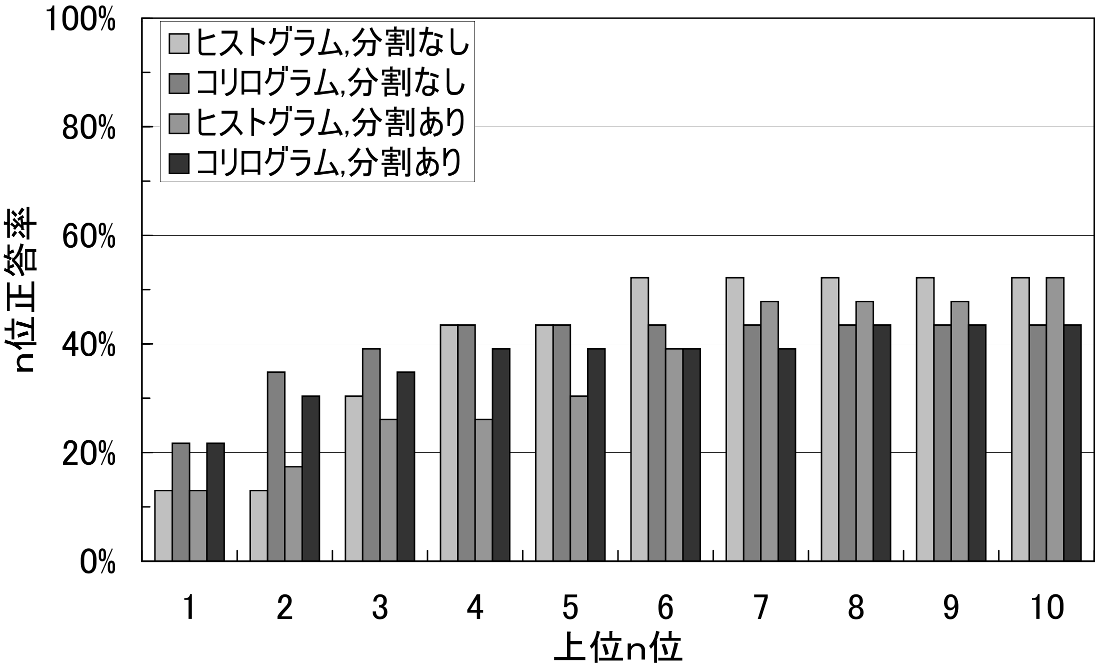


図 6.23: 場面索引付け上位 n 位正答率（法廷：人物なし，字幕属性：解析結果）

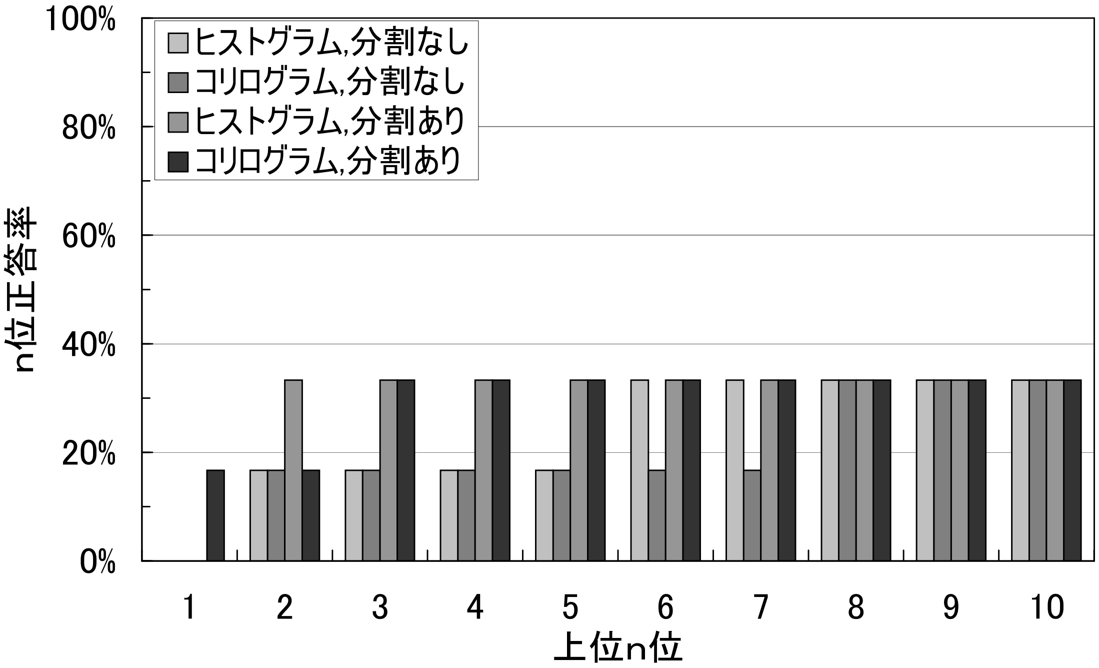


図 6.24: 場面索引付け上位 n 位正答率（法廷：人物あり，字幕属性：解析結果）

6.2 まとめ

本章では，テキストと画像の属性を考慮した最終的な索引付け結果を，人物及び場面に分けて示した．

その結果，人物に関して，同一話題内において画像特徴量の類似度が高いショット中の人物に関する字幕の利用のみで，全体として 8 割程度の 1 位正答率，4 割程度の 1 位成功率が得られた．一方，場面に関して，画像特徴量の類似度が高いショット中の場所・組織に関する字幕の利用により，上位 1 位から 10 位の範囲で，全体として 2 割から 5 割程度の正答率が得られた．

また，便宜上設定した頻出場面別にみると，法廷において場面に関する索引付けが比較的成功している反面，人物に関する索引付けには完全に失敗しているのを筆頭に，場面により出現する字幕の形態が異なったり，顕著な画像特徴が得にくいことが明らかになった．

第 7 章

結論と今後の課題

本章では，以上の議論の結論とそれから導かれる今後の課題をまとめる．

7.1 結論

本節では，本論文で述べた手法の提案と実装及びその評価について簡潔にまとめる．

7.1.1 手法の提案

本論文では，画像内容を考慮したニュース映像への自動索引付け手法を提案し，実現に必要な主要な要素技術を開発し，実装を通じてその評価を行った．実現方法として，統合メディア処理による複数メディアを統合的に処理する手法を採用した．複数メディアのうちテキストメディアとしては字幕を用い，字幕の語義属性解析に必要な独自の語義解析手法を提案し，辞書作成を経て，評価を行った．一方，画像メディアに関しては，画像特徴量と画像内容の対応関係を利用した．ここでは，ルールベースの典型的ショット分類と，事例ベースの類似度判定による画像内容推定の 2 つを行った．さらに，以上の要素技術を用いて，提案手法による索引付けを実施し，性能の評価を行った．

7.1.2 実装と評価

具体的な実験として，以下の 4 つを行った．

- 実験 4.1：ニュース字幕に対する語義属性解析
  高再現率低適合率の結果を得た．
- 実験 5.1：典型的ショット分類
  高適合率低再現率の結果を得た．

- 実験 5.2：背景領域の画像特徴量に基づく場面推定
国内報道関連の場面の推定において，領域分割による場面推定の効果を確認するとともに，対象に応じて用いるべき特徴量の適性があることが分かった．また，事例に基づく推定性能の向上には，指数規模の事例数の拡大が必要であることが明らかになった．
- 実験 6.1：ニュース映像に対する自動索引付け
人物に関する索引付けでは，単純な対応付けでありながら，高い正答率が得られた．一方，場面に関する索引付けでは，1 位正答率は必ずしも高くなかったものの，上位 10 位でみると，比較的高い正答率が得られた．総合的にみて 5 割程度の索引付け成功率が得られた．

これらの結果から総合的に判断して，提案手法により，テキストと画像の属性を考慮して，5 割程度の正答率で索引付けが実現できる可能性が示された．この性能は決して高いものではないが，類似事例や索引候補の情報源の絶対的な量の不足に起因する失敗が多く，指数倍規模での事例の追加や字幕以外のテキスト情報源からの具体的画像内容を反映した索引候補の抽出などにより，実用的水準に達することができるものと考えられる．

7.2 今後の課題

本節では，以上の実験結果から得られた知見に基づき，提案手法の抱える問題点の解決法について，各処理毎にまとめる．これらの課題の多くは，各分野において現在も活発に研究がなされているものであり，それらの進展により実現可能となるものが多い．

その一方で，以下に述べるような解決法を適用してもなお人手を介する必要があるもの，あるいは介する方が効果的である処理もあり，映像製作段階から索引付けを意識した手順を経る必要がある．このような意識は，MPEG4 などの内容に踏み込んだ映像記述方式による映像製作とも無縁ではなく，今後の製作者側の努力にも期待したい．

7.2.1 自然言語処理

- 字幕以外のテキスト情報源の利用
  本手法を欧米のニュース映像など，字幕の出現頻度が少ない映像に適用する際や，実験 6.1 で示されたように，字幕から得られる情報では不十分な場合には，クローズドキャプションや音声など，他のテキスト情報源の利用を検討する必要がある．その際には，既存の文章要約技術 [157, 176, 184] や，話題（タイトル，見出しなど）に応じた重要文あるいは語句抽出手法 [28, 146, 186] を利用することになる．

- より高度な語義属性の解析
  接尾辞のみによらず，名詞句を構成する固有名詞との関連や，共起する字幕間の関係を考慮した総合的な属性解析により，字幕単独での解析性能の向上が期待される．また，語義属性解析の段階から画像解析結果を統合的に利用することも考えられる．

7.2.2 画像処理

- 様々な画像特徴量の導入
  提案機構では様々な特徴量を総合的に扱えるようになっているが，本論文では，一部の特徴量を用いたのみである．表 3.3に挙げた，テクスチャ，エッジ，直線成分など，特徴量の種類を増やすほど弁別可能な画像内容は増えるはずであり，汎用性をもたせるためには，様々な特徴量を導入すべきである．

- 画像特徴量の選択
  実験 5.2 で示されたように，画像内容に応じて解析に適した画像特徴量がある．このような画像特徴量の選別や重み付けの処理を行う必要があり，場合によってはテキストから得られる情報を反映することも検討すべきである．

- 処理の自動化
  本手法の実現に必要な顔領域抽出，領域分割など，現時点では依然自動化が困難な処理がいくつも存在する．それらが自動化されることにより，提案機構の自動化がはかられるため，今後の要素技術の発達や性能向上を期待したい．

- 学習事例数の拡大
  本論文では，国内報道関連の一部の場面のみを対象としたが，提案手法によるより汎用的でロバストな画像内容解析を実現するためには，事例数を増やし，様々な画像内容毎に多様性に対応し得るだけの事例を用意する必要がある．また，実験 5.2 で示されたように，画像内容推定性能は事例数の対数に比例するとみられ，推定性能向上には指数規模で事例数を増やす必要がある．そのために，単純により多くの事例を投入するだけでなく，ショットの代表フレーム以外のすべてのフレームも事例として利用することにより，多様な状況に対応することを考えるべきである．

7.2.3 統合処理

- 各メディア処理の相互補完
  本研究では，各メディア処理の結果を最終段階で統合することにより，単一メディア処理による索引付けの弊害を解消したが，各処理のより早い段階で相互補完を実現するような処理を行う必要がある．具体的には，7.2.1及び 7.2.2で述べたような，属性解析や認識の段階での相互補完が望まれる．

- これからの統合メディア処理
  本研究を含め，従来の統合メディア処理は，既存の研究成果により計算機が扱えるようになったメディアの統合に過ぎなかった．しかし，今後は，人間の五感，すなわち，視覚（画像，テキスト），聴覚（音声・音響，テキスト），嗅覚，味覚，触覚のうち，いまだほとんど取り組まれていない，嗅覚，味覚，触覚及び，これらから二次的に派生するメディアを統合的に扱い，より人間による状況の知覚に近い処理が望まれるようになるだろう．

以上の課題を提示し，本論文の結びとする．

謝辞

本論文にまとめた研究内容は，基本的には標準修了年限を 1 年越えた 4 年間の博士課程在学中に行ったものだが，統合メディア処理による画像への索引付けという点では，修士課程における研究の延長線上にある．従って，本来ならば通算 6 年間にお世話になった全ての方々に謝意を表すべきであるが，修士課程において特にお世話になった方々については修士論文 [1] の謝辞に譲り，ここでは主に博士課程在学中にお世話になった方々への謝辞を述べたい．

指導教官である田中英彦教授には学部 4 年時の卒業論文の時より 7 年間一貫して御指導頂いた．相談会や SIGAP といった定例会合もさることながら，御多忙にもかかわらず幾度も長い時間を割いて個人的に研究の方向性や進路について相談にのって頂いた．御自身の専門とは異なる分野であるにもかかわらず，失礼を顧みず申せば，時に驚くほど鋭い指摘をされ，その造詣の深さに感服したものである．また，研究のみならず，様々な状況において，相当な範囲での自由な行動を認めて下さった寛容さと教育的配慮の深さには感謝の言葉もない．更には，情報処理の分野の発展のために日々奔走される姿には頭が下がるのみである．

一方，後半の 2 年間田中教授と共同で研究室の運営に当たられた坂井修一助教授には，教官としてのみならず，同じ研究室の卒業生として，より身近に様々なことを実践的に教えて頂いた．

両先生には，今後も様々な局面で御迷惑をかけることになろうかと思うが，宜しく御指導頂きたい．

平成 11 年 7 月に行われた予備審査では，御多忙中にもかかわらず，田中英彦教授（電気工学専攻），原島博教授（電子情報工学専攻），坂内正夫教授（電気工学専攻，生産技術研究所），安達淳教授（電気工学専攻，文部省学術情報センター），坂井修一助教授（電気工学専攻），相澤清晴助教授（電子情報工学専攻）の 6 名の先生方に審査して頂き，各々の専門的見地から様々な有益な御指摘を頂いた．

以上の先生方にお世話になった一方で，所属専攻外の以下の方々には日常的な研究上の相談や見学，学会における交流など，様々な局面でお世話になった．自然言語処理や情報抽出の分野では，相澤彰子（文部省学術情報センター研究開発部助教授），柏野和佳子（富士通（株）→国立国語研究所），神門典子（文部省学術情報センター研究開発部助教授），

角田達彦（医科学研究所助手），永松健司（情報工学専攻博士課程→（株）日立製作所中央研究所）の各氏に，画像処理やマルチメディアデータベースの分野では，佐藤真一（文部省学術情報センター研究開発部助教授），佐藤隆（日本電信電話（株）サイバースペース研究所），杉山善明（龍谷大学理工学部研究生），舘村純一（電子工学専攻講師，生産技術研究所），中村裕一（筑波大学電子・情報工学系助教授），堀修（（株）東芝研究開発センター），森靖英（（株）日立製作所中央研究所（技）新情報処理開発機構つくば研究センタ），渡辺靖彦（龍谷大学理工学部助手）の各氏に再三にわたって初歩的な疑問や，研究の方向性に関する質問や相談に付き合って頂いた．これらの方々が与えて下さった知識，情報，そして何よりも同じ分野の研究者としての懇切丁寧な御指導と温かい眼差しなくしては，本研究は成立し得なかったであろう．

また，研究室における共同研究者として，以下の各氏とは日常的に互いに助け合いながら，様々な成果を生み出してきた．

- 平成 8 年度：
  林 Michael 建宇（研究生），上坂英樹（修士課程→日本電気（株））[189]

- 平成 9 年度：
  陳廸傑（学部生→電子情報工学専攻修士課程）[199]，山本晃司（修士課程→（株）東芝研究開発センター）[201]

- 平成 10 年度：
  瀬戸喜巳（学部生→電子情報工学専攻修士課程）[196]，宮崎誠也（学部生→電子情報工学専攻修士課程）[190]，浜田玲子（電気工学専攻修士課程）

- 平成 11 年度：
  佐久間丈貴（学部生）[194]，下萩原勉（学部生）[197]，浜田玲子 [192]

- 研究補助：
  織田友恵（御茶の水女子大学大学院人間文化研究科数理・情報科学専攻修士課程），平博司（学部生→電子情報工学専攻修士課程），中村実（情報工学専攻修士課程）

特に，平成 10 年度より該当年度の共同研究者と共にほぼ毎週開催した私的会合 sigMM（special interest group in Multi-Media の略称）は，実験的試みから始まったが，予想以上に効果的かつ，非常に多くのことを心から楽しみながら学ばせて頂いた．この体験を，研究教育活動に当たる際の原点として，今後の指針としたいと思う．

また，上坂，山本の両氏とは修士課程のうち最後の 1 年間の共同研究ではあったものの，様々な便利なツールやデータの遺産を残して頂き，本論文完成直前まで使用させて頂いた．特に，山本氏の修士論文 [201] における研究は本論文中の実験 5.2 の大部分として利用させていただいた．

そして，浜田氏とは，初めて修士課程 2 年間を通じての共同研究者として，時にあらぬ偏見を抱かれがちな料理映像を題材とした新たな研究分野の開拓を行い，互いに非常に有

益な時間を過ごすことができた．丁度，博士課程最後の 2 年間と重複していたため，本論文の完成にも直接間接に多大なる貢献をして頂き，感謝の言葉もない．

また，研究補助として字幕書き下し，目視によるカット検出，映像のディジタル化及び記録に骨身を惜しまず多大な労力を費して下さった研究補助の方々にも深く感謝する．

研究室所属の職員であった，古宇田フミ子（電子情報工学専攻助手），清水修（電気工学専攻助手），菅原利枝（事務補佐員），田丸あや（同），小林美智（同）の各氏には日々の事務手続きや研究室運営において，陰に日向に，お世話になった．

清水氏は，研究室の備品の整備や改造により，居住性の向上に尽力して下さった．また，ご自宅で採れた様々な作物を差入れて下さり，とかく時節を忘れがちな研究室生活に季節感を与えて下さった．

田丸氏には，旅費などの支払を中心とした事務手続きのみならず，個人的な相談にも度々時間を割いて頂いた．

修士課程 1 年時に博士課程 3 年であった角田達彦（医科学研究所助手），毛利隆夫((株）富士通研究所），柏野邦夫（日本電信電話（株）コミュニケーション科学基礎研究所）の 3 氏には，博士課程修了後も研究の節目節目に相談にのって頂き，その都度親身に的確な指導をして頂いた．

そもそも本研究の初期の方向付けは角田氏の助言によるところが大きく，その先見の明に感銘を受けている．また，毛利氏には論文の最終段階において，泊りがけで押しかけた際に，研究のまとめ方について忙しい中を夜を徹して共に考えて下さり，このような先輩の存在を幸せに思う．

研究生や留学生の方々には，他の学生とは異なる刺激を非常に強く受け，彼らとの日常的な接触を通じて多くのことを学んだ．とりわけ，顧乃杰（中国科学技術大学），王明会（北京郵電大学からの研究員），李盛熲（情報工学専攻修士課程），林 Michael 建宇（当時，University of California, Berkeley からの学部交換留学生），Alexandre Essayan（Crédit Lyonnais）の各氏とは多くの時間を費して各国の文化・社会について語り合い，一部の方々とは帰国後も折に触れて交流が続いている．

研究室で博士課程の 3 年目まで同期であった，荒木拓也（情報工学専攻博士課程→日本電気（株）C&C メディア研究所），中村友洋（電気工学専攻博士課程→（株）日立製作所中央研究所），Behrens, Harry，渡辺正泰（情報工学専攻博士課程→日本電気通信システム（株）技術本部），4 年目まで同期であった，五十嵐健夫（情報工学専攻博士課程），吉瀬謙二（同），木下智義（同），馬場恒彦（電気工学専攻博士課程），Magnaghi, Antonio（同）の各氏とは，様々な局面において公私にわたって交流を深めた．

以上に触れた方々以外にも，博士課程在学中に居室のあった工学部 13 号館に在籍した以下の各氏との思い出は枚挙に遑がないが，ここでは氏名を列挙するにとどめる．

青木秀行	五十嵐健夫	石井円力	王明会
内田友幸	Essayan, Alexandre	大久保千英	春日和之
鎌野篤	上坂英樹	金澤輝一	木下智義
顧乃杰	佐久間丈貴	下萩原勉	白石知之
瀬戸喜巳	陳廸傑	Chantngarm, Peerasak	永松健司
野尻大祐	浜田玲子	半田伊吹	Bateman, Gregory
Behrens, Harry	本多正明	宮崎誠也	武藤誠
村岡秀哉	山本晃司	李盛頴	林 Michael 建宇
林國恩			

学部時代からの同期である，熊田亜紀子（電気工学専攻博士課程→新領域創成科学研究科先端エネルギー工学専攻助手），児玉和也（電子情報工学専攻博士課程→文部省学術情報センター研究開発部助手），高瀬亮（電子情報工学専攻博士課程），日高宗一郎（電子情報工学専攻博士課程→文部省学術情報センター研究開発部助手），森野博章（電子情報工学専攻博士課程→日本学術振興会リサーチアソシエート）の各氏とは，博士課程在学中多くの時間を共に過ごした．

児玉，高瀬，日高の 3 氏とはティーチングアシスタントとして計算機管理やレポート採点で苦楽を共にした．森野氏とは更に遡り，大学入学時からの同窓で，既に 10 年もの付き合いになる．彼らとは非常に近い分野の研究者として今後も互いに切磋琢磨していきたいと思う．

熊田氏とは，全く異なる分野の研究に携わりながらも，互いに研究室が入居している工学部 13 号館内における実質的に唯一の同期生として，研究室外の友人としては最も長い時間を共に過ごした．その中で，氏の研究に対する集中力には幾度となく秘かに感嘆して刺激を受けたものである．また，一人旅の楽しみ方や，トーマス・ハーディ作「ダーバヴィル家のテス」を始めとする様々な良書を教わり，旅行の仕方や読書の幅を広げて下さった．一方で，落葉掃き，鼠退治など様々な活動を共にし，同研究室助手であった千葉政邦氏（工学系研究科助手）とも様々な珍味を肴に杯を交わし合った．

一方，4 年にわたり家庭教師をさせて頂き，その後も家族の一員のように受け入れて下さっている齊藤忠夫，万里子夫妻，兄弟姉妹のように慕って下さる，Junko Saito Cattelain，恭子，教え子の玲子の各氏との不定期ながら絶え間ない交流は，これまでも，そしてこれからも，非常に大きな意味をもつものであり，大切にしていきたい．特に，Junko 氏とは，氏の渡仏以来今日に至るまで隔日のように交わした電子メールの数はのべ 700 余通にも達し，人生のあらゆる問題に関して率直に語り合った経験は何物にも代えがたいものである．

大学入学時以来の付き合いである小林知樹氏（国際協力事業団）は修士課程修了後，青年海外協力隊員の理数科教師としてタンザニア連合共和国イリンガ市へ赴任した．氏を訪ったアフリカの大地で見聞きしたものは，それまで知っていた西欧諸国や北東アジアの各国とは良くも悪しくも全く異質なものであった．僅かな期間に垣間見たに過ぎないが，自分の研究の意義についても考えさせられ，このような経験をする機会を与えてくれた小林氏に感謝したい．

書家の冨永秋山師に師事して3年になるが，日々電子化されたテキストや活字に埋もれて暮らしている者にとって，研究の合間を縫っての稽古は新鮮かつ良い気分転換であった．
数年来気管支炎の治療にあたって下さっている藤澤道夫氏（保健センター内科医師）とは医師と患者としての信頼を築けたこともさることながら，ときに医学と工学の世界の違いについて話し合い，大変興味深かった．
また，年間50回以上の美味しい食事を作って下さったフレンディ及びダージリンの皆さんには，食事という生活面で非常に重要な要素を愉しいものにして下さった．

顧みれば，これまでに幸運にも巡り合えた数多くの良き先達や書物から得た有形無形の知識や思想は計りしれない．また，残念ながら具体的な知識の伝授こそなかったものの，旧帝国海軍航空工廠の技術士官であった祖父の鈴木健士との幼少時からの接触は，工学者としての基本的思想の形成に多大な影響があったように思う．今後，研究者あるいは教育者として生きていくにあたり，このような影響を次世代の後輩達に及ぼすことの重要性を絶えず肝に銘じて実践していきたいと思う．
最後に，幼少時よりこのような分野に関わる様々なきっかけや環境を与えてくれ，20数年の長きにわたる学生生活を経済的，心理的に支えてくれた両親，英文学から特殊教育学に転じてますます興味深い話題を提供してくれる妹の幸子，そしていつも無言で側にいてくれた，今は亡き飼い犬のハニーに深く感謝する．

平成11年12月17日

文献一覧

発表文献

学位論文

[1] 井手一郎：
“画像内容の記述に基づく階層的キーワードを用いた画像データベースシステムの提案”，
修士論文，東京大学大学院工学系研究科（Jan 1996）．

雑誌論文・書籍

[2] 井手一郎，田中英彦：
“末尾の名詞に着目したテレビニュース字幕の語義解析”，
情報処理学会論文誌 Vol.39，No.8，pp.2543-2546（Aug 1998）．

[3] Ichiro Ide，Koji Yamamoto，Hidehiko Tanaka：
“Automatic Video Indexing Based on Shot Classification”，
Advanced Multimedia Content Processing –First Intl. Conf. AMCP’98, Osaka, Japan–，Nishio, S.，Kishino, F. eds.，*Lecture Notes in Computer Science* Vol.1554，pp.87-102，Springer-Verlag（Mar 1999）．

[4] 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”，
電子情報通信学会論文誌（D-II）Vol.J82-D-II，No.10，pp.1543-1551（Oct 1999）．

[5] 井手一郎，浜田玲子，坂井修一，田中英彦：
“接尾名詞に着目したテレビニュース字幕の語義属性解析”，
情報処理学会論文誌（投稿中）．

国際会議・ワークショップ

[6] Ichiro Ide，Reiko Hamada，Hidehiko Tanaka，Shuichi Sakai：
"News Video Classification Based on Semantic Attributes of Captions"，
Proc. Sixth ACM Intl. Multimedia Conf. (ACM-MM'98) –Art Demos, Technical Demos, Poster Papers–，pp.60-61，Intellect（於：Univ. of Bristol［Bristol, UK］，Sep 1998）．

[7] Ichiro Ide，Hidehiko Tanaka：
"Automatic Semantic Analysis of Television News Captions"，
Proc. Third Intl. Workshop on Information Retrieval with Asian Languages (IRAL'98)，pp.56-60（於：Kent Ridge Digital Labs［Singapore］，Oct 1998）．

[8] Ichiro Ide，Koji Yamamoto，Hidehiko Tanaka：
"Automatic Video Indexing Based on Shot Classification"，
Proc. First Intl. Conf. on Advanced Multimedia Content Processing (AMCP'98)，pp.99-114（於：大阪大学国際会議場，Nov 1998）．

[9] Ichiro Ide，Reiko Hamada，Shuichi Sakai，Hidehiko Tanaka：
"Relating Graphical Features with Concept Classes for Automatic News Video Indexing"，
Proc. IJCAI-99 Workshop on Intelligent Information Integration，pp.94-98（於:Stockholm City Conference Center［Stockholm, Sweden］，Jul 1999）．

[10] Ichiro Ide，Reiko Hamada，Shuichi Sakai，Hidehiko Tanaka：
"Identification of Scenes in News Video from Image Features of Background Region"，
First Intl. Workshop on Multimedia Intelligent Storage and Retrieval Management (MISRM'99)（於：Grosvenor Resort［Orlando FL, USA］，Nov 1999）．

[11] Ichiro Ide，Reiko Hamada，Shuichi Sakai，Hidehiko 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（於：臺灣中央研究院［臺北, 臺灣］，Nov 1999）．

[12] Reiko Hamada，Ichiro Ide，Shuichi Sakai，Hidehiko Tanaka：
"Associating Video with Related Documents"，
Proc. Seventh ACM Intl. Multimedia Conf. (ACM-MM'99)，Vol.2，pp.7-11（於：Grosvenor Resort［Orlando FL, USA］，Nov 1999）．

[13] Reiko Hamada，Ichiro Ide，Shuichi Sakai，Hidehiko Tanaka：
"Structural Analysis of Preparation Steps on Supplementary Documents of Cultural TV Programs"，
Proc. Fourth Intl. Workshop on Information Retrieval with Asian Languages

(IRAL'99)，pp.43-47（於：臺灣中央研究院［臺北, 臺灣］, Nov 1999）．

研究会・シンポジウム

[14] 井手一郎，浜田玲子，坂井修一，田中英彦：
"言語情報を伴う画像の画像的特徴量と語義の統計的対応付け"，
情報処理学会コンピュータビジョンとイメージメディア研究会報告 99-CVIM-114，Vol.99，No.3，pp.137-143（於：大阪大学銀杏会館，Jan 1999）．

[15] 井手一郎，浜田玲子，坂井修一，田中英彦：
"ニュース映像における人物・背景領域を分割した特徴量解析による内容推定"，
第 5 回電子情報通信学会知能情報メディアシンポジウム（IIM'99），pp.45-51（於：慶應義塾大学三田校舎，Dec 1999）．

[16] 浜田玲子，井手一郎，坂井修一，田中英彦：
"料理番組とテキスト教材の対応付け"，
第 5 回電子情報通信学会知能情報メディアシンポジウム（IIM'99），pp.69-74（於：慶應義塾大学三田校舎，Dec 1999）．

全国大会

[17] 井手一郎，田中英彦：
"マルチメディア・データベースに要求される画像の描写に関する研究"，
第 51 回情報処理学会全国大会 No.1S-8，Vol.3，pp.303-304，大会奨励賞受賞（於：富山大学，Sep 1995）．

[18] 井手一郎，田中英彦：
"画像内容の記述に基づいた階層的キーワードの生成と利用に関する研究"，
第 52 回情報処理学会全国大会 No.6Q-3，Vol.4，pp.321-322（於：電気通信大学，Mar 1996）．

[19] 井手一郎，田中英彦：
"画像・言語情報の統合的利用による映像データの自動的インデクシングの試み"，
第 3 回言語処理学会年次大会 No.A5-2，pp.485-488（於：京都大学工学部，Mar 1997）．

[20] 井手一郎，田中英彦：
"末尾の名詞に着目した TV 放送中の字幕の意味属性解析"，
第 55 回情報処理学会全国大会 No.6J-4，Vol.2，pp.91-92（於：福岡工業大学，Sep 1997）．

[21] 井手一郎，山本晃司，田中英彦：
"ショットの分類に基づく映像データへの自動的索引付け"，
第 56 回情報処理学会全国大会 No.5Q-6，Vol.2，pp.263-264（於：中央大学理工学部，Mar 1998）.

[22] 井手一郎，山本晃司，田中英彦：
"字幕の語義に基づくニュース映像の分類・索引付け手法"，
1998 電子情報通信学会総合大会 No.SD-3-4，Vol. 情報・システム 1，pp.403-404（於：東海大学湘南校舎，Mar 1998）.

[23] 井手一郎，浜田玲子，田中英彦，坂井修一：
"字幕の語義に基づくニュース映像分類"，
第 57 回情報処理学会全国大会 No.4V-4，Vol.3，pp.259-260（於：名古屋大学工学部，Oct 1998）.

[24] 井手一郎，浜田玲子，坂井修一，田中英彦：
"ニュース映像における人物領域と背景領域の個別特徴量解析による人物と場面の追跡・索引付け手法"，
第 59 回情報処理学会全国大会 No.5M-8，Vol.2，pp.311-312（於：岩手県立大学，Oct 1999）.

[25] 山本晃司，井手一郎，田中英彦：
"ニュース映像中のショット分類に関する研究"，
第 55 回情報処理学会全国大会 No.4K-7，Vol.2，pp.160-161（於：福岡工業大学，Sep 1997）.

[26] 山本晃司，井手一郎，田中英彦：
"言語情報を基準にした学習による映像分類手法"，
1998 電子情報通信学会総合大会 No.D-12-69，Vol. 情報・システム 2，p.268（於：東海大学湘南校舎，Mar 1998）.

[27] Tan Tit-Keat，井手一郎，田中英彦：
"副音声利用による映像への索引付け"，
第 12 回人工知能学会全国大会 No.23-2，pp.362-365（於：早稲田大学国際会議場，Jun 1998）.

[28] 瀬戸喜巳，井手一郎，坂井修一，田中英彦：
"見出しからの制約による新聞記事からの重要文抽出"，
第 58 回情報処理学会全国大会 No.1E-9，Vol.2，pp.73-74（於：早稲田大学，Mar 1999）.

[29] 浜田玲子，井手一郎，坂井修一，田中英彦：
"料理番組における映像とテキスト情報の対応付け"，

第 58 回情報処理学会全国大会 No.4U-7，Vol.3，pp.191-192（於：早稲田大学，Mar 1999）．

[30] 宮崎誠也，井手一郎，坂井修一，田中英彦：
"対象の構成に着目した画像認識手法による画像検索"，
第 58 回情報処理学会全国大会 No.5D-8，Vol.2，pp.51-52（於：早稲田大学，Mar 1999）．

[31] 浜田玲子，井手一郎，坂井修一，田中英彦：
"教養番組のテキスト教材における手順の構造化"，
第 59 回情報処理学会全国大会 No.2N-7，Vol.2，pp.353-354（於：岩手県立大学，Oct 1999）．

[32] 佐久間丈貴，井手一郎，坂井修一，田中英彦：
"出現単語の概念関係を用いたテキストの誤り訂正"，
第 60 回情報処理学会全国大会 No.2K-3，Vol.2（於：拓殖大学八王子校舎，Mar 2000 発表予定）．

[33] 下萩原勉，浜田玲子，井手一郎，坂井修一，田中英彦：
"テンプレート切出しによる不特定話者対応のワードスポッティング"，
第 60 回情報処理学会全国大会 No.2L-1，Vol.2（於：拓殖大学八王子校舎，Mar 2000 発表予定）．

[34] 浜田玲子，井手一郎，坂井修一，田中英彦：
"料理素材における手順の構造化"，
第 60 回情報処理学会全国大会 No.1F-5，Vol.2（於：拓殖大学八王子校舎，Mar 2000 発表予定）．

参考文献

論文

[101] Aoki, H.，Shimotsuji, S.，Hori, O.："A Shot Classification Method of Selecting Effective Key-Frames for Video Browsing"，*Proc. Fourth ACM Intl. Multimedia Conf.*，pp.1-10（Nov 1996）．

[102] 青木　恒，堀　修："繰り返しショットの統合による階層化アイコンを用いたビデオ・インターフェース"，情処学論，Vol.39，No.5，pp.1317-1324（May 1998）．

[103] 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出し"，信学論（D-II），Vol.J80-D-II，No.9，pp.2421-2427（Sep 1997）．

[104] 有木康雄，杉山善明，石川則之，寺西俊裕，櫻井光康："ニュース映像中の記事に対する音声・文字・映像を用いた索引付けと分類"，信学技報 PRMU-96，Vol.96，No.385，pp.31-38（Nov 1996）．

[105] Brown, M. G.，Foote, J. T.，Jones, G. J. F.，Jones, K. S.，Young，S. J.："Automatic Content-Based Retrieval of Broadcast News"，*Proc. Third ACM Intl. Multimedia Conf.*，pp.35-43（Nov 1995）．

[106] Brunelli, R.，Roggio, T.："Face Recognition: Features versus Templates"，*IEEE Trans. on Pattern Analysis and Machine Intelligence*，Vol.15，No.10，pp.1042-1052（Oct 1993）．

[107] Chang, Y. L.，Zeng, W.，Kamel, I.，Alonso, R.："Integrated Image and Speech Analysis for Content-Based Video Indexing"，*Proc. IEEE Intl. Conf. on Multimedia Computing and Systems '96*（Jun 1996）．

[108] Chinchor, N.："MUC-7 Named Entity Task Definition, Version 3.5"，*Proc. Seventh Message Understanding Conf.*（Sep 1997）．

[109] Fischer, S.，Lienhart, R.，Effelsberg, W.："Automatic Recognition of Film Genres"，*Proc. Third ACM Intl. Multimedia Conf.*，pp.367-368（Nov 1995）．

[110] 藤田悦郎，安部伸治，小倉健司："パターン情報から言語表現を連想想起する手法および連想モデル獲得手法　–映像へのインデクス生成を目指して–"，人工知能学研報 SIG-J-9702，pp.39-44（Dec 1997）．

[111] Govindaraju, V.，Srihari, S. N.，Sher, D.："A Computational Model for Face Location Based on Cognitive Principles"，*Proc. Tenth National Conf. On Artificial Intelligence*，pp.350-355（Jul 1992）．

[112] 顧　海松，浅田　稔，白井良明：“動き情報に基づくエッジセグメントの最適分割”，信学論（D-II），Vol.J76-D-II，No.8，pp.1544-1553（Aug 1993）．

[113] 平田恭二，原　良憲：“概略画像を用いた画像検索”，信学技報 DE-92，Vol.92，No.51，pp.9-16（May 1992）．

[114] 堀　修：“統合的手法を用いた様々な背景からの顔領域の抽出”，情処学研報 97-CVIM-106，Vol.97，No.70，pp.139-146（Jul 1997）．

[115] 堀　修：“テロップ認識のための映像からの文字部抽出法”，情処学研報 99-CVIM-114，Vol.99，No.3，pp.121-136（Jan 1999）．

[116] Hsu, W.，Chun, T. S.，Pung, H. K.：“An Integrated Color-Spatial Approach to Content-Based Image Retrieval”，*Proc. Third ACM Intl. Multimedia Conf.*，pp.305-313（Nov 1995）．

[117] Huang, J., Kumar, S. R.，Mitra, M.，Zhu, W. J.，Zabih, R.：“Image Indexing Using Color Correlograms”，*Proc. IEEE Conf. on Computer Vision and Pattern Recognition '97*，pp.762-768（Jun 1997）．

[118] Huang, J.，Kumar, S. R.，Zabih, R.：“An Automatic Hierarchical Image Classification Scheme”，*Proc. Sixth ACM Intl. Multimedia Conf.*，pp.219-228（Sep 1998）．

[119] 糸井清晃，久保田哲也，小林幸雄：“見出し文字列の地紋除去”，信学論（D-II），Vol.J82-D-II，No.4，pp.763-770（Apr 1999）．

[120] 和泉直樹，柳沼良知, 中川裕志，坂内正夫：“シナリオ文書の解析による存在・行動マップの作成”，信学論（D-II），Vol.J79-D-II，No.11，pp.1993-1996（Nov 1996）．

[121] 金出武雄，佐藤真一：“Informedia： CMU ディジタルビデオライブラリプロジェクト”，情処学誌，Vol.37，No.9，pp.841-847（Sep 1996）．

[122] 金子敏充，堀　修：“尤度比検定による MPEG 圧縮動画像からのカット検出”，情処学研報 98-CVIM-109，Vol.98，No.5，pp.105-112（Jan 1998）．

[123] Kass, M.，Witkin, A.，Terzopoulos, D.：“Snakes: Active Contour Models”，*Intl. J. of Computer Vision*，Vol.1，No.4，pp.321-331（Jan 1988）．

[124] 加藤俊一，下垣弘之，藤村是明：“画像対話型商標・意匠データベース TRADE-MARK”，信学論（D-II），Vol.J72-D-II，No.4，pp.535-544（Apr 1989）．

[125] 君山博之，清末悌之，大庭有二：“動画像の自動記述方法の検討 –動画像構造化へのアプローチ–”，信学技報 IE-91，Vol.91，No.419，pp.25-32（Jan 1992）．

[126] 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（Mar 1999）．

[127] 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）.

[128] Kuwano, H.，Kurakake, S.，Odaka, K.：“Telop Character Extraction from Video Data”，*Proc. Workshop on Document Image Analysis '97*，pp.82-88（Jun 1997）.

[129] Ma, W. A.，Manjunath, B. S.：“Edge Flow: A Framework of Boundary Detection and Image Segmentation”，*Proc. IEEE Intl. Conf. On Computer Vision and Pattern Recognition '97*（Jun 1997）.

[130] 松橋　聡，藤本研司，中村　納，南　敏：“顔領域抽出に有効な修正 HSV 表示系の提案”，テレビ誌，Vol.49，No.6，pp.787-797（Jun 1995）.

[131] 松村季樹，安部伸治，小倉健司：“画像特徴と時間制約の記述による映像シーンの分類手法”，人工知能学研報 SIG-J-9702，pp.27-32（Dec 1997）.

[132] 松浦克海，鷹尾誠一，杉山善明，有木康雄：“ニュース映像中のテロップ・フリップフレームの検出と文字抽出”，信学技報 PRMU-98-188，Vol.98，No.527，pp.91-98（Jan 1999）.

[133] Merlino, A.，Morey, D.，Maybury, M.：“Broadcast News Navigation Using Story Segmentation”，*Proc. Fifth ACM Intl. Multimedia Conf.*，pp.381-391（Nov 1997）.

[134] 南　憲一，阿久津明人，浜田　洋，外村佳伸：“音情報を用いた映像インデクシングとその応用”，信学論（D-II），Vol.J81-D-II，No.3，pp.529-537（Mar 1998）.

[135] 美濃導彦，岡崎　洋，坂井利之：“対象物の属性特徴による画像検索法”，情処学論，Vol.32，No.4，pp.513-522（Apr 1991）.

[136] 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論文集，pp.53-60（Dec 1999）.

[137] 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論文集，pp.127-132（Dec 1998）.

[138] 茂木裕治，有木康雄：“ニュース映像中の文字認識に基づく記事の索引づけ”，信学技報 PRU-95，Vol.95，No.584，pp.33-40（Mar 1996）.

[139] 村瀬　洋，Vinod, V. V.：“局所色情報を用いた高速物体探索 —アクティブ探索法—”，信学論（D-II），Vol.J81-D-II，No.9，pp.2035-2042（Sep 1999）.

[140] 長坂晃朗，宮武孝文：“時系列フレーム特徴の圧縮符号化に基づく映像シーンの高速分類手法”，信学論（D-II），Vol.J81-D-II，No.8，pp.1831-1837（Aug 1998）.

[141] 長坂晃朗，田中　譲：“カラービデオ映像における自動索引付け法と物体探索法”，情処学論，Vol.33，No.4，pp.543-550（Apr 1992）.

[142] Nakamura, Y.，Kanade, T.：“Semantic Analysis for Video Contents Extraction –Spotting by Association in News Video–”，*Proc. Fifth ACM Intl. Multimedia Conf.*，pp.393-402（Nov 1997）.

[143] 中村祐一，村山正司，大田友一：“図的メディアと言語メディアの統合による知識の解析と提示”，第 4 回信学知能情報メディアシンポ論文集，pp.31-38（Dec 1998）.

[144] 中村裕一，外村佳伸：“見たい部分を簡単に短時間で –気の利いた映像メディア技術を目指して–”，信学誌，Vol.82，No.4，pp.346-353（Apr 1999）.

[145] 中村裕一，古川　亮：“概念図理解を目的としたパターン情報と自然言語情報の統合”，情報学論，Vol.36，No.1，pp.196-206（Jan 1995）.

[146] 仲尾由雄：“自動要約技術を利用した文書選別支援機能の試作”，第 1 回信学知識発見のための自然言語処理シンポオンライン論文集（Nov 1999）.

[147] 那須川哲哉：“文脈情報を利用したキーワード語義決定”，第 11 回人工知能学全大 17-01，pp.348-349（Jun 1997）.

[148] 新倉康巨，阿久津明人，谷口行信，佐藤　隆，外村佳伸，浜田　洋：“MPEG 符号化映像ショットチェンジ検出のための動き補償解析ハイブリッド法の提案”，信学論（*D-II*），Vol.J81-D-II，No.8，pp.1836-1846（Aug 1998）.

[149] 西田昌史，緒方　淳，有木康雄：“アナウンサー発話の自動抽出とディクテーションによるニュース分類”，情処学論，Vol.40，No.3，pp.1482-1490（Apr 1999）.

[150] 小野敦史，天野督士，斗谷充宏，佐藤　隆，坂内正夫：“状態遷移モデルとシーン記述言語による自動キーワード附与機能をもつ画像データベースとその評価”，信学論（D-II），Vol.J79-D-II，No.4，pp.476-483（Apr 1996）.

[151] 大辻清太，外村佳伸：“映像カット自動検出方式の検討”，テレビ学技報 ICS-92-44，Vol.16，No.43，pp.7-12（Jul 1992）.

[152] 大辻清太，外村佳伸，大庭有二：“輝度情報を使った動画ブラウジング”，信学技報 IE-90，Vol.90，No.450，pp.55-61（Feb 1991）.

[153] Ponceleon, D.，Sriniivasan, S.，Amir, A.，Petkovic, D.，Diklic, D.：“Key to Effective Video Retrieval: Effective Cataloging and Browsing”，*Proc. Sixth ACM Intl. Multimedia Conf.*，pp.99-107（Aug 1998）.

[154] Rowe, L. A.，Boreczky, J. S.，Eads, C. A.：“Indexes for User Access to Large Video Databases”，*SPIE Symp. on Electronic Imaging Science and Technology*，pp.150-161（Feb 1994）.

[155] Rowley, H. D.，Baluja, S.，Kanade, T.："Neural Network-Based Face Detection"，*IEEE Trans. on Pattern Analysis and Machine Intelligence*，Vol.20，No.1，pp.23-38（Jan 1998）.

[156] 坂内正夫，佐藤真一："画像データベースにおけるモデル形成"，信学論（D-I），Vol.J47-D-I，No.8，pp.455-466（Aug 1991）.

[157] 佐藤理史，奥村　学："電脳文章要約術 –計算機はいかにしてテキストを要約するか–"，情処学誌，Vol.40，No.2，pp.157-161（Feb 1999）.

[158] Satoh, S.，Nakamura, Y.，Kanade, T.："Name-It: Naming and Detecting Faces in News Videos"，*IEEE MultiMedia*，Vol.6，No.1，pp.22-35（Mar 1999）.

[159] 佐藤真一："ドラマ映像における登場人物のアノテーションシステム"，第5回信学知能情報メディアシンポ論文集，pp.75-82（Dec 1999）.

[160] 佐藤　隆，児島治彦，阿久津明人，外村佳伸："映像コーパスの構築と分析"，信学論（D-II），Vol.J82-D-II，No.10，pp.1552-1560（Oct 1999）.

[161] 佐藤俊雄，金出武雄："文字情報と異種情報の対応関係に基づいたニュース放送からの情報抽出"，情処学論，Vol.40，No.12，pp.4266-4276（Dec 1999）.

[162] 佐藤　隆，坂内正夫："ライブハイパメディアにおける映像情報の獲得"，信学論（D-II），Vol.J79-D-II，No.4，pp.559-567（Apr 1996）.

[163] 佐藤　隆，新倉康巨，谷口行信，阿久津明人，外村佳伸，浜田　洋："MPEG符号化映像からの高速テロップ領域検出法"，信学論（D-II），Vol.J81-D-II，No.8，pp.1847-1855（Aug 1998）.

[164] Sawaki, M.，Hagita, N.："Text-line Extraction and Character Recognition of Document Headlines with Graphical Designs Using Complementary Similarity Measure"，*IEEE Trans. on Pattern Analysis and Machine Intelligence*，Vol.20，No.527，pp.91-98（Jan 1998）.

[165] 柴田正啓："映像の内容記述モデルとその映像構造化への応用"，信学論（D-II），Vol.J78-D-II，No.5，pp.754-764（May 1995）.

[166] Smith, M. A.，Kanade, T.："Video Skimming and Characterization Through the Combination of Image and Language Understanding Techniques"，*CMU Tech. Rep. CMU-CS-97-111*（Feb 1997）.

[167] Smoliar, S. W.，Zhang, H. J.："Content-Based Video Indexing and Retrieval"，*IEEE MultiMedia*，Vol.1，No.2，pp.62-72（Aug 1994）.

[168] Srihari, R. K.：“Automatic Indexing and Content-Based Retrieval of Captioned Images”，*IEEE Computer*，Vol.28，No.9，pp.49-56（Sep 1995）．

[169] 杉山善明，西田昌史，有木康雄：“Video Assembler：動画象処理と音声処理の統合によるニュース記事内容の分割”，第 4 回信学知能情報メディアシンポ論文集，pp.23-30（Dec 1998）．

[170] 杉山善明，有木康雄：“多重部分空間法に基づくテレビスポーツニュース映像の自動分類”，信学論（D-II），Vol.J81-D-II，No.9，pp.2112-2119（Sep 1998）．

[171] Sundheim，B. M.：“Named Entity Task Definition, Version 2.1”，*Proc. Sixth Message Understanding Conf.*，pp.317-332（Nov 1995）．

[172] Sundheim，B. M.：“Overview of Results of the MUC-6 Evaluation”，*Proc. Sixth Message Understanding Conf.*，pp.13-31（Nov 1995）．

[173] Wactlar, H. D.，Christel，M. G.，Gong, Y.，Hauptmann, A. G.：“Lessons Learned from Building a Terabyte Digital Video Library”，*IEEE Computer*，Vol.32，No.2，pp.66-73（Feb 1999）．

[174] Wactlar, H. D.，Kanade, T.，Smith, M. A.，Stevens, S. M.：“Intelligent Access to Digital Video: Informedia Project”，*IEEE Computer*，Vol.29，No.3，pp.46-52（May 1996）．

[175] Wactler, H. D.，Hauptmann, A. G.，Witbrock, M. J.：“Informedia News-on-Demand: Using Speech Recognition to Create a Digital Video Library”，*CMU Tech. Rep. CMU-CS-98-109*（Mar 1998）．

[176] 若尾孝博，江原暉将，村木一至，白井克彦：“テレビニュース番組電子化原稿を題材とした自動要約手法の大規模評価”，情処学研報 97-NL-119，Vol.97，pp.31-36（Nov 1997）．

[177] 若尾考博，江原暉将，白井克彦：“テレビニュース番組の字幕に見られる要約の手法”，情処学研報 97-NL-122，Vol.97，No.109，pp.83-89（Nov 1997）．

[178] 渡辺靖彦, 長尾　眞：“パターン情報と自然言語情報の統合による植物図鑑の図の理解”，人工知能学誌，Vol.11，No.6，pp.888-895（Nov 1996）．

[179] 渡辺靖彦，岡田至弘，角田達彦，長尾　眞：“TV ニュースと新聞記事の対応付け”，人工知能学誌，Vol.12，No.6，pp.921-927（Nov 1997）．

[180] 渡辺靖彦，岡田至弘，長尾　眞：“TV ニュースで用いられるテロップの意味解析”，情処学研報 96-NL-116，Vol.96，No.89，pp.107-114（Nov 1996）．

[181] 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”，信学論（D-II），Vol.J79-D-II，No.5，pp.747-755（May 1996）.

[182] 柳沼良和, 和泉直樹，坂内正夫：“同期されたシナリオ文書を用いた映像編集方法の一提案”，信学論（D-II），Vol.J79-D-II，No.4，pp.547-558（Apr 1996）.

[183] 山田剛一，杉山一成，与那嶺靖典，中川裕志：“新聞記事における顔写真と文書表現との自動対応づけ”，第 4 回信学知能情報メディアシンポ論文集，pp.39-46（Dec 1998）.

[184] 山田一郎，金　淵培，柴田正啓，浦谷則好：“ニュース原稿を利用した話題抽出とニュース話題選択の個人化の検討”，第 1 回信学知識発見のための自然言語処理シンポオンライン論文集（Nov 1999）.

[185] 横山太郎，呉　海元，谷内田正彦：“色彩画像からの顔の発見と顔部品の同定”，情処学研報 96-CVIM-100，Vol.96，No.47，pp.85-92（May 1996）.

[186] 吉見毅彦，奥西稔幸，山路孝浩，福持陽士：“表題へのつながりに基づく文の重要度評価”，自然言語処理，Vol.6，No.1，pp.43-57（Jan 1999）.

[187] Zabih, R.，Miller, J.，Mai, K.：“A Feature-Based Algorithm for Detecting and Classifying Scene Breaks”，*ACM J. of Multimedia Systems*，Vol.7，No.2，pp.119-128（Mar 1999）.

書籍・学位論文

[188] 浜田玲子：“料理テキスト教材における調理手順の構造化手法”，修士論文，東京大学大学院工学系研究科（2000 予定）.

[189] 上坂英樹：“カメラ操作を考慮した静止領域と動領域の分離”，修士論文，東京大学大学院工学系研究科（1997）.

[190] 宮崎誠也：“対象物の階層的認識手法による画像検索”，卒業論文，東京大学工学部（1999）.

[191] 長尾　眞：“自然言語処理”，岩波講座ソフトウェア科学 15，岩波書店（1996）.

[192] 国立国語研究所：“テレビ放送の語彙調査 I –方法・標本一覧・分析–”，国立国語研究所報告，国立国語研究所（1995）.

[193] 西尾章治郎，大田友一，横田一正，西田豊明，佐藤哲司：“情報の共有と統合”，岩波講座マルチメディア情報学 7，岩波書店（1999）.

[194] 佐久間丈貴：“単語の概念関係を用いた誤り訂正補助手法”，卒業論文，東京大学工学部（2000 予定）.

[195] 佐藤　隆：“ライブ映像の認識と利用を統合するハイパーメディアシステムの研究”，博士論文，東京大学大学院工学系研究科（1995）．

[196] 瀬戸喜巳：“文間の概念的類似度を用いた重要文抽出”，卒業論文，東京大学工学部（1999）．

[197] 下萩原勉：“テンプレート切り出しによる不特定話者対応のワードスポッティング”，卒業論文，東京大学工学部（2000 予定）．

[198] 高木幹雄，下田陽久：“画像解析ハンドブック”，東京大学出版会（1991）．

[199] 陳　廸傑：“副音声利用による映像への索引付け”，卒業論文，東京大学工学部（1998）．

[200] Wechsler, H.，Phillips, P. J.，Bruce, V.，Soulie, F. F.，Huang，P. S. eds.：“Face Recognition from Theory to Application”，*NATO ASI Series F: Computer and System Sciences*，Springer（1998）．

[201] 山本晃司：“ニュース映像の分類及び索引付けに関する研究”，修士論文，東京大学大学院工学系研究科（1998）．

WWW・ツール等

[202] The Robotics Institute, Carnegie Mellon University：“Informedia Digital Library Project”，http://www.informedia.cs.cmu.edu/．

[203] （社）日本放送協会：“NHK 文字放送”，http://www.nhk.or.jp/pr/bangumi/moji/moji.htm．

[204] 京都大学大学院情報学研究科知能情報学専攻言語メディア研究室：“日本語構文解析システム KN パーザ第 2.0 β 6 版”，http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/knp.html．

[205] 京都大学大学院情報学研究科知能情報学専攻言語メディア研究室：“日本語形態素解析システム JUMAN 第 3.6 版”，http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/juman.html．

[206] 京都大学大学院情報学研究科知能情報学専攻言語メディア研究室：“京都大学コーパス第 2.0 版”，http://www-lab25.kuee.kyoto-u.ac.jp/nl-resource/corpus.html．

[207] 奈良先端科学技術大学院大学情報科学研究科情報処理学専攻自然言語処理学講座：“日本語形態素解析システム茶筅第 2.0 β 6 版”，http://cactus.aist-nara.ac.jp/lab/nlt/chasen.html．

[208] 国立国語研究所：“分類語彙表［フロッピーディスク版］”，国立国語研究所言語処理データ集第 5，秀英出版．

[209] 国立国語研究所："「分類語彙表」形式による語彙分類表(増補版)[増補モニタ用電子化データ]".

[210] (技)新情報処理開発機構："RWC テキストデータベース第 1 版".

[211] Rowley, H. D.："Facedetector".

付録

RGB 表色系と HSI 表色系

特徴量の定義と性質について述べる前に，RGB 表色系と HSI 表色系について述べる．

一般に，計算機上で画像を扱う際，最終的にディスプレイに表示する必要があるために，赤（R, Red），緑（G, Green），青（B, Blue）の 3 原色の強さの和で色を表示する加法混色系である RGB 表色系を用いている．RGB 表色系は，画像を比較的現実に忠実な色（心理物理色）でディスプレイ上に再現するのには適しているものの，例えば，ある色の領域が照明条件により異なる色として扱われるなど，人間による直感的な色の認識（知覚色）とは乖離している．

そこで，より人間による直感的な色表現を行うために，知覚色を表示する代表的な顕色系であるマンセル表色系を用いることが考えられる．マンセル表色系では，色を色相（H, Hue），彩度（S, Saturation），明度（I, Intensity）の 3 要素に分離して扱う．各々の要素は，色相は知覚色，彩度は色の鮮やかさ，明度は明るさにほぼ対応する．

RGB 表色系と HSI 表色系相互間の変換式の例として，HSI 六角錐色モデルによるものを以下に示す．

- HSI 変換（$R, G, B \rightarrow H, S, I$）

$$I = \max\{R, G, B\}, S = \begin{cases} \frac{I-i}{I} & (\text{when } I = 0) \\ 0 & (\text{when } I > 0) \end{cases}, H = \begin{cases} \frac{\pi}{3}(b-g) & (\text{when } R = I) \\ \frac{\pi}{3}(2+r-b) & (\text{when } G = I) \\ \frac{\pi}{3}(4+g-r) & (\text{when } B = I) \\ \text{不定} & (\text{when } I = 0) \end{cases}$$

ただし，

$$i = \min\{R, G, B\}, r = \frac{I-R}{I-i}, g = \frac{I-G}{I-i}, b = \frac{I-B}{I-i}$$

であり，$H < 0$ のときは，Hの値に 2πを加える．

- HSI 逆変換（ $H, S, I \rightarrow R, G, B$ ）

$$R = \begin{cases} I & (\text{when } S > 0 \text{ and } h = 0, 5 \text{ or } S = 0) \\ Q & (\text{when } S > 0 \text{ and } h = 1) \\ P & (\text{when } S > 0 \text{ and } h = 2, 3) \\ T & (\text{when } S > 0 \text{ and } h = 4) \end{cases}$$

$$G = \begin{cases} I & (\text{when } S > 0 \text{ and } h = 1, 2 \text{ or } S = 0) \\ Q & (\text{when } S > 0 \text{ and } h = 3) \\ P & (\text{when } S > 0 \text{ and } h = 4, 5) \\ T & (\text{when } S > 0 \text{ and } h = 0) \end{cases}$$

$$B = \begin{cases} I & (\text{when } S > 0 \text{ and } h = 3, 4 \text{ or } S = 0) \\ Q & (\text{when } S > 0 \text{ and } h = 5) \\ P & (\text{when } S > 0 \text{ and } h = 0, 1) \\ T & (\text{when } S > 0 \text{ and } h = 2) \end{cases}$$

ただし，$H = 2\pi$のときは $H = 0$ と読みかえ，$h = \lfloor \frac{3H}{\pi} \rfloor$ としたとき，

$$P = I(1 - S), Q = I\{1 - S(H - h)\}, T = I\{1 - S(1 - H + h)\}$$

である．

以下の特徴量を抽出する際には，RGB 表色系，HSI 表色系のいずれの値を用いても原理的には可能な処理である．しかし，HSI 表色系を用いた場合，色相と彩度に関する頻度分布は知覚色に関する分布を，明度に関する頻度分布は明るさの分布を示すことになり，単純な RGB 表色系の頻度分布とは異なる意味合いをもつ．

また，HSI 表色系よりも毛髪や陰影など低明度領域の彩度を低く抑える修正 HSI 表色系は以下のように定義される [130]．なお，Iの定義は通常の HSI 表色系と同じである．

$$S_m = \sqrt{\left(\frac{(R-G)+(R-B)}{2}\right)^2 + \left(\frac{\sqrt{3}(G-B)}{2}\right)^2}, H = \begin{cases} \arccos \frac{(R-G)+(R-B)}{2S_m} & (\text{when } I > 0) \\ \text{不定} & (\text{when } I = 0) \end{cases}$$

人物名詞辞書該当分類番号一覧（旧分類語彙表）

項目番号	見出し	該当分類段落
1.104	本体	3, 5
1.1300	様相	1
1.131	風・観・姿	2
1.1960	単位	3
1.200	われ・なれ・かれ・だれ	2, 5
1.201	自他	1
1.202	人間	1, 5, 6
1.204	男女	2
1.205	老少	2, 5
1.210	家族	5
1.212	親・先祖	2
1.213	子・子孫	1, 7
1.220	相手・仲間	4
1.221	友・なじみ	1
1.222	主客	2
1.2301	国民・住民	3
1.232	君主	3
1.234	人物	1, 4, 17, 23, 27
1.240	成員・職	1, 2, 3, 4
1.2410	専門的・技術的職業	2, 3, 4, 7, 9
1.2411	支配的・管理的・書記的職業	1, 2, 3, 4, 5, 7
1.2412	販売など	1
1.2415	運輸	2
1.242	軍人	1, 2, 6
1.243	長	1, 2, 3, 4, 5, 6, 7, 9
1.244	相対的地位	1, 2
1.2450	臨時的地位（役・役員）	5
1.2451	臨時的地位（作者・仕手・主役・観客など）	3, 6
1.2452	臨時的地位（被告・仲人・審判・使者・持主など）	1, 4, 5, 7
1.251	家	1
1.3081	方法	1
1.3400	義務	5
1.344	実行・犯罪	1
1.3800	仕事	4
1.440	住居	1
1.550	生	3

（次頁へ続く）

項目番号	見出し	該当分類段落
1.573	手足・指	2
1.580	生命	1
1.581	生死・生	2
3.100	こそあど	12, 17
3.112	相互・異同	1
非収録語	さま, エリート, シェフ, ダイバー, ディーラー, 委員長, 幹事長, 観光客, 関係者, 宮司, 経営者, 経験者, 剣士, 検事正, 高校生, 高等弁務官, 佐, 市議, 室長, 書記長, 障害者, 組員, 総書記, 総務審議官, 代表者, 大人, 町議, 調査官, 通産相, 頭取, 秘書官, 副総裁, 民兵, 労相	

（以上）

場所・組織名詞辞書該当分類番号一覧（旧分類語彙表）

項目番号	見出し	該当分類段落
1.100	こそあど	3
1.101	事柄	3
1.103	真・実	2
1.104	本体	7, 8
1.1100	類・例	1, 3, 5
1.1101	等級・系列	1, 2, 5
1.1110	関係	1
1.1112	因果	3
1.1211	（見出しなし）	1, 4, 6
1.122	確立	1, 2
1.1240	存在	3
1.1251	除去	3, 4
1.1300	様相	2
1.1301	情勢	2
1.131	風・観・姿	1
1.132	内容・組織	1, 2
1.1330	性質	2
1.1341	調節	1
1.1400	力	1
1.1401	動力・圧力など	2
1.1502	同化・強化など	1
1.1505	開始・終了	1, 5
1.1510	動き	2
1.1520	過程・経路	1, 2
1.1521	移動・発着	2
1.1526	進退	2
1.1528	往復	2, 4
1.1530	出入り	1, 2, 3, 4, 9
1.1550	離合・集合	1, 3
1.1552	分離・分散	1
1.1553	開閉・封	4
1.1563	止め・防ぎ・妨げ・打破など	4, 7
1.1571	刺し・切り・削り	2
1.1572	破壊	2
1.1582	伸縮	2
1.160	位置・地点・場合	1

（次頁へ続く）

項目番号	見出し	該当分類段落
1.1610	時間	3
1.1611	時機	3
1.1620	期間	2
1.1623	時代	1, 4
1.1630	年	2
1.1633	週・週日	2
1.1634	日	1, 2
1.1642	過去	1
1.1650	順序	1, 3, 4, 5
1.1651	終始	6
1.1652	途中・盛り・次ぎ	1
1.1670	前後	1
1.1700	空間・場所	3, 4, 6
1.171	点・目・許	1, 2, 7
1.1720	範囲・席・跡	1, 2, 3, 4, 5, 6
1.1721	境	1
1.1730	方向・たてよこ	1
1.1731	方面・方角	1, 3, 6, 7
1.1741	上下	1, 2, 3, 7
1.1742	中・かみ・しも・頂・すみ	1, 2, 3, 6
1.175	面・側・表裏	1, 2, 4
1.176	前後・間・端	1, 2, 6
1.1770	内外	1, 2, 3
1.178	ふち・そば・まわり・沿い	2, 5, 6, 7, 8
1.180	形・型・姿・構え	2, 6
1.181	点・線	2
1.182	角・円・三角形など	3
1.184	穴・口	1, 2
1.1910	数	1, 2
1.1911	値・額	1
1.193	角度・軽重・寒暖など	3
1.1960	単位	1, 2, 4, 5, 8, 9
1.1961	円・尺・坪・合・匁など	1, 2
1.1962	年・時・ワット・馬力など	1
1.1963	割・前など	1
1.1970	和・差・比・率など	1
1.1980	一般・全体・部分	2, 4, 10
1.1981	（見出しなし）	1, 3, 4, 7
1.199	程度・限度	3

（次頁へ続く）

項目番号	見出し	該当分類段落
1.210	家族	1
1.220	相手・仲間	1, 3
1.2301	国民・住民	1, 2, 4
1.231	人民	2
1.232	君主	2
1.233	社会階層	1, 6
1.234	人物	15
1.240	成員・職	7
1.2410	専門的・技術的職業	12, 15
1.2411	支配的・管理的・書記的職業	9
1.2412	販売など	3
1.2416	生産工程	1, 5
1.2450	臨時的地位（役・役員）	1
1.2451	臨時的地位（作者・仕手・主役・観客など）	1, 4
1.2452	臨時的地位（被告・仲人・審判・使者・持主など）	6
1.252	郷里	2
1.253	国	1, 4, 5, 6, 7
1.254	都市・集落	1, 2, 3, 4, 5
1.255	政治的区画	1, 3, 4
1.259	個有地名	1
1.260	社会	1
1.261	社会分野	1, 3, 4, 5
1.262	現場	1, 2, 3, 4
1.263	社寺・学校	2, 4, 5, 6, 7
1.264	事務所・市場・駅など	1, 2, 3, 5, 6
1.265	店・旅館・病院・劇場など	1, 2, 5, 6, 9, 10, 11, 12
1.266	宅・宿・教室など	1, 2, 3
1.270	機関	1
1.271	政府機関	1, 2, 3, 4
1.272	公共機関	1, 2, 3, 4
1.273	議会	1, 2
1.274	軍	1, 4
1.280	同盟・団体	1, 2, 3, 4, 5
1.281	クラス	1
1.3004	感情・気分	7
1.3045	研究・実験・調査・検査など	2
1.3047	解決・決定・きまり	4
1.3050	練習・まね・学習	1

（次頁へ続く）

項目番号	見出し	該当分類段落
1.3051	（見出しなし）	6
1.3061	（見出しなし）	10, 11
1.3062	（見出しなし）	6, 9
1.3063	（見出しなし）	2, 6, 7, 9
1.3065	（見出しなし）	6
1.3066	（見出しなし）	2
1.3067	（見出しなし）	1, 3, 4, 5, 7
1.3068	（見出しなし）	1
1.3070	意味・問題・趣旨・大綱など	1, 3, 6
1.3072	論理・うそ・誤り・訂正など	1, 3
1.3074	学問・学科	2, 7
1.3075	説・論	1, 3
1.3080	原理・規則・主義など	1, 2, 4, 5, 6, 7, 9
1.3081	方法	3, 4
1.3082	制度・慣例	1
1.3083	見当・準備	1
1.3084	計画・案	4, 5, 6
1.3101	言語	1, 2
1.3102	名	2
1.3111	文法	2
1.3113	文字	3
1.3115	表・図・譜・式	7
1.3122	通信	1, 2
1.3123	伝達・報知	4, 7
1.3131	談話	2
1.3133	会議・論議	1, 2, 4
1.3135	言論	1
1.3140	宣言・発表	3
1.3151	書き	1
1.3153	記名など	1
1.3154	文章	2
1.3155	文書	1
1.3160	文献・図書	1
1.3161	教科書・古典・集・新聞雑誌	7
1.3162	辞書・目録・暦	2, 4
1.320	創作・著述	1
1.324	演劇・映画	3, 4
1.330	文化・歴史・風俗	1
1.3310	人生・祝福	6

（次頁へ続く）

項目番号	見出し	該当分類段落
1.3312	出処進退	1
1.332	勤務・労働	3
1.3331	住	9
1.336	祭儀・式・宗教的行為	1, 6
1.3371	旅・行楽	1
1.3374	スポーツ	3, 4, 9
1.338	騒ぎ	1
1.3400	義務	4
1.3401	権利	1, 3
1.341	身上	1, 2
1.345	活動	3
1.346	用事	1
1.347	成功・失敗	1
1.3500	交わり	2, 3
1.3501	争い	2
1.3510	集会	1, 3, 5
1.3511	出欠	4
1.3520	対面・訪れ・招き	4
1.3523	紹介・周旋	1, 5
1.3530	約束	1
1.3531	交渉	2, 3, 5, 8
1.3532	賛成・許容	1, 3, 5
1.3540	協力・参加	4, 5
1.3551	競争・戦争	1, 3, 5
1.356	攻防	1, 3, 5
1.357	勝敗	1, 2
1.360	支配・政治・革命	5, 7
1.3611	裁判	1, 2
1.3612	刑	1
1.362	設立・運営	1, 2, 3
1.363	人事	8, 12
1.3640	教育	4
1.365	救護・世話	1, 3
1.366	請求	2
1.367	命令・制約	2, 10, 12
1.3682	賞罰	1
1.3701	所有	1, 5
1.3710	経済・収支	1, 6
1.3721	資本・金銭	8

（次頁へ続く）

項目番号	見出し	該当分類段落
1.373	価格・費用	8
1.374	給与・料金・利子	3
1.375	損得	4
1.3761	売買	1
1.378	貸借	3
1.3802	産業・生産	1, 2
1.3822	土木	1, 7
1.3830	運輸・交通	1
1.3832	出版・興行	1, 2, 8
1.3852	使用	3
1.386	製造・荷造り	1
1.402	産物	1
1.404	貨幣	1, 3
1.410	資材	2
1.411	紙	1
1.4150	輪・車・棒・管など	2, 3
1.4153	ばね・栓・キーなど	3
1.4161	網	1
1.4200	衣料・綿・皮・糸	5
1.423	下着・羽織・ズボン・コートなど	7
1.424	袖・衿・身頃・ポケットなど	2
1.4251	ネクタイ・帯・手袋・靴下など	3
1.4320	米・糠・小麦粉など	2
1.440	住居	4, 8, 9
1.441	家屋	4, 10
1.443	へや・廊下・床など	1
1.4511	びん・壷・皿・膳など	10
1.452	食器	3
1.4550	刃物	2
1.457	おもちゃなど	2, 3, 18
1.458	標識	3, 8
1.461	鏡・レンズ	1
1.462	電気機具	1
1.463	機械	1, 2, 3
1.465	乗り物（陸上）	3, 8, 9
1.470	地類（土地利用）	1, 10, 12, 13, 14
1.471	道路・橋	1, 4, 5, 6, 13
1.472	その他の土木施設	4, 6, 14
1.502	色	1, 6

（次頁へ続く）

項目番号	見出し	該当分類段落
1.510	自然・物体・物質	1
1.514	天災	4
1.5150	気象	7
1.5151	風	1
1.5152	雲	9
1.520	宇宙・空	3
1.521	天体	2
1.523	地	2
1.524	地形・山野	1, 5, 10, 13
1.525	川・湖	1, 10, 14
1.526	海・島	1, 3, 6, 7, 8, 9, 11, 14
1.528	地帯	1
1.552	植物名	20, 27
1.553	枝・葉・花など	5, 13
1.564	魚	2
1.571	頭・目鼻・顔	5, 12
1.573	手足・指	14
1.577	血・孔・涙・汗・尿など	10
1.581	生死・生	3, 10
1.582	死	6
1.586	四百四病	11
3.100	こそあど	13
3.114	さながら	1
3.120	在不在	2
3.131	普通・非凡	1
3.132	特別・異様	2
3.167	翌・次	1
3.182	丸い・平たい・荒らい・まっすぐ・水平など	2
3.1920	長い・広い	4
3.1921	厚い・太い・大きい	3, 4
3.198	ひとり・みんな	4
3.304	かしこい・おろか	4
3.330	風俗	3
非収録語	ＩＣ, ＯＢ, ＳＡ, ごろ, そのもの, づくり, ふ頭, べり, アイスアリーナ, アイランド, アジア大会, イルクーツク, インタチェンジ, ウイング, キャンパス, ギャラリー, コンバインド, サイド, シアトル, ショップ, ジャズマン, スタジアム, スペースシャトル, スポークスマン, スポーツセンター, セミナー, ターミナル, ツアー, テニスコート, データベース, ナイン, ニュータウン, ノーマルヒル, フォーラム, ヘリ, マスコミ, モーターショー, ユー	

（次頁へ続く）

項目番号	見出し	該当分類段落
	ス, リスト, ルート, 安保, 安保理, 駅伝, 会派, 回り, 海域, 海兵団, 開会式, 外相・蔵相会議, 外相会談, 各紙, 核実験, 学芸員, 環礁, 簡裁, 観念論, 球界, 競技場, 競輪場, 共和国, 橋上, 区内, 区役所, 軍司令官, 係, 経済界, 警備隊, 建設相, 研究所, 県警, 県内, 厚生年金会館, 高裁, 合弁, 国税局, 懇談会, 再選挙, 作り, 三国, 山ろく, 産米, 市域, 支店長, 事務局, 事務総長, 実験線, 主導, 取引所, 州知事, 諸語, 症候群, 礁, 訟廷事務所, 上り線, 乗用車, 侵攻, 審議会, 新幹線, 新法, 進攻, 陣内, 水俣病, 生態調査, 西岸, 西口, 跡, 跡地, 切り崩し, 先進国, 選管, 選挙区, 選手権, 全域, 全国ネット, 早版, 総領事館, 速道, 卒, 打線, 体育館, 大リーグ, 大学院, 大使館, 大社, 大震災, 大地震, 大島, 地下街, 中華街, 駐車場, 庁舎, 展示場, 電, 都立高, 東京本社, 特捜部, 特派員, 南署, 年会, 飛行場, 府警, 府庁, 府内, 武道館, 部会, 副首相, 福祉会館, 弁護士会, 放送局, 北端, 本会議, 本港, 万博, 門跡, 予算委, 理事会, 留保, 竜馬, 両社, 領事館, 和平	

（以上）

時相名詞辞書該当分類番号一覧（旧分類語彙表）

項目番号	見出し	該当分類段落
1.100	こそあど	4
1.104	本体	8
1.1101	等級・系列	2
1.1111	本	1
1.1112	因果	3
1.1113	理由・目的・証拠	1, 2
1.1520	過程・経路	3
1.160	位置・地点・場合	4
1.1610	時間	1, 3, 4
1.1612	時刻	1, 2, 4
1.1613	毎日・平生	1, 2, 3, 6
1.1620	期間	1, 2, 4
1.1621	永久・一生	3
1.1622	年配	3
1.1623	時代	1, 2, 5
1.1624	季節	2, 3, 5
1.1630	年	1, 2
1.1631	節・節日	1
1.1632	月	1
1.1633	週・週日	2
1.1634	日	1, 3
1.1635	朝晩	1, 2, 3, 5
1.1641	現在	1, 2
1.1642	過去	1, 2, 4
1.1643	未来	1, 3, 4
1.1651	終始	2, 3, 6
1.1652	途中・盛り・次ぎ	1
1.1670	前後	1, 3, 5
1.1671	翌・後日・当面など	1, 2, 4
1.1700	空間・場所	4
1.1720	範囲・席・跡	1, 6
1.1730	方向・たてよこ	1
1.1731	方面・方角	1
1.1741	上下	1
1.1742	中・かみ・しも・頂・すみ	1, 2
1.176	前後・間・端	1, 6, 8

（次頁へ続く）

項目番号	見出し	該当分類段落
1.1770	内外	1, 2
1.1950	一二三	7, 11, 14
1.1960	単位	2
1.1961	円・尺・坪・合・匁など	2
1.1962	年・時・ワット・馬力など	1
1.1963	割・前など	1
1.1980	一般・全体・部分	2
1.1981	（見出しなし）	7
1.199	程度・限度	1, 3, 4
1.259	個有地名	1
1.264	事務所・市場・駅など	1
1.3001	感覚	1
1.3004	感情・気分	5
1.3062	（見出しなし）	6
1.3081	方法	3
1.375	損得	1
1.434	菓子	6
1.443	へや・廊下・床など	1
1.471	道路・橋	6
1.510	自然・物体・物質	5
1.521	天体	2, 3, 5
3.100	こそあど	1, 9, 13
3.114	さながら	1, 2
3.120	在不在	2
3.130	繁簡	1
3.131	普通・非凡	2
3.133	良不良・適不適	1
3.160	時	1
3.162	一度に・再び・毎度など	7
3.164	過現未	5
3.165	まだ・もう・もっと・また	12
3.167	翌・次	1
3.1921	厚い・太い・大きい	4
非収録語	うえ, うち, かわり, こんど, その後, たび, とき, なか, まえ, ゆえ, 為, 一国, 一章, 一町, 一発, 一歩, 一夕, 月比, 頃, 昨夏, 史上, 世界一, 全日, 中期, 長年, 冬期, 特段, 日日	

（以上）

人物名詞辞書該当分類番号一覧（新分類語彙表）

項目番号	見出し	該当分類段落
1.1000	こそあど	7
1.1040	本体	9, 10, 12, 15
1.1113	理由・目的・証拠	8
1.1131	連絡・連結	7A
1.1300	様相	2
1.1310	風・観・姿	4
1.1401	弾力・動力・圧力など	2B
1.1622	年配	1
1.1960	単位	48, 53, 65
1.2000	われ・なれ・かれ	2, 5
1.2010	自他	2, 4
1.2020	人間	1, 6, 7
1.2030	神仏・精霊	13
1.2040	男女	2
1.2050	老少	4A, 7, 8
1.2100	家族	6
1.2110	夫婦	2
1.2120	親・先祖	2
1.2130	子・子孫	1, 9
1.2200	相手・仲間	8
1.2210	友・なじみ	1
1.2220	主客	2
1.2301	国民・住民	3, 4
1.2320	君主	5
1.2330	社会階層	3
1.2340	人物	1, 2, 4, 6, 22, 34, 36, 39, 51, 58
1.2400	成員・職	1, 2, 3, 5, 6, 7, 10
1.2410	専門的・技術的職業	2, 5, 6, 7, 11, 15, 25, 27
1.2411	支配的・管理的・書記的職業	1, 2, 4, 5, 7, 9, 11, 13
1.2412	販売など	1
1.2415	運輸	1A, 2
1.2416	生産工程	2
1.2420	軍人	1, 2, 6
1.2430	長	1, 2, 3, 4, 5, 7, 9, 10, 12, 13, 14

（次頁へ続く）

項目番号	見出し	該当分類段落
1.2440	相対的地位	1, 9, 14
1.2450	臨時的地位（役・役員）	1, 5
1.2451	臨時的地位（作者・仕手・主役・観客など）	1, 2, 3, 6
1.2452	臨時的地位（被告・仲人・審判・使者・持主など）	1, 6, 8, 12
1.2510	家	2
1.2650	店・旅館・病院・劇場など	1
1.3081	方法	1, 2
1.3102	名	5
1.3151	書き	4
1.3220	美術	11
1.3374	スポーツ	1
1.3400	義務	9
1.3430	行為	10
1.3440	実行・犯罪	1
1.3640	教育	5
1.3800	仕事	6
1.4400	住居	1
1.4470	家具	14
1.5500	生	3
1.5730	手足・指	2
1.5800	生命	1
1.5810	生死・生	1
3.1000	こそあど	19
3.1120	相互・異同	2, 14
非収録語	さま, ディーラー, 幹事長, 観光客, 関係者, 経験者, 検事正, 高等弁務官, 佐, 書記長, 障害者, 組員, 総書記, 総務審議官, 調査官, 通産相, 秘書官, 副総裁, 労相	

（以上）

場所・組織名詞辞書該当分類番号一覧 (新分類語彙表)

項目番号	見出し	該当分類段落
1.1000	こそあど	12
1.1010	事柄	7
1.1020	事項	1
1.1030	真・実	6
1.1040	本体	2, 17
1.1100	類・例	1, 7, 12
1.1101	等級・系列	1, 2, 9, 15
1.1110	関係	1, 2
1.1112	因果	9
1.1210	出現	3
1.1211	復活・前兆・興亡	1, 6
1.1220	確立	1, 2
1.1240	存在	6, 8
1.1251	除去	7, 11
1.1300	様相	5, 6
1.1301	情勢	2
1.1310	風・観・姿	1, 13
1.1320	内容・組織	1, 4, 6, 7
1.1330	性質	4
1.1341	調節	1, 2
1.1400	力	3B
1.1401	弾力・動力・圧力など	4
1.1402	物力・権力・体力など	1A
1.1501	変化	1, 5
1.1502	同化・強化など	1
1.1504	改新	1
1.1505	開始・終了	2, 7
1.1510	動き	4
1.1520	過程・経路	3, 4A, 7
1.1521	移動・発着	6
1.1524	通過・普及など	1
1.1525	導き・追い・逃げなど	4
1.1526	進退	3
1.1528	往復	2, 7, 9
1.1530	出入り	2, 4, 7, 8, 18
1.1550	離合・集合	2, 3, 10

(次頁へ続く)

項目番号	見出し	該当分類段落
1.1551	統一・組合せ	4
1.1552	分離・分散	1, 2
1.1563	止め・防ぎ・妨げ・打破など	5, 13
1.1571	刺し・切り・削り	7
1.1572	破壊	8
1.1582	伸縮	6
1.1600	位置・地点・場合	3, 4, 7
1.1610	時間	1, 4, 6
1.1611	時機	3, 4, 8
1.1620	期間	2
1.1623	時代	1, 12
1.1630	年	3
1.1633	週・週日	3
1.1634	日	1, 6
1.1635	朝晩	2
1.1642	過去	2
1.1643	未来	2
1.1650	順序	4, 8, 9, 10, 12
1.1651	終始	4, 14
1.1652	途中・盛り・次ぎ	1, 1A, 2, 4A, 8
1.1670	前後	1
1.1700	空間・場所	9, 10, 11, 12, 18, 19, 21A
1.1710	点・目・許	1, 4, 14, 17, 21
1.1720	範囲・席・跡	1, 2, 4, 5, 6, 7, 8, 9, 14, 17, 18, 25, 27, 28
1.1721	境	2
1.1730	方向・たてよこ	1, 2, 3, 14
1.1731	方面・方角	5, 9, 16, 20, 21
1.1740	左右	5
1.1741	上下	1, 8, 9, 17
1.1742	中・かみ・しも・頂・すみ	1, 3, 4, 5, 9, 12, 13, 14, 18
1.1750	面・側・表裏	1, 8, 12, 13
1.1760	前後・間・端	1, 3, 8, 10, 12
1.1770	内外	1, 2, 3, 4, 6, 8, 11
1.1780	ふち・そば・まわり・沿い	8, 9, 11, 13, 14, 18
1.1800	形・型・姿・構え	3, 12, 13
1.1810	点・線	2, 12, 20

（次頁へ続く）

項目番号	見出し	該当分類段落
1.1820	角・円・三角形など	8A, 14
1.1830	玉・凹凸・うず・しわなど	6
1.1840	穴・口	1, 4
1.1860	束・片・列・わく	13
1.1910	数	4, 8
1.1911	値・額	4
1.1920	長短・広狭	4
1.1930	角度・軽重・寒暖など	8
1.1950	一二三	26, 36
1.1951	多少	9
1.1960	単位	1, 3, 7, 10, 13, 14, 17, 22, 27, 28, 30, 43, 45, 49, 50A, 51, 52, 54, 55, 57, 59, 61, 63, 64, 71
1.1963	割・前など	2
1.1970	和・差・比・率など	1, 2, 4, 12
1.1980	一般・全体・部分	2, 7, 11, 20
1.1981	群・組・対	3, 4, 9, 10, 16
1.1990	程度・限度	1, 7
1.2050	老少	4
1.2100	家族	1
1.2200	相手・仲間	1, 5, 7
1.2301	国民・住民	1, 2, 5
1.2310	人民	2, 3
1.2320	君主	3
1.2330	社会階層	4, 7
1.2340	人物	7, 32
1.2410	専門的・技術的職業	13, 28
1.2416	生産工程	1, 6
1.2418	サービス	11
1.2420	軍人	4
1.2440	相対的地位	13
1.2450	臨時的地位（役・役員）	4
1.2451	臨時的地位（作者・仕手・主役・観客など）	4
1.2452	臨時的地位（被告・仲人・審判・使者・持主など）	11, 14
1.2510	家	1
1.2520	郷里	2
1.2530	国	1, 3, 4, 5, 6, 7

（次頁へ続く）

項目番号	見出し	該当分類段落
1.2540	都市・集落	1, 2, 3, 4, 6, 8
1.2550	政治的区画	1, 4, 5
1.2590	個有地名	1, 1A, 4
1.2600	社会	1
1.2610	社会分野	1, 6, 7, 9, 10
1.2620	現場	1, 2A, 3, 3A, 4
1.2630	社寺・学校	1, 2, 8, 9, 10, 12, 14, 14A
1.2640	事務所・市場・駅など	1, 2, 3, 4, 5, 6, 7
1.2650	店・旅館・病院・劇場など	2, 3, 14, 17, 20, 22, 23, 24
1.2660	宅・宿・教室など	1, 2, 6, 7, 8
1.2700	機関	1, 3, 4
1.2710	政府機関	1, 2, 3, 5, 6, 7
1.2720	公共機関	1, 2, 3, 4
1.2730	議会	1, 2
1.2740	軍	1, 2, 6
1.2800	同盟・団体	1, 2, 8, 10, 11, 12, 14
1.2810	クラス	1
1.3004	感情・気分	8
1.3021	対人感情（敬意・信頼など）	11
1.3042	欲望・期待など	1
1.3045	意志	2
1.3046	誠・徳・義・信・恩など	4
1.3047	信仰・宗教	5
1.3050	練習・まね・学習	1, 5, 6
1.3051	習慣・記憶・記念	6
1.3061	思考・感想・疑い	13, 14
1.3062	注意・認識・了解	8, 12
1.3063	比較・区別・選択・参考	5, 11, 12, 19
1.3065	研究・実験・調査・検査など	4, 7, 8
1.3067	解決・決定・きまり	2, 3, 5, 6, 7, 9
1.3068	承認・肯定など	1
1.3070	意味・問題・趣旨・大綱など	1, 5, 9, 13
1.3072	論理・うそ・誤り・訂正など	1, 2
1.3074	学問・学科	2, 7, 10
1.3075	説・論	1, 3, 5, 6
1.3080	原理・規則・主義	2, 3, 3A, 4, 7, 8, 9A, 12

（次頁へ続く）

項目番号	見出し	該当分類段落
1.3081	方法	4, 6
1.3082	制度・慣例	1
1.3084	計画・案	8, 9, 10, 11, 12
1.3101	言語	1
1.3102	名	4
1.3111	文法	3, 5
1.3113	文字	9
1.3115	表・図・譜・式	1, 9
1.3120	発言・沈黙	14
1.3122	通信	2, 5, 6
1.3123	伝達・報知	5, 10, 11
1.3131	談話	4
1.3132	問答	4
1.3133	会議・論議	1, 2, 4
1.3135	言論	1, 5
1.3140	宣言・発表	1
1.3151	書き	1
1.3153	記名など	1
1.3154	文章	1
1.3155	文書	1
1.3160	文献・図書	1
1.3161	教科書・古典・集・新聞雑誌	14, 16
1.3162	辞書・目録・暦	5, 12
1.3200	創作・著述	5
1.3220	美術	10
1.3240	演劇・映画	3, 4
1.3300	文化・歴史・風俗	1
1.3310	人生・禍福	3, 9, 10, 12
1.3312	出処進退	1A
1.3320	勤務・労働	3
1.3331	住	16
1.3340	学事・兵事	4
1.3360	祭儀・式・宗教的行為	15
1.3371	旅・行楽	1
1.3374	スポーツ	3, 4A, 9
1.3380	騒ぎ	1
1.3390	身振り	1
1.3391	立ち居	0A
1.3393	手の動作	1

（次頁へ続く）

項目番号	見出し	該当分類段落
1.3400	義務	8
1.3401	権利	1, 3B
1.3410	身上	1, 3
1.3440	実行・犯罪	4
1.3450	活動	4
1.3460	用事	1
1.3470	成功・失敗	1
1.3500	交わり	1, 2, 3
1.3501	争い	2
1.3510	集会	1, 2, 3, 4, 5, 7
1.3511	出欠	1, 2
1.3520	対面・訪れ・招き	6
1.3522	―	1, 5
1.3530	約束	5
1.3531	交渉	3, 6
1.3532	賛成・許容	1, 4, 8A
1.3540	協力・参加	5, 7A
1.3542	競争	4
1.3551	戦争	1, 6
1.3560	攻防	1, 5, 9, 12
1.3570	勝敗	2
1.3580	軍事	8
1.3600	支配・政治・革命治乱	9, 13, 14
1.3610	国務	2
1.3611	裁判	1, 3
1.3612	刑	1
1.3620	設立・運営	1, 2, 3
1.3630	人事	12, 18
1.3650	救護・世話	1, 8
1.3660	請求	2A, 4
1.3670	命令・制約	3, 17, 22, 25, 26
1.3682	賞罰	1A, 5
1.3701	所有	1, 10, 13
1.3710	経済・収支	1, 4, 16
1.3721	資本・金銭	6, 18
1.3730	価格・費用	12
1.3740	給与・料金・利子	10
1.3750	損得	9
1.3761	売買	1

（次頁へ続く）

項目番号	見出し	該当分類段落
1.3780	貸借	5
1.3801	産業	1
1.3802	生産	2
1.3822	建設・土木	1, 10
1.3823	建築	2
1.3830	運輸・交通・通信	2, 9
1.3832	出版・放送・興行	1, 5, 8, 9, 12, 15
1.3850	設備・作業・手当て・処理	18
1.3852	使用	5
1.3860	製造・荷造り	3
1.4010	持物・売物・みやげなど	9
1.4020	産物	5
1.4040	貨幣	1, 4
1.4100	資材	4
1.4110	紙	1
1.4150	輪・車・棒・管など	2, 4, 5, 8
1.4153	ばね・栓・キーなど	7
1.4154	型	1
1.4160	コード・なわ・綱など	1
1.4161	網	1
1.4200	衣料・綿・皮・糸	5
1.4230	下着・羽織・ズボン・コートなど	10, 11
1.4240	袖・衿・身頃・ポケットなど	6
1.4251	ネクタイ・帯・手袋・靴下など	5
1.4320	米・糠・小麦粉など	2
1.4323	さかな・鰹節・肉	4
1.4400	住居	6, 7, 10, 12, 14
1.4410	家屋	7, 9, 11, 12, 13, 15, 17
1.4430	へや・廊下・床など	1, 2, 19
1.4450	棚・戸棚・台・壇など	6
1.4511	びん・壷・皿・膳など	14
1.4520	食器	3
1.4570	おもちゃなど	2, 5, 28
1.4580	標識	4, 6, 13
1.4610	鏡・レンズ	1
1.4620	電気機具・装置	3, 13
1.4630	設備・機械	2, 4, 7, 9
1.4650	乗り物（陸上）	3, 8, 14, 15, 17

（次頁へ続く）

項目番号	見出し	該当分類段落
1.4660	乗り物（海上）	2
1.4670	乗り物（空中・宇宙）	3
1.4700	地類（土地利用）	2, 19, 22, 24, 25, 26, 27
1.4710	道路・橋	1, 9, 10, 12, 15, 19, 25
1.4720	その他の土木施設	9, 11, 12, 13
1.5010	光	9
1.5020	色	2
1.5100	自然・物体・物質	2
1.5140	天災	4
1.5150	気象	9
1.5151	風	2
1.5200	宇宙・空	4
1.5210	天体	2
1.5230	地	1, 3
1.5240	地形・山野	2, 3, 13
1.5250	川・湖	1, 14, 17
1.5260	海・島	1, 4, 5, 9, 12, 13, 14, 15, 19, 22
1.5280	地帯	1
1.5500	生	11
1.5520	植物名	26
1.5530	枝・葉・花など	10, 23
1.5640	魚	4
1.5710	頭・目鼻・顔	11, 30
1.5730	手足・指	30
1.5740	膜・筋・神経・内臓	1, 5
1.5770	血・孔・涙・汗・尿など	14
1.5810	生死・生	3, 21
1.5820	死	12
1.5850	病気	14
1.5860	四百四病	14
3.1000	こそあど	8
3.1120	相互・異同	10A, 11
3.1320	特別・異様	3
3.1610	すぐ・次第になど	4, 5B
3.1640	過現未	14
3.1670	翌・次	2, 3

（次頁へ続く）

項目番号	見出し	該当分類段落
3.1710	相対的位置	7A, 9
3.1920	長い・広い	5
3.1921	厚い・太い・大きい	3, 4, 5
3.1950	多い・少ない	2
3.1980	ひとり・みんな	5
3.1992	およそ・かつがつ・最も・もっと	2, 13
3.3040	かしこい・おろか	13
3.3100	ことば	1
3.3300	風俗	8
3.3470	熱心・努力	2
3.3700	経済	6
非収録語	ＳＡ, あて, かぜ, ごろ, づくり, なし, ふ頭, べり, もの, アイスアリーナ, アイランド, アジア大会, イルクーツク, インタチェンジ, キャンパス, コンバインド, シアトル, ジャズマン, スポーツセンター, データベース, ノーマルヒル, フォーラム, ヘリ, モーターショー, ユース, 安保理, 委, 霞, 会派, 海兵団, 開会式, 外相・蔵相会議, 外相会談, 各紙, 学芸員, 簡裁, 競輪場, 橋上, 区内, 警備隊, 建設相, 県内, 厚生年金会館, 国税局, 再選挙, 山ろく, 市域, 支店長, 事務総長, 実験線, 州知事, 諸語, 礁, 訟廷事務所, 陣内, 生態調査, 跡地, 切り崩し, 全国ネット, 早版, 速道, 体育館, 大リーグ, 大地震, 大島, 地下街, 中華街, 展示場, 電, 都立高, 東京本社, 特捜部, 特派員, 南署, 年会, 府警, 府内, 武道館, 副首相, 福祉会館, 平, 並み, 弁護士会, 本港, 予算委, 藍, 理事会, 竜馬, 両社	

（以上）

時相名詞辞書該当分類番号一覧（新分類語彙表）

項目番号	見出し	該当分類段落
1.1000	こそあど	8, 10, 12
1.1101	等級・系列	2, 5
1.1110	関係	2
1.1111	本	3
1.1112	因果	7, 9
1.1113	理由・目的・証拠	2, 6
1.1300	様相	1
1.1520	過程・経路	8
1.1524	通過・普及など	1
1.1527	上り下り	7
1.1528	往復	9
1.1600	位置・地点・場合	1, 4
1.1610	時間	1, 4, 6, 7
1.1611	時機	4, 10
1.1612	時刻	1, 2, 3, 5, 7
1.1613	毎日・平生	1, 4, 5, 6, 10, 12
1.1620	期間	1, 2, 7, 8
1.1621	永久・一生	3
1.1622	年配	6
1.1623	時代	1, 6, 10
1.1624	季節	4, 5, 7
1.1630	年	1, 2, 3
1.1631	節・節日	1
1.1632	月	1
1.1633	週・週日	3
1.1634	日	1, 10
1.1635	朝晩	1, 2, 3, 4, 6, 12, 13, 17
1.1641	現在	1, 2, 4
1.1642	過去	1, 2, 5, 6, 7
1.1643	未来	1, 2, 3, 5, 6, 7
1.1650	順序	8
1.1651	終始	2, 3, 4, 12, 15
1.1652	途中・盛り・次ぎ	1, 1A, 2, 3, 4, 4A, 7, 8, 9
1.1670	前後	1, 3, 4, 6

（次頁へ続く）

項目番号	見出し	該当分類段落
1.1671	翌・後日・当面など	1, 3, 5, 6
1.1700	空間・場所	11, 15
1.1720	範囲・席・跡	1, 2
1.1730	方向・たてよこ	1
1.1731	方面・方角	1
1.1740	左右	5
1.1741	上下	1
1.1742	中・かみ・しも・頂・すみ	1, 5, 9
1.1760	前後・間・端	1, 8, 10, 13, 14
1.1770	内外	2, 3, 6
1.1780	ふち・そば・まわり・沿い	3, 13
1.1820	角・円・三角形など	14
1.1950	一二三	33, 36
1.1951	多少	12
1.1960	単位	7, 11, 12, 13, 14, 17, 27, 28, 35, 40, 51, 59, 71
1.1963	割・前など	2
1.1980	一般・全体・部分	2, 7
1.1981	群・組・対	16
1.1990	程度・限度	1, 7, 10
1.2010	自他	7
1.2020	人間	1
1.2100	家族	4
1.2200	相手・仲間	5
1.2530	国	1
1.2590	個有地名	1, 4
1.2640	事務所・市場・駅など	1
1.2710	政府機関	1
1.3001	感覚	1
1.3004	感情・気分	5, 6
1.3031	声	6
1.3046	誠・徳・義・信・恩など	4
1.3062	注意・認識・了解	8
1.3081	方法	4
1.3300	文化・歴史・風俗	2
1.3310	人生・禍福	3
1.3360	祭儀・式・宗教的行為	18
1.3392	足の動作	2

（次頁へ続く）

項目番号	見出し	該当分類段落
1.3721	資本・金銭	2
1.3730	価格・費用	3
1.3750	損得	1
1.4150	輪・車・棒・管など	6
1.4152	柄・杖・へらなど	9
1.4400	住居	1
1.4430	へや・廊下・床など	1
1.4710	道路・橋	12
1.4720	その他の土木施設	16
1.5010	光	9
1.5100	自然・物体・物質	9
1.5210	天体	2, 3, 5
1.5530	枝・葉・花など	25
3.1000	こそあど	1, 8, 12
3.1120	相互・異同	10A, 11, 12
3.1200	在不在	4
3.1310	普通・非凡	3, 7A
3.1320	特別・異様	1
3.1600	時	2, 22
3.1610	すぐ・次第になど	6A
3.1620	一度に・再び・毎度など	11, 12
3.1640	過現未	7, 8, 9, 11A
3.1650	まだ・もう・もっと・また	12, 24
3.1670	翌・次	2, 3
3.1710	相対的位置	9C
3.1921	厚い・太い・大きい	5
3.1980	ひとり・みんな	13
3.1991	くらい・ほど	2
3.1992	およそ・かつがつ・最も・もっと	2, 16
3.3460	みずから・あえて・ぬけぬけ	10
4.1110	累加	8
4.3100	〔間投〕	1
非収録語	うえ, かわり, きのう, ことし, たび, ところ, なか, はじめ, ゆえ, 為, 一章, 一町, 一発, 一夕, 月比, 頃, 昨夏, 代々, 当初, 半年	

（以上）