

## 蓄積されたニュース番組からの映像クイズ生成手法の検討

佐野 雅規<sup>1,2,3</sup> 八木 伸行<sup>1</sup> 片山 紀生<sup>3,2</sup> 佐藤 真一<sup>3,2</sup>

<sup>1</sup>NHK放送技術研究所 〒157-8510 東京都世田谷区砧 1-10-11

<sup>2</sup>総合研究大学院大学 複合科学研究科 情報学専攻 〒101-8430 東京都千代田区一ツ橋 2-1-2

<sup>3</sup>国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: <sup>1</sup>{sano.m-fo, yagi.n-iy}@nhk.or.jp, <sup>3</sup>{katayama, satoh}@nii.ac.jp

あらまし 本稿では、大量に蓄積されたニュース番組から、映像（画像）付きクイズを自動生成する手法について提案する。クイズの形式には多種多様なものが存在するが、本稿で対象とするのは、映像とそれに関連する説明文とのマッチングを問うクイズである。このようなクイズを生成するためには、映像または説明文同士の「似て非なる」関係を見つけ出す必要がある。また、通常の文字だけのクイズと違い、映像がクイズとして適切である必要があり、どのような画像も利用できる訳ではない。本稿では、これら映像クイズを自動生成する際の問題点を整理し、表層的な解析だけで生成する手法を提案する。また、その実験結果についても報告する。

キーワード ニュース解析, クイズ生成, クラスタリング, 色レイアウト, Local Binary Pattern, tf-idf

### A study on generation of image-based quiz from news video archive

Masanori SANO<sup>1,2,3</sup> Nobuyuki YAGI<sup>1</sup> Norio KATAYAMA<sup>3,2</sup> and Shin'ichi SATOH<sup>3,2</sup>

<sup>1</sup>NHK Science and Technical Research Laboratories 1-10-11 Kinuta, Setagaya-ku, Tokyo, 157-8510

<sup>2</sup>The Graduate University for Advanced Studies, School of Multidisciplinary Sciences,  
Department of Informatics 1-2-1 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430

<sup>3</sup>National Institute of Informatics 1-2-1 Hitotsubashi, Chiyoda-ku, Tokyo, 101-8430  
E-mail: <sup>1</sup>{sano.m-fo, yagi.n-iy}@nhk.co.jp, <sup>3</sup>{katayama, satoh}@nii.ac.jp

**Abstract** This paper proposes a method of generating image-based quiz from news video archive. Although there are various types of quiz, our target is a kind of matching quiz between images and their explanations. To make the quiz, a special relationship that is “close but not the same” should be found among images and/or their explanations. In addition, not all images from news videos can be used to form the image-based quiz because image should have a certain roll as the quiz. In this paper, we clear the problems of making quiz and propose the automatic generating method by using some statistical analyses, then we give experimental results of our method.

**Keyword** News analysis, quiz generation, clustering, color-layout, local binary pattern, tf-idf

#### 1. はじめに

近年の放送業界における急速なデジタル化に伴い、NHKでも番組の有効な活用方法について様々な検討が行われている。従来の有効活用と言えば、一度放送した番組を、ほとんど手を加えずに再び放送する「再放送」を意味した。これは番組全体を利用する活用方法にあたり、この他にも番組のジャンルにもよるが、ビデオ/DVD化、国際市場への番組提供などがある。また、番組のシーン単位での活用としては、新しく番組を制作する際の素材としての再利用や、近々予定されているサーバ型放送を挙げることができる。また、2003年には、NHKが保有する番組や映像を多角的に活用していくための施設としてアーカイブス[1]がオ

ープンしている。更に、インターネットと連携した番組を活用する試みとして、「おこめ」[2]、「南極キッズ」[3]などが公開されており、番組で放送した映像や、放送内容を補足する約350語の用語解説、約200本の映像クリップを中心に、ゲームや掲示板の機能が提供されている。現在NHKでは、放送番組を利用したマルチメディア百科事典の自動生成をめざし、その要素技術の研究[4]を進めている。

本研究も、番組を利用した学習支援の1つとして、ニュース番組からクイズの自動生成を試みる。ここで目標とするクイズは、単なる文字だけのクイズでなく、番組の一番の特徴である映像（画像）を用いたクイズである。今回は、人手によるクイズ生成の予備実験を行い、その知見に基づいた自動生成の手法を検討し、実験を行ったので報告する。

## 2. クイズの自動生成

### 2.1. クイズ生成の目的

一般的に、ニュースを題材にしたクイズは、いろいろなところで見ることが出来る。新聞や雑誌をはじめ、最近では Web 上に最新の時事クイズが掲載されたり [5]、携帯電話に配信されるサービスまである。また、就職試験のための時事問題集などもこれらのも一種と考えることができる。このようにいろいろな媒体で提供されているが、映像を伴ったクイズは見かけない。百聞は一見に如かずという諺にもあるように、映像の役割は大きく、これを利用したクイズは、文字だけのクイズに比べ大きなインパクトがある。また、クイズそのものは、単なる娯楽にとどまらず、楽しみながら学ぶ手段として教育的にも大変有効である。

ニュースを題材にした研究の 1 つに、QA システムがある。QA タスクは、古くから TREC[6]のトラックにもあり、未だホットな話題である。応用アプリケーションとしては、ニュース番組を蓄積しておき、ユーザが質問を投げかけると、その答えとなりそうな映像部分を検索して提示するシステム [7]などが報告されている。一見、クイズ生成のタスクは、この QA システムのタスクと似ているように感じるかもしれないが、実は大きく違う点がある。それは、QA システムの場合は、Q が予め与えられるということである。そして、与えられた Q から、質問の意図を解析し、該当する答えを抽出しなければならない。一方、クイズ生成の場合は、ユーザからの何らかのトリガーはあるかもしれないが、それを基に Q を作成し、その答えも同時に求めておく必要がある。別な表現をすれば、QA のペアを探し出すタスクであり、Q の中身を理解しなくても生成できる。従って、ニュース番組の中で、クイズになりそうな素材（映像、文章）はどういうものであるのか、また、どのような手順で抽出できるのかをモデル化する。

### 2.2. クイズの種類とアプローチ

一言でクイズといっても、いろいろなバリエーションが存在し、対象とするものにより生成のアプローチは異なる。図 1 は、ニュースを題材にした映像付きクイズのパターン例を示している。①は、映像の中身に関することを問う記述式の問題例である。②も同様であるが、前後の説明文をヒントに空欄を埋める方式である。③④は、どちらも選択問題に分類され、③は説明文の真偽を問う問題、④は映像にふさわしい説明文を選ぶ問題である。

次に、各クイズの生成アプローチについて考えてみる。我々が素材として用いるのは、ニュース番組（映像と音声データ）と字幕（アナウンサの喋ったテキストデータ）である。その際、もし意味内容を正確に扱うことができれば、図に示すようなクイズは比較的容易に作成できると考える。しかし、現在の技術では、意味を正確に理解するのは難しい。そこで、統計など表層的な特徴解析を用いるアプローチで、それぞれのクイズの生成手法を考えてみる。

#### ① 次の問に答えなさい（5 WITH クイズ）



これは誰ですか？  
役職は何ですか？



ここはどこですか？  
何が行われますか？

#### ② 括弧の中に単語をいれなさい（穴埋めクイズ）



皇居では、元旦のきょう、  
( ) の儀が行われました。

#### ③ 説明の真偽を答えなさい（〇×クイズ）



第 84 回天皇杯全日本サッカー選手権大会、決勝戦、優勝したのは、東京ヴェルディでした。 ○or×

#### ④ 最も適している説明文は？（選択クイズ）



I. これは鴨川市の海岸に現れた「かもちゃん」です。  
II. これは多摩川に現れた「たまちゃん」です。

図 1：映像クイズの例

①の場合は、映像の被写体についての質問文を生成する必要がある。被写体が人物の場合は、Name-It[8]などの手法を用いれば、人物（顔）と対応する人物名の対応がとれるため、図に示したようなクイズは生成できると考える。被写体が人物以外の場合は難しく、今後の課題である。②の場合は、映像の説明文は字幕をそのまま利用し、クイズになりそうな部分を空欄にすればよい。ニュースの多くの場合、先頭に要約的な説明があるので、その付近の字幕文を選択し、重要と思われる固有名詞の箇所を空欄にする方法が考えられる。③についても、字幕を利用して説明文を作成し、その一部分を入れ替えることで、真偽の問題は作成可能と考える。④の場合は、映像に関連した複数の説明文が必要となるが、これらも字幕を利用することが可能である。この場合、複数選択した説明文を、単に並べるだけでは不十分であり、更に一步踏み込んで「似て非なる」という関係を見つけ出すことが必要である。この「一見すると似ているが、実は全然別のものである」という関係は、クイズの選択肢を生成する上で重要な鍵となる。この意味においては、③は単語レベルで、④は文章レベルで似て非なるものを抽出する必要がある。①と②とは少し違うアプローチが必要となる。

本研究では、最初に目標とするクイズを④のクイズに限定し、選択肢の鍵となる「似て非なる」関係を抽出しクイズの生成を行う。

### 3. 生成手法の検討

図1の④をベースとするクイズとしては、映像と説明文の数により、3タイプが考えられる(図2)。左端は、映像1枚に対し、映像を説明しているとされる文章が複数あるもの、真ん中はその逆のパターン、右端は映像も説明文も複数あるタイプである。どのパターンにおいても、複数の映像間、もしくは説明文の間には、似て非なる関係が必要となる。今回はまず、図2左端の、映像1枚に対し複数の説明文が選択肢となっているクイズの生成を行う。

はじめに、人間がこのようなクイズを作成する場合の手順を考えてみる。ニュース番組から興味をひきそうな映像、言い換えるとクイズになりそうな、「これは何だろう? どういうニュースだろう?」と思う映像を探す。次にその映像に関するニュース(字幕に相当)を聞き、被写体は何であって、どういう意図でニュースになっているのかを理解する。被写体そのものと、ニュースの意図を考えつつ、選択した映像に対する正解の文章とおもしろくて間違えやすい不正解の文章を作成し、クイズを完成させる。この作業をフローチャートにすると図3のようになる。以下に、各手順について、手法の検討と今回の実装方法について説明する。

#### 3.1. クイズに使う映像の選択

はじめに、どのような映像(画像)がクイズになり得るのかを調べるため、予備実験として人手によりクイズの生成を試みた。クイズに利用する映像(画像)は、あらかじめニュース番組をショット分割し、各ショットの先頭フレームを用いた。ある1日分のニュースの映像(画像)全てに対し、字幕を参考にどのようなクイズが作成できるのかを試した。その結果、内容によって映像クイズにふさわしいものと、ふさわしくないものがあることがわかった。

例えば、ニュース番組の各項目の最初は、スタジオでのアナウンサのバーストショットが多く、この映像はクイズにする意味がない。同様に、中継現場で取材にあっている記者のバーストショットもふさわしくない。また、人混みが映っている映像や会議の映像などは、選択肢(説明文)をつくるのが難しい。なぜなら、実際のニュースの中で人混みの映像は、その日の暑さをテーマにした街頭を歩き交う人を映した場面であったり、サラリーマンの給料についての項目での労働力を表す映像であったり、その意図するものが様々なものである。何通りもの解釈が容易にできる映像は、説明文との対応も曖昧になりやすく、自動で生成するクイズには不向きである。このような映像は、政治、事件、スポーツなどの項目に比較的多い。また、人物アップ(顔アップ)の映像についても、今回のクイズ対象からは外す。ニュース中の人物アップの映像は、その項目でのキーマンが発言をしている様子か、またはインタビューに答えている場面が多い。このような映像は、図1の①のような「これは誰ですか?」というクイズには向いているが、④のような内容に関わる説明文との対応を問うクイズには向いていない。従って、大勢の人が映っている、もしくは特定の人が大きく映って



図2: マッチングクイズ (映像: 説明文)

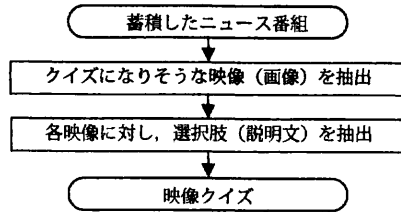


図3: 映像クイズ生成の手順

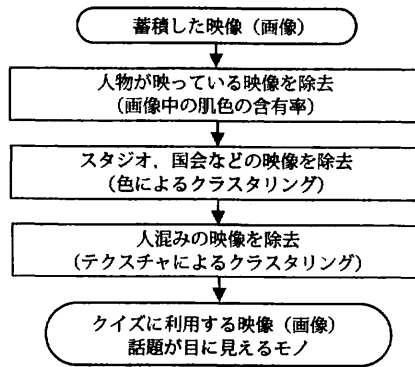


図4: クイズになりそうな映像抽出

いる、どちらの場合も、今回の映像クイズには適さないと判断した。これらに対し、あるモノの紹介、イベントの紹介などの場合は、映像にそのモノ、もしくはシンボリックな場面が映っている場合も多く、映像クイズを作りやすいと考える。別の表現をすれば、ニュースで話題となっているものが目に見えるものといえる。

次に、予備実験で得た知見を基に、クイズになりやすい映像抽出の手法を図4に示す。基本的なアイデアは、大量に蓄積した映像(画像)から、クイズになりにくい映像を除外していく方法である。図4に示すように、まず人物が映っている映像を除き、続いてスタジオや国会などの映像を除く、更に人混みの映像を除くことで、クイズになりやすい映像が得られると仮定している。

各種画像の除去処理のために、今回実装して用いた方法について説明する。採用した方法は、図4の括弧内に示されている。まず、人物が映っている映像については、本来であれば顔の検出を行い人物かどうかの判断するのが理想である。しかし、顔が小さかったり、横などを向いている場合も多く、完全に検出することは難しい。そこで、手荒な手法ではあるが、今回は肌色という色だけに注目し、その画像中に占める割

合が小さいものを、人物が映っていない映像と仮定する。

次に、スタジオや国会、人混みの映像については、大量に蓄積した画像から、ある特徴を用いてそれぞれのクラスタを形成し、一度に取り除く方法を採用する。これと対照的な方法としては、1つ1つの映像を判別していく方法が考えられるが、判別すべき映像について全ての手法を開発することは困難である。我々は、ニュース番組内の映像にもパターンがあることに注目し、大量の映像データからそのパターンを見つけ、クイズになりやすいかどうかの判断に用いる。ここでは、ある意味をもった映像は、似た映像特徴を持っているであろうという仮定をしている。例えば、扱うニュース内容によっても異なるが、ある意図をもって人や建物を映すときはこのように撮る、といった撮影のパターンや、取材場所の制約などから、ある映像については常に決まった背景をもっているなどの事例をあげることができる。このような例から、映像の低レベル特徴に注目するだけでも、いくつかの意味をもった集合が得られると考える。今回、スタジオショットや国会などの映像については、色の位置による特徴（以後色レイアウトと呼ぶ）でクラスタリングを行う。同様に、人混みの映像については、画像に含まれる空間周波数に注目する。特徴の候補としてDCTやガボールなどが挙げられるが、今回はテクスチャ解析でもよく用いられるLBP(Local Binary Pattern)[9]を用いる。

ここまでの処理で、映像クイズに利用できそうな映像群を得ることができる。

### 3.2. 選択肢の抽出(似て非なるモノ)

クイズ用に抽出された各映像に対し、選択肢となる説明文を自動生成する手法について述べる。この説明文には、ニュース番組に付随する字幕データを利用する。選択肢として複数の説明文を抽出するが、そのうちの1つは映像に対応する説明文であり、選択クイズの正解となるため正解文と呼ぶ。他の説明文はクイズを形成するためのもので、偽り文と呼ぶ。

はじめに、映像と字幕データの対応関係について説明する。図5は典型的なニュース項目を示しており、スタジオショットではじまる複数の映像と、複数の字幕データから構成されていることを示す。字幕データは、時刻情報と文章がペアになったものであるが、必ずしも1文1文に時刻情報があるわけではなく、複数文にまとめて付与されていることもある。図5では、字幕(2)の時刻が、映像No.2のショット中に含まれていて、3つの文章が含まれていることを示す。この場合の映像と字幕の対応は、映像No.2~4が字幕(2)に対応すると考える。すなわち、ある映像に対応する字幕データを探す場合、自分の映像区間に字幕データがなければ、時間的にさかのぼって探す。また、ニュース項目の冒頭がスタジオショットではなく、更に最初の数カットにアナウンサのコメントがないような演出もある。このような映像については、そのニュース項目で最初に現れる字幕データを対応する字幕と考える。

次に選択肢の生成について説明する。図6に示すように、はじめに正解文を抽出し、その後、偽り文(正

(No.)	映像	(番号) 字幕データ
No.1		(1) 今日、〇〇××が行われました。
No.2		(2) 気温40度近い中で整列する人々。気を緩めることなく、一心不乱で訓練に励みました。中には熱中症にかかり、倒れる人まで現れました。
No.3		
No.4		
No.5		(3) 夕方には全てが終わり、参加者達は世界各国へ帰っていきました。

図5：典型的なニュース項目

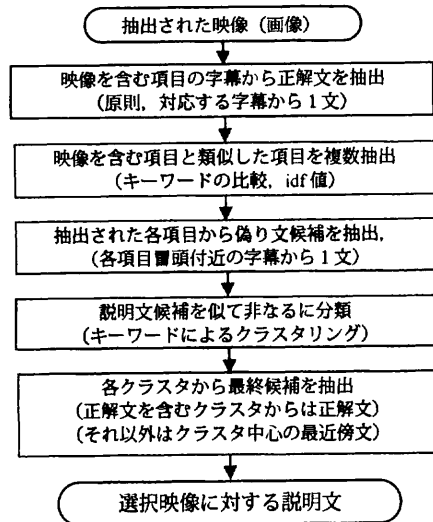


図6：説明文の抽出手法

解文に似て非なる文)を複数抽出する。この処理では、説明文の選び方と説明文同士の「似て非なる」の判断方法が鍵となる。説明文の選び方には、構文解析を利用した文の構造を考慮する方法から、単純なキーワードの出現により選択を行うものまでいろいろと考えられる。また、文と文の類似度を計算する場合にも、文から構文木を作成し、その類似性を考慮するものから、先と同様キーワードにより比較する方法がある。今回はどの処理においてもキーワードに注目して処理を行う。以下に今回の手法を説明する。

(1) 正解文の抽出：ニュースの映像は、そのときアナウンサが説明している内容に何らかの関連があるため、原則、時間的に一番近い字幕データから正解文を選ぶ。その際、文末の表現による選択条件を課す。これは、最初に行った予備実験の結果、文末によりクイズの説明文としての向き不向きがあることがわかったからである。例えば、疑問文や「～としています」、

「～していきましょう」などの文は、映像クイズの説明文には向かないのである。もし、映像に対応する字幕データに条件にあう文がなかった場合は、そのニュース項目の冒頭の字幕から再検索し、最初にみつかった文を正解文とする。冒頭から探す理由は、ニュースの伝え方が、最初に概要を説明し、その後、重要な順に説明をしていく傾向があるからである。

(2) 偽り文の抽出： 選択された映像を含む項目の字幕データからキーワードを抽出する。そして、検索対象の全項目に対し、そのキーワードに注目した類似度を計算し、類似している複数の候補を抽出する。次に各項目から偽り文の候補を1つ抽出する。抽出ルールは先の正解文を抽出する場合と同じで、各項目の冒頭から探す。更に、この複数の偽り文候補に正解文を加え、各文から生成したキーワードベクトルでクラスタリングを行う。このときのクラスタ数は、最終的に取得したい説明文の数と同じにする。クラスタリングの後、正解文を含むクラスタからは正解文を、その他のクラスタからはそれぞれ1文を偽り文として抽出し、クイズのための説明文抽出が完了する。

ここで、「似て非なる」偽り文を抽出するクラスタリング処理について、図7を使って説明をする。似て非なる文とは、出現する単語は似ているが、主題が異なると定義する。まず、抽出された説明候補文（正解文を文1とする）からそれぞれキーワードを抽出する。次にキーワードの出現頻度を調べ、ある閾値以上出現する単語にのみ注目し、各候補文の特徴ベクトルを生成する。このとき、正解文に含まれるキーワードを1つも持たない候補文があれば除く。図7の例では、文4は、Cという単語が1回出現するが、文1（正解文）が含んでいるA、B、D、Eという単語は1つも含んでいない。このような場合、文4は候補から除く。最終的に残った候補文の特徴ベクトルでクラスタリングを行う。この一連の処理により、対象データは全て類似した単語を含んでいるが、文毎の微妙なキーワードの含み方の違いによりクラスタが形成され、これが似て非なるに相当する。

## 4. 実験と考察

### 4.1. 映像の選択の実験

映像選択の実験は、NHKのニュース7、2005年の6月分を対象に行った。今回は、あらかじめ気象情報のコーナーは除外している。一日分のニュースでは、約200枚の映像（画像）候補があり、1ヶ月では約6500枚となった。これらに対し図4の処理を行った。

人物が映っている映像を除くための肌色は、あらかじめ無作為に選んだ人物画像10枚から、肌の部分を人手で切り取り、HSB空間で和をとったものを用いた（ $0 < H < 42, 0.08 < S < 0.97, 55 < B < 230$ ）。この肌色の画像中における含有率を計算し、含まない方から全体の約1割（画像中の1%が肌色）約600枚を選択した。

次に、同じ1ヶ月分の画像に対し、色レイアウトによるクラスタリングを行った。画像は縦横3等分し、9ブロック毎に色ヒストグラムを生成した。色は、HSBをそれぞれ8.3.3に分割したので、1ブロックから72

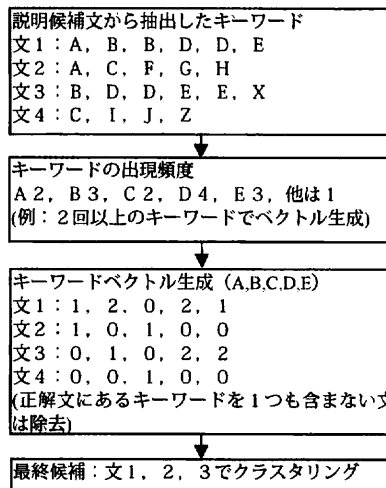


図7：キーワードによるクラスタリング

次元のベクトルが生成され、画像全体では648次元のベクトルとなった。クラスタリングは、K-means法を用い、初期値には6月1日の画像（約200枚）を用いた。クラスタが形成された後、各クラスタ内に含まれる画像を調べ、その分類精度や、連想されるコンセプトを調べた（表1）。図8に実際のクラスタ内の映像例を示した。スタジオショットには、色の特徴が安定しているため、そのパターン毎にほぼ間違いを含まないクラスタを形成していた（図8の上段と下段は別々のクラスタ）。同様に、画像中の緑色の含み具合で形成されたクラスタのいくつかは、スポーツ映像の集合となっていた。その他、茶色を基調としてクラスタの中には、国会の映像のクラスタがあった。これらのクラスタリング結果を利用して、人物が映っていない約600枚の画像の中で、表1の「屋外・風景」と「その他」のクラスタに含まれる画像を残し、次の処理の入力とした。

最後は、テキスト特徴を用いたクラスタリングによる、人混みの映像除去である。色レイアウトの場合と同様に、9ブロックに分割し、1ヶ月分の画像に対して行った。テキスト特徴として用いたLBPは、各画素においてその8近傍との輝度差により0～255の値を出力するものを用いた。この0～255を15等分に分け、そのヒストグラムをブロックごとに作成した。従って、1枚の画像から115次元の特徴ベクトルが生成され、これらをクラスタリングした。表2は、LBPによるクラスタリングの結果である。結果からもわかるように、スタジオショットのクラスタも形成されていた。これは背景がほぼ同一のため、その部分のテキスト特徴が類似したためである。ここでは、人物の細かい画像（人混み）とスタジオのクラスタに属する画像が除外され、最終的に388枚の画像がクイズの候補として残った。実際に取り除いた人混みクラスタの映像例を図9に示した。LBPは色を扱っていないため、人混みが細かい場合は、自然の木々などの細かいテキストと混同されることがわかる。図10は最終的に残ったクイズの映像候補例である。

表1：色レイアウトによるクラスタリング結果

分類	クラスタ数	画像枚数	割合[%]
スタジオ	11	413	6.4
人混み	10	710	11.0
スポーツ	13	522	8.0
顔アップ	10	294	4.6
屋外・風景	4	123	1.8
国会	2	63	0.9
バドジョット	1	39	0.5
その他	174	4340	66.8

表2：LBP レイアウトによるクラスタリング結果

分類	クラスタ数	画像枚数	割合[%]
細かい(人)	4	241	3.7
細かい(風景)	4	370	5.7
無地, 荒い	13	413	6.4
スタジオ	10	408	6.3
その他	194	5072	77.9

#### 4.2. 説明候補分の抽出実験

4.1でクイズ用に抽出された映像388に対し、図6に示した説明文の抽出を行った。説明文の候補抽出には、2005年1年分の字幕を用いた。まず、映像に対する正解文を抽出する際の文末表現は、「～ました」及び「～です」で終わる文章に限定した。次にキーワードに注目した類似項目の抽出処理であるが、名詞一般、名詞-固有名詞-[全て]、名詞-サ変接続（これらは茶筌の結果で用いられている表現）を用いて行った。項目間の類似度は、検索元となる項目中のキーワードが、検索対象の各項目の中にどれほど出現するかを数えたもので、キーワードの重複は数えず、あらかじめ検索対象全体で計算した各キーワードのidf値を重みとしてかけた。実験では15個の類似項目から、それぞれ偽り文候補を抽出し、更に似て非なるものを抽出するためのクラスタリング処理を行った。クラスタリングに用いるベクトル生成では、2回以上出現したキーワードを用いた。1回しか出現しないキーワードを含めると、そのキーワードが「非なるもの」として強く作用してしまい、3回以上にすると、ベクトルの次元が小さく分解能が低下するためである。実験では、ベクトルそのものが生成できない例が46件あった。最後に、形成されたクラスタから偽り文を抽出する方法であるが、そのクラスタ中心が一番近い文を選択した。これは中心に近い文ほど、そのクラスタの共通の特徴をもっており、代表として選出するのに良いと考えたからである。最終的には342の映像クイズが生成された。

#### 4.3. 自動生成されたクイズの評価と考察

自動生成された342個のクイズを調べたところ、重複したクイズが多くあった。これは今回の手法では、映像が同じ項目に含まれていると、その偽り文の抽出処理が全く同じとなることに起因している。最終評価を行うために、人手により一番良さそうなものを選択し、最終的に172項目についてのクイズが生成された。

ここで筆者が「ほぼクイズになっているだろう」と考えることができたものは6個で、もう少しでなりそうというものは10個であった。図11に生成したクイズを示した。クイズ1～3は、ほぼ成功といえる例、クイズ4.5は惜しいもの、クイズ6.7は失敗例を示している。クイズ4の場合は、映像をよくみるとチンパンジーが映っており、答えがわかってしまう。もし絵画だけが映っているとすれば、選択肢のどれもが絵画について説明しておりクイズになった例である。クイズ5は逆のケースで、選択肢が各航空会社でわかれているので、映像に映っている飛行機がもう少し横向きで会社名がわかれば、明確な対応付けができクイズになった例である。クイズ6は、映像そのものがふさわしくない例である。これは番組制作者が「暑さ」をイメージさせるために用いた「太陽」の映像をであるが、この映像から選択肢を特定するのは難しい。また、クイズ7は説明文の文末が「伝えました」「打診していました」となっており、映像からは判断できない不適切な選択肢の例である。もう少し文末付近の特徴を検討する余地があると考え。また、今回の手法では、映像選択には映像だけの特徴を用い、説明文抽出には説明文における特徴のみを用いており、両者の情報をお互い用いて絞り込むことで、よりクイズの質を向上させることができると考える。

#### 5. おわりに

本稿では、蓄積されたニュース番組から、映像クイズを自動生成する手法について報告した。今回の実験では、生成に成功したクイズの数は少ないが、もう少し手法を改善することで、表層的な解析でも多くのクイズが自動生成できそうなのがわかった。今後はクラスタリングに用いる特徴を再検討し、Co-Clustering手法やマルチモーダルな処理も検討したい。また、他のタイプのクイズについても自動生成を試していく。

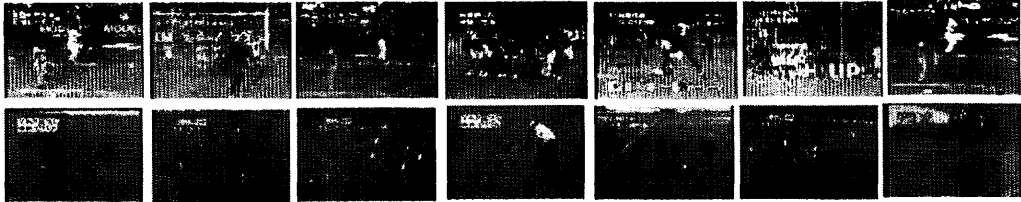
#### 文 献

- [1] <http://www.nhk.or.jp/nhk-archives/>
- [2] <http://www.nhk.or.jp/okome/>
- [3] <http://www.nhk.or.jp/nankyoku-kids/>
- [4] 三浦, 山田, 住吉, 八木, “クローズドキャプションを利用した映像主被写体の推定方法,”
- [5] <http://www.kentei.com/fjji/week/index.php>
- [6] <http://trec.nist.gov/>
- [7] H.Yang, L.Chaisorn, Y.Zhao, S.-Y.Neo, T.-S.Chua, “VideoQA: question answering on news video,” ACM Multimedia, pp.632-641, 2003
- [8] Shin'ichi Satoh, Takeo Kanade, Name-It: Association of Face and Name in Video. In Proc. IEEE Conference on Computer Vision and Pattern Recognition, pp.368-373, 1997.
- [9] T.Ojala, M.Pietikainen and D.Harwood, “A Comparative Study of Texture Measures with Classification Based on Feature Distributions,” Pattern Recognition, 29 (1), pp.51-59, 1996.
- [10] I.S.Dhillon, S.Mallela and D.S.Modha, “Information-Theoretic Co-clustering,” Proc. The 9th ACM SIGKDD, pp.89-98, 2003.

■スタジオショット (スタジオセットの青色) 上段：番組冒頭、 下段：ニュース項目の冒頭



■スポーツ (上段：緑+他色、 下段：ほぼ緑色) 主に、ゴルフ、サッカー、野球のシーン



■国会のシーン (茶色が基調)



■人物シーン (中央に肌色)



■風景のシーン (上側に青) 空や海に相当



図8：色レイアウトによる形成されたクラスタ

■人混みのシーン (細かいテクスチャ) 人だけでなく、自然の木々も混じっている



図9：テクスチャ・レイアウトにより形成された人混みクラスタ



図10：クイズ用の映像として抽出されたもの

### ■クイズ1■



- (1) 人間国宝で上方歌舞伎の第一人者の、中村雁治郎さん、上方歌舞伎の祖と呼ばれる、坂田藤十郎を231年ぶりに襲名することになりました。
- (2) 落語の大名跡の柳家小さん、先代は、落語家として初めて人間国宝になるなど、円熟した芸で親しまれました。
- (3) 上方落語の第一人者で、桂三枝さんや、桂文珍さんなど多くの弟子を育てた桂文枝さんが、きょう亡くなりました。

### ■クイズ2■



- (1) 江戸時代前期の僧、円空が彫った仏像21体が盗まれました。
- (2) 江戸時代に建てられ国の重要文化財にも指定されている、香川県琴平町の通称金丸座で、きょうから恒例の四国こんびら歌舞伎大芝居が始まりました。
- (3) 世界初の全身麻酔手術に成功した、江戸時代の外科医華岡青洲その青洲が考案した薬が見つかりました。

### ■クイズ3■



- (1) 大手電機メーカーのシャープは、市販されるものとしては世界で最大の65型の液晶テレビを発表しました。
- (2) 東京都内の大手デパートに早くもお中元専用のコーナーが登場しました。
- (3) 急速に普及する液晶テレビその下になった液晶ディスプレイの試作品を、37年前に世界で初めて開発したのは、アメリカの技術者ジョージハイルマイヤー氏です。

### ■クイズ4■



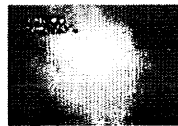
- (1) オークションに出品されたのは、イギリスの研究者に訓練を受けていた、コンゴという名前のチンパンジーが、半世紀前に描いた3点の抽象画です。
- (2) プロ野球巨人の長嶋終身名監督がマヒしている右手で書いたサイン入りカレンダーがチャリティーオークションに出品されることになりました。
- (3) 光と影を鮮やかに描き出す、レンブラント特有の筆使い、17世紀を代表するオランダの画家レンブラントの作品かどうか意見が分かれていた、4点の肖像画は、いずれも本物とわかりました。

### ■クイズ5■



- (1) 今月5日長崎から羽田に向かっていた全日空機が、高度計の故障に機長の判断ミスも加わって、管制官の指示より1600mも高い高度をおよそ40分間飛行していたことがわかりました。
- (2) 次に、石川県の小松空港で先月、エア・ニッポンが運航を受託している全日空の便が管制官の許可がないのに離陸しようとしたトラブルで、会社側は社内調査の結果を報告しました。
- (3) おととい新潟空港で管制官が日本航空機のパイロットに、航空管制に必要な旅客機の認識番号を伝え忘れ、パイロットもそれに気づかないまま離陸していたことがわかりました。

### ■クイズ6■



- (1) さて、ネクタイを外して上着も脱いで勤務するノーネクタイノー上着、温暖化対策の一貫としてきょうから、国のすべての省庁など、総理大臣官邸で始まりました。
- (2) 夏場の軽装として、沖縄の人たちに親しまれているかりゆしウェア、地球温暖化対策として、各省庁で来月1日からノーネクタイノー上着の運動が始まるのを前に、細田官房長官らが、展示即売会を訪れ取り組みへの理解を呼びかけました。
- (3) 小泉総理大臣は各閣僚に率先した取り組みを求めました。

### ■クイズ7■



- (1) 河野衆議院議長は、先週総理大臣経験者5人と会談した際に出た意見を伝えました。
- (2) 政府は小泉総理大臣の靖国神社参拝などでぎくしゃくしている中国や韓国との関係改善に向けまして、首脳会談を打診していました。
- (3) 公式的には、今言われましたように緊急の公務ができたからだと説明しているんですけども、しかし、首脳との会談を土壇場でキャンセルして帰国しなければならない公務があったいんなのかはまったく明らかにされていないんです。

図11：自動生成したクイズ（問題：映像にもっとも関連する説明文はどれでしょう？）