

ニュース映像構造化のための
視覚情報の役割解析とその応用に関する研究

佐野 雅規

博士（情報学）

総合研究大学院大学
複合科学研究科
情報学専攻

平成 20 年度
(2008)

2009 年 3 月

本論文は総合研究大学院大学複合科学研究科情報学専攻に
博士（情報学）授与の要件として提出した博士論文である。

審査委員：

主査	佐藤 真一 教授	国立情報学研究所 / 総合研究大学院大学
	片山 紀生 准教授	国立情報学研究所 / 総合研究大学院大学
	曽根原 登 教授	国立情報学研究所 / 総合研究大学院大学
	相澤 彰子 教授	国立情報学研究所 / 東京大学
	井手 一郎 准教授	名古屋大学 / 国立情報学研究所

A study on structuring news video
based on visual information and its applications

Masanori Sano

DOCTOR OF
PHILOSOPHY

Department of Informatics
School of Multidisciplinary Sciences
The Graduate University for Advanced Studies (SOKENDAI)

March 2009

A dissertation submitted to
the Department of Informatics,
School of Multidisciplinary Sciences,
The Graduate University for Advanced Studies (SOKENDAI)
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Advisory Committee:

Prof. Shin'ichi Satoh (Chair)	National Institute of Informatics/ The Graduate University for Advanced Studies
Assoc.Prof. Norio Katayama	National Institute of Informatics/ The Graduate University for Advanced Studies
Prof. Noboru Sonehara	National Institute of Informatics/ The Graduate University for Advanced Studies
Prof. Akiko Aizawa	National Institute of Informatics/ The University of Tokyo
Assoc.Prof. Ichiro Ide	Nagoya University/ National Institute of Informatics

論文要旨

昨今の放送と通信の融合により、様々な情報伝達手段が整備され、コンテンツの視聴スタイルも大きく変化してきている。放送局においても、従来の番組だけでなく、多種多様で魅力的なコンテンツの制作が必要となっており、放送済みの番組を再利用した新しいサービス展開への模索が行われている。本論文は、この番組を再利用したコンテンツ制作を、より効率的に進めるための支援技術について提案するものである。本提案技術の活用場所は、放送局内だけにとどまらず、一般の家庭においても、ハードディスクレコーダの出現により規模は違うが同様な環境が生まれつつあり、適用することができる。これにより、放送局には、番組を活用した新サービスを効率よく研究開発する環境を提供し、一般家庭には蓄積番組を利用した新しい視聴スタイルを提供することができる。また、ビジネス的な要素も取り込むことで、研究開発機関、サービス提供機関、ユーザの3者にメリットがあるような環境構築を目指すものである。

第1章では、本研究の背景として、大規模な映像コンテンツ提供組織である放送局を例に挙げ、昨今のコンテンツ制作環境について述べる。次に、本論文でニュース番組に着目した理由を述べた後、ニュース番組の再利用の現状について、主に従来研究を中心に整理を行う。そして、そこに存在する技術的問題点を明らかにする。

第2章では、蓄積したニュース番組を効率よく再利用するための共通基盤(フレームワーク)として EvenTank(イベタンク)を提案する。この EvenTank は、研究開発の段階から、実際のサービス提供までを、複数の機関が連携して作業を進められるよう、共通のテストベッドを提供することが目的である。はじめに、関連研究として、既存のメタデータ規格や共通基盤を取り上げ、それぞれの問題点を明らかにする。次に、これら既存技術をベースとし、ニュース番組の効率よい活用のための仕組みを加えることで EvenTank を提案する。具体的には、扱うデータモデルの変更とそのデータをやり取りするインタフェースの変更である。そして、従来から行われている様々なニュース活用の研究が、この EvenTank において、脱着可能なモジュールとして実装可能になっていることを示し、本提案の有効性を確認しつつ、その将来への拡張性と柔軟性についても確認する。

第3章では、ニュース番組活用の観点から、その映像をより適切に再利用するために必要なメタデータについて提案する。はじめに、従来おこなわれてきた映像に対するメタデータ付与について整理をし、適切な再利用のためには映像の役割に注目する必要があることを明らかにする。そして、この役割抽出のために、ニュース番組そのものの役割と、その中での映像の役割について再考し、番組制作手法を考慮したアプローチを提案する。最後に、実際のニュース番組を用いて、いくつかの映像の役割情報が、制作手法を考慮した手法で抽出できることを確認する。

第4章と第5章では、第2章で提案するニュース番組再活用のための共通基盤の有効性について実証することを第1の目的とする。その上で、ニュース番組の新しい活用例として、第4章では時事に関するクイズの自動生成手法を提案し、第5章では時事に関するインタラクティブ・サマリの自動生成手法を提案する。どちらのコンテンツにおいても映像や画像を再利用しており、第3章で提案した映像の役割に注目したメタデータの有効性についても同時に検証を行う。

第4章において生成対象とするクイズは、ニュース番組内の発話データだけでなく画像も再利用したクイズであり、ニュース映像が伝えようとしている「主被写体」に注目して自動生成される。この画像付きクイズの生成手法については、3つのサブタスクに分割し、それぞれに対する工学的アプローチを提案する。そして、実際のニュース番組を用いて生成実験を行い、各サブタスクへのアプローチと、それを総合したアプローチについて評価を行い、最後に EvenTank の枠組みの中でどのように実現しているのかについて検証する。

第5章で生成するインタラクティブ・サマリは、あるニューストピックについて、画像もしくは映像付きのサマリを提供するものである。ユーザへの表示として、伝えたいこと、人物別、時系列の3つの表示方法を備え、注目したトピックに対し、どのような事が起こっていて、誰が関係し、どのような流れなのかをわかりやすく画像や映像付きで閲覧できるよう設計されている。本章において特に注目した問題点は、各ニュース項目を表現するための代表となる画像や映像、また文章について、どのように適切なものを選択するかである。このために第3章で提案した映像の撮影ルールや、番組編集のルールなどを利用した役割に注目したアプローチを提案し、その有効性を確認する。

第6章は、本論文の成果をまとめる。本論文の成果は、コンテンツ制作を支援するために、蓄積された放送済みニュース番組を効率よく利用するためのフレームワークを提案し、このフレームワークに沿った形で新しいサービスのコンテンツ生成実験を行い、その有効性を示したことである。さらに、映像の適切な再利用を高めるために、その映像の役割に注目すべきことを提案したこと、また、そ

の自動抽出について、番組制作におけるルールを利用した手法を提案し、新しいサービスのコンテンツ生成実験の中で検証を行い、その有用性について確認したことにある。従って、本論文で提案した支援技術は、放送局をはじめ、一般家庭においても、蓄積したニュース番組をより効率的に活用する共通基盤を提供し、更に、映像の役割に注目したメタデータの抽出・活用を示したことで、今後の新しいコンテンツの開発制作に大きな貢献をしたと考える。

Abstract

Thanks to the integration of broadcasting and communication, a variety of methods to transmit information have been created, and the styles of experiencing contents are also undergoing great changes. For broadcasters, there is the now the need to not just to produce conventional programs, but also to create diverse kinds of attractive related contents. They are seeking development of new services for reusing programs after the programs have been broadcast. In this thesis, we propose an assistive technology to more efficiently advance the production of reusable program contents. The proposed technology is not limited for use in broadcast stations. It can be applied to general consumer use, thanks to the emergence of hard disk recorders. While the scale is different, the situation being able to store many programs and access them instantly, so the technology can be applied. The technology provides an environment for more efficiently researching and developing new services to reuse programs for broadcasters, while offering new styles of viewing and listening to the new services for general households. By incorporating business considerations, we seek to create an environment that has benefits for R&D institutions, service providers, and users.

In Chapter 1, as background of this research, we describe the current content production environment of large-scale video content providers, using broadcasters as an example. Next, we describe reasons for focusing on news programs in this thesis. We then review the current state of reuse of news programs, with a focus on previous research, and clarify existing technological issues.

In Chapter 2, we propose EvenTank, a common framework for efficiently reusing accumulated news programs. EvenTank was created to provide a common test bed so multiple institutions could collaborate from R&D to the delivery of actual services. First, for collaborative research, we discuss existing metadata specifications and common frameworks, and clarify problem points in each of these areas. Next, we propose EvenTank by adding a framework to the base of these existing technologies in order to efficiently reuse news programs. Specifically, Eventank modifies

the handling data model and the interface for exchanging the data. Next, we show that previous research on reusing various news contents can be implemented as removable modules in EvenTank. We test the effectiveness of the proposed framework, and confirm its expandability and flexibility.

In Chapter 3, we propose metadata required to more appropriately reuse video from the standpoint of reusing news programs. We review the previous assigning of metadata to video, and clarify the need to look up the role played by a video footage for appropriate reuse. Next, to extract this role, we re-consider the role of a news program and the role of video footages within a news program, and propose an approach that takes into account the methods of program production. Finally, using an actual news program, we confirm that the role-based metadata of several video footages can be extracted with this method.

In Chapter 4 and 5, the first goal is to experimentally prove the effectiveness of the common framework for reusing news program, as proposed in Chapter 2. Next, as an example of the new utilization of news programs, in Chapter 4 we propose a method for automatically generating a quiz related to current events. In Chapter 5, we propose a method to automatically generate an interactive summary of current events. For both of these contents, video and images are reused, so we also test at the same time the effectiveness of metadata that pays attention to the role played by video footages, as proposed in Chapter 3.

The generated quiz in Chapter 4 reuses not only spoken text from the news program but also images. It is automatically generated by focusing on the principal object shown in a news video. The method of generating this image-based quiz is divided into three subtasks, and we propose engineering approaches for each of these subtasks. Next we carry out a test using an actual news program, and evaluate the approaches to each subtask and the overall approach. Finally, we test the implementation of the approaches within the EvenTank framework.

The generated interactive summary described in Chapter 5 provides image and video-added summaries for news topics. Three methods of display — by subject to be conveyed, by human subjects, and by time — are shown to the user. The summary is designed to allow the user to view video footages and images that make it easy to understand a news event, the persons involved, and the flow of events surrounding the news. An issue of special importance here is how to select the appropriate representative image and video, as well as text, for each news

topic. Therefore we propose an approach that utilizes the rules of capturing video proposed in Chapter 3 and the rules of program editing, and test the effectiveness of this approach.

In Chapter 6, we organize the results of this thesis. To support the production of contents, we tested a new service of generating contents according to the framework we proposed for efficiently using stored, already-broadcast news programs. The results demonstrated the effectiveness of this framework. Furthermore, to increase the appropriate reusability of video, we performed verification during the experiment on the new service's content generation to test the proposal of looking up the roles of video footages and the proposed method of using program production rules to automatically extract the roles. The results show the effectiveness of these proposals. Based on these results, we believe that the assistive technology proposed in this thesis is a major contribution to the development and production of new contents in the future, not just for broadcasters, but also for general households. It accomplishes this by offering a common framework that utilizes more efficiently accumulated news programs, and furthermore, by demonstrating the extraction and use of metadata that describes the roles played by video footages.

目次

論文要旨	i
Abstract	iv
第1章 序論	1
1.1 背景	1
1.2 ニュース番組再利用の現況と従来研究	3
1.2.1 サマリ機能	4
1.2.2 質問応答機能	7
1.2.3 統合検索機能	9
1.3 技術的課題	11
1.3.1 ニュースコンテンツを扱う共通基盤の整備	11
1.3.2 再利用のための映像とテキストの関係抽出	12
1.4 本論文の構成と概要	12
第2章 EvenTank の提案	14
2.1 はじめに	14
2.2 EvenTank とは	14
2.3 要求項目の整理	17
2.4 既存の関連メタデータとフレームワーク	18
2.4.1 ダブリンコア (Dublin Core Metadata Element)	18
2.4.2 MPEG-7	19
2.4.3 NewsML	21
2.4.4 MPF (Metadata Production Framework)	23
2.5 EvenTank の提案	27
2.5.1 EvenTank のデータモデル	28
2.5.2 インタフェース	32
2.6 フレームワークの従来研究への適合性	34
2.7 まとめ	36

第3章	ニュース番組における映像のメタデータ	38
3.1	はじめに	38
3.2	従来研究と問題点	38
3.2.1	従来研究	38
3.2.2	再利用における問題点	40
3.3	ニュース番組における映像の役割解析	41
3.3.1	ニュース番組の役割とその制作	41
3.3.2	映像(ショット)の役割メタデータ	44
3.4	制作ルールを利用した役割抽出	48
3.4.1	顔情報のみを利用した「主張・訴え」の役割抽出	48
3.4.2	ナレーションのタイミングを利用した「導入・見せる」役割抽出	51
3.5	まとめ	54
第4章	クイズコンテンツの自動生成	55
4.1	はじめに	55
4.2	クイズ生成について	56
4.3	従来研究との比較	59
4.4	提案手法	61
4.4.1	クイズに適した画像の選択	61
4.4.2	選択肢の生成	63
4.5	実験結果	66
4.5.1	クイズに適した画像選択の結果	66
4.5.2	選択肢生成の結果	67
4.5.3	生成されたクイズの評価	68
4.6	考察と評価	76
4.7	まとめ	77
第5章	インタラクティブ・サマリの生成	78
5.1	はじめに	78
5.2	サマリ生成について	78
5.3	提案手法	79
5.3.1	ニュース項目の代表静止画の抽出	80
5.3.2	ニュース項目の代表文とキーワードの抽出	80
5.3.3	3種類のインタラクティブサマリの生成手法	81

5.4	実験結果と考察	86
5.4.1	ニュース項目の代表静止画・代表文の抽出結果の評価	87
5.4.2	3種類のインタラクティブサマリの生成結果と評価	90
5.5	まとめ	103
第6章	結論	104
	謝辞	107
付録A	NewsMLのサンプルデータ	109
付録B	EvenTankデータモデル	112
付録C	EvenTankデータ操作関数	126
付録D	EvenTankモジュール制御関数	147
付録E	EvenTankモジュール情報	150
	参考文献	159
	研究業績	168

目次

1.1	放送局業務の多様化	2
1.2	ニュース番組を対象とした研究	4
1.3	サマリ生成手法の研究	7
1.4	質問応答システムの研究	8
1.5	工夫を凝らしたユーザインターフェイス	9
1.6	統合検索機能	10
2.1	EvenTank のイメージ	15
2.2	EvenTank のユースケースシナリオ	16
2.3	NewsML のデータ構造	22
2.4	MPF のアーキテクチャ	24
2.5	EvenTank における 2 つの基本データモデル	28
2.6	SourceItem の基本構造	29
2.7	Key-Value による情報記述	30
2.8	Topic の基本構造	31
2.9	EvenTank のアーキテクチャ	35
3.1	ニュース番組の役割	42
3.2	ニュース番組の制作	42
3.3	ショットの役割分類	45
3.4	「主張・訴え」を伝えるショットの抽出手法	48
3.5	「導入・見せる」役割をもつショットの例	52
3.6	「場所・導入/見せる・様子」を伝えるショットの抽出手法	53
4.1	画像付きクイズの例	57
4.2	画像選択クイズの生成手法	62
4.3	スポットライト画像の例	62
4.4	スポットライト画像の選択手法	62
4.5	無地度の例	63
4.6	画像を説明している字幕文の選択手法	64

4.7	似て非なる選択肢の生成手法	65
4.8	似て非なる文の選択手法	66
4.9	主被写体が具象物の場合の失敗例	70
4.10	生成された人物に関するクイズの例	73
4.11	生成された具象物に関するクイズの例	74
4.12	画像と選択肢の明確な対応はできないがクイズと見なした例	75
5.1	インタラクティブ・サマリの生成手法	80
5.2	代表静止画群の抽出手法	81
5.3	代表文の抽出手法	82
5.4	伝えたいことサマリの生成	82
5.5	「伝えたいこと文字列」の抽出手法	83
5.6	人物サマリの生成	85
5.7	時系列サマリの生成	86
5.8	ニュース項目の代表画像と代表文を選択するツール	88
5.9	北朝鮮拉致の伝えたいことサマリ (デフォルト表示)	90
5.10	北朝鮮拉致の人物サマリ (デフォルト表示)	91
5.11	北朝鮮拉致の時系列サマリ ((デフォルト表示)	91
5.12	北朝鮮拉致の伝えたいことサマリ (全展開後)	93
5.13	北朝鮮拉致の人物サマリ (全展開後)	94
5.14	福岡県の中2男子自殺の伝えたいことサマリ (全展開後)	96
5.15	福岡県の中2男子自殺の人物サマリ (全展開後)	96
5.16	福岡県の中2男子自殺の時系列サマリ (全展開後)	96
5.17	安倍総理の日中関係に関する活動の伝えたいことサマリ (全展開後)	97
5.18	安倍総理の日中関係に関する活動の人物サマリ (全展開後)	98
5.19	安倍総理の日中関係に関する活動の時系列サマリ (全展開後)	98
5.20	項目を並べ各映像を視聴できるシステム (チャプタ画面)	99

表 目 次

2.1	Dublin Core Metadata Element Set	19
2.2	MPEG-7 の高次特徴メタデータ	20
2.3	IPTC での規格化項目	21
2.4	EvenTank のシステムモデル構成要素の概要	25
2.5	高レベルメタデータ操作インタフェース	33
2.6	モジュール制御インタフェースの関数	34
3.1	ニュース映像の制作過程による分類	45
3.2	正面顔情報のみによる主張・訴えショットの抽出	49
3.3	サイレントショットによる導入・見せるショットの抽出	53
4.1	ニュース番組を対象とした QA システムと要約システムとクイズ生成タスクの比較	60
4.2	正解文選択の評価	68
4.3	似て非なる選択肢生成の評価	68
4.4	生成したクイズの有効性の評価	69
4.5	一般人による生成したクイズの有効性の評価	70
5.1	ニュース項目のイントロ文の文末解析	84
5.2	2006 年 10 月のニュースから作成したトピック	87
5.3	自動抽出した代表静止画・代表文の評価	89
5.4	インタラクティブ・サマリとチャプタ画面の評価結果 (11 人の平均)	100

第1章 序論

1.1 背景

近年、さまざまな分野の技術が加速度的に進歩したことで、放送と通信/インターネットという一昔前まで親和性が低かった分野においても、双方の利点を併せ持ったサービスが次々と生まれてきている。これは、いわゆる放送と通信の融合といわれる変化であり、端的な例としては、テレビに通信機能を持たせ、番組に参加したり、Webのコンテンツと連携したり、逆に、パソコンにテレビ番組を録画して見たい時にノンリニア的に視聴することなどが一般的になってきている。また、これらに加え、携帯等モバイル端末にも同様な機能が搭載され、それに向けたサービスが次々と始まっている。

これらの変化を放送局の視点でみると、単にアナログがデジタルに変わっただけではすまされない、大きな変化を伴っている。それはコンテンツを制作する体系部分である(図 1.1 参照)。一昔前までの作業は、番組を企画し、取材、編集を行い、完成した番組を放送して、1サイクルという基本サイクルがあった。もちろん、この基本的な流れは今でも変わりはないが、その途中の工程に様々な副次的作業が追加され、出力が1つではなくなっている。例えばそれは、番組に関連するホームページであったり、番組連動型のデータ放送コンテンツであったり、また、通信を利用した双方向の視聴者参加型番組への仕掛けや、携帯向けのワンセグ放送、番組放送後のアーカイブ化やDVD化など多岐にわたっている。

これら従来の番組以外のコンテンツの制作については、ほとんどがそのための専用のシステムを備えているわけではなく、人を雇うか、外部に発注するなどを行っており、多くのコストがかかっている。この部分をシステム化しない理由は大きく2つ考えられる。1つは創造性を必要とする作業を含むため、現在の技術では完全自動化は困難であること。もう1つは、これら番組に付随するコンテンツは、視聴者からの反応により、その体系や内容が流動的であり、時間とお金をかけてシステム化したとしても、すぐ改修せざるを得ないなど、柔軟性の高い人間が行ったほうが結果的にコストパフォーマンスが良いかもしれないという考えがある。しかし、このように人手による柔軟性や正確性が要求される一方で、経営的にはコスト削減が必要であり、ワンソースマルチユースなどのコンセプトが

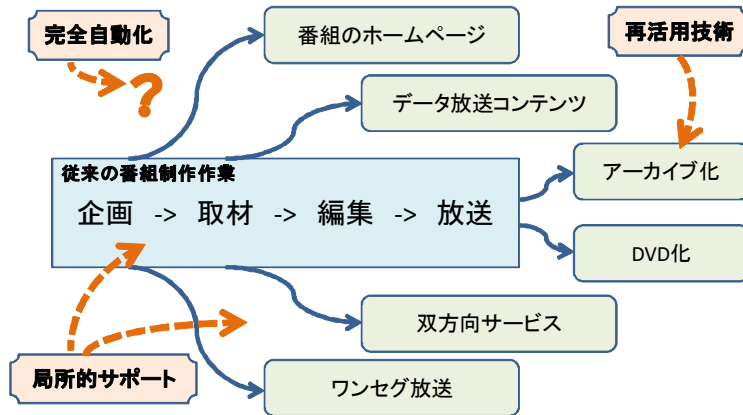


図 1.1: 放送局業務の多様化

奨励されているのが現実である。

このような流れの中、技術はどのような役割を果たすべきかを考えてみる。まず、放送局の主軸である、人間が英知を結集して制作する番組については、この制作自体を完全に自動化することは非現実的である。しかし、番組制作を進めていく中での、個々のサブタスクに対しては、その作業を支援する諸技術を提供することはできる。例えば、制作者のための素材検索に関する技術や、バーチャルスタジオなど実写とCGを合成した演出手法などである。これらは、番組を含めた既存のコンテンツ制作を、サポートする技術として位置づけることができる。

一方、これまでにない、新しいサービスへの足がかりを示すことも、技術の役割として重要である。放送局にとっては、少ないリソースの中で、定着するかどうか先行きが不透明な新しいコンテンツの制作は、なかなか踏み出せないのが現状である。そこで、半自動でもよいので、新しいコンテンツの制作手法を示し、そのサービスイメージまで示唆することは、経営にとっても今後の方針を決める上での議論の土台を得ることに相当し、重要な役割を担っている。このような位置づけとしては、放送後の番組コンテンツなどを再利用する技術が挙げられる。最近では、放送局のアーカイブとは規模は違うものの、一般の家庭においても、大容量のハードディスクに多くの番組を蓄積できるようになってきている。このような環境では、蓄積された放送済み番組のどの部分にも瞬時にアクセスが可能であり、これらの特質を活かした新しいサービスの展開にも大きな期待が高まっているのが現状である。

本研究では、後者の立場に立って、蓄積された放送済みニュース番組を、再利用するためのフレームワークとして「EvenTank」(イベタンク)を提唱し、その構造化のために映像の役割に注目した解析手法を提案し、いくつかのサービス生成

実験を行い、その特徴及び有効性を明らかにする。

1.2 ニュース番組再利用の現況と従来研究

ニュース番組の再活用について考えてみる。一般的には、ニュースはある出来事についての最新情報を提供するという性格上、視聴者にとって一度見てしまうと、それを再利用するということはあまりない。個人的な興味で、特定の話題等を録画保存し、ライブラリを作成している人はいるかもしれないが、非常に限られた特殊なケースといえる。また、放送局におけるニュース番組の再活用について考えると、こちらは個人的な場合に比べ、かなり頻繁に行われている。それは、ある出来事を報道する際に、背景をわかりやすく説明するため、過去の映像を再利用して、これまでの経緯を振り返ったり、また、関連する映像が得られない場合などは資料映像として流用することが多いのである。しかしながら、これ以外の主だった再利用というものは、現時点ではみあたらない。

一方、学術的な分野に注目すると、ニュース番組を利用した様々な研究報告がされている。その範囲については、非常に多岐に渡るが、大きくサービスとして分類すると、図 1.2 に示す 3 つの機能に分けることができる。1 つ目はニュースの要約を提供する「サマリ機能」、2 つ目は質問に答えてくれる「質問応答 (QA: Question Answering) 機能」、3 つ目は必要なコンテンツを探すための検索機能である。従来研究には、これら複数のサービスを統合的に組み合わせて提供する、ライブラリ (アーカイブ) システムとしての報告も多い。

これらニュース番組を対象とした研究を振り返ると、かなり古くから行われている。ただし、当初は、テレビのコンテンツではなく、新聞等のニュース記事が解析対象となっていた。ここ十数年になって、コンピュータの処理能力や蓄積装置等の飛躍的な発展により、比較的容易に映像を扱うことができるようになったため、映像を蓄積したアーカイブシステムとして、多種多様なブラウジング機能の提案や、そのためのインデックス自動生成についての報告等が盛んになった。サマリや質問応答については、基本的には意味をなしているかどうかが一番重要であり、言語処理の研究としての色合いがまだ強いのが現状である。しかしながら、映像を扱うサマリやQAシステムの報告も数は少ないがいくつか報告されているので、これら従来研究の中で、ニュース番組の映像も扱った報告について、カテゴリ別に概観する。

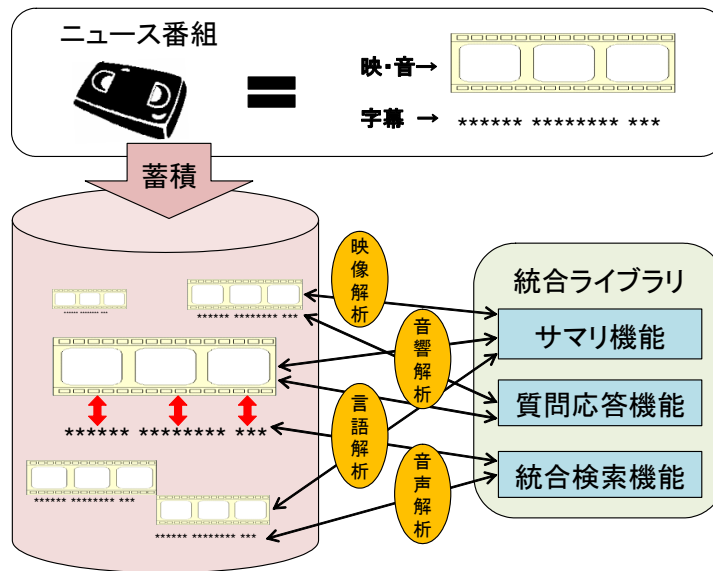


図 1.2: ニュース番組を対象とした研究

1.2.1 サマリ機能

ニュースを対象としたサマリ生成は，その対象とするデータに注目すると，大きく2つに分類することができる．1つは，通常のニュース番組など，話題の異なるニュース項目を含んでいるものからサマリを生成するもの，もう1つは，ある1つの話題に関する多くの情報からそのサマリを作成するものである．一般的には，どちらのサマリを生成するにしても，その生成手法のエッセンスは，以下の二段階で構成される．

- 重要部分の抽出/不要部分の削除
- サマリ生成

はじめに，情報の絞り込みを行い，第2ステップでそれを最終的なサマリに編集するのである．情報の絞り込みを行う場合には，重要な部分を抽出していく方法と，不要な部分を削除していく方法があり，両者を組み合わせた手法もある．ここにおける核心の問題は，重要/不要の定義である．従来研究で報告された簡単な手法としては，ニュース番組のアナウンサショットだけを抽出してつなげる手法 [1] や，ショット長が長いものだけを集める手法 [2] がある．前者は，アナウンサショットが，各ニュース項目の冒頭に必ずあり，そこは内容を簡単に述べた導入部分に相当するため，それらを抽出することでサマリを生成する手法である．ただし，アナウンサショットは，各ニュース項目の冒頭以外にも出現したり，逆に冒頭であって

も出現しない演出方法をとるニュース番組もあるため、万能とは言えないが、サマリとしてある程度機能すると言える。また、生成されるサマリは、抽出したアナウンサショットをつなぎ合わせるだけであるので、対象とした番組によってサマリの長さが異なることとなり、ユーザが指定した時間(例えば5分)のサマリを作成することはできない。一方、後者の手法では、ニュースの重要な部分はアナウンサショットとインタビューショットであるという仮定をし、これらが他のショットに比べ比較的長さが長いという統計的な観測を基にしたものである。この手法では、最終的なサマリを5分と決めると、その累計が5分となるまで長いショットから順に選択し、それらを番組内に出現した順に並べるものであるため、ユーザが希望する長さのサマリが生成可能である。しかし、この場合は、最終的なサマリの長さを調節できる代わりに、続くショット間のつながりが不明確なサマリが生成される傾向にあり、サマリとしては問題が残る。

これまでの例は、重要部分を抽出する手法であったが、次に逆方向のアプローチである不要部分を削除する手法の例を挙げる。各ショットから抽出した代表静止画の色レイアウト特徴を利用し、類似する静止画をまとめる、すなわち重複部分を不要部分とみなし削除するというサマリ生成手法[3]である。よく行われる手法としては、抽出した静止画をある画像特徴でクラスタリングする方法であるが、この論文では同様のことを、行列の特異値分解を利用して行っている点が主張点となっている。彼らは、同様なショットを同じグループに分類できたことで、サマリを行ったとしているが、実際のサービスコンテンツとしては更なる加工が必要である。

次に、より意味内容を考慮したサマリ生成手法について考察する。一般的に、現在の技術を使っても、コンピュータで意味を扱うことは、まだかなり難しいといわれている。多くの研究報告は、字幕データ(クローズドキャプション)や音声認識結果などのテキストデータを利用し、言語の統計的なパターンなどを利用する、いわゆる表層的な解析でサマリを生成している。その手法についてみると、ある程度メタデータとして人間が情報を付与したものを出発点として、そこから自動でサマリ生成を行うものと、最初から最後まで自動で行う手法に分類することができる。

前者は、現在人手により付与しているデータはいずれ自動的に付与されることを想定し、その先の手法についての研究報告である。例えば[4]は、発話単位をショットとし、ショット間の関係を人手で入力する。付与する関係は、因果、後続、要素、具現化、状況、換言、背景の7つであり、これら付与された関係から各ショットの重要度を計算し、また各ショット間の距離により理解のしやすさを計算し、最終的

なサマ리를自動生成する手法である．また，同じ発話文を用いて，形態素レベルに詳細なデータを付与し，これを出発点とする報告もある [5, 6]．これらの報告では，MAML[7]により動画の転記情報を記述し，GDA[8]により文法機能，修辞関係や照応関係を記述し，発話テキストの中の不要部分を削除しつつ，重要部分を選択する手法である．この手法では，抽出した重要部分から最終的なサマ리를生成する手法についても新規性を有している．それは，抽出した重要部分を音声合成で出力し，その際の映像については，対応するショットから，抽出した発話部分の重要度により長さを決定してサマリビデオを生成する点である．従来単につながり合わせるだけであったサマリでは，しばしば音声の不連続が問題となり制約が大きかったが，この手法により，音声はとぎれることなく，また，指定した時間でサマリを作成できる．

最後に，人手によるメタデータ付与を想定しないサマリ生成手法について述べる．基本的には字幕データや音声認識結果を解析し，必要なメタデータを抽出しつつ，区間毎の重要度や類似度を計算してサマ리를生成する．例えば [9] では，キーワード間の意味的関連までを含む概念グラフを生成し，それを用いて記事間の類似度計算を行うことで，単なるキーワードを用いるより精度良く類似記事を収集可能としている．そして，収集した記事の中での記事の重要度を，その重複回数や放送時間の長さ，また番組の冒頭で放送されるほど重要といった，いくつかのルールに基づいてサマ리를生成する手法である．また，[10] では，キーフレームやロゴなど映像の特徴をつかって，繰り返し放送されているニュース項目を自動的に抽出し，そのニュース内容を表現するアイテム (代表的な画像とキーワードの組み合わせ) も自動抽出してサマ리를生成する．

これまで，様々なサマリ手法について述べてきたが，これらを図 1.3 にまとめる．左端は，冒頭でも述べたサマリ生成の基本手法である．図中央は，その処理段階によって扱うデータの単位を整理したもので，ショット，項目 (一般的にトピックとも呼ばれる)，スレッド (つながりを持った項目群) に分けることができる．従来研究に照らし合わせて説明すると，重要部分や不要部分は，システムによってはショット単位で行われたり，項目単位で行われ，出力されるサマリは，各項目内でのサマリであることもあれば，スレッドのサマリの場合もある．一番右端はそれぞれ必要とされる主な技術を示しており，それぞれの単位に，映像解析を利用して，もしくは字幕など言語処理を用いて分割し，ある尺度をもってその重要さを判断し，あるいは類似を判定することで重複する不要部分を判断して，最終的なサマ리를生成する．最後に，サマリにおける映像の選択方法 (映像と発話テキストの関係) に注目する．一般に，ニュースにおける映像は，その時の発話内容と何

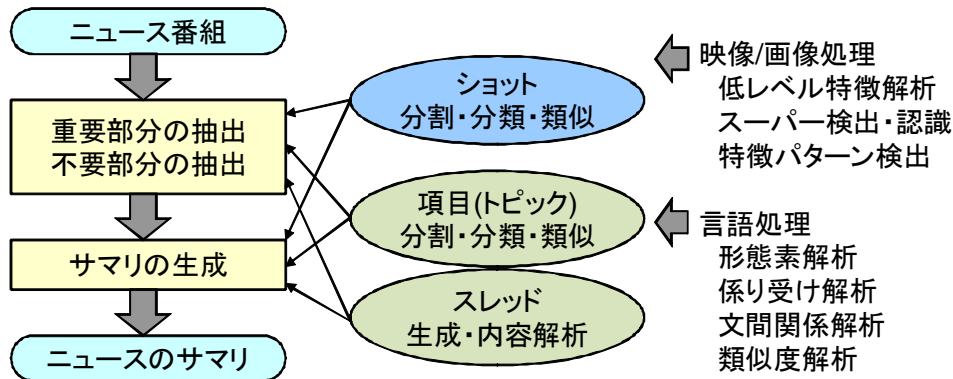


図 1.3: サマリ生成手法の研究

らかの関係があることを利用し，時刻で同期している部分を，そのまま対応している映像として扱っていることが多い。

1.2.2 質問応答機能

ニュースを対象とした質問応答システムに関する研究は，米国の国立標準技術研究所 (NIST: National Institute of Standards Technology) が主催する TREC (Text REtrieval Conference) [11] や，国内では国立情報学研究所 (NII: National Institute of Informatics) が中心となって行っている NTCIR (NII-Test Collection for IR Systems) プロジェクト [12] などがあり，1つの代表的な分野を形成しているが，それらは基本的にテキストベースである．最近になって，映像も扱うようになったシステムも数は少ないが報告されている．例えば，[13] では，音声認識結果を用いて，従来のテキストベースのアプローチで質問に対する解答文を3文選択し，各文に最も適切なショットを割り当てて，ビデオ形式で解答を生成するものである．この時の字幕と映像の関係は，基本的には時刻において同期していたものであるが，更に，アナウンサのショットは除き，質問されているジャンルによって，あらかじめ13種類に分類したショットタイプの中からふさわしいと定義したショットを選択している．また，[14] では，提案システムの一部にQA機能を実装しており，コンテキストベースとビジュアルベースの2種類の質問を扱い，それぞれ解答として適切なショットを表示するものである．この論文での，コンテキストベースの質問とは「アラファト議長は何歳で亡くなったのか？」や「いつ亡くなったのか」といったもので，ビジュアルベースの質問とは「アラファト議長の顔が写っているショット」や「反米デモのシーン」といった所望の映像についての記述である．コンテキストベースについては，その答えをナレーションに含むショットをそのまま

出すという手法をとっており、ビデオを扱う QA システムのほとんどは、この手法にとどまっている。ビジュアルベースの質問については、質問応答というより、検索に分類されるかもしれないが、提示された映像に関する記述を理解しないと、該当する映像を探すこともできないため、彼らは質問として扱っている。その手法は、関連するトピックを検出したあとに、その中のどのショットを解答として選択すべきかを、画像から抽出した HLF (High Level Feature) の情報を利用して決定している。HLF には、屋外、海、空、車、顔などたくさんの高次特徴(コンセプトと表現されることもある)が定義されており、それぞれの高次特徴が画像中に写っているか否かの情報を利用するものである。簡単な例を挙げれば、人の写真を要求された時は、顔コンセプトが存在すると判断された画像に絞り込むということをするのである。[15] なども同様に HLF を利用して、目的のショットを決定している例である。

ここまで、ビデオを扱う質問応答システムの従来研究について説明してきたが、図 1.4 に、その典型的な質問応答の解析手法とそのデータ単位、主要要素技術についてまとめた。手法としては、質問に対する解答文を選択するところまではテキストベースで処理が行われ、その後にショットとの対応をとっている。映像のような多義性のメディアから意味を抽出する場合は、ある条件がなければ人間でも統一した意味を抽出するのは難しいため、この手法は自然なものといえる。質問から解答文取得までの典型的な手順をまとめると、はじめに、テキスト解析を行い、質問のタイプについて解析し、何を解答すれば良いのかを決定する。次に、質問文の中から解答をさがすためのキーワードなどを抽出し、そのキーワードを基に蓄積されたデータから関連する情報を抽出する。最後に、そこに含まれるキーワードの関係や質問タイプから最終的な解答候補を絞りだす。更に、ビデオを扱う QA システムでは、この得られた解答文に対して、ふさわしい映像を探すのであるが、通常はその解答文が述べられているショットをそのまま最終出力とするも

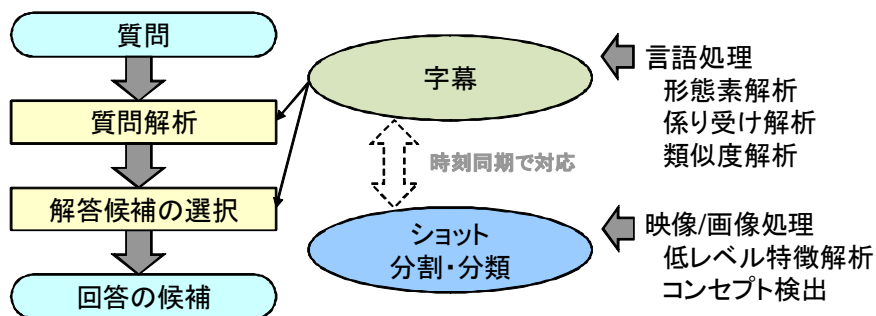


図 1.4: 質問応答システムの研究

のがほとんどである．これら一連の処理の中で扱うデータに注目すると，図 1.4 の中央に示すように，テキスト解析では字幕の1文1文が，また，映像についてはショットを単位とすることが多いようである．そしてテキストとショット間のいわゆるセマンティックギャップを，最近ではHLFなどによるショットのある種タイプ情報を利用して対応付けを試みている．

1.2.3 統合検索機能

最後に，検索機能を含めた統合システムとしての研究報告について述べる．ニュース番組を蓄積したライブラリ(アーカイブ)システムとして有名なものには，Infomedia プロジェクト [16] やコロンビア大学のデジタルニュースシステム [17] を挙げる事ができる．Infomedia プロジェクトは，1994年から5年間のプロジェクトとして始まり，デジタルライブラリの自動構築を目指したものである．自動解析処理としては，音声認識によるトランスクリプト生成，言語処理によるキーワード抽出，画像処理としては，カラーヒストグラムやDCT係数，カメラモーション，テクスチャ特徴量などを用いてショット境界の検出や，顔検出など，様々な情報を抽出している．そして，抽出した情報を統合的に利用することで，ユーザのリクエストに対して，コンテンツをスキミングして表示するような機能を実装した，デジタルライブラリの研究の草分け的存在である．現在は，第二期のプロジェクトに遷っており，複数のビデオ素材からの自動要約生成やデータの視覚化手法について研究を続けている．同様に，コロンビア大学のシステムも，様々な



(a)MediaMillのインターフェイス

(b)Collagesのインターフェイス

図 1.5: 工夫を凝らしたユーザインターフェイス

メディア処理を組み合わせることで、ユーザがある興味のあるトピックを検索して追跡したり、その要約を得たり、類似した画像を取得するなど総合的なシステムの実装を行った。これらにみる大規模なシステムは、多くの研究者により、それぞれの問題に対する研究成果を統合接続して実現している。規模はちがうものの、その検索手法やデータ表示手法に工夫を凝らしたシステムも多く報告がある。

例えば、MediaMill[18] は、491 のコンセプトを自動的に抽出し、これをもとに視覚的な類似、意味的な類似、意味のつながりを考慮し、ユーザの目的にあったブラウジングを実装したものである(図 1.5(a))。また、Video Collages[19] は、蓄積された様々なニュース項目を、画像とテキストを用いてユーザにわかりやすく、インタラクティブに提示するシステムである。あるトピックについてのサマリ生成や、関連項目へのリンクなど、ユーザの興味が向くままにニュースコンテンツを利用できるように設計されている。図 1.5(b) は、Video Collages の地図とニュース画像を用いて、直感的にどの場所で何が起きたかを表示するインターフェースで、ズーム操作によりその粒度での表示すべき画像を替えて見やすさ等を確保している。地図の下には、関連するキーワードが、画像同様にその粒度に応じて表示される。また、ニュースの場合は、あるニュースがどのように変化していったのかを閲覧する要望も高く、これらは TDT (Topic Detection and Tracking) といわれる 1 つの分野を形成している。このデータの閲覧に特化したシステムの報告もたくさんある [20]。

このようにライブラリの基本機能である検索・閲覧については、どのようなインデックスをどのように生成して、どのように構造化しておき、それを利用してどのように見せるか、様々な研究要素がある。もちろん、生成したインデックスから一致するデータを、できるだけ素早く正確に見つけ出すための検索アルゴリズムも重要である。図 1.6 は、関連する要素技術について簡単にまとめたものである。一般的に、ユーザの検索要求が与えられると、それを検索のためのキー(イン

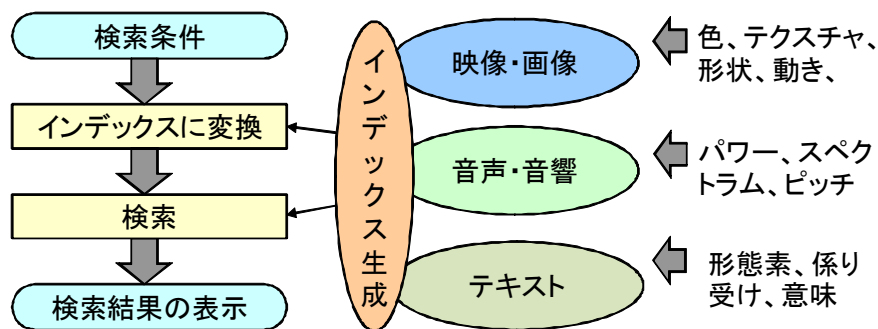


図 1.6: 統合検索機能

デックス)に置き換え、大量のデータの中から該当する部分を抽出し、ユーザに提示するというタスクである。この際に利用するインデックスとして何が適切であるかが問題であり、用途によって様々な特徴が利用されるのである。例として、ある夕日の映像を検索したい場合に、そのサンプルとなる夕日の画像をシステムに与えて、これと類似した画像を含む映像区間を探すという手法が考えられる。与えられた画像から位置を考慮した色特徴を抽出し(例えば、上が赤っぽく、下は黒いという特徴)、色特徴が同じような映像区間を抽出することが考えられる。この例では、色と位置の情報がインデックスとなったが、先にも述べたように、実際には様々なインデックスが利用されており、これらを生成するために映像、音声、テキストなど様々な情報処理が組み合わせて用いられている。

1.3 技術的課題

前節で説明したように、ニュース番組を対象とした研究は数多く行われており、これはニュース番組が様々な情報を含んでおり、その再利用価値がそれだけ高いことを示唆しているといえる。更に、この分野の研究開発を進め、実用化していくには、次の2点を解決する必要がある。

1.3.1 ニュースコンテンツを扱う共通基盤の整備

従来研究をみると、ニュースという多くの情報を含んだコンテンツを解析するためには、様々なメディア処理を統合的に扱う必要がある。番組そのものを形成する映像、音声だけでなく、その内容を理解するためには、人間と同じように知識を蓄積活用することなど、あらゆる方面における技術を組み合わせる必要がある。しかし、一研究者、もしくは一組織にとっては、各分野での最新の技術を全て保有することは難しい。実際には、ある特定の分野には詳しいが、それ以外についてはあまり得意ではないことが多い。また、実際の研究成果を実用化する際にも、ピンポイントの技術だけでのビジネスは難しく、それをシステム化するなど別のノウハウが必要になることが多い。

そこで、このような種々の問題を解決するために、オープン仕様のニュース解析用の共通基盤が必要と考える。従来研究を見てもわかるように、質問応答、サマリ生成、検索のためのインデックス生成と、目的とするサービスは異なるものの、そのための映像、音声、テキストなどの基礎的な解析技術などは共通するものも多い。また、解析対象とする単位にも共通する部分があるため、多くの研究者が互いに協調しながら効率よく研究を進めるための共通の切り口が存在すると

考える．これらを，ニュース番組の再利用のための共通基盤としてある程度統一し，これに沿った形で皆が協調できれば，この分野における研究開発を加速して進めることができると考える．

1.3.2 再利用のための映像とテキストの関係抽出

視覚情報は，人間が生活する上で，最も重要な情報元となっている．ニュース番組における映像も，視聴者の内容理解を更に促進するなど，重要な役割をもっている．理想的には，ニュース番組の再利用の際には，これら映像も再利用できることが望ましい．しかしながら，多くの従来研究においては，字幕などのテキストに同期した部分をそのまま用いるなどの利用方法にとどまっている．これはニュースにおけるサマリや質問応答では，その意味が大事であるため，出力されるテキスト部分に注目が集まり，映像についてはあまり重要視されていないためである．ニュース番組において，伝えられているテキストと映像は，ある程度関係があることは事実であるが，その関係が明確でないものも少なくない．この映像とテキストの対応付け問題は，いわゆるセマンティックギャップといわれ，難しい問題である．先に述べたように HLF などを用いて，徐々にそのギャップを狭める方向に来ているのであるが，今後，ニュース番組を再利用して，さらに別のサービス，例えば時事を学ぶ教育用コンテンツなどへの自動変換などを考えると，映像の意味についての情報を少しでも正確に引き出し，構造化して蓄積する必要があると考える．

1.4 本論文の構成と概要

次章以降の本論文の構成とその概要について以下に示す

2章 EvenTank の提案

ニュース番組再利用のための研究開発および応用のための共通基盤として EvenTank(イベントンク)を提案する．はじめに，従来研究から得られる要求条件について検討し，EvenTank の設計指針について説明する．次に，データ構造とその表現について，既存の関連フォーマットと比較を行いながら説明する．最後に，従来研究のそれぞれが，EvenTank においてどの位置にあるかについて説明し，設計指針の有効性を確認しつつ，その将来への拡張性と柔軟性についても確認する．

3章 ニュース番組における映像のメタデータ

ニュース番組の映像再利用のためのメタデータとして、映像の役割に注目し、番組の制作ルールを利用した抽出手法を提案する。はじめに、従来おこなわれてきた映像に対するメタデータ付与について整理をする。次に、ニュース番組そのものの役割と、その中での映像の役割について、番組制作手法からの観点を考慮した議論を行う。この仮定に基づき、映像の役割を考慮したメタデータ生成について提案し、実証実験を行いその有効性を確認する。

4章 クイズコンテンツの自動生成

第2章で述べた EvenTank の枠組みを利用し、ニュース番組の新しい活用の方向として、クイズの生成実験を行う。第3章で述べた映像の役割としては、「主被写体」を伝えるショットに注目し、画像付き選択クイズを自動生成する。はじめに、クイズ生成というタスクを、3つのタスクに分割し、それぞれに対する工学的アプローチを提案する。実際のニュース番組を用いて生成実験を行い、各サブタスクへのアプローチと、それを総合したアプローチについて評価を行い、その有効性について検証する。

5章 インタラクティブ・サマリの生成

第2章で述べた EvenTank の枠組みを利用し、ニュース番組の新しい活用の方向として、インタラクティブ・サマリの生成実験を行う。第3章で述べた映像の役割の中で、主張・訴えを伝えるモノローグショット、導入・見せる役割をもつサイレントショット、前章でも用いた主被写体が明確であるスポットライト画像に注目する。各項目毎に、その項目を代表する画像群と文やキーワードを抽出し、それらを用いてインタラクティブ・サマリを自動生成する手法を提案する。実際のニュース番組を用いて生成実験を行い、代表画像や文の選択手法の評価をはじめ、インタラクティブ・サマリそのものへの主観的評価を行い、その有用性について確認する。

6章 結論

最後に本論文で述べた研究成果についてまとめる。

第2章 EvenTankの提案

2.1 はじめに

本章では、ニュース番組を効率よく再利用するための共通基盤として「EvenTank」(イベントンク)を提案する。EvenTankとは、一言で表現すれば、世の中の出来事(Event)についての情報が、映像付きで整理/蓄積されているデータベース(Tank)を中心とし、様々な時事に関するサービスを効率よく研究/開発、および提供するためのルール群から構成される一種のフレームワークである。本章ではまず、EvenTankの概略について述べた後、EvenTankの要求条件について整理する。次に、これら要求条件を満たすよう設計したフレームワークを提案する。最後に従来の研究それぞれも、本フレームワーク上で動作し、更に今後への拡張性/柔軟性を有していることを確認する。

2.2 EvenTankとは

EvenTankは、ニュース番組を効率よく再利用するための共通基盤、いわゆるフレームワークの名称である。具体的には、世の中の出来事についての情報が、映像付きで整理/蓄積されているデータベースを中心とし、データの入出力や加工処理について、大枠のルールを共通化し、複数の研究開発機関などが互いに協調し効率よくニュース番組の解析/構造化を進められるようにしたものである。図 2.1 にそのイメージを示すが、核となるデータベース(Eventが蓄積されているTank)は、ある種の映像付き情報辞典ということもできる。

このような映像付き情報辞典のような存在は、次の2つの点において、今後ますます重要になってくる。1つは、情報が整理されており、必要な部分だけを利用することができるという点である。近年、急速に普及したインターネットが、多くの情報を提供してくれる一方で、「情報の洪水」とも表現されるように、その量は人間の処理範囲を大幅に超え、所望する情報にたどり着くことが難しくなった。この問題解決のためには、これまでさまざまな検索の手法が研究開発され、実用化されている。このような効率よい検索手法を提供することも1つの解決方法であるが、EvenTankで提案するように、あらかじめデータを整理する枠組みを提

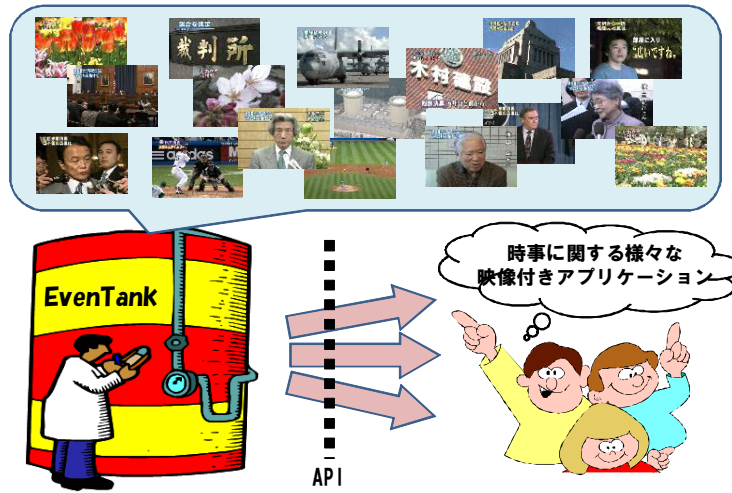


図 2.1: EvenTank のイメージ

供しておくことも、有効な解決手段と考える。もう1つは、映像付きの情報であるという点である。現在でもなお、検索と言えば、テキストベースの処理が主流である。意味を扱うという観点からみると、処理がテキストベースであることはやむを得ないところであるが、その検索結果については、百聞は一見に如かずという諺にもあるように、ある物事についての説明が、テキストだけでなく映像も付随していれば、その物事に対する理解度はより深まるはずである。最近になって、Google[21] や Yahoo[22] などから、Web 上の画像や映像の検索が提供されるようになったが、今のところ、抽出された映像/画像を構造化して、更にこれらを利用するという枠組みはない。また実際のところ、インターネット上の情報は、その内容が多岐にわたり、表現方法も千差万別であるため、自動的な構造化は難しい。そこで、本論文では、対象をニュース番組に絞り、番組のパターンなどを利用することで、映像付き時事情報の自動構造化を試みる。また、研究/開発の効率性、実用化についても考慮したフレームワークを提案する。

次に、EvenTank の利用シナリオについて述べる。図 2.2 は、想定するシナリオを示している。EvenTank の目的は、ニュース番組を蓄積して効率よく再利用するための共通基盤の提供である。従って、放送局がその主たる場所であるが、図にも示すように、一般の家庭においても、DVD/HDD レコーダ内で EvenTank の小規模版を実現できると考える。

まず、放送局でのシナリオを考えると、大きく3つの場合が考えられる。1つ目は、番組制作者がアーカイブとして利用するケースで、素材の検索はもちろんのこと、あるトピックに関するサマリの自動生成や、逆にある質問を投げかける

と、該当する映像があれば、その解答を映像付きで示してくれるような機能があると、大変役に立つ。現状でも、日々放送されたニュースは蓄積されていて、イントラネットを介してアクセスできるが、再利用を想定したデータの構造化はされていない。従って、放送日で検索し、その後は再生ソフトを制御しながら実際の番組を見て探さなければならず、解決する必要がある。2つ目は、新しいコンテンツの研究/開発を行うためのテストベッドとして利用するケースである。ここには、運用している EvenTank の機能アップなども含まれている。3つ目は、局内向けでなく、実際の一般向けサービスのための基盤としての利用である。例えば、映像付き情報辞典としてサービスを行い、ユーザが何か調べものをする際に利用してもらうなどが考えられる。また、局内向けと同様に、ユーザが注目しているある事件に関して、これまでの経緯を、映像付きサマリとして提供することも考えられる。

一方、一般の家庭でもこの EvenTank を構築し、利用する可能性は十分にあると考える。それは、最近普及している DVD/HDD レコーダは、大量の番組を蓄積でき、しかもノンリニアに、容易にコンテンツ中の特定場所にアクセスできるため、家庭内でも同じようなシナリオが考えられる。例えば、1週間の世の中の動きをкаいつまんで見せるだけでなく、クイズ形式で見せるなど、多種多様な活用方法が考えられる。これらは、データの構造とその入出力を共通化することで、そのサービスを生成するプログラムだけを DVD/HDD レコーダに加えるだけで実現できるのである。

基本的には、ニュース番組は日々の生活に役立てるための情報を伝えているので、放送後も再利用する価値が高いコンテンツである。従って、これらニュース番組の再利用は、放送局にとっては、更に効率よいコンテンツ制作が、またユーザにとっては、より便利で豊かな生活を送ることができるようになる、1つの有用な方法と考える。

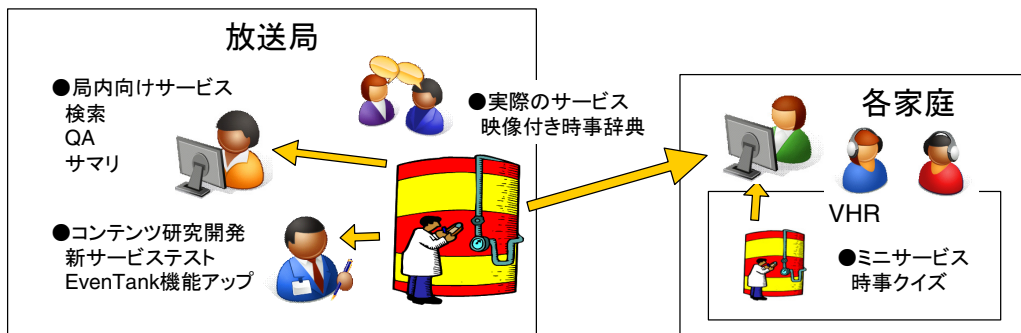


図 2.2: EvenTank のユースケースシナリオ

2.3 要求項目の整理

ここでは EvenTank に求められる仕様について述べる。EvenTank を構築する目的は、ニュース番組再利用のための共通基盤の提供であり、その核となる部分は、再利用のためのメタデータの自動構築とその管理である。技術的課題でも述べたように、ニュース活用の研究シナリオは多種多様であるが、用いられている様々なメディア処理は重複する部分が多い。また、現在は、ビジネス的な戦略等の理由で、個別に独自のシステムを開発することが多いが、今後は、複数の組織/人による協調作業を確保しつつ、新しい視点でのビジネス展開が重要になると考える。以下に、EvenTank に対するこれらの要求項目を挙げ、各々について述べていく。

- 1) 共通したメタデータ形式
- 2) 様々なメディア処理技術を組み合わせ利用することができる
- 3) 将来技術の取り入れやすさ (拡張性, 柔軟性)
- 4) 自動技術と手動技術の容易な置き換え, もしくは協調作業を可能とする
- 5) 誰でも参入できるオープンな仕様
- 6) 独自の技術は秘匿しつつ, 機能だけを提供できるしくみ
- 7) ユーザ, 企業, 大学/研究機関の3者にメリットがあるように

はじめに要求項目 1) について述べる。共通したメタデータは、共通基盤には不可欠なものであり、これまでもたくさんのメタデータが規格化されている。この部分については、次節で詳しく述べることとし、ここでは共通したデータ構造が必要であることだけを述べておく。

要求項目 2) は、様々なメディア処理を組み合わせ利用できるようにすることである。これは従来の研究も示しているように、様々なメディア解析処理を利用して情報抽出/加工を行うため、これら複数の処理の組み合わせを容易に行えることが望ましい。

要求項目 3) は、将来技術の取り入れやすさとして、システムの柔軟性や拡張性の確保である。これまで世の中で開発されたシステムは、時が経つにつれそのパフォーマンスが問題となり、ある時点で更新されるのが一般的である。これは、新しい技術が出現した場合に、その技術を古い技術と差し替えられるようなアーキテクチャになっていないことが原因の1つとして挙げられる。EvenTank では、様々

なメディア解析処理があつてこそ、その存在が優位になるため、次々と研究開発される技術を容易に取り込めることは必須の条件と考える。

要求項目 4) は、自動技術と手動操作の置き換え、もしくは協調作業が可能でなければならないというものである。これは、実際に EvenTank を通してメタデータを生成し、サービス等を始める際には重要な機能となる。例えば、サービス開始当初に、自動処理では難しく、人手によるデータ付与を行っていたものが、1年後には完璧ではないが、ある程度自動で処理できるようになった場合である。自動で生成したものを手動でチェックし、必要があれば修正を行うなど協調作業が必要となるまた、更に技術が進歩し、間違いのない自動処理が開発されれば、完全に置き換えてしまえばよく、このような仕組みは是非取り込むべきである。

要求項目 5) は、誰でも参入できるオープンな仕様であることである。これは文字通り、EvenTank における仕様は完全にオープンとし、皆で共有し、誰でも参加しビジネスにすることができるよう環境であるべきと考える。

要求項目 6) は、独自の技術は秘匿しつつ、機能だけを提供できる仕組みを挙げる。独自の技術はビジネスに直結するものであり、その秘匿性を確保できなければ、誰もこの EvenTank 上で研究開発などは行わないであろう。しかしながら、一方でその技術による「処理機能」は、容易に扱うことが要求される。従って、この2つの要求を同時に解決することは、EvenTank 普及の重要な鍵となる。

要求項目 7) は、ユーザ、企業、大学/研究機関の3者にメリットがあるようなフレームワークにする必要がある。メリットがなければ誰も利用しないことは明白であり、一般に研究フェーズから実用化への閾を低くするなどの工夫が必要と考える。

2.4 既存の関連メタデータとフレームワーク

メタデータの共通化については、誰もがまず考えることであり、これまでも多くのメタデータ規格が作られている。また、メタデータの共通化だけでなく、その運用についてもある一定のワークルーチンを設定したフレームワークもいくつか提案されている。以下に、関連するメタデータ2つと、フレームワーク2つを、従来技術として述べる。

2.4.1 ダブリンコア (Dublin Core Metadata Element)

古くから存在し、現在でも一番よく使われているといわれるメタデータである。この規格は、リソースを記述するために、わずか15項目の属性を規定した単純な

表 2.1: Dublin Core Metadata Element Set

要素名	説明
Title	タイトル
Creator	制作者, 著作者
Subject	主題, キーワード
Description	内容記述
Publisher	発行者, 公開者
Contributor	貢献者
Date	日時
Type	カテゴリ, ジャンル
Format	データフォーマット
Identifier	識別 ID
Source	出展
Language	言語
Relation	関係
Coverage	対象範囲
Rights	権利

ものであるが、書誌情報をはじめ、様々な用途に利用されている。DCMI (Dublin Core Metadata Initiative[23]) によりメンテナンスをされており、現在も拡張を続けている。表 2.1 に基本の 15 項目を示す。

このように様々なコンテンツに対して共通して付与できる項目が規定されている。しかしながら、コンテンツの内容を詳しく構造化して記述することはできない。ニュース番組であれば、まず個々のニュース項目に分割し、更に各項目内をショットに分割し、各ショットの内容、例えば某総理大臣の発話という記述は、含めるとすると「Description:内容記述」の部分に、文章で書くことしかできず、コンピュータでは扱いにくい。

2.4.2 MPEG-7

ダブリンコアに対し、比較的新しく、複雑な構造を記述できるメタデータとして、ISO/IEC[24] がマルチメディアコンテンツの記述のために規格化した、MPEG-7[25] がある (MPEG と名前はつくが、映像音声の圧縮とは全く関係がなく、メタデータの国際規格である)。MPEG-7 は、コンテンツを細かく構造化して表現することができ、その部分部分に詳細な記述を付与することができるようになっている。現

表 2.2: MPEG-7 の高次特徴メタデータ

ツール名	説明
Basic description tools	テキスト注釈などの基本ツール
Linking, identification and localization tools	メディア参照に関する記述
Media description tools	フォーマットなどのメディア記述
Creation and production description tools	タイトル, 制作者などの制作に関する記述
Usage information tools	コンテンツの使用に関する記述
Structure description tools	コンテンツの構造記述
Semantics description tools	コンテンツの意味内容の記述
Navigation and access tools	コンテンツの要約, 派生コンテンツの記述
Content organization tools	複数コンテンツをまとめた集合の記述
User interaction tools	ユーザの嗜好, 動作履歴の記述

在, パート 12 まで規定されているが, メタデータの記述方法を規定する中心となるのは, パート 3~5 である. MPEG-7 では, ある 1 つの特徴を記述するための構造を記述子 (D: Descriptor) として定義し, いくつかの記述子間に構造を持たせて, より大きな記述を実現するために記述スキーム (DS: Description Scheme) を規定している. これら D, DS は規格の中では「記述ツール」と呼ばれており, その構造は, XML Schema[26] をベースに, MPEG が独自拡張した DDL (Description Definition Language)[27] を用いて定義される. 従って, MPEG-7 のメタデータは, XML データ形式をとる.

MPEG-7 の中でのメタデータは, 低次特徴メタデータと, 高次特徴メタデータに整理して規格化されている. 低次特徴メタデータは, 映像や音声の物理的特徴量をさし, 主に自動で抽出できるものであり, 映像/画像用の低次特徴メタデータはパート 3[28] に, 音響音声に関する低次特徴メタデータはパート 4 [29] に規定されている. これら以外のメタデータは, 高次特徴メタデータに分類され, すべてパート 5 [30] に規定されている. 表 2.2 は, パート 5 に規定されている高次特徴メタデータをグループ別に表したものである.

このように MPEG-7 では, 幅広いマルチメディアコンテンツを記述できるよう設計したため, その仕様書は最初のバージョンだけでも約 2000 ページに及ぶ膨大なものとなっている. このため, 完全な MPEG-7 データを扱うシステムの実装は大変であり, いくつかの一般的な用途に限って, ツールのサブセットを作成し, 必要に応じてその使用方法や構造に制約を加えたプロファイル [31] が策定されている. しかしながら, その自由度はまだ高く, データの相互操作性 (Interoperability) の問題等もあり, あまり普及しているとはいえない.

2.4.3 NewsML

NewsML は、ニュース管理のために作られた標準フォーマットである。通信社の国際的な協議機関である IPTC(国際新聞電気通信評議会) [32] が 2000 年 10 月に発表した XML 形式のフォーマットで、テキスト、写真、音声、動画などのマルチメディア素材を管理配信するのに適している。また、属性情報を決められた形で付けることにより、ニュースの分類、権利処理などを容易に扱えるようになっている。現在は、IPTC G2-Standards として、モジュールのように利用できる、規定項目と XML コンポーネントのセットとして開発が進んでいる。2008 年 12 月現在、G2-Standards として、NewsML-G2 と EventsML-G2 がリリースされている。これ以外にも、表 2.3 に示した規格が現在策定されている。

NewsML は、元々ニュース原稿の管理用のために制定され、原稿のバージョン、開示条件、利用可能範囲など、様々な項目から構成されている。最近策定された EvenTank も同様に、あるイベントに対する主催者側の情報はもちろん、参加者に向けた情報など、イベントに関連する情報を構造化して管理できるようになっている。

図 2.3 に、NewsML のデータ構造を表す。また、サンプルデータを付録に掲載する。これらを見比べるとわかりやすいが、大きな記述項目としては、カテゴリ情報をはじめとし、権利情報、そのニュースの管理情報(メディアの種類、制作者、バージョン等)、内容情報(発生場所、制作者、言語、タイトル、そして内容等)、関連する素材(例では、代表静止画とキャプション)などが記述できる。

NewsML では、いろいろなニュース素材を扱えるようになっているが、その内容記述については〈description〉というタグ内にフリーテキストで記述するのが普通である。従って、意味的な記述であり、そのメディアの物理的特徴などを書くには不向きである。また、〈description〉以外の部分は、主に素材交換のための詳細な情報を記述できるよう特化しており、これ以外の用途には大きなオーバーヘッド

表 2.3: IPTC での規格化項目

名前	Version	内容
BNewsML-G2	Ver. 2.2	ニュースアイテムのラッパー
EventsML-G2	Ver. 1.1	イベントに関する情報を記述
SportsML	Ver. 2.0	スポーツに関する情報を記述
NewsML 1	Ver. 1.2	構造化してニュースアイテムを記述
IIM	Ver. 4.1	写真のためのメタデータフォーマット
IPTC 7901	Ver. 5	テキスト伝送のためのフォーマット

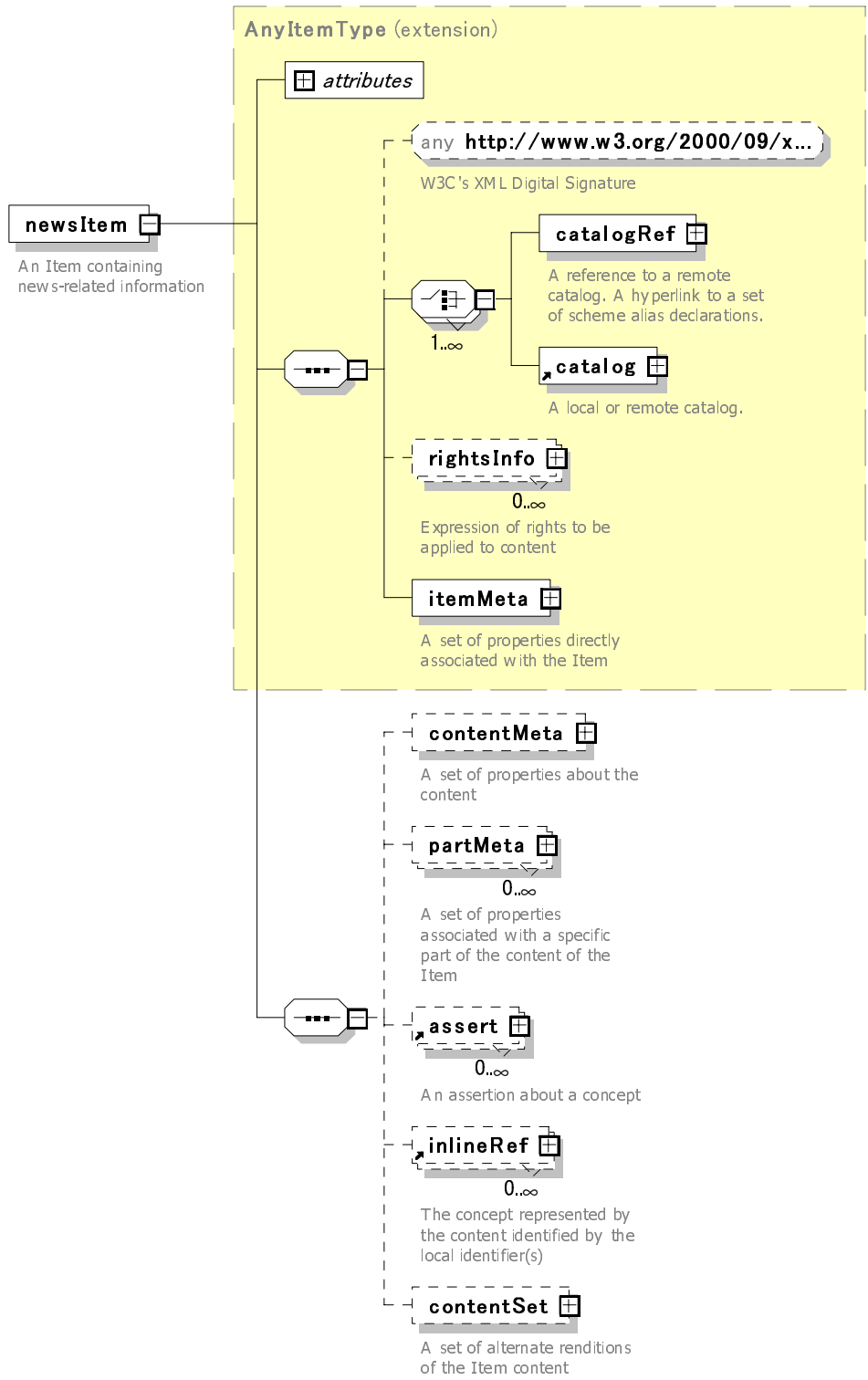


図 2.3: NewsML のデータ構造

となっている。

2.4.4 MPF (Metadata Production Framework)

最後に番組など、ビデオコンテンツにメタデータを付与するための共通基盤としてNHKが提案してきた「メタデータ制作フレームワーク」、通称MPF (Metadata Production Framework)[33, 34]について述べておく。番組制作全体をフレームワークとらえたものとしては、[35]がある。これは番組制作を、9つの正規処理 (Canonical Process) で表現するものであり、各処理はUML (Unified Modeling Language)[36]を用いて定義されている。9つの正規処理は、Premeditate, Create, Annotate, Package, Query, Construct, Organise, Publish, Distributeであり、どの処理も別に定義した9つのステレオタイプにより表現されている。[35]では、番組制作全体のワークフローに重点をおき、各処理におけるシステムの設計や細かい仕様などは全く規定していない。NHKが提案してきたMPFは、この中の「Annotation」部分について、実際の処理を行うための細かい仕様を規定したものと位置づけることができる。現在、このMPFは、ネットワーク対応を終え、新たにVer.2.0として一般に公開されている[37]。このMPFの設計指針は、先に述べたEvenTankへの要求項目とほぼオーバーラップしている。従って、このMPFをニュース番組の再利用のための共通基盤として利用できればよいが、いくつかの点で問題がある。例えば、メタデータの対象が1つのビデオコンテンツ(番組)であるため、ニュース番組の中のある特定の項目だけを複数抽出したようなコンテンツの集合などは扱えないのである。このほかにも、意味的な記述をする際に「参照」という仕組みをよく使うのであるが、この参照をすることが許されていなかったり、メタデータの記述自身はMPEG-7に準拠するため、若干複雑で扱いにくい部分がある。本論文では、この部分を解決したものをEvenTankとして提案する。まずはMPFの基本的な部分について以下に説明をする。

MPFの概要とシステムモデル

MPFでは、メタデータ制作に必要な抽出、編集などのメタデータ処理機能とメタデータ蓄積機能を独立性の高いモジュールによって実装するメタデータ制作システムモデルを定義している。また、モジュール間でのメタデータの共有、交換方法やモジュール制御の方法を規定している。

このフレームワークの利点は、モジュールの組み合わせによる構成方法をとることで、さまざまなOSやプログラム開発環境などで分散的に開発されたモジュール

を自由に組み合わせてメタデータ制作システムを実装することができる点である。また、各モジュールが独立に動作するような実装も可能になり、複雑な処理を複数のハードウェアで分担処理することにより処理速度を向上させ、スポーツ中継のような生放送番組でのリアルタイム (もしくはそれに近い速度) のメタデータ生成にも拡張できる。さらに、複数のメディアから得られるメタデータの相互補完、統合により、信頼度の高いメタデータを効率的に制作することができるのである。

図 2.4 は、MPF のシステムモデルを示している。本モデルは、2 種類のモジュールとそれらを接続するコントローラにより構成される。以下に、各構成要素について簡単に説明する。

蓄積モジュール

メタデータ DB を内包して素材メタデータを蓄積、管理し、コントローラ、および、他の処理モジュールからのメタデータ操作インタフェースの規約に準拠した要求に応じてメタデータ DB の検索、更新を行う機能を実装する。メタデータ制作システムには、少なくとも 1 式の蓄積モジュールが常に接続される。

処理モジュール

モジュール制御インタフェースおよび以下の機能を、個別または複合的に実装する。任意の数の処理モジュールをメタデータ制作システムに接続でき、また、着脱も可能とする。

- メタデータ抽出機能：内容に基づいてメディアデータ (映像、音声など) を時間軸上のセグメントに分割するセグメンテーション、および、セグメントの

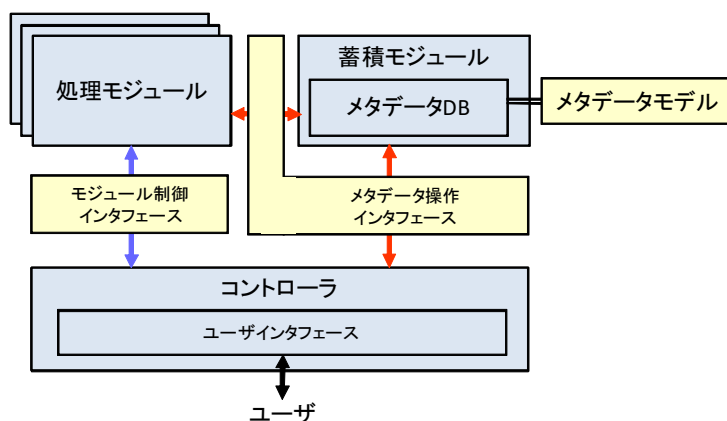


図 2.4: MPF のアーキテクチャ

内容情報を取り出すインデクシングを行う機能。また、他の処理モジュールによる抽出結果を参照してよりを精緻化されたメタデータを抽出する連携抽出機能も実装に含めることが可能である。

- メタデータ統合機能：複数のモジュールが抽出したメタデータを統合して、より豊富な内容記述を含むメタデータを作成する機能。
- メタデータ利用機能：メタデータ DB に蓄積された特定の素材メタデータを読み込み、アプリケーションで必要となる情報の付加、形式の変換、編集などを行い、アプリケーションメタデータを制作する機能。

コントローラ

メタデータ操作インタフェース、モジュール制御インタフェース、および、ユーザインタフェースを実装し、ユーザの操作とコントローラに接続された各モジュールの動作との仲介をする。ユーザインタフェースでは処理モジュールの制御のほか、メタデータ DB に蓄積された特定のメタデータを読み込み、セグメントメタデータの区間情報、内容情報の表示、変更、追加、削除などの編集操作を行う。

表 2.4: EvenTank のシステムモデル構成要素の概要

構成要素	概要
蓄積モジュール	<ul style="list-style-type: none"> ● メタデータ操作インタフェースの実装 ● メタデータの蓄積管理/操作 (追加, 変更, 削除, 検索) 機能
処理モジュール	<ul style="list-style-type: none"> ● モジュール制御インタフェースの実装 ● メタデータ抽出/メタデータ統合/メタデータ利用機能
コントローラ	<ul style="list-style-type: none"> ● モジュール制御インタフェース/メタデータ操作インタフェース/ユーザインタフェースの実装 ● 処理モジュールの制御/メタデータの表示/編集機能

メタデータモデル

MPF で規定するメタデータモデル，すなわちメタデータのフォーマット構造は，MPEG-7 のサブセットを採用している．従って，XML で記述されるため，木構造をなし，1 つの映像コンテンツに対し，1 つ XML メタデータが対応する．映像コンテンツに対するメタデータは，コンテンツ全体に関する情報と，コンテンツの一部分に対する部分的情報にわけることができる．MPF では，前者を基本情報とよび，その内容は，メディア情報（ファイルやテープ等，映像コンテンツの物理的形態や保存場所に関する情報），制作情報（タイトルや制作者などの情報），利用情報（放送日などの情報）の 3 つからなる．また，後者はセグメントメタデータと呼び，映像コンテンツの内容を時間軸上のセグメントに分割して記述する．そして，このセグメントメタデータをある観点でまとめたものを，セグメントブロックと呼び，まとめて管理できるようになっている．MPF の目的は，この様々なセグメントメタデータを効率よく生成するための共通の仕組みということになる．

モジュール間インタフェース

MPF では，図 2.4 のように，様々な処理をモジュール化し，それをコントローラから制御し，生成したメタデータは蓄積モジュールに蓄えて利用する．従って，生成したメタデータを送受するための約束事を「メタデータ操作インタフェース」として，またモジュールの制御のための約束事を「モジュール制御インタフェース」として規定している．

メタデータ操作インタフェースは，メタデータ蓄積モジュールに蓄えられるメタデータの操作を行うためのインタフェースであり，処理モジュールからのメタデータ操作を簡便なものとするため，高レベルと低レベルの 2 つのインタフェースを用意している．そして，MPF における蓄積モジュールは，高レベル操作関数の実装を必須としている．また，ネットワークを介した複数のプロセスによる 1 つのメタデータの更新を可能にするため，操作するメタデータ領域の書き込み権限の管理を行う機能の実装を必須とする．

モジュール制御インタフェースは，処理モジュールが必ず実装するものであり，他のプログラムからの制御を受け付ける窓口となる．MPF で規定するのは 4 種類の関数群のみで，シンプルな構成となっている．モジュールのプロパティ情報や，処理に必要なパラメータ情報などは，MPF で独自に定める 1 つの XML 文書で扱い，これを処理モジュール情報と呼ぶ（以降では，paramXML と表現）．この処理モジュール情報は，決められた関数の中で引数として渡され，必要な情報交換

を実現する．以下，はじめに MPF で規定する関数群について説明し，次に処理モジュール利用の手順について説明し，最後に処理モジュール情報の構造について説明する．

従来技術の問題点

はじめに紹介したダブリンコアは，非常に広く使われてはいるものの，その記述能力があまり高くなく，ニュースの内容を構造化して記述することができない．逆に MPEG-7 は，非常に細かい部分まで規定されているが，複雑すぎて使いづらく，規定項目以外のデータを扱うことができない．また，NewsML は，素材の管理交換には細かい情報を付与できるが，内容については通常の文章による記述であり，やはり意味的な解析をした結果などを記述するには不向きといえる．最後の MPF は，MPEG-7 のサブセットを採用することで，必要最低限の複雑さに抑えている．しかしながら，まだ MPEG-7 の複雑さを継承しており，そのタグ等のオーバーヘッドが問題といえる．また，MPF では，映像コンテンツを単位としており，それらの集合体を扱うようにはなっていない．ニュースなどは，複数の映像コンテンツをまとめて扱うことも多いため，MPF では扱いにくい．また，付与されるメタデータは，メディアのある部分に直接結びつけられるが，ニュースを対象とした従来研究をみると，抽出した情報が必ずしもメディアの特定部分と直接関係があるとはいえない場合も多く，そのような情報を柔軟に扱うことができる必要がある．

このような状況を考えると，その構造はできるだけシンプルで，かつ拡張性を有することが重要であると考えられる．また，研究や開発段階を考えると，様々なメタデータが存在することは明白であり，細かい部分までを最初から規定することは，避けるべきである．実際に，研究開発が進み，プロトタイプサービスとして動き出す段階になった時にはじめて，将来を考えた適切なフォーマットを決めるべきと考えられる．従って，EvenTank におけるデータ構造には，シンプルな構造で，かつ様々なデータを保持できるような場所を用意し，それらをまとめたデータ処理単位で流通させることが必要と考える．

2.5 EvenTank の提案

ここでは，前節までに整理した要求項目を満たすニュース番組再利用のためのフレームワークとして，EvenTank を提案 [38] する．基本的には，MPF をベースとし，必要な追加と変更を加えることで実現している．従って，そのシステムモデ

ルはMPFと同一であり、様々なモジュールと、そのモジュールを制御するコントローラによりシステムは構成される。提案する変更は、メタデータモデルをニュース番組を扱い易いよう変更する点と、これに該当するデータ送受のためのインタフェース部分の変更である。以下では、この変更部分について詳しく説明し、最後に、従来研究がこの EvenTank のシステムモデルのどの部分に該当するのかについて、いくつかの具体例を挙げて説明し、EvenTank の適用の可能性について議論する。

2.5.1 EvenTank のデータモデル

ここでは、EvenTank で扱うメタデータのモデルについて規定する。設計方針としては、前節でも述べたが、シンプルな構造で、かつ様々なデータを保持できる共通の記述場所を用意し、それらのあるまとまった単位で、やり取りできるように設計した。データ構造は、XML スキーマを用いて定義したので、スキーマそのものと、各データタイプなどの詳細は付録に載せ、ここでは基本的な概念構造だけについて述べる。

本フレームワークでは、ニュース番組を扱うデータ単位として、2つのレベルを規定する(図 2.5)。1つは、ニュース番組を構成する1つ1つの「ニュース項目」レベルのもの、もう1つは「あるトピックに関するニュース項目の集合体」レベルのものである。本フレームワークでは、前者を「SourceItem」と呼び、いくつかの関連する SourceItem を集合体としたものを「Topic」と呼ぶ。なお、前者の SourceItem データについては、従来から行われている TDT (Topic Detection and Tracking) 分野においては、Story と呼ばれてきたデータに相当する。以下に、この2つのデータについて特徴を述べる。

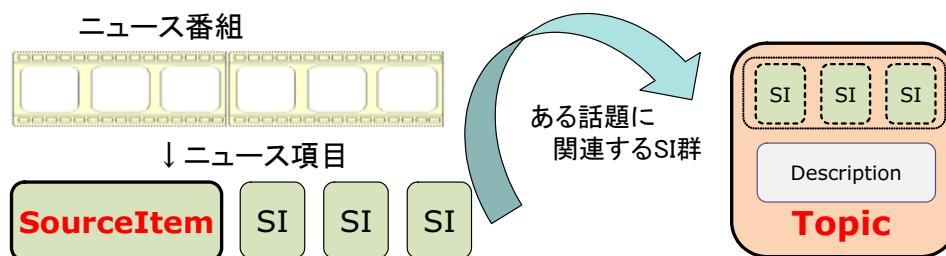


図 2.5: EvenTank における 2つの基本データモデル

データモデル : SourceItem

SourceItem は、ニュース番組を構成する各項目にあたる。この項目に対するメタデータ構造は、図 2.6 で示される。この SourceItem を識別するための ID をはじめ、この SourceItem の作成者情報、ヘッダ情報、代表メタデータ、セグメントメタデータの5つからなる。ヘッダ情報には、メディアの情報、タイトル、放送日などが記述される。代表メタデータは、この項目全体に関わるメタデータであり、その記述方法は、文章、キーワード、後述する Key-Value の3通りを用意している。従って、項目の要約を文章で記述したり、項目のいわゆる 5W1H (who, when, where, what, why, how) をキーワードで表現したりできるようになっている。これに対し、セグメントメタデータは、この項目の中のある部分的な区間(セグメント)に対するメタデータ付与のために設けられている。これは、MPF のセグメントメタデータと同じ位置づけであり、区間と、その区間に対するメタデータで形成される。この場合のメタデータも Key-Value を採用した。ここには、ショット分割

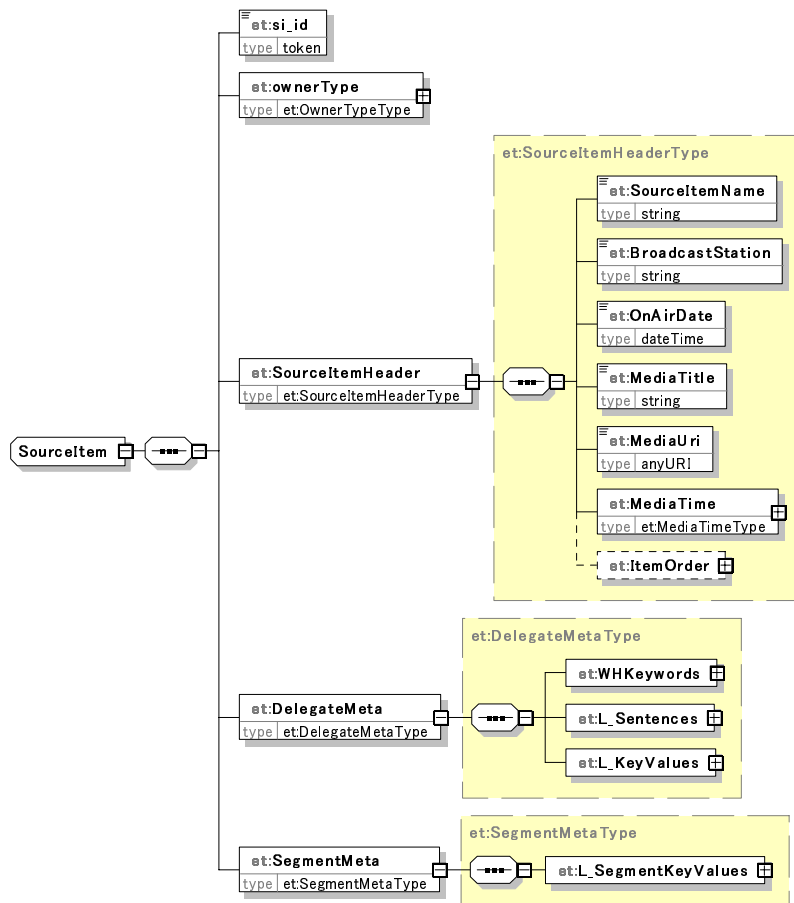


図 2.6: SourceItem の基本構造

情報や、アナウンサの発話情報、顔出現情報等、様々なメタデータが格納される。

次に、メタデータの主たる表現方法として採用した、Key-Value 方式について述べる。Key はメタデータの項目であり、Value は文字通りその値である。簡単な例を示せば、Key = 色合い、Value=赤という単純な構造である。この情報に加え、このデータの制作者情報を加え、区別しやすいようにしている。図 2.7 が、この Key-Value データの構造である。Key-Value の Key は、ownerKey タグに、その Key を定義した組織/人物の情報とともに格納される。その Key に対する Value(値)については、3 種類の記述方法から選択することができ、それぞれ、XML データ内にそのまま記述できる文字列、XML データ内にそのまま埋め込めない文字列、外部ファイルとなっている。また、それぞれの Value に対しても、そのデータ生成者の情報を付与することになっており、これは ownerType タグに埋め込まれる。

このような構造により、柔軟にメタデータを蓄積・利用することができる。例を挙げれば、Key として「顔情報」を作成し、これに対し A 氏は、自分自身で確認しながら作った情報を ownerType = 「A 氏：マニュアル」という情報を加えて保存し、同じ場所に、B 氏は自作の画像処理システムでパラメータを変えて 2 つの抽出結果を出力し、それぞれ ownerType = 「B 氏：パラメータ 1」、「B 氏：パラメータ 2」として保存する。これらと比較検討することもあれば、ある特定の 1 つを

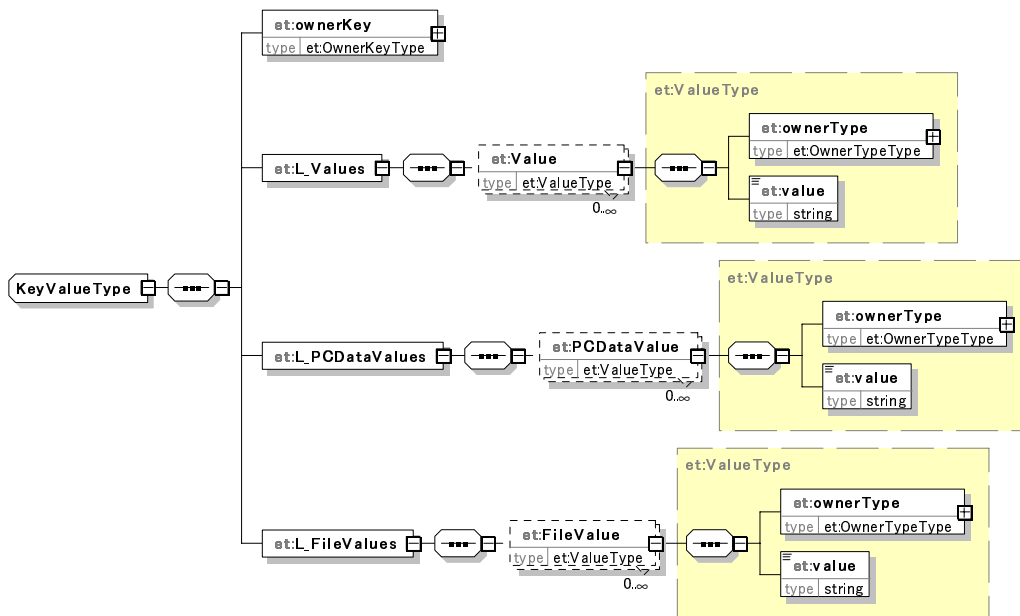


図 2.7: Key-Value による情報記述

選択し、違うシステムの入力として利用することもあり得る。この Key-Value に、

その制作者情報を加えておくことで、最低限の共通環境が確保できると考える。

データモデル：Topic

SourceItem がニュースの項目に相当するのに対し、いくつかの関連する SourceItem をまとめたものが Topic である。ある大きな事件に関する情報を集めたものも Topic であるし、ある日のニュース項目を全て集めたものも「ある日」という観点でまとめられた Topic である。この Topic は、1 つ以上の SourceItem から形成される。図 2.8 は、この Topic の構造を示したものである。まず、Topic を識別するための ID、Topic の制作者情報、Topic のヘッダ情報（トピックの名前、概要、含んでいる SourceItem の放送日の機能）、Topic が含む SourceItem へのリンク、Topic に関するメタデータの記述の 5 つから構成される。Topic に関するメタデータは、〈Descriptions〉タグの下に、0 個以上の Key-Value 方式で記述される。従って、この〈Descriptions〉には、様々なメタデータを格納することができる。例えば RDF や OWL など記述された意味情報などを、XML 内に直接挿入することも、外部ファイルに出力し、そこへのリンクを埋め込む方法を用いることもできる。現段階での規約としては、このデータの汎用的な記述場所までを規定しておき、実際の Topic に関する詳細な情報記述形式については、更に共通形式などについて EvenTank の利用者での合意が得られたところで再考すべきと考える。

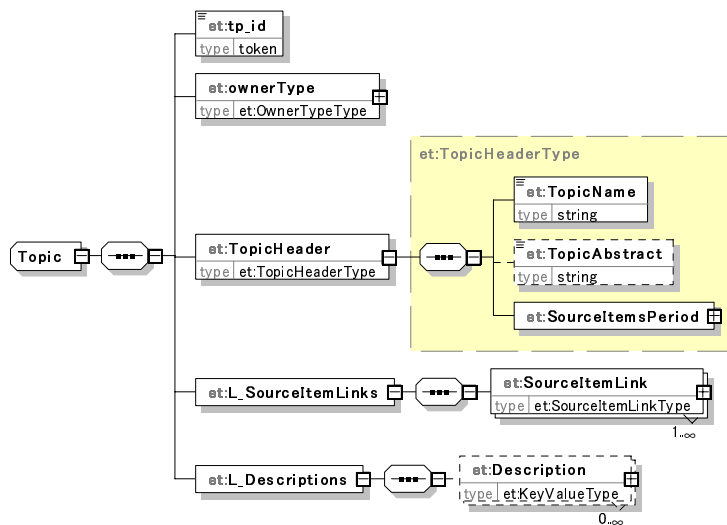


図 2.8: Topic の基本構造

2.5.2 インタフェース

インタフェースについても，MPF の2種類のインタフェースをベースとし，扱うデータモデルの変更に係る部分だけを変更する．

メタデータ操作インタフェース

メタデータ操作インタフェースの変更点は，高レベル操作関数の部分だけである．低レベル操作関数については，MPF 同様に XQuery と，現在 W3C にて規格化が進められている XQuery Update Facility を採用する．一方，高レベル操作関数については，扱うデータが SourceItem と Topic の2種類となり，それぞれの内部における主要部分への入出力が容易になるよう，下記のように定義する．なお，インタフェースの詳細については付録参照のこと．

モジュール制御インタフェース

モジュール制御のためのインタフェースについても，MPF で規定されたものをほぼそのまま利用する．変更点は，モジュール制御関数とモジュール情報の2つについて，扱うデータモデル変更によるものである．モジュール制御関数については，EvenTankSetMoudleInfo 関数内の引数のうち，処理対象を時間で指定する部分を外す．これはMPF の処理対象が，番組などある程度の時間長を持つセグメントブロックに限定されているのに対し，EvenTank においては，ニュース項目である SourceItem やその集合を扱う Topic であり，時間的には短いものが多く時間指定の利点がありませんと判断したからである．なお，EvenTank で規定する4つのモジュール制御関数の詳細は付録に記す．

モジュール情報 (paramXML) については，以下に挙げる変更を行う．主には，セグメントブロック要素がなくなり，代わりに新しいデータモデルに関する要素を追加する．モジュール情報の詳細については付録のそのスキーマを記す．

- ProgramMedia , SegmentBlock 要素以下の削除
- SourceItem, Topic 要素の追加
- SegmentKeyValue, SegmentKeyValueInSpecificSI 要素の追加
- KeyValue, KeyValueInSpecificSI 要素の追加
- Value, ValueInSpecificSI 要素の追加

表 2.5: 高レベルメタデータ操作インタフェース

扱うデータ	関数	モード
SourceItem	searchSourceItem	read
	getSourceItem	read
	insertSourceItem	write
	deleteSourceItem	write
	replaceSourceItem	write
SourceItemHeader	getSourceItemHeader	read
	replaceSourceItemHeader	write
DelegateMeta	getDelegateMeta	read
	replaceDelegateMeta	write
SegmentMeta	getSegmentMeta	read
	replaceSegmentMeta	write
SegmentKeyValue	getSegmentKeyValue	read
	insertSegmentKeyValue	write
	deleteSegmentKeyValue	write
	replaceSegmentKeyValue	write
KeyValue	getKeyValue	read
	insertKeyValue	write
	deleteKeyValue	write
	replaceKeyValue	write
Topic	searchTopic	read
	getTopic	read
	insertTopic	write
	deleteTopic	write
	replaceTopic	write
TopicHeader	getTopicHeader	read
	replaceTopicHeader	write
SourceItemLink	getSourceItemLink	read
	replaceSourceItemLink	write
Description	getDescription	read
	insertDescription	write
	deleteDescription	write
	replaceDescription	write

表 2.6: モジュール制御インタフェースの関数

関数 1	int EvenTankInitializeModule()
	処理モジュールの初期化
関数 2	int EvenTankGetModuleInfo()
	処理モジュールの情報取得
関数 3	int EvenTankSetModuleInfo()
	処理モジュールの情報設定
関数 4	int EvenTankDoProcess()
	処理モジュールの実行制御

2.6 フレームワークの従来研究への適合性

ここでは、MPF をベースにデータモデルと、そのデータを扱うインタフェースを変更することで実現する EvenTank について、その適合性や拡張性について検証する。はじめに、これまで定義したデータモデルとインタフェースを用いて、EvenTank という共通基盤の全体像を図 2.9 に示すとともに、実際のシナリオに沿って説明をする。

まず、ニュース番組は、各項目 (SourceItem) に分割された後、データベースに SourceItem を扱うデータ操作インタフェースを通して登録される。ここでの項目への分割 (丸矢印 1) は、人手で行われるかもしれないし、自動処理で行われるかもしれない。次に、データベース内の SourceItem を対象として、ショット分割や顔検出、付随する字幕データを解析して固有表現を抽出するなど、様々なメディア解析処理が適用され、それらはメタデータとして SourceItem に付与される (丸矢印 2)。また、あるニュースについての情報をまとめて扱うために、Topic データが生成される。この Topic は、ある観点に注目して集めた SourceItem 群を 1 つのグループとして扱うためのもので、ニュース解析の研究では、トピック追跡と呼ばれるものに相当する (丸矢印 3)。例としては、大きな事件の場合などに、事件発生から、日にちがたつにつれ様々な事実が明らかになり、最終的な解決に至るまでの、一連のニュース項目を集めて、1 つのトピックとする場合である。このような場合、集めてきた SourceItem を解析し、全体の流れを再構築 (例えば、サマリ生成など) するような研究もある (丸矢印 4)。さらには、構成要素の SourceItem 群から抽出された複数の情報を、更に統合させることで、より高度な意味情報の記述を試みる場合も考えられる (丸矢印 5)。最近では、複数の情報源を対象にしたサマリ生成もテキストの分野では行われており、複数の Topic を活用して、さらに別の視点からの Topic を生成することも考えられる (丸矢印 6)。

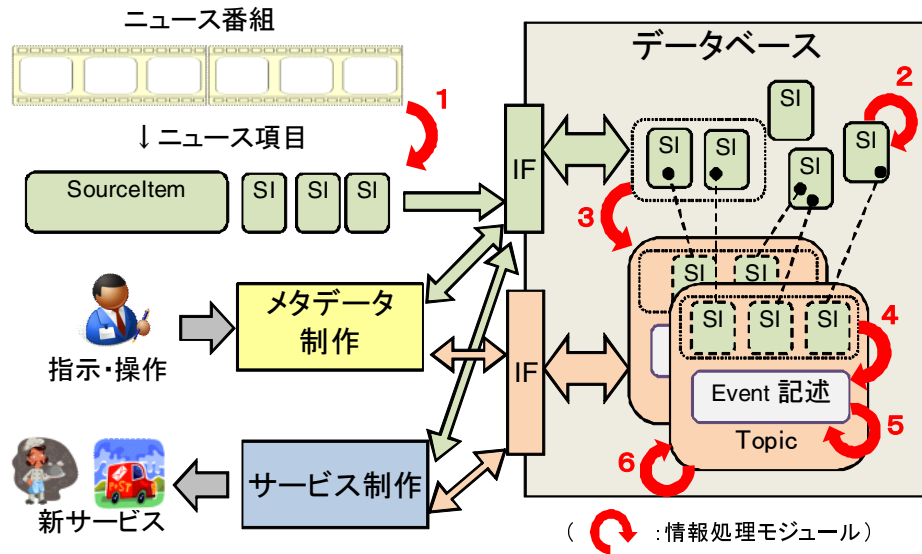


図 2.9: EvenTank のアーキテクチャ

このように、図 2.9 の丸矢印の部分は、実は、従来からも行われている様々な解析処理に相当しており、モジュールと成り得る部分なのである。従って、EvenTank のシステムモデルは、基本的には MPF のシステムモデル (図 2.4) と同じであり、コントローラと 2 種類のモジュール (蓄積モジュールと解析処理モジュール) から構成される。図 2.9 における蓄積モジュールとは、データベースそのもののことであり、SourceItem と Topic データを基本単位としたデータの入出力機能を備えている。一方、解析処理モジュールは、これまで説明したように、丸矢印の部分がこれにあたる。これらモジュール群を制御して所望のメタデータを生成かつ利用するのがコントローラであり、図中では、メタデータ制作とサービス制作という四角部分がこれにあたる。

次に、従来研究からいくつかを選択し、この EvenTank の考え方が適用できるかについて述べる。はじめに、丸矢印 1 と 3 の部分の研究として、TDT (Topic Detection and Tracking)[39] について考えてみる。例えば、字幕のテキスト情報のみによる解析としては [40] がある。字幕からキーワードベクトルを作成し項目の区切れを検出、その後で項目間の類似度によりスレッド (関連項目がリンクされたもの) を抽出、これらを「Topic Finder」と「Topic Tracker」というアプリケーションを通してユーザがある特定のトピックを追跡できるとしている。この研究を例に挙げれば、処理モジュールは 2 つで、1 つ目は字幕解析による項目 (SourceItem) 検出 (丸矢印 1) で、もう 1 つは SourceItem 間の類似度を計算し、Topic を生成し (丸矢印 2)、かつその SourceItem 間のリンク情報を形成する (丸矢印 4) モジュールであ

る。Topic をクエリで検索する Topic Finder と、選択したある Topic についてその関連項目をたどる Topic Tracker は、両者ともコントローラという位置づけであり、Topic データのデータ操作インタフェースを通してデータを取得し、コントローラ内でデータを加工してユーザに提供することに相当する。この Topic Tracker を更に発展させたものに mediaWalker[41] があるが、これもコントローラの位置づけとなる。もう1つ例として、テキスト情報だけでなく、映像の情報として重複する画像とコンセプトを利用したトピック追跡手法 [42] を挙げる。この報告では、画像の重複度を色とテキスト情報により計算し、コンセプトは TRECVID 2005 でおこなわれた 39 のコンセプトを検出し、これらを項目間の類似度計算にテキスト情報と一緒に用いて実験を行っているトピックの検出はあらかじめ人手でおこなっているため、図 2.9 では、メタデータ制作コントローラなるもので SourceItem を生成、データベースに登録する。登録した SourceItem に対し、丸矢印 2 で示されるモジュール処理により色特徴、テキスト特徴、またコンセプト抽出を行う。これら複数の特徴をつかい、項目間の類似度を計算し、Topic を形成する (丸矢印 3)。ちなみに、彼らは実験結果として、画像の重複度は Topic 形成に重要であったが、コンセプト情報はそれほど役に立っていないという結論を出している。

更に、序章で述べた 3 つのニュース番組を利用した従来研究についても簡単にその EvenTank 内での位置づけについて確認しておく。まず、サマリ生成については、Topic に含まれる複数の SourceItem を解析して、サマリを出力することに相当する。このサマリ生成の処理途中における様々な中間データは、Topic の Description(Event 記述) に保存しておくことができる。そして、質問応答など別のシステムで、この解析した中間データを用いることも十分考えられるのである。また、検索については、必要なインデックスを丸矢印 2 の処理で生成し、サービス制作というコントローラでその所望の検索システムを実装することになる。このように EvenTank では、ある特定部分だけについての研究を行う場合にも、また素材加工から最後のアプリケーションまでの一貫した処理についても、どちらにも対応することができ、様々な抽出データのある程度の枠組みを持って共通に再利用する環境を提供している。

2.7 まとめ

本章では、ニュース番組を効率よく再利用するための共通基盤として「EvenTank」を提案した。はじめに、従来の関連研究として、規格化された既存のメタデータや、フレームワークについて、その適合性を検証し、蓄積されたニュース番組を、

効率よく再利用し、研究開発を進めるには不十分であることを示した。そして、既存技術の優れた点をベースとし、必要な拡張を行うことで、EvenTank の提案を行った。そして、提案した EvenTank の適合性を検証するために、個々の従来研究が、本フレームワーク上で動作可能であることを示し、また今後への拡張性/柔軟性を有していることを確認した。

第3章 ニュース番組における映像のメタデータ

3.1 はじめに

本章では、ニュース番組の映像に付与する新しいメタデータとして、映像の役割に注目し、その必要性和有用性を述べる。そして、その役割のうちのいくつかについて、番組の制作ルールなどを利用する手法を提案し、工学的な実現性について検証するとともに、今後の方向性についても述べる。

はじめに、従来行われてきた映像に対するメタデータ付与について整理をし、映像の再利用時に発生する問題について述べる。そして、この問題を解決するために、映像の「役割」というメタデータを付与することを提案する。まず、ニュース番組そのものの役割を考え、その中での映像の役割について、制作手法なども考慮にいれながら議論する。最後に、これら役割のうちの2つについて、番組の制作ルールを利用する手法を提案し、その有効性について検証する。最後に、本章で述べた「映像の役割」についてのメタデータについてまとめる。

3.2 従来研究と問題点

3.2.1 従来研究

ニュース番組のショット映像に対するメタデータ抽出の研究は、一般的にニュース映像の構造化、もしくはショット分類を目的としたものが多い。ニュース番組のショットで、特によく検出対象とされたのはアナウンサショットである。ニュース番組内の各項目の最初に現れることが多いことから、検出して番組の構造化をする最初の手がかりとして注目されたのである。他に、アナウンサショットだけでなく、レポートショットや、CG ショット、天気予報のショットなど、様々なカテゴリについてその分類手法が提案されている。

これらショット属性とその抽出手法について簡単にまとめる。初期の研究では、あらかじめ用意されたテンプレートをを用い、似ている画像をそのカテゴリに分類するものが提案された。[43] は、CNN のニュースにおいてアナウンサショットが

決まった構図であることに注目し、これをテンプレートとして利用し、番組を各項目を単位として構造化する手法を提案した。しかし、この手法では、たとえアナウンサショットであっても、各局各番組でその演出スタイルは違うため、それぞれのテンプレートを作成する必要があり、汎用性が高くない。そこで、アナウンサショットを構成する画像部品を利用したフレームモデル[44]が提案された。この部品領域としては、アナウンサ(人)、画面上部左右に出現するアイコン、項目のタイトル、アナウンサの名前、背景を用意し、これらを組み合わせるいくつかの典型的なアナウンサショットのモデルを構築し、検出に利用している。更に汎用性を高めた手法としては、ショットの遷移統計に基づく手法[45]も提案されており、あらかじめショットを6種類、その境界のタイプを4種類定義し、フレーム間差分、色、音などの特徴を利用し、HMM (Hidden Markov Model) で分類するものである。また、ショットへのメタデータとしては、画面上に現れるキャプション(スーパー、オープンキャプションともいう)を利用する手法[46]や、音声認識により作成したテキストデータ、もしくは耳が不自由な方のための字幕(クローズドキャプション、CCともいう)をあわせて利用する提案[47, 48]もある。

音声認識の結果や字幕情報などのテキストと、映像の特徴を利用して、その間の対応関係を抽出することも一種のショットへのメタデータ付与と考えられる。このような研究としては、ショットを会見・報告、集合、屋外の3通りに分類し、一方で字幕を会見、会議、集合、訪問、場所の5つに分類し、それぞれの適切な対応関係を見つける手法[49]、ショットを会見・報告、キャスト、歩行、集合、CGの5つに分類し、このうち前者4つについて、字幕における特定の品詞の語を付与する手法[50]、画像から顔を検出し、字幕から人物名を抽出し、両者の共起関係を利用して、画像中の人物に対し自動的に名前を付与するシステム[51]、画像中に一定以上の大きさの顔がある場合に、その背景領域の特徴を解析することで、閣議前室、記者会見室、スタジオ、その他という場所を推定する手法[52]など様々な報告がある。また、ニュース番組の構造化には、やはりアナウンサショットが各ニュース項目の区切れ目となりやすいため、現在でも様々な手法が提案されている。例えば、グラフ理論を用いた教師データなしで自動的に判断する手法[53]や、この手法と他の2つの教師データなしの手法[54, 55]を比べ、その長所短所を報告[56]した上で、これら3つを組み合わせた統合的なアナウンサショット検出システムについて報告[57]したものもある。

最近では更に細かい情報をショットに付与することが行われている。例えば、米国立標準技術研究所(NIST)が主催する映像版の評価型ワークショップ TRECVID (Text Retrieval Conference Video retrieval evaluation)[58]にはいくつかのタスク

があり、そのうちの1つに、高次特徴抽出 (HLF: High Level Feature) がある。これはあらかじめ定義されたいくつかのコンセプトが各ショットに含まれているかどうかの判断を競うもので、毎年多くの研究者が参加している。2004年に行われたタスクを例に挙げれば、10のコンセプト(船、電車、海岸、バスケットのスコア、マデレーン・オルブライト、ビル・クリントン、離陸、歩く/走る人、暴力、道)について、それぞれのショットに各コンセプトが映っているか否かを判定し、その正解率を競った [59]。また、同様のカテゴリ分類としては、ワードネットや人手により構築した意味構造に、低レベル特徴量からのマッピングを試みた報告 [60] もある。分類のための意味構造の例をあげると、政治、スポーツ、経済など内容に沿ったカテゴリで、スポーツは更にバスケットボールやフットボールに分かれ、バスケットボールは更にフリーショットや得点ゴールなどに分割されている。

3.2.2 再利用における問題点

これらの研究にみるように、ショットに対して様々なメタデータが付与されるようになってきており、その再活用が期待される場所である。ニュース映像の再利用という視点では、報告 [61] が、全てのショットをアナウンサショットかそれ以外、また、実写かCGかを判別し、アナウンサショットとCG以外は再利用可能としている。しかしながら、サマリなどの生成を考えると、アナウンサショットのイントロ部分は1つの候補部分となり得るし、CGなども言葉の説明ではわかりにくい部分を映像で説明しているのだから、一概に再利用には不適切とは言えないと考える。また、[61] では、再利用には2通りあり、1つは過去を振り返るときの素材として利用する場合で、もう1つは汎用的なショットの流用としている。更に、アナウンサの発話内容と映像が必ずしも一致しないことも述べており、従来よく行われている映像に同期した発話をメタデータとして利用する手法には問題があることも述べている。

ここで、[61] が述べている2つの再利用を想定し、従来の映像のメタデータの有用性について考えてみる。特に、最近多くの研究者がその付与を試みている「車」や「空」などの高次特徴について考えてみる。これらは、汎用的なショットの流用という場合には有用と言える。汎用的なショットの流用とは、とにかく「車」や「空」など、特定の高次特徴が映っている映像が必要な場合に、その映像がニュースの話題の中で持つ意味的な観点は無視して、単なる高次特徴が映っている映像として利用することである。一方、過去を振り返るときの素材としてや、元のニュースに沿って別のコンテンツに展開する場合の再利用においては、高次特徴だけでは適切な再利用はできないと考える。例えば、街角の道路を映した映像があった

とする。高次特徴では「道路」「屋外」「街」などのメタデータが付与される。そして、実際のニュースの中でのこの映像の位置づけは、事故現場を伝えるものだったとする。この映像を元のニュースに沿って扱う場合には、前者の「道路」というメタデータでは不十分であり、「事故現場」もしくは、もう少し抽象化したイベントの「発生場所を伝える」役割というメタデータが有用であると言える。従って、その映像がその時に用いられた理由が重要であり、それを抽出・利用することで適切な再利用が実現できると考える。別な言葉で表現すれば、「映像の役割」がわかることが必要なのである。

これを従来の映像特徴のみからコンピュータに判断させるのは難しい。しかし、ニュース番組という対象に限定すれば、その制作のパターンなどを参照することで、この「映像の役割」が抽出できると考える。従って、今後は、再利用のためのメタデータという視点での情報抽出を行う必要がある。

3.3 ニュース番組における映像の役割解析

ここでは、ニュース番組における映像の役割について考えるため、まずニュース番組そのものの役割について考える。次に、その役割の中で、映像メディアが担う役割について述べ、制作側に視点をおいた映像の役割について整理を行う。そして、工学的にどこまで検出可能なのかについて議論し、最後に2つの役割について制作ルールを利用した抽出手法について検証する。

3.3.1 ニュース番組の役割とその制作

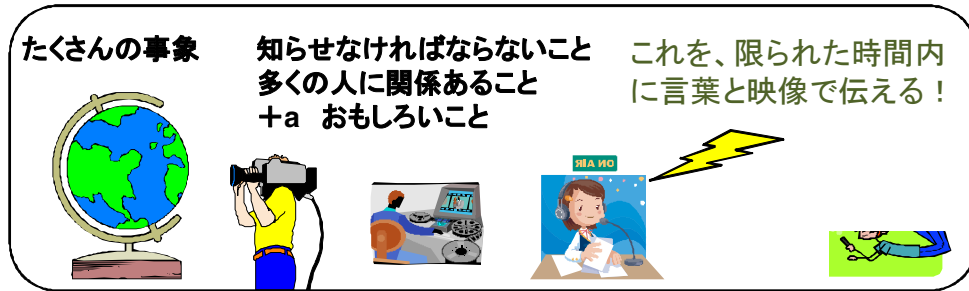
世界には多くの放送局があり、専門チャンネルを除く大きな放送局には、かならずニュースの時間枠がある。このニュースという普段何気なく見ているコンテンツを、放送局はどのような目的でどのように制作し、視聴者はそこから何をどのように得ているか、以下で改めて考えることにより、ニュース番組における映像の役割について考える。

一般に、ニュースの役割とは、世の中の出来事をわかりやすく伝え、視聴者の生活に少しでも役に立つ情報を提供することである(図 3.1)。

世の中の出来事とは、特別な場所で起こっているものだけでなく、森羅万象、全ての人々の活動が出来事になり得るのであり、その数は無限である。放送局は、このように無限に近い出来事の中から、多くの視聴者の生活に関係がありそうな、また有益とおもわれる出来事を選択し、それを編集という作業を通して、限られた短い時間にできるだけわかりやすく伝えようと努力をしていると言える。ここに

ニュースは、どうして放送するの？

> 世の中で起こっていることを伝えて、視聴者の生活に少しでも役に立ててもらえればと...



何をどう伝えれば良い？ 制作のノウハウ＝パターン

図 3.1: ニュース番組の役割

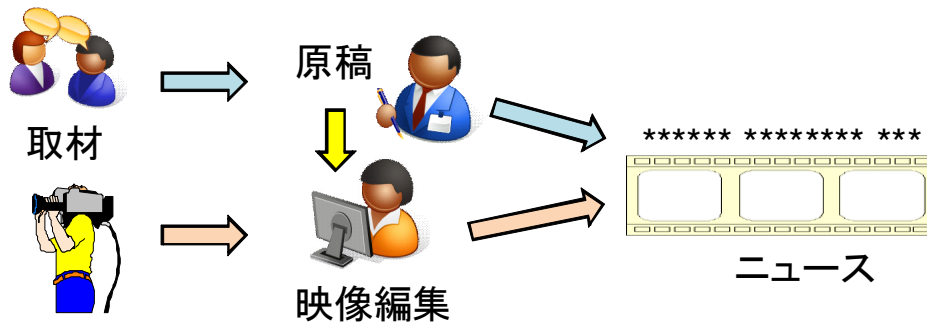


図 3.2: ニュース番組の制作

は、長年にわたり培われたある種のノウハウがあるはずであり、それらは限られた時間内に、言葉と映像でどのように伝えればよいかというパターンを形成していると考えられる。従って、そこには「映像の役割」というものがあると考えられる。以降では、ニュース番組の制作手法から、映像の役割を考えてみる。

図 3.2 は、ニュース番組の制作について簡単に示したものである。まず、伝えるべき出来事について、記者が取材を行い、同時にカメラマンが必要な映像を記録する。これらを放送局に持ち帰り、記者は時間枠にそって原稿を作成する。できあがった原稿は、映像編集者に渡され、取材映像は原稿に沿って編集される。最後に、原稿はアナウンサーによって読まれ、編集映像は VTR から送出することで、一般的に TV で視聴するニュース番組となる。これらの制作手順を見てもわかるように、ニュースというのは原稿が主体であり、映像は付属要素とも考えられる。TV そのものがなかったラジオの時代には、全ては言葉で伝えられていた。では、

ニュースにおける映像の役割とは何であるのかについて、ここで改めて考えてみる。この疑問への答えの1つは、TVのニュース番組を目を閉じて視聴した場合と、普通に視聴した場合の差分について考えることで見つけることができる。映像を見ないで視聴した場合、多くの人々がニュースの内容についてはおおむね理解できると思われる。これは、人間がコミュニケーションを言葉を中心としておこなってきたため自然なことだと考えられる。次に、普通にニュースを視聴する。映像なしの場合からの変化としては、ニュース内容の理解度が増したと感ずるのではないかと想像する。これは、ニュース中の言葉では表現されていない部分があった、それが映像により伝えられたためと考えられる。この例のように、実は映像にしか伝えられないものも世の中には多く存在すると言える。例えば、秋の紅葉について伝える場合、言葉を使ってどんな表現したとしても、その美しさを正確に伝えることは難しい。しかし、映像を用いれば、それだけで十分であり、わざわざ言葉を使って、その美しさについて表現などすることなど必要がない場合もある。また、ニュース内の解説で使われる図などを考えると、映像には別の役割もあるといえる。それは、言葉だけによる説明では、人間の短時間記憶能力に限界があり、全体を把握しながらの詳細部分の理解は、かなりの注意力を必要とする。しかし、映像により全体図を示すことで、全体は視覚により把握しつつ、言葉による説明で、その伝えようとしている本質部分に注意を向けることが可能となる。このように考えると、ニュース映像は、言葉では表現できないもの、もしくは表現しにくいものを瞬時に伝えるとともに、百聞は一見にしかずという諺にもあるように、人間の理解を大きく手助けする役割があるといえる。

次に、もう少し細かな役割について考えてみる。今度は、ニュース番組を音なしの映像だけ、しかもスーパーなどの情報もないと仮定をして視聴した場合について考えてみる。ニュース映像によっては、何のニュースなのか想像つくものもあれば、全くつかないものもあり、詳細な内容についてはほぼ把握できないと思われる。ここで注目すべき点は、「想像がつく」と思われる映像であり、次のような場合があると考えられる。1つ目は、映っているものが明確であり、その映っているものに関する何らかのニュースであると推測する場合で、多くは人物が大きく映っており発話している場合である。このような映像からは、その人物がそのニュースに深く関わっているということだけでなく、その表情等をみることでその人物の考えや気持ちについても推測できる可能性もある。また、建物が大きく映っている、もしくはある物体が大きく映っている映像も、同様の「想像がつく」と思われる映像である。2つ目は、過去に視聴した映像と照らし合わせて内容が推測できる場合である。簡単な例としては、その映っている場所から推測できるもので、国

会の映像であれば、国会で何かが審議されたことに関するニュースであることが、容易に想像することができる。また、どこかの交差点の映像が映っていて、路上にタイヤの跡などが映っている場合は、多くの人は交通事故のニュースだと想像し、そこに手向けられた花が映っていれば、誰かが亡くなった現場であることを類推すると思われる。これらの例が示すように、映像をみて、そのニュース内容の一部を類推できるということは、現在制作されているニュースには、ある種のパターンが存在することを示唆しているといえる。そして、前者の例では、映像における「主被写体」に注目することで、また、後者の例では、ニュースの項目別の映像パターンを利用することで、ニュース映像再利用のための有用なメタデータが抽出できると考える。以上の議論により、映像にも物事を伝える有用な情報が含まれており、その映像の役割の一部分は自動抽出ができそうであることがわかる。

3.3.2 映像(ショット)の役割メタデータ

ここではショットの役割について考え、映像特徴や制作ルールを利用した抽出手法について検討する。ニュース映像をその制作場所やその制作形態により分類したものを、図 3.3 に示す。映像の制作場所としては、放送局内で生成されるものか、放送局外、すなわち現場で生成されるものかで分類することができる次に、映像の生成時刻に注目すると、放送時点で生成する生演出と、予め用意しておく録画とに分けることができる。現場での生演出とは、一般的に中継と呼ばれるものに相当し、生演出以外は、テープなどに録画されたもので、放送局で最終的な素材映像に編集される。これら現場からの映像の役割は、生演出/録画には関係なく、現場における何かを伝える役目を担っている。代表的な例を挙げれば、ある人物の「主張や訴え」をはじめとし、ニュースが対象としている人物、もの、イベント、現象、更にはそれが発生した「場所」や「様子」などである。一方、放送局内での生演出は、スタジオショットがこれに相当し、それ以外はCG等による映像である。スタジオにおけるショットは、大きな事件などの場合には記者やゲスト、模型等を使った説明の場合もあるが、ほとんどの場合が各ニュース項目の導入やまとめなど、進行に関わるものである。また、CG等は、ある調査の結果や、スポーツの結果、また気象情報などで用いられ、「説明の補助」という役目と考えることができる。

表 3.1 は、2006年10月1ヶ月間のショット 6673 個を人手により分類した結果である。まず、制作場所別にみると、放送局内で制作される映像は全体の約 1 割であり、残りの 9 割は現場の映像であることがわかる。また、制作時刻別にみると、生演出によるものは放送局内外あわせて約 1 割、録画によるものが約 9 割となっ

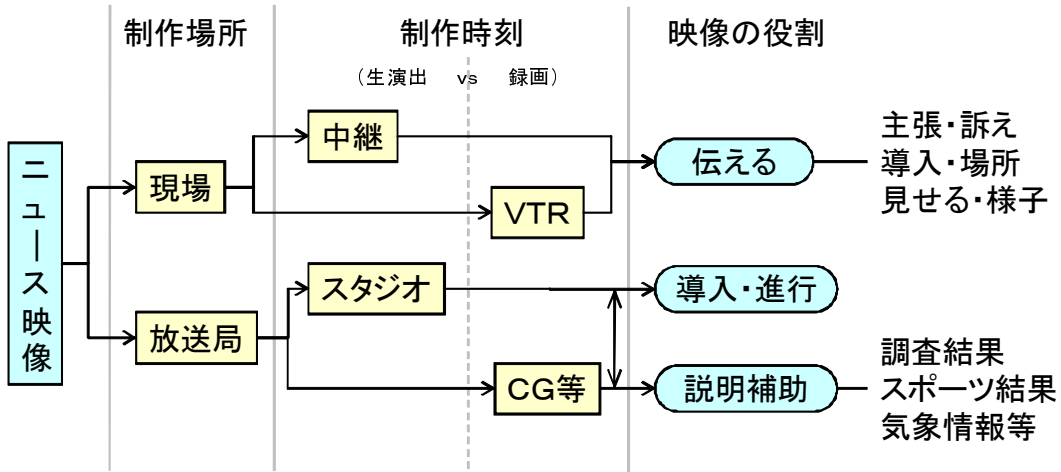


図 3.3: ショットの役割分類

表 3.1: ニュース映像の制作過程による分類

分類	ショット数	割合 [%]
現場	5995	89.8
中継	99	1.5
録画	5896	88.3
放送局内	678	10.2
スタジオ	500	7.5
CG等	178	2.7
合計	6673	100.0

いる。現場で生成される映像にのみ注目すると、そのうち 1.7% が生中継によるものであり、ほとんどが VTR など録画によるものと言える。この録画による映像の全体に占める割合は 89.9% と、ほぼ 9 割であり、この部分の映像をどう再利用するのが今後の鍵であることがわかる。一方、放送局で生成される映像にのみ注目すると、そのうち 73.7% がスタジオショットであり、残りの 2 割強が CG 等になっている。映像の再利用という点では、図 3.3 の「映像の役割」で示した「伝える」と「説明補助」という部分に相当する映像を抜き出し、更にそれらを細かな役割に分類する必要がある。例えば「伝える」映像は、人々の主張や訴え、場所、様子など、その伝えるもの別に、「説明補助」映像は、何かの説明、調査の結果、スポーツの結果、気象情報などである。

伝える役割

図 3.3 の「伝える」役割について更に細かい役割について考えてみる。この中に分類される映像で、極めて重要と考えられるのは、ニュース項目内で発言をしている人物の映像といえる。この映像は、その人物の意見や考え、すなわち「主張・訴え」を伝えるための映像であり、そのニュースを形成する重要な映像要素となる。このような映像は、一般的にはモノローグショットと呼ばれ、従来研究も多くおこなわれている [62, 63, 64]。例えば [64] におけるモノローグの定義は、映像中の人物が肉声で発話している連続した映像としており、ニュース映像においては、アナウンサー、記者、天気レポーターをはじめ、各ニュース項目内で発言をしている人々もこれに相当する。従って、ニュースの内容に深く関わる発言を抽出するためには、放送局関係者のモノローグを除く必要がある。これらについても、画面中の顔情報に加え、音声データを利用した話者特定を行い検出する手法などが提案されており、「主張・訴え」の役割を付与することは十分可能と考える。

その他の映像については、様々な役割が考えられる。なぜなら、役割自体、ある視点から見たものであり、別の視点から見ると同じ映像が違う役割を果たすことは多々あると考えられる。一般的に、出来事などを単純に記述する場合に、いわゆる 5W1H (when, who, where, what, why, how) を用いるが、これが役割として当てはまるかについて考えてみる。先の主張や訴えの役割は、who と what の両方を有していると考えられる。また、現場の where と how を伝える映像なども存在すると言える。残りの when, why については、映像によっては暗にそれを伝えていることは十分あり得ると考える。従って、ニュースを構成する映像については、それぞれ何らかの役割、言い換えるとそれは制作者の意図、があることは確かであると言える。しかしながら、映像の多義性という問題もからみ、この映像は、この役割であると 1 つに決めることは難しい。また、画像の特徴からだけでは、解決できない問題と思われる。

説明補助の役割

次に、CG 映像や、全面にわたってスーパーが付加された映像の役割について考える。これらの映像は、先に挙げた 5W1H には該当せず、ある事柄を説明する際の補助的な役割を有すると考えられ、再利用する場合にも有用であることが多いと考える。例えば、複数の人間や組織でのやり取りを説明したもの、また事件現場における見取り図などで、言葉による説明だけでは全体が把握しにくい場合に用いられる。また、宇宙における探査機の動作や、ミクロの世界での現象など、一

般的には想像し難い内容を解説する場合にも多用されている。この他、調査の結果や、重要な項目の箇条書きなどを一定時間みせることで、視聴者のペースによる内容把握という役目もあると考える。これは、一方的なペースで言葉を中心に進行するニュース番組に対し、新聞などの読者のペースによる把握という利点を、少しでも映像により補完しようとしているものと考えられる。スポーツの結果や、気象情報などについても同様な理由が考えられる。更に、気象の場合の天気図を考えると、その役割は大きい。それは、言葉では全てを短時間に伝えられないのであるが、図として現在の気象状況を描画して見せることで、多くの情報を視聴者に提供しているのである。従って、視聴者は、言葉としては伝えられていないが、画像をみることにより自分なりの推測が可能になるなど、言葉にはない映像だけの情報伝達の利点を最大限に活かしたものといえる。これら CG 映像や全面スーパーの映像検出についても、従来研究として CG 検出を行うものや、スーパー検出を利用したインデックス生成の報告がいくつもあり、それらを組み合わせることで、ある程度検出が可能と考えられる。

役割の自動抽出について

先にも述べたように、ニュース映像の再利用という観点では、その映像の数も考慮すると、現場から何かを「伝える」映像について、いかに細分類を行うことができるかが鍵となる。「主張・訴え」の映像については、先モノログ検出という手法である程度分別することが可能であるが、残りの「場所・様子」をはじめとする、何かを伝える映像については、その分類はかなり難しい。この原因の1つは、先に挙げた映像の多義性に加え、同じ「場所」を伝える映像であっても、内容や場合によりその映像内容の変化の程度が大きいからである。初期の TRECVID におけるコンセプト抽出のように、各ショットにおける代表静止画の色やテクスチャなどの幾何学的特徴だけに注目しても、そこから多くのパターンを検出するのは難しい。更に、これらの映像は、映像を撮影する際の制約により、パターン抽出を難しくさせている。カメラマンは、後々の映像編集作業を考えて、その出来事を伝えるのに適切と思われる映像を、その場で瞬時に判断し記録する。その際に、何をどのように映すべきかについては、マニュアルがあるわけではなく、口頭で受け継がれ、またこれまでに制作された映像を見て暗に取得していくものであり、ある程度のパターンがある。しかし、実際の取材時には、撮影場所や撮影時間など様々な制限が加わることがあり、所望の映像が撮れない場合も多々ある。その時に、代わりに何をどのように撮影するのかなどは、カメラマンの判断に完全に委ねられており、放送局ではこれら得ることができた映像を元に編集作業を行う。

このように臨機応変な対応が要求される制作現場であるがゆえに、同じような目的であっても多彩な変化に富む映像が撮影されるのである。しかしながら、そこにはある程度の暗黙のルールが存在すると考えられたため、以降では、色やテクスチャなどではなく、より制作ルールに近い特徴を利用した役割抽出手法について検討する。

3.4 制作ルールを利用した役割抽出

これまで議論してきたショットの役割の内、2つの役割ショットについて、制作ルールを利用した手法を提案し抽出を試みる。1つ目は、ニュースの中では重要な役割を担う、主張や訴えのショット、一般的にはモノログといわれるショットの検出である。2つ目は、場面の説明をするための導入の位置づけ、もしくはユーザに対して言葉でなく「映像で見せる」という役割を持つショットの検出である。

3.4.1 顔情報のみを利用した「主張・訴え」の役割抽出

提案手法

ここでは図 3.3 の「主張・訴え」を伝えるショットの抽出を、その映像の撮影方法、別な表現を用いれば演出方法から抽出を試みる。利用する情報は、画面中の顔の情報のみとし、この表層的な情報だけによる抽出性能を検証する。抽出目標となるショットは、モノログショットに属し、かつアナウンサや記者、天気レポータなど放送局関係者でない人物のショットと定義する。図 3.4 に提案する手法を示

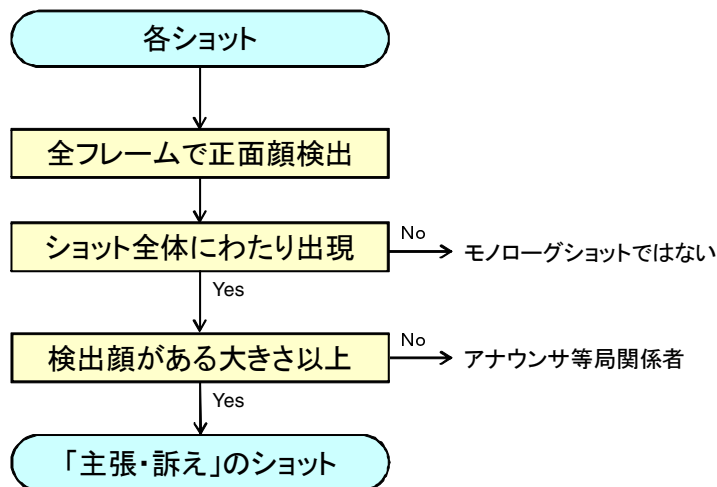


図 3.4: 「主張・訴え」を伝えるショットの抽出手法

す。各ショットを入力とし、その全フレームに対し正面顔の検出を行う。そして、ショット全体にわたり顔が出現した場合をモノログショットとし、更に、検出した顔の大きさにより、放送局関係者かそうでない人物かを分類する。すなわち、放送局関係者の顔を画面上に大きく映すことはしない、というある種のルールが存在すると仮定している。そして、映像というメディアを通して、人々の主張や訴えをその感情も含めて伝える場合に、どのように制作することが望ましいのかという問いに対する1つの答えとして、できるかぎり正面向きで、顔の表情がよく見えるよう、画面上に大きく撮影すると考えられる。

実験

実際に、2006年10月の全ショット6,799個に対し、このアルゴリズムを適用した。ショット分割については予め人手により正解を付与し、正面顔検出はOpenCV[65]を利用した。ショット全体にわたり出現したか否かの判断は、次のように行った。ショットを等間隔に10の小区間に区切り、各小区間において顔が検出されたフレームの割合を求める。この割合がある閾値(τ_1)以上の小区間を、顔ありの小区間とし、これが全体のうちある一定以上の割合(τ_2)で存在するとき、ショット全体にわたり顔が出現するとした。ショット全体にわたり顔が出現したものをモノログシーンと判断し、検出された顔の縦の大きさが画面のある一定以上(τ_3)である場合を「主張・訴え」のショットとした。なお、3つの閾値は、それぞれ $\tau_1 = 0.0$, $\tau_2 = 5$, $\tau_3 = 1/3$ とした。 $\tau_1 = 0.0$ が意味するところは、小区間内に1フレームでも顔が検出されたら、その区間は顔ありと判断することに相当する。

実験結果を表3.2に示す。全ショットのうちの959ショットが主張・訴えのショットとして抽出された。これをあらかじめ人手により付与した正解データと比較したところ、再現率は59%、適合率も59%となり、F値も0.59となった。過剰検出について分析すると、27.4%が局内関係者のモノログであった。これにはアナウンサー以外に、主に中継先での記者や気象情報のリポーターなどが含まれていた。更に、過剰検出で局関係者ではないショットについては、その93.0%が人物は大きく映っているが本人がしゃべっていない場合であり、残りの7%については写真など静止した映像であった。

表 3.2: 正面顔情報のみによる主張・訴えショットの抽出

正解数	検出数	検出中の正解	過剰検出	未検出	再現率	適合率	F 値
964	959	564	359	605	0.59	0.62	0.60

一方、検出できなかったショットについて調べると、主に顔が横向きや、傾いている、また大きく動いているなどの原因で検出できていない場合である。また、めがねをかけている人、帽子をかぶっている場合の検出もできていない場合が多かった。特に、スポーツ選手などは、サングラス、帽子の着用が多く、かつ首の動きも多い傾向にあり、検出が難しかった。

考察

本実験では、正面顔の情報のみという極めて単純な特徴のみを用いて、主張・訴えのショット抽出を行った。単純な情報ではあったが、F値で0.6を超える精度を得た。更に精度を上げるために、従来研究 [64] でも用いられている音声情報を利用すべきと考える。これによりアナウンサであるか否かを判断し、過剰検出の93%を占める、画像に映っている主張や訴えをする人物とアナウンサの発話が重なるショットを除去することができる。従って、その精度も実用的なものと考えることができる。

更に、本手法で設定する3つの閾値を変化させ、検出される局関係者と一般の人物数の変化をみることで、次の3つの点が明らかになった。1つ目は、局内関係者の顔はあまり大きく映さないということである。これは顔の大きさの閾値である τ_3 を変化させ、画面の縦方向に7/15以上ある大きさの顔に絞ると、局関係者ショットは1つも検出されなくなった。また、10分割した各小区間において、顔ありの小区間と判断するために必要な顔検出の割合の閾値(τ_1)を変化させたところ、一般人物の検出は、条件を厳しくするとともに検出数は減少したが、局関係者の検出は一定であった。これは、局関係者は、ほぼ正面を向き、動かずカメラに向かって発話していることを示す。更に閾値 τ_2 は、条件を緩くすれば検出数も増えるが、その増加率は緩やかで、顔の大きさほどの変化はない。1ヶ月分の全ショットに対し、顔の大きさが1/3以上という条件で、ショット中に正面顔が検出された割合でヒストグラムを作成すると、両端で最も頻度が高く、中間は少ない結果となった。これは、顔のある程度大きく映すショットについては、全く顔が出現しないか、するのであれば、最初から最後まで出現していることが多いことを意味する。

本実験を通し、顔情報という表層的な情報だけであっても、暗に存在する制作ルールを仮定することで、ショットの役割付与のために有用な情報となり得ることを確認した。

3.4.2 ナレーションのタイミングを利用した「導入・見せる」役割抽出

サイレントショット

ここでは図 3.3 の「伝える」という役割の中で、どのような場所での出来事かを示すような導入的役割をもつショットや、ある出来事そのものを映像を通してユーザに「見せる」ことで様子等を伝えるショットの抽出を試みる。注目するのは、誰も発言をしていないショットで、本論文では、このショットのことを「サイレントショット」と呼ぶ。いくつかのサイレントショットの典型例を図 3.5 示す。(a)(b) は記者会見の例で、(c)(d) はスポーツにおけるサイレントショットの例である。(a) は、「フランスの凱旋門賞に出走したディープインパクトから、フランスで禁止されている薬物が検出された」ニュースである。はじめに、記者会見場全体の様子を伝えるサイレントショットがあり、その後に薬物検出に関する発表のショットが続いた例である。(b) は、「安倍首相の発言」を伝えるのが主目的と考えられるニュースであるが、安倍総理がどこでこの発言をしたのかを映像でも示すべく、国会全体を俯瞰するサイレントショットが挿入されている例である。(c) は、「ゴルフのパット」シーンである。これは勝敗を決めるパットで、選手がパットを打つ瞬間がサイレントショットになっており、視聴者に「見せる」ためのシーンと考えられる。続くシーンでは、映像でパットの行方を見せると同時に、アナウンサによる解説が入っている。(d) も野球における同様の演出である。ピッチャーが投げて打つシーンがサイレントショットで、続くショットでその結果について伝えている。このようにみると、注目しているサイレントショットには「導入・見せる」という役割があることがわかる。

提案手法

サイレントショットの定義は、誰も発話していないショットであるので、利用する情報は、ショットの切り替わりとアナウンサの発話情報である。これは、一種の演出方法を利用した手法である。図 3.6 に提案手法の処理の流れを示す。入力は、ショット境界の情報と、時刻付き発話情報であり、全て人手により付与したデータを用いた。ただし、ショット境界の検出技術は現在かなり安定したものとなっているため、これを自動で行わせることは可能であると考えられる。また、発話区間についても、話者認識までは難しいが、誰かが発話しているという情報だけで十分であるため、実現可能と考える。入力されたショット情報に対し、その区間内で発話があるか否かを調べ、発話がない場合、そのショットをサイレントショットとして出力する。

(a) 記者会見の例1



記者会見会場



発表 (モノローグ)

(b) 記者会見の例2



議論会場 (国会)



発言 (モノローグ)

(c) スポーツの例1



パットを打つ瞬間



パットの行方

(d) スポーツの例2



野球打つ瞬間



打ったたまの行方

図 3.5: 「導入・見せる」役割をもつショットの例

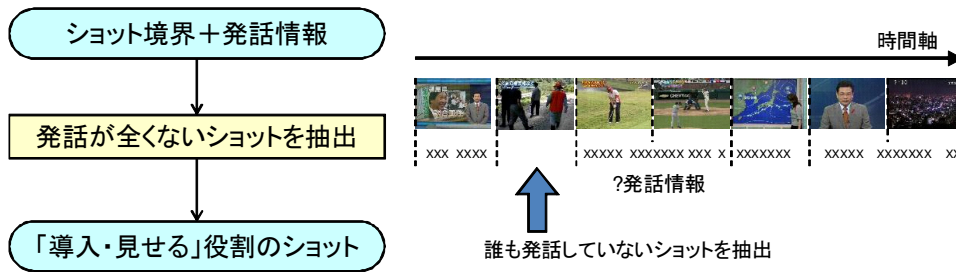


図 3.6: 「場所・導入/見せる・様子」を伝えるショットの抽出手法

実験結果と考察

実験結果を表 3.3 に示す．全ショットのうちの 475 ショットがサイレントショットであった．これは全体の約 7.0%にあたる．各々についてショットの役割を確認したところ，約 8 割強のショットは，仮定した通りの導入，もしくは様子を見せるショットであることがわかった．残りの 2 割のショットは，ほぼ等しく 3 つに分類することができ，それぞれ観客のシーン，タイトル，捨てカットであった．観客のシーンについては，現場の雰囲気を与えているシーンであると解釈できなくもないが，内容を特定するための情報にはかけると考えたため，本実験では除外した．また，捨てカットとは，画面全体に文章をスーパーすることで内容を伝えるショットの一部であり，アナウンサが説明した後に，視聴者に読んでもらうためにしばらくそのままの状態を保持したものである．

本実験により，誰も発話していないサイレントショットは，確かに意図的に存在し，その多くは次のショットへの導入や，イベントそのものを見せるという重要な役割を担っていることが確認できた．導入という役割についても，場所や雰囲気を映像で示すという役割はもちろんのこと，前の話題からの区切りという役目も果たしていると考えられる．以上のことから，これら暗に存在する制作のルールを仮定することで，ショットの役割付与が可能になることを確認した．

表 3.3: サイレントショットによる導入・見せるショットの抽出

導入・見せる	その他			合計
	観客	タイトル	捨てカット	
385 (81.1%)	29 (6.1%)	33 (6.9%)	28 (5.9%)	475

3.5 まとめ

本章では、ニュース番組の映像に付与する新しいメタデータとして、映像の役割に注目し、その必要性和有用性を述べた。そして、その役割のうち「主張・訴え」を伝えるショット（一般的にはモノローグショット）と「導入・見せる」役割をもつショットについて、暗に存在する番組の制作ルールを利用した手法を提案し、工学的な実現性について検証した。主張・訴えを伝えるショットについては、顔の大きさやショット中の出現頻度に注目し、撮影の方法にパターンが存在することがわかり、本ショットの抽出の際には有用な情報になることを確認した。また、導入・見せる役割をもつショットについては、誰も発話していないサイレントショットを導入し、抽出したところ、約8割強のショットが、そのような役割を有していることを確認し、工学的にもこの役割メタデータ付与は実現可能であることを示した。

本章で提案した役割というメタデータは、付与する観点により様々なものが考えられ、完全に整理することは難しい。本論文では、上記2つについて実験を行い検証をしたが、今後は、これ以外のものについても抽出を試みていく必要がある。その際には、映像解析のみからその役割を抽出することは難しいと考えられ、本章で示したような制作側に存在するルール等を利用することが有効であると考える。

第4章 クイズコンテンツの自動生成

4.1 はじめに

本章では、EvenTank の枠組みを利用し、ニュース番組の新しい活用の方向として、クイズの生成実験を行う。はじめに、生成対象とするクイズについて説明し、クイズ生成というタスクを、3つのサブタスクに分割した工学的アプローチを提案する。最後に、実際のニュース番組を用いた生成実験を行い、各サブタスクとそれらを組み合わせたアプローチの評価を行うとともに、EvenTank の有用性について確認する。

この十数年の間に、情報通信のインフラが急速に発達し、コンテンツを取得視聴する方法は、従来のテレビやラジオ、新聞に加え、携帯電話やパーソナルコンピュータ、ハードディスクレコーダなど多種多様に変化している。一方で、同じ内容であっても、それぞれに向けて差別化されたコンテンツを用意する必要があるなど、その制作環境についても効率化を求められるようになってきた。放送局においても、従来の番組制作に加え、ネットワークと連携したコンテンツ [66, 67] を実際に制作/公開したり、サーバ型放送 [68] などに必要なメタデータの制作環境 [69] について既に研究を進めている。現在、番組に関連するコンテンツ制作は、人手による作業がほとんどであり、例えばデータ放送用コンテンツの編集、またはウェブページの作成など、それぞれ専用のスタッフを雇用する、もしくは外部に発注するなどして制作を行っている。このため、この作業の省力化を行う自動あるいは半自動の制作支援システムが注目されている。例を挙げれば、放送番組からのマルチメディア百科事典の自動生成 [70] や、番組の紹介ビデオの自動生成 [71] などである。本章の目的も、これらと同じ位置づけであり、番組を活用するサービスの一例として、ニュース番組から画像付きクイズの生成を試みる。これが実現できれば、放送局では、新しいコンテンツが低コストで制作可能となり、一般の人々は、世の中の時事について楽しみながら学ぶことができるようになる可能性がある。昨今のゲーム機による学習、いわゆる「脳トレ」の流行にもみるように、教育的コンテンツへの接触のモチベーションを高めることにもつながることが期待される。また、インターネット上には、自分の得意分野についてのクイズを生成/公開し、皆で楽しめるようなサイト [72] も出現しており、クイズ生成は、今後

のコンテンツ制作の1つの大きな柱となる可能性もある。

ここでの目標は、全自動を目指すものではなく、半自動を目標としている。もし全自動が可能となれば、大規模なアーカイブや、家庭のハードディスクレコーダなどに搭載することも考えられるが、クイズの生成は、ある種高度な創作物であるため、自動処理は容易ではない。コンテンツの制作/公開を考えると、どのような形で制作が行われたとしても、最終的には再度人手による確認が行われるため、半自動システムであってもその有用性は見いだすことができる。特に、放送局等においては、間違った情報を発信することは許されず、コンテンツの人手によるチェックは必須のものとなっている。このような背景から、はじめに計算機により大量に蓄積された番組からクイズ候補を自動生成し、その中から人手により最終的に使用するものを選択することを想定する。そして、ここでの狙いは、前段のクイズの自動生成をいかに高品質かつ効率的に実現するかにある。

4.2 クイズ生成について

クイズには様々な種類が存在し、対象とするものにより生成のアプローチは異なる。本論文では、ニュース番組を素材とし、画像付きのクイズ生成を行う。はじめに、いくつかの画像付きクイズの例を挙げ、必要な要素技術について述べる。図4.1は、画像と文章を組み合わせて作成できるいくつかのクイズパターンを示している。(a)は、映像の中身に関することを問う記述式の問題である。(b)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。(c)は説明文の真偽を問う問題、(d)は画像にふさわしい説明文を選ぶ問題である。これらのクイズ生成手順は、大きく画像の選択と、画像に関連する文章生成の2つに分けることができる。画像選択における問題は、クイズに適した画像を選ぶ点であり、このようなクイズの場合は、画像中に映っているものが明解であることが望ましい。特に(a)や(d)のクイズは、画像がないと解くことができないものであり、文章との対応関係がより明確である画像でなければならない。一方、文章の生成については、次のように整理できる。(a)の場合は、問題文そのものを生成する必要があり、これはいくつかのテンプレートを用意することである程度生成が可能である。しかし、画像に何が映っているのかがわからないとテンプレートを選択できないため、この点が問題となる。(a)以外の文章については、字幕を基に生成することができる。(b)の場合は、括弧抜きにする語彙の選択がポイントであり、ランダムでは好ましくなく、字幕文中の重要な語彙を判定するなどの処理が必要である。(c)についても同様に、文中の重要な語を判定し、それを正解と間違いやすい語に置き換

(a)(5W1Hクイズ) 次の問に答えなさい



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？

(b)(穴埋めクイズ) 括弧の中に単語をいれなさい



今日、8月9日は、()の日です。午前11時2分になると、
黙祷が捧げられました。

(c)(×クイズ) 説明の真偽を答えなさい



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、ヴェルディでした。

(d)(選択クイズ) 最も適切な説明文を選びなさい



- A. 鴨川市に現れたアザラシです。
- B. 怪我をしたトドが発見されました。
- C. セイウチの人工飼育が始まりました。

図 4.1: 画像付きクイズの例

えることが重要である。(d)は、画像と文章の対応関係に加え、複数ある選択肢が、解答に迷いやすい関係にあることが重要である。これらを整理すると、図 4.1 に示したクイズ生成には、クイズに適した画像の選択、被写体の分類特定、文章のテーマや重要部分の抽出、画像と文章の関係抽出、解答に迷いやすい関係抽出などが必要であることがわかる。

ここでは、図 4.1 の (d) に示した選択クイズを対象とした。なぜなら、その生成手法が、他に比べると簡略化可能であり、また、クイズ以外の役割も付与することができるからである。まず生成手法について考察をすれば、(a)は画像中の被写体の特定が、また (b)(c)は文章中の重要語特定がクイズ生成の鍵であり、それぞれが難しい問題である。(d)についても、画像と選択肢との対応関係は画像の被写体に深く関係し、選択肢間の似て非なる関係は各選択肢の話題等がある程度扱わ

なくてはならず、同様に難しい問題である。しかしながら、その生成手法に注目すると、(d)のクイズは、素材のある一部分(画像もしくは字幕文)を選択し、その組合せにより生成することが可能であるが、(a)から(c)までの場合は、文章そのものの生成や、文の中身をより細かく加工処理する必要があり、この点で(d)よりは難しい問題といえる。

一方、クイズ以外の役割については、素材となった番組へのナビゲーションという役割を考えている。これはクイズ生成時に、複数の選択肢を用意する際に、実際に起こった様々なニュースの字幕をそのまま利用することで、「このようなニュースもあったのか。もう少し詳しくみてみたい」と、クイズ解答者の興味を引くことができると考えている。各選択肢に、該当するニュース項目へのリンクを張っておけば、ニュースの新しい種類の閲覧機能が提供できるのである。本研究でクイズ生成を目的とした理由として、制作した番組を少しでも多く見てもらいたいという意図もあり、この点からも(d)を選択した。

最後に、クイズ生成の難しさについて考えてみる。クイズ生成という処理を簡略化すると、ある事実が存在するとき、その一部分を隠し、そこを問う問題を作ることである。例えば、AとBにCという関係がある場合、「AとBの関係は？」もしくは、「BとCという関係にあるものは？」という質問を作ることである。この時、AとBが文章であったり、画像、音であってもよく、Cという関係も様々な定義ができる。従って、A、B、Cの組合せによってはその事実関係の抽出が難しいこともある。また、このような問いを生成しても、それがあまりに一般的な常識で明らかすぎる場合や、おもしろくない場合には、クイズの質を満たさないこともあり、主観的要素も考慮に入れると、その生成はより難しくなる。更に、生成手法について考えてみると、クイズの役割によって大きく異なる。例えば、単に知っているかどうか(知識の有無)で解くクイズもあれば、クイズを解くことによる学習を目的としたクイズもある。また、おもしろさを埋め込み、笑いや納得を誘うエンタテインメント的なものなどは、同じ素材を用いたとしても、それぞれのアプローチは大きく異なると考えられる。

このようにクイズ生成は、さまざまな要素が絡み合う複雑で奥深い難しい問題を含んでいるが、以下では、クイズの半自動生成への足がかりとして、比較的簡易な処理を用いた場合の限界と可能性を明らかにするものである。

4.3 従来研究との比較

クイズや質問の生成を扱った論文を探してみると、数は少ないがいくつか存在する。文献 [73] は、ある人物に関連する文章を提示し、そこからその人物を連想させるクイズについて報告したものである。提示される文章はヒントと考えることができ、たくさんあるヒントをその連想の難しい順に並べ替える部分の手法を提案している。従って、この論文では、並べ替えの対象である「人物に関連する文章」は、予め与えられており、また扱うデータもテキストのみとなっている。他の例としては、ユーザの質問に答えるだけでなく、システム側が自発的にユーザに質問を行う情報提示システムが報告されている [74]。ここでの質問生成は、あらかじめ京都に関する説明文が存在し、その中の固有表現を隠す形で質問文を生成しており、やはりテキストのみを扱っている。これら以外には、ユーザに情報を提示して理解させる場合は、単に辞書的なものを提示して読ませるのではなく、クイズ形式にして提示する方がより効果的であるという報告 [75] や、職業別案内システムにおいて、お店の表現方法のぶれを学習するためのデータとして、クイズという枠組みを使い、ゲーム感覚で人々の参加を促しデータを収集するという報告 [76] など、クイズというコンテンツの実社会への幅広い応用例が研究として報告されている。本研究は、素材から最終的なクイズ候補生成までの一貫した自動処理の提案であり、対象とするクイズもテキストだけでなく画像も扱い、新しい試みと言える。

次に、同じニュース素材を対象に含む QA システムと要約システムを例に挙げて比較を行う。表 4.1 は、各システムにおける処理手順や入出力、評価について比較したものである。一般的に、QA システムの研究は、ニュースや新聞記事を対象とした入出力がテキストであるものが多く、海外では TREC の Question Answer Task [77]、国内では NTCIR の評価型ワークショップ [78] が有名である。最近では、対象とする情報も多岐にわたり、例えば百科事典を対象としたもの [79] や、Web の情報源を対象としたもの [80]、また取り扱い説明書を対象としたもの [81] まである。また、ニュースの QA として、解答を映像付きで表示するシステムの報告 [82] もあるが、これは解答を字幕文から抽出し、そこに付随している映像を表示しているだけで、映像の意味解析を行っているものではない。このようにシステムは多様化しているが、QA 部分の基本的処理手順 [83] は、表 4.1 のようにまとめることができる。

要約システムについても、多くの報告があり、例えばニュースに関しては、発話単位に区切った映像に、人手によりラベル付けをして要約手法を検討したもの [84]、字幕文から同様な概念グラフを自動生成して要約を試みるもの [85] がある。

表 4.1: ニュース番組を対象とした QA システムと要約システムとクイズ生成タスクの比較

	QA システム	要約システム	画像付きクイズ生成システム
処理手順とサブタスク	質問の解析 1. 質問タイプの分類 解答の生成 2. 解答を引き出す検索キー抽出 3. 解答候補の絞り込み	重要部分の抽出 1. 重要文/映像の選択 要約の生成 2. 編集/再構成	質問の生成 1. クイズに適した画像選択 解答の生成 2. 画像を説明した字幕文選択 3. 似て非なる選択肢の生成
入力	質問文	-	-
出力	解答候補	要約	質問(問題) 解答(正解+不正解)
評価	質問に対する解答 (正解が明確) 答えなければいけない	絶対的正解はない (重要な定義による)	絶対的正解はない (クイズタイプによる)

ニュース以外では、スポーツ映像の要約として、人手によるアノテーションを用いるもの [86]、画像や音声の低レベル特徴を用いて重要部分を抽出する報告 [87] などがある。他にも、料理番組を動き情報から要約するシステム [88] や、ドラマを心理的内容に注目して要約するシステム [89] が報告されている。要約についても同様に、共通するサブタスク [90] を、表 4.1 にまとめた。クイズ生成については、生成対象とした画像付き選択クイズについて記した。

はじめに、各処理手順において要求される、よりブレイクダウンしたサブタスクについて比較を行う。表 4.1 を見ると、「質問タイプの分類」や「重要文/映像の選択」、「クイズに適した画像選択」には、質問タイプの定義や、重要な定義、クイズに適するの定義など、まだ漠然とした部分が含まれており、これらは、より工学的で具体的な処理に次々と置き換えられていく。これら置き換えられた処理を比較すると、一般的には、表中のどのシステムにおいても用いられる基本処理は似ていることが多い。例えば、どのシステムにおいても、言語処理は必要であり、用いる形態素解析など基本的な処理は同じである。しかし、先の漠然とした要求に応えるためには、これらの処理結果を、どのようにそこに結びつけるのかという点が全く違うである。このように考えると、表 4.1 にあげた各サブタスクは、それぞれ全く違う研究要素であるといえる。

次に、各システムの入出力や評価について、類似点、相違点を述べる。但し、ここでの入力は、解析対象となる素材データは除いている。QA システムと、クイズ生成システムでは、両者とも「質問」と「解答」を扱う点で似ているが、ここには大きな違いがある。それは、QA システムは質問が与えられるのに対して、クイズ生成では質問も解答も両方生成する必要がある、これは QA にはない新しいタスクと言える。また、解答については、QA システムでは正解だけを用意すれば良い

が、クイズ生成では、正解のほか間違えやすい不正解も用意する必要があり、この点も新規といえる。更に、評価について考えると、QA システムは正解が明確であり、質問に対して必ず答えなければならない。一方、要約システムやクイズ生成システムでは絶対的な正解はなく、特にクイズ生成では、ニュースの1項目から必ず1つ生成するという必要はなく、たくさんの中からいくつか生成でき、それらがクイズになっていけばよい点も異なる点である。このように比較してみると、研究要素としても、システムとしても、クイズ生成というタスクは、これまでにない新しい研究要素を含んでいることがわかる。

4.4 提案手法

本手法で生成対象とする画像付き選択クイズは、画像1枚と複数の文からなる選択肢で構成される。選択肢内の1つの文は、画像を説明した文であり、これを正解文と呼ぶ。それ以外の文は偽り文と呼び、正解文と似て非なる関係にある。このクイズの生成については、これまで [91, 92] で報告した手法が存在するが、これらの手法で自動生成したクイズ候補は、画像と選択肢間の対応がとりにくい傾向にあり、クイズ生成の精度はあまりよくないものであった。これは、画像により表現しているものと選択肢で述べられているものが明確でないことが大きな原因と考え、本論文ではこの点を考慮し、図 4.2 に示すように、3つのサブタスクに分割して行うこととした。以降、各サブタスクをより具体的な工学的タスクに置き換え、その解決手法について述べる。

4.4.1 クイズに適した画像の選択

クイズに適した画像として、背景が比較的無地であり、中央に何かが映っている画像(以後、スポットライト画像と呼ぶ)を選択することとした。これは、選択クイズは、画像と選択肢間の対応関係を用いて解くため、解答者が画像をみて、その主被写体を容易に判断できることが望ましいからである。図 4.3 は、スポットライト画像の例であり、このような映像はニュース内に良く現れる。これは視聴者の意識を主被写体に向けるためと考えられる。(a) は、ある機密文書が暗い部屋の中におかれ、そこにスポットライトがあたっている映像である。それ以外は実際にはスポットライトがあたっているわけではないが、同等の演出効果がある画像と考えられる。

スポットライト画像を選択する手法を図 4.4 に示す。はじめに、画像をサブブロックに分割し、各ブロックの無地の度合いを数値化する(図 4.5 参照)。その数

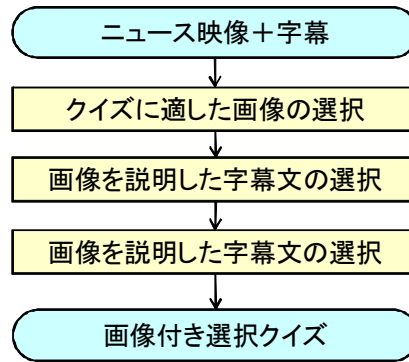


図 4.2: 画像選択クイズの生成手法



図 4.3: スポットライト画像の例

値を左上から右下まで順にならべたもの(以降, 無地度ベクトルと呼ぶ)を作成し, このベクトルを特徴とした学習を用いて, スポットライト画像を判別する. アナウンサのバーストショット(以降, アナバスと略す)は, ニュースの冒頭によく出現ショットであるが, 内容とは関係ないため, スポットライト画像中に含まれていれば除去する. 以上の処理により出力された画像を, クイズに適したスポットライト画像とする.

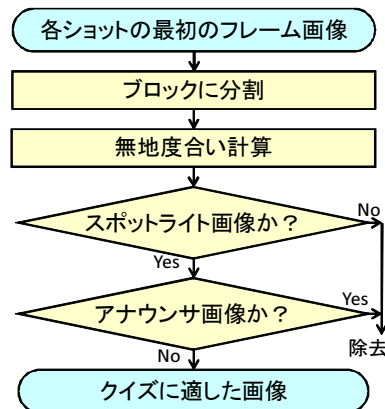


図 4.4: スポットライト画像の選択手法

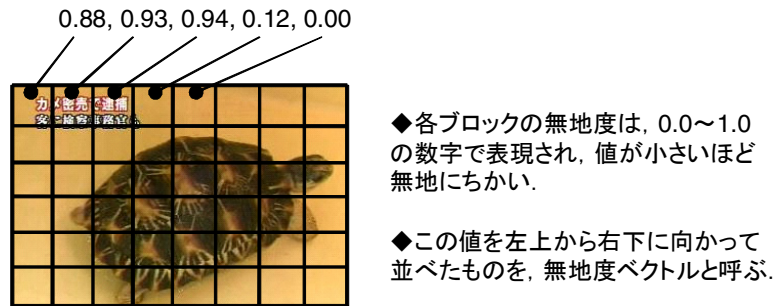


図 4.5: 無地度の例

4.4.2 選択肢の生成

画像を説明した字幕文の選択

クイズに適した画像1枚1枚について、ニュースの字幕文を利用して選択肢を生成する。正解文の選択は、その画像を含むニュース項目 (EvenTank における SourceItem に相当) 内の全字幕文から、画像と字幕文の両特徴を用いて選択する。図 4.6 に、その選択手順を示す。画像中に顔が検出された場合とそうでない場合で字幕文の選択方法を分け、前者には必ず人名を含む字幕文を選択し、後者にはこの制約を加えずに選択する。正解文は、画像が表示されている区間に発話されている字幕文を探す。これは、字幕文はその性質上、発話されているときに表示されている画像と何らかの関係があるからである。存在しない場合は、画像の表示時刻を起点に、前後の字幕文について、特徴的言い回しとして「このように」や「これが」という文節を含む字幕文があれば選択する。これらは、アナウンサーが視聴者に映像をみるよう促す場合に用いられるもので、字幕と画像の間に、より密接な関係があることを示すからである。例えば、「仕組みは、このようになっています」や「これが問題のゴミです」というような文章である。ここまでの処理で該当する字幕文がない場合は、項目の先頭から探す。理由は、項目冒頭では、その項目の簡潔な説明 (イントロ) が述べられることが多く、この部分は項目内の全ての画像に関連があるからである。以上の処理により、正解文が生成される。

似て非なる選択肢の生成

似て非なる選択肢の生成手法を図 4.7 に示す。ここでは、選択肢間の似て非なる関係の抽出が鍵であり、これは各選択肢文の主被写体と文全体の2つに注目して処理を進める。

はじめに、正解文から主被写体を特定する。正解文の中に人名が含まれていれ

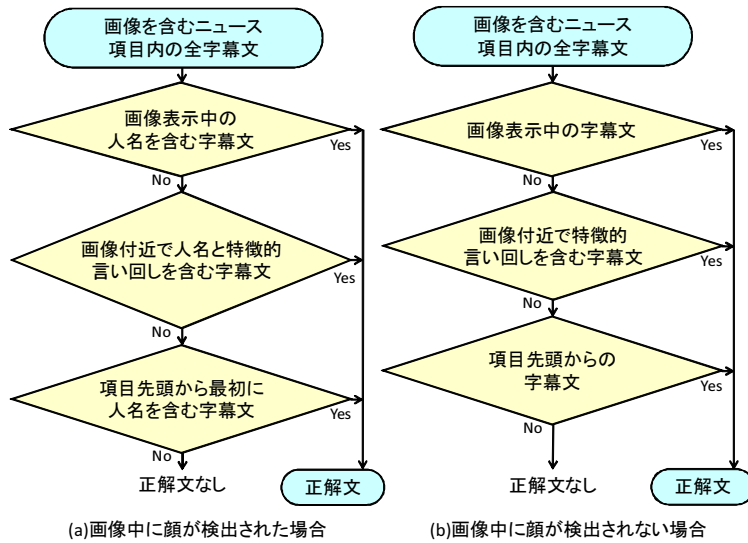


図 4.6: 画像を説明している字幕文の選択手法

ば主被写体は人物とし、含まれていなければ、文中の具象物 (具体的に目に見えるもの) 名詞の中で、最も IDF (Inverse Document Frequency) 値が高い単語を主被写体と特定する。IDF 値が高い単語とは、ある特定の文書だけに頻出する単語であり、一般的にその文章を特徴づけるものとされている。なお、主被写体の抽出方法として、固有表現の「固有物名 (Artifact)」を用いることも考えられるが、画面内の主被写体が必ずしも固有物名を持つとは限らないため、本手法ではこれらを内包するより幅の広い「具象物名詞」という枠で抽出を試みる。

次に、主被写体に注目した似て非なる偽り文の選択を行う。この偽り文は、指定された期間中の全ての項目内の字幕文を対象として、次の手順で選択する。主被写体が人物である場合は、「似て非なる」の定義を、同じジャンルの人物であるが、別の人物であることとする。例えば、同じスポーツ選手であるが違う人物や、同じ政治家であるが違う人物である。ジャンルについては、項目毎にあらかじめスポーツ、国家関連 (政治、国防、皇室など)、その他、の3つのいずれかを人手で付与した。一方、主被写体が具象物の場合は、まず、指定された期間中の全ての字幕文について、同様に各文から主被写体を抽出し、その主被写体の概念が「似て非なる」関係の文を複数選択する。この場合の似て非なる定義は、語を意味によって分類・整理したシソーラス (類義語集) である分類語意表 [93] を用いて判定する。例えば、「イタリア」は、分類語彙表の中で固有地名に分類されており、同じ分類項目内で違う単語である、「スペイン」や「ポーランド」などが類似に該当する。以上の手順により、主被写体が人物である場合、ない場合、それぞれ偽り文を複数抽出する。

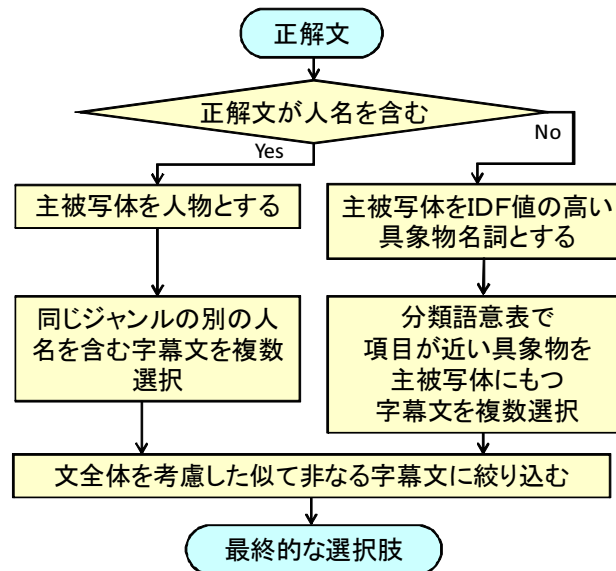


図 4.7: 似て非なる選択肢の生成手法

最後に文全体に注目して、複数選択された偽り文を、更に似て非なる関係の文に絞り、最終的な選択肢を生成する。文に注目した似て非なるの定義は、出現する単語が似ているが、話題が違っていると定義した。ここでの話題とは、各ニュース項目が扱う1つ1つの事件やイベントに相当する。従って、同じ車の事故のニュースでも、大阪で起きたものと東京で起きたものは違う話題として扱う。具体的には、出現する単語の共起パターンが異なれば、話題も異なるという仮説をたて、図 4.8 に示す手順で実現した。まず処理 (1) では、正解文と偽り文の全ての文からキーワードとして名詞を抽出する。抽出する名詞は、下位分類が一般や固有名詞のもので、代名詞や数、非自立、接尾、接続詞的といったものは抽出しない。処理 (2) では、各名詞の出現頻度を調べ、処理 (3) で、2 回以上出現したキーワードを要素とするベクトル (そのキーワードが出現した回数) を各文から生成する。処理 (4) では、正解文に含まれるキーワードを1つも含まない文 (図では偽 3) を、偽り文候補から除外することで、出現する単語が似ている偽り文だけを残す。処理 (5) では、正解文と、必要な数の偽り文全ての組合せに対して、ベクトル距離の総和を求め、それが最大となる組合せで選択肢を形成する。距離が大きいということは、単語の共起パターンが最も異なることを意味し、すなわち話題が異なるものと考えられる。以上の手順で、似て非なる選択肢を生成する。

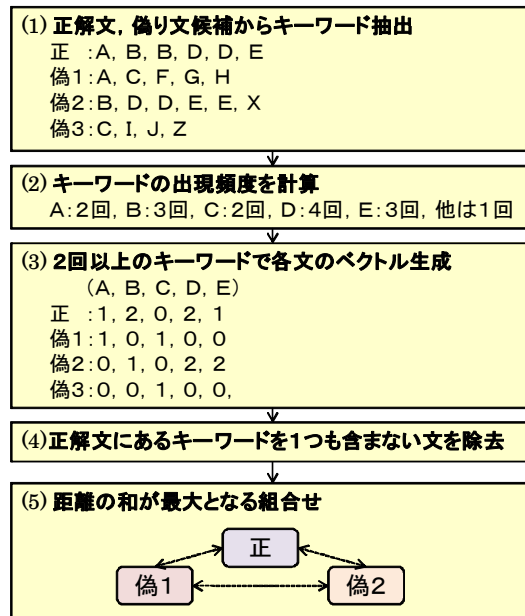


図 4.8: 似て非なる文の選択手法

4.5 実験結果

4.5.1 クイズに適した画像選択の結果

実験は、NHKの「ニュース7」、2005年1年分を用いて行った。素材画像としては、番組内の各ショットの最初のフレーム78,325枚を用意した。最初のフレームを用いた理由は、ニュース番組のショット内では、あまり構図が変化しないからである。以降、図4.4の手順に従って説明する。

はじめに、各画像(320×208画素)を、16×16画素のブロック260個に分割し、各ブロックで4.4.1で定義した無地度を計算した。無地度の計算には、LIBSVM[94]を利用し、カーネル関数にはRBF(Radial Basis Function)を用いた。あらかじめ2004年12月よりランダムに選択した45枚の画像から、無地ブロック6,080個、非無地ブロック5,620個をラベル付けし学習を行った。無地度は、LIBSVMが出力する、その評価対象のブロックが、無地ブロックのクラスに属する確率を利用した。用いた特徴は、ブロックにおけるDCT係数のAC成分のみ(255次元のベクトル)である。これは、DC成分はブロックの平均輝度値を指し示すだけであり、実際の無地度はAC成分の含有量によって決まるからである。

スポットライト画像の判定にも、同じLIBSVMを用いた。学習は、2004年12月の画像5,791枚について、スポットライト画像249枚、非スポットライト画像5,542枚をラベル付けし行った。特徴は、先に計算した各ブロックの無地度を、画像左

上から右下に順にならべた無地度ベクトルを用いた。これにはレイアウト情報が含まれるため、画像のどのあたりに、どの程度の無地が含まれているかを示す情報に相当する。

テストデータ 78,325 枚を用いて実験を行ったところ、スポットライト画像として 3,457 枚が選択された。あらかじめ複数の人手によりスポットライト画像かそれ以外かをラベル付けしたものでこの結果を評価すると、再現率は 66.3%、適合率は 63.8% となった。検出漏れや過剰検出された画像を調べると、検出されたスポットライト画像とあまり変わらないものであった。これは、判別器のための学習データ、また評価用の正解データを生成する際の主観評価にバラツキがあったのではないかと考える。特に、背景の部分が比較的無地の画像を正解とする場合に、「比較的無地」と判断する基準が作業者によって異なっている事例を確認した。

最後のアナバス判定にも同じ LIBSVM を用いた。スポットライト画像と判別された 3,457 枚の画像に、アナバス (1,027 枚) か、それ以外 (2,430 枚) というラベルを付与し、これを 2 分割の交差法で判断した。用いた特徴は、各ブロック内の RGB の平均値を、画像左上から右下まで順に並べた、色の分布情報である。アナバスの判別結果は、再現率が 98.9%、適合率が 99.5% であり、高い精度で除去することができた。最終的に、クイズに適した画像として 2,436 枚が得られた。

4.5.2 選択肢生成の結果

画像を説明した字幕文の選択

選択肢生成の実験は、スポットライト画像として得られた 2,436 枚から、ランダムに 250 枚 (約 10 分の 1) を選択して行った。図 4.6 に示すように、正解文の選択は、画像中に顔が検出されたかどうかで処理が異なる。画像内の顔検出は、OpenCV[65] を用いた。実際の人物の有無に対する結果は、再現率が 81.5%、適合率は 98.0% であり、本手法の実現に足ると考える。次に条件にあう字幕文を探す際に、人名などの判断には、南瓜 [95] の固有表現抽出を利用した。表 4.2 は、正解文の選択結果を示しており、画像 250 枚に対し、212 枚について正解文が選択できた。正解文が選択できないケースは、画像に人物が映っていると判断されたにもかかわらず、人名を含む字幕文がその項目の中に 1 つもない場合である。

この正解文が画像を説明した文になっているかを調べると、99 件 (46.7%) で正しい対応がとれていた。なお、この評価は、図 4.7 で説明した主被写体が画像中に確認できることを判断基準とした。表 4.2 にも示すように、人物については約半分以上が対応していたが、人物以外の場合は 27.9% しか対応していなかった。

表 4.2: 正解文選択の評価

	合計数	人物	人物以外
正解文選択成功	212	151	61
正解文が画像と対応	99 (46.7)	82 (54.3)	17 (27.9)

() 内は割合%

表 4.3: 似て非なる選択肢生成の評価

	合計数	主被写体別	
		人物	具象物
選択肢生成成功	193	138	55
似て非なる成功	134 (69.4)	108 (78.3)	26 (47.3)

() 内は割合%

似て非なる選択肢の生成

正解文が選択できた212枚の画像について、図4.7に示したように、抽出した主被写体(人物か具象物)に注目し、選択肢生成(本実験では、選択肢の数を3とした)を行った結果を表4.3に示す。全部で193件について選択肢が生成できていた。生成に失敗した理由は、人名も具象物も取り出せなかった場合と、図4.8の処理4により、正解文に似ていない文を除去したときに、偽り文が全てなくなるもしくは1つになってしまう場合であった。選択肢が生成できた193件の中で、似て非なる関係が認められたものは134件であり、人物については78.3%、具象物については47.3%の割合だった。人物の場合で、似て非なる関係が得られなかった原因の1つは、固有表現抽出による人名の判断を誤った場合である。また、具象物の場合の失敗は、主に次の2パターンであった。1つ目は、複数の選択肢が同じ話題を含む場合で、本手法では話題を厳密に区別できないことに起因する。もう1つは、3つの選択肢の各話題に、全く共通する部分がなく、互いに「似ている」という関係がない場合であり、主被写体と特定した具象物が話題に直接関係ないことが原因と考える。

4.5.3 生成されたクイズの評価

表4.4は、生成された193件のクイズ候補について、人手により画像付き選択クイズとして有効かを評価した結果である。はじめに、選択した正解文が画像と対

表 4.4: 生成したクイズの有効性の評価

	合計数	主被写体別	
		人物	具象物
クイズとしての出力数	193 (100)	138 (71.5)	55 (28.5)
正解文・選択肢がっている	82 (42.5)	70 (36.3)	12 (6.2)
上段の内クイズになっている	74 (38.3)	70 (36.2)	4 (2.1)
最終的にクイズになっていた	78 (40.4)	71 (51.4)	7 (12.7)

() 内は生成されたクイズ 193 件に対する割合%

応がとれており、かつ、生成した選択肢同士が似て非なる関係にあるクイズを数えると、人物が70件、具象物が12件あった。更に、これらについて、画像付き選択クイズを評価すると、人物については70件全てが有効であったが、具象物の場合は4件のみが有効であった。クイズとして有効か否かを判断する基準は、選択文間に類似した関係がみられ、かつ画像と対応する選択文が1つだけの場合と定義した。人物の場合の一例を挙げれば、画像中の人物の名前が1つの選択肢のみに含まれており、他の選択肢には同職業の他の人名が含まれている場合などである。主被写体が具象物の場合も同じ基準ではあるが、表 4.4 に示すように、クイズとして有効でない場合が多かった。これは、選択肢の似て非なるという関係の、似ている部分が映し出されており、対応が1つに絞れないものである。図 4.9 にその例を示すが、選択肢それぞれは、ノルウェーの国について、ノルウェーの特定の人物、ノルウェーのレスキューチームについて記述しているが、画像にはノルウェーの国旗が映っており、どの選択肢とも対応が取れてしまうような場合である。表 4.4 の最下段は、正解文の対応が不明確ではあったが、選択クイズとして成功していると考えられる場合(図 4.12)を加えた結果である。全部で78件のクイズが生成されており、自動生成したクイズの約4割が、有効なクイズになっていたことを示す。人物に関するものは、自動生成した約半分が成功しているが、具象物については12.7%と、生成が難しいことがわかる。

次に、同じ基準で、一般人5人によりクイズの有効性の評価を行った。ただし、可能な限り通常感覚で解答してもらうため、主被写体は明示せず、画像と文章との対応関係と、選択肢間の似て非なる関係にだけ注目する評価とした。表 4.5 に評価結果を示す。被験者5人中3人以上、ならびに4人以上がクイズとなっていると判断した割合は、それぞれ50.8%と33.7%であった。

また、いずれもそのうちの約8割が人物に関するクイズであった。これは、表 4.4 の最下段の評価と同じ傾向を示しており、本論文で用いたクイズの有効性判断の基準は、ある程度妥当であったと言える。また、主観的評価のバラツキを数



- A. 両陛下の両国への訪問は、20年ぶりで、特に ノルウェー とは外交関係樹立 100 周年になることから友好親善をいっそう深める訪問になります
- B. ノルウェー人 の 2 歳の男の子、ラグナーちゃん。
- C. ほぼ同じころ、ノルウェー の救助チームが別の建物で救助に当たっていました。

図 4.9: 主被写体が具象物の場合の失敗例

表 4.5: 一般人による生成したクイズの有効性の評価

	合計数	主被写体別	
		人物	具象物
クイズとしての出力数	193 (100)	138 (71.5)	55 (28.5)
5人中3人以上がクイズと判断	98 (50.8)	79 (57.2)	19 (34.5)
5人中4人以上がクイズと判断	65 (33.7)	50 (36.2)	5 (9.1)

() 内は生成されたクイズ 193 件に対する割合%

値で示す指標である 値を計算すると、この評価者 5 人の 値は 0.45 であり、評価結果は中程度一致していると言える。また、人物だけに限ると 値は 0.54、具象物だけに限ると 0.12 となっており、具象物の場合の評価のブレが大きいことがわかる。

ここでクイズとして成功した実例を示し説明をする。なお、ニュース番組では、そのニュース項目に関連する情報がスーパーインポーズされており、それが答えになっていることも多い。実例で示す画像にもこの情報が含まれているが、ここでは無視することとする。いずれこの部分は、自動もしくは人手によりマスクをかけるか、もしくは放送局の場合は編集前の画像を使用することで解決できると考える。

図 4.10 には、主被写体が人物として生成されたクイズ例を 4 つ示した。なお、これらは説明の都合上、全て選択肢 A が正解となっている。

クイズ 1 は、画像にはフィギュア選手が、選択肢にはそれぞれ違ったフィギュア選手の名前が含まれており、生成に成功した。解答にあたっては、画像には女性が映っており、選択肢 C は男子の選手名であることから、正解は選択肢 A か B のどちらかという絞り込み方ができる。これは、ある意味「ヒント」に相当するものであり、似て非なるの定義によりクイズの難易度を操作できることを意味している。例えば、似て非なる関係の類似が「フィギュア選手」でなく、「女性のフィ

ギョア選手」であれば、より難しいクイズとなる。

クイズ2についても、画像には政治家が、選択肢にもそれぞれ違う政治家の名前が含まれており成功した例である。このクイズのおもしろいところは、画像には後ろ姿しか映っていないため、通常では誰なのか判断することは難しい。しかし、別の特徴(この場合は髪型)から、特定の人物と対応付けができるため、クイズとなった例である。これは、遊び心があるクイズ制作者が、答えが明らかではなく、先に述べたヒントを利用することで解けるようなクイズを生成した場合に相当する。今回は偶然ではあるが、用いた表層的な解析技術の曖昧さ(この場合は、画像に何となく何かが中央に映っていて、字幕文から人物と特定したこと)が、結果的に良い方向に働いたといえる。

クイズ3についても、各選択肢に登場する人物が全て異なる政治家でありクイズとして有効と判断した。偽り文を選択する際に、正解文の人名がカタカナであることを考慮し、Bの選択肢にもカタカナの名前を含む字幕を選択すれば、更に難しいクイズとなり得る。このように、似て非なる関係を形成する類似部分と異なる部分には、様々なレベルが存在し、これらを調整することで選択クイズの難易度が調整できることがわかる。

クイズ4については、選択肢には日本の伝統文化に関わるそれぞれ違う人物の名前が含まれている。このクイズのジャンルは「スポーツ、国家関連(政治、国防、皇室など)、その他」の3つの中で「その他」にあたり、このジャンルには様々なニュース項目が含まれているが、全ての選択肢において伝統文化に関わる人物を含んでおり、似て非なる関係を抽出する4.4.2で提案した手法が有効であった結果と言える。

図4.11には、主被写体が具象物として生成されたクイズを4つを示した。これらも説明の都合上、全て選択肢Aが正解となっている。

クイズ5は、具象物クイズとしての成功例である。具象物としては「素材」という単語が抽出されている。3つの選択肢はそれぞれ、サッカーユニフォームの素材、電車の車両素材、スペースシャトルの断熱素材についてのものであり、素材という点で似ているが、それぞれ別物である。画像は、素材のある部分をクローズアップしているため、3つの選択肢のどれにも当てはまるように感じられ、クイズになっているといえる。実際は、この画像の被写体は青色をしており、選択肢Aの「着心地」とこの「青色」で、勘の良い解答者であれば、「サッカー日本代表のユニフォーム」との関係を見いだすことができる。

クイズ6は、主被写体が「動物」であった場合の生成例である。3つの選択肢は、それぞれ異なる動物園名を含んでおり、画像に現れているアカエリマキキツネザ

ルは選択肢 A にしか含まれないため、クイズとして有効と判断した。

クイズ7は、抽出された主被写体が「火山」であり、画像にも火山の噴火の瞬間が映っている。選択肢は、それぞれコリマ火山、海底火山、シエラネグラ火山と、どれも火山であるが異なるものが含まれていた。

クイズ8の生成例は、主被写体が「ロボット」の場合であり、画像には海底調査用のシャチホコのロボットが映っている。それぞれの選択肢は、ロボットに関する異なる説明をしておりクイズとして有効と判断した。一方で、選択肢 A だけに「シャチホコ」という単語が現れているため、画像と選択肢の対応付けが、少し簡単すぎるとも考えられる。この場合、主被写体が「シャチホコ」であれば、もう少しクイズの難易度が上がったかもしれない。

最後に、画像と選択肢間の明確な対応がとれないが、クイズとなっていると判断した2例を図 4.12 に示した。いずれも正解選択肢は A となっている。

クイズ9は、正解文が明確に画像と対応しているとはいえないが、クイズになっていると判断したケースである。画像には、何かわからないものが映っており、選択肢をみると、A は「今人気のもの」、B は「尾瀬」、C は「打ち水」を説明している。確かに、「尾瀬」や「打ち水」でもなさそうであるが、クイズという見方をすると、この画像中のものが「尾瀬」や「打ち水」と関係がないとは断言できない。もしかすると、「尾瀬に出現した新種の何か?」、「打ち水の効果を更に高める新素材?」かもしれないと解答者は考えるからである。実際は、イルカの耳の骨であり、アクセサリとして今人気を呼んでいるというニュースの断片であった。これは興味をかき立てているようなクイズの例になっているといえる。

クイズ10についても、正解文と画像の明確な対応をとることのできる組合せとはなっていないものである。画像には、絵画が映っており、選択肢は、A は「巨大なスクリーンに映し出されたハイビジョン映像」、B は「世界初の映像システム」、C は「ゲーム機」を説明している。正解を知ること初めて画像中の絵画は、ハイビジョンの16倍というきめ細かい映像であることがわかるものである。これらの選択肢などは、どれも興味が沸くような話題を含んでおり、クイズを解いた後に、実際に「世界初の映像システムとはどんなものなのか?」など、そのニュース項目にリンクを張って視聴できるようにすることで、新しいサービスとしての展開も可能と考える。このようなものは数は少ないが、数件あった。

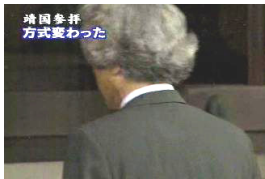
クイズ1 (人物)



(注) 権利関係で映像を加工しています

- A. 女子シングルのフリーが行われ、前半2位の中野友加里選手が初出場で初優勝を決めました。
- B. フィギュアスケート、グランプリシリーズの第2戦、スケートカナダは女子シングルのショートプログラムが行われ、村主章枝選手が2位でスタートしました。
- C. 男子シングルの後半のフリーが行われ、18歳の織田信成選手が逆転で初優勝しました。

クイズ2 (人物)



- A. また、内閣総理大臣、小泉純一郎と記帳をしていたのも今回は行いませんでした。
- B. 佐々江局長は、今月17日にも予定されている、町村外務大臣と中国の李肇星外相との会談で、事態打開のための話し合いを行えるよう、中国側に強く促したいという立場を示しました。
- C. 協議は、麻生総務大臣や竹中経済財政郵政民営化担当大臣ら関係閣僚が出席して午後5時から始まりました。

クイズ3 (人物)



- A. アナン事務総長は日本など4カ国に対して幅広い意見に配慮するよう促しました。
- B. 岡田代表はこのように述べて、法案を廃案にするため自党内の反対派は、衆議院会議の採決で反対票を投じるよう促しました。
- C. アメリカは採択を考え直すよう呼びかけていただけに、今月下旬、北京を訪れるアメリカのライス国務長官と中国の指導者との会談でこの問題が議論されるのは確実です。

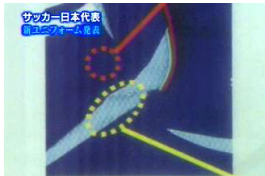
クイズ4 (人物)



- A. 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。
- B. 桂文枝本名長谷川多持さんは、昭和5年大阪市に生まれ章を22年に、仙台の桂文枝に弟子入りして初舞台を踏みました。
- C. 色鮮やかな人形が演じる人形浄瑠璃文楽その部隊でストーリーを語る太夫の竹本住大夫さん、人間国宝の竹本さんは、情感豊かな人物描写で活躍しテレビ番組にも数多く出演、文楽の普及拡大に貢献してきました。

図 4.10: 生成された人物に関するクイズの例

クイズ5 (具象物=素材)



A. その上で、従来のものより軽く、伸びたり縮んだりする 素材 を使い、通気性をよくして着心地が悪くならないよう、新たなくふうが加えられたということです。

B. また、今回の事故で車両がここまで大きく壊れたことについて電車の車両 素材 の強度や構造について詳しい立命館大学の坂根政男教授は次のように話しています。

C. 機体の耐熱タイルの傷が少なくとも26か所、また、操縦室の窓の近くの表面の傷やタイルの間から断熱 素材 が飛び出している箇所、翼の前の部分の損傷などもありました。

クイズ6 (具象物=動物)



A. アカエリマキキツネザルは、アフリカのマダガスカル島にしか生息していない、珍しいサルで、日本では横浜市の、野毛山動物公園 で飼育されているだけです。

B. 旭山動物園 は日本で最も北にある動物園で、動物の生態を生かしたユニークな施設で知られています。

C. 全国各地で2本足で立ち上がるレッサーパンダが話題になっていますが、よこはま動物園ズーラシアのデールは、さらに2本足のまま歩くことで人気です。

クイズ7 (具象物=火山)



A. メキシコの、コリマ火山 が爆発した瞬間です。

B. 気象庁によりますと現場付近にある福得岡ノ場と呼ばれる、海底火山 が噴火したとみられるということで分析を急いでいます。

C. 噴火したのは、標高およそ1500mの、シエラネグラ火山、22日、日本時間のきのう朝噴火しました。

クイズ8 (具象物=ロボット)



A. 泳ぐ金のシャチホコ、海底探査用に開発された ロボット です。

B. 今月25日に愛知県で開幕する愛・地球博、会場で活躍するさまざまな ロボット が、きょう公開されました。

C. また、近未来をイメージした、1人乗りの電気自動車 i - u n i t、人を乗せて2本足で歩行できる、i - f o o t と呼ばれる ロボット によるダンスショーも公開されます。

図 4.11: 生成された具象物に関するクイズの例

クイズ9 (具象物=都会)



- A. 実は今、都会の若者などに人気なんです。
- B. 通りかかった買い物客などが、都会に出現した小さな尾瀬を楽しんでいました。
- C. 打ち水で都会の暑さを和らげようという試みが東京浅草で行われました。

クイズ10 (具象物=スクリーン)



- A. 今のハイビジョンの16倍というきめ細かい映像が横7m80cmの巨大なスクリーンに映し出されます。
- B. この部屋、内側がすべてスクリーンになっている世界初の映像システムを導入しました。
- C. ソニーは、画像の処理速度を従来の35倍に高めたゲーム機を公開し、大型スクリーンで、なめらかな動きをアピールしていました。

図 4.12: 画像と選択肢の明確な対応はできないがクイズと見なした例

4.6 考察と評価

クイズ生成のために具体的に設定した3つのサブタスクと、それらの組合せによる生成手法について考察する。はじめに、クイズに適した画像選択手法について考察する。本手法では、スポットライト画像と表現した、比較的無地の背景の中央に何か映っている画像に注目した。結果的に人物に関するものが多かったが、いくつものクイズが生成できており、主被写体がわかりやすいスポットライト画像が、クイズに適した画像であるという仮定は妥当だったと言える。しかし、これ以外にもクイズに適した画像は存在する。例えば同様に中央に何か映っているが、背景が無地でないものもある。このような場合を考えると、画像中の主被写体の切り出し、言い換えるとセグメンテーションが正確にできるようになると、クイズ生成の可能性が広がる。また、ある特定の物体が主題ではなく、例えば「祭り」のように、多くの人と特徴ある儀式などにより象徴される画像もクイズとして考えることができる。このような場合には、もはや画像だけで解析することは難しく、付随する字幕文など他の特徴を利用して解かなければならない。

次に、画像を説明している字幕文の選択手法について考察する。本手法では、画像の被写体情報として、人物であるかどうかを利用した。人物の場合は、字幕文選択の際に人名に注目して選択することで、半分以上の場合について対応がとれた。対応がとれなかった場合は、画像中の人物がその項目の話題に直接関係なかった場合である。例えば、街角インタビューであったり、人物を象徴的に映している場合である。従って、その話題に直接関係あるか区別をすることも、今後の課題である。一方、具象物については、話題に関係なく IDF 値など統計的手法で決めため精度が低くなった。画像に映っている被写体の情報を、カテゴリに分類する程度でも、クイズ生成の精度は上がると考えられ、今後、画像と言語処理の両面から、解決したい問題である。

似て非なる選択肢の生成手法については、人物の場合は有効であったが、具象物の場合は再検討が必要と考える。選択肢の生成という点では、具象物の場合は画像と文章(主被写体)との対応が正しく抽出できなかったことが問題の根底にある。しかしながら、仮に対応が正しく抽出できたとしても、次のような問題点もあることがわかった。それは、最適な似て非なる関係を考慮できなかったことである。これは、似て非なる関係を抽出する際に、類似点をどの程度広げるのかという問題である。例えば、「さる」の似て非なる関係は、類似点を「生き物」と考えたときは、「人」「ライオン」「マグロ」「鶏」「細菌」など多種にわたるが、「ほ乳類」として考えれば、「人」や「ライオン」だけになる。どちらが良いのかは、そのクイズのコンテキストに依存する。これについては、画像と言語に加え、ニュース番

組という対象データの特徴を用いて解決すべきと考える。

最後に、組合せによる生成手法について考察すると、「画像」と「選択肢の似て非なる関係」の間関係を考慮すべきだった。これは、似て非なるの「似ている」というものが画像に現れていると、クイズにはならない。似て非なるの「非なる」部分が画像に映っている必要がある。人物の場合は、被写体が特定されたことで、この部分が明確に手順化されていたといえる。この観点からも、画像の主被写体の特定という技術は重要であり、ニュース番組という特徴、言い換えると映像の修辞構造に注目することも1つの手法と考える。

4.7 まとめ

本章では、EvenTankを基盤とし、ニュース映像を活用した新しいコンテンツ生成として、画像付き選択クイズの生成手法について提案した。クイズに用いる画像については、3.1節において提案した映像の役割に注目し、その中の「主被写体」を伝えるという役割を持つショットを利用した。この主被写体が明確な画像を、改めてスポットライト画像と定義し、画像の特徴を用いて抽出を行い、クイズ生成に用いた。クイズ生成そのものについては、自動生成した画像付きクイズ候補のうち、約4割が実際に有効なクイズになっていることを確認した。また、本実験を通し、主被写体を伝える画像には、スポットライト画像ではないものも存在することがわかり、そのような画像についても考慮すれば、更にクイズの幅が広がることがわかった。また、生成システムは、EvenTankが提供する枠組みの中で実装できており、その適合性、有用性についても同時に確認をした。

第5章 インタラクティブ・サマリの生成

5.1 はじめに

本章では、EvenTank の枠組みを利用し、蓄積されたニュースの中のあるトピックに関するサマリを生成し、それをインタラクティブに閲覧するシステムを構築する。はじめに、生成対象とするサマリについて説明し、その工学的アプローチを提案する。最後に、実際のニュース番組を用いた生成実験を行い、提案したアプローチの評価を行うとともに、EvenTank の有用性について確認する。

近年、IT 技術の発達とインフラ整備に伴い、次々と新しいサービスが提供されはじめている。前章でも述べたように、今や放送局においても、従来の番組だけでなく、様々な関連コンテンツを制作する状況にあり、これら制作作業の効率化を模索しているのが現状である。このような中、一度は放送されて蓄積されている番組を再利用し、別の形のサービスとして提供するための研究が進んでいる。本章における提案も、この1つとして位置づけられ、蓄積されたニュース番組からインタラクティブ・サマリの生成を試みる。これが実現できれば、放送局では新しいサービスが低コストで実現できることになり、また一般の人々は、世の中の出来事について、現在より更に関連性などに着目しながら理解をすることができるようになる可能性がある。また、改めてトピック全体を見直すことで、新しい発見をしたり、新しい疑問を持ち、更に調べるなど、その利用についての発展性は大きいと考える。

本研究では、可能な限り全自動化を目指す。従来技術により精度良く処理可能な部分については、人手による正確なデータを用い、アルゴリズム自体の有効性について検証を行う。将来的には、序論で述べたような QA 機能や、知識オントロジなどと連携することで、様々なサービス展開の核となる「ニュースの映像付き辞書」の自動構築を目指しており、本研究はその第一歩と位置づける。

5.2 サマリ生成について

サマリには様々な種類が存在し、対象とするものにより生成のアプローチは異なる。例えば、テレビ番組でもドラマそのものを要約するものから、ドラマを制

作する際に撮影する素材映像，一般的にはラッシュと呼ばれるものを対象にするものまでその幅は広い．後者については，TRECVID などにおいて，1つのタスクとして設定されており，例えば [96] のように繰り返し撮影された部分を検出・除去することで，サマリを生成する手法が提案されている．本研究は，放送したニュース番組を素材とするものであり，そこから映像/画像付きのインタラクティブ・サマリの生成を行う．

ニュース番組で報道される内容には，長期にわたるものがあり，そのような話題は，時間が経つと，その全体像や，内容の遷移を把握することが難しい．また，短期間に報道される話題であっても，その内容が多岐に渡っている場合も，全体像を把握することが難しい．放送局側は，このようなことを考慮に入れ，可能な限り関連性をもたせてニュースの制作を行っているが，限られた時間内でこれを行いきれないのが現状である．一方，受け手の我々について考察してみると，普段はニュース番組をおおざっぱに視聴していると言える．単純に1つ1つを見ているだけでは気が付かないが，蓄積したニュースから関連項目を抽出し，改めて連続して視聴すると，それまで気づかなかった事実関係が明らかになってくることがある．

本章で紹介するインタラクティブ・サマリは，言い換えれば，新聞スクラップ帳の映像メディア版を目指すものとも言える．ここでは，まずその出発点として，ある観点で集められたニュース項目を入力とし，そこから各項目を表現するメディアを抽出し，それを組み合わせることで，トピック全体の流れも把握できるようにすることが目的である．

5.3 提案手法

インタラクティブ・サマリの生成手法の概略を図 5.1 に示す．入力は，ある視点に基づいて集められたニュース項目群であり，これをトピックデータ，または単にトピックと呼ぶことにする．このトピックデータは，人手により収集される場合もあれば，自動で生成される場合も考えられる．最も単純な自動生成例としては，あるキーワードをクエリとして入力し，該当する言葉を含むニュース項目全てを用いるものである．入力されたトピックデータは次に，1つ1つの項目に注目し，その項目を代表する静止画群と文章，関連するキーワードや画像が抽出される．ここでは，抽出されたデータをまとめて，項目の代表データと呼ぶことにする．そして，各項目の代表データをキーに，3種類のグラフ構造をもつ表示データを形成する．また，あるノードからは，実際のニュースコンテンツへのリンクを張るこ

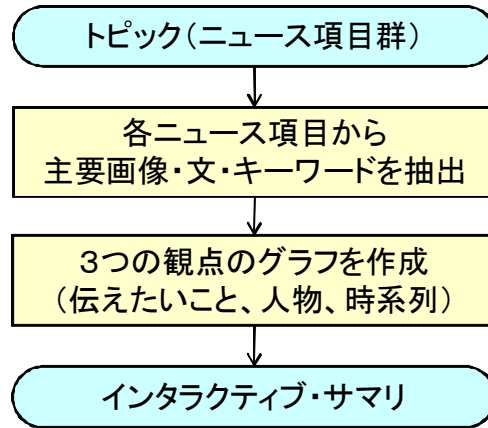


図 5.1: インタラクティブ・サマリの生成手法

とで、インタラクティブ性を付与する。以降では、項目の代表静止画群を抽出する手法、代表文やキーワードを字幕文より抽出する手法、インタラクティブ・サマリを構成する3種類のグラフ構造データの生成手法について述べる。

5.3.1 ニュース項目の代表静止画の抽出

ニュース項目を代表する静止画を抽出する手順を図 5.2 に示す。選択元となる静止画像については、各ショットの最初のフレームを切り出したものとした。その中から代表画像候補として、3.4.1 で抽出実験をした「主張・訴え」を伝えるモノローグショット、3.4.2 で導入した「導入・見せる」役割を持つサイレントショット、そして 4.4.1 で導入したスポットライト画像を抽出する。これら3種類の画像を抽出した後に、この中から1枚を項目の代表静止画として選択し、残りは「関連ショット」という位置づけでデータを保持する。代表静止画を選択する手順は、まずサイレントショットがあれば、サイレントショットで最も早く出現するものとする。サイレントショットがない場合は、モノローグショットの中で最も早く出現するもの、なければスポットライト画像で最も早く出現するものとする。ここまでの処理で該当する画像が見つからなかった場合は、項目の先頭に最も近いアナウンスのバーストショットではない画像を、代表静止画とする。

5.3.2 ニュース項目の代表文とキーワードの抽出

次に、ニュース項目内での発話から代表文を選択し、キーワードを抽出する手順について述べる。代表文の選択は、図 5.3 に示すように、項目内の全発話から放

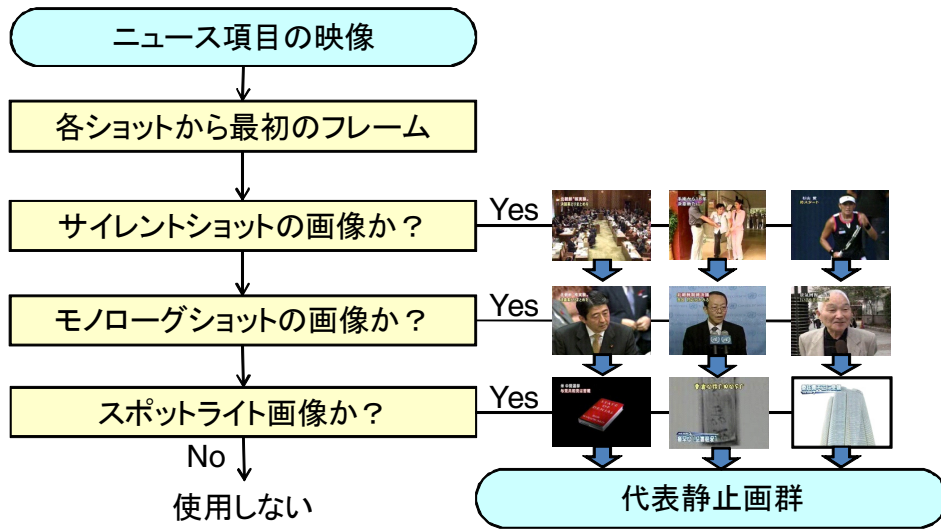


図 5.2: 代表静止画群の抽出手法

送局関係者の発話を抽出し，文末が「ました」「でした」「います」で終了している最初の文を代表文とする．ここでの放送局関係者とは，アナウンサ，記者，レポーターなどを指す．また，キーワードについては，代表文中に含まれている人物名を除く固有名詞とした．従って，組織名や，場所の名前などが抽出される．

5.3.3 3種類のインタラクティブサマリの生成手法

ここまでの処理で，各項目について，代表静止画と代表文，項目を表現するのに重要な画像とキーワードが抽出されている．これらについて3つの視点で情報を整理し，それをグラフ構造で表示し，更に各ノードから映像などへのリンクを張ることで，インタラクティブ・サマ리를生成する．グラフ構造で情報を視覚化するものとしては，意味ネットワークや概念地図，トピックマップ，マインドマップなどがある．これらは外見同じように見えるが，ノードとリンクの定義や利用の方法が異なっている．この中では，トピックマップだけが標準化されており，その交換用標準形式は，XML Topic Maps (XTM)[97]と呼ばれている．マインドマップ以外は，ノード間に自由にリンクを張ることができる傾向にあり，複雑な関係も表現できる反面，その視覚化には工夫を必要とする．一方，マインドマップは，表現したい概念の中心となるキーワードやイメージを図の中央に置き，そこから放射状に関連するアイデアなどを，キーワードや図を用いて広げていくものであり，表現の構造が制限される代わりにユーザにはわかりやすいという傾向がある．そこで，マインドマップの亜種である FreeMind[98]を利用して視覚化

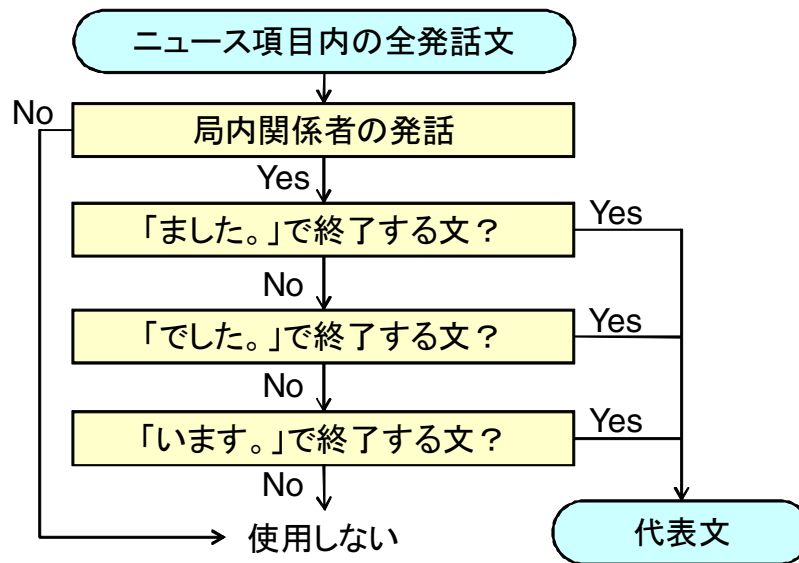


図 5.3: 代表文の抽出手法

する。どの視点のサマリにおいても，中心にはそのトピックを表す文字列のノードがあり，その周りに各視点をベースに作成された様々なノードが接続される。以下に3つの視点とそのグラフ構造のデータ生成について述べる。

伝えたいことサマリ

伝えたいことサマリは，各項目で伝えなかった主題を，1枚の画像と，数個の文字で表現して1つのノードとし，トピックの中心ノードの周りに接続していき，一覧性よく見せることを目的とするものである。従って，1つのニュース項目が1つのノードを形成し，その項目のノードに対し，更に様々な付属情報がやはりノードとしてぶら下がる構造をもっている。その作成手順を図 5.4 に示す。項目を代表

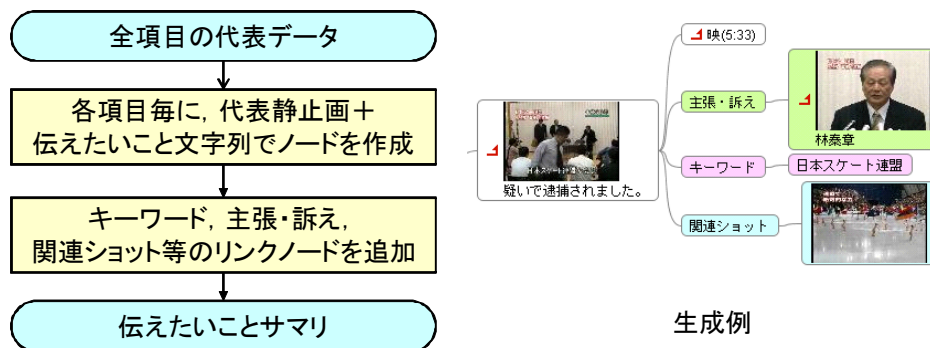


図 5.4: 伝えたいことサマリの生成

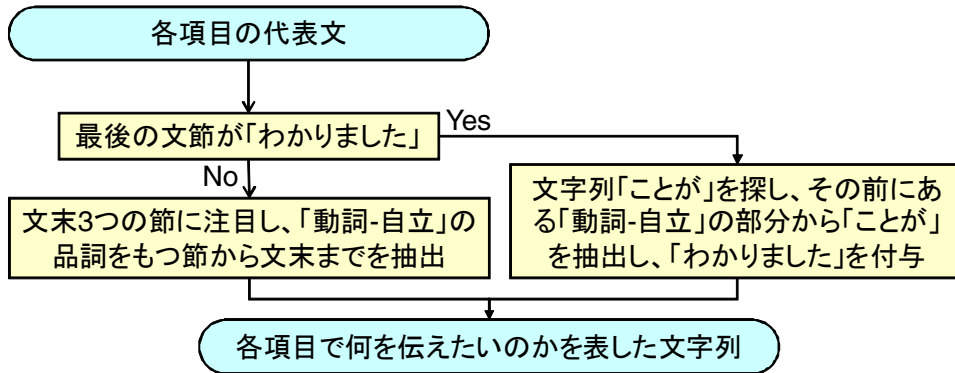


図 5.5: 「伝えたいこと文字列」の抽出手法

する1枚の画像については、既に抽出してある代表静止画を利用し、主題を表す数個の文字については、代表文の述部に注目して抽出する。また、項目全体を視聴できるようにしたリンクをはじめ、キーワード、関連画像等の付加情報も項目ノードの子ノードとして生成する。主題を表す数個の文字列の抽出手法は、図 5.5 のようになっている。まず、文末が「わかりました」の場合は、その前に出現する「ことが」という文字列を探す。そして、その文字列から更に前に「動詞-自立」の品詞を持つ部分を探し、そこから「ことが」までの文字列に、「わかりました」を付与し、主題を示す文字列とする。一方、文末が「わかりました」でない場合は、文末から文頭に向かい、同様に「動詞-自立」の品詞を持つ部分を探し、そこから文末までを抜き出して伝えたいことを示す文字列とする。

これらのルールは事前に、ニュース番組内のアナウンサのスタジオショットで発話された文章と、スタジオショットがないニュース項目については、最初のアナウンサの発話文を集め、その文末にどのような表現が多いのかを調査して決めた。表 5.1 は、その調査結果を示したもので、文末が「～た。」で終わる文が全体の 47.4% と約半分を占めている。次に多いのが「～です。」の 20.8%，次に「～ます。」の 14.4%，残りの 17.4% は、ほぼ名詞で終わる体言止めとなっている。更に、「～た。」の直前の文字列もあわせて調べると、その大半は「～ました。」であり、残りは「～でした。」となっていた。これらの前には動詞が接続し、「逮捕されました。」や「訪問しました。」というような文末になる。「次です。」や「スポーツです。」などは、ニュース番組内での項目の遷移を表す文章で、短く情報量も少ないことから、「～です。」で終わる文章については、項目の代表文を選択する際に取り除いた。「～ます。」の場合には、「～います。」が約 8 割であり、ほとんどが「～しています。」という、現在進行していることを説明している文章が多いため、代表文として選択する。

表 5.1: ニュース項目のイントロ文の文末解析

た .	です .	ます .	その他
47.4%	20.8%	14.4%	17.4%
ました . (97.0%) でした . (2.7%)		います . (77.0%)	[体言止め] (93.4%) [疑問文] (6.6%)

ここで、文末の述部に注目した理由についても述べておく。3.3.1でも述べたように、ニュース番組は、世界で起こっている出来事からいくつかを選択し、更に短い時間でその出来事を伝えるため、編集という一種のサマリ作業を行い、放送している。従って、アナウンサによるニュースの説明そのものも、一種のサマリの結果と考えられる。例えば、下に示す2つの文章は、練馬区で起きた事件についてのイントロと呼ばれる導入部分での発話文である。

- きょう昼過ぎ、東京・練馬区で 包丁を振り回していた男を取り押さえようとした警察官3人が 重軽傷を負いました。
- この際、警察官が 拳銃を発砲して男も怪我をし、殺人未遂などの疑いで 逮捕されました。

それぞれの文を注意深く見てみると、複数の事実を含んでいることがわかる。下線は、その事実を表しているが、最初の文では、「男が包丁を振り回していた」「警察官が取り押さえようとした」「警察官が重軽傷を負った」という3つの事実が、2つ目の文にも3つの事実が含まれている。しかし、それぞれの文において、最初の2つの事実は、経緯を説明する位置づけにあり、最後の述部が、各文で伝えたい一番の出来事になっていると考えられる。そこで、ここではある項目の主題は、その項目の代表文の述部に現れていると仮定した。

図 5.4 に示した生成例では、最も左のノードが、あるニュース項目を表しており、誰かが逮捕されたことがわかる。そしてこの項目に関係する4種類の情報等がノードとして接続されており、上から順に、項目の全再生、主張訴えのショット、キーワード、関連ショットとなっている。なお、ノードの中に含まれる鍵型の矢印は、メディアへのリンクを示しており、これをクリックすると該当する映像や音声再生されるようになっている。これをみると、日本スケート連盟が関係していて、誰かが逮捕され、そのための記者会見が開かれ、該当者による謝罪などがあったことが想像できる。更に知りたい場合は、リンク先のメディアをクリックし、主張部分だけを聞いたりすることができるよう設計した。

人物サマリ

このサマリは、人物に注目したサマリという位置づけで、トピック内の全ての主張・訴えの役割をもつショットを集め、人物毎に、その日付で整理し、一覧性よく見せることことを目的とするものである。生成手法を、図 5.6 に示す。主張・訴えの役割をもつショットを抽出した後、その人物の名前を利用してグループを作成し、各グループを1つのノードとして、人物サマリの中心ノードに接続する。その際には、一覧性をよくするために、グループを表すノードには、サムネイルと人物名を加える。更に、各グループ内において、同一人物の発言の日付別にノードを生成し、それぞれの発言ショットをその子ノードとして加える。この場合にも、発言ショットの最初のフレームをノードに含め、視覚的な雰囲気を表示するようにする。

図 5.6 の生成では、当時外務大臣であった麻生太郎氏の発言が集められており、それぞれ発言日の日付毎にわかれている。末端のノードにある赤色の鍵型矢印はリンクを示しており、この部分をクリックすると、発言ショット部分が再生される。

時系列サマリ

このサマリは、伝えたいことサマリでも用いた、代表静止画と主題を表す文字列に、日付を加えたノードを形成し、それを時系列に追えるよう配置しなおしたものである。その生成手法は、図 5.7 に示したとおりである。生成する時系列は、3種類あり、トピック内の全ての項目を直線的に接続したもの、年毎に区切ったもの、月毎に区切ったものを用意し、年毎もしくは月毎の時系列は、全項目をつなげたものと違いがある場合のみ生成を行う。従って、あるトピックを形成する全てのニュース項目が、ある1ヶ月間に報道されたものであった場合、上述した3つのサマリは全て同じものになるため、この場合は年毎、月毎のサマリは生成され

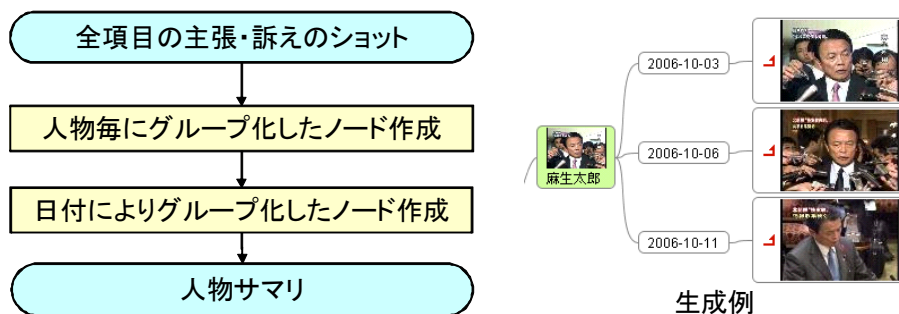


図 5.6: 人物サマリの生成



図 5.7: 時系列サマリの生成

ない。

図 5.7 に示した生成例では、3つのニュース項目が時系列に接続されたものである。各ノード、すなわちニュース項目をみると、それがいつのニュースであるか、何が起こったのか、またそのときの雰囲気や静止画像から得ることができるのがわかる。これらのノードにもリンクが張られているが、これは各項目の代表文の音声だけが再生されるようになっている。従って、何が起こったのかを表した文字列をみて、もう少し知りたい場合は、代表文を聞くことで概要をつかむことができるよう設計されている。

5.4 実験結果と考察

実験は、NHKの「ニュース7」、2006年10月1ヶ月分を用いて行った。本実験では、ショット分割情報、発話者とその発話内容の書き起こしテキスト、3.4.1で示した手法により自動抽出されたモノログショット中の一般人でない人物の名前を手で付与した。更に、入力ピックを形成するニュース項目の選択を一部人手で行った。上記以外のデータの作成は、全て自動で行った。

まず、ショット分割については、ニュース番組は特に暫時切替が多いため、現在の技術でも高精度で自動検出が可能である。また、ニュース内での発話者が放送局関係者であれば、比較的落ち着いてはっきり発話する傾向にあるため、高い精度で認識できることがわかっている。また、放送局関係者でない一般人の発話内容認識についても、最近の研究で精度が高くなりつつある。本実験で用いるのは、発話のタイミングと、放送局関係者だけであるため、自動化が実現可能と考える。最後の自動抽出したモノログショットに対し、一般人以外の人物の名前を付与する部分については、自動付与の研究[51]で、ある程度の自動化は可能であると報告されている。また、現在の放送局内でのニュース制作過程においても、主要人物のショットには人物名を付与しており、このデータを用いることを考えれば、本手法の自動化について特段難しい部分はないと考える。

表 5.2: 2006年10月のニュースから作成したトピック

番号	トピックの話題	ニュース項目の数
1	高校未履修科目問題	11
2	福島県：公共工事談合	10
3	北朝鮮拉致問題	7
4	福岡県：中2男子自殺	6
5	耐震強度偽装問題	5
6	愛媛県：臓器移植問題	5
7	スケート連盟：元会長経費流用	4
8	北海道：小6女子自殺	4
9	岐阜県：中2女子自殺	3
10	ディープインパクト	3
11	番号ポータビリティ制度	3
12	新潟県中越地震関連	3
13	北朝鮮核実験国内影響	3
14	ノーベル平和賞	3
15	自民党復党問題	3
16	安倍総理：日中関連	3

入力トピックのデータについては、あらかじめ著者が内容の関連性を判断し、2006年10月中の537件のニュース項目から作成した。表5.2は、著者が作成した3つ以上のニュース項目からなるトピックの一覧である。ただし、プロ野球など、スポーツは除いてある。

実験としては、この16のトピックについて5.3で述べた手法によりインタラクティブサマリを生成した。以降では、まず各ニュース項目からの代表静止画の抽出、代表文の抽出の抽出性能について、主観評価を行った結果を示し、その後に、3つのトピックについて、生成したインタラクティブ・サマリを示す。また、インタラクティブ・サマリそのものについても主観的评价を行った結果を示す。

5.4.1 ニュース項目の代表静止画・代表文の抽出結果の評価

各ニュース項目から抽出した代表静止画と代表文を評価するために、次のような主観的評価実験を行った。図5.8は、項目毎に全てのショットにおける最初のフレームを切り出し、その時の発話内容を右側に表示したもので、この画面を用いて被験者に代表画像と代表文を選択してもらった。

作業は、項目毎に、その全ての画像と発話内容を読んでもらい、その項目を一

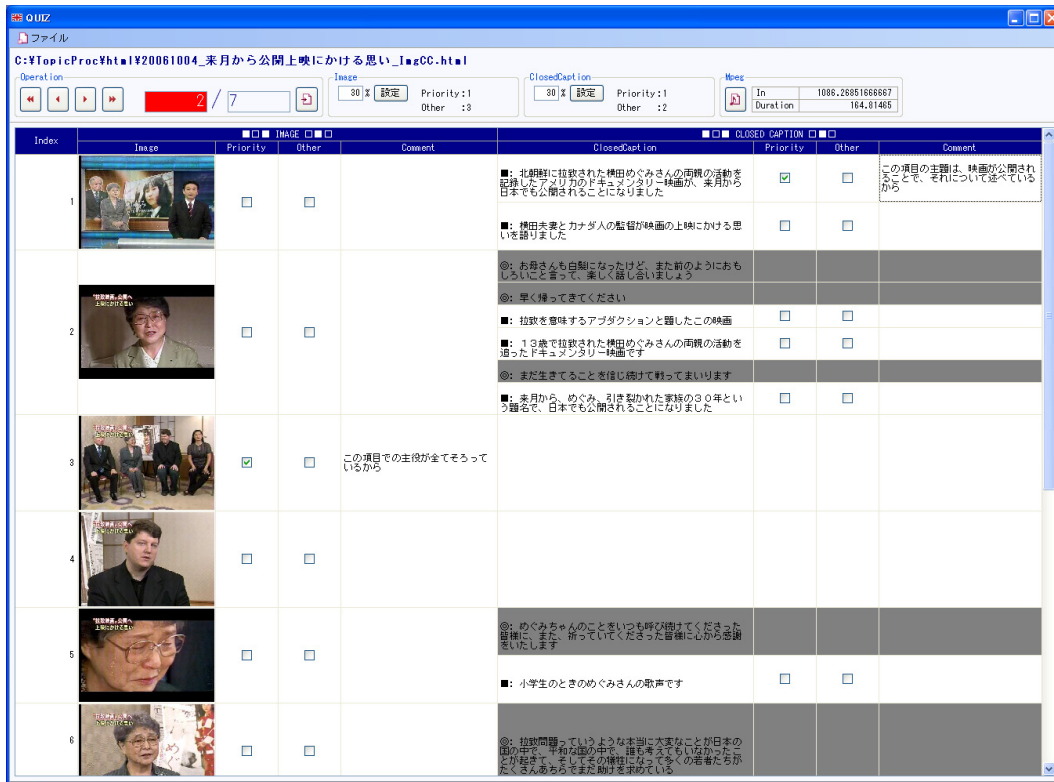


図 5.8: ニュース項目の代表画像と代表文を選択するツール

番よく表現していると思われる静止画と、発話文を選択してもらうというものである。ただし、次に示すいくつかの制約を設けた。

- 1) 画像の選択数は全画像数の30%の数まで選択し、スタジオのアナウンサショットは選択しない。
- 2) 字幕については、放送局関係者の発話文から全発話文数の30%の数選択する
- 3) 画像にスーパーとして重畳されている情報は無視し、映像だけに注目して選択をすること。
- 4) 画像と字幕の選択は、その日のニュースとして何が伝えたかったかに注目すること

制約4)を加えた理由は、大事件のニュースなどを考えると、毎回繰り返し使われる映像があると考えられるが、サマリなどのアプリケーションでは、新情報や内容の進展に関する映像や文章に注目するべきであると考えからである。

表 5.3 に評価結果を示す。表に示した数値は、被験者6名に選択してもらった代表画像と代表文について、項目毎に集計をし、それぞれトップから30%に自動選

表 5.3: 自動抽出した代表静止画・代表文の評価

種別	自動抽出が 30%に含まれた割合
代表静止画	50.6%
代表文	95.8%

択した画像，もしくは文が入っていた割合である．代表画像については，50.6%と決して高い数字ではないが，ランダムに選択するよりは良い結果が得られた．選択の際の理由を見てみると，同じ話題であっても人により注目する観点が違うことが最も大きく影響しており，代表静止画を選択する難しさを表している．例えば，あるダムの公共工事についての談合事件に関するものでは，ダムの風景がこの事件をシンボリックに表すと考える人もいれば，談合に関わった当事者の人物を選択する人，また，会社の建物を選択する人など様々であった．

提案手法では，導入・見せる役割のショット，モノローグ，スポットライト画像を利用しており，それぞれ次のようなことがわかった．導入・見せる役割のショットでは，会場全体の画像を映していることから，この画像が選択される場合もよく見られた．しかし，例えば，記者会見のような場合には，この導入ショットの次に，主となる人物の発言がある場合が多く，この場合には後者を選択する被験者も多かった．これは，文字通り「導入」であって，代表静止画という役割には十分でない場合である．モノローグに関しては，発話者がトピック内でどのように重要なかを判断することが必要であることがわかった．端的な例を示せば，街頭でインタビューされている一般人の画像は，代表静止画には不適切であるというような場合である．最後のスポットライト画像については，クイズ生成の場合と同様に，背景が無地ではないが話題になっている事柄が映っている場合を抽出できるようになれば，代表静止画への貢献度がかなり高まる．これらについては今後の課題としたい．

次に代表文の評価についてしてみると，自動抽出した 95.8%が含まれており，高い精度で抽出できた．特に，アナウンサによる導入部分の説明は，全体をまとめているため，代表文としてふさわしいことがわかる．また，本実験を通し，代表になり得る文が，項目最後の部分にも集中することがわかった．これは，その項目のまとめである場合もあるが，今後の予定など，どのように話題の出来事が進むかについて言及しているものが多い．これらは前述したニュースの役割の中の「今後の展開」を伝える部分であり，複数の項目間の関係やサマリ生成そのものにも有用な情報になると考える．

5.4.2 3種類のインタラクティブサマリの生成結果と評価

ここでは3件のトピックを例に取り上げ、生成したインタラクティブ・サマリを示す。

北朝鮮拉致についてのインタラクティブ・サマリ

実験で用いた2006年10月の1ヶ月間に生じた北朝鮮拉致に関するニュース項目は7つであり、これを入力してインタラクティブ・サマリを生成した。図5.9, 5.10, 5.11は、それぞれ「伝えたいことサマリ」「人物サマリ」「時系列サマリ」のデフォルト表示状態である。ただし、時系列サマリだけは通常横一列に長いため、折りたたんで示した。この状態から各ノードをクリックしていくと、様々な情報が展開される。インタラクティブ・サマリの「インタラクティブ」とは、最初の情報提示は最低限にすることでその一覧性を確保し、更に情報が必要な部分については、ユーザがクリックして展開していく、というコンセプトである。

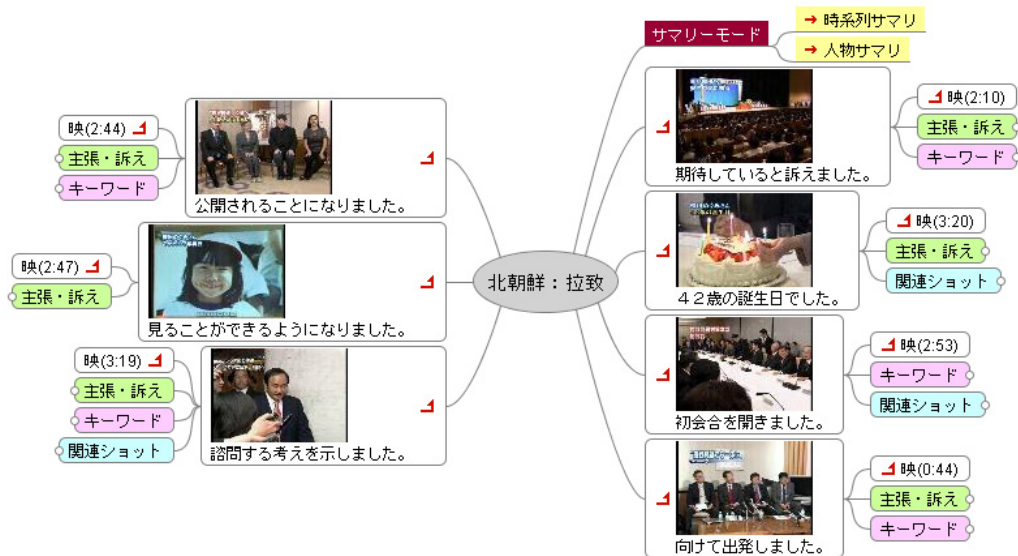


図 5.9: 北朝鮮拉致の伝えたいことサマリ (デフォルト表示)

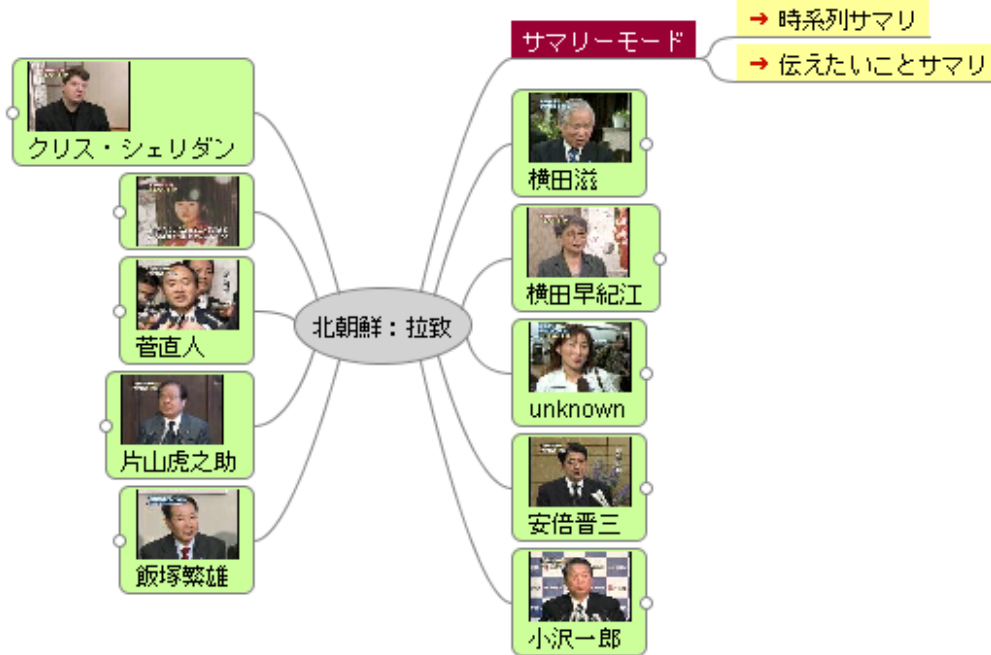


図 5.10: 北朝鮮拉致の人物サマリ (デフォルト表示)



図 5.11: 北朝鮮拉致の時系列サマリ ((デフォルト表示)

次に、どのような情報が抽出されているかを確認するため、伝えたいことサマリと人物サマリの全ノードを展開したものを図示する。図 5.12 は、伝えたいことサマリであり、注目は中心の「北朝鮮：拉致」というノードの7つの子ノードである。これらは1つ1つは、ニュース項目に相当し、1枚の代表画像と、この項目で伝えたいことを短い文字列で表現している。ここで左上から下へ、次に右上から下の順番に、項目内容を説明する。左上は、拉致問題が映画化されそれが公開されるというニュースである。画像には、横田夫妻をはじめ、映画を制作した監督一同がそろっているものである。その下の娘の顔写真は、全国で展開してきた横田めぐみさんの写真展が、Web 上でもみられるようになったというニュースである。左の一番下の項目は、北朝鮮の拉致問題解決の1つの策として、法律に基づくNHKへの「留意命令」を検討しているというニュースである。これについては、政治の放送への介入の恐れがあるなど、様々な人物がこの件に関して意見を述べており、主張・訴えのノードに該当するショットが抽出提示されているのがわかる。右上のノードは、北朝鮮拉致解決に向けて政府に訴える集会を鹿児島で開いたニュースである。その下は、拉致されためぐみさんの42回目の誕生日であることを伝えるニュースであり、代表静止画にもケーキが映っており適切な画像が選択されている。この画像は、導入・見せる役割で自動抽出されたものである。その下の項目は、拉致問題対策本部の初会合が開かれたというニュースであり、代表静止画も会議全体の画像であり、その雰囲気伝わってくると言える。最後、右下の項目は、拉致被害者の家族らが、更なる拉致問題解決のため、ニューヨークの国連に出発するというニュースである。このように眺めてみると、対象とした1ヶ月間の中に起こった北朝鮮拉致に関する出来事が簡潔に表現されていると言える。

人物サマリについては、図 5.13 に示すように、多くの人物が関係しており、更に発言の多い人物については、時間を追ってその内容を確認できるようになっている。なお、左の上から2番目のように、名前のないノードは、主張・訴えを伝えるモノログとして自動抽出されたが、実はモノログではなかったものが集まったものである。本実験では、自動抽出されたモノログにのみ人手により名前を付与したため、名前が付与されていないショットは全て、モノログでないショットであり、上記ノードにまとめられている。

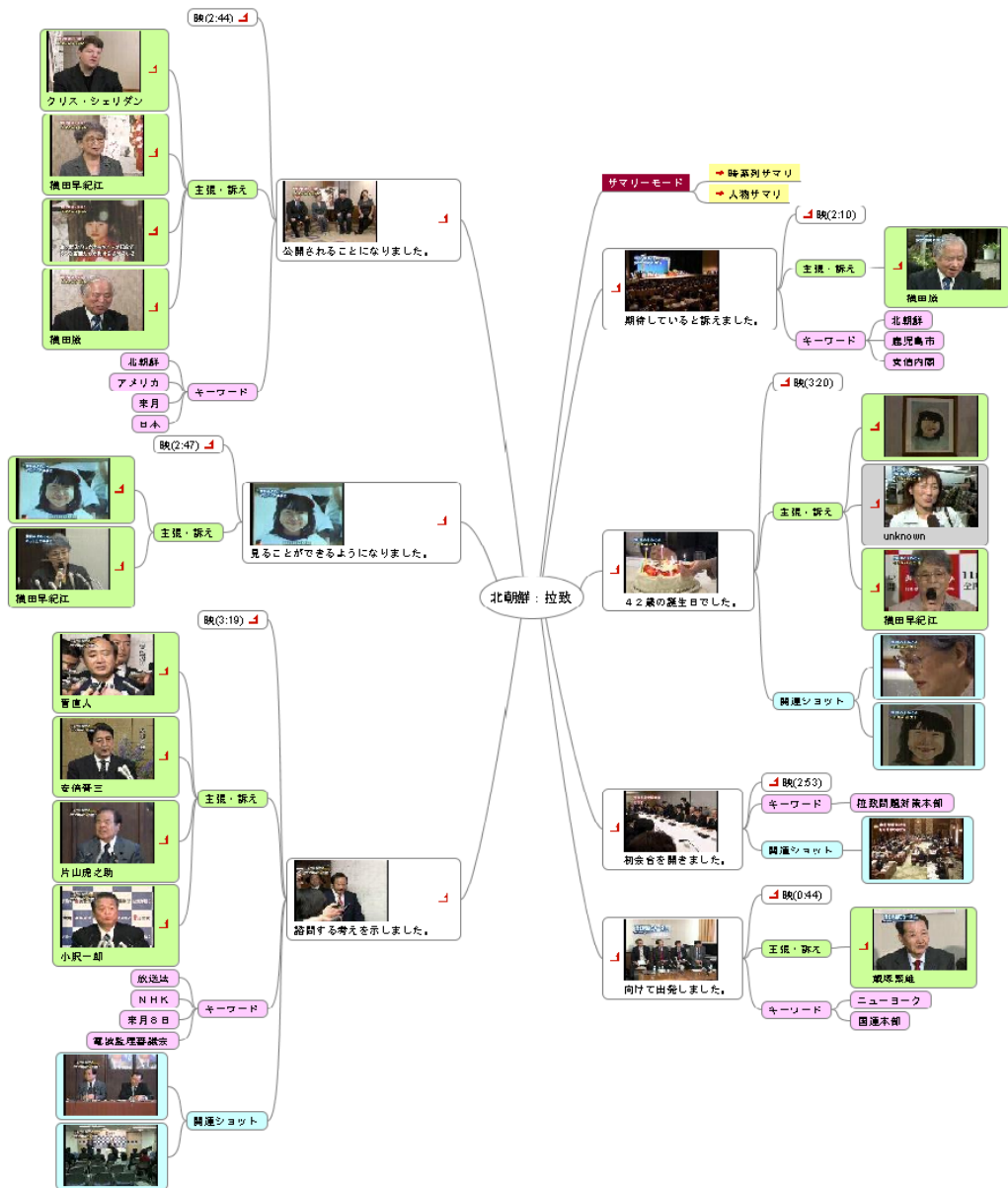


図 5.12: 北朝鮮拉致の伝えたいことサマリ (全展開後)

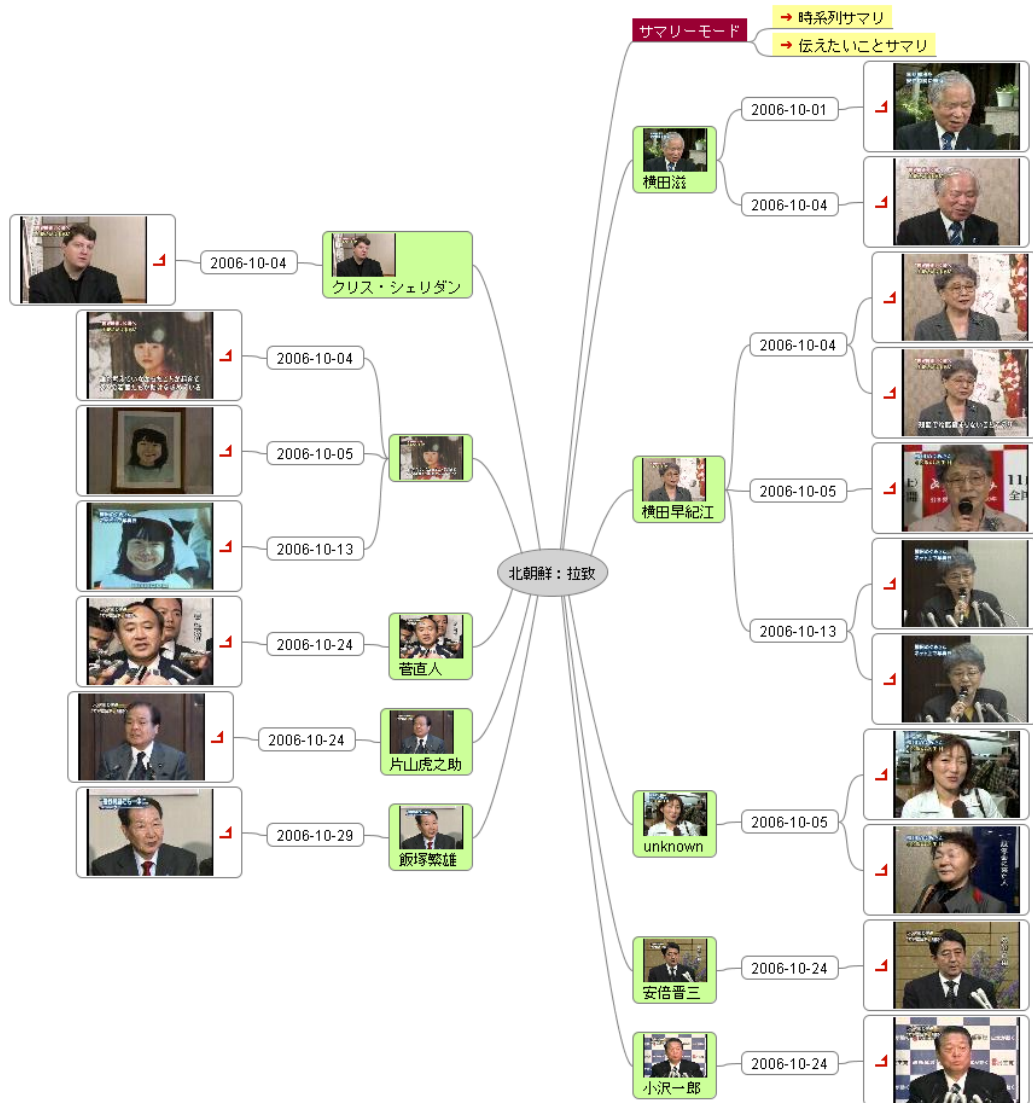


図 5.13: 北朝鮮拉致の人物サマリ (全展開後)

福岡県の中2男子自殺についてのインタラクティブ・サマリ

このトピックに関するニュース項目は6つであり、これを入力して生成したインタラクティブ・サマリを図 5.14, 5.15, 5.16 に示す。はじめに図 5.14 を用いて、このトピックを形成するニュース項目について説明する。前節と同様に左上から右下の順に説明する。左上の項目は、事件後に、学校側が全校集会を開き、教職員側が手を抜いてしまった、と謝罪したというニュースである。代表画像は、謝罪する校長のモノローグが自動選択されている。その下は、この問題を調査しに来た文部科学省の担当者に、男子生徒の父親が、いじめの実態把握のあり方について問いただしたというニュースであり、画像は、父親が問いただしている場面が選ばれている。左下の項目は、事件後の調査により、自殺する直前の状況があきらかになったことを伝えるニュースである。画像は、調査によって収集されたアンケート用紙が選択されている。右上の項目は、自殺した生徒が母親に話していた内容から、いじめのきっかけになったとも考えられる事実がわかったことを伝えるニュースである。その下の項目は、学校側がおこなったアンケートにより、いじめがあったことがわかったというニュースである。画像は、項目最後の方の文部科学大臣の導入ショットが選択されているが、ここでは適切でないと思われる。最後に、右下の項目は、文部科学省の小淵優子政務官が自殺した生徒の自宅を訪れ、遺影の前で焼香をしたニュースである。選択された画像は、小淵政務官が車から降りてくるところであるが、切り取った画像ではちょうど前面に関係者が映ってしまっており、本人はその陰に隠れてしまっている。

人物サマリ(図 5.15)については、自殺のあった学校の校長と、いじめの問題に詳しい専門家のモノローグが抽出されていた。実際には、このほかに2人の登場人物が意見を述べているショットがあったが、顔の向きが横向きだったためモノローグとして自動検出はできていなかった。今後、音声などの特徴を使うことで、これらは抽出可能と考える。

また、図 5.16 の時系列サマリをみると、このトピックについても、各項目の代表文の述部だけに注目するだけで、事実の判明、謝罪、問いただし、焼香、新たな事実というように、その出来事の移り変わりを把握できることがわかる。

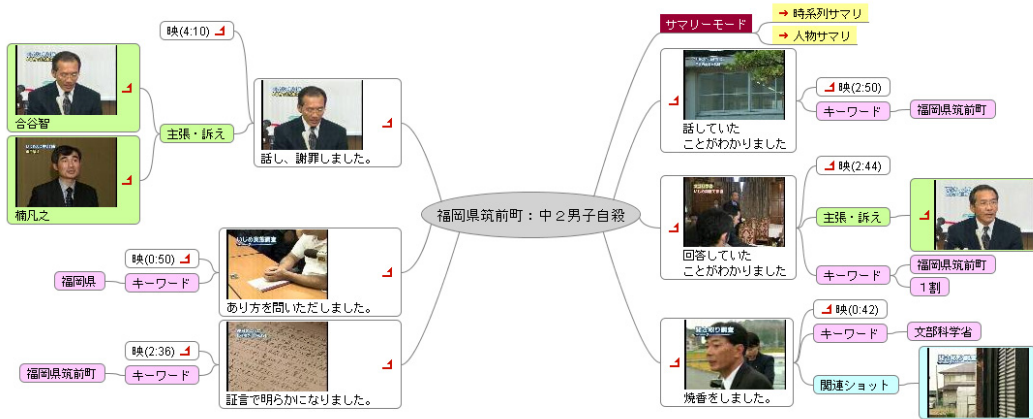


図 5.14: 福岡県の中2男子自殺の伝えたいことサマリ (全展開後)

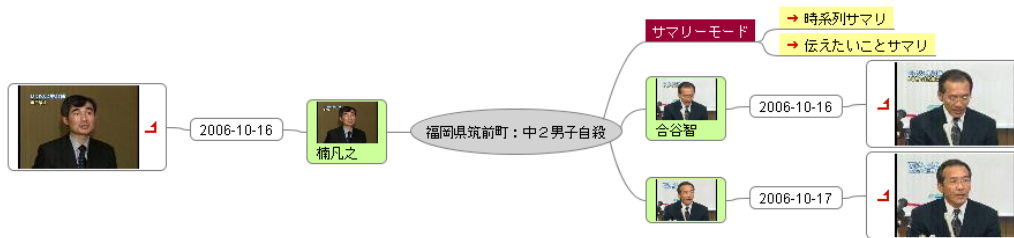


図 5.15: 福岡県の中2男子自殺の人物サマリ (全展開後)

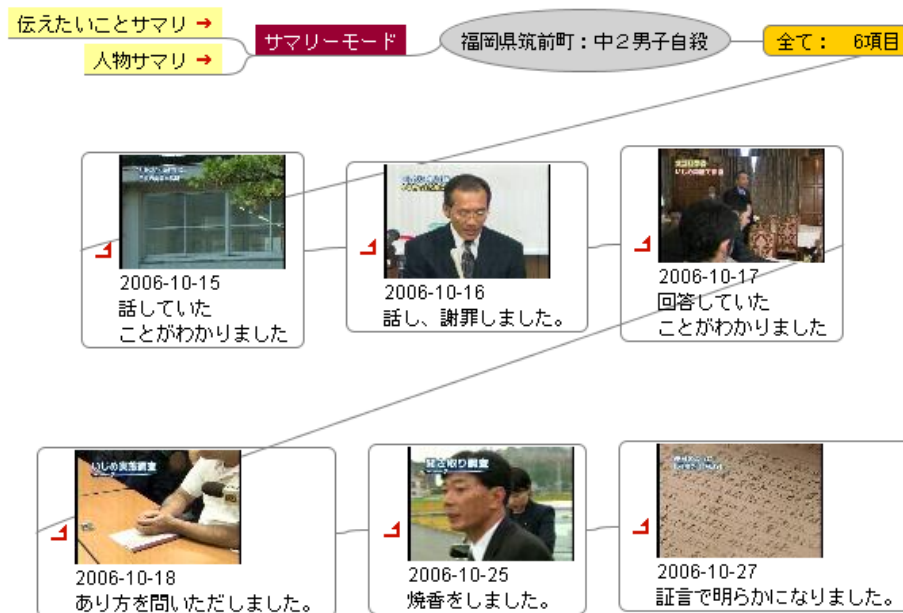


図 5.16: 福岡県の中2男子自殺の時系列サマリ (全展開後)

安倍総理の日中関係に関する活動についてのインタラクティブ・サマリ

このトピックに関するニュース項目は3つであり、これを入力して生成したインタラクティブ・サマリを図 5.17, 5.18, 5.19 に示す。安倍総理の日中関係に関する活動としては、3つという数が少ないトピックではあるが、サマリそのものは、これまでと同様に生成されていることがわかる。図 5.17 により、各ニュース項目を説明すると、左の項目は、安倍総理の昭恵夫人も中国を訪れ、外交デビューをはたしたというニュースである。選択された画像も、昭恵夫人がよくわかる画像が自動選択されている。右上の項目は、中国を公式訪問し、胡錦濤国家主席との首脳会談を行っているというニュースである。選択された画像は、導入・見せる役割のショットとして抽出されたもので、首脳会談の全景がわかる適切な画像であると言える。右下の項目は、上記首脳会談のことについて、国会で報告をしているものである。選択された画像からは、国会において安倍総理が何かを説明したことがニュースになっていることが一目でわかる。

他の人物、時系列サマリにおいても、関係している人物などが画像付きで提示されており、このトピックに関する様々な出来事の表面的な部分は、十分把握できると考える。

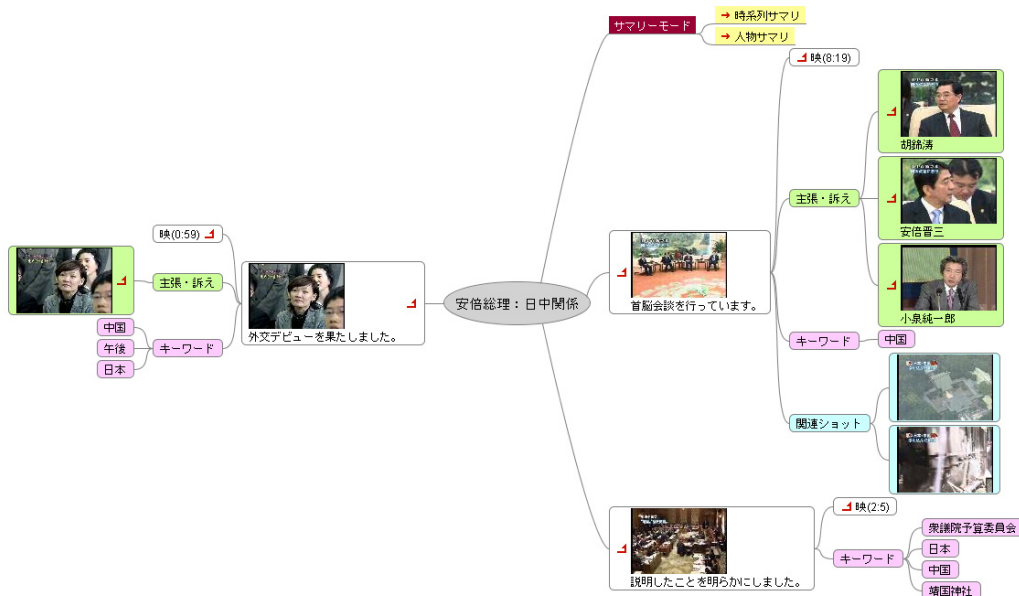


図 5.17: 安倍総理の日中関係に関する活動の伝えたいことサマリ (全展開後)

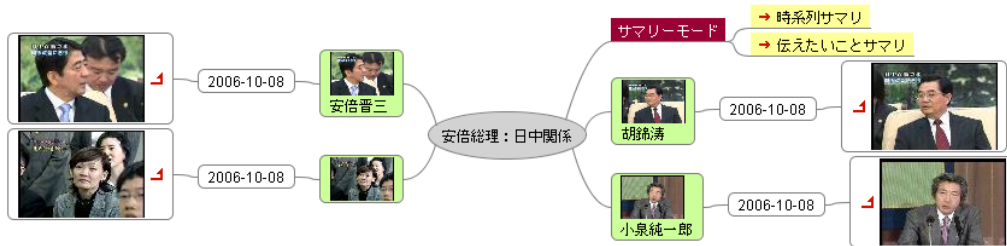


図 5.18: 安倍総理の日中関係に関する活動の人物サマリ (全展開後)



図 5.19: 安倍総理の日中関係に関する活動の時系列サマリ (全展開後)

生成したインタラクティブ・サマリへの評価

一般的にサマリの評価は難しいといわれている。それは、主観的要素が極めて強く、どこに主眼を置くのかによりその評価は大きく変わるからである。今回のインタラクティブ・サマリの評価は、要素毎に分けており、代表画像と代表文の選択についての評価は、既に 5.4.1 で示した。従って、ここでは、インタラクティブ・サマリそのものの評価結果について述べる。この評価のために、比較用として、項目を並べて映像を視聴できるようにした HTML(図 5.20) を作成した。これは DVD のチャプタ画面に似ているので、以降「チャプタ画面」と表現する。

被験者を 11 名を雇い、表 5.2 の 16 のトピックサマリについて、インタラクティブ・サマリとチャプタ画面を 1 時間自由に操作してもらい、表示の仕方、扱っているデータ、操作性、要望等について意見を求めた。なお、求めた意見は以下の通りであり、10 段階評価 (1~10) とし、評価値 5 はいずれも「どちらともいえない」とした。また、各項目とも記述欄を設け、補足説明をしてもらった。

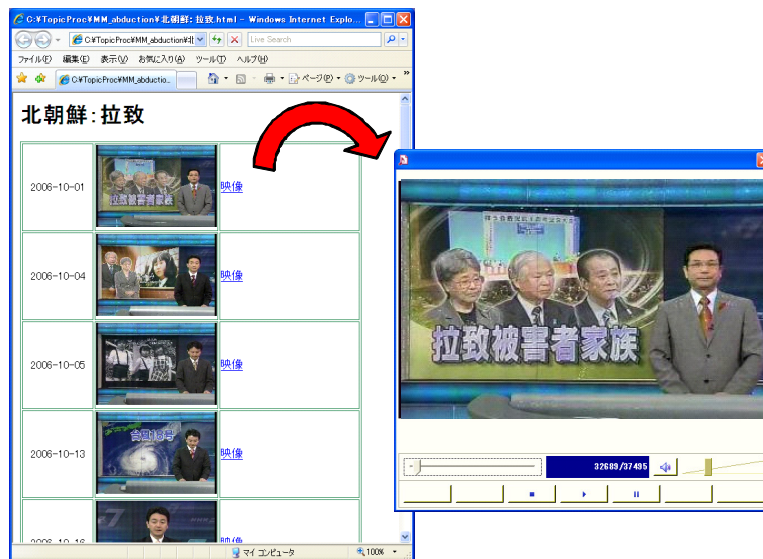


図 5.20: 項目を並べ各映像を視聴できるシステム (チャプタ画面)

- 操作性について (1. 操作しにくい 10. 操作しやすい)
- 実生活で使ってみたいか? (1. 使いたくない 10. 使ってみたい)
- 何か調べる時に便利だと思うか? (1. そう思わない 10. 便利だと思う)
- トピックの全体把握はしやすいか? (1. しにくい 10. しやすい)

- サマリ (要約) としてどう思うか? (1. なっていない 10. なっている)
- おもしろさはあるか? (1. ない 10. おもしろさを感じる)
- 総合的な満足度は? (1. 大変不満 10. 大変満足)
- 各システムの良い点を1つだけ挙げると? (記述のみ)
- 各システムの悪い点を1つだけ挙げると? (記述のみ)
- その他, 意見, 要望等 (記述のみ)

これらの項目に対する評価結果を表 5.4 に示す。はじめに, インタラクティブ・サマリの方が評価値が高かった項目を挙げると「調べる際に便利」, 「サマリになっている」, 「おもしろさがある」, 「総合評価」の4項目であった。逆に, チャプタ画面が高評価であった項目は「操作性」, 「実生活での使用希望」であり「トピック全体の把握」については同程度の結果となった。以下に, 各項目順に評価結果の詳細を述べる。

「操作性」については, 画面の構成からも予想されるように, チャプタ画面が「お年寄りにも見えそう」, 「リモコンでも操作可能」という理由で, インタラクティブ・サマリより2.3ポイント高い値が付いている。逆に, インタラクティブ・サマリは, マウスによる操作が必要で, クリックが有効な場所や, クリックにより発生する動作などが複雑であるという理由で低い値となった。

「実生活で使ってみたいか?」という問いに対しても, チャプタ画面の方が0.5ポイント高得点となった。これは, 一般的に機器やシステムの使用の際には, その操作性は重要な要素であり, 前項の評価結果が大きく影響していると考える。また, その用途を整理すると, チャプタ画面の方は「見逃したニュースをみたい」が多く, インタラクティブ・サマリの方は「ニュースの内容の検索, 概要把握」とい

表 5.4: インタラクティブ・サマリとチャプタ画面の評価結果 (11人の平均)

評価項目	インタラクティブ・サマリ	チャプタ画面
操作性について	6.1	8.4
実生活での使用希望	6.9	7.4
調べる際に便利か	7.7	6.3
トピックの全体把握	6.3	6.3
サマリ (要約) になっているか	6.1	4.3
興味/おもしろさ	7.4	4.8
総合評価	6.7	5.7

う意見が多かった。また、実生活から想像すると、自分と直結するニュースでない限り、詳しく調べようとは思わないため、単純に興味ある項目だけを視聴できれば満足する傾向にあると考えられる。この観点では、現在のハードディスクレコーダでは、番組単位での瞬時再生はできるが、1つの番組(ニュース)内の個々の項目を再生することはできず、チャプタ画面はこれが実現されており、かつ使いやすいため、高評価になったと考える。従って、特定のニュースについて調べたいという状況になれば、インタラクティブ・サマリの利用を望むと考えられ、補足説明にも同様な意見は複数みられた。

「調べる際に便利か？」に対しては、インタラクティブ・サマリが1.4ポイント高評価であり、最も多かった理由は「整理されているから」であった。本実験では、各項目の主題、時系列遷移、人物により整理をしているため、全てをみなくても良い点が高評価につながったと考える。また、チャプタ画面はアナウンサのショット画像であるのに対し、インタラクティブ・サマリでは、その項目を代表する静止画を利用しているため、視覚的にもわかりやすいと多くの被験者が評価をしていた。

「トピック全体の把握」については、どちらも同じ評価値となった。しかしながら、補足説明を整理すると次のようなことがわかった。チャプタ画面の方は、映像全部を視聴することで全体把握はできるが、チャプタ画面自体にはほとんど情報は無い。一方、インタラクティブ・サマリは、様々な情報が画像付きで一覧できて良いが、2次元的に広がって複雑になる傾向があり、やはり把握はしづらい。ただし、時系列トピックに関しては、1次元にデータが並べられているため、それを順に追うことで把握はできるという意見が多かった。

「サマリ(要約)になっているか？」に対しては、インタラクティブ・サマリの方が1.8ポイント高かった。特に、時系列サマリについては、各項目の主題が現れ、時間順に追えるため高評価であった。また、人物サマリについても、トピックに関連する人物が一覧でき、またその人物の主張もビデオで視聴できるため高評価であった。チャプタ画面の方は、多くの被験者がサマリ的な要素はないと判断していたが、数人は「大きな要約」「おおざっぱな要約」として評価していた。これは、チャプタ画面に、あるトピックに関するニュース項目が集められているということの評価していたと考えられる。これらの意見からは、改めて「要約」にも様々なレベルが存在することを確認した。

「おもしろさがあるか？」に対しても、インタラクティブ・サマリの方が2.6ポイント高かった。これは新しいサービスとして展開する場合に、おもしろさは重要な要素であるため設定した問いである。設問の中では、2つの評価点に最も差が

生じた項目でもあった。多く被験者は、ニュースの違う側面をみることができそうな点をおもしろいと感じていた。これは、ニュースがある点で整理されていると、そこを起点として、新しい見方や疑問が沸き、ニュース内容のネットサーフィンを想定しているようであった。また、ニュース番組に限らず、他のドラマやクイズ番組にも適用したらどうかという提案も複数あった。いずれにしても、蓄積した番組を視聴する新しい糸口になりえると考える。

「総合的な評価」については、インタラクティブ・サマリの方が1.0ポイント高い結果となった。主な意見としては、チャプタ画面の方は、とにかく単純で使いやすい点が良いというものであった。インタラクティブ・サマリの方は、もう少し画面構成や動き、抽出される情報が改善されれば、映像辞書的であり便利なサービスになりそうという意見であった。

以上の数値による評価に加え、記述のみによる評価の結果を含めて、2つのサービスについて全体をまとめると、次のようになる。チャプタ画面は単純で使いやすく良いが、全てを見なくてはならない点が問題である。インタラクティブ・サマリは、情報が整理されている分、特定の情報にメディアを含めてアクセスできるため便利であるが、情報の提示を工夫しないと複雑さが増加し問題となる。従って、これらの評価実験を通し、サマリにも様々な要求があり、それぞれに適した表示を考えることが重要であることがわかった。これらについては今後の課題とするが、少なくとも、ある観点でニュース項目を解析し情報を整理し表示・アクセスできるような環境は、今後の我々の生活には有用であると考えられる。

最後に、インタラクティブ・サマリについて記述してもらったものを整理しまとめる。提案した3つのサマリの中では、時系列サマリが一番好評であり、映像や画像など視覚的情報が付与されている点が高く評価されていた。また、不完全ではあるが、各項目の「簡潔な」表現についても肯定的な意見が多かった。特に、事件などの場合は、「逮捕されました」「罪を認めています」「～ということがわかりました」「起訴されました」など、関連する出来事がどのように展開しているのかがわかりやすいという意見が多かった。一方で、文章の述部だけでは、誰が主体なのかという疑問がわき、主語も付与した方が良いという意見も多かった。また、文章をそのまま提示するのは、冗長で長すぎるため、「誰がどうした」という「主語+述語」のタイトル的な文字列生成が重要であることがわかった。

人物サマリについても、そのトピックに関連する人物が一目でわかり、しかもその発言を時間順に追えるため、肯定的な意見が多かった。また、名前だけでなく、職業などや、トピックの中での位置づけについて情報があると良い、更には人物同士の関係などの記述を要望する意見があり、今後の課題として考えていきたい。

最後に、このアプリケーションの存在意義についての意見を整理する。日々のニュースは、限られた短時間での情報提供であり、更に前回の報道から日にちが経過した場合には、これまでの話題についての遷移や、場合によってはその存在さえも忘れていくことが多い。従って、このインタフェースのように時系列で遡って、その発端について知る、また経緯を知ることができることは大変良いとの評価が大半であった。また、その項目を全部みなくても、ポイントが画像/映像付きで提示され、知りたい人は自由に更に詳細を知ることができるような設計も望ましいという意見が多かった。以上の主観的評価により、本章で提案したインタラクティブ・サマリの存在価値は十分にあると考える。

5.5 まとめ

本章では、EvenTank を基盤とし、ニュース映像を活用した新しいコンテンツ生成として、インタラクティブ・サマリの生成手法を提案した。このサマリでは、ある関連するニュース項目群、すなわちトピックについて、「何がおこったのか」「誰がかかわっているのか」「これまでの経緯は？」の3つの視点によりデータを整理し、簡潔に提示をした。主観評価によれば、本インタラクティブ・サマリの存在価値は大きく、本手法で注目した、適切な画像と文章の簡潔な表現についても、その方向性で更に精度を高めていくことが必要であることがわかった。

また、本生成システムの構築にあたっては、EvenTank を基盤としているため、代表静止画を抽出する際のスタジオアナウンスショットの検出機能、スポットライト画像の検出機能、また言語処理については、形態素解析や構文解析などの解析結果から情報を抽出する部分などは、前章のクイズ生成で開発したものをそのまま流用した。これは、EvenTank の設計思想である、処理のモジュール化と再利用、また、生成したデータの再利用を活用しており、本実験を行うことで、EvenTank の実用性も同時に実証した。

第6章 結論

本論文では、コンテンツ制作支援を目的とし、放送済み番組を効率よく再利用するための共通基盤として、ニュース番組を蓄積し活用するフレームワーク提案した。また、これらニュース番組の映像を再利用するためのメタデータについて解析し、今後のメタデータ生成への方向性を示した。

第1章では、本研究の背景として、大きなコンテンツ提供組織である放送局を例に挙げ、そこにおける昨今のコンテンツ制作環境について述べた。次に、本論文でニュース番組に着目した理由を述べた後、現状におけるニュース番組の再利用について、主に従来研究を中心に整理を行った。そして、そこに存在する技術的問題点を明らかにした。

第2章では、蓄積したニュース番組を効率よく再利用するための共通基盤(フレームワーク)として EvenTank を提案した。はじめに、従来研究から得られる要求条件について検討し、これを基にした EvenTank の設計指針について説明した。次に、EvenTank におけるデータ構造とその表現について、既存の関連フォーマットと比較を行いながら提案を行った。最後に、これまでの個々の従来研究について、EvenTank における位置付けを説明し、設計指針の有効性を確認しつつ、将来への拡張性と柔軟性についても確認した。

第3章では、ニュース番組の映像を、より適切に再利用するために必要なメタデータと、その付与手法について提案した。はじめに、従来おこなわれてきた映像に対するメタデータ付与について整理をし、適切な再利用のためには更なる情報が必要であることを例を挙げて説明した。そして、今後は映像の役割に注目したメタデータを抽出する必要性を提案し、このためには、映像等の特徴だけでなく、ニュース制作の手順などの情報も利用すべきことを提案した。実際に、この役割メタデータの一部について、番組制作手法を考慮したアプローチを提案し、ニュース番組を用いた実験を行い、この仮定の妥当性について確認した。

第4章では、第2章で述べた EvenTank の枠組みを用いて、ニュース番組の新しい活用の方向として、クイズの生成実験を行った。生成対象とするクイズは、テキストだけでなく画像(映像)も利用するクイズであり、第3章で述べる番組制作のルールを利用した手法を提案した。具体的には、ニュース映像が伝えようとして

いる「主被写体」に注目し、この主被写体が明確なものをスポットライト画像と定義して、自動抽出を行いクイズに活用した。クイズ生成手法については、3つのサブタスクにより構成し、それぞれに対する工学的アプローチを提案した。実際のニュース番組を用いて生成実験を行い、各サブタスクへのアプローチと、それを総合したアプローチについて評価を行ったところ、自動生成した約4割が有効なクイズであることを確認した。また、主被写体が明確な画像として、スポットライト画像を導入し、クイズ生成に有用であることを確認した。同時に、スポットライト画像の定義に当てはまらない、主被写体が明確な画像も多く存在することが明らかとなったので、今後の課題としたい。最後に、第2章で提案した EvenTank の枠組みの中で実際のシステムが構築できたことを確認した。

第5章では、前章同様に、EvenTank の枠組みを用いて、あるトピックに関するインタラクティブ・サマリの生成実験を行った。ある視点により集められた複数のニュース項目を入力とし、各項目から第3章で提案した番組制作ルールを利用した映像の役割抽出を行い、エッセンスとなる画像(ショット)を抽出した。このエッセンスとなり得る画像やショットの候補として、モノローグショット、スポットライト画像、サイレントショットに注目し、評価実験によりこれらが確かにその役割を果たす場合が多いことを確認した。しかし、項目の代表文の自動抽出の評価結果と比べると、その一致度は低い。これは、映像の多義性と、サマリの視点の違いなどいくつかの主観的要素が大きくからんでいるため、更にニュースの内容を深く理解することでその精度をあげることができると考える。次に、各項目から抽出した代表画像や文章を用いて、インタラクティブ・サマリの生成を行った。このサマリについても主観的評価を行い、その存在意義と利便性について多くの肯定的な意見を得た。同時に、更に内容に踏み込んだ適切な映像選択と、抽出したエッセンス同士の関係抽出など、今後の進むべき方向性も明らかになった。最後に、本システムで用いている処理モジュールのいくつかは、前章のクイズ生成で開発したものであり、EvenTank が提供する共通基盤上でそれらを流用できている点からも、その実用性について実証できたと考える。

本論文の成果は、コンテンツ制作を支援するために、蓄積された放送済みニュース番組を効率よく利用するためのフレームワークを提案し、このフレームワークに沿った形で新しいサービスコンテンツの生成実験を行い、その有効性を示したことと、さらに、映像の適切な再利用のためのメタデータについて、番組の制作ルールとニュース内での映像の役割を考慮した手法を提案し、実際のデータで検証するとともに、今後の映像再利用へ向けたメタデータ付与の1つの進むべき方向性を示したことにある。

今後については、コンテンツ制作とメタデータ付与の作業が協調して効率よく行える環境整備が必要と考える。これにより、既存のコンテンツを活用した新しいサービス展開などは、低いコストで行えるようになる。また、メタデータについては、更に深い意味的情報を整理して付与することが重要になると考える。更には、これらメタデータの流通についても考える必要があり、コンテンツ制作やその再利用を行うコンテンツ提供者、そのコンテンツを伝送もしくは視聴するための端末を開発する企業や大学、そしてコンテンツを楽しむユーザ、皆が有意義に参加できるプロジェクトを設定し、具体的目標を定めて連携して進めていければ理想的と考える。

謝辞

本論文は、様々な方々のお力添えのもとに完成いたしました。

はじめに、本論文の主査であり、主任指導教官でもありました国立情報学研究所/総合研究大学院大学の佐藤真一教授におかれましては、本研究全般に関して多大なご指導とご鞭撻を賜りました。本研究をまとめるにあたり幾度となく挫けそうになったときも叱咤激励とともに我慢強くご支援くださったことなど、先生のご指導なくしては本研究は成し得ないものでありました。心より厚くお礼申し上げます。

また本論文をまとめるにあたり、論文審査委員の先生方からも多大なご指導とご鞭撻を賜りました。片山紀生准教授におかれましては、わたくしの副指導教官を務めて頂き、マルチメディア処理の講義をはじめ、いろいろとご指導ご鞭撻を賜りました。また、研究素材となった大量のニュース番組と字幕データは、片山先生を中心とした多くの方々の多大なる努力によって整備された映像データベースがあったからこそ着手できた研究であり、厚くお礼申し上げます。曾根原登教授におかれましては、非常にお忙しい中、お時間を割いて頂き様々なご指摘、ご鞭撻を頂きましたこと深く感謝致します。特に、放送局と視聴者の両方の視点を常に考慮しつつ、研究の位置づけを明確化する点など、多くの有用なご指摘を賜りました。感謝致します。相澤彰子教授におかれましては、言語処理の講義を通し、テキストの扱い方の基本から多くのことをご指導頂きました。また、中間発表などでも言語解析分野からの違った視点での方向性などをご指摘頂くなど、大変感謝しております。井手一郎准教授におかれましては、同じニュース解析における先行研究を多くされており、そこから得られる知見、ヒントはもちろん、本研究への鋭く重要なご指摘を頂きましたこと、深く感謝致します。皆様のご指導、ご鞭撻により本論文をまとめ、より質の高いものとすることができました。重ねてお礼申し上げます。

また、本研究の機会を与えてくださり、ご指導、ご鞭撻を賜った、日本放送協会の多くの方々に心より感謝致します。特に、放送技術研究所の榎並和雅元所長(現NICT けいはんな研究所所長)、谷岡健吉元所長、久保田啓一所長のご理解とご鞭撻は、本論文をまとめるうえで大きな支えとなりました。厚くお礼申し上げます。

八木伸行氏には、約3年に渡り、公私を問わず、いろいろな相談に懇切丁寧にのって頂き、本研究を進める上での大きなよりどころの1つでありました。深く感謝致します。更に、柴田正啓氏、藤井真人氏、住吉英樹氏にも、いろいろな面でサポートをして頂きましたことお礼申し上げます。クリピングデル氏、後藤淳氏、山田一郎氏、望月貴裕氏、松井淳氏、宮崎勝氏、三須俊枝氏、高橋正樹氏、河合吉彦氏、三浦菊佳氏、Yeonhee Oh, Frank Nack, Duy-Dinh Le, Jean Martinet, 山岸史典、瀧本政雄、武小萌、杜黙林、Ngo Duc ThanhをはじめとしたNHK放送技術研究所の皆様、並びに国立情報学研究所/総合研究大学院大学佐藤真一研究室の皆様には、日頃のご討論、ご指導、ご鞭撻を賜りましたこと、深く感謝致します。最後に、心の支えとなり元気つけてくれた家族に感謝します。

付録A NewsMLのサンプルデータ

NewsMLのサンプルデータ

```

<newsItem xmlns="http://iptc.org/std/nar/2006-10-01/"
  xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xsi:schemaLocation="http://iptc.org/std/nar/2006-10-01/
  http://www.ebu.ch/metadata/EBUnewsml-g2/sample/NewsML-G2_2.2-spec-NewsItem-Power.xsd"
  xmlns:ebu="http://ebu.org/nar-extensions/" standardversion="2.2" guid="tag:ebu.ch,2008:407729"
  version="1" standard="NewsML-G2" conformance="power" xml:lang="en">

  <catalogRef href="http://www.iptc.org/std/catalog/catalog.IPTC-G2-Standards_6.xml"/>
  <catalogRef href="http://www.ebu.ch/metadata/EBUnewsml-g2/catalog.EBU-NewsML-G2.xml"/>
  <rightsInfo>
    <usageTerms> no access Belgium </usageTerms>
  </rightsInfo>

  <itemMeta>
    <itemClass qcode="ninat:video"/>
    <provider qcode="EBUProviderCode:EBU">
      <name>European Broadcasting Union - EVN</name>
      <definition>Eurovision Exchange Network</definition>
      <organisationDetails>
        <contactInfo>
          <web>http://www.eurovision.net</web>
          <phone>+41 22 717 2850</phone>
          <email>evn@eurovision.net</email>
          <address role="AddressType:Office"><line>Eurovision News Exchanges</line>
          <line>L Ancienne Route 17 A</line><line>CH-1218</line>
          <locality literal="Grand-Saconnex"/>
          <country qcode="ISOCountryCode:ch"><name xml:lang="en">Switzerland</name></country>
        </address>
        </contactInfo>
      </organisationDetails>
    </provider>
    <versionCreated>2008-11-07T12:23:29Z</versionCreated>
    <firstCreated>2008-11-07T11:12:50Z</firstCreated>
    <pubStatus qcode="NMSStatusCode:TRS"/>
    <profile>TBA</profile>
    <service qcode="EBUServicerCode:EUROVISION">
      <name>Eurovision services</name>
    </service>
    <edNote>EBU NewsML DRAFT release</edNote>
    <link href="/News/07Nov2008/Items/20081107-PNN-1132-407729/index.html"/>
  </itemMeta>

  <contentMeta>
    <contentCreated>2008-11-05T00:00:00Z</contentCreated>
    <located type="cptype:city">
      <name>KIBUMBA</name>
      <broader type="cptype:province">
        <name>KIVU</name>
      </broader>
    </located>
  </contentMeta>

```

```

</broader>
<broader type="cptype:country" qcode="country:CG">
  <name>CONGO</name>
</broader>
</located>
<creator qcode="EBUNMSSRVCodeSource:BEVRT">
  <name>Vlaamse Radio en Televisieomroep </name>
  <organisationDetails>
    <location literal="BRUX"/>
  </organisationDetails>
</creator>
<contributor qcode="EBUNMSSRVCodeOrigin:BEVRT" role="EBURoleCode:TechnicalOrigin">
  <name>Vlaamse Radio en Televisieomroep </name>
</contributor>
<altId/>
<language tag="en" role="EBULanguageUsageCode:PartNarrated">
  <name>French</name>
</language>
<subject qcode="EBUNMSCategoryCode:EXCH"/>
<subject qcode="EBUNMSWebcatCode:AFRICA">
  <name>Africa</name>
</subject>
<headline>Congo rebel area</headline>

<description role="EBUNMSDescription:dopesheet"> For over a week now Kibumba has been in
  the hands of the CNDP, the rebel movement of Laurent Nkunda. His soldiers are in
  control of the roads and the markets. 75\% of the local population has fled.
  Those who stayed or have come back, have to be very careful. <br/><br/>
  The propagandists of the CNDP are doing their job. The story is simple: we have the
  power now, and it is better to accept that. Everyone seems to be complying
  obediently. <br/><br/>
  But only 1km away from the village, on a hill, some people tell another side of the
  story, about plundering and how they had to flee. They say that immediately after
  the attack, gangs from neighbouring country Rwanda came in and stole everything that
  the villagers had left behind, even windows and doors from the best houses. <br/><br/>
  And more villages in the rebel area are now completely empty. Those we stayed behind,
  mistrust the new rulers. Less than one kilometre away, in a small strip of no no man?s
  land, people are still being killed.<br/>
</description>

<description role="EBUNMSDescription:shotlist">
  1. various of people listening while CNDP soldiers are talking. <br/>
  2. Soundbite CNDP man: You don't have to be scared, because wherever the CNDP is,
    there will be peace and security so you can live in peace. If the great leader
    Nkunda asks to be patient, don't be impatient. Ok?<br/>Answer: Yes!<br/>
</description>

<ebu:dt_log_start>07-NOV-2008T11:33:34Z</ebu:dt_log_start>
<ebu:dt_log_end>07-NOV-2008T11:36:06Z</ebu:dt_log_end>
<ebu:soundType>Natural</ebu:soundType>
</contentMeta>

<partMeta partid="Part1_ID" seq="1">
  <icon href="/News/07Nov2008/Keyframes/20081107-PNN-1132-407729
    //20081107-PNN-1132-407729-11333418.jpeg"/>
  <timeDelim start="1040368" end="" timeunit="EBU:FPS-EditUnits"/>
  <language tag="en" role="EBULanguageUsageCode:VoiceOver"/>
  <description>
    1. various of people listening while CNDP soldiers are talking.
  </description>
</partMeta>

```

```
<partMeta partid="Part2_ID" seq="2">
  <icon href="/News/07Nov2008/Keyframes/20081107-PNN-1132-407729
    //20081107-PNN-1132-407729-11334617.jpeg"/>
  <timeDelim start="1040667" end="" timeunit="EBU:FPS-EditUnits"/>
  <language tag="en" role="EBULanguageUsageCode:VoiceOver"/>
  <description>
    2. Soundbite CNDP man: You don't have to be scared, because wherever the CNDP
      is, there will be peace and security so you can live in peace. If the great leader
      Nkunda asks to be patient, don't be impatient. Ok? Answer: Yes!
  </description>
</partMeta>

<contentSet>
  <remoteContent href="destination/path/filename(407729)" duration="152" id="_b07A1132"/>
</contentSet>
</newsItem>
```

付録B EvenTank データモデル

データモデル

EvenTank のデータモデルは、SourceItem と Topic の 2 つのデータ型が基本処理単位となっており、それぞれ SourceItemType、TopicType として XML Schema で定義されている。以下に、EvenTank のスキーマを示し、その後に全複型定義 (Complex Type Definition) についてそのデータ構造を示す。

EvenTank のスキーマ

```
<schema xmlns="http://www.w3.org/2001/XMLSchema" xmlns:et="http://www.nhk.or.jp/str1/evenTank"
xmlns:ns1="http://www.nhk.or.jp/str1/et" targetNamespace="http://www.nhk.or.jp/str1/evenTank"
elementFormDefault="qualified" attributeFormDefault="unqualified">
  <complexType name="SourceItem">
    <sequence>
      <element name="si_id" type="token"/>
      <element name="ownerType" type="et:OwnerTypeType"/>
      <element name="SourceItemHeader" type="et:SourceItemHeaderType"/>
      <element name="DelegateMeta" type="et:DelegateMetaType"/>
      <element name="SegmentMeta" type="et:SegmentMetaType"/>
    </sequence>
  </complexType>
  <complexType name="SourceItemHeaderType">
    <sequence>
      <element name="SourceItemName" type="string"/>
      <element name="BroadcastStation" type="string"/>
      <element name="OnAirDate" type="dateTime"/>
      <element name="MediaTitle" type="string"/>
      <element name="MediaUri" type="anyURI"/>
      <element name="MediaTime" type="et:MediaTimeType"/>
      <element name="ItemOrder" minOccurs="0">
        <complexType>
          <sequence>
            <element name="ThisItem" type="nonNegativeInteger"/>
            <element name="TotalItems" type="nonNegativeInteger"/>
          </sequence>
        </complexType>
      </element>
    </sequence>
  </complexType>
  <complexType name="DelegateMetaType">
    <sequence>
      <element name="WHKeywords">
        <complexType>
          <sequence>
            <element name="When">
```



```
<complexType>
  <sequence>
    <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
  </sequence>
</complexType>
</element>
<element name="Where">
  <complexType>
    <sequence>
      <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>
</element>
<element name="Who">
  <complexType>
    <sequence>
      <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>
</element>
<element name="WhatAction">
  <complexType>
    <sequence>
      <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>
</element>
<element name="WhatObject">
  <complexType>
    <sequence>
      <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>
</element>
<element name="Why">
  <complexType>
    <sequence>
      <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>
</element>
<element name="How">
  <complexType>
    <sequence>
      <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>
</element>
</sequence>
</complexType>
</element>
<element name="L_Sentences">
  <complexType>
    <sequence>
      <element name="Sentence" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
  </complexType>
</element>
<element name="L_KeyValues">
  <complexType>
    <sequence>
```

```

        <element name="KeyValue" type="et:KeyValue" minOccurs="0" maxOccurs="unbounded"/>
    </sequence>
</complexType>
</element>
</sequence>
</complexType>
<complexType name="KeyValue" type="et:KeyValue">
    <sequence>
        <element name="ownerKey" type="et:OwnerKeyType"/>
        <element name="L_Values">
            <complexType>
                <sequence>
                    <element name="Value" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
                </sequence>
            </complexType>
        </element>
        <element name="L_PCDataValues">
            <complexType>
                <sequence>
                    <element name="PCDataValue" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
                </sequence>
            </complexType>
        </element>
        <element name="L_FileValues">
            <complexType>
                <sequence>
                    <element name="FileValue" type="et:ValueType" minOccurs="0" maxOccurs="unbounded"/>
                </sequence>
            </complexType>
        </element>
    </sequence>
</complexType>
<complexType name="MediaTime" type="et:MediaTime">
    <sequence>
        <element name="inPoint" type="long"/>
        <element name="duration" type="long"/>
    </sequence>
</complexType>
<complexType name="Topic" type="et:Topic">
    <sequence>
        <element name="tp_id" type="token"/>
        <element name="ownerType" type="et:OwnerType"/>
        <element name="TopicHeader" type="et:TopicHeader"/>
        <element name="L_SourceItemLinks">
            <complexType>
                <sequence>
                    <element name="SourceItemLink" type="et:SourceItemLink" maxOccurs="unbounded"/>
                </sequence>
            </complexType>
        </element>
        <element name="L_Descriptions">
            <complexType>
                <sequence>
                    <element name="Description" type="et:KeyValue" minOccurs="0" maxOccurs="unbounded"/>
                </sequence>
            </complexType>
        </element>
    </sequence>
</complexType>
<complexType name="TopicHeader" type="et:TopicHeader">
    <sequence>

```

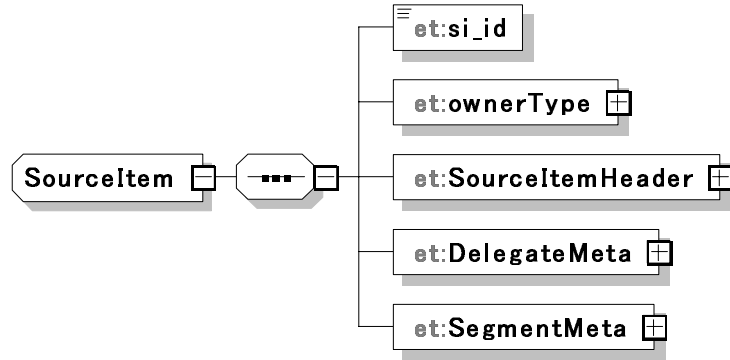
```

<element name="TopicName" type="string"/>
<element name="TopicAbstract" type="string" minOccurs="0"/>
<element name="SourceItemsPeriod">
  <complexType>
    <sequence>
      <element name="From" type="dateTime"/>
      <element name="To" type="dateTime"/>
    </sequence>
  </complexType>
</element>
</sequence>
</complexType>
<complexType name="OwnerKeyType">
  <sequence>
    <element name="owner" type="string"/>
    <element name="key" type="string"/>
  </sequence>
</complexType>
<complexType name="OwnerTypeType">
  <sequence>
    <element name="owner" type="string"/>
    <element name="type" type="string"/>
  </sequence>
</complexType>
<complexType name="ValueType">
  <sequence>
    <element name="ownerType" type="et:OwnerTypeType"/>
    <element name="value" type="string"/>
  </sequence>
</complexType>
<complexType name="SegmentKeyValueType">
  <sequence>
    <element name="ownerType" type="et:OwnerTypeType"/>
    <element name="L_KeyValues">
      <complexType>
        <sequence>
          <element name="KeyValue" type="et:KeyValue" minOccurs="0" maxOccurs="unbounded"/>
        </sequence>
      </complexType>
    </element>
    <element name="MediaTime" type="et:MediaTimeType"/>
  </sequence>
</complexType>
<complexType name="SegmentMetaType">
  <sequence>
    <element name="L_SegmentKeyValues">
      <complexType>
        <sequence>
          <element name="SegmentKeyValue" type="et:SegmentKeyValue" minOccurs="0" maxOccurs="unbounded"/>
        </sequence>
      </complexType>
    </element>
  </sequence>
</complexType>
<element name="SetOfSourceItems">
  <complexType>
    <sequence>
      <element name="L_SourceItems">
        <complexType>
          <sequence>
            <element name="SourceItem" type="et:SourceItem" minOccurs="0" maxOccurs="unbounded"/>
          </sequence>
        </complexType>
      </element>
    </sequence>
  </complexType>

```

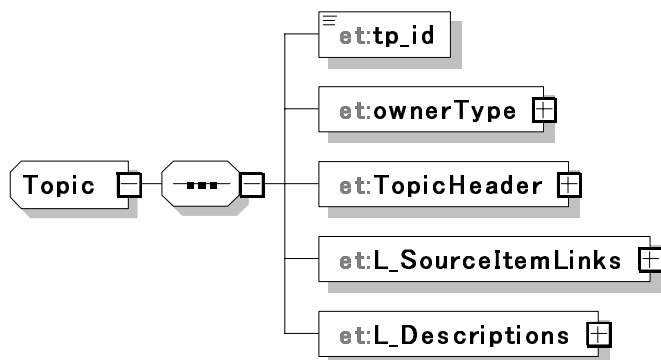
```
        </sequence>
      </complexType>
    </element>
  </sequence>
</complexType>
</element>
<element name="SetOfTopics">
  <complexType>
    <sequence>
      <element name="L_Topics">
        <complexType>
          <sequence>
            <element name="Topic" type="et:Topic" minOccurs="0" maxOccurs="unbounded"/>
          </sequence>
        </complexType>
      </element>
    </sequence>
  </complexType>
</element>
<complexType name="SourceItemLinkType">
  <sequence>
    <element name="sourceItemID" type="string"/>
    <element name="ownerType" type="et:OwnerTypeType"/>
  </sequence>
</complexType>
</schema>
```

【SourceItem】の構造

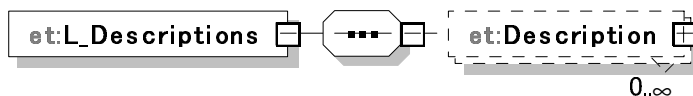


要素名称	データ型	意味
si_id	string	SourceItem の ID
ownerType	OwnerTypeType	タイプ/作成者情報
SourceItemHeader	SourceItemHeaderType	SourceItem のヘッダー情報
DelegateMeta	DelegateMetaType	SourceItem の代表メタデータ
SegmentMeata	SegmentMetaType	SourceItem のセグメントメタデータ

【Topic】の構造

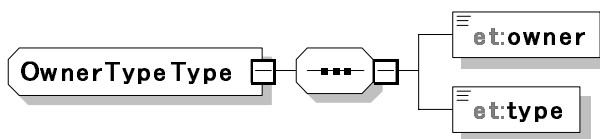


要素名称	データ型	意味
tp_id	string	Topic の ID
ownerType	OwnerTypeType	タイプ/作成者情報
TopicHeader	TopicHeaderType	Topic のヘッダ情報
L_SourceItemLinks	SourceItemLinkType	Topic を構成する SourceItem 情報
L_Descriptions	-	Topic を記述するメタデータ



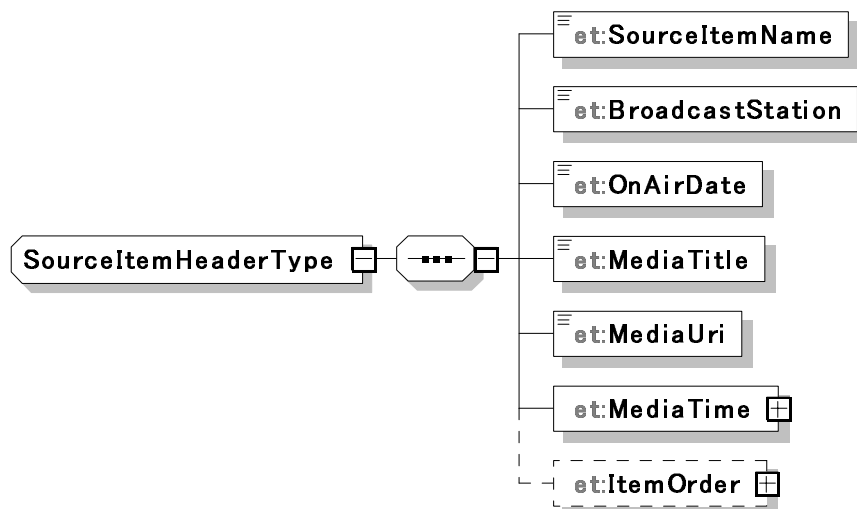
要素名称	データ型	意味
Description	KeyValueTypes	Topic を記述するメタデータ

【OwnerTypeType】の構造

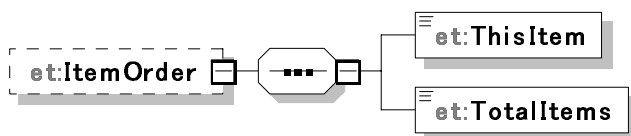


要素名称	データ型	意味
owner	string	作成者情報
type	string	タイプ情報

【SourceItemHeaderType】の構造

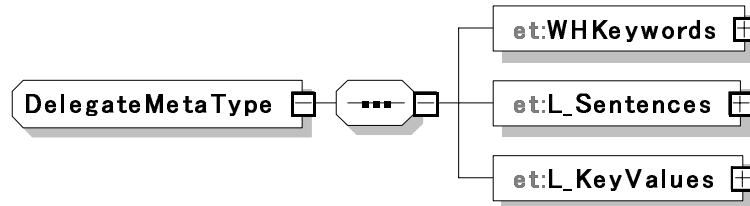


要素名称	データ型	意味
SourceItemName	string	SourceItem のタイトル
BroadcastStation	string	放送局
OnAirDate	dateTime	放送日
MediaTitle	string	番組のタイトル
MediaUri	anyURI	メディア情報
MediaTime	MediaTimeType	SourceItem のメディア上での時刻情報
ItemOrder	-	ニュース番組内での項目の位置

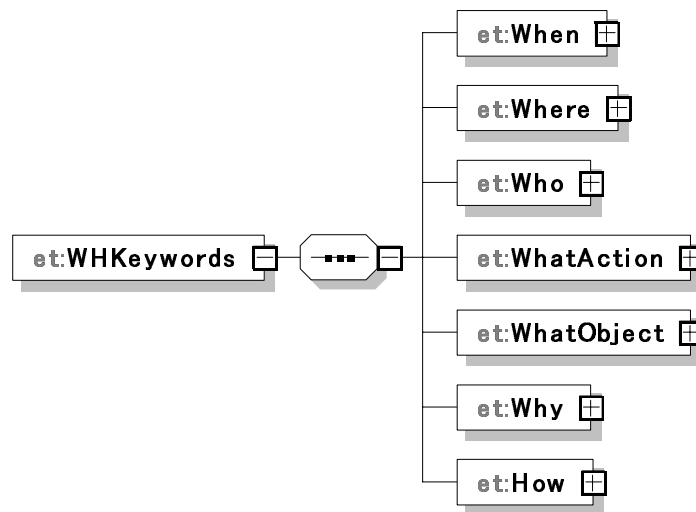


要素名称	データ型	意味
ThisItem	nonNegativeInteger	本項目の順番
TotalItems	nonNegativeInteger	ニュース番組内での全項目数

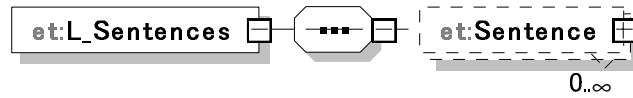
【DelegateMetaType】の構造



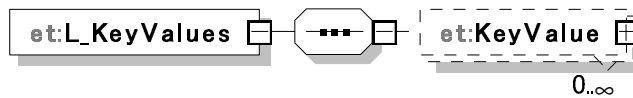
要素名称	データ型	意味
WHKeywords	-	代表キーワード
L_Sentences	-	代表文章
L_KeyValues	-	代表 Key-Value データ



要素名称	データ型	意味
When	ValueType	「いつ」に関する情報
Where	ValueType	「どこで」に関する情報
Who	ValueType	「誰」に関する情報
WhatAction	ValueType	「動作」に関する情報
WhatObject	ValueType	「オブジェクト」に関する情報
Why	ValueType	「なぜ」に関する情報
How	ValueType	「どのように」に関する情報

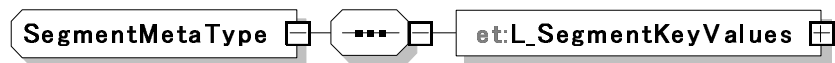


要素名称	データ型	意味
Sentence	ValueType	文章

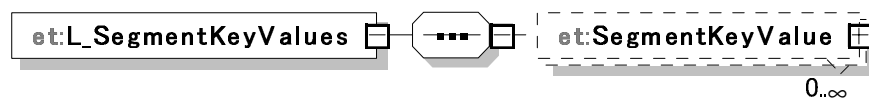


要素名称	データ型	意味
KeyValue	KeyValueType	Key-Value データ

【SegmentMetaType】の構造

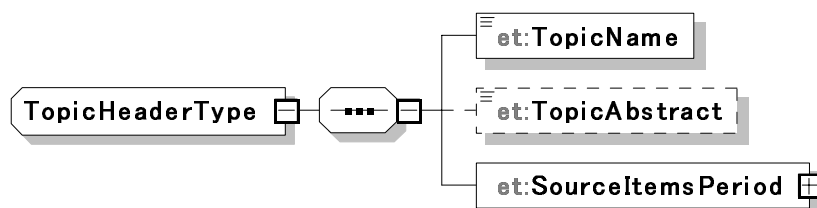


要素名称	データ型	意味
L_SegmentKeyValues	-	セグメント Key-Value データ群

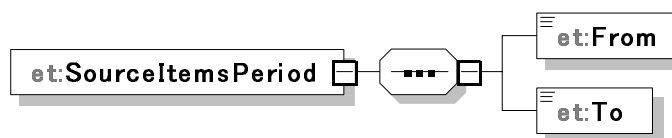


要素名称	データ型	意味
SegmentKeyValue	SegmentKeyValueType	セグメント Key-Value データ

【TopicHeaderType】の構造

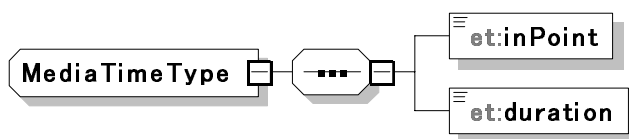


要素名称	データ型	意味
TopicName	string	Topic のタイトル
TopicAbstract	string	トピックの概要
SourceItemsPeriod	-	トピックを構成する SourceItem の期間



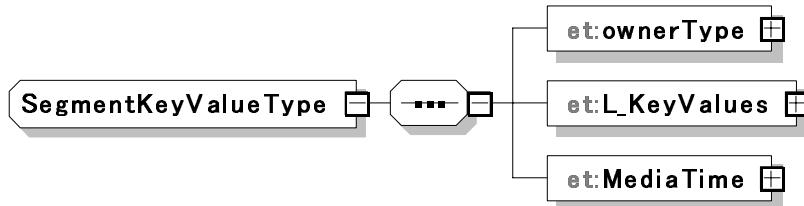
要素名称	データ型	意味
From	dateTime	Topic を構成する SourceItem の中で一番過去の日時
To	dateTime	Topic を構成する SourceItem の中で一番最近の日時

【MediaTimeType】の構造

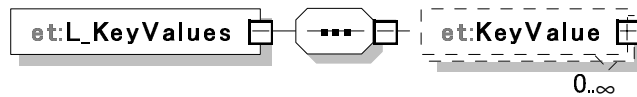


要素名称	データ型	意味
inPoint	long	区間のイン点 [ミリ秒]
duration	long	区間の長さ [ミリ秒]

【SegmentKeyValueType】の構造

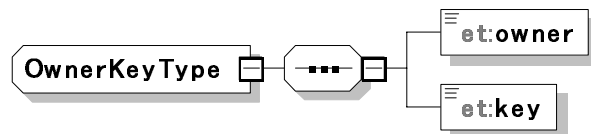


要素名称	データ型	意味
ownerType	OwnerTypeType	タイプ/作成者情報
L_KeyValues	-	Key-Value データ群
MediaTime	MediaTimeType	セグメント Key-Value データの区間情報



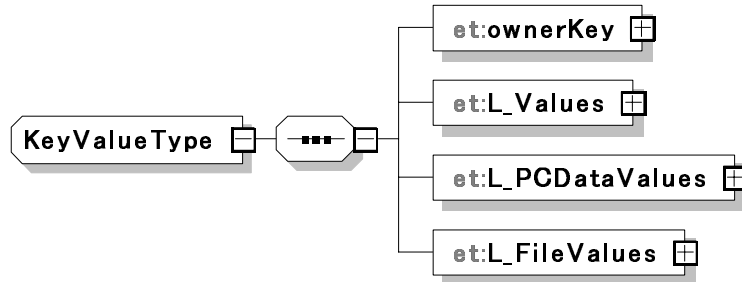
要素名称	データ型	意味
KeyValue	KeyValueType	Key-Value データ

【OwnerKeyType】の構造

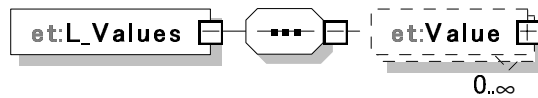


要素名称	データ型	意味
owner	string	作成者情報
key	string	Key 情報

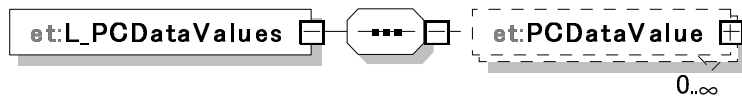
【KeyValueType】の構造



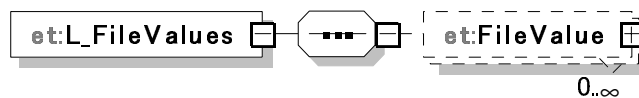
要素名称	データ型	意味
ownerKey	OwnerKeyType	Key/作成者情報
L_Values	-	Value データ (文字列)
L_PCDataValues	-	Value データ (XML 内禁則文字を含む文字列)
L_FileValues	-	Value データ (外部ファイルへの URI)



要素名称	データ型	意味
Value	ValueType	Value データ

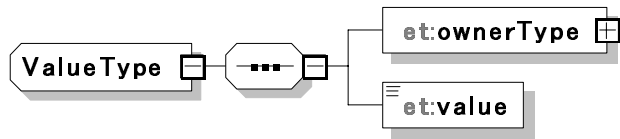


要素名称	データ型	意味
PCDataValue	ValueType	Value データ



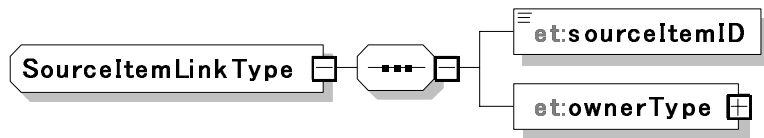
要素名称	データ型	意味
FileValue	ValueType	Value データ

【ValueType】の構造



要素名称	データ型	意味
ownerType	OwnerTypeType	タイプ/作成者情報
value	string	Value データ

【SourceItemLinkType】の構造



要素名称	データ型	意味
SourceItemID	string	リンクする SourceItem の ID
ownerType	OwnerTypeType	タイプ/作成者情報

付録C EvenTank データ操作関数

EvenTank の高レベルデータ操作関数

SourceItem 操作関数

SourceItem を扱う高レベル操作関数として、以下の 19 個の関数を定義する。

扱うデータ	関数	モード
SourceItem	searchSourceItem	read
	getSourceItem	read
	insertSourceItem	write
	deleteSourceItem	write
	replaceSourceItem	write
SourceItemHeader	getSourceItemHeader	read
	replaceSourceItemHeader	write
DelegateMeta	getDelegateMeta	read
	replaceDelegateMeta	write
SegmentMeta	getSegmentMeta	read
	replaceSegmentMeta	write
SegmentKeyValue	getSegmentKeyValue	read
	insertSegmentKeyValue	write
	deleteSegmentKeyValue	write
	replaceSegmentKeyValue	write
KeyValue	getKeyValue	read
	insertKeyValue	write
	deleteKeyValue	write
	replaceKeyValue	write

関数定義	searchSourceItem(in userID, in password, in condition)		
機能	SourceItem の検索		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：SourceItemID (複数の場合には、<ID> タグ内に複数記述)</p> <p>検索結果 0 件の場合は、空タグ (</ReturnValue>)</p> <pre><Result> <ReturnValue> <si_id>2b0d7115-5ee4-4012-aac0-f51e484cfe53</si_id> <si_id>014353b5-bff9-41d1-acf5-3eb2359dda9a</si_id> </ReturnValue> </Result></pre> <p>失敗：エラー情報</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	condition	<p>以下の 4 つの条件を、必要に応じて <Conditions> 要素内に記述する。それぞれ出現回数は高々 1 回とし、指定された複数条件は And として処理する。キーワードはスペース区切りにより複数指定可能とし、その検索条件も常に And とする。</p> <ul style="list-style-type: none"> ・日付による検索 (From, To 要素はどちらか 1 方を省略することができる) <pre><Date> <From>yyyy-mm-dd</From> <To>yyyy-mm-dd</To> </Date></pre> <ul style="list-style-type: none"> ・タイトル内を検索 <pre><InSouceItemName> keyword </InSouceItemName></pre> <ul style="list-style-type: none"> ・SourceItem の代表メタデータ内を検索 <pre><InDelegate> keyword </InDelegate></pre> <ul style="list-style-type: none"> ・SourceItem のセグメントメタデータ内を検索 <pre><InSegment> keyword1 keyword2 </InSegment></pre>

関数定義	getSourceItem(in userID, in password, in si_id)		
機能	SourceItem の取得		
戻り値	string	戻り値は下記に示す XML 書式 成功：指定された SourceItem データ <pre> <Result> <ReturnValue> <SourceItem> . . . </SourceItem> </ReturnValue> </Result> </pre> 失敗：エラー情報 「指定された SourceItem は存在しません」 <pre> <Result> <Error> エラー情報 </Error> </Result> </pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	SourceItem の ID

関数定義	insertSourceItem(in userID, in password, in sourceItem)		
機能	SourceItem の追加		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre> <Result> <ReturnValue> 0 </ReturnValue> </Result> </pre> 失敗：エラー情報 <pre> <Result> <Error> エラー情報 </Error> </Result> </pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	sourceItem	追加する SourceItem の XML データ

関数定義	deleteSourceItem(in userID, in password, in si_id)		
機能	SourceItem の削除		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> 失敗：エラー情報 「指定された SourceItem は存在しません」 <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	削除する SourceItem の ID

関数定義	replaceSourceItem(in userID, in password, in si_id, in sourceItem)		
機能	SourceItem の置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> 失敗：エラー情報 「指定された SourceItem は存在しません」 <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	置き換えられる SourceItem の ID
	string	sourceItem	置換後の SourceItem の XML データ

関数定義	getSourceItemHeader(in userID, in password, in si_id)		
機能	SourceItem のヘッダ情報取得		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> 失敗：エラー情報 「指定された SourceItem は存在しません」 <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	ヘッダ情報を取得する SourceItem の ID

関数定義	replaceSourceItemHeader(in userID, in password, in si_id, in sourceItemHeader)		
機能	SourceItem のヘッダ情報を置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> 失敗：エラー情報 「指定された SourceItem は存在しません」 <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	SourceItemHeader が置き換えられる SourceItem の ID
	string	sourceItemHeader	置換後の SourceItemHeader の XML データ

関数定義	getDelegateMeta(in userID, in password, in si_id)		
機能	SourceItem の代表メタデータ取得		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された SourceItem は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	代表メタデータを取得する SourceItem の ID

関数定義	replaceDelegateMeta(in userID, in password, in si_id, in delegateMeta)		
機能	SourceItem の代表メタデータを置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された SourceItem は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	代表メタデータが置き換えられる SourceItem の ID
	string	sourceItemHeader	置換後の代表メタデータの XML データ

関数定義	getSegmentMeta(in userID, in password, in si_id)		
機能	SourceItem のセグメントメタデータ取得		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された SourceItem は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	セグメントメタデータを取得する SourceItem の ID

関数定義	replaceSegmentMeta(in userID, in password, in si_id, in delegateMeta)		
機能	SourceItem のセグメントメタデータを置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された SourceItem は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	セグメントメタデータが置き換えられる SourceItem の ID
	string	sourceItemHeader	置換後のセグメントメタデータの XML データ

関数定義	getSegmentKeyValue(in userID, in password, in si_id, in segOwner, in segType)		
機能	指定された OwnerType の SegmentKeyValue の取得		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：指定された OwnerType の SegmentKeyValue データ</p> <pre><Result> <ReturnValue> <SegmentKeyValue>...</SegmentKeyValue> <SegmentKeyValue>...</SegmentKeyValue> </ReturnValue> </Result></pre> <p>失敗：エラー情報 「指定された OwnerType は存在しません」</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	取得対象とする SourceItem の ID
	string	segOwner	取得したい SegmentKeyValue の owner 情報
	string	segType	取得したい SegmentKeyValue の type 情報

関数定義	insertSegmentKeyValue(in userID, in password, in ListSKV)		
機能	SegmentKeyValue データの追加		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：0</p> <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> <p>失敗：エラー情報</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	ListSKV	追加する SegmentKeyValue の XML データ

関数定義	deleteSegmentKeyValue(in userID, in password, in si_id, in segOwner, in segType)		
機能	指定された OwnerType の SegmentKeyValue の削除		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> 失敗：エラー情報 「指定された OwnerType は存在しません」 <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	削除対象とする SourceItem の ID
	string	segOwner	削除したい SegmentKeyValue の owner 情報
	string	segType	削除したい SegmentKeyValue の type 情報

関数定義	replaceSegmentKeyValue(in userID, in password, in si_id, in segOwner, in segType, in ListSKV)		
機能	SourceItem の置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> 失敗：エラー情報 「指定された SourceItem は存在しません」 <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	置き換えられる SourceItem の ID
	string	segOwner	置き換えられる SegmentKeyValue の owner 情報
	string	segType	置き換えられる SegmentKeyValue の type 情報
	string	ListSKV	置換後の SegmentKeyValue の XML データ

関数定義	getKeyValue(in userID, in password, in si_id, in segOwner, in segType, in kvOwner, in kvKey)		
機能	指定された OwnerType の SegmentKeyValue 内で、指定された OwnerKey の Key-Value の取得		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：該当する KeyValue データ</p> <pre> <Result> <ReturnValue> <KeyValue>...</KeyValue> <KeyValue>...</KeyValue> </ReturnValue> </Result> </pre> <p>失敗：エラー情報</p> <p>「指定された OwnerType は存在しません」</p> <p>「指定された OwnerKey は存在しません」</p> <pre> <Result> <Error> エラー情報 </Error> </Result> </pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	取得対象とする SourceItem の ID
	string	segOwner	取得したい SegmentKeyValue の owner 情報
	string	segType	取得したい SegmentKeyValue の type 情報
	string	kvOwner	取得したい KeyValue の owner 情報
	string	kvKey	取得したい KeyValue の key 情報

関数定義	insertKeyValue(in userID, in password, in si_id, in segOwner, in segType, in ListKV)		
機能	指定された OwnerType の SegmentKeyValue 内に , KeyValue データを追加		
戻り値	string	戻り値は下記に示す XML 書式 成功 : 0 <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> 失敗 : エラー情報 <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	追加対象とする SourceItem の ID
	string	segOwner	追加したい SegmentKeyValue の owner 情報
	string	segType	追加したい SegmentKeyValue の type 情報
	string	ListKV	追加する KeyValue の XML データ

関数定義	deleteKeyValue(in userID, in password, in si_id, in segOwner, in segType, in kvOwner, in kvKey)		
機能	指定された OwnerType の SegmentKeyValue の削除		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された OwnerType は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	削除対象とする SourceItem の ID
	string	segOwner	削除したい SegmentKeyValue の owner 情報
	string	segType	削除したい SegmentKeyValue の type 情報
	string	kvOwner	削除したい KeyValue の owner 情報
	string	kvKey	削除したい KeyValue の key 情報

関数定義	replaceKeyValue(in userID, in password, in si_id, in segOwner, in segType, in kvOwner, in kvKey, in ListSKV)		
機能	KeyValue の置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された SourceItem は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	si_id	置き換えられる SourceItem の ID
	string	segOwner	置き換えられる SegmentKeyValue の owner 情報
	string	segType	置き換えられる SegmentKeyValue の type 情報
	string	ListSKV	置換後の SegmentKeyValue の XML データ
	string	kvOwner	置き換えられる KeyValue の owner 情報
	string	kvKey	置き換えられる KeyValue の key 情報
	string	ListKV	置換後の KeyValue の XML データ

Topic 操作関数

Topic を扱う高レベル操作関数として、以下の 13 個の関数を定義する。

扱うデータ	関数	モード
Topic	searchTopic	read
	getTopic	read
	insertTopic	write
	deleteTopic	write
	replaceTopic	write
TopicHeader	getTopicHeader	read
	replaceTopicHeader	write
SourceItemLink	getSourceItemLink	read
	replaceSourceItemLink	write
Description	getDescription	read
	insertDescription	write
	deleteDescription	write
	replaceDescription	write

関数定義	searchTopic(in userID, in password, in condition)		
機能	Topic の検索		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：TopicID (複数の場合には, <ID> タグ内に複数記述)</p> <p>検索結果 0 件の場合は, 空タグ (</ReturnValue>)</p> <pre><Result> <ReturnValue> <si_id>2b0d7115-5ee4-4012-aac0-f51e484cfe53</si_id> <si_id>014353b5-bff9-41d1-acf5-3eb2359dda9a</si_id> </ReturnValue> </Result></pre> <p>失敗：エラー情報</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	condition	<p>以下の 4 つの条件を, 必要に応じて <Conditions> 要素内に記述する. それぞれ出現回数は高々 1 回とし, 指定された複数条件は And として処理する. キーワードはスペース区切りにより複数指定可能とし, その検索条件も常に And とする.</p> <ul style="list-style-type: none"> ・日付による検索 (From, To 要素はどちらか 1 方を省略することができる) <pre><Date> <From>yyyy-mm-dd</From> <To>yyyy-mm-dd</To> </Date></pre> <ul style="list-style-type: none"> ・Topic の名前内を検索 <pre><InTopicName>keyword</InTopicName></pre> <ul style="list-style-type: none"> ・Topic の概要メタデータ内を検索 <pre><InAbstract>keyword1 keyword2</InAbstract></pre> <ul style="list-style-type: none"> ・Topic を形成している SourceItemID 内を検索 <pre><InSouceItemLinks>SourceItemID1</InSouceItemLinks></pre>

関数定義	getTopic(in userID, in password, in tp_id)		
機能	Topic の取得		
戻り値	string	戻り値は下記に示す XML 書式 成功：指定された Topic データ <pre> <Result> <ReturnValue> <SourceItem> . . . </SourceItem> </ReturnValue> </Result> </pre> 失敗：エラー情報 「指定された Topic は存在しません」 <pre> <Result> <Error> エラー情報 </Error> </Result> </pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	Topic の ID

関数定義	insertTopic(in userID, in password, in topic)		
機能	Topic の追加		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <pre> <Result> <ReturnValue> 0 </ReturnValue> </Result> </pre> 失敗：エラー情報 <pre> <Result> <Error> エラー情報 </Error> </Result> </pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	topic	追加する Topic の XML データ

関数定義	deleteTopic(in userID, in password, in tp_id)		
機能	Topic の削除		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された Topic 存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	削除する Topic の ID

関数定義	replaceTopic(in userID, in password, in tp_id, in sourceItem)		
機能	Topic の置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された Topic 存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	置き換えられる Topic の ID
	string	topic	置換後の Topic の XML データ

関数定義	getTopicHeader(in userID, in password, in tp_id)		
機能	Topic のヘッダ情報取得		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された Topic は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	ヘッダ情報を取得する Topic の ID

関数定義	replaceTopicHeader(in userID, in password, in tp_id, in topicHeader)		
機能	Topic のヘッダ情報を置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された Topic は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	TopicHeader が置き換えられる Topic の ID
	string	topicHeader	置換後の TopicHeader の XML データ

関数定義	getSourceItemLink(in userID, in password, in tp_id)		
機能	Topic の SourceItemLink 情報取得		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：SourceItemLink (複数の場合には , <ReturnValue> 内に複数記述)</p> <pre> <Result> <ReturnValue> <SourceItemLink>...</SourceItemLink> <SourceItemLink>...</SourceItemLink> </ReturnValue> </Result> </pre> <p>失敗：エラー情報 「指定された Topic は存在しません」</p> <pre> <Result> <Error> エラー情報 </Error> </Result> </pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	SourceItemLink 情報を取得する Topic の ID

関数定義	replaceSourceItemLink(in userID, in password, in tp_id, in ListSIL)		
機能	Topic の SourceItemLink 情報を置換		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：0</p> <pre> <Result> <ReturnValue> 0 </ReturnValue> </Result> </pre> <p>失敗：エラー情報 「指定された Topic は存在しません」</p> <pre> <Result> <Error> エラー情報 </Error> </Result> </pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	TopicHeader が置き換えられる Topic の ID
	string	ListSIL	置換後の SourceItemLink 情報の XML データ

関数定義	getDescription(in userID, in password, in tp_id, in dscpOwner, in dscpType)		
機能	指定された OwnerType の Description の取得		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：指定された OwnerType の Description データ</p> <pre><Result> <ReturnValue> <Description>...</Description> </ReturnValue> </Result></pre> <p>失敗：エラー情報 「指定された OwnerType は存在しません」</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	取得対象とする Topic の ID
	string	dscpOwner	取得したい Description の owner 情報
	string	dscpType	取得したい Description の type 情報

関数定義	insertDescription(in userID, in password, in description)		
機能	Description データの追加		
戻り値	string	<p>戻り値は下記に示す XML 書式</p> <p>成功：0</p> <pre><Result> <ReturnValue> 0 </ReturnValue> </Result></pre> <p>失敗：エラー情報</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	description	追加する Description の XML データ

関数定義	deleteDescription(in userID, in password, in tp_id, in dscpOwner, in dscpType)		
機能	指定された OwnerType の Description の削除		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された OwnerType は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	削除対象とする Topic の ID
	string	dscpOwner	削除したい Description の owner 情報
	string	dscpType	削除したい Description の type 情報

関数定義	replaceDescription(in userID, in password, in tp_id, in dscpOwner, in dscpType, in description)		
機能	Description の置換		
戻り値	string	戻り値は下記に示す XML 書式 成功：0 <Result> <ReturnValue> 0 </ReturnValue> </Result> 失敗：エラー情報 「指定された SourceItem は存在しません」 <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	tp_id	置き換えられる Topic の ID
	string	dscpOwner	置き換えられる Description の owner 情報
	string	dscpType	置き換えられる Description の type 情報
	string	description	置換後の Description の XML データ

付録D EvenTank モジュール制御関数

EvenTank のモジュール制御関数

EvenTank のモジュールを制御するための関数 4 つを以下に定義する。

関数定義	EvenTankInitializeModule(in userID, in password, in returnPoint)		
機能	処理モジュールの初期化		
戻り値	string	戻り値は XML 書式による記述 成功：処理モジュールが管理する処理プロセス ID <pre><Result> <ReturnValue>proc0021-223</ReturnValue> </Result></pre> 失敗：エラーコード <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	Any	returnPoint	コントローラへの処理データ返却ポイント
備考	<ul style="list-style-type: none"> • EvenTankInitializeModule が呼び出されたら、処理を識別するプロセス ID を返すこと。 • プロセス ID は EvenTankDoProcess で呼び出された処理が終了、もしくは中止されるまで、モジュール自身が保持管理し、処理プロセスの識別を可能にすること。 • 複数の処理を同時に行うことができない処理モジュールの場合には、1 つの処理が終わるまで新たなプロセス ID は発行せず、処理中に EvenTankInitializeModule() 関数が呼び出された場合は、エラーを返すこと。 • 第 3 引数は、モジュールの実装形態によりパラメータの型が異なる。WebServices による実装の場合には、コントローラが実装するデータ操作の WebServices ポイント、すなわち URL(string) を記述する。Windows-DLL による実装の場合には、コントローラが実装するコールバック関数の delegate ポイント (List<delegate>) を記述する。 		

関数定義	EvenTankDoProcess(in userID, in password, in procID, in operation)		
機能	処理モジュールの実行制御		
戻り値	string	<p>戻り値は XML 書式による記述</p> <p>成功：0</p> <pre><Result> <ReturnValue>0</ReturnValue> </Result></pre> <p>失敗：エラーコード</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	procID	EvenTankInitializeModule() の戻り値として取得した処理プロセス ID
	string	operation	[start—stop] 処理制御種別
備考	<ul style="list-style-type: none"> • EvenTankDoProcess で呼び出された処理が終了，もしくは中止されたら，プロセス ID の解放を行うこと． 		

関数定義	EvenTankGetModuleInfo(in userID, in password)		
機能	処理モジュールの情報取得		
戻り値	string	<p>戻り値は XML 書式による記述</p> <p>成功：処理モジュールの情報，必要とするパラメータ情報が記述された XML データ</p> <pre><Result> <ReturnValue> <ModuleInfo> <Property> ... </ReturnValue> </Result></pre> <p>失敗：エラーコード</p> <pre><Result> <Error> エラー情報 </Error> </Result></pre>	
引数	string	userID	ユーザ ID
	string	password	パスワード

関数定義	EvenTankSetModuleInfo(in userID, in password, in paramXML, in outputPoint)		
機能	処理モジュールに情報設定		
戻り値	string	戻り値は XML 書式による記述 成功：0 <Result> <ReturnValue>0</ReturnValue> </Result> 失敗：エラーコード <Result> <Error> エラー情報 </Error> </Result>	
引数	string	userID	ユーザ ID
	string	password	パスワード
	string	paramXML	EvenTankInitializeModule() の戻り値として取得した処理プロセス ID
	string	operation	EvenTankGetModuleInfo() で取得した XML 文字列に、実行に必要となる各種パラメータなどの値をセットした XML 文字列
	string	outputPoint	処理結果の戻し先 (WebServices の URL 文字列, もしくは文字列 “controller”)
備考	<ul style="list-style-type: none"> 第 5 引数が “controller” の場合は, EvenTankInitializeMoudle 関数の第 3 引数で指定されたポイントヘデータを出力し, それ以外の URL 文字列の場合は, 該当するデータ操作関数の WebServices にデータを出力する. 		

付録E EvenTank モジュール情報

EvenTank 処理モジュール情報

EvenTank の処理モジュール情報は、モジュールの書誌的情報や、処理に必要なパラメータ情報などを、実装環境に依存せず利用できるよう EvenTank 独自の XML 形式で規定したものであり、以下に示すスキーマで定義される。

EvenTank の処理モジュール情報のスキーマ

```
<xs:schema xmlns:xs="http://www.w3.org/2001/XMLSchema" xmlns="http://www.nhk.or.jp/str1/evenTankModule"
xmlns:ns1="http://www.nhk.or.jp/str1/evenTankModule"
targetNamespace="http://www.nhk.or.jp/str1/evenTankModule" elementFormDefault="qualified"
attributeFormDefault="unqualified">
  <xs:element name="ModuleInfo">
    <xs:annotation>
      <xs:documentation>EvenTank モジュールの情報</xs:documentation>
    </xs:annotation>
    <xs:complexType>
      <xs:sequence>
        <xs:element name="Property" type="PropertyType"/>
        <xs:element name="Arguments" type="ArgumentsType"/>
      </xs:sequence>
    </xs:complexType>
  </xs:element>
  <xs:complexType name="PropertyType">
    <xs:annotation>
      <xs:documentation>モジュールに関する書誌的情報を記述</xs:documentation>
    </xs:annotation>
    <xs:sequence>
      <xs:element name="Description" type="DescriptionStringType"/>
      <xs:element name="Creator" type="DescriptionStringType" minOccurs="0"/>
      <xs:element name="Date" type="DescriptionStringType" minOccurs="0"/>
      <xs:element name="Rights" type="DescriptionStringType" minOccurs="0"/>
      <xs:element name="ReferencePaper" type="DescriptionStringType" minOccurs="0" maxOccurs="unbounded"/>
      <xs:element name="Version" type="DescriptionStringType"/>
      <xs:element name="InputData">
        <xs:complexType>
          <xs:attribute name="SourceItem" type="xs:boolean"/>
          <xs:attribute name="Topic" type="xs:boolean"/>
          <xs:attribute name="others" type="xs:boolean"/>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:schema>
```

```

<xs:element name="OutputData">
  <xs:complexType>
    <xs:attribute name="SourceItem" type="xs:boolean"/>
    <xs:attribute name="Topic" type="xs:boolean"/>
    <xs:attribute name="others" type="xs:boolean"/>
  </xs:complexType>
</xs:element>
<xs:element name="Interface">
  <xs:complexType>
    <xs:attribute name="webServices" type="xs:boolean" default="true"/>
  </xs:complexType>
</xs:element>
</xs:sequence>
</xs:complexType>
<xs:complexType name="ArgumentsType">
  <xs:annotation>
    <xs:documentation>モジュールを制御するための引数を記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="ParameterList" type="ParameterListType" minOccurs="0"/>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="ParameterListType">
  <xs:annotation>
    <xs:documentation>モジュールの処理に必要な全パラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:choice maxOccurs="unbounded">
    <xs:element name="Parameter" type="ParameterType"/>
    <xs:element name="Group" type="GroupType"/>
  </xs:choice>
</xs:complexType>
<xs:complexType name="GroupType">
  <xs:annotation>
    <xs:documentation>パラメータをグループ化する記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Label" type="LabelStringType"/>
    <xs:element name="ParameterList" type="ParameterListType" maxOccurs="unbounded"/>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="ParameterType">
  <xs:annotation>
    <xs:documentation>個別のパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Label" type="LabelStringType"/>
    <xs:element name="Description" type="DescriptionStringType"/>
    <xs:choice>
      <xs:element name="MediaTimeRange" type="MediaTimeRangeType"/>
      <xs:element name="MediaTimePoint" type="MediaTimePointType"/>
      <xs:element name="MediaTimeDuration" type="MediaTimeDurationType"/>
      <xs:element name="URI" type="URIType"/>
      <xs:element name="String" type="StringType"/>
      <xs:element name="Integer" type="IntegerType"/>
      <xs:element name="Double" type="DoubleType"/>
      <xs:element name="Select" type="SelectType"/>
      <xs:element name="SourceItem" type="SourceItemType"/>
      <xs:element name="Topic" type="TopicType"/>
      <xs:element name="SegmentKeyValue" type="SegmentKeyValueType"/>
      <xs:element name="SegmentKeyValueInSpecificSI" type="SegmentKeyValueInSpecificSIType"/>
      <xs:element name="KeyValue" type="KeyValueType"/>
    </xs:choice>
  </xs:sequence>
</xs:complexType>

```

```

        <xs:element name="KeyValueInSpecificSI" type="KeyValueInSpecificSIType"/>
        <xs:element name="Value" type="ValueType"/>
        <xs:element name="ValueInSpecificSI" type="ValueInSpecificSIType"/>
    </xs:choice>
</xs:sequence>
<xs:attribute name="pid" type="xs:ID" use="required"/>
</xs:complexType>
<xs:complexType name="DescriptionStringType">
    <xs:annotation>
        <xs:documentation>日英によるテキスト記述</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="DescriptionStringJa" type="xs:string" minOccurs="0"/>
        <xs:element name="DescriptionStringEn" type="xs:string"/>
    </xs:sequence>
</xs:complexType>
<xs:complexType name="LabelStringType">
    <xs:annotation>
        <xs:documentation>日英による 3 2 文字以内のテキスト記述</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="LabelStringJa" type="labelType" minOccurs="0"/>
        <xs:element name="LabelStringEn" type="labelType"/>
    </xs:sequence>
</xs:complexType>
<xs:simpleType name="labelType">
    <xs:annotation>
        <xs:documentation>3 2 文字以内のテキスト</xs:documentation>
    </xs:annotation>
    <xs:restriction base="xs:string">
        <xs:maxLength value="32"/>
    </xs:restriction>
</xs:simpleType>
<xs:complexType name="MediaTimeRangeType">
    <xs:annotation>
        <xs:documentation>時間区間を記述</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="MediaTimePoint" type="MediaTimePointType"/>
        <xs:element name="MediaTimeDuration" type="MediaTimeDurationType"/>
    </xs:sequence>
</xs:complexType>
<xs:complexType name="MediaTimePointType">
    <xs:annotation>
        <xs:documentation>時刻を記述</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="Value" nillable="true">
            <xs:simpleType>
                <xs:restriction base="xs:long">
                    <xs:minInclusive value="0"/>
                </xs:restriction>
            </xs:simpleType>
        </xs:element>
        <xs:element name="InitValue" minOccurs="0">
            <xs:simpleType>
                <xs:restriction base="xs:long">
                    <xs:minInclusive value="0"/>
                </xs:restriction>
            </xs:simpleType>
        </xs:element>
    </xs:sequence>

```



```
</xs:sequence>
</xs:complexType>
<xs:complexType name="MediaTimeDurationType">
  <xs:annotation>
    <xs:documentation>継続期間を記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value" nillable="true">
      <xs:simpleType>
        <xs:restriction base="xs:long">
          <xs:minInclusive value="0"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:element>
    <xs:element name="InitValue" minOccurs="0">
      <xs:simpleType>
        <xs:restriction base="xs:long">
          <xs:minInclusive value="0"/>
        </xs:restriction>
      </xs:simpleType>
    </xs:element>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="StringType">
  <xs:annotation>
    <xs:documentation>文字列タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value" type="xs:string" nillable="true"/>
    <xs:element name="InitValue" type="xs:string" minOccurs="0"/>
  </xs:sequence>
  <xs:attribute name="minChar" type="xs:integer"/>
  <xs:attribute name="maxChar" type="xs:integer"/>
</xs:complexType>
<xs:complexType name="URIType">
  <xs:annotation>
    <xs:documentation>U R I タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value" type="xs:anyURI" nillable="true"/>
    <xs:element name="InitValue" type="xs:anyURI" minOccurs="0"/>
    <xs:element name="Filter" minOccurs="0">
      <xs:complexType>
        <xs:simpleContent>
          <xs:extension base="extensionListType">
            <xs:attribute name="video" type="xs:boolean"/>
            <xs:attribute name="audio" type="xs:boolean"/>
            <xs:attribute name="text" type="xs:boolean"/>
          </xs:extension>
        </xs:simpleContent>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
  <xs:attribute name="type" use="required">
    <xs:simpleType>
      <xs:restriction base="xs:string">
        <xs:enumeration value="file"/>
        <xs:enumeration value="directory"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>

```

```

<xs:attribute name="direction" use="required">
  <xs:simpleType>
    <xs:restriction base="xs:string">
      <xs:enumeration value="input"/>
      <xs:enumeration value="output"/>
    </xs:restriction>
  </xs:simpleType>
</xs:attribute>
</xs:complexType>
<xs:simpleType name="extensionListType">
  <xs:annotation>
    <xs:documentation>拡張子のリストタイプ</xs:documentation>
  </xs:annotation>
  <xs:list itemType="xs:string"/>
</xs:simpleType>
<xs:complexType name="IntegerType">
  <xs:annotation>
    <xs:documentation>整数タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="UnitName" type="LabelStringType" minOccurs="0"/>
    <xs:element name="Value" type="xs:integer" nillable="true"/>
    <xs:element name="InitValue" type="xs:integer" minOccurs="0"/>
  </xs:sequence>
  <xs:attribute name="min" type="xs:integer"/>
  <xs:attribute name="max" type="xs:integer"/>
</xs:complexType>
<xs:complexType name="DoubleType">
  <xs:annotation>
    <xs:documentation>小数タイプのパラメータ記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="UnitName" type="LabelStringType" minOccurs="0"/>
    <xs:element name="Value" type="xs:double" nillable="true"/>
    <xs:element name="InitValue" type="xs:double" minOccurs="0"/>
  </xs:sequence>
  <xs:attribute name="min" type="xs:double"/>
  <xs:attribute name="max" type="xs:double"/>
</xs:complexType>
<xs:complexType name="SelectType">
  <xs:annotation>
    <xs:documentation>選択タイプのパラメータ記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Option" type="OptionType" minOccurs="2" maxOccurs="unbounded"/>
    <xs:element name="Value" type="xs:string" nillable="true" maxOccurs="unbounded"/>
    <xs:element name="InitValue" type="xs:string" maxOccurs="unbounded"/>
  </xs:sequence>
  <xs:attribute name="minSelect" type="xs:nonNegativeInteger" use="required"/>
  <xs:attribute name="maxSelect" type="xs:nonNegativeInteger" use="required"/>
</xs:complexType>
<xs:complexType name="OptionType">
  <xs:annotation>
    <xs:documentation>選択肢を記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Label" type="LabelStringType"/>
    <xs:element name="DelegateValue" type="xs:NMTOKEN"/>
    <xs:element name="ParameterList" type="ParameterListType" minOccurs="0" maxOccurs="unbounded"/>
  </xs:sequence>
</xs:complexType>

```

```

<xs:complexType name="SourceItemType">
  <xs:annotation>
    <xs:documentation>SourceItem タイプのパラメータをを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <!--
    <xs:element name="TargetWS" type="URIType"/>
    -->
    <xs:element name="Value">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="si_id" type="xs:token"/>
          <xs:element name="locatedWS" type="xs:string" minOccurs="0"/>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="TopicType">
  <xs:annotation>
    <xs:documentation>Topic タイプのパラメータをを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <!--
    <xs:element name="TargetWS" type="URIType"/>
    -->
    <xs:element name="Value">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="tp_id" type="xs:token"/>
          <xs:element name="locatedWS" type="xs:string" minOccurs="0"/>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="ValueLocatorType">
  <xs:annotation>
    <xs:documentation>Value データの位置を記述</xs:documentation>
  </xs:annotation>
  <xs:choice>
    <!--
    <xs:element name="TargetWS" type="URIType"/>
    -->
    <xs:element name="DelegateMeta">
      <xs:complexType>
        <xs:choice>
          <xs:element name="WHS">
            <xs:complexType>
              <xs:sequence>
                <xs:element name="owner" type="xs:string"/>
                <xs:element name="type" type="xs:string"/>
              </xs:sequence>
              <xs:attribute name="elementName">
                <xs:simpleType>
                  <xs:restriction base="xs:string">
                    <xs:enumeration value="When"/>
                    <xs:enumeration value="Where"/>
                    <xs:enumeration value="Who"/>
                    <xs:enumeration value="WhatAction"/>
                    <xs:enumeration value="WhatObject"/>
                  </xs:restriction>
                </xs:simpleType>
              </xs:attribute>
            </xs:complexType>
          </xs:element>
        </xs:choice>
      </xs:complexType>
    </xs:element>
  </xs:choice>
</xs:complexType>

```

```

        <xs:enumeration value="Why"/>
        <xs:enumeration value="How"/>
        <xs:enumeration value="Sentence"/>
      </xs:restriction>
    </xs:simpleType>
  </xs:attribute>
</xs:complexType>
</xs:element>
<xs:element name="KV">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="key" type="xs:string"/>
      <xs:element name="type" type="xs:string"/>
      <xs:element name="V">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="owner" type="xs:string"/>
            <xs:element name="type" type="xs:string"/>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
</xs:choice>
</xs:complexType>
</xs:element>
<xs:element name="SKV">
  <xs:complexType>
    <xs:sequence>
      <xs:element name="owner" type="xs:string"/>
      <xs:element name="type" type="xs:string"/>
      <xs:element name="KV">
        <xs:complexType>
          <xs:sequence>
            <xs:element name="key" type="xs:string"/>
            <xs:element name="type" type="xs:string"/>
            <xs:element name="V">
              <xs:complexType>
                <xs:sequence>
                  <xs:element name="owner" type="xs:string"/>
                  <xs:element name="type" type="xs:string"/>
                </xs:sequence>
              </xs:complexType>
            </xs:element>
          </xs:sequence>
        </xs:complexType>
      </xs:element>
    </xs:sequence>
  </xs:complexType>
</xs:element>
</xs:choice>
</xs:complexType>
<xs:complexType name="ValueType">
  <xs:annotation>
    <xs:documentation>Value タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value" type="ValueLocatorType"/>
  </xs:sequence>
</xs:complexType>

```

```

<xs:complexType name="ValueInSpecificSIType">
  <xs:annotation>
    <xs:documentation>特定の SI 内の Value タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value">
      <xs:complexType>
        <xs:complexContent>
          <xs:extension base="ValueLocatorType">
            <xs:attribute name="locatedWS" type="xs:anyURI"/>
            <xs:attribute name="si_id" type="xs:string" use="required"/>
          </xs:extension>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="KeyValueLocatorType">
  <xs:annotation>
    <xs:documentation>Key-Value データの位置を記述</xs:documentation>
  </xs:annotation>
  <xs:choice>
    <!--
    <xs:element name="TargetWS" type="URIType"/>
    -->
    <xs:element name="DelegateMeta">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="KV">
            <xs:complexType>
              <xs:sequence>
                <xs:element name="key" type="xs:string"/>
                <xs:element name="type" type="xs:string"/>
              </xs:sequence>
            </xs:complexType>
          </xs:element>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
    <xs:element name="SKV">
      <xs:complexType>
        <xs:sequence>
          <xs:element name="owner" type="xs:string"/>
          <xs:element name="type" type="xs:string"/>
          <xs:element name="KV">
            <xs:complexType>
              <xs:sequence>
                <xs:element name="key" type="xs:string"/>
                <xs:element name="type" type="xs:string"/>
              </xs:sequence>
            </xs:complexType>
          </xs:element>
        </xs:sequence>
      </xs:complexType>
    </xs:element>
  </xs:choice>
</xs:complexType>
<xs:complexType name="KeyValueType">
  <xs:annotation>
    <xs:documentation>Key-Value タイプのパラメータを記述</xs:documentation>
  </xs:annotation>

```

```

<xs:sequence>
  <xs:element name="Value" type="KeyValueLocatorType"/>
</xs:sequence>
</xs:complexType>
<xs:complexType name="KeyValueInSpecificSIType">
  <xs:annotation>
    <xs:documentation>特定の SI 内の Key-Value タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value">
      <xs:complexType>
        <xs:complexContent>
          <xs:extension base="KeyValueLocatorType">
            <xs:attribute name="locatedWS" type="xs:anyURI"/>
            <xs:attribute name="si_id" type="xs:string" use="required"/>
          </xs:extension>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="SegmentKeyValueLocatorType">
  <xs:annotation>
    <xs:documentation>セグメント Key-Value データの位置を記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <!--
    <xs:element name="TargetWS" type="URIType"/>
    -->
    <xs:element name="owner" type="xs:string"/>
    <xs:element name="type" type="xs:string"/>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="SegmentKeyValueInSpecificSIType">
  <xs:annotation>
    <xs:documentation>セグメント Key-Value タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value" type="SegmentKeyValueLocatorType"/>
  </xs:sequence>
</xs:complexType>
<xs:complexType name="SegmentKeyValueInSpecificSIType">
  <xs:annotation>
    <xs:documentation>特定の SI 内のセグメント Key-Value タイプのパラメータを記述</xs:documentation>
  </xs:annotation>
  <xs:sequence>
    <xs:element name="Value">
      <xs:complexType>
        <xs:complexContent>
          <xs:extension base="SegmentKeyValueLocatorType">
            <xs:attribute name="locatedWS" type="xs:anyURI"/>
            <xs:attribute name="si_id" type="xs:string" use="required"/>
          </xs:extension>
        </xs:complexContent>
      </xs:complexType>
    </xs:element>
  </xs:sequence>
</xs:complexType>
</xs:schema>

```

参考文献

- [1] Yasuyuki Nakajima, Kiyono Ujihara, and Tamotsu Kanoh, “Video Structure Analysis and Its Application to Creation of Video Summary,” 信学技報 IE95-66, pp.35-42, Oct. 1995
- [2] 工藤大樹, 西川博文, 加藤嘉明, “ニュース番組の要約作成に関する検討,” 2006年信学ソ大(基礎・境界), A-4-4, p.71, Sep. 2006
- [3] Yihong Gong, and Xin Liu, “Video Summarization Using Singular Value Decomposition,” Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’00), Volume 2, pp.2174-2180, Hilton Head, SC, USA, Jun. 2000
- [4] 伊藤一成, 藤原司, 橋田浩一, “セマンティックビデオオーサリングによるニュース動画群からのダイジェスト生成,” 人工知能学 KBS 研報 DBS-A601, Jul. 2006
- [5] 伊藤一成, 酒井康旭, 齋藤博昭, “メタデータ解析と自然言語処理を併用した動画要約,” 情処学 DBS 研報 2004-DBS-132, pp.41-48, Jan. 2004
- [6] 伊藤一成, 酒井康旭, 齋藤博昭, “音声と映像の一貫性を考慮したニュース動画の要約,” 日本データベース学会 Letter, Vol.3, No.1, pp.1-4, Jun. 2004
- [7] 伊藤一成, 齋藤博昭, “汎用アノテーション記述言語 MAML,” 情処学 FI 研報 FI74-9, pp.63-69, Mar. 2004
- [8] 橋田浩一, “GDA 意味的修飾に基づく多用途の知的コンテンツ,” 人工知能学誌, Vol.13, No.4, pp.528-535, Jul. 1998
- [9] 林英俊, 李龍, 上林弥彦, “概念グラフを用いたニュース映像要約システムの構築,” データ工学ワークショップ (DEWS), 4-A-03, Mar. 2003
- [10] P. Duygulu, J.-Y. Pan, and D.A. Forsyth, “Towards Auto-Documentary: Tracking the Evolution of News Stories,” Proceedings of ACM International Conference on Multimedia 2004, pp.820-827, New York, NY, USA, Oct. 2004

- [11] NIST (National Institute of Standards, and Technology,) TREC (Text RE-trival Conference) Home Page, <http://trec.nist.gov/>
- [12] NII (National Institute of Informatics,) NTCIR (NII Test Collection For IR Systems) Project, <http://research.nii.ac.jp/ntcir/>
- [13] H. Yang, L. Chaisorn, Y. Zhao, S.-Y. Neo, and T.-S. Chua, “VideoQA: Question Answering on News Video,” Proceedings of ACM International Conference on Multimedia 2003, pp.632-641, Berkeley, CA, USA, 2003
- [14] S.-Y. Neo, Y. Ran, H.-K. Goh, Y. Zheng, T.-S. Chua, and J. Li, “The Use of Topic Evolution to Help Users Browse and Find Answers in News Video Corpus,” Proceedings of ACM International Conference on Multimedia 2007, pp.198-207, Augsburg, Bavaria, Germany, Sep. 2007
- [15] Rong Yan, Jun Yang, and Alexander G. Hauptmann, “Learning Query-Class Dependent Weights in Automatic Video Retrieval,” Proceedings of ACM International Conference on Multimedia 2004, pp.548-555, New York, NY, USA, Oct. 2004
- [16] Carnegie Mellon University, Informedia Project, <http://www.informedia.cs.cmu.edu/>
- [17] A. Aho, S.-F. Chang, K. McKeown, D. Radev, J. Smith, and K. Zaman, “Columbia Digital News System: An Environment for Briefing and Search over Multimedia Information,” Proceedings of IEEE International Forum on Research and Technology Advances in Digital Library, pp.82-94, Washington, D.C., USA, May 1997
- [18] M. Worring, C.G.M. Snoek, O. de Rooij, G.P. Nguyen, and A.W.M. Smeulders, “The MediaMill: semantic video search engine,” Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol.4, pp.1213-1216, Honolulu, Hawaii, USA, Apr. 2007
- [19] M.G. Christel, A.G. Hauptmann, H.D. Wactlar, and T.D. Ng, “Collages as Dynamic Summaries for News Video,” Proceedings of ACM International Conference on Multimedia 2002, pp.561-569, Juan-les-Pins, France, Dec. 2002

- [20] 井手一郎, 木下智義, 高橋友和, 孟洋, 片山紀生, 佐藤真一, 村瀬洋, “大量ニュース映像を対象とした時系列意味構造に基づく情報編纂手法の提案” 人工知能学会論文誌, Vol.23, No.5, SP-A, pp.282-292, 2008
- [21] Google Image Search, <http://images.google.co.jp/imghp?hl=ja&tab=wi>
- [22] YAHOO 画像検索, <http://image-search.yahoo.co.jp/>
- [23] Dublin Core Metadata Initiative, <http://dublincore.org/>
- [24] ISO/IEC Information Centre,
<http://www.standardsinfo.net/info/livelihood/link/fetch/2000/148478/6301438/index.html>
- [25] “Information technology – Multimedia content description interface –,” ISO/IEC 15938
- [26] XML Schema, <http://www.w3.org/XML/Schema>
- [27] “Information technology – Multimedia content description interface – Part 2: Description definition language,” ISO/IEC 15938-2:2002
- [28] “Information technology – Multimedia content description interface – Part 3: Visual,” ISO/IEC 15938-3:2002
- [29] “Information technology – Multimedia content description interface – Part 4: Audio,” ISO/IEC 15938-4:2002
- [30] “Information technology – Multimedia content description interface – Part 5: Multimedia description schemes,” ISO/IEC 15938-5:2003
- [31] “Information technology – Multimedia content description interface – Part 9: Profiles and levels,” ISO/IEC 15938-9:2005
- [32] International Press Telecommunications Council (IPTC,) <http://www.iptc.org/cms/site/index.html>
- [33] M. Sano, Y. Kawai, H. Sumiyoshi and N. Yagi, “Metadata Production Framework and Metadata Editor, ” Proceedings of ACM International Conference on Multimedia 2006, pp.789-790, Santa Barbara, CA, USA, Oct. 2006

- [34] 佐野雅規, 河合吉彦, 住吉英樹, 八木伸行, “メタデータ制作フレームワーク ネットワークシステムへの拡張,” 映情学技報 ME2007-31, Vol.30, No.10, p.171-176, Feb. 2007
- [35] L. Hardman, Z. Obrenovic, F. Nack, B.Kerherve and K. Piersol, “Canonical processes of semantically annotated media production,” Multimedia Systems, Vol.14, No.6, pp.327-340, Dec. 2008
- [36] “Information technology – Open Distributed processing – Unified Modeling Language (UML) Version 1.4.2,” ISO/IEC 19501:2005
- [37] NHK Science and Technical Research Laboratories, Metadata Production Framework, <http://www.nhk.or.jp/strl/mpf/>
- [38] 佐野雅規, 柴田正啓, 八木伸行, 佐藤真一, “蓄積したニュース番組を活用するフレームワーク ~ EvenTank への構想 ~,” 映情学冬大, 7-7, Dec. 2008
- [39] J. Allan, Topic Detection and Tracking: Event-based Information Organization, Springer, Feb. 2002
- [40] 井手一郎, 孟洋, 片山紀生, 佐藤真一, “トピックに基づく大規模ニュース映像コーパスの構造化,” 第2回情報科学技術フォーラム一般講演論文集 (FIT 2003), D-12, pp.23-25, Sep. 2003
- [41] 井手一郎, 木下智義, 高橋友和, 佐藤真一, 村瀬洋, “mediaWalker: 時系列意味構造に基づく映像アーカイブ探索インターフェース,” 画像の認識・理解シンポジウム (MIRU 2008), DS-5, pp.1670-1671, Jul. 2008
- [42] W. H. Hsu and S.-F. Chang, “Topic Tracking Across Broadcast News Video with Visual Duplicates and Semantic Concepts,” Proceedings of IEEE International Conference on Image Processing, pp.141-144, Atlanta, GA, USA, Oct. 2006
- [43] D. Swanberg, C.-F. Shu, and R. Jain, “Knowledge Guided Parsing in Video Databases,” Proceedings of SPIE, Vol.1908, pp.13-24, San Jose, CA, USA, Feb. 1993
- [44] H. Zhang, Y. Gong, S.W. Smoliar, and S.Y. Tan, “Automatic Parsing of News Video,” Proceedings of IEEE International Conference on Multimedia Computing and Systems, pp.45-54, Boston, MA, USA, May 1994

- [45] S. Eickeler, and S. Muller, "Content-based Video Indexing of TV Broadcast News Using Hidden Markov Models," IProceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, pp.2997-3000, Phoenix, AZ, USA, Mar. 1999
- [46] T. Sato, T. Kanade, E.K. Hughes, and M.A. Smith, "Video OCR for Digital News Archive," Proceedings of IEEE International Workshop on Content-Based Access of Image and Video Databases, pp.52-60, Bombay, India, Jan. 1998
- [47] M.J. Witbrock, and A.G. Hauptmann, "Speech Recognition for a Digital Video Library," Journal of the American Society for Information Science. Vol.49, Issue 7, pp.619-632, May 1998
- [48] A. Merlino, D. Morey, and M. Maybury, "Broadcast news navigation using story segmentation," Proceedings of ACM International Conference on Multimedia 1997, pp.381-391, Seattle, WA, USA, Nov. 1997
- [49] Y. Nakamura, and T. Kanade, "Semantic Analysis for Video Contents Extraction — Spotting by Association in News Video," Proceedings of ACM International Conference on Multimedia 1997, pp.393-401, Seattle, WA, USA, Nov. 1997
- [50] 井手一郎, 山本晃司, 浜田玲子, 田中英彦, "ショット分類に基づく映像への自動的索引付け手法," 信学論 D-II, Vol.J82-D-II, No.10, pp.1543-1551, Oct. 1999
- [51] S. Satoh, Y. Nakamura, and T. Kanade, "Name-It: Naming and Detecting Faces in News Videos," IEEE Multimedia, Vol.6, Issue 1, pp.22-35, Jan. 1999
- [52] 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," 信学論 D-II, Vol.84-D-II, No.8, pp.1856-1863, Aug. 2001
- [53] X. Gao, and X. Tang, "Unsupervised Video-Shot Segmentation and Model-Free Anchorperson Detection for News Video Story Parsing," IEEE Transactions on Circuits and Systems for Video Technology, Vol.12, No.9, pp.765-776, Sep. 2002

- [54] M. Bertini, A. Del Bimbo, and P. Pala, "Content-based indexing and retrieval of TV news," *Pattern Recognition Letters*, Vol.22, Issue 5, pp.503-516, Apr. 2001
- [55] A. Hanjalic, R.L. Lagendijk, and J. Biemond, "Semi-Automatic News Analysis, Indexing and Classification System based on Topic Preselection," *Proceedings of SPIE*, Vol.3656, pp.86-97, San Jose, CA, USA, Jan. 1999
- [56] M.D. Santo, G. Percannella, C. Sansone, and M. Vento, "A Comparison of Unsupervised Shot Classification Algorithms for News Video Segmentation," *Proceedings of International Workshop on Structural and Syntactic Pattern Recognition, Lecture Note in Computer Science*, Vol.3138, pp.233-241, Lisbon, Portugal, Aug. 2004
- [57] M.D. Santo, G. Percannella, C. Sansone, and M. Vento, "A Multi-expert Approach for Shot Classification in News Videos," *Proceedings of International Workshop on Image Analysis and Recognition, Lecture Note in Computer Science*, Vol.3211, pp.564-571, Porto, Portugal, Sep.-Oct. 2004
- [58] NIST (National Institute of Standards, and Technology,) TRECVID (TREC Video Retrieval Evaluation) Home Page, <http://www-nlpir.nist.gov/projects/trecvid/>
- [59] T.-S. Chua, S.-Y. Neo, K.-Y. Li, G. Wang, R. Shi, M. Zhao, and H. Xu, "TRECVID 2004 Search and Feature Extraction Task by NUS PRIS," *TREC Video Retrieval Evaluation Online Proceedings*, <http://www-nlpir.nist.gov/projects/tvpubs/tv.pubs.org.html>
- [60] J. Fan, A.K. Elmagarmid, X. Zhu, W.G. Aref, and L. Wu, "Class View: Hierarchical Video Shot Classification, Indexing, and Accessing," *IEEE Transactions on Multimedia*, Vol.6, No.1, pp.70-86, Feb. 2004
- [61] M. Bertini, A.D. Bimbo, and P.Pala, "Indexing for reuse of TV news shots," *Pattern Recognition*, Vol.35, Issue 3, pp.581-591, Mar. 2002
- [62] H. J. Nock, G. Iyengar and C. Neti, "Assessing Face and Speech Consistency for Monologue Detection in Video," *Proceedings of ACM International Conference on Multimedia*, pp.303-306, Juan Les Pins, France, Dec. 2002

- [63] C. G. M. Snoek and M. Worring, "Detection of TV News Monologues by Style Analysis, " Proceedings of IEEE International Conference on Multimedia and Expo, Vol.2, pp.1103-1106, Taipei, Taiwan, Jul. 2004
- [64] 關岡直城, 高橋友和, 井手一郎, 村瀬洋, "ニュース映像中のモノローグシーン検出による発言集の自動作成, "信学技報 PRMU2005-301, Vol.105, No.674, pp277-282, Mar. 2006
- [65] OpenCV, <http://opencv.willowgarage.com/wiki>
- [66] NHK , おこめ , <http://www.nhk.or.jp/okome/>
- [67] NHK , 南極キッズ , <http://www.nhk.or.jp/school/nankyoku-kids/ja/frame.html>
- [68] TV-Anytime Forum, <http://www.tv-anytime.org/>
- [69] NHK , メタデータ制作フレームワーク (MPF: Metadata Production Framework,) <http://www.nhk.or.jp/str1/mpf/>
- [70] K. Miura, I. Yamada, H. Sumiyoshi and N. Yagi, "Automatic Generation of a Multimedia Encyclopedia from TV Programs by Using Closed Captions and Detecting Principal Video Objects," Proceedings of IEEE International Symposium on Multimedia, pp.873-880, San Diego, CA, USA, Dec. 2006
- [71] Y. Kawai, H. Sumiyoshi and N. Yagi, "Automated Production of TV Program Trailer using Electronic Program Guide," Proceedings of ACM International Conference on Image and Video Retrieval, pp.49-56, Amsterdam, The Netherlands, Jul. 2007
- [72] YAHOO , みんなの検定 , <http://minna.cert.yahoo.co.jp/>
- [73] R. Higashinaka, K. Dohsaka and H. Isozaki, "Learning to Rank Definitions to Generate Quizzes for Interactive Information Presentation," Proceedings of Annual Meeting of the Association for Computational Linguistics, pp.117-120, Prague, Czech Republic, Jun. 2007
- [74] T. Misu and T. Kawahara, "An Interactive Framework for Document Retrieval and Presentation with Question-Answering Function in Restricted Domain," Proceedings of IEA/AIE International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems, pp.126-134, Kyoto, Japan, Jun. 2007

- [75] R. Higashinaka, K. Dohsaka, S. Amano and H. Isozaki, "Effects of Quiz-style Information Presentation on User Understanding," Proceedings of Interspeech, pp.2725-2728, Antwerp, Belgium, Aug. 2007
- [76] T. Paek, Y.-C. Ju and C. Meek, "People Watcher: A Game for Eliciting Human-Transcribed Data for Automated Directory Assistance," Proceedings of Interspeech, pp.1322-1325, Antwerp, Belgium, Aug. 2007
- [77] NIST (National Institute of Standards, and Technology,) TREC Question Answering Track, <http://trec.nist.gov/tracks.html>
- [78] NII (National Institute of Informatics,) NTCIR-6 Question Answering Challenge (QAC-4) Home Page <http://www.nlp.is.ritsumei.ac.jp/qac/>
- [79] 関根聡, "百科事典を対象とした質問応答システムの開発," 言語処理学会第9回年次大会, pp.637-640, Mar. 2003
- [80] 山田一郎, 柴田正啓, 金淵培, "Webを情報源としたQ&Aシステムの検討," 言語処理学会第9回年次大会, pp.633-636, Mar. 2003
- [81] 浦田耕二, 福井美佳, 藤井寛子, 鈴木優, 酒井哲也, 齋藤佳美, 市村由美, 佐々木寛, "質問応答技術に基づくマルチモーダルヘルプシステム," 情処学研報, 2004-FI-74-(4), Vol.2004, No.36, pp.23-29, Mar. 2004
- [82] H. Yang, L. Chaisorn, Y.Zhao, S.-Y. Neo, T.-S. Chua, "VideoQA: question answering on news video," Proceedings of ACM International Conference on Multimedia, pp.632-641, Berkeley, CA, USA, Nov. 2003
- [83] 福本淳一, 榎井文人, "質問応答技術 - 大量のデータを基に任意の質問に答える - ," 情報処理, Vol.45, No.6, pp.580-585, Jun. 2004
- [84] 伊藤一成, 藤原司, 橋田浩一, "セマンティックビデオオーサリングによるニュース動画群からのダイジェスト生成," 人工知能学会 第74回知識ベースシステム研究会 SIG-KBS-A601, Jul. 2006
- [85] 林英俊, 李龍, 上林弥彦, "概念グラフを用いたニュース映像要約システムの構築," データベース学会年次大会, DEWS2003, 4-A-03, Mar. 2003
- [86] 大平茂輝, 長尾確, "アノテーションに基づくスポーツ映像要約とナレーション生成," 第67回情処全大, 2K-4, Mar. 2005

- [87] 河合吉彦, 馬場口登, 北橋忠宏, “個人適応を指向したスポーツ要約映像の生成法,” 信学技報, PRMU2000-171, Jan. 2001
- [88] 三浦宏一, 浜田玲子, 井手一郎, 坂井修一, 田中英彦, “動きに基づく料理映像の自動要約,” 情処学論 コンピュータとイメージメディア研究会論文誌 Vol.44, No.SIG9, pp.21-29, Jul. 2003
- [89] 森山剛, 坂内正夫, “ドラマ映像の心理的内容に基づいた要約映像の生成,” 信学論 (D-II), Vol.J84-D-II, No.6, pp.1122-1131, June 2001
- [90] 奥村学, “要約技術,” 情報処理, Vol.45, No.6, pp.574-579, Jun. 2004
- [91] 佐野雅規, 八木伸行, 片山紀生, 佐藤真一, “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,” 信学技報 PRMU2006-72, Vol.106, No.229, pp.89-96, Sep. 2006
- [92] M. Sano, N. Yagi, J. Martinet, N. Katayama and S. Satoh, “Image-based Quizzes from News video Archives,” Proceedings of IEEE International Conference on Multimedia and Expo, Th-L6.1, pp.1547-1550, Beijing, China, Jul. 2007
- [93] 国立国語研究所 . 分類語彙表, 1993 年版
- [94] C.-C. Chang and C.-J. Lin, “LIBSVM: a library for support vector machines,” 2001, Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm/>
- [95] CaboCha/南瓜, <http://www.chasen.org/~taku/software/cabocha/>
- [96] M. Sano, Y. Kawai, N. Yagi and S. Satoh, “Video Rushes Summarization Utilizing Retake Characteristics,” Proceedings of ACM International Conference on Multimedia Workshop on TRECVID BBC Rushes Summarization, pp.95-99, Vancouver, Canada, Oct. 2008
- [97] “Information technology – SGML applications – Topic maps,” ISO/IEC 13250:2003
- [98] FreeMind, http://freemind.sourceforge.net/wiki/index.php/Main_Page

研究業績

本論文を構成する論文

【学術論文】

1. 佐野雅規, 八木伸行, 片山紀生佐藤真一, “蓄積されたニュース番組からの画像付きクイズ生成手法の提案,” 信学論 D-II, Vol. J92-D, No.1, pp.141-152, Jan.2009

【国際会議論文】(査読有)

2. M. Sano, N. Yagi, N. Katayama and S. Satoh, “Image-based Quiz Generation from News Video Archives Based on Principal Object,” Proceedings of ACM International Conference on Multimedia, pp.929-932, Vancouver, Canada, Oct. 2008
3. M. Sano, Y. Kawai, N. Yagi and S. Satoh, “Video Rushes Summarization Utilizing Retake Characteristics,” Proceedings of ACM International Conference on Multimedia Workshop on TRECVID BBC Rushes Summarization, pp.95-99, Vancouver, Canada, Oct. 2008
4. M. Sano, N. Yagi, J. Martinet, N. Katayama and S. Satoh, “Image-based Quizzes from News video Archives,” Proceedings of IEEE International Conference on Multimedia and Expo, Th-L6.1, pp.1547-1550, Beijing, China, Jul. 2007
5. M. Sano, Y. Kawai, H. Sumiyoshi and N. Yagi, “Metadata Production Framework and Metadata Editor,” Proceedings of ACM International Conference on Multimedia, pp.789-790, Santa Barbara, CA, USA, Oct. 2006

【研究会】

6. 佐野雅規, 八木伸行, 片山紀生, 佐藤真一, “主被写体に注目した画像付き選択クイズの生成手法,” 映情学技報 ME2008-60, Vol.32, No.10, pp.39-44, Feb. 2008

7. 佐野雅規，河合吉彦，住吉英樹，八木伸行，“メタデータ制作フレームワークネットワークシステムへの拡張”，映情学技報 ME2007-31, Vol.30, No.10, p.171-176, Feb. 2007
8. 佐野雅規，八木伸行，片山紀生，佐藤真一，“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”，信学技報 PRMU2006-72, Vol.106, No.229, pp.89-96, Sep. 2006

【全国大会】

9. 佐野雅規，柴田正啓，八木伸行，佐藤真一，“蓄積したニュース番組を活用するフレームワーク～EvenTankへの構想～”，映情学冬大，7-7 Dec. 2008

学術論文 (全て査読有)

【総研大入学以後の主著】

1. 佐野雅規, 八木伸行, 片山紀生, 佐藤真一, “蓄積されたニュース番組からの画像付きクイズ生成手法の提案,” 信学論 D-II, Vol. J92-D, No.1, pp.-, Jan.2009

【上記以外】

2. M. Sano, I. Yamada, H. Sumiyoshi and N. Yagi, “Automatic Real-Time Selection and Annotation of Highlight Scenes in Televised Soccer,” IEICE Transactions on Information & Systems, Vol.E90-D, No.1, pp.224-232, Jan. 2007
3. 山田一郎, 佐野雅規, 住吉英樹, 柴田正啓, 八木伸行, “アナウンサーと解説者のコメントを利用したサッカー番組セグメントメタデータ自動生成,” 信学論 D-II, Vol.J89-D, No.10, pp.2328-2337, Oct. 2006
4. 住吉英樹, 有安香子, 望月祐一, 佐野雅規, 井上誠喜, “階層化した情報管理構造を用いた番組制作,” 映情学誌, Vol.55, No.3, p.397, Mar. 2000

国際会議・査読付き講演・レター等 (全て査読有)

【総研大入学以後の主著】

1. M. Sano, N. Yagi, N. Katayama and S. Satoh, "Image-based Quiz Generation from News Video Archives Based on Principal Object, " Proceedings of ACM International Conference on Multimedia, pp.929-932, Vancouver, Canada, Oct. 2008
2. M. Sano, Y. Kawai, N. Yagi and S. Satoh, "Video Rushes Summarization Utilizing Retake Characteristics," Proceedings of ACM International Conference on Multimedia Workshop on TRECVID BBC Rushes Summarization, pp.95-99, Vancouver, Canada, Oct. 2008
3. M. Sano, N. Yagi, J. Martinet, N. Katayama and S. Satoh, "Image-based Quizzes from News video Archives," Proceedings of IEEE International Conference on Multimedia and Expo, Th-L6.1, pp.1547-1550, Beijing, China, Jul. 2007
4. M. Sano, Y. Kawai, H. Sumiyoshi and N. Yagi, "Metadata Production Framework and Metadata Editor," Proceedings of ACM International Conference on Multimedia, pp.789-790, Santa Barbara, CA, USA, Oct. 2006

【上記以外】

5. K. Adistambha, M. Doeller, R. Tous, M. Gruhne, M. Sano, C. Tsinaraki, S. Christodoulakis, K. Yoon, C.H. Ritz, I. Burnett, "The MPEG-7 Query Format: A New Standard in Progress for Multimedia Query by Content," Proceedings of IEEE International Symposium on Communications and Information Technologies, p.479-484, Sydney, Australia, Oct. 2007
6. M. Sano, H. Sumiyoshi, M. Shibata and N. Yagi, "Generating Metadata from Acoustic and Speech Data in Live Broadcasting," Proceeding of IEEE International Conference on Acoustics, Speech, and Signal Processing, MSP-P.2.4, pp.1145-1148, Philadelphia, PA, USA, Mar. 2005

学術講演・研究報告等

1. 佐野雅規, 柴田正啓, 八木伸行, 佐藤真一, “蓄積したニュース番組を活用するフレームワーク～EvenTankへの構想～,” 映情学冬大, 7-7, Dec. 2008
2. 住吉英樹, 佐野雅規, 柴田正啓, 八木伸行, “メタデータ制作フレームワーク～データモデルとAPI機能の拡張～,” 第7回情報科学技術フォーラム講演論文集 [FIT 2008] H-011, pp.85-86, Aug. 2008
3. 佐野雅規, 八木伸行, 片山紀生, 佐藤真一, “主被写体に注目した画像付き選択クイズの生成手法,” 映情学技報 ME2008-60, Vol.32, No.10, pp.39-44, Feb. 2008
4. 佐野雅規, 河合吉彦, 住吉英樹, 八木伸行, “メタデータ制作フレームワークネットワークシステムへの拡張,” 映情学技報 ME2007-31, Vol.30, No.10, p.171-176, Feb. 2007
5. 住吉英樹, 佐野雅規, 八木伸行, “メタデータ制作フレームワーク,” 映情学誌, Vol.61, No.2, p.152-157, Feb. 2007
6. N. Yagi, H. Sumiyoshi, M. Sano, “Metadata Production Framework and Metadata Editor,” ABU Technical Committee, 2007 Annual Meeting, Doc T-6/44-7, Beijing, China, Nov. 2007
7. 佐野雅規, 八木伸行, 片山紀生, 佐藤真一, “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,” 信学技報 PRMU2006-72, Vol.106, No.229, pp.89-96, Sep. 2006
8. 佐野雅規, 河合吉彦, 住吉英樹, 八木伸行, “メタデータ制作フレームワークの提案とメタデータエディタの開発,” 第5回情報科学技術フォーラム一般講演論文集 (FIT 2006), I-067, pp.165-166, Aug. 2006
9. 住吉英樹, 佐野雅規, 山田一郎, 松井淳, サイモン・クリッピングゲール, 望月貴裕, 三須俊彦, 佐藤庄衛, 小林彰夫, 今井亨, 松村欣司, 八木伸行, “メタデータ制作・活用システムの試作,” 映情学技報 ME2005-205, Vol.29, No.70, pp.13-18, Nov. 2005
10. 佐野雅規, 住吉英樹, 八木伸行, “サッカー中継における会場音とスピーチを利用したメタデータ生成,” 信学技報 PRMU2005-120, Vol.105, No.415, pp.33-38, Nov. 2005

11. 佐野雅規, 住吉英樹, 八木伸行, “メタデータ生成のための歓声と応援の分類手法,” 第4回情報科学技術フォーラム一般講演論文集 (FIT 2005), I-029, pp.71-72, Aug. 2005
12. 佐野雅規, 住吉英樹, 八木伸行, “情報統合機能を実装したメタデータエディタ,” 第4回情報科学技術フォーラム一般講演論文集 (FIT 2005), I-028, pp.69-70, Aug. 2005
13. 住吉英樹, 佐野雅規, 山田一郎, 松井淳, サイモン・クリッピングゲール, 望月貴裕, 三須俊彦, 佐藤庄衛, 小林彰夫, 今井亨, 松村欣司, 八木伸行, “メタデータ制作・活用システムの試作,” 2005年映情年次大, 11-3, Aug. 2005
14. 佐藤庄衛, 小林彰夫, 尾上和穂, 山田一郎, 佐野雅規, 今井亨, “メタデータ生成のための音声認識の改善,” 2005年映情年次大, 9-1, Aug. 2005
15. 山田一郎, 佐野雅規, 住吉英樹, 柴田正啓, 八木伸行, “アナウンスコメントを利用したサッカー番組メタデータ自動生成,” 信学技報, NLC2004-122, pp.37-42, Feb. 2005
16. 佐野雅規, 山田一郎, 有安香子, 住吉英樹, 柴田正啓, 八木伸行, “メタデータエディタの試作 メタデータ制作活用のプラットフォームの提案,” 信学技報 PRMU2004-147, pp.71-76, Dec. 2004
17. 住吉英樹, 佐野雅規, 八木伸行, “スポーツダイジェスト番組の制作支援を目的としたスロー再生区間検出手法,” 2004年映情冬季大, 9-6, Dec. 2004
18. 山田一郎, 佐野雅規, 住吉英樹, 柴田正啓, “アナウンスコメントを利用したサッカー番組のメタデータ自動生成の検討,” 第3回情報科学技術フォーラム一般講演論文集 (FIT 2004), E-030, pp.177-178, Aug. 2004
19. 佐野雅規, 山田一郎, 有安香子, 住吉英樹, 柴田正啓, “メタデータエディタの開発,” 2004映情学年次大, 7-9, Aug. 2004
20. 宮崎勝, 佐野雅規, 金淵培, 柴田正啓, “MPEG-7を用いた情報提示システムに関する一検討,” 第1回情報科学技術フォーラム一般講演論文集 (FIT 2002), D-23, p.45, Sep. 2002
21. 佐野雅規, 宮崎勝, 住吉英樹, 柴田正啓, “MPEG-7を利用したアプリケーションシステムの一検討,” 第1回情報科学技術フォーラム一般講演論文集 (FIT 2002), D-22, p.43, Sep. 2002

22. 佐野雅規, 住吉英樹, 柴田正啓, 八木伸行, “サッカーゲームにおける会場音からの重要部分抽出の検討,” 信学技報, PRMU2001-166, pp.87-92, Dec. 2001
23. 佐野雅規, 住吉英樹, 柴田正啓, “映像番スコアブックの相撲への適用,” 2001年映情学年次大, 17-8, pp.250-251, Aug. 2001
24. 宮崎勝, 佐野雅規, 山田一郎, 住吉英樹, 柴田正啓, “MPEG-7を用いた番組情報検索システムの開発,” 2001年映情学年次大, 16-5, p.233, Aug. 2001
25. 佐野雅規, 住吉英樹, 柴田正啓ほか, “スポーツ映像における人物インデックス生成の一検討,” 信学総合大, D-12-51, p.218, Mar. 2001
26. 佐野雅規, 住吉英樹, 柴田正啓, 井上誠喜, “映像番スコアブックの提案,” 映情学技報, MIP2001-48, pp.7-12, Mar. 2001
27. 住吉英樹, 有安香子, 望月祐一, 佐野雅規, 井上誠喜, “番組制作システムDTPPによる番組情報のデータベース化,” 情報科学技術協会 INFOSTA シンポ, ISSN1343-2044, pp.60-65, Jun. 2000
28. 佐野雅規, 住吉英樹, 井上誠喜, “映像版スコアブックの検討～新しいスポーツ映像管理を目指して～,” 信学技報, PRMU99-257, pp.83-90, Mar. 2000
29. 住吉英樹, 有安香子, 望月祐一, 佐野雅規, 井上誠喜, “データベースを中心とした番組制作支援システム,” 情処学 DBS 研報 DBS-120-12, Vol.2000, Mo.10, p.83-90”, Jan. 2000
30. 佐野雅規, 住吉英樹, 井上誠喜, “映像版スコアブック,” 1999年映情学冬季大, 2-14, p.57, Dec. 1999
31. 佐野雅規, 大槻一博, 井上誠喜, “データ放送コンテンツの制作,” 画電学研資, 99-02-03, p.11, Sep. 1999
32. 有安香子, 住吉英樹, 望月祐一, 佐野雅規, 井上誠喜, “番組制作情報の構造文章化と検索システム,” 1999年映情学年次大, 21-2, pp.277-278, Aug. 1999
33. 有安香子, 住吉英樹, 望月祐一, 佐野雅規, 井上誠喜, “番組制作情報を用いた高機能検索システム,” 1999信学総大, SMD-1, pp.327-328, Mar. 1999

34. 住吉英樹, 有安香子, 望月祐一, 佐野雅規, 井上誠喜, “階層化番組制作手法における番組情報構造,” 1999 信学総大, SMD-2, pp.329-330, Mar. 1999
35. 佐野雅規, 望月祐一, 住吉英樹, 井上誠喜, 福井一夫, “ISDB データ放送における気象サービスの一検討,” 映情学技報, BCS-98-38, Vol.22, pp.1-6, Sep. 1998
36. 望月祐一, 佐野雅規, 住吉英樹, 福井一夫, “ISDB におけるニュースサービスの検討,” 1998 年映情学年次大, 7-4, pp.94-95, Jul. 1998
37. 佐野雅規, 望月祐一, 住吉英樹, 福井一夫, “ISDB データ放送における気象サービスの検討,” 1998 年映情学年次大, 7-5, pp.96-97, Jul. 1998
38. 佐野雅規, 伊藤崇之, 中川俊夫, 松本隆, “視覚前注意過程におけるテクスチャー境界検出モデル,” テレビ学技報, VAI'94-29, p.31, Mar. 1994
39. 佐野雅規, 伊藤崇之, 中川俊夫, 松本隆, “視覚前注意過程におけるテクスチャー境界検出モデル,” 1994 信学春季大, D-77, Vol.6, p.77, Mar. 1994
40. 佐野雅規, 伊藤崇之, 中川俊夫, 松本隆, “多重解像度表現によるテクスチャー境界の検出,” 信学技報, NC91-110, Vol.9, pp.89-96, Mar. 1992

標準化寄与文書等

1. **M. Sano**, H. Sumiyoshi, N. Yagi, “Filtering function in Join operation,” ISO/IEC JTC1/SC29/WG11 (MPEG) 84th meeting, MPEG2008/M15401, Archamps, France, Apr. 2008
2. **M. Sano**, H. Sumiyoshi, N. Yagi, “The function signaling desirable page in MPQF,” ISO/IEC JTC1/SC29/WG11 (MPEG) 84th meeting, MPEG2008/M15399, Archamps, France, Apr. 2008
3. **M. Sano**, H. Sumiyoshi, N. Yagi, “Solutions for inconsistency of interpretation in SpatialQuery and TemporalQuery,” ISO/IEC JTC1/SC29/WG11 (MPEG) 84th meeting, MPEG2008/M15398, Archamps, France, Apr. 2008
4. **M. Sano**, H. Sumiyoshi, N. Yagi, “Paging function in MPQF,” ISO/IEC JTC1/SC29/WG11 (MPEG) 83rd meeting, MPEG2008/M15176, Antalya, Turkey, Jan. 2008
5. **M. Sano**, H. Sumiyoshi, N. Yagi, “Interpretation consistency for Spatial-Query and TemporalQuery,” ISO/IEC JTC1/SC29/WG11 (MPEG) 83rd meeting, MPEG2008/M15177, Antalya, Turkey, Jan. 2008
6. **M. Sano**, H. Sumiyoshi, N. Yagi, “Some thoughts on data model for MPQF,” ISO/IEC JTC1/SC29/WG11 (MPEG) 82nd meeting, MPEG2007/M14946, Shenzhen, China, Oct. 2007
7. **M. Sano**, H. Sumiyoshi, N. Yagi, “A proposal for MP7QF based on QueryCondition part,” ISO/IEC JTC1/SC29/WG11 (MPEG) 81st meeting, MPEG2007/M14641, Lausanne, Switzerland, Jul. 2007
8. **M. Sano**, H. Sumiyoshi, N. Yagi, “Test report of CEs on MP7QF,” ISO/IEC JTC1/SC29/WG11 (MPEG) 80th meeting, MPEG2007/M14537, San Jose, U.S.A., Apr. 2007
9. **M. Sano**, H. Sumiyoshi, N. Yagi, “A Proposal for MPEG-7 Query Format Framework,” ISO/IEC JTC1/SC29/WG11 (MPEG) 79th meeting, MPEG/M14211, Marrakech, Morocco, Jan. 2007
10. **M. Sano**, “Some thoughts on MPEG-7 Profiling,” ISO/IEC JTC1/SC29/WG11 (MPEG) 69th meeting, MPEG2004/M11023, Redmond, U.S.A., Jul. 2004

11. **M. Sano**, “MPEG-7 Core Description Profile Update,” ISO/IEC JTC1/SC29/WG11 (MPEG) 69th meeting, MPEG2004/M11024, Redmond, U.S.A., Jul. 2004
12. **M. Sano**, B. Gandhi, A. Melby, R. Pinho, “Minor Editing to CD of MPEG-7 Part 9,” ISO/IEC JTC1/SC29/WG11 (MPEG) 67th meeting, MPEG2003/M10336, Waikaloa, U.S.A., Dec. 2003
13. **M. Sano**, B. Gandhi, A. Melby, R. Pinho, “Supporting Descriptions for the Core Content Description Profile (CCDP) and a Proposed Grouping Within CCDP,” ISO/IEC JTC1/SC29/WG11 (MPEG) 66th meeting, MPEG2003/M10066, Brisbane, Australia, Oct. 2003
14. V.P. Beek, T. DeMartini, B. Gandhi, A. Melby, **M. Sano**, J.R. Smith, A. Vetro, “Signaling of profile & level in MPEG-7 descriptions,” ISO/IEC JTC1/SC29/WG11 (MPEG) 66th meeting, MPEG2003/M10154, Brisbane, Australia, Oct. 2003
15. **M. Sano**, M. Shibata, “Levels and templates for Core Content Description Profile,” ISO/IEC JTC1/SC29/WG11 (MPEG) 65th meeting, MPEG2003/M9825, Trondheim, Norway, Jul. 2003
16. **M. Sano**, M. Shibata, “An additional issue on MPEG-7 Video Program Profile,” ISO/IEC JTC1/SC29/WG11 (MPEG) 64th meeting, MPEG2003/M9593, Pattaya, Thai, Mar. 2003
17. **M. Sano**, R.V. Possato, M. Shibata, “A Study on MPEG-7 Video Program Profile,” ISO/IEC JTC1/SC29/WG11 (MPEG) 63rd meeting, MPEG2002/M9223, Awaji, Japan, Dec. 2002
18. **M. Sano**, M. Shibata, “Results of experimental implementation of Content Retrieval System by using MPEG-7 compliant metadata,” ISO/IEC JTC1/SC29/WG11 (MPEG) 61st meeting, MPEG2002/M8542, Klagenfurt, Austria, Jul. 2002
19. **M. Sano**, K. Hashida, “Amendments to MPEG-7 MDS for Linguistic Descriptions,” ISO/IEC JTC1/SC29/WG11 (MPEG) 60th meeting, MPEG2002/M8397, Fairfax, U.S.A., May 2002

紀要等

1. 佐野雅規 , メタデータ制作フレームワーク , 技研だより第 45 号 , p.3, Dec. 2008
2. 八木伸行 , 佐野雅規 , メタデータの規格 , NHK 技研 R&D, No.95, pp. 12-21, Jan. 2006
3. 住吉英樹 , 山田一郎 , 佐野雅規 , メタデータ抽出・付与技術の研究動向 , NHK 技研 R&D, No.95, pp. 22-31, Jan. 2006
4. M. Sano , Metadata Editor, Broadcast Technology No.22, p.21, Spring 2005
5. 佐野雅規 , メタデータエディター , 新技研だより , No.67, p.3, Feb. 2005

解説記事等

1. M. Doller, R. Tous, M. Gruhne, K. Yoon, M. Sano, I.S. Burnett, “The MPEG Query Format: Unifying Access to Multimedia Retrieval Systems,” IEEE MultiMedia, pp.82-95, October-December 2008
2. 佐野雅規 , メタデータエディター , ケーブル新時代 2007 年 11 月号 , p.49
3. 佐野雅規 , メタデータエディター , 電波受験界 2005 年 6 月号 , pp.66-67

書籍等

1. 映像メディア技術 , 八木伸行監修 , オーム社 , 第 9 章 映像検索・活用技術 , pp.169-190, 2008

登録特許

【国内】

1. 特許第 4256077 号，代表静止映像自動生成装置
2. 特許第 4177608 号，カスタマイズ料理番組配信装置およびカスタマイズ料理番組配信プログラム
3. 特許第 3983532 号，場面抽出装置
4. 特許第 4172904 号，映像・音声検索装置
5. 特許第 3914643 号，番組情報管理編集システムとそれに使用される階層化番組情報蓄積管理装置
6. 特許第 3550300 号，マルチメディア型情報サービス方式，マルチメディア型情報サービス送信装置およびマルチメディア型情報サービス受信装置

その他，出願多数．

表彰等

1. 電子情報通信学会 平成 13 年度学術奨励賞, 2001 年. (学術講演・研究報告等の項番 25 について)
2. NHK 放送技術研究所所長賞，平成 19 年.
3. NHK 放送技術研究所長特賞，平成 12 年.