

# 日常生活オントロジーに基づくマルチメディアコンテンツのための操作インタフェース

## Operation Interface for Multimedia Content Based on Everyday Life Ontology

大西可奈子\*1  
Kanako ONISHI

渡邊裕子\*2  
Yuko WATANABE

和泉憲明\*3  
Noriaki IZUMI

小林一郎\*2  
Ichiro KOBAYASHI

橋田浩一\*3  
Kôiti HASIDA

\*1お茶の水女子大学大学院 人間文化研究科 数理・情報科学専攻  
Graduate School of Humanities and Sciences Ochanomizu University

\*2お茶の水女子大学 理学部 情報科学科  
Dept. of Information Sciences Ochanomizu University

\*3産業技術総合研究所  
National Institute of Advanced Industrial Science and Technology (AIST)

In this paper, we propose a method to share and collaborate on multimedia content data and to access the multimedia content using natural language. To manage and share multimedia content, we introduce an ontology concerning events in people's everyday lives, called 'Everyday Life Ontology' (ELO). We discuss a data management application using ELO, and show a concrete example of how we can collaboratively organize multimedia content data and operate on them using natural language queries with actual multimedia content obtained from an actual complex, but realistic, event.

## 1. 研究背景と目的

近年、コンピュータの価格低下やインターネットの普及により、コンピュータを使用する機会は格段に増加している。また、ハードディスクの小型・大容量化や Podcasting などの技術の進歩にともなう様々なマルチメディアプレーヤを屋外で楽しむこともできるようになった。様々なサービスが楽しめる一方で、そのようなコンピュータやデジタル機器は高性能化が進み、操作も複雑になってきており、また、我々が管理すべき電子データは日を追うごとに増えている。

このような背景を踏まえて、本研究では、それらをより簡単に操作・管理することができるようにするために、自然言語によって容易に操作ができるインタフェースを持ち、身近にある生活情報を使って膨大なコンテンツを容易に管理できるシステムを提案する。具体的にはまず、日常生活の中に存在する人々の間で共有されているオントロジー（日常生活オントロジーと呼ぶ）を利用しデータを管理する。また、自然言語インタフェースを実現しようとするとき問題になる、人が使用する自然言語とコンピュータが理解できるプログラムの間に存在するギャップは、ユーザの操作に関するオントロジー（ユースケースオントロジーと呼ぶ）や日常生活オントロジーによるデータ構造などを利用し吸収する。

## 2. オントロジーに基づいたコンテンツ管理手法の提案

### 2.1 概要

本研究では、ユーザの自然言語による操作命令に対して、直接コンテンツを指定することにより、その操作内容を決定する。操作指示されたコンテンツは操作内容により、異なるアプリケーションソフトウェアを使用する。本研究では外部アプリケーションソフトウェアとして、iTunes[1]、MS Word、MS Outlook を使用している。このような自然言語によるコンテ

ンツの指定を可能にするため、まずはオントロジーを利用したコンテンツの管理法を提案する。

### 2.2 日常生活オントロジーを用いた対象の特定 日常生活オントロジー

静止画や動画といったマルチメディアコンテンツを管理する際、“どこで誰とどうして撮ったのか”などといったメタデータを一枚一枚に付与することができれば、精度のよい検索が可能となる。しかし、実際にすべての動画や静止画にそういったメタデータを手動で付与することは、大変な労力と時間が必要となる。そこで本研究では、日常生活に存在する様々な制約を使って、ある程度のメタデータを自動で付与することを考える。

例えば、人の一日の行動スケジュールにおいて人と会う場合には、約束の時間や場所などの制約が存在し、会う人とそれらの知識を共有することにより目的を達成することが出来る。このように、日常生活の中に自然な形で存在し他人と共有している制約を用いて情報の管理を行う。これを用いて画像を管理することを考えると、画像を取得するモチベーションとなる日常生活オントロジーとして「行事に基づくイベント」と「地理情報に基づくイベント」の二つが考えられる。日常生活の中に存在する制約の中で、今回はこの二つと参加者の情報を用いる。本研究において、この三つを日常生活オントロジーと呼ぶ。

#### ● 行事に基づくイベント

プログラムやスケジュールが予めわかっているものについては、プログラム中の小イベント名をそのまま画像を関連づけるスケジュールの箱として利用する。同じイベントに参加した複数の人が投稿して共有された画像の取得時間、およびそれらにつけられたコメントから判断して、スケジュールの箱と画像情報を結びつける。

#### ● 地理情報に基づくイベント

地理情報が画像と密接な関係をもつ場合には、地理情報をイベントとして捉え、それに対して画像を関連づけることで画像の整理・管理を行う。例えば、遊園地で遊ぶ場合、あらかじめ行動に対するプログラムが定められて

連絡先: 大西可奈子, お茶の水女子大学人間文化研究科数理・情報科学専攻, 東京都文京区大塚 2-1-1, 03-5978-5709, 03-5978-5709, onishi@koba.is.ocha.ac.jp

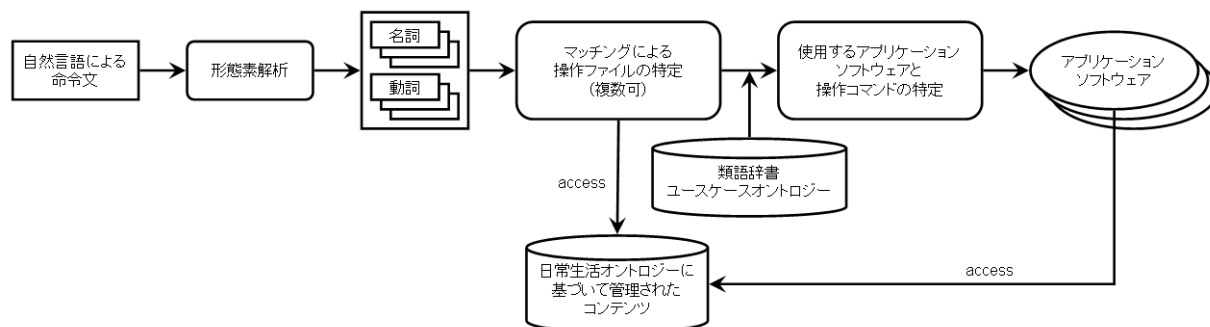


図 1: システム概要

いるわけではないので、取得された画像情報のほとんどが、アトラクションに参加した際のものになる。アトラクション、すなわち地理情報がイベント情報となり画像と関連づけられることになる。

● 参加者情報

さらに日常生活オントロジーとして、画像の内容に関連する情報として被写体の情報がある。被写体の情報は各画像情報には予め付与することはできないが、イベントの種類やスケジュールによっては、参加者が時間帯ごとに予めわかっている場合がある。特定の被写体が写っている画像などを検索する際には、イベントのスケジュールと参加者の情報を照らし合わせて、特定の被写体が写っている画像を絞り込むことができる。

マルチメディアコンテンツのファイル属性の利用

以上のような日常生活オントロジーに基づいて得た情報とは別に、取得した画像情報が保持する情報も管理・検索に使用する。具体的にはファイルが撮影された日付やファイルサイズなどがある。またそれとは別に、画像を撮影した人が誰なのかは比較的取得が簡単であるので後ほど登録してもらいファイル属性として使用することとする。

これらの情報は動画、静止画の基本情報として、上記に説明したイベントに関連付けを行う際に用いられ、また、特定の画像を検索する際にも用いられる。

2.3 一般的なファイルの構造化

現在は扱えるコンテンツを動画と静止画に限っているためスケジュールに基づくイベントと地理情報に限って論じているが、これは一般的なファイルを特定する際にも同様のことがいえる。例えば、ワードで作成されたファイルには、作成日時などのファイル属性の他に、そのファイルを作成した動機となったイベントがあり、そのファイルを作成した場所（地理情報）がある。これらの情報を使えばこのような一般的なファイルを特定することも可能である。特定できるファイルの一般化については現在考察中である。

3. 自然言語による操作インタフェース

システムの概要を図 1 に示す。

3.1 入力文解析

奈良先端科学技術大学院大学で開発された日本語係り受け解析器 CaboCha/南瓜 [2] を使って、入力文の構文解析および形態素解析を行い、操作項目と操作対象につながる述部と目的語を特定する。このとき述部は目的語とセットで抽出される。

3.2 マッチングによる操作ファイルの特定

対象となるファイルを決めるために、問い合わせ文から抽出した名詞ごとに下記の 4 つのチェックを行い、その結果をユーザの指定するファイルの条件とする（図 2 参照）。最終的に最大で名詞の数に相当する条件が出る。

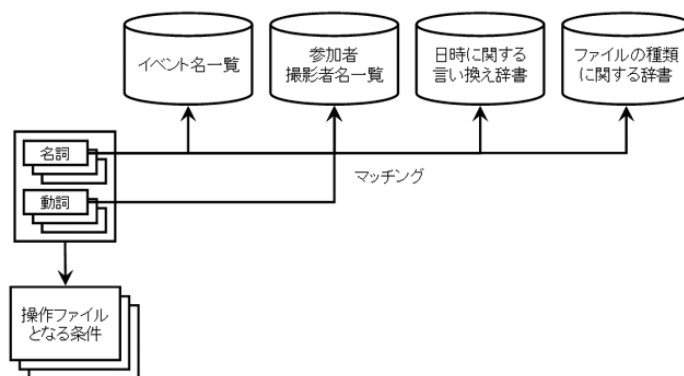


図 2: マッチングによる操作ファイルの特定法

● イベント名とのマッチング

ある名詞が、日常生活オントロジーにより管理されているファイルが持つイベント名と一致した場合は、それをユーザが指定したイベントと特定し、その条件を保持する。

● 参加者名とのマッチング

ある名詞が、日常生活オントロジーにより管理されているファイルが持つ参加者名と一致した場合は、まずその名詞に付随する動詞が否定形でないかを確認する。否定形でなければそれをユーザが指定した参加者と特定し、その条件を保持する。否定形であった場合は、その参加者以外の参加者をユーザが指定した参加者と特定し、その条件を保持する。

● 日付に関する単語

解析によって抽出された名詞に日付を特定するような単語がないかをチェックする。その際、言葉の揺らぎ等は日付に関する言い換え辞書を用いて解消する。

● ファイルの種類

ファイルの種類を特定するような単語がないかを調べる。例えば“ムービー”という単語が入っていた場合は、対象となるファイルは動画であると判断する。

以上の条件を満たさない名詞は、この条件抽出には使用されない。

### 3.3 類語辞書とユースケースの特定

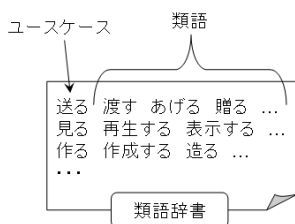


図 3: 類語辞書

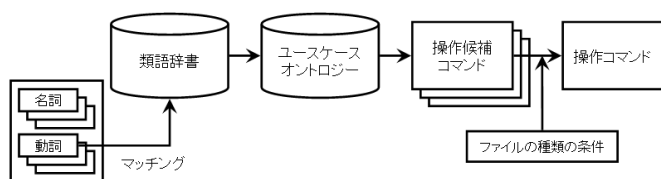


図 4: 類語辞書とユースケースの特定法

本研究では、類語辞書（図 3 参照）を用いて述部からユースケースを抽出する。その概要を図 4 に示す。類語辞書とは、同義語や類義語、さらに本研究で使用したアプリケーションソフトウェアを使用する際に、同等の意味で使用されると思われる述部をひとつの述部に集約するための辞書で、その集約先の述部がアプリケーションソフトウェアの操作のユースケースとなっている。ユースケースはユーザの操作に関するオントロジー（ここではユースケースオントロジーと呼ぶ）に基づいて決定される（図 5 参照）。

また、ユースケースはアプリケーションソフトウェアのコマンドに直接結びついているため、ユースケースを決定することにより実行候補となるコマンドも決定する（図 6 参照）。結びつくコマンドは複数であることがあるため、ユースケースを決定することにより操作コマンドが一意に決定できるわけではないが、最終的にこれらのコマンドのいずれかが実行される。

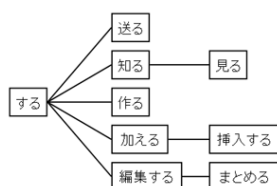


図 5: ユースケースオントロジー

### 3.4 使用するアプリケーションソフトウェアと操作コマンドの特定

これらコマンド候補から、操作コマンドを決定するために用いる条件が操作ファイルの種類である。例えば、操作対象が動画であればユースケース「見る」は静止画を表示するのではなく、動画を再生するコマンドの実行を行えばよいのだと判断する。これに関しては、時節で詳しく説明する。また、操作コマンドを決定することにより使用するアプリケーションソフトウェアも決まる。このようにして決定した操作アプリケーション

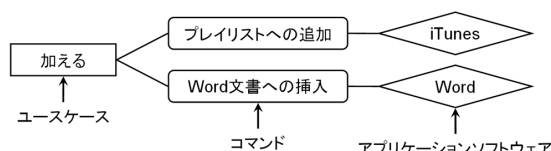


図 6: ユースケース「加える」に対する操作コマンド

ソフトウェアが外部アプリケーションソフトウェアだった場合、システムは外部アプリケーションソフトウェアにアクセスし、対象ファイルに対し、特定の操作を行う。

## 4. 自然言語によるコンテンツ操作例

「裕子さんが昼ごろ富士急ハイランドで撮った動画を見せて」を例に、システムの動きを具体的に説明する。

### 4.1 入力文解析

命令文を形態素解析することにより、以下の動詞と名詞のセットを抽出することができる。

- 動詞「撮った」にかかる名詞  
裕子、さん、昼、ごろ、富士急ハイランド
- 動詞「見せて」にかかる名詞  
動画

### 4.2 マッチングによる操作ファイルの特定

次に、形態素解析で抽出した名詞を順に、特定の語とのマッチングを行う。

- 裕子  
まず日常生活オントロジーに基づいて管理している動画・静止画データベースにアクセスして、イベント名とのマッチングを行う。しかしイベント名に“裕子”は存在しないので、次に撮影者または参加者に“裕子”がいるかどうかを調べる。今回の例では“裕子”という名前の撮影者が見つかったので、操作ファイルのひとつの条件は「裕子が撮影したものであるもの」となる。
- 昼  
同様にチェックを行っていくと、日時に関する言い換え辞書とマッチングを行ったところでこの単語は日時を表す単語であるということがわかる。したがって操作ファイルのひとつの条件は「撮影時間が昼であるもの」となる。
- 富士急ハイランド  
イベント名とマッチングしたところ同様の名前のイベントが見つかったので、操作ファイルのひとつの条件は「イベント“富士急ハイランド”で撮影されたもの」となる。また、イベント“富士急ハイランド”は、イベントの木構造よりイベント“ええじゃないか”、“ドドンパ”、“メリーゴーラウンド”などを含んでいることがわかる（図 7 参照）。
- 動画  
同様のチェックを行っていった結果、この名詞はファイルの種類を“動画”に限定する単語であることがわかる。

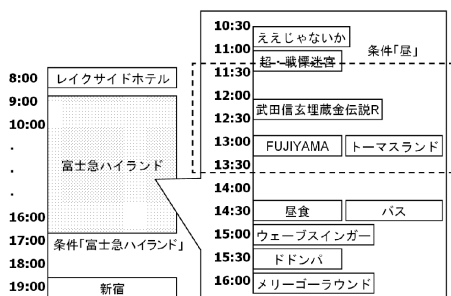


図 7: スケジュールと操作ファイル条件

よって操作ファイルのひとつの条件は「動画であるもの」となる。

以上の四項目が、ユーザが指定するファイルである条件となり、この条件をすべて満たす動画ファイルが操作対象ファイルである。

### 4.3 類語辞書とユースケースの特定

操作に関する動詞は全体の最後に出てくる動詞としているので、この場合は「見せて」がユーザの操作命令であると判断する。ここで、動詞「見せて」は基本形が「見る」であることから、ユースケースは「見る」となる。「見る」には、「動画の再生」と「静止画の表示という二つのコマンドが結びついている。

### 4.4 使用するアプリケーションソフトウェアと操作コマンドの特定

4.2 節で、今回ユーザが指定したファイルはすべて動画であることがわかっているので、操作コマンドは「動画の再生」であることがわかる。

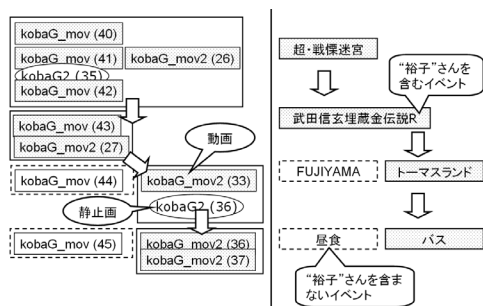


図 8: すべての条件を満たすファイル

よって、すべての条件を満たす9つのファイルが iTunes により再生される (図 8 参照)。

### 4.5 考察

本研究では、日々増え続ける膨大なデータを対象にしたコンテンツ管理と、それを容易に行えるための自然言語インタフェースの開発を目標としている。その中で、現在は対象とするデータを動画や静止画に限定している。

本研究で提案する日常生活オントロジーにより管理されたデータは、イベントに基づいて塊で取り出すことができたり、イベントや日時などの条件を付与することにより個別に取り出すこともできる。また本研究ではそれらの操作を自然言語で行

える。これによりユーザは、続けて自然言語による命令を出していくことにより、求めるファイルへ少しずつ焦点を絞っていくことができる。また、「次へ」と指定することにより、日時的に「次」に撮影されたファイルを指定するのではなく、次に起こったイベントそのものを指定することも可能である。これは、その時システムが指しているコンテンツが何であるかということ、システムが自動的に判断しユーザに提供するものである。ただし、現在のユースケースオントロジーでは自然言語による複雑な命令には対応しきれていない。複雑な命令を認識するためには、さらにオントロジーの充実が必要と思われる。

### 4.6 関連研究

時系列を中心にメディアの情報を管理する手法として、Yale 大学の D.Gelernter によって主導されている Lifestreams プロジェクト [3][4] がある。このプロジェクトにおいて、Freeman ら [5][6] は、様々な電子ファイルを時系列に並べ、5つの基本的な操作コマンド(新規作成、コピー、移動、検索、要約)によって、電子ファイルを管理している。ファイルを時間軸に沿って管理するという点においては、本研究と共通しているが、その時間軸の中に、人にとって親和性の高い生活オントロジーを導入し、その制約に基づきファイルを管理するという点において、我々の提案する手法は彼らの提案よりもより高度な電子ファイルの管理を実現している。また、この研究はデータを管理するという点に主眼を置いているため、我々が提案するようなコンテンツを操作するという複雑な操作は対象にせず、また、自然言語での柔軟な問い合わせも行えない。

## 5. まとめと今後の課題

本研究では、日常生活オントロジーに基づくコンテンツの管理法と、そのようにして管理されたコンテンツを自然言語で操作できるインタフェースを提案した。人とコンピュータが意思疎通しようとする場合、そこには必ずギャップが存在する。このギャップをうまく吸収し、人が伝えたい命令をコンピュータに理解させるために、本研究では日常生活オントロジーにより管理されているイベントとのマッチングや、類語辞書、ユースケースオントロジーを使用した結び付け方を提案した。

また今後は、ユースケースオントロジーの拡充とともに自然言語処理部の改良を行い、より複数のアプリケーションソフトウェアとの連携を図ることを目指した柔軟な操作インタフェースの開発を進めていくつもりである。

## 参考文献

- [1] iTunes: <http://www.apple.com/jp/itunes/>
- [2] 奈良先端科学技術大学大松本研究室: CaboCha/南瓜 <http://chasen.org/taku/software/cabocha/>
- [3] <http://www.cs.yale.edu/homes/freeman/lifestreams.html>
- [4] Scott Fertig, Eric Freeman and David Gelernter. "Finding and Reminding" Reconsidered. In SIGCHI Bulletin, volume 28, January, 1996.
- [5] Eric T.Freeman, The Lifestreams Software Architecture, Ph.D. Dissertation, Yale University, Dept. of Computer Science, May 1997.
- [6] Eric Freeman and David Gelernter. Lifestreams: A Storage Model for Personal Data, ACM SIGMOD Bulletin, March, 1996.
- [7] YouTube: <http://www.youtube.com/>
- [8] Flickr: <http://www.flickr.com/>