

# 博物館におけるユーザの学習行動に基づく 事後学習のための映像視聴支援方式の提案

小川 航<sup>†</sup> 王 元元<sup>††</sup> 河合由起子<sup>†††,††††</sup> 角谷 和俊<sup>†</sup>

<sup>†</sup> 関西学院大学メディア情報学科 〒669-1337 兵庫県三田市学園2丁目1

<sup>††</sup> 山口大学大学院創成科学研究科 〒775-8611 山口県宇部市常盤台2-16-1

<sup>†††</sup> 京都産業大学情報理工学部 〒603-8555 京都府京都市北区上賀茂本山

<sup>††††</sup> 大阪大学サイバーメディアセンター 〒567-0047 大阪府茨木市美穂ヶ丘5番1号

E-mail: <sup>†</sup>{fkg44895,sumiya}@kwansei.ac.jp, <sup>††</sup>y.wang@yamaguchi-u.ac.jp, <sup>†††</sup>kawai@cc.kyoto-su.ac.jp

あらまし 近年、博物館のデジタル化に伴い、来館者の行動やデバイスのログデータといった学習行動を取得できるようになった。また、博物館の映像資料は詳細な情報を多く含んでおり、展示を補足するためのデバイスや教育機関への貸し出し資料に利用されている。しかし、博物館に来館したユーザが事後学習として視聴するには、映像が長すぎるために利用しにくい状況にある。そのため、本研究では博物館の映像資料を事後学習に利用するために、UI上にてユーザに映像のタグを操作してもらい、そのログデータからタグに対してスコアリングを行い、館内での学習行動と組み合わせてユーザの興味に沿ったシーン区間を提示することで、効果的に事後学習を行う手法を提案する。

キーワード 事後学習, 博物館, 学習行動, ハイパーメディア, Wikipedia カテゴリ構造

## 1 はじめに

近年、博物館のデジタル化に伴い、来館者の行動やデバイスのログデータといった学習行動を取得できるようになった。また、博物館の豊富な映像資料は詳細な情報を多く含んでおり、展示を補足するためのデバイスや教育機関への貸し出し資料に利用されている。また、そのような映像を教育に利用することの有用性は Wu ら [1] の研究により明らかとなっている。しかし、映像資料は比較的長時間のものが多く、映像のトピックも多岐にわたる。そのため、ユーザが事後学習のために視聴する際、多くの動画の中から自分の見たい部分を探し出さなければならない。しかし、博物館の事後学習システムは利用者が初めてそのシステムを利用する機会が多いため、利用者が視聴したいシーンを適切に検索・視聴することは難しい。そのため、本研究ではユーザの興味に即した博物館の映像資料の各シーンを横断的に視聴させることで、短時間で効果的な事後学習の支援を行うことを目的としている。

人は博物館に来館した際、さまざまな学習行動をとる。本研究は、図1のような、来館後に自分の興味があったトピックについて学習を行うようなシチュエーションを想定している。ユーザの館内での行動やデバイス操作で興味を抽出する研究は以前にも行われている。本研究は、それらに加え、事後学習中の興味喚起を反映したシーン区間を提示することで、効果的な視聴支援を行う。本システムでは、事後学習中の興味喚起をリアルタイムで取得するために、UI上でタグの操作を行ってもらい、タグの付与は映像中のナレーションをテキスト化したデータを形態素解析し、名詞を抽出することでシーンのトピックとなっているキーワードを抽出する。付与されたタグや、それに

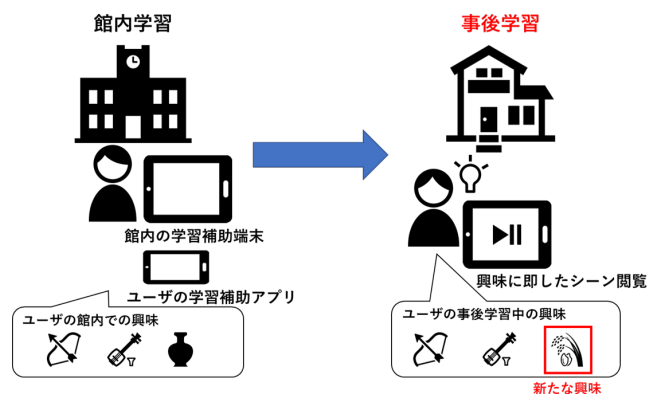


図1 想定しているシチュエーション

関連するタグなどを事後学習画面に表示し、タグの操作ログを利用して各タグに興味スコアを付与する。これらの処理と館内での興味スコアを組み合わせ、適切なシーンを提示することで効果的な事後学習のための映像視聴支援を実現する。

## 2 システム概要と関連研究

### 2.1 システム概要

本研究で使用する事後学習 UI は図2のようなものを想定している。このUIでは、画面右側にタグが配置されており、2種類のタグ表示形式がある。画面左側にシーン映像を表示し、左側下部に再生/停止ボタンと、次のシーンへ進むボタン、戻るボタンを配置している点は共通している。ユーザはこれらのボタンで映像の操作を行う。

図2の左の操作画面では、8つのタグが表示されている。このうち4つは再生中のシーンに対して付与されたタグであり、



図2 事後学習 UI の操作画面

残りの4つは、視聴時点で後述する手法に基づき推薦されたタグである。シーン視聴中のユーザは、いずれかのタグをタップすることでタグ操作を行う。すると、右の操作画面のようにタップしたタグが画面右側中心に移動し、その周りに当該タグに関連するタグが配置される。このようなタグ操作を行いながら、ユーザは自分の興味のあるタグを見つけたのち、検索ボタンでシーンを検索する。検索ボタンがタップされると、システムはユーザのタグ操作を考慮した検索クエリを生成し、シーンデータベースから関連度の最も高い映像を表示する。

システム側では、タグ操作のログを元に各タグに対してスコアリングを行い、その結果からタグに興味スコアを付与していく。このスコアリングでは、タグ操作の回数と Wikipedia カテゴリ構造上でのタグのトランジションからスコアを算出する。このスコアを元に検索クエリを生成し、ユーザの興味に即したシーン区間を表示することで映像視聴を支援する。

## 2.2 関連研究

近年、博物館での体験をテーマとした研究が盛んに行われている。Misu ら [2] は、ボストン科学博物館内で使用されている音声ガイドキャラクターと、来館者の会話ルールを強化学習を用いて改善する手法について提案している。また、神門ら [3] は、みんなく博物館内で使用するための展示物検索アプリケーションを提案している。このアプリケーションでは、展示物の提示型検索を行う UI を提案しており、展示物の詳細情報や手紙メモ機能、位置情報に基づく展示物推薦を実装している。2020年に発表された荘司ら [4] の研究では、それらの機能に加え、事前に館内の展示物鑑賞コースと鑑賞の観点を作成するウェブアプリケーションと、事後学習支援のための展示物の写真で構成されたポストカード自動生成アプリケーションを提案している。Roberts ら [5] の研究では、展示物の推薦やメモ機能を搭載することで鑑賞体験の向上を図っている。しかし、これらの研究は博物館での体験のうち、事前学習と館内での鑑賞行動に着目したものが多く、そのため、本研究は、事後学習自体に焦点を当てている。

ユーザの興味、趣向を利用した博物館の研究も数多く存在する [6] [7]。Hatala ら [8] は博物館向けの音声解説システムであ

る "ec(h)o" を利用して、ユーザが興味を持つであろう展示物を推薦するシステムを提案している。瀧平ら [9] は、国立民族博物館の館内アプリケーションから取得した来館者の行動ログを利用し、観覧体験を記念品化するポストカードを生成するシステムを提案している。Kakimoto ら [10] は、来館者が館内で操作したデバイスのログデータから興味を抽出し、事後学習に生かす手法を提案している。これらの研究は、博物館内での学習行動のみをインプットとしているが、本研究では博物館内の興味に加え、事後学習中に喚起した新たな興味にも対応するシステムを提案する。

また、博物館の事後学習に関する研究も盛んに行われている。Spence ら [11] は、来館者が友人にシェアしたいと感じた展示物の写真を撮り、なぜシェアしたいと感じたのかを説明する音声とともに共有できるアプリケーションを提案している。しかし、事後学習に必要な要素は新たな興味の喚起に加え、より専門的な知識を提供することである。そのため、本研究では事後学習中の操作ログが映像提示に影響を与えるものにした。竹島ら [12] は、博物館内で利用されているガイドシステムの利用履歴から、来館者の履歴に関連する Web ページを提示するシステムを提案している。しかし、Web ページの情報は正確性や詳細性に欠ける場合があり、ユーザの求める事後学習コンテンツを提示できるか不確実なものとなっている。しかし、本研究では博物館が所有している映像資料を利用するため、提示する情報の正確性や詳細性を担保できると考えられる。

e ラーニングのための動画利用に関する研究では、Huh ら [13] が e ラーニングに利用されている動画に Web のリソースをタグ付けすることで、提供する情報を拡張する提案を行っている。また、スマートビデオを e ラーニングに適用することで、ユーザが動画からさまざまな情報へ容易にアクセスできるとしている。本研究でも、ユーザの興味に即したさまざまな情報をリンクした事後学習システムを目指すという点で参考にした。

Wikipedia の活用という観点では、楊ら [14] が来館者の観覧行動を能動的にするという目的で、Wikipedia ページのリンク構造を利用して、展示物と来館者の興味の関連性を見つける手法を提案している。事後学習でユーザが起こすであろう学習行

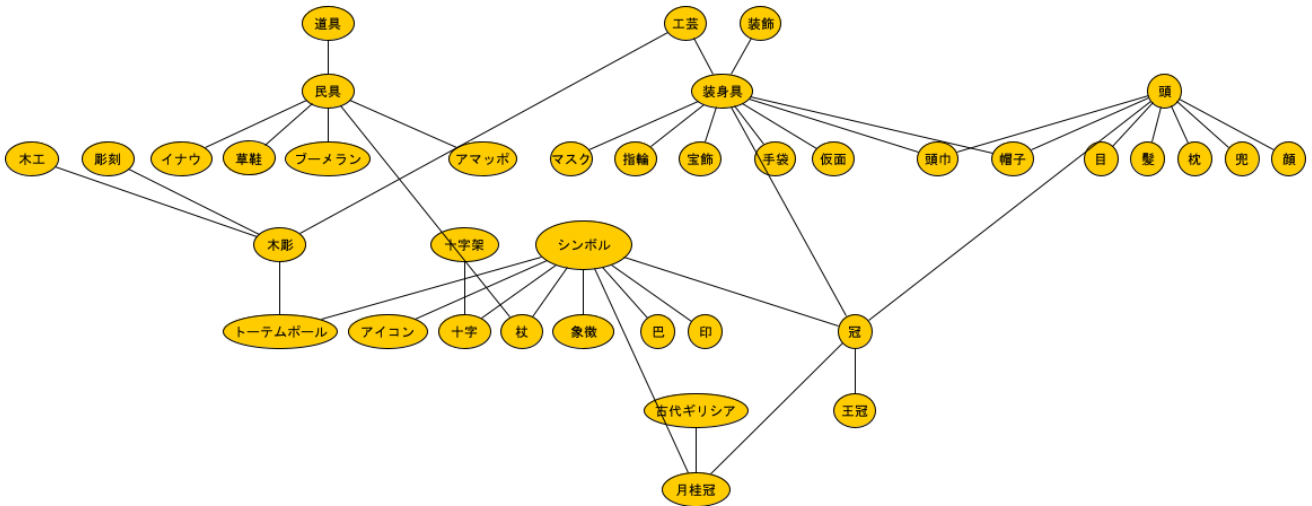


図3 タグ「シンボル」から3エッジ以内のWikipediaカテゴリとページ

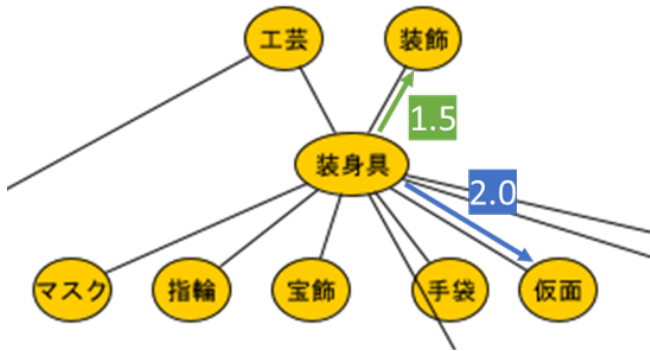


図4 T算出の例

動としては、新たな興味の喚起と、現在抱いている興味の深掘りである。Wikipedia ページのリンク構造で抽出する情報は関連性が低いものも多く、興味とコンテンツの概念的つながりを視覚化しにくい。そのため、本研究ではWikipedia カテゴリ構造のシソーラスを利用することで、タグ間の包含関係を加味した興味の抽出を行う。

### 3 提案手法

まず、本システムでは博物館の資料映像をシーンに分割し、各シーンの音声をテキスト化したデータをもとにシーンに対してタグを付与していく。次に、付与されたタグとWikipedia カテゴリ構造を利用し、タグへのスコアリングを行う。その結果から、ユーザーに適した短いシーンを視聴させる形式をとっている。

#### 3.1 シーン分割

博物館が所有している映像をシーンに分割し、事後学習での映像視聴を容易にする。本研究では、Microsoft Azure Video Indexer [15] というツールでシーンに分割する。本システムで利用する博物館映像は、ナレーション音声があてられているという想定のため、音声のテキスト化もこのツールを利用する。

実際に、7分程度のナレーション付き映像は5シーンに分割された。また、ニュース番組を利用し、テキスト化の精度も測定したところ、93.6%だった。この結果から、おおむね正確にテキスト化できていると考えられるため、こちらのツールを利用する。

次に、シーンのトピックとなっているキーワードを抽出し、タグとして付与する。まず、テキスト化した音声データを形態素解析し、名詞のみを抽出する。今回はMeCab [16] という形態素解析ツールを利用した。また、解析に使用する辞書は、更新頻度が高く、新語・専門用語を多く含んでいるMeCab-ipadic-NEologd [17] を使用する。解析された名詞は、Wikipedia カテゴリ、またはWikipedia ページが存在するもののみが抽出され、各シーンで最もTF値の高い名詞の上位4件を該当シーンへタグとして付与する。多くの動画で、シーン中出现する名詞群の多くがTF値1の単語だったため、ここでタグとして付与される名詞にTF値1のものは含まないものとする。以上の処理をもって博物館の所有する映像資料をシーンに分割し、タグを付与する工程としている。

#### 3.2 UI上での関連タグの表示

前章でも説明したように、本研究の事後学習UIではユーザーが映像の視聴を行いながら、画面右側でタグの操作を行うことを想定している。そのため、まずWikipedia カテゴリ構造を利用し、ユーザーがタップしたタグに関連するタグの提示を行う。

ユーザーがタップしたタグは、シーン分割の工程により、Wikipedia のカテゴリないしページとして存在するものに限定されている。そのため、タップしたタグがカテゴリ名だった場合、該当カテゴリの上位/下位カテゴリと、該当カテゴリに分類されているページ名をエッジで接続し、関連タグとして表示する。一方、タップしたタグがページ名だった場合、該当ページが分類されているカテゴリ名をエッジで接続し、関連タグとして表示する。例えば、「シンボル」というワードが最初にタップされた場合、図3のような構造がWikipedia から抽出

$X_t$	シンボル, トーテムポール, 木彫, 彫刻, 木彫, 工芸, 装身具, 冠, 頭, 帽子, 装身具, 工芸 2.0            2.0            1.5 1.5    1.0 1.5    2.0 2.0 1.5 2.0    3.0 3.0
$Y_t$	シンボル, トーテムポール, 木彫, 彫刻, 木彫, 工芸, 装身具, 冠, 頭, 帽子, 装身具, 工芸 0.50            0.50            0.25 0.25    0 0.25    0.50 0.50 0.25 0.50    1.00 1.00
$Interest_t$	シンボル, トーテムポール, 木彫, 彫刻, 工芸, 装身具, 冠, 頭, 帽子 0.50            0.50            0.13 0.25    0.63 0.75 0.50 0.25 0.50

図 5 興味スコア算出の例



図 6 システムの実行例

される。これは、wikipedia カテゴリ「シンボル」をノードとして、3エッジ以内に存在するカテゴリとページを抽出したものである。各ノードの上位にエッジで接続されているノードが、上位カテゴリもしくはページが分類されているカテゴリであり、下位にエッジで接続されているノードが、下位カテゴリもしくは当該カテゴリに分類されているページである。エッジで接続されているすべてのカテゴリとページを抽出すると、あまりに多くのノードが抽出されてしまうため、ほかの動画内で出現している名詞という条件で抽出を行う。今回は博物館の映像資料の実データを用意できなかったため、国立民族学博物館 HP の映像資料目録データベース [18] にて、該当カテゴリ/ページ名を全て検索し、ヒットしたものを抽出している。本システムでは、シーン再生時に表示されている 8 つのタグから、3エッジ以内のカテゴリとページをタグとして抽出し、操作されたタグにエッジでつながっているタグを、関連タグとして表示する。これにより、ユーザに対して新たな興味の喚起と興味の詳細化の両方を促す。

### 3.3 学習行動からの興味抽出

事後学習 UI 上でタグ操作を行った際に行われる、タグに対するスコアリング手法について説明する。前章で説明した関連タグの表示により、「タップしたタグ」、「戻る/進むボタンで移動した先のタグ」、「検索ボタンをタップした直前にタップしたタグ」という 3 種類のタグ操作ログを取得することができる。これらのログを利用してスコアリングを行っていく。

まず、各ログ地点でのタグに対するスコア  $X_t$  を式 (1) で

算出する。このスコアは、当該タグへの操作回数である  $I$  と、Wikipedia カテゴリ構造上のトランジションである  $T$  を掛けて計算している。ここでは、図 4 に示しているように、上位へのタグ操作は 1.5、下位へのタグ操作は 2.0 を  $T$  としている。なぜなら、上位への移動は上位概念への興味の変遷と同義であり、概念を広くすることで新たな興味を喚起するタグ操作だと定義できるからである。一方、下位への移動は、ユーザが抱いている興味を詳細化するタグ操作であり、より強い興味を示したと定義できるため、上位への移動よりも高い数値である 2.0 を与えている。また、戻る/進むボタンの操作で画面を遷移させた場合、タグへの興味を喚起したとは言えないため、 $X_t$  には一律で 1.0 を与えている。その結果を式 (2) で、0 から 1 に正規化した数値である  $Y_t$  を算出し、式 (3) でそれぞれのタグの  $Y$  の平均を算出したものを興味スコアとしてタグに付与する。

$$X_t = I \times T \quad (1)$$

$$Y_t = \frac{X_t - X_{min}}{X_{max} - X_{min}} \quad (2)$$

$$Interest_t = \frac{\sum_{i=1}^n Y_{t,i}}{f_t} \quad (3)$$

これらの計算をログの例を用いて実際に行い、興味スコアを算出したものが図 5 である。緑で示されているタグ操作ログが下位への操作、青で示されているタグ操作ログが上位への操作、黒で示されているログが戻る/進むの操作ログである。ログでは、最終的に「工芸」と「装身具」のタグで 0.5 より高いスコアが出ており、ほかのタグより興味を強く示した結果となって

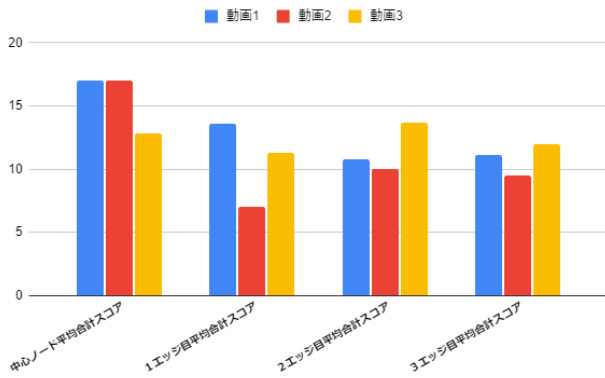


図 7 動画ごとの各項目平均合計スコア

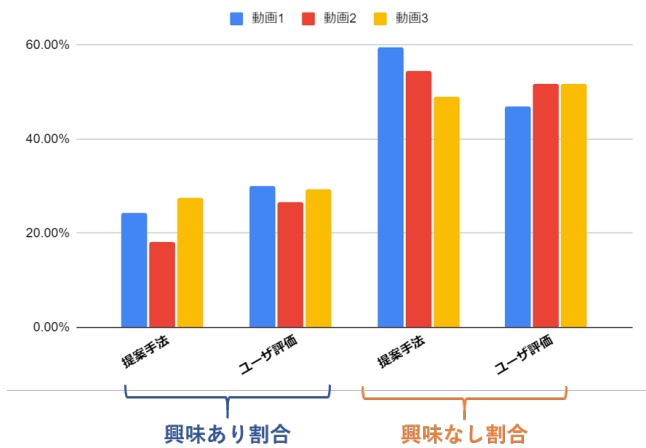


図 8 動画別の興味の割合

いる。

本システムは、館内でのデバイス操作からも興味を抽出し、事後学習中のシーン検索に影響を与える。しかし、館内で使用したデバイスからの興味抽出は、Kakimoto ら [9] の手法を利用することで、本研究の手法と同じように正規化された数値を算出できる。そのため、本システムではこの興味抽出手法を用いてユーザの興味を事後学習に利用する。

### 3.4 シーン検索

以上のスコアリングをもとに、興味スコアが付与されたタグを利用してシーン検索クエリを生成する。本システムの UI では、ユーザが興味のあるタグを見つけるためにタグ操作を行い、検索ボタンをタップする手順となっている。そのため、シーンを検索するクエリは最後に操作したタグ *or* スコアの一番高いタグに設定する。これらの処理を繰り返しながら、ユーザの興味抽出と視聴支援を同時に行う。

## 4 評価実験

### 4.1 システムの実行例

図 6 に、システムを実際に実行した際の遷移図を示している。この遷移図は、図 5 のログと同じ動きである。まず、ユーザは最終的に工芸というタグに興味を持ち、工芸タグをタップした

のち検索ボタンをタップしたとする。すると、システムはシーン検索クエリの生成を行う。今回は、最後にタップした「工芸」と、検索ボタンをタップした時点で一番興味スコアの高いタグという想定である「北アメリカ」で *or* 検索を行っている。その結果、ヒットした「北アメリカ 北西海岸民族の工芸」というタイトルの映像のシーンが再生される。

図 6 の実行例では、ユーザが「工芸」というタグに興味を示し、検索を行った。その結果、「工芸」に加え、現在興味のある「北アメリカ」を検索クエリに加えることで、ユーザにパーソナライズした検索が行える。また、ユーザがさらに当該トピックについて知りたい場合はそのまま類似度が 2 番目に高いシーンへ移動することができる。本システムを利用することで、ユーザはタグ操作と視聴行動という最低限の操作で映像を視聴することができる。

### 4.2 wikipedia カテゴリ利用の評価実験

#### 4.2.1 予備実験と評価コンテンツ

ユーザの興味を喚起するために、Wikipedia カテゴリ構造を利用することの有効性を検証するための評価実験を行った。評価実験を実施するにあたり、まず、ユーザが映像視聴中にどのような興味を喚起するのかを調査する予備実験を行った。この実験の結果は、比較実験にて提案手法である Wikipedia カテゴリ構造を利用して抽出した名詞群と比較を行うために用いる。使用した 3 つの映像 [19] [20] [21] は、いずれもナレーションがついたものであり、文化・民族に関するトピックを扱っている。また、予備実験と評価実験では、提案手法で分割されたシーンの中から我々が指定したシーンを視聴してもらった。動画の詳細は表 1 に示している。

予備実験では、10 代から 40 代の被験者 8 名が、3 つの映像について以下の設問に回答した。設問 1 では、動画中に出現した音声のテキスト化データを形態素解析し、名詞のみを抽出したリストの中から 1 - 3 位を選んでもらうという方式をとっている。設問 2, 3 では、一つの設問で 2 つの必須回答項目を設定できなかったため、同じ質問を両方必須項目として二つ回答を求める形をとっている。結果として、各動画で選択式から 3 語、自由記述式から 2 語の合計 5 語をユーザ抽出の名詞として得られる。

(1) 動画を視聴し、あなたが次に興味のある動画を探すときに適切だと思う検索ワードを 3 語選んでください。

(2) 動画を視聴し、あなたが次に興味のある動画を探すときに適切だと思う検索ワードのうち、選択肢にないもので”名詞”を 1 つ記述してください。

(3) 動画を視聴し、あなたが次に興味のある動画を探すときに適切だと思う検索ワードのうち、選択肢にないもので”名詞”を 1 つ記述してください。

予備実験の結果、動画 1 では 20 個、動画 2 では 24 個、動画 3 では 27 個の名詞を収集できた。個数の増減は、複数人が同じ名詞を回答したことによるものである。結果の定性的考察は以下のとおりである。

- 設問 1 の回答は、いずれの動画も TF 値の高い名詞が極

表 1 実験に使用した動画の詳細

動画番号	内容	視聴してもらうシーン長	予備実験名詞数	wikipedia から抽出されたタグ数
動画 1	ラマレラのクジラ漁	1 分 58 秒	20	74
動画 2	奈良の墨職人	2 分 17 秒	24	11
動画 3	アイヌの食事	3 分 38 秒	27	54

端に選ばれていた。一方、ほかの名詞はまばらに選ばれていたことから、シーンのトピックとなる単語を確実に抽出し、そこに関連する単語を Wikipedia カテゴリ構造から抽出することが重要だと考えられる。

- 自由記述の名詞は、大きく 2 種類に分類することができた。
  - － 出現した名詞と概念的につながりのあるもの。  
例：クジラ漁の動画で、「捕鯨」と回答された
  - － 出現した名詞や、関連した事柄から想起・連想したもの。  
例：クジラ漁の動画で、「社会問題」など。（捕鯨関連のトピックが国際的に問題になっていることから想起したと考えられる）
  - － 前者のパターンは、Wikipedia カテゴリ構造上で比較的少ないエッジで接続されているものであり、後者は比較的多くのエッジを経て抽出できる名詞だと考えられる。

また、予備実験で回答された名詞の内、自由記述式で 2 名以上が記述した名詞は以下ようになった。

- 動画 1: 捕鯨, 民族, 鮭, 海
- 動画 2: 伝統工芸
- 動画 3: 文化

以上の名詞は、各動画において多くの人が興味を喚起する名詞と考えられる。

#### 4.2.2 評価実験の結果と考察

評価実験では、予備実験で得られた名詞と、提案手法により Wikipedia カテゴリ構造から抽出された名詞とを比較することで、Wikipedia カテゴリ構造を利用することの有効性を検証する。提案手法では TF 値の高いタグから 3 エッジ以内のカテゴリとページ名をタグとして抽出しているが、この評価実験ではユーザが構造上で抽出の起点となったノードから、どの程度のエッジ数まで興味を示すのかも併せて検証する。

評価は、10 代から 30 代の 5 人の被験者が各動画の指定シーン区間を視聴したのち、Wikipedia カテゴリ構造から抽出された名詞と、予備実験にて集めたユーザが想起した名詞すべてを 5 段階で行った。使用した動画は予備実験と同じものである。

図 7 は、各名詞の 5 段階評価を 1-5 のスコアとして計算し、様々な条件で平均スコアを算出したグラフである。

このグラフでは、カテゴリ構造上で、実験に使用したシーンに付与されたタグを起点ノードとして、何エッジ目までが興味を喚起するタグとなるのかを分析する。まず、動画に付与され、カテゴリ抽出の際に起点のノードとして扱われる名詞の平均スコアは、動画 1 と動画 2 において高い数値を示した。この要因として、提案手法では起点のノードに、シーン中で一番出現率の高い単語を利用したため、被験者が強く興味を残したのではないかと考えられる。一方、動画 3 のスコアが低くなっている

のは、同じ TF 値の名詞が多く出現したため、中心ノードが多くなり、シーンのトピックとして扱えない単語まで抽出されたためだと考えられる。

次に、起点のノードから 1 エッジの距離にある名詞について考察する。動画 1 では、1 エッジ目の名詞が 2 エッジ、3 エッジ目の名詞に比べて高い平均スコアを示した。それに対して、動画 2 の平均スコアは非常に低い平均スコアとなった。この原因として、動画 2 における 1 エッジの距離にある名詞が 1 つしか抽出できなかったことで、被験者 5 人の経験や知識的趣向が色濃く表れたのではないかと考えられる。起点のノードから 2 エッジの距離にある名詞は、動画 3 が特に高い平均スコアを示した。中心ノードから 3 エッジ目の名詞の平均スコアでは、比較的大きな差は見られなかった。しかし、1 から 3 エッジの全距離を通して動画 2 のスコアが 3 つの動画中最も低くなっている原因は、表 1 から確認できるように、Wikipedia カテゴリ構造から抽出された名詞数がほかの動画に比べて大幅に少ないことが原因だと考えられる。中心ノードや 1 エッジ目に抽出された名詞が少なければ、接続されているエッジ数自体が少なくなる。そのため、2 エッジ目や 3 エッジ目の名詞抽出数も少なくなってしまう。この欠点を解消するため、シーン中に出現する名詞数があまりに少ない場合は、前後のシーンと結合するなどの処理を行うことで、一定数の抽出数を確保する必要があると考える。全体として、何エッジ目までが有効だと断定できる結果は得られなかったが、ユーザが自分で想起した名詞と評価が大きくは変わらないことを確認できた。また、Wikipedia カテゴリ構造のグラフ上で距離が遠い場合でも、ユーザの興味が薄れることはなく、少なくとも 3 エッジ目まではある程度の興味を喚起する名詞を抽出できることを確認できた。

次に、5 段階評価を行った名詞群のうち、1 と 2 の評価を興味なし、4 と 5 のものは興味ありとしてみなし、その割合を提案手法の名詞群と、予備実験にて収集したユーザ評価の名詞群で、別々に算出したグラフが図 8 である。このグラフから、提案手法で抽出された名詞群の方が、ユーザ評価の名詞群より興味あり割合が少し低くなっていることがわかる。特に、動画 2 においては興味ありの割合が低い。この原因として、前述したように動画 2 の名詞抽出数が少なく、そのいずれれもがユーザの興味に合致しなかったためだと考えられる。これを解消するためには、動画によって Wikipedia カテゴリ構造からタグを抽出する際のルールを緩くし、必ず一定数のタグを抽出できる方法を検討する必要がある。また、動画 1 では提案手法の方が、ユーザ評価の名詞より興味なし割合が高くなっている。この原因として、提案手法では名詞を抽出しすぎる場合もあり、動画を視聴して喚起する興味の範囲外に存在する名詞まで推薦してしまったことが原因として考えられる。そのため、名詞の抽出

を行う際は、最低・最高抽出数の閾値を設定する必要がある。

最後に、予備実験と提案手法で抽出された名詞の比較を行う。予備実験の自由記述形式で回答された名詞と、提案手法により抽出された名詞の内、重複しているものがいくつか見られた。動画1では、捕鯨・鮫・海の3語、動画3では、少数民族、民族の2語が重複していた。動画2では重複した名詞は見られなかった。これらの名詞は、ユーザが抱く興味を反映してきた例である。再現率は、動画1では自由記述で記入された10種類の名詞の内3語で、33.3%、動画3では15種類の内2語で、13.3%であった。両動画ともそこまで高い数値を記録したわけではないが、興味はユーザの知識や経験に深くかかわっているため、予備実験の被験者を増やし、多様な興味と比較することで、再現率は高まる可能性がある。

## 5 おわりに

本研究は、博物館が所有する映像資料をユーザの興味をリアルタイムに反映しながら、シーンを横断的に視聴する事後学習システムの提案を目的としている。事後学習中のデバイス操作を考慮することにより、ユーザが事後学習中に喚起した興味の変化を、映像視聴に反映できると考える。評価実験では、Wikipediaから抽出される単語の量によって興味喚起の度合いに差が出るのがわかった。現在の手法では、シーンに出現する名詞の量にタグ推薦が大きく左右される。そのため、動画によってタグ抽出手法を変えることが必要である。

また、今回はWikipediaカテゴリ構造のシソーラスを用いた手法の提案と実験を行った。今後は、WordNetなどの別のシソーラスや、LSTMといったモデルの学習を伴う分散表現の取得によるタグ推薦も視野に入れ、手法を改善していく。また、プロトタイプを作成し、スコアリングやUIの評価実験を実施し、システム自体の評価を行うことも必要だと考える。

## 謝 辞

本研究の一部は、2020年度国立情報学研究所公募型共同研究(20FC04)の助成を受けたものである。ここに記して謝意を表す。

## 文 献

- [1] Shao-Chun Wu. 2019. Online learning and opinions of educator: a quantitative study of museum educational video platform's user. In Proceedings of the 8th International Conference on Informatics, Environment, Energy and Applications (IEEEA '19). Association for Computing Machinery, New York, NY, USA, 258–262. DOI:https://doi.org/10.1145/3323716.3323756
- [2] Teruhisa Misu, Kallirroi Georgila, Anton Leuski, and David Traum. 2012. Reinforcement learning of question-answering dialogue policies for virtual museum guides. In Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL '12). Association for Computational Linguistics, USA, 84–93.
- [3] 神門典子, 大島裕明, 相原建朗, 荘司慶行, 白石晃一, 瀧平土夫, 中島悠太, 山本岳洋, 山本祐輔, 楊澤華, 提示型検索に基づくミュージアム電子ガイドを中核とした事前・事後学習

支援, じんもんこん 2020 論文集, 81-88, 2020.

- [4] 荘司慶行, 大島裕明, 神門典子, 相原建朗, 白石晃一, 山本岳洋, 山本祐輔, 楊澤華, 提示型検索モデルに基づくミュージアム鑑賞体験の提案, じんもんこん 2019 論文集, 127-132, 2019.
- [5] Jessica Roberts, Amartya Banerjee, Annette Hong, Steven McGee, Michael Horn, and Matt Matcuk. 2018. Digital Exhibit Labels in Museums: Promoting Visitor Engagement with Cultural Artifacts. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems (CHI '18). Association for Computing Machinery, New York, NY, USA, Paper 623, 1–12. DOI:https://doi.org/10.1145/3173574.3174197
- [6] B. Fabian, "Personalising the museum experience," *Pervasive User Modeling and Personalization (PUMP 2010)*, p.33, 2010.
- [7] Y. Taizo, A. Kenro, K. Noriko, F. Satoko, U. Yusuke, B. Takayuki, N. Shigemi, T. Takashi, H. Yuko, and J. Adachi, "Ceax's learning support system to explore cultural heritage objects without keyword search," *Current Developments in Technology-Assisted Education*, pp.318–322, 2006.
- [8] Hatala, M., Wakkary, R. *Ontology-Based User Modeling in an Augmented Audio Reality System for Museums*. *User Model User-Adap Inter* 15, 339–380 (2005). https://doi.org/10.1007/s11257-005-2304-5
- [9] 瀧平土夫, 荘司慶行, 山本岳洋, 山本祐輔, 大島裕明, 博物館における鑑賞体験の記念品化を目的とするナビゲーション端末操作ログからの印象深い展示物推定. *DEIM Forum 2020*, D4-5, 2020年.
- [10] Honoka Kakimoto, Yuanyuan Wang, Yukiko Kawai, Kazutoshi Sumiya, *User Interests Extraction by Analyzing User Operations on Multimedia Museum Devices*, *DEIM Forum 2020*, D4-5, 2020.
- [11] Jocelyn Spence, Benjamin Bedwell, Michelle Coleman, Steve Benford, Boriana N. Koleva, Matt Adams, Ju Row Farr, Nick Tandavanitj, and Anders Sundnes Løvlie. 2019. Seeing with New Eyes: Designing for In-the-Wild Museum Gifting. In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). Association for Computing Machinery, New York, NY, USA, Paper 5, 1–13. DOI:https://doi.org/10.1145/3290605.3300235
- [12] 竹島さや香, 中川千種, 守屋和幸, 大野照文, 博物館での見学履歴を活用したWebによる事後学習支援システムの提案. *情報処理学会第68回全国大会*, 5V-3, 2006年.
- [13] S. Huh, Y. Park, J. Jang and W. Choi, "Making a video smart for smart e-learning," 2015 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, 2015, pp. 858-863, doi: 10.1109/ICTC.2015.7354686.
- [14] 楊澤華, 山本祐輔, 山本岳洋, 神門典子, 大島裕明, 博物館の展示物と見学者の興味を関連付ける情報の発見. *DEIM Forum 2020*, E4-3, 2019年.
- [15] Microsoft Azure Video Indexer: https://vi.microsoft.com/ja-jp/.
- [16] MeCab: https://taku910.github.io/mecab/.
- [17] MeCab-ipadic-NEologd: https://github.com/neologd/mecab-ipadic-neologd
- [18] 国立民族学博物館 映像資料目録: https://htq.minpaku.ac.jp/databases/av/movcat.html
- [19] 実験に使用した動画 1: https://youtu.be/ljB6sfvdc?t=115
- [20] 実験に使用した動画 2: https://youtu.be/hP5J9FKxeBY?t=70
- [21] 実験に使用した動画 3: https://youtu.be/kDH4itkYJs0?t=31