

# 視聴操作を考慮した地理的關係性に基づく ビデオ視聴支援システムの提案

橋本 樹<sup>†</sup> 王 元元<sup>††</sup> 河合由起子<sup>†††,††††</sup> 角谷 和俊<sup>†</sup>

<sup>†</sup> 関西学院大学総合政策学部 〒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>{ewd16738,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、動画配信サービスの普及により、ビデオ映像の関連情報を推薦する研究が盛んに行われている。たとえば、NHKの次世代の放送通信連携サービス「NHK Hybridcast」[20]では、テレビ番組の放送中に同画面上で関連情報の推薦を行っている。また、YouTubeではチャプター機能が実装され、ユーザの視聴操作の有用性が向上してきている。しかし、現在はユーザの視聴操作を考慮して補足情報推薦を行うサービスは実装されていない。そこで本研究では、ユーザの選択シーンを抽出可能なユーザインタフェースを用いてユーザの視聴動作を抽出する。それにより、シーン選択から得られた各シーンの地理的情報を持つキーワードと各シーンでの重要語を用いて空間的距離および意味的距離を算出することで、補足情報の検索クエリを生成する。本論文では、視聴者の視聴シーン履歴を用いた視聴者の視聴傾向抽出に基づくビデオ視聴支援システムの提案とその検証を行う。

キーワード マルチメディア、トピック抽出、地理情報、情報推薦

## 1. はじめに

Tver や GyaO など見逃し配信を行う動画配信サービスの普及や Netflix や hulu などのサブスクリプションサービスの普及により、放送時間などを気にして動画を視聴することが減っている。そのため、興味のあるシーンのみを視聴する「飛ばし見」を行うユーザが増え、動画全体を視聴する「流し見」を行うユーザが減ってきていると考えられる。本研究では、そのような研究背景に基づいて、主に「飛ばし見」を行うユーザを対象としている。上記の前提に基づき、ユーザの視聴操作を取得可能なユーザインタフェースにより、「飛ばし見」を行うユーザが選択したシーン履歴を抽出する。抽出された選択シーン履歴はユーザが求める情報であると仮定し、選択シーン履歴内に出現するキーワードを用いてクエリ式を生成する。これにより、ユーザの視聴操作によって推薦する情報が変化するため、個人化された情報推薦を行うことが可能である。また、従来のビデオ映像を用いた研究ではクロズドキャプション内に出現するキーワードからトピックを抽出し情報推薦を行う場合が多い。しかし、本研究では、キーワード群より得られるトピックに加えて、地名や地域など地理的情報を持つキーワードを使用する。

本稿では、地名など地理的情報を持つキーワードを空間的キーワードと定義する。空間的キーワードを使用することで地名や地域は集合記号を用いて表現することが可能であるため、シーンごとで得られる空間的キーワードの地理的關係性を用いて、推薦する情報を判定することが可能である。たとえば、比較対象のシーンで「兵庫県」に関する内容を放送し、視聴中のシーンで「関西学院大学」に関する内容を放送している場合、

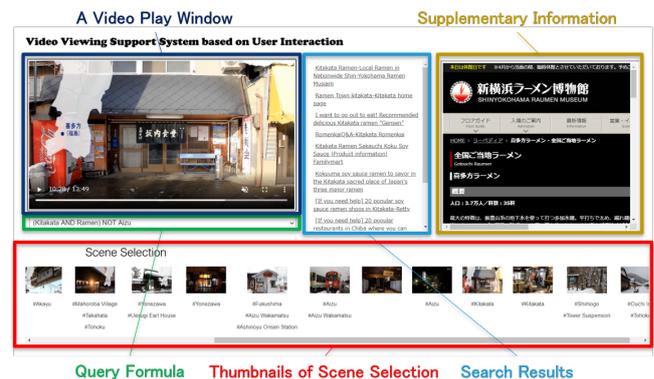


図1 本研究で使用するユーザインタフェース

「兵庫県」に対して「関西学院大学」は包含関係にあり、「兵庫県 $\supset$ 関西学院大学」と表現される。この関係性により、シーン間の地理的關係性が「兵庫県」から「関西学院大学」に範囲が狭まっているため、番組内容では「兵庫県の関西学院大学」という流れで、より詳細な内容を放送されることが考えられる。このように、シーン間の地理的關係性を使用することにより、推薦すべき情報の種類を判定することが可能である。本研究では、それらに対応したクエリ式の生成方式を提案し、これにより、ユーザの視聴傾向を考慮した補足情報として適切な Web ページの推薦を目指す。

## 2. 関連研究

近年、ビデオ映像のクロズドキャプションなどのメタデータを用いて、番組内容に関連する情報を推薦する研究が多く行

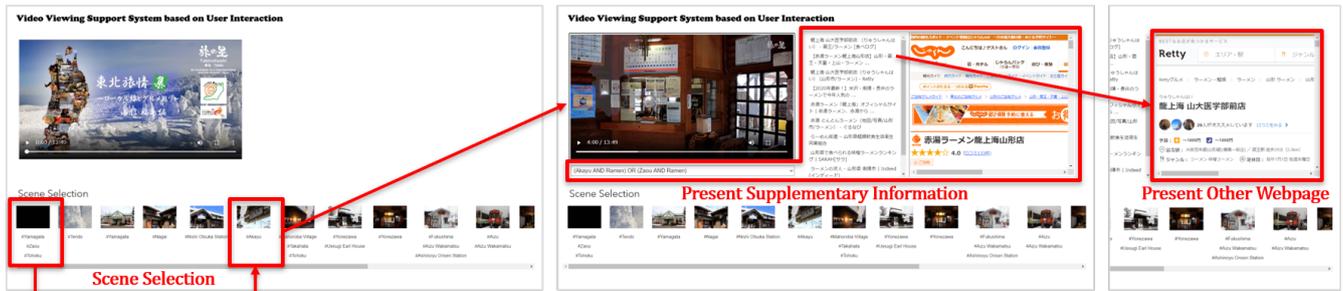


図 2 ユーザインタフェースの画面遷移図

われている。

先行研究として、王ら [1] [2] [3] や柿本ら [4] はテレビ番組のクローズドキャプション内に出現する地理的な情報を持つキーワードを用いて地理情報や関連する情報を推薦するシステムを提案している。これにより、地理的関係を考慮した情報推薦は有効であると考えられる。また、クリペンデルら [5] はクローズドキャプションに出現する単語ごとに重要度を付与する手法を提案し、手法として  $TF_iDF$  を使用して重要度を算出した。三浦ら [6] は単語間の意味的関係の抽出手法を提案した。これらの研究によりビデオ映像のクローズドキャプションからクエリキーワードを生成し、情報推薦を行うことが有効であると考えられる。よって、本研究で行っている地理的関係性に基づくクエリ生成手法の提案も有効性が高いと思われる。

また、動画を対象とした研究として、トピックの変化に基づくシーン分割に関する研究は盛んに行われている。Baraldiら [7] はビデオフレーム内に存在するトピックの時間的変化を利用したシーン分割手法の提案を行い、Sidiropoulosら [8] はBGMやナレーションなどの音声データを用いたシーン分割手法の提案を行った。他にも、Rasheedら [9] はビデオフレーム内のオープンキャプションやクローズドキャプションなどのテキストデータを用いて、映画やテレビ番組の構造の分析を行った。対照的に、Liu [10] らは動画のショットに注目し、ショットに表示されるオープンキャプションやナレーション、BGMなどのテキストデータと音声データの2つを用いたシーン分割手法の提案を行った。

さらに、ビデオフレームを用いた研究では、ビデオ要約に関する研究も盛んに行われている。ビデオ要約では、シーン分割と同様にトピックの変化や場面の切り替わりを判定し、個人化された短いビデオの推薦を行う。本研究ではユーザのシーン選択を用いているため、ビデオ映像全体をユーザが視聴しないという点などは内容的に近いと考えられる。しかし、動画の要約は動画制作者の意図を考慮する必要があるため、非常に困難であると考えられる。[11] [12] [13]

また、本研究では地理情報に関連する内容も含むため、ビデオ映像を対象とした研究以外も参考している。北山ら [14] や小林ら [15]、倉田ら [16] は地理情報を推薦する手法を提案している。これらの研究では、地理情報を推薦するために用いる特徴ベクトルの抽出が必要不可欠であり、倉田ら [16] は地理情報推薦に Topological Relationships [23] を応用する手法を提案し

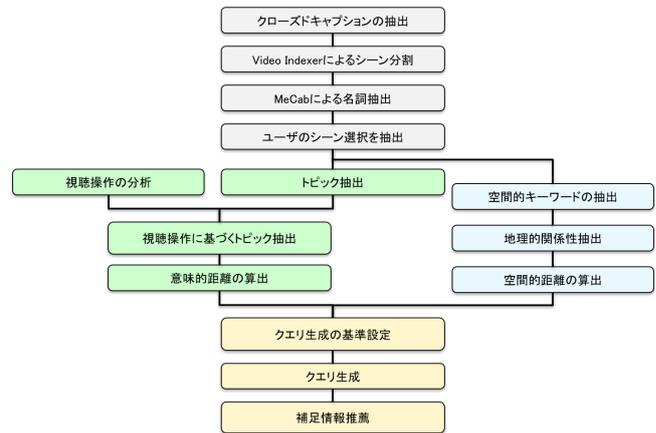


図 3 システム概要図

ている。本研究においても Topological Relationships [23] を用いた手法は有効であると考え、地理的関係性抽出に本手法を使用している。

### 3. システム概要

本研究では、シーン選択可能なユーザインタフェース (図 1) により得られるユーザの視聴操作を用いて、ユーザの視聴操作に基づく映像の補足情報提示手法を提案する。図 2 はユーザインタフェースの画面遷移図であり、ユーザのシーン選択に基づいて表示する内容が変化する。システム概要は図 3 に示している。本手法はテレビ番組のクローズドキャプションを抽出し、MeCab [21] を用いて形態素解析を行い、番組内に出現する名詞のみを抽出する。シーン分割には Microsoft Azure の Video Indexer [22] を使用し、分割されたシーンごとのサムネイルをユーザインタフェースの下部に表示する。サムネイルをクリックすると該当シーンの開始時間にスキップして動画が再生される。ユーザインタフェースにより得られるユーザのシーン選択をユーザの視聴操作と定義し、ユーザの視聴操作により得られる選択シーン内のキーワードと空間的キーワードを用いて、トピック抽出と地理的関係性抽出を行う。抽出された地理的関係性の種類によって、ブーリアン検索モデルに基づいて生成されるクエリの種類を決定する。地理的関係性を用いたクエリ生成手法については 5 章で述べる。生成されるクエリ式をユーザインタフェース (図 1) の再生画面下部に表示する。また、クエリ式を用いた検索結果の上位 10 件を中央部に表示し、検索結

表1 LDA-U, LDA, TFiDFを用いたトピック抽出の例

TV Program	Scene Selection	LDA-U	LDA	TFiDF
A	7 → 9 → 13	Becky, back, sudden, first, Kagetsutei, seafood, hot spring	Becky, back, sudden, first, Kagetsutei, seafood, book	certificate, incinerator, final destination, reservation
B	12 → 16 → 22	Arima, hot spring, town, seawater, research, together, fluffy	Arima, hot spring, town, scent, seawater, research, center, together, fluffy, hot water, earlier	expressway, main street
C	4 → 3 → 5	birds, Fujimae, food, mud, tidal flat, nutrients, location, snipe, feathers	birds, Fujimae, food, mud, tidal flat, nutrients, location	fish

表2 トピック抽出に使用したテレビ番組

TV program	Name	Time	#Scenes
A	Soda tabi ni iko	45 min	22
B	Buratamori #97	45 min	28
C	Sawayaka shizen hyakei #2	15 min	8

果の上位1件のWebページを右上部で表示する。

#### 4. トピック抽出

本研究では、補足情報を推薦する際にトピックと空間的キーワードを用いてクエリ式を生成する。そのため、該当シーンにおけるクエリ式で使用するトピックの抽出手法について説明する。

##### 4.1 トピック抽出手法

トピック抽出手法には、潜在的ディリクレ配分法 (Latent Dirichlet Allocation, LDA) [17] を使用し、番組内に出現するキーワードに対して重み付けを行う。これにより、シーンごとにトピックとなるキーワードの判定を行うことが可能となる。次に具体的な処理の流れについて説明する。番組におけるカメラの切り替わりをショットと定義し、ショットの意味的な集合をシーンと定義し、各ショット内で発言されたナレーションやセリフを1つの文書として使用する。番組全体でのショットの集合を入力として、MeCab [21] による形態素解析を行い、それにより得られる名詞のみをキーワードとして使用する。また、形態素解析の際に不要語を除去するために数字1文字やひらがな1文字などをStopwordsとして設定した。そして、抽出されたキーワードの出現頻度を用いてTFiDFによるベクトル化を行い、それらを用いてLDAによるトピック分布を求める。

ハイパーパラメータ  $\alpha$ ,  $\beta$  の設定はZhaoら [18] に従い、 $\alpha = 50/T$ ,  $\beta = 0.01$  とした。また、トピック数の設定は渡邊ら [19] に従い、トピック数=5から5刻みで30まで実行し、人手で調べた結果最も適切にトピックが構成されたと判断されたトピック数を使用する。適切なトピック数は番組の長さによって異なるため、45分の番組ではトピック数=15、15分の番組ではトピック数=10とする。シーンごとのトピックを判定するために、該当シーン内に含まれる各ショットごとの各トピックに属する確率の平均値を用いて、シーンごとで上位3件に含まれるトピックを抽出する。そして、上位3件のトピックにおける上位キーワード10件を抽出し、各キーワードがトピックに属する確率の平均値を閾値として、閾値以上のキーワードをトピック

ワードとする。

##### 4.2 視聴動作に基づくトピック抽出

本研究では、ユーザの視聴操作を考慮した補足情報推薦を目的としているため、トピック抽出においても視聴操作を考慮した手法としてLDA-Uを提案する。LDA-UではLDAと同様に番組全体でのショットの集合を入力として処理を行うが、シーンごとのトピックを判定する際に該当シーンと2つ前までのシーン選択履歴の3シーンを用いて各ショットごとの各トピックに属する確率の平均値を算出する。また、平均値の算出を行う際に定数を重みとして付与する。重みの設定としては該当シーンにおける確率には0.5、1シーン前のシーンにおける確率には0.3、2シーン前のシーンにおける確率には0.2を乗算する。これにより、視聴履歴でのトピックも考慮したトピック分布の構成が可能となる。

##### 4.3 各手法におけるトピック抽出結果の比較

実際の番組を用いてLDA・LDA-Uにより抽出されるトピックワードの比較を行った結果を表1に示す。また、比較手法として、従来手法であるTFiDFの結果も示す。TFiDFの計算は以下の式により算出され、番組全体における総シーン数を  $N$  とし、該当するシーンを  $s$ 、該当するシーンに出現するキーワードを  $k$  と定義する。

$$TFiDF = TF(k, s) \times \log \frac{N}{DF(k)} \quad (1)$$

結果として、テレビ番組A, CにおけるLDA-Uを用いたトピック抽出では、1シーン前のシーンでトピックワードとして抽出された”hot spring”や”snipe”などのトピックワードを抽出することが可能である。また、TFiDFにより抽出されたトピックワードとLDA-U, LDAにより抽出されたトピックワードを比較するとTFiDFでは出現頻度のみを用いて計算されるため、該当シーンで出現頻度の高いキーワードのみが抽出されるが、LDA-U, LDAでは該当シーンに出現していないが関連のあるキーワードも抽出することが可能である。以上の結果より、提案手法であるLDA-Uを用いることでユーザの視聴操作を考慮したトピック抽出が可能であるということが確認された。

#### 5. クエリ生成手法

本章では、補足情報を推薦するためにWebページを使用する。よって、補足情報となるWebページを抽出可能なクエリ式を生成する。クエリ式の生成には情報検索システムで使用されているブーリアン検索モデルを使用する。

表 3 空間的・意味的距離を用いたクエリ生成の基準

Geographical Relationship	Disjoint	Equal	Inside	Contain	Disjoint	Equal	Inside	Contain
	Geographical Distance : Near				Geographical Distance : Far			
Semantic Distance : Near	Generate	Generate	Generate	Generate	Generate	Not Generate	Generate	Generate
Semantic Distance : Far	Generate	Generate	Generate	Generate	Not Generate	Not Generate	Not Generate	Not Generate

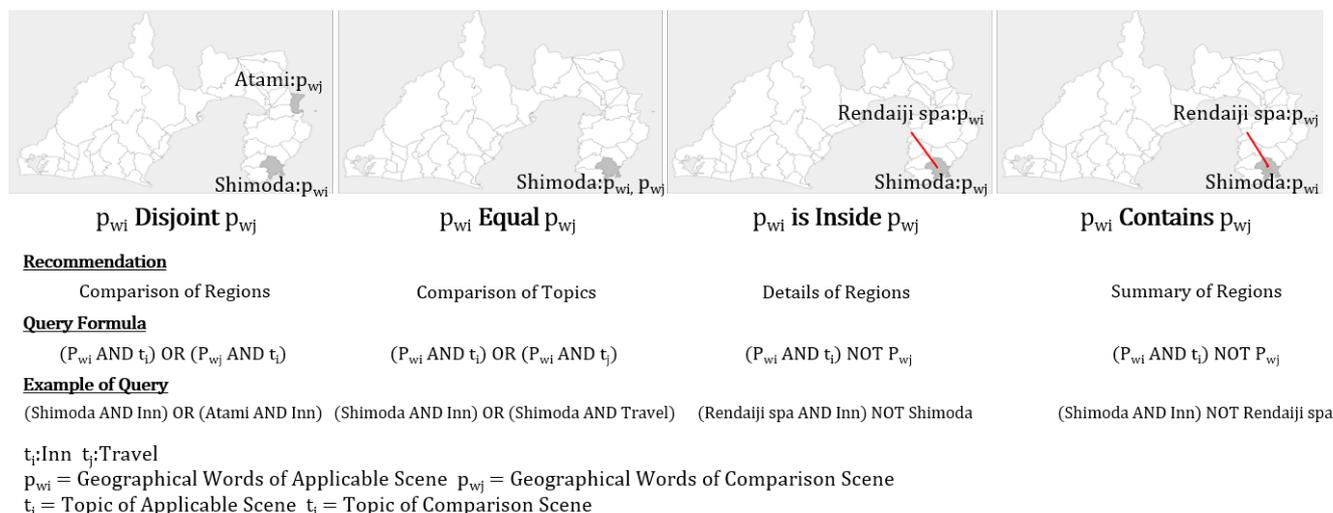


図 4 地理的關係性の例

表 4 地理的關係性に基づくクエリ生成

Geographic Relationship	Recommendation	Query	Example of Query
(1) Disjoint	Comparison of Regions	$(p_{w_i} \text{ AND } t_i) \text{ OR } (p_{w_j} \text{ AND } t_i)$	$(\text{Shimoda AND Inn}) \text{ OR } (\text{Atami AND Inn})$
(2) Equal	Comparison of Topics	$(p_{w_i} \text{ AND } t_i) \text{ OR } (p_{w_i} \text{ AND } t_j)$	$(\text{Shimoda AND Inn}) \text{ OR } (\text{Shimoda AND Travel})$
(3) Inside	Details of Regions	$(p_{w_i} \text{ AND } t_i) \text{ NOT } p_{w_j}$	$(\text{Rendaiji Spa AND Inn}) \text{ NOT } \text{Shimoda}$
(4) Contain	Summary of Regions	$(p_{w_i} \text{ AND } t_i) \text{ NOT } p_{w_j}$	$(\text{Shimoda AND Inn}) \text{ NOT } \text{Rendaiji Spa}$

### 5.1 空間的・意味的距離を用いたクエリ生成の基準

シーン間の地理的關係性とトピックを用いたクエリ生成では、シーン間の内容が明らかに異なる場合や実距離が遠く離れている場合がある。よって、関係の無いシーン間でのクエリ生成を制御するためにクエリ生成の基準を設定する必要があると考えられる。本研究では、空間的キーワードの郵便番号を用いた空間的距離とシーン間の内容の類似度を用いた意味的距離の2つの距離を定義し、これらの距離に基づき、クエリ式を生成するための基準を表3に示している。

#### 5.1.1 空間的距離の算出

空間的距離の算出には、上3桁と下4桁で構成される郵便番号を使用し、実距離の近似値を算出する。これにより、実距離を使用するコストを大幅に削減することが可能である。また、郵便番号を使用することにより、空間的キーワード間の地理的關係性を判定することが可能である。空間的距離の算出には以下の式により算出される。

$$GeoD(s_i, s_j) = |\text{Zipcode}(s_i) - \text{Zipcode}(s_j)| \quad (2)$$

#### 5.1.2 意味的距離の算出

意味的距離を算出するために、シーン間での類似度をユークリッド距離により算出する。これにより、関連の無いシーン間での補足情報の推薦を防止する。ユークリッド距離の算出には

特徴ベクトルとしてシーンごとの  $TFiDF$  値を使用する。また、意味的距離の計算は以下の式により算出される。

$$SemD(s_i, s_j) = \sqrt{\sum_{i=1}^N (s_j - s_i)^2} \quad (3)$$

$$s_i = \{TFiDF(k_1), TFiDF(k_2), \dots, TFiDF(k_j)\}$$

### 5.2 地理的關係性に基づくクエリ生成

本手法では、該当シーンに対して1シーン前のシーン履歴を比較シーンと呼ぶ。また、該当シーンと比較シーンに出現する空間的キーワードを用いて、シーン間での地理的關係性を抽出し、生成されるクエリの種類を判定する。使用される地理的關係性のパターンとして Topological Relations の9種類の空間的関係より、4種類の関係性を使用する。使用する関係性は **Disjoint**, **Equal**, **Inside**, **Contains** の4種類である。以下に地理的關係性によって推薦される情報を説明する。

**Disjoint** 該当シーンに出現する地名と比較シーンに出現する地名が分離しているため、地域の比較情報を推薦する。

**Equal** 該当シーンに出現する地名と比較シーンに出現する地名が一致しているため、トピックの比較情報を推薦する。

**Inside** 該当シーンに出現する地名が比較シーンに出現する地名に対して内在しているため、地域の詳細情報を推薦する。

表 5 SUS Score を用いたユーザインタフェースの評価結果

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	SUS Score	Grade	Adjective Rating
With scene selection	3.7	2.7	3.9	2.8	4.2	1.9	4.1	2.0	4.0	2.9	<b>69.2</b>	C	Okay
Without scene selection	3.2	2.9	3.7	2.0	4.1	2.3	3.3	2.0	3.8	3.1	64.4	D	Poor

表 6 クエリ生成の評価実験で使用するテレビ番組

TV Program	Title	Time Length	Scenes	Percentage	Respondents	Scene Selection
Travel Program	Tabi no hoshi	14m	17	24%	11	2 → 6 → 11 → 13 → 14
News Program	News kyo-ichiniti	15m	29	7%	11	1 → 2 → 6 → 7 → 13 → 19

表 7 三手法を用いて抽出された Web ページの評価

Method	User Rating (Travel)	User Rating (News)
Proposed Method	<b>3.56</b>	<b>3.00</b>
<i>TFiDF</i>	2.38	2.75
<i>TF</i>	1.76	2.36

**Contains** 該当シーンに出現する地名が比較シーンに出現する地名に対して内包しているため、地域の概要情報を推薦する。

それぞれの地理的関係性の例は図 4 で示している。また生成されるクエリ式は表 4 で示している。

## 6. 評価実験

本章では、本研究で使用するユーザインタフェースのユーザビリティに関する評価と補足情報を抽出する際に使用するクエリ式の評価を行う。ユーザインタフェースの評価には、システムのユーザビリティを評価する際に使用される指標である SUS Score を用いる。また、クエリ式の評価には実際の番組を用いてクエリ生成を行い、生成されたクエリ式により得られる Web ページを *TFiDF* と *TF* により得られる Web ページと比較することで評価を行う。

### 6.1 ユーザインタフェースのユーザビリティに関する評価

ユーザインタフェースのユーザビリティに関する評価実験では、プロトタイプシステムのユーザインタフェース (図 1) を使用し、ユーザインタフェースの操作の例は図 2 で示している。上記のユーザインタフェースを用いて、シーン選択を行う場合と行わない場合の 2 つのパターンでどちらの方が有用性が高いかを検証する。また、SUS Score の算出に使用する質問は以下の項目で行う。

- Q1 しばしば利用したいと思う
- Q2 利用するには説明が必要となるほど複雑だと感じた
- Q3 容易に使いこなす事が出来ると思った
- Q4 利用するには専門家のサポートが必要だと感じる
- Q5 コンテンツやナビゲーションは十分に統一感があると感じた
- Q6 一貫性のないところが多々あったと感じた
- Q7 たいていの人は利用方法をすぐに理解すると思う
- Q8 とても操作しづらいと感じた
- Q9 利用できる自信がある
- Q10 利用し始める前に知っておくべきことが多くあると思う

以上の質問に対して、被験者は 5 段階のリッカート尺度で回

答を行い、SUS Score の評価は以下の 5 段階で評価する。A) Excellent: >80.3, B) Good: >68.0-80.3, C) Okay: >68.0, D) Poor: >51.0-68.0, E) Awful: >51.0. 本実験では、シーン選択の有無に関わらず、SUS Score を使用し、各パターンを被験者 9 名ずつ、総被験者数 18 名によって評価する。各質問の評価値や各パターンの SUS Score の平均値を表 5 で示す。表 5 の結果より、シーン選択を行う場合の評価は *Okay*、行わない場合の評価は *Poor* とであることが確認された。また、SUS Score の平均値は比較手法であるシーン選択を行わない場合に比べて、提案手法であるシーン選択を行う場合の方が評価値が高いため、提案手法の方が有用性が高いことを確認された。

### 6.2 クエリ式の評価

クエリ式の評価実験では、提案手法により生成されるクエリ式と比較手法である *TFiDF* と *TF* により抽出されるキーワードを用いて比較を行い、クエリ式の有効性の検証を行う。データセットには、表 6 に示している旅番組とニュース番組の 2 つのテレビ番組を使用し、各手法により生成されるクエリを用いて、上位 10 件の Web ページを抽出する。

旅番組では、提案手法により (喜多方 AND ラーメン) NOT 会津というクエリ式が生成され、比較手法である *TFiDF* では「喜多方」、*TF* では「ラーメン」というクエリキーワードが生成された。ニュース番組では、提案手法により (中国 AND 護衛艦) OR (アメリカ AND 護衛艦) というクエリ式が生成され、比較手法である *TFiDF* では「護衛艦」、*TF* では「防衛省」というクエリキーワードが生成された。

以上のクエリ式とクエリキーワードを用いて、上位 10 件の Web ページを抽出し、抽出された Web ページを被験者 5 名ずつにより評価を行った。評価には 5 段階のリッカート尺度を使用し、得られた評価値の平均を表 7 に示している。表 7 の結果より、提案手法により得られた Web ページは比較手法と比べて、補足情報としての有効であることが確認された。

## 7. おわりに

本研究では、視聴操作を考慮した地理的関係性に基づくビデオ視聴支援システムの提案を行った。提案手法としてユーザの視聴操作に基づく抽出されるトピックワードと空間的キーワード間の地理的関係性を用いたクエリ生成手法の提案を行った。また、他分野においても空間的キーワードのようにキーワード同士の関係性を抽出可能なキーワードがあれば、本研究の提案手法は応用可能であると考えられる。

## 謝 辞

本研究の一部は、JSPS 科研費 JP19H04118, および 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] Yuanyuan Wang, Daisuke Kitayama, Yukiko Kawai, and Kazutoshi Sumiya, "Automatic Street View System Synchronized with TV Program using Geographical Metadata from Closed Captions", in Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces (AVI 2014), pp. 383-384, May 27-30, 2014.
- [2] Yuanyuan Wang, Daisuke Kitayama, Yukiko Kawai, Kazutoshi Sumiya, and Yoshiharu Ishikawa, "An Automatic Video Reinforcing System for TV Programs using Semantic Metadata from Closed Captions", in International Journal of Multimedia Data Engineering and Management, Volume 7, Issue 1, pp. 1-21, January-March 2016.
- [3] Yuanyuan Wang, Maho Nishizawa, Yukiko Kawai, and Kazutoshi Sumiya, "Location-based Image Viewing System Synchronized with Video Clips", in Proceedings of the 13th International Conference on Location-Based Services, LBS 2016, pp. 233-238, November 14-16, 2016.
- [4] Honoka Kakimoto, Toshinori Hayashi, Yuanyuan Wang, Yukiko Kawai, and Kazutoshi Sumiya, "Query Keyword Extraction from Video Caption Data based on Spatio-Temporal Features", in Proceedings of the International MultiConference of Engineers and Computer Scientists 2018, Volume I, IMECS 2018, pp. 405-408, March 14-16, 2018.
- [5] サイモン クリピングデル, 高橋 正樹, 山内 結子, 苗村 昌秀: 視聴者の興味内容推定における EPG や字幕中の単語への重みの付与。映像情報メディア学会年次大会, 2015 年。
- [6] 三浦 菊佳, 山田 一郎, 宮崎 太郎, 加藤 直人, 田中 英輝: 単語間の意味的關係を用いたテレビ番組マップ生成。情報処理学会第 76 回全国大会, 5C-4, 2014 年。
- [7] Lorenzo Baraldi, Costantino Grana, and Rita Cucchiara, "Scene Segmentation using Temporal Clustering for Accessing and Re-using Broadcast Video", in Proceedings of the 2015 IEEE International Conference on Multimedia and Expo (ICME 2015), 6 pages, June 2015.
- [8] Panagiotis Sidiropoulos, Vasileios Mezaris, Ioannis Kompatsiaris, Hugo Meinedo, Miguel Bugalho, and Isabel Trancoso, "Temporal Video Segmentation to Scenes Using High-Level Audiovisual Features", in IEEE Transactions on Circuits and Systems for Video Technology, Volume 21, Issue 8, pp. 1163-1177, 2011.
- [9] Z. Rasheed and M. Shah, "Detection and Representation of Scenes in Videos", in IEEE Transactions on Multimedia, Volume 7, Issue 6, pp. 1097-1105, December 2005.
- [10] Cailiang Liu, Dong Wang, Jun Zhu, and Bo Zhang, "Learning a Contextual Multi-Thread Model for Movie/TV Scene Segmentation", in IEEE Transactions on Multimedia, Volume 15, Issue 4, pp. 884-897, June 2013.
- [11] Shayok Hakraborty, Omesh Tickoo, and Ravi Iyer, "Adaptive Keyframe Selection for Video Summarization", in Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision (WACV 2015), pp. 702-709, 2015.
- [12] Ba Tu Truong, and Svetha Venkatesh, "Video abstraction: A systematic review and classification", in Proceeding of ACM Transactions on Multimedia Computing, Communications, and Applications, February 2007.
- [13] Aidean Sharghi, Boqing Gong, Mubarak Shah, "Query-Focused Extractive Video Summarization", European Conference on Computer Vision (ECCV, 2016) pp 3-19
- [14] 北山 大輔, 宮本 節子, 角谷 和俊, "ユーザ操作履歴と地名出現パターンに基づくオンライン地図の動的カスタマイズ方式", 情報処理学会論文誌データベース (TOD), 3, 4, pp. 65-81, 2010.
- [15] K. Kobayashi, D. Kitayama, and K. Sumiya, "Cinematic Street: Automatic Street View Walk-through System Using Characteristics of Modified Maps", Proc. of the 10th International Symposium on Web & Geographical Information Systems (W2GIS 2011), pp. 142-158, 2011.
- [16] 倉田 陽平, "地理情報科学における位相的空間関係の研究と今後の課題", GIS-理論と応用, 18, 2, pp. 149-159, 2010.
- [17] David M. Blei, Andrew Y. Ng, and Michael I. Jordan, "Latent Dirichlet Allocation", Journal of Machine Learning Research 3, pp. 993-1022, 2003.
- [18] Wayne Xin Zhao, Jing Jiang, Jianshu Weng, Jing He, and Ee Peng Lim, "Comparing Twitter and Traditional Media using Topic Models", Proc. 33rd European Conference on Advances in Information Retrieval, pp. 338-349, 2011.
- [19] 渡邊 恵太, 加藤 昇平, "トピックモデルと協調フィルタリングに基づくユーザ興味を反映した情報推薦システム", 2014 年度人工知能学会全国大会, 2M3-4, 2014.
- [20] NHK Hybridcast:  
<https://www.nhk.or.jp/hybridcast/online/>.
- [21] MeCab:  
<https://taku910.github.io/mecab/>.
- [22] Microsoft Video Indexer:  
<https://vi.microsoft.com/ja-jp/>.
- [23] Topological Relationships:  
[http://www.gitta.info/SpatialQueries/en/html/TopoBasedOps\\_learningObject1.html](http://www.gitta.info/SpatialQueries/en/html/TopoBasedOps_learningObject1.html)