

ユーザの視聴操作を考慮した地理的関係と重要語判定に基づく ビデオ映像の補足情報提示方式

橋本 樹[†] 王 元元^{††} 河合由起子^{†††,††††} 角谷 和俊[†]

[†] 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

^{††} 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

^{†††} 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

^{††††} 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: [†]{ens47564,sumiya}@kwansei.ac.jp, ^{††}y.wang@yamaguchi-u.ac.jp, ^{†††}kawai@cc.kyoto-su.ac.jp

あらまし 近年、インターネットテレビの普及により、テレビ番組を対象として関連情報を推薦する研究が活発に行われている。NHK の Hybridcast ではテレビ番組放送中に同画面上で関連情報を推薦するというサービスを行っている。また、最近ではユーザの検索履歴や視聴履歴を用いて個人化された情報推薦を行うことが重要視されている。しかし、映像やテレビ番組などにおいて視聴者の視聴動作や嗜好などを考慮した情報推薦は現在行われていない。そのため、視聴者の視聴動作や嗜好などを考慮した情報推薦を映像やテレビ番組などで行うことが今後必要であると思われる。そのため、本研究では視聴者の視聴動作を考慮した情報推薦を行うことを目的としている。

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

近年、インターネットテレビの普及により、テレビ番組を対象として関連情報を推薦する研究が活発に行われている。NHK の Hybridcast [1] ではテレビ番組放送中に同画面上で関連情報を推薦するというサービスを行っている。また、最近ではユーザの検索履歴や視聴履歴を用いて個人化された情報推薦を行うことが重要視されている。しかし、映像やテレビ番組などにおいて視聴者の視聴動作や嗜好などを考慮した情報推薦は行われていないため、視聴者の視聴動作や嗜好などを考慮した情報推薦を映像やテレビ番組などで行うことが必要であると思われる。そのため、本研究ではテレビ番組の出演者の声やナレーションに付与されるクローズドキャプションからシーンごとに得られる重要語と地理的な関係性を抽出する。また、地理的な関係性の取得を行うために、地名や地域など地理的な情報を持つキーワードの抽出を行う。

本研究では、地理的な情報を持つキーワードを空間的キーワードと定義し、論文内で使用する。ユーザインタフェースには動画視聴前に表示するユーザインタフェースと動画視聴中に表示するユーザインタフェースの 2 つを使用する。動画視聴前に表示するユーザインタフェースは視聴者が興味のあるトピックを選択可能な仕様になっている。動画視聴中に表示するユーザインタフェースは、従来のコントロールバーを用いて動画を早送りや巻き戻しを行うユーザインタフェースではなく、視聴者がシーン選択可能なユーザインタフェースの提案を行う。それら 2 つのユーザインタフェースを用いて視聴者の視聴動作を取得することで、視聴者の視聴動作に合わせて重要語と空間的キーワードの地理的な関係性を変化させる。視聴者のシーン選択によって情報推薦の内容が変化するため、視聴動作を考慮した

情報推薦が可能になる。上記で提案する手法の有効性を実際のテレビ番組を使用して検証する。本研究では「そうだ旅に行こう【やっぱり夏の伊豆は最高！ポンコツ極上温泉宿の旅】」という 45 分の旅番組を使用する。評価実験では上記の旅番組に追加で 2 つの旅番組を使用している。

2. 関連研究

近年、ビデオ映像のクロー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われている。王ら [2] は番組内に出現する地理的な情報を持つキーワードをクエリキーワードとして、Web により写真やルート推薦を行うシステムの提案を行った。同じく西澤ら [3] もテレビ番組のクローズドキャプションを用いてクエリキーワードを生成し、Wikipedia の階層構造により、関連するキーワードの抽出手法を提案した。クリピンデルら [4] 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、手法として本研究と同じく *TFiDF* を使用して重要度を算出した。三浦ら [5] 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。また、クローズドキャプションを用いた重要語判定や情報推薦は、他のデータを対象にした研究においても応用されている。

田原ら [6] や長谷川ら [7] は SNS の投稿からテキスト情報を抽出し、抽出されたテキスト情報を用いて特徴語の判定を行い、アウトプットとして観光情報検索や地域ユーザの検索に応用する手法を提案した。原田ら [8] は Twitter を用いてテキストマイニングを行い、語彙の概念を可視化し、関係性を目視で確認できるシステムを提案した。SNS におけるテキストマイ



図1 視聴前のトピック選択画面

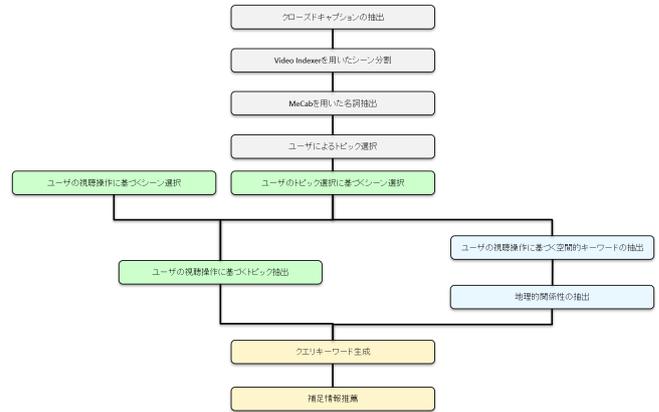


図3 システム概要図



図2 視聴中のユーザインタフェース

とで出現回数の多い空間的キーワードを表示している。トピックの抽出方法としては $TFiDF$ を使用し、番組内で重要であるキーワードを表示している。

図2は視聴中のユーザインタフェースであり、4つのパートに分かれたユーザインタフェースを提案する。画面左上部分では動画の再生画面を表示する。画面下部分では全シーンのサムネイルを表示し、クリックすることにより該当シーンにスキップすることが可能である。画面右上部分では視聴者のシーン選択の履歴から得られた視聴動作と選択シーンの内容に基づく補足情報を表示し、該当シーンの内容に即した情報を同画面上で閲覧することが可能となる。画面右中部分では空間的キーワードを表示し、番組視聴中にどの地名に関する補足情報を推薦しているのかを表示することにより、視聴者自身が検索を行うことも可能となる。上記のユーザインタフェースにより、視聴者の選択シーンの履歴を抽出することが可能となる。また、本研究では視聴者のトピック選択と選択シーン履歴を視聴者の視聴動作と定義し、視聴動作に基づいてシーンごとに抽出される重要語や空間的キーワードの地理的関係性が変化する。よって、視聴者の視聴意図や選択シーン順序の関係性を用いた情報推薦を行うことが可能になるため、本研究の目的である視聴者の視聴動作を考慮した情報推薦を行うことが可能になると考えられる。

4. 提案手法

本研究では図3のシステム概要図に記述しているように、視聴者の視聴動作に基づく重要語判定と空間的キーワードを用いた地理的関係性抽出により、視聴者の視聴動作に基づいてクエリキーワードを生成する。生成されたクエリキーワードを用いて視聴者の視聴動作を考慮した補足情報推薦を行う。

4.1 シーン分割

テレビ番組のクローズドキャプションを抽出し、抽出されたクローズドキャプションを用いて MeCab [12] により形態素解析を行い、番組内に出現する名詞のみを抽出する。次に Microsoft Video Indexer [13] を用いてシーン分割を行う。本研究で使用する45分の旅番組では22個のシーンに分割された。シーン分割が正常に行われたため、本研究ではこのツールを使用した。

3. システム概要

本研究では、図1のトピック選択画面と図2のシーン選択画面を使用する。図1は番組視聴前のユーザインタフェースであり、シーンごとの空間的キーワードとトピックを表示し、視聴者の興味があるシーンの判定を行う。空間的キーワードの抽出手法としては TF による計算を使用しているため、シーンご

ningとクローズドキャプションにおけるテキストマイニングは類似しているため、本研究でも参考としている。また、本研究では視聴者の視聴動作を考慮した情報推薦を目的としているため、個人化された情報推薦を行っている研究を参考文献に挙げている。近年では Web の閲覧履歴や EC サイトの購買履歴などを用いて、ユーザの嗜好や行動傾向を抽出し、個人化された情報推薦を行う研究が盛んに行われている。酒井ら [9] と高間ら [10] は SNS やレビューの情報を用いて個人化された情報推薦を行う手法を提案している。井口ら [11] はテレビ番組のクローズドキャプションを用いて視聴者の視聴傾向を判定し、個人化された番組表の生成を行っている。これらの研究により、個人化された情報推薦がさまざまな領域で重要であると考えられる。よって、本研究ではユーザの視聴動作を用いて、個人化された情報推薦を可能とする手法を提案する。

通常の視聴履歴



視聴者の視聴動作から得られる視聴履歴



図4 視聴操作の例



トピック
空間的キーワード
宿: t , 温泉: u
下田: a , 熱海: b , 蓮台寺温泉: c

図5 Topological Relationships で使用する4つの関係性

4.2 重要語判定手法

番組内容に即した情報推薦を行うために、時系列で変化するトピックを判定し、判定されたトピックの中から重要となるキーワードを抽出し、それらを用いて情報推薦を行う必要がある。そのため、番組内のシーンごとにトピックを判定を行う。トピック判定の方法として各シーンごとでの重要語を判定し、それらをトピックに関係のあるキーワードであると仮定し、情報推薦を行うためのクエリキーワードとして抽出する。クエリキーワードの生成方法として、重要語とシーンごとで抽出した空間的キーワードを用いて、検索を行い、検索結果数が一番多い組み合わせを関係性のある重要語と空間的キーワードの組み合わせであると定義する。

テレビ番組から抽出したクロードキャプションを用いて、MeCab [12] による形態素解析を行い名詞のみを抽出し、式 (1) シーンごとの名詞を用いて $TFiDF$ を行うことで、重要語の判定を行う。 TF はシーン内の出現キーワード総数におけるシーン内のキーワードごとの出現頻度、 iDF は全シーンにおいて該当シーン内に該当キーワードが何回出現したかを表している。 \log を追加して計算を行うことにより、ビデオ映像において出現回数の多いキーワードが重要であるという条件を満たす重みを付与する。

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

$k = \text{Keyword}, s = \text{Scene}, N = \text{All Scenes}$

4.3 視聴動作に基づく重要語判定

図4の下部は視聴者がシーン1, 3, 4, 7を選択して視聴している様子を表している。視聴者の選択したシーンを用いて局所

的な重要語判定を行い、視聴者の視聴傾向に合わせた情報推薦を行う。視聴者の2回分のシーン選択履歴、視聴シーンを用いて $TFiDF$ により計算を行う。の場合、シーン7を視聴中のためシーン3, 4, 7を用いて重要語判定を行う。それにより、従来の重要語判定手法として用いられる全文書を用いた $TFiDF$ では得られない視聴者の視聴動作を考慮した重要語が得られると考えられる。たとえば、レストランの紹介など食に関するシーンを多く選択している視聴者には、食に関するの重要語、温泉巡りなど観光に関するのシーンを多く選択している視聴者には観光に関する重要語を推薦することが可能となる。

4.4 出現間隔を考慮した重み付け

上記の提案手法で得られたクエリキーワードでは空間的キーワードの範囲によって検索結果数の大小が大きく分かれてしまう問題が発生する。しかし、実際に使用したクロードキャプションでは空間的キーワードと重要語の出現間隔が近いほど関係があるということが確認されたため、出現間隔の近い重要語と空間的キーワードに重みを追加することにより、関係度が向上するという仮説を立てることが可能である。仮説により、得られたクエリキーワード生成の際に得られた検索結果数に出現間隔を考慮した重みを追加し、再計算を行う。出現間隔を考慮した重み付けの計算方法は得られた重要語と空間的キーワード群のショットが出現した時間を抽出し、検索結果数を出現ショット間の秒数で割り、同じショットに出現するキーワードの数値を残すために1を足す。これにより、検索結果数と出現間隔を考慮した関係性の抽出が可能である。

4.5 地理的關係性抽出

本研究では、視聴者の視聴履歴から得られた視聴順序に基づ

表 1 地理的関係性を用いたクエリキーワード生成

地理的関係性	情報推薦の種類	クエリ	クエリキーワード例
(1) Disjoint	地域の比較情報	(t_a, t_b)	(宿下田, 宿熱海)
(2) Equal	トピックの比較情報	(t_a, u_a)	(宿下田, 温泉下田)
(3) Inside	地域の詳細情報	(t_c, u_c)	(宿運台寺温泉, 温泉運台寺温泉)
(4) Contains	地域の概要情報	(t_a, u_a)	(宿下田, 温泉下田)

き、シーン間での地理的関係性を抽出する。関係性のパターンとして Topological Relationships [14] の 4 つの関係性を使用する。使用する関係性は Disjoint, Equal, Inside, Contains の 4 つである。これらの関係性によって、推薦する情報を変化させ、視聴者のシーン選択に合わせて情報推薦を行う。情報推薦で使用する 2 つの空間的キーワードと 2 つの重要語には、視聴中のシーンで生成されたクエリキーワードに用いられる空間的キーワードと重要語、視聴順序において前のシーンで生成されたクエリキーワードに使用されている空間的キーワードと重要語を使用する。それぞれの処理については図 5 と表 1 で示す。例には前述のテレビ番組「そうだ旅に行こう【やっぱり夏の伊豆は最高！ボンコツ極上温泉宿の旅】」を使用する。図 5 で示している関係性は視聴中のシーンと 1 つ前のシーン履歴で得られるクエリキーワードの重要語と空間的キーワードを使用する。クエリキーワードで使用される空間的キーワードの前後関係により、トピックの比較情報、地域の比較情報、地理的な詳細情報、地理的な概要情報の 4 種類の情報推薦から選択する。

5. 評価実験

「そうだ旅に行こう【やっぱり夏の伊豆は最高！ボンコツ極上温泉宿の旅】」から抽出したクローズドキャプションを用いて、本研究の提案手法の有効性を検証する。4 章で述べた提案手法を用いて、最終的に推薦される補足情報を抽出し、結果から有効性を確認する。

5.1 重要語判定の検証

本研究で使用する旅番組のデータを Video Indexer [13] で解析し、22 個のシーンに分割を行った。抽出されたシーン群に対してユーザインタフェースにより、シーン選択が行われる。シーン選択により、視聴者の視聴動作として 12 回のシーン選択を取得される。これは観光に関係するシーンに興味がある視聴者が選択を行った。抽出されたシーン群を用いて、提案手法で説明した局所的な $TFiDF$ を行い、シーンごとの重要語を抽出する。抽出された重要語群は表 3 に示している、全体的 $TFiDF$ と局所的 $TFiDF$ の差異は観光関係のキーワードである「温泉」、「宿」、「花月亭」などの全体的に頻出しているキーワードの重要度が上がり、重要語として抽出される。

5.2 クエリキーワード生成方式の検証

本研究で使用しているクローズドキャプションを用いて、視聴者がシーン 13 を視聴中の場合で計算を行う。前までのキーワードを用いて局所的 $TFiDF$ を行うため、シーン 7、シーン 8、シーン 13 の 3 つのシーンを用いて計算を行う。これにより、直前の視聴者の視聴傾向を考慮することが可能である。視聴中のシーンまでの視聴履歴全てを使用してしまうと、番組視聴中に視聴傾向が変化した場合に視聴中の視聴傾向を考慮すること

表 2 全体的な $TFiDF$ を用いた重要語群

シーン	全体的な $TFiDF$ により得られる重要語
シーン 1	極上
シーン 2	極上
シーン 4	ヒゲ
シーン 5	老舗, 癒し, 贅沢, 一泊, 二日
シーン 7	天気, 安心, 顔, 席, 出発, 方面, 所有, 長旅
シーン 8	熱海, うなぎ
シーン 13	最終目的地, 証, 焼却炉, 予約
シーン 14	花月亭
シーン 18	新鮮, 伊勢海老, 真鯛, マグロ, アジ, ボタン海老, 季節, 魚介類
シーン 19	刺身, 海鮮丼, しゃぶしゃぶ, 意識, 魅力, イメージ, 我々
シーン 21	真ん中
シーン 22	朝

表 3 局所的な $TFiDF$ を用いた重要語群

シーン	局所的な $TFiDF$ により得られる重要語
シーン 1	温泉, 極上
シーン 2	極上, 開湯, 1300 年, 物語
シーン 4	ベッキー
シーン 5	老舗, 癒し, 贅沢, 家族, 一泊, 二日
シーン 7	車, 天気, 安心, 顔, 席, 出発, 方面, 所有, 長旅
シーン 8	順調
シーン 13	宿, 花月亭, 奥座敷, 最終目的地, 証, 焼却炉, 予約
シーン 14	散歩
シーン 18	新鮮, 伊勢海老, 真鯛, マグロ, アジ, ボタン海老, 季節, 魚介類
シーン 19	刺身, 海鮮丼, しゃぶしゃぶ, 意識, 魅力, イメージ, 我々
シーン 21	真ん中
シーン 22	朝

表 4 出現間隔を考慮した重みづけ

空間的キーワード	検索結果数	出現間隔 (秒)+1	検索結果数/出現間隔 (秒)+1
白浜大浜海水浴場	451,000	192	2,349
伊豆半島	4,990,000	192	25,990
下田	8,150,000	192	42,448
白浜	9,520,000	174	54,713
下田	8,150,000	57	142,982
運台寺温泉	163,000	57	2,860
花月亭	266,000	1	266,000

が出来ないため、3 つのシーンのみを用いて重要語判定を行う。シーン 13 のクローズドキャプションが図 6 であり、赤字が得られた重要語群、緑字が得られた空間的キーワードである。これらの重要語群と空間的キーワード群を総当たりで組み合わせ検索を行った結果、「白浜 予約」の検索数が 9,520,000 件で最大であった。しかし、この 2 つのキーワードは番組内で関係していない。そのため、出現間隔を考慮した重みづけを行う必要がある。

5.3 出現間隔を考慮した重み付けの検証

前節で得られたクエリキーワードに用いられる「予約」というキーワードと関係性があるキーワードはシーン内の内容から「花月亭」というキーワードであると判定されると考えられる。しかし、「花月亭 予約」の検索結果数は 266,000 件である。そのため、提案手法を用いて重み付けを行う。計算結果は表 4 であり、「花月亭 予約」の数値が最大値となっている。これにより、重要語と空間的キーワードが関係しているクエリキーワードを生成することが可能となる。

6. おわりに

本研究では、テレビ番組のクローズドキャプションを使用し、

02339	白浜大浜海水浴場スカイプルの海と美しい白砂が全国的に知られた伊豆半島屈指の海水浴場、多くの海水浴客やサーファーが訪れる下田で人気のスポット。
02357	この気温がいいな写真ならば良いでしょう、ちょっといいなあ。白浜、バックショットをいただきたい。
02408	ちよっと塵って。
02409	うわあ。
02411	いいね撮影会みたいになってるよ。
02415	自然にしてて、これもうジャケットの表紙だもんね。
02424	ベッキーねえ、これね、これやらないほうがいいな。
02428	そうそうそうそう多い。
02433	それでもきつと今ちょっととシャツこしたほうがいいね
02445	次、おじさんたちの場合。
02450	これいい。
02454	最高。
02456	このシルエットの人たち、なにかやりそうだもんね。
02502	じゃあちよっと海ばちやばちやってやつ、大好きパターンでやっつく
02523	早いな、波があるから
02529	男が二人とも彼女のことが好きなんだね
02538	行きます。
02540	行きましょうベッキーの言う取れ高は、これでだいぶあったかな。
02554	今、内容全然なかったですよ、それでは最終目的地、下田の奥座敷、蓮台寺温泉の宿へ向かいます。
02610	温泉の証だね。
02613	この辺の人たちは自然に温泉に浸かっているということだね。
02624	分らない。
02627	これ凍却病があるね、なんか燃やしてるね
02641	人が何かを燃やしてる
02650	これだ、急に来たね、花月亭だ。こんにちは、予約したベッキーですけど、よろしくおねがいます。

図 6 シーン 13 のクローズドキャプション

視聴者の視聴動作に基づく情報推薦を行うことを目的としている。評価実験の結果により、生成すべきクエリキーワードを生成することが可能であると考えられる。また、本研究では地理的情報を持つ空間的キーワードを利用して得られる地理的關係性と視聴者のシーン選択により得られる視聴履歴の順序により、推薦する範囲を変化させるという手法を提案したが、他分野において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、Topological Relationships [14] のように関係ごとに情報推薦を変化させ、他分野の映像においても応用することが可能になると考えられる。そのため、今後は旅番組以外の映像でも本研究で提案した手法を用いて実験を行いたいと考えている。

- SIG-AM-14-05, 2014 年.
- [10] 高間 康史, 清水 涼人: レビュー閲覧履歴からの価値観モデリングを用いた情報推薦システム. 人工知能学会インタラクティブ情報アクセスと可視化マイニング研究会 (第 13 回), SIG-AM-13-02, 2013 年.
- [11] 井口 晃一, 土方 嘉徳, 西田 正吾: ユーザプロファイルの仮想的個別化によるテレビ番組推薦. 人工知能学会論文誌, Vol. 30, No. 1, pp. 71-83, 2015 年.
- [12] MeCab:
<https://taku910.github.io/mecab/>.
- [13] Microsoft Video Indexer:
<https://vi.microsoft.com/ja-jp/>.
- [14] Topological Relationships:
http://www.gitta.info/SpatialQueries/en/html/TopoBasedOps_learningObject1.html

文 献

- [1] NHK Hybridcast:
<https://www.nhk.or.jp/hybridcast/online/>.
- [2] Yuanyuan Wang, Daisuke Kitayama, Yukiko Kawai, Kazutoshi Sumiya, Yoshiharu Ishikawa: An Automatic Video Reinforcing System for TV Programs using Semantic Metadata from Closed Captions. International Journal of Multimedia Data Engineering and Management, Volume 7, Issue 1, January-March 2016.
- [3] 西澤 真帆, 王 元元, 河合 由起子, 角谷 和俊: 映像に付帯する地域情報を用いた投稿型写真共有サイトにおける画像抽出方式. DEIM Forum 2017, H7-4, 2017 年.
- [4] サイモン クリッピングデル, 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.
- [5] 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.
- [6] 田原 琢士, 馬 強: Twitter から有益な日常情報を発見するための特徴語による地域ユーザの検索. DEIM Forum 2014, B3-3, 2014 年.
- [7] 長谷川 馨亮, 馬 強, 吉川 正俊: Twitter からの地域特徴語辞書の構築とその観光情報検索への応用. DEIM Forum 2014, B3-4, 2014 年.
- [8] 原田 真喜子, 渡邊 英徳: 特徴語抽出と感情メタデータ付与によるウェブ上の語彙の概念の視覚化. 映像情報メディア学会誌, Vol. 68, No. 2, pp. J78-J86, 2014 年.
- [9] 酒井 勇人, 熊野 雅仁, 木村 昌弘: Flickr データに基づいたインタラクティブ観光スポット推薦システム. 人工知能学会インタラクティブ情報アクセスと可視化マイニング研究会 (第 14 回),