

修士論文

直観的な観光計画を可能にする 観光動画要約システムの提案と評価

金谷 勇輝

2019 年 3 月 15 日

奈良先端科学技術大学院大学
情報科学研究科

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学) 授与の要件として提出した修士論文である。

金谷 勇輝

審査委員：

安本 慶一 教授 (主指導教員)

中村 哲 教授 (副指導教員)

荒川 豊 准教授 (副指導教員)

諏訪 博彦 助教 (副指導教員)

直観的な観光計画を可能にする 観光動画要約システムの提案と評価*

金谷 勇輝

内容梗概

効率よく観光を行うために、事前の観光計画は重要である。しかし、検索や計画に関わるコスト（労力や時間など）を考慮した場合、観光計画が億劫になることがある。観光計画を容易にするために、ユーザの要望に応じた観光ルートを自動で作成する研究は多数存在するが、そのルートの表示方法は地図やテキストに限られており、直観的ではない。直観的かつ素早くルートを理解するためには、動画による表示が好ましい。そこで本研究では、観光計画のために、任意の観光ルートに沿った動画を自動で作成する観光動画要約システムを提案する。システムは、ユーザの情報に従って観光ルートを作成し、そのルートに沿った動画を作成する。動画は、観光経路を移動する移動動画と、観光スポットの写真を組み合わせたスライドショー動画によって構成される。ユーザは、観光計画が終了するまで観光スポットや移動経路を細かく変更しつつ、動画で繰り返し確認しながら、観光計画を立てることができる。動画を何度も確認する場合、動画が長時間になると、計画効率やユーザの視聴意欲が低下してしまう。短時間かつ観光計画に必要な情報を十分に得られる動画の作成が本システムにおける課題の一つである。

この課題解決のために、長時間になりやすい移動動画の要約手法を考案し、実装・評価を行なう。本手法は、移動動画の同じ景観が続く場面はユーザにとって価値が低い（不要）と考え、カラーヒストグラムを用いて不要場면을省略し、さらに早送り再生を併用することで要約を実現する。要約手法を評価するために、被験者 20 名が手動で動画要約を行なった場合と本手法による要約の結果を比較

*奈良先端科学技術大学院大学 情報科学研究科 修士論文, 2019 年 3 月 15 日.

したところ、F 値 62.22%で必要部分を抽出できた。また、観光計画に必要な情報を得るために、観光動画の再生速度と理解度の関係を調査したところ、4 倍速から 8 倍速が適切であるという知見を得た。これらの結果を元に、自動要約を行い、再生速度を 4 倍速にして動画を作成したところ、6 分 58 秒の動画が 41 秒に短縮された。またその動画を同じ被験者 20 名に視聴してもらったところ、70%の被験者が観光計画に有用であると回答したことから、本手法による動画作成は、観光計画に有用であることが示唆された。

キーワード

観光計画，ルート推薦，動画要約，キュレーション，再生速度

Tour Video Summarization System toward Intuitive Sightseeing Planning*

Yuki Kanaya

Abstract

Sightseeing planning is important for tourists so that they can go sightseeing efficiently. Meanwhile, people sometimes find it hard to plan the sightseeing tour because it is a time consuming task. There have been many studies regarding sightseeing plan creation, with the goal of automatically creating a plan adapted to the user's preferences.

However, it is difficult to intuitively understand the recommended plans because these plans are typically represented by texts or on maps. If the recommended plans are shown with means of video, users can easily and quickly understand them. Videos have richer visual information than simple text and maps, and so many tourists watch videos about the sightseeing spots they want to visit in order to plan their sightseeing tours.

In this thesis, we propose a video summarization system that creates a summarized video along any sightseeing route. This system creates a tour route according to user information and a summarized video showing this route. A summarized video consists of route videos showing the movement between sightseeing spots, and slideshows of photos of the sightseeing spot. Users can intuitively understand the recommended plan by watching the summarized video and can then make a desirable plan quickly. Users may check the summarized video

*Master's Thesis, Graduate School of Information Science, Nara Institute of Science and Technology, March 15, 2019.

repeatedly because they might make minor adjustments to their routes and/or their desired sightseeing spots before settling on a route. Users will then be bored if the summarized video is too long, while the video will not be informative if it is played back too fast. One of the challenges for realizing the system is creating short and informative video.

In order to solve this challenges, we proposed a video summarized method to summarize route videos. This method identifies the scene where the same view continues (unnecessary scenes) from a route video by using color histogram, and cuts or fast-forwards these scenes to make short summarized videos. To evaluate the effectiveness of our video summarized method, we conducted experiments with 20 participants using a sightseeing video taken in Kyoto. When comparing the video summarized with our method and one that was summarized manually (by voting for necessary/unnecessary scenes), our method identified the important scenes with an F-measure of 62.22%. In addition, we investigated the appropriate playback speed of sightseeing videos. Using a questionnaire, we found that playback speed between $\times 4$ and $\times 8$ was the most effective for viewers to understand the sightseeing information for tour route planning. Moreover, 70% of the participants answered that the summarized video automatically generated by our method was effective for planning a tour.

Keywords:

sightseeing planning, route recommendation, video summarization, curation, playback time

目次

1. 序論	1
2. 関連研究	4
2.1 観光経路の推薦に関する研究	4
2.2 コンテンツキュレーションに関する研究	5
2.3 動画の要約手法に関する研究	5
3. 観光動画要約システムの概要	7
3.1 ユースケース	7
3.2 システム構成	8
3.2.1 ユーザ情報収集	9
3.2.2 観光ルート作成	9
3.2.3 CGM 収集	9
3.2.4 要約動画作成	10
3.2.5 動画視聴	11
3.3 本研究で扱う研究課題	13
4. 景観の変化に基づく動画要約手法	15
4.1 変化点抽出アルゴリズム	15
4.1.1 ショットの分割	16
4.1.2 隣接ショットのヒストグラム相関係数	17
4.1.3 変化点の決定	17
4.1.4 動画作成	17
4.2 要約手法の評価	18
4.2.1 実験概要	19
4.2.2 実験結果	20
4.2.3 考察	23
5. 要約動画のユーザスタディ	26

5.1	動画の再生速度による理解度の調査	26
5.1.1	実験概要	26
5.1.2	実験結果	26
5.2	要約動画の有用性評価	27
5.2.1	実験概要	28
5.2.2	実験結果	28
6.	結論	30
	謝辞	32
	参考文献	33
	研究業績	36

図 目 次

1	システム利用時の概要図	7
2	観光動画要約システムの概要図	8
3	要約動画作成の概要	11
4	要約動画視聴時の画面遷移	12
5	観光動画要約手法の簡略図 (フレームシフト δ が 3 の場合)	14
6	動画撮影者の移動経路	18
7	動画撮影時のカメラの配置	19
8	正解データ (70%agree) と提案手法の比較	20
9	2 値分類の結果	22
10	提案手法による抽出場面の一部	24
11	動画 No.2 の場面の一部	25
12	動画中の情報の理解度と再生速度の関係	27
13	動画の再生速度と再生時間の関係	28

表 目 次

1	別の動画による要約手法の評価結果	21
2	アンケート内容	29
3	アンケートの回答結果	29

1. 序論

現在日本では、2020年に東京オリンピック・パラリンピックの開催が予定されており、情報技術を利用した観光業を活性化させる取り組みが期待されている。インバウンドの増加に伴う訪日外国人旅行客の増加も予想され、すべての観光客に対して、ストレスなく快適に観光を満喫できる環境の整備が必要不可欠である。観光客にとって、見知らぬ土地で観光するときにあらかじめ計画を立てておくことは、効率よく観光するために重要であり、ガイドブックをはじめ、TripAdvisor¹などの口コミを含む Web サイトを用いた観光計画手法が一般的である。これらの多くは、観光スポットを単体で紹介しているものが多く、それぞれの周遊プランや観光ルート全体を紹介しているものは少ない。そのため、観光客は、どのように観光スポットを巡るかを一から考えなければならず、時として観光計画が億劫になることがある。ルート別のガイドが存在すれば、それに従って行動するだけで観光が行えるので、観光計画の手間はかからない。そこで、観光計画を効率よく行うために、ユーザの状況や要望に応じた観光ルートを推薦する研究が存在する。ユーザの嗜好に応じて観光ルートを推薦する手法 [1] や、観光地のリアルタイム情報を元に、快適な観光を可能にする観光ルート推薦手法 [2] などがあげられる。しかしながら、既存手法の多くは、推薦結果を地図やテキストを用いて表示することを想定しており、そこから移動経路の雰囲気や各観光スポット付近の様子を確認するのは難しい。推薦結果を直観的に理解するためには、動画を用いた表示方法が好ましい。動画は、地図やテキストに比べてリッチな情報を含んでいるので、観光地の雰囲気や様子を理解することができる。事実、Google の調査によると約 40% の観光客が観光に関する動画を見ながら観光計画を行なっているという結果が出ている [3]。ただし、ここでの観光動画とは、有名な観光スポットを網羅した動画で、観光ルート全体や移動経路は全く考慮されていない。

そこで本研究では、事前の観光計画を想定し作成された任意の観光ルートを動画で表示する観光動画要約システムの提案を行う。観光動画要約システムは、ユーザの嗜好や観光条件（観光場所・観光予定時間など）を収集するユーザ情報収集ステップ、収集情報に基づき観光ルートを作成する観光ルート作成ステップ、

¹<https://www.tripadvisor.jp/>

観光ルートに沿った動画や写真 (Consumer Generated Media) を収集する CGM 収集ステップ, 収集された動画や写真を用いて観光ルートに沿った動画を作成する要約動画作成ステップ, ユーザが要約動画を視聴しながら観光計画を立てる動画視聴ステップの 5 つのステップで構成される. 既存研究と同じようにユーザの要望に応じて観光ルートを決定し (ユーザ情報収集ステップ, 観光ルート作成ステップ), そのルートの移動経路と観光スポットを含む動画を作成して (CGM 収集ステップ, 要約動画作成ステップ), ユーザがその動画を視聴しながら観光計画を立てる (動画視聴ステップ). 本システムにより, 計画の手間を省きながら観光ルートを直観的に理解しつつ, 理想的な観光計画を立てることができる. 動画は, 観光スポット間を移動する移動動画と観光スポットの写真をもとめたスライドショー動画によって構成される. 筆者の事前実験では, 提示された観光ルートに対して, インターネットで検索を行うよりも, 動画でそのルートを表示した方が直観的にルートの良し悪しを判断でき, 観光意欲が向上するという結果を得た [4].

動画視聴後に観光ルート内の移動経路や観光スポットを細かく変更しながら, 再び動画を視聴する場合, 観光計画が終了するまで作成された動画を何度も視聴する必要がある. よって, 動画が長時間になってしまうと, 効率よく観光計画を行うことができないことに加え, ユーザの視聴意欲を損なう可能性がある. 短時間で視聴可能かつ, 観光計画に必要な情報を十分に得られる動画を作成することが, システム実現のための課題の一つである.

そこで, 上記の課題を解決するために, 長時間になりやすい移動動画の自動要約手法を考案し, 実装・評価する. 動画中の同じ景観が続く場面は冗長で, 観光計画を行う上で価値が低い. 景観が変化する場面のみを残し, 変化が少ない場면을省略する景観の変化に着目した動画要約手法によって, 移動動画を要約する. 要約には, 動画のシーン判別に用いられる一般的な特徴量であるカラーヒストグラムを用いる. 数フレームごとの平均カラーヒストグラムを隣接するフレームの前後で比較し, 相関によって必要か不要かを判断する.

要約手法を評価するため, 被験者 20 名を対象に, 手動による要約の結果と本手法による要約の結果を比較したところ, F 値 62.22% で必要部分を抽出できた.

また，要約後の動画を早送りすることで，より高速に動画が視聴できる．しかし，動画の速度が速くなりすぎると，視聴者が観光計画を行うために必要な情報を正しく理解できない可能性がある [5]．そこで，ユーザスタディによって観光情報を十分に得ることができる最大の再生速度を調査した．その結果，4 倍速から 8 倍速が観光動画の再生に適しているという知見を得た．評価実験により得た結果を元に動画作成を行なったところ，6 分 58 秒の移動動画を 41 秒に短縮できた．さらに，短縮後の動画を被験者 20 名に視聴してもらった結果，70%の被験者が観光計画に有用であると回答しており，本手法による動画作成が観光計画に有用であることが示唆された．

以下，2 章では，過去における関連研究について述べ，3 章では，観光動画要約システムの概要と本研究で取り組むシステム実現への課題を述べる．4 章では，課題解決のための動画要約手法の詳細とその評価結果について述べる．5 章では，作成した要約動画のユーザスタディの結果を述べる．6 章で結論および今後の課題を述べ本論文をまとめる．

2. 関連研究

2.1 観光経路の推薦に関する研究

慣れない土地で満足度の高い観光を行うためには，事前に計画を立てておくことが重要である．しかし，多くの観光客は，主要な観光スポットを調べて観光の目的地として設定するのみで，どのようにそのスポットを回るかといった緻密な計画は立てない．なぜなら，計画に時間がかかることや計画自体が面倒だと考えるためである．そこで，観光計画を容易にするために，ユーザの嗜好や状況に応じた観光経路を推薦する研究が存在する．倉田らは，誰でも簡単に計画を立てられる観光プランニング支援ツール CT-Planner を開発した [1]．またソーシャルサイトに投稿されたデータに基づいて経路を推薦する手法も存在する [6, 7]．投稿された写真には位置情報が付随しており，それらを利用することで人気スポットを特定し，そのスポットを効率よく巡回する経路を推薦することが可能である．最も単純な観光経路作成の 1 つに，観光スポット間を最短経路を利用して移動する手段が挙げられる．時間的制約を考慮した場合，最短経路を利用することでより多くの観光スポットを巡回できる．しかしながら，観光スポット間の移動を最短経路のみで構成した場合，いくつかの問題点がある．例えば，最短経路は，景観や路面店の情報を考慮していないので，観光気分を味わえない道を通ってしまう可能性がある．観光スポットへ向かう経路を含めて理想的な観光を行うべきだとして，Gavalas らは，景観を考慮した観光ルートプランニングを提案している [8]．

既存研究では，推薦結果を地図やテキストを用いて表示することを想定しているものが多く，本研究のように動画を用いて表示することに言及している研究は無い．本研究のように，動画で推薦結果を表示することができれば，ルートの景観を素早く判断できるので，効率よく理想的な観光計画を立てることができる．

2.2 コンテンツキュレーションに関する研究

観光地の様子を知ることができる動画や写真は、インターネット上に多く存在する。動画編集のプロが作成したものから、一般の観光客が自身の体験を撮影したものなど様々である。本研究では、一般の観光客が撮影した観光地の動画や写真 (Consumer Generated Media : CGM) をキュレーションしながら動画を作成することを想定している。キュレーションでは、キュレータと呼ばれる情報の編纂者が、さまざまな情報を独自の判断で収集・整理し、新しい価値を持たせて共有し、ユーザは整理された価値の高い情報を受け取ることができる [9]。近年、ニュースのキュレーションを行う SmartNews²や Gunosy³が普及しており、キュレーションという概念も浸透しつつある。カメラの小型化や高性能化に伴い、多くの人が動画を簡単に撮影できるようになったことから、複数のカメラから撮影されたスポーツ映像をキュレーションする研究も存在する [10]。このように、キュレーションの対象は動画や写真といったコンテンツの違いやジャンルの違いを問わない。しかしながら、観光情報に特化したキュレーションの既存研究は、2.1 節で述べた経路推薦の他に、Flickr⁴に投稿された位置情報付き写真から人気スポットを抽出する手法 [11, 12] や都市の写真を収集して観光計画に役立つスライドショー動画作成の提案 [13] など、その多くが写真を扱っており動画を扱う研究は少ない。その中でも Zhang らは、観光地で撮影された複数の動画から観光スポットがよく見える経路を推定する研究を行なっている [14, 15]。事前計画を支援するために動画を作成することを目的としている点は同じであるが、1つの観光スポットに着目した動画作成を行なっており、観光スポット間を移動する動画を扱う本研究とは異なる。

2.3 動画の要約手法に関する研究

観光動画要約システムの動画作成において、短編動画を作成することができれば短時間かつ効率的に観光計画が可能になる。筆者は以前、観光スポット間を移

²<https://www.smartnews.com/ja/>

³<https://gunosy.com/>

⁴<https://www.flickr.com/>

動する動画それぞれに対してユーザの嗜好に基づき動画の短縮率を決定することで動画の要約を行なった(短縮率が低いほど重要な動画とみなす)[4]. このような短編動画作成手法に加えて、動画中の不要な部分を省略する手法を加えることで、より視認性の高い短編動画の作成が可能になる. 不要なシーンを検出するためには、シーンの重要度を算出する必要がある. スポーツやニュースは、特徴的な動きや歓声などの音声、映像の切り替えといった特徴的な部分を処理することで重要度を算出し、映像が自動要約される [16]. しかし、CGMのようにワンカットで撮影されている観光動画からは、特徴的な音声や画面の切り替えが現れないことから同様の手法を用いることができない. 一方で、観光動画の位置情報を用いて主要な観光スポットが出現するシーンを重要と判断する研究 [14] や道案内動画の重要なシーンは曲がり角だとし、オプティカルフローを用いて曲がり角を検出する研究 [17] が存在する. 森下らは、入力映像から得られるカラーヒストグラム等の情報を用いて、車載カメラの映像から桜が綺麗に映るシーンを抽出する手法を提案している [18]. これらは全て重要と判断する部分が異なっており、本研究においても景観の違いに着目した重要度の算出を行うので目的が異なる. 機械学習や強化学習を用いて観光動画のシーン重要度を算出する手法 [19] も存在するが、提案するシステムにおいては、扱う映像が多いので、計算コストが低いことが望ましく、それらは適用しづらい.

本研究の要約動画作成では、動画の高速鑑賞を可能にするために、要約後の動画を早送りする. この時、再生速度が速くなるとユーザの理解度が低下してしまうことから [5], 重要なシーンを理解できる速度での再生が必要である. そこで、本研究では、情報を取得しやすい再生速度についても考察を行なっている.

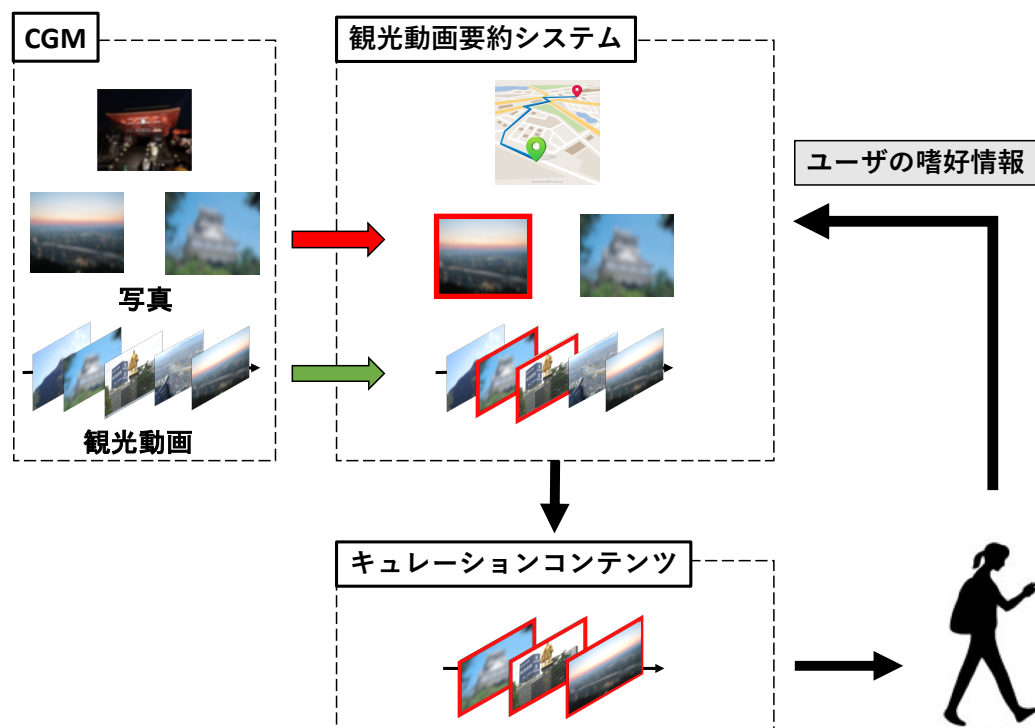


図 1 システム利用時の概要図

3. 観光動画要約システムの概要

3.1 ユースケース

観光動画要約システムは、ユーザがある1日の観光計画を立てる場面での利用を想定している。ただし、本システムは徒歩による観光のみを想定しており、交通機関を用いての移動は無いとする。図1にシステム利用時の概要を示す。ユーザが嗜好情報や観光目的をシステムに入力すると、その情報に従ってオススメの観光ルートが作成され、そのルートに沿った動画がキュレーションされる。このキュレーション動画(以下、要約動画)を視聴することで、従来の観光ルート推薦では分からなかった移動中の情報を直観的に把握できる。さらに、ルートの微調整を繰り返しながら動画を視聴することで、従来の観光計画手法より満足度が高



図 2 観光動画要約システムの概要図

い観光計画を立てられる。以下 3.2 節から観光動画要約システムのシステム構成について詳述する。

3.2 システム構成

システム概要図を図 2 に示す。観光動画要約システムは、1. ユーザ情報収集ステップ、2. 観光ルート作成ステップ、3. CGM 収集ステップ、4. 要約動画作成ステップ、5. 動画視聴ステップの 5 つのステップによって構成される。

まず、ユーザ情報収集ステップで、観光目的や観光を行う場所など観光ルート作成に必要な情報をユーザから収集する。次に、観光ルート作成ステップでは、入力情報や嗜好情報に基づき、観光ルートを作成する。CGM 収集ステップでは、動画作成に必要な素材（写真や動画）を観光ルートに沿って収集する。そして、要約動画作成ステップでは、観光ルート作成ステップで作成された観光ルートに沿った動画を作成する。要約動画視聴ステップで、ユーザが要約動画を視聴する。ここで、観光ルートが気に入らなければ、観光ルート作成ステップに戻り、観光

ルートを再び作成する。この流れを繰り返すことで、ユーザは、動画を視聴しながら理想的な観光ルートを作成できる。

3.2.1 ユーザ情報収集

ユーザ情報収集ステップでは、観光の目的地や観光に対する嗜好（ユーザが好む観光スポットの特徴）を収集する。収集方法として、ユーザが直接入力する方法と過去のデータに基づきユーザの特性をシステム自身で収集する方法があるが、ここでは前者を想定している。その場合、特に嗜好の収集に関しては、ユーザにとって負担が少なく正確であることが望ましい。もしユーザがシステムを利用している場所の位置情報を用いることができれば、周辺の情報を取り入れることで、事前計画だけでなく観光中に次の計画を立てることも可能である。

3.2.2 観光ルート作成

観光ルート作成ステップでは、ユーザに推薦する観光ルートを作成する。日高らは、各観光スポットの持つ特性とユーザの嗜好情報をそれぞれ特徴ベクトルとして表現し、2つのベクトルの内積の値が高い順に観光スポットを推薦している[2]。本システムでも同様の手法を用いて、観光可能な全てのスポットから推薦順位が高い観光スポットを観光予定時間を考慮して組み合わせる。観光スポット間の移動は、最短経路を利用する。その他、2.1節で述べたような別の手法を用いることで、多様なルート作成が可能になる。

ユーザが動画視聴後に別のルートを希望する場合、観光スポットの変更、移動経路の変更の2通りが考えられる。観光スポットを変更する場合、変更を考慮してシステム自身で観光ルートの再計算を行う。移動経路を変更する場合には、ユーザ自身で経路の微調整を手動で行えると、より多様で詳細な計画が可能になる。

3.2.3 CGM 収集

要約動画作成には、観光ルートを移動する動画と観光スポットの写真が必要となる。本システムは、一般の観光客が撮影した Consumer Generated Media (CGM)

の利用を想定している。観光客が撮影した動画や写真は、綺麗ではないがその観光地のリアルな状況を反映したものとなる。また、数が多く、同じ場所でも様々な時間や季節の情報を得ることができ、本システムには適している。動画や写真は、SNS 等から収集する方法のほかに、参加型センシング [20] の利用による収集を想定している。参加型センシングによって、リアルタイムに近い情報を収集でき、経路中の混雑度も大まかに把握できる。現状では、全ての経路に関する動画が CGM や参加型センシングの利用によって入手できるという前提が必要である。しかし今後、ウェアラブルカメラや車、ドローンに搭載されたカメラなどの様々な IoT デバイスが網羅的に各所をキャプチャできる環境が整えば、動画収集が容易になるので、広範囲な地域でのシステムの利用が可能になる。

1 つの観光スポットに対して、投稿される写真は多数存在するのでより適切な写真の選択が必要である。SNS には、投稿写真に対してコメント機能や評価機能が実装されているので、それらを元に写真を選択することが可能である。この要約動画作成に必要なデータ収集は、観光ルート作成ステップ後にシステムが自動的に行う。そのため収集される動画や写真は、観光ルートに沿ったもののみである。

3.2.4 要約動画作成

要約動画作成ステップでは、観光ルートを擬似的に体験できるような要約動画を作成する。要約動画は観光スポットの写真をもとめたスライドショー動画 V_s と、観光スポット間を移動する移動動画 V_r の 2 種類を用いる。要約動画の作成方法の概要を図 3 に示す。作成された観光ルートに沿って、観光スポットのスライドショー動画と観光スポットを移動する移動動画を交互にまとめることで要約動画を作成する。

要約動画は、計画を効率化するため再生時間を短くする必要がある。そこで、景観の変化に基づいて移動動画の要約を行なっている。この動画要約手法については、4 章にて詳述する。また、移動動画は、早送り (タイムラプス) にすることで高速鑑賞を可能にする。

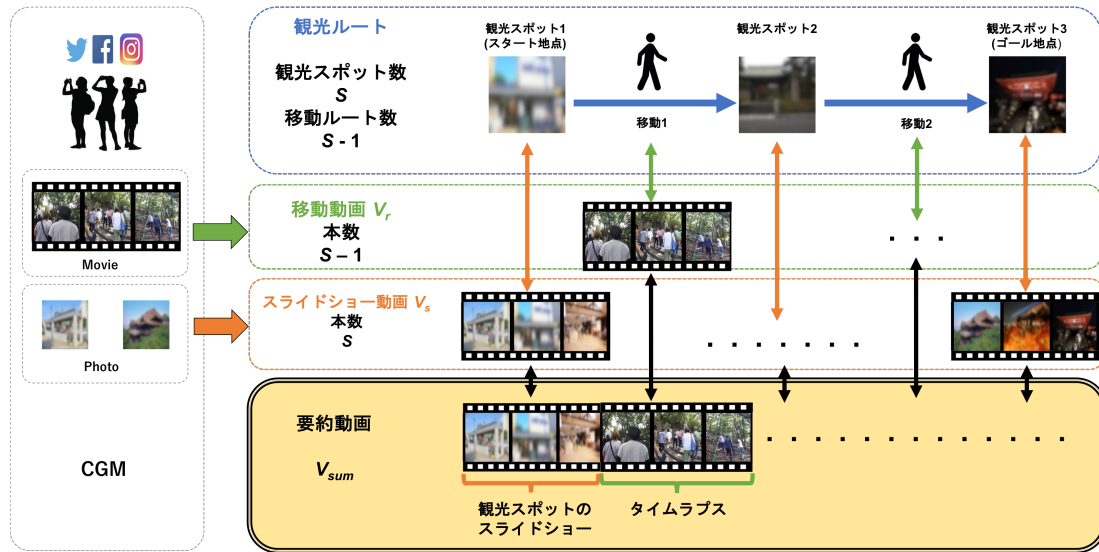


図 3 要約動画作成の概要

3.2.5 動画視聴

動画視聴ステップで、ユーザは、要約動画を視聴しながら観光計画を行う。要約動画を視聴してもらう際に表示される画面は、図4のように3種類のUIを想定している。まず、画面(A)で観光ルートの全体を確認できる。画面(A)には旅行予定日、予定日の天気予報、観光ルート、観光の所要時間が観光ルートの地図と一緒に表示されている。ここで表示される観光ルートは、観光ルート作成ステップで作成されたものと同じである。ここから画面(B)に遷移し移動動画を表示させる。

画面(B)右下には、どこを移動しているかがわかるように地図とアイコンが要約動画と連動して移動する。画面左上には移動中の絶対時間と天気を示すアイコンを表示する。ここで絶対時間は、移動時間や観光スポットでの所要時間に応じて動的に変化する。移動動画が観光スポットに到達した時点で、画面(C)に切り替わる。

画面(C)は、スライドショー動画を表示している画面である。このように画面全体にスライドショー動画を表示すると同時に、画面右下には、観光スポット名を表示する。画面左上には、画面(B)と同様に、観光スポット訪問時の絶対時間



図 4 要約動画視聴時の画面遷移

と天気アイコンに加えて、観光スポットでの所要時間を表示する。要約動画を視聴し終わると、画面 (A) に再び遷移する。

動画視聴後にユーザがルートを気に入らないと感じた場合、画面 (A) 上の観光スポット名の右側にある変更ボタンを押すことで、観光スポットを変更することができる。この動作によって、観光ルートが再計算され、その結果が再び画面 (A) の観光ルートに表示される。

季節が近い日時で撮影された動画像を収集することは可能であるが、UI に表示されている予定日や時間と、動画や写真が撮影された日時が必ずしも一致しているわけではない。そこで、こうした日時や天候を表示することで、動画の不足部分を補う。これによりユーザの擬似観光体験をサポートできるので観光計画への有用性が増す。ユーザは、一連の視聴を繰り返しながら計画を行う。

3.3 本研究で扱う研究課題

観光ルート内の移動経路や観光スポットを細かく変更する場合、観光計画が終了するまで作成された動画を何度も視聴する必要がある。この時、動画が長時間になってしまうと、効率よく観光計画を行うことができないことに加え、ユーザの視聴意欲を損なう可能性がある。短時間で視聴可能かつ、観光計画に必要な情報を十分に得られる動画を作成することが、システム実現のための課題の一つである。ここでの再生時間は、繰り返し視聴することを考慮し、30秒から60秒程度を想定している。動画を短時間にするためには、単純に早送りする方法があるが、動画の速度が速すぎると視認性が低下し、観光計画に必要な情報を得ることができない。

筆者は以前、この関係を観光動画要約問題として、最適な再生速度をユーザの嗜好から求めることで、この問題を解決する手法を提案した [4]。しかしながら、想定された再生時間を設定した場合、動画の視聴者からは十分な満足度を得ることができなかった。以上のことを踏まえて、本研究では観光動画要約システムの要約動画作成を想定し、以下の2つの課題に取り組む。

課題1：得られる情報量を変えずに移動動画を要約する

設定された再生時間による動画視聴後の満足度が低くなってしまったのは、移動動画が長時間であったために再生速度が速くなりすぎたことが考えられる。特に、移動動画は、同じ景観や景色が続く場面が多く、冗長なシーンは視聴する必要が無い。そこで、冗長で価値の低いシーンを要約すれば、視聴時間を短時間にできる可能性がある。

課題2：観光動画の再生速度と理解度の関係を明らかにする

要約後の動画を早送りすることで、より高速な動画鑑賞を実現する。しかし、視聴者が観光情報を理解できる再生速度には限界がある。ユーザが観光計画に必要な情報を理解できる速度で、再生時間を短時間にする必要がある。

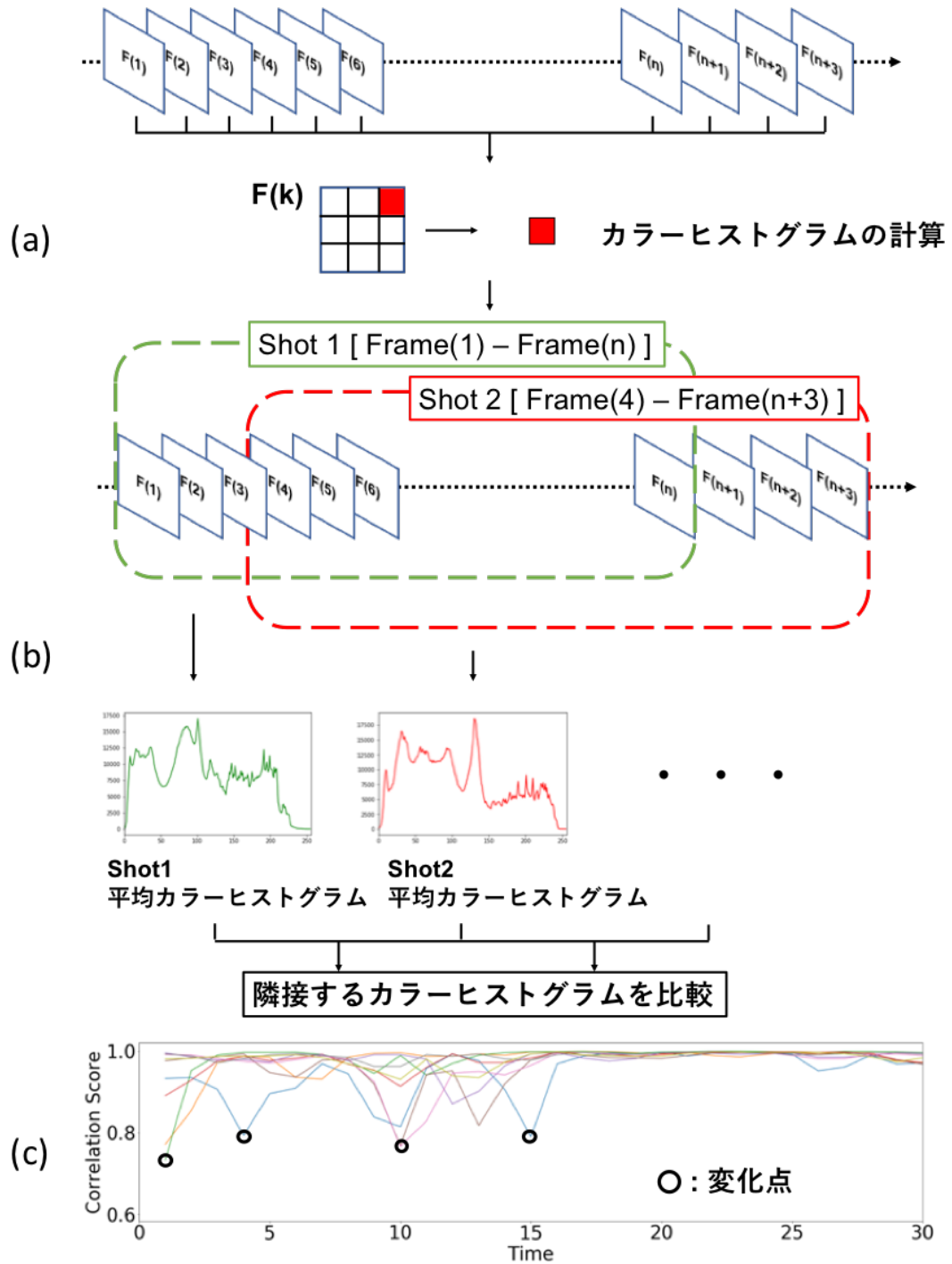


図 5 観光動画要約手法の簡略図 (フレームシフト δ が 3 の場合)

4. 景観の変化に基づく動画要約手法

3.3 節で述べた課題 1 を解決するために、移動動画を要約する必要がある。そこで、移動動画中には同じ景観や景色が長時間続くことに着目し、景観の変化に基づく動画要約を行う。動画を早送りしたり、タイムラプスにするときに視聴を快適にするために、機械学習や強化学習を用いてフレームを省略する研究が存在する [21, 22] が、本システムでは多数の移動動画を扱うことやモバイルデバイスでの利用を想定して、比較的計算コストの低いカラーヒストグラムのみで動画要約を行う。カラーヒストグラムは、動画要約に一般的に用いられる特徴量で、本研究でも [23] の手法を用いて算出する。本手法では、景観が変化する場面を変化点、2つの変化点間の映像をシーンと定義する。図 5 に動画要約の簡略図を示す。まず、図 5(a) のように、動画の全てのフレームに対して、 3×3 のエリアに分解し、それぞれのエリアのカラーヒストグラムを計算する。次に、図 5(b) のように、 n フレームごとのまとまり (ショット) の平均カラーヒストグラムを算出する。さらに、ショットを δ フレームずつシフトしながら、動画終了まで平均カラーヒストグラムを計算する。このとき、 δ と n は事前に決められており、 $\delta < n$ である。最後に、図 5(c) のように、隣接するショットの平均カラーヒストグラムを比較して、映像の変化点を抽出する。この変化点付近の映像のみを残すことで、同じ景観が続く価値の低い場面を省略した動画が作成できる。

4.1 変化点抽出アルゴリズム

動画要約時の変化点抽出アルゴリズムをアルゴリズム 1 に示す。動画の全フレームを 3×3 に分割する (1 行目)。分割された 9 つのエリアそれぞれに対してカラーヒストグラムを計算する (2 行目)。ここでは、動画に含まれる全てのフレームを N で表す。 δ は、フレームをシフトするときのシフト数を表す。全フレームに到達するまで、 δ フレームだけシフトしながら、 n フレームのまとまりごとにヒストグラムの平均値 H_s^i を計算する (3–9 行目)。 $C_s (s = 1, 2, \dots, \lfloor (N - n) / \delta \rfloor)$ は、隣接するショット H_s^i と H_{s+1}^i の相関係数を表す (10 行目)。相関係数の下位 p % が 2 つ以上のエリアで抽出された点を、変化点として抽出する (11–12 行目)。

Algorithm 1 変化点抽出アルゴリズム

Require: 観光動画 : v , 全てのフレーム数 : N , 1 ショットに含まれるフレーム数 $shot : n$, 抽出率 : p , シフトするフレーム数 : δ

Ensure: 変化点 : CP

- 1: v の全てのフレームを 9 分割する.
- 2: 全フレーム j ($1 \leq j \leq N$) における 9 つのエリア i ($1 \leq i \leq 9$) それぞれのカラーヒストグラム h_j^i を計算
- 3: $t \leftarrow 1$
- 4: $s \leftarrow 1$
- 5: **while** $t < N$ **do**
- 6: 9 つのエリア i ($1 \leq i \leq 9$) ごとに, $h_t^i, \dots, h_{t+n-1}^i$ の平均値である H_s^i を算出
- 7: $t \leftarrow t + \delta$
- 8: $s \leftarrow s + 1$
- 9: **end while**
- 10: 分割エリア i ($1 \leq i \leq 9$) ごとに, 平均値 H_s^i と H_{s+1}^i の相関係数 C_s^i を算出 . ただし s ($1 \leq s \leq \lfloor N/n \rfloor$).
- 11: $CP \leftarrow \{s \mid C_s^i \text{ の下位 } p\% \text{ を抽出した時に 2 つ以上のエリアが当てはまる点} \}$
- 12: 変化点 CP を出力

4.1.1 ショットの分割

1 ショットごとのカラーヒストグラムの平均値を用いることで, 人混みや撮影者のブレに大きな影響を受けない. フレームの分割数は, 処理速度向上のために少ない方が望ましい. しかし, 予備実験の結果, フレームを分割せずにカラーヒストグラムを計算し, 同様に変化点を抽出した場合, 小さな変化に影響されてしまい適切な抽出が行えないことが分かった. 様々な分割数を試みた結果, 3×3 の分割が処理速度と精度の点で優れていると判断した. ショットのフレーム数 n やフレームシフト数 δ は, 元の動画の 1 秒間のフレーム数 (FPS) や撮影者の動きの速度に応じて変化させる必要がある. 例えば, 元の動画が徒歩により撮影されていれば, 一般的な人の歩行速度を考慮して, 1 秒から 5 秒程度であれば映像が大きく変化することはないので, その範囲内で n と δ を設定すればよい.

4.1.2 隣接ショットのヒストグラム相関係数

各フレームのヒストグラムの値を $h_j^i (j = 1, 2, \dots, N)$ とすると, 1 ショットに含まれるフレームのカラーヒストグラムの平均値 H_s^i は, 次の式で求められる.

$$H^i = \frac{\sum_{j=1}^n h_j^i}{n} \quad (1)$$

ここで n は, 1 ショットに含まれるフレームの数である. 同様に全てのショットにおいて H^i を求める. i はフレームを 9 つに分割した際のエリアを示している. 提案手法において, ヒストグラム H_s^i, H_{s+1}^i が与えられた時, 相関係数 c は以下のよう求める. 隣接するヒストグラムにおける k フレームの l 番目の要素の値をそれぞれ, $x_{k,l}^i, y_{k,l}^i$ とすると,

$$c^i = \frac{\sum_l \sum_{k=1}^n (x_{k,l}^i - \bar{x}_k^i)(y_{k,l}^i - \bar{y}_k^i)}{\sqrt{\sum_l \sum_{k=1}^n (y_{k,l}^i - \bar{y}_k^i)^2} \sqrt{\sum_l \sum_{k=1}^n (x_{k,l}^i - \bar{x}_k^i)^2}} \quad (2)$$

ここで, $\bar{x}_k^i, \bar{y}_k^i$ は, それぞれ $x_{k,l}^i, y_{k,l}^i$ の相加平均である. 式 2 を用いて全ての隣接するショットから c の値を計算する.

4.1.3 変化点の決定

4.1.2 節で求めた c を用いて変化点を検出する. c は, フレームの分割数である 9 つの値が時間軸に沿って存在している. それぞれの c^i の数値が小さいものから順に, p % だけ抽出し, 変化点の候補とする. ある時間帯において, 2 つ以上の c^i が候補として抽出されていた場合, その点を変化点とする. この時の p は, 全体を要約したい時間に合わせて変化するものとする.

4.1.4 動画作成

変化点の抽出に伴い, 動画を作成する. 変化点の出現後, 次の変化点が登場するまでは同様の景色が続くとして, 変化点付近の映像以外は価値が低く不要なため, 変化点出現後すぐの 5 秒程度を必要部分と定義し, その部分のみを用いて動画を作成する.

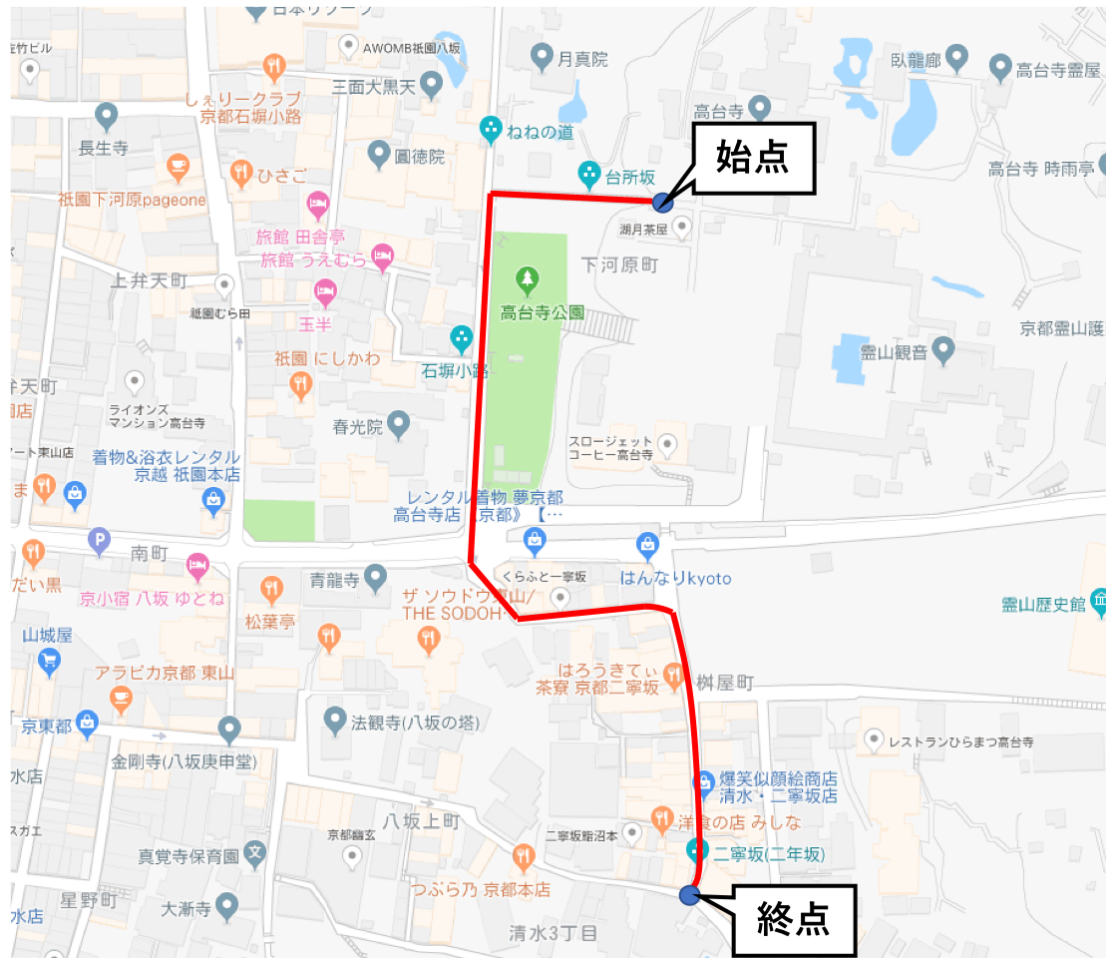


図 6 動画撮影者の移動経路

4.2 要約手法の評価

観光客を想定し，ユーザが必要だと感じる場面の考察と，提案する要約手法による必要場面の抽出精度を確かめるため評価実験を行なった．京都府京都市東山区で撮影された，高台寺から二年坂へ向かう 6 分 58 秒の動画を用いた．撮影者の移動経路を図 6 に示す．この動画には，曲がり角や路面店，人混みが映る場面があり，評価に適している．

本実験で使用される動画は，全て GoPro HERO6 により撮影されており，FPS60 に設定されている．なお，図 7 のように，肩マウンタとスタビライザを用いて，



図 7 動画撮影時のカメラの配置

撮影者の体の正面にカメラのレンズが来るように撮影された。

4.2.1 実験概要

動画は、徒歩により撮影されているので、5秒間で大きく場面が変化することは無いとし、用意した動画を5秒ごとに区切り、動画を84本に分割した。その動画を見ながら、被験者20名（20代，男性：16名，女性：4名）に元の動画を要約することを想定したときに、観光計画において必要(1)か不要(0)かを2値で判断してもらった。提案手法におけるパラメータは、今回は使用する動画がFPS60で撮影されていることを考慮し、 $s = 60$ とした。

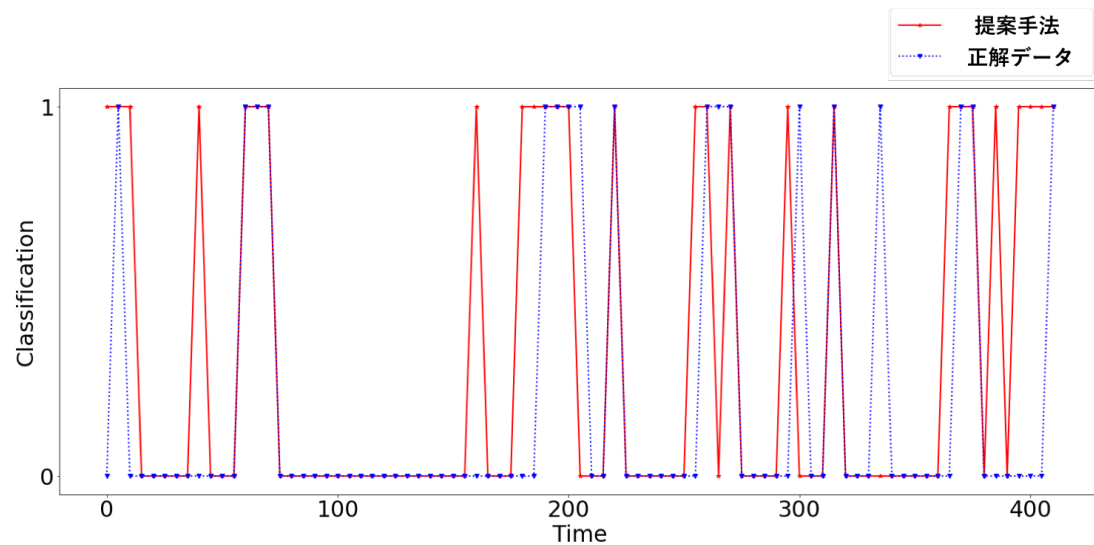


図 8 正解データ (70%agree) と提案手法の比較

4.2.2 実験結果

被験者による 2 値分類の結果を図 9 に示す．黒く塗りつぶされている点が必要と判断された動画である．被験者によってばらつきはあるが，動画の後半は特に必要と判断する被験者が多かった．これは，後半に二年坂の路面店が多く写っているため，観光用に必要と判断する被験者が多かったためだと考えられる．

さらに，被験者に必要な動画と不要な動画の違いを聞いたところ，曲がり角を曲がる動画，路面店の外観が多く見られる動画は必要と感じるという回答が多かった．一方で，人混みが多くて道路がよく見えない動画や景色に変化がなく冗長な動画は不要であるという意見が多かった．

次に，この結果を元に，被験者の 7 割以上が必要と回答した動画を正解データとして，提案手法による抽出結果との比較を図 8 に示す．動画要約により抽出された場面と正解データを比較したところ，F 値 62.22% の精度で必要部分を抽出できた．いくつかのパラメータを試した結果， $n = 300$ ， $p = 15$ と設定した場合の精度が最も高くなった．

精度が最も高くなったパラメータを用いて，別の動画でも要約手法の評価を行った．京都市東山区で撮影された別の動画を 5 本用意し，20 代の被験者 11 名

表 1 別の動画による要約手法の評価結果

動画 No.	時間	Accuracy	Precision	Recall	F-measure	ランダム抽出 の期待値
1	8m11s	0.714	0.621	0.514	0.563	0.296
2	8m28s	0.549	0.383	0.514	0.439	0.461
3	8m41s	0.615	0.439	0.514	0.474	0.394
4	8m22s	0.663	0.560	0.378	0.452	0.248
5	8m19s	0.670	0.475	0.613	0.535	0.400

(男性:10名, 女性:1名)に同じ方法で必要か不要かを判断してもらい, そのデータの多数決の結果を正解データとして, 要約手法による結果との比較を行なった. その結果を表1に示す.

表1のランダム抽出の期待値とは, 被験者が必要と判断した5秒の動画数だけランダムで抽出し, 動画要約を行なった場合の精度である (ランダム抽出の期待値 = 被験者が必要と判断した動画の数/5秒に区切った動画の総数). 正解データと比較した結果, どの動画もF値50%程度の結果となり, パラメータ推定を行なった動画以上の精度を得ることはできなかった. しかしながら, 必要箇所をランダムに決定した場合の期待値と比較すると, 5本中4本の動画は, 8%から20%の精度向上を確認できた.

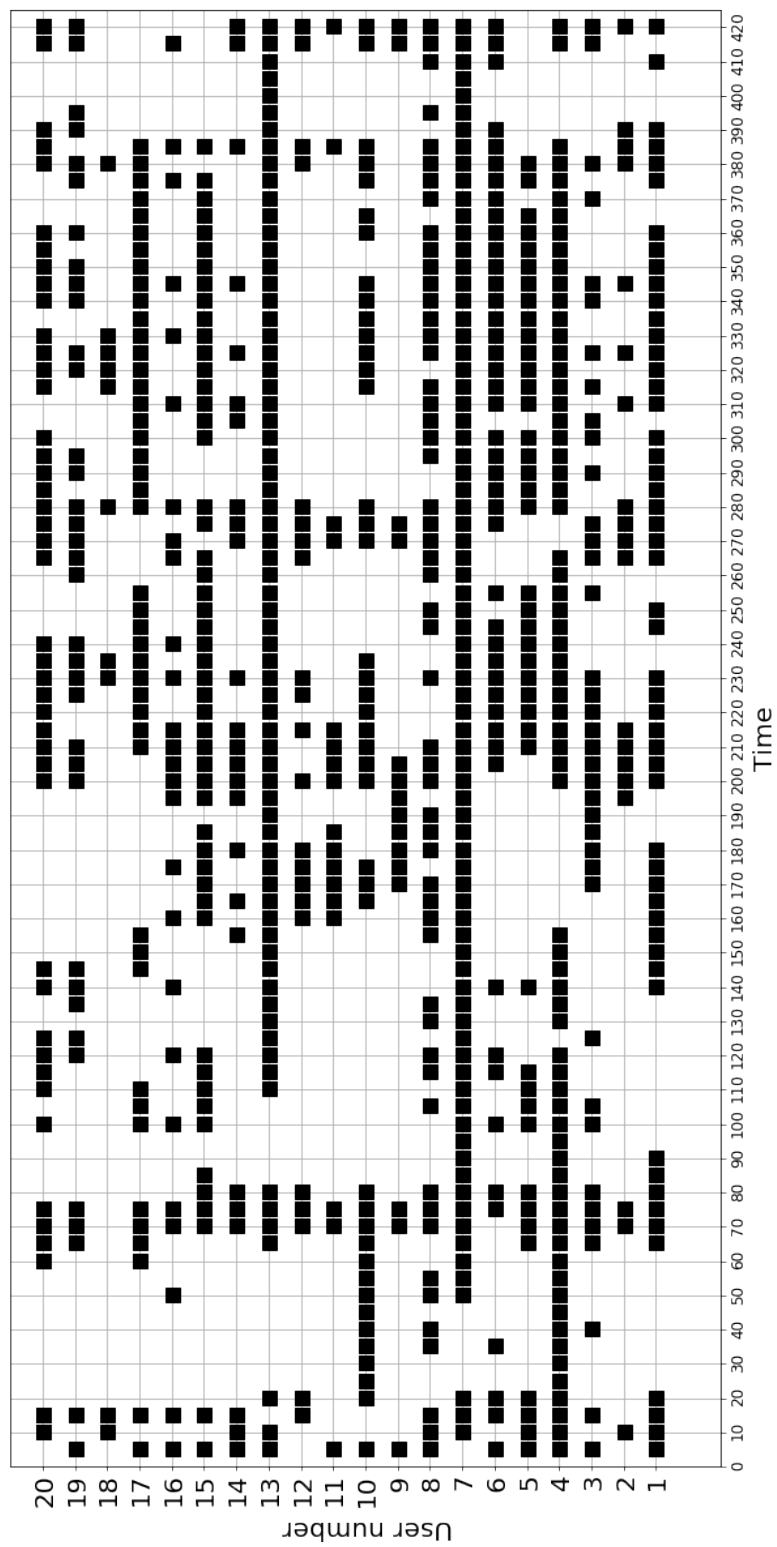


図 9 2 値分類の結果

4.2.3 考察

提案手法による抽出場面の一部を図 10 に示す．図 10(A1) のように曲がり角を写す場面や図 10(A2) のように路面店が出現する場面は，被験者の意見と同じく必要部分として抽出できていた．しかしながら，図 10(B1) のように，動画中にバスが大きく横切る場面も画面の変化が大きいので，変化点として検出していたが，観光動画としては不要であるため，誤検出となり，精度低下の要因になった．さらに，図 10(B2) のように人混みによる色彩の変化が多い場面では，カラーヒストグラムのみでは誤検出してしまうことが多かった．

また，精度確認のために要約を行なった別の 5 本の動画について，5 本中 1 本の動画のみ，ランダム抽出の期待値が要約手法による F 値を上回ってしまった．この動画 No.2 内に映る場面の一部を図 11 に示す．

この動画は，商店街の中 (図 11(A)) や八坂神社の中 (図 11(B), (C)) が撮影されており，被験者がほとんどの場面を必要と判断していたことが，ランダム期待値の上昇につながったと考えられる．つまり，景観の変化が少ない場面でも，観光スポットの内部など観光気分が味わえるような動画は，必要と判断されやすいことがわかった．

以上のことから，提案手法は，移動動画の要約には十分な精度を持つが，人混みや障害物に弱く，ロバスト性を高くする要約手法が求められる．さらに，今回パラメータ推定に用いた動画が 1 つであったため，別の動画の評価実験で要約精度が大きく上昇することはなかったが，パラメータの推定手法を確立し変更することで，精度の向上する可能性がある．

動画が撮影された日時によって，経路の混雑状況等が変化するため，動画データのみから得ることができる観光情報には，限界がある．景観の変化の有無に関わらず，日時を含めた観光スポット周辺の情報を用いた動画要約を行うことで，観光動画として有用になることが予想される．そのためには，観光客の動向を元にした Point of interest (POI) の推定手法などを組み合わせた動画作成の検討が必要である．その他観光だけでなく，様々な POI に着目することができれば，観光以外の動画にも適用できる可能性がある．

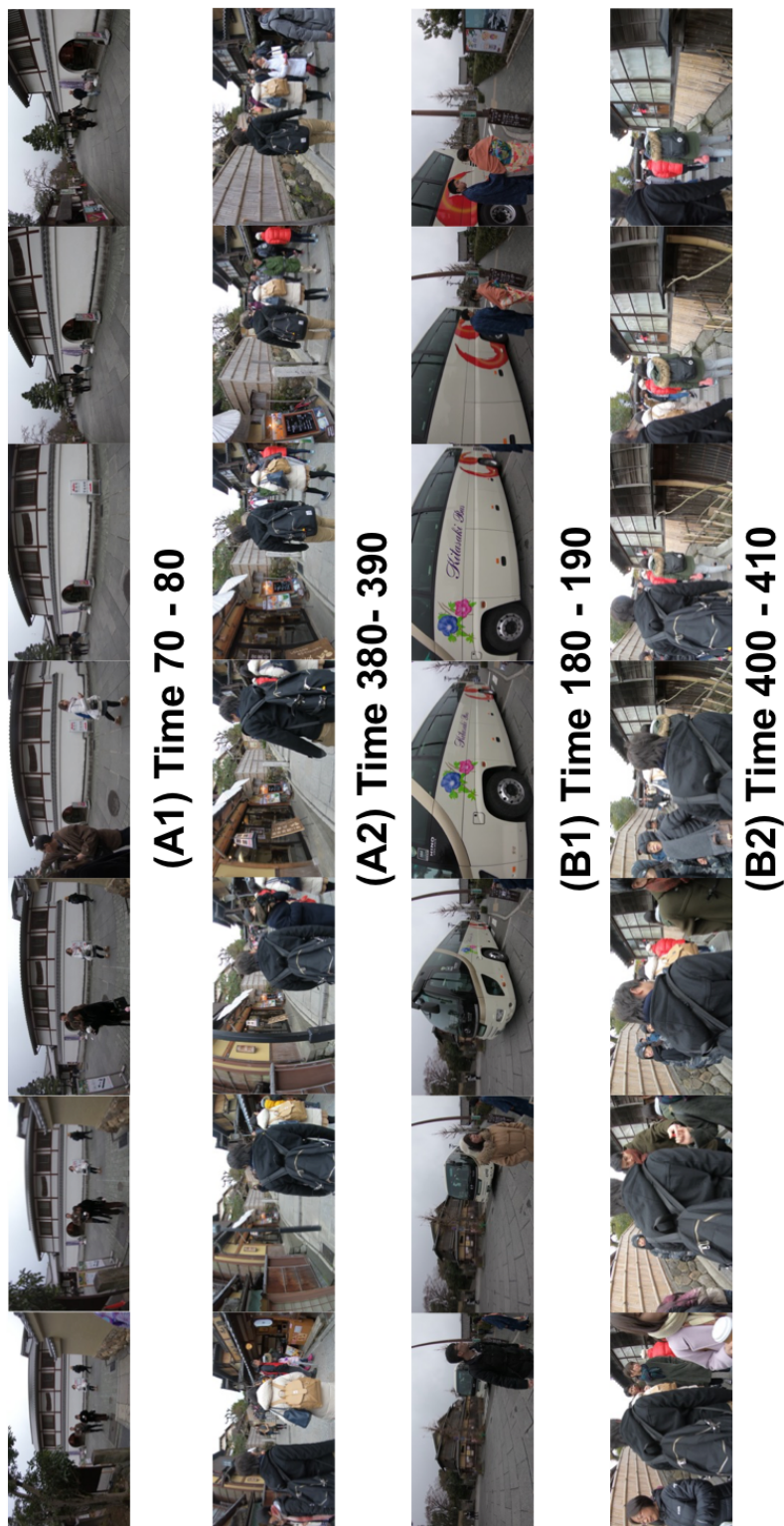


図 10 提案手法による抽出場面の一部



(A) 商店街の内部



(B) 八坂神社の内部 1



(C) 八坂神社の内部 2

図 11 動画 No.2 の場面の一部

5. 要約動画のユーザスタディ

5.1 動画の再生速度による理解度の調査

本実験は、3.3 節で述べた課題 2 を解決するための調査実験である。要約後の動画を高速に視聴可能にするため、早送りする時の最大の再生速度を求める。

5.1.1 実験概要

本実験でも、4.2 節と同じ高台寺から二年坂へ向かう動画を使用した。被験者 20 名 (20 代, 女性 : 4 名, 男性 : 16 名) に、再生速度を変化させた動画を見比べてもらった。被験者は、1) 距離感の理解度, 2) 曲がり角の理解度, 3) 路面店の理解度, 4) 移動中の雰囲気の理解度の 4 項目についてそれぞれ、1 (全くわからない) から 7 (とてもよくわかる) まで 7 段階のリッカート尺度を用いて評価した。加えて、それぞれの再生速度による再生時間をどう感じるかを、1 (短すぎる) から 7 (長すぎる) までの 7 段階のリッカート尺度で評価した。今回用意した動画の再生速度は、1 (変化無し), 4, 8, 16, 25, 40, 50, 75 倍速の 8 種類である。なお、被験者は、それぞれの動画をランダム順に視聴している。

5.1.2 実験結果

まず、4 項目の理解度と動画の再生速度についての回答結果を図 12 に示す。ここでは、1 倍速の動画を視聴した時に得る理解度が最も大きいと仮定して、1 倍速動画視聴後の回答の平均を 1.0 として各項目の結果を示している。どの項目においても、速度が向上することで理解度が低下していることが分かる。しかし距離感や曲がり角は、1 倍速と比較したとき、4 倍速でも理解度の変化は無く、8 倍速でも 90% 近い理解度を確認できた。路面店の有無や経路の雰囲気に関しても、4 から 8 倍速で 70% 以上の理解度を確認できた。本結果を応用すれば、それぞれの理解度と速度の関係を算出することができ、ユーザが理解したい情報に合わせて動画を作成することが可能である。

次に、動画の速度と再生時間に関する実験結果を図 13 に示す。グラフは、回答

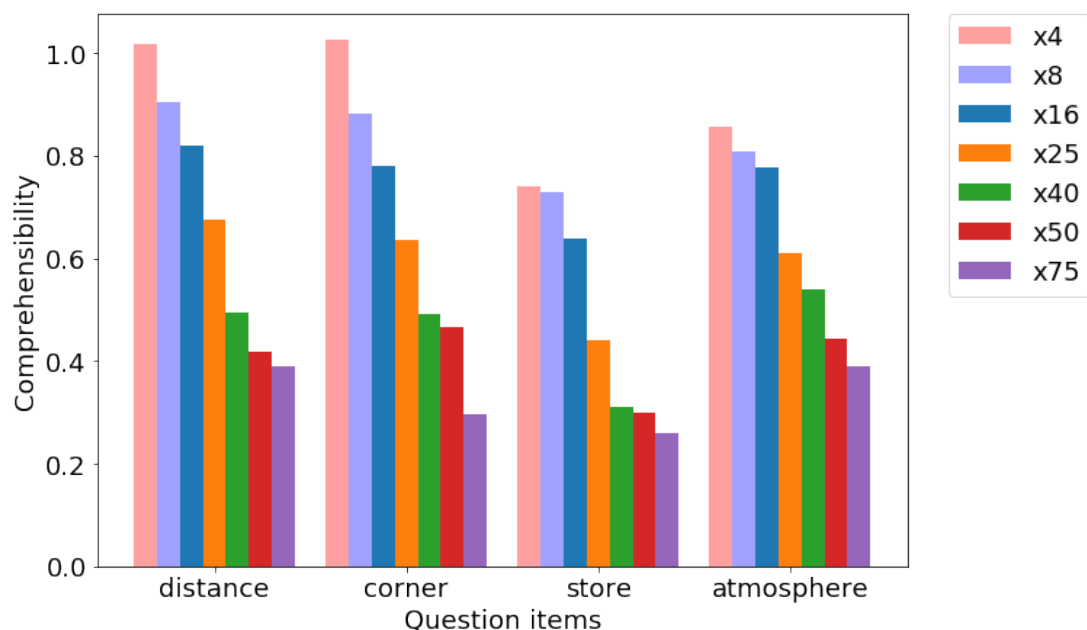


図 12 動画中の情報の理解度と再生速度の関係

結果の平均値を表している。グラフ中のバーは、回答の幅を示す。25 倍速以上になると視認性の観点からも理解が追いつかなくなり、1 に近い結果となった。しかし、再生時間の観点から速ければ速いほど良いと回答する被験者もいて、再生時間の短さの重要性がうかがえる結果となった。平均回答結果を見ると、4 倍速か 8 倍速が 4 に近い結果を示しており、多くの被験者が見やすいと判断した。

これらの結果を踏まえて、理解度と再生時間を考慮すると、動画の再生速度として最も適していると考えられるのは 4 倍速から 8 倍速の間であると考えられる。

5.2 要約動画の有用性評価

4.2 節の要約手法の評価結果と再生速度調査結果を踏まえて、観光動画要約システムが要約動画を作成したとき、観光計画にどれだけ有用であるかを評価するために実験を行なった。

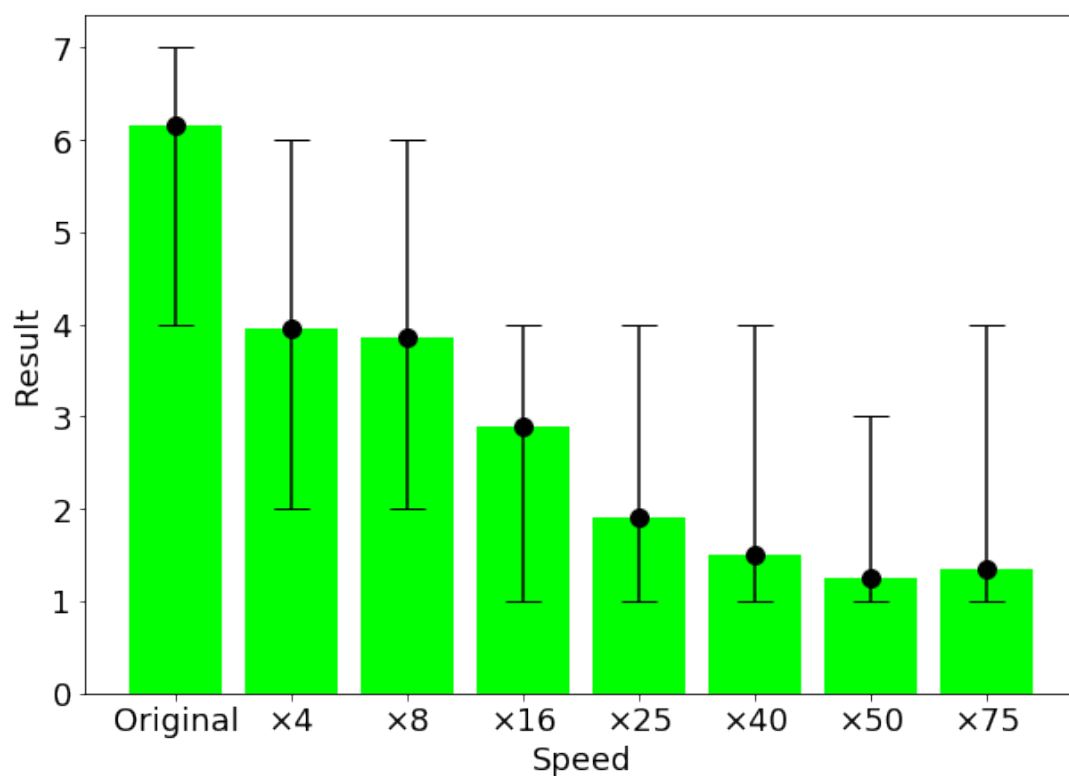


図 13 動画の再生速度と再生時間の関係

5.2.1 実験概要

高台寺と二年坂間で撮影された動画に対して、提案手法により要約を行い、必要部分を4倍速にして動画を作成し、同じ被験者20名に動画を視聴後、7段階のリッカート尺度を用いて評価してもらった。アンケート内容を表2に示す。

5.2.2 実験結果

実験結果を表3に示す。アンケートは7段階の評価になっており、数値が高いほど良い結果だと言える。Q1, Q2, Q3, Q4, Q5それぞれの回答の平均値は、5.65, 5.2, 6.7, 6.7, 4.65という結果になった。

75%の被験者は、本手法によって必要と判断されたシーンが適切であると回答した(Q2)。動画の再生速度(Q3)および動画全体の長さ(Q4)は、事前の実験結果

表 2 アンケート内容

Q1	移動経路を直観的に理解できましたか？
Q2	必要部分の選択は正しかったですか？
Q3	必要部分の再生速度は適切だと思いますか？
Q4	動画全体の長さは適切だと思いますか？
Q5	この動画は観光計画に有用だと思いますか？

表 3 アンケートの回答結果

Item	回答 (1: worst, 7: best)							平均	分散
	1	2	3	4	5	6	7		
Q1	0	0	0	2	7	7	4	5.65	0.93
Q2	0	2	1	2	4	8	3	5.20	1.51
Q3	0	0	1	0	1	0	18	6.70	0.51
Q4	0	0	0	0	3	0	17	6.70	0.39
Q5	0	3	2	1	8	5	1	4.65	1.50

を考慮しているため、適切と回答する被験者が多くなった。また、観光地の様子の理解度 (Q1) と観光計画への有用性 (Q5) を考慮しても、本手法によって作成された要約動画は、観光計画に有用となり得ると考えられる。

有用だと思わなかった被験者からは、店が出てくる部分はもっとゆっくり見せる、テロップを表示するなど、より詳細な情報を加えた動画が見たいという意見があった。また、動画要約時に誤検出している部分が計画に不要だとして、低く評価した被験者もあり、今後要約アルゴリズムの改善により、有用性が向上する可能性がある。

6. 結論

本研究では、事前に観光計画を立てる際に、動画を視聴しながら直観的に計画可能な観光動画要約システムを提案した。観光動画要約システムとは、ユーザが観光計画を立てる際に、嗜好情報や観光目的を入力すると、オススの観光ルートとそのルートに沿った要約動画を表示するシステムで、ユーザはその動画を見ながら、観光ルートの微調整を繰り返すことで理想的な観光計画を立てることができる。観光動画要約システムは、1. ユーザ情報収集ステップ、2. 観光ルート作成ステップ、3. CGM 収集ステップ、4. 要約動画作成ステップ、5. 動画視聴ステップの5つのステップで構成されている。動画視聴ステップでは、ユーザが観光ルートを決定するまで繰り返し動画を視聴しなければならないので、動画の再生時間は、短いことが望ましい。しかし、動画を短くするために早送りしすぎると、動画の理解度が低下してしまい、観光計画に有用な情報を得られない。以上のことから、要約動画の短縮に着目し、本システムにおける研究課題を2つ(課題1: 得られる情報量を変えずに移動動画を要約する、課題2: 観光動画の再生速度と理解度の関係を明らかにする)設定した。

まず課題1を解決するために、カラーヒストグラムを用いた景観の変化に基づく動画要約を実装した。要約動画は、観光スポットを移動する移動動画と観光スポットの写真を組み合わせたスライドショー動画から構成されるが、特に長時間になりやすい移動動画に着目し、要約を行なった。同じ景観が続く場面はユーザにとって価値が低い(不要)と定義し、カラーヒストグラムを用いて不要場面を省略し、さらに早送り再生を併用することで要約を実現した。まず、要約手法を評価するために、被験者20名に動画の必要・不要を手動で判断してもらった結果を正解データとし、提案手法による結果と比較したところ、F値62.22%の精度で必要部分を抽出できた。その他の動画でも、同じ手法で要約手法の評価を行なったところ、ランダムに必要箇所を抽出した場合の精度に比べて8%から20%の精度上昇が確認できた。ただし、人混みや障害物による色彩の変化に影響を受けやすいことが明らかになった。次に、課題2を解決するために、速度を変えた動画を用意し、視聴後にアンケート調査を実施した。観光計画に必要な情報としてスポット間の距離、曲がり角、路面店、経路の雰囲気の種類を設定し、それらを

理解できる最大の速度を調査したところ、4から8倍速が適切であることが分かった。再生時間を考慮した場合も同様に、4から8倍速による動画再生の評価が最も高くなった。また、評価実験により得られた知見から、移動動画を提案手法で要約し、4倍速にして要約動画を作成し、被験者に視聴してもらったところ、70%の被験者が観光計画に有用であると回答した。また、90%の被験者がルートを直観的に理解できると回答したことから、本手法による動画作成が観光計画に有用であることが示唆された。

今後の課題

今後の課題として、動画要約の精度を向上させる必要がある。精度を向上させるためには、現在用いているカラーヒストグラム他に、比較的計算コストの低いフラクタル次元などの画像特徴量を用いる方法が考えられる。また、計算コストについても既存手法との比較を行いながら評価する必要がある。さらに今回、要約手法のパラメータ推定に用いた動画が1つであったため、別の動画を考慮したパラメータ推定や推定手法の検討が必要である。

動画の再生速度について、要約前の動画のみで情報の取得量調査を行っており、要約後の動画による最適な再生速度を評価する必要がある。さらに観光動画における重要な情報として、今回定義した4つの他に、階段の有無や道幅、自動車の多少など別の情報も考慮する必要がある。

要約動画は移動動画とスライドショー動画で構成されており、今回は移動動画の評価のみを行なった。そのため、スライドショー動画の有効性を評価する必要がある。そのためには、3.2.5項で提案したUIの実装やその評価、写真や動画の適切な選択方法の検討などが必要である。さらに、写真や動画だけでなく、テキストで情報を付与することで、観光計画に向けた有用性が向上する可能性がある。テキストは、インターネット上の口コミや観光スポットのホームページから引用でき、要約動画に載せることができる。また、観光動画要約システム全体の評価のために、本研究により作成した要約動画による計画終了までの時間や、実観光における満足度の調査が必要である。

謝辞

本研究を進めるにあたり、安本慶一 教授には、研究全般に関し、多大なるご指導・ご助言を賜りました。また、充実した研究環境の整備など、研究活動を手厚くご支援いただきました。感謝の意を表すとともに、心より厚く御礼申し上げます。

中村哲 教授には、ご多忙の中、論文審査委員を引き受けてくださった上で、副指導教官として様々なご助言をいただきました。本研究関連のプロジェクトでも、ご指導いただき、的確なアドバイスを賜りました。感謝の意を表すとともに、心より厚く御礼申し上げます。

荒川豊 准教授には、ご多用にも関わらず論文の添削などで、的確なご指導およびご指摘をいただきました。また、本研究を進めるにあたり様々な視点から有用な知見やご助言を賜りました。感謝の意を表すとともに、心より厚く御礼申し上げます。

諏訪博彦 助教には、本研究を進めるにあたり、使用するデータの分析やアンケート方法を始めとした専門的な立場からのご指導・ご助言をいただきました。研究に関する相談にも、丁寧に回答してくださいました。感謝の意を表すとともに、心より厚く御礼申し上げます。

藤本まなと 助教には、学生に近い立場から、論文執筆時の基礎や本研究に関するご助言を賜りました。感謝の意を表すとともに、心より厚く御礼申し上げます。

金岡恵 事務補佐員、山内奈緒 事務補佐員、尾川恵理 事務補佐員、南あずさ 事務補佐員には、学会や出張に関する事務処理を始め、研究生生活の様々な場面でご支援いただきましたこと、謹んで感謝申し上げます。

また研究全般において、的確なアドバイスをくださった松田裕貴 先輩、中村優吾 先輩、河中祥吾 先輩を始めとする先輩の皆様、共に研究生生活を過ごしたユビキタスコンピューティングシステム研究室の同輩、後輩には、公私ともにお世話になりました。心より感謝申し上げます。

最後に、今日まで学生生活を様々な面から支えてくださった母を始め、家族に心より感謝申し上げます。

参考文献

- [1] Yohei Kurata, Yasutaka Shinagawa, and Tatsunori Hara. Ct-planner5: a computer-aided tour planning service which profits both tourists and destinations. *Workshop on Tourism Recommender Systems*, 2015.
- [2] Masato Hidaka, Yuki Matsuda, Shogo Kawanaka, Yugo Nakamura, Manato Fujimoto, Yutaka Arakawa, and Keiichi Yasumoto. A system for collecting and curating sightseeing information toward satisfactory tour plan creation. The Second International Workshop on Smart Sensing Systems (IWSSS ’17), 2017.
- [3] Ipsos MediaCT, et al. The 2014 traveler’s road to decision’ *Google Travel Study*, 2014.
- [4] Yuki Kanaya, Shogo Kawanaka, Masato Hidaka, Hirohiko Suwa, Yutaka Arakawa, and Keiichi Yasumoto. Preference-aware video summarization for virtual tour experience. *The Third International Workshop on Smart Sensing Systems (IWSSS’18)*, 2018.
- [5] Haojian Jin, Yale Song, and Koji Yatani. Elasticplay: Interactive video summarization with dynamic time budgets. In *Proceedings of the 2017 ACM on Multimedia Conference*, pp. 1164–1172. ACM, 2017.
- [6] Xin Lu, Changhu Wang, Jiang-Ming Yang, Yanwei Pang, and Lei Zhang. Photo2trip: generating travel routes from geo-tagged photos for trip planning. In *Proceedings of the 18th ACM international conference on Multimedia*, pp. 143–152. ACM, 2010.
- [7] Michalis Korakakis, Phivos Mylonas, and Evaggelos Spyrou. Xenia: A context aware tour recommendation system based on social network metadata information. In *Semantic and Social Media Adaptation and Personalization (SMAP), 2016 11th International Workshop on*, pp. 59–64. IEEE, 2016.

- [8] Damianos Gavalas, Vlasios Kasapakis, Charalampos Konstantopoulos, Grammati Pantziou, and Nikolaos Vathis. Scenic route planning for tourists. *Personal and Ubiquitous Computing*, Vol. 21, No. 1, pp. 137–155, 2017.
- [9] Keiichi Yasumoto, Hirozumi Yamaguchi, and Hiroshi Shigeno. Survey of real-time processing technologies of iot data streams. *Journal of Information Processing*, Vol. 24, No. 2, pp. 195–202, 2016.
- [10] Kazuki Fujisawa, Yuko Hirabe, Hirohiko Suwa, Yutaka Arakawa, and Keiichi Yasumoto. Automatic live sport video streams curation system from user generated media. *International Journal of Multimedia Data Engineering and Management (IJMDEM)*, Vol. 7, No. 2, pp. 36–52, 2016.
- [11] Yutaka Arakawa. Discovering popular point of interests for tourism with appropriate names from social data analysis. In *Proceedings of the 2014 International Workshop on Web Intelligence and Smart Sensing*, pp. 1–2. ACM, 2014.
- [12] Chih-Yuan Sun and Anthony JT Lee. Tour recommendations by mining photo sharing social media. *Decision Support Systems*, Vol. 101, pp. 28–39, 2017.
- [13] Feng Jing, Lei Zhang, and Wei-Ying Ma. Virtualtour: an online travel assistant based on high quality images. In *proceedings of the 14th ACM international conference on multimedia*, pp. 599–602. ACM, 2006.
- [14] Ying Zhang, He Ma, and Roger Zimmermann. Dynamic multi-video summarization of sensor-rich videos in geo-space. In *International Conference on Multimedia Modeling*, pp. 380–390. Springer, 2013.
- [15] Ying Zhang, Luming Zhang, and Roger Zimmermann. Aesthetics-guided summarization from multiple user generated videos. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, Vol. 11, No. 2, p. 24, 2015.

- [16] 滝嶋康弘. 映像の自動要約技術. 映像情報メディア学会誌, Vol. 62, No. 5, pp. 714–716, 2008.
- [17] Masaya Okamoto and Keiji Yanai. Summarization of egocentric moving videos for generating walking route guidance. In *Pacific-Rim Symposium on Image and Video Technology*, pp. 431–442. Springer, 2013.
- [18] Shigeya Morishita, Shogo Maenaka, Daichi Nagata, Morihiko Tamai, Keiichi Yasumoto, Toshinobu Fukukura, and Keita Sato. Sakurasensor: quasi-realtime cherry-lined roads detection through participatory video sensing by cars. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 695–705. ACM, 2015.
- [19] Shuyue Lan, Rameswar Panda, Qi Zhu, and Amit K Roy-Chowdhury. Ffnet: Video fast-forwarding via reinforcement learning. In *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 6771–6780, 2018.
- [20] Yutaka Arakawa and Yuki Matsuda. Gamification mechanism for enhancing a participatory urban sensing: survey and practical results. *Journal of Information Processing*, Vol. 57, No. 1, pp. 31–38, 2016.
- [21] Neel Joshi, Wolf Kienzle, Mike Toelle, Matt Uyttendaele, and Michael F Cohen. Real-time hyperlapse creation via optimal frame selection. *ACM Transactions on Graphics (TOG)*, Vol. 34, No. 4, p. 63, 2015.
- [22] Yair Poleg, Tavi Halperin, Chetan Arora, and Shmuel Peleg. Egosampling: Fast-forward and stereo for egocentric videos. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 4768–4776, 2015.
- [23] Rainer W Lienhart. Comparison of automatic shot boundary detection algorithms. In *Storage and Retrieval for Image and Video Databases VII*, Vol. 3656, pp. 290–302. International Society for Optics and Photonics, 1998.

研究業績

国際会議

1. Yuki Kanaya, Shogo Kawanaka, Masato Hidaka, Hirohiko Suwa, Yutaka Arakawa, Keiichi Yasumoto: “Preference-aware Video Summarization for Virtual Tour Experience”, International Workshop on Smart Sensing Systems (IWSSS ’ 18), Jun. 2018.
2. Yuki Kanaya, Shogo Kawanaka, Hirohiko Suwa, Yutaka Arakawa, Keiichi Yasumoto: “Automatic Tour Video Summarization Focusing on Scene Change for Advance Touristic Experience”, IEEE 2nd International Conference on Multimedia Information Processing and Retrieval (IEEE MIPR 2019), Mar. 2019.

国内会議

1. 金谷勇輝, 中村優吾, 諏訪博彦, 荒川豊, 安本慶一: “観光動画キュレーションの実現に向けたハイライト抽出手法の検討”, 2017 年度 情報処理学会関西支部 支部大会 講演論文集, Vol.2017.
2. 徳田博行, 高橋雄太, 音田恭宏, 金谷勇輝, 荒川豊, 安本慶一: “アイウェアによる集中力センシングに基づいた行動変容誘発システムの設計”, 2017 年度 情報処理学会関西支部 支部大会 講演論文集, Vol.2017.
3. 金谷勇輝, 河中祥吾, 日高真人, 諏訪博彦, 荒川豊, 安本慶一: “直感的な観光ルート計画のための観光動画キュレーションシステムの提案”, 第 86 回モバイルコンピューティングとパーベシブシステム研究会 (MBL), 2018 年 2 月.