

360 度動画における POV 情報を用いた 興味推定アルゴリズムの提案

齊藤義仰¹ 塚原敬之佑¹ 西岡大¹

概要：近年，新たな動画コンテンツの視聴方法として，360 度動画が注目されている．従来の非 360 度動画では，一般的なカメラによって撮影環境を矩形に切り取って撮影するため，視聴者は常に動画投稿者が意図した方向でのみ視聴することができなかった．一方，360 度カメラは全方位カメラにより，カメラを中心とした 360 度全体を撮影することで，視聴者は動画視聴時に視線を向ける方向を自由に変更することが出来る．また，360 度動画では従来の非 360 度動画にて用いられる時間軸を用いた平均視聴時間等の指標に加えて，視聴者がどの方向を視聴したかを示す POV 情報（視聴方向の情報）も取得し活用することができる．本研究では，360 度動画を POV 情報に基づき分析し，視聴者が興味を持ったショットを推定する手法を提案する．

A Study of Scene Analysis of 360-degree Videos using POV Information

YOSHIA SAITO¹ KYOUNOSUKE TSUKAHARA¹ DAI NISHIOKA¹

1. はじめに

近年，新たな動画コンテンツとして，360 度動画が注目されている．従来の非 360 度動画では，一般的なカメラによって撮影環境を矩形に切り取って撮影するため，視聴者は常に動画投稿者が意図した方向でのみ視聴することができなかった．一方，360 度カメラは全方位カメラにより，カメラを中心とした 360 度全体を撮影することで，視聴者は動画視聴時に視線を向ける方向を自由に変更することが出来る．また，360 度動画では従来の非 360 度動画にて用いられる時間軸を用いた平均視聴時間等の指標に加えて，視聴者がどの方向を視聴したかを示す POV 情報（視聴方向の情報）も取得し活用することができる．

一方で，360 度動画を用いた観光分野におけるメディア活用が増加傾向である．例として，ハワイ州観光局では若年層に向けた取り組みとして 360 度動画の活用，VR の活用 [1] が行われている．360 度動画に関する研究も増加傾向である．藤原ら [2,3] では観光利用を想定した，360 度動画を映像を用いたコンテンツ作成システムの開発を行っている．360 度動画を観光に活用する取り組みや研究が増加している一方，視聴者が 360 度動画のどのタイミングに興味を持ったのかについての分析は進んでいない．

本研究では，360 度動画を POV 情報に基づき分析し，視聴者が興味を持ったショットを推定する興味推定アルゴリズムを提案する．興味推定アルゴリズムを作成するにあたり，360 度動画における視聴者の POV 情報収集と，視聴者が興味を持ったショットの情報を収集する必要がある．本

稿では，実験により視聴者の POV 情報と興味を持ったショットの情報を収集し，関連性の有無について分析を行う．

本稿の構成を以下に示す．2 章では，本研究で提案する興味推定アルゴリズムについて述べる．3 章で，視聴者の POV 情報と興味を持ったショットの情報を収集するために実施した実験について記述する．4 章では，実験結果を分析し，5 節で考察を述べる．6 章で本稿をまとめ，今後の課題を述べる．

2. 提案

本研究では，360 度動画の視聴方向の情報をを用いて，視聴者の興味を分析し，分析結果から視聴者らが興味を持ったタイミングを予測する興味推定アルゴリズムを提案する．

視聴者が 360 度動画の視聴中に，どの方向を視聴したのかについてのデータを，ここでは POV 情報または視聴方向と呼ぶ．360 度動画では，視聴方向を自由に変更できる特性上，視聴者が視聴したい方向を能動的に視聴することが想定される．また視聴方向の移動量に関しては，興味を持っている被写体を視聴している場合，視聴方向を動かさずに視聴する．このため，視聴している方向を探している場合に比べて，視聴方向の移動量は少ないと想定される．一方で，視聴者が興味を持つ場所を探している場合，視聴方向を動かして視聴するため，興味をもっている被写体を視聴している場合に比べて，視聴方向の移動量は増える想定される．このような視聴者の興味関心と視聴方向の移動量の関係を活用し，視聴者が興味を持ったショットを推定する．

¹ 岩手県立大学ソフトウェア情報学部
Faculty of Software and Information Science, Iwate Prefectural University

本研究における分析対象としては、観光に活用される 360 度動画を用いた。将来的な活用に、ダイジェスト動画の作成やマーケティングへの応用が見込まれるためである。観光に活用される 360 度動画は各地域の観光協会、旅行代理店や放送局、沿線に観光客を誘致したい鉄道会社等が制作することが多く、視聴者の興味を分析する一定の需要があると見込まれる。

提案システムは、視聴者と動画投稿者が利用することを想定した。提案システムのモデルを図 1 に示す。動画投稿者は、提案システムを利用して 360 度動画を投稿する。投稿された 360 度動画は、視聴者らによって視聴される。視聴された際の視聴方向のデータは提案システムへ送られる。その後、興味推定アルゴリズムによって視聴者らが興味を持ったタイミングが分析され、その結果が動画投稿者に対して提示される。

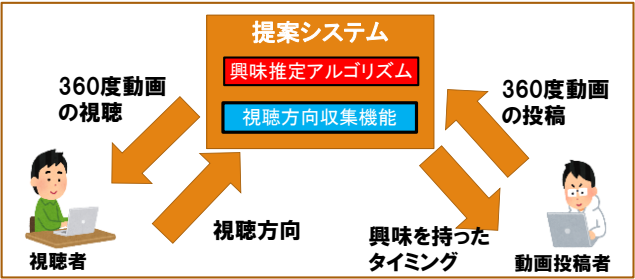


図 1: 提案システムのモデル図

3. 実験

分析アルゴリズムを作成するために調査実験を行った。この調査では、360 度動画視聴時の視聴方向の情報を調査協力者より収集する。本調査における目的は、分析アルゴリズムを作成に必要なデータの収集である。

調査実験にて視聴方向の情報を取得するために、視聴方向収集機能を実装した。視聴方向収集機能のシステム図を図 2 に示す。360 度動画の視聴ビューアである VR View[4] に改造を施し、10 ミリ秒毎に視聴方向の中心の x,y,z 三次元座標を取得できるようにした。360 度動画の視聴機能の他に、興味を持ったシーンでキーボードを押下することで、その時間と期間を記録できるようにした。また、視聴終了後に視聴ビューア上にある Write ボタンを押すことで、視聴中に取得された視聴方向の中心の x,y,z 三次元座標の情報を、csv ファイルにて書き出すことができるようにした。

実装した視聴方向収集機能を用いて、視聴時の視聴方向と視聴者が興味を持ったタイミングのデータを収集した。実験協力者は岩手県立大学ソフトウェア情報学部の学生で各動画 10 人、のべ 20 人。実験に用いたのは合計 4 本（動画 A～D）である。実験内容は、観光ジャンルの 360 度動画を視聴してもらい、協力者が動画の視聴中に興味を持ったと感じたタイミングでキーボードを押下してもらうよう

にお願いし、協力者が興味を持ったタイミングと視聴方向のデータを収集した。実験で用いた動画を表 1 に示す。これらの動画は観光 PR 用の 360 度動画であり、固定カメラと可搬カメラを用いた 2 つのカメラワークによって撮影されている。

また、動画の開始時点をショット番号 1 とし、動画終了までナンバリングを実施した。調査実験にて協力者が興味ありと解答したタイミングをショット毎に分類した。ショットとは 1 台のカメラで連続的に撮影された映像区間のことである。フェードインやフェードアウトといった特殊効果によって、カメラ視点が変わる場所のことをショット境界と呼ぶ。ショット境界の間の映像区間のことをショット境界と呼ぶ。

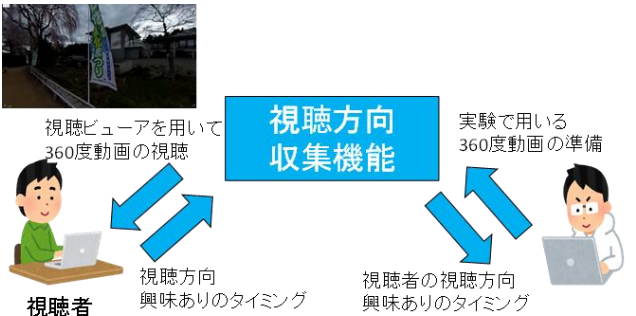


図 2: 視聴方向収集機能

表 1: 実験に用いた動画

動画のタイトル	URL
Kamishihoro VR winter (動画 A)	[5]
「360° ハワイ×男旅編×オアフ島ダイジェスト by ハワイ州観光局 360 Oahu VR Digest by Hawaii Tourism Japan」(動画 B)	[6]
VR ハワイ島 X 女子旅 by ハワイ州観光局 (動画 C)	[7]
走りたい、道がある。[北海道北広島市に行った気になれる VR コンテンツ] (動画 D)	[8]

4. 分析

調査実験にて実験協力者の視聴方向のデータを収集した。調査実験にて実験協力者の視聴方向のデータを収集した。このデータは 10 ミリ秒間隔で画面の中心の x 座標、y 座標、z 座標を示すデータである。視聴者が視聴中にどれだけ視聴方向を動かしたのかわかるために、得られた座標から式 4.1 を用いてベクトル移動量を計算する。ここにおいて x_i, y_i, z_i は i 番目に取得された視聴方向のデータの各座標、 M_i は i 番目のベクトル移動量を示す。このベクトル移動量を用いて分析を実施する。

$$M_i = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2 + (z_{i+1} - z_i)^2} \cdots \text{式 1}$$

今回の分析において「視聴者が興味あり」と判断するシ

ショットの基準は、実験協力者のうち3割以上が興味ありと回答したショットであることかつ、興味ありと判断する動画の全ショットのうち5割以下の数とした。これは3割以上が興味ありと回答したショットが全体の8割を超える数になってしまうことがあったことからである。

収集した個々のデータを10件分合算し、興味関心の分析を行った。分析に用いた指標は、移動量のデータをショット毎に合算したデータの分散、積算、平均を用いた。分散はショット内での移動量の変化の幅を示す。積算はショット内での移動量そのものを示す。平均はショット内での移動量の変化を表す。興味があると視聴者が解答した人数が多いショット順に並べた表を、動画毎に表2から表5に示す。

動画はショットの集合体である。そこで動画の各ショットを分析する際に、連続するショットの関係性を考慮する必要があると考えられる。関係性を考慮しない単独のショットでは、視聴者が興味を持っているかどうかの比較の対象が無いからである。このことから、ショット毎に10件分データを合算したデータにおいて直後のショットとの関係性を考慮して時間軸的な意味での直前のショットと比較して興味を持っているかどうかという視点で分析を試みた。

表 2: 動画 A の興味ありショットとその回答数

順位	ショット番号 (興味ありと回答したデータ数)
1	16 (9)
2	12 (7)
3	14 (6)
4	7,17 (5)
5	9,20,22,23,24,30 (4)
6	19 (3)
7	3,8,10,11,13,15,27,28,29 (2)
8	2,4,5,6,25,26,31,32,33 (1)
9	1,21,18 (0)

表 3: 動画 B の興味ありショットとその回答数

順位	ショット番号 (興味ありと回答したデータ数)
1	3,4 (9)
2	8 (8)
3	5,7 (7)
4	2 (6)
5	6,11,13 (5)
6	10,12,14 (4)
7	9 (3)
8	15 (1)
9	1,16 (0)

表 4: 動画 C の興味ありショットとその回答数

順位	ショット番号 (興味ありと回答したデータ数)
1	18 (9)
2	9,10 (8)
3	3,15,22 (7)
4	7 (6)
5	5,8,12,14 (5)
6	2,4,11,13,16,19,23 (4)
7	6,17 (3)
8	1,21 (2)
9	20 (1)

表 5: 動画 D の興味ありショットとその回答数

順位	ショット番号 (興味ありと回答したデータ数)
1	8(8)
2	3,6,7(7)
3	4,5(5)
4	2,10(4)
5	9(2)
6	11(1)
7	1,12(0)

4.1 直前のショットとの変化量を用いて分析した結果

まず、収集した個々のデータを10件分合算したものを、時間軸的に直前のショットと変化量を算出したものの活用することで興味関心の分析を行った。分析に用いた指標は、移動量のデータをショット毎に合算したデータの分散 σ^2 の変化量、積算 Sum の変化量、平均 Ave の変化量を用いた。直前のショットとの分散、積算、平均の変化量の計算式を式2から式4に示す。 d はショット番号を示す。

直前のショットとの分散の変化量

$$\sigma^2_{(d-1)-d} = \sigma^2_{(d-1)} - \sigma^2_d \quad \cdots \text{式 2}$$

直前のショットとの積算の変化量

$$Sum_{(d-1)-d} = Sum_{(d-1)} - Sum_d \quad \cdots \text{式 3}$$

直前のショットとの平均の変化量

$$Ave_{(d-1)-d} = Ave_{(d-1)} - Ave_d \quad \cdots \text{式 4}$$

分散の変化量は、ショット内での移動量の変化の幅の差を示す。積算の変化量は、ショット内での移動量そのものの差を示す。平均の変化量は、ショット内での移動量の変化の差を表す。これらのデータを用いて、サポートベクターマシンによる興味ありショットの線形分類を実施した。サポートベクターマシンによって線形分類を実施する際に、分散、積算、平均の変化量の3つより2つを選択し、二次元の分類を作成する。図3から図5は、それぞれの組み合わせ毎に、散布図で表したものである。

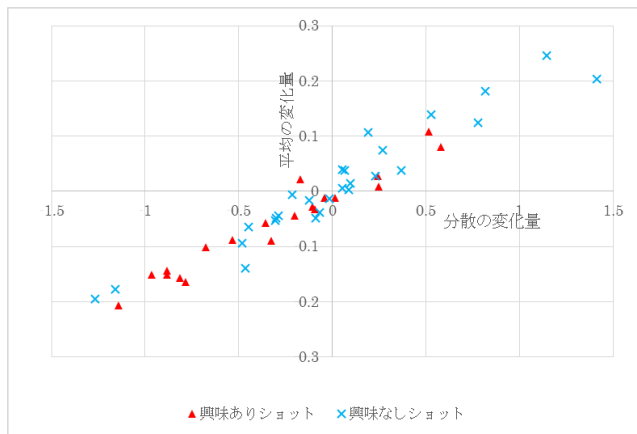


図 3: 直前のショットとの分散, 平均の変化量の散布図

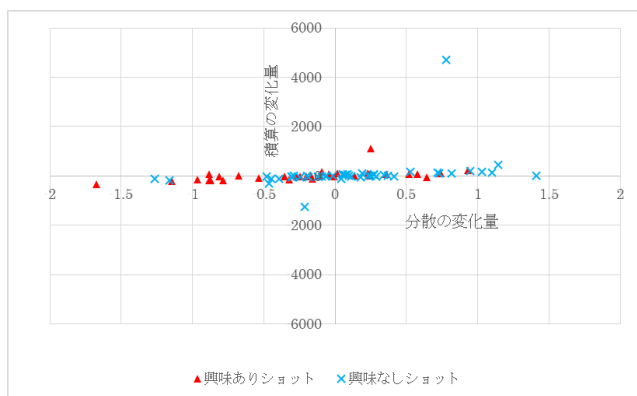


図 4: 直前のショットとの分散, 積算の変化量の散布図

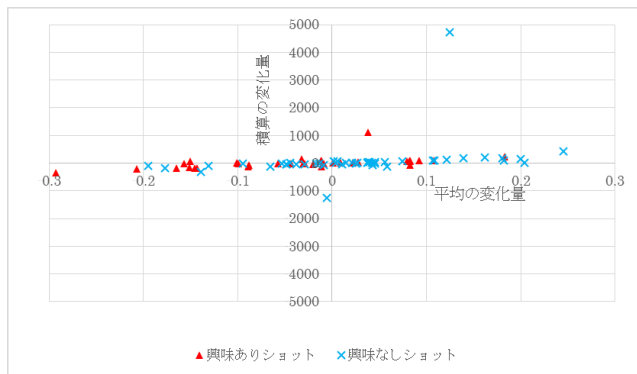


図 5: 直前のショットとの平均, 積算の変化量の散布図

この3つの組み合わせについてそれぞれサポートベクターマシンにて線形分類を実施し, 性能比較した. 表 6 に, それぞれの組み合わせの F 値を記す. この結果より, 一番性能のよい結果は, 分散の変化量と平均の変化量を用いた組み合わせであった. 図 6 に線形分類した結果と, 表 7 に分散の変化量と平均の変化量を用いた場合の, 動画 A-D における再現率, 適合率, F 値を示す. 以上の結果から, 興味ありショットの推定にはサポートベクターマシンの活用が有効であると考えられる. しかし, 動画 A 以外の場合, 特に動画 B に関しては, 性能が低い結果となった.

表 6: SVM を用いた各組み合わせの性能比較

	分散, 平均 の変化量	分散, 積算 の変化量	平均, 積算 の変化量
F 値	0.6027	0.492	0.543

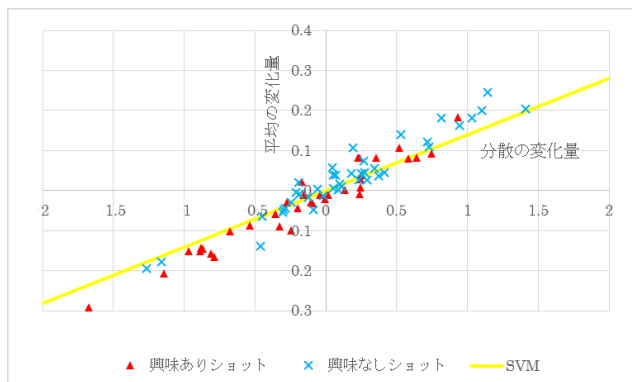


図 6: 直前のショットとの分散, 平均の変化量を用いた場合の線形分類の結果

表 7: 直前のショットとの分散, 平均の変化量を用いた場合の動画 A-D における再現率, 適合率, F 値

	再現率	適合率	F 値
動画 A	0.75	0.563	0.642
動画 B	0.33	0.33	0.33
動画 C	0.636	0.7	0.667
動画 D	0.5	0.5	0.5

4.2 直後のショットとの変化量を用いて分析した結果

直後のショットとの変化量をサポートベクターマシンにより, 興味ありショットの線形分類を実施した. 利用するデータは直前のショットとの変化量と同様, 直後のショットとの変化量のうち分散の変化量と平均の変化量である. 直前のショットとの分散, 平均の変化量の計算式を式 5 と式 6 に示す. d はショット番号を示す.

直後の分散の変化量

$$\sigma^2_{d-(d+1)} = \sigma^2_d - \sigma^2_{d+1} \quad \cdots \text{式 5}$$

直後の平均の変化量

$$Ave_{(d-1)-d} = Ave_d - Ave_{d+1} \quad \cdots \text{式 6}$$

直後のショットとの変化量なので最後のショットに関しては計算ができない. そのため最後のショットに関しては分析から除外した. 線形分類を用いた結果を図 7 に示す. 動画毎の再現率, 適合率, F 値を用いて性能評価を実施した結果を表 8 に示す. 直前のショットとの変化量を利用したものに比べ, 動画 B に関しては再現率, 適合率, F 値ともに大幅に上昇した. しかし, 他の動画の性能に関しては直前のショットを用いたものと比べて性能が低下した. こ

れより直前のショット，直後のショット双方を考慮し分析をする必要があると考えられる。

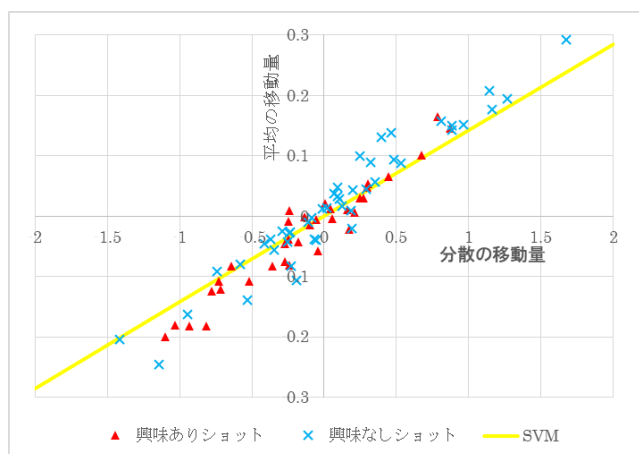


図 7: 直後のショットとの分散，平均の変化量を用いた場合の線形分類の結果

表 8: 直後のショットとの分散，平均の変化量を用いた場合の動画 A-D における再現率，適合率，F 値

	再現率	適合率	F 値
動画 A	0.5	0.4	0.444
動画 B	0.833	0.625	0.71
動画 C	0.54	0.4615	0.5
動画 D	0.66	0.8	0.727

4.3 直前のショットとの変化量と直後のショットとの変化量を用いて分析した結果

直前のショットとの変化量と直後のショットとの変化量を合算したものの活用し，サポートベクターマシンによる興味ありショットの線形分類を実施した．利用するデータは，分散の変化量と平均の変化量である．直前・直後のショットとの分散，平均の変化量の計算式を式 7 と式 8 に示す． d はショット番号を示す．

直前・直後のショットとの分散の変化量

$$\Delta\sigma^2_d = \sigma^2_{(d-1)} - \sigma^2_{(d+1)} \quad \cdots \text{式 7}$$

直前・直後のショットとの分散との平均の変化量

$$\Delta Ave_d = Ave_{(d-1)} - Ave_{(d+1)} \quad \cdots \text{式 8}$$

線形分類を用いた結果を図 8 に示す．動画毎の再現率，適合率，F 値を用いて性能評価を実施した結果を表 9 に示す．直前のショットとの変化量を利用したものに比べ，性能が向上した．一方，直後のショットとの変化量を利用したもので性能が比較的良好だった動画 B に関しては，直前のショットとの変化量を利用したものと同様に性能が低かった．動画 B 以外に関しては良好な結果が得られているこ

とから，総合的に直前・直後のショットとの分散，平均の変化量を用いて線形分類する手法が，興味推定アルゴリズムとして最も有効であると考えられる。

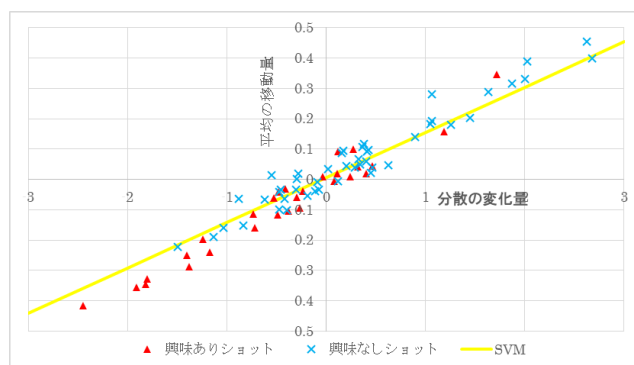


図 8: 直前・直後のショットとの分散，平均の変化量を用いた場合の線形分類の結果

表 9: 直前・直後のショットとの分散，平均の変化量を用いた場合の動画 A-D における再現率，適合率，F 値

	再現率	適合率	F 値
動画 A	0.91	0.647	0.785
動画 B	0.334	0.4	0.364
動画 C	0.909	0.833	0.870
動画 D	0.5	0.6	0.55

5. 考察

平均の変化量と分散の変化量がどちらもマイナスの範囲，平均の変化量と分散の変化量がどちらもプラスの範囲，平均の変化量と分散の変化量がどちらも 0 付近の範囲の 3 つの範囲に分類して考察する．考察する 3 つの範囲を図 9 にて記す．

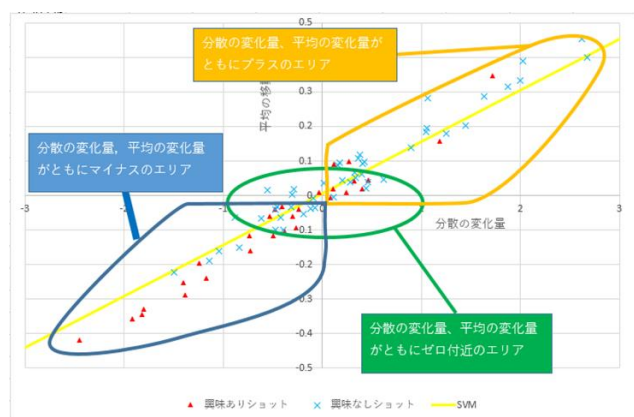


図 8: 考察する 3 つの範囲

5.1 平均，分散の変化量がどちらもマイナスの範囲

平均の変化量がマイナスなので移動量をより少なくし

ていると考えられる。分散の変化量がマイナスなので、より移動量の幅を狭くしていると考えられる。つまり、視聴者が視聴方向を動かさずに見ていると考えられる。360 度動画は見たい視点を視聴者が能動的に探しながら視聴しなければいけない特性から考えると、視聴者は動かさずに興味のある所を視聴していると考えられることから、この範囲は視聴者の興味があるショットであると推定される。

5.2 平均、分散の変化量がどちらもプラスの範囲

平均の変化量がプラスなので、移動量をより少なくしていると考えられる。分散の変化量がプラスなので、より移動量の幅を狭くしていると考えられる。つまり、視聴者が視聴方向を動かして見ていると考えられる。360 度動画は見たい視点を視聴者が能動的に探しながら視聴しなければいけない特性から考えると、視聴者は視聴方向を動かし、興味のある所を探しながら視聴していると考えられる。よって、この範囲は視聴者の興味がないショットであると推定される。

5.3 平均、分散の変化量がどちらも 0 付近の範囲

この範囲は、前述した視聴者の興味があるショットと視聴者の興味がないショットに加え、視聴者が動画の内容に飽きて視聴を放棄したショットも存在すると考えられる。この 3 つのショットが混在し、検出精度の低下が発生している。分析において性能が伸び悩んだ動画 B と動画 D に関しては、この 0 付近の範囲にあるショットがほかの動画に比べて多く、検出精度の低下の原因となったと考えられる。

6. おわりに

本稿では観光分野に活用される 360 度動画の視聴方向を用いて、視聴者の興味のあるショットを推定する興味推定アルゴリズムを作成するため、実験および分析を行った。実験では、360 度動画の視聴方向のデータと視聴者の興味のあるタイミングを収集し、視聴者の興味を推定する方法に関して分析を行った。直前のショットとの変化量と直後のショットとの変化量を合算した平均の変化量と分散の変化量を活用し、サポートベクターマシンにて線形分類を実施することで、視聴者の興味をある程度推定できることがわかった。今後は、さらなる推定精度の向上のため、分析を続けていく。

参考文献

- [1] ハワイ州観光局のプレスリリース:
<https://prtimes.jp/main/html/rd/p/000000037.000009504.html>
(最終アクセス日:2018 年 2 月 19 日).
- [2] 藤原佑歌子, 吉野孝: 観光予定者のためのパノラマ画像を介した情報共有システムの提案, DICOMO2014, pp.430-437 (2014).
- [3] 藤原佑歌子, 吉野孝, 児玉康宏, 吉住千亜紀, 尾久土正己: パノラマ画像および動画を用いた観光支援システムの提案, 情報処理学会関西支部支部大会講演論文集, 2010G- 10 (2014).
- [4] VR View for the Web:
<https://developers.google.com/vr/develop/web/vrview-web>

- [5] KamishihoroVR winter – YouTube:
<https://www.youtube.com/watch?v=pK6SlhXnd10>
- [6] 360°ハワイ×男旅編×オアフ島ダイジェスト by ハワイ州観光局 360 Oahu VR Digest by Hawaii Tourism Japan – YouTube:
<https://www.youtube.com/watch?v=93MSuK2AX14>
- [7] VR ハワイ島 X 女子旅 by ハワイ州観光局 – YouTube:
<https://www.youtube.com/watch?v=twGBaCcmT8M>
- [8] 走りたい、道がある。[北海道北広島市に行った気になれる VR コンテンツ] – YouTube:
<https://www.youtube.com/watch?v=glmTayrIGmw>