

一人称視点映像による実環境の記憶可能性推定

大泉 建人^{1,a)} 中澤 篤志^{1,b)} 西田 豊明^{1,c)}

概要: 実環境中の建物や看板、掲示物などのオブジェクトに対して、人がどのようなものを良く記憶できるかという情報は、まちづくりやマーケティングにおいて有用である。我々はこの、実環境中のオブジェクトの記憶されやすさを「オブジェクトの記憶可能性」と呼び、一人称視点映像を基にオブジェクトの記憶可能性を推定する。このため、一人称視点映像を撮影するためのウェアラブルカメラを装着した状態で、被験者に指定したルート上で散策を行うタスクを課し、得られた映像とアンケートによって得られたオブジェクトの記憶可能性との関係を調査した。その結果、一人称視点映像における出現回数が多く、停留時間の長いオブジェクトについて記憶可能性が高くなるという関係が示唆された。今後の課題として、オプティカルフローやフレーム間の輝度値の差を用いて頭部運動を推定する処理を改善する事が挙げられる。

キーワード: 一人称視点映像, 記憶可能性, 注視行動

KENTO OIZUMI^{1,a)} ATSUSHI NAKAZAWA^{1,b)} TOYOAKI NISHIDA^{1,c)}

Abstract: Information of whether the person can well remember the objects in real environment such as buildings, signs and notices, are useful in city planning and marketing. We defined the stored ease of objects in real environment as “environmental memorability”. It is considered that environmental memorability is effected by parameters which express status of an individual. For example, the person looks at what he have an interest. First-person video data is very good as the data for recording the gaze and it is associated with objects easily. Then we propose a method to estimate the environmental memorability with first-person video. We did experiment to find out the relation between a first-person video data and the environmental memorability. The experiment is which subjects walk the designed route with wearable camera and answer a questionnaire. Then we analyze the first-person video and compare with the result of questionnaire. As a result, we found that the more times and the longer an object appear, the higher environmental memorability is. Improving the process for estimating the head movement by using the optical flow or the difference in luminance value between frames is considered as a future issue.

Keywords: first-person video, memorability, gaze

1. 序論

実環境中の建物や看板、掲示物などのオブジェクトに対して、人がどのようなものを良く記憶できるかという情報は、まちづくりやマーケティングにおいて有用である [7][9]。例えば、災害対策に重要な避難情報を覚えやすい所に掲示したり、広告をより記憶されやすいように設置するなどの利用方法があり、新たな地図情報として幅広い用途に活用できる。我々はこの、実環境オブジェクトの記

憶されやすさを「オブジェクトの記憶可能性」と定義し、一人称視点映像を基に推定する手法を提案する。

近年、一人称視点映像が様々な分野で用いられている。一人称視点映像とは、頭部等に装着されたカメラから撮影される映像である。人が実際に見るシーンとほぼ同じシーンが映像として得られることから、一人称視点映像は日常生活における行動の記録や分析に用いられている。例えば、相澤ら [8] はウェアラブル機器により取得された長時間の生活体験映像の情報量について論じた上で、ウェアラブルカメラ映像の構造化と要約を行っている。この研究では、人の動きの情報という客観的特徴を用いることで構造化を行い、脳波情報を用いることで個人の主観を考慮した要約を行っている。また、樋口ら [10] は一人称視点映像か

¹ 京都大学大学院情報学研究科
Graduate school of Informatics, Kyoto University

a) oizumi@ii.ist.i.kyoto-u.ac.jp

b) nakazawa.atsushi@i.kyoto-u.ac.jp

c) nishida@i.kyoto-u.ac.jp

ら頭部の動きを高精度に計測し、社会的地位の推定を行っている。

実環境中のオブジェクトの記憶可能性は、そのオブジェクトを視認できる範囲を通過した人の中で一定時間後に再び思い出す事の出来る人の割合と定義できる。これは、ある未知シーンにユーザを歩かせ、その後アンケート等を用いることで求めることができる。しかし、アンケートに基づく方法は大きなコストがかかるため、自動化する必要がある。そこで、本稿では人の前方向を撮影する一人称視点映像を基に記憶可能性を推定することを考える。一人称視点映像を用いる理由として、まず近年のウェアラブルカメラの普及が挙げられる。また、一人称視点映像中から注視行動が検出可能であることが挙げられる。人は目の前に興味を惹かれるものがあつた場合、注視を行う。興味を惹かれるものは記憶に残りやすいため、記憶可能性は注視行動に現れると考えられる。よって、一人称視点映像から注視行動を検出することで記憶可能性が推定できる。

本稿では、一人称視点映像を基にオブジェクトに対する個人の記憶可能性を推定するためのモデルを構築する。集められた一人称視点映像それぞれに対してこのモデルを用いることでオブジェクトに対する撮影者個人の記憶可能性を推定し、個人の記憶可能性を統合することでオブジェクトの記憶可能性を推定することを目指す。

2. 関連研究

本章では記憶や一人称視点映像に関する関連研究を挙げ、本研究の位置づけを明確にする。

人の記憶に関しては古くから研究が為されている。オブジェクトの記憶の関連研究として Brady らの研究が挙げられる。Brady ら [2] は、長期記憶が多くのオブジェクトの詳細な蓄積に秀でていることを基に、視覚的情報を蓄積する長期記憶能力に関する実験を行っている。この実験では、まず 2500 枚のオブジェクト画像を被験者に見せ、記憶させる。その後、事前に見せたオブジェクト画像とそのオブジェクト画像によく似た画像のペア画像に対し、事前に見せられたと思う画像を被験者に選ばせている。また、Konkle ら [4] は視覚的長期記憶におけるシーン描写の忠実性を調査している。Konkle らは学習フェーズとテストフェーズの二つのフェーズに分けられた実験を行った。学習フェーズでは、128 の異なるシーンカテゴリから選んだ 2912 枚のシーン画像を被験者に見せ、記憶させる。学習フェーズ後のテストフェーズでは、Brady らの実験と同様の二者択一の質問を被験者に行い、質問の正答率からシーン描写をどれだけ忠実に記憶しているかを導出している。

また、Isola ら [3] は画像の記憶のしやすさには共通の特徴が存在することを調査している。Isola らはまず、画像の記憶されやすさを数値化するために、画像を被験者が一度見ただけで覚えることができるかを調べるゲームを作成し、

実験を行った。そして画像の覚えやすさ (memorability) を記録した画像データベースを構築し、記憶のしやすさに影響を与える画像特徴を分析している。その結果を基に、画像記述子を元に予測機を学習することで memorability の予測を行なうことが可能である事を示した。

一方本研究では、画像の記憶しやすさではなく、実環境中のオブジェクトに対する記憶しやすさを扱う。これは、単なる画像特徴のみならず、頭部運動に基づくオブジェクトを見た回数等の個人の状況を表すパラメータにも影響を受けていると考えられる。

人の行動を観察するために一人称視点映像を用いる技術に注目が集まっている。Berry ら [1] は、記憶に障害を抱えた人にウェアラブルカメラで一人称視点映像を記録させ、介護者と共に閲覧することによって出来事を覚え、想起できる度合いに良い効果があつたと述べている。また、山田ら [6] は、視覚的顕著性マップモデルを用いて一人称視点映像に対して視覚的注意を推定する手法を提案している。本研究においてもこのような一人称視点映像の特性を個人の状況を記録するデータとして用いる。一人称視点映像中には、視界とオブジェクトの関係等、視線情報から得られる情報の一部が含有されているはずである。これを用い、一人称視点映像から記憶可能性を推定することができれば、アンケート等を用いる方法と比較して低コストで記憶可能性を知ることができる。そこで本研究では一人称視点映像を用いて得られる情報と記憶可能性の関係を調査する。

3. 記憶可能性モデル

実環境中のオブジェクトの記憶可能性は、興味を惹かれて長時間眺めたり、気になって振り向いたりすると高くなると考えられる。また、見る対象となるオブジェクト自身の持つ特徴も影響する。その他にも、記憶可能性はオブジェクトを見ているときの人の状態にも影響される。

以上より、オブジェクト O_i の記憶可能性 $P_{remember}$ は、フレーム数 T の一人称視点映像 $\mathbf{I} = (I_1, \dots, I_T)$ 、撮影者の状態 $\mathbf{S} = (S_1, \dots, S_T)$ を用いて以下のように表されると仮定する。

$$\underbrace{P_{remember}(O_i, \mathbf{I}, \mathbf{S})}_{\text{記憶可能性}} = \underbrace{P_{personality}}_{\text{個人差}} \cdot \underbrace{P_{memorability}(O_i)}_{O_i \text{ の記憶されやすさ}} \cdot \underbrace{P_{seen}(O_i, \mathbf{I})}_{I \text{ からわかる情報}} \cdot \underbrace{P_{status}(\mathbf{S})}_{\text{撮影者の状態}}$$

ここで、 $P_{personality}$ は個人差を表す項である。また、 $P_{memorability}(O_i)$ は対象オブジェクト O_i の持つ記憶されやすさを表す項である。これは Isola らによって提案されている静的シーンでの記憶可能性である memorability を表す [3]。 $P_{seen}(O_i, \mathbf{I})$ は一人称視点映像 \mathbf{I} に関する項である。この項に関連する要素として、 \mathbf{I} における O_i の出現回数、停留時間、出現位置や \mathbf{I} のオプティカルフロー等が

挙げられる。 P_{seen} の各要素について、第 3.1 節に述べる。 $P_{status}(S)$ は時間によって変化する、撮影者の状態を表す項である。 S には撮影者の興味や集中等、 I からは推測し難い要素が含まれるが、現状では人の状態を測定することは状態の複雑さから困難であるため、本稿では常に P_{status} は一定であると仮定している。

3.1 一人称視点映像から得られる情報

$P_{seen}(O_i, I)$ は一人称視点映像から得られる情報に関する項である。本稿では以下の要素が関係すると考えるが、ここに示す以外にも関連する要素が存在する可能性がある。

出現回数、停留時間

よく見たオブジェクトほど記憶に残りやすいと考えられるため、一人称視点映像におけるオブジェクトの出現回数が多いほど、また停留時間が長いほど記憶可能性が高くなると考えられる。

画像中の出現位置

人の視野は大きく中心視と周辺視に分けられる。中心視とは視線方向の中心に位置する部分であり、周辺視と比較して高解像度の映像を受容することができる [5]。中心視で捉えたものは周辺視でとらえたものに比べて詳細に観察できていると考えられるため、覚えられやすいと推測できる。よって一人称視点映像の中心位置が注視点と重なると仮定するとき、撮影された画像の中心に出現したオブジェクトの記憶可能性が高いと考えられる。

オプティカルフロー

オプティカルフローは 1 フレームの間に対象点が撮影画像内でどの程度移動しているかを表す。一人称視点映像ではカメラは頭部の動きに合わせて移動するため、撮影対象が運動を行わない場合、オプティカルフローには頭部運動が現れる。オプティカルフローの絶対値が小さいときは頭部が静止していると判断でき、頭部静止中は一点を注視していると考えられる。従って、オプティカルフローの絶対値が小さいときに一人称視点映像内に現れたオブジェクトは、記憶可能性が高いことが考えられる。

また、撮影者が一定の速さで一定の方向に正面を向いて歩いているとき、オプティカルフローは放射状に広がる。このオプティカルフローの拡大中心を FOE (Focus of Expansion) という。FOE は進行方向の無限遠点となる。FOE は視覚的注意を強く引くことが確かめられているため、記憶可能性を推定する際に有用な要素となると考えられる。

フレーム間の輝度値の差分

オプティカルフローについての議論と同様に、フレーム間の輝度値の差分に頭部運動が現れる。画像内の輝度値の変化が緩やかな場合、フレーム間の頭部の運動が小さいならば差分の値も小さくなる。よって、フレーム間の輝度値

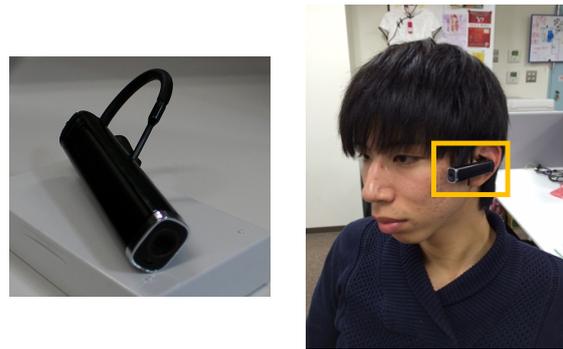


図 1 ウェアラブルカメラ (Looxcie LX2)

Fig. 1 Wearable camera(Looxcie LX2)

の差分が小さいときに見ていたオブジェクトは注視されたと推測されるため、記憶可能性が高くなることが考えられる。

3.2 提案モデルの検証

このモデルを実験を通して検証することを考える。被験者が未知であるシーンに対しては $P_{memorability}$ は一定であるとみなすことができる。また、 $P_{remember}$ 、 $P_{personality}$ はアンケートを用いることで分かる。従って、 P_{seen} について以下の式が成り立つ。

$$P_{seen} = \frac{\overbrace{P_{remember}}^{\text{アンケートから分かる}}}{\underbrace{P_{personality}}_{\text{アンケートから分かる}} \cdot \underbrace{P_{memorability}}_{\text{const.}} \cdot \underbrace{P_{status}}_{\text{const.}}} = f_{(\text{出現回数, 停留時間, 画像中の出現位置, オプティカルフロー})}$$

すなわち、被験者に $P_{memorability}$ が一定であると見なせる環境下で一人称視点映像を撮影しながら特定ルートを散策するタスクを課し、その後アンケートによって特定オブジェクトについての記憶の有無を確認することで、一人称視点映像と P_{seen} との関係が得られ、モデルの妥当性を検証できる。

4. 実験

4.1 予備実験

4人の被験者に、一人称視点映像を撮影するためのウェアラブルカメラ (Looxcie LX2, 図 1) を装着した状態で大学構内の指定したルートを通り 15 分程度の散策を行うタスクを課し、その後アンケートによって特定オブジェクトについての記憶の有無を確認した。指定したルートを図 2 に橙矢印で示す。ルートは大学構内の被験者が普段訪れない場所を指定した。アンケートの内容は、そのルートから視認できるオブジェクトの画像 1 枚と、そのオブジェクトとよく似た、ルート上からは視認できない画像 2 枚の計 3 枚の画像をランダムに並べ、その中から散策中に見たと思うもの 1 枚を選択する形式とし、全 10 題出題した。アン



図 2 実験ルート。予備実験では橙矢印のみ、実験では加えて青矢印の部分もルートに指定している。

Fig. 2 the pathway of experiments. The orange route is used for the preliminary experiment, and orange and blue routes are used additionally for the main experiment.



図 3 予備実験アンケート質問例。図中の丸印は正解画像を表す。

Fig. 3 Examples of the questionnaire for the preliminary experiment. Red circles indicate correct answers.

アンケートの用いた質問の例を図 3 に示す。撮影終了後、一人称視点映像から特定オブジェクトの出現回数、停留時間の情報を取り出し、アンケート結果との関係を調査した。

4.1.1 実験結果

アンケート結果を表 1 に示す。質問番号 2 については、すべての被験者が同じ不正解画像を選択した。また、質問番号 4 については、誤答した 3 人の被験者は全員同じ画像を選択した。我々はこの結果から、三者択一のアンケートでは不正解画像の影響により、オブジェクトの記憶の正確な測定が困難な場合があると考えた。そこで、画像の比較が行えない方式のアンケートを用いて実験を行った。この実験について第 4.2 節で述べる。

表 1 予備実験アンケート結果。背景色青が正答、赤が誤答。

Table 1 Result of the preliminary experiment. Correct answers have blue background, and incorrect answers have red background.

質問番号	1	2	3	4	5	6	7	8	9	10
正答	3	1	1	2	3	2	3	1	2	3
被験者1	3	2	1	3	3	2	3	2	2	3
被験者2	3	2	1	3	3	3	2	1	3	1
被験者3	3	2	1	3	2	2	3	3	3	2
被験者4	3	2	1	2	3	1	2	3	2	1

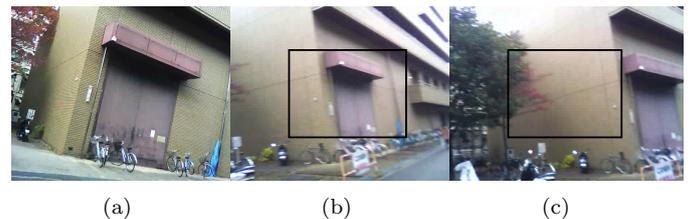


図 4 (a) アンケートに用いたオブジェクトの画像 (b), (c) 一人称視点映像の画像。(b) では対象オブジェクトが中心部分に、(c) では周辺部分に出現。

Fig. 4 (a) Example of the questionnaire.

(b), (c) Pictures taken with wearable camera. The target object appears (b) in center, (c) in margin.

4.2 実験

予備実験で行ったものと同様のタスクを 5 人の被験者に課し、異なる方式のアンケートを実施した。画像を一枚ずつ提示してアンケートを行うことで他のアンケート画像から受ける影響を減らし、対象オブジェクトに対する記憶の有無をより正確に調査することを目的とした。ただし、散策のルートを一部変更した。アンケートは、特定オブジェクトが撮影された画像を一枚ずつ提示し記憶しているか否かを回答するものとし、提示画像としてルート上から視認できるオブジェクト（以下、視認可能オブジェクト群と呼ぶ）の画像 47 枚とルート上からは視認することのできないオブジェクト（以下、視認不可能オブジェクト群と呼ぶ）の画像 47 枚を用意した。アンケートに用いたオブジェクトの画像の一例を図 4(a) に示す。撮影終了後、一人称視点映像から特定オブジェクトの出現回数、停留時間の情報を取り出し、アンケート結果との関係を調査した。

4.2.1 アンケートによる記憶可能性調査

視認可能オブジェクト群の画像に対して、一人称視点映像に 1 回以上現れたオブジェクトの内記憶していると回答した割合 (true positive rate), 視認不可能オブジェクト群の画像に対して記憶していると回答した割合 (false positive rate) を表 2 に示す。この表から、false positive rate の平均値と比較して true positive rate の平均値が明らかに高くなっていることがわかる。また、被験者 C, D について

表 2 実験アンケート結果

Table 2 Result of the main experiment.

	true positive rate	false positive rate
被験者 A	24/42(57.1%)	5/47(10.6%)
被験者 B	31/44(70.5%)	5/47(10.6%)
被験者 C	10/28(35.7%)	2/47(4.3%)
被験者 D	18/37(48.6%)	5/47(10.6%)
被験者 E	23/39(59.0%)	14/47(29.8%)
平均	54.2%	13.2%

は true positive rate がチャンスレベルである 50.0%を下回っている。被験者 E については false positive rate が他の被験者と比較して高い、すなわち実際には見ていないにも関わらず見たと回答した割合が高くなっている。以上より、被験者 C, D, E についてはアンケート結果の信頼性が低いと判断できる。この結果、回答の信頼性の高い被験者 A, B のデータを解析に用いる。

4.2.2 一人称視点映像の解析

$P_{memorability}$ が一定であると仮定すると、アンケート結果は P_{seen} の項のみに依存していると考えられる。 P_{seen} を推定するために一人称視点映像からどのような情報を取り出し、特徴量として用いることが適当であるかを実験によって得られた結果より検討する。

(a) 出現回数・停留時間

図 5, 6 にアンケート結果と出現回数、停留時間との関係を示す。x 軸は出現回数、y 軸は停留時間であり、o はアンケートで記憶していると回答したもの、x は記憶していないと回答したものである。点のプロットは、全 47 個の視認可能オブジェクト群について行った。

図 5, 6 から、一人称視点映像における出現回数が多く、停留時間の長いオブジェクトの記憶可能性が高いことが推測される。これはすなわち、何度も見たもの、長く見たものが覚えられやすいことを表しており、我々の仮説と一致する結果である。

(b) オブジェクトの出現位置

出現位置として、一人称視点映像の上下左右から、それぞれ 1/4 の範囲を除いた部分（以下、中心部分）、その他の部分の 2 種類のみを考える。一人称視点映像の画像の例を、図 4(b), (c) に示す。被験者 A のアンケートについて、一人称視点映像全体に対する出現回数、停留時間を基に行った推定と比較して、中心部分への出現回数、停留時間を基に行った推定の方がより高い精度で推定を行うことができた。しかし、被験者 B からはそのような結果は得られなかった。この理由として、中心視の範囲と今回用いた一人称視点映像の中心部分が一致しなかったことが考えられる。実際、頭部方向と視線方向は必ずしも一致しないため、注視点は一人称視点映像の中心位置とは一致しない。より正確に注視対象のオブジェクトを判断するためには、

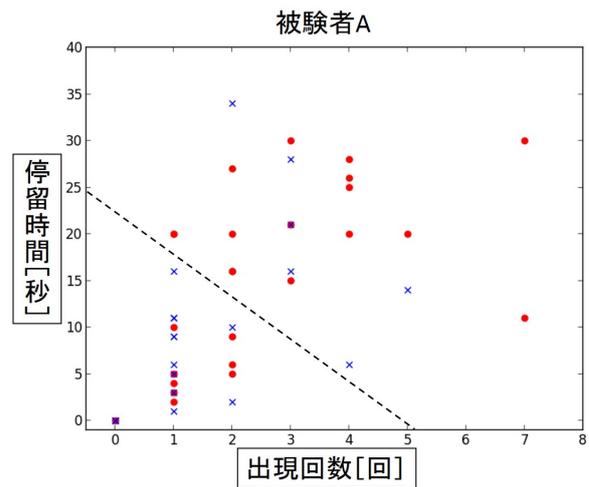


図 5 被験者 A. x 軸は出現回数、y 軸は停留時間であり、o はアンケートで記憶していると回答したもの、x は記憶していないと回答したもの。

Fig. 5 Subject A. This figure is plotted with retention time as the vertical axis and number of occurrences as the horizontal axis. “o” indicate the objects which subject answered that he watched, “x” indicate the objects which subject answered that he did not watch.

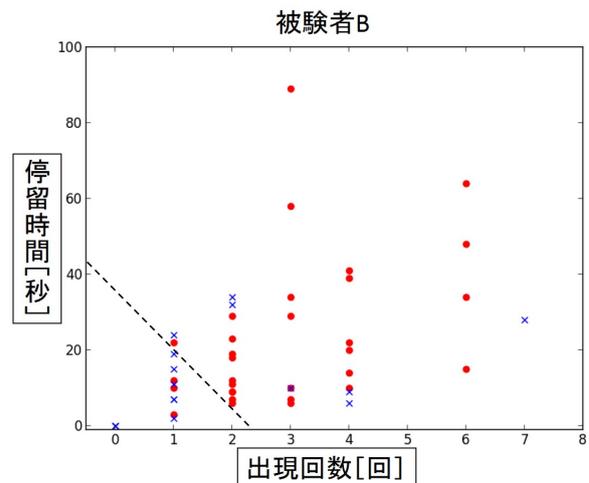


図 6 被験者 B. x 軸は出現回数、y 軸は停留時間であり、o はアンケートで記憶していると回答したもの、x は記憶していないと回答したもの。

Fig. 6 Subject B. This figure is plotted with retention time as the vertical axis and number of occurrences as the horizontal axis. “o” indicate the objects which subject answered that he watched, “x” indicate the objects which subject answered that he did not watch.

注視点推定を用いたアプローチをとることが考えられる。

(c) オプティカルフロー・フレーム間の輝度値の差分

頭部の運動を抽出するための指標として、実験により得られた一人称視点映像からオプティカルフローを計算した。一人称視点映像にオプティカルフローを重ねて描画した画像を図 7(a) に示す。また、x 軸を一人称視点映像開始



図 7 オプティカルフロー
Fig. 7 Optical flow

からのフレーム数, y 軸をオプティカルフローの絶対値の平均としたグラフに, そのフレームに現れていたオブジェクトを記憶していた割合を重ねて描画したものの一部を図 8 に示す.

第 3.1 節ではオプティカルフローの絶対値が小さい場合には頭部運動が少なく, すなわちあるオブジェクトに注視しておりそのオブジェクトの記憶可能性が高いと推定されると考察した. しかし, 図 8 を見るとオプティカルフローの絶対値が小さい場合であっても正答率が高い部分と低い部分があることがわかる.

また, 頭部の運動を検出するための指標としてフレーム間の輝度値の差分を用いることを考える. 各ピクセル毎にフレーム画像の輝度値の差を求め, その差の絶対値の総和を計算する. この値が小さい場合にはあまり頭部が動いていないと判断する. フレーム間の輝度値の差分についても, オプティカルフローと同様にその絶対値が低い場合であっても正答率が高い部分と低い部分があることがわかった.

そこで, 次にオプティカルフローの絶対値の変動から, オブジェクトを注視するために行った頭部運動による変動のみを抽出することを考える. ローパスフィルタをかけた絶対値の時間変化の中で極小値をとる点で興味を持ったオブジェクトの注視を行ったと考え, その時点で一人称視点映像に映っているオブジェクトの記憶可能性が高くなるという仮説を立てた. 1Hz のローパスフィルタをかけたグラフの極小値をとる時点と一人称視点映像を見比べたとき, 我々が注視を行ったと判断した時点は概ね検出された. しかし, 我々が注視を行っていないと判断するような点, 例えば頭部をほとんど運動させずに足を踏み出した時点などについても, 極小値をとることが確認できた. フレーム間の輝度値の差分についてもローパスフィルタをかけたグラフの極小値と一人称視点映像を比較したが, 同様の結果が得られた.

4.2.3 一人称視点映像の特徴とアンケート結果の関係

図 5 及び図 6 より, 被験者 A, B に関して, 出現回数が

多いほど, また停留時間が長いほど記憶可能性が高いことが分かる. そこで, このデータに対して SVM を用いた 2 クラス分類を行った. 得られた境界面を図 5, 6 に点線で示す. 被験者 A の (b) について, この SVM を用いることで正解率 70.2%, 適合率 73.9%, 再現率 68.0% の精度で分類を行うことができた. 被験者 B の (a) については正解率 78.7%, 適合率 80.0%, 再現率 90.3% の精度で分類を行うことができた.

5. 結論

一人称視点映像から対象オブジェクトの記憶可能性を推定するモデルを構築した. 本稿ではこの内 P_{seen} 項を一人称視点映像を基に推測することを試みた. P_{seen} 項を推測するため, オブジェクトの一人称視点映像における出現回数, 停留時間, 出現位置の他, 一人称視点映像のオプティカルフローやフレーム間の輝度値の差分について, その扱いを検討した. その結果, 対象オブジェクトの一人称視点映像における出現回数, 停留時間は記憶可能性の推定に有用であることが示唆された. 一方, オブジェクトの出現位置や, 一人称視点映像のオプティカルフロー, フレーム間の輝度値の差分については, どのように記憶可能性推定に用いることが適切であるか結論付けることができなかった. 出現回数や停留時間について, 追実験を行い今回考察された記憶可能性との関係が普遍的なものである検討を行うことを今後の課題とする. また, 同時に出現位置をどのように利用するか, オプティカルフローやフレーム間の輝度値の差分からどのように記憶可能性の推定に有用と思われる頭部運動を抽出するかについて, さらに検討を行う. 今回考慮していない要因についても, 記憶可能性との関係を調査することを今後の課題とする.

記憶可能性を推定する手法が確立されると, 実世界上のどこが記憶されやすいのかを知ることができるようになる. この地図からは, 多くの人に記憶してもらうためにはどこに情報を掲示すればよいのか等がわかるため, まちづくりやマーケティングにおいて有用である. また, 一人称視点映像を基に個人の記憶可能性を推定することも可能になる. これにより個人の記憶に基づいたサービスを行うことができ, 例えばより精度の高いサジェストを行えるようになることや, ロボットが人と会話を行う際に相手の記憶を考慮して話題を選ぶことが可能になることなどが考えられる.

参考文献

- [1] Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B. and Wood, K.: The use of a wearable camera, SenseCam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis: A preliminary report, *Neuropsychological Rehabilitation*, Vol. 17, No. 4-5, pp.

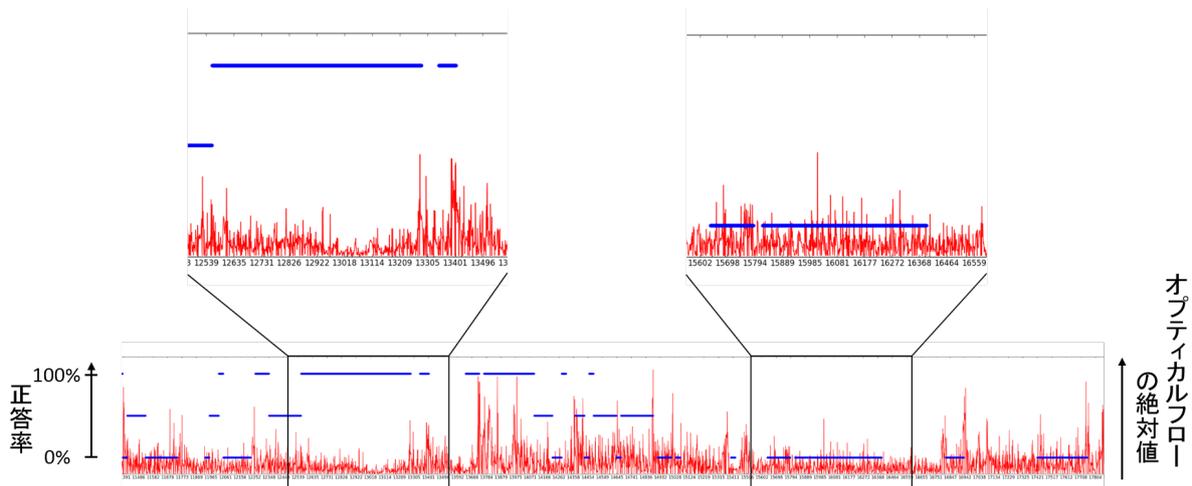


図 8 オプティカルフローの大きさの時間変動 (赤線) と正答率 (青線).

x 軸は一人称視点映像の開始からのフレーム数

Fig. 8 Time fluctuation of the size of optical flow (red) and the rate of correct answer. This figure is drawn with the number of frame from start of the first-person video as the vertical axis.

- 582–601 (2007).
- [2] Brady, T. F., Konkle, T., Alvarez, G. A. and Oliva, A.: Visual long-term memory has a massive storage capacity for object details, *Proceedings of the National Academy of Sciences*, Vol. 105, No. 38, pp. 14325–14329 (2008).
 - [3] Isola, P., Xiao, J., Torralba, A. and Oliva, A.: What makes an image memorable?, *Computer Vision and Pattern Recognition (CVPR)*, IEEE, pp. 145–152 (2011).
 - [4] Konkle, T., Brady, T. F., Alvarez, G. A. and Oliva, A.: Scene memory is more detailed than you think the role of categories in visual long-term memory, *Psychological Science*, Vol. 21, No. 11, pp. 1551–1556 (2010).
 - [5] Turvey, M. T.: On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli., *Psychological review*, Vol. 80, No. 1, p. 1 (1973).
 - [6] 山田健太郎, 菅野裕介, 岡部孝弘, 佐藤洋一, 杉本晃宏, 開一夫: 一人称視点における視覚的顕著性マップモデルの性能評価, 電子情報通信学会技術研究報告. HIP, ヒューマン情報処理, Vol. 110, No. 422, pp. 81–86 (2011).
 - [7] 松本創, 岩田伸一郎, 古賀利郎: 7180 中心市街地の景観画像における注視行動特性と色彩分布に関する研究 (夜間景観, 都市計画), 学術講演梗概集. F-1, 都市計画, 建築経済・住宅問題, Vol. 2009, pp. 429–430 (2009).
 - [8] 相澤清晴, 石島健一郎, 椎名誠: ウェアラブル映像の構造化と要約: 個人の主観を考慮した要約生成の試み, 電子情報通信学会論文誌 D, Vol. 86, No. 6, pp. 807–815 (2003).
 - [9] 谷岡誠一, 横田史郎: 洪水ハザードマップの認知と理解の向上を目指して, 平成 16 年度河川情報シンポジウム講演集, pp. 40–47 (2004).
 - [10] 樋口未来, 木谷クリス真実, 佐藤洋一: 一人称視点映像を用いた頭部動作からの社会的地位の推定, *MIRU* (2014).