

位置情報と地理情報を用いたウェアラブルカメラ映像のダイジェスト作成

上田 隆正† 天笠 俊之† 吉川 正俊 †‡ 植村 俊亮†

† 奈良先端科学技術大学院大学 情報科学研究科
‡ 国立情報学研究所 ソフトウェア研究系

† 〒 630-0101 奈良県生駒市高山町 8916-5
‡ 〒 101-8430 東京都千代田区一ツ橋 2-1-2

{takama-u, amagasa, yosikawa, uemura}@is.aist-nara.ac.jp

あらまし 本稿では、人間が身体に装着したビデオカメラ（ウェアラブルカメラ）を用いて撮影された映像に対して、地理オブジェクトに着目してダイジェストを作成する手法を提案する。利用者はウェアラブルカメラに加えてGPSやジャイロセンサなどを装着し、ビデオ映像の他に撮影位置やカメラの向きなどの情報を取得する。これらの情報とあらかじめ用意した地理データを基に、撮影位置の近くに存在する地理オブジェクトやカメラが地理オブジェクトを向いているかどうかを調べ、その距離や向きに応じた重みを地理オブジェクトに付与する。これにより、映像中における地理オブジェクトの重要度を算出する。この重要度を基に映像の要約を行い、ダイジェストを作成し利用者に呈示する。

キーワード 映像データベース, ウェアラブルカメラ, GPS, ジャイロセンサ, ダイジェスト

Digest Creation of Video Data recorded by Wearable Cameras using Locational and Geographical Information

Takamasa Ueda†, Toshiyuki Amagasa†, Masatoshi Yoshikawa †‡, and Shunsuke Uemura†

† Graduate School of Information Science, Nara Institute of Science and Technology
‡ Software Research Division, National Institute of Informatics

†8916-5 Takayama, Ikoma, Nara 630-0101, Japan
‡ 2-1-2 Hitotsubashi, Chiyoda, Tokyo 104-8430, Japan

{takama-u, amagasa, yosikawa, uemura}@is.aist-nara.ac.jp

Abstract This paper proposes a method for digest creation of video data recorded by wearable cameras, which enable us to make digests focused on geographical objects. A user brings a wearable camera, a GPS, and a gyro-compass, and records changes of locations and angles of the camera in addition to video data. When making a digest, we give importance factors to every geographical object according to the distance and angle between the object and the user. As a result, we can create digests based on the importance factors.

Key words Video Database, Wearable Camera, GPS, Gyro-Compass, Digests

1 はじめに

近年、ハードウェア技術の目覚ましい発展を背景に、身に付けて使用することを前提としたコンピュータ（ウェアラブルコンピュータ）が開発され、その研究も盛んに行われるようになった[1, 2]. ウェアラブルコンピュータでは、情報の表示はHMD (Head Mounted Display)を通して行われることが多く、利用者の位置や状況に応じた情報をリアルタイムに取得、処理し、HMDに表示されることで日常生活をより快適にするという応用が期待されている。

身に付けて使用する機器の一つに、ウェアラブルカメラと呼ばれる超小型のデジタルビデオカメラがある[8]. ウェアラブルカメラを眼鏡の縁などに装着して使用することで、利用者は「撮影する」という意識なしに連続して視点から見た風景を撮影することができる。また、磁気ディスクの小型化と大容量化が進み、数年後には超小型で数TBの容量を持つ製品が実用化されると予測されている。このような環境下では、利用者の行動を四六時中映像として記録することが可能となる。

このようにして蓄積された映像には、利用者の過去の行動がすべて記録されている。例えば、旅行に行ったときの風景を後からもう一度見たり、駅前で友人とすれ違ったことにその時は気付いてなくても後で確認することなどが可能となる。しかし、ウェアラブルカメラ映像のような膨大な情報に対し、参照したいシーンを逐次的に確認して探すことは大変な労力と時間を要する。そこで、シーン検索や、その日の出来事の代表的なシーンを自動抽出するようなシステムが重要になる。

前者に関して、我々は、ウェアラブルカメラで撮影された映像の索引として地理情報を利用する手法を提案した[10]. この手法では、GPS (Global Positioning System)を用いて得られた撮影位置の情報を基に地名やランドマーク名といった地理情報を取得し、映像データの索引として利用する。これにより、利用者は地理情報を用いた映像検索を行うことができる。

後者はいわゆる映像の要約である。長時間に渡る映像でも、全てが利用者にとって重要というわけではなく、利用者にとって重要な部分とそうでない部分が存在する。このため、その重要な区間を抜き出して利用者に表示する。本稿では、映像の要約を実現するための手法として、利用者が長く滞在した場所に着目する。すなわち、ある地理オブジェクトの近くに長時間滞在している場合、その地理オブジェクトは利用者にとって重要であると判定するのである。地理オブジェクトを映像の索引として付与する際に、利用者との距離を反映した重みを付与する。また、カメラが地理オブジェクトの方向を向いているかどうかとも一つの指標とし、それを重みとして索引に付与する。これらの重みと、地理オブジェクトの近くに滞在した時間を考慮して地理オブジェクトの重要度を定義し、

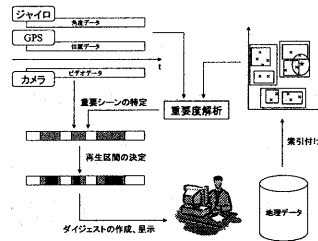


図1: 処理の流れ

ダイジェスト作成の際のシーン選択の判定に利用する。

具体的には、利用者はウェアラブルコンピュータとウェアラブルカメラに加えてGPSとジャイロセンサを常時携帯する。GPSでは撮影時の位置情報（緯度、経度）を、ジャイロセンサではカメラの向きの情報を記録する。このように取得された情報は、利用者が一日の行動を終えてからホストコンピュータに転送される。これらの情報とあらかじめ用意された地理データから、システムは単位時間ごとに利用者の近くに存在した地理オブジェクトを計算し、ビデオデータに索引として付与する。その際に、利用者との距離を基にした重みと、カメラの向きを考慮した重みを算出し、索引に加える。ダイジェストを作成する際には、単位時間ごとの重みの総和をその地理オブジェクトの重要度とみなし、重要と思われるシーンを決定する。

距離に関する重みは、利用者との地理オブジェクトの距離が近いほど値が大きくなるように設定する。また、カメラの向きに関する重みは、地理オブジェクトの方向を向いているときほど値が大きくなるように設定する。このような指標を用いることで、地理オブジェクトに主眼を置いたウェアラブル映像の要約が実現でき、ダイジェストの作成が可能となる。

本稿の構成は以下の通りである。まず2章で提案手法である地理オブジェクトの重要度解析とダイジェストの作成方法について述べる。3章で実験と考察、4章で関連研究を述べ、5章で結論と今後の課題について述べる。

2 提案システムの概要

本章では、提案システムの概要について述べる。全体の処理の流れを図1に示す。

2.1 想定する環境

利用者は、ウェアラブルコンピュータを常時携帯する。ウェアラブルコンピュータには、ウェアラブルカメラ、GPS、ジャイロセンサが接続されており、ウェアラブルカメラでは視点から見える風景の映像を、GPSでは屋外での位置情報を、ジャイロセンサではカメラ動作の角速

度情報を取得する。GPSを用いるため、利用者は屋外を移動するという状況を想定しているが、将来的にはGPS以外の手段を用いて屋内でも位置情報を取得できる環境が整うことが予想される。このため、本手法は将来的には屋内でも利用可能となることが期待できる。

地理オブジェクトの重要度を算出する際のシステムの処理手順を以下に示す。

1. 利用者が記録した映像、撮影位置、カメラの角度の情報を取り込む。
2. 単位時間ごとに、撮影位置の近くに存在する地理オブジェクトを特定し、映像へ索引付けを行う。その際に、利用者と地理オブジェクトの距離を基にした重みと、カメラの向きと地理オブジェクトの位置を考慮した重みを索引に付与する。
3. 地理オブジェクトごとの重要度を算出する。

利用者の要求によりダイジェストを作成する際には、以下のように処理が行われる。

1. 解析した重要度のランキングから、ダイジェストに含める地理オブジェクトを決定する。
2. 採用された地理オブジェクトが撮影された映像から実際に再生する映像区間を決定する。
3. ダイジェストを作成し、利用者に呈示する。

ダイジェストの再生時間や重要度に関するパラメータは利用者がシステムと対話的に決定し、要求に合わせたダイジェストを作成することができる。

2.2 地理データ

本研究では地理情報として、地理オブジェクトの緯度、経度を利用する。具体的には、国土交通省から無償で提供されている地理情報の数値データを利用する[7]。なお、本研究では地理情報の中でも地名や広い面積を持つ地理オブジェクト（奈良公園など）は対象から除外した。これは本手法がこのようなオブジェクトに対応していないためであり、これらへの対応は今後の課題である。地理データに対して高速なアクセスを実現するために、R-Tree[3]を用いて索引付けを行う。

2.3 位置情報とカメラの向きを利用した地理オブジェクトの重要度算出

2.3.1 地理オブジェクトと利用者の距離に基づく重み付け

ここでは、利用者が近くを通った地理オブジェクトに対してダイジェスト作成時に用いられる重要度を与える方法について述べる。ある地理オブジェクトの近くに長く滞在したことは、一日の行動の中でその地理オブジェクトが重要であったと考えられる。そこで、ある時刻における利用者の位置と地理オブジェクトとのユークリッド

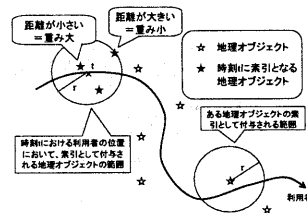


図 2: 地理オブジェクトと利用者のそれぞれから見た「近い」範囲

ド距離を算出し、その距離が短い程値が大きくなるように重み付けする。

ある時刻 t における利用者の座標を (x_t, y_t) 、ある地理オブジェクト O_k の座標を (X_k, Y_k) としたとき、これらのユークリッド距離 $d(P_t, O_k)$ は次のように算出される。

$$d(P_t, O_k) = \sqrt{(X_k - x_t)^2 - (Y_k - y_t)^2}$$

$d(P_t, O_k)$ がある一定の値 r より小さい場合、すなわち、 $d(P_t, O_k) < r$ を満たす場合に、映像にその地理オブジェクト名を索引として重みを付けて付与する(図2)。ただし、求められた距離をそのまま重みとするのではなく、あるオブジェクトと利用者の位置が近いときに、より大きな値となるようにしたい。そこで、上式の逆数を重みとして付与する。よって、時刻 t における、地理オブジェクト O_k の距離に関する重み $w_{distance}(P_t, O_k)$ を次のように定義する。

$$w_{distance}(P_t, O_k) = \frac{1}{\sqrt{(X_k - x_t)^2 - (Y_k - y_t)^2}}$$

ここで、 $d(P_t, O_k) < r$ を常に満たす時間を $s_i \leq t \leq e_i$ (ただし、 s_i, e_i はそれぞれ、利用者の位置がある地理オブジェクトに近いと判断される区間の最初の時刻と最後の時刻で、 $s_i \leq e_i$ を満たす)としたとき、ある地理オブジェクト O_k の重要度 $W_{distance}(O_k)$ を次のように定義する。

$$W_{distance}(O_k) = \sum_{t=s_i}^{e_i} w_{distance}(P_t, O_k)$$

なお、 $d(P_t, O_k) < r$ を満たす区間が複数箇所存在する場合は、それぞれに対し上式を計算し、その和を重要度とする。すなわち、

$$W_{distance}(O_k) = \sum_{i=1}^n \sum_{t=s_i}^{e_i} w_{distance}(P_t, O_k)$$

で表される。 n は $d(P_t, O_k) < r$ を満たす区間の数である。 $n=2$ の場合の例を図3に示す。

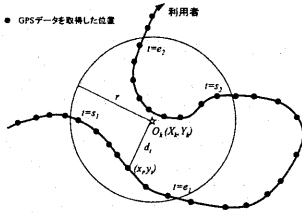


図3: 近い範囲と判断される区間が複数存在する場合

2.3.2 カメラの向きに対する重み

ある地理オブジェクトの近くに長く滞在したといっても、その地理オブジェクトに注目していなければ、それは利用者にとってはあまり重要でないと考えられる。そこで、カメラがどこを向いているのかを計測し、地理オブジェクトの方角を向いているとき、重要度が高くなるような重みを導入する。

時刻 t におけるカメラの向きを α_t 、利用者からみた見た地理オブジェクト O_k の角度を $\theta(P_t, O_k)$ とする。カメラの向きと地理オブジェクトの関係を図4に示す。このとき、 $\theta(P_t, O_k)$ は次式で表される。

$$\theta = \tan^{-1} \frac{Y_k - Y_t}{X_k - X_t}$$

重みを付ける範囲を $\theta(P_t, O_k)$ を中心に 2β とし、 α_t が $\theta(P_t, O_k)$ に近い程大きな重みが付くようにする。すなわち、 $-\beta < \alpha_t - \theta(P_t, O_k) < \beta$ を満たす範囲に、正規分布 $N(0, \sigma^2)$ にしたがるように重み $w_{direction}(P_t, O_k)$ を付与する。このとき、 $0 < w_{direction}(P_t, O_k) \leq 1$ となるように調節し、時刻 t における、地理オブジェクト O_k のカメラの向きに関する重み $w_{direction}(P_t, O_k)$ を次のように定義する。

$$w_{direction}(P_t, O_k) = \frac{1}{1-C} \times \exp\left(-\frac{(\alpha_t - \theta(P_t, O_k))^2}{2\sigma^2}\right) - \frac{C}{1-C}$$

$$\text{ただし } C = \exp\left(-\frac{\beta^2}{2\sigma^2}\right)$$

なお、 $|\alpha_t - \theta(P_t, O_k)| \geq \beta$ のとき、 $w_{direction}(P_t, O_k) = 0$ とする。

ここで、 $d(P_t, O_k) < r$ を常に満たす時間を $s_i \leq t \leq e_i$ としたとき、ある地理オブジェクト O_k の重要度 $W_{direction}(O_k)$ を次のように定義する。

$$W_{direction}(O_k) = \sum_{t=s_i}^{e_i} w_{direction}(P_t, O_k)$$

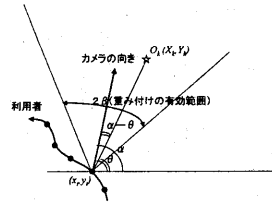


図4: カメラの向きと地理オブジェクトの関係

なお、 $d(P_t, O_k) < r$ を満たす区間が n 箇所存在する場合は、それぞれに対し上式を計算し、その和を重要度とする。すなわち、

$$W_{direction}(O_k) = \sum_{i=1}^n \sum_{t=s_i}^{e_i} w_{direction}(P_t, O_k)$$

と表される。

2.3.3 重みの統合

地理オブジェクトと利用者の距離に基づく重み $w_{distance}$ と、カメラの向きと地理オブジェクトの方向に基づく重み $w_{direction}$ の2つを反映した重要度を提案する。ある地理オブジェクト O_k に対して、ある時刻 t における両者を統合した重み $w(P_t, O_k)$ を次式で与える。

$$w(P_t, O_k) = z w_{distance}(P_t, O_k) \times (1-z) w_{direction}(P_t, O_k)$$

z ($0 \leq z \leq 1$) はどちらの重みを重視するかを決定するパラメータで、利用者が指定するものである。

ここで、 $d(P_t, O_k) < r$ を常に満たす区間を $s_i \leq t \leq e_i$ としたとき、ある地理オブジェクト O_k の重要度 $W(O_k)$ を次のように定義する。

$$W(O_k) = \sum_{t=s_i}^{e_i} w(P_t, O_k)$$

なお、 $d(P_t, O_k) < r$ を満たす区間が n 箇所存在する場合は、それぞれに対し上式を計算し、その和を重要度とする。すなわち、

$$W(O_k) = \sum_{i=1}^n \sum_{t=s_i}^{e_i} w(P_t, O_k)$$

と表される。

2.3.4 地理オブジェクトの高さの考慮

地理オブジェクトの中には、東京タワーのように高い建物も存在する。このような地理オブジェクトは、立ち寄ること以外に、遠くから眺めることも重要であり、そのような場面もダイジェストに反映させることを考える。そこで、高い地理オブジェクトに対しては、ある程度離れた場所でもそちらの方角を向いていれば重要度に反映させることが考えられる。

2.4 ダイジェストの作成

ここでは、前節で得られた地理オブジェクトの重要度を基に、ダイジェストを作成する手法について説明する。基本的には、重要度のランキングの上位から順に地理オブジェクトを選び、それに対応する映像区間からダイジェストを作成する。処理手順を以下に示す。

1. 利用者がダイジェストの再生時間、再生シーン数を指定する。
2. 地理オブジェクトの重要度ランキングの上位から再生シーンの数だけ取り出し、利用者に提示する。利用者はその中から自分の興味から外れたものをシステムにフィードバックすると、システムはそれらを再生シーンの候補から外す。
3. 指定された再生シーン数になるまでランキング上位のシーンを再生候補として採用する。
4. 候補となる地理オブジェクトが含まれる映像をデータベース検索する間を特定する。
5. 4で得た部分映像から、さらにダイジェストとして再生する区間を絞る。
6. 以上で得られた映像を時系列順に利用者に呈示する。

2.4.1 ダイジェストの再生時間と再生シーン数

ダイジェストの総時間は利用者が指定するが、再生シーンの数と各シーンの再生時間の与え方には次のような手法が考えられる。

- 利用者が各シーンの再生時間を指定。総時間と各シーンの再生時間が決まれば、再生シーンの数は自ずと決定される。
- 利用者が再生シーンの数を指定。各シーンの再生時間が同じだとすれば、総時間と再生シーンの数が決まれば、各シーンの再生時間は自ずと決定される。
- 利用者が再生シーンの数を指定。ランキング上位から指定された数だけ再生のシーンとし、その重要度の比から各シーンの再生時間を決定する。

利用者は自分の好みに応じてこれらの中から作成方式を選択することができる。

2.4.2 再生シーンの特定

一日の行動において、ある地理オブジェクトの近くに滞在した時間というのは長時間に渡る場合があり、時には数時間に及ぶことも考えられる。このように、重要だと判断された地理オブジェクトに関わっている時間すべてをダイジェストに盛り込むことは現実的ではない。そこで、重要シーンの中からダイジェストとして再生するシーンを特定する必要がある。

一般に、ある地理オブジェクトに近付いていく場面やその場所で何かをしている場面は重要であるが、離れていく場面は重要ではないことが多い。すなわち、再生すべき映像区間は、最も重みが高いと判断される時刻を基準として、それより後には少なく、前に多く存在すると考えられる。このようなことを考慮して、再生シーンの特定方法として次のような手法を提案する。

- ある時間 $s_i \leq t \leq e_i$ において、時刻 $t = t'$ における地理オブジェクト O_k の重み $w(P_i, O_k)$ がその区間で最大であったとき、時刻 t' を基準に再生区間を決定するが、あらかじめ t' より前と後の再生時刻の比を決めておき、再生時間から再生区間を算出する。例えば、 t' の前と後を 5:1 の割合で再生すると決めた場合、再生時間が 1 分ならば、 $t = t' - 50$ 秒から $t = t' + 10$ 秒のシーンが再生される。

また、ある地理オブジェクトが複数区間に索引付けされている場合は、それぞれの区間ごとに地理オブジェクトの重要度を計算し、先に計算した重要度ランキングと比較する。その結果、上位にランクされた場合は全てダイジェストに取り込まれる。なお、その再生区間の計算方法は前述の通りである。

2.4.3 重要でない地理オブジェクトの排除

重要度のランキングを行っても、利用者にとって重要でない地理オブジェクトが上位にきてしまうことも考えられる。そこで、このような問題に対処するため、以下の機能をシステムに実装する。

- システムと利用者が対話的にダイジェストを作成する。すなわち、ダイジェストを作成する前に地理オブジェクトのランキング付けの結果を利用者に提示する。利用者は、重要でないと思う地理オブジェクトをシステムに通知し、システムはそれらを除いたダイジェストを作成する。
- 地理データにはもともと利用者にとって重要でない地理オブジェクトが多く存在していると考えられる。また、時と場合に応じて利用者の興味の対象も変化する。そこで、地理オブジェクトごとに重みを設けることを提案する。一つ一つの地理オブジェクトに対して重みを設定することは現実的

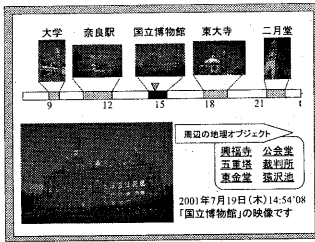


図 5: 利用者インターフェースの様式図

ではないので、あらかじめ地理オブジェクトをカテゴリ分けしておき、利用者が興味のあるカテゴリを選択しておけば、そのカテゴリに属する地理オブジェクトの重みをあらかじめ大きくできるようにすることができる。

2.5 利用者インターフェース

ダイジェストを呈示するには、利用者の利便性を考えた見やすいインターフェースを提供する。以下に示す機能を実装する。

- 再生している映像の場所、時刻の表示
- 再生している映像の近くにある地理オブジェクトの表示とそのリンク
- ダイジェストとして再生される各シーンのサムネイルと場所、時刻を時系列にまとめたサマリーの表示
- 時刻バーから見たい場所の直接指定

インターフェースのイメージを図5に示す。画面にはダイジェスト映像が大きく再生され、その横に映像の地理オブジェクト名とその再生時刻が記されている。画面上部にはダイジェストとして再生される各区間のサムネイルが表示されている。それ以外の場所の映像を見たいときには、その下の時刻バーをクリックすることにより、任意の時刻の映像を再生することができる。また、周辺の地理オブジェクト名をクリックすると、その地理オブジェクトの映像が再生される。

3 提案手法の評価

提案手法のうち、地理オブジェクトと利用者の距離に基づく重みを用いた重要度の解析に関する実験を行った。

3.1 実験の概要

本実験の概要は以下の通りである。

1. ウェアラブルカメラとGPSを装着して屋外を移動し、視点映像とその撮影位置を記録する。

2. 移動したルートを記録する
3. 撮影終了後、あらかじめ用意した地理データを基に地理オブジェクトの重要度解析を行い、そのランキング付けを行う
4. 得られたランキングと移動したルートの比較を行う

ウェアラブルカメラはSony社製のカラービデオカメラ「まめカム」“CCD-MC100”を、GPSはSony社製のハンディGPSレシーバ“PCQ-GPS3SUG”を使用した。PCQ-GPS3SUGにおける測定誤差は±10m程度である。GPSデータ取得のサンプリング周期は1秒にして実験を行った。

地理データは、国土交通省がWeb上で無料配布している奈良県の公共施設の数値データを利用した。これは、公共施設の緯度、経度や住所、種類などが納められたデータで、記述されている地理オブジェクトの数は5020である。しかし、このデータには寺社仏閣や商店などは記述されていないほか、データが最新のものではないといった問題がある。そこで、記述されていない地理オブジェクトの情報に関しては、ゼンリン社の“電子地図帳Z III”[11]から手動で重要と思われる地理オブジェクトを追加した。緯度、経度は0.1秒単位までを使用した。緯度、経度はそのままでは距離が計算できないので、変換式を用いて整数値に変換し、利用者と地理オブジェクトの位置を求めた。東経を a_e 度 b_e 分 c_e 秒、北緯を a_n 度 b_n 分 c_n 秒とすると、変換後の座標 (e', n') は次式で求まる。

$$e' = (a_e \times 3600 + b_e \times 60 + c_e) \times 10$$

$$n' = (a_n \times 3600 + b_n \times 60 + c_n) \times 10$$

本実験では、 $r=30$ として重要度の解析を行った。これにより、利用者の位置から半径約90m以内に存在する地理オブジェクトが索引として付与される。

3.2 結果と考察

前節で述べた環境下で、次に示す2種類の移動経路に関して実験を行った。

1. 大学の近くのスーパーでの買物
2. 奈良公園での観光

3.2.1 実験1: スーパーでの買物

駐車場に停めてある車で近くのスーパー3軒(コスト、ライフ、サカエ)に立ち寄り、また駐車場に帰ってくるというルートである。このデータに対し、重要度解析を行った結果を表1に示す。

結果からわかる通り、意図した通りの重要度を得ることができた。ライフの方がコストよりも長時間いたにもかかわらず、重要度の値はコストの方が大きくなった。これは、ライフの登録している緯度、経度と、実際にGPSで取得された利用者の緯度、経度が少し離れていたため、

表 1: 索引付けされた地理オブジェクトと重要度 (実験 1)

順位	地理オブジェクト名	重要度	時間 [秒]
1	サカエ	92.295	1100
2	コスト	44.661	467
3	ライフ	32.662	795
4	学生宿舎	14.884	144
5	北大和第 5 緑地	5.372	118
6	ガスト	2.363	38
7	北大和第 2 公園	2.358	51
8	北大和第 6 公園	1.510	28
9	AM/PM	1.261	11
10	北大和第 4 緑地	1.030	11
11	北大和第 3 緑地	0.723	10
12	北大和第 2 緑地	0.504	12
13	コーナン	0.498	11
14	中登美ヶ丘 2 丁目第 2 号児童公園	0.241	7
15	登美ヶ丘高校	0.237	6
16	北大和第 6 緑地	0.229	6
17	生駒市公民館鹿畑南中分館	0.102	3
18	奈良先端大	0.068	2

単位時間における距離の重みが小さくなったことが原因であると考えられる。

3.2.2 実験 2: 奈良公園での観光

奈良観光を想定して、奈良公園周辺を散策してデータを収集し、重要度の解析を行った。移動したルートと、どの時間にどこにいたかをまとめたものを表 2 に示す。なお、表中の開始時刻、終了時刻は地理オブジェクトと利用者の実際の距離に関係なく、その地理オブジェクトを主題として意識していた時間を示している。例えば、五重塔に近付いてきたと思った時刻を開始時刻とし、五重塔から離れようとした時刻を終了時刻としている。

このデータに対し、重要度解析を行った結果を表 3 に示す。また、この結果を基にした再現率-適合率グラフを図 6 に示す。考察を以下に述べる。

- 上位 6 位までは意図した地理オブジェクトがランキングされている。これは、提案システムが有効であることを示している。
- 「二月堂」、「鐘楼」は下位にランキングされた。これは、GPS がデータ取得できなかった時間があったため、本来ならもっと上位になったものと考えられる。正確なデータを得るためには、GPS 以外の機器による位置情報の取得を行う必要がある。例えば、PHS を用いることや、ジャイロセンサと加速度計を用いて欠損した位置情報を補間することなどが考えられる。
- 「四月堂」、「二月堂下一貫堂」、「法華堂」、「三味堂」など、二月堂周辺に密集している地理オブジェクトには立ち寄っていないにも関わらず高い重要度

表 2: 奈良観光のルート (実験 2)

開始時刻	終了時刻	時間 [秒]	滞り場所
17:38:08	17:43:04	296	駐車場
17:46:42	17:50:20	212	興福寺五重塔
17:50:40	17:52:50	130	興福寺東金堂
17:55:15	17:57:10	105	南円堂
18:09:00	18:14:00	300	奈良国立博物館
18:16:50	18:19:30	160	氷室神社
18:27:20	18:28:30	110	東大寺南大門
18:31:00	18:34:22	202	東大寺中門
18:37:40	18:40:38	178	鐘楼
18:43:10	18:50:00	410	二月堂

表 3: 索引付けされた地理オブジェクトと重要度 (実験 2)

順位	地理オブジェクト名	重要度	時間 [秒]
1	東大寺中門	73.310	252
2	若草モータープール (駐車場)	57.232	380
3	奈良国立博物館	51.439	464
4	南円堂	38.995	625
5	興福寺五重塔	38.277	623
6	興福寺東金堂	37.514	602
7	猿沢池派出所	35.347	415
8	氷室神社	28.915	396
9	四月堂	25.094	452
10	二月堂下一貫堂	22.316	182
11	法華堂	20.999	282
12	三味堂	20.772	298
13	高円山 OL パーマネントコース	19.333	346
14	奈良市東寺林連絡所	16.150	320
15	金珠院	15.869	284
16	二月堂	15.802	187
17	念仏堂	14.045	199
18	鐘楼	12.858	191
19	東大寺南大門	11.356	116
20	興福寺国宝館	9.571	151
21	興福寺本坊	8.990	153
22	春日野派出所	6.997	125
23	辛国神社	6.927	129
24	奈良国立博物館新館	0.405	12

がついた。これは、利用者は二月堂に滞在していると認識していても、二月堂の近くにある地理オブジェクトに対しても索引が付与されたためである。単位時間当たりの重みは大きくはないが、二月堂に滞在している時間が長いことそれらの重要度が大きくなったものと考えられる。地理オブジェクトが密集している地域への索引付けは、近傍の閾値 r を小さくすることが必要であると考えられる。

- 「猿沢池派出所」、「高円山 OL パーマネントコース」など、一般には重要ではないと考えられる地理オブジェクトもランキングされた。しかし、このようなことは 2.4.3 節で提案した、対話的なダイジェスト作成にて対応できると考えられる。
- 表 2 と表 3 の時間を比べると、GPS データが欠損してしまった二月堂を除くすべてのオブジェクトで実際に滞在していたと感じた時間より、システ

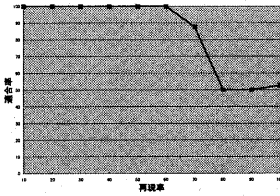


図 6: 実験 2 の再現率-適合率グラフ

ムが索引付けした時間の方が長くなった。これは、地理オブジェクトから離れる際にも、距離が近い場合索引付けされることに起因すると考えられる。このことから、地理オブジェクトから離れているような状況のときには重みを小さくするなどの対応が求められる。

4 関連研究

飯島等 [5] は、映像中に現れる人物像を検出し、それを基にエピソード映像を自動生成する手法を提案している。中村等 [9] は利用者の注目動作を分析し、その映像特徴と画像の動き推定により得られる動きベクトルを用いて日常行動から注目シーンを検出する方法を試みている。Healey 等 [4] は発汗、石島等 [6] は脳波という生体情報をセンサーで捉え、それらを利用して映像を自動的に構造化する手法を提案している。これらは、それぞれのアプローチにより効果的な映像の構造化を実現しているが、位置情報を用いたものではない。これらのような手法で構造化された映像から重要度を算出し、ダイジェストに盛り込むことも有用であると考えられる。これは今後の課題である。

5 まとめと今後の課題

本稿では、ウェアラブルカメラにより得られる映像に対し、地理オブジェクトに着目したダイジェストを作成する手法を提案した。提案システムでは、ダイジェスト作成のために映像中の重要シーンを特定するために、GPSにより得られる位置情報と地理オブジェクトの距離やカメラの向きからオブジェクトの重要度を解析する手法を提案した。また、実験では、複数の経路上においてウェアラブルカメラにより得られた映像に対し、実際に距離に基づく地理オブジェクトの重要度解析を行った。その結果、地理オブジェクトと利用者の距離に基づく重要度解析を行うことにより、有効なダイジェスト作成が可能であることを確認した。

今後の課題としては

- 地理オブジェクトの影響範囲の動的決定法の考案

- 地理オブジェクト以外の指標に基づくシーンの重要度解析とダイジェストへの反映
- 粒度の低い地理オブジェクトへの対応
- 全方位カメラを用いたダイジェストの作成
- ダイジェスト作成システムの実装

が挙げられる。

謝辞

本研究の一部は、文部科学省科研費基盤研究(11480088, 12680417, 12208032) および科学技術振興事業団 戦略的基礎研究推進事業の支援を受けている。ここに記して謝意を表す。

参考文献

- [1] M. Billingham, S. Bee, J. Bowskill, and H. Kato : "Asymmetries in collaborative wearable interfaces", Proc. ISWC, pp. 133-140, 1999.
- [2] Bradley J., Rhodes, Nelson Minar, and Josh Weaver : "Wearable Computing Meets Ubiquitous Computing: Reaping the best of both worlds", Proc. ISWC, pp. 141-149, 1999.
- [3] Guttman A. : "R-Tree: A Dynamic Index Structure for Spatial Searching", Proc. SIGMOD, pp.45-57, 1984.
- [4] Jennifer Healey and Rosalind Picard : "SturtleCam: A Cybernetic Wearable Camera", Proc. ISWC, 1998.
- [5] 飯島俊匡, 石上陽一, 川嶋稔夫, 青木由直 : "日常生活映像から検出された人物像によるエピソード想起", PRMU, 1998
- [6] 石島健一郎, 相澤清晴 : "ウェアラブルによる長時間個人体験記録の編集 -脳波を利用した映像の自動編集の試み-", PRMU, pp.85-92, 2001.
- [7] 国土交通省, 国土数値情報ダウンロードサービス : <http://nlftp.mlit.go.jp/ksj/>
- [8] Mayol, W. W., Tordoff, B., and D.W. : "Wearable visual robots", Proc. ISWC, pp. 95-102, 2000.
- [9] Yuichi Nakamura, Jun'ya Ohde, and Yuichi Ohta : "Structuring personal experiences -Analyzing views from a head-mounted camera", Proc. IEEE, 2000.
- [10] 上田隆正, 天笠俊之, 吉川正俊, 植村俊亮 : "位置情報と時刻情報を利用した映像データの索引付け手法", 電子情報通信学会第 12 回データ工学ワークショップ, 2001
- [11] ゼンリン, 電子地図 "ゼンリン電子地図帳 Z III" : http://www.zenrin.co.jp/prod/03_01.html