

歩行者ナビゲーションを目的とした経路映像の画像解析

樽見 佑亮[†] 亀田 能成[†] 大田 友一[†]

[†] 筑波大学 大学院システム情報工学研究科, 茨城県つくば市天王台 1 丁目 1-1

E-mail: [†]s1320844@u.tsukuba.ac.jp, ^{††}{kameda,ohta}@iit.tsukuba.ac.jp

あらまし 我々は、歩行者ナビゲーションを目的とした経路上での事前撮影映像の映像解析の研究を行っている。我々の考えている歩行者ナビゲーションとは特定の経路を歩行者に辿ってもらうことである。数分程度の歩行で移動可能な長さで、かつ屋内屋外であることを制限しない経路を想定する。実現方法としては事前に経路に沿って撮影したカメラ位置にカメラレジストレーションを利用して歩行者のカメラを誘導する。亀田らによってカメラのみを利用して経路上での位置推定を行う手法が提案されている。この手法は画像シーン全体の局所特徴量を利用した類似画像検索を用いることで歩行者の経路上での位置推定を行う。この手法では経路上での位置推定として事前撮影した映像をデータベース化し、歩行者が撮影したスナップショットを問い合わせ画像とした類似画像検索を利用する。本稿では事前撮影映像と問い合わせ画像との間でその撮り方に差があるとき、その差が類似画像検索の性能にどのような影響があるかを調べ、その結果を報告する。

キーワード 局所特徴量, 画像検索, 位置推定, 歩行者ナビゲーション

Computer Vision Based Analysis of Route Video for Pedestrian Navigation

Yusuke TARUMI[†], Yoshinari KAMEDA[†], and Yuichi OHTA[†]

[†] Graduate School of Systems and Information Engineering, University of Tsukuba,
1-1-1 Tenoudai, Tsukuba, Ibaraki, Japan

E-mail: [†]s1320844@u.tsukuba.ac.jp, ^{††}{kameda,ohta}@iit.tsukuba.ac.jp

Abstract We are working on walker navigation by using a camera only. Location is estimated by taking a snapshot on route. We estimate location of the snapshot by retrieving pre-recorded route video. Before performing the location estimation of the snapshot, we make the database of image key-features in the pre-recoded route video. The most similar video frame is selected by counting the number of key-pairs between the snapshot and every frame of the pre-recoded route video, and the frame that gives the largest number of key-pairs indicates the location of the snapshot. In this paper, we have taken videos in various camera motions along a navigation path and we confirmed the performance of image retrieval of our method.

Key words local feature, photo retrieval, position estimation, pedestrian navigation

1. ま え が き

我々は、特定の経路を移動しながら事前撮影した一人称映像をもとに経路上で撮影したスナップショットを利用した歩行者ナビゲーションの研究を行っている。我々の想定している歩行者ナビゲーションとは特定の経路を歩行者に辿ってもらうことである。具体的には事前に経路に沿って撮影したカメラ位置にカメラレジストレーションを利用して歩行者のカメラを誘導する。我々の想定している経路は、人物が数分程度で移動可能な長さで、かつ屋内屋外であることを制限しないものである。具体

例を挙げると駅から店舗までの移動や大学の教室から教室の移動における経路である。

亀田ら [1] [2] によってカメラのみを利用して経路上での位置推定を行う手法が提案されている。この手法では画像シーン全体で見つかる局所特徴量を利用した類似画像検索を用いることで歩行者の経路上での位置推定を行っている。以後、特別に断りがない限り本研究で扱う歩行者の位置とは経路上での位置を表すものとする。

亀田らの手法では位置推定として事前撮影した映像をデータベース化し、歩行者が撮影したスナップショットを問い合わせ

画像とした類似画像検索を行う。事前撮影映像が経路に対して十分望ましい間隔で撮影されていれば、経路上における高精度な位置推定が可能である。

我々はこれまで画像検索を用いた位置推定を行う際の日照等による影響を報告してきた [3] [4]。本稿では歩行者が事前撮影映像中でのカメラの動きと異なる動作をしながら撮影した問い合わせ画像について検索精度がどのように変化するかを検証した。本稿で取り扱う歩行者の動作としては経路上を蛇行しながら撮影する動作と経路上で周りを見回す動作を取り上げる。

2. 関連研究

GPS を利用した歩行者ナビゲーションはすでに広く普及しており、商業的なサービスとして一般に利用されている。しかし、GPS の性質上利用できない環境も多く、そのような環境での GPS 以外のセンサを利用した位置推定やナビゲーション手法について多くの研究が進められている。

電波以外の情報を手がかりにする研究としては、例えば歩行者の位置推定を目的とした研究として Kourogi ら [5] [6] は加速度変化や地磁気変動を利用して歩行者の位置推定を行っている。

一方、カメラを用いたナビゲーションや定位の研究としてロボットの分野では Morales [7] らの取り組みのように自己位置推定と環境地図作成を行う SLAM の研究が進められている。

この他に画像を利用した粗い位置推定の手法としては Hays ら [8] の提案する手法が挙げられる。この手法では予め画像に位置情報を付加して記憶した大量の画像を地球規模で用意する。用意した画像と位置推定を行いたい位置で撮影した画像とのマッチングを取ることで、大まかな位置推定を実現する。

定位の精度を高めた取り組みとしては、例えば池田ら [9] は事前に GPS の位置情報と全方位カメラ映像を元に Structure-from-Motion によって作成したランドマークデータベースを用意し、このランドマークデータベースと現在撮影した映像との対応をとることで現在位置の推定を行っている。

しかし、これらのアプローチでは屋内外でのシームレスな運用を想定していないため、我々の想定する経路上での位置推定には適さないと考えられる。

亀田ら [1] [2] は事前に経路に沿って一人称で撮影した事前撮影映像があることを前提にして、歩行者がその場で撮影した画像と事前撮影映像とを局所特徴量を用いて比較、検索することで歩行経路上での位置を推定できることを示している。

亀田らの手法では経路上での位置を推定するために画像検索を利用している。現在、画像検索や一般物体認識の分野では局所特徴量やそれを利用した研究が多くなされている [10] [11] [12]。本研究はシーン全体に散らばる keypoint を利用して 2 つのシーンが類似しているかを検証する。

3. 経路に沿った事前撮影映像を用いた位置推定手法の概要

本節では本研究で用いる位置推定手法の概要を述べる。

本研究での歩行者ナビゲーションとは予め定められた経路を歩行者がその経路を辿るように誘導するものとする。本研究

では位置推定手法として亀田ら [1] [2] の手法を用いる。

本研究ではナビゲーションの事前準備として経路上に沿って位置を記録しながら事前撮影映像を撮影する。事前撮影映像は google ストリートビューなどに代表されるような経路上で全周囲を撮影した映像ではなく、経路を移動しながら歩行者の一人称視点と近くなるように撮影した映像である。このように撮影した映像をフレーム毎に分割し、事前撮影画像群を作成する。以後、事前撮影画像を $\{R_i\}$ と表記する。映像が n フレームあるとし、各フレームを以後 R_i と表記する ($i = 1, \dots, n$)。次に歩行者が経路上でスナップショットを撮影する。このスナップショットを以後問い合わせ画像 Q と表記する。 Q と $\{R_i\}$ との間で類似画像検索を行うことで Q の撮影位置を推定する。

続いて、位置推定時に用いる類似画像検索について説明する。まず各 R_i から局所特徴量を抽出してデータベースを作成する。局所特徴量として画像中から複数の特徴的な keypoint を検出する。keypoint の属性として座標、方向、輝度勾配などの情報を多次元ベクトルやバイナリ値で記述する。 R_i から m_i 個の keypoint が検出されたとして、1 つの keypoint を K_{jR_i} と表記する ($j = 1, \dots, m_i$)。そして、 K_{jR_i} に記述されている座標、方向、輝度勾配情報をそれぞれ $k_{jR_i}.pt$, $k_{jR_i}.orient$, $k_{jR_i}.des$ と表記する。

続いて、 Q から局所特徴量を抽出する。このとき Q から l 個の keypoint が検出されたとして、1 つの keypoint を K_{hQ} と表記する ($h = 1, \dots, l$)。 K_{hQ} の座標、方向、輝度勾配情報をそれぞれ $K_{hQ}.pt$, $K_{hQ}.orient$, $K_{hQ}.des$ と表記する。

ここで 2 つの特徴量間の相違度を出力する関数 D を定義する。 D は式 (1) の通り 2 つの特徴量を入力として相違度 d を出力するものである。

$$d = D(K_{hQ}.des, K_{jR_i}.des) \quad (1)$$

次に Q 中に含まれるある $K_{hQ}.des$ に対して最も相違度 d の小さい $k_{jR_i}.des$ を有する keypoint K_{jR_i} を探索する関数を定義する。

$$kmatch_h = \underset{i,j}{\operatorname{argmin}} (D(K_{hQ}.des, K_{jR_i}.des)) \quad (2)$$

この $kmatch_h$ は $K_{hQ}.des$ と d が最も小さい $k_{jR_i}.des$ を持つ K_{jR_i} の j, i を $kmatch_h.j$, $kmatch_h.i$ という形で持っている。また、この時の K_{hQ} と K_{jR_i} のペアを key ペアと呼称する。式 (2) で求めた $kmatch_h$ を利用して、 R_i に含まれる key ペアの数 $kmatchnumber_i$ を与える式 (3) を定義する。

$$kmatchnumber_i = \sum_{h=1}^l \begin{cases} 1 & \text{if } dmatch_h.i = i \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

最後に $kmatchnumber_i$ が最大であるフレーム $\hat{i}$ を式 (4) から得る。

$$\hat{i} = \underset{i}{\operatorname{argmax}} kmatchnumber_i \quad (4)$$

R_i が Q に対する類似画像検索の結果となる． Q と R_i の画像のペアを検索画像ペアと呼称する．

4. 位置推定精度の検証

本節では本研究で用いている位置推定手法に対して行った精度検証の方法及びその結果に対する考察について述べる．

4.1 検証方法

実際に本研究が歩行者ナビゲーションを行おうと想定しているシーンでの経路を複数用意し，そこで撮影したシーケンスを用いて実際に画像検索を行いその精度を検証した．いずれも映像は Panasonic 製の DMC-3D1 の単眼撮影モードを用いて 30fps の mp4 で取得した．画像サイズは 320x240 とした．

今回の検証では歩行者が $\{R_i\}$ 撮影時とは異なる動作をしながら問い合わせ画像となるスナップショット画像を撮影した場合にどの程度検索精度に影響を与えるかを検証した．本稿で取り上げる動作は以下の二つである．

- 経路上を蛇行動作しながらのスナップショット撮影（動作 1）
- 経路上の分岐点で周りを見回し動作をしながらスナップショット撮影（動作 2）

また今回の実験で用いる経路として以下の 2 つを用意した．

- 筑波大学の第三エリアの F 棟前から A 棟内に入り中央図書館方面に移動する経路（経路 A）．
- 筑波大学の大学会館エリアの大学会館前バス停から大学会館内売店まで移動する経路（経路 B）．

これらの経路は経路上に屋内外を含み，2 分程度で移動可能な長さである．各経路の開始地点，中間地点，終了地点近くの画像を図 1,2 に示す．

今回の実験では予め進行方向正面にカメラを保持しながら，経路から大きく外れることがないように注意して撮影した映像を $\{R_i\}$ として用意する．このとき先頭フレームから順に $R_0, R_1, \dots$ とする．この $\{R_i\}$ に対して経路上で動作 1，2 を行いながら撮影した映像の各フレームを問い合わせ画像 Q として画像検索を行い検索精度について検証した．

4.1.1 通常歩行でのスナップショット撮影

本節では動作 1，2 との比較のために $\{R_i\}$ と同一の経路で $\{R_i\}$ 撮影時と同じく通常の動作でほぼ同じ速度で歩行しながら映像を撮影し，この映像の各フレーム s をスナップショット Q_s として $\{R_i\}$ との画像検索の結果を検証する．画像検索に用いる二つの映像は同一経路上で同程度の速度で撮影された映像といえる．そのため位置推定を行う上で十分な精度を有する検索精度を持つなら Q_s と類似した $\{R_i\}$ 中の画像のフレーム R_i について $s \cong i$ となることが期待される．

検索結果を図 3 から 6 に示す．図 3,4 で横軸は問い合わせ画像 Q_s のフレーム番号 s ，縦軸は $\{R_i\}$ から検索されたフレーム番号 i である．図 3,4 のグラフでは $s \cong i$ となる直線上に検索結果が集中していることが分かる．この結果は先の仮定と一致することから，十分な検索精度を有していると考えられる．



図 1 経路 A の一部

また，図 3,4 とを比較すると図 3 では $s \cong i$ とならない検索結果が散見される．これは経路 A は経路 B と比較して類似した見た目となる長い廊下がシーケンス中に多く含まれることが影響したと考えられる．

図 3,4 に対して画像検索時のマッチ数 $kmatchnumber_i$ を追加したものを図 5,6 に示す．図 5 を見ると $s \cong i$ 近傍にある検索結果は高いマッチ数を持ち，それ以外のところはほとんどの場合においてマッチ数が小さくなっていることからマッチ数を利用した閾値処理によって検索精度の向上が見込めると考えられる．

4.1.2 蛇行動作しながらのスナップショット撮影

経路上で進行方向に対して左右に 2 m 程度蛇行しながら撮影した映像の各フレームを問い合わせ画像 Q_s として用いる．4.1.1 節と同じ $\{R_i\}$ に対して画像検索を行った結果を図 7 から 10 に示す．図 7,8 は図 3,4 と比較して検索結果が傾きの小さい直線関係になっていることが分かる．これはスナップショットが蛇行移動しながら撮影をしているため，単位時間あたりの経路上での移動距離が遅くなっていることが原因である．図 7,8 は図 3,4 と比較して誤検索が増加しているものの直線上に検索結果が集中することから蛇行していても類似画像を検索できていることが分かる．また図 9,10 を見ると図 5,6 と同様に類似画像を検索できている場合はマッチ数が高い値を持っていること



図 2 経路 B の一部

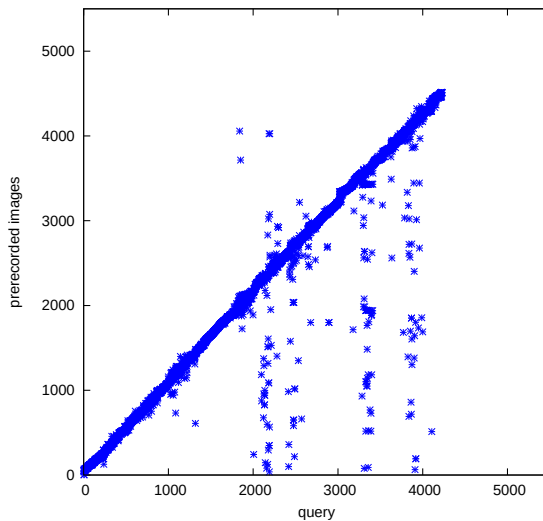


図 3 通常歩行動作 (経路 A)

が分かる。このことからスナップショットを撮影した位置姿勢が異なる場合でもマッチ数による閾値処理によって誤検索を低減できると考えられる。

4.1.3 分岐点で周りを見回し動作をしながらスナップショット撮影

歩行者が経路上の分岐点に差し掛かった際に周りを見回しな

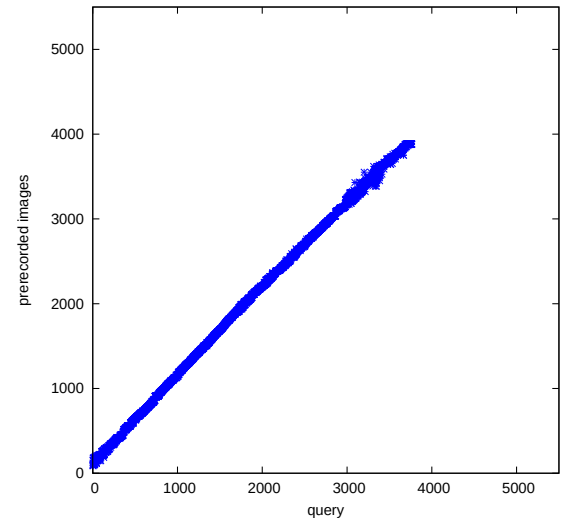


図 4 通常歩行動作 (経路 B)

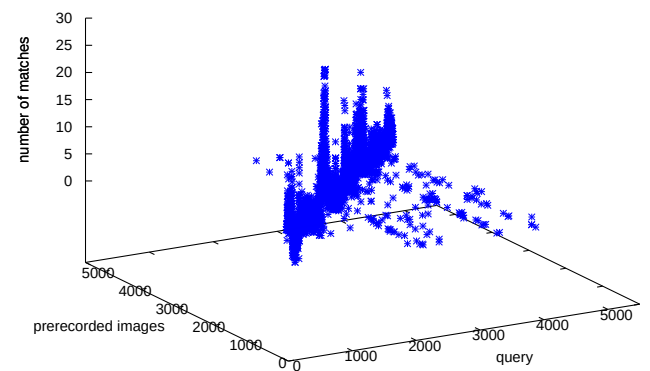


図 5 通常歩行動作におけるマッチ数 (経路 A)

がら撮影した映像の各フレームを問い合わせ画像 Q_i として用いる。4.1.1 節と同じ $\{R_i\}$ に対して画像検索を行った結果を図 11 から 14 に示す。経路 A (図 11) で見回し動作をした 1500, 2100, 3800, 4000 フレーム付近で検索結果にばらつきが多く見られる。また経路 B (図 12) で見回し動作をした 1000, 2500, 3500 フレーム付近で検索結果にばらつきが多く見られる。見回し動作を行うと、そこでの問い合わせ画像に対する正解画像が $\{R_i\}$ に含まれないためこのような結果になると考えられる。しかし、それ以外のフレームでは検索結果は直線上に集中していることから見回し動作を行わなかった箇所では類似画像が検索されていることが分かる。また、図 13,14 をから類似画像を検索できている場合はマッチ数が高い値を持っていることが分かる。このことから経路上での分岐点などで進行方向を見失った場合でもマッチ数による閾値処理によって誤検索を低減し、正しい進行方向を提示できると考えられる。

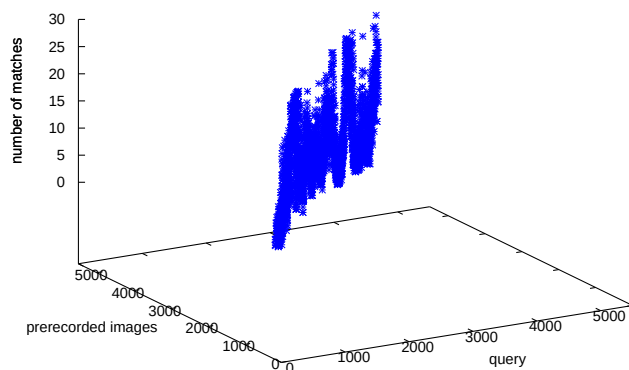


図 6 通常歩行動作におけるマッチ数（経路 B）

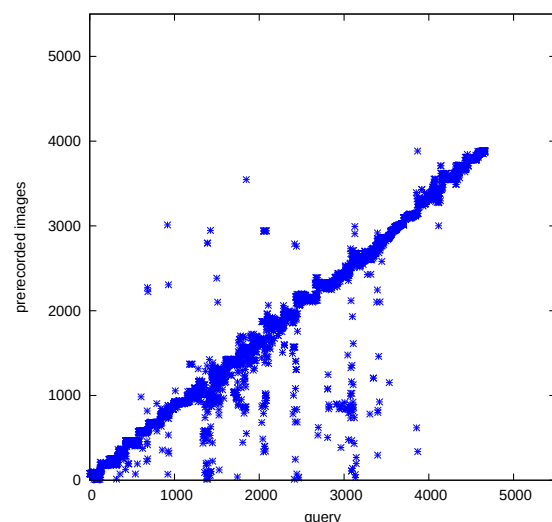


図 8 蛇行動作（経路 B）

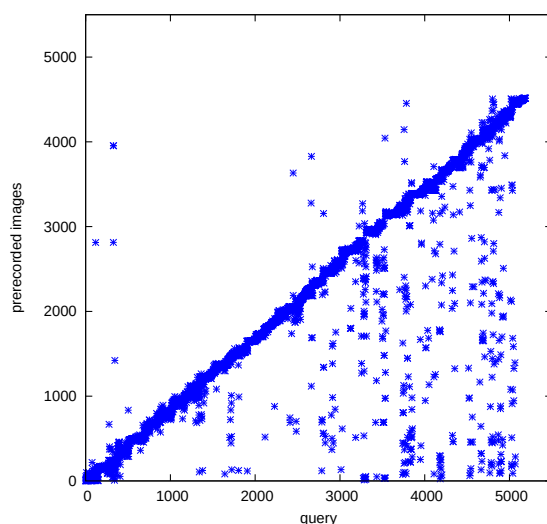


図 7 蛇行動作（経路 A）

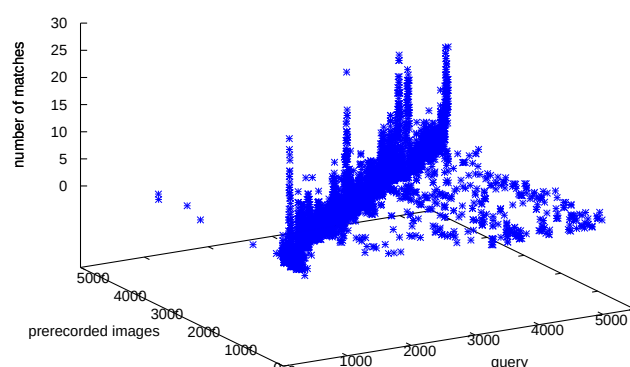


図 9 蛇行動作にマッチ数（経路 A）

5. おわりに

本稿では類似画像検索を用いた経路上での歩行者の位置推定手法において、歩行者が事前撮影映像を撮影したときとは異なる動作で問い合わせ画像を撮影した際にそれがどの程度検索精度に影響を与えるかを検証した。

実験結果から歩行者が経路上を蛇行、見回し動作をしたりしても十分に位置推定が可能であることが分かった。さらに、マッチ数を用いた閾値処理により誤検索の低減も可能であることが分かった。

本研究は科研費 23300064 の助成を受けたものである。

文 献

- [1] 亀田能成, 大田友一, “歩行者視点カメラによる歩行者位置オンライン推定の取り組み,” 信学会 PRMU, vol.110, no.27, pp.67-72, 2010.
- [2] Y. Kameda, and Y. Ohta, “Image Retrieval of First-Person Vision for Pedestrian Navigation in Urban Area,” ICPR, pp.364-367, 2010.
- [3] 樽見佑亮, 亀田能成, 大田友一, “歩行経路上の位置推定のため

の画像検索性能評価,” 信学会 総合大会講演論文集 (情報・システム講演論文集 2), pp. 144, 2013.

- [4] 樽見佑亮, 亀田能成, 大田友一, “歩行経路記録映像における画像検索性能評価,” 第 19 回画像センシングシンポジウム, 2013.
- [5] M. Kourogi and T. Kurata, “Personal Positioning based on Walking Locomotion Analysis with Self-Contained Sensors and a Wearable Camera,” Int. Symposium on Mixed and Augmented Reality, pp.103-112, 2003.
- [6] M. Kourogi, T. Ishikawa, and T. Kurata, “A Method of Pedestrian Dead Reckoning Using Action Recognition,” In Proc. Position Location and Navigation Symposium (PLANS), pp.85-89, 2010.
- [7] Y. Morales, E. Takeuchi and T. Tsubouchi, “Vehicle Localization in Outdoor Woodland Environments with sensor fault detection,” Proc. of IEEE Int. Conf. on Robotics and Automation, pp.449-454, 2008.
- [8] J. Hays and A. A. Efros, “IM2GPS: Estimating Geographic Information from a Single Image,” CVPR, pp.1-8, 2008.
- [9] 池田 聖, 佐藤 智和, 山口 晃一郎, 横矢 直和, “全方移動画像と GPS 位置情報からのランドマークデータベースの構築,” MIRU, pp.1129-1134, 2007.
- [10] J.M. Morel and G.Yu “ASIFT: A New Framework for Fully Affine Invariant Image Comparison,” SLAM Journal on Imaging Sciences, vol. 2, issue 2, 2009.

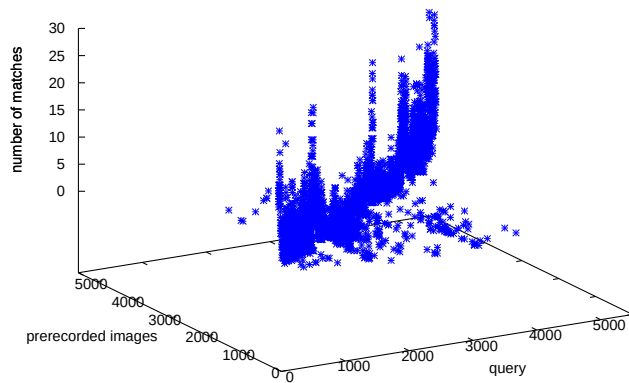


図 10 蛇行動作にマッチ数 (経路 B)

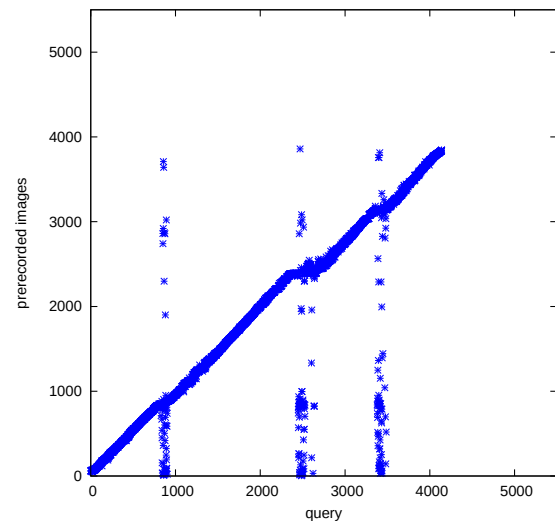


図 12 見回し動作 (経路 B)

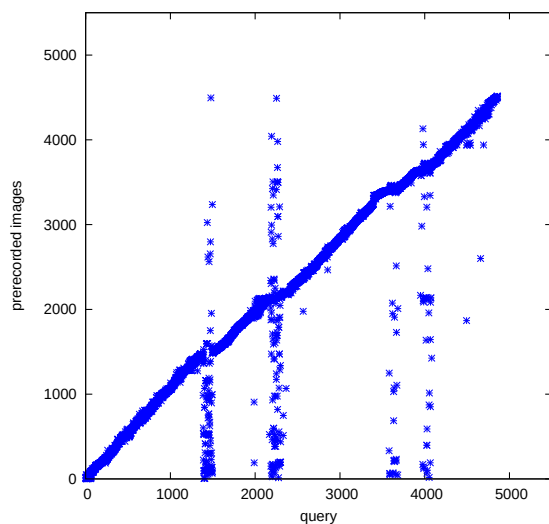


図 11 見回し動作 (経路 A)

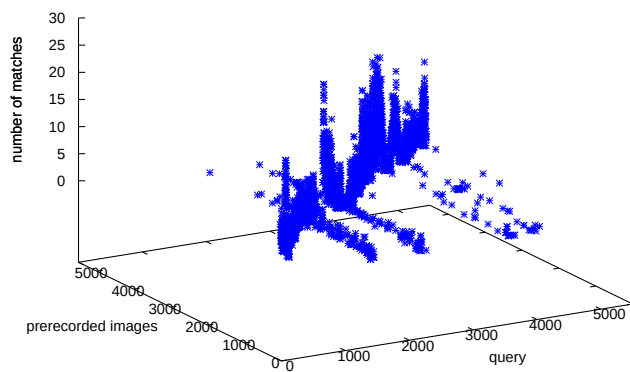


図 13 見回し動作におけるマッチ数 (経路 A)

- [11] Ke Gao, Shouxun Lin, Yongdong Zhang, Sheng Tang, and Huamin Ren “Attention Model Based SIFT Keypoints Filtration for Image Retrieval,” Seventh IEEE/ACIS International Conference on Computer and Information Science, pp191-196, 2008.
- [12] 柳井 啓司 “一般物体認識の現状と今後,” CVIM, vol.48. no. 16. pp.1-4, 2007.

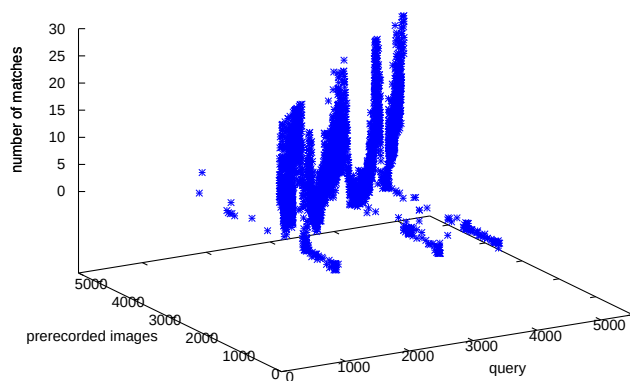


図 14 見回し動作におけるマッチ数 (経路 B)