

# 人物歩行経路の学習と固有空間を用いた予測

Predicting Paths of a Pedestrian by using Eigenspace

山本遼<sup>†</sup>, 玉木徹<sup>‡</sup>, 山本正信<sup>†</sup>

Ryo Yamamoto<sup>†</sup> and Toru Tamaki<sup>‡</sup> and Masanobu Yamamoto<sup>†</sup>

<sup>†</sup>: 新潟大学工学部, yamamott@vision.ie.niigata-u.ac.jp

<sup>‡</sup>: 新潟大学自然科学研究科, tamaki@ie.niigata-u.ac.jp

**概要** 本論文では、学習された人物の移動軌跡をもとに、ある時点までの人物の歩行経路から、それ以降の歩行経路を予測する手法を提案する。まず、あらかじめ固定カメラで撮影した多数の移動経路を取得し、正規化する。次に、それぞれの歩行者の移動経路を各歩行経路に分類し、各経路の固有空間を生成する。そして、新たに人物を追跡する場合には、追跡中にどの経路に対応するのかを決定し、固有空間への射影・逆射影からその人物の将来の移動軌跡の予測を行う。

## 1 まえがき

近年、監視カメラによって撮影された映像が犯罪の早期解決に役立っており、また軽犯罪防止のために商店街などにも防犯カメラを設置する動きも出てきている [1]。このように犯罪の抑止などを目的としているカメラ映像は、高い解像度ではっきりと人物を写していることが望まれる。しかし、複数のカメラ映像が多分割画面としてビデオに録画されている場合も多く、(いわゆる電子ズームによって)映像を拡大しても人物を特定するには不十分なものが多い。

最近では、映像中に動きのある場合にのみ録画を行い、長時間の監視映像の保存を可能にする映像サーバーや、動きのある部分を撮影するようなカメラシステムも登場している [2]。しかし、常に複数の人物が同時に出現する場合には、これらの方法では対応できない。そこで、複数の人物を個別に拡大し撮影することが必要になってくる。同一視野を撮影する複数のカメラを用いることで、複数人物の追跡を行う手法 [3, 4] は研究されているが、これらは拡大撮影することを目的としたものではない。また一台のカメラで通常監視と拡大撮影の両方を行う手法はほとんどなく、これが実現できれば、カメラシステムのコストを抑え、かつ頑健にすることができる。

ここで解決すべき問題は次の二つである。まず一つ目は、ズームされる人物の移動先の予測である。ズームカメラを制御してある人物を拡大する場合、カメラがズームをしている最中にも人物が移動するため、ズーム終了後の人物の行き先をあらかじめ予測しておく必要がある。ズーム中にも追跡を行うことは可能だが、背景が移動することになり、容易ではない。そして二つ目は、ズームされる人物以外の移動先の予測である。ある人物を拡大する場合、ズームを開始する時点で映像に写っている他の人物は、一旦映像から消え去り、拡大終了後にズームが元に戻った時にまた映像に現れる。しかし、これらの人物も既に移動しているため、再び追跡を行うためには以前の位置からどこへ移動しているのかを予測しておく必要がある。

そこで本論文では、あらかじめ人物の歩行経路を固有空間によって学習し、ある時点までの人物の移動軌跡をもとに、それ以降の移動軌跡を予測する手法を提案する。本手法は、撮影場面を固定した時、人々の移動軌跡(画面上での二次元座標の時系列)はそのシーンに依存したいくつかの歩行経路に分けられるということを仮定する。まず経路の学習のために、あらかじめ固定カメラで撮影した映像から多数の歩行者の移動軌跡を取得する。次に、それぞれの歩行者の移動軌跡をクラスタリングによりいくつ

かの経路に分類し、各経路毎に移動軌跡の長さを正規化して、固有空間を生成する。そして、新たに人物を追跡する場合には、その人物の移動軌跡に基づいてどの経路に属するのかを決定し、その固有空間への射影・逆射影からその人物の将来の移動軌跡の予測を行う。

## 2 学習する移動軌跡の分類

まず、学習データとして多数の移動軌跡をテンプレートマッチングなどの追跡手法により取得し、それらを歩行経路に分類する。ある人物  $i$  の、画像上での二次元座標の時系列（移動軌跡）を  $(x_j^i, y_j^i)$  ( $j = 1, \dots, l_i$ ) とする。ここでは移動軌跡が多項式で表せると仮定し、移動軌跡の  $x$  成分と  $y$  成分のそれぞれを（ここでは二次式で）最小二乗法により近似する。この近似多項式の 6 パラメータ ( $x, y$  について各 3) を用いて、 $k$  平均クラスタリングにより移動軌跡をクラスタに分類する。このクラスタを歩行経路と呼ぶ。撮影するシーンに応じて、人々が歩く歩行経路やその数は大まかに決められることができるので、ここではクラスタの数  $K$  は既知であるとする。

## 3 移動軌跡の正規化と歩行経路固有空間の生成

各歩行経路  $k$  ( $k = 1, \dots, K$ ) に属するすべての移動軌跡  $i$  ( $i = 1, \dots, N_k$ ) について正規化を行い、歩行経路を表現する固有空間を生成する。

ここで、移動軌跡  $i$  を構成する  $l_i$  個の二次元座標を次のように長さ  $2l_i$  のベクトル  $x_i$  で表す（図 1）。

$$x_i = [x_1^i, y_1^i, x_2^i, y_2^i, \dots, x_{l_i}^i, y_{l_i}^i]^t \quad (1)$$

移動軌跡はテンプレートマッチングなどを用いた追跡手法によって取得するが、人物によって追跡される軌跡は異なるため、それらを正規化する必要がある。たとえば歩行速度が遅い人物の場合、同じ経路を歩いている足の早い人物に比べて二次元座標の時系列は多くなる。また、追跡の開始と終了場所は一定ではないため、同じ歩行経路を歩いている人物でも、画像上で移動した軌跡の道のりは異なる。

学習データである移動軌跡間のこれらの差異を吸収するために、まず  $x_i$  を線形補間で間引き、長さ

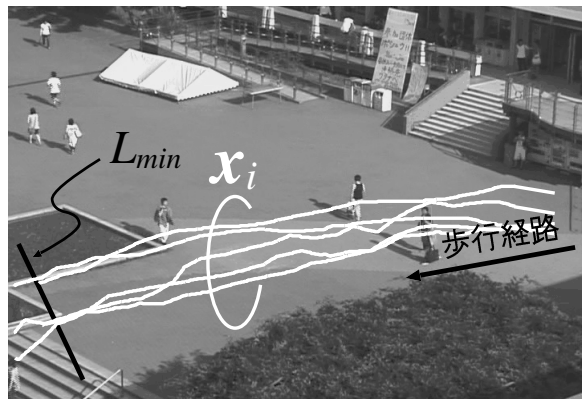


図 1: ある歩行経路の移動軌跡ベクトル  $x_i$  と長さ  $L_{min}$  での正規化。

$m$  ( $< 2l_i$ ) のベクトル  $x'_i$  を作成する。ここで、 $x'_i$  の長さ  $m$  は  $x_i$  の長さ  $2l_i$  よりも十分に小さいとする。

次に  $x'_i$  の画面上での長さ  $L_i$ （つまり各要素  $(x_s^i, y_s^i)$  と隣接する要素  $(x_{s+1}^i, y_{s+1}^i)$  との距離の総和）を計算する。

$$L_i = \sum_{s=1}^{m-1} \sqrt{(x_{s+1}^i - x_s^i)^2 + (y_{s+1}^i - y_s^i)^2} \quad (2)$$

そして  $x'_i$  を線形補間して、長さ  $L_i$  ( $> m$ ) のベクトル  $x''_i$  を求める。元の移動軌跡  $x_i$  の特徴を十分に保ち、かつ小さな  $m$  を選ぶことで、画像上で移動した道のりに比例した長さのベクトル  $x''_i$  を得ることができ、歩行速度の差異は吸収される。

しかし、まだ各移動軌跡間で長さは異なっている。そこで各  $x''_i$  の長さ  $L_i$  を、各歩行経路  $k$  に属するすべての移動軌跡の中でもっとも短い長さ  $L_{min}^k$ 、つまり

$$L_{min}^k = \min_{1 \leq i \leq N_k} L_i \quad (3)$$

にあわせて切りそろえる（図 1）。

そして、各ベクトル  $x''_i$  の  $L_{min}^k + 1$  番目以降の要素を捨て、先頭から  $L_{min}^k$  番目までの要素だけを並べたベクトル  $x'''_i$  を作る。

この正規化された移動軌跡  $x'''_i$  ( $i = 1, 2, \dots, N_k$ ) を用いて、歩行経路  $k$  を表す固有空間  $E_k$  を求める。求めた固有ベクトルを  $e_1^k, \dots, e_{n_k}^k$  とすると、固有空間からの投影を表す行列  $E_k$  は次のように表せる。

$$E_k = [e_1^k, e_2^k, \dots, e_{n_k}^k] \quad (4)$$

ただし、 $n_k$  ( $\leq N_k$ ) は固有空間の次元である。

## 4 固有空間を用いた人物歩行経路予測

学習された固有空間  $E_k$  ( $k = 1, \dots, K$ ) を用いて、ある時点までの人物の移動軌跡から、それ以降の全体の移動軌跡を推定する。

歩行経路が推定されるべき新たな人物  $y$  の移動軌跡  $y$  は、その人物が画面に現れてから消え去るまでの間画像上で追跡することで得られる。しかし、その人物が画像に写っている間は  $y$  の一部分しか観測することができない。したがって、真の移動経路  $y$  を、その一部分  $y'$  から推定（予測）することになる。本研究では、学習した固有空間を元にその予測を行うために BPLP[5](Back Projection for lost pixels) を用いる。BPLP は一部が欠損している画像の復元を行うために提案されたが、本研究の場合は一部が欠損している移動軌跡の復元とみなすことができる。

ある人物  $y$  が追跡されている時、追跡開始後の各時刻で  $y'$  が得られる。まず、前節で学習データに対して行った正規化と同様にして、 $y'$  を間引き、その要素数が画面上での長さ  $L_{y'}$  に一致するように正規化する。

$L_{y'}$  は追跡開始からその時刻までの移動距離であるが、これは（追跡開始後当分の間）歩行経路  $k$  を学習した移動軌跡の中の最も短い移動距離  $L_{min}^k$  よりもさらに短い。つまり  $L_{y'} < L_{min}^k$  であるので、 $y'$  の末尾に  $(L_{min}^k - L_{y'})$  個の 0 を追加し、次式のような長さ  $L_{min}^k$  のベクトル  $y''$  を作る（図 2）。

$$y'' = [y', \overbrace{0, \dots, 0}^{L_{min}^k - L_{y'}}]^t \quad (5)$$

ここで、0 を付加した部分を予測部分、それより前の部分を観測部分とする。またベクトルの長さに対する予測部分の割合  $\frac{L_{min}^k - L_{y'}}{L_{min}^k}$  を欠損率とよぶ。追跡を行うにしたがって人物の移動距離も長くなるため、各時刻で追加する 0 の個数（つまり欠損率）は減少する。

真の移動軌跡  $y$  を用いて  $y''$  を表すと、 $y''$  に付加した 0 を表現する対角行列  $\Sigma$  を用いて次式のように表せる。

$$y'' = \text{diag}[\overbrace{1, \dots, 1}^{L_{y'}}, \overbrace{0, \dots, 0}^{L_{min}^k - L_{y'}}] y \equiv \Sigma y \quad (6)$$

ここで BPLP[5] を用いると、固有空間  $E_k$  を用い

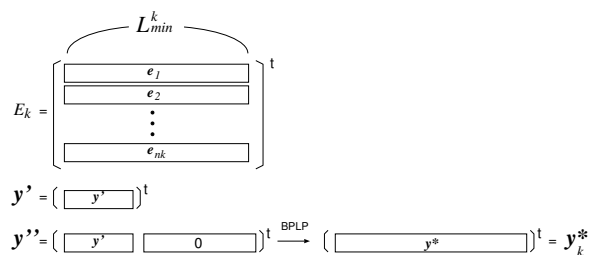


図 2:  $y'_i, y''_i, y^*$  ベクトルの長さ

た移動軌跡の推定値  $y_k^*$  は次式で求められる。

$$y_k^* \cong E_k T^{-1} E_k^t y'', \quad T = E_k^t \Sigma E_k \quad (7)$$

ただし、 $\det(T) \neq 0$  でなければならないので、 $T$  がフルランクである必要がある。これは付加した 0 の割合に依存するが、ここでは、観測されたベクトル  $y'$  の長さ  $L_{y'}$  が固有空間の次元  $n_k$  よりも大きければよい。つまり  $n_k < L_{y'}$  である必要があるが、人物  $y$  の追跡開始直後は、 $L_{y'}$  は非常に小さいため、この制約を満たさない。したがって、追跡開始後しばらく時間が経過して、人物の移動距離  $L_{y'}$  が固有空間の次元を越えた時点から移動軌跡の予測を行う。

## 5 予測誤差と経路選択

各時刻において、移動軌跡の推定値  $y_k^*$  は各歩行経路毎に計算される。ここでは、追跡している人物がどの歩行経路に沿って移動しているのかを判別する。

$y''$  を固有空間  $E_k$  に投影し、さらに逆投影したものが  $y_k^*$  であるため、 $y''$  と  $y_k^*$  との間には差が生じる。ここで  $y''$  の観測部分（つまり  $y'$ ）は真の経路  $y$  の観測部分と等しい。しかし推定値  $y_k^*$  の観測部分は、正しい歩行経路に対応する固有空間で生成された場合には  $y'$  に近い値を持ち、そうでなければ  $y'$  との差が大きくなるはずである。

そこで、観測部分である先頭から  $L_{y'}$  個の要素について、 $y''$  と各  $y_k^*$  との差の二乗和を求める。この予測誤差がもっとも小さい歩行経路、つまり

$$k^* = \underset{k}{\operatorname{argmin}} \frac{\|\Sigma y'' - \Sigma y_k^*\|^2}{\operatorname{trace}(\Sigma)} \quad (8)$$

なる歩行経路  $k^*$  を、人物  $y$  が移動している経路であると選択する。ここで  $\Sigma$  は前節で導入した対角行列であり、その要素数（つまり  $L_{y'}$ ）で正規化している。

## 6 実験

本手法を計算機上で実装し、実画像に対して実験を行った。市販ビデオカメラ (Sony DCR-TRV900, 画像サイズ  $720 \times 480$ ) で実験用映像を撮影し、学習用の人物の移動軌跡データを 163 人分取得した。これらを 12 の歩行経路に分類し、それぞれの固有空間を求めた。

まず、追跡する人物の歩行経路を手動で判断し、欠損率が変化するにしたがってその予測精度がどう変化するのかを確かめる実験を行った。図 4~6(a) はそれぞれの固有空間を生成するために使用した学習用の歩行経路 (白線) を示す。図 4~6(b)(c) に予測結果を示す。それぞれ白線が予測経路、黒線が実際の経路であり、追跡中の人物は矩形で示してある。追跡初期の欠損率が大きい場合、図 5(b) のように真の移動軌跡と予測との差が大きい。しかし、いずれも欠損率が低くなるにしたがって、予測された移動軌跡は真の軌跡に近付いていることが分かる。

次に、予測誤差による歩行経路選択がどのように行われるのかを示す。まず図 3 に、ある人物の追跡中の、各歩行経路の予測誤差の変化を示す。で示されている歩行経路が一貫して選択されているが、それに似た (で示される) 歩行経路との差は 200 フレームまでは非常に小さい。この理由は、画面上で二つの歩行経路が似ているため、ある程度追跡を行わなければその人物がどちらに進のかが判断できないためである。

図 7~9 (a)(b)(c) はそれぞれ、追跡開始後 50, 100, 150 フレーム後の予測された移動軌跡を表し、予測誤差を並記してある。図 7 では選択される歩行経路は一貫しているが、図 8 では時間が経つにつれて

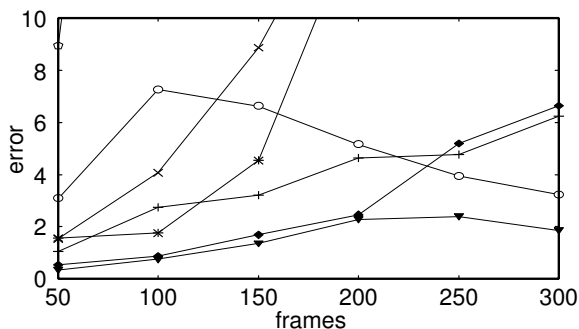


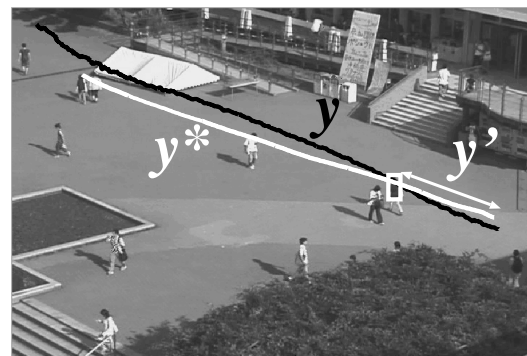
図 3: 各歩行経路についての予測誤差の変化。縦軸は予測誤差、横軸は追跡開始からのフレーム数。

(欠損率が低くなるので) 予測される移動軌跡が変化し、歩行経路の選択が変わっている。これは図 3 で示したように、二つの歩行経路が似通っているためである。

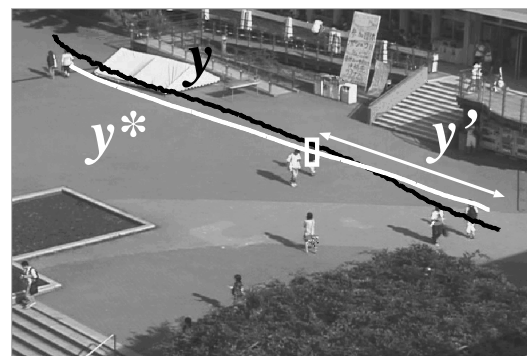
図 9 では、追跡される人物は右から現れて画面右上の階段を登っていくが、その歩行経路の予測誤差が一貫して小さく、正しい経路が選択されている。しかしそれ以外の歩行経路については、適切でない



(a)



(b)

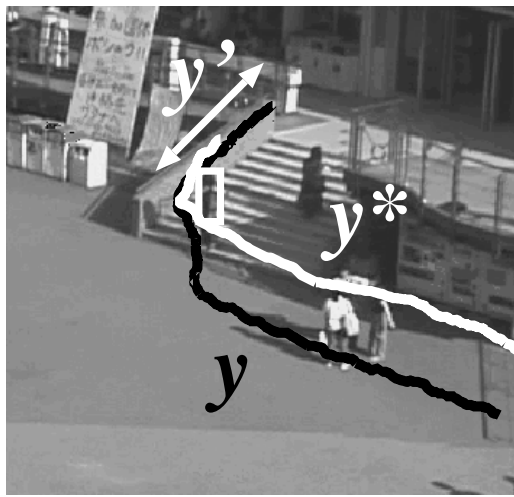


(c)

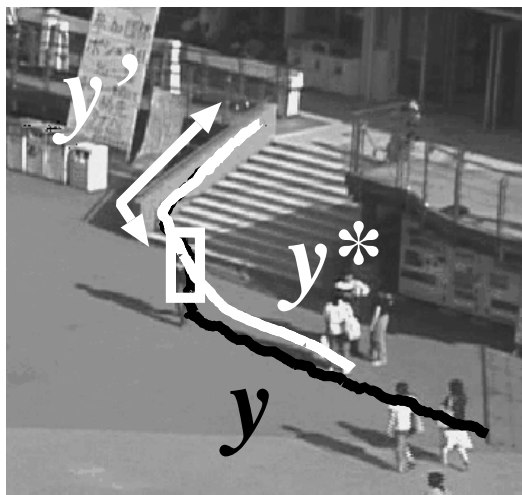
図 4: 画面右から左上への歩行経路。  $L_{min}^k = 1116$   
 (a) 学習した移動軌跡。(b)  $L_{y'} = 262$ 、欠損率 76%。  
 (c)  $L_{y'} = 462$ 、欠損率 57%。



(a)



(b)



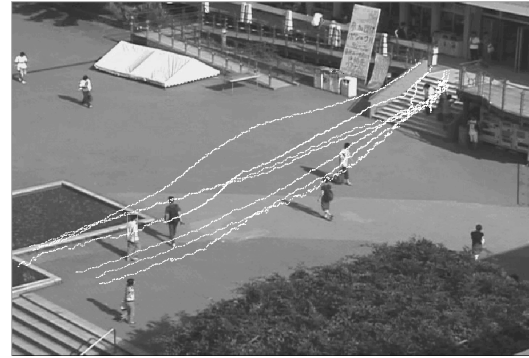
(c)

図 5: 画面上から右への歩行経路。  $L_{min}^k = 458$  (a) 学習した移動軌跡。(b)  $L_{y'} = 138$ 、欠損率 70%。(c)  $L_{y'} = 50$ 、欠損率 89%。

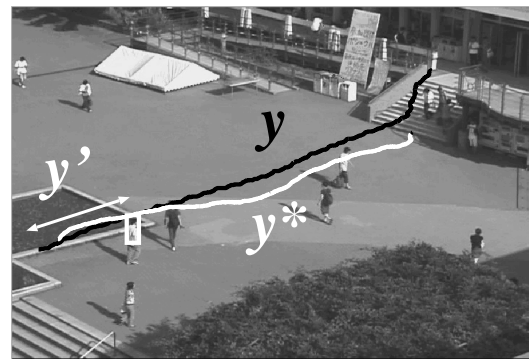
固有空間への射影の影響のために、予測誤差は非常に大きくなっている。

## 7 むすび

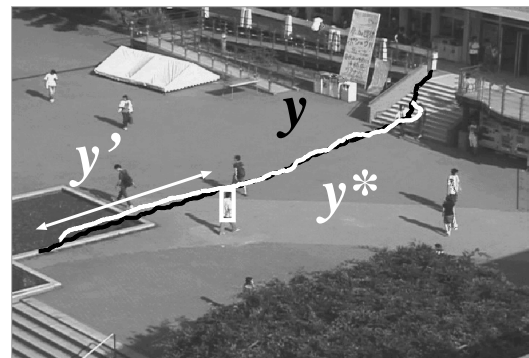
本論文では、あらかじめ人物の歩行経路を学習し、ある時点までの人物の移動軌跡をもとに、それ以降



(a)



(b)



(c)

図 6: 画面左下から右上への歩行経路。  $L_{min}^k = 994$  (a) 学習した移動軌跡。(b)  $L_{y'} = 194$ 、欠損率 80%。(c)  $L_{y'} = 452$ 、欠損率 55%。

の移動軌跡を予測する手法を提案した。実画像を用いた実験を行い、欠損率が小さくなるにつれて予測軌跡は真の移動軌跡に近くなり、また正しい歩行経路が選択されることを確認した。

今後は、さらに学習する移動軌跡と予測実験を増やし、本手法の定量的評価を行う。また現在の手法は空間的な予測のみを行っているので、時間的な予測を行う手法について検討する。また本手法では、 $y'$  から  $y''$  へベクトルを拡張する時に単に 0 を追加している。この拡張の予測への影響や拡張手法の改善は今後の課題である。

謝辞 本研究の一部は文部科学省科学研究費補助金若手研究 (B)(15700151) の助成を受けて行われた。

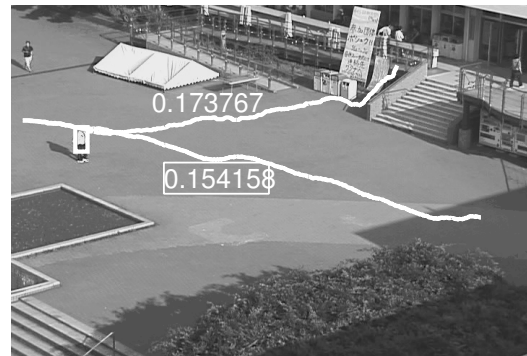
## 参考文献

- [1] 江下雅之, 監視カメラ社会, 講談社プラスアルファ新書, 2004.
- [2] 映像情報インダストリアル編集部: 「セキュリティのためのデジタルビデオレコーダを知る」, 映像情報インダストリアル, Vol.35, No.3, pp.13-16, 2003.
- [3] 鄭文涛, 三須俊彦, 苗村昌秀, 和泉吉則, 金次保明: 「複数カメラの協調動作による動オブジェクトの抽出」, 電子情報通信学会論文誌, Vol.J86-D-II, No.3 pp.397-408, 2003.
- [4] 森大樹, 内海章, 大谷淳, 谷内田正彦, 中津良平: 「非同期多視点画像による人物追跡システムの構築」, 電子情報通信学会論文誌, Vol. J84-D-II, No.1, pp.102-110, 2001.
- [5] 天野敏之, 佐藤幸男: 「固有空間を用いた BPLP による画像補間」, 電子情報通信学会論文誌, Vol.J85-D-II, No.3, pp.457-465, 2002.

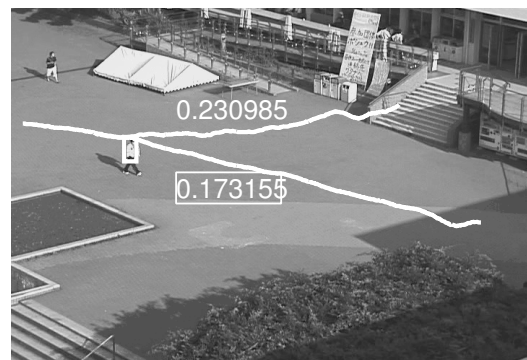
山本遼: 新潟大学工学部情報工学科 4 年生。

玉木徹: 2001 年名大大学院工学研究科情報工学専攻修了、同年新潟大学工学部助手、2003 年新潟大学自然科学研究科助手。工学博士。URL: <http://www.vision.ie.niigata-u.ac.jp/~tamaki/>

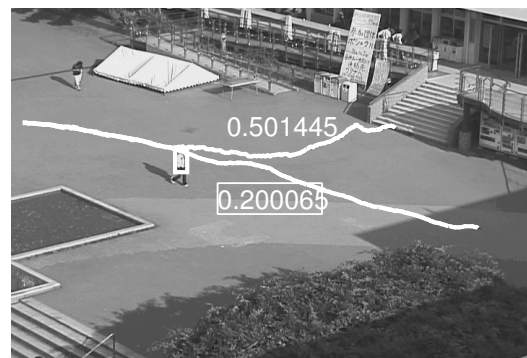
山本正信: 昭和 48 年九工大・工・制御卒、昭和 50 年東工大大学院修了、同年電総研(現産総研)入所。平成 4 年より新潟大学工学部教授。工学博士。著書「グラフィックスとビジョン」(オーム社, 1996) 等。URL: <http://www.vision.ie.niigata-u.ac.jp/>



(a)

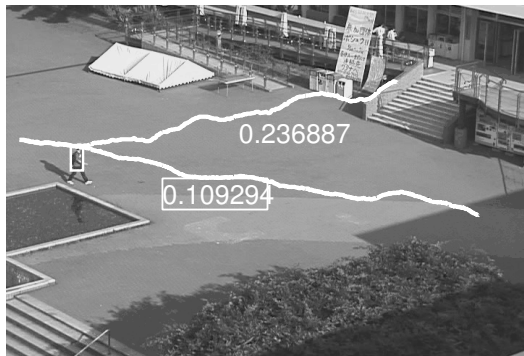


(b)

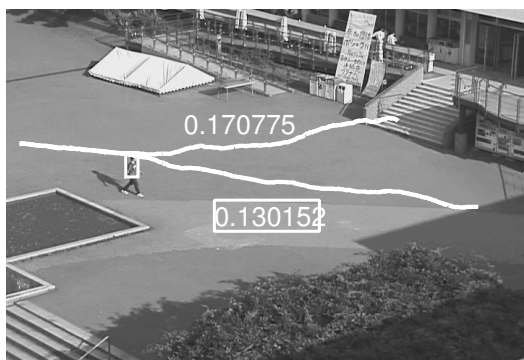


(c)

図 7: 経路選択の変化の様子。白線は歩行経路毎の予測経路を、数字は各経路の予測誤差を表す。選択された歩行経路の数字を矩形で示してある。(a) 追跡開始 50 フレーム後。(b) 追跡開始 100 フレーム後。(c) 追跡開始 150 フレーム後。いずれも左上から右下への経路が選択されている。



(a)

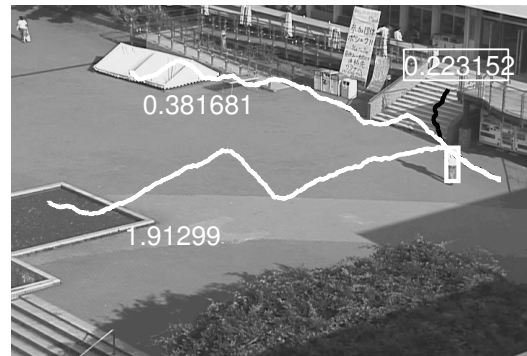


(b)



(c)

図 8: 経路選択の変化の様子。(a) 追跡開始 50 フレーム後。(b) 追跡開始 100 フレーム後。ここまででは左上から右下への経路が選択されている。(c) 追跡開始 150 フレーム後。右上への経路に選択が変更される。



(a)



(b)



(c)

図 9: 経路選択の変化の様子。いずれも右から右上への経路（黒線で示す）が選択されている。(a) 追跡開始 50 フレーム後。(b) 追跡開始 100 フレーム後。(c) 追跡開始 150 フレーム後。予測誤差は非常に大きくなり、選択された経路以外の予測は意味を持たない。