

クリギングを用いた人物移動予測

戸村 恵之^{1,a)} 川本 一彦^{2,b)} 岡本 一志^{3,c)}

概要: カメラを用いた人物移動予測のために、観測値間の空間相関に基づく面補間法の一つであるクリギングを用いた移動予測モデルを提案する。提案モデルでは、事前に収集した移動軌跡データに対してクリギングを適用し、未観測位置の移動方向と速度を補間し、任意の位置での予測を実現する。一般にクリギングでは、variogramに基づく面補間を行うが、方向データに対しては必ずしも適していない。本研究では、方向データに適したcosinogramに基づいて移動方向を補間する。評価実験では、200人の移動軌跡データから予測モデルを構築し、同じ空間で収集された別の人物データに対する予測成功率を算出し、クリギングの適用可能性を示す。

キーワード: クリギング, 移動予測, 補間

1. はじめに

カメラを用いた群集解析技術は重要性を増し、多くの手法が提案されている [1]。群集解析は精度の高い人物移動予測により高度なものとなる。移動予測を考える際に、相互作用を組み込んだ人物移動モデルが必要となる。代表的な従来手法に Social Force Model(SFM)[2] があり CV の分野では [3],[4]、および [5] などで適用されている。SFM は人間間にはたらく力に基づくモデルであり、偏微分方程式により表現される。現時点での位置関係に依存するため移動履歴の参照には適していない。別な手法には Floor Field Model(FFM)[6] があり [7] で適用されている。FFM は局所的なルールに基づくモデルであり、セルオートマトンにより表現される。空間構造を取り入れた予測に適しているが、予測の位置が各セル上であるため量子化誤差が生じてしまう。

本研究では対象空間における人物の過去の移動履歴から、任意の位置での移動予測を行うことを目的とする。空間上の相関を仮定し、統計的なモデルを用いて未観測位置の補

間を行うことで実現する。面補間法の一つであるクリギング [8] を用いて移動モデルを作成する手法を提案する。移動データとしては方向と速度を扱い、独立して予測する。

2. クリギング

いくつかの地点データが与えられたとき、それらを用いてデータの得られない地点の値を推定する方法にクリギングがある。これは地下資源の埋蔵量の推定など、地球統計学の分野で用いられている手法である。観測された地点データからバリオグラムモデルを作成し、非観測地点のデータを補間する。

2.1 バリオグラム

クリギングにおいて重要となるのがバリオグラムである。バリオグラムとは観測地点 s_i, s_j での観測データ $Z(s_i), Z(s_j)$ の共分散、

$$\gamma(h_{ij}) = \frac{1}{2} E[(Z(s_i) - Z(s_j))^2], \quad (1)$$

であり、距離の関数として表す^{*1}。 h_{ij} は観測地点間の距離である。このとき観測地点数の二乗個のバリオグラムが得られる。地点間の距離 h を階級区分 \hat{h} に細分化し、各区分に含まれるバリオグラムの平均値、

$$\gamma^*(\hat{h}) = \frac{1}{N_h} \sum_{i=1}^{N_h} \gamma_i(h), \quad (2)$$

をその区分の代表値、すなわち標本バリオグラムとする。

^{*1} 本来の共分散を両辺 2 で割っているため正式名称はセミバリオグラムであるが、しばしば混同して使われるためバリオグラムと表現する。

¹ 千葉大学大学院融合科学研究科
Graduate School of Advanced Integration Science, Chiba University

² 千葉大学統合情報センター
Institute of Management and Information Technologies, Chiba University

³ 電気通信大学大学院情報理工学研究科
Graduate School of Informatics and Engineering, The University of Electro-Communications

a) tomura@chiba-u.jp

b) kawa@faculty.chiba-u.jp

c) kazushi@uec.ac.jp

N_h は各区分中にあるバリオグラムの個数である。求めた標本バリオグラムを既存のバリオグラムモデルに当てはめ理論バリオグラムとする。

2.2 重み係数の算出

当てはめた理論バリオグラムを用いた予測方法を述べる。非観測地点 u_0 のデータ $Z^*(u_0)$ を、観測データと重み係数 w_i の重み付け平均、

$$Z^*(u_0) = \sum_{i=1}^N w_i Z(s_i), \quad (3)$$

として予測する。このとき重み係数には、

$$\sum_{i=1}^N w_i = 1, \quad (4)$$

といった拘束条件が与えられる。最適解を与えるために式 (3) での予測誤差、

$$\varepsilon(u_0) = Z^*(u_0) - \sum_{i=1}^N w_i Z(s_i), \quad (5)$$

を最小とする重み係数の組み合わせを求める。

これは平均二乗予測誤差、

$$\sigma^2(u_0) = - \sum_{i=1}^N \sum_{j=1}^N w_i w_j \gamma(h_{ij}) + 2 \sum_{i=1}^N w_i \gamma(h_{i,u_0}), \quad (6)$$

を最小にすることと等しい。 $\gamma(h_{ij}), \gamma(h_{i,u_0})$ はバリオグラムである。 σ^2 を最小にする重み係数は式 (4) の拘束条件があるためラグランジュ乗数法を用いて求める。重み係数とラグランジュ乗数を未知変数とした目的関数 $f(w_1, w_2, \dots, w_N, \lambda)$ は、

$$f = - \sum_{i=1}^N \sum_{j=1}^N w_i w_j \gamma^*(\hat{h}_{ij}) + 2 \sum_{i=1}^N w_i \gamma^*(\hat{h}_{i,u_0}) - 2\lambda \left(\sum_{i=1}^N w_i - 1 \right), \quad (7)$$

となる。式 (7) を各変数で偏微分することで連立方程式、

$$\begin{bmatrix} \gamma^*(\hat{h}_{11}) & \cdots & \gamma^*(\hat{h}_{1N}) & 1 \\ \vdots & \ddots & \vdots & 1 \\ \gamma^*(\hat{h}_{N1}) & \cdots & \gamma^*(\hat{h}_{NN}) & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} w_1 \\ \vdots \\ w_N \\ \lambda \end{bmatrix} = \begin{bmatrix} \gamma^*(\hat{h}_{1u_0}) \\ \vdots \\ \gamma^*(\hat{h}_{Nu_0}) \\ 1 \end{bmatrix} \quad (8)$$

が得られる。式 (8) を解くことで重み係数の組合せを得る。重み係数を式 (3) に代入することで予測値が算出される。

2.3 コサインオグラム (cosineogram)

従来のクリギングではバリオグラムモデルを用いて予測

を行うが本研究では移動データとして角度データを扱う。バリオグラムは角度データの扱いには適さないためコサインオグラム [9] を導入する。コサインオグラムとは観測地点 s_i, s_j での角度データ $Z_\theta(s_i), Z_\theta(s_j)$ の差の余弦値である。バリオグラムと同様に距離の関数として考え、

$$C(h_{ij}) = \cos |Z_\theta(s_i) - Z_\theta(s_j)|, \quad (9)$$

と表す。予測の手順はバリオグラムと同様である。

3. クリギングを用いた予測

人物移動予測を行うにあたり、対象となる空間を事前に撮影した動画から人物移動データを抽出しておくことを前提とする。移動データとしては方向と速度の2つを扱い、方向データはコサインオグラム、速度データはバリオグラムを用いる。

全方向への移動を同時に扱うと、コサインオグラムの分布は一様になると考えられる。一様分布に対して一つの代表値を当てはめることは適切でない。そこで本手法では簡易化のために移動軌跡がクラスタリングされていることを仮定する。撮影空間上にいくつかの目的地を設定し、目的地別に予測モデルを構築する。目的地は各人物が最後に観測された地点からの距離が最も近い場所とする。

3.1 予測モデルの構築

重み係数の推定用の人物の移動データを用いて学習を行う。各人物の移動軌跡から座標、及びその位置での移動方向と移動速度の3つのデータを抽出する。ある人物の n フレーム目での座標が (x_n, y_n) であったとき、フレーム間差分として移動方向 Z_θ を、

$$Z_\theta = \tan^{-1} \left(\frac{y_{n+1} - y_n}{x_{n+1} - x_n} \right) \quad (-\pi \leq Z_\theta \leq \pi), \quad (10)$$

移動速度 $Z_v(s_i)$ を、

$$Z_v = \sqrt{(y_{n+1} - y_n)^2 + (x_{n+1} - x_n)^2}, \quad (11)$$

と求める。角度データは x 軸方向を0とする。ただし、移動距離が0となる地点データは取り除くこととする。

得られたデータからバリオグラム、コサインオグラムを求める。地点間距離は小数点以下を四捨五入することで整数値に細分化する。理論モデルは、標本バリオグラム、標本コサインオグラムともに球形モデル、

$$\gamma_s^* = \begin{cases} a + b \left(\frac{3}{2} \left(\frac{h}{c} \right) - \frac{1}{2} \left(\frac{h}{c} \right)^3 \right) & (h \neq 0), \\ 0 & (h = 0), \end{cases} \quad (12)$$

$$C_s^* = \begin{cases} 1 - d - (1 - e - d) \left(\frac{3}{2} \left(\frac{h}{f} \right) - \frac{1}{2} \left(\frac{h}{f} \right)^3 \right) & (h \neq 0), \\ 1 & (h = 0), \end{cases} \quad (13)$$

またはガウスモデル、



図 1 学習用データの例

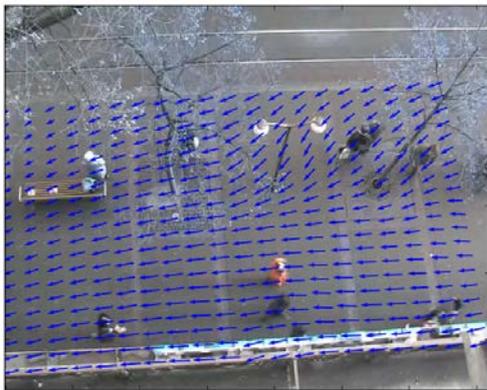


図 2 予測モデルの例

$$\gamma_g^* = \begin{cases} a - b \exp(-(\frac{h}{c})^2) & (h \neq 0), \\ 0 & (h = 0), \end{cases} \quad (14)$$

$$C_g^* = \begin{cases} d + (1 - e - d) \exp(-3(\frac{h}{f})^2) & (h \neq 0), \\ 1 & (h = 0), \end{cases} \quad (15)$$

のうち誤差が小さくなるほうのモデルを採用する。 a, b, c, d, e, f はパラメータであり、非線形最小二乗法を解くことにより最適値を求める。学習用データの例を図 1、作成する予測モデルの例を図 2 に示す。

3.2 移動予測

重み係数の推定に用いなかった人物の移動予測を行う。予測する人物の初期地点を u_0 とする。初期地点での移動方向と移動速度を予測し、移動先の座標を求める。この行程を繰り返すことにより人物の移動経路を予測する。なお、移動方向 Z_θ^* は一度 x 方向 $\cos(Z_\theta^*)$ 、 y 方向 $\sin(Z_\theta^*)$ にわけて予測を行ったのち、

$$Z_\theta^* = \tan^{-1} \frac{\sin(Z_\theta^*)}{\cos(Z_\theta^*)}, \quad (16)$$

として求める。これは、たとえば $-\pi$ と π のデータがあったとき、本来ならば平均値は π となるはずが 0 と求まってしまふのを防ぐためである。

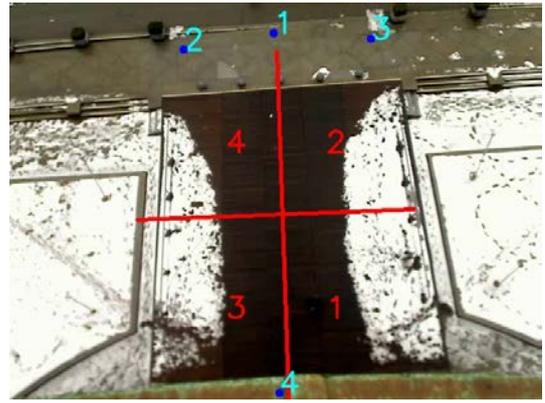


図 3 データセット 1 の目的地と領域の分割。縦方向が下向きが x 軸正の向き、横方向右向きが y 軸正の向き。

表 1 データセット 1 での目的地の座標。

目的地	x 座標 (m)	y 座標 (m)
1	-7.45	5.86
2	-6.59	0.07
3	-6.56	11.87
4	15.11	5.57

4. 予測性能の評価実験

文献 [5] の著者らが公開している動画から作成された 2 種類のデータセットを用いて実験を行う。どちらのデータセットも人物座標の取得は 2.5fps 間隔で行われている。データセットには、ホモグラフィ行列により画像上の人物位置を真上から見た画像上の位置に変換したものが格納されている*2。

4.1 評価方法

前半 200 人のデータを学習用、残りを予測、評価用として扱う。予測回数は予測人物の座標が取得されたフレーム数と同じ数だけ行う。フレーム度に真値と予測値の誤差を求める。誤差が、あらかじめ定めたしきい値以内であればそのフレームでの予測は成功とする。全フレームでの成功率を求める。予測人物全員の成功率の平均を予測精度とする。

人物の移動軌跡は空間構造より少なからず影響を受けているはずである。障害物があれば迂回し、建物の入り口があればそこへ向かう。このような、場所による差を予測に取り入れるため、動画内の空間をいくつかの領域に分割する。各領域別に予測モデルを構築し、同様に評価する。空間分割による予測精度の変化を観察する。

4.2 データセット

[5] で用いられている 2 種類のデータセットを用いる。seq_eth という名称のデータをデータセット 1 とし、seq_hotel という名称のデータをデータセット 2 とす

*2 ホモグラフィ行列は共に与えられているが、画像上の位置と変換後の位置が一致しない場合があるため、その場合は再計算した。

表 2 データセット 1 での領域分割.

領域	範囲 (m)
1	$x > 5.00, y > 6.00$
2	$x < 5.00, y > 6.00$
3	$x > 5.00, y < 6.00$
4	$x < 5.00, y < 6.00$



図 4 データセット 2 の目的地と領域の分割. 縦方向下向きが x 軸正の向き, 横方向右向きが y 軸正の向き.

表 3 データセット 2 での目的地の座標.

目的地	x 座標 (m)	y 座標 (m)
1	1.75	-10.25
2	1.75	4.32
3	4.38	-2.76
4	-3.29	-6.00
5	-3.29	2.00

表 4 データセット 2 での領域分割.

領域	範囲 (m)
1	$y < -3.00$
2	$y > -3.00$

る. データセット 1 は 6 フレームごと, データセット 2 は 10 フレームごとに座標の取得が行われている.

データセット 1 の目的地を図 3 と表 1 に示し, 空間分割の仕様を表 2 に示す. データセット 2 の目的地を図 4 と表 3 に示し, 空間分割の仕様を表 4 に示す. 図 3 と図 4 は動画の 1 フレームを切り取ったものに目的地と空間分割の仕様を描画したものである. 表 1 から表 4 に示す座標は真上から見た画像の座標であるため, 画像上にはホモグラフィ行列の逆行列により変換された座標で表示している. 縦方向が x 軸, 横方向が y 軸である. 目的地は本来データセットにより与えられているが, 動画空間内から外れていたり, 過剰数存在したりしているため修正を加えている. 各人物の目的地は最後に観測された地点からの距離が最小となる地点とするが, 観測が途中で中断されている人物データは目的地不明と見なし除くこととする. データセット 1 においては目的地 1, 2, 3 に向かう人物データはひとまとめとして予測モデルを構築している.

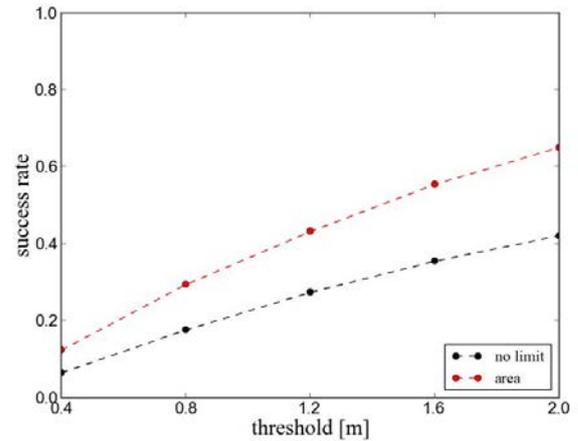


図 5 データセット 1 の予測成功率. 黒の破線が空間の分割なし. 赤の破線が空間分割あり.

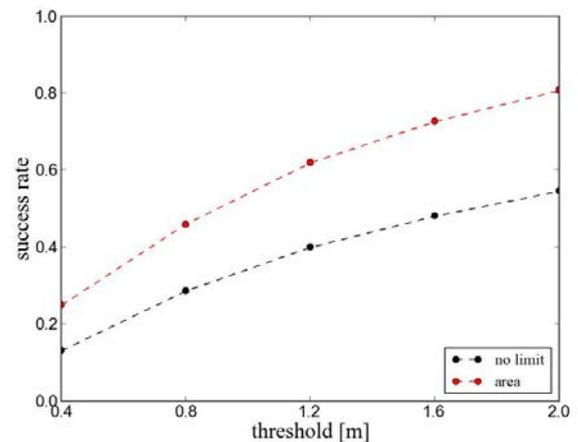


図 6 データセット 2 の予測成功率. 黒の破線が空間の分割なし. 赤の破線が空間分割あり.

4.3 結果

各データセットにおいて領域分割の有無による予測成功率の違いを図 5, 図 6 に示す. 横軸は予測成功と見なす誤差の許容範囲, 縦軸は予測成功率を表している. データセット 1 では 167 人, データセット 2 では 220 人分の予測を行っている.

例として何人かの真の経路と予測経路を図 7, 図 8 に示す. 赤の点が真の経路, 緑の点が空間分割なしの予測経路, 青の点が空間分割ありの予測経路である.

4.4 考察

図 5, 図 6 より, 空間をいくつかの領域に分割して予測を行うことで精度が向上することがわかる. 本手法において空間の分割を行うことはコサインオグラム, バリオグラム生成の際に, 遠方の地点でのデータを参照しないということと同意である. つまり, より近くの, 似たような傾向を持った移動データを用いて予測を行ったため精度が向上し

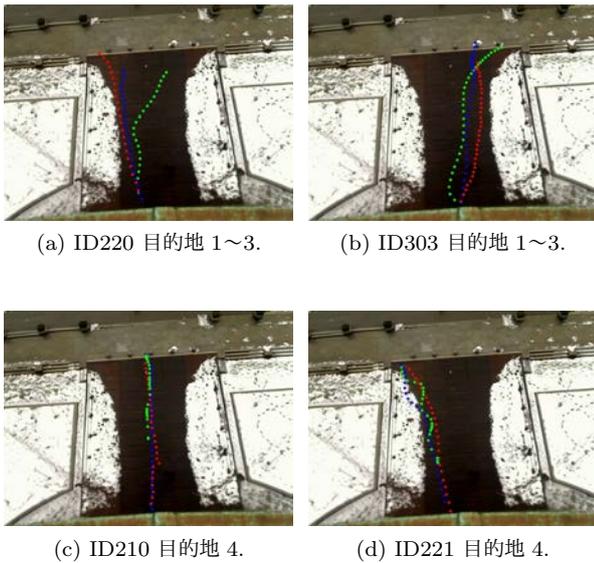


図 7 データセット 1 での予測結果. 赤が真の経路, 緑が空間分割なしの経路, 青が空間分割ありの経路.

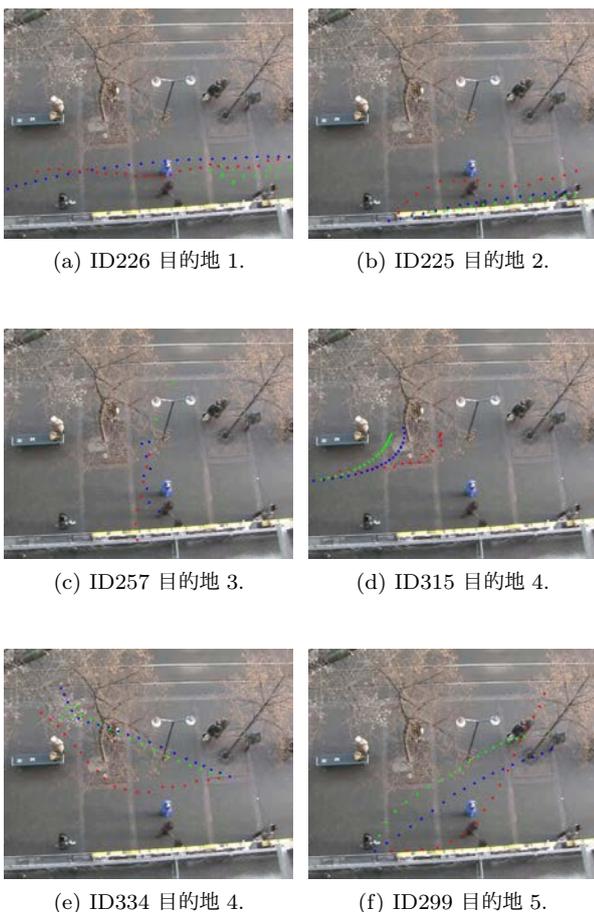


図 8 データセット 2 での予測結果. 赤が真の経路, 緑が空間分割なしの経路, 青が空間分割ありの経路.

たのだと思われる. 空間の分割をより細かくすることでさらに精度は向上すると考えられる.

図 5, 図 6 よりデータセット 2 のほうが予測精度が高いことがわかる. これはデータセット 2 のほうが動画空間の範囲が狭いためであると考えられる. 誤差の許容範囲を等しく設定するとデータセット 2 の成功判定基準が甘くなるためである.

予測がうまくいかない場合を挙げる. 予測対象者の中には移動中に静止を含むデータが存在する. しかし予測モデル作成の段階で移動がゼロのデータは除いている. そのためこういった人物の予測成功率はさがってしまう.

5. おわりに

本研究ではクリギングを用いて人物移動予測を行う手法を提案した. 実動画から得られたデータセットを用いて移動予測を行い, 真値と予測値の誤差を求めることで評価を行った. 動画上の空間をいくつかの領域に分割し, 領域ごとに予測モデルを構築することで予測精度が向上することを示した.

本研究では空間分割の際に単に空間の中央付近での分割を行った. 実際の人物移動は空間構造により影響をうけるはずである. そのような構造を考慮した分割を行うことでさらなる精度の向上が期待できると考えられる.

本手法では予測人物の目的地も既知であるとして予測を行っているが, これは必ずしも現実的ではない. データセット 1 の場合であれば初期地点により目的地は絞られるが, データセット 2 の場合は明確ではない. 初期地点より目的地の候補を選び, 複数の経路予測を同時に行うことで予測が可能であると考えられる.

データセット 2 の動画では電車の発着が確認できる. こういった場所での人物移動軌跡は時間帯によって大きく変動することが予想できる. 時間差を考慮したモデルの構築による精度の変化も検証していきたい.

謝辞

本研究は JSPS 科研費 25330186 の助成を受けたものである.

参考文献

- [1] J. C. S. Jacques, Jr., S. R. Musse and C. R. Jung, "Crowd Analysis Using Computer Vision Techniques," *IEEE Signal Processing Magazine*, Vol. 27, No. 5, pp. 66-77 (2010).
- [2] D. Helbing and P. Molnar, "Social force model for pedestrian dynamics," *Phys. Rev. E*, Vol. 51, No. 5, pp. 4282-4286, (1995).
- [3] R. Mehran, A. Oyama and M. Shah, "Abnormal crowd behavior detection using social force model," *CVPR*, (2009).
- [4] M. Luber, J. Stork, G. Tipaldi and K. Arras, "People tracking with human motion predictions from social forces," *ICRA*, (2010).
- [5] S. Pellegrini, A. Ess, K Schindler and L. V. Gool, "You'll

- never walk alone: Modeling social behavior for Multi-target Tracking," *ICCV*, (2009).
- [6] C. Burstedde, K. Klauck, A. Schadschneider and J.Zittartz, "Simulation of pedestrian dynamics using a 2-dimensional cellular automaton," *Phys. A Statistical Mechanics and its Applications*, Vol. 295, No.3-4, pp.507-525(2001).
- [7] S. Ali and M. Shah. "Floor fields for tracking in high density crowd scenes," *ECCV*, (2008)
- [8] 古谷知之, "R による空間データの統計分析", 朝倉書店, 2011
- [9] Morphet, W.James. Simulation, "Kriging and Visualization of Circular-Spatial Data". *Graduate Theses and Dissertations*. Paper 386, (2009).