

UniversalSAX を用いた人流可視化

宮城 優里 (指導教員: 伊藤 貴之)

1 はじめに

カメラで記録されている映像データを分析して抽出した人の歩行パターンは混雑の原因究明や、より良い避難経路の策定など様々な分野で役立てられる。しかし現在、人流の解析は人が映像を直接観察するのが主流である。そのため、大規模人流データの全体像を把握し重要な部分を発見することが困難である。計算機を用いた分析にあたり、特にデータの圧縮とマイニングについて検討する必要がある。本報告では時系列データを文字列に変換する SAX(Symbolic Aggregate Approximation) を拡張した UniversalSAX を用いて人流データを圧縮し、自然言語処理手法によって特徴的な情報を抽出する方法およびこれらの情報を可視化する手法を提案する。

2 関連研究

人流データを文字列化し処理する方法としては、動線の形状と絶対的な位置を基に分類する手法 [1] や独自の文法に基づいてパターン検出を行う手法 [2] が挙げられる。これらの研究では、取得した人流データをすべて可視化しているため、データの増加に伴って注目すべき要素を見つけにくくなる傾向がある。この点に関し、本手法では人流データを変換して生成した文字列から特徴的なものを選択し、それらを強調表示するように可視化する点が既存手法と異なる。

3 提案手法

本手法ではカメラで撮影した人流データを文字列に変換して圧縮したうえで、文字列用アルゴリズムを適用して人流の特徴を抽出し可視化する。

3.1 人流データの取得

現時点ではモーションキャプチャデバイス Xtion によって取得した人流データを使用している。撮影中は微小時間ごとに歩行者の頭部座標 (3次元) と個人識別子を記録するが、座標に関しては床平面上の2次元系を使用している。つまり本研究における人流データは、時刻 t 、識別子 id 、座標値 (x,y) の4変数からなるレコードの集合であると考えられる。この集合から同じ識別子を有する座標を時系列順に連結することで各歩行者の歩行経路を取得する。

3.2 人流データの文字列化

続いて本手法では SAX を多次元データに適用できるように拡張した UniversalSAX[3] を用い、実数値座標で表現された歩行経路を文字列に変換する。まず多次元空間 (今回は x,y の2次元) に領域分割を適用し、各領域に文字を割り当てる。同時に各領域間の距離表を作成する。さらに人流データの座標にアフィン変換を適用し、カメラの視野中央の座標が原点となるように調整する。これらの下準備をしたのち、領域分割結果に従って人流データを文字列に変換し、撮影した人数と同数の文字列を生成する。

最後に、作成した文字列にランレングス圧縮を適用

し、データ量を削減する。これらの処理により歩行者の各々について、滞在場所を文字の種類、滞在時間を文字の連続出現回数として取得できる。

3.3 人流の特徴抽出

3.2節までの処理によって得られた文字列に対し、以下の情報を得るためのマイニング処理を適用する。

- 立ち止まる人が多い場所
- 各領域の平均滞在時間
- 典型的な歩行経路

歩行者が立ち止まりやすい場所は、ランレングス符号の中の連続出現回数が多い文字として検出できる。同時に各領域での滞在時間についても、文字ごとの連続出現回数の平均値をとることで検出可能である。典型的な歩行経路の検出は次のような手順で行う。

まずランレングス符号から連続出現回数を示す数値を除去して文字だけを残した文字列を作成し、これに n -gram を適用する。以上の処理によって部分的な歩行経路を n 文字の単語として検出し、それぞれの出現回数を取得する。この時、含んでいる文字の種類数 (歩行範囲) が一定値以下の単語は除外する。

次にこれらの単語の中から代表経路を m 種類選択する。まず最も出現回数の多かった単語を1つめの代表経路として決定し、2つめ以降は出現数が多くかつ選択済みの代表経路との距離が一定以上あるものを追加する。2つの動線同士の距離は i 番目の文字同士の距離を距離表から取得し、その合計値として算出する。このため同じ場所を通ったものでも向きが正反対だった歩行ルートは別々に代表経路として採用されることがある。これらの操作は代表経路数が m 種類に達するか、全ての単語の判定が終わるまで繰り返す。

代表経路として選択されなかった残りの単語については、領域分割時の距離表を用いて各代表経路との距離を計算し、最も近いものと同一とみなす。いずれの代表経路とも距離が大きかった場合は例外的な歩行経路とみなして個別に記録する。

3.4 歩行経路描写とグラフ化

最後に2通りの可視化を行う。まず図1(左)のように、取得した歩行経路を動線として平面上に描写するアニメーションを実行する。画面右下に時刻を表示し、その時刻に歩行者がいる位置に○、IDを表示する。動線はIDごとに異なる色で着色し、同一時刻に複数の歩行者がいた場合と同じ色の動線が交差して区別しづらくなることを避ける。色の明るさは歩行速度と対応し、高速で歩いている時ほど明るくなる。これによって特定の時間帯における各歩行者の具体的な行動が把握できる。それと同時に本手法では、マイニング処理によって得られる集計結果、例えば各領域で立ち止まった人の数、典型的な歩行経路の出現頻度、なども棒グラフを用いて可視化する (図1(右))。これによって動線アニメの観察だけでは把握しづらい人流の長期的な

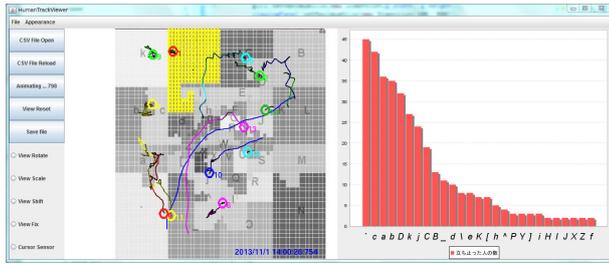


図 1: 歩行経路アニメ (左) と特徴集計結果グラフ (右)

傾向や歩行者集団の特徴に関するユーザーの理解を促進する。これらの可視化画面を連動させることで、動線アニメ画面において特徴的な人流や場所を選択または強調して表示する。例えば棒グラフの中で一定値を超えた要素や、ユーザーがクリックした要素に該当する動線や領域を色付けする。

4 分析例

ある施設での展示会において取得した人流データの分析結果を図 2 に示す。

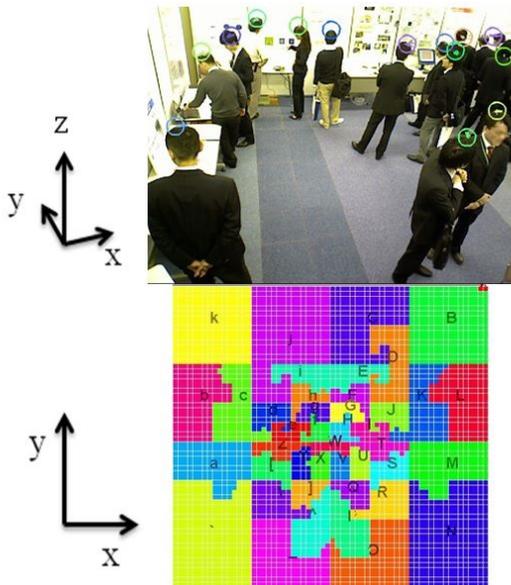


図 2: 撮影場所風景 (上) と領域分割結果 (下)

この実行例では展示が行われていた部屋の一部を図 2(上)の構図で撮影した。撮影時間は 1 時間で、のべ 645 人分の歩行を取得した。カメラの視野の左端から上端にかけては展示スペースとなっている。カメラの視野には入っていないが下方と右上にはそれぞれ出入口がある。この実空間を真俯瞰した平面を図 2(下)に対応づけ、歩行者の通過座標を文字列化した。図 3 は、作成した文字列に $n=6$ として n -gram を適用し、典型的な歩行経路を単語として選出した結果である。

最も多く出現した単語 “\XVUT” や “UTIJKL” は ‘_’ 付近の出入口から ‘L’ 付近の出入口へ直行した場合の部分経路と考えられる。これを逆向きに移動した経路が “VX[_” として別に集計された。前者に比べて出現数が少なくなっている理由は、展示会の受付に近い ‘_’ 側の出入口からの入室者の方が多かったためと考察される。それに対して “jdcZ[\” と

“djhiC” は出入口と展示スペースの間を往来した人の動きに相当する。

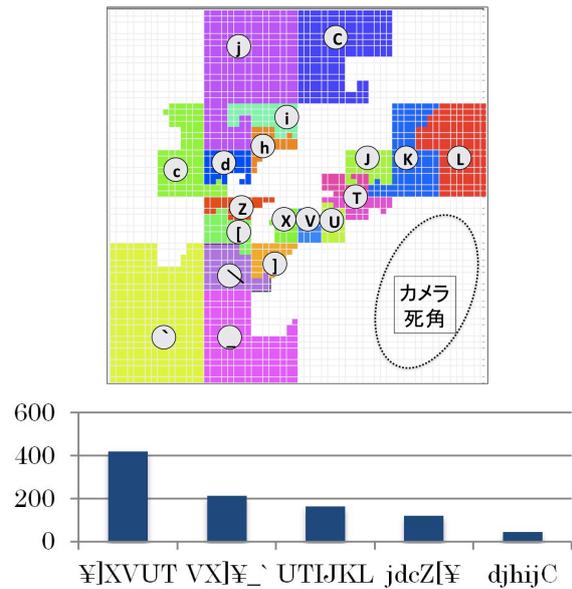


図 3: 典型的な歩行経路とその出現頻度

5 まとめと今後の課題

本報告では、UniversalSAX により文字列化した大規模人流データに、マイニングを適用して特徴的な人流を抽出し、可視化する手法を提案した。今後の課題は以下の通りである。

- 歩行経路と場所の強調表示手法拡張
- より長時間分の人流データを用いた実験
- 異なる場所で同時刻に取得した人流データ間の相関性発見

6 謝辞

本研究に関してご指導いただきました産業技術総合研究所大西正輝様、プログラムを提供していただきました筑波大学渡辺知恵美先生に感謝いたします。

参考文献

- [1] Oates, T., Boedihardjo, A. P., Lin, J., Chen, C., Frankenstein, S. and Gandhi, S.: Motif Discovery in Spatial Trajectories using Grammar Inference, in *ACM International Conference on Information and Knowledge Management*, pp. 1465–1468 (2013).
- [2] Thach, N. H. and Suzuki, E.: A Symbolic Representation for Trajectory Data, 2010 年度人工知能学会全国大会, 1A2-2 (2010).
- [3] 大西史花, 渡辺知恵美: Universal SAX:空間充填曲線を利用した SAX の多次元時系列データへの適用, *日本データベース学会論文誌*, Vol. 11, No. 1, pp. 43–48 (2012).