

[研究助成 (A)]

一人称カメラを用いた歩行時の姿勢推定と 視覚障がい者の感覚代行への応用

Posture Estimation during Walking using a First-Person Camera and Its Application to Sensory Substitution for the Visually Impaired

2201014



研究代表者

福井大学 学術研究院
工学系部門

講師

張

潮

[研究の目的]

視覚障がい者の点字ブロックを用いた歩行は、白杖で確認しながら歩く必要があり、円滑とは言えない。一方、一人称カメラの小型化に伴い、常時カメラや情報を処理できる装置を身に付けることが可能になった。そこで、本研究では、人間の視覚システムが担う姿勢推定を機械の自動化で代替し、視覚情報を聴覚情報へマッピングすることで感覚代行の実現を目指す。

そこで、一人称カメラから点字ブロックを検出することで、利用の円滑化を進め、視覚障がい者の歩行支援に繋がることが期待できる。点字ブロック検出に関する研究はいくつか行われており、センサを利用して点字ブロックの表面の凹凸を検出して識別する手法や、畳み込みニューラルネットワークを使用した検出が行われているが、これらは特別なセンサや大量の学習データを必要とする。しかし、単純な形状や色のついた物体においては、大量の学習データやそれに伴う手作業を必要とせず、特徴を定式化することで検出が可能であると考えている。そこで、本研究では点字ブロック検出を多目的最適化問題とし、色と形状特徴を利用した検出方法を提案し、一人称視点画像から点字ブロック検出を行う（結果例を図1に示す）。本研究



図1 一人称視点の画像における点字ブロックの検出

では、大量の学習データや手作業の必要の無く、様々な状況下でも検出が可能な点字ブロック検出の実現を目的とする。

[研究の内容, 成果]

1. 提案手法の概要

本研究では多目的最適化アルゴリズムを用いた一人称視点画像における点字ブロックの検出手法を提案した。提案手法では、点字ブロックの色と形状の2つの特徴を用いて、点字ブロック上から撮影した一人称視点画像から点字ブロックを検出する。

本研究では、図1に示す一般的に使用されている直線的な輪郭を持つ黄色の点字ブロックを対象とする。一人称視点映像では、点字ブロックは常に画像の下側から現れる。点字ブロックは2本の直線で囲まれた領域として表現できるため、点字ブロック検出は2本の直線で囲まれた内側が黄色の領域を見つける問題と言い換え

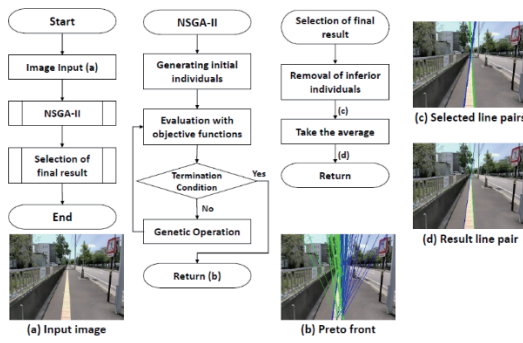


図2 提案手法の概要

ることができる。ここで、画像から直線を検出することは、幾何学形状を検出することと理解することができる。提案手法である多目的最適化を用いた点字ブロック検出について、以下に詳細に説明し、その概要を図2示す。

2. 提案手法

2.1 解個体の表現方法

本研究では点字ブロックの領域を2本の直線で表現し、これらの個体を line pair (LP) とする。各個体である LP は、画像中の上端と下端の点の x 座標を表す実数を染色体として使用している。LP = $\{x^l, x^b, x^r, x^t\}$ は、点字ブロックの輪郭の左側の線を x^l と x^b で、右側の線を x^r と x^t を通る直線として個体を符号化し、4つの点で表現する。このような表現の下で、初期母集団の各解候補 LP の4点は、画像の上端と下端から無作為に選択される。この時、点字ブロックの領域を明確に示さない LP は候補として生成されない。

2.2 目的関数の設計および最適化

個々の LP が点字ブロックの領域を正確に表しているかどうかを評価するために、色と形状の特徴を用いた2つの目的関数を同時に最適化する。どちらの目的関数も、反復的にサンプリングされた LP 候補の端点の座標を計算し、画像の下端から LP の交点まで（交点が画像中に存在しなければ上端まで）、テストラインごとに評価を行い、それらの評価値の総和を目的値としている。図3に目的関数で用いられる変数

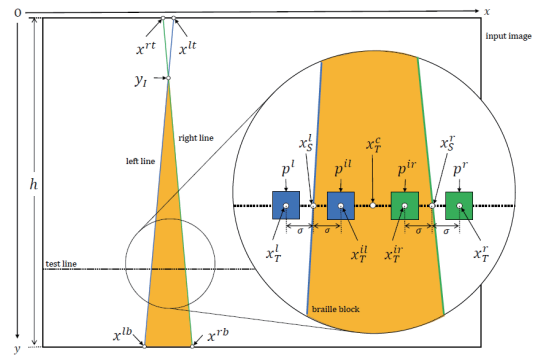


図3 点字ブロックにおける決定変数

を示す。

点字ブロックの色が黄色であることに着目し、HSV 色空間のカラーヒストグラムを利用した。目的関数1は LP が示す領域が黄色であることを、(1) LP の示す領域の内側と外側でカラーヒストグラムが大きく異なること、(2) LP の内側のテストポイントの HSV 色空間における値が予め設定している黄色の閾値内に収まることの2つの条件から判断している。入力画像の HSV 画像を $I_{HSV}(x, y)$ 座標における画素値を $I_{HSV}(x, y)$ と表現する。カラーヒストグラムを調べるためのテストポイントの x 座標を $x_T = [x_T^l, x_T^b, x_T^r, x_T^t]$ とし、テストポイントを中心座標とする 3×3 画素のパッチを $p = [p^l, p^b, p^r, p^t]$ 、そのパッチ内のカラーヒストグラムを $x_T = [x_T^l, x_T^b, x_T^r, x_T^t]$ とする。これらの座標は各テストラインごとに毎回計算され決定される。また、 x_T は LP と y から算出され、 p は x_T と I_{HSV} から作成され、 $hist$ は p から算出される。LP の交点の y 座標を y_I とする。 y_I は、LP から算出され、 y_I が画像内にあればテストラインの y 座標の初期値とし、画像内になければテストラインの y 座標の初期値は画像の上端としている。 y から画像の高さ h までの各テストラインにおいて、次の2つの条件が同時に満たされている場合、カウンタ変数である c_a と c_b をカウントする。(A) パッチ内のカラーヒストグラム $hist^l$ と $hist^b$ 、 $hist^r$ と $hist^t$ に大きな差がある場合。ここでは、2つのパッチ内のカラーヒストグラムを比較するために、

Bhattacharyya 距離を使用して2つのヒストグラムがどれだけ一致していないかを計算している。(B) テスト点 x_T の値が事前に設定した黄色の閾値内にある場合。特に(B-1)ではテスト点 x_T^h と x_T^l を、(B-2)では x_T^h と x_T^l 、もしくは x_T^h と x_T^l が(B)の条件を満たす場合カウントを行う。カウンタを2つ用いることで、テストポイントの場所によって重みを変えることができ、画像に対するノイズ耐性を向上させている。点字ブロックの輪郭のすぐ内側に位置するテストポイント(x_T^h や x_T^l)を用いれば、個体をより正確に評価することができるが、画像中にノイズが多い場合にはLPの中央に位置するテストポイント(x_T^c)を代替として使用することで、ノイズがある場合でもカウントが行えるようにしている。 c_a は条件(B-1)を満たした場合のみカウントを行うため、より完全な点字ブロックに適応させることができるが、画像中にノイズが多い場合には上手く機能しない。そこで、条件(B-1)または(B-2)のいずれかが満たされた場合にカウントを行う c_b を追加することで、テストポイントにノイズがある場合であっても、個体を適切に評価できるようにした。目的値 v_{1a} と v_{1b} には c_a と c_b の負数が設定される。これは、点字ブロック検出の最適化問題が最小化問題として設計されているためである。

目的関数2は点字ブロック上から撮影した一人称視点画像において、点字ブロックが直線的な形状であることを特徴として用いた。目的関数2では、直線に対する類似性と距離を使用した最適化を行う。この手法では、特定のエッジ抽出方法で得られた点字ブロックの輪郭を含むエッジ画像から直線を検出する。この時、エッジ画像を距離変換画像に変換することで、各画



図4 (a) 入力画像, (b) エッジ画像, (c) 距離変換画像

素の値は最近傍のエッジからのユークリッド距離となる(図4)。交点以下のLP上の画素値の総和を目的値とすることで、個体の直線に対する類似度と距離を評価することができる。また、距離変換画像をLP上の画素値の総和を目的値とすることで、個体の直線に対する類似度と距離を評価することができる。距離変換画像を I_{DT} とし、 (x, y) 座標の画素値を $I_{DT}(x, y)$ とし、距離変換画像の値をサンプリングするLP上のサンプル点の x 座標を $x_s=[x_s^l, x_s^r]$ としている。目的関数1と同様に y から h までの各テストラインにおいて、LPと y から x_s を算出し、 $I_{DT}(x_s^l, y)$ と $I_{DT}(x_s^r, y)$ の総和を目的関数2の値(v_2)としている。

2.3 遺伝的操作および結果選択

提案手法では遺伝的操作として選択、交叉、突然変異が使用される。NSGA-IIで用いられる遺伝的操作である選択には、入れ替えのないCrowded binary tournament selectionを使用している。これは、前回の探索で見つかった優良個体を次の世代に伝播させるために、親集団の非優越解を非優越ランクと混雑度距離に基づいて選択を行う。また子集団は、選択された染色体の交叉と突然変異によって生成される。交叉にはsimulated binary crossover(SBX)、突然変異にはpolynomial mutation(PM)が使用される。SBXはバイナリエンコードされた実数の決定変数の一点交叉を再現し、PMは突然変異としてビット反転を行い、これらの遺伝的操作によって、親に近い子孫を生成している。提案手法では、NSGA-IIの交叉確率を1.0、突然変異確率を0.25に設定して実行している。

NSGA-IIを実行して得られた非優越解から最終的な解を選択するにあたって、点字ブロックの領域を明らかに示さない個体を取り除き、残りの個体の平均を取っている。各個体は、LPの交点が画像下部に位置している場合か、画像下部のほとんど同じ座標からLPの2本の線が始まっている場合に、点字ブロックの領域

を明らかに示さないと判断される。

3 実験と結果

3.1 実験環境

実験では、実際に提案手法を用いて画像データセットに対して点字ブロック検出を行い、手作業でアノテーションした正解のLPとの誤差を比較することで検出精度を評価する。画像テストデータセットは、点字ブロック上から一人称カメラ（Vuzix M400）で撮影された一人称画像で50枚で、照度変化、影、欠損、遮蔽、撮影角度変化の5つの状況が含まれており、それぞれ10枚ずつで構成されている。いずれも320×240画素のRGBカラー画像である。また、全ての実験結果は、異なる乱数シード値を用いた10回試行結果の平均値を用いている。

評価指数には、手作業でアノテーションした正解のLPとの交点以下の平均位置誤差とした（画像内に交点が位置していない場合には、LP全体の平均位置誤差としている）。平均位置誤差には下記の式で定義され、Root Mean Squared Error (RMSE)を用いた。

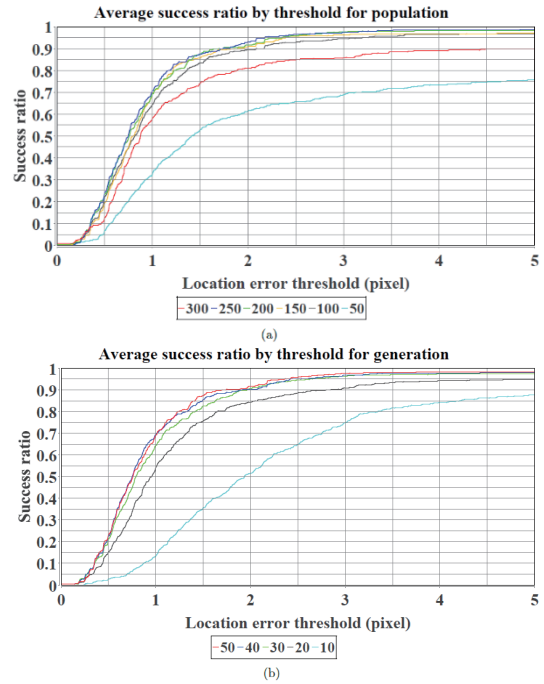
$$RMSE = \sqrt{\frac{\sum_{y=y_i}^h (x_R - x_C)^2}{h - y_i}}$$

ここで、 x_R と x_C はそれぞれ検出されたLPと正解のLPのx座標であり、 y_i は正解のLPの交点のy座標、 h は画像の高さを表す。

3.2 パラメータチューニング

NSGA-IIでは、個体数と最大世代数をパラメータとして設定する必要がある。そのため、まず個体数と世代数ごとに実験を行った。

図5に個体数と世代数別の結果を示す。図5(a)は個体数を50から300までの範囲で50ごとに变化させた結果であり、世代数はすべて50としている。個体数200と250で最良の結果が得られ、この結果から実験に使用した画像の大きさでの特徴空間の探索には、個体数は



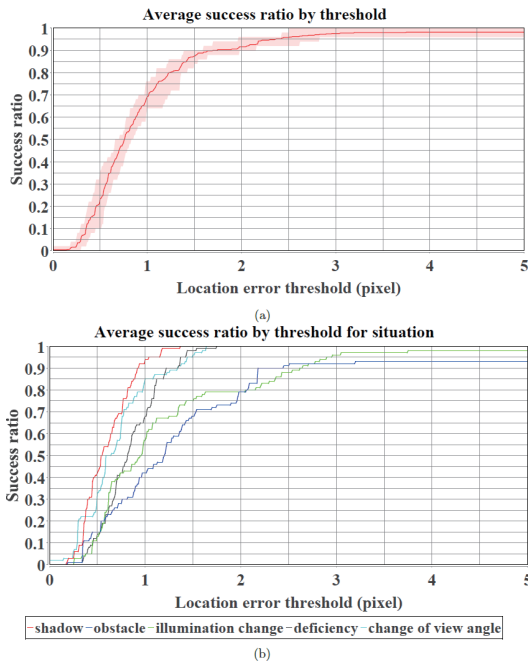
(a) 実線は個体数ごとの10回試行結果の平均値を示しており、凡例は個体数を表している。
(b) 実線は世代数ごとの10回試行結果の平均値を示しており、凡例は世代数を表している。

図5 画像データセットにおける平均位置誤差ごとの検出成功率

200で十分であることがわかる。また、個体数300で検出成功率が低下しており、過剰な個体数は誤検出を増加させている。図5(b)には最大世代数を10から50まで10世代ごとに变化させた結果を示す。個体数はすべて200で行った。検出精度は30世代でほとんど上限に到達し、50世代では必要十分な最良の結果が得られることが分かる。

3.3 検出精度の評価実験

次に、得られた最適なパラメータ設定を用いて得られた実験結果を示す。NSGA-IIは個体数200、世代数50で実行した。図6(a)が示すように、平均位置誤差が3画素以下の場合、提案手法による点字ブロックの検出成功率は95%以上であることがわかる。ここで、平均位置誤差が3画素以内であれば、画像中から点字ブロックの識別するには十分である。図6(b)に各状況ごとの数値結果を、図7に出力結果の画像を示す。これらの結果が示すように、



- (a) 実線は10回試行結果の平均値を、赤の領域は最大値から最小値までの範囲を示している。
- (b) 実線は状況ごとの10回試行結果の平均値を、凡例は各状況を表している。

図6 画像データセットにおける平均位置誤差ごとの検出成功率

提案手法は、撮影角度変化、影、欠損に強く、障害物や照度変化においても十分な精度があり、様々な状況下で点字ブロック検出が可能であることがわかる。ただし、点字ブロック上に障害物がある場合は色や形状の特徴が十分に得られないことがあり、照明変化が大きい場合には色の特徴が全く得られないために、いずれの場合も検出精度が少し低下している。両方の特徴が十分に得られない場合や、片方の特徴が全く得られない場合に検出が難しくなることが分かる(図8)。以上のように、提案手法を用いた点字ブロック検出は、様々な状況下で柔軟に対応ができ、ロバストな検出が可能であることが得られた結果からわかった。



図7 各状況ごとの出力結果

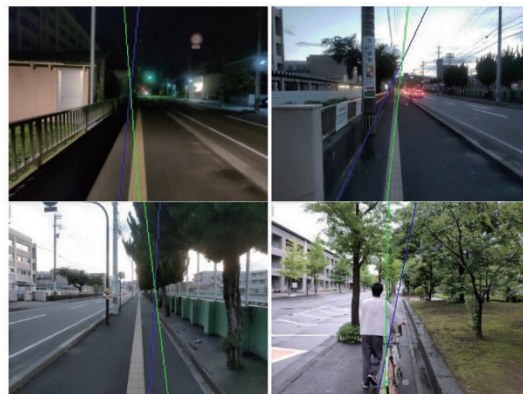


図8 検出に失敗した場合の出力結果

[今後の研究の方向, 課題]

本年度は主に多目的最適化を用いた点字ブロック検出手法を提案した。実験の結果, 提案手法は様々な状況を含む画像データセットにおいて高い検出成功率を示し, 実シーンの一人称視点の画像から点字ブロック検出するのに有効であることがわかった。点字ブロックの色と形状の特徴を利用した2つの目的関数に対する最適化により, 一方の特徴量が十分に利用できない場合でも, もう一方の特徴空間から適切な解を導き出すことができ, ロバストな検出が可能になった。点字ブロックの検出結果に基づいて, 歩行者の歩行姿勢を判定することが可能であり, 今後の研究で歩道から外れていることを実際に聴覚で感じてもらうことで感覚代行の実現を目

指す。なお, 本研究の詳細は成果発表 [1] にて論じており, 関連研究の調査結果は成果発表 [2] にまとめられている。

[成果の発表, 論文等]

- [1] Tsubasa Takano, Takumi Nakane, Takuya Akashi, Chao Zhang, "Braille Block Detection via Multi-Objective Optimization from an Egocentric Viewpoint", *Sensors*, Vol. 21, No. 8, article number 2775, 15 pages, 2021
- [2] Takumi Nakane, Naranchimeg Bold, Haitian Sun, Xuequan Lu, Takuya Akashi, Chao Zhang, "Application of Evolutionary and Swarm Optimization in Computer Vision: A Literature Survey", *IPSN Transactions on Computer Vision and Applications*, Vol. 12, No. 3, pp. 1-34, 2020