

カメラ映像を用いた地下街歩行者の分布推定方法に関する研究

○高木 尚哉*¹ (1411767)
瀧澤 重志*² (952332)

キーワード：避難計画 歩行者 地下街 ニューラルネットワーク カメラ映像 背景差分法

1. はじめに

大阪府の発表¹⁾によると、南海トラフ巨大地震により梅田地区では最大2mの浸水が予想されている。また、昨今の局地的豪雨による河川の氾濫や大規模都市施設への浸水被害が後を絶たない。このような災害に備え、国内でも有数の巨大な地下街が存在する大阪駅前梅田地区においても地下街を含めた大規模な避難計画を立案する必要性に迫られている。近年、避難計画の立案に際し避難シミュレーションを行うことが必須となりつつある。避難シミュレーションを実行するには、避難者の初期配置を決定する必要があるが確定した方法があるわけではなく、均等に避難者を分散させるといった方法がとられているのが現状である。しかし、空間内での人の分布は様ではなく偏っていることが多い。また、スマートシティの発展により、人の分布を把握してスムーズな避難につなげるといった技術の可能性も考えられる。

本研究は、大阪市が管轄する大阪駅前梅田地区の地下街に設置された監視カメラの映像から、顔認識などの個人を特定する情報を用いず、ややマクロなレベルで空間内の歩行者数を推定する方法を提案し、検証を行う。関連する研究として林田ら²⁾は、カメラの死角になっている場所の歩行者数を推定する方法を提案しているが、撮影された範囲の歩行者自体は手動でカウントされている点で本研究と異なる。

また、同映像の色空間をRGBだけでなく、非線形色空間であるHSV色空間とL*a*b*色空間への変換を行い、歩行者数の推定精度の検証を行う。

2. データの概要

本研究では、図1に示す箇所に設置された大阪市が管理する監視カメラの撮影データを用いる。監視カメラは25台稼働しており、2013年6月16日(日)と17日(月)の8:00, 10:00, 12:00, 15:00, 18:00の最初10分間の動画データを取得した。元々の動画は画像レート30 fps, QVGA, AVI形式のフォーマットで圧縮保存されたものであるが、解析を行うために動画から1秒毎の静止画を切り出して使用する。

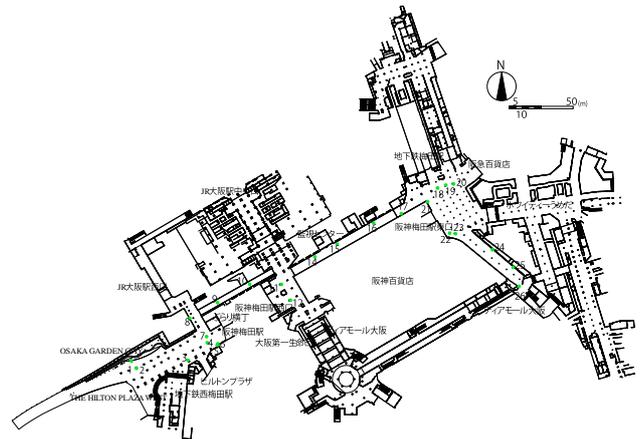


図1 対象となる梅田地下街の平面図³⁾

3. 歩行者数の推定方法

3.1. 歩行者流動によるカメラ映像の分類

人の流動の複雑さによって推定結果が変化すると考え、25台のカメラに映る歩行者の流動のパターンを分類した。本研究はまずは基礎的検討を行うために、流動方向が単純な直線の通路を撮影したカメラ9の映像を推定対象とする。

3.2. 色空間の変換による特徴量の取得

色空間の変換によって推定結果に変化が見られると考え、画像に対して複数の色空間から特徴量を取得した。まず、線形色空間として与えられるRGB色空間における特徴量R, G, B値を取得した。さらに、非線形色空間であり、色相(hue), 彩度(saturation), 明度(value)で与えられるHSV色空間における特徴量H, S, V値, 均質な色空間として色彩方面で広く使われており、「座標空間の距離」が色差として正当な指標になるとされるL*a*b*色空間における特徴量L*, a*, b*値を取得した。

3.3. 画像の類似性の評価

本研究では背景差分法を用いて人数の推定を行う。画像の計測範囲内に歩行者が映っていない画像を背景画像とし、それ以外の画像(歩行者画像)との類似度によって、約5×5メートル程度の局所的な対象範囲の歩行者数を推定する。地下街の映像は季節や日時を通じて屋外を撮影した映像よりは安定しているが、それでも周辺の店舗の開店等により、空間の状況が変化するので、その状況が変化す

る。このことから表1に示すように複数の背景画像を作成し、歩行者画像の撮影時間に対応した背景画像を用いる。

背景画像と歩行者画像の類似性は、両画像の同一画素の色成分の差で評価する。差が大きいほど背景とは異なるもの(歩行者)が映っていると判断できる。具体的に説明すると、両画像の画素における色成分としての要素について画素数×色成分の次元のベクトルとみなし、この2枚の類似性を画像のベクトル間のユークリッド距離の総和(SSD)とマンハッタン距離の総和(SAD)で表すことにする。RGB色空間においては、2枚の画像の各画素に対する色成分の差の総和を式(1), (2), (3)のようにR, G, Bとして与え、ユークリッド距離の総和は式(4)に示すように2点の各次元の要素同士の差の二乗の和、マンハッタン距離の総和は式(5)に示すように2点の各次元の要素同士の差の絶対値の和である。ここで、 $p \in P$ は画像の各画素、 s は歩行者画像、 t は背景画像、 $R_x(p), G_x(p), B_x(p)$ はそれぞれ画像 $x \in \{s, t\}$ の画素 p のRGB値を示す。HSV空間、 $L^*a^*b^*$ 空間においても以下に示す同指標、 $H, S, V, L, a, b, HSV_SSD, HSV_SAD, Lab_SSD, Lab_SAD$ を算出した。

$$R = \sum_{p \in P} |R_s(p) - R_t(p)| / 255 \quad (1)$$

$$G = \sum_{p \in P} |G_s(p) - G_t(p)| / 255 \quad (2)$$

$$B = \sum_{p \in P} |B_s(p) - B_t(p)| / 255 \quad (3)$$

$$RGB_SSD = \sum_{p \in P} \left((R_s(p) - R_t(p))^2 + (G_s(p) - G_t(p))^2 + (B_s(p) - B_t(p))^2 \right) \quad (4)$$

$$RGB_SAD = \sum_{p \in P} (|R_s(p) - R_t(p)| + |G_s(p) - G_t(p)| + |B_s(p) - B_t(p)|) \quad (5)$$

3.4. 推定モデルの作成

各歩行者画像における歩行者人数を0.1人単位で目視により計測を行い目的変数とする。まず各色空間別にSSDとSADを説明変数とする重回帰モデルと中間層が1層の階層型ニューラルネットワークを用いてモデルを作成し、推定モデル①とする。次に色空間をまとめ、各SSD, SADの計6変数を説明変数とする重回帰モデルとニューラルネットワークによりモデルを作成し、推定モデル②とする。また、各色空間における要素の差分値を正規化したもの($R, G, B, H, S, V, L, a, b$)を説明変数とし、属性選択を経て選

表1 時間による背景画像の変化

	期間	状態	背景画像
①	8:00 以降	画面右側の店舗のシャッターが降りている。	
②	10:00 以降	画面右奥のシャッターが上がり、光が漏れている。提灯は出されていない。	
③	10:00 以降	画面右奥のシャッターが上がり、明かりがつかうとしている。右手前の提灯は出されているが灯されていない。	
④	10:00 以降	画面右奥のシャッターは上がり光が漏れ、右手前の提灯は出されているが灯されていない。	
⑤	12:00 以降	④の状態から、右手前の提灯が灯される。	

択された変数による重回帰モデルとニューラルネットワークを推定モデル③とする。

3.5. 説明変数の選択

各推定モデルは複数の説明変数で構成されている。重回帰モデルのこれら説明変数間に多重共線性が生じている可能性や回帰係数が0であるような変数の存在が考えられることから重回帰モデルの再現性を高めるため、統計解析

ソフトウェアである JMP11⁴⁾を用いてステップワイズ法 ($p=0.05$)により変数の選択を行う。

さらに、これにより選択された変数間の多重共線性の有無を確認するため、それら変数における VIF (分散拡大係数) が 10 を下回るまで、最大の VIF 値を示す変数を 1 ステップにおいて 1 変数ずつ減らし残った変数を各推定モデルにおける説明変数として採用する。

4. 検証

4.1. 準備

本研究では歩行者の重なりがそれほど多くない日曜日の 10 時台のデータを用いて検証を行う。歩行者画像 971 枚において学習セットを作成し、交差検証にて精度の検証を行う。画像の指標の計算は、オープンソースの画像処理ライブラリ OpenCV 2.3.1⁴⁾と Visual C++ 2010 を用いてプログラムを実装して行い、得られた指標のデータから、フリーのデータマイニングソフト Weka3.6.10⁵⁾を用いて推定を行った。

4.2. 推定モデル①結果

推定モデル①は、色空間別に歩行者数を推定するものである。この場合、説明変数は各色空間において SSD と SAD の 2 変数であるが、重回帰モデルにおいて 3.5 節の手順で変数の選択を行った。

表 2 には、色空間別に選択した変数の回帰係数を示す。これら変数を説明変数として選択し、重回帰モデルを作成する。また、ニューラルネットワークにおいて多重共線性の影響は大きく出ないと考え、両変数ともに用い、歩行者数の推定を行った。

表 2 色空間別説明変数の回帰係数

RGB				
項	推定値	標準誤差	t 値	p 値
切片	0.406	0.0231	17.6	<.0001
SAD	1.43E-06	4.96E-08	28.8	<.0001
HSV				
項	推定値	標準誤差	t 値	p 値
切片	-0.23514	0.021014	-11.19	<.0001
SSD	0.000483	8.59E-06	56.18	<.0001
Lab				
項	推定値	標準誤差	t 値	p 値
切片	-0.1584	0.020671	-7.66	<.0001
SAD	0.000745	1.38E-05	53.89	<.0001

重回帰モデルを用いた場合の教師データの補正決定係数と、この重回帰モデルを検証データに適用した決定係数、中間層が 1 層 3 ユニットの階層型ニューラルネットワーク

の教師データの決定係数と、そのニューラルネットワークによる検証データの決定係数を表 3 に示す。ニューラルネットワークによる推定結果において比較的高い精度で歩行者数を推定できることがわかった。図 3 に、検証データにおける実測値と色空間別のニューラルネットワークモデルによる推定歩行者数の時系列変化を示す。下段に示された色空間別の歩行者数の実測値との誤差から、RGB 色空間においてニューラルネットワークを用いた推定結果が実データに比較的良好に追従した結果を得た。

表 2 色空間別歩行者数推定精度

	重回帰モデル		ニューラルネットワーク	
	教師	検証	教師	検証
RGB	0.648	0.645	0.825	0.754
HSV	0.765	0.763	0.773	0.690
L*a*b*	0.735	0.581	0.764	0.681

4.3. 推定モデル②結果

推定モデル②は推定モデル①における色空間別の変数を一つにまとめたモデルである。よって説明変数は 3 種の色空間において 2 変数所持していることから、6 変数有するモデルとなる。この変数から 3.5 節における手順を経て RGB_SSD, HSV_SSD の 2 変数でモデルを作成した。

表 4 から重回帰モデルを用いた場合の教師データの補正決定係数は 0.777 となった。この重回帰モデルを検証データに適用したところ、その決定係数は 0.771 と結果を得た。

また中間層 1 層の階層型ニューラルネットワークの教師データの決定係数は 0.772、その推定値での決定係数は 0.682 となり、今モデルにおいて重回帰モデルの方が高い精度で歩行者数を推定できていることがわかった。

図 4 に検証データにおける実測値、重回帰モデルによる推定値とニューラルネットワークによる推定歩行者数の時系列変化を示す。精度に表れているように重回帰モデルにて追従性のよさを確認できた。

表 4 選択 2 変数における歩行者数の推定精度

	重回帰モデル		ニューラルネットワーク	
	教師	検証	教師	検証
2 変数	0.777	0.771	0.772	0.683

4.4. 推定モデル③結果

推定モデル③は、歩行者画像と背景画像における各色空間の同一色成分による差分絶対値を説明変数にとるモデルである。つまり、説明変数は全色成分の 9 種で構成されたモデルとなっている。これら変数に対しても 3.5 節同様の変数選択を行い、各色空間の色成分の差分絶対値のうち R

と b を説明変数としたモデルを作成した。

表5から重回帰モデルを用いた場合の教師データの補正決定係数は0.771となった。この重回帰モデルを検証データに適用したところ、その決定係数は0.768と結果を得た。

またニューラルネットワークの教師データの決定係数は0.774、その推定値での決定係数は0.681となり、今モデルにおいても重回帰モデルにて高い精度で歩行者数を推定できていることがわかった。図5に、検証データにおける実測値、差分値 R 、 b^* による重回帰モデルとニューラルネットワークモデルによる歩行者数の時系列変化を示す。推定モデル②において追従できている箇所が推定モデル③では追従しきれていないなど、考察が必要であると考えられるが、精度にあまり差は出ていない結果となった。ニューラルネットワークを用いた場合、推定モデル①、②の中ではよくない結果となっている。

表5 推定モデル③による歩行者数推定精度

R,b	重回帰モデル		ニューラルネットワーク	
	教師	検証	教師	検証
	0.771	0.768	0.774	0.681

5. まとめ

本研究では、監視カメラの映像から個人を特定することなく、歩行者数を大まかに推定する方法を提案し、実用的な精度を得られる事を確認した。また、複数の色空間による特徴量を用い、多角的に検証することで、ある程度の精度において正確さを担保できたのではないかととらえる。結果として推定モデル②における重回帰モデルによる推定で最も高い精度が得られることを確認できた。しかしながら他のモデルと突出して精度の良いモデルとは言えな

いことから、他のカメラ映像においても検証が必要であると考えられる。今後は歩行者の分布を推定するために時空間の相関を考慮したモデルへの発展や、他の混雑した時間帯や複雑な歩行者流動を示すカメラでの検証を行う予定である。

謝辞

データを提供していただいた大阪市建設局、大阪駅前地下街総合共同防火管理協議会、大阪市立大学大学院谷口与史也研究室の皆様へ感謝いたします。また本研究の一部は、科学研究費補助金基盤研究(A)とCRESTの補助を受けて行われました。

[参考文献]

- 1) 大阪府：南海トラフ巨大地震災害対策等検討部会，第3回資料，2013.8，<http://www.pref.osaka.lg.jp/kikikanri/bukai/> (2014.4.6 確認)
- 2) 林田和人ほか：ニューラルネットワークを用いた駅構内の旅客概数の推定，2013年日本建築学会大会(北海道)学術講演梗概集，シミュレーション・数理計画，pp.85-86，2013
- 3) 合田祥子，谷口与史也，吉中進，瀧澤重志，“大阪駅前地下街の津波避難計画に関する研究”，日本建築学会大会学術講演梗概集（近畿）(2014)，pp.147-148.
- 4) OpenCV ホームページ，<http://opencv.org/> (2014.4.6 確認)
- 5) Weka ホームページ，<http://www.cs.waikato.ac.nz/ml/weka/> (2014.4.6 確認)
- 6) JMP ホームページ，<http://www.jmp.com/japan/> (2014.10.10 確認)
- 7) 高木尚哉，瀧澤重志：カメラ映像を用いた地下街歩行者の分布推定方法に関する研究，2014年日本建築学会大会(近畿)学術講演梗概集，図形処理・画像処理，11059

*1 大阪市立大学大学院工学研究科

*2 大阪市立大学大学院工学研究科 准教授

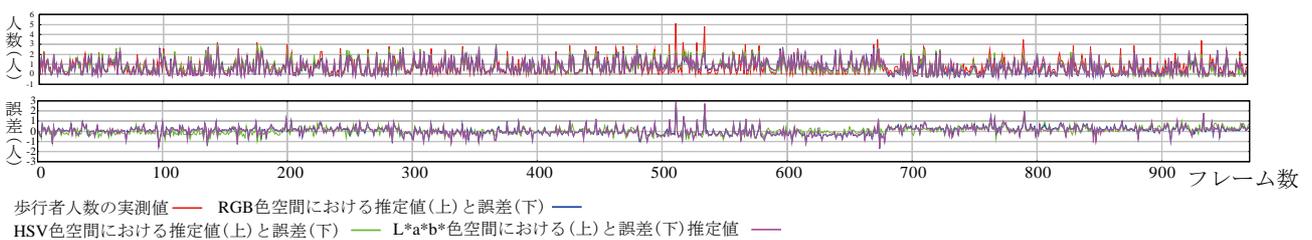


図3 歩行者数の実測値と推定モデル①における色空間別ニューラルネットワークによる推定値の推移

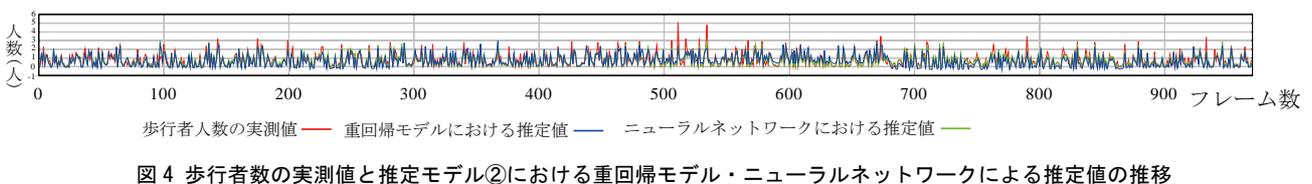


図4 歩行者数の実測値と推定モデル②における重回帰モデル・ニューラルネットワークによる推定値の推移

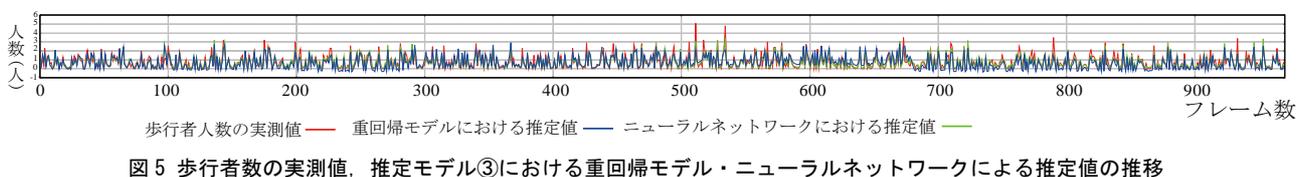


図5 歩行者数の実測値，推定モデル③における重回帰モデル・ニューラルネットワークによる推定値の推移