

映像解析による大規模避難誘導システムの設計

寺西 裕一¹ 地引 昌弘¹ 西永 望¹

概要：広範囲で複雑な構造を持つ大規模複合施設においては、災害時に施設内の歩行者に対する避難誘導を安全かつ迅速に行なうための対策が大きな課題である。我々は施設内に設置された複数のネットワーク接続されたカメラ映像を解析し、避難経路ごとの歩行者の移動速度を推定することで、広範囲の避難状況を把握するとともに、施設内の場所ごとに適切な経路を避難経路として歩行者へ案内し誘導するシステムを提案する。本研究では、映像解析技術として顔解析技術を適用した提案システムをテストベッド上に構築し、実際の大規模複合施設内で取得した映像データセットを用いてその実現可能性を検証した。

Design of a Large-scale Evacuation Navigation System Utilizing Video Analysis

YUUCHI TERANISHI¹ MASAHIRO JIBIKI¹ NOZOMU NISHINAGA¹

1. はじめに

大型化し、広範の複数の建物にわたる近年の大規模複合施設において災害が発生した場合、歩行者が危険な状況にある場所から避難する際、特定の経路の混雑によって避難が遅延してしまう危険がある。たとえば、南海トラフ地震では、津波によって大都市の中心部が2メートル程度浸水するとの予測もある[1]。この場合、津波が到達するまでの間に、数万人規模の歩行者を安全な場所へ避難させる必要がある。しかし、大規模複合施設では、ある避難経路に容量以上の数の歩行者が集中してしまう状況が容易に生じ得る。このような環境下では、混雑によって歩行速度の低下や滞留が生じて避難完了が遅延し、要求される時間内に完了することができなくなる可能性がある。

上記事態を避けるには、施設内の歩行者の避難状況をリアルタイムに把握し、避難計画にフィードバックできることが望ましい。たとえば、混雑している経路があれば、歩行者を迂回させ、滞留を緩和するよう誘導すれば、避難完了を早めることができると考えられる。しかし、大規模な施設においては、目視等により人手で混雑状況を調査する方法では、多くの人員を割く必要がある。また、避難先までの距離が長くなるため、利用しようとしている避難経路の先で混雑しているかどうか等を、目視のみで把握することは困難である。

そこで本研究では、近年技術進歩がめざましい映像解析技術、とくに、人物照合技術を用いて大規模施設内の歩行者の移動経路や移動時間等を人の流れに関する情報として取得し、場所ごとに適切な避難経路へ歩行者を誘導する大規模避難誘導システムを提案する。映像解析のもとになる映像は、施設内の複数箇所に配備され、ネットワーク接続

されたカメラを用いて取得することを想定する。映像解析処理は負荷が高い処理であるため、提案システムとしては、複数の計算機を同時並列で利用し、リアルタイムに実行可能とする設計とした。従来、映像解析処理を用いて閉空間の混雑度を取得する方法[2]、映像解析結果の避難誘導への適用可能性[3]に関する研究はなされてきたが、映像解析処理技術の大規模複合施設を想定した避難誘導システムへの応用についての検討は、従来十分になされていない。とくに、人物照合を用いて人の流れに関する情報を実用上どの程度取得できるか、避難誘導においてどの程度有用性があるかについて、これまで明らかとはなっていない。

我々は、人物照合において既に定評がある映像解析技術の一つである顔解析技術を適用した提案システムをテストベッド上に構築し、実際の大規模複合施設内で取得した映像データセットを用いて、人の流れに関する情報をどの程度の精度および性能で取得可能となるかを検証した。また、得られた人の流れに関する情報を元に避難誘導を行なうシミュレーションを行ない、大規模施設における有用性を確認した。

2. 映像解析を用いた避難誘導システム

2.1 避難誘導システムの動作

図1は、提案システムの動作の概要を示している。提案システムでは、対策部門における避難計画立案と現場の防災担当者(避難誘導者)への指示を支援するため、避難誘導に必要な次の人の流れに関する情報を取得可能とする。

- 領域内の歩行者数
- 移動経路と移動時間
- 歩行者の属性

指定領域の歩行者数は、避難先など特定の領域の人数の変動を把握する目的に利用する。対策部門は、領域面積に対する滞在人数が、許容密度を超過する場合に、警告したり、

¹ 情報通信研究機構
NICT Koganei, Tokyo, Japan

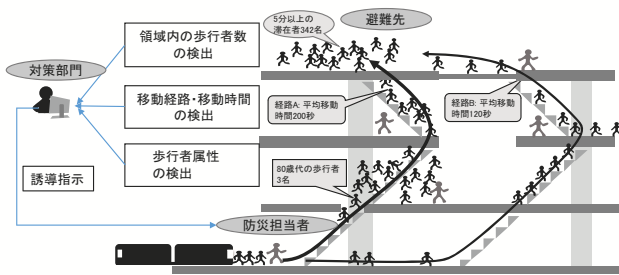


図 1 避難誘導システムの動作

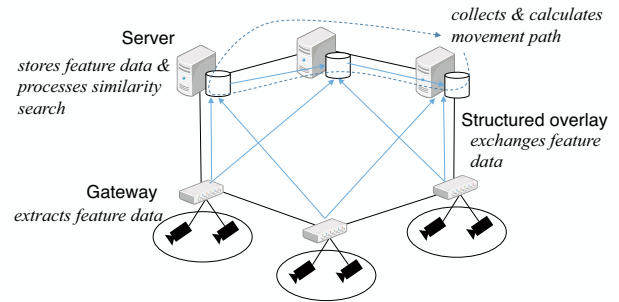


図 2 映像解析システムの基本構成

別の避難先を指示したりするなどの対策を行なう。また、歩行者の移動経路と移動時間を得ることで、利用されている避難経路、およびその混雑状況を把握する。対策部門は、避難完了を早めるため、混雑している経路があれば、迂回するよう現場の避難誘導者へ誘導指示を行なう。歩行者の属性は、年齢や性別等の情報であり、避難行動の支援が必要な高齢者等を把握し、優先的に誘導する対策等の必要性を判断するために用いる。

2.2 人の流れに関する情報の取得法

提案システムでは、2.1 で示した人の流れに関する情報を、施設内に設置され、ネットワーク接続された複数のカメラの映像を解析することで取得する。

領域内の歩行者数は、映像内の顔領域を検出する方法 [4] や人物領域を検出する方法 [5] を用いて、歩行者として検出された数を数えることによって得ることができる。

移動経路およびその移動時間を得る方法としては、映像内の歩行者の動きを追跡する方法 [6] がある。しかし、この方法では全領域の歩行者の移動経路や移動時間を得るには、カメラを大規模施設内に死角なく設置しなければならず、広域にわたる大規模複合施設全体の状況を把握するには、相当数のカメラが必要となってしまう。また、観測領域以外の場所において滞留が発生し、避難経路の移動時間が長くなることを検知できない。

したがって提案システムでは、観測領域を共有しない複数のカメラに映った映像間で人物照合を行うことで、移動経路およびその移動時間を得る方法を採用する。カメラの視野は通常固定されており、歩行者が映ったカメラが分かれば、どの位置に歩行者が居たかが分かる。すなわち、歩行者毎に、いずれのカメラに、どの時刻に映ったかを記録していけば、移動経路および移動時間を得ることができる。複数カメラ間における人物照合に関する研究は数多く、近年の計算機技術の発達に伴ない実用化が進んでいる分野である。例えば、顔解析 [7]、歩容解析 [8]、[9] を用いる方法の他、服装の色や模様を用いる方法 [10] などがある。人物照合が正しく行なえれば、たとえば、カメラの観測点で囲まれた領域内に入り出した人を計測することができ、領域内の歩行者数の推定も可能となると考えられる。

歩行者の属性については、顔画像から年齢や性別を推定する方法 [11]、歩行特性から年齢を推定する方法 [12] などがある。

本研究では、歩行者検出、人物照合、および、歩行者属性を取得できる既存の映像解析技術を利用する前提とする。情報を取得する映像解析方法そのものについては、多数の既存研究があるため本研究のスコップ外とする。また、システムとしては特定の映像解析技術に依存しない、汎用性のある構成とする。

2.3 映像解析システムの基本構成

図 2 は、提案システムにおいて、映像解析を用いて人の流れに関する情報を得る映像解析システムの基本構成を示している。

既存の映像解析技術では、同一人物がどうかの判定を、映像から人物ごとに抽出される特徴データ (feature data) 間の照合処理によって行なう。従来の同様のシステムでは、複数のカメラから得られる映像データ上の歩行者の検出、特徴データの抽出を単一サーバで実行する形態が取られている [13]。しかし、この構成では、多くのカメラを処理対象とする必要がある環境では、ネットワーク負荷、処理負荷ともに集中し、データの伝送や処理に時間がかかってしまう。

提案システムでは、カメラを接続するネットワークエッジに計算リソースを配置し、映像中の歩行者の検出および特徴データを抽出する処理を行なう。この特徴データの数が、検出された歩行者数に対応することになる。図におけるゲートウェイ (Gateway) がこの計算リソースに対応する。ゲートウェイは複数配置し、一つのゲートウェイは一定数のカメラのみを担当する。一般に、特徴データのサイズは、映像データのサイズよりも小さいため、この構成によって、ネットワークの負荷を下げるができる。映像データは特徴データの抽出処理後は不要となるため、消去する。

特徴データは、一般に多次元のベクトルデータとなり、その照合処理においては、距離 (類似度) の計算が必要となる。特徴データの照合処理は、形式的にはある特徴データ $q \in \mathcal{D}$ 、に対し、1 つの $a \in \mathcal{S}$ 、 $\mathcal{S} = \{x \in \mathcal{D} | d(q, x) \leq \mathcal{T}\}$ を得る処理である。 \mathcal{D} は特徴データのドメイン空間、 d はドメイン空間 \mathcal{D} における距離、 \mathcal{T} は、特徴データ同士が照合されたときのみならず距離のしきい値である。この照合処理は、計算負荷が高く、照合対象となる特徴データの数が多ければ時間がかかる。提案システムでは、この照合処理を複数のサーバを用いた並行処理によって実行する。

担当サーバのいずれで処理を実行するかは、特徴データのドメイン空間 \mathcal{D} をサーバの台数で分割して決定する。特徴データの照合をどのサーバに担当させるかは、次の通り決める。まず、次の条件を満たす \mathcal{D} から 1 次元空間 I への写像 f を規定する。

$$\forall x \in \mathcal{D}, \forall y \in \mathcal{D}, |f(q) - f(x)| < |f(q) - f(y)| \Rightarrow P(d(q, x) < d(q, y)) > P(d(q, x) > d(q, y))$$

P は確率である。上記はすなわち、 f によって写像される 1 次元空間 I 上の距離が小さいほど、ドメイン空間 \mathcal{D} における距離が小さくなる確率が高くなる必要があるということである。そのような条件を満たす f は解析技術に依存するが、例えば、観測された歩行者の身長、動き幅、推定された年齢などは、正確であれば値に近いほど特徴データとして照合される可能性は高く、それらを特徴データから抽出する手続きは f として利用し得る。

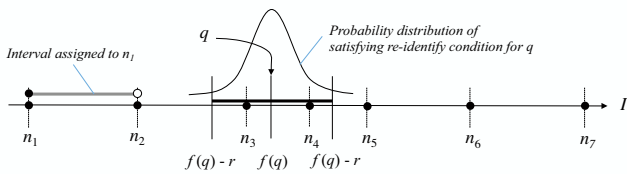


図 3 サーバ間の担当領域の分担

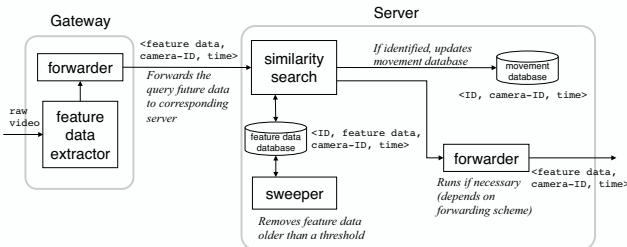


図 4 映像解析システムのモジュール構成

一次元空間 I をサーバ数に分割するため、各サーバには I 上の値 $k \in I$ をキーとして割り当てる。各サーバは、 I 上の $[k..l]$ (l は k より大きい最小のキー) を担当範囲とする。図 3 において、 n_1 がキーとして割り当てられたサーバの担当範囲は、 $[n_1..n_2]$ である。

特徴データ q に対応する $f(q)$ は、 I 上の近い値に対応する特徴データが照合する可能性が高い。提案システムでは、照合成功率を高く保ちつつ全サーバで照合処理を行なうことを回避する方策として、特徴データ間の距離が T となる可能性が高い I 上の範囲 (例えば、90%) を、あらかじめ計測しておき、その範囲に対応するサーバのみで照合処理を行なうものとする。この I 上の範囲を r とするとき、 $[f(q) - r \dots f(q) + r]$ の範囲を担当するサーバへ q を転送し、各サーバにおいて照合処理を行なう。 r の範囲内で照合されなかった場合は、 $f(q)$ を担当するサーバに特徴データを保存する。各サーバに割り当てるキーは、計測に基づき各サーバが均等に特徴データを保持するよう割り当てる。特徴データは、テキストデータ等と比べるとサイズが大きく、単純にロードバランスを用いてサーバへ振り分ける方法では、大量の歩行者が対象なる場合、ボトルネックが生じ得る。提案システムでは、担当サーバへの特徴データの転送を通信トラヒックが一部に集中することなく行なえるよう、ゲートウェイとサーバから構成される構造化オーバーレイを用いたデータ転送を行なう。

2.4 映像解析システムの実装

図 4 は、我々が実装した映像解析システムのモジュール構成を示している。入力として、映像データがカメラからゲートウェイに送信される。特徴データ抽出モジュール (feature data extractor) が特徴データを映像データから抽出する。特徴データは、属性情報 (年齢、性別) を含む。映像データは、映像内の歩行者すべての特徴データの抽出が完了した時点で消去される。抽出された特徴データは、データ転送 (forwarder) プロセスによって、 q として担当サーバへ送られる。 $f(q)$ を含むキーを持つサーバへのデータ転送には、構造化オーバーレイ実装の一つである Chord# [14] を用いた。転送データは、特徴データのほか、撮影されたカメラの識別子 (Camera-ID)、観測時刻 (time) を含むエントリとなる。サーバ上の特徴データベース (feature data database) には、各エントリに、ID 付与のうえ保持されており、 q との照合に成功するエントリがデータベース内に既にあれば、移動経路データベース (movement database) 上に、ID、Camera-ID、time、および属性情報の組が記録

される。存在しなければ、 $[f(q) - r \dots f(q) + r]$ の範囲内で次に $f(q)$ に近いキーを持つサーバへエントリを転送する。 $[f(q) - r \dots f(q) + r]$ の範囲に既に存在しない q のエントリは、新規 ID を持つエントリとして $f(q)$ を担当するサーバ上の特徴データベースに保存する。また、当該エントリの ID、Camera-ID、time、属性情報は、移動経路データベース (movement database) にも記録される。サーバ上では、掃除プロセス (sweeper) が常駐し、定期的の特徴データベースから古いエントリを削除する。

避難誘導システムとしては、移動経路データベースの情報をもとに、歩行者数、経路毎の移動時間、属性情報を得られるようにする。対策部門が、これらを元に避難経路を判断し、誘導指示を行なう想定である。その際に取得するのは、各歩行者に対応する移動経路の情報そのものではなく、時刻範囲、Camera-ID を指定した歩行者数 (記録されている特徴データ数)、時刻範囲、Camera-ID の組を指定した移動経路毎の移動時間の平均、および、時刻範囲、Camera-ID を指定した属性情報毎の歩行者数である。

3. 評価

本研究で提案する避難誘導システムの実現可能性を評価するため、テストベッド JOSE [15] 上に映像解析システムを構築し、どの程度の精度および性能で人の流れに関する情報を取得できるかを検証した。

本評価では、歩行者検出、人物照合、および、歩行者属性を取得できる映像解析技術として、既に確立し、実用化がなされている顔解析技術を適用した。顔解析技術としては、精度が高いことで知られ、多くの適用事例がある Glory 社のモジュール [16] を用いた。

顔解析技術は、一定以上の品質、かつ、正面に近い歩行者映像が必要であるが、混雑した群集映像に対しても適用可能である。従来、顔解析技術を、静止状態あるいはそれに近い状態で個別に歩行者を撮影するカメラ映像を対象に、犯罪防止目的や本人確認目的等で人物照合を行なっている事例は多く、高い照合性能が得られている [17]*1。しかし、同技術を大規模複合施設等の実際の施設内に設置されたカメラのライブ映像を対象に、群集の人の流れに関する情報の取得に適用した場合、どの程度の精度および性能が得られるかについては、未知数であった。大規模複合施設においては、カメラを設置できる位置に物理的な制限が生じる場合があり、どのようなカメラ視野が適しているかについても検証が必要である。

3.1 データセット

映像データとして、実際の大规模複合施設内 (大阪ステーションシティ; OSC) において、2014 年および 2015 年に、営業時間外の深夜撮影にて取得した映像データセット (以下、OSC データセット) を用いた。OSC データセットの概要は表 1 に示す通りである。撮影対象となった被験者は、映像データ取得の同意を得たエキストラのみであり、一般の通行者は含まない。本データセットは、被験者が避難行動を模擬して動作した群集映像を含んでおり、提案システムの評価に適している。また、一部の領域については、画面内に同時に映る人数に変化を持たせた。本評価では、映像品質として、10fps の HD 画像を用いた。

対象データとして、顔解析に適用できるカメラ視野を持つ 9 つのエリアの映像を用いた。図 5 は各エリアからピックアップした典型的なカメラ視野の例である。各エリアには、数台~数十台のカメラが設置されている。ただし、各

*1 実際、同モジュールを用いて OSC データセットの映像から静止画を切り出して逐一処理を行なった場合、後述の R_{op} が 0.97 (Area-A) と高い値となる。



図 5 エリアごとのカメラ視野の例

パラメータ	値
のべ参加者数	97
カメラ台数	92
エリア数	9
視野範囲	3m×3m ~ 7m×7m
視野内人数	1 - 40
画質	HD, VGA
フレームレート	10fps, 30fps
総時間	253 時間

エリア	カメラ視野	照度
エリア A	標準的	高
エリア B	標準的	中
エリア C	標準的	低
エリア D	遮蔽物	高
エリア E	遮蔽物	中
エリア F	遮蔽物	低
エリア G	アングル低	高
エリア H	アングル低	中
エリア I	アングル低	低

画像は頭部領域にぼかし加工、および、誌面上の見やすさのための明るさ調整加工を施している。また、それぞれ、カメラ上でゲイン調整などが行なわれており、実際の肉眼で見たときの明るさや見え方とは異なる。

表 2 は、それぞれのエリアの諸元である。各エリアは、大まかには、障害物が無く良好なカメラ視野が得られるエリア A~C、ドア、看板、柱等の遮蔽物がカメラ視野に存在するエリア D~F、カメラ設置角度が浅く（アングル低）、歩行者が他歩行者に隠蔽されやすい視野エリア G~I に分類される。また、データ取得は、深夜の時間帯に営業時間帯と同様の照明を確保して実施しているが、営業時間帯と同様の照明であってもエリア C, F, I はそれぞれ屋外に近

い通路の領域であり照度は低い（250lx 以下）。エリア A, D, G は商店街または改札付近であり比較的照度が高く（照度 350lx 以上）、エリア B, E, H はそれらの間の照度を持つ領域である。

3.2 精度評価

OSC データセットを用いて歩行者数、移動経路、移動時間、歩行者属性を、それぞれどの程度の精度で取得可能であるかを検証した。

3.2.1 評価指標

精度評価において、まず、領域内の歩行者数については、カウント率 (Count Rate) R_c 、観測可能カウント率 (Observable Count Rate) R_{oc} を評価指標とする。 R_c, R_{oc} の定義は次の通りである。

$$R_c = \frac{m}{n}, \quad R_{oc} = \frac{m}{s}$$

m は観測された人数、 n は、実際的人数である。 s は、映像を目視で確認できた人数である。 R_c は、実際に通った人数のうち、計測できた人数の比率であり、映像解析を用いた人数計測の精度を表す指標となる。例えば、隠蔽物が多くカメラ視野に映らない歩行者が多い場合、低い値となる。また、 R_{oc} は、カメラ視野内に現れた人数のうち、観測できた人数の比率であり、映像解析技術そのものの精度を表す指標となる。例えば、隠蔽物が多くカメラ視野に映らない歩行者が多い場合であっても、カメラ視野に映った歩行者を検知できていれば高い値となる。 R_c, R_{oc} ともに 1 に近いほど観測の精度が高いことを表す。

避難経路の検出については、経路検出率 (Path-detection Rate) R_p 、観測可能経路検出率 (Observable Path-detection Rate) R_{op} を評価指標とする。 R_p, R_{op} の定義は、次の通りである。

$$R_p = \frac{m_p}{n_p}, \quad R_{op} = \frac{m_p}{n_p R_{oc}}$$

m_p は観測された経路の通過人数, n_p は、実際に通過した人数である. R_p は、映像解析を用いた計測の精度を表す指標, R_{op} は、映像解析技術そのものの精度を表す指標となる. とともに 1 に近いほど観測の精度が高いことを表す.

移動時間については、得られた経路の移動時間の観測結果を用いて、 t 分布による平均区間推定 (95%信頼区間) を行ない、推定区間が実際の移動時間の平均を含んでいるか否か、および、推定区間の長さを指標とした. 推定区間が実際の移動時間の平均を含んでいる場合、推定区間が短いほど精度が良いと言える.

属性については、推定年齢の 10 歳ごとの頻度分布を推定性別ごとに求め、実際の頻度分布とのユークリッド距離 d を評価指標とする. d は属性の分布推定における誤差が大きいほど大きな値となり、0 に近いほど正しく属性を推定できていることを表す.

3.2.2 精度評価の結果

歩行者数

図 6 は、エリアごとの R_c, R_{oc} をそれぞれ示している. カメラ視野や照度によって観測の精度には変化が見られた. まず、エリア A~C のうち、照度が高いエリア A, B では、 $R_c = 0.6, R_{oc} = 0.7$ 程度となった. カメラ視野内での歩行者の露出時間が短時間である場合、歩行者が映っている領域の切り出しができない(歩行者の切り出し失敗). これらのエリアに遮蔽物は無いが、人の影に隠れたりよそ見をしたりしたために頭部が解析処理に十分な時間映らなかった場合や、埃等がカメラに付着して映像の画質が十分でなかった場合などに、歩行者の切り出し失敗が起きると考えられる. R_{oc} のほうが高い値となっているのは、通過したにもかかわらず一度もカメラ視野内に現れなかった歩行者が存在したためである. 一方、照度が低いエリア C では、 R_{oc} が 1 を越えている. これは、これは歩行者が一旦視野から消えてたのち再度現れた場合に、照度が低いために同一の歩行者と認識できず、実際の数よりも多く観測されてしまったこと(視野内照合の失敗)が原因と考えられる.

遮蔽物があるエリア D~F のうち、エリア D は比較的精度は高く、 $R_c = 0.7, R_{oc} = 0.85$ となった. 一方、エリア E は、 R_c, R_{oc} ともに 1 を越えた. これはドアなど遮蔽される領域が多いうえ照度も十分でなく、視野内照合の失敗が起きたためである. エリア F は、 R_c, R_{oc} ともに 0.5 程度であり、歩行者の切り出し失敗と視野内照合の失敗の双方が起きていると考えられる.

エリア G~I はアングルが低く、他歩行者による隠蔽が多く起きており、エリア G, H において R_c, R_{oc} が 1 を越えている. エリア I は、照度が低く歩行者の切り出し処理が十分に行なえない状況が起きていると考えられ、 R_c, R_{oc} ともに低い.

図 7 は、遮蔽物がなく、照度も高いエリア A において、カメラ視野に同時に現れる歩行者数を変化させて観測した場合の R_c, R_{oc} を示している. 人数が増えるに従い、 R_c, R_{oc} ともに低下した. これは人数が増えるに従って人の影に隠れる状況が生じやすくなったため、歩行者の切り出し失敗が起きやすくなったことが要因と考えられる. ただし、歩行者の切り出し失敗は、視野内の人数が 30 名から 40 名に増えてもあまり変化しておらず、人の影に隠れることによる影響は概ね一定に抑えられている.

以上から、カメラの設置条件が十分であれば、混雑した状況において、把握可能な歩行者数は、実際の概ね 6 割から 7 割程度となることが分かった. したがって、避難時等の混雑した状況下で、人数を厳密に計測する用途で用いるには、顔解析技術を用いる方法は適しておらず、何らかの別の方策が必要である. ただし、歩行者数の割合や変動を知る等の用途に用いることができる可能性はある.

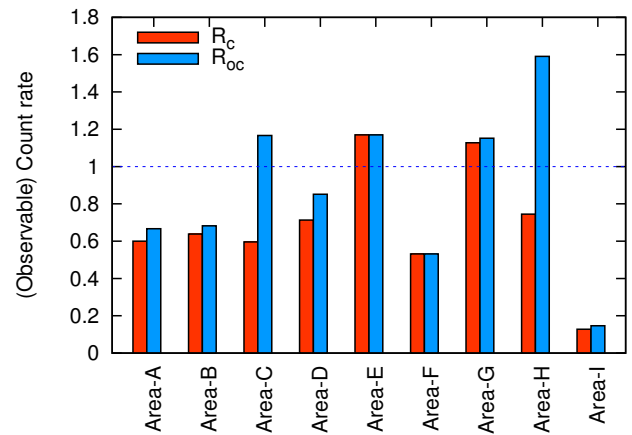


図 6 エリア毎の R_c, R_{oc}

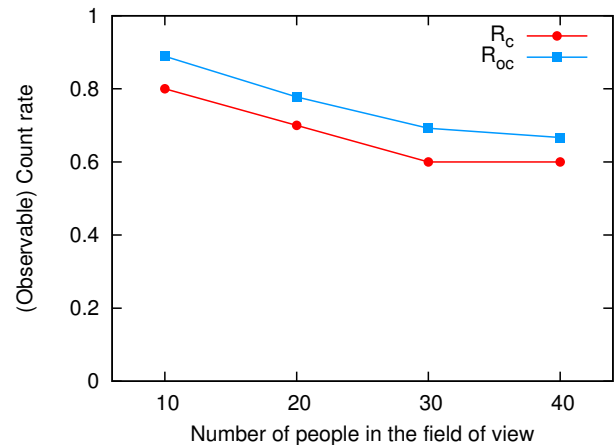


図 7 カメラ視野内の人数に応じた R_c, R_{oc} (エリア A)

移動経路と移動時間

OSC データセットにおけるいくつかの移動経路について、経路検出精度を評価した. 図 8 は、移動経路ごとの R_p, R_{op} をそれぞれ示している. FAR(False Acceptance Rate) は、誤受入率または他人受け入れ率とも呼ばれ、歩行者照合の際に誤った照合がどの程度発生するかを示している. 本評価では、 $FAR=10^{-2}$, $FAR=10^{-1}$ に設定した場合の結果を示した. $FAR=10^{-2}$ よりも $FAR=10^{-1}$ のほうが照合に成功する数は多くなるが、誤認識の可能性は上がる. $FAR=10^{-1}$ の場合、10 人に 1 人の割合で異なる人物を照合してしまう可能性がある.

まず、照度が高く、人物による隠蔽が少ないエリア A, D 内の経路では $FAR=10^{-2}$ の場合 $R_p = 0.3, R_{op} = 0.4 \sim 0.5$, $FAR=10^{-1}$ の場合、 $R_p = 0.4, R_{op} = 0.6$ 程度が得られている. R_{op} が 1 とならない要因としては、同人物が映った視野角度の違いやカメラ間の画質の違い等が考えられる. 照度が低いエリアを含む B, C 間の移動経路では R_p, R_{op} ともに低くなり、 R_p, R_{op} ともに 0.1~0.2 前後となった. 障害物があるエリア E 内の経路、および、他歩行者による隠蔽が多い G, H 間の経路では、前節で述べた通り、視野内照合の失敗により検出数が多く、 R_{op} の値が高くなっているが、経路検出数としての信頼性は低い.

図 9 は、遮蔽物がなく、照度が高いエリア A 内の経路にお

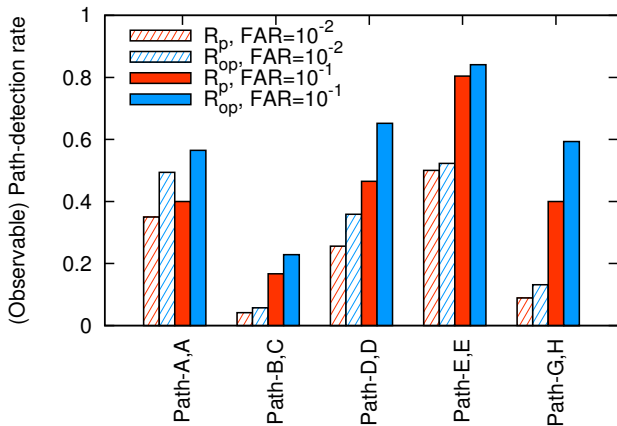


図 8 移動経路毎の R_p, R_{op}

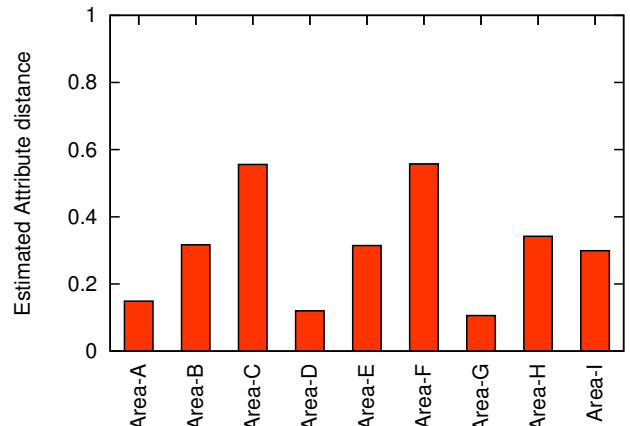


図 11 属性分布推定距離 d

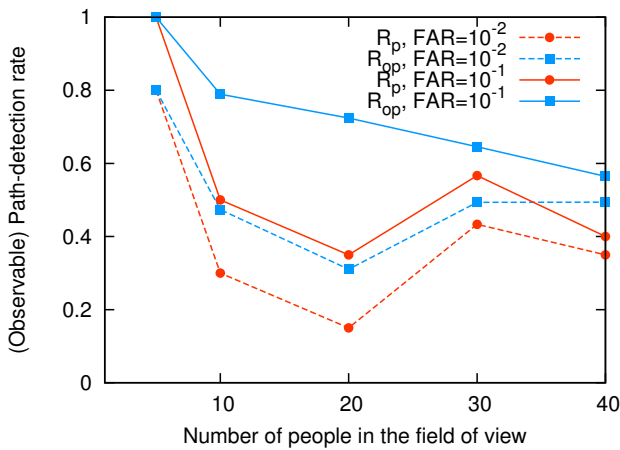


図 9 カメラ視野内の人数に応じた R_p, R_{op} (経路 A,A)

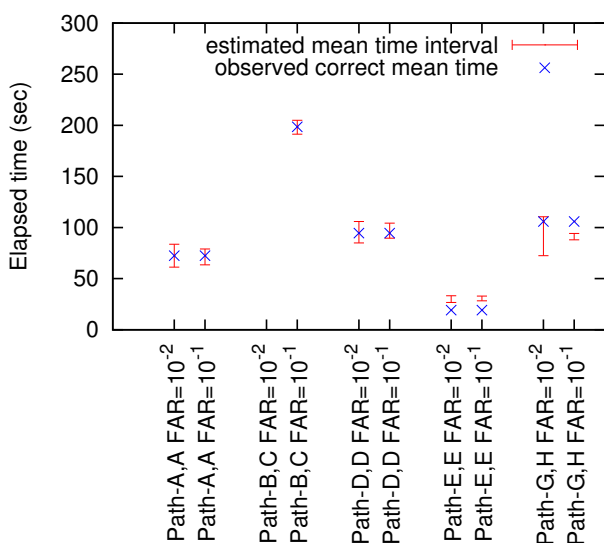


図 10 移動時間の推定

けるカメラ視野内の人数に応じた R_p, R_{op} である。全体としては人数が増えるに従い、 R_p, R_{op} ともに低下する傾向にあるが、 $FAR=10^{-2}$ のとき、概ね $R_p = 0.2 \sim 0.3$, $R_{op} = 0.4$ 程度、 $FAR=10^{-1}$ のとき、概ね $R_p = 0.4$, $R_{op} = 0.6$ 以上の値が得られた。

図 10 は、各経路における移動時間の平均区間推定結果と、実際の移動時間の平均を示している。エリア A 内、および、エリア D 内においては、移動時間の平均を、実際の時間に対して 10 秒前後の範囲で推定できている。E 内の経路では、推定区間が実際よりも長い時間に、G, H 間の経路では、推定区間が短い時間となった。これは、映像の前半に検出された歩行者と映像の後半に検出された歩行者が照合された、あるいは、その逆が起きたことが要因と考えられる。G, H 間の経路では、アングルが低いために、遠くの歩行者が映っているが、映像解析によって検出できる大きくなっておらず、実際の映像上で目視で確認できる移動時間よりも、検出時間が短くなる現象が見られた。

以上から、カメラの設置条件がそろえば、混雑した状況においても、1%の誤照合に抑える場合はおよそ 3 割、10%の誤照合を認めれば、概ね 4 割程度の経路が映像解析によって計測可能であることが分かった。これらは、数万単位の歩行者を対象とした統計処理を行なう上では十分な精度と考えられる。また、移動時間の平均は、10 秒前後の誤差で推定できることが分かった。

歩行者の属性

図 11 は、エリアごとの属性分布推定距離 d を示している。照度が高いエリア A, D, G では属性推定の精度は比較的高く、ほぼ正確に属性の分布を得られていた (距離 0.1 前後)。また、エリア B, C, E, F, H において、照度が下がるとして属性の取得精度が下がる傾向が読み取れる。エリア I は属性取得の精度が比較的高くなっているが、 R_c が低いために、サンプル数が少ない。総じて、照度が高ければ、検出歩行者に対する属性推定は、比較的高い精度で行うことができ、避難行動に支援が必要な高齢の歩行者が多いエリア等の傾向は得ることができると言える。

3.3 性能評価

次に、提案システムが、大規模複合施設を対象としたとき、どの程度の性能、すなわち、処理遅延で、人の流れに関する情報を取得可能であるかを検証した。表 3 は、JOSE テストベッド上に構築したシステムにおける評価の環境設定を示している。

評価で用いた映像解析の製品では、特徴データの一部に

表 3 性能評価環境設定・パラメータ

項目	値
出現頻度	10,000 人/分
f	推定年齢値 (連続値)
r	10
照合処理	FAR= 10^{-1} の場合の計測結果に基づく
特徴データサイズ	1M (bytes)
掃除プロセス閾値	60 分
疑似ゲートウェイ数	40
特徴データ抽出処理時間	10 秒 (計測に基づく)
オーバーレイアルゴリズム	Chord#
ネットワークプロトコル	TCP/IP
CPU(仮想マシンあたり)	1 core / 2.1 GHz
仮想マシン/物理マシン	1
メモリ	2G (bytes)
物理ネットワーク	1000BASE-T
OS	Ubuntu 12.04
Hypervisor	KVM 3.5

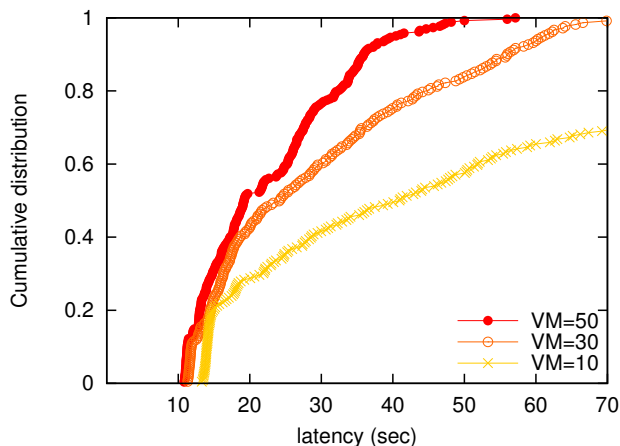


図 12 処理完了時間の累積分布

推定年齢値を連続値で含み、この値に近いほど照合に成功する確率が高い。したがって、 f として特徴データにおける推定年齢を割り当てた。また、 r 上の検索範囲 r は、事前計測の結果、90%以上照合に成功した範囲となった 10 を割り当てた。したがって、実際の照合成功率は $0.9 \times R_p$ となる。

人の出現頻度は、大規模複合施設を想定した性能を検証するため、最大 10,000 人/分とした。OSC データセットの被験者は最大でも 97 名であるため、ゲートウェイの動作を疑似的に再現するソフトウェアを用いて、ダミーの特徴データを生成させた。生成したダミーデータは、OSC データセットの処理結果と同等の推定年齢値の分布、データサイズ、照合成功率となるようにした。ゲートウェイ上で特徴データ抽出の処理実行にかかる時間は、実際の計測値に基づき、10 秒とした。

ゲートウェイ、サーバはそれぞれ、テストベッドの仮想マシン (VM) 上で動作する。物理マシンのネットワークインタフェースの負荷がボトルネックとならないよう、物理マシンあたりの VM 数は 1 としている。掃除プロセス閾値は、60 分とした。これは 60 分以上経過した特徴データを特徴データベースから削除するということである。その他のパラメータは、使用した製品、テストベッド環境の標準的な値である。

図 12 は、映像上に歩行者が出現し終わってから、その歩行者に対応する特徴データが生成され、照合処理が完了す

るまでにかかった時間の累積分布を示している。図は、全体で 100,000 人の歩行者の処理を実行した場合の結果である。処理にあたっては、照合処理を行なうサーバの数 (VM 数) を 10, 30, 50 に変化させた。図から分かるとおり、照合処理の並列度が上がるにしたがい、処理完了時間は短縮されている。これは、サーバ上の単位時間あたりの特徴データ受信数、および、照合処理数が減少するためである。サーバ台数 10 台では、全ての処理が完了するまでに 200 秒以上かかったが、サーバ台数 30 台の場合、約 70 秒、50 台の場合、約 60 秒と短縮された。

以上の評価から、構築したシステム構成において、50 台のサーバを用いた場合、概ね 1 分以内に照合処理が完了でき、人の流れに関する情報の元になるデータが得られることが分かった。

3.4 避難シミュレーション

以上の精度評価、性能評価の結果を踏まえ、複数路線の駅を含む大規模複合施設である OSC を模した 1 階から 3 階までの建築物モデルを対象として、避難シミュレーションを行なった。

歩行者のシミュレーションは、artisoc[18] のエンジンをベースにカスタマイズしたマルチエージェントシミュレータを用いて評価を実施した。本シミュレーションでは、災害として津波を想定し、2,500 人の歩行者全員を 2 階以上の高さにある特定の領域 (避難場所) へ避難させる動作を再現した。誘導方法として、次の 3 通りの方法で誘導を行ない、全員が避難完了するまでの時間を計測した。

- 方法 1(None): 誘導無し
- 方法 2(Simple): 静的な誘導のみ
- 方法 3(Proposal): 提案システムを用いた誘導

本シミュレーションは、方法 3 において、提案システムにより経路毎の移動時間を計測することの有効性の検証を主目的とした。提案システムにより得られる歩行者数、属性情報は、今回利用していない。

表 4 はシミュレーションの諸元である。方法 2、方法 3 では、シミュレーション環境上、誘導者が 22 箇所の別れ道において経路誘導を行なう。方法 2 では、ある避難経路に避難者の滞留が発生した場合、事前に効率的な避難が見込めると想定された経路 (広い通路や階段など) へ固定的に誘導する。方法 3 では、21 の観測エリアを設定し、各誘導者の位置から 10 箇所の避難先へ向かう合計 231 個の経路に対する移動時間を取得する。各誘導者は、得られた移動時間をもとに、所要時間が最も短い経路を案内する。誘導指示の周期は 1 分とし、経路の移動時間の差が 20 秒以上生じた場合に案内経路を変更する設定とした。これらの設定は、前章の性能評価結果 (処理遅延 1 分、移動時間平均の推定誤差 10 秒程度) に基づいている。避難先の避難者数が収容数を越えたかどうかは、目視等により誤差なしで判定する動作とした。方法 2、方法 3 では、避難先のキャパシティを越えた場合、別の避難先を目的地として誘導する。

方法 2,3 を、避難経路の効率的な利用という観点から比較すると、静的な誘導では、避難者の滞留が発生した時点で固定的な経路に誘導するため、避難に用いる経路毎の滞留に偏りが生じやすくなる。一方、提案システムによる誘導では、迂回する時間を含めた所要時間を動的に検出し誘導するため、避難に用いる経路間で避難者の滞留の偏りが小さくなることが期待できる。

図 13 は、シミュレーションの結果 (試行回数 10 回の平均値) である。横軸は避難開始からの時間 (秒)、縦軸は避難完了率である。提案システムを用いることで、何も行なわない場合と比べ、避難時間を平均でおよそ 5 分程度短縮できることが分かった。また、静的な避難誘導を行なう場合と比べても、4 分近く避難時間を短縮できている。

表 4 シミュレーション諸元

パラメータ	値
床面積	82,000 m ²
階段数	7
避難先	2F 2 箇所, 3 階 8 箇所 (階段含)
誘導者数 (方法 2,3)	22
計測経路数 (方法 3)	231
観測エリア数 (方法 3)	21
判定移動時間差 (方法 3)	20 秒
誘導指示周期 (方法 3)	1 分

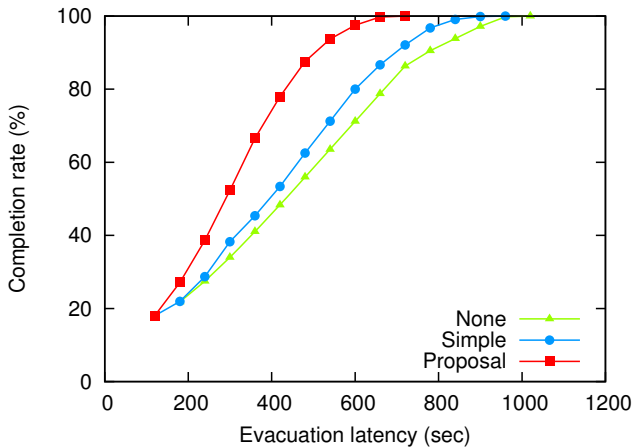


図 13 避難シミュレーションの結果

4. おわりに

本稿では、施設内に設置された複数のネットワーク接続されたカメラ映像を解析し、歩行者の人の流れに関する情報を得ることで、広範囲の避難状況を把握するとともに、施設内の場所ごとに適切な経路を避難経路として歩行者へ案内し誘導するシステムを提案した。また、提案システムの評価のため、映像解析技術として顔特徴量解析技術を適用したシステムをテストベッド上に構築し、実際の大規模複合施設内で取得した映像データセットを用いてその実現可能性を検証した。

検証の結果、現状の映像解析技術によって、大規模複合施設において、条件がそろえば、混雑した状況においても、概ね 6 割程度の歩行者数、4 割程度の経路把握が可能であること、移動時間の推定は 10 秒前後の誤差で可能であること、また、歩行者の属性分布推定が比較的高い精度で行なえることが分かった。またそれらの処理は、40 台のゲートウェイ、50 台の計算サーバを用いれば、施設全体で 1 万人/分の歩行者が観測される状況において、1 分以内で結果が得られることが分かった。さらに検証結果を踏まえた避難シミュレーションを行なった結果、2,500 人程度の歩行者を避難させる際、提案システムを用いることで、何も行なわれない場合と比べ、避難時間を平均でおよそ 5 分程度短縮できることが分かった。

他方、顔解析技術を人の流れに関する情報の取得に用いる場合、歩行者数の正確な把握が困難であることや、精度を保つことが困難となる環境がある等の課題があることも分かった。今後の課題としては、提案システムへの他の映像解析技術の適用や併用による精度の向上や、定常的に人の流れを観測する必要がある用途に向けた省エネルギーで長期安定運用可能なシステムの構成法の検討などがあげられる。

謝辞 本研究の実施にあたっては多くの方々にご協力いただいた。とくに、研究利用目的のデータ取得ポリシーに同意いただき、被験者として協力いただいたエキストラの方々、および、データ取得にあたり場所を提供いただいた施設管理者に深謝する。

参考文献

- [1] 内閣府, 南海トラフの巨大地震に関する津波高、浸水域、被害想定公表について, http://www.bousai.go.jp/jishin/nankai/nankaitrough_info.html (2016.01.22).
- [2] 庭川誠, 恩田寿和: 監視カメラ映像中の局所的な動き検出とイベント累積による時間的かつ空間的な混雑度調査, 電気学会論文誌 D, Vol. 124, No. 10, pp.1060-1066 (2004).
- [3] 目黒公郎: リアルタイム最適避難誘導システム構築に向けた基礎的研究, リアルタイム地震防災シンポジウム論文集, Vol. 1, pp. 111-116 (1999).
- [4] H. Wechsler, et al., eds.: *Face recognition: From theory to applications*, Springer Science & Business Media, Vol. 163 (2012).
- [5] G. Philip and G. Schneider: *A two-staged approach to vision-based pedestrian recognition using Haar and HOG features*, IEEE Intelligent Vehicles Symposium, pp. 554-559 (2008).
- [6] 布施孝志, 中西航: 歩行者拳動モデルを統合した人物自動追跡手法の構築, 土木学会論文集 D3, Vol. 68, No.2, pp. 92-104 (2012).
- [7] M. Bingpeng, et al.: Covariance descriptor based on bio-inspired features for person re-identification and face verification, *Image and Vision Computing* 32 (6), pp.379-390 (2014).
- [8] T. H. W. Lam, et al.: *Gait Flow Image: A Silhouette-based Gait Representation for Human Identification*, Elsevier Pattern Recognition, Vol. 44, pp. 973-987 (2011).
- [9] D. Muramatsu, et al.: *Gait-based Person Recognition Using Arbitrary View Transformation Model*, IEEE Trans. Image Processing, Vol.24, Issue 1, pp. 140-154 (2015).
- [10] 井尻善久, et al.: サーベイ論文: 視野を共有しない複数カメラ間での人物照合, 信学技報, PRMU2011-21 (PRMU), pp.117-124 (2011).
- [11] 滝本裕則, et al.: 姿勢変動に影響されない顔画像からの性別年齢推定, 電気学会論文誌, C 127.7, pp. 1022-1029 (2007).
- [12] 万波秀年, et al.: 歩容における性別・年齢の分類と特徴解析, 電子情報通信学会論文誌, D 92.8, pp. 1373-1382 (2009).
- [13] Y. Raja and S. Gong: *Scalable Multi-camera Tracking in a Metropolis*, Person Re-Identification. Springer London, Chapter 20, pp.413-438, 2014.
- [14] T. Schütt, et al.: *Chord#: Structured Overlay Network for Non-Uniform Load-Distribution*, ZIB-Report 05-40 (2005).
- [15] JOSE, <http://www.nict.go.jp/en/nrh/nwgn/jose.html> (accessed Dec. 1, 2015).
- [16] Face Recognition, <http://www.glory-global.com/groupinfo/technologies/others.html> (2016.01.22).
- [17] 「日本人出帰国審査における顔認証技術に係る実証実験結果(報告)」について. http://www.moj.go.jp/nyuukokukanri/kouhou/nyuukokukanri04_00044.html (2016.1.22).
- [18] artisoc, <http://mas.kke.co.jp/modules/tinyd0/index.php?id=11> (2016.01.22).