

深度情報を利用した頭上方向画像からの人物識別

香野 大地[†] 嶋田 和孝[†] 遠藤 勉[†]

[†] 九州工業大学大学院情報工学府

〒 820-8502 福岡県飯塚市川津 680-4

E-mail: [†]{d_kouno,shimada,endo}@pluto.ai.kyutech.ac.jp

あらまし 画像を用いた人物識別では、人物の顔やシルエットから特徴量を抽出し用いる手法が一般的に利用されている。しかし顔やシルエットを用いる場合、人物の重なりにより隠れが生じ、正しく特徴量を抽出できない場合がある。その結果、人物識別の精度が大きく低下する。この問題に対処するため、本論文では頭上方向に設置したカメラにより撮影を行う人物識別の手法を提案する。頭上方向から撮影することで設置の環境に関する制限も少なく、人物の重なりの問題も解消することができる。更に、本手法では人物を識別する際に深度情報を利用する。この深度情報を利用して (1) 人物の身長、(2) 人物の面積、(3) 人物の体型、(4) 深度ヒストグラムの 4 つを特徴量として抽出する。これらの特徴量を学習アルゴリズムである AdaBoost に適用することで人物を識別する。実験により、我々の提案手法が有効であることを示す。

キーワード 頭上カメラ、人物識別、深度情報

1. はじめに

現在、人物識別はセキュリティなど多くの場面に応用可能である。人物識別の手段には指紋や虹彩、筆跡、シルエット、画像を用いるものなどさまざまなものがあるが、画像を用いた人物識別には機器に非接触で実現できるため対象者にかかる負担が小さいというメリットがある。多くの研究者が画像を用いた人物識別の手法を提案しており、その中でも顔画像を用いた人物識別の手法が多く提案されている。Kanade [6] は鼻や目などを用いた手法を提案している。CLAFIC 法 [12] や EigenFace 法 [10] も顔から抽出した特徴量を用いて人物識別を行う有名な手法である。これらのような顔画像を用いた人物識別では、混雑している場所などで人物の重なりによって隠れが生じた場合、顔画像を正しく取得することが困難になる。

そこで本研究では人物の重なりにより隠れが生じるという問題を解決するために、頭上方向から撮影されたカメラ画像に着目した手法を提案する。図 1 は人物の重なりにより隠れが生じる問題を解決する方法を表している。正面から人物を撮影した場合では隠れが生じ、正しく顔画像を取得できないような場面においても、頭上方向から撮影したカメラ画像を用いることで隠れが生じることなく人物を識別することができる。この手法には人物の重なりによる隠れの問題を解消できること以外にもメリットがある。1 つ目は、対象者の視野に入らない位置にカメラを設置して画像を取得することができるため、撮影されているという心理的負担を軽減することができるということである。2 つ目は、顔を撮影する必要がないためカメラの設置場所に関する制限が軽減されるという

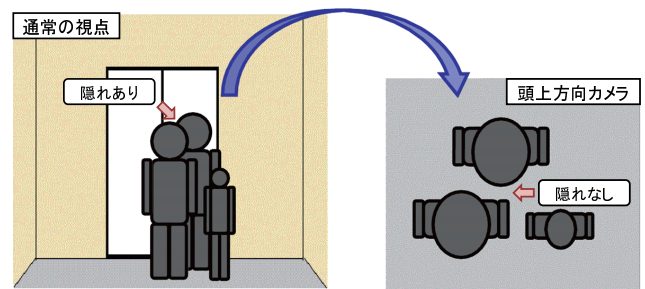


図 1 隠れによる問題と解決法

ことである。

中谷ら [7] はこれらの利点をふまえ、頭上方向からの画像を使用した人物識別手法を提案している。この先行研究では背景差分による人物抽出を行い、体型や髪型などの人物の特徴を得ることで人物を識別している。しかし、背景差分を用いた人物抽出では、衣服と背景の色が類似している場合や照明の条件によっては適切な差分画像が得られず、その結果、人物識別に失敗してしまうという問題がある。また、人物の顔のような特徴のあるパーツを使用できないため高精度な人物識別を行うには特徴量が不足しているという問題もある。

そこで、本研究ではカメラから被写体までの距離の情報を持つ深度情報を利用することで人物領域を抽出し、深度情報を持つ特徴量を追加する。深度情報を利用した

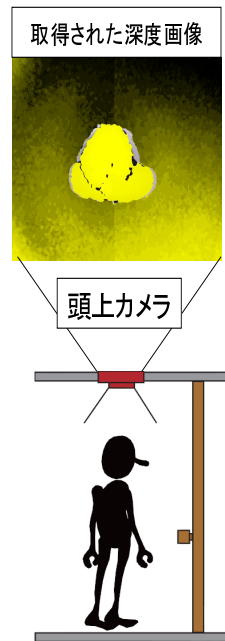


図2 カメラの設置と取得画像

場合、色や照明のような環境情報の影響を受けることなく人物と背景の切り離しが容易になるため人物領域の抽出精度も向上すると考えられる。また、頭上方向から撮影された人物画像中にはカメラから頭部までの距離、頭部から肩までの距離など様々な情報が含まれる。この情報は個人によって異なるため人物識別において有効な特徴になると考えられる。つまり、頭上方向からの人物識別において深度情報は人物を抽出する際の重要な情報であり、人物識別に用いる有効な特徴量の1つにもなる。図2は深度情報を含んだ頭上画像の取得例を表している。本手法ではこの深度画像から(1)人物の身長、(2)人物の面積、(3)人物の体型、(4)深度ヒストグラムの4つの特徴量を抽出する。このようにして得られた4つの特徴量を学習アルゴリズムであるAdaBoostに適用することで人物を識別する。

2. 関連研究

内田ら[11]はエレベータホールにおける頭上方向カメラを用いた人物検出手法を提案している。彼らの研究の目的は人物と車椅子の計測であった。Onishi&Yoda[8]は頭上方向からの画像を利用した人流の可視化手法を提案している。彼らの研究は頭上方向からの画像を利用しているが、目的は人物の計測と追跡であった。Iwashita&Stoica[5]は頭上方向からの画像を利用した歩行者の個人識別手法を提案しているが、彼女らは歩行者の識別に影を利用していた。本論文では人物の体型や面積などの4つの特徴量を用いている。福添ら[3]は人物の輪郭を用いた人物識別の手法を提案している。しかし彼らは頭上方向カメラを用いていないため、システムの有効性については画像中の隠れが問題となっている。本

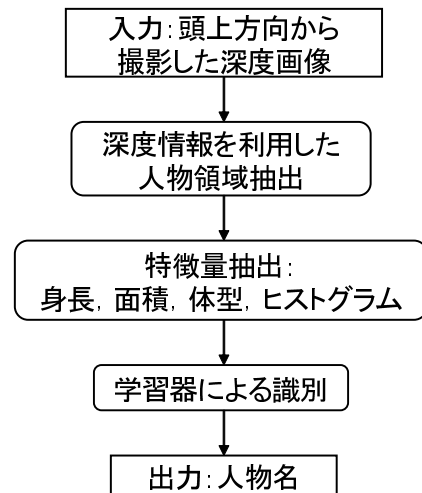


図3 提案手法の概要

手法では頭上方向からの画像を使用することで隠れの問題を解消することができる。更に、提案手法では深度情報を利用することで人物領域抽出の精度向上、特徴量の追加による識別精度向上を目指す。

3. 提案手法

この節では、提案手法について述べる。図3は手案手法の大まかな流れを表しており、大きく(1)深度情報を利用した人物領域の抽出、(2)特徴量抽出、(3)人物識別、により構成される。本論文では、深度情報に注目している。この深度情報はマイクロソフト社のKinectを利用することで取得する^(注1)。深度情報を利用することでより有効に人物領域を取得することができ、有効な特徴量を取得することができる。

3.1 人物領域の抽出

人物を識別するためには、撮影した画像から人物領域を抽出する必要がある。通常、人物抽出を行う際には背景差分を利用したものが多いが、照明の変化など環境による影響を強く受けてしまうことが問題として挙げられる。

この問題を解決するために深度情報を利用して人物と背景を切り離す。本論文で想定している状況では、人物の頭部、体、床が写るような画像が撮影される。これらの間にはカメラからの距離という重要な情報がある。そこで、人物の深度値と床の深度値の境界になる値を基準に切り離しを行うことで環境条件に頑健な人物領域の抽出が可能になる。図4は深度情報に基づいた人物領域抽出の例である。この図においては、色の明るさが深度情報を意味している。

(注1) : <http://www.xbox.com/en-US/Kinect>

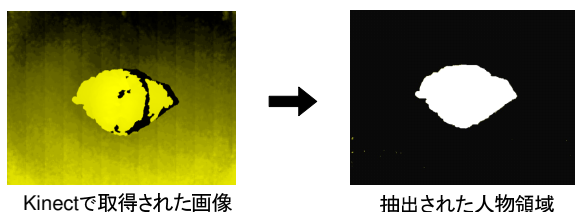


図 4 人物領域抽出例

3.2 特徴量抽出

この節では人物識別に用いる特徴量について述べる。本手法では、(1) 人物の身長、(2) 人物の面積、(3) 人物の体型、(4) 深度ヒストグラムの 4 つを特徴量を扱っている。

3.2.1 人物の身長

人物の身長は人物間の区別において最も重要な特徴量の一つになる。本手法では深度情報を利用することで人物の身長に関する特徴量を求める。カメラから最も近い位置にある人物の頭部までの距離は人物ごとに異なる。そこで、それぞれの深度画像においてカメラから最も近いピクセルまでの距離を抽出する。このようにして人物の頭部までの距離は得ることができるが、立ち位置による距離値の差という問題がある。図 5 はその問題の例を表している。この図において、下の値はそれぞれ Kinect から得られた距離を示している。この値から分かるように人物の立っている位置によって同一人物においても特徴量に大きく差が出てしまう^(注2)。

この問題を解決するために人物の立ち位置により出力値を正規化する。本手法では深度情報を利用することで概ね正しく人物領域を推定することができる。こうして求められた人物領域の中心位置を基に正規化することで同一人物内における誤差を減少させることができる。図 6 は正規化の例を表している。このようにして得られた値を人物の身長特徴とする。

3.2.2 人物の面積

人物の面積も人物識別において直感的で特有の特徴量となる。そこで本手法では人物領域における面積の値を利用する。この特徴量は 3.1 節において抽出された人物領域のピクセル数の総和を算出することにより求める。

3.2.3 人物の体型

人物の身体の大きさ^(注3)は人物識別において最も効果的な特徴量の一つになる。本手法では、体型特徴として

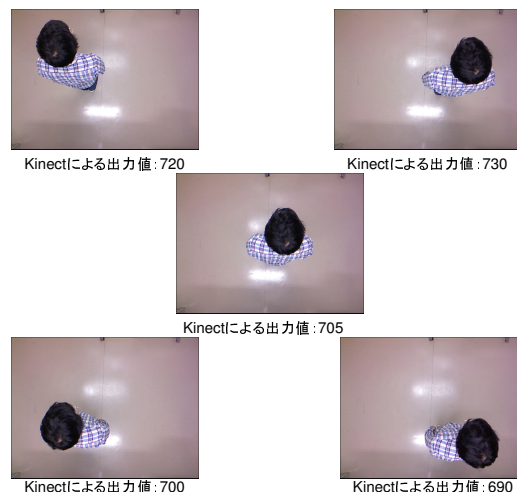


図 5 人物の立ち位置による Kinect の出力結果の違い

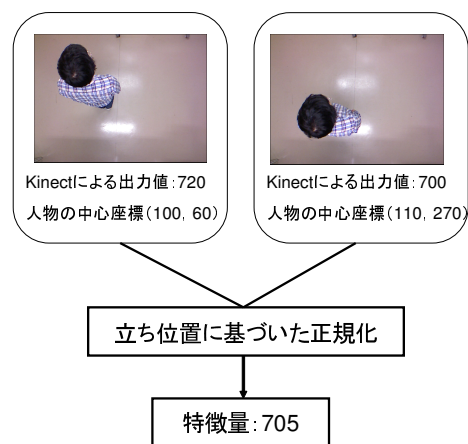


図 6 正規化の例

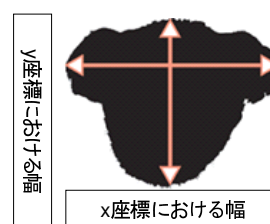


図 7 人物の体型特徴の取得例

x 座標と y 座標における幅の大きさを使用している。これらの値は人物領域画像の周辺分布から抽出する。人物領域の x 座標と y 座標において幅が最大となる値を使用する。図 7 は体型特徴の取得例を表している。

(注2)：本手法における身長特徴は Kinect から距離であり、実際の人物の身長ではないことに注意。

(注3)：本論文における人物の体型特徴は頭部や肩の領域の幅。

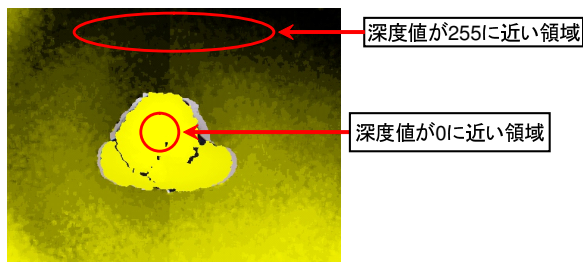


図 8 深度値の例

3.2.4 深度ヒストグラム

4 つ目はそれぞれの画像の深度ヒストグラムに基づいた特徴量である。深度画像において頭部領域、肩領域、背景である床領域までの間にはそれぞれ大まかな高さの差の情報が含まれる。

まず、Kinect カメラからの距離の情報を基に深度画像を生成する。ここでは深度画像の深度値は 0 から 255 までの範囲で表される。例えば頭部のようなカメラから最も近い部分は深度値が 0、逆に床のようなカメラから最も遠い領域は 255 に近い値が入ることになる。図 8 は領域による深度値の例を表している。次に 1 枚の深度画像において深度値 0 のピクセル数から深度値 255 までのピクセル数をそれぞれカウントする。これにより計 256 個の値を持つ深度値のヒストグラムが生成される。

図 9 は深度ヒストグラム生成の例を表している。0 に近い深度値のピクセル数は身長の高い、または身体の大きな人物ほど多くなる。床領域の深度値については、身長の高い人物が写ると 255 に近い値を持つピクセルが多くなる傾向がみられ、逆に身長の高い人物が写ると 255 に近い値を持つピクセルが少なくなる傾向がみられた。これは、深度画像が画像内におけるそれぞれの領域の距離の差を相対的に表しているため人物の身長が高くなるほど床領域が暗くなる傾向がみられたためであると考えられる。この特徴量はカメラからの距離、領域の面積など様々な情報を含む。深度値ごとにピクセル数をカウントしているため、人物領域から頭部と肩までの距離の情報や人物領域と床領域の深度値を比較することで頭部、あるいは肩と床までの距離の情報も取得できる。そこで、この深度ヒストグラム特徴には人物領域のみの深度値ではなく床領域の深度値も扱うことが有効となる。

3.3 分類器による識別

最後に抽出した 4 つの特徴量に基づく人物識別を行う。本論文では、機械学習アルゴリズムとして AdaBoost [2] を用いる。AdaBoost は機械学習アルゴリズムの中で多く用いられているものの 1 つであり、多数の弱識別器が

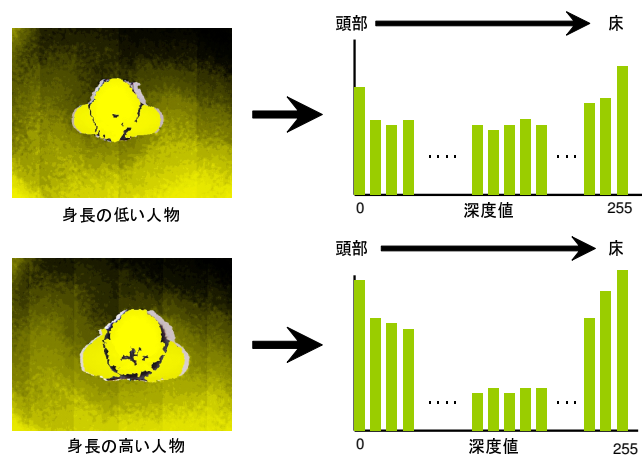


図 9 深度ヒストグラムの生成

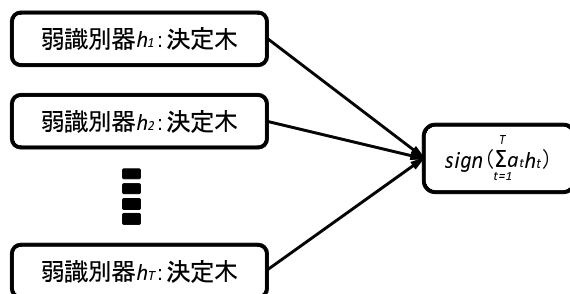


図 10 AdaBoost

ら 1 つの強識別器が構成されている 2 クラス分類器である。弱識別器としては C4.5 アルゴリズム [9] を用いる^(注4)。C4.5 は決定木により分類を行うアルゴリズムであり、これもまた機械学習アルゴリズムとして有名なものの 1 つである。図 10 は AdaBoost アルゴリズムの概要を表している。本論文では、実装にオープンソースソフトウェアである Weka を使用する^(注5)。

4. 実験

本実験では 8 名の被験者に対して人物識別を行った。画像サイズが 640×480 ピクセルのものを各被験者 20 枚、計 160 枚の画像を使用し、leave-one-out 法により識別精度を評価した。

4.1 実験結果

本実験では 4 つの特徴量の組み合わせすべてについて評価実験を行った。表 1 は実験結果の一部である。この表において、“身長”、“面積”、“体型”そして“深度”はそれぞれ 3.2 節における“人物の身長”、“人物の面積”、“人

(注4)：本論文で使用するフリーソフト Weka では“J48”。

(注5)：<http://www.cs.waikato.ac.nz/ml/weka/>

表 1 実験結果

特徴量	精度 [%]
身長	74.4
面積	40.6
体型	73.8
深度	90.6
身長+深度	94.4
身長+体型	91.3
面積+深度	92.5
体型+深度	90.6
身長+面積+体型	90.0
身長+体型+深度	93.1
all	93.1

物の体型”そして“深度ヒストグラム”特徴を示している．“+”は特徴量の組み合わせを表している．例えば，“身長 + 深度”は“人物の身長”と“深度ヒストグラム”を特徴量として用いたときの結果である．“all”は4つの特徴量をすべて用いたときの結果である．

精度は

$$\frac{\text{正解枚数}}{\text{全画像枚数}} \times 100[\%]$$

とした．表1からわかるように，“身長+深度”の特徴量を組み合わせた際に精度が最も高かった(94.4%)．“人物の身長”，“人物の体型”そして“深度ヒストグラム”または4つの特徴量を全て組み合わせた場合も93.1%の高精度を示した．この結果から特徴量を組み合わせることが効果的であることが確認できた．単体での特徴量としては“深度ヒストグラム”特徴量が他の特徴量と比較して高精度な結果が得られた．

4.2 考 察

本節では実験結果と本手法の有効性について考察する．一つ目に，識別精度の最も良かった組み合わせである人物の身長+深度ヒストグラムについて詳しく調べる．表2はこの2つの特徴量を組み合わせた識別結果の詳細を表している．“A”から“H”はそれぞれ被験者を表し，この表におけるそれぞれの被験者は身長順に並んでいる．つまり被験者Aの人物が一番身長が低く，被験者Hの人物が全ての被験者の中で一番背の高い人物である．誤識別は身長の近い人物間において発生する傾向がみられた．例えば，被験者Eと被験者Gの間で誤識別が3件発生している．この問題を解消するためには，身長とは本質的に異なる新たな特徴量を取り入れる必要があると考えられる．

続いて，単体の特徴量についての考察を行う．人物の身長特徴は単体での特徴量としては，精度が74.4%であり，身体の大きさを利用した面積特徴と比較して高精度な結果が得られた．この身長特徴は人物により差が出やすいため本質的には最も有効な特徴量の一つであるとい

える．この特徴量は画像内における人物の立ち位置の影響を受けてしまうこともあり正規化を行っていた．同一人物内の頭部までの距離において Kinect による出力では5~6cm 程度あった誤差を正規化することで2~3cm までに減らすことができた．正規化する前の身長特徴単体で人物識別を行ったところ，その精度は52.5%であった．すなわち，正規化によって，より正しい身長特徴が算出されたことで，21.9%の精度向上がみられたことになる．これは正規化の有効性を示している．身長特徴は組み合わせによる識別精度に貢献していることもあり，正規化をより高精度にすることで全体の精度向上に繋がると考えられる．

人物の面積特徴もまた人物領域において誤差の出やすい特徴量である．単体での精度は40.6%と最も低い結果となった．人物の身長特徴の正規化が精度向上に貢献したことから，面積特徴も身長特徴と同様に正規化することで精度改善が期待できると考えられる．

人物の体型特徴は人物の面積特徴と比較して立ち位置による影響を受けにくい特徴量であった．先行研究[7]でも同様の傾向がみられた．これは，人物領域において x 座標と y 座標の幅のみを利用するため，立ち位置による人物領域の大きさや体の傾きによる影響を抑えることができたからだと考えられる．しかし，体型特徴も姿勢などを考慮した正規化することによって更なる精度の向上が見込めるため，今後改善が必要である．

人物の面積特徴と体型特徴は衣服の変化による影響を受けてしまう．季節の変化により被験者の着る服の厚さが変化してしまうと画像内における人物の面積や体型も変化し，識別精度に悪影響を与えてしまう可能性がある．一方で，衣服の情報は人物識別におけるコンテキスト情報として最も効果的な特徴量でもある．Gallagher&Chen [4] は衣服の情報が画像中の人物を識別する際の有効性を示している．山口ら [13] は部分的に顔に隠れが生じた人物識別に対して衣服をコンテキスト情報として適用している．本手法においては，服装の情報をコンテキスト情報として追加することにより精度向上させることが今後の課題となる．

深度ヒストグラムは単体での利用において最も有効な特徴量であった．この特徴量はカメラから頭部，頭部から肩やカメラから床までなどさまざまな距離の情報を含む．つまり，深度ヒストグラムはさまざまな領域から求められた距離の差を含んでいる．深度ヒストグラムは深度値の範囲が0から255までの範囲に圧縮されているため絶対的な距離を表す身長特徴を組み合わせることで高精度な結果が得られたのだと考えられる．

5. ま と め

本論文では頭上方向から撮影した画像と深度情報を利用した人物識別の手法を提案した．頭上方向からの画像を利用することで人物の重なりによる隠れの問題とカメ

表 2 “身長 + 深度” の識別結果の詳細

被験者	A	B	C	D	E	F	G	H
A	20	0	0	0	0	0	0	0
B	1	19	0	0	0	0	0	0
C	0	1	19	0	0	0	0	0
D	0	0	1	18	0	0	1	0
E	0	0	1	0	19	0	0	0
F	0	0	0	0	0	20	0	0
G	0	0	0	0	3	0	17	0
H	0	0	0	0	1	0	0	19

ラが視界に入ることによる心理的拘束性の問題を解消することができた。更に、頭上方向画像を用いた人物識別の手法において深度情報の有効性を確認した。人物識別の提案手法として、(1) 人物の身長、(2) 人物の面積、(3) 人物の体型、(4) 深度ヒストグラムの4つの特徴量を抽出した。学習アルゴリズムとしてはAdaBoostとC4.5を適用している。本手法において特に有効である特徴量は身長情報を含んだ(1) 人物の身長、(4) 深度ヒストグラムであった。この2つの特徴量を用いることで、94.4%の識別精度を得ることができ、提案手法の有効性を確認した。

今後は面積特徴と体型特徴の正規化による精度改善が人物識別全体の精度向上として期待できる。人物の面積と体型の特徴量は季節の変化などによる衣服の影響を受けてしまう。頭上方向の人物識別において、衣服の情報がコンテキスト情報として重要な役割を持つと考えられるため、このコンテキスト情報を追加することが今後の課題として挙げられる。本手法では人物がドアの前で立ち止まっている状況のみしか扱っていない。この実験環境は自然ではあるが^(注6)、対象者が歩いている場合や動いている場合など、より複雑な環境に対応した手法に拡張することも今後の課題として挙げられる。人物の監視において、人物追跡は最も重要な課題の一つである。人物同定においては別の場所を撮影しているカメラで個人を認識することが課題となる[1]。提案手法である頭上方向から撮影した画像を人物同定に適用することも今後の課題として興味深いものである。

文 献

- [1] M. Farenzena, L. Bazzani, A. Perina, V. Murino and M. Cristani. “Person re-identification by symmetry-driven accumulation of local features” Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2360-2367, 2010.
- [2] Y. Freund and R. E. Schapier. “Experiments with a new boosting algorithm,” Proc of ICML pp. 148-156, 1996.
- [3] 福添孝明, 伊藤雅人, 水戸大輔, 渡邊 睦, “体型特徴と習慣性特徴の確率的統合認識に基づく非拘束状態下での人物同定,” 電子情報通信学会 2008, Vol.J91-D, No. 5,

pp. 1369-1379, 2008.

- [4] A. C. Gallagher and T. Chen. “Using Context to Recognize People in Consumer Images,” IPSJ Transactions on Computer Vision and Applications, Vol. 1, pp. 115-126, 2009.
- [5] Y. Iwashita and A. Stoica. “Gait Recognition using Shadow Analysis,” Proc. of Symposium on Bio-inspired, Learning, and Intelligent Systems for Security 2009, pp. 26-31, 2009.
- [6] T. Kanade. “Picture processing by computer complex and recognition of human face,” Technical report, Kyoto University, Dept. of Information Science, 1973.
- [7] R. Nakatani, D. Kouno, K. Shimada and T. Endo. “A Person Identification Method Using a Top-view Head Image from an Overhead Camera,” Proc. of 2nd International Workshop on Advanced Computational Intelligence and Intelligent Informatics (IWACIII2011), SS4-1, 2011.
- [8] M. Onishi and I. Yoda. “Visualization of Customer Flow in an Office Complex over a Long Period,” Proceedings of International Conference on Pattern Recognition (ICPR), pp. 1747-1750, 2010.
- [9] J. R. Quinlan. “C4.5 Programs for Machine Learning,” Morgan Kaufmann Publishers, 1993.
- [10] M. Turk and A. P. Pentland. “Eigenfaces for recognition,” Journal of Cognitive Neuroscience, Vol. 3, No. 1, pp. 71-86, 1991.
- [11] 内田 光, 猪田良介, 辻 俊明, 阿部 茂, “実時間画像処理によるエレベータ乗場の人数計測と車椅子の識別,” 電気学会論文誌, 129, No. 6, pp.578-584, 2009.
- [12] S. Watanabe and N. Pakvasa. “Subspace method in pattern recognition,” Proc. of 1st Int. J. Conf on Pattern Recognition, pp. 2-32, 1973.
- [13] 山口 純平, 嶋田 和孝, 榎田 修一, 江島俊朗, 遠藤 勉, “顔特徴とコンテキスト情報に基づく人物識別,” 日本知能情報ファジィ学会誌, 知能と情報, Vol. 23, No. 2, pp. 13-21, 2011.

(注6) : 他の設置環境としてエスカレーター上やエレベーターの中などが人物の立ち止まる状況として考えられる。