

人物全身ボリウムデータを用いた Depth-based Gait Feature の認証性能評価

池田 太郎^{1,a)} 満上 育久^{2,b)} 八木 康史^{2,c)}

概要：従来よりカメラで撮影された人物のシルエット時系列から歩容特徴を抽出しそれによって個人認証を行う研究が行われているが、近年、Microsoft Kinect 等の安価なレンジセンサが登場したことにより、人物レンジデータを用いた歩容特徴も提案されてきている。我々の研究グループが提案した Depth-based Gait Feature (DGF) もそのような歩容特徴の 1 つだが、同一のデータを用いた他手法との性能評価比較は充分に行われていなかった。そこで本研究では、歩行者の全身ボリウムのデータセットを用い、そこから仮想的なレンジデータを生成することによって、同一の歩行データを用いた他手法との性能比較を行った。また、原理上歩行者の前方からの DGF と後方からの DGF の運動成分が類似することに着目し、前後間での DGF による歩容認証性能についても調査した。

キーワード：ボリウムデータ, Depth-based Gait Feature, 歩容認証, レンジデータ

1. はじめに

従来より、個人を識別・認証する目的で、様々なバイオメトリクス（生体情報）が用いられている。犯罪捜査や鑑定など、極めて高い識別精度が求められる場合には DNA、日常生活におけるセキュリティのためには指紋、静脈、虹彩などがすでによく利用されている。しかし、これらの認証は、細胞の採取やセンサへの接触・接近が要求されるため、意図的に回避することも可能であり、また遠方の人物の識別は原理上不可能である。一方で、歩容（歩き方）は遠方からでも観測可能なバイオメトリクスであり、これを用いた認証技術、歩容認証が近年注目を集めている。

通常、歩容はカメラによって観測され、従来の多くは撮影された人物のシルエットを抽出し、その時系列から特徴量を得る。最も多く利用される特徴量は Han らによって提案された Gait Energy Image (GEI) [1] であり、これは歩行 1 周期分のシルエット画像列の平均画像に相当する。楨原らは、同じく 1 周期分のシルエット時系列に対して、時間軸方向にフーリエ変換を適用し、平均画像だけでなく、その中に現れる非対称および対象な運動成分を抽出する Fourier Domain Feature (FDF) [2] を提案している。さら

に、これらの派生として、Gait Entropy Image (GENI) [5], Masked GEI based on GENI (MGEI) [6], また、Gait Flow Image (GFI) [7], Chrono-Gait Image (CGI) [8] など提案されているが、認証精度の面は GEI や FDF が高い精度を示すことが知られている。

一方で、近年 Microsoft Kinect に代表されるような安価なレンジセンサが普及するようになり、カメラ映像からよりもシルエットが安定・容易に抽出できるようになっただけでなく、シルエット内部の三次元形状までも容易に得られるようになったため、これを積極的に活用した特徴量も提案されるようになった。Gait Energy Volume (GEV) [3] はその 1 つであり、これはシルエットにおける GEI を単純に三次元に拡張したものである。すなわち、1 周期分の人物全身ボリウムの時系列に対する平均ボリウムを特徴量としたものである。一方、中島らは、FDF の考え方をレンジデータに拡張し、1 周期分のレンジデータ列に対してフーリエ変換を適用した Depth-based Gait Feature (DGF) [4] を提案している。この DGF は、既存の他の歩容特徴よりも高い認証精度を示すことが実験的に確認されているが、その被験者データ数は必ずしも十分なものではなく、同じく三次元的な歩容特徴である GEV との比較は行われていない。また、この DGF は単なる既存手法の拡張ではなく、人物の直立姿勢形状の成分と手足等の運動に相当する成分を分離した特徴量であるという点で既存の歩容特徴と異なっているが、文献 [4] ではこの点に関する客

¹ 大阪大学大学院 情報科学研究科

² 大阪大学 産業科学研究所

a) ikeda@am.sanken.osaka-u.ac.jp

b) mitsugami@am.sanken.osaka-u.ac.jp

c) yagi@am.sanken.osaka-u.ac.jp

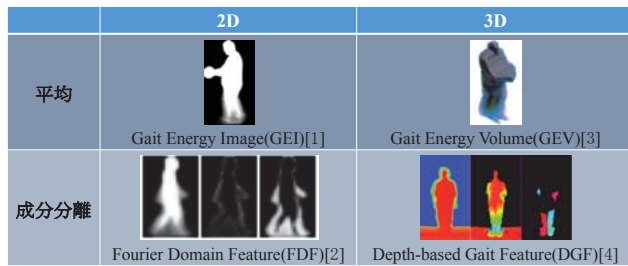


図 1 様々な歩容特徴

観的実験評価は行われていない。

そこで本研究ではまず、DGF と GEV を多数の被験者データを用いて公平に性能比較する。公平を期するためには同一の歩行者データを用いる必要があるが、DGF は歩行者正面からのレンジデータ、GEV は全身のボリウムデータと、両者は異なるデータ形式を前提としている点が問題となる。本研究では、トレッドミルの周囲に 25 台のカメラを設置した実験環境を用いて収集した約 100 名の歩行者の全身ボリウム系列のデータセットを用い、そこから仮想的にレンジデータを生成することで、同一の歩行者データを用いた両特徴量の性能比較を実現した。また、DGF により身体形状成分と運動成分が分離できることに着目し、歩行者の前方からのレンジデータ観測による DGF と後方からの DGF の間での認証について評価した。歩行時の手足の運動は主に前後方向であることを考慮すると、前方からの DGF と後方からの DGF では身体形状成分は異なるものの運動情報は高い相関性を示すことが期待されるため、この点について上述のデータを用いて実験を行った。

本稿では、まず第 2 章で DGF の概要について述べる。第 3 章で全身ボリウムデータからレンジデータを生成する手法について述べ、第 4 章で生成したレンジデータを用いた DGF の認証性能評価実験について述べる。

2. Depth-based Gait Feature

中島らが提案した DGF [4] について概説する。二次元シルエットベースの GEI と FDF、三次元形状ベースの GEV と DGF は、およそ図 1 のような対応関係になっている。GEI では、歩行 1 周期分のシルエット画像列の平均画像を特徴量としている。FDF は、同じく 1 周期分のシルエット時系列に対して、各ピクセルにおける時間軸方向のシルエットの有無の系列にフーリエ変換を適用し、その直流成分と 1 次、2 次周波数成分を取り出したものである。なお、この直流成分は GEI と等価である。GEV は、GEI の単純な三次元への拡張であり、1 周期分の人物全身ボリウムの時系列に対する平均ボリウムを特徴量としたものである。一方、DGF は FDF の三次元への拡張と捉えることもできるが、ただし、三次元ボリウムの各ボクセルにおけるボリウムの有無についてフーリエ変換するのではなく、レンジデータの各ピクセルの奥行き値の系列に対して

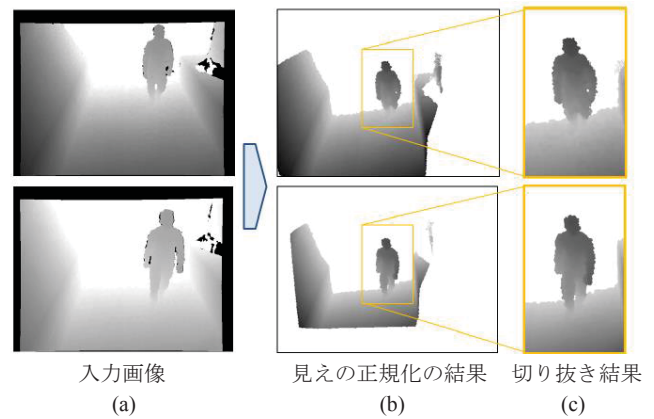


図 2 入力画像の見えの正規化と切り抜き結果 [4]

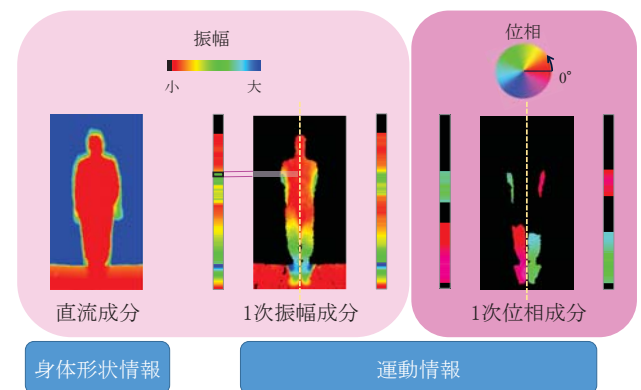


図 3 Depth-based Gait Feature[4]

フーリエ変換を適用する。

具体的な DGF の算出方法について述べる。まず、図 2(a) のような観測データに対し、見えの正規化、人物領域の切り抜き（同図 (b)）を行うことによって同図 (c) のような画像を得る。これを時間方向に歩行 1 周期分積み重ねた系列に対し離散フーリエ変換を適用し、直流成分と 1 次周波数成分を取り出す。この直流成分と 1 次周波数成分は、それぞれ身体形状情報と主要な運動情報に対応する。このうち 1 次周波数成分については、振幅と位相に分離するとともに、身体形状成分との冗長性を除くために、それぞれ左半身と右半身に分離し、計 4 つの 1 次元ベクトルとして表現する（図 3）。なお、このように DGF の運動成分には身体形状の情報が含まれていないが、この運動成分だけでも Equal Error Rate (EER) で約 90% の認証精度を示すことが文献 [4] で報告されている。

3. 全身ボリウムデータからのレンジデータ生成

本研究は、DGF による認証精度を GEV によるものと比較することを目的としているが、その際、十分な被験者数のデータを用いることと、公平な比較を行うために両手法で同一の歩行者データを用いることが求められる。しか

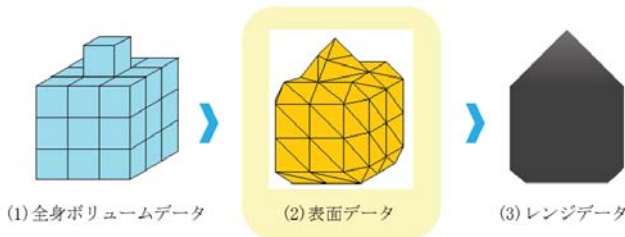


図 4 レンジデータ生成までのデータ形式

し、DGF は歩行者正面からのレンジデータ、GEV は全身のボリュームデータと、両者は異なる観測を前提としているため、これらの異なるデータを同時に計測可能な環境が必要となり、またこの環境下で多数の歩行者データを収集するのは、時間と労力のかかる作業となる。

そこで本研究では、我々の研究グループで保有している 25 眼歩行者同期撮影システムによって撮影された画像データセットを用いて全身ボリュームデータを作成し、さらにそこから仮想的にレンジデータを生成することで、同一の歩行者データを用いた両特徴量の性能比較を行う。

この処理の概要を図 4 に示す。まず、25 眼歩行者同期撮影システムにより得られた画像群からボクセル表現された全身ボリューム（同図 (1)）を作成する。この全身ボリュームデータに対して表面データ（同図 (2)）を作成した後、それをある視点に投影することで仮想レンジデータ（同図 (3)）を生成する。各ステップの処理の詳細を以降の節で述べる。

3.1 全身ボリュームデータ

我々が保有する 25 眼歩行者同期撮影システムは、図 5 に示すように、歩行者が歩行するトレッドミルの周囲に同期撮影が可能な 25 台のカメラを設置したシステムである。これらのカメラは事前にキャリブレーションされており、得られた画像群に対して視体積交差法を適用することによって各撮影時刻における歩行者の全身ボリュームを復元する。

図 6 に視体積交差法の処理の概念を示す。カメラで物体を撮影している状況を考えると、三次元空間中の物体は、画像中の物体領域とカメラ中心によって作られる錐体中に内包される。物体を複数のカメラで観測している場合には、この錐体の積領域 (Visual hull) に内在することとなる。カメラの台数を増やすことにより、この積領域は次第に物体そのものの形状に近づく。視体積交差法は、この考え方に基き、三次元形状を復元する手法である。

具体的には、まずカメラキャリブレーション情報を用いて、三次元ボクセルを各カメラの画像平面に投影し、全ての画像の人物領域の内部に投影されるボクセルには人物内部にあるとして 1 (形状内部)、それ以外のボクセルには 0 (形

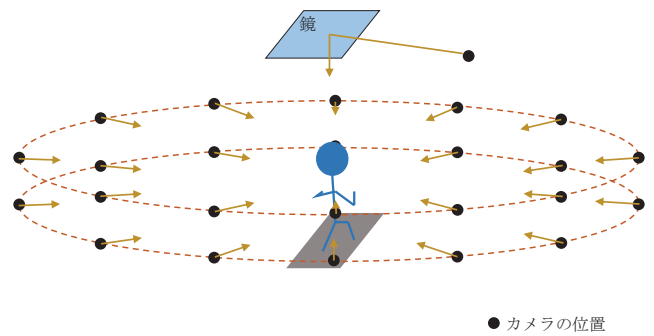


図 5 25 眼歩行者同期撮影システム

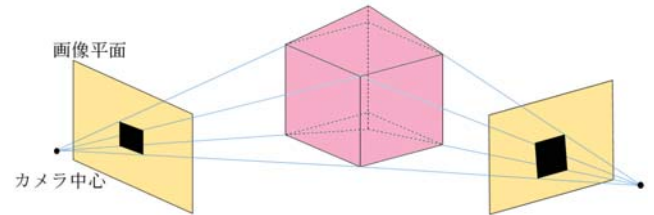


図 6 視体積交差法

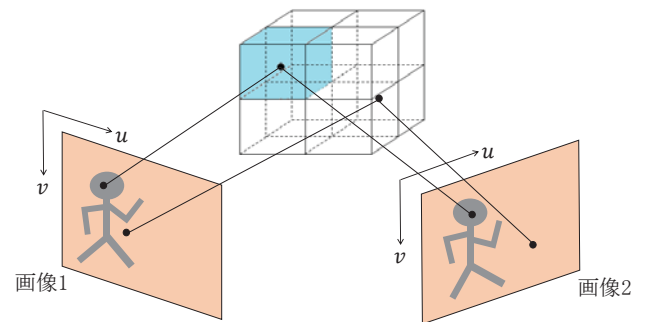


図 7 ボクセルの画像平面への投影

状外部) を付与する (図 7)。すなわち、位置 (x, y, z) にあるボクセルの値を $V(x, y, z)$ 、カメラ $i (i = 1, 2, \dots, 25)$ で観測された画像 i 上のボクセルが投影されたピクセルを (u_i, v_i) 、画像 i 中の人物領域のマスク画像を $INSIDE_i(u, v)$ とすると、 $V(x, y, z)$ を以下のように求める。

$$V(x, y, z) = \begin{cases} 1 & (INSIDE_1(u_1, v_1) = 1 \wedge \dots \\ & \wedge INSIDE_{25}(u_{25}, v_{25}) = 1), (1) \\ 0 & \text{otherwise,} \end{cases}$$

3.2 表面データへの変換

3.1 節で作成した人物ボリュームデータは、離散的な格子点の集合として表現されているため、これらを仮想レンジセンサの画像平面に投影するだけでは、穴のない連続的なレンジデータを復元することができない。そこで、この人物ボリュームデータに対してマーチングキューブ法を適用し、連続的な表面データを生成する。マーチングキューブ法とは、ボクセル表現されたボリュームの隣接 $8(2 \times 2 \times 2)$ ボクセルには、回転対称や 0/1 の反転を無視すると等値面

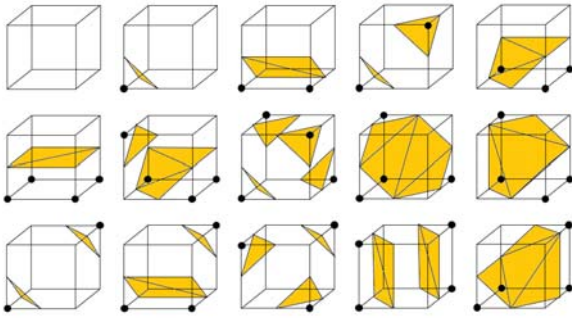


図 8 マーチングキューブ法による等値面での分割パターン

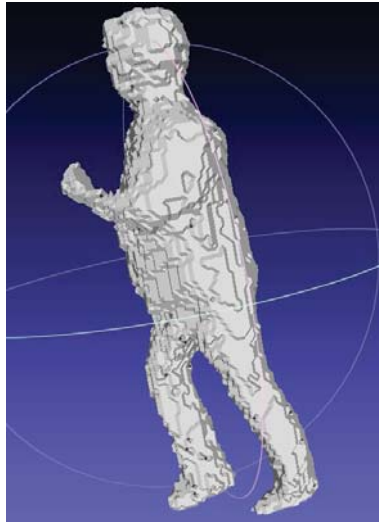


図 9 三角形メッシュデータのイメージ

での分割パターンは図 8 に示す 15 パターンしかないことに着目し、各パターンについて等地面のメッシュを定義しておく手法である。これをボリューム全体に対して適用することにより、図 9 のような全身表面データが得られる。

3.3 レンジデータの作成

前節で作成した表面データを用いてレンジデータを作成する。まず三角形メッシュデータに対し、仮想レンジセンサを配置する位置・姿勢を決定する。次に、図 10 のように表面データをすべて表面データの座標系からレンジセンサの座標系へ変換する。センサの向きを決めたので適切な回転行列 R 、並進ベクトル t を与えることによって、表面データの点の座標 M_w は式 (2) によってセンサ座標 M_c に変換される。

$$M_c = RM_w + t \quad (2)$$

そして、レンジセンサ座標系へ変換された各三角形メッシュの 3 頂点を画像平面に射影し、射影された 3 点から形成される三角形内部に画素が存在すれば、その奥行き値をその画素値に設定する。射影する点の座標を (X, Y, Z) 、射影された画像平面上の点の座標を (u, v) 、センサの内部パラメータ行列を A とすると、式 (3) によって画像平面に射

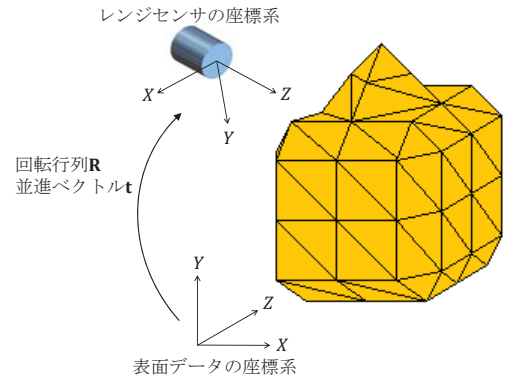


図 10 レンジセンサ座標系への変換

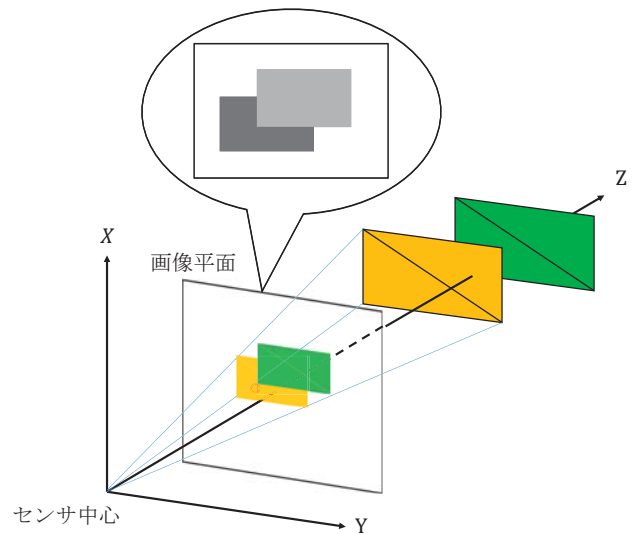


図 11 画像平面へのメッシュの射影 (黄面先の場合)

影される。

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = A \begin{pmatrix} X/Z \\ Y/Z \\ 1 \end{pmatrix} \quad (3)$$

ただし、一般に各方向には複数の三角形メッシュが存在するため、その投影の順序に注意を払わなければならない。例えば、図 11, 12 のように緑色の面が黄色の面より奥にあり、その一部が重なっている場合を考える。黄色の面を先に射影したとすると、緑色の面が射影されたときに重なっている部分の奥行きが上書きされてしまい、図 11 の吹き出しの中の画像のような誤ったレンジデータを生成してしまう。逆に、緑色の面を先に射影したとすると、黄色の面の奥行きが後で上書きされるため図 12 の吹き出しの中の画像のようなレンジデータが得られる。したがって、メッシュのオクルージョンが発生した場合に手前のメッシュの奥行きが最後に残るように、メッシュをレンジセンサの中心から遠い順に射影しなければならない。以上の処理により生成されたレンジデータの例を図 13 に示す。

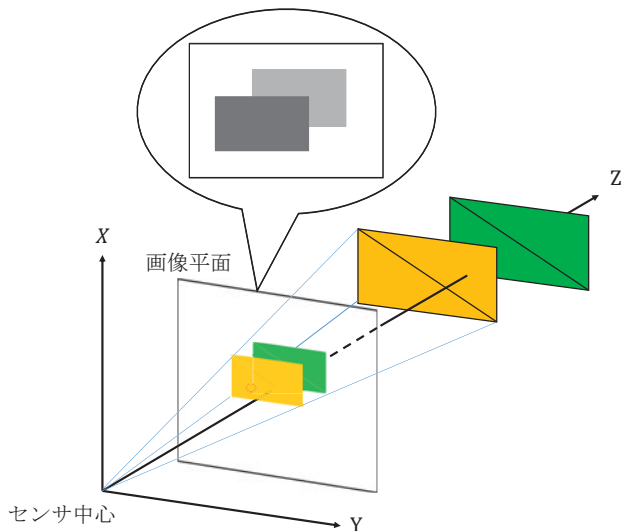


図 12 画像平面へのメッシュの射影 (緑面先の場合)

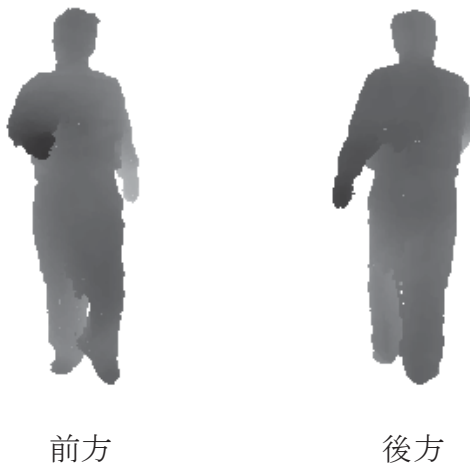


図 13 レンジデータ

4. 認証性能評価

4.1 実験環境

我々の研究グループでは、3章で述べた25眼歩行者同期撮影システムによって撮影された99名の歩行者の画像データを保有している。本研究では、各歩行者データから3周期分のデータを取り出し、そのそれぞれから全身ボリュームデータと仮想レンジデータを生成した。そして、全身ボリュームデータからはGEV、仮想レンジデータからはDGFをそれぞれ算出した。すなわち、各歩行者につきそれぞれ3つのGEVとDGFが得られており、そのうちの1つを辞書データ、残り2つを入力データとして認証に用いた。入力データと辞書データの距離尺度には、L2ノルムを用いた。

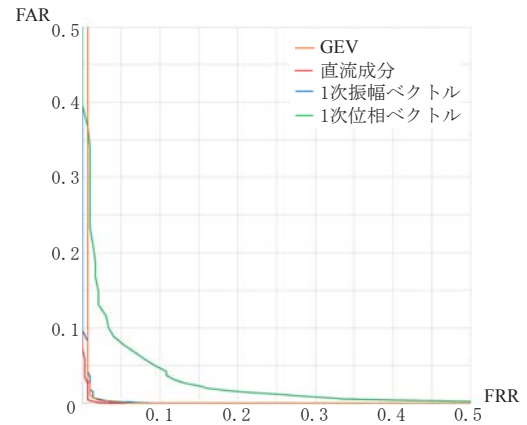


図 14 GEV と DGF の認証性能比較

4.2 実験結果

4.2.1 GEV と DGF の認証性能比較

GEV と DGF を用いた際の個人認証における ROC (Receiver Operating Characteristic) 曲線を図 14 に示す。この結果より、GEV、DGF の直流成分、DGF の 1 次振幅成分はほぼ同等の性能を示すことが確認できる。より詳細に見ると、DGF の直流成分は GEV よりもわずかに高い認証性能を示している。ここで考慮すべきなのは、GEV が全周の形状を用いておりしかも運動の情報も含んでいるのに対して、DGF が前方からの形状のみに基づきかつ運動の情報を含んでいないという点である。このことを踏まえると、GEV の方が高い認証性能を示すことが予想され、この実験結果はその予想とは一致しない。この理由としては、これは文献 [4] でも示唆されているが、歩容認証においてはやはり身体形状情報が支配的であるということが考えられる。すなわち、GEV は運動情報も含んではいるもののそれによって身体形状がぼやけてしまうのに対し、DGF の直流成分は身体形状を鮮明に表現しているということが、このような結果になった理由の 1 つではないかと考えられる。

4.2.2 前後の DGF 間の認証性能比較

DGF により身体形状成分と運動成分が分離できること、および、歩行時の手足の運動が主に前後方向であることを考慮すると、DGF の運動成分を用いることにより、歩行者の前方からのレンジデータ観測による DGF と後方からの DGF の間でも個人認証が可能であるのではないかと考えられる。そこで、3章で述べたのと同様の処理で歩行者後方の仮想レンジデータも作成し、前後 DGF 間での認証性能を評価した。なお比較のため、前方からの DGF 同士 (前節の結果と同一)、後方からの DGF 同士の認証性能も評価した。

前方 DGF 同士、後方 DGF 同士、前後 DGF 間の ROC 曲線をそれぞれ図 15, 16, 17 に示す。また、これらの Equal Error Rate (EER) を表 1 に示す。これらの結果より、前方同士、後方同士では直流成分と 1 次振幅成分が高い認証

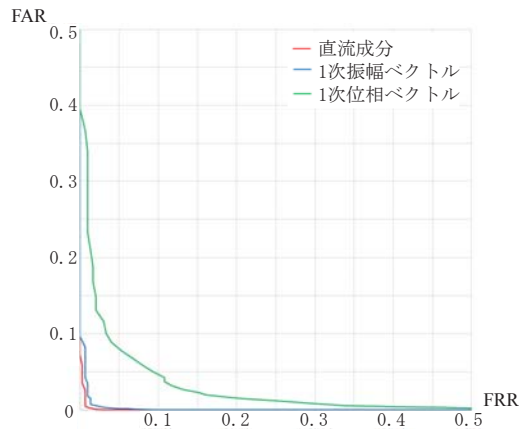


図 15 前方 DGF 同士

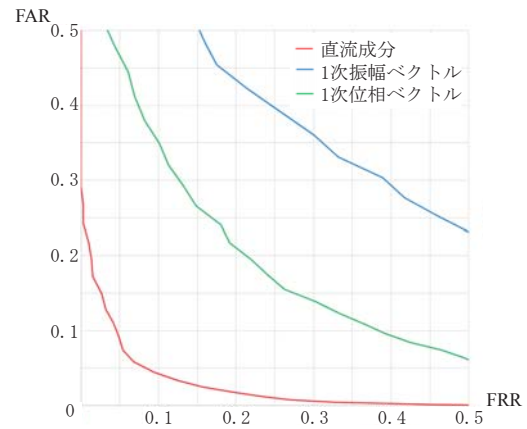


図 17 前後 DGF 間

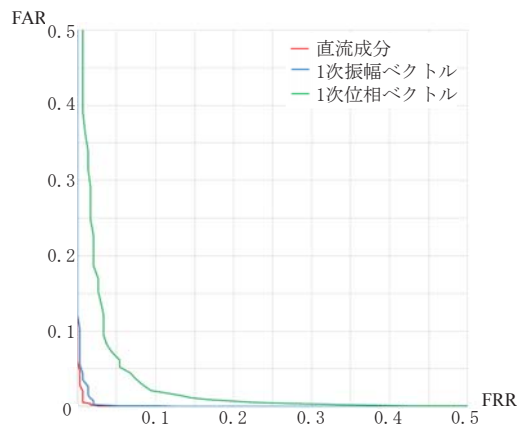


図 16 後方 DGF 同士



図 18 腕振り運動の遮蔽 (左：前方，右：後方)

表 1 各成分の EER

	直流成分	1 次振幅 ベクトル	1 次位相 ベクトル
前方データ vs 前方データ	0.6	1.3	6.7
後方データ vs 後方データ	0.7	1.4	5.4
前方データ vs 後方データ	6.2	33.0	20.3

性能を示しているのに対し，前後間ではこれらの認証精度が大きく低下しているのが確認できる．前後間で身体形状が異なることから，直流成分で認証精度が低下するのは妥当だが，運動成分である 1 次振幅が極端に悪化していることは意外な結果であると言える．この原因としては，腕振り運動における腕の遮蔽の有無が考えられる．すなわち，通常多くの歩行者は腕を前後に振る際，必ずしも前後方向のみに振るのではなく，図 18 の水色部で示すように先端を体の内側に入れるように振っており，そのため，図 19 に示すように，前方からの DGF では捉えられるが後方からの DGF では自己遮蔽により捉えられていないということが起こり，認証精度を低下させているのではないかと考えられる．

5. おわりに

本研究では，レンジセンサによる歩容認証のために提案

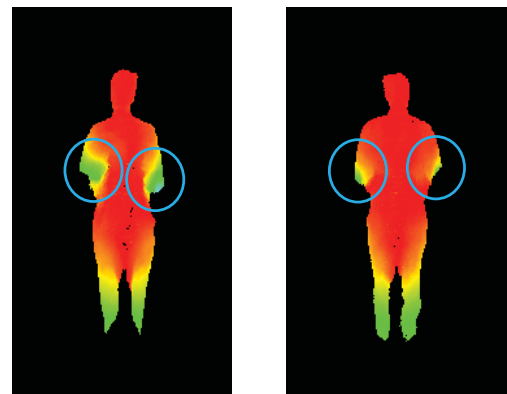


図 19 同一人物の 1 次振幅の比較 (左：前方，右：後方)

された Depth-based Gait Feature (DGF) に着目し，従来研究では十分に評価されていなかった Gait Everyy Volume (GEV) との認証性能比較を行った．これらの特徴量を，同一の歩行者データで公平に比較するために，25 眼歩行者同期撮影システムによって撮影された約 100 名の歩行者データから全身ボリュームとレンジデータを生成し，両手法を適用した．また，DGF により身体形状成分と運動成分が分離できることに着目し，歩行者の前方からのレンジデータ観測による DGF と後方からの DGF の間での認証につ

いて評価した。実験結果より、DGF は GEV とほぼ同等あるいはやや高い認証性能を示すことが確認された。前後 DGF 間の認証については、歩行者の腕振りが必ずしも前後方向だけではないために、認証精度が大きく低下することを確認した。

今後の課題として、まず腕振りにおける自己遮蔽を考慮した前後間の DGF 比較の改良が挙げられる。本研究では前後 DGF を単純に L2 ノルムによって比較したが、遮蔽の有無による大きな差分に影響を受けにくい L0, L1 ノルムを用いることにより、性能の向上が期待される。また、本研究で実現した仮想レンジデータの生成を、他の向きについても行うことで、前後間だけではなく任意の異なる 2 観測方向間での歩容認証性能評価も行うとともに、そのような状況に則した認証手法の検討することも重要な課題である。

参考文献

- [1] J. Han and B. Bhanu. Individual recognition using gait energy image, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.28, pp.316–322, 2006.
- [2] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi. Gait recognition using a view transformation model in the frequency domain, Proc. 9th European Conference on Computer Vision, pp.151–163, 2006.
- [3] S. Sivapalan, D. Chen, S. Denman, S. Sridharan, and C. Fookes. Gait Energy Volumes and Frontal Gait Recognition using Depth Images, Proc. of the International Joint Conference on Biometrics, pp.1–6, 2011.
- [4] H. Nakajima, I. Mitsugami, Y. Yagi. Depth-Based Gait Feature Representation, IPSJ Transactions on Computer Vision and Applications, Vol.5, pp.94–98, 2013.
- [5] K. Bashir, T. Xiang and S. Gong. Gait Recognition Using Gait Entropy Image, Proc. the 3rd International Conference on Crime Detection and Prevention, pp.1–6, 2009.
- [6] K. Bashir, T. Xiang and S. Gong. Gait recognition without subject cooperation, Pattern Recognition Letters, Vol.31, No.13, pp.2052–2060, 2010.
- [7] T. H. Lam, K. Cheung and J. N. Liu. Gait flow image: A silhouette-based gait representation for human identification, Pattern Recognition, Vol.44, No.4, pp.973–987, 2011.
- [8] C. Wang, J. Zhang, J. Pu, X. Yuan and L. Wang. Chrono-Gait Image: A Novel Temporal Template for Gait Recognition, Proc. the 11th European Conference on Computer Vision, pp.257–270, 2010.