

demonstrate the effectiveness of the proposed method in terms of not only the temporal super resolution quality but also gait-recognition performances.

低フレームレート映像からの 周期画像列復元による歩容認証

赤江 直樹^{†1} アル モンスール^{†1}
榎原 靖^{†1} 八木 康史^{†1}

近年、広域監視や犯罪捜査等を目的として人の歩き方の個性に基づく個人認証である歩容認証が注目を集めている。一般的に、防犯カメラによる映像は通信帯域幅や記録装置の容量の制限から、低フレームレートで撮影されていることが多く、映像中に含まれる歩容の位相（姿勢）が疎になるため、低フレームレート映像同士の照合は困難である。そこで本研究では、再構成型の時間超解像と事例に基づく時間超解像を組み合わせて、低フレームレートの歩容画像列から高フレームレートの歩容周期画像列を生成することで、照合を可能にする。具体的には、歩容画像列を固有空間における位相をパラメータとした固有空間中の歩容軌跡として捉え、複数の認識対象外被験者の歩容軌跡を学習データとして、標準モデルを作成する。次いで、入力の高フレームレート映像が与えられると、標準モデルを取り入れたエネルギー最小化の枠組みにおいて入力画像の位相推定と歩容軌跡の復元を繰り返し行うことで、高フレームレートの歩容周期画像列を生成する。100人の被験者の歩容データを用いた認証実験により、本手法の有効性を確認した。

Periodic Image Sequence Reconstruction from Low Frame-rate Videos for Gait Recognition

NAOKI AKAE,^{†1} AL MANSUR,^{†1} YASUSHI MAKIHARA^{†1}
and YASUSHI YAGI^{†1}

This paper describes a method of periodic temporal super resolution and its application to low frame-rate gait recognition. The existing reconstruction-based methods of periodic temporal super resolution often suffer from two problems: wagon-wheel effect and stroboscopic effect. To overcome these two problems, we introduce a standard gait manifold derived from a training set as an example-based cue and formulate the periodic temporal super resolution as an energy minimization framework considering both reconstruction-based and example-based cues. Experiments with side-view gait videos of 100 subjects

1. はじめに

近年、監視システムやアクセスコントロールのための個人認証の必要性が高まっており、人間の生体情報（バイオメトリクス）を利用した個人認証技術が注目されている。バイオメトリクスには、指紋、静脈、虹彩、顔などがあるが、最近、人の歩き方の個性を表す歩容を利用した個人認証方法として歩容認証が関心を集めている。歩容認証は、カメラから離れた人物を認証可能であるという点で、街中に設置された監視カメラを利用した広域監視への応用が期待されている。

歩容認証には、フレームレートの低さ¹⁾、歩行速度の違い^{2),3)}、観測方向の違い^{4),5)}、服装変化⁶⁾、経年変化^{7),8)}などが原因で認証精度が低下するといった様々な問題が存在する。一般的な監視カメラで撮影される映像は記憶容量等の制限によりフレームレートが低い（例えば1~5 fps）ことが多いため、中でも、低フレームレート映像への対応は必須と言える。低フレームレート映像では、映像中に含まれる歩容の位相（姿勢）が疎になり、プローブ（入力データ）とギャラリー（登録データ）に含まれる歩容の位相が一般には異なることから単純な照合が困難となるため、標準フレームレートを想定した従来手法では認証精度が大きく低下する。

そこで、低フレームレート映像を対象とした認証手法^{1),9)}が提案されているが、その数は少ない。Moriら¹⁾は、低フレームレートのプローブの位相を標準フレームレートのギャラリーの位相に同期させることで認証する方法を提案しているが、プローブとギャラリーの両方が低フレームレートである場合は照合することができない。また、低フレームレート映像に対する別の解決策として、モーフィングを用いた時間補間手法が挙げられる。Al-Huseinyら⁹⁾は、Level-set法によるモーフィングを歩容認証に適用し、中程度のフレームレート（5 fps程度）で有効であることを示しているが、極めて低いフレームレート（1~2 fps）では入力画像間の姿勢変化が非常に大きくなり、モーフィングによる補間はできない。

さらに、別の解決策として時間超解像¹⁰⁾⁻¹³⁾が提案されている。Shechtmanら¹⁰⁾は、複

^{†1} 大阪大学産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

数台のカメラから複数の低時空間解像度の映像を取得し、時空間の位置合わせに基づく再構成型の超解像手法を提案している。しかし、広域監視のようなカメラの台数が限られる環境では、複数台のカメラを利用することはできない。Shimano ら¹¹⁾ は、自己類似性を利用して、低フレームレートの単一映像から時間超解像を行っている。しかしながら、この方法では2倍程度の超解像が限界であり、歩容のように動きの速い運動を極めて低いフレームレートで撮影した映像には適用できない。

周期画像列の復元を目的とした時間超解像として、Makihara ら¹²⁾ は単一の準周期画像列からの時間超解像をエネルギー最小化の枠組みによって実現している。この方法では、推定された位相に基づいて、複数周期の画像を含む低フレームレートの映像から高フレームレートの1周期の映像を再構成している。しかし、フレームレートが極めて低いときに、ワゴンホイール現象が原因で初期位相の推定に失敗し、時間超解像の結果として逆転再生された映像が出力される可能性がある。さらに、カメラのサンプリング間隔と周期が一致するとき、観測される映像が止まって見えるストロボ効果が発生し、時間超解像に失敗する。

ワゴンホイール現象による逆転再生を解決するために、Akae ら¹³⁾ は、高フレームレートの学習データを用いて頑健に初期位相を推定する方法を提案している。しかし、この手法でもストロボ効果が発生している場合は時間超解像に失敗する。

そこで本論文では、事例に基づく手法と再構成に基づく手法を組み合わせることで、ストロボ効果が発生している歩容映像からも、個人性を保ちながら周期画像列を復元できる時間超解像手法を提案する。事例に基づく手法では、予め学習した認識対象外の複数被験者の歩容映像からなる標準モデルを利用する。一方で、再構成に基づく手法には、位相合わせされた複数周期分の歩容画像を利用する。これにより、ストロボ効果によって人物が静止しているような映像を入力として与えた場合でも、入力映像に存在する位相は入力映像に近く、欠けている位相は標準モデルによる推定結果に近くなるような高フレームレートの周期画像列を復元することが可能である。さらに、本手法によって時間超解像された高フレームレート映像を用いて歩容認証を行うことで、従来手法に比べて高い認証性能を目指す。

本論文の構成を以下に示す。2章では関連研究を、3章ではシルエットベースの歩容認証の前処理について説明する。続いて4章では、提案手法である時間超解像について述べる。5章で実験結果について説明し、最後に6章で結論と今後の課題を述べる。

2. 関連研究

時間超解像: 時間超解像は、事例に基づく時間超解像と再構成に基づく時間超解像の二つの

手法に分類される。

事例に基づく手法として、一般的な映像における微小な時空間パッチが、複数の時空間スケール映像中において繰り返し現れるという傾向を利用して、単一映像から時空間超解像を行う方法¹⁴⁾ が提案されている。また、時空間パッチの自己類似性を利用した方法¹¹⁾ も提案されている。しかしながら、これらの方法における時間超解像の倍率は高々2倍程度である。

再構成に基づく手法は、同一シーンを複数台のカメラで撮影し複数の映像を統合することで、サブピクセル・サブフレーム単位で時空間超解像を行う^{10),15)}。Agrawal ら¹⁵⁾ は、 N 台のカメラを利用して N 倍の時間超解像を行っているが、街頭カメラによる広域監視を考えると、多数のカメラを利用することは困難である。複数台のカメラを同時に使って撮影する手法の他に、繰り返される周期運動を一台のカメラで撮影し、複数周期の映像から1周期分の映像を再構成する手法^{12),13),16)} が提案されている。Xu ら¹⁷⁾ は、周期運動をする同種の動物が複数体含まれるシーンにおいて、位相の違う各個体の画像を統合し、その動物の周期運動を復元している。しかし、人間の歩容は被験者ごとに違っており、そもそも歩容の違いを利用して認証を行う歩容認証では、このような方法を利用することは不相当である。また、Ding ら¹⁸⁾ は、映像中から消失した一部のフレームを復元する方法を提案しているが、消失部分が観測された映像に比べて小さいことを仮定しているため、極めてフレームレートが低い場合に適用できない。

多様体学習: 低次元空間における多様体表現は、歩容の解析や姿勢追跡、運動合成において有効である。Cheng ら¹⁹⁾ は、Gaussian Process Latent Variable Model (GP-LVM) によって低次元空間で学習を行い、隠れマルコフモデル (HMMs) を利用して動きをモデリングすることで人物の照合システムを提案している。また、それを拡張した Scaled GP-LVM を用いた三次元人物追跡手法も提案されている²⁰⁾。Lee ら²¹⁾ は、関節角等の運動学的なパラメータに基づく多様体と見えに基づく多様体を組み合わせたモデルを用いて三次元人物追跡を行う手法を提案している。Elgammal ら²²⁾ は、人物の歩容を表す多様体を用いて動きと個人の見えを分離する方法を提案している。しかしながら、これらの方法は全て、多様体学習のために標準フレームレート映像を利用している。

歩容認証: 歩容認証に用いる特徴としては、周期単位のものと同フレーム単位のものも提案されており、周期単位の歩容特徴が比較的高い認証性能を示している。中でも、Okumura ら²³⁾ が大規模歩容データベースを用いた性能評価において報告しているように、平均シルエット²⁴⁾ (または GEI²⁵⁾) が最も高い認証性能を示している。その他にも、時間方向の振幅スペクトルを用いた周波数領域特徴⁴⁾ や、Fourier descriptors²⁶⁾、Gabor features²⁷⁾ が



図1 歩容シルエット画像列 (GSV)
Fig. 1 Normalized gait silhouette sequence

提案されている。しかしながら、低フレームレート映像では1歩行周期に含まれる画像枚数が極端に少なくなり、抽出される歩容特徴が高フレームレートの映像を用いたときに比べて大きく変化し、認証精度は低下してしまう。

3. 前処理

まず、背景差分に基づくグラフカットによる領域分割手法²⁸⁾により、歩容シルエットを抽出する。次に、歩容シルエットの高さが一定になるようにアスペクト比を保ちながらスケーリングし、歩容シルエットの重心に基づいて位置合わせを行うことで、画像の大きさを正規化した歩容シルエット画像列（以下 GSV）を生成する（図1参照）。

さらに、GSV に対して主成分分析 (PCA) を適用し、歩容シルエット画像の次元を圧縮する。本論文では、固有値の累積寄与率が99%以上になるように次元圧縮を行う。また図2に示すように、歩容シルエット画像の固有空間において、1枚の歩容シルエット画像は1点として表され、歩容シーケンスは各点を結んだ歩容軌跡として表現される²⁹⁾。

4. 時間超解像

4.1 準周期画像列の定義

本論文では、歩容シルエット画像列を準周期画像列として定義する。まず、歩容シルエット画像列が周期 P の周期画像列である場合は、時刻 t における画像を多次元ベクトル $\mathbf{x}(t)$ で表現すると、以下の性質を満たす。

$$\mathbf{x}(t + P) = \mathbf{x}(t) \quad \forall t \quad (1)$$

次に、時間の無次元量である位相 s と相対位相 \tilde{s} を以下のように導入する。

$$s = s_P(t) = \frac{t}{P} \quad (2)$$

$$\tilde{s} = s - \lfloor s \rfloor \quad (3)$$

ここで、 s_P は位相進行関数、 $\lfloor \cdot \rfloor$ は床関数である。周期画像列をこの位相によって表現すると以下ようになる。

$$\mathbf{x}_s(s) = \mathbf{x}(s_P^{-1}(s)) \quad (4)$$

これより、周期画像列は相対位相 $\tilde{s} \in [0, 1]$ をパラメータとするパラメトリック固有空間法³⁰⁾の多様体として表現され、またその多様体は $\mathbf{x}_s(1) = \mathbf{x}_s(0)$ を満たす。

一方、 N^{in} 枚の離散的に観測された歩容シルエット画像列を $\mathbf{X}^{in} = \{\mathbf{x}_i^{in}\} (i = 0, \dots, N^{in} - 1)$ として定義する。歩容シルエット画像列が完全に周期的で、かつフレームレート f も完全に一定である場合には、歩容シルエット画像列 \mathbf{X}^{in} 、各フレームの時間列 $t = \{t_i\}$ 、位相列 $s_P = \{s_{P,i}\}$ はそれぞれ以下で表現される。

$$\mathbf{x}_i^{in} = \mathbf{x}(t_i) = \mathbf{x}_s(s_{P,i}) \quad (5)$$

$$t_i = t_0 + \frac{i}{f} \quad (6)$$

$$s_{P,i} = s_P(t_i) = s_0 + \frac{i}{fP} \quad (7)$$

しかし、フレームレートの揺らぎや歩容の時間に関する揺らぎによって、周期画像列の仮定は満たされないことがある。その場合には、歩容シルエット画像列は以下で定義される準周期画像列 $\mathbf{X}_Q^{in} = \{\mathbf{x}_{Q,i}^{in}\}$ になる。

$$\mathbf{x}_{Q,i}^{in} = \mathbf{x}_s(s_{Q,i}) \quad (8)$$

$$s_{Q,i} = s_{P,i} + \Delta s_i \quad (9)$$

ここで、 $\{\Delta s_i\}$ は周期位相列からの揺らぎであり、 $s_Q = \{s_{Q,i}\}$ はその揺らぎが加わった準周期位相列である。

よって、本論文の問題設定は、与えられた入力準周期歩容シルエット画像列 \mathbf{X}_Q^{in} から、入力の位相列 s_Q の推定と多様体 \mathbf{x}_s の復元を行う問題として定義される。ここで、多様体 \mathbf{x}_s は先にも述べた通り周期画像列に対応し、適当な位相間隔で多様体上の点を再標本化することで高時間解像度の周期画像列を作成できる点に注意されたい。

4.2 多様体表現

前節で定義した周期画像列の多様体 \mathbf{x}_s を、パラメトリック固有空間法²⁹⁾により表現する。パラメトリック固有空間法を用いると、入力画像から多様体への投影により対応する位相が求まることから、位相合わせを効果的に行うことができる。具体的には、固有空間中における多様体を、位相をパラメータとする3次自然スプライン関数により表現する。まず、 M 次元の固有空間における N^{cp} 点の制御点 $\{\mathbf{y}_j^{cp}\}$ とそれに付随する位相を $\{s_j^{cp} (= j/N^{cp})\}$ ($j = 0, \dots, N^{cp} - 1$) とする。次に、 j 番目の区間 $[s_j^{cp}, s_{j+1}^{cp}]$ の k 次項 ($k = 0, 1, 2, 3$) に対するスプラインパラメータベクトルを $\mathbf{a}_{j,k}^{sp} \in \mathbb{R}^M$ として定義し、 j 番目の区間に対するスプラインパ

ラメータベクトルを並べた部分行列 A_j^{sp} とその部分行列を並べた全体行列 A^{sp} を、各々 $A_j^{sp} = [\mathbf{a}_{j,0}^{sp}, \dots, \mathbf{a}_{j,3}^{sp}]^T \in \mathbb{R}^{4 \times M}$, $A^{sp} = [A_0^{spT}, \dots, A_{N^{cp}-1}^{spT}]^T \in \mathbb{R}^{4N^{cp} \times M}$ として定義する。すると、 j 番目の区間に含まれる相対位相 \tilde{s} に対する固有空間中の補間点 $\hat{\mathbf{y}}(\tilde{s})$ は以下によって表現される。

$$\hat{\mathbf{y}}(\tilde{s}) = A^{spT} \mathbf{w}(\tilde{s}) \quad (10)$$

$$\mathbf{w}(\tilde{s}) = [0, \dots, 0, 1, w, w^2, w^3, 0, \dots, 0]^T \quad (11)$$

$$w = \frac{\tilde{s} - s_j^{cp}}{s_{j+1}^{cp} - s_j^{cp}} \quad (s_j^{cp} \leq \tilde{s} \leq s_{j+1}^{cp}) \quad (12)$$

ここで、 $\mathbf{w}(\tilde{s})$ は、 $4j$ 成分から $(4j+3)$ 成分までが $[1, w, w^2, w^3]$ で、その他が 0 となる補間係数ベクトルであり、 w は区間両端の制御点からの補間比率である。

一方、制御点ベクトルを並べた行列 $Y^{cp} = [\mathbf{y}_0^{cp}, \dots, \mathbf{y}_{N^{cp}-1}^{cp}]^T$ とスプラインパラメータ行列 A^{sp} の間には、制御点における 2 次微分までの連続性の制約¹²⁾ より以下の関係が成り立つ。

$$A^{sp} = DY^{cp} \quad (13)$$

ここで、 $D \in \mathbb{R}^{4N^{cp} \times N^{cp}}$ は重み行列である。したがって、制御点列 Y^{cp} が与えられると、式 (10)(13) を用いることで、任意の相対位相 \tilde{s} に対する固有空間中の補間点 $\hat{\mathbf{y}}(\tilde{s})$ が求まる。よって、次節からは、周期画像列の多様体 $\mathbf{x}_s(s; Y^{cp})$ の復元問題を、制御点列 Y^{cp} の復元問題として見なす。

4.3 標準モデルの構築

本節では、標準フレームレートで撮影された N^{tr} 人の歩容シーケンスから標準モデルの構築を行う方法について述べる。まず、認識対象外の N^{tr} 人の歩容シルエット画像列から、従来の時間超解像手法¹³⁾ によって N^{cp} [frame / period] の高フレームレートの周期画像列を被験者ごとに作成する。次に、 N^{tr} 人の周期画像列全てに対して PCA を適用して次元圧縮を行い、その次元数を M とする。これにより、被験者 N^{tr} 人の N^{cp} 個の制御点から成る多様体が同一の固有空間で表現される。その際、各被験者の多様体の位相が揃っていることに注意されたい。また、 i 番目の被験者の多様体を Y_i^{tr} と表す。

次に、各被験者の多様体の制御点行列 $Y_i^{tr} \in \mathbb{R}^{N^{cp} \times M}$ を制御点ベクトル $\mathbf{z}_i^{tr} \in \mathbb{R}^{N^{cp}M}$ として展開し、PCA による制御点ベクトルの次元圧縮を行う。PCA によって得られた制御点ベクトルの平均ベクトルと M_m 個の固有ベクトルを、それぞれ制御点行列の形に戻し、平均行列 \bar{Y}^{tr} と固有行列 $\{E_j^{tr}\} (j=0, \dots, M_m-1)$ と表現する。これより、標準モデルによって表現される多様体の制御点行列は、以下の式で表される。

$$\hat{Y}^{tr} = \bar{Y}^{tr} + \sum_{j=0}^{M_m-1} \alpha_j E_j^{tr} \quad (14)$$

なお、 $\boldsymbol{\alpha} = [\alpha_0, \dots, \alpha_{M_m-1}]^T$ は固有行列 E_j^{tr} の重みとなる主成分であり、以降では、これを標準モデルの係数と呼ぶ。

4.4 エネルギー最小化の枠組み

従来の周期画像列の時間超解像の手法^{10),11)} と同様に、本論文ではエネルギー最小化の枠組みによる定式化を行う。まず、入力として与えられる歩容シルエット画像列の固有空間における既知のベクトル列を $\mathbf{Y}_Q^{in} = \{\mathbf{y}_{Q,i}^{in}\}$ とする。それに対する未知の位相列が $s_Q = \{s_{Q,i}\}$ であり、その i 番目の位相 $s_{Q,i}$ に対する補間係数ベクトルを式 (11) と同様にして $\mathbf{w}(s_{Q,i})$ として定義する。この時、 i 番目の位相 $s_{Q,i}$ に対応する多様体上の点、即ち補間点は以下となる。

$$\hat{Y}(Y^{cp}, s_{Q,i}) = A^{spT} \mathbf{w}(s_{Q,i}) = Y^{cpT} D^T \mathbf{w}(s_{Q,i}) \quad (15)$$

エネルギー関数は以下の 4 項を考慮して、式 (16) のように定義する。

- (1) データ項 (第 1 項) : 補間点 $\hat{\mathbf{y}}(Y^{cp}, s_{Q,i})$ と入力点 $\mathbf{y}_{Q,i}^{in}$ の距離
- (2) 標準モデル項 (第 2 項) : 多様体 Y^{cp} と推定された標準多様体 \hat{Y}^{tr} の距離
- (3) 多様体の平滑化項 (第 3 項) : 固有空間における多様体 $\mathbf{y}_s(s; Y^{cp})$ の曲率
- (4) 位相進行の平滑化項 (第 4 項) : 位相列 s_Q の進行速度と大域的な位相進行速度の差

$$E(Y^{cp}, \boldsymbol{\alpha}, s_Q) = \frac{1}{N^{in}} \sum_{i=0}^{N^{in}-1} \|Y^{cpT} D^T \mathbf{w}(s_{Q,i}) - \mathbf{y}_{Q,i}^{in}\|^2 + \lambda_t \frac{1}{N^{cp}} \|Y^{cp} - (\bar{Y}^{tr} + \sum_{j=0}^{M_m-1} \alpha_j E_j^{tr})\|^2 + \lambda_m \frac{1}{N^{cp}} \|BY^{cp}\|^2 + \lambda_s \frac{1}{N^{in}} \sum_{i=1}^{N^{in}-1} \left(s_{Q,i+1} - s_{Q,i} - \frac{1}{P'} \right)^2 \quad (16)$$

ここで、 B は多様体の曲率を算出するための係数行列であり、 P' はフレームを単位とする大域的な周期であり、フレームレート f と時間単位の周期 P の積として $P' = fP$ で表さ

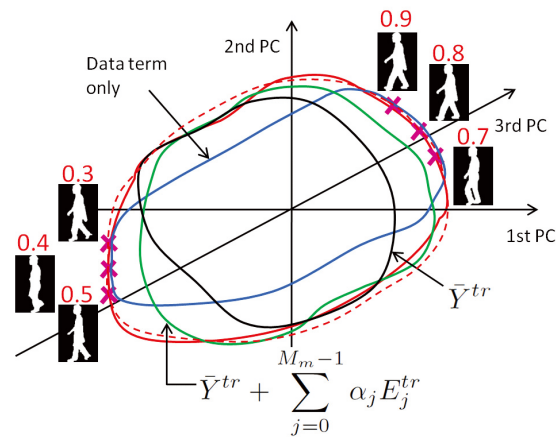


図2 データ項と標準モデル項の有無による復元される多様体の違い (桃色バツ印と赤色数値は入力画像列とその相対位相を固有空間 (第1主成分から第3主成分まで) で表している。赤色点線は本来の多様体、青色実線は標準モデル項が無い場合、緑色実線はデータ項が無い場合、赤色実線は全ての項を含めた場合の復元結果である。) Fig.2 Effect of the data term and manifold fitness

れる。

エネルギー関数におけるデータ項と標準モデル項の働きを示すために、各項の有無による復元される多様体の違いを図2に示す。エネルギー関数に標準モデル項が含まれない場合、多様体復元は入力画像と平滑化制約にのみ依存するため、多様体は図中の青色実線のようになり、特に観測から欠けている位相において、本来の多様体 (赤色点線) とは異なるものが復元結果として得られる。一方、入力画像が標準モデルで表現可能な範囲外に存在するにも拘わらず、標準モデルの範囲内で多様体を表現する場合、復元される多様体 (緑色実線) は入力画像から離れてしまう可能性があり、本来の多様体とは異なるものになる。

エネルギー関数の四つの項全てを考慮した場合は、入力画像が存在する位相付近では入力画像の近くを通り、観測が欠落した位相付近では標準モデルに近づくように多様体が復元される。このように、データ項と標準モデル項の二つを組み合わせることで、図中の赤色実線のように、本来の多様体にできるだけ近い多様体を復元することが可能となる。

4.5 反復解法によるエネルギー最小化

エネルギー関数 $E(Y^{cp}, \alpha, s_Q)$ は、多様体の制御点列 Y^{cp} と標準モデルの係数 α に関する2次形式になっており、ある固定の位相列 s_Q の下では、多様体の制御点列 Y^{cp} と標準

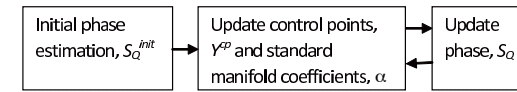


図3 反復解法
Fig.3 Iterative solution

モデルの係数 α が線形解法によって求まる。一方、位相列 s_Q については、スプライン曲線が位相により区分的に切り替わり、かつ、ある区間における補間比率 w がデータ項において6次式として現れることから、エネルギー関数中では非常に複雑な形で現れることになる。この強い非線形最適化問題を解くために、図3に示す反復解法によってエネルギー最小化を行う。

まず、Akaike¹³⁾ の手法と同様に、動的計画法³¹⁾ によって初期位相 $s_Q^{init} = \{s_{Q,i}^{init}\}$ の推定を行う。さらに、大域的な周期 P' を平均位相進行として算出する。そして、初期位相 s_Q^{init} を s_Q^0 として再定義し、エネルギー最小化の反復を行う。 r 回目の繰り返しにおける多様体の制御点 $Y^{cp,r}$ と標準モデルの係数 α^r は、固定された位相進行 s_Q^r の下、線形解法によって算出される。 $r+1$ 回目の繰り返しにおける位相進行 s_Q^{r+1} は、固定された制御点行列 $Y^{cp,r}$ と標準モデルの係数 α^r の下で、エネルギー関数のデータ項の位相項を Taylor 展開によって2次近似することで、線形解法によって計算される。このように多様体の制御点 $Y^{cp,r}$ 、標準モデルの係数 α^r 、及び位相進行 s_Q^r を、反復解法によって交互に更新し、それを解が収束するまで繰り返す。

5. 実験

5.1 データセット

OU-ISIR Gait Database³²⁾ に含まれる185人の被験者の歩容データを用いて評価を行った。185人のうち、85人を標準モデル構築のための学習セットに、100人をテストセットに用いた。被験者は、時速4kmで動く歩行機上を歩行し、1人につき2回撮影を行い、それぞれの映像をギャラリー及びプローブとした。なお、フレームレートは60fps、撮影時間は6秒、歩容シルエット画像のサイズは 22×32 pixel とした。低フレームレート映像は、60fpsの歩容シルエット画像列から一定間隔で画像を間引くことで作成した。

実験における各パラメータは、歩容シルエット画像の固有空間の次元数を $M = 112$ 、エネルギー関数における各項の係数を $\lambda_t = 1$ 、 $\lambda_m = 5$ 、 $\lambda_s = 1$ 、多様体の制御点数を $N^{cp} = 100$ とした。

5.2 時間超解像の精度評価

本節では、時間超解像によって得られた周期画像列の精度評価を行った。レベルセット法によるモーフィング⁹⁾ (Morph), 及び多様体復元に標準モデルを用いない時間超解像¹³⁾ (TSR) との比較を行った。図4は、1 fpsの歩容シルエット画像列を入力として与えたときの入力画像列と、各手法による時間超解像結果を示している。この被験者の場合、歩行周期がサンプリング間隔とほぼ一致してストロボ効果が起き、複数周期間で観測される歩容の位相がほぼ等しくなっている。したがって、Morph (図中第2行) によって入力画像間を補間しても、結果として同じ姿勢の歩容画像列が出力されてしまう。また、TSR (図中第3行) においても、ストロボ効果が原因で時間超解像に失敗していることが分かる。一方、提案手法 (図中第4行) では、自然な周期運動をしている歩容画像列が復元されていることが分かる。さらに、復元された周期画像列は60 fpsで撮影された真値 (図中第5行) に比較的近いことが分かる。

定量的な評価を行うために、各手法によって復元された周期画像列と真値間の平均二乗誤差を求めた。平均二乗誤差 E は、以下の式によって定義される。

$$E = \frac{1}{NWH} \sum_{k=0}^{N-1} \sum_{i=0}^{W-1} \sum_{j=0}^{H-1} [I(i, j, k) - \hat{I}(i, j, k)]^2 \quad (17)$$

なお、 N は画像枚数、 W 及び H は画像の幅と高さ、 I は真値、 \hat{I} は復元された周期画像列である。図5は、各手法における平均二乗誤差を示しており、提案手法の誤差が比較手法に比べて小さいことが分かる。

5.3 歩容認証の認証精度評価

本節では、レベルセット法によるモーフィング⁹⁾ (Morph), 時間超解像¹³⁾ (TSR) に加えて、低フレームレートの映像から直接特徴を抽出する手法 (NoTSR) に対する認証性能を比較する。歩容特徴として平均シルエット^{24),25)} と周波数領域特徴⁴⁾ を用い、評価指標として Equal Error Rate (EER)³³⁾ と Rank-1 認証率を用いて、認証精度の評価を行った。

図6、図7は、各歩容特徴を用いたときの EER と Rank-1 認証率をフレームレートごとに示している。同図より、提案手法が比較手法に比べて高い認証精度を示していることが分かる。特に、フレームレートが極めて低い時に、認証精度が大きく改善している。なお周波数領域特徴は4 fps 以下の場合には抽出できないため、NoTSR は5 fps 以上でしか認証できないことに注意されたい。



図4 時間超解像の結果 (第1行: 入力画像列 (1 fps), 第2行: Morph⁹⁾, 第3行: TSR¹³⁾, 第4行: 提案手法, 第5行: 真値 (60 fps))

Fig.4 Results of temporal super resolution

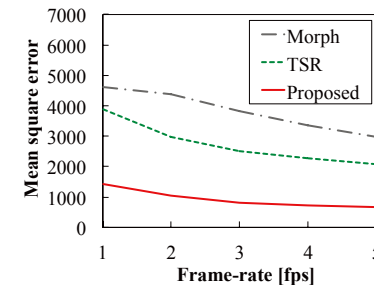


図5 平均二乗誤差による時間超解像の精度評価
Fig.5 Mean squared error of reconstructed images

5.4 ハイパーパラメータに対する感度評価

本節では、エネルギー関数におけるパラメータ変化に対する認証精度の評価を行う。図8は、エネルギー関数 (式 (16)) の標準モデル項の係数 λ_t の値を変化させた時の EER と Rank-1 認証率を示している。他項の係数に比べて λ_t を小さくしたとき (例えば 0.001 から 0.01), 標準モデル項の影響は無視出来るほど小さくなる。その結果、入力画像に含まれない位相付近では多様体を正しく復元することができず、認証精度は低下する。一方、

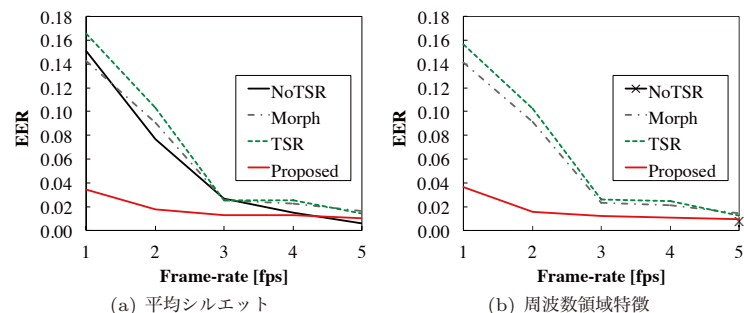


図 6 EER による認証精度の比較
Fig.6 EER with different features

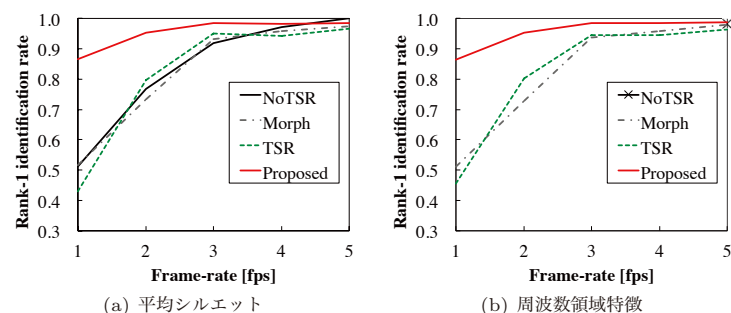


図 7 Rank-1 認証率による認証精度の比較
Fig.7 Rank-1 identification rate with different features

$0.1 \leq \lambda_t \leq 100$ の比較的広い範囲で高い認証精度を保っており、認証精度は λ_t に対して安定であることが分かる。

6. おわりに

本論文では、事例に基づく手法と再構成に基づく手法を組み合わせた時間超解像を提案し、歩容認証に適用した。学習セットから構築した標準モデルをエネルギー最小化の枠組みに導入することで、ストロボ効果が起きている歩容画像列からも高フレームレートの周期画像列を生成することを可能にした。実験では、時間超解像の精度と認証精度の点で、提案手法の有効性を示した。

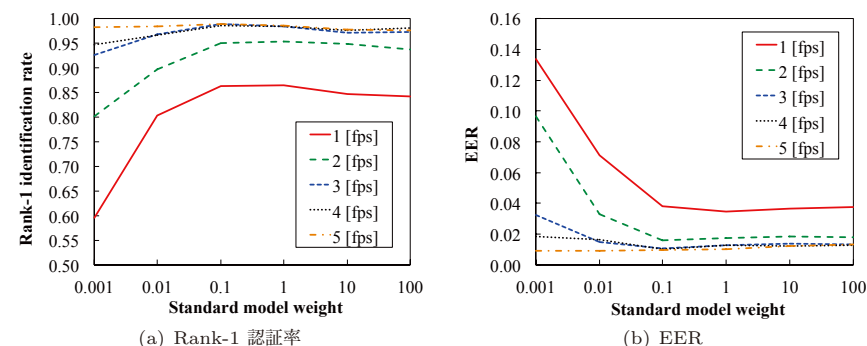


図 8 標準モデル項の係数による認証精度への影響
Fig.8 influence of standard manifold weight on EER and Rank-1 identification rate

また、本手法では複数周期に渡って同じ運動をしていることを前提としているため、周期間で見えが異なる場合（歩容の空間的な揺らぎ、歩行速度の変化、観測方向の変化）は、本手法を適用できない。そこで、今後の課題として、周期間で見えの変化がある場合への対応が考えられる。

参考文献

- 1) Mori, A., Makihara, Y. and Yagi, Y.: Gait Recognition using Period-based Phase Synchronization for Low Frame-rate Videos, *Proc. of the 20th Int. Conf. on Pattern Recognition*, Istanbul, Turkey, pp.2194–2197 (2010).
- 2) Makihara, Y., Tsuji, A. and Yagi, Y.: Silhouette Transformation based on Walking Speed for Gait Identification, *Proc. of the 23rd IEEE Conf. on Computer Vision and Pattern Recognition*, San Francisco, CA, USA (2010).
- 3) Tanawongsuwan, R. and Bobick, A.: A Study of Human Gaits across Different Speeds, Technical report, Georgia Tech (2003).
- 4) Makihara, Y., Sagawa, R., Mukaigawa, Y., Echigo, T. and Yagi, Y.: Gait Recognition Using a View Transformation Model in the Frequency Domain, *Proc. of the 9th European Conf. on Computer Vision*, Graz, Austria, pp.151–163 (2006).
- 5) Wang, Y., Yu, S., Wang, Y. and Tan, T.: Gait Recognition Based on Fusion of Multi-view Gait Sequences, *Proc. of the IAPR Int. Conf. on Biometrics 2006*, pp. 605–611 (2006).
- 6) Hossain, M.A., Makihara, Y., Wang, J. and Yagi, Y.: Clothing-Invariant Gait Identification using Part-based Clothing Categorization and Adaptive Weight Control,

- Pattern Recognition*, Vol.43, No.6, pp.2281–2291 (2010).
- 7) Sarkar, S., Phillips, J., Liu, Z., Vega, I., Grother, P. and Bowyer, K.: The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis, *Trans. of Pattern Analysis and Machine Intelligence*, Vol.27, No.2, pp.162–177 (2005).
 - 8) Matovski, D., Nixon, M., Mahmoodi, S. and Carter, J.: The effect of time on the performance of gait biometrics, *Proc. of the 4th IEEE Int. Conf. on Biometrics: Theory Applications and Systems*, Washington D.C., USA, pp.1–6 (2010).
 - 9) Al-Huseiny, M.S., Mahmoodi, S. and Nixon, M.S.: Gait Learning-Based Regenerative Model: A Level Set Approach, *The 20th Int. Conf. on Pattern Recognition*, Istanbul, Turkey, pp.2644–2647 (2010).
 - 10) Shechtman, E., Caspi, Y. and Irani, M.: Space-Time Super-Resolution, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.27, No.4, pp.531–545 (2005).
 - 11) Shimano, M., Okabe, T., Sato, I. and Sato, Y.: Video temporal super-resolution based on self-similarity, *Proc. of the 10th Asian Conf. on Computer Vision*, Vol.1, Queenstown, New Zealand, pp.93–106 (2010).
 - 12) Makihara, Y., Mori, A. and Yagi, Y.: Temporal Super Resolution from a Single Quasi-Periodic Image Sequence Based on Phase Registration, *Proc. of the 10th Asian Conf. on Computer Vision*, Queenstown, New Zealand, pp.107–120 (2010).
 - 13) Akae, N., Makihara, Y. and Yagi, Y.: Gait Recognition using Periodic Temporal Super Resolution for Low Frame-rate Videos, *Proc. of the International Joint Conference on Biometrics*, Washington, D.C., USA, pp.107–120 (2011).
 - 14) Shahar, O., Faktor, A. and Irani, M.: Space-Time Super-Resolution from a Single Video, *Proc. of the IEEE Int. Conf. on Computer Vision and Pattern Recognition*, Colorado Springs, USA (2011).
 - 15) Agrawal, A., Gupta, M., Veeraraghavan, A. and Narasimhan, S.: Optimal coded sampling for temporal super-resolution, *Proc. of of IEEE computer society conference on Computer Vision and Pattern Recognition 2010*, San Francisco, CA, USA, pp.599–606 (2010).
 - 16) Singh, M., Basu, A. and Mandal, M.K.: Event dynamics based temporal registration, *IEEE Trans. on Multimedia*, Vol.9, No.5, pp.1004–1015 (2007).
 - 17) Xu, X., Wan, L., Liu, X., Wong, T.-T., Wang, L. and Leung, C.-S.: Animating animal motion from still, *ACM Trans. Graphics*, Vol.27, No.5 (2008).
 - 18) Ding, T., Sznajder, M. and Camps, O.I.: A rank minimization approach to video inpainting, *Proc. of the IEEE Int. Conf. on Computer Vision*, Rio de Janeiro, Brazil (2007).
 - 19) Cheng, M.-H., Ho, M.-F. and Huang, C.-L.: Gait analysis for human identification through manifold learning and HMM, *Pattern Recognition*, Vol.41, No.8, pp.2541–2553 (2008).
 - 20) Urtasun, R., Fleet, D.J., Hertzmann, A. and Fua, P.: Priors for People Tracking from Small Training Sets, *Proc. of of IEEE computer society International conference on Computer Vision 2005*, Beijing, China, pp.403–410 (2005).
 - 21) Lee, C.S. and Elgammal, A.: Coupled Visual and Kinematic Manifold Models for Tracking, *International Journal of Computer Vision*, Vol.87, No.1-2 (2010).
 - 22) Elgammal, A. and Lee, C.S.: Separating Style and Content on a Nonlinear Manifold, *Proc. of the 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, Vol.1, Washington, DC, USA, pp.478–485 (2004).
 - 23) Okumura, M., Iwama, H., Makihara, Y. and Yagi, Y.: Performance Evaluation of Vision-based Gait Recognition using a Very Large-scale Gait Database, *Proc. IEEE 4th Int. Conf. on Biometrics: Theory, Applications and Systems*, Washington, D.C., USA, pp.1–6 (2010).
 - 24) Liu, Z. and Sarkar, S.: Simplest Representation Yet for Gait Recognition: Averaged Silhouette, *Proc. of the 17th Int. Conf. on Pattern Recognition*, Vol.1, pp.211–214 (2004).
 - 25) Han, J. and Bhanu, B.: Individual Recognition Using Gait Energy Image, *Trans. on Pattern Analysis and Machine Intelligence*, Vol.28, No.2, pp.316–322 (2006).
 - 26) Zhao, G., Chen, R., Chen, G. and Li, H.: Recognition of human periodic movements from unstructured information using a motion-based frequency domain approach, *Image and Vision Computing*, Vol.24, pp.795–809 (2006).
 - 27) Tao, D., Li, X., Wu, X. and Maybank, S.: Human Carrying Status in Visual Surveillance, *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, Vol.2, New York, USA, pp.1670–1677 (2006).
 - 28) Makihara, Y. and Yagi, Y.: Silhouette Extraction Based on Iterative Spatio-Temporal Local Color Transformation and Graph-Cut Segmentation, *Proc. of the 19th Int. Conf. on Pattern Recognition*, Tampa, Florida USA (2008).
 - 29) Murase, H. and Nayar, S.K.: Parametric eigenspace representation for visual learning and recognition, *Proc. of SPIE*, 2031 (1993).
 - 30) Murase, H. and Sakai, R.: Moving Object Recognition in Eigenspace Representation: Gait Analysis and Lip Reading, *Pattern Recognition Letters*, Vol.17, pp.155–162 (1996).
 - 31) Oka, R.: Spotting method for classification of real world data, *Computer Journal*, Vol.41, No.8, pp.559–565 (1998).
 - 32) : OU-ISIR Gait Database, <http://www.am.sanken.osaka-u.ac.jp/GaitDB/index.html>.
 - 33) Phillips, P., Moon, H., Rizvi, S. and Rauss, P.: The FERET Evaluation Methodology for Face-Recognition Algorithms, *Trans. of Pattern Analysis and Machine Intelligence*, Vol.22, No.10, pp.1090–1104 (2000).