

## [特別講演] 事前知識を利用した人体の形状・姿勢・運動推定

浮田 宗伯†

† 奈良先端科学技術大学院大学 〒 630-0192 奈良県生駒市高山町 8916-5

E-mail: [†ukita@is.naist.jp](mailto:†ukita@is.naist.jp)

あらまし 時系列画像から人体の3次元形状・姿勢・運動を推定する手法を紹介する。これらの手法は、予め学習しておく「人体に関する事前知識」を参照することにより以下のような特徴を備えることができる。1) 動きの事前知識を利用した追跡：複雑かつ多様な形状・姿勢変化が伴う運動においても高速かつ安定に人体を追跡可能、2) 画像特徴からの写像による人体推定：撮影画像や復元形状と人体姿勢の同期データを学習しておくことによって、着衣の大きな形状変化や遮蔽が生じている対象（スカートや着物を着た人間）でも安定な推定が可能。本稿ではこうした事前知識を参照した手法を紹介しながら、特に「画像からの復元形状誤差を考慮した高精度推定」や「複数種類の動き（歩行、ジョギングなど）の統一的なモデル化による汎用性の向上」という特徴を実現した著者らの手法について述べる。キーワード 人体計測，形状復元，姿勢推定，運動推定，事前知識

## Human Shape, Pose, and Motion Estimation using Prior Models

Norimichi UKITA†

† Nara Institute of Science and Technology Takayama 8916-5, Ikoma, Nara, 630-0192 Japan

E-mail: [†ukita@is.naist.jp](mailto:†ukita@is.naist.jp)

**Abstract** This paper introduces methods for estimating 3D human shape, pose, and motion from temporal image sequences. These methods have the following properties with prior models, which are learned in advance; they can handle with even 1) complexly moving shapes and poses and 2) loose-fitting clothing that covers a human body robustly to its significant shape deformation. While a variety of these methods are introduced in the paper, the authors' works that enable "accurate estimation with explicit handling of 3D reconstruction error" and "improving the versatility using a unified model of multiple actions" are described in detail.

**Key words** Human body measurement, Shape reconstruction, Pose estimation, Motion estimation, Prior models

### 1. ま え が き

人体の姿勢・運動解析に関する数多くの研究が行われている。これらの研究により、ヒューマンモデリング [1]、ロボットの運動制御 [2], [3]、CG アニメーション [4] など、さまざまな応用が可能となる。こうした解析に必要な人体の情報は、一般的にモーションキャプチャシステムによって取得されることが多い。モーションキャプチャは高精度であるが、被観測者はマーカや機器を身につけていなければならない、利用できる場面は限られたリアルで自然な動きを計測できないといった問題がある。

これに対して、カメラ映像からのモーションキャプチャはマーカや機器を人体に取り付ける必要が無く、マンマシンインタラクション、サーベイランス、大量映像の検索など、様々な応用の拡大が見込めるため、数多くの研究 [5] ~ [7] が行われている。

さらに、計算機性能の向上に伴ってより豊かな情報として3次元形状まで実時間で復元することが可能になってきてい

る [8], [9]。3次元形状の参照により、2次元画像からのみでは困難な複雑な姿勢・運動の解析も可能となる。また、パフォーマンスキャプチャ [10] ~ [12] や CG アニメーション [13] ~ [15] などの応用では3次元形状まで復元されていることが必須である。こうした理由から、着衣も含んだ人の形状・動きまでもカメラ映像から3次元的に復元することへの期待は高い。

画像からの推定において問題となるのは、撮影・復元誤差、手足や障害物による遮蔽、人体の対称性に起因する曖昧性、予測が難しい高速で複雑な動き、などである。これらの問題に対処するため、「人体の動きに関する事前知識」の利用が注目を集めている。

本稿では、こうした事前知識を参照した手法を2種類に大別して紹介した後、著者らによる「画像からの復元形状誤差への対処を伴った高精度推定」や「複数種類の動き（歩行、ジョギングなど）の統一的な事前知識のモデル化による汎用性の向上」という特徴を実現した手法について述べる。

## 2. 人体の姿勢・運動推定における事前知識

人体の姿勢や運動の推定において事前知識を利用する際には、モーションキャプチャシステムによって解となる人体骨格構造（各関節の3次元位置・角度から構成されるスケルトン構造）の静的・動的特徴を事前に計測・モデル化しておく<sup>(注1)</sup>。推定時には、このモデルと画像特徴量や推定履歴を比較することによって推定の頑健化・高精度化を実現する。

ただし、事前のモーションキャプチャシステムによる計測・モデル化には以下のような問題がある。

個人間の差：撮影対象ごとにモデルを用意することは現実的に不可能なため、モデル化した人間と推定対象である人間は異なることが一般的である。

個人内の差：同じ人間の動きでも、その時々によって動きは微妙に異なる。

不自然な計測：現存するいずれのモーションキャプチャシステムでも、自然でリアルな動きの計測は難しい：モーションキャプチャシステムが1) 光学式であればマーカを人体に貼り付ける必要があるために服装が限定されてしまい、2) センサ式であればその重量や締め付けで自然な動きが阻害されてしまう。

よって、モデル化された動きをそのまま推定値とみなすことが許されないのは当然であり、より汎化されたモデルを学習しておくことが重要な課題の一つとなる。

以下、本章では従来法を2種類に大別して紹介する。

### 2.1 運動予測：動きの事前知識

各時刻における姿勢推定ではなく、過去の姿勢を参照した追跡（運動推定）を行うことでより安定に姿勢を得る研究が数多く行われている。追跡問題においては、追跡履歴からの現時刻における状態予測の精度が重要な要素である。追跡履歴のみを参照した手法では、カルマンフィルタなどの予測器を利用して現時刻における状態予測を行っている（文献[18]など）。

履歴のみを参照した予測は、追跡対象の種類やその動きに依存することなく適用可能である一方で、複雑で急激な動きの変化に対応するのが難しいという問題がある。この問題に対して「歩行やジョギングなどの動作ごとに、それぞれにおける身体の動きを事前知識として学習」することによって安定かつ正確に動きを予測する手法が提案されてきている。

動きの事前知識を利用する手法は、まず学習サンプルになる動きデータをモーションキャプチャシステムによって獲得し、その動きデータの中から撮影された動きに最も適合するものを探索することによって動きの予測を行っている。この種の手法においては、どのように大量の学習サンプルを記録して、どのように撮影データに適合する学習サンプルを探索するかが技術的課題となる。

一般的な方法は、大量データからそのパラメトリックな近似表現を事前に獲得しておくことである。正確かつ効率的な予測のために次のように様々なモデルが利用されている：線

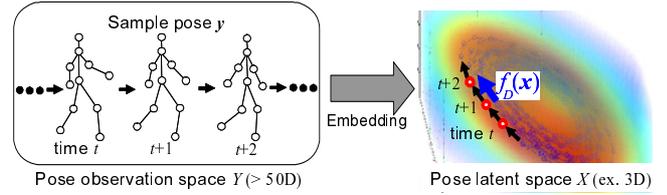


図1 姿勢パラメータ空間  $Y$  とその低次元潜在空間  $X$ 。  $X$  中の小円と矢印は、それぞれ各姿勢の潜在変数とそれらの間の時系列遷移の写像  $f_D(x)$  を示す。

Fig. 1 Pose space  $Y$  and its latent space  $X$ . Circles and arrows in  $X$  depict latent variables and temporal mapping,  $f_D(x)$ , respectively.

形補間[19]、可変線形システム[20]、ガウス混合モデル[21]~[23]、隠れマルコフモデル[24]、可変長マルコフモデル[25]、多様体[26]、自己回帰モデル[27]、混合自己回帰モデル[28]。

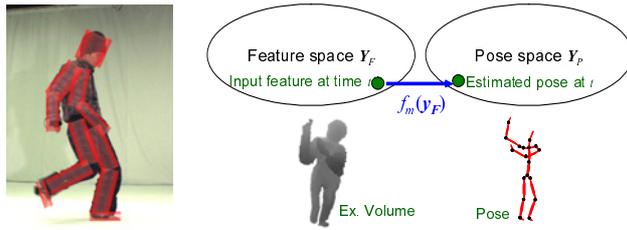
これらの手法では学習サンプルから得られるそのパラメトリックな表現のみを学習データとして記録しているが、学習サンプルをそのまま（またはいくつかの代表サンプルを）記録しておくことによって、より正確な予測を実現する事例ベースモデルも提案されている。これらのモデルでは、どのようにして大量かつ高次元（人体全身で50程度の自由度）の学習サンプルから適切な動きデータを正確かつ効率的に探索するかが主な技術的課題として挙げられる：主成分分析で得られる低次元空間における探索木[29]、ギブスサンプリングと勾配上昇法の組み合わせ[30]、Locally Linear Embedding (LLE)[31]で得られた低次元空間における Relevance Vector Machine (RVM) による回帰[32]。

パラメトリックモデルと事例ベースモデルのいずれにおいても、観測された高次元の学習サンプルを低次元潜在空間において効率的に学習しているモデルが多い。図1に例を示す：高次元の全身姿勢の学習サンプル（Sample pose sequences:  $y$ ）から低次元空間  $X$  における潜在変数  $x$  を求め、その低次元潜在空間における時系列遷移  $f_D(x)$  を推定している。

従来は線形の主成分分析が主に利用されていたが、複雑非線形な全身関節の動きを表現するためには LLE や Isomap[33] が利用されている。更に近年では、非線形性に加えて学習サンプルの確率的な要素も表現できるモデルが提案され（Gaussian Process Latent variable Models, GPLVM[34]など）、その低次元潜在空間における運動予測関数の最適化まで実現されており（Gaussian Process Dynamical Models, GPDM[35]）、安定な姿勢追跡への応用も提案されている[36]。

これらの低次元潜在空間化は、高次元観測空間と低次元潜在空間の間の滑らかな双方向写像化[37]、データの階層化[38]、低次元潜在空間におけるデータの構造を既知の構造に合わせる空間最適化（例：歩行がサイクリックな動きであるなど）[39]、低次元潜在空間の次元数の最適化[40]など、数多くの拡張が研究されている。

(注1): 近年、このような身体のモーションキャプチャデータは動きのモデル化や手法の性能評価を目的として多数配信されている[11], [16], [17]。



(a) Model matching: generative approach. (b) Feature-to-pose mapping: discriminative approach.

図 2 撮影画像からの姿勢パラメータ推定 .

Fig. 2 Pose parameter estimation from an observed image.

## 2.2 画像から姿勢への写像：画像特徴量と人体姿勢の間の事前知識

追跡の有無に関係なく、従来は「撮影画像から得られた特徴量と姿勢パラメータに相当する特徴量（例：図 2(a) に示すような全身シルエットと姿勢パラメータにあわせて変形したシリンダモデル）が一致する」ように姿勢パラメータが推定されていた（文献 [41], [42] など）。このような特徴量と姿勢パラメータの関係（写像）を、運動予測における前時刻と現時刻の関係のように事前に学習しておく手法がある [43] ~ [48]。図 2(b) の例では、画像からの復元形状  $y_F$  を特徴量として、そこから姿勢パラメータへの写像  $f_m(y_F)$  が学習されている。

特徴量のマッチングによる姿勢推定は、ある姿勢パラメータにおいて観測されるはずの特徴量（シルエットなど）が必要なため、全身の各体節を剛体近似している手法がほとんどである。スカートのような緩い着衣の形状変化までパラメトリックに表現することにより特徴量を得る手法も存在するが [49], [50]、着衣の大きな形状変化の正しいモデル化が難しく、また計算コストが高いという問題を含んでいる。一方、写像ベースの手法では着衣の大きな形状変化も容易に取り扱うことができる [48]。

こうした手法における技術的課題は、高次元な画像特徴量と姿勢パラメータ間の写像を大量の学習サンプルからどのようにしてモデル化するかという点である。データの次元性に対しては、観測画像や復元形状からの効率的かつノイズや遮蔽に対して頑健な低次元特徴量抽出に加えて（例：画像からの特徴量 [51] や復元形状からの特徴量 [52] ~ [55]）、運動予測でも利用されている低次元潜在空間化が広く利用されている：特徴量空間と姿勢パラメータ空間の低次元潜在空間の間の写像を推定する。特徴量空間と姿勢パラメータ空間の間の写像の推定においては、高速性と汎化性能が重要な課題である。高速性実現のためには、ハッシング（画像からのハッシング [44]、復元形状からのハッシング [46]）や回帰関数（画像からの回帰 [43], [45]、復元形状からの回帰 [47], [48]）が利用されている。手法 [56] のように、複数観測空間の低次元潜在空間とそこを介した写像の最適化を同時に行っている [57] 手法も存在する。

## 3. 運動予測と写像を統合した手法

本章では、ここまで述べてきた事前知識に基づいた手法のより詳細な研究例として、著者らによる運動予測と写像による姿勢推定を統合した手法を紹介する。



図 3 復元形状の誤差修正と体節分割 . 1 段目：撮影画像，2 段目：体節分割結果 .

Fig. 3 Volume refinement and body-part labeling. 1st row: observed images, 2nd row: Body-part labeled voxels.

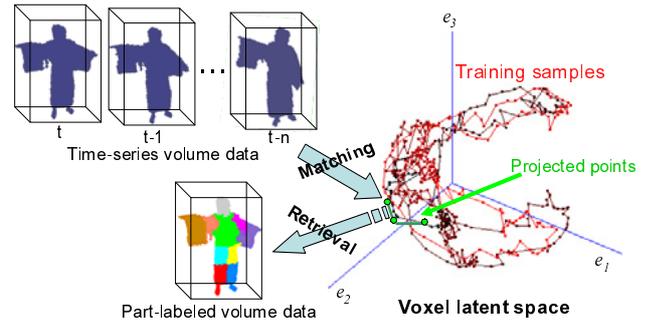


図 4 低次元潜在空間における時系列形状のマッチング，およびマッチング結果からの体節分割データの検索

Fig. 4 Matching time-series volumes in their latent space and retrieving body-part labeled voxels from the matching result.

### 3.1 復元形状の誤差修正と体節分割

人体の復元形状に対して、復元誤りの修正と各体節領域の特定を同時に実時間で実現する手法を提案した [58]。図 3 に実験結果を示す。この手法は、形状変化の大きな非剛体着衣にも適用であるという特徴を持つ。この手法では、運動予測の代わりに時系列マッチングを、写像の代わりに直接検索を行っている。

学習形状の生成にはオフラインで時間をかけても問題無いので、Photo consistency を利用した形状復元 [59], [60] などによって正しい対象形状に近い形状を獲得しておく。形状中の体節ラベル（頭、胴体など）は、特定したい体節ごとに着色された着衣の 3 次元復元結果から直接獲得する。

各学習形状にはその体節分割結果も対応付けて記録しておくことにより、入力形状に類似する学習形状が探索されると（図 4 中の "Matching"）それに対応する形状誤差の少ない体節分割結果も得られる（図 4 中の "Retrieval"）：実際には複数の類似データの重み付き投票で最終的な体節分割結果を得る。

ただしオンライン処理のため、入力形状は高速な Shape-from-Silhouette により得られる近似形状（視体積）である。よって、人体の姿勢が同じでもその入力形状と学習形状は異なる。この問題に対処するため、各学習形状とそこから計算される視体積集合との差分を誤差候補とみなし、学習形状と入力形状のマッチングにおいて重みを下げる。この重み付けによって、視体積の誤差に影響されにくいマッチングを実現した。

探索も実時間で行うため、学習データである時系列形状は主成分分析され、形状変化はその低次元潜在空間中に記録される。体節分割結果は、その形状に対応付けて記録しておく。

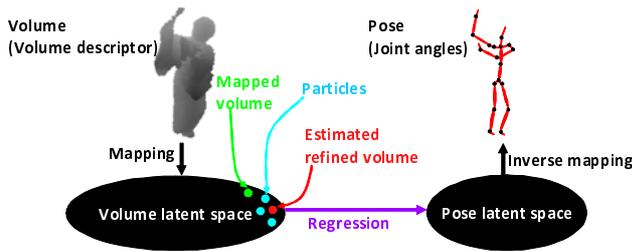


図 5 低次元潜在空間を介した復元形状から姿勢パラメータへの写像  
Fig. 5 Volume-to-pose regression via their latent spaces.

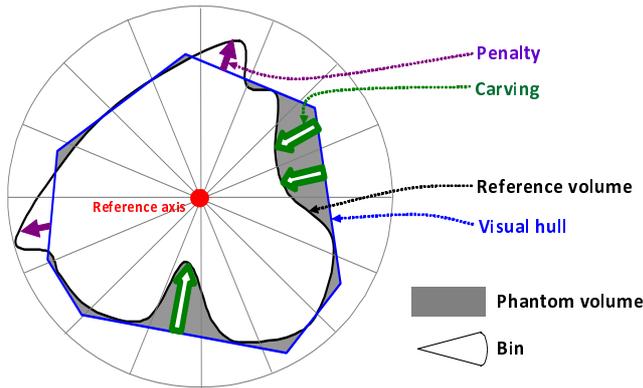


図 6 形状マッチングのための視体積制約：図は水平面に平行な形状の断面図。  
Fig. 6 Visual hull constraints for volume matching.

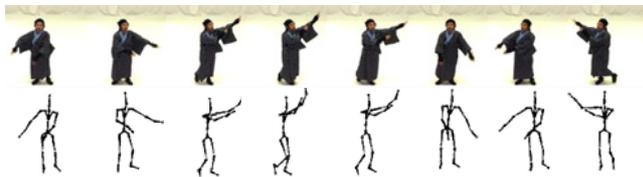


図 7 復元形状誤差を考慮した形状追跡と人体姿勢推定．1 段目：撮影画像，2 段目：姿勢推定結果．  
Fig. 7 Human volume tracking with error handling and pose estimation. 1st row: observed images, 2nd row: Estimated poses.

### 3.2 復元形状誤差を考慮した形状追跡と人体姿勢推定

着衣の大きな形状変化や自己遮蔽に対して頑健な 3 次元形状からの人体姿勢推定を実現した [48] .

学習時は、手法 [58] と同様に正しい対象形状に近い形状とそれに同期した人体姿勢を獲得して、それらの低次元潜在空間の間で形状<sup>(注2)</sup>から姿勢への写像 (図 5 中の "Regression") を学習しておく。手法 [58] の検索ベースの学習データ参照と異なり、回帰ベースの写像によって汎化性能を向上できる。また、運動予測に基づいた形状追跡によって姿勢推定を安定かつ高精度化させるため、形状の低次元潜在空間でパーティクルフィルタリングを行う：各時刻において全パーティクル (図 5 中の "Particles") と入力形状 (図 5 中の "Mapped volume") の尤度を計算し、それらの尤度重み付き平均により各瞬間の推定

形状 (図 5 中の "Estimated refined volume") を得る。各時刻のパーティクルは、学習サンプルから推定された運動予測にしたがって前時刻のパーティクルを移動させて得られる。この手法 [48] では、低次元潜在空間と運動予測を同時に獲得するため GPDM [35] を利用した。

この手法 [48] においても、オンライン処理は高速性実現のために対象形状は視体積として得る。よって、手法 [58] と同じく入力形状と正しい姿勢に写像される学習形状、すなわち各パーティクルとの差を考慮する必要がある。手法 [58] では「各学習形状における誤差確率」を事前に学習しているため、学習形状のいずれかと入力形状との間の誤差評価は高速に行うことができる一方、任意の形状と入力形状との間の誤差評価ができない。この問題は、学習形状からの検索ベースの手法 [58] と任意形状からの回帰ベースの本手法 [48] との違いに起因する。

そこで、以下の基準に従いオンラインで任意形状 (パーティクル) とその入力形状 (視体積) 間の誤差を評価する：図 6 参照。

- Carving: 形状の重心から見て入力形状表面のほうがパーティクル表面よりも外側にある場合、その入力形状表面は誤差とみなしてパーティクルの表面に近づける。

- Penalty: 逆にパーティクルの表面のほうが外側にある場合、視体積の性質上ありえない形状なのでそのパーティクルの尤度にペナルティを与える。

こうして誤差修正またはペナルティを与えた上で、入力形状とパーティクルの尤度をユークリッド距離に基づいて計算する。

こうして得られた各瞬間の推定形状から、対応する姿勢パラメータが推定される。撮影画像と推定結果を図 7 に示す。視体積が誤差を多く含み、着衣が人体を覆い隠しているような画像からでも適切な姿勢推定が実現できていることがわかる。

### 3.3 複数動作の動きの事前知識の統合モデル

著者らの手法 [48], [58] も含めて、類似従来法は単一の動作 (歩行, ジョギングなどのいずれか) のみの事前知識をモデル化している。複数の動作をそれぞれ独立に記録した学習サンプル (例: [11], [16], [17] などのモーションキャプチャデータベース) をそのまま 1 つにモデル化しても、動作間の遷移中の状態はモデル化されていないので、1) その間の学習サンプルが無い場合写像による姿勢推定が難しい、2) その間のダイナミクスが未知である (運動予測ができない) ため追跡追従性が落ちる、という問題がある。

この問題は、動作ごとの既存の計測データから遷移動作を生成して学習モデルに組み込んでおくことにより回避できる。遷移動作を自動的に生成する手法はキャラクターアニメーションの分野で数多く研究されている [61], [62]。こうした手法を応用して、複数動作をそれぞれモデル化した低次元空間の集合を生成する際、それらの間の遷移経路を学習データに加えることで姿勢追跡における動作切替時の追従性および姿勢追跡精度を向上させる手法を提案した。

形状復元、低次元潜在空間におけるパーティクルフィルタリング、追跡形状からの写像による姿勢推定、というフレームワークは手法 [48] と同様である。手法 [48] との違いは、1) 動作ごとに低次元潜在空間が用意され、それらの間では類似した潜在

(注2): 実際には、復元形状から得られる効率的な特徴量 (文献 [46], [47] などで利用されている特徴量と類似した形状記述子) を利用した。

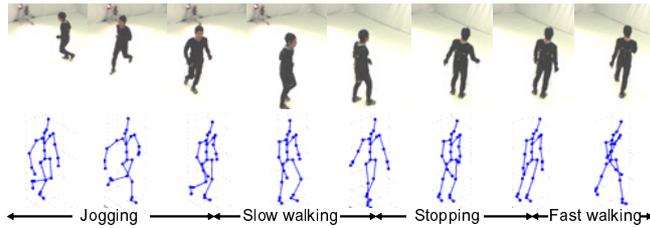


図 8 複数動作シーケンスにおける姿勢推定 . 1 段目 : 撮影画像 , 2 段目 : 姿勢推定結果 .

Fig.8 Pose estimation with multiple actions. 1st row: observed images, 2nd row: Estimated poses.

変数同士が遷移動作生成 [61] によって結ばれており, 2) 入力形状は全潜在空間に投影されてその中に分布しているパーティクルの尤度計算に利用され, 3) 保有する全パーティクルの尤度の和が最大となる動作空間でその時刻における推定形状が得られている, という点である. 歩行に類する 6 種類の動作からなる統合モデルを利用した姿勢推定結果を図 8 に示す.

#### 4. む す び

事前知識を利用して人の形状, 姿勢, 運動を安定かつ精度よく推定する手法を紹介した. 数ある事前知識の中から, 本稿では特に運動予測と画像特徴量から人体の姿勢・体節形状パラメータへの写像を学習する手法について述べた.

さらなる研究の方向性としては, より多様な事前知識 (例: 関節可動域 [63], 物理運動制約 [64], [65]) の利用や, より豊かな特徴量を利用したマッチング (例: メッシュマッチング [11], [66] など) などによる推定の安定性と精度の向上などが挙げられる.

形状修正 [60] については京都大学延原章平助教に, GPDM [35] については Dr. Neil D. Lawrence にソフトウェアを提供して頂いた. 深謝致します.

#### 文 献

- [1] K. Yamane, A. Murai, S. Takaya, and Y. Nakamura, "Muscle tension database for contact-free estimation of human somatosensory information," In *ICRA*, 2009.
- [2] N. S. Pollard, J. K. Hodgins, M. J. Riley, and C. G. Atkeson, "Adapting Human Motion for the Control of a Humanoid Robot," In *ICRA*, 2002.
- [3] S. Nakaoka, A. Nakazawa, K. Yokoi, and K. Ikeuchi, "Leg Motion Primitives for a Dancing Humanoid Robot," In *ICRA*, 2004.
- [4] K. Grochow, S. Martin, A. Hertzmann, and Z. Popovic, "Style-Based Inverse Kinematics," In *SIGGRAPH*, 2004.
- [5] D. M. Gavrilu, "The Visual Analysis of Human Movement: A Survey," *CVIU*, Vol.73, No.1, pp.82–98, 1999.
- [6] T. B. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture," *CVIU*, Vol.81, No.3, pp.231–268, 2001.
- [7] R. Poppe, "Vision-based human motion analysis: An overview," *CVIU*, Vol.108, No.2, pp.4–18, 2007.
- [8] G. Cheung, T. Kanade, J. Bouguet, and M. Holler, "A real time system for robust 3D voxel reconstruction of human motions," *CVPR*, 2000.

- [9] X. Wu, O. Takizawa, and T. Matsuyama, "Parallel Pipeline Volume Intersection for Real-Time 3D Shape Reconstruction on a PC Cluster," *The 4th IEEE International Conference on Computer Vision Systems (ICVS)*, 2006.
- [10] T. Matsuyama, X. Wu, T. Takai, and S. Nobuhara, "Real-Time 3D Shape Reconstruction, Dynamic 3D Mesh Deformation, and High Fidelity Visualization for 3D Video," *CVIU*, Vol.96, No.3, pp.393–434, 2004.
- [11] D. Vlastic, I. Baran, W. Matusik, and J. Popovic, "Articulated Mesh Animation from Multi-view Silhouettes," *ACM Transactions on Graphics*, Vol.27, No.3, 2008.
- [12] E. de Aguiar, C. Stoll, C. Theobalt, N. Ahmed, H.-P. Seidel, and S. Thrun, "Performance Capture from Sparse Multi-view Video," *ACM Transactions on Graphics*, Vol.27, No.3, 2008.
- [13] R. W. Sumner, J. Schmid, and M. Pauly, "Embedded deformation for shape manipulation," *ACM Transactions on Graphics*, Vol.26, No.3, 2007.
- [14] S. Kircher and M. Garland, "Free-form Motion Processing," *ACM Transactions on Graphics*, Vol.27, No.2, 2008.
- [15] I. Baran, D. Vlastic, E. Grinspun, J. Popovic, "Semantic Deformation Transfer," *ACM Transactions on Graphics*, Vol.28, No.3, 2009.
- [16] L. Sigal and M. J. Black, "HumanEva: Synchronized Video and Motion Capture Dataset for Evaluation of Articulated Human Motion," Technical Report CS-06-08, Brown University, 2006. <http://vision.cs.brown.edu/humaneva/>
- [17] CMU Graphics Lab Motion Capture Database: <http://mocap.cs.cmu.edu/>
- [18] D.-S. Jang, S.-W. Jang, and H.-I. Choi, "2D human body tracking with Structural Kalman filter," *Pattern Recognition*, Vol.35, No.10, pp.2041–2049, 2002.
- [19] R. Urtasun, D. Fleet, and P. Fua, "Temporal motion models for monocular and multiview 3D human body tracking," *CVIU*, Vol.104, No.2, pp.157–177, 2006.
- [20] V. Pavlovic, J. M. Rehg, and J. Maccormick, "Learning Switching Linear Models of Human Motion," *NIPS*, 2000.
- [21] N. How, M. Leventon, and W. Freeman, "Bayesian Reconstruction of 3D Human Motion from Single-Camera Video," *NIPS*, 1999.
- [22] N. Huazhong, T. Tan, L. Wang, and W. Hu. "People tracking based on motion model and motion constraints with automatic initialization," *Pattern Recognition*, Vol.37, No.7, pp.1423–1440, 2004.
- [23] C. Sminchisescu and A. Jepson, "Variational Mixture Smoothing for Non-Linear Dynamical Systems," *CVPR*, 2004.
- [24] M. Brand, "Shadow Puppetry," *ICCV*, 1999.
- [25] S. Hou, A. Galata, F. Caillette, N. Thacker, and P. Bromiley, "Real-time Body Tracking Using a Gaussian Process Latent Variable Model," *ICCV*, 2007.
- [26] A. Elgammal and C.-S. Lee, "Inferring 3D Body Pose from Silhouettes using Activity Manifold Learning," *CVPR*, 2004.
- [27] X. Zhao and Y. Liu, "Tracking 3D Human Motion in Compact Base Space," *IEEE Workshop on Applications of Com-*

- puter Vision*, 2007.
- [28] A. Agarwal and B. Triggs, “Tracking Articulated Motion using a Mixture of Autoregressive Models,” *ECCV*, 2004.
- [29] H. Sidenbladh, M. J. Black, and L. Sigal, “Implicit Probabilistic Models of Human Motion for Synthesis and Tracking,” *ECCV*, 2002.
- [30] A. Fathi and G. Mori, “Human Pose Estimation using Motion Exemplars,” *ICCV*, 2007.
- [31] S. Roweis and L. Saul, “Nonlinear Dimensionality Reduction by Locally Linear Embedding,” *Science*, Vol.290, No.5500, pp.2323–2326, 2000.
- [32] T. Jaeggli, E. Koller-Meier, and L. “Multi-Activity Tracking in LLE Body Pose Space,” *2nd Workshop on HUMAN MOTION Understanding, Modeling, Capture and Animation*, 2007.
- [33] J. Tenenbaum, V. de Silva, and J. Langford, “A Global Geometric Framework for Nonlinear Dimensionality Reduction,” *Science*, Vol.290, No.5500, pp.2319–2323, 2000.
- [34] N. D. Lawrence, “Probabilistic non-linear principal component analysis with Gaussian process latent variable models,” *Journal of Machine Learning Research*, Vol.6, pp.1783–1816, 2005.
- [35] J. M. Wang, D. J. Fleet, A. Hertzmann, “Gaussian Process Dynamical Models for Human Motion,” *PAMI*, Vol.30, No.2, pp.283–298, 2008.
- [36] R. Urtasun, D. Fleet, and P. Fua, “3D People Tracking with Gaussian Process Dynamical Models,” *CVPR*, 2006.
- [37] N. D. Lawrence, “Local distance preservation in the gp-lvm through back constraints,” *ICML*, 2006.
- [38] N. D. Lawrence and A. J. Moore, “Hierarchical Gaussian process latent variable models,” *ICML*, 2007.
- [39] R. Urtasun, D. J. Fleet, A. Geiger, J. Popovic, T. Darrell, and N. D. Lawrence, “Topologically-Constrained Latent Variable Models,” *ICML*, 2008.
- [40] A. Geiger, R. Urtasun, and T. Darrell, “Rank Priors for Continuous Non-Linear Dimensionality Reduction,” *CVPR*, 2009.
- [41] I. Mikić, M. Trivedi, E. Hunter, and P. Cosman, “Human Body Model Acquisition and Tracking using Voxel Data,” *IJCV*, Vol.53, No.3, pp.199–223, 2003.
- [42] G. Cheung, S. Baker, and T. Kanade, “Shape-from-silhouette of articulated objects and its use for human body kinematics estimation and motion capture,” *CVPR*, 2003.
- [43] A. Agarwal and B. Triggs, “3D Human Pose from Silhouettes by Relevance Vector Regression,” *CVPR*, 2004.
- [44] G. Shakhnarovich, P. Viola, and T. Darrell, “Fast Pose Estimation with Parameter Sensitive Hashing,” *ICCV*, 2003.
- [45] R. Urtasun and T. Darrell, “Sparse probabilistic regression for activity-independent human pose inference,” *CVPR*, 2008.
- [46] Y. Sagawa, M. Shimosaka, T. Mori, and T. Sato, “Fast online human pose estimation via 3D voxel data,” *IROS*, 2007.
- [47] Y. Sun, M. Bray, A. Thayananthan, B. Yuanand, and P. H. S. Torr, “Regression-based human motion capture from voxel data,” *BMVC*, 2006.
- [48] N. Ukita, M. Hirai, and M. Kidode, “Complex Volume and Pose Tracking with Probabilistic Dynamical Models and Visual Hull Constraints,” *ICCV*, 2009.
- [49] B. Rosenhahn, U. Kersting, K. Powell, R. Klette, G. Klette, H.-P. Seidel, “A system for articulated tracking incorporating a cloth model,” *Machine Vision and Applications*, Vol.18, No.1, pp.25–40, 2007.
- [50] A. O. Balan and M. J. Black, “The Naked Truth: Estimating Body Shape Under Clothing,” *ECCV*, 2008.
- [51] S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” *PAMI*, Vol.24, No.4, pp.509–522, 2002.
- [52] M. Kortgen, G.-J. Park, M. Novotni, and R. Klein, “3D shape matching with 3D shape contexts,” *7th Central European Seminar on Computer Graphics*, 2003.
- [53] A. Frome, D. Huber, R. Kolluri, T. Bulow, and J. Malik, “Recognizing Objects in Range Data Using Regional Point Descriptors,” *ECCV*, 2004.
- [54] M. Kazhdan, T. Funkhouser, and S. Rusinkiewicz, “Rotation invariant spherical harmonic representation of 3D shape descriptors,” *Eurographics/SIGGRAPH SGP*, 2003.
- [55] B. Bustos, D. A. Keim, D. Saube, T. Schreck, and D. V. Vranic, “An experimental effectiveness comparison of methods for 3D similarity search,” *International Journal on Digital Libraries*, Vol.6, No.1, pp.39–54, 2006.
- [56] C. H. Ek, P. H. S. Torr, and N. D. Lawrence, “Gaussian Process Latent Variable Models for Human Pose Estimation,” *4th International Workshop on Machine Learning for Multimodal Interaction*, 2007.
- [57] A. P. Shon, K. Grochow, A. Hertzmann, and R. P. N. Rao, “Learning Shared Latent Structure for Image Synthesis and Robotic Imitation,” *NIPS*, 2005.
- [58] N. Ukita, R. Tsuji, and M. Kidode, “Real-time Shape Analysis of a Human Body in Clothing using Time-series Part-labeled Volume,” *ECCV*, 2008.
- [59] S. M. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski, “A comparison and evaluation of multi-view stereo reconstruction algorithms,” In *CVPR*, 2006.
- [60] T. Tung, S. Nobuhara, and T. Matsuyama. “Simultaneous Super-resolution and 3D Video Using Graph-cuts,” In *CVPR*, 2008.
- [61] L. Kovar, M. Gleicher, and F. H. Pighin, “Motion graphs,” *SIGGRAPH*, 2002.
- [62] L. Zhao and A. Safonova, “Achieving good connectivity in motion graphs,” *Graphical Models Journal*, Vol.71, No.4, pp.139–152, 2009.
- [63] L. Herda, R. Urtasun, and P. Fua, “Hierarchical implicit surface joint limits for human body tracking,” *CVIU*, Vol.99, No.2, pp.189–209, 2005.
- [64] A. Safonova, J. K. Hodgins, N. S. Pollard, “Synthesizing Physically Realistic Human Motion in Low-Dimensional, Behavior-Specific Spaces,” *SIGGRAPH*, 2004.
- [65] M. Vondrak, L. Sigal, and O. C. Jenkins, “Physical Simulation for Probabilistic Motion Tracking,” *CVPR*, 2008.
- [66] J. Gall, C. Stoll, E. de Aguiar, C. Theobalt, B. Rosenhahn, and H.-P. Seidel, “Motion Capture Using Joint Skeleton Tracking and Surface Estimation,” *CVPR*, 2009.