

A data-driven kinematic model of the human hand with soft-tissue artifact compensation mechanism for grasp synergy analysis

Marco Gabiccini* and Georg Stillfried*, Hamal Marino, Matteo Bianchi

Abstract—This paper presents a methodology to accurately record human finger postures during grasping. The main contribution consists of a kinematic model of the human hand reconstructed via magnetic resonance imaging of one subject that (i) is fully parameterized and can be adapted to different subjects, and (ii) is amenable to in-vivo joint angle recordings via optical tracking of markers attached to the skin. The principal novelty here is the introduction of a soft-tissue artifact compensation mechanism that can be optimally calibrated in a systematic way. The high-quality data gathered are employed to study the properties of hand *postural synergies* in humans, for the sake of ongoing neuro-science investigations. These data are analyzed and some comparisons with similar studies are reported. After a meaningful mapping strategy has been devised, these data could be employed to define robotic hand postures suitable to attain effective grasps, or could be used as prior knowledge in lower-dimensional, real-time avatar hand animation.

I. INTRODUCTION

In the field of robotic grasping with anthropomorphic hands, robots can learn from the way humans grasp. For example, *postural synergies* [1] utilized by humans can be mapped to a robotic hand, so that hand postures can be more easily commanded. This paper deals with the challenge of how to suitably record human finger postures during grasps. To combine high accuracy with low data dimensionality, optical position measurements need to be combined with an accurate kinematic model of the hand. Commonly, such models are constructed of serial chains of rotation axes, and their kinematics is quite suitable for modelling the movement of the skeleton. In [2], a 22 Degrees of Freedom (DoFs) model of the hand has been employed to reconstruct hand posture based on optical markers data. However, what was not considered is that the skin moves relatively to the bones, introducing what is called in literature a *soft-tissue artifact* (STA) when measuring skeletal movement by tracking skin markers.

This work is supported by the European Commission under CP grant no. 248587, “THE Hand Embodied”, within the FP7-ICT-2009-4-2-1 program “Cognitive Systems and Robotics” and by ERC Advanced Grant no. 291166 “SoftHands: A Theory of Soft Synergies for a New Generation of Artificial Hands”.

M. Gabiccini is with the DIC1 and Res. Center “E. Piaggio” University of Pisa, 56122 Pisa, Italy, and with the Dep. of Advanced Robotics, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova (phone: +39-050-221.80.77, fax: +39-050-221.80.65, email: m.gabiccini@ing.unipi.it)

G. Stillfried is with the Institute of Robotics and Mechatronics, German Aerospace Center (DLR), Münchner Straße 10, 82234 Weßling, Germany (Georg.Stillfried@dlr.de)

H. Marino and Matteo Bianchi are with the Res. Center “E. Piaggio”, University of Pisa, 56122 Pisa, Italy, and with the Dep. of Advanced Robotics, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova (hamal.marino@gmail.com, matteo.bianchi@centropiaggio.unipi.it)

*) Both starred authors contributed equally to the paper.

In [3], magnetic resonance imaging (MRI) was used to measure the STA in different hand postures. Based on the model by Zhang et al. [4], we propose an extension to 3D and validate it using the aforementioned MRI data as ground truth.

The skin movement model is incorporated into a 26 DoFs kinematic hand model, which allows to take into account the differences between hands of different subjects with a calibration phase.

An algorithm, based on [5] where a probabilistic inference framework is introduced to concurrently estimate joint angles and model parameters, is developed; still, the calibration phase is kept separated in order to avoid parameters drift.

A similar approach has been used by [6], with a 29 DoFs model (23 + hand position and orientation), static geometric calibration, an assumption for STA compensation, and linearizing state estimation to retrieve joint angles. The algorithm here developed uses instead an optimization to calibrate model parameters, the used STA compensation mechanism has been validated, and postures are identified via an iterative state estimation procedure for improved results. The proposed methods are used to reconstruct hand postures during grasps of imagined objects and analyze them by means of Principal Component Analysis (PCA). Results are compared with the ones described in [7], where authors used a glove-based hand pose reconstruction system (CyberGlove; Virtual Technologies, Palo Alto, CA) to record hand postures of five subjects, who were asked to shape their right hand as if to grasp and use a large number of familiar objects.

II. DESCRIPTION OF THE KINEMATIC MODEL

The kinematic model of the human hand with respect to the forearm is devised as a *kinematic tree*, whose root node consists of the Cartesian reference frame $\{B\}$ (rigidly attached to a bracelet fastened to the forearm) and whose leaves are the frames fixed to the distal phalanxes (PDs) of the five fingers, as depicted in Figure 1. The five paths to the PDs have a common segment through the wrist joint, centered at point W , and then branch out from the MC_2 (metacarpal bone of the index) as serial kinematic chains.

A. Hand posture parametrization

To efficiently parameterize the posture of the j -th phalanx in the i -th finger chain, we employ the *Product of Exponentials* (POE) formula [8], i.e.

$$g_{BF_{ij}}(\theta_i) = \left[\prod_{k=1}^j e^{\hat{\xi}_{ik}\theta_{ik}} \right] g_{BF_{ij}}(0). \quad (1)$$

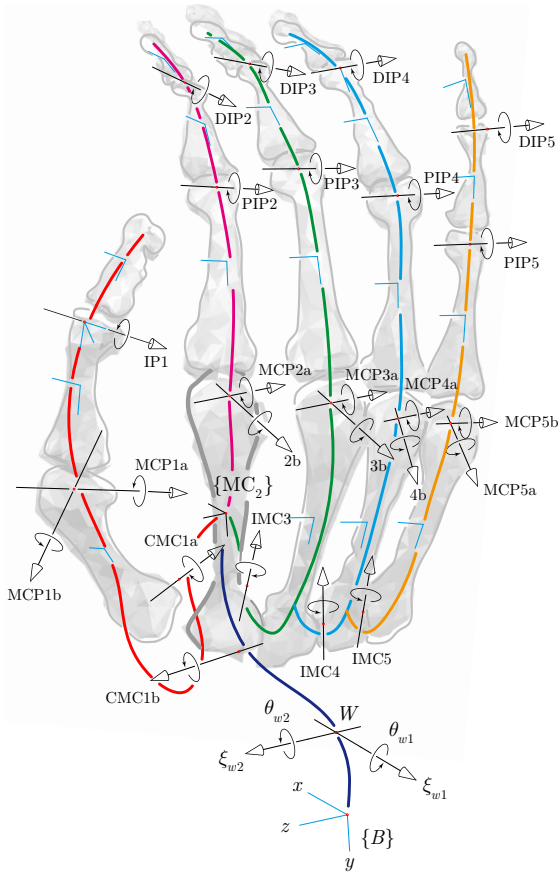


Fig. 1: Human hand kinematic model.

Here, the $\hat{\xi}_{ik}$'s are the twists of the joints defining the kinematic chain of the i -th finger, $\theta_i = [\theta_{i1} \dots \theta_{ik} \dots \theta_{ij}]^T$ are the exponential coordinates of the 2nd kind for a local representation of SE(3) (Special Euclidean group, 4×4 rototranslation matrices) [9] for the j -th phalanx in the i -th finger, and $g_{PB_{ij}}(0)$ is its initial configuration.

Twist components have to be expressed in the common base frame $\{B\}$ and in the reference posture for the hand, i.e. when $\theta_{ik} = 0$. Therefore, it is possible to write more explicitly $\xi_{ij} = {}^b\xi_{ij}$, where the left superscript describes the *reading* frame, in accordance with [10]. To express it in a different frame, one must employ

$${}^b\xi_{ij} = \text{Ad}_{g_{ba}(0)} {}^a\xi_{ij}, \quad (2)$$

where $g_{ba}(0)$ is the posture of $\{A(0)\}$ w.r.t. $\{B\}$ and the *Adjoint* transformation Ad_g is a map between different expressions of the *same* twist in different *reading* reference frames.

B. Hand velocity parameterization

The POE parameterization of the hand posture can be profitably employed for computing the *rigid-body* velocity of each phalanx. As we shall see, this quantity is key for systematically calculating the linear velocity of the optical markers attached to the bone or to the skin.

The rigid-body velocity $\widehat{V}_{BF_{ij}}^{F_{ij}}$ of $\{F_{ij}\}$ in the moving frame $\{F_{ij}\}$ is given (as a 4×4 matrix) by the following formula

$$\widehat{V}_{BF_{ij}}^{F_{ij}} := g_{BF_{ij}}^{-1} \dot{g}_{BF_{ij}} = \begin{bmatrix} \widehat{\omega}_{BF_{ij}}^{F_{ij}} & v_{BF_{ij}}^{F_{ij}} \\ 0 & 0 \end{bmatrix} \quad (3)$$

where, given $R_{BF_{ij}}$ the 3×3 rotation matrix from $\{B\}$ to $\{F_{ij}\}$, $\widehat{\omega}_{BF_{ij}}^{F_{ij}} := R_{BF_{ij}}^T \dot{R}_{BF_{ij}}$ is the skew-symmetric matrix of the angular velocity components (in $\{F_{ij}\}$) of $\{F_{ij}\}$ w.r.t. $\{B\}$, and $v_{BF_{ij}}^{F_{ij}} = R_{BF_{ij}}^T \dot{d}_{BF_{ij}}$ are the components (in $\{F_{ij}\}$) of the velocity of the origin $O_{F_{ij}}$ with respect to O_B . Equation (3) can be rewritten (as a 6×1 vector) in a convenient form by factoring out the joint velocities $\dot{\theta}_i$ of the i -th finger as follows

$$V_{BF_{ij}}^{F_{ij}} = J_{BF_{ij}}^{F_{ij}}(\theta_i) \dot{\theta}_i, \quad (4)$$

where the *distal Jacobian* $J_{BF_{ij}}^{F_{ij}}$ can be computed as ($k = 1, \dots, j$)

$$J_{BF_{ij}}^{F_{ij}}(\theta_i) = [\xi_1^\dagger \dots \xi_j^\dagger]; \quad \xi_k^\dagger = \text{Ad}_{g_{k+1, F_{ij}}}^{-1} \xi_k, \quad (5)$$

where we defined $g_{k+1, F_{ij}} := e^{\widehat{\xi}_{k+1}\theta_{k+1}} \dots e^{\widehat{\xi}_j\theta_j} g_{BF_{ij}}(0)$, with $g_{j+1, F_{ij}} = g_{BF_{ij}}(0)$.

Since we will be interested in calculating the velocity $v_{BF_{ij}}^{B, [O_{F_{ij}}]}$ of the origin $O_{F_{ij}}$ w.r.t. O_B in $\{B\}$ components¹, we also define the *hybrid* rigid-body velocity

$$V_{BF_{ij}}^{B, [O_{F_{ij}}]} := \begin{bmatrix} \widehat{\omega}_{BF_{ij}}^{B, [O_{F_{ij}}]} & v_{BF_{ij}}^{B, [O_{F_{ij}}]} \\ 0 & 0 \end{bmatrix} = \text{Ad}_{(R_{BF_{ij}}, 0)} V_{BF_{ij}}^{F_{ij}}. \quad (6)$$

A convenient form where the joint velocities $\dot{\theta}_i$ of the i -th finger are factored out is given by the following expression

$$V_{BF_{ij}}^{B, [O_{F_{ij}}]} = J_{BF_{ij}}^{B, [O_{F_{ij}}]}(\theta_i) \dot{\theta}_i, \quad (7)$$

where, defining $\xi_k^\# = \text{Ad}_{(R_{BF_{ij}}, 0)} \text{Ad}_{g_{F_{ij}, k+1}} \xi_k$, the *hybrid Jacobian* $J_{BF_{ij}}^{B, [O_{F_{ij}}]}$ can be computed as ($k = 1, \dots, j$)

$$J_{BF_{ij}}^{B, [O_{F_{ij}}]}(\theta_i) = [\xi_1^\# \dots \xi_j^\#];, \quad (8)$$

C. Modelling bone markers

To use the *local* frame $\{M_{ij}\}$ with the origin coincident with the *optical marker* attached to the same phalanx, we could simply add a *local* transformation such that

$$g_{BM_{ij}}(\theta_i) = g_{BF_{ij}}(\theta_i) g_{F_{ij}M_{ij}}. \quad (9)$$

¹It is worth observing that the apparently cumbersome expression $v_{BF_{ij}}^{B, [O_{F_{ij}}]}$ is simply $d/dt(p_{BF_{ij}})$. The indirect route followed for its derivation is due to the different meanings of the linear velocities $v_{BF_{ij}}^{F_{ij}}$ and $v_{BF_{ij}}^B$.

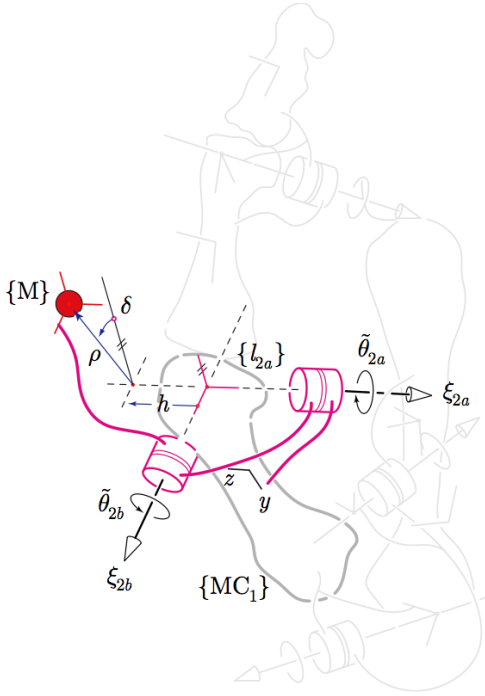


Fig. 2: Model for markers on a joint moving with the skin.

D. Modelling joint markers

For optical markers close to areas where the skin stretch is not negligible, a kinematic model considering the marker frame *fixed* to some bone would give poor results because of STA, as shown by the experiments presented in Sec. III. In order to tackle this, we include a skin stretch compensation mechanism specific for *joint markers*.

Joint marker displacement is influenced by the rotations of the joints above which it is positioned, besides the rotations of the proximal joints. However, the rotations of the last joints in the virtual chain $\tilde{\theta}(\theta)$ are (possibly nonlinear) functions of the joint angles θ . To keep things relatively simple we modeled this with linear functions

$$\tilde{\theta}_i(\theta_i) = c_i \theta_i. \quad (10)$$

Then, there is a constant *offset* transformation to account for the posture of the marker frame in the initial configuration: this can be recovered by direct inspection of Figure 2. Each joint marker is characterized by the parameters ρ, δ, h, c_i . ρ, δ and h are the cylindrical coordinates of the marker position in the initial configuration, while c_i describe the skin displacement ($c_i = 0$, marker fixed to the proximal bone; $c_i = 1$, marker fixed to the distal bone). These parameters can be calibrated once, as described in Sec. IV.

It is important to note that, given the joint marker direct kinematics, the *rigid-body velocity* of the marker frame can be computed by combining the contributions of (i) the kinematic chain of the bones $J_{l0,m}^m$, and (ii) the Jacobian $\partial\tilde{\theta}/\partial\theta$ of the joint speeds $\tilde{\theta}$ of *virtual chain* with respect to

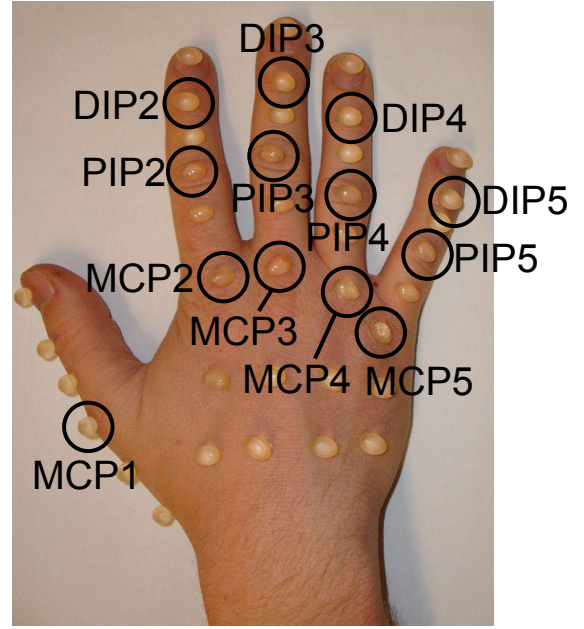


Fig. 3: Placement of the MRI-sensitive markers on the skin. For validating the marker movement model, markers near the MCP, PIP and DIP joints are considered.

the independent joint speeds θ , as follows

$$V_{l0,m}^m(\theta, \dot{\theta}) = J_{l0,m}^m[\theta_{1a}, \theta_{1b}, \tilde{\theta}_{2a}(\theta_{2a}), \tilde{\theta}_{2b}(\theta_{2b})] \frac{\partial\tilde{\theta}}{\partial\theta} \Big|_{\tilde{\theta}(\theta)} \dot{\theta} \quad (11)$$

When the linear model in eq. (10) is adopted for the description of the skin displacement, the explicit expression for $\partial\tilde{\theta}/\partial\theta$ is given by

$$\frac{\partial\tilde{\theta}}{\partial\theta} \Big|_{\tilde{\theta}(\theta)} = \text{diag}(1, 1, c_{2a}, c_{2b}). \quad (12)$$

Then, the *hybrid* version of eq. (11), can be recovered as shown in eq. (7), and will be useful in the computation of the *residual Jacobian*, as it will be explained in Sec. V, eq. (28).

III. VALIDATION OF THE SKIN MOVEMENT MODEL USING MAGNETIC RESONANCE IMAGING

Here we use the MRI data measured in [3] as ground truth to validate the skin movement model described in Sec. II-D.

A. Measurement of skin movement

MRI-sensitive Soledum oil capsules (Casella-med, Cologne, Germany; spheroids with diameter 7 mm and long axis 10 mm) are attached to the skin on the dorsal side of the hand of one subject (Fig. 3) over the metacarpophalangeal (MCP), proximal interphalangeal (PIP) and distal interphalangeal (DIP) joints.

One posture (flat hand) is designated as the reference posture.

The soft tissue artefact (STA) is quantified as the distance between the marker in the reference posture and the marker in another posture:

$$s_k = \|\|{}^P p_{m,k} - {}^B p_{m,\text{ref}}\|\|, \quad (13)$$

where s_k is the amount of uncompensated STA in posture k , ${}^B p_{m,k}$ is the position of the marker with respect to bone B in posture k and ${}^B p_{m,\text{ref}}$ is the position of the marker with respect to the bone in the reference posture. Without loss of generality, we choose to calculate with respect to the proximal bone. The mean STA and its standard deviation are shown in the second column of Table I.

B. Identification of joint axes

Our choice is to model PIP and DIP joints with one axis of rotation (1 DoF), and the MCP joints with two intersecting orthogonal axes of rotation (2 DoFs).

In 1-DoF joints, the modelled pose ${}^P T_{D,\text{mod},k}$ of the distal bone D with respect to the proximal bone P in posture k can be obtained via an homogeneous transformation:

$${}^P T_{D,\text{mod},k} = H({}^P p_{a1}, {}^P a_1, q_{1k}) {}^P T_{D,0}, \quad (14)$$

where ${}^P p_{a1} \in \mathbb{R}^3$ is a point on the rotation axis, ${}^P a_1 \in \mathbb{R}^3$ is the rotation axis, q_{1k} is the joint angle in posture k and the operator $H(p, a, q)$ produces an homogeneous matrix that describes the general rotation:

$$H(p, a, q) = \begin{pmatrix} R_a(q) & (p - R_a(q)p) \\ 0 & 1 \end{pmatrix}, \quad (15)$$

where $R_a(q)$ is a matrix that describes a rotation around an axis a that passes through the origin of the coordinate system, and q is the rotation angle [9].

As initial pose, the measured pose in the reference posture is taken:

$${}^P T_{D,0} = {}^P T_{D,\text{ref}}. \quad (16)$$

In 2-DoFs joints, the rotation takes place around two rotation axes:

$${}^P T_{D,\text{mod},k} = H({}^P p_{a1}, {}^P a_1, q_{1k}) H({}^P p_{a2}, {}^P a_2, q_{2k}) {}^P T_{D,0}. \quad (17)$$

The assumption of intersecting orthogonal axes imposes:

$${}^P p_{a1} = {}^P p_{a2} \text{ and } {}^P a_1 {}^P a_2 = 0. \quad (18)$$

The residual transformation $T_{\text{diff},k}$ between the modelled bone pose ${}^P T_{D,\text{mod},k}$ and the measured bone pose ${}^P T_{D,k}$ is

$$T_{\text{diff},k} = {}^P T_{D,k} ({}^P T_{D,\text{mod},k})^{-1} = \begin{pmatrix} R_{\text{diff},k} & t_{\text{diff},k} \\ 0 & 1 \end{pmatrix}. \quad (19)$$

Rotational $r_{r,k}$ and translational $r_{t,k}$ residuals are:

$$r_{r,k} = \text{angle}(R_{\text{diff},k}) \quad \text{and} \quad r_{t,k} = \|t_{\text{diff},k}\|. \quad (20)$$

| joint name | mean STA (mm) \pm SD | | | | optimized parameters | |
|------------|------------------------|----------------|------------------|-----------------|----------------------|-----|
| | uncompensated | default values | optimized values | cross-validated | c1 | c2 |
| MCP1 | 3.2 \pm 2.6 | 2.3 \pm 1.5 | 2.0 \pm 1.4 | 2.5 \pm 1.5 | 0.5 | 0.4 |
| MCP2 | 6.8 \pm 4.0 | 3.7 \pm 2.3 | 3.2 \pm 1.7 | 3.7 \pm 1.7 | 0.5 | 0.8 |
| PIP2 | 2.5 \pm 3.3 | 1.4 \pm 1.0 | 0.8 \pm 0.4 | 1.1 \pm 0.6 | 0.8 | |
| DIP2 | 1.7 \pm 1.9 | 1.1 \pm 0.6 | 0.9 \pm 0.4 | 1.1 \pm 0.5 | 0.6 | |
| MCP3 | 6.6 \pm 5.0 | 2.8 \pm 2.1 | 2.0 \pm 1.2 | 2.2 \pm 1.3 | 0.6 | 1.2 |
| PIP3 | 3.6 \pm 4.4 | 1.7 \pm 1.1 | 1.4 \pm 0.9 | 1.6 \pm 0.9 | 0.6 | |
| DIP3 | 1.7 \pm 1.6 | 1.7 \pm 1.2 | 1.2 \pm 0.9 | 1.1 \pm 0.9 | 0.4 | |
| MCP4 | 7.2 \pm 5.4 | 3.1 \pm 2.5 | 2.0 \pm 1.3 | 2.4 \pm 1.2 | 0.8 | 0.7 |
| PIP4 | 3.0 \pm 3.6 | 1.7 \pm 1.2 | 1.5 \pm 1.1 | 1.6 \pm 1.2 | 0.6 | |
| DIP4 | 1.7 \pm 1.7 | 1.2 \pm 0.8 | 1.0 \pm 0.7 | 1.1 \pm 0.7 | 0.6 | |
| MCP5 | 6.9 \pm 4.5 | 4.6 \pm 2.4 | 2.8 \pm 1.6 | 3.5 \pm 1.8 | 0.5 | 0.2 |
| PIP5 | 2.9 \pm 3.1 | 2.3 \pm 1.4 | 1.7 \pm 1.2 | 1.9 \pm 1.3 | 0.7 | |
| DIP5 | 2.1 \pm 1.7 | 1.7 \pm 1.1 | 1.3 \pm 1.0 | 1.4 \pm 1.0 | 0.7 | |
| mean | 3.8 \pm 3.3 | 2.3 \pm 1.5 | 1.7 \pm 1.1 | 1.9 \pm 1.1 | | |

TABLE I: Validation of the skin movement model using MRI measurements as ground truth.

The optimal axis positions ${}^P p_{a1,\text{opt}}$ and ${}^P p_{a2,\text{opt}}$, the optimal axis orientations ${}^P a_{1,\text{opt}}$ and ${}^P a_{2,\text{opt}}$, and the optimal joint angles $Q_{1,\text{opt}} = (q_{11,\text{opt}}, \dots, q_{1n_p,\text{opt}})$ and $Q_{2,\text{opt}} = (q_{21,\text{opt}}, \dots, q_{2n_p,\text{opt}})$, are identified by minimising the mean weighted sum of rotational and translational residuals.

C. Compensation of skin movement

The amount $s_{\text{residual},k}$ of residual STA is the distance between the modelled and the measured marker position:

$$s_{\text{residual},k} = \|\|{}^P p_{m,k} - {}^P p_{m,\text{mod},k}\|\|. \quad (21)$$

As described in Sec. II-D, the initial marker position ${}^P p_{m0}$ and the proportionality factors c_1 and c_2 can be optimized to best describe the skin movement. As a starting point, we set

$$\begin{aligned} {}^P p_{m0,\text{start}} &= {}^P p_{m,\text{ref}}, \\ c_{1,\text{start}} &= 0.5, \\ c_{2,\text{start}} &= 0.5. \end{aligned}$$

The mean STA and its standard deviation obtained with these values are shown in the third column of Table I.

The optimal skin movement parameters minimize the mean amount of residual STA. The simplex algorithm by Nelder and Mead [11], as implemented in Matlab, is used to find the optimal parameters.

The mean STA and its standard deviation using the optimized parameters, as well as the optimal parameters values, are shown in the fourth, sixth and seventh column of Table I, respectively.

A leave-one-out cross-validation is performed, and the resulting mean and standard deviation of the STA are shown in column five of Table I.

Notice that, although optimizing reduces the STA and generalizes well (cross-validation results are close to optimized ones), even the starting point we chose is a substantial improvement over the uncompensated values.

IV. CALIBRATION OF SUBJECT-SPECIFIC STATIC PARAMETERS

In order to reconstruct realistic values of the joint angles from markers data, the kinematic model employed should mimic as closely as possible the actual kinematics of the subject being recorded, identifying geometric parameters a_G , and the parameters a_B and a_J controlling the location of the *bone* and the *joint* markers, respectively.

At time instant k it is possible to define the residual

$$r_k = r(x_k; a_G, a_B, a_J) := y_k - f(x_k; a_G, a_B, a_J), \quad (22)$$

where y_k is the vector of markers coordinates measured by the marker-based optical tracking motion capture system at time k , and $f(x_k; a_G, a_B, a_J)$ represent the corresponding coordinates when the joint angle values are set to x_k and the static parameters are set to the values a_G , a_B and a_J . Considering a number N_p of *training* hand postures, being the residuals in a unique vector R and $x = [x_1^T \ x_2^T \ \dots \ x_{N_p}^T]^T$, the following *scalar* cost function is a measure of misfit:

$$g(x; a_G, a_B, a_J) := \frac{1}{2} R^T R = \frac{1}{2} \sum_{k=1}^{N_p} r_k^T r_k. \quad (23)$$

The calibration can thus be framed as the following *constrained least-squares minimization* problem

$$(x^*; a_G^*, a_B^*, a_J^*) = \underset{\substack{x_k \in \mathcal{D}_x \\ a \in \mathcal{D}_a}}{\operatorname{argmin}} g(x; a_G, a_B, a_J), \quad (24)$$

where $\mathcal{D}_\#$ represents a feasibility region for the associated quantity (e.g. *box constraints* for state variables or tolerance regions around some initial guesses obtained, for example, via caliper measurements for geometric parameters). In our implementation, the solution of (24) is obtained employing a *primal-dual interior point method* based on [12].

It is worth observing that the problem to be solved for the calibration works on a high number of variables as it also concerns finding the value of x^* , which is then discarded, but this is not avoidable as we do not have any *a priori* estimate of joint angles.

V. POSTURE RECONSTRUCTION

Since there are uncertainties in the model description and noisy measurements, it is appropriate to consider the framework of *probabilistic inference*.

Let x_k denote the vector of all joint angles at time k , and all parameters are considered known from the calibration. In the following, let \hat{x} indicate an estimated quantity.

Let then y denote the coordinates of all markers measured by the marker-based optical tracking motion capture system.

For the *process model* we consider a *random walk* in the state

$$x_{k+1} = x_k + v_k, \quad v_k \sim \mathcal{N}(0; V_k) \quad (25)$$

where $\mathcal{N}(\mu; M)$ is a normal distribution with mean μ and covariance M , while, for the observation model of the measurements y_k , we use the *direct kinematics* relationship

$$y_k = f(x_k) + w_k, \quad w_k \sim \mathcal{N}(0; W_k). \quad (26)$$

Hence the residual is:

$$r(y_k, x_k) = y_k - f(x_k) = w_k, \quad r_k \sim \mathcal{N}(0; W_k). \quad (27)$$

By differentiating eq. (27), the equation relating the Jacobian of the residual $J_r(x)$ and the kinematic (positional) Jacobian $J_f(x)$ is obtained as

$$J_r(x) = -J_f(x), \quad (28)$$

which can be composed using the elemental *hybrid Jacobians* $J_{BM_{ij}}^{B, [O_{M_{ij}]}}$, computed as in (7), for all the markers.

For the estimation phase, an Extended Kalman Filter (EKF) could be applied, which would linearize the residual and converge in one iteration. However, neglecting $O(\|x\|^2)$ terms in the observation model can lead to inaccurate results. With this in mind, we implement a *recursive estimation* procedure based on an *iterative* EKF. The function we minimize for both *explaining the measurements* and *staying close to the prior* is:

$$h(x) = \frac{1}{2} r^T(y_{k+1}, x) W_{k+1}^{-1} r(y_{k+1}, x) + \frac{1}{2} (x - \hat{x}_{k+1|k})^T P_{k+1|k}^{-1} (x - \hat{x}_{k+1|k}). \quad (29)$$

Due to the nonlinearity of the residual $r(y_{k+1}, x)$ w.r.t. x , the minimization of (29) leads to a *nonlinear least-squares problem*. This problem can be easily handled by modern *trust-region* methods [13], [14]. These methods progressively minimize (29) by iteratively minimizing quadratic models of it built at each step and following a policy of acceptance/rejection of the step based on the agreement between the reduction in the nonlinear function and its approximant.

VI. EXPERIMENTS AND RESULTS

The models and techniques described in the previous sections have been used to reconstruct hand poses in experiments conducted by subjects with an optical tracking system. Although a complete movement can be reconstructed, we focus on “static” grasping poses. To analyze reconstruction outcomes, we consider only 24 out of 26 DoFs of the kinematic model previously described (wrist DoFs are neglected); as a possible application, PCA has been performed and results are compared with the ones presented in [7].

A. Materials and Methods

Experimental task: Subjects were instructed to shape their right hand as to grasp a certain amount of objects ($n = 20$, Table II) which were not physically present during the experiment. Pictures of the objects were instead shown on a

computer screen with a specific timing (3 s with the image displayed, 2 s waiting for a “beep” sound, then perform the grasping and come back before, in 7 s, a new image is shown - total cycle length 12 s), and the subject was asked to grasp them as if they want to use them.

| | |
|-------------------|-----------------------|
| 1. Bucket | 11. Hammer |
| 2. Calculator | 12. Ice cube |
| 3. Chalk | 13. Jar lid |
| 4. Cherry | 14. Light bulb |
| 5. Computer mouse | 15. Pen |
| 6. Dinner plate | 16. Rope |
| 7. Espresso cup | 17. Telephone handset |
| 8. Fishing rod | 18. Tennis racket |
| 9. Frisbee | 19. Toothpick |
| 10. Hair dryer | 20. Wrench |

TABLE II: List of objects used in the task

Each subject performed a total of six trials for each of the objects, preceded by a few training trials; all trials were presented in random order.

Subjects: Four right-handed subjects (two males and two females, age in the interval 20 to 30 years) participated in this study. All subjects gave informed consent, and the protocols were approved by the ethical committee of the University of Pisa.

B. Experimental Procedure and Data Analysis

Hand posture was obtained with the identification procedure described in Section V, measuring the position of optical markers using an optical motion capture system (Phase Space, San Leandro, CA - USA) composed of 10 stereo-cameras at 480 Hz, undersampling in post-processing down to 15 Hz. Problems related to marker occlusions are treated directly via the EKF estimation, which considers only its prior to generate the estimate, achieving satisfying results for short-time missing markers. To the authors’ best knowledge, no means is present to solve the problem of long lasting occlusions in optical tracking systems. The reconstructed postural angles were then smoothed with a 5 samples moving average filter. Subsequently, a selection of the actual frame to consider for a specific grasp has been carried out, based on the timing of object pictures presentation, via visual inspection.

Markers were present on a bracelet fastened to the subject’s forearm, to define the local system of reference; markers were placed on each bone of the hand and on a selected group of joints (Thumb CMC, MPC, IP; Index and Middle MCP; all PIP’s) (see Figure 4).

Data Analysis: The data were then studied using PCA: the percentage of variance accounted for by each principal component (table III) and the cumulative variance (table IV) are lower than the values reported in [7] (where, e.g., cumulative variance explained by the first 2 PCs ranges between 77 and 90%), possibly because we do not reduce intra-object variability via averaging, considering instead all the repetitions for each object grasp in data analysis.

However, the main result is that with only five synergies $\sim 80\%$ of data variance is taken into account for all subjects (except for T.C.), thus suggesting a reduction of the 24 DoFs to be recorded. This increased number of PCs w.r.t. [7] needed to account for 80% of total variance is probably due to the higher number of DoFs. Nonetheless, it is possible that the kinematic model of the thumb, known to have complex non-orthogonal base joints (see [15]), is a major factor: this issue remains however to be studied.

Considering the covariation matrix as reported in Figure 5 (for the sake of space only the covariance matrix for one subject is reported), what is noticeable is that MCP angles of adjacent fingers as well as the PIP angles are highly related to each other, with the extent of correlation decreasing with the separation between pairs of fingers. A similar result is attained also in [7]. In Figure 7 the distribution of the angular differences for all joint angles between hand posture reconstructed from the first two PCs and the actual posture recorded is shown; for a large percentage of poses ($\sim 70\%$) the angular difference is within $\pm 5^\circ$. This can be seen also from Figure 6, where the distribution of normalized amplitudes of the first five PCs are reported. The amplitudes are normalized to the maximum (absolute) value of the first PC. Notice that the amplitude to the third through the fifth PCs are uniformly small, although not as small as in [7] where they almost never (less than 1%) reach a normalized value of 0.5. Finally, in Figure 8, postural synergies defined by the first two principal components in 24 DoFs are reported. The central hand posture is the average over 120 postures (20 different objects 6 times each) for one subject (A.C.) (for the sake of space). The postures to the *right* and *left* are for the maximum and minimum values of the first principal component (*PCI*), while other principal components have

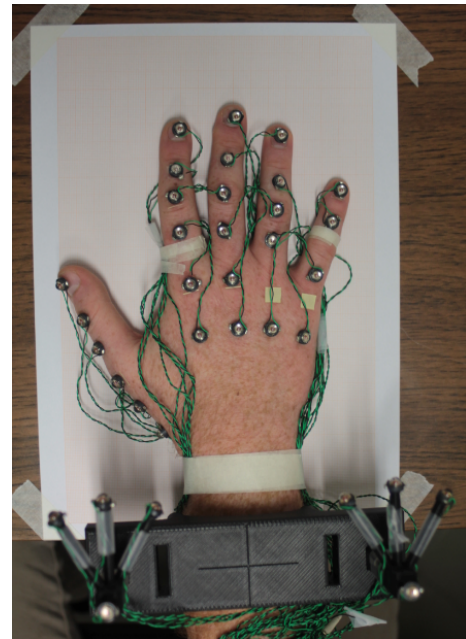


Fig. 4: The hand of one subject after all markers have been placed, showing the protocol for marker placement.

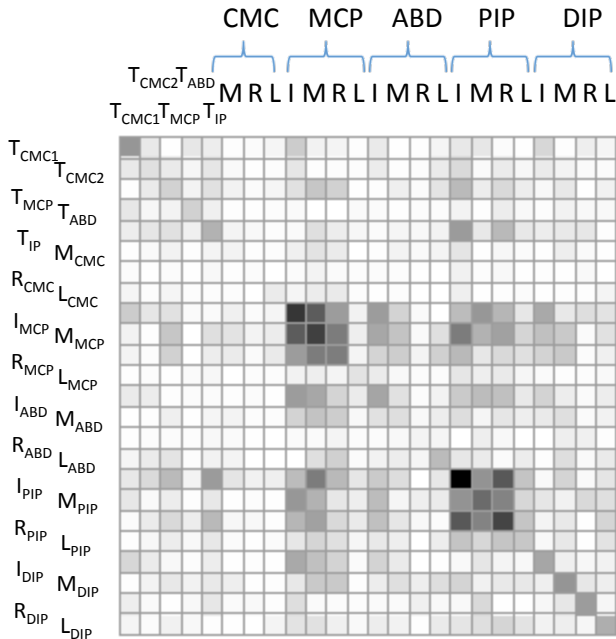


Fig. 5: Gray scale plot of the covariance matrix for one subject (M.B.) analyzing 24 DoFs. For sake of clarity and to enhance result comparison, the same naming in [7] is adopted; the term ABD refers to the adduction/abduction degrees of freedom.

been set to zero. The postures at the *top* and *bottom* are the same for the second principal component (PC_2). Similarly to what has been obtained in [7], along the PC_1 (horizontal) axis, it is possible to observe a gradual closure of the hand: at one extreme, it is noticeable an extension and abduction of the fingers at the MCP joints (PC_1 max), while at the other extreme, MCP joints flexion and adduction is observed (PC_1 min). Along the PC_2 (vertical) axis, we detect a PIP joints flexion and MCP joints extension while moving towards PC_2 min, while the thumb abduction and internal rotation decrease towards PC_2 max. This result suggests that the model of synergy still works in spite of the dimension of the kinematic model in use, with many similarities across models with a different number of DoFs.

Noteworthy, similar results could be seen also if considering a different subset of the total 26 DoFs to perform the PCA, e.g. 15 joints chosen as in [7], where neither palm arch joints nor DIP joints were present. For the sake of space we do not report this other consideration, which anyway leads to very similar conclusions.

| Subjects | PC_1 | PC_2 | PC_3 | PC_4 | PC_5 |
|----------|--------|--------|--------|--------|--------|
| AC | 55.6 | 13.2 | 7.3 | 5.3 | 4.5 |
| MB | 40.0 | 16.1 | 12.3 | 5.9 | 5.2 |
| TC | 27.9 | 17.3 | 12.1 | 7.4 | 6.3 |
| DR | 45.7 | 13.6 | 9.4 | 7.4 | 5.0 |

TABLE III: Percent variance accounted for by each principal component in 24 DoFs

| Subjects | PC_1 | PC_2 | PC_3 | PC_4 | PC_5 |
|----------|--------|--------|--------|--------|--------|
| AC | 55.6 | 68.7 | 76.1 | 81.3 | 85.8 |
| MB | 40.0 | 56.1 | 68.4 | 74.3 | 79.5 |
| TC | 27.9 | 45.3 | 57.4 | 64.8 | 71.1 |
| DR | 45.7 | 59.3 | 68.7 | 76.1 | 81.1 |

TABLE IV: Cumulative percent variance accounted for by 24 DoF PCs

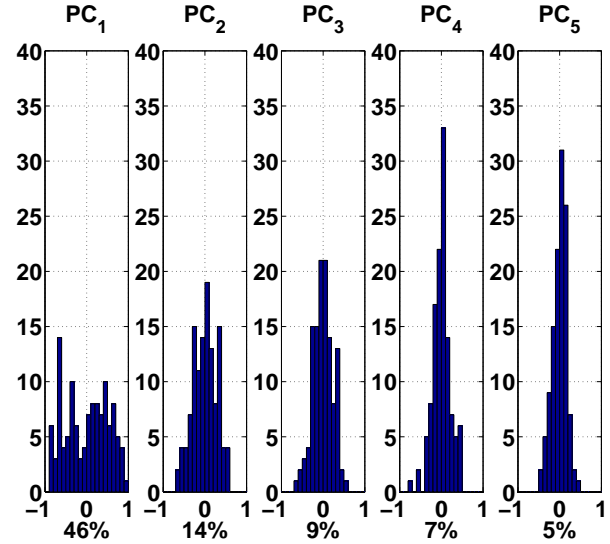


Fig. 6: Distribution of normalized amplitudes of the first five principal components in 24 DoFs. These amplitudes have been normalized to the maximum (absolute) value of the first PC. For the sake of space, the data shown are for one subject (D.R.).

VII. CONCLUSIONS AND FUTURE WORK

In this work a complete procedure to reconstruct human hand posture is presented. The kinematic hand model in use is fully parameterized and allows to take into account differences in subjects hands. Furthermore, the introduction of a soft-tissue artifact compensation mechanism makes the procedure amenable to in-vivo joint angle recordings via optical tracking of markers attached to the skin. The parametrization structure of the model has then been exploited to reconstruct hand poses in experiments, aiming at analyzing postural synergies in grasping, using kinematic models with a different number of degrees of freedom. Results are compared to the ones reported in [7], showing a significant coherence in the synergies obtained, independently from the dimensionality of the kinematic description in use. In conclusion, this paper offers a useful tool that can be used to investigate more in depth and with higher accuracy the synergy concept; the final results, after the definition of a meaningful mapping strategy, might drive the definition of robotic hand postures able to realize effective grasps, or be used as *a priori* information

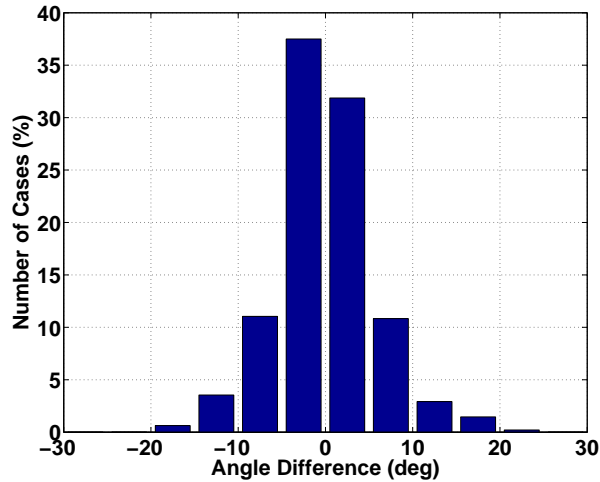


Fig. 7: Distribution of the angular differences for all joint angles between hand postures reconstructed from the first 2 PCs in 24 DoFs and the actual postures recorded. For the sake of space, the data are for all objects from one subject (D.R.).

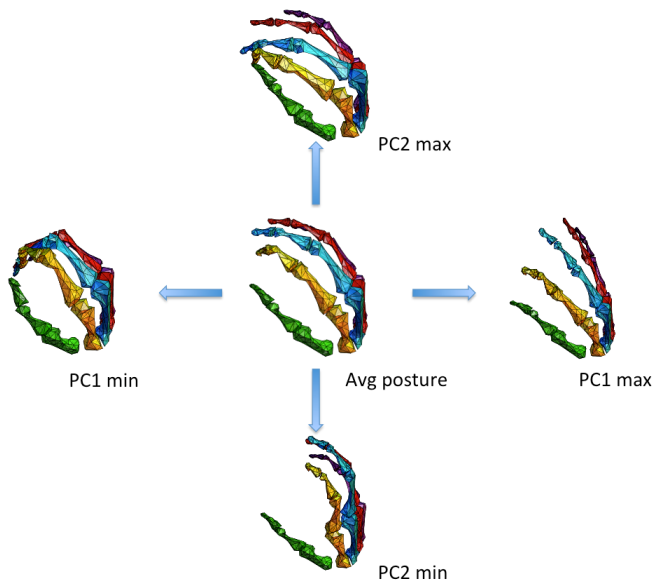


Fig. 8: Postural synergies defined by the first two principal components. The central hand posture is the average over 120 postures (20 different objects 6 times each) for one subject (A.C.). The postures to the *right* and *left* are for the maximum and minimum values of the first principal component (*PC1*), while other principal components have been set to zero. The postures at the *top* and *bottom* are the same for the second principal component (*PC2*).

for hand avatar animation [16] or for the improvement of the design and the performance of glove-based Hand Pose Reconstruction (HPR) systems [17], [18].

REFERENCES

- [1] M. H. Schieber and M. Santello, "Hand function: peripheral and central constraints on performance," *Journal of Applied Physiology*, vol. 96, no. 6, pp. 2293–2300, 2004.
- [2] P. Cerveri, E. De Momi, N. Lopomo, G. Baud-Bovy, R. Barros, and G. Ferrigno, "Finger kinematic modeling and real-time hand motion estimation," *Annals of biomedical engineering*, vol. 35, no. 11, pp. 1989–2002, 2007.
- [3] A. Gustus, G. Stillfried, J. Visser, and H. Jörntell, "Human hand modelling: kinematics, dynamics, applications," *Biological Cybernetics*, vol. 106, pp. 741–755, 2012. [Online]. Available: <http://dx.doi.org/10.1007/s00422-012-0532-4>
- [4] X. Zhang, L. Sang-Wook, and P. Braido, "Determining finger segmental centers of rotation in flexion-extension based on surface marker measurement," *Journal of Biomechanics*, vol. 36, pp. 1097–1102, 2003.
- [5] E. Todorov, "Probabilistic inference of multijoint movements, skeletal parameters and marker attachments from diverse motion capture data," *Biomedical Engineering, IEEE Transactions on*, vol. 54, no. 11, pp. 1927–1939, 2007.
- [6] Q. Fu and M. Santello, "Tracking whole hand kinematics using extended kalman filter," in *Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE*, 2010, pp. 4606–4609.
- [7] M. Santello, M. Flanders, and J. F. Soechting, "Postural hand synergies for tool use," *The Journal of Neuroscience*, vol. 18, no. 23, pp. 10 105–10 115, 1998.
- [8] R. W. Brockett, "Robotic manipulators and the product of exponentials formula," in *Mathematical Theory of Networks and Systems*, P. A. Fuhrman, Ed. New York: Springer-Verlag, 1984.
- [9] R. M. Murray, Z. Li, and S. Sastry, *A Mathematical Introduction to Robotic Manipulation*. Boca Raton, FL: CRC Press, 1994.
- [10] M. Gabbicini, "A twist exponential approach to gear generation with general spatial motions," *Mechanism and Machine Theory*, vol. 44, no. 2, pp. 382–400, 2009.
- [11] J. A. Nelder and R. Mead, "A simplex method for function minimization," *The Computer Journal*, vol. 7, no. 4, pp. 308–313, 1965.
- [12] S. Mehrotra, "On the implementation of a primal-dual interior point method," *SIAM Journal on Optimization*, vol. 2, pp. 575–601, 1992.
- [13] J. Nocedal and S. J. Wright, *Numerical Optimization*, ser. Springer Series in Operations Research and Financial Engineering, T. V. Mikosch, S. I. Resnick, and S. M. Robinson, Eds. New York, NY, USA: Springer, 2006.
- [14] A. R. Conn, N. I. M. Gould, and P. L. Toint, *Trust-region methods*. Philadelphia: SIAM, 2000.
- [15] V. J. Santos and F. J. Valero-Cuevas, "Reported anatomical variability naturally leads to multimodal distributions of denavit-hartenberg parameters for the human thumb," *Biomedical Engineering, IEEE Transactions on*, vol. 53, no. 2, pp. 155–163, 2006.
- [16] S. Mulatto, A. Formaglio, M. Malvezzi, and D. Prattichizzo, "Animating a synergy-based deformable hand avatar for haptic grasping," in *International Conference EuroHaptics*, vol. 2, 2010, pp. 203–210.
- [17] M. Bianchi, P. Salaris, and A. Bicchi, "Synergy-based hand pose sensing: Reconstruction enhancement (in press)," *International Journal of Robotics Research*, 2013.
- [18] —, "Synergy-based hand pose sensing: Optimal design (in press)," *International Journal of Robotics Research*, 2013.