On the Use of Postural Synergies to Improve Human Hand Pose Reconstruction

Matteo Bianchi[◇] [*] Paolo Salaris[◇] [†] Armando Turco[◇] [‡] Nicola Carbonaro[◇] [§] Antonio Bicchi[◇+] [¶]

[◇]Interdept. Research Center "Enrico Piaggio", University of Pisa, via Diotisalvi, 2, 56126, Pisa, Italy.
[+]Department of Advanced Robotics, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy

ABSTRACT

In this paper we consider the problem of estimating the posture of a human hand using sensing gloves, and how to improve their performance by exploiting the knowledge on how humans most frequently use their hands. We consider low-cost gloves providing measurements which are limited under several regards: they are generated through an imperfectly known model, are subject to noise, and are less than the number of degrees of freedom of the hand. Under these conditions, direct reconstruction of the hand pose is an ill-posed problem, and performance is very limited. To obtain an acceptable level of accuracy without modifying the glove hardware, hence basically at no extra cost, we propose to exploit the information on most frequent human hand poses, as represented in a database of postural synergies built beforehand. We discuss how such an *a priori* information can be fused with glove data in a consistent way, so as to provide a good hand pose reconstruction in spite of insufficient and inaccurate sensing data. Simulations and experiments on a low-cost glove are reported which demonstrate the effectiveness of the proposed techniques.

Index Terms: H.5.2 [User Interfaces]: Haptic I/O—Input devices and strategies; I.2.9 [Robotics]: Sensors—; G.1.6 [Optimization]: Constrained optimization—

1 INTRODUCTION

In recent years the issue for a correct hand tracking and pose reconstruction has widely grown together with the diffusion of sensing gloves for gesture measurement [20]. These devices provide useful interfaces for human-machine and haptic interaction in many applications fields like, for example, virtual reality, musical performance, video games, teleoperation and robotics [16]. Different technologies have been exploited to capture fingers movements and positions [9, 22].

At the same time, a widespread commercialization of electronic gloves imposes limits on the production costs in terms of the amount and the quality of the sensors adopted. As a consequence, the correctness of the hand pose reconstruction available by low cost devices themselves might be compromised. On the other hand, several studies and observations on the complex role of human hand in grasping tasks show that it is possible to individuate a reduced number of coordination patterns, called *postural synergies*, which constrains the motion of multiple fingers [19]. From the observability point of view, this suggests that simultaneous movements of fingers follow coordination patterns that reduce the number of independent degrees of freedom to be measured.

In this paper, the information embedded in a known grasps set, which expresses the postural constraints for multi-finger joints, is exploited to reconstruct the hand posture in static grasping tasks, when only a limited and inaccurate number of measures are provided by a low cost sensing glove.

To solve this problem we follow an optimal estimation approach based on Bayes'inference and *a priori* normal data distribution assumption. Two closed-form solutions are presented and tested. The first one, which solves a constrained optimization problem of multinormal probability density function (pdf), is mainly adopted with accurate measured data. The second solution deals with noisy measured data and relies on classic Minimum Variance Estimation (MVE).

In order to validate our reconstruction procedures an optical tracking system has been used to acquire the *a priori* set in addition with a large number of grasp poses. We have performed some simulations where only the measurements of metacarpal joints from the acquisitions have been considered (we refer to metacarpal joint as the mathematical modelling of the rotation degree of freedom (DoF) of metacarpophalangeal joint, MCP). Gaussian noise has been also added and the estimates we get in these cases indicate that, with increasing noise level, MVE method has to be preferred. As a final step we have applied the reconstruction techniques to the five metacarpal measurements of a low cost electronic glove, built for general hand aperture/closure monitoring but not particularly suitable for fine gestural and manipulation recognition. Statistical analyses demonstrate the effectiveness of the proposed techniques w.r.t. the one achievable by simply applying the pseudo-inverse of the measurement matrix.

2 SYNERGIES

The main idea of this work is to exploit the *a priori* information embedded in previously acquired grasps set. Numerous studies have underlined the complex role of human hand in motor organization, with particular attention to grasping, where simultaneous control of many degrees of freedom is requested. It was shown that individuated finger motions was phylogenetically superimposed on basic grasping movements [8]. On this basis, it is possible to individuate a reduced number of coordination patterns (*synergies*) which constrain both joints motion and force exertion of multiple fingers [19]. These constraints may be related to biomechanic factors [3] and synchronization between different motor units [12]. For example, during reaching movements significant covariation schemes at multiple finger joints were observed [4]. Coordination patterns were analyzed by means of multivariate statistical methods, revealing that a limited amount of so-called *eigenpostures* or principal components (PCs) [14], i.e. "statistically identified kinematic coordination patterns" [19], are sufficient to reconstruct a great amount of hand poses. Moreover, a gradient in *eigenpostures* was identified [18]; lower order PCs take into account covariation patterns for MCP and interphalangeal (IP) joints, which are mainly responsible for coarse hand opening and closing, while higher order PCs are used for fine hand shape adjustments. This concept may be interpreted in terms of inner hand representations of increasing complexity, which allow to reduce the number of DoFs according to desired level of approximation. In robotics this idea was then adopted to define simplified manners for the design and control of artificial hands [1, 2]. The dual issue of *controllability*, i.e. *observability*, was considered in [15]. In this case, the whole hand avatar animation problem was studied, when only two haptically-enabled con-

[*]e-mail: m.bianchi@centropiaggio.unipi.it
[†]e-mail: p.salaris@centropiaggio.unipi.it
[‡]e-mail: armando.tur@gmail.com
[§]e-mail: n.carbonaro@centropiaggio.unipi.it
[¶]e-mail: bicchi@centropiaggio.unipi.it

tact points for the thumb and the index were considered, leading to knowledge of 6 DoFs (over the 20 DoFs used to define mesh skeleton) by inverse kinematics algorithm. The key of this approach was to exploit *synergy variables* [17] to approximate principal motion components, thus improving the realism of animation. *Virtual springs*, which couple the avatar tips to the corresponding operator's tips, are exploited to compute the forces necessary (for the avatar fingertips) to track real finger trajectories. These forces are then transformed in the synergy-space general forces and used to recursively estimate the new synergy variables, which characterize the whole-hand configuration. This kind of estimation, which is dependent on the number of synergy variables, was proved to be good enough in qualitative terms but no numerical evaluation was furnished.

In this paper, we address the whole-hand reconstruction issue from a limited amount of measurements following an optimal estimation approach and by taking into account all the information available in the covariance patterns of multi-finger joint positions. This information is summarized with the covariance matrix obtained from the *a priori* grasps set and used for static pose reconstructions. Moreover, our method does not require any interaction with avatars to estimate the hand configuration.

3 THE HAND POSTURE ESTIMATION ALGORITHM

Let us consider a set of measures $y \in \mathbb{R}^m$ given by a sensing glove. By using a n degrees of freedom kinematic hand model, we shall assume a linear relationship between joint variables $x \in \mathbb{R}^n$ and measurements y given by

$$y = Hx + \nu\,, \tag{1}$$

where $H \in \mathbb{R}^{m\times n}$ $(m < n)$ is a full rank matrix which represents the relationship between measures and joint angles, and $v \in \mathbb{R}^m$ is a vector of measurement noise. The goal is to determine the hand posture, i.e. the joint angles x, by using a set of measures y whose number is lower than the number of DoFs describing the kinematic hand model in use. Equation (1) represents a system where there are fewer equations than unknowns and hence leads to an infinite number of least-square solutions. The solution resulting from the pseudo-inverse of matrix H for a such kind of system is a vector $\hat{x} = H^\dagger y$ of minimum Euclidean norm, where y is the vector of noisy measurements. If noise ν is negligible, point $\hat{x}$ is on the hyperplane parallel to the null space and passing through real solution x and such that $\|\hat{x}\|_2$ is minimum. In case of H is a full rank selection matrix, i.e. it consists only of ones and zeros entries with maximum row norm equals to one or equivalently with each sensor measuring exactly one hand kinematic model DoF, the solution is simply given as $\hat{x} = H^T y$. However, depending on both how many joints are measured and the shape of matrix H (e.g. selection matrix), the hand pose reconstruction could be very inaccurate and hence different from the real one. Hereinafter we will refer to the H pseudo-inverse based method as Pinv.

For this reason, our purpose is to improve hand pose reconstruction by using postural synergies information.

All techniques developed in the following sections are based on well established design tools for multinormal data set which usually represents a good approximation and limit to of many other distributions. For this reason, in the following we assume that the *a priori* information we use is multinormally distributed with covariance matrix P_o and mean μ_o. The basic justification, in addition to the procedural advantages in data handling, relies on the central limit theorem [10].

3.1 Pdf Maximization and Mahalanobis Distance

If measurement noise is negligible ($\nu \approx 0$), the best estimation of the hand posture may be obtained by choosing as optimality criterion the maximization of the *probability density function* (pdf) [21] expressed by

$$f(x) = \frac{1}{\sqrt{2\pi\|P_o\|}} \exp\left\{-\frac{1}{2}(x-\mu_o)^T P_o^{-1}(x-\mu_o)\right\}. \tag{2}$$

This is equivalent to solve the following optimal problem:

$$\begin{cases} \hat{x} = \arg\min\limits_{\hat{x}} \frac{1}{2}(x-\mu_o)^T P_o^{-1}(x-\mu_o) \\ \text{Subject to} \quad y = Hx\,. \end{cases} \tag{3}$$

Equation (1) with $\nu = 0$ represents a system where there are fewer equations than unknowns, leading to an infinite number of least-square solutions given by

$$x = H^\dagger y + N_h \xi\,, \tag{4}$$

where $H^\dagger$ is the pseudo-inverse of matrix H, N_h is the null space basis of matrix H and $\xi \in \mathbb{R}^{(n-m)}$ is a free vector of parameters.

As a consequence, the optimal problem defined in (3) becomes

$$\begin{cases} \hat{\xi} = \arg\min\limits_{\hat{\xi}} (H^\dagger y + N_h\xi - \mu_o)^T P_o^{-1}(H^\dagger y + N_h\xi - \mu_o) \\ \text{Subject to} \quad y = Hx\,. \end{cases} \tag{5}$$

By using classical optimization procedures we obtain $\hat{\xi} = (N_h^T P_o^{-1} N_h)^{-1} N_h^T P_o^{-1}(\mu_o - H^\dagger y)$ and hence substituting in (4), after some algebras, the estimation of the hand joint angles is

$$\hat{x} = [I - N_h(N_h^T P_o^{-1} N_h)^{-1} N_h^T P_o^{-1}] H^\dagger y + N_h(N_h^T P_o^{-1} N_h)^{-1} N_h^T P_o^{-1}\mu_o \tag{6}$$

The same result for the optimal problem reported in (3) can be also obtained by using the classic method of Lagrange multipliers. We introduce a new variable $\lambda \in \mathbb{R}^m$ called Lagrange multiplier and study the Lagrange function defined by

$$L = \frac{1}{2}(x-\mu_o)^T P_o^{-1}(x-\mu_o) + \lambda^T(Hx - y)\,, \tag{7}$$

and, by imposing $\frac{\partial L}{\partial x} = \frac{\partial L}{\partial \lambda} = 0$, we have

$$\hat{x} = \mu_o - P_o H^T (H P_o H^T)^{-1}(H\mu_o - y)\,. \tag{8}$$

This solution is equivalent to the one obtained in (6).

If the measurement matrix H is a selection matrix, it is possible to easily maximize $E[x|y]$ in terms of multinormal conditional distribution [10]. Indeed, in this case vector y defines a precise subset of the state variables, being X_1, whose values are known by means the measurement process, while X_2 indicates the rest of state variables to be estimated. This definition allows to partition the *a priori* covariance matrix as

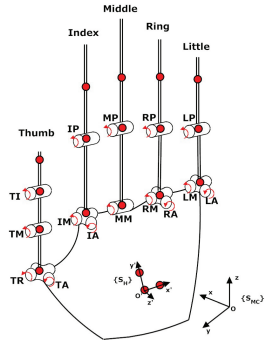
$$\left(\frac{X_1}{X_2}\right) \Longrightarrow P_o = \left(\frac{P_{o11}|P_{o12}}{P_{o21}|P_{o22}}\right) \tag{9}$$

as well as the *a priori* mean $\mu_o = (\mu_{o1}|\mu_{o2})$. The estimation of X_2 is easily derived as

$$\hat{X}_2 = E[X_2|X_1 = y] = \mu_{o2} + P_{o21}P_{o11}^{-1}(y - \mu_{o1})\,. \tag{10}$$

It is interesting to give a geometrical interpretation of the cost function in (3), which expresses the squared of the Mahalanobis distance [13]. The concept of Mahalanobis distance, which takes into account data covariance structure, is widely exploited in statistics, e.g. in PCs Analysis, mainly for outliers detection [11]. According to it, to assess if a test point belongs to a known data set, whose distribution defines an hyper-ellipsoid, we have to take into account both its closeness to the centroid of data set and the direction of the test point w.r.t. the centroid itself. In other words, more samples are distributed along this direction, more probably the test point likely belongs to data set even if it is further from the center.

Hereinafter, we will refer to the aforementioned constrained p.d.f. maximization as Mp algorithm.



DoFs	Description
TA	Thumb Abduction
TR	Thumb Rotation
TM	Thumb Metacarpal
TI	Thumb Interphalangeal
IA	Index Abduction
IM	Index Metacarpal
IP	Index Proximal
MM	Middle Metacarpal
MP	Middle Proximal
RA	Ring Abduction
RM	Ring Metacarpal
RP	Ring Proximal
LA	Little abduction
LM	Little Metacarpal
LP	Little Proximal

Figure 1: Kinematic model of the hand with 15 DoFs. Markers are reported as red spheres.

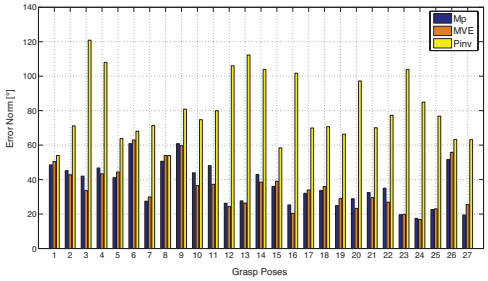


Figure 2: Euclidean error norm [°] in pose reconstruction for Mp, MVE and pseudo-inverse based algorithm, for 27 random elements from the grasps set. Performance is evaluated with 10% noise amplitude.

3.2 MVE and MAP

Let us assume now that sensors used in a sensorized glove are affected by noise ν. If sensor noise is relevant, measures coming from the glove are not much reliable and the hand posture estimation obtained by previous algorithms might be very different from the real one. For this reason, in this section we propose an other algorithm in case of noise: the Minimum Variance Estimation (MVE) technique. This method minimizes a cost functional which expresses the weighted Euclidean norm of deviations, i.e. cost functional $J = \int_X (\hat{x}-x)^T S(\hat{x}-x)dx$, where S is an arbitrary, positive-semidefinite matrix.

Under the hypothesis that ν has zero mean and gaussian distribution with covariance matrix R, we get the solution for the minimization of J as $\hat{x} = E[x|y]$, where $E[x|y]$ represents the *a posteriori* pdf expectation value. The estimation $\hat{x}$ may be achieved by easily generalizing the expression in [7] to nonzero mean *a priori* distributions as

$$\hat{x} = (P_o^{-1} + H^T R^{-1} H)^{-1}(H^T R^{-1} y + P_o^{-1}\mu_o), \qquad (11)$$

where matrix $P_p = (P_o^{-1} + H^T R^{-1} H)^{-1}$ is the *a posteriori* covariance matrix, which has to be minimized to increase information about the system. This result represents a very common procedure in applied optimal estimation, where usually there is redundant sensor information. Here we adopt it for an under-determined problem (i.e. $m < n$), with nonzero mean *a priori* distribution. When covariance matrix R tends to assume very small values, the solution described in equation (11) might encounter numerical problems and hence it can not be exploited. For this reason, in these cases it is preferable to use the estimation formulas given by (6), (8) or (10).

4 Model and Data Capture

Without loss of generality, a 15 DoFs model is adopted for hand pose reconstruction, which is the same used in [18, 6] and reported

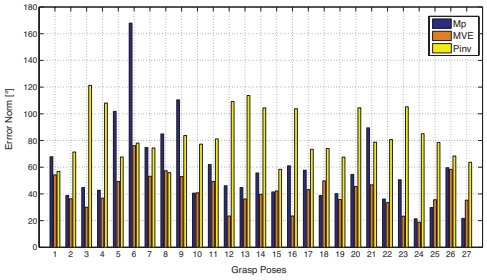


Figure 3: Euclidean error norm [°] in pose reconstruction for Mp, MVE and pseudo-inverse based algorithm, for 27 random elements from the grasps set. Performance is evaluated with 45% noise amplitude.

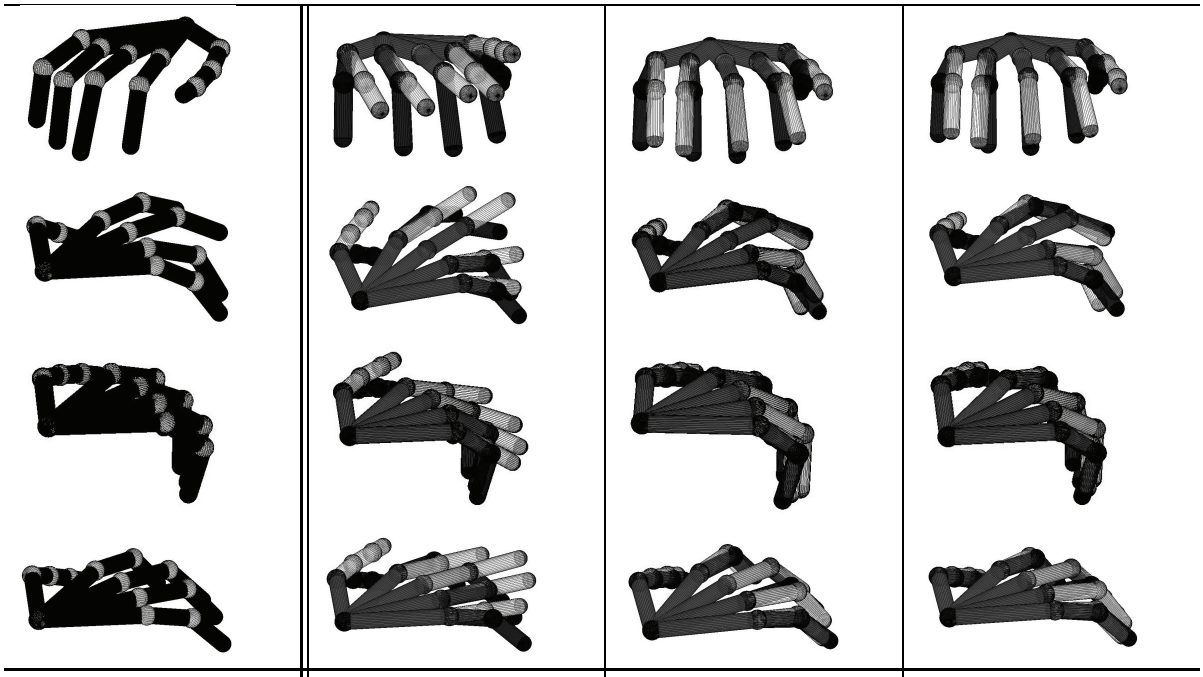
in figure 1. An optical motion capture system (Phase Space, San Leandro, CA - USA) with 19 active markers is used to collect a large number of static grasp positions. Subject AT (M,26) has performed all the grasps of the 57 imagined objects described in [18]; these data have been acquired twice to define a set of 114 *a priori* data. LC (M,26) has executed 54 grasp poses of a wide range of different imagined objects, which have been recorded in parallel with the sensing glove and the Phase Space system. In this manner both a glove calibration and reliable reference values for whole hand positions have been achieved. Indeed, given the high accuracy provided by the optical system to detect markers (the amount of static marker jitter is inferior than 0.5 mm, usually 0.1 mm) and assuming a linear correlation (due to skin stretch) between marker motion around the axes of rotation of the joint and the movement of the joint itself [23], we can consider the processed hand poses acquired with Phase Space as a good approximation of real hand positions. This assumption is still valid when the sensing glove is worn because it perfectly adapts to hand shape. None of the subjects had physical limitations that would affect the experimental outcomes. Data collection from subjects in this study was approved by the University of Pisa Institutional Review Board.

The disposition of the markers on the glove refers to [5] and it is reported in figure 1 and 6. We have used four markers for the thumb and three markers for each of the rest of the fingers. Three markers have been also placed on the dorsal surface of the palm to define a local reference system S_H. The positions of the markers, which have been sampled at 480 Hz, are given referring to the global reference system S_{MC} (which is directly defined during the calibration of the acquisition system).

The computation of joint angles w.r.t. S_H has been executed using the *ikine* function of Robotics Toolbox, which implements an iterative algorithm of kinematic inversion suitably modified by adapting computational tolerance to guarantee numerical convergence. Data have been suitably pre-filtered with a moving average filter to enhance Signal Noise Ratio (SNR). This application needs a preliminary phase, where the hand is posed in a reference position and fingers flexion-extension is nearly zero. During this phase phalanges length and eventual offset angles have been calculated.

5 Simulation Results

In simulations, only metacarpal joint measurements from Phase Space are used with reconstruction algorithms. Estimation results are compared with corresponding reference values. An additional normal noise with different amplitudes (1%, 10% and 45% measures amplitude, respectively) is also added, and algorithms performance is evaluated. These noise levels have been chosen in order to offer a reasonable workspace to test the robustness of the algorithm; indeed 10% amplitude represents a common value for sensors while 1% and 45% identify, respectively, a reasonable upper-quality case and worst-quality case. Two different types of analysis have been conducted. The first one evaluates the estimation performance for the three techniques over all the poses in terms of absolute esti-



Real hand posture (from optical position tracking system).	Hand reconstruction by Pinv estimation algorithm.	Hand reconstruction by MVE estimation algorithm.	Hand reconstruction by Mp estimation algorithm.

Figure 4: Hand pose reconstructions obtained with MVE, Mp algorithms and by pseudo-inverse of selection matrix H. Matrix H allows to measure TM, IM, MM, RM and LM (see figure 1). Measures are affected by 1% of noise. In black the "real" hand posture whereas in white the estimated one.

	Noise Level(%)			Max Error [°]		
	1%	**10%**	**45%**	**1%**	**10%**	**45%**
Mp	6.74±2.40	7.25±2.51	10.11±4.70	13.50	13.58	26.08
MVE	6.74±2.41	7.23±2.56	8.44±2.77	13.51	13.76	15.95
Pinv	13.93±3.09	14.36±3.05	15.87±3.23	20.85	21.56	23.75

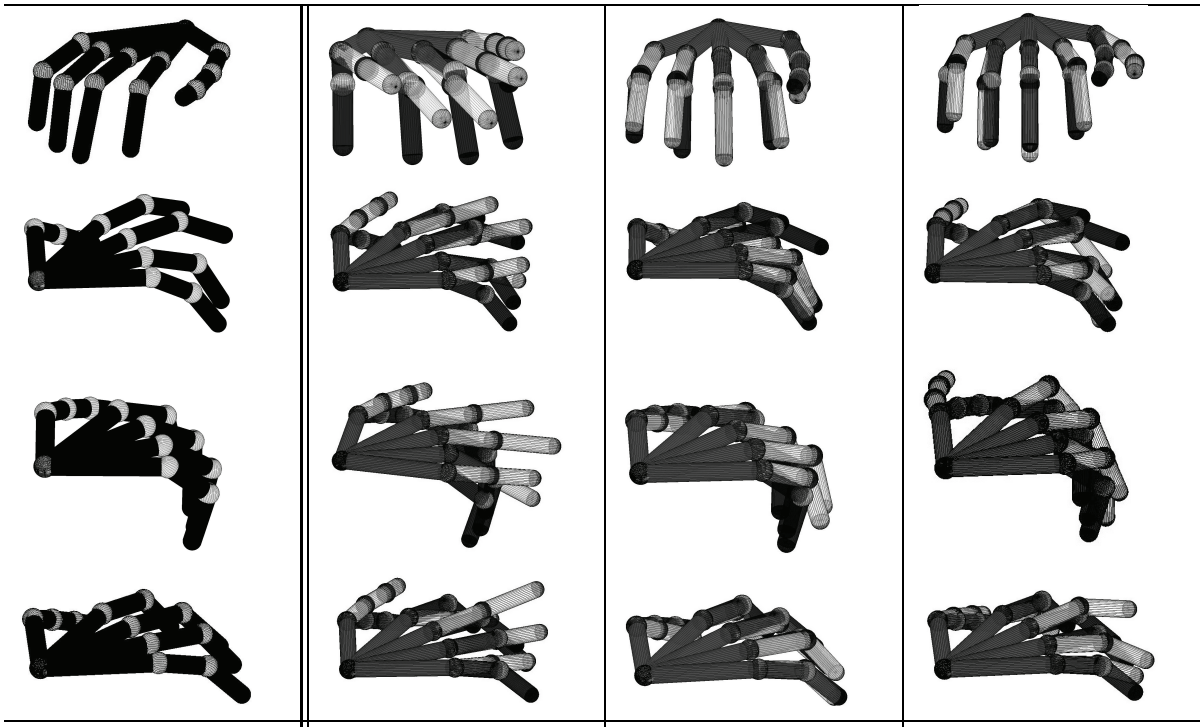
Table 1: Mean pose estimation errors and standard deviations ([°]) for the three methods with different noise levels. Maximum errors are also reported.

mation error for each pose (i.e. the average of the DoF absolute estimation errors). The second analysis is focused on the estimation accuracy on each DoF. In table 1 the mean estimation errors over all the poses with the corresponding standard deviations and maximum errors are reported.

In order to demonstrate the effectiveness of our methods, we have tested for statistical differences between estimation poses and joint errors obtained with above described techniques. For this purpose, various types of tools have been used. More specifically, if normality and homogeneity of variances assumption on samples are verified (through Lilliefors' composite goodness-of-fit test and Levene's test, respectively), a classic two-tailed t-test (hereinafter referred as T_{eq}) has been used. If the variances assumption is not met, a modified two-tailed T-test has been exploited instead (Behrens-Fisher problem, using Satterthwaite's approximation for the effective degrees of freedom, hereinafter referred as T_{neq}). When the assumption of normality is not verified, a non parametric test has been adopted for the comparison (Mann-Whitney U-test, hereinafter referred as U). Significance level of 5% has been considered whereas p-values less than 10^{-4} are assumed equal to zero.

Referring to pose estimation errors, results obtained show no statistical differences between Mp and MVE for noise levels of 1% ($p = 0.99$, T_{eq}), 10% ($p = 0.97$, T_{eq}) and 45% ($p = 0.09$, U). In the latter case the significance is strongly related to the randomness of the noise, while for low levels of noise p-values for the comparison between Mp and MVE methods are much larger than the significance level. On the contrary, when noise increases Mp performance appears to be degraded w.r.t. MVE, as clearly reported in figures 4 and 5, where same reconstructed poses are displayed in comparison with the reference values, considering 1% and 45% noise respectively. Under a qualitative point of view, in case of low level noise, Mp and MVE reconstructions are quite identical and hardly distinguishable. Estimated joint angles and finger positions are very close to the reference ones. When noise is at 45% level, Mp algorithm reconstructions are worst than those obtained with MVE algorithm. Nevertheless, Mp algorithm maintains the likelihood with common grasp poses because of the *a priori* information unlike Pinv method. Same trends are observable also in figures 3 and 2, where Euclidean reconstruction error norm is plotted for 10% and 45% noise levels, respectively. In all cases, Pinv performance in pose reconstruction is significantly different from Mp and MVE (at 1% and 10% noise: $p = 0$, T_{eq}, for Mp vs. Pinv and MVE vs. Pinv; at 45% noise: $p = 0$, T_{eq}, for Mp vs. Pinv and $p = 0$, U, for MVE vs. Pinv). Pinv method gives the worst average results as appeared in table 1. Maximum errors for pose reconstruction show that MVE method offers the best performance for all the levels of noise. For 45% level Mp method produces the largest maximum error, because this technique is the most sensitive to noise effects (see figure 3).

Since the average pose errors analysis gives only a global performance description, a second type of analysis w.r.t. the average absolute estimation error for each DoF over the poses has been conducted. In tables 2, 3 and 4 the average values of each DoF absolute estimation error with their corresponding standard deviations are reported, with level of noise of 1%, 10% and 45% respectively. The



Real hand posture (from optical position tracking system).	Hand reconstruction by Pinv estimation algorithm.	Hand reconstruction with MVE estimation algorithm.	Hand reconstruction with Mp estimation algorithm.

Figure 5: Hand pose reconstructions obtained with MVE, Mp algorithms and by pseudo-inverse of selection matrix H. Matrix H allows to measure TM, IM, MM, RM and LM (see figure 1). Measures are affected by 45% of noise. In black the "real" hand posture whereas in white the estimated one.

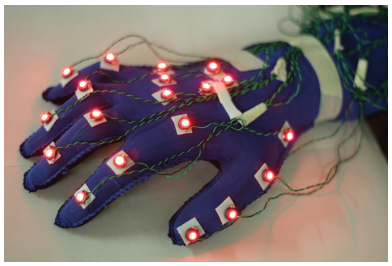


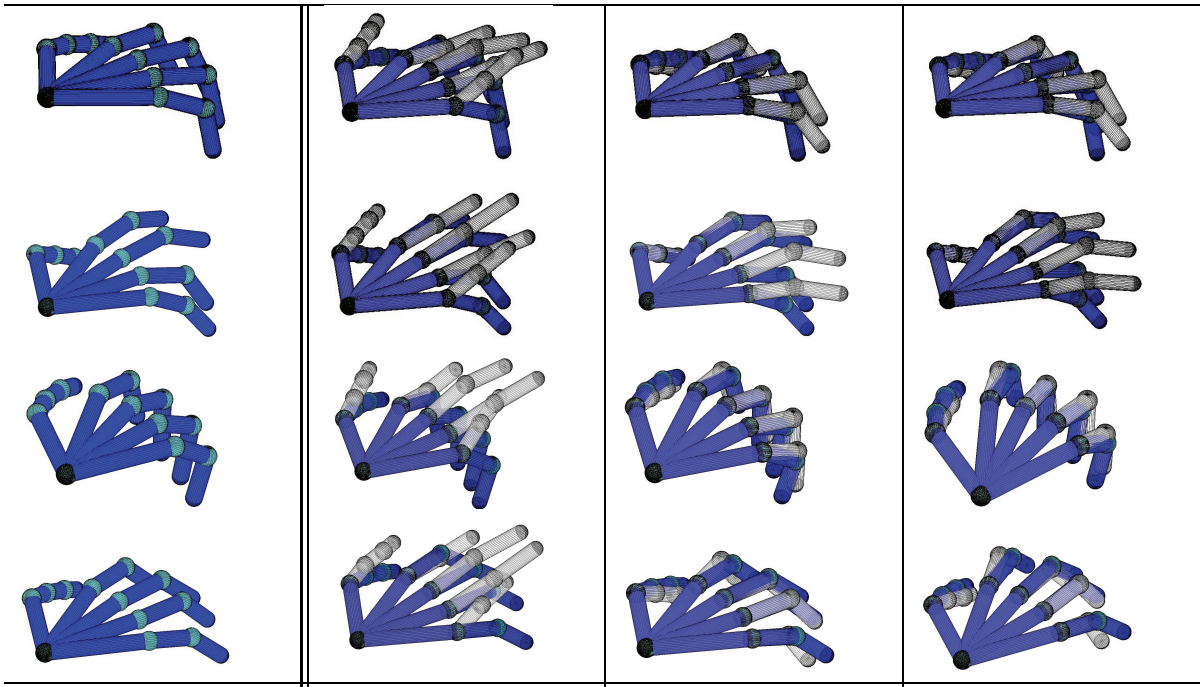
Figure 6: The sensing glove with added markers.

DoF	Noise 1%			Max Error [o]		
	Mp	**MVE**	**Pinv**	**Mp**	**MVE**	**Pinv**
TA	10.71±8.52	10.58±8.42	14.04±11.10	31.73	31.35	32.74
TR	7.17±4.54	7.15±4.54	27.62±10.24	19.40	19.45	45.65
TM∗	0.07±0.06	0.07±0.06	0.07±0.06	0.20	0.20	0.20
TI	4.86±3.68	4.84±3.68	6.74±5.54	19.55	19.52	23.16
IA	11.96±5.31	11.94±5.32	6.27±3.27	26.16	26.16	14.90
IM∗	0.17±0.24	0.17±0.24	0.17±0.24	1.22	1.26	1.22
IP	13.29±7.10	13.32±7.10	28.87±13.79	28.19	28.18	59.41
MM∗	0.14±0.14	0.15±0.13	0.14±0.14	0.61	0.57	0.61
MP	12.36±7.77	12.39±7.77	29.84±13.64	30.77	30.82	57.78
RA	3.44±2.44	3.43±2.43	10.17±3.78	9.56	9.54	16.45
RM∗	0.19±0.18	0.22±0.17	0.19±0.18	0.61	0.69	0.61
RP	13.41± 9.67	13.46±9.69	34.00±13.88	39.61	39.73	65.43
LA	11.33±5.97	11.26±5.85	24.28±5.18	24.6	24.46	37.89
LM∗	0.14±0.15	0.13±0.14	0.12±0.15	0.80	0.76	0.80
LP	11.94± 9.55	11.98±9.55	26.50±13.65	36.74	37.85	63.64

∗ indicates a measured DoF.

Table 2: Average estimation errors and standard deviations for each DoF [°] with 1% noise level. Mp, MVE and Pinv methods are considered. Maximum errors are also reported.

three reconstruction methods are considered and maximum errors are also indicated. What is noticeable is that for 1% level Mp and MVE methods furnish the best results for all the DoFs. As it can be seen from table 5, no statistical difference is observed between Mp and MVE. Moreover, Mp, MVE and Pinv show no differences in TA, TI and TM estimations, as well as for the measured DoFs. This may be due to the fact that the thumb phalanges are the hardest to be modeled under a kinematic point of view. Same behavior between Mp and Pinv is observed for all the considered noise levels, while MVE and Pinv exhibit no differences for TM, TI and LM DoFs at 10% noise levels and TI, IM, MM, RM and LM DoFs at 45% level. When noise increases, Mp and MVE method differentiate each other in terms of the estimation errors for the measured DoFs and TA; at 45% level they present significant different reconstruction performance for TA, TM and LA. From this analysis, Mp and MVE methods exhibit a performance which is better or comparable with the one provided by Pinv, except for the IA DoF which is a not measured DoF. Maximum errors for the three methods are observed for RP and LP DoFs. However, for a proper evaluation they should be weighted w.r.t. the maximum variation produced by each DoF. Referring to Mp and MVE, a possible explanation for



Real hand posture (from optical position tracking system).	Hand reconstruction by Pinv estimation algorithm.	Hand reconstruction by MVE estimation algorithm.	Hand reconstruction by Mp estimation algorithm.

Figure 7: Hand pose reconstructions with MVE, Mp and Pinv algorithms, with measures given by sensing glove. In blue the "real" hand posture whereas in white the estimated one.

DoF	Noise 10%			Max Error [°]		
	Mp	**MVE**	**Pinv**	**Mp**	**MVE**	**Pinv**
TA	11.22±8.68	6.50±5.59	14.04±11.10	35	19.41	32.74
TR	7.06±4.74	6.96±5.11	27.62±10.24	20.85	21.54	45.65
TM∗	0.70±0.61	0.70±0.55	0.70±0.60	2.48	2.66	2.58
TI	4.91±3.61	4.83±3.52	6.74±5.54	19.80	19.20	23.16
IA	11.95±5.24	11.34±5.33	6.27±3.27	26.55	26.36	14.90
IM∗	1.66±2.31	2.08±1.90	1.66±2.32	12.68	9.43	12.68
IP	13.49±7.10	14.34±7.19	28.87±13.79	27.55	28.23	59.41
MM∗	1.55±1.77	2.67±1.49	1.55±1.77	7.82	6.54	7.82
MP	12.38±7.54	13.20±7.36	29.84±13.64	31.04	31.74	57.78
RA	3.53±2.48	3.25±2.29	10.17±3.78	9.59	8.99	16.45
RM∗	1.64±1.77	3.12±1.79	1.64±1.77	10.43	8.13	10.43
RP	13.80±9.60	15.41±10.00	34.00±13.88	43.70	42.89	65.43
LA	11.18±6.16	9.07±5.29	24.28±5.18	24.28	19.39	37.89
LM∗	1.53±1.61	1.44±1.53	1.53±1.61	7.07	7.49	7.07
LP	12.13±9.62	13.59±9.24	26.50±13.65	38.64	37.24	63.64

∗ indicates a measured DoF.

Table 3: Average estimation errors and standard deviations for each DoF [°] with 10% noise level. Mp, MVE and Pinv methods are considered. Maximum errors are also reported.

DoF	Noise 45%			Max Error [°]		
	Mp	**MVE**	**Pinv**	**Mp**	**MVE**	**Pinv**
TA	18.05±13.73	7.46±4.73	14.04±11.10	64.10	20.79	32.74
TR	7.66±5.76	7.31±5.70	27.62±10.24	27.57	26.04	45.66
TM∗	3.16±2.60	2.16±1.85	3.16±2.60	12.17	7.96	12.17
TI	6.60±5.07	5.50±4.33	6.74±5.54	20.79	18.17	23.16
IA	11.73±6.16	10.94±5.42	6.27±3.27	30.61	26.81	14.90
IM∗	7.71±8.75	6.33±4.96	7.71±8.75	41.46	21.98	41.46
IP	13.87±10.52	15.05±7.80	28.87±13.79	45.19	32.89	59.41
MM∗	5.83±5.56	5.30±3.42	5.83±5.56	26.72	18.33	26.72
MP	14.43±12.15	13.96±8.29	29.84±13.64	52.75	38.31	57.78
RA	3.84±2.59	3.10±2.28	10.17±3.78	10.55	8.91	16.45
RM∗	6.72±7.84	5.48±4.07	6.72±7.84	35.53	15.33	35.53
RP	18.38±18.79	17.11±11.38	34.00±13.88	85.71	51.62	65.43
LA	10.53±6.60	7.85±5.16	24.28±5.18	26.68	21.22	37.89
LM∗	6.26±6.34	4.91±4.67	6.26±6.34	27.35	20.36	27.35
LP	16.72±17.77	15.09±9.94	26.50±13.65	72.30	49.89	63.64

∗ indicates a measured DoF.

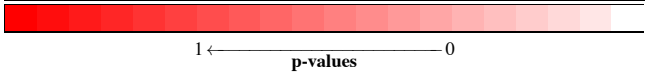
Table 4: Average estimation errors and standard deviations for each DoF [°] with 45% noise level. Mp, MVE and Pinv methods are considered. Maximum errors are also reported.

the errors occurring at RP and LP may be related to the fact that the ring-proximal and little-proximal estimated positions tend to be more *closed* in common grasping tasks, as it is contained in the *a priori* set. Notice the growth of the maximum error for Mp with noise level. On the contrary, Pinv method errors are always the same except for those DoFs referring to measures. This fact may be easily derived from the pseudo-inverse definition.

6 Experimental Results

The sensorized glove used in this study is manufactured by printing a strip of a Conductive Elastomer (CE) following the contour of the hand on a Lycra®/cotton fabric, see figure 6. Additional elements of the conductive elastomer are also printed on the dorsal side of the glove to create the connection to the 20 different sensor segments of the polymeric strip [22].

DoF	p Mp vs. MVE			p Mp vs. Pinv			p MVE vs. Pinv		
	N-1%	**N-10%**	**N-45%**	**N-1%**	**N-10%**	**N-45%**	**N-1%**	**N-10%**	**N-45%**
TA	**0.88**	0.0025	0‡	**0.16**	**0.36**	**0.15**	**0.14**	0.0004	0.0173
TR	**0.99**◇	**0.92**◇	**0.57**	0	0	0	0	0	0
TM∗	**0.97**	**0.66**	0.0332	**1**	**1**	**1**	**0.97**	**0.66**	0.0332
TI	**0.96**	**0.91**◇	**0.17**	**0.14**	**0.18**	**0.8**	**0.14**	**0.11**	**0.23**
IA	**0.82**	**0.31**	**0.54**	0	0	0	0	0	0
IM∗	**0.78**	0.0159	**0.27**	**1**	**1**	**1**	**0.78**	0.0159	**0.27**
IP	**0.98**◇	**0.54**◇	**0.15**	0	0	0	0	0	0
MM∗	**0.43**	0	**0.62**	**1**	**1**	**1**	**0.43**	0	**0.62**
MP	**0.98**◇	**0.57**◇	**0.64**	0	0	0	0	0	0
RA	**0.94**	**0.44**	**0.1**	0	0	0	0‡	0	0
RM∗	**0.23**	0	**0.61**	**1**	**1**	**1**	**0.23**	0	**0.61**
RP	**0.93**	**0.34**	**0.34**	0	0	0	0	0	0
LA	**0.95**‡	**0.07**	0.0207◇	0◇	0◇	0◇	0◇	0	0◇
LM∗	**0.78**	**0.93**	**0.44**	**1**	**1**	**1**	**0.78**	**0.93**	**0.44**
LP	**0.90**	**0.25**	**0.26**	0	0	0.0001	0	0	0‡



∗ indicates a measured DoF.

Table 5: p-values from the evaluation of DoF estimation errors between Mp and MVE, Mp and Pinv, MVE and Pinv. 1%, 10% and 45% noise levels have been considered. ◇ indicates T_{eq} test. ‡ indicates T_{neq} test. When no symbol appears near the tabulated values, U test is used. **Bold** value indicates no statistical difference between the two methods under analysis at 5% significance level. When the difference is significative, values are reported with a 10^{-4} precision. p-values less than 10^{-4} are considered equal to zero.

The CE materials show piezoresistive properties. So, as the hand moves, the sensor elements corresponding to different segments of the contour of the hand are subject to changes in length, thus causing changes in the electrical properties of the material. Such changes can be detected by reading the voltage drop across such segments. Sensors and connections made by means of the same material avoided the use of obtrusive metallic wires, which may interfere with hand movements. Moreover this solution allows to maintain the textile original elasticity placing electrical contacts in the periphery of the garment, resulting in an advantage in terms of user comfort.

The sensors are connected in series thus forming a single sensor line while the connections intersect the sensor line in the appropriate points. An *ad hoc* electronic front-end was designed to compensate the resistance variation of the connections, made by the same material of the sensors, using an high input impedance stage.

In the present study, metacarpophalangeal (MCP) flexion-extension measures are provided by the glove.

6.1 Results and Discussion

In the calibration phase, we firstly get $\hat{H}$, which is an estimation of the measurement matrix H. For this purpose, number of poses N to be collected in parallel with the glove and the position optical tracking system has to be larger or equal than the dimension of the state to estimate, i.e. $N \geq 15$. The reference poses have been collected in matrix $X_c \in \mathbb{R}^{15\times 15}$.

At the same time measures of the glove have been averaged over the last 50 acquired samples and organized in matrix $Z_c \in \mathbb{R}^{5\times 15}$.

Given the relation $Z_c = \hat{H}X_c$, matrix $\hat{H}$ is obtained as

$$\hat{H} = Z_c((X_c^T)^\dagger)^T . \tag{12}$$

Measurement noise is calculated in terms of fluctuations w.r.t. the aforementioned average values of the measures and noise covariance matrix R is obtained. In this case noise is less than 10%

	Mean ± std	**Max Error** [°]
Mp	11.17±4.32	25.70
MVE	10.94±4.24	25.18
Pinv	19.00±3.66	30.30

Table 6: Mean pose estimation errors and standard deviations ([°]) for the three methods with different noise levels. Maximum errors are also reported.

DoF	Mean±Std			Max Error		
	Mp	**MVE**	**Pinv**	**Mp**	**MVE**	**Pinv**
TA	12.12±9.98	12.10±9.88	14.37±10.78	36.95	36.63	34.28
TR	9.20±7.13	9.20±7.04	26.46±10.49	26.61	26.34	46.43
TM∗	4.36±3.73	4.32±3.71	6.43±4.44	13.34	13.25	18.50
TI	14.56±9.96	14.46±9.86	7.84±5.47	33.61	33.25	22.38
IA	9.82±6.89	9.91±6.81	7.10±5.08	30.01	29.60	21.18
IM∗	15.27±11.86	15.06±11.76	16.48±12.62	47.92	46.76	43.58
IP	9.60±7.65	9.67±7.62	31.47±14.70	27.28	27.40	61.11
MM∗	14.40±12.84	14.01±12.54	19.88±14.58	55.18	53.03	51.47
MP	6.80±6.49	6.92±6.58	24.36±9.85	24.49	24.74	43.72
RA	6.20±4.31	5.94±4.11	5.69±4.72	16.31	15.72	20.90
RM∗	19.00±13.44	18.21±12.95	19.22±11.81	64.80	61.98	46.32
RP	8.98±8.91	9.11±9.03	31.51±13.98	31.66	32.24	60.62
LA	11.42±8.50	10.72±8.03	32.24±6.98	31.56	29.59	48.11
LM∗	17.37±12.51	16.03±11.79	17.98±11.81	62.44	58.40	45.05
LP	8.43±6.36	8.48±6.33	23.90±12.53	25.74	26.07	56.21

∗ indicates a measured DoF.

Table 7: Average estimation errors and standard deviations for each DoF [°], for the sensing glove acquisitions. Mp, MVE and Pinv methods are considered. Maximum errors are also reported.

but the errors in the measurement matrix estimation are consistent due to intrinsic non-linearities and hysteresis of sensing glove.

In table 6 the absolute average pose estimation errors are reported with the corresponding standard deviations. Same trends already observed for simulation results can be noticed here. There is no statistical difference between Mp and MVE ($p = 0.78$, T_{eq}) and significant difference between Mp and Pinv and MVE and Pinv ($p = 0$, T_{eq}). Mp and MVE exhibit best pose reconstruction performance, also in terms of maximum errors. In table 7 the average reconstruction errors for each DoF are reported. No statistical difference is present between Mp and MVE performance as appeared from table 8. Mp and MVE methods exhibit results statistically different w.r.t. Pinv method except for IM, RM and LM DoFs, which are provided by measures, RA DoF, for which the average estimation error is limited ($\approx 6°$), and finally TA. Notice that the best TI average estimation is achieved with Pinv; this result, together with the aforementioned observation about TA, may be explained in terms of the difficulties in modelling thumb kinematics (as previously observed in Section 4). Also IA DoF presents the smallest average estimation error using Pinv algorithm, even if in this case p-values resulting from the comparisons between the three techniques are close to the significance value and mean error values comparable. Looking at maximum DoF reconstruction errors, we can see that they occur at those measured DoFs which present the maximum variation during grasping tasks. This is especially noticeable for Mp and MVE methods and it may be explained in terms of an inaccurate estimation of the measurement matrix due to non-linearities in glove behavior. However, except for some poses, both Mp and MVE procedures globally exhibit best performance and a good robustness to modelling errors. It is important to notice that the main goal achieved by Mp and MVE algorithms is the *likelihood* with real poses which is always preserved as reported in figure 7. Same fact can not be obtained with pseudo-inverse technique. Euclidean norm errors between the real hand poses and the reconstructed ones for the above described techniques are shown in figure 8.

DoF	p Mp vs. MVE	p Mp vs. Pinv	p MVE vs. Pinv
TA	**0.98**	**0.27**	**0.28**
TR	**0.99**	0	0
TM∗	**0.90**	0.0107	0.0093
TI	**0.93**	0.0008	0.0008
IA	**0.95**◇	0.0448	0.0381
IM∗	**0.92**	**0.63**	**0.58**
IP	**0.87**	0	0
MM∗	**0.85**	0.0348	0.0232
MP	**0.88**	0	0
RA	**0.76**	**0.37**	**0.51**
RM∗	**0.76**◇	**0.93**	**0.67**
RP	**0.90**	0	0
LA	**0.66**	0	0
LM∗	**0.48**	**0.60**	**0.26**
LP	**0.99**	0	0

1 ← **p-values** → 0

∗ indicates a measured DoF.

Table 8: p-values from the evaluation of DoF estimation errors between Mp and MVE, Mp and Pinv, MVE and Pinv. 1%, 10% and 45% noise levels have been considered. ◇ indicates T_{eq} test. ‡ indicates T_{neq} test. When no symbol appears near the tabulated values, U test is used. **Bold** value indicates no statistical difference between the two methods under analysis at 5% significance level. When the difference is significative, values are reported with a 10^{-4} precision. p-values less than 10^{-4} are considered equal to zero.

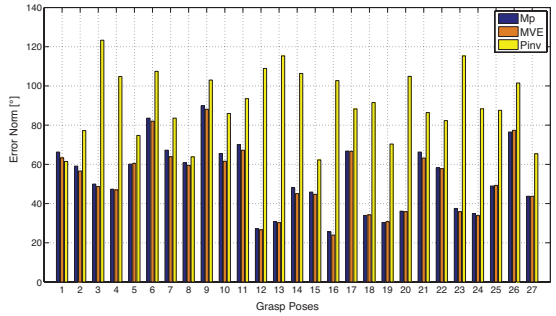


Figure 8: Euclidean norm error [°] for 27 random poses acquired with the glove. All three techniques are considered.

7 Conclusions and Future Work

In this work reconstruction techniques to estimate static hand poses from a reduced number of measures given by an input glove-based devices are presented. These techniques are based on classic optimization and applied optimal estimation methods. The main innovation relies on the exploitation of the *a priori* information embedded in the covariance structure of a grasps set. This covariance individuates some coordination patterns, defined as *postural synergies*, which reduce hand DoFs to be measured and controlled.

Simulation results, where noise effects are also considered, and experiments with a low-cost sensing glove are reported. Performance is compared with the one obtained with a simple pseudo-inverse based algorithm. Statistical analyses demonstrate the effectiveness of the here proposed hand pose reconstructions.

The achieved results may be useful to improve a large class of human-interfaces in many application fields, e.g. video-games or tele-robotics, where fine hand position individuation and low cost devices are crucial features to allow a reliable haptic experience.

Future work aims at individuate both optimal criteria for sensors design and optimal procedures for a good calibration of the sensorized glove arrangement on the basis of the *a priori* information.

Acknowledgements

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. Authors gratefully acknowledge Marco Gabiccini for the inspiring discussion, Alessandro Tognetti for his assistance in sensing glove configuration and Stefano Mastria for the help in setting up data acquisition.

References

[1] A. Bicchi. Hands for dextrous manipulation and robust grasping: a difficult road towards simplicity. *IEEE Trans. on Robotics and Automation*, 16(6):652 – 662, 2000.

[2] C. Brown and H. Asada. Inter-finger coordination and postural synergies in robot hands via mechanical implementation of principal component analysis. In *IEEE-RAS International Conference on Intelligent Robots and Systems*, pages 2877 – 2882, 2007.

[3] M. Fahrer. *The hand*, chapter Interdependent and independent actions of the fingers, pages 399 – 403. Philadelphia, PA: Saunders, 1981.

[4] M. Flanders and J. F. Soechting. Kinematics of typing: parallel control of the two hands. *J Physiol*, 67:1264 – 1274, 1992.

[5] 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*, pages 4606 – 4609, 2010.

[6] M. Gabiccini and A. Bicchi. On the role of hand synergies in the optimal choice of grasping forces. In *Robotics Science and Systems*, 2010.

[7] A. Gelb. *Applied Optimal Estimation*. M.I.T. Press, Cambridge, USMA, 1974.

[8] A. Gordon. *Handbook of Brain and Behaviour in Human Development*, chapter Development of hand motor control, pages 513 – 537. The Netherlands: Kluwer Academic, 2001.

[9] G. J. Grimer. Digital data entry glove interface device, Nov. 1983.

[10] W. Hardle and L. Simar. *Applied Multivariate Statistical Analysis*. Springer-Verlag Berlin Heidelber New York, 2007.

[11] D. Hawkins. *Identification of Outliers*. Chapman and Hall, London, 1980.

[12] S. Kilbreath and S. Gandevia. Limited independent flexion of the thumb and fingers in human subjects. *J Physiol*, 543:289 – 296, 2002.

[13] P. C. Mahalanobis. On the generalised distance in statistics. *Proceedings of the National Institute of Sciences of India*, 2(1):49 – 55, 1936.

[14] C. R. Mason, J. E. Gomez, and T. J. Ebner. Hand synergies during reach-to-grasp. *J Neurophysiol*, 86:2896 – 2910, 2001.

[15] S. Mulatto, A. Formaglio, M. Malvezzi, and D. Prattichizzo. Animating a synergy-based deformable hand avatar for haptic grasping. In *International Conference EuroHaptics*, 2010.

[16] L. Pao and T. Speeter. Transformation of human hand positions for robotic hand control. In *Robotics and Automation, 1989. Proceedings., 1989 IEEE International Conference on*, volume 3, pages 1758 – 1763, 1989.

[17] D. Prattichizzo, M. Malvezzi, and A. Bicchi. On motion and force controllability of grasping hands with postural synergies. In *Proceedings of Robotics: Science and Systems*, volume 2, 2010.

[18] M. Santello, M. Flanders, and J. F. Soechting. Postural hand synergies for tool use. *The Journal of Neuroscience*, 18(23):10105 – 10115, 1998.

[19] M. H. Schieber and M. Santello. Hand function: peripheral and central constraints on performance. *Journal of Applied Physiology*, 96(6):2293 – 2300, 2004.

[20] D. J. Sturman and D. Zeltzer. A survey of glove-based input. *Computer Graphics and Applications, IEEE*, 14(1):30 – 39, 1994.

[21] A. Tarantola. *Inverse Problem Theory and Model Parameter Estimation*. SIAM, 2005.

[22] A. Tognetti, N. Carbonaro, G. Zupone, and D. De Rossi. Characterization of a novel data glove based on textile integrated sensors. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC06, Proceedings.*, pages 2510 – 2513, 2006.

[23] X. Zhang, S. Lee, and P. Braido. Determining finger segmental centers of rotation in flexion-extension based on surface marker measurement. *Journal of Biomechanics*, 36:1097 – 1102, 2003.