Synergy–Based Hand Pose Sensing: Optimal Glove Design*

Matteo Bianchi[†] Paolo Salaris[†] Antonio Bicchi[†‡]

Abstract

In this paper we study the problem of optimally designing glove–based sensing devices for hand pose reconstruction to maximize their potential for precision. In a companion paper we studied the problem of maximizing the reconstruction accuracy of the hand pose from partial and noisy data provided by any given pose sensing device (a sensorized "glove") taking into account the knowledge on how humans most frequently use their hands in grasping tasks. In this paper we consider the dual problem of how to design pose sensing devices, i.e. how and where to place sensors on a glove, to get maximum information about the actual hand posture. We study the optimal design of gloves of different nature, according to a classification of current sensing technologies adopted in the domain. The objective is to provide, for given *a priori* information and fixed number of sensor inputs, the optimal design minimizing the reconstruction error statistics (assuming that optimal reconstruction algorithms are adopted). Finally, an experimental evaluation of the proposed method for optimal design is provided.

1 Introduction

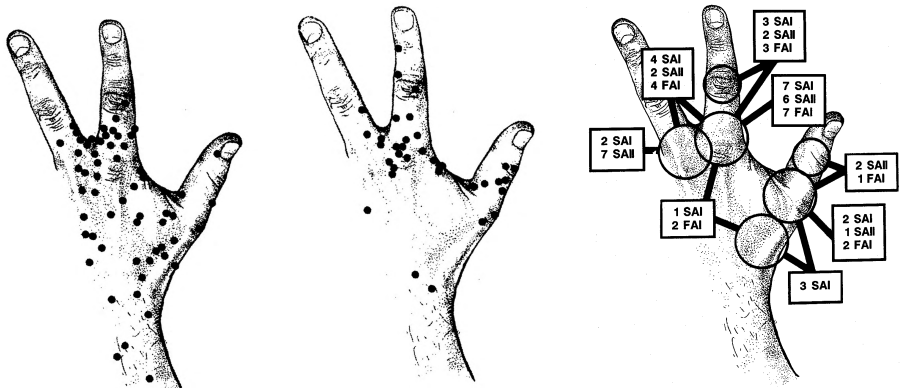
Hand Pose Reconstruction (HPR) systems are gaining an increasing importance, since they provide useful human–machine interfaces in many applications ranging from tele–robotics, to virtual reality, entertainment and rehabilitation [Dipietro et al., 2008]. To enable a widespread use of low–cost glove–based HPR, it is crucial to address the problem of correct hand pose estimation despite the many non–idealities arising, for example, from the complexity of human hand biomechanics and from measurement inaccuracies.

In [Bianchi et al., 2012b] we describe how to optimize HPR system accuracy — for a given hardware configuration — so as to provide optimal hand pose estimation from incomplete and imperfect glove data. In the present paper we extend the analysis to the optimal sensing glove design. The issue is to choose the optimal sensor distribution

*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 the ERC Advanced Grant no. 291166 "SoftHands": "A Theory of Soft Synergies for a New Generation of Artificial Hands".

[†]The Interdept. Research Center "Enrico Piaggio", University of Pisa, Largo Lucio Lazzarino 1, 56126 Pisa, Italy. m.bianchi,p.salaris,bicchi@centropiaggio.unipi.it

[‡]Department of Advanced Robotics, Istituto Italiano di Tecnologia, via Morego, 30, 16163 Genova, Italy



(a) Slowly adapting (SA) units location.

(b) Fast adapting (FA) units location.

(c) Mechanoreceptor afferent units responding to ≥ 1 joint.

Figure 1: Location of cutaneous mechanoreceptive units in the dorsal skin of the human hand. Adapted from [Edin and Abbs, 1991], courtesy of the authors.

which maximizes the information on the actual posture. This information, used with the algorithms proposed in [Bianchi et al., 2012b], will lead to the minimization of the reconstruction error statistics.

That the optimal distribution of sensitivity for HPR is not trivial is strongly suggested by the observation of the human example. Indeed, the role of cutaneous information in kinesthesia and movement control in human hands and fingers has been extensively studied in the neuro-physiology literature. For example, [Edin and Abbs, 1991] describe the response to finger movements of cutaneous mechanoreceptors in the dorsal skin of human hand. The involved mechanoreceptors mainly consist of two types: Fast Adapting afferents of the first type (FAI), and Slow Adapting afferents, of both the first and second type (SAI and SAII, respectively). The non–uniform distribution of these receptors is shown in figure 1(a) and 1(b). FA units, whose response is localized to movements about one or, at most, two nearby joints, are found primarily close to joints; on the other hand SA units, whose discharge rate is influenced by several joints at the same time, are more uniformly distributed (see figure 1(c)).

These observations suggest that in the human hand sensory system, different typologies of proprioceptive sensors are distributed in the dorsal skin with different densities. This produces a non–uniform map of sensitivities to joint angles, whose functional motivation is unclear, but might correspond to the different importance of different elementary percepts in building an overall representation of the hand pose. Motivated by this biological evidence, in this paper we deal with the problem of searching for a preferential distribution and density of different typologies of sensors, so as to optimize the accuracy of glove–based HPR systems.

The design of sensorized gloves for HPR has received already some attention in the literature. For example, in [Sturman and Zeltzer, 1993] an investigation of "whole-hand" interfaces for the control of complex tasks is presented, along with the descrip-

tion, design, and evaluation of whole-hand inputs, based on empirical data from users. In [Edmison et al., 2002] authors discussed the properties, advantages, and design aspects associated with piezoelectric materials for sensing glove design, in an application where the device is used as a keyboard. Finally, [Chang et al., 2007] explored how to methodically select a minimal set of hand pose features from optical marker data for HPR. The objective is to determine marker locations on the hand surface that is appropriate for classification of hand poses. All the aforementioned approaches rely on ad–hoc experimental or qualitative observations: from actual sensor data, locations that provide the largest and most useful information on the system are chosen.

More analytical approaches to the problem can be based on the existing literature on optimal experimental design (see e.g. [Pukelsheim, 2006]). Among optimal design criteria, Bayesian methods are ideally suited to contribute to experimental design for error statistics minimization (see e.g. [Chaloner and Verdinelli, 1995, Ghosh and Rao, 1996, Bicchi and Canepa, 1994] for a review). On the contrary, non–Bayesian criteria are adopted when the linear Gaussian hypothesis is not fulfilled and/or when the designer's primary concern is to minimize worst-case errors rather than error statistics. Criteria on explicit worst-case/deterministic bounds on the errors and tools from the theory of optimal worst–case/deterministic estimation and/or identification are discussed e.g. in [Helmicki et al., 1991, Tempo, 1988, Bicchi, 1992].

It is noteworthy that most results in optimal experimental design refer to the case where the number of measurements is redundant or at least equal to the number of variables to be estimated. However, the opposite case that fewer sensors are available than the hand variables is of main concern in our problem. To circumvent this problem, it is natural to think of exploiting *a priori* knowledge to disambiguate poses from scarce data.

In this paper, partially based on [Bianchi et al., 2012a], optimal HPR system design is obtained by relying on results from neuro–scientific studies about hand postures in grasping tasks [Santello et al., 1998, Schieber and Santello, 2004]. The main finding of these studies is that, despite the hand complexity, simultaneous motion and force of the fingers occur in a consistent fashion and is characterized by coordination patterns ("synergies") that effectively reduce the number of independent hand degrees of freedom to be controlled. In [Bianchi et al., 2012b], this prior knowledge on how humans most frequently use their hands is fused with partial and noisy data provided by any given HPR device, to maximize reconstruction accuracy. Here, the goal is to characterize a design which enables for optimally exploiting – in a Bayesian sense – such an *a priori* information.

The optimization goals of this paper become particularly relevant when restrictions on the production costs limit both the number and the quality of sensors. In these cases, a careful design is instrumental to obtain good performance. Different technologies and sensor distributions can be considered to realize an HPR device. At the physical level, sensors for HPR gloves can be classified as either *lumped* (as e.g. a mechanical angular encoder about a joint) or *distributed* (e.g. a flexible optic fiber running along a finger from base to tip, as in the DataGlove [Zimmerman, 1988]). At the signal level, glove sensors can be *coupled* (if more than one hand joint angle influences the reading) or *uncoupled*. Of course, all distributed sensors are coupled, but also lumped sensors can be coupled due to cross–coupling.

An example of a distributed, coupled glove is the one described in [Tognetti et al., 2006] (cf. figure 2, on the left). On the other hand, the Humanglove (Humanware s.r.l., Pisa, Italy) shown in figure 2, on the right, is an example of lumped glove.

Figure 2: Examples of sensorized gloves. On the left, a sensing glove based on conductive elastomer sensor strips printed on fabric, each measuring a linear combination of joint angles [Tognetti et al., 2006]. On the right, the Humanglove (image courtesy by Humanware s.r.l. (www.hmw.it/)), using individual joint angle Hall effect sensors. Depending on the level of cross–talk between joint angles and measurements, this device can result in either a coupled or uncoupled model.

To validate our optimal designs we use an experimental setup consisting of a highly accurate optical tracking system. We use these data to measure a number of hand postures completely, providing reliable ground truth, and to simulate low-accuracy sensors of different design. Optimal solutions referring to different sensing strategies are compared in terms of HPR performance. Statistical analyses on the estimated hand postures obtained by using measurements coming from simulated optimal gloves demonstrate the effectiveness of the proposed method.

2 Problem Definition

For reader's convenience we summarize here the definitions and results of [Bianchi et al., 2012b]. Let us assume a n Degrees of Freedom (DoFs) kinematic hand model and let $y \in \mathbb{R}^m$ be the measures provided by a sensing glove. The relationship between joint variables $x \in \mathbb{R}^n$ and measurements y is

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

where $H \in \mathbb{R}^{m\times n}$ $(m < n)$ is a full row rank matrix, and $v \in \mathbb{R}^m$ is a vector of measurement noise which has zero mean and Gaussian distribution with covariance matrix R. Different sensor arrangements generate different measurement matrices H: the row corresponding to a lumped, uncoupled sensor has a non–zero element only in correspondence to the measured joint, hence (up to rescaling) it is a binary "selection" matrix. We will call such a matrix *discrete*, i.e. $H_{ij} \in \{0,1\}$ — a *discrete* set of values. Conversely, a coupled sensor with general weights, i.e. a distributed sensor or a lumped sensor with not negligible cross–coupling, produces a matrix whose row elements are

real numbers, i.e., up to rescaling, $H_{ij} \in [-1, 1] \subset \mathbb{R}$ — a *continuous* set of values. In the following, we will call such a matrix *continuous*. Finally, a glove employing both lumped (uncoupled and coupled) and distributed sensors will generate a *hybrid* measurement matrix which consists of a continuous part and a discrete one. Notice that, according to the discussion in figure 1, the human hand sensing distribution could be considered to belong to the latter glove class.

In [Bianchi et al., 2012b], 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. To improve the hand pose reconstruction, we used postural synergy information embedded in the *a priori* grasp set, which is obtained by collecting a large number N of grasp postures x_i, consisting of n DoFs, into a matrix $X \in \mathbb{R}^{n \times N}$. This information can be summarized with a covariance matrix $P_o \in \mathbb{R}^{n \times n}$, which is a symmetric matrix computed as $P_o = \frac{(X-\bar{x})(X-\bar{x})^T}{N-1}$, where $\bar{x}$ is a matrix $n \times N$ whose columns contain the mean values for each joint angle arranged in vector $\mu_o \in \mathbb{R}^n$.

Based on the Minimum Variance Estimation (MVE) technique, in [Bianchi et al., 2012b] we obtained the hand pose reconstruction as

$$\hat{x} = (P_o^{-1} + H^T R^{-1} H)^{-1} (H^T R^{-1} y + P_o^{-1} \mu_o), \tag{2}$$

where matrix $P_p = (P_o^{-1} + H^T R^{-1} H)^{-1}$ is the *a posteriori* covariance matrix. When R tends to assume very small values, the solution described in (2) might encounter numerical problems. However, by using the Sherman-Morrison-Woodbury formulae, (2) can be rewritten as

$$\hat{x} = \mu_o - P_o H^T (H P_o H^T + R)^{-1} (H \mu_o - y), \tag{3}$$

and the *a posteriori* covariance matrix becomes $P_p = P_o - P_o H^T (H P_o H^T + R)^{-1} H P_o$.

The *a posteriori* covariance matrix, which depends on measurement matrix H, represents a measure of the amount of information (i.e. observed information) that an observable variable carries about unknown parameters. In this paper we explore the role of the measurement matrix H on the estimation procedure, providing the optimal design of a sensing device able to obtain the maximum amount of the information on the actual hand posture.

Let us preliminary introduce some useful notations. If M is a symmetric matrix with dimension n, let its Singular Value Decomposition (SVD) be $M = U_M \Sigma_M U_M^T$, where Σ_M is the diagonal matrix containing the singular values $\sigma_1(M) \geq \sigma_2(M) \geq \cdots \geq \sigma_n(M)$ of M and U_M is an orthogonal matrix whose columns $u_i(M)$ are the eigenvectors of M, known as Principal Components (PCs) of M, associated with $\sigma_i(M)$. For example, the SVD of the *a priori* covariance matrix is $P_o = U_{P_o} \Sigma_{P_o} U_{P_o}^T$, with $\sigma_i(P_o)$ and $u_i(P_o)$, $i = 1, 2, \ldots, n$, the singular values and the principal components of matrix P_o, respectively.

3 Optimal Sensing Design

Referring to model (1), we first analyze the case that individual sensing elements in the glove can be designed to measure a linear combination of joint angles (i.e. a glove

which can be modeled with a continuous matrix), and provide, for given *a priori* information and fixed number of measurements, the optimal continuous measurement matrix, minimizing the reconstruction error statistics. We then consider the case where each measure provided by the glove corresponds to a single joint angle (i.e. a glove which can be modeled with a discrete matrix). For these types of gloves we determine the optimal discrete measurement matrix, i.e. which joints should be individually measured in order to optimize the design. Finally, we also define a procedure to obtain the optimal hybrid measurement matrix.

In the ideal case of noise–free measures ($R = 0$), P_p becomes zero when H is a full rank n matrix, meaning that available measures contain a complete information about the hand posture. In the real case of noisy measures and/or when the number of measurements m is less than the number of DoFs n, P_p can not be zero. In these cases, the following problem becomes very interesting: find the optimal matrix H^* such that the hand posture information contained in the few number of measurements is maximized. Without loss of generality, we assume H to be full row rank and we consider the following problem.

Problem 1. *Let H be an $m \times n$ full row rank matrix with $m < n$ and $V_1(P_o, H, R) : \mathbb{R}^{m\times n} \to \mathbb{R}$ be defined as $V_1(P_o, H, R) = \|P_o - P_o H^T (HP_o H^T + R)^{-1} HP_o\|_F^2$, find*

$$H^* = \arg\min_H V_1(P_o, H, R)$$

where $\|\cdot\|_F$ denotes the Frobenius norm defined as $\|A\|_F = \sqrt{\mathrm{tr}(AA^T)}$, for $A \in \mathbb{R}^{n\times n}$.

To solve problem 1 means to minimize the entries of the *a posteriori* covariance matrix: the smaller the values of the elements in P_p, the greater is the predictive efficiency.

In order to simplify the analysis, in the following we will consider separately the optimal procedure to define the continuous, discrete and hybrid measurement matrix.

3.1 Continuous Measurement Matrix

For this case, each row of the measurement matrix H is a vector in $\mathbb{R}^n$ and hence can be given as a linear combination of a $\mathbb{R}^n$ basis. Without loss of generality, we can use the principal components of matrix P_o, i.e. the columns of the previously defined matrix U_{P_o}, as a basis of $\mathbb{R}^n$. Consequently the measurement matrix can be written as $H = H_e U_{P_o}^T$, where $H_e \in \mathbb{R}^{m\times n}$ contains the coefficients of the linear combinations. Given that $P_o = U_{P_o} \Sigma_{P_o} U_{P_o}^T$, the *a posteriori* covariance matrix becomes

$$P_p = U_{P_o} \left[\Sigma_o - \Sigma_o H_e^T (H_e \Sigma_o H_e^T + R)^{-1} H_e \Sigma_o\right] U_{P_o}^T, \tag{4}$$

where, for simplicity of notation, $\Sigma_o \equiv \Sigma_{P_o}$.

Next sections are dedicated to how to obtain the optimal the continuous measurement matrix both in a numerical and analytical manner. For this purpose, let us introduce the set of $m \times n$ (with $m < n$) matrices with orthogonal rows, i.e. satisfying the condition $HH^T = I_{m\times m}$, and let us denote it as $\mathscr{O}_{m\times n}$.

3.1.1 Analytical Solutions

We first consider the case of noise–free measures, i.e. $R = 0$. Let A be a non–negative matrix of order n. It is well known (cf. [Rao, 1964]) that, for any given matrix B of rank m with $m \leq n$,

$$\min_B \|A - B\|_F^2 = \alpha_{m+1}^2 + \cdots + \alpha_n^2 \,, \tag{5}$$

where α_i are the eigenvalues of A, and the minimum is attained when

$$B = \alpha_1 w_1 {w_1}^T + \cdots + \alpha_m w_m {w_m}^T \,, \tag{6}$$

where w_i are the eigenvector of A associated with α_i. In other words, the choice of B as in (6) is the best fitting matrix of given rank m to A. By using this result we are able to show when the minimum of (4), hence of

$$\|\Sigma_o - \Sigma_o H_e^T (H_e \Sigma_o H_e^T)^{-1} H_e \Sigma_o\|_F^2 \,, \tag{7}$$

can be reached. Let us preliminary observe that the row vectors $(h_i)_e$ of H_e can be chosen, without loss of generality, to satisfy the condition $(h_i)_e \Sigma_o (h_j)_e = 0,\ i \neq j$, which implies that the measures are uncorrelated ([Rao, 1964]). Let $\mathscr{O}_{m\times n}$ denotes the set of $m \times n$ matrices, with $m < n$, whose rows satisfy the aforementioned condition, i.e. the set of matrices with orthonormal rows ($H_e H_e^T = I$). By using (5), the minimum of (7) is obtained when (cf. [Rao, 1964])

$$\begin{aligned} \Sigma_o H_e^T (H_e \Sigma_o H_e^T)^{-1} H_e \Sigma_o = \sigma_1(\Sigma_o) u_1(\Sigma_o) u_1^T(\Sigma_o) + \cdots + \\ + \sigma_m(\Sigma_o) u_m(\Sigma_o) u_m^T(\Sigma_o) \,. \end{aligned} \tag{8}$$

Since Σ_o is a diagonal matrix, $u_i(\Sigma_o) \equiv e_i$, where e_i is the *i-th* element of the canonical basis. Hence, it is easy to verify that (8) holds for $H_e = [I_m \,|\, 0_{m\times(n-m)}]$. As a consequence, row vectors (h_i) of H^* are the first m principal components of P_o, i.e. $(h_i) = u_i(P_o)^T$, for $i = 1, \ldots, m$.

From these results, a principal component can be defined as a linear combination of optimally–weighted observed variables meaning that the corresponding measures can account for a maximal amount of variance in the data set. As reported in [Rao, 1964], every set of m optimal measures can be considered as a representation of points in the best fitting lower dimensional subspace. Thus the first measure gives the best one–dimensional representation of data set, the first two measures give the best two–dimensional representation, and so on.

In the noisy measurement case, (8) can be rewritten as

$$\begin{aligned} \Sigma_o H_e^T (H_e \Sigma_o H_e^T + R)^{-1} H_e \Sigma_o - \sigma_1(\Sigma_o) u_1(\Sigma_o) u_1^T(\Sigma_o) + \cdots + \\ + \sigma_m(\Sigma_o) u_m(\Sigma_o) u_m^T(\Sigma_o) = \Delta \end{aligned}$$

In this case, $\Delta = 0$ can not be attained for any finite H: indeed, for unconstrained H, $\inf_H V_1(P_0, H, R)$ would be attained for $\|H\| \to \infty$, i.e. for infinite signal–to–noise ratio. The problem can be recast in a well–posed form by imposing a constraint on the magnitude of the measurement matrix. Up to a possible re–normalization of R, we

can search the optimum design in the set $\mathscr{A} = \{H : HH^T = I_m\}$. This problem was discussed and solved in [Diamantaras and Hornik, 1993], showing that, for arbitrary noise covariance matrix R,

$$\min_{H\in\mathscr{A}} V_1(H) = \sum_{i=1}^{m} \frac{\sigma_i(P_o)}{1+\sigma_i(P_o)/\sigma_{m-i+1}(R)} + \sum_{i=m+1}^{n} \sigma_i(P_o)\,, \tag{9}$$

which is attained for

$$H^* = \sum_{i=1}^{m} u_{m-i+1}(R) u_i^T(P_o)\,. \tag{10}$$

Hence, if $\mathscr{A}$ consists of all matrices with mutually perpendicular, unit length rows, the first m principal components of P_o are still the optimal choice for H rows only in case of uncorrelated noise (i.e. R is a diagonal matrix). For generic noise covariance matrices, the optimal choice of matrix H depends also on the principal components of noise matrix R. The alternative case that the solution is sought under a Frobenius norm constraint on H, i.e. $\mathscr{A} = \{H : \|H\|_F \leq 1\}$ is discussed in [Diamantaras and Hornik, 1993].

3.1.2 Numerical Solution: Gradient flows on $\mathscr{O}_{m\times n}$

In this subsection we describe a different approach to the solution of problem 1, which consists of constructing a differential equation whose trajectories converge to the desired optimum. The method lends itself directly to efficient numerical implementations. Although a closed–form solution has been proposed in the previous subsection, the numerical solution considered here is very useful when constraints are imposed on the measurement structure (as they will be for instance in the hybrid case), where closed form solutions are not applicable.

The following proposition describes an algorithm that minimizes the cost function $V_1(P_o,H,R)$, providing the gradient flow which will be useful in the method of steepest descent.

Proposition 1. *The gradient flow for the function $V_1(P_o,H,R) : \mathbb{R}^{m\times n} \to \mathbb{R}$ is given by,*

$$\dot{H} = -\nabla \|P_p\|_F^2 = 4\left[P_p^2 P_o H^T \Sigma(H)\right]^T\,, \tag{11}$$

where $\Sigma(H) = (HP_oH^T + R)^{-1}$.

Proof. See Appendix. □

Let us observe that rows of matrix H can be chosen, without loss of generality, such that $H_iP_oH_j^T = 0$, $i \neq j$ which imply that measures are uncorrelated, i.e. satisfying the condition $HH^T = I_m$. Of course, in case of noise–free sensors, this constraint is not strictly necessary. On the other hand, in case of noisy sensors, the minimum of $V_1(P_o,H,R)$ can not be obtained since it represents a limit case that can be achieved when H becomes very large (i.e. an infimum) and hence increasing the signal–to–noise ratio.

A reasonable solution for the constrained problem will be provided by using the Rosen's gradient projection method for linear constraints [Rosen, 1960], which is based

on projecting the search direction onto the subspace tangent to the constraint. Hence, given the steepest descent direction for the unconstrained problem, this method consists on finding the direction with the most negative directional derivative which satisfies the constraint on the structure of the matrix H, i.e. $HH^T = I_m$. This can be obtained by using the projection matrix

$$W = I_n - H^T(HH^T)^{-1}H\,, \tag{12}$$

and then projecting the unconstrained gradient flow (11) into the subspace tangent to the constraint, obtaining the search direction

$$s = \nabla \|P_p\|_F^2 \; W\,. \tag{13}$$

Having the search direction for the constrained problem, the gradient flow is given by

$$\dot{H} = -4\left[P_p^2 P_o H^T \Sigma(H)\right]^T \; W \tag{14}$$

where $\Sigma(H) = (HP_oH^T + R)^{-1}$. The gradient flow (11) guarantees that the optimal solution H^* will satisfy $H^*(H^*)^T = I_m$, if $H(0)$ satisfies $H(0)H(0)^T = I_m$, i.e. $H \in \mathscr{O}_{m\times n}$.

Notice that both $\mathscr{O}_{m\times n}$ and $V_1(P_o,H,R)$ are not convex, hence the problem could not have a unique minimum. However, in case of noise–free measures, the invariance of the cost function w.r.t. changes of basis, i.e. $V_1(P_o,H,0) = V_1(P_o,MH,0)$, with $M \in \mathbb{R}^{m\times m}$ a full rank matrix, suggests that there exists a subspace in $\mathbb{R}^{m\times n}$ where the optimum is achieved. Hence, considering results reported in previous section (see (8)), we can conclude that gradient (11) with $R = 0$ becomes zero when rows of matrix H are any linear combination of the first m principal components of the *a priori* covariance matrix P_o. This does not happen in case of noisy measures with generic noise covariance matrix and gradient (14) becomes zero only for a particular matrix H which depends also on the principal components of the matrix R (see (10)). Notice that, according to previous analytical results, when the noise covariance matrix is diagonal, then gradient (14) becomes zero only if matrix H consists of the first m principal components of P_o.

3.2 Discrete Measurement Matrix

When each measure y_j, $j = 1,\ldots,m$ provided by the glove corresponds to a single joint angle x_i, $i = 1,\ldots,n$, the problem is to find the optimal choice of m joints or DoFs to be measured. Measurement matrix becomes in this case a full row rank matrix where each row is a vector of the canonical basis, i.e. matrices which have exactly one nonzero entry in each row.

Let $\mathscr{N}_{m\times n}$ denote the set of $m\times n$ element–wise non–negative matrices, then $\mathscr{P}_{m\times n} = \mathscr{O}_{m\times n} \cap \mathscr{N}_{m\times n}$, where $\mathscr{P}_{m\times n}$ is the set of $m\times n$ permutation matrices (see lemma 2.5 in [Zavlanos and Pappas, 2008]). This result implies that if we restrict H to be orthonormal and element–wise non–negative, we get a permutation matrix. In this paper we extend this result in $\mathbb{R}^{m\times n}$, obtaining matrices which have exactly one nonzero entry in each row. Hence, the problem to solve becomes:

Problem 2. *Let H be a $m \times n$ matrix with $m < n$, and $V_1(P_o, H, R) : \mathbb{R}^{m \times n} \to \mathbb{R}$ be defined as $V_1(P_o, H, R) = \|P_o - P_o H^T (H P_o H^T + R)^{-1} H P_o\|_F^2$, find the optimal measurement matrix*

$$H^* = \arg\min_H V_1(P_o, H, R)$$
$$s.t. \quad H \in \mathscr{P}_{m \times n}.$$

In this case a closed–form solution is not available. Nonetheless, as the model used to describe the kinematics of the hand has usually a low number of DoFs, the optimal choice H^* can be computed by exhaustion, substituting all possible sub–sets of m vectors of the canonical basis in the cost function $V_1(P_o, H, R)$. In the next section, a more general approach to compute the optimal matrix will be provided in order to obtain a result also when a model with a large number of DoFs is considered.

3.2.1 Numerical Solution: Gradient Flows on $\mathscr{P}_{m \times n}$

In this section, we describe an alternative approach to the solution of problem 2 based on a gradiental method. Once again, although the enumeration approach can solve the problem in practical cases, the numerical solution based on the method here presented will be useful in the hybrid case.

A numerical solution for problem 2 can be obtained following a method presented in [Zavlanos and Pappas, 2008], which consists in defining a function $V_2(P)$ with $P \in \mathbb{R}^{n \times n}$ that forces the entries of P to be as positive as possible, thus penalizing negative entries of H. In this paper, we extend this function to measurement matrices $H \in \mathbb{R}^{m \times n}$ with $m < n$. Consider a function $V_2 : \mathscr{O}_{m \times n} \to \mathbb{R}$ as

$$V_2(H) = \frac{2}{3} \operatorname{tr}\left[H^T (H - (H \circ H))\right], \tag{15}$$

where $A \circ B$ denotes the *Hadamard* or element–wise product of the matrices $A = (a_{ij})$ and $B = (b_{ij})$, i.e. $A \circ B = (a_{ij} b_{ij})$. The gradient flow of $V_2(H)$ is given by ([Zavlanos and Pappas, 2008])

$$\dot{H} = -H\left[(H \circ H)^T H - H^T (H \circ H)\right], \tag{16}$$

which minimizes $V_2(H)$ converging to a permutation matrix if $H(0) \in \mathscr{O}_{m \times n}$.

The two gradient flows given by (11) and (16), both defined on the space of orthogonal matrices, tend to minimize their cost functions, respectively. By combining these two gradient flows we can achieve a solution for Problem 2. An interesting result applies to the dynamics of the convex combination of these gradients, which can be stated as follows.

Theorem 1. *Let $H \in \mathbb{R}^{m \times n}$ with $m < n$ be the measurement process matrix and assume that $H(0) \in \mathscr{O}_{m \times n}$. Moreover, suppose that $H(t)$ satisfies the following matrix differential equation,*

$$\begin{aligned} \dot{H} = {} & 4(1-k)\left[P_p^2 P_o H^T \Sigma(H)\right]^T W + \\ & + k\, H\left[(H \circ H)^T H - H^T (H \circ H)\right], \end{aligned} \tag{17}$$

where $k \in [0,\,1]$ is a positive constant, $W = I_n - H^T(HH^T)^{-1}H$ and $\Sigma(H) = (HP_oH^T + R)^{-1}$. For sufficiently large k, $\lim_{t\to\infty} H(t) = H_\infty$ exists and approximates a permutation matrix that also (locally) minimizes the squared Frobenius norm of the a posteriori covariance matrix, $\|P_p\|_F^2$.

The proof of this theorem is a direct extension of results in [Zavlanos and Pappas, 2008], and is omitted for brevity.

As in most numerical optimization algorithms, the non–convex nature of the cost function and of the support set implies the need for multi–start approaches. A possible technique to help converge towards the global optimum consists in increasing k during the search procedure (cf. [Zavlanos and Pappas, 2008]).

3.3 Hybrid Measurement Matrix

Up to re–arranging the sensor numbering, we can write a hybrid measurement matrix $H_{c,d} \in \mathbb{R}^{m\times n}$ as

$$H_{c,d} = \left[\frac{H_c}{H_d}\right],$$

where $H_c \in \mathbb{R}^{m_c\times n}$ defines the m_c rows of the continuous part, whereas $H_d \in \mathscr{P}^{m_d\times n}$ describes the m_d single–joint measurements of the discrete part, with $m_c + m_d = m$. Neither the closed–form solution valid for the continuous measurement matrix, nor the exhaustion method used for discrete measurements are applicable in the hybrid case. Therefore, to optimally determine the hybrid measurement matrix, we will recur to gradient–based iterative optimization algorithms.

By combining the continuous and discrete gradient flows, previously defined in (11) and (16), respectively, and constraining the solution in the sub–set $\mathscr{H}_{c,d} = \{H_{c,d} : H_{c,d}H_{c,d}^T = I_m\}$, we obtain

$$\begin{aligned}\dot{H}_{c,d} = 4\,(1-k)\;\left[P_p^2 P_o H_{c,d}^T \Sigma(H_{c,d})\right]^T\; W + \\ + k\,\bar{H}_d\left[(\bar{H}_d\circ\bar{H}_d)^T\bar{H}_d - \bar{H}_d^T(\bar{H}_d\circ\bar{H}_d)\right],\end{aligned} \tag{18}$$

where $k \in [0,\,1]$ is a positive constant, $P_p = P_o - P_oH_{c,d}^T(H_{c,d}P_oH_{c,d}^T + R)^{-1}H_{c,d}P_o$, $W = I_n - H_{c,d}^T(H_{c,d}H_{c,d}^T)^{-1}H_{c,d}$, $\Sigma(H_{c,d}) = (H_{c,d}P_oH_{c,d}^T + R)^{-1}$, and

$$\bar{H}_d = \left[\frac{0_{m_c\times n}}{H_d}\right].$$

Starting from any initial guess matrix $H_{c,d} \in \mathscr{H}_{c,d}$, the gradient flow defined in (18) remains in the sub–set $\mathscr{H}_{c,d}$ and, on the basis of Theorem 1, it converges toward a hybrid measurement matrix, (locally) minimizing the squared Frobenius norm of the *a posteriori* covariance matrix. Multi–start strategies have to be used to circumvent the problem of local minima.

When noise is not negligible, without constraining the solution in $\mathscr{H}_{c,d}$ by W, the gradient search method of (18) would tend to produce measurement matrices whose continuous parts, H_c, are very large in norm. This is an obvious consequence of the fact that, for a fixed noise covariance R, larger measurement matrices H would produce an apparently higher signal–to–noise ratio in (1).

4 On the Practical Realization of Optimal Sensing Devices

In this section we describe some feasible solutions to realize a device which can be modeled by the previously obtained optimal continuous, discrete and hybrid measurement matrices, to achieve a trade–off between accuracy, cost, usability and ease to made, on the basis of the current technology.

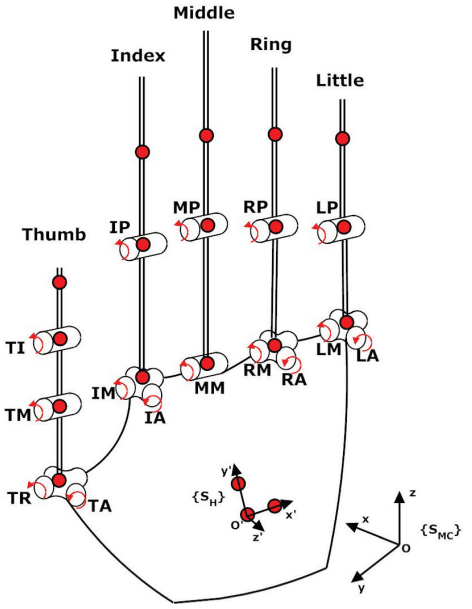
Lumped, uncoupled sensing devices, which generate a discrete measurement matrix, are probably the easiest to be implemented, as they require to individually measure single joints according to the optimal measurement matrix. Common sensing strategies include Hall–effect (e.g. Humanglove) or piezoresistive sensors (e.g. CyberGlove, by CyberGlove System LLC, San Jose, CA – USA), directly placed on the joints to be measured, hence obtaining a lumped device. However some difficulties can occur due to coupling between non–measured and measured joints and cross–coupling between sensors. This last problem becomes particularly relevant when the optimality requires to place sensors on adjacent joints. To circumvent these drawbacks, attention must be paid on the ergonomics of sensor physical support as well as appropriate sensor shielding (e.g. with Hall–effect transducer).

On the other hand, distributed sensors, which generate an optimal continuous matrix, should provide measurements in terms of optimally weighted linear combinations of the contributions of different DoFs, according to the principal components of the *a priori* data and of noise covariance matrix. Of course, under a technological viewpoint, the weight of each measured DoF can be approximated with a given level of accuracy due to the particular application and sensor technology. The literature ([Dipietro et al., 2008]) describes at least two main technologies to implement distributed sensing strategies. A first one is based on resistive ink printed on flexible plastic bends that follow the movement of hand joints (e.g. PowerGlove by Mattel Inc., El Segundo, CA–USA). In this case a change in the configuration of the joints results in a change of the overall resistance. It is possible to weight each joint contribution in a different manner by suitably modifying its electrical resistance (e.g. increasing or decreasing the section of the printed segments corresponding to different joints). A second technology adopts capacitive sensors (as e.g. in the Didjiglove by Dijiglove Pty. Ltd., Melbourne, AUS), i.e. two layers of conductive polymer separated by a dielectric, which overlap different joints. By bending the sensors, a change in the overall capacitance will be produced. Also in this case we can weight each joint differently by varying the thickness and the type of dielectric surface as well as the electrode surface overlapping. Finally, the above discussed technologies (lumped, uncoupled and distributed) can be adopted and combined in an efficient manner to optimally realize devices which can be modeled by a hybrid measurement matrix.

5 Experimental Setup for A Priori and Validation Data

To validate our optimal design, we performed some simulations using measurements from human posture data acquired by means of an optical tracking system. As described in the following, this data contains joint angles for a large set of grasping poses,

and can be regarded as a reasonable reference to compare reconstruction outcomes. These values were used to simulate optimally designed gloves, i.e. they were combined each time to provide a measurement outcome according to the continuous, hybrid and discrete measurement matrix, i.e. H_c^*, $H_{c,d}^*$ and H_d^*, respectively, as well as to other non–optimal measurement matrix. Without loss of generality, for hand pose reconstruction



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 3: Kinematic model of the hand with 15 DoFs.

we adopt the 15 DoFs model also used in [Santello et al., 1998, Gabiccini et al., 2011] and reported in figure 3. The model DoFs are: 4 DoFs for the thumb, i.e. TR, TA, TM, TI (Thumb Rotation, Abduction, Metacarpal, Interphalangeal); 3 DoFs for the index, i.e. IA, IM, IP (Index Abduction, Metacarpal, Proximal Interphalangeal); 2 DoFs for the middle, i.e. MM, MP (Middle Metacarpal, Proximal Interphalangeal); 3 DoFs for the ring, i.e. RA, RM, RP (Ring Abduction, Metacarpal, Proximal Interphalangeal); 3 DoFs for the little, i.e. LA, LM, LP (Little Abduction, Metacarpal, Proximal Interphalangeal). Notice that the middle finger has no abduction since it is considered the "reference finger" in the sagittal plane of the hand.

An optical motion capture system (Phase Space, San Leandro, CA – USA) with 19 active markers was used to collect a large number of static grasp positions (see figure 4). Subject AT (M,26) performed all the grasps of the 57 imagined objects described in [Santello et al., 1998]; these data were acquired twice to define a set of 114 *a priori* data.

The disposition of the markers on the hand refers to [Fu and Santello, 2010]. We used four markers for the thumb and three markers for each of the rest of the fingers. Three markers were also placed on the dorsal surface of the palm to define a local reference system S_H (see figure 3). The positions of the markers, which were sampled

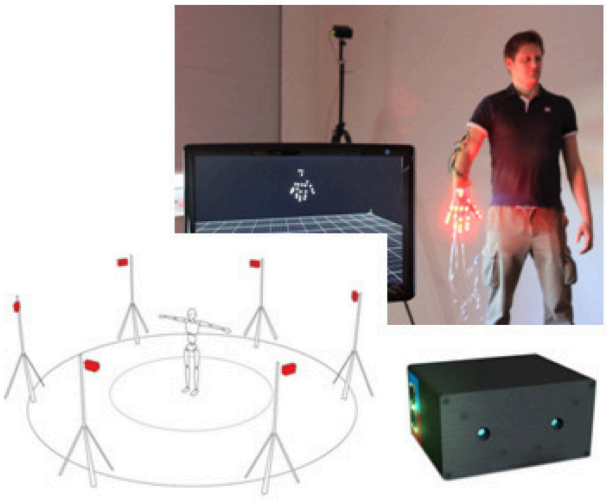


Figure 4: Experimental setup for hand pose acquisition with Phase Space system.

at 480 Hz, are given referring to the global reference system S_{MC} (which is directly defined during the calibration of the acquisition system).

An additional set of $N_p = 54$ grasp poses was performed by subject LC (26,M). Subject was asked to perform some imagined grasped object poses contained in the *a priori* data set and also some new postures which identify basic grasping configurations of the hand (e.g. precision and power grasp) [1]. 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 latter set of poses will be referred hereinafter as *validation set*, since these poses can be assumed to represent accurate reference angular values for hand pose configurations, 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 [Zhang et al., 2003]. The validation set was then used to simulate optimal gloves. According to the number of measures, we considered from the postures in this set only the joint values resulting from the optimization procedure, assuming to select them in a linear weighted combination (continuous measurement matrix), individually (discrete measurement matrix) or combining discrete and continuous measurements (hybrid measurement matrix). Since all the DoFs of the postures in the validation set are known, we compare the reconstructed hand configurations obtained from simulated optimal glove measures with the reference ones. The same procedure is used for hand pose reconstruction achieved starting from non optimal measures provided by H_s, which is described in the following section.

[1] These hand posture acquisitions are available at http://handcorpus.org/

6 Application of Optimal Glove Design Techniques

In this section we analyze the main features of hand pose sensing devices, whose design is determined on the basis of the optimal procedures proposed in previous sections. First, based on the covariance matrix of the *a priori* data set (cf. Section 5), we give examples of optimal distributions of sensors in terms of continuous, hybrid and discrete measurement matrix in case of noise-free measures. Furthermore, we characterize how the information available by measurement process increases with the minimization of the squared Frobenius norm of the *a posteriori* covariance matrix as well as with an increasing number of measures.

Second, we show how the optimal design improves the estimation algorithms proposed in the companion paper [Bianchi et al., 2012b] and briefly described in Section 2 of this manuscript. In [Bianchi et al., 2012b], we tested the proposed estimations algorithms using discrete matrix H_s, which models a lumped, uncoupled sensing device, providing the individual measures of five metacarpal joints, i.e. TM, IM, MM, RM and LM. In this section, we compare the hand posture reconstruction obtained by H_s with the one obtained by using the optimal matrix H_d^* with the same number of measurements. Improvements on the estimation of hand postures in case of both noise–free and noisy measures are shown. In case of noisy measures, an additional random noise was artificially added on each measure. A zero–mean, Gaussian noise with standard deviation 0.122 rad (7°) was chosen based on data about common technologies and tools used to measure hand joint positions [Simone et al., 2007], thus obtaining a noise covariance matrix $R \approx \mathrm{diag}(0.0149)$.

Finally, for the sake of completeness, we compare hand pose estimation performance achievable with the optimal design of corresponding to continuous, hybrid and discrete measurement matrix, considering the same number of measures. The purpose of this comparison is to analyze and discuss benefits of each type of optimal sensor distribution w.r.t. the other ones.

6.1 Sensor Distributions Based on Continuous, Discrete and Hybrid Measurement Matrices

As shown in Section 3, the optimal choice H_c^* of the measurement matrix $H \in \mathbb{R}^{m\times n}$ is represented by the first m principal components (synergies) of the *a priori* covariance matrix P_o. Figure 5 shows the optimal sensor distribution related to a continuous measurement matrix which furnishes a number of measures up to three. In this case, each measure is given as $y_k = \sum_{i=1}^{n} w_i x_i$ with $k = 1,\ldots,m$. In figure 5 a representation, in grayscale, of the absolute value of w_i, is reported: the greater is the absolute coefficient w_i, the darker is the color of that joint.

The optimal measurement matrix H_d^*, for a number of noise–free measures m ranging from 1 to 14, is reported in table 1. Notice that, H_d^* does not have an incremental behavior, especially in case of few measures. In other words, the set of DoFs which have to be chosen in case of m measures does not necessarily contain all the set of DoFs chosen for $m-1$ measures. Moreover, noise randomness can slightly change which DoFs have to be measured compared with the noise–free case.

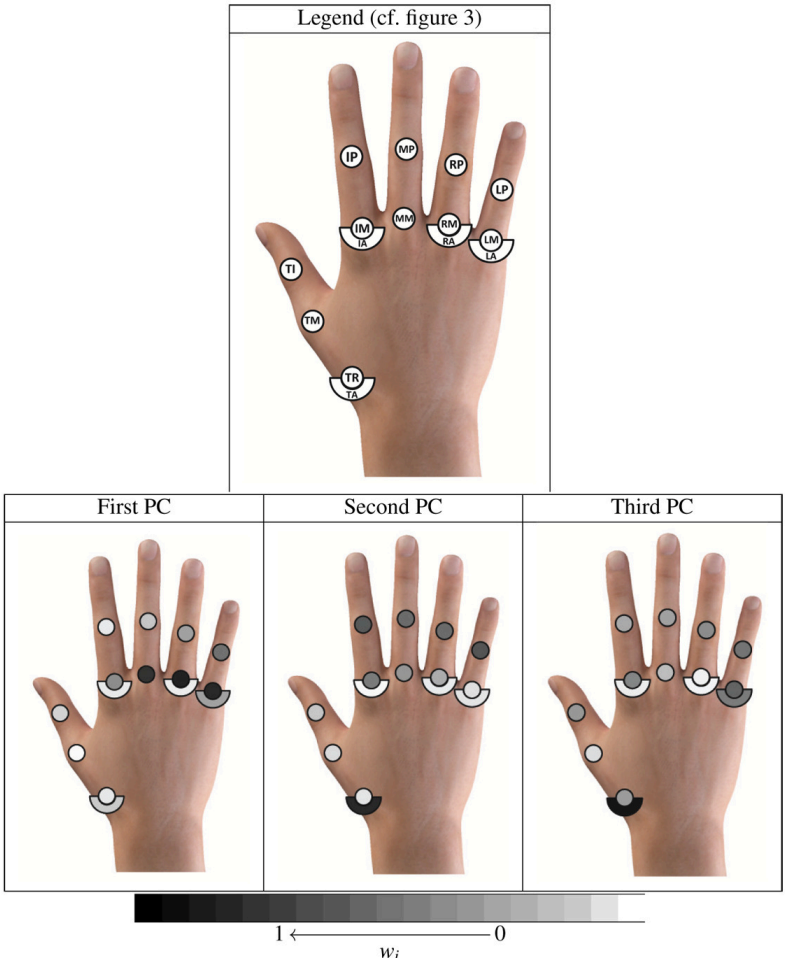


Figure 5: Optimal sensing distribution according to the first three PCs of P_o, which correspond to the rows of the optimal continuous measurement matrix up to three measures. The greater is the absolute coefficient w_i of the joint angle in the PC and hence in each measure $y_k = \sum_{i=1}^{n} w_i x_i$ with $k = 1,2,3$, the darker (in grayscale) is the color of that joint. For representation purposes only, we consider the absolute value of the coefficient of the *i-th* joint in the PC to be normalized w.r.t. the maximum absolute value of the coefficients that can be achieved all over the joints.

For the sake of completeness, table 2 shows also an example of optimal measurement matrix $H^*_{c,d}$ for a number of noise–free measures m ranging from 1 to 14, where only the first measure is continuous (i.e. $m_c = 1$). Of course, table 2 shows only the selected DoFs of each discrete measure. However, results obtained with the gradient

m	TA	TR	TM	TI	IA	IM	IP	MM	MP	RA	RM	RP	LA	LM	LP	V_1
1											X					$7.12 \cdot 10^{-2}$
2							X				X					$2.39 \cdot 10^{-2}$
3	X										X	X				$6.59 \cdot 10^{-3}$
4	X							X				X		X		$3.30 \cdot 10^{-3}$
5	X							X				X	X	X		$1.90 \cdot 10^{-3}$
6	X			X				X				X	X	X		$5.32 \cdot 10^{-4}$
7	X			X				X	X				X	X	X	$2.92 \cdot 10^{-4}$
8	X			X	X			X	X				X	X	X	$1.98 \cdot 10^{-4}$
9	X			X	X		X	X				X	X	X	X	$1.30 \cdot 10^{-4}$
10	X			X	X	X	X	X				X	X	X	X	$6.86 \cdot 10^{-5}$
11	X	X		X	X	X	X	X				X	X	X	X	$2.70 \cdot 10^{-5}$
12	X	X		X	X	X	X	X	X			X	X	X	X	$1.40 \cdot 10^{-5}$
13	X	X	X	X	X	X	X	X	X			X	X	X	X	$3.39 \cdot 10^{-6}$
14	X	X	X	X	X	X	X	X	X	X		X	X	X	X	$1.32 \cdot 10^{-6}$

Table 1: Optimal measured DoFs for H_d^* with an increasing number of noise–free measures m (cf. figure 3).

$m = 1 + m_d$	TA	TR	TM	TI	IA	IM	IP	MM	MP	RA	RM	RP	LA	LM	LP	V_1
2											X					$1.81 \cdot 10^{-2}$
3											X	X				$5.49 \cdot 10^{-3}$
4								X				X		X		$2.68 \cdot 10^{-3}$
5								X				X	X	X		$1.27 \cdot 10^{-3}$
6	X			X				X					X	X		$3.64 \cdot 10^{-4}$
7	X			X				X					X	X	X	$2.46 \cdot 10^{-4}$
8	X			X	X			X					X	X	X	$1.58 \cdot 10^{-4}$
9	X			X	X	X		X					X	X	X	$9.10 \cdot 10^{-5}$
10	X	X		X	X		X					X	X	X	X	$4.95 \cdot 10^{-5}$
11	X	X		X	X	X	X	X					X	X	X	$2.09 \cdot 10^{-5}$
12	X	X	X	X	X	X	X	X					X	X	X	$8.70 \cdot 10^{-6}$
13	X	X	X	X	X	X	X		X			X	X	X	X	$3.02 \cdot 10^{-6}$
14	X	X	X	X	X	X	X		X	X		X	X	X	X	$4.59 \cdot 10^{-7}$

Table 2: Optimal measured DoFs for the discrete part of the optimal hybrid matrix $H_{c,d}^*$ with only one continuous measure ($m_c = 1$) and with an increasing number of noise–free measures m (cf. figure 3).

flow (18) proposed in Section 3.3 show that the elements of the continuous part of $H_{c,d}^*$ corresponding to DoFs measured by the discrete part, tend to be zero. For instance, if $m = 2$ with $m_c = 1$, matrix $H_{c,d}^*$ furnishes two measures: the discrete one gives the angle of RM DoF (see table 2), while the continuous part is a linear combination of all joints angle but RM one. Indeed, the RM angle is perfectly known, i.e. we have the maximum information available on that DoF and hence, the continuous measure has to furnish information only on the other DoFs.

Let us quantify the information made available by the optimal measurement pro-

cess through the square Frobenius norm of the *a posteriori* covariance matrix (V_1). Figure 6 shows the values of V_1 for increasing number m of noise–free measures. The best performance is obtained by the continuous measurement matrix, as expected and the observed information is always the greatest one. Indeed, principal components are considered the optimal measures for the representation of points in the best fitting lower dimensional subspace [Rao, 1964]. The optimal hybrid measurement matrix we

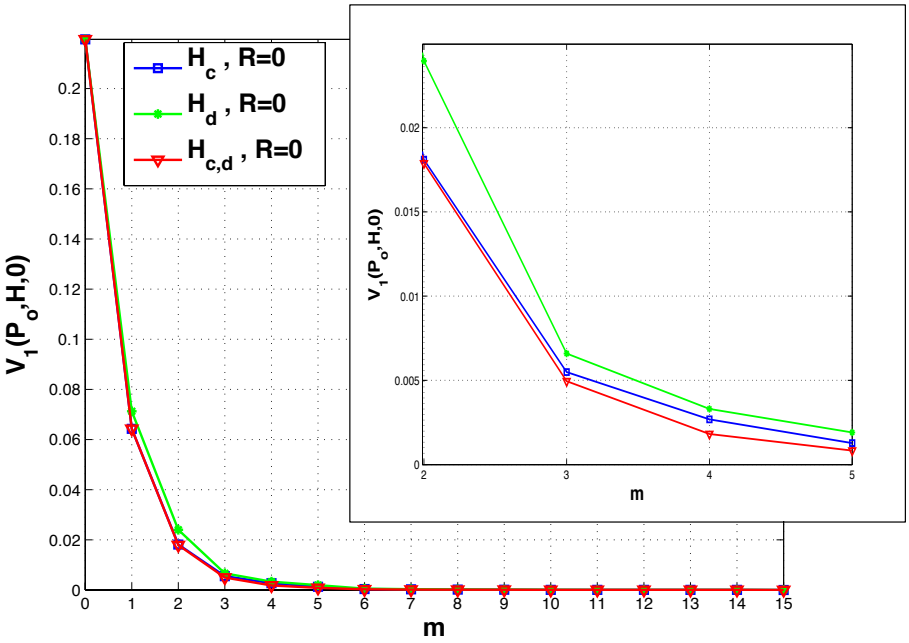


Figure 6: Squared Frobenius norm of the *a posteriori* covariance matrix with noise–free measures in case of H_c^*, H_d^* and $H_{c,d}^*$ ($m_c = 1$) with an increasing number of noise-free measures. A zoomed detail of the graph is shown for $m = 2,3,4,5$ measures.

consider for this analysis has only a continuous measure, the first one. For this reason, in figure 6, the same value for V_1 is observed for both hybrid and continuous case when only one measure ($m = 1$) is considered, hence leading to same performance. Then, increasing the number of discrete measures, performance of hybrid matrix tends to be close to the discrete one (see the zoomed detail in figure 6), even if observed information for hybrid measurement matrix is always greater than the observed information for the discrete measurement matrix. Notice that V_1 values decrease with the number of measures, tending to be zero (cf. figure 6) for all three cases. This fact is trivial because increasing the measurements, the uncertainty on the measured variables is reduced. When all the measured information is available V_1 assumes zero value with perfectly accurate measures in all three cases. In case of noisy measures, V_1 values decrease with the number of measures tending to a value which is larger, depending on the level of noise.

For noise–free measures, we analyze how much V_1 reduces with the number of measurements w.r.t. the value it assumes for zero measures ($P_p \equiv P_0$). In terms of measurement process, i.e. from the *observability* viewpoint, a reduced number of measures coinciding with the first three principal components enable for $\simeq 97\%$ reduction of the squared Frobenius norm of the *a posteriori* covariance matrix. An analogous result can be found also under the *controllability* point of view. In [Santello et al., 1998] authors state that three postural synergies are crucial in grasp pre–shaping since they take into account for $\simeq 90\%$ of pose variability in grasping tasks. The above reported result seems logic considering the duality between observability and controllability.

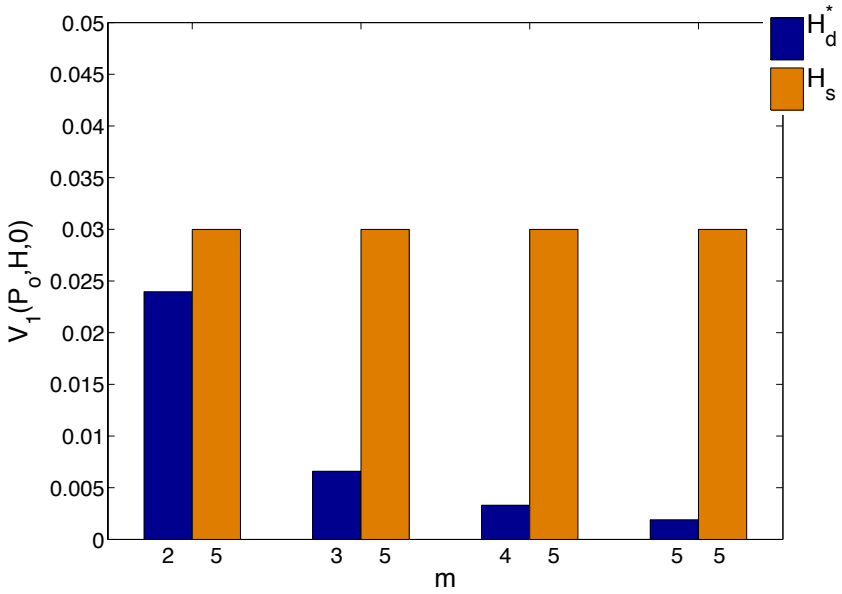


Figure 7: Squared Frobenius norm for the *a posteriori* covariance matrix of H_s which models a sensing device, providing the individual measures of five ($m = 5$) metacarpal joints, i.e. TM, IM, MM, RM and LM, and H_d^* with $m = 2, 3, 4, 5$ measures, in case of noise–free measures (see table 2).

Finally, let us consider figure 7 and 8, where squared Frobenius norm for the *a posteriori* covariance matrix of H_s with $m = 5$ measures, and H_d^* with $m = 2,3,4,5$ measures, in case of noise-free and noisy measures, respectively, is reported. Notice that, V_1 is significantly smaller in the optimal case, even when a reduced number of measures is considered, leading to a better estimation performance also with an inferior number of measured DoFs w.r.t. H_s. For sake of space, in Subsection 7.1, we report only the estimation comparison when both H_s and H_d^* provide five measures.

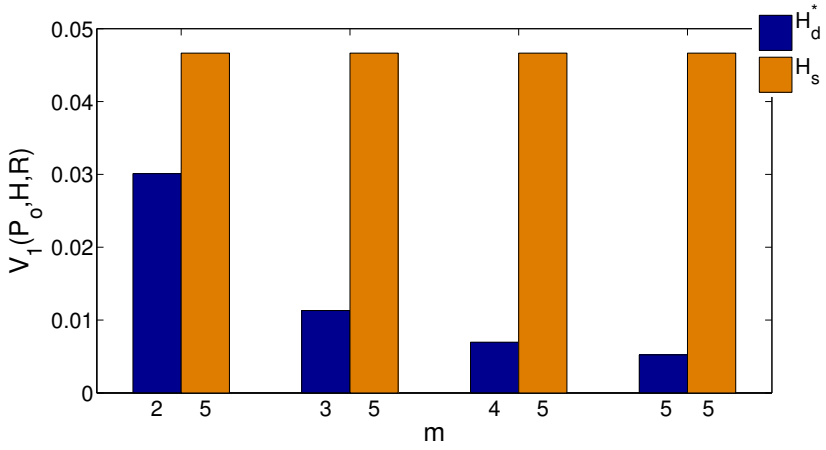


Figure 8: Squared Frobenius norm for the *a posteriori* covariance matrix of H_s which models a sensing device, providing the individual measures of five ($m = 5$) metacarpal joints, i.e. TM, IM, MM, RM and LM with $m = 5$ measures, and H_d^* with $m = 2, 3, 4, 5$ measures, in case of noisy measures.

7 Experimental Results

In this section, we compare the hand posture reconstructions, considering $m = 5$ measures provided by matrix H_s and by optimal matrix H_d^* in case of both noise–free and noisy measures. Second, we analyze HPR performance for the tree sensing typologies of optimal sensor distributions corresponding to the three optimal measurement matrices.

7.1 H_s vs. H_d^* with $m = 5$

In figure 9, sensor locations related to matrix H_s and H_d^* (with noise–free and noisy measures), are represented. For the comparison, measures are provided by grasp data acquired with the optical tracking system as described in Section 5 (see also [Bianchi et al., 2012b] for more details), where degrees of freedom to be measured are chosen on the basis of optimization procedure outcomes, while the entire pose is recorded to produce accurate reference posture.

In figure 10, some reconstructed poses with MVE algorithm are reported by using both H_s and H_d^*, with and without additional noise. Under a qualitative point of view, what is noticeable is that reconstructed poses are not far from the real ones for both measurement matrices, even if the pose error $e_i = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i|$ is smaller for the optimal case. Moreover, it may happen that some poses are estimated in a better manner using H_s and vice versa, even if, as we will show in next subsections, from a statistical point of view, H_d^* provides best average performance. Indeed, MVE methods are thought to minimize error statistics rather than worst–case sensing errors related to particular poses [Bicchi and Canepa, 1994].

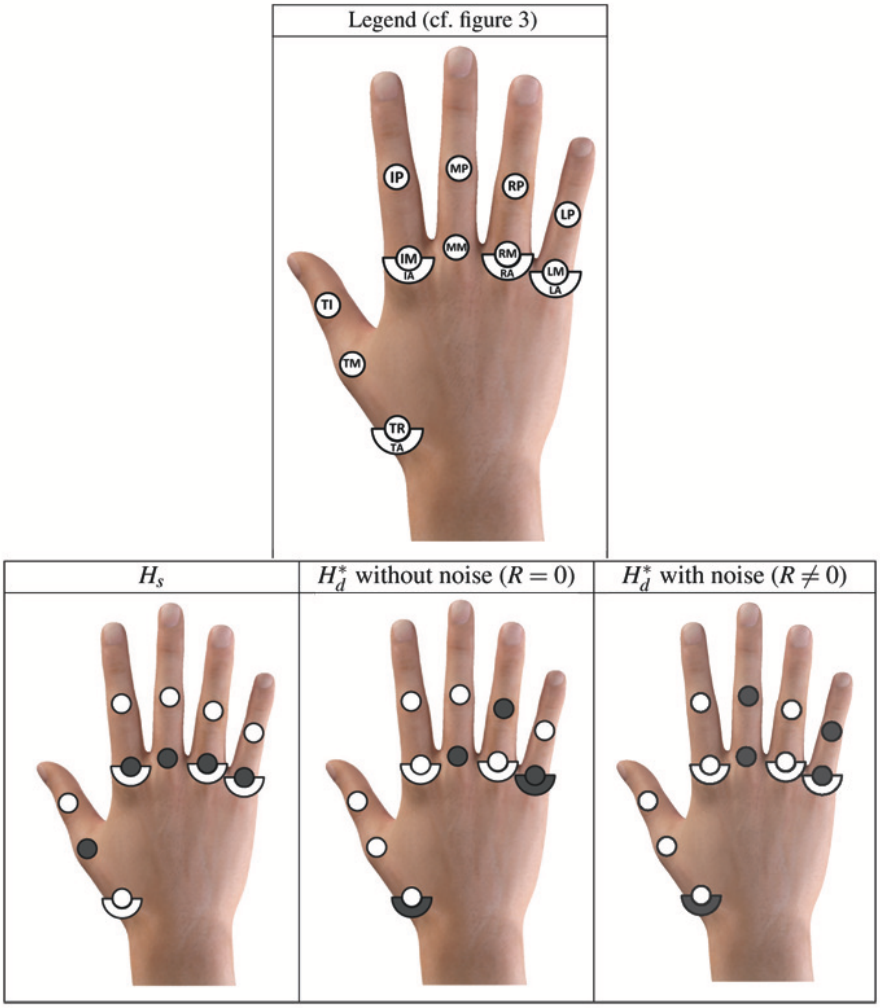


Figure 9: Measured joints (highlighted in color) for matrix H_s, and H_d^*, with and without noise (cf. figure 3).

In the following, as evaluation indices, the average pose estimation error and average estimation error for each estimated DoF are used. Maximum errors are also reported. These errors as well as statistical tools are chosen according to the ones considered in [Bianchi et al., 2012b]. Statistical differences between estimated pose and joint errors obtained with different glove designs are computed by using classic tools, after having tested for normality and homogeneity of variances assumption on samples (through Lilliefors' composite goodness–of–fit test and Levene's test, respectively). Standard two–tailed t–test (hereinafter referred as T_{eq}) is used in case of both the as-

sumptions are met, a modified two–tailed T–test is exploited (Behrens–Fisher problem, using Satterthwaite's approximation for the effective degrees of freedom, hereinafter referred as T_{neq}) when variance assumption is not verified and finally a non parametric test is adopted for the comparison (Mann–Whitney U–test, hereinafter referred as U) when normality hypothesis fails. Significance level of 5% is assumed and p-values less than 10^{-4} are posed equal to zero.

7.1.1 Noise–Free Measures

In terms of average absolute estimation pose errors ([°]), performance obtained with H_d^* is always better than the one exhibited by H_s (3.67±0.93 vs. 6.69±2.38). Moreover, H_d^* exhibits smaller maximum error than the one achieved with H_s (i.e. 8.25° for H_d^* vs. 13.18° for H_s). Statistical difference between results from H_s and H_d^* is found ($p \simeq 0$, T_{neq}). In table 3 average absolute estimation errors with their corresponding standard deviations for each DoF are reported. For the estimated DoFs, performance with H_d^* is always better or not statistically different from the one referred to H_s. Maximum estimation errors underline cases where H_s furnishes smaller values and vice versa, since they strictly depend on particular poses.

7.1.2 Noisy Measures

In case of noise, performance in terms of average absolute estimation pose errors ([°]) obtained with H_d^* is better than the one exhibited by H_s (5.96±1.42 vs. 8.18±2.70). Moreover, maximum pose error with H_d^* is the smallest (9.30° vs. 15.35° observed with H_s). Statistical difference between results from H_s and H_d^* is found (p=0.001, T_{neq}). In table 4 average absolute estimation error with standard deviations are reported for each DoF. Also in this case, for the estimated DoFs, performance with H_d^* is always better or not statistically different from the one referred to H_s. Maximum estimation errors with H_d^* are usually inferior to the ones obtained with H_s. The MVE method seems to give more accurate estimates for certain DoFs (e.g. TA, IA, RA, LA, for H_s and MP, RA, RM, LP for H_d^*) with noise than without noise even if in the latter case, the average absolute pose estimation error is the smallest one.

It should be noticed that the MVE method guarantees that the mean squared norm of the joint error vector (i.e. the Mean Squared Error, $MSE = \frac{1}{N}\sum_{i=1}^{N} \|\hat{x} - x\|^2$, where N represents the number of predictions) is minimized, but not necessarily the value of each single component. Same applies with noise: indeed, some particular joints have a lower error with noise than without noise, yet the overall error norm (across all joints) is always higher if noise is present. Indeed, if the joint angles of both the estimations and the reference poses are in degrees, the MSE with the optimal measurement matrix H_d^* (H_s) is 522 (1583) in case of noise–free measures and 948 (1992) in case of noise, hence it increases with noise. The fact that noise happens to reduce the error in some joints is a statistically insignificant case, that has occurred with the validation set reported in the paper. Using other validation sets, we have obtained estimates where noise reduces the error of different individual joints, or increases all components: however, as theory predicts, the overall mean square error vector is always increased by noise. The whole argument rests on the fact that the validation sets are samples from the same

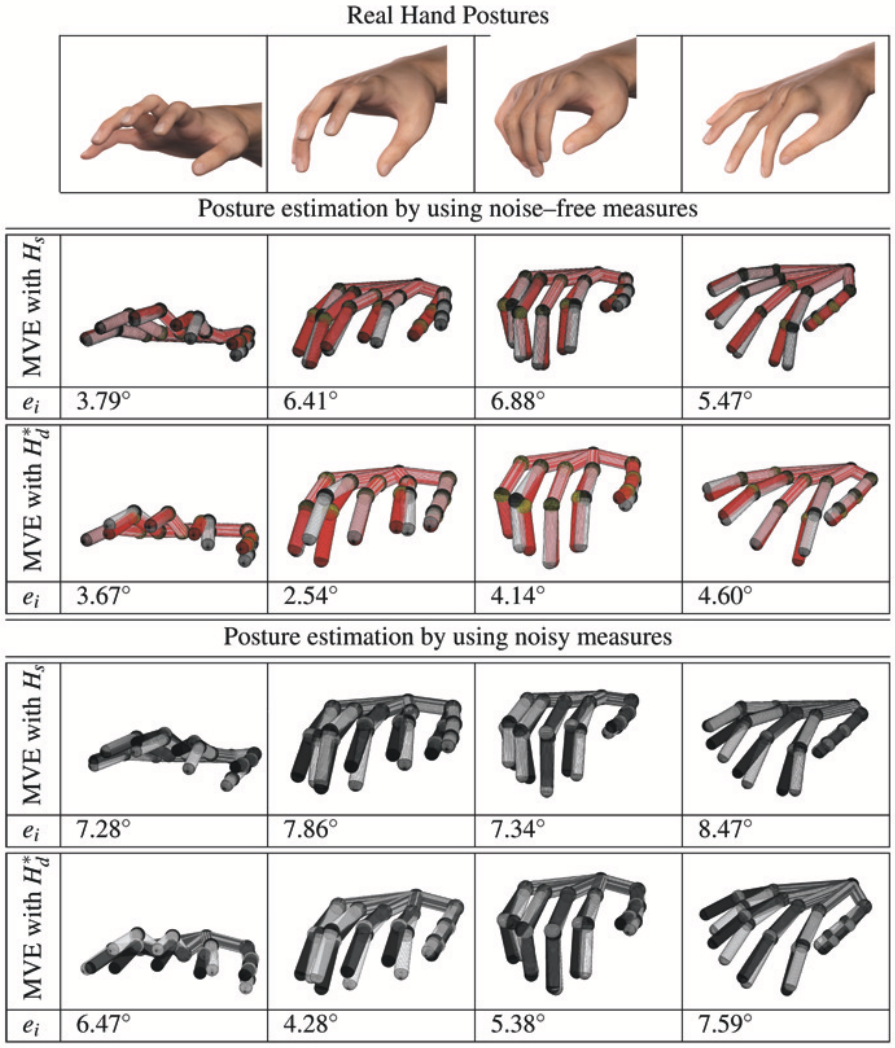
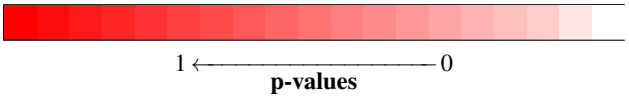


Figure 10: Hand pose reconstructions MVE algorithm by using matrix H_s which allows to measure TM, IM, MM, RM and LM and matrix H_d^* which allows to measure TA, MM, RP, LA and LM (cf. figure 3). In color the real hand posture whereas in white the estimated one. The pose error is given by $e_i = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{x}_i|$.

distribution, of which the a priori set is assumed to provide a statistically accurate description.

DoF	Mean Error [°]		H_s vs. H_d^*	Max Error [°]	
	H_s	H_d^*	**p-values**	H_s	H_d^*
TA⊗	10.74±8.45	0	0	31.65	0
TR	7.16±4.54	6.84±4.75	**0.72** ⋄	19.50	20.13
TM○	0	2.17±2.21	0	0	13.04
TI	4.81±3.68	5.33±4.16	**0.64**	19.68	15.15
IA	11.96±5.33	10.55±5.65	**0.14**	26.35	26.15
IM○	0	4.02±3.43	0	0	16.01
IP	13.26±7.06	5.42±6.44	0	27.46	43.86
MM○⊗	0	0	–	0	0
MP	12.35±7.75	4.90±2.91	0 ‡	29.94	9.91
RA	3.45±2.43	3.82±2.94	**0.73**	9.51	12.68
RM○	0	6.68±3.68	0	0	16.01
RP⊗	13.40±9.65	0	0	39.33	0
LA⊗	11.33±5.87	0	0	24.47	0
LM○⊗	0	0	–	0	0
LP	11.94±9.52	6.27±3.97	0.0002	36.58	16.63



○ indicates a DoF measured with H_s
⊗ indicates a DoF measured with H_d^*

Table 3: Average estimation errors and standard deviations for each DoF [○] for the simulated acquisition considering H_s and H_d^* both with five noise–free measures. Maximum errors are also reported as well as p-values from the evaluation of DoF estimation errors between H_s and H_d^*. ⋄ 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. Symbol "–" is used for those DoFs which are measured by both H_s and H_d^*.

7.2 Estimation Performance Comparison for Optimal Sensor Designs

As previously shown, the observed information quantified through V_1 (squared Frobenius norm of the *a posteriori* covariance matrix) is greatest for continuous case, while hybrid case provides better performance than the discrete one. Here, we analyze how these differences affect the reconstruction pose accuracy. To accomplish this goal, we consider as an example the case of three noise–free measures ($m = 3$). For the hybrid case, only the first measure is continuous (i.e. $m_c = 1$ and $m_d = 2$).

DoF	Mean Error [°]		H_s vs. H_d^*	Max Error [°]	
	H_s	H_d^*	**p-values**	H_s	H_d^*
TA⊗	6.7±5.62	4.87±3.57	**0.19**	23.35	15.93
TR	7.65±5.57	7.54±5.00	**0.91** ⋄	27.46	22.73
TM○	2.81±1.75	2.63±1.90	**0.61** ⋄	7.2	8.78
TI	6.08±4.63	5.42±4.74	**0.32**	19.6	19.10
IA	10.74±5.6	11.52±5.81	**0.32**	27.31	28.46
IM○	4.15±3.17	6.91±5.00	0.003	11.66	21.49
IP	14.61±7.93	6.61±6.01	0	31.85	38.07
MM○⊗	4.59±3.08	4.71±3.19	**0.77**	11.43	15.72
MP⊗	13.71±8.07	4.08±2.98	0 ‡	37.61	13.71
RA	3.12±2.37	3.28±2.45	**0.71**	9.18	9.37
RM○	4.03±3.07	6.30±4.72	0.01 ‡	12.94	12.91
RP	16.78±11.07	6.89±3.82	0 ‡	50.66	16.34
LA	8.97±5.11	9.86±5.45	**0.38** ⋄	20.86	21.48
LM○⊗	3.82±3.05	4.82±4.30	**0.44**	11.33	14.26
LP⊗	14.64±9.68	3.94±2.95	0	48.61	11.03

1 ←——————— 0

p-values

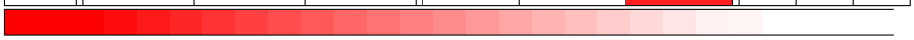
○ indicates a DoF measured with H_s

⊗ indicates a DoF measured with H_d^*

Table 4: Average estimation errors and standard deviations for each DoF [○] for the simulated acquisition considering H_s and H_d^* both with five noisy measures. Maximum errors are also reported as well as p-values from the evaluation of DoF estimation errors between H_s and H_d^*. ⋄ 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. Symbol "–" is used for those DoFs which are measured by both H_s and H_d^*.

To enable for a correct comparison, we compute average absolute estimation pose errors ([°]) only on estimated DoFs, disregarding each time those joints whose values are perfectly known since individually measured (i.e. RM, RP for both $H_{c,d}^*$ and H_d^* and TA for H_d^*). Performance obtained with H_c^* is always better than the one exhibited by $H_{c,d}^*$ and H_d^* (5.09±1.6 vs. 6.16±1.64 and 6.61±1.89, respectively). Statistical difference between H_c^* and H_d^* and between H_c^* and $H_{c,d}^*$ is found ($p \simeq 0, U$). No statistical difference is observed between $H_{c,d}^*$ and H_d^* ($p = 0.19, T_eq$). Moreover, H_c^* exhibits the smallest maximum error: 9.95 vs. 11.30 and 11.73 for $H_{c,d}^*$ and H_d^*, respectively.

DoF	Mean Error [°]			p-values			Max Error [°]		
	H_c^*	$H_{c,d}^*$	H_d^*	H_c^* **vs.** H_d^*	H_c^* **vs.** $H_{c,d}^*$	$H_{c,d}^*$ **vs.** H_d^*	H_c^*	$H_{c,d}^*$	H_d^*
TA$\circ$	4.44±4.85	4.60±4.48	0	0	**0.69**	0	17.19	15.83	0
TR	5.64±3.96	5.79±3.86	7.28±4.68	**0.08**	**0.84**$\diamond$	**0.13**	17.11	16.65	18.86
TM	2.69±1.89	2.40±1.90	2.07±1.88	0.04	**0.34**	**0.35**	8.61	8.77	10.39
TI	3.72±2.80	4.02±3.03	4.36±3.34	**0.28**$\diamond$	**0.70**	**0.73**	13.95	12.83	14.18
IA	12.04±6.00	12.13±6.23	12.05±6.40	**0.83**	**1**	**0.87**	29.32	30.24	30.56
IM	4.43±4.09	6.84±5.64	9.81±7.35	0	0.02	**0.05**	16.79	21.01	25.22
IP	4.39±4.64	5.17±6.63	5.09±6.53	**0.77**	**0.64**	**0.98**	30.81	45.43	44.63
MM	3.68±2.70	7.36±4.92	8.76±5.40	0	0	**0.07**	11.04	21.40	23.34
MP	3.66±2.57	4.63±2.87	4.54±2.94	**0.1**$\diamond$	**0.07**$\diamond$	**0.87**$\diamond$	14.26	10.80	10.88
RA	3.40±2.41	3.42±2.56	3.49±2.54	**0.91**	**0.96**	**0.91**	9.44	10.00	9.88
RM$\circ\otimes$	5.41±2.58	0	0	0	0	–	11.85	0	0
RP$\circ\otimes$	4.62±3.76	0	0	0	0	–	16.29	0	0
LA	10.12±4.41	12.26±4.59	10.96±6.52	**0.43**‡	0.01$\diamond$	**0.24**‡	19.74	20.89	25.85
LM	3.84±2.96	5.26±3.44	4.79±4.72	**0.55**	0.01	**0.07**	10.76	18.10	19.64
LP	3.79±3.02	6.18±4.32	6.09±4.24	0.004	0.003	**0.88**	13.53	17.29	16.94



$\circ$ indicates a DoF measured with H_d^*

$\otimes$ indicates a DoF measured with the discrete part of $H_{c,d}^*$.

Table 5: Average estimation errors and standard deviations for each DoF [$\circ$] for the simulated estimation considering H_c^*, $H_{c,d}^*$ and H_d^*, with three noise-free measures. Maximum errors are also reported as well as p-values from the evaluation of DoF estimation errors between the continuous, hybrid and discrete design. $\diamond$ 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. Symbol "–" is used for those DoFs which are measured by both H_d^* and the discrete part of $H_{c,d}^*$.

In table 5 average absolute estimation errors with their corresponding standard deviations for each DoF are reported. For the estimated DoFs, performance with H_c^* is always better or not statistically different from the one referred to $H_{c,d}^*$ or H_d^*. Only for TM joint performance with H_d^* is better than H_c^* even if p–value is close to the significance level. Finally, no statistical difference is observed between $H_{c,d}^*$ and H_d^*. Same considerations still work for maximum estimation errors.

Conclusions we can drawn are that, even if the values for V_1 (see figure 6) differ not so much for the three cases, the continuous measurement matrix provides the best pose reconstruction performance. Of course, the trade–off between performance, cost and ease to design should be taken into account to determine whether and how well a particular design suits a particular application.

8 Conclusions

In this paper, optimal design of sensing glove has been proposed on the basis of the minimization of the *a posteriori* covariance matrix as it results from the estimation procedure described in [Bianchi et al., 2012b]. Optimal solution are described for the continuous, discrete and hybrid case.

In the continuous measurement matrix case, basically corresponding to distributed sensing devices, optimal measures are individuated by principal components of the *a priori* covariance matrix, thus suggesting the importance of postural synergies not only for hand control.

The reconstruction performance obtained by combining the estimation technique proposed in [Bianchi et al., 2012b] and the optimal design proposed in this paper is significantly improved if compared with non–optimal measure case. Therefore, [Bianchi et al., 2012b] and this paper provide a complete procedure to enhance the performance and for a more effective development of sensorization systems for robotic hands and active touch sensing systems. These techniques can be useful in a wide range of applications, ranging from virtual reality to tele-robotics and rehabilitation. Moreover, by optimizing the number and location of sensors the production costs can be further reduced without loss of performance, thus increasing device diffusion.

Further work will be dedicated to the physical implementation of the here proposed optimal designs as well as defining optimal calibration strategies.

Acknowledgments

Authors gratefully acknowledge Marco Santello and Lucia Pallottino for the inspiring discussion and useful suggestions.

References

[Bianchi et al., 2012a] Bianchi, M., Salaris, P., and Bicchi, A. (2012a). Synergy–based optimal design of hand pose sensing. In *IEEE-RSJ International Conference on Intelligent Robots and Systems*.

[Bianchi et al., 2012b] Bianchi, M., Salaris, P., and Bicchi, A. (2012b). Synergy-based hand pose sensing: Performance enhancement. *The International Journal of Robotics Research*. Submitted.

[Bicchi, 1992] Bicchi, A. (1992). A criterion for optimal design of multiaxis force sensors. *Journal of Robotics and Autonomous Systems*, 10(4):269–286.

[Bicchi and Canepa, 1994] Bicchi, A. and Canepa, G. (1994). Optimal design of multivariate sensors. *Measurement Science and Technology (Institute of Physics Journal "E")*, 5:319–332.

[Chaloner and Verdinelli, 1995] Chaloner, K. and Verdinelli, I. (1995). Bayesian experimental design: A review. *Statistical Science*, 10:273–304.

[Chang et al., 2007] Chang, L. Y., Pollard, N. S., Mitchell, T. M., and Xing, E. P. (2007). Feature selection for grasp recognition from optical markers. In *Intelligent Robots and Systems, 2007. IROS 2007. IEEE/RSJ International Conference on*, pages 2944–2950.

[Diamantaras and Hornik, 1993] Diamantaras, K. and Hornik, K. (1993). Noisy principal component analysis. *Measurement'93*, pages 25 – 33.

[Dipietro et al., 2008] Dipietro, L., Sabatini, A., and Dario, P. (2008). A survey of glove-based systems and their applications. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 38(4):461 –482.

[Edin and Abbs, 1991] Edin, B. B. and Abbs, J. H. (1991). Finger movement responses of cutaneous mechanoreceptors in the dorsal skin of the human hand. *Journal of neurophysiology*, 65(3):657–670.

[Edmison et al., 2002] Edmison, J., Jones, M., Nakad, Z., and Martin, T. (2002). Using piezoelectric materials for wearable electronic textiles. In *Wearable Computers, 2002. (ISWC 2002). Proceedings. Sixth International Symposium on*, pages 41 – 48.

[Fu and Santello, 2010] Fu, Q. and Santello, M. (2010). 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.

[Gabiccini et al., 2011] Gabiccini, M., Bicchi, A., Prattichizzo, D., and Malvezzi, M. (2011). On the role of hand synergies in the optimal choice of grasping forces. *Auton. Robots*, 31(2-3):235–252.

[Ghosh and Rao, 1996] Ghosh, S. and Rao, C. R. (1996). Review of optimal bayes designs. In *Design and Analysis of Experiments*, volume 13 of *Handbook of Statistics*, pages 1099 – 1147. Elsevier.

[Helmicki et al., 1991] Helmicki, A. J., Jacobson, C. A., and Nett, C. N. (1991). Control oriented system identification: a worst-case/deterministic approach in h^{∞}. *Automatic Control, IEEE Transactions on*, 36(10):1163 –1176.

[Pukelsheim, 2006] Pukelsheim, F. (2006). *Optimal Design of Experiments (Classics in Applied Mathematics) (Classics in Applied Mathematics, 50)*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

[Rao, 1964] Rao, C. R. (1964). The use and interpretation of principal component analysis in applied research. *The Indian journal of statistic*, 26:329 – 358.

[Rosen, 1960] Rosen, J. B. (1960). The gradient projection method for nonlinear programming. part i. linear constraints. *Journal of the Society for Industrial and Applied Mathematics*, 8(1):181 – 217.

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

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

[Simone et al., 2007] Simone, L. K., Sundarrajan, N., Luo, X., Jia, Y., and Kamper, D. G. (2007). A low cost instrumented glove for extended monitoring and functional hand assessment. *Journal of Neuroscience Methods*, 160(2):335–348.

[Sturman and Zeltzer, 1993] Sturman, D. J. and Zeltzer, D. (1993). A design method for "whole-hand" human-computer interaction. *ACM Trans. Inf. Syst.*, 11(3):219–238.

[Tempo, 1988] Tempo, R. (1988). Robust estimation and filtering in the presence of bounded noise. *Automatic Control, IEEE Transactions on*, 33(9):864 –867.

[Tognetti et al., 2006] Tognetti, A., Carbonaro, N., Zupone, G., and De Rossi, D. (2006). 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.

[Zavlanos and Pappas, 2008] Zavlanos, M. M. and Pappas, G. J. (2008). A dynamical systems approach to weighted graph matching. *Automatica*, 44(11):2817 – 2824.

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

[Zimmerman, 1988] Zimmerman, T. G. (1982). Optical flex sensor. *Patent US 4542291, Sep. 29.*

A Appendix

This appendix is devoted to the derivation of the gradient equation given in proposition 1.

Proof of Proposition 1 The Frobenius norm of a matrix $A \in \mathbb{R}^{n\times n}$ is given as

$$\|A\|_F = \sqrt{\mathrm{tr}(A^T A)} = \sqrt{\sum_{i=1}^{n} \sigma_i^2},$$

and hence,

$$\|P_o - P_o H^T (HP_o H^T + R)^{-1} HP_o\|_F^2 = \mathrm{tr}(P_p^T P_p) \tag{19}$$

where $P_p = P_o - P_o H^T (HP_o H^T + R)^{-1} HP_o$. To find the gradient flow, we need to compute

$$\frac{\partial \operatorname{tr}(P_p^T P_p)}{\partial H} = \operatorname{tr}\left(\frac{\partial (P_p^T P_p)}{\partial H}\right) = \operatorname{tr}\left(\frac{\partial P_p^T}{\partial H} P_p + P_p^T \frac{\partial P_p}{\partial H}\right) = \\ = \operatorname{tr}\left(\frac{\partial P_p^T}{\partial H} P_p\right) + \operatorname{tr}\left(P_p^T \frac{\partial P_p}{\partial H}\right) = 2\operatorname{tr}\left(P_p^T \frac{\partial P_p}{\partial H}\right), \tag{20}$$

as $\partial(\mathbf{XY}) = (\partial \mathbf{X})\mathbf{Y} + \mathbf{X}(\partial \mathbf{Y})$ and $\operatorname{tr}(A^T) = \operatorname{tr}(A)$. Moreover, from differentiation rules of expressions w.r.t. a matrix $\mathbf{X}$, we have $\partial \mathbf{X}^{-1} = -\mathbf{X}^{-1}(\partial \mathbf{X})\mathbf{X}^{-1}$ and hence, assuming $\Sigma(H) = (HP_o H^T + R)^{-1}$, we obtain

$$\frac{\partial P_p}{\partial \mathbf{H}} = -P_o \left[(\partial H)^T \Sigma(H) H + H^T \left(\frac{\partial \Sigma(H)}{\partial H} H + \Sigma(H)\, \partial H\right)\right] P_o = \\ = -P_o \left[(\partial H)^T \Sigma(H) H - H^T \left(\Sigma(H) \left(\partial H P_o H^T + \right.\right.\right. \\ \left.\left.\left. + H P_o (\partial H)^T\right) \Sigma(H) H + \Sigma(H)\, \partial H\right)\right] P_o . \tag{21}$$

Substituting (21) in (20) and by using a well note trace property ($\operatorname{tr}(A+B) = \operatorname{tr}(A) + \operatorname{tr}(B)$) we obtain

$$\frac{\partial \operatorname{tr}(P_p^T P_p)}{\partial H} = 2\left[-\operatorname{tr}(P_p^T P_o (\partial H)^T \Sigma(H) H P_o) + \operatorname{tr}(P_p^T P_o H^T \Sigma(H) \partial H P_o H^T \Sigma(H) H P_o) + \right. \\ \left. + \operatorname{tr}(P_p^T P_o H^T \Sigma(H) H P_o (\partial H)^T \Sigma(H) H P_o) - \operatorname{tr}(P_p^T P_o H^T \Sigma(H) \partial H P_o)\right] . \tag{22}$$

As $\operatorname{tr}(AB) = \operatorname{tr}(BA)$, we obtain

$$\frac{\partial \operatorname{tr}(P_p^T P_p)}{\partial H} = 2\left[-\operatorname{tr}((\partial H)^T \Sigma(H) H P_o P_p^T P_o) + \operatorname{tr}(P_o H^T \Sigma(H) H P_o P_p^T P_o H^T \Sigma(H) \partial H) + \right. \\ \left. + \operatorname{tr}((\partial H)^T \Sigma(H) H P_o P_p^T P_o H^T \Sigma(H) H P_o) - \operatorname{tr}(P_o P_p^T P_o H^T \Sigma(H) \partial H)\right] \tag{23}$$

and as $\operatorname{tr}(A^T) = \operatorname{tr}(A)$ we have

$$\frac{\partial \operatorname{tr}(P_p^T P_p)}{\partial H} = 2\left[-\operatorname{tr}(P_o^T P_p P_o^T H^T \Sigma(H)^T \partial H) + \operatorname{tr}(P_o H^T \Sigma(H) H P_o P_p^T P_o H^T \Sigma(H) \partial H) + \right. \\ \left. + \operatorname{tr}(P_o^T H^T \Sigma(H)^T H P_o^T P_p P_o^T H^T \Sigma(H)^T \partial H) - \operatorname{tr}(P_o P_p^T P_o H^T \Sigma(H) \partial H)\right] , \tag{24}$$

whence,

$$\frac{\partial \operatorname{tr}(P_p^T P_p)}{\partial H} = 2\left[-P_o^T P_p P_o^T H^T \Sigma(H)^T + P_o H^T \Sigma(H) H P_o P_p^T P_o H^T \Sigma(H) + \right. \\ \left. + P_o^T H^T \Sigma(H)^T H P_o^T P_p P_o^T H^T \Sigma(H)^T - P_o P_p^T P_o H^T \Sigma(H)\right] = \\ = 2\left[(P_o H^T \Sigma(H) H - I) P_o P_p^T P_o H^T \Sigma(H) + (P_o^T H^T \Sigma(H)^T H - I) P_o^T P_p P_o^T H^T \Sigma(H)^T\right] . \tag{25}$$

Matrices P_p, P_o and $\Sigma(H)$ are symmetric, and hence, for this particular case we obtain

$$\frac{\partial \operatorname{tr}(P_p^T P_p)}{\partial H} = -4 \left[P_p^2 P_o H^T \Sigma(H) \right]^T , \tag{26}$$

with $\Sigma(H) = (H P_o H^T + R)^{-1}$.