Synergy–Based Hand Pose Sensing: Optimal Glove Design*

Matteo Bianchi,† Paolo Salaris,† Antonio Bicchi†‡

Abstract

In this paper we study the problem of improving human hand pose sensing device performance by exploiting the knowledge on how humans most frequently use their hands in grasping tasks. 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 statistical *a priori* information. 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 continuous case, whereas individual sensing elements in the glove measure a linear combination of joint angles, the discrete case, whereas each measure corresponds to a single joint angle, and the most general hybrid case, whereas both continuous and discrete sensing elements are available. The objective is to provide, for given *a priori* information and fixed number of measurements, the optimal design minimizing in average the reconstruction error. Solutions relying on the geometrical synergy definition as well as gradient flow-based techniques are provided. Simulations of reconstruction performance show the effectiveness of the proposed optimal design.

1 Introduction

This paper investigates the problem of estimating the posture of human hands using sensing devices, and how to improve their performance based on the knowledge on how humans most frequently use their hands. Similarly to the companion paper [Bianchi et al., 2012b], this work is motivated by studies on the human hand in grasping tasks [Santello et al., 1998] suggesting hand posture representations of increasing complexity ("synergies"), which allow to reduce the number of Degrees of Freedom (DoFs) to be used according to the desired level of approximation. In [Bianchi et al., 2012b], we analyzed the role of the *a priori* information for pose hand reconstructions by using given sensing devices, and showed that acceptable reconstruction results can be obtained even in presence of insufficient and inaccurate sensing data.

*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.

†The Interdept. Research Center "Enrico Piaggio", University of Pisa, via Diotisalvi 2, 56100 Pisa, Italy. m.bianchi,p.salaris,bicchi@centropiaggio.unipi.it

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

Figure 1: Examples of continuous and discrete 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 sensors.

In this work, we extend the analysis to consider the optimal design of sensing "gloves", i.e. devices for hand pose reconstruction based on measurements of few geometric features of the hand. The problem we consider is to find the distribution of a number of sensing elements of limited accuracy so as to provide, together with the *a priori* information, the optimal design which minimizes in probability the reconstruction error. The problem becomes particularly relevant when limits on the production costs of sensing gloves introduce constraints limiting both the number and the quality of sensors. In these cases, a careful design of sensor distribution is instrumental to obtain good performance.

Optimal experimental design represents a challenging, widely discussed topic in literature [Pukelsheim, 2006]. Among all optimal design criteria, Bayesian methods are ideally suited to contribute to experimental design and error statistics minimization, when some information is available prior to experimentation (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 a linear Gaussian hypothesis is not fulfilled and/or when the designer's primary concern is to minimize worst-case sensing 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 and Canepa, 1994, Bicchi, 1992].

However, most of these approaches refer to cases with a number of basic sensors which is redundant or at least equal to the number of variables to be estimated. Moreover, no previous example of application to the peculiar problem of exploiting the a priori psychophysical information on the structure of human hand embodiment for under–sensorized gloves has ever been reported. In [Sturman and Zeltzer, 1993] an investigation of "whole-hand" interfaces for the control of complex tasks is presented, along with the description, 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] authors explored how to methodically select a minimal set of hand pose features from optical marker data for grasp recognition. The objective is to determine marker locations on the hand surface that is appropriate for grasp classification of hand poses. All the aforementioned approaches rely on experimental or qualitative observations: from actual sensor data, locations that provide the largest and most useful information on the system are chosen.

In this paper, we investigate in depth the problem of obtaining the optimal distribution of sensors minimizing in probability the reconstruction error of hand poses. We adopt a classical Bayesian approach to minimize the *a posteriori* covariance matrix norm and hence, to maximize the information on the real hand posture available by the glove measurement.

The *a posteriori* covariance matrix, $P_p = P_o - P_o H^T (H P_o H^T + R)^{-1} H P_o$, which directly depends on the sensor design through the measurement matrix H, its noise covariance R, and on the *a priori* information P_o, represents a measure of the amount of information that the observable variables carry about the unknown pose parameters. Here we explore the role of the measurement matrix H on the estimation procedure, providing the optimal design of a sensing device able to get the maximum amount of the information on the actual hand posture.

We first consider the *continuous* sensing case, where individual sensing elements in the glove can be designed so as to measure a linear combination of joint angles. An example of this type is the sensorized glove developed in [Tognetti et al., 2006] (cf. figure 1(a)), or the 5DT Data Glove (5DT Inc., Irvine, CA - USA). Other devices, such as e.g. the Cyberglove (CyberGlove System LLC, San Jose, CA - USA), or the Humanglove (Humanware s.r.l., Pisa, Italy) shown in figure 1(b), provide instead *discrete* sensing, i.e. each sensor provides a measure of a single joint angle.

Finally, for the sake of generality, we consider the optimal design of hybrid sensing devices, which combine continuous and discrete sensors. It is interesting to note that human hands represent, to some extent, examples of such hybrid sensing: among the cutaneous mechanoreceptors in the dorsal skin of the hand that were demonstrated to be involved in the responses to finger movements, [Edin and Abbs, 1991] includes both Fast Adapting (mainly FAI) afferents, with localized response to movements about one or, at most, two nearby joints; and Slow Adapting (SA) afferents, whose discharge rate is influenced by several joints interactively. Note also that FA units are found primarily close to joints, while SA units are more uniformly distributed.

To validate our technique we consider hand posture reconstruction using a limited number of measurements from a set of grasp postures acquired with an optical tracking system, providing accurate reference poses. Experiments and statistical analyses demonstrate the improvement of the estimation techniques proposed in [Bianchi et al., 2012b] by using the optimal design proposed in this paper.

2 Problem Definition

For reader's convenience we summarize here the definitions and results of [Bianchi et al., 2012b] used in the following. Let us assume a n degrees of freedom 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 $\nu \in \mathbb{R}^m$ is a vector of measurement noise. 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 in 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 that an observable variable carries about unknown parameters. In this paper we will 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

We first analyze the case that individual sensing elements in the glove can be designed to measure a linear combination of joint angles (continuous sensing devices), and provide, for given *a priori* information and fixed number of measurements, the optimal design, minimizing in average the reconstruction error. We then consider the case where each measure provided by the glove corresponds to a single joint angle (discrete sensing devices). For these types of gloves we determine which joint should be individually measured in order to optimize the design. Finally, we will consider the case that both continuous and discrete sensor elements are used in the achieve sensing devices, defining a procedure to obtain the optimal hybrid sensing glove design.

In the ideal case of noiseless 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 fewer 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 $V_1(P_o, H, R)$: $\mathbb{R}^{m\times n} \rightarrow \mathbb{R}$ 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*

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

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 analyze separately the design of continuous, discrete and hybrid sensing devices.

3.1 Continuous Sensing Design

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 describe the optimal continuous sensing design both in a numerical and analytical way. 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 noiseless 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 renormalization 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

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

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. 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 sensor design), 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 $H_i P_o H_j^T = 0,\ i \neq j$ 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 into 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

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

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

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

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

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

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 $M \in \mathbb{R}^m$ a full rank matrix, suggests that there exists a subspace in $\mathbb{R}^n$ where the optimum is achieved. Indeed, gradients become zero when rows of matrix H are any linear combination of a subset of m principal components of the *a priori* covariance matrix. Unfortunately, this does not happen in case of noisy measures and gradients become zero only for a particular matrix H which depends also on the principal components of the noise covariance matrix.

3.2 Discrete Sensing Design

When each measure y_j, $j = 1,\dots,m$ provided by the glove corresponds to a single joint angle x_i, $i = 1,\dots,n$, 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 $V_1(P_o,H,R) = \|P_o - P_oH^T(HP_oH^T + R)^{-1}HP_o\|_F^2$, find the optimal measurement matrix*

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

In this case a closed-form solution is not available. Nonetheless, as the model hand adopted 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 next section, a more general approach to computing 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 design of hybrid sensors.

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

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

where $A \circ B$ denotes the *Hadamard* or elementwise 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 respectively minimize their cost functions. 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,*

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

where $k \in [0,\,1]$ is a positive constant 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 Sensing Design

In this section we analyze the sensing device with both continuous and discrete sensors. Up to rearranging 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 continuous sensing elements, whereas $H_d \in \mathscr{P}^{m_d\times n}$ describes the m_d single-joint measurements, with $m_c + m_d = m$. Neither the closed-form solution valid for continuous sensing design, nor the exhaustion method used for discrete measurements are applicable in the hybrid case. Therefore, to optimally design hybrid pose sensing systems, we will recur to gradient-based iterative optimization algorithms.

We first consider the case that noise is negligible ($R \approx 0$). By combining the continuous and discrete gradient flows, previously defined in (11) and (16), respectively, we obtain

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

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

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

On the basis of Theorem 1, the gradient flow defined in (18) converges toward a hybrid sensing system (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, 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).

This problem can be circumvented by constraining the solution in the sub-set $\mathscr{H}_{c,d} = \{H_{c,d} : H_c H_c^T = I_{m_c}\}$. A solution for this problem can be obtain by the following gradient flow

$$\begin{aligned} \dot{H}_{c,d} = 4\,(1-k)\, W_{c,d}\ \left[P_p^2 P_o H_{c,d}^T \Sigma(H_{c,d})\right]^T + \\ + k\,\bar{H}_d \left[(\bar{H}_d \circ \bar{H}_d)^T \bar{H}_d - \bar{H}_d^T (\bar{H}_d \circ \bar{H}_d)\right], \end{aligned} \tag{19}$$

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

$$W_{c,d} = \begin{bmatrix} I_{m_c} - H_c (H_c^T H_c)^{-1} H_c^T & 0_{m_c \times m_d} \\ 0_{m_d \times m_c} & I_{m_d \times m_d} \end{bmatrix}$$

for the projection matrix, and starting from any initial guess matrix $H_{c,d} \in \mathscr{H}_{c,d}$, the gradient flow remains in the sub-set $\mathscr{H}_{c,d}$, and converges to a (local) minimum for the problem. Also in this case, multistart strategies can circumvent the problem of local minima.

4 Results

In this section we will describe 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 increasing the number of measures, leading to better estimation performance.

First, based on the *a priori* covariance matrix obtained with the *a priori* data set described in section 3 of [Bianchi et al., 2012b], we will show the optimal distribution of sensors on the hand in case of continuous and discrete sensing devices. We will also show that, although the number of measures used with the optimal matrix is less than the five measures available by matrix H_s (cf. [Bianchi et al., 2012b]), the hand posture information achievable with the optimal measurement matrix H_d^* related to a discrete sensing device, is greater, i.e. $V_1(H_d^*) < V_1(H_s)$, leading to a better hand pose estimation performance.

Second, we will compare the hand posture reconstruction obtained by means of matrix H_s with the one obtained by using the optimal matrix H_d^* with the same number of measures. Additional random normal noise ν with standard deviation of $7°$ on each measure is also considered to evaluate the performance in case of noisy measures.

4.1 Continuous, Discrete and Hybrid Sensing Distribution

As shown in section 3, in case of continuous sensing design, 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 2 shows the hand sensor distribution related to each synergy.

In case of discrete sensing, the optimal measurement matrix H_d^*, related to a discrete sensing device, 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 behaviour, 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 4 shows the values of the square Frobenius norm of the *a posteriori* covariance matrix for increasing number m of noise–free measures. The best performance

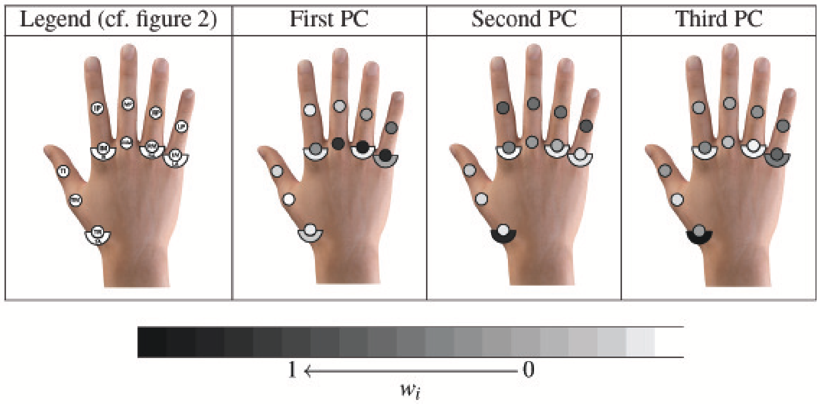
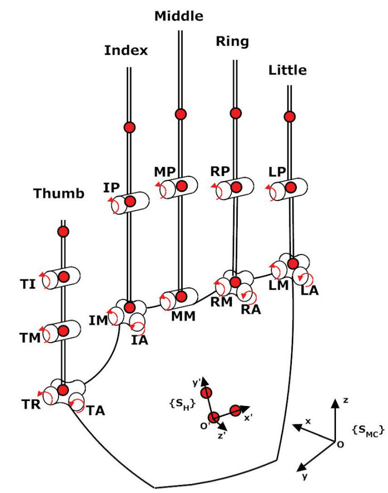


Figure 2: Optimal continuous sensing distribution for the first PCs of P_o. The greater is the absolute coefficient w_i of the joint angle in the PC, the darker is the color of that joint. We assume the coefficient of the *i-th* joint in the PC to be normalized w.r.t. the maximum absolute value of the coefficients that can be achieved all over the joints.

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 increasing number of noise–free measures m (cf. figure 3).

is obtained by the continuous sensing design, as aspected. Indeed, principal components are considered the optimal measures for the representation of points in the best fitting lower dimensional subspace [Rao, 1964]. The hybrid performance is better than the discrete one, thus representing a trade-off between the quality of estimation of the continuous sensing design and feasibility and costs of the discrete one. Moreover, V_1 values decrease with the number of measures, tending to be zero (cf. figure 4). This fact is trivial because increasing the measurements, the uncertainty on the measured



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. Markers are reported as red spheres.

variables is reduced. When all the measured information is available V_1 assumes zero value with perfectly accurate measures. 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, if we analyze how much V_1 reduces with the number of measurements w.r.t. the value it assumes for one measure, reduction percentage with three measured DoFs is greater than 80%. This result suggests that with only three measurements, the optimal matrix can furnish more than 80% of uncertainty reduction. This is equivalent to say that a reduced number of measurements is sufficient to guarantee a good hand posture estimation. In [Santello et al., 1998] and [Gabiccini and Bicchi, 2010], under the *controllability* point of view, authors state that three postural synergies are crucial in grasp pre-shaping as well as in grasping force optimization since they take into account for more than 80% of variance in grasp poses. Here, the same result can be obtained in terms of measurement process, i.e. from the *observability* point of view: a reduced number of measures coinciding with the first three principal components enable for more than 80% reduction of the squared Frobenius norm of the *a posteriori* covariance matrix.

The above reported result seems logic considering the duality between observability and controllability. Moreover, under an engineering point of view, it is reasonable that those actuators which are used the most being also the most monitored and hence the most sensor endowed.

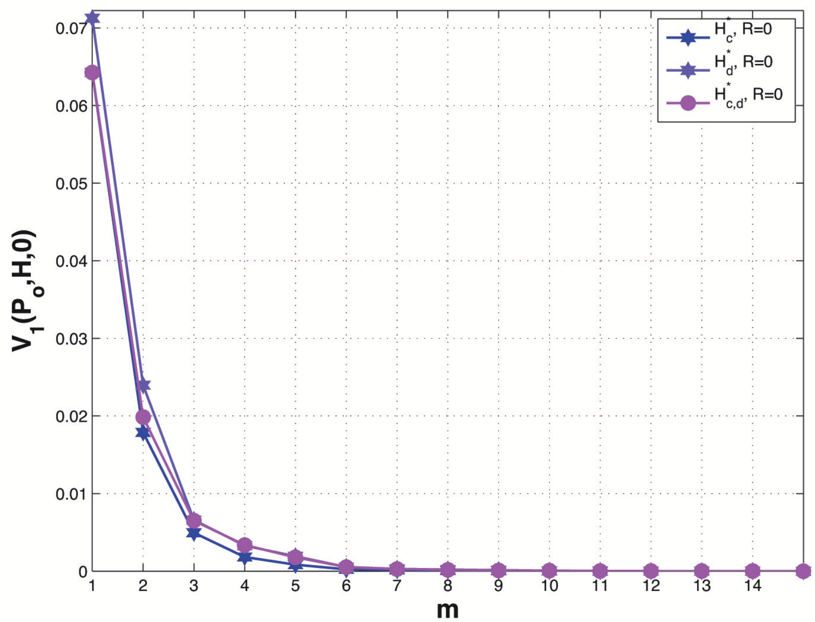


Figure 4: Squared Frobenius norm of the *a posteriori* matrix with noise–free measures in case of H_c^*, H_d^* and $H_{c,d}^*$ ($m_c = 1$).

5 Discussion

In this section, we will compare the hand posture reconstruction obtained by applying the hand pose reconstruction techniques described in [Bianchi et al., 2012b] to $m = 5$ measures provided by matrix H_s and by optimal matrix H_d^*.

5.1 Estimation Results with Optimal Discrete Sensing Devices

Measures are provided by grasp data acquired with the optical tracking system as in [Bianchi et al., 2012b], 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 5 sensor locations related to matrix H_s and H_d^* are represented. In order to compare reconstruction performance achieved with H_s and H_d^* we use as evaluation indices the average pose estimation error and average estimation error for each estimated DoF. Maximum errors are also reported. These errors as well as statistical tools are chosen according to the ones considered in [Bianchi et al., 2012b], where it is possible to find a complete description of the here adopted. Both noise-free and noisy measures are analyzed.

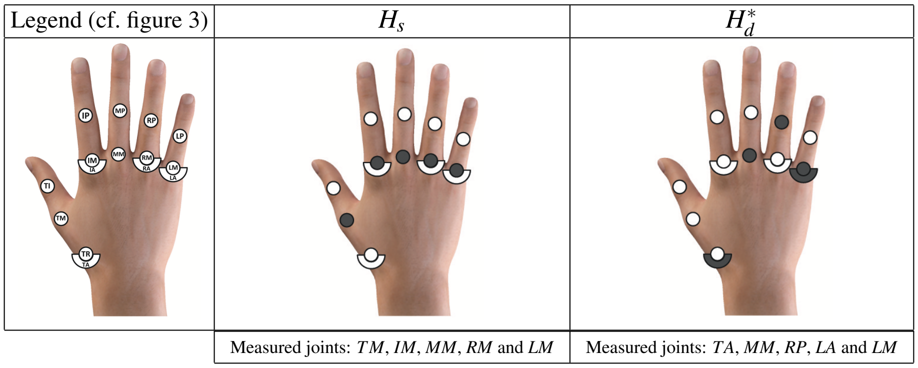


Figure 5: Discrete sensing distributions for matrix H_s, on the left, and H_d^*, on the right (cf. figure 3). The measured joints are highlighted in color.

5.1.1 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 differences between results from H_s and H_d^* are found (p $\simeq$ 0, T_{neq}). In table 2 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 peculiar poses; however, results from the two matrices are globally comparable.

In figure 7, 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 measures is reported. Notice that squared Frobenius norm is significantly smaller in the optimal case, even when a reduced number of measures is considered.

5.1.2 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^* are found (p=0.001, T_{neq}).

In table 3 average absolute estimation error with standard deviations are reported for each DoF. 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.

Figure 8 shows the 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. Also

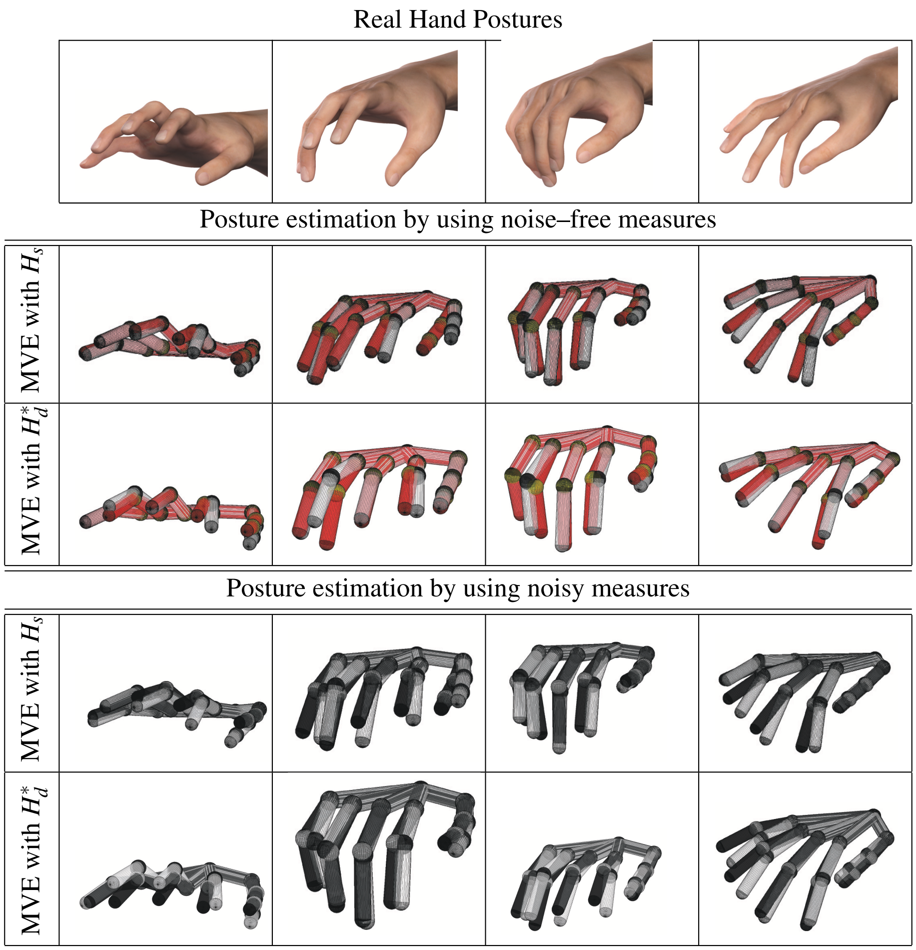


Figure 6: Hand pose reconstructions MVE algorithm by using matrix H_s which allows to measure TM, IM, MM, RM and LM and matrix H_d^* which allows to measure TA, MM, RP, LA and LM (cf. figure 3). In color the real hand posture whereas in white the estimated one.

in this situation, squared Frobenius norm is significantly smaller in the optimal case, even if a reduced number of measures is considered, thus suggesting that an optimal design leading to error statistics minimization can be achieved using optimal matrix with an inferior number of measured DoFs w.r.t. H_s. Notice that in this case, squared Frobenius norm values are larger than the corresponding ones obtained in absence of noise, as expected.

Finally, in figure 6 some reconstructed poses with MVE algorithm are reported by using both H_s and H_d^* measurement matrix, with and without additional noise. Under a qualitative point of view, what is noticeable is that reconstructed poses are not far from

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

1 ⟵ 0
p-values

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

Table 2: Average estimation errors and standard deviation 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^*.

the real ones for both measurement matrices. Moreover, it is not surprising that some poses seem to be estimated in a better manner using H_s and vice versa, even if from the previously described statistical results H_d^* provides best average performance. Indeed, MVE methods are thought to minimize error statistics rather than worst-case sensing errors related to peculiar poses [Bicchi and Canepa, 1994].

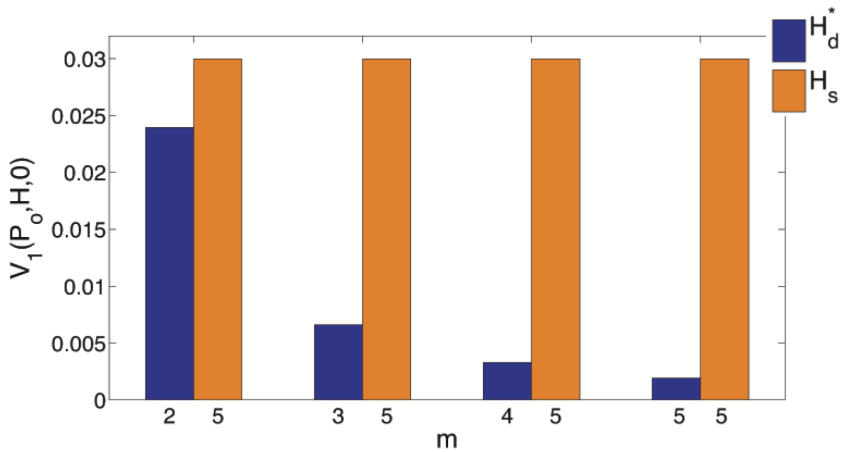


Figure 7: Squared Frobenius norm for the *a posteriori* covariance matrix of H_s with $m = 5$ measures, and H_d^* with $m = 2, 3, 4, 5$ measures, in case of noise–free measures.

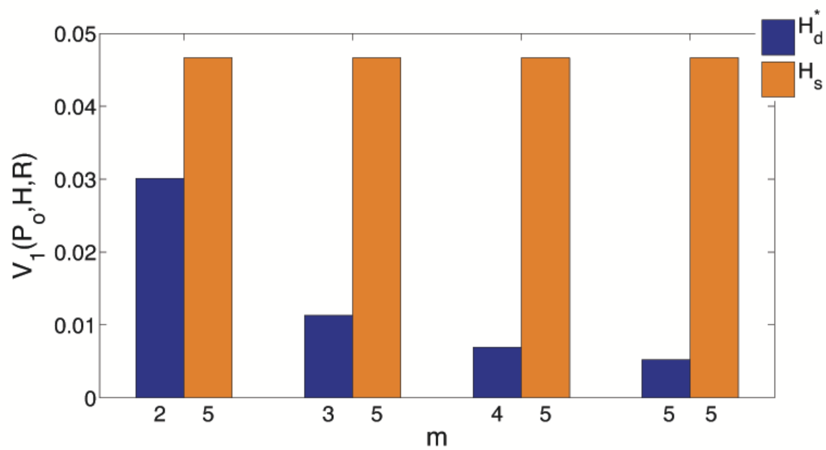


Figure 8: Squared Frobenius norm for the *a posteriori* covariance matrix of H_s with $m = 5$ measures, and H_d^* with $m = 2, 3, 4, 5$ measures, in case of noisy measures.

6 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 sensing case, optimal measures are individuated by principal components of the *a priori* covariance matrix, thus suggesting the importance of pos-

DoF	Mean Error [°]		H_s vs. H_d^*	Max Error [°]	
	H_s	H_d^*	**p-values**	H_s	H_d^*
TA$\otimes$	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** $\diamond$	27.46	22.73
TM$\circ$	2.81±1.75	2.63±1.90	**0.61** $\diamond$	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$\circ$	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$\circ\otimes$	4.59±3.08	4.71±3.19	**0.77**	11.43	15.72
MP$\otimes$	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$\circ$	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** $\diamond$	20.86	21.48
LM$\circ\otimes$	3.82±3.05	4.82±4.30	**0.44**	11.33	14.26
LP$\otimes$	14.64±9.68	3.94±2.95	0	48.61	11.03

1 ⟵ 0

p-values

$^\circ$ indicates a DoF measured with H_s

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

Table 3: Average estimation errors and standard deviation for each DoF [$\circ$] 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^*. $\diamond$ 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_s and H_d^*.

tural 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 [Bianchi et al., 2012a] provide a complete procedure to enhance the performance and for a more effective development of both sensorization systems for robotic hands and active touch sensing systems, which can be used 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.

Acknowledgment

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 hand pose sensing: Optimal glove design. *The International Journal of Robotics Research*. Submitted.

[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.

[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.

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

[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$^\infty$. *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.

[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.

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,

$$\|P_o - P_o H^T (H P_o H^T + R)^{-1} H P_o\|_F^2 = \mathrm{tr}(P_p^T P_p) \tag{20}$$

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

$$\frac{\partial\,\mathrm{tr}(P_p^T P_p)}{\partial H} = \mathrm{tr}\left(\frac{\partial(P_p^T P_p)}{\partial H}\right) = \mathrm{tr}\left(\frac{\partial P_p^T}{\partial H}P_p + P_p^T\frac{\partial P_p}{\partial H}\right) =$$
$$= \mathrm{tr}\left(\frac{\partial P_p^T}{\partial H}P_p\right) + \mathrm{tr}\left(P_p^T\frac{\partial P_p}{\partial H}\right) = 2\,\mathrm{tr}\left(P_p^T\frac{\partial P_p}{\partial H}\right), \tag{21}$$

as $\partial(\mathbf{XY}) = (\partial\mathbf{X})\mathbf{Y} + \mathbf{X}(\partial\mathbf{Y})$ and $\mathrm{tr}(A^T) = \mathrm{tr}(A)$. 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 $\Sigma(H) = (HP_oH^T + R)^{-1}$, we obtain

$$\frac{\partial P_p}{\partial\mathbf{H}} = -P_o\left[(\partial H)^T\Sigma(H)H + H^T\left(\frac{\partial\Sigma(H)}{\partial H}H + \Sigma(H)\,\partial H\right)\right]P_o =$$
$$= -P_o\left[(\partial H)^T\Sigma(H)H - H^T\left(\Sigma(H)\left(\partial HP_oH^T + \right.\right.\right.$$
$$\left.\left.\left.+HP_o(\partial H)^T\right)\Sigma(H)H + \Sigma(H)\,\partial H\right)\right]P_o\,. \tag{22}$$

Substituting (22) in (21) and by using a well note trace property ($\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B)$) we obtain

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

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

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

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

$$\frac{\partial\,\mathrm{tr}(P_p^T P_p)}{\partial H} = 2\left[-\mathrm{tr}(P_o^T P_pP_o^T H^T\Sigma(H)^T\partial H) + \mathrm{tr}(P_oH^T\Sigma(H)HP_oP_p^T P_oH^T\Sigma(H)\partial H) + \right.$$
$$\left. + \mathrm{tr}(P_o^T H^T\Sigma(H)^T HP_o^T P_pP_o^T H^T\Sigma(H)^T\partial H) - \mathrm{tr}(P_oP_p^T P_oH^T\Sigma(H)\partial H)\right], \tag{25}$$

whence,

$$\frac{\partial\,\mathrm{tr}(P_p^T P_p)}{\partial H} = 2\left[-P_o^T P_pP_o^T H^T\Sigma(H)^T + P_oH^T\Sigma(H)HP_oP_p^T P_oH^T\Sigma(H) + \right.$$
$$\left. + P_o^T H^T\Sigma(H)^T HP_o^T P_pP_o^T H^T\Sigma(H)^T - P_oP_p^T P_oH^T\Sigma(H)\right] =$$
$$= 2\left[(P_oH^T\Sigma(H)H - I)P_oP_p^T P_oH^T\Sigma(H) + (P_o^T H^T\Sigma(H)^T H - I)P_o^T P_pP_o^T H^T\Sigma(H)^T\right]. \tag{26}$$

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

$$\frac{\partial \operatorname{tr}(P_p^T P_p)}{\partial H} = -4 \left[P_p^2 P_o H^T \Sigma(H)\right]^T , \tag{27}$$

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