

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/282631892>

# Recognizing the grasp intention from human demonstration

ARTICLE *in* ROBOTICS AND AUTONOMOUS SYSTEMS · JULY 2015

Impact Factor: 1.26 · DOI: 10.1016/j.robot.2015.07.006

CITATION

1

READS

27

4 AUTHORS:



Ravin de Souza

5 PUBLICATIONS 2 CITATIONS

SEE PROFILE



S. El-Khoury

École Polytechnique Fédérale de Lausanne

17 PUBLICATIONS 205 CITATIONS

SEE PROFILE



José Santos-Victor

Technical University of Lisbon

233 PUBLICATIONS 3,556 CITATIONS

SEE PROFILE



Aude Billard

École Polytechnique Fédérale de Lausanne

267 PUBLICATIONS 6,354 CITATIONS

SEE PROFILE



# Recognizing the grasp intention from human demonstration



Ravin de Souza<sup>a,b,\*</sup>, Sahar El-Khoury<sup>a</sup>, José Santos-Victor<sup>b</sup>, Aude Billard<sup>a</sup>

<sup>a</sup> Learning Algorithms and Systems Laboratory, École Polytechnique Fédéral de Lausanne, Switzerland

<sup>b</sup> VisLab, Institute for Systems and Robotics, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal

## HIGHLIGHTS

- We propose a general method to interpret human grasp behavior in terms of opposition primitives.
- A primitive model consisting of 41 oppositions for the hand is defined.
- The most likely primitive combination is inferred from tactile and configuration data.
- An 87% recognition rate is achieved over a wide range of human grasp behavior.

## ARTICLE INFO

### Article history:

Received 6 August 2014

Received in revised form

2 July 2015

Accepted 10 July 2015

Available online 17 July 2015

### Keywords:

Grasp representation

Grasp recognition

Learning from demonstration

Tactile sensing

## ABSTRACT

In human grasping, choices are made on the use of hand-parts even before a grasp is realized. The human associates these choices with end-functionality and is confident that the resulting grasp will be able to meet task requirements. We refer to these choices on the use of hand-parts underlying grasp formation as the *grasp intention*. Modeling the grasp intention offers a paradigm whereby decisions underlying grasp formation may be related to the functional properties of the realized grasp in terms of quantities which may be sensed/recognized or controlled. In this paper we model grasp intention as mix of oppositions between hand parts. Sub-parts of the hand acting in opposition to each other are viewed as a basis from which grasps are formed. We compute a set of such possible oppositions and determine the most likely combination from the raw information present in a demonstrated grasp. An intermediate representation of raw sensor data exposes interactions between elementary grasping surfaces. From this, the most likely combination of oppositions is inferred. Grasping experiments with humans show that the proposed approach is robust enough to correctly capture the intention in demonstrated grasps across a wide range of hand functionality.

© 2015 Elsevier B.V. All rights reserved.

## 1. Introduction

Significant research has been conducted into the design and development of anthropomorphic hands to enable robotic systems [1] as well as prosthetic devices [2] to interact with real-world environments. However the additional flexibility associated with these hands comes at the cost of complexity in control due to the increased degrees of freedom. While these hands can be tele-operated to perform complex grasping and manipulation tasks, autonomous control in response to the demands of a task scenario remains a difficult problem. Humans however are extremely

adept at controlling the high degree of freedom hand-wrist-arm musculo-skeletal system and are able to grasp and manipulate objects according to task requirements with a minimum of effort. It is of interest therefore to study human grasping behavior in order to extract underlying principles which may be transferred to robotic or prosthetic devices [3–8]. This paper focuses on an aspect of human grasp behavior which we will refer to as the *grasp intention*.

Even before a grasp is realized, choices have been made regarding parts of the hand that will be engaged and the manner of their application against object surfaces. These choices stem from a perception of task demands and are therefore related to functionality that is brought to the grasp in terms of generating and controlling force, torque and motion. This is evident from the four task scenarios shown in Fig. 1. Tasks requiring dexterity (turning a dial, writing), make use of the finger tips which open up degrees of freedom and bring into play required manipulability for in-hand motion. Also, greater sensitivity associated with finger

\* Corresponding author at: Learning Algorithms and Systems Laboratory, École Polytechnique Fédéral de Lausanne, Switzerland.

E-mail addresses: [ravin.desouza@epfl.ch](mailto:ravin.desouza@epfl.ch) (R. de Souza), [sahar.elkhoury@epfl.ch](mailto:sahar.elkhoury@epfl.ch) (S. El-Khoury), [jav@isr.ist.utl.pt](mailto:jav@isr.ist.utl.pt) (J. Santos-Victor), [aude.billard@epfl.ch](mailto:aude.billard@epfl.ch) (A. Billard).



**Fig. 1.** Four task scenarios: turning a dial, writing, opening a tightly closed bottle-cap, screw-driving show the different use of hand-parts to provide grasp functionality as demanded by a task.

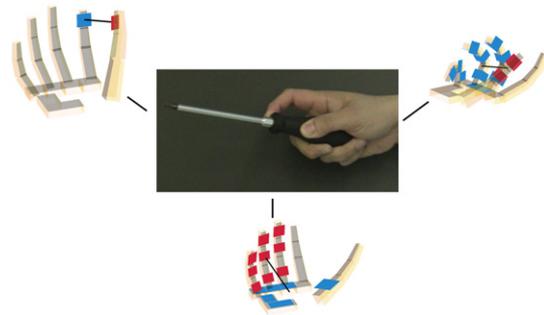
tips is essential for controlling the manipulation [9]. In contrast, tasks requiring power (opening a tight bottle cap, screw-driving) make use of finger surfaces and the palm. Use of these hand-parts is directly related to transmission of torque and motion generated by the wrist–arm system.

We refer to these choices on the use of hand-parts underlying grasp formation as the *grasp intention*. Grasp intention guides the realization of a stable grasp, determining outwardly visible posture and also the grasp function or the particular quality in which force, torque and motion can be delivered to the grasped object. Thus, modeling grasp intention provides a way to relate the decisions underlying grasp formation to end-functionality. This has applications in constructing an appropriate grasp for a task. From a learning from demonstration perspective, we are better positioned to decide on the more important aspects of a grasp to transfer to the robotic platform under consideration.

Is it possible to characterize grasp intention in some general way and recognize this from a grasp demonstration? Grasp taxonomies such as [3,10], based on studies of human grasp behavior, have been proposed as a means to capture grasp ability. The authors in [11–13] attempt to recognize a taxonomy category from human demonstration using cues such as visual features of the grasp or joint angles from a data glove. In [14,15] tactile information is incorporated as well. While it is useful to identify a taxonomy category, key information is lacking on how to recreate the grasp or adapt it to a different object while preserving underlying functional roles of the fingers involved. For example if the object is perturbed or used in a task context, are all hand surfaces equally important for applying pressure or are some more important than others. Similarly, if the properties of the object change how can we purposefully change the hand configuration and object contacts made while remaining confident that the essential meaning of the grasp is preserved. Heuristics have to be designed on a case by case basis to encode the meaning of each grasp. A more general approach defines a set of grasp components from which a wide range of grasps may be constructed. The problem is then identifying and prioritizing the appropriate set of components present in a grasp demonstration.

In this paper we adopt the hypothesis of Iberall et al. in [4] that opposition between hand-parts, while engaging the hand in a well-defined manner, is also correlated with the end-function to be delivered on a grasped object. Thus it is well suited to model grasp intention. Accordingly, a grasp is interpreted as a mix of oppositions between hand-parts. Each opposition serves a particular functional end. For example, the grasp of screw-driving in Fig. 2 may be interpreted as a combination of 3 components: action of the thumb against side of the fingers which supports the action of fingers against the palm in order to keep the tool gripped firmly, while use of the thumb-tip against the finger-tip enables the tool to be directed appropriately during the task. We infer this mix of oppositions from the hand configuration and tactile information in a grasp demonstration.

Although the Opposition Space model admits oppositions where the hand is working against external forces, this paper relies only on opposition between two hand-parts. Consequently, we are not able to recognize non-prehensile grasps, such as the hook or flat-palm grasps, which work against gravity. Similarly,



**Fig. 2.** A screw-driving grasp may be interpreted in terms of 3 oppositions between hand-parts. Each opposition serves a particular functional role. Action of the thumb against side of the fingers supports the action of fingers against the palm in order to keep the tool gripped firmly. Use of the thumb-tip against the finger-tip enables the tool to be directed appropriately during the task.

with prehensile grasps, components where hand-parts work solely against task forces cannot be recognized. Examples of these would be finger extension for applying cutting force, pressing a button or resting the side of the palm against a surface during writing.

This paper makes the following contributions:

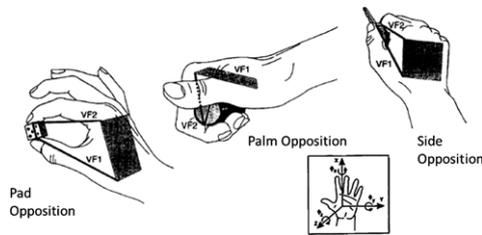
- We extend the definition of Opposition Space in [4] so that the full flexibility of the thumb – in opposing finger surface, palm and finger sides – may be recognized as separate components of a demonstrated grasp. Additionally, Opposition Space concepts are redefined so that they can be more readily applied in a demonstration context.
- We propose a new way to look at raw information from a grasp demonstration (configuration, tactile) by quantifying the importance of pair-wise interactions between elementary grasping surfaces of the hand. This intermediate representation integrates both hand configuration and interaction force, highlighting the multiple roles that a single sensor patch may have in a grasp. This representation is better able to discriminate among different kinds of oppositions and serves as a basis from which their presence in a grasp may be inferred.
- Inference of the most likely oppositions is done automatically without the use of heuristics. The method is not tied to any functional category of grasps. Experiments with human demonstrations show that the proposed method allows for recognition of a wide range of human grasp intentions with a recognition rate of 87%.

## 2. Related work

Different approaches are adopted in the literature when seeking to represent information from demonstrated human prehensile posture. Here we consider three approaches that are commonly encountered: joint angles and joint synergies, discrete classification of hand function through grasp taxonomies, virtual fingers in opposition.

Studies in human motion [6,16] have shown that there exists significant correlation among finger joint movements in prehensile postures for everyday tasks. Grasp configurations may therefore be represented by a low-dimensional subspace of a few principal components, known as hand synergies. The small number of dimensions makes it feasible to search for grasp configurations using metrics for overall grasp stability as presented in [17]. However, synergy representations face problems in task related scenarios where specific hand configurations appropriate for the functional requirements of the task are required, such as the examples shown in Fig. 1.

Alternatively, one may start with a set of grasps representing the functional categories of interest. Functional taxonomies such



**Fig. 3.** Categories of the Opposition Space framework [21]. Each category is made up of two virtual fingers [22] in opposition. The Virtual-to-Real (or V-R) mapping specifies real fingers which group together.

as those proposed by Cutkosky [3] and Kamakura [10] are employed for this purpose. In [7,18] separate synergy subspaces are found for each category which then enable human-like grasp formation for the identified grasp types. Another strategy commonly adopted first classifies the demonstrated grasp into a discrete category which is then mapped to a robotic system. Grasp recognition is accomplished by approximate nearest neighbor techniques on visual features in [11], while [12] discuss the use of neural networks trained on hand configuration obtained from a data glove. The authors in [13] improve classification by also incorporating information from demonstrated hand trajectories during the reaching phase. These approaches rely solely on visual appearance or hand configuration and hence suffer from the fact that similar hand shapes have entirely different functions depending on the particular hand surfaces employed. For example [13] confuse a power grasp (power sphere) with a precision grasp (precision disc) as both have similar hand shapes but the former makes use of finger and palm surfaces while the later uses only the finger-tips. To alleviate this problem, contact information can be used to improve the recognition of grasp categories. This is shown in [14] where contact information from virtual reality simulations forms the base for classification heuristics, whereas in [15] information from tactile sensors is combined with hand configuration to train HMMs. Murakami et al. [19] focus on the precise placement of tactile sensors to improve classification, while [20] learn a set of tactile-templates to interpret in-hand manipulation.

The representations discussed so far, do not capture how fingers may close to form the grasp or adapt it to different objects while preserving their functional role. A strategy which closes fingers of the hand uniformly until contact with the object is made or joint limits are reached does not preserve multiple axes of oppositional pressure. These constitute important task related components of the grasp as was explained earlier with the example of screw-driving Fig. 2.

A general way by which this problem may be addressed is through the Opposition Space framework as presented by Iberall et al. in [23,4]. Here, a prehensile posture is viewed in terms of oppositions between hand-parts. A set of 3 basic ways in which hand-parts may oppose are defined. These are shown in Fig. 3. Each opposition is associated with a different functional role. Palm opposition provides strong oppositional forces at the cost of dexterity. In contrast, pad opposition commands lighter forces but opens up degrees of freedom for manipulability. Finally, side-opposition provides intermediate forces while leaving some room for dexterous ability. Combining these oppositions together brings into play, in the coordinate frame of the hand, different abilities for the generation and control of force, torque and motion on a grasped object. Further, while correlated to how hand parts are employed functionally, these oppositions can also serve as a guide for lower level controllers to complete a grasp by driving the opposing hand-parts together [22,23].

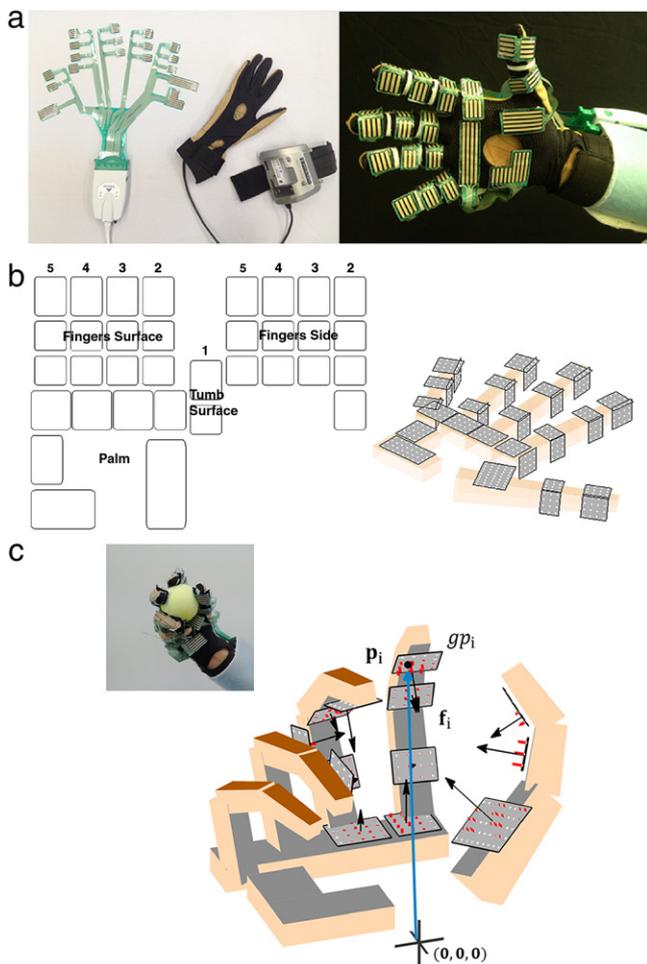
Using concepts from Opposition Space, Iberall et al. [21] explain hand postures across several functional taxonomies in the

literature. In [24,25] they present a rule based scheme to construct task-oriented grasps for robotic and prosthetic hands. The authors in [14] base their heuristics for grasp recognition on opposition between hand-parts. In the computational neuroscience literature, this has been used to explain behaviors in infant grasp learning [26] and imitation of grasping from visual stimuli [27]. Kang and Ikeuchi [28,5] use oppositions to characterize a demonstrated grasp for transfer to robotic systems. In [28] they propose a method to decompose a grasp into a set of oppositions between hand-parts. This is done by maximizing force coupling (or similarity between force vectors) among the real fingers assigned to each hand-part while simultaneously favoring a small number of functionally distinct hand-parts. This method makes the inherent assumption that entire fingers are involved in dedicated oppositions. We adopt a different perspective wherein fingers have grasping surfaces which can be simultaneously involved in multiple oppositional roles, for which they oppose different parts of the hand in varying degrees. We extend the definition of Opposition Space to identify a set of 41 oppositions for the hand. The extension allows for different oppositional roles of the thumb to be recognized as separate parts of a grasp. We propose a general method to detect the most likely oppositions present from tactile and configuration information in a demonstrated grasp. The method evaluated by human grasp experiments achieves an 87% recognition rate over a wide range of grasp behavior.

The rest of the paper is organized as follows. Section 3 describes the sensing infrastructure we use to capture human demonstrations of grasping. Section 4 outlines a model of grasp intention based on opposition between hand-parts. For this, the framework described in [4] is extended and relevant terminology is introduced leading to the definition of a grasp signature characterizing grasp intention. Section 5 proposes a method by which the grasp signature can be inferred from the raw sensor information obtained from a grasp demonstration. Section 6 reports on empirical evaluation of the proposed approach using human demonstrations of grasping conducted over a wide range of hand function. Sections 7 and 8 discuss directions for future research and conclude the paper.

### 3. Hardware setup

Human grasp demonstrations are captured using the hardware setup shown in Fig. 4(a). Hand configuration is measured through a data glove (Cyberglove [29]). The Cyberglove has 22 bend sensors strategically located over the hand joints. Since bending can be detected anywhere along the sensor length, the glove can adapt well to different hands sizes. The glove needs to be calibrated in order to transform raw sensor output to hand joint angles. Raw data from the glove is of dimension  $\mathbb{R}^{22}$ . Interaction forces from the grasping surfaces of the hand (including the sides of the fingers) are obtained through a tactile sensory array (TekScan [30]). The Tekscan sensor array consists of 18 sensors patches which are matrices of pressure sensitive sensing elements or sensels. The patches in one array are strategically located so as to cover the grasping surfaces of the human hand. Two tactile arrays are employed in an overlapping configuration in order to cover the frontal grasping surfaces of the hand as well as all finger-sides which are able to oppose the thumb. We make use of uncalibrated tactile response as only relative force levels are necessary for analyzing synergistic use of grasping surfaces. Raw data from the tactile sensory array is of dimension  $\mathbb{R}^{581}$ . A careful calibration of the data glove is conducted so that the relative geometry of the patches in the kinematic model reconstruction corresponds to the grasp demonstration. Data streams from the hand configuration and tactile response are synchronized. The combined data is obtained at a frequency of 200 Hz and is averaged over a pre-determined time interval over which the grasp demonstration is maintained. Further details on the hardware setup, the



**Fig. 4.** (a) shows the hardware setup to capture human grasp demonstrations. The raw sensory data is interpreted in terms of 34 sensor units with the large palm patches divided into subunits (b). (c) shows a grasp demonstration and the raw information captured using the tactile glove. Each sensor unit can be represented by a force vector  $\mathbf{f}_i$  equal to the sum of all sensel activations and acting at the centroid of pressure  $\mathbf{p}_i$ , defined in a coordinate frame centered at the wrist.

hand kinematic model employed and its calibration may be found in [31].

For the purpose of analyzing how the instrumented grasping surfaces of the hand are employed in a human grasp demonstration, they are subdivided into a set of elementary **grasping patches** as shown in Fig. 4(b). A total of 34 grasping patches are identified. These are denoted by

$$\mathcal{G}\mathcal{P} = \{gp_i\}_{i=1}^{34}. \quad (1)$$

Each grasping patch is viewed as a single unit of grasping force. The force  $\mathbf{f}_i \in \mathbb{R}^3$  associated with a grasping patch  $gp_i$  is obtained as the sum of all the sensel activations associated with it and is assumed to be acting normally to the patch at the centroid of sensel activations  $\mathbf{p}_i \in \mathbb{R}^3$  (Fig. 4c). The position and orientation of each patch is expressed with respect to a coordinate frame centered at the wrist. Data from a grasp demonstration can therefore be summarized as

$$D = \{\mathbf{p}_i, \mathbf{f}_i\}_{i=1}^{34}. \quad (2)$$

We consider only hand surfaces that are actively engaged in applying force on the object. Other sources of tactile response arise from artifacts induced due to glove construction. Prior to analysis, the active patches are identified by applying a threshold on the tactile response normalized by the maximum  $\|\mathbf{f}_i\|$  detected.

While this method works for frontal surfaces of the fingers and the palm it does not always work with finger sides. Due to artificial enlargement of the finger, the tactile signal may be quite large even when the finger side is not actively engaged with the object. For this paper, sides of fingers actually impacting the object are identified by visual analysis of the grasp. This could also be achieved automatically by fitting an object approximation given the tactile information and patch geometry.

#### 4. A model for grasp intention

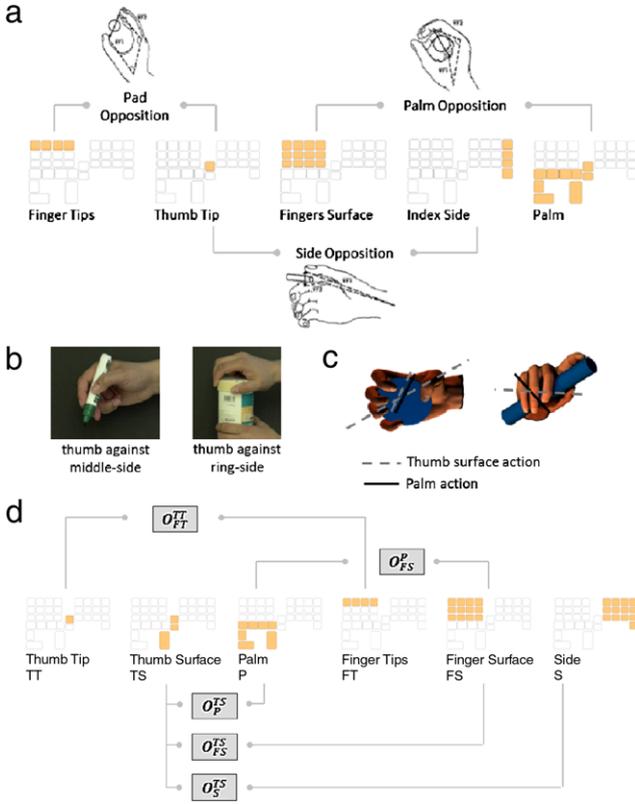
Even before a grasp is realized the human has an intention in mind regarding the grasping patches to be engaged and how these patches will be employed against an object. We have referred to this earlier as the *grasp intention*. In this section we define a model to recognize grasp intention from human demonstrations of grasping. The model builds on the Opposition Space framework introduced by Iberall and Arbib in [4]. We revisit these concepts in terms of the information provided by the sensing infrastructure described in Section 3.

The authors in [4] were motivated by the incompleteness of all existing prehensile posturing taxonomies in the literature. While taxonomies explain broad categories of grasps from a functional perspective, they do not explain how grasps are formed and consequently lack the ability to explain the myriad of variations that occur in practice. This motivated the authors in [4] to look for a functional basis with which grasps could be explained in a bottom-up manner. After analyzing a large number of grasp taxonomies in the literature, they reached the conclusion that all taxonomies investigated could be explained by 3 types of oppositions between hand-parts. The oppositions identified were: *pad*, the thumb-tip opposing the set of finger-tips, *palm*, the set of frontal finger surfaces opposing the palm surface, and *side*, the thumb-tip opposing the side of the index finger. Fig. 5(a) depicts this visually.

Our view is that while this model may not explain all variations in human grasp behavior, it presents a means by which we may interpret the high level intention governing the way hand-parts are committed for use even before the grasp is formed. This has implications for control of robot hands. The high level oppositional intention mandates a coupling between sub-parts of each finger and other parts of the hand which can be expressed in terms of opposing and cooperating constraints. Controlling for constraints can be used to preshape the hand as well as conform it to object profile and adapt it to perturbation while serving some high level intention. This could be advantageous over controlling each finger independently.

##### 4.1. Extended definition for opposition space

Underlying the concept of oppositions is a sub-division of the hand's grasping patches into **hand-parts** based on similar functional roles, such as the set of finger-tips. Each hand-part can be seen as the maximal set of grasping patches which can be committed together for use with the same oppositional intention. The original set of hand-parts (Fig. 5(a)), has some limitations in properly representing the full flexibility of the thumb, leaving some grasping ability unaccounted for. *side opposition* accounts only for opposition of the thumb against the index finger side. This implies that thumb usage against other finger sides cannot be recognized. Fig. 5(b) shows some examples where this forms an important component of commonly encountered grasps. Another issue is that the oppositional intention of the thumb surfaces is always clubbed with that of the palm. While this is true for some grasps (where the thumb acts as an extension of the palm), in many instances, such as the examples shown in Fig. 5(c), thumb



**Fig. 5.** (a) shows a hand-part decomposition as prescribed by the Opposition Space framework [23]. 3 oppositions are formed when the hand-parts are combined in kinematically feasible ways. (b) and (c) show common scenarios illustrating limitations of this framework in capturing full flexibility of thumb surface usage. (d) shows the new hand-part decomposition proposed along with feasible oppositions between them.

action has quite a different functional meaning and should form a separate component of the grasp. That the original set of hand-parts leaves some grasping capability unaccounted for may also be seen from the fact that the union of oppositions between them is not equal to the entire set of grasping patches.

To address these limitations the original framework is modified by adding another hand-part definition, *Thumb Surface*, which separates out the action of thumb surfaces from that of the palm. Also, the *Index Side*, now called just *Side*, is enlarged to cover sides of all fingers. The new set of hand parts is shown in Fig. 5(d). They are collectively referred to as  $\mathcal{H}$ , below, where each hand-part is abbreviated by its starting letter (e.g. Fingers Surface = FS).

$$\mathcal{H} = \{TT, TS, P, FT, FS, S\}. \quad (3)$$

Note that for all hand-parts  $h \in \mathcal{H}$ ,

$$h \subset \mathcal{G}, \quad \bigcup_{h \in \mathcal{H}} h = \mathcal{G}, \quad \bigcap_{h \in \mathcal{H}} h \neq \emptyset.$$

The union of these hand-parts now covers the entire grasping patch set, but their intersection is not empty. The overlapping hand-parts model the fact that individual grasping patches can play multiple functional roles when different oppositions cooperate in delivering the overall functionality of a grasp. There are 5 ways in which oppositions between these hand-parts are kinematically feasible. These are shown in Fig. 5(d). Let  $O_y^x$  be a notation to represent an opposition between the hand-parts  $x$  and  $y$ . Using this notation, the set of hand-part oppositions can be denoted as:

$$\mathcal{O}_{\mathcal{H}} = \{O_{FT}^{TT}, O_{FS}^P, O_P^{TS}, O_{FS}^{TS}, O_S^{TS}\}. \quad (4)$$

This is a new definition for Opposition Space. The set  $\mathcal{O}_{\mathcal{H}}$  cannot be used directly as a model for grasp intention, due to ambiguities

that exist. To resolve these ambiguities, we use the concept of a virtual finger and impose constraints on how virtual fingers may be formed. This leads to a set of 41 opposition primitives. The following sub-sections explain this in more detail.

#### 4.2. Opposition primitives

In practice, the elements of  $\mathcal{O}_{\mathcal{H}}$ , with the exception of  $O_P^{TS}$ , do not uniquely determine a grasp intention. This is because of the different ways fingers may group together in response to object properties or functional requirements of the task. For example, in the turning-a-dial task of Fig. 1, which is a case of opposition  $O_{FT}^{TT}$ , 3 fingers (index–middle–ring) group together to act against the thumb. In general however, this number depends on the diameter of the dial. Thus each element of  $\mathcal{O}_{\mathcal{H}}$  is actually a category of oppositions.

The virtual finger concept is used by [22,4] to represent a set of real fingers that act together with the same oppositional intention. In the context of a recognition framework, we define a **virtual finger** as a cooperating set of grasping patches belonging to a hand-part. Examining  $\mathcal{H}$  closely, we see that with the hand-parts  $TT$ ,  $TS$  and  $P$ , there is no ambiguity, as they identify a set of grasping patches that are constrained to be used in their entirety. Thus the hand-part name is sufficient to identify the virtual finger. With  $FT$ ,  $FS$ ,  $S$  however, ambiguity exists, as several combinations are possible based on the number of real-fingers that act together with the same oppositional intention. To denote these possibilities let the index, middle, ring, little fingers be identified by numbers 2–5 and let  $\mathcal{F}$  represent all their combinations.

$$\mathcal{F} = \{f \mid f \subset \{2, 3, 4, 5\}, f \neq \emptyset\}. \quad (5)$$

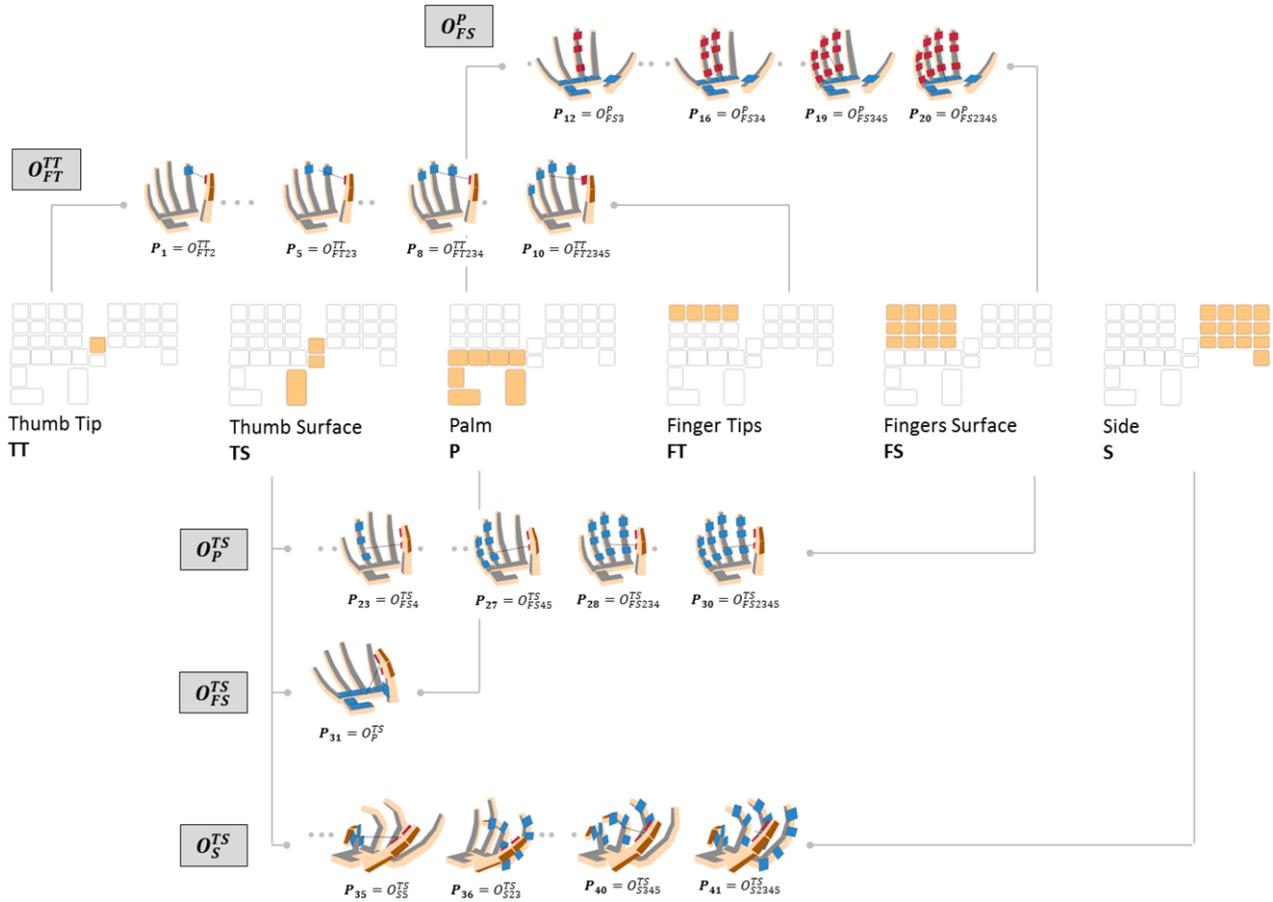
A virtual finger is denoted by concatenating a hand part  $h \in \{FT, FS, S\}$  with a real-finger grouping  $f \in \mathcal{F}$ . So,  $FT234$  denotes the set of grasping patches belonging to tips (distal phalanges) of the index, middle and ring fingers, whereas with  $FS234$  the set of cooperating grasping patches is extended to cover also the middle and proximal phalanges of the same fingers.

An opposition made between two virtual fingers belonging to different hand parts, provided opposition between the hand-parts is kinematically feasible, is termed as an **opposition primitive** (or just primitive for brevity). It is denoted by  $O_{VF_1}^{VF_2}$ . Using this terminology, we can explain the turning-a-dial task of Fig. 1 as making use of the opposition primitive  $O_{FT234}^{TT}$ .

#### 4.3. Primitive set for recognizing grasp intention

The cardinality of  $\mathcal{F}$  is 15 ( $C_1^4 + C_2^4 + C_3^4 + C_4^4$ ). In conjunction with  $\mathcal{O}_{\mathcal{H}}$ , this gives a total of  $4 * 15 + 1 = 61$  opposition primitives. Many of these however are never employed in practice. This can be seen by examining grasp taxonomies in the literature, such as [32], made from studies of human grasping behavior. Motivated by the same studies, we find that it is almost always the case that virtual finger span is contiguous. For example, this means that if fingers 2 and 4 are being used with the same oppositional intention, say  $O_{FS2}^P$  and  $O_{FS4}^P$ , finger 3 is required to cooperate with them and the primitive being used is actually  $O_{FS234}^P$ . With this simplifying assumption there are 10 valid real-finger groupings: 2, 3, 4, 5, 23, 34, 45, 234, 345, 2345 and a total of 41 opposition primitives. These are indicated in Fig. 6.

Let this primitive set be denoted by  $\mathcal{P} = \{P_1, \dots, P_{41}\}$ . As seen from Fig. 6, each element of  $\mathcal{P}$  is associated with a pre-shape configuration from which the opposing hand-parts can be brought together to manifest the opposition on an object. Thus selecting a primitive as part of a grasp in effect commits grasping patches for



**Fig. 6.** Shows the set of opposition primitives with which *grasp intention* is interpreted. For the categories  $O_{FT}^{TT}$ ,  $O_{FS}^P$ ,  $O_{FS}^{TS}$ ,  $O_S^{TS}$ , there are several ways in which real-fingers may group together with the same oppositional intention. Using the simplifying assumption of contiguous virtual finger span, 10 possibilities may be identified: 2, 3, 4, 5, 23, 34, 45, 234, 345, 2345. For clarity only a selected number of these are shown above.

use with a well defined intention, cooperating with some and opposing others, even before a grasp is formed. The set  $\mathcal{P}$  is therefore able to represent grasp intention. The difference of this primitive set with the original Opposition Space framework is seen in the additional primitives in the bottom half of Fig. 6 (primitives no. 21–41). These primitives model flexibility of thumb usage (against finger surfaces, against palm, against sides of all fingers) all of which play an important role towards overall hand functionality. The addition of these primitives implies that they can now be recognized as separate intentions in a grasp demonstration.

#### 4.4. Grasp signature from opposition primitives

A **grasp signature**,  $GS$ , is defined as an importance distribution over the set of opposition primitives.

$$GS = \left\{ \mathbf{x} \in \mathbb{R}^{41} \mid x_i \geq 0, \sum x_i = 1 \right\}.$$

The grasp signature characterizes the grasping intention underlying a demonstrated grasp.

### 5. Inferring a grasp signature

In this section we see how the presence of opposition primitives along with the importance with which they are manifested can be inferred from hand configuration and contact force information present in a grasping demonstration. For this the interactions between active grasping patches in a grasp demonstration are quantified. If a primitive is an important component of a grasp,

it becomes more likely to find “strong” patch level interactions between the opposing groups of grasping patches by which it is defined.

We first define a metric of opposition strength which quantifies the importance of pairwise oppositions between grasping patches in the context of the demonstrated grasp. This is defined in terms of normal force and geometrical opposition of grasping patch pairs, Section 5.1. It is then applied to all grasping patch pairs for which opposition is kinematically feasible (as defined in Section 5.2) using information from the grasp demonstration  $D = \{\mathbf{p}_i, \mathbf{f}_i\}_{i=1}^{34}$  where  $\mathbf{p}_i$  and  $\mathbf{f}_i$  denote the position and force vector associated with grasping patch  $gp_i$ . This results in an intermediate representation exposing the contribution towards different primitives in which a single grasping patch may be involved and serves as the basis from which they can be inferred (Fig. 7).

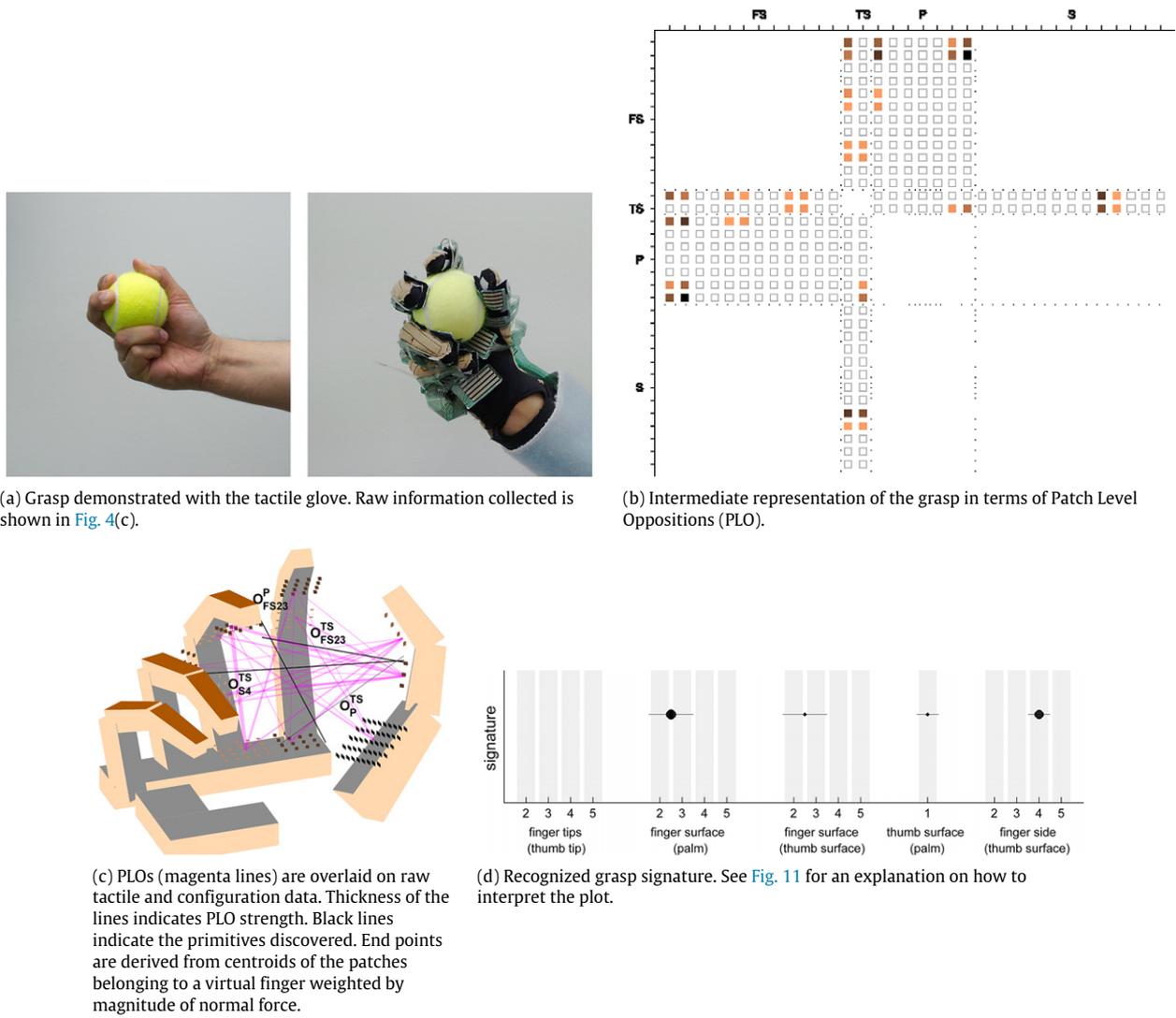
#### 5.1. A metric for patch level opposition

Given a pair of grasping patches  $gp_i, gp_j$ , we wish to quantify how relevant is the opposition of  $gp_i$  against  $gp_j$  to the demonstrated grasp. We propose a metric of opposition strength based on two measures:

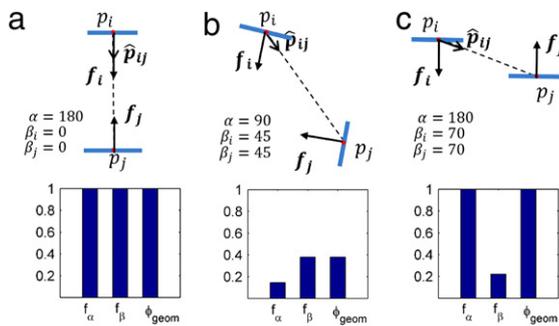
(1) *The normal force.* The minimum of the two forces  $\mathbf{f}_i$  and  $\mathbf{f}_j$  is taken.

$$\phi_{force}(\mathbf{f}_i, \mathbf{f}_j) = \min \{ \|\mathbf{f}_i\|, \|\mathbf{f}_j\| \}. \quad (6)$$

(2) *Quality of geometrical opposition.* Two angles arising from the relative geometry of the patches are considered and the one having the greater influence is used.



**Fig. 7.** Inferring a grasp signature (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



**Fig. 8.** Angles arising from relative geometry of patches are used to quantify opposition quality.  $\alpha$  is angle between the normal force vectors  $\mathbf{f}_i$ ,  $\mathbf{f}_j$ .  $\beta_i$ ,  $\beta_j$  are the angles that normal force vectors make with the line joining patch centroids  $\hat{\mathbf{p}}_{ij}$ .

The first is the angle between the normal force vectors.

$$\alpha = \cos^{-1} \left\langle \frac{\mathbf{f}_i}{\|\mathbf{f}_i\|}, \frac{\mathbf{f}_j}{\|\mathbf{f}_j\|} \right\rangle.$$

Patches oppose the best when  $\alpha = 180$ , Fig. 8(a). The quality of opposition decreases with decreasing  $\alpha$ . Once  $\alpha$  crosses a threshold i.e.  $\alpha < \alpha_t$ , opposition between the patches is deemed not relevant.

The second is the angle that force vectors make with the line joining patch centroids.

$$\beta_i = \cos^{-1} \left\langle \frac{\mathbf{f}_i}{\|\mathbf{f}_i\|}, \hat{\mathbf{p}}_{ij} \right\rangle \quad \beta_j = \cos^{-1} \left\langle \frac{\mathbf{f}_j}{\|\mathbf{f}_j\|}, -\hat{\mathbf{p}}_{ij} \right\rangle$$

$$\beta = \max \{ \beta_i, \beta_j \}.$$

In contact models used for analytical grasping analysis [33], this angle is related to the maximum force that can be applied to a surface before incurring the risk of slipping.  $\beta$  should lie within the friction cone of the surface. We use it here to indicate opposition capability even when the angle between the normal force vectors decreases significantly such as in Fig. 8(b). In such cases, opposition is still possible if enough friction exists or if the object is immobilized by other parts of the hand. Smaller the angle  $\beta$  greater is the force which can be applied.

As long as opposition is deemed relevant i.e.  $\alpha \geq \alpha_t$ , the influence of  $\alpha$  and  $\beta$  on opposition strength can be quantified using the functions  $f_\alpha$  and  $f_\beta$  as follows:

$$f_\alpha = e^{\left(-\frac{\alpha-\alpha_c}{\alpha_c}\right)^\gamma}$$

$$f_\beta = e^{\left(-\frac{\beta}{\beta_c}\right)^\gamma}.$$

Values for all parameters are indicated below.

$\alpha_t = 1.48$	Opposition is considered only if angular separation between normals is greater than $85^\circ$ .
$\alpha_c = 1.22$	Reduces the effect of $\alpha$ once $\alpha \leq 110^\circ$ .
$\beta_c = 1.22$	Conservative estimate of a friction cone taken at $70^\circ$ .
$\gamma = 1.5$	Determined empirically.

$f_\alpha$  and  $f_\beta$  are dominant in different situations. In Fig. 8(b),  $f_\alpha$  is low but opposition is still possible,  $f_\beta$  is the better indicator. In Fig. 8(c),  $f_\beta$  is low but the patches are well opposed,  $f_\alpha$  is the better indicator. The effect of patch geometry on the quality of opposition is defined considering both functions as follows:

$$\phi_{geom}(\hat{\mathbf{p}}_{ij}, \mathbf{f}_i, \mathbf{f}_j) = \max \begin{cases} 0 & \alpha < \alpha_t \\ \{f_\alpha, f_\beta\} & \alpha \geq \alpha_t. \end{cases} \quad (7)$$

Finally, the metric of opposition strength,  $\phi_{plo}$ , is defined as:

$$\phi_{plo} = \phi_{geom} \cdot \phi_{force}. \quad (8)$$

Since  $0 \leq \phi_{geom} \leq 1$ ,  $\phi_{plo}$  may be seen as the normal force modulated by the geometrical quality of the opposition.

### 5.2. Feasible patch level oppositions (PLO-space)

There are  $C_2^{34} = 561$  pairwise combinations for the 34 grasping patches in  $\mathcal{G}\mathcal{P}$ . Of these, the set of grasping patch pairs for which opposition is kinematically feasible is termed as the PLO-space. Let  $O_{PLO} \in \mathbb{R}^{34 \times 34}$  represent all valid pairwise interactions between grasping patches.

$$O_{PLO}(j, k) = \begin{cases} 1 & \text{patch } gp_j \text{ can oppose } gp_k \\ 0 & \text{otherwise.} \end{cases} \quad (9)$$

Each primitive  $P_i = O_{VF_2^i}^{VF_1^i} \in \mathcal{P}$  defines a set of valid pairwise interactions between grasping patches as a consequence of opposition between its virtual fingers.

$$O_{PLO}^i(j, k) = \begin{cases} 1 & j \in VF_1^i \text{ and } k \in VF_2^i \\ 0 & \text{otherwise.} \end{cases} \quad (10)$$

The PLO-space can be computed as a union of  $O_{PLO}^i$  and is shown in Fig. 9.

$$O_{PLO} = \bigvee_{i \in \mathcal{P}} O_{PLO}^i.$$

This method of determining a PLO-space captures the fact that it is infeasible for each major hand-part to oppose its own self. Further, surfaces of the fingers (excluding the thumb) cannot oppose their sides and distal patches of the fingers cannot oppose the intermediate patches.

### 5.3. Grasp signature

The grasp signature is a distribution over the primitive set corresponding to the importance with which each primitive is manifested in the grasp demonstration. To discover this from hand configuration and tactile force in a grasp demonstration, the oppositional roles possible for each finger are examined. These correspond to the set of primitives listed below.

$$\mathcal{X} = \{O_{FT2}^{TT}, O_{FT3}^{TT}, O_{FT4}^{TT}, O_{FT5}^{TT}, O_{FS2}^P, O_{FS3}^P, O_{FS4}^P, O_{FS5}^P, O_{FS2}^{TS}, O_{FS3}^{TS}, O_{FS4}^{TS}, O_{FS5}^{TS}, O_P^{TS}, O_{S2}^{TS}, O_{S3}^{TS}, O_{S4}^{TS}, O_{S5}^{TS}\}$$

To obtain primitive likelihood we use the notion of a primitive template. The studies by Kamakura [10], using real world

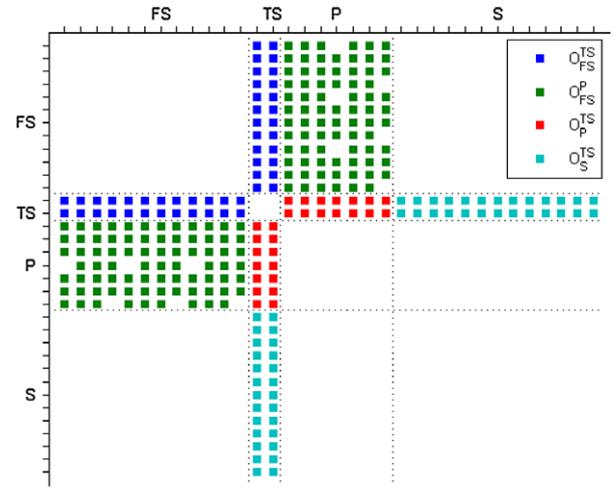


Fig. 9. The space of kinematically feasible patch level oppositions. Each axis represents the set of 34 grasping patches grouped by major hand-part: FS—Finger Front Surface, TS—Thumb Surface, P—Palm Surface, S—Finger Side Surface. Patch level oppositions are color-coded according to the hand parts between which they occur. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

objects, identified tactile signatures commonly encountered when employing certain finger pre-shapes. Similarly [20] use tactile templates of grasping regions to characterize 7 grasps with which to interpret in-hand manipulations. In our case, the template for each opposition primitive is based on the pairwise oppositions that could be generated by it  $O_{PLO}^i$ , see (10). Since the oppositional intention for any given primitive is known, we may identify patches on each opposing hand part where we may expect to see oppositional pressure focused if the primitive is being used. These are termed as primary patches. The surrounding patches, acting in support of these are termed secondary patches. Following the reasoning that interactions between primary patches have the most importance followed by primary–secondary interactions and then secondary–secondary interactions, a relevance mask or template for a primitive can be defined as in Eq. (11).

$$M_{i \in \mathcal{X}}^i(j, k) = \begin{cases} 1 & gp_j, gp_k \text{ is a primary patch pair} \\ 0.7 & gp_j, gp_k \text{ is a primary–secondary patch pair} \\ 0.3 & gp_j, gp_k \text{ is a secondary patch pair} \\ 0 & gp_j, gp_k \notin P_i \text{ i.e. } O_{PLO}^i(j, k) = 0. \end{cases} \quad (11)$$

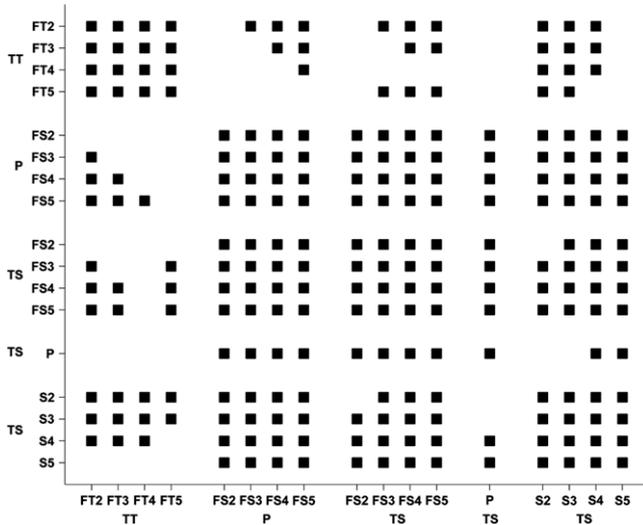
This may then be used as a prior knowledge filter to determine a value proportional to the likelihood of a primitive's presence given the observed patch level oppositions in a grasp demonstration Eq. (12).

$$\Phi(i) = \sum_{j,k=1}^{34} M_{i \in \mathcal{X}}^i(j, k) \cdot \phi_{plo}(\hat{\mathbf{p}}_{jk}, \mathbf{f}_j, \mathbf{f}_k) \quad (12)$$

$\Phi$  once normalized represents primitive likelihood.

Due to the kinematic coupling that exists between the finger joints, choosing one of these as the dominant intention in a grasp makes others infeasible. Coexistence between these oppositional intentions can be pre-analyzed and is recorded in a primitive compatibility matrix, Fig. 10.

Discovering a set of cooperating primitives proceeds in an iterative fashion. First the most likely primitive  $\psi = \max_{i \in \mathcal{X}} \Phi$  is chosen. The span of the virtual finger is expanded by selecting all primitives in  $\mathcal{X}$  having a non-zero likelihood of opposition with the same hand-part as  $\psi$ . Normal force of patches contributing to the selected primitives are reduced by strength of the contributing PLOs and  $\Phi$  is recomputed. The new likelihood thus incorporates



**Fig. 10.** The opposition compatibility matrix captures co-existence between oppositional intentions at the finger level. Black squares indicate cooperation is possible. White squares indicate incompatibility. Categories  $O_{FS}^P, O_{FS}^S, O_P^S, O_S^S$  mostly cooperate but conflicts arise between these and finger-tip opposition  $O_{FT}^T$ .

an explanation of the raw data due to the selected primitives. The selected primitives as well as those that are not compatible with the ones selected are excluded from consideration and the process is iterated. The process terminates when there are no more single finger oppositions likely i.e.  $\Phi(i) = 0 \quad \forall i \in \mathcal{X}$ .

The grasp signature discovered above is modified to express contiguity of virtual finger span. For example, if fingers 2, 3 and 4 are found to be opposing the palm i.e.  $O_{F2}^P, O_{F3}^P$  and  $O_{F4}^P$  are present, then these are combined and reported as  $O_{F234}^P$ . If  $O_{F3}^P$  is absent, then to ensure contiguity of the virtual finger, palm opposition for finger 3 is added and the primitive  $O_{F234}^P$  is reported. Finally, each primitive reported is assigned an importance by summing the PLO strength for each participating finger and taking the average.

## 6. Experimental results

The system for recognizing grasp intention from tactile and configuration data is evaluated using human grasp demonstrations carried out over several grasp scenarios. We first establish the intention that will be demonstrated and then vary other variables in the system to evaluate recognition performance. A grasp scenario consists of an object–grasp pair. The grasp chosen is motivated by how the object should be handled in order to perform a particular task. Alternatively, it can be picked from a grasp taxonomy. Identifying a grasp scenario implies that a specific intention for grasping is communicated to the human subject who will manifest this on an object. Performance of the system is based on whether the recognized signature corresponds to the pre-identified grasp intention. All grasps are demonstrated using the tactile glove described earlier. We differentiate between expert and naïve demonstrator. An expert is one who has a lot of experience with using the tactile glove to grasp objects and is well versed in grasp taxonomy. A naïve demonstrator has no knowledge of either.

Several parameters can be varied to examine the generality and reliability of the system. These include the grasp intention itself, the object on which the grasp is manifested or the hand (subject) making the grasp. A set of experiments are designed to vary the different parameters. An underlying theme behind the experiments is that if the intention is the same then regardless of whether the grasp is demonstrated by different hands or on different objects, the grasp signature recognized should remain

unchanged. However, if the intention changes then this change should be correctly reflected in the grasp signature. The change can be small such as the importance given to different grasp components or the number of fingers employed, or large, as when employing a different combination of oppositions.

### 6.1. Single demonstrator

The performance of grasp intention recognition is evaluated using a single expert demonstrator. By using an expert demonstrator we minimize the possibility of misunderstanding the grasp intention or improperly manifesting it on the object using the tactile glove.

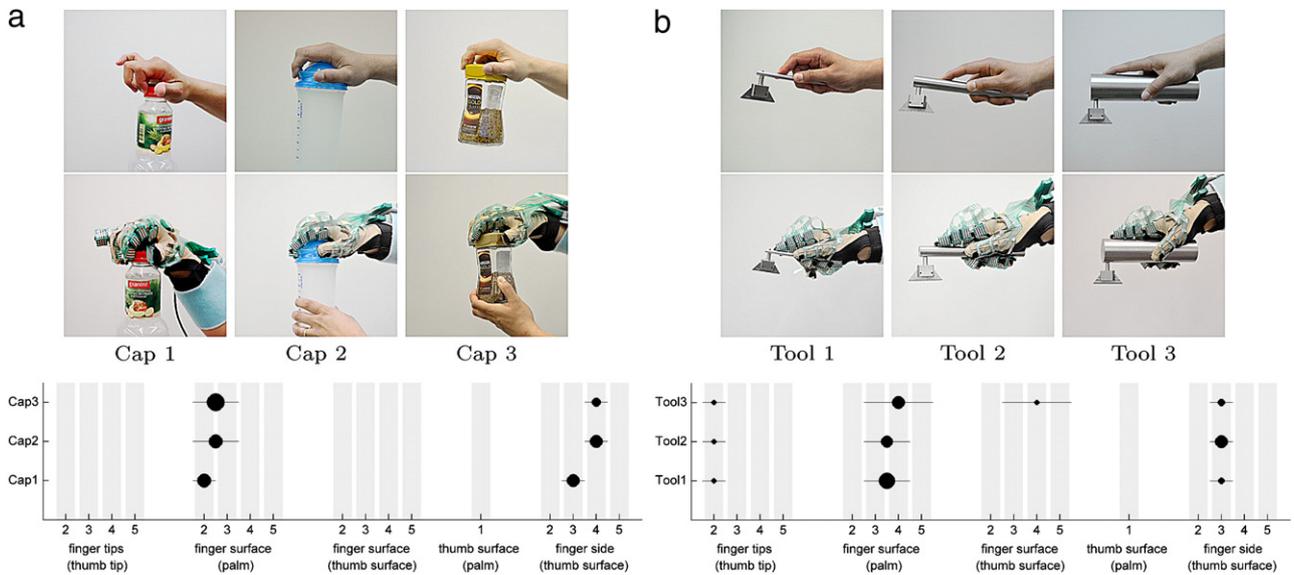
We test first the outcome of exercising same intention on different objects. The first scenario involves the task of opening a tight bottle cap. Three different caps are used having different size and/or shape (round, square). The grasp does not need to deliver any motion rather it needs to grip the cap firmly in order to transmit the strong torques and coarse motion generated by the wrist–arm without allowing any slippage. This is done by using the thumb surface against the side of the ring finger supported by the action of finger-tips (index–middle) against the fleshy part of the thumb (Fig. 11(a)). Grasp intention remains the same for cap 2 and 3 but is changed slightly for cap 1 to accommodate the smaller diameter. Referring to the figure, the recognized signature for each grasp–object pair corresponds to the grasp intention and also correctly captures the change in intention for the small diameter cap. A grasp scenario involving a cutting tool is also taken. Three handles of different diameter and weight are used. The grasp chosen grips the handle firmly using action of fingers (middle–ring) against the palm while directional ability is provided by a tripod grip between thumb, middle-side and index-tip (Fig. 11(b)). This intention remains the same for tool 1 and 2 but is changed slightly to accommodate the larger diameter and weight of tool 3. The grasp signatures plotted shows that this intention is well recognized including the change in intention for tool 3.

Next we test different intention on the same object. Opening a tight bottle cap scenario is taken. The grasp for this was explained earlier. However, when the cap becomes loose a new grasp is employed. The strong action of palm opposition is no longer required and is replaced instead by use of finger tips. Fig. 12(a) shows that this different intention is correctly detected from the demonstration. A cutting tool scenario is also examined. Here the position of the cutting blade is changed to the middle (tool 2) and the top (tool 3) of the handle as shown in Fig. 12(b). Entirely different grasps are now required in each case. The grasp for tool 2 uses exclusively side-opposition, whereas grasp for tool 3 uses a combination of palm opposition and side opposition. Examining the recognized grasp signatures, we see that these intentions are well detected.

From the above experiment it is seen that the method for grasp recognition performs well when intention is kept the same, is changed in a small way or when completely different. However, only 1 hand is used and relatively few number of intentions are demonstrated. In the next experiment we widen the set of grasps and objects considered. Also, grasps are demonstrated by several naïve subjects who do not have experience using the tactile glove nor are knowledgeable about grasp taxonomy.

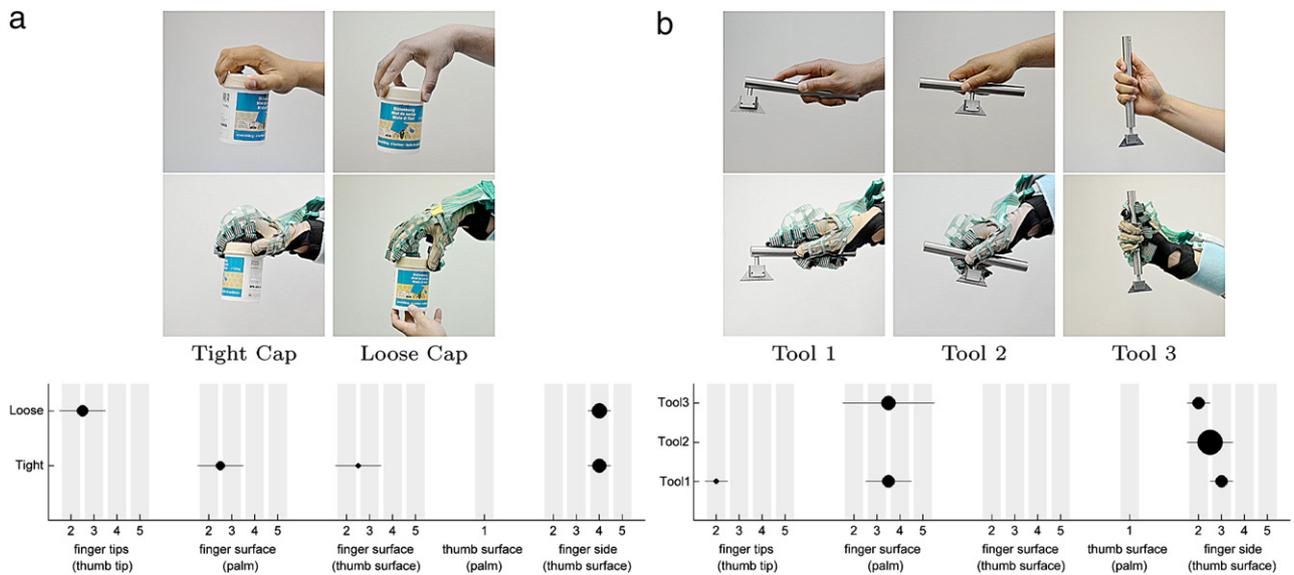
### 6.2. Multiple demonstrators

The performance of grasp intention recognition is evaluated using a 10 naïve demonstrators. The subjects were male (between 21–30 years of age) with different hand sizes as summarized in Fig. 14. Ten grasp scenarios are selected representing a broad range of grasp intention. Grasps are taken from the taxonomy from



**Fig. 11.** Similar grasp intention on different objects. Single demonstrator. (a) Examines opening a tight bottle-cap with different size/shape. Cap1 and Cap2 are round with diameter 4 and 8.5 cm, while Cap3 is square with side 7 cm. Grasp intention is changed slightly for Cap1 to accommodate smaller diameter. (b) Examines cutting with different diameter/weight handles: Tool 1 ( $\varnothing = 0.8$  cm, 23 gm), Tool 2 ( $\varnothing = 2$  cm, 500 gm), Tool 3 ( $\varnothing = 5$  cm, 3136 gm). Grasp intention is changed slightly for Tool3 to accommodate larger diameter and weight.

The recognized grasp signatures are plotted below the figures in each case. The y-axis denotes the grasp scenario and the x-axis denotes the different opposition classes. Fingers thumb–index–middle–ring–little are numbered 1–5. For each grasp scenario, primitives detected are denoted by filled circles. The circle diameter corresponds to the importance of the primitive in the grasp. The horizontal line in each circle indicates the virtual finger span i.e. the real fingers detected as having the same oppositional intention. This representation allows any subset of the 41 primitives to be presented in a compact manner. For example, recognized signature for scenario Cap2 and Cap3 comprise of the primitives  $O_{FS23}^p$  and  $O_{S4}^{TS}$ , whereas for Cap1 the primitives recognized are  $O_{FS2}^p$  and  $O_{S3}^{TS}$ .



**Fig. 12.** Different grasp intention on the same object. Single demonstrator. Fig. 12(a) examines opening a bottle-cap when it is tight and when it is loose. Entirely different grasps are required for each case. Fig. 12(b) examines cutting with different tools. Tools 1–3 use the same handle but have the cutting blade positioned differently requiring entirely different grasps for each case. See Fig. 11 for an explanation on how to interpret the plots.

Feix et al. [32] which represents a comprehensive summary of several grasp taxonomies proposed in the literature. Grasps are selected to cover a wide range of hand functions. These range from high precision grasps, which allow fine resolution of motion, to intermediate grasps mixing power with the ability to control force and torque at a tool tip, to high power grasps with strong resistance to external wrenches from all directions. Grasps selected are presented in Table 1. Other than wide range of intention, we may note that scenarios 5–7 test the case where different intention is manifested on the same object. Also, scenarios 7, 8, 9 are all examples of power grasp with directional ability. Different grasp components provide the directional capability in each case.

For each grasp scenario, the grasp intention is communicated to the human subject using a picture of the grasp and a high level description (as recorded in Table 1). The subject first tries out the grasp with the ungloved hand and then with the gloved hand. Tactile and configuration data are recorded once the subject is comfortable with creating the grasp with the gloved hand. Fig. 13 shows an example of the grasps in Table 1 being demonstrated with the tactile glove.

A total of 100 trials were conducted across all scenarios and all subjects. Fig. 15 and 16 present the recognition results. Based on a visual analysis of the taxonomy category an expected signature may be identified for each grasp scenario (indicated in red in

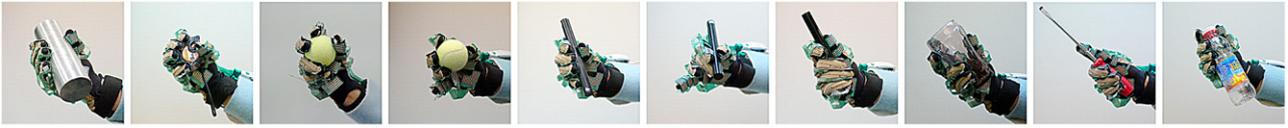


Fig. 13. Grasp scenarios in Table 1 demonstrated on objects using the tactile glove.

Table 1

Grasp scenarios covering a range of different grasp intentions. Scenarios and figures (except for 8, 10) are taken from [32].

Scenario	Description
1	 <b>Power grasp.</b> Object is encaged and grasped firmly.
2	 <b>Power grasp.</b> Object is encaged and grasped firmly.
3	 <b>Power grasp including finger side.</b> Object is encaged and grasped firmly.
4	 <b>Fingertip grasp.</b>
5	 <b>Fingertip grasp with side support.</b>
6	 <b>Tripod grasp.</b> Tripod is made using fingertip and finger side.
7	 <b>Directional power grasp.</b> Directional ability is provided using thumb and finger side.
8	 <b>Directional power grasp.</b> Directional ability is provided using thumb and finger tips.
9	 <b>Directional power grasp.</b> Directional ability is provided using thumb, finger side and finger tips.
10	 <b>Power grasp with dexterous ability.</b> Dexterous ability is provided using thumb and finger tip.

Figs. 15 and 16). However, it was noticed that subjects did not always adhere to the communicated intention. Clear variations exhibited on the communicated intention were noted at the time of demonstration and are reported in the results (indicated in blue). Out of 100 trials, 10 variations were noticed.

We see that in 87 trials the recognized signature matches the expected exactly, thus indicating that the system is good at detecting expressed intention from tactile and configuration data over a wide range of ways in which the hand may be utilized to generate grasps. Cases where mismatch occurred were investigated and are noted below. In the following (X-a.b.c) should be read as scenario-X, subjects-a,b,c.

(a) In 9 trials although the grasp was demonstrated correctly, tactile signal was too weak. This resulted in the virtual finger

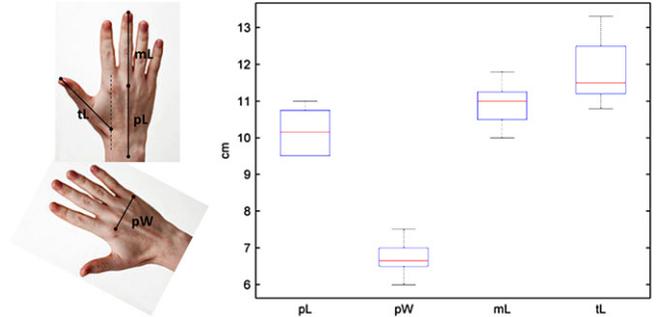


Fig. 14. Dimensions for measuring hand size and dispersion of hand size data across the 10 subjects. For each dimension, the vertical line indicates extremes, box covers 25th–75th percentiles and the red line denotes the median.

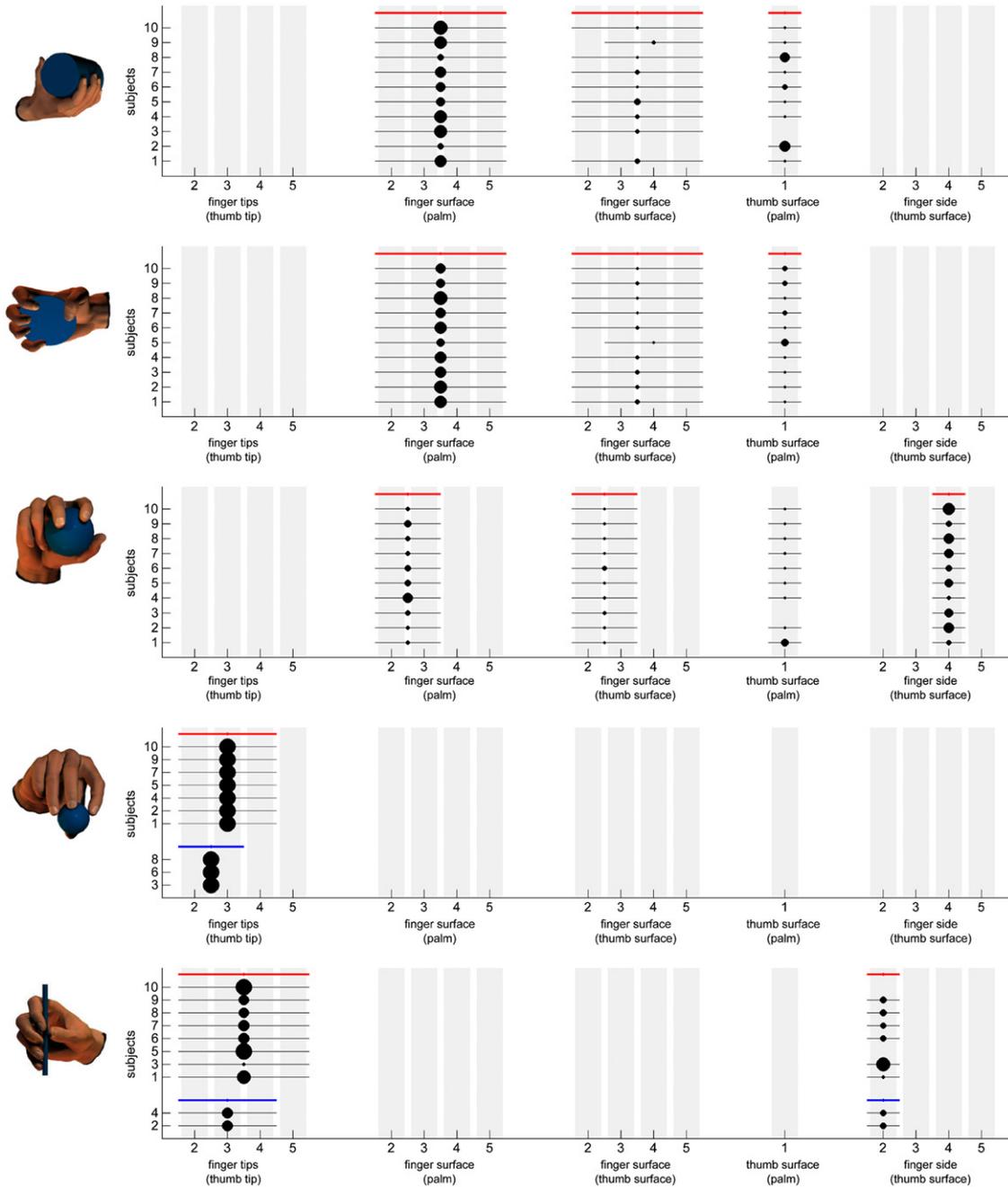
- span being smaller than expected (1–9, 2–5, 6–1.6, 7–10), or certain components not getting detected at all (1–2.3, 5–5.10).
- (b) In 4 trials, confusion occurs where oppositions are detected which have clearly not been demonstrated (6–9, 7–4.6.8). In all these cases,  $O_{FS2}^p$  is detected but it is clear that the index finger is employed differently.
- (c) In 23 trials, an additional component of the type  $O_F^{FS}$  is detected when not expected. This is seen in scenarios 3 and 7–10. This can be explained due to the natural tendency to include this component when the power grasp  $O_F^p$  is being exercised as a strong intention, which is the case in all these scenarios.

Examination of the confusions detected (point b) showed that these were caused due to patches which exhibited geometrical opposition but whose tactile response came from some other involvement. When two patches oppose geometrically and also exhibit tactile response, the system assumes mutual opposition and quantifies opposition strength. The system cannot tell if the tactile response is due to other causes. Recognition of the correct signature relies on the fact that the grasp intention being demonstrated results in oppositions that are stronger, causing the correct primitives to be prioritized. However relying solely on geometry and interaction force can result in confusion and other indicators of opposition would need to be considered.

Although the system reports on importance of the grasp components recognized, we have no basis for examining the detected importance. For static grasping there is no cause for giving importance to different grasp components. Importance only becomes relevant in a task context when the capabilities brought to the grasp by a component are exercised in response to task demands that occur. This may also be partly responsible for tactile signal being absent or very weak for some components even though they were demonstrated. For example, side stabilization in scenario 5.

## 7. Discussion

Let us revisit again the problem posed in the introduction where we have selected a taxonomy category for performing a given task. A demonstration of the grasp now emphasizes the important grasp components that are cooperating. Using the method described in



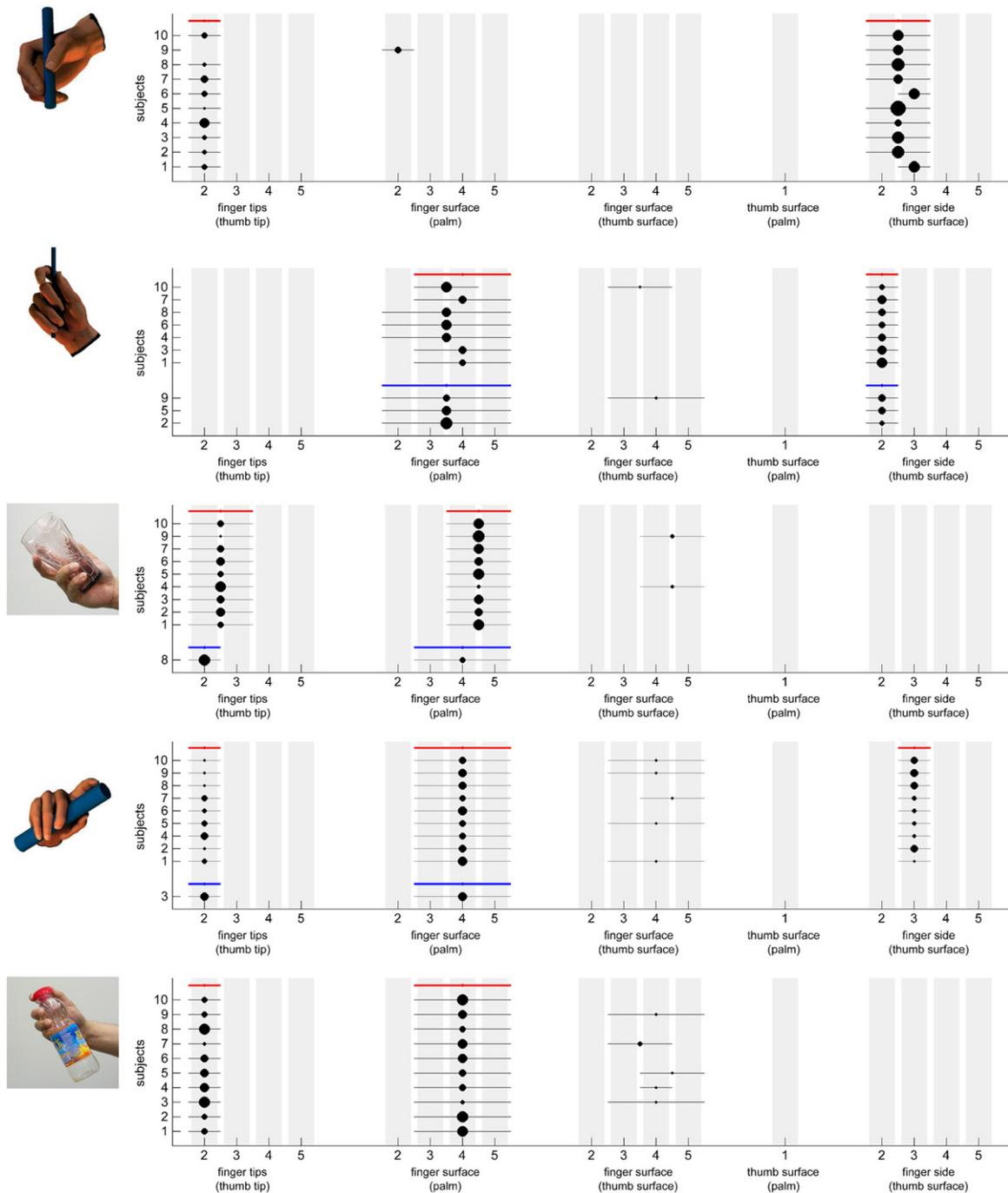
**Fig. 15.** Recognized signatures for scenarios 1–5 (Table 1) across all 10 subjects. Based on a visual analysis of the taxonomy category an expected signature is identified for each grasp scenario. This is indicated in red. For some grasp scenarios, certain subjects executed variations on the communicated intention. These were noted at the time of demonstration are indicated in blue. Results are grouped according to the variation demonstrated. See Fig. 11 for an explanation on how to interpret the plots. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

this paper, this can be recognized for a wide range of grasps in the space of human prehensile ability. The grasp signature becomes input to an Opposition Space controller which does not focus on achieving similar configuration nor contacts but recreates the relevant oppositions in order of the importance that was demonstrated. In response to perturbations or task demands, the controller can apply pressure purposefully by emphasizing the oppositions that comprise the grasp. When adapting the grasp to an object with different properties the problem lies with positioning oppositions appropriately. A grasp controller focuses on finding configuration and contacts to serve this high level intention.

To map the grasp to a robot hand, a correspondence problem must be solved. This can be tackled at the level of Opposition Space. We first define an Opposition Space for the robot hand under con-

sideration and then establish a mapping between the primitives of the human hand to that of the robot hand. This mapping may impose additional constraints on the real finger groupings that are allowed (Section 4.2). These constraints can be used to find a new set of primitives which map one-to-one to the robot. Thus the proposed method discovers a signature in terms of the robot hand under consideration.

The method proposed a new way to look at raw information from a grasp demonstration (configuration, tactile) by combining them into an intermediate patch-level opposition (PLO) representation. This representation highlights different roles of a single sensor patch in the grasp which otherwise get overlapped when raw information for patch is considered alone. It was shown to be useful in identifying the most likely primitives present in the



**Fig. 16.** Recognized signatures for scenarios 6–10 (Table 1) across all 10 subjects. Results are grouped according to grasp intention demonstrated. Communicated intention represented in red, variations on this in blue. See Fig. 11 for an explanation on how to interpret the plots. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

grasp. Seen objectively, the PLO is a 144 dimensional feature. In future, data-driven models could be used to capture PLO correlation with other components of a task scenario namely, object properties, force/torque and motion. In conjunction with the method proposed in the paper, this allows for automatic generation of grasp signature in response in task requirements.

The present work is limited to static grasping. However hand-parts become more actively engaged and the importance of grasp components becomes relevant only when actually performing a task. Examining recognition over an entire task duration is also a direction for further work. The signatures detected may be used as a means to segment task sequences based on the grasp employed. Also, a set of grasp signatures detected over several trials and subjects can be used as basis for grasp adaptation.

## 8. Conclusion

In a grasp demonstration, individual hand-parts are employed so as to leverage their particular functional qualities in order to provide overall grasp function in terms of generating and controlling forces, torques and motion. This is termed as the *grasp intention*. Modeling the grasp intention provides ways to relate the decisions underlying grasp formation to hand function in terms of quantities which may be sensed/recognized or controlled. Further, it allows for a demonstrated grasp to be reconstructed or adapted while preserving its functional properties.

This paper characterizes *grasp intention* using the concepts of Opposition Space proposed in [4], where virtual fingers in opposition are used to form a set of opposition primitives. These con-

cepts were expressed in a manner suitable for a demonstration framework. Extensions were proposed to cover the additional oppositional roles assumed by the thumb frequently encountered in everyday grasps. A general method was proposed by which the specific combination of primitives present in a grasp demonstration could be identified. Using a single expert demonstrator, scenarios testing the same grasp intention expressed on different objects, and different grasp intentions on the same object were recognized successfully. A recognition rate of 87% was achieved with multiple naive demonstrators over a wide range of categories taken from a grasp taxonomy.

## Acknowledgments

This research was funded by a doctoral grant (SFRH/BD/51071/2010) from the portuguese Fundação para a Ciência e a Tecnologia, by the Swiss National Science Foundation through the National Centre of Competence in Research (NCCR) in Robotics and by EU project POETICON+.

## References

- [1] L. Biagiotti, F. Lotti, C. Melchiorri, G. Vassura, *How Far Is the Human Hand? A Review on Anthropomorphic Robotic Endeffectors*, University of Bologna, Italy, 2015.
- [2] J.T. Belter, A.M. Dollar, *Performance characteristics of anthropomorphic prosthetic hands*, in: IEEE International Conference on Rehabilitation Robotics, Department of Mechanical Engineering and Materials Science, Yale University, New Haven, CT 06511, USA, IEEE, 2011, pp. 1–7.
- [3] M.R. Cutkosky, *On grasp choice, grasp models, and the design of hands for manufacturing tasks*, IEEE Trans. Robot. Autom. 5 (3) (1989) 269–279.
- [4] T. Iberall, *The nature of human prehension: three dexterous hands in one*, in: IEEE Int. Conf. on Robotics and Automation, vol. 4, IEEE, 1987, pp. 396–401.
- [5] S.B. Kang, K. Ikeuchi, *Robot task programming by human demonstration: mapping human grasps to manipulator grasps*, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, vol. 1, 1994.
- [6] M. Santello, M. Flanders, J.F. Soechting, *Postural hand synergies for tool use*, J. Neurosci. 18 (23) (1998) 10105–10115.
- [7] H. Ben Amor, O. Kroemer, U. Hillenbrand, G. Neumann, J. Peters, *Generalization of human grasping for multi-fingered robot hands*, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, IEEE, 2012, pp. 2043–2050. URL <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6386072>.
- [8] U. Prieur, V. Perdureau, A. Bernardino, *Modeling and planning high-level in-hand manipulation actions from human knowledge and active learning from demonstration*, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2012, pp. 1330–1336.
- [9] R.S. Johansson, J.R. Flanagan, *Coding and use of tactile signals from the fingertips in object manipulation tasks*, Nat. Rev. Neurosci. 10 (5) (2009) 345–359.
- [10] N. Kamakura, M. Matsuo, H. Ishii, F. Mitsuboshi, Y. Miura, *Patterns of static prehension in normal hands*, Amer. J. Occup. Therapy 34 (7) (1980) 437–445.
- [11] H. Kjellstrom, J. Romero, D. Kragic, *Visual recognition of grasps for human-to-robot mapping*, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, 2008.
- [12] H. Friedrich, V. Grossmann, M. Ehrenmann, O. Rogalla, R. Zollner, R. Dillman, *Towards cognitive elementary operators: grasp classification using neural network classifiers*, in: International Conference on Intelligent Systems and Control, 1999.
- [13] S. Ekvall, D. Kragic, *Grasp Recognition for Programming by Demonstration*, in: IEEE International Conference on Robotics and Automation, 2005.
- [14] J. Aleotti, S. Caselli, *Grasp recognition in virtual reality for robot pregrasp planning by demonstration*, in: IEEE International Conference on Robotics and Automation, 2006.
- [15] K. Bernardin, K. Ogawara, K. Ikeuchi, R. Dillmann, *A sensor fusion approach for recognizing continuous human grasping sequences using hidden Markov models*, IEEE Trans. Robot. 21 (1) (2015).
- [16] P.H. Thakur, A.J. Bastian, S.S. Hsiao, *Multidigit movement synergies of the human hand in an unconstrained haptic exploration task*, J. Neurosci. 28 (6) (2008) 1271–1281.
- [17] M.T. Ciocarlie, P.K. Allen, *Hand posture subspaces for dexterous robotic grasping*, Int. J. Robot. Res. 28 (7) (2009) 851–867.
- [18] A. Bernardino, M. Henriques, N. Hendrich, J. Zhang, *Precision grasp synergies for dexterous robot hands*, in: ROBOT, 2013.
- [19] K. Murakami, K. Matsuo, T. Hasegawa, R. Kurazume, *A decision method for placement of tactile elements on a sensor glove for the recognition of grasp types*, IEEE/ASME Trans. Mechatronics 15 (1) (2015).
- [20] D.R. Faria, R. Martins, J. Lobo, J. Dias, *Extracting data from human manipulation of objects towards improving autonomous robotic grasping*, Robot. Auton. Syst. 60 (3) (2012) 396–410.
- [21] T. Iberall, *Human prehension and dexterous robot hands*, Int. J. Robot. Res. 16 (3) (1997) 285–299.
- [22] M.A. Arbib, T. Iberall, D. Lyons, *Coordinated control programs for movements of the hand*, Exp. Brain Res. 10 (1985) 111–129.
- [23] T. Iberall, G. Bingham, M.A. Arbib, *Opposition space as a structuring concept for the analysis of skilled hand movements*, Exp. Brain Res. 15 (1986) 158–173.
- [24] T. Iberall, J. Jackson, L. Labbe, R. Zampano, *Knowledge-based prehension: Capturing human dexterity*, in: IEEE International Conference on Robotics and Automation, Philadelphia, 1988, pp. 82–87.
- [25] T. Iberall, G. Sukhatme, D. Beattie, G. Bekey, *On the development of EMG control for a prosthetic hand*, in: IEEE International Conference on Robotics and Automation, 1994.
- [26] E. Oztop, N.S. Bradley, M.A. Arbib, *Infant grasp learning: a computational model*, Exp. Brain Res. 158 (4) (2004) 480–503.
- [27] E. Oztop, M.A. Arbib, *Schema design and implementation of the grasp-related mirror neuron system*, Biol. Cybernet. 87 (2) (2002) 116–140.
- [28] S.B. Kang, K. Ikeuchi, *Toward automatic robot instruction from perception-recognizing a grasp from observation*, IEEE Trans. Robot. Autom. 9 (4) (2015).
- [29] Cyberglove, URL: [www.cyberglovesystems.com](http://www.cyberglovesystems.com).
- [30] TekScan, URL: [www.tekscan.com](http://www.tekscan.com).
- [31] R.L. De Souza, S. El Khoury, J. Santos-Victor, A. Billard, *Towards comprehensive capture of human grasping and manipulation skills*, in: 13th International Symposium on the 3-D Analysis of Human Movement, Lausanne, Switzerland, 2014.
- [32] T. Feix, R. Pawlik, H. Schmiedmayer, J. Romero, D. Kragic, *A comprehensive grasp taxonomy*, in: Robotics, Science and Systems: Workshop on Understanding the Human Hand for Advancing Robotic Manipulation, 2009. URL: <http://grasp.xief.net>.
- [33] R. Murray, S. Sastry, Z. Li, *A Mathematical Introduction To Robotic Manipulation*, vol. 29, CRC Press, 1994.



**Ravin de Souza** did his undergraduate degree in Computer Engineering from the University of Mumbai in 1997 and an M.S. in Computer Science from Duke University, North Carolina in 1999. He is currently enrolled in the joint doctoral program between IST and EPFL. His research interests are focused on grasping and manipulation.



**Sahar El-Khoury** is a postdoctoral fellow at EPFL (Switzerland). She has an Electrical Engineering Degree from Ecole Supérieure d'Ingénieurs de Beyrouth in 2004. She received her M.Sc. and PhD degree in Robotics and Intelligent Systems from Université Pierre et Marie Curie (Paris 6) in 2005 and 2008. Her research interests have been focused on grasp planning and manipulation.



**José Santos-Victor** (S'86-A'88-M'00) received the Ph.D. degree in electrical and computer engineering from the Instituto Superior Técnico (IST), Lisbon, Portugal, in 1995. He is currently Professor in the Department of Electrical and Computer Engineering, IST, and director of the Computer and Robot Vision Laboratory (VisLab), Institute of Systems and Robotics (ISR). He is also responsible for the participation of the IST in various European and National research projects in the areas of computer vision and robotics. His current research interests include areas of computer and robot visions, particularly in the relationship between visual perception and the control of action, biologically inspired vision and robotics, cognitive vision and visual controlled (land, air, and underwater) mobile robots. Dr. Santos-Victor was an Associate Editor of the IEEE Transactions on Robotics.



**Aude Billard** is Professor of Micro and Mechanical Engineering, and the head of the LASA Laboratory at the School of Engineering at the Swiss Federal Institute of Technology in Lausanne. She received a M.Sc. in Physics from EPFL (1995), a M.Sc. in Knowledge-based Systems (1996) and a Ph.D. in Artificial Intelligence (1998) from the University of Edinburgh. She was the recipient of the Intel Corporation Teaching award, the Swiss National Science Foundation career award in 2002, the Outstanding Young Person in Science and Innovation from the Swiss Chamber of Commerce and the IEEE-RAS Best Reviewer award in 2012. Aude Billard served as an elected member of the Administrative Committee of the IEEE Robotics and Automation society (RAS) for two terms (2006–2008 and 2009–2011) and is the chair of the IEEE-RAS Technical Committee on Humanoid Robotics. Her research interests focus on machine learning tools to support robot learning through human guidance. This extends also to research on complementary topics, including machine vision and its use in human-robot interaction and computational neuroscience to develop models of motor learning in humans.