A Framework for Recognizing Grasps

Sing Bing Kang and Katsushi Ikeuchi

CMU-RI-TR-91-24

Robotics Institute
Carnegie Mellon University
Pittsburgh, Pennsylvania 15213

November 1991

Contents

List of Figures

List of Tables

Abstract

Many of the tasks that are potential candidates for automation involve grasping. We are interested in the programming of robots to perform grasping tasks. To do this, we propose the notion of "perceptual programming," where the key idea is to enable a system to observe a human performing a task, understand it, and perform the task with minimal human intervention. This allows the programmer to easily choose the grasp strategy.

A grasping task is composed of three phases: pre-grasp phase, static grasp phase, and manipulation phase. The first step in recognizing a grasping task is identifying the grasp itself (within the static grasp phase).

We propose to identify the grasp by mapping the low-level hand configuration to increasingly more abstract grasp descriptions. To this end, we introduce a grasp representation called the *contact web* which is composed of a pattern of effective contact points between the hand and the object. We also propose a grasp taxonomy based on the contact web to systematically identify a grasp. Results from grasping experiments show that it is possible to distinguish between various types of grasps.

Chapter 1 Introduction

Robot programming is an essential component of task automation. The current methods for robot programming include teaching (e.g., [1], [2]), textual programming (e.g., [3], [5]), and automatic programming (e.g., [4], [6], [8], [11]). The first two methods are by far the most pervasive in both the industrial and academic environments. In teaching methods, the robot or manipulator learns its trajectory either through a teach pendant or actual guidance through the sequence of operations ("teach-by-guiding" or less appropriately "teach-by-showing"). This method is the easiest to use since the implicit knowledge of the task is not necessary. On the other hand, because "teach-by-showing" involves some degree of repetition owing to errors, it can be tiring, and possibly risky. Furthermore, this method is not easily transferable to a different system. Textual programming, while flexible, requires expertise and often a long development time. Several of the robot programming languages include AL, AML, RAIL, RPL, and VAL. A summary of these programming languages can be obtained from [5]. These problems can be alleviated by automatic programming, where conceptually the only inputs to the robot system required for generating the control command sequences are the description of the objects involved in the task, and the task specifications. However, realization of a practical system with automatic programming is difficult since important issues remain relatively unresolved in a satisfactory manner. Such issues include: How does one generate a sequence of operations? How can tasks be described unambiguously? If a task involves grasping, how can a stable grasp be effected - should it be optimal from the "human" point of view or a purely analytic point of view?

Although these problems exist for traditional approaches to task programming, we could avoid such problems by using a different approach. We are particularly interested in programming the robot to perform grasping tasks. Because most tasks that are performed by humans, especially manufacturing tasks [18], involve grasping, automation of such tasks would certainly involve knowledge of grasps. This makes the analysis of grasps and their purposes important. Much work on grasps concentrates on grasp synthesis, i.e., the determination of the "optimal" grasp given knowledge about the objects and the task. Either the task specifications need to be explicitly enumerated [6], or the grasp chosen is to be optimal according to some grasping quality metrics [7], or the grasp chosen is to be stable (from the "human" point of view [8] or from the analytic point of view [9]). The issues of task specification and grasping strategy are complicated and difficult.

The task programming approach that we propose to adopt is *perceptual programming*. In perceptual programming, task programming is performed by demonstrating the task to the system rather than by the traditional method of hand-coding. The realization of perceptual programming would entail the understanding of hand grasping motions. The key idea is to enable a system to do the following: observe a human performing a task, understand it, and perform the task with minimal human intervention. Perceptual programming would obviate the need to explicitly describe the required task, since the system is able to understand the task based on observation of the task performance by a human.

Work in this area would result in a greater understanding of grasping motions, to the extent that recognition by a robotic system would be possible. The areas in which this body of knowledge is potentially useful include planning and automation, and teleoperation.

The ideas and goals of our work are very similar to those that embody the Assembly Plan from Observation (APO) paradigm proposed by Ikeuchi and Suehiro [10]. They describe a system that observes a human performing an assembly task while a geometric reasoner analyzes and recognizes the task from observation, and generates the same assembly sequence for a robot. In this paradigm, the human operator does all the thinking - the system "understands" what needs to be done based on what is observed and performs the task or tasks. A similar approach was taken by Kuniyoshi et al. [11] who developed a system which emulates the performance of a human operator. However, their system is restricted to pick-and-place operations.

This report is an account of our work on grasp identification. In Chapter 2, we describe a 3-D structure called the *contact web* which can be used to classify a grasp. It also facilitates higher level descriptions of a grasp by using a special objective function. The experiments conducted to justify the uses of the contact web and the objective function are described along with their results in Chapter 3. We summarize our findings in Chapter 4.

Chapter 2 Grasp Identification

Grasp identification is central to the recognition of grasping tasks. In order to identify grasps, we need a hand model, which is explained in the next subsection.

2.1 Hand Model

An articulated hand model whose finger movements closely resemble those of an actual hand is used to infer the grasp observed in the scene. The human hand is depicted in Figure 1.
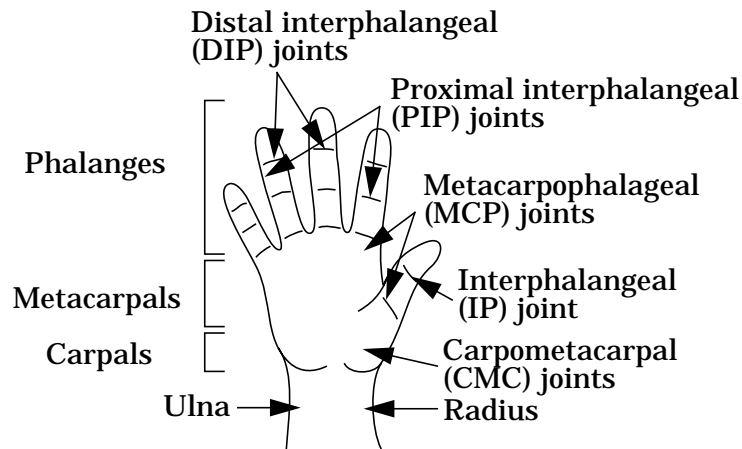
Figure 1 **Bones and joints of the human hand (adapted from [23]).**

Each finger is approximated to have four degrees of freedom. The four angular parameters associated with finger movement (except the thumb) are directly associated with the degree of finger abduction and the degrees of flexion at the metacarpophalangeal, proximal interphalangeal and distal interphalangeal joints (please refer to Figure 1). For the thumb, they are the angular flexions in the carpometacarpal joints (two parameters), the metacarpophalangeal joint and the interphalangeal joint. In this hand model, limits of these angular parameters which are consistent with anatomical and physiological studies of the hand (e.g., [12], [13], and [14]) are imposed. Flexion angles are defined with respect to the hand frontal plane[1] while the abduction angles are defined with respect to the sagittal planes[2].

Finger segments are modeled by cylinders. Buchholz and Armstrong [28] model hand segments with ellipsoids for ease of analytic determination of contact points between the hand and held object, which is also modeled by an ellipsoid. The ellipsoid-ellipsoid con-

---

1. The frontal plane is the plane parallel to a flat hand with fingers extended.

2. The reference sagittal plane of a finger is the plane perpendicular to the frontal plane passing through the principal long axis of the fully adducted and extended finger.

tact algorithm allows modelling of soft tissue penetration. While the ellipsoidal modelling is suitable for simulation, it is less suitable for tracking.

2.2 Classification of Grasps

There has been a lot of study in the medical community on the grasping capabilities of the human hand, from the anatomical and functional points of view. Schlesinger [15] and Taylor and Schwarz [14] associate human grasps primarily with the object shape in their categorization of six grasps (cylindrical, fingertip, hook, palmar, spherical and lateral). Griffiths' [32] grasp classification is also based on objects of varying form. He partitions the functions of the hand into cylinder grasp, ball grasp, ring grasp, pincer grasp and plier grasp. McBride [33] took a different approach in dividing the function of the hand: his classification depends on the parts of the hand which participate in the grasp (grasp with the whole hand, grasp with thumb and fingers, and grasp with finger and palm). These classifications , while expressive and intuitively informative, do not reflect a fundamental analysis of the hand as an entity. They are either too dependent on the shape of the held object ([14], [15], [32]), or arbitrary without any particular functional basis ([33]).

Napier [26], on the other hand, dichotomized grasps into precision grasps and power grasps[1]. His classification of grasps is based on the purpose of the task, shape and size of the object, and the posture of the fingers. This division of grasps into precision and power grasps is the most widely accepted and used by researchers in the medical, biomechanical and robotic fields. A power grasp is used for higher stability and security at the expense of object maneuverability, while the converse is true for a precision grasp. A precision grasp is characterized by a small degree of contact between the hand and the object. In this type of grasp, the object is normally pinched between the thumb and the flexor aspects of at least one finger. In a power grasp, however, the object is held tight by the fingers and the palm[2]. The major classifications of a power grasp are the cylindrical power grasp and the spherical power grasp. In a cylindrical power grasp, the thumb can either be adducted for some element of precision, or abducted for more clamping action on the object. Henceforth the cylindrical power grasp refers to the former type while the "coal-hammer" grasp refers to the latter type. Sollerman [29] uses a different terminology for power grasps; he refers to the cylindrical power grasp, "coal-hammer" grasp, and spherical power grasp as the diagonal volar grasp, transverse volar grasp, and spherical volar grasp respectively.

Cutkosky and Wright [16] construct a hierarchical tree of grasps beginning with Napier's distinction between precision and power grasps. At the lowest level, a grasp is chosen based on object geometric details and task requirements. However, not only is the taxonomy incomplete, but because the grasp classification is discrete, there may exist problems in categorizing grasps in intermediate cases (e.g., the shape of the object is somewhere between being strictly prismatic and strictly spherical). In these cases, determination of

---

1. He actually referred to them as precision grips and power grips. The term "grasp" is used throughout this document for consistency.

2. An exception is the lateral pinch, which is the grasp employed when turning a key in a lock. This grasp involves the thumb and the radial side of the index finger.

the type of grasp will then be dependent mostly on human judgment rather than on reasoning.

In our effort to automate the recognition of grasps, we require a grasp taxonomy which could provide a systematic way of identifying grasps based on the hand configuration and object shape. Cutkosky and Wright's grasp taxonomy is not suitable for use in our work because, in addition to its limitations mentioned above, it presumes a priori knowledge of the task requirements which are not available in our problem domain. We propose a grasp taxonomy based on the analysis of the effective contact points of the hand with the grasped object. The effective contact point of a finger segment represents the surface contact of that segment with the object. The resultant spatial pattern of contact points forms what we call a *contact web*.

2.3 The Contact Web

2.3.1 Definitions

A *contact web* is defined as a 3-D graphical structure connecting the effective points of contact between the hand and the object grasped. When parts of a finger or palm make contact with the grasped object, the actual contact area is finite. A point contact is useful in representing the contact between the phalangeal segments and palm, and the object because of ease of representation and analysis, and accommodation of uncertainty in grasping. The shape and cardinality of the contact web yield important information about the type of grasp effected.

Intradigital contact points are contact points along the same finger. *Interdigital contact points* are those located at different fingers. The contact notation adopted is illustrated in Figure 2.

P_I^i is the intradigital contact point set for the ith finger ($i = 0$ (the thumb), 1, 2, 3, 4);

E.g., $P_I^1 = \{C_{13}\}$ refers to the finger tip contact point set of the index finger.

$P_C = \{\text{fingers in contact with object}\}$;

$P_H = \{ P_I^i : i \in P_C \}$ (the contact point set);

$N_0(P_H) = $ cardinality of $P_H = $ number of fingers in contact with object;

$$N_1(P_H) = N_0 \left(\bigcup_{i \in P_C} P_I^i \right) = \text{total number of contact points.}$$

Note: $N_1 (P_H) \big|_{max} = 15$.
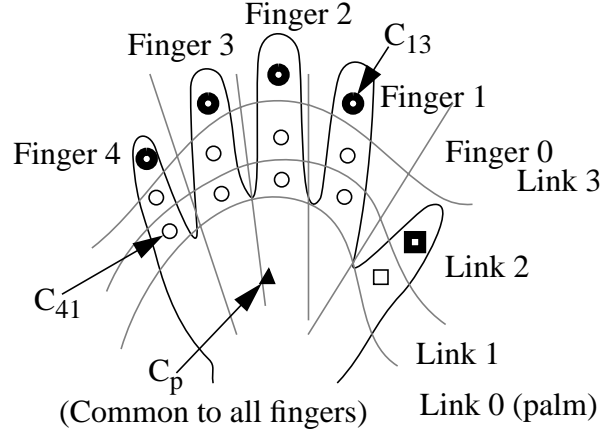
Figure 2 **Contact Notation on the right hand (palmar side)**

2.3.2 A Taxonomy based on the Contact Web

Cutkosky and Wrights' taxonomy [16] provides a systematic way of finding a grasp that will satisfy a particular set of task requirements and a particular object shape. However, such a taxonomy is difficult to use if (as in our case) the task requirements are not known a priori. We propose a grasp taxonomy which is based on the contact web. It provides a systematic way of recognizing grasps from the hand configuration and the object shape. In addition, it provides a more continuous classification of grasps by not restricting the classification to discrete grasp groups.

Figure 3 shows the proposed major classifications of grasps and the type of contact web associated with each major grasp group. The contact web provides an easy, compact and convenient framework for describing grasps, giving insights into the type and shape of the object grasped, and possibly, the tasks involved. Note that the contact point of the palm is merely an artifact which represents the effective contact of the palm with the grasped object. This point may or may not be physically in contact with the object.

```
                          Grasps
                            │
          ┌─────────────────┴─────────────────┐
   Volar Grasps                      Non-Volar Grasps
   (Non-Planar                              │
   Contact Web)                   ┌─────────┴─────────┐
                          Fingertip Grasps       Composite
                                             Non-Volar Grasps
                             (Planar            (Non-Planar
                             Contact Web)        Contact Web)
```
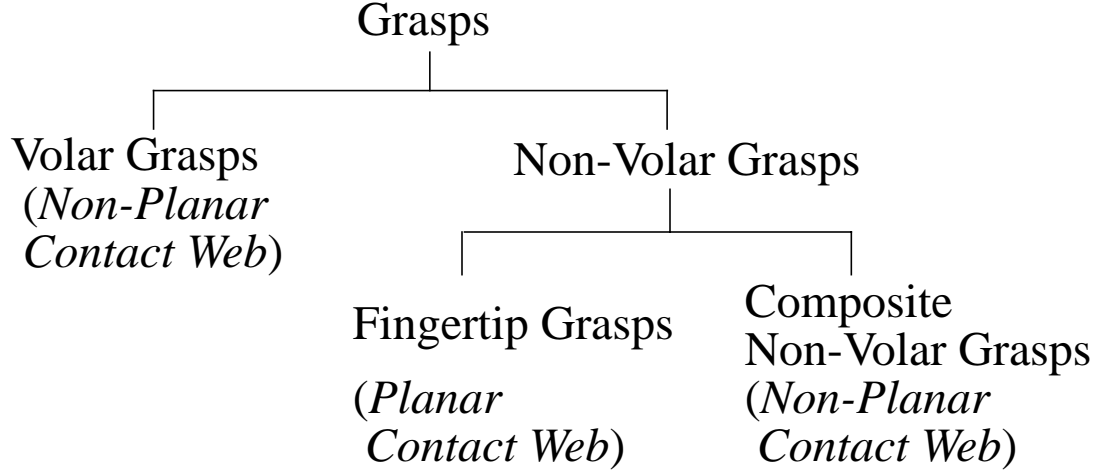
Figure 3 **Major classifications of grasps for recognition**

A power grasp is characterized by a high degree of contact between the hand and the held object. This allows high clamping forces on the object. A feature that we use to first distinguish between grasps is the involvement of the palm surface in the grasp. Grasps which involve the palm surface are called *volar grasps* while others are called *non-volar grasps*. All volar grasps are power grasps. All but one type of non-volar grasps are precision grasps[1]. The exception mentioned is the lateral pinch, which is the grasp assumed by the hand when turning a key in a lock. Although the lateral pinch does not involve the palm surface, it emphasizes on the security of the object rather than its dexterity; hence it is a power grasp. This is the reason why volar and non-volar grasps are not equivalent to power and precision grasps respectively, and are not labeled as such in our taxonomy. The non-volar grasps are further classified as *fingertip grasps* and *composite non-volar grasps*. The fingertip grasp involves only the fingertips while the composite non-volar grasp involves surfaces of other segments of the fingers in addition to the fingertips. The major grasp classifications are shown in Figure 3, while the effective contact point notation is depicted in Figure 2.

One interesting feature of a category of grasps is whether the contact web associated with that category is planar or non-planar. The contact web formed by a volar grasp is spatially non-planar (except for the platform push[2]). In most non-volar grasps where the areas of contact between the object and the hand are those of the fingertips (fingertip grasps), the associated contact web is approximately planar. However, there are at least two identifi-

---

1. Napier [26] regards volar grasps as power grasps and non-volar grasps as precision grasps. This view is shared by various other researchers in the medical and biomechanical fields (e.g., [28], [31]). However, this would not be *strictly* true if we adhere to the position that the type of grasp should be classified according to the predominance of either power or precision in the grasp (which, interestingly, Napier [26] adheres to as well). The demand for power is higher than the demand for precision in the lateral pinch; though it is a non-volar grasp, it is, by the power predominance definition, a power grasp. See text.

2. The platform push is a *non-prehensile* grasp; non-prehensile grasps are not considered in this work.

able cases of non-volar grasps where the contact web is non-planar, namely the lateral pinch and the pinch grasp. These are separately grouped as composite non-volar grasps.

The contact web enables a more continuous categorization of grasps as shown in Figure 4 and Figure 5. The degree of membership to a strictly prismatic grasp or spherical/disc grasp lies in the degree of fit of the contact points to the respective shapes. In addition, the contact web facilitates a mathematical framework for the recognition of grasps as described in Subsection 2.5. Finally, as discussed in this subsection, this grasp taxonomy provides a systematic means of grasp discrimination from observation.

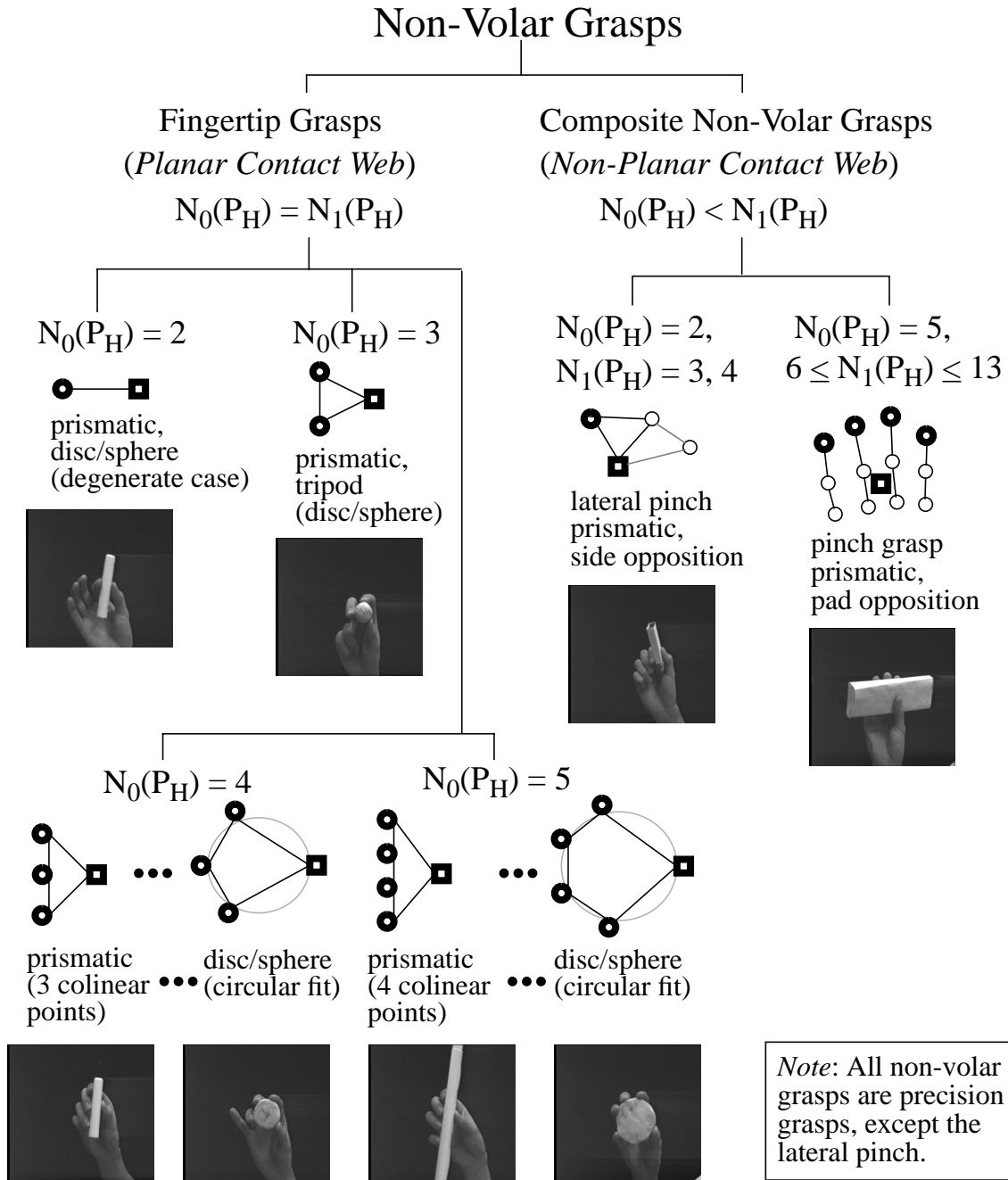### 2.3.3 Comparisons with Other Grasp Frameworks

Cutkosky discusses how to choose grasps based on task requirements and object shape, and he implemented an expert system called "Grasp-Exp" to do this [18]. Cutkosky and Howe [19], in a similar vein, relate grasp attributes such as dexterity and precision to analytic measures such as manipulability and isotropy in their grasp analysis. Iberall and MacKenzie [23] concentrate on finding a grasp solution for the controller given anticipated object properties and predictable interaction outcome in terms of opposition space and virtual fingers. Iberall [24] describes a neural network that maps qualitative task and object properties (surface length in terms of finger span, object width, amount of forces, and level of task precision) onto a desired prehensile posture. The mapping is based on empirical evidence. In this paper, the posture is restricted to possess only one opposition vector (either pad or palm opposition).

In each of these cases, an explicit analytical framework is not provided to *identify* a given grasp. In other words, they do not answer questions such as: Given a task scene, what grasp is being used? On a lower level, what opposition space is present and what is the number and composition of the available virtual fingers? On a higher level, what task do the grasping actions achieve? In our scenario, object description is available while task description is not available. The system has to somehow be able to determine what grasping strategy has been employed based on viewing a multiple sequence of the task.

Nguyen and Stephanou [25] describe a topological algorithm for continuous grasp planning. Their paper proposes a topological algorithm to determine a grasp expressed in terms of low-level joint variables, given a high-level task description. Again, the assumptions made and the domain are different. In our framework, no task description is available. Instead, it is inferred.

Pao and Speeter [30] determine the matrix transformation linking the joint parameters of the human hand and a multifingered robotic hand based on a predefined set of hand poses. The robotic hand posture corresponding to any given hand configuration is then interpolated using this matrix transformation. There is no higher level of abstraction in their method as correspondence between the human and robotic hand configurations is established based on low-level joint angles. As a result, it may be difficult to generalize this scheme to less anthropomorphic robotic hands and for complicated tasks.

Our framework provides a direct mathematical means to use higher-level conceptual terms for describing the physical configuration of the hand. As mentioned earlier, it also provides a more continuous taxonomy of grasps.

Non-Volar Grasps

Fingertip Grasps
(*Planar Contact Web*)
$N_0(P_H) = N_1(P_H)$

$N_0(P_H) = 2$

prismatic, disc/sphere (degenerate case)

$N_0(P_H) = 3$

prismatic, tripod (disc/sphere)

Composite Non-Volar Grasps
(*Non-Planar Contact Web*)
$N_0(P_H) < N_1(P_H)$

$N_0(P_H) = 2$, $N_1(P_H) = 3, 4$

lateral pinch prismatic, side opposition

$N_0(P_H) = 5$, $6 \leq N_1(P_H) \leq 13$

pinch grasp prismatic, pad opposition

$N_0(P_H) = 4$

prismatic (3 colinear points)

disc/sphere (circular fit)

$N_0(P_H) = 5$

prismatic (4 colinear points)

disc/sphere (circular fit)

Note: All non-volar grasps are precision grasps, except the lateral pinch.

Figure 4 **Classification of non-volar grasps**

Volar Grasps (*Non-Planar Contact Web*)

$N_1(P_H) = 15$ $N_1(P_H) < 15$

planar/almost planar
(degenerate case)
platform push

prismatic/
cylindrical

••• spherical

Note: All volar
grasps are
power grasps
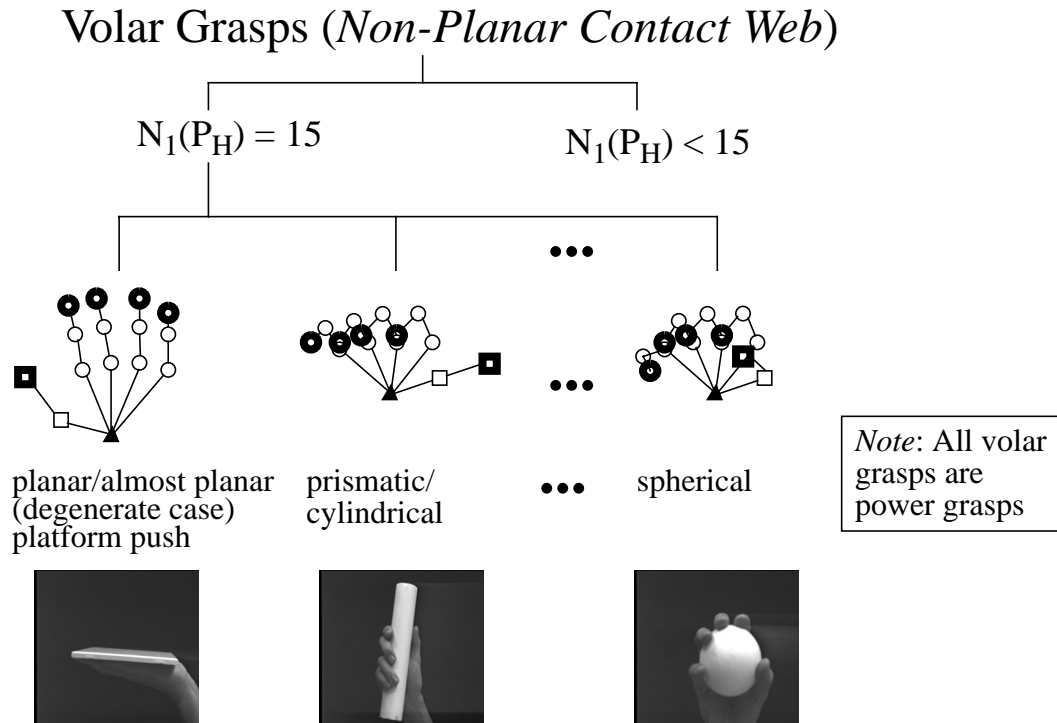
Figure 5 **Classification of volar grasps**

2.4 Virtual Fingers and Opposition Space

By analyzing the contact web, medium level grasp concepts such as virtual fingers and opposition space, can be described. These two concepts, in turn, are the key elements in characterizing the type of grasp and indicating the functionality of the grasp.

Arbib et al. [20] introduced the concept of the virtual finger: a functional unit which comprises at least one real physical finger (which may include the palm). The real fingers composing a virtual finger act in unison to apply an opposing force on the object and against the other virtual fingers in a grasp. This concept replaces the analysis of the mechanical degrees of freedom of individual fingers by the analysis of the functional roles of forces being applied in a grasp.

Cutkosky and Howe [19] suggest that virtual fingers correspond to independently controlled contact sites, and oppositions correspond to internal grasp forces. While Iberall [22] indicates that the precision tripod grasp consists of two virtual fingers (thumb as one and the index and third fingers as the other), Cutkosky and Howe [19] state that it may have either two or three virtual fingers, depending on the amount of coupling between the index and third fingers. We agree with this view. Finger force interaction is the basis for virtual finger composition determination in our work.

Iberall et al. [21] define opposition space as "the area within the coordinates of the hand where opposing forces can be exerted between virtual finger surfaces in effecting a stable

grasp." They show that prehensile grasps involve combinations of the three basic opposi-tions shown in Figure 6. Pad opposition is used in precision grasps, while palm opposition provides a more powerful hold on the object. Finally, side opposition is a compromise between these two oppositions in terms of dexterity and strength of the grasp. Opposition space is an important concept in characterizing grasps.

(a)　　　　　(b)　　　　　(c)
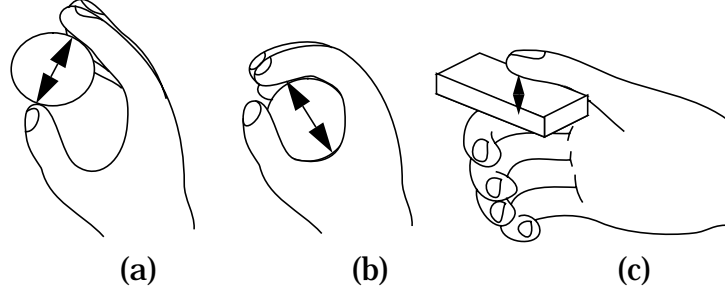
Figure 6 **Types of opposition (adapted from [23]). (a) Pad opposition, (b) Palm opposition, (c) Side opposition**

2.5 Recognizing Grasps from the Contact Web

We now illustrate how the contact web can be used to identify the grasp. As mentioned earlier, there are three different types of opposition: pad opposition, palm opposition, and side opposition. We start with the simplest type of opposition, namely, pad opposition, and then proceed to side opposition. The detailed analyses involving these oppositions in the next two subsections constitute the main ideas embodied in the mathematical framework for grasp recognition.
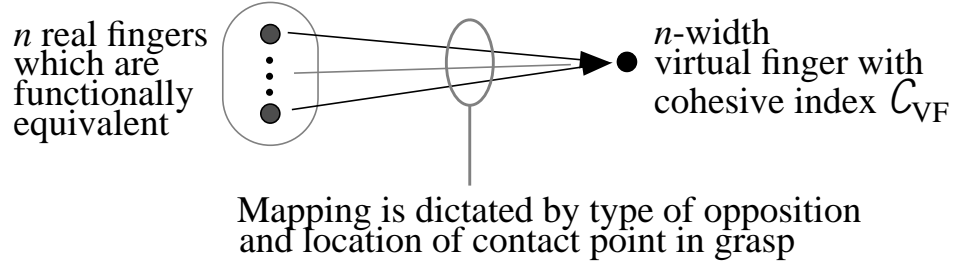
2.5.1 Pad Opposition Only

There are at least two virtual fingers to effect this opposition. The *degree of force coupling* between any two given forces $\mathbf{f}_{i,cpp}$ and $\mathbf{f}_{j,cpp}$ is defined to be their normalized dot product:

$$D_c(i,j) \;=\; \frac{\mathbf{f}_{i,cpp} \bullet \mathbf{f}_{j,cpp}}{\left|\mathbf{f}_{i,cpp}\right|\left|\mathbf{f}_{j,cpp}\right|}$$

The following is a proposed analytical method of determining the mapping of all the fin-gers touching the grasped object into either one, two or three virtual fingers. Note that this method does not presume the mapping of the thumb into one virtual finger.

n real fingers which are functionally equivalent

n-width virtual finger with cohesive index C_{VF}

Mapping is dictated by type of opposition and location of contact point in grasp

Figure 7 **Virtual finger mapping under the influence of opposition space and point contact placement**

This method is based on the premise that the mapping and number of virtual fingers depend on the location and degree of coupling between the fingers in contact with the grasped object. This philosophy is illustrated in Figure 7. It quantifies the degree of coupling between fingers and introduces the concept of the *cohesive index* of a virtual finger. The cohesive index of a virtual finger indicates the degree to which the real fingers mapped into it are functionally equivalent. Let the normal forces on the contact polygon plane (CPP) be denoted as $\mathbf{f'}_{i,cpp}$ (i=1, ..., n), and the actual internal projected forces acting on the object be represented by $\mathbf{f}_{i,cpp}$ (i=1, ..., n). Assume that there are n_{RF} real fingers in contact with the grasped object and that $\mathbf{f'}_{i,cpp}$ are known for i = 1, ..., n_{RF}. Assume also, that $\mathbf{f}_{i,cpp} \approx \mathbf{f'}_{i,cpp}$. The virtual finger membership index between fingers i and j (each with only one contact point for the moment) is defined as:

$$m_{ij} = \frac{\min (|\mathbf{f}_{i,cpp}|, |\mathbf{f}_{j,cpp}|)}{\max (|\mathbf{f}_{i,cpp}|, |\mathbf{f}_{j,cpp}|)} \frac{1 + D_c(i,j)}{2}$$

It can be seen that $0 \leq m_{ij} \leq 1$. Two real fingers are more likely to be members of the same virtual finger if the force vectors at the contact points are similar (i.e., both in terms of force direction and magnitude). Obviously $m_{ii} = 1$ and $m_{ij} = m_{ji}$. Let VF_k denote the set of real fingers hypothetically belonging to the kth virtual finger. Then the cohesive index for that virtual finger is defined as the geometric mean of all the pairwise virtual membership indices:

$$C_{VF,k} = \prod_{\substack{i,j \in VF_k \\ j \geq i}} m_{ij}^{\xi}$$

where

$$\xi = \left(\begin{matrix} N(VF_k) \\ 2 \end{matrix} \right)^{-1}$$

$N(VF_k)$ being the number of real fingers in virtual finger VF_k and ξ is the reciprocal of the number of possible pairs of real fingers in virtual finger VF_k. $C_{vf,k}$ characterizes the similarity of action of the real fingers in VF_k. If all the fingers in VF_k act in unison, i.e., exert

forces equal in magnitude and direction, then $C_{vf,k} = 1$. However, if any two fingers in VF_k exert forces in opposite directions, then $C_{vf,k} = 0$.

The problem of determining the number of virtual fingers and the constituents of each virtual finger can be described as a non-linear mixed program:

Maximize

$$C_{\text{eff}} = \left(\frac{1}{n_{VF}!} \prod_{i=1}^{n_{VF}} C_{VF,i} \right)^{\frac{1}{n_{VF}}}$$

subject to

$$n_{VF} \in \{1, 2, 3\}$$

$$\bigcup_{i=1}^{n_{VF}} VF_i = RF$$

$$VF_i \cap VF_j = \varnothing, \qquad i \neq j; \qquad (1 \leq i, j \leq n_{VF})$$

The product term with the exponent in the objective function C_{eff} is the geometric mean of the membership indices of the hypothesized virtual fingers. This ensures that the division of real fingers into virtual fingers is done in such a way that the real fingers in each virtual finger act on the object in as similar a manner as possible. C_{eff} is called the *grasp cohesive index*. The remaining factor in the objective function is a contrived one to favor a smaller number of virtual fingers should there exist equivalency in the objective function (without this factor) for different hand configurations comprising different numbers of virtual fingers. RF is the set of real fingers in contact with the grasped object. For the following examples, for simplicity, it is assumed that all the forces exerted are of unit magnitude (this assumption does not detract from the basic principle). This assumption is altered somewhat for the experiments which are described in subsequent sections.

Example 1 (see Figure 8)

$$m_{12} = 0$$
$$m_{13} = 0$$
$$m_{23} = 1$$
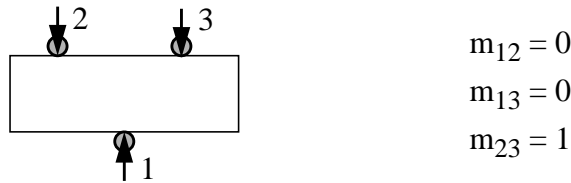
Figure 8 **Illustration for Example 1**

For the simple case of the thumb and two fingers holding a prismatic object in place, the highest value for C_{eff} is obtained for $VF_1 = \{1\}$ and $VF_2 = \{2, 3\}$. $C_{VF,1} = 1$ and $C_{VF,2} = 1$, giving $C_{\text{eff, max}} = 0.707$. For $VF_1 = \{1\}$, $VF_2 = \{2\}$ and $VF_3 = \{3\}$, $C_{\text{eff}} = 0.550$.
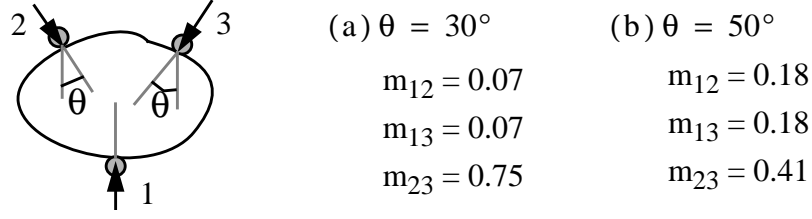
Example 2(a) and (b) (see Figure 9)

(a) $\theta = 30°$

$m_{12} = 0.07$

$m_{13} = 0.07$

$m_{23} = 0.75$

(b) $\theta = 50°$

$m_{12} = 0.18$

$m_{13} = 0.18$

$m_{23} = 0.41$

Figure 9 **Illustration for Example 2(a) and (b)**

In this example, the object held is no longer prismatic but roughly resembles an ellipse. For case (a), the highest value for C_{eff} is again obtained for $VF_1 = \{1\}$ and $VF_2 = \{2, 3\}$ (here $C_{eff, max} = 0.612$). However, in case (b), the highest value for C_{eff} is obtained for $VF_1 = \{1\}$, $VF_2 = \{2\}$, and $VF_3 = \{3\}$ ($C_{eff, max} = 0.550$). (Note that θ is the angle measured with respect to the vertical line to which the force direction at contact point 1 is anti-parallel.)

2.5.2 Side Opposition Only

Side opposition involves two fingers - the thumb and the index finger. Contact points that are part of the same finger (i.e., intradigital contact points) are automatically grouped together. This means that either one or two virtual fingers exist in this type of grasp configuration. The kth "composite" finger comprising l ($l = 2$ or 3) intradigital contact points is denoted by $k \rightarrow \{1, ..., l\}$. The mixing rule employed for the "composite" finger is

$$m_{i, j \rightarrow \{1, ..., l\}} = \frac{1}{2} \left(\prod_{p = 1}^{l} m_{i, j \rightarrow p} + \max_{p \in \{1, ..., l\}} m_{i, j \rightarrow p} \right)$$

This rule makes it more difficult for "composite" fingers to be grouped together as virtual fingers. The steps for determining the number of virtual fingers and their constituents proceed as for the previous examples. An example of how the mixing rule is used is shown in Figure 10.
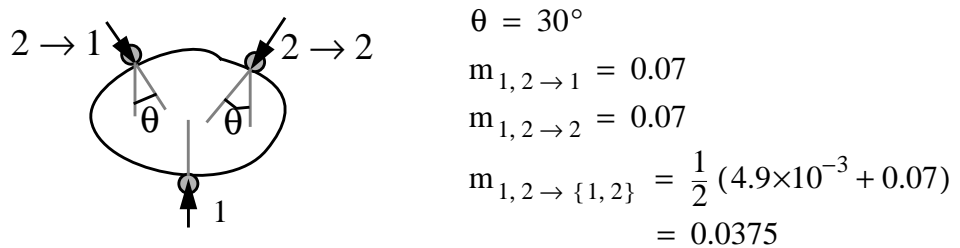

$\theta = 30°$

$m_{1, 2 \rightarrow 1} = 0.07$

$m_{1, 2 \rightarrow 2} = 0.07$

$m_{1, 2 \rightarrow \{1, 2\}} = \frac{1}{2} (4.9 \times 10^{-3} + 0.07)$

$= 0.0375$

Figure 10 **Illustration for mixing rule application**

2.5.3 Hand Configuration Notation

A shorthand notation for the configuration of the hand, say, $VF_1 = \{1\}$, $VF_2 = \{2, 3, 4\}$ is $\{(1)(234)\}$. This notation will be used in the remainder of this report.

2.5.4 General Mixing Rule for "Composite" Fingers

The virtual finger membership index between two "composite" fingers i and j is given by the expression

$$m_{i \rightarrow \{1, ..., l_i\}, j \rightarrow \{1, ..., l_j\}} =$$

$$\frac{1}{2} \left(\prod_{q=1}^{l_j} \max_{p \in \{1, ..., l_i\}} m_{i \rightarrow p, j \rightarrow q} + \prod_{p=1}^{l_i} \max_{q \in \{1, ..., l_j\}} m_{i \rightarrow p, j \rightarrow q} \right)$$

The expression on the right is easily seen to be commutative in "composite" fingers i and j. This mixing rule has a simple physical interpretation: Each contact point on one "composite" finger is only matched to its nearest equivalent (in terms of force vector) and not to all other points in the other "composite" finger. For each contact point of finger i, the largest membership index with finger j is found and then multiplied together; the other term is determined by interchanging these fingers.

We see that this expression reduces to the equation in subsection 2.5.2 when $l_i = 1$. If the two "composite" fingers are fully compatible, i.e., $l_i = l_j = l$ (say) and $m_{i \rightarrow p, j \rightarrow q} = 1$ for p, q = 1, ..., l, then $m_{i \rightarrow \{1, ..., l_i\}, j \rightarrow \{1, ..., l_j\}} = 1$, as to be expected.

Chapter 3 Experiments and Results

3.1 Analysis of Grasps by Human Subjects

3.1.1 Experiments Involving Precision Grasps

Several experiments were conducted to illustrate the use of the mathematical framework for both precision and power grasps. Since force is not measured for these simple experiments, a strong assumption is made here: The force exerted at each finger is the same. This means that the sum of the forces at the contact points of each each finger is equal. So, if there are two contact points at the thumb, then the force at each thumb contact is equal to 1.5 times the force at each contact point on the other fingers. Also, note that only the contact points of the fingers and/or palm with the grasped object are considered in the analysis. Contact points of the finger on the hand itself are not taken into consideration; e.g., thumb contact with the index finger in the "coal-hammer" grasp are disregarded. In practice, these "invalid" contact points can be determined by analyzing the hand configuration.

Table 1 **Description of Experiments involving Precision Grasps**

Experiment #	Description
1	Four fingers and thumb on flat rectangular object (3.0 cm x 16.0 cm)
2	Three fingers and thumb on flat rectangular object (3.0 cm x 16.0 cm)
3	Two fingers and thumb on flat small circular object (diameter = 2.4 cm)
4	Four fingers and thumb on flat circular object (diameter = 5.8 cm)
5	Three fingers and thumb on flat circular object (diameter = 5.8 cm)
6	Three fingers and thumb on flat small right triangle (sides 5.2 cm, 6.7 cm , and 8.4 cm)
7	Four fingers and thumb on flat right triangle (sides 10.0 cm, 8.8 cm, and 13.3 cm)
8	Three or four fingers and thumb on flat elliptical object (a=3.4cm, b=2.0cm)[1]
9	Three or four fingers and thumb on flat elliptical object (a=4.0cm, b=2.0cm)

---

1. a and b are the semi-major and semi-minor axis lengths respectively.

Table 2 **Results of Experiments involving Precision Grasps**

Experiment #	$C_{eff,max}$	$Config_{max}$	$C_{eff,2}$	$Config_2$
1	0.707	{(1)(2345)}	0.550	{(1)(2)(345)}
2	0.707	{(1)(234)}	0.550	{(1)(2)(34)}
3	0.550	{(1)(2)(3)}	0.528	{(1)(23)}
4	0.538	{(1)(2345)}	0.514	{(1)(234)(5)}
5	0.578	{(1)(234)}	0.518	{(1)(23)(4)}
6	0.561	{(1)(234)}	0.550	{(1)(23)(4)}
7	0.595	{(1)(2345)}	0.550	{(1)(2)(345)}
8	0.681	{(1)(234)}	0.545	{(1)(23)(4)}
9	0.680	{(1)(2345)}	0.550	{(1)(234)(5)}

In the tables of results, the numbers 0 and 1 refer to the palm and thumb respectively, while the numbers 2, 3, 4, and 5 refer to the other four fingers in order (2 being the index finger and 5 being the little finger). Table 2 shows the results of the experiments described in Table 1 for a subject. $C_{eff,max}$ is the maximum grasp cohesive index while $C_{eff,2}$ is the second largest grasp cohesive index found for the hand configuration. Experiments 3, 4, and 5 (with flat circular objects) are repeated for five other subjects, and their results are shown in Table 3. As Table 3 indicates, the best cohesive indices do not exhibit a strong consistency in values across different subjects. The optimal hand configuration depends on the manner upon which the object is grasped. For example, in the tripod grasp, if the object is grasped such that the index and middle fingers are separated relatively far apart, as were most of the subjects' hands in the experiments, then they are regarded as separate virtual fingers. If, on the other hand, these fingers are kept close to one another, then they will be grouped as one virtual finger, as for subject 5 (Table 3).

Table 3 **Best Cohesive Indices for Precision Grasps on flat circular objects**

Experiment #	3	4	5
Subject #	$C_{eff,max}$	$C_{eff,max}$	$C_{eff,max}$
0	0.550 {(1)(2)(3)}	0.538 {(1)(2345)}	0.578 {(1)(234)}
1	0.550 {(1)(2)(3)}	0.498 {(1)(2)(345)}	0.497 {(1)(23)(4)}
2	0.550 {(1)(2)(3)}	0.523 {(1)(2345)}	0.526 {(1)(234)}
3	0.550 {(1)(2)(3)}	0.492 {(1)(2)(345)}	0.512 {(1)(23)(4)}
4	0.550 {(1)(2)(3)}	0.483 {(1)(2)(345)}	0.510 {(1)(23)(4)}
5	0.574 {(1)(23)}	0.498 {(1)(2)(345)}	0.514 {(1)(23)(4)}

Table 3 lists the results for Experiments 8 and 9 which involves flat elliptical objects of different eccentricities. In both these experiments, the thumb is considered as one virtual finger and the other fingers as the other virtual finger. While most subjects used five fingers in handling these objects, several of them used only four fingers.

Table 4 **Best Cohesive Indices for Precision Grasps on flat elliptical objects**

Experiment #	8	9
Subject #	$C_{eff,max}$	$C_{eff,max}$
0	0.681 {(1)(234)}	0.680 {(1)(2345)}
1	0.607 {(1)(2345)}	0.655 {(1)(2345)}
2	0.660 {(1)(2345)}	0.680 {(1)(2345)}
3	0.666 {(1)(2345)}	0.684 {(1)(2345)}
4	0.679 {(1)(234)}	0.658 {(1)(2345)}
5	0.664 {(1)(2345)}	0.687 {(1)(2345)}
6	0.620 {(1)(234)}	0.673 {(1)(234)}
7	0.655 {(1)(234)}	0.680 {(1)(2345)}
8	0.663 {(1)(234)}	0.676 {(1)(234)}
9	0.648 {(1)(2345)}	0.676 {(1)(2345)}
10	0.578 {(1)(2345)}	0.670 {(1)(2345)}
11	0.657 {(1)(234)}	0.645 {(1)(2345)}

3.1.2 Experiments Involving Power Grasps

Eight experiments were performed to illustrate the use of the mathematical framework in determining the number and composition of virtual fingers in power grasps. The experiments involved marking the centers of each phalangeal segment and palm, and applying different types of power grasps on rods of different thicknesses and a sphere. Description of the experiments are listed in Table 5. It is assumed, for simplicity of this analysis, that the effective forces at each contact point are normal to the object surface, and that they are equal in magnitude.

The "coal-hammer" grasp is a special case of the cylindrical power grasp, and is identified by the high degree of thumb abduction. We define the type 1 "coal-hammer" grasp to be one in which the thumb does not touch the held object, while the type 2 "coal-hammer" grasp refers to one in which the thumb touches the object. The type 2 grasp normally occurs for a thick object, as in the case of experiment 2 described in Table 5.

Table 5 **Description of Experiments involving Power Grasps**

Experiment #	Description
1	Spherical power grasp. Radius of sphere = 3.26 cm.
2	Type 2 cylindrical "coal-hammer" grasp (thick cylinder). Radius of circular cross-section = 3.30 cm.
3	Type 1 cylindrical "coal-hammer" grasp (medium-thick cylinder). Radius of circular cross-section = 1.47 cm. (Note: the thumb does not touch the cylinder)
4	Cylindrical power grasp (medium-thick cylinder). Radius of circular cross-section = 1.47 cm.
5	Type 1 cylindrical "coal-hammer" grasp (thin cylinder). Radius of circular cross-section = 0.97 cm. (Note: the thumb does not touch the cylinder)
6	Cylindrical power grasp (thin cylinder). Radius of circular cross-section = 0.97 cm.
7	Cylindrical power grasp (elliptical cross-section with a=3.3cm, b=1.9cm)[1]
8	Cylindrical power grasp (elliptical cross-section with a=4.1cm, b=1.9cm)

Table 6 **Results of Experiments involving Power Grasps**

Experiment #	$C_{eff,max}$	$Config_{max}$	$C_{eff,2}$	$Config_2$
1	0.407	{(0)(1)(2345)}	0.318	{(0)(1234)(5)}
2	0.542	{(0)(1)(2345)}	0.270	{(0)(12345)}
3	0.666	{(0)(2345)}	0.544	{(02345)}
4	0.531	{(0)(1)(2345)}	0.337	{(0)(12345)}
5	0.650	{(0)(2345)}	0.528	{(0)(34)(25)}
6	0.522	{(0)(1)(2345)}	0.351	{(0)(12345)}
7	0.538	{(0)(1)(2345)}	0.427	{(0)(12345)}
8	0.537	{(0)(1)(2345)}	0.415	{(0)(12345)}

It is interesting to note from Table 2 that, despite the differences in cylinder thickness, the maximum grasp indices for the power grasps in experiments 2, 4 and 6 do not differ very much from one another. It is also interesting to note that the grasp cohesive index remains about the same despite changes in the cross-sectional shapes, as evidenced in the results of Experiments 7 and 8. The corresponding values for the "coal-hammer" grasps for experiments 3 and 5 do not seem to be significantly different from each other.

---

1. a and b are the semi-major and semi-minor axis lengths respectively.

Table 7 **Best effective cohesive indices for power grasps (including type 2 "coal-hammer" grasps) on cylinders of different thicknesses**

Subject #	$C_{eff,max}$ (thick cylinder)	$C_{eff,max}$ (medium cylinder)	$C_{eff,max}$ (thin cylinder)
0	0.542	0.531	0.522
1	0.534	0.531	0.526
2	0.543	0.534	0.505
3	0.541	0.533	0.535
4	0.536	0.531	0.529
5	0.543	0.538	0.528
6	0.542	0.524	0.522
7	0.544	0.535	0.453
8	0.543	0.534	0.532
9	0.544	0.511	0.515
10	0.537	0.535	0.513
11	0.545	0.516	0.497

Table 7 shows the best grasp cohesive index for the power grasps on cylinders of different thickness. Note that the type 2 "coal-hammer" grasp is virtually unidentifiable as a special case of the power grasp on the basis of the grasp cohesive index alone. This is due to the thumb touching the object. The virtual configuration and composition for all these grasps and for all the different subjects is the same, namely, {(0)(1)(2345)}. The average value is 0.528, with the standard deviation of 0.017 (3.2% of the average value).

Table 8 **Best effective cohesive indices for type 1 "coal-hammer" grasps on cylinders of different thicknesses**

Subject #	$C_{eff,max}$ (medium cylinder)	$C_{eff,max}$ (thin cylinder)
0	0.666	0.650
1	0.681	0.665
2	0.676	0.661
3	0.677	0.640
4	0.672	0.672
5	0.680	0.645
6	0.681	0.661
7	0.678	0.657
8	0.659	0.683
9	0.671	0.683
10	0.683	0.676
11	0.646	0.627

Table 8 lists the results for the type 1 "coal-hammer" grasps on two cylinders of differing thicknesses. Again the best effective cohesive index is relatively independent of the thickness of the cylinder grasped and the person holding the object. The mean best grasp cohesive index is 0.666 with the standard deviation of 0.016 (2.4% of the average index). The configuration associated with all the indices is {(0)(2345)}.

Table 9 **Best effective cohesive indices for power grasps on cylinders with elliptical cross-section of different eccentricities**

Subject #	$C_{eff,max}$ (a=3.3cm; b=1.9cm)	$C_{eff,max}$ (a=4.1cm; b=1.9cm)
0	0.538	0.537
1	0.538	0.538
2	0.537	0.540
3	0.535	0.543
4	0.535	0.535
5	0.536	0.536
6	0.539	0.540
7	0.537	0.538
8	0.534	0.524
9	0.538	0.538
10	0.539	0.534
11	0.537	0.528

The virtual finger configuration and composition for all the subjects is {(0)(1)(2345)}. The mean grasp cohesive index is 0.536 and the standard deviation is 3.8×10^{-3}, which is about 0.7% of the mean. (Note: a and b in Table 7 are the semi-major and semi-minor axis lengths respectively.) The mean grasp cohesive index for these grasp experiments is virtually the same as that for grasp experiments involving cylinders of circular cross-sections. This further strengthens our claim that the grasp cohesive index can be used to identify prismatic power grasps, regardless of the cross-sectional shape of the object held. The similarity in the grasp cohesive indices can be attributed to the proximity of the four fingers to each other and the high similarity in the finger configuration in this type of power grasp.

Table 10 **Best effective cohesive indices for spherical power grasps**

Subject #	$C_{eff,max}$	$C_{eff,2}$
0	0.407 {(0)(1)(2345)}	0.318 {(0)(1234)(5)}
1	0.400 {(0)(1)(234)}	0.329 {(0)(123)(4)}
2	0.375 {(0)(1)(2345)}	0.294 {(0)(1234)(5)}
3	0.343 {(0)(123)(45)}	0.329 {(0)(1)(2345)}
4	0.361 {(0)(1234)(5)}	0.310 {(0)(1)(2345)}
5	0.447 {(0)(1)(2345)}	0.364 {(0)(1234)(5)}
6	0.326 {(0)(123)(45)}	0.311 {(0)(1234)(5)}
7	0.444 {(0)(1)(2345)}	0.353 {(0)(1234)(5)}
8	0.493 {(0)(1)(2345)}	0.449 {(0)(12345)}
9	0.488 {(0)(1)(2345)}	0.349 {(0)(12345)}
10	0.393 {(0)(1)(2345)}	0.366 {(0)(1234)(5)}
11	0.426 {(0)(1)(2345)}	0.312 {(0)(1234)(5)}

Similar experiments were conducted using a sphere. The results for the spherical power grasps are shown in Table 10. The mean best grasp cohesive index is 0.409 with the standard deviation of 0.051 (12.5% of the mean). As can be seen, the best grasp cohesive indices differ markedly from person to person, and even then, the configuration associated with the best cohesive index is different for subjects 1, 3, 4, and 6. For the others, the configuration is {(0)(1)(2345)}. Note that for subject 1, the last or little finger barely touched the ball and hence was not considered in the analysis, and the configuration is {(0)(1)(234)} instead. These results are based on the assumption that the force exerted by each finger is the same. If the force data were available and that the thumb exerted a higher force than other fingers (a possible scenario), then the optimal configuration {(0)(1)(2345)} may have been more consistent for all subjects. This prediction has yet to be tested.

The reason for the higher disparity in the grasp cohesive index in spherical grasps than in cylindrical grasps is that in cylindrical grasps, the fingers (excluding the thumb) are normally kept very close together. The relative inter-phalange arrangements in the cylindrical grasps are consistent, despite the differing sizes of the cylinders handled. The amount of finger flexion (due to the different cylinder sizes) has little effect on the grasp cohesive index. For spherical grasps, however, relative inter-phalange arrangements do differ markedly from subject to subject, causing the higher range of grasp cohesive indices observed. It is interesting to note that the grasp cohesive indices for the spherical grasps are all lower than those for the cylindrical grasps. This is to be expected, because in a spherical grasp, the amount of force interaction between fingers is higher than that in a prismatic power grasp. The higher force interaction is, in turn, attributed to the more widely separated fingers in a spherical grasp and the significant curvature of the sphere.

3.2 Procedure for Grasp Recognition

From this study, a grasp can be identified from the following general steps:

1. *Compute the real finger to virtual finger mapping which yields the virtual finger compositions and the grasp cohesive index.*

2. *If the palm surface is not involved in the grasp, classify it as a non-volar grasp.*

3. *Otherwise, by checking the grasp cohesive index and, if necessary, the degree of thumb abduction, classify it either as a spherical, cylindrical or coal-hammer (type 1 or type 2) power grasp.*

A grasp can be classified as a volar grasp or non-volar grasp according to whether there is volar-object interaction or not (Figure 11). If it is a non-volar grasp, further classification can be done by checking if only the fingertips are involved in the grasp, and the contact points' closeness of fit to a circle or rectangle. This is illustrated in Figure 12. Unless the grasp is a lateral pinch (in which case the grasp is a power grasp), the grasp is classified as a precision grasp.

Contact Web — Map real fingers to virtual fingers

Is the palm involved in the grasp?

Yes — Grasp is Volar Grasp (Figure 13)

No — Grasp is Non-Volar Grasp (Figure 12)
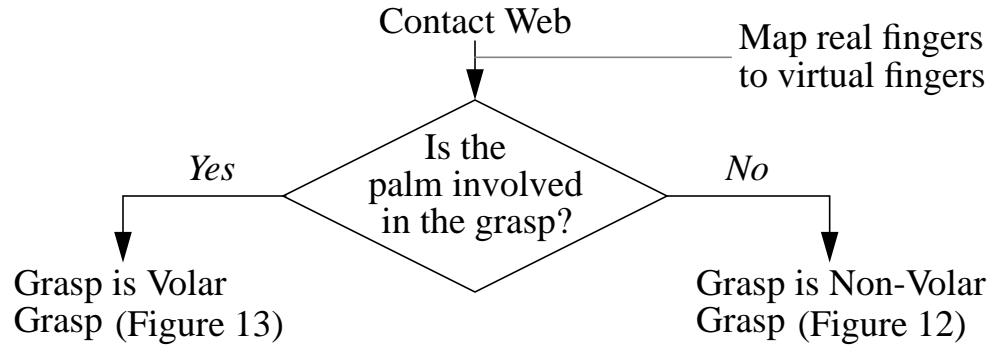
Figure 11 **Recognition of major type of grasp**

Non-Volar Grasp

$N_0(P_H) = N_1(P_H)$

$N_0(P_H) < N_1(P_H)$

Fingertip Grasp

Composite Non-Volar Grasp

Classify according to $N_0(P_H)$ and fit of contact points to circle or rectangle

Classify according to $N_0(P_H)$ and $N_1(P_H)$
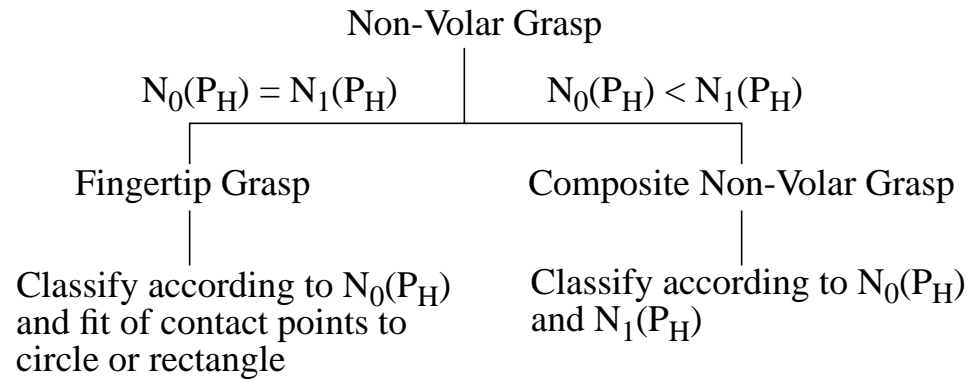
Figure 12 **Discrimination graph for non-volar grasps**

Volar Grasp

Grasp Cohesive Index

Spherical
Power Grasp

Cylindrical Power and
Type 2 "Coal-hammer"
Grasps

Type 1 "Coal-hammer"
Grasp

Degree of Thumb Abduction

Cylindrical
Power Grasp

Type 2 "Coal-hammer"
Grasp
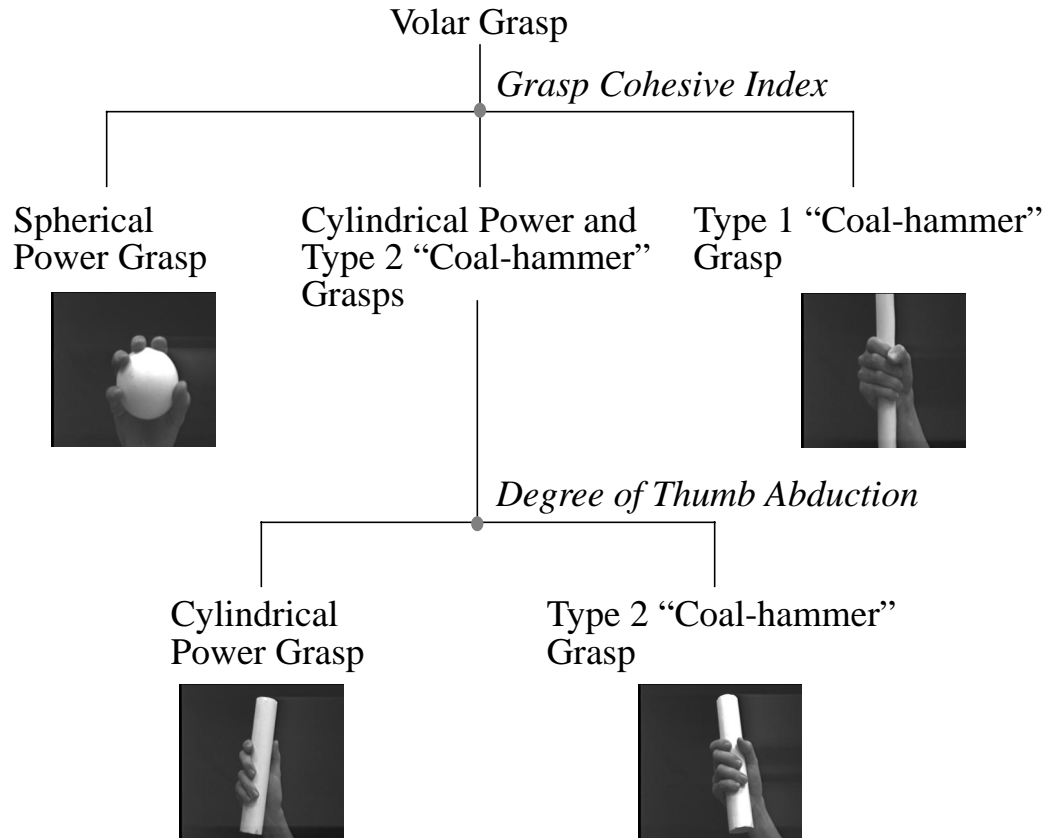
Figure 13 **Discrimination graph for power grasps**

The volar grasp discrimination procedure in step 3 is graphically depicted in Figure 13. (Note that all volar grasps are power grasps.) The first level of classification is performed using the following discrimination function:

$$\tau_i = e^{-\frac{1}{2}\left(\frac{x - \mu_i}{\sigma_i}\right)^2}$$

where μ_i is the mean value of the grasp cohesive index for the ith power grasp category and σ_i is the associated standard deviation. The power grasp category is identified by the largest value of the discrimination function. Should the cylindrical and type 2 "coal-hammer" grasps need to be discriminated, we would then determine the degree of thumb abduction. The type 2 "coal-hammer" grasp is associated with a high degree of thumb abduction. We use the thumb in a standard position (fully extended and in the plane of the palm) as a reference line in determining the degree of thumb abduction. We consider deviations greater than 45° to be significant enough to be categorized as a type 2 "coal-hammer" grasp.

Note that the object shape has not been directly taken into consideration here; the local object shape (i.e., part of the object within the compass of the hand) has been implicitly taken care of by the contact web.

3.3 Grasp Recognition from Range and Intensity Images

Three range and intensity image sequences of finger movements leading to different power grasps were taken and then analyzed using the grasp recognition scheme described earlier. In each sequence, the fingers are tracked to the final grasp configuration before the grasp is identified. However, prior to this, the hand model needs to be initialized. This is done using a separate program.

3.3.1 Hand Model Initialization

Each finger segment is modeled by a cylinder. The purpose of the hand model initialization is to determine the finger segment cross-sectional radius and length, and the relative positions of the fingers. The assumed hand posture is with the fingers fully extended, and such that the plane of the palm is approximately perpendicular to the camera viewing direction. These fingers are taken to be the reference positions, and abduction angles are measured relative to these positions.

In addition, to facilitate the measurement of finger segment lengths, dark lines were drawn across the finger at the distal and proximal interphalangeal joints (except for the thumb, where a line was drawn across it at the interphalangeal joint). The proximal finger segment length is calculated from empirical anthropometric studies of human finger segment length ratios [13].

The steps involved in hand model initialization are:

1. *Thresholding and hand boundary extraction*
   The image is first thresholded by assigning background values to intensity pixels whose corresponding range values are inadmissible and assigning foreground values if they are admissible. Small regions are eliminated, leaving the hand in the resulting binary image. Subsequently, the hand boundary is extracted using a simple 8-connected boundary following algorithm.

2. *Curvature analysis of hand boundary*
   Using the convention that a convex portion of a body has negative curvature, the tips of the fingers are located at points where curvature minima are observed. Similarly, the five grooves between fingers are located at positions of curvature maxima.

3. *Identification of finger regions*
   The finger regions are identified by noting that if the fingers are cyclically ordered anti-clockwise, the thumb is farthest away from the fingers immediately preceding and following it.

4. Location of interphalangeal joints and calculation of finger segment lengths

The position of the interphalangeal joints are approximated by using the Hough transform to locate dark lines. The finger segment lengths are then calculated from the distances between joints or between fingertips and the nearest joints.

5. Determination of best fit cylinders

3-D cylinders are fitted to the fingers from the hand range data using an iterative least-square formulation.

Figure 14 shows the intensity image of a hand and three snapshots of the hand model initialization program. Note that this program is run only once for a subject's hand.
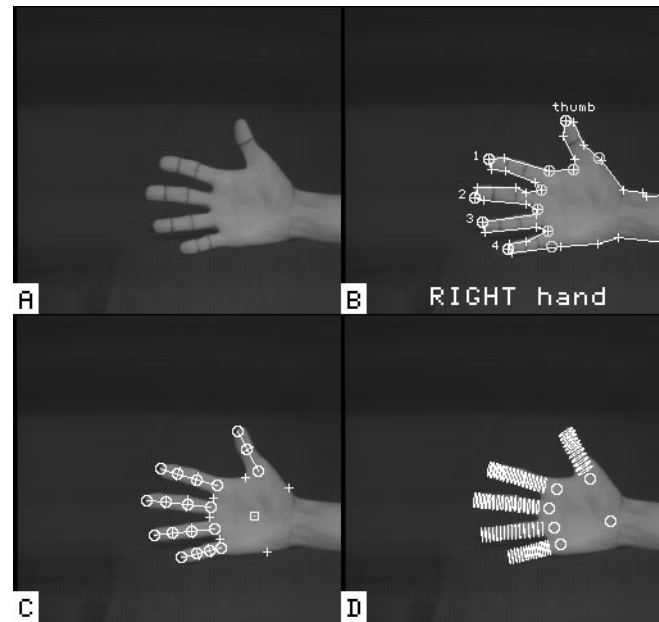
Figure 14 **Hand model initialization. (a) Intensity image; (b) Identification of fingers; (c) Localization of finger joints; (d) Cylindrical fitting of fingers.**

3.3.2 Finger Tracking

The basic method used in finger tracking is local search of the minimum sum of two types of matching errors: error in range data fitting, and error in matching the hypothesized finger 2-D projection to the actual finger position in the image. The following assumptions made are:

1. The hand does not move; only the fingers move (via flexion, extension, abduction and adduction).

2. The first frame features a hand with fingers fully or nearly fully extended.

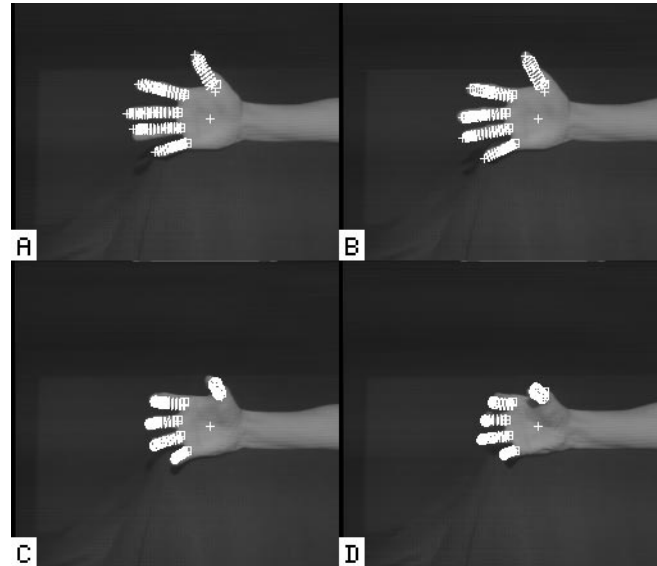*3. There is no significant interphalangeal occlusion.*

3.3.3 Grasp Recognition Results

Three range and intensity images were recorded and analyzed.

Example 1 (Figure 15 and Figure 16)

Four of the eight frames for this grasp sequence is shown in Figure 15.

Figure 15 **Finger tracking sequence for Example 1. (a) Frame 1; (b) Frame 3; (c) Frame 6; (d) Frame 8**
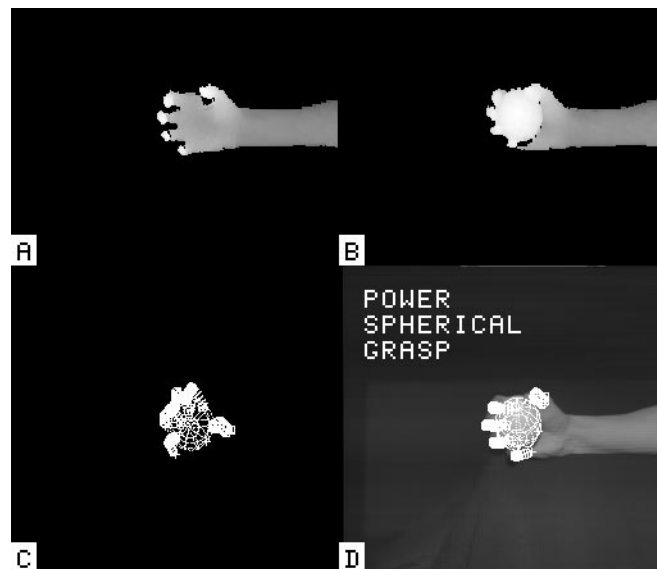
Figure 16 **Recognition results for a spherical power grasp. (a) Range image of last frame of sequence; (b) Range image of hand and object; (c) Alternate view of tracked fingers with object; (d) Classification of grasp**

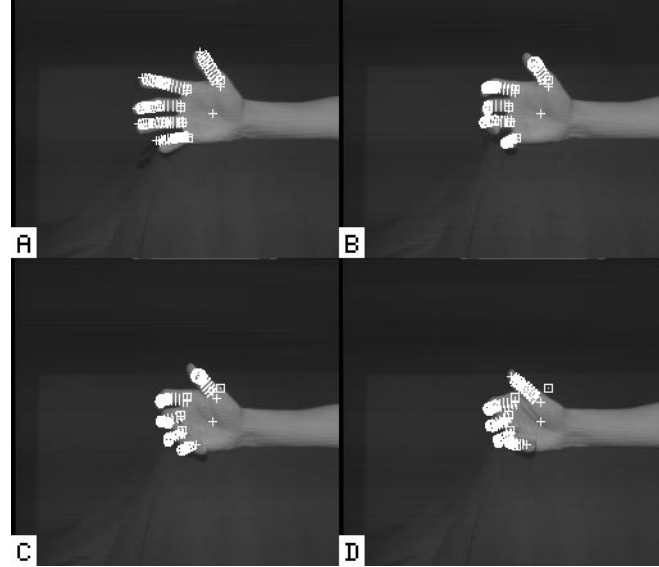The grasp cohesive index for this example is 0.356, with the following virtual finger compositions: $VF_1 = \{0\}$, $VF_2 = \{1\}$, and $VF_3 = \{2, 3, 4, 5\}$. This grasp is classified as a spherical power grasp (Figure 16).

Example 2 (Figure 17 and Figure 18)

Part of the frame sequence for this example is shown in Figure 17.

Figure 17 **Finger tracking sequence for Example 2. (a) Frame 1; (b) Frame 3; (c) Frame 6; (d) Frame 8**

Figure 18 **Recognition results for a cylindrical power grasp. (a) Range image of last frame of sequence; (b) Range image of hand and object; (c) Alternate view of tracked fingers with object; (d) Classification of grasp**
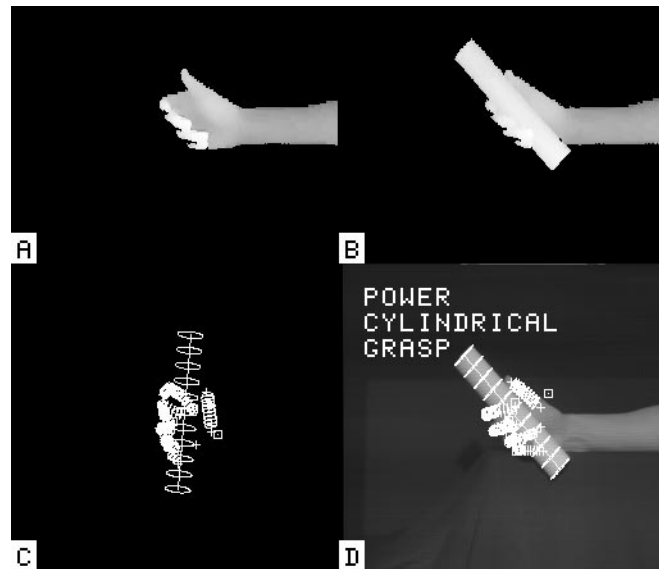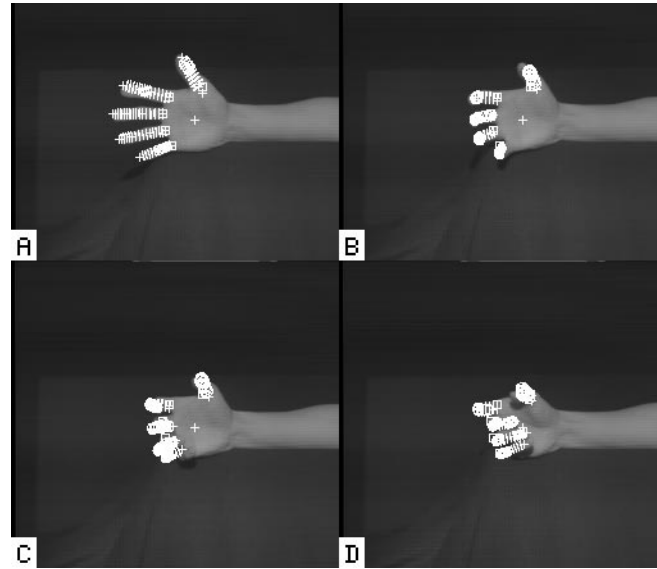
The grasp cohesive index for the second example is 0.508, with the following virtual finger compositions: $VF_1 = \{0\}$, $VF_2 = \{1\}$, and $VF_3 = \{2, 3, 4, 5\}$. From the grasp cohesive index, this grasp can either be a cylindrical power or type 2 "coal-hammer" grasp. Since the angle of the thumb subtends only 23° with the standard (original) thumb posture, it is classified as a cylindrical power grasp.

Example 3 (Figure 19 and Figure 20)

Figure 19 shows four frames of the grasp sequence for Example 3.

Figure 19 **Finger tracking sequence for Example 3. (a) Frame 1; (b) Frame 3; (c) Frame 6; (d) Frame 8**

Figure 20 **Recognition results for a type 2 "coal-hammer" grasp. (a) Range image of last frame of sequence; (b) Range image of hand and object; (c) Alternate view of tracked fingers with object; (d) Classification of grasp**
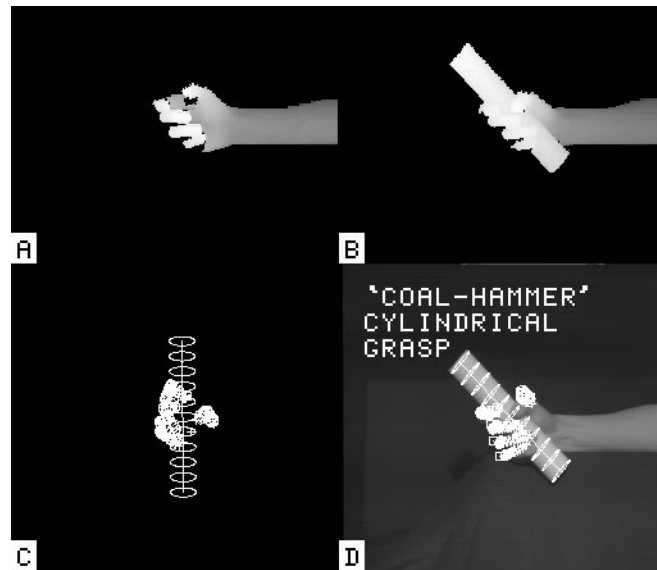
The grasp cohesive index for this example is 0.527, with the following virtual finger compositions: $VF_1 = \{0\}$, $VF_2 = \{1\}$, and $VF_3 = \{2, 3, 4, 5\}$. As with example 2, just from the grasp cohesive index alone, this grasp could be either a cylindrical power or a type 2 "coal-hammer" power grasp. From the high degree of thumb abduction (subtending 77° with the original thumb configuration), it is thus classified as a type 2 "coal-hammer" grasp.

The experiments and their results described in Subsection 3.1 and Subsection 3.3 indicate that it is possible to categorize grasps by using the contact web and real finger to virtual finger mapping. This mapping is instrumental in characterizing the type of grasp demonstrated in the scene.

Chapter 4 Conclusions

A framework for recognizing a grasp has been described in this report. A 3-D structure comprising a network of effective contact points of the hand with the grasped object is proposed as a tool for grasp analysis. We call this 3-D structure the *contact web*. It enables the grasp to be classified in a more continuous manner. In addition, by employing a particular real finger to virtual finger mapping, the grasp can be described in higher level conceptual terms such as virtual finger composition and opposition space. Another important consequence of this mapping is an index called the *grasp cohesive index*, which can be used to identify the grasp.

The grasp is actually one of the three identifiable phases in a grasping task. The other two phases are the pre-grasp and manipulation phases. Future work will be devoted to the analysis of these two phases in our effort to automate the recognition of a grasping task. All this is in line with our proposed notion of "perceptual programming," which epitomizes the capability of a robotic system to replicate a task by observing and understanding the same task performed by a human operator.

Acknowledgments

References

[1] Lammineur, P., and Cornillie, O., *Industrial Robots*, Pergammon Press, 1984, pp. 43-54.

[2] Asada, H., and Asari, Y., "*The Direct Teaching of Tool Manipulation Skills via the Impedance Identification of Human Motions*," Proc. IEEE Int'l Conf. on Robotics and Automation, 1988, pp. 1269-1274.

[3] Finkel, R., Taylor, R., Bolles, R., Paul, R., and Feldman, J., "*AL: A Programming System for Automation*," Tech. Rep. AIM-177, Stanford University, Artificial Intelligence Lab., 1974.

[4] Lozano-Perez, T., "*Automatic Planning of Manipulator Transfer Movements*," IEEE Trans. on Systems, Man and Cybernetics, SMC-11(10), 1981, pp. 681-689.

[5] Gruver, W.A., Soroka, B.I., Craig, J.J., Turner, T.L., "*Evaluation of Commercially Available Robot Programming Languages*," Proc. 13th Int'l Symp. on Industrial Robots, 1983, pp. 12-58.

[6] Iberall, T., Jackson, J., Labbe, L., and Zampano, R., "*Knowledge-Based Prehension: Capturing Human Dexterity*," Proc. IEEE Int'l Conf. of Robotics and Automation, 1988, pp. 82-87.

[7] Li, Z., and Sastry, S., "*Task Oriented Optimal Grasping by Multifingered Robot Hands*," Proc. IEEE Int'l Conf. of Robotics and Automation, 1987, pp. 389-394.

[8] Tomovic, R., Bekey, G.A., and Karplus, W.J., "*A Strategy for Grasp Synthesis with Multifingered Robot Hands*," Proc. IEEE Int'l Conf. of Robotics and Automation, 1987, pp. 83-89.

[9] Nguyen, V., "*Constructing Stable Grasps in 3-D*," Proc. IEEE Int'l Conf. of Robotics and Automation, 1987, pp. 234-239.

[10] Ikeuchi, K., and Seuhiro, T., "*Towards an Assembly Plan from Observation: Task Recognition with Polyhedral Objects*," CMU Tech. Rep. CMU-CS-91-167, Aug. 1991.

[11] Kuniyoshi, T., Inaba, M., and Inoue, H., "*Teaching by Showing: Generating Robot Programs by Visual Observation of Human Performance*," Proc. 20th Int'l Symp. on Industrial Robots, 1989, pp. 119-126.

[12] *Joint Motion: Method of Measuring and Recording*, American Academy of Orthopedic Surgeons, Chicago, 1965.

[13] An, K.N., Chao, E.Y., Cooney, W.P., and Linscheid, R.L., "*Normative Model of Human Hand for Biomechanical Analysis*," Journal of Biomechanics, Vol. 12, 1979, pp. 775-788.

[14] Taylor, C.L., and Schwarz, R.J., "*The Anatomy and Mechanics of the Human Hand*," Artificial Limbs, No. 2, 1955, pp. 22-35.

[15] Schlesinger, G., "*Der Mechanische Aufbau der Kunstlichen Glieder*," Borchardt, et al. (eds.), Ersatzglieder und Arbeitshilfen fur Kriegsbeschadigte und Unfallverletzte, Springer, 1919, pp. 321-699.

[16] Cutkosky, M.R., and Wright, P.K., "*Modeling Manufacturing Grips and Correlations with the Design of Robotic Hands*," Proc. IEEE Int'l Conf. on Robotics and Automation, 1986, pp. 1533-1539.

[17] Nguyen, T.N., and Stephanou, H.E., "*A Topological Model of Multifingered Prehension*," Proc. IEEE Int'l Conf. on Robotics and Automation, 1989, pp. 446-451.

[18] Cutkosky, M.R., "*On Grasp Choice, Grasp Models and the Design of Hands for Manufacturing Tasks*," IEEE Trans. on Robotics and Automation, 5(3), 1989, pp. 269-279.

[19] Cutkosky, M.R., and Howe, R.D., "*Human Grasp Choice and Robotic Grasp Analysis*," Dextrous Robot Hands, eds. Venkataraman, S.T., and Iberall, T., Springer-Verlag, 1990, pp. 5-31.

[20] Arbib, M.A., Iberall, T., and Lyons, D.M., "*Coordinated Control Programs for Movements of the Hand*," Hand Function and the Neocortex, eds. Goodwin, A.W., and Darian-Smith, I., Springer-Verlag, 1985, pp. 111-129.

[21] Iberall, T., Bingham, G., and M.A. Arbib, "*Opposition Space as a Structuring Concept for the Analysis of Skilled Hand Movements*," Experimental Brain Research Series 15 - Generation and Modulation of Action Patterns, eds. Heuer, H., and Fromm, C., Springer-Verlag, 1986, pp. 158-173.

[22] Iberall, T., "*The Nature of Human Prehension: Three Dextrous Hands in One*," Proc. IEEE Int'l Conf. of Robotics and Automation, 1987, pp. 396-401.

[23] Iberall, T., and MacKenzie, C.L., "*Opposition Space and Human Prehension*," Dextrous Robot Hands, eds. Venkataraman, S.T., and Iberall, T., Springer-Verlag, 1990, pp. 32-54.

[24] Iberall, T., "*A Neural Network for Planning Hand Shapes in Human Prehension*," Proc. of American Control Conf., 1988, pp. 2288-2293.

[25] Nguyen, T.N., and Stephanou, H.E., "*A Topological Algorithm for Continuous Grasp Planning,*" Proc. IEEE Int'l Conf. on Robotics and Automation, 1990, pp. 670-675.

[26] Napier, J., "*The Prehensile Movements of the Human Hand*," Journal of Bone and Joint Surgery, Vol. 38B, No. 4, Nov. 1956, pp. 902-913.

[27] Lyons, D.M., "*A Simple Set of Grasps for a Dextrous Hand*," Proc. IEEE Int'l Conf. on Robotics and Automation, 1985, pp. 588-593.

[28] Buchholz, B., and Armstrong, T.J., "*A Kinematic Model of the Human Hand to Evaluate its Prehensile Capabilities*," Journal of Biomechanics, Vol. 25, No. 2, 1991, pp. 149-162.

[29] Sollerman, C., "*Grip Function of the Hand: Analysis, Evaluation and a New Method*," Hand Surgery, Dept. of Orthopaedic Surgery, Sahlgren Hospital, University of Goteberg, Goteberg, Sweden, 1980.

[30] Pao, L., and Speeter, T.H., "*Transformation of Human Hand Positions for Robotic Hand Control*," Proc. IEEE Int'l Conf. on Robotics and Automation, 1989, pp. 1758-1763.

[31] Long, C., Conrad, P.W., Hall, E.A., and Furler, S.L., "*Intrinsic-Extrinsic Muscle Control of the Hand in Power Grip and Precision Handling*," Journal of Bone and Joint Surgery, Vol. 52-A, No. 5, 1970, pp. 853-867.

[32] Griffiths, H.E., *Treatment of the Injured Workman*, Lancet, 1943.

[33] McBride, E.D., *Disability Evaluation*, T.B. Lippincott Co., 1942.

[34] Ikeuchi, K., and Hebert, M., "*Task Oriented Vision*," CMU Tech Rep CMU-CS-91-163, July 1991.