Data Driven Grasp Synthesis using Shape Matching and Task-Based Pruning

Ying Li, Jiaxin L. Fu, and Nancy S. Pollard, *Member, IEEE*

Abstract— **Human grasps, especially whole-hand grasps, are difficult to animate because of the high number of degrees of freedom of the hand and the need for the hand to conform naturally to the object surface. Captured human motion data provides us with a rich source of examples of natural grasps. However, for each new object, we are faced with the problem of selecting the best grasp from the database and adapting it to that object. This paper presents a data-driven approach to grasp synthesis. We begin with a database of captured human grasps. To identify candidate grasps for a new object, we introduce a novel shape matching algorithm that matches hand shape to object shape by identifying collections of features having similar relative placements and surface normals. This step returns many grasp candidates, which are clustered and pruned by choosing the grasp best suited for the intended task. For pruning undesirable grasps, we develop an anatomically based grasp quality measure specific to the human hand. Examples of grasp synthesis are shown for a variety of objects not present in the original database. This algorithm should be useful both as an animator tool for posing the hand and for automatic grasp synthesis in virtual environments.**

Index Terms— **Grasp synthesis, hands, shape matching, grasp quality.**

I. INTRODUCTION

ANIMATED characters in games and virtual environments must be able to interact with their world in a realistic way. One requirement is the ability to grasp objects in order to manipulate them. If a grasp is not carefully constructed, it may appear impossible or awkward. In fact, the anatomical complexity of the human hand makes it challenging to pose the hand in ways that appear natural.

The problem of posing the human hand in a natural way can be addressed by making use of a database of successful grasps. Even though the hand has a large number of degrees of freedom, it takes on a similar shape in many grasps, and it may be easy to find an example grasp that is a good match to the current problem. This paper presents a data driven algorithm for selecting hand poses to grasp an object in a natural way. The key insight presented in this paper is that when a grasp database is available, grasping can to a large extent be treated as a shape matching problem (Figure 1). As intuition for why shape matching may be an effective way to select a grasp from a database, consider Figure 2, which shows the hand pose for a grasp of a mouse. The shape of the inside surface of the hand gives us a great deal of information about the shape of the mouse. We represent this shape information by extracting from
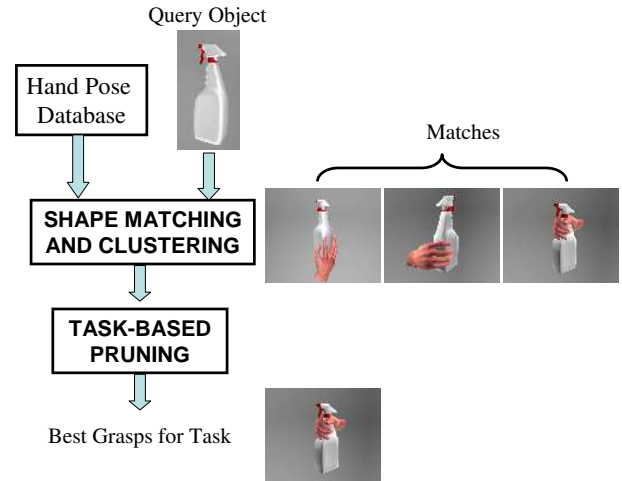
The authors are affiliated with the School of Computer Science, Carnegie Mellon University, 5000 Forbes Avenue, Pittsburgh, PA 15213-3890. Email: [liyingus@gmail.com] [jfu | nsp @ cs.cmu.edu]

Fig. 1. Grasp synthesis as a shape matching problem. Offline, the user creates a database of hand poses. Online, a user or a program loads a query— a three-dimensional model of the object to be grasped. The shape matching system searches the database to find the hand poses that best match the query object. Representative hand poses for this example are shown in the middle of the figure. The poses displayed are from a grasp of a mouse, a jelly jar, and a lightbulb (left to right). Finally, unacceptable grasps are pruned from consideration and the best grasps are chosen using a quality metric tailored to the desired task.
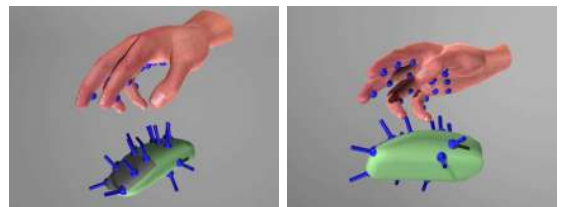

Fig. 2. Hand pose for the mouse grasp. The figure shows contact points on the hand and object, and contact normals on the object surface. Note that the inside surface of the hand contains a great deal of information about the shape of the mouse. If similar features can be found on a new object, it may be possible to use the same grasp for the new object.

each example grasp a set of representative contact points and normals.[1] If we find a similar arrangement of contact points and normals on the surface of a different object, then we may be able to reuse the same hand pose to grasp the new object. For example, the mouse grasp is used in the leftmost image in the middle row of Figure 1 to hold the spray bottle.

Shape matching is commonly used in other fields— examples include object recognition in computer vision [1]

---

[1]Specifically, the contact points are selected by hand from a discrete set of possibilities each time the user enters a new grasp into the database. Normals are obtained by taking the normal from the closest point on the object surface.

and example-based retrieval of geometric models in computer graphics [2]. However, the use of shape matching for grasp synthesis presents some new challenges. First of all, the hand surface gives us information about only a portion of the surface of the object. Because of this, global measures of shape such as overall volume and principle directions cannot be used to match the shape of the inner surface of the hand to the shape of the object. This problem also arises in the vision community–where there may be only partial shape information available from range scans, for example. In those circumstances, discriminative local features are typically used to compensate for the absence of global shape information (e.g., [3]). Discriminative local features are local regions of three-dimensional geometry (curves, corners, angles, etc.) that are useful for distinguishing one shape from another. However, in shape matching for grasp synthesis, our local features are typically not discriminative. They generally consist of contact patches that are nearly planar and ellipsoidal in shape. As such, they are not helpful for distinguishing between different grasped objects or different hand shapes.

In this paper, we present a novel algorithm that allows us to overcome the problem of partial shape information that does not contain discriminative local features. The main idea behind our approach is to randomly generate a set of global shape features from the contact points and normals representing a grasp. These features are global in the sense that they capture arrangements of contacts with respect to one another. If a similar collection of features can be located on an object surface, then we hypothesize that the hand shape may match this object surface and attempt to find alignments of the hand to the object that achieve the desired overall arrangement of contacts and normals. We tailor our solution to the domain of grasp synthesis by making use of shape features that contain contact normal information, which is important for grasping.

Our shape matching algorithm is extremely successful in that it returns a large number of candidate grasps. However, many of these grasps are not appropriate for the desired task (e.g., lift the object to place it on a shelf). To obtain a single best grasp, or a small set of candidate grasps to present to the user, we cluster the results and then prune them using a grasp quality metric. For effective pruning, we develop a novel grasp quality measure that compares the ability of the hand to apply forces to the grasped object to those forces needed to accomplish the task. Our model of grasp quality captures anatomical features of the human hand (e.g., the fact that flexors are stronger than extensors) and distinguishes grasps having a good arrangement of contacts over the object surface from those that do not.

Portions of this work have appeared separately in [4] and [5]. In this paper, we bring the ideas of shape matching and task-based quality metrics together into a single system for humanlike grasp synthesis. Specifically, we add a refinement algorithm to the shape matching portion of the algorithm to achieve solid contact with the object, and we present many new results and observations from evaluating the resulting grasps based on the intended task.

II. BACKGROUND

Algorithms for grasp synthesis that consider complex hand kinematics are typically procedural or rule-based [6], [7], [8], [9], [10], [11], [12]. These systems involve treating all or part of an object as a primitive shape (e.g., box, cylinder, sphere) for which a grasp synthesis strategy is available. This approach is supported by classifications and taxonomies of observed human grasps [13], [14], [15]. However, creating a rule-based grasping system based on these classifications requires a significant amount of skill on the part of the designer, and it may not always be clear how to fit irregularly shaped objects into a given classification system. Our goals in this paper are to reduce the knowledge and tuning required of the algorithm designer by (1) making use of a grasp database to capture variation in hand shape both across and within families of grasps, and (2) developing a grasp synthesis algorithm that allows an appropriate grasp to be selected from the database based on object geometry and task definition.

Data driven grasp synthesis has been used by Pollard and Zordan [16], who automatically construct a controller for grasping by fitting control setpoints from measured human example grasps and similarly by Kry and Pai [17], who create grasping controllers to match measured stiffness in measured hand-object interactions. However, in both cases, the grasp synthesis algorithms cannot form appropriate controllers for objects that differ substantially from those objects in the database. The work in the current paper could be combined with that in [16] or [17] to create a more robust and flexible grasping system. In other data-driven work related to grasping, ElKoura and Singh [18] use a data driven approach to animate the hand for guitar playing. They use a database of human grasps to filter results produced using an inverse kinematics algorithm so that a natural coupling between joint angles is expressed.

Data driven approaches are now quite popular for tasks such as reaching, kicking, and locomotion (e.g., [19], [10], [20]). In particular, Yamane and colleagues [21] make use of a database to obtain full body posture for characters that are grasping and moving objects. However, these systems often work by identifying poses that match constraints such as end effector position. Extending this approach to work for non-trivial grasps is challenging due to the large amount of contact between the hand and object. This difficulty motivated us to consider an alternative approach based on matching the shape of the inner surface of the hand to the geometry of the object.

A. Shape Matching Background

Shape matching algorithms have been studied for related applications such as example-based retrieval of three-dimensional objects from databases and object identification from range scans. (See [22], [1], [23] for surveys of techniques in these fields.) Most of these techniques, however, cannot be applied directly to our problem, because they assume problem characteristics that are not applicable here. Some techniques require information about the entire surface of the two shapes that are to be compared, either to measure global features (e.g., [22]) or to compare probability distributions of features (e.g.,

[2]). We have only partial information about shape available from the hand pose. Some techniques require dense and discriminative information about local features, such as may be available from range images in cluttered environments (e.g., [24], [25]). The shape information associated with the hand pose has very uninteresting local features: locally flat contact patches with their normals. Finally, some techniques require registration / alignment of objects or feature correspondence, for example to build histograms characterizing the space occupied by the object [26]. We have no *a priori* way to align the query object to a hand pose.

For the shape matching portion of this paper, we combine and modify ideas from two bodies of work. First, we take inspiration from work on shape descriptors that randomly sample a global shape function (e.g., [2], [27]). An example of a global shape function is the probability distribution of distances between pairs of points on an object surface [2]. A descriptor for this shape function could be computed by randomly sampling pairs of surface points and recording their distances. Such a descriptor is useful because it gives us information about how points on the surface are distributed with respect to one another. We adapt the shape descriptors used in previous work to provide more detailed information about the relative orientations of contact normals, which is important for grasping.

Second, we take inspiration from work on partial matching. In particular, we use a representative descriptor technique (i.e., a nearest neighbor search) to determine whether hand pose features can be found in the query object (e.g., [3], [28]). We modify this technique so that the descriptors are not based on local features but instead are based on a sample of a global shape function. We also introduce a weighting term to capture the intuition that some features are more important than others.

B. Grasp Quality Background

Research on grasp quality measures has primarily been in the field of robotics. Within robotics, a great deal of research has focused on selecting good points of contact on the object surface (e.g., [29], [30], [31], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43]). This research, however, makes the assumption that contact points are identical and independent, which is a very poor approximation to the actual mechanical constraints of the human hand [5].

Quality metrics that include mechanical constraints for robot hands and parallel manipulators have been investigated (e.g., [44], [45], [46], [47], [48], [49], [50]). However, ellipsoidal metrics such as force manipulability ellipsoids are most commonly considered. Ellipsoidal metrics may be appropriate for robotic hands, which are typically highly symmetrical. However, the human hand has strong asymmetries, such as the large difference between strength of the flexors and extensors [5]. To account for these asymmetries, we introduce a quality metric that more accurately models anatomical constraints such as maximum available muscle force and the specific tendon layout of the human hand. Our algorithm is unique in using a fast algorithm (linear programming) to compute a grasp quality metric that includes a biologically
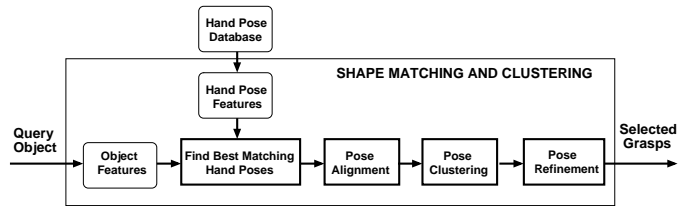

Fig. 3. Block diagram of the shape matching and clustering portion of our system. The online portion of this algorithm contains four main steps: find the hand poses that seem likely matches to this object, find good alignments of these poses to the object surface, cluster the results, and locally refine poses to conform to the object surface.

plausible tendon layout and muscle force capability obtained using measurement results from the biomechanics literature.

The core of our algorithm involves computing maximum force that can be applied to a grasped object. When developing this portion of the algorithm, we drew heavily upon work on optimizing contact forces and analyzing force capabilities, such as [51], [52], [53], [46], [54], [55], [39], [40].

III. SHAPE MATCHING AND CLUSTERING

Figure 3 shows a block diagram of the shape matching and clustering portion of our system. The input to the system is a geometric description of an object—the query object. The first step of the algorithm is to compute a collection of features that will be used to match the object against different hand poses. A hand pose database is available, and a similar feature set has been computed for each hand pose. There are then four main steps to the algorithm: (1) find the best hand poses to match the query object by comparing object features to hand pose features, (2) align these poses to the object geometry to obtain a set of possible grasps, (3) cluster results into a few representative grasps, and (4) refine each representative hand pose to conform to object geometry. The output of this portion of the system is a set of candidate grasps that will then be sorted and pruned based on effectiveness for the intended task (Section IV).

A. A Feature Set for Grasps

This section describes the feature set used for shape comparison. Our goal in choosing a feature set is to find a representation of object geometry that will allow us to quickly determine whether a given grasp can be matched to some portion of the object surface. In particular, we would like the feature set to capture important information about the relative configurations of contact positions and contact normals in the grasp.

We begin with an oriented point representation for both the hand pose and the object geometry. The feature sets are computed from these oriented point representations.

The oriented point representation for the query object is computed by randomly sampling points on the object surface. For point sampling, we use the algorithm outlined in [2], which randomly selects points on the surface of a triangulated object that are unbiased with respect to the surface area of the triangulated mesh. We use a fixed number of samples per
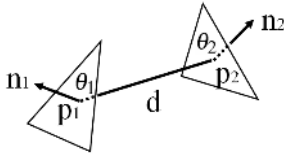
Fig. 4. We compute a three-dimensional feature value for pairs of points on the object or hand surface. This feature value consists of the distance d between points, and the angles $\theta 1$ and $\theta 2$ between the surface normals and the line passing through both points.

unit area (100 sample points per square centimeter in our experiments). Normals at those points are determined based on local surface geometry.

For the hand pose, the oriented point representation is the complete set of contact points and normals taken from the grasp. Each time a new grasp is added to the database, the relevant contact points are selected by hand from a discrete set of typical contact points that represent the contact patches on the hand surface that are commonly used in grasping. We select from a set of 38 total potential contact points distributed over the inner surface of the hand and the side of the index finger. Some of these points can be seen indicated on the surface of the hand in Figures 2, 6, and 7, for example. The normals for each of these contact points are taken from the nearest point on the surface of the grasped object. The nearest point on the object surface is approximated as the nearest neighbor from the oriented point representation computed for the object.

A variety of considerations drove our particular choice of a feature set that differs from those already available in the shape matching literature. First, global features cannot be computed due to the sparse shape information available from hand pose. For example, aligning the "center of mass" of the hand pose with the center of mass of the object would not make sense. Second, local feature information is not helpful here, as regions of contact between the hand and object are typically devoid of discriminatory features such as edges, corners, and areas of high curvature. Third, features that are important are not necessarily those that are visually salient or visually discriminative as discussed by Shilane and colleagues, for example [56]. Instead, we want to capture how contact points and normals are distributed relative to one another in three-dimensional space, because this arrangement of contacts is what allows us to create a solid grasp.

Because of these considerations, we develop a feature set based on random samples of pairs of contact points, and we include distance and normal information about each sampled pair. Specifically, we store for each pair of contact points a three-dimensional value that includes the distance between the points and the angles formed between the surface normals and the line passing through those points (Figure 4). This feature value is similar to that used by Ohbuchi and colleagues [27], but contains an extra dimension to locate the normals relative to the line between the contact points (not just relative to each other). We considered this to be the feature set of minimal dimension to distinguish between contact arrangements that have very different properties for
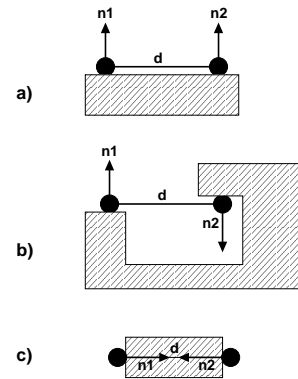

Fig. 5. Considering distance d only (a 1D feature), all three contact pairs look the same. Considering distance d and the relative angle between normals $n1$ and $n2$ (a two-dimensional feature), pairs b and c look the same. The three-dimensional feature illustrated in Figure 4 allows us to distinguish between these three situations. This discriminatory power is important for grasping, because it is important to be able to balance forces and torques appliied to the object through combinations of contacts.
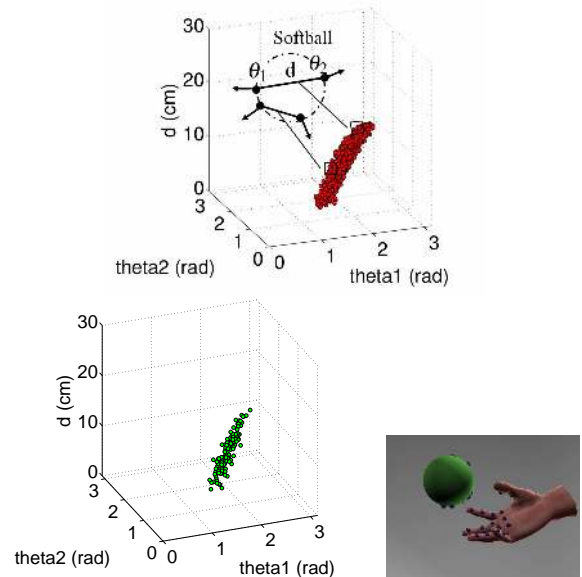

Fig. 6. (Top) Feature set for the softball. This figure shows 1% of the features actually used in our system. (Bottom Left) Feature set for the hand pose in this grasp. This figures shows half of the features available from the hand pose. (Bottom Right) Hand shape for the softball grasp.

the purposes of stable grasping (Figure 5). Higher-dimensional features would likely have even more discriminatory power, but increasing feature dimensionality would also add to the computational complexity of the algorithm and make it more difficult to visualize feature sets.

Example feature sets are shown in Figure 6, which illustrates feature sets for a softball and for the hand pose when the hand is grasping the softball. Note that although the hand pose feature set is more sparse than that of the softball, it captures its overall shape very well.

For good discriminatory power, we require that these feature sets differ substantially for different objects. As one example, Figures 7 show the feature sets for a book and for a grasp of that book. In this example, note that the feature set for the hand pose captures only a portion of the information contained in
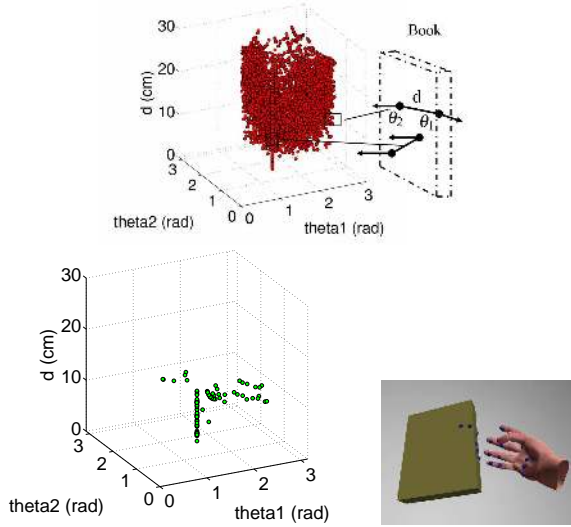
Fig. 7. (Top) Feature set for the book. This figure shows 1% of the features actually used in our system. (Bottom Left) Feature set for the hand pose in this grasp. This figure shows half of the features available from the hand pose. (Bottom Right) Hand pose for the book grasp.

D	database size = number of hand poses
P_i	set of features for hand pose i
$P_{i,k}$	feature k of hand pose i, represented as $[d \ \theta_1 \ \theta_2]^T$
Q_i	set of features for object i
$Q_{i,k}$	feature k of object i
$C(P_i)$	count of features for hand pose i
$C(Q_i)$	count of features for object i
$NN(A, b)$	nearest neighbor to b in feature set A
α	weights angular vs. linear terms in features

TABLE I

DEFINITION OF TERMS USED TO DESCRIBE THE SHAPE MATCHING
ALGORITHM.

the feature set for the book. Yet this information is sufficient to distinguish this grasp from the softball grasp.

The number of feature values to calculate is a practical consideration for the algorithm. We compute feature values for all pairs of points on the hand surface. For the object, we compute feature values for randomly selected pairs of points on the object surface. A sampling of 1000 features per unit centimeter of surface area has worked well for the sizes and types of objects in our database. The point pairs from which features are computed are drawn in an unbiased manner from the original set of points computed at the time the object was converted to an oriented point representation.

B. Matching Hand Pose to Object

We compare a hand pose and a query object using their feature sets. For a matching hand pose, we expect all features present in the hand feature set to appear in the object feature set, while the converse is not generally true. We therefore use a representative descriptor approach for evaluating a match [28][3].

The representative descriptor approach is designed for comparing a partial shape such as that obtained from a range image to a complete shape (a geometric model). The quality of the match is based on the average distance from features in the partial shape to their nearest neighbors in the complete shape. In this case, we compute distance from hand pose i to object j. This distance $E_{i,j}$ is expressed as follows, where Table I contains notation for this section.

$$E_{i,j} = \frac{1}{C(P_i)} \sum_{k=1}^{C(P_i)} Dist(P_{i,k} - NN(Q_j, P_{i,k})) \quad (1)$$

$$Dist(x) = \left[d^2 + \alpha^2 \theta_1^2 + \alpha^2 \theta_2^2 \right]^{\frac{1}{2}} \quad (2)$$

The first expression is just distance from each feature k in hand pose i to its nearest neighbor on object j, averaged over all hand pose features $P_{i,k}$. The distance metric $Dist$ is designed to provide a reasonable weighting between linear and angular values. In our case, we use $\alpha = 1 \ cm/rad$.

Although initial results were promising, this distance expression was not good enough. In particular, it would sometimes identify matches that were unacceptable because the object lacked certain key features (e.g., a place to put the thumb for the book grasp). To address this problem, we introduce a scheme to weight feature values more highly if they are relatively plentiful on the grasped object and relatively rare in other hand poses. The weighted distance from hand pose i to object j is

$$E_{i,j}^w = \frac{1}{w_i \, C(P_i)} \sum_{k=1}^{C(P_i)} w_{i,k} \, Dist(P_{i,k} - NN(Q_j, P_{i,k})) \quad (3)$$

where the normalizing term w_i is the sum of all weights for hand pose i:

$$w_i = \sum_{k=1}^{C(P_i)} w_{i,k} \quad (4)$$

and the weight for feature k of hand pose i is computed as

$$w_{i,k} = \frac{matchCount(Q_i, P_{i,k})}{\sum_{l \neq i} (matchCount(P_l, P_{i,k}))} \quad (5)$$

The numerator of this equation is a count of the number of times a feature similar to $P_{i,k}$ is found on object i, where object i is the object that was originally grasped using hand pose i. The denominator contains a count of the number of times a feature similar to $P_{i,k}$ is found in any of the other hand poses in the database. In other words, $w_{i,k}$ is high when the feature is relatively common on the grasped object and/or when it is uncommon in other hand poses. It is important to note that when computing the value in the numerator, we check only the portion of the object surface relevant for this grasp. For example, if a person was grasping a handle to a pitcher, only features on that handle would be relevant. The portion of an object relevant to the grasp is selected by the user during construction of the grasp database (Figure 16).

We can make a Bayesian argument for this weighting function. Let $P(pose_i)$ indicate the probability of encountering the current hand pose i. Let $P(feature_{i,k})$ indicate the probability of encountering the current feature $P_{i,k}$. Then the probability

Fig. 8. It is not enough to identify matching hand poses. We must also find an alignment of the hand pose to the object surface.
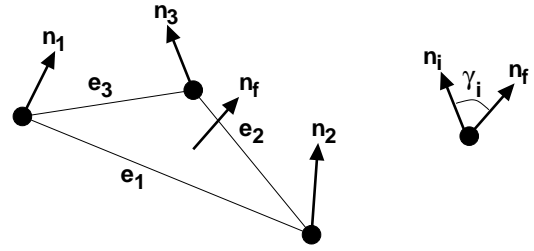

Fig. 9. A single triangle is used to find a candidate set of alignments of hand pose to object surface. This alignment triangle is created from three contact points selected from the hand pose.

that hand pose i is a good match for object j based on finding the current feature $P_{i,k}$ on object j can be expressed as

$$P(pose_i|feature_{i,k}) = \frac{P(feature_{i,k}|pose_i)P(pose_i)}{P(feature_{i,k})} \quad (6)$$

In weighting term $w_{i,k}$, we estimate $P(feature_{i,k}|pose_i)$ by counting the number of times the feature appears on the object originally grasped using this hand pose. $P(pose_i)$ can be considered uniform and we ignore it. $P(feature_{i,k})$ is estimated by counting the occurrences of that feature over all hand poses.

The function $matchCount$ used in computing weight $w_{i,k}$ can be expressed as follows:

$$matchCount(A, b) = \sum_{k=1}^{C(A)} isMatch(A_k, b) \quad (7)$$

with

$$isMatch(A_k, b) = \left\{ \begin{array}{ll} 0 & Dist(A_k - b) > \epsilon \\ 1 & Dist(A_k - b) \le \epsilon \end{array} \right\} \quad (8)$$

A match is counted if the distance between corresponding features is less than a value ϵ. In our experiments, ϵ is set to 10% of the maximum distance between features in the hand pose under consideration.

During runtime, pose selection proceeds as follows. Given a new object O, features of that object are extracted and value $E_{O,j}^{w}$ is computed for each hand pose j in the database. Those poses with the lowest values of $E_{O,j}^{w}$ are identified as candidate poses and passed on to the pose alignment, clustering, and refinement algorithms described in the next section.

C. Pose Alignment, Clustering, and Refinement

Once a hand pose is selected from the set of candidates, it must be aligned to the object shape. The goal of the alignment process is to find a transformation to be applied to the hand pose so that the contact points in the hand pose are brought into correspondence with points on the object having similar normals (Figure 8). Our alignment process has three parts. First, we narrow the set of transforms by identifying good alignments for a single triangle from the hand pose. Then, for each of these candidate transforms, we check all other contact points from the hand pose, discarding transforms that result in a poor match for any one of those contacts. Finally, we cluster and refine the results. The next paragraphs provide more detail.

1) Finding Candidate Transforms: In the first step of the alignment process, coordinate transforms are collected by finding good matches between a triple of points from the hand pose and triples of points on the object. We indicate an alignment triangle formed by three contact points on the

hand pose as shown in Figure 9. The alignment triangle is manually defined as a specific triple of contact points for each hand pose at the time the grasp is added to the database. The primary goal in selecting an alignment triangle is to provide early rejection of unlikely matches; in practice, we simply select one of the largest possible triangles from the contacts that made up the grasp.

For the alignment triangle, we compute edge lengths e_i and angles γ_i between the normals n_i and the triangle normal n_f. Randomly sampled triangles from the object surface are then tested for a match to the alignment triangle. Matching triangles must have edge lengths within value ϵ_e and angles within value ϵ_γ, where $\epsilon_e = 0.5\ cm$ and $\epsilon_\gamma = 0.5\ rad$ in our experiments. We check the alignment triangle against a randomly generated set of triples of points that can be constructed from the oriented point representation of the object. [2] All matching triangles are recorded and passed on to the next step.

2) Pruning Poor Transforms: Once a number of candidate transforms are generated using the alignment triangle, each of these transforms is examined to test whether it represents a plausible grasp. For an alignment to be plausible, it must meet two conditions. The first condition is that all desired contact points must be present in the new grasp. To check this condition, all contact points are first transformed by aligning the corresponding triangles, and the pose is checked to see whether there are good matches for all contact points on the object surface. For each contact point, its nearest neighbor on the surface of the object is identified by selecting the closest point from the object's oriented point representation. Then the entire transform is discarded if the distance from the contact point to the object surface is greater than value ϵ_d, or the angle between the contact normal and the object surface normal is greater than value ϵ_n, with $\epsilon_d = 1\ cm$ and $\epsilon_n = 1\ rad$ in our experiments.

The second condition that must be met is that there are no major collisions with the object surface. For this test, a sampling of points representing the entire inner surface of the hand are transformed and checked for penetration into the object surface. For collision checking, we use the same set of 38 points that represent potential contact points on the hand surface (Section III-A). If penetration exceeds a threshold, the

---

[2]In our experiments, we check as many alignment triangles as there are object features. Specifically, each pair of oriented points that were used to create a feature for shape matching is joined with a third, randomly selected point on the object surface to form a test triangle.

alignment is rejected. We used a threshold of $1.5\ cm$ in our experiments.

3) Pose Clustering: At the end of the alignment process, there may be many possible grasps available for a single hand pose and object. These grasps are automatically clustered for use in the next stage of the algorithm. We use K-means clustering [57], with the number of clusters selected using the criterion proposed in Krzanowski and Lau [58]. Clustering was performed on the transformation used to align the hand pose to the object surface. We experimented with a variety of different subspaces that can be formed from this transformation, including using Euler angles or quaternions for the orientation of the hand, and found that for our problem the most intuitive results were produced when clustering in the nine-dimensional space representing the rotation portion of the transformation matrix (i.e., we formed a nine-dimensional vector from the 3x3 rotation matrix of each transformation and performed clustering in the resulting nine-dimensional space). We did not consider translation of the hand during clustering, and we found that the results were still intuitive for the types of objects in our database. If objects are more complex, it may be important to consider translations as well as rotations.

4) Pose Refinement: One grasp from each cluster (the grasp closest to the cluster mean) is selected for further processing. In particular, we refine the grasp to remove collisions and obtain more exact contact locations. We use a local refinement algorithm for this purpose. Jacobian based inverse kinematics (e.g., [59]) is used to iteratively eliminate collisions of points inside the surface and achieve contact for points that are outside the surface. In both cases, contact points are moved toward their nearest neighbors on the object surface. These nearest neighbor points are selected from the object's oriented point representation. The are selected once, at the beginning of the iterative process, and are not changed. We found that keeping these points fixed resulted in more natural looking results. In addition, we impose joint limits on the hand to avoid unnatural poses. Finally, if a contact is too far outside the object surface at the start of the iterative process, we assume that it is not likely to successfully make contact, and we do not attempt to move that point to the surface. These choices were made to avoid unnatural hand poses, but one side effect is that in some cases not all desired contacts are actually achieved. For this reason it is especially important to use a quality metric to prune and rank grasps resulting from the pose alignment, clustering, and refinement processes.

IV. TASK-BASED PRUNING

After shape matching and clustering, a small number of candidate grasps remain. Some of these grasps may be inappropriate for the intended task. They may fail to support the object securely, for example, or the main power of the grasp may be aligned in the wrong direction for the task.

For our system, we choose a force based grasp quality measure to prune and rank candidate grasps. In other words, the quality of a grasp is computed by comparing the forces that can be achieved by a grasp to the forces required for a task. For intuition on a force-based grasp quality measure,
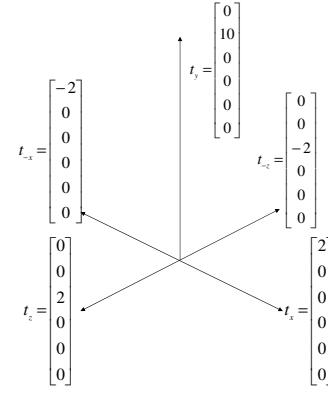

Fig. 10. Example of a task expressed as a set of wrenches that must be exerted on the object. In this case the wrenches are pure forces to be applied through the object center of mass, such as might be required to support the object against gravity.

consider the difference between placing a bowl on a high shelf and holding a pitcher to pour water. In the first case, it may be fine to support the object from the bottom only, as the expected forces are primarily to support the object against gravity. In the second case, however, a secure grasp of the handle is needed to support the pitcher at different orientations as it is tipped for pouring. Similarly, a sander is grasped differently from a drill because of the different forces that must be applied to support the tool and to perform the intended task.

When developing a grasp quality measure for the human hand, we must consider the fact that the ability of our hand to apply forces is different for different force directions. This is caused by the asymmetric structure of the hand mechanism. For example, the flexors of the human hand are typically much stronger than the extensors. Thus, our grasp quality metric is based on a biomechanically plausible estimate of the hand's ability to apply forces, so that task requirements and hand capabilities can be compared in an equitable manner.

A. Grasp Quality

To compute a quality metric that considers both task requirements and hand abilities, we need to quantify these two variables.

1) Task Requirements: A force-based task can be represented generally as a set of wrenches that must be exerted on the object, where a wrench w is a six-dimensional vector of forces f and torques τ:

$$w = \begin{bmatrix} f \\ \tau \end{bmatrix} \qquad (9)$$

We represent a task as the set of wrenches t_i, i=1,...,t which are the extreme points of the convex hull that bounds the space of wrenches required for the task as in [40] [41]. This set of extreme wrench requirements can be user defined when the task requirement is known to the user. For example, Figure 10 shows the task definition used in some of our experiments. This example task consists of pure forces that
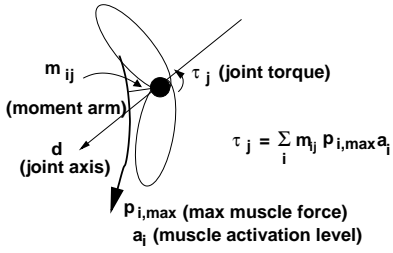
Fig. 11. This figure shows the mapping from muscle activation levels to joint torques.
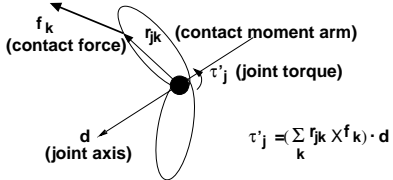

Fig. 12. There is also a well-defined mapping from contact forces to joint torques. For a static grasp, these joint torques must match exactly those generated by muscle activation (Figure 11).

the hand should be able to apply easily to the grasped object. These forces correspond, for example, to supporting the object against gravity when it is vertical or at a relatively small angle from vertical. Torque requirements could be added to the task without changing the algorithm.

2) Hand capabilities: To compute the hand capabilities, we start from an anatomical human hand model. The details of the hand model are described in the Appendix. Given anatomical details derived from the biomechanics literature, we can estimate the maximum forces that can be transmitted from muscles through tendons to joints to contact points and finally to be applied to the object.

To compute the set of wrenches that the hand can apply to a grasped object, we need the following equations. All anatomical parameters for these equations can be found in the Appendix, along with their original references.

First, the joint torques generated by the tendons due to muscle activation can be computed as follows (Figure 11):

$$\tau_J = MPa, \tag{10}$$

where P is $diag(p_{1,max}, p_{2,max}, ..., p_{n,max})$, where $p_{i,max}$ is the maximum force that can be generated along tendon i. Matrix M contains joint moment arm information and converts tendon forces to joint torques. Parameter a is an nx1 vector of activation levels, ranging from 0 (inactive) to 1 (at maximum force), for n tendons.

Then when the hand grasps an object, we can also map from the contact forces to the joint torques as follows (Figure 12):

$$\tau'_J = J^T f, \tag{11}$$

where f is a 3mx1 vector of contact forces with m as the number of contacts, and J^T is the contact Jacobian, which maps contact forces to joint torques.
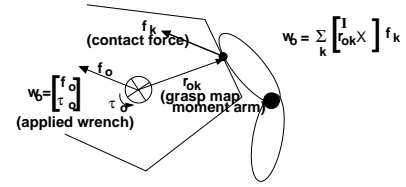

Fig. 13. Finally, we need the mapping from contact forces to forces and torques applied to the object.

Finally we can map the contact forces to wrenches on the object using the grasp matrix G (Figure 13):

$$W = Gf, \tag{12}$$

$$G = \begin{pmatrix} I_3 & I_3 & ... & I_3 \\ R_{o1} & R_{o2} & ... & R_{om} \end{pmatrix}, \tag{13}$$

$$R_{ok} = \begin{pmatrix} 0 & -r_{ok_z} & r_{ok_y} \\ r_{ok_z} & 0 & -r_{ok_x} \\ -r_{ok_y} & r_{ok_x} & 0 \end{pmatrix}, \tag{14}$$

where m is the number of contacts and $r_{ok}=[r_{ok_x} \quad r_{ok_y} \quad r_{ok_z}]^T$ is the grasp map moment arm for the kth contact.

3) Grasp Quality as Optimization: To take both the hand grasp ability and the task requirement into account, we consider the following grasp quality metric:

$$Q = \min_i \frac{\|w_{i,max}\|}{\|t_i\|}, i = 1, ..., t \tag{15}$$

where $w_{i,max}$ is the maximal wrench that can be applied to the object in the direction i. This minimum ratio between the grasp's wrench capability and the task requirement can be interpreted as a safety index of how strong the grasp is when compared to the most difficult task requirement. Figure 14 shows examples of $w_{i,max}$ and t_i in various directions i. Note that this is in the 6D wrench space and $w_{i,max}$ and t_i are all 6D vectors. However, the key computation to be performed is quite simple: repeatedly identify the maximal wrench $w_{i,max}$ that can be applied by the hand to the object in a given direction defined by a task vector t_i.

We note that maximizing the applied wrench in a given direction can be efficiently solved as a linear optimization problem. Let $\alpha_i = \|w_{i,max}\|$ be the value we wish to maximize (Figure 14). In other words, our objective is

$$\text{maximize} (\alpha_i) \tag{16}$$

Let $s_i = \frac{t_i}{\|t_i\|}$ be the unit direction vector for $w_{i,max}$, as shown in Figure 14.

We can specify the following equality constraints for our problem:

$$\begin{pmatrix} -s_i & G & 0 \\ 0 & J^T & -MP \end{pmatrix} \begin{pmatrix} \alpha_i \\ f \\ a \end{pmatrix} = 0 \tag{17}$$

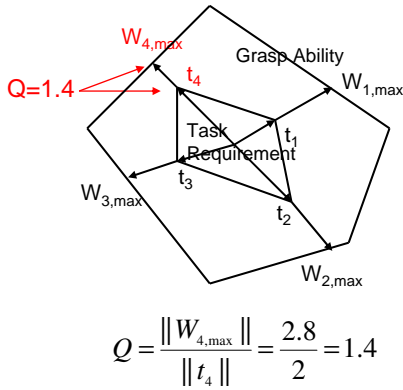$$Q = \frac{\| W_{4,max} \|}{\| t_4 \|} = \frac{2.8}{2} = 1.4$$

Fig. 14. Example of $w_{i,max}$ and t_i on direction i in 6D force / torque (wrench) space. The outer convex hull is the grasp's capability and the inner convex hull is the task requirement.

The first row ensures that the applied wrench Gf is in the desired direction. Specifically, the applied wrench must equal $\alpha_i s_i$, which by definition is $w_{i,max}$. The second row ensures zero acceleration at the finger joints by equating torques produced by tendon activation (MPa) to those required to generate the contact forces ($J^T f$).[3] Note that this equation is very similar to that in [46], extending that work to include tendon activation levels.

To take the friction between the finger and the object at the contact points into account to ensure that the object does not slip, we first approximate the Coulomb friction cone with a friction pyramid to keep the system linear (Figure 15). In practice, we use a friction pyramid having four sides (i.e., defined by four halfplanes), as shown in the figure. Let the normals of the friction pyramid halfplanes be n_{jk}, where j is the index of the contact point and k is the index of the halfplane. Now we can define the following inequality constraints which ensure that contact forces remain within the friction cones at the contact points:

$$n_{jk} \cdot f_j < 0 \qquad j = 1, \ldots, m; \quad k = 1, \ldots, h \qquad (18)$$

where m is the number of contact points and h is the number of halfplanes in each friction pyramid.

To compute the grasp quality metric in Equation 15, we perform the optimization problem defined in Equations 16 through 18 for $i=1, \ldots, t$, setting $\| w_{i,max} \| = \alpha_i$ at the end of each optimization step, and taking the minimum $\frac{\| w_{i,max} \|}{\| t_i \|}$ as our quality metric.

V. THE GRASP DATABASE

We illustrate our results with a database consisting of 17 hand poses obtained from motion capture data. Objects that were grasped to form the database are shown in silhouettes in the right hand portion of Figure 16. The objects on the left of

---

[3] A more complete analysis would consider the distinction between sets of contact forces that can be resisted by the hand and those that can be actively applied. In our experiments and in most practical grasps, the two sets of forces are identical and the constraint equations in Equation 17 are sufficient. Please see [5] for more detail.
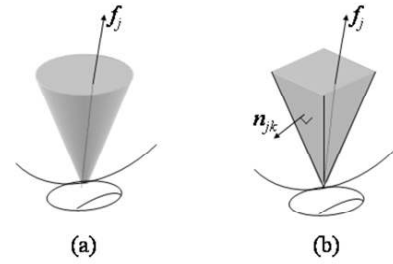

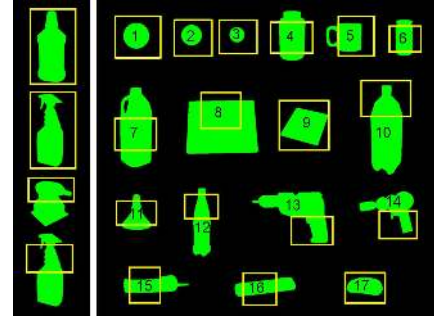Fig. 15. (a)Friction cone. (b) Friction pyramid. The contact force must fall within the friction cone/pyramid.

Fig. 16. (Right) Silhouettes of the 17 objects in our database. The highlighted "partial objects" show the regions of those objects that were grasped to form the hand pose database. (Left) New objects that were not in the database were used as query objects to test our algorithm. Specifically, we created four queries pictured from top to bottom as (1) the entire surface geometry of a mouthwash bottle, (2) the entire surface geometry of a spray bottle, (3) a portion of the surface of a sander, and (4) a portion of the surface of a spray bottle. Figure 21 shows results from these four queries.

Figure 16 are test objects that were not a part of the database. Grasps of the 17 objects in the database were captured using a Vicon optical motion capture system using the skeleton model described in the Appendix. Geometric models for the objects were created manually for each of the objects grasped. For each hand pose, contact points were identified manually by selecting from a discrete set of contact points on the hand surface, and the normals at those contact points were obtained by taking the normal from the closest point on the object surface.

In terms of the standard grasp hierarchy (e.g. [15]), our database focuses on power grasps, because these grasps tend to have a large number of contacts between hand and object. The database includes power sphere and cylinder grasps for objects of a variety of sizes and having a variety of differences from the primitive shapes. For example, the grasp of the lightbulb is more like a cone grasp than a cylinder grasp, and the drill handle has a roughly ellipsoidal cross section. In some grasps such as the phone and remote control grasp, the thumb does not wrap around the object in opposition to the fingers, but lies along the main axis of the object. The database also contains grasps of flat objects–a book and a CD case–where the object makes contact at many places on the fingers and has a single area of contact on the thumb.

VI. RESULTS

Results are presented to illustrate the discriminative power of our features, to show the effectiveness of the shape match-
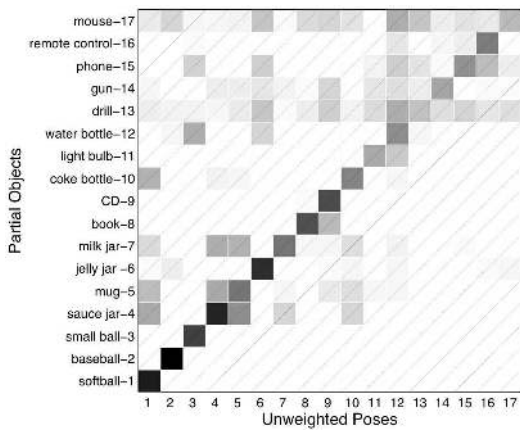
Fig. 17. Evaluation using the *unweighted* distance metric $E_{i,j}$ (Equation 1). All hand poses in the database are compared to all of the objects for which grasps were collected. Darker colors indicate lower values of $E_{i,j}$, or closer matches. A perfect discriminator would have black boxes on the diagonal and white everywhere else.
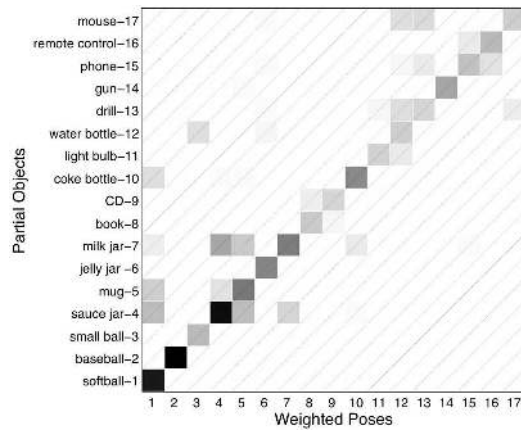

Fig. 18. Evaluation using the *weighted* distance metric $E_{i,j}^w$ (Equation 3). All hand poses in the database are compared to all of the objects for which grasps were collected.
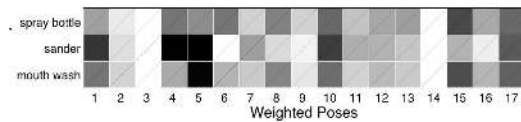

Fig. 19. Test of three new objects against the hand pose database. The *entire object* was considered. The figure shows weighted results, using $E_{i,j}^w$.

ing, clustering, and refinement algorithms, and to demonstrate use of our grasp quality metric.

A. Discriminative Power of Features

We first tested whether our matching algorithm would select appropriate grasps for the objects used to form the database. The motivation behind this experiment is that if our algorithm does not work on the original data, then we must conclude that our feature set is not sufficiently discriminative.

Figures 17 and 18 show results from this initial experiment. Figure 17 illustrates results using the unweighted distance $E_{i,j}$ from Equation 1, and Figure 18 illustrates weighted results using distance $E_{i,j}^w$ from Equation 3. Hand poses 1 through 17 are shown on the x-axis, and query objects 1 through 17 are shown on the y-axis. The list of objects is:

- 1-Softball, 2-Baseball, 3-SmallBall, 4-SauceJar, 5-Mug, 6-JellyJar, 7-MilkJar, 8-LargeBook, 9-CD, 10-CokeBottle, 11-LightBulb, 12-WaterBottle, 13-Drill, 14-BubbleGun, 15-Phone, 16-RemoteControl, 17-Mouse

So, for example, column 15, row 15 represents the hand pose from the phone grasp compared to the object geometry of the phone, and column 10, row 15 represents the hand pose from the coke bottle grasp compared to the object geometry of the phone. All of these results are run with the user highlighting the relevant region of the object to consider. For example, for the drill geometry, the user selected the handle of the drill as shown in Figure 16 and other parts of the drill were not considered. In these two figures, a purely black cell would represent zero distance. The diagonal terms are not exactly zero due to errors in estimating contact positions from the motion capture data and due to discrete sampling of points on the object surface. In some cases (e.g., the grasp of the mouse and the drill), the motion capture data was quite poor, resulting in less good matches for the original object. Better motion capture technology or post-processing to improve our estimates of the hand poses in the database would solve this problem.

Clearly the weighting term adds a great deal to our ability to discriminate between different grasps based on shape. However, note that even in Figure 18 some poses / objects show quite similar results. In particular, objects 4, 5, and 7 group together, as do objects 8 and 9 and objects 15 and 16. These groupings are not surprising. Objects 4, 5, and 7 are all medium sized objects of approximately cylindrical shape. Objects 8 and 9 are flat objects (a book and a CD). Objects 15 and 16 are a phone and a remote control, which are likewise grasped in a similar manner. This experiment shows that the weighting term used in Equation 3 is effective for pruning spurious matches.

B. Results for Feature Matching on New Objects

Our second set of experiments was to find good matches for objects that were not in the database by comparing the feature sets computed from those objects to the feature sets computed from hand poses. We selected three objects of moderate complexity for this experiment: 1-MouthwashBottle, 2-Sander, and 3-SprayBottle. Figures 19 and 20 illustrate results for these three objects. Figure 19 contains results of searching for matches using the entire object geometry. Figure 20 contains results in a similar format when the user has highlighted the portion of the object to be grasped. Only results using weighted metric $E_{i,j}^w$ are shown. From these images we can see that the feature matching process can reduce the number of candidate poses to be considered, especially when the user has indicated a portion of the object to be grasped, as in Figure 20.

C. Results for Aligning, Clustering and Refinement

Based on the similarity matrices in Figures 19 and 20, we selected the best hand poses for further processing. Specifi-
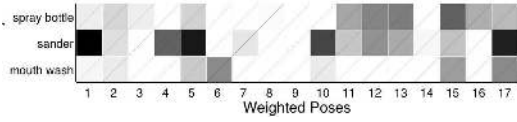
Fig. 20. Test of three new objects against the hand pose database. Only *a portion of the object* was considered. The figure shows weighted results, using $E_{i,j}^{w}$.

WHOLE OBJECT		PARTIAL OBJECT	
Mouthwash		**Mouthwash**	
5-Mug	399 (4)	17-Mouse	0 (-)
15-Phone	15 (4)	6-JellyJar	218 (4)
17-Mouse	2 (-)	15-Phone	7 (-)
10-CokeBottle	380 (4)	5-Mug	360 (4)
1-Softball	721 (2)	10-CokeBottle	11 (4)
Sander		**Sander**	
5-Mug	2 (-)	1-Softball	31 (4)
4-SauceJar	7 (-)	5-Mug	1 (-)
1-Softball	59 (2)	17-Mouse	0 (-)
SprayBottle		**SprayBottle**	
15-Phone	4 (-)	15-Phone	0 (-)
17-Mouse	13 (4)	13-Drill	0 (-)
6-JellyJar	163 (4)	12-WaterBottle	243 (2)
		11-Lightbulb	13 (3)

TABLE II

RESULTS FOR THREE NEW OBJECTS. (LEFT) THE ENTIRE OBJECT
SURFACE WAS USED. (RIGHT) A PORTION OF THE OBJECT WAS SELECTED.
BOTH ALIGNMENTS AND CLUSTERS (IN PARENTHESES) ARE SHOWN.
HAND POSES WITH 8 OR FEWER ALIGNMENTS WERE NOT CLUSTERED.

cally, for each hand pose / object pair, we identified candidate alignments of hand to object, clustered the results, and refined the median pose from each cluster to obtain representative grasps, which were displayed to the user.

Table II shows the number of alignments and the number of clusters returned for the three new objects in each of the conditions tested. For example, the top section of the left column shows alignments for the mouthwash container when the entire object was considered. The five best matching hand poses based on their feature sets were the mug, phone, mouse, coke bottle, and softball poses. The mug pose resulted in 399 valid alignments with the object surface, which could be grouped into 4 clusters based on orientation of the hand pose. When only a portion of the mouthwash container was considered, the best matching hand poses were from the mouse, jelly jar, phone, mug, and coke bottle grasps. Although the mouse grasp was best in terms of its feature set, no valid alignments could be found on the object surface. In other words, even though similar features to all pairs of points in the grasp could be found on the object surface, none of these matching pairs led to a complete matching grasp.

Figure 21 shows some results of the shape matching and clustering phase. Each collection of images shows all cluster results where a single hand pose matches the given object. For example, the top two rows show the median cluster representative for each cluster matching the coke bottle pose to the mouthwash object.
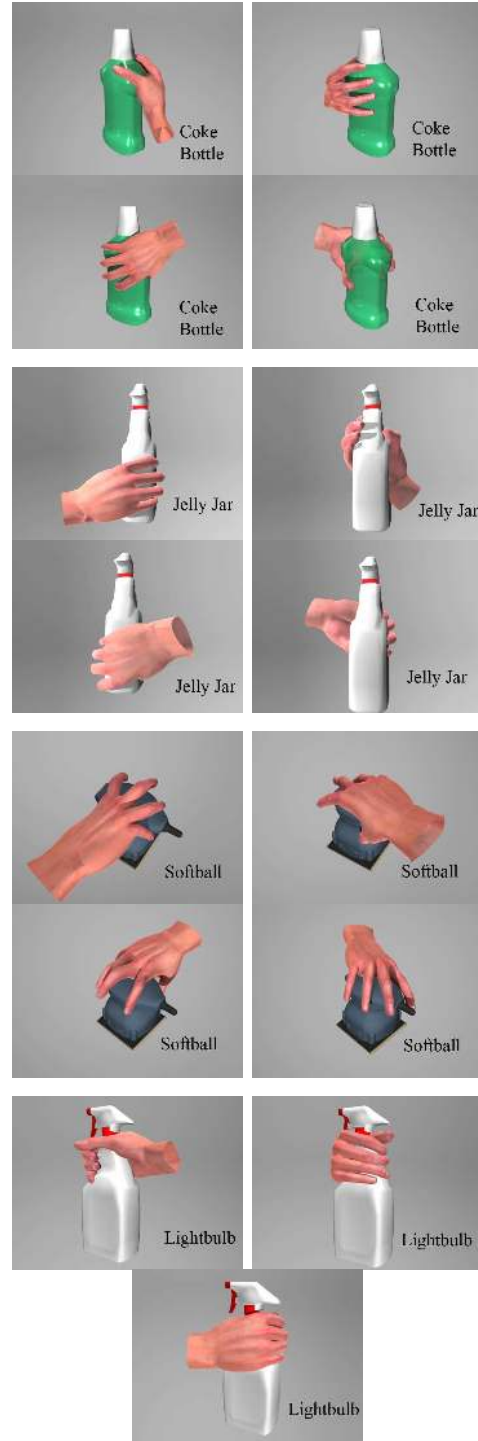
Fig. 21. Cluster results for four queries presented to our system. The left hand portion of Figure 16 shows the portions of the object geometry considered for these queries. (1) For the entire surface geometry of a mouthwash bottle, 4 clusters were found using the coke bottle pose. (2) For the entire surface geometry of a spray bottle, 4 clusters were found using the jelly jar pose. (3) For a portion of the surface of a sander, 4 clusters were found using the softball pose. (4) For a portion of the surface of a spray bottle, 3 clusters were found using the lightbulb pose.
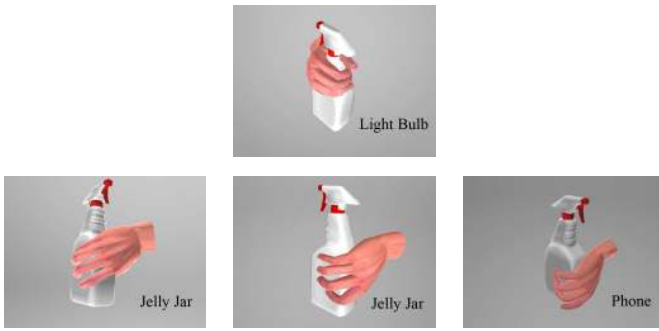
Fig. 22.   Best hand poses for the spraybottle. The first grasp (lightbulb grasp) is the best among the five candidates of water bottle and light bulb grasps, considering the partial spraybottle shape. The next three grasps are chosen from the twelve candidates (phone, jellyjar, and mouse grasps) considering the entire spraybottle shape. Among these three grasps, the first two are jellyjar grasps and the last one is a phone grasp.

Fig. 23.   Best hand poses for the mouthwash bottle. The first grasp (phone grasp) is the best among nineteen candidates (phone, jellyjar, cokebottle, and mug grasps) considering the partial mouthwash shape. The last two grasps are chosen from the sixteen candidates (phone, mouse, cokebottle, mug, and softball grasps) considering the entire mouthwash shape. The first one is a cokebottle grasp and the second one is a softball grasp.

D. Results from Task-Based Pruning

We define a task for our experiments as statically grasping the object with the ability to support it against gravity in vertical and near vertical orientations (Figure 10). We compute the quality metric using this task requirement and choose the best grasps from the grasp candidates obtained from shape matching and clustering. Figures 22 to 24 show results for the spray bottle, mouthwash bottle, and sander. Note that if there is more than one grasp with a similar, high quality metric, all such grasps are shown.

We note that some grasps are typically better for a given object and some are generally worse. For example, among the candidates for the mouthwash bottle (considering the whole object), the mug grasp candidates have very poor performance with an average metric of 0.602, compared to the phone grasp of 6.01, mouse grasp of 1.34, cokebottle grasp of 8.065, and softball grasp of 5.537. Figure 25 shows an example of the mug grasp. Note that the index finger and middle finger do not

Fig. 24.   Best hand poses for the sander. The first two softball grasps are the best among the five candidates (mug and softball grasps) considering the partial sander shape. The last grasp is the saucejar grasp, the best among the eleven candidates (saucejar, mug, and softball grasps) considering the whole sander shape.

Fig. 25.   Mug grasps in general have worse performance.

Fig. 26.   Two softball grasp candidates with different metric values, caused by the difference in contact points.

contact the object, which is the primary reason for the poor performance of this grasp.

We notice that the shape matching algorithm sometimes cannot achieve the necessary contact points for a stable grasp. Grasps that cannot be made stable also tend to appear implausible, and we would like to remove them from consideration. In our system, the grasp candidates having this problem are screened out by the grasp quality metric. For example, one of the softball candidates for the mouthwash bottle (whole) only has contacts on the index, ring, and little finger (shown as the first grasp in Figure 26.) With the locations of the contacts on the object all at one side of the object, this grasp cannot be stabilized. As a result, the grasp earns a metric of zero. However, another softball grasp candidate (the second grasp in Figure 26) has an extra contact on the thumb that can apply torque to the object to balance the contacts on the other fingers. The second grasp has a metric of 12.56.

Another example is given in Figure 27. The first grasp has no thumb contact and has a metric of 1.55, while the other three all have thumb contacts and the metrics are 24.3, 13.2, and 13.58, respectively.

Because of the much stronger flexors in the fingers and the wrist, grasps with the palm facing upward, which cause the forces in the y-direction (upward) to be generated by the stronger flexors, usually have better performance. For example, the last three grasps in Figure 27 are similar phone grasp
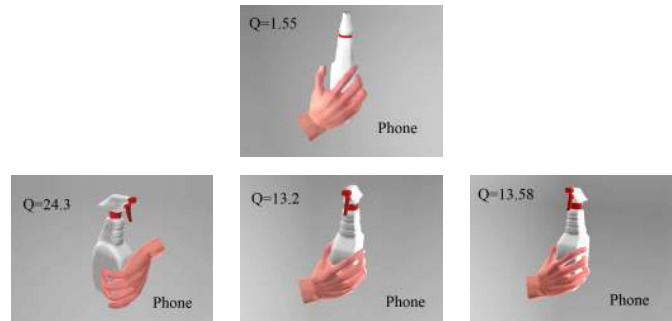

Fig. 27.   Four phone grasp candidates for the spraybottle (whole). The metrics are different because of the different contact points and hand orientation.

candidates with very similar contact points for the spraybottle (whole object), but with different orientations. The first one is facing palm-upward and has a metric of 24.3, while the other two are facing palm-downward and the metrics are only 13.2 and 13.58. Of course, other factors such as reachability or comfort of the arm may cause the grasps on the right side of Figure 27 to be preferred over the grasp on the left side of that figure.

VII. CONCLUSIONS AND FUTURE WORK

This paper describes a shape matching algorithm for synthesizing humanlike enveloping grasps that is inspired by work in geometric model matching. We have described an approach that handles the specific challenges of applying shape matching to grasp synthesis. The availability of only partial information is handled using a representative descriptor approach, and the lack of good local features is handled by sampling a global shape function. We introduce a novel global shape function that emphasizes the relations between relative positions and normals of pairs of oriented points and we introduce a mechanism to weight features relative to one another to improve the discriminatory power of our shape matching algorithm. To prune unsuitable matches from consideration, we introduce a grasp quality metric that compares forces that can be exerted by the hand to those required for the task.

There are several limitations to the current algorithm that must be addressed in future work. First of all, our algorithm assumes that the animated hand model is very similar to the hand of the actor whose motion was originally captured. If the same joint angles are assumed for a similar hand, the database should still represent plausible grasps. However, adapting grasps in the database to extremely different hand kinematics may produce unnatural results. Retargeting grasps in the database to different hands while preserving the "intent" of the grasp is an interesting topic of future work.

The algorithm uses a number of sampling steps. The sampling process is straightforward, easy to implement, and effective. However, it is inefficient, and processing a new object to obtain a small set of grasps to present to the user requires several minutes of running time. There are many ways this process could be made more efficient (e.g., with fast nearest neighbor algorithms), although we expect that obtaining real-time performance may be challenging with today's computers. In addition to efficiency considerations, sampling has the potential to poorly sample high frequency or very thin features. Objects for which power grasps are applicable, however, do not typically have such features; we did not have any such difficulties with the objects in our database.

We considered power grasps almost exclusively, because the large number of contacts between the hand and object provided a great deal of useful shape information. Precision grasps could also be addressed using this framework. However, if there are only two, three, or four contacts between the hand and object, the possibility of more efficient algorithms should be considered.

Our current refinement algorithm has some difficulty in achieving all of the desired contacts (e.g., see Figures 25 and 27). There are two primary reasons for this difficulty. The first is that the set of algorithms we use are threshold based. For example, if a contact is too far from the surface, we do not attempt to move it to a surface point. However, there is no one threshold that works in all cases, and although we attempted to set thresholds that produced generally good results, there were still a number of failures, the most extreme of which are shown here. The second reason for difficulty in achieving desired contacts is that the standard Jacobian based IK algorithm we were using produced highly unnatural grasps in our first implementation (e.g., unnatural twisting of the fingers) and we chose to impose stringent joint limits in order to avoid these problems. A more satisfying solution may be obtained by developing an IK algorithm that takes hand synergies into account or formulates the pose refinement task as a more general optimization or control problem. Improving the refinement algorithm is a topic of future research.

Our results show that the shape matching algorithm can find very good matches between the hand and the object surface. In fact, we were surprised at the variety of good matches found using our small database of 17 grasps, even after functionally poor grasps were pruned using our quality metric. At present, a user must select the desired grasp from among the possibilities presented by the system, because some of the grasps appear unintuitive. For example, the hand may appear to be grasping the object "backwards" or "upside down." Objects such as tools with buttons or levers may require customized shape matching algorithms that can exploit these features (e.g., see Miyata et al. [60]), and for a fully automatic system, aspects of natural grasps other than ability to apply forces must be taken into account. Much research remains to be done to better understand human grasping (e.g., see [61]) and to develop algorithms that quantify the "naturalness" of grasps. For example, comfort and reachability of the hand and arm pose must certainly be considered.

APPENDIX

Based on the biomechanics work by Valero-Cuevas et al. [62] [63], we use a hand model that has the following configuration (see Figure 28): The fingers have four degrees of freedom (DOF). The metacarpophalangeal (MCP) joint has two DOF of flexion/extension and adduction/abduction. The proximal interphalangeal (PIP) and the distal interphalangeal (DIP) joints both have one DOF of flexion/extension. The thumb has five DOF in total. The carpometacarpal(CMC), MCP and thumb interphalangeal (IP) joint each has one DOF of flexion/extension. The CMC and MCP joint both have another DOF of adduction/abduction. In the actual human hand, the ring and little finger have two additional DOFs of flexion/extension and adduction/abduction at the CMC joint. Since they only provide a relatively small range of motion, we consider these extra DOFs fixed.

To obtain the maximum muscle force for the index finger tendons, Valero-Cuevas et al. [62] uses physiological cross-sectional area (PCSA) numbers and a conversion from PCSA
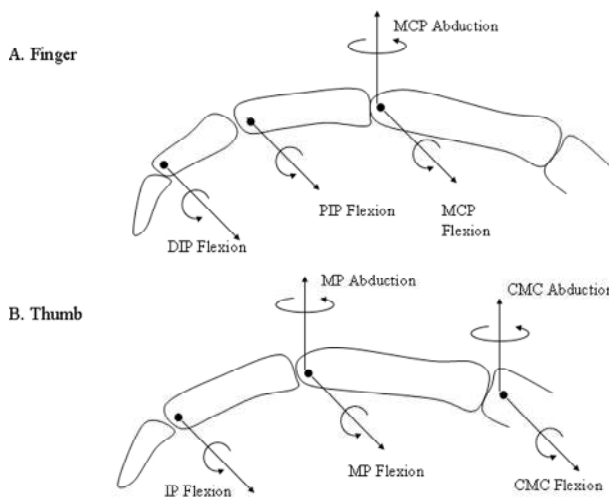
Fig. 28.   A. Finger model; B. Thumb model.

to force of 35 N/cm^2. Lacking PCSA data for all tendons, we combine this computation with the relative tension data from Brand and Hollister [64] to get the maximum muscle force for each tendon. First we compute the maximum muscle force of the FDP tendon for the index finger: $FDPI = 4.1 * 35 = 143.5\ N$, where 4.1 is the PCSA number for the FDP tendon of the index finger. Then the maximum muscle forces for other tendons can be generated with the relative tension data. For example, the maximum force of the FDS tendon for the middle finger is $143.5 * 3.4/2.7 = 180.7\ N$, where 2.7 and 3.4 are the relative tension data for the FDP and the FDS tendon, respectively. Moment arm and relative force data for each tendon are listed in Table III.

ACKNOWLEDGMENTS

REFERENCES

[1] G. Mamic and M. Bennamoun, "Representation and recognition of 3D free-form objects," *Digital Signal Processing*, vol. 12, no. 1, pp. 47–76, 2002.

[2] R. Osada, T. Funkhouser, B. Chazelle, and D. Dobkin, "Shape distributions," *ACM Transactions on Graphics*, vol. 21, no. 4, 2002.

[3] A. Frome, D. Huber, R. Kolluri, T. Buelow, and J. Malik, "Recognizing objects in range data using regional point descriptors," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2004.

[4] Y. Li and N. S. Pollard, "A shape matching algorithm for synthesizing humanlike enveloping grasps," in *IEEE-RAS International Conference on Humanoid Robots*, 2005.

[5] J. L. Fu and N. S. Pollard, "On the importance of asymmetries in grasp quality metrics for tendon driven hands," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

[6] B. Douville, L. Levison, and N. Badler, "Task level object grasping for simulated agents," *Presence*, vol. 5, no. 4, pp. 416–430, 1996.

[7] H. Rijpkema and M. Girard, "Computer animation of knowledge-based human grasping," in *Computer Graphics (SIGGRAPH 91 Proceedings)*, vol. 25, 1991, pp. 339–348.

[8] R. Mas Sanso and D. Thalmann, "A hand control and automatic grasping system for synthetic actors," in *Eurographics '94*, 1994.

[9] Z. Huang, R. Boulic, N. Magnenat Thalmann, and D. Thalmann, "A multi-sensor approach for grasping and 3D interaction," in *Proc. Computer Graphics International*, 1995.

[10] Y. Aydin and M. Nakajima, "Database guided computer animation of human grasping using forward and inverse kinematics," *Computers and Graphics*, vol. 23, pp. 145–154, 1999.

[11] G. A. Bekey, H. Liu, R. Tomovic, and W. J. Karplus, "Knowledge-based control of grasping in robot hands using heuristics from human motor skills," *IEEE Transactions on Robotics and Automation*, vol. 9, no. 6, pp. 709–722, 1993.

[12] T. Iberall, "Human prehension and dextrous robot hands," *International Journal of Robotics Research*, vol. 16, no. 3, June 1997.

[13] G. Schlesinger, "Der mechanische Aufbau der künstlichen Glieder (the mechanical building of artifical limbs)," in *Ersatzglieder und Arbeitshilfen für Kriegsbeschädigte und Unfallverletzte (Prosthetic and Occupational Help for War Victims and Accident Victims)*, M. Borchardt, Ed.   Springer, 1919.

[14] J. Napier, *Hands*.   Princeton, NJ: Princeton Univ. Press, 1980.

[15] M. R. Cutkosky and R. D. Howe, "Human grasp choice and robot grasp analysis," in *Dextrous Robot Hands*, S. Venkataraman and T. Iberall, Eds.   New York: Springer-Verlag, 1990.

[16] N. S. Pollard and V. B. Zordan, "Physically based grasping control from example," in *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 2005.

[17] P. G. Kry and D. K. Pai, "Interaction capture and synthesis," *ACM Transactions on Graphics*, vol. 25, no. 3, 2006.

[18] G. ElKoura and K. Singh, "Handrix: Animating the human hand," in *ACM SIGGRAPH / Eurographics Symposium on Computer Animation*, 2003.

[19] D. J. Wiley and J. K. Hahn, "Interpolation synthesis of articulated figure motion," *IEEE Computer Graphics and Applications*, vol. Nov/Dec, pp. 39–45, 1997.

[20] L. Kovar and M. Gleicher, "Automated extraction and parameterization of motions in large data sets," *ACM Transactions on Graphics*, vol. 23, no. 3, 2004.

[21] K. Yamane, J. Kuffner, and J. K. Hodgins, "Synthesizing animations of human manipulation tasks," *ACM Transactions on Graphics*, vol. 23, no. 3, 2004.

[22] R. J. Prokop and A. P. Reeves, "A survey of moment-based techniques for unoccluded object representation and recognition," *Graphical Models and Image Processing*, vol. 54, pp. 438–460, 1992.

[23] J. W. H. Tangelder and R. C. Veltkamp, "A survey of content based 3D shape retrieval methods," in *Proceedings Shape Modeling International*, 2004, pp. 145–156.

[24] A. E. Johnson and M. Hebert, "Using spin-images for efficient multiple model recognition in cluttered 3-D scenes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 433–449, 1999.

[25] C. Dorai and A. K. Jain, "COSMOS—A representation scheme for 3D free-form objects," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 10, pp. 1115–1130, 1997.

[26] M. Ankerst, G. Kastenmller, H.-P. Kriegel, and T. Seidl, "Nearest neighbor classification in 3D protein databases," in *Proc. 7th Int. Conf. on Intelligent Systems for Molecular Biology*, 1999.

[27] R. Ohbuchi, T. Minamitani, and T. Takei, "Shape-similarity search of 3D models by using enhanced shape functions," in *Proceedings of the Theory and Practice of Computer Graphics*, 2003.

[28] G. Mori, S. Belongie, and J. Malik, "Shape contexts enable efficient retrieval of similar shapes," in *IEEE Computer Vision and Pattern Recognition (CVPR)*, 2001.

[29] X. Markenscoff and C. H. Papadimitriou, "Optimum grip of a polygon," *International Journal of Robotics Research*, vol. 8, no. 2, pp. 17–29, 1989.

[30] D. G. Kirkpatrick, B. Mishra, and C. K. Yap, "Quantitative steinitz's theorems with applications to multifingered grasping," in *Proc. 20th ACM Symposium on Theory of Computing*, Baltimore, Maryland, May 1990, pp. 341–51.

[31] C. Ferrari and J. Canny, "Planning optimal grasps," in *IEEE Int. Conference on Robotics and Automation*, Nice, France, May 1992, pp. 2290–2295.

[32] Q. Lin, J. Burdick, and E. Rimon, "A stiffness-based quality measure for compliant grasps and fixtures," *IEEE Transactions on Robotics and Automation*, vol. 16, no. 6, pp. 675–88, 2000.

[33] B. Mishra, "Grasp metrics: optimality and complexity," in *Proc. 1st International Workshop on the Algorithmic Foundations of Robotics (WAFR)*, Stanford, CA, 1994, pp. 137–165.

[34] V.-D. Nguyen, "Constructing force-closure grasps," *International Journal of Robotics Research*, vol. 7, no. 3, pp. 3–16, 1988.

[35] J. Ponce, D. Stam, and B. Faverjon, "On computing two-finger force-closure grasps of curved 2d objects," *International Journal of Robotics Research*, vol. 12, no. 3, pp. 263–73, 1993.

[36] J. Ponce and B. Faverjon, "On computing three-finger force-closure grasps of polygonal objects," *IEEE Transactions on Robotics and Automation*, vol. 11, no. 6, pp. 868–81, 1995.

[37] J. Ponce, S. Sullivan, A. Sudsang, J.-D. Boissonnat, and J.-P. Merlet, "On computing four-finger equilibrium and force-closure grasps of polyhedral objects," *International Journal of Robotics Research*, vol. 16, no. 1, pp. 11–35, 1997.

[38] Z. Li and S. Sastry, "Task-oriented optimal grasping by multifingered robot hands," *IEEE Journal of Robotics and Automation*, vol. 4, no. 1, pp. 32–44, 1988.

[39] X. Zhu, H. Ding, and H. Li, "A quantitative measure for multi-fingered grasps," *2001 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. Proceedings*, pp. 213–19, 2001.

[40] X. Zhu and J. Wang, "Synthesis of force-closure grasps on 3-d objects based on the q distance," *IEEE Transactions on Robotics and Automation*, vol. 19, no. 4, pp. 669–79, 2003.

[41] N. S. Pollard, "Closure and quality equivalence for efficient synthesis of grasps from examples," *International Journal of Robotics Research*, vol. 23, no. 2, pp. 595–614, June 2004.

[42] G. Liu, J. Xu, X. Wang, and Z. Li, "On quality functions for grasp synthesis, fixture planning, and coordinated manipulation," *IEEE Transactions on Automation Science and Engineering*, vol. 1, no. 2, pp. 146–62, 2004.

[43] Y. Zhang, F. Gao, Y. Zhang, and W. Gruver, "Evaluating the quality of grasp configurations for dextrous hands," *Proceedings. 1997 IEEE International Conference on Robotics and Automation*, pp. 100–105, 1997.

[44] Z. Li, P. Hsu, and S. Sastry, "Grasping and coordinated manipulation by a multifingered robot hand," *International Journal of Robotics Research*, vol. 8, no. 4, pp. 33–50, 1989.

[45] J. Salisbury and J. Craig, "Articulated hands: force control and kinematic issues," *International Journal of Robotics Research*, vol. 1, no. 1, Spring 1982.

[46] C. Melchiorri, "Multiple whole-limb manipulation: an analysis in the force domain," *Robotics and Autonomous Systems*, vol. 20, no. 1, pp. 15–38, 1997.

[47] R. Kurtz and V. Hayward, "Dexterity measure for tendon actuated parallel mechanisms," *91 ICAR. Fifth International Conference on Advanced Robotics. Robots in Unstructured Environments*, pp. 1141–6, 1991.

[48] P. Voglewede and I. Ebert-Uphoff, "Overarching framework for measuring closeness to singularities of parallel manipulators," *IEEE Transactions on Robotics*, vol. 21, no. 6, pp. 1037–45, 2005.

[49] E. Chinellato, A. Morales, R. Fisher, and A. del Pobil, "Visual quality measures for characterizing planar robot grasps," *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)*, vol. 35, no. 1, pp. 30–41, 2005.

[50] S. Zechloul, B. Blanchard, and J. Pamanes, "Optimization of kinematics performances of manipulators under specified task conditions," in *Theory and Practice of Robots and Manipulators:Proceedings of RoManSy 10, the Tenth CISM-IFTOMM Symposium*, 1995, pp. 247–255.

[51] J. Kerr and B. Roth, "Analysis of multifingered hands," *International Journal of Robotics Research*, vol. 4, no. 4, pp. 3–17, Winter 1986.

[52] M. Buss, H. Hashimoto, and J. Moore, "Dextrous hand grasping force optimization," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 3, pp. 406–18, 1996.

[53] S. B. Nokleby, R. Fisher, R. P. Podhorodeski, and F. Firmani, "Force capabilities of redundantly-actuated parallel manipulators," *Mechanism and Machine Theory*, vol. 40, no. 5, pp. 578–99, 2005.

[54] A. Bicchi and D. Prattichizzo, "Analysis and optimization of tendinous actuation for biomorphically designed robotic systems," *Robotica*, vol. 18, pp. 23–31, 2000.

[55] L. Han, J. Trinkle, and Z. Li, "Grasp analysis as linear matrix inequality problems," *IEEE Transactions on Robotics and Automation*, vol. 16, no. 6, pp. 663–74, 2000.

[56] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, "The Princeton shape benchmark," in *Proceedings Shape Modeling International*, 2004.

[57] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*. New York: John Wiley & Sons, Inc., 2001.

[58] W. J. Krzanowski and Y. T. Lai, "A criterion for determining the number of groups in a data set using sum-of-squares clustering," *Biometrics*, vol. 44, pp. 23–34, 1988.

[59] J. Zhao and N. I. Badler, "Inverse kinematics positioning using nonlinear programming for highly articulated figures," *ACM Transactions on Graphics*, vol. 13, no. 4, pp. 313–336, 1994.

[60] N. Miyata, M. Kouchi, T. Kurihara, and M. Mochimaru, "Analysis of human hand posture to estimate mouse-operation," in *IEEE Int'l. Conf. on Systems, Man, and Cybernetics*, 2003, pp. 3446–3451.

[61] G. Baud-Bovy, P. D., and B. N, "Does torque minimization yield a stable grasp?" *Workshop on Multi-point interaction in Robotics and Virtual Reality, ICRA*, pp. 1037–45, 2004.

[62] F. J. Valero-Cuevas, F. E. Zajac, and C. G. Burgar, "Large index-finger forces are produced by subject-independent patterns of muscle excitation," *Journal of Biomechanics*, vol. 31, pp. 693–703, 1998.

[63] F. Valero-Cuevas, M. Johanson, and J. Towles, "Towards a realistic biomechanical model of the thumb: the choice of kinematic description may be more critical than the solution method or the variability/uncertainty of musculoskeletal parameters," *Journal of Biomechanics*, vol. 36, no. 7, pp. 1019–30, 2003.

[64] P. W. Brand and A. M. Hollister, *Clinical Mechanics of the Hand*, 3rd ed. St. Louis, MO: Mosby - Year Book, Inc., 1999.

[65] K. An, Y. Ueba, E. Chao, W. Cooney, and R. Linscheid, "Tendon excursion and moment arm of index finger muscles," *Journal of Biomechanics*, vol. 16, no. 6, pp. 419–25, 1983.

[66] I. Albrecht, J. Haber, and H. Seidel, "Construction and animation of anatomically based human hand models," *ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 98–109, 2003.

[67] W. Smutz, A. Kongsayreepong, R. Hughes, G. Niebur, W. Cooney, and K.-N. An, "Mechanical advantage of the thumb muscles," *Journal of Biomechanics*, vol. 31, no. 6, pp. 565–70, 1998.

Ying Li received the BS an MS degrees in mechanical engineering from Jilin University of China. She earned PhD. Degree in system information engineering from Kagoshima University of Japan in 2002. She worked at the School of Computer Science, Carnegie Mellon University as postdoctoral fellow from 2004–2006. Her research focuses on graphics and robotics hand grasping. Recent publications are listed at www.cs.cmu.edu/ yingli.
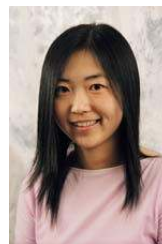
Jiaxin Fu received the bachelor's degree in computer science from Middlebury College in 2003. She is pursuing the PhD degree in Robotics at School of Computer Science at Carnegie Mellon University. Her main research interest is to model realistic hand and arm motion interacting with the environment for both graphics and robotics.

Nancy Pollard is an Associate Professor in the Robotics Institute and the Computer Science Department at Carnegie Mellon University. She received her PhD in Electrical Engineering and Computer Science from the MIT Artificial Intelligence Laboratory in 1994, where she performed research on grasp planning for articulated robot hands. Before joining CMU, Nancy was an Assistant Professor and part of the Computer Graphics Group at Brown University. Her primary research objective is to understand how to create natural motion for animated human characters and humanoid robots.

Finger	Anatomical Name	Relative Tension [a]	Joints/DOF	| Moment Arm | [b]
Index [c]	Flexor Digitorum Superficialis	2.0	MCP flex.	11.9
			MCP add.	1.7
			PIP flex.	6.2
	Flexor Digitorum Profundus	2.7	MCP flex.	11.1
			MCP add.	1.1
			PIP flex.	7.9
			DIP flex.	4.1
	Extensor Digitorum Communis	1.0	MCP ext.	8.6
			MCP abd.	0.2
			PIP ext.	2.8
			DIP ext.	2.2
	Extensor Indicis	1.0	MCP ext.	9
			MCP add.	1.3
			PIP ext.	2.6
			DIP ext.	1.9
	Lumbrical I	0.2	MCP flex.	9.3
			MCP abd.	4.8
			PIP ext.	1.8
			DIP ext.	0.7
	Palmar Interosseus I	1.3	MCP flex.	6.6
			MCP add.	5.8
			PIP ext.	2.6
			DIP ext.	1.6
	Dorsal Interosseus I	3.2	MCP flex.	3.7
			MCP abd.	6.1
Middle [d]	Flexor Digitorum Superficialis	3.4	MCP flex.	11.9
			MCP add.	1.7
			PIP flex.	6.2
	Flexor Digitorum Profundus	3.4	MCP flex.	11.1
			MCP add.	6
			PIP flex.	7.9
			DIP flex.	4.1
	Extenxor Digitorum	1.9	MCP ext.	8.6
			MCP abd.	0.2
			PIP ext.	2.8
			DIP ext.	2.2
	Lumbrical II	0.2	MCP flex.	5
			MCP abd.	4.8
			PIP ext.	1.8
			DIP ext.	0.7
	Dorsal Interosseus II	2.5	MCP flex.	3.7
			MCP add.	6.1
			PIP ext.	2.6
			DIP ext.	1.6
	Dorsal Interosseus III	2.0	MCP flex.	3.7
			MCP add.	6.1
			PIP ext.	2.6
			DIP ext.	1.6
Ring [d]	Flexor Digitorum Superficialis	2.0	MCP flex.	11.9
			MCP add.	1.7
			PIP flex.	6.2
	Flexor Digitorum Profundus	3.0	MCP flex.	11.1
			MCP add.	6.0
			PIP flex.	7.9
			DIP flex.	4.1
	Extensor Digitorum	1.7	MCP ext.	8.6
			MCP abd.	0.2
			PIP ext.	2.8
			DIP ext.	2.2
	Lumbrical III	0.1	MCP flex.	5
			MCP abd.	4.8
			PIP ext.	1.8
			DIP ext.	0.7
	Palmar Interosseus II	1.2	MCP flex.	6.6
			MCP add.	5.8
			DIP ext.	2.6
			PIP ext.	1.6

Finger	Anatomical Name	Relative Tension	Joints/DOF	| Moment Arm |
Pinky [d]	Flexor Digitorum Superficialis	0.9	MCP flex.	11.9
			MCP add.	1.7
			PIP flex.	6.2
	Flexor Digitorum Profundus	2.8	MCP flex.	11.1
			MCP add.	6
			PIP flex.	7.9
			DIP flex.	4.1
	Extensor Digitorum	0.9	MCP ext.	8.6
			MCP abd.	0.2
			PIP ext.	2.8
			DIP ext.	2.2
	Extensor Digiti Minimi	1.0	MCP ext.	8.6
			PIP ext.	2.6
			DIP ext.	1.9
	Abductor Digiti Minimi	1.4	CMC opp.	6
			MCP abd.	4
			PIP ext.	2.5
			DIP ext.	2
	Flexor Digiti Minimi Brevis	0.4	CMC opp.	6
	Opponens Digiti Minimi	2.0	CMC opp.	6
	Lumbrical IV	0.1	MCP flex.	5
			MCP abd.	4.8
			PIP ext.	1.8
			DIP ext.	0.7
	Palmar Interosseus III	1.0	MCP flex.	6.6
			MCP add.	5.8
			DIP ext.	2.6
			PIP ext.	1.6
	Dorsal Interosseus IV	1.7	MCP flex.	3.7
			MCP abd.	6.1
			PIP ext.	2.6
			DIP ext.	1.6
Thumb [e]	Flexor Pollicis Longus	2.7	CMC abd.	0.2
			CMC flex.	14.3
			MCP add.	0.1
			MCP flex.	13.6
			IP flex.	8.7
	Extensor Pollicis Longus	1.3	CMC ext.	8.1
			CMC add.	9.5
			MCP ext.	8.5
			MCP add.	4.4
			IP ext.	4.1
	Abductor Pollicis Longus	3.1	CMC ext.	7.1
			CMC abd.	10.5
	Extensor Pollicis Brevis	0.8	CMC ext.	13.0
			CMC abd.	3.2
			MCP ext.	8.6
			MCP abd.	1.4
	Abductor Pollicis Brevis	1.1	CMC flex.	3.9
			CMC abd.	16.5
			MCP abd.	11.1
			MCP flex.	2.6
	Flexor Pollicis Brevis	1.3	CMC flex.	13.4
			CMC abd.	10.5
			MCP abd.	8.7
			MCP flex.	8.8
	Opponens Pollicis	1.9	CMC flex.	12.9
			CMC abd.	4.8
	Adductor Pollicis(t)	3.0	CMC flex.	36.9
			CMC add.	20.6
			MCP flex.	9.7
			MCP add.	6.0
	Adductor Pollicis(o)	3.0	CMC flex.	27
			CMC add.	17
			MCP flex.	8.2
			MCP add.	4.0

TABLE III

TENDON DATA

[a] Relative Tension data are collected from [64]

[b] Adductors or flexors have positive moment arms and abductors or extensors have negative moment arms.

[c] Index finger data are collected from [65]

[d] Middle, ring, and little finger data are collected from [66]

[e] Thumb data are collected from [67].