Finding Small, Versatile Sets of Human Grasps to Span Common Objects

Ian M. Bullock, Student Member, IEEE, Thomas Feix, Member, IEEE and Aaron M. Dollar, Member, IEEE

Abstract—Robotic and prosthetic hand designers are challenged to replicate as much functionality of the human hand as possible, while minimizing cost and any unnecessary complexity. Selecting which aspects of human hand function to emulate can be difficult, especially when little data is available on unstructured human manipulation behavior. The present work analyzes 19 hours of video with over 9000 grasp instances from two housekeepers and two machinists to find small sets of versatile human grasps. A novel grasp span metric is used to evaluate sets of grasps and pick an optimal grasp set which can effectively handle as many different objects as possible. The results show medium wrap and lateral pinch are both important, versatile grasps for basic object handling. The results suggest that three-fingertip precision grasps such as thumb-2 finger, tripod, or lateral tripod can be used to handle dexterous manipulation of a wide range of objects. The recommended grasp sets can help aid difficult design decisions for robotic and prosthetic hands, as well as suggesting important human hand functionality to restore during hand surgery or rehabilitate in an impaired hand.

I. INTRODUCTION

For robotic and prosthetic hand design, there are a number of reasons why the incredibly complex structure of actuators and sensors in the human hand cannot or should not be copied. Current engineered systems cannot replicate the full human capabilities in a similar packaging size, and with added complexity comes lower durability and increased cost. Furthermore, as evidenced by the widely used single DOF split hook [1], even simple, well-designed devices can have great utility. A number of simplified hands, such as underactuated hands, have been developed to leverage the benefits of lower complexity devices (e.g. [2–5]). For the design of prosthetic hands, simple designs are even more beneficial since many devices are limited to the space distal to the wrist, due to the variety of amputation points (e.g. [6], [7]), and there is a premium placed on light weight.

This work develops the concept of grasp span, a metric designed to assess the versatility of a set of grasps to handle a wide variety of objects. The concept of grasp span is then applied to 19 hours of grasp-object data from four subjects to select small sets of versatile grasps to emulate in a simplified prosthetic or robotic hand. We hope the resulting grasp sets and discussion will help designers to create effective hands that can pick up and manipulate a wide variety of objects.

In this paper, we will first discuss the grasp classification used and other related work. We will then describe the experimental methodology, as well as the procedure for computing grasp span. We then present the grasp-object matrix along with the grasp sets which maximize the grasp span metric for each profession and overall. We then discuss the characteristics and implications of these grasp sets and provide recommendations for a final set of grasps to emulate in a robotic or prosthetic hand. Finally, we discuss application of the results in robotic and medical domains, as well as addressing limitations and future work.

II. BACKGROUND AND RELATED WORK

Previous grasp studies have primarily focused on hand postures used for pre-selected objects, as opposed to recording unstructured human manipulation. An early study related to prosthetics [8] photographed 12 subjects to determine hand prehension shapes used in picking up 27 objects and the “hold-for-use” posture for 57 objects. Santello et al. asked subjects to imagine grasping fifty seven test objects while a motion capture system recorded 15 finger joint angles [9]. The first two principal components of the kinematic hand movements were shown to account for ~80% of the variance.

Cutkosky studied the grasps utilized by machinists using single-handed operations in working with metal parts and hand tools [10]. Kemp created a wearable system including a head-mounted camera and orientation sensors mounted on the body to learn body kinematics (not including the hand) and record manipulation tasks. A large amount of manipulation video was recorded but was never analyzed for details of grasp and object type [11]. While these previous efforts have helped better understand human grasp behavior, none have formally recorded and evaluated grasp type and frequency over a large time span of daily use.

Schlesinger et al. made the first major attempt to organize human grasping behavior into distinct categories: cylindrical, tip, hook, palmar, spherical, and lateral [12]. In 1956, Napier suggested a scheme that would divide grasps into power and precision grasps [13]. In studying the grasps required for manufacturing tasks, Cutkosky provided a much more comprehensive and detailed organization of human grasps [10]. For a more comprehensive review of grasp taxonomies, please see [14].

This work was supported in part by the National Science Foundation grant IIS-0953856.

I.M. Bullock, T. Feix, and A. M. Dollar are with the Department of Mechanical Engineering & Materials Science, Yale University, New Haven, CT, USA (e-mail: {ian.bullock, thomas.feix, aaron.dollar}@yale.edu).
In this paper we utilize a slightly extended version of the taxonomy presented by Feix [15]. This taxonomy is the most complete in existence, in the authors’ opinion, but lacks non-prehensile grasps, which we have added, specifically the platform push (“platform”) from Cutkosky’s taxonomy [10]. We kept the original Cutkosky naming for the grasps, such as using thumb-n finger instead of the Feix taxonomy’s “prismatic n finger.” Feix’s “adduction grip” is simply called “adduction.” The final taxonomy used in this paper is given in Fig. 1. All of Feix’s grasps were observed in the data set except for the distal type, a specialized grasp for scissors. Note that no grasps were seen that did not fall into this taxonomy, but some license was taken to fit some into it, such as with certain compliant objects.

Though there have been a number of efforts focused on classifying types of human grasps, the authors are not aware of any studies examining a large set of human grasping data in a real life setting, aside from a conference paper presenting initial results for our project [16]. The present work differs markedly from these previous works, in that it focuses on a specific type of grasp-object analysis, while these previous works only considered grasping data. A related study looks at the how frequently certain high-level manipulation tasks are used [17], but uses a very broad classification and a fairly small data set.

III. METHODS

Two machinists and two housekeepers participated in the study discussed in this paper. The first machinist (Machinist 1) is a 41 year old right-handed male with more than 20 years of professional machining experience, and the second machinist (Machinist 2) is a 50 year old right-handed male with about 30 years of experience. The first housekeeper (Housekeeper 1) is a 30 year old right-handed female with five years of housekeeping experience, and the second housekeeper (Housekeeper 2) is a 20 year old right-handed female.

![Modified Feix grasp taxonomy](image-url)
female with eight months of experience. None of the four subjects had any history of injury or any disability that would affect grasping and manipulation behavior. Machinist 2 did report prior shoulder injury due to repetitive overhead reaching, but did not experience any issues during the duration of the study.

The following enrollment criteria were used to screen potential subjects for the study: significant experience as professionals in their field, of normal physical ability, right-handed, able to participate to the extent to generate eight hours of useful data, and performing a wide variety of tasks representative of their profession during the span of their participation. For example, one machinist subject who almost entirely used a CNC lathe was excluded in favor of another subject who performed a wider variety of manual machining tasks on several different machines. Subjects were paid $10 per hour for participation on top of their normal salary.

A. Experimental Procedure and Apparatus

Full details of the experimental protocol can be found in [16], but a brief summary follows. Subjects meeting the enrollment criteria wore a head-mounted camera that recorded their hand use during normal work, for at least eight hours per subject. Fig. 2 shows an example image from the camera. One of two trained raters then tagged the right-handed grasps and objects in the video whenever the grasp changed, with inter-rater grasp tag agreement on an hour of test data giving Cohen’s $\kappa = 0.78$ [18]. For this study, rarely encountered objects with less than 40 instances were trimmed from the data set. This threshold was chosen to ensure a representative distribution of grasps is available for each object. Correlations between objects expected to have similar or dissimilar grasping patterns were used to confirm that the minimum number of instances for each object is sufficiently large, and $R > 0.9$ (Pearson’s $R$) was often observed for similar objects. After trimming, the data set contains 9748 grasp instances and 19 hours of data, with 59 unique objects.

B. Grasp Spanning Metrics

The overall process of selecting a small, versatile set of grasps using the relationships between grasps and objects in the dataset can be broken up into three different steps which will be discussed separately. The first part is to estimate how well an object is suited to a particular grasp based on the available data. The second part is to use these individual estimates to estimate the object handling capability of a set of grasps, using a grasp span metric. Finally, a versatile grasp set can be selected by applying the grasp span metric to all possible sets of grasps and choosing the set with the highest span score.

First, we will introduce the concept of a grasp-object matrix $M$. In this matrix, the rows correspond to different grasps, and the columns correspond to different objects. A heat map can be used to visualize the grasp-object matrix, as seen in Fig. 3. Each individual cell of the heat map shows how much data is present for a particular grasp being used for a specific object. A column illustrates the grasping pattern for a particular object. For example, the first column on the right shows that the towel was grasped more with precision disk and power sphere grasps, while the light tool grasp was hardly used at all for that object. Similarly, rows show the pattern of objects grasped by a particular grasp. The top row shows, for example, that the medium wrap grasp is used frequently with the vacuum and spray bottle objects, but rarely with the sponge and hex wrench objects.

Our goal is for each entry in the matrix, $m_{ij}$, to represent how suitable the $i^{th}$ grasp is for handling the $j^{th}$ object. This data set has durations and number of instances (counts) that each grasp is used with each object, which we can use to estimate how suitable a grasp is for an object.

In particular, we assume that if the subjects choose to use a particular grasp with a particular object for many instances or for a long duration, that grasp is well suited to handling that object. First, we calculate each element $c_{ij}$ by a sum of log duration and count measures:

$$c_{ij} = \frac{ldur_{ij}}{\text{mean}(ldur)} + \frac{lcount_{ij}}{\text{mean}(lcount)}$$  (1)

$ldur_{ij}$ and $lcount_{ij}$ are the log of the duration and log of the counts of each particular grasp-object combination: $ldur_{ij} = \log_2(1 + \text{dur}_{ij})$, and $lcount_{ij} = \log_2(1 + \text{count}_{ij})$. The logarithm is used to avoid overemphasizing certain frequent objects in the data set, such as cleaning tools used by the housekeeper. One could introduce weighting factors in that sum to change the importance of either duration or frequency of a grasp, but we currently choose not to. Now, we normalize by dividing by the sum of all the $c_{ij}$ to get our final grasp-object matrix:

$$M = [m_{ij}] = \left[ \frac{1}{\sum_p \sum_q c_{pq}} c_{ij} \right]$$  (2)

The dimension of $M$ is determined by the number of grasps and objects in the data. For our data, $M$ is a 33x59 matrix, since there are 33 grasps and 59 objects. The grasp-object matrix for both professions is found by taking the average of the normalized matrices for the machinists and housecleaners.

Having estimated the suitability of each grasp to each object, we can move on to assessing the versatility of a set of grasps. We first calculate the capability of the grasp set to
handle each object in the data set, to be discussed below. Then, the overall grasp span score is calculated as a simple sum of the handling capability for each object. Therefore, a grasp set with a high span score is a versatile set which is able to handle a wide variety of different objects.

The details of this calculation are as follows: First, for a given grasp set \( G = \{g_a, g_b, \ldots, g_n\} \), we create a new matrix \( M(G) \) containing only the rows corresponding to each grasp in the grasp set. For example, if we have \( G = \{g_1, g_3, g_4\} = \{\text{medium wrap}, \text{lateral pinch}, \text{tripod}\} \), with the top six objects shown (see upper right corner of Fig. 3):

\[
M(G) = \begin{bmatrix}
\ldots & 0 & 0.001 & 0.007 & 0.009 & 0.002 & 0.003 \\
\ldots & 0.003 & 0.002 & 0.0008 & 0.002 & 0.003 & 0.007 \\
\ldots & 0.001 & 0.005 & 0.002 & 0.002 & 0 & 0.004
\end{bmatrix}
\]

Note that the magnitude of each element is small, since it represents a small proportion of the entire matrix of 59 objects and 33 grasps. Next, we sort each column of the new matrix in descending order:

\[
M_{\text{sort}}(G) = \begin{bmatrix}
\ldots & 0.003 & 0.005 & 0.007 & 0.009 & 0.003 & 0.007 \\
\ldots & 0.001 & 0.002 & 0.002 & 0.002 & 0.002 & 0.004 \\
\ldots & 0.001 & 0.0008 & 0.002 & 0.002 & 0 & 0.003
\end{bmatrix}
\]

The matrix is then premultiplied by a vector of elements of a geometric sequence, \( S = [1 \ 1/2 \ 1/4 \ \ldots \ 1/2^N] \), where \( N \) is the number of rows in \( M_{\text{sort}}(G) \):

\[
SM_{\text{sort}}(G) = \begin{bmatrix}
\ldots & 0.003 & 0.006 & 0.008 & 0.01 & 0.004 & 0.009
\end{bmatrix}
\]

This geometric sequence scaling reflects the intuitive idea that each additional grasp that can handle a given object will add to the hand’s dexterity in handling that object, but with diminishing returns. This sequence weighting emphasizes the
ability to grasp a large set of different objects over being able to grasp the same object with multiple different grasp types.

Finally, this resulting vector of the handling scores for each object is summed to produce the final grasp span score:

$$\text{Grasp span}(G) = \sum S M_{\text{sort}}(G)$$

(3)

For the example grasp set, $\text{Grasp span}(G_{ex}) = 0.175$. Note that the grasp span is normalized by dividing by the span of the set of all grasps, so $\text{Grasp span}(G) \in [0,1]$. Thus, we want to find sets of grasps for which the span is as close to the maximum value of 1 as possible.

Once this grasp span score is defined as an estimate of the object handling capability and versatility of a set of grasps, we can find an optimal grasp set $G_k$ which maximizes $\text{Grasp span}(G_k)$ by testing all grasp set combinations for a given $N$ (number of grasps in the grasp set).

IV. RESULTS

A. Versatile Grasp Sets With Maximum Span Score

Grasp span scores for $N = 1$ grasp are presented in Fig. 4. This illustrates the versatility and overall capability of individual grasps. Fig. 5 shows the span scores for the best grasp sets, as optimized for the data from all subjects. For example, the dark blue bar for three grasps indicates that for the best combination of three grasps from all the possible grasp types gives a span score of just under 0.6. The span scores are compared with the average score of all possible grasp sets (shown in the light grey colored bar). These span scores show asymptotic behavior, with a much sharper increase in overall grasp span score when the number of grasps is small. This asymptotic behavior is expected from the grasp span, since the first few grasps ought to be most important in determining the overall functionality of a hand. Because of this effect, results only up to five grasps are presented below. While the choice of the first two or three grasps is quite important, after that it is possible to exchange a few different grasps and achieve a very similar span score.

The versatile grasp sets with maximum score for the housekeeper, machinist, and all subjects are shown in Fig. 6, 7, and 8, respectively. It should be noted that by the optimization method used (testing all combinations), it is possible for the larger sets of grasps to include grasps not present in the smaller sets, depending on the way that the object handling capabilities of the grasps complement one another. For this data set, this occurs only once in the top five grasps – for the five grasp machinist set, tripod and light tool are added while lateral tripod is removed.

B. Grasp Frequency

While the focus of the present work is on the grasp span method, it is worth also considering the overall prevalence of each grasp in the data set, as calculated by a mean of the duration proportion and frequency of each grasp. This data is presented for each subject type in Fig. 9. This figure helps to illustrate overall differences in grasping between the subjects which can help to explain the different spanning sets for each subject type. For example, medium wrap and precision disk are used much more by the housekeeper subjects, while tripod and lateral pinch are used more by the machinists. This grasp frequency data is analyzed in more detail in [16].
V. DISCUSSION

The grasp sets for each profession and for the combined data will now be discussed in detail, followed by a discussion of metric sensitivity.

A. Housekeeper Grasp Sets

The housekeeper data involves grasping of cleaning objects as well as many other household objects. There is more emphasis on power grasping and simple object transport motions in the housekeeper data than in the machinist data.

The grasp sets with maximum span score for the housekeeper data reflects these characteristics of the underlying data. Medium wrap is chosen first as a general purpose grasp which is especially well suited to larger, cylindrical objects. This is followed by power sphere, which the housekeeper subjects often use with soft, compliant objects. Lateral pinch is third, and is often used with locally flat objects or cords. Finally, index finger extension and precision disk are the fourth and fifth grasps.

B. Machinist Grasp Sets

The machinist data includes significantly more dexterous object handling and thus provides a nice complement to the housekeeper data. The machinist data focuses on manipulating small parts to be machined, handling various small tools such as calipers, files, and hex keys, and turning of various knobs on the machine tools used. While the machinist data may lack some household objects the dexterous handling capabilities should also apply to many everyday objects of similar geometry, and the housekeeper data can help to fill in this gap.

The grasp sets chosen for the machinist data reflect the higher prevalence of precision manipulation, with only two power grasps appearing in the five grasps. Lateral pinch is selected first and is commonly used with small knobs and flat objects. Next is thumb-2 finger, which is often used with
small parts and tools. Third is medium wrap, showing that this is an important power grasp for both professions. Fourth is lateral tripod, which is used for manipulation of small parts, tools, and knobs. For the five grasp set, lateral tripod is swapped out for tripod and light tool. tripod serves a similar function to lateral tripod, while light tool is used with various locally cylindrical objects such as wrenches, machine handles, and calipers. Overall, the machinist grasps appear to include lateral pinch and medium wrap for larger objects and less precise manipulation, whereas thumb-2 finger, lateral tripod, and tripod provide versatile object manipulation capability. Light tool provides extra power grasp capability, especially with certain tools that the machinist uses.

C. Grasp Sets for Both Professions Combined
The grasp sets chosen for the data of both professions combined follow well from the individual profession grasp sets. The top three grasps from each profession show up in the top four grasps for the combined data, and tripod is the final fifth grasp.

The first two grasps for the combined data are medium wrap and lateral pinch. These two grasps appeared consistently throughout the data analysis process, even when various variants of the ultimately chosen grasp span metric were used. Overall, these two grasps appear to be very useful, versatile grasps which seem to serve a complementary function. Medium wrap and lateral pinch are suited to different object types, and also grasp mainly along different axes relative to the hand. In particular, the main grasping axis for the medium wrap (radial-ulnar) is normal to the plane lateral pinch grasps can be easily performed in, where the thumb pad intersects with the lateral surface of the index finger.

Beyond the basic grasping capabilities of medium wrap and lateral tripod, various three-fingertip precision grasps can be used to enhance precision manipulation ability. The third grasp, thumb-2 finger, and the fifth grasp, tripod, fit this category. Looking at the top few grasp sets for each number of grasps (N) suggests that lateral tripod also has similar capability to thumb-2 finger and tripod.

Power sphere (grasp 4) and tripod (grasp 5) can be seen as adaptations of medium wrap and thumb-2 finger, respectively, with additional finger abduction. This suggests that adding abduction capability to a hand is another way to enhance its dexterity. Increased finger abduction in the power sphere may be especially useful when trying to enclose compliant objects of irregular geometry.

These grasp sets can then be compared with the raw grasp frequency data (see Fig. 9). Of the top five selected grasps, lateral pinch, medium wrap, and tripod do appear within the top five by frequency, though in a different ordering. The biggest difference is that precision disk, which was mostly associated with a single object, is considered much less important using the grasp span metric.

D. Metric Sensitivity
The specific choice of metric will affect the final grasp sets. Ultimately, log scaling was used to reflect the intuition that certain common cleaning objects should be deemphasized to improve the score for more versatile grasps, and the geometric series weighting reflects the diminishing returns for having multiple grasps to handle the same object. However, grasp sets were also calculated without log scaling and with other weightings, such as a maximum weighting which kept only the best grasp score for each object.

Omitting log scaling has the largest effect on the results. If log scaling is not used, medium wrap is still the top overall grasp, but grasps such as precision disk that are heavily associated with only one or two frequently used cleaning objects become highly scored, a result that is likely not generally valid.

Using a maximum weighting, where only the best grasp score for each object is kept, has less of an effect, but still produces some differences with the chosen geometric series weighting. Medium wrap and lateral pinch are still the top two grasps for the combined data, but they are followed by lateral tripod, precision disk, and light tool.

Overall, when testing other metrics and weightings, medium wrap and lateral pinch showed up quite consistently.
VI. CONCLUSIONS AND FUTURE WORK

The general approach of grasp-object span and the specific results obtained in this case can be applied both in the robotics domain and in medical domains. In the robotics domain, the grasp sets described above can be helpful in choosing a versatile grasp set to emulate in a robotic or prosthetic hand, especially with anthropomorphic designs. Specifically, the strongest result, most consistent across variations of the chosen metrics, is that both medium wrap and lateral pinch appear to be important, versatile grasps that can handle a wide range of objects. Precision manipulation ability can be effectively added through some sort of three fingertip precision grasp, such as the thumb-2 finger grasp. Adding finger abduction capability may help access another two useful grasps – power sphere and tripod.

In addition to picking a small set of grasps to emulate in a robotic or prosthetic hand design, these grasp sets have implications for the medical community. Our results suggest that a general purpose power grasp (e.g. medium wrap), a lateral pinch, and a precision fingertip grasp are all important for versatile object handling in the human hand. Thus, these results are an additional data point for making difficult decisions either about what hand function to restore in tendon transfer surgery or to rehabilitate in an impaired hand.

Various future work is possible and would address certain limitations of the present effort. While our present data set does provide information about object transport of household objects and various precision manipulation tasks, it could be further strengthened and made more general by adding other types of subjects. A slightly larger data set would allow an even wider range of objects to be considered. The existing data could also be used for various other types of analysis. For example, clustering and distance metrics based on the objects associated with each grasp could be used to understand functionally similar groups of grasps. Applying an object classification might provide a clearer view of the functionality each grasp provides. The idea of grasp span could be adapted to a similar concept of task span, where a set of grasps is selected and scored based on its involvement in effectively completing a wide range of tasks. This would help to address one final limitation of the present work – while our method should help ensure that some grasp is available to handle an object, it does not guarantee that the correct grasp will be available to perform a particular task with the object.

Despite certain limitations of the current work, we anticipate that these small, versatile grasp sets found by maximizing the grasp span metric will be useful to many researchers in the robotics and medical communities. We hope that the grasp sets will help in further understanding some of the essential functionality of the incredibly complex and dexterous human hand.

ACKNOWLEDGMENT

The authors thank Joshua Z. Zheng, Sara De La Rosa, and Charlotte Guertler for their assistance in gathering and tagging video data used in this study. The authors also thank Kayla Matheus for her help in creating the hand illustrations used in this paper.

REFERENCES