Search Procedures during Haptic Search in an Unstructured 3D Display

Alexandra Moringen[1], Robert Haschke[2] and Helge Ritter[3]

Abstract— **In this work, we focus on finding and characterizing stereotypical behavioral modes during haptic search in an unstructured 3D display. To this end, we introduce the notion of *search procedures* and present a machine learning approach for their identification. *Search procedures* are derived from the haptic exploratory procedures (EPs) through their parameterization and, in some cases, their mixture.**

In order to identify representative *search procedures*, we have evaluated data of eight individuals who performed a one-handed haptic search task in nine different scenarios, blindfolded. In all cases, they were asked to localize a specific target object within a tactile scene. Both the target object and the scene were formed from wooden bricks of various tactile shapes arranged in a configurable haptic display to exhibit a specific 3D tactile pattern.

By performing t-SNE-based dimensionality reduction and subsequent *k-means* clustering, we could identify three representative types of search procedures during manual haptic search: 1) one finger performs one EP 2) several fingers and the hand perform a differently parameterized EPs in a parallel manner 3) a mixture of EPs is performed by the hand and fingers in a parallel manner.

We exemplify the notion of search procedures and discuss the corresponding mixtures based on the results of the clustering.

I. Introduction

Haptic search is a process connecting at least three different mechanisms in a very efficient way: the movement and the search strategy of the hand and the arm, the haptic and kinematic activity and configuration of the fingers, and the haptic and kinematic perception. In our endeavor to understand this process we can build upon a large amount of previously conducted research in the above-mentioned three directions. Previously conducted modeling of the search behavior of bacteria, fish or foraging animals are commonly used for computational models and robot control (e.g. [12], [10]). Models such as Levy walk, Brownian motion, persistent search models and composite correlated random walk seem to yield optimal search strategies depending on the exact conditions of the search, i.e. the number, the distribution and type of the targets [8], [7]. Other complex movement strategies, like those of the marching crickets attempting to bite the ones in front and to run away from the approaching ones [1], have not been reused so far. Altogether, it is not yet clear, whether one of the existing models explains the human hand movement strategy during haptic search.

Several features, such as velocity and hand configuration, provide us with insights into some aspects of haptic exploration. According to Morash [9], hand configuration has an influence on the higher-level strategy of the haptic search.

[1,2,3] Neuroinformatics, CITEC, Bielefeld University, Germany {abarch,rhaschke,helge}@techfak.uni-bielefeld.de

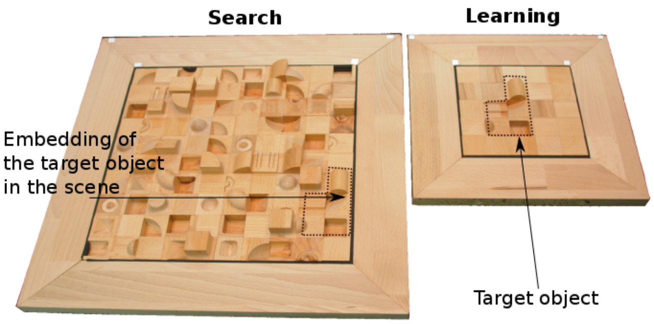

Fig. 1. Blindfolded participants are presented with two stationary wooden frames containing 10×10 (left) and 5×5 (right) differently shaped bricks. The task is to memorize the complex shape on the right-hand side and to find it again within the larger frame on the left-hand side, always using a single hand only.

In her recent work dedicated to analysis of trajectories in an unstructured display, Morash established a connection between the hand configuration and the choice of exploration strategy. She showed a relation between the detection radius, represented by a hand configuration, and the choice of strategy during haptic exploration: systematic vs. non-systematic.

On the lower-level, Wing et al. [16] suggest that during haptic exploration, velocity reflects a low level sensorimotor process. Based on the analysis of finger velocity, they showed that different modes can be observed even during haptic exploration of the simplest elliptic shapes with a single finger. A fundamental link between the kinematic and haptic characteristics of the finger and the haptic perception is provided in the work of Hayward [3]. It gives an account of the so called kinematic haptic perceptual invariants, the laws that are invoked in the context of haptic tasks. The author provides an example of a task during which one has to size a coin with a contact movement, and explains which invariants are employed to successfully conduct it [3].

The content of the present paper strongly builds on exploratory procedures, and can be roughly positioned between the higher-level movement strategies and the lower-level sensorimotor processes. Our main contributions are as follows:

- we introduce the notion of *search procedures*, typical realizations of EPs during haptic search and
- we show that during haptic search EPs are commonly performed together.

To clarify the first point, EPs are defined as movement patterns without a particular parameterization. Tanaka at el. [14] demonstrated that the parameterization is modulated by the type of haptic exploration task, discrimination or identification. This has motivated us to explore characteristic parameterization of EPs in a haptic search task, resulting in *search procedures*.

Here we show that during the search process different types of EP mixtures are used, i.e. different EPs are performed together, or same EP is performed by different fingers in a parallel manner. According to Klatzky et al. [5] this is done in order to simultaneously access different object properties. In the future text we skip the word "realization" and use *EP mixture* to describe the observed data, instead of *mixture of EP realizations* for simplicity.

Because in this work we focus on shape exploration and search, we assume that only three exploratory procedures associated with shape exploration can be observed in our experiment: contour following, enclosure and lateral motion [6]. An example of a typical mixture of EPs during haptic search/exploration is a mixture of EPs *enclosure* and *contour following*, which is characterized by an encompassing hand posture, haptic pressure and a different level of finger-activity per finger (see Figure 5a, posture 5). The other five search procedures will be discussed in detail in the Section III.

In Section II we describe the data acquisition and experimental design:

- A 3D display containing a great diversity of shapes, specifically designed to invoke a high number of search modes
- A feature extractor that generates a simple multimodal characteristic of the search procedures based on the highly dimensional recorded data
- A dimensionality reduction and clustering procedure that yields a preliminary representation of the *search procedures*.

Inspired by [9], [16], we aim to integrate both the dynamic and the shape-specific characteristics of the haptic search data, in a multimodal feature space. In this space, the dynamic component is represented by the velocity of the fingers and the hand; the static component is represented by the hand configuration.

The first, preparatory step of our exploratory data analysis serves the purpose of noise reduction and visualization of the feature space. In this step, we conduct a dimensionality reduction with Barnes-Hut SNE [15] from the feature space into a 2D space. In the second step, the identification of the *search procedures*, we perform simple k-means clustering on the resulting 2D values. Finally, by mapping back from the resulting 2D cluster centers into the feature space and into the raw data space, we exemplify the resulting search procedures with the help of the data in the corresponding spaces.

Similar to the work of Morash, we are targeting spontaneous and untrained haptic search behavior. However, in this work we do not investigate trajectory pieces, but focus on analysis of single data-points. Our experimental setup and the acquired data can be also used for investigation of both the lower- and the higher-level aspects. Therefore, the long-term challenge of our future work is to link the search procedures described in this work with the higher-level strategy on the one hand and the lower-level haptic perceptual invariants on the other.

II. METHOD

A. Participants and Apparatus

Thirteen right-handed sighted individuals aged 22 to 30 participated in the study. The participants were all students of the Bielefeld University, apart from one. The protocol was approved by the Bielefeld University Ethics Committee, and informed consent was obtained from all participants prior to their participation. None of the participants had any prior knowledge of the experimental design or the stimuli. An individual recording session took between two and three hours. During this time a subject conducted search in ten experimental conditions: a training phase, followed by nine iterations used for the evaluation. The training phase has been necessary for the blindfolded participants to learn about the previously unknown positioning of the boards and their size.

The collection of available tactile stimuli consisted of 360 rigid wooden bricks, 3 × 3 cm in size, containing 55 distinct types of primitive shapes. The collection of primitive shapes also contained blank wooden bricks. Subtle and extensive shape dissimilarities between the bricks were chosen to evoke different types of exploration strategies, particularly different types of the finger-specific exploration.

Two wooden frames (see Figure 1), containing and fixating the rigid bricks within the display were designed to hold 25 and 100 bricks, respectively. The main purpose of this experimental design was to produce rigid and stationary previously unknown shape patterns for both the learning and the search part of the task. The combination of basic building blocks allowed to create an arbitrary complex shape.

For recording of the haptic interaction forces, the subjects were wearing a thin elastic glove with integrated fabric-based flexible sensors[1] [2]. However, in this work we don't show evaluation of this data. We use it only to manually examine the binary presence/absence of contact for a given search procedure.

The position of the hand as well as the position of the boards has been tracked by 13 Vicon cameras [4] at the rate of 200 Hz. Two Vicon markers have been positioned on each finger to enable a rough hand posture identification. The hand pose (position and orientation) has been tracked with extra three Vicon markers positioned on the back of the hand (see Figure 2c). Additionally, a Basler camera has been recording the video logs of the experiment, top-shot perspective, a microphone has been used to capture the feedback of the participants during the experiment as well as the point of time in the search task at which the target object has been located by the study participant. A snapshot of the experimental setup can be seen in Figure 2b. A video

[1]It has not been tested yet, what effect the glove had on the search and exploration strategy. Nevertheless, we have decided to use it in our experiment due to anticipated advantages of the information gain.

Fig. 2. (a): The nine employed search stimuli differing by the number, shape and positioning of the non-flat bricks. (b): A student wearing the haptic glove during the setup test: 13 Vicon cameras. (c): 13 blue points schematically visualize the placement of Vicon markers during the recording of the hand trajectories. Only the markers on the fingertips are used for the finger-relevant feature extraction. The second row of marker is recorded for technical reasons, e.g. in order to improve the quality of the 3D data generation by the Vicon system.

of one experimental run is available under the link below[2].

B. Task and Procedure

The small board on the right of the participant contained a target pattern that the subject was asked to learn in the first part of the experiment (see the right part of Figure 1). The large board on the left side of the subject (see the left part of Figure 1) contained the target pattern in the same orientation integrated in an array of distractor shapes. Each trial has been characterized by a different search target and a different shape density of the search field (50%, 60%,75%, 90% and 100% of non-flat bricks). Target patterns of different complexity are displayed in Figure 2a.

Both displays, the target object display and the search field, are a strong simplification of a general 3D complex shape and a search environment in a 3D space. To focus the experiment on the hand and the finger dynamics, rather than the arm control, we have restricted the 3D shape to a plane. Thus, we neglected the effects on the hand posture in such scenarios, as in-hand object manipulation, or manipulation of objects in other orientations in the 3D space.

C. Data Postprocessing and Feature Extraction

Preliminary postprocessing of the Vicon data (labeling, gap filling) has been very time-consuming due to a high temporal resolution, a high number of markers, self-occlusion, and a comparatively high resolution of markers on the fingertips. From thirteen recorded subjects, eight have been completely postprocessed and used for the evaluation in this work. Previous to the calculation of the features and subsampling, we have applied Savitzky-Golay filter for smoothing with window size $n = 51$ and order of the polynomial $d = 2$. After the calculation of features, we have conducted subject-specific mean-value normalization.

The calculation of the features is based on the 3D marker trajectories as depicted in Figure 2c. The feature set $\Phi \in \mathbb{R}^{13}_{\geq 0}$ is defined as follows:

[2]Video of one experimental run `http://www.techfak.uni-bielefeld.de/~abarch/s2t7.avi`.

$$\Phi := \{(d_{1,2}, d_{2,3}, d_{3,4}, d_{4,5}, d_{5,1}, v_{f,1}, \ldots, v_{f,5}, v_{h,1}, v_{h,2}, v_{h,3}) \mid d_{i,j}, v_{f,i}, v_{h,k} \in \mathbb{R}_{\geq 0}\}, \quad (1)$$

where

1) $d_{i,j}$ represents the **distance between the adjacent fingertips** i and j, e.g. the distance between the thumb and the index finger is denoted by $d_{1,2}$, the distance between the index finger and the middle finger is denoted by $d_{2,3}$
2) $v_{f,i}$ represents the Euclidean length of the **velocity of the fingertip** $i \in \{1, \ldots, 5\}$ w.r.t. the hand coordinate system. Hence, the velocity of the fingers does not contain the component resulting from the rotation or translation of the hand in 3D space.
3) $v_{h,i}$ with $i \in \{1, 2, 3\}$, corresponding to the three hand markers, represent the Euclidean length of the **velocity of the hand**. We have used three values in order to capture the dynamics of the hand in the 3D space.

For the preliminary evaluation conducted in this work we have chosen the smallest representative set of features to describe the dynamics and the shape of the hand and the fingers. The feature vector can be extended by e.g. including the distances from the fingertips to the back of the hand, and, therefore, resulting in an approximate representation of the volume of the hand configuration.

Figure 3 (a,b) visualizes the complex intermodal dependencies in the feature space based on the modality-specific averages before the subject-specific mean-value normalization. The left plot is a plot of the averages of finger distances $\overline{d} = \sum d_{i,j}/5$ against the averages of the finger velocities $\overline{v_f} = \sum v_{f,i}/5$ forming a triangle with the top peak lying close to the mean value of the distribution of $\overline{d}$s. The middle plot displays a V-shaped dependency between the averages of the finger velocities $\overline{v_f}$ and the averages of the hand velocities $\overline{v_h} = \sum v_{h,j}/3$.

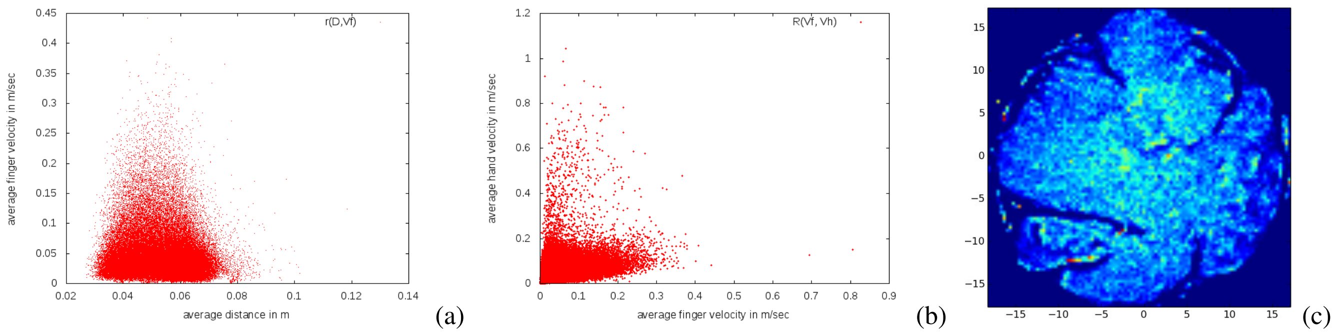

Fig. 3. (a): Relation between the average finger distances and the average finger velocities. (b): Relation between the average finger velocities and the average hand velocities. Number of samples $n = 350000$. (c): Heat map of the haptic exploration data in the SNE space. Number of samples $n = 350000$, $\theta = 0.3$ and perplexity $p = 30$.

Even in this simplified form, the feature space exhibits a complicated structure (Figure 3 (a,b)). This fact motivated us to use clustering versus linear regression as an approach to modeling.

D. Dimensionality reduction with Barnes-Hut SNE

In order to roughly estimate the structure of the multi-dimensional feature space primary to clustering, to denoise the data and to visualize it, we have used the Barnes-Hut SNE [15]. t-SNE is a dimensionality reduction method that aims to preserve structural similarity of data in the projection space. However, due to its computational complexity, it can not be used with large data sets like ours. With the help of an approximation during the calculation of the similarity matrix, Barnes-Hut SNE allows to apply t-SNE to a large number of samples [15]. Figure 3c presents an example of the converted feature set, which shows approximately six partially separated regions. For our calculations we have used the software available on Github [13].

E. Clustering with k-means

In order to obtain a representative set of search procedures, in this step we performed a *k-means* clustering on the 2D data set Φ_{sne}. Φ_{sne} results from the mapping of the feature set Φ (see Definition 1) into the 2D space with the Barnes-Hut SNE as described in the previous Section II-D. For the clustering we have used every 20th sample of the data set Φ_{sne}. Based on the empirical tests, the resulting data frequency of 10 Hz seemed to yield a good tradeoff between efficiency and the quality of movement representation. The choice of the number of cluster centers $k = 6$ has been motivated by the data structure also discussed in the previous section. However, it is only a preliminary value. In order to find an appropriate value, a deeper investigation of the data, the feature space and the 2D t-SNE space has to be conducted. This question is not the focus of this work.

In this work the resulting cluster centers represent the search procedures. In order to find a close representation for each search procedure in the recorded data we conduct the following two simple steps. Firstly, for each cluster center we look for the closest point in Φ_{sne} data set based on the Euclidean distance. Secondly, in order to find its correspondence in the target space (i.e. trajectory space, feature space), we simply look for a point in the target space with the same meta data description. In our work timestamps, trial number, subject number serve as such meta data. This way we are able to obtain a representation for each search procedure in the space of the original data and in the feature space.

III. Results and Discussion

The goal of this section is to characterize search procedures based on k-means clustering. Due to the design of the k-means clustering algorithm the cluster centers are highly suitable for this purpose. Firstly, the cluster centers should be pairwise dissimilar. Secondly, each cluster center should be representative of the corresponding data subset. Therefore, we limit the final outcome of this work to the set of cluster centers illustrating representative states of the corresponding search procedures.

Figure 4 provides a quantitative overview of the cluster centers and Figure 5a provides a qualitative overview of the cluster centers. In Figure 4, each sixfold bin is used for visualization of a specific feature across all six clusters. The first five bins correspond to the first five dimensions of the feature vector, the finger distances $(d_{1,2}, d_{2,3}, d_{3,4}, d_{4,5}, d_{5,1})$. The next five bins correspond to the fingertip velocities $(v_{f,1}, \ldots, v_{f,5})$. The last three bins correspond to the three velocities of the hand $(v_{h,1}, v_{h,2}, v_{h,3})$.

Cluster $C4$ (see Figure 4) is characterized by the lowest velocity of the hand and all fingers, apart from the index finger, as well as the highest distance between fingers. Based on the inspection of the haptic data (see Figure 5b-bottom), we could verify that only one finger is taking part in the haptic exploration. However, one can observe a certain level of velocity in the adjacent fingers. The velocity of the adjacent fingers seems to be slightly lower and decreasing with the distance to the active finger. This effect can be commonly observed in the data and, in particular, in all following examples of the cluster centers. We further refer to this kind of velocity profile as the *center* of finger activity.

The type of a search procedure represented by the $C4$ (Figure 5a, bottom row, left) is used for the exploration of small-scale local shape structures like edges, holes, etc.

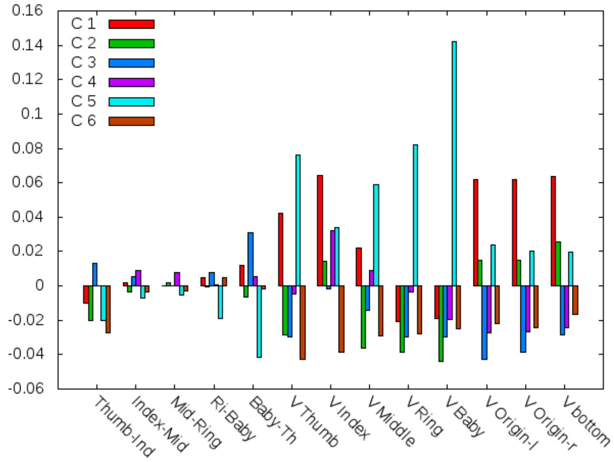

Fig. 4. Six colors correspond to six mean-value normalized feature vectors representing the cluster centers. Mean-value normalization is subject-specific.

Therefore, we assume that this is a realization of the **contour following** EP, which, in contrast to all other search procedures exemplified in this work, is not executed in a mixture with any other EP.

Cluster $C1$ (Figure 5a, top row, left) is characterized by the highest speed of the hand. In this example a high level of hand velocity is not compatible with a high finger velocity. Same as in the previous cluster, the velocity profile is *centered* around one finger, and is declining in the adjacent fingers with the distance.

Based on the characteristics of the hand-posture, the presence of the haptic feed-back along with a high velocity of the index finger and the thumb, we assume that this search procedure is a mixture of two EPs: **enclosure** and **contour following**. Haptic and kinematic activity of both fingers lets us assume that in this search procedure the contour following is conducted with two fingers in a parallel manner.

Contrary to the previous clusters, cluster $C2$ is characterized by low finger velocity (apart from the index finger) and a slightly above average hand velocity. In this cluster the index, the middle and the ring fingers are in contact with the surface (see Figure 5b - top). Based on the above-described characteristics, in particular the activity of several fingers, we assume that this search procedure realizes **contour following** with several fingers in a parallel manner.

Cluster $C3$ (Figure 5a) is characterized by low velocity in all components apart from the index finger, which exhibits almost an average velocity. As expected, all fingers, apart from the thumb are in contact with the surface (see Figure 5b - middle). This search procedure may be a mixture of **lateral motion** and **contour following** EPs. Based on similar considerations of velocity and haptic feedback as in clusters $C1$ and $C2$, we assume that in cluster $C3$ the contour following is performed by several fingers.

Cluster $C5$ is characterized by an extremely high velocity in all five fingers, accompanied by the smallest distances between the fingers and by a relatively high velocity of the hand. In the example of Figure 5a before normalization its values range from 0.09 m/sec to 0.17 m/sec. The velocity profile suggests that the exploration is *centered* around the fifth finger. Analogously to the previous cluster centers and based on the feature characteristics, we assume the cluster center to represent a mixture of **enclosure** and **contour following** EPs.

In comparison to cluster $C5$, cluster $C6$ is characterized by an opposite velocity profile and the lowest velocity of the index finger and the thumb among all cluster centers. In the example in Figure 5 the center of activity is the middle finger. We assume that $C6$ is a mixture of **enclose** and a very slow **contour following** in a detailed local search.

IV. Summary and Outlook

This work shows our first approach to identification of common behavioral modes in a haptic search of an unstructured 3D display.

The method is based on unsupervised learning of features typically used for representation of a manual exploration process. Therefore, feature extraction has not required any deep expert knowledge. The presented method axiomatically relies on the fact that only three EPs can be used during shape exploration. This results in a relatively simple access to a complex data. The main findings consists of representative search modes, named search procedures, and their representations as different types of mixtures of one of two exploratory procedures.

Our approach to identification of search procedures consisted, firstly, of a particular experimental design and a freely configurable haptic display. This allowed us to form an extensive set of previously unknown complex shapes. With the help of a large number of simple shape types, we have hoped to encompass a broad set of search strategies. Even in this simplified form, such a modular design provides us with rich opportunities for the future investigation of the haptic search process. For example, it is possible to approach the following questions: How does the exploration during the learning phase differ from the exploration during the search? How does the search strategy depend on the shape of the target object, the content of the haptic search field? What happens when the orientation of the target object in the search field is different to the one previously learned? What difference does it make, when the parts of the search target in the search field are presented in a permuted order?

In order to obtained a comprehensive description of the search and exploration process, a detailed kinematic and haptic capture of the hand with 13 Vicon markers and 54 haptic sensors of the haptic glove has been conducted.

Secondly, we have been looking for a possibility not only to analyze the dynamics of the hand, but to connect it with its configuration. This resulted in a multimodal feature space consisting of both the velocities of the hand and the hand posture. In the next step we plan to integrate the haptic profile in the feature vector.

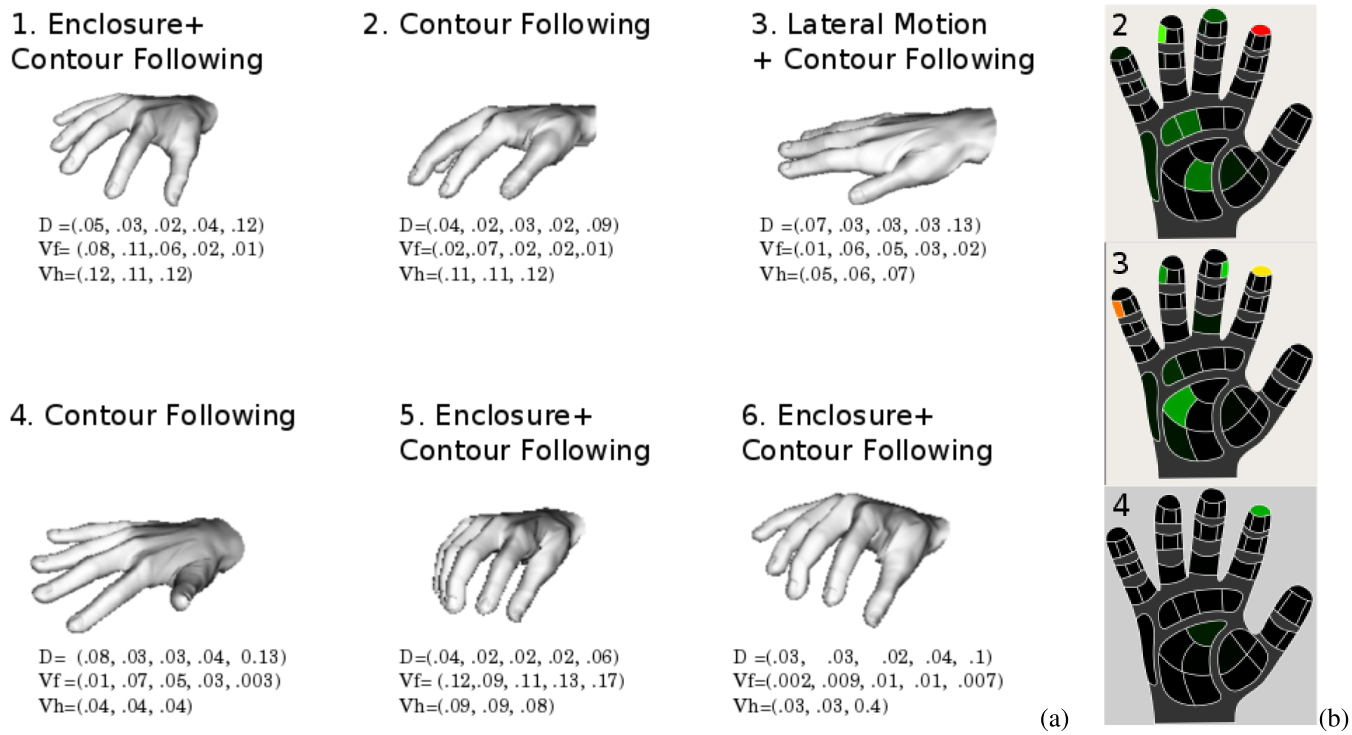

Fig. 5. (a): An approximate visualization of the cluster centers based on the procedure described in [11]. The values of D in meters, V_f and V_h in m/sec are given without the mean-value normalization. (b): Haptic data corresponding to the cluster centers 2, 3 and 4 resp. The color (from green to red) illustrates calibrated value of the corresponding haptic sensor.

Finally, after the dimensionality reduction and the preliminary partitioning of the 2D data with k-means, we have received a set of cluster centers representing the *search procedures*. Based on the properties of the cluster centers we could infer the following representative types of search behaviour:

1) single EP realization (cluster $C4$)
2) mixtures of different EPs (e.g. cluster $C1$)
3) mixture of same EP parameterized in a different way during a parallel execution with multiple fingers (e.g. cluster $C2$).

The current clustering with k-means is not optimal and does not well reflect the structure of the 2D data. An independent investigation needs to be conducted about an appropriate choice of k. In the future work we would like to further explore the structure of the 2D data w.r.t. our mixture assumption and further continue to evaluate the dynamic structure of the search process.

V. Acknowledgments

Big thanks to Kathrin Krieger who put lots of work into the organization and the data recording. Thanks to Alex Schulz for his advice on t-SNE. We would also like to thank Roberta Klatzky for fruitful discussions that preceded this project.

References

[1] S. Bazazi, C. Ioannou, S Simpson, G. Sword, C. Torney, P. Lorch, and I. Couzin. The social context of cannibalism in migratory bands of the mormon cricket. *Plos One*, 2010.

[2] G. Buscher, R. Koiva, C. Schurmann, R. Haschke, and H.J. Ritter. Tactile dataglove with fabric-based sensors. In *International Conference on Humanoid Robots*, 2012.

[3] M. Grunwald, editor. *Human Haptic Perception*, chapter Haptic Shape Cues, Invariants, Priors, and Interface Design. Birkhaeuser Verlag, 2008.

[4] Information available at: http://www.vicon.com/.

[5] R. L. Klatzky, S. J. Lederman, and C. L. Reed. There's more to touch than meets the eye: the salience of object attributes for hpatics with and without vision. *Journal of Experimental Psychology*, 1987.

[6] S. Lederman and Klatzky R. Hand movements: a window into haptic object recognition. *Cognitive Psychology*, 1987.

[7] Mark A. Lewis, Philip Maini, and Sergei V Petrovskii, editors. *Dispersal, Individual Movement and Spatial Ecology*. Springer-Verlag Berlin Heidelberg, 2013.

[8] V. Mendez, D. Campos, and Bartumeus F. *Stochastic Foundations in Movement Ecology*. Springer Berlin Heidelberg, 2014.

[9] Valerie S. Morash. Detection radius modulates systematic strategies in unstructured haptic search. In *World Haptics*, 2015.

[10] S. Nurzaman, Y. Matsumoto, Y. Nakamura, K. Shirai, S. Koizumi, and H. Ishiguro. From lévy to brownian: A computational model based on biological fluctuation. *Plos One*, 2011.

[11] Matthias Schroeder. *Articulation estimation and real-time tracking of human hand motions*. PhD thesis, Bielefeld University, 2015.

[12] D. Sutantyo, P. Levi, C. Moslinger, and M. Read. Collective-adaptive lévy flight for underwater multi-robot exploration. In *Mechatronics and Automation (ICMA)*, 2013.

[13] https://github.com/lvdmaaten/bhtsne.

[14] Y. Tanaka, Bergmann Tiest W. M., A. Kappers, and A. Sano. Contact force and scanning velocity during active roughness perception. *Plos One*, 2014.

[15] Laurens van der Maaten. Barnes-hut-sne. *CoRR*, 2013.

[16] Alan Wing, Chris Christou, and Adam Waller. Touching trajectories: the relation between speed and curvature in exploring shape. In *Euro Haptics*, 2001.