Motion Primitive Recognition in Tooth Removal Surgery

B. (Bram) Bornhijm, BSc. (4692721)
Master thesis: Cognitive Robotics (3ME), Delft University of Technology
Supervised by: Dr. Ing. J. Kober & Drs. T.C.T van Riet
Period: November 2021 - December 2022

***Abstract*— Whilst the extraction of teeth (exodontia) remains one of the oldest and most performed surgeries on earth, very little is understood about the procedure itself. Especially in the area of the required movements, torques and forces to remove specific teeth and how these interact with existing tissue. This knowledge gap has been hypothesized to contribute to an increasing referral rate towards Oral and Maxillofacial Surgery (OMFS) practices in the Netherlands for simple extractions due to low confidence in young dentists. The objective of this project is to apply techniques used in imitation-learning in the field of robotics to deconstruct complex movement into smaller fundamental building blocks called movement primitives (MPs) to deepen the understanding of exodontia. To achieve this an existing dataset was used consisting of high resolution force-, torque-, position- and rotation-data of extractions. This dataset was collected using a measurement setup consisting of fresh frozen cadaver material, a force torque sensor and a robotic arm in gravity compensation mode. In this paper a novel iterative two staged method for identifying movement primitives is introduced and employed to extract movement primitive information from these extractions in the dataset.**

I. Introduction

One of the most performed surgeries on earth, tooth extraction [1] (or exodontia) is seemingly not well understood in terms of the reaction forces and torques present in the extraction process, as well as for the exact pattern of movement required to remove a tooth. Most research seems focused rather on the maximum force necessary to remove the tooth [2], [3]. This leads to a knowledge gap as to the fundamental properties of tooth removal.

At this moment, most training is limited to textual instructions on motion patterns to apply to specific teeth due to this knowledge gap [4]. Therefore, the quality of education is also limited and has remained mostly unchanged in the last centuries. Besides this, due to the increase in standards of preventive care in most western countries, also little clinical training is possible for dental students [5], [6].

Hanson et al. showed that practice on cadaver material is closest to the clinical tooth extraction procedure [5]. However, as this is not a resource that is sufficiently available to most dental schools combined with the ethical concerns of processing large amounts of cadaver material, most education consists of gaining experience from practicing directly on patients. Plastic models are rarely used because of their ill representation of clinical extractions [5], [6].

Research from 56 different dental schools across Europe shows supporting evidence that dental students suffer from low confidence to perform these procedures. Parallel to this an increase in referrals to Oral and Maxillofacial Surgery (OMFS) practices for simple tooth extraction procedures in recent years is observed [7], which is both more expensive as well as a greater strain on care resources [6].

This project aims to improve the understanding of exodontia by analyzing movement (patterns) present during this procedure by extracting motion primitives. It is theorized that human movement is comprised of multiple smaller basic movements, or motion primitives, which are combined to form complex movements [8]. By analyzing these basic building blocks of complex motions makes that they can be used to deepen our understanding of said complex motions and make it much easier to teach these to a robot. This method has been extensively researched and used successfully in the field of robotics for modelling and synthesizing complex movement by breaking down said complex movements into multiple, smaller fundamental building blocks [9], [10], [11].

It is hypothesized that if these motion-primitives can be used to teach a robot to complete a complex task / movement, theoretically these could then also be used to analyze complex movement to better understand this complex movement.

The objective of this research project is hence to improve the understanding of tooth removal through motion-primitive-based methods. This research could be used to on one side to improve the quality of education, secondly to improve the quality of tooth extractions, and thirdly make it possible for better methods of treatment to be developed.

To achieve this, this project will employ methods from the field of imitation learning, more specifically the field of movement primitive based learning, in the field of exodontia using a dataset comprising movement data of clinically representative extractions.

The main research question of this research is hence:

Can motion-primitive based learning be used to deepen the understanding of tooth extraction procedures?

A. Background on Exodontia

This section will discuss a basic background on dental anatomy and exodontia. Section I-A.1 will discuss the general anatomy of the human denture, specifically that of a canine and Section I-A.2 discusses exodontia and its associated challenges.

1) Dental Anatomy: The typical adult human dentition consists of 32 teeth arranged as shown in Figure 1. This contains four incisors (blue), two canines (orange), four

premolars (green), and six molars (red) (those placed furthest at the rear being wisdom teeth) in each jaw [12]. Each tooth is often clinically referred to using its coordinates, as designed by the Fédération Dentaire Internationale (FDI). In this system, the denture is split into four quadrants, each holding eight teeth. Each coordinate is comprised of two single integers, where the first integer refers to the quadrant and the second to the location of the specific tooth inside that quadrant, the origin is fixed at the incisors. As an example, "11" refers to the incisor in the upper right and "33" refers to the canine in the lower left (from the perspective of the patient) [4].

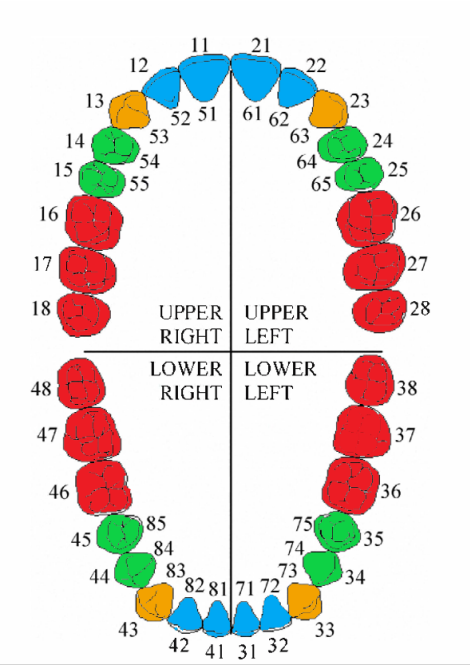

Fig. 1. Dental anatomy of upper jaw (top of Figure) and lower jaw (bottom of Figure). Here, the incisors are marked in blue, canines are marked orange, premolars are marked in green and molars are marked in red. Adopted and modified from [13].

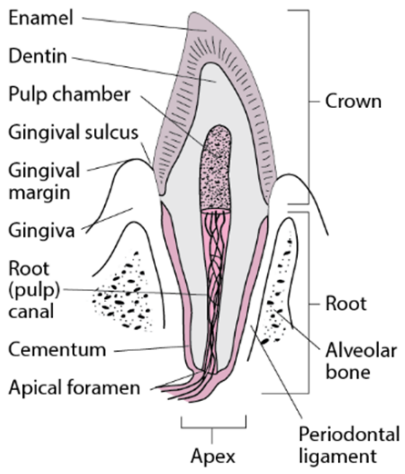

Fig. 2. Cross section of a canine tooth. Adopted from [12]

Each tooth consists of a crown (top, exposed section) and a root (area that, through a layer of cementum, is fixed to the jaw), the number of and size of the root(s) differ per type of tooth, for example, incisors and canines have a singular root whereas premolars and molars have multiple [12]. Figure 2 shows a cross section of a canine tooth. Here it can be seen that the tooth is fixated to the jaw using a layer of cementum, hard connective tissue covering the root, connecting the root to the alveolar bone through a layer of periodontal ligament [14].

2) Tooth Removal Procedures: During exodontia the objective is to weaken and break the bonds between the roots and the periodontal ligament so that the tooth can be removed. This is typically achieved by clamping the tooth in forceps, an example of such can be seen in Figure 3, while manually applying a movement and force pattern [4]. Frequently, also an elevator is used to start moving the tooth, however as the used dataset only contains data from using forceps, as is described more in detail in Section I-B, this project focuses solely on the use of forceps.

Fig. 3. Example of forceps used in exodontia. Adopted from [15].

The most commonly used techniques for removing a tooth during exodontia using forceps is to apply a combination of rocking and twisting motions to the tooth. The used technique depends on the morphology, number and shape of the roots the specific tooth has. For example, single-rooted teeth can be twisted around their vertical axis whilst multi-root teeth like molars require a rocking motion as their rotational asymmetry makes twisting impossible [3], [4].

By forcefully moving the tooth in its respective cavity, it is believed that this space is enlarged, creating space for the tooth to be removed whilst at the same time destroying connective tissues that holds the tooth in place, allowing for its removal [4]. The recommended pattern, as commonly taught to dental students, per tooth can be seen in Table I [4], [16] [1].

[1]Referring to the table: Lingual means towards the tongue in the lower jaw, Palatinal means towards the palatinum in the upper jaw. In practice these refer to the same direction, namely opposite to buccal which is towards the cheek.

TABLE I

EXTRACTION PATTERN PER TOOTH. THIS TABLE IS ADAPTED FROM TABLE 2-1 IN [16]

Main movement	Tooth	Root shape
Twisting	11,21 24,25,44,45	Round singular shape Sometimes dual roots
Rocking	12,22	Oval shaped
Rocking and Twisting	13,23 15,25	Long oval shaped Flat shaped
Rocking [Lingual /Palatinal]	31,32,33,41,42,43 14,24 16,17,26,27	Long oval root Two thin roots Three roots
Rocking [Mostly Lingual]	36,37,46,47	Two flat shaped roots

The complexity of the procedure can vary depending on the number of roots, clinical conditions such as restorative state of a tooth or periodontal health, and the position in the jaw [4]. Typical challenges associated with exodontia are the limited access available to the dentist performing the operation (keyhole surgery) and that little fixation is possible during the procedure, especially for the lower jaw.

B. Dataset

This section will discuss the details of the dataset that is used in this project. Section I-B.1 will discuss the method of collecting the dataset, Section I-B.2 will discuss the used method for transforming the collected data and how this data has been preprocessed, lastly Section I-B.3 will discuss the challenges and limitations associated with using this dataset.

1) Dataset Collection: For analysis of tooth extraction, ideally, an extensive dataset of in-vivo exodontia would be collected. Such a dataset would need to contain high resolution data of both the applied forces, moments, and movements during the procedure.

However, this poses a set of unique challenges, one of which being little space and limited line of sight for any measuring equipment, also little fixation to measure the forces is possible, as both the used forceps and the tooth itself cannot be fixed directly [16].

For this reason, an alternative approach was chosen. As shown by Hanson et al., cadaver material is closest to its clinical equivalent [8] but does allow for fixation. The approach that was chosen was one where fresh frozen cadaver material, specifically the lower and upper jaw would be used to perform the extractions. The jaw would be fixated to a force-torque sensor measuring the applied moments and forces. The surgeon would use forceps attached to a Kuka LBR iiwa7 R800 robotic arm in gravity compensation mode to measure the movements at a frequency of 100 Hz [16]. The complete setup with associated labels can be seen in Figure 4.

To collect the raw dataset, 3 experienced oral and maxillofacial surgeons performed 116 successful and, in their experience, representative extractions out of a total of 127 extractions. A student also performed 15 successful extractions out of a total of 22. Hence data on a total of 132 successful and representative extractions have been collected, 61 of which are from the lower jaw, and 71 from the upper jaw [16]. However, for the final analysis only the successful extractions performed by experienced surgeons will be analyzed, this limits the size of the dataset to a total 110 extractions.

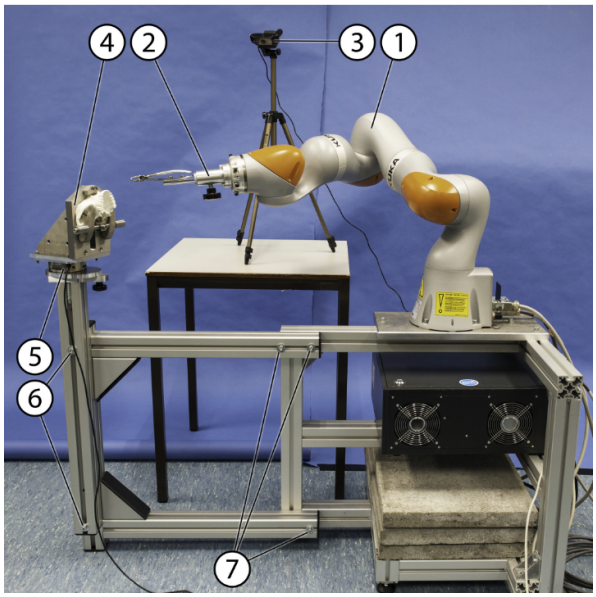

Fig. 4. Overview of measurement setup, (1) robot arm, (2) forceps holding device, (3) video camera, (4) upper jaw holding device, (5) force-torque sensor, (6) bolts for vertical adjustment, (7) bolts for horizontal adjustment. Adopted from Figure 1 in [17].

2) Transformation: All performed extractions yield a seven dimensional time series for position and orientation (position and quaternion), in addition, these also each yield a six dimensional force/torque timeseries. This data was then transformed into the coordinate frame of the respective tooth for analysis, a visualization of this coordinate system can be seen in Figure 5.

Fig. 5. Local coordinate system for the 32 (lower left lateral incisor).

The dataset has further been preprocessed by first eliminating the failed extractions and failed measurements, secondly the remaining extractions have been resampled to obtain a constant sample rate. The resampled dataset was then passed through a low pass butterworth filter with a cutoff frequency of 1 Hz to filter out high frequency noise from the measurement, this cutoff frequency was selected based on an expectation of the maximum speed of movements during exodontia. Finally the data was smoothed using a Savitzky-Golayfilter. A comparison of the original and preprocessed data can be seen in Figure 6 and 7.

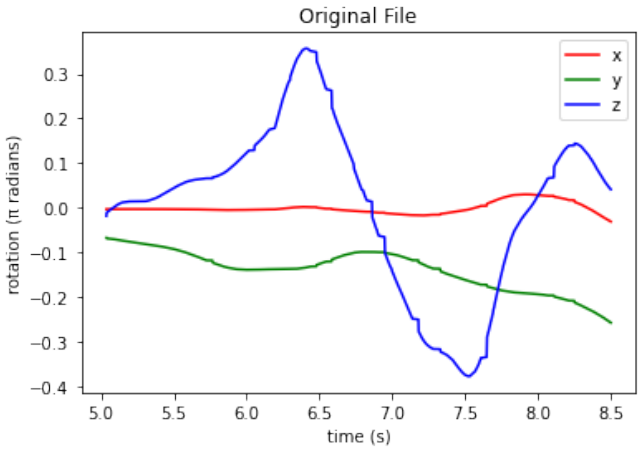

Fig. 6. Sample extraction from original dataset. The horizontal axis shows time in seconds and the vertical axis shows rotation in rad/s. Each line represents a respective dimension being either x, y, or z.

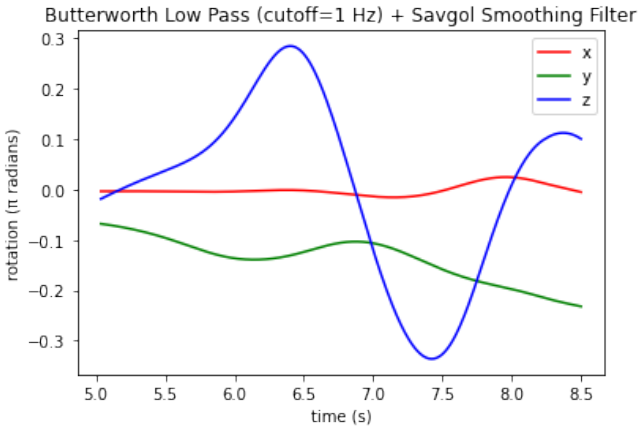

Fig. 7. Sample fragment from preprocessed dataset. The horizontal axis shows time in seconds and the vertical axis shows rotation in rad/s. Each line represents a respective dimension being either x, y, or z.

Figure 8 shows two example extractions from the raw dataset. When comparing the extractions of incisors, certain dominant sinusoid-shaped movement patterns seem to be visible. For example in Figure 8, the recommended movement pattern as taught to dental students for a 21-exodontia is a rotational one due to the shape of the root, as can be seen in table I, this is clearly visible as the dominant movement during the procedure with little movement on the other axes.

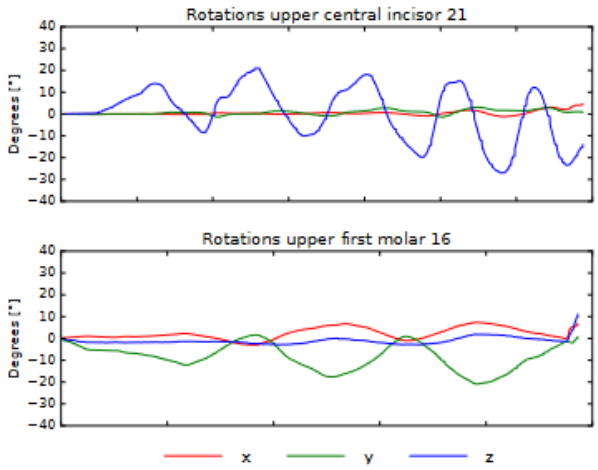

Fig. 8. Comparison of rotations of an upper incisor (21) and upper first molar (16). Adopted from Figure 6 in [17].

For removal of a 16, as seen in Figure 8, a rocking buccal movement is recommended, see table I, what is also observed as a negative sinusoidal (lingual is defined as positive) movement with increasing divergence as the tooth becomes looser with peaks at zero. During this procedure rotations around the longitudinal axis remain mostly absent.

From both figures we initially observe given features to be more descriptive of the performed action than others, which seems to be the result of a focused motion on a single axis during exodontia, for example rotation during the extraction of an incisor or the rocking of a molar.

3) Dataset Limitations: Due to the expensive nature of the available data (in terms of time, financial and ethical), only a very limited amount is available. As a result, this dataset does not allow for certain types of (unsupervised) analysis like deep learning and other in data-intensive methods of machine learning for extracting motion primitives from the available data.

Besides its limited size, the proposed dataset remains also biased, as most extractions have been performed by a single surgeon, but also because most of the cadaver material belonged to senior individuals mostly with prior conditions, so very limited to no data is available on younger patients with healthier teeth. However, at the same time most exodontia is performed on patients with older dentures with pre-existing conditions rather than younger healthier ones.

Another concern is related to the fact that all extractions have been performed using a robotic arm in gravity compensation mode in a static measurement setup. It cannot be guaranteed that the arm moved without any resistance or that it did not influence the behavior of the surgeon as he tried to accommodate the robot. All these factors impact the feasibility, reliability and applicability of the learnt movement primitives.

C. Related Work

This section will discuss related work. Section I-C.1 will discuss different ways of modelling movement primitives

found in literature and section I-C.2 will discuss different ways of learning movement primitive libraries in an unsupervised manner.

Whilst no related work seems to exist in the field of employing movement primitive based learning to analyze movements during exodontia specifically, related work exists in the use of movement primitives for analysis of movements in an unsupervised or semi supervised manner in the field of imitation learning where the objective is to teach a robot an action by means of a series of demonstrations. In general this prior art can be separated into two core processes consisting of the representation of a movement primitive and the learning / bootstrapping of the movement primitive library (how to extract these movement primitives from the data).

1) Movement Primitives: The current state provides several alternatives for representing and/or modelling a movement primitive, these mainly differ in whether these allow for dynamic or geometric modelling. The advantage of dynamic modelling over geometric modelling is that this allows for flexibility as the start and end points can be varied but this often comes at the cost of added (computational) complexity. An example of geometric modelling are Bezier Curves [18] that rely on multiple points to fit a curve, thus compressing more complex geometric shapes into a set of discrete points in space [19].

Examples of dynamic modelling of movement primitives are Dynamic Movement Primitives (DMP) [8], Extreme Learning Machine (ELM) [20], Stable Estimator of Dynamical Systems (SEDS) [21], Task Parameterized Gaussian Mixture Models (TP-GMM) [22], and Conditional Neural Movement Primitives (CNMP) [23].

In DMP the movement primitive is modelled as a dynamic system, but is not necessarily stable [24]. ELM on the other hand uses an extreme learning machine neural network to model a dynamic movement in a supervised manner, a downside of this method is that it more complex and that the method can be relatively data intensive and thus not suited to a sparse dataset [20]. SEDS encodes movement primitives as a finite mix of Gaussian functions using Gaussian mixture regression [21]. This method is more data efficient than ELM but remains quite computationally expensive due to the repeated calculation of kernel densities. An alternative to SEDS is TP-GMM where prior knowledge of tasks and the usage of multiple reference frames is leveraged to increase the generalizability of movement primitives [22]. Lastly, Conditional Neural Movement Primitives (CNMP) makes use of Conditional Neural Processes (CNPs) as proposed in [23] that acts as a probabilistic approximator to infer the distribution parameters, mean and variance, from observations of an unknown trajectory [25]. This method is the best at generalizing movement primitives but is very computationally intensive.

As it is not required to be able to generalize the learned movement primitives to novel situations, as the objective of this study is not imitation learning, it makes more sense to select a geometric representation for the movement primitives due to their low complexity and computational cost.

2) Movement Primitive Library Learning: After modelling movement primitives remains the component of learning / bootstrapping the movement primitive library in an unsupervised / semi supervised manner. The proposed methods in the prior art seem to suggest either a probabilistic approach or an iterative approach. A probabilistic approach aims to extract movement primitives based on expectation maximization of Gaussian Models whilst an iterative approach aims to extract movement primitives based on a recursive process.

Examples of a probabilistic approach are Probabilistic Segmentation (ProbS) [26], Real Time Unsupervised Segmentation (RUS) [27] and Gaussian Process-Hidden Semi-Markov Model (GP-HSMM) [28]. An example of iterative segmentation is self supervised bootstrapping [29].

Probablistic Segmentation (ProbS), as proposed in [26] starts by oversegmenting the observed trajectory into multiple segments, to then limit the number of movement primitives based on expectation maximization. ProbS aims to achieve this by learning a locally optimal MPL using probabilistic inference methods by assessing all possible (sequential) segmentations. This method is limited because the computational cost increases quickly with the number of initial segmentations.

An alternative approach, referred to by the authors as Real-time Unsupervised Segmentation (RUS) is presented by Tanako et al. [27]. The method consists of three phases, in the first phase continuous movements are discretized into short movements that are encoded into feature Hidden Markov Models (HMMs). Secondly, the causality between these feature HMMs is extracted by means of a correlation matrix. This data makes it possible to predict movement from observation. Lastly, movements that have a large prediction uncertainty are designated as the boundaries of MPs, i.e., the uncertainty of the following discrete feature following the current sequence of discrete features. This way, the boundaries of MPs can be found.

This method differentiates itself from other methods because it learns an MPL implicitly rather than explicitly by estimating the segmentation points of the data. Unlike ProbS, it requires no initial setup and can hence operate completely unsupervised. At the same time, like ProbS, it cannot create new segmentation points and has increased computational complexity due to the amount of recursive calculations required to train a large amount of HMMs, albeit less so than ProbS. This allows RUS to start with more (equidistant) initial segmentation points.

Because this method learns an MPL implicitly, there initially is no flexibility to use a different representation for MPs other than HMMs as this method partially relies on the probability of a segment following a sequence.

To address some of these concerns [28] proposes GP-HSMM (Gaussian Process-Hidden Semi-Markov Model) by segmenting the time-series data based on a stochastic model to estimate the length and class of the different segments. The proposed methods makes use of a combination of Gaussian process regressions and probability optimization using a

Gibbs sampler, more specifically the model makes use of a combination of a hidden semi-Markov model (HSMM) in combination with a Gaussian process (GP) as emissions where each state represents a unit action or MP. Segmentation points and classes are estimated by learning the model parameters. A downside to this method is that it is very computationally intensive.

In contrast, an approach as suggested by [29], called Self Supervised Bootstrapping (SSB), makes use of an iterative decomposition algorithm for decomposing the learned trajectories to extract an MPL. This can be broken down in three general steps, these are the decomposition of existing trajectories into multiple trajectory segments, then comparing these to the existing movement primitive library and either adopting the trajectory segment as a novel movement primitive or matching it to an existing one, lastly to consolidate the existing library based on the recurrence of learned movement primitives.

An advantage of an iterative approach like SSB is that it does not require an initial assumption for the shape or number of motion primitives. However the method could also be used in a semi supervised manner by initializing the MPL using a series of pre-trained MPs and hence offers flexibility.

For these reasons an iterative approach was selected because of the reduced computational cost, better generalizability and flexibility of movement primitive representation. A downside of iterative approaches like SSB is that these do not generalize well to multiple dimensions because of their iterative segmentation and clustering of time series data, for this reason a novel iterative method is proposed that generalizes better to multiple dimensions.

The proposed method, henceforth referred to as multi-dimensional self supervised bootstrapping (MDSSB), combines SSB with RUS to obtain a computationally efficient method that generalizes well to multiple dimensions and requires little to no assumptions about the number of primitives, the size of the primitives or the location of primitives.

This is achieved by first (over-)segmenting the dataset and iteratively clustering the extracted segments by implementing a custom distance function for clustering the segmented trajectories that simultaneously compares k n-dimensional segments at time t_1 to t_2 with k n-dimensional segments at time t_3 to t_4, wherein $t_2 - t_1 = t_4 - t_3$. Similar to SSB, MDSSB either adopts a segment as a new cluster or matches it to an existing one. Secondly, like SSB, MDSSB also completes with a consolidation step wherein the learned library is post-processed to remove outliers and combine similar or sequential clusters. Then, like in RUS, the labelled dataset is converted into a MPL based on the transition likelihood / occurrence of a given sequences of labels. This results in a computationally efficient method for extracting movement primitives from an n-dimensional dataset.

Given that within this project's scope synthesis of novel trajectories is not required, a strictly geometric (discrete) representation of a movement primitive was selected to reduce computational complexity.

II. Method

This section will discuss the method employed in this project. Firstly the segmentation of the dataset is explained in Section II-A after which the algorithm for clustering these segments is introduced in Section II-B, next the algorithm for extracting MPs from the clustered segments is introduced in Section II-C. Lastly, a summary of the complete process is provided in Section II-D.

A. Segmenting Data

MDSSB consists of first time-segmenting the dataset consisting of k extractions into discrete segments $s_{k,d,n}$ of respective length l_{seg}, wherein k refers to the extraction, d refers to the respective dimension either being x, y or z, and n refers to the respective segment number. Thus creating the four-dimensional dataset $\boldsymbol{S}$, where $|s_{k,d}|$ depends on k as not all extractions have the same number of segments. An example segmentation of a one dimensional time series is shown in Figure 9.

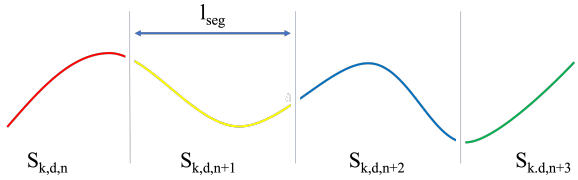

Fig. 9. Example segmentation of a 1-dimensional time series using the proposed method

B. Clustering Segments

Each segment is then organized in clusters, c^n. Each dimension in each segment $s_{k,d,n}$ is separately mean-centered and compared against each other segment using the mean squared error (MSE) based on a similarity threshold, M_{sim}, using the formula from Equation (1) for calculating the MSE for two segments $s^1_{k,d,n}$ and $s^2_{k,d,n}$, where p denotes the index of discrete points in the segment. All segments are labelled by iterating over k and n.

$$MSE(s^1, s^2) = \frac{1}{3 \cdot l_{seg}} \sum^{d\in(x,y,z)} \sum_{p=0}^{l_{seg}} (s^1_{k,d,n}[p] - s^2_{k,d,n}[p])^2 \tag{1}$$

Should the MSE between the mean of an existing cluster and a segment exceed the similarity threshold, a new cluster c^{n+1} is formed, should the MSE remain less than the similarity threshold, the segment is added to that cluster c^n and the respective cluster's mean is updated. This process repeats until each segment is assigned a cluster. This clustering algorithm is similar to that of SSB, with the difference being that this uses a modified distance function. This method has the advantage that all segments in the found clusters are guaranteed to be within a predefined spread and that the method does not require an initial assumption or guess regarding the location, size or number of clusters.

An inherent tradeoff exists between the accuracy and computational complexity of segmenting the extraction and l_{seg}. Should l_{seg} be selected too large, critical compound movements might be completely enclosed in a single segment and thus missed during segmentation. Should l_{seg} be selected too small each segment might not contain sufficient detail to be able to accurately cluster each segment. After repeated experimentation, an optimal l_{seg} of 100 samples (approx. 1s) was found to be optimal.

After each segment is assigned a cluster, a consolidation step takes place. During the consolidation phase each mean of each cluster is compared to the mean of each other cluster. If two clusters' means show a similarity less than the similarity threshold then these clusters are combined. The result of this initial segmentation is shown for an example extraction of a 12-incisor in Figure 10. From this it can be observed that this initial segmentation has distinguished an upward motion (shown in dark blue), a downward motion (shown in red) and a transitional phase (shown in pink).

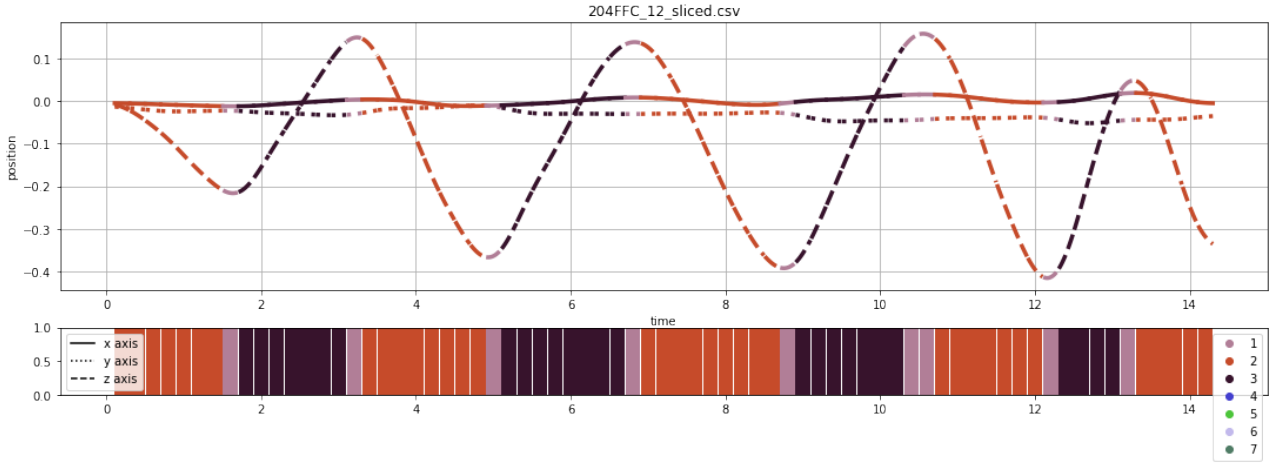

Fig. 10. Exemplary segmentation into clusters of a extraction according to the proposed method. The vertical axis represents rotation in π radians and the horizontal axis represents time in seconds. Each color represents a respective cluster and each line style represents a respective dimension. The measurement has been segmented into multiple blocks of a fixed length of 100 samples, as can be seen in the vertical stripes in the bottom figure.

C. Extracting Movement Primitives

After the consolidation phase has completed, the clusters are organized into movement primitives. A movement primitive in this context is a repeating sequential pattern of labelled segments that contains no repeating patterns on its own. The algorithm uses a greedy-like approach to identify movement primitives in the labelled segments, wherein it seeks to maximize the data explained by a single movement primitive. An example would be the labelled / encoded series $[1, 2, 3, 1, 2, 3, 4]$ where the movement primitive, or shortest possible non-repeating sequence that explains the most data, would be $[1, 2, 3]$. As a result, this sequence would be re encoded as a movement primitive, resulting in the following sequence $[A, A, 4]$. This process repeats until either a given threshold of labelled data has been encoded as movement primitives or no more movement primitives can be found that explain more than a given threshold of data.

For this reason, the algorithm only considers movement primitives of a length equal to a prime number, as any found movement primitive of a non-prime length would by definition contain at least two shorter movement primitives in sequence. The algorithm also only considers prime numbers up to half the length of the array, as any primitive of a larger length would only appear once and would therefore be trivial.

This algorithm can be run multiple times to further consolidate the set of found movement primitives. For example if multiple movement primitives always appear in sequence, then in a new iteration these are consolidated into a new, larger movement primitive. A downside of this method however, is that the more iterations of this algorithm are performed, the more it will converge to the trivial solution of a single movement primitive for the entire time series. The complete algorithm is shown in Algorithm 1.

Algorithm 1 Algorithm for identifying movement primitives

Input: $\boldsymbol{S}$, $\boldsymbol{C} = \{c^1, ..., c^n\}$, $\boldsymbol{P}^* = \{p \in \boldsymbol{P} : 2 \leq p < \frac{1}{2}max(|s_{k,d,n} \in S|)\}$, $0 \leq q \leq 1.0$ **Output:** $\boldsymbol{MP}$

$\boldsymbol{S}_{collapsed} = Collapse(\boldsymbol{S})$
$\boldsymbol{S^*} = \boldsymbol{S}_{collapsed}$
$\boldsymbol{MP} = \{\}$
while $|\boldsymbol{S^*}| \geq q \cdot |\boldsymbol{S}|$ **do**
  $N = \{\}$
  for $\forall p \in \boldsymbol{P}^*$ **do**
    $C = Comb(\boldsymbol{C}, p)$
    for $\forall c \in C$ **do**
      $N \leftarrow \{c : Lookup(\boldsymbol{S^*}, c)\}$
    end for
  end for
  $c_{max} = \max(N)$
  $MP \leftarrow c_{max}$
  $\boldsymbol{S^*} = \text{Delete}(\boldsymbol{S^*}, c_{max})$
end while

In Algorithm 1, $\boldsymbol{C}$ is the collection of all found clusters and $\boldsymbol{P^*}$ is the collection of all prime numbers larger or equal to two and smaller or equal to half the size of the largest set of segments in dataset $\boldsymbol{S}$, q is a search parameter between 0 and 1 determining how much data needs to be explained before recursion stops, and $\boldsymbol{MP}$ is the collection of all learned movement primitives.

First, in Algorithm 1, the dataset $\boldsymbol{S}$ is collapsed using the $Collapse$-function that returns the inputted array wherein each entry is different from its neighbors, it achieves this by removing sequences of the same entry. An example would be the array $[1, 1, 3, 1, 2, 2, 4]$ where the function would return $[1, 3, 1, 2, 4]$. After this, a parallel dataset S^* is initialized by setting this equal to the original dataset S and $\boldsymbol{MP}$ is initialized as an empty array.

The next code block is executed as long as S^* remains larger than S. Firstly an empty dictionary, N, is initialized. Then for each prime number, p, in $\boldsymbol{P^*}$, C is calculated using the $Comb$-function that takes in an array and an integer value k and returns a a two dimensional array containing all combinations of the inputted array of length k. Then, for each combination, c stored in C the previously initialized dictionary N is appended with the result of the $Lookup$-function, that returns the maximum number of non-overlapping instances that c appears in all of $\boldsymbol{S}$ multiplied with its length. An example would be the array $[1, 2, 3, 1, 2, 3, 4]$ where the combination $[1, 2]$ would return 4.

Having computed this value for each combination, c, of each length, p, the combination with the maximum occurrence is selected as a movement primitive, as out of all possible movement primitives this explains the most data. This movement primitive is added to the array of movement primitives, $\boldsymbol{MP}$, and subtracted from the parallel dataset, $\boldsymbol{S^*}$ using the $Delete$-function that returns a new dataset segmented using the selected movement primitive as a breakpoint. An example would be the array $[1, 2, 3, 1, 2, 3, 4]$ where the combination $[1, 2]$ would return $[[3], [3, 4]]$.

This process is then repeated until the amount of explained data as a fraction of the total data no longer exceeds the threshold q. A q that is too high would induce non-representative movement primitives, as movement primitives become more sparse the more iterations are required and a q that is too low would leave out key movement primitives. For this dataset an optimal q was found to be 0.5.

An example result of this algorithm for an exemplary 12-incisor extraction is shown in Figure 11. From this figure it can be observed that this method has extracted the sequence of 3,1 and 1,2 as movement primitives. From this a repeating sequence of learned movement primitives can be observed. It is noted that from the visualization in Figure 12 and the example in Figure 11 only MPs of length 2 are extracted, but the proposed method allows for MPs of any (viable) length.

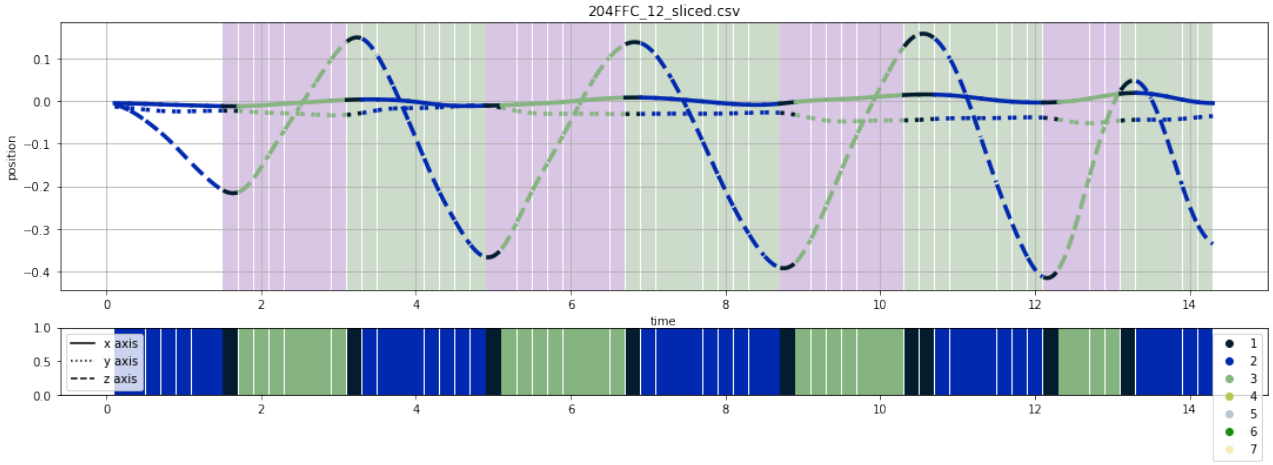

Fig. 11. Exemplary segmentation into clusters of a extraction according to the proposed method. The vertical axis represents rotation in π radians and the horizontal axis represents time in seconds. Each color in the top figure represents a respective movement primitive, each color in the bottom figure represents a respective cluster (see also legend in bottom right of the figure), and each line style represents a respective dimension (see also legend in bottom left corner). The measurement has been segmented into multiple blocks of a fixed length of 100 samples, as can be seen in the vertical stripes in the bottom figure. In this figure, two movement primitives can be seen; These are marked purple and green. Both movement primitives consist of two sequential labelled segments. The movement primitive marked in purple consists of labelled segments 1 and 3, and the movement primitive marked in green consists of labelled segments 1 and 2.

D. Summary

A summary of the complete method is visualized in Figure 12. Here a one-dimensional movement is first segmented into discrete segments, $s_1 ... s_n$, after which segments are labelled using MSE-comparison, resulting in the encoded sequence $[c_{x0}, ..., c_{xn}]$, wherein $x_n \in C$. The resulting segmented labelled data is thus encoded into a sequence of labels, after which sequential identically labelled segments are merged resulting in the encoded data.

From the encoded data a common repeating element, $c_2 - c_3$, is observed, this sequence is then extracted as a new movement primitive and added to the library. This last step is repeated until a given threshold of the data has been explained or when any new movement primitive explains

less than a given threshold of new data. Each segment of the resulting labelled time series is either labelled as a movement primitive or as unlabeled / leftover data.

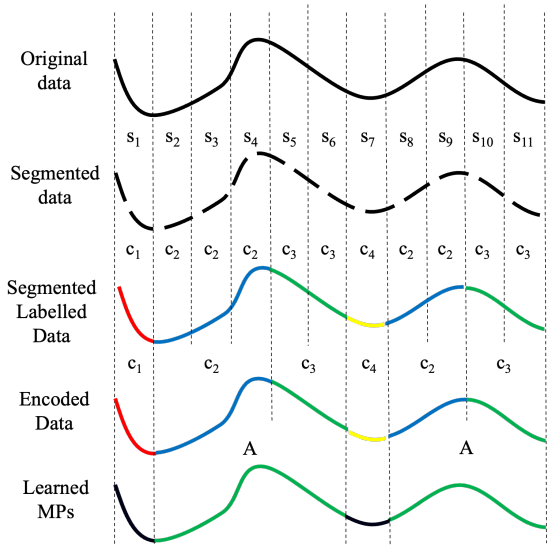

Fig. 12. Visualization of method, wherein a one-dimensional movement is segmented into discrete segments, each of which is labelled. Afterwards sequences of repeating labels are merged into single sections after which movement primitives are extracted based on common repeating sequences.

This method is applied to the entire dataset. The resulting movement primitives are then used for further analysis.

III. Results

This section will discuss the results from applying MDSSB to the complete dataset. The found movement primitives are shown in Section III-A, the prevalence of each movement primitive in each major tooth group in the dataset is shown in Section III-B, lastly an alternative method of movement primitive extraction, namely Fourier analysis, is applied and shown in Section III-C as a baseline comparison.

A. Movement Primitives

Using the method described in section II four movement primitives were obtained from the complete dataset. These are shown in Figure 14 using an overlay plot wherein each instance of the movement primitive in each dimension is separately and jointly overlayed. Next to this, a pie chart showing the occurrence (i.e. the number of times the primitive appears in each major tooth group) of each respective primitive as a fraction of each primitive's total occurrence. For example, an occurrence of thirty percent would mean that thirty percent of all instances of that movement primitive occurred within that group. Each of the found movement primitives can be observed in Figure 14.

Movement primitive zero seems to be a general rotation comprising an parabolic rotation in the z-dimension and a similar rotation but with smaller amplitude in the y-dimension with a relatively flat rotation in the x-dimension. This may be interpreted as a general rocking and twisting motion of the tooth, which seems like a general type of rotation. When observing the occurrence of primitive zero in Figure 14, it is noted that primitive zero seems almost equally present in all major tooth groups, with the exception of premolars being seemingly overrepresented at the cost of the canines.

Primitive one, shown in the second row of Figure 14, seems more expressed in the z-dimension relative to the other dimensions which seem to remain mostly flat. When observing the pie chart, this type of movement primitive seems much more common in the incisors, canines and premolars in contrast to the molars. A dominant rotation in the Z-dimension might be interpreted as a strong twisting motion. This primitive seems to be mostly present in the incisors and premolars compared to the molars and canines.

Primitive two, shown in the third row of Figure 14, seems even more expressed in the z-dimension when compared to primitive one. The difference seems to be that primitive one describes an upward curved rotation and primitive two a rapid downward rotation. This movement primitive, even more so than primitive one, seems to describe a fast singular rotation of the tooth about the longitudinal axis. This primitive seems more common in the the incisors and the canines compared to the molars and premolars

Primitive three, shown in the fourth row of Figure 14, seems a much more complex rotation when compared to primitive one or two, as general rotation / movement is observed in each dimension, but most movement seems concentrated in the X and Y-dimension. This primitive also seems to be highly present in the molars and premolars (72 percent) compared to the canines and the incisors (28 percent), even more so than primitive two. This primitive could be interpreted as a general compound movement / rocking motion of the tooth.

B. Prevalence

The prevalence of the learned movement primitives is shown in Figure 13.

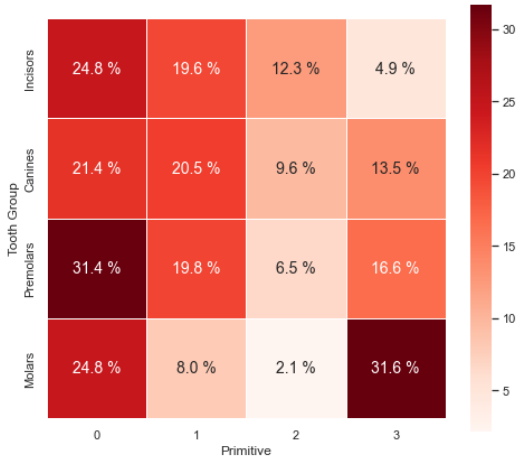

Fig. 13. Heatmap of found movement primitives their prevalence in each of the major tooth groups. Each fraction shows how much of a specific group can be explained by a single primitive, for instance primitive three explains 32 percent of the molars whilst only explaining approximately 5 percent of the incisors.

Prevalence in this context is meant to indicate the amount of data a movement primitive explains within a specific tooth group. For example, primitive zero and primitive three together explain 56.4 percent of molar extractions in the dataset (24.8 + 31.6). These rows do not sum to 100 as certain parts of the data remain unlabelled. Prevalence differs from occurrence in two ways, firstly prevalence measures what fraction of data is explained by each movement primitive whereas occurrence counts the number of instances of the primitives, secondly prevalence counts data within tooth groups whereas occurrence counts within primitives.

From Figure 13 can be observed that the seemingly more complex movement primitives (zero and three) explain a larger fraction of the data in the molars and premolars compared to the rotation dominant movement primitives (one and two). At the same time, it can be seen that the more rotation focused movement primitives explain a comparatively larger fraction of data of the incisors, canines and premolars compared to the molars. At the same time, it can be observed that primitive 2 explains a relatively smaller portion of each major tooth group compared to the other primitives, but still remains relatively dominant in the incisors and canines compared to its prevalence in the premolars and molars.

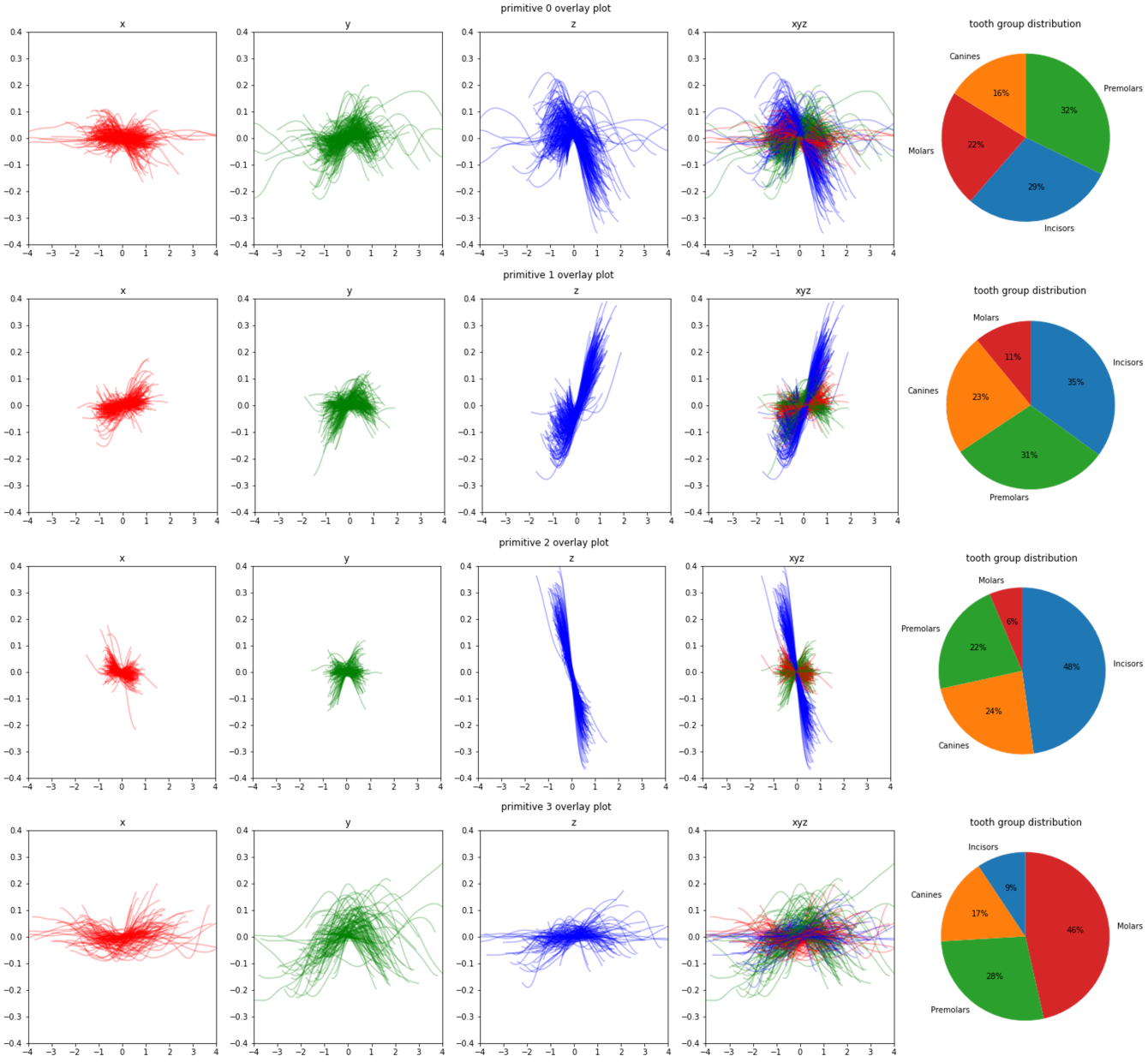

Fig. 14. Visualization of movement primitives zero to three from top to bottom, from left to right: the first four figures include the X, Y and Z-limited and total overlay plot of the found movement primitives wherein the horizontal axis represents mean centered time and the vertical axis rotation in π radians. Lastly a distribution pie chart of the occurrence of each primitive in each of the major tooth groups is included.

C. Fourier analysis

The dataset has also been analyzed using Fourier analysis as a baseline. When observing Table I it can be noted that the current state of exodontia seems to rely on periodic repeating rotations, so theoretically Fourier analysis of the performed extractions is relevant in distinguishing the speed and complexity of the rotations during exodontia.

Each dimension in each extraction in the dataset has been analyzed using Fast Fourier Transform (FFT), subsequently for each major tooth group and dimension these discrete results have been overlayed, after which a Kernel Density Estimation (KDE) analysis yielded the distributions of the Fourier analyses of each dimension in each major tooth group. The results are shown below in Figure 15. In these figures negative frequencies have been disregarded as these do not seem to have real world relevance in dental extraction. Additionally all frequencies whose amplitude was less than 5 percent of the amplitude of an extraction its maximum amplitude in the respective dimension have been disregarded to make the resulting figures more legible.

Figure 15 can be interpreted as a representation / distribution of the speed and magnitude of sinusoidal movements present during exodontia. The horizontal axis showing the frequency in rad/s, where a low frequency corresponds to a large, slow movement and a high frequency to a fast, short movement. The vertical axis represents magnitude, showing how dominant / present these movements at a given frequency are.

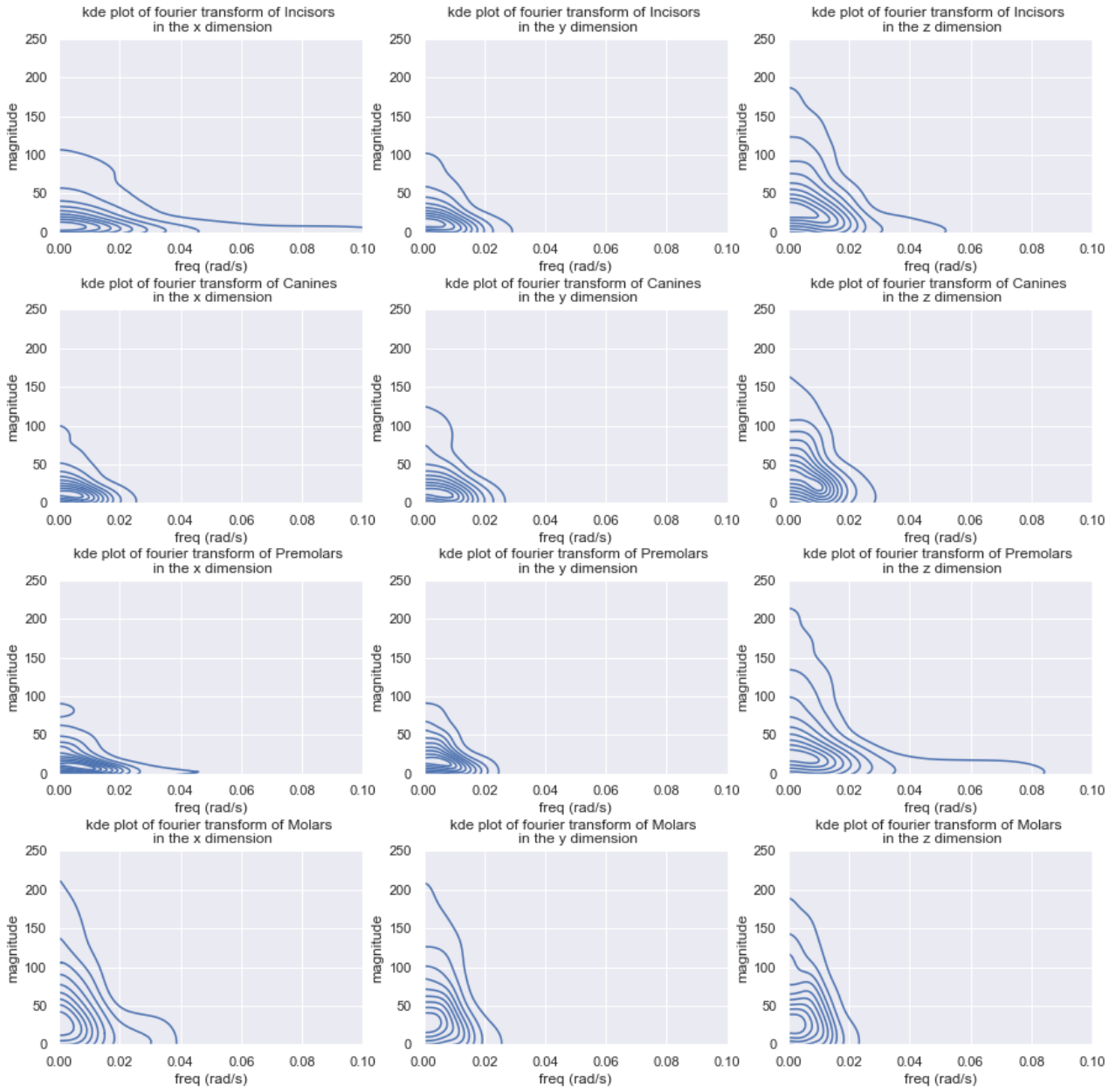

Fig. 15. Fourier analysis of all extractions in the dataset grouped on direction (left to right) and tooth group (top to bottom). Each direction in each extraction has been analyzed using Fast Fourier Transform (FFT). All these discrete Fourier transforms have then been overlayed after which KDE analysis has been employed to estimate a distribution of found extractions in the frequency domain. To improve legibility of the figures, frequencies with amplitudes less than 5 percent of that analysis' maximum have been excluded.

Figure 15 seems to show the molars having an approximately equal distribution about each respective axis compared to the other major tooth groups, suggesting a more complex nature of (large) rotations required to remove these teeth.

From Figure 15 a distinctive difference between each major tooth group does not become apparent. It is observed that each direction for each tooth group is highly concentrated within low frequencies between zero and 0.02 rad/s. Differences do become apparent in the relative amplitude of rotations in the local Z-dimension in the incisors compared to the other axes. The same can be mentioned of the premolars. In the canines a similar pattern can be observed but this seems less distinctive compared to the premolars.

IV. Discussion

The research question of this project was to identify whether movement primitive based analysis as employed in imitation learning could be used to deepen our understanding of exodontia. To achieve this, a novel method for identifying movement primitives was proposed and applied to a dataset of clinically representative extractions performed by OMFS surgeons.

Using this method four unique movement primitives have been identified (see Figure 14 and Figure 13) that show distinctive occurrence and prevalence in each of the major tooth groups. These seem to describe a general rocking and twisting movement (primitive zero), a strong positive rotation in the Z-dimension (primitive one), a fast negative rotation in the Z-dimension (primitive two), and a general low amplitude compound rotation (primitive three).

When comparing the found movement primitives with literature (see also Section I-A, Table I), it is noted that rotation focused movement primitives (primitives one and two) appear mostly in incisors and canines (61 and 66 percent respectively). This makes sense as these teeth only have a singular root allowing for such rotation whereas premolars and molars have multiple roots preventing such rotations.

Furthermore, it can be seen that primitive one and two explain 31.9 percent and 30.1 percent of the incisor and canine movement data respectively, whilst only explaining 10.1 percent of the molars, see Figure 13. This appears also in line with literature, see Table I, wherein a twisting motion is recommended for incisors and canines and a rocking and twisting motion for premolars.

It is also observed that smaller compound rotations / movements (primitive four) mostly seem to occur within the molars and premolars (74 percent), this seems in line with literature, wherein a rocking motion is recommended for molars and a rocking and twisting motion for premolars, see also Table I.

It is also observed that a general type of movement primitive, primitive zero, seems mostly omnipresent in each tooth group. This is reflected in both its occurrence and prevalence in each major tooth group, see Figure 13 and Figure 14. This movement primitive could be interpreted as a general type of rocking and twisting motion.

It can also be argued that this type of analysis has yielded a more in depth result compared to the simpler method of Fourier analysis that is limited to periodic sinusoidal patterns for movement primitives. This assumption was based on the contents of table I which merely describes various oscillating patterns for exodontia depending on the tooth type in question.

A. Limitations of research

Several comments related to the research could be made. Some of the most important are summarized.

The used dataset was both limited and biased, as described in par. I-B. This could have influenced the reliability of the made conclusion and analysis, as the provided method might not be representative of exodontia in general but rather the specific method of the OMFS surgeon performing the extractions in the dataset.

Furthermore, the proposed method has been tuned specifically to the provided dataset, which in turn has been preprocessed based on expectations of the characteristics of movement primitives in exodontia. In that sense the method has not been objective, as the author might have tuned the method to provide results in line with his expectation rather than allow a natural result to occur which might have yielded more or different movement primitives.

More specifically, given the limited availability of literature, little comparison material is available. Therefore, it remains unclear what exactly an unsupervised method like the one in this paper is expected to return. For that same reason, it remains unclear as to what exactly has been achieved as an improvement over the prior art.

Furthermore, the employed method is limited to sequential movement primitives, movement primitives that overlap or combine to create new movement can not be detected using this method. Therefore, it remains unclear whether the discovered movement primitives are a result of a combination of more fundamental underlying movement primitives or are unique.

In an attempt to evaluate whether a more complex method like one based on movement primitives has yielded an improvement over more simple movement decomposition, a Fourier decomposition has been employed. It can be argued that this method of analysis is limited in its ability to (dis-)prove the effectiveness of the provided method as more possible methods for deconstructing such movement data exist, therefore the comparison could be argued to be incomplete.

B. Future Work

To make the provided method more reliable and/or to increase the reliability of the provided dataset, it is recommended to further research methods for collecting data on exodontia on a larger scale. To achieve this a method needs to be researched that lowers the cost of data (both in a financial and ethical sense).

The learned movement primitives could be further analyzed and evaluated for their educational value and inter-

pretability by providing these to a panel of OMFS professionals and students. In line with this proposed research, the learned movement primitives could be further analyzed.

The provided method relies on a simplified geometric representation of movement primitives, in a further research a more complex method could be employed. A prime candidate could be SEDS for its stability and flexibility. This allows for dynamic start and endpoints by allowing synthesis of trajectories within a movement primitives, allowing for study of movement primitive dynamics in exodontia, and whether this would be of added value.

Additionally, the current method is unable to distinguish combined or overlapping movement primitives. Future research could comprise developing a novel method capable of discovering such movement primitives.

C. Conclusion

A movement primitive library containing four distinctive movement primitives could be constructed using the proposed method from a limited and biased dataset. These movement primitives seem to correlate with what is described in literature (see Table I) and seem to outperform simpler methods like Fourier analysis. This seems to suggest merit to the proposed method.

Whilst more research is required to judge and/or evaluate the value of the learned movement primitives for dental professionals and/or in education, the results presented in this project do not seem to discourage from further pursuing a movement primitive based method of analysis of exodontia, therefore it could be argued that the research question could be answered in the affirmative. Primitive based learning seems promising for deepening our understanding of exodontia.

REFERENCES

[1] Gaballah and Kamis, "Avoiding common problems in tooth extractions," *DentalTribuneUK Edition*, 10 2015.

[2] M. Cicciù, E. Bramanti, S. F., A. Cicciù, and F. Sortino, "Experimental study on strength evaluation applied for teeth extraction: an in vivo study," *The open dentistry journal*, p. 20–26, 2013.

[3] T. Ojala, "Rocking and twisting moments in extraction of teeth in the upper jaw," *International journal of oral surgery*, p. 367–372, 1980.

[4] R. Bos, B. Stegenga, A. Vissink, L. d. Bont, F. Spijkervet, M. Doff, A. Hoekema, J. H. Slater, J. Jansma, H. Meijer, and et al., *extractieleer en dento-alveolaire chirurgie*. Van Gorcum, 2013.

[5] H. S. Brand, C. van der Cammen, S. Roorda, and J. A. Baart, "Tooth extraction education at dental schools across europe." *BDJ open*, 2015.

[6] R. Lehtinen and T. Ojala, "Recente toename van reguliere extracties in mka-chirurgie onderzocht," *Nederlands Tijdschrift voor Tandheelkunde*, pp. 365–372, 2016.

[7] C. Hanson, T. Wilkinson, and M. Macluskey, "Do dental undergraduates think that thiel-embalmed cadavers are a more realistic model for teaching exodontia?" *European journal of dental education*, p. e14–e18, 2018.

[8] H. Kimura, K. Tsuchiya, A. Ishiguro, and H. Witte, *Adaptive Motion of Animals and Machines (p. 261-275)*, 1st ed. Springer, Tokyo, 2006.

[9] D. Kulić, t. C. Ott, D. Lee, J. Ishikawa, and Y. Nakamura, "Incremental learning of full body motion primitives and their sequencing through human motion observation." *he International Journal of Robotics Research*, pp. 330–345, 2012.

[10] Pfeiffer, Sammy, Angulo, and Cecilio, "Gesture learning and execution in a humanoid robot via dynamic movement primitives," *Pattern Recognition Letters*, vol. 67, 08 2015.

[11] A. Lemme and A. Soltoggio, "Movement primitives as a robotic tool to interpret trajectories through learning-by-doing," *International Journal of Automation and Computing*, vol. 10, pp. 375–386, 10 2013.

[12] R. Sulyanto, "Dental anatomy and development - dental disorders," Dec 2021. [Online]. Available: https://www.msdmanuals.com/professional/dental-disorders/approach-to-the-dental-patient/dental-anatomy-and-development

[13] A. R. Hussain, K. Kachymalay, and A. A. Aziz, "Interactive dental charting: Towards an electronic dental information system," pp. 794–797, 2010.

[14] "Tooth." [Online]. Available: https://www.mouthhealthy.org/en/az-topics/t/tooth

[15] Kay4yk180718, "Dental tool for removing teeth. close top view. removal of a tooth. dental pliers, tooth extraction (stock image)." [Online]. Available: https://www.dreamstime.com/dental-tool-removing-teeth-close-top-view-removal-tooth-pliers-extraction-image1652

[16] W. de Graaf, "Understanding tooth removal procedures using feature engineering and classification models," 11 2020.

[17] T. C. van Riet, W. M. de Graaf, R. van Antwerpen, J. van Frankenhuyzen, J. de Lange, and J. Kober, "Robot technology in analyzing tooth removal - a proof of concept," *International Conferences of the IEEE Engineering in Medicine and Biology Society (EMBC 2020)*.

[18] C. Bitter, T. Thun, and T. Meisen, "Bézier curve based continuous and smooth motion planning for self-learning industrial robots," *Procedia Manufacturing*, vol. 38, pp. 423–430, 01 2019.

[19] J. Faraway, M. Reed, and J. Wang, "Modelling three-dimensional trajectories by using b??zier curves with application to hand motion," *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, vol. 56, pp. 571 – 585, 11 2007.

[20] J. Duan, Y. Ou, J. Hua, Z. Wang, S. Jin, and C. Xu, "Fast and stable learning of dynamical systems based on extreme learning machine," *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS*, vol. 49, pp. 1175–1185, 2019.

[21] S. M. Khansari-Zadeh and A. Billard, "Learning stable nonlinear dynamical systems with gaussian mixture models," *IEEE Transactions on Robotics*, vol. 27, no. 5, pp. 943–957, 2011.

[22] T. Alziadeh and B. Saduanov, "Robot programming by demonstration of multiple tasks within a common environment," 11 2017, pp. 608–613.

[23] M. Garnelo, D. Rosenbaum, C. J. Maddison, T. Ramalho, D. Saxton, M. Shanahan, Y. W. Teh, D. J. Rezende, and S. M. A. Eslami, "Conditional neural processes," 2018.

[24] M. Saveriano, F. J. Abu-Dakka, A. Kramberger, and L. Peternel, "Dynamic movement primitives in robotics: A tutorial survey," *CoRR*, vol. abs/2102.03861, 2021. [Online]. Available: https://arxiv.org/abs/2102.03861

[25] M. Y. Seker, M. Imre, J. H. Piater, and E. Ugur, "Conditional neural movement primitives." in *Robotics: Science and Systems*, vol. 10, 2019.

[26] R. Lioutikov, G. Neumann, G. Maeda, and J. Peters, "Learning movement primitive libraries through probabilistic segmentation," *The International Journal of Robotics Research*, vol. 36, no. 8, pp. 879–894, 2017. [Online]. Available: https://doi.org/10.1177/0278364917713116

[27] W. Takano and Y. Nakamura, "Real-time unsupervised segmentation of human whole-body motion and its application to humanoid robot acquisition of motion symbols," *Robotics and Autonomous Systems*, vol. 75, pp. 260–272, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092188901500216X

[28] T. Nakamura, T. Nagai, D. Mochihashi, I. Kobayashi, H. Asoh, and M. Kaneko, "Segmenting continuous motions with hidden semi-markov models and gaussian processes," *Frontiers in Neurorobotics*, vol. 11, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fnbot.2017.00067

[29] A. Lemme, R. F. Reinhard, , and J. J. Steil, "Self-supervised bootstrapping of a movement primitive library from complex trajectories," *IEEE-RAS International Conference on Humanoid Robots (Humanoids)*, 2014.