UNIVERSITY OF CALIFORNIA

Los Angeles

Time-series Classification for Monitoring Toothbrushing Behavior Using Motion Sensors

A dissertation submitted in partial satisfaction

of the requirements for the degree

Doctor of Philosophy in Electrical and Computer Engineering

by

Mahmoud Essalat

2023

ABSTRACT OF THE DISSERTATION

Time-series Classification for Monitoring Toothbrushing Behavior Using Motion
Sensors

by

Mahmoud Essalat

Doctor of Philosophy in Electrical and Computer Engineering

University of California, Los Angeles, 2023

Professor Gregory J. Pottie, Chair

Digital health technologies, such as smart toothbrushes, hold great potential for unobtrusively monitoring brushing behaviors in home settings by utilizing motion-sensors, including accelerometers, gyroscopes, and magnetometers. Some studies have attempted to modify toothbrushes by attaching these sensors to the brush handle, while others have employed external devices, such as wrist-watches to identify the dental regions being brushed during a brushing session. Although these research studies show promising preliminary results, they were conducted under structured toothbrushing assumptions performed in controlled laboratory settings (e.g. limited head and body movement, predefined sequence of brushing) and not the free-form brushing observed in real-world settings.

To address the aforementioned issue, we collected a dataset of 187 brushing sessions, including free-form brushing, and present, to the best of our knowledge, the first motion-sensor dataset obtained during free-form brushing. We label these brushing session using concurrent video recordings as the ground truth. Given the fast and frequent transitions of the brush, labeling is prone to errors, known as the noisy label scenario in machine learning, and to address this challenge, we propose a relabeling algorithm for dataset cleaning. To perform a statistical analysis of brushing behavior for the first time, we utilized a zero-inflated mixed-effect generalized linear regression model to examine the brushing time for each dental region. Pre-

vious studies frequently depended on non-parametric tests, such as the Wilcoxon test, which are not suitable for repeated measurement studies like ours, involving multiple participants each brushing for multiple sessions. Our results show that individuals generally brushed their buccal teeth surfaces (i.e. outer teeth surfaces) more than twice as long as the occlusal (i.e. flat teeth surfaces) (2.18 times longer (95% CI 1.42, 3.35; p leq 0.001)) and lingual surfaces (i.e. inner teeth surfaces) (2.22 times longer (95% CI 1.62, 3.10; p leq 0.001)). Additionally, we also propose a three-stage method (i.e. pre-processing, change point detection (CPD), and time-series classification) to detect the teeth surfaces brushed during a session using motion-sensor data. We present CluMing, a novel change point detection (CPD) algorithm based on clustering, which outperforms other CPD algorithms in our application. We compare the results of multiple time-series classification methods in machine learning such as the feature engineering method, LSTM, and Transformer models. Our findings indicate that high classification accuracy can be achieved using a random train-test split of the data (i.e. k-fold cross-validation); however, considering the high variations in brushing data, generalization beyond the participants in the training set (i.e. one-subject-out cross-validationin), may need additional aid such as domain knowledge transfer or personalization. We validate our findings by applying our proposed method to our provided dataset, as well as the datasets of toothbrushing in controlled settings.

The dissertation of Mahmoud Essalat is approved.

Oscar Hernan Madrid Padilla

Xiang (Anthony) Chen

Lieven Vandenberghe

Vivek Shetty

Gregory J. Pottie, Committee Chair

University of California, Los Angeles

2023

To my mom who put her heart into my education.

Contents

List of Figures

List of Tables

Acknowledgements

Completing this dissertation signified a journey abundant with diverse experiences, ranging from trials to triumphs. It was a testament to perseverance, steadfast dedication, and resource-fulness. Throughout this endeavor, my family, friends, and advisors emerged as invaluable pillars, providing support during challenging times, lifting my spirits, and redirecting me toward success.

I am immensely grateful to Prof. Pottie for his consistent support during my PhD journey. His encouragement cultivated an environment conducive to my growth and the completion of this dissertation. His guidance emphasized focusing on 'what I can achieve' in life, steering me away from fixating on limitations. I deeply appreciate his dedication in extracting the best from this dissertation, devoting time and energy to help me reach this milestone. I extend my gratitude to Prof. Padilla for his unwavering support and availability for both technical and non-technical discussions. His contributions to the statistical analysis of this dissertation and his mentorship remain invaluable. I also express heartfelt appreciation to Prof. Shetty for his commitment to advancing the project and his guidance on the dental aspects. I am thankful for his provision of resources, facilitating connections and collaborations through the mHealth program, and financial support, which were instrumental in progressing through this dissertation. Additionally, my thanks go to Prof. Vandenberghe and Prof. Chen for their guidance on my PhD committee.

I extend my gratitude to Deeona Columbia, the graduate student counselor, for her genuine care and support within the department. Her professional approach to handling various issues, despite the influx of requests, has been invaluable.

Furthermore, I appreciate the collaboration and contributions of my former labmate, Dr.

Curriculum Vitae

Education

- **Ph.D. Candidate in Electrical and Computer Engineering:** University of California, Los Angeles (UCLA), U.S., 2017 - 2023

- **M.Sc. in Electrical and Computer Engineering:** University of California, Los Angeles (UCLA), U.S., 2017 - 2020

- **B.Sc. in Electrical Engineering:** Sharif University of Technology, Iran, 2012 - 2017

Publications

1. Essalat, M., Mashhadi, M. B., & Marvasti, F. (2016, December). Supervised heart rate tracking using wrist-type photoplethysmographic (PPG) signals during physical exercise without simultaneous acceleration signals. In 2016 IEEE global conference on signal and information processing (GlobalSIP) (pp. 1166-1170). IEEE.

2. Mashhadi, M. B., Essalat, M., Ahmadi, M., & Marvasti, F. (2016, November). An improved algorithm for heart rate tracking during physical exercise using simultaneous wrist-type photoplethysmographic (ppg) and acceleration signals. In 2016 23rd Iranian Conference on Biomedical Engineering and 2016 1st International Iranian Conference on Biomedical Engineering (ICBME) (pp. 146-149). IEEE.

3. Mashhadi, M. B., Farhadi, M., Essalat, M., & Marvasti, F. (2018, April). Low complexity heart rate measurement from wearable wrist-type photoplethysmographic sensors robust

to motion artifacts. In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 921-924). IEEE.

4. Farhadi, M., Mashhadi, M. B., Essalat, M., & Marvasti, F. (2016). Realtime heart rate monitoring using photoplethysmographic (PPG) signals during intensive physical exercises. bioRxiv 092627.

5. Essalat, M., Morrison, D., Kak, S., Chang, E. J., Penso, I. R., Kulchar, R. J., ... & Shetty, V. (2022). A naturalistic study of brushing patterns using powered toothbrushes. Plos one, 17(5), e0263638.

6. Essalat, M., Abolhosseini, M., Le, T. H., Moshtaghion, S. M., & Rezai Kanavi, M. (2022). An Interpretable Deep Learning Approach for Automatic Diagnosis of Fungal Keratitis and Acanthamoeba Keratitis Based on an Introduced Database of in Vivo Confocal Microscopic Images.

7. Essalat, M., Madrid Padilla, O. H., Shetty, V., & Pottie, G. (2023). Monitoring Brushing behaviors using Toothbrush Embedded Motion-Sensors.

Chapter 1

Introduction

As a majority of Americans have experienced or will experience injurious oral diseases, such as caries or gingival disease, that can negatively impact their quality of life, healthcare professionals have made it a priority to improve the oral health of this population [24]. The primary cause of periodontal (gum) disease and caries is dental plaque [29, 57]. Plaque, which is a sticky film of bacteria that covers the teeth, can cause tooth and gum decay by producing acids that break down enamel and irritate the gums if it is not removed [56].

Brushing one's teeth is the primary mechanism for removing plaque; however, proper brushing technique is crucial to ensure effective plaque removal [33, 59]. Unfortunately, many people are not aware of the proper brushing techniques, resulting in suboptimal oral health practices. In fact, studies have shown that 46% of children only brush their teeth once a day [3]. It is widely recognized that systematic, twice-a-day tooth brushing is essential for maintaining good oral health and preventing dental diseases [82]. Despite this knowledge, this fundamental behavior is not as widely and fully practiced as dental professionals and health organizations would like.

The challenges of promoting optimal oral health behaviors (OHBs) are numerous and complex. They include a confusing array of brushing recommendations, with some being too intricate to perform consistently; expecting generic oral hygiene instructions (OHIs) to work equally well among dissimilar populations; viewing OHI as a one-time event rather than an ongoing process that extends beyond the dental clinic; inability to monitor and influence actual OHBs in home settings; and failing to detect emerging non-adherence to prescribed OHBs and provide

adequate timely support. Ultimately, good oral health depends on an individual's ability and willingness to carry out self-care behaviors. Engaging individuals in their own care is vital to reducing disease burden, costs, and improving satisfaction with the care experience for both patients and clinicians [46].

The widespread availability of mobile devices and the increasing intersection with behavioral science have given rise to the use of digital approaches, known as digital health interventions (DHIs), for reinforcing health behaviors and self-care on a large scale [58, 68]. The advent of low-cost sensors has enabled unobtrusive monitoring, thereby providing an accurate representation of a patient's behaviors in real-world settings, thus enabling the use of digital health interventions (DHIs) on a large scale [58,68]. To leverage the potential of this pervasive technology, researchers have proposed a variety of models over the years to monitor activities of daily living (ADL), including categories of personal hygiene such as hand washing [27], flossing [5], mobility-related activities like climbing stairs, walking, and functional transfers such as standing up from a chair or lying down on a bed [12–14]. These activity recognition systems can be employed in the realm of behavioral change, such as smoking cessation [62, 62, 67] or cessation of obsessive eating [72]. While these methods have demonstrated accuracy and reliability for their intended purposes, the systems are often designed and optimized for specific applications. As a result, the techniques used to recognize some ADL may not be easily transferable to other domains, thus accounting for the diversity of approaches used for activity recognition.

The area of toothbrushing monitoring is a newly emerged field of study that presents unique challenges not encountered in recognizing activities such as walking, standing, and running. In contrast to these activities, which have easily observable sensor patterns, toothbrushing activities, such as brushing the inner and outer surfaces of teeth, may exhibit much more subtle differences in acceleration patterns captured by attached sensors. Therefore, accurate toothbrushing sensing poses significant challenges. [52]

Studies that have attempted to study monitor brushing have been conducted in controlled clinical settings (e.g. limited head and body movement, predefined sequence of brushing) and does not represent free-form brushing in naturalistic settings.

In this dissertation, our objective is to monitor and analyse brushing behavior in naturalistic

settings, utilizing motion sensors. We study the detection of the regions being brushed during a brushing session using motion-sensors attached to a brush. Additionally, we conduct a comprehensive analysis of brushing behaviors in at-home settings to elucidate both the similarities and variabilities observed between- and within- individuals during brushing.

Specifically, our contributions are as follows:

1. We delve into the challenges associated with brushing region detection and elucidate why achieving accurate position tracking of the brush within the mouth poses inherent difficulties. Furthermore, we explore the limitations of employing a physics-based approach solely relying on the 3D geometry of orientation to address this problem. Consequently, we discuss how to use machine learning techniques as a more viable solution.

2. We provide the first publicly available dataset of 9-axis motion-sensor (including accelerometer, gyroscope, and magnetometer) attached to the handle of the toothbrush. This dataset is unique because it includes both manual and electronic toothbrushes and involves free-form brushing sessions. Unlike previous studies, which only included brushing performed under constraints such as no head/body movement or a predefined sequence of brushing regions, our dataset offers a more realistic representation of brushing behavior. Also, we include two different facing direction during brushing in our dataset; a factor not considered in the previous studies or bypassed by modifying the brush as in [36].

3. We will address the challenges associated with labeling the brushing dataset and shed light on the reasons why the labels are susceptible to errors. To mitigate these challenges and improve the quality of the dataset, we propose a relabeling procedure which was able to correct up to 10% of the mislabeled data. The relabeling procedure aims to enhance the accuracy and reliability of the dataset annotations.

4. We propose a three-stage framework (i.e., pre-processing, brush transition times estimation, and time-series classification) to tackle the toothbrushing region detection problem. By applying our algorithm, both to our dataset as well as the dataset collected under the structured brushing assumption [37], we show that region detection during free-form brushing is more challenging than detection in controlled settings.

3

5. We show that a high classification accuracy can be achieved using a random train-test split (i.e. k-fold cross validation), but generalization beyond the participants in the training set (i.e. leave-one-subject-out cross-validation) poses difficulties. Our observation should guide other researchers in reporting the results of their brushing detection studies.

6. We compare different classification methods based on feature engineering and deep learning techniques. We apply the Transformer Encoder model to our dataset and show that additional customization can be done for different types of brushes (i.e. manual and electronic) and the attributes of the user (i.e. left-handed and right-handed) by introducing feature embedding technique used in natural language processing applications (NLP). We show that our proposed classifier can achieve a higher accuracy compared to the method proposed in [38] under all evaluation settings.

7. We present the first statistical model for analyzing brushing behavior, specifically focusing on the duration of brushing and the application of pressure on each dental region. Previous studies commonly relied on the Wilcoxon test [28], a non-parametric test, for statistical deductions. However, the applicability of this test is limited in the context of repeated measurement studies, such as our study where multiple subjects engage in multiple brushing sessions each. Moreover, the inability to utilize multiple measurements per subject would require incorporating hundreds of participants (e.g. in [81]), resulting in a costly and time-consuming study design. Therefore, we conducted a systematic analysis of brushing behavior for the first time using mixed-effect generalized linear regression models. Recognizing that many regions may be skipped during brushing, we advocate for the suitability of a zero-inflated model. The count submodels employ a log-link function and a negative binomial distribution to effectively account for count-based outcomes. On the other hand, the zero-inflation submodels utilize a logistic link function and a Bernoulli distribution to address the excess zeros observed in the data. By employing this model, we found that participants generally brushed their buccal (i.e., outer) tooth surfaces more than twice as long as the occlusal (i.e., flat) surfaces, with a factor of 2.18 times longer (95% CI 1.42, 3.35; $p \leq 0.001$). Similarly, participants spent 2.22 times longer (95% CI

4

1.62, 3.10; p ≤ 0.001) brushing their lingual (i.e., inner) surfaces compared to the occlusal surfaces. Interestingly, the lingual surfaces of the maxillary molars (i.e., upper jaw) were frequently neglected (p ≤ 0.001). Additionally, the occlusal surfaces were more prone to being brushed with excessive pressure (95% CI 0.10, 0.98; p = 0.015).

8. We propose a statistical analysis that encompasses both between-individual and within-individual variability in brushing behavior using our zero-inflated mixed-effect regression model. Previous studies, which relied on non-parametric tests like the Wilcoxon test, were inadequate in analyzing intra-subject variability. Our model incorporates fixed effects for tooth surface, mouth side, and jaw in both the zero-inflation and count submodels. To capture between- and within-participant variabilities, random effects were included on the intercept by session nested within each participant. This accounts for person-to-person and session-to-session differences in overall brushing duration. In the count submodel for region-specific brushing duration, participant-specific random effects were also included for tooth surface, mouth side, and jaw. Our analysis revealed that all dental regions that were brushed exhibited a coefficient of variation exceeding 20%, indicating significant between-individual variability in brushing duration for each dental surface. Additionally, the coefficients of between-individual and within-individual variability for total brushing duration were 0.19 and 0.17, respectively, highlighting substantial variability between and within individuals in brushing duration.

In Chapter 2, we provide the necessary background information to understand this dissertation, including the definition of brushing regions and methods for estimating object orientation using motion sensors. We also explain why absolute position tracking is not feasible in our application in Section 2.3. In Chapter 3 we discuss the challenges of the toothbrushing regions detection problem. We start by discussing previous research conducted on this problem in Section 3.1. We then explain our experimental setup for data collection in Section 3.2. In Subsection 3.2.2 we explore the challenges in labeling our dataset and propose a relabeling method to deal with the noisy labels. Section 3.4 discusses the challenges of our problem and why machine learning is necessary to address it accurately. In Chapter 4, we present our pro-

posed method for brushing region detection, starting with data preprocessing in Section 4.1, followed by the introduction of CluMing, a novel change point detection algorithm, in Subsection 4.2.1. Finally, in Section 4.3, we introduce the machine learning methods used for time-series classification. We demonstrate the results of our algorithm on our provided dataset and controlled toothbrushing datasets in Section 4.4. In Chapter 5, we analyze the brushing behaviors of individuals using zero-inflated mixed-effect generalized linear regression models. In particular, we analyze the total brushing time, time spent brushing each dental region, and the time brushed with excessive pressure on each dental region. We also demonstrate between- and within-individual variability in each category. We conclude and discuss potential future research directions in Chapter 6.

Chapter 2

Background and Definitions

In this chapter, we will provide some necessary background information to facilitate a better understanding of the problem we are attempting to solve, which will be used throughout this dissertation.

2.1   Toothbrushing Regions

There is no universally accepted standard for dividing the mouth into different brushing regions in the dental community. However, two common approaches are the use of four brushing regions known as quadrants, or six brushing regions known as sextants.

In this dissertation, we adopt the standard developed under the Bass brushing technique, as outlined in [49]. According to this approach, the mouth area is divided into 16 regions, as demonstrated in Figure 2.1. This is the same division that has been used by researchers in [35, 52] for the sake of consistency. It is worth noting, however, that the Bass technique includes the tongue as an additional region of the mouth, which we did not consider in our data collection due to participant reluctance.

There have been other studies that have used varying numbers of brushing regions. For instance, the authors of [80] divided the mouth into 18 regions by adding two extra regions to the 16 regions mentioned above. In [16], 24 brushing regions were proposed based on the 20 milk teeth that children have.
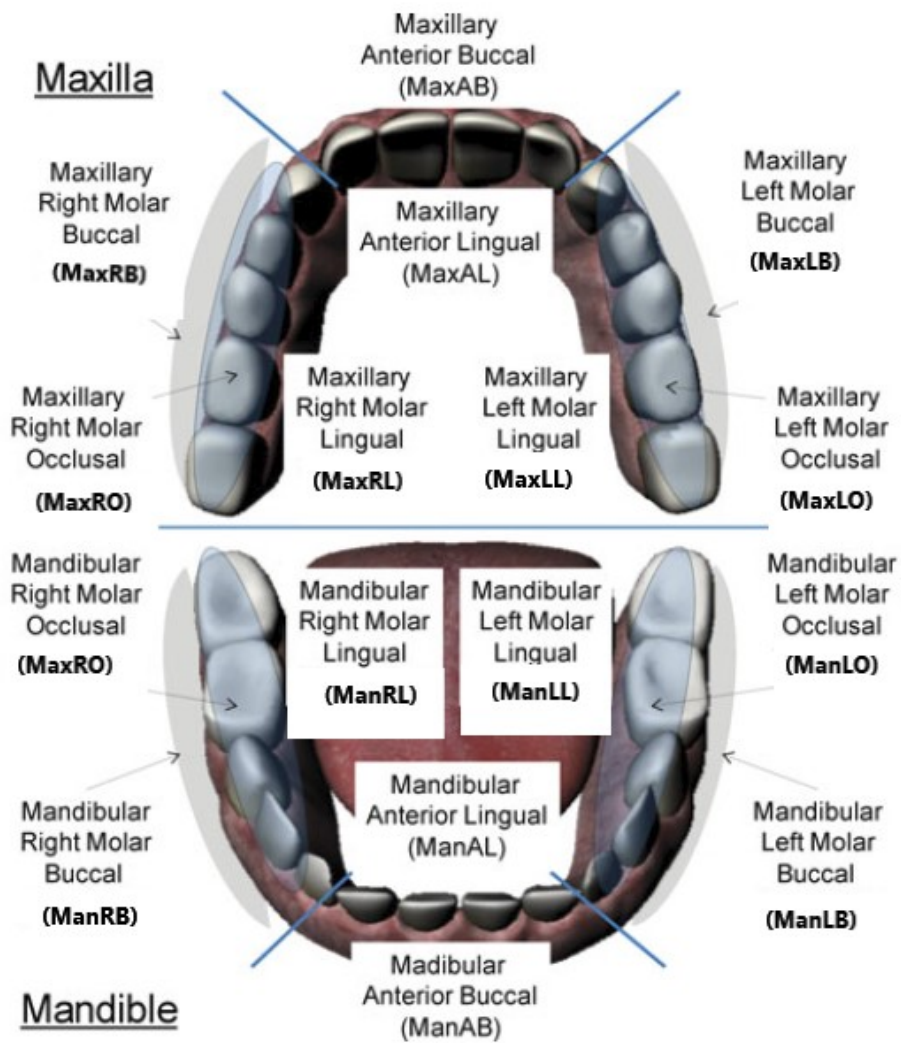
Figure 2.1: Defined 16 brushing regions of mouth [49]

2.1.1 Bass brushing technique

The Bass brushing technique is a widely accepted method for toothbrushing that involves positioning the bristles of the toothbrush at a 45-degree angle towards the gum line and making short, gentle back-and-forth strokes. The technique requires the user to gently vibrate the brush head in small circular motions while applying light pressure against the gums. The brushing movement should be performed for about 20 seconds on each tooth, focusing on a few teeth at a time before moving on to the next set. The method also involves gentle brushing of the tongue and the roof of the mouth to remove bacteria and freshen breath.

The Bass brushing technique has been shown to be effective in removing plaque and preventing gum disease when performed correctly. However, its effectiveness depends on proper positioning of the toothbrush bristles and consistent use over time.

2.2 Motionsensors and Orientation measurement

In this section, we will first introduce the various sensors commonly used for orientation estimation. Then, we will present several methods for representing the orientation of an object. Finally, we will suggest a few techniques for estimating the orientation and compare them by applying them to the data obtained by motion sensors during a brushing session.

In order to accurately detect the brushing regions, it is necessary to track the position of the toothbrush head within the mouth. Motion sensors can be employed for this purpose. However, there are challenges associated with that which we discuss in Subsection 2.3.

2.2.1 Motion sensors

Motion sensors are available in various types and forms, and can be compared based on several factors such as size, cost, energy consumption, response time, and sensitivity to external factors. In our study, we utilize micro-electro-mechanical system (MEMS) sensors, which are small and lightweight enough to be mounted on a brush grip. Additionally, they are affordable, have low power consumption, and have short start-up times [44]. Typically, motion sensors include one or more of the following sensors:

1. 3-axis accelerometer sensor, which measures the acceleration applied to the sensor along its x, y, and z axes. During brushing, an accelerometer attached to the brush can detect the acceleration of the brushing strokes and the effect of gravity along the sensor axes.

2. 3-axis gyroscope sensor, which measures the rotational velocity of the sensor around its x, y, and z axes. During a brushing session, the gyroscope signal has almost zero magnitude when the brush performs motions that do not involve much rotation, such as back and forth brushing strokes. However, it changes rapidly during transitions from one brushing region to another.

3. 3-axis magnetometer sensor, which measures the magnetic flux intensity at the sensor location along its x, y, and z axes. The magnetometer sensor is designed to measure the earth's magnetic field at the sensor location but is also affected by other ferromagnetic components nearby.

Motion sensors that include accelerometer and gyroscope sensors are referred to as inertial sensors or Inertial Measurement Units (IMUs). Those that have a magnetometer in addition to an accelerometer and gyroscope are called Magnetic, Angular Rate, Gravity (MARG) sensors.

2.2.2 Orientation Representation

The orientation of an object is measured with respect to a reference frame. As a result, the reference frame should be defined before computing orientation. In motion sensor applications, the orientation of the sensor frame, which is also called the body frame, is represented in the world reference frame. The world reference frame is defined as follows:

1. The X-axis of the world frame is in the direction of the north magnetic pole, which has a declination angle from the north pole.

2. The Z-axis of the world frame is in the opposite direction of gravity.

3. The Y-axis of the world frame is defined using the right-hand rule based on the X and Z axes of the world frame.

10

3D orientation is represented in three main ways:

1. Euler Angles (EAs) represent orientation with three angles known as roll or bank angle ϕ, pitch or elevation angle θ, and yaw or heading angle ψ. These angles measure the orientation angle with respect to the x, y, and z axes, respectively. However, there are multiple ways of incorporating EAs, such as by changing the order of the rotation around the x, y, and z axes, using body frame coordinates or world frame coordinates as the axes of rotation, and using active or passive rotations. In this dissertation, we use the Tait-Bryan EA convention, which means consecutive active rotations with respect to the world frame.

   We also use the convention of a leading subscript for the frame being described and a leading superscript for the frame that the orientation is referenced to. For example, ${}_{B}^{W}\psi$ is the heading angle of the brush in the world reference frame.

2. Quaternions are expressions with four constituents, one of which is a real number and the other three are imaginary components [32]. They can be interpreted as an arbitrary rotation by an arbitrary angle θ around an arbitrary axis of rotation r, as shown in Figure 2.2. They are computed using Equation 2.1.

$$
{}_{B}^{A}\hat{\boldsymbol{q}} = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 \end{bmatrix} = \begin{bmatrix} \cos\frac{\theta}{2} & -r_x\sin\frac{\theta}{2} & -r_y\sin\frac{\theta}{2} & -r_z\sin\frac{\theta}{2} \end{bmatrix} \tag{2.1}
$$

   A three-dimensional vector can be rotated by a quaternion using the relationship described in Equation 2.2. ${}^{A}\boldsymbol{v}$ and ${}^{B}\boldsymbol{v}$ are the same vector described in frame A and frame B, respectively, where each vector contains a 0 inserted as the first element to make them 4 element row vectors [53].

$$
{}^{B}\boldsymbol{v} = {}_{B}^{A}\hat{\boldsymbol{q}} \otimes {}^{A}\boldsymbol{v} \otimes {}_{B}^{A}\hat{\boldsymbol{q}}^{*} \tag{2.2}
$$
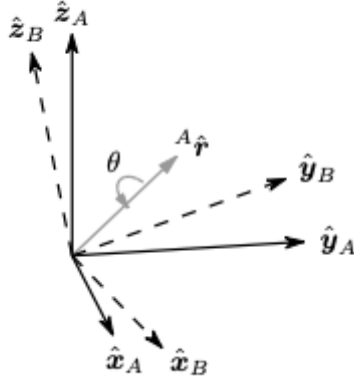
Figure 2.2: Axis-Angle representation of orientation.

2.2.3 Orientation Estimation

The process of orientation estimation entails the fusion of information from all motion sensors, a procedure known as sensor fusion, by means of an orientation filter. Relative estimation of orientation with respect to the initial orientation, can be obtained by integrating gyroscope values, which represent angular velocity. Nevertheless, this integration process can trigger drift in orientation estimation, caused by the accumulation of minor errors in gyroscope measurements, referred to as the gyroscope drift problem [53]. To establish an absolute reference, an accelerometer is utilized to provide a reference to the z-axis direction of the world reference frame. In the case of IMU sensors, both the accelerometer and gyroscope are used in tandem to determine the absolute roll angle ϕ, absolute pitch angle θ, and heading angle ψ relative to the sensor's initial orientation. However, to ascertain the absolute heading angle, a magnetometer sensor (such as that found in MARG sensors) is required to provide an absolute reference to the x-axis of the world frame.

To obtain an accurate estimation of orientation in relation to the world reference frame, it is imperative to use accelerometer and magnetometer sensors in conjunction. While the accelerometer can provide the roll ϕ and pitch angle θ, the magnetometer can furnish the heading angle ψ. However, to achieve a smoother estimation of orientation, it is necessary to fuse the gyroscope with the accelerometer and magnetometer in MARG sensors using orientation filters,

12

which are commonly referred to as complementary filters. This approach leverages the faster dynamic response of the gyroscope and leads to a more refined orientation estimation.

A multitude of methods exist for orientation estimation, ranging from Kalman-based approaches [26, 51, 55] to complementary filters [23, 54]. However, Kalman-based methods can be complex, require extensive parameter tuning, and necessitate a high sampling rate, often in the order of kHz [53].

In 2010, Madgwick et al. [53] introduced a method for orientation estimation that is applicable to both IMU and MARG sensors. In our study, we employed this filter and compared its performance to another complementary filter proposed by Mahony et al. [54], which has similar objectives. We also evaluated the results against the Cho filter, a method proposed by Cho et al. [17], which we refer to as such in this dissertation.

Madgwick filter

The orientation estimation filter we employed is a lightweight computational model based on a complementary fusion algorithm that can be implemented with just a couple of hundred arithmetic operations. The fusion algorithm combines the orientation representation obtained from the gyroscope with the one derived from the accelerometer and magnetometer, resulting in a more accurate orientation estimation.

Orientation estimation from gyroscope

It is shown in [18] that the quaternion derivative describing the rate of change of orientation of the earth frame relative to the sensor frame ${}^{S}_{E}\dot{\boldsymbol{q}}_{\omega,t}$ can be calculated from the gyroscope measurements $[\omega_x \ \omega_y \ \omega_z]$ as follows:

$$
S\boldsymbol{\omega} = \begin{bmatrix} 0 & \omega_x & \omega_y & \omega_z \end{bmatrix}
$$
$$
{}^{S}_{E}\dot{\boldsymbol{q}} = \tfrac{1}{2}{}^{S}_{E}\hat{\boldsymbol{q}} \otimes {}^{S}\boldsymbol{\omega}
$$

(2.3)

By integrating these measurements at time t, given the sample period Δt and the orientation at time $t-1$, we can compute the orientation at time t as follows:

$$
{}_E^S\boldsymbol{q}_{\omega,t} = {}_E^S\hat{\boldsymbol{q}}_{est,t-1} + {}_E^S\dot{\boldsymbol{q}}_{\omega,t}\Delta t \tag{2.4}
$$

Here, the subscript ω indicates that the orientation computed from gyroscope measurements.

Orientation estimation from accelerometer and/or magnetometer

If the direction of an earth's field (e.g. gravitational or magnetic) is known in the earth frame, a measurement of the field's direction within the sensor frame will allow an orientation of the sensor frame relative to the earth frame to be calculated. This may be achieved through the formulation of an optimization problem where the orientation of the sensor, ${}_E^S\hat{\boldsymbol{q}}$, is that which aligns a predefined reference direction of the field in the earth frame, ${}^E\hat{\boldsymbol{d}}$, with the measured direction of the field in the sensor frame, ${}^S\hat{\boldsymbol{s}}$, using the rotation operation described by the equation 2.2 as follows [53]:

$$
\min_{S_\varepsilon \in T^4} \boldsymbol{f}\left({}_E^S\hat{\boldsymbol{q}}, {}^E\hat{\boldsymbol{d}}, {}^S\hat{\boldsymbol{s}}\right)
$$
$$
\boldsymbol{f}\left({}_E^S\hat{\boldsymbol{q}}, {}^E\hat{\boldsymbol{d}}, {}^S\hat{\boldsymbol{s}}\right) = {}_E^S\hat{\boldsymbol{q}}^* \otimes {}^E\hat{\boldsymbol{d}} \otimes {}_E^S\hat{\boldsymbol{q}} - {}^S\hat{\boldsymbol{s}}
$$
$$
{}_E^S\hat{\boldsymbol{q}} = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 \end{bmatrix} \tag{2.5}
$$
$$
{}^E\hat{\boldsymbol{d}} = \begin{bmatrix} 0 & d_x & d_y & d_z \end{bmatrix}
$$
$$
{}^S\hat{\boldsymbol{s}} = \begin{bmatrix} 0 & s_x & s_y & s_z \end{bmatrix}
$$

which can be solved with gradient descent as follows:

$$
{}_E^S\boldsymbol{q}_{k+1} = {}_E^S\hat{\boldsymbol{q}}_k - \mu \frac{\nabla \boldsymbol{f}\left({}_E^S\hat{\boldsymbol{q}}_k, {}^E\hat{\boldsymbol{d}}, {}^S\hat{\boldsymbol{s}}\right)}{\left\|\nabla \boldsymbol{f}\left({}_E^S\hat{\boldsymbol{q}}_k, {}^E\hat{\boldsymbol{d}}, {}^S\hat{\boldsymbol{s}}\right)\right\|} \quad , k = 0, 1, 2 \ldots n \tag{2.6}
$$

Assuming the direction of gravity as z-axis of the earth frame, we can write:

$$
{}^E\hat{\boldsymbol{g}} = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}
$$
$$
{}^S\hat{\boldsymbol{a}} = \begin{bmatrix} 0 & a_x & a_y & a_z \end{bmatrix} \tag{2.7}
$$

14

Also, the earth's magnetic field can be considered to have components in one horizontal axis and the vertical axis considering the inclination of it.

$$
\begin{aligned}
{}^E\hat{\boldsymbol{b}} &= \left[\begin{array}{cccc} 0 & b_x & 0 & b_z \end{array}\right] \\
{}^S\hat{\boldsymbol{m}} &= \left[\begin{array}{cccc} 0 & m_x & m_y & m_z \end{array}\right]
\end{aligned}
\tag{2.8}
$$

In order to use both sensor measurements, we can stack them as a larger vector:

$$
\boldsymbol{f}_{g,b}\left({}^S_E\hat{\boldsymbol{q}}, {}^S\hat{\boldsymbol{a}}, {}^E\hat{\boldsymbol{b}}, {}^S\hat{\boldsymbol{m}}\right) = \left[\begin{array}{c} \boldsymbol{f}_g\left({}^S_E\hat{\boldsymbol{q}}, {}^E\hat{\boldsymbol{g}}, {}^S\hat{\boldsymbol{a}}\right) \\ \boldsymbol{f}_b\left({}^S_E\hat{\boldsymbol{q}}, {}^E\hat{\boldsymbol{b}}, {}^S\hat{\boldsymbol{m}}\right) \end{array}\right]
\tag{2.9}
$$

Filter fusion

The final orientation estimation can be obtained by a linear combination of the orientation estimations from the equations 2.6 and 2.3 as follows:

$$
{}^S_E\,\boldsymbol{q}_{est,t} = \gamma_t {}^S_E\boldsymbol{q}_{\nabla,t} + (1-\gamma_t){}^S_E\,\boldsymbol{q}_{\omega,t} \quad ,0 \leq \gamma_t \leq 1
\tag{2.10}
$$

By setting the value of γ_t as that which ensures the weighted divergence of ${}^S_E\boldsymbol{q}_\omega$ is equal to the weighted convergence of ${}^S_E\boldsymbol{q}\nabla$, and some further simplifications, the following final estimation is obtained:

$$
\begin{aligned}
{}^S_E\boldsymbol{q}_{est,t} &= {}^S_E\,\hat{\boldsymbol{q}}_{est,t-1} + {}^S_E\,\dot{\boldsymbol{q}}_{est,t}\Delta t \\
{}^S_E\,\dot{\boldsymbol{q}}_{est,t} &= {}^S_E\,\dot{\boldsymbol{q}}_{\omega,t} - \beta {}^S_E\dot{\hat{\boldsymbol{q}}}_{\epsilon,t} \\
{}^S_E\,\dot{\hat{\boldsymbol{q}}}_{\epsilon,t} &= \frac{\nabla\boldsymbol{f}}{\|\nabla\boldsymbol{f}\|}
\end{aligned}
\tag{2.11}
$$

It can be seen from these equations that the filter calculates the orientation ${}^S_E\boldsymbol{q}_{\text{est},t}$ by numerically integrating the estimated orientation rate ${}^S_E\dot{\boldsymbol{q}}_{est}$. The filter computes ${}^S_E\dot{\boldsymbol{q}}_{est}$ as the rate of change of orientation measured by the gyroscopes, ${}^S_E\dot{\boldsymbol{q}}_\omega$, with the magnitude of the gyroscope measurement error, β, removed in the direction of the estimated error, ${}^S_E\dot{\hat{\boldsymbol{q}}}_\epsilon$, computed from accelerometer and magnetometer measurements.

As an extra element of their proposed filter, they suggest a method to compensate for
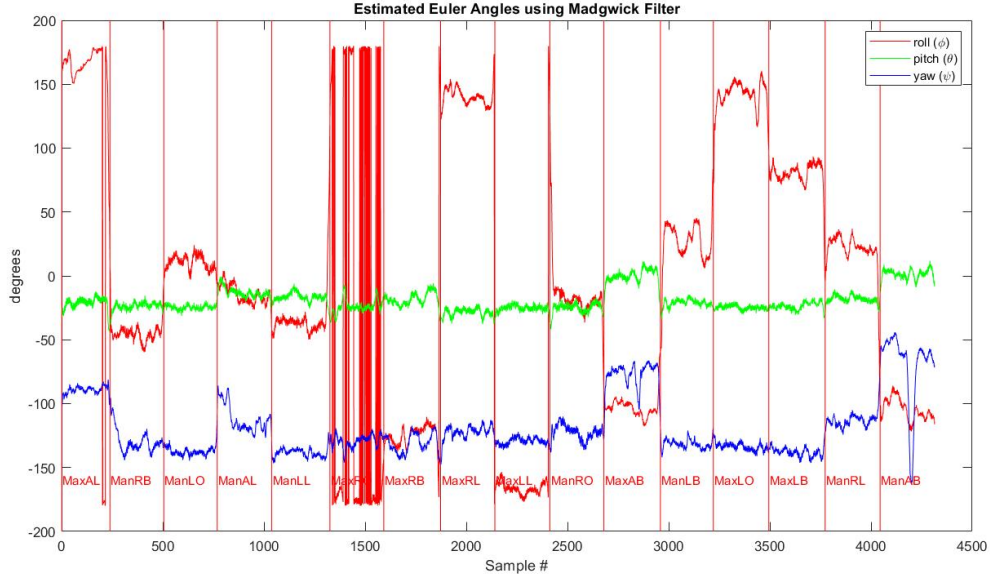
Figure 2.3: Estimated Euler Angles in a brushing session using Madgwick filter

magnetic distortions and also gyroscope drift, which is only applicable to MARG sensors.

Cho filter

They use an accelerometer to calculate the roll and pitch angle θ and they use a magnetometer to derive the yaw angle ψ by using the earth's magnetic field inclination angle or dip angle. They propose a method to find the dip angle and they use that to calibrate magnetometer values (see Subsection 4.1.3 for calibration).

Comparison between Orientation Filters

Below, we consider the motion data that was captured by the MARG sensor attached to the brush in a brushing session with a manual toothbrush and we compare the EAs that were estimated using Madgwick, Mahony, and Cho filters.

As observed in Figures 2.3, 2.4, and 2.5, the estimated EAs from all three methods exhibit remarkable similarity, and they all demonstrate low computational complexity. The run times for Madgwick, Mahony, and Cho filters are 0.056, 0.063, and 0.002 milliseconds per sample,

16
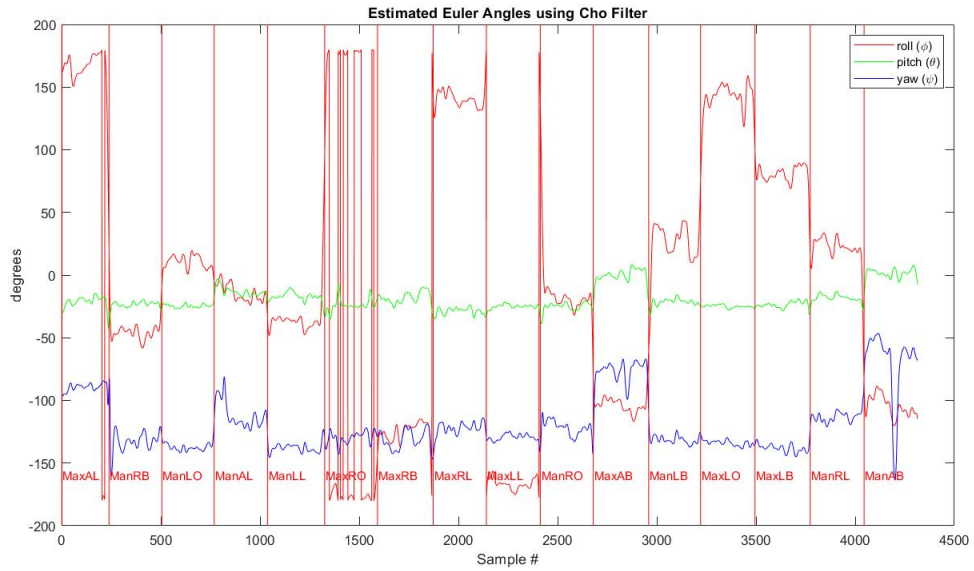
Figure 2.4: Estimated Euler Angles in a brushing session using Cho filter
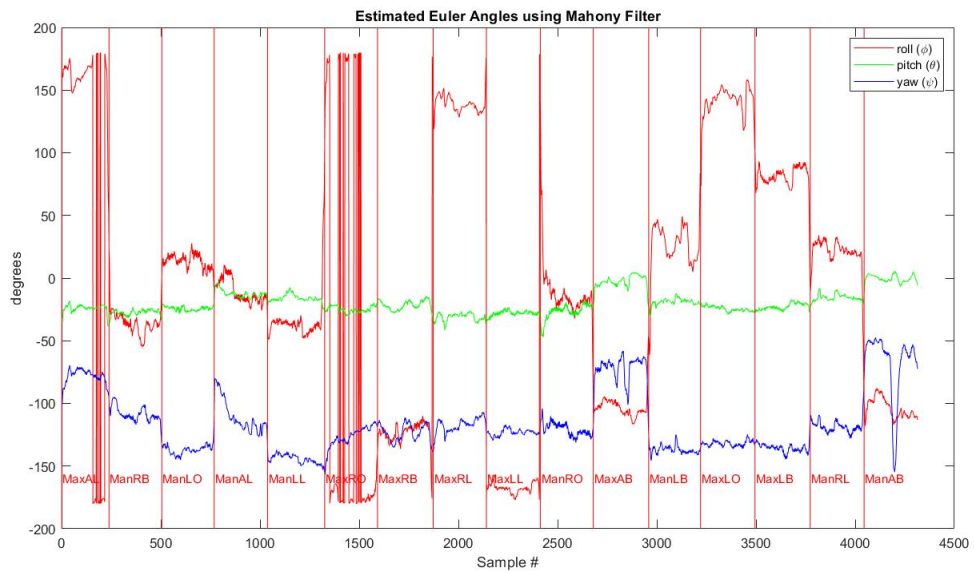
Figure 2.5: Estimated Euler Angles in a brushing session using Mahony filter

17

respectively, on an Intel(R) $Core^{(TM)}$ i7-6700T CPU @ 2.8 GHz. However, it should be noted that the Cho filter lacks the ability to measure the ψ angle when magnetometer measurements are unavailable. In contrast, both the Madgwick and Mahony filters can integrate gyroscope measurements to capture changes in the ψ angle. Therefore, for orientation estimation in this dissertation, we rely on the Madgwick filter.

In order to derive these EA estimations we used the magnetometer calibration method described in Subsection 4.1.3. However, for Cho filter we can use its own magnetometer calibration technique (min-max normalization), which would generate the results shown in Figure 2.6. A comparison of the estimated angles reveals that the ϕ and θ angles are consistent between Figures 2.4 and 2.6 since they are derived from the accelerometer, which is the same in both cases. However the ψ angle is differs because it is calculated based on the magnetometer measurements, which are calibrated differently.

Figure 2.4 demonstrates a more accurate estimation of the ψ angle compared to Figure 2.6. In Figure 2.6, the ψ angle estimated for the MaxRB and ManRB regions diverges significantly, and the gap between them even encompasses the ManLB and MaxLB regions. However, in reality, the ManRB and MaxRB regions should exhibit similar ψ angles due to their proximity in brushing orientation, as depicted in Figure 2.4. Moreover, Figure 2.4 indicates that the MaxAB and ManAB regions have an angular difference of approximately 70 degrees with respect to the ManRB, MaxRB, ManLB, and MaxLB regions on average. This aligns more closely with a realistic scenario compared to the approximately 100 degrees difference depicted in Figure 2.6. These findings have been validated through additional experiments involving measured orientations.

2.3 Position Tracking

Challenges associated with accurate position tracking of the brush in the mouth is as follows:

1. Hardware limitations for sensing and transmitting motion signals: Due to the rapid tooth-brushing movements and quick transitions of the toothbrush head between different brushing regions, accurate estimation of the toothbrush head position requires a high sampling
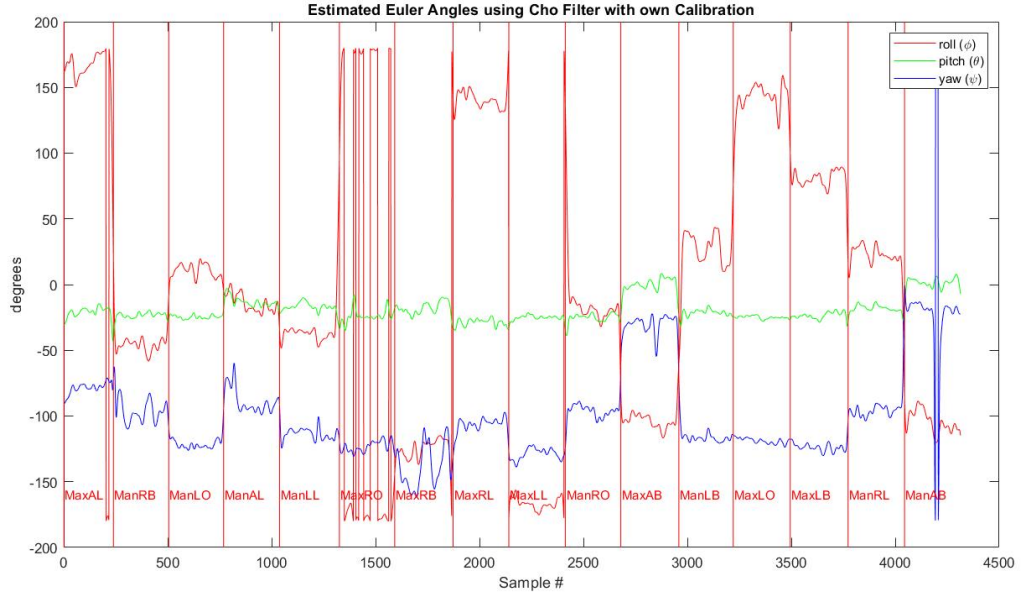
Figure 2.6: Estimated Euler Angles in a brushing session using Cho filter and Cho calibration method

rate. However, motion sensors are currently limited in their ability to provide such a high sampling rate. Additionally, the transmission of large amounts of motion sensor data, which typically occurs via Bluetooth, is not possible due to communication limitations.

2. Size of mouth: The mouth is a relatively small cavity that is divided into 16 different regions in our study. Precise measurements of the toothbrush head position within the mouth are required for accurate tracking. However, motion-sensors do not currently possess the level of precision required by the geometric dimensions of the mouth. In particular, to estimate the position of the toothbrush in mouth, we should do a double discrete integration of the gravity compensated accelerometer vector defined as follows:

$$A_{GC} = R_B^W \times A - \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix}. \tag{2.12}$$

Here, g is the gravitational acceleration which is approximately 9.8 m/s^2, A is the three-

19

dimensional acceleration vector measured by the accelerometer, and R_B^W is the rotation matrix representing the orientation of the brush with respect to the world reference frame. The world reference frame is a coordinate frame with its X-axis in the direction of the north magnetic pole, its Z-axis in the opposite direction of the gravity, and its Y-axis in the direction of the outer product $Z \times X$.

The double integration of A_{GC} can amplify any measurement errors in the accelerometer values. More importantly, the errors that exist in estimating the rotation matrix R_B^W can introduce too much error in location estimation. This makes location estimation using motion-sensors not a suitable approach in our application.

Figure 2.7 shows the location of the toothbrush estimated from the motion sensors attached to it while it was left idle on a table. As can be seen, the location of the brush changes about 19 meters in two seconds. Even when ignoring the z-axis position change (since it might not be valuable for estimating the brush location in the mouth, provided that the dental surfaces are located at similar height approximately), the brush position changes about 1 meter both along the X and Y directions. This value is higher than the almost 4cm dimensions of the mouth in the X, Y, and Z directions, making precise tracking of the brush position challenging in such a small space.

To obtain a more accurate estimation of the orientation of the brush, one can consider using a low-pass filter to remove the noise appearing as high-pass components of the accelerometer before calculating R_B^W in Equation 2.12. This approach may result in less noise being involved during the double integration of the accelerometer. Additionally, after each integration, it may be beneficial to apply band-pass filtering to the velocity and position estimation obtained to obtain a better estimation of the position considering that the integrations may accumulate noise drifts as biases in the measurements.

In our experiment, we used a Butterworth filter of order 3 with cutoff frequencies of 0.5 and 4 Hz as a band-pass filter, and a low-pass filter of order 3 with a cutoff frequency of 0.5 Hz. This resulted in a position estimation shown in Figure 2.8. As can be seen, this estimation yielded an error of about 12 cm along the X and Y directions, which is almost triple the size of the mouth along each of those dimensions. Also, the filters side lobes (in the time-domain)
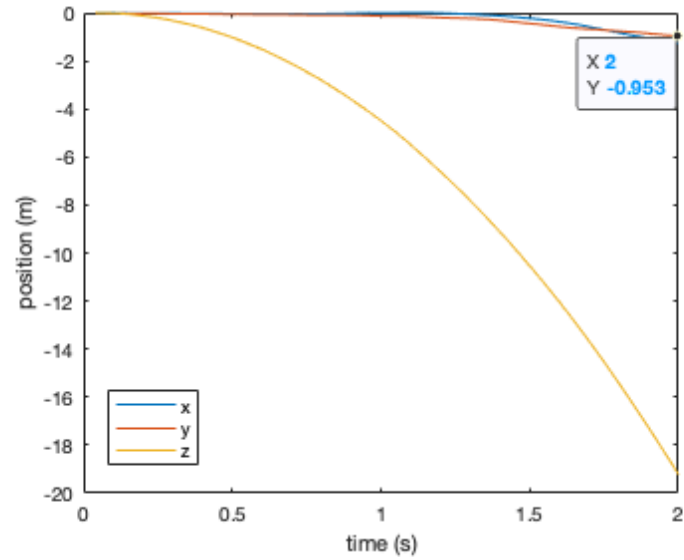
Figure 2.7: Position of the toothbrush calculated from Equation 2.12
without low-pass filtering of acceleremoter

would appear as oscillations in the output signal (estimated position).

Therefore, the focus of the study is on estimating the orientation of the brush head using
motion-sensors to detect the brushing regions.
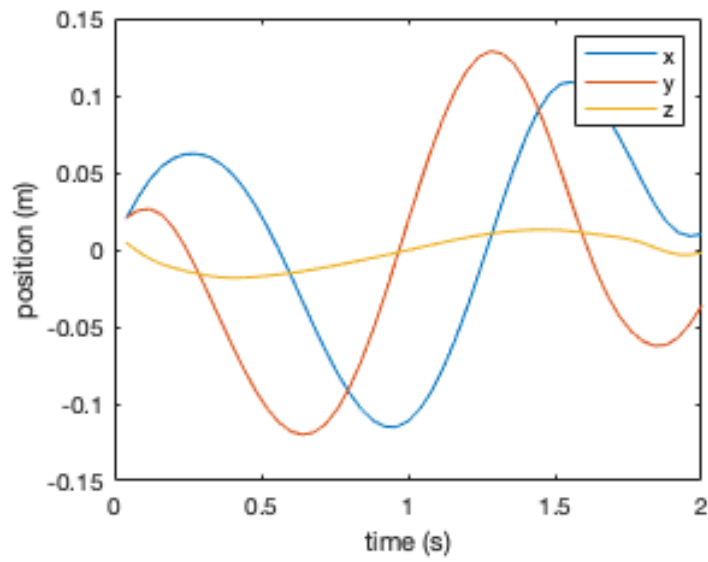
Figure 2.8: Position of the toothbrush calculated from Equation 2.12
with low-pass filtering of acceleremoter, and band pass filtering velocity and position

Chapter 3

Toothbrushing Region Detection - Data Collection and Analysis

In this chapter, we delve into the related work on brushing region detection, elucidate our data collection process, discuss the challenges encountered during dataset labeling, present our proposed method for improving the accuracy of data labels, and delve into the challenges associated with brushing region detection.

3.1 Related Work

In this section, we will introduce the methods that have been proposed for brushing region detection. These methods employ external devices such as wrist-worn sensors or sensors attached to the brush.

3.1.1 ToothBrushing monitoring using wrist-worn sensors

In 2016, Huang et al. [35] introduced a system for detecting brushing regions using wrist-watch embedded Magnetic, Angular Rate, Gravity (MARG) sensors, specifically for manual toothbrushing. The data collected from the sensors was sent to a cloud system for storage and further processing. The authors attached multiple small magnets to the toothbrush grip to modify it, claiming that this modification was necessary for the algorithm to work effectively against rolling motions of the brush in the user's hand. In this scenario, the rotational movements of the brush in the user's hand, which would have no effect on the motion signals sensed

by the user's wristwatch (as the wrist remains unchanged in orientation), would be detected by the motion sensor in the wristwatch. The authors relied on the change in the magnetic field sensed at the wrist-worn sensor location caused by the added magnets to detect the brushing regions. Additionally, they used the wristwatch microphone to record the sound of the brushing strokes to count the number of brush strokes. The study lasted for 3 weeks and involved 12 participants. It was conducted in two phases: In the first phase, the participants were trained on the Bass brushing technique (refer to Section 2) for some period of time, and in the second phase, they were asked to brush using the learned Bass brushing technique.

A machine learning algorithm was proposed based on time and frequency domain features extracted from the accelerometer and magnetometer signals, achieving an accuracy of 85.6% in detecting 15 brushing regions using the Naive Bayes classifier. The authors later showed that using a two-layer shallow neural network, they could achieve an accuracy of 91.2% [36]. The authors combined MaxAB and ManAB brushing regions due to similar brushing patterns (refer to Section 3.3). Although these studies by Huang et al. exhibited superior performance compared to previous research, it is important to note that their instruction to participants was solely focused on the Bass-brushing technique. Furthermore, participants were asked to limit their head and body movements, which hinders the scalability of their proposed model. The algorithm can only be applied to detecting brushing regions in sessions conducted under the specific technique. Therefore, their model lacks generalizability to real-world scenarios, where individuals may move their head and body in various ways and hold their brushes at different angles that may not conform to the recommended 45 degrees angle for the Bass-brushing method. This limitation makes their algorithm less applicable to free-style brushing sessions and real-world scenarios. Furthermore, the authors used the transition matrix between consecutive brushing regions on their tested population of 12 subjects to gain an extra boost in the algorithm's performance, which is not applicable to the general population of brush users. It is also unclear whether their method is applicable to the use of electronic toothbrushes.

Another study that employed a similar setup is presented in [52]. In this study, researchers conducted a month-long investigation with ten subjects using brushes equipped with IMU sensors. Unlike the previous study, they did not modify the toothbrush or use audio signal recording

in their setup, and they only used accelerometer signals. They claimed to have eliminated the differences in signals obtained from sensors attached to manual and electronic toothbrushes using a low-pass filter, but there is no explanation of the amount of data obtained from sensors attached to either of the brushes. Additionally, the Bass technique was taught to the participants, and they were asked to brush with that restriction. The researchers demonstrated that using an attention-based LSTM network [78], they could achieve a 97.3% accuracy on 15 brushing activities using raw accelerometer data after applying a rotation matrix. They did not include MaxAB and ManAB in their categories and added "raising hand to brush" as an extra brushing activity while replacing ManAL with brushing the tongue (refer to Subsection 2.1). Furthermore, they claimed to classify brushing activity from other ADLs such as running, walking, and dining, using a dataset containing three minutes of doing those activities by each user during the study. However, it is not clear how big their dataset was. Their data was collected by restricting users to brush according to the Bass technique, which, similar to the previous study, limits the generalizability of their method. Additionally, they reported that signal annotation was performed manually by the researchers, which raises concerns about the usefulness of this method of annotating data in real-life brushing scenarios, considering the fast brushing behavior of users in transitioning the brush between different brushing regions in a short period of time. Moreover, the results reported use a random train-test split of the data (i.e. k-fold cross validation) and the study does not consider the generalization of their classification algorithm beyond the participants in their training set.

3.1.2 ToothBrushing monitoring using sensors attached to the toothbrushes

The authors in [47–49] have proposed a method that involves attaching a MARG sensor to manual toothbrushes to detect 16 brushing regions. In a series of studies, they first tested the feasibility of tooth brushing monitoring and brushing pattern classification algorithm schemes in [47] and [48] using a limited dataset that they had collected. While they showed proof of concept and how different orientations are represented in the accelerometer and magnetometer signals collected, they encountered difficulties in distinguishing between left and right brushing

regions due to inaccurate estimation of the heading angle ψ of the brush (refer to Section 3.3). To overcome this challenge, they planned to use magnetometer data with proper calibration in their later studies. In [49], they claimed to have solved the heading angle problem provided that the heading angle ψ of the front mirror is given and users remain very still. In their study, they included 15 subjects (13 right-handed and 2 left-handed) and collected about 3 minutes of brushing time in total from each participant using only accelerometer and magnetometer sensors. They proposed a method that involves classifying the posture of the brush into one of four categories (up, down, left, or right) based on the rolling angle ϕ, and then performing K-means clustering [41] on a stack of accelerometer and magnetometer data and Euler angles (EA) as feature vectors to identify the brushing region. They reported a classification accuracy of 97.1% on 15 brushing regions (where ManAB and MaxAB were considered as one region) (refer to Subsection 2.1). While they reported high accuracy, they restricted participants from any head or body movements and assumed that the heading angle ψ of the front mirror was measured accurately. They also asked participants to follow a predefined sequence of brushing regions and spend five seconds on each region, and to use the Bass brushing technique, which uses rolling strokes with specific motion on the side teeth, in developing their proposed algorithm. These assumptions generated a highly biased dataset and depended on factors that do not exist in real-life brushing scenarios. Hence, their dataset and results have limited generalizability. In addition, the reported results are based on a random division of the data into training and testing sets using k-fold cross-validation. However, the generalizability of their classification algorithm beyond the individuals included in the training set has not been taken into account.

Recently, Hussain et al. [37, 38] published a dataset collected from accelerometer and gyroscope sensors attached to the brush handle during toothbrushing. They used 17 participants in their study, each brushing their teeth for one to five brushing sessions for a total of 64 brushing sessions. Three participants used electronic toothbrushes while the rest used manual brushes. They calculated pitch and roll angles using a complementary filter [63] and used a random forest classifier to detect the brushing regions. The participants were instructed to brush each dental region for approximately seven seconds, following the Bass brushing technique, and to tackle the teeth surfaces in a specified order. However, these instructions impose limitations

on the generalizability of their method to brush users beyond those included in their dataset. Furthermore, the restrictions on head and body movements, as well as the predefined sequences of brushing, render their study not applicable to real-world scenarios of free-form brushing.

3.1.3 ToothBrushing monitoring using sound

The authors in [45] have proposed a method that utilizes sound data recorded from the user's phone to capture toothbrushing sound data. They extracted spectrograms from the audio signals and applied a Hidden Markov Model (HMM) to classify 4 brushing regions. In contrast, the researchers in [60] utilized an innovative approach by capturing audio signals of a user's brushing strokes from an earphone equipped on the neck and ear of the user. They applied classifiers on the feature set obtained from the spectrogram of the signals and achieved a notable 85.69% accuracy in classifying seven brushing strokes. However, it should be noted that in their study, they only performed fine (gentle) brushing strokes for some dental regions and rough brushing strokes for other regions. Moreover, their approach was unable to differentiate between strokes on the upper or lower sides of the mouth, as well as the left or right sides of the mouth.

3.2 Data Collection

3.2.1 Experimental Setup

While there are multiple studies conducted on toothbrush monitoring, there is no publicly available dataset that we could use for real-world conditions. Hence, we designed a study that includes 12 participants (2 left handed and 10 right handed) brushing their teeth for 187 brushing sessions total where each brushing session includes a random combination of the following factors:

1. Brush type: In our study, we have included both manual and electric toothbrushes as the brushing style with these two types of brushes differs. Brushing with an electric toothbrush involves slower and smoother brushing strokes, whereas brushing with a manual

toothbrush involves faster and stronger brushing strokes. Additionally, transitions from one brushing region to another in electric toothbrushing include more dragging motions of the brush on the teeth surfaces of each jaw, compared to the lifting motion that occurs in manual toothbrushing. While previous studies have focused only on manual toothbrushes, [52] claimed that their algorithm is robust to brush type. However, they did not investigate the differences in brushing behavior using the two types of brushes.

2. Direction: We arranged two mirrors at a 90-degree angle from each other to enable participants to brush in front of them in different directions. Our dataset will examine the impact of face direction on the magnetometer, considering its effect. This factor occurs naturally because users face different directions while brushing in different locations. In [35], the authors claimed that facing different directions does not affect their algorithm, as the brush grip has strong magnets attached to it, which is at the cost of modifying the manual toothbrush.

3. Brushing Method: Brushing typically involves quick transitions of the brush between different regions, and the brush stays in one region for a very short period of time. This makes the process of labeling data difficult and time-consuming. Furthermore, some brushing sessions are short, which is not advantageous for Machine Learning algorithms that require a large amount of labeled data. Hence, our dataset provides two brushing methods:

    (a) Scripted: Participants should follow a random sequence of 16 brushing regions (without any repetitive region) from a slideshow projected on a screen beside the mirror that they use for brushing. Each slide displays a figure of the mouth with the specific region pointed at for 10 seconds. To prepare the participants for the next brushing region, the next brushing region is displayed on the side of the screen in the last 3 seconds of each slide.

    (b) Freestyle: Participants are allowed to brush freely as they do on their own. Including freestyle brushing in our dataset is essential due to the variations in individual brushing styles and the need to develop a brushing region detection algorithm that

28
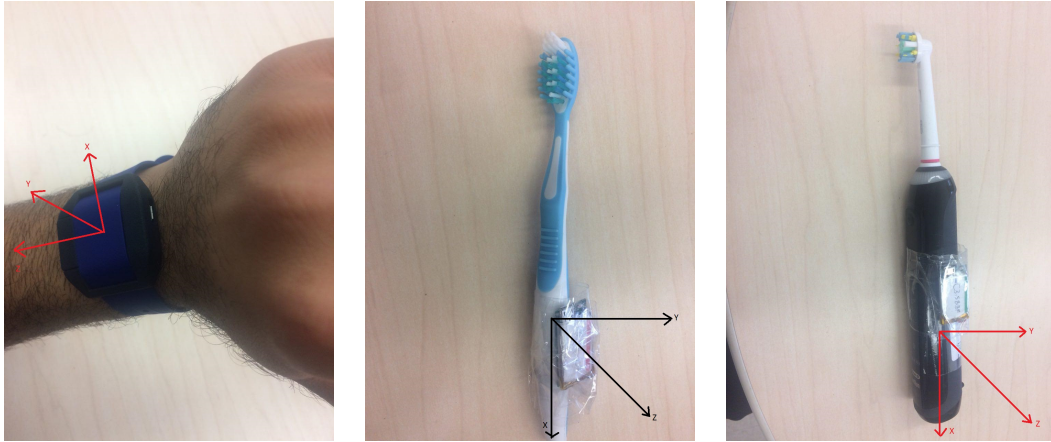
Figure 3.1: The sensors used in the study and their coordinates

can work for real-life brushing scenarios, unlike other datasets collected in previous studies [35, 49, 52].

All participants provided written informed consent and the study protocol was reviewed and approved by the Institutional Review Board of the University of California, Los Angeles (IRB#18–000874).

In our study, we used two MARG sensors. The first sensor was attached to the participants' toothbrush, and the second sensor was embedded in the wristband that participants wore during brushing sessions.

All sensors in our study collected data at a 25Hz sampling rate and transferred their data to a nearby phone via Bluetooth. We made an effort to replicate the exact attachment settings for attaching the sensors to all the brushes used in our study, including the same location on the brush and the same orientations. The figures in 3.1 demonstrate the axes of the brushes, and the wristband sensors. These axes were defined in a manner that avoids the gimbal lock problem when representing EAs, as described in [20].

To establish ground truth, we deployed two cameras to record the participants' faces from bottom and top angles. This was done to capture the brushing region in case their hand obscures one camera's angle of sight.

29

3.2.2   Labeling

In order to label the signal, we undertook a preprocessing step to synchronize the signals from both sensors with the accompanying video. To accomplish this, we had to estimate the signal values precisely at the times the Bluetooth packets were transmitted. Unfortunately, we only had access to the timestamps for when the Bluetooth packets were received by the phone, which could deviate by a few milliseconds. Furthermore, certain packets were lost during transmission. Consequently, we spline interpolated the signals by using the timestamps of the Bluetooth packets that were received by the phone to synchronize both signals.

Labeling toothbrushing regions can be a challenging task due to the following reasons:

1. Frequent brush transitions: During toothbrushing, the brush transitions frequently between different regions, with only brief stays at each region that may last only half a second. This makes the labeling process time-consuming and prone to inaccuracies.

2. Obscured brushing regions: The brushing regions can often become obscured by the participant's hand or the brush, even though we employed two cameras to minimize this effect.

3. Brushing of neighboring brushing regions together: Occasionally, participants may brush two neighboring brushing regions together, such as (ManLB, MaxLB), (ManRB, MaxRB), or (ManAB, MaxAB).

Due to the aforementioned challenges in labeling, the resulting labels may contain noise and cannot be entirely relied upon. Therefore, we implemented a two-step procedure to enhance the labeling as illustrated in Figure 3.2, which involved:

1. In the first step, we diligently endeavored to label our dataset using the video data and subsequently applied our proposed algorithm, as described in Chapter 4, to train a classifier based on these initial labels. To accomplish the primary labeling step, we assigned labels to each epoch of toothbrushing that lasted for more than 0.5 seconds after synchronizing the videos with the signals captured by the two sensors. Specifically, we labeled
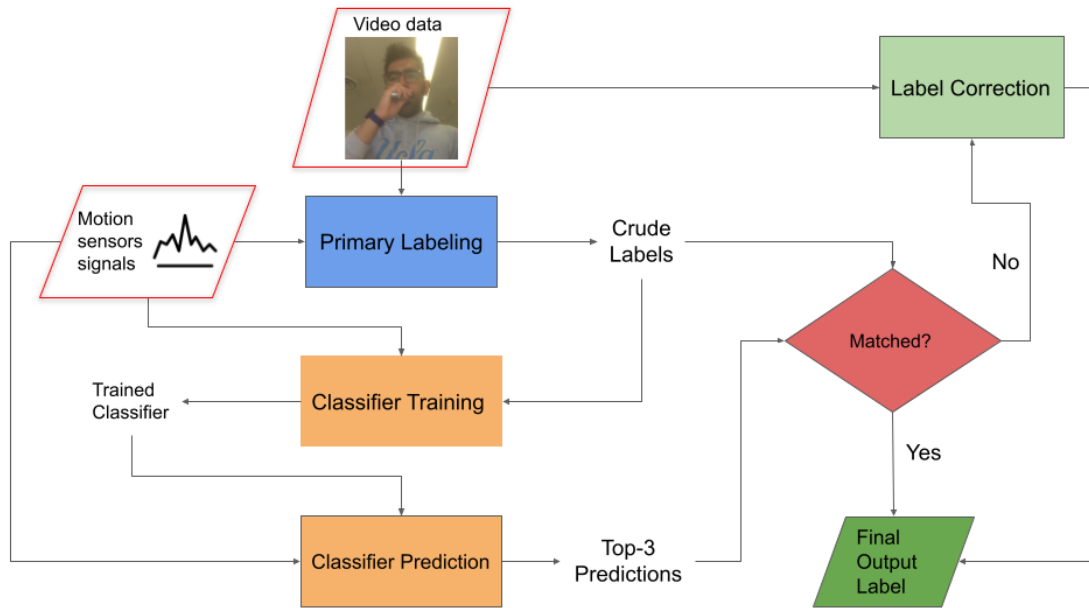
Figure 3.2: Labeling Procedure. The implemented labeling procedure involves two steps. Firstly, labels are generated manually utilizing video recordings of the brushing sessions. Secondly, the labels are corrected based on the top-three predictions of a classifier trained based on the labels output from the first step.

the starting and ending samples for each brushing region within the brushing session, excluding the periods when the brush was transitioning between brushing regions or when the user was not actively brushing any region.

2. In the second step, we employed our trained classifier to predict the brushing regions and identified regions that did not rank among the top three predictions made by the classifier. Subsequently, we rectified these regions using the brushing video.

Through this method, we were able to correct roughly 10% of the labels produced in the first labeling step.

In this dissertation, our focus is solely on the detection of brushing regions using the toothbrush sensor. However, we do plan to explore the potential of utilizing the wristband sensor as a future research direction.

Figure 3.3: Two brushing regions that have similar associated brush orientation [49]

3.3 Challenges of Toothbrushing Region Detection

The challenges in using orientation of the toothbrush head to detect the brushing regions are as follows:

1. Uniquely mapping toothbrushing orientations to brushing regions: As discussed in Section 2.2.3, tracking the position of the toothbrush is currently not feasible. Therefore, we aim to detect brushing regions using the orientation of the toothbrush head. However, mapping toothbrush orientation to the 16 defined brushing regions is a challenging task due to the similarity in brush orientation between many of the brushing regions. Figure 3.3 illustrates two brushing regions (ManLO and ManRO) that have similar brushing orientations. The difficulty in detecting left or right brushing regions was also observed and reported in previous studies [48, 49]. In Figure 3.4, we present the EAs of all 16 brushing regions obtained during a brushing session with a manual toothbrush. As can be seen, many regions have very similar orientations and overlapping EAs, making their detection challenging.

2. Intra-subject variations in brushing styles: Even when limiting the brushing behavior study to a single subject, there are variations in the orientation of the brush between different brushing sessions. Figure 3.5 compares the EAs of the brush obtained from another brushing session of the same subject using a manual toothbrush.

32
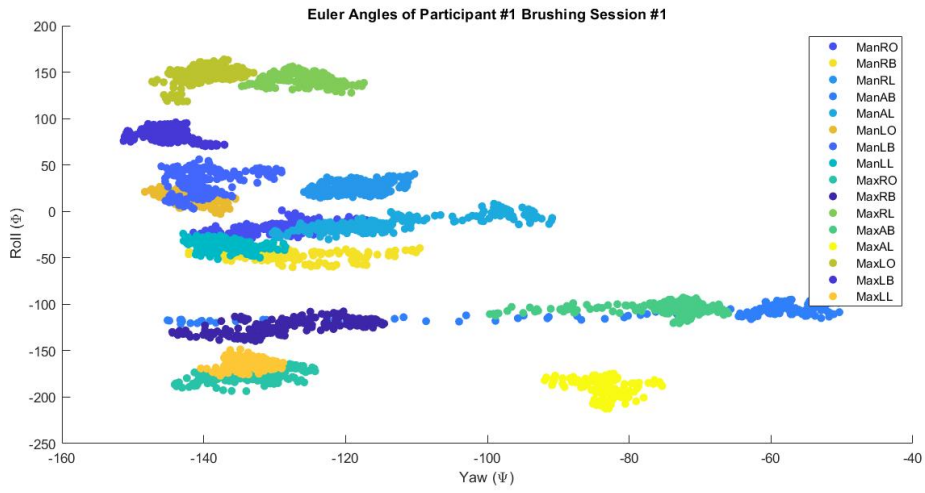
Figure 3.4: Euler Angles of 16 brushing regions of participant #1 session #1
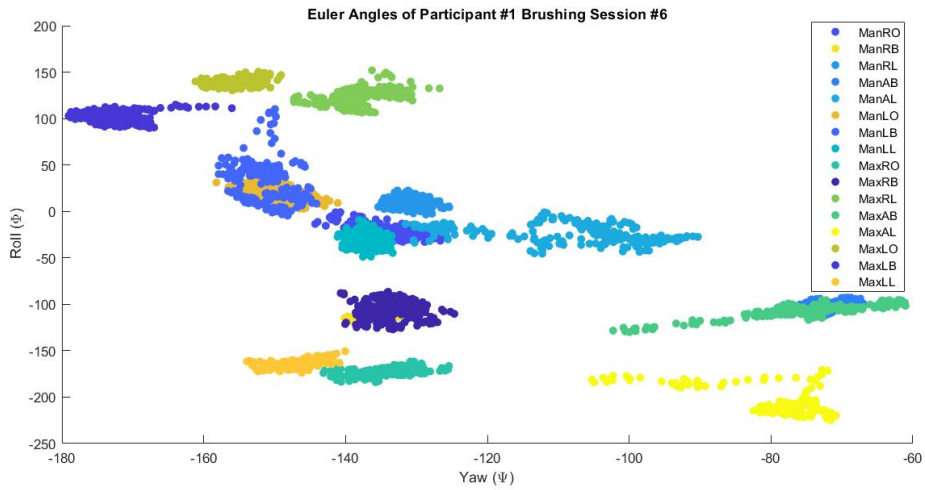
Figure 3.5: Euler Angles of 16 brushing regions of participant #1 session #6
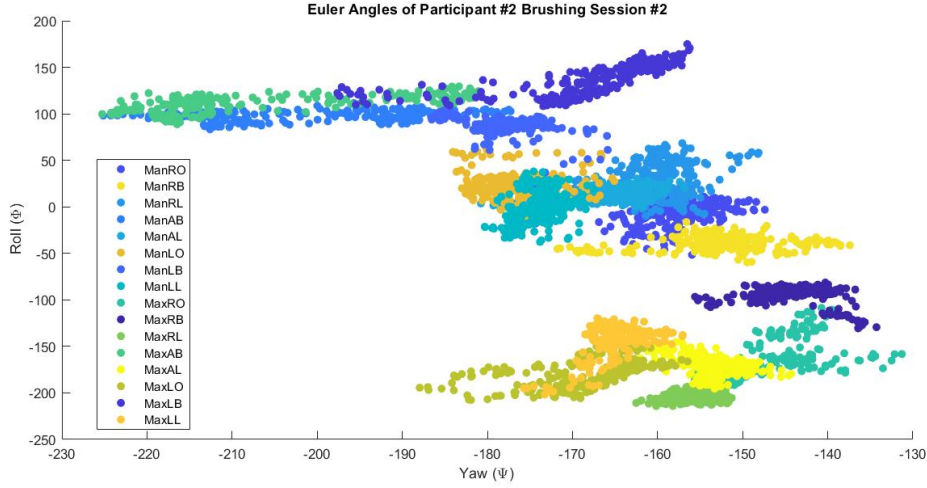
Figure 3.6: Euler Angles of 16 brushing regions of participant #2 session #2

Another factor that can further complicate this problem is the potential change in brushing habits over time.

3. Inter-subject variations in brushing styles: Although not well studied in the literature, our observations suggest that individuals have different brushing styles, as reflected in variations in toothbrushing stroke type, brushing orientations, brushing pressure, and habits such as the time spent on each brushing region and the order of visiting brushing regions. Figure 3.6 illustrates how brushing orientations differ among subjects brushing the same regions.

We study some of the variation in the brushing behavior mentioned in Chapter 5.

4. Head Movements: During a brushing session, individuals tend to move their head, which can cause artifacts in detecting brushing regions as the orientation of the toothbrush with respect to the user's head is desired. Relying solely on brush orientation without any reference to the head orientation can lead to errors in estimating brushing regions. Therefore, any proposed method for estimating brushing regions should be robust to head movements. However, head movement during toothbrushing has not been allowed in the previous studies [49]. In our experimental setup, we allowed for free head movements to

34

propose an algorithm that is robust to this factor when detecting toothbrushing regions.

5. Facing direction: There are two components that should be studied about facing direction:

   (a) Initial facing direction, $_B^W\psi_0$, is an important factor to consider when detecting the orientation of a toothbrush with respect to the user's mouth. As toothbrush users in different locations face different directions, it is essential to determine their initial facing direction. To achieve this, we aim to calculate the heading direction of the brush with respect to the user at each time t, denoted as $_B^U\psi_t$. This can be achieved using the following equation:

   $$_B^U\psi_t = {}_B^W\psi_t - {}_U^W\psi_t \tag{3.1}$$

   Here, $_B^W\psi_t$ represents the heading direction of the toothbrush in the world reference frame, while $_U^W\psi_t$ represents the facing direction of the user in the world frame. $_U^W\psi_0$ can be calculated using two methods:

   i. Using a magnetometer for an initial measurement: A method mentioned in [49] is to calculate $_B^W\psi_0$ by placing a magnetometer on the front mirror of the user at the beginning of the brushing session. However, this method has limited practicality due to its inconvenience for the user.

   ii. Using anchor regions: Another method to find $_B^W\psi_0$ is to estimate it when we are confident that the user has brushed a particular region, which we call the anchor regions, using ϕ and θ angles calculated from the accelerometer. Authors in [52] have used a similar approach, using IMU sensors and having ManLB and MaxLB regions as their anchor regions. However, the disadvantage of this method is that anchor regions may never occur in a brushing session or may occur at the end of a brushing session, which limits their usability for online detection of brushing regions and just-in-time feedback for brushing behavior change. The advantage of using a magnetometer is that it is more accurate than a gyroscope in finding the heading angle ψ, since a gyroscope is prone to the drift problem as mentioned in Section 2.2.3.

(b) Change in facing direction during a brushing session is a challenge even if $_U^W\psi_0$ is obtained using one of the methods described above, because users may move and turn around during the session. Therefore, any method that uses a magnetometer to find the heading angle of the brush is prone to error.

On the other hand, if a gyroscope is used to calculate $_U^W\psi$, it is less sensitive to changes in facing direction during the session, provided $_U^W\psi$ is initialized using anchor regions. This robustness is due to the fact that the gyroscope measures angular velocity with respect to the body frame and therefore does not sense significant changes in facing direction, even when the user turns around. However, as previously discussed in Section 2.2.3, gyroscope measurement of heading angle ψ is unreliable due to the drift problem. This issue has not been discussed in the previous studies. But perhaps this is why [38] does not include ψ angle in their proposed method.

3.4   Limitation of Physics Based Methods and the Need for Machine Learning

The challenges outlined in Section 3.3 demonstrate that while EAs are useful in detecting brushing regions, they are not comprehensive. There are many variations in brushing behavior at both population and subject levels that cannot be captured solely by relying on EAs. In other words, we can identify "groups of regions" that are either separable using EAs or have some overlaps, but further classification of the "regions inside the groups" is not feasible using EAs alone. Therefore, we propose incorporating machine learning and time series techniques in addition to the 3D geometry of orientation to address the region detection problem more effectively.

Our premise is that there exist certain time or frequency domain features that can be extracted from the brushing data at both population and individual levels to enable the classification of regions inside the groups. Indeed, some studies have yielded promising results. For instance, the authors in [39] found that the frequency of brushing is higher on the left side of the mouth than on the right side, based on motionsense data collected from dentists during

36
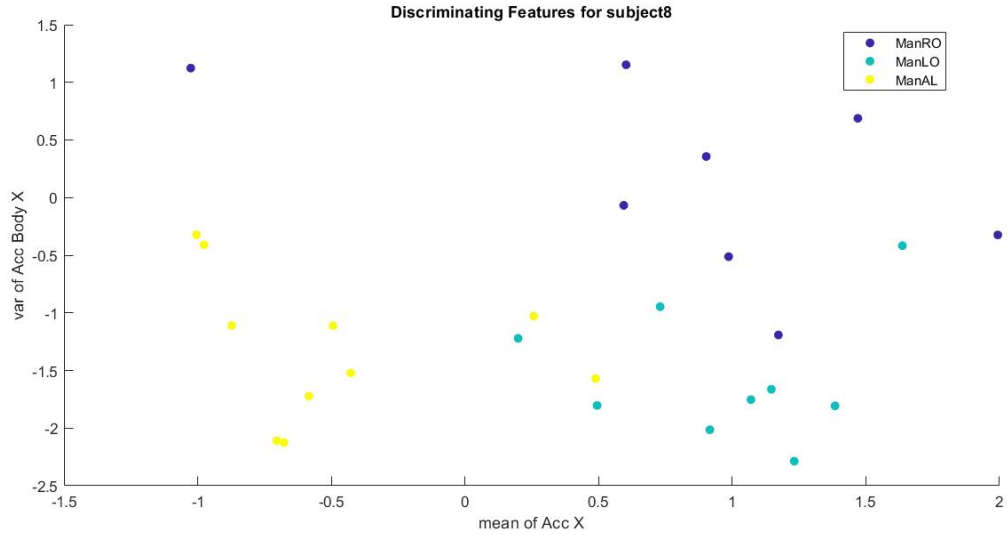
Figure 3.7: Discriminating features for classifying some brushing regions of subject #8

brushing. Furthermore, [61] demonstrated that some statistical features differ between brushing regions, such as ManLL/MaxLL and ManAB/MaxAB regions.

To provide an example of such existing features, Figure 3.7 demonstrates that ManAL, ManRO and ManLO can be distinguished using two features extracted from a manual brushing session of a subject, which relate to the force and speed of the back-and-forth brushing strokes.

To elaborate, it should be noted that although the features extracted from a manual brushing session of a subject were successful in discriminating between ManAL, ManRO, and ManLO regions, the same features may not be able to classify the regions inside the same group for a new subject. This point is highlighted in Figure 3.9.

However, it is possible that other features could be used to classify these regions for the new subject. For example, we applied new features that relate to the sideway motions of the brush to the same regions for the new subject, and the results are shown in Figure 3.8.

This indicates that personal level features extracted from individual brushing sessions may have the potential to improve the estimation of brushing regions at the population level.
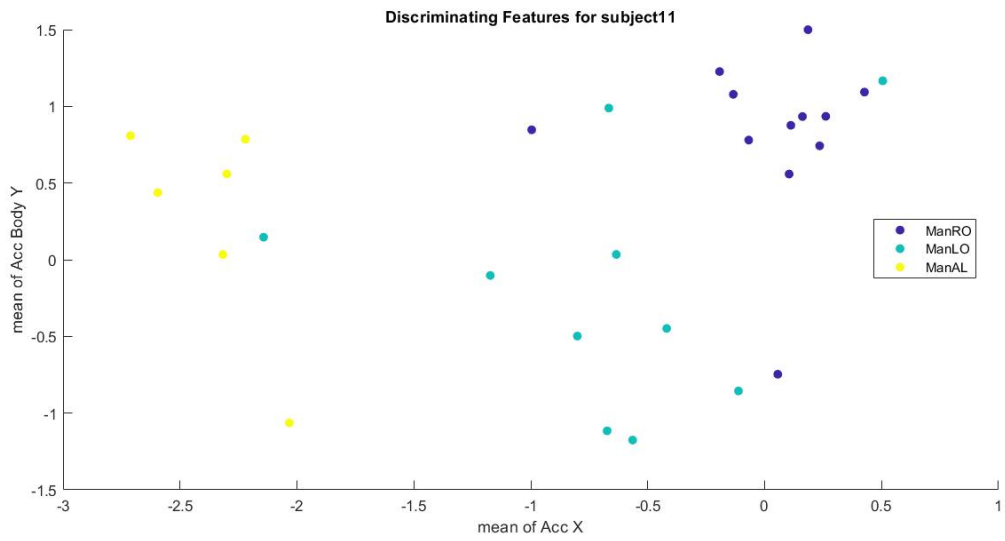
Figure 3.8: Discriminating features for classifying the same brushing regions of subject #11
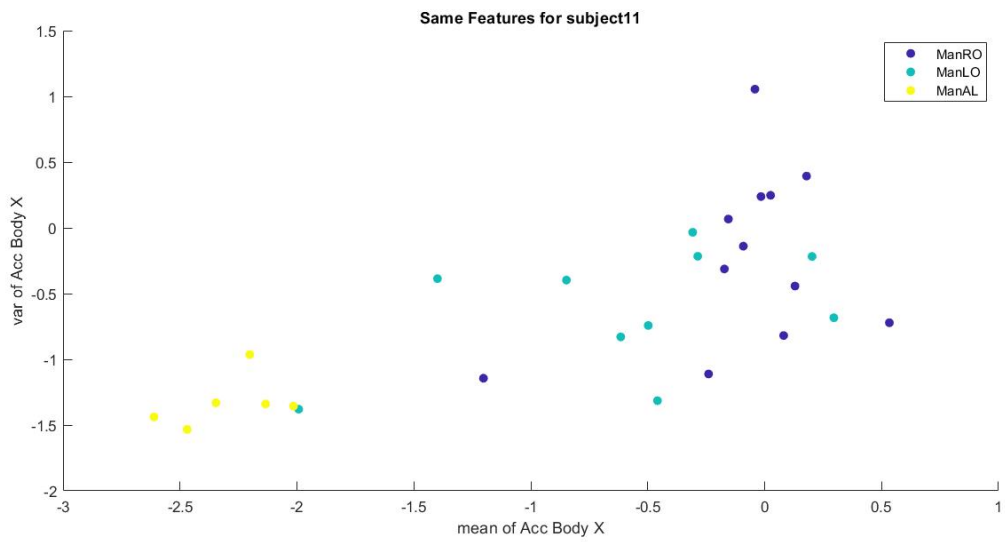
Figure 3.9: Applying the discriminating features of subject #8 to subject #11

Chapter 4

Toothbrushing Region Detection - Proposed Method

In this chapter, we propose a method to solve the brushing region detection problem, which consists of three main stages. Firstly, we preprocess the data. Secondly, we find the brush transition points, which are the moments when the brush moves from one brushing region to another. Lastly, we classify the samples between two transition points, which we refer to as segments to detect the region being brushed during that segment.

4.1 Preprocessing

In this stage, we begin by interpolating all received signals to synchronize them (as explained in Section 3.2). Then, we calibrate the magnetometer signals to compensate for the effect of any magnetic source that may affect the measurement of the earth's magnetic field. Next, we filter the accelerometer and magnetometer signals before feeding them into the orientation filter.

4.1.1 Time-stamp synchronization

To accurately label the signal, we implemented a preprocessing step to synchronize the signals from the two motion sensors used in our study: one attached to the brush and the other embedded in the wristband. This synchronization was also helpful in aligning the signals with the corresponding video footage. The motionsensors transmitted their data to a nearby phone using Bluetooth. However, we only had access to the timestamps for when the Bluetooth packets

were received by the phone, which could introduce a slight deviation of a few milliseconds compared to the time they actually been collected. To address this, we employed a linear regression approach to modify the motion sensor timestamps based on their corresponding counters (which was stored and sent in the Bluetooth packets), allowing us to achieve a nominal sampling rate of 25Hz. Additionally, some packets were lost during transmission, for which we used spline interpolation to fill in the gaps in the signals. This interpolation was performed using the timestamps of the Bluetooth packets received by the phone, ensuring synchronization of both signals.

4.1.2 Low Pass filtering

In our work, we apply a Butterworth low-pass filter of order 5 with a cutoff frequency of 2Hz. The reason for applying low-pass filtering is that orientation filters utilize accelerometer and magnetometer signals as a reference with a low dynamic response to compensate for the erroneous high-frequency components in the estimated orientation generated by the gyroscope. Specifically, the accelerometer measurements are affected by the acceleration of the brushing strokes, which appear as high-frequency components in the signals. Therefore, by applying low-pass filtering to the accelerometer signals, we can remove those components and retain only the components that relate to orientation, which have only low-frequency components generated by gravity.

4.1.3 Magnetometer Calibration

In this section, we discuss how magnetometer measurements can be affected by the geomagnetic field and the ferromagnetic sources in the vicinity of the magnetometer. These sources can cause distortion in the magnetometer measurements and can be divided into two categories:

1. Hard-iron distortions refer to the ferromagnetic components that are present on the printed circuit board (PCB) on which the magnetometer is mounted [61]. These components are fixed and will remain with the magnetometer after it is built. Hard-iron sources create a constant bias in the magnetometer measurements and can be modeled by an

additive constant vector to the magnetometer measurements.

2. Soft-iron distortions are caused by magnetic components present in the magnetometer environment, including electrical appliances, metal furniture, and metal structures within a building's construction [53]. These sources not only affect the magnetometer measurements directly but also induce a temporary magnetic field into normally unmagnetized ferromagnetic components, such as steel shields and batteries, on the PCB [61]. The effect of soft-iron sources can be modeled by a 3x3 matrix multiplication.

Hence, magnetometer measurements can be modeled as:

$$B_{\text{measured}} = W_{\text{soft-iron}} \times B_{\text{sensor}} + B_{\text{hard-iron}} \tag{4.1}$$

We include any distortion factors that can be modeled as matrix multiplication, such as accelerometer and magnetometer axis misalignment or magnetometer axis non-orthogonality, in $W_{\text{soft-iron}}$. Similarly, $B_{\text{hard-iron}}$ includes any zero field offset in the magnetometer factory calibration.

Our goal in calibrating the magnetometer is to determine $B_{\text{hard-iron}}$ and $W_{\text{soft-iron}}$ so that we can correct for magnetic distortions and obtain B_{sensor}, which represents the magnetometer measurements without any distortions. As demonstrated in [61], it can be shown that the distorted magnetometer measurements under arbitrary orientations form an ellipsoid, but if the distortions are corrected through calibration, they will form a sphere. The equation for the distorted magnetometer values is as follows:

$$\left\{W^{-1}(B_{\text{measured}} - B_{\text{hard-iron}})\right\}^{T} \left\{W^{-1}(B_{\text{measured}} - B_{\text{hard-iron}})\right\} =$$

$$(R \times B \begin{bmatrix} \cos\delta \\ 0 \\ \sin\delta \end{bmatrix})^{T} (R \times B \begin{bmatrix} \cos\delta \\ 0 \\ \sin\delta \end{bmatrix}) =$$

$$B^2$$

Here, B represents the geomagnetic field strength, δ represents the geomagnetic inclination angle, R is the rotation matrix that represents the world frame in the body frame, and $W^{-1}(B_{\text{measured}} - B_{\text{hard-iron}})$ represents the calibrated magnetometer measurements after compensating for distortion. The equation can be written in the form of an ellipsoid locus as shown below:

$$(R - R_0)^T A (R - R_0) = \text{const} \tag{4.2}$$

where $A = W^{-1} W^{-T}$ and $R_0 = B_{\text{hard-iron}}$.

By fitting an ellipsoid to distorted magnetometer measurements we can find A and R_0. Under the constraint of symmetric W^{-1} [61], W^{-1} can be found uniquely.

To fit an ellipsoid to the magnetometer measurements, we use the method proposed in [7] based on a least squares fitting method. Using the calculated W^{-1} and $B_{\text{hard-iron}}$ we can calibrate our magnetometer measurements.

4.2 Transition Points detection

In the second stage of our proposed algorithm, the transition points are estimated. Transition points are the moments that the brush transits from one region to another one. As it can be seen in figure 4.2, at these transition points, often a change in the values of some of the brush motionsense signals can be observed, which originates from the change in the orientation of the brush. This change in the signal regime is referred to as change point detection (CPD) in the literature [2].

There are several change point detection methods available in the literature [42]. However, upon examination, we found that most of these methods either have a high computational cost or have been designed for specific applications, limiting their usability in our problem.

Here, we propose CluMing, a novel CPD method based on K-means clustering. Next, we describe a nonparametric Bayesian CPD method. Lastly, we propose another CPD method, which, unlike the previous two methods, is an online algorithm. It will be demonstrated in Section 4.4 that CluMing outperforms the other methods in detecting brushing transition points.
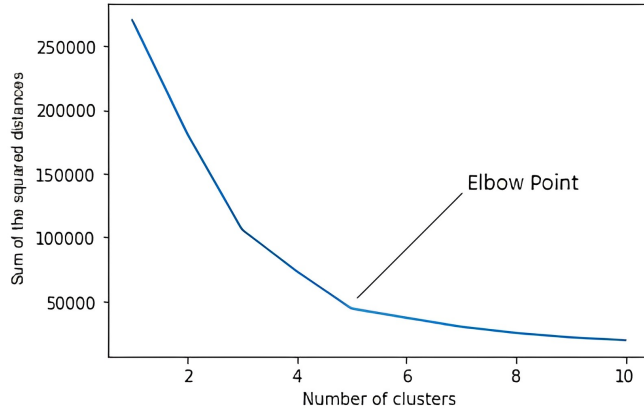
Figure 4.1: Choosing the optimal K for K-means clustering method by using the Elbow point on the graph

4.2.1 CluMing algorithm

CluMing consists of a two stages:

1. Clustering stage: In this stage, we cluster the sample points of a brushing session using accelerometer and magnetometer values as features by the K-means++ algorithm [6]. To determine the optimal number of clusters, we employ the Elbow method [73]. This technique involves calculating the sum of squared Euclidean distances between the samples and the center of their assigned cluster, serving as a measure of unexplained variation. By plotting the explained variation against the number of clusters, we identify the elbow point on the graph. The Elbow point (Figure 4.1 represents the value of k, the number of clusters, at which the explained variation exhibits a significant change. To ensure a meaningful reduction in unexplained variation, we select the largest k value that results in at least a 20% decrease in this measure.

2. Merging stage: After implementing the previous stage, all samples in the brushing session are assigned to specific clusters. Any change in the cluster assignment between consecutive samples is considered a transition point, indicating a transition of the brush from one region to another. As previously mentioned, the samples between two transition points are referred to as a segment. However, segments from the clustering stage can sometimes

be as small as 5 samples (one fifth of a second), which does not align with the fact that a brushing session, as a time series, cannot have such rapid transitions from one region to another. To address this issue, we propose a method for merging segments with their left or right neighbor segments by recursively changing their cluster assignments to their neighbor's cluster label. This approach effectively eliminates most of the falsely detected change points and results in smoother cluster assignments with fewer abrupt transitions.

In the merging step, we begin with the segment with the shortest length as our reference segment and attempt to merge it with either its left or right neighbor segment, based on the following criteria. We continue this process until there are no segments left with a length less than a threshold τ. We set τ to 25 samples, which equals one second considering the 25Hz sampling rate of our sensors. This means we expect the subject to brush for at least one second in each region before transitioning to a new dental region. The criteria for selecting the left or right neighbor segment to merge with the reference segment is based on two factors: the relative length of the neighboring segments and their mean Euclidean distance to the reference segment in the feature space (in our case 6-dimensional signal resulted from the concatenation of accelerometer and magnetometer signals). Specifically, the following criterion is used:

$$\text{S}_{LR} = \frac{l_L}{l_R} - \lambda \frac{||\bar{F} - \bar{F_L}||_2}{||\bar{F} - \bar{F_R}||_2}$$

$$\begin{cases} S > 0 & \rightarrow \text{the considered segment will merge with the left segment} \\ S \leq 0 & \rightarrow \text{the cosidered segment will merge with the right segment} \end{cases}$$

Here, $F = \begin{bmatrix} Acc & Mag \end{bmatrix}$ represents the 6-dimensional vector resulted from the concatenation of three-axis accelerometer and three-axis magnetometer signals. S_{LR} denotes the score of the left neighbor segment compared to the right neighbor segment. $\bar{F}$ represents the mean of the features of the samples in the considered segment. The subscripts L and

R indicate left and right segments, respectively. λ is a constant that weighs the length factor over the euclidean distance factor, which we set to 1 in our work.

This criterion favors longer segments as they are more likely to be correctly clustered as well as the segments that are closer in feature space, indicating that they are more likely to belong to the same cluster.

The CluMing algorithm is presented in pseudocode in Algorithm 4.2.1. An example of applying the CluMing algorithm to a segment of a brushing signal is illustrated in Figure 4.2.

Algorithm 1 CluMing algorithm

    Input F
    Output A
1: k-means++ on $F \rightarrow A^{Old}$
2: **while** $\min_i l_{Seg_i} \leq \tau$ **do**
3:     Merging Step on $A^{Old} \rightarrow A^{New}$
4:     $A^{New} \rightarrow A^{Old}$
5: **end while**
6: $A^{New} \rightarrow A$

Here, A represents the cluster assignment of samples in the brushing session.

4.2.2 Sticky HDP-HMM

A method is proposed in [25] to use Hidden Markov Models (HMMs) with an infinite number of states in which the states generate a Hierarchical Dirichlet Process (HDP) with a bias towards self-transition.

By choosing states based on a Dirichlet Process (DP), the previously visited states have a chance to be revisited, and there is also a possibility that a new state is visited that was never visited before. This model is based on an analogy to the Chinese restaurant process, where each customer that enters the restaurant will either sit at an occupied table with a probability proportional to the number of customers already sitting at that table or choose to sit at a new unoccupied table. This is a non-parametric model that assumes the number of states is not known a priori and can be potentially infinite. Therefore, a new table can always be opened up in this imaginary restaurant.
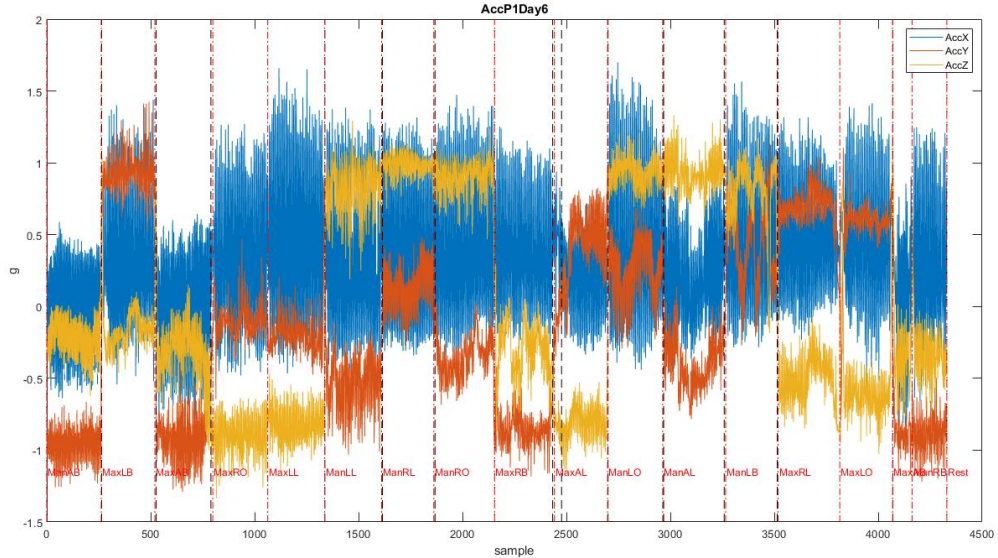
Figure 4.2: Applying CluMing algorithm to a brushing session (CluMing change point estimations are in black dashed lines and red lines are true change points

Under the assumption that observations in each state come from a multimodal distribution, the states are organized into a Hierarchical Dirichlet Process. Each parent state, called a group, has associated children states, called atoms, that have a chance of being selected based on a Dirichlet distribution. In other words, a global DP prior $DP(\alpha, G_0)$ is placed on Θ (atoms), and group-specific distributions are drawn from a global prior, $G_j \sim DP(\alpha, G_0)$. The base measure G_0 acts as an "average" distribution across all groups, and $E[G_j|G_0] = G_0$. The base measure G_0 is distributed according to a Dirichlet process $G_0 \sim DP(\gamma, H)$, implying that atoms are shared not only within groups but also between groups [70].

The HDP-HMM has been applied to the speaker diarization problem, where the goal is to recognize the correct speaker from an unknown and possibly infinite number of speakers in an unsupervised manner. To address the non-interruptive nature of speech in group settings, the authors in [25] introduced a modification to the state transition probability in the conventional HMM, by adding a bias towards self-transition, and coined the term "sticky HMM" to describe it. Through their experiments on a real speaker diarization problem, they demonstrated that sticky HDP-HMM outperforms conventional HDP-HMM, which often suffers from too many

46

switches between the estimated states.

The brushing region detection problem bears striking similarities to the speaker diarization problem, where a user brushes one region for some time before transitioning to another region, making the sticky nature of the HMM a suitable model for both. Moreover, as discussed in section 3.3, the brushing regions tend to form groups, thus justifying the use of a hierarchical structure as employed in sticky HDP-HMM. The overlap in brushing groups further strengthens the suitability of using HDP, as it allows for sharing of atoms between groups.

However, although using a nonparametric Bayesian model where the number of states can potentially be infinite may seem inappropriate for our problem with only 16 brushing states, we can still confine the number of states to an upper bound of L by using the degree L weak limit approximation to the DP, as described in [70].

Specifically, in accurate HDP, the discrete probabilities are defined as follows:

$$
\begin{aligned}
v_k \mid \gamma &\sim \text{Beta}(1, \gamma) & k &= 1, 2, \ldots \\
\beta_k &= v_k \prod_{\ell=1}^{k-1} (1 - v_\ell) & k &= 1, 2, \ldots
\end{aligned}
\tag{4.3}
$$

Here, the stick-breaking construction works by dividing a unit-length stick into segments with lengths determined by the weights β_k. Specifically, the kth weight is a random proportion v_k of the remaining stick after the previous $(k-1)$ weights have been established. This construction is commonly referred to as the $GEM(\gamma)$ distribution.

In approximated HDP, the following discrete probabilities are used:

$$
\text{GEM}_L(\alpha) \triangleq \text{Dir}(\alpha/L, \ldots, \alpha/L)
\tag{4.4}
$$

By employing this approximation, we not only reduce the computational complexity in sampling posterior probabilities, but also obtain an accurate representation of the limited state space of brushing regions by setting $L = 16$.

The authors of [25] describe two methods of inference in sticky HDP-HMM, which involve sampling posterior probabilities through either direct assignment or blocked sampling of state sequences. For our purposes, we have applied the latter method to infer the state sequences in

a brushing session, and the results of this analysis are presented in section 4.4.

4.2.3 Local Change point Detection

In this subsection, we propose a simple and efficient online CPD method called Local CPD. This method operates by using a sliding window of length L that overlaps by P percent over the motionsense signals. It detects a change in signals by observing the change in their mean values. We have used $L = 50$ and $P = 20$ heuristically to optimize the performance of the method.

To detect a change in a window, we employ the following criterion:

$$\left\| \bar{F}_i - \bar{F}_{i-1} \right\|_2 > \left\| \sigma_{i-1} \right\|_2 * \alpha + \beta \tag{4.5}$$

The criterion we use for detecting a change in a window is based on comparing the Euclidean norm of the change in $\bar{F}$, which is the mean feature vector in the considered window, to a linear threshold that is determined based on σ, the Euclidean norm of the vector of standard deviation of the features in the previous window.

Using the standard deviation to set the threshold allows for adaptability to the recent range of signal variations, which outperforms a fixed threshold in our experience.

4.3 Segment Classification

In this section, we introduce several algorithms that can be used for time series classification, ranging from traditional methods that involve feature engineering which uses hand-tailored features to deep learning frameworks that automatically select features. Later in Section 4.4, we compare the results obtained from using these algorithms for classifying brushing regions.

4.3.1 Hand-Tailored Feature Extraction Method

In this method, we extract time and frequency domain features from each segment in our training set. Our framework employs features commonly used in activity recognition problems

using motion sensors.

The features that we extracted from the brush sensor signals comprise of various measures such as the mean and variance of the motionsense signals, Auto-Regressive (AR) coefficients of those signals, mean and variance of EAs, mean quaternion representation of the orientation, mean frequency of the motionsense signals extracted from Yule-Walker [15] estimation of their spectrum, cross correlation of the motionsense signals, and mean and variance of the gravity compensated linear acceleration of the brush. These features were specifically chosen as they are commonly used in activity recognition problems using motion sensors.

Out of the 120 features extracted in this stage, we selected the top 20 features by applying linear discriminant analysis [50]. This involved finding the optimal projection matrix Φ that maximizes the Rayleigh quotient, which is defined as follows:

$$\mathcal{J}(\Phi) = \frac{\left|\Phi^T \hat{\Sigma}_b \Phi\right|}{\left|\Phi^T \hat{\Sigma}_w \Phi\right|} \tag{4.6}$$

In this optimization problem, $\mathcal{J}(\Phi)$ represents the ratio of between-class variance to within-class variance. Therefore, $\hat{\Sigma}b$ and $\hat{\Sigma}w$ denote the between-class and within-class scattering matrices, respectively. Solving the generalized eigenvalue problem below yields the solution to this optimization problem:

$$\hat{\Sigma}_b \Phi = \lambda \hat{\Sigma}_w \Phi \tag{4.7}$$

Then, based on the selected features, a classifier is used to detect the brushing region. In section 4.4, we only report the performance of the Naive Bayes classifieres since other classifiers (e.g. random forest, multiclass support vector machine (SVM), and multilayer perceptron) had similar performance.

4.3.2 Dynamic Time Warping (DTW)

DTW is a well-known method for time series classification, which can identify specific patterns in time series regardless of their shift in time or frequency. DTW finds an optimal match

between two arbitrary-length time series and can therefore yield a distance measure between them. By ignoring the length of a signal, DTW would focus on similarities in the similar patterns exist in two signals. An introduction to DTW can be found in [9].

We use the DTW distance, which is based on the euclidean distance, to classify a segment of motion sense data in the test set. This is done by finding the minimum mean DTW distance to the n random segments of the classes in the training set, as follows:

$$j^{opt} = argmin_{j \in \{1,\ldots,16\}} \frac{\Sigma_{i=1}^{n} DTW(s, t_j^i)}{n} \tag{4.8}$$

where s is the test segment and t_j^i is the ith train segment from class j. To avoid extra computational complexity, we chose the subset of $n = 100$.

The underlying assumption of our approach is that different users exhibit different brushing habits and patterns while brushing different regions. Therefore, we hypothesize that DTW can capture the distinctive time series patterns associated with these habits and aid in the detection of the brushing regions.

4.3.3   Wasserstein Distance (WD)

The WD, also known as Earth's mover distance (EMD), is a distance metric used for computing the distance between distributions. It calculates the minimum cost needed to transform one distribution into another.

In particular, let $P = \{(\mathbf{p}_1, w_{\mathbf{p}_1}), \ldots, (\mathbf{p}_m, w_{\mathbf{p}_m})\}$ be the first signature with m clusters, where $\mathbf{p}_i$ is the cluster representative and $w_{\mathbf{p}_i}$ is the weight of the cluster; Q be defined as $Q = \{(\mathbf{q}_1, w_{\mathbf{q}_1}), \ldots, (\mathbf{q}_n, w_{\mathbf{q}_n})\}$ the second signature with n clusters; and $\mathbf{D} = [d_{ij}]$ the ground distance matrix where d_{ij} is the ground distance between clusters $\mathbf{p}_i$ and $\mathbf{q}_j$. We want to find a flow $\mathbf{F} = [f_{ij}]$, with f_{ij} the flow between $\mathbf{p}_i$ and $\mathbf{q}_j$, that minimizes the overall cost [65]:

$$\text{WORK}(P, Q, \mathbf{F}) = \sum_{i=1}^{m} \sum_{j=1}^{n} d_{ij} f_{ij} \tag{4.9}$$

subject to the following constraints:

$$f_{ij} \geq 0 \qquad \text{for } 1 \leq i \leq m \quad , \quad 1 \leq j \leq n \tag{4.10}$$

$$\sum_{j=1}^{n} f_{ij} \leq w_{\mathbf{p}_i} \qquad \text{for } 1 \leq i \leq m \tag{4.11}$$

$$\sum_{i=1}^{m} f_{ij} \leq w_{\mathbf{q}_j} \qquad \text{for } 1 \leq j \leq n \tag{4.12}$$

$$\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij} = \min \left(\sum_{i=1}^{m} w_{\mathbf{p}i}, \sum_{j=1}^{n} w_{\mathbf{q}_j} \right) \tag{4.13}$$

Constraint 4.10 allows moving "supplies" from P to Q and not vice versa. Constraint 4.11 limits the amount of supplies that can be sent by the clusters in P to their weights. Constraint 4.12 limits the clusters in Q to receive no more supplies than their weights; and constraint 4.13 forces to move the maximum amount of supplies possible. We call this amount the total flow.

This problem can be solved by linear programming and once we find the optimal flow $\mathbf{F}$, the earth mover's distance is defined as the resulting work normalized by the total flow:

$$\text{EMD}(P,Q) = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} d_{ij} f_{ij}}{\sum_{i=1}^{m} \sum_{j=1}^{n} f_{ij}} \tag{4.14}$$

The above equations demonstrate that EMD takes into account not only the Euclidean distance between distributions but also the difference in their shapes. This justifies the use of EMD when dealing with the challenges illustrated in Section 3.3, as different brushing regions may have distributions with similar shapes in their corresponding feature space.

4.3.4   Long Short-Term Memory (LSTM) networks

LSTM networks [34] are specific types of Recurrent Neural Networks (RNNs) that can overcome the exploding and vanishing gradient problems happening during training of the Neural Networks. They are being used for classifying time-series data and are powerful at remembering the long term dependencies in time-series data. By using the LSTM networks, we intended

51

to capture not only the features related to the specific orientations of the brush but also any motion patterns that exist during brushing a particular region.

We concatenated the accelerometer, magnetometer, and EAs to obtain 9-dimensional feature vectors and then fed them to our LSTM model. We used two layers of bidirectional LSTM networks, with 64 hidden units and a drop-out layer with a drop-out probability of 0.1 after each layer. At the end, we connected the LSTM output to a fully-connected layer and then calculated the Softmax probabilities of the brushing regions. Our model hyper-parameters were selected heuristically by a coarse-to-fine approach.

For training, we used multi-class cross entropy loss and Adam [43] optimizer with a learning rate of 1e-3. Our network was trained for 10 epochs with a batch-size of 1024. The number of epochs was chosen based on the training and evaluation accuracy to avoid overfitting.

We further describe our training process in Subsection 4.3.6.

4.3.5 Transformer Encoder

Transformers made a revolution in natural language processing (NLP). These models that were initially used in language translation tasks, consist of two parts: an encoder and a decoder. These models use the attention mechanism [77] to learn the long-term dependencies in a corpus and are shown to outperform the RNN models in almost all NLP tasks.

While both the encoder and decoder parts of the Transformers are needed for language translation tasks, the Transformer Encoder can be used by itself to learn a vector representation of the time-series data for classification tasks. This is accomplished by adding a classification token to the beginning of the time-series segment.

In this chapter, we used the Transformer Encoder model to classify the segments of motion-sense signals during a brushing session. The architecture of our model is shown in Figure 4.3. Similar to the Subsection 4.3.4, we concatenated the accelerometer, magnetometer, and EAs as 9-dimensional feature vectors and then fed them into our Transformer Encoder model. We first used a fully-connected layer of size 32 to project our feature vectors into a feature space with higher dimensions. Then we added a positional embedding vector as well as brush-type and left-handedness embedding vectors of size 32 to the feature vectors. The addition of the latter

embedding vectors will allow further personalization to be possible via Transformer Encoder models.

Then we used a stack of four attention layers each containing a feed-forward (fully-connected) layer with 256 hidden units. Each attention layer has two attention heads with a size of 128 for key, query and value vectors. We chose a drop-out probability of 0.1 for both the attention scores as well as the feed-forward layer inside each attention layer. We used the Gelu activation function for the feed forward layers. Here again, our model hyper-parameters were selected heuristically by a coarse-to-fine approach.

Our training configuration was similar to the LSTM network discussed in Subsection 4.3.4, with the exception of using the BERT-Adam optimizer instead of the Adam optimizer.

4.3.6 Majority voting technique

Since we know that each segment of the motion-sense data which lies in between two consecutive brush transition times corresponds to only one brushing region, any subsegment of each segment should also correspond to the same brushing region. Therefore, we can split each segment into multiple subsegments using a sliding window of size l with a stride of p percent of the window length. At the training time, splitting our segments into multiple subsegments is useful due to the increase in the training set size which helps with reducing the generalization error of our classifiers. At the test time, to predict the class that each segment belongs to, we can split the segment into multiple subsegments and perform majority voting over the class predictions of the subsegments. For majority voting, we treated the subsegments as independent measurements. Therefore, we summed the log probabilities of the subsegments' distributions over the classes (i.e. brushing regions) to find the class with the highest probability as the final prediction. We applied this technique to predict segments using both the LSTM and the Transformer Encoder models and observed a performance improvement of approximately 10% in accuracy compared to not using the technique. We set the values of $l = 32$ and $p = 0.25$ heuristically.
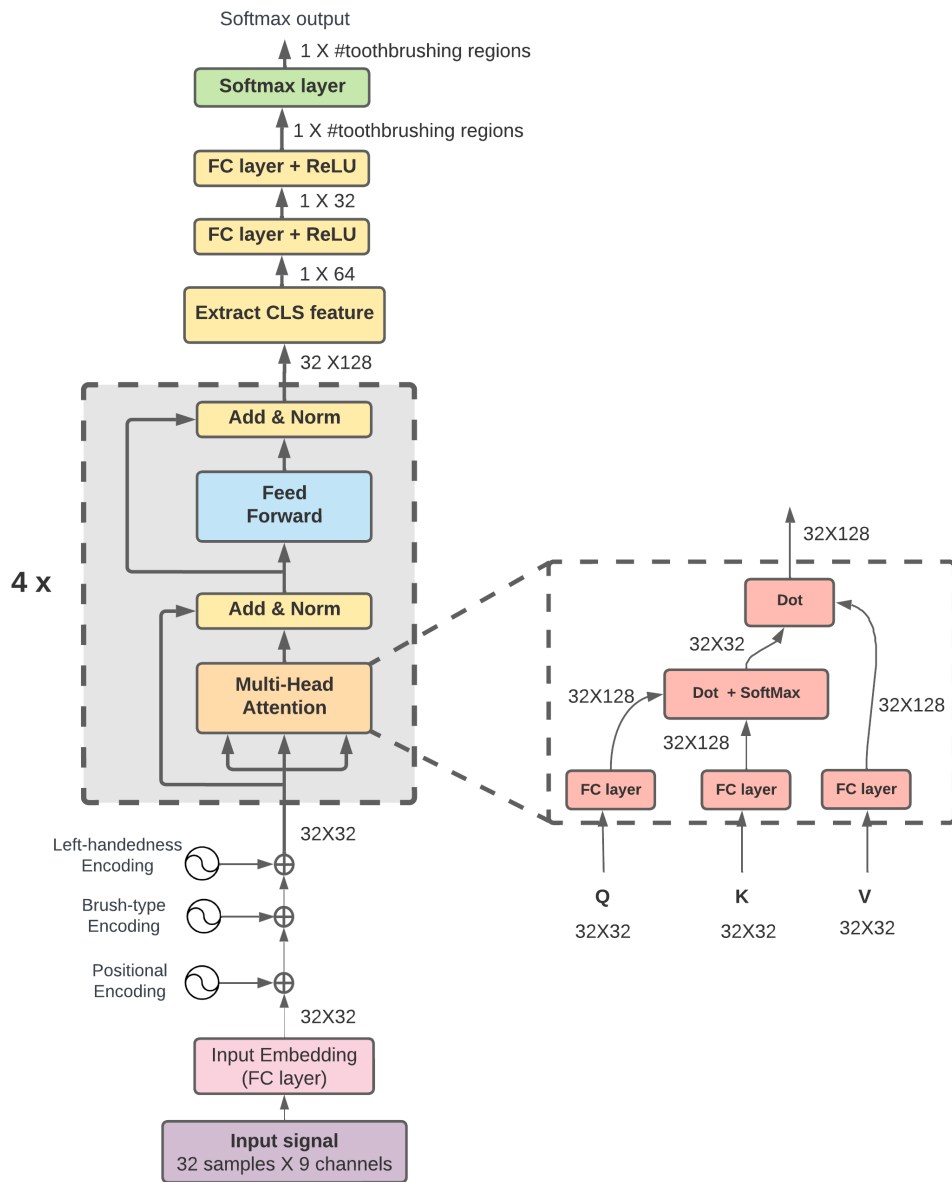
Figure 4.3: Our Transformer Encoder model configuration

4.4 Results

In this section, we present the performance of our proposed algorithm using various segment classification algorithms described in Section 4.3. First, we describe the evaluation metrics employed in our study, followed by a comparison of the performance of different classifiers based on these metrics.

4.4.1 Evaluation metrics

We report three cross-validation accuracy metrics defined as follows:

- One-subject-out accuracy: To calculate this metric, we train our model on brushing data from all participants except for one and evaluate its performance on the left-out participant. We repeat this process for each participant and then average the results. This metric evaluates the model's generalization power to new participants outside the training set and measures how well it can perform at the population level.

- One-session-out accuracy: For this cross-validation metric, we train our model on all brushing sessions of a participant except for one, which we use as the validation set. We repeat this process for all sessions of the participant and all participants, and then average the results. This method evaluates the personalization power of the models for each individual since the training set includes data from one participant.

- K-fold accuracy: For k-fold (we set k=5), cross-validation [8] all brushing sessions of all participants are used. We used this metric since it is conventionally used to evaluate the performance of the machine learning models.

4.4.2 Classification results

To evaluate the performance of our proposed algorithm, we report micro-F1 score [30] for sample-based and segment-based classification methods in Table 4.1. In the sample-based classification method, we classify each sample of motion-sense signals based on the method proposed

in [38]. The Random Forest classifier is applied to the feature vectors generated by concatenation of accelerometer, magnetometer, and roll and pitch angles for each time-sample. The roll and pitch angles are estimated using the proposed complimentary filter in [38]. In contrast to classifying feature vectors of each time-sample in sample-based methods, in the segment-based classification methods, we classify the segments that lie in between the consecutive brush transition times. In segment-based methods, we ultimately assign the predicted class of each segment to all samples in the corresponding segment. Hence, the reported accuracy of segment-based methods is calculated based on the correct predictions of the samples. This makes the comparison of the performance of the segment-based and the sample-based methods valid.

For classification, since some regions are brushed together or the accurate location of the brush is not clear in the videos, we merged some regions which were often ambiguous during labeling following the convention proposed in [4]. Hence, we ended up with 9 brushing regions as follows:

$$R = \{ManRO/ManRL, ManLO/ManLL,$$
$$MaxLO/MaxLL, MaxRO/MaxRL,$$
$$MaxRB/ManRB, MaxLB/ManLB,$$
$$MaxAB/ManAB, ManAL, MaxAL\}.$$

As it can be seen in Table 4.1, our proposed segment-based methods outperform the sample-based method proposed in [38] when applied on both their published dataset as well as our provided dataset. The difference in the performances of our model when applied on our dataset versus the dataset in [37], shows the challenging nature of brushing region detection when performed free-form versus under constraints (such as prescribed sequence of brushing, structured Bass brushing technique, etc.). While our models can achieve high k-fold cross-validation accuracy, one-subject-out classification accuracy is much more challenging. This large gap in the models' performances, manifest the vast variation in brushing styles and the other challenges discussed in Section 3.3. Hence, it is important to assess models' performances based on one-

56

Table 4.1: F1 score of our proposed classification methods

			Our Dataset	Dataset in [37, 38]
Sample-Based (method in [37, 38])		K-fold	81.5	91.2
		Subject-out	52.4	61.5
		Session-out	60.7	67.3
Segment-Based	Feature Engineering	K-fold	80.3	89.5
		Subject-out	50.7	58.1
		Session-out	60.4	66.5
	LSTM	K-fold	89.5	93.3
		Subject-out	53.4	63.7
		Session-out	65.3	70.5
	Transformer Encoder	K-fold	95.3	97.2
		Subject-out	58.6	67.9
		Session-out	67.3	74.7

subject-out evaluation metric. This should guide other researchers in the field when reporting their models' performance metrics.

Furthermore, it can be observed that one-session-out accuracy is higher than the one-subject-out accuracy. This may be because of the fact that the individuals' brushing habits might differ from each other and hence a personalized model would perform better than a general model for the whole population.

Overall, the Transformer Encoder model performed slightly better than the other segment-based models. The one-subject-out and one-session-out accuracy metrics for the Transformer Encoder model are shown in Figures 4.5 and 4.6, respectively. In the one-subject-out histogram, each bar represents the accuracy of the model when the data from that participant constituted the validation set. In the one-session-out histogram, each bar represents the accuracy result of the corresponding participant averaged over cases where each session of that participant constituted the validation set and the rest of that participant's sessions constituted the training set. Figure 4.4 displays the confusion matrix of the Transformer Encoder model applied to our provided dataset, with the one-subject-out accuracy calculated for the case where participant #1 is left out. Our model performed well in distinguishing regions with different brush-head orientations (e.g., ManRO and MaxRO). However, it had some difficulties discerning regions with similar brush orientations, such as ManRO and ManLO regions.
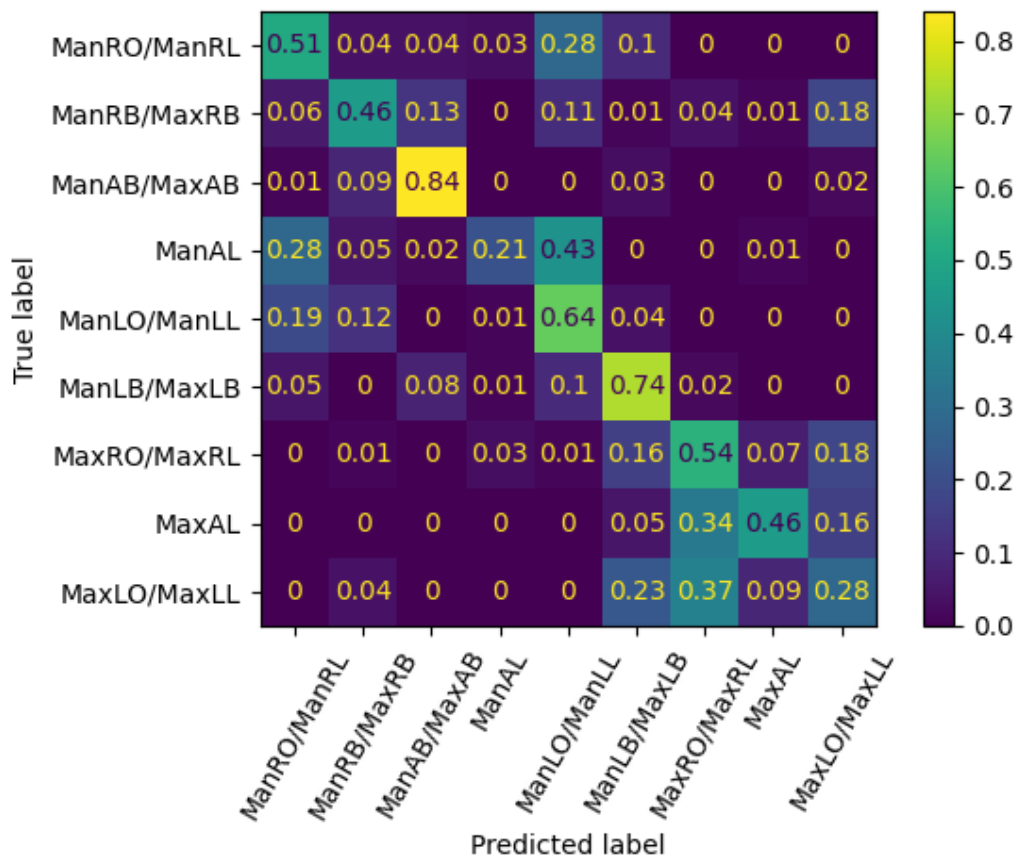
Figure 4.4: Confusion matrix of the Transformer Encoder model applied to our provided dataset, with the one-subject-out accuracy calculated for the case where participant #1 is left out.
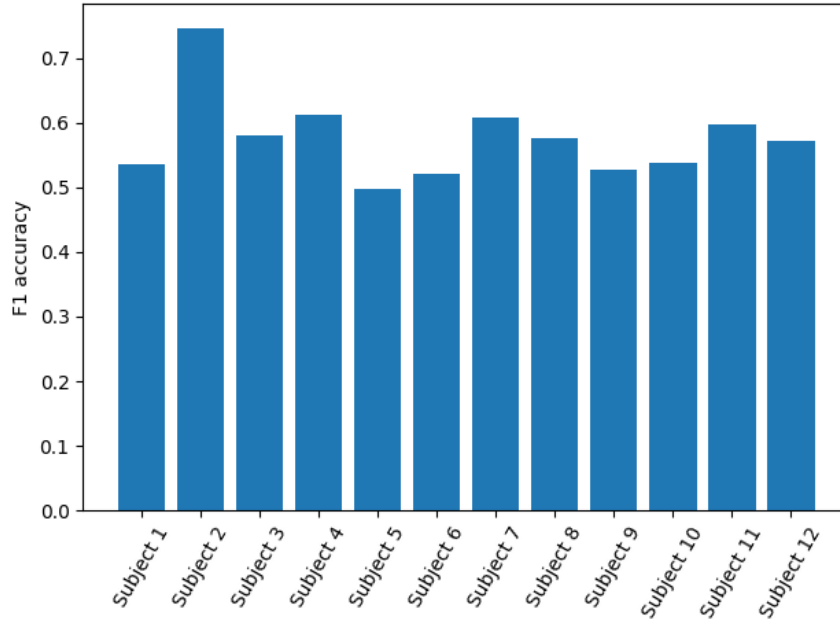
Figure 4.5: F1 score of one-subject-out evaluation metric using the Transformer Encoder model
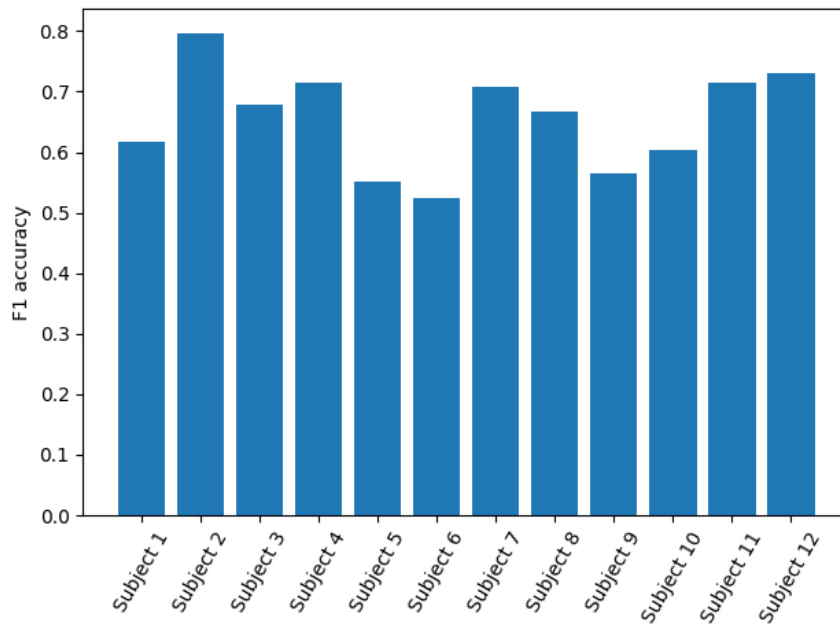
Figure 4.6: F1 score of one-session-out evaluation metric using the Transformer Encoder model

4.4.3 Change Point Detection

Evaluation metrics

To compare the performance of CPD methods, we use the Hausdorff distance [10], which measures how far two subsets of a metric space are from each other. Let $\widehat{\mathcal{T}}$ be the set of estimated change points and $\mathcal{T}^*$ be the set of true change points. The Hausdorff distance is defined as follows:

$$\text{HD}\left(\mathcal{T}^*, \widehat{\mathcal{T}}\right) := \max\{H\mathcal{T}^*, H\widehat{\mathcal{T}}\} := \max\left\{\max_{\hat{t} \in \widehat{\mathcal{T}}} \min_{t^* \in \mathcal{T}^*} \left|\hat{t} - t^*\right|, \max_{t^* \in \mathcal{T}^*} \min_{\hat{t} \in \widehat{\mathcal{T}}} \left|\hat{t} - t^*\right|\right\} \quad (4.15)$$

This can be viewed as the greatest distance from any point in either one of the sets to the closest point in the other set.

We also introduce three new metrics for quantifying the performance of change point detection methods: the number of extra estimated change points, denoted by E; the number of missing change points, denoted by M; and the distance error, denoted by D, which measures the discrepancy between the estimated change points and the true change points.

To achieve this, we find the maximum matching between the sets of bipartite graphs $\mathcal{T}^*$ and $\widehat{\mathcal{T}}$ with the minimum sum of distances, as follows:

$$\min_x \sum_{\forall t^* \in \mathcal{T}^*, \forall \hat{t} \in \widehat{\mathcal{T}}} d(t^*, \hat{t}) x(t^*, \hat{t}) \qquad \text{subject to:}$$
$$\sum_t x(t^*, \hat{t}) = \min\{|\mathcal{T}^*|, |\widehat{\mathcal{T}}|\} \qquad \forall t^* \in \mathcal{T}^*$$
$$\sum_{t^*} x(t^*, \hat{t}) = \min\{|\mathcal{T}^*|, |\widehat{\mathcal{T}}|\} \qquad \forall \hat{t} \in \widehat{\mathcal{T}}$$
$$x(t^*, \hat{t}) \in \{0, 1\} \qquad \forall t^* \in \mathcal{T}^*, \hat{t} \in \widehat{\mathcal{T}}$$

where $|\cdot|$ denotes the cardinality of a set. For the distance metric d, various distance measures can be employed, and we use the $\ell 1$ norm as a straightforward distance measure.

By solving a proxy to this problem and obtaining x^{opt}, we can define M, E, and D based on the set of matched estimated change points, denoted by $\mathcal{T}^{\mathcal{M}}$, as follows:

	E $\left(\text{probability}\right)$	M $\left(\text{probability}\right)$	D $\left(\text{samples}\right)$	HP $\left(\text{samples}\right)$	HQ $\left(\text{samples}\right)$	HD $\left(\text{samples}\right)$
Local CPD	0.29	0.16	7.84	217.12	255.68	285.89
CluMing	0.25	0.14	5.59	209.07	256.35	281.68
HDP-HMM	0.12	0.54	3.29	1591.1	254.22	1593.5

Table 4.2: Accuracy of change point detection methods

$$\mathcal{T}^{\mathcal{M}} = \{\forall t^* \in \mathcal{T}^* \ | \ \exists \hat{t} \in \widehat{\mathcal{T}} \text{ such that } x^{opt}(t^*, \hat{t}) = 1\}$$
$$E = \max\{0, |\mathcal{T}^{\mathcal{M}} - \widehat{\mathcal{T}}|\}$$
$$M = \max\{0, |\widehat{\mathcal{T}} - \mathcal{T}^{\mathcal{M}}|\}$$
$$D = \sum_{\forall t^* \in \mathcal{T}^*, \forall \hat{t} \in \mathcal{T}} d(t^*, \hat{t}) x^{opt}(t^*, \hat{t})$$

Results

According to Table 4.2, cluMing outperforms HDP-HMM in terms of both accuracy and run-time by a significant margin among the CPD methods. However, Local CPD is a faster algorithm, making it more suitable for online CPD. We heuristically tuned the hyper-parameters α and β to 0.5 and 1, respectively, although further hyper-parameter tuning is possible.

4.5   Discussion and Conclusion

In this chapter, we present a comprehensive dataset designed for the detection of dental regions being brushed during a brushing session. Our dataset comprises motion-sense signals obtained from accelerometers, gyroscopes, and magnetometers attached to the toothbrush. We collected a total of 187 brushing sessions from 12 participants. Unlike previous studies that focused on constrained brushing conditions, such as following a structured Bass brushing technique, prescribed brushing transition sequences, and fixed head/body positions, our data collection encompasses free-form brushing, reflecting real-world naturalistic settings.

We discussed the challenges associated with accurately tracking the position of the toothbrush during brushing, considering the geometric dimensions of the mouth. To address this, we proposed a three-stage algorithm consisting of pre-processing, brush transition time detection, and time-series classification. This algorithm enables the prediction of nine brushing regions by identifying the moments when the brush transitions between different dental regions and classifying the segments between consecutive brush transitions. Notably, we observed that free-style brushing poses greater challenges compared to the dataset presented in [37], which was collected under structured brushing assumptions.

Furthermore, we demonstrated that classifiers can achieve high classification accuracy using a random train-test split approach, such as k-fold cross-validation. However, when it comes to generalizing the algorithm to new toothbrush users, particularly in leave-one-subject-out cross-validation, significant challenges arise due to variations in brushing styles and other factors, such as head movement during brushing.

The outcomes, as presented in Section 4.4, were achieved via a hyperparameter tuning process. To optimize the hyperparameters of our classifiers, we employed a coarse-to-fine approach consisting of two steps:

1. Initially, a coarse grid of values was selected to evaluate the overall performance of the classifiers over a wide range of hyperparameter values.

2. Subsequently, in the second step, a finer grid of values was explored around the range of values that yielded the best results on the desired metric, particularly one-subject-out cross-validation accuracy.

In addition, in the case of neural network classifiers, we did not allow the time-consuming training process to run for multiple epochs. Instead, we terminated the training process after the first few epochs. This approach is in line with the concept of the critical learning period in deep learning [1], where the initial phase of training neural networks is shown to be critical in achieving the final performance of the network.

In general, we observed that for neural networks, regardless of the type of the three cross-validation metrics mentioned in Section 4.4, increasing model complexity enables attaining high

training accuracy. However, this usually results in a decline in validation accuracy, commonly known as overfitting in machine learning. Below are the ranges of hyperparameters that we found to be effective for the classifiers we employed. Hyperparameter values falling outside these ranges may cause underfitting or overfitting for the classifiers we utilized:

- Transformer Encoder:

  - Hidden size for the key, query, and value: [128, 256]

  - number of stacked-attention layers: [3, 5]

  - number of hidden units in the attention layer: [128, 512]

  - number of attention heads: [2, 4]

  - dropout probabilities: [0.1, 0.4]

- LSTM:

  - hidden size: [64, 128]

  - dropout probability: [0.1, 0.4]

- Random-Forrest (for the method proposed in [38]):

  - number of trees: [70, 120]

  - trees' depths: [12, 18]

Regarding the optimization process, we discovered the subsequent range of values that prevent underfitting or overfitting:

- learning rate: [5e-4, 5e-3]

- number of epochs: [5, 10] for one-subject- and one-session- out cross-validation, and [20, 30] for k-fold cross-validation

- batch size: [256, 2048]

For the majority-voting technique:

- segment length: $[0.4, 1.5] \times$ sampling rate

- stride length: $[0.1, 0.4] \times$ sampling rate

Nevertheless, although we attempted to find the optimal values within the aforementioned ranges, additional experiments within these ranges may yield further optimization.

Regarding the labeling process, we implemented a two-step algorithm that involves relabeling some of the inaccurate labels using the trained classifiers' predictions. However, the literature provides various methods to handle partially accurate labels, commonly referred to as noisy labels [69]. Some methods tackle this problem by modifying the loss function, such as Bootstrapping, Forward/Backward Loss Correction, and Mixup, while others rely on relabeling the data based on the trained classifiers' predictions. Certain approaches utilize semi-supervised learning after discarding less confident labels. However, in our study, we proposed a relabeling approach that uses the classifiers' predictions in conjunction with the video recordings of the brushing sessions as the ground truth. Our approach is particularly suitable for our work due to the relatively small size of our dataset and the lack of brushing datasets in the field. Moreover, for regions that can be brushed together (e.g., ManAB, MaxAB), we could label them as ManAB/MaxAB, which could pave the way for multi-label classification [76].

Chapter 5

Toothbrushing Behavior Analysis

In this chapter we explain the heterogeneity of brushing behaviors. We gathered accurate data on habitual brushing patterns using powered/electronic toothbrushes, with a focus on the duration of each session, tooth surfaces covered, and episodes of excessive brushing pressure per session. By examining habitual brushing patterns at the individual and session-level, we sought to clarify the similarities as well as the between- and within- person variability in brushing patterns using powered toothbrushes.

5.1   Data collection

As part of the parent study, 12 healthy college students with no evident dental disease, provided their brushing data in the home setting over three weeks (10 sessions each). The 12 participants comprised of eight females and four males with ages ranging from 18 to 23 years (20.77 ± 1.59).

Basic instruction on the use of the brush and setting up the data collection system was given and the participants were instructed to freely brush their teeth in a manner most natural to them. All participants provided written informed consent and the study protocol was reviewed and approved by the Institutional Review Board of the University of California, Los Angeles (IRB#18–000874).

To allow objective, individual-level, and ecologically-valid data on oral hygiene behaviors, we deployed the Remote Oral Behaviors Assessment System (ROBAS) described previously in [68]. Briefly, ROBAS builds on a broadly available consumer-grade powered toothbrush

(Oral-B Genius X; Procter & Gamble) as the data source for brushing behaviors (timing, duration, pressure applied). The Oral-B brush employs a rotational-oscillation mode of action and brushing movements are captured by an accelerometer, gyroscope, and pressure sensor contained within the powered brush. Captured data is transmitted over BLE (Bluetooth Low Energy) to a paired smartphone running the companion data collection app. Collected data is then uploaded to a secure cloud server for remote monitoring of data yields and analytics. Visualization of time-series data streams of brushing episodes and remote monitoring of sensor function and participant compliance is accomplished through an adaptation of the open platform GrafanaTM dashboard [31].

Upon enrollment, each participant was provided an Oral-B powered toothbrush, a suction-cup phone mount, charger, and quick-start instructions. Participants downloaded the study-specific app onto their own smartphone and paired it to the powered brush. Participants were instructed on the operation of the powered brush and on how to mount the smartphone to their bathroom mirror during a brushing session for the duration of the study (3 weeks). Participants were instructed to brush twice daily for two minutes each time. At the start of each brushing session, participants launched the study app and activated the smartphone camera. Data from the embedded sensors was buffered by the brush and transmitted to the study phone via Bluetooth. Turning off the toothbrush ended data collection and triggered the app to save the timestamped brushing data. The brushing session data and the corresponding video clip were then uploaded via the ROBAS platform to a secure cloud server for subsequent analysis. The ROBAS platform with integrated GrafanaTM dashboard allowed research staff to remotely monitor data feeds and conduct quality checks.

For labeling the brushing sessions, we performed the same procedure explained in Section 3.2.2.

5.2 Analyses

Several participants skipped some regions altogether in some or all brushing sessions, and most participants avoided using excessive pressure most of the time. Therefore, the durations of

region-specific brushing and region-specific brushing with excessive pressure were often equal to zero, a statistical phenomenon referred to as zero-inflation [11]. Accordingly, we analyzed these outcomes using statistical models that account for zero-inflation. Specifically, we modeled these outcomes (each measured in counts of 25Hz samples) using zero-inflated generalized linear mixed-effects regression models, with a log-link and a negative binomial outcome distribution for the count submodels and a logistic link and a Bernoulli outcome distribution for the zero-inflation submodels.

Here, we provide a more detailed explanation of our analysis regarding three quantities of interest: the duration of brushing on each dental surface, the duration of excessive brushing pressure on each dental surface, and the total active brushing duration of each brushing session.

We calculated the total active brushing duration per-session by excluding pauses in brushing and the epochs of transitioning the brush head to different dental surfaces.

To examine the data on subject and session levels, we used boxplots. Data points were labeled as outliers if they were not in the range of $[q1 - w \times (q3 - q1), q3 + w \times (q3 - q1)]$; in which w is the Whisker value and q1 and q3 are the 25th and 75th percentiles of the sample data, respectively. We used a whisker value of $\pm 2.7\sigma$ (σ is the standard deviation of the sample data) that corresponds to the coverage of 99.3% of the data, if the data is normally distributed. The significance level α was set at 5%.

All statistical analysis was performed in R version 4.1.1 [64], using the regression modeling package "glmmTMB" [11]. The boxplots are generated using Matlab R2021a [71]. The dataset used as well as the code to generate the results are publicly available publicly at a GitHub repository [22].

5.2.1 The duration of brushing on each dental surface

Several participants skipped some regions altogether in some or all brushing sessions. Therefore, the brushing durations of any dental surface was often equal to zero, and hence forming a zero-inflated distribution. Zero-inflation for brushing duration distribution of MaxRO is shown in Fig 5.1.

As a result, we fit a zero-inflated negative binomial regression model for the amount of
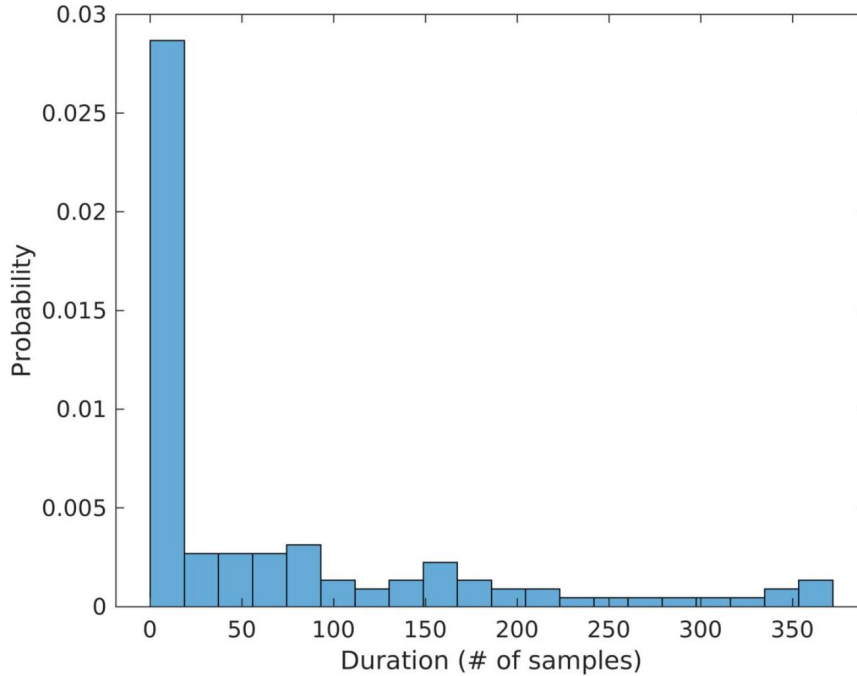
Figure 5.1: Zero-inflated distribution of brushing duration on MaxRO.

time (measured in counts of 25 Hz samples) spent brushing each dental surface, with a log-link and a negative binomial outcome distribution for the count submodel and a logistic link and a Bernoulli outcome distribution for the zero-inflation submodel. Both submodels had fixed effects for tooth surface, mouth side, and jaw, and random effects on the intercept by session nested in participant, to account for participant-to-participant and session-to-session differences in overall brushing duration, and participant-specific overdispersion parameters to account for participant-to-participant differences in residual variance. The count submodel also included participant-specific random effects for tooth surface, mouth side, and jaw. We attempted to add participant-specific random effects for tooth surface, mouth side, and jaw in the zero-inflation submodel, but the estimation algorithm failed to converge for that extended model.

Fig 5.2 summarizes the brushing duration of each region for all participants. MaxAB, MaxLB, and MaxRB were the areas brushed the longest with a median of 10.68, 8.78, and 8.22 seconds respectively. In contrast, MaxLL and MaxLO were frequently skipped during brushing.
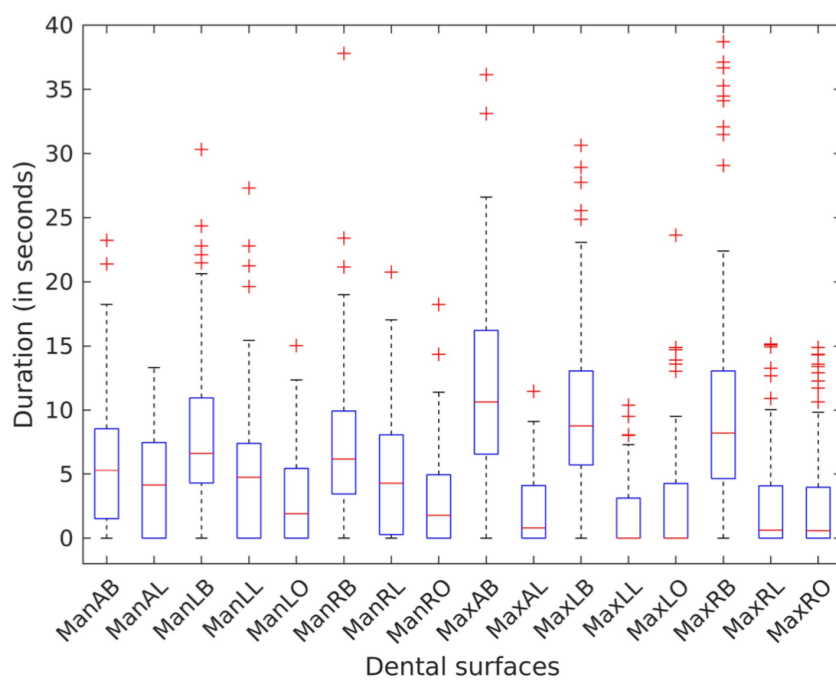
Figure 5.2: Brushing time of all dental surfaces. MaxAB, MaxLB, and MaxRB were the areas brushed the longest and in contrast, MaxLL and MaxLO were frequently skipped during brushing.

Modeling

The brushing duration Y_{ijk} of participant $i \in \{1 : 12\}$, in session $j \in \{1 : 10\}$, of dental surface $k \in \{$MaxRO, MaxRB, MaxAB, MaxLB, MaxLO, MaxRL, MaxAL, MaxLL, ManRO, ManRB, ManAB, ManLB, ManLO, ManRL, ManAL, ManLL $\}$ is modeled in Equation 5.1:

$$
\begin{aligned}
& Z_{ijk} \sim \text{Bernoulli}\,(\pi_{ijk}) \\
& Y_{ijk} \sim
\begin{cases}
0 & \text{if } Z_{ijk} = 1 \\
\text{NB}\,(\lambda_{ijk}, \kappa_i) & \text{if } Z_{ijk} = 0
\end{cases}
\end{aligned}
\tag{5.1}
$$

We used the negative binomial distribution parameterized as described in Equation 5.2:

$$
p\,(Y_{ijk} = y \mid Z_{ijk} = 0) = \frac{(\lambda_{ijk})^y}{y!} \cdot \frac{\Gamma\,(y + \kappa_i)}{\Gamma\,(\kappa_i)\,(\kappa_i + \lambda_{ijk})^y} \cdot \left(1 + \frac{\lambda_{ijk}}{\kappa_i}\right)^{\kappa_i}
\tag{5.2}
$$

Hence:

$$
\begin{aligned}
\mathbb{E}\,(Y_{ijk} \mid Z_{ijk} = 0) &= \lambda_{ijk} \\
\text{Var}\,(Y_{ijk} \mid Z_{ijk} = 0) &= \lambda_{ijk}\left(1 + \frac{\lambda_{ijk}}{\kappa_i}\right)
\end{aligned}
$$

It should be noted that larger values of κ_i correspond to smaller variances.

We modeled $\lambda_{ijk} = \mathbb{E}\,(Y_{ijk} \mid Z_{ijk} = 0)$ as shown in Equation 5.3:

$$\log\left(\lambda_{ijk}\right) = \beta_{0ij} + \beta_{\mathcal{M}i}1\{k \in \mathcal{M}\} + \beta_{\mathcal{O}i}1\{k \in \mathcal{O}\} + \beta_{\mathcal{L}i}1\{k \in \mathcal{L}\} + \beta_{\mathcal{A}i}1\{k \in \mathcal{A}\} + \beta_{\mathcal{G}i}1\{k \in \mathcal{G}\}$$

$$\beta_{0ij} = \beta_0 + \alpha_{0i} + \gamma_{ij}$$

$$\beta_{\mathcal{M}i} = \beta_{\mathcal{M}} + \alpha_{\mathcal{M}i}$$

$$\gamma_{ij} \sim N\left(0, \sigma_\gamma^2\right)$$

$$\begin{pmatrix} \alpha_{0i} \\ \alpha_{\mathcal{M}i} \\ \vdots \end{pmatrix} \sim N\left(\mathbf{0}, \begin{bmatrix} \sigma_{\alpha_0}^2 & \sigma_{\alpha_0}\sigma_{\mathcal{M}}\rho_{\alpha_0\mathcal{M}} & \cdots \\ \sigma_{\alpha_0}\sigma_{\mathcal{M}}\rho_{\alpha_0\mathcal{M}} & \sigma_{\mathcal{M}}^2 & \ddots \\ \vdots & \ddots & \ddots \end{bmatrix}\right)$$

$$\log\left(\kappa_i\right) = \eta_i$$

$$(5.3)$$

We modeled π_{ijk} as in Equation 5.4:

$$\text{logit}\left(\pi_{ijk}\right) = \theta_{0ij} + \theta_{\mathcal{M}}1\{k \in \mathcal{M}\} + \theta_{\mathcal{O}}1\{k \in \mathcal{O}\} + \theta_{\mathcal{L}}1\{k \in \mathcal{L}\} + \theta_{\mathcal{A}}1\{k \in \mathcal{A}\} + \theta_{\mathcal{G}}1\{k \in \mathcal{G}\}$$

$$\theta_{0ij} = \theta_0 + \delta_i + \epsilon_{ij}$$

$$\delta_i \sim N\left(0, \sigma_\delta^2\right)$$

$$\epsilon_{ij} \sim N\left(0, \sigma_\epsilon^2\right)$$

$$(5.4)$$

In expressions 5.3 and 5.4, $\mathcal{M}, \mathcal{O}, \mathcal{L}, \mathcal{A}, \mathcal{G}$ denote the sets of maxillary, occlusal, lingual, anterior, and gauche (left) regions, respectively; e.g., $\mathcal{M} = \{$ MaxRO, MaxRB, MaxAB, MaxLB, MaxLO, MaxRL, MaxAL, MaxLL $\}$ and $\mathcal{O} = \{$ MaxRO, ManRO, MaxLO, ManLO $\}$.

Results

We initially considered using a Poisson distribution for the outcome of the count submodel. However, upon comparing the results, we found that the negative binomial distribution is a more suitable choice. The Akaike Information Criterion (AIC) for the zero-inflated negative binomial model was 18,598.76, which was 69,183.57 lower than the AIC of the zero-inflated Poisson model with the same fixed and random effects. Additionally, the Bayesian Information Criterion (BIC) for the negative binomial model was 18,865.64, which was 69,116.85 lower than

Table 5.1: Estimated fixed effects of count submodel for brushing duration of each dental surface.

Parameter	Log-Mean	SE	95%CI	p
β_0 : Intercept	5.33	0.11	$(5.12, 5.54)$	$< .001$
$\beta_{\mathcal{L}}$: Surface (Lingual)	-0.78	0.22	$(-1.21, -0.35)$	$< .001$
$\beta_{\mathcal{O}}$: Surface (Occlusal)	-0.80	0.16	$(-1.13, -0.48)$	$< .001$
$\beta_{\mathcal{A}}$: Side (Anterior)	-0.01	0.09	$(-0.19, 0.17)$	0.892
β_G : Side (Left)	$-3.29e-03$	0.08	$(-0.15, 0.15)$	0.966
$\beta_{\mathcal{M}}$: Jaw (Maxillary)	0.09	0.14	$(-0.18, 0.37)$	0.508

Table 5.2: Estimated fixed effects of zero-inflation submodel for brushing duration of each dental surface.

Parameter	Log-Odds	SE	95% CI	p
θ_0 : (Intercept)	-4.25	0.43	$(-5.09, -3.41)$	$< .001$
$\theta_{\mathcal{L}}$: Surface (Lingual)	3.04	0.23	$(2.59, 3.49)$	$< .001$
$\theta_{\mathcal{O}}$: Surface (Occlusal)	3.47	0.25	$(2.98, 3.96)$	$(2.59, 3.49)$
$\theta_{\mathcal{A}}$: Side (Anterior)	0.17	0.19	$(-0.20, 0.53)$	0.363
θ_{G} : Side (Left)	0.13	0.14	$(-0.16, 0.41)$	0.381
$\theta_{\mathcal{M}}$: Jaw (Maxillary)	0.68	0.13	$(0.42, 0.94)$	$< .001$

the BIC of the zero-inflated Poisson model with the same fixed and random effects. These findings demonstrate that the negative binomial distribution is a better choice as it mitigates the risk of overfitting.

The estimated parameters for the count and zero-inflated submodels are summarized in the tables below .

Table 5.6: Estimated participant-level random effects of count sub-model for brushing duration of each dental surface.

Participant #	Parameter	Estimate	Std. Error	Pr($>$ |z|)	2.5%	97.5%
1	α_{01} : (Intercept)	0.132	0.141	0.347	-0.143	0.408
2	α_{02} : (Intercept)	0.363	0.140	0.010	0.087	0.638
3	α_{03} : (Intercept)	-0.008	0.152	0.958	-0.306	0.290
4	α_{04} : (Intercept)	0.506	0.167	0.003	0.178	0.834
5	α_{05} : (Intercept)	-0.305	0.181	0.091	-0.660	0.049
6	α_{06} : (Intercept)	0.091	0.166	0.584	-0.234	0.416

Participant #	Parameter	Estimate	Std. Error	Pr(> |z|)	2.5%	97.5%
7	α_{07} : (Intercept)	-0.225	0.206	0.275	-0.629	0.179
8	α_{08} : (Intercept)	-0.273	0.162	0.092	-0.589	0.044
9	α_{09} : (Intercept)	-0.473	0.177	0.007	-0.820	-0.127
10	α_{010} : (Intercept)	-0.114	0.155	0.459	-0.417	0.188
11	α_{011} : (Intercept)	0.496	0.185	0.007	0.134	0.858
12	α_{012} : (Intercept)	-0.267	0.158	0.091	-0.575	0.042
1	$\alpha_{\mathcal{L}1}$: SurfaceLingual	0.188	0.239	0.431	-0.280	0.657
2	$\alpha_{\mathcal{L}2}$: SurfaceLingual	0.498	0.236	0.034	0.037	0.960
3	$\alpha_{\mathcal{L}3}$: SurfaceLingual	0.401	0.240	0.095	-0.070	0.872
4	$\alpha_{\mathcal{L}4}$: SurfaceLingual	-0.593	0.264	0.025	-1.110	-0.076
5	$\alpha_{\mathcal{L}5}$: SurfaceLingual	-0.276	0.257	0.282	-0.779	0.227
6	$\alpha_{\mathcal{L}}$ SurfaceLingual	-0.794	0.425	0.061	-1.627	0.038
7	$\alpha_{\mathcal{L}}$: SurfaceLingual	-0.951	0.869	0.274	-2.654	0.752
8	$\alpha_{\mathcal{L}8}$: SurfaceLingual	0.547	0.251	0.029	0.055	1.039
9	$\alpha_{\mathcal{L}9}$: SurfaceLingual	0.940	0.262	< 0.001	0.426	1.453
10	$\alpha_{\mathcal{L}10}$: SurfaceLingual	0.496	0.245	0.043	0.016	0.975
11	$\alpha_{\mathcal{L}11}$: SurfaceLingual	-1.015	0.304	0.001	-1.610	-0.420
12	$\alpha_{\mathcal{L}12}$: SurfaceLingual	0.504	0.249	0.043	0.015	0.992
1	α_{O1} : SurfaceOcclusal	0.312	0.194	0.108	-0.069	0.693
2	α_{O2} : SurfaceOcclusal	0.091	0.225	0.686	-0.350	0.533
3	α_{O3} : SurfaceOcclusal	-0.996	0.257	< 0.001	-1.499	-0.492
4	α_{O4} : SurfaceOcclusal	-0.545	0.253	0.032	-1.042	-0.048
5	α_{O5} : SurfaceOcclusal	-0.824	0.259	0.001	-1.332	-0.315
6	α_{O6} : SurfaceOcclusal	0.567	0.218	0.009	0.140	0.994
7	α_{O7} : SurfaceOcclusal	0.109	0.297	0.714	-0.474	0.692
8	α_{O8} : SurfaceOcclusal	0.390	0.214	0.068	-0.029	0.809
9	α_{O9} : SurfaceOcclusal	-0.311	0.259	0.229	-0.819	0.196

Continued on next page

Participant #	Parameter	Estimate	Std. Error	Pr(> |z|)	2.5%	97.5%
10	α_{O10} : SurfaceOcclusal	0.569	0.208	0.006	0.161	0.977
11	α_{O11} : SurfaceOcclusal	0.301	0.236	0.201	-0.161	0.764
12	α_{O12} : SurfaceOcclusal	0.266	0.209	0.204	-0.144	0.677
1	$\alpha_{\mathcal{M}1}$: JawMaxillary	-0.148	0.164	0.365	-0.469	0.172
2	$\alpha_{\mathcal{M}2}$: JawMaxillary	-0.046	0.164	0.781	-0.366	0.275
3	$\alpha_{\mathcal{M}3}$: JawMaxillary	-0.084	0.168	0.615	-0.414	0.245
4	$\alpha_{\mathcal{M}4}$: JawMaxillary	-0.340	0.193	0.079	-0.719	0.039
5	$\alpha_{\mathcal{M}5}$: JawMaxillary	-0.476	0.193	0.014	-0.854	-0.097
6	$\alpha_{\mathcal{M}6}$: JawMaxillary	0.190	0.195	0.330	-0.192	0.573
7	$\alpha_{\mathcal{M}7}$: JawMaxillary	1.404	0.236	< 0.001	0.941	1.866
8	$\alpha_{\mathcal{M}8}$: JawMaxillary	-0.033	0.177	0.850	-0.380	0.313
9	$\alpha_{\mathcal{M}9}$: JawMaxillary	0.001	0.198	0.996	-0.387	0.389
10	$\alpha_{\mathcal{M}10}$: JawMaxillary	-0.097	0.172	0.571	-0.434	0.240
11	$\alpha_{\mathcal{M}11}$: JawMaxillary	-0.264	0.212	0.212	-0.679	0.150
12	$\alpha_{\mathcal{M}12}$: JawMaxillary	-0.107	0.177	0.544	-0.453	0.239
1	$\alpha_{\mathcal{A}1}$: SideAnterior	0.177	0.139	0.203	-0.095	0.448
2	$\alpha_{\mathcal{A}2}$: SideAnterior	-0.031	0.130	0.813	-0.285	0.223
3	$\alpha_{\mathcal{A}3}$: SideAnterior	-0.009	0.143	0.951	-0.289	0.271
4	$\alpha_{\mathcal{A}4}$: SideAnterior	0.020	0.162	0.901	-0.297	0.338
5	$\alpha_{\mathcal{A}5}$: SideAnterior	0.335	0.186	0.072	-0.030	0.699
6	$\alpha_{\mathcal{A}6}$: SideAnterior	-0.426	0.271	0.116	-0.958	0.105
7	$\alpha_{\mathcal{A}7}$: SideAnterior	-0.464	0.230	0.043	-0.914	-0.014
8	$\alpha_{\mathcal{A}8}$: SideAnterior	0.024	0.148	0.872	-0.265	0.313
9	$\alpha_{\mathcal{A}9}$: SideAnterior	0.088	0.168	0.600	-0.241	0.418
10	$\alpha_{\mathcal{A}10}$: SideAnterior	0.236	0.151	0.118	-0.060	0.533
11	$\alpha_{\mathcal{A}11}$: SideAnterior	-0.116	0.184	0.529	-0.477	0.245
12	$\alpha_{\mathcal{A}12}$: SideAnterior	0.170	0.148	0.248	-0.119	0.460

Continued on next page

Participant #	Parameter	Estimate	Std. Error	Pr($>$ |z|)	2.5%	97.5%
1	$\alpha_{\mathcal{G}1}$: SideLeft	-0.009	0.112	0.934	-0.229	0.210
2	$\alpha_{\mathcal{G}2}$: SideLeft	-0.204	0.120	0.090	-0.439	0.032
3	$\alpha_{\mathcal{G}3}$: SideLeft	0.086	0.125	0.489	-0.159	0.331
4	$\alpha_{\mathcal{G}4}$: SideLeft	-0.034	0.149	0.817	-0.326	0.257
5	$\alpha_{\mathcal{G}5}$: SideLeft	0.293	0.158	0.065	-0.018	0.603
6	$\alpha_{\mathcal{G}6}$: SideLeft	-0.074	0.154	0.631	-0.375	0.227
7	$\alpha_{\mathcal{G}7}$: SideLeft	-0.313	0.200	0.118	-0.705	0.080
8	$\alpha_{\mathcal{G}8}$: SideLeft	0.091	0.122	0.458	-0.148	0.329
9	$\alpha_{\mathcal{G}9}$: SideLeft	0.290	0.134	0.030	0.028	0.552
10	$\alpha_{\mathcal{G}10}$: SideLeft	-0.025	0.120	0.836	-0.261	0.211
11	$\alpha_{\mathcal{G}11}$: SideLeft	-0.321	0.171	0.061	-0.657	0.015

Participant #	Parameter	Estimate	Std. Error	Pr($>$ |z|)	2.5%	97.5%
12	$\alpha_{\mathcal{G}12}$: SideLeft	0.249	0.137	0.068	-0.019	0.518

Brushing time categorized by different regions is shown in Figure 5.3. Participants did not vary significantly in the brushing times spent on the maxillary and mandibular regions or different sides (right, anterior, and left). However, participants differed in the times spent brushing various teeth surfaces with buccal surfaces brushed significantly more than the lingual and occlusal surfaces. As can be seen in Table 5.1, on average, buccal surfaces were brushed 2.18 times longer than the lingual surfaces (95% CI 1.42, 3.35; p $\leq$ 0.001) and 2.22 times longer than the occlusal surfaces (95% CI 1.62, 3.10; p $\leq$ 0.001).
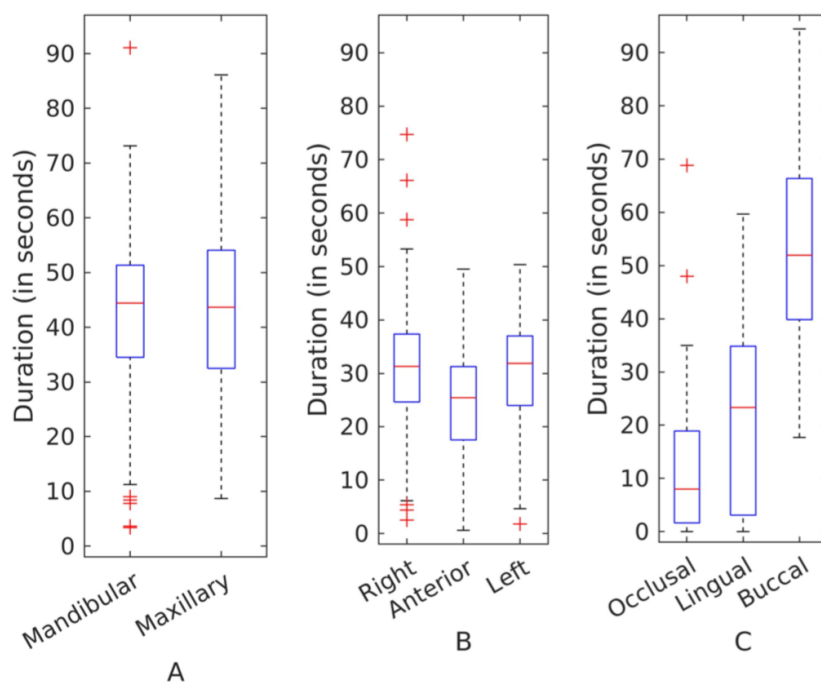
Figure 5.3: Group-level brushing times of different dental regions. Categorized by (A) jaw, (B) side, and (C) surface. Buccal surfaces were brushed twice more than the lingual and occlusal surfaces.

Table 5.3: Estimated participant-specific overdispersion parameters of count submodel for brushing duration of each dental surface.

Parameter	Coefficient	SE	95% CI	p
η_1 : Participant 1	1.23	0.12	$(1.00, 1.46)$	$< .001$
η_2 : Participant 2	1.55	0.14	$(1.29, 1.82)$	$< .001$
η_3 : Participant 3	1.16	0.13	$(0.91, 1.40)$	$< .001$
η_4 : Participant 4	0.84	0.14	$(0.58, 1.11)$	$< .001$
η_5 : Participant 5	0.64	0.13	$(0.38, 0.90)$	$< .001$
η_6 : Participant 6	0.74	0.16	$(0.43, 1.05)$	$< .001$
η_7 : Participant 7	1.05	0.19	$(0.68, 1.43)$	$< .001$
η_8 : Participant 8	0.77	0.12	$(0.54, 1.00)$	$< .001$
η_9 : Participant 9	0.56	0.12	$(0.31, 0.80)$	$< .001$
η_{10} : Participant 10	0.98	0.11	$(0.75, 1.20)$	$< .001$
η_{11} : Participant 11	0.72	0.14	$(0.44, 0.99)$	$< .001$
η_{12} : Participant 12	0.94	0.12	$(0.70, 1.18)$	$< .001$

Table 5.4: Estimated standard deviations of random effects of count submodel for brushing duration of each dental surface.

Parameter	Coefficient
σ_{α_0} : SD (Intercept: Participant)	0.34
σ_{γ} : SD (Intercept: Participant:Session)	$9.56e - 05$
$\sigma_{\mathcal{L}}$: SD (SurfaceLingual: Participant)	0.68
$\sigma_{\mathcal{O}}$: SD (SurfaceOcclusal: Participant)	0.54
$\sigma_{\mathcal{M}}$: SD (JawMaxillary: Participant)	0.47
$\sigma_{\mathcal{A}}$: SD (SideAnterior: Participant)	0.26
$\sigma_{\mathcal{G}}$: SD (SideLeft: Participant)	0.22

Table 5.5: Estimated standard deviations of random effects of zero-inflated submodel for brushing duration of each dental surface.

Parameter	Coefficient
σ_{ϵ} : SD (Intercept: Session:Participant)	0.12
σ_{δ} : SD (Intercept: Participant)	1.20

Table 5.7: Estimated participant-level random effects of zero-inflated submodel for brushing duration of each dental surface

Participant #	Parameter	Estimate	Std. Error	$\Pr(> |z|)$	2.5%	97.5%
1	δ_1 : (Intercept)	-1.467	0.449	0.001	-2.347	-0.587
2	δ_2 : (Intercept)	0.430	0.396	0.278	-0.347	1.207
3	δ_3 : (Intercept)	-0.306	0.405	0.450	-1.101	0.488
4	δ_4 : (Intercept)	0.541	0.396	0.172	-0.235	1.318
5	δ_5 : (Intercept)	0.151	0.399	0.704	-0.631	0.933
6	δ_6 : (Intercept)	1.162	0.397	0.003	0.385	1.939
7	δ_7 : (Intercept)	2.496	0.414	< 0.001	1.684	3.308
8	δ_8 : (Intercept)	-1.389	0.445	0.002	-2.261	-0.516
9	δ_9 : (Intercept)	-0.091	0.402	0.821	-0.879	0.697
10	δ_{10} : (Intercept)	-1.765	0.471	< 0.001	-2.688	-0.842
11	δ_{11} : (Intercept)	0.796	0.396	0.044	0.021	1.572
12	δ_{12} : (Intercept)	-0.399	0.407	0.327	-1.197	0.399

Between-individual variability in brushing duration

There was considerable between-individual variability in terms of brushing time devoted to different regions (Figure 5.4). Coefficient of variation for all the regions brushed during a session was greater than 20%. Some (e.g. participant 7) brushed their maxillary teeth much more than their mandibular teeth (see Table 5.6; $p \leq 0.001$). Others (e.g. participant 11), paid less attention to their lingual surfaces and focused primarily on the buccal surfaces (see Table 5.6; $p = 0.001$).

Within-individual variability in brushing duration

Participants varied greatly in their brushing of different regions across their brushing sessions. As exemplified by the brushing patterns of participant 2, the brushing duration for the lingual surfaces across sessions varied from none to 60 seconds (Figure 5.5).

From the ratio of the estimated standard deviations for participant and session level random effects on the intercept, $\frac{\widehat{\sigma_\gamma}}{\hat{\sigma}_{\alpha_0}} = \frac{9.56e-05}{0.34} < 0.1\%$ and $\frac{\hat{\sigma}_\epsilon}{\hat{\sigma}_\delta} = \frac{0.12}{1.20} = 10\%$, we can see that within-participant variability (session-to-session) is much smaller than between-participant variability, on the logmean scale.

We summarized coefficient of variation of all 16 dental regions of all 12 participants in Table 5.8 and illustrated in Fig 5.6.
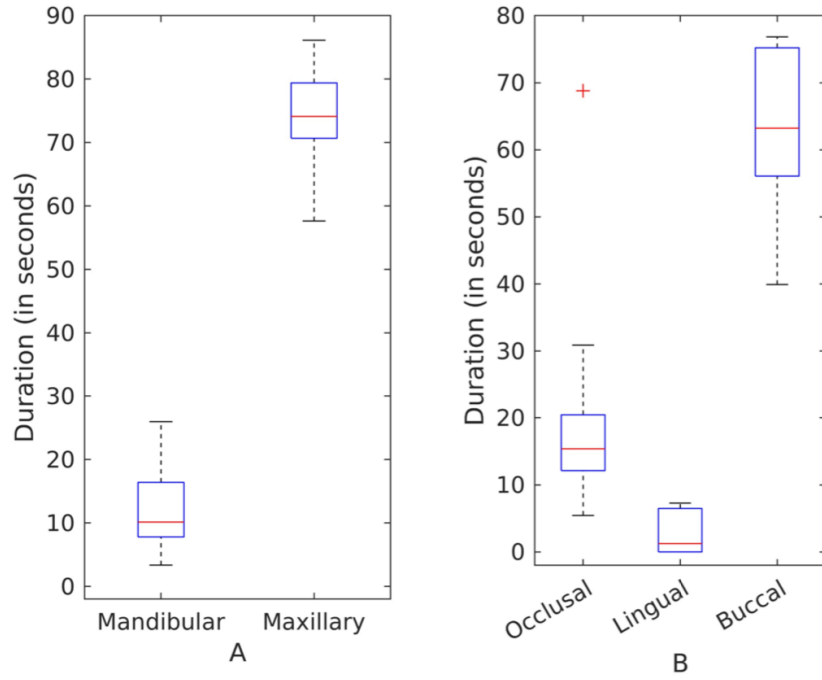
Figure 5.4: Between-individual variability in brushing duration by region. (A) by jaw and (B) by tooth surface.

Table 5.8: Coefficient of variation of all 16 dental regions for all participants

Participant	ManAB	ManAL	ManLB	ManLL	ManLO	ManRB	ManRL	ManRO	MaxAB	MaxAL	MaxLB	MaxLL	MaxLO	MaxRB	MaxRL	MaxR0
1	1.072	0.328	0.399	0.503	0.443	0.291	0.735	0.355	0.465	0.575	0.377	0.714	0.583	0.384	1.001	0.518
2	0.468	0.437	0.482	0.456	1.436	0.880	0.622	2.894	0.383	0.770	0.370	1.795	3.162	0.449	0.611	2.500
3	0.275	0.562	0.242	0.500	2.111	0.300	0.346	2.539	0.254	0.579	0.302	0.583	1.917	0.441	0.261	1.807
4	0.337	0.554	0.284	0.663	2.037	0.391	0.845	1.250	0.409	1.457	0.493	2.110	2.742	0.310	3.162	0
5	0.686	1.153	0.669	1.104	1.240	0.813	1.026	0	0.326	1.224	0.891	0.994	1.476	0.995	1.091	0.835
6	1.158	2.108	0.830	0	0.610	0.399	0	1.302	0.495	3.162	0.907	0	0.639	0.718	3.162	0.376
7	2.250	0	0.721	0	0	1.192	3.162	0.688	0.397	0	0.214	0	0	0.249	0	0
8	0.981	0.473	0.567	0.403	0.568	0.552	0.580	0.699	0.616	0.554	0.394	0.573	1.639	0.667	0.645	0.912
9	0.839	0.860	0.400	0.382	1.673	0.635	0.414	1.369	0.336	1.753	0.600	0.636	2.279	0.544	0.496	1.417
10	0.544	0.394	0.474	0.263	0.680	0.473	0.266	0.622	0.204	0.593	0.300	1.203	0.508	0.663	0.745	0.543
11	0.373	2.482	0.460	1.656	0.506	0.524	1.629	1.203	0.278	2.061	0.518	0	2.182	0.519	0	1.181
12	0.610	0.286	0.609	0.470	0.751	0.710	0.671	0.979	0.370	1.093	0.330	1.758	0.911	0.466	1.7623	1.050
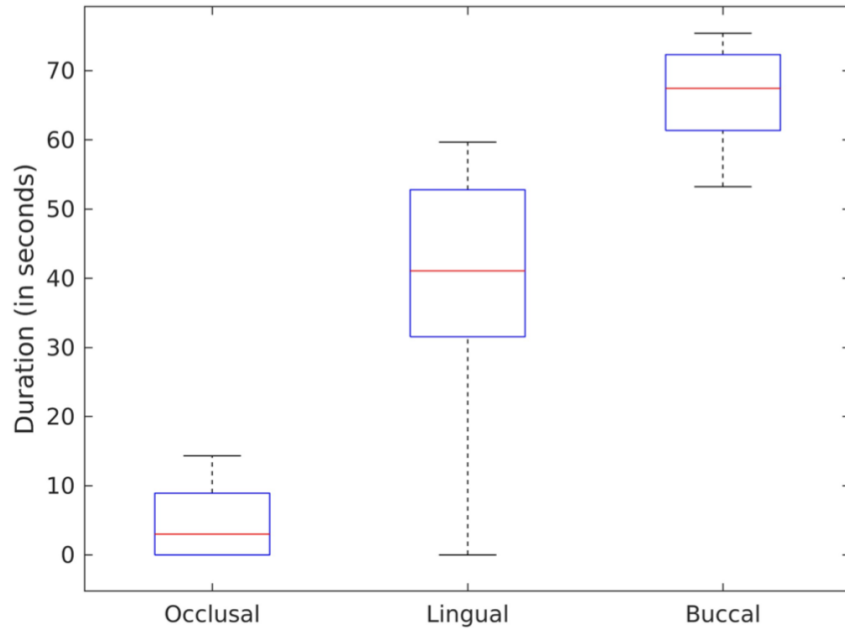
Figure 5.5: Within-individual variability in brushing duration for dental surfaces. Participant 2's brushing duration for the lingual surfaces varied greatly across sessions.

As it can be seen all the dental regions which are brushed at all, have a coefficient of variation of more than 20%.

5.2.2 Duration of excessive brushing pressure on each dental surface

Figure 5.7 summarizes the episodes of excess brushing pressure by region. About 16.7% of the participants exerted excessive brushing pressure of more than one second duration during each brushing session.

As in Subsection 5.2.1, we fit a zero-inflated negative binomial model; both the zero-inflation and count submodels include fixed effects for tooth surface, side, and jaw, and random effects on the intercept by participant and session. Extended models adding participant-level random effects on tooth surface, side and jaw and participant-specific overdispersion parameters failed to converge.
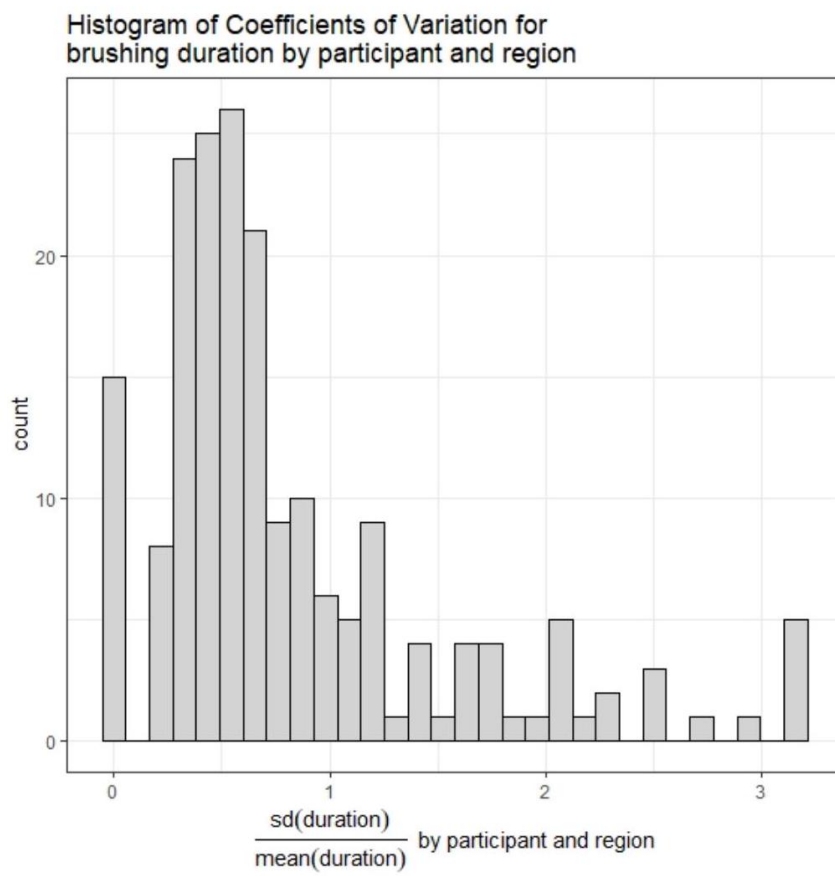
Figure 5.6: Coefficient of variation of brushing duration of all regions brushed by all participants
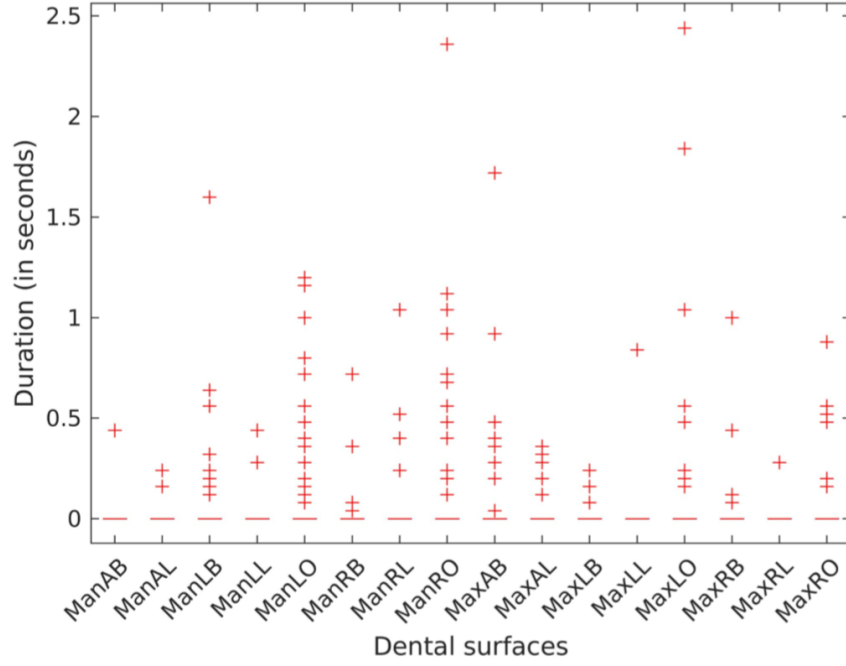
Figure 5.7: Episodes of excessive brushing pressure by region. Occlusal surfaces were most frequently brushed with excessive pressure.

Modeling

Our model is described in Equation 5.5

$$Z_{ijk} \sim \text{Bernoulli}\left(\pi_{ijk}\right)$$

$$Y_{ijk} \sim \begin{cases} 0, & Z_{ijk} = 1 \\ NB\left(\lambda_{ijk}, \kappa\right) & Z_{ijk} = 0 \end{cases} \tag{5.5}$$

Hence:

$$\lambda_{ijk} = \mathbb{E}\left(Y_{ijk} \mid Z_{ijk} = 0\right)$$

$$\text{Var}\left(Y_{ijk} \mid Z_{ijk} = 0\right) = \lambda_{ijk}\left(1 + \frac{\lambda_{ijk}}{\kappa}\right)$$

λ_{ijk} is modeled as:

$$\log\left(\lambda_{ijk}\right) = \beta_{0ij} + \beta_{\mathcal{M}}1\{k \in \mathcal{M}\} + \beta_{\mathcal{O}}1\{k \in \mathcal{O}\} + \beta_{\mathcal{L}}1\{k \in \mathcal{L}\} + \beta_{\mathcal{A}}1\{k \in \mathcal{A}\} + \beta_{\mathcal{G}}1\{k \in \mathcal{G}\}$$

$$\beta_{0ij} = \beta_0 + \alpha_{0i} + \gamma_{ij}$$

$$\gamma_{ij} \sim N\left(0, \sigma_\gamma^2\right)$$

$$\alpha_{0i} \sim N\left(0, \sigma_{\alpha_0}^2\right)$$

$$\log(\kappa) = \eta$$

$$(5.6)$$

We modeled π_{ijk} as in Equation 5.7

$$\text{logit}\left(\pi_{ijk}\right) = \theta_{0ij} + \theta_{\mathcal{M}}1\{k \in \mathcal{M}\} + \theta_{\mathcal{O}}1\{k \in \mathcal{O}\} + \theta_{\mathcal{L}}1\{k \in \mathcal{L}\} + \theta_{\mathcal{A}}1\{k \in \mathcal{A}\} + \theta_{\mathcal{G}}1\{k \in \mathcal{G}\}$$

$$\theta_{0ij} = \theta_0 + \delta_i + \epsilon_{ij}$$

$$\delta_i \sim N\left(0, \sigma_\delta^2\right)$$

$$\epsilon_{ij} \sim N\left(0, \sigma_\epsilon^2\right)$$

$$(5.7)$$

Results

Initially, we contemplated utilizing a Poisson distribution as the outcome for the count sub-model. Nevertheless, upon conducting a thorough comparison, we determined that the negative binomial distribution is more appropriate in this context. Our evaluation, based on the Akaike Information Criterion (AIC), revealed that the zero-inflated negative binomial model achieved an AIC of 1,252.2, representing a noteworthy 274.6 decrease compared to the AIC of the zero-inflated Poisson model with identical fixed and random effects. Moreover, when considering the Bayesian Information Criterion (BIC), the negative binomial model attained a BIC of 1,346.7, which exhibited a significant reduction of 269.0 in comparison to the BIC of the zero-inflated Poisson model with the same fixed and random effects. These results unequivocally demonstrate the superiority of the negative binomial distribution as it effectively addresses the potential issue of overfitting.

Table 5.9: Estimated fixed effects of count submodel for excessive brushing pressure duration on each dental surface.

Parameter	Log-Mean	SE	95% CI	p
β_0 : (Intercept)	2.12	0.30	$(1.53, 2.71)$	$< .001$
$\beta_{\mathcal{L}}$: Surface (Lingual)	0.01	0.29	$(-0.56, 0.58)$	0.971
$\beta_{\mathcal{O}}$: Surface (Occlusal)	0.54	0.22	$(0.10, 0.98)$	0.015
$\beta_{\mathcal{A}}$: Side (Anterior)	0.04	0.32	$(-0.58, 0.66)$	0.889
$\beta_{\mathcal{G}}$: Side (Left)	-0.04	0.20	$(-0.44, 0.37)$	0.862
$\beta_{\mathcal{M}}$: Jaw (Maxillary)	-0.10	0.21	$(-0.51, 0.30)$	0.612

Table 5.10: Estimated fixed effects of zero-inflation submodel for excessive brushing pressure duration on each dental surface.

Parameter	Log-Odds	SE	95%CI	p
θ_0 : Intercept)	4.87	0.93	$(3.04, 6.69)$	$< .001$
$\theta_{\mathcal{L}}$: Surface (Lingual)	0.93	0.34	$(0.27, 1.60)$	0.006
$\theta_{\mathcal{O}}$: Surface (Occlusal)	-1.01	0.29	$(-1.59, -0.44)$	$< .001$
$\theta_{\mathcal{A}}$: Side (Anterior)	-0.20	0.37	$(-0.93, 0.52)$	0.583
$\theta_{\mathcal{G}}$: Side (Left)	-0.38	0.27	$(-0.91, 0.14)$	0.151
$\theta_{\mathcal{M}}$: Jaw (Maxillary)	0.59	0.24	$(0.11, 1.07)$	0.016

The estimated parameters for the count and zero-inflated submodels for excessive brushing pressure duration are summarized in the tables in this subsection. Additionally, the estimated value for overdispersion parameter η was 0.72 .

As it can be seen in Table 5.9, the occlusal surfaces were most frequently brushed with excessive pressure (estimated log relative duration = 0.54; 95% CI 0.10, 0.98; p = 0.015).

5.2.3 Total active brushing duration of each brushing session

Modeling

As demonstrated in Equation 5.8, we fit a negative binomial regression model for the total active duration of each brushing session (measured in counts of 25 Hz samples), with a log-link

Table 5.11: Estimated standard deviations of random effects of count submodel for excessive brushing pressure duration on each dental surface.

Parameter	Coefficient
σ_{γ} : SD (Intercept: Session:Participant)	0.27
$\sigma_{\alpha 0}$: SD (Intercept: Participant)	0.38
κ : Overdispersion parameter	2.06

Table 5.12: Estimated standard deviations of random effects of zero-inflation submodel for excessive brushing pressure duration on each dental surface.

Parameter	Coefficient
σ_ϵ : SD (Intercept: Session:Participant)	0.69
σ_δ : SD (Intercept: Participant)	2.39

Table 5.13: Estimated participant-level random effects of count submodel for excessive brushing pressure duration on each dental surface.

Participant #	Parameter	Estimate	Std. Error	$Pr(> |z|)$	2.5%	97.5%
1	α_{01} : (Intercept)	0.289	0.232	0.213	-0.166	0.743
2	α_{02} : (Intercept)	-0.002	0.380	0.996	-0.747	0.743
3	α_{03} : (Intercept)	-0.002	0.380	0.996	-0.747	0.743
4	α_{04} : (Intercept)	-0.002	0.380	0.996	-0.747	0.743
5	α_{05} : (Intercept)	-0.158	0.302	0.600	-0.750	0.434
6	α_{06} : (Intercept)	-0.599	0.349	0.086	-1.283	0.084
7	α_{07} : (Intercept)	-0.002	0.380	0.996	-0.747	0.743
8	α_{08} : (Intercept)	0.206	0.346	0.552	-0.473	0.885
9	α_{09} : (Intercept)	-0.002	0.380	0.996	-0.747	0.743
10	α_{010} : (Intercept)	-0.080	0.254	0.753	-0.578	0.418
11	α_{011} : (Intercept)	0.277	0.254	0.276	-0.221	0.775
12	α_{012} : (Intercept)	-0.022	0.324	0.947	-0.656	0.613

Table 5.14: Estimated participant-level random effects of zero-inflated submodel for excessive brushing pressure duration on each dental surface.

Participant #	Parameter	Estimate	Std. Error	$Pr(> |z|)$	2.5%	97.5%
1	δ_1 : (Intercept)	-3.737	0.938	< 0.001	-5.576	-1.898
2	δ_2 : (Intercept)	1.594	1.576	0.312	-1.494	4.682
3	δ_3 : (Intercept)	1.594	1.576	0.312	-1.494	4.682
4	δ_4 : (Intercept)	1.594	1.576	0.312	-1.494	4.682
5	δ_5 : (Intercept)	-1.269	1.002	0.205	-3.233	0.695
6	δ_6 : (Intercept)	-2.172	0.963	0.024	-4.060	-0.284
7	δ_7 : (Intercept)	1.594	1.576	0.312	-1.494	4.682
8	δ_8 : (Intercept)	-0.329	1.089	0.763	-2.463	1.805
9	δ_9 : (Intercept)	1.594	1.576	0.312	-1.494	4.682
10	δ_{10} : (Intercept)	-2.207	0.956	0.021	-4.082	-0.333
11	δ_{11} : (Intercept)	-2.842	0.944	0.003	-4.693	-0.992
12	δ_{12} : (Intercept)	-0.717	1.044	0.492	-2.762	1.328

and random intercepts by participant ID. We also included a participant-specific overdispersion parameters to account for participant-to-participant differences in residual variance.

$$Y_{ij} \sim NB\left(\lambda_{ij}, \kappa_i\right) \tag{5.8}$$

Hence:

$$\lambda_{ij} = \mathbb{E}\left(Y_{ij}\right)$$

$$\mathrm{Var}\left(Y_{ij}\right) = \lambda_{ij}\left(1 + \frac{\lambda_{ij}}{\kappa_i}\right)$$

We modeled λ_{ij} as follows:

$$\log\left(\lambda_{ij}\right) = \beta_{0i}$$

$$\beta_{0i} = \beta_0 + \alpha_{0i}$$

$$\alpha_{0i} \sim N\left(0, \sigma_{\alpha_0}^2\right)$$

$$\log\left(\kappa_i\right) = \eta_i$$

Results

Figure 5.8 summarizes the active brushing duration for all participants. Most of the participants (91.67%) brushed less than the prescribed two minutes in all sessions. The mean brushing duration for a participant was 89.22 seconds.

Some participants (e.g. # 1 and 2) brushed for almost two minutes in most sessions, whereas others (e.g. participant 5) brushed for less than a minute. Some participants (e.g., # 2, 4, and 7) brushed consistently for nearly the same duration each time; others (e.g., # 5, 9, and 12) varied greatly (≥ 70 seconds) in their brushing duration.

Initially, we considered employing a Poisson distribution for the count submodel. However, upon analysis, we discovered that the negative binomial distribution is a more suitable choice. This determination was supported by the significant improvement observed in the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values. Specifically, the AIC for our chosen model was 1,802.112, which was substantially lower by 8,098.558 compared

86
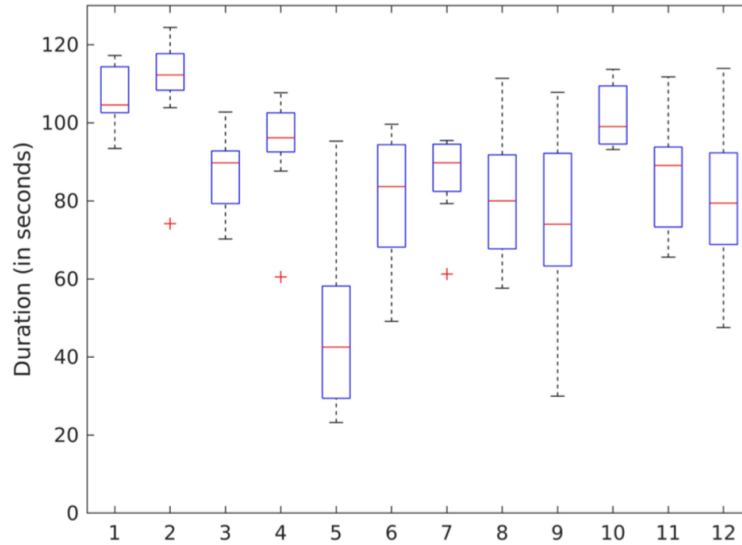
Figure 5.8: Brushing duration for the 12 participants. Active brushing duration is calculated by excluding pauses in brushing and the times transitioning the brush head to different regions.

Table 5.15: Estimated fixed effects of total active brushing in each brushing session

Parameter	Log-Mean	SE	95% CI	p
β_0 : (Intercept)	7.71	0.04	$(7.63, 7.79)$	$< .001$

to the zero-inflated Poisson model with identical fixed and random effects. Similarly, the BIC for our model yielded a value of 1,841.137, significantly reduced by 8,065.108 in comparison to the zero-inflated Poisson model with the same fixed and random effects. These findings indicate that our decision to opt for the negative binomial distribution was crucial to mitigate the risk of overfitting.

The estimated parameters are summarized in the tables of this subsection.

To find out the between- and within-person variabilities in active brushing duration, we have summarized the mean and standard deviations estimated from the model and calculated empirically from the samples in Table 5.19.

As it can be seen in Table 5.19, the empirical values and the model estimates are close. By calculating the standard deviation of the second column and the mean of the fourth column, we report the between- and within- person variability in brushing duration to be 16.69 and 14.50 seconds, respectively. By dividing the calculated between- and within-person variability

Table 5.16: Estimated participant-specific overdispersion parameter of total active brushing in each brushing session.

Parameter	Coefficient	SE	95%CI	p
η_1 : Participant1	5.34	0.51	$(4.34, 6.34)$	$< .001$
η_2 : Participant2	3.95	0.49	$(3.00, 4.91)$	$< .001$
η_3 : Participant3	4.32	0.48	$(3.37, 5.26)$	$< .001$
η_4 : Participant4	3.78	0.47	$(2.86, 4.71)$	$< .001$
η_5 : Participant5	0.97	0.52	$(-0.05, 1.99)$	0.064
η_6 : Participant6	2.97	0.47	$(2.06, 3.89)$	$< .001$
η_7 : Participant7	4.10	0.48	$(3.17, 5.04)$	$< .001$
η_8 : Participant8	3.19	0.47	$(2.27, 4.10)$	$< .001$
η_9 : Participant9	2.21	0.46	$(1.30, 3.12)$	$< .001$
η_{10} : Participant10	5.24	0.51	$(4.24, 6.23)$	$< .001$
η_{11} : Participant11	3.68	0.47	$(2.76, 4.60)$	$< .001$
η_{12} : Participant12	2.86	0.46	$(1.95, 3.77)$	$< .001$

Table 5.17: Estimated random effect standard deviation of total active brushing in each brushing session.

Parameter	Coefficient
σ_{α_0} : SD (Intercept: Participant)	0.12

Table 5.18: Estimated participant-level random effects of total active brushing per brushing session.

Participant #	Parameter	Estimate	Std. Error	$\Pr(> |z|)$	2.5%	97.5%
1	α_{01} : (Intercept)	0.172	0.048	< 0.001	0.078	0.266
2	α_{02} : (Intercept)	0.185	0.061	0.003	0.065	0.305
3	α_{03} : (Intercept)	-0.021	0.052	0.684	-0.123	0.081
4	α_{04} : (Intercept)	0.045	0.059	0.438	-0.069	0.160
5	α_{05} : (Intercept)	-0.145	0.154	0.346	-0.446	0.156
6	α_{06} : (Intercept)	-0.085	0.069	0.222	-0.220	0.051
7	α_{07} : (Intercept)	-0.028	0.054	0.604	-0.133	0.078
8	α_{08} : (Intercept)	-0.064	0.065	0.322	-0.192	0.063
9	α_{09} : Intercept)	-0.095	0.088	0.277	-0.267	0.076
10	α_{010} : (Intercept)	0.124	0.048	0.010	0.030	0.219
11	α_{011} : (Intercept)	-0.029	0.058	0.617	-0.143	0.085
12	α_{012} : (Intercept)	-0.071	0.071	0.318	-0.210	0.068

Table 5.19: Participant-specific estimates of active brushing duration.

	sample mean (seconds)	sample var (seconds ^)	sample sd (seconds)	lambda (samples)	lambda (seconds)	kappa	model var (seconds 2)	model sd (seconds)
1	107	58.2	7.63	2646	106	209	57.8	7.6
2	110	195	14	2680	107	52.2	224	15
3	87	103	10.1	2180	87.2	75.1	105	10.2
4	93.9	171	13.1	2330	93.2	44	201	14.2
5	46.7	477	21.8	1927	77.1	2.63	2263	47.6
6	79.3	292	17.1	2046	81.9	19.6	345	18.6
7	86.3	113	10.6	2166	86.6	60.6	127	11.3
8	81.9	287	16.9	2088	83.5	24.2	291	17.1
9	74.9	486	22	2025	81	9.15	720	26.8
10	101	59.9	7.74	2522	101	188	58.2	7.63
11	86.1	192	13.8	2163	86.5	39.6	192	13.9
12	80.5	368	19.2	2074	83	17.4	399	20

to the mean of the second column, we can calculate the coefficients of between- and within-person variability to be 0.19 and 0.17 which is more than 5%. Therefore, there was substantial between- and within- individual variability in brushing duration.

5.3 Discussion and Conclusion

In this chapter, we conducted an in-depth analysis of brushing behavior among a sample of 12 subjects, each of whom participated in 10 brushing sessions. Our primary objective was to model various aspects of brushing behavior, including the duration of brushing on individual dental surfaces, the total brushing duration, and the duration of excessive pressure applied to each dental surface.

Previous studies on brushing behavior analysis were limited in scope, as separate investigations were conducted to address specific questions regarding brushing behavior. For example, previous studies focused on determining the total duration of brushing or comparing the duration of brushing on specific dental regions within the mouth. These separate studies were necessary because they did not incorporate a statistical model to comprehensively analyze brushing behavior. Instead, the results were predominantly derived using non-parametric tests, such as the Wilcoxon test, which does not make assumptions about the underlying data distribution.

However, relying on the Wilcoxon test for each brushing behavior question had several drawbacks. Firstly, it required running separate tests for each question regardings brushing behavior, such as comparing the brushing of lingual (inner) teeth surfaces to occlusal (flat) teeth

surfaces. Moreover, since the Wilcoxon test does not assume a specific distribution, there was no exploration of a relevant distribution to model brushing behavior. Furthermore, the Wilcoxon test is not specifically designed for repeated measurement studies, such as brushing scenarios involving multiple participants, each participating in multiple brushing sessions. Attempting to compensate for this limitation by conducting studies with hundreds of participants, but with only one brushing session per participant (e.g., as observed in [81]), would result in considerable expenses and time consumption. In contrast, opting for a smaller group of participants with multiple sessions not only renders the analysis more feasible but also enables the examination of intra-subject brushing behavior, a crucial aspect that has been lacking in previous literature.

To overcome these limitations, we undertook a systematic modeling of brushing behavior in a free-form brushing scenario, utilizing generalized mixed-effect linear regression models. Our analysis unveiled that the distribution of brushing times on each dental surface demonstrated zero-inflation, indicating the necessity of incorporating a zero-inflation submodel into our proposed framework. To capture the outcomes, measured as counts of 25Hz samples, we employed zero-inflated generalized linear mixed-effects regression models. The count submodels were fitted with a log-link and a negative binomial outcome distribution, while the zero-inflation submodels utilized a logistic link and a Bernoulli outcome distribution.

Our model offers several advantages, including its ability to handle multiple measurement studies and incorporate random factors to analyze inter- and intra-subject variability in brushing behavior. For instance, we incorporated random effects on the intercept by session nested in participant, allowing us to account for differences in overall brushing duration between individuals and sessions. Furthermore, we introduced person-specific over-dispersion parameters to account for variations in residual variance among participants. In the count submodel for region-specific brushing duration, we also included participant-specific random effects for tooth surface, mouth side, and jaw.

Moreover, in a novel contribution, we analyzed brushing behavior when excessive pressure was applied using powered/electric toothbrushes. This aspect had not been explored in previous studies, further enhancing the comprehensiveness of our analysis.

Our study revealed that brushing patterns with powered toothbrushes in the home setting

varied greatly between individuals as well as within individuals. Most participants brushed their teeth for less than the prescribed two minutes. Most spent less than 85 seconds per brushing session, a finding consistent with other studies that indicate that patients often fail to adhere to the recommended brushing time of two minutes [19]. Even individual participants were inconsistent in the total duration of time they spent brushing over different days. The times devoted to brushing different dental regions are quite unequal at the individual level with certain regions receiving more attention than the others. Participants generally brushed the buccal tooth surfaces twice as long as the occlusal and lingual surfaces; the lingual surfaces of the maxillary molars were often neglected.

Our observation that participants tended to brush tooth surfaces unequally substantiates other studies [19, 28, 66, 81]. Whereas some studies found that lingual surfaces are brushed significantly less than both occlusal and buccal surfaces [19,66], we did not find any statistically significant difference between the times spent brushing the occlusal and lingual surfaces (see Table 5.1).

Also, participants varied in the epochs of excessive brushing pressure and the regions to which they were applied. In general, the occlusal surfaces were more likely to be brushed with excessive pressure. Only a subset (16.7%) of our participants exerted excessive brushing pressures and these were mostly transient ($\sim$ 1 second). This finding contrasts Janusz et al. [40] who reported that 46.3% of their participants exerted excess brushing pressures for more than four seconds within a two-minute brushing session.

Our technology-facilitated study has several strengths. Unlike previous observational studies with limited external validity, our naturalistic study is more representative of real-world brushing behaviors and patterns. Study limitations include the fact that the study participants were recruited as a convenience sample and involved a young well-educated and dentally-aware group of college students. This may limit the generalizability of our findings to an older and more socioeconomically diverse population.

Our study uncovers substantial variations in the application of powered toothbrushes among users, highlighting a departure from recommended brushing practices. With our framework, we are able to derive statistically significant insights about brushing behavior at both the

population and individual levels, thereby paving the way for personalized brushing behavior recommendations. The identification of inconsistent brushing patterns also creates an opportunity for providing tailored feedback and behavioral nudges to individuals who neglect to brush specific dental quadrants for a sufficient time or overlook certain dental regions. This personalized approach can enhance brushing efficiency. In related work, the authors in [74,75] explored the application of reinforcement learning to assess brushing behavior, specifically focusing on the total duration of brushing sessions. However, further studies are warranted to incorporate more specific brushing behavior recommendations tailored to individuals, addressing aspects that are most important for their oral hygiene.

Chapter 6

Conclusion and Future Research Directions

In this dissertation, we presented a method to detect the regions of the mouth that are being brushed during a brushing session and analysed brushing behaviors using motion sensors (i.e. accelerometer, gyroscope, and magnetometer sensors). Previous research studies show promising preliminary results; however, they were conducted under structured toothbrushing assumptions performed in controlled laboratory settings (e.g. limited head and body movement, predefined sequence of brushing) and not the free-form brushing observed in real-world settings. To address the aforementioned issue, we collected a dataset of 187 brushing sessions, including free-form brushing, and present, to the best of our knowledge, the first motion-sensor dataset obtained during free-form brushing. The labeling of these brushing sessions is done using concurrent video recordings as the ground truth. However, due to the rapid and frequent transitions of the brush, labeling is susceptible to errors, which is known as a noisy label scenario in machine learning. To address this, we introduced a relabeling algorithm aimed at cleaning the dataset.

We conducted an extensive statistical analysis of brushing behavior, focusing on the duration of brushing on individual dental surfaces, total brushing duration, and excessive pressure applied to each dental surface. Previous studies on brushing behavior were limited as they relied on separate investigations and non-parametric tests like the Wilcoxon test [28, 81], lacking a comprehensive statistical model. This approach had drawbacks, requiring separate tests for each brushing behavior related question and neglecting the exploration of a suitable distribution for modeling brushing behavior. Additionally, the Wilcoxon test was unsuitable for

repeated measurement studies like ours, involving multiple participants brushing for multiple sessions. To address these limitations, we developed a systematic framework using generalized mixed-effect linear regression models. Our analysis revealed zero-inflation in the distribution of brushing times, necessitating the inclusion of a zero-inflation submodel. We employed zero-inflated generalized linear mixed-effects regression models, with log-link and negative binomial distribution for count submodels, and logistic link and Bernoulli distribution for zero-inflation submodels. Our model offers several advantages, accommodating multiple measurement studies and incorporating random factors to capture inter- and intra-subject variability in brushing behavior. We accounted for variations in overall brushing duration and residual variance through participant-specific random effects. Additionally, we examined brushing behavior under excessive pressure using powered/electric toothbrushes, a novel contribution to the field. Our findings revealed that individuals tended to spend over twice as much time brushing their buccal teeth surfaces (outer teeth surfaces) compared to the occlusal (flat teeth surfaces) (2.18 times longer, 95% CI 1.42, 3.35; p ≤ 0.001) and lingual surfaces (inner teeth surfaces) (2.22 times longer, 95% CI 1.62, 3.10; p ≤ 0.001). The study highlighted substantial variations in brushing patterns with powered toothbrushes in a home setting, both between individuals and within individuals. Most participants fell short of the recommended two-minute brushing time, typically spending less than 85 seconds per session. These findings align with previous studies that indicate patients often fail to adhere to the recommended brushing duration. Interestingly, even individual participants exhibited inconsistencies in the total duration of their brushing sessions across different days. Moreover, there was a notable imbalance in the time allocated to brushing different dental regions, with certain regions receiving more attention than others. Specifically, participants dedicated twice as much time to brushing buccal surfaces compared to occlusal and lingual surfaces, with the lingual surfaces of the maxillary molars often being neglected. Additionally, the occlusal surfaces were more likely to be brushed with excessive pressure. The findings of our study highlight significant variations in the application of powered toothbrushes among users, deviating from recommended brushing practices. The observed inconsistencies in brushing patterns emphasize the need for personalized feedback regarding individual brushing habits.

We thoroughly discussed the limitations associated with accurate position tracking of the toothbrush due to the geometric dimensions of the mouth. Additionally, we conducted a comprehensive analysis and comparison of various orientation estimation algorithms. Consequently, we concluded that relying solely on brush head orientation is not a reliable method for accurately detecting brushing regions, mainly due to challenges such as the similarity of brush head orientations in certain regions.

To overcome these challenges, we proposed a three-stage algorithm, consisting of preprocessing, brush transition times detection, and time-series classification, to infer brushing regions. Our algorithm effectively identifies the instances when the brush transitions between dental regions and accurately classifies the segments between consecutive brush transition times. In addition, we introduced CluMing, a novel method for change point detection based on time-series clustering, which exhibited superior performance in our application.

Furthermore, we compared the results obtained from multiple time-series classification methods in machine learning, including feature engineering methods, LSTM, and Transformer models. We observed that free-style brushing poses more challenges compared to the dataset presented in [37], which was collected under structured brushing assumptions. Additionally, we demonstrated that classifiers can achieve high classification accuracy using a random train-test split (i.e., k-fold cross-validation). However, generalization to new toothbrush users (i.e., leave-one-subject-out cross-validation) presents significant challenges due to variations in brushing styles and other factors, such as head movement during brushing.

Our results signifies that generalization to specific manifolds of data (e.g. one-subject-out cross-validation), is a more challenging problem in cases such as toothbrushing region prediction which involves significant between- and with- subject variability. Considering the complexity and the variations included in our problem, we think that a bigger dataset (our speculation is at least 20 times bigger) is needed for the models to be able to generalize to the subjects outside our datasets' participants. In small data scenarios such as our case, we also hypothesize that other avenues of machine learning which are suitable for out-of-distribution genralization, such as transfer learning [79], can possibly be effective. This may need training models on other datasets that involve IMU sensors for detection of other activities other than toothbrushing.

Another potential avenue for future work involves applying continuous Hidden Markov Models (HMMs) to the brushing region detection problem. HMMs learn from the sequence of states and their corresponding observations using the Baum-Welch algorithm [21] to be able to predict the most probable sequence of states based on the observations recorded using the Viterbi algorithm. Figures 6.1 and 6.2 present the transition matrix and transition graph for the sequence of brushing regions in our dataset, respectively. To generate this graph, we used all the data collected during the free-style brushing sessions and removed any transition probabilities below a threshold of 0.1 to simplify the representation. In the context of HMMs for this problem, the states correspond to brushing regions, and there are two sets of observations:

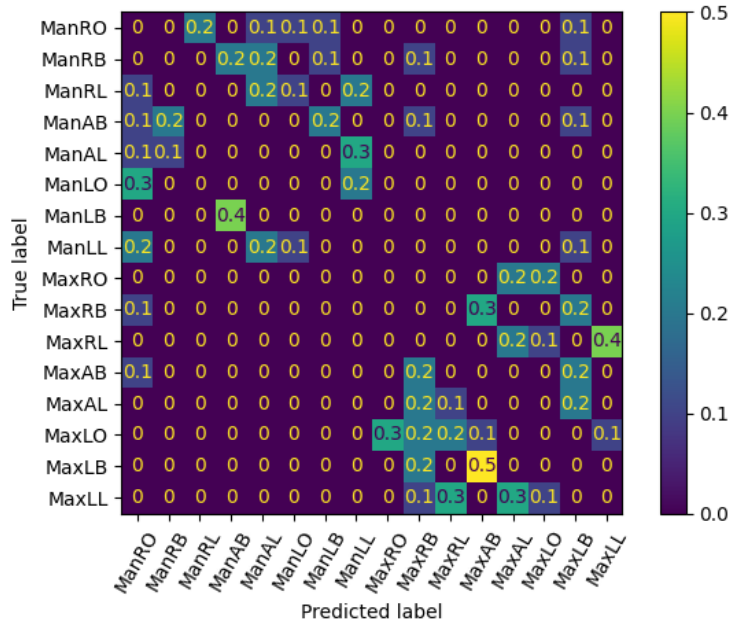| True label \ Predicted label	ManRO	ManRB	ManRL	ManAB	ManAL	ManLO	ManLB	ManLL	MaxRO	MaxRB	MaxRL	MaxAB	MaxAL	MaxLO	MaxLB	MaxLL
ManRO	0	0	0.2	0	0.1	0.1	0.1	0	0	0	0	0	0	0	0.1	0
ManRB	0	0	0	0.2	0.2	0	0.1	0	0	0.1	0	0	0	0	0.1	0
ManRL	0.1	0	0	0	0.2	0.1	0	0.2	0	0	0	0	0	0	0	0
ManAB	0.1	0.2	0	0	0	0	0.2	0	0	0.1	0	0	0	0	0.1	0
ManAL	0.1	0.1	0	0	0	0	0	0.3	0	0	0	0	0	0	0	0
ManLO	0.3	0	0	0	0	0	0	0.2	0	0	0	0	0	0	0	0
ManLB	0	0	0	0.4	0	0	0	0	0	0	0	0	0	0	0	0
ManLL	0.2	0	0	0	0.2	0.1	0	0	0	0	0	0	0	0	0.1	0
MaxRO	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.2	0	0
MaxRB	0.1	0	0	0	0	0	0	0	0	0	0	0	0.3	0	0.2	0
MaxRL	0	0	0	0	0	0	0	0	0	0	0	0	0.2	0.1	0	0.4
MaxAB	0.1	0	0	0	0	0	0	0	0	0.2	0	0	0	0	0.2	0
MaxAL	0	0	0	0	0	0	0	0	0	0.2	0.1	0	0	0	0.2	0
MaxLO	0	0	0	0	0	0	0	0	0.3	0.2	0.2	0.1	0	0	0	0.1
MaxLB	0	0	0	0	0	0	0	0	0	0.2	0	0	0.5	0	0	0
MaxLL	0	0	0	0	0	0	0	0	0	0.1	0.3	0	0.3	0.1	0	0

Figure 6.1: Transition matrix of brushing region sequences. Probabilities less than 0.1 are removed for better representation.

1. Motion sensor signals during brushing of a single region. Each observation is associated with only the corresponding brushing region.

2. Motion sensor signals at the time of a brush transition from one dental region to another. Each observation is associated with two brushing regions: the region from which the brush
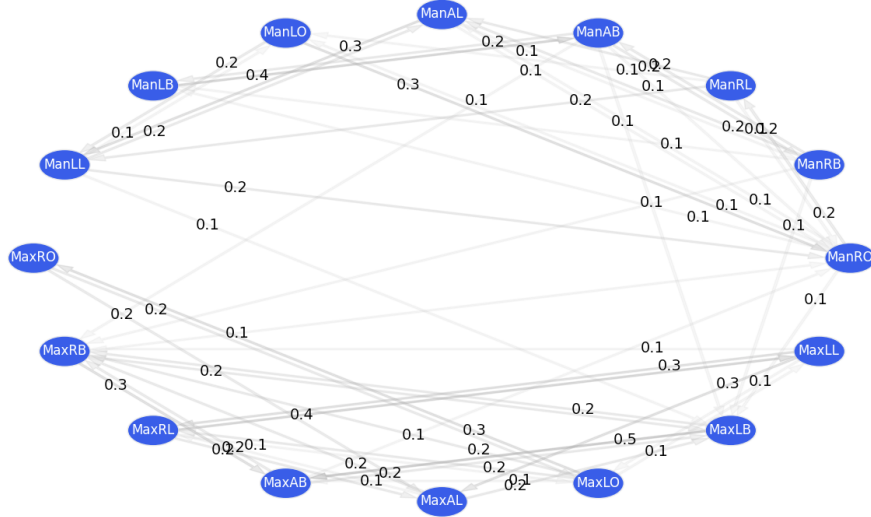
96

Figure 6.2: Transition graph of brushing region sequences. Probabilities less than 0.1 are removed for better representation.

is transitioning and the region to which it is transitioning.

In this dissertation, we mainly used the first set of observations to detect brushing regions. However, estimating brushing regions based solely on the first set of observations may not be effective due to challenges such as person-to-person and session-to-session variations, as well as head movement, as mentioned in Section 3.3. To address these challenges, we hypothesize that the second set of observations, which consider the relative changes in brush orientation/position, can be useful in detecting brushing regions by identifying the type of brush movements, such as brush flips (up-down or down-up rotations) and side-to-side transitions (left-to-right or right-to-left transitions). To support this hypothesis, we extracted the Euler Angles changes of the brush orientation during transitions from left-to-right and right-to-left, as shown in Figure 6.3. As observed, the psi angle change for transitions from left to right is positive, while it is negative for the reverse transition. We estimated the EA change by integrating the gyroscope values for the motion sense samples during the transition period, which we refer to as the transition span, as marked in Figure 6.4. We hypothesize that while the absolute value of the psi angle may not be relevant as mentioned in Section 3.3, the relative change of psi can be useful in estimating brushing regions. Considering the confusion matrix presented in Section 4.4 and the

97

confusion between detecting the regions on the left or right side of the mouth, we hypothesize that side-to-side transitions can aid in brushing region detection. For instance, the confusion in estimating ManRO and ManLO can be resolved if the next brushing region is ManAL because one region consists of a left-to-right transition and the other consists of a right-to-left transition. However, there are some challenges in using HMMs for this problem, such as:

1. Accurately estimating the transition span is critical for the success of this algorithm. Integrating gyroscope values outside the transition span can lead to incorrect estimation of the EA change, which can in turn result in inaccurate estimation of the brushing regions and a cascade of errors in estimating other brushing regions in the session.

2. Observations in HMMs are typically associated with a single state, whereas the second set of observations in our problem includes two states (i.e., the region that the brush transitions from and the one brush is transitioning to) and thus requires an extension of HMMs to model the second set of observations discussed earlier. One potential approach is to stack the first and second set of observations as the observations associated with the states.

3. However, it should be noted that not all confusions can be resolved by considering side-to-side transitions. For example, if the confusion is between ManRO and ManAL regions and the subsequent region is ManLO, the side-to-side transition from both ManRO and ManAL to ManLO is the same.

In this dissertation, we presented a novel approach to detect brushing regions using motion sensors and extract valuable brushing habit data. Building upon this foundation, future studies can focus on determining the information that resonates most positively with users and effectively motivates them to improve their brushing habits. By exploring the key elements of brushing habit information that hold the greatest significance for brush users, we can pave the way for personalized interventions that foster positive behavior change and promote optimal oral hygiene practices.
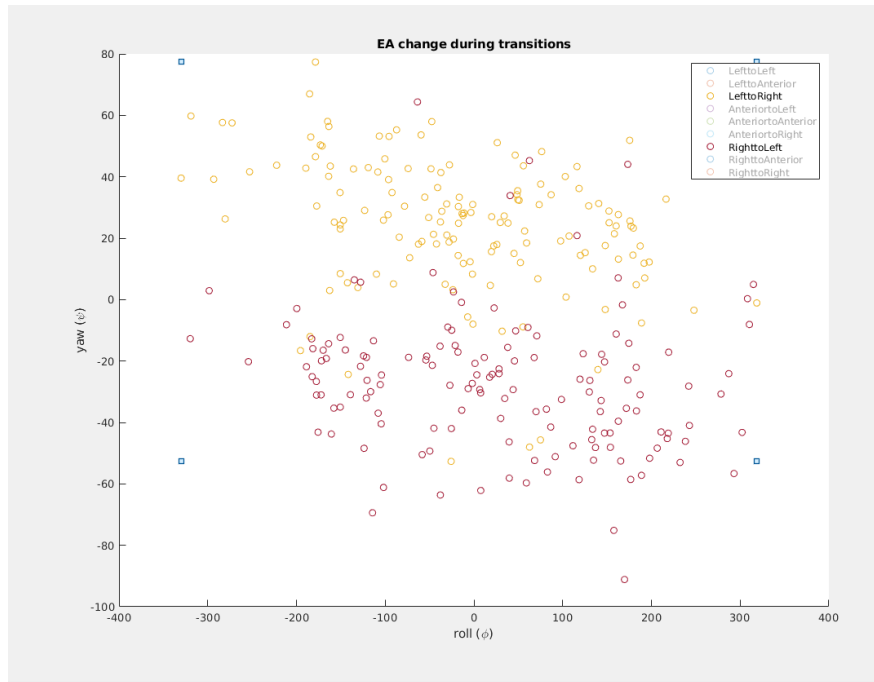
Figure 6.3: EAs change during the brush transitions. Left-to-right transitions and right-to-left transitions indicate different ψ changes.
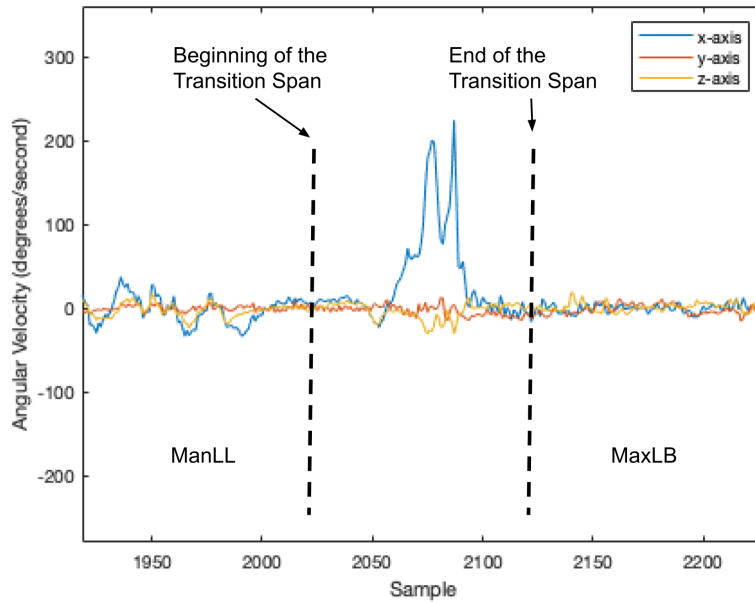

Figure 6.4: Transition Span for a transition of the brush from ManLL to MaxLB

Bibliography

[1] Alessandro Achille, Matteo Rovere, and Stefano Soatto. Critical learning periods in deep neural networks. *arXiv preprint arXiv:1711.08856*, 2017.

[2] Ryan Prescott Adams and David JC MacKay. Bayesian online changepoint detection. *arXiv preprint arXiv:0710.3742*, 2007.

[3] M. Ahad and Gheena Sukumaran. Awareness of tooth brushing techniques and proper oral hygiene among school children. *Journal of Pharmaceutical Sciences and Research*, 7:367–372, 01 2015.

[4] Sayma Akther, Nazir Saleheen, Mithun Saha, Vivek Shetty, and Santosh Kumar. mteeth: Identifying brushing teeth surfaces using wrist-worn inertial sensors. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, 5(2):1–25, 2021.

[5] Sayma Akther, Nazir Saleheen, Shahin Alan Samiei, Vivek Shetty, Emre Ertin, and Santosh Kumar. moral: An mhealth model for inferring oral hygiene behaviors in-the-wild using wrist-worn inertial sensors. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 3(1):1–25, 2019.

[6] David Arthur and Sergei Vassilvitskii. k-means++: The advantages of careful seeding. Technical report, Stanford, 2006.

[7] Sebahattin Bektas. Orthogonal distance from an ellipsoid. *Boletim de Ciências Geodésicas*, 20(4):970–983, 2014.

[8] Yoshua Bengio and Yves Grandvalet. No unbiased estimator of the variance of k-fold cross-validation. *Journal of machine learning research*, 5(Sep):1089–1105, 2004.

[9] Donald J Berndt and James Clifford. Using dynamic time warping to find patterns in time series. In *KDD workshop*, volume 10, pages 359–370. Seattle, WA, USA:, 1994.

[10] Leif Boysen, Angela Kempe, Volkmar Liebscher, Axel Munk, Olaf Wittich, et al. Consistencies and rates of convergence of jump-penalized least squares estimators. *The Annals of Statistics*, 37(1):157–183, 2009.

[11] Mollie E Brooks, Kasper Kristensen, Koen J van Benthem, Arni Magnusson, Casper W Berg, Anders Nielsen, and Hans J Skaug. glmmtmb balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *The R Journal*, 9(2):378–400, 2017.

[12] Barbara Bruno, Fulvio Mastrogiovanni, and Antonio Sgorbissa. A public domain dataset for adl recognition using wrist-placed accelerometers. In *the 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 738–743. IEEE, 2014.

[13] Barbara Bruno, Fulvio Mastrogiovanni, Antonio Sgorbissa, Tullio Vernazza, and Renato Zaccaria. Human motion modelling and recognition: A computational approach. In *2012 IEEE International Conference on Automation Science and Engineering (CASE)*, pages 156–161. IEEE, 2012.

[14] Barbara Bruno, Fulvio Mastrogiovanni, Antonio Sgorbissa, Tullio Vernazza, and Renato Zaccaria. Analysis of human behavior recognition algorithms based on acceleration data. In *2013 IEEE International Conference on Robotics and Automation*, pages 1602–1607. IEEE, 2013.

[15] Y Chan and Randy Langford. Spectral estimation via the high-order yule-walker equations. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(5):689–698, 1982.

[16] Yu-Chen Chang, Jin-Ling Lo, Chao-Ju Huang, Nan-Yi Hsu, Hao-Hua Chu, Hsin-Yen Wang, Pei-Yu Chi, and Ya-Lin Hsieh. Playful toothbrush: ubicomp technology for teaching tooth brushing to kindergarten children. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 363–372, 2008.

[17] Seong Yun Cho and Chan Gook Park. A calibration technique for a two-axis magnetic compass in telematics devices. *ETRI journal*, 27(3):280–288, 2005.

[18] Joseph M Cooke, Michael J Zyda, David R Pratt, and Robert B McGhee. Npsnet: Flight simulation dynamic modeling using quaternions. *Presence: Teleoperators & Virtual Environments*, 1(4):404–420, 1992.

[19] Renate Deinzer, Stefanie Ebel, Helen Blättermann, Ulrike Weik, and Jutta Margraf-Stiksrud. Toothbrushing: to the best of one's abilities is possibly not good enough. *BMC Oral Health*, 18:1–7, 2018.

[20] James Diebel. Representing attitude: Euler angles, unit quaternions, and rotation vectors. *Matrix*, 58(15-16):1–35, 2006.

[21] Sean R Eddy. Hidden markov models. *Current opinion in structural biology*, 6(3):361–365, 1996.

[22] Mahmoud Essalat and Douglas Morrison. Brushing-Behavior, January 2022.

[23] Mark Euston, Paul Coote, Robert Mahony, Jonghyuk Kim, and Tarek Hamel. A complementary filter for attitude estimation of a fixed-wing uav. In *2008 IEEE/RSJ international conference on intelligent robots and systems*, pages 340–345. IEEE, 2008.

[24] Centers for Disease Control, Prevention (CDC), et al. Division of oral health. chronic disease–oral health–at a glance, 2014.

[25] Emily B Fox, Erik B Sudderth, Michael I Jordan, and Alan S Willsky. The sticky hdp-hmm: Bayesian nonparametric hidden markov models with persistent states. *Arxiv preprint*, 2007.

[26] Eric Foxlin. Inertial head-tracker sensor fusion by a complementary separate-bias kalman filter. In *Proceedings of the IEEE 1996 Virtual Reality Annual International Symposium*, pages 185–194. IEEE, 1996.

[27] Valerie Galluzzi, Ted Herman, and Philip Polgreen. Hand hygiene duration and technique recognition using wrist-worn sensors. In *Proceedings of the 14th International Conference on Information Processing in Sensor Networks*, pages 106–117, 2015.

[28] Carolina Ganss, R Duran, Tobias Winterfeld, and N Schlueter. Tooth brushing motion patterns with manual and powered toothbrushes—a randomised video observation study. *Clinical oral investigations*, 22:715–720, 2018.

[29] William V Giannobile, Thomas Beikler, Janet S Kinney, Christoph A Ramseier, Thiago Morelli, and David T Wong. Saliva as a diagnostic tool for periodontal disease: current state and future directions. *Periodontology 2000*, 50:52, 2009.

[30] Cyril Goutte and Eric Gaussier. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *Advances in Information Retrieval: 27th European Conference on IR Research, ECIR 2005, Santiago de Compostela, Spain, March 21-23, 2005. Proceedings 27*, pages 345–359. Springer, 2005.

[31] Grafana Labs. Grafana. https://grafana.com/, Accessed 2023.

[32] William Rowan Hamilton. Lxxviii. on quaternions; or on a new system of imaginaries in algebra: To the editors of the philosophical magazine and journal. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 25(169):489–495, 1844.

[33] Everett B Hancock. Question set supports that gingivitis. *Ann Periodontol*, 1:223–249, 1996.

[34] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[35] Hua Huang and Shan Lin. Toothbrushing monitoring using wrist watch. In *Proceedings of the 14th ACM Conference on Embedded Network Sensor Systems CD-ROM*, pages 202–215, 2016.

[36] Hua Huang and Shan Lin. Toothbrushing recognition using neural networks. In *2017 IEEE/ACM Second International Conference on Internet-of-Things Design and Implementation (IoTDI)*, pages 309–310. IEEE, 2017.

[37] Zawar Hussain, David Waterworth, Murtadha Aldeer, Wei Emma Zhang, and Quan Z Sheng. Toothbrushing data and analysis of its potential use in human activity recognition

applications: dataset. In *Proceedings of the Third Workshop on Data: Acquisition To Analysis*, pages 31–34, 2020.

[38] Zawar Hussain, David Waterworth, Murtadha Aldeer, Wei Emma Zhang, Quan Z Sheng, and Jorge Ortiz. Do you brush your teeth properly? an off-body sensor-based approach for toothbrushing monitoring. In *2021 IEEE International Conference on Digital Health (ICDH)*, pages 59–69. IEEE, 2021.

[39] Emi Inada, Issei Saitoh, Yong Yu, Daisuke Tomiyama, Daisuke Murakami, Yoshihiko Takemoto, Ken Morizono, Tomonori Iwasaki, Yoko Iwase, and Youichi Yamasaki. Quantitative evaluation of toothbrush and arm-joint motion during tooth brushing. *Clinical oral investigations*, 19(6):1451–1462, 2015.

[40] Karen Janusz, Bruce Nelson, Robert D Bartizek, Patricia A Walters, and Aaron R Biesbrock. Impact of a novel power toothbrush with smartguide technology on brushing pressure and thoroughness. *J Contemp Dent Pract*, 9(7):1–8, 2008.

[41] Xin Jin and Jiawei Han. K-means clustering., 2010.

[42] Rebecca Killick, Kaylea Haynes, Idris Eckley, Paul Fearnhead, and Jamie Lee. Package 'changepoint'. *R package version 0.4.-2011.-http://cran. rproject. org/web/packages/changepoint/index. html*, 2016.

[43] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[44] Manon Kok, Jeroen D Hol, and Thomas B Schön. Using inertial sensors for position and orientation estimation. *arXiv preprint arXiv:1704.06053*, 2017.

[45] Joseph Korpela, Ryosuke Miyaji, Takuya Maekawa, Kazunori Nozaki, and Hiroo Tamagawa. Evaluating tooth brushing performance with smartphone sound data. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 109–120, 2015.

[46] Alex H Krist, Sebastian T Tong, Rebecca A Aycock, and Daniel R Longo. Engaging patients in decision-making and behavior change to promote prevention. *Information Services & Use*, 37(2):105–122, 2017.

[47] Jeong-Whan Lee, Kang-Hwi Lee, Kyeong-Seop Kim, Dong-Jun Kim, and Kyungho Kim. Development of smart toothbrush monitoring system for ubiquitous healthcare. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6422–6425. IEEE, 2006.

[48] Kang-Hwi Lee, Jeong-Whan Lee, Kyeong-Seop Kim, Dong-Jun Kim, Kyungho Kim, Heui-Kyung Yang, Keesam Jeong, and Byungchae Lee. Tooth brushing pattern classification using three-axis accelerometer and magnetic sensor for smart toothbrush. In *2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4211–4214. IEEE, 2007.

[49] Young-Jae Lee, Pil-Jae Lee, Kyeong-Seop Kim, Wonse Park, Kee-Deog Kim, Dosik Hwang, and Jeong-Whan Lee. Toothbrushing region detection using three-axis accelerometer and magnetic sensor. *IEEE Transactions on Biomedical Engineering*, 59(3):872–881, 2011.

[50] Tao Li, Shenghuo Zhu, and Mitsunori Ogihara. Using discriminant analysis for multi-class classification: an experimental investigation. *Knowledge and information systems*, 10(4):453–472, 2006.

[51] Henk J Luinge, Peter H Veltink, and Chris TM Baten. Estimation of orientation with gyroscopes and accelerometers. In *Proceedings of the First Joint BMES/EMBS Conference. 1999 IEEE Engineering in Medicine and Biology 21st Annual Conference and the 1999 Annual Fall Meeting of the Biomedical Engineering Society (Cat. N*, volume 2, pages 844– vol. IEEE, 1999.

[52] Chengwen Luo, Xingyu Feng, Junliang Chen, Jianqiang Li, Weitao Xu, Wei Li, Li Zhang, Zahir Tari, and Albert Y Zomaya. Brush like a dentist: Accurate monitoring of toothbrushing via wrist-worn gesture sensing. In *IEEE INFOCOM 2019-IEEE Conference on Computer Communications*, pages 1234–1242. IEEE, 2019.

[53] Sebastian Madgwick. An efficient orientation filter for inertial and inertial/magnetic sensor arrays. *Report x-io and University of Bristol (UK)*, 25:113–118, 2010.

[54] Robert Mahony, Tarek Hamel, and Jean-Michel Pflimlin. Nonlinear complementary filters on the special orthogonal group. *IEEE Transactions on automatic control*, 53(5):1203–1218, 2008.

[55] João Luís Marins, Xiaoping Yun, Eric R Bachmann, Robert B McGhee, and Michael J Zyda. An extended kalman filter for quaternion-based orientation estimation using marg sensors. In *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No. 01CH37180)*, volume 4, pages 2003–2011. IEEE, 2001.

[56] Anita M Mark. Options for dealing with tooth decay. *The Journal of the American Dental Association*, 149(10):927–928, 2018.

[57] Philip D Marsh. Microbiology of dental plaque biofilms and their role in oral health and caries. *Dental Clinics*, 54(3):441–454, 2010.

[58] Susan Michie, Lucy Yardley, Robert West, Kevin Patrick, and Felix Greaves. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. *Journal of medical Internet research*, 19(6):e232, 2017.

[59] Timothy J O'Leary. Oral hygiene agents and procedures. *Journal of Periodontology*, 41(11):625–629, 1970.

[60] Zhenchao Ouyang, Jingfeng Hu, Jianwei Niu, and Zhiping Qi. An asymmetrical acoustic field detection system for daily tooth brushing monitoring. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*, pages 1–6. IEEE, 2017.

[61] Talat Ozyagcilar. Calibrating an ecompass in the presence of hard and soft-iron interference. *Freescale Semiconductor Ltd*, pages 1–17, 2012.

106

[62] Abhinav Parate, Meng-Chieh Chiu, Chaniel Chadowitz, Deepak Ganesan, and Evangelos Kalogerakis. Risq: Recognizing smoking gestures with inertial sensors on a wristband. In *Proceedings of the 12th annual international conference on Mobile systems, applications, and services*, pages 149–161, 2014.

[63] Mark Pedley. Tilt sensing using a three-axis accelerometer. *Freescale semiconductor application note*, 1:2012–2013, 2013.

[64] R Core Team. *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, 2021. R version 4.1.1.

[65] Yossi Rubner, Carlo Tomasi, and Leonidas J Guibas. The earth mover's distance as a metric for image retrieval. *International journal of computer vision*, 40(2):99–121, 2000.

[66] AJ Rugg-Gunn and IDM MacGregor. A survey of toothbrushing behaviour in children and young adults. *Journal of Periodontal Research*, 13(4):382–389, 1978.

[67] Nazir Saleheen, Amin Ahsan Ali, Syed Monowar Hossain, Hillol Sarker, Soujanya Chatterjee, Benjamin Marlin, Emre Ertin, Mustafa Al'Absi, and Santosh Kumar. puffmarker: a multi-sensor approach for pinpointing the timing of first lapse in smoking cessation. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 999–1010, 2015.

[68] Vivek Shetty, Douglas Morrison, Thomas Belin, Timothy Hnat, and Santosh Kumar. A scalable system for passively monitoring oral health behaviors using electronic toothbrushes in the home setting: Development and feasibility study. *JMIR mHealth and uHealth*, 8(6):e17347, 2020.

[69] Hwanjun Song, Minseok Kim, Dongmin Park, Yooju Shin, and Jae-Gil Lee. Learning from noisy labels with deep neural networks: A survey. *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[70] Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei. Hierarchical dirichlet processes. *Journal of the american statistical association*, 101(476):1566–1581, 2006.

[71] Inc. The MathWorks. Matlab. https://www.mathworks.com/products/matlab.html, 2021.

[72] Edison Thomaz, Irfan Essa, and Gregory D Abowd. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1029–1040, 2015.

[73] Robert L Thorndike. Who belongs in the family? *Psychometrika*, 18(4):267–276, 1953.

[74] Anna L Trella, Kelly W Zhang, Inbal Nahum-Shani, Vivek Shetty, Finale Doshi-Velez, and Susan A Murphy. Designing reinforcement learning algorithms for digital interventions: pre-implementation guidelines. *Algorithms*, 15(8):255, 2022.

[75] Anna L Trella, Kelly W Zhang, Inbal Nahum-Shani, Vivek Shetty, Finale Doshi-Velez, and Susan A Murphy. Reward design for an online reinforcement learning algorithm supporting oral self-care. *arXiv preprint arXiv:2208.07406*, 2022.

[76] Grigorios Tsoumakas and Ioannis Katakis. Multi-label classification: An overview. *International Journal of Data Warehousing and Mining (IJDWM)*, 3(3):1–13, 2007.

[77] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *arXiv preprint arXiv:1706.03762*, 2017.

[78] Yequan Wang, Minlie Huang, Xiaoyan Zhu, and Li Zhao. Attention-based lstm for aspect-level sentiment classification. In *Proceedings of the 2016 conference on empirical methods in natural language processing*, pages 606–615, 2016.

[79] Karl Weiss, Taghi M Khoshgoftaar, and DingDing Wang. A survey of transfer learning. *Journal of Big data*, 3(1):1–40, 2016.

[80] Lahiru NS Wijayasingha and Benny Lo. A wearable sensing framework for improving personal and oral hygiene for people with developmental disabilities. In *2016 IEEE Wireless Health (WH)*, pages 1–7. IEEE, 2016.

[81] Tobias Winterfeld, N Schlueter, Daniela Harnacke, Jörg Illig, Jutta Margraf-Stiksrud, Renate Deinzer, and Carolina Ganss. Toothbrushing and flossing behaviour in young adults—a video observation. *Clinical oral investigations*, 19:851–858, 2015.

[82] Munirah Yaacob, Helen V Worthington, Scott A Deacon, Chris Deery, A Damien Walmsley, Peter G Robinson, and Anne-Marie Glenny. Powered versus manual toothbrushing for oral health. *Cochrane Database of Systematic Reviews*, (6), 2014.