

# Analysis and Enhancement of Interpersonal Coordination Using Inertial Measurement Unit Solutions

Von der Fakultät für Elektrotechnik und Informatik  
der Gottfried Wilhelm Leibniz Universität Hannover zur Erlangung des Grades

Doktor der INGENIEURWISSENSCHAFTEN

(abgekürzt: Dr.-Ing.)

genehmigte Dissertation

von

M.S. Tonghun Hwang

geboren am 19.10.1982

in Anyang, Republik Korea

2020

Referentin/ Referent: Prof. Dr.-Ing. Holger Blume  
(Institut für Mikroelektronische Systeme)

Korreferentin/ Korreferent: Prof. Dr.-Ing. Jörn Ostermann  
(Institut für Informationsverarbeitung)

Korreferentin/ Korreferent: Prof. Dr. Alfred Effenberg  
(Institut für Sportwissenschaft)

Tag der Promotion: 29.01.2020

## **Abstract**

Today's mobile communication technologies have increased verbal and text-based communication with other humans, social robots and intelligent virtual assistants. On the other hand, the technologies reduce face-to-face communication. This social issue is critical because decreasing direct interactions may cause difficulty in reading social and environmental cues, thereby impeding the development of overall social skills. Recently, scientists have studied the importance of nonverbal interpersonal activities to social skills, by measuring human behavioral and neurophysiological patterns. These interdisciplinary approaches are in line with the European Union research project, "Socializing sensorimotor contingencies" (socSMCs), which aims to improve the capability of social robots and properly deal with autism spectrum disorder (ASD). Therefore, modelling and benchmarking healthy humans' social behavior are fundamental to establish a foundation for research on emergence and enhancement of interpersonal coordination. In this research project, two different experimental settings were categorized depending on interactants' distance: distal and proximal settings, where the structure of engaged cognitive systems changes, and the level of socSMCs differs. As a part of the project, this dissertation work referred to this spatial framework. Additionally, single-sensor solutions were developed to reduce costs and efforts in measuring human behaviors, recognizing the social behaviors, and enhancing interpersonal coordination. First of all, algorithms using a head worn inertial measurement unit (H-IMU) were developed to measure human kinematics, as a baseline for social behaviors. The results confirmed that the H-IMU can measure individual gait parameters by analyzing only head kinematics. Secondly, as a distal sensorimotor contingency, interpersonal relationship was considered with respect to a dynamic structure of three interacting components: positivity, mutual attentiveness, and coordination. The H-IMUs monitored the social behavioral events relying on kinematics of the head orientation and oscillation during walk and talk, which can contribute to estimate the level of rapport. Finally, in a new collaborative task with the proposed IMU-based tablet application, results verified effects of different auditory-motor feedbacks on the enhancement of interpersonal coordination in a proximal setting.

This dissertation has an intensive interdisciplinary character: Technological development, in the areas of sensor and software engineering, was required to apply to or solve issues in direct relation to predefined behavioral scientific questions in two different settings (distal and proximal). The given frame served as a reference in the development of the methods and settings in this dissertation. The proposed IMU-based solutions are also promising for various future applications due to widespread wearable devices with IMUs.

**Keywords:** *Auditory-motor feedback, Kinematic analysis, Real-time monitoring, Joint action, Socializing sensorimotor contingency*

## **Abstract**

Die heutigen mobilen Kommunikationstechnologien haben den Umfang der verbalen und textbasierten Kommunikation mit anderen Menschen, sozialen Robotern und künstlicher Intelligenz erhöht. Auf der anderen Seite reduzieren diese Technologien die nonverbale und die direkte persönliche Kommunikation, was zu einer gesellschaftlichen Thematik geworden ist, weil die Verringerung der direkten persönlichen Interaktionen eine angemessene Wahrnehmung sozialer und umgebungsbedingter Reizmuster erschweren und die Entwicklung allgemeiner sozialer Fähigkeiten bremsen könnte. Wissenschaftler haben aktuell die Bedeutung nonverbaler zwischenmenschlicher Aktivitäten als soziale Fähigkeiten untersucht, indem sie menschliche Verhaltensmuster in Zusammenhang mit den jeweiligen neurophysiologischen Aktivierungsmustern analysiert haben. Solche Querschnittsansätze werden auch im Forschungsprojekt der Europäischen Union "Socializing sensori-motor contingencies" (socSMCs) verfolgt, das darauf abzielt, die Leistungsfähigkeit sozialer Roboter zu verbessern und Autismus-Spektrumsstörungen (ASD) adäquat zu behandeln. In diesem Zusammenhang ist die Modellierung und das Benchmarking des Sozialverhaltens gesunder Menschen eine Grundlage für theorieorientierte und experimentelle Studien zum weiterführenden Verständnis und zur Unterstützung interpersoneller Koordination. In diesem Zusammenhang wurden zwei verschiedene empirische Kategorien in Abhängigkeit von der Entfernung der Interagierenden zueinander vorgeschlagen: distale vs. proximale Interaktionssettings, da sich die Struktur der beteiligten kognitiven Systeme zwischen den Kategorien ändert und sich die Ebene der erwachsenden socSMCs verschiebt. Da diese Dissertation im Rahmen des socSMCs-Projekts entstanden ist, wurden Interaktionssettings für beide Kategorien (distal und proximal) entwickelt. Zudem wurden Ein-Sensor-Lösungen zur Reduzierung des Messaufwands (und auch der Kosten) entwickelt, um eine Messung ausgesuchter Verhaltensparameter bei einer Vielzahl von Menschen und sozialen Interaktionen zu ermöglichen. Zunächst wurden Algorithmen für eine kopfgetragene Trägheitsmesseinheit (H-IMU) zur Messung der menschlichen Kinematik als eine Ein-Sensor-Lösung entwickelt. Die Ergebnisse bestätigten, dass die H-IMU die eigenen Gangparameter unabhängig voneinander allein auf Basis der Kopfkinematik messen kann. Zweitens wurden—als ein

distales socSMC-Setting—die interpersonellen Kopplungen mit einem Bezug auf drei interagierende Merkmale von „Übereinstimmung“ (engl.: rapport) behandelt: Positivität, gegenseitige Aufmerksamkeit und Koordination. Die H-IMUs überwachten bestimmte soziale Verhaltensereignisse, die sich auf die Kinematik der Kopforientierung und Oszillation während des Gehens und Sprechens stützen, so dass der Grad der Übereinstimmung geschätzt werden konnte. Schließlich belegten die Ergebnisse einer experimentellen Studie, die zu einer kollaborativen Aufgabe mit der entwickelten IMU-basierten Tablet-Anwendung durchgeführt wurde, unterschiedliche Wirkungen verschiedener audio-motorischer Feedbackformen für eine Unterstützung der interpersonellen Koordination in der Kategorie proximaler sensomotorischer Kontingenzen.

Diese Dissertation hat einen intensiven interdisziplinären Charakter: Technologische Anforderungen in den Bereichen der Sensortechnologie und der Softwareentwicklung mussten in direktem Bezug auf vordefinierte verhaltenswissenschaftliche Fragestellungen entwickelt und angewendet bzw. gelöst werden—und dies in zwei unterschiedlichen Domänen (distal, proximal). Der gegebene Bezugsrahmen wurde als eine große Herausforderung bei der Entwicklung der beschriebenen Methoden und Settings wahrgenommen. Die vorgeschlagenen IMU-basierten Lösungen könnten dank der weit verbreiteten IMU-basierten mobilen Geräte zukünftig in verschiedene Anwendungen perspektiv reich integriert werden.

**Schlagworte:** *Audio-motorisches Feedback, Kinematische Analyse, Echtzeitüberwachung, Gemeinsame Aktion, Sozialisierung sensomotorischer Kontingenz*

## **Acknowledgments**

The author acknowledges support from European Commission HORIZON2020-FETPROACT-2014 No. 641321.

## Table of Contents

List of Tables.....	x
List of Figures .....	xi
List of Abbreviations.....	xvi
List of Terminology.....	xviii
Chapter 1. Introduction.....	1
1.1. Interpersonal Coordination.....	4
1.1.1. Analysis of Interpersonal Coordination .....	5
1.1.2. Assessment Methods of Interpersonal Coordination.....	6
1.1.3. Head and Body Kinematics for Interpersonal Coordination.....	8
1.1.4. Coordination between Head and Body Kinematics.....	8
1.1.5. Coordination between Head Pose and Emotions .....	9
1.2. Inertial Measurement Unit.....	10
1.2.1. Motivation of Using IMU-Based Motion Sensor.....	10
1.2.2. MEMS Sensors: Accelerometer, Gyroscope, and Magnetometer.....	11
1.2.3. Sensor Fusion Technology and Error Compensation.....	16
1.3. Background of Goals.....	20
1.3.1. Interventions for Autism Spectrum Disorder (ASD) .....	20
1.3.2. Interaction with Robots .....	21



Chapter 2. Measurement of Gait Parameters.....	24
2.1. Introduction .....	24
2.2. Related Work.....	25
2.2.1. Measurement of Human Movement.....	25
2.2.2. Motion Capture System.....	26
2.2.3. Reduced Number of Sensors and Sensor Placement.....	28
2.2.4. Previous Work Using a Head-Worn Sensor.....	28
2.3. Gait Events in Gait Cycle .....	29
2.4. Methodology of Gait Event Detection Using H-IMU .....	31
2.4.1. System Overview .....	31
2.4.2. Peak Detection.....	34
2.4.3. Digital Filters.....	38
2.4.4. Thresholding.....	42
2.5. Gait Parameters with H-IMU .....	43
2.5.1. Temporal Gait Parameters .....	43
2.5.2. Spatial Gait Parameters .....	47
2.6. Results .....	52
2.6.1. Step Counting .....	53
2.6.2. Spatiotemporal Parameters .....	55
2.6.3. System Validity .....	56
2.7. Discussion .....	60
2.7.1. Advantages of Using Head Kinematics in Real Time Gait Event Detection.....	61
2.7.2. Limitation of H-IMU in Gait Analysis.....	63

2.7.3. Technical Issues of H-IMU in Gait Analysis .....	64
2.8. Conclusion.....	65
Chapter 3. Rapport Monitoring System During Walk and Talk .....	66
3.1. Introduction .....	66
3.2. Related Work.....	67
3.2.1. Molecular and Molar Measurement of the Three Rapport Components .....	67
3.2.2. Head Pose Estimation.....	68
3.2.3. Gait Event Detection .....	69
3.3. Methods .....	71
3.3.1. Head Nod Detection .....	71
3.3.2. Estimation of Mutual Head Orientation and Its Coordination .....	71
3.3.3. Gait Coordination .....	76
3.4. Results .....	77
3.4.1. Head Nod Detection .....	77
3.4.2. Mutual Head Orientation and Coordination of Head Orientation.....	78
3.4.3. Correlation of Gait Patterns.....	80
3.5. Discussion .....	82
3.6. Conclusion.....	83
Chapter 4. Joint Task using IMU-based Application .....	84
4.1. Introduction .....	84
4.2. Materials and Methods .....	90
4.2.1. Participants .....	90
4.2.2. System Specifications.....	91

4.2.3. Design and Stimuli .....	91
4.2.4. Procedure .....	95
4.2.5. Data Analysis .....	97
4.3. Results .....	98
4.4. Discussion .....	105
4.5. Conclusion.....	109
Chapter 5. Discussion and Conclusion.....	110
5.1. Advantages and Comparisons .....	111
5.1.1. H-IMU: Advantages in Measuring Large Groups.....	111
5.1.2. Tetherball Study: Advantage of IMU-Based Tablet Task.....	112
5.2. Tetherball Study: Comparison with Other Interpersonal Coordination Studies ...	113
5.3. Technical Review .....	114
5.3.1. H-IMU: System Optimizations for Spread of the Solution.....	114
5.3.2. H-IMU: Compatibility for more Analytical Parameters .....	115
5.3.3. Tetherball Study: Sound Systems.....	116
5.4. Future Applications .....	116
5.4.1. Walking Intervention.....	117
5.4.2. Walk and Talk Intervention.....	117
5.4.3. Social Intervention with Physical Collaboration Results .....	119
5.5. Concluding Remarks .....	120
<b>List of Publications</b> .....	122
<b>Bibliography</b> .....	124

## List of Tables

Table 2.1. Comparison of Single IMU Methods for Gait Analysis .....	33
Table 2.2. Comparison of Three Methods in Step Counting.....	52
Table 2.3. Spatiotemporal Parameters Computed by H-IMU .....	54
Table 2.4. Spatial Parameters in Different Methods and Conditions .....	56
Table 2.5. MAPE of Spatial Parameters of PSL and ESL in Different Conditions .....	58
Table 2.6. Comparison of Single IMU Methods for Step Counting Accuracy .....	59
Table 3.1. The Nod Recognition Using H-IMU.....	78

## List of Figures

- Figure 1.1. A capacitive accelerometer with (a) a modeling diagram of the basic concept (damping factor is ignored) and (b) an example readout circuit with a charge amplifier, and (c) an example layout with comb structure of variable capacitors. .... 13
- Figure 1.2. A vibrating gyroscope with (a) a modeling diagram (damping factor is ignored) and (b) an example layout with the dual mass tuning fork design..... 15
- Figure 1.3. A magnetometer with (a) a modeling diagram (damping factor is ignored) and (b) an example layout with comb structure of variable capacitors. .... 17
- Figure 1.4. Local coordinate systems of sensors to express (a) gravity factor ( $g$ ) of measured acceleration in the initial (as the same as the global coordinate system) and the rotated position, and (b) the magnetic fields measured in initial ( $m_0$ ) and the rotated position ( $m_x, m_y, m_z$ )..... 19
- Figure 2.1. Time diagram of a gait cycle with example pictures at walking gait events which are from the right side view on the sagittal plane. Abbreviations of ‘HS’ is heel strike, ‘TO’ is toe off, ‘R’ is right, ‘L’ is left, (i) is initial phase, (m) is mid-phase, and (t) is terminal phase. .... 30
- Figure 2.2. Comparison of two peak detection algorithms: (a) the windowed peak detection (WPD) which analyzes every 1 step time (optimally 700 ms), and (b) the proposed peak detection which analyzes every frame (16.7 ms at 60 Hz)..... 32
- Figure 2.3. Vertical acceleration of body segments of the left foot, right foot, pelvis and head which are measured by an XSENS MVN system at 60 Hz. .... 36

Figure 2.4. A block diagram of the general digital filter, where  $x[n]$  and  $y[n]$  are the input and output signal, respectively. As integer numbers,  $N$  and  $M$  are the filter orders. The values,  $a_i$  and  $b_j$ , are the impulse response for  $i$ -th and  $j$ -th components. ....38

Figure 2.5. Block diagram of (a) the overall system with an IMU sensor module, and (b) the real time low pass filter at one frame ( $x[n_{fr}][n_s]$ ;  $n_s = 0-15$ ), which implements FFT and IFFT with 16 samples.....40

Figure 2.6. Head vertical acceleration,  $a[n_{fr}]$ , is shown with (a) 46 frames ( $n_{fr} = 35-80$ ) near an HS and depicted with the filtering area at Frame 56 ( $n_{fr} = 41-56$ ; frame rate: 60 Hz). With the 16 samples, (b) the input of the real time low pass filter (RLF),  $x[56][n_s]$ , and (c) the output of RLF,  $y[56][n_s]$ , are depicted where  $n_s = 0-15$ . ....41

Figure 2.7. Waveforms of the head acceleration after real time low pass filter (RLF) between frames  $n_{fr} = 51-56$ , demonstrating that the peak is over the threshold value ( $2.0 \text{ m/s}^2$ ) at frame  $n_{fr} = 54$ . ....42

Figure 2.8. The sequence of propulsion motions during walking, which causes head vertical acceleration peaks (a) at foot flat (Peak 2), (b) toe off (Peak 3), and (c) the beginning of the stance phase (Peak 4). ....44

Figure 2.9. Acceleration of the right foot and the head (frames  $n_{fr} = 50-69$ ) and peak count data with before and after real-time low-pass filter (RLF). After HS detection, Peak 3 is labeled as TO. The velocity of the head is also depicted as another evidence for TO detection, which can prevent skipping TO detection. ....45

Figure 2.10. Sequences of a walking avatar is depicted (a) on the sagittal plain along with the walking vector ( $V_d$ ) including labels of SL, PSL and ESL, (b) on the transverse plane with SL and PSL, and (c) on the transverse plane with SL and ESL. ....49

Figure 2.11. Constant errors between H-IMU and foot worn IMU data (reference) with Block and Whisker diagrams in terms of (a) timestamp of gait events HS and TO as well as (b) CT (Hwang et al., 2019b).....58

Figure 3.1. Geometrical definitions: (a) the yaw (shake), pitch (nod) and roll of the head; (b) the walking vector, the head orientation, the rotation matrix, as well as the global and local coordinate system.....70

Figure 3.2. RMS value (window width: 1 second) of head pitch acceleration, and the threshold values for obvious (T1) and subtle (T2) nod detection. ....70

Figure 3.3. Example diagram of the expressions of the rotated local axes ( $x'$ ,  $y'$ ,  $z'$ ) by  $\theta_z$  in the global coordinate system, and the new expressions of the axes and the walking vector,  $W'$ , in the local coordinate system .....73

Figure 3.4. Examples of walking vector, head direction and angles. In global Cartesian coordination, the origin and direction of participants are demonstrated. ....74

Figure 3.5. The four different scenarios of walking with conversation partners, and angle differences of two walkers. ....75

Figure 3.6. Individual head orientation status with three statuses; looking left (-1.0), forward (0.0), and right (1.0). P1 and P2 are participants walk on the left and right, respectively.....75

Figure 3.7. The binary signals of two participants during walk. The status value is 1 between HS and TO when two legs are in contact to the ground, called the double support period. After TO and before HS, one leg supports the body, so that the status value is 0.....77

Figure 3.8. Evaluation results of mutual head orientation and coordination of head orientation at the beginning 1,540 samples at a 60 Hz. At (a), high values stand for turn their head more toward their partner. At (b), there are three status values; 1: interactants look in the same direction; 0.5: look forward; and 0.0: look in the different directions.....79

Figure 3.9. Pearson correlation results between two participants’ binary signals of the dual support period in two conditions: (a) during the first lap, when they walked alone and (b) during the second lap, when they walked together. ....81

Figure 4.1. Top view of the tablet screen and levers fixed to the casing. ....88

Figure 4.2. Side view of the mechanical structure which supports the tablet and restricts to 2-DOF movement of the tablet. ....92

Figure 4.3. Illustrations of (a) the top view of the mechanical structure and the seats for left-handed (LH) and right-handed (RH) participants (P1 and P2), as well as (b) the exemplary situation of the experiment with two right-handed participants and a tester.....93

Figure 4.4. The feedback loop of perceptual information during the joint task.....94

Figure 4.5. Top view of a visuo-motor test application to measure the individual visuo-motor ability which is an important skill for the performance at the tetherball-application. ....96

Figure 4.6. Across subject means and standard deviations of the absolute error over the trials in four groups, (a) the VFG, (b) the EAFG, (c) the PAFG, and (d) the CAFG. Between-subject means and standard deviations are illustrated. The first 8.3 s (500 samples) of every 1-minute trial was eliminated.....99



Figure 4.7. Across subject means and standard deviations about the cross correlation coefficient of a pair of participants over the trials in four groups, (a) the VFG, (b) the EAFG, (c) the PAFG, and (d) the CAFG. To take maximal performance of pairs, the maximal coefficient in each trial was chosen with lags equivalent about 90° phase delay. The first 8.3 s (500 samples) of every 1-minute trial were eliminated. .... 101

Figure 4.8. Box and whisker plots of subjective ratings when participants were asked about how much they felt (a) that their own action helped their partners, (b) that their partners' action helped their own action, (c) how pleasant their experience was during the collaboration, and (d) how effectively they felt that they handled the apparatus together at the initial time and at the end. .... 104

## List of Abbreviations

Abbreviation	Explanation
2-D	2-Dimensional
3-D	3-Dimensional
95-CI	95% Confidence Interval
AI	Artificial Intelligence
ANOVA	ANalysis Of VAriance
APA	American Psychiatric Association
ARM	Advanced RISC (Reduced Instruction Set Computer) Machine
ASD	Autism Spectrum Disorder
BCE	Before Common Era
CAF	Combined Auditory Feedback with EAF and PAF
CAFG	CAF Group
CNS	Central Nervous System
CT	foot-ground Contact Time
CTR	Contact Time Ratio
CW	Clock Wise
DOF	Degree Of Freedom
DSM	Diagnostic and Statistical Manual of mental disorders
dps	<u>d</u> egree per <u>s</u> econd ( $^{\circ}/s$ )
DTFT	Discrete Time Fourier Transform
EAF	Effect-based Auditory Feedback
EAFG	EAF Group
EEG	Electro-Encephalo-Graphy (or Electroencephalography)
EMG	Electro-Myo-Graphy (or Electromyography)
ESL	Estimated Step Length
EU	European Union
FFT	Fast Fourier Transform
FIR	Finite Impulse Response
GPS	Global Positioning System
HHI	Human-Human Interaction
H-IMU	Head-worn Inertial Measurement Unit
HMD	Head Mounted Display
HMM	Hidden Markov Model
HRI	Human-Robot Interaction

Abbreviation	Explanation
HS	Heel Strike
IAAF	International Association of Athletics Federations
IDTFT	Inverse Discrete Time Fourier Transform
IFFT	Inverse Fast Fourier Transform
IIR	Infinite Impulse Response
IMU	Inertial Measurement Unit
IoT	Internet of Things
IR	Infrared Radiation
KP	Knowledge of Performance
KR	Knowledge of Result
LGPL	Lesser General Public License
MEMS	Micro-Electro-Mechanical Systems (or Microelectromechanical Systems)
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
PAF	Performance-based Auditory Feedback
PAFG	PAF Group
px	pixels
PSL	Pseudo Step Length
RLF	Real-time Low-pass Filter
RMS	Root Mean Square
Sci-Fi	Science Fiction
SL	Step Length
SCER	Step Count Error Rate
SNR	Signal to Noise Ratio
socSMCs	socializing SensoriMotor Contengencies
TO	Toe Off
ToM	Theory of Mind
UWB	Ultra-WideBand
VF	Visual Feedback
VFG	VF Group
VSA	Vehicle Stability Asist
WHO	World Health Organization
Wi-Fi	Wireless Fidelity
WLAN	Wireless Local Area Network
WPD	Windowed Peak Detection
XGA	eXtended Graphics Array: 1,024 × 768 pixels

## List of Terminology

Content	Explanation
Theory of Mind	An ability to attribute mental states (e.g., beliefs, desires, perspectives, intentions) to oneself and others, and to understand that others' mental states are different from one's own (Premack and Woodruff, 1978). This is important when individuals socialize and communicate with others.
Internet of Things	A concept of interconnection between everyday physical objects through the internet. In this idea, computing systems or devices embedded in the objects can transfer data and identify one another. This concept includes interdisciplinary technology (artificial intelligence, real time sensors, embedded systems, and communication) and will contribute to future technology (autonomous cars, smart home, and smart city).
Frames per second (fps)	A unit of the frame rate which is the frequency of refreshing images (or frames) in display or camera systems, also expressed in Hertz (Hz). In this dissertation, it is used as frame rate of a motion capture system.
Sample	Data point of discrete-time signal in data analysis. It is normally used for input and output of digital filters or processed sensor data. In chapter 4, statistical meaning of the sample is also used to refer to participants when using these terms: independent, dependent, across, and between samples.
Entrainment	Synchronization of organisms to perceived external cues or rhythms such as music or dance (Clayton, Sager, and Will, 2005), which comes from entrainment in biomusicology. In this dissertation, it is used for psychological and physical synchronization between two or more people.
Coordination	Combination of movement which comes from motor coordination. It is used for physical synchronization between interactants, in this dissertation.
Complexity	Degree of complexity of the algorithm in computer science, also called computational complexity. In this dissertation, the complexity indicates the worst-case complexity in time, which is expressed by a function of input size $n$ (unit: bit). This term is related to the speed of algorithm in discrete signal processing.
Rapport	A close and harmonious relationship between two or more people. Based on friendliness and warmth, rapport helps people to understand others' ideas and feelings and to communicate smoothly (Tickle-Degnen and Rosenthal, 1990).
Molecular method	An approach to analyze unit objects or behaviors, which is a start for analysis or science. In psychology, discrete units of activities in a complex chain of performances are analyzed in terms of a specific period of time or the number of the activities. This approach is suitable for the beginning stage of analysis or the short term analysis of behaviors (Baum, 2004).
Molar method	An approach to analyze objects or behaviors considering their environment. Whereas the same atoms have different properties in different structures or compositions in chemistry or ecology, the discrete behaviors have different meanings in different contexts in psychology. This approach is required in long term analysis of behaviors (Baum, 2004).

Content	Explanation
Dead reckoning	An estimation method using previously determined data. It is widely used for positioning systems that find the displacement by integrating the past velocity or acceleration information. The integration inherently causes error accumulation, which is dead reckoning effect. Today, many algorithms have been developed to compensate for the error accumulation.
Kinematic analysis	An analysis which focuses on the geometry of motion of the body and systems. The kinematic analysis provides kinematic parameters, such as acceleration, velocity, angular velocity, and angle. Kinematic analysis excludes force analysis, whereas kinetic analysis includes force analysis.
Kinesthetic system	A sensory system for awareness of the position and movement of the body part in physiology, also called kinesthetic sense, proprioceptive sense or proprioception. For example, individuals are aware of where the arm is located and its speed, even though the eyes are closed. Kinesthetic system can receive feedback information about moving the arms, legs, and head.
Likert scale	A rating scale widely used in questionnaires to measure participants' opinions or attitudes (Likert, 1932). In this dissertation, a 7-point Likert scale from 1 to 7 was used, which provides ordinal data.
Mauchly's sphericity test	A statistical procedure for scrutinizing the sphericity in a repeated-measures analysis of variances (ANOVA). The sphericity is met when the variances between the individual factor levels as well as the correlations between the factor levels are homogeneous (Bortz, 2005). The violation of sphericity results in too low p-values and thus inflates the error rate. The F-test of an ANOVA, therefore, might be false positive.
Huynh-Feldt correction	In case of a violation of sphericity, the Huynh-Feldt correction can be applied. It adjusts the degrees of freedom of the numerator and the denominator of the F-ratio by multiplying them with epsilon ( $\epsilon$ ). $\epsilon$ is estimated from the sample. It is $< 1$ if the sphericity is violated (Lane, 2016; Davis, 2002).
Levene's test	A test for assessing variance homogeneity across groups (Levene, 1960; Bortz, 2005). It tests the null hypothesis ( $H_0$ ) that the population variances are homogeneous. $H_0$ is rejected if $p < 0.05$ , indicating that variances differ between the samples.
Tukey's HSD test	The Tukey's honestly significant difference (HSD) test is performed as a post hoc test if an ANOVA confirms a significant overall effect. It compares all possible pairs of means and corrects for the increased probability of Type I errors, so that it corrects for an increased probability of significant results (e.g., p-value) by chance due to multiple testing.
Mann-Whitney-U-test	A nonparametric test for ordinal scaled variables which compares the medians of two independent samples (Bortz, 2005; Mann & Whitney, 1947). In this dissertation, this test was used for the analysis of questionnaire results.
Wilcoxon test	A nonparametric test for ordinal scaled variables which compares the medians of two dependent samples (Bortz, 2005; Wilcoxon, 1992). In this dissertation, this test was used for the analysis of questionnaire results.

## **Chapter 1. Introduction**

Socializing and integrating into the community are important because it supports the sense of community-belonging, which is highly correlated with one's happiness and health. Social and communication skills contribute to co-performance in group and team-based activities, which are required by today's collaborative societies. On the one hand, widespread mobile devices with up-to-date technology have recently increased indirect communication with others through the internet connection and social network services. On the other hand, face-to-face interactions have gradually decreased during last decades due to the mobile technology. Indirect communication with mobile devices causes users to have less experiences in reading nonverbal cues and environmental context, which reduces one's social skills (Munoz, 2013). Today, therefore, it is much more important to study how to enhance nonverbal interpersonal coordination and how to learn social skills efficiently and effectively.

For this topic, several attempts have been made to understand when and where interpersonal coordination emerges. Recent evidence suggests that interpersonal coordination is ubiquitous, which can emerge whenever the opportunities for interpersonal connections appear, such as a physical connection, visual or auditory cues, linguistic information, even without one's awareness (Shockley, and Riley, 2015). In football, a ball can be a coupling medium as a way of interaction. In orchestra, a conductor gives musicians visual cues that make performance synchronized (D'Ausilio, Badino, Li, Tokay, Craighero, Canto, Aloimonos, and Fadiga, 2012). Auditory and visual information shared in musical ensembles can be media for coordination. Linguistic information can be a medium on conversation between two individuals and arise common understanding of audience, which can be a means of exchanging multiple individuals' feelings, attitudes, and actions (Shockley, and Riley, 2015; Port, 1981).

Depending on the medium, recognition of social cues can give rise to one's reactions in a physical and neuropsychological manner, which can affect interpersonal coordination positively or negatively, acting as feedback. Researchers have studied short-

term and long-term effects of the feedback on human cognition and behaviors during interaction. They have used high-end technology to scan the human brain and measure the whole human body movements, which can be realized usually in laboratory settings at high costs. On the other hand, simple solutions are also considered to measure the specific biological events (e.g., the heart rate, eye reaction, body gestures), which can indicate a response to social stimuli during interaction. In this field, mobile and wearable technologies are emerging and wide-spreading as the simple behavioral measurement devices due to the up-to-date sensor technology, such as inertial measurement unit (IMU).

The IMU is a sensor device that measures kinematic parameters using a body's inertial force, and that is embedded in positioning systems, such as motion capture and navigation systems. Embedded in IMU, accelerometers and gyroscopes measure the acceleration and the rate of turn. Magnetometers, barometers, or global positioning system (GPS) receivers can be added in IMU to compensate for error accumulation in positioning. Nowadays, the applications of IMU have been broadened to consumer electronics, such as smartphones, smart watches, and wireless earbuds. One of the main interests of this dissertation is measuring one's independent meaningful movements with a single IMU. Despite the simple unit, the IMU can provide information about one's gait and head gestures, as well as joint actions during a joint task.

The analysis and enhancement of interpersonal coordination are investigated in this dissertation, as a part of the EU research project, "Socializing sensorimotor contingencies" (socSMCs). This research project aims at analyzing and stimulating sensorimotor functions in correlated activities, in order to initiate and enhance interpersonal coordination. In this project, different interaction conditions, distal and proximal settings, are considered in terms of the spatial factors between interactants because phenomena of action-effect contingencies and the level of SMCs are changed, which can be demonstrated by measuring brain activation and body dynamics. Another reason is that these conditions are regarded as a suitable experimental frame for artificial agents and autistic people. This dissertation is similarly categorized into three environments considering one individual setting, and two settings of distal and proximal interpersonal coordination.

Before describing main studies of this dissertation, the background knowledge of interpersonal coordination and IMU is described in the following subsections in Chapter 1.

In Chapter 2, to measure a basic human behavior, gait analysis is considered in a daily life setting. It is demonstrated how a single head-worn IMU (H-IMU) supports gait analysis, which is normally measured by a number of sensors or cameras. The single IMU solution, however, enables users to use the simple measurement system in daily life. The system validation was done based on a foot-worn IMU solution and manual assessments, which are valid methods in gait analysis. This chapter does not include the empirical study on interpersonal coordination; however, gait analysis using the simple H-IMU solution lays the groundwork for the study on group and dyadic interpersonal coordination.

In Chapter 3, distal interpersonal coordination (i.e., with the physical distance more than about 1.5 m) is taken into account in a walk and talk setting, as an everyday-life setting. The structured measurement of rapport between interactants during walking together is suggested with a combined analysis of the head direction, the number of head nods, and gait patterns. The rapport measurement, thus, provides information of positivity, mutual attentiveness, and coordination. This solution analyzes discrete behaviors without a consideration of social context. Nevertheless, the structured measurement is valuable as a reference for the rapport measurement, and the H-IMU reduces time-consuming manual works. This chapter indicates that, combined with social context analysis from human examiners, the solution can provide more efficient and effective rapport measurements.

In Chapter 4, finally, proximal interpersonal coordination (i.e., with the distance less than about 1.5 m) is described to investigate independent and joint actions of two interactants during a joint task. A tablet with an IMU measures the joint outcomes and synchronization with a joint task application. In this task, human behavior is not directly measured; however, the angular velocity and posture of the tablet are measured, which is manipulated by participants' finger movements. With this setting, it is demonstrated that certain kinds of additional auditory feedback can affect interpersonal coordination.

In Chapter 5, general findings, contributions, and applications are discussed, and also concluding remarks are given.



## **1.1. Interpersonal Coordination**

Interpersonal coordination entails physical and psychological interaction between two or more individuals. Interpersonal coordination is categorized depending on the number of agents: dyadic interpersonal coordination between two individuals, and group interpersonal coordination. Dyadic interpersonal coordination is researched for joint action (Clark, 1996), communication, and rapport, which can explain the formation of relationships between two individuals (Tickle-Degnen and Rosenthal, 1990; Fiske, 1992). Dyadic interaction was studied with a rocking chair task (Demos, Chaffin, Begosh, Daniels, and Marsh, 2012), a duet dance (Waterhouse, Watts, and Bläsing, 2014), a tetherball paradigm (Hwang, Schmitz, Klemmt, Brinkop, Ghai, and Stoica et al., 2018a), posture mirroring, joint attention (Tickle-Degnen and Rosenthal, 1990), as well as gaze and emotional coordination (Shockley, and Riley, 2015). On the other hand, group interpersonal coordination might be the base of group cohesiveness, groupness, group entitativity and social identification (Richardson, Garcia, Frank, Gergor, and Marsh, 2012). Those are important concepts to understand social phenomena such as social influence, social identity process, and intergroup conflict (Richardson, Garcia, Frank, Gergor, and Marsh, 2012). Examples of group interpersonal coordination are that a speaker or a movie arouses audience clapping (Néda, Ravasz, Brechet, Vicsek, and Barabási, 2000), or musicians make a harmony in quartets and orchestras (D'Ausilio et al., 2012). The coordination is also observed in team sporting games (e.g., football, basketball) (Duarte, Araújo, Correia, Davids, Marques, and Richardson, 2013), in military marching (McNeill, 1995), and in group dancing (Ellamil, Berson, Wong, Buckley, and Margulies, 2016).

Interpersonal coordination is also categorized by physical distance between interactants: distal and proximal interpersonal coordination. Distal interpersonal coordination is observed normally when interactants have a distance greater than around 1.5 m, in which interactants can communicate with verbal and nonverbal social cues, such as gestures, facial expressions, languages, and clapping sounds. Visual and auditory information are used in this interaction, which is observed during conversations in work places, shops, and schools, as well as during group activity such as team sports, and musical

co-performances. On the other hand, proximal interpersonal coordination emerges when the distance between interactants is normally less than around 1.5 m. Interactants might directly touch each other or manipulate an object together. In addition to visual and auditory information, tactile information can be used in this interaction, which is found during interaction between mother and child or during cooperation to carry a table or a sofa.

### **1.1.1. Analysis of Interpersonal Coordination**

The objective of interpersonal coordination study is not only to optimize joint outcomes in musical ensembles, team sports, and workplaces, but also to support individuals with social deficits, such as autism spectrum disorder (ASD), schizophrenia, and dementia. Many interventions have been developed for athletes, patients and ordinary people to enhance the ability to predict others' behaviors. These interventions include promoting physical activity, and also enhancing the Theory of Mind (ToM), which is an ability to understand that mental states (e.g., beliefs, desires, perspectives, intentions) are caused by oneself and others, and that different persons have different mental states (Premack and Woodruff, 1978). Deficits of ToM are regarded as mind-blindness, which is a cognitive disorder resulting from impaired or delayed neural networking (Gallagher and Frith, 2003; Carruthers, 1996; Bird, Castelli, Malik, Frith, and Husain, 2004). In terms of ASD, however, some phenomena cannot be explained with the theory of mind-blindness because of delayed motor skills, frequent repetitive behaviors, and high functioning memory skills (Frith, 2001; Baron-Cohen, 2004). Nevertheless, these affected individuals need therapies relying on behavioral interventions (Frith, 2001; Laghi, Lonigro, Levanto, Ferraro, Baumgartner, and Baiocco, 2016).

The analysis of one's behaviors is necessary to arrange successful interventions and to improve outcomes. Monitoring kinematic and linguistic behaviors during interventions plays an important role in feedback. Long-term and real-time feedback can help to enhance the participants' interpersonal skills, and to inform therapists of participants' progress in the intervention. In addition, observation of different groups of participants is needed considering occupation, religion, gender and nationality, as well as the affected individuals.

By comparing data on ordinary versus affected people, therapists can decide the future direction of interventions. Furthermore, recent developments in the field of human-robot interaction have also led to a renewed interest in the research on interpersonal coordination because the analysis of human interactions should be a prerequisite for human-like robots (Scassellati, 2002; Muthugala, Munasinghe, Lakshan, Madurangi, and Jayasekara, 2013). The studies on human-human interactions are also helpful in social interaction studies for therapeutic robots (Robinson, MacDonald, Kerse, and Broadbent, 2013) and empathy between humans and robots (Darling, Nandy, and Breazeal, 2015).

### **1.1.2. Assessment Methods of Interpersonal Coordination**

In social psychology, the level of interpersonal coordination has usually been measured by subjective methods. For example, according to meta-analysis of dyadic interactions, the rapport was assessed by participants' self-reports and outside observers (Tickle-Degnen and Rosenthal, 1990). The degree of groupness or group entitativity was measured by participants' subjective rating (Lickel, Hamilton, Wierzchowska, Lewis, Sherman, and Uhles, 2000). These methods, however, allows various factors to change results, which causes difficulty in a comparison of two or more different studies. Researchers tried to realize objective psychological measurement by categorizing meaningful parameters during interpersonal coordination, such as the number of nodding, smiling, or arm crossing (Tickle-Degnen and Rosenthal, 1990). Nevertheless, the manual measurement is time consuming. Researchers, therefore, have found objective methods and standard models using computational systems. The numbers related to head gestures, facial expressions, and body postures are counted. The period of synchronized motion, coupled posture, mutual gaze or common attention is also taken into account. In data analysis phase, a number of computational models between two data sets have been used for quantifying the degree of interpersonal coordination, such as cross-correlation, and cross recurrence analysis, phase difference, frequency difference of coherence as well as entropy measurement and mutual information (Richardson, Garcia, Frank, Gergor, and Marsh, 2012). These methods are appropriate for analyzing dyadic interpersonal coordination. For group interpersonal coordination, on the other hand, a few methods of analysis have been

reported due to a large set of data and high complexity (Ellamil, Berson, and Margulies, 2016). One simple method is to measure the decibel of audience clapping sound, which estimates how many people in the audience are clapping (Néda, Ravasz, Brechet, Vicsek, and Barabási, 2000). A cluster phase method was reported in a study on group synchronization of six participants sitting on rocking chairs, which computed the average of each phase difference from the cluster phase, ranging from 0 to 1. The numerical results were regarded as indicators of the degree of group interpersonal synchronization (Richardson, Garcia, Frank, Gergor, and Marsh, 2012).

Those factors are based on social behaviors, which can be measured with various systems, such as microphones, cameras, and inertial sensors. For instance, the auditory measurement using microphones was suggested when an audience is clapping (Néda et al., 2000). The complex patterns of coordination that arise during a string quartet can also be analyzed based on the audio signals (Chang, Lee, Choe, and Lee, 2017; Volpe, D'Ausilio, Badino, Camurri, and Fadiga, 2016). Acoustic measurement, however, can be used only when the movement results in auditory events. Consequently, different types of measurement are needed to directly assess various social behaviors. For example, vision technology can detect head gestures (e.g., nodding, shaking) and recognize facial expressions (e.g., smile, frown, angry), indicating current interactants' emotions.

For gross motor behaviors, motion capture systems are preferred. Optical motion capture systems based on cameras and markers are used in tasks for spatiotemporal synchronization, such as hand mirroring tasks (Llobera, Charbonnier, Chagué, Preissmann, Antonietti, Ansermet, and Magistretti, 2016). Motion capture using inertial sensors were implemented for the measurement of participants' nonverbal behaviors during verbal communication (Feese, Arnrich, Troster, Meyer, and Jonas, 2012). Motion capture systems are also suitable for monitoring human interactions in education, training and clinical intervention settings, which can measure behavioral expressions, such as gestures and postures. Compared to 2-dimensional (2-D) video analysis, a 3 dimensional (3-D) motion capture system can provide more movement information in terms of geometry aspects. Thus, researchers can visualize necessary parameters from different viewpoints.

### **1.1.3. Head and Body Kinematics for Interpersonal Coordination**

People shake hands, dance and gesture to interact with others. Audiences clapping (Néda et al., 2000), drummers marching (Butzin, Hochendoner, Ogle, Hill, and Mather, 2015), and instruments' quartets (Wing, Endo, Bradbury, and Vorberg, 2014) are also frequently discussed in studies on audio-visual coupling which is based on inter-personal entrainment. On the one hand, kinematic measurements of arm gesture (Junker, Amft, Lukowicz, and Tröster, 2008), gait (Sabatini, Martelloni, Scapellato, and Cavallo, 2005) and trunk movement (Pfau, Ferrari, Parsons, and Wilson, 2008) are promising for the research on multi-agent interaction. Measuring extremity movement, however, could face limitations because individuals derive different meanings from gestures depending on their cultural background and individual characteristics. The cost also increases in the body kinematic measurement with more individuals because the measurement systems require more resources in the number of professionals and facilities, as well as the measurement time and space. On the other hand, measurement of head movement has no comparable limitations because it is based much more on basic instincts rather than on cultural environment and individual habits. For instance, humans start to move their head by themselves early in their lives (Gordon and Browne, 2013) and the head responds to basic stimuli, such as light and sound (Land, 1999; Murray, Lillakas, Weber, Moore, and Irving, 2007; Thurlow, and Runge, 1967). The measurement of head movements, in addition, needs only a few sensor devices per person, enabling a kinematic measurement of group behavior and entrainment. Head movement, nonetheless, is still lacking information to describe whole body movement. Therefore, measuring both extremity and head movement can compensate each other and can provide more reliable data for the research on interpersonal entrainment.

### **1.1.4. Coordination between Head and Body Kinematics**

Research on intrapersonal coordination has demonstrated that the head movement predicts body movements in specific cases. Walking speed can be estimated from the head velocity, and gait parameters are analyzed by head vertical kinematics (Hwang, Reh,

Effenberg, and Blume, 2018b). The head kinematics can also indicate that one's attention shifts because head orientation correlates to gaze position, where one's attention lies (Land, 1999). Several studies found that gaze position especially correlates to horizontal head movement (yaw, see Figure 3.1) during a gaze shifting task in front of a screen (Fang, Nakashima, Matsumiya, Kuriki, and Shioiri, 2015; Nakashima, Fang, Hatori, Hiratani, Matsumiya, Kuriki, and Shioiri, 2015). During straight and curved walking, it was reported that the gaze direction led the head yaw rotation and walking directions (Imai, Moore, Raphan, and Cohen, 2001; Grasso, Prévost, Ivanenko, and Berthoz, 1998), which means the walking direction is predictable from head orientation, when a person walks towards an interaction partner.

In addition, eyes-head-hand coordination is researched using several visuo-motor tasks: making tea (Land, 1999), and a pointing task (Vercher, Magenes, Prablanc, and Gauthier, 1994). In these tasks, the hand movement follows the gaze direction. The authors showed their sequential movement; eyes are shifted to the target, the head rotated towards the target, and then the hand moves to the target point. This sequence was observed in the optimal conditions without any visual or physical restrictions. The results indicate that the aimed target of hand's movement can be predicted from the head movement, which is applicable to interacting scenarios such as shaking hands and high five. Similarly, with activities for eye-foot coordination, the head orientation might support the estimation of foot positions, even the shooting directions in football (a.k.a. soccer) (Wood, and Wilson, 2010; Savelsbergh, Van der Kamp, Williams, and Ward, 2005).

### **1.1.5. Coordination between Head Pose and Emotions**

Behavioral scientists and neuropsychologists attempted to recognize emotion by measuring human behaviors. When subjects walk, gesticulate, and talk (Troje, 2002; Pollick, Paterson, Bruderlin, and Sanford, 2001; Bente, Krämer, Petersen, and de Ruitter, 2001; Troje, Westhoff, and Lavrov, 2005), body movement and posture can be referred to assess the emotional states. In the last decades, research has been conducted on relationships between nonverbal behavior and emotion. In terms of head movement,

oscillating rotational movement of the head expresses agreement or disagreement. Interlocutors and the audience nod their heads to agree and shake to disagree. The head pose can indicate one's emotion. Happiness causes people's head up and makes the movement faster, whereas sadness keeps the head down and makes the movement slower (Troje, 2002; Pollick et al., 2001). For more information, scientists visualize inherent body movement and posture by utilizing 3-dimensional (3-D) animation (Bente et al., 2001), point-lights (Troje, 2002; Pollick, et al., 2001; Troje et al., 2005), and even sonification (Effenberg, 2005; Schmitz, Mohammadi, Hammer, Heldmann, Samii, Münte, and Effenberg, 2013). Therefore, the relationship between emotion and body movement have been evidenced by data from the various animation visualizations. Consequently, these methods can also prove the dependency between emotion and the head movement. Compared to other body parts, the measurement of head kinematics can provide various kinds of information about emotion because the head motion is strongly related to emotion itself and coordinates other body movement during psychological changes. For example, the balance of the head is closely related to the whole body balance and posture (Paloski, Wood, Feiveson, Black, Hwang, and Reschke, 2006; Clark and Iltis, 2008). The tilt of one's head results from the bent spine which determines the balance of the body. The longitudinal axis from the head to the pelvis also indicates body balance and posture. Therefore, the database of physical and psychological studies might lead to various interpretations and analyses on human behaviors during interpersonal coordination, which helps people with mind blindness to be sociable.

## **1.2. Inertial Measurement Unit**

### **1.2.1. Motivation of Using IMU-Based Motion Sensor**

An IMU-based motion sensor is suitable for measuring human kinematics with wearable applications. The IMU has been used for sensing motions of vehicles. For example, IMU sensors can provide trajectory information, supporting navigation system of aircraft, watercraft, and automobiles. Drivers can also be supported by vehicle stability assist (VSA) when they steer a vehicle on a curved road. The IMU is also used for safety

and security purposes, by detecting movements to trigger an airbag or robbery alarm system. Nowadays, the sensor size is reduced because of microelectromechanical systems (MEMS), and thereby the IMU sensors can be embedded in smartphones and wearable devices. This trend allows mobile applications to monitor the health status in everyday life. Furthermore, the motion sensors have wireless connections to other devices and systems, establishing the health monitoring systems with internet of things (IoT) technology. With these advantages of mobility, simplicity and connectivity, IMU motion sensors are preferable for the wearable solutions, compared to fixed camera solutions. As the growth of wearable sensor market leads to more technology and price competitions, consumers can have more high quality and low cost products. These wearable sensors also support easy-to-access motion and posture detectors in everyday life conditions, which is suitable for physical interventions and social therapies including indoor and outdoor settings.

### 1.2.2. MEMS Sensors: Accelerometer, Gyroscope, and Magnetometer

Microelectromechanical systems (MEMS) is a technology for micro-scale devices, including mechanical structures, such as springs. These devices are manufactured with semiconductor fabrication technology, which includes molding, plating, etching (wet and dry etching) and machining (electro discharging machining: EDM), and mainly focuses on shaping metal for MEMS. Nowadays, MEMS can be integrated with other electronic circuits, even motion processing units (MPU) (Seeger, Lim, and Nasiri, 2010), which leads to a smaller size of MEMS. During the fabrication process, the vacuum level is controlled, and the MEMS structures are hermetically sealed for a high quality sensor (Nasiri, 2009). The accelerometers of IMU measure the acceleration using inertial force. When the micro system accelerates or decelerates, the inertial force moves the proof mass connected by springs in MEMS. This movable proof mass can cause capacitance changes as shown in Figure 1.1. In a time invariant system, capacitive accelerometers use the relationship inertial force ( $F_i(t)$ ) and two variable capacitors ( $C_1(t)$  and  $C_2(t)$ ) as in equations below:

$$C_1(t) = C_2(t) = \varepsilon \frac{A(t)}{d(t)}, (t = 0). \quad (1.1)$$



$$\begin{aligned}
C_1(t) &= \varepsilon \frac{A(t)}{d(t) - x(t)}, (t > 0). \\
C_2(t) &= \varepsilon \frac{A(t)}{d(t) + x(t)}, (t > 0).
\end{aligned}
\tag{1.2}$$

The capacitors are defined with a dielectric permittivity ( $\varepsilon$ ), area ( $A(t)$ ), and distance between electrodes ( $d(t)$ ). When inertial force occurs at  $t > 0$ , the capacitance is changed due to the displacement ( $x(t)$ ) of the proof mass. The displacement results from the inertial force, which is expressed by kinematics of the proof mass ( $m$ : mass,  $\mathbf{a}(t)$ : acceleration) and the spring ( $k$ : spring constant) (see also Figure 1.1(a)).

$$\mathbf{F}_i(t) = m\mathbf{a}(t) = -kx(t).
\tag{1.3}$$

$$x(t) = -\frac{m}{k}\mathbf{a}(t).
\tag{1.4}$$

The ratio of two variable capacitances can be measured by various types of readout circuits (Aszkler, 2005; Liang, Xiaowei, Weiping, and Zhiping, 2011; Langfelder and Tocchio, 2014; Qu, Yu, Zhou, Peng, Peng, and He, 2016). For example, the variable capacitances can be placed as an input capacitance of the charge amplifier circuit as shown Figure 1.1(b). The amplitude of an oscillating input voltage of the circuit can change the amplitude of output voltage depending on the ratio between the variable input capacitance ( $C_1$  and  $C_2$ ) and the feedback components ( $R_f$  and  $C_f$ ) (Qu et al., 2016). For the high sensitivity, the comb structure (see Figure. 1.1(c)) is mostly used to increase  $A(t)$  because it increases  $C(t)$ ; as a result, it is easy to measure micro-scale  $x$ -displacement (Xie and Fedder, 2000). The variable capacitance can be differently designed. It can depend on varying  $A(t)$  instead of  $d(t)$ . There are also piezoelectric types of accelerometers, which use changing electric displacement field of a piezoelectric material. When this material is under strain or stress due to the movable proof mass, output voltage (Aszkler, 2005) or its resistance is changed.

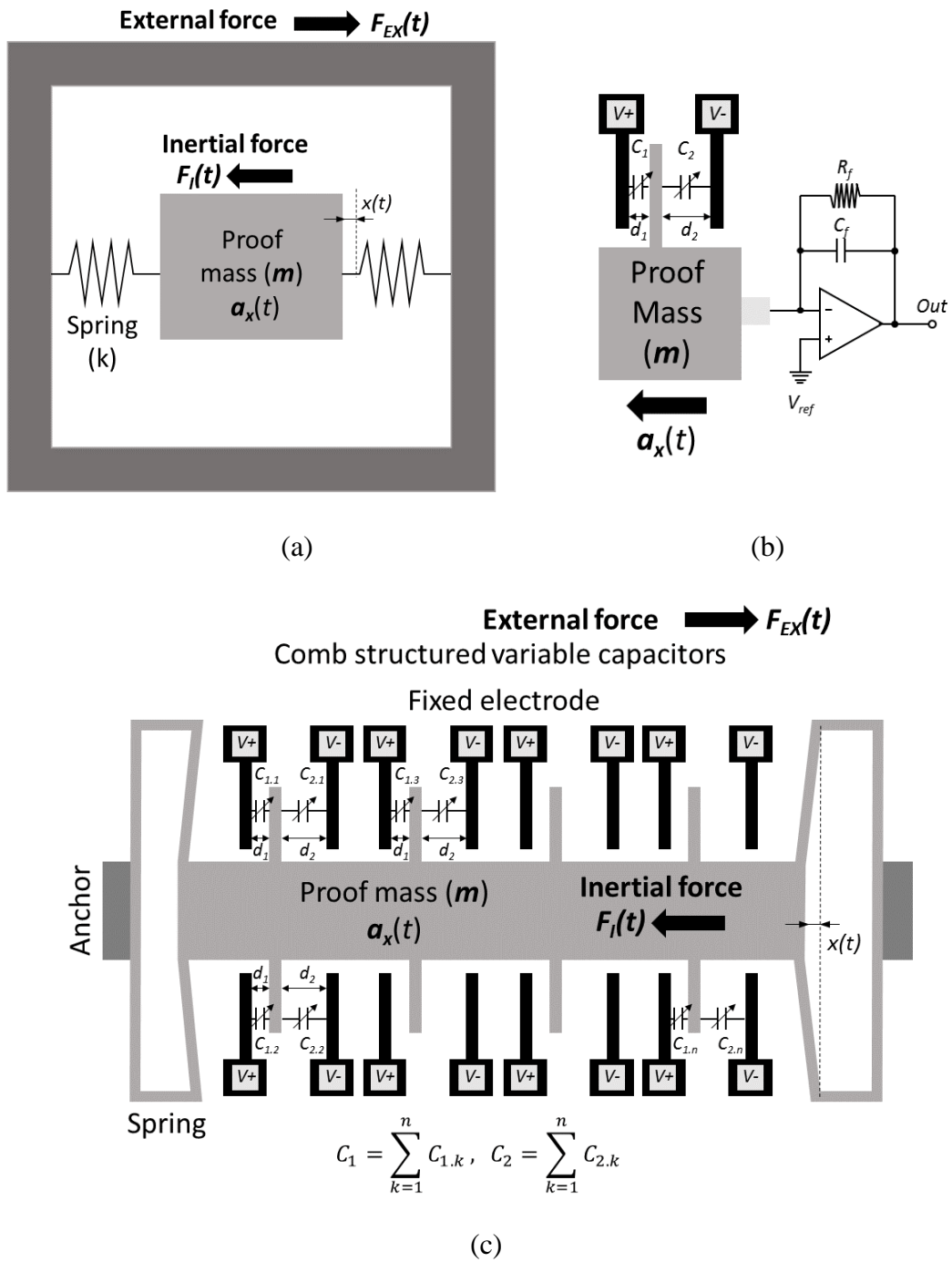


Figure 1.1. A capacitive accelerometer with (a) a modeling diagram of the basic concept (damping factor is ignored) and (b) an example readout circuit with a charge amplifier, and (c) an example layout with comb structure of variable capacitors.

A gyroscope measures the rate of turn based on the characteristics of Coriolis force. Currently, a vibrating structure gyroscope is broadly applied in smartphones, cameras and wearable devices because it is simpler and less expensive than the conventional rotating gyroscope structure—both structures have nearly the same level of accuracy. The vibrating structure involves the dual mass tuning fork design (Bernstein, Cho, King, Kourepenis, Maciel, and Weinberg, 1993), in which two proof masses are initially vibrating only in a vibrating plane along an axis (e.g.,  $x$ -axis) at frequency  $\omega_r$  (see Figure 1.2). When the frame rotates at a rate of turn ( $\Omega(t)$ ), Coriolis force ( $F_C(t)$ ) affects the velocity ( $\mathbf{v}(t)$ ) of the vibrating proof masses, generating a vector component outside of the vibrating plane (e.g.,  $y$ -axis component) as in equations below:

$$\mathbf{F}_C(t) = -2m\boldsymbol{\Omega}(t) \times \mathbf{v}(t) = -k_{op}y(t). \quad (1.5)$$

$$\begin{pmatrix} x(t) \\ v_x(t) \end{pmatrix} = \begin{pmatrix} X_{ip}\omega_r \sin(\omega_r t) \\ X_{ip}\omega_r \cos(\omega_r t) \end{pmatrix}. \quad (1.6)$$

$$y(t) = \frac{2m\Omega(t)X_{ip}\cos(\omega_r t)}{k_{op}}. \quad (1.7)$$

In (1.5),  $k_{op}$  is the spring constant for the out-plane (along the  $y$ -axis), and in (1.6),  $X_{ip}$  is the in-plane amplitude of the displacement  $x(t)$  (along the  $x$ -axis). According to (1.5), the direction of the Coriolis force changes, when the vibrating proof mass is moving inside ( $t_{in}$ ) and outside ( $t_{out}$ ) (see Figure 1.2). When the  $x(t)$  is given in (1.6), the  $x$ -velocity ( $v_x(t)$ ) is calculated by differentiation. Finally, the  $y$ -displacement  $y(t)$  is calculated as shown in (1.7). This  $y(t)$  causes capacitance changes, piezoelectric effects, or electromagnetic characteristics. These are detected by readout circuits, resulting in varying output voltage. Compared to the accelerometer, this gyroscope consumes more power to drive the vibration of proof masses. For consumer electronics, therefore, gyroscopes provide different modes for low power operation (e.g., ST microelectronics, 2013). Low-power gyroscopes also have been researched.

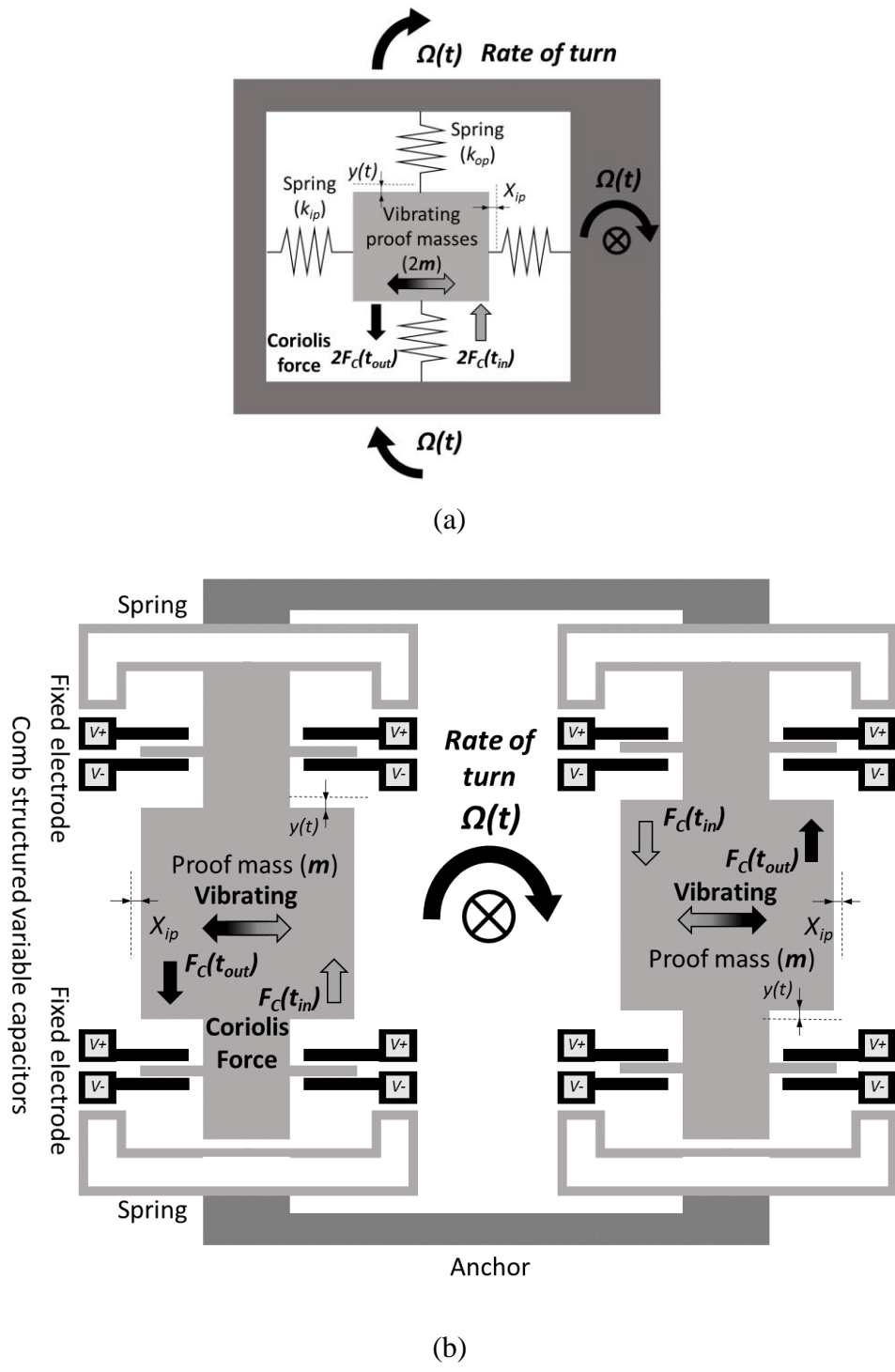


Figure 1.2. A vibrating gyroscope with (a) a modeling diagram (damping factor is ignored) and (b) an example layout with the dual mass tuning fork design.

In Figure 1.3, a magnetometer that uses Lorentz force ( $F_L(t)$ ) is depicted. When current ( $I(t)$ ) flows in magnetic field ( $B(t)$ ), Lorentz force occurs as in equation below:

$$F_L(t) = I(t) \times B(t) = -kx(t). \quad (1.8)$$

The  $F_L(t)$  pushes a movable proof mass, which changes capacitance of comb structure capacitors due to  $x$ -displacement  $x(t)$  in Figure 1.3 (Emmerich and Schofthaler, 2000). Finally, the varying capacitance is sensed by a readout circuit.

### 1.2.3. Sensor Fusion Technology and Error Compensation

Errors between measured and true parameters should be considered. For example, noise occurs when varying capacitance is measured by readout circuits (measurement noise: e.g., the offset of amplifier, process variations of capacitance and springs in MEMS), and when the analog output voltage is digitalized and filtered (processing noise: e.g., quantization error). In IMU, accelerometers have problems with gravity factor ( $g$ ) and accumulated errors when positions ( $x$ ) and velocities ( $v$ ) are calculated by integrating accelerations as below:

$$a_{motion} = a_{measured} - g. \quad (1.9)$$

$$v = \int_0^t a_{motion} + \int_0^t n_a. \quad (1.10)$$

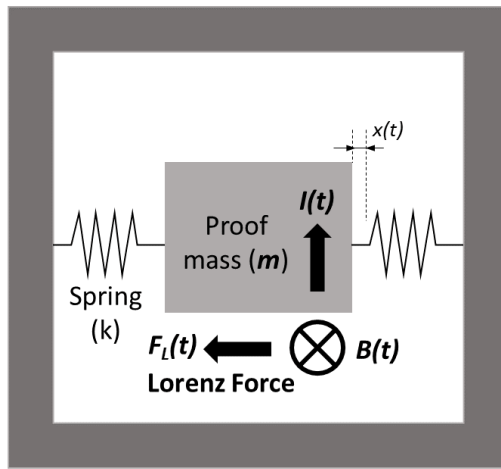
$$x = \iint_0^t a_{motion} + \iint_0^t n_a.$$

In the equations,  $a_{motion}$  and  $a_{measured}$  are the pure motion and measured acceleration, respectively, and  $n_a$  is the noise from the accelerometers. Gyroscopes also have the problems with accumulated error with the angular position ( $\theta$ ) and the earth speed ( $w_{earth}$ ):

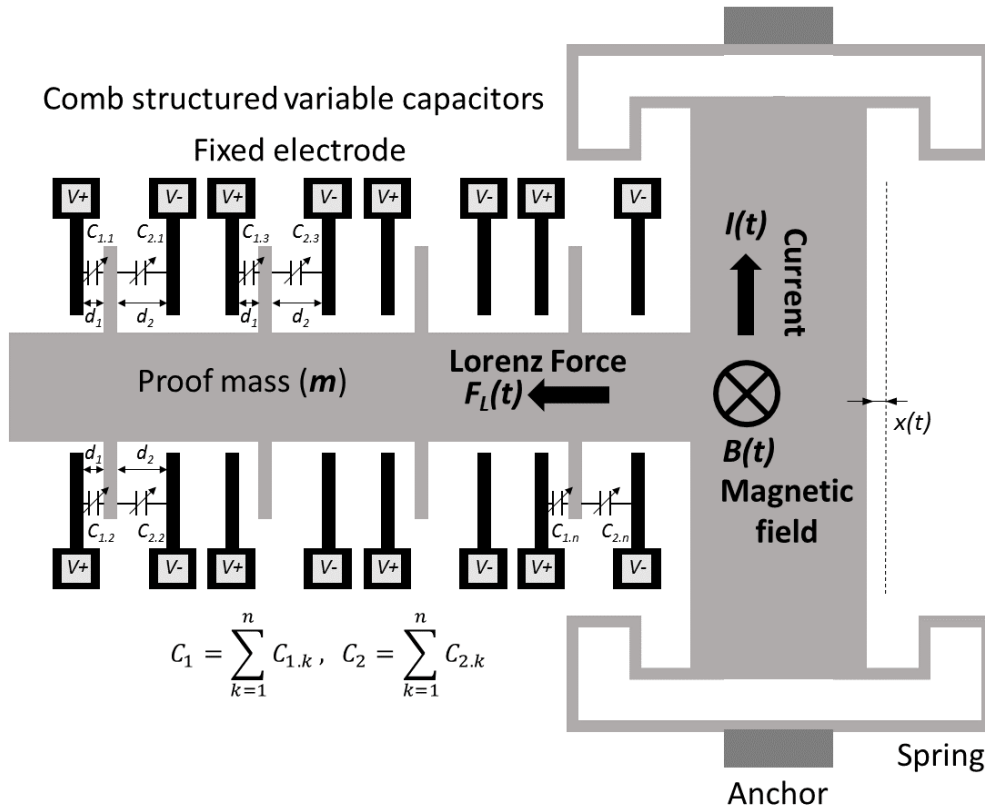
$$w_{motion} = w_{measured} - w_{earth}. \quad (1.11)$$

$$\theta = \int_0^t w_{motion} + \int_0^t n_w. \quad (1.12)$$

In the equations,  $w_{motion}$  and  $w_{measured}$  are the pure motion and measured angular velocity, respectively, and  $n_w$  is the noise of the gyroscopes. These accumulated errors cause drift



(a)



(b)

Figure 1.3. A magnetometer with (a) a modeling diagram (damping factor is ignored) and (b) an example layout with comb structure of variable capacitors.

effect or dead reckoning effect. However, fusion sensor technology can compensate for these errors. Usually, an IMU has accelerometers, gyroscopes, and magnetometers for the 3-D representation. To eliminate the gravity factor, the accelerometers can be supported by gyroscopes and magnetometers. A rotation matrix ( $\mathbf{R}$ ) is generated by a combination of accelerometers and magnetometers as below:

$$\begin{aligned} \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} &= \mathbf{R} \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix} \\ \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} &= \begin{pmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -g \end{pmatrix} \quad (1.13) \\ \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix} &= \begin{pmatrix} -g \sin \theta_x \\ g \sin \theta_x \cos \theta_y \\ -g \cos \theta_x \cos \theta_y \end{pmatrix} \end{aligned}$$

where  $a_x$ ,  $a_y$ , and  $a_z$  are measured accelerations along the  $x$ -,  $y$ -, and  $z$ -axis of the local coordinate system of the sensor, respectively. With three equations, two variables, roll ( $\theta_x$ ) and pitch ( $\theta_y$ ), can be obtained. Yaw ( $\theta_z$ ) can be estimated by measured magnetic fields ( $m_x$ ,  $m_y$ ,  $m_z$ ) with the estimated roll and pitch as below:

$$\begin{aligned} \begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} &= \mathbf{R} \begin{pmatrix} m_0 \\ 0 \\ 0 \end{pmatrix} \\ \begin{pmatrix} m_x \\ m_y \\ m_z \end{pmatrix} &= \begin{pmatrix} m_0 \cos \theta_y \cos \theta_z \\ m_0 \sin \theta_z \\ -m_0 \sin \theta_y \cos \theta_z \end{pmatrix} \quad (1.14) \end{aligned}$$

where  $m_0$  is the initial magnetic field toward the magnetic north as shown in Figure 1.4. However, the estimation equations are more accurate in static position of sensors than in moving scenario of sensors, so that gyroscopes should be combined with the estimation results to generate an accurate rotation matrix in both static and dynamic situations. The rotation matrix can eliminate the gravity factor from measured accelerations, and thereby accelerometers can obtain pure motion acceleration (zero-gravity). Gyroscopes also can

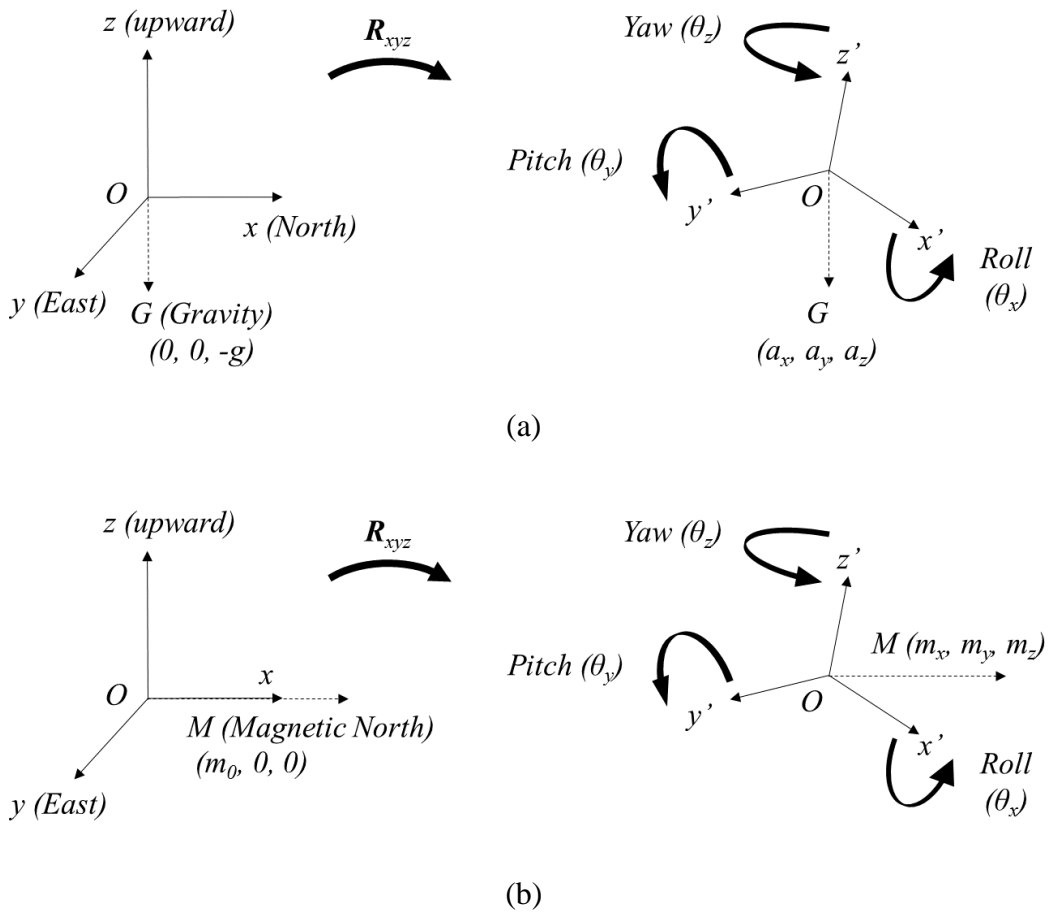


Figure 1.4. Local coordinate systems of sensors to express (a) gravity factor ( $g$ ) of measured acceleration in the initial (as the same as the global coordinate system) and the rotated position, and (b) the magnetic fields measured in initial ( $m_0$ ) and the rotated position ( $m_x, m_y, m_z$ ).

eliminate the earth rotation speed due to the rotation matrix. Nevertheless, the rotation matrix can be influenced by adjacent magnetic disturbances resulting from metal or other electronic devices. Although these disturbances affect magnetometers, the position information of the magnetic object can be recorded in calibration time, which can compensate for the magnetometer errors near the magnetic objects.

To improve the accuracy of the combination of devices in IMU, accumulated errors should be taken into account by referring noise information at the initial states (during



calibration). The estimation process based on Kalman filters can estimate a more accurate position or rotation matrix despite process and measurement noise (Li, and Wang, 2013). In addition, IMU can also be combined with other sensors. For instance, GPS receivers can help to compensate for inaccuracy of the position (Dixon, 1991; Obradovic, Lenz, and Schupfner, 2004), and cameras can support the correction of the direction of IMU (Hesch, Kottas, Bowman, and Roumeliotis, 2014). In terms of indoor positioning system, communication network infrastructures (e.g., wireless local area network (WLAN), ultra-wideband (UWB)) can improve IMU-based positioning systems, using received signal strength, time of arrival, and other temporal properties of the signal (Wang, Lenz, Szabo, Bamberger, and Hanebeck, 2007).

### **1.3. Background of Goals**

In line with the socSMCs project, this dissertation aims to contribute to the enhancement of interpersonal coordination in two special conditions: people with Autism spectrum disorder (ASD) and human-robot interaction conditions. This section will provide the background knowledge, issues, and possible applications in working on the interaction with people with ASD and robots.

#### **1.3.1. Interventions for Autism Spectrum Disorder (ASD)**

This dissertation suggests a single IMU solutions to apply to the interventions for people with ASD, by enhancing physical and social activities. Children and adults with ASD have difficulties in socializing and communication, according to the “Diagnostic and Statistical Manual of Mental Disorders, 5<sup>th</sup> edition (DSM-5)” released by the American Psychiatric Association (APA) in 2013. The symptom is inappropriate interaction, such as delayed reaction, repetitive or one-sided conversation, and lack of eye contact, to name a few. These symptoms are mainly based on works of Autism Disorder (Kanner, 1943) and Asperger’s disorder (Asperger, 1944), respectively. The term “Autism” firstly appeared as a pervasive developmental disorder in DSM-3, which is previous version of DSM issued in 1980. Asperger’s disorder was firstly defined as a separate disorder from autism in DSM-

4 issued in 1994. As a majority of professionals regarded Asperger's disorder as a mild form of autism, DSM-5 (2013) introduced a unified definition, ASD, which combines autistic disorders, Asperger's disorders, and other pervasive developmental disorders. From observation, people with ASD are affected due to different conditions. Delayed or impaired development of childhood is involved. Even if one's development is not problematic, isolation can cause ASD. Interventions, therefore, differ depending on causes of ASD. Although the symptoms are defined, there are no specific criteria to distinguish people with ASD. Many practitioners can, therefore, diagnose ASD by observation, which is usually a partially subjective assessment.

The rapport analysis of H-IMU solution (proposed by Chapter 3, see also Hwang, Effenberg, and Blume, 2019a) can be helpful for the diagnosis of ASD, suggesting an objective reference. The reason is that the rapport monitoring solution includes the number of nods, and head orientations, which can support analysis of one's verbal and nonverbal head reactions. Gait analysis can aid in analysis of the motor development of children. The single H-IMU solution is easy to access and to wear, even for young children. In walk and talk intervention, language development can be also analyzed with the audio recordings during conversation. In terms of the tetherball paradigm in Chapter 4 (see also Hwang et al., 2018a), an IMU-based tablet solution can monitor behaviors of people with ASD during a joint task. This tablet solution can measure one's high level of interactive skills because it involves proximal settings and object sharing between two interactants—one of the most difficult conditions for individuals with ASD. Based on these analyses, the diagnosis can be more objective, and the progress of intervention can be monitored. The measurement data can be used for long-term and real-time feedback, which can lead to a successful intervention.

### **1.3.2. Interaction with Robots**

Robots have been investigated and developed to communicate with humans. Human-robot interaction is involved in an interdisciplinary field: human-computer interface, robotics, artificial intelligence (AI), and social science. In the manufacturing

field, automation systems and software provide intuitive interfaces for technicians. Industrial robots should coordinate with other human workers to share the same work space and to enable fluent cooperation. For consumer electronics, AI technology supports translation and autonomous driving, as well as virtual assistant applications. Nowadays, consumer social robots are emerging to work for service, rehabilitation, caregiving and education industries (Muthugala et al., 2013). Robots can be understood as human-like social agents when robots are able to imitate human body shapes as well as mimic human facial expressions (Breazeal and Scassellati, 2002). For proper responses, the robots need to understand human gestures, cues and emotions. The robots are also supposed to identify customers and provide an individualized service.

To support the promising applications, robots should have a database of human gestures and social cues, thereby developing their social intelligence. Compared to other body gestures, the head movements provide rich information related to agreement, attentions, desires, beliefs, and emotions. To be specific, the head orientation is related to gaze position where one's interest lies. The head pose can also indicate that users are curious, happy, or sad, as well as if they agree or not. Studies on coordination with the eyes, the neck, the trunk, the hand and the foot would inform human-like movement (Muthugala et al., 2013, Kavanagh, Barrett, and Morrison, 2006, Vercher et al., 1994). Relationship between the head and the linguistic context can also lead robots to move like humans during conversation with robot users (Cabibihan, Javed, Ang, and Aljunied, 2013). With accumulated head kinematic data, robots can distinguish others' mental states based on the emergent reactions (e.g., facial expressions, body postures, tone of voice; Cabibihan et al., 2013). Based on these observations, robots can properly react, which can build a trustful relationship with human co-actors. The rapport and gait monitoring system equipped with IMUs can accumulate human head kinematics. When the social robots are released in the mass market in the future, their rapport building abilities can be also tested simply using H-IMU sensors during walk and talk with humans.

In proximal settings, robots can participate in joint action-settings. Their interactive skills can also be tested in a joint task such as the tetherball paradigm in Chapter 4 (see

also Hwang et al., 2018a). During this type of joint tasks, robots can be assessed in terms of the high level of interactive skills. Firstly, the robot has to discriminate their own and partner's effect on the ball, respectively, which is a basic interactive skill—important to support a high level of interactions. Secondly, the robot has to consider the movement of the partner and the ball, simultaneously. With the ability in handling the complicated situation, the robot is regarded as having the highest level of joint attention ability. Lastly, the robot has to understand not only the partner's intention, but also what the partner predicts that the robot is going to do, which requires the second order Theory of Mind decoding (Mehta, Bhagyavathi, Kumar, Thirthalli, and Gangadhar, 2014; Scassellati, 2002). The robot can accumulate the database of immediate reactions of humans in response to the robot movement. These nonverbal communication interactive skills can enable the robot to co-operate with humans at work places and at home, especially with a reference to physical work.

## **Chapter 2. Measurement of Gait Parameters**

### **2.1. Introduction**

Gait analysis can inform us of the number of steps, travel trajectory, and calorie consumption. In everyday life, gait is a key feature in monitoring health. Gait speed is an important vital sign, such as blood pressure, pulse, respiration, temperature, pain, and heart-beat rate (Fritz and Lusardi, 2009), which is supported by the strong correlation between seniors' motility and their gait speed (Studenski, Perera, Patel, Rosano, Faulkner, and Inzitari et al., 2011). Additionally, gait balance can contribute to the prevention of injuries. For instance, monitoring gait balance control can prevent seniors from fall injuries (Bridenbaugh and Kressig, 2011), and help to manage patients after concussion (Howell, Osternig, and Chou, 2015). Foot-ground contact time (CT) can be applied to the postoperative rehabilitation for hip replacement (Reh, Hwang, Michalke, and Effenberg, 2016; Reh, Hwang, Schmitz, and Effenberg, 2019). The variability of gait stride time can aid in preventing overuse injuries of loaded and strenuous walking during military training (Springer, Gottlieb, and Lozin, 2016). These evidences support that a number of gait parameters play an important role in the prevention of injuries and the management of health status.

Today, smartphones, wearable devices or internet of things (IoT) systems can provide gait monitoring system in global market (Brajdic, and Harle, 2013; Yamada, Aoyama, Mori, Nishiguchi, Okamoto, and Ito et al., 2012; Atallah, Aziz, Lo, and Yang, 2009; Majumder, Saxena, and Ahamed, 2016). The gait monitoring systems can predict one's fall (Majumder et al., 2016) and recognize a pathologic gait (Yamada et al., 2012; Atallah et al., 2009). A key issue is that most of the wearable devices fail to maintain long-term engagement for consumers because of less connection to one's behaviors and habits (Ledger and McCaffrey, 2014). Nevertheless, pedometers are still investigated to enhance the sustainability of impact on the elderly (Harris, Kerry, Victor, Ekelund, Woodcock, and Iliffe et al., 2015) and on the young (Inchley, Cuthbert, and Grimes, 2007; Rye, Zizzi, Vitullo, and Tompkins, 2005). Pedometers are appropriate for interventions that encourage

participants walk more because they measure habitual activity and provides daily motivations (Harris et al., 2015, Inchley et al., 2007; Rye et al., 2005). Furthermore, wireless earbuds with IMUs, which can support walking detection, are already penetrated to the mass market of health monitoring device (Hunn, 2016). When well-known and widely-used products (e.g., earbuds) serve as human health monitoring systems, the users can feel a seamless experience in everyday life due to their friendliness. The simple IMU-based solutions can be combined with cameras, navigation applications, smart insoles and force plates, providing gait and health information with high quality and quantity.

This chapter demonstrates that the H-IMU has the ability to provide more gait parameters than the pedometer. The proposed system of gait analysis works in a wireless communication setting and in real-time. The system can provide temporal detection of the heel strike (HS) and the toe off (TO). Mid-stance phase can be also estimated. Based on the gait events, spatiotemporal gait parameters are measured, such as the stride length, step time, foot-ground contact time (CT) and contact time ratio (CTR). To demonstrate the applicability of everyday life settings, the experiment was performed outdoors. In the following section, the related work is described. In the third section, biomechanical terminologies and the name of events are explained during a gait cycle. The detection methods of the gait events are described in the fourth section, followed by the computation method of gait parameters in the fifth section. The sixth section deals with the results, the limitations and contributions is discussed in the seventh, and finally the conclusion follows.

## **2.2. Related Work**

### **2.2.1. Measurement of Human Movement**

Thousand years ago, Aristotle (384–322 BCE) reported anatomical theory of animals and humans, including human gait. Although most of the parts were proven as false later, the particular observation of the movement was proven as a truth, which is the first written reference of the assessment of human movement. In the nineteenth-century, measurement of human movement was developed with electrophysiology. For gait analysis, pressure sensors were used for the localization and force analysis of the feet (Nilsson,

Stokes, Thorstensson, 1985). Electromyography (EMG) is also used to analyze muscle kinetics (Muro-De-La-Herran, Garcia-Zapirain, and Mendez-Zorrilla, 2014). In addition, photography contributes to the measurement of human kinematics, being developed to video analysis technology (Davis, Ounpuu, Tyburski, and Gage, 1991). In film industry, 3 dimensional (3-D) motion capture technology was developed for 3-D animation or science fiction (Sci-Fi) movies. This technology was adopted for study on human movement, gait analysis (Pfister, West, Bronner, and Noah, 2014) and rehabilitation (Brock, Schmitz, Baumann, and Effenberg, 2012; Schmitz, Kroeger, and Effenberg, 2014).

### **2.2.2. Motion Capture System**

Motion capture is used to record human movement normally in 3-D space. Motion capture has been developed for different uses in film making, gaming, and clinical training. Motion capture systems are mainly divided into two types: optical and non-optical systems. Optical systems are utilized with vision technology. A combination of cameras and markers is the most well-known optical systems. The markers on the actor reflect the light emitted from the external optical source, or emit their own light. The light can be in different spectrum range of optical waves: visible light, infrared ray (IR) (Kirk, O'Brien, and Forsyth, 2005). In a certain distance from the actor, a multiple number of cameras, normally more than two, record the spatial position of the illuminating markers. Motion capture software computes the motion of the markers, and finally regenerates an avatar imitating the actor's motion in 3-D space. Recently, motion capture systems are available even without markers, called the marker-less system. One of popular applications is using dual vision technology which provides image with depth information with only two cameras; however, it fails to regenerate virtual characters in full 3-D space. The simplicity is, nevertheless, beloved in robotics, gaming industry and a number of human movement research (Moeslund, Hilton, and Krüger, 2006; Wandt, Ackermann, and Rosenhahn, 2016; Wandt and Rosenhahn, 2019). On the other hand, straightness of light results in hidden places. Cameras can fail to recognize makers hidden by parts of human body, which should be compensated for high quality 3-D animation (Silaghi, Plänkner, Boulic, Fua, and

Thalmann, 1998). A multiple number of cameras causes high cost, and the actors movement is restricted to the stage where the cameras' angles are reached. Marker-less system can reduce cost but its accuracy depends on ambient light.

Development of sensor technology allows non-optical motion capture systems. Mechanical sensor systems and magnetic marker systems were developed earlier and contributed to the study of human movement. Mechanical sensor systems measure bending angles of wires associated to the joints. Disadvantage of mechanical systems is the limitation of human motion because of the dynamic range of wires and apparatus. Magnetic marker systems utilized the level of adjacent magnetic field of the actor. This system is, however, vulnerable to magnetic distortion that stems from metal or electromagnetic devices. With development of microelectromechanical systems (MEMS), IMU was invented (Brodie, Walmsley, and Page, 2008), which is used in motion capture. The inertial sensors provide linear acceleration or angular velocity, regenerating kinematic parameters using integration or derivation. For instance, the global and relative positions of the body segments are estimated from double integration of acceleration. Joint angles are also computed from integration of the rate of turn. Integration of inertial sensor data, however, causes error accumulation (dead reckoning effect), which causes drift effect.

The IMU has recently been developed to compensate for the previous disadvantage, such as the drift effect. Cutting edge MEMS technology improves signal to noise ratio (SNR) itself. Sensor fusion technology supports the compensation for accumulated errors. Magnetometer (Rios and White, 2002), global positioning system (GPS) (Hellmers, Norrdine, Blankenbach, and Eichhorn, 2013), or camera (Hesch et al., 2014) is combined with inertial sensors (accelerometers, gyroscopes) and reduce drift effects. Unlike with optical systems, in addition, motion capture with an IMU system is free from fixed reference frames such as fixed stages and cameras, thereby supporting the kinematic measurement of human movement not only indoors (Hwang, Reh, Effenberg, and Blume, 2016), but also outdoors (Brodie et al., 2008, Hwang et al., 2018b). Motion capture technology based on IMU has been quickly matured and commercialized, which is suitable for this study because of the accuracy, and the freedom from the spatial limitation.



### **2.2.3. Reduced Number of Sensors and Sensor Placement**

Inertial measurement units are easily found in daily life because of the spread of smartphones and wearable devices. Consumers can download navigation and health monitoring applications. Accuracy and precision of positioning and movement tracking functions are improved with map applications. A pedometer function provides the number of steps, which leads to estimating the amount of daily activity and calorie consumption. With smart devices, remarkable reduction in the number of sensors is shown in gait analysis applications. Research on gait analysis is typically conducted with integrated sensor insoles of shoes or with both the camera-marker-based and sensor-based motion capture systems by being fixed onto the pelvis, thigh, tibia, and foot (Rueterbories, Spaich, Larsen, and Andersen, 2010). Lower limb joint angles and foot orientation are normally analyzed. Regarding a real time solution, IMU can be used with pressure sensors on the feet or ground that directly provides information about foot places and force during walk. These settings with additional sensors and wired devices are usable only within laboratory conditions, but not in daily life settings. Therefore, simple solutions with an IMU in mobile application have been developed. Advanced gait analysis with mobile applications can recognize pathologic gait from healthy gait and distinguish run and walk. The IMU embedded in a smartphone can detect the number of steps, when placed in a hand, a handback, a backpack, and trouser pockets (Brajdic, and Harle, 2013), which demonstrates the simplicity of a single sensor solution and the independency from the measurement location.

### **2.2.4. Previous Work Using a Head-Worn Sensor**

A head-worn IMU has been researched regarding head stabilization during walking (Cromwell and Wellmon, 2001; Kavanagh, Morrison, and Barrett, 2005; Kavanagh et al., 2006). Head acceleration was analyzed to measure a coordination of the neck and the trunk in 3-D space (Cromwell and Wellmon, 2001), which found reliably regular patterns in head acceleration during walking. Although head acceleration is affected by the stabilizing effect of the trunk and neck (Cromwell and Wellmon, 2001; Kavanagh et al., 2005; Kavanagh et al., 2006), it still contains information of gait events, which allows the head

acceleration to support gait pattern analysis (Hirasaki, Kubo, Nozawa, Matano, and Matsunaga, 1993). A wireless ear-worn sensor was reported in the study on gait pattern analysis, which also recognizes whether gait patterns are pathologic or not (Atallah et al., 2009). In another study, the ear-worn sensor can detect heel strike and toe off. Data sets of gait cycle are compared to classifying gait events by using singular spectrum analysis (SSA) and longest common subsequence (LCSS). These methods intrinsically require a several number of gait cycle delay because it needs to compare with kinematic data in previous gait cycles. Nevertheless, at least heel strike could be detected in real time by monitoring certain dominant oscillations of the signal (Jarchi, Wong, Kwasnicki, Heller, Tew, and Yang, 2014). Finally, with an advanced peak detection algorithm, it has been reported that a head worn IMU can detect heel strike and toe off in real time with minimal delays (Hwang et al., 2018b). This real time TO detection algorithm has been first reported, thereby allowing the real time measurement of more gait parameters, such as contact time (CT) and contact time ratio (CTR).

### **2.3. Gait Events in Gait Cycle**

Gait events were detected by head worn IMU (H-IMU). For better understanding, terminology of gait analysis is introduced. A gait cycle is a repetitive sequence of specific events during walking. The gait cycle is mainly divided into two phases (stance phase, swing phase) by two gait events (heel strike (HS), toe off (TO)). The stance phase is a period after the heel strikes the ground before the toe gets off. During the stance phase, a leg supports the upper body and applies forces for locomotion. Swing phase is the other period when a foot swings in the air moving forward after the toe gets off the ground. The swing phase starts at TO and ends at HS, whereas the stance phase starts at HS and ends at TO. Other spatiotemporal gait parameters are defined by the time points of two gait events. A gait cycle is illustrated in Figure 2.1 with exemplary gait motions and the time diagram. The gait motions are depicted in the right side view, being projected on the sagittal plane. The time points of gait events, the period of gait phases, and gait parameters are described in Figure 2.1. The gait cycle in Figure 2.1 starts at the time point of a right HS, which is

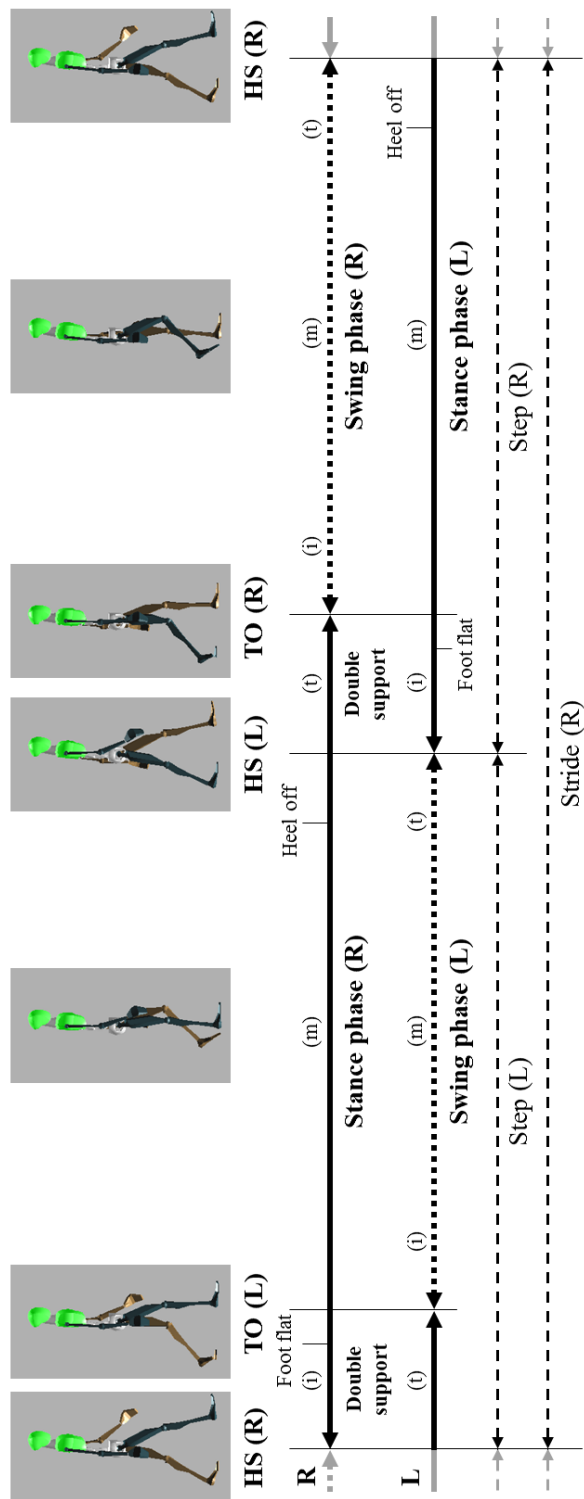


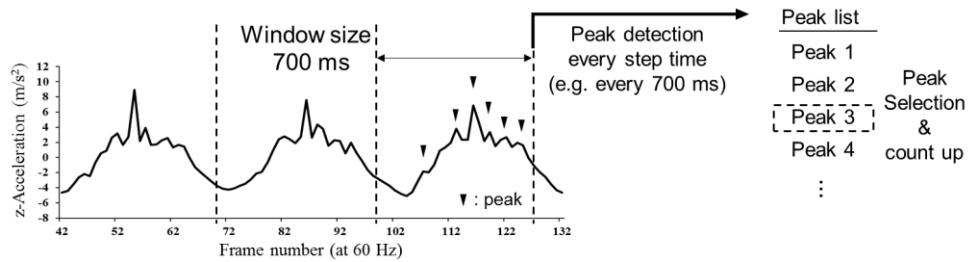
Figure 2.1. Time diagram of a gait cycle with example pictures at walking gait events which are from the right side view on the sagittal plane. Abbreviations of ‘HS’ is heel strike, ‘TO’ is toe off, ‘R’ is right, ‘L’ is left, ‘(i)’ is initial phase, ‘(m)’ is mid-phase, and ‘(t)’ is terminal phase.

the start of the right stance phase. Directly after the first right HS, the right foot flat comes, which means the whole foot from heel to toe lies on the ground. Then a left TO follows, terminating the left stance phase, and initiating the left swing phase. This is the beginning of the mid-stance phase for the right foot, which includes the mid- and terminal swing phase for the left foot. When right heel gets off the ground, the terminal stance phase of right foot starts. The left HS follows and the left stance phase starts. Before the right TO, both legs support the body, which is defined as a double support phase. The time period between a right HS and the next left HS is a left step time. At the left HS, a right step starts with the same sequence as described before, but with the contralateral foot. At the right TO, the right stance phase is finished. The stance phase is from an HS of one foot to the next TO of the same foot, also called the foot-ground contact time (CT) which corresponds to the sum of a step time and the following double support time. The CT also depends on the gait velocity, but the ground contact time ratio (CTR), defined as a ratio of the CT to the stride time, is approximately 60%. For the left foot, foot flat follows, heel off appears at the end of the left stance phase, and at the followed right HS a gait cycle is finished. The right step time is defined between a left HS and the next right HS. The period from an HS of one foot to the next HS of the same foot is a stride time. This gait cycle is continuously repeated during walking.

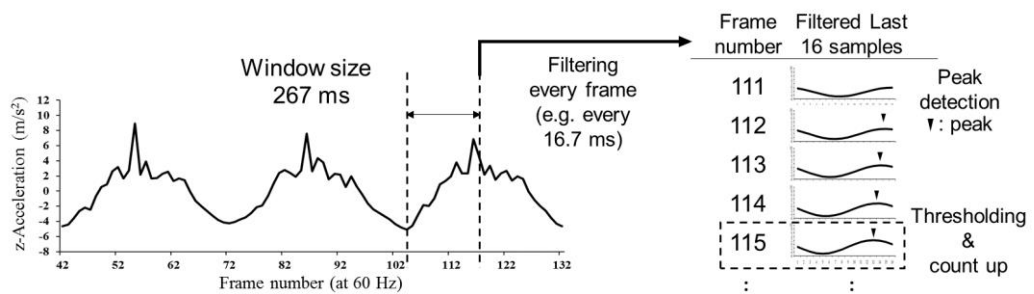
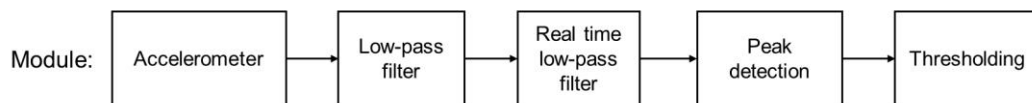
## **2.4. Methodology of Gait Event Detection Using H-IMU**

### **2.4.1. System Overview**

A single H-IMU facilitates gait analysis with a peak detection algorithm. Peak detection is applicable to gait analysis, when IMU modules are attached on various body parts, such as the feet, tibia, thigh, hip, and pelvis (Yamada et al., 2012). In smartphone applications as a single IMU solution, a peak detection showed the highest accuracy in step counting (Brajdic and Harle, 2013). The peak detection is, furthermore, given with a high reliability for different smartphone placement such as being held in hand, placed in the pockets, and even located in the handbag and the backpack (Brajdic and Harle, 2013).



(a)



(b)

Figure 2.2. Comparison of two peak detection algorithms: (a) the windowed peak detection (WPD) which analyzes every 1 step time (optimally 700 ms), and (b) the proposed peak detection which analyzes every frame (16.7 ms at 60 Hz).

Table 2.1. Comparison of Single IMU Methods for Gait Analysis

Comparison	H-IMU (XSENS; head)	E-AR	Smartphones
Reference	Hwang et al. 2018b	Jarchi et al. 2014	Brajdic et al. 2013
Number	1	1	1
Type	IMU	Accelerometers	IMU
Degree of freedom	9	3	9
Update rate at sensor output (Hz)	60	100	60
Place	The head	The head	Various
Algorithm	<sup>a</sup> RLF, <sup>a</sup> PD & thresholding	<sup>a</sup> SSA, <sup>a</sup> LCSS & <sup>a</sup> PD	<sup>a</sup> WPD
Delay	HS	Real time (30 ms + <sup>b</sup> $e_1$ )	1 gait cycle (optimally 0.7 s)
	TO	Real time (30 ms + <sup>b</sup> $e_2$ )	-

a: RLF= Real time low pass filter; PD= Peak detection; SSA= Singular spectrum analysis; LCSS= Longest common subsequence; WPD= Windowed peak detection.

b:  $e_1, e_2$  = temporal errors (10–55 ms); in section 2.6.3, it was calculated between data from foot worn IMUs and the H-IMU.

The proposed H-IMU system should be supported by several error canceling technologies which are applied in IMU-based motion capture systems or smart devices. For instance, with sensor fusion technology, accurate kinematic data can be generated by reducing error accumulation resulted from integration. The gravity and earth rotating velocity are eliminated from measured data due to the rotation matrix, which results in pure motion data. Various types of Kalman filters play an important role in combining two or more sensors and reducing offset and white noise for more accurate results in sensor positioning systems. Different kinds of FIR filters are also applied to reduce different kinds of noise. Nevertheless, unwanted peaks are detected during walking because of movements of the walkers such as head shakes, nods, and direction changes.

Windowed peak detection (WPD) can be one of the solutions to handle unwanted actions. In a study of step counting (Brajdic and Harle, 2013; see also Figure 2.2(a)) using a smartphone with an IMU placed in the hand, handbag, backpack, and pockets of trousers

during different activities, the WPD demonstrated the most accurate results, compared to other methods on time domain (Mean Crossings Counts, MCC; Normalized Autocorrelation Step Counting, NASC), on frequency domain (Short Term Fourier Transform, STFT; Continuous/Discrete Wavelet Transform, CWT/DWT), and of feature clustering (Hidden Markov Models, HMMs; K-Means Clustering, KMC). In WPD, the highest peak is detected even though other activities occur in a period of window size (walkers' average step time, optimally 700 ms). This window size, however, becomes the time delay of the step counting, which is problematic for real time solutions because a step is detected at the next step, and also it takes time to measure the average step time in the beginning of the step counting. Therefore, in this dissertation, a peak detection algorithm is used with real time low-pass filter (RLF) for the real time solutions. The system diagram is shown in Figure 2.2(b). The algorithm detects peaks on vertical head acceleration (z-axis; see Figure 2.3) to recognize the influence of the foot on the head. The impacts are transmitted from the foot along the longitudinal body axis, some of which are regarded as HS or TO. Table 2.1 shows the comparison between H-IMU system, E-AR (Jarchi et al. 2014), and smartphone methods (Barjdic and Harle 2013). In a comparison of systems, H-IMU has a delay at 30 ms, resulted from the XSENS MVN system. Temporal errors ( $e_1$  and  $e_2$ ) are added on it because the detection results of H-IMU is temporally different from the ground truth. In the section 2.6.3, the temporal errors are calculated between results from foot worn IMUs and the H-IMU. For E-AR, most of the gait parameters take several gait-cycle time, except for HS which is detected in real time and delay is not reported by Jarchi et al. Various methods with smartphones are tested and WPD is reported as the most optimal solution in terms of simplicity and accuracy, which needs 1 gait cycle (optimally 700 ms). So far, TO detection with WPD was not reported in the research literature.

#### 2.4.2. Peak Detection

A peak in sensor signals is detected from the comparison between differential values near the sample of interest. When two signs of differential values before and after the sample are different, the sample is the peak. Sensor signals, however, include many

positive or negative peaks, which might be the signals of interest or noise. To eliminate unwanted peaks or noise, the signals are filtered and windowed as shown in Figure 2.2(a) (Brajdic and Harle, 2013). For the gait event detection in this dissertation, peaks of interest were distinguished by using peak detection with real time low-pass filter (RLF) and thresholding algorithm. Whereas the WPD (Figure 2.2(a)) detects several peaks and recognizes an HS with the highest peak in a window (around 700 ms), the RLF (Figure 2.2(b)) detects one high peak after smoothing all other small peaks in smaller temporal window (16 samples, 267 ms at 60 fps). As a random signal processing, RLF runs at every frame. The details of the operation will be explained in the next subsections. In this subsection, the peak detection on the head accelerometer is explained and the comparison with acceleration data in other body parts are described.

Vertical head accelerations ( $z$ -axis) of four body parts are depicted in Figure 2.3, which are measured by an IMU-based motion capture system, named XSENS MVN system. From the top, the four graphs are arranged in order for the left foot, right foot, pelvis, and head. These four body parts provide different evidence of gait event detection at HS and TO. In the first and the second graph for the feet acceleration shown in Figure 2.3, the highest peak appears at HS. During a stance phase, the acceleration keeps stable without oscillation, and then several peaks occur again before TO. The last negative peak before a dramatic acceleration drop is the time point at TO. In this dissertation, data from foot-worn IMU were regarded as reference because feet are the nearest to detect the impact between the foot and ground, and foot-worn IMU-based methods were used and validated by previous studies (Reh et al., 2019; Bailey and Harle, 2015; Mariani, Hoskovec, Rochat, Büla, Penders, and Aminian, 2010). In the third graph for the acceleration of the pelvis, normally several large positive peaks appear, which can cause confusion to detect the peak for the HS. Near HS, double peaks can appear. In this case, the center of mass is gradually higher until the terminal swing phase, and goes lower at HS, which forces the pelvis down as well. After HS, the grounded leg supports the trunk, which forces the pelvis to shift up again, which was observed in this study when walkers had a long step length or a fast step time. On the other hand, a single peak can also appear at HS. In this case, the acceleration



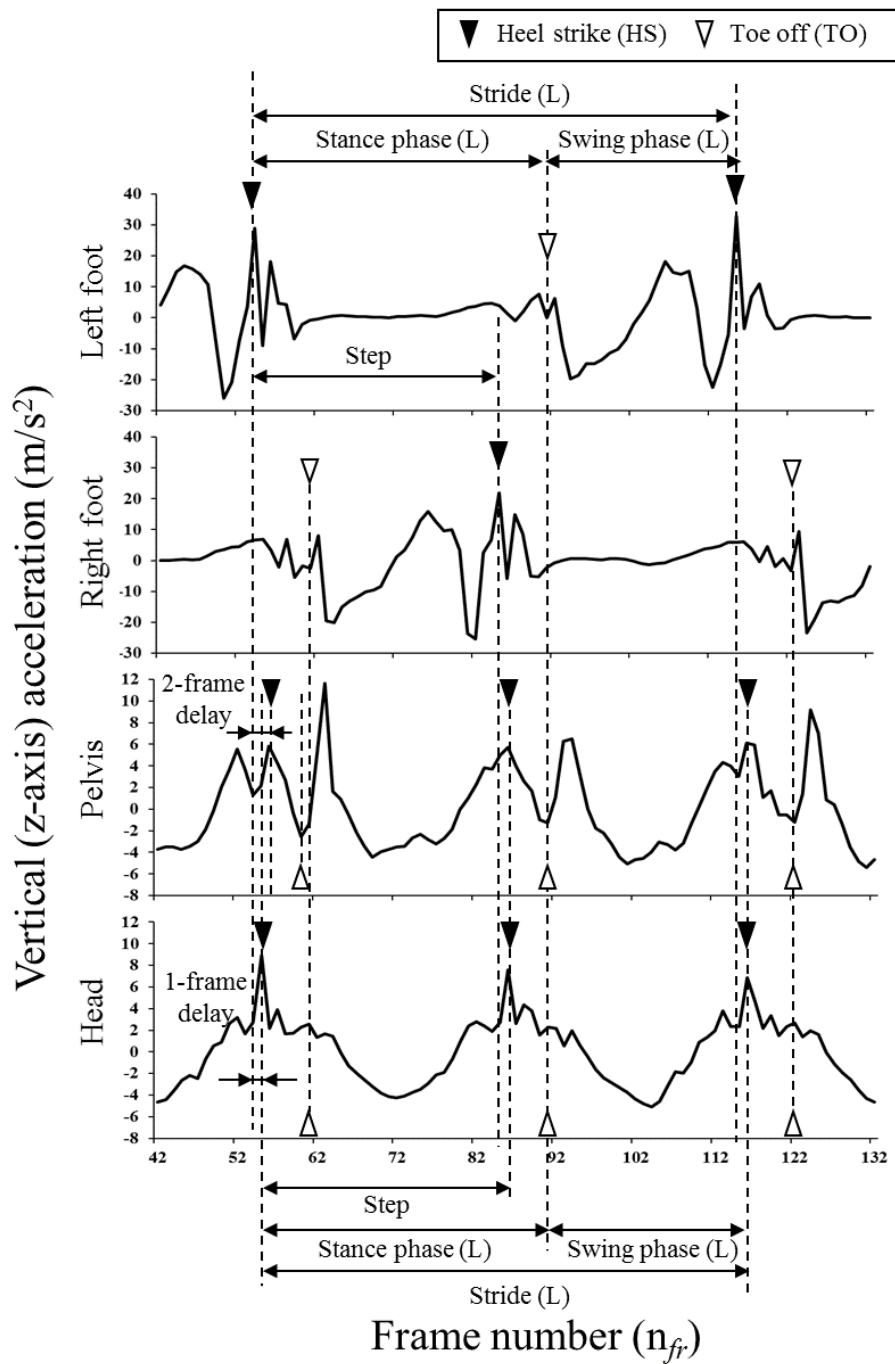


Figure 2.3. Vertical acceleration of body segments of the left foot, right foot, pelvis and head which are measured by an XSENS MVN system at 60 Hz.

stops increasing or a small negative peak appears at HS, and then pelvis moves up, which appears on the pelvis acceleration one or two frames later than HS. The HS detection based on the pelvis acceleration is, as a result, difficult because the two cases are varied depending on the individual physical condition and the gait styles. When it comes to a TO, another large peak appears, following a large negative peak at the TO. The negative peak appears because the knee bends to generate the propulsion force, during which the pelvis position is down just before the TO. After TO, the propulsion force pushes up the pelvis and a high positive peak appears. With pelvis acceleration, labeling HS and TO is not simple because the amplitude of their peaks are similar, and thereby it is hard to define an optimal threshold value. In the last graph for the head acceleration, the first large peak appears at HS. This large peak is easier to detect because it is usually the maximum value in a gait cycle, so that the threshold value can be optimized easily. For the TO, a positive peak appearing after HS and the foot flat peak is regarded as TO. With head acceleration, the variation of the waveforms is smaller than the pelvis acceleration because the head is stabilized by the trunk and neck (Cromwell and Wellmon, 2001; Kavanagh et al., 2005; Kavanagh et al., 2006).

Regarding the foot acceleration in Figure 2.3, the relatively high peaks appear at the moment of gait events, and the peak values are over  $30 \text{ m/s}^2$  at HS and around  $10 \text{ m/s}^2$  at TO, which are the distinguishable level of values. At HS and TO, however, other high peaks appear, and could create confusion about which is the peak of interest. During the terminal swing phase, high peaks appear, which negatively affects the accuracy in detecting gait events. Moreover, at least two IMUs are needed to detect two feet's gait events, which can be a disadvantage, compared to a single IMU on the head or the pelvis that can detect two feet's gait events as shown in Figure 2.3. Concerning peak detection for gait event recognition, head acceleration has a higher accuracy than the pelvic acceleration (Hwang et al. 2018b) because head acceleration shows less double peaks compared to pelvic acceleration, which is less confusing in detecting HS and TO. Although the pelvis is localized closer to the foot, the head peak acceleration is temporally the same or closer to that of the foot, compared to the pelvis peak acceleration as shown in Figure 2.3.

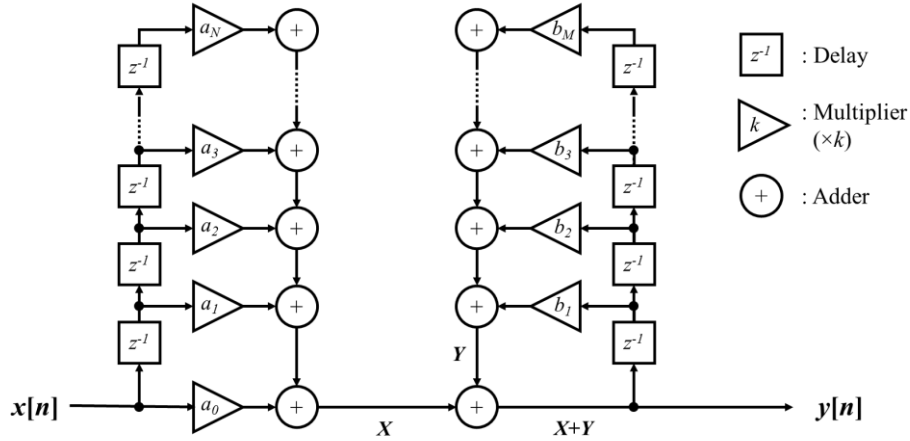


Figure 2.4. A block diagram of the general digital filter, where  $x[n]$  and  $y[n]$  are the input and output signal, respectively. As integer numbers,  $N$  and  $M$  are the filter orders. The values,  $a_i$  and  $b_j$ , are the impulse response for  $i$ -th and  $j$ -th components.

### 2.4.3. Digital Filters

As a positioning system, IMU-based systems need to reduce noise and estimate the accurate position. These systems implement two types of digital filters: a finite impulse response (FIR) (Bialkowski, 1988a) and an infinite impulse response (IIR) filter (Bialkowski, 1988b). An FIR filter responds to the recent finite length input, whereas an IIR filter indefinitely responds to the input, utilizing feedback loops. Figure 2.4 demonstrates the block diagram of the general digital filter (Oppenheim, Buck, and Schaffer, 2001). The output signal is computed as below:

$$\begin{aligned}
 X &= \sum_{i=0}^N a_i x[n-i], & Y &= \sum_{j=1}^M b_j y[n-j]. \\
 y[n] &= X + Y = \sum_{i=0}^N a_i x[n-i] + \sum_{j=1}^M b_j y[n-j]
 \end{aligned}
 \tag{2.1}$$

where  $X$  is the weighted sum of the current and past input values and  $Y$  is the weighted sum of the past output values. Output signal,  $y[n]$ , is sum of  $X$  and  $Y$ . The integer numbers,  $N$  and  $M$ , are the order of the filter for intermediate results,  $X$  and  $Y$ , respectively. For the

coefficients,  $a_i$  is the impulse response of an  $i$ -th component of the resulting  $X$ , whereas  $b_j$  is the  $j$ -th component of the resulting  $Y$ . If only  $X$  exists, it is an FIR filter because the output,  $y[n]$ , results from the finite number of past input signals. With FIR filters, the output is inherently stable, and has no phase shift due to the linear phase property, which is suitable for phase-sensitive applications. In the case of an IIR filter,  $y[n]$  includes sum of the past output signals,  $Y$ . As  $Y$  includes a series of input signals,  $y[n]$  is finally a result of the infinite series of input signals. Due to the feedback structure, the IIR filter is sometimes unstable, but it has a higher efficiency than FIR filters for the same results.

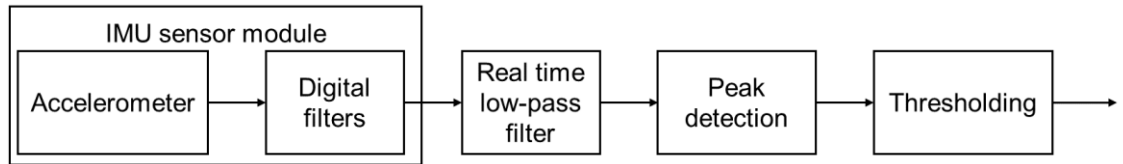
In this dissertation, the RLF was designed based on an FIR filter, which takes 16 samples at every frame. In (2.1),  $N$  is 15, so that the range of  $n$  is from 0 to 15 (16 samples), and the feedback result  $Y$  is 0 because it is an FIR filter. The number of samples was decided to correspond to about half of a step time for optimization.

As the proposed H-IMU uses the sensor output data of an IMU-based motion capture system (Xsens MVN BioMech System), the input data of RLF are already filtered by the motion capture system including Kalman filters, which are one well-known type of IIR filters. These filters use statistical information of input signals, which is desirable when uncertainty exists in measurements, even in non-linear systems, thereby reducing the drift effect and statistical noise in measurements (e.g., white noise). The Kalman filters are designed with a system modeling based on state equations expressing a set of differential equations, which is suitable for positioning systems because of the relationships of acceleration, velocity, and displacement. Various Kalman filters have been developed and used to navigate and to control the movement of aircrafts, watercrafts, and automobiles.

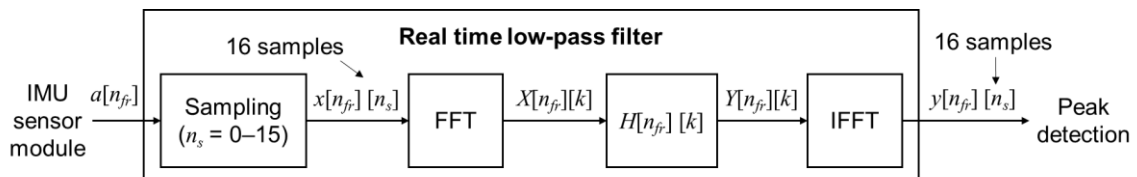
As shown in Figure 2.5(a), the output of the IMU sensor module is the input data of RLF. For the 3-D rendering animation of the motion capture system, the frame update rate is 60 Hz, which is the update rate of RLF input data. The convolution of the RLF input signal and the impulse response of a low pass filter is computed in the frequency domain. By using fast Fourier transform (FFT), the convolution is computed as a multiplication of  $X[n_{fr}][k]$  and  $H[n_{fr}][k]$ , and the inverse fast Fourier transform (IFFT) results in  $y[n_{fr}][n_s]$  (Figure 2.5(b)). The method using FFT and IFFT can reduce the computational complexity

of the convolution from  $O(N_s^2)$  to  $O(N_s \log N_s)$ , thereby reducing the load of processors ( $N_s$ : the number of samples of RLF input). Despite the small number of samples, it is important because the filter has to be computed at every frame in the limited frame time (e.g., in 16.7 ms at 60 Hz).

The RLF runs when the 3-D rendering frame is updated at 60 Hz. At every frame, 16 samples ( $x[n_{fr}][n_s]; n_s = 0-15$ ) of the latest input (Figure 2.6(b)) result in 16 samples of output ( $y[n_{fr}][n_s];$  Figure 2.6(c)). To be specific, Figure 2.6(a) shows the head z-acceleration at frames  $n_{fr} = 35-80$ . At frame 56 ( $n_{fr} = 56$ ), acceleration data,  $a[41] - a[56]$ , are taken as input data of RLF ( $x[56][n_s]; n_s = 0-15$ ) as shown in Figure 2.6(b). For the next step, after RLF, the output has 16 samples ( $y[56][n_s]; n_s = 0-15$ ) as the results of the circular convolution ( $x[56][n_s] * h[56][n_s]$ ). The low pass filter eliminates small peaks as shown in Figure 2.6(c). Figure 2.7 shows the increasing amplitude of the output signals, when the frame is closer to HS. When the peak is higher than the threshold, the frame number  $n_{fr} = 54$  is timestamped as an HS. Although the head peak appears at Frame 55, HS is detected at Frame 54 because threshold value is lowered considering the foot-head delay at HS.

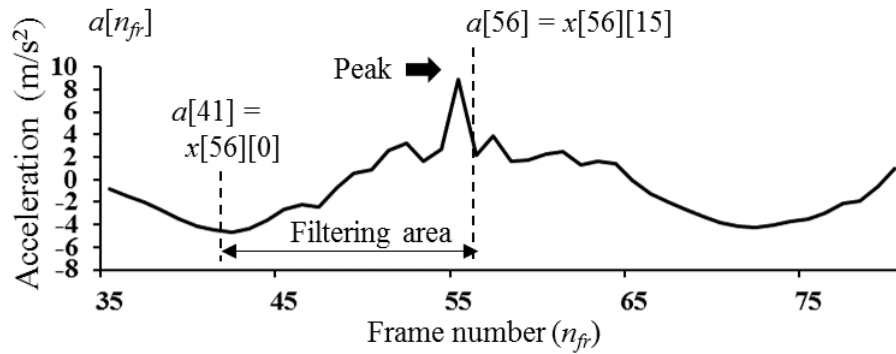


(a)

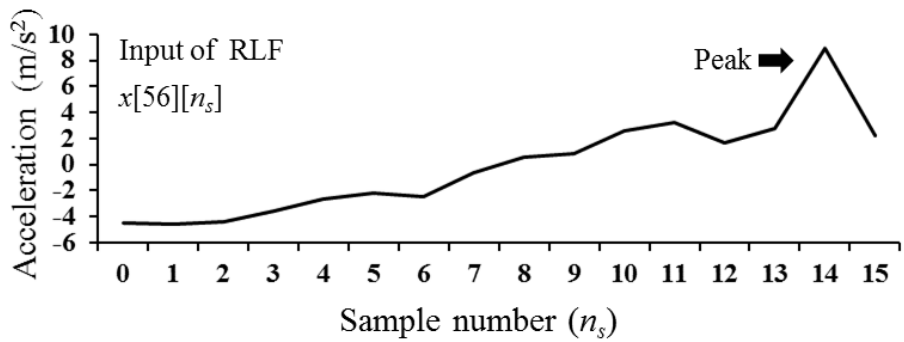


(b)

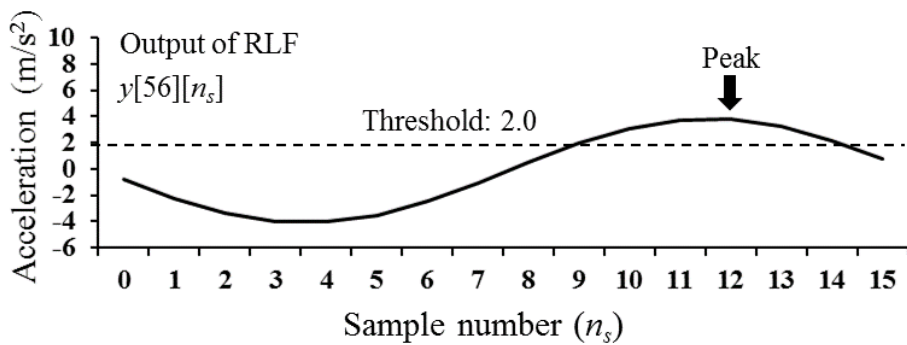
Figure 2.5. Block diagram of (a) the overall system with an IMU sensor module, and (b) the real time low pass filter at one frame ( $x[n_{fr}][n_s]; n_s = 0-15$ ), which implements FFT and IFFT with 16 samples.



(a)



(b)



(c)

Figure 2.6. Head vertical acceleration,  $a[n_{fr}]$ , is shown with (a) 46 frames ( $n_{fr} = 35-80$ ) near an HS and depicted with the filtering area at Frame 56 ( $n_{fr} = 41-56$ ; frame rate: 60 Hz). With the 16 samples, (b) the input of the real time low pass filter (RLF),  $x[56][n_s]$ , and (c) the output of RLF,  $y[56][n_s]$ , are depicted where  $n_s = 0-15$ .

#### 2.4.4. Thresholding

In data processing, thresholding is usually used to eliminate unwanted data. A threshold value can be decided to determine whether the signal is of interest or not. In the proposed H-IMU system, thresholding was implemented after RLF to recognize the peak at HS. In Figure 2.7, a growing peak is depicted over time, and it crosses the threshold value at frame  $n_{fr} = 54$ , which is one frame earlier than the peak of the head acceleration ( $n_{fr} = 55$ ; see Figure 2.6(a)); however, it is controllable to report a HS at the same frame of the peak of foot acceleration ( $n_{fr} = 54$ ; see Figure 2.3). The level of threshold values can also be altered depending on gait speed because faster gait speed might result in higher overall acceleration including both meaningful and noisy peaks. It is more effective to apply a thresholding method using the head acceleration than both the foot and pelvis. This is because outstanding peaks on head acceleration appear only at HS, whereas several high peaks on foot and pelvis acceleration appear, causing a confusion in detecting the HS exactly (see Figure 2.3). Therefore, threshold values can be easily found with head acceleration, which can minimize the confusion and improve accuracy in detecting HS.

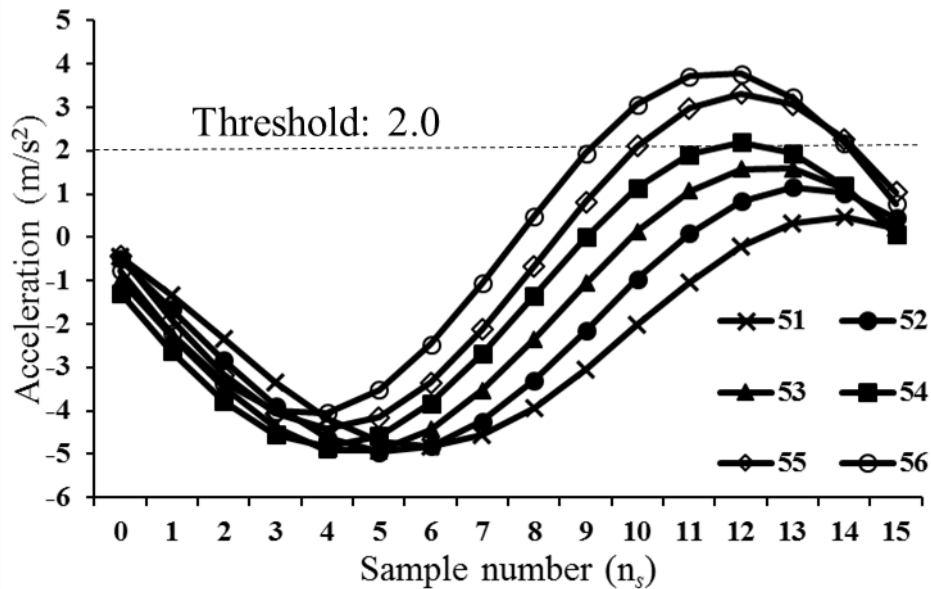


Figure 2.7. Waveforms of the head acceleration after real time low pass filter (RLF) between frames  $n_{fr} = 51-56$ , demonstrating that the peak is over the threshold value ( $2.0 \text{ m/s}^2$ ) at frame  $n_{fr} = 54$ .

## **2.5. Gait Parameters with H-IMU**

Gait parameters include spatial and temporal (also called spatiotemporal) gait parameters, angles of joints, and pressure of the foot. As a simple solution, a single H-IMU can provide spatiotemporal gait parameters, such as stride length, stride time, cadence, velocity, and CT. These parameters can be computed by a certain period between gait events (HS or TO). Time accuracy of gait event detection is, therefore, strongly related to the accuracy of the gait parameters. The exact time point of HS can provide step time, stride time, and the number of steps. Foot-ground contact time needs also the time point of TO. Spatial gait parameters can be calculated by displacement of the head at HS and TO.

### **2.5.1. Temporal Gait Parameters**

Temporal gait parameters are related to the gait velocity (m/s) and the cadence (steps/min). Step time and stride time are also temporal gait parameters. These parameters are obtained by the timing of gait events. The H-IMU can detect the time point of two gait events, such as HS and TO, which can provide temporal gait parameters.

First of all, head vertical acceleration is collected from the IMU sensor module to detect the impact at HS as shown in Figure 2.5. A real time low pass filter is applied to the head acceleration. Peak detection and thresholding are implemented to recognize HS. In Figure 2.3, head's peak acceleration at HS has one frame delay (16.7 ms at 60 Hz), compared to foot's peak acceleration, which is the transmission time of the impact from the foot to the head at HS.

Secondly, a reliable TO detection is needed for computation of temporal gait parameters because it divides the stance phase (foot-ground contact time) and swing phase. As no previous research has investigated the use of IMU for real time TO detection, this dissertation provides new insights into the real time TO detection using H-IMU, which is related to the propulsion motion of gait. After HS peak, named Peak 1, other peaks appear when a foot is getting off the ground because the propulsion force pushes the body to move forward. The propulsion motion results from movements of joints: the ankles, metatarsophalangeal joints of toes, knees, and hips as shown in Figure 2.8. The angular



movements of the joints push the body forward and upward. Right after HS peak, three peaks appear on the vertical head acceleration at the foot flat (Peak 2), toe off (Peak 3), and the beginning of the stance phase (Peak 4; see Figure 2.8 and 2.9). At Peak 2, while one foot (the right foot in Figure 2.8) makes the foot flat, the ankle of the contralateral foot (the left foot in Figure 2.8) mainly makes plantarflexion, pushing down on the ground. This propulsion force moves up the heel, and pushes the body and head forward and upward. At Peak 3, plantarflexions of the left toes are made, where the toes push the ground backward and downward, and the reaction force pushes the body and head forward and upward again, which helps the mass of the body move forward. At Peak 4, finally, extensions of the right knee and hip are made to move the mass of the body forward over the right leg, pushing the body and head, which also causes the outstanding peak on the pelvic acceleration as shown in Figure 2.3. However, it is harder to identify the TO peak with the same method as an HS peak because the amplitude of the three peaks after HS are similar. For detection of TO, therefore, the number of peaks is counted as shown in Figure 2.9. The peak at HS is labeled as the first peak, the foot flat as the second peak, the TO as the third peak, and the knee extension as the fourth peak. In TO detection, this algorithm deals with other two

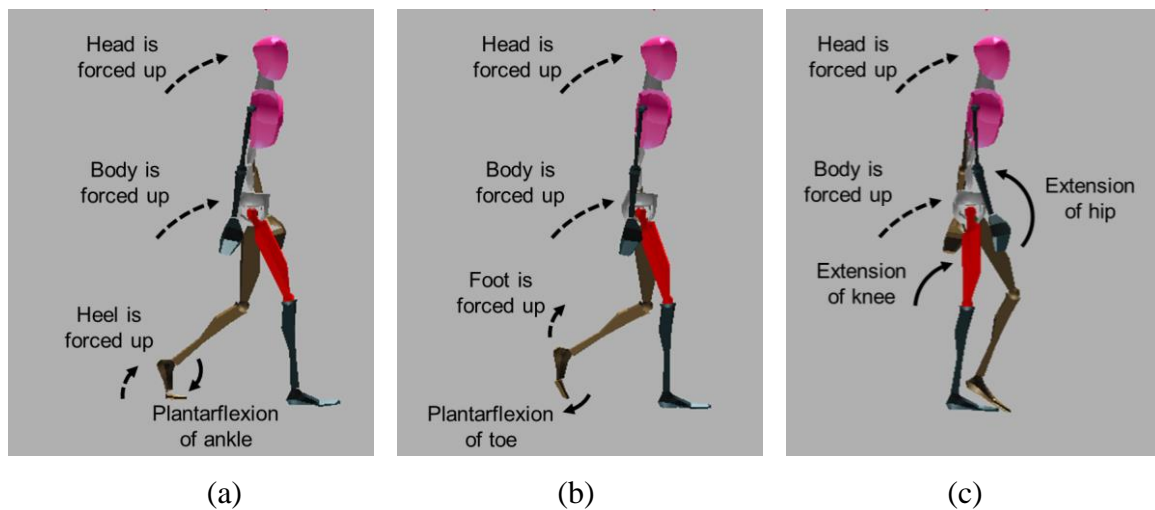


Figure 2.8. The sequence of propulsion motions during walking, which causes head vertical acceleration peaks (a) at foot flat (Peak 2), (b) toe off (Peak 3), and (c) the beginning of the stance phase (Peak 4).

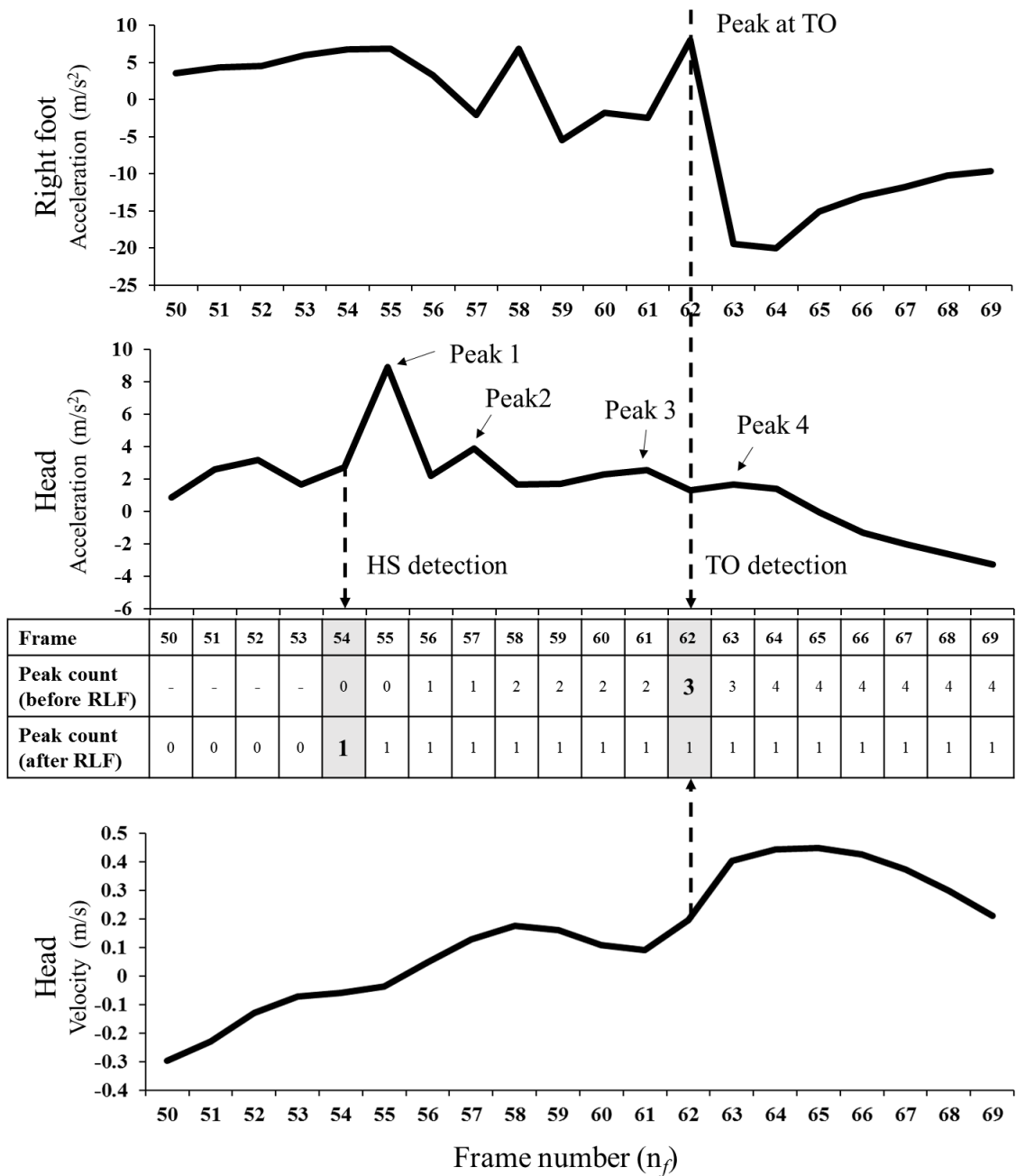


Figure 2.9. Acceleration of the right foot and the head (frames  $n_{fr} = 50-69$ ) and peak count data with before and after real-time low-pass filter (RLF). After HS detection, Peak 3 is labeled as TO. The velocity of the head is also depicted as another evidence for TO detection, which can prevent skipping TO detection.

different cases. First, one more peak can appear between the foot flat and TO, which normally happens in the five frames (83.3 ms) after HS. The additional peak is ignored, and the next one is recognized as a TO. Second, when the TO peak does not appear at right timing, the head velocity is analyzed as other evidence of TO. The negative peak of head velocity is monitored, as the time point when the head starts to move upward after the foot pushes off the ground, which prevents skipping and delaying TO detection. This algorithm detecting HS and TO is finally implemented with an analysis software. The pseudo code of the algorithm is described below (Hwang, Reh, Effenberg, and Blume, 2019b):

**Algorithm** Gait\_analysis\_using\_a\_head\_worn\_IMU is

**input:** z-axis acceleration at 60 Hz ( $a[n_{fr}][n_s]$ ):  $acc[16]$   
z-axis velocity at 60 Hz :  $vel[3]$   
FILTER = 1100 0000 0000 0011 /\* 0 to 15/16 \* 2 $\pi$  rad \*/

**while** true:

$acc$  = get acceleration data array ( $n_s$ : 0 to 15) at a frame ( $n_{fs}$ )  
 $acc\_fft$  = Fast Fourier Transform of  $acc$   
 $acc\_fft\_filtered$  =  $acc\_fft$  \* FILTER /\* real and imaginary numbers \*/  
 $acc\_filtered$  = Inverse Fast Fourier transform of  $acc\_fft\_filtered$   
call function **heel\_strike\_finder**( $acc\_filtered.Real$ ) /\* only real numbers \*/  
call function **toe\_off\_finder**( $acc, vel$ )

**function** heel\_strike\_finder( $acc\_filtered$ )

**if** a positive peak of  $acc\_filtered$  is detected **and** the peak of  $acc\_filtered$   $\geq 2$   
    **and**  $peak\_counter\_after\_RLF$  is 0,  
     $step\_counter = step\_counter + 1$   
    **print**  $step\_counter$   
     $peak\_counter\_after\_RLF = 1$   
     $peak\_counter\_before\_RLF = 0$   
    **print** timestamp for heel strike  
**else if** (a positive peak of  $acc\_filtered$  is not detected  
    **or** the peak of  $acc\_filtered$   $< 2$ ) **and**  $peak\_counter\_after\_RLF$  is 1,  
     $peak\_counter\_after\_RLF = 0$

```

function toe_off_finder(acc, vel)
  if a positive peak of acc is detected,
    peak_counter_before_RLF = peak_counter_before_RLF + 1
    if peak counter_before_RLF is 3,
      print timestamp for toe off
    else if peak counter_before_RLF >= 2 and a negative peak of vel is detected,
      print timestamp for toe off

```

In this algorithm, the filter in the frequency domain, FILTER, is 1100 0000 0000 0011. The first coefficient, FILTER[0], corresponds to the mean value of the input signal. Although FILTER is asymmetric, it yields the same results as a symmetric filter (e.g., [1, 1, 0.5, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0.5, 1]) because a convolution with FFT and IFFT is computed as a circular convolution. In other words, the filtering results in frequency domain are treated as a periodic function, so that negative and positive frequency components are added up. These filtered signals are used for HS detection, whereas TO detection uses the raw signals before filtering as shown in Figure 2.9.

Accurate temporal parameters are computed by accurate time points of gait events. The stance phase is defined from an HS of a foot to the next TO of the same foot, whereas the swing phase is from a TO of a foot to the next HS of the same foot again. The step time is defined from an HS of one foot to the following HS of the opposite foot. A stride time can be calculated from an HS of one foot to the following HS of the same foot. The cadence is obtained when counted step numbers which are identical to the number of HSs is divided by walking time in minute. The walking time is easily measured by a timer embedded in an IMU system.

### 2.5.2. Spatial Gait Parameters

Spatial gait parameters are related to the distance and position, such as step length (SL) and stride length. An IMU can estimate the displacement in 3-D space, which can generate a global and relative position of the body segments during walking. Related to the gait events, the gait parameters are obtained as shown in Figure 2.10. The step length and stride length are typically computed as a distance between foot prints. In Figure 2.10(a),

the sequence of steps is projected to the sagittal plane, and the distance is calculated between the HS positions of one foot and the other foot along with the walking vector ( $\mathbf{V}_d$ ). Although H-IMU provides head position instead of foot position, it can compute the average step length that is obtained when the walking distance is divided by the number of steps. Each step length is not correctly obtained by H-IMU; however, the average distance between two head positions at HS would be similar to the distance between heel positions, which is called pseudo step length (PSL; Hwang et al., 2018b) in this dissertation. The head is located between two feet, which divides the step length into two parts:  $SL_n.a$  and  $SL_n.b$  where the  $n$ -th step length is  $SL_n$  ( $n = 1, 2, 3, \dots$ ) as shown in Figure 2.10(b) and (c). A step length is defined as (2.1). The PSL is defined as the summation of the second part of an SL ( $SL_n.b$ ) and the first part of the next SL ( $SL_{n+1}.a$ ) as shown in Figure 2.10 and (2.2).

$$SL_n = SL_n.a + SL_n.b. \quad (2.1)$$

$$PSL_n = SL_n.b + SL_{n+1}.a. \quad (2.2)$$

Between the SL and PSL, an error should occur, which can be found from the differences of the average of SL and PSL as below:

$$\begin{aligned}
E[SL] &= \frac{1}{N} \sum_{k=1}^N SL_k \\
&= \frac{1}{N} (SL_1.a + SL_1.b + SL_2.a + \dots + SL_{N-1}.b + SL_N.a + SL_N.b) \\
&= \frac{1}{N} \left( SL_1.a + \sum_{k=1}^{N-1} PSL_k + SL_N.b \right) \\
&= \frac{1}{N-1} \sum_{k=1}^{N-1} PSL_k + \frac{1}{N} \left( SL_1.a + SL_N.b - \frac{1}{N-1} \sum_{k=1}^{N-1} PSL_k \right) \\
&= E[PSL] + \frac{1}{N} \{ (SL_1.a + SL_N.b) - E[PSL] \} \\
&= E[PSL] + \varepsilon_p
\end{aligned} \quad (2.3)$$



where  $E[*]$  is the expectation of  $*$  that computes the average value,  $N$  is the number of steps, and  $\varepsilon_P$  is the difference between  $E[SL]$  and  $E[PSL]$ . As  $N$  increases,  $\varepsilon_P$  decreases, finally reaches zero as shown (2.4).

$$\varepsilon_P = \lim_{N \rightarrow \infty} \frac{1}{N} \{(SL_1 \cdot a + SL_N \cdot b) - E[PSL]\} = 0. \quad (2.4)$$

Although  $N$  is not infinite, the error can be nearly zero when a user has a constant step length and all the first parts,  $SL_n \cdot a$ , are nearly the same. Imagine  $SL_{N+1}$  and  $PSL_N$  exist, then

$$SL_1 \cdot a + SL_N \cdot b \approx SL_{N+1} \cdot a + SL_N \cdot b = PSL_N \approx E[PSL]. \quad (2.5)$$

With regular gait, a single  $PSL_N$  and  $E[PSL]$  would be similar. The  $\varepsilon_P$  is nearly zero because

$$SL_1 \cdot a + SL_N \cdot b - E[PSL] \approx 0. \quad (2.6)$$

For another approach of the step length estimation, the distance between head horizontal positions at every vertical peak position of the head in the mid-stance phase is measured as demonstrated in Figure 2.10(c), called the estimated step length (ESL) in this dissertation (Hwang et al., 2018b). When the leg is straight and orthogonal to the ground, the head vertical position has the peak value. This time point is when the head position would be directly above a foot position in the side view on the sagittal plane as shown in Figure 2.10(c). Therefore, the distance between the head horizontal positions can be used for the estimation of the distance between every foot position which is the SL. The equation below shows the relationship between the  $n$ -th step length,  $SL_n$ , and the  $n$ -th estimated step length,  $ESL_n$  ( $n = 1, 2, 3, \dots$ ):

$$\begin{aligned} SL_n &= w_{foot.n+1} - w_{foot.n} \\ &= (w_{head.n+1} - \xi_{n+1}) - (w_{head.n} - \xi_n) \\ &= (w_{head.n+1} - w_{head.n}) + (\xi_n - \xi_{n+1}) = ESL_n + \varepsilon_n \end{aligned} \quad (2.7)$$

where  $w_{foot.n}$  is foot position on the walking vector ( $V_d$ ) at HS as shown in Figure 2.10(c),  $w_{head.n}$  is head position on  $V_d$  at the mid stance, and  $\xi_n$  is the error between each  $w_{foot.n}$  and

*W<sub>head,n</sub>*. The estimation error between  $SL_n$  and  $ESL_n$  is  $\varepsilon_n$ . The average of  $\varepsilon_n$  can be the representative estimation error  $\varepsilon_E$ , which is expressed by the average of errors of one foot,  $E[\xi_{2n}]$ , and the other foot,  $E[\xi_{2n-1}]$  as below:

$$\varepsilon_E = E[\varepsilon_n] = E[\xi_{2n-1}] - E[\xi_{2n}]. \quad (2.8)$$

When an H-IMU user has symmetric gait,  $\varepsilon_E$  would converge to zero. Otherwise,  $\varepsilon_E$  is not zero. When one's gait is biased a certain direction, the average error of left and right step lengths can be regularly different. From (2.7) and (2.8), two equations for the right and left step length can be defined as below:

$$SL_{2n-1} \approx ESL_{2n-1} + \varepsilon_E. \quad (2.9)$$

$$SL_{2n} \approx ESL_{2n} - \varepsilon_E \quad (2.10)$$

where  $SL_{2n-1}$  is the right step length and  $SL_{2n}$  is the left step length, which is measured by foot-worn IMU. The right and left step lengths estimated by head position are defined as  $ESL_{2n-1}$  and  $ESL_{2n}$ . With difference between odd and even values of  $ESL_n$  and  $\varepsilon_E$ , it can be found which is the right step or the left step. The estimated step length error,  $\varepsilon_E$ , also can be used as an indicator of symmetric gait.

The stride length is normally defined as the distance between a foot position at an HS and the same foot at the next HS. For example, the measurement of the left stride length starts at the left HS and ends at the next left HS, which is the summation of right SL and left SL. The right stride length is measured in the opposite order. Twice of the average step length is approximately the average stride length. A stride length can be estimated by head position by referring to the sum of the right and left ESL.

In terms of calculating the travel distance, two methods are utilized. The first method is to integrate acceleration twice, which can provide the total travel distance at every sample point. The second method is to sum the step lengths, which updates the travel distance at every HS. Between HS points, the former method records horizontal oscillation of head movement, whereas the latter one considers the straight distance. Although a



Kalman filter of the IMU (XSENS MVN System) compensated for the accumulated error in positioning, both methods would contain errors resulting from the unexpected head movement during walking. They need compensation by projecting the trajectory onto the walking vector.

## 2.6. Results

Measurement of gait was implemented using IMUs (MVN MTw) and an MVN Awinda station (router) for wireless communication (base frequency: 2.4 GHz), which is a part of the MVN Awinda system. The whole package of MVN XSENS Biomech includes a hardware and software solution for whole body motion capture. Seventeen wireless IMUs measure the whole body kinematics, such as local and global parameters of acceleration, velocity, position, and orientation of IMUs. Through Awinda station, data are transmitted to the host computer. The sensor movement data are reaching the computer after about 30 ms, and are immediately processed by the software so that the real-time motion capture in 3-D space can be realized with this system. With the parameters, gait analysis was implemented by using Python 2.7. Head kinematic data in this study refer to the head sensor of this motion capture system. Data update rate is 60 Hz for the analysis. Gait patterns of seven participants were analyzed in total; however, data of a male participant (M4) who walked without a pedometer was eliminated in the step counting analysis (Table 2.2).

Table 2.2. Comparison of Three Methods in Step Counting

Participant (Age; year)	Manual count	Pedometer		H-IMU	
		Steps	SCER (%)	Steps	SCER (%)
M1 (30)	1,105	1,117	1.09	1,100	-0.45
M2 (24)	1,866	1,851	-0.80	1,857	-0.48
M3 (34)	1,799	1,801	0.11	1,796	-0.17
F1 (30)	2,699	2,697	-0.07	2,697	-0.07
F2 (35)	1,221	1,223	0.16	1,221	-0.00
F3 (28)	1,868	1,866	-0.11	1,862	-0.32
Total	10,558	10,555	-0.03	10,533	-0.24
Mean absolute error (MAE)		35	0.33	25	0.24

### 2.6.1. Step Counting

The number of HSs was counted for step counting by the H-IMU. Kinematic data of physiological gait patterns were collected with six participants (male: 3; female: 3; age:  $30.2 \pm 3.7$  years; height:  $174.3 \pm 9.0$  cm). For the comparison with the H-IMU, a commercial pedometer fixed onto the waist was used, and steps were manually counted for ground truth. In this work, a trial is defined as walking a lap of the third lane of a whole 400 m track and another 25 m. Participants were asked to walk four trials with constant speed; however, some trials provided invalid data because sensors on other body parts were out of the initial places or dropped down, especially on the tibia or thigh, affecting head kinematic data. The results of step counting with manual counting, a pedometer, and an H-IMU are compared in Table 2.1. As the ground truth, total 10,558 steps were manually counted referring to 3-D motion capture recordings. The pedometer and H-IMU counted 10,555 and 10,533 steps, respectively. The step counting error ratio (SCER; Brajdic and Harle, 2013) is calculated with estimated steps ( $c_{est}$ ) and manually counted ground truth ( $c_{gt}$ ) as below:

$$SCER (\%) = \frac{c_{est} - c_{gt}}{c_{gt}} \times 100 (\%). \quad (2.11)$$

The SCER of both devices are very small, resulting in -0.33% for the pedometer and -0.24% for the H-IMU, where the constant error of the step counting is applied. However, SCER with the constant error is unable to represent the accuracy of a device because it might include over-counting and skip-counting. In terms of mean absolute error (MAE) for all participants, the SCER of the H-IMU was 0.24% (25 steps), whereas that of the pedometer was higher, 0.33% (35 steps). This result suggested that the H-IMU was normally more accurate than the pedometer in step counting. In Table 2.2, the first female participant (F1) walked 2,699 steps in four trials. Both pedometer and H-IMU counted 2,697 steps (-0.07% SCER), which is the most accurate results over all participants. In male participants, M3 who walked 1,799 steps in three trials provided the most accurate results with both the pedometer (1,801 steps, 0.11% SCER) and H-IMU (1,796 steps, -0.17% SCER).

Table 2.3. Spatiotemporal Parameters Computed by H-IMU

Participant (Age; year)	Height (cm)	Valid steps (Total)	Cadence (steps/min)	CT <sup>a</sup> (ms) (Avg. ±Std.)	CTR <sup>a</sup> (%)	PSL <sup>a</sup> (mm) (Avg. ±Std.)	ESL <sup>a</sup> (mm) (Avg. ±Std.)	Number of trials	Total distance (m)	Distance error per trial <sup>b</sup> (m)
M1 (30)	190	893 (1,110)	112.9	645.6 ±53.5	60	763.7 ±60.1	768.5 ±105.1	2	853	-13.8
M2 (24)	179	1,836 (1,857)	124.4	579.8 ±26.4	60.1	729.5 ±44.3	727.9 ±110.2	3	1351.7	10.2
M3 (34)	177	832 (1,189)	117.4	616.6 ±34.0	60.3	754.4 ±42.6	754.2 ±67.3	2	896.7	-2.2
M4 (26)	184	1,162 (1,176)	109.5	667.6 ±28.5	60.9	715.2 ±57.5	713.6 ±54.8	2	839.2	-20.8
F1 (30)	162	2,669 (2,697)	117.6	618.1 ±20.1	60.6	658.4 ±25.7	655.2 ±32.7	4	1,767.1	1.4
F2 (35)	170	1,207 (1,221)	106.8	676.6 ±33.7	60.2	720.9 ±61.9	717.6 ±51.1	2	876.2	-2.3
F3 (28)	168	1,841 (1,862)	113	644.5 ±26.3	60.7	686.6 ±49.0	684.1 ±79.8	3	1,273.8	-15.7

a: CT= contact time; CTR= contact time ratio; PSL= pseudo step length; ESL= estimated step length

b: Distance error per trial is calculated based on distance differences from the inside length of the third lane on 400 m track plus 25 m (ground truth: 440.33 m).

### 2.6.2. Spatiotemporal Parameters

Table 2.2 shows temporal gait parameters of seven participants (age:  $29.6 \pm 3.7$  years; height:  $175.7 \pm 9.0$  cm). Total 10,454 steps were analyzed for temporal gait parameters. Data at the beginning and the end of walking were excluded for analysis because of relatively low accuracy in temporal parameters, which negatively affects the spatial parameters. In the case of participant M1 and M3, several hundred samples were excluded because one of the sensors slipped down too much, and motion capture recording is distorted, which makes it difficult to compare it with other participants' data. The highest cadence was 124.4 steps per minute, recorded with the participant M2. The shortest CT as  $579.8 \pm 26.4$  ms has been also recorded from M2. The lowest cadence was 106.8 steps per minute measured on F2; however, F2 showed the longest CT as  $676.6 \pm 33.7$  ms. The range of participants' cadence varied from 112.9 to 124.4 steps per minute. The highest CTR was 60.9% (M4), whereas the lowest CTR was 60.0% (M1).

Regarding spatial gait parameters, two estimation methods of step length were implemented: PSL and ESL. All participants' average PSLs were higher than their average ESL, except for M1. Compared to the PSL, the ESL mostly showed large standard deviations. The possible explanation is that ESL contains information of differences between left and right step length, whereas PSL focusses on the average step length.

For the measurement of the travel distance, participants walked along the third lane of the standard 400 m track. This lane allows the deviation from 415.33 m to 423.0 m because of the differences of the inner and outer line, according to the standard of international association of athletics federations (IAAF). Participants walked another 25 m for one trial, which means the walking distance can be 440.33–448 m for one trial, 880.66–896 m for two, 1,320.99–1344 m for three, and 1,761.32–1,792 m for four trials. If the measured distance lies out of these ranges, it would indicate a measurement error, such as the accumulated error of IMUs. Total distance in Table 2.2 was calculated by multiplying the total number of steps and the ESL. Although ESLs of participant M1 and M3 were obtained from the part of their trials, we used the ESLs to estimate the total travel distance during the whole trials because they walked in constant manner. The distance error per trial

of F1 is in the range of the ground truth as 1.4 m, the results of M3 and F2 lie approximately in the range of -2.2 m and -2.3 m, respectively. Other participants show less than 20.8 m of the distance error per trial.

### 2.6.3. System Validity

The spatial and temporal analysis were respectively validated. With data from Table 2.3, the spatial reliability was validated. Table 2.4 describes error rates of the PSL (P = 6), ESL (P = 6), and distance error (P = 7). The average of the SL was calculated as below:

$$SL = \frac{(Walking\ distance)}{(Manual\ count\ of\ steps)}. \quad (2.12)$$

$$Distance\ error\ (\%) = \frac{(Distnace\ error\ per\ trial)}{(Distance\ per\ trial)}. \quad (2.13)$$

For SL in Table 2.4, the walking distance results from the multiplication of the number of trials and the distance per trial (440.33 m). The number of trials and the distance error per

Table 2.4. Spatial Parameters in Different Methods and Conditions

Parameters	SL Avg.	PSL		ESL		Distance error (%)
		Avg. (m)	Error (%)	Avg. (m)	Error (%)	
M1 (30)	797.0	763.7	-4.18	768.5	-3.57	-3.13
M2 (24)	707.9	729.5	3.05	727.9	2.82	2.32
M3 (34)	740.7	754.4	1.85	754.2	1.83	-0.50
M4 (26)	-	715.2	-	713.6	-	-4.72
F1 (30)	652.6	658.4	0.89	655.2	0.40	0.32
F2 (35)	721.3	720.9	-0.05	717.6	-0.51	-0.52
F3 (28)	707.2	686.6	-2.91	684.1	-3.26	-3.57
Mean absolute percentage error (MAPE)			2.15			2.15

trial are shown in Table 2.3. For participant M4, the average of SL was not calculated because his steps were not counted. The mean absolute percentage error (MAPE) of ESL and PSL were similar, at 2.07% and 2.15%, respectively. For distance error, MAPE was also similar, at 2.15% (without M4: 1.73%).

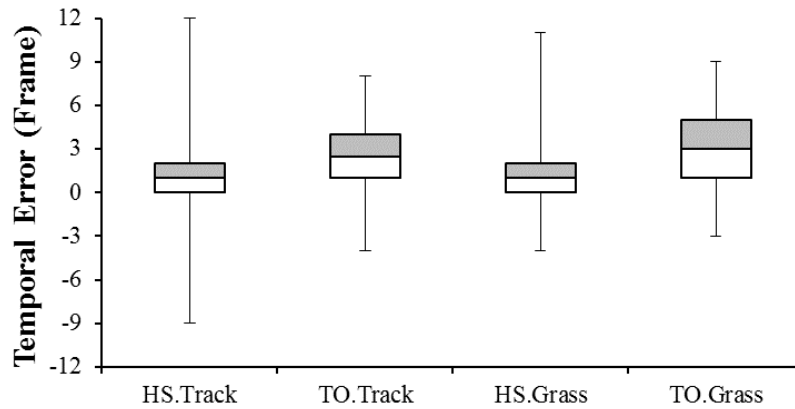
The conditions of the ground can influence the spatial accuracy. Thirteen participants (female: 6, male: 7; age:  $29.8 \pm 6.5$  years; height:  $174.2 \pm 8.7$  cm) were asked to walk 100 steps on the track (cadence:  $109.8 \pm 4.92$ ) and grass (cadence:  $108.0 \pm 5.29$ ). Two outliers were, however, removed in the dataset. Table 2.5 displays real numbers for the manual assessment and MAPE for PSL and ESL in each condition. For step counting, MAPE is lower on the track, at 1.09%, than the grass, at 1.37%. Both are higher than the results of the 6 participants in Table 2.2 because error occurs normally at the start and end points, so that when the number of steps increases, the error rate decreases. For the step length and total distance, MAPEs are also higher on the grass. In a comparison of PSL and ESL, MAPE of PSL is normally higher than ESL on the track and grass with exception of the step length on the track. Their MAPEs are normally less than data in Table 2.4 because spatial errors might be accumulated for a long pathway.

The temporal analysis was validated with gait patterns of the same 13 participants. In Figure 2.11 (Hwang et al., 2019b), Block and Whisker diagrams are depicted about the constant errors (including positive and negative differences between the measured and the reference data; c.f., absolute errors) in a comparison between H-IMU data and the foot-worn IMU data at each point of HS and TO on different ground floors. Figure 2.11(a) shows the constant errors of both HS and TO, whereas Figure 2.11(b) shows a comparison of CT. On the track, the median temporal errors of HS and TO are 1 frame and 2.5 frames, respectively. The 95% confidence interval of mean absolute errors (MAE-95CI) of HS and TO are  $12.32 \pm 1.34$  ms and  $43.77 \pm 1.94$  ms, respectively. On the grass, the temporal median errors are 1 frame for HS and 3 frames for TO. The MAE-95CI is  $26.43 \pm 1.66$  ms for HS and  $51.33 \pm 2.07$  ms for TO. Figure 2.11(b) displays the constant errors of CT on the track (median error: 1 frame; MAE-95I:  $31.23 \pm 2.37$  ms) and on the grass (median error: 2 frames; MAE-95I:  $41.23 \pm 2.68$  ms).

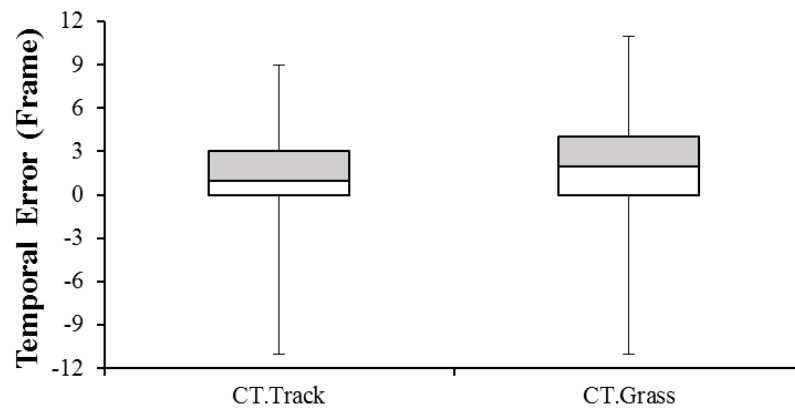
Table 2.5. MAPE of Spatial Parameters of PSL and ESL in Different Conditions

Parameters	Manual assessment	Track		Manual assessment	Grass	
		MAPE <sup>a</sup> (%)			MAPE <sup>a</sup> (%)	
		PSL	ESL		PSL	ESL
Steps	99.8 ± 4.7	1.09	1.09	98.5 ± 3.4	1.37	1.37
SL (cm)	807.1 ± 66.5	0.99	0.38	824.7 ± 60.5	1.14	1.40
Walking distance (m)	80.8 ± 5.7	0.71	1.42	81.3 ± 7.7	2.38	2.85

a: MAPE = mean absolute percentage error



(a)



(b)

Figure 2.11. Constant errors between H-IMU and foot worn IMU data (reference) with Block and Whisker diagrams in terms of (a) timestamp of gait events HS and TO as well as (b) CT (Hwang et al., 2019b).

In terms of SCER of the median error, the H-IMU system made 0.25% (P = 6 participants, N = 10,558 steps, MAE = 0.24%), whereas smartphones made less than 1.3 % (P = 27; Brajdic and Harle, 2013), which is shown in Table 2.6 (Hwang et al., 2019b). For CT (stance phase), 95% confidence interval (CI) of MAEs are compared with E-AR (Jarchi et al. 2014), which are  $31.2 \pm 2.37$  with H-IMU (P = 13, N = 1,316; track) and  $36.9 \pm 3.84$  ms with E-AR (see Table 2.6). Considering the validity of each ground truth system, the IMU-based 3-D motion capture system has the time delay of around 30 ms. In previous research, the time delays of IMU-based gait analysis were compared to opto-electric systems and foot switches, resulting in 40–50 ms online (Taborri, Palermo, Rossi, and Cappa, 2016) and under 10 ms offline (Kotiadis, Hermens, and Veltink, 2010). On the other hand, the force plate has the average time delay of 3 ms (< 8 ms; Taborri et al., 2016).

Table 2.6. Comparison of Single IMU Methods for Step Counting Accuracy

	<b>H-IMU</b>	<b>E-AR</b>	<b>Smartphones</b>
<b>Step count error SCER (%)</b>	0.25 (Median) 0.24 (MAE)	-	< 1.3 (Median)
<b><sup>a</sup>CT error <sup>b</sup>MAE-95CI (ms)</b>	$31.2 \pm 2.37$	$36.9 \pm 3.84$	-
<b>Speed</b>	Constant	Increasing	various
<b>Floor</b>	400 m track	Treadmill	various
<b>Ground truth</b>	3-D motion capture recording (XSENS)	Force plate in treadmill	Video recording

a: CT = foot-ground contact time

b: MAE-95CI = 95% confidence interval (95CI) of the mean absolute error (MAE)

The lower accuracy on the grass stems from the reduction of the impact from a foot to the head along the vertical direction (z-axis). The grass is softer than the track, thereby absorbing the impact, especially at HS. For the improvement of accuracy, the threshold values for the grass condition should be lower than the track condition because the



amplitude of the sinusoidal signal after RLF is lower. In addition, compared to the track, the grass is also uneven, and thereby the impact occurs in a more diagonal direction, so that less force is transferred in the vertical direction compared to walking on the track. Therefore, it is difficult to detect an HS with a single threshold value, so that the threshold values should be adaptable to more varied situations, referring to more parameters.

## **2.7. Discussion**

This Chapter described the measurement of individual gait patterns using a single IMU. Prior studies have noted the importance of gait analysis in monitoring one's physical health (Fritz and Lusardi, 2009; Studenski et al., 2011) and in preventing injuries (Bridenbaugh and Kressig, 2011; Howell et al., 2015), as well as in rehabilitation (Reh et al., 2016; Reh et al., 2019). An initial objective of this work was to reduce cost and to simplify the human gait measurement system. This measurement system enables gait rehabilitation with artificial auditory-motor feedback, which can enhance motor learning and relearning process of walking. Previous studies revealed that the real time auditory feedback is also applicable to gait rehabilitation of patients after hip or knee replacement surgery (Reh et al., 2016; Reh et al., 2019) and Parkinson's patients (Ghai, Ghai, Schmitz, and Effenberg, 2018). The real time auditory feedback can even support learning or relearning on healthy young people obviously based on multisensory integration effects (Effenberg, Fehse, Schmitz, Krueger and Mechling, 2016). In other words, this work initially aims at providing real time gait parameters with H-IMU, especially CT, which is measured by foot IMUs in previous study (Reh et al., 2019).

As a result, this study finally shows that foot-ground contact time (i.e., period between HS and TO) can be measured with an H-IMU in real time. Previous studies implementing a device with an IMU demonstrated that real time HS detection is possible, but not real time TO detection. Therefore, real time TO detection using H-IMU has been first reported in this work. As enabling HS and TO detection, H-IMU can measure most of the spatiotemporal gait parameters, whereas a device with only HS detection is restricted to counting steps and measuring step/stride time and length. These single IMU solutions are inherently unable to measure angular parameters, such as knee and hip angles.

Nevertheless, the proposed H-IMU solution can provide foot-ground contact time (the stance phase) and contact time ratio in real time, which is applicable to online rehabilitation methods using sensorimotor feedback. For instance, the proposed H-IMU solution with earbuds or an HMD can measure the foot-ground contact time and immediately provide users with auditory feedback of the parameters, enhancing the motor learning and relearning process. In addition, if H-IMU can distinguish the left and right step based on the accurate definition of gait phases, it can support gait balance analysis. These gait parameters are helpful for seniors' fall prevention, post-concussion management, and post-operational rehabilitation. Furthermore, H-IMU can be applied to group interpersonal coordination studies in marching scenario, and also dyadic interaction studies as shown in Chapter 3. These are applicable to walking interventions for various purposes.

### **2.7.1. Advantages of Using Head Kinematics in Real Time Gait Event Detection**

One of the advantages is that the head acceleration correlates to gait events and it is easy to detect the outstanding peaks appearing right after HS. The peak on head acceleration appears normally the same timing as on the pelvis because of head-trunk coordination in oscillation during walking (Cromwell and Wellmon, 2001; Kavanagh et al., 2005; Kavanagh et al., 2006). Interestingly, the peaks on head acceleration at HS is often faster than on pelvis, even though the head is further than the pelvis from the foot. A possible explanation is the principle of Newton's cradle, which is invented to explain the conservation of momentum. The cradle is a series of balls as rigid bodies hung on by strings, and when the first ball swings and hits the second ball, then the last ball swings out. The rest of balls stay at the same position, which indicates that momentum is transmitted through the balls; however, their kinematic changes are limited by each other. Compared to the series of pure rigid bodies, the human body structure can differentiate the kinematic properties because human body is a series of dampers, springs and rigid bodies in terms of biomechanical modeling. Nevertheless, the results in this study found that the measurement of head kinematics can be free from interference from other body parts than that of pelvic

kinematics, by showing the stable time lag between peaks on the head and feet at gait events.

When it comes to a question whether the head-feet time lag is acceptable to the real time solution, the results demonstrated that head-feet time delay of peaks was 10–55 ms at around 0.7 s of the step time (see section 2.6.3). A previous study also reported the head-feet time lag of oscillations is 30–200 ms at 1.0–5.0 s of step time (De Nunzio, Nardone, and Schieppati, 2005). According to recommendation from telecommunication standardization sector of international telecommunication union (ITU-T, R. G., 1997), these time delays can be regarded as “real time” solution in fields of telecommunication, which is 300 ms; however, these are longer than the threshold for audio and video (A/V) synchronization, which is 22 ms (Kudrle, Proulx, Carrieres, and Lopez, 2010). Furthermore, additional system delays, such as wireless communication delay (around 30 ms for XSENS system) and algorithm delay (16.7 ms for peak detection at 60 Hz of update rate), are already higher than the recommended delay for a real time solution of the A/V systems, which indicates that real time auditory-motor feedback for motor learning seems unavailable with H-IMU. However, embedding the H-IMU, head-worn wearable products, such as earbuds and HMDs, can measure the gait events and generate sound simultaneously without the wireless communication delays. In addition, the H-IMU using RLF and thresholding can detect HSs before HS peaks actually appear on head acceleration as shown in Figure 2.9. This early detection indicates that the HS is predictable with head acceleration because outstanding peaks on head acceleration appear normally once at HS, whereas other high peaks appear more times on foot and pelvic acceleration in a gait cycle. If accumulated individual gait data and artificial intelligence support the proposed algorithm with H-IMU, more accurate early detection is available in the near future. Thus, H-IMU can reduce not only head-foot delay at HS, but also the algorithm delay of peak detection.

In terms of another advantage in hardware system, the proposed H-IMU solution minimizes the calibration routines, compared to other wearables fixed onto the thigh, tibia or feet with bands. With earbuds or HMDs, the IMU can be fixed onto nearly the same

place on the ears or head. The H-IMU, consequently, has less risk of slipping down and moving from the initial place, which can reduce recalibration routines during use, and thereby provides seamless experience of the wearable devices.

### **2.7.2. Limitation of H-IMU in Gait Analysis**

Regarding the range of the gait parameters, joint angular parameters cannot be measured by using a single IMU because they need at least two sensors for one angle. Thus, H-IMU is not able to measure angles of knees and hips.

In this dissertation, the reported accuracy of the gait analysis is based on a relatively small number of participants (7 for step counting for long distance; 13 for the system validation), so there is still the need of validating the H-IMU solution with a larger sample of participants. To enhance the accuracy of the gait detection, larger sample size can be used to optimize threshold values and the real time low pass filter depending on personal gait patterns. For example, threshold values can be reduced to detect smaller peak values on the soft floors or while walking slowly. The low pass filter can set the higher cut-off frequency, so that raw information is not lost. Machine learning technology can detect if the users walk on the soft or hard floor, and faster or slower. The modification of detection algorithm can also refer to measurement history of different walkers and the optimal values can be found depending on the individualized gait patterns.

Another limitation of this work is that H-IMU was tested normally on the flat and hard floor. Although steps on the hard floor are compared to that on the grass-covered floor, a limited number of steps were analyzed. In addition, jogging and sprinting are not considered. Therefore, more test in different conditions are required in the future. Jogging and sprinting would be analyzed considering lofting time in the air and stronger impact during landing, which can be different points in data analysis. Gait analyses on various environments (e.g., inclines, stairs) and surfaces (e.g., the grass, sand), can be taken into account as well. The accuracy of peak detection would decrease because the acceleration peaks would be weaker and smoother, compared to walking on the hard floor.

### **2.7.3. Technical Issues of H-IMU in Gait Analysis**

One issue of H-IMU for gait analysis is the power consumption. For the zero gravity technology, the proposed H-IMU solution relies on not only accelerometers (using several hundred  $\mu\text{A}$ ), but also the gyroscopes that cause high power consumption (using several mA; Zhu, Anderson, and Wang, 2012). However, it was reported that gyroscopes embedded in smartphones can consume power levels similar to that of accelerometers, if the gyroscopes are not running as the fastest mode (Brajdic and Harle, 2013). In other words, H-IMU can reduce the gyroscope power consumption when it uses the same IMU modules as mobile applications and when it does not use the fastest mode in a special condition. The proposed algorithm with the H-IMU does not need the fastest mode, which indicates that the proposed method is applicable to consumer electronics market immediately. Although the optimization for smart devices was not dealt within this work, the power consumption of this H-IMU solution can be further optimized.

Frame drop is also one of the reasons for decreasing accuracy. Data are not always received at the right timing because of network congestion, which stems from limited network bandwidth or heavy data traffic. The network congestion causes queuing delay because a receiver system would wait for the missed data. In media data transmission, however, the receiver would drop the delayed frames because the regular interval is required for video or audio data streaming. Real time 3-D rendering animation in motion capture is also affected by frame drop, especially linear and angular positioning errors, which result from integration of sensor data. The single IMU solution has the advantage to reduce the data traffic, compared to the whole body system—17 sensors are used for XSENS MVN Awinda system. Nevertheless, if the data are transmitted on 2.4 GHz carrier frequency, the accuracy would decrease at places where high Wi-Fi or Bluetooth traffic occurs. In contrast, this issue would hardly occur, when the H-IMU is applied to earbuds or HMD.

## **2.8. Conclusion**

The proposed method with H-IMU provides a real time measurement solution for spatiotemporal gait parameters, such as foot-ground contact time, contact time ratio, and step length. This real time kinematic data are not only used for monitoring health and preventing injuries, but also for the enhancement of rehabilitation. Current health monitoring systems have issues, such as battery life, sustainable engagement for users. The proposed simple H-IMU can, nevertheless, improve sustainability of engagement for end-users, by guaranteeing a seamless experience, thereby encouraging users to walk. With IMU-based wearable applications, such as earbuds, smart glasses and HMDs, H-IMU method can easily reach the mass market of the health monitoring and home-based rehabilitation systems. In future, movement analysis with H-IMU also would cover not only running and jumping, but also swimming and skating. In addition, gait analysis with H-IMU is applicable to study on group interpersonal coordination during marching in military or a band. In a combination of head gesture recognition, H-IMU is applicable on dyadic interpersonal interactions in walking interventions, which will be discussed more comprehensively in Chapter 3.

## Chapter 3. Rapport Monitoring System During Walk and Talk

### 3.1. Introduction

When there is some distance between individuals in verbal or nonverbal communication, it is called distal interpersonal coordination. This type of interpersonal coordination is realized with physical distance between the interactants. For instance, there is a social distance in verbal conversation, gesture, and facial expressions. Distal interaction can be found manifold in sports games. Football or basketball players give a sign between players to pass each other. Auditory cues are, in addition, used in musical ensembles such as duets, quartets, and orchestras. To assess the quality of interaction, it is necessary to evaluate the degree of interpersonal coordination. The evaluation can be different depending on different cases. In sports, the number of completed passes can be an example to indicate the level of collaboration. In musical ensemble, aligned onset and keeping the tempo and rhythm during the performance can be evidences of well synchronized performance.

In daily communication, how many times and how long interactants nod, contact eyes, and imitate postures can be a standard of the measurement of interaction (Tickle-Degnen and Rosenthal, 1990). These factors indicate the level of rapport that is one's friendliness and warmth to each other, which is a familiar term in clinics, education and business, improving satisfaction and overall outcomes. One's ability to build good rapport supports better relationship in schools, occupations, and parties. The rapport, however, have been measured by participants' self-reports and external observers, which are usually subjective and unstructured. One of the greatest challenges is the proof of the research validity of their own methods because such approaches have limitations to compare results of each study. Therefore, for structured rapport measurement, three concepts were suggested to measure rapport: positivity, mutual attentiveness, and coordination (Tickle-Degnen and Rosenthal, 1990), which can be standardized in rapport measurement. When rapport is built between interactants, *positivity* is observed, which is one's agreement or approval to each other. *Mutual attentiveness* is established when individuals have a

cohesiveness and a unified feeling. Interactants also have *coordination* which stands for harmony, balance, and synchronization between individuals.

Another challenge is that the manual work for measuring the number and duration of behaviors is exhausting and time consuming. For reducing manual efforts, a combination of vision and machine learning technology has recently supported to detect nod and shake for agreement (Brunelli and Poggio, 1993; Rowley, Baluja, and Kanade, 1998; Kapoor and Picard, 2001). Motion capture has been used to measure participants' meaningful gesture or position.

In this chapter, the molecular measurement of rapport using H-IMU is demonstrated. The proposed H-IMU solution would enhance the efficiency in rapport measurement during walk and talk interventions and should lead to better outcomes in educational, clinical and social contexts. The related work is described in the next session, followed by methods and results. The discussion and conclusion appears in the last sections.

## **3.2. Related Work**

### **3.2.1. Molecular and Molar Measurement of the Three Rapport Components**

For rapport measurement, molecular and molar methods are described by Tickle-Degnen and Rosenthal (1990) (see also Baum, 2004). In *molecular methods*, specific behaviors are counted and timed during interacts. For instance, positivity is measured with the numbers of head nodding or smiling to one another, as well as judged with and durations of body orientation toward talking partners. In terms of mutual attentiveness, the number and duration of eye contacts and crossed arms are measured. Durations of posture and gesture mirroring also can be factors of coordination. These factors are easily conceptualized and immediately reported. On the other hand, *molar methods* more focus on contextual information to judge the development of rapport. Fake smiles, for example, cannot be counted as an evidence of good rapport. A seamless conversation between two individuals would be highly regarded as a signal of developed rapport, even though the individuals are crossing their own arms. For high level of rapport, therefore, individual imagination, impression or perception under social context should be measured rather than



counting the number of smiles, eye contacts, or crossed arms (Tickle-Degnen and Rosenthal, 1990). This measurement relies on the participants' self-reports and impressions of outside observers, and thereby the judgement can be subjective and time consuming. It is reported that the molecular method is more appropriate to assess positivity and mutual attentiveness, whereas the molar method is more suitable for the measurement of coordination. Both methods are, nevertheless, important to all three rapport components.

The importance of both methods varies over time during an interaction. The molecular method weighs more strongly in the initial stage of interaction when they interact with new acquaintances or strangers (Tickle-Degnen and Rosenthal, 1990) because individuals would behave politely and circumspectly with unfamiliar interactants under awkward conditions. They also tend to follow basic norms of social propriety, which leads to circumscribed and stereotypical behavior. These structured behaviors can be measured by molecular methods (Tickle-Degnen and Rosenthal, 1990). Later, interactions are loosely structured and show more diversity in the ways of communication. Under familiar conditions, interactants communicate by developing their own conventions rather than following the cultural social norms because it can enhance communication efficiency and reduce misunderstandings during conversation. However, the molecular method is difficult to analyze rapport in the later phase of interactions because the diverse types of stimuli and responses between interactants exist, and are even interwoven. The molar method, therefore, becomes more significant in later interaction. Nonetheless, the three components of rapport are still present regardless of time. Although the importance of the components is varying, the molecular methods can provide evidences to assess the rapport development. Thus, molecular and molar methods contribute to the measurement of the components.

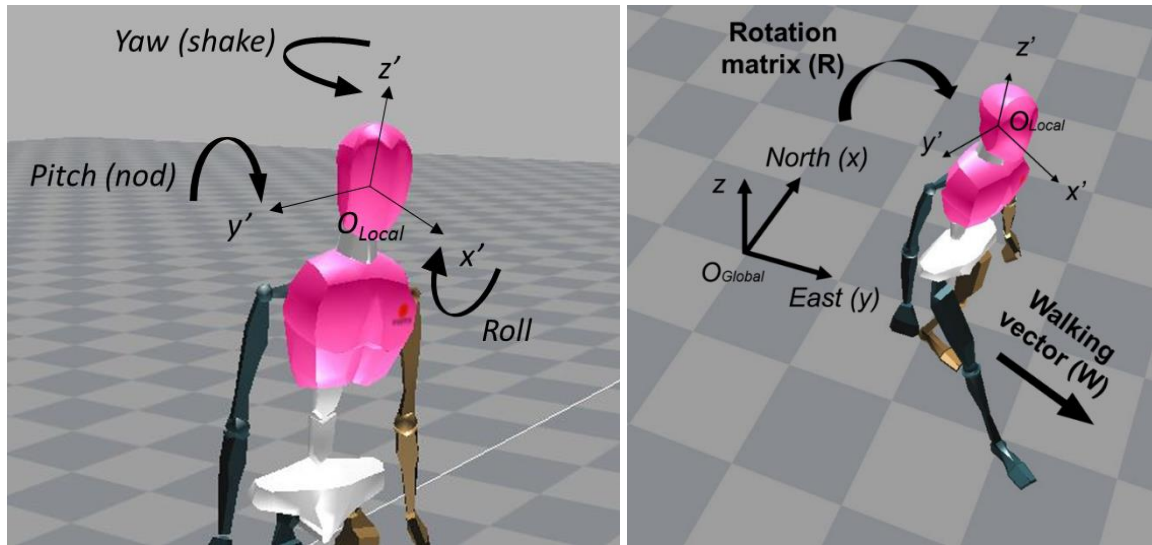
### **3.2.2. Head Pose Estimation**

Head pose estimation technology supports nod and shake detection as well as gaze estimation. Researchers have applied cameras, sensors and machine learning technology to estimate head posture and movement. First of all, face recognition systems were developed. Two main methods were realized: feature-based methods and template

matching methods (Brunelli and Poggio, 1993). Feature-based procedures are implemented with geometric information of the nose, mouth, and the eyebrows. For example, nose vertical position, mouth height and eyebrow thickness can be considered to compare the unknown face with a prepared database. Template matching is also realized with geometric information of the mouth, eyes, and eyebrows; however, normalized cross correlation between the unknown target and the database is directly computed with the area of the features. Secondly, previous research has suggested that rotation and posture of the face can enhance the recognition rate when the head is rotated (Maurer and von der Malsburg, 1996), and has established 3-dimensional (3-D) representation of the head by using a single camera (Chen, Wu, Fukumoto, and Yachida, 1998). The angular kinematic information in 3-D space contributes to nod and shake detection. In the case of nodding, the head rotates up and down. When shaking, the head rotates left and right. Finally, improvement of the recognition rate is studied by using machine learning technology, such as a neural network and a Hidden Markov Model (HMM) (Rowley et al., 1998). The rotated pose is also related to the head direction and gaze, which would be applicable in measuring interactants' mutual gaze and attentiveness.

### **3.2.3. Gait Event Detection**

The rapid development of microelectromechanical systems (MEMS) leads to the development of inertial sensors (e.g., accelerometers and gyroscopes). A combination of inertial and magnetic sensors improves accuracy and robustness. IMU-based motion capture systems are free from spatial limitations and applicable indoors and outdoors. Smartphones and wearable devices have been investigated for simple health monitoring systems in everyday life settings. Concerning health monitoring, the IMU-based solutions can provide information of trajectory, calorie consumptions, and the number of steps. As mentioned in Chapter 2, a head-worn IMU (H-IMU) is more informative than commercial pedometers because it can detect detailed gait events (e.g., heel strike and the toe off) and more gait parameters (e.g., foot-ground contact time) (Hwang et al., 2018b). The gait parameters can be utilized in behavioral studies on everyday-life walking settings.



(a)

(b)

Figure 3.1. Geometrical definitions: (a) the yaw (shake), pitch (nod) and roll of the head; (b) the walking vector, the head orientation, the rotation matrix, as well as the global and local coordinate system.

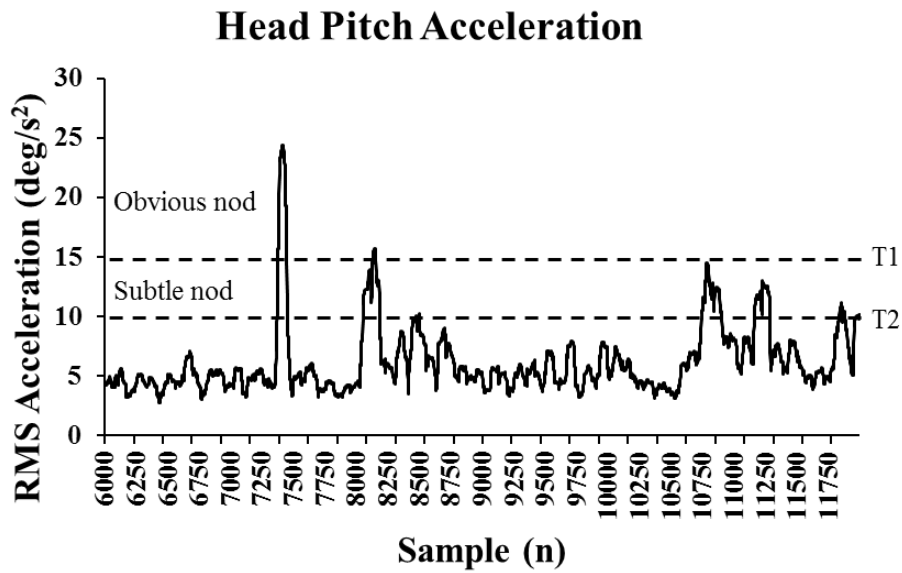


Figure 3.2. RMS value (window width: 1 second) of head pitch acceleration, and the threshold values for obvious (T1) and subtle (T2) nod detection.

### 3.3. Methods

#### 3.3.1. Head Nod Detection

From the notations of angular movement, the head pitch rotation is defined as nodding movement and the head yaw rotation is defined as shaking movement in Figure 3.1(a). The gyroscope in an IMU can estimate the rate of turn and angular acceleration of the head pitch rotation. When the head is nodding, a pitch torque is applied to the head, and the angular acceleration increases. In this dissertation, the nod detection method refers to the root mean square (RMS) of the pitch angular acceleration with 1 second window width. Adequate threshold values are decided to detect head nods. When the head nod is detected, the average value, duration, and maximum value are measured. Depending on RMS value, the head nod can be categorized into two classes: an obvious and a subtle nod (Chen et al., 1998). The higher threshold, T1, is for obvious nod detection, whereas the lower threshold, T2, is for subtle nod detection as shown in Figure 3.2. Another condition is the relationship with the yaw rotation; the head nod is recognized when pitch RMS values should be higher than the yaw RMS values, even though the pitch value is over T2, which is to avoid detecting other head gestures.

#### 3.3.2. Estimation of Mutual Head Orientation and Its Coordination

Gaze position is predictable from head orientation due to head-gaze coordination (Freedman and Sparks, 2000; Fang et al., 2015), which is confirmed especially when the head can naturally move like in daily conditions such as during walking (Nakashima, Fang, Hatori, Hiratani, Matsumiya, Kuriki, and Shioiri, 2015). From the gaze estimation, mutual gaze is also estimated from head orientation in walking settings with their conversation partners. The angle of the head orientation ( $x'$ ) from the walking vector ( $\mathbf{W}$ ) is measured as shown in Figure 3.1(b). The head orientation ( $x'$ ) is obtained by referring to the earth-fixed reference, such as directions of the north and the east measured from the magnetometer. The walking vector ( $\mathbf{W}$ ) is calculated from head displacement for 1 second. The calculation of the angles is performed in the 3-D Cartesian coordinate system. The global coordination (origin:  $\mathbf{O}_{Global}$ ) has the forward (north) as the positive direction of the

$x$ -axis, the right (west) side as the  $y$ -axis, and the upward as the  $z$ -axis (see Figure 3.1(b)). The  $x'$ -,  $y'$ -, and  $z'$ -axis in the local coordinate system (origin:  $\mathbf{O}_{Local}$ ) are defined by the rotation matrix ( $\mathbf{R}$ ) and the direction of original axes in the global coordination system as below:

$$\begin{aligned}
 \begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} &= \mathbf{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{R}_z \mathbf{R}_y \mathbf{R}_x \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
 \mathbf{R}_x(\theta_x) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta_x & -\sin \theta_x \\ 0 & \sin \theta_x & \cos \theta_x \end{pmatrix} \\
 \mathbf{R}_y(\theta_y) &= \begin{pmatrix} \cos \theta_y & 0 & \sin \theta_y \\ 0 & 1 & 0 \\ -\sin \theta_y & 0 & \cos \theta_y \end{pmatrix} \\
 \mathbf{R}_z(\theta_z) &= \begin{pmatrix} \cos \theta_z & -\sin \theta_z & 0 \\ \sin \theta_z & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix}
 \end{aligned} \tag{3.1}$$

where  $\mathbf{R}_x$ ,  $\mathbf{R}_y$  and  $\mathbf{R}_z$  are the basic rotation matrices for the rotations by  $x$ -,  $y$ -, and  $z$ -axis, respectively. The angle differences between two coordinate systems are expressed by Euler angles ( $\theta_x$ ,  $\theta_y$ ,  $\theta_z$ ). Figure 3.3 simply shows the expressions of the rotated axes ( $x'$ ,  $y'$ ,  $z'$ ) and the walking vector,  $\mathbf{W}'$ , in the new coordinate system, which is rotated by  $\theta_z$ .

When  $\mathbf{W}$  is expressed in the local coordinate system, the angle between  $x'$ -axis and the walking vector is obtained as  $\theta - \theta_z$ , which is the same as being inversely rotated by  $\theta_z$ . The local walking vector,  $\mathbf{W}'$ , is computed as below:

$$\begin{aligned}
 \mathbf{R}_z^{-1} \mathbf{W} &= \mathbf{R}_z(-\theta_z) \mathbf{W} = \begin{pmatrix} \cos \theta_z & -\sin(-\theta_z) & 0 \\ \sin(-\theta_z) & \cos \theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos \theta \\ \sin \theta \\ 0 \end{pmatrix} \\
 &= \begin{pmatrix} \cos \theta \cos \theta_z + \sin \theta \sin \theta_z \\ \sin \theta \cos \theta_z - \cos \theta \sin \theta_z \\ 0 \end{pmatrix} = \begin{pmatrix} \cos(\theta - \theta_z) \\ \sin(\theta - \theta_z) \\ 0 \end{pmatrix} = \mathbf{W}'.
 \end{aligned} \tag{3.2}$$

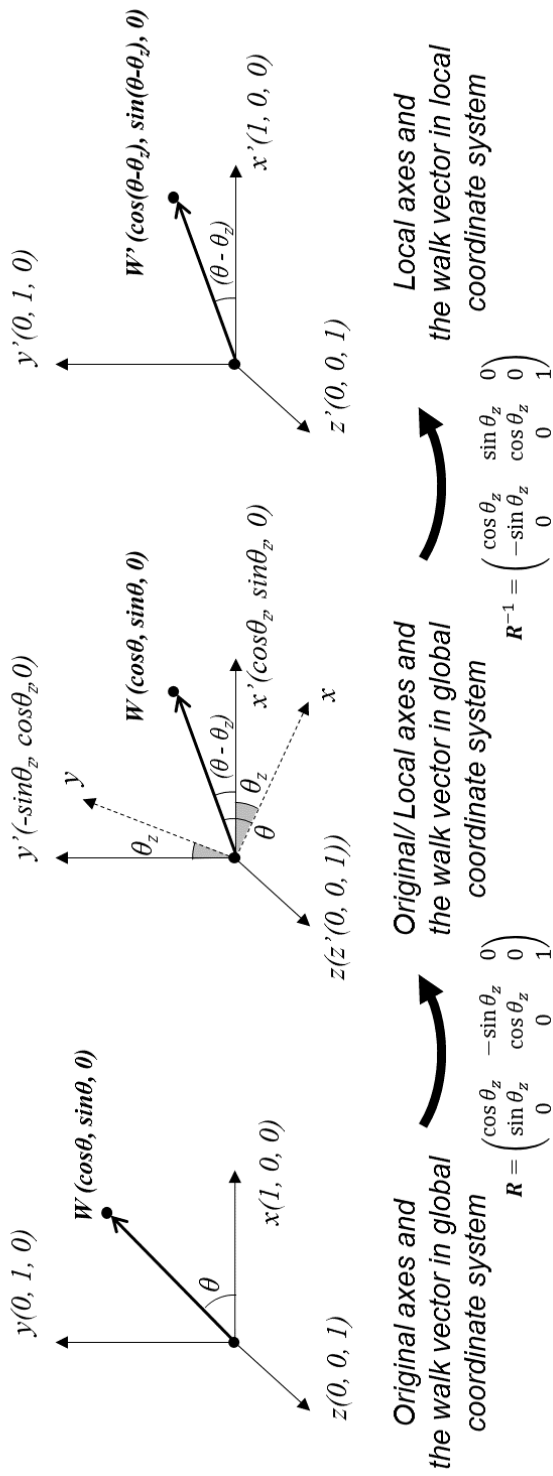


Figure 3.3. Example diagram of the expressions of the rotated local axes ( $x'$ ,  $y'$ ,  $z'$ ) by  $\theta_z$  in the global coordinate system, and the new expressions of the axes and the walking vector,  $W'$ , in the local coordinate system

The inverse rotation matrix ( $\mathbf{R}_z^{-1}$ ) results in  $\mathbf{W}'$ . With rotation by  $\theta_x$ ,  $\theta_y$ , and  $\theta_z$ , the general equation is obtained using  $\mathbf{R}$  and  $\mathbf{R}^{-1}$  as below:

$$\mathbf{R}^{-1}\mathbf{W}_n = \mathbf{W}'_n \quad (3.3)$$

where  $\mathbf{W}_n$  and  $\mathbf{W}'_n$  are walking vectors in the global and local coordinate system, respectively, and  $n$  is the index of participants. When two participants ( $\mathbf{P}_1$ ,  $\mathbf{P}_2$ ) walk together, there are two walking vectors ( $\mathbf{W}'_1$ ,  $\mathbf{W}'_2$ ). For analysis of turning left, right, and forward, horizontal angles ( $\theta_z$ ) of participants are considered on x-y plane, which will be expressed as  $\theta_1$  and  $\theta_2$  (see Figure 3.4). When the head turns right, the angle has a positive value like  $\theta_1$ , and when the head turns left, the angle has a negative value like  $\theta_2$ .

Several scenarios are anticipated when  $\theta_1$  and  $\theta_2$  are ranged  $\pm 90^\circ$  (from turning  $90^\circ$  right to left). First, when two partners look at one another, the angle difference ( $\theta_d = \theta_1 - \theta_2$ ) can be  $180^\circ$  (see Figure 3.5), which is the maximum case. When both participants look forward,  $\theta_d$  is  $0^\circ$ . During walking and talking together,  $\theta_d$  is usually between  $180^\circ$  and  $0^\circ$  because participants try to rotate their head toward their partner. The minimum case is that  $\theta_d$  is  $-180^\circ$  when both participants look at exactly opposite directions. The angle difference varies between  $-180^\circ$  and  $180^\circ$ , which is rescaled from -1 to 1 to generate a simple indicator.

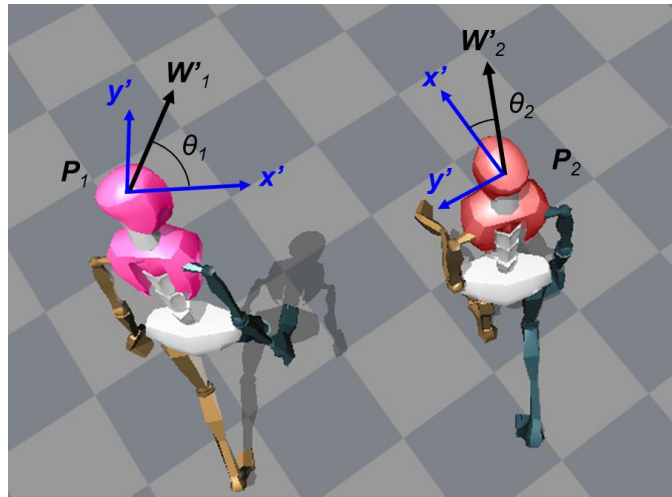


Figure 3.4. Examples of walking vector, head direction and angles. In global Cartesian coordination, the origin and direction of participants are demonstrated.

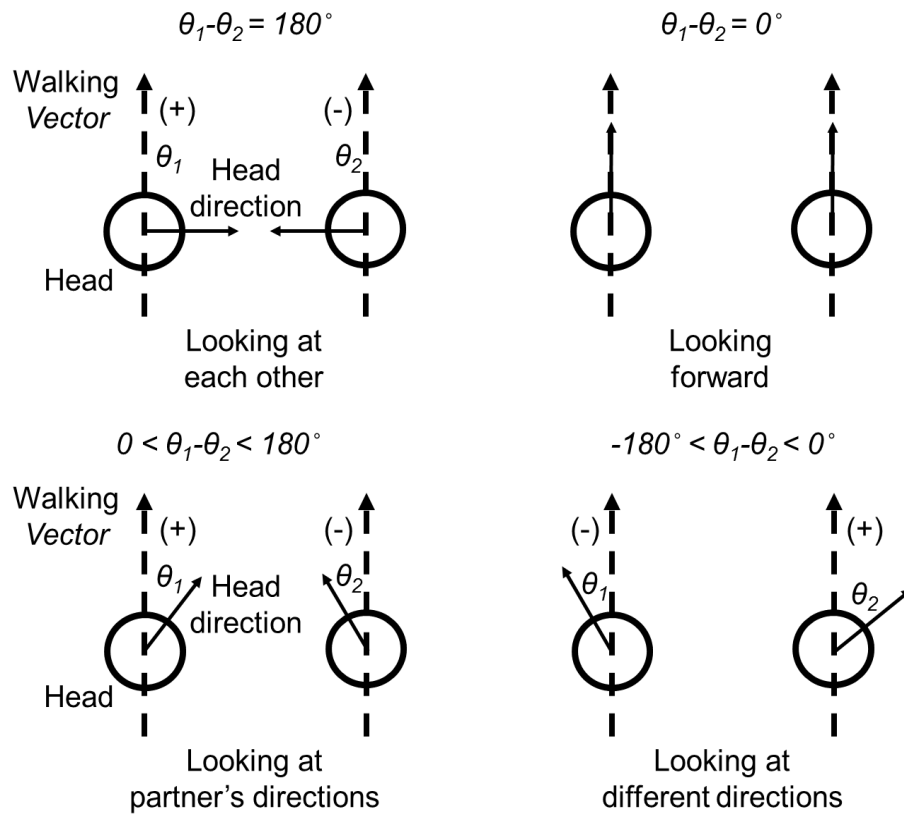


Figure 3.5. The four different scenarios of walking with conversation partners, and angle differences of two walkers.

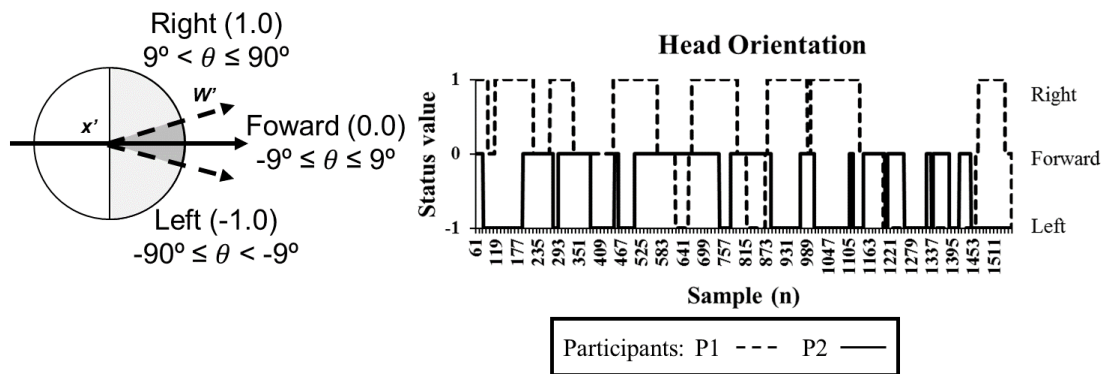


Figure 3.6. Individual head orientation status with three statuses; looking left (-1.0), forward (0.0), and right (1.0). P1 and P2 are participants walk on the left and right, respectively.



For the level of gaze coordination, head orientation is divided into three scenarios: Looking right is regarded as the range of the horizontal head angle,  $\theta$ , from  $90^\circ$  to  $9^\circ$ , forward is  $\pm 9^\circ$ , and left is from  $-9^\circ$  to  $-90^\circ$  as shown in Figure 3.6. Looking right, forward, and left have the orientation index ( $OI$ ), at 1, 0 and -1, respectively. Then, gaze correlation is also evaluated as below:

$$f(OI_1, OI_2) = \begin{cases} 0.5 & (OI_1 = 0 \text{ and } OI_2 = 0) \\ 1.0 - \frac{1}{2}|OI_1 - OI_2| & (OI_1 \neq 0 \text{ or } OI_2 \neq 0) \end{cases} \quad (3.4)$$

This is the evaluation function of the orientation correlation where  $OI_1$  and  $OI_2$  are the orientation indexes of P1 and P2. When two participants look in the same direction, it has the highest value 1.0. This case is regarded as emergence of the high level of rapport because they look at the same place or object that might attract their common interest. In contrast, when they look in different directions, it is 0. When both participants look forward, the gaze correlation is 0.5 as a neutral value because looking forward is the head's normal posture in natural gait. The average of correlation values or the duration of the highest correlation can automatically indicate the level of rapport. If outside examiners are involved in the rapport measurement, they can easily find the meaningful period for their judgement. The evaluation results of the gaze correlation are shown in Figure 3.8(b).

### 3.3.3. Gait Coordination

As mentioned in Chapter 2, a single head-worn IMU (H-IMU) is applicable to the detection of gait events (Hwang et al, 2018b). The detection method includes the real-time low-pass filter, peak detection, and thresholding. With these data analysis methods, the system can detect the heel strike (HS) and the toe off (TO) in real time (Hwang et al, 2018b). To evaluate gait coordination, a binary signal is generated with the double support period, which is defined as the period when both feet support body, exactly between HS of one foot and TO of the other foot. During the double support period, the binary value is 1, otherwise it is 0. Figure 3.7 demonstrates the binary signals from both participants (P1, P2). For the results of gait coordination, the Pearson correlation between two binary signals is

computed during the walking period of interest. For real-time monitoring, the correlation coefficients are calculated as the status value at every frame with 10 seconds window, which is displayed in Figure 3.9 in the results section.

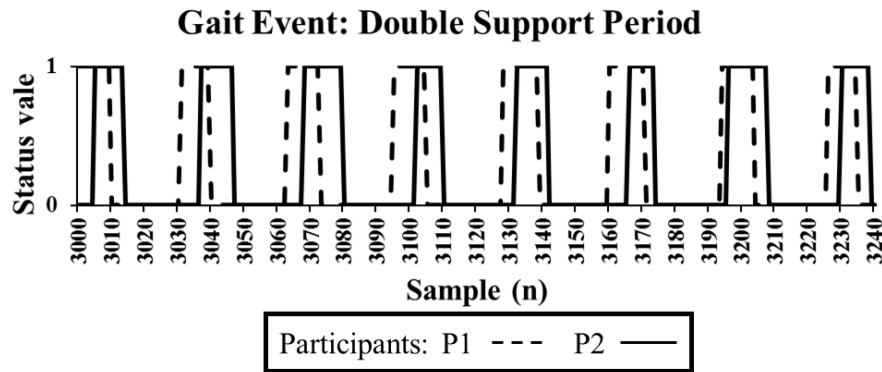


Figure 3.7. The binary signals of two participants during walk. The status value is 1 between HS and TO when two legs are in contact to the ground, called the double support period. After TO and before HS, one leg supports the body, so that the status value is 0.

### 3.4. Results

Two healthy male (age:  $26 \pm 0.5$  years, height:  $184.5 \pm 0.5$  cm) participated in the experiment. They wore a suit with 17 IMUs from 3-D motion capture system (XSSENS MVN Awinda system). Each IMU includes a 3-axis accelerometer, 3-axis gyroscope, and 3-axis magnetometer. On the standard 400-m track, participants walked one lap alone each, and one lap by talking together. Head kinematic data were collected from the head IMUs. A software solution (MVN Studio) rendered their gait in 3-D space and organized data. From the database, the all real time detection and analysis were performed by Python 2.7. Participants' conversation is recorded by a digital voice recorder (Olympus WS-852).

#### 3.4.1. Head Nod Detection

In Table 3.1, the results of nod recognition using H-IMU is demonstrated. The recognition rate is calculated in a comparison to a reference that is manually counted referring to the voice recording and the motion capture animation. In the manual count, the clear nods were counted when one responds to the partner. The nod events frequently

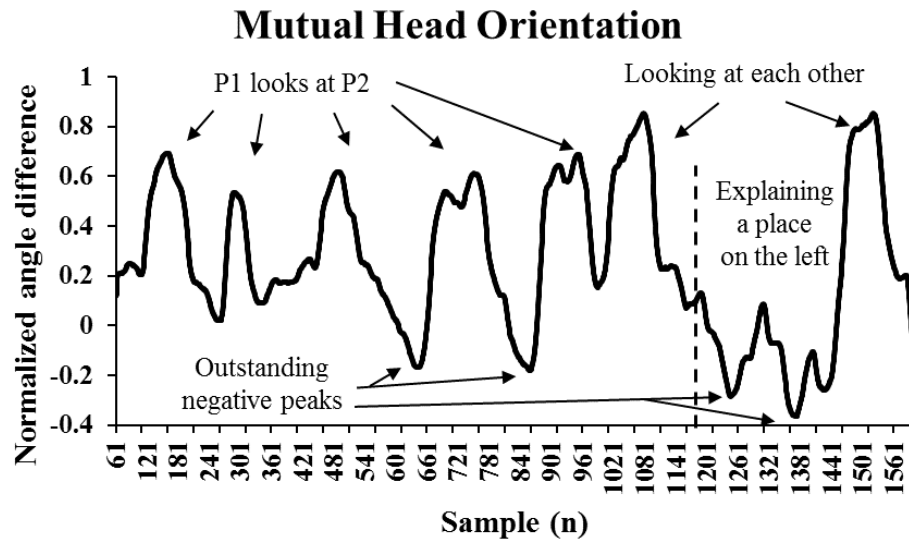
included linguistic responses, such as “yes,” “OK,” “good,” and “nice” as well as other positive words or repetitions to partner’s expressions. In total, the nod recognition rate was 94.34%, regarding the analysis with 3-D motion capture recordings as ground truth. In the walking setting, P1 nodded 16 times and 15 nods was recognized, whereas P2 nodded 37 times and 35 nods were recognized. For P1, 10 subtle nods were detected, which is twice as much as his obvious nods. In terms of P2, 27 subtle nods were counted, which is more than threefold his obvious nods.

Table 3.1. The Nod Recognition Using H-IMU

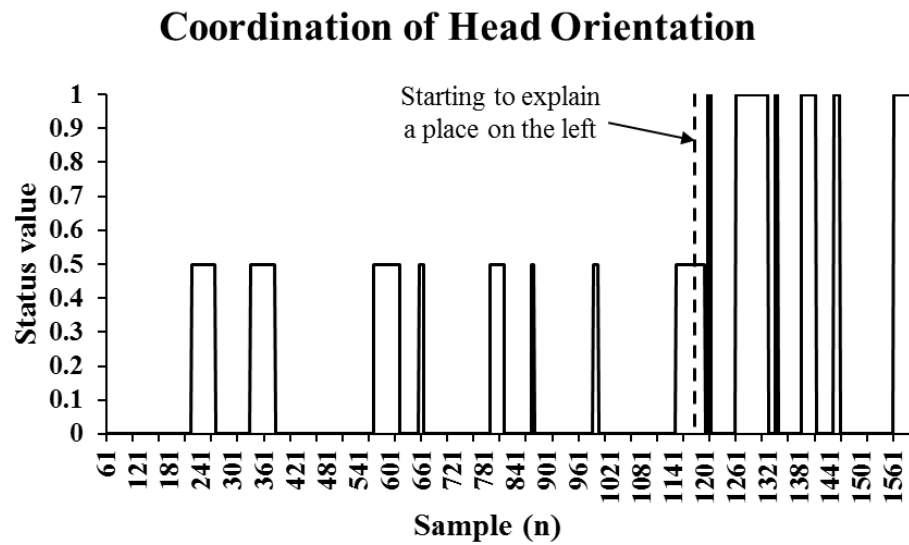
<i>Participant (age, height)</i>	<i>H-IMU</i>			<i>Manual count</i>	<i>Recognition rate (%)</i>
	<i>T1</i>	<i>T2</i>	<i>Total</i>		
P1 (27, 184 cm)	5	10	15	16	93.8
P2 (25, 185 cm)	8	27	35	37	94.6
Total	13	37	50	53	94.3

### 3.4.2. Mutual Head Orientation and Coordination of Head Orientation

The mutual head orientation resulted from the angle differences between both participants’ head direction was monitored in real time. Figure 3.8(a) demonstrated the mutual head orientation at the beginning of the walking for around 1,540 sampled data (25.7 s at 60 Hz). The mutual orientation is fluctuated over time because they repeatedly look at the partner, forward, and other directions. From the analysis of the motion capture recording, P2 speaks and rarely looks at P1, whereas P1 listens to P2 and frequently looks at P2. At the positive peaks before  $n = 1,021$ , the values from 0.5 to 0.8 correlate to the moments when P1 looks at P2, as evidenced by the motion capture recording and the data shown in Figure 3.6. The amplitude is higher than 0.5, which means P2’s head is slightly oriented to P1. At  $n = 1050$  and  $1520$ , they look at their partners together, so that the value peaks higher than 0.8. At the last two peaks, coordination evaluated as higher, which means both participants look at the same direction, whereas, at the former peaks, they look to different directions. In contrast, negative peaks are observed around at  $n = 663, 878, 1265,$  and  $1394$ . The negative peaks occur when P1 looks at the opposite direction (left) of P2 (right) as shown in Figure 3.8(b), which depicts the coordination of head orientation.



(a)



(b)

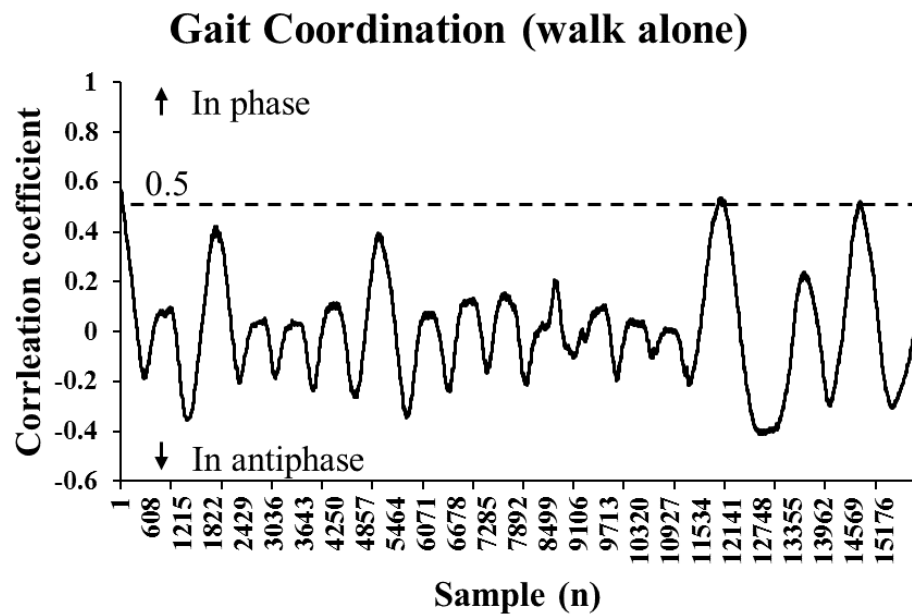
Figure 3.8. Evaluation results of mutual head orientation and coordination of head orientation at the beginning 1,540 samples at a 60 Hz. At (a), high values stand for turn their head more toward their partner. At (b), there are three status values; 1: interactants look in the same direction; 0.5: look forward; and 0.0: look in the different directions.

### 3.4.3. Correlation of Gait Patterns

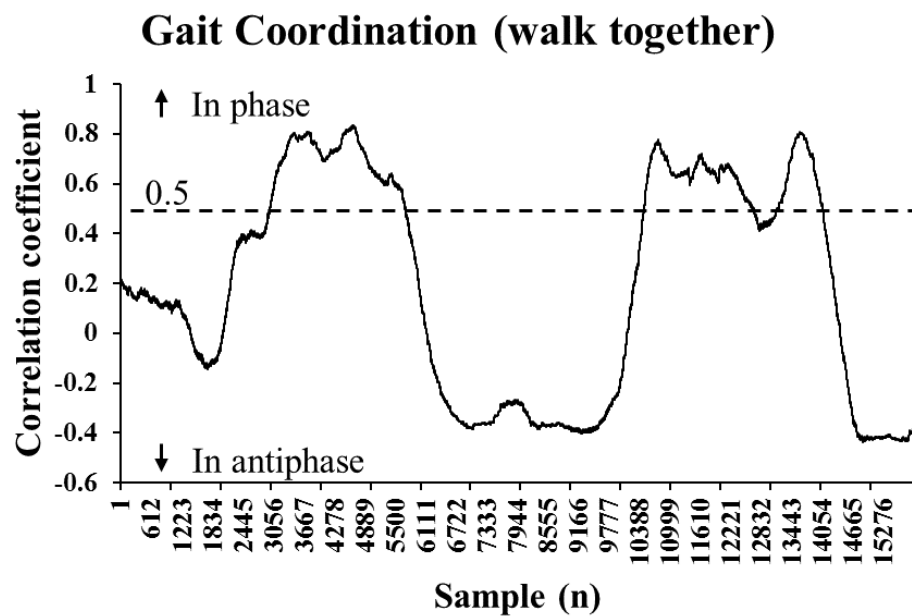
The Pearson correlation coefficient between double support periods of two participants ( $\rho_{X,Y}$ ) was evaluated as below, and uses data in a 10-s time window:

$$\rho_{X,Y} = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y} \quad (3.5)$$

where  $E[*]$  is expectation, X and Y is the status value of the double support period for two participants. For X and Y, the means are  $\mu_X$  and  $\mu_Y$ , and the standard deviations are  $\sigma_X$  and  $\sigma_Y$ , respectively. Figure 3.9(a) shows the correlation results of the first lap data when participants walk alone. Figure 3.9(b) displays the results of the second lap data, where both participants talk to each other during walking. Two diagrams show definitely different patterns. In Figure 3.9(a), the coefficient is oscillating but the peaks are normally less than 0.5. This graph indicates that both participants usually walked with temporally different gait cycles, (e.g., step time and cadence), so that the values oscillate up- and down-wards when they partially correlated by chance. Regarding Figure 3.9(b), two periods with high correlation values and one low valued period are demonstrated, which are observed in relatively long time (around 1 min. each). The high correlated periods (over 0.5) are when participants walked on the curved lanes of the track, showing their HS and TO are almost synchronized. In the lower plateau (around -0.4), participants walked on the straight lane between the curved lanes. Although the correlation coefficient is low, two participants had the same step time and cadence. In other words, their gait cycles are temporally synchronized, but a delay exists between them. The delay probably occurred when the path conditions are changed from the curved to straight lane. Depending on which lanes participants walked on, their gait speeds should be different on the curved lane. Entering the straight lane, participants should modify their gait speed to walk side by side. A delay of their gait cycles can occur, which changes the coefficient from 1.0 to -1.0. Regardless of the coefficient, they maintain the temporal synchronization of their gait cycles with the same delay. It is interesting that the delay is smaller on the curved lane because one on the outer lane might have additional visual cues of the partner's gait, whereas both walkers might have less visual cues of partners' gait on the straight lane, by looking forward.



(a)



(b)

Figure 3.9. Pearson correlation results between two participants' binary signals of the dual support period in two conditions: (a) during the first lap, when they walked alone and (b) during the second lap, when they walked together.

### 3.5. Discussion

In this chapter, another solution of H-IMU is described to directly investigate human behaviors in communicative interactions. Previously, researchers have measured rapport during interpersonal coordination, relying on onsite observation or video recordings (Tickle-Degnen et al., 1990). This method is, however, time consuming and labor intensive for examiners to count the numbers and to measure the duration of discrete meaningful behaviors, by watching videotapes for a long time. Technologies have been able to estimate one of the components at a time, such as with head pose estimation (Bruneli and Poggio, 1993; Maurer and von der Malsburg, 1996; Chen et al., 1998). In reviewing the literature, no study was found on measuring all the rapport components using one device at the same time. This dissertation, therefore, has first proposed a rapport measurement method by implementing H-IMU, which can automatically measure all the three rapport components, reducing measurement time and efforts of examiners.

The key advantage is that H-IMU enables gait analysis and head pose estimation together during walk and talk. However, H-IMU solution for rapport measurement is restricted to only the molecular method, which considers counting or analyzing interactants' discrete behaviors, whereas previous studies use both molar and molecular method, and consider analysis of the contextual information as well. In addition, it is reported that importance of the molecular method decreases over time because the discrete social behaviors can have different meanings depending on the social context and individuality.

Nevertheless, the proposed H-IMU solution cannot be devaluated because, regardless of the time, the molecular method contributes to the rapport measurement, providing the structured and objective evidence (Tickle-Degnen and Rosenthal, 1990). Counting specific behaviors with videotapes is also one of the most time-consuming works. Therefore, as an automatic and simple solution, H-IMUs still contribute to rapport measurement, reducing time and cost for human resources.

In terms of technological contribution, although only a single IMU is implemented, the proposed H-IMU solution can avoid drift effect resulting from error accumulation, which normally occurs when an IMU tracks the position for a long time. The accumulated

errors are caused by distorted sensor data from electromagnetic disturbances and data drops from network congestion as well as white noise. Compared to single-agent motion capture, drift effect is more problematic in measuring two or more agents because kinematic measurement of their mutual activity can be wrongly recorded, especially orientation (or rotation matrix) and position. In this work, a walking vector is considered as a reference direction. Although the sensor's position drifts during long time measurement, data of head turning angles are hardly affected because a reference direction is measured in short time. To be specific, the walking vector was calculated as a displacement of 1 second. The RMS of nodding angular velocity was also measured with 1 second window. For such a short time, the accumulated error is negligible.

### **3.6. Conclusion**

This chapter demonstrated that a single H-IMU can support the molecular measurements of a rapport between dyads during walk and talk. Based on the H-IMU, head pose estimation and gait analysis were carried out, which can support in monitoring physical and mental behaviors during interactions, such as walk and talk interventions. The method provides simplicity and convenience to users who care for the level of rapport in business, training, educational and clinical settings. The proposed method supports molecular methods, so that the importance of the proposed evaluation methods decreases over time because the meaning of behavioral gestures changes during interactions and the types of stimuli and responses increased and are interwoven. The evaluated values itself cannot fully represent the level of rapport; however, H-IMUs can provide a quick reference for overall judgements when combined with molar methods. Reducing molecular measurement efforts, H-IMU users can also improve efficiency in overall rapport measurement. In addition, studies on socializing and interactions in daily outdoor settings can be widely supported by consumer devices with IMUs (e.g., smartphones, earbuds and HMD). With a combination with outdoor counseling or real time feedback of movement, the simple and easy-to-access solution can finally contribute to public health, education, business, and even service robot industry in the future.



## **Chapter 4. Joint Task using IMU-based Application**

### **4.1. Introduction**

Concerning social interaction, various modes of nonverbal communication have been recently researched, such as gestures and facial expressions (Vicaria and Dickens, 2016), which contain spatiotemporal and affective components (Phillips-Silver and Keller, 2012). During nonverbal communication, perceptual cues are exchanged between interactants, which is fast enough to instantly react to; however, verbal or linguistic cues would be too slow to immediately react to other interactants (Knoblich and Jordan, 2003). Verbal communication is, as a result, used to prepare action plans and strategies before joint actions, whereas nonverbal communication induces emergent reactions between the interactants during joint actions. The importance of the emergent reaction is observed in mother-infant dyads, which is seen as the basic nonverbal communication. Two individuals display coaction such as facial expressions, voice, and social gaze when the baby is about three months old, and show gestures and joint attention to objects when the baby is about six months of age (Feldman, 2007). The mother-infant entrainment stems from motor resonance (Meltzoff and Decety, 2003) and emergent coordination, which can drive affective entrainment, cooperation, and rapport (Feldman, 2007).

Recent evidence suggests that the emergent coordination can be caused by stimulation on various perceptual modalities, such as visual, auditory, kinesthetic, and tactile systems (Marsh et al., 2009). For example, Waterhouse et al. (2014) demonstrated choreography performance of two dancers who synchronized their performance or aligned the onsets of their own performance. The nonverbal communication and the synchronization are relied on visual cues from their body movement as well as on auditory cues from breath and step sound. These phenomena can be observed in musical ensembles where musicians perform in a joint action setting. During piano duet performance, it is demonstrated that two pianists monitor auditory feedback from their own, their partners' and the joint action outcomes, which allows successful performance (Loehr, Kourtis, Vesper, Sebanz, and Knoblich, 2013). In addition, when the number of addressed perceptual modalities increases, temporal synchronization is normally enhanced during

interpersonal coordination (Knoblich and Jordan, 2003; Schmidt and Richardson, 2008). Goebel and Palmer (2009) reported that auditory and visual information might be complementary to one another: When the auditory information was reduced during ensemble performance, leaders and pianists show exaggerated movement for successful co-performance. Relationship between the perceptual modalities and temporal synchrony has been repeatedly reported in musical performance (Demos, Carter, Wanderley, and Palmer, 2017) and dance performance (Vicary et al., 2017). Demos et al. (2017) reported that, during the duet performance, asynchronies increased once the auditory feedback was removed, which confirms that removing one of the addressed perceptual modalities can increase temporal asynchrony in a joint task.

Demos et al. (2012) directly compared effects on interpersonal synchrony under different conditions with auditory and visual information in a rocking chair task. Authors showed that spontaneous coordination was enhanced in an audio-only (rocking-chair sound) and a visual-only condition, compared to a condition with neither vision nor sound. In audio-visual condition (rocking-chair sound and vision), the participants' synchrony was highest. These findings confirm again that a multiple number of perceptual modalities addressed by task related information usually enhance spontaneous coordination (see also Schmidt and Richardson, 2008). On the other hand, non-task-related music has no effect in audio only condition; however, the music in an audio-visual condition results in less synchrony than a vision-only condition (Demos et al., 2012). This finding indicates that a multiple number of perceptual modalities cannot be independent, but can cause interference, thereby leading to a negative effect on sensorimotor tasks. Demos et al. (2012) concluded that emergent perceptuo-motor couplings should occur with task-related information to enhance instantaneous coordination (see also Kelso, 1995). Under the multiple number of perceptual information, a study (Allerdissen, Gldenpenning, Schack, and Blsing, 2017) reported that the response to the perceptual feedback can be different, depending on personal experience. The authors measured timing of the occlusion of fencing novices and experts when they watched opponent's fencing attacks in a video, which demonstrated that the novices showed the negative effect in the audio-visual

condition, compared to vision-only condition. The experts, however, show no difference between the visual-audio and the visual-only condition because auditory information provided together with vision can be non-task-related for novices, but task-related for experts. Allerdissen et al. (2017) explained that non-task-related perceptual feedback might cause cognitive overload, which negatively influences novices.

In a study on auditory perception and movements, it has been previously observed that auditory-motor co-activation in individuals' brain is emerged during short training in a musical condition (Bangert, Peschel, Schlaug, Rotte, Drescher, and Hinrichs et al., 2006), and a team sports condition (Schmitz and Effenberg, 2017). One of the significant current discussions is which factors cause the enhancement of auditory and motor processing during interpersonal coordination. Vesper et al. (2013) also conducted an experiment with a jump task. A pair of participants were asked to jump forward next to one another and to land at the same time, while their landing positions are informed with auditory and visual feedback before they jump. Although online sensory feedback was not provided, the participants adjusted preparing and executing motion of jumps, relying on given spatiotemporal information. Authors indicated that spatiotemporal properties of the sensory feedback can enable interpersonal coordination as well as intrapersonal coordination. The successful spontaneous coordination can emerge with auditory feedback when participants distinguish their own sound from the partner's (Loehr et al., 2013). Another study (Murgia, Hohmann, Galmonte, Raab, and Agostini, 2012) demonstrated that participants can identify their own golf swing sounds, which were naturally recorded during their performance. The authors insisted that temporal factors are important in self-other-discrimination because participants wrongly perceived a swing sound as their own swing, when other's swing are similar temporal components in terms of the overall duration and the relative timing of the swing. However, Kennel et al. (2014) failed to confirm the correlation between self-other-discriminations and temporal factors of natural movement sounds previously recorded in hurdling performance. The authors concluded that the self-and other-movement sound is distinguishable because the individuality of sounds activates listener's own sensorimotor memory. In addition, the ability to reproduce perceived

movement increases with more appropriate internal models (sensorimotor, visual, or auditory perceptions) (Kennel et al., 2014). When perceptual information is provided in real-time, it is beneficial for participants to anticipate their own and co-performers' actions (Keller and Appel, 2010; Keller, 2012). From a neurophysiological analysis, Keller et al. (2014) suggested that the additional auditory information improves the phase correction via a neural pathway between subsections of the cerebellum which are linked to motor and auditory cortices. Additional auditory feedback also can enhance the periodic correction by additionally recruiting a corticothalamic network that includes the prefrontal, medial frontal, and parietal cortex, as well as basal ganglia (Keller et al., 2014; Repp and Su, 2013).

As a result, researchers have underlined the positive effects of real-time kinematic auditory feedback on motor control and learning (Effenberg et al., 2016; Effenberg, 2005). In single-agent settings, Effenberg et al. (2016) demonstrated the faster development of motor representations and the enhancement of preciseness with the presence of real-time auditory feedback. The additional real-time sound, artificially mapped onto human kinematic and dynamic patterns, should enhance the development of sensorimotor representations without consciousness (Effenberg, 2005; Effenberg et al., 2016). For the study on motor learning, this auditory movement feedback can be implemented by effect-based auditory feedback (EAF) and performance-based auditory feedback (PAF), which stems from "Knowledge of performance" (KP) and the "Knowledge of result" (KR) (Schmidt and Wrisberg, 2008). Performance-based feedback is related to the movement quality, whereas the effect-based feedback provides the movement result (Magill and Anderson, 2007; Schmidt and Wrisberg, 2008). The benefits of performance-based feedback on learning were reported by many researchers (Sharma, Chevidikunnan, Khan, and Gaowgzeh, 2016; Nunes, Souza, Basso, Monteiro, Corrêa, and Santos, 2014; Weeks and Kordus, 1998) as well as the impact of effect-based information (Sharma et al., 2016; Schmidt and Wrisberg, 2008; Winstein, 1991). Two types of feedback can be compared in a number of studies on motor learning. Above findings indicate that additional auditory feedback can have effects on motor learning and interpersonal coordination, which can differ because of the relationship with KP and KR as well as the task, one's experience,

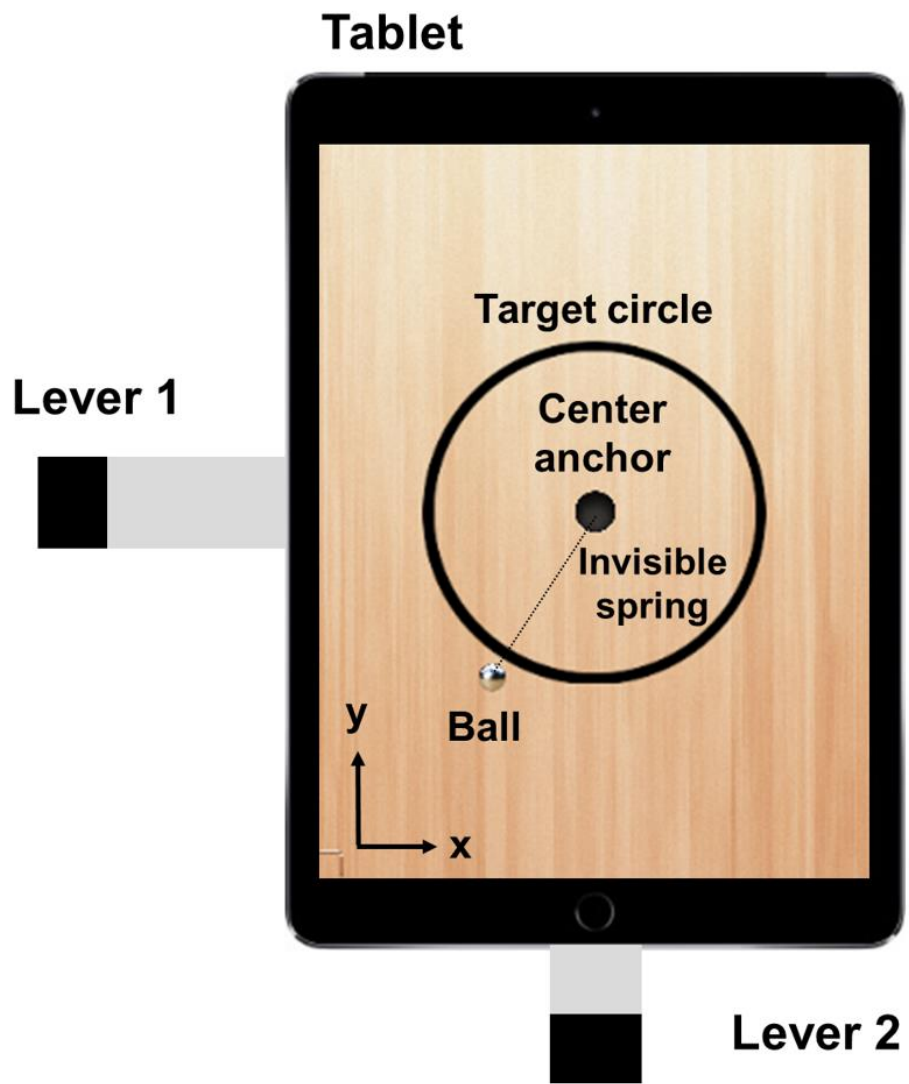


Figure 4.1. Top view of the tablet screen and levers fixed to the casing.

and gestalt movement. Feedbacks based on KP and KR were implemented to explore the impact on the interpersonal setting with a collaborative task in this study.

A novel paradigm was developed by programming a new iPad-application and creating a mechanical apparatus as shown in Figure 4.1 and Figure 4.2, called the *tetherball paradigm* (Hwang et al. 2018a). The mission of the task is to manage the tetherball to travel on a given circular track. The ball is tethered to the center on the tablet screen by spring that has an elastic force, so that the ball tends to be close to the center when tablet stays flat. A pair of participants, therefore, has to tilt the tablet and accelerate the metal ball to travel on the target circle (Figure 4.1). Additionally, a participant allowed to move the tablet by the index finger along the only x- or y-axis by using one of the levers. Two participants, as a result, have to collaborate with temporal synchronization to make a circular trajectory of the ball. This paradigm artificially mapped real-time kinematic auditory feedbacks onto tablet tilt rate (mapped to PAF), and moving ball position (mapped to EAF). Each feedback is divided into x- and y-axis to inform both participants of their own and partners' actions as well as the joint outcomes. With a presence of visual feedback of the movements, the artificial auditory feedbacks are increasing the number of perceptual modalities. Although kinesthetic and tactile feedback might affect this task, visual and auditory feedback play dominant roles in these kind of tasks. It is demonstrated that visual information might be most important in perceiving action and its effects (Demos et al., 2012). Auditory feedback might remarkably alter the participants' perception because the human auditory system is powerful especially in perceiving the temporal factors of acoustic events, and in identifying pace and rhythm specification (Collier and Logan, 2000; Murgia, Prpic, McCullagh, Santoro, Galmonte, 2017). In addition, the human auditory system is highly effective not only in being aware of smoothness and regularity of motions, but also in synchronizing temporally and adjusting actions to external events (Repp and Penel, 2002). Therefore, better joint outcomes, temporal synchronization, and collaboration experience are expected with the auditory perceptual system involved additionally.

In this study (Hwang et al., 2018a), three different visual-audio conditions (EAF, PAF, and combined EAF and PAF (CAF)) are compared to a visual-only condition (VF;

no audio). In the EAF condition, a melodic sound informs a pair of participants of ball dynamics that is the joint outcomes of the task. In the PAF condition, a rhythmical sound is provided for feedback of rhythmical movement, which is similar to the sound in a research by Demos et al. (2012). EAF is designed to provide the participants with non-rhythmical sound because rhythmical sound can contain PAF components, affecting the task results. Unlike these two separated conditions, two types of sound are mixed in the CAF condition. For all conditions, positive hypotheses were defined referring to previous literature (Vesper et al., 2013; Effenberg et al., 2016). In each condition, the mean absolute error (MAE) was measured to evaluate the improvement of joint outcomes over time, and the cross-correlation between the tablet tilt rates in *x*- and *y*-axis is assessed for the synchronization of their own and partners' actions. Participants self-reports were asked to analyze their subjective experience during interpersonal coordination. According to these data, the hypotheses described below were tested (Hwang et al., 2018a):

H1: "Faster error reduction in the task is achieved when participants are provided with either additional (a) effect-based, (b) performance-based, or (c) both combined auditory feedback."

H2: "Cross-correlation in the participants' actions is stronger when participants are provided with either additional (a) effect-based, (b) performance-based, or (c) both combined auditory feedback."

H3: "Subjective ratings of the sense of interpersonal coordination are more positive when participants are additionally provided with (a) effect-based, (b) performance-based, or (c) both combined auditory feedback."

## **4.2. Materials and Methods**

### **4.2.1. Participants**

In total 72 healthy participants (female: 30, male: 42, age:  $24.8 \pm 3.3$  years) were tested. They have normal eyesight and hearing abilities. Participants were divided into four

different groups for the four different conditions, which means nine pairs are allocated to each group. In pairing participants, only one criterion is that they should be a “same-sex pair.” Participants are also asked to use their dominant hand. The ethical approval of the study was issued by the Ethics Committee of Leibniz University Hannover.

#### **4.2.2. System Specifications**

The application of this paradigm was implemented using Objective-C with iPad Air (Apple Inc.) operating on iOS 10.2. Display resolution was  $2,048 \times 1,536$  pixels (px), and the frame rate is 60 Hz. An accelerometer and gyroscope in the tablet were also sampled at 60 Hz. The sound was designed with the Csound 6 (free software under LGPL; Boulanger, 2000) and a physical model was implemented by Chipmunk2D Pro (Howling Moon Software). During the task, participants wore headphones, (Beyerdynamic DT 100). The auditory signal was divided by a 4-channel stereo headphone amplifier (Behringer MicroAMP HA400) to supply a pair of participants with the real time audio signal.

#### **4.2.3. Design and Stimuli**

The main screen of the tablet application is shown in Figure 4.1. The ball is tethered to the center of the screen by an invisible spring, and the circular track is surrounding the center. At the edge of the screen, it is shown that two levers are orthogonally fixed on the tablet. The application runs on XGA ( $1,024 \times 768$ ) resolution, where the ball’s radius is 30 px, whereas the radius of the circular track is 232.5 px (inner: 225 px, outer: 240 px) and the thickness of the track line is 15 px. In  $x$ - $y$  Cartesian coordinate system, the center of the ball position is considered in the analysis. The ball is tethered to the center by an invisible spring whose elastic force makes the ball closer to the center when the tablet stays flat. For the mission of the task, participants have to tilt the tablet to let the ball travel on the circular track surrounding the center. Each participant uses the index finger of the dominant hand to manipulate the tablet, which allows move up and down along only  $x$ - or  $y$ -axis. The long lever at the  $x$ -axis is designed compared to the  $y$ -axis because of the compensation for the different lengths of the tablet edge. The tablet is restricted to two degrees of freedom (DOF)



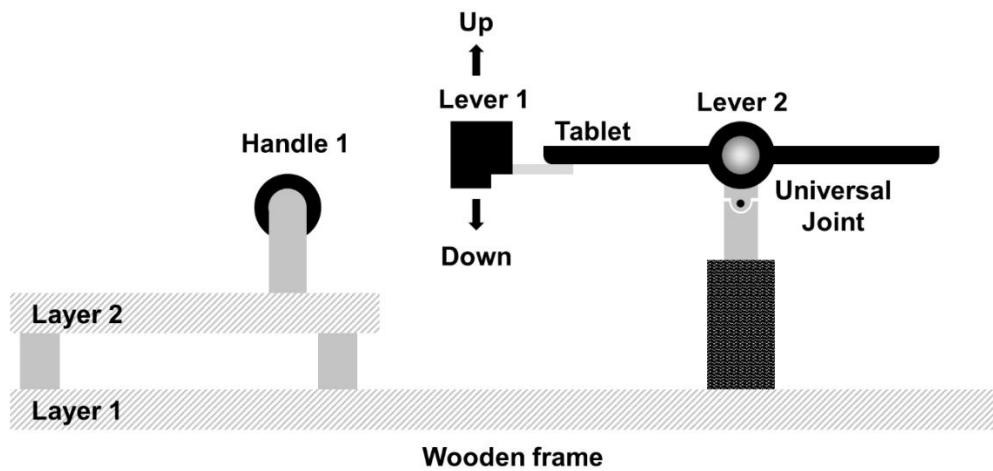


Figure 4.2. Side view of the mechanical structure which supports the tablet and restricts to 2-DOF movement of the tablet.

along the  $x$ - and  $y$ -axis, which means  $z$ -axis rotation is not allowed (see Figure 4.2). For the successful joint outcome, it is assessed how close the ball trajectory is located to the circular track. The tetherball's circular trajectory can be realized when both participants synchronize to tilt the tablet in terms of movement amplitude and frequency, but with a  $90^\circ$  phase difference.

Mechanical structure of the experimental apparatus is depicted in Figure 4.2 in the side view. Participants can manipulate the tablet using the levers attached to the casing. A vertical bar supports the tablet with a universal joint that restricts the movement to the  $x$ - and  $y$ -axis (roll and pitch), which prevents the yaw rotation around the  $z$ -axis. Participants were guided to hold the horizontal handle bar (Figure 4.2) that prevents from redundant hand movement. Participants' elbows and wrists are comfortably rested on the wooden frame (the layer 2 in Figure 4.2).

The top view of the mechanical wooden structure is shown in Figure 4.3(a). Each participant is asked to take a seat depending on their dominant hand. Right-handed (RH) participants take a position on the left of one of the wooden wings, and then they let their right arm lie on the layer 2 on their right side. Left-handed (LH) participants can sit on the right of a wooden wing to put their left arm on the wood table. The participants were not

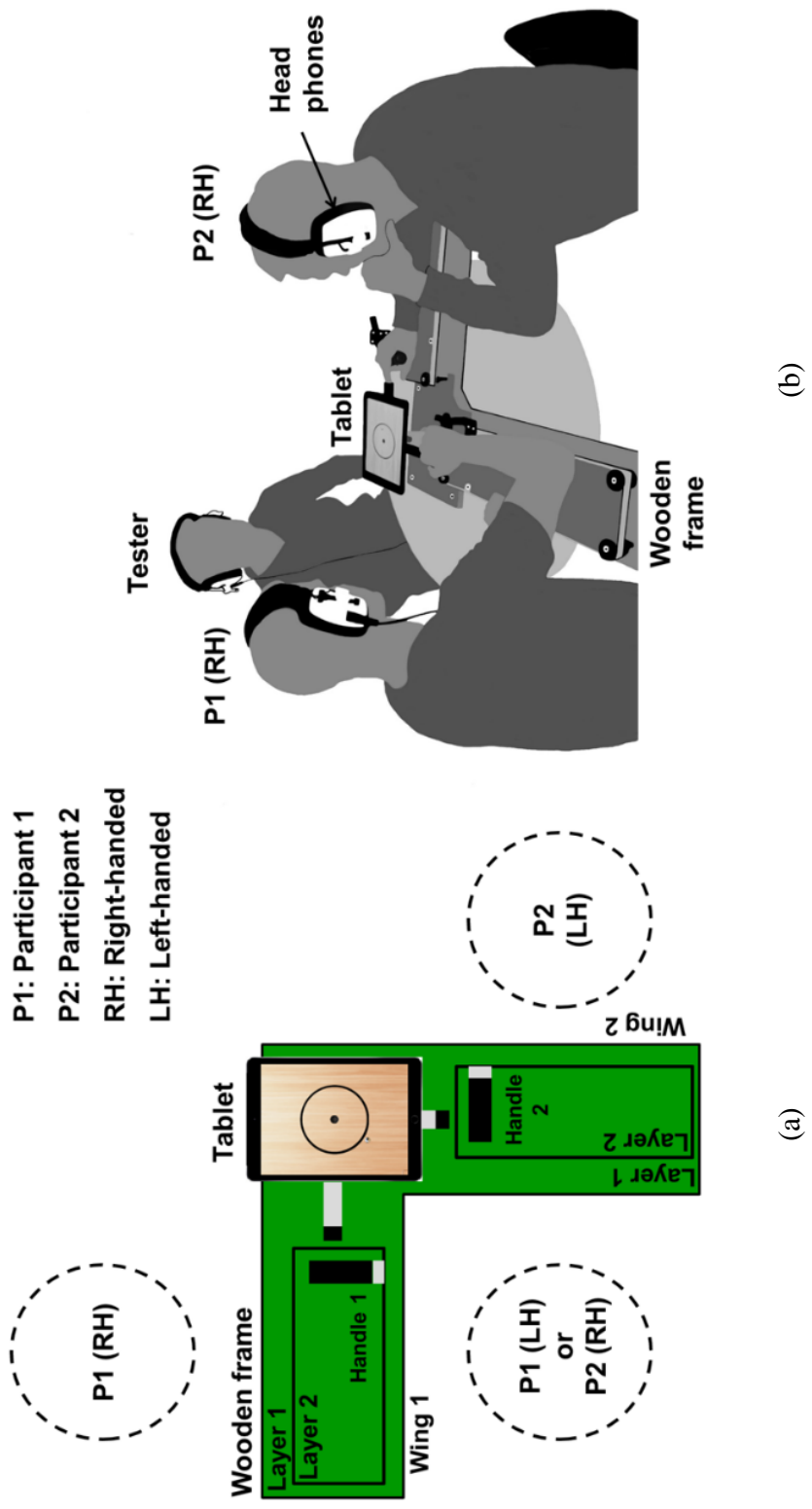


Figure 4.3. Illustrations of (a) the top view of the mechanical structure and the seats for left-handed (LH) and right-handed (RH) participants (P1 and P2), as well as (b) the exemplary situation of the experiment with two right-handed participants and a tester.

allowed to swap their positions during the whole task. The horizontal handle bars can be adjusted to the individual hand size and handedness, which allows participants to manipulate the tablet comfortably. In addition, participants can watch the tablet display from almost the same distance, which guarantees that participants have visual information in nearly the same condition. The audio output signal from the tablet was split into both performers, the same auditory information reaches them via headphones. The headphones can monitor their own and partners' actions as well as joint outcomes by auditory feedback.

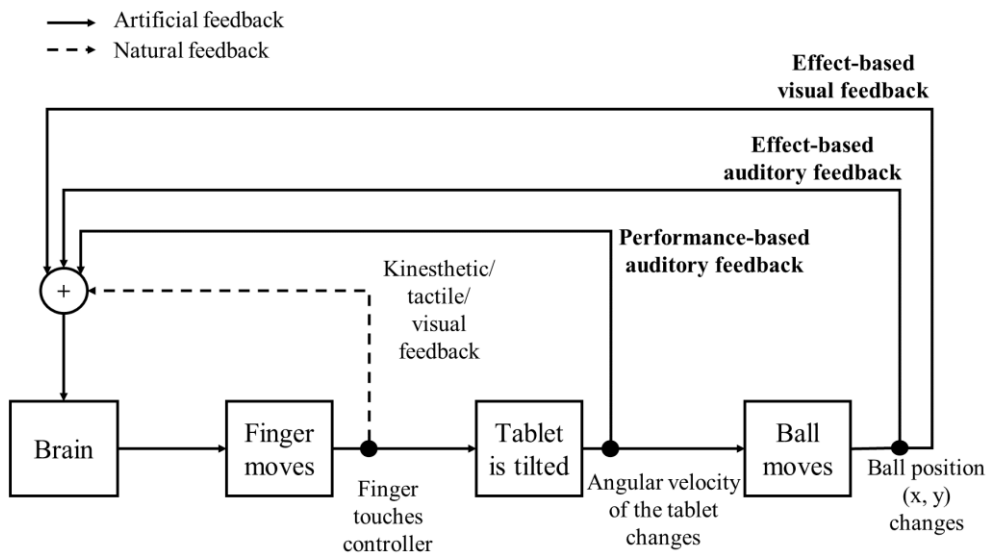


Figure 4.4. The feedback loop of perceptual information during the joint task.

Figure 4.4 shows the feedback loop in terms of the perceptual information during the task. Movements of participants and their effects provide kinesthetic, visual, tactile, and auditory information. Effect-based visual feedback (VF), EAF, and PAF are artificially generated by the tablet application in the measurement, whereas the kinesthetic, tactile, and visual feedback of finger movements are provided by the environment independently from the measurement. Visual information of the ball movement is defined as VF, which is congruent to EAF that is driven by the  $x$ - $y$  ball position (audio frequency) and the velocity (amplitude). Continuous synthesized violin sound is used for EAF, which is appropriate to inform of the ball's continuous movement. The EAF violin sound is distinguishable between  $x$ - and  $y$ -position. In addition, it is familiar to human ears, so that participants can

hear it comfortably for a long time. The base audio frequency ( $f_b$ ) is between 250 and 427 Hz along the  $x$ -axis and between 600 and 835 Hz along the  $y$ -axis.

The PAF sound is characterized by rhythmical tilt movements. Referring to the embedded gyroscope data, the amplitude of PAF is driven by the rate of turn of the tablet, resulting in a certain congruence with the kinesthetic feedback of the finger movements. The sound of PAF is created by a white noise generator and band-pass filter (cutoff frequency:  $f_b \pm 25$  Hz), which is similar to is “wind sound” when broom sweeps on the floor. The base audio frequency is separated referring to each participant. The centrifugal force measured by the built-in accelerometer changed  $f_b$  from 700 to 1,700 Hz (lever 1) and from 100 to 1,100 Hz (lever 2). By using a noisy wind sound, the PAF sound timbre can imitate natural wind sound, so that the sound is comfortable for the participants to hear. The PAF is clearly distinguishable from the EAF, which allows the participants to have two types of auditory information together in the CAF condition. These two types of augmented auditory feedback (PAF, EAF) are designed for three audio-visual conditions, which is hearable or perceivable over the given natural feedbacks such as kinesthetic, visual and tactile feedback,

#### **4.2.4. Procedure**

At the beginning of the test, participants were given a questionnaire to complete their personal backgrounds such as music and sports experience. Eyesight and hearing abilities were tested by the Landolt rings chart (Jochen Meyer-Hilberg) and the HTTS audio test (SAX GmbH), which confirms that the participants can properly perform with the tetherball paradigm. Visio-motor ability is also tested by the tablet application, in which a participant let a ball placed on a target that moves along only one axis but changing the direction randomly as shown in Figure 4.5. Two participants performed this visuo-motor test together, but each participant is responsible for their own ball moving along their corresponding axis.

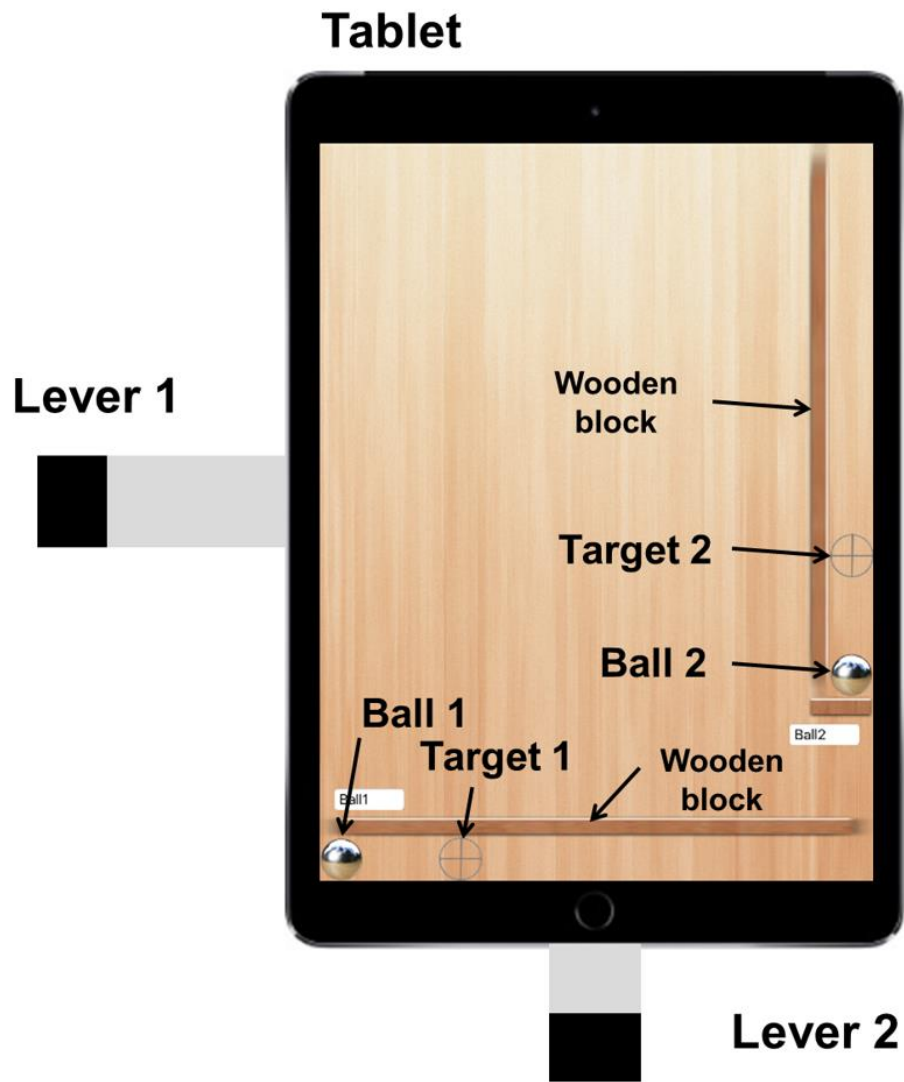


Figure 4.5. Top view of a visuo-motor test application to measure the individual visuo-motor ability which is an important skill for the performance at the tetherball-application.

The participants' visuo-motor skills were assessed in the visuo-motor test application for 2 minutes, which also makes participants familiarized in manipulating the tablet. When participants perform the visuo-motor test, the mean absolute error between the moving target and ball was measured for 30 s at the end of the 2-minute period. Participants' mean absolute error and gender are parallelized in four groups. The visual group (VFG) allows only VF without any types of auditory feedback, which works as a reference condition. Additional to the VF, the EAF group (EAFG), PAF group (PAFG), and CAF group (CAFG) receives EAF, PAF, and CAF, respectively. All four groups also got natural perceptual feedback: kinesthetic, visual, and tactile feedback. The results of the visuo-motor test are well parallelized in four groups, which have nearly the same level in the average and standard deviation [VFG:  $75 \pm 23$  px, EAFG:  $72 \pm 29$  px, PAFG:  $70 \pm 20$  px, CAFG:  $78 \pm 16$  px] and in analysis of variance (ANOVA) [ $F(3, 36) = 0.24$ ,  $p = 0.872$ ,  $\eta_p^2 = 0.02$ ].

Pairs of participants performed each trial for 1 minute, and 15 trials in total. Five trials are bound to one set, resulting in a total of three sets. Between the sets, a 2-minute break was given to each participant. In the main test, participants were asked not to talk about the possible strategies, which leads to the nonverbal joint performance as well as the concentration on the task. They were also instructed to let the ball revolve in clockwise direction (CW). After the test, the second questionnaires were given to the participants for subjective ratings of their experience during interpersonal coordination in the tetherball paradigm. The questionnaire included four questions with 7-point Likert scale (Likert, R., 1932), which subjectively evaluates the own, partners', and joint outcomes during the test.

#### **4.2.5. Data Analysis**

The tablet application recorded the ball trajectories (from screen) and the angular velocity of the tablet (from gyroscope) at a 60 Hz sampling rate. In each trial, mean absolute errors and the cross correlations were submitted to three-way ANOVAs of a between-subject factor Group (VFG vs. EAFG, VFG vs. PAFG, VFG vs. CAFG) and the within-subject factors Set (set I–III) and five trials in each set. The Mauchly's sphericity test was

conducted (with Huynh–Feldt correction for significance of ANOVA) for the sphericity assumption, which met the variances between the individual factor levels as well as the correlations between the factor levels are homogeneous. Levene's test was carried out to assess homogeneity of variances and Tukey's honestly significant different (HSD) tests were implemented for the post-hoc comparisons. Participants' self-report of the interpersonal coordination were compared with Mann–Whitney-U-Tests for across groups and a Wilcoxon test for within-groups. The overall significance level was set at 5%.

### 4.3. Results

Expertise in sports, music, and computer game were asked, as well as visuo-motor test performance were taken into account in the data analysis because they could affect the overall results in the tetherball paradigm. In comparisons between these parameters of the visual group (VFG) and those of the audio-visual groups (EAFG, PAFG, CAFG), there were no significant differences except on sport expertise between the VFG and PAFG [ $F(1, 16) = 6.38, p = 0.022, \eta_p^2 = 0.29$ ]. Therefore, sport specific expertise was considered as a possible covariate in the future analyses.

For the main analysis, the joint outcomes were assessed by the mean absolute error between the radius of the circular track and the ball's trajectory. A mean absolute error during a 1-minute trial was calculated, but 500 samples (8.3 s) at the beginning were omitted, which is the average time about to initiate the circling ball's movement. With the mean absolute error, the means and standard deviations were calculated across the subjects for each trial and in each group as shown in Figure 4.6.

In Figure 4.6, comparing VFG with EAFG, the mean absolute error across trials decreased significantly from Set I to Set II and to Set III as shown by the significant main effect for set [ $F(2, 30) = 3.95, p = 0.043, \eta_p^2 = 0.21$ ]. This is also confirmed in the post-hoc test between Set I to Set II ( $p < 0.001$ ) and Set I to Set III ( $p < 0.001$ ), resulting in the significant differences. Within each set, moreover, the mean absolute error is reduced from Trial 1 to Trial 5 [ $F(4, 60) = 4.58, p = 0.005, \eta_p^2 = 0.23$ ]. A post-hoc comparison confirmed

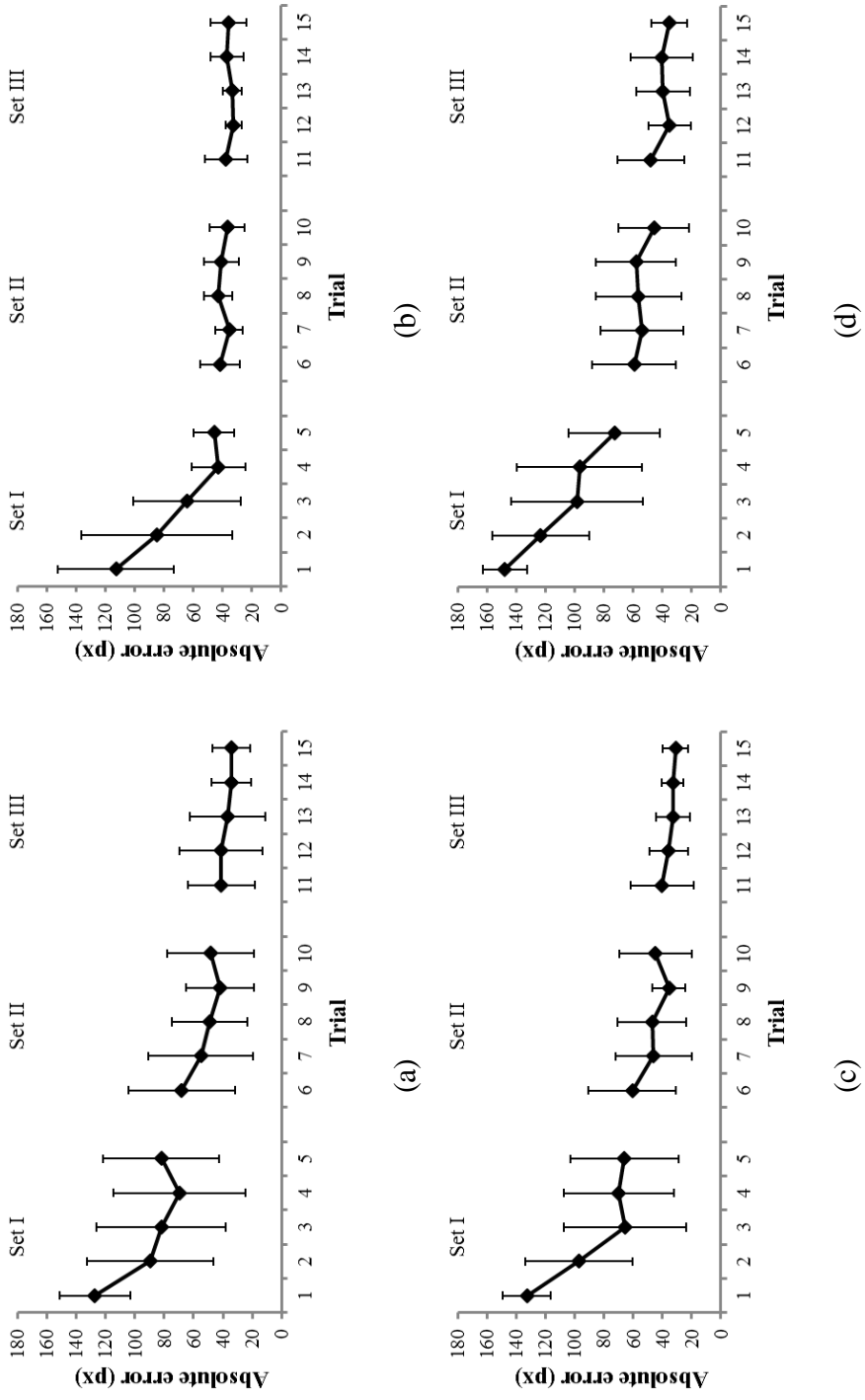


Figure 4.6. Across subject means and standard deviations of the absolute error over the trials in four groups, (a) the VFG, (b) the EAFG, (c) the PAFG, and (d) the CAFG. Between-subject means and standard deviations are illustrated. The first 8.3 s (500 samples) of every 1-minute trial was eliminated.



that Trial 1 was significantly different from all the other trials (all  $p < 0.001$ ), and Trial 2–4 from Trial 5 (both  $p < 0.01$ ). The reduction trend of the mean absolute error across trials was significantly correlated with the sport specific expertise [ $F(4, 60) = 3.84$ ,  $p = 0.013$ ,  $\eta_p^2 = 0.20$ ].

The three-way interaction Set\*Trial\*Group confirmed significant difference in the reduction trend of the mean absolute error [ $F(8, 120) = 2.63$ ,  $p = 0.030$ ,  $\eta_p^2 = 0.15$ ] (due to  $p < 0.05$ ). In EAFG, the mean absolute error decreased dynamically in the first four trials and then reached a stable plateau at the fourth trial. A post-hoc test confirmed significant differences from Trial 1–3 to the last trial (at least  $p < 0.05$ ); however, no significant differences from Trial 4 to all remaining trials became significant (all  $p > 0.05$ ). Compared to EAFG, the results in VFG slowly reached a plateau as the same level as that in EAFG as shown in Figure 4.6. The post-hoc test showed significant differences from Trial 1–6 (the first six trials) to Trial 13–15 (the last three trials) (at least  $p < 0.05$ ). Levene's test revealed that variances significantly differed between groups in Trials 4–8 and Trial 12 (at least  $p < 0.05$ ).

In contrast to the EAFG, the PAFG failed to confirm a significant difference from the VFG across group and in their interactions. Analysis across groups, nevertheless, showed significant the main effects [Set:  $F(2, 32) = 56.66$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.78$ ; Trial:  $F(4, 64) = 40.81$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.72$ ], as well as interactions [ $F(4, 64) = 10.19$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.39$ ]. In a post-hoc test, the latter interaction is significantly different between Trial 1–2 and Trial 3–5 in Set I (at least  $p < 0.05$ ) as well as between Trial 6 and Trial 9–10 in Set II (at least  $p < 0.05$ ), but not in Set III (all  $p > 0.05$ ). Obviously the mean absolute error decreased mainly in Set I and the beginning of Set II but plateaued in Set III. The Levene's test failed to show significant differences in the other trials.

An ANOVA of VFG and CAFG revealed the same overall effects as the other groups' results [Set:  $F(2, 32) = 67.26$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.81$ ; Trial:  $F(4, 64) = 35.76$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.69$ ], and demonstrated a significant interaction of Set\*Trial [ $F(4, 64) = 10.56$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.40$ ]. The CAFG, moreover, had a significant difference on the reduction course of the mean absolute error, which is supported by significant interactions in

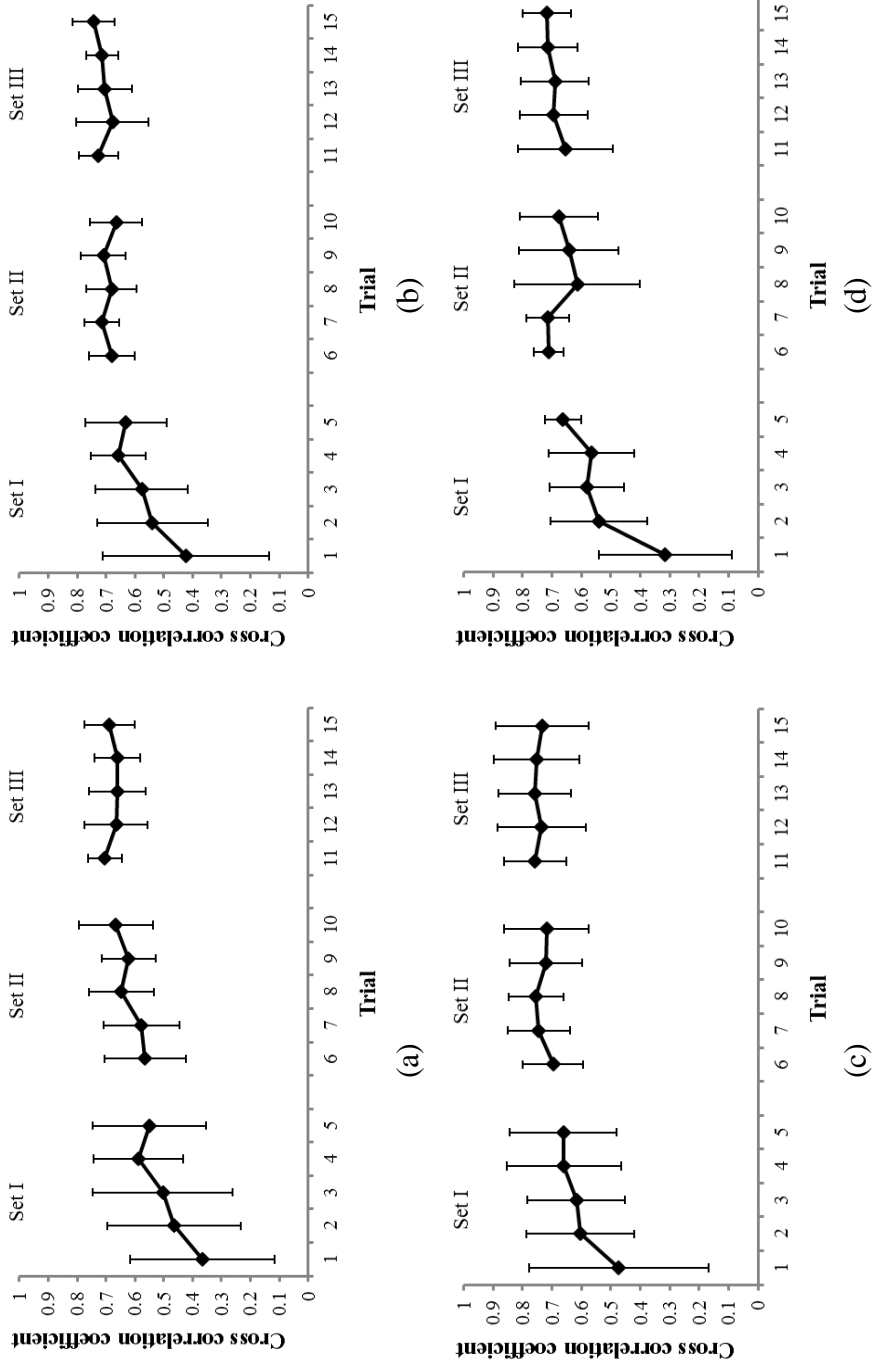


Figure 4.7. Across subject means and standard deviations about the cross correlation coefficient of a pair of participants over the trials in four groups, (a) the VFG, (b) the EAFG, (c) the PAFG, and (d) the CAFG. To take maximal performance of pairs, the maximal coefficient in each trial was chosen with lags equivalent about  $90^\circ$  phase delay. The first 8.3 s (500 samples) of every 1-minute trial were eliminated.

ANOVA with Trial\*Group [ $F(4, 64) = 3.70, p = 0.021, \eta_p^2 = 0.19$ ]. According to the post-hoc test of the CAFG, the mean absolute error further decreased between the second last and the last trial ( $p = 0.02$ ), compared to VFG ( $p > 0.05$ ). In addition, the significant interaction in three-way ANOVA with Set\*Trial\*Group [ $F(8, 128) = 2.45, p = 0.031, \eta_p^2 = 0.13$ ] revealed that the mean absolute error was differently reduced between groups over time. Compared to VFG, the mean absolute error reached a plateau on an early trial in CAFG, which is supported by Tukey's post-hoc test: the mean absolute error was decreasing during the first five trials (Trial 1–5) that shows significant differences from the last trial in CAFG (all  $p < 0.001$ ). In VFG, however, the mean absolute error is getting reduced during the first six trials (Trial 1–6), that displayed significant differences from the last trial (all  $p < 0.001$ ). This indicated that the mean absolute errors in CAFG were decreasing faster than in VFG. Levene's test did not reveal significant differences in any of the trials.

For another analysis, cross correlation is applied for the measurement similarity of two series in numerical data analysis and statistics (Proakis and Manolakis, 1996). Data processing using cross correlation can also analyze time delay between two data series (Rhudy, Bucci, Viperman, Allanach, and Abraham, 2009), which is applicable to temporal parameter analysis of interpersonal coordination (Cornejo, Cuadros, Morales, and Paredes, 2017). In this experiment, cross correlation between angular velocities of  $x$ - and  $y$ -axis of the tablet was used to measure the degree of temporal synchronization between finger movements of both participants, which is calculated with all other pairs (Figure 4.7).

First of all, 1,000 samples were taken into account for a unit of cross correlation, and three units in three different periods were included in a 1-minute trial. In total 3,610 samples (60.2 s), 3,000 samples are analyzed. At the beginning, the ball improperly revolved around the center, so that 500 samples were eliminated. Another 100 samples are placed between the cross correlation units, and 10 samples were placed at the end. Each cross correlation unit resulted in a series of coefficients along with lags. Coefficients with a lag between  $n = 8$  and 15 were used, which indicates synchronization with around  $90^\circ$  phase difference for the circling ball. The range of the lags served as an empirical standard,

which was obtained by data from 12 pairs (three pairs per group) who achieved the lowest mean absolute error in Trial 11–15. The data from the pairs showed 11.65 ( $\pm 4.53$ ) samples differences, when they were performing with their maximum joint outcomes. From the numerical result, the range of the lags between  $n = 8$  and 15 can be focused to assess the level of synchronization. The mean coefficient in the range of a unit (1,000 samples) is compared with the other two units in each trial. The highest value was regarded as a representative cross correlation result of the trial, which can indicate the maximum joint outcome in each trial. Dividing into three units can also avoid the average effect of the data that can be altered by participants' mistakes. It is observed that cross correlation results improved over time according to across subject's means and standard deviations (Figure 4.7). This was also supported by ANOVAs which confirmed the significance of the factor "Set" with the VFG and audio-visual groups [VFG and EAFG:  $F(2, 32) = 26.81$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.63$ ; VFG and PAFG:  $F(2, 32) = 21.17$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.57$ ; VFG and CAFG:  $F(2, 32) = 26.82$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.63$ ] as well as the factor "Trial" [VFG and EAFG:  $F(4, 64) = 5.49$ ,  $p = 0.003$ ,  $\eta_p^2 = 0.26$ ; VFG and PAFG:  $F(4, 64) = 5.48$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.26$ ; VFG and CAFG:  $F(4, 64) = 8.68$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.35$ ]. The improvement of cross correlation in these groups was also indicated by the significant interactions Set\*Trial [VFG and EAFG:  $F(8, 128) = 5.25$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.25$ ; VFG and PAFG:  $F(8, 128) = 4.68$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.23$ ; VFG and CAFG:  $F(8, 128) = 6.20$ ,  $p < 0.001$ ,  $\eta_p^2 = 0.28$ ]. Interestingly, cross correlations in CAFG increased significantly faster than in VFG. The significance is confirmed by ANOVA with the three-way interaction Set\*Trial\*Group [ $F(8, 128) = 2.53$ ,  $p = 0.014$ ,  $\eta_p^2 = 0.14$ ]. A Tukey's post-hoc test revealed significant differences between Trial 1–3 and Trial 15 (the last trial) within VFG (each  $p < 0.05$ ). In CAFG, on the other hand, only Trial 1–2 significantly differed from the last trial (each  $p < 0.05$ ).

Results of participants' self-reports are shown with the box and whisker plots (Figure 4.8). Participants were asked to answer a questionnaire with 7-point Likert scales (the most negative: 1, the most positive: 7). The results of the first question "How much did you feel your movement helps the collaborator's performance?" resulted in no

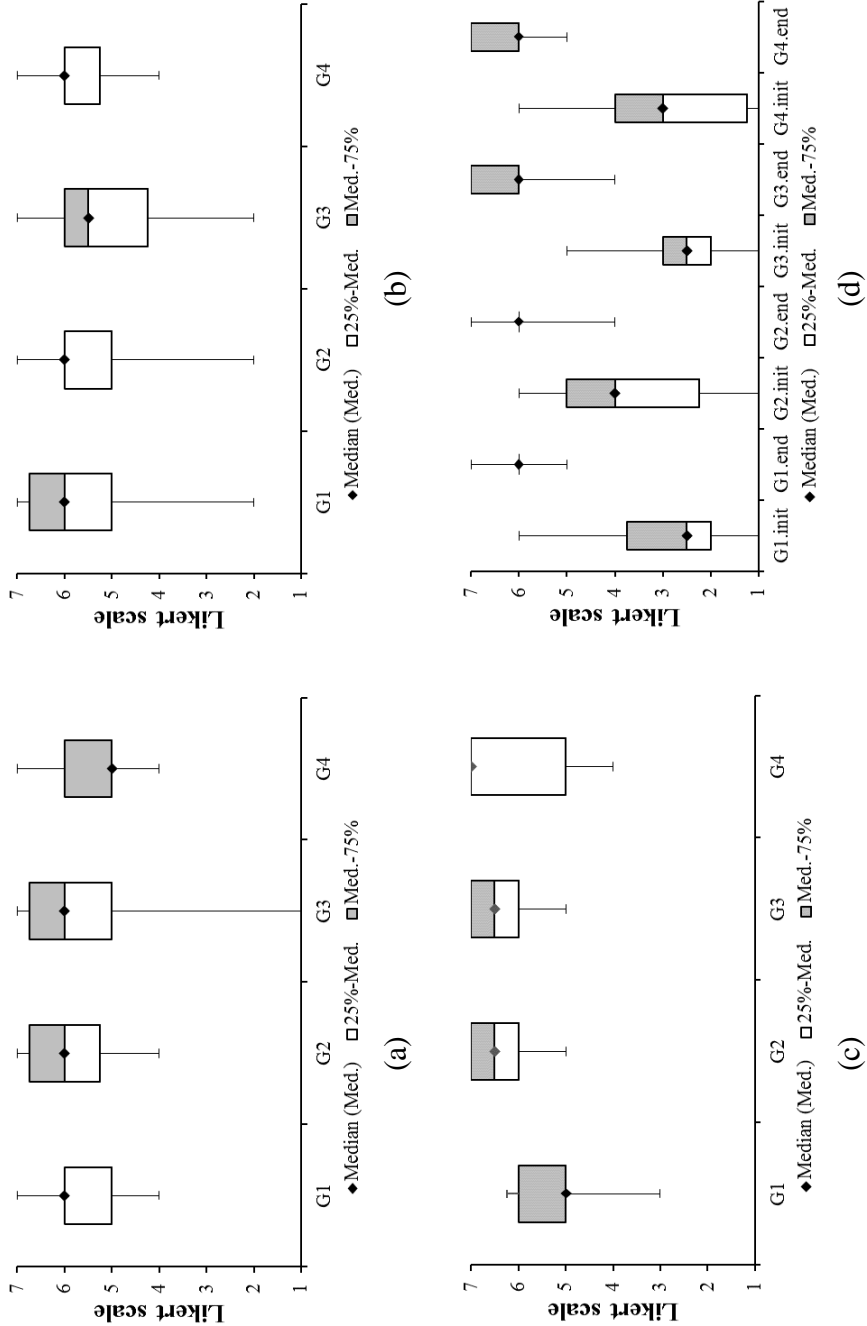


Figure 4.8. Box and whisker plots of subjective ratings when participants were asked about how much they felt (a) that their own action helped their partners, (b) that their partners' action helped their own action, (c) how pleasant their experience was during the collaboration, and (d) how effectively they felt that they handled the apparatus together at the initial time and at the end.

significant differences between VFG and other audio-visual groups, which is confirmed by Mann–Whitney U-tests [VFG vs. EAFG:  $U = 150.0$ ,  $p = 0.719$ ; VFG vs. PAFG:  $U = 161.0$ ,  $p = 0.988$ ; VFG vs. CAFG:  $U = 120.0$ ,  $p = 0.192$ ]. The second question “How much did you feel the collaborator's movement helps your performance?” also showed no significant differences [VFG vs. EAFG:  $U = 159.5$ ,  $p = 0.938$ ; VFG vs. PAFG:  $U = 136.0$ ,  $p = 0.424$ ; VFG vs. CAFG:  $U = 159.0$ ,  $p = 0.938$ ]. However, in the third question “How did you experience the collaboration with your partner?” (unpleasant: 1, very pleasant: 7), all audio-visual groups showed significant differences to VFG [VFG vs. EAFG:  $U = 66.0$ ,  $p = 0.002$ ; VFG vs. PAFG:  $U = 60.0$ ,  $p = 0.001$ ; VFG vs. CAFG:  $U = 90.5$ ,  $p = 0.022$ ]. In the fourth question “How effectively did you feel that you managed to do the task?” (not effectively at all: 1, very effectively: 7), two sub-questions were asked to answer their feelings “at the beginning” and “at the end” of the task. None of audio-visual groups showed significant differences [VFG vs. PAFG:  $U = 151.0$ ,  $p = 0.743$ ; VFG vs. EAFG:  $U = 103.0$ ,  $p = 0.064$ ; VFG vs. CAFG:  $U = 156.0$ ,  $p = 0.864$ ]; however, only EAFG showed a tendency of difference between VFG and EAFG ( $p = 0.064$ ). Participants in all groups felt that they had an improvement in managing the tablet at the end, compared to the beginning, which is confirmed by a significant difference in the Wilcoxon-Test ( $z = -7.28$ ,  $p < 0.001$ ). The improvement calculated as pre-post difference was not significant between groups [VFG vs. EAFG:  $U = 112$ ,  $p = 0.118$ ; VFG vs. PAFG:  $U = 156$ ,  $p = 0.864$ ; VFG vs. CAFG:  $U = 128$ ,  $p = 0.293$ ].

#### **4.4. Discussion**

Data were analyzed in the tetherball paradigm to assess interpersonal coordination in terms of the mean absolute error, the cross correlation results, and participants' self-reports. In comparisons of VFG to three audio-visual groups, results revealed that the mean absolute error decreased significantly faster in EAFG and CAFG than VFG, but no statistical differences were observed between PAF and VFG. These results support the first hypothesis, H1(a) and H1(c) but not H1(b). Regarding the second hypothesis, only H2(c) is confirmed by the cross correlation results because a significant difference was observed

only between VFG and CAFG. In terms of the third hypothesis, most of the participants in all groups seemed to perceive that they affected each other during the joint task; however, no significant differences were observed between groups. Interestingly, participants in audio-visual groups felt pleased more than VFG during the main test. Although all participants seemed to feel the improvement in manipulating the tablet together, no significant differences were reported between VFG and audio-visual groups. Therefore, the third hypothesis (H3) is partially confirmed regarding a pleasant feeling during the joint task.

In the cyclic task, it is expected that the participants predicted their partner's actions and the joint outcomes. The results demonstrated that real-time audio-visual feedback supports the improvement of joint outcomes. According to Stein and Stanford (2008), the integration of auditory and visual information in multisensory areas of the central nervous system (CNS) usually improves one's perception. This phenomenon might lead participants to have better perception regarding movements of their own and partner's actions as well as joint outcomes, which can bring positive effects on the joint task. Additionally, previously referred literature (Schmidt and Richardson, 2008; Keller et al., 2014; Lang et al., 2016; Loehr and Vesper, 2016) underlines that rhythmical movement components are significantly related to interpersonal coordination. The rhythmic movement components during interpersonal coordination can decrease practice efforts and errors (Lang et al., 2016; Loehr and Vesper, 2016). When perceptual information of the rhythmical movement is shared by two or more individuals, spatiotemporal entrainment is enhanced by the same dynamic principles of the movement (Knoblich, Butterfill, and Sebanz, 2011; Phillips-Silver and Keller, 2012). Schmidt and Richardson (2008) also supported that additional perceptual information can improve the degree of action coupling, enabling interactants to align their movements. In the tetherball paradigm, sound of EAF was non-rhythmical but finally provides a periodic melody. This sound can aid the participants to predict their partner's and ball's dynamics, enhancing the precision of their own movement. After reaching the plateau, the mean absolute error in EAFG showed less standard deviation than in VFG. This possibly means that participants maintained their

joint performance well, after an audio-visual-motor network in the brain was established regarding a specific task.

In contrast, PAF failed to significantly affect the reduction of the mean absolute error and the improvement of the cross correlation, compared to VF. This suggests that different effects can be observed with different types of feedback on the joint task. One possible explanation is the effect of integration of auditory information with other perceptual modalities, which is called multisensory integration. Allerdissen et al. (2017) demonstrated that fencing experts showed almost the same pattern of results with both audio-visual and visual information. Demos et al. (2012) also suggested that the level of interpersonal coordination can be improved by audio-visual feedback, but can be decreased if auditory feedback is not task-related like music. Another possibility in the tetherball paradigm is that the types of sounds might influence the results. To be specific, EAF sound (“synthesized violin”) might be more dominant than PAF sound (“wind noise”), because EAF sound was a continuous sound with high audio frequency and bright timbre. Both sounds, nonetheless, were well-balanced and perceivable.

Another approach was made with CAFG by increasing the usual feedback in a motor learning setting. As reported by Effenberg et al., 2016, it was indicated that the motor performance is enhanced without conscious attention drawn to both types of auditory feedback which had been applied simultaneously. Also in the joint task of the tetherball paradigm, CAF confirmed a significant effect on interpersonal coordination, regarding not only the joint performance, but also the temporal synchronization. When it comes to the joint performance, the improvement might stem from the presence of EAF because PAF did not show any significant effect in the statistical analysis. However, in terms of the temporal synchronization, the positive effect might be supported by a synergy between EAF and PAF. In other words, although PAF-only and EAF-only condition failed to make an effect on interpersonal coordination, better joint performance was observed with a combination of PAF and EAF.

Participants’ self-report suggested that a joint task can be more pleasant with audio-visual feedbacks than only VF. The pleasant feeling with EAF might motivate the



collaboration, and cause significant improvement of the joint performance. Although PAF alone failed to have significant effect on physical coordination, it might significantly influence the psychological connectedness. This argument can be supported by Demos et al. (2012) who reported that participants felt connected, no matter whether they listen to task-related or non-task-related sound. In other words, the pleasant feelings might be related to the auditory component than to the task performance. For future research, it might be worth researching on how much participants' pleasant feeling during the joint task depends on different types of sound: task-related and non-task-related sound; or rhythmical and non-rhythmical sound. This approach would be in line with a research on the affective entrainment by Phillips-Silver and Keller (2012) who investigated the relationship between the pleasantness of participants and the task-relatedness of sound.

In interactions between humans and machines, auditory feedback can be applied in future. Humans move and react based on an ecological audio-motor network in brain which results from every-day experiences (Carello, Wagman, and Turvey 2005): When a car moves faster, the engine sounds louder. Such correlation between sound and movement are also observed when using a vacuum cleaner, a washing machine, and a printer, which might be regarded as PAF. On the other hand, humans also experience with ecological or artificial EAF: A car equipped with sensors can inform the driver of the distance to obstacles during parking by auditory feedback. A radar also sonifies the displacement and velocity of nearby objects. In this context, machines can transfer auditory information to humans. The tetherball paradigm represents the case of human-human interaction in a first step. The following studies might research on which types of sound are the best to improve human-machine interactions. It is important how the appropriate sound is chosen to map to physical performance and events. As shown with the tetherball paradigm, the results suggest that the joint performance between humans can benefit from EAF, at least, if the common goal is known already. In the case of interaction between human and humanoid robot, two agents might not share a common expected goal and perception because they did not experience other's movement behaviors before. For the novel interaction scenarios, PAF might be useful in a first step because both agents have to try to understand their

partners' movements initially. Although PAF did not show a significant effect here on human-human interaction, humanoid robot-human interaction might benefit more by initiating audio-visual-motion networks when initiating joint performance: According to Schmitz et al. (2013), appropriate auditory feedback from humanoid robot's movements might address biological motion perception mechanisms in the brain, even though these mechanisms are usually not addressed by non-human motions.

#### **4.5. Conclusion**

This research on auditory feedback in the motor domain is directed to the two categories: "Knowledge of performance" (KP) and "Knowledge of result" (KR). Within multisensory interplay, different types of auditory information can affect interpersonal coordination in different ways, which supports studies on the proximal sensorimotor contingency in socSMCs research project. For more observation of changing action-effect contingencies in different conditions, more variations of the joint task (e.g., sound synthesis, task period, task difficulty) can be of interest for the future studies. Furthermore, it is important to investigate the relationship between motor learning and the emergence of interpersonal coordination because both refer closely to the perception of kinematics, especially partners' or objects' movements (e.g., a sofa, a tetherball). Although research on both motor learning and interpersonal coordination is a challenging approach, for the future studies, continued efforts are needed by focusing on how to support the kinematic perception for improving individual and joint behavior in many different fields of research.

## Chapter 5. Discussion and Conclusion

In this dissertation, IMU-based basic measurement solutions are described to improve interpersonal coordination from Chapter 2 to Chapter 4. In Chapter 2, spatiotemporal gait parameters were measured using a single H-IMU. The single IMU solution can detect the exact time point of HS and TO. Measuring angular gait parameters are inherently impossible with a single IMU; however, the proposed solution has enabled real-time TO detection with head kinematic data for the first time, thereby increasing capability of single head-worn IMU solutions (e.g., earbuds, HMD) to measure spatiotemporal gait parameters. Furthermore, despite a small number of participants, the H-IMU shows reliability and robustness compared to pedometers, E-AR, and smartphones.

In Chapter 3, the simple H-IMU system monitors head motion and gait simultaneously when two individuals walk and talk. The proposed system enables an easy assessment of three interrelating components: *positivity*, *mutual attentiveness*, and *coordination*. These components vary during rapport development over time, and are interesting to researchers of interpersonal relationship (Tickle-Degnen and Rosenthal, 1990). Head pose estimation with an IMU is implemented, which facilitates the analysis of head nod and head orientation, indicating positivity and mutual attentiveness. In terms of coordination, the time correlation of interactants' gait events is also monitored in terms of the double support period, which is the time duration when both feet are on the ground. The proposed IMU method is, however, restricted to the molecular method, which measures rapport with discrete behaviors independently from social contexts. For example, the H-IMU system is able to only measure behavioral mirroring coordination, but not collaborative and interactive synchrony (Tickle-Degnen and Rosenthal, 1990). Nevertheless, the IMU solution provides structured and objective rapport components, and reduces measurement time and cost.

In Chapter 4, finally, a simple IMU solution is realized for a psychological study. An IMU based tablet (iPad Air) was used for a joint task, named "Tetherball paradigm." The study is about how different types of auditory feedback affect interpersonal

coordination. In the study, two participants performed the task with different auditory conditions. The joint outcomes and collaborative coordination were measured by the IMU-based tablet application. Each participant's feelings and experiences about interpersonal activity were also assessed by a questionnaire. The results depend on different types of auditory feedback with effect- and performance-based information. Interpersonal coordination was significantly influenced in terms of three categories: the joint performance (significantly influenced by CAF, and EAF), the temporal synchrony (by CAF), and pleasant feelings during the task (by CAF, EAF, and PAF). The tablet with an IMU monitored not only individual movements, but also the joint outcomes. Both types of data were used for the sound synthesis for auditory-motor feedback.

Each chapter of this dissertation includes new findings tightly related to human movement measurement using a single IMU solution, which can contribute to the analysis and enhancement of interpersonal coordination. As a part of the socSMCs project supported by EU, this work is framed with distal and proximal interpersonal coordination, considering group and dyad settings. In line with studies in the EU project, this work aims to contribute to the promotion of one's social well-being as well as public health, especially for individuals with social deficits. Human-robot interaction research may also benefit by collecting behavioral data during interaction between human and non-human agents, which remains as future works. General findings, contributions, and applications are discussed in the following sections.

## **5.1. Advantages and Comparisons**

### **5.1.1. H-IMU: Advantages in Measuring Large Groups**

The proposed H-IMU solution enables contribution to the measurement of the kinematics in group interactions. In this case, the H-IMU has a remarkable advantage in terms of cost. For example, when ten agents perform, the XSENS system (whole body mode) needs 170 sensors; however, H-IMU needs 10 sensors. During the measurement of football players, optical motion capture systems need a number of high resolution cameras,

which are fixed on the edges of the football field. These whole body motion capture systems can provide more kinematic data from a person, whereas H-IMU solutions can easily collect kinematic data from a larger number of people because it is very simple to set up the sensors on all participants. In terms of simplicity and cost, surveillance cameras and drones can be comparable to the H-IMU. Image processing technology enables cameras to measure gait and head orientation. Compared to the H-IMU, however, the camera systems need higher computational cost because they need additional computations to obtain human kinematic data, such as down sampling, image filtering, and facial feature detection. The additional computations, nevertheless, do not result in higher measurement resolution than the H-IMU. Including error canceling, XSENS MTw sensor used by this work provides higher sensitivity up to 0.50 dps (degree per second, °/s; Xsens MTw Awinda), which can be even higher with a gyroscope from ST Microelectronics (2013) that is embedded in smartphones and sensitive up to 0.00875 dps with raw data. An FHD or 4K camera can monitor multiple agents' interaction, but it might not provide gait information as much as H-IMU.

### **5.1.2. Tetherball Study: Advantage of IMU-Based Tablet Task**

Unlike methodological studies in other chapters, Chapter 4 describes a behavioral study with a novel intervention, called the “Tetherball paradigm.” This paradigm includes not only kinematic measurement, but also sensorimotor feedback using auditory information. The self-reports and the statistical analysis of 96 participants are included. The tetherball paradigm is designed to assess how different types of auditory information affect interpersonal coordination in a joint task. It is one of the simplest systems for measuring behaviors during collaborative interactions. Although a tablet PC equipped with an IMU is implemented, it can measure joint outcomes and synchronization simultaneously, which includes three advantages. First of all, joint outcomes are simply measured by a single IMU solution. In musical ensembles, the level of successful joint outcomes relies on co-performers' auditory perceptions of partner's performance, which is measured by electroencephalography (EEG) (Loehr et al., 2013). Secondly, whereas the previous

research has regarded the synchrony as joint outcomes (Vicaria and Dickens, 2016), joint outcomes and synchrony are measured independently in this study. To be specific, an effect-based and performance-based movement are measured for joint outcomes and synchrony, respectively. Thirdly, it is the first time that joint outcomes are measured in a collaborative interaction scenario while a pair of performers have continuous contact with an object. Previously, the joint outcomes were measured in the auditory or visual domain without physical contact. For example, joint outcomes have been assessed with the musical harmony played by two separated musicians (Loehr et al., 2013), as well as a visual synchrony between two separated dancers (Waterhouse et al., 2014).

Based on the separated measurement of both participants' movements, two types of auditory information can be generated: effect-based (EAF) and performance-based auditory feedback (PAF). Combined auditory feedback (CAF: EAF + PAF) influenced joint outcomes and interpersonal synchronization; however, EAF-only affects joint outcomes, and PAF-only did not significantly affect both outcomes and synchronization. To figure out the reason, it is required to test more with different sound in the PAF condition. For instance, the noise-based wind sound used for PAF must be replaced by the continuous violin sound (previously used in EAF). Another approach is changing the task levels. The joint task might not have been difficult enough to show a possible effect of PAF on interpersonal coordination. From participants' self-report, nevertheless, all auditory feedbacks significantly affect participants' pleasant feeling during the joint task, which is the affective component of interpersonal coordination (Phillips-Silver and Keller, 2012). In the future, the paradigm can be a reference for studies on the affective aspect of the entrainment, by comparing task related sound with non-task-related sound.

## **5.2. Tetherball Study: Comparison with Other Interpersonal Coordination Studies**

Despite the simplicity of the solution, the tetherball paradigm enables the measurement of a higher level of joint attention. For studies on interpersonal coordination with sensorimotor feedback, not only cameras, but also keyboards, buttons, and pressure sensors have been used (Waterhouse et al., 2014; Vesper and Richardson, 2014; Vesper et

al., 2013). Recently, motion capture systems and HMDs have been also used in virtual environment (Waltemate, 2018). These experiments consider the alignment of the onset and end of movements as well as the similarity of movement trajectories (e.g., arm, legs) during joint tasks or mirroring tasks. It was also considered how a participant predicts another participants' actions. With the motion capture systems and HMDs, participants' input and response of nonverbal actions are easily measured; however, most of the systems are designed for distal settings and the molecular level of measurement by analyzing discrete movements. Unlike these solutions, the tetherball paradigm is designed to measure interpersonal coordination in proximal settings because participants are asked to manipulate one object together. This paradigm also considers the molar level of measurement. For example, one participant should modify the reaction to the partner's action considering the ball's movement, even though the partner moves in the same pattern. This means participants should differently respond to the partner's input depending on the contextual information resulting from their joint actions. Their joint activities can be analyzed with information about the rate of turn of the tablet and velocity of the ball. The analysis of their input and response can be separately measured because a participant is allowed to move the tablet along only one axis. For professional measurement, the tetherball paradigm can be combined with camera systems or EEG systems, thereby monitoring participants' facial expressions or brain activities. For commercial uses, the apparatus that supports the tablet should be simplified.

### **5.3. Technical Review**

#### **5.3.1. H-IMU: System Optimizations for Spread of the Solution**

On the other hand, it is necessary to optimize for real applications, such as wireless earbuds or HMDs. For example, there are various wireless standards (e.g., Bluetooth, Wi-Fi, and Zig-Bee) and some devices have their own host controller units (e.g., ARM processors), which have different electrical properties and would require optimization of threshold values and cut off frequency in the proposed algorithm. In addition, since the study is performed with only one pair of participants, rapport analyses with more participants are needed for real intervention scenarios. For the future works, the system

also needs optimal settings of algorithm for different individuals considering age, height, and gender, to name a few.

### **5.3.2. H-IMU: Compatibility for more Analytical Parameters**

To develop a professional application with more Analytical parameters, H-IMU can be combined with other sensor systems, which can provide kinematic information in independent and interpersonal conditions. The H-IMU system is compatible with other motion sensors and cameras. For motion sensors, it has been proven in Chapter 2 where the kinematic data were obtained with the foot and pelvis IMUs at the same time. If knee and hip angles are measured by additional IMUs on the lower body, more gait parameters are obtained with optimal number of sensors.

A camera sensor has been embedded with an IMU to improve localization systems by compensating for IMU's accumulate errors (Hesch et al., 2014). Based on the improved positioning system, a combination of IMU sensors and cameras can be used for not only single agent, but also multi-agents. In addition, if cameras are placed in different position from the IMU, it can provide a different type of information. From the real time video recordings, facial expression, hand gestures, and contextual information, can be observed during interaction, which can enable more accurate estimation of interactants' rapport and emotion (Tickle-Degnen and Rosenthal, 1990).

Furthermore, other assessment methods can improve independent and interactive analysis with H-IMU. Machine learning algorithms such as hidden Markov model (HMM) can increase accuracy for different individuals referring to collected individual data. Results of participants' self-report and the judges' impressions can strengthen the combination of a molecular and a molar measurement.

No matter which system is combined with H-IMU, additional analytical parameters are supportive to overcome the limitation of single H-IMU solutions. On the other hand, analysis of head kinematics with H-IMU provides a great deal of information about gait patterns, walking direction, and social cues. The simplicity of H-IMU is also advantage to establish a combination with other sensors or systems. The proposed H-IMU solution also



can be easily combined with various health monitoring systems for one's physical and social activity.

### **5.3.3. Tetherball Study: Sound Systems**

As a real time sound synthesizer solution, Csound 6 runs well with Objective C for an iPad application. On the other hand, Csound 6 requires programming details of sound properties. Compared to the self-programming, it is easy to use a hardware synthesizer box, such as SonicCell (Rolland Corp.) that easily synthesizes various sound effects with several buttons and wheels. However, the SonicCell system causes more complicated experiment settings because of the additional device and wire connections. When a wireless solution is used to reduce the wire connections, sound delay can increase to around 116 ms (from video and audio analysis with Kdenlive 19.04.0). In this work, therefore, the iPad Air is responsible for both auditory and visual output with Csound 6 and Objective C, as a one-device solution. The tablet has enough computing power for physical modeling of the ball and the spring as well as synthesizing sound. The sound delay is about 83 ms (Kdenlive 19.04 and Audacity 2.3.1), by reducing Wi-Fi communication delay, which leads to the simple measurement solution for interpersonal coordination.

## **5.4. Future Applications**

Single IMU solutions enable the measurement of human behaviors during interpersonal coordination at low costs. Compared to motion capture systems and EEG, users have easy access to the devices with IMUs, which can facilitate studies on emotional and kinematic interactions, even outside the laboratory. Researchers of joint actions have instant results by using smart devices with IMUs, such as smartphones, tablets, smart watches, pedometers and earbuds in every-day life settings. The researchers can also collect more valuable data in participants' real life, compared to experimental laboratory settings. The single IMU solutions also facilitate both dyadic and group interpersonal coordination because they can be implemented with a bigger number of participants at

lower costs, compared to motion capture or EEG solutions. The single IMU solutions can directly support researchers and practitioners interested in one's physical and social health.

#### **5.4.1. Walking Intervention**

The walking intervention is currently performed with mainly two objectives: rehabilitation and disease prevention. First, elderly people and patients can relieve chronic and acute pains caused by musculo-skeletal problems of the knee, back, and hip (Ferrell, Josephson, Pollan, Loy, and Ferrell, 1997; Sitthipornvorakul, Klinsophon, Sihawong, and Janwantanakul, 2017). A walking intervention can also help to rehabilitate gait dysfunctions and support post-operative patient care. Secondly, World Health Organization (WHO) reported in 2018 that walking is one of key means to increase physical activity on a daily basis (see also Carr, Karvinen, Peavler, Smith, and Cangelosi, 2013). The increasing physical activity can reduce sedentary behaviors, and prevent lifestyle diseases, such as obesity, diabetes, hypertension, and heart disease, thereby promoting public health. Both types of walking interventions can be supported by technology. Pedometers and smartphones can aid in setting a daily goal and monitoring physical activity (Mansi, Milosavljevic, Baxter, Tumilty, and Hendrick, 2014). Therapists can manage patients via internet, which enables walking interventions at home, workplaces and neighborhood. With these advantages, H-IMU can additionally monitor the balance control, provide auditory-motor feedback, and implement virtual/augmented reality. These functions of H-IMU enable high quality interventions based on advanced gait analysis in everyday life settings. To sum up, H-IMU can contribute to public health and quality of life by increasing physical activity.

#### **5.4.2. Walk and Talk Intervention**

Psychological therapies with outdoor activities have been developed, such as nature therapy (Berger and McLeod, 2006), outdoor therapy (Revell, Duncan, and Cooper, 2014; Jordan, 2015), and adventure therapy (Gass, Gillis, and Russell, 2012). Walk and talk interventions are in line with outdoor-based psychological therapies. Walking itself

contributes to psychological health: enhancement of psychological processing (Hays, 1999), alleviation of depressive symptoms (Pickett, Yardley and Kendrick, 2012) as well as improvement of self-esteem and mood (Barton, Hine and Pretty, 2009). These effects result from increasing physical activities, which are encouraged by primary care settings of walk and talk interventions. In primary care settings, nurse practitioners can increase clients' physical activity by counseling the intervention. The practitioners help clients to set short- and long-term goals, as well as monitor their progress—comparable to pedometer-driven interventions. They can also provide clients' clear understanding about target behaviors of the therapy, based on the health behavioral theory. These therapies enable individualized interventions, which are usually more effective. Therefore, walk and talk intervention can be recommended in primary care settings due to its contribution to mental health through increased physical activity.

In addition, walk and talk interventions play an important role in psychological therapy as well. When walking outside, therapists and clients can create a casual and meditative atmosphere, which aids in opening clients' mind compared to a counselling office. Clients can, therefore, directly talk about their business, relationships, and family issues where they are suffering from anxiety, stress, and “stuckness”. For young children and adolescents, walk and talk therapy allows developmental counselling as well as psychotherapy. They can talk to teachers, parents or practitioners about their school works, socializing, relationships, as well as body changes. The youths can understand their own characteristics and society more, and learn coping strategies and experience interpersonal connections, which supports encouragement, self-efficacy, and prosocial behaviors (Doucette, 2004).

With rapport monitoring, walk and talk interventions can be more successful. A database of rapport building history of therapists can also improve the matching between a therapist-client dyad, depending on background, topics, and personalities. In addition, the development of social skills can be monitored by rapport changes over intervention periods for children, adolescents, and even autistic individuals. This H-IMU-based rapport monitoring solution is easily applicable to wireless smart devices, and currently many

therapies are developed with HMDs and the wearable devices, which are gradually familiar to people. Thus, walk and talk intervention with H-IMU solution can easily involve parents, teachers, and friends, with whom the participants prefer to freely talk in preferable places, such as at work places, homes, and neighborhoods.

#### **5.4.3. Social Intervention with Physical Collaboration Results**

The proposed tetherball paradigm can support a social intervention that promotes social health using a collaboration setting. Social health involves one's ability to maintain satisfying interpersonal relationships. To foster the ability, social intervention has been performed in collaborative settings, such as musical ensembles, group dance, team sports, and simple joint tasks. For everyday settings, moving a sofa or table together at home and division of labors in workplaces are also considered. During collaboration, individuals can foster multi-faceted interactive skills. While individuals interact with each other to pursue a common goal, they should adapt to the given situation and others' behaviors. For a successful collaboration, individuals have to communicate with each other by using verbal and nonverbal cues. Verbal cues are taken into account in turn-taking settings (e.g., dialog), whereas nonverbal cues play an important role in temporal synchronizing settings. This difference results from the response to verbal cues is not fast enough for emergent interpersonal coordination, whereas it is fast enough to nonverbal cues (Knoblich and Jordan, 2003). In tetherball paradigm, the nonverbal cues are focused, mainly including visual, auditory, and tactile cues. These cues establish the sensorimotor feedback loop and thereby influence the behaviors during collaboration. By steering a common object, two interactants can experience multimodal sensory feedback from one's own and the partner's movement.

This collaboration setting can be especially difficult for children with ASD because of the proximity (McConnell, 2002; Schleien, Heyne, and Berken, 1988), fine movement, (Lloyd, MacDonald, and Lord, 2013) and object sharing (Quilitch, and Risley, 1973). These joint tasks require collaborative socio-motor skills which need complicated cognitive brain function (McConnell, 2002; Schleien et al., 1988; Lloyd et al., 2013).

However, the proposed joint task can let children experience collaboration in peer to peer settings (DeKlyen and Odom, 1989; McConnell, 2002), including musical components (Kern and Aldridge, 2006). Although the tetherball paradigm has not yet been tested with people with social deficits, it can strongly support the improvement of autistic children's interactive skills.

Furthermore, the proposed IMU-based tablet solution can measure the joint outcomes and the synchronization in the tetherball task over experimental periods, which can provide objective references to the development of collaborative activity. Monitoring the development process can also support successful intervention by providing long-term feedback. In terms of real-time feedback, only auditory information is taken into account in tetherball paradigm; however, other modalities of sensorimotor feedback can be considered with other possible tablet solutions. In the joint task, moreover, both interactants' roles are divided into the movement referring to one of both orthogonal axes. The individual movement during the interaction can be independently analyzed. The tablet also provides the database related to effect- and performance-based movement information. Therefore, the proposed tablet solution is applicable to a collaborative intervention for developing one's interactive skill, as well as an analyzing tool of how participants manage the collaborative interaction. Therefore, this tablet solution may also help to diagnose autism by assessing one's ability to collaborate with others (McConnell, 2002; Schleien et al., 1988; Lloyd et al., 2013; Quilitch and Risley, 1973).

### **5.5. Concluding Remarks**

This dissertation has an intensive interdisciplinary character: Technological requirements in the fields of sensor technology and software engineering had to be fulfilled, in order to measure predefined social behaviors in two different experimental settings (distal, proximal). These requirements on the reference frame were regarded as a major challenge in the development of the experimental methods and settings. For the future studies, the proposed IMU-based kinematic measurement solutions are needed to be applied to real consumer electronics, such as wireless earbuds, HMDs, smart phones, or

tablet PCs. More people can, therefore, use the IMU solution in various physical and social settings in daily life, such as training, therapies and interventions. During walk and talk interventions, therapists or counselors can provide clients with feedback based on physical, mental, and linguistic aspects. These interventions are available for all ages and even people with a broad range of social deficits. In addition, a large set of data with more participants and different conditions are needed to appropriate applications. In terms of the tetherball paradigm, social deficits of the participants (e.g., ASD) can be assessed with the tasks requiring a high level of interpersonal skills in correlated environments with proximity, fine movement, and object sharing. In a next step, the development of enhanced interactive abilities could be supported by different types of additional auditory feedback. On the other hand, robots also could be trained with joint tasks and accumulate data related to human interactive behaviors. From the research on the coordination of the head and other body parts, developers of humanoid robots might properly design mechanical systems for the human-like movement. In addition, well-trained robots could be used to provide social therapies for people with social deficits in the future. Although the IMU-based solutions cannot measure the whole body kinematics, they can provide effective and efficient approaches for the research on human-human and human-robot interaction. The single IMU solution would be more powerful in a combination with other sensors, such as camera, EEG, and another IMU at other places. Overall, single IMU-based solutions lay an interesting and efficient groundwork for future research on technology-driven socio-motor interventions for public health and quality of life.

## List of Publications

### As the first author:

- Hwang, T.-H., Reh, J., Effenberg, A. O., & Blume, H. (2016, Sep.). Real-time gait event detection using a single head-worn inertial measurement unit. In *IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)* (pp. 28–32). IEEE. Berlin, Germany.
- Hwang, T.-H., Schmitz, G., Klemmt, K., Brinkop, L., Ghai, S., Stoica, M., Maye, A., Blume, H., & Effenberg, A. O. (2018). Effect- and performance-based auditory feedback on interpersonal coordination. *Frontiers in Psychology*, 9:404. doi: 10.3389/fpsyg.2018.00404.
- Hwang, T.-H., Reh, J., Effenberg, A. O., & Blume, H. (2018a). Real-Time Gait Analysis Using a Single Head-Worn Inertial Measurement Unit. *IEEE Transactions on Consumer Electronics*, 64(2). 240–248. doi:10.1109/TCE.2018.2843289.
- Hwang, T.-H., Gerd, S., Ghai, S., & Effenberg, A. O. (2018b). Additional auditory information alters interpersonal coordination. In *Arbeitsgemeinschaft für Sportpsychologie (ASP) 2018*, Cologne, Germany.
- Hwang, T.-H., Effenberg, A. O., & Blume, H. A. (2019a, Jan.). Rapport and Gait Monitoring System Using a Single Head-Worn IMU During Walk and Talk. In *IEEE International Conference on Consumer Electronics (ICCE 2019)* (pp. 1–5). IEEE. Las Vegas, NV.
- Hwang, T.-H., Reh, J., Effenberg, A. O., & Blume, H. (2019b, July). Validation of Real Time Gait Analysis Using a Single Head-Worn IMU. In *Proceedings of the 2019 Europe-Korea Conference on Science and Technology* (pp. 87–97). Springer.

### As one of authors:

- Reh, J., Hwang, T.-H., Michalke, V., & Effenberg, A. O. (2016, September). Instruction and real-time sonification for gait rehabilitation after unilateral hip arthroplasty. In *Human Movement and Technology: Book of Abstracts, Proceedings of the 11th Joint Conference on Motor Control & Learning, Biomechanics & Training*.

- Reh, J., Hwang, T.-H., Schmitz, G., & Effenberg, A. O. (2019). Dual Mode Gait Sonification for Rehabilitation after unilateral Hip Arthroplasty. *Brain Sciences*, 9(3), 66.
- Ghai, S., Schmitz, G., Hwang, T.-H., & Effenberg, A. O. (2017). Effects of real time auditory feedback on proprioception. In *22nd Annual Congress of the European College of Sport Science*, Essen, Germany.
- Ghai, S., Schmitz, G., Hwang, T.-H. & Effenberg, A. O. (2018). Auditory Proprioceptive Integration: Effects of Real-Time Kinematic Auditory Feedback on Knee Proprioception. *Frontiers in Neuroscience*, 12:142. doi: 10.3389/fnins.2018.00142.
- Schmitz, S., Bergmann, J., Effenberg, A. O., Krewer, C., Hwang, T.-H. and Mueller, F. (2018). Movement sonification in stroke rehabilitation. *Frontiers in Neurology*. 9:389. doi: 10.3389/fneur.2018.00389.
- Effenberg, A. O., Hwang, T.-H., Ghai, S., & Gerd, S. (2018). Auditory Modulation of Multisensory Representations. In *International Symposium on Computer Music Multidisciplinary Research* (pp. 284–311). Springer, Cham.
- Ghai, S., Schmitz, G., Hwang, T.-H., & Effenberg, A. O. (2019). Training proprioception with sound: effects of real-time auditory feedback on intermodal learning. *Annals of the New York Academy of Sciences*, 1438(1), 50–61.



## Bibliography

- Allerdissen, M., Güldenpenning, I., Schack, T., & Bläsing, B. (2017). Recognizing fencing attacks from auditory and visual information: a comparison between expert fencers and novices. *Psychology of Sport and Exercise*, *31*, 123–130. doi: 10.1016/j.psychsport.2017.04.009.
- American Psychiatric Association. (2013). *Autism Spectrum Disorder. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®)* (pp. 50–59). American Psychiatric Pub.
- Apple iPad Air (1st generation). Available online: <http://www.apple.com/ipad/compare>
- Asperger, H. (1944). Die „Autistischen Psychopathen“ im Kindesalter. *Archiv für Psychiatrie und Nervenkrankheiten*, *117*(1), 76–136.
- Aszkler, C. (2005, May). The Principle of Acceleration, Shock, & Vibration Sensors, *Sensors online*, Available: <https://www.sensorsmag.com/components/principles-acceleration-shock-and-vibration-sensors/>
- Atallah, L., Aziz, O., Lo, B., & Yang, G. Z. (2009, June). Detecting walking gait impairment with an ear-worn sensor. In *2009 Sixth International Workshop on Wearable and Implantable Body Sensor Networks* (pp. 175–180). IEEE.
- Bachmann, E. R. (2000). *Inertial and Magnetic Tracking of Limb Segment Orientation for Inserting Humans into Synthetic Environments*. Naval Postgraduate School Monterey CA.
- Badino, L., D'Ausilio, A., Glowinski, D., Camurri, A., & Fadiga, L. (2014). Sensorimotor Communication in professional quartets. *Neuropsychologia*, *55*, 98–104.
- Bailey, G. P., & Harle, R. K. (2015). Measuring Temporal Parameters of Gait with Foot Mounted IMUs in Steady State Running. In *icSPORTS* (pp. 24–33).
- Baker, R. (2007). The history of gait analysis before the advent of modern computers. *Gait & Posture*, *26*(3), 331–342.
- Bangert, M., Peschel, T., Schlaug, G., Rotte, M., Drescher, D., Hinrichs, H., Heinzed, HJ. Altenmüller, E. (2006). Shared networks for auditory and motor processing in professional pianists: evidence from fMRI conjunction. *Neuroimage* *30*, 917–926. doi: 10.1016/j.neuroimage.2005.10.044.
- Baron-Cohen, S. (2004). The cognitive neuroscience of autism. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*, 945–948. doi: 10.1136/jnnp.2003.018713.
- Barton, J., Hine, R., & Pretty, J. (2009). The health benefits of walking in greenspaces of high natural and heritage value. *Journal of Integrative Environmental Sciences*, *6*(4), 261–278.

- Baum, W. M. (2004). Molar and molecular views of choice. *Behavioural Processes*, 66(3), 349–359.
- Bellusci, G., Roetenberg, D., Dijkstra, F., Luinge, H., & Slycke, P. (2010). Xsens MVN Motiongrid: Drift-free human motion tracking using tightly coupled ultra-wideband and miniature inertial sensors. *Xsens Technologies White Paper*.
- Bente, G., Krämer, N. C., Petersen, A., & de Ruiter, J. P. (2001). Computer animated movement and person perception: Methodological advances in nonverbal behavior research. *Journal of Nonverbal Behavior*, 25(3), 151–166.
- Bernstein, J., Cho, S., King, A. T., Kourepenis, A., Maciel, P., & Weinberg, M. (1993, February). A micromachined comb-drive tuning fork rate gyroscope. In *Proceedings of the 1993 IEEE Micro Electro Mechanical Systems* (pp. 143–148). IEEE.
- Bialkowski, S. E. (1988a). Real Time Digital Filters: Finite Impulse-Response Filters. *Analytical Chemistry*, 60(5), 355A–361A. doi: 10.1021/ac00156a743.
- Bialkowski, S. E. (1988b). Real-time Digital Filters: Infinite Impulse Response Filters. *Analytical Chemistry*, 60(6), 403A–413A. doi: 10.1021/ac00157a003.
- Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on ‘Theory of Mind’ and cognition. *Brain*, 127(4), 914–928.
- Bortz, J. (2005). *Statistik für Sozialwissenschaftler*. Springer, Berlin.
- Boulangier, R. (2000). *The Csound Book: Perspectives in Software Synthesis, Ssound Design, Signal Processing, and Programming*. MIT press.
- Brajdic, A., & Harle, R. (2013, September). Walk detection and step counting on unconstrained smartphones. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (pp. 225–234). ACM.
- Breazeal, C., & Scassellati, B. (2002). Robots that imitate humans. *Trends in Cognitive Sciences*, 6(11), 481–487.
- Bridenbaugh, S. A., & Kressig, R. W. (2011). Laboratory review: the role of gait analysis in seniors’ mobility and fall prevention. *Gerontology*, 57(3), 256–264.
- Brock, H., Schmitz, G., Baumann, J., & Effenberg, A. O. (2012). If motion sounds: Movement sonification based on inertial sensor data. *Procedia Engineering*, 34, 556–561.
- Brodie, M., Walmsley, A., & Page, W. (2008). Fusion motion capture: a prototype system using inertial measurement units and GPS for the biomechanical analysis of ski racing. *Sports Technology*, 1(1), 17–28.
- Brown, S., Martinez, M. J., & Parsons, L. M. (2006). The neural basis of human dance. *Cerebral Cortex*, 16(8), 1157–1167.

- Bruckner, H. P., Theimer, W., & Blume, H. (2014, January). Real-time low latency movement sonification in stroke rehabilitation based on a mobile platform. In *2014 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 264–265). IEEE.
- Brunelli, R., & Poggio, T. (1993). Face recognition: Features versus templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(10), 1042–1052.
- Butzin, N. C., Hochendoner, P., Ogle, C. T., Hill, P., & Mather, W. (2015). Marching along to an offbeat drum: Entrainment of synthetic gene oscillators by a noisy stimulus. *ACS Synthetic Biology*, *5*(2), 146–153.
- Cabibihan, J. J., Javed, H., Ang, M., & Aljunied, S. M. (2013). Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. *International Journal of Social Robotics*, *5*(4), 593–618.
- Carello, C., Wagman, J. B., & Turvey, M. T. (2005). Acoustic specification of object properties. In J. D. Anderson and B. F. Anderson (Eds.), *Moving Image Theory: Ecological Considerations*. Carbondale (pp. 79–104), IL: Southern Illinois University Press.
- Carr, L. J., Karvinen, K., Peavler, M., Smith, R., & Cangelosi, K. (2013). Multicomponent intervention to reduce daily sedentary time: a randomised controlled trial. *BMJ Open*, *3*(10), e003261.
- Carruthers, P. (1996). 16 Autism as Mind-Blindness: an elaboration and partial defence. *Theories of Theories of Mind*, 257.
- Chen, Q., Wu, H., Fukumoto, T., & Yachida, M. (1998, April). 3D head pose estimation without feature tracking. In *Proceedings 1998 Third IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 88–93). IEEE.
- Chang, S., Lee, J., Choe, S. K., & Lee, K. (2017). Audio Cover Song Identification using Convolutional Neural Network. *arXiv Preprint arXiv:1712.00166*.
- Clark, H. H. (1996). *Using language*. Cambridge university press.
- Clark, S., & Iltis, P. W. (2008). Effects of dynamic head tilts on sensory organization test performance: a comparison between college-age athletes and nonathletes. *Journal of Orthopaedic & Sports Physical Therapy*, *38*(5), 262–268.
- Clayton, M., Sager, R., & Will, U. (2005, January). In time with the music: the concept of entrainment and its significance for ethnomusicology. In *European Meetings in Ethnomusicology*. (Vol. 11, pp. 1–82). Romanian Society for Ethnomusicology.
- Collier, G. L., & Logan, G. (2000). Modality differences in short-term memory for rhythms. *Memory & Cognition*, *28*(4), 529–538. doi: 10.3758/BF03201243.
- Cornejo, C., Cuadros, Z., Morales, R., & Paredes, J. (2017). Interpersonal coordination: Methods, achievements, and challenges. *Frontiers in Psychology*, *8*, 1685.

- Cromwell, R., & Wellmon, R. (2001). Sagittal plane head stabilization during level walking and ambulation on stairs. *Physiotherapy Research International*, 6(3), 179–192.
- D'Ausilio, A., Badino, L., Li, Y., Tokay, S., Craighero, L., Canto, R., Aloimonos, Y., & Fadiga, L. (2012). Leadership in orchestra emerges from the causal relationships of movement kinematics. *PLoS One*, 7(5), e35757.
- Darling, K., Nandy, P., & Breazeal, C. (2015, August). Empathic concern and the effect of stories in human-robot interaction. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 770–775). IEEE.
- Davis, C. S. (2002). *Statistical Methods for the Analysis of Repeated Measurements*. Springer Science & Business Media.
- Davis III, R. B., Ounpuu, S., Tyburski, D., & Gage, J. R. (1991). A gait analysis data collection and reduction technique. *Human Movement Science*, 10(5), 575–587.
- DeKlyen, M., & Odom, S. L. (1989). Activity structure and social interactions with peers in developmentally integrated play groups. *Journal of Early Intervention*, 13(4), 342–352.
- Demos, A. P., Carter, D. J., Wanderley, M. M., & Palmer, C. (2017). The Unresponsive Partner: roles of social status, auditory feedback, and animacy in coordination of joint music performance. *Frontiers in Psychology*. 8:149. doi: 10.3389/fpsyg.2017.00149.
- Demos, A. P., Chaffin, R., Begosh, K. T., Daniels, J. R., & Marsh, K. L. (2012). Rocking to the beat: effects of music and partner's movements on spontaneous interpersonal coordination. *Journal of Experimental Psychology: General*. 141, 49–53. doi: 10.1037/a0023843.
- De Nunzio, A. M., Nardone, A., & Schieppati, M. (2005). Head stabilization on a continuously oscillating platform: the effect of a proprioceptive disturbance on the balancing strategy. *Experimental Brain Research*, 165(2), 261–272.
- Dixon, T. H. (1991). An introduction to the Global Positioning System and some geological applications. *Reviews of Geophysics*, 29(2), 249–276.
- Doucette, P. A. (2004). Walk and talk: An intervention for behaviorally challenged youths. *Adolescence*, 39(154), 373.
- Duarte, R., Araújo, D., Correia, V., Davids, K., Marques, P., & Richardson, M. J. (2013). Competing together: Assessing the dynamics of team–team and player–team synchrony in professional association football. *Human Movement Science*, 32(4), 555–566.
- Effenberg, A. O. (2005). Movement sonification: effects on perception and action. *IEEE Multimedia* 12(2), 53–59. doi: 10.1109/MMUL.2005.31.
- Effenberg, A. O., Fehse, U., Schmitz, G., Krueger, B., & Mechling, H. (2016). Movement sonification: effects on motor learning beyond rhythmic adjustments. *Frontiers in Neuroscience*. 10:219. doi: 10.3389/fnins.2016.00219.

- Effenberg, A. O., Hwang, T.-H., Ghai, S., & Gerd, S. (2018a). Auditory Modulation of Multisensory Representations. In *International Symposium on Computer Music Multidisciplinary Research* (pp. 284–311). Springer, Cham.
- Effenberg, A. O. & Schmitz, G. (2018b). Acceleration and deceleration at constant speed: systematic modulation of motion perception by kinematic sonification. *Annals of the New York Academy of Sciences*, *1425*, 52–69. doi:org/10.1111/nyas.13693.
- Ellamil, M., Berson, J., & Margulies, D. S. (2016). Influences on and measures of unintentional group synchrony. *Frontiers in Psychology*, *7*, 1744.
- Ellamil, M., Berson, J., Wong, J., Buckley, L., & Margulies, D. S. (2016). One in the dance: musical correlates of group synchrony in a real-world club environment. *PloS One*, *11*(10).
- Emmerich, H., & Schofthaler, M. (2000). Magnetic field measurements with a novel surface micromachined magnetic-field sensor. *IEEE Transactions on Electron Devices*, *47*(5), 972–977.
- Fang, Y., Nakashima, R., Matsumiya, K., Kuriki, I., & Shioiri, S. (2015). Eye-head coordination for visual cognitive processing. *PloS One*, *10*(3), e0121035.
- Feese, S., Arnrich, B., Tröster, G., Meyer, B., & Jonas, K. (2012, September). Quantifying behavioral mimicry by automatic detection of nonverbal cues from body motion. In *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Conference on Social Computing* (pp. 520–525). IEEE.
- Feldman, R. (2007). Parent–infant synchrony and the construction of shared timing; physiological precursors, developmental outcomes, and risk conditions. *Journal of Child Psychology and Psychiatry*, *48*(3-4), 329–354.
- Ferrell, B. A., Josephson, K. R., Pollan, A. M., Loy, S., & Ferrell, B. R. (1997). A randomized trial of walking versus physical methods for chronic pain management. *Aging Clinical and Experimental Research*, *9*(1–2), 99–105. doi: 10.1007/BF03340134.
- Fiske, A. P. (1992). The four elementary forms of sociality: framework for a unified theory of social relations. *Psychological Review*, *99*(4), 689.
- Fritz, S., & Lusardi, M. (2009). White paper: “walking speed: the sixth vital sign”. *Journal of Geriatric Physical Therapy*, *32*(2), 2–5.
- Freedman, E. G., & Sparks, D. L. (2000). Coordination of the eyes and head: movement kinematics. *Experimental Brain Research*, *131*(1), 22–32.
- Frith, U. (2001). Mind blindness and the brain in autism. *Neuron*, *32*(6), 969–979.
- Foxlin, E. (1996, March). Inertial head-tracker sensor fusion by a complementary separate-bias Kalman filter. In *Proceedings of the 1996 IEEE Virtual Reality Annual International Symposium* (pp. 185–194). IEEE.

- Gallagher, H. L., & Frith, C. D. (2003). Functional imaging of ‘theory of mind’. *Trends in Cognitive Sciences*, 7(2), 77–83.
- Gass, M. A., Gillis, L., & Russell, K. C. (2012). *Adventure Therapy: Theory, Research, and Practice*. Routledge.
- Ghai, S., Ghai, I., Schmitz, G., & Effenberg, A. O. (2018). Effect of rhythmic auditory cueing on parkinsonian gait: A systematic review and meta-analysis. *Scientific Reports*, 8(1), 506.
- Goebel, W., & Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Perception: An Interdisciplinary Journal*, 26(5), 427–438. doi: 10.1525/mp.2009.26.5.427.
- Gordon, A., & Browne, K. (2013). *Beginnings & Beyond: Foundations in Early Childhood Education*. Cengage Learning.
- Grasso, R., Prévost, P., Ivanenko, Y. P., & Berthoz, A. (1998). Eye-head coordination for the steering of locomotion in humans: an anticipatory synergy. *Neuroscience Letters*, 253(2), 115–118.
- Harris, T., Kerry, S. M., Victor, C. R., Ekelund, U., Woodcock, A., Iliffe, S., et al. (2015). A primary care nurse-delivered walking intervention in older adults: PACE (pedometer accelerometer consultation evaluation)-Lift cluster randomised controlled trial. *PLoS Medicine*, 12(2), e1001783.
- Hays, K. F. (1999). *Working it out: Using Exercise in Psychotherapy*. American Psychological Association.
- Heinz, E., Kunze, K. S., Gruber, M., Bannach, D., & Lukowicz, P. (2006, May). Using wearable sensors for real-time recognition tasks in games of martial arts-an initial experiment. In *2006 IEEE Symposium on Computational Intelligence and Games* (pp. 98–102). IEEE.
- Heinzmann, J., & Zelinsky, A. (1998, April). 3-D facial pose and gaze point estimation using a robust real-time tracking paradigm. In *Proceedings of the 1998 Third IEEE International Conference on Automatic Face and Gesture Recognition* (pp. 142–147). IEEE.
- Hellmers, H., Norrdine, A., Blankenbach, J., & Eichhorn, A. (2013, October). An IMU/magnetometer-based Indoor positioning system using Kalman filtering. In *the 2013 International Conference on Indoor Positioning and Indoor Navigation (IPIN)* (pp. 1–9). IEEE. doi: 10.1109/IPIN.2013.6817887.
- Hesch, J. A., Kottas, D. G., Bowman, S. L., & Roumeliotis, S. I. (2014). Camera-IMU-based localization: Observability analysis and consistency improvement. *The International Journal of Robotics Research*, 33(1), 182–201.

- Hinton, P. R., McMurray, I., & Brownlow, C. (2014). *SPSS Explained*. Routledge.
- Hirasaki, E., Kubo, T., Nozawa, S., Matano, S., & Matsunaga, T. (1993). Analysis of head and body movements of elderly people during locomotion. *Acta Oto-Laryngologica*, *113*(sup501), 25–30.
- Howell, D., Osternig, L., & Chou, L. S. (2015). Monitoring recovery of gait balance control following concussion using an accelerometer. *Journal of Biomechanics*, *48*(12), 3364–3368.
- Hunn, N. (2016). The market for hearable devices 2016–2020. *Technical Report November*, WiFore Consulting, London, UK.
- Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, *1*(1), 69–82.
- Hwang, T.-H., Effenberg, A. O., & Blume, H. A. (2019a, Jan). Rapport and Gait Monitoring System Using a Single Head-Worn IMU During Walk and Talk. In *2019 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1–5). IEEE. Las Vegas, NV.
- Hwang T.-H., Schmitz G., Klemmt K., Brinkop L., Ghai S., Stoica M., Maye A., Blume H., & Effenberg A. O. (2018a). Effect- and Performance-Based Auditory Feedback on Interpersonal Coordination. *Frontiers in Psychology*. *9*:404.
- Hwang, T.-H., Reh, J., Effenberg, A., & Blume, H. (2016, September). Real-time gait event detection using a single head-worn inertial measurement unit. In *2016 IEEE 6th International Conference on Consumer Electronics-Berlin (ICCE-Berlin)* (pp. 28–32). IEEE. Berlin, Germany.
- Hwang, T.-H., Reh, J., Effenberg, A. O., & Blume, H. (2018b). Real-Time Gait Analysis Using a Single Head-Worn Inertial Measurement Unit. *IEEE Transactions on Consumer Electronics*. *64*(2), 240–248. doi:10.1109/TCE.2018.2843289.
- Hwang, T.-H., Reh, J., Effenberg, A. O., & Blume, H. (2019b, July). Validation of Real Time Gait Analysis Using a Single Head-Worn IMU. In *Proceedings of the 2019 Europe-Korea Conference on Science and Technology* (pp. 87–97). Springer.
- Imai, T., Moore, S. T., Raphan, T., & Cohen, B. (2001). Interaction of the body, head, and eyes during walking and turning. *Experimental Brain Research*, *136*(1), 1–18.
- Inchley, J., Cuthbert, L., & Grimes, M. (2007). An Investigation of the use of Pedometers to Promote Physical Activity and Particularly Walking Among School-Aged Children. in *Report to Paths to Health and the Scottish Health Promoting Schools Unit*.
- ITU-T, R. G. (1997). Series G: Transmission systems and media digital systems and networks. *Ginebra, Suiza*. ITY.

- Jarchi, D., Wong, C., Kwasnicki, R. M., Heller, B., Tew, G. A., & Yang, G. Z. (2014). Gait parameter estimation from a miniaturized ear-worn sensor using singular spectrum analysis and longest common subsequence. *IEEE Transactions on Biomedical Engineering*, *61*(4), 1261–1273.
- Jordan, M. (2015). *Nature and Therapy: Understanding Counselling and Psychotherapy in Outdoor Spaces*. Routledge.
- Junker, H., Amft, O., Lukowicz, P., & Tröster, G. (2008). Gesture spotting with body-worn inertial sensors to detect user activities. *Pattern Recognition*, *41*(6), 2010–2024.
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, *2*(3), 217–250.
- Kapoor, A., & Picard, R. W. (2001, November). A real-time head nod and shake detector. In *Proceedings of the 2001 Workshop on Perceptive User Interfaces* (pp. 1–5). ACM.
- Kavanagh, J., Barrett, R., & Morrison, S. (2006). The role of the neck and trunk in facilitating head stability during walking. *Experimental Brain Research*, *172*(4), 454.
- Kavanagh, J. J., Morrison, S., & Barrett, R. S. (2005). Coordination of head and trunk accelerations during walking. *European Journal of Applied Physiology*, *94*(4), 468–475.
- Keller, P. E. (2012). Mental imagery in music performance: underlying mechanisms and potential benefits. *Annals of the New York Academy of Sciences*, *1252*, 206–213. doi: 10.1111/j.1749-6632.2011.06439.x.
- Keller, P. E., & Appel, M. (2010). Individual differences, auditory imagery, and the coordination of body movements and sounds in musical ensembles. *Music Perception*, *28*, 27–46. doi: 10.1525/mp.2010.28.1.27.
- Keller, P. E., Novembre, G., & Hove, M. J. (2014). Rhythm in joint action: psychological and neurophysiological mechanisms for real-time interpersonal coordination. *Philosophical Transactions of Royal Society B: Biological Sciences*, *369*:20130394. doi: 10.1098/rstb.2013.0394.
- Kelso, J. S. (1995). *Dynamic Patterns: The Self-organization of Brain and Behavior*. Cambridge, MA: MIT press.
- Kennel, C., Pizzera, A., Hohmann, T., Schubotz, R. I., Murgia, M., Agostini, T., et al. (2014). The perception of natural and modulated movement sounds. *Perception* *43*, 796–804. doi: 10.1068/p7643.
- Kertesz, A. E. (1994). *Localization and neuroimaging in neuropsychology*. Academic Press.
- Kern, P., & Aldridge, D. (2006). Using embedded music therapy interventions to support outdoor play of young children with autism in an inclusive community-based child care program. *Journal of Music Therapy*, *43*(4), 270–294.



- Kirk, A. G., O'Brien, J. F., & Forsyth, D. A. (2005, June). Skeletal parameter estimation from optical motion capture data. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 2, pp. 782–788). IEEE. San Diego, CA.
- Knoblich, G., Butterfill, S., & Sebanz, N. (2011). Psychological research on joint action: theory and data. In B. Ross (Eds.), *Psychology of Learning and Motivation*, (Vol. 54, pp. 59–101). Burlington: Academic Press.
- Knoblich, G., & Jordan, J. S. (2003). Action coordination in groups and individuals: learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 29(5), 1006–1016. doi: 10.1037/0278-7393.29.5.1006.
- Kotiadis, D., Hermens, H. J., & Veltink, P. H. (2010). Inertial Gait Phase Detection for control of a drop foot stimulator: Inertial sensing for gait phase detection. *Medical Engineering & Physics*, 32(4), 287–297.
- Kudrle, S., Proulx, M., Carrieres, P., & Lopez, M. (2010, October). Fingerprinting for solving A/V synchronization issues within broadcast environments. In *2010 SMPTE Annual Tech Conference & Expo* (pp. 1–18). SMPTE.
- Laghi, F., Lonigro, A., Levanto, S., Ferraro, M., Baumgartner, E., & Baiocco, R. (2016). The role of nice and nasty theory of mind in teacher-selected peer models for adolescents with autism spectrum disorders. *Measurement and Evaluation in Counseling and Development*, 49(3), 207–216.
- Land, M. F. (1999). Motion and vision: why animals move their eyes. *Journal of Comparative Physiology A*, 185(4), 341–352.
- Lang, M., Shaw, D. J., Reddish, P., Wallot, S., Mitkidis, P., and Xygalatas, D. (2016). Lost in the rhythm: effects of rhythm on subsequent interpersonal coordination. *Cognitive Science*, 40(7), 1797–1815. doi: 10.1111/cogs.12302.
- Langfelder, G., & Tocchio, A. (2014). MEMS integrating motion and displacement sensors. In *Smart Sensors and Mems* (pp. 366–401). Woodhead Publishing.
- Lane, D. (2016). The assumption of sphericity in repeated-measures designs: What it means and what to do when it is violated. *Quantitative Methods for Psychology*, 12, 114–122.
- Ledger, D., & McCaffrey, D. (2014). Inside wearables: How the science of human behavior change offers the secret to long-term engagement. *Endeavour Partners*, 200(93), 1.
- Levene, H (1960). Robust tests for equality of variances. In Olkin I., et al. (Eds.) *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling* (pp. 278–292). Stanford University Press, Palo Alto, CA.
- Li, W., & Wang, J. (2013). Effective adaptive Kalman filter for MEMS-IMU/magnetometers integrated attitude and heading reference systems. *The Journal of Navigation*, 66(1), 99–113.

- Liang, Y., Xiaowei, L., Weiping, C., & Zhiping, Z. (2011). High resolution interface circuit for closed-loop accelerometer. *Journal of Semiconductors*, 32(4), 045005.
- Lickel, B., Hamilton, D. L., Wierzchowska, G., Lewis, A., Sherman, S. J., & Uhles, A. N. (2000). Varieties of groups and the perception of group entitativity. *Journal of Personality and Social Psychology*, 78(2), 223.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Llobera, J., Charbonnier, C., Chagué, S., Preissmann, D., Antonietti, J. P., Ansermet, F., & Magistretti, P. J. (2016). The subjective sensation of synchrony: an experimental study. *PLoS One*, 11(2), e0147008.
- Lloyd, M., MacDonald, M., & Lord, C. (2013). Motor skills of toddlers with autism spectrum disorders. *Autism*, 17(2), 133–146.
- Loehr, J. D., Kourtis, D., Vesper, C., Sebanz, N., & Knoblich, G. (2013). Monitoring individual and joint action outcomes in duet music performance. *Journal of Cognitive Neuroscience*, 25(7), 1049–1061. doi: 10.1162/jocn\_a\_00388.
- Loehr, J. D., & Vesper, C. (2016). The sound of you and me: novices represent shared goals in joint action. *Quarterly Journal of Experimental Psychology*. 69(3), 535–547. doi: 10.1080/17470218.2015.1061029.
- Magill, R. A., & Anderson, D. I. (2007). *Motor Learning and Control: Concepts and Applications, 11*. New York, NY: McGraw-Hill.
- Majumder, A. J. A., Saxena, P., & Ahamed, S. I. (2016, June). Your walk is my command: gait detection on unconstrained smartphone using IoT system. In *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)* (Vol. 1, pp. 798–806). IEEE.
- Mariani, B., Hoskovec, C., Rochat, S., Büla, C., Penders, J., & Aminian, K. (2010). 3D gait assessment in young and elderly subjects using foot-worn inertial sensors. *Journal of Biomechanics*, 43(15), 2999–3006.
- Marsh, K. L., Richardson, M. J., & Schmidt, R. C. (2009). Social connection through joint action interpersonal coordination. *Topics in Cognitive Science*, 1(2), 320–339. doi: 10.1111/j.1756-8765.2009.01022.x.
- Mann, H. B., & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 50–60.
- Mansi, S., Milosavljevic, S., Baxter, G. D., Tumilty, S., & Hendrick, P. (2014). A systematic review of studies using pedometers as an intervention for musculoskeletal diseases. *BMC Musculoskeletal Disorders*, 15(1), 231.
- Maurer, T., & von der Malsburg, C. (1996, October). Tracking and learning graphs and pose on image sequences of faces. In *Proceedings of the 1996 Second International Conference on Automatic Face and Gesture Recognition* (pp. 176–181). IEEE.

- McConnell, S. R. (2002). Interventions to facilitate social interaction for young children with autism: Review of available research and recommendations for educational intervention and future research. *Journal of Autism and Developmental Disorders*, 32(5), 351–372.
- McNeill, W. H. (1995). *Keeping Together in Time: Dance and Drill in Human History*. Harvard University Press.
- Mehta, U. M., Bhagyavathi, H. D., Kumar, C. N., Thirthalli, J., & Gangadhar, B. N. (2014). Cognitive deconstruction of parenting in schizophrenia: the role of theory of mind. *Australian & New Zealand Journal of Psychiatry*, 48(3), 249–258.
- Meltzoff, A. N., & Decety, J. (2003). What imitation tells us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 358(1431), 491–500.
- Moeslund, T. B., Hilton, A., & Krüger, V. (2006). A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104(2–3), 90–126.
- Munoz, M. J. (2013). *Preference for Online Communication and its Effect on Perceived Social Skills and Academic Performance* (Doctoral dissertation, California State University, Northridge).
- Murgia, M., Hohmann, T., Galmonte, A., Raab, M., & Agostini, T. (2012). Recognising one's own motor actions through sound: the role of temporal factors. *Perception* 41(8), 976–987. doi: 10.1068/p7227.
- Murgia, M., Prpic, V., O, J., McCullagh, P., Santoro, I., Galmonte, A., et al. (2017). Modality and perceptual-motor experience influence the detection of temporal deviations in tap dance sequences. *Frontiers in Psychology*, 8:1340. doi: 10.3389/fpsyg.2017.01340.
- Muro-De-La-Herran, A., Garcia-Zapirain, B., & Mendez-Zorrilla, A. (2014). Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications. *Sensors*, 14(2), 3362–3394.
- Murray, K., Lillakas, L., Weber, R., Moore, S., & Irving, E. (2007). Development of head movement propensity in 4–15 year old children in response to visual step stimuli. *Experimental Brain Research*, 177(1), 15–20.
- Muthugala, M. A. V. J., Munasinghe, M. K. B. S., Lakshan, M. H. C., Madurangi, L. H. H., & Jayasekara, A. G. B. P. (2013, December). Design of an interactive robotic head with human-like movements. In *2013 8th IEEE International Conference on Industrial and Information Systems (ICIIS)* (pp. 355–360). IEEE.
- Myers, S. M., & Johnson, C. P. (2007). Management of children with autism spectrum disorders. *Pediatrics*, 120(5), 1162–1182.

- Nakashima, R., Fang, Y., Hatori, Y., Hiratani, A., Matsumiya, K., Kuriki, I., & Shioiri, S. (2015). Saliency-based gaze prediction based on head direction. *Vision Research*, *117*, 59–66.
- Nasiri, S. (2009). A critical review of MEMS gyroscopes technology and commercialization status. Available online: [http://blog.ncue.edu.tw/sys/lib/read\\_attach.php?id=17430](http://blog.ncue.edu.tw/sys/lib/read_attach.php?id=17430)
- Néda, Z., Ravasz, E., Brechet, Y., Vicsek, T., & Barabási, A. L. (2000). Self-organizing processes: The sound of many hands clapping. *Nature*, *403*(6772), 849–850.
- Nilsson, J., Stokes, V. P., & Thorstensson, A. (1985). A new method to measure foot contact. *Journal of Biomechanics*, *18*(8), 625–627.
- Nunes, M. E., Souza, M. G. T., Basso, L., Monteiro, C. B., Corrêa, U. C., & Santos, S. (2014). Frequency of provision of knowledge of performance on skill acquisition in older persons. *Frontiers in Psychology*, *5*:1454. doi: 10.3389/fpsyg.2014.01454.
- Obradovic, D., Lenz, H., & Schupfner, M. (2004, September). Sensor fusion in Siemens car navigation system. In *Proceedings of the 2004 14th IEEE Signal Processing Society Workshop Machine Learning for Signal Processing* (pp. 655–664). IEEE.
- Oppenheim, A. V., Buck, J. R., & Schafer, R. W. (2001). *Discrete-Time Signal Processing*. Upper Saddle River, NJ: Prentice Hall.
- Paloski, W. H., Wood, S. J., Feiveson, A. H., Black, F. O., Hwang, E. Y., & Reschke, M. F. (2006). Destabilization of human balance control by static and dynamic head tilts. *Gait & Posture*, *23*(3), 315–323.
- Parsey, C. M., & Schmitter-Edgecombe, M. (2013). Applications of technology in neuropsychological assessment. *The Clinical Neuropsychologist*, *27*(8), 1328–1361.
- Pfau, T., Ferrari, M., Parsons, K., & Wilson, A. (2008). A hidden Markov model-based stride segmentation technique applied to equine inertial sensor trunk movement data. *Journal of Biomechanics*, *41*(1), 216–220.
- Pfister, A., West, A. M., Bronner, S., & Noah, J. A. (2014). Comparative abilities of Microsoft Kinect and Vicon 3D motion capture for gait analysis. *Journal of Medical Engineering & Technology*, *38*(5), 274–280.
- Phillips-Silver, J., & Keller, P. (2012). Searching for roots of entrainment and joint action in early musical interactions. *Frontiers in Human Neuroscience*, *6*:26. doi: 10.3389/fnhum.2012.00026.
- Pickett, K., Yardley, L., & Kendrick, T. (2012). Physical activity and depression: A multiple mediation analysis. *Mental Health and Physical Activity*, *5*(2), 125–134.
- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*(2), B51–B61.
- Port, R. F. (1981). Linguistic timing factors in combination. *The Journal of the Acoustical Society of America*, *69*(1), 262–274.

- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind?. *Behavioral and Brain Sciences*, 1(4), 515–526.
- Proakis, J. G., & Manolakis, D. G. (1996). *Digital Signal Processing: Principles, Algorithms, and Applications*. Prentice Hall, Upper Saddle River, NJ.
- Qu, H., Yu, H., Zhou, W., Peng, B., Peng, P., & He, X. (2016). Effects of dielectric charging on the output voltage of a capacitive accelerometer. *Journal of Micromechanics and Microengineering*, 26(11), 115001.
- Quilitch, H. R., & Risley, T. R. (1973). The Effects of Play Materials on Social Play 1. *Journal of Applied Behavior Analysis*, 6(4), 573–578.
- Rapport, M. D., Chung, K. M., Shore, G., Denney, C. B., & Isaacs, P. (2000). Upgrading the science and technology of assessment and diagnosis: Laboratory and clinic-based assessment of children with ADHD. *Journal of Clinical Child Psychology*, 29(4), 555–568.
- Reh, J., Hwang, T. H., Michalke, V., & Effenberg, A. O. (2016, September). Instruction and real-time sonification for gait rehabilitation after unilateral hip arthroplasty. In *Human Movement and Technology: Book of Abstracts, Proceedings of the 11th Joint Conference on Motor Control & Learning, Biomechanics & Training*.
- Reh, J., Hwang, T.-H., Schmitz, G., & Effenberg, A. O. (2019). Dual Mode Gait Sonification for Rehabilitation after unilateral Hip Arthroplasty. *Brain Sciences*, 9(3), 66.
- Repp, B. H., & Keller, P. E. (2004). Adaptation to tempo changes in sensorimotor synchronization: effects of intention, attention, and awareness. *Quarterly Journal of Experimental Psychology Section A*, 57(3), 499–521. doi: 10.1080/02724980343000369.
- Repp, B. H., & Penel, A. (2002). Auditory dominance in temporal processing: new evidence from synchronization with simultaneous visual and auditory sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 28(5), 1085–1099. doi: 10.1037/0096-1523.28.5.1085.
- Repp, B. H., & Su, Y. H. (2013). Sensorimotor synchronization: a review of recent research (2006–2012). *Psychonomic Bulletin & Review*, 20(3), 403–452. doi: 10.3758/s13423-012-0371-2.
- Revell, S., Duncan, E., & Cooper, M. (2014). Helpful aspects of outdoor therapy experiences: An online preliminary investigation. *Counselling and Psychotherapy Research*, 14(4), 281–287.
- Rhudy, M., Bucci, B., Viperman, J., Allanach, J., & Abraham, B. (2009, January). Microphone array analysis methods using cross-correlations. In *ASME 2009 International Mechanical Engineering Congress and Exposition* (pp. 281–288). American Society of Mechanical Engineers.

- Richardson, M., Garcia, R. L., Frank, T. D., Gregor, M., & Marsh, K. L. (2012). Measuring group synchrony: a cluster-phase method for analyzing multivariate movement time-series. *Frontiers in Physiology*, *3*, 405.
- Rios, J. A., & White, E. (2002). Fusion filter algorithm enhancements for a MEMS GPS/IMU. *Crossbow Technology, Inc*, 1–12.
- Robinson, H., MacDonald, B., Kerse, N., & Broadbent, E. (2013). The psychosocial effects of a companion robot: a randomized controlled trial. *Journal of the American Medical Directors Association*, *14*(9), 661–667.
- Roetenberg, D., Luinge, H., & Slycke, P. (2009). Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors. *Xsens Motion Technologies BV, Tech. Rep. 1*.
- Rowley, H. A., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(1), 23–38.
- Rueterbories, J., Spaich, E. G., Larsen, B., & Andersen, O. K. (2010). Methods for gait event detection and analysis in ambulatory systems. *Medical Engineering & Physics*, *32*(6), 545–552.
- Rye, J. A., Zizzi, S. J., Vitullo, E. A., & Tompkins, N. O. H. (2005). The pedometer as a tool to enrich science learning in a public health context. *Journal of Science Education and Technology*, *14*(5–6), 521–531.
- Sabatini, A. M., Martelloni, C., Scapellato, S., & Cavallo, F. (2005). Assessment of walking features from foot inertial sensing. *IEEE Transactions on Biomedical Engineering*, *52*(3), 486–494.
- Savelsbergh, G. J., Van der Kamp, J., Williams, A. M., & Ward, P. (2005). Anticipation and visual search behaviour in expert soccer goalkeepers. *Ergonomics*, *48*(11–14), 1686–1697.
- Scassellati, B. (2002). Theory of mind for a humanoid robot. *Autonomous Robots*, *12*(1), 13–24.
- Schleien, S. J., Heyne, L. A., & Berken, S. B. (1988). Integrating physical education to teach appropriate play skills to learners with autism: A pilot study. *Adapted Physical Activity Quarterly*, *5*(3), 182–192.
- Schleien, S. J., Mustonen, T., & Rynders, J. E. (1995). Participation of children with autism and nondisabled peers in a cooperatively structured community art program. *Journal of Autism and Developmental Disorders*, *25*(4), 397–413.
- Schmitz, G., & Effenberg, A. O. (2017). Coxswain 2.0 – movement–acoustic dimensions of interpersonal coordination in team sports [Schlagmann 2.0 - Bewegungsakustische Dimensionen interpersoneller Koordination im Mannschaftssport.] *German Journal of Exercise and Sport Research*, *47*(3), 232–245. doi: 10.1007/s12662-017-0442-7

- Schmitz, G., Kroeger, D., & Effenberg, A. O. (June, 2014). A mobile sonification system for stroke rehabilitation. In *The 20th International Conference on Auditory Display*.
- Schmitz, G., Mohammadi, B., Hammer, A., Heldmann, M., Samii, A., Münte, T. F., & Effenberg, A. O. (2013). Observation of sonified movements engages a basal ganglia frontocortical network. *BMC Neuroscience*, *14*(1), 32.
- Schmidt, R. C., & Richardson, M. J. (2008). Dynamics of interpersonal coordination. In A. Fuchs and V. K. Jirsa (Eds.) *Understanding Complex Systems* (pp. 281–308). Springer, Berlin, Heidelberg.
- Schmidt, R., & Wrisberg, C. (2008). *Motor Learning and Performance: A Situation-based Learning Approach*, 4th (Eds.). Human Kinetics, Champaign, IL.
- Seeger, J., Lim, M., & Nasiri, S. (2010, June). Development of high-performance, high-volume consumer MEMS gyroscopes. In *Solid-State Sensors, Actuators, and Microsystems Workshop* (pp. 61–64).
- Sharma, D. A., Chevidikunnan, M. F., Khan, F. R., & Gaowgzeh, R. A. (2016). Effectiveness of knowledge of result and knowledge of performance in the learning of a skilled motor activity by healthy young adults. *Journal of Physical Therapy Science*, *28*(5), 1482–1486. doi: 10.1589/jpts.28.1482.
- Shockley, K., & Riley, M. A. (2015). Interpersonal couplings in human interactions. In *Recurrence Quantification Analysis* (pp. 399-421). Springer, Cham.
- Silaghi, M. C., Plänklers, R., Boulic, R., Fua, P., & Thalmann, D. (1998). Local and global skeleton fitting techniques for optical motion capture. In *Modelling and Motion Capture Techniques for Virtual Environments* (pp. 26–40). Springer, Berlin, Heidelberg.
- Sitthipornvorakul, E., Klinsophon, T., Sihawong, R., & Janwantanakul, P. (2017). The effects of walking intervention in patients with chronic low back pain: A meta-analysis of randomized controlled trials. *Musculoskeletal Science and Practice*, *34*, 38–46.
- Springer, S., Gottlieb, U., & Lozin, M. (2016). Spatiotemporal gait parameters as predictors of lower-limb overuse injuries in military training. *The Scientific World Journal*, 2016.
- Stanton, C., Bogdanovych, A., & Ratanasena, E. (2012, December). Teleoperation of a humanoid robot using full-body motion capture, example movements, and machine learning. In *Proceedings of Australasian Conference on Robotics and Automation* (pp. 3–5). Victoria University of Wellington New Zealand.
- Stein, B. E., & Stanford, T. R. (2008). Multisensory integration: current issues from the perspective of the single neuron. *Nature Reviews Neuroscience*, *9*(4), 255–266. doi: 10.1038/nrn2331.

- ST Microelectronics (2013), MEMS motion sensor: three-axis digital output gyroscope. L3GD20H datasheet, Mar. 2013 [DocID023469 Rev 2]. Available: <https://www.st.com/resource/en/datasheet/l3gd20h.pdf>
- Studenski, S., Perera, S., Patel, K., Rosano, C., Faulkner, K., Inzitari, M., Brach, J., Chandler, J., Cauthon, P., Conner, E. B., & Nevitt, M. (2011). Gait speed and survival in older adults. *Jama*, *305*(1), 50–58.
- Taborri, J., Palermo, E., Rossi, S., & Cappa, P. (2016). Gait partitioning methods: A systematic review. *Sensors*, *16*(1), 66.
- Thaut, M. H., Tian, B., & Azimi-Sadjadi, M. R. (1998). Rhythmic finger tapping to cosine-wave modulated metronome sequences: Evidence of subliminal entrainment. *Human movement science*, *17*(6), 839–863.
- Thurlow, W. R., & Runge, P. S. (1967). Effect of induced head movements on localization of direction of sounds. *The Journal of the Acoustical Society of America*, *42*(2), 480–488.
- Tickle-Degnen, L., & Rosenthal, R. (1990). The nature of rapport and its nonverbal correlates. *Psychological Inquiry*, *1*(4), 285–293.
- Troje, N. F. (2002). Decomposing biological motion: A framework for analysis and synthesis of human gait patterns. *Journal of Vision*, *2*(5), 2.
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: Effects of structural and kinematic cues. *Perception & Psychophysics*, *67*(4), 667–675.
- Tukey, J. W. (1949). Comparing individual means in the analysis of variance. *Biometrics*, *5*(2), 99–114.
- Vercher, J. L., Magenes, G., Prablanc, C., & Gauthier, G. M. (1994). Eye-head-hand coordination in pointing at visual targets: spatial and temporal analysis. *Experimental Brain Research*, *99*(3), 507–523.
- Vesper, C., van der Wel, R. P., Knoblich, G., & Sebanz, N. (2013). Are you ready to jump? Predictive mechanisms in interpersonal coordination. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 48–61. doi: 10.1037/a0028066.
- Vesper, C., & Richardson, M. J. (2014). Strategic communication and behavioral coupling in asymmetric joint action. *Experimental Brain Research*, *232*(9), 2945–2956.
- Vicaria, I. M., & Dickens, L. (2016). Meta-analyses of the intra-and interpersonal outcomes of interpersonal coordination. *Journal of Nonverbal Behavior*, *40*(4), 335–361. doi: 10.1007/s10919-016-0238-8.
- Vicary, S., Sperling, M., von Zimmermann, J., Richardson, D. C., & Orgs, G. (2017). Joint action aesthetics. *PLoS One*, *12*:e0180101. doi: 10.1371/journal.pone.0180101.



- Volpe, G., D'Ausilio, A., Badino, L., Camurri, A., & Fadiga, L. (2016). Measuring social interaction in music ensembles. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1693), 20150377.
- Waltemate, T. (2018). *Creating a Virtual Mirror for Motor Learning in Virtual Reality* (Doctoral dissertation, Universität Bielefeld, Bielefeld, Germany). doi:10.4119/unibi/2932705.
- Wandt, B., Ackermann, H., & Rosenhahn, B. (2016). 3d reconstruction of human motion from monocular image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(8), 1505–1516.
- Wandt, B., & Rosenhahn, B. (2019). Repnet: Weakly supervised training of an adversarial reprojection network for 3d human pose estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 7782–7791).
- Wang, E., Lignos, C., Vatsal, A., & Scassellati, B. (2006, March). Effects of head movement on perceptions of humanoid robot behavior. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction* (pp. 180–185). ACM.
- Wang, H., Lenz, H., Szabo, A., Bamberger, J., & Hanebeck, U. D. (2007, March). WLAN-based pedestrian tracking using particle filters and low-cost MEMS sensors. In *2007 4th Workshop on Positioning, Navigation and Communication* (pp. 1–7). IEEE.
- Waterhouse, E., Watts, R., & Bläsing, B. E. (2014). Doing duo—a case study of entrainment in William Forsythe's choreography “Duo”. *Frontiers in Human Neuroscience*, 8:812. doi: 10.3389/fnhum.2014.00812.
- Weeks, D. L., & Kordus, R. N. (1998). Relative frequency of knowledge of performance and motor skill learning. *Research Quarterly for Exercise and Sport*, 69(3), 224–230.
- Wilcoxon, F. (1992). Individual comparisons by ranking methods. In *Breakthroughs in Statistics* (pp. 196–202). Springer, New York, NY.
- Wing, A. M., Endo, S., Bradbury, A., & Vorberg, D. (2014). Optimal feedback correction in string quartet synchronization. *Journal of The Royal Society Interface*, 11(93), 20131125.
- Winstein, C. J. (1991). Knowledge of results and motor learning—implications for physical therapy. *Physical Therapy*, 71(2), 140–149. doi: 10.1093/ptj/71.2.140.
- Wood, G., & Wilson, M. R. (2010). Gaze behaviour and shooting strategies in football penalty kicks: Implications of a "keeper-dependent" approach. *International Journal of Sport Psychology*, 41(3), 293.
- World Health Organization. (2018). *Global Action Plan on Physical Activity 2018–2030: more Active People for a Healthier World*. Available online: <http://apps.who.int/iris/bitstream/handle/10665/272722/9789241514187-eng.pdf>

- Xie, H., & Fedder, G. K. (2000, January). A CMOS z-axis capacitive accelerometer with comb-finger sensing. In *Proceedings IEEE Thirteenth Annual International Conference on Micro Electro Mechanical Systems (Cat. No. 00CH36308)* (pp. 496–501). IEEE.
- Xsens MVN Biomech. Available online: <https://www.xsens.com/products/mvn-biomech/>
- Xsens MTw Awinda, “Wireless motion tracker.” Available online: <https://www.xsens.com/download/pdf/documentation/mtw-awinda/mtw-awinda.pdf>
- Yamada, M., Aoyama, T., Mori, S., Nishiguchi, S., Okamoto, K., Ito, T., Muto, S, Ishihara, T., Yoshitomi, H., & Ito, H. (2012). Objective assessment of abnormal gait in patients with rheumatoid arthritis using a smartphone. *Rheumatology International*, 32(12), 3869–3874.
- Zhu, S., Anderson, H., & Wang, Y. (2012, December). Reducing the power consumption of an imu-based gait measurement system. In *Pacific-Rim Conference on Multimedia* (pp. 105–116). Springer, Berlin, Heidelberg.

# Tonghun Hwang

Address: Leibniz Universität Hannover, Institut für Sportwissenschaft, Am Moritzwinkel 6, 30167 Hannover  
Tel: +49-511-762 17290 (Office), +49-176-3705 1088 (Mobile), Email: tonghun.hwang@sportwiss.uni-hannover.de

## EDUCATION & RESEARCH

---

- |                     |  |                           |
|---------------------|--|---------------------------|
| Aug 2015 –          | <b>Leibniz University Hannover</b>   | <b>Hannover, Germany</b>  |
|                     | <ul style="list-style-type: none"><li>◆ PhD student in the Institute of Microelectronic Systems (Academic Advisor: Prof. Dr.-Ing. Holger Blume)</li><li>◆ Research assistant in the Institute of Sports Science (Academic Advisor: Prof. Dr. Alfred Effenberg)</li></ul> |                           |
| Mar 2002 – Aug 2010 | <b>Hanyang University</b>  | <b>Seoul, South Korea</b> |
|                     | <ul style="list-style-type: none"><li>◆ M.S. (2010) in the Department of Electronics and Computer Engineering (Academic Advisor: Prof. Oh-Kyong Kwon)</li><li>◆ B.S. (2008) in the Department of Electrical and Computer Engineering</li></ul>                           |                           |

## PROFESIONAL EXPERIENCE

---

- |                     |  |                           |
|---------------------|--|---------------------------|
| Aug 2010 – Aug 2013 | <b>Samsung Electronics</b>   | <b>Seoul, South Korea</b> |
|                     | <ul style="list-style-type: none"><li>◆ Camera software engineer for Android platform<ul style="list-style-type: none"><li>• AP: Exynos 3470, 4412, 5250, 5410.</li><li>• Mobile product: Samsung Galaxy S3/S4, Google Nexus 10</li></ul></li><li>◆ Linux camera drivers for Android platform<ul style="list-style-type: none"><li>• Camera drivers for embedded ISPs and image sensors (5MP, 8MP, 13MP)</li><li>• MIPI-CSI drivers for camera sensor interfaces for Android systems</li></ul></li><li>◆ Camera hardware attraction layer (HAL) for Android platform</li><li>◆ Technical customer service<ul style="list-style-type: none"><li>• Samsung Electronics Mobile (Exynos5410, Galaxy S4, 13MP)</li><li>• Samsung Semiconductor China R&amp;D (Exynos5250, 5MP)</li></ul></li><li>◆ Manager of Android build server.</li></ul> |                           |

## RESEARCH PROJECTS

---

- |                     |   |                           |
|---------------------|---|---------------------------|
| Aug 2015 –          | <b>Leibniz University Hannover</b>  | <b>Hannover, Germany</b>  |
|                     | <ul style="list-style-type: none"><li>◆ “Development Disabilities of School Skills” supported by Federal Ministry of Education and Research</li><li>◆ “Socializing Sensorimotor Contingencies (socSMCs)” supported by the European Union’s Horizon 2020 Research</li></ul>  |                           |
| Mar 2008 – Aug 2010 | <b>Hanyang University</b>   | <b>Seoul, South Korea</b> |
|                     | <ul style="list-style-type: none"><li>◆ “Design Driving IP for Capacitive Touch Screen Panel” supported by the Electronics and Telecommunications Research Institute (ETRI)</li><li>◆ “Design Driver ICs and Modeling Devices for Oxide Thin Film Transistor (TFT) for Transparent Display” supported by ETRI</li></ul> |                           |

## TEACHING EXPERIENCE

---

	<b>Leibniz University Hannover</b>	<b>Hannover, Germany</b>
Oct 2019 –	♦ Seminar: “Kinematic assessment and human motion”	
Apr 2016 – Jul 2016	♦ Seminar: “Kinematic assessment and human movement”	
Apr 2016 – Jul 2016	♦ Joint seminar: “Bewegungslehre Schwerpunkt Bewegung und Kognition B: Koordinative und kognitive Leistungsfähigkeit”	

## EXTRACURRICULAR ACTIVITIES

---

Sep 2016 – Sep 2017	<b>President of Korean Badminton Club Hannover</b>	
Sep 2006 – Feb 2008	<b>Volunteer at the Korea Support Center for Foreign Workers</b>	
	♦ Computer teacher for Asian foreign workers	
Mar 2005 – Feb 2008	<b>Inventions and Demonstrations in BARAMI Club</b>	
	♦ “3D motion capture system for human arm using accelerometers”	
	♦ “Laser pointer to control a mouse cursor on a beam screen using camera”	
	♦ “Display system to inscribe fonts in the air”	

## PUBLICATIONS: JOURNAL

---

May 2018	<b>“Real-Time Gait Analysis Using a Single Head-Worn Inertial Measurement Unit”</b> <i>IEEE Transactions on Consumer Electronics</i> , 62(2), pp.240–248.
Mar 2018	<b>“Effect- and performance-based auditory feedback on interpersonal coordination”</b> <i>Frontiers in Psychology</i> , 9:404.
Mar 2011	<b>“Inverters Using Only N-Type Indium Gallium Zinc Oxide Thin Film Transistors for Flat Panel Display Applications”</b> <i>Japanese Journal of Applied Physics</i> , 50, 03CB06.
May 2010	<b>“A Highly Area-Efficient Controller for Capacitive Touch Screen Panel System”</b> <i>IEEE Transactions on Consumer Electronics</i> , 56(2), pp.1115–1122.

## PUBLICATIONS: CONFERENCE

---

Jan 2019	<b>“Rapport and Gait Monitoring System Using a Single Head-Worn IMU During Walk and Talk”</b> <i>IEEE International Conference on Consumer Electronics (ICCE 2019)</i> , pp. 1–5
Oct 2009	<b>“A Level Shifter Using Aluminum-Doped Zinc Tin Oxide Thin Film Transistors with Negative Threshold Voltages”</b> <i>IMID 2009 Digest</i> , pp. 464-465.
May 2009	<b>“A Scan Driver Circuit Using Transparent Thin Film Transistors”</b> <i>SID Symposium Digest of Technical Papers</i> , 40(1), pp. 1136–1139.

## LANGUAGE

---

Korean, English

## TOOLS

---

Java, Swift, Objective C, C/C++, Python, VBA, HSPICE, PSPICE, MATLAB