

Article

Real-Time and Embedded Detection of Hand Gestures with an IMU-Based Glove [†]

Chaithanya Kumar Mummadı ¹, Frederic Philips Peter Leo ¹, Keshav Deep Verma ¹,
Shivaji Kasireddy ¹, Philipp M. Scholl ¹, Jochen Kempfle ² and Kristof Van Laerhoven ^{2,*} 

¹ Faculty of Engineering, University of Freiburg, 79085 Freiburg im Breisgau, Germany;
chaithu0536@gmail.com (C.K.M.); frederic.peter@merkur.uni-freiburg.de (F.P.P.L.);
keshavdeepverma@gmail.com (K.D.V.); shivaji.kasireddy@venus.uni-freiburg.de (S.K.);
pscholl@ese.uni-freiburg.de (P.M.S.)

² Ubiquitous Computing, University of Siegen, 57068 Siegen, Germany; Jochen.Kempfle@uni-siegen.de

* Correspondence: kvl@eti.uni-siegen.de; Tel.: +49-271-740-2312

† This article builds upon an original conference paper [1], targeting the use of the IMU-based glove for generic gestures, and including additional experiments regarding the fusion of IMU data.

Received: 28 February 2018; Accepted: 6 June 2018; Published: 11 June 2018

Abstract: This article focuses on the use of data gloves for human-computer interaction concepts, where external sensors cannot always fully observe the user's hand. A good concept hereby allows to intuitively switch the interaction context on demand by using different hand gestures. The recognition of various, possibly complex hand gestures, however, introduces unintentional overhead to the system. Consequently, we present a data glove prototype comprising a glove-embedded gesture classifier utilizing data from Inertial Measurement Units (IMUs) in the fingertips. In an extensive set of experiments with 57 participants, our system was tested with 22 hand gestures, all taken from the French Sign Language (LSF) alphabet. Results show that our system is capable of detecting the LSF alphabet with a mean accuracy score of 92% and an F1 score of 91%, using complementary filter with a gyroscope-to-accelerometer ratio of 93%. Our approach has also been compared to the local fusion algorithm on an IMU motion sensor, showing faster settling times and less delays after gesture changes. Real-time performance of the recognition is shown to occur within 63 milliseconds, allowing fluent use of the gestures via Bluetooth-connected systems.

Keywords: gesture recognition; data gloves; inertial sensing; hand articulation tracking

1. Introduction

A large number of systems that rely on hand gestures as input technologies for wearable computers were proposed during the past two decades of research. The advantages of these gestures include that they are easy to learn and enable the adoption of common gestures used in everyday life or existing alphabets for sign languages, relying on a significant user base. However, as these gestures tend to use the full articulation of the hand, requiring the detection of the exact position and motion of all fingers, such systems are not straightforward to implement. Systems that are placed in the environment tend to suffer from occlusion and only cover a selected region. Previously-proposed wearable approaches on the other hand, have used reduced alphabets, less-accurate but small-enough sensors, and mostly rely on offline processing of the sensor readings on other 'back end' devices.

A large amount of data gloves and glove-based systems have been proposed for interaction methods, though most of these have been focusing on sensors that measure bending [2–4]. These sensors are generally straightforward to integrate in fabric, but offer less precision as they measure changes in the bending of the fingers, and suffer from hysteresis effects and accumulating measurement errors [5–7]. Few systems have thus far utilized fully integrated sensor chips that contain

tri-axial accelerometers, gyroscopes, and magnetometers on all fingers, while none have been designed or evaluated for distinguishing different hand articulations on the glove itself in real-time.

In this article, we introduce a novel sensing glove design (Figure 1) that detects its wearer's hand postures and motion at the fingertips with high accuracy, while keeping a small size with respect to the amount and size of hardware components. Our proposed system is designed to handle the data from the fingertip-located inertial measurement units (IMUs) as efficiently as possible to achieve real-time behaviour. The system is evaluated by detecting all hand shapes from the French sign language alphabet and a background class. Unlike other glove-based systems, this detection is performed *on a minimal embedded sensing system in the glove itself*, allowing it to be used as a stand-alone gesture recognition system, and with *a high accuracy (overall 92% accuracy for 22 finger-signs)* when dealing with large collections of signs to be recognized. We furthermore illustrate possible applications by connecting the glove via Bluetooth to a smartphone to immediately display the glove's detected gestures.



Figure 1. This article's self-contained data glove is capable of detecting and distinguishing between fine-grained signs in real-time through a set of miniature IMU units, one per finger. Detection is performed on an integrated system placed in the back of the glove, powered from a flat Li-Po battery integrated under the wrist. Shown are four different gestures (out of the 22 we consider in the evaluation) including those that for most gloves are not easily distinguishable from each other.

The remainder of this article is structured as follows: After situating our work within similar past research in data glove systems, the system design choices we made in both hardware and software are discussed with particular emphasis on the sensor data fusion. Furthermore, the design is evaluated for accuracy and speed of the hand shape detection on 57 study participants, as well as a more in-depth look into the fusion of the inertial data, and subsequently the results are discussed. After an illustrative example how our system works, the final chapter concludes with our main findings in this work and an outlook on future research.

2. Related Work

Glove-based systems have been popular for more than three decades and are already being used in numerous applications (e.g., [8] or [2]). Research into augmented gloves can be broadly divided into vision-based and sensor-based approaches. Most vision-based systems require external infrastructure, rendering them unfit for mobile application. The first sensor gloves were primarily used as input devices, though haptic gloves later were introduced as well that provide the user with vibration-based feedback. In this article, the focus lies on using the glove as a pure input device, capable of recognizing any short gestures [9] that the user might perform locally. The earliest sensor glove was developed in 1977 by de Fanti and Sandin [3], based on flexible tubes with a light source and photocell. Finger movements modulated the amount of light passing to the photocell causing a change in voltage. By the end of the decade, advancements in camera-based approach that can track the LEDs placed on the hand were designed at the MIT Media Lab [3].

The idea of sensing the inflection of fingers then was continued by Zimmerman et al. [4] in 1987, who commercialized the first data glove. Five to fifteen resistive flex sensors measured each finger and provided a novel interface to PCs. Kuroda et al., 2004 [10] proposed the StrinGlove equipped with 24 inductcoders (sensing displacement based on the principle of magnetic induction) and 9 contact sensors on the fingertips of both hands. Calibration is performed automatically after obtaining minimum and maximum ranges of values from the sensor readings. Around 48 finger characters were examined in this work and the accuracy was found to be about 85%. Khambaty et al. [11] had presented Ges-TALK in 2008—a glove to capture and process the finger movements to distinguish 24 sign language static gestures and to output the corresponding voice of the respective gesture. An assembly of 11 resistive bend sensors were setup to gather the readings. Here, a potentiometer is used to calibrate the system for every new user. Results showed an accuracy of 90% with a response rate of 750 ms based on template matching along with statistical pattern recognition. In 2011, Huang et al. [12] used a 5DT (Dimensional Technologies) data glove with 5 flex sensors for each finger. They proposed a concept grounding method using a data fusion approach called Clustering Ensembles and multi-class classifiers [13] for the gesture recognition. A total of 32 gestures were chosen by limiting to two states for each finger and a classification performance was achieved with 72.2% accuracy.

Different materials allowed better integration into gloves, and sign language applications were targeted increasingly to evaluate such systems. In 2011, Jeong et al. [14] made a gesture recognition glove with 'Velostat' material, which can measure the finger inflection through changes in resistance. The glove also embedded a 5-Degrees of Freedom (DOF) IMU to measure the motion angle for American Sign Language (ASL) and Korean Sign Language. Voltage differences are recorded for the finger movements and used to recognize the gestures. The sign language letters were recognized using the 'Velostat', which after a short delay delivers stable readings. In 2014, Park et al. [15] used 10 linear potentiometers, flexible wires and linear springs on each finger to calculate the joint angles and thereby measuring the motion of fingers. In 2012, Kadam et al. [16] made a bend sensor-based glove that emulates a sign language teacher for those who would like to learn the language. For this purpose, a bend sensor-based glove was designed to be able to learn 26 English sign gestures from the American Sign Language (ASL) and stored them into the EEPROM of its microcontroller. The finger movements of a user are recorded from flex sensors mounted on the glove. The study was examined with 86% accuracy on 14 selected gestures.

A system using inertial Micro-Electro-Mechanical Systems (MEMS) sensors to estimate individual finger motion and position has been argued for as a wearable interaction system in as early as 1999 in [17]. Most proposed systems, however, relied on combining multiple modalities to achieve higher accuracies. Jiangqin et al. [18] recognized in 1998 for instance 26 Chinese sign language words using a Cyberglove with 18 sensors and a 3D-Tracker, employing a multi-layer perceptron to remove noise from the sensor data that can be passed as input to Hidden Markov Models to classify the words. The reported recognition performance was over 90%. In [19], Hrabia et al. studied the relationship between

bio-mechanically modeled finger joints and report that it is possible to track the hand gesture using eight motion sensors with 9 DOF sensors comprising accelerometer, gyroscope and magnetometer.

Several projects have combined an inertial unit with bend sensors. In [5], the 5DT Data Glove 5 Ultra along with an accelerometer was used to obtain each finger's flexion degree and information about wrist orientation. These readings afterwards were processed offline with an artificial neural network to test classification of 24 ASL static hand gestures for fingerspelling. A classification rate of 94.07% accuracy was obtained on 1200 test patterns provided the network was trained on 5300 patterns. In [6], Vutinuntakasame et al. recognized fingerspelling with 5 flex sensors and a 3-DOF accelerometer connected to a Body Sensor Network (BSN). They proposed a hierarchical framework using multivariate Gaussian distribution coupled with bigram and set of rules to detect a particular padgram with an accuracy of about 72.7–73.6%. Tanyawiwat et al. [7] designed in 2012 a glove to recognize ASL fingerspelling hand gestures with 5 Contact sensors, 5 Flex Sensors and a 3-DOF accelerometer installed on it and also presented a concept of combined sensory channels. A combination of Multivariate Gaussian Distribution and the multi-objective Bayesian Frame network is used to classify the gestures and helped to improve the recognition accuracy rate to 77.4%. Here, six calibration steps were defined to adapt to different hand sizes. More recently, Tubaiz et al. [20] worked on the DG5-VHand data glove system, which consists of 5 bend sensors and a 3-DOF accelerometer for both hands to recognize 40 sentences in the Arabic sign language. The data glove communicates with a computer via Bluetooth and a camera is used to collect the data. Sentences were classified with K-Nearest Neighbour with resampled feature vectors, achieving a 98.9% accuracy.

Glove-based accelerometers and inertial units have also been combined with several other modalities to achieve more accurate detection of gestures. In [21], a 3-axis accelerometer and Electromyogram (EMG) was examined to figure out their complementary functionality and their potential in recognizing a small set of 7 German sign language words collected from 8 subjects. KNN Classification results produced an average accuracy of 98.9% on subject-dependent recognition whereas subject-independent recognition resulted in 54.82% accuracy. A vision-based sign language recognition using a hat-mounted camera coupled with the integration of an accelerometer was proposed by Brashears et al. [22]. The noise in recognizing hand gestures on 5 sets of words representing a vocabulary along with calibration gestures was ruled out. Using these sensors, it was shown that the accuracy is 90.48%. In 2007, Oz et al. [23] translated 60 ASL words into English using the Cyberglove and a 3D motion tracker. Gesture classification was performed by two artificial neural net classifiers obtaining a 95% accuracy. Later systems such as [24] have displayed similar gains in accuracy.

In contrast to the above approaches, we focus on designing a data glove to primarily sense the articulated hand via single IMUs placed on the fingertips. As IMUs have become extremely small, self-contained, and power-efficient, such a system lends itself well to be miniaturized. Unlike other, wrist-worn solutions for wearable gesture recognition that target a few gestures, such as [25], we focus on the recognition of sizable sets of gestures (more than 20). Furthermore, rather than capturing the raw data and processing it elsewhere, we developed an embedded sensor fusion technique that combines the data of each of the five fingers' IMUs to obtain stable and accurate information about the articulated hand. The classification is carried out through on-board computation as well, making the system function in real-time without a need for any external processing units. Using a trained classifier on a factory-provided calibrated IMU sensor also means that no personal calibration is required for this approach as we rely on the local frame of reference.

3. System Design

3.1. Hardware Design

Our goal is to design our glove with on-board computation capability, while being comfortable to wear and modularly built. The hardware components of the glove are shown in Figure 2 and described as below:

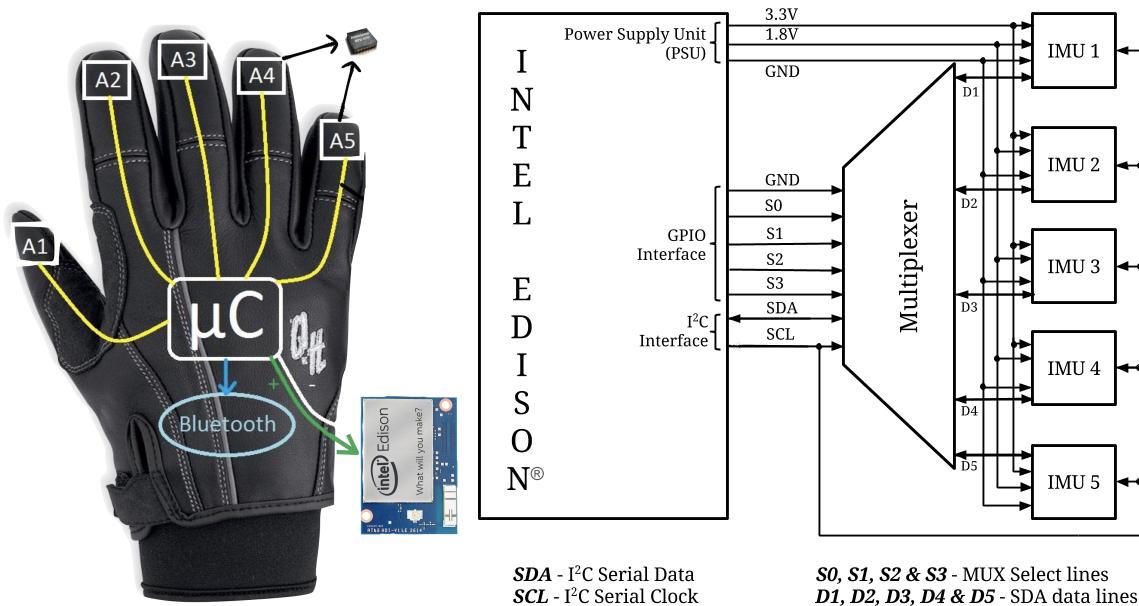


Figure 2. (Left): The presented prototype, equipped with Intel’s Edison module, connects serially with five IMUs at all finger tips that are bridged with a multiplexer. The glove has a 1000 mAh rechargeable battery that lasts at least 8 h with both processors, Bluetooth and WiFi on. **(Right):** The hardware design of the system, showing the combination of the Edison module and five IMU sensors, through the I^2C interface bus. Connections for the IMU are as follows: V_{CC} —3.3 V, V_{LOGIC} —1.8 V, SDA , SCL and Ground GND . The Edison’s SDA is shared among the IMUs through a multiplexer, where all IMUs have a common address. Power is sourced by the Edison’s internal Power Supply Unit.

Microprocessing Unit (MPU): A powerful on-board computation unit is the key component to make the glove functionally independent: Intel’s Edison was chosen, as it provides both processing and integrated wireless capabilities. Delivering high performance, the Edison features a dual-core CPU along with integrated WiFi and Bluetooth support. The MPU runs Ubilinux, a stripped-down version of the Debian Linux distribution for embedded applications. These characteristics enable the MPU to perform gesture recognition in real-time.

Inertial Measurement Unit (IMU): The IMU comprises a 3D accelerometer (ACC), 3D gyroscope (GYRO) and 3D magnetometer, which can measure acceleration, rate-of-turn and magnetic field respectively. For signs, the orientation of each finger relative to magnetic north are less critical, so the magnetometer readings are ignored and only orientation in the vertical planes are used. The magnetometer data might, on the other, have benefits for detecting gestures in certain environments where the user location and orientation is crucial for discerning gestures; This is left for future work. ACC and GYRO readings are 16-bit in resolution and are fused using a Complementary Filter to generate Euler angles (Roll and Pitch) as will be specified in the Sensor Fusion section. The MPU-9250 by InvenSense is used to collect the raw 9D inertial data from each finger.

Multiplexer (MUX): The IMU has only 2 selectable I^2C addresses and on the MPU’s side there are two I^2C buses available for communication, allowing at most 4 IMUs to be interfaced. Introducing 5 IMUs thus requires to multiplex the I^2C data line of the MPU to transmit data between the IMUs. To facilitate this, a 16-bit bi-directional multiplexer is used.

Power: The power consumption of the glove at 5 V and in full-speed mode is approximately 130 mA without an active WiFi connection, and 250 mA with an active WiFi connection. Bluetooth connectivity with a throughput of 0.27 Mbit/s is used to send the detected signs to the smartphone whereas WiFi currently just serves debugging purposes. The Lithium Polymer battery with 7.2 V @1000 mAh therefore can provide continuous power for at least 8 h. In real-time detection mode,

in which raw sensor data is not stored locally and all processing units are put into low-power modes when idle, this figure is much lower.

Glove: To accommodate the IMUs, a cotton glove was extended with elastic pockets on the finger tips and cable hooks along the fingers to manage the cables.

3.2. Sensor Fusion Approach

Our goal is to obtain precise orientation data of each finger. When used independently from each other, the raw readings from the ACC and GYRO are not reliable. The ACC unit measures any acceleration due to gravity or motion. In steady state, the earth's gravitational field causes a 1 g (or 9.81 m/s^2) acceleration on the axes to which the ACC is oriented. Moreover, the ACCs on the glove are under constant minor fluctuations due to involuntary quivering of the hand, registered as high-frequency noise. The GYRO measures angular rotations and, unlike the ACC, is hardly affected by external forces. However, the GYRO tends to drift from its mean value in the long run as it lacks a reference frame, which is visualized in Figure 3.

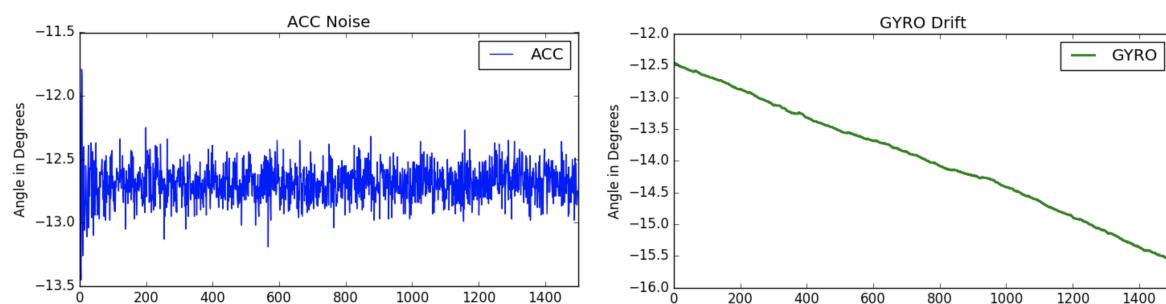


Figure 3. Accelerometer noise and gyroscope drift in static conditions (y axis), plotted over time (x axis, in equidistant sensor samples).

Sensor fusion combines the sensor data from different sources to reduce their uncertainty and improves the overall quality of any individual series of readings. In our case, the ACC and GYRO data are fused with a digital filter to minimize noise and drift respectively. Among different available filters, the Complementary Filter (CF) [26] is best suited for the problem at hand, as it primarily operates on signals having opposite noise levels. When applying the CF, high frequency noise in the ACC data and low frequency drift of the GYRO data are filtered out simultaneously. The Complementary Filter thus acquires each finger's acceleration and rotation readings from the sensors and delivers less noisy data with low drift.

Among the various ways of representing an orientation in 3D Euclidean space, Euler angles (EA) are chosen as a simple and expressive alternative. A body's orientation is described by a sequence of rotations around the main axes, namely Pitch θ around the x -axis, Roll ϕ around the y -axis and Yaw ψ around the z -axis respectively. This also is illustrated in Figure 4. The Yaw angle in this case is neglected as the z -axis, being parallel to the yaw axis of reference, is not affected by the gravitational force.

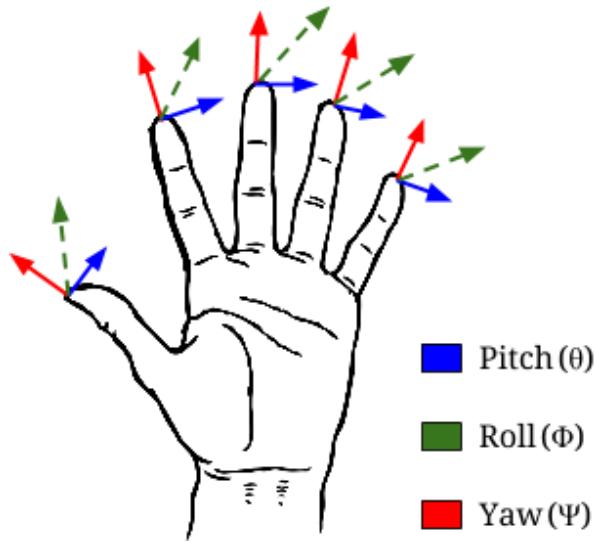


Figure 4. Euler Angle representation of each finger for measuring the absolute orientation. The IMUs are placed on the finger tip of the glove and track the hand motion.

The Euler angles can be calculated from the inverse tangent of the ACC data and the angular position or Gyro rate is obtained by integrating the angular velocity over time:

$$Pitch(\theta) = \text{atan} \left(\frac{ACC_x}{\sqrt{ACC_y^2 + ACC_z^2}} \right) \times \left(\frac{180}{\pi} \right)$$

$$Roll(\phi) = \text{atan} \left(\frac{ACC_y}{\sqrt{ACC_x^2 + ACC_z^2}} \right) \times \left(\frac{180}{\pi} \right)$$

$$Gyro_X_Rate = \int_0^t Gyro_X(t) \times dt$$

$$Gyro_Y_Rate = \int_0^t Gyro_Y(t) \times dt$$

With the sampling period dt , the raw GYRO angular velocities in the θ and ϕ axes $Gyro_X(t)$ and $Gyro_Y(t)$, and the uncompensated θ and ϕ angular positions $Gyro_X_Rate$ and $Gyro_Y_Rate$ respectively.

The Complementary Filter (Figure 5) introduces a coefficient factor α that controls the amount of influence of ACC and GYRO against one another. The high frequency fluctuations of the accelerometer data is attenuated by multiplying the unfiltered accelerometer value with a fraction of the time constant (Figure 6). This value was empirically determined to be $\alpha = 0.93$ (compare Figure 7).

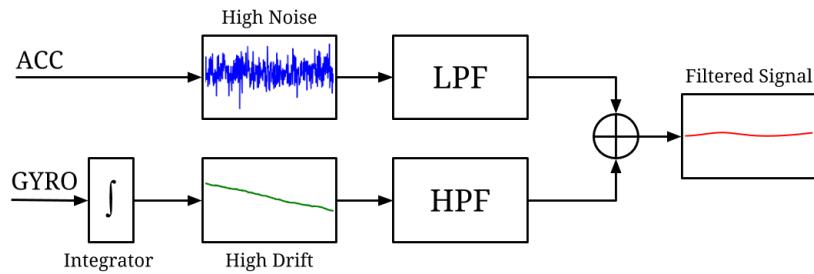


Figure 5. Working model of the Complementary Filter: Sensor fusion is performed on low-pass filtered accelerometer data and, after being integrated over time, high-pass filtered gyroscope data to obtain the filtered signal.

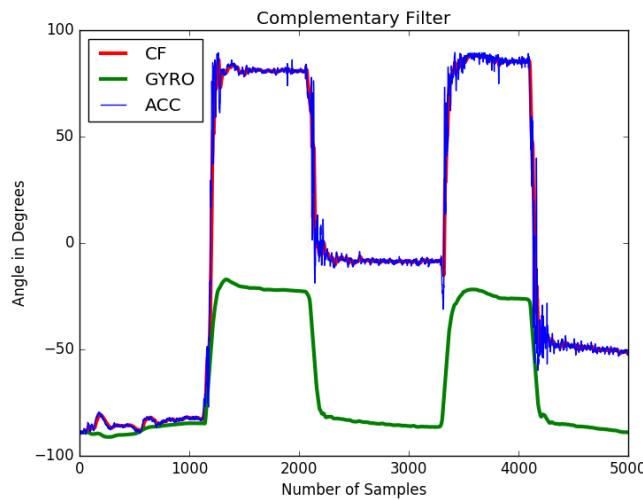


Figure 6. The Complementary Filter returns a smooth and consistent signal, by fusing both accelerometer (ACC) and gyroscope (GYRO) data.

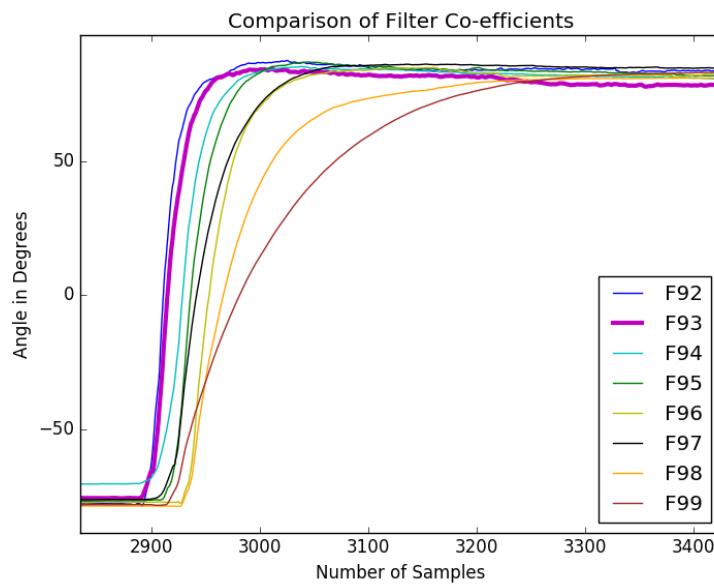


Figure 7. Comparison of eight different gyroscope to acceleration ratios for the complementary filter. The ratio is computed in the range $91 < X < 100$ with $\frac{X \cdot \text{Gyro}}{(100-X) \cdot \text{Acc}}$.

$$\theta_t = \underbrace{\alpha \times (\theta_{t-1} + \text{Gyro}_X(t) \times dt)}_{HPF} + \underbrace{(1 - \alpha) \times (\text{ACC}\theta_t)}_{LPF}$$

$$\phi_t = \underbrace{\alpha \times (\phi_{t-1} + \text{Gyro}_Y(t) \times dt)}_{HPF} + \underbrace{(1 - \alpha) \times (\text{ACC}\phi_t)}_{LPF}$$

With the current estimate of Pitch θ_t and Roll ϕ_t , the previous estimate of Pitch θ_{t-1} and Roll ϕ_{t-1} , and the current estimate of ACC Roll $\text{ACC}\theta_t$ and Pitch $\text{ACC}\phi_t$.

Pitch and Roll angles are continuously calculated from the readings of each finger with a sampling rate of approximately 3 ms (mostly due to the more processing-intensive use of the arctangent function, which we did not optimize). For every newly polled GYRO and ACC data, the angles are iterated using the CF. The HPF only allows the GYRO readings to pass through if their rate of change is large enough to be captured within the sampling period and likewise the LPF only lets ACC data pass whose rate of change is small. The Complementary Filter with its coefficient factor $\alpha = 0.93$ hereby relies to 93% on the previous angle integrated with the current angular position given by GYRO and to 7% of the current ACC angle to filter out the short-term noise and long-term drift.

As an alternative to any external filtering, InvenSense provides a proprietary sensor fusion code in form of a Binary Blob (Blob = Binary Large Object, closed-source binary piece of software) that is called Dynamic Motion Processing (DMP) and can be loaded onto the MPU-9250. The Binary Blob (BB) will be executed on the sensor itself and has a sampling rate of 20 ms. It requires up to 20 s settle time initially, whereas the CF is available instantaneously (see Figure 8, left). Furthermore, the BB has a slightly higher gesture change delay and tends to produce less stable results when compared to the CF (see Figure 8, right). As the Complementary Filter outperforms the Binary Blob, the CF is chosen as final filtering and sensor fusion approach.

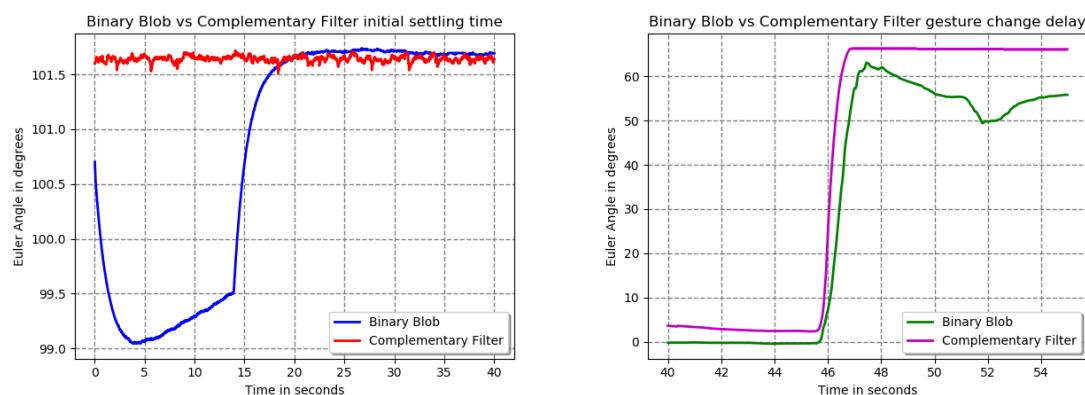


Figure 8. (Left): The Complementary Filter instantaneously provides a stable signal, whereas the Binary Blob needs up to 20 s settle time. (Right): The Complementary Filter reacts faster to a gesture change than the Binary Blob and returns more stable results.

3.3. Data Communication and Interface Design

The glove's intention is to use it as a novel input device to a computer system. To show its capabilities, an Android app integration was made that utilizes the glove to receive letters from the LSF alphabet. The glove detects the handshapes as described above, encodes them to a character and communicates them via Bluetooth to the smartphone. The app then visualizes the detected letters on screen provided the certainty of prediction is high (if the predicted probability for the classification crosses a threshold of 0.7), otherwise the data is classified as belonging to a null class, and a null

character is shown to indicate the undefined gesture. The flowchart presented in Figure 9 depicts a detailed description of the working model for the whole arrangement in real-time scenarios. The glove and a Smartphone (SP) are paired with each other at the beginning of the control routine via Bluetooth, where the glove acts as a server and waits for a Bluetooth-enabled glove to start sending.

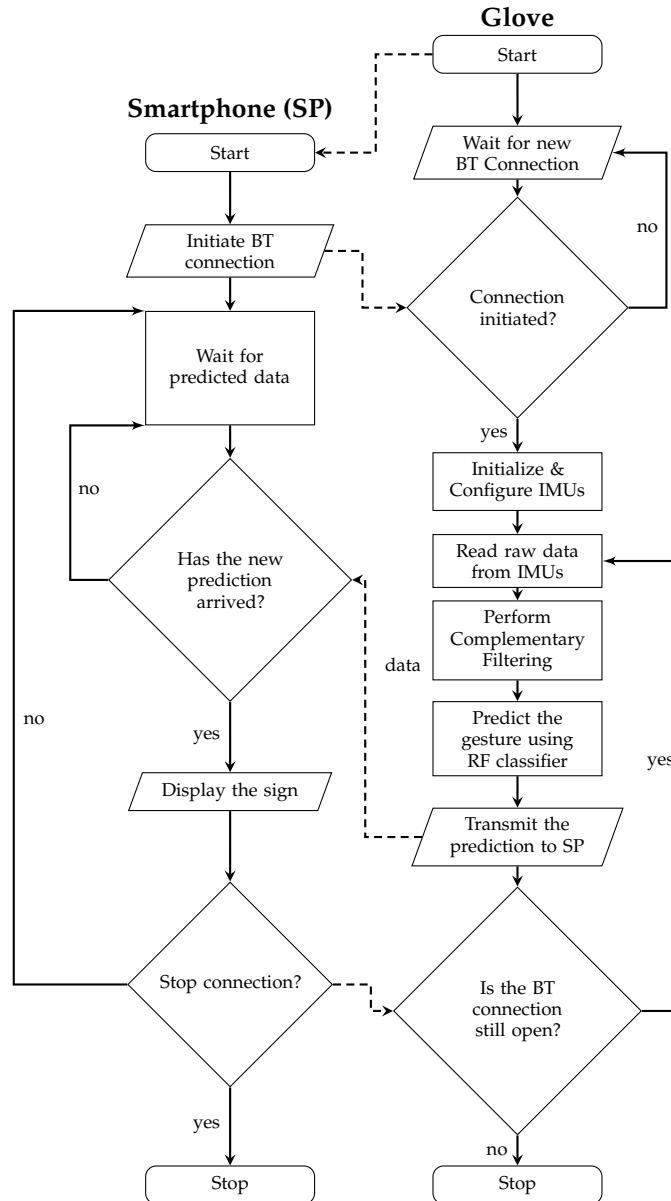


Figure 9. The flowchart describing the entire setup and communication process between smartphone and our glove prototype (BT—Bluetooth, IMU—Inertial Measurement Unit, RF—Random Forest classifier).

The following steps illustrate the sequence of actions presented in the flowchart:

- Step 1:** The glove acts as a server and waits for a Bluetooth-enabled device to initialize the program. Thereafter, the smartphone can connect to the glove and will wait for the first set of predicted data.
- Step 2:** After the connection prompt, the glove starts to initialize and configure the IMUs.
- Step 3:** Once the sensors are set, the raw data (ACC & GYRO) for each finger is read in cyclic manner.
- Step 4:** The sampling time to fetch the readings from each IMU is measured precisely and the CF is performed.

- Step 5:** The sensor-fused data from the above step is fed to RF classifier to predict the gesture from the obtained readings.
- Step 6:** The corresponding sign of the predicted gesture is transmitted to the smartphone via Bluetooth.
- Step 7:** Smartphone displays the currently received sign if it is not same as the previous one.
- Step 8:** This cycle is repeated indefinitely until the connection is terminated by the smartphone.

4. Evaluation

The proposed glove system's gesture recognition capabilities are in this section examined in detail on data collected from 57 study participants performing gestures from the French sign language (LSF) alphabet, in order to validate how well such a system would work for larger sets of gestures. The following will detail the experiment's scope, after which it covers the setup of the data collection, the experiment design, and the experiment results from all classifiers and classification parameters tested.

4.1. Experiment Scope and Goals

The following experiment intends to investigate the classification aspects of the system that was discussed thus far. First and foremost, the aims of our experiments are (1) to find the classification algorithm that performs best for this paper's fingerspelling hand gestures; as well as (2) to investigate which gestures perform well and (3) the impact of the amount of training data on the classification performance. The best-performing classification algorithm needs to be measured from the point of view of accuracy in the detection of performed gestures, as well as the speed at which the classification model can be trained and used during testing. The accuracy performance of all signs' gestures is required to obtain insights into the confusion that might occur between individual, hard-to-distinguish gestures. In addition to the single performance figures for all gestures, an empirical investigation in the influence of the amount of training data on the performance of the model was performed. The goal of this is to identify how much data, in particular in terms of the number of participants that is required to build up a sufficiently strong training dataset.

4.2. Data Collection and Experiment Design

Choice of hand gestures: The finger spelling alphabet. For validation of the gesture recognition capabilities of the embedded glove, 22 out of the 26 possible gestures from the French sign language alphabet were chosen to obtain a significantly large amount of gestures. The four gestures for "J", "P", "Y", "Z" were left out due to their non-static nature. The sensor readings from all five IMUs are mapped to one of the gestures that are uniquely described by a set of angles assumed by the fingers. More precisely, each data sample is represented by a 10-dimensional vector, which contains the roll and pitch Euler angles for each finger tip. Each gesture hereby will have minor offsets from the ideal finger positions, especially when compared between different users. For the training of the classification models, it therefore is crucially important to repeat the capture of each gesture for a sufficient amount of times and for a high variety of users in order to capture these slight deviations.

Experiment design: leave-one-user-out, 57-fold cross validation. To obtain experiment results that include the high variability of gesture instances, a large amount of data samples including all possible orientations in which a user could perform a particular sign language gesture were recorded from 57 participants in total. Of these 57 participants, one participant who was a native LSF instructed the hand gestures and movements of each gesture to the remaining 56 participants before the start of the study, such that each participant was shown a short demonstration to make them familiar with the French sign language signs before gathering the glove's sensor samples.

We created the dataset so that classification results could be created from using 57-fold cross validation, more specifically by following a leave-one-user-out approach (as suggested by [27] as a proper methodology): For each fold, the model was trained with the gesture data of 56 subjects and the

gesture data of the remaining subject was used for testing. Each subject's gesture data hereby amounts to 22,000 samples (1000 samples for each of the previously discussed handshapes).

It is important to note that the samples involving transition between the gestures were ignored and no calibration was required as the sensors were initialized with their default factory calibration. To support the variability of the training and testing data in the dataset, the participants were instructed to assume a resting position for about 10 s between performing all gestures, and each gesture was performed five times by the same user. For each of the single gesture instances that was performed, 200 data samples were retained for the dataset to ensure that all classes in the evaluation would be balanced. This thus results in every gesture having training data from 1000 samples (since every gesture was performed five times) per participant, or 57,000 data points across all participants. The total dataset contains approximately 1.25 million samples in total. The data collection from each participant took about 15 min and the sensor readings were stored as a CSV file on the onboard Intel Edison module. To prepare the dataset for the Random Forest classifier, as detailed in the next section, manual data annotation was performed afterwards of all gestures.

4.3. Classification Algorithms, Parameters, and Performance

Classifiers. Different machine learning approaches like Support Vector Machines (SVM), Naive Bayes (NB), Multi Layer Perceptrons (MLP) and Random Forest (RF) models [28] were examined to classify the hand gestures as shown in Table 1. Among them, MLP and RF showed the most promising results, whereas the SVM with different kernels took a substantial amount of time to train with all the data and was not further evaluated. The MLP had been reconfigured with a different number of hidden layers and neurons and also by tuning various hyperparameters to achieve satisfactory results. A stochastic gradient descent algorithm with a learning rate of 1.0×10^{-5} and 0.5 momentum along with the total of 100 neurons in each of the two or three hidden layers was shown as a best MLP classifier for our task. The performance of RF with a varying number of 5 to 75 trees was tested. The accuracy of validation results on the dataset showed minor improvements with the increase in number of trees. As a trade-off, the computation time of the RF was gradually increased with increasing number of trees resulting in a delay in performing the gesture classification. Since the aim of our work is to design a real-time gesture recognition glove, an RF with 15 trees was found to be a reasonable decision to balance accuracy and response rate. The performance and accuracy produced by both, MLP with 100 neurons and RF with 15 trees on the validation data was comparable. The rate of false positives for MLP was relatively high compared to RF and thus the RF with 15 trees was chosen as a final classifier.

Table 1 summarizes the training times and the accuracy results for these different classifiers. The accuracy of our gesture recognition approach with the random forest (RF) algorithm is detailed further in the confusion matrix in Table 2. Several 1000 samples representing a gesture is predicted in each fold of the cross-fold validation, the confusion matrix depicts the average prediction count of each sign over 57 iterations. The overall mean accuracy of all the gestures was 92.4% with a standard deviation of 0.042, the F1 score was 91.3% with a standard deviation of 0.048. Although our study's focus is limited to validating the embedded classification performance, an additional step that could have been carried out is to measure the accuracy of the Euler angles that were obtained from the IMU data. This is however notoriously hard to obtain for hand articulation. Given the results in this section, we furthermore argue that these estimates are exact enough for the task of gesture classification.

Table 1. Overall comparison of the performances, in training time and in classification accuracy, from different types of classifiers on the glove's IMU data, by training on the data from 56 participants and testing it on the remaining 57th participant. Based on these results from this cross-validation study, the Random Forest (RF) classifier with 15 sub-trees was chosen as it gave the highest accuracy (92.95%) for a negligible training time (105 s per fold ¹).

Classifier	Parameters	Training Time (Seconds/Fold ¹)	Testing Time (Seconds/Sample)	Accuracy (%)	F1 Score (%)
Naive Bayes	-	2.2	0.002	89.1	87
<i>No. of layers</i>					
MLP	2 layers with 50–50 units and final layer with 10 classes	3144.8	0.16	90.9	89.7
	2 layers with 100–50 units and final layer with 10 classes	3630	0.19	91.8	90.3
RF	2 layers with 150–50 units and final layer with 10 classes	3860	0.23	92.2	91.1
	<i>No. of sub-trees</i>				
RF	5	20	0.12	90.3	89.0
	10	50	0.13	92.3	91.3
	15	105	0.14	92.95	91.98

¹—Timings are averaged across all folds in the cross-validation procedure.

Individual gestures' results. The per-class accuracy data show good results for most hand shape signs, with 12 of the hand shapes having an average accuracy of more than 95%, six of them account in the range of 89% to 95% and four hand shapes vary between 75% to 80% (Table 3). With the evidence provided by the confusion matrix, the reason for the accuracy drop found in the gestures 'F', 'T', 'L', 'U' can be analyzed more thoroughly: The trained model has most difficulties with the 'F'-'T' and 'L'-'U' hand shapes. The gestures of the signs 'F' and 'T' find more similarity with each other resulting in a mean prediction count of 'F' as 'T' for 184 times and 'T' as 'F' for 163 times out of 1000 in the confusion matrix. While measurements from all the other fingers remain the same, pitch angle readings from the thumb are the only apparent measurements that would differentiate both of the gestures. On the other hand, a similar behavior can also be observed for 'L' and 'U'. Here, uncertainty is introduced since the difference of orientation can be noticed only for the thumb and middle fingers. As we gathered data from distinct people having different hand structures varying in length, width and thickness, the pitch of the thumb and the roll of middle finger for a few subjects would not reveal great variance in the measurements, with the same holding for 'F' and 'T'. Hence, the random forest classifier exhibited some ambiguity in recognizing these four hand shapes during cross validation.

Impact of training set size. Table 4 shows accuracy and standard deviation over 10 folds, as the training set is increased from 5 participants to 40. As is shown in this table and Figure 10, our study data indicates that at least 30 participants were needed to achieve an accuracy of at least 90%.

The evaluation results in Table 4 on these subsets of our full dataset, when compared with the accuracy scores in Table 3, shows that a model with only five subjects' data resulted in a significant drop in the accuracy and F1 score to 79% and 77% respectively. After obtaining these validation results on the total dataset of 57 volunteers, the glove-based system has also been tested on further people in real-time studies: the time taken to predict the gesture was found to be performed within 65 ms, where 23 ms were spent on gathering the sensor readings from all 5 IMUs, and the random forest classifier predicting the hand shape from the obtained readings within 42 ms.

Table 2. The confusion matrix for the random forest classifier's 24 recognized fingerspelling signs. Visible are minor confusions between 'L' and 'U', and 'T' and 'F'. Overall accuracy is 92.4% (with a standard deviation of 0.042), the overall F1-score is 91.3% (with a standard deviation of 0.048).

A	B	C	D	E	F	G	H	I	K	L	M	N	O	Q	R	S	T	U	V	W	X
A 986.63	0.00	0.02	0.00	0.00	0.54	0.00	0.00	12.25	0.56	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
B 0.00	900.46	16.86	22.35	56.68	0.32	0.00	0.00	0.05	0.00	0.00	0.09	0.00	0.00	0.42	0.00	0.02	0.19	0.00	1.25	0.00	1.32
C 0.02	1.58	978.89	0.79	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	1.05	0.14	0.00	17.49	0.00	0.00	
D 0.00	9.51	1.79	953.98	34.32	0.32	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.07	0.00	0.00	0.00	0.00
E 0.00	33.21	0.11	36.70	928.79	0.88	0.00	0.02	0.00	0.26	0.00	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
F 12.84	13.56	0.05	0.21	0.79	754.32	0.51	0.00	13.67	0.00	0.05	0.00	0.00	0.00	1.54	11.07	3.33	183.19	0.00	3.44	0.04	1.39
G 0.02	0.07	0.00	0.11	0.00	0.37	998.67	0.35	0.04	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.26	0.05	0.00	0.00	0.00	0.02
H 0.00	0.05	0.00	0.00	1.49	0.00	0.74	893.67	3.93	0.00	0.00	94.30	3.56	0.00	2.26	0.00	0.00	0.00	0.00	0.00	0.00	0.00
I 35.33	0.16	0.00	0.00	0.21	4.05	0.00	0.61	956.21	0.02	0.00	0.00	0.04	0.00	0.05	0.00	0.00	3.32	0.00	0.00	0.00	0.00
K 0.86	0.00	0.00	0.04	0.00	0.05	0.00	0.00	0.19	993.35	0.04	0.00	0.00	0.00	0.02	5.35	0.00	0.05	0.00	0.04	0.02	0.00
L 0.09	0.00	0.00	0.04	0.02	0.04	0.00	0.00	1.09	0.23	791.82	0.00	0.00	0.00	0.00	0.00	0.00	0.02	206.54	0.09	0.00	0.04
M 0.00	0.00	0.00	0.00	1.28	0.00	0.00	23.16	0.04	0.02	0.72	932.30	38.56	0.00	3.75	0.02	0.00	0.00	0.16	0.00	0.00	0.00
N 0.00	0.00	0.00	0.00	0.00	0.00	0.00	2.54	0.00	0.00	0.00	34.60	961.51	0.07	1.28	0.00	0.00	0.00	0.00	0.00	0.00	0.00
O 0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.02	0.00	0.16	1.23	996.33	2.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Q 0.00	4.70	0.02	0.00	0.00	0.07	0.00	0.84	11.32	0.00	0.00	10.84	8.65	1.07	962.33	0.12	0.00	0.02	0.00	0.02	0.00	0.00
R 2.74	0.18	0.00	0.00	0.00	64.82	0.12	0.00	0.00	19.53	0.12	0.00	0.00	0.00	0.05	897.40	0.89	13.82	0.02	0.00	0.30	0.00
S 0.00	0.51	43.86	0.00	0.00	0.47	0.00	0.00	0.00	1.67	0.00	0.65	0.02	0.00	0.11	13.35	933.93	1.98	0.00	0.00	3.46	0.00
T 0.00	10.19	0.16	0.23	0.00	163.46	3.91	0.00	25.82	5.96	0.02	0.00	0.00	0.00	0.04	21.26	11.86	752.79	0.02	0.00	3.02	1.26
U 0.00	0.00	0.00	0.00	0.00	0.67	0.00	0.00	0.12	0.19	236.28	0.00	0.00	0.00	0.02	0.18	0.00	0.68	761.77	0.09	0.00	0.00
V 0.00	1.58	5.89	0.00	0.00	0.07	0.00	0.00	0.00	0.04	0.58	0.02	0.00	0.00	0.00	0.04	0.00	0.00	0.58	987.46	3.75	0.00
W 0.04	0.14	1.25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.07	0.07	0.05	0.00	0.86	994.65	2.88
X 0.00	4.63	0.00	0.00	0.00	0.58	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	1.00	993.72		

Table 3. Mean accuracy scores (denoted as acc, in %) and standard deviations (denoted as σ) of all 57-fold cross validation results given for each gesture in the french sign language (LSF) alphabet.

	A	B	C	D	E	F	G	H	I	K	L
acc	98.6	90.0	97.8	95.3	92.8	75.4	99.8	89.3	95.6	99.3	79.1
σ	0.0632	0.2526	0.0875	0.1703	0.1712	0.3156	0.0026	0.2070	0.1390	0.0241	0.3162
	M	N	O	Q	R	S	T	U	V	W	X
acc	93.2	96.1	99.6	96.2	89.7	93.3	75.2	76.1	98.7	99.4	99.3
σ	0.1019	0.1043	0.0102	0.1463	0.2677	0.1972	0.3674	0.3304	0.0502	0.0200	0.0326

Table 4. The following table shows the relation between the accuracy and the number of participants chosen for evaluation. The number of participants was selected for 10 folds, and performance was evaluated on test data from a fixed number of 17 participants.

No. of Participants	Accuracy in %	σ	No. of Participants	Accuracy in %	σ
5	79	0.02	30	90	0.0055
10	84	0.0141	35	91	0.0025
20	88	0.011	40	91	0.0024

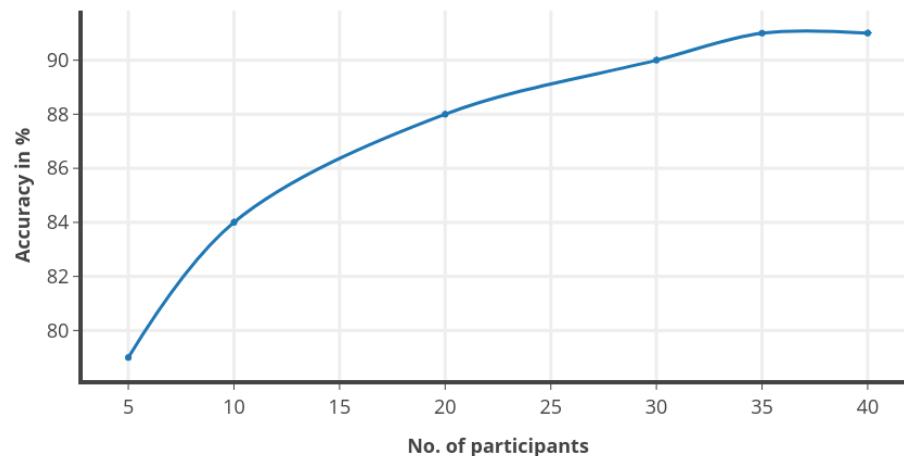


Figure 10. The relation between the number of participants' data and the accuracy. For an accuracy of 90% or above, a minimum of 30 participants were needed for our study.

5. Conclusions and Future Work

This article has contributed with a data glove design that enables the on-glove detection of fine-grained hand shapes based on IMU sensors on all fingertips. It takes advantage of recent System-on-Chip designs that are powerful enough to perform data fusion and classification routines in real-time within the glove. The total number of components with 5 IMUs, a multiplexer, and the built-in microprocessor accounts for entire configuration of the glove. To improve the noisy and drift-prone sensor readings from the IMUs, a Complementary Filter was employed to generate a smooth and consistent signal. Fusing the data moreover yields a precise measure of the finger orientation and movements. To evaluate the glove's capabilities, it was trained to 22 different hand gestures from the french sign language alphabet. It was shown that the gesture recognition performs poorly when there is not enough variance in the training data, especially in a scenario where the wearer did not train the glove. So to account for the differences in individuals' hand movements, a dataset from 57 people was collected to capture the variations of gestures made while fingerspelling. The system

is trained and executed fully on the glove, making it thus capable of recognizing the non-dynamic hand gestures from the LSF signs. The performance was evaluated using 57-fold (leave-one-user-out) cross-validation, resulting in 92% mean accuracy and an F1 score of 91%. Some complications however were found on similar signs that are not easily distinguishable for the glove. The proposed system has shown a real-time detection performance with gesture recognition being performed on-board within 65 milliseconds.

The evaluation results lead to the conclusion that the glove is able to learn and thereafter recognize various different hand gestures in real-time. As advise for a transferable system, the hand gestures should not be too similar and it is essential to train the glove with highly diverse data possibly coming from multiple people. Furthermore the developed glove can easily be utilized as an input device to any Bluetooth-enabled system as shown on the Android app example that was developed to immediately visualize the recognized gestures. A demo highlighting the glove's fingerspelling functionality can be found here (<https://youtu.be/RW4J5zOLeDg>).

Although our focus was to investigate the use of on-glove fusion and classification of gestures, the proposed idea of on-glove gesture recognition can further be investigated to make the model more reliable and energy-efficient. Examining different variants of sensor fusion might produce more accurate measurements than the Complementary Filter. Training the system to also recognize dynamic gestures would not only allow to detect all signs in sign language but would also enhance the space of possible gestures and helps resolving ambiguity in similar hand shapes. To achieve better classification results for the gestures used in this article's study, we expect that well-selected touch or proximity sensors in the glove would enable an improved detecting of the hand articulation. The prototype's form factor can drastically be reduced by shrinking the PCBs and replacing the wires by conductive ribbons. The glove-embedded microprocessor is capable of allowing a more interactive design, that could be controlled by the system the glove currently is utilized for as input device. In various realistic applications, such as home-automation or multimedia interfaces, it would be beneficial to calculate quaternions from the IMU readings to better represent the fingertip orientations and thus any articulated hand model. Finally, the presented system can be complemented with hand articulation tracking in 3D, to widen such a system to gestures that use relative distances between the hands (e.g., zooming in gestures where the distance between hands is decreased).

Both data set and source code for the system and evaluations in this paper are publicly available online on: <http://ubicomp.eti.uni-siegen.de/home/datasets> to facilitate reproduction of our results.

Author Contributions: The development of the glove system, the carrying out of the study, initial writing, and the analysis of all data were done by C.K.M., F.P.P. L., K.D. V. and S.K. P.M.S. and K.V.L. have assisted in the design of the user study and the classification design, as well as the writing. J.K. has assisted in the writing of this paper and the expansion and restructuring of the related work section and the sensor fusion approach.

Funding: Funding for this research was supplied by the FFG's MiniAttention Project (BRIDGE Frühphase, 3. Ausschreibung 2014, Projekt 851227).

Acknowledgments: We would like to thank all 57 study participants for wearing and testing our system.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mummadi, C.K.; Leo, F.P.P.; Verma, K.D.; Kasireddy, S.; Scholl, P.M.; Laerhoven, K.V. Real-time Embedded Recognition of Sign Language Alphabet Fingerspelling in an IMU-Based Glove. In Proceedings of the 4th international Workshop on Sensor-based Activity Recognition and Interaction—iWOAR, Rostock, Germany, 21–22 September 2017.
2. Dipietro, L.; Sabatini, A.M.; Dario, P. A survey of glove-based systems and their applications. *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* **2008**, *38*, 461–482.
3. Sturman, D.J.; Zeltzer, D. A survey of glove-based input. *IEEE Comput. Graph. Appl.* **1994**, *14*, 30–39.

4. Zimmerman, T.G.; Lanier, J.; Blanchard, C.; Bryson, S.; Harvill, Y. A Hand Gesture Interface Device. In Proceedings of the SIGCHI/GI Conference on Human Factors in Computing Systems and Graphics Interface, Toronto, ON, Canada, 5–9 April 1987; pp. 189–192.
5. Cabrera, M.E.; Bogado, J.M.; Fermin, L.; Acuna, R.; Ralev, D. Glove-based gesture recognition system. In Proceedings of the 15th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines, Baltimore, MD, USA, 23–26 July 2012; pp. 747–753.
6. Vutinuntakasame, S.E.A. An assistive body sensor network glove for speech-and hearing-impaired disabilities. In Proceedings of the IEEE 2011 International Conference on Body Sensor Networks (BSN), Dallas, TX, USA, 23–25 May 2011; pp. 7–12.
7. Tanyawiwat, N.; Thiemjarus, S. Design of an assistive communication glove using combined sensory channels. In Proceedings of the 2012 Ninth International IEEE Conference on Wearable and Implantable Body Sensor Networks (BSN), London, UK, 9–12 May 2012; pp. 34–39.
8. Hein, A.; Kirste, T. A Hybrid Approach for Recognizing ADLs and Care Activities Using Inertial Sensors and RFID. *Proceedings of the 5th International Conference Universal Access in Human-Computer Interaction, Intelligent and Ubiquitous Interaction Environments (UAHCI '09)*, San Diego, CA, USA, 19–24 July 2009; Springer: Berlin/Heidelberg, Germany, 2009; pp. 178–188.
9. Zinner, A.; van Laerhoven, K.; Schiele, B. Toward Recognition of Short and Non-repetitive Activities from Wearable Sensors. In *Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 2007; pp. 142–158.
10. Kuroda, T.; Tabata, Y.; Goto, A.; Ikuta, H.; Murakami, M. Consumer price data-glove for sign language recognition. In Proceedings of the 5th International Conference on Disability, Virtual Reality and Associated Technologies (ICDVRAT 2004), Oxford, UK, 20–22 September 2004; pp. 253–258.
11. Khambaty, Y.; Quintana, R.; Shadaram, M.; Nehal, S.; Virk, M.A.; Ahmed, W.; Ahmedani, G. Cost effective portable system for sign language gesture recognition. In Proceedings of the IEEE International Conference on System of Systems Engineering (SoSE '08), Singapore, 2–4 June 2008; pp. 1–6.
12. Huang, Y.; Monekosso, D.; Wang, H.; Augusto, J.C. A concept grounding approach for glove-based gesture recognition. In Proceedings of the 7th IEEE International Conference on Intelligent Environments (IE), Nottingham, UK, 25–28 July 2011; pp. 358–361.
13. Huang, Y.; Monekosso, D.; Wang, H. Clustering ensembles based on multi-classifier fusion. In Proceedings of the 2010 IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS), Xiamen, China, 29–31 October 2010; Volume 3, pp. 393–397.
14. Jeong, E.; Lee, J.; Kim, D. Finger-gesture recognition glove using velostat (ICCAS 2011). In Proceedings of the 2011 11th IEEE International Conference on Control, Automation and Systems (ICCAS), Gyeonggi-do, South Korea, 26–29 October 2011; pp. 206–210.
15. Park, Y.; Lee, J.; Bae, J. Development of a finger motion measurement system using linear potentiometers. In Proceedings of the 2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics, Besacon, France, 8–11 July 2014; pp. 125–130.
16. Kadam, K.; Ganu, R.; Bhosekar, A.; Joshi, S. American Sign Language Interpreter. In Proceedings of the 2012 IEEE Fourth International Conference on Technology for Education, Hyderabad, India, 18–20 July 2012; pp. 157–159.
17. Perng, J.K.; Fisher, B.D.; Hollar, S.; Pister, K.S.J. Acceleration Sensing Glove (ASG). In Proceedings of the Third International Symposium on Wearable Computers (ISWC 1999), San Francisco, CA, USA, 18–19 October 1999; pp. 178–180.
18. Wu, J.; Gao, W.; Song, Y.; Liu, W.; Pang, B. A simple sign language recognition system based on data glove. In Proceedings of the 1998 Fourth IEEE International Conference on Signal Processing Proceedings (ICSP'98), Beijing, China, 12–16 October 1998; Volume 2, pp. 1257–1260.
19. Hrabia, C.E.; Wolf, K.; Wilhelm, M. Whole hand modeling using 8 wearable sensors: Biomechanics for hand pose prediction. In Proceedings of the 4th Augmented Human International Conference, New York, NY, USA, 7–8 March 2013; pp. 21–28.
20. Tubaiz, N.; Shanableh, T.; Assaleh, K. Glove-based continuous Arabic sign language recognition in user-dependent mode. *IEEE Trans. Hum. Mach. Syst.* **2015**, *45*, 526–533.

21. Kim, J.; Wagner, J.; Rehm, M.; André, E. Bi-channel sensor fusion for automatic sign language recognition. In Proceedings of the 2008 8th IEEE International Conference on Automatic Face & Gesture Recognition, Amsterdam, The Netherlands, 17–19 September 2008; pp. 1–6.
22. Brashear, H.; Starner, T.; Lukowicz, P.; Junker, H. *Using Multiple Sensors for Mobile Sign Language Recognition*; ISWC 2003; IEEE Press: New York, NY, USA, 2003.
23. Oz, C.; Leu, M.C. Linguistic properties based on American Sign Language isolated word recognition with artificial neural networks using a sensory glove and motion tracker. *Neurocomputing* **2007**, *70*, 2891–2901.
24. Georgi, M.; Amma, C.; Schultz, T. Recognizing Hand and Finger Gestures with IMU Based Motion and EMG Based Muscle Activity Sensing. *Proceedings of the International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2015), Lisbon, Portugal, 12–15 January 2015*; SCITEPRESS—Science and Technology Publications, Lda: Setúbal, Portugal, 2015; Volume 4, pp. 99–108.
25. Gong, J.; Yang, X.D.; Irani, P. WristWhirl: One-handed Continuous Smartwatch Input Using Wrist Gestures. In Proceedings of the 29th Annual Symposium on User Interface Software and Technology (UIST ’16), Tokyo, Japan, 16–19 October 2016; pp. 861–872.
26. Gui, P.; Tang, L.; Mukhopadhyay, S. MEMS Based IMU for Tilting Measurement: Comparison of Complementary and Kalman Filter Based Data Fusion. *Proceedings of the 2015 10th IEEE Conference on Industrial Electronics and Applications (ICIEA 2015), Auckland, New Zealand, 15–17 June 2015*; IEEE: Piscataway, NJ, USA, 2015; pp. 2004–2009.
27. Hammerla, N.Y.; Ploetz, T. Let’s (Not) Stick Together: Pairwise Similarity Biases Cross-validation in Activity Recognition. In Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing (UbiComp ’15), Osaka, Japan, 7–11 September 2015; pp. 1041–1051.
28. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).