

Data-driven Glove Calibration for Hand Motion Capture

Yingying Wang *

Michael Neff †

Department of Computer Science, University of California, Davis



Figure 1: Glove Captured Hand Poses

Abstract

Hand motion is an important component of human motion, playing a central role in communication. However, it is difficult to capture hand motion optically, especially in conjunction with full body motion. Due to a lack of appropriate calibration methods, data gloves also do not provide sufficiently accurate hand motion. In this paper, we present a novel glove calibration approach that can map raw sensor readings to hand motion data with both accurate joint rotations and fingertip positions. Our method elegantly handles the sensor coupling problem by treating calibration as a flexible mapping from sensor readings to joint rotations. A sampling process collects data tuples according to accuracy requirements, and organizes all the tuples in a training set. From these data, a specially designed Gaussian Process Regression model is trained to infer the calibration function, and the learned model can be used to calibrate new sensor readings. For real-time hand motion capture, a sparse approximation of the model is used to enhance performance. Evaluation experiments demonstrate that our approach provides significantly better results that have more accurate hand shapes and fingertip positions, compared to other calibration methods.

CR Categories: I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation, Virtual Reality

Keywords: Human Motion, Motion Capture, Calibration, Character Animation

Links:  DL  PDF

*e-mail: yiwang@ucdavis.edu

†e-mail: neff@cs.ucdavis.edu

1 Introduction

Hand motion plays a key role in manipulating objects and conveying information. Collecting hand motion data is crucial for research in robotics, sign language, character animation and human computer interaction. It is difficult to capture finger movement using optical systems due to the occlusion issue, especially when whole body motion is also required. Data gloves are an appealing alternative, but have seen relatively little research compared to well-developed body motion capture. One fact that inhibits the popularization of glove-based hand motion capture is the inaccurate raw data provided by the sensors. Raw sensor readings are a vector of signals which require a non-trivial calibration process to convert them into an accurate hand pose.

In this paper, we present a novel and sophisticated glove calibration method that accommodates multiple undesirable sensor factors (nonlinearity, cross-coupling and noise) and is capable of achieving both FK hand shape accuracy and IK fingertip position accuracy for captured hand motion.

We construct the hand model convenient and appropriate for glove sensor layout, based on which a careful examination is performed to check sensor linearity and coupling effects. Correspondence between sensor readings and hand joint rotations are categorized into several types. The mapping function that maps sensor readings to accurate rotations is the goal we need to infer. Essentially, we treat the glove calibration as a data-driven prediction problem, where the training data is collected from a separate sampling process.

The sampling process is further divided into FK independent sampling, FK cross-couple sampling and IK sampling, according to different sensor-rotation correspondence types and accuracy concerns. Key hand poses with measured joint rotations are predefined to assist FK independent sampling; widgets are built to help efficiently collect dense cross-coupled data. During IK sampling, a probabilistic distribution in end effector space is constructed to automatically calculate the most probable fingertip positions. Our constrained IK solver solves for the ideal joint rotations that can reach the accurate fingertip positions while maintaining plausible hand shapes.

A novel Linear Mean Composite Gaussian Process Regression (LMC-GPR) model with high evidence and low time cost is selected for glove calibration within the general GPR framework. The latent calibration function is inferred by maximizing the marginal likelihood, when training this LMC-GPR model on sampled data. To make our method more affordable for real-time glove calibration, Sparse Pseudo-input Gaussian Processes (SPGPs) are applied

to enhance the training and prediction performance. While testing on CyberGloves, we do not restrict our method to specific hand capture hardware, instead, we generalize our method to flexible calibration for different possible glove hardware.

Glove calibration research faces several challenges: sensor non-linearity, noise, reading differences w.r.t. hand sizes, influences from sensor stretching as well as other factors that cause reading variance. For example, the abduction sensors between fingers are stretched by finger flexion movements, and thus can not be accurately calibrated by methods that assume sensor independence [Kahlesz et al. 2004]. The false assumption of exact sensor linearity restricts flexible sampling which either causes over-fitting or fails to get solutions [Griffin et al. 2000; Hu et al. 2004].

Another challenge of glove calibration is the difficulty of obtaining the ground truth of real hand joint rotations. Manual measurement of hand poses is time-consuming and usually inaccurate. To avoid pose measurement, [Huenerfauth and Lu 2010; Menon et al. 2003] use limited samples, while [Chou et al. 2000; Hu et al. 2004] utilizes an expensive vision tracking system. Without enough ground truth data, there is no sound evaluation method for glove calibration accuracy.

The number of sensors on the glove and DOFs in the hand model make glove calibration a high dimensional problem, thus manual adjustment of parameters for each sensor on-site is intolerable. Accuracy requirements for captured hand motion depend on the applications: [Huenerfauth and Lu 2010; Kahlesz et al. 2004; Chou et al. 2000; Menon et al. 2003] emphasize plausible hand shapes; [Griffin et al. 2000; Hu et al. 2004] focus on exact fingertip positions. No current calibration method can satisfy both FK and IK accuracy requirements simultaneously.

Due to the difficulties of glove calibration, we generalize the fundamental contributions of our paper as follows:

1. We performed an extensive examination of sensor properties based on defined hand kinematic structure, and categorized sensor-rotation correspondences into different types.
2. We present an independent sampling process, separate from the calibration to collect training data meeting varied accuracy concerns, where predefined key poses, widgets and an IK probabilistic model are used to accelerate the sampling.
3. We introduce a novel LMC-GPR model designed specifically for glove data that can achieve both FK and IK accuracy.
4. We generalize our calibration method so that it can be adapted to other hand motion capture hardware.

2 Related Work

Previous glove calibration work can be divided into FK calibration and IK calibration, where the former aims to achieve accurate joint rotation angles and believable hand shapes, and the latter calibrates the glove sensor readings to achieve accurate fingertip positions.

The default calibration application [Yazadi 2009] shipped with CyberGlove uses a simple linear method to fit a calibration line to 2 sampled hand poses for each sensor. However, the sampled “flat” and “okay” poses do not involve enough reading changes to calibrate all sensors. Thus another manual parameter tuning process is required to improve calibration accuracy (36 parameters for 18 sensors). CyberGlove calibration only happens in joint rotation thus it can only achieve FK accuracy. [Huenerfauth and Lu 2010] also uses a FK linear method to make their glove calibration more accessible to deaf subjects. Glove calibration in [Chou et al. 2000; Menon et al. 2003] uses linear regression and thus is less subject

to noise in the sampled hand poses. The problems with most of the FK calibration methods are: 1. No concern for the IK fingertip accuracy. 2. All sensors are assumed to be independent of each other and calibrated separately, which falsely lowers the calibration problem from HD space to 1-D space. 3. Due to the difficulty of hand pose measurement, usually sample quantity is insufficient.

The official CyberGlove documentation [Yazadi 2009] provides sensor resolution, linearity and noise descriptions, which explains the possible inaccuracies caused by linear calibration. Evaluation work in [Kessler et al. 1995] reports that sensor readings are complicated: readings vary based on sensor resolution, hand size, pressure on the sensors, human variations and noise in the system. Their experiments verified that stretching of neighboring sensors can cause errors of more than 20° for abduction rotation angles. Previous work by [Kahlesz et al. 2004; Steffen et al. 2011] also reports severe coupling problems in their experiments. Their calibration methods synthesize training data using parabolic functions, to avoid too much sampling work. In [Kahlesz et al. 2004] and [Jin et al. 2010], calibration for cross-coupling effect is treated as an interpolation problem with no consideration for the possible noise in the samples. Calibration research [Chou et al. 2000] also addresses the multiple correspondences between sensors and joint rotations. They use an external vision system to extract relevant features, and calibrate sensor readings by a multiple linear regression method. Fundamentally, these cross-coupling calibration methods only adjust joint rotations to achieve better hand shape, but offer no guarantee of IK fingertip accuracy.

Glove calibration for tele-operation of a robot hand [Griffin et al. 2000; Hu et al. 2004] uses the pseudo inverse of Jacobian as the IK solver to achieve accurate fingertip positions for object manipulation. [Griffin et al. 2000] approximates the ideal fingertip positions without manual measurement, while a vision system is applied in [Hu et al. 2004] to track the real fingertip positions for calibration. Both work assume exact sensor linearity and independence: each sensor’s linear calibration function is embedded into the Jacobian matrix, which causes frequent generation of trivial solutions. Both methods use an unconstrained IK solver, thus there is no guarantee of plausible hand shapes. Motion smoothness is not considered either.

Acquiring hand motion has been discussed in [Zhao et al. 2012], which does not utilize gloves, but combines two optical systems for the capture. In [Ye and Liu 2012], wrist movements and the motion of handled objects are optically captured to reconstruct the real hand motion. [Kang et al. 2012] reconstructs high quality hand motion by using a reduced marker set for the hand motion capture. The recent hand motion research [Jörg et al. 2012] synthesizes realistic finger motion for gesturing characters from a pre-captured hand motion database.

3 Glove Hardware and Kinematic Hand Model

3.1 Glove and Sensors

The glove hardware used in this paper is the CyberGlove II, an 18-sensor glove where sensors are located over or near the joints to measure the hand poses. The detailed sensor layout is shown in Figure 2. For the sensor and rotation naming, T, I, M, R, P, W before the underscore indicate thumb, index, middle, ring, pinky and wrist respectively; the joint name is after the underscore.

Ideally sensors should provide readings only due to changes in the joint angle it is designed to measure, not other movements. Based on this independence assumption, calibration for each sensor can

Sensor Layout:

1. s_{T_TMC}
2. s_{TL_ABD}
3. s_{T_MCP}
4. s_{T_IP}
5. s_{L_MCP}
6. s_{L_PIP}
7. s_{LM_ABD}
8. s_{M_MCP}
9. s_{M_PIP}
10. s_{R_MCP}
11. s_{R_PIP}
12. s_{MR_ABD}
13. s_{P_MCP}
14. s_{P_PIP}
15. s_{RP_ABD}
16. s_{PALM_ARCH}
17. s_{W_FLEX}
18. s_{W_ABD}

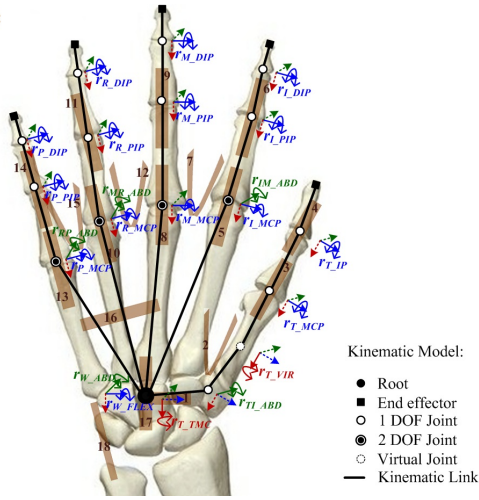


Figure 2: Sensor Layout and Kinematic Hand Model. All 18 sensors are listed and kinematic joint rotations are marked.

be conducted in 1-D space. Unfortunately in practice, neighboring sensors stretch each other and create the cross-coupling effect. The coupling problem, according to [Huenerfauth and Lu 2010; Kahlesz et al. 2004; Chou et al. 2000; Jin et al. 2010], is not peculiar to CyberGlove, but common to data gloves on the market due to the similar sensor layout. We completed a thorough examination of all the sensors using a similar method introduced in [Kahlesz et al. 2004]: keep the joint rotation unchanged and record reading variations from its measuring sensor (ideally variations should be 0 for independent sensors) when freely moving other speculative affecting sensors to the largest possible extent. In practice, reading variations from measuring sensors never remain exactly 0, so sensors are regarded as cross-coupled with other affecting sensors if variation $\text{std.dev} \geq 5\%$ or $\text{max.dev} \geq 12\%$ of the joint’s active range. Otherwise, the variations are negligible and the sensor is defined as independent. The results from our independence examination are listed in Table 1. Note that the coupling problem is not bi-directional.

For sensor linearity and noise, the official CyberGlove specification [Yazadi 2009] claims sensor resolution is approximately 1° , repeatability $\pm 3^\circ$ and maximum nonlinearity of 0.6%. Other research by [Chao 2001] reports mean measurement errors of $1.7^\circ \pm 1.5^\circ$, and [Kessler et al. 1995] reports mean sensor noise to be 1.1° , and a max noise of 3° .

Independence and exact linearity assumptions with no concern for coupling effects and noise cause inaccuracies in linear calibrations.

3.2 Kinematic Hand Model

We design the kinematic hand model to be consistent with the actual hand skeletal structure and also to match the glove sensor layout. Each finger is modeled as a kinematic chain that roots at the palm base. Most of the measuring sensor is conveniently located at a joint and can directly measure the joint rotation. There are some exceptions we need to pay attention to. The 18-sensor glove does not place distal interphalangeal (DIP) sensors on the 4 fingers, thus the finger DIP rotations are not measured. Also to account for the thumb pronation, we adopt a common solution as described in [Kahlesz et al. 2004; Griffin et al. 2000; Hu et al. 2004; Turner 2001]: a virtual joint is added to the thumb kinematic chain to ensure anatomical correctness, which does not have a direct measuring sensor either. The palm-arc sensor measures palm movement that can be formed by more than one joint rotation at the roots of

Table 1: Results from Independence Examination. To prove the individual relevance to the coupling effect, for each affecting sensor, its own contribution is also listed, in the format (% of std.dev, % of max.dev).

Measured Rotation r_m	Measuring Sensor s_m	Affecting Sensors $\{s_{a1}, \dots, s_{aK}\}$	%vari std.dev	%vari max.dev
r_{TL_ABD}	s_{TL_ABD}	s_{T_TMC}	9.6%	25.4%
r_{IM_ABD}	s_{IM_ABD}	s_{L_MCP} (22.2%,33.8%) s_{M_MCP} (13.2%,54.1%)	45.7%	114.2%
r_{MR_ABD}	s_{MR_ABD}	s_{M_MCP} (34.4%,92.1%) s_{R_MCP} (42.1%,115.5%)	42.5%	120.4%
r_{PM_ABD}	s_{PM_ABD}	s_{R_MCP} (27.8%,69.2%) s_{P_MCP} (36.0%,80.5%)	22.7%	93.8%
Palmar movement ¹	s_{PALM_ARC}	s_{T_TMC} (11.5%,37.5%) s_{TL_ABD} (27.4%,67.0%) s_{L_MCP} (4.3%,15.5%) s_{IM_ABD} (5.3%,14.2%) s_{M_MCP} (8.6%,36.5%) s_{R_MCP} (20.9%,49.7%) s_{MR_ABD} (6.7%,24.3%) s_{P_MCP} (33.9%,73.0%) s_{RP_ABD} (4.2%,16.2%)	21.2%	76.1%
r_{W_FLEX}	s_{W_FLEX}	s_{W_ABD}	13.6%	37.5%
r_{W_ABD}	s_{W_ABD}	s_{W_FLEX}	6.5%	26.3%

all the finger kinematic chains.

3.3 Sensor-Rotation Correspondence

We generalize the mappings between sensor readings and joint rotations in

$$\mathbf{r} = \tau(\mathbf{s}) \quad (1)$$

where \mathbf{s} denotes readings from all involved sensors, \mathbf{r} is the actual joint rotations, the mapping function τ can be regarded as the calibration function. The purpose of glove calibration is to find the appropriate form of τ and finalize all parameters in it so that it can produce accurate rotation angles. Under this definition, linear calibration has prematurely fixed a linear form for τ even before observations of the data.

In our research, \mathbf{r} and \mathbf{s} in (1) are defined differently according to the type of Sensor-Rotation correspondence²:

- 1 Sensor to 1 Joint Rotation(**ISIR**): Measuring sensor s_m independently measures joint rotation r_m . In this case, $\mathbf{r} = r_m$, $\mathbf{s} = s_m$.
- Multiple Sensors to 1 Joint Rotation(**MSIR**): Measuring sensor s_m is cross-coupled with a set of K affecting sensors $\{s_{a1}, s_{a2}, \dots, s_{aK}\}$ when measuring rotation r_m . In this case, $\mathbf{r} = r_m$, $\mathbf{s} = \{s_m, s_{a1}, s_{a2}, \dots, s_{aK}\}$ and the latent calibration function τ we need to infer is in a higher dimensional space. Typically for abduction between index and middle finger, $r_m = r_{IM_ABD}$, $s_m = s_{IM_ABD}$, $K = 2$ where $s_{a1} = s_{L_MCP}$ and $s_{a2} = s_{M_MCP}$, thus the mapping is in the form $r_{IM_ABD} = \tau(s_{IM_ABD}, s_{L_MCP}, s_{M_MCP})$. Our experiments also prove that calibrating thumb joint rotations in HD space $\{s_{T_TMC}, s_{TL_ABD}, s_{T_MCP}, s_{T_IP}\}$ can generate more accurate results, due to the complexity of thumb structure.

¹Based on the exam, palm-arc sensor is affected by other 9 sensors. The severe coupling effects suggest it is unlikely to provide accurate data. Thus in this paper, we choose not to use this data for calibration. The palm is treated as a rigid body.

²Under this categorization, the palm-arc sensor should belong to Multiple Sensors to Multiple Joint Rotations(**MSMR**), if considered.

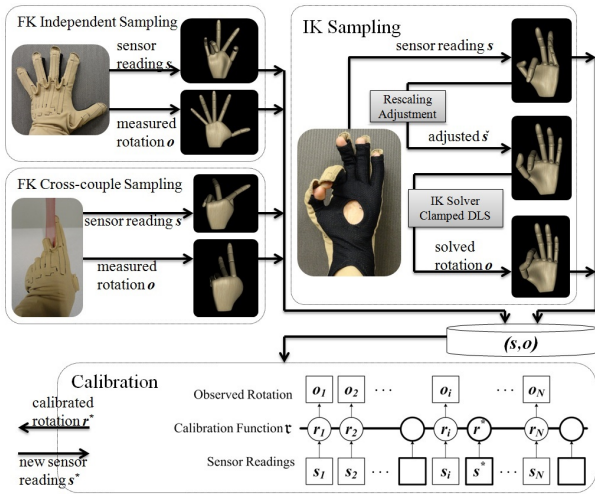


Figure 3: Overall process of our data-driven glove calibration.

- 0 Sensor to 1 Joint Rotation(**0S1R**): Joint rotation r_m has no measuring sensors. In this case, $\mathbf{r} = r_m$, \mathbf{s} uses pseudo sensor data \dot{s}_m simulated from existing sensor readings, based on well recognized biomechanics and anatomical principles. Then **0S1R** can be treated as **1S1R**. Typically for index DIP, $\mathbf{r} = r_{I_DIP}$, $\mathbf{s} = \dot{s}_{I_DIP} = \frac{2}{3} * s_{I_PIP}$ according to findings in [Chou et al. 2000; Kuch and Huang 1994; Lee and Kunii 1995; Pavlovic et al. 1997]. And for the added virtual joint on thumb, $\mathbf{r} = r_{T_VIR}$, $\mathbf{s} = \dot{s}_{T_VIR} = a * s_{T_TMC} + b * s_{T_LABD}$ based on practices from [Kahlesz et al. 2004; Griffin et al. 2000; Hu et al. 2004], where a and b are constant coefficients.

4 Overview

The goal of glove calibration is to infer the latent mapping function τ in (1), given 2 accuracy requirements:

- FK accuracy defines joint rotation accuracy in the joint rotation space(θ -space). The calibrated \mathbf{r} should reflect accurate joint rotations and sensible hand shapes.
- IK accuracy defines fingertip position accuracy in the end-effector space(l -space), when the hand is configured as the calibrated rotations \mathbf{r} . This is important for finger-touching motion and object manipulation.

Our calibration method is data-driven and initiated with a sampling process (Figure 3). The goal of the sampling process is to collect data for the calibration process to handle all sensor-rotation correspondence types and achieve both accuracy requirements. The data is collected in the format of tuple (\mathbf{s}, \mathbf{o}) where $\mathbf{o} = \mathbf{r} + \epsilon$, is the observed joint rotations with error ϵ from measurement or calculation. The sampling process makes no premature assumption of a fixed form of τ . Based on sensor-rotation correspondences and accuracy requirements, the sampling process includes: FK Independent Sampling, FK Cross-couple Sampling and IK Sampling.

During the calibration process, our LMC-GPR model is trained based on the collected data tuples (\mathbf{s}, \mathbf{o}) , and the unobserved calibration function τ is inferred. For new incoming sensor reading \mathbf{s}_* , τ can predict the real joint rotation \mathbf{r}_* . The general GPR framework can handle more flexible forms of τ in higher dimensional space with consideration of errors in the observation \mathbf{o} to achieve the dual accuracy goals. Sparse approximation SPGPs is used for real-time calibration.

5 Data Sampling

5.1 FK Sampling

5.1.1 FK Independent Sampling

FK independent sampling collects data from the independent sensors with **1S1R** correspondence. Thus the (\mathbf{s}, \mathbf{o}) tuples it collects are (s_m, o_m) , which makes the sampling very straightforward. We predefine several key poses (Figure 4), record readings from the independent sensors when performing these poses as s_m , and the corresponding measured joint rotation as o_m . The key poses we currently use are: “flat”, “spread”, “fist” and “overbend”, which ensure at least 2 different rotation angles for each independent sensor. More key poses can be added for specific accuracy requirements. Adding a sampled key pose tuple (s_m, o_m) will cause the calibration function to yield joint rotation close to o_m for readings s_m .

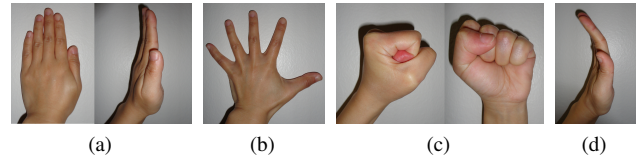


Figure 4: Predefined Key Poses for FK Independent Sampling.(a)“flat” (b)“spread” (c)“fist” (d)“overbend”

5.1.2 FK Cross-couple Sampling

FK cross-couple sampling is designed for cross-coupled sensors with **MS1R** correspondence. Readings from both the measuring sensor and the affecting sensors need to be recorded. Thus the (\mathbf{s}, \mathbf{o}) tuples collected are $(s_m, s_{a1}, \dots, s_{aK}, o_m)$. Joint rotation o_m is controlled by using extra widgets. The general idea of the cross-couple sampling is:

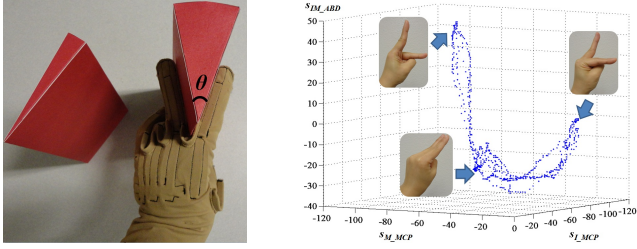
- Fix o_m at rotation angle θ using facilitating widgets.
- Keep o_m still while freely move the rest of the hand to change sensors readings of $\{s_m, s_{a1}, \dots, s_{aK}\}$.
- Record sensor readings during the free movement together with joint rotation in tuples $(s_m, s_{a1}, \dots, s_{aK}, o_m)$
- Change o_m to another rotation angle, goto step a and repeat.

This sampling method applies to all cross-coupled cases listed in Table 1. Figure 5 illustrates the details of sampling s_{IM_ABD} , which is affected by s_{I_MCP} and s_{M_MCP} . A widget with a known angle θ is put inbetween the index and middle finger to keep abduction o_{IM_ABD} fixed; freely flexing the index and middle finger produces different readings from s_{I_MCP} , s_{M_MCP} , and due to the stretching effect, readings from s_{IM_ABD} will also change. Samples are recorded and the process is repeated for 0° , 10° , 20° and 30° as the fixed o_{IM_ABD} value.

5.2 IK Sampling

5.2.1 Sampling Process

The goal of IK sampling is to collect training data that allows the calibration process to achieve fingertip accuracy. The process begins with the user performing a hand motion sequence in which fingertip accuracy is regarded as important, such as the finger touching motion (shown in Figure 6). For each frame in the motion, the ideal fingertip touching positions \mathbf{d} is automatically calculated



(a) Collecting sensor readings from s_{IM_ABD} , s_{I_MCP} , s_{M_MCP} with fixed θ_{IM_ABD} .

(b) Plot of collected tuples $(s_{IM_ABD}, s_{I_MCP}, s_{M_MCP}, \theta_{IM_ABD})$ in one motion clip with θ_{IM_ABD} controlled by wigget $\theta = 0$.

Figure 5: Cross-couple sampling for abduction between index and middle.

based on the probability distribution in l -space. Clamped Damped Least Square(DLS) with hand shape constraints is used to solve for joint rotations θ that can reach d . Original sensor readings s and solved joint rotations θ are recorded in tuple (s, θ) . s includes readings from the measuring sensor and affecting sensors for coupled cases.



Figure 6: IK Sampling - finger touching motion sequence

5.2.2 IK Solver Clamped DLS with Constraints

The IK solver should ensure accurate fingertip positions, sensible hand shapes, and smooth solutions for the sampled motion sequence. Based on previous IK research [Buss 2004; Aristidou and Lasenby 2009; Buss and Kim 2005; Nakamura and Hanafusa 1986; Wampler 1986; Chan and Lawrence 1988], we use the Clamped DLS to satisfy these requirements.

To provide a sensible initial configuration for the IK solver, sensor readings s will be linearly rescaled to the joint's active range as \tilde{s} . Finger configurations θ are initialized from \tilde{s} , under which fingertip positions are l . Then θ will be adjusted by (2) in an iterative process until fingertips reach the ideal positions d .

$$\Delta\theta = J^{+\lambda}e + (I - J^+J)z \quad (2)$$

$J^{+\lambda}$ is defined as $J^T(JJ^T + \lambda^2I)^{-1}$, based on Jacobian matrix J . We set the damping factor λ proportional to $\|d - l\|^2$. e is the distance from current fingertip positions l to the ideal d clamped by Eq.(3) to suppress radical adjustment in one iteration step. For each joint, we define hand shape constraints as $z = \nabla c(\theta - \bar{\theta})^2$, where $\bar{\theta}$ is set to be the middle value of joint's active range to penalize θ 's excessive deviation. c is the stiffness coefficient. We assign relatively lower stiffness to the thumb than to other fingers. J^+ is the pseudo inverse of Jacobian. Thus z is projected to the nullspace of J . For θ that is already out of range, we enforce a hard constraints as introduced in [Meredith and Maddock 2005] through $\Delta\theta = w\Delta\theta, 0 \leq w < 1$ to effectively suppress further deviation from its active range. Final θ solutions are recorded as the observed rotations θ in tuple (s, θ) .

$$\text{clamp}(d-l, e_{max}) = \begin{cases} d-l & \text{if } \|d-l\| < e_{max} \\ e_{max} * \frac{d-l}{\|d-l\|} & \text{otherwise.} \end{cases} \quad (3)$$

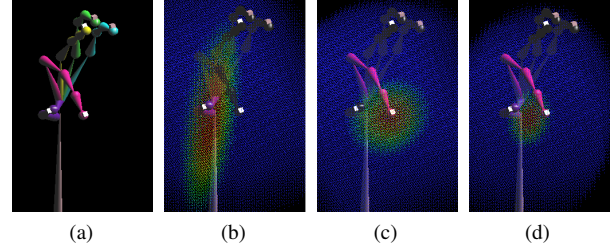


Figure 7: Visualization of probability distribution in l -space for thumb-index touching pose.(a) The configuration from \tilde{s} . (b) The visualization of thumb tip probability distribution in l -space with color map. (c) The visualization of index tip probability distribution. (d) The visualization of the joint probability distribution in l -space for the common thumb-index touching position.

5.2.3 Automatic Calculation of Fingertip Touching Position

We avoided the use of an expensive external vision system to track fingertips in l -space, nor did we manually measure fingertip positions. Instead, we automatically calculate the fingertip touching positions based on the probability distribution in l -space.

First, for each finger u , we define the probability distribution of joint rotations in θ -space as p_u in (4):

$$p_u(\theta) = \prod_{v=1}^V p_v(\theta_v) \quad (4)$$

where θ_v is the rotation of joint v on the finger u . p_v is modeled as a Gaussian distribution $\mathcal{N}(m_v, \sigma_v)$. m_v is the initial configuration from \tilde{s} , σ_v is set to 3° for independent sensors and 6° or higher for coupled sensors and thumb sensors (based on the official technical specifications [Yazadi 2009] and the experimental results in [Kessler et al. 1995]).

The probability distribution of u 's fingertip position in l -space p'_u can be calculated from its joint rotation distributions in θ -space in (5). θ_{l_u} is the output solution from our IK solver to make u 's fingertip reach l_u in the l -space.

$$p'_u(l_u) = p_u(\theta_{l_u}) \quad (5)$$

For the finger touching pose, all involved fingertips reach the common position d . We find d by maximizing the joint probability in l -space for all the involved fingers (6). Typically $U = 2$. Figure 7 illustrates the thumb-index touching case.

$$\arg \min_d \left\{ - \sum_{u=1}^U \log p'_u(l_u = d) \right\} \quad (6)$$

6 Glove Calibration

We train the LMC-GPR calibration model using a training set $T = \{(s_n, \theta_n)\}_{n=1}^N$ of N tuples, consisting of the three sample types described above. Input sensor reading s_n is in D dimensional space. We reorganize T by aggregating the N sensor inputs as column vectors in an $D \times N$ matrix s . The corresponding N observations are in vector $\theta = (\theta_1, \theta_2, \dots, \theta_N)$. GPR infers the latent calibration function τ , and provides calibrated rotation r_* for new sensor reading s_* . SPGPs proposed by [Snelson and Ghahramani 2006; Lawrence 2007] are further used to enhance the real-time performance of our calibration.

6.1 Glove Calibration using LMC-GPR

The LMC-GPR model is specially designed for glove data based on general Gaussian Process Regression (thoroughly discussed in [Rasmussen and Williams 2006]). It utilizes the sensor linearity yet does not restrict τ to be of linear form, or any fixed form from a limited number of basis functions.

6.1.1 GPR Model Selection

The Gaussian process $\tau \sim GP(m(\mathbf{s}), k(\mathbf{s}, \mathbf{s}'))$ can be specified by its mean function $m(\mathbf{s}) = E[\tau(\mathbf{s})]$ and covariance function $k(\mathbf{s}, \mathbf{s}') = E[(\tau(\mathbf{s}) - m(\mathbf{s}))(\tau(\mathbf{s}') - m(\mathbf{s}'))]$. As closer \mathbf{s} are more likely to have similar \mathbf{r} , $k(\mathbf{s}, \mathbf{s}')$ measures the similarity between \mathbf{s} and \mathbf{s}' in the input space and thus \mathbf{r} from closer \mathbf{s} are regarded as more informative during prediction.

The general Gaussian process commonly uses zero mean m_0 and Radial Basis Function k_{RBF} as the covariance function for lack of prior knowledge, which is not an appropriate model for glove calibration for several reasons: 1. k_{RBF} is translation-invariant in input space and depends only on the distance between \mathbf{s} , which is inconsistent with sensor linearity. For coupled input $\mathbf{s} = (s_m, s_{a1}, \dots, s_{aK})$, the stretching effect also changes when \mathbf{s} translates in the input space. 2. k_{RBF} is isotropic using one length-scale hyperparameter to describe the relevance for all dimensions in the input space, but the coupling problem is anisotropic. For joint rotation r (e.g. r_{IM_ABD}), the relevance of readings from measuring sensor s_{IM_ABD} and affecting sensors s_{LM_CP} , s_{SM_MCP} are different. 3. According to our experiments, k_{RBF} takes longer to train and to predict. Thus we use the LMC-GPR model for glove calibration. Detailed GPR model comparisons for glove calibration are listed in Table 2.

6.1.2 LMC-GPR Mean and Covariance

Based on the general idea of sensor linearity, we explicitly specify the mean function of LMC-GPR as linear mean m_l

$$m_\tau = m_l(\mathbf{s}) = \sum_{d=1}^D \alpha_d s_d \quad (7)$$

where α_d are constant hyperparameters.

A non-trivial composite covariance function k_τ is designed by using linear kernel k_{LARD} and squared exponential kernel k_{seARD} as building blocks:

$$\begin{aligned} k_\tau &= k_{LARD} + k_{seARD} \\ &= \sum_{d=1}^D \beta_d s_d s'_d + \eta \exp\left(-\frac{1}{2} \sum_{d=1}^D \gamma_d (s_d - s'_d)^2\right) \end{aligned} \quad (8)$$

β_d , η and γ_d are hyperparameters in k_τ . k_{LARD} corresponds to the covariance function of linear regression and thus can be used to model the basic sensor linearity. k_{seARD} is more flexible to model the nonlinear part in the training data, including the coupling effect and noise. The Automatic Relevance Determination (ARD) technique uses different coefficient β_d and length-scale γ_d for dimension d to determine the relevance of each input dimension.

6.1.3 LMC-GPR Training

Based on the (m_τ, k_τ) model, the Gaussian prior of LMC-GPR is defined in (9), where K_τ is the $N \times N$ Gram matrix with entry $K_\tau[i, j] = k_\tau(\mathbf{s}_i, \mathbf{s}_j)$ for all pairs of \mathbf{s} in the training set T .

$$p(\mathbf{r}|\mathbf{s}) = \mathcal{N}(\mathbf{r}|m_\tau, K_\tau) \quad (9)$$

Table 2: Model selection experiments applied to 2000+ glove training data (s_{IM_ABD} , s_{LM_MCP} , s_{SM_MCP} , o_{IM_ABD}) collected from FK cross-couple and IK sampling. We record average training time \bar{t}_{tr} and prediction time \bar{t}_{pred} to predict the r_{IM_ABD} for a 267-frame clip. To measure model fitness, the negative log marginal likelihood (NLML) and noise's log standard deviation $\log \sigma$ are also recorded, which for a well-fit model should be minimized. The general purpose model (m_0, k_{RBF}) takes longer to train and to predict, and achieves acceptable fitness. Change m_0 to m_l , (m_l, k_{RBF}) has smaller NLML and thus fits better. Using linear kernel k_{LARD} corresponds to linear regression. With either m_0 or m_l , k_{LARD} has shorter \bar{t}_{tr} and \bar{t}_{pred} , but the large NLML and $\log \sigma$ indicate poor fitness and the unfitted parts are interpreted as noise. ($m_l, k_{LARD} + k_{seARD}$) is the LMC-GPR model we use for glove calibration. It appropriately integrates the prior knowledge of sensor linearity into a flexible model, fits the glove data better, and has \bar{t}_{tr} and \bar{t}_{pred} at relatively the magnitude of a simple model.

m_τ	k_τ	\bar{t}_{tr}	\bar{t}_{pred}	NLML	$\log \sigma$
m_0	k_{RBF}	752.3s	7.1s	5.795e+03	-1.79
m_0	k_{LARD}	197.5s	0.55s	7.641e+03	2.39
m_0	$k_{seARD} + k_{seARD}$	4569.2s	16.5s	4.962e+03	-1.79
m_0	$k_{LARD} + k_{seARD}$	333.8s	0.97s	4.545e+03	-1.81
m_l	k_{RBF}	2615.5s	23.4s	4.973e+03	-1.80
m_l	k_{LARD}	219.2s	0.68s	7.635e+03	2.39
m_l	$k_{seARD} + k_{seARD}$	3760.3s	8.2s	4.600e+03	-1.80
m_l	$k_{LARD} + k_{seARD}$	331.2s	0.81s	4.543e+03	-1.81

Noise ε in the observations has the Gaussian distribution $\mathcal{N}(\varepsilon|0, \sigma^2 I)$ where σ is the hyperparameter indicating the variance of the noise. Thus the likelihood of observing \mathbf{o} given \mathbf{r} is defined as

$$p(\mathbf{o}|\mathbf{r}) = \mathcal{N}(\mathbf{o}|\mathbf{r}, \sigma^2 I) \quad (10)$$

The training process learns the values of all hyperparameters $\mathbf{h} = \{\alpha_d, \beta_d, \eta, \gamma_d, \sigma\}_{d=1}^D$ in the LMC-GPR model from the training set T . Based on the definition of prior and likelihood, integrating out the latent \mathbf{r} , we can obtain the log marginal likelihood as:

$$p(\mathbf{o}|\mathbf{s}, \mathbf{h}) = \mathcal{N}(\mathbf{o}|m_\tau, K_\tau + \sigma^2 I) \quad (11)$$

The marginal likelihood tells the probability of the observations given the assumptions of the mean and covariance and thus it measures the fitness of the model. We find the value of \mathbf{h} that can best explain the observations in T by maximizing (11) i.e. minimizing the negative log-marginal likelihood (NLML). Partial derivatives of the NLML w.r.t \mathbf{h} are calculated for the gradient-based optimization.

$$\frac{\partial \ln p(\mathbf{o}|\mathbf{s}, \mathbf{h})}{\partial h_j} = \frac{1}{2} \mathbf{y}^T K_o^{-1} \frac{\partial K_o}{\partial h_j} K_o^{-1} \mathbf{y} - \frac{1}{2} \text{tr}(K_o^{-1} \frac{\partial K_o}{\partial h_j}) \quad (12)$$

where $\mathbf{y} = \mathbf{o} - m_\tau$, $K_o = K_\tau + \sigma^2 I$. The complexity of this training process is dominated by the inversion of K_o , which is $O(N^3)$.

6.1.4 LMC-GPR Prediction

After all hyperparameters have been determined, given new incoming sensor readings \mathbf{s}_* as inputs, the trained LMC-GPR can predict the calibrated joint rotation \mathbf{r}_* based on training set T . The predictive distribution is

$$p(\mathbf{r}_*|\mathbf{s}_*, T) = \mathcal{N}(\mathbf{r}_*|m_{\tau*}, K_{\tau*}) \quad (13)$$

where

$$m_{\tau*} = K_\tau(\mathbf{s}_*, \mathbf{s}) [K_\tau(\mathbf{s}, \mathbf{s}) + \sigma^2 I]^{-1} \mathbf{o} \quad (14)$$

$$K_{\tau*} = K_\tau(\mathbf{s}_*, \mathbf{s}_*) - K_\tau(\mathbf{s}_*, \mathbf{s}) [K_\tau(\mathbf{s}, \mathbf{s}) + \sigma^2 I]^{-1} K_\tau(\mathbf{s}, \mathbf{s}_*)$$

(14) is the calibration function τ we need for glove calibration. Because of the matrix inversion operation, this prediction process also requires $O(N^3)$ computation. Cholesky decomposition is used for matrix inversion.

6.2 Real-time Optimization

One difficulty of applying the LMC-GPR to real-time glove calibration is its $O(N^3)$ complexity. This performance problem can be overcome by a sparse approximation method using a pseudo training set \hat{T} of size M , $M \ll N$. Pseudo sensor inputs $\hat{\mathbf{s}}$ decide \hat{T} 's location in T and an optimization process will help find $\hat{\mathbf{s}}$ for \hat{T} to best represent the entire training set. As the values of hyperparameters \mathbf{h} in m_τ and k_τ depend on pseudo-set's location $\hat{\mathbf{s}}$, separating hyperparameter learning from pseudo-set selection causes fluctuations in the marginal likelihood and its gradients, and leads to unsmooth convergence during the optimization process. Thus we use sparse approximation method SPGPs proposed in [Snelson and Ghahramani 2006], which finds $\hat{\mathbf{s}}$ and \mathbf{h} together in one smooth joint optimization.

Given $\hat{\mathbf{s}}$, the prior and likelihood are defined as (15) and (16)

$$p(\hat{\mathbf{r}}|\hat{\mathbf{s}}) = \mathcal{N}(\hat{\mathbf{r}}|m_\tau(\hat{\mathbf{s}}), K_\tau(\hat{\mathbf{s}}, \hat{\mathbf{s}})) \quad (15)$$

$$p(\mathbf{o}|\mathbf{s}, \hat{\mathbf{s}}, \hat{\mathbf{r}}) = \mathcal{N}(\mathbf{o}|K_\tau(\mathbf{s}, \hat{\mathbf{s}})K_\tau(\hat{\mathbf{s}}, \hat{\mathbf{s}})^{-1}\hat{\mathbf{r}}, \Lambda + \sigma^2 I) \quad (16)$$

where Λ is $N \times N$ diagonal matrix with entry $\lambda_{nn} = K_\tau(\mathbf{s}_n, \mathbf{s}_n) - K_\tau(\mathbf{s}_n, \hat{\mathbf{s}})K_\tau(\hat{\mathbf{s}}, \hat{\mathbf{s}})^{-1}K_\tau(\hat{\mathbf{s}}, \mathbf{s}_n)$. Marginalizing over $\hat{\mathbf{r}}$, we obtain the marginal likelihood:

$$p(\mathbf{o}|\mathbf{s}, \hat{\mathbf{s}}, \mathbf{h}) = \mathcal{N}(\mathbf{o}|m_\tau, K_\tau(\mathbf{s}, \hat{\mathbf{s}})K_\tau(\hat{\mathbf{s}}, \hat{\mathbf{s}})^{-1}K_\tau(\hat{\mathbf{s}}, \mathbf{s}) + \Lambda + \sigma^2 I) \quad (17)$$

Location $\hat{\mathbf{s}}$ together with the value of \mathbf{h} can be obtained by maximizing the marginal likelihood (17) w.r.t $\{\hat{\mathbf{s}}, \mathbf{h}\}$.

The predictive distribution of this sparse model is:

$$p(\mathbf{r}_*|\mathbf{s}_*, T, \hat{\mathbf{s}}) = \mathcal{N}(\mathbf{r}_*|\hat{m}_{\tau*}, \hat{K}_{\tau*}) \quad (18)$$

where

$$\hat{m}_{\tau*} = K_\tau(\mathbf{s}_*, \hat{\mathbf{s}})Q^{-1}K_\tau(\hat{\mathbf{s}}, \mathbf{s})(\Lambda + \sigma^2 I)^{-1}\mathbf{o} \quad (19)$$

$$\hat{K}_{\tau*} = K_\tau(\mathbf{s}_*, \mathbf{s}_*) - K_\tau(\mathbf{s}_*, \hat{\mathbf{s}})(K_\tau(\hat{\mathbf{s}}, \hat{\mathbf{s}})^{-1} - Q^{-1})K_\tau(\hat{\mathbf{s}}, \mathbf{s}_*)$$

$Q = K_\tau(\hat{\mathbf{s}}, \hat{\mathbf{s}}) + K_\tau(\hat{\mathbf{s}}, \mathbf{s})(\Lambda + \sigma^2 I)^{-1}K_\tau(\mathbf{s}, \hat{\mathbf{s}})$. Thus (19) is the approximated calibration function using a size M pseudo-set. SPGPs has $O(M^2N)$ training cost and $O(M^2)$ prediction cost. Our experiments show that using \hat{T} of size $M = 100$ to approximate T of size $N = 2000+$ can still produce excellent calibration results. Details of these results are discussed in Section 7.

7 Results and Applications

7.1 Experimental Evaluation of Accuracy

We evaluate the effectiveness of our approach by using the LMC-GPR calibration model trained on a data set consisting of 2000+ tuples. For sparse approximation, a pseudo-set of size $M = 100$ is used for SPGPs. To demonstrate the advantage of LMC-GPR calibration, we compare its results with the most commonly used linear calibration (regressed among the key poses defined in Section 5).

7.1.1 Perceptual Evaluation

Perceptual evaluation focuses on checking the plausibility of hand shapes. For key poses, all calibration methods can provide ideal results as the key poses are part of the training set. In Figure 8, the calibrated hand shapes for some commonly used hand poses (not in the key pose set) are shown. In Figure 9, we illustrate thumb-finger touching poses where both fingertip positions and hand shapes are important.

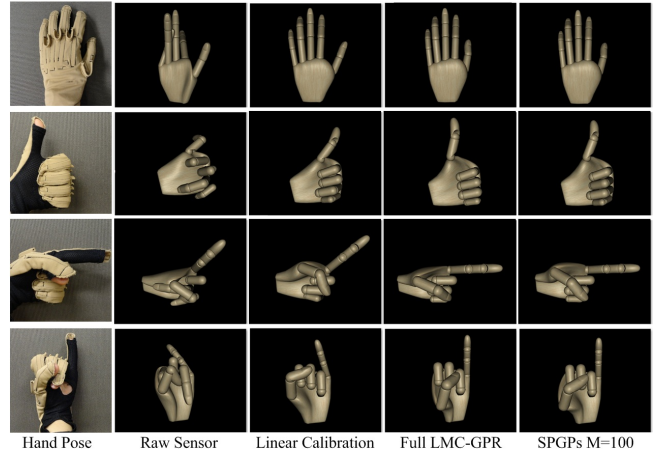


Figure 8: FK hand shape accuracy evaluation using different calibration methods. 4 commonly used hand poses are shown. The cross-coupled stretching effects (e.g. r_{IM_ABD} in "pointing" pose) are fixed by LMC-GPR and its SPGPs approximation.



Figure 9: IK fingertip accuracy evaluation using different calibration methods.

We put 12 hand motions calibrated in 4 different ways (totally 48 clips) on Mechanical Turk. 30 users watched the ground truth first and then rated the calibrated motions, shown in random order (1 indicating the lowest resemblance, and 7 highest). The ratings demonstrate that LMC-GPR best resembles the ground truth. The result from a one-way ANOVA run between linear, LMC-GPR and SPGPs on the ratings is significant ($p < 0.001$, $F = 154.03$, $df = 2$), as shown in Figure 10.

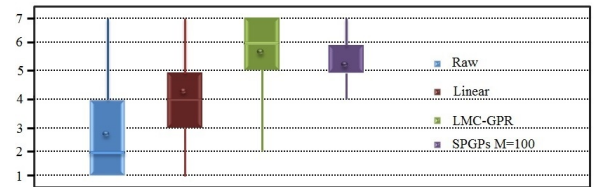


Figure 10: Resemblance ratings of calibrated hand motion clips.

7.1.2 Quantitive Evaluation

To quantify the FK accuracy of different calibration methods, we compare the calibrated joint rotations against the measured rotation

angles for key hand poses and other error-prone poses (Table 3). We also designed experiments to further evaluate the correction for cross-coupling problem, where users are asked to perform free finger movements with fixed index-middle abduction. Figure 11 shows the calibrated r_{IM_ABD} with the real θ_{IM_ABD} fixed to 0° (in the training set), and 25° (not in the training set), respectively. Raw sensor readings and linear calibration produces unstable abduction values due to the sensor stretching, which is fixed by full LMC-GPR and SPGPs. For $\theta_{IM_ABD} = 25^\circ$, the predicted values are slightly smaller than real abduction.

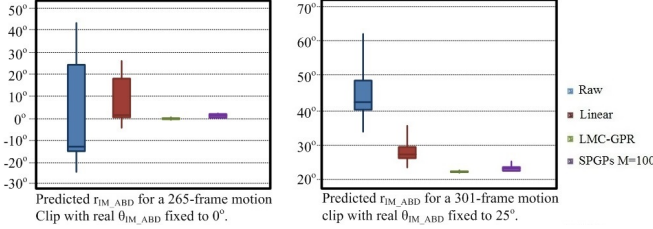


Figure 11: Correction of cross-coupling effects.

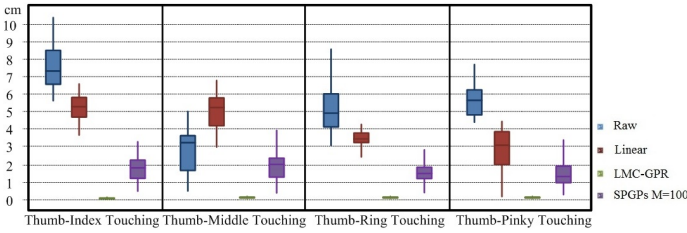
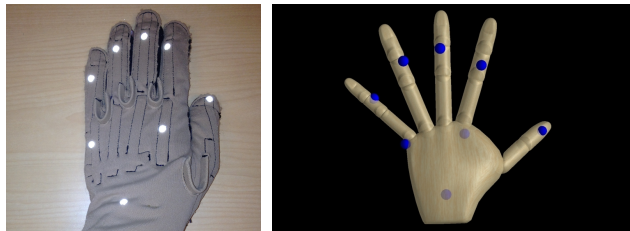


Figure 12: Boxplot of involved fingertip distances in a continuous touching motion, using different calibration methods.

To quantify IK fingertip accuracy, we record the distances between the involved fingertips when they are performing continuous touching motions. The ideal distance should be 0. Figure 12 compares the fingertip distances resulting from different calibration methods, e.g. for the 176-frame thumb-index touching motion, the average distance between thumb tip and index tip is about 7.53cm, 5.25cm, 0.08cm and 1.75cm using raw data, linear calibration, LMC-GPR and SPGPs respectively. Similarly for thumb-mid, thumb-ring and thumb-pinky touching motions, the full LMC-GPR provides significantly better results than linear calibration. Also both Figure 9 and Figure 12 show that SPGPs can yield improved results compared to linear calibration.



(a) Marker placement for the capture (b) Glove captured hand motion v.s. optically captured markers.

Figure 13: Comparison against optical motion capture.

We ran an experiment that compared the differences between glove captured hand motion and optically captured marker positions. 8

Table 3: Rotation differences against measured values among all DOFs.

Pose	Raw	Linear	LMC-GPR	SPGPs 100
Flat	$10.2^\circ \pm 11.8^\circ$	$1.5^\circ \pm 2.0^\circ$	$0.9^\circ \pm 1.1^\circ$	$0.9^\circ \pm 1.2^\circ$
Fist	$16.1^\circ \pm 24.8^\circ$	$3.2^\circ \pm 3.5^\circ$	$1.5^\circ \pm 2.5^\circ$	$1.6^\circ \pm 2.5^\circ$
Spread	$13.4^\circ \pm 17.6^\circ$	$2.5^\circ \pm 4.7^\circ$	$0.9^\circ \pm 1.6^\circ$	$0.9^\circ \pm 1.6^\circ$
Overbend	$13.6^\circ \pm 19.3^\circ$	$2.4^\circ \pm 3.7^\circ$	$1.7^\circ \pm 2.6^\circ$	$1.8^\circ \pm 2.4^\circ$
ThumbUp	$17.3^\circ \pm 21.8^\circ$	$4.5^\circ \pm 8.3^\circ$	$2.0^\circ \pm 3.3^\circ$	$2.2^\circ \pm 3.3^\circ$
Pointing	$15.1^\circ \pm 19.1^\circ$	$4.5^\circ \pm 6.0^\circ$	$2.2^\circ \pm 3.7^\circ$	$2.5^\circ \pm 3.8^\circ$
PinkyUp	$13.9^\circ \pm 15.6^\circ$	$5.4^\circ \pm 8.0^\circ$	$2.5^\circ \pm 3.8^\circ$	$2.7^\circ \pm 3.9^\circ$
Rest	$12.0^\circ \pm 15.8^\circ$	$4.5^\circ \pm 6.6^\circ$	$2.4^\circ \pm 3.5^\circ$	$2.6^\circ \pm 3.6^\circ$
Ok	$12.8^\circ \pm 14.5^\circ$	$3.3^\circ \pm 6.1^\circ$	$1.5^\circ \pm 2.1^\circ$	$1.5^\circ \pm 2.2^\circ$

markers are put over the glove on root, index MCP, pinky MCP, thumb tip, and finger DIP joints (Figure 13(a)). The result shows general consistency between the two capturing methods and the differences are expected due to the following reasons: 1. The kinematic hand model is different from the real hand. 2. Placement of markers is not exact on the joint location. Markers slide on the glove. 3. The palm-arc is not considered. 4. Minor cross-coupling effects are ignored. Even though we use a small marker set, it still requires significant effort to correct mis-labeling. It is insufficient to reconstruct the entire hand motion from this marker set, and the markers are frequently occluded during the capture. The whole clip is included in the video.

7.2 Accuracy vs. Performance

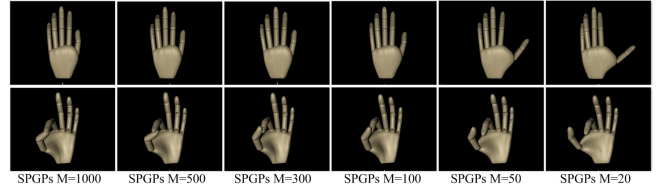


Figure 14: FK/IK accuracy losses using SPGPs with different pseudo-set size M .

To make LMC-GPR available for real-time hand motion capture, its sparse approximation is used. This subsection discusses the selection of an ideal M , and the trade-offs between accuracy and time efficiency. Table 4 lists the training and prediction time of SPGPs using different size M for glove calibration. It is obvious that the smaller M is, the faster the \bar{t}_{tr} and \bar{t}_{pred} will be. However, calibration accuracy could be compromised by discarding informative data in the training set. Figure 14 shows the cost of FK and IK accuracy when shrinking the size of pseudo-set for SPGPs.

Table 4: Average training Time \bar{t}_{tr} and prediction time \bar{t}_{pred} using SPGPs with different pseudo-set size run on a 2.80 GHz Intel Core 2 Duo CPU.

SPGPs size	\bar{t}_{tr} per glove	\bar{t}_{pred} per frame
$M = 1000$	657.572s	0.745s
$M = 500$	67.439s	0.187s
$M = 300$	32.027s	0.078s
$M = 100$	9.721s	0.018s
$M = 50$	7.557s	≤ 0.016 s
$M = 20$	6.224s	≤ 0.016 s

Based on the performance and accuracy tests, we chose pseudo-set size $M = 100$ for SPGPs to approximate LMC-GPR in real-time motion capture. Its calibration time per frame is about 0.0180s, which is sufficient for real-time use and it still maintains fairly good results. Users can run SPGPs $M = 100$ onsite and re-calibrate

data with the full LMC-GPR offline, providing timely feedback and greater accuracy if needed.

7.3 Applications

This LMC-GPR calibration method has been applied to body/hand motion capture for gesture research. Body motion was captured using the VICON system and hand motion was collected from the CyberGloves. The overall quality of captured gestures is greatly enhanced by our calibration method, e.g. with deictic gestures (“pointing”) and emblematic gestures (“okay”). In addition to gesture research, our glove calibration method can be applied to robotics, sign language, human computer interaction and so forth.

8 Generality and Flexibility

8.1 Generality

Unlike body motion capture, hand motion capture lacks unified hardware and standards. The goal of this paper is to present a flexible glove calibration method that can be applied to other hand capture hardware (with or without prior knowledge of sensor linearity) rather than being restricted to the CyberGlove.

Our calibration method mainly consists of two processes: sampling and calibration, which can both be customized for new hardware. To sample a sensor with uncertain properties, FK independent sampling may require more key poses to get denser samplings. FK cross-couple sampling and IK sampling in our method are already dense. The form of tuple set (s, r) allows for flexible mappings from s to r .

The LMC-GPR model can provide calibration for all hardware like the CyberGlove with known sensor linearity. For hardware with other properties (known or unknown), GPR is a flexible framework that allows different mean and covariance functions, including linear, polynomial, squared exponential, exponential and their combinations. For example, given a glove with polynomial sensors, we can replace the linear kernel term k_l with a polynomial kernel k_p in the model to better fit the data. For sensors with unknown properties, a more general and flexible GPR model $(m_0, k_{seARD} + k_{seARD})$ is needed, using the zero mean m_0 and the composite of two k_{seARD} . The experimental results listed in Table 2 show that treating linear glove data as a general signal, $(m_0, k_{seARD} + k_{seARD})$ yields relatively the same NLML and $\log \sigma$ as the LMC-GPR model. The only disadvantage is that this general-purpose model requires more time to train and predict.

8.2 Flexibility

Our approach has a sampling process that is independent from calibration, which separates the data from the algorithm. This structure has the following advantages:

Multiple Wearings: Sampling and the subsequent training, need to be performed only once for the same subject. The trained GPR can be used in different mocap sessions over multiple wearings of the gloves, for the same subject. Evaluation experiments in Section 7 are performed from different sessions using the same trained model.

Data Reuse: According to [Kessler et al. 1995], for different subjects, data from FK sampling can be reused. The calibration problem is thus treated as “calibrating sensors” rather than “calibrating subjects”. Such calibration may be “rough”, but can still maintain perceptually plausible hand shapes. IK sampling depends on hand dimensions, and thus is subject specific.

Accuracy Adaptation: Our calibration method can achieve both FK and IK accuracy, but some applications may not have dual accuracy concerns. A single accuracy requirement can be achieved by merely sampling the relevant data. For example, a robotic hand may only need fingertip accuracy, in which case, only collecting tuples from IK sampling to train the calibration model will suffice.

9 Conclusion and Future Work

We have presented a novel calibration approach for hand motion capture. It uses the general GPR framework to accommodate sensor linearity, coupling effects and noise. The flexible calibration function is inferred by a special LMC-GPR model trained on data collected from a separate sampling process. Results confirm that our approach can ensure better calibrated FK and IK accuracy.

Despite the effectiveness of our method for calibrating data gloves, it is worth noting that LMC-GPR is a statistical model. Its predictive mean function provides calibrated values not by interpolating through every sampled point in the input space. Thus we should also care about its predictive variance as this indicates the uncertainty of the prediction. Any chosen model whose variance is too large should not be applied in calibration.

In addition, there are several limitations to our calibration method that could benefit from future work:

- First, CyberGlove includes a palm arch sensor, but due to the severe cross-coupling effects, we ignore this sensor and treat the palm as a rigid body. Although GPR can handle high dimensional inputs, the real difficulty comes from the sampling process. It would be very time-consuming to collect sufficient samples to describe the calibration function in the high dimensional space. On the other hand, the fact that the palm arc is highly related to rotations of some hand joints can be utilized for simulating palm shapes using a biomechanic rule based approach.
- Second, although our approach can ensure fingertip accuracy for an “ok” sign or a grabbing motion, as the IK sampling process only samples finger-touching motions, we cannot guarantee IK accuracy when the fingertip distances are larger. One possible solution to ensure accuracy for a given fingertip distance is sampling motions when fingers are manipulating a rod with the given length, but more general positioning remains a challenge.
- Third, as a data driven calibration approach, the achieved accuracy is related to the sampling quality. More training data with accurate information is preferred, but the aforementioned sampling difficulties restrict us from collecting too much data. Therefore, in the future, a more efficient sampling process, or a sampling method that collects most representative data, or a non data-driven method are possible solutions to improve the calibration approach.

We are hoping that our work can provide a new thread on glove calibration and further promote the use of hand motion capture techniques.

10 Acknowledgment

We would like to thank Jonathan Graham for making the MAYA hand model, and Scott Bishop for his precious review feedback. Financial support for this work was provided by the National Science Foundation through grants IIS 0856084, IIS 1115872 and IIS 0845529.

References

- ARISTIDOU, A., AND LASENBY, J., 2009. Inverse kinematics: a review of existing techniques and introduction of a new fast iterative solver. F-INFENGIfr. 632, CUED.
- BUSS, S., AND KIM, J. 2005. Selectively damped least squares for inverse kinematics. *journal of graphics, gpu, and game tools* 10, 3, 37–49.
- BUSS, S. 2004. Introduction to inverse kinematics with jacobian transpose, pseudoinverse and damped least squares methods. *University of California, San Diego, Typeset manuscript, available from <http://math.ucsd.edu/~sbuss/ResearchWeb>*.
- CHAN, S., AND LAWRENCE, P. 1988. General inverse kinematics with the error damped pseudoinverse. In *Robotics and Automation, 1988. Proceedings., 1988 IEEE International Conference on*, IEEE, 834–839.
- CHAO, A. 2001. *Uni-dimensional evaluation of Wristsystem and CyberGlove*. Master's thesis, Department of Industrial and Manufacturing Engineering, Wichita State University, Los Alamitos, CA.
- CHOU, T., GADD, A., AND KNOTT, D. 2000. Hand-eye: A vision based approach to data glove calibration. *Proceedings of Human Interface Technologies 2000*.
- GRIFFIN, W., FINDLEY, R., TURNER, M., AND CUTKOSKY, M. 2000. Calibration and mapping of a human hand for dexterous telemanipulation. In *ASME IMECE 2000 Symposium on Haptic Interfaces for Virtual Environments and Teleoperator Systems*, 1–8.
- HU, H., GAO, X., LI, J., WANG, J., AND LIU, H. 2004. Calibrating human hand for teleoperating the hit/dlr hand. In *Robotics and Automation, 2004. Proceedings. ICRA'04. 2004 IEEE International Conference on*, IEEE, vol. 5, 4571–4576.
- HUENERFAUTH, M., AND LU, P. 2010. Accurate and accessible motion-capture glove calibration for sign language data collection. *ACM Transactions on Accessible Computing (TACCESS)* 3, 1, 2.
- JIN, S., LI, Y., AND CHEN, W. 2010. A novel dataglove calibration method. In *Computer Science and Education (ICCSE), 2010 5th International Conference on*, IEEE, 1825–1829.
- JÖRG, S., HODGINS, J., AND SAFONOVA, A. 2012. Data-driven finger motion synthesis for gesturing characters. *ACM Transactions on Graphics (TOG)* 31, 6, 189.
- KAHLESZ, F., ZACHMANN, G., AND KLEIN, R. 2004. 'visual-fidelity' dataglove calibration. In *Computer Graphics International, 2004. Proceedings*, IEEE, 403–410.
- KANG, C., WHEATLAND, N., NEFF, M., AND ZORDAN, V. 2012. Automatic hand-over animation for free-hand motions from low resolution input. In *Motion in Games*. Springer, 244–253.
- KESSLER, G., WALKER, N., AND HODGES, L. 1995. Evaluation of the cyberglove as a whole hand input device.
- KUCH, J., AND HUANG, T. 1994. Human computer interaction via the human hand: a hand model. In *Signals, Systems and Computers, 1994. 1994 Conference Record of the Twenty-Eighth Asilomar Conference on*, IEEE, vol. 2, 1252–1256.
- LAWRENCE, N. 2007. Learning for larger datasets with the gaussian process latent variable model.
- LEE, J., AND KUNII, T. 1995. Model-based analysis of hand posture. *Computer Graphics and Applications, IEEE* 15, 5, 77–86.
- MENON, A., BARNES, B., MILLS, R., BRUYNS, C., TWOMBLY, A., SMITH, J., MONTGOMERY, K., AND BOYLE, R. 2003. Using registration, calibration, and robotics to build a more accurate virtual reality simulation for astronaut training and telemedicine. In *Proceedings of the 11th International Conference in Central Europe on Computer Graphics, Visualization, and Computer (WSCG03)*.
- MEREDITH, M., AND MADDOCK, S. 2005. Adapting motion capture data using weighted real-time inverse kinematics. *Computers in Entertainment (CIE)* 3, 1, 5–5.
- NAKAMURA, Y., AND HANAFUSA, H. 1986. Inverse kinematic solutions with singularity robustness for robot manipulator control. *Journal of Dynamic Systems, Measurement, and Control* 108, 163.
- PAVLOVIC, V., SHARMA, R., AND HUANG, T. 1997. Visual interpretation of hand gestures for human-computer interaction: A review. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 19, 7, 677–695.
- RASMUSSEN, C. E., AND WILLIAMS, C. K. I. 2006. *Gaussian Processes for Machine Learning*. MIT Press.
- SNELSON, E., AND GHAHRAMANI, Z. 2006. Sparse gaussian processes using pseudo-inputs.
- STEFFEN, J., MAYCOCK, J., AND RITTER, H. 2011. Robust dataglove mapping for recording human hand postures. *Intelligent Robotics and Applications*, 34–45.
- TURNER, M. 2001. *Programming dexterous manipulation by demonstration*. PhD thesis, Citeseer.
- WAMPLER, C. 1986. Manipulator inverse kinematic solutions based on vector formulations and damped least-squares methods. *Systems, Man and Cybernetics, IEEE Transactions on* 16, 1, 93–101.
- YAZADI, F., 2009. *Cyberglove systems cyberglove ii wireless data glove user guide*. CyberGlove Systems LLC.
- YE, Y., AND LIU, C. K. 2012. Synthesis of detailed hand manipulations using contact sampling. *ACM Transactions on Graphics (TOG)* 31, 4, 41.
- ZHAO, W., CHAI, J., AND XU, Y.-Q. 2012. Combining marker-based mocap and rgb-d camera for acquiring high-fidelity hand motion data. In *Eurographics/ACM SIGGRAPH Symposium on Computer Animation*, 33–42.