Estimation of Continuous Multi-DOF Finger Joint Kinematics from Surface EMG using a Multi-output Gaussian Process

Jimson Ngeo[1], Tomoya Tamei[1], Tomohiro Shibata[2]

***Abstract*— Surface electromyographic (EMG) signals have often been used in estimating upper and lower limb dynamics and kinematics for the purpose of controlling robotic devices such as robot prosthesis and finger exoskeletons. However, in estimating multiple and a high number of degrees-of-freedom (DOF) kinematics from EMG, output DOFs are usually estimated independently. In this study, we estimate finger joint kinematics from EMG signals using a multi-output convolved Gaussian Process (Multi-output Full GP) that considers dependencies between outputs. We show that estimation of finger joints from muscle activation inputs can be improved by using a regression model that considers inherent coupling or correlation within the hand and finger joints. We also provide a comparison of estimation performance between different regression methods, such as Artificial Neural Networks (ANN) which is used by many of the related studies. We show that using a multi-output GP gives improved estimation compared to multi-output ANN and even dedicated or independent regression models.**

I. Introduction

Recently, more and more advanced multifingered, high degree-of-freedom (DOF) robotic hands are being developed. More particularly, research in myoelectric control has led to the development of robot prosthetic devices and exoskeletons that are able to mimic the function of the human hand and can do highly dexterous tasks. However, current myoelectric-based control schemes have been limited to on-off binary control, pattern recognition-based sequential control and proportional control of a few DOFs [1].

While many have mapped surface EMG signals to upper limb kinematics and dynamics, few have focused on mapping to fine finger joint actuations. This is because of the many muscles, both extrinsic and intrinsic, involved in the control of complex coordination and high DOFs present in the hand. Many studies have modeled the human hand as a complex and highly articulated mechanism having more than 20 DOFs, considering all fingers and the wrist [2]. However, many studies have shown that extrinsic muscles in the forearm have enough information to estimate finger kinematics (e.g. joint angles) [3][4] or dynamics (e.g. pinch force). It has also been suggested that the effective dimensionality of hand postures can be represented in a lower dimensional space [5] possibly due to the existence of synergies, biomechanical coupling and limitations in neural control.

This work was supported by Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (No. 23240028).

[1] J. Ngeo and T. Tamei are with the Grad. School of Information Science, Nara Institute of Science and Technology, Ikoma-shi, Nara 630-0192, Japan. {jimson-n, tomo-tam}@is.naist.jp

[2] T. Shibata is with the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology, 2-4 Hibikino, Wakamatsu-ku, Kitakyushu 808-0196, Japan. tom@brain.kyutech.ac.jp

Jiang et al. have shown the feasibility of doing online simultaneous and proportional myoelectric control of the wrist using EMG signals taken from both healthy subjects and amputees [1]. On relating to finger movements, Smith et al. were able to asynchronously decode individual metacarpophalangeal (MCP) joint angles of all five fingers using an artificial neural network (ANN) [4]. Hioki et al. estimated the five proximal interphalangeal (PIP) finger joints and considered some dynamical relationship between the EMG and finger actuation by adopting time delay factors and feedback stream into a recurrent ANN [6]. As in our previous study, we also estimated 15 finger joint DOFs simultaneously using ANN [7]. However, in all the previous studies mentioned, all the output joint kinematics were estimated either independently or whose relationship were heuristically considered through the interaction of the hidden neurons. Not to mention that in almost all the studies related to continuous kinematics estimation from EMG, regression using ANN with different configurations seems to be the default nonlinear estimator used [1,3-4,6-7].

However, highly dexterous hand and finger movements are done in a coordinated fashion. It has been shown in previous studies that finger joints in the hand are highly correlated. By using principal component analysis, Ingram et al. found that the actual dimension in hand movements was much less than the total DOFs available [8]. Driven by the existence of highly correlated finger joints, our work aims to improve current finger joint kinematics estimation from EMG by implicitly considering correlations between multiple outputs. In this study, we estimate finger joint kinematics from muscle activation inputs using a full-covariance dependent output Gaussian Process (GP) for modelling correlated outputs. A convolution process and a shared latent function framework approach is used to establish dependencies between output variables. We also compare accuracies of mapping EMG to the corresponding multi-DOF hand/finger kinematics using different regression models, namely, a dedicated or independent NN for each DOF, a multi-output NN, a dedicated or independent GP and a multi-output full GP.

II. Method

A. Subjects

Five participants (3 Male, 2 Female, aged 23-31 years old), who gave informed consent, volunteered in this study. All the subjects were healthy and had no known physical impairments. Except for two of the participants, all the rest had no previous experience with myoelectric control nor with any 3D motion capture experiments.

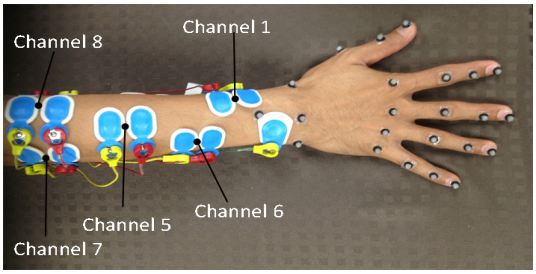

Fig. 1. The surface EMG electrodes and the reflective markers are placed on the target muscles in the forearm and on the hand, respectively.

B. EMG Recordings

Eight bipolar active-type Ag-AgCl electrodes from Ambu, with inter-electrode distance of 20 mm were placed on eight extrinsic muscles in the forearm that are known to contribute to wrist and finger movements (Figure 1). These muscles are the Abductor Pollicis Longus (APL), Flexor Carpi Radialis (FCR), Flexor Digitorum Superficialis (FDS), Flexor Digitorum Profundus (FDP), Extensor Digitorium (ED), Extensor Indices (EI), Extensor Carpi Ulnaris (ECU), and Extensor Carpi Radialis (ECR). These target muscles were mostly found by palpation and chosen from known anatomical landmarks. A single electrode was also placed on the subject's olecranon to serve as ground. The EMG signals were measured using a compact BA1104 pre-amplifier with active-type (Ag/AgCl) electrodes having interelectrode distance of 20 mm, and a telemetry unit TU-4 (Digitex Lab. Co. Ltd). The hardware provided a high frequency filter of 1 kHz during the EMG data acquisition process. The EMG signals were sampled at 2 kHz, and input to an A/D converter.

C. Finger Kinematics

Twenty-two small reflective markers (6 mm diameter) were attached on the subject's hand, with a marker located on each joint of the finger and 3 in the wrist area. While finger movements were made, the hand and finger motions were recorded simultaneously using a MAC3D motion capture system (Motion Analysis Corp.). The marker trajectories were sampled at 200 Hz with units in millimeters. The joint kinematics, namely the MCP, PIP and DIP finger joints, were then calculated from the 3D positions of each marker. Because the thumb does not have a DIP joint, the carpometacarpal (CMC) joint was considered before the MCP joint. All in all 15 DOFs were computed.

D. Experimental Protocol

Each participant was comfortably seated with their hands positioned on a table centered on the motion capture area. They were tasked to do the following tasks: (1) move one finger at a time periodically, (2) move all five fingers simultaneously and (3) move any finger freely in any direction. All finger movements were limited to flexion and extension movements. Irregular periodic movements and different finger combinations were encouraged on the third task. The first task consisted of 5 sets of movements, one for each finger. While the second and third tasks consisted of 1 set each. Each set consisted of 5 trials lasting 20 seconds each. All the trials were sequentially done and the participants were allowed to rest anytime throughout the experiment. The subjects were also instructed to, as much as possible, maintain the position of the wrist and the whole arm in a fixed position.

E. EMG Processing

The raw EMG signals were first rectified, normalized, lowpass filtered (with 4 Hz cut-off frequency, zero-phase 2nd-order Butterworth filter), and downsampled to match the frequency of the motion data. We used an EMG-to-Muscle Activation model to transform the preprocessed EMG to a form that considers some muscle activation dynamics. It is known that there occurs some time delay, known as electromechanical delay (EMD), between the onset of the EMG and the exerting tension in the muscles. To learn a suitable filtered signal that considers EMD and nonlinearity between EMG and muscle activations, we use an EMG-to-Muscle Activation model that transforms the EMG to muscle activation (force). Buchanan et al. created a second-order filter that works efficiently to model EMG and muscle activation [9]. In this study, we employed their filter:

$$u_j(t) = \alpha e_j(t-d) - \beta_1 u_j(t-1) - \beta_2 u_j(t-2) \quad (1)$$

$$v_j(t) = \frac{e^{A_j u_j(t)} - 1}{e^{A_j} - 1}, \quad (2)$$

where e_j is the pre-processed EMG and v_j is the muscle activation of muscle j at time t. Parameters α, β_1, β_2 are recursive coefficients, d is the EMD parameter and A_j introduces nonlinearity between EMG and muscle activation. Filter stability is guaranteed by putting on constraints:

$$\beta_1 = \gamma_1 + \gamma_2 \quad (3)$$

$$\beta_2 = \gamma_1 \cdot \gamma_2 \quad (4)$$

$$|\gamma_1| < 1, |\gamma_2| < 1 \quad (5)$$

$$\alpha - \beta_1 - \beta_2 = 1. \quad (6)$$

A_j is constrained between -3 (highly exponential) and 0.

We used the muscle activation as the chosen input feature because it considers muscle dynamics and have been shown to perform better than time-domain (TD) features or filtered EMG signals (without delay considerations) [7]. Parameters of the muscle activation model were obtained using Matlab's Optimization Toolbox and by minimizing the mean-square-error between the measured and estimated finger kinematics.

F. Multi-output Regression

Using artificial neural networks has been the common choice by many of the previous studies because computation is fast and inexpensive. However, this choice is heuristic, gives a black-box approach, and requires a lot of training data to effectively model complex nonlinear functions. In this study, we chose to use a Bayesian inference framework in the form of a convolved multi-output Gaussian Process that considers correlation between outputs [10]. Let $X \in \mathbf{R}^{m \times N}$ be the N-channel muscle activation input vectors, with the multi-output finger DOF output $Y \in \mathbb{R}^{m \times Q}$. Given the convolution formalism, where a convolution process between a smoothing kernel and latent function is done in order to establish dependencies between output. Alvarez and

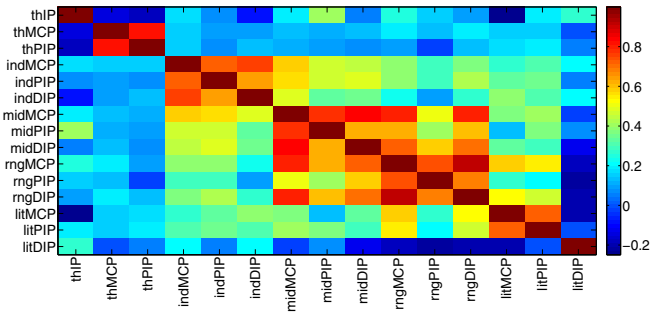

Fig. 2. A representative correlation matrix from Subject 5 shows the pairwise correlation between all finger joints.

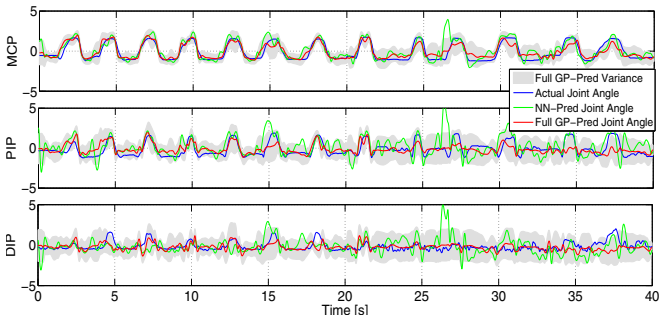

Fig. 3. One representative data from an index finger flexion and extension task. The estimation performance along each finger joint (MCP, PIP and DIP) joint is shown.

Lawrence showed a full Gaussian Process model over the set of outputs and defined the likelihood of the model as

$$p(y|X,\phi) = \mathcal{N}(0, K_{f,f} + \Sigma), \tag{7}$$

where ϕ is the set of parameters of the covariance matrix, Σ is a diagonal matrix with elements $\{\sigma_q^2\}_{q=1}^{Q}$ from independent Gaussian noise processes and $K_{f,f} \in \mathbb{R}^{QN \times QN}$ is the covariance matrix relating all data points at all outputs with elements taken from the correlation between outputs

$$cov[f_q(x), f_s(x')] = \sum_{r=1}^{R} \int_{-\infty}^{\infty} k_{qr}(x-z) \times \int_{-\infty}^{\infty} k_{sr}(x'-z') k_{u_r u_r}(z-z') dz' dz \tag{8}$$

where k_{qr} and k_{sr} are the kernel smoothing functions at different output functions f, and $k_{u_r u_r}$ is the covariance function for the latent function $u_r(z)$.

The predictive distribution for new input X_* is given by

$$p(y_*|y, X, X_*, \phi) = \mathcal{N}(K_{f_*,f}(K_{f,f} + \Sigma)^{-1} y, K_{f_*,f}(K_{f,f} + \Sigma)^{-1} K_{f,f_*} + \Sigma) \tag{9}$$

Because learning from the log-likelihood involves the computation of the inverse of $K_{f,f} + \Sigma$ whose complexity grows as the size input or output matrix increases. Rather than using the sparse approximation scheme proposed in the original paper [10], we reduced the full covariance matrix by using only a subset of the full training data [11]. A fixed interval sampling was used to reduce the number of samples and this drastically improved hyperparameter learning and the training time needed. In this study, a Matlab implementation of the multi-output Full GP [10] was used.

We compared the estimation performance of using the multi-output Full GP compared to the standard baseline of using a 3-layer feedforward multi-output neural network. For the full GP, we used a squared exponential with automatic relevance determination (SEard) covariance or kernel function given by $k_{u_r u_r}(x, x') = exp[-\frac{1}{2}(x-x')^\top \mathbf{L_r}(x-x')]$, where $\mathbf{L_r}$ is a diagonal matrix with different length parameters along each dimension. The smoothing kernel also had the same form, $k_{qr}(\tau) = \frac{S_{qr}|L_{qr}|^{1/2}}{(2\pi)^{\frac{1}{2}}} exp[-\frac{1}{2}(\tau)^\top \mathbf{L_{qr}}(\tau)]$, where $S_{qr} \in \mathbb{R}$ and $\mathbf{L_{qr}}$ is a symmetric positive definite matrix. The GP hyperparameters and NN neuron weight parameters were learned by minimizing the negative marginal log likelihood and the mean-square-error (MSE), respectively. A scaled conjugate gradient (SCG) algorithm was used in learning the parameters. The maximum iteration was set to 500 and for the neural network the hidden layer was composed of 250 hidden neurons with tan-sigmoidal activation functions.

Eighty percent of the data from each task set were concatenated together to form a larger training set while the remaining twenty percent, also concatenated together were used for testing. The training data were standardized to have zero mean and unit variance on each dimension while the test data where standardized to have a mean around the training mean. A five-fold cross validation was also done.

To measure the overall estimation performance of the regression method, we measured the accuracy of the method's predictions on an unseen test data using the Standardized Mean Squared Error (SMSE) and the Pearson's Correlation Coefficient (R) performance indices. The SMSE is the mean squared error normalized by the MSE of a dumb predictor that always predicts the mean of the training set [11].

Rather than getting the full multi-output model of all 15 finger DOFs, which also has some scalability issues with using the GP, we make use of some prior knowledge and assume that correlations in finger joints exists only in some joints. We validate this in the next section. In this study, we simplify our GP model and only consider multi-output within each finger (3 DOF) and within adjacent MCP finger joints (4 DOF) without the thumb.

III. Results and Discussion

Figure 2 shows the correlation matrix between the actual measured finger joint angles of Subject 5 across all set of finger movements concatenated together. It can be seen in the figure that there is high correlation between joints within the same finger (0.5804 ± 0.3512) and between joints in adjacent fingers (0.5206 ± 0.3060). This result shows the existence of coupled finger joints, such as the MCP joints within adjacent finger and the PIP and DIP joints within the same finger. It is interesting to see that among all the fingers, the thumb is the most independent and has the least correlation with other fingers. When getting the mean correlation matrix across all 5 subjects, almost all the paired joints showed significant correlations ($P < 0.05$) except for the paired joints, litMCP-indPIP and litDIP-thMCP ($P = 0.07$ and $P = 0.19$, respectively).

The different regression models were trained using two configurations, one training for 3 output joints in each

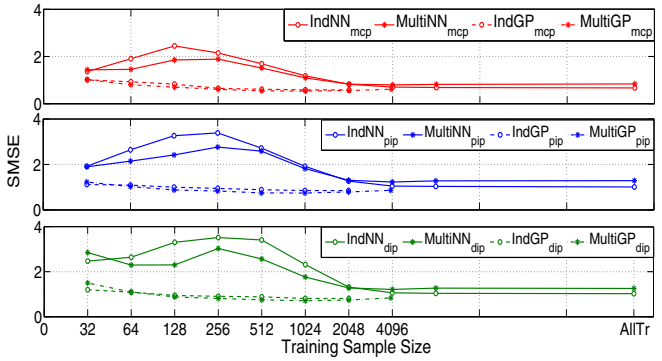

Fig. 4. Error Learning Curve comparison between different configurations of the NN and GP regression models.

individual finger, and another across all 4 MCP joints of the adjacent fingers. Figure 3 shows a representative estimation result taken from an index finger flexion and extension test trial. In this representative data, the NN was fully trained with all the training samples available, while the full-GP was trained using only 2048 induced samples obtained from fixed-interval sampling. As shown in this test trial, full-GP ($R_{mcp} = 0.93, R_{pip} = 0.59, R_{dip} = 0.54$) that considers correlation between outputs performed better than multi-output NN ($R_{mcp} = 0.83, R_{pip} = 0.52, R_{dip} = 0.12$).

We also plotted the error learning curve involving the four regression models that we compared. Figure 4 shows the mean standardized mean-square-error (SMSE) of the predicted and measured finger joint when the models were trained using various number of samples across all subjects. In can be shown in this figure that the use of GP regression gives significant reduction in error given that only a few training samples are used. This shows that we can have shorter training data collection periods yet still obtain relatively good estimation performance as GP is shown to be able to handle missing data more readily.

In figure 5, the overall estimation performance using different regression models across different finger output configurations are shown. In the case of estimating the joint kinematics within each individual finger, multi-output GP significantly outperformed ($P < 0.05$) the independent NN, multi-output NN and the independent GP model. In this individual finger configuration, the mean across the 3 DOF in each finger and across all fingers were obtained. Compared to using the neural network, estimation even for the PIP and DIP angles improved slightly when compared with the independent GP. Creating an independent GP model for each DOF gives only a slightly lower performance compared to the multi-output GP. The latter model however considers correlations between outputs which gives a more natural model that reflects more closely the correlations between finger joints, rather than having separate models for each DOF. In the case of the adjacent fingers, where we considered only the 4 finger MCP joints, no significant difference ($P > 0.05$) was found between the independent GP model compared to the independent NN model. The multi-output GP, however, consistently showed improved estimation performance compared to using any NN configuration and a slight improvement compared to using an independent GP.

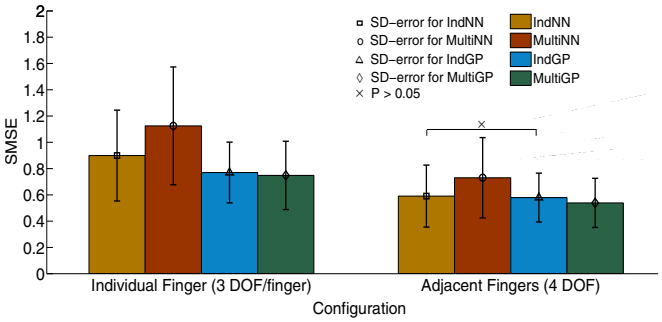

Fig. 5. Comparison of the SMSE between different regression models using the number of training samples that gave the best estimation performance.

IV. Conclusion

This study presents a solution on how complex, multi-DOF and correlated finger joint kinematics can be estimated from EMG or muscle activation inputs. We have systematically compared different regression methods, such as those conventionally used by many related studies, and those that do not consider any existing relationship between joint kinematics. We have shown that using multi-output GP or in general GP regression method gives better estimation performance compared to using neural networks, even when training data is few. To truly validate the usability of our system, future work would include investigating motor coordination and muscle activation patterns involved in doing more coordinated, skillful and activities of daily living (ADL) hand tasks.

V. Acknowledgments

This work was supported by the Grant-in-Aid for Scientific Research from Japan Society for the Promotion of Science (No. 23240028).

References

[1] N. Jiang, J. Vest-Nielsen, S. Muceli, and D. Farina, "EMG-based simultaneous and proportional estimation of wrist/hand kinematics in uni-lateral trans-radial amputees," *J. Neuroeng. & Rehab.*, vol. 9, 2012.

[2] H. Liu, "Exploring human hand capabilities into embedded multifingered object manipulation," *Industrial Informatics, IEEE Transactions on*, vol. 7, pp. 389–398, Aug 2011.

[3] P. Afshar and Y. Matsuoka, "Neural-based control of a robotic hand: evidence for distinct muscle strategies," in *Robotics and Automation, IEEE International Conference on*, vol. 5, pp. 4633–4638, 2004.

[4] R. Smith, F. Tenore, D. Huberdeau, R. Etienne-Cummings, and N. Thakor, "Continuous decoding of finger position from surface EMG signals for the control of powered prostheses," in *30th Annual International IEEE EMBS Conference*, pp. 197–200, August 2008.

[5] E. Todorov and Z. Ghahramani, "Analysis of the synergies underlying complex hand manipulation," in *26th Int. IEEE EMBS Conf.*, 2004.

[6] M. Hioki and H. Kawasaki, "Estimation of finger joint angles from sEMG using a neural network including time delay factor and recurrent structure," *Int. Scholarly Research Network (ISRN) Rehab.*, 2012.

[7] J. Ngeo, T. Tamei, and T. Shibata, "Continuous estimation of finger joint angles using muscle activation inputs from surface EMG signals," in *34th International IEEE EMBS Conference*, pp. 2756–2759, 2012.

[8] J. Ingram, K. Kording, I. Howard, and D. Wolpert, "The statistics of natural hand movements," *Exp. Brain Research*, vol. 188(2), 2008.

[9] T. Buchanan, D. Lloyd, K. Manal, and T. Besier, "Neuromusculoskeletal modeling: Estimation of muscle forces & joint moments and movements from measurements of neural command," *Journal of Applied Biomechanics*, vol. 20(4), pp. 367–395, November 2004.

[10] M. Alvarez and N. Lawrence, "Sparse convolved Gaussian processes for multi-output regression," in *NIPS*, vol. 21, pp. 57–64, 2008.

[11] K. Chalupka, C. Williams, and I. Murray, "A framework for evaluating approximation methods for Gaussian process regression," *The Journal of Machine Learning Research*, vol. 14(1), pp. 333–350, 2013.