Neural Network Methods for Error Canceling in Human-Machine Manipulation

Wei Tech Ang, Cameron N. Riviere
The Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15213, USA

Abstract – A neural network technique is employed to cancel hand motion error during microsurgery. We trained and tested a cascade neural network with Kalman filtering on 15 hand movement data files collected from 4 surgeons. The neural network reduces the root mean square error (rmse) of the surgeons’ erroneous motion by an average of 39.5%. Our experimental approach shows a higher rmse reduction by 9.6% and better motion profile preservation than a previous work by Riviere and Khosla [15] that worked on the same set of data. Our investigation on a neural network trained on one surgeon’s file in predicting the movement of the others has drawn inconclusive results, but they suggest that there might exist a generalized neural network for a group of surgeons.

I. INTRODUCTION

The human ability to perform micromanipulation is hampered by inherent erroneous hand motion. This manual imprecision affects the performance of microsurgery [1], it complicates some surgical procedures and makes certain delicate procedures impractical and often impossible [2].

The most familiar type of involuntary or erroneous movement affecting microsurgery is physiological tremor [3]. Tremor is defined as any involuntary, approximately rhythmic, and roughly sinusoidal movement [4]. Physiological tremor is a type of tremor that is inherent in the movement of healthy subjects. As exhibited during vitreoretinal microsurgery, it is essentially an oscillation at 8-12 Hz whose frequency is independent of the mechanical properties of the hand and arm [4]. The resulting tool tip oscillation is typically 50 µm peak-to-peak (p-p) or greater [5]. Besides physiological tremor, measurements of the hand motion of surgeons have shown significant non-tremorous components of motion such as jerk (i.e., normal myoclonus), drift, and certain vaguely defined and poorly understood low-frequency undesired components [6], which are often larger than physiological tremor.

There has been an ongoing research interest in enhancing human positioning accuracy during microsurgery. The first of these is the use of telerobotic technology [7], where the unstable human hand is replaced by a robotic arm. Taylor et al. have used a "steady hand" approach, in which a robot and a surgeon directly manipulate the same tool [8], with the robot having high stiffness, and moving along with only those components of the manual input force that are deemed desirable.

In order to further reduce cost, and to maximize ease of use, user acceptance, and compatibility with current surgical practice, the present authors are implementing accuracy enhancement within a completely hand-held tool, seeking to keep the instrument size and weight as close as possible to those of existing passive instruments. This device should sense its own motion, estimate the undesired component of the sensed motion, and manipulate its own tip to nullify the erroneous motion in real-time as shown in Fig. 1.

Fig. 1. Active hand-held instrument for error compensation in microsurgery

For this approach to work, it is of paramount importance to accurately model both tremor and various types of non-tremorous involuntary movement, so as to enable online canceling. Several techniques have been developed for tremor modeling and suppression. Riley and Rosen [9], among others, have investigated lowpass filtering; Gonzalez et al. [10] proposed an equalizer to suppress pathological tremor. Riviere et al. [11] developed an adaptive filter to cancel physiological tremor during surgery, using an artificial frequency-modulated sinusoid as a reference. However, other significant sources of error, e.g. jerk and drift, or low-frequency error, have yet to be substantially suppressed. Since little is known about these components, and since reference signals for adaptive noise canceling, are unavailable, suppression is difficult.

The mapping from human intention to human movement output is nonlinear. Neural networks model nonlinear processes well, and have been used in modeling of human control strategies [12]. The complexity and multiplicity of involuntary hand motion components, and the paucity of knowledge about components such as drift, makes a neural network approach well suited to modeling of human movement error processes. Riviere and Khosla [13] used a cascade neural network for noise canceling in human hand motion. Their experiments showed that the neural network successfully modeled and reduced the errors on recorded hand movement files of four surgeons. In this paper, we are
presenting a different approach in employing the same cascade neural network technique on the same set of data and compare our results with those obtained by Riviere and Khosla. We also go a step further to investigate the effectiveness of the trained networks on surgeons other than those on which they were trained.

Though the experiments presented here focus on surgery, the concepts demonstrated are directly relevant to a wide range of manipulation applications with small signal-to-noise ratio, e.g. helping rehabilitation patients with pathological tremors to use a computer mouse, enhancing manual accuracy in cell manipulation in the biotech industry, etc.

II. NEURAL NETWORK WITH EXTENDED KALMAN FILTERING

Extended Kalman filtering (EKF) is an extension of Kalman filter to deal with non-linear systems via linearization about the current parameter estimates. In neural network training, learning is cast as an identification problem for a non-linear dynamic system. The neural network weights represent the state of the non-linear system. The EKF theory is then used to derive a recursion for the weight updates. This work uses NDEKF, in which the network weights are grouped such that each group contains the input nodes, the output nodes and one hidden node. For each group, elements of the error covariance matrix estimate corresponding to other groups can be ignored, greatly reducing the computational complexity.

III. EXPERIMENTAL METHODS

Hand movement data of surgeons were recorded in Wilmer Eye Institute of Johns Hopkins University. Each surgeon held a microsurgical instrument with the tip inserted in a sclerotomy in the eye of a mannequin face. A Hall effect sensor mounted inside the mannequin eye detected the position, in one dimension, of a 0.26g permanent magnet mounted on the tip of the instrument. Data were recorded for 16s at a sampling rate of 250Hz. The surgeons attempted to hold the instrument motionless for the duration of each test, therefore any motion in these recordings is considered to be error. A total of 15 files were obtained from four surgeons (5, 5, 3, and 2 files, respectively).

To ease disambiguation of erroneous movement from desired movement for purposes of evaluation, surgeons were given fixed targets at which to point, and tried to keep the instrument motionless, thus ensuring that all recorded motion is error. To make the experiments more realistic we generate low frequency pseudo-voluntary motions and add them to the recorded still hand error movement. The pseudo-voluntary motions thus serve as the target motions in our experiments. The magnitude of the randomly generated pseudo-voluntary motions has a ratio of roughly 1:1 to the mean RMS error of the 15 data files. Gaussian white noise sequences are also generated and then low-pass filtered before adding to the randomly generated low frequency motions. Two different pseudo-voluntary motions are generated in this manner, one for the training data files and the other for testing data files.

Separate neural networks were used for each of the four surgeons. For each surgeon, a cascade-architecture neural network, using extended Kalman filtering for learning, was trained using one of the training data files described above. The remaining data files from each surgeon were used for testing of the trained network. The rmse with respect to the pseudo-voluntary motion was calculated for each file, both before and after processing by the neural network.

The input to the neural network was a window of data in the time series, i.e. the number of input nodes depended on the length of the window. The output of the neural network was the error-compensated motion, and since we were looking only at one-dimensional data, there will only be 1 output node.

![Diagram of the cascade neural network architecture. The diagram shows a network with three hidden nodes. As each hidden node is added, it is connected to the input and output nodes, as well as each of the preceding hidden nodes.](image-url)
Different combinations of number of input nodes and hidden nodes were tested to obtain the best net architecture for each surgeon. Riviere and Khosla [13] used the same set of data but chose the error estimate as the network output, so that the output of the neural network could be used directly by other downstream actions to cancel this error. In addition, Riviere and Khosla fixed the number of input nodes to be 100 and the maximum number of hidden nodes to be 10. The results of our experiment will be compared with those reported in [13].

We also investigate the existence of a generalized neural network for all surgeons by testing how well a neural network trained on one surgeon’s file perform in predicting the time series of the others.

IV. RESULTS

The neural network reduced the rmse with respect to the randomly generated pseudo-voluntary motion for all the testing data files. Table I shows for each of the surgeon the mean raw rmse of the data files (with the number of the testing files in parenthesis), the mean rmse of the output of the neural network and their standard deviations, and the neural network architecture that gives the best result. Table II shows that our approach (NN_{vm}) of using the pseudo-voluntary motion as the training targets outperforms Riviere and Khosla’s method (NN_{err}) that uses the erroneous motion as the training target.

**TABLE I**

Performance of the neural network and the network architecture that produces the best result.

<table>
<thead>
<tr>
<th>Surgeon # (no. of testing files)</th>
<th>Mean raw rmse (mm)</th>
<th>Mean rmse of neural network output (mm)</th>
<th>Standard deviation (mm)</th>
<th>Best Network Architecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (4)</td>
<td>0.112</td>
<td>0.055</td>
<td>0.005</td>
<td>75 input, 3 hidden</td>
</tr>
<tr>
<td>2 (4)</td>
<td>0.046</td>
<td>0.033</td>
<td>0.002</td>
<td>60 input, 6 hidden</td>
</tr>
<tr>
<td>3 (2)</td>
<td>0.048</td>
<td>0.037</td>
<td>0.001</td>
<td>100 input, 3 hidden</td>
</tr>
<tr>
<td>4 (1)</td>
<td>0.127</td>
<td>0.056</td>
<td>-</td>
<td>50 input, 6 hidden</td>
</tr>
</tbody>
</table>

**TABLE II**

Percentage rmse reduction comparison between ours (NN_{vm}) and Riviere/Khosla’s (NN_{err}) method

<table>
<thead>
<tr>
<th>Surgeon #</th>
<th>% rmse reduction (NN_{vm})</th>
<th>% rmse reduction (NN_{err})</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50.9</td>
<td>44.6</td>
</tr>
<tr>
<td>2</td>
<td>28.3</td>
<td>26.1</td>
</tr>
<tr>
<td>3</td>
<td>22.9</td>
<td>18.8</td>
</tr>
<tr>
<td>4</td>
<td>55.9</td>
<td>29.9</td>
</tr>
<tr>
<td>Average</td>
<td>39.5</td>
<td>29.9</td>
</tr>
</tbody>
</table>

Figures 3a and 3b depict the sample results of our approach and Riviere and Khosla’s approach respectively. Table III shows how well a neural network trained on one surgeon’s file perform in predicting the movement of the others. Only 1 file of each surgeon is cross-tested here.
The results show the feasibility of the basic approach of neural network error canceling in human-machine control. The neural network reduced the rmse of the surgeons’ erroneous motion by an average of 39.5%.

NN_{vol,mot} outperforms NN_{err} by 9.6% in rmse reduction. During training, both methods terminate at reaching the maximum number of hidden nodes. Riviere and Khosla used 100 input nodes throughout their experiment while we explore the effect of different input-hidden node combination on the performance of the neural networks. NN_{vol,mot} has better result may be because of the optimization of the network architecture in the number of input and hidden nodes, not necessary the superiority of one method over the other.

We observe that there are distinct differences in the motion profile between the outputs of NN_{err} and those of NN_{vol,mot}. Firstly the NN_{err} amplifies the high frequency noise (tremor) while NN_{vol,mot} attenuates the high frequency noise. This is because NN_{err} is trained to learn the high frequency error, while NN_{vol,mot} is trained to learn the smooth voluntary motion. In addition, qualitatively, NN_{vol,mot}'s estimation of voluntary motion has shown a better preservation of motion profile than NN_{err}.

The experiments on a neural network trained on one surgeon’s file in predicting the movement of the others has drawn inconclusive results. Out of the 12 runs, 3 has equal or better predictions than the NN trained on files from the same surgeon, 5 has worse rmse reduction result and 4 has worse rmse than the original error data files. Generally, as one would expect, the results show more worse (9/12) than better (3/12). However, the fact that the neural network of surgeon 1 in predicting the motion of surgeons 2 and 3 outperforms their own NN suggests that there might exist an optimal NN for a group of surgeons. More meaningful conclusion can only be drawn with more in-depth studies.

In the near future, we will be extending the current work to conduct experiments involving three dimensional hand movement data containing both voluntary and erroneous motion. The data collection will include more testing subjects performing tasks like tracing a predetermined path or tracking a predetermined trajectory.

VI. CONCLUSION

We trained and tested a cascade neural network with Kalman filtering on 15 hand movement data files collected from 4 surgeons. The neural network reduced the rmse of the surgeons’ erroneous motion by an average of 39.5%. NN_{vol,mot} outperforms NN_{err} by 9.6% in rmse reduction because its network architecture is optimized in terms of number of input and hidden nodes.

ACKNOWLEDGMENTS

We would like to express our gratitude to R. S. Rader for providing data and Y. Xu and M. Nechyba for software.

REFERENCES