33rd Annual International Conference of the IEEE EMBS
Boston, Massachusetts USA, August 30 - September 3, 2011

Dynamic Time Warping as a Spatial Assessment of Sensorimotor Impairment resulting from Stroke

Nathan Chalmers, Janice Glasgow, Stephen Scott

Abstract— **Robotic assessment of sensorimotor impairment began in the mid 1990s as a means to address some of the issues regarding inter-rater reliability and the lack of precision associated with traditional measures of sensorimotor impairment. Robotic measures of postural control, reaction time, movement smoothness, and movement error associated with robotic assessment of the upper-limb fail to recognize the inherent spatial and geometric differences between stroke and control hand path trajectories. In this study we propose the application of a class of algorithms, Dynamic Time Warping, designed to quantify the spatial difference and skew between hand written characters and vocal waveforms as a means for identifying individuals exhibiting sensorimotor impairment. In order to achieve this 85 stroke subjects, and 54 age, gender, and handedness matched control subjects, underwent robotic assessment of the upper-limb. Subjects were identified as either stroke or control using a K Nearest Neighbour classifier with a Dynamic Time Warping distance metric. Classification accuracy, sensitivity, and specificity in excess of %80 percent was achieved.**

I. Introduction

A cerebrovascular accident (CVA), or stroke, is a localized disturbance of blood flow to the tissues of the brain resulting from thrombosis, embolism, or hemorrhage and concluding in sudden cell death and tissue damage. Patient outcomes vary based on both the size and location of the lesion. These may include, but are not limited to, patient death or chronic conditions of varying severity such as aphasia and sensorimotor impairment in either the upper or lower limbs [1][2].

In order to address these impairments, stroke patients may receive multiple interventions from an occupational or physiotherapist targeted towards the areas which are believed to be affected. The decision to continue the application of said interventions and the overall success of the rehabilitative protocol is determined through observation of continued response on a set of traditional assessment metrics. Metrics such as the Functional Independence Measure [3], the Chedoke-McMaster Scale of Stroke Recovery [4], and the Fugl-Meyer Motor Assessment Scale [5] currently represent the gold standard of assessment of either function or impairment within the clinic. Patients are scored upon their perceived capacity to complete a set of predetermined tasks pertaining to their ability to perform daily tasks of living or to move, or resist movement of, the limb under examination. The results gained by manner of a qualitative survey are subsequently recorded on an ordinal scale. As a consequence of definition, the potential for inter-rater bias is introduced by the qualitative nature of the measurement. Conversely, in order to avoid such potential inter-rate bias the groupings of the ordinal scale may be made wide enough to encompass rater variability at the cost of reduced precision [6].

In order to address these potential failings, robotic platforms capable of capturing high frequency kinematic data regarding hand speed, acceleration, and position in a quantitative manner have been developed for the purpose of assessment of sensorimotor impairment. Though multiple platforms with subtle differences in construction and feature definitions exist, a common set of kinematic measures pertaining to the subject's postural control, reaction time, movement smoothness, and movement error have been derived [7]. Though such metrics have been successfully demonstrated to delineate between subject groups [7], they neglect to quantify the inherent spatial and geometric differences in hand path trajectories between stroke and control subjects illustrated in Figure 1.

Dynamic Time Warping (DTW) may function as measure of similarity suitable for comparing the spatial and geometric properties of two hand path trajectories. DTW is a distance metric which equates the similarity between two signals as a function of the degree to which the time axis of each signal must be skewed in order to construct an alignment such that the naive pairwise euclidean distance between the signals is minimal [8]. Though originally developed to measure the similarity between two waveforms in a speech recognition task, the algorithm has also successfully been applied to hand writing and gesture recognition [9]. Additionally, DTW has been shown to outperform other similar similarity measures such as Earth Mover's Distance, Frechet Disance, and Hausdorff Distance when applied to classifying time series data in the classical cylinder-bell-funnel problem [10].

Consequently, given the inherent spatial differences between control and stroke subject hand trajectories and the demonstrated ability of DTW to cluster and classify based upon the spatial difference in signal pairs, it is our belief that DTW is capable of distinguishing between stroke and control hand trajectories with high sensitivity and specificity as part of a K Nearest Neighbours classification protocol.

II. Methodology

A. Subjects

A population consisting of 85 stroke subjects from St. Mary's of the Lake Hospital in Kingston. Ontario, Canada and from Foothills Medical Centre in Calgary, Alberta,

The authors would like to thank Helen Bretzke, Troy Herter, Kim Moore, Justin Peterson, and Marie-Jo Demers for their support both collecting data and maintaining the backend systems which made this study possible

978-1-4244-4122-8/11/$26.00 ©2011 IEEE

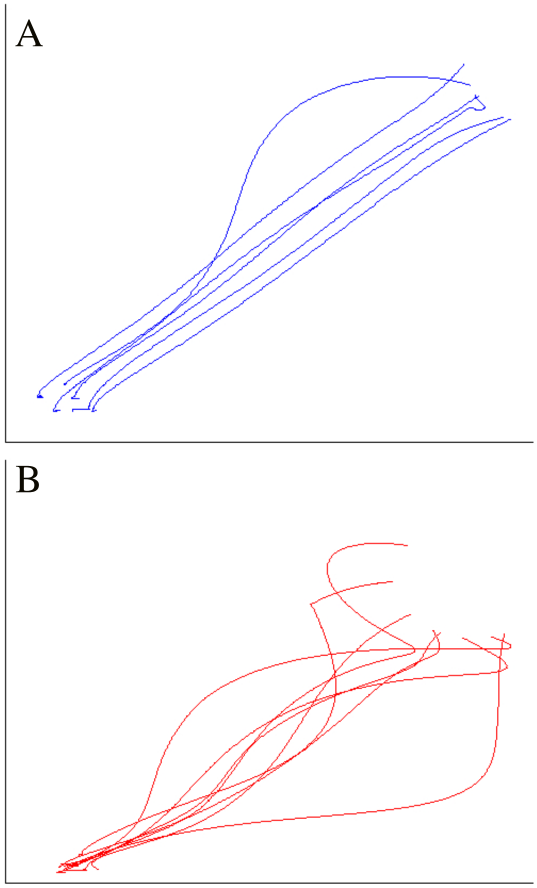

Fig. 1. A) Hand trajectories of a control subject, male age 55, from the central target to the 45 degree peripheral target. B) Hand trajectories of a stroke subject of identical gender and age exhibiting sensorimotor impairment in the right hand.

Canada were selected for robotic assessment of the upper-limbs. In order to qualify, all stroke subjects had to possess the ability to understand verbal commands, suffered only a singular cerebrovascular accident, had to have no additional history of neurological disorder, and had to have no history of musculoskeletal disorders which could impair their performance on the upper-limb assessment task. A similar age, gender, and handedness matched control population consisting of 54 healthy individuals with no known history of neurological disorder was selected from the surrounding areas.

Robotic assessment of stroke subjects was conducted between 14 and 45 days post-CVA. For some members of the population, additional robotic assessments were performed between 60 and 120, and 170 and 270 days post-CVA providing a total of 124 assessment points for the 85 stroke subjects. Prior to robotic assessment the upper-arm and hand portions of the Chedoke-McMaster stroke assessment scale were administered by a physiotherapist.

B. Task

Robotic assessment was conducted by recording subject performance on a centre-out reaching task within the horizontal plane using the Kinesiological Instrument for Normal and Altered Reaching Movement (KINARM), a bimanual exoskeleton robot and augmented reality display system manufactured by BKIN Technologies [11]. Subjects undergoing assessment were placed in the device such that their arms were both supported within the horizontal plane by the robotic exoskeleton and occluded from view by the horizontal virtual reality display. The centre target was placed such that it was located in the centre of the subject's natural workspace at $90°$ elbow flexion and $30°$ shoulder flexion. Eight peripheral objective targets were placed relative to the location of the centre target at $45°$ increments at a distance of 10cm.

Subjects were instructed that upon presentation of a peripheral target on the horizontal display, they were to move their hand from the centre target to the peripheral target indicated and stabilize. Subjects were allotted three seconds to complete the task, in which a randomized waiting period at the centre target between 1250ms and 1750ms was included. Eight trials to each of the 8 targets were conducted in random order for a total of 64 trials per hand. The task was replicated for both hands for a total of 128 trials per session.

C. Data

Four kinematic metrics pertaining to postural control, reaction time, movement smoothness, and movement error were derived from the recorded hand position, velocity, and acceleration data [7]. A subject's postural control was defined as their average hand speed during the random length postural window preceding the onset of the centre-out reaching task. Similarly, reaction time time was defined as the amount of elapsed time in milliseconds between presentation of the desired peripheral target and movement onset. In order to quantify the smoothness of the movement, the number of local hand speed maxima was determined. Finally, movement error was quantified as the amount of direction error resulting from the first sub-movement. The first sub-movement was defined as the period of time bounded by movement onset and the time at which the first local hand speed maxima occurred. Subsequently, the direction error inherent to the first sub-movement was calculated to be the angular difference between a vector spanning from the hand's position at movement onset and first sub-movement end time, and a vector spanning from the centre to the peripheral target. Only hand position data recorded during the period of time between presentation of the peripheral target and the time at which the subject's hand first entered the peripheral target was utilized to calculate the DTW distance between hand trajectories.

D. Dynamic Time Warping

Let A and B represent a pair of hand trajectories such that $A = (a_1, a_2, ..., a_n)$ and $B = (b_1, b_2, ..., b_m)$ where each point a_i or b_j represents the X and Y coordinate position of the hand within the global coordinate frame of KINARM at time i or j. Similarly, let $D(a_i, b_j) = \mid a_i - b_j \mid$ represent a distance function equivalent to the euclidean distance between trajectory A at time i and trajectory B

at time j. In order to isolate a subset of values $\hat{a}_i$ form A and $\hat{b}_j$ from B such the cumulative distance between matched points $\hat{a}_i$ and $\hat{b}_j$ is minimal, a warping function $W(i,j)$ can be introduced. Let $W(i,j)$ define a warping function such that $W(i,j) = 1$ for pairs (a_i, b_j) which are located on the warping path and produce a minimal sum of distances $D(a_i, b_j)$ and let $W(i,j) = 0$ otherwise. Therefore, the DTW distance between two hand trajectories can be calculated by deriving the warping function $W(i,j)$ such that Equation 1 is minimal [8][9].

$$DTW(A,B) = \left(\frac{\sum_i^n \sum_j^m W(i,j) D(a_i, b_j)}{\sum_i^n \sum_j^m W(i,j)} \right) \quad (1)$$

The warping function which provided an optimal solution for W in Equation 1 was found via the dynamic program outlined in Equation 2. Let $C(i,j)$ represent a cost matrix in which each element is equivalent to the total cost of creating a warping function which contains $D(a_i, b_j)$. Each point in $C(i,j)$ can be reached by either pairing the preceding points $C(i-1, j-1)$, removing a time step from trajectory A $C(i-1, j)$, or by removing a time step from trajectory B $C(i, j-1)$. Through observation of the principle of optimality, one notes in order to build an optimal warping function of length $k+1$ an optimal warping function of length k can be extended. In order to derive W, C was evaluated iteratively for all values such that $0 > i < n$ and $0 > j < m$, resulting in the warping function being extended to include the pair (a_n, b_m).

$$C(i,j) = D(A_i, B_j) + min \begin{pmatrix} C(i-1, j-1) \\ C(i, j-1) \\ C(i-1, j) \end{pmatrix} \quad (2)$$

E. Classification

K Nearest Neighbours (KNN), a classifier in which class inference is achieved by identifying the closest K data points in the feature space to a probe datapoint and then assigning the probe the class which comprises the majority of the closest neighbours, was selected to differentiate between stroke and control hand trajectories [12]. The use of KNN lent itself to the classification of hand trajectory data for two reasons, the first of which was the use of an explicit distance metric. Typically the distance between two feature vectors in a KNN classifier is determined by calculating the euclidean distance. In order to calculate the distance between two hand trajectories the euclidean distance was replaced with the DTW distance. In order to compare traditional robotic assessment features, a euclidean distance metric was used as part of the KNN classification protocol.

The second aspect of the KNN classifier which lent itself towards the classification of hand trajectory data was the training process. A KNN classifier is unique in that it is a supervised learning system which does not rely on a training process, and consequently does not rely upon class prototypes in order to base its classifications. Instead, the KNN classifier compares the current probe vector to the entirety of the feature space. Consequently, the use of a KNN classifier allowed for the use of stroke and control hand trajectories as recorded without the need to create stroke and control prototype hand trajectories; a process which would have been required for the use of a probabilistic artificial neural network.

TABLE I

CONFUSION MATRICES: DTW KNN CLASSIFIER

		Predicted				
Actual		Affected			Unaffected	
		S	C		S	C
	S	108	16	S	110	14
	C	6	48	C	23	33

TABLE II

CONFUSION MATRICES: TRADITIONAL KNN CLASSIFIER

		Predicted				
Actual		Affected			Unaffected	
		S	C		S	C
	S	110	14	S	120	4
	C	22	27	C	30	24

The K value of the KNN classifier implemented was determined through experimentation. A K value of 5 was found to give results with both high sensitivity and specificity.

III. RESULTS

Given the limited amount of data, classification was conducted using a leave one out verification strategy. Hand trajectory data from affected and non-affected arms was classified against control data separately. Control data sets for the affected and unaffected arm classification tasks were selected such that proportion of left and right hand dominant subjects remained identical between the stroke and control populations. Traditional robotic measures of impairment were classified in an identical matter.

A pair of confusion matrices describing the classification of the affected and non-affected arms of stroke subjects and a population of control subjects using a DTW similarity metric are presented in Table I. Similarly, Table II contains a pair of confusion matrices for the same subject classes using a euclidean distance metric to calculate the similarity between a 4 element vector describing the traditional robotic features postural control, reaction time, movement smoothness, and movement error. Table III illustrates additional classification performance parameters including percent accuracy, sensitivity, and specificity.

A binomial test of classifier performance indicated that the probability that the DTW KNN classifier outperformed the traditional classifier on affected hand trajectories by random chance was 0.11. Similarly, the probability that the Traditional KNN classifier outperformed the DTW KNN classifier on unaffected hand trajectories by random chance was found to be 5.2e-10. As illustrated in Table III, the DTW KNN classifier was also found to exhibit greater specificity for affected hand trajectories.

TABLE III

SPECIFICITY AND SENSITIVITY: DTW AND TRADITIONAL CLASSIFIERS

	DTW		Traditional	
	Affected	Unaffected	Affected	Unaffected
% Accuracy	0.88	0.79	0.79	0.81
Specificity	0.89	0.59	0.55	0.44
Sensitivity	0.87	0.89	0.89	0.96

IV. DISCUSSION

Though the DTW KNN classifier did not statistically outperform the KNN classifier based on traditional robotic assessment metrics, it did demonstrate an ability to produce statistically comparable results. This is particularly noteworthy when the amount of information available to each classifier is considered. The KNN built upon traditional robotic assessment measures had information pertaining to postural control, reaction time, movement smoothness, and movement error to base classification decisions. In contrast, the DTW KNN classifier had access to only one measure of similarity to base its classifications upon. If given access to additional sources of geometrical and spatial similarity, performance of the DTW KNN classifier may substantially improve.

However, despite a similar overall performance to the traditional KNN classifier, Table III illustrates the DTW KNN classifiers increased specificity when classifying the affected arms of stroke subjects. Specificity, is a term directly related to the true negative and false positive rates inherent to the classifier. This suggests that the DTW KNN classifier has a lower probability of identifying an affected stroke hand trajectory as a control hand trajectory by a margin of %30 when compared to similar analysis by the DTW KNN classifier on unaffected hand trajectories and the traditional KNN classifier on both affected and unaffected hand trajectories. Such a property may suggest that strong spatial and geometric differences exist between hand trajectories in the affected arms of stroke subjects and the hand trajectories of control subjects and the unaffected arms of stroke subjects. The KNN classifier based upon traditional robotic assessment metrics does not posses this spatial information, and in turn may explain why the false positive rate is high in proportion to the true negative rate for both affected and unaffected hand trajectories.

Furthermore, resulting from its high specificity rate, a DTW metric of sensorimotor impairment in stroke subjects may find applications outside of the clinic and inside an emergency room environment. The diagnosis of a CVA immediately after onset is difficult given that the resulting lesions do not typically appear on Computer Aided Topography or Magnetic Resonance Imaging scans until a period of 24 to 48 hours after the accident [1][2]. Given the DTW KNN classifiers low false positive and high true negative rates, such a system may be able to separate individuals who have suffered a stroke from those who have been suspected of suffering a stroke with a high degree of accuracy.

V. CONCLUSION

A significant number of robotic assessment platforms designed to quantitatively record kinematic data for the purposes of measuring sensorimotor impairment record limb position data in addition to traditional measures of robotic assessment such as postural control, reaction time, movement smoothness, and movement error throughout the course their evaluation. Dynamic Time Warping as a similarity metric as part of a K Nearest Neighbours classification protocol effectively demonstrates that this limb position data can provide classification results that are comparable to traditional robotic assessment metrics and in some instances outperform traditional robotic assessment metrics on measures of specificity. However, further research is needed to model how spatial and geometric similarity as measured by DTW changes in sub-acute, acute, and chronic stroke subjects in comparison to traditional robotic assessment metrics. Ultimately, the performance of the implemented DTW KNN classifier demonstrates that this is an area of sensorimotor impairment assessment measures that should not continued to be ignored.

REFERENCES

[1] J. Adams, HP, B. Bendixen, L. Kappelle, J. Biller, B. Love, D. Gordon, and d. Marsh, EE, "Classification of subtype of acute ischemic stroke. Definitions for use in a multicenter clinical trial. TOAST. Trial of Org 10172 in Acute Stroke Treatment," *Stroke*, vol. 24, no. 1, pp. 35–41, 1993.

[2] R. P. Donahue, R. D. Abbott, D. M. Reed, and K. Yano, "Alcohol and Hemorrhagic Stroke: The Honolulu Heart Program," *JAMA*, vol. 255, no. 17, pp. 2311–2314, 1986.

[3] K. Ottenbacher, Y. Hsu, C. Granger, and R. Fiedler, "The reliability of the functional independence measure: a quantitative review," *Arch Phys Med Rehabil*, vol. 77, no. 12, pp. 1226–1232, 1996.

[4] C. Gowland, P. Stratford, M. Ward, J. Moreland, W. Torresin, S. Van Hullenaar, J. Sanford, S. Barreca, B. Vanspall, and N. Plews, "Measuring physical impairment and disability with the chedoke-mcmaster stroke assessment," *Stroke*, vol. 24, no. 1, pp. 58–63, 1993.

[5] D. J. Gladstone, C. J. Danells, and S. E. Black, "The fugl-meyer assessment of motor recovery after stroke: A critical review of its measurement properties," *Neurorehabil Neural Repair*, vol. 16, no. 3, pp. 232–240, Sep 2002.

[6] P. S. Lum, D. J. Reinkensmeyer, R. Mahoney, W. Z. Rymer, and C. G. Burgar, "Robotic devices for movement therapy after stroke: current status and challenges to clinical acceptance," *Top Stroke Rehabil*, vol. 8, no. 4, pp. 40–53, Oct 2003.

[7] A. M. Coderre, A. A. Zeid, S. P. Dukelow, M. J. Demmer, K. D. Moore, M. J. Demers, H. Bretzke, T. M. Herter, J. I. Glasgow, K. E. Norman, S. D. Bagg, and S. H. Scott, "Assessment of upper-limb sensorimotor function of subacute stroke patients using visually guided reaching," *Neurorehabil Neural Repair*, vol. 24, no. 6, pp. 528–41, Jan 2010.

[8] P. Somervuo and T. Kohonen, "Self-organizing maps and learning vector quantization for feature sequences," *Neural Processing Letters*, vol. 10, no. 2, pp. 151–159, 1999.

[9] H. Sakoe and S. Chiba, "Dynamic programming algorithm optimization for spoken word recognition," *Acoustics, Speech and Signal Processing, IEEE Transactions on*, vol. 26, no. 1, pp. 43 – 49, 1978.

[10] P. Sung, Z. Syed, and J. Guttag, "Quantifying morphology changes in time series data with skew," *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech and Signal Processing-Volume 00*, pp. 477–480, 2009.

[11] S. H. Scott, "Apparatus for measuring and perturbing shoulder and elbow joint positions and torques during reaching," *Journal of Neuroscience Methods*, vol. 89, no. 2, pp. 119 – 127, 1999.

[12] T. Cover and P. Hart, "Nearest neighbor pattern classification," *Information Theory, IEEE Transactions on*, vol. 13, no. 1, pp. 21 – 27, jan 1967.