An Analytic Approach to Decipher Usable Gestures for Quadriplegic Users

Hairong Jiang
School of Industrial Engineering
Purdue University
West Lafayette, USA
jiang115@purdue.edu

Bradley S. Duerstock
School of Industrial Engineering
Weldon School of Biomedical Engineering
Purdue University
West Lafayette, USA
bsd@purdue.edu

Juan P. Wachs*
School of Industrial Engineering
Purdue University
West Lafayette, USA
jpwachs@purdue.edu

***Abstract*— With the advent of new gaming technologies, hand gestures are gaining popularity as an effective communication channel for human computer interaction (HCI). This is particularly relevant for patients recovering from mobility-related injuries or debilitating conditions who use gesture-based gaming for rehabilitation therapy. Unfortunately, most gesture-based gaming systems are designed for able-bodied users and are difficult and costly to adapt to people with upper extremity mobility impairments. While interface customization is an active area of work in assistive technologies (AT), there is no existing formal and analytical grounded methodology to adapt gesture-based control systems for quadriplegics. The goal of this work is to solve this hurdle by developing a mathematical framework to project the patterns of gestural behavior designed for existing gesture systems to those exhibited by quadriplegic subjects due to spinal cord injury (SCI). A key component of our framework relied on Laban movement analysis (LMA) theory, and consisted of four steps: acquiring and preprocessing gesture trajectories, extracting feature vectors, training transform functions, and generating constrained gestures. The feasibility of this framework was validated through user-based experimental paradigms and subject validation. It was found that 100% of gestures selected by subjects with high-level SCIs came from the constrained gesture set. Even for the low-level quadriplegic subject, the alternative gestures were preferred.**

Keywords— hand gesture-based interfaces; assistive technology; mobility impairments; Laban space.

I. Introduction

In the last few years, gesture-based interfaces have become increasingly popular for applications, such as entertainment [1], healthcare [2], robotics [3], communication [4], and transportation [5]. The application that has received most traction is arguably "gaming" [6]. Recent studies have also shown that playing games can substantially improve the well-being [7] and recovery of function [8] in stroke [9], multiple-sclerosis [10] and Parkinson's disease [11] rehabilitation patients. Unfortunately, commercial gesture-based consoles, such as the Wii® and Xbox®, have been developed without considering users' motor limitations. While there have been individual, spontaneous, and unstructured customizing gesture-based interfaces developed for people with disabilities (PWDs) [12]. For most, these systems lead to suboptimal solutions, and adopted ad-hoc methods rather than generalizable solutions. There is no existing methodology to convert a gesture-based interface designed for able-bodied individuals to a usable and effective interface for PWDs without redesigning the interface from scratch. The aim of this work is to project existing patterns of gestural behavior to correspond to those of users with upper extremity mobility impairments, thereby making commercial gesture-based interfaces widely usable by quadriplegics.

Previous work leveraged on the theory of Laban movement analysis (LMA) proposed by Laban [13] to characterize gestures. This theory can be of paramount importance for finding the common patterns in gestures performed by PWDs. The LMA method has four major components (*Body, Space, Shape and Effort*) [14]. To simplify its representation, Norman Badler developed a special notation called "Labanotation" [15] to describe human movements using LMA. Rett and Dias [16] discussed the modeling and implementation of LMA. Santos and Dias [17] focused on converting and interpreting human motion signals into a series of features based on the study of body trajectories. The main contribution of their work was the design of the gesture lexicon consisting of many motion-entities which were defined though LMA parameters. To analyze the relationships between these motion entities, Bayesian networks can be applied [18].

The use of LMA for characterizing gesture sets is part of a more general approach for determining gestural vocabularies, called "analytical-based" [19]. This type of approach builds on mathematical models to determine an optima gesture set (lexicon). There are also the "technology-based" [20] and "human-based" approaches [21]. The gestures selected by the technology-based approach were easily "recognizable" by the machine, however may be difficult to perform and remember by quadriplegic users. In contrast, the human-based approach established the gesture vocabulary by maximizing usability-based metrics (e.g. such as satisfaction and comfort) [22].

This paper proposes three main contributions: (a) propose a new analytical approach based on transforming gestures from different manifold spaces, called the Laban Transform; (b) project existing gesture lexicons from gaming applications into a new set of gestures suitable for users with quadriplegia; and (c) validate and determine the usability of the constrained gestures with quadriplegic users.

This project was supported by the National Institute of Health Director's ARRA Pathfinder Award to Promote Diversity in the Scientific Workplace (1DP4GM096842-01).

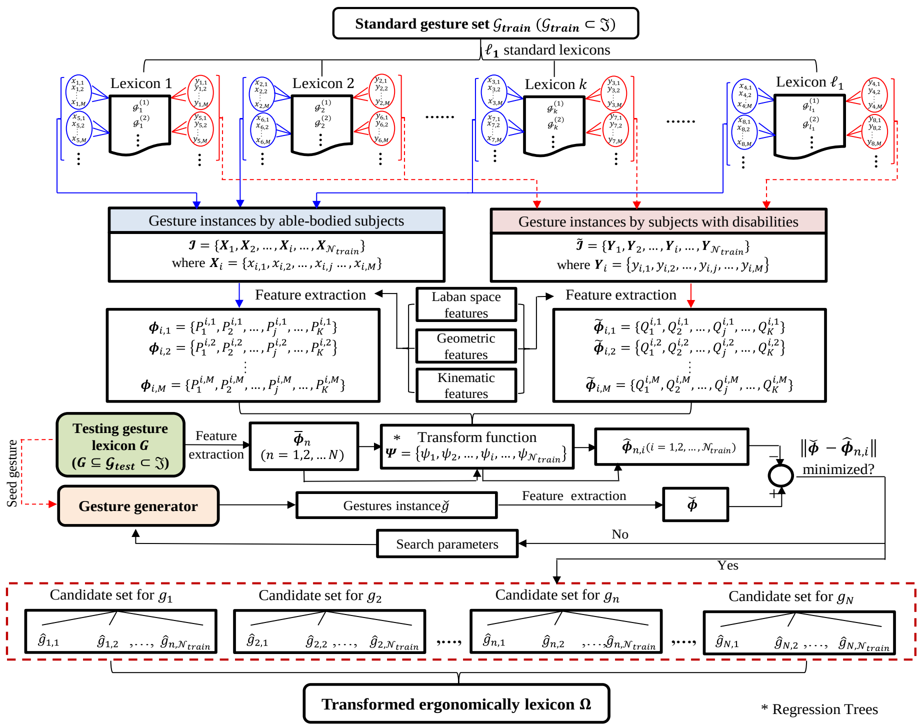

Fig. 1 System architecture

II. Problem Definition

The main problem in this work addressed how to project standard gestures from a known manifold to a constrained (unknown) manifold that corresponds to the space and effort that persons with quadriplegia can perform. The term "standard gestures" is denoted as gestures designed for able-bodied individuals. A "standard gesture lexicon" is referred to a set of standard gestures used for a gesture-based interface. To meet the goal of making commercial consoles available for users with quadriplegia, L standard gesture lexicons (denoted as $\boldsymbol{Q}_1, \boldsymbol{Q}_2, \dots \boldsymbol{Q}_L$) are selected. The union is denoted as $\mathfrak{J}$ (Eq. 1).

$$\mathfrak{J} = \boldsymbol{Q}_1 \cup \boldsymbol{Q}_2 \dots \cup \boldsymbol{Q}_L \tag{1}$$

Let $\boldsymbol{G}$ represent a standard lexicon with N gestures, where $\boldsymbol{G} \subset \mathfrak{J}$. $\widetilde{\boldsymbol{G}}$ is a constrained gesture lexicon corresponding to $\boldsymbol{G}$. g_n and $\tilde{g}_n$ (n=1,2,...,N) denote the nth gesture in $\boldsymbol{G}$ and $\widetilde{\boldsymbol{G}}$, respectively (Eq. 2 and Eq. 3). Let $\breve{g}$ denote an arbitrary gesture, $\mathcal{L}$ represents a mapping from a gesture trajectory to a feature vector, and $\boldsymbol{\Psi}$ be a pre-trained transform function between the feature vector of a standard gesture and that of a constrained gesture (see Section III for details). The problem is interpreted as: finding a constrained gesture lexicon to satisfy (Eq. 4 and 5).

$$\boldsymbol{G} = \{g_1, g_2, \dots, g_n, \dots, g_N\} (n = 1,2, \dots, N) \tag{2}$$

$$\widetilde{\boldsymbol{G}} = \{\tilde{g}_1, \tilde{g}_2, \dots, \tilde{g}_n, \dots, \tilde{g}_N\} (n = 1,2, \dots, N) \tag{3}$$

$$\tilde{g}_n = \arg\min_{\breve{g}} \|\mathcal{L}(\breve{g}) - \boldsymbol{\Psi}(\mathcal{L}(g_n))\| \tag{4}$$

$$\text{s.t. } n \leq N, n \in \mathbb{Z}^+, g_n \in \boldsymbol{G}, \text{ and } \tilde{g}_n \in \widetilde{\boldsymbol{G}} \tag{5}$$

In this paper, an analytic approach is presented to solve this problem (minimize Eq. 4). A set of gestures are collected to train the model and once the model is trained, it is tested using a testing lexicon. The union of the standard gesture lexicons $\mathfrak{J}$ is further divided into two subsets: one is used to collect the gesture instances for training (denoted as $\boldsymbol{\mathcal{G}}_{train}$) and the other is used for testing (denoted as $\boldsymbol{\mathcal{G}}_{test}$), where Eq. 6 is satisfied. $\bar{g}_i$ and $\bar{\bar{g}}_j$ represent the gesture in $\boldsymbol{\mathcal{G}}_{train}$ and $\boldsymbol{\mathcal{G}}_{test}$. $\mathcal{N}_{train}$ and $\mathcal{N}_{test}$ are the number of gestures in $\boldsymbol{\mathcal{G}}_{train}$ and $\boldsymbol{\mathcal{G}}_{test}$ (Eq. 7 and Eq. 8).

$$\boldsymbol{\mathcal{G}}_{train} \cup \boldsymbol{\mathcal{G}}_{test} = \mathfrak{J}, \qquad \boldsymbol{\mathcal{G}}_{train} \cap \boldsymbol{\mathcal{G}}_{test} = \emptyset \tag{6}$$

$$\boldsymbol{\mathcal{G}}_{train} = \{\bar{g}_1, \bar{g}_2, \dots, \bar{g}_i, \dots, \bar{g}_{\mathcal{N}_{train}}\} (i = 1,2, \dots, \mathcal{N}_{train}) \tag{7}$$

$$\boldsymbol{\mathcal{G}}_{test} = \{\bar{\bar{g}}_1, \bar{\bar{g}}_2, \dots, \bar{\bar{g}}_j, \dots, \bar{\bar{g}}_{\mathcal{N}_{test}}\} (j = 1,2, \dots, \mathcal{N}_{test}) \tag{8}$$

III. Methods

The architecture of the analytic gesture generation approach to solve the problem in Section II is shown in Fig. 1. This approach consists of the following four steps: *A-D*.

A. Acquiring and Preprocessing Gesture Trajectories

To collect the gesture instances (trajectories) for training, both able-bodied and quadriplegic subjects were recruited. Each gesture ($\bar{g}_i$) in $\boldsymbol{\mathcal{G}}_{train}$ was presented to subjects via slideshows. The subjects were then asked to perform each gesture M times and to follow the presented gesture trajectory as much as possible. While the subject performed a given gesture, the 3D coordinates of the hands were acquired using a Kinect camera. Each gesture instance (j) obtained from a trial (i) is denoted as $x_{i,j}$ for able-bodied subjects, and $y_{i,j}$ for subjects with quadriplegia (Eq. 9 and Eq. 10). Here, one trial corresponds to the gestures generated from one slide in the slideshow. The function $\mathcal{f}$ represents the mapping from the subjects' performance of a gesture to the corresponding trajectory. The set of instances for each standard gesture is denoted as $\boldsymbol{X}_i$ and $\boldsymbol{Y}_i$ (Eq. 11 and Eq. 12). Following this procedure, the set of gesture instances collected from able-bodied individuals (denoted as $\boldsymbol{\mathcal{I}}$) and subjects with quadriplegia (denoted as $\tilde{\boldsymbol{\mathcal{I}}}$) is obtained (Eq. 13 and Eq. 14). The union ($\boldsymbol{\mathcal{S}}$) of all the gesture instances is expressed in Eq. 15.

Two steps (outlier removal and smoothing) were employed for the acquired gesture instances to reduce noise and the variability exhibited by the users. Outliers were those trajectory points further than 3σ from the mean. A Kalman filter is employed to smooth the 3D gesture trajectories.

$$x_{i,j} = \not{f}(\bar{g}_i) \tag{9}$$

$$y_{i,j} = \not{f}(\bar{g}_i) \tag{10}$$

$$\boldsymbol{X}_i = \{x_{i,1}, x_{i,2}, \dots, x_{i,j}, \dots, x_{i,M}\} \tag{11}$$

$$\boldsymbol{Y}_i = \{y_{i,1}, y_{i,2}, \dots, y_{i,j}, \dots, y_{i,M}\} \tag{12}$$

$$\boldsymbol{\mathcal{I}} = \{\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_i, \dots, \boldsymbol{X}_{\mathcal{N}_{train}}\} \tag{13}$$

$$\tilde{\boldsymbol{\mathcal{I}}} = \{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \dots, \boldsymbol{Y}_i, \dots, \boldsymbol{Y}_{\mathcal{N}_{train}}\} \tag{14}$$

$$\boldsymbol{\mathcal{S}} = \{\boldsymbol{\mathcal{I}}, \tilde{\boldsymbol{\mathcal{I}}}\} \tag{15}$$

B. Feature Extraction

Each gesture trajectory is encoded into a feature vector with dimensionality *K* (number of features per gesture). Two principles were followed for feature selection; (a) *generable*: representative of the user target population (e.g. quadriplegics); and b) *separable*: differentiable between standard gestures and those within the constrained gesture space. To satisfy the aforementioned requirements, a union made of Laban space, and kinematic and geometric based features was created.

The Laban space features can provide a good representation of the limitations experienced by people with upper extremity physical impairments. Features based on Space, Effort, and Shape were adopted. The symbolic representation developed by Longstaff et al., [23] is used to extract features representing the Space component. The Effort component is expressed by directness, inertia, and duration of a gesture trajectory [19]. The volume of the trajectory is used to quantify the Shape component. The kinematic characteristics of a given gesture trajectory are described by the velocity, acceleration, and jerk component of the motion. The average, maximum and minimum value of these three parameters are selected to construct the kinematic feature set. Each of them is extracted from the gesture trajectory and treated as a component of the feature vector. Since the gesture trajectory is a curve, its geometric characteristics can be represented using four features often used for curve representation: arc length, curvature, torsion, and number of inflection points. These features are adopted as a complement to the kinematic features, and they are key differentiators of the standard and constrained gestures. The extracted features are normalized to lie within the 0-1 range.

C. Transform Functions Computation

This section describes the process of acquiring a set of transform functions associated with the set of gesture instances $\boldsymbol{\mathcal{S}}$. Let $\boldsymbol{\phi}_{i,j} \in \mathbb{R}^K$ (Eq. 16) denote a vector comprised by all the features extracted from a gesture instance $x_{i,j}$. Similarly, $\tilde{\boldsymbol{\phi}}_{i,j} \in \mathbb{R}^K$ (Eq. 17) is a vector consisting of all the features extracted from a constrained gesture instance $y_{i,j}$ $(i = 1,2,\dots,\mathcal{N}_{train}; j = 1,2,\dots,M)$. $\mathcal{L}$ represents the projection from a gesture instance to a feature vector. Let the set consisting of all the feature vectors associated with a given gesture $\bar{g}_i$ for able and disabled bodied individuals be $\boldsymbol{\Phi}_i$ and $\tilde{\boldsymbol{\Phi}}_i$, respectively (Eq. 18 and 19). The transform function (ψ_i) for each gesture $\bar{g}_i$ in $\boldsymbol{\mathcal{G}_{train}}$ is then computed using regression trees [24] in the following way: for each transform function ψ_i, a binary regression tree is obtained based on the input and output variables $\boldsymbol{\Phi}_i$ and $\tilde{\boldsymbol{\Phi}}_i$ (Eq. 20) so a regression error is minimized. The set of transformation functions ($\boldsymbol{\Psi}$) for all the gestures in the standard lexicon is given by $\boldsymbol{\Psi} = \{\psi_1, \psi_2, \dots, \psi_i, \dots, \psi_{\mathcal{N}_{train}}\}$.

$$\boldsymbol{\phi}_{i,j} = \mathcal{L}(x_{i,j}) \tag{16}$$

$$\tilde{\boldsymbol{\phi}}_{i,j} = \mathcal{L}(y_{i,j}) \tag{17}$$

$$\boldsymbol{\Phi}_i = [\boldsymbol{\phi}_{i,1}, \boldsymbol{\phi}_{i,2}, \dots, \boldsymbol{\phi}_{i,M}] \tag{18}$$

$$\tilde{\boldsymbol{\Phi}}_i = [\tilde{\boldsymbol{\phi}}_{i,1}, \tilde{\boldsymbol{\phi}}_{i,2}, \dots, \tilde{\boldsymbol{\phi}}_{i,M}] \tag{19}$$

$$(\tilde{\boldsymbol{\Phi}}_i)_{K\times M} = (\psi_i)_{K\times K}(\boldsymbol{\Phi}_i)_{K\times M} \tag{20}$$

D. Constrained Gesture Generation

A two-step iterative process is proposed to generate a candidate gesture set using the acquired transform function $\boldsymbol{\Psi}$ and a gesture generator. The first step consisted of projecting the feature vector of a gesture from the standard to the constrained space using $\boldsymbol{\Psi}$. The second step consisted of generating gestures in the vicinity space of the given arbitrary gesture through a gesture generator. The generated gesture's feature vector is then compared to the constrained feature vector. If the distance between the two vectors is minimum (the distance does not decreases more than *ε*), then the gesture is kept as a candidate gesture. Otherwise, the gesture is discarded and a new gesture is generated. This process is iteratively conducted until a complete candidate set is obtained for all the gestures in the testing lexicon.

In the first step, a gesture lexicon $\boldsymbol{G} \subseteq \boldsymbol{\mathcal{G}_{test}}$ is selected for testing (see Section II). Able-bodied subjects are asked to perform *M* times each gesture g_n in $\boldsymbol{G}$. The set of collected gesture instances for g_n is converted to trajectories following a similar process as the one explained in Section III.A, and is denoted as $\check{X}_n$. Then, the gesture encoding approach proposed by Calinon et al. [25] is applied to obtain the mean gesture trajectory from the set of trajectories $\check{X}_n$. This consists of building a Gaussian Mixture Model (GMM) from 3D trajectories' data points of all the gesture instances in $\check{X}_n$. To determine the parameters of the Gaussians, the Expectation Maximization algorithm [26] is used. The K-means clustering technique was used to give the initial estimate of these parameters. Then the mean gesture trajectory (denoted as $\breve{g}_n$) is obtained using Gaussian Mixture Regression (GMR). To obtain the GMR, the joint density is computed using the parameters estimated before, from the GMM. This way, GMM and GMR are used to encode the gesture trajectories collected from able-bodied subjects and obtain a mean standard gesture trajectory. The feature vector denoted as $\overline{\boldsymbol{\phi}}_n (n = 1,2, \dots, N)$ (with features presented as in Section *B*) is computed for each mean gesture trajectory $\breve{g}_n$ (Eq. 21). The transform function $\boldsymbol{\Psi} = \{\psi_1, \psi_2, \dots, \psi_2, \dots, \psi_{\mathcal{N}_{train}}\}$ is then applied to map $\overline{\boldsymbol{\phi}}_n$ to a set of constrained feature vector $\widehat{\boldsymbol{\phi}}_{n,i} (i = 1,2, \dots, \mathcal{N}_{train})$ (Eq. 22). Thus, for each gesture g_n, $\mathcal{N}_{train}$ constrained feature vectors $(\widehat{\boldsymbol{\phi}}_{n,1}, \widehat{\boldsymbol{\phi}}_{n,2}, \dots, \widehat{\boldsymbol{\phi}}_{n,i}, \dots, \widehat{\boldsymbol{\phi}}_{n,\mathcal{N}_{train}})$ are projected using $\boldsymbol{\Psi}$. The feature vectors acquired in this step represent the characteristic constrained gesture trajectories. The goal is to determine the constrained gestures from the constrained feature vectors' available information. However, since the trajectories possess more information than their corresponding feature vectors, the process of obtaining a gesture trajectory from its inverse Laban transform $\mathcal{L}^{-1}(\overline{\boldsymbol{\phi}}_n) = g_n$ is not analytically possible.

$$\overline{\boldsymbol{\phi}}_n = \mathcal{L}(\breve{g}_n)(n = 1,2, \dots, N) \tag{21}$$

$$\widehat{\boldsymbol{\phi}}_{n,i} = \psi_i(\overline{\boldsymbol{\phi}}_n)(i = 1,2, \dots, \mathcal{N}_{train}) \tag{22}$$

To solve this hurdle, the second step incorporated a pseudo-random gesture generation process (Fig. 2) using a combination of the gesture encoding approach (as described in the first step) and a neighborhood search method. This search procedure starts by an initial solution (or seed gesture). This seed gesture, denoted as $\breve{g}$, is obtained through the following procedure: 3D data points of each trajectory in $\breve{X}_n$ are projected onto a 2D space by using principal component analysis (PCA) (denoted as $\breve{\xi}_n$). Then, the same gesture encoding approach explained earlier (applying GMM and GMR) is used to obtain a mean gesture trajectory, which acts as the seed gesture $\breve{g}$. In the first iteration of the search procedure, the generated gesture equals to the seed gesture. A feature vector $\breve{\boldsymbol{\phi}}$ (see Section *B* and *C*) is then computed from the generated gesture and compared with the constrained feature vector $\widehat{\boldsymbol{\phi}}_{n,i}$ (Eq. 23 and 24). Since $\widehat{\boldsymbol{\phi}}_{n,i}$ characterize the constrained gestures, we need to find a gesture trajectory that can minimize a distance metric between $\breve{\boldsymbol{\phi}}$ and $\widehat{\boldsymbol{\phi}}_{n,i}$. A parameter search (a neighborhood search) [27] is conducted to tune the parameters of the Gaussian and generate a new gesture trajectory, $\breve{g}$, and the comparison process is repeated. When the distance between $\breve{\boldsymbol{\phi}}$ and $\widehat{\boldsymbol{\phi}}_{n,i}$ is minimized, the mean trajectory resulting from GMR is kept as a candidate gesture $\hat{g}_{n,i}$ (Eq. 24). This gesture generation process is conducted for all the gestures in $\boldsymbol{G}$ (refer to Algorithm 1). For each gesture g_n, $\mathcal{N}_{train}$ constrained gestures are obtained to constitute the set $\widehat{\boldsymbol{G}}_n$ (Eq. 25). The union of all the constrained gesture set $\widehat{G}_n$ is denoted as $\boldsymbol{\Omega}$ (Eq. 26). Sample results for the gesture generation step are shown in Fig. 3.

$$\breve{\boldsymbol{\phi}} = \mathcal{L}(\breve{g}) \tag{23}$$

$$\hat{g}_{n,i} = \arg\min_{\breve{g}} \|\breve{\boldsymbol{\phi}} - \widehat{\boldsymbol{\phi}}_{n,i}\| \tag{24}$$

$$\widehat{\boldsymbol{G}}_n = \{\hat{g}_{n,1}, \hat{g}_{n,2}, \dots, \hat{g}_{n,i}, \dots, \hat{g}_{n,\mathcal{N}_{train}}\} \tag{25}$$

$$\boldsymbol{\Omega} = \widehat{\boldsymbol{G}}_1 \cup \widehat{\boldsymbol{G}}_2 \cup \widehat{\boldsymbol{G}}_n \cup \widehat{\boldsymbol{G}}_N \tag{26}$$

Algorithm 1 Constrained Gesture Generation

Input: a standard gesture lexicon $\boldsymbol{G} = \{g_1, g_2, \dots, g_n, \dots, g_N\}$
Output: constrained candidate gesture set $\boldsymbol{\Omega} = \{\widehat{\boldsymbol{G}}_1, \widehat{\boldsymbol{G}}_2, \dots, \widehat{\boldsymbol{G}}_n, \dots, \widehat{\boldsymbol{G}}_N\}$, where $\widehat{\boldsymbol{G}}_n = \{\hat{g}_{n,1}, \hat{g}_{n,2}, \dots, \hat{g}_{n,i}, \dots, \hat{g}_{n,\mathcal{N}_{train}}\}$
for $n = 1{:}\,N$
// Feature vector projection
  // Feature extraction
  $\overline{\boldsymbol{\phi}}_n = \mathcal{L}(g_n)$
  for $i = 1{:}\,\mathcal{N}_{train}$
    // Laban transform $\boldsymbol{\Psi} = \{\psi_1, \psi_2, \dots, \psi_i, \dots, \psi_{\mathcal{N}_{train}}\}$
    $\widehat{\boldsymbol{\phi}}_{n,i} = \psi_i(\overline{\boldsymbol{\phi}}_n)$
    // Feature extraction for a generated trajectory $\breve{g}$
    $\breve{\boldsymbol{\phi}} = \mathcal{L}(\breve{g})$
    // Neighborhood search and gesture generation
    $\hat{g}_{n,i} = \arg\min_{\breve{g}} \|\breve{\boldsymbol{\phi}} - \widehat{\boldsymbol{\phi}}_{n,i}\|$
  end
  $\widehat{\boldsymbol{G}}_n = \{\hat{g}_{n,1}, \hat{g}_{n,2}, \dots, \hat{g}_{n,i}, \dots, \hat{g}_{n,\mathcal{N}_{train}}\}$
end
$\boldsymbol{\Omega} = \{\widehat{\boldsymbol{G}}_1, \widehat{\boldsymbol{G}}_2, \dots, \widehat{\boldsymbol{G}}_n, \dots, \widehat{\boldsymbol{G}}_N\}$

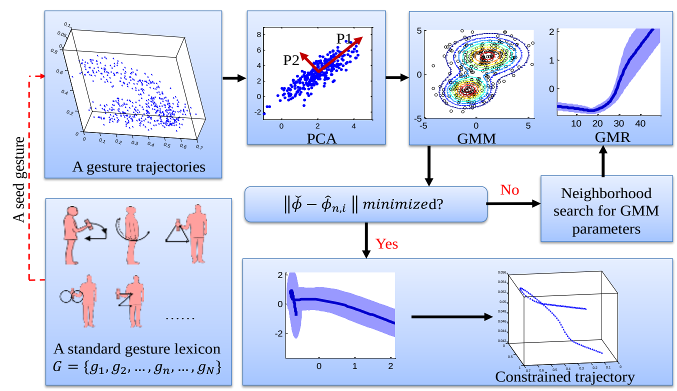

Fig. 2 Gesture generator

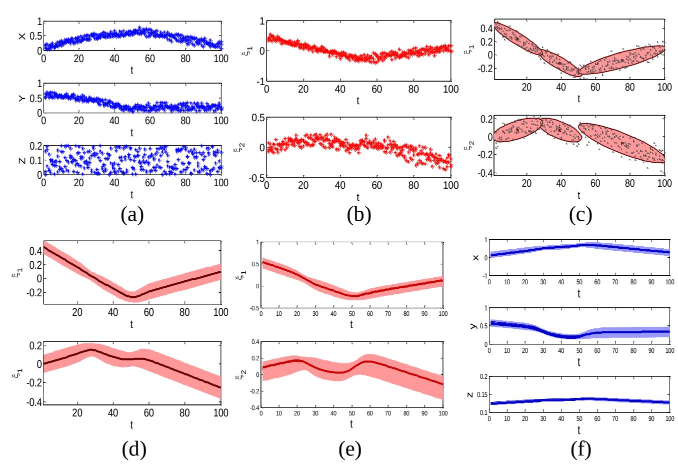

Fig. 3 Sample results of gesture generation. (a) 3D data; (b) 2D data using PCA; (c) GMM model; (d) GMR results; (e) neighborhood search results; (f) 3D data from back-projecting of 2D data after neighborhood search.

IV. Results & Discussion

A. Experimental Results

Four able-bodied subjects and three subjects with Cervical 4 (C4) to Cervical 5 (C5) SCIs were recruited to train the set of transform functions. The framework explained in Section III was applied (see Fig. 1) to obtain the candidate constrained gesture set $\widehat{\boldsymbol{G}}_n (n = 1,2, \ldots, N)$ for each gesture g_n in the testing lexicon. The standard gesture lexicons used in this experiment is $\mathfrak{J}$ ={"*Xbox*", "*PointGrab*", "*Wisee*", "*Win8*"} (see Appendix). The set of gesture lexicons for training is $\boldsymbol{\mathcal{G}}_{train}$ ={"*Xbox*", "*PointGrab*", "*Win8*"} (Fig. 4(a), (b), and (c)) and for testing is $\boldsymbol{\mathcal{G}}_{test}$ ={"*Wisee*"} (Fig. 4(d)). Note that each lexicon included a number of gestures. Given $\boldsymbol{G} = \boldsymbol{\mathcal{G}}_{test}$, the objective is to generate the constrained gesture set $\widetilde{\boldsymbol{G}}$ corresponding to $\boldsymbol{G}$ (as explained in Section II). The number of gestures in "*Xbox*", "*PointGrab*", and "*Win8*" was five, four, and eight, respectively. Since for each gesture in $\boldsymbol{\mathcal{G}}_{train}$, a pre-trained transform function set $\boldsymbol{\Psi}$ is computed, the number of transform functions obtained is seventeen (5+4+8). Thus, by projecting each gesture g_n in $\boldsymbol{G}$ using the set of transform functions $\boldsymbol{\Psi}$, seventeen candidate gestures were obtained.

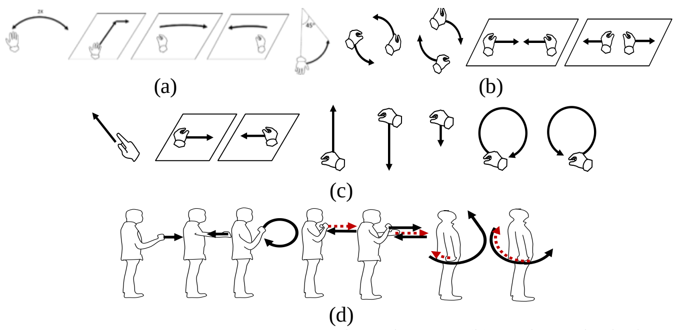

Fig. 4 Standard gesture lexicons (a) "Xbox" lexicon; (b) "PointGrab" lexicon; (c) "Win8" lexicon; (d) "Wisee" lexicon.

Fig. 5 illustrates the set of candidate gestures ($\widehat{\boldsymbol{G}}_n$) resulting from the proposed approach. The figures displayed present varied forms of the original gestures. Most of the gestures exhibit more curvature than the original ones $g_i \in \boldsymbol{G}$. Based only on appearance, it is not possible to assess their usability. To further evaluate the constrained gestures, a subjective validation was conducted with users with quadriplegia in the next section.

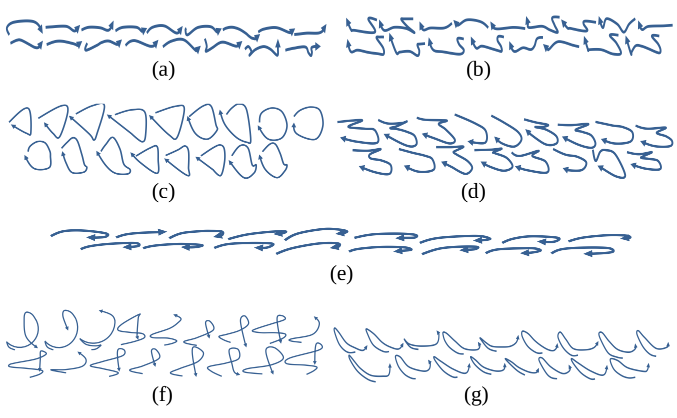

Fig.5 Candidate gestures for the "Wisee" lexicon

B. Gesture Validation

Four subjects with upper extremity mobility impairments (one with Neurofibroma, two with C4 to C5 SCIs and one with a C7 SCI) were recruited in a subjective validation experiment to evaluate the constrained gestures generated by the proposed approach (Fig. 5). The subjects were asked to respond to two questions: (1) how confident you feel you can perform the given gesture? (gestures in Fig. 4 (d)) (Q1); (2) choose one alternative gesture better than the gesture in Q1 (Q2). For Q1, a standard gesture in the "*Wisee*" lexicon was shown to the subjects via a slideshow. The subjects were required to use the Borg scale [28] (0-10) to measure the difficulty of the given gesture. The higher the score, the more difficult the gesture was to perform. For Q2, the gesture illustrated in Q1 as well as its corresponding constrained gestures were presented to the subjects. The subjects can either select the standard gesture shown in Q1 or select an alternative gesture.

Unpaired T-test with a statistically significant value of P=0.05 tested whether there was a significant difference in effort (represented by the Borg scale) among quadriplegic subjects. The effort reported by subjects with high-level C4 and C4/5 SCIs were significantly lower than subjects with Neurofibroma (P=0.004; P=0.017) and greater than the effort reported by the subject with a low-level C7 SCI (P=0.016; P=0.005) when performing gestures in the "*Wisee*" lexicon (Fig. 6).

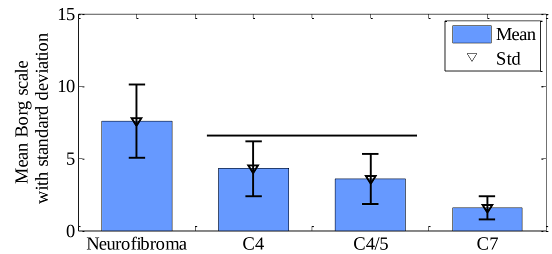

Fig. 6 Average Borg scale ranking, unpaired t-test, p<0.05

From the gesture selection results of Q2, 100% of the gestures selected by the subjects with C4 and C4/5 quadriplegia were from the constrained gestures generated by our approach. The stem graph (lower part) in Fig. 7 illustrates the index of constrained gestures selected by the subjects (see Fig. 5 for the gestures corresponding to the index). If there is no rectangle under the bar graph, it means that the standard gesture was selected rather than a constrained gesture (this occurred with the subject with C7 SCI). Even for the subject with C7 quadriplegia, who has more residual hand/arm functions than the other subjects, three out of seven constrained gestures were selected.

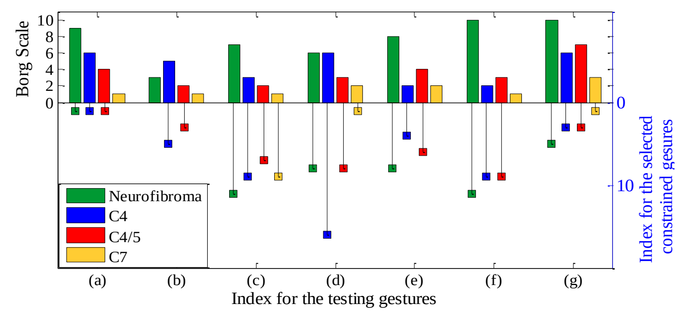

Fig. 7 Borg scale and gesture selection results

V. Conclusion & Future Work

An analytic method was proposed to address the problem of projecting standard gestures from a known manifold to an unknown constrained manifold that corresponds to the types of upper limb gestures that quadriplegics (due to middle to lower level (C4-C7) SCIs) are able to make. For each standard gesture in a set of lexicons, seventeen alternate constrained gestures with varied shape and curvature were generated using the pre-trained transform function (that we coined the *Laban Transform*).

A user-based validation test was conducted with four quadriplegic subjects with impaired upper extremity mobility to evaluate the usability of the constrained gestures. The results demonstrated that subjects reported larger effort when using a gesture from the standard group and thus preferred using a gesture from our generated alternatives. For subjects with higher level (C4 and C4/5) quadriplegia, each of the selected gestures came from the constrained gesture set. For the less paralyzed subject (C7 SCI), the alternative gestures were mostly preferred. These single subject assessments independently validated that the generated gestures were more usable and sufficient for individuals with quadriplegia to engage in widespread gesture recognition technologies.

Future work will: (1) expand the experiments to recruit more subjects with quadriplegia due to SCI and (2) further evaluate the optimal constrained gesture set by having subjects use the gestures during the performance of tasks, such as playing video games or robotic control.

References

[1] J. Shotton, T. Sharp, A. Kipman, A. Fitzgibbon, M. Finocchio, A. Blake, M. Cook, and R. Moore, "Real-time Human Pose Recognition in Parts from Single Depth Images," *Commun ACM*, vol. 56, no. 1, pp. 116–124, Jan. 2013.

[2] M. Jacob, C. Cange, R. Packer, and J. P. Wachs, "Intention, Context and Gesture Recognition for Sterile MRI Navigation in the Operating Room," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, L. Alvarez, M. Mejail, L. Gomez, and J. Jacobo, Eds. Springer Berlin Heidelberg, 2012, pp. 220–227.

[3] J. P. Wachs, M. Kölsch, H. Stern, and Y. Edan, "Vision-based Hand-gesture Applications," *Commun ACM*, vol. 54, no. 2, pp. 60–71, Feb. 2011.

[4] T. Ende, S. Haddadin, S. Parusel, T. Wusthoff, M. Hassenzahl, and A. Albu-Schaffer, "A human-centered approach to robot gesture based communication within collaborative working processes," in *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2011, pp. 3367–3374.

[5] J. Rico, A. Crossan, and S. Brewster, "Gesture-Based Interfaces: Practical Applications of Gestures in Real World Mobile Settings," in *Whole Body Interaction*, D. England, Ed. Springer London, 2011, pp. 173–186.

[6] T. C. Davies, T. Vinumon, L. Taylor, and J. Parsons, "Let's Kinect to Increase Balance and Coordination of Older People: Pilot Testing of a Balloon Catching Game," *Int. J. Virtual Worlds Hum. Comput. Interact.*, vol. 2, no. 1, pp. 37–46, 2014.

[7] K. Gerling, I. Livingston, L. Nacke, and R. Mandryk, "Full-body Motion-based Game Interaction for Older Adults," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2012, pp. 1873–1882.

[8] M. Pirovano, P. L. Lanzi, R. Mainetti, and N. A. Borghese, "IGER: A Game Engine Specifically Tailored to Rehabilitation," in *Games for Health*, B. Schouten, S. Fedtke, T. Bekker, M. Schijven, and A. Gekker, Eds. Springer Fachmedien Wiesbaden, 2013, pp. 85–98.

[9] G. Alankus, "Motion-Based Video Games for Stroke Rehabilitation with Reduced Compensatory Motions," *Electron. Theses Diss.*, Jan. 2011.

[10] W. Akram, L. Tiberii, and M. Betke, "A Customizable Camera-Based Human Computer Interaction System Allowing People with Disabilities Autonomous Hands-Free Navigation of Multiple Computing Tasks," in *Universal Access in Ambient Intelligence Environments*, C. Stephanidis and M. Pieper, Eds. Springer Berlin Heidelberg, 2007, pp. 28–42.

[11] A. Hyvärinen, P. O. Hoyer, and M. Inki, "Topographic Independent Component Analysis," *Neural Comput.*, vol. 13, no. 7, pp. 1527–1558, Jul. 2001.

[12] Y. Zhang, J. Zhang, and Y. Luo, "A novel intelligent wheelchair control system based on hand gesture recognition," in *2011 IEEE/ICME International Conference on Complex Medical Engineering (CME)*, 2011, pp. 334–339.

[13] R. Laban and F. C. Lawrence, *Effort*. Macdonald & Evans, 1947.

[14] D. Chi, M. Costa, L. Zhao, and N. Badler, "The EMOTE Model for Effort and Shape," in *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, New York, NY, USA, 2000, pp. 173–182.

[15] B. Webber, C. Phillips, and N. Badler, "Simulating Humans: Computer Graphics, Animation, and Control," *Cent. Hum. Model. Simul.*, Jun. 1993.

[16] J. Rett and J. Dias, "Laban Movement Analysis using a Bayesian model and perspective projections," *Brain Vis. AI*, vol. 4, no. 6, pp. 953–978, 2008.

[17] L. Santos and J. Dias, "Motion Patterns: Signal Interpretation towards the Laban Movement Analysis Semantics," in *Technological Innovation for Sustainability*, L. M. Camarinha-Matos, Ed. Springer Berlin Heidelberg, 2011, pp. 333–340.

[18] D. Swaminathan, H. Thornburg, J. Mumford, S. Rajko, J. James, T. Ingalls, E. Campana, G. Qian, P. Sampath, and B. Peng, "A Dynamic Bayesian Approach to Computational Laban Shape Quality Analysis," *Adv. Hum.-Comput. Interact.*, vol. 2009, May 2009.

[19] L. Zhao and N. Badler, "Synthesis and Acquisition of Laban Movement Analysis Qualitative Parameters for Communicative Gestures," *Tech. Rep. CIS*, Jan. 2001.

[20] S. Padam Priyal and P. K. Bora, "A robust static hand gesture recognition system using geometry based normalizations and Krawtchouk moments," *Pattern Recognit.*, vol. 46, no. 8, pp. 2202–2219, Aug. 2013.

[21] A. J. Quinn and B. B. Bederson, "Human Computation: A Survey and Taxonomy of a Growing Field," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, New York, NY, USA, 2011, pp. 1403–1412.

[22] C. J. C. Chan, M. A. M. Morada, M. R. Solamo, and R. Feria, "Measuring the Usability of a Low-Cost 3D Infrared Tracking and Wiimote-Based Interface," in *Theory and Practice of Computation*, S. Nishizaki, M. Numao, J. Caro, and M. T. Suarez, Eds. Springer Japan, 2012, pp. 90–100.

[23] C. Longstaff, C. Thelwell, S. C. Williams, M. M. C. G. Silva, L. Szabó, and K. Kolev, "The interplay between tissue plasminogen activator domains and fibrin structures in the regulation of fibrinolysis: kinetic and microscopic studies," *Blood*, vol. 117, no. 2, pp. 661–668, Jan. 2011.

[24] G. De'ath and K. E. Fabricius, "CLASSIFICATION AND REGRESSION TREES: A POWERFUL YET SIMPLE TECHNIQUE FOR ECOLOGICAL DATA ANALYSIS," *Ecology*, vol. 81, no. 11, pp. 3178–3192, Nov. 2000.

[25] S. Calinon, F. Guenter, and A. Billard, "On Learning, Representing, and Generalizing a Task in a Humanoid Robot," *IEEE Trans. Syst. Man Cybern. Part B Cybern.*, vol. 37, no. 2, pp. 286–298, 2007.

[26] A. P. DEMPSTER, "Maximum likelihood from incomplete data via the EM algorithm," *J R. Stat. Soc*, vol. 39, pp. 1–38, 1977.

[27] J. P. Wachs, H. Stern, and Y. Edan, "Cluster labeling and parameter estimation for the automated setup of a hand-gesture recognition system," *IEEE Trans. Syst. Man Cybern. Part Syst. Hum.*, vol. 35, no. 6, pp. 932–944, 2005.

[28] G. A. Borg, "Psychophysical bases of perceived exertion," *Med. Sci. Sports Exerc.*, vol. 14, no. 5, pp. 377–381, 1982.