Analysis of human performance using physiological data streams

Gaurav N. Pradhan
University of Texas at Dallas
P.O. Box 75083
Richardson, Texas 75083
gaurav@utdallas.edu

Balakrishnan Prabhakaran
University of Texas at Dallas
P.O. Box 75083
Richardson, Texas 75083
praba@utdallas.edu

ABSTRACT
Advancement in technology has led to measure the human performance using sophisticated multiple systems such as motion capture and physiological data monitoring systems. These systems together, represent the human activity in various physiologic and motoric streams that forms a multi-dimensional framework. The immediate requirement that rises is, analyzing these data streams to quantify the human performance. In this paper, we have proposed an efficient, multi-dimensional factor analysis technique that quantifies the multiple observations of data streams across different participants. In our approach, we extract characteristic parameters from the streams and conduct a separate global analysis on the data sets of each stream. The individual data sets are then projected onto the respective global analysis to analyze the differences in the responses of the participants. Next, we integrate these global analysis spaces of all streams, to get a *compromise* structure that represents the aggregate effect of all streams on the performance of each participant.

Categories and Subject Descriptors
H.2.8 [**Database Management**]: Database Applications; J.3 [**Computer Applications**]: Life and Medical Sciences

Keywords
Motion capture, electromyogram, multi-dimensional, factor analysis, principal component analysis.

1. INTRODUCTION
In the fields of medical, sports and training, human performance is a broad term that includes physical/muscular functions, body joint movements, as well as perceptual and cognitive abilities. Evidence from multiple systems involved in human performance may provide: 1) Crucial clues to guide identification, remediation, and ultimately prevention of a variety of medical conditions and behavioral deficits. 2) The necessary information to diagnose the problems in sports training that can improve the performance. With rapid advances in multiple fields (such as semiconductors, optics, and material sciences), the above fields uses a plethora of human performance measuring devices. Instruments such as 3D motion capture systems, EKG/ECG, EMG are increasingly being used for a variety of clinical studies and sports performance analysis. In this work, we focus on two systems,

- *3D Motion Capture :* It gives the positional and orientational information of all human joints in 3D space with a speed of 120 readings per joint per second.
- *Electromyogram (EMG) :* These are biomedical electrodes which gives information on the electrical activity generated during muscle contraction while doing motion.

On synchronizing these systems, we get the detailed information regarding the internal muscular contraction activity corresponding to the external joint movements. Together, it forms a multi-dimensional time series framework. The main challenge is identifying and extracting the set of attributes or characteristics from each multi-dimensional time series that can evaluate the performance aspect of each individual. Data streams coming from different sensors have different resolution/characteristics, and hence the extracted attributes for each data stream are different in nature. Our main objective is to analyze these heterogenous sets of attributes from multi-dimensional streams and to "model" the performance measure between individuals. This analysis will derive an integrated picture of the observations and relationships between the different data streams.

To achieve our objective, we explore two-stage multidimensional factor analysis approach. The first stage is intra-structure stage, wherein we analyze the structure of the individual multi-dimensional streams based on corresponding extracted attributes/parameters and derive a global analysis for each stream. In this global analysis, the differences between the participants' performance can be noticed with respect to each stream. The second stage is inter-structure stage, where we integrate the global analyses of all streams to get the inter-global analysis structure known as *compromise structure* that represents the aggregate effect of all sensors. It is useful to quantify the overall performance of the participants.

BodyNets 2008, March 13-15 Tempe, Arizona, USA

DOI 10.4108/ICST.BODYNETS2008.2937

2. RELATED WORK

Many multivariate analytical techniques have been developed to analyze several kinds of data. Some of the well-established techniques are like principal component analysis (PCA), canonical correlation analysis (CC), multiple factor analysis (MFA) [5], STATIS [6] etc. Most of these techniques on multivariate analysis are discussed in [1]. These techniques are mainly used in areas like sensory profiling [9], chemistry [10], food quality research [8], [7], [3] etc. As the data sets go more complex and multi-dimensional, these techniques need to be extended. The multiple factor analysis was extended to hierarchical MFA in [4], to handle hierarchical structure of data set. In [11], two multi-blocks tables were analyzed using STATIS and Tucker inter-battery method. In this paper, we are analyzing multi-dimensional data streams in the form of positional information of joints and muscular activity from multiple participants across multiple trials.

3. PARAMETER EXTRACTION

To quantify the participant's performance for a motion, we need to analyze the joints and muscles that are actively involved in motion. We call them ***prominent*** joints and muscles for a motion. *Example: For a motion "jump", the three leg joints namely, 'tibia', 'foot', and 'toe' and two leg muscles namely, 'tibialis anterior', and 'gastrocnemius' are prominent to consider for analysis. Depending on needs and requirement, other joints and muscles can be included in analysis.*

3.1 Motion Capture streams (joints)

To analyze the behavior of the prominent joint movements involved in the motion, we extract the crucial positional information at regular time intervals during the span of action. *In (Figure 1 (d)), we have chosen* $j = \{0(=L_{ON}), 25, 50, 75, 100(=L_{OFF})\}$. *Figures 1 (a), (b), and (c) shows the corresponding discrete points of 3D trajectories for every joint in 2D-space.*

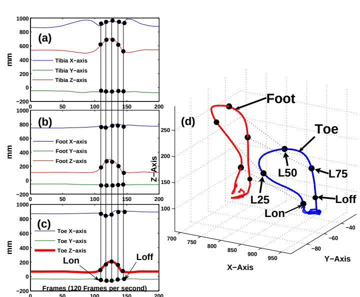

Figure 1: Extraction of features for each prominent human joints in motion (in this case, it is "jump").

3.2 EMG streams (muscles)

As EMG streams are non-stationary in nature, the previous discretization approach on joint streams is not effective. As a result, to analyze the prominent EMG streams, we need to extract the parameters that indicate the characteristics and temporal relationships between the consecutive peaks. *Figure 2 shows the general parameters that we extract from prominent EMG streams,*

- *The mean of all peak amplitudes* a_1, a_2, a_3, *and* a_4.

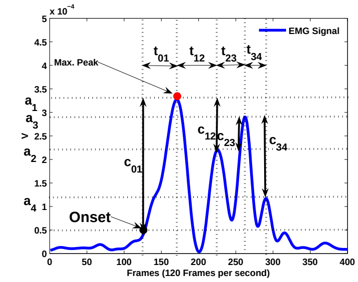

Figure 2: Extraction of features from the post-processed EMG signal.

- *The average duration between the consecutive peaks, i.e. mean of* t_{01}, t_{12}, t_{23} *and* t_{34}.
- *Average change in the consecutive peaks, i.e. mean of* c_{01}, c_{12}, c_{23} *and* c_{34}.
- *Number of peaks and maximum amplitude.*

Symbol	Explanation
n	Total number of participants $q_1, q_2, \cdots, q_n$
k	Total number of repetitive trials for similar motion
s	Total number of prominent joints and muscles i.e. sensors
$S_1, \cdots, S_s$	Label for s sensors
m_i	Total number of parameters associated with Sensor S_i
$P_i^1, \cdots, P_i^{m_i}$	Label for m_i parameters
$T_{i1}, \cdots, T_{ik}$	k trial tables for Sensor S_i

Table 1: Notations

4. MULTIDIMENSIONAL FACTOR ANALYSIS

First, we conduct separate global analysis on each stream in which we can find differences in the responses of the participants corresponding to each stream. The intra-structure stage is necessary because the behavior of each sensor and nature of its extracted parameters are different. In order to integrate them, we need to make these parameters comparable to each other so that we can quantify the overall performance of the individuals. The notations used further in this section, are summarized in Table 1.

4.1 Intra-structure Stage: Global analysis on individual stream

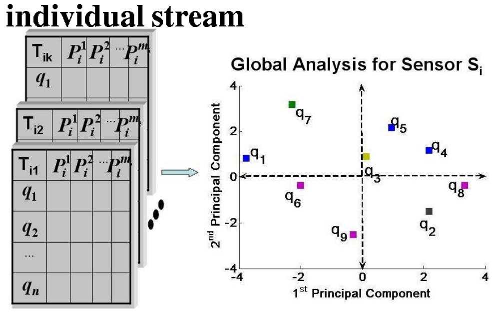

Figure 3: Construction of the global analysis space for the Sensor S_i**.**

To analyze the data-streams related to all s sensors, we form a multi-block structure for each sensor (as shown in left hand side of Figure 3). It organizes the parameters P_i^1, P_i^2, $\cdots$, $P_i^{m_i}$ of the sensor stream S_i ($\forall i : 1 \leq i \leq s$) for all n participants in all k trials. Thus, in one multi-block of S_i sensor, we have k tables/matrices ($n \times m_i$) each corresponding to one trial. The advantageous reason for this organization is that, in the global analysis space we can represent and evaluate the differences between the participants' responses/motion with respect to sensor S_i across k trials.

The following steps explain the approach for multidimensional factor analysis applied to multi-block structure corresponding to sensors S_i ($\forall i : (1 \leq i \leq s)$),

1. *Preprocessing:* The trial matrices $T_{i1}, \cdots, T_{ik}$ are centered, normalized and are denoted as X_{i1}, $\cdots$, X_{ik} respectively.

2. *Scalar Products:* Each matrix X_{ik} defines inherently a structure for the performance of the participants with respect to sensor S_i, which can be derived by computing the scalar products between participants. Hence, the preprocessed matrices X_{i1}, $\cdots$, X_{ik} are transformed into $n \times n$ scalar product matrices denoted as M_{i1}, $\cdots$, M_{ik} respectively. Thus, $\forall z : 1 \leq z \leq k$,

$$M_{iz} = X_{iz} \cdot X_{iz}^T \tag{1}$$

3. *Computing the Global Analysis Matrix:* The global analysis matrix for sensor S_i is given as follows,

$$M_{iC} = \sum_{x=1}^{k} \alpha_x \cdot M_{ix} \tag{2}$$

   where α_t denotes the weight for the t^{th} trial. The weights are chosen so that the trial structures agreeing most with other trial structures will have larger weights.

4. *Analyzing the Global Analysis Matrix:* The principal component analysis of the global analysis matrix M_{iC} explores the overall performance of the participant with respect to Sensor S_i. Since, global analysis matrix is also a scalar product matrix, its PCA is given as

$$M_{iC} = Q \wedge Q^T \tag{3}$$

   The factor scores (i.e. the projection of the rows on the principal components of the analysis of M_{iC}) are obtained as,

$$F_i = Q \wedge^{\frac{1}{2}} \tag{4}$$

   In this matrix F_i, each row corresponds to the participant and each column corresponds to the component. Figure 3 shows the example where few participants are represented in the global analysis space of first two principal components of the factor score matrix that carry total variance of $85 - 90\%$.

4.2 Inter-structure Stage: Inter-Global Analysis

In this stage, our goal is to integrate these global analysis spaces of all sensors, to get a *compromise* structure that represents the aggregate effect of all sensors on the performance of each participant. To make global analysis of all sensors comparable, we convert them into corresponding dimension-invariant, $n \times n$ *distance matrices*. That is, for each global analysis for sensor S_i, we get the corresponding distance matrix D_i for n participants. But distance matrices D_1, $\cdots$, D_s cannot be analyzed directly using eigen decomposition and need to be transformed into corresponding cross-product matrices $D_1^c, \cdots, D_s^c$ using *metric multidimensional scaling* [2]. On comparing and analyzing these s cross product matrices using steps 3 and 4 from Section 4.1, we find a final compromise inter-structure that quantifies the overall performance of the participants.

5. RESULTS

5.1 Test Environment and Data-sets

The experiments were conducted in 3D motion capture laboratory equipped with 16 high-resolution Vicon cameras capturing at the rate of 120 frames/second. EMG Ag-Cl surface electrodes provided by Delsys were synchronized with 3D motion cameras to pick the corresponding muscle activity of limbs while performing motions. The similar set of actions were captured from 24 participants (i.e. $n = 24$), and each participant performed 10 trials (i.e. $k = 10$) for each action.

In Sections 5.2 and 5.3, we will discuss the results of our approach by analyzing the leg segments for the "jump" action. The prominent joints and muscles for the "jump" activity are three leg segments, 'tibia', 'foot', and 'toe' and two leg muscles namely, 'tibialis anterior', and 'gastrocnemius'.

5.2 Global Analysis

The multi-variate structure to analyze the data-streams for "jump" response consists of 5 multi-blocks, each for prominent sensors. Figure 4(a) and (c) shows the projections of factor scores for all participants on the first two dimensions of global analysis spaces of tibialis anterior, and gastrocnemius respectively. For 'Tibialis Anterior', first two dimensions explains 89.4% of high variance, which is sufficient to interpret the results. The corresponding loadings i.e. correlations between the original parameters and the principal components (Figure 4(b)) shows that the first dimension has high correlation with 'mean of peaks', and 'maximum amplitude of peaks'. The first principal component, having 71.4% variance, differentiates the participants having high and low muscular activity. Similarly, with all other parameters related to 'Tibialis Anterior', we can interpret the behavior within the participants on same response. In global analysis for 'Gastrocnemius', first dimension explains 84.2% of variance. The corresponding loadings (Figure 4(d)) are almost similar to the previous muscle. Similarly we can also show the factor scores for joints projected on the first few dimensions of the corresponding global analysis.

5.3 Inter-Global Analysis

The PCA of the compromise matrix reveals the inter-structure of the participants by combining the effect of all prominent sensors. The projections of the participants on the first four dimensions are shown in Figure 5. Together, these four dimensions explain 83% of the variance of the compromise matrix.

In order to facilitate the interpretation of the Figure 5, we

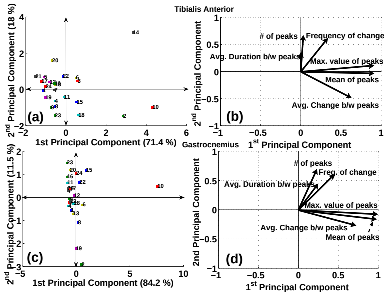

Figure 4: (a),(c): Global Analysis of tibialis anterior, and gastrocnemius; (b),(d): Respective loadings i.e. correlations between parameters and components.

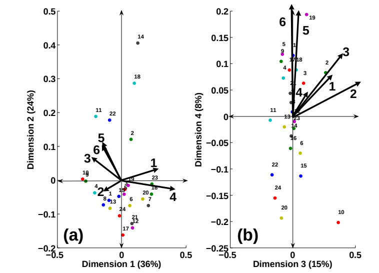

Figure 5: Analysis of compromise: Plot of the participants in the compromise space defined by dimensions 1-2(a) and 3-4(b) of the compromise matrix.

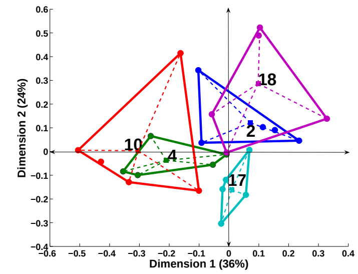

Figure 6: Compromise map of the participants with the performance of the sensors.

(1)	Height of the jump.(Recognized from the toe segment)
(2)	Time duration between onset and peak of the jump.
(3)	Total duration of jump.
(4)	Onset time difference between toe segment and EMG muscle 'tibialis anterior'.
(5)	Onset time difference between toe segment and EMG muscle 'gastrocnemius'.
(6)	Onset time difference between two EMG muscles 'tibialis anterior' and 'gastrocnemius'.

Table 2: Supplementary variables specific to the action "jump"

projected the supplementary variables (discussed in Table 2) specific to the "jump" action in compromise space. This was done by computing the loadings (i.e. correlations between these variables and factor scores) and then re-scaling these loadings by multiplying them by the square root of the eigenvalue associated with the dimension. The first dimension, which explains 36% of variance is highly correlated with variable 1 and 4. Similarly, variables 2 and 3 are highly correlated with dimension 3, and variables 6 and 5 are highly correlated with dimensions 4.

Figure 6 shows the first two principal components of the compromise space with the projections of some participants for each sensor. The position of the participant is the centroid for the corresponding prominent sensors. In order to facilitate the interpretation, we have drawn the convex hull for the projections of the sensors. The partial overlapping between convex hulls of participants 2 and 18 indicates that some sensors of the corresponding participants are in consensus. If there is no overlap with other participants for e.g. 17, then the performance of the sensors is clearly different from other participants. In this way, we can quantify the performance of each participant with others.

6. CONCLUSIONS

In this paper, we have proposed an efficient multidimensional factor analysis technique that quantifies the human performance. We have represented the human activity / responses using two advanced systems: 3D Motion capture to get positional information of joints and electromyography sensors to get information on muscle contractions. We extracted important parameters from the streams and conducted global analysis on each stream to find the differences in the responses of the participants with respect to that stream. We integrated these global analysis spaces of all streams, to get a *compromise* structure that represented the aggregate effect of all streams on the performance of each participant. The advantage of our approach was to provide the compromise space that gives valuable information on the prominent sensors and on the consensus between the sensors when projected in the compromise space. Hence, our work builds a robust platform for integrating, evaluating, and quantifying the evidences from the multiple systems involved to measure human performance.

7. REFERENCES

[1] H. Abdi. *Encyclopedia for research methods for the social sciences*, chapter Multivariate analysis, pages 699–702. SAGE, Thousand Oaks (CA), 2003.

[2] H. Abdi. *Encyclopedia for research methods for the social sciences*, chapter Metric multidimensional scaling, pages 598–605. SAGE, Thousand Oaks (CA), 2007.

[3] C. Chaya, C. Perez-Hugalde, L. Judez, C. Wee, and J. Guinard. Use of the STATIS method to analyze time-intensity profiling data. *Food quality and preference*, 15:3–12, 2003.

[4] S. L. Dien and J. Pagés. Hierarchical multiple factor analysis: application to the comparison of sensory profiles. *Food quality and preference*, 14:397–403, 2003.

[5] B. Escofier and J. Pagés. Multiple factor analysis. *Computational Statistics and Data Analysis*, 18:121–140, 1990.

[6] C. Lavit, Y. Escoufier, R. Sabatier, and P. Traissac. The ACT (STATIS method). *Computational Statistics and Data Analysis*, 18(1):97–119, 1994.

[7] M. Meyners, J. Kunert, and E. Qannari. Comparing generalized procrustes analysis and statis. *Food quality and preference*, 11:77–83, 2000.

[8] E. Qannari, I. Wakeling, and H. MacFie. A hierarchy of models for analysing sensory data. *Food quality and preference*, 6:309–314, 1995.

[9] P. Schlich. Defining and validating assessor compromises about product distances and attribute correlations. *Multivariate Analysis of Data in Sensory Science*, 6:229–306, 1996.

[10] I. Stanimirova, B. Walczak, D. Massart, V. Simeonov, C. Saby, and E. D. Crescenzo. STATIS, a three-way method for data analysis. application to environmental data. *Chemometrics and intelligent laboratory systems*, 73:219–233, 2004.

[11] M. Vivien and R. Sabatier. A generalization of STATIS-ACT strategy: DO-ACT for two multiblocks tables. *Computational Statistics and Data Analysis*, 46:155–171, 2004.