BIOMECHANICAL DATA CHARACTERIZING FOR MULTIFACTOR AND MULTIVARIATE ANALYSIS

P. Loslever & J.J. Flahaut
Laboratoire d'Automatique et de Mecanique Industrielles et Humaines, Universite de Valenciennes

INTRODUCTION

Empirical studies in sport biomechanics are mainly organized in experimental designs which combine factor modalities. The a0 individuals are the modalities of the first factor (FO), and the other controlled factors are F1, F2,... with a1, a2,... modalities, respectively. Thus, the data structure looks like a hyperparallelepiped (HP) with a0*a1*a2*... modality factor combinations (MFCs) or cells, Fig. 1. All the MFCs can be tested or not (Latin square, for instance). For each MFC or cell c, the same variables are considered: V1, V2,... These are generally time varying variables describing movements, forces, pressures...From such multifactor and multivariate experimental designs, the main aim is the investigation of the notion of dependence i.e. the influence of factors onto variable and connections between variables. Nevertheless, these two statistical aspects are seldom considered both at the same tie. For instance, the analysis of variance exists (Johnson & Wichern, 1992). Whatever the statistical method used to get results from an experimental design, there is an essential stage: the way to go from empirical raw data to data that are compatible with the statistical method i.e. the data characterizing stge. This paper deals with such a stage but in the prospect that monodimensional or multidimensional statistical approaches can be used. *An* example about rock climbing is considered.

METHODOLOGY

The data characterizing stage means a data reduction but does not involve a structural simplification: the phenomena included in each cell c of the HP are described without eliminating interesting information but the input and output of the characterizing method are Hps. Variables V1, V2,... can be characterized into specific and new variables according to two points of view: the space coding level and the time integration level.

Space coding level: The lowest coding level consists in keeping the scale of each variable (or to merely use a translation and multiplication of the scale) and the highest level consists in considering new variables that underscore the variable semantics. For instance: *variable V1 is low and variable V2 is high...* or *patterns* in the signals. In that case, numeric data are changed into symbolic data, that needs 1) to consider modalities in each variable scale (binary or fuzzy membership patterns) and 2) to associate these modalities.

Time integration level, Fig 1.: The lowest level consists in considering a chronology of time windows (the signal is more or less smoothed) and the highest in summarizing the time samples. Between these two extreme levels, there are intermediate ones such as the magnitude distribution, the transition matrix or the power spectrum.

Combining the levels of space coding and time integration, Fig. 2: Let us consider two levels for coding a quantitative scale i.e. the raw scale and the scale modalities. Four combinations can be considered: with the highest time integration level, the time average (for instance, arithmetic mean) and magnitude distribution; with the lowest integration level, the chronology of raw data or space modality membership values.

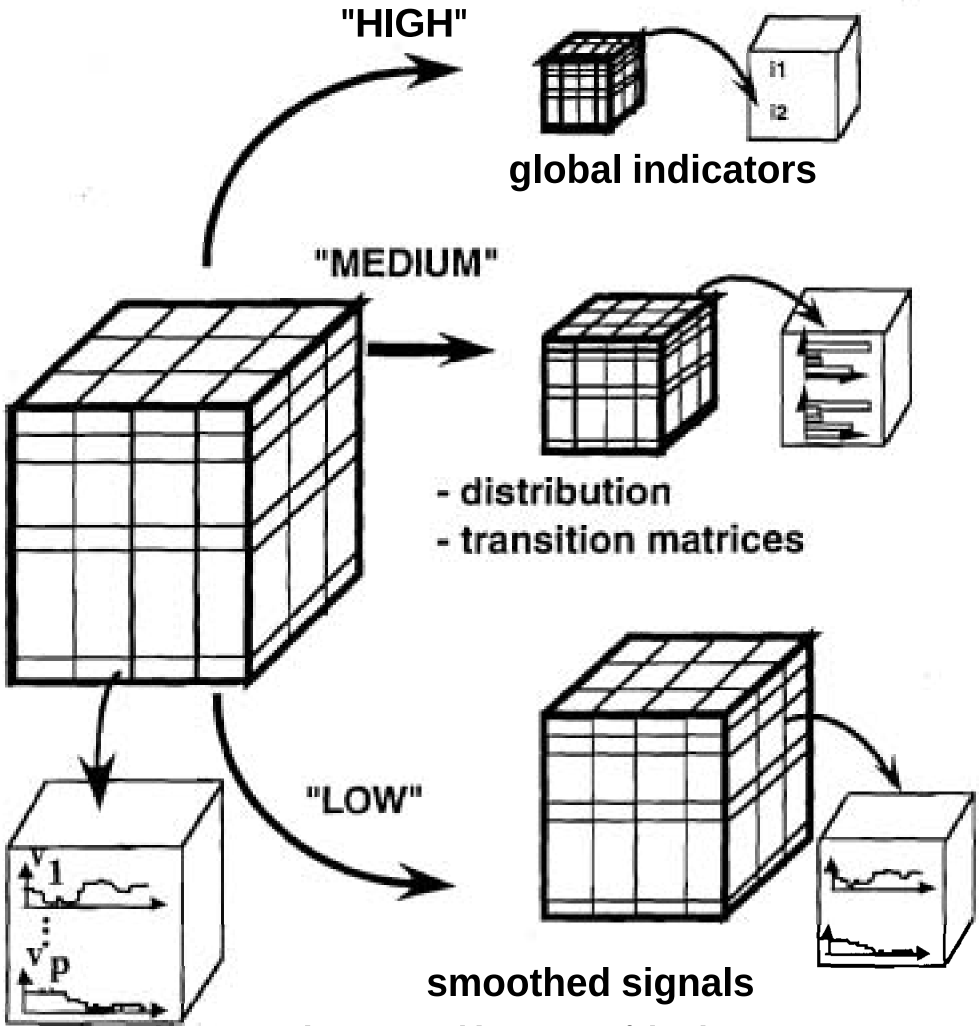

Fig. 1 : The input and some possible outputs of the characterizing stage.

Characterizing assessment: It is obvious that 1) many levels of space coding and time integration can be considered and 2) many new variables combing space and time aspects can be built. Thus it is necessary to assess the data characterization performance. Three main propositions can be formulated:

1) To represent a semantic notion within the signals through one or several variables it si necessary to check if they can summarize this notion. Example: Does the average behaviour characterized through the arithmetic mean have a meaning?

2) To consider the data reduction level DR which is the ratio between the sizes of the output data set and the input data set. Example: Let us consider a signal contain 1000 times samples. DR=1/1000 with the arithmetic means, DR=10/1000 with the 5 intervals magnitude histogram, DR=5*5/1000 with the transition matrix.

3) To check if the output variables of the characterizing stage are compatible with the statistical technique that will be used. Example: Does a variable fulfill the Gaussian model in the perspective of the variance analysis application?

The characterizing problem is illustrated with a generic example, rock climbing, which is considered because of its high levels in both physical and mental aspect.

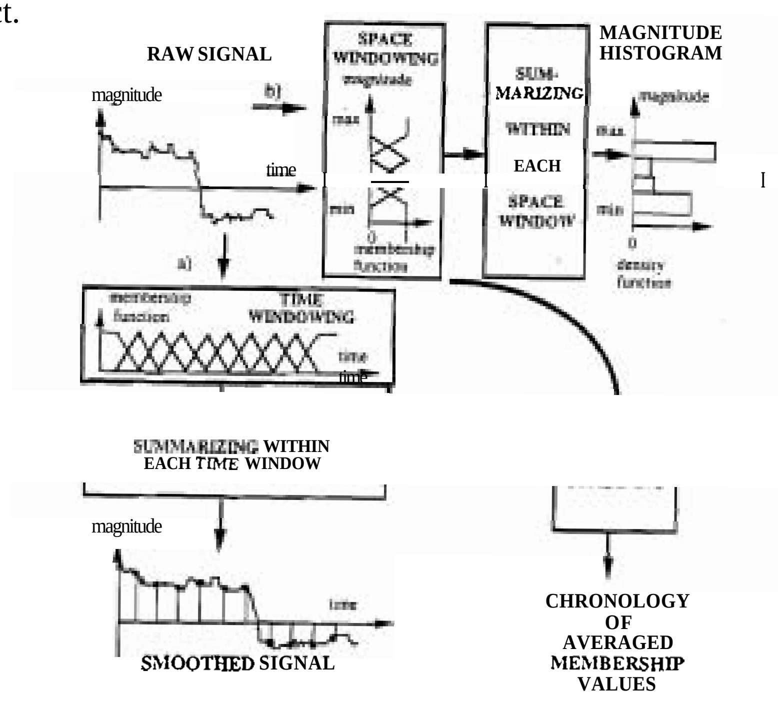

Fig. 2 : Examples of signal characterizing :
a) time windowing, b) space windowtng, c) combining space and time windowing.

Fig. 2: Examples of signal characterizing: a) time windowing, b) space windowing, c) combining space and time windowing.

EXAMPLE: CLIMBING

The data structure is a parallelepiped where the three directions tally with the a) individuals, the a1 expertise levels, and a2 climbing situations. Variables are vertical (V) and horizontal (H) positions of left and right hands and feet (the segment number is s=1to4) and gravity center (s=5). The characterizing stage aims at describing how the climbing activity i.e. describing the signals X(c,s,t), and Y(c,s,t), where c is a cell of the parallelepiped (c=1...nc), s is a segment and t the time. There are many ways to achieve this aim. Three of them are considered here in.

Method 1: Entropy Hc of the gravity center trajectory Y=f(X) (for s=5) (Cordier, 1992), according to (Stewart 90). This indicator involves no space coding and an integration over the time. The statistical analysis shows that H decreases from the first to the last trial and is larger with experts than with beginners (Dupuy, Ripoll, & Flahaut, 1992).

Method 2: **Distribution of motor actions Dc** where the actions are: static phase (crabbing, magnesia taking, segment displacement, equilibrium reaching) and dynamic phases (body motion). The space variable becomes qualitative and there is an integration over the time. Statistical analysis shows that the experts optimize the number of movements for each of the a2 climbing situations and the beginners don't achieve this optimization even in the last situation (Dupuy, Ripoll, & Flahaut, 1992).

Method 3: Inter-limb transition matrix Tc. Transitions between two successive stable postures p and p+1 are considered, a stable posture p being obtained when the four limbs are a four grabbing posture (p=1,..., npc). The generic term of the matrix, Tc(s,s'), contains the number of times a segment s' moves after a segment s between two postures p and p+1. This indicator represents a qualitative variable (with 4x4=16 categories) and is obtained when integrating over the npc postures. Nevertheless the notion of chronology is partly kept. The statistical analysis, shows that the experts prefer diagonal transitions (hand to foot) and beginners lateral transitions (hand to hand or foot to foot) (Flahaut, Loslever, 1995).

DISCUSSION

Many paths can be used to reach 'latent' results from empirical and raw data. 'Latent' means that interesting results can exist but they are found only if the appropriate characterizing method (CM) and statistical analysis method (SM) are used. That is why it is necessary to test several Cms and Sms. For each of these tow stages, we believe that 1) the problem must be states, 2) some ways to solve it are to be proposed, 3) one or two ways must be chosen and tested according to specific criteria. This three-point approach must be put against assertions such as: *"the global indicators within each cell of the HP are..."*, and *"the analysis of variattce shows that..."*. With this "simplistic" approach, questions such as *"why summarizing data*

in this way" and ***"why using the analysis of variance"*** emerge.

REFERENCES

Dupuy, C., Ripoll, H., & Flahaut, J.J. (1992). Organisation spatiotemporelle de la motricite en escalade sportive. Recherche en A.P.S., ***3,*** 233-253.

Flahaut, J.J. & Loslever, P. (1995). Quantification des associations intersegmentaires en escalade sportive et comparaison entre grimpeurs "experts" et "non experts". Note interne LAMIH mars 1995.

Johnson, A.R. & Wichern, D.W. (1992). Applied multivariate statistical analysis. Prentice-Hall Ed.

Stewart, I. (1990). La physique des courbes. Pour la science, 155, 106-117.