

SOFTWARE PACKAGE FOR BIOMECHANICAL DATA ANALYSIS

D'AMICO, M.; FERRIGNO, G.
Centro di Bioingegneria - Politecnico di Milano
Milano
Itali

In human movement studies, in order to determine the way in which the various muscles cooperate to generate the motion, the estimates of the inertial and gravitational force contributions are required. The first step toward a quantitative description of a phenomenon is always represented by the use of an adequate model. The design of such a model necessitates both the assessment of the mass distribution of the body segments and the estimation of the translational and rotational displacement, velocities and accelerations of these latter. Also the external forces involved in the motion must be known.

Once the aforementioned variables are assessed and the body model has been designed, the solution of the indirect dynamic problem is possible and other variables, such as load distribution within musculoskeletal components, joint moments, energy, power and so on, can be computed.

The multifactorial analysis of the time course of all these variables can lead to motor control interpretation and neural activity investigation. In order to carry out such analysis, several body landmarks positions, according to the model, are measured at discrete time instants. The measuring process, whatever the means used, is quite error prone. Because of these errors, the biomechanist tackles everyday the problem of a proper signal processing. Since data are acquired, in fact, the recorded signal must be processed in some way in order to obtain correct information about the event to be analysed (i.e. the space calibration performed to obtain the parameters to be used for space intersection from the recorded 2D images). Moreover signal processing is mandatory to compute further variables from a data set (i.e. the computing of velocities and accelerations of displacement or angular data). Finally, more information may be extracted from the data by a proper processing such as for example the Power Spectrum Density (PSD) estimate describing the pulsation course of signal power and which can be used to determine the frequency content of the signal.

The aim of this paper is to present a personal computer program based on an original algorithm which allows to obtain from a set of data, in a single run, the following:

- the PSD of the signal via autoregressive (AR) modelling
- the frequency which bounds the signal to a given power or to a given signal to noise ratio (SNR).
- the filtered signal
- the interpolated signal (to whatever sampling rate different from the original)
- the integral and derivatives (first and second order).

All these results are obtained without operator intervention, i.e. the procedure is completely automatic. Particular attention has been paid to the processing speed in order to minimize the time spent for the computations, feature very important in routine analyses.

METHOD

As previously introduced, the instrumentation used to collect data introduces some amount of error corrupting the useful signal. This fact obliges to treat the data in order to remove, as possible, the noise superimposed on the signal in order to perform subsequent processing. Some critical operations, such as differentiation, magnifies, in fact, noise in this case, the measurement error can be amplified to such an extent which floods the useful signal. More details about this problem can be found in Wood (1982), Woltring (1985), D'Amico and Ferrigno (1988, 1990).

The sources of errors can be various (Wood, 1982, Lanshammar, 1982), but in general can be subdivided in systematic ones (introduced by image distortions or marker ill-positioning as happens at knee joint level) and random ones (mainly due to quantisation noise). The latter are the more insidious because of their wideband spectrum, particularly dangerous for the computation of derivatives.

Several methods to solve the problem of noise magnification in processing biomechanical data have been described in literature, relying on time and frequency domains approach. In D'Amico and Ferrigno, 1988 an

original technique characterised by a frequency domain approach was proposed. A new and extended version of this was later described in D'Amico and Ferrigno, 1990.

The idea underlying the proposed new technique is to perform the noise suppression by using a batch-adaptive filter. The cut-off frequency of this latter is set by assessing the PSD of the measurement (signal plus noise) via AR modelling.

Figure 1 shows the algorithm block scheme describing the phases through which the data are processed and the various features are obtained. This algorithm represents an extended version of the procedure presented in D'Amico and Ferrigno, 1990. The integral computation has been, in fact, added, as well as the time domain filter windowing has been modified. The interpolation has been also explored by using a continuous inverse Fourier Transformation.

By examining in more detail the figure, the first step of the procedure is to fit an AR model to the measurement in order to perform either the PSD estimate, which leads to the bandwidth selection, or the signal prediction (extrapolation) to avoid edge effects. After these preliminary operations the data are Fourier Transformed by the FFT algorithm. Successively, a low-pass linear-phase finite impulse response (FIR) filter is designed by windowing in time domain the impulse response of the ideal low-pass (Oppenheim and Schaffer, 1975). The filtering is performed in frequency domain by multiplying the Fourier Transform of the impulse response by the FFT of the preprocessed data. The result of this multiplication gives the filtered data by inverse fast Fourier Transformation (IFFT). The previous multiplication of the FFT by $1/j\omega$, $j\omega$ and $-\omega^2$ leads to the integral, the first and second derivatives.

Hereafter we shall briefly describe each step in detail

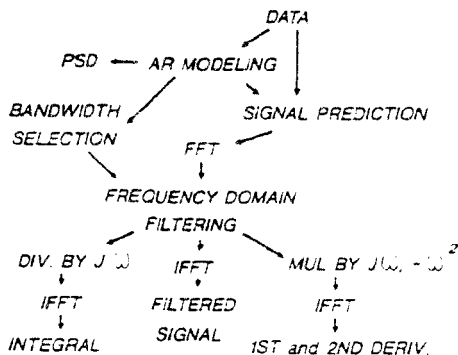


Figure 1: Algorithm Block Scheme

AR MODELLING

Very often to better understand the characteristics underlying a physical process it is useful to introduce the spectral analysis of the data representing the phenomenon under study.

There are several methods to obtain a PSD estimate and they can be divided in two large classes: Classical Spectral Estimation and Parametric Spectral Estimation (see Kay and Marple, 1981, for a review). The first one is basically based on Fourier Transform either of the data (Periodogram method) or on the correlation function (Correlation method). However this approach presents several problems related to the poor frequency resolution for short data records (as in the case of sports movements), and the undesirable effect of leakage due to tacit time domain windowing which distorts the PSD. We remember that in this case the frequency resolution is approximately the reciprocal of the observation interval depending on the main lobe of the window transform, while the leakage phenomenon (i.e. the signal power spread into adjacent frequency region) is due to the

sidelobes.

The second method relies on signal modelling. Three important classes of such models are known as AR, moving average (MA) and autoregressive moving average (ARMA). Each class is suitable for different kind of signals, but AR models show better performances from a computational point of view. In fact, differently from the other models, they require only the solution of linear equations to obtain the parameters estimate. Moreover the Wold decomposition theorem relates (Kay and Marple, 1981) these three kinds of model asserting that AR model can be used instead of ARMA and MA model at the cost of an increased order. For all these reasons our choice fell in AR field and, among the various available techniques for parameters estimation, we used the least square version of the forward-backward linear prediction, also called modified covariance algorithms (Ulrych and Clayton, 1976, Nuttall, 1976). This technique allows to obtain a very sharp PSD estimate, without showing the problem of spectral line splitting (Kay and Marple, 1981), but the computational costs are greater than for other methods. For this reason we optimised requirements taking advantage from the symmetries (real signals) and from the relationships between the elements of the matrix of the linear system to be solved.

The use of a parametric technique implies the problem of the model order choice, i.e. the definition of the proper number of parameters for the model necessary to obtain a reliable PSD estimate. Too few parameters lead to a low detailed PSD estimate, while an overabundant order introduces spurious peaks. In D'Amico and Ferrigno, 1990, a discussion on this topic has been presented, however for the aim of this paper, it is sufficient to report that, after having analysed many biomechanical displacement data, we set the order at 15. In this way we tested the possibility to perform reliable PSD estimates for a broad class signals.

Once the PSD of the signal has been obtained, not only it is used for the further steps of the algorithm, but it is useful to analyse the phenomenon under study from a frequency domain point of view. An example of such an utilisation is reported in D'Amico et al., 1989, 1990.

BANDWIDTH SELECTION

The bandwidth selection procedure, sets the cut-off frequency F_u of the low-pass filter. The selection is automatically accomplished. The noise (supposed white) power is initially estimated by the PSD. The power is computed as the average value of the PSD in high frequency regions, namely between $0.8 F_s/2$ and $F_s/2$ where $F_s/2$ is the Nyquist frequency (F_s is the sampling frequency). In this region the contribution of the signal to the PSD is negligible with respect to the noise. This latter assumption is not a constraint if the error formula presented by Lanshamer, 1982 is considered. According to this relation, a sampling rate higher than twice the frequency content of the signal must be chosen in order to avoid dramatic effects on the derivatives. Once the noise power (N) estimate is known, the frequency at which the PSD falls below $50 N$ is found and taken as cut-off frequency.

SIGNAL PREDICTION

In order to avoid the edge distortions due to the periodic continuation implicit in the computation of the FFT, an extrapolation of the data beyond the record ends is performed. For this purpose, the parameters of the AR model are used. In order to optimise the procedure, the forward and backward parameters are separately used for the data after and before the record respectively. The length of the extrapolation depends on the selected filter bandwidth. It equals the number of samples of the impulse response of the filter in order to assure that all the data in the record are not distorted by edge effect due to filter loading.

The prediction is performed according to the standard formulas of AR models:

$$\sum_{k=1}^p s_t = -\sum_{k=1}^p a_k^f s_{t-k} + e_t^f \quad (1a)$$

$$\sum_{k=1}^p s_t = -\sum_{k=1}^p a_k^b s_{t+k} + e_t^b \quad (1b)$$

where a_k^f , a_k^b are the model parameters (forward and backward respectively), s is the signal and e^f and e^b are known as modelling errors. The standard deviation of these latter is estimated by applying the eqns. 1 to the

available data and computing the residual error.

FILTERING

The impulse response of an ideal low pass filter set at P_u is expressed by the function $\sin(ka)/k$, where k is the sample (kT with $T=1/P_s$) and a is the normalised P_u pulsation ($2 P_u/P_s$). The zeroes of this function lie at $k = \pm\pi/a, \pm 2\pi/a, \dots, \pm n\pi/a$. The number of samples bounded between the first two (positive and negative) zeroes are P_s/P_u . We have chosen to use for the filter a number of samples multiple of this ratio. In this way it is possible to use the desired number of lobes of the impulse response. This number has been defined as sharpening factor (PS) because, the higher it is, the sharper is the transition band behaviour. SP equal one is indicated for rapid transient analysis. In this case, in fact, only the (all positive) main lobe is used, avoiding the arising of oscillations or ringing on the second derivatives. For standard biomechanical data an SP of 3 has shown good results.

Once the impulse response of the ideal low pass filter has been designed, a windowing to the desired finite number of samples is required. In a previous work (D'Aiico and Ferrigno, 1990) a Hamming window was used. A better behaviour of the derivatives has been obtained by using a new window expressly designed. In order to have a sharp transition band behaviour (given SP) it is necessary to have a narrow frequency domain main lobe (Hamming window satisfies this condition), and in order to avoid problems of noise magnification at high frequencies it is desirable to have sidelobes continuously rolling off with a sufficient slope (this is not the case of Hamming window which presents constant value sidelobes). These characteristics have been obtained designing a polynomial window as follows:

$$w(t) = (15t^4 - 8t^2 + 1)^L \quad (2)$$

L trims the width of the main lobe and the roll off slope of the sidelobes. For our application L has been fixed at 2. Figure 2 shows the behaviour of two windows designed with eqn. 2, the Hamming window and a Kaiser window with the parameter equal to 3.5.

Once the design of the filter has been completed, it is transformed in the frequency domain and multiplied by the extended signal FFT.

INTEGRAL AND DERIVATIVES COMPUTATION

Integral and derivatives are computed in frequency domain by using the properties of the Fourier transform as described above. IFFT completes the algorithm.

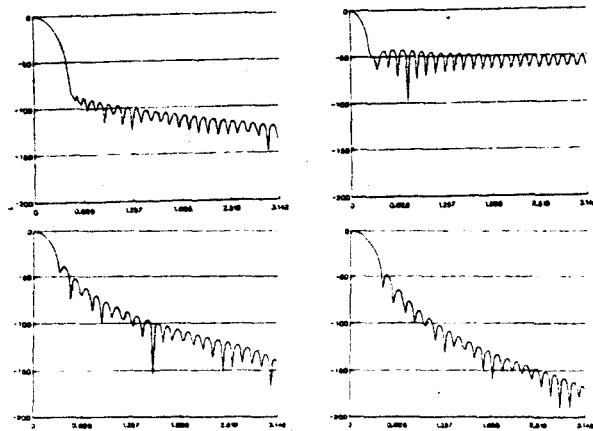


Figure 2: The upper two panels show the frequency domain transform of the polynomial window (eqn. 2) for $L=2$ and $L=3$ respectively. The lower left panel shows the Kaiser window ($p=3.5$) and the lower right the Hamming window. All the window transforms have been computed from 51 samples in time domain

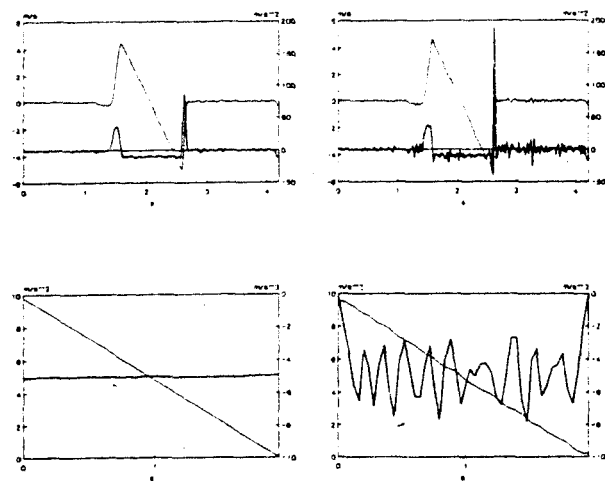


Figure 3: The upper panels show the first (thin line) and second (thick line) derivatives of the vertical displacement of a free falling ball including the throwing phase obtained with our procedure (left) and with GCVQS (right). The lower panels report second (thin line) and third (thick line, true value=-5) derivatives of a synthetic data set

EXAMPLE OF RESULTS

Figure 3 shows a comparison between the results obtained with the procedure herein described and by the generalised cross validatory quintic splines (GCVQS) (Woltring, 1986), believed to be the best smoothing and differentiating technique (Murphy and Mann, 1987). The upper panels refer to a free falling ball sampled at 50 Hz with the ELITE system. The second derivative obtained with our algorithm are more clean and do not show ringing near the peak. The lower panels show the results obtained from simulated data reported by Woltring, 1985. In this case the difference is dramatic on the third derivative. Moreover the time spent by the GCVQS for the processing was of 110 and 27 seconds for the ball and the simulated data respectively, while our algorithm spent 3 and 2.5 seconds respectively.

REFERENCES:

- D'AMICO, M. and FERRIGNO, G. \1988\ Estimation of Velocities and Accelerations from noisy Kinematic Data. In proc. VI Int. Symp. ISBS, in press
- D'AMICO, M. and FERRIGNO, G. \1990\ A Technique for Evaluation of Derivatives from Noisy Biomechanical Displacement Data by a Model-Based Bandwidth-Selection Procedure. In press on Medical & Biol. Eng & Comput.
- D'AMICO, M., FERRIGNO, G. and RODANO, R. \1989\ Frequency Content of Different Track and Field Basic Movements. In Proc. of VII Int. Symp. ISBS, ed. by W.E. Morrison, Footscray Institute of Technology, Victoria, Australia, 3011, 176-193
- KAY, S.W. and MARPLE, S.L. \1981\ Spectrum Analysis - A Modern Perspective. Proc. of IEEE, 69, 1380-1419
- LANSHAMMAR, H. \1992\ On Precision Limits for Derivatives Numerically Calculated from Noisy Data. J. Biomech., 15, 459-470
- MURPHY, M.C. and MANN, R.W. \1987\ A Comparison of Smoothing and Digital Filtering\Differentiation of Kinematic Data. IEEE IX Ann. Conf. of Eng. in Med. and Biol. Soc., 852-853
- NUTTALL, A.H. \1976\ Spectral Analysis of a Univariate Process with Bad Data Points, via Maximum Entropy, and Linear Predictive Techniques. Naval Underwater System Centre, Tech. Rep. 5303, New London, CT
- OPPENHEIM, A.V. and SCHAFER, R.W. \1975\ Digital Signal Processing. Prentice-Hall Inc., Englewood Cliffs, N.J.
- ULRICH, T.J. and CLAYTON, R.W. \1976\ Time Series Modelling and Maximum Entropy. Phys. Earth Planetary Interiors, 12, 188-200
- WOLTRING, H.J. \1985\ On Optimal Smoothing and Derivative Estimation from Noisy Displacement Data. Human Movement Science, 4, 229-245
- WOLTRING, H.J. \1986\ A PORTMAN Package for Generalised Cross- Validatory Spline Smoothing and Differentiation. Adv. Eng. Software, 8, 104-113
- WOOD, G.A. \1982\ Data Smoothing and Differentiation Procedures in Biomechanics. Exercise and Sport Science Rev., 10, 308-362