PII: S0021-9290(97)00042-0

J. Biomechanics, Vol. 30, No. 8, pp. 847–850, 1997

Printed in Great Britain
0021-9290/97 $17.00 + .00

ELSEVIER

TECHNICAL NOTE

A COMPARISON OF AUTOMATIC FILTERING TECHNIQUES APPLIED TO BIOMECHANICAL WALKING DATA

Giannis Giakas* and Vasilios Baltzopoulos†

* Division of Sport, Health and Exercise, Staffordshire University, Stoke-on-Trent ST4 2DF, U.K.; and
† Department of Exercise and Sport Science, Manchester Metropolitan University, Alsager ST7 2HL, U.K.

Abstract—The purpose of this study was to compare and evaluate six automatic filtering techniques commonly used in biomechanics for filtering gait analysis kinematic signals namely: (1) power spectrum (signal-to-noise ratio) assessment; (2) generalised cross validation spline; (3) least-squares cubic splines; (4) regularisation of Fourier series; (5) regression model and (6) residual analysis. A battery of 1440 signals representing the displacements of seven markers attached upon the surface of the right lower limbs and one marker attached upon the surface of the sacrum during walking were used; their original signal and added noise characteristics were known *a priori*. The signals were filtered with every technique and the root mean square error between the filtered and reference signal was calculated for each derivative domain. Results indicated that among the investigated techniques there is not one that performs best in all the cases studied. Generally, the techniques of power spectrum estimation, least-squares cubic splines and generalised cross validation produced the most acceptable results. © 1997 Elsevier Science Ltd

Keywords: Smoothing; Optimal; Gait analysis; Differentiation.

INTRODUCTION

Automatic filtering techniques estimate the optimal amount of filtering, based usually on statistical or power spectrum information calculated from the recorded signal. Despite the general tendency towards the use of automatic filtering methods (Hatze, 1990), there are only a few studies that have evaluated and compared these methods (Burkholder and Lieber, 1996; DAmico and Ferrigno, 1992). However, these studies compared only three methods each using a limited number of signals, and none has yet focused on gait analysis.

Automatic filtering methods cannot be applied universally. Every method is based on a number of assumptions that have to be satisfied by the data for the filtering to be performed appropriately. The fact that a certain method produces acceptable results when a specific signal is examined does not indicate that this method will also be adequate for other signals with different original signal or noise characteristics.

The purpose of the present study, was to compare and evaluate all of the most commonly used automatic filtering techniques in biomechanics, with walking input signals with a wide range of noise levels. The present study aims to establish the most appropriate automatic filtering techniques for gait analysis purposes.

METHODS

Six automatic filtering methods were examined and evaluated:

1. Power spectrum assessment (PSA). The algorithm for the selection of the optimal cut-off frequency and for signal extrapolation was based on the method by DAmico and Ferrigno (1990). Filtering was then performed using a recursive second-order Butterworth digital filter (Vaughan, 1982) applied in both forward and backward directions (zero phase-lag).

2. Generalised cross validation spline (GCV) using dynamic programming (Dohrmann *et al.*, 1988). The algorithm was provided by the author of the above paper. This algorithm was preferred because it does not require zero accelerations at the end points and because it performs the computations faster than the one presented by Woltring (1986).

3. Least-squares cubic splines (LSCS) (Simons and Yang, 1991). The software was provided and used with the default settings as suggested by the author.

4. Regularised Fourier series (RFS). The software was based on the algorithm by Anderssen and Bloomfield (Anderssen and Bloomfield, 1974; Cullum, 1971; Wood, 1996). The optimal regularisation was performed in each derivative domain which effectively eliminated the process of numerical differentiation and all its associated problems (Hatze, 1979; Hatze, 1981).

5. Regression model (RM) (Yu, 1989). The optimal cut-off frequencies were calculated by using the two equations (RM_1 and RM_2, respectively) provided by Yu (1989, 1996). The first model (RM_1) is a non-linear model that predicts the cut-off frequency from the amplitude of the first 'overlap' harmonic and the sampling frequency; the second (RM_2) uses only the sampling frequency as predicted also from a non-linear model. A reverse mirror extrapolation (Smith, 1989) with half the data length at each end was applied, and filtering was performed as for PSA.

6. Residual analysis (RA) (Winter, 1990). The optimal cut-off frequency was calculated following guidelines in Winter (1990) and recommendations in *biomch-l* (31 May 1996). Extrapolation and filtering were performed as for RM and PSA, respectively.

The higher derivatives of the filtered signals using the PSA, $RM_{1,2}$ and RA techniques were calculated using first-order finite differences since filtering was performed using the Butterworth filter. This technique has no smoothing effects (Miller and Nelson, 1973). The higher derivatives of the other techniques were computed by differentiating the resulting polynomials (GCV and LSCS).

Twenty-four signals (3 × 8) simulating the X, Y and Z displacements of eight markers attached to the lower segments were used as reference signals ($N = 48$). Two types of noise, with 30 levels each, were generated: one in the time domain ('random' noise) and the second in the frequency domain ('white' noise). Because of the small number of data points in each signal the two types of noise had slightly different characteristics. The 60 different combinations of noise type and level were superimposed onto the 24 reference walking signals giving a total

Received in final form 1 *April* 1997.

Address correspondence to: Giannis Giakas, Division of Sport, Health and Exercise, Staffordshire University, Stoke-on-Trent ST4 2DF, U.K.

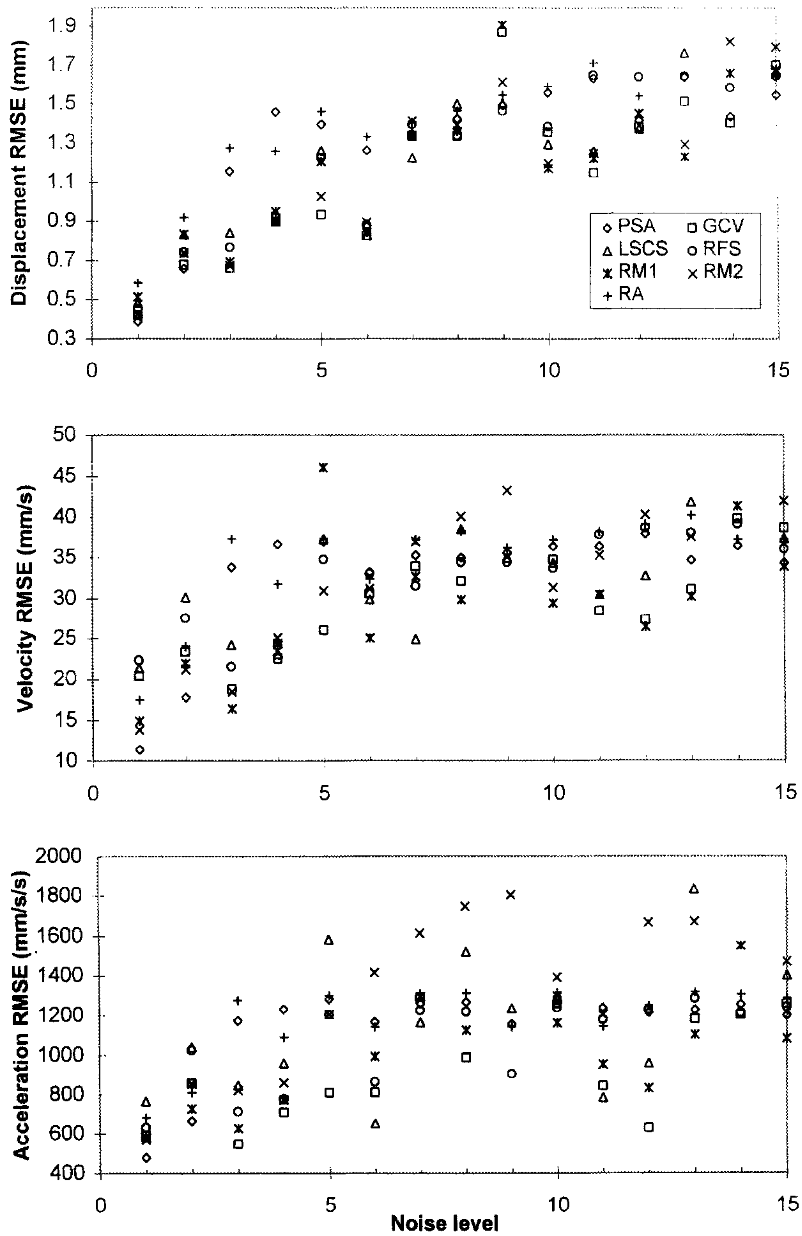

Fig. 1. The root mean square error (RMSE) in the medial–lateral displacement (top), velocity (centre) and acceleration (bottom) of the marker attached on the right tibial tubercle. Refer to the text (Methods section) for explanation of abbreviations. Any symbols not shown in this graph have higher values.

number of 1440 noisy signals. The characteristics of these signals are described elsewhere (Giakas and Baltzopoulos, 1997)*.

RESULTS

Figure 1 shows the RMSE calculated by using all methods in the displacement, velocity and acceleration domains respectively, of the right tibial tubercle marker in the medial–lateral direction throughout the first 15 levels of 'random' noise. For the same reference signal, depending on the characteristics of noise (in this example there are 15 different cases-levels), a certain method might produce the best results in some cases but the worst in some others. No method was consistently better than all the other methods, i.e. it produced the lowest RMSE values throughout all noise levels. Generally, the RMSE tended to increase as the noise level increased. The effects of the derivative order are also shown; e.g. in level 14, the best two in displacement were GCV and PSA, in velocity PSA and RA and in acceleration GCV and RFS. The results indicated similar conclusions for all the other markers and levels of noise.

Each method was ranked according to the RMSE results. The methods with the lowest and next lowest RMSE were ranked as 'best', the next two methods as 'medium' and those with the highest three RMSE as 'worst'. This procedure was repeated for all 1440 signals for each derivative domain and created a 'rank score' for each filtering technique. Only the number of best cases was examined.

The performance of each automatic filtering technique is demonstrated in Fig. 2. The results indicated that among the investigated techniques there was not one that would be optimal for all signals. For

* All signals as well as additional information on signal development can be downloaded from the International Society of Biomechanics WWW (http://www.kin.ucalgary.ca/isb/giakas).

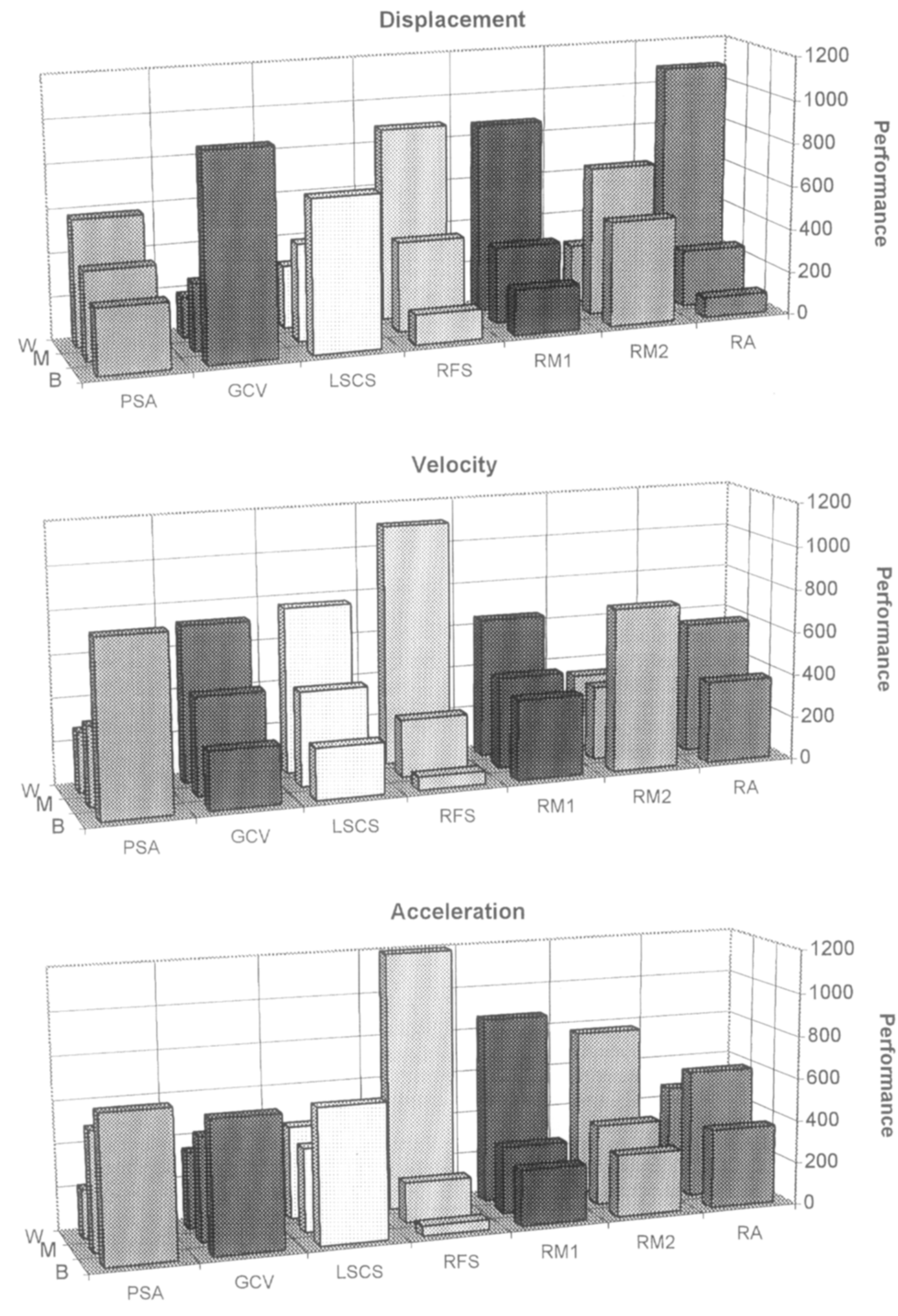

Fig. 2. The total number of the best, medium and worst ranking for each method for displacement (top), velocity (centre), and acceleration (bottom) measurements. Refer to the text (Methods section) for explanation of abbreviations.

displacement data, the two time domain methods (GCV and LSCS) were superior to all the other methods. These two produced the lowest root mean square error in 955 cases out of 1440. For velocity data the above two methods (GCV and LSCS) presented problems and in most cases they were classified as medium and worst. The best two methods for calculating the velocity were PSA and RM_2. These two produced the lowest RMSE in 884 cases out of 1440. In the acceleration domain, despite the improvement of performance of GCV, the best were PSA and LSCS. These two were the best in 796 of the cases.

DISCUSSION

DAmico and Ferrigno (1992) showed the superiority, in some cases, of their method LAMBDA when compared with GCV. According to their results, the behaviour of GCV in the zeroth derivative was sometimes even better than LAMBDA, but LAMBDA was much better in the higher derivatives. These results are in agreement with the findings of the present study. Computation speed was not generally assessed by the present study because all methods needed less than 3 s to filter and differentiate the signals in a 486 based computer, and therefore timing differences were considered insignificant.

Differentiation alters the performance of the automatic filtering methods. Therefore if a certain method is optimal for smoothing a certain signal, then it is not guaranteed that this method will be optimal for calculating the first and second derivatives of that signal. Therefore, if a certain technique has been ranked as 'best' in the displacement domain, it is not guaranteed that will be also 'best' in the higher derivatives (Giakas and Baltzopoulos, 1997).

The second regression model (RM_2) performed much better than the first one (RM_1). The calculation of cut-off frequency using RM_1 is based not only on the sampling frequency as in RM_2, but also on information from the power spectrum of the signal recorded. It is therefore a more complex method that is supposed to distinguish between original signal and noise harmonics, however, its performance was poor. The $RM_{1,2}$ were developed based on a particular type of signal collected by filming a falling ball. Therefore, it might not be appropriate for other signals which are collected with different equipment or have different signal patterns and characteristics. The cut-off frequency calculated by RM_2 for all signals by the present study was 6.6 Hz since the sampling frequency was 50 Hz. For a sampling frequency of 100 Hz, the cut-off frequency calculated by RM_2 would be 9.2 Hz. This is probably too high for gait analysis especially when the signal-to-noise ratio is low.

The regularised Fourier series (RFS) method uses optimisation techniques similar to generalised cross validation. It was expected that these two optimisation methods would achieve approximately the same results. However, the performance of RFS was, among the investigated techniques, relatively poor in all derivative domains for most signals examined by the present study. This might have been caused by the violation of several assumptions. The fact that the data used represented not exactly one walking cycle might have been one of those reasons.

Poor performance was also shown by the RA method compared to the other methods. In most cases, this method underestimated the cut-off frequency calculated for the displacement. However an improvement was demonstrated in the velocity and acceleration domains (Fig. 2).

In conclusion the present study supports the hypothesis that there is no optimal solution or all-purpose automatic method to filter biomechanical data. Therefore the method used by each investigator should not simply be based on a 'black box' approach. Knowledge of the assumptions and limitations of every method along with careful examination of the signal characteristics both in the time and frequency domain are fundamental to optimal filtering. Some methods allow the experimenter to tune their parameters whenever information concerning the signal examined is known. Moreover, the examination of data in each derivative domain before and after filtering is essential. Results demonstrated that PSA, LSCS and GCV were the most appropriate methods for filtering kinematic data derived from gait analysis, however none of these was ranked as the best for all derivative domains.

The conclusions of the present study are not generally applicable to other types of biomechanical data. The focus of this study was on gait analysis, in which data are usually sampled using video-based systems at 50 Hz and filtering is applied on the position patterns of various markers attached upon various surface landmarks of the lower segments, and therefore the conclusions may not be applicable to different types of signals. Further investigation is required using a variety of signals with different total number of data, collected with different sampling rates and having different signal and noise characteristics.

Acknowledgments—The authors would like to express their gratitude to Dr Graeme A. Wood for his help with the filtering subroutines, Dr Lampros Stergioulas for assisting with the PSA software development, Prof. W. Simons and Dr Clark R. Dohrmann for providing software implementing their techniques, and to Dr Bing Yu for his on-line help. Finally, we appreciate the comments of Prof. Roger M. Bartlett, Mr John Buckley and journal referees on the manuscript.

REFERENCES

Anderssen, R. S. and Bloomfield, P. (1974) Numerical differentiation procedures for non-exact data. *Numererische Mathematik* **22,** 157–182.

Burkholder, T. J. and Lieber, R. L. (1996) Stepwise regression is an alternative to splines for fitting noisy data. *Journal of Biomechanics* **29,** 235–238.

Cullum, J. (1971) Numerical differentiation and regularisation. *SIAM Journal on Numerical Analysis* **8,** 254–265.

DAmico, M. and Ferrigno, G. (1990) Technique for the evaluation of derivatives from noisy biomechanical displacement data using a model-based-bandwidth-selection procedure. *Medical Biology and Engineering Computing* **28,** 407–415.

DAmico, M. and Ferrigno, G. (1992) Comparison between the more recent techniques for smoothing and derivative assessment in biomechanics. *Medical Biology and Engineering Computing* **30,** 193–204.

Dohrmann, C., Busby, H. and Trujillo, D. (1988) Smoothing noisy data using dynamic programming and generalised cross-validation. *Journal of Biomechanical Engineering* **110,** 37–41.

Giakas, G. and Baltzopoulos, V. (1997) Optimal digital filtering requires a different cut-off frequency strategy for the determination of the higher derivatives. *Journal of Biomechanics* **30,** 851–855.

Hatze, H. (1979) An algorithm for computing higher order derivatives of noisy experimental data sequences. Report Number TWISK 123. Nasionale Navorsingsinstituut vir Wiskundige Wetenskappe, Pretoria.

Hatze, H. (1981) The use of optimally regularized Fourier series for estimating higher-order derivatives of noisy biomechanical data. *Journal of Biomechanics* **14,** 13–18.

Hatze, H. (1990) Data conditioning and differentiation techniques. In *Biomechanics of Human Movement: Applications in Rehabilitation, Sports and Ergonomics*, eds N. Berme and A. Cappozzo, pp. 237–248, Bertec Corporation, Worthington, OH.

Miller, D. and Nelson, R. (1973) *Biomechanics of Sport*. Lea and Febiger, Philadelphia.

Simons, W. and Yang, K. (1991) Differentiation of human motion data using combined spline and least squares concepts. *Journal of Biomechanical Engineering* **113,** 348–351.

Smith, G. (1989) Padding point extrapolation techniques for the butterworth digital filter. *Journal of Biomechanics* **22,** 967–971.

Vaughan, C. L. (1982) Smoothing and differentiation of displacement-time data: an application of splines and digital filtering. *International Journal of Biomedical Computing* **13,** 375–386.

Winter, D. (1990) *Biomechanics and Motor Control of Human Movement* (2nd Edn.) Wiley-Interscience, Toronto, Ontario.

Woltring, H. (1986) A Fortran package for generalised, cross-validatory spline smoothing and differentiation. *Advanced Engineering Software* **8,** 104–107.

Wood, G. (1996) Personal communication.

Yu, B. (1989) Determination of the optimum cutoff frequency in the digital filter data smoothing procedure. *Proceeding of the 12th International Congress of Biomechanics*, University of California, Los Angeles.

Yu, B. (1996) Personal communication.