

Repetitive Motion Analysis: Segmentation and Event Classification

ChunMei Lu, *Member, IEEE*, and
Nicola J. Ferrier, *Member, IEEE*

Abstract—Acquisition, analysis, and classification of repetitive human motion for the assessment of postural stress is of central importance to ergonomics practitioners. We present a two-threshold, multidimensional segmentation algorithm to automatically decompose a complex motion into a sequence of simple linear dynamic models. No a priori assumptions were made about the number of models that comprise the full motion or about the duration of the task cycle. A compact motion representation is obtained for each segment using parameters of a damped harmonic dynamic model. Event classification was performed using cluster analysis with the model parameters as input. Experiments demonstrate the technique on complex motion.

Index Terms—Motion analysis, motion classification, event detection.

1 INTRODUCTION

THIS paper presents a method, suited for use in ergonomic studies, for segmenting and classifying repetitive motion. Quantification of the relationship between posture, repetitiveness, and risk of injury while performing a workplace task requires measurement and analysis of human motion during task execution. In ergonomic analysis, the current approaches for analyzing video sequences involves either trained observers [1] to numerically score task “repetitiveness,” or manually partitioning the video and annotating the video with text and estimated joint angles [2], [3], [4]. Manual segmentation is an expensive, time consuming process, and the scoring is subjective. Video-based machine analysis of human motion has potential to improve the data available for ergonomic studies by eliminating errors caused by human subjectivity, improving the accuracy of joint motion measurement, and increasing the quantity of data through automation of both data acquisition and analysis. To achieve this potential, automated systems must be able to provide data suited to ergonomic analysis in terms of both accuracy and representation; here, we focus on the latter.

Ergonomic studies require a motion representation that permits the comparison of motion from different persons and from the same person at different times. The variability found in common workplace tasks (e.g., assembly or packaging) precludes the development of specialized procedures to analyze each task. Dynamic systems representations of human activity have been used in numerous studies of human motor behavior [5], [6]. In recognition systems, motions have been represented as a sequence of simple movements, typically linear dynamic systems [7], [8], [9], [10]. Tracking and recognition techniques have used switching dynamic system models, but require the number of motion classes to be known (e.g., [7], [11], [9]).

Recognition of human activities [12], [13] does not typically extract the level of detail required for our application. Segmentation, as described in these works, pertains to “events” (e.g., walking, hopping) where the motion description is not of sufficient detail to compare individuals. In ergonomics, the issue is not whether a motion is, say, a reach or a wave, but rather, how rapidly is a task performed, how much rest between movements, how awkward is the posture, etc. However, these approaches may be suitable for decomposing long sequences then applying the methods presented

here to the shorter sequences. Likewise, the measure of self-similarity used to segment periodic motion [14] could be used to identify cycles. Such a hybrid approach would have the advantage that shorter sequences are easier to analyze because the space of possible segmentation times is greatly reduced.

Our focus here will be on determining the number and locus of motion segments, and in classifying and comparing motion based on observed patterns of joint activity. Others have sought to segment motion data without prior knowledge of the expected number of segments. “Via points” for cursive handwriting [15] were determined based on optimization of the trajectory (minimum jerk). Their objective function is “critically related to arm dynamics” [15], requiring a complete forward and inverse dynamic model of the arm. Nonlinear dynamic models are used by Ijspeert et al. [8] as the basic motion primitive useful for trajectory formation (“pattern generators”). Complex motions are compactly represented as weighted sums of the primitive motions. This method has the advantage that segmentation of motion is based on easily extracted zero velocity points. Although not yet explored by the authors, their representation (the weights associated with each primitive) lends itself to classification techniques. Additional effort will be required to map these weights into familiar ergonomics task descriptive statistics (e.g., frequency and amplitude). In work performed concurrently with ours, Vecchio et al. [10] presents a formal framework for human motion representation. A cost function is defined to jointly minimize errors in segmentation times and in parameter fitting. They classify motion using the dynamic model parameters. They prove, for ideal signals, that segmentation is unique and show that the error in segmentation time for a noisy system is bounded. Their algorithm finds a *single* segmentation point on a time interval, if it exists. Because we anticipate multiple segments, careful extensions of this work is required to avoid a combinatorial problem in searching the set of segmentation points.

We have developed a complete system, from video processing to obtain motion data through classification of motion patterns. The focus of this paper is on our method to automate the segmentation and classification of human motion based on motion parameters. We will assume that the observed human is performing a repetitive task. Segmentation is required primarily because the length of the task cycle (the time to execute the task) is not known a priori, nor will it be a constant interval and, the “component” motions that comprise the cycle will vary. Task motion will be described by sequences of simple linear dynamic models. Multidimensional segmentation methods [16] are used with a refinement step to automatically segment complex motion into a sequence of simple models where no a priori assumptions are made about the number of simple models that comprise the full motion. A compact motion representation is obtained in each motion segment using parameters obtained from damped harmonic dynamic model. Event detection and classification are performed using cluster analysis with the dynamic model parameters as input. Experimental evaluation demonstrates that complex motion (karate) can be handled with our methods. A discussion the on limitations and future directions are presented.

2 MULTIDIMENSIONAL SIGNAL SEGMENTATION

Human joint motion can be described as a discrete N-dimensional time series, e.g., joint angles and their derivatives. In this paper, a second order AR model was selected to model the motion data because of its simplicity and the clear interpretation of its parameters with respect to human motion.¹ Let $\mathbf{x}_k \in \mathbf{R}^n$ denote the n-dimensional state vector at time k . Using an extended state vector $\mathcal{X}_k = (\mathbf{x}_{k-1}^T, \mathbf{x}_k^T)^T$, the second-order AR process in matrix form is

$$\mathcal{X}_k - \bar{\mathcal{X}} = \mathbf{A}(\mathcal{X}_{k-1} - \bar{\mathcal{X}}) + \mathbf{B}\mathbf{e}_k, \quad \bar{\mathcal{X}} = \begin{pmatrix} \bar{\mathbf{x}} \\ \bar{\mathbf{x}} \end{pmatrix}, \quad (1)$$

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ \mathbf{A}_2 & \mathbf{A}_1 \end{pmatrix}, \quad \text{and } \mathbf{B} = \begin{pmatrix} \mathbf{0} \\ \mathbf{B}_0 \end{pmatrix}.$$

1. First-order models can describe complex motions using a (possibly prohibitive) large number of segments.

• The authors are with the Center for Mathematical Sciences and the Departments of Mechanical and Biomedical Engineering, University of Wisconsin, Madison, WI 53706.
E-mail: chunmei@cae.wisc.edu, ferrier@enr.wisc.edu.

Manuscript received 19 Nov. 2001; revised 9 Feb. 2003; accepted 10 Aug. 2003.
Recommended for acceptance by L. Quan.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 115412.

\mathbf{A}_2 , \mathbf{A}_1 , and \mathbf{B}_0 are $n \times n$ matrices, \mathbf{e}_k is zero mean Gaussian noise, and $\bar{\mathbf{x}}$ is the mean of \mathbf{x} . The system matrix of the process, \mathbf{A} , fully describes the deterministic part of the model, and matrix \mathbf{B} describes the stochastic part of the model. We consider each motion segment as one element in the task cycle, characterized by the parameters \mathbf{A} and \mathbf{B} .

Using the above model, the *motion segmentation problem* is to detect the times at which the dynamical parameters of \mathbf{A} (\mathbf{A}_1 and \mathbf{A}_2) change significantly. There are a number of algorithms for change detection in dynamic systems, typically based on minimizing the error between the data and the dynamic model given a choice of segmentation (e.g., [16], [15]). We introduce a two threshold algorithm for segmentation, an extension of our work detailed in [17], summarized as follows:

MULTIDIMENSIONAL SEGMENTATION ALGORITHM

A.1 Smooth the data with a low-pass filter and subtract the mean value of the data (1).

A.2 Apply recursive least squares [16] to the data to estimate the dynamic parameters.

A.3 Compute the Frobenius norm of the difference of the parameter matrices between the current and previous time step. Treat the value of this norm as a one-dimensional time series and apply RLS to determine locations where the norm deviates from zero.

A.4 For each segment, perform model fitting using system identification techniques [16].

If the fitting error is greater than a predetermined threshold, T_1 , **go to step 2** for a finer segmentation, **else, output** the times of the segmentation points (and **stop**).

This segmentation algorithm produces an initial set of possible segmentation points. To avoid oversegmentation and/or to control the number of segments obtained, the initial segmentation is refined as follows (where n is the number of segmentation points sought):

MERGE SEGMENT ALGORITHM

B.1 **Input** two thresholds T_1, T_2 with $T_1 < T_2$ and initialize: $n = 1$.

B.2 Apply the segmentation algorithm above on motion data using fitting error threshold, T_1 . Let $S_1 = \{p_1, \dots, p_M\}$ denote the temporally ordered set of M segmentation points output.

B.3 Pick n distinct points $\{p_1, \dots, p_n\}$ in S_1 that maximize the inter-point distance:

a. compute the distance between the segmentation points

$$d_i = |p_i - p_{i-1}|, \text{ for } i = 2, \dots, M$$

b. sort by d_i and select the n points with the highest d_i values.

B.4 Measure the fitting error using these segmentation points.

If the fitting error is less than T_2 , **stop**;

else increase the number of segmentation points, $n = n + 1$. **Go to step 3**.

This algorithm has a number of useful alternatives: In step B.3, search all $\binom{M}{n}$ combinations for the best fit, or terminate when a prespecified n is reached, or search only for a specific n (this is relevant for motion comparison). The intuition behind step B.3 is that a large number of segmentation points in a short period of time was usually produced by noisy data (which can be ignored), and such rapid changes are hard to detect visually.

Example results of segmenting a reaching motion are shown in Fig. 1. Three segments were found for each reach. In [18], we compared reach motions for multiple subjects, all resulted in three segments. Similarly, most "walking" models use four segments. Ideally, segmentation should produce the same segmentation for each task cycle. Unfortunately, in the presence of noise, this cannot be guaranteed [10].

3 MOTION REPRESENTATION

Previous human motion studies used statistics such as average and peak values in time and frequency domain to represent univariate motion data on each segment of a task cycle [19]. To be consistent with previous ergometric analysis, after segmentation, we considered the motion of each joint independently. Thus, for each

segment and each joint, we have a time series to which we fit a damped harmonic model²

$$\alpha \exp(-\beta t) \cos(2\pi f t + \phi) + T, \quad (2)$$

where α , β , f , ϕ , and T represent the amplitude, damping rate, frequency, phase, and offset information of the damped harmonic oscillation motion, respectively. The different harmonic modes are determined by the eigenvalues, λ_i of the fitting parameter matrix³ \mathbf{A} . The damping rate β_i and frequency f_i , given the sampling frequency τ , are

$$\beta_i = \frac{1}{\tau} \log \frac{1}{|\lambda_i|} \text{ and } f_i = \frac{1}{2\pi\tau} \arg \lambda_i.$$

Amplitude, phase, and offset information can be obtained by minimizing the error between harmonic model and the original motion data. The phase information can be determined from the amplitude in damped harmonic motion model and thus is ignored from the motion representation.

The four-dimensional vector $[\alpha_{ij}, \beta_{ij}, f_{ij}, T_{ij}]$ is a compact representation of the motion data $x_{ij}(t_k)$ (the i th component of the state vector on the j th segment). This representation has two advantages for our application. First, the frequency and amplitude are physically intuitive for measuring repetitive motion stress. Second, this description of the motion is suitable for classification techniques, as shown below.

Fig. 1 shows the computed feature values for three joint angle dynamics on each segment for 10 cycles of a reaching motion. The feature value on each task cycle is represented as a step function whose height is the value of that parameter.

4 EVENT DETECTION IN REPETITIVE SEQUENCES

Motion model parameters were treated as features and data from sample motions of the same task were used to learn the motion pattern. Noise in the data and human variation in performing a repetitive task produce segments that do not in general have identical lengths. Combining or comparing the sample motions requires a definition of "similar" and a metric to quantify the distance between two motions. Our definitions are presented below, as we construct our classifier. Our input was a set of n -dimensional motion data measured on M distinct task cycles. Each cycle was segmented and on each segment we represented joint motion using model parameters (a 4-vector for each segment/joint combination). Additionally, an estimate of the error in fitting the dynamic model (the RMSE value over the segment) and the length of the segment are known. These values were used to quantify an representative value for each feature, on each segment, for the motion. Once obtained, classical pattern recognition techniques can be used.

To compare motions, we define a metric based on the dynamic model parameters. For each of the n joints, we define a feature vector, $\mathbf{d} = [\alpha, \beta, f, T]$, and we measure $\mathbf{d}(k)$ on each segment $k = 1, \dots, s$, for M cycles. Consider one feature $v = v^\ell$ (we will drop the superscript ℓ for clarity). Let $\mathbf{v} = [v_1, \dots, v_s]$ be the vector containing feature values of a specific feature on each segment. For task cycles indexed by i and j , let \mathbf{v}_i and \mathbf{v}_j denote the vector obtained by concatenating the vectors from each segment, then we define the distance between \mathbf{v}_i and \mathbf{v}_j with respect to feature ℓ as the weighted sum of the differences of each of its components for each segment:

2. The second order AR process is a class of Gaussian Markov processes of order 2. A Markov process in continuous time is governed by the Fokker-Planck or Kolmogorov equation [20]. The sample path of such a process in continuous time in one dimension is a damped harmonic oscillation.

3. Note that, when estimated from one dimensional data, \mathbf{A} has two eigenvalues λ_1 and λ_2 which correspond to one mode if both eigenvalues are complex, and two modes otherwise. Harmonic modes with a high damping rate influence the motion only for a very short time, and the least damped modes characterize the motion. Hence, we always consider the first mode of the motion in the two mode case.

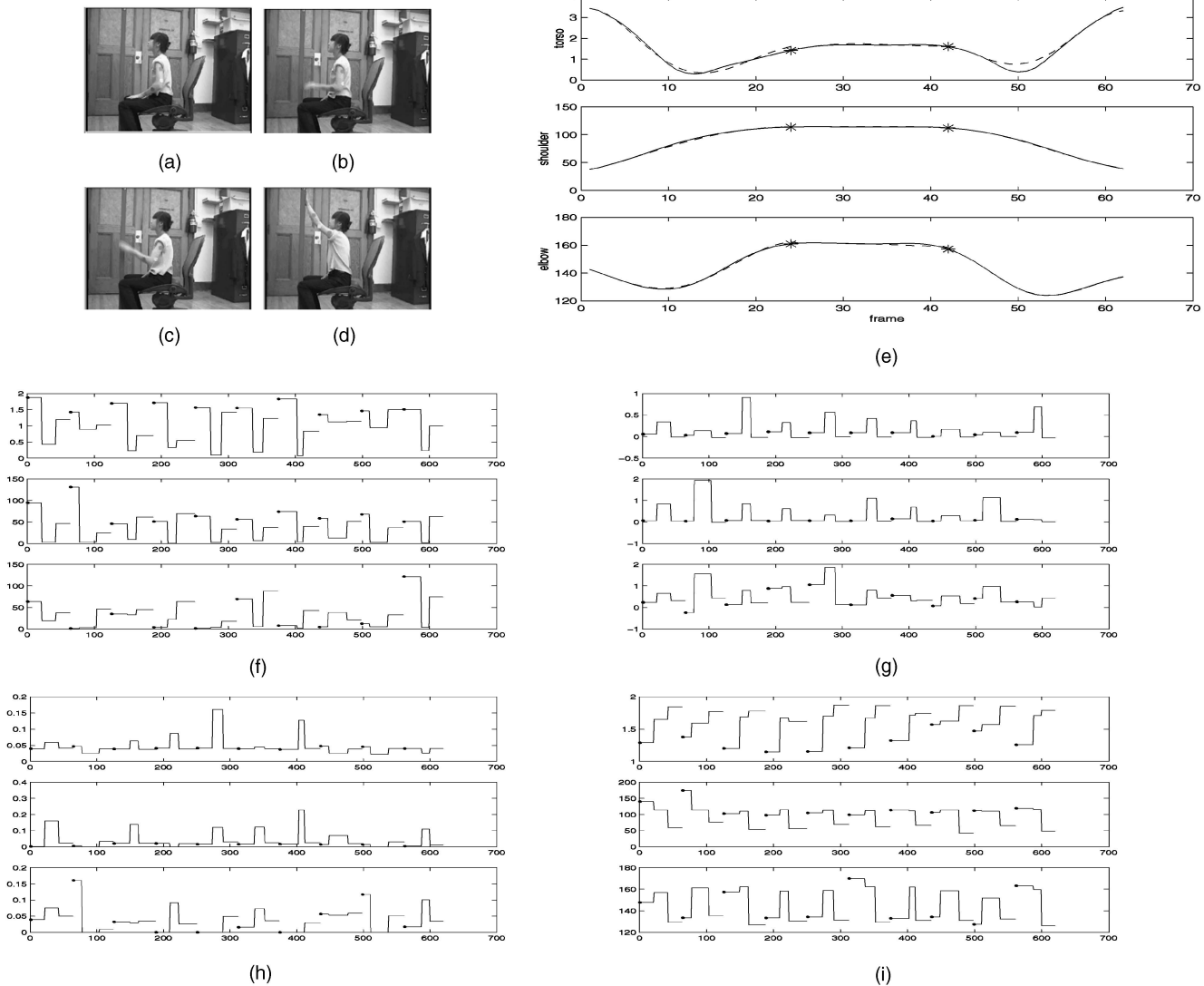


Fig. 1. (a)-(d) Four images from the motion sequences of a human subject performing a reaching task. (e) Three segments (delineated with asterisks) found for reaching motion. Original data shown with a solid line, harmonic model fit with a dashed line. Step function representation of the motion parameters for 10 cycles of reaching: (f) amplitude; (g) damping rate; (h) frequency; (i) offset for three joint angles (top to bottom). “.” indicates the start of a cycle.

$$D_{ij}^{\ell} = |\mathbf{v}_i - \mathbf{v}_j| \triangleq \sum_{k=1}^s w_k |\mathbf{v}_{i,k} - \mathbf{v}_{j,k}|, \quad w_k^2 = e_{i,k} \cdot e_{j,k} \cdot \frac{|l_{i,k} - l_{j,k}| + 1}{\max(l_{i,k}, l_{j,k})},$$

where $e_{i,k}$ is the residual fitting error, and $l_{i,k}$ is the measured length for the k th segment of the i th cycle. This weight penalized the segments with larger fitting errors and different lengths. Clearly, performance will vary with alternate definitions of feature difference. Comparison of motion depends on the application. While the actual motion in space may be identical for two persons, in ergonomic studies, this is *not* important. This particular weight was designed to achieve properties of prior ergonomic analysis (motions whose frequencies match on an interval are expected to contribute to postural stress equivalently [19]).

We build distance matrices of dimension $M \times M$, D_{ij}^{ℓ} for each feature for each joint. We defined a representative value for each feature, μ^{ℓ} , as the feature value of the sample motion that minimizes the sum of its distance to the other motion samples

$$\mu^{\ell} = \operatorname{argmin}_j \sum_i D_{ij}^{\ell}$$

(i.e., sum up either the rows or columns of the distance matrix (it is symmetric) and find the minimum, take the corresponding feature value). Because the feature values are obtained from segments with

different lengths and fitting errors, it does not make sense to average the sample values. Instead, μ^{ℓ} is a representative value for those features observed in the sample motion pattern, weighted based on fitting error and variation in segment lengths.

Fig. 2c gives a grayscale representation of the distance matrix, D_{ij}^{ℓ} , where low values of D are dark and high values are white. The motion consists of 20 cycles: 10 cycles of reach a low target, followed by 10 cycles of reaching a high target, thus each matrix contains 20×20 blocks and each block, D_{ij}^{ℓ} , represents the distance between the parameter estimated for the motion on cycle i and cycle j . The distance matrix is shown for the shoulder joint for amplitude, damping factor, frequency, and offset from left to right.

Given a new observation of a motion, the motion can be compared to the representative motion parameters μ^{ℓ} for known motion classes to determine which motion matches best. As an example, we classified reach motions using three joint angles, into two groups (high versus low target). Ten samples of each motion were used to determine μ^{ℓ} for each joint/feature combination. Fig. 2a plots motion data obtained when a subject reached for a low target 10 times, then for a 20cm higher target 10 times. The segmentation and fitting results are illustrated in Fig. 2b. The distance matrix of the motion data for the shoulder joint with respect to the four features is shown in grayscale in Fig. 2c.

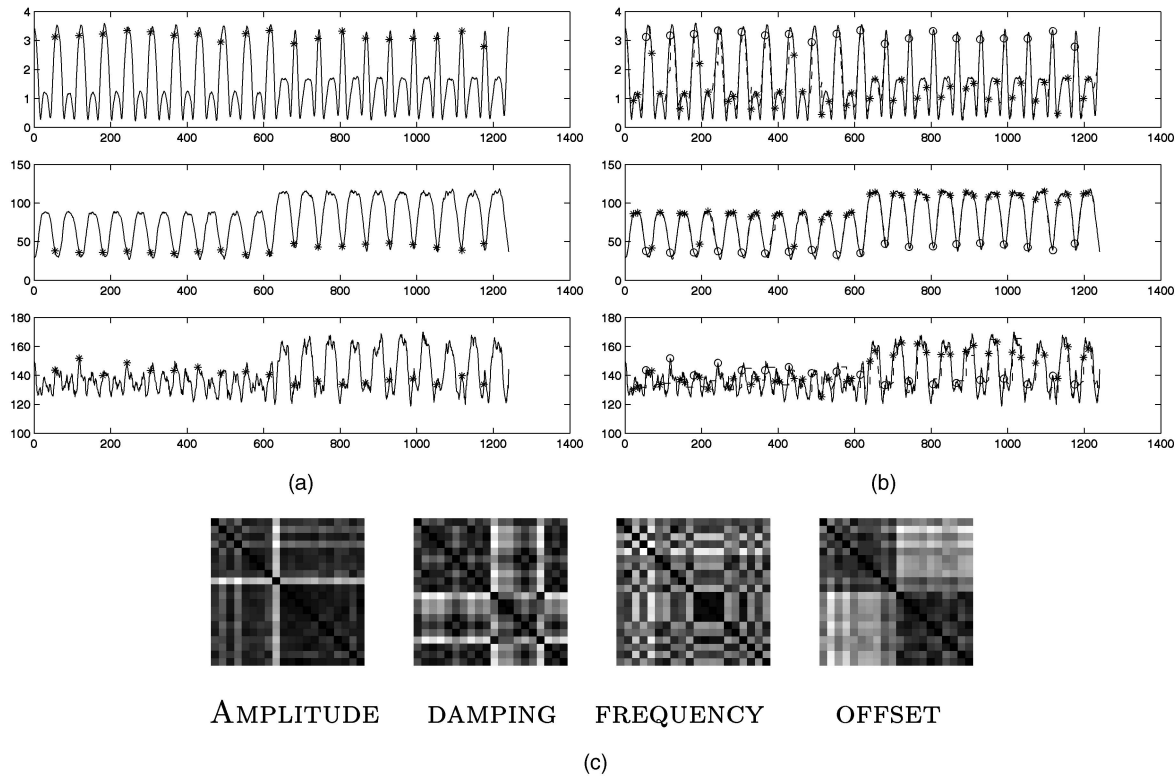


Fig. 2. (a) Motion data of 10 cycles of a subject reaching a lower target followed by 10 cycles of reaching a higher target shown for elbow, shoulder, and torso. (b) Segmentation and fitting result: "o" indicates end points of each cycle. Three segments (marked with an asterick) were found for each cycle. (c) Grayscale representation of the distance matrices, D_{ij} , of the shoulder motion data with respect to the motion parameters amplitude, damping rate, frequency, and offset. Each block, D_{ij} , represents the distance between the parameter estimated for the motion on cycle i and cycle j with dark representing low values (small distances). Notice that the offset parameter distances are smaller for the similar motions (the upper 10 by 10 block and lower 10 by 10 block are dark), making this feature useful for classification. In contrast, the frequency parameter is similar for both reach motions and this is reflected in the distance matrices which show no identifiable clustering of values.

The distance matrices of the amplitude and offset parameters of the shoulder motions (columns 1 and 4 in Fig. 2) show distinct cluster of low values for each cycle, i.e., the upper 10×10 block and lower 10×10 block are dark indicating small distances between features for these cycles. In classical pattern recognition techniques [21], the discrimination power of a certain feature is a quantitative measure of the degree to which that feature, as distinct from all other features, is able to discern a given pattern. It is an expression of the probability of pattern misclassification (hence, the smaller the feature power, the better feature discriminates between the patterns). For the example, in Fig. 2, the computed feature power is lowest for the offset and amplitude (approximately one-tenth of the other two features), consistent with the clustering observed in the distance matrix for these features. Thus, for these particular reaching tasks, classification can be performed using only these features.

5 EXPERIMENTS

An analysis of a karate sequence (Fig. 3) is presented to illustrate applicability to complex motion. In this motion, a martial artist performs three swift kicks. The motion data is six-dimensional, comprised of upper leg angle, knee angle, and ankle angle from both left and right legs. Given the complexity of the data, it is not clear how to determine the number of motion classes. Fig. 3 shows the segmentation and fitting result of our multidimensional segmentation routine. Three joint angles from left and right legs are shown in Figs. 3e and 3f, respectively. It was found that there were a total of 10 segments. Fig. 3 shows representative images, one from each segment. Five sample images from third segment and fourth segment are shown in Figs. 3c and 3d, where the motions appear to correspond, respectively, to leg raising and leg lowering motions.

Another experiment involved classification between three reach motions (low, medium, and high targets). Each motion was found to

consist of three segments (reach, push, retract). The distances between targets was approximately 10 cm. In the video images, it was hard for a human to distinguish the motions. Our training data consisted of 60 sequences with 10 reaches to each target and 30 idle phases between the reaches. An independent test set consisted of 19 high, 17 medium, and 18 low target sequences and 50 idle phases. Pattern classification was performed as outlined above. The confusion matrix for classification of the independent test motions is tabulated below where each entry corresponds to the probability that a motion which is in the *row* category was classified as the *column* category. The idle phase, and distinguishing between high and low targets were easily separated from other motions. There was a 5 percent probability of misclassification between reaching the middle and the other two targets.

Confusion matrix for classification of reach				
	low	medium	high	idle
low	1.0	0.0	0.0	0.0
medium	0.06	0.94	0.0	0.0
high	0.0	0.05	0.95	0.0
idle	0.0	0.0	0.0	1.0

6 DISCUSSION

We have presented a method to automatically segment complex motion data such that each individual segment can be described by a simple damped harmonic model. A pattern classifier was trained using the parameters of the dynamic model. Our approach assumes no a priori knowledge about the number of dynamic models. The automatic segmentation is important for analyzing human motion during performance of repetitive tasks due to the variety and complexity of tasks. The dynamic model parameters form a compact

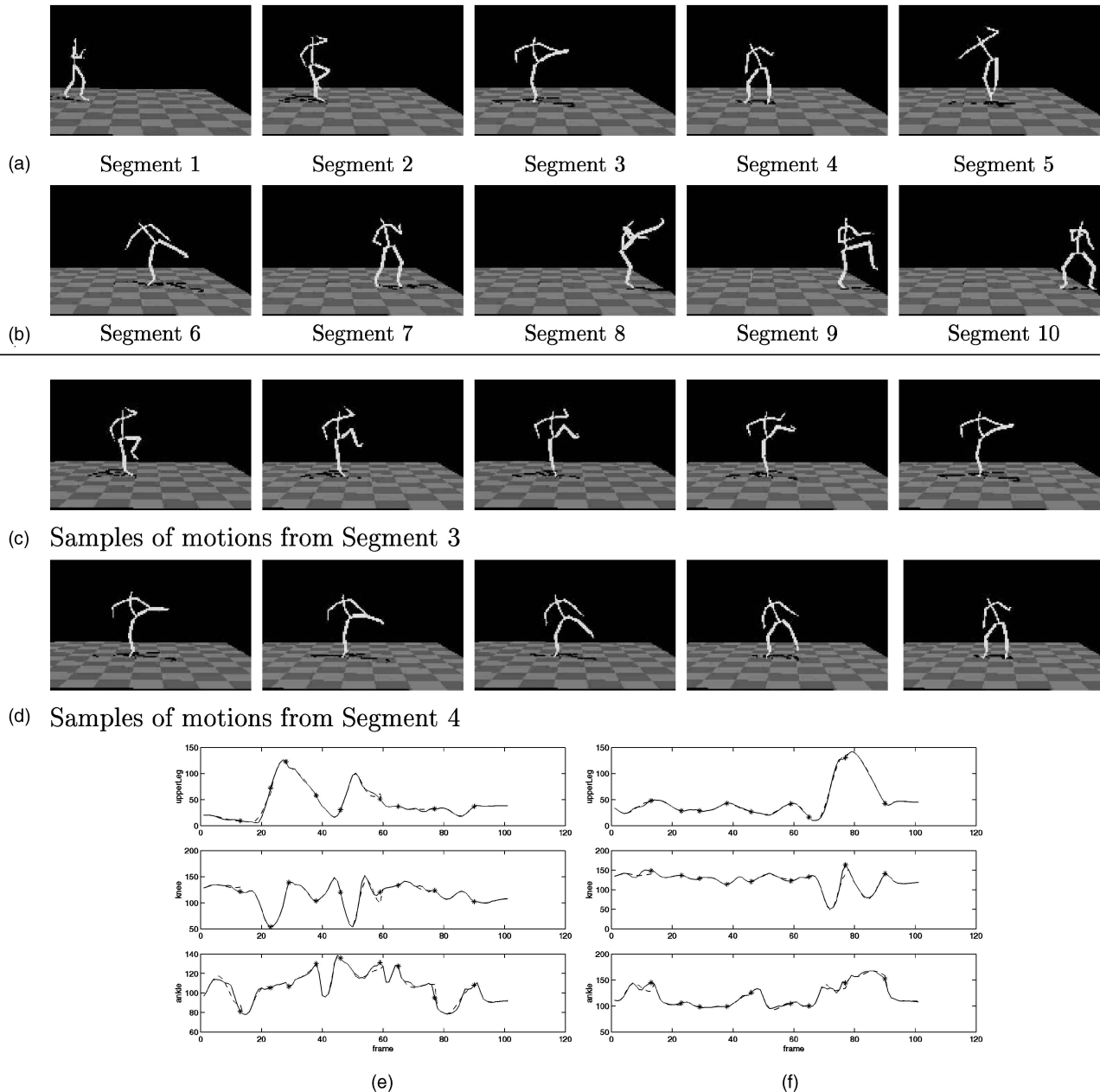


Fig. 3. (a) and (b) Representative images from each of the 10 segments in the karate kick sequence. (c) Samples from third segment and (d) fourth segment (bottom) in karate kick sequences. 6d karate kick data: 10 segments (astericks) and joint angles fit for (e) left leg, and (f) right leg. Original data is shown with solid line, the harmonic model with a dashed line.

representation of the motion data amenable to cluster analysis for classification of events. Our segmentation and classification experiments have demonstrated promising results in both aspects.

There are some drawbacks to the techniques presented here. First, there are a number of parameters, most are associated with segmentation, that effect the system performance and must be tuned. Filtering for noise reduction is a standard procedure where the parameter is determined based on the sensor/algorithm used. Gaussian smoothing with a σ of 1 frame was used. The thresholds, T_1 and T_2 bound the error for the fit to the linear model on each segment and cycle, respectively. Lower thresholds yield a better fit with more segments. The application might offer some insight to fitting tolerances. The RLS includes: a forgetting factor that effects how much the systems matrices must change to force a segmentation, a window length over which to fit, and a threshold for declaring sufficient deviation from zero in the univariate time series. We used a window of 30 frames and 0.98 for the forgetting

factor. The peak univariate time series over the window was used. Clearly our values will not hold for all motion series and these parameters values must be determined.

The pattern classifier works without tuning parameters. However, changing the weight function is expected to change the performance because this function defines "similar" motions. Some applications may, however, wish to classify motions as similar even if one parameter varies. Variations in defining the distance would enable such flexibility in comparing motion. Distance was defined for motions with the same number of segments, thus errors in segmentation will lead to misclassification.

Clustering and subsequent classification techniques can only recognize variations of motions *known* to the classifier. For systems trained on repetitive motion, this is not a problem. For other applications, using the simple "best fit" matching, the classification computation grows with the size of the known motions.

Notwithstanding the number of parameters required for segmentation, our technique works on complex noisy data and

over repeated trials. Failures during our experiments most frequently involved the acquisition of data. Tasks with subtle differences were classified with minimal error. No initial estimates of the dynamic model were used (as is often required with switching linear systems [7]).

The system in its current form would integrate well with existing algorithms. For example, a recognition system [12], [13] could be used to perform preliminary segmentation into subtasks ("lift box from conveyor to pallet"). Then, the techniques presented here could be applied to determine motion segments. Further, data acquired while observing the repetitive task could be used to further refine the motion segments. The dynamic models obtained could be used to initialize more effective classifiers [7], [11], [9].

ACKNOWLEDGMENTS

This work was supported by the US National Science Foundation (IRI-9703352) and by a University of Michigan Center for Occupational Health and Safety Engineering Pilot Research Award. Thanks to House of Moves, Inc. for providing the karate data.

REFERENCES

- [1] W. Latko, T. Armstrong, J. Foulke, G. Herrin, R. Rabourn, and S. Ulin, "Development and Evaluation of an Observational Method for Assessing Repetition in Hand Tasks," *Am. Industrial Hygiene Assoc. J.*, vol. 58, no. 4, pp. 278-285, 1997.
- [2] G. Ahanger and T. Little, "A Survey of Technologies for Parsing and Indexing Digital Video," *J. Visual Comm. and Image Representation*, vol. 7, no. 1, pp. 28-43, 1996.
- [3] H. Wactlar, T. Kanade, M. Smith, and S. Stevens, "Intelligent Access to Digital Video: The Imformedia Project," *IEEE Computer*, vol. 29, no. 5, pp. 46-52, 1996.
- [4] T. Yen and R. Radwin, "A Video-Based System for Acquiring Biomechanical Data Synchronized with Arbitrary Events and Activities," *IEEE Trans. Biomedical Eng.*, vol. 42, no. 9, pp. 944-948, 1995.
- [5] *Human Motor Behavior*. J.A.S. Kelso ed., Lawrence Erlbaum Associates, 1982.
- [6] F.A. Mussa-Ivaldi and E. Bizzi, "Learning Newtonian Mechanics," *Self-organization, Computational Maps, and Motor Control*, Morasso and Sanguietti, eds., pp. 491-501, Elsevier, 1997.
- [7] C. Bregler, "Learning and Recognizing Human Dynamics in Video Sequences," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, June 1997.
- [8] A.J. Ijspeert, J. Nakanishi, and S. Schaal, "Movement Imitation with Nonlinear Dynamical Systems in Humanoid Robots," *Proc. IEEE Int'l Conf. Robotics and Automation*, 2002.
- [9] V. Pavlovic and J. Rehg, "Impact of Dynamic Model Learning on Classification of Human Motion," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 788-795, 2000.
- [10] D. Del Vecchio, R. Murray, and P. Perona, "Primitives for Human Motion: A Dynamical Approach," *15th IFAC World Congress on Automatic Control*, 2002.
- [11] N. Howe, M. Leventon, and W. Freeman, "Bayesian Recognition of 3D Human Motion from Single-Camera Video," *Proc. Neural Information Processing Systems Conf. (NIPS '99)*, 1999.
- [12] M. Brand and V. Kettner, "Discovery and Segmentation of Activities in Video," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 844-851, Aug. 2000.
- [13] Y.A. Ivanov and A.F. Bobick, "Recognition of Visual Activities and Interactions by Stochastic Parsing," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 852-872, Aug. 2000.
- [14] R. Cutler and L. Davis, "Robust Real-Time Periodic Motion Detection Analysis and Applications," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 781-796, Aug. 2000.
- [15] Y. Wada and M. Kawato, "A Theory for Cursive Handwriting Based on the Minimization Principle," *Biological Cybernetics*, vol. 73, pp. 3-13, 1995.
- [16] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*. MIT Press, 1983.
- [17] C.M. Lu, H.Z. Liu, and N.J. Ferrier, "Multidimensional Motion Segmentation and Identification," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 629-636, 2000.
- [18] C.M. Lu and N.J. Ferrier, "Automatic Analysis of Repetitive Joint Motion," *IEEE Trans. Information Technology in Biomedical Systems*, (to appear), 2003.
- [19] R. Radwin, M. Lin, and T. Yan, "Exposure Assessment of Biomechanical Stress in Repetitive Manual Work Using Frequency-Weighted Filters," *Ergonomics*, vol. 37, no. 12, pp. 1984-1998, 1994.
- [20] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*. Springer, 1991.
- [21] R. Duda and P. Hart, *Pattern Classification and Scene Analysis*. John Wiley and Sons, 1973.