A Model (In)Validation Approach to Gait Classification

María Cecilla Mazzaro, *Member*, *IEEE*, Mario Sznaier, *Member*, *IEEE*, and Octavia Camps, *Member*, *IEEE*

Abstract—This paper addresses the problem of human gait classification from a robust model (in)validation perspective. The main idea is to associate to each class of gaits a nominal model, subject to bounded uncertainty and measurement noise. In this context, the problem of recognizing an activity from a sequence of frames can be formulated as the problem of determining whether this sequence could have been generated by a given (model, uncertainty, and noise) triple. By exploiting interpolation theory, this problem can be recast into a nonconvex optimization. In order to efficiently solve it, we propose two convex relaxations, one deterministic and one stochastic. As we illustrate experimentally, these relaxations achieve over 83 percent and 86 percent success rates, respectively, even in the face of noisy data.

Index Terms—Gait classification, activity recognition, model (in)validation, risk-adjusted (in)validation.

1 Introduction

THE problem of human activity recognition has received considerable attention within the Computer Vision community (see, for instance, the surveys [1] and [9]). Existing approaches to this problem can be divided into those recasting activity recognition as a classification problem, and those that seek to exploit context information (see [9], [19] and references therein).

Classification approaches are based on the general idea of matching experimental data provided by an earlier segmentation stage against a set of suitable representations of different actions of interest. Roughly speaking, most of these approaches accomplish this by finding the "nearest neighbor," in some suitable distance, either stochastic or deterministic. Among the several approaches available in the literature that fall in this general category, one can distinguish the use of spatio and/or temporal templates [10], [3]; principal component analysis [24], [12]; Bayesian-based stochastic methods [17], [21]; Hidden Markov Models [5], [6], [20], [11]; and the use of dynamical black or gray box models underlying the activities under consideration [2].

While these approaches have proven successful in many scenarios, in others they may fail due to the effects of measurement noise and errors in extracting the features under consideration. In addition, methods that exploit underlying models (either Markovian or deterministic) are susceptible to failure due to modeling errors, arising from the fact that, in practical situations, the parameters of these models are identified from a finite set of noisy measurements. Finally, stochastic approaches, due to their nature, cannot unequivocally falsify the experimental information, to indicate cases where the observed data does not match any of the available models.

In order to circumvent these difficulties, in this paper we propose a method for robust gait classification based upon recasting the problem into a robust model (in)validation form. Specifically, we will associate to each class of gaits a family of models, represented by a nominal model and bounded model uncertainty, and a class of admissible inputs, representing measurement and process noise. In this context, the problem of determining whether or not a given experimental sequence corresponds to a particular activity type reduces to establishing whether this sequence could have been generated by a combination of the given model, and some elements from its associated classes of model uncertainty and inputs: a model (in)validation problem. However, contrary to the standard invalidation case discussed in the Robust Control literature [7] where the input is known, here the only information available is a set membership characterization, based on spectral data. This leads to a nonconvex, generically NP-hard problem. To solve this difficulty, we propose two relaxations: one convex and one stochastic. These relaxations achieved over 83 percent and 86 percent success rates, respectively, with real data.

- *M.C. Mazzaro is with GE Global Research, Niskayuna, NY 12309. E-mail: mazzaro@research.ge.com.*
- *M. Sznaier and O. Camps are with the Department of Electrical Engineering, The Pennsylvania State University, University Park, PA 16802. E-mail: {msznaier@frodo.ee.psu.edu, camps@psu.edu}.*

Manuscript received 8 Apr. 2004; revised 22 Mar. 2005; accepted 23 Mar. 2005; published online 14 Sept. 2005.
Recommended for acceptance by M. Pietikainen.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-0169-0404.

2 Preliminaries

2.1 Notation

$\ell^2[0,N]$ denotes the space of square summable, real-valued sequences $\{x_i\}_{k=0}^N$ equipped with the norm $\|x\|_{\ell^2}^2 \doteq \sum_{i=0}^N x_i^2$. $\mathcal{L}$ denotes the space of causal, linear time invariant (LTI) operators bounded in $\ell^2[0,\infty)$. It is well-known that this space is isomorphic to $\mathcal{H}_\infty$, the space of matrix functions essentially bounded on the unit circle and with bounded analytic continuation inside the unit disk, equipped with the norm: $\|G\|_\infty \doteq ess\sup_{|z|<1} \overline{\sigma} G(z)$, where $\overline{\sigma}$. denotes maximum singular value. In the sequel, to any finite sequence $\{\mathbf{x}_k\}$, we will associate the following Toeplitz matrix:

$$\mathbf{T}_x^n = \begin{bmatrix} \mathbf{x}_0 & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{x}_1 & \mathbf{x}_0 & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{x}_n & \mathbf{x}_{n-1} & \dots & \mathbf{x}_0 \end{bmatrix}.$$

By a slight abuse of notation, to each LTI system S we will associate the matrix $\mathbf{T}_S^n$ obtained from the first $n+1$ elements of its impulse response sequence.

2.2 Modeling Human Gait

The problem of *modeling* human gait has been extensively researched (see, for instance, the surveys [1], [9], [14]) leading to several different approaches that can be roughly divided into biomechanical motivated [8], [15] and input/ouput (or gray/black box) box modeling [2], [6]. In the sequel, we will concentrate on input/output models since we are interested in classifying gaits, rather than explaining the internal processes involved. In this situation is advantageous to use smaller, simpler models since model order is directly related to the computational complexity of the algorithms involved. However, we want to stress the fact that the approach pursued here is completely general and can be applied to other models and/or activities.

Specifically, following [2], we will represent each of the activities under consideration as realizations of a second order stationary stochastic process. Thus, to each activity we can associate a discrete LTI system $\mathcal{S}_i$, driven by white zero-mean Gaussian noise $\mathbf{e}$:

$$\mathbf{x}_{k+1} = \mathbf{A}^i\mathbf{x}_k + \mathbf{K}^i\mathbf{e}_k^i, \quad \mathbf{y}_k^S = \mathbf{C}^i\mathbf{x}_k + \mathbf{w}_k^i, \tag{1}$$

where $\mathbf{w}$ denotes measurement noise. The output of the model, $y^S = \{\mathbf{y}_k^S\}$, is a vector sequence containing the angular displacements, relative to their mean values, of the shoulder, elbow, hip, and knee joints of the target. In this formulation, the actual values of the matrices $\{\mathbf{A}^i, \mathbf{K}^i, \mathbf{C}^i\}$ can be obtained from the experimental data using subspace identification methods [23], [2].

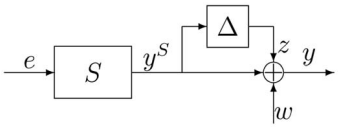

Fig. 1. The gait classification set-up.

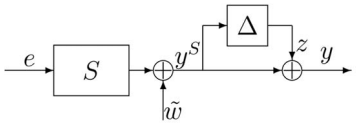

Fig. 2. Alternative, jointly convex gait-classification setup.

2.3 Gait Classification as a Model (In)Validation Problem

The goal of this paper is to, given 1) a set of gaits $\mathcal{G} = \{\mathcal{G}_i\}$, each represented by a set of exemplars, and 2) a sequence of frames from an unknown activity, determine whether the latter corresponds to one of the known gaits in $\mathcal{G}$. We propose to solve this problem by postulating that all time sequences corresponding to realizations of a given gait $\mathcal{G}_o$ can be obtained as the output of some LTI system S_o to an unknown input signal $e \in \ell^2$, $\|e\|_{\ell^2} \leq 1$.[1] This leads to the set-up shown in Fig. 1, where S is one of the models in the set $\mathcal{S}$, and where y and w represent the experimental data and (energy bounded) measurement noise, respectively. The block Δ represents (dynamic) model uncertainty to account for the fact that the models S were identified using finite data sequences, corrupted by noise (see the Appendix for a discussion of the role of Δ). Motivated both by theoretical and practical considerations we will use the ℓ^2 norm[2] to measure the distance between two given experimental sequences $d(y_i, y_j) = \|y_i - y_j\|_{\ell^2}$. Thus, since all other signals involved are also in ℓ^2, in the sequel the operator Δ will be characterized in terms of its ℓ^2-induced (or $\mathcal{H}_\infty$) norm. In this context, the gait classification problem can be precisely recast in the following model (in)validation form:

Problem 1. *Given 1) experimental data $\{\mathbf{y}_k\}$, $k = 0, \ldots, n$, consisting of $(n+1)$ measurements of the angles of the shoulder, elbow, hip, and knee joints of a person, and 2) some a priori information consisting of the set of models $\mathcal{S}$ and the following set descriptions $\mathcal{E}$, $\mathcal{N}(\eta)$, and $\mathbf{\Delta}$ of admissible inputs, noise, and uncertainty blocks:*

$$\begin{aligned} \mathcal{E} &= \Big\{ e \in \ell^2 \colon \|e\|_{\ell^2[0,n]} \leq 1 \Big\} \\ \mathcal{N}(\eta) &= \Big\{ w \in \ell^2 \colon \|w\|_{\ell^2[0,n]} \leq \eta \Big\}, \\ \mathbf{\Delta} &= \big\{ \Delta \in \mathcal{H}_\infty \colon \|\Delta\|_{\ell^2 \to \ell^2} \leq \delta \big\}, \end{aligned} \tag{2}$$

where the constants η and δ characterize the noise and uncertainty levels, determine if there exists at least one quadruple $(e_i, w_i, \Delta_i, S_i) \in \mathcal{E} \times \mathcal{N} \times \mathbf{\Delta} \times \mathcal{S}$ that can reproduce the available experimental evidence, that is:

$$y = (I + \Delta_i) S_i * e_i + w_i, \tag{3}$$

where $*$ denotes convolution.

If the answer to this problem is negative, then the experimental sequence does not correspond to any of the known gait types $\mathcal{S}_i$. On the other hand, if more than a feasible quadruple exists, the unknown sequence can be assigned to the class corresponding to the smallest uncertainty norm, $\|\Delta_i\|_\infty = \delta_i$. In this sense, the proposed approach belongs to the class of "nearest-neighbor" classifiers, with a metric selected to improve robustness against model uncertainty and noise.

3 Main Results

In this section, we show that Problem 1 can be transformed into a finite-dimensional convex optimization and efficiently solved with commercially available software.

1. This is a deterministic, set-membership based equivalent to (1).
2. Using the ℓ^2 rather than the peak (ℓ^∞) norm minimizes the effects of outliers in the experimental data.

Theorem 1. *Problem 1 has an affirmative answer if and only if there exists at least one pair of finite sequences $e = \{\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_n\} \in \mathcal{E}$ and $w = \{\mathbf{w}_0, \mathbf{w}_1, \cdots, \mathbf{w}_n\} \in \mathcal{N}$ such that*

$$\mathbf{T}_z^n \mathbf{T}_z^n \leq \delta^2 (\mathbf{T}_S^n \mathbf{T}_e^n)^T \mathbf{T}_S^n \mathbf{T}_e^n, \tag{4}$$

where $\mathbf{T}_z^n \doteq \mathbf{T}_y^n - \mathbf{T}_{y^S}^n - \mathbf{T}_w^n$, and where $\mathbf{T}_y^n$, $\mathbf{T}_e^n$, $\mathbf{T}_w^n$, and $\mathbf{T}_S^n$ are defined as in Section 2.

Proof. Begin by noting that the interconnection $\{\mathcal{S}, \Delta\}$ shown in Fig. 1 could have generated the experimental evidence y if and only if there exists an operator $\Delta \in \mathbf{\Delta}$ and signals $e \in \mathcal{E}$, $w \in \mathcal{N}$, z, and y^S satisfying the following equations:

$$y^S = S * e, \; z = \Delta * y^S, \; z = y - y^S - w \tag{5}$$

or, equivalently,

$$\begin{aligned} &\mathbf{T}_{y^S}^n = \mathbf{T}_S^n \mathbf{T}_e^n, \; \mathbf{T}_z^n = \mathbf{T}_\Delta^n \mathbf{T}_{y^S}^n, \\ &\mathbf{T}_z^n = \mathbf{T}_y^n - \mathbf{T}_{y^S}^n - \mathbf{T}_w^n. \end{aligned} \tag{6}$$

From Lemma 2 in the Appendix, there exists $\Delta \in \mathbf{\Delta}$ mapping the sequence y^S to z if and only if $(\mathbf{T}_z^n)^T \mathbf{T}_z^n \leq \delta^2 (\mathbf{T}_{y^S}^n)^T \mathbf{T}_{y^S}^n$. Substituting back in (6) and reordering yields the desired result. □

Unfortunately (4) is not jointly convex on all the variables involved, due to the cross-terms $\mathbf{T}_w^{n^T} \mathbf{T}_S^n \mathbf{T}_e^n$. Rather, it can be shown that it is equivalent to a Bilinear Matrix Inequality (BMI) optimization. These problems are known to be generically NP-hard [22] and, thus, computationally expensive to solve. In order to obtain computationally tractable algorithms, in the sequel we propose two convex relaxations of Problem 1: one deterministic and one stochastic.

3.1 A Deterministic Convex Relaxation

Consider the alternative setup shown in Fig. 2, where the measurement noise is also affected by the unknown error dynamics Δ:

$$y = (I + \Delta)(S * e + \tilde{w}). \tag{7}$$

When compared to the original setup shown in Fig. 1, it can be easily seen that the only difference is in the measurement noise level. Specifically, if there exists a triple $(e, \tilde{w}, \Delta)$ satisfying (7) with $\|\tilde{w}\|_{2[0,n]} \leq \tilde{\eta} \doteq \frac{\eta}{1 + \|\Delta\|_\infty}$, then $w = (I + \Delta)\tilde{w} \in \mathcal{N}(\eta)$ and the triple (e, w, Δ) satisfies (3). Thus, one can attempt to find a solution to the original problem by searching for a solution to the model (in)validation problem shown in Fig. 2, with noise level $\tilde{\eta}$. As we show next, this leads to a convex optimization problem. In addition, one will expect that if $\|\Delta\|_\infty \ll 1$, then this approximation is not too conservative. This conjecture will be experimentally substantiated in Section 4.

Theorem 2. *There exists a quadruple $(e, \tilde{w}, \Delta, S) \in \mathcal{E} \times \tilde{\mathcal{N}} \times \mathbf{\Delta} \times \mathcal{S}$ that satisfies (7) if and only if there exists at least one pair of finite sequences $e = \{\mathbf{e}_0, \mathbf{e}_1, \cdots, \mathbf{e}_n\} \in \mathcal{E}$ and $\tilde{w} = \{\tilde{\mathbf{w}}_0, \tilde{\mathbf{w}}_1, \cdots, \tilde{\mathbf{w}}_n\} \in \tilde{\mathcal{N}}$ such that the following LMI holds:*

$$\mathbf{A}_1(e, \delta) \doteq \begin{bmatrix} \mathbf{X}(e) & (\mathbf{T}_{\tilde{w}}^n + \mathbf{T}_S^n \mathbf{T}_e^n)^T \\ \mathbf{T}_{\tilde{w}}^n + \mathbf{T}_S^n \mathbf{T}_e^n & (\delta^2 - 1)^{-1} \mathbf{I} \end{bmatrix} \leq 0, \tag{8}$$

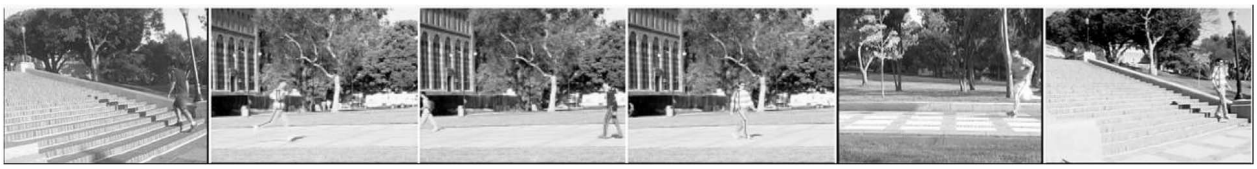

Fig. 3. Sample frames of four different activities. These images are courtesy of Professor S. Soatto, UCLA.

where

$$\mathbf{X}(e) \doteq (\mathbf{T}_y^n)^T \mathbf{T}_y^n - (\mathbf{T}_y^n)^T (\mathbf{T}_S^n \mathbf{T}_e^n + \mathbf{T}_{\tilde{\boldsymbol{w}}}^n) - (\mathbf{T}_{\tilde{\boldsymbol{w}}}^n + \mathbf{T}_S^n \mathbf{T}_e^n)^T \mathbf{T}_y^n$$

and $\mathbf{Y}(\mathbf{e}) \doteq \begin{bmatrix} \mathbf{e_0^T} & \mathbf{e_1^T} & \cdots & \mathbf{e_n^T} \end{bmatrix}^{\mathbf{T}}$.

Proof. From (7) (see also Fig. 2), we have that

$$\begin{aligned} \mathbf{T}_{y^S}^n &= \mathbf{T}_{\tilde{\boldsymbol{w}}}^n + \mathbf{T}_S^n \mathbf{T}_e^n, \quad \mathbf{T}_z^n = \mathbf{T}_\Delta^n \mathbf{T}_{y^S}^n, \\ \mathbf{T}_y^n &= \mathbf{T}_z^n + \mathbf{T}_{y^S}^n. \end{aligned} \tag{9}$$

As in the proof of Theorem 1, there exists $\Delta \in \boldsymbol{\Delta}$ mapping the input-output sequences (y^S, z) if and only if $(\mathbf{T}_z^n)^T \mathbf{T}_z^n \leq \delta^2 (\mathbf{T}_{y^S}^n)^T \mathbf{T}_{y^S}^n$. Substituting in (9) yields:

$$(\mathbf{T}_y^n - \mathbf{T}_{y^S}^n)^T (\mathbf{T}_y^n - \mathbf{T}_{y^S}^n) - \delta^2 (\mathbf{T}_{y^S}^n)^T \mathbf{T}_{y^S}^n \leq 0.$$

The LMI (8) follows now by reordering terms and using Schur complements (see [4], Chapter 2). □

3.2 A Risk-Adjusted Convex Relaxation

As mentioned before, the convex relaxation discussed in Section 3.1 is not too conservative for cases where $\|\Delta\|_\infty$ is small. On the other hand, if this condition does not hold, (3) might be infeasible in cases where the original problem has a solution. In order to handle these cases, in this section, we propose a stochastic relaxation of the original problem that has polynomial, rather than exponential, computational complexity growth with the problem data [13].

The main idea of the method is to uniformly sample the set of admissible uncertainties $\boldsymbol{\Delta}$, in an attempt to find at least one element Δ_i so that model $(I + \Delta_i)S$ together with an admissible input $e \in \mathcal{E}$ and output noise $w \in \mathcal{N}$ can explain the experimental data y. This removes the interpolation constraint that makes the problem nonconvex in (e, w) since for a fixed Δ_i Problem 1 has an affirmative answer if and only if there exists at least one feasible solution $e \in \mathcal{E}$ to the convex problem $y - (I + T_{\Delta_i}) T_S e \in \mathcal{N}$. This observation leads to the following gait classification algorithm:

Algorithm 1. Given an experimental sequence y, the nominal model of a gait type S and a value of δ, take N_t samples of $\boldsymbol{\Delta}$, $\{\Delta_i(z)\}_{i=1}^{N_t}$ using Algorithm 2 described in the Appendix.

1. For each Δ_i, solve the following convex feasibility problem in e:

$$\text{CPe}: \ \text{Find } e \in \mathcal{E} \text{ such that} \qquad y - (I + T_{\Delta_i}^n) T_S^n e \in \mathcal{N}.$$

TABLE 1
Experimental Data

Person	Walking	Running	Staircase
A	1, 2	16 to 18	25 to 27
B	3 to 8	11 to 15	21 to 24
C	9, 10	none	28 to 30
D	none	19	none
E	none	20	none

2. If there exists at least one feasible e, stop. Otherwise, consider next sample $\Delta_{i+1}(z)$ and go back to Step 1.

The algorithm finishes, either by finding one admissible e or after N_t steps. In the former case, we assign the given sequence to the gait represented by model S. In the latter, there exists a (small) risk of incorrectly concluding that the sequence does not correspond to the gait represented by S. The next result, adapted from [16] gives a bound on this risk:

Lemma 1. *Let (ϵ, ν) be two positive constants in $(0, 1)$. If*

$$N_t \geq \frac{\ln(1/\nu)}{\ln(1/(1-\epsilon))}, \tag{10}$$

then

$$\text{Prob}(\exists \Delta : \text{CPe feasible} | \text{CPe infeasible for } \Delta_i, i = 1 \ldots, N_t) < \epsilon$$

and this event occurs with probability greater than $(1 - \nu)$. This implies that, with confidence greater that $1 - \nu$ the algorithm above has a probability smaller than ϵ of not finding Δ, if one exists.

4 Experimental Validation

In this section, we illustrate the effectiveness of the proposed algorithms using experimental data. We begin by outlining a method to compute suitable nominal models for each gait class.

4.1 The Experimental Data

The experimental data, the same used in [2], consists of 30 vector sequences y_k, taken from five different persons, labeled A, B, C, D, and E. Each sequence contains measurements of the angles of the shoulder, elbow, hip, and knee joints of a person walking, running, or walking a staircase, and are assumed to be corrupted by additive noise of bounded energy. Fig. 3 shows sample frames from the videos used to generate the data sets. The trajectories of the joints over time were extracted by using a variant of the algorithm proposed in [6] where the skeleton of the human is represented as a kinematic chain supporting ellipsoidal texture patches. For illustrative sake, the data sequences are numbered from 1 to 30 so that 1-10 correspond to walking, 11-20 to running and 21-30 to walking a staircase, as shown in Table 1.

4.2 The Nominal Models

In the sequel, we will use models of the form (1), with $n_s = 4$ states,[3] $n_i = 4$ inputs, and $n_o = 4$ outputs. In order to use these models, the experimental sequences should be normalized to have zero mean, and scaled so that the corresponding input e_i has unit energy. Under the assumption that gaits are second order stationary, mean ergodic random processes, an unbiased estimate of the mean of each exemplar is given by the temporal average: $\boldsymbol{\mu}_i = E(\mathbf{y}_{\text{raw}}^i) \approx \sum_{k=0}^{n-1} \mathbf{y}_{k,\text{raw}}^i / n$. The data used to compute the model S_i associated with the ith gait is given by $\mathbf{y}^{S_i} = \mathbf{y}_{\text{raw}}^i - \boldsymbol{\mu}_i$. Similarly, an estimate of $\epsilon_i^2 \doteq (n+1)\text{E}(\mathbf{e}_k^T \mathbf{e}_k)$, the energy of an input sequence e can be estimated by $(n+1) \sum_{k=0}^{n^*-1} \mathbf{e}_k^T \mathbf{e}_k / n^*$. A normalized system having unity-energy input is then given by $\hat{\boldsymbol{S}}_i \doteq \epsilon_i S_i$.

Finally, to minimize the computational time, rather than representing each class of activities by all the models $\{S_i^j\}$

3. This represents a good compromise between model quality and complexity.

corresponding to different experimental realizations y_i^j of the gait $\mathcal{G}^j$, we will choose as a *single* representative for this class the model $S^j \in \{S_i^j\}$ that is closest to each other element in the class, in the sense of minimizing the norm of the (multiplicative) uncertainty required to map the two models under consideration, i.e.,

$$S^j = \arg \min_{\hat{S}_i^j, \hat{S}_k^j \in \mathcal{S}^j} \left\{ \| (\hat{S}_i^j - \hat{S}_k^j) \hat{S} j_k^{-1} \|_\infty \right\}. \tag{11}$$

In order to avoid introducing biases, we will compute the nominal models using all sequences in a given class while testing a sequence *outside* that class; and excluding the test sequence while matching it with its class (*leave-one-out* method). For instance, when classifying the sequence y_1, the walking, running, and climbing stairs models where obtained using the sequences $y_2 - y_{10}$, $y_{11} - y_{20}$, and $y_{21} - y_{30}$, respectively.

4.3 The Results

4.3.1 Using the Deterministic Convex Relaxation

Table 1 shows the results of applying Theorem 1, using 20 sample points per sequence and noise level of $\tilde{\eta} = 0.10$.,[4] i.e. $\|\tilde{w}\|_{\ell^2[0,19]} \leq \tilde{\eta}$. In all cases, the first column contains the experimental sequences to be recognized; the second, third, and fourth columns display the minimum size of the uncertainty block Δ measured in the $\mathcal{H}_\infty$ norm, so that nominal models S_{walk}, S_{run}, and S_{stair} can reproduce the given data. Each sequence is assigned to the class corresponding to the smallest uncertainty value. As shown in Table 2a, the proposed method can successfully recognize 25 out of the 30 sequences under consideration; with the misclassifications likely due to the relatively short length of the data record used for classification. Indeed, using 25 sample data points allows for correctly classifying also the sequence y_{21} and considering 30 data points further extends these results to include the sequences y_{29} and y_{30}. However, these additional successes come at the price of a computational complexity increase since the computational complexity of conventional LMI solvers grows as the fifth power of the number of decision variables.

For comparison, we also classified the experimental sequences by finding the nearest neighbor in the Martin distance sense, proceeding as in [2]. This approach misclassified eight out of the 30 sequences (y_3, y_5, y_6, y_7, y_{15}, y_{16}, y_{17}, and y_{27}).

4.3.2 Using the Risk-Adjusted Relaxation

As indicated before, the main idea of this relaxation is to use a combination of sampling and convex optimization to determine the minimum size of the uncertainty, δ, required for each of the nominal models—either S_{walk}, S_{run}, or S_{stair}—to reproduce a given experimental sequence. This was accomplished by taking an initial value of δ (a good upper bound is the value furnished by the noise free case) and using Algorithm 1 to check whether there exists at least one sample $\Delta_n \in \mathbf{\Delta}$ so that Problem CPe has a feasible solution. If the answer was affirmative, then the value of δ was decreased. For each experimental sequence, we carried out this procedure for each nominal model S_i, until the minimum value of δ_i for which the associated problem CPe is feasible was found. Proceeding as before, we then assigned the given sequence y to the gait represented by the nominal model corresponding to the minimum value of δ_i.

Table 2b shows the results obtained using the procedure outlined above using $N_t = 6,000$ samples over the set $\mathbf{\Delta}$ and an output noise level of $\eta = 0.06$. This value of N_t guarantees, with confidence 0.9975, a probability smaller than 0.001 of incorrectly determining δ. Compared to the convex relaxation, this approach has an 86.67 percent success rate, but at the price of increased computational requirements. Note, however, that the required samples $\{\Delta_i\}$ can be generated offline and, thus, not all of this additional computational burden needs to be carried out in real time.

4. This value was selected so that, with an actual noise level $\eta = 0.06$, infeasibility of the convex relaxation is a sufficient condition for infeasibility of the original problem for values of $\|\Delta\|$ up to 0.4.

5 Conclusions

This paper approaches the problem of human gait classification from a model (in)validation viewpoint. The main idea is to associate to each class of activities a nominal model and a class of bounded energy inputs. In this context, the problem of recognizing a sequence of frames can be formulated as the problem of determining whether this sequence could have been generated by a combination of the given model, some admissible uncertainty and its associated class of inputs. By exploiting norm constrained interpolation theory results, this problem can be recast as a finite dimensional, albeit generically nonconvex, NP-hard, optimization problem. In order to efficiently solve this problem, in the paper, we propose two different convex relaxations, one deterministic and one stochastic. As we illustrate with experimental sequences, the deterministic relaxation achieves good success rate (in this case, over 83.33 percent), even in the face of noisy data. Better success rates, at the price of additional computational complexity, can be obtained by resorting to a stochastic, risk-adjusted relaxation. This stochastic relaxation has a computational complexity that grows polynomially (rather than exponentially as in the original nonconvex problem) with the number of frames. Moreover, the computational-time requirements can be reduced by using the deterministic relaxation as a first step to improve the a priori available bounds on the uncertainty size. Finally, it is worth noticing that the sequences misclassified by our algorithm are different from the ones misclassified by the algorithm proposed in [2]. This suggests that robustness could potentially be improved by running both algorithms and then reexamining the sequences where the answers differ, for instance, by considering longer data sequences only for these cases.

Appendix A

A.1 Background Results

In this section, we recall, for ease of reference, the interpolation theory results used to recast the gait recognition problem into a tractable convex finite dimensional optimization form.

Lemma 2 ([7], Chapter 2). *Given matrix valued sequences $\mathbf{U}_i \in \mathbf{C}^{l\times j}$ and $\mathbf{V}_i \in \mathbf{C}^{k\times j}$, $i = 0, 1, \ldots n$, there exists a matrix interpolant $L(z) \in \mathcal{BH}_\infty^{k\times l}$ such that*

$$L(z) = \mathbf{L}_0 + \mathbf{L}_1 z + \mathbf{L}_2 z^2 + \ldots + \mathbf{L}_n z^n + \ldots$$
$$\mathbf{T}_L^n \left[\mathbf{U_o^T} \ldots \mathbf{U_n^T}\right]^T = \left[\mathbf{V_o^T} \ldots \mathbf{V_n^T}\right]^T$$

if and only if $(\mathbf{T}_U^n)^T \mathbf{T}_U^n - (\mathbf{T}_V^n)^T \mathbf{T}_V^n \geq 0$.

The following algorithm, described in [16], generates uniformly distributed finite impulse responses $\{h^i\}_{i=1}^{N_t}$ with $h^i = \{\mathbf{H}_0^i, \mathbf{H}_1^i, \ldots, \mathbf{H}_N^i\}$, $\mathbf{H}_k^i \in \mathbf{R}^{m\times s}$, so that the function $H^i(z) \doteq \sum_{k=0}^N \mathbf{H}_k^i z^k$ *can be completed* to belong to $\mathcal{BH}_\infty$, the unit ball in $\mathcal{BH}_\infty$.

Algorithm 2. Let $k = 0$. Generate N_t samples uniformly distributed over the set $\{\mathbf{H}_0 : \bar{\sigma}(\mathbf{H}_0) \leq 1\}$.

1. Let $k := k + 1$. For every generated sample $(\mathbf{H}_0^i, \mathbf{H}_1^i, \ldots, \mathbf{H}_{k-1}^i)$, consider the partition

$$\begin{bmatrix} \mathbf{H}_k^i & \cdots & \mathbf{H}_1^i & \mathbf{H}_0^i \\ \mathbf{H}_{k-1}^i & \cdots & \mathbf{H}_0^i & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{H}_0^i & \mathbf{0} & \cdots & \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_k^i & \mathbf{B} \\ \mathbf{C} & \mathbf{A} \end{bmatrix} \tag{12}$$

and let the matrices $\mathbf{Y}$ and $\mathbf{Z}$ be a solution of the linear equations:

$$\mathbf{B} = \mathbf{Y}(\mathbf{I} - \mathbf{A}^T\mathbf{A})^{\frac{1}{2}}, \quad \mathbf{C} = (\mathbf{I} - \mathbf{A}\mathbf{A}^T)^{\frac{1}{2}}\mathbf{Z}.$$

TABLE 2
Classification Results

Sequence	$\delta_{\text{opt}}^{\text{walk}}$	$\delta_{\text{opt}}^{\text{run}}$	$\delta_{\text{opt}}^{\text{stair}}$
y_1	0.2465	0.8643	0.5449
y_2	0.3269	0.6986	0.4356
y_3	0.1442	0.8237	0.4328
y_4	0.1379	0.7122	0.5203
y_5	0.2003	0.8535	0.4067
y_6	0.0928	0.8292	0.3702
y_7	0.2872	0.7973	0.4138
y_8	0.3427	0.6450	0.3542
y_9	0.0151	0.8487	0.6091
y_{10}	0.4252	0.8547	0.5399
y_{11}	1	0.5077	0.9207
y_{12}	1	0.4061	0.8282
y_{13}	1	0.2792	0.5522
y_{14}	1	0.2498	0.5208
y_{15}	1	0.2502	0.5578
y_{16}	1	0.2158	0.3328
y_{17}	0.8224	0.2179	0.5769
y_{18}	1	0.1519	0.2866
y_{19}	1	0.5097	0.8586
y_{20}	1	0.3441	0.6613
$y_{21}{}^{\dagger}$	1	$0.2351^{\dagger}$	0.2379
y_{22}	0.6648	0.1557	0.0091
y_{23}	0.6978	0.2395	0.1159
$y_{24}{}^{\dagger}$	0.8741	$0.1738^{\dagger}$	0.3204
y_{25}	0.5831	0.5501	0.3040
y_{26}	0.7155	0.9091	0.3535
y_{27}	0.4143	0.5868	0.1475
$y_{28}{}^{\dagger}$	$0.1229^{\dagger}$	0.8705	0.7358
$y_{29}{}^{\dagger}$	$0.0160^{\dagger}$	0.6144	0.4095
$y_{30}{}^{\dagger}$	0.8085	0.3053	0.4034

(a)

Sequence	$\delta_{\text{opt}}^{\text{walk}}$	$\delta_{\text{opt}}^{\text{run}}$	$\delta_{\text{opt}}^{\text{stair}}$
y_1	0.2502	**	**
y_2	0.3292	**	**
y_3	0.1552	**	**
y_4	0.1385	**	**
y_5	0.2302	**	**
y_6	0.0995	**	**
y_7	0.2771	**	**
y_8	0.3368	**	**
y_9	0.0004	**	**
y_{10}	0.4350	**	**
y_{11}	**	0.5145	**
y_{12}	**	0.4380	**
y_{13}	**	0.2861	**
y_{14}	**	0.2694	**
y_{15}	**	0.2656	**
y_{16}	**	0.2218	**
y_{17}	**	0.2301	**
y_{18}	**	0.1612	**
y_{19}	**	0.5202	**
y_{20}	**	0.4486	**
y_{21}	**	**	0.2346
y_{22}	**	**	0.0001
y_{23}	**	**	0.1111
$y_{24}^{\dagger}$	**	0.1828	**
y_{25}	**	**	0.3084
y_{26}	**	**	0.3600
y_{27}	**	**	0.1567
$y_{28}^{\dagger}$	0.1239	**	**
$y_{29}^{\dagger}$	0.0145	**	**
$y_{30}^{\dagger}$	**	0.3125	**

(b)

(a) Using the deterministic convex relaxation: success rate 83.33 percent. (b) Using the stochastic relaxation ($**$ and $\dagger$ denote infeasibility and a misclassified sequence, respectively). Success rate: 86.67 percent.

2. Let

$$\mathbf{J}(\mathbf{H}_0, \mathbf{H}_1, \ldots, \mathbf{H}_{k-1}) \doteq |(\mathbf{I} - \mathbf{Y}\mathbf{Y}^T)^{\frac{1}{2}}|^m |(\mathbf{I} - \mathbf{Z}^T\mathbf{Z})^{\frac{1}{2}}|^s.$$

Generate $\lfloor N_t \mathbf{J}(\mathbf{H}_0^i, \mathbf{H}_1^i, \ldots, \mathbf{H}_{k-1}^i) \rfloor$ samples uniformly distributed over the set $\{\mathbf{W} : \overline{\sigma}(\mathbf{W}) \leq 1\}$. For each sample $\mathbf{W}^i$, compute

$$\mathbf{H}_k^i = -\mathbf{Y}\mathbf{A}^T\mathbf{Z} + (\mathbf{I} - \mathbf{Y}\mathbf{Y}^T)^{\frac{1}{2}}\mathbf{W}^i(\mathbf{I} - \mathbf{Z}^T\mathbf{Z})^{\frac{1}{2}}. \quad (13)$$

3. If $k \leq N$ go to Step 1. Otherwise, stop.

A.2 Model Identification and Uncertainty Descriptions

The problem of extracting models from experimental data has been extensively researched in the control community [23], [7]. In the past few years, a large portion of the effort has been directed toward obtaining worst-case identification error bounds, to account for the fact that any model identified from a finite set of noisy measurements is likely to be inaccurate. Next, we illustrate the issues involved and motivate the uncertainty description used in Section 2.3.

Consider the problem of identifying a LTI system with a transfer function $G_o(z) = 1$ from noisy measurements of its step

response: $y_k = (G * u)_k + d_k$. Since $G_o(z) = 1$, it follows that $y_k = 1 + d_k$. Consider now the following two systems:

$$G_1(z) = \frac{z - 0.8578}{z - 0.9},\ G_2(z) = \frac{z + 0.8578}{z + 0.9}.$$

Simple algebra shows that the corresponding step response sequences are given by:

$$y_{1k} = \{1, 1.0422, 1.0802, 1.1144, 1.1451, \ldots\},$$
$$y_{2k} = \{1, 0.9578, 0.996, 0.9616, 0.9924, \ldots\}.$$

Thus, if the measurement noise d can take values above 0.15, all of these systems are *undistinguishable*. If the noise level is 0.09, an experiment of length 4 will establish that G_1 is no longer consistent with the experimental measurements, but still cannot distinguish between G_o and G_2. This illustrates the concept of consistency set $T(y) = \{G: y = G * u + d\}$, e.g., the set of all plants compatible with the experimental outcome. From an identification standpoint, all plants in $T(y)$ could have generated the observed experimental data. Thus, its "size" gives an upper bound of the worst-case error incurred when choosing any one element of $T(y)$ as the identified systems. For instance, in the simple example above, assuming a noise level of 0.15, one could choose as the "identified" system $G_{id} = 0.5[G_1(z) + G_2(z)] = 0.5\frac{z^2 - 0.81}{z^2 - 0.1544}$. In this case, every element of the consistency set can be (conservatively) represented by $G(z) = G_{id}[1 + \Delta(z)]$, where $\Delta(z)$ accounts for the mismatch between the dynamics of the systems. For instance, if $\Delta_1(z) = \frac{z^2 - 1.4656}{z^2 - 0.1544}$, then the formula above yields $G(z) = 1$. Similarly $\Delta = \frac{G_1 - G_2}{G_1 + G_2}$ yields $G_{id}(1 + \Delta) = G_1$. A compact (albeit conservative) representation of the set of all systems compatible with the available experimental information is:[5]

$$G = G_{id}(1 + \Delta),\ \Delta \in \mathcal{H}_\infty,\ \|\Delta\|_\infty \leq \delta.$$

In the example above, $\|\Delta\|_\infty \leq \|0.5(G_1 - G_2)\|_\infty = 0.22$.

ACKNOWLEDGMENTS

The authors are indebted to Professor Stefano Soatto and Mr. Alessandro Bissacco, Deparment of Computer Science, University of California at Los Angeles, for providing the experimental data, the Matlab m-files to compute the Martin distances, and many discussions and suggestions concerning this manuscript. This work was supported by the US National Science Foundation, under grants IIS-0117387, ECS-0221562, and ITR-0312558 and by the US Air Force Office of Scientific Research, under grant FA 9550-05-1-0437.

5. Roughly speaking, this amounts to covering this set with a disk centered at G_{id}, with radius δ in the $\mathcal{H}_\infty$ topology.

REFERENCES

[1] J.K. Aggarwal and Q. Cai, "Human Motion Analysis: A Review," *Computer Vision and Image Understanding*, vol. 73, no. 3, pp. 428-440, 1999.
[2] A. Bissacco, A. Chiuso, Y. Ma, and S. Soatto, "Recognition of Human Gaits," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 52-57, Dec. 2001.
[3] A.F. Bobick and J.W. Davis, "The Recognition of Human Movement Using Temporal Templates," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 257-267, Mar. 2001.
[4] S. Boyd, L. ElGhaoui, E. Feron, and V. Balakrishnan, "Linear Matrix Inequalities in System and Control Theory," *SIAM Studies in Applied Math.*, vol. 15, 1994.
[5] M. Brand, N. Oliver, and A. Pentland, "Coupled Hidden Markov Models for Complex Action Recognition," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 994-999, June 1997.
[6] C. Bregler, "Learning and Recognizing Human Dynamics in Video Sequences," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 568-574, June 1997.
[7] J. Chen and G. Gu, *Control-Oriented System Identification*. John Wiley & Sons, 2000.
[8] D. Cunado, M.S. Nixon, and J.N. Carter, "Automatic Extraction and Description of Human Gait Models for Recognition Purposes," *Computer Vision and Image Understanding*, vol. 90, no. 1, pp. 1-41, 2003.
[9] D.M. Gavrila, "The Visual Analysis of Human Movement: A Survey," *Computer Vision and Image Understanding*, vol. 73, no. 1, pp. 82-98, Jan. 1999.
[10] M.A. Giese and T. Poggio, "Quantification and Classification of Locomotion Patterns by Spatio-Temporal Morphable Models," *Proc. IEEE Int'l Workshop Visual Surveillance*, pp. 27-34, July 2000.
[11] J. Hoey and J.J. Little, "Representation and Recognition of Complex Human Motion," *Proc. IEEE CS Conf. Computer Vision and Pattern Recognition*, pp. 752-759, June 2000.
[12] P.S. Huang, "Automatic Gait Recognition via Statistical Approaches for Extended Template Features," *IEEE Trans. Systems, Man, and Cybernetics-Part B: Cybernetics*, vol. 31, no. 5, pp. 818-824, Oct. 2001.
[13] P. Khargonekar and A. Tikku, "Randomized Algorithms for Robust Control Analysis and Synthesis have Polynomial Complexity," *Proc. IEEE Conf. Decision and Control*, pp. 3470-3475, Dec. 1996.
[14] T. Moeslund and E. Granum, "A Survey of Computer Vision-Based Human Motion Capture," *Computer Vision and Image Understanding*, vol. 81, no. 3, pp. 231-268, 2001.
[15] R. Plankers and P. Fua, "Articulated Soft Objects for Multiview Shape and Motion Capture," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 9, pp. 1182-1187, Sept. 2003.
[16] M. Sznaier, C. Lagoa, and M.C. Mazzaro, "An Algorithm for Sampling Subsets of $\mathcal{H}_\infty$ with Applications to Risk-Adjusted Performance Analysis and Model (In)Validation," *IEEE Trans. Automatic Control*, pp. 410-416, Mar. 2005.
[17] A. Madabhushi and J.K. Aggarwal, "A Bayesian Approach to Human Activity Recognition," *Proc. IEEE Workshop Visual Surveillance*, pp. 25-32, June 1999.
[18] R.J. Martin, "A Metric for Arma Processes," *IEEE Trans. Signal Processing*, vol. 48, no. 4, pp. 1164-1170, Apr. 2000.
[19] I.E.D. Minnen and T. Starner, "Expectation Grammars: Leveraging High-Level Expectations for Activity Recognition," *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 626-632, June 2003.
[20] B. North, A. Blake, M. Isard, and J. Rittscher, "Learning and Classification of Complex Dynamics," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 9, pp. 1016-1034, Sept. 2000.
[21] N.M. Oliver, B. Rosario, and A.P. Pentland, "A Bayesian Computer Vision System for Modeling Human Interactions," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 831-843, Aug. 2000.
[22] H.D. Tuan and P. Apkarian, "Low Nonconvex Rank Bilinear Matrix Inequalities: Algorithms and Applications," *Proc. Conf. Decision and Control*, pp. 1001-1006, Dec. 1999.
[23] P. VanOverschee and B. DeMoor, "Subspace Algorithms for the Stochastic Identification Problem," *Automatica*, vol. 29, no. 3, pp. 649-660, May 1993.
[24] Y. Yacoob and M.J. Black, "Parameterized Modeling and Recognition of Activities," *Proc. Int'l Conf. Computer Vision*, pp. 120-127, Jan. 1998.