Temporal Alignment Improves Feature Quality: an Experiment on Activity Recognition with Accelerometer Data

Hongjun Choi[1,2], Qiao Wang[1,2], Meynard Toledo[3], Pavan Turaga[1,2], Matthew Buman[3], Anuj Srivastava[4]

[1]Geometric Media Lab, School of Arts, Media and Engineering, Arizona State University
[2]School of Electrical, Computer and Energy Engineering, Arizona State University
[3]School of Nutrition and Health Promotion, Arizona State University
[4]Department of Statistics, Florida State University

Abstract

Activity recognition has been receiving significant attention from a variety of research areas such as human performance enhancement, health promotion, and human computer interaction. However, recognizing activities from accelerometer data still remains a challenging problem due to sensitivity to sampling rates, misalignment of data, and increased variability in activities among clinically relevant populations. In order to solve these issues, we adopt methods from functional analysis, which consider non-elastic rate variations in movement. The overall framework factors out temporal variability within activity classes, before leveraging robust machine learning pipelines for a given end-use. The proposed approach has been evaluated on 7 *classes of everyday activities with* 50 *subjects. The results indicate that proposed approach achieves improved performance with the improvements observed in separating similar classes that differ in temporal rates, and also demonstrate higher robustness to change in window lengths. These results suggest that temporal alignment should be considered a core part of activity recognition pipelines.*

1. Introduction

Activity recognition from wearable sensors has been an active area of research for the past several years due to its wide range of applications [2, 14]. In clinical applications, self-measures were often used to assess daily activities, having subjects report activity at the end of day. However, the results of this approach suffer from bias, inaccuracy and discrepancies. With advances in machine learning, automated assessment of activities has been widely adopted, even in clinical interventions [5, 6]. New devices with small form-factors like the FitBit, Jawbone, etc. in conjunction with advanced machine learning techniques can provide effective and scalable solutions.

However, developing recognition algorithms purely from classical statistical learning frameworks has limitations for clinical interventions. Challenges include: a) a much higher variability in execution of actions in clinically relevant populations, both within the same subject and across subjects, compared to healthy subjects, b) the activities of interest in clinical studies often show similarities in feature-space (e.g. walking at 1mph and walking at 4mph), c) signals may be temporally mis-aligned due to variability in sampling rates, and temporal windowing operations as well.

Due to these issues, one needs to more explicitly model temporal variability, while also leveraging popular machine learning frameworks. We argue that *comparing, summarizing, and modeling* a set of activity based time-series data are needed for invariance to temporal mis-alignment which can distort features and distance measures. The most common way to deal with the temporal mis-alignment problem is to use dynamic time warping (DTW) which was originally introduced in the speech processing community [4]. Despite its usefulness in a wide range of fields, in general, DTW has limitations: it is not a true distance metric, and does not naturally allow the estimation of statistical measures such as mean and variance of time-series. It also might not provide smooth differentiable time-warping needed in various applications. Therefore, in this work, we adopt more recent methods from the field of functional analysis [22] that counter some of these limitations.

Temporal variability and misalignment in activity recognition: Factoring out temporal variability is important for classifying activity signals. In addition, signals may be mis-aligned due to variability in sampling rates, and windowing operations as well. There is an inherent non-linearity of the domain of analysis, where temporal variability of activities can distort traditional features and met-

rics between time-series. Simple methods such as linear time-scaling do not work well in practice. Sensor noise also makes feature extraction difficult. To deal with these issues, we adopt a recent framework for signal alignment based on square-root velocity function (SRVF) [20]; this framework uses the Fisher-Rao Riemmannian metric to derive a proper distance on the quotient space of functions modulo the time-warping action. Exploiting this framework has several desirable properties, including a rate-invariant metric and vector-space representations.

For the proposed application of interest, using the SRVF approach, the raw time-series can be decomposed into a composition of time-warping functions and aligned time-series signals, as illustrated in Figure 1. The rest of the paper builds on this basic framework, and develops a complete algorithmic pipeline that integrates the strengths of SRVF with classical machine learning approaches, delivering improved performance.

Contributions: We develop a complete pipeline for activity recognition that integrates results from functional analysis with robust statistical learning approaches. Our approach results in the following advantages: (1) improvements in activity recognition, with maximum improvements in separating activities that differ in temporal rates (walking at 1mph vs 4mph) (2) increased robustness to window lengths compared to classical methods.

1.1. Related Work

Clinical Application: A wide range of classical methods for activity recognition have been developed, typically using wearable sensors on specific body parts. For instance, data from a sensor on the hip, worn over a 7-day period was used to classify various activities of daily living [15]. A recent trend, however, is to move the activity monitor from the hip to the wrist location to increase wear-time compliance and capture sleep-related behavior. Current interest in clinical studies is aimed at 24-hour monitoring and automatic recording of activities while being low-cost and wearable.

Classical Methods: Classical approaches to wearable-based activity analysis include extracting features from overlapping windows in time, such as wavelets, Fourier features etc, c.f. [2, 14, 7]. This is then followed by standard machine learning approaches for classification. Classical approaches do not provide special consideration to the issue of temporal mis-alignment.

DTW-based time-series analysis: In order to account for temporal misalignment, the DTW-based method has been used to compare and measure the distance between a given template and observed signal. By means of the DTW-based template comparison, the template database is selected from individual training samples and continuously updated to capture the particular activities performed (c.f. [17]). After computing and sorting the DTW scores for all templates, the set of the smallest scores is compared to the activity-specific thresholds. In such conventional methods, DTW is used only to measure similarity between the testing data and template database, without additional statistical modeling capabilities. Additionally, classical DTW has inherent limitations in scaling to large training-sets. Using a small fraction of training samples as templates is also not optimal because they may not well represent their classes. Thus, we seek a generic framework which can be applied where the requirement is to *compare, summarize, and model* a set of activity based time-series data. In this study, we exploit SRVF to compute Fréchet means of given time-series signals. It provides well-defined mathematical properties where distance and statistics are straightforward to compute.

Functional Analysis: In this work, we adopt recent functional analysis methods based on the SRVF representation that would allow dealing with phase variability or elastic temporal misalignment, by separating phase and amplitude components. The space of analysis here becomes the space of diffeomorphism that map the unit interval $[0, 1]$ to $[0, 1]$. The SRVF method allows development of proper Riemannian metrics over time-series. It also allows for developing statistical summaries, and avoids some limitations of DTW-based methods such as the 'pinching' effect [16]. More broadly, exploiting differential geometry of the constrained space leads to more efficient solutions, as well as provides faster and stable solutions based on the geometry of the underlying manifolds with nice mathematical structure. For these reasons, there has been an increased interest in understanding differential geometric properties that can be applied on a variety of fields, including activity recognition [1, 24], movement quality assessment [18], medical image analysis [10], and shape analysis [19]. In particular, functional analysis from differential geometry and statistics has been exploited for shape-based activity recognition with high performance in [21, 13]. Furthermore, the proposed approach can serve as a pre-processing step for machine learning classifiers like deep neural networks [12, 9].

2. Proposed approach for action recognition

A brief overview of the proposed approach for action recognition is shown in Figure 2. During the training phase, we compute energy signals from original x, y, z signals, and convert energy signals into SRVF representation to compute the Fréchet mean for each class. Next, the original x, y, z signals can be aligned by a composition with the estimated warping functions as: $[\tilde{x}, \tilde{y}, \tilde{z}] = [x, y, z] \circ \gamma$. Consequently, a new training data set is created from which features can be extracted in a sliding window procedure (non-overlapping) with different window lengths. During the testing phase, we align test signal to the Fréchet means for all classes, and choose the warping function that results in the best fit across

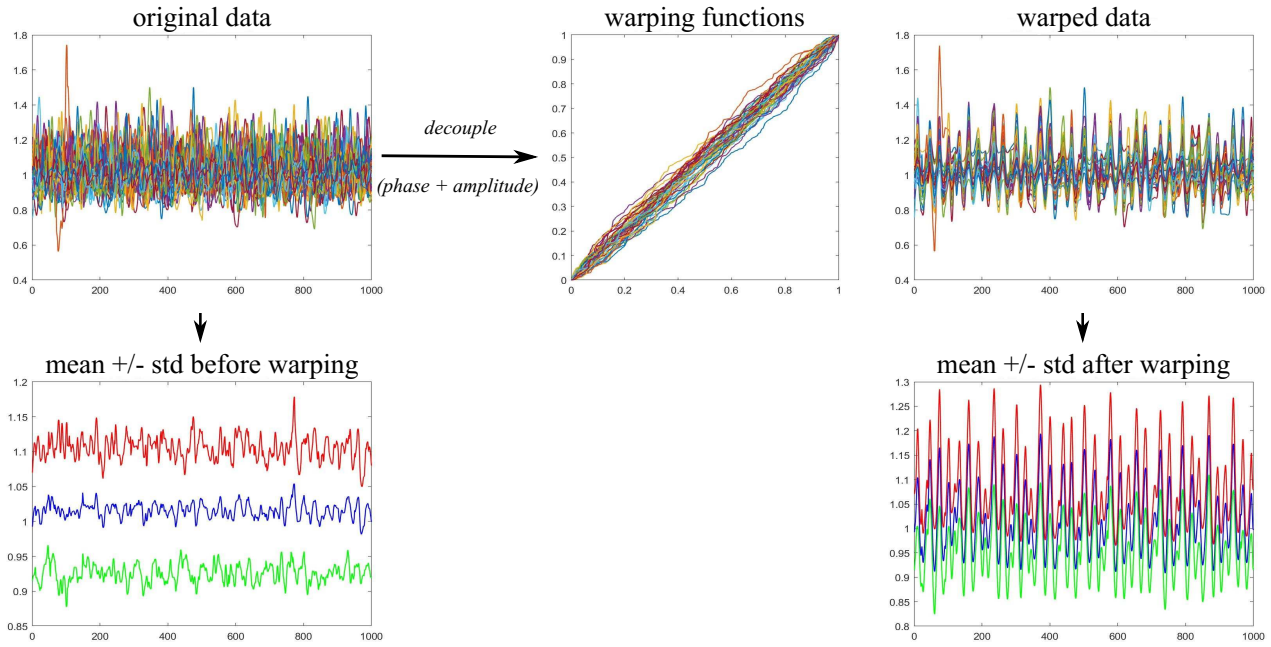


Figure 1. Here we illustrate how temporal variability and misalignment issues can significantly affect activity data. We show several instances of walking on a treadmill at 1mph, using the energy of accelerometer signals with a 1000 sample window. The SRVF framework allows decomposing the raw data into a composition of warping functions and aligned signals. In the bottom row, we illustrate the qualitative difference in computing an average signal and standard deviation signals for the raw signal set and the aligned signal set. Note that the summary signal statistics on after alignment show much more meaningful pattern, whereas the signal summaries before alignment have lost all patterns of the walking activity.

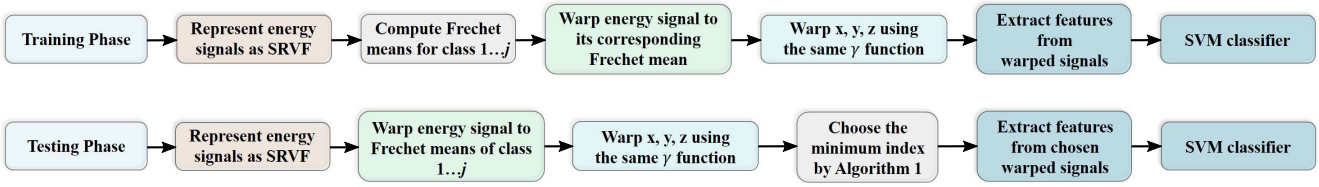


Figure 2. Illustration of the proposed framework's pipeline: The top row corresponds to the training phase and the bottom row corresponds to the test phase. The raw energy signals can be represented as SRVF representation to compute Fréchet means for each class. This allows original time-series, x, y, z, to be aligned using the same warping function. The new warped signals are then created to leverage SVM classifier by extracting features from them.

classes. Using the best-fit warping function, we align the test signals for x, y, z. Note that for final classification, we use a separate support vector machine (SVM) classifier that operates on a set of classical features extracted from this aligned test signal.

We describe the testing phase in more details in Algorithm 1. While testing, we represent a test signal and each Fréchet mean as SRVF representations (denoted by q, u respectively) to find the best warping function, and test signals are then aligned with this warping function. Here, the warping function can be computed by solving $\gamma^* = \text{arg}min_{\gamma \in \Gamma} ||u - (q \circ \gamma)\sqrt{\dot{\gamma}}||$ where u is the Fréchet mean which is obtained from training phase, Γ is referred to as the warping group, and q is the SRVF representations of given functions. The solution to this optimization comes from a dynamic programming algorithm. Through this step, we can perform warping of test signals by a composition: $[\tilde{x}, \tilde{y}, \tilde{z}] = [x, y, z] \circ \gamma^*$. Next, we calculate the Euclidean distance to measure the similarities between the Fréchet means of all classes, $[\bar{x}_j, \bar{y}_j, \bar{z}_j]$, and warped test signals, $[\tilde{x}, \tilde{y}, \tilde{z}]$, where j is the class index. The distance values from the Euclidean distance indicate how test signals can be warped with best-fit warping function and we denote minimum index by *Prediction_label* in Algorithm 1. Note, the final classification results could be affected based on prediction label. Finally, we can extract features from

the warped signals corresponding with the minimum class index. We describe each of the major steps involved in the pipeline in more details next after first briefly describing the dataset used.

Algorithm 1 Algorithm for computing warped signals during testing phase.

Input : Test signals x, y, z, E and the Fréchet means for all classes, $\bar{x}_j, \bar{y}_j, \bar{z}_j, \bar{E}_j$, where $j = 1, 2, ...,$number of classes
Output : Warped signals $\tilde{x}, \tilde{y}, \tilde{z}, \tilde{E}$.

1: **for** $j = 1, 2, ..., n$ **do**
2: $[\tilde{E}, \gamma_j]$ = SRVF warp$(\bar{E}_j, E)$
3: $[\tilde{x}, \tilde{y}, \tilde{z}] = [x, y, z] \circ \gamma_j$
4: Dist$[j]$ = Euclidean distance$([\bar{x}_j, \bar{y}_j, \bar{z}_j], [\tilde{x}, \tilde{y}, \tilde{z}])$
5: Temp data$[j]$ = $[\tilde{x}, \tilde{y}, \tilde{z}, \tilde{E}]$
6: **end**
7: Min_index = min(Dist)
8: Prediction_label = Min_index
9: Warped signals = Temp data[Min_index]

2.1. Data collection and activity classes

We collected data from subjects of male and female adults, in the age range 18-64 years. All of them were recruited by fliers, emails, or social networks. The consenting participants were requested to complete an online screener and visit a lab to determine eligibility. All procedure for this study were approved by the Institutional Review Board at Arizona State University.

The participants were requested to wear comfortable clothing and were fitted with a GENEActiv sensor (Activinsights Ltd., Kimbolton, UK) on their non-dominant wrist along with other activity monitors. The GENEActiv wrist-worn sensor provides us raw accelerometer data and advantages of being lightweight and waterproof. The subjects then performed a set of ambulatory and daily living activities from a predetermined pool of activities.

Table 1 lists the 7 activity classes collected from this study and the description of each activity. Note that the acceleration signals were first trimmed to remove the resting periods at the beginning and the end of each activity (if such periods existed).

Class	Activity
1	Walking (treadmill at 1mph)
2	Walking (treadmill at 4mph)
3	Running (treadmill at 6mph, 5% grade)
4	Seated/folding/stacking laundry
5	Standing/fidgeting with hand while talking
6	Hard surface walking holding filled coffee cup
7	Walking down stairs (5 floors)

Table 1. The activities of daily living of interest in this paper.

3. Detailed description of approach

3.1. Converting accelerometer data to features

Extracting features succinctly capture the important information embedded in the signal, while reducing the amount of resources needed to describe the data accurately. We note here that using the SRVF framework, it is actually possible to build a full Bayesian model on the product space of phase and amplitude functions, in a manner similar to the one used by Veeraraghavan et. al. [23]. In this paper, however, we use the framework only as a pre-processing step, so as not to preclude the use of any end-classifier one wishes to use. This choice ensures that one can specifically evaluate the impact of pre-alignment on final classification performance, while conforming to popular classification pipelines.

A variety of methods have been used to extract features from accelerometer data – mean, variance, RMS value of signal, Pearson correlation coefficients, frequency domain entropy features, and power spectral density (PSD) from FFT values (e.g. [25]). In this work, complementary to the triaxial raw accelerometer data, an additional time-series, called energy signal C_E, is obtained by computing the magnitude of the acceleration vector: $C_E = \sqrt{C_x^2 + C_y^2 + C_z^2}$. We also extend the feature space by extracting features from energy signals. Finally, a 37-dimensional feature vector is used and calculated for each window for SVM training and prediction. The final features include:

1) Mean, variance, and root mean square values of the acceleration values on each of C_x, C_y, C_z and C_E respectively.
2) Pearson correlation coefficients between all combinations of C_x, C_y, C_z and C_E, frequency domain entropy, and the PSD from FFT values.
3) Difference between maximum and minimum accelerations on each axis (dx, dy, dz, dE), and the following 5 features [11]:

$$\begin{aligned} dC_{x,y} &= \sqrt{dx^2 + dy^2}, \\ dC_{y,z} &= \sqrt{dy^2 + dz^2}, \\ dC_{x,z} &= \sqrt{dx^2 + dz^2}, \\ dC_{x,y,z} &= \sqrt{dx^2 + dy^2 + dz^2}, \\ dC_{x,y,z,E} &= \sqrt{dx^2 + dy^2 + dz^2 + dE^2}. \end{aligned} \tag{1}$$

Classifier Validation : In this work, we extracted features from non-overlapping windowing. In particular, we used different window length sizes to evaluate robustness to

window lengths. We employed leave-one-subject-out cross-validation method, i.e., we trained the classifier using data from 49 subjects, and tested on the remaining subject. This process was repeated for all 50 subjects and results were averaged.

3.2. Pre-alignment with SRVF framework

In this section we describe the SRVF representation, which was developed in a series of papers [20, 22] as a novel geometric framework for separating the phase and amplitude variability in given functional data. One important reason for using SRVF representation is to transform the Fisher-Rao metric into the standard $\mathbb{L}^2$ metric, simplifying the computations and thus the Fisher-Rao distance between two functions becomes the standard $\mathbb{L}^2$ between their SRVFs. This metric could help define a Fréchet mean template, among other summaries. For completeness of this paper, we reiterate the essence of the SRVF representation below.

We denote a real-valued function on the interval $[0, 1]$ denote by f, and let $\mathcal{F}$ denote the set of all such functions, we then define the mapping: $Q : \mathbb{R} \to \mathbb{R}$ according to: $Q(x) \equiv \begin{cases} x/\sqrt{|x|}, & \text{if } |x| \neq 0 \\ 0 & \text{otherwise} \end{cases}$. Note that for further analysis of function f, we redefine it as a SRVF representation satisfying the following formulation: $q : [0,1] \to \mathbb{R}$, where $q(t) \equiv Q(\dot{f}(t)) = \dot{f}(t)/\sqrt{|\dot{f}(t)|}$. In fact, if the function f is absolutely continuous, then the resulting SRVF is square-integrable. Thus, the set of all SRVFs are given by $\mathbb{L}^2([0, 1], \mathbb{R})$ (or simply $\mathbb{L}^2$) and we can define that the SRVF of $f \circ \gamma$, which means that warp a function f by γ is given by [22]:

$$\tilde{q}(t) = \frac{\frac{d}{dt}(f \circ \gamma)(t)}{\sqrt{|\frac{d}{dt}(f \circ \gamma)(t)|}} = (q \circ \gamma)(t)\sqrt{\dot{\gamma}(t)}. \qquad (2)$$

For convenience, we will denote this transformation by $(q, \gamma) = (q \circ \gamma)\sqrt{\dot{\gamma}}$.

Elastic Riemannian metric and distance on quotient space : We use the nonparametric form of the Fisher-Rao metric for analyzing SRVFs. This metric provides several fundamental advantages, including invariant property to domain warping.

The next step is to define an elastic distance between functions. Let $[q] = \text{closure}\{(q, \gamma)|\gamma \in \Gamma\} = \text{closure}\{(q \circ \gamma)\sqrt{\dot{\gamma}}|\gamma \in \Gamma\}$ denote the orbit of an SRVF $q \in \mathbb{L}^2$, that is, the set of SRVFs involved with all the warpings of a function, and their limit points. For any two orbits, the elastic distance d on the quotient space $\mathcal{S}$ is given by:

$$d([q_1], [q_2]) = \inf_{\gamma \in \Gamma} ||q_1 - (q_2, \gamma)||. \qquad (3)$$

The solution to this minimization problem over Γ is solved by using dynamic programming. This distance d is a proper distance (i.e. with the properties of non-negativity, symmetry, and the triangle inequality) on $\mathcal{S}$.

In this work, we compute the Fréchet mean of energy signals for each class. The Fréchet means are used as representative templates with which testing signals are compared. To compute the Fréchet means, one needs to apply warping and averaging iteratively. For further details of computing Fréchet mean, we refer the reader to [22].

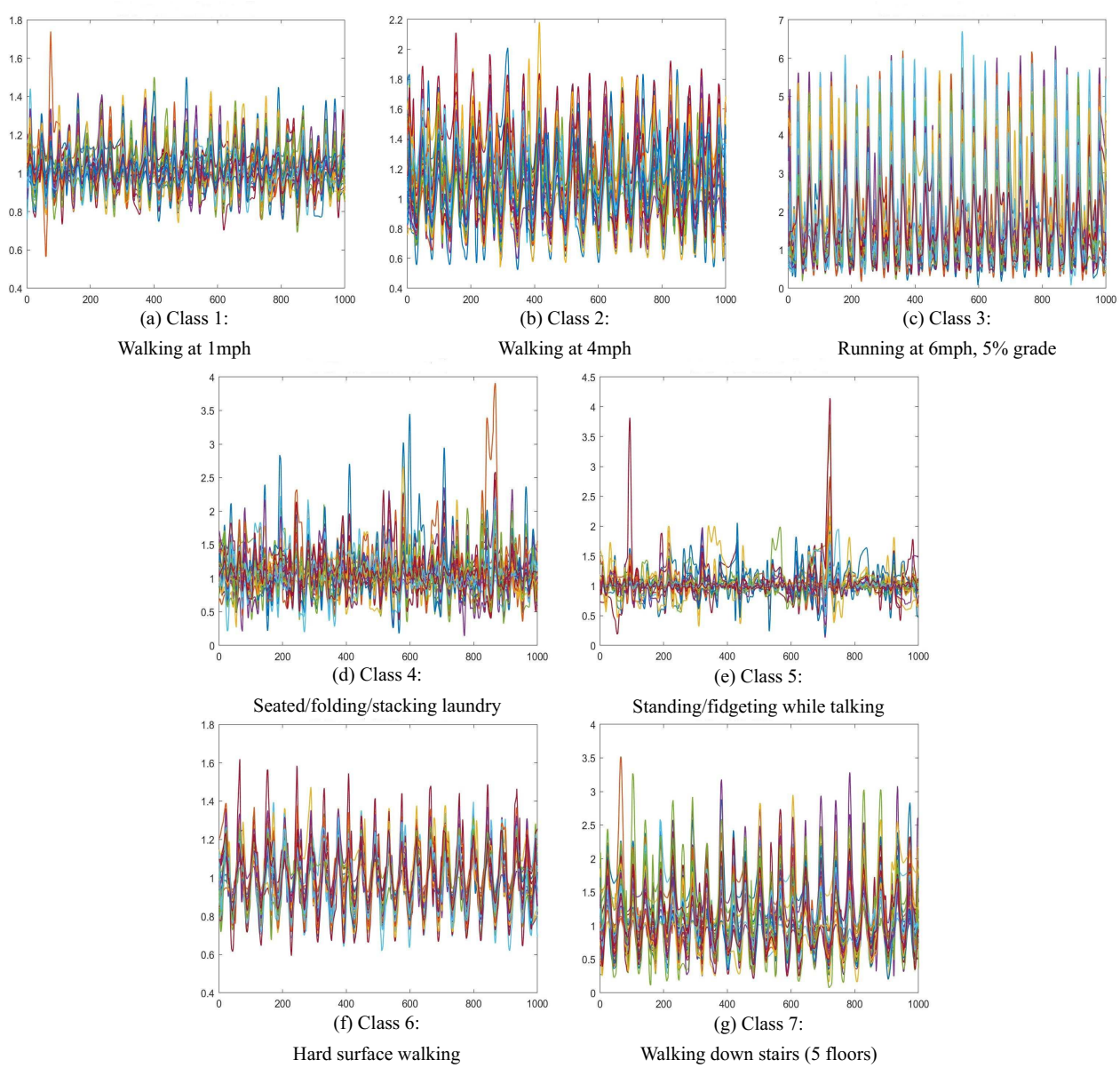
3.3. Post-alignment activity classification

After a given test signal is first aligned to the best-fit mean energy signal from all classes, the warping function thus estimated is used to align the x, y, z time-series as well, and then we extract the features described in Section 3.1 from the aligned signal. After this step, a trained SVM classifier is used for final classification. We generally observed that the highest classification rates were achieved using SVM classifiers with a radial basis function kernel. SVM classifiers are desirable because the optimization criteria are convex, which implies that a global optimal solution exists, and many toolboxes exist that simplify application of the algorithms. In this study, the SVM implementation from the LibSVM toolbox [8] was used.

4. Experiments

We evaluate our proposed approach on data captured with a wrist-worn accelerometer, the details of which are described in Section 2.1. The 7 activity classes with 50 subjects listed in Table 1 are investigated in this study. First, we show in Figure 3, how temporal pre-alignment can reveal significant patterns of quasi-periodicities across all classes. Without pre-alignment, none of this structure is revealed, and any further statistical summary computed from unaligned data simply loses all signal structure.

Comparison with and without SRVF: We tabulate and compare the activity recognition performance without and with SRVF, referred to as – Original Data (SVM) Baseline and Warped Data (SVM) SRVF respectively, as shown in Table 2. We observe that using the proposed pipeline that integrates the SRVF framework with classical machine learning approaches can deliver improved performance for all window lengths. Our approach combines both approaches, aiming to preserve feature information in feature domain, as well as reducing variability due to temporal misalignment.



(a) Class 1: Walking at 1mph

(b) Class 2: Walking at 4mph

(c) Class 3: Running at 6mph, 5% grade

(d) Class 4: Seated/folding/stacking laundry

(e) Class 5: Standing/fidgeting while talking

(f) Class 6: Hard surface walking

(g) Class 7: Walking down stairs (5 floors)

Figure 3. These figures show how temporal alignment can reveal patterns of activity across all classes, whereas the signal summaries without any alignment would have lost all meaningful patterns of activity, as illustrated in Figure 1 for one class. Note that the length of signals for each class were 1000 samples.

Approach	**Window length**					**Avg. testing time**
	1000	**500**	**200**	**100**	**50**	**(for window length =** 500)
Original Data(SVM) Baseline	86.29%	85.86%	84.74%	79.54%	73.81%	$\approx 1.3s$
Warped Data(SVM) SRVF	**89.43**	**87.86%**	**87.14%**	**85.74%**	**80.71%**	$\approx 22.1s$

Table 2. Comparison of proposed approach (bottom) with classical method (top). We observe that using the proposed approach can deliver improved performance for all window lengths, demonstrating increasing robustness to variability in window lengths. The recognition accuracy (%) is reported for activities using leave-one-subject-out validation over 50 subjects.

Increased separation of similar classes: As another advantage, we show that the activities that cause significant confusion are much better classified using the proposed ap-

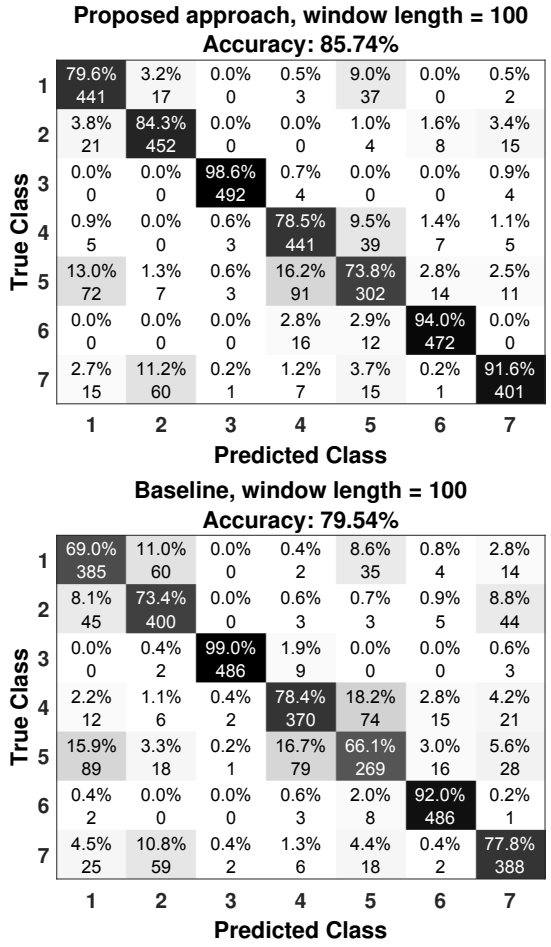


Figure 4. Comparison of confusion matrices with precision rates of proposed approach (top) and baseline (bottom) with **window length 100**. Please note that the confusion between the first two classes – which differ in temporal rates – is reduced using the proposed approach.

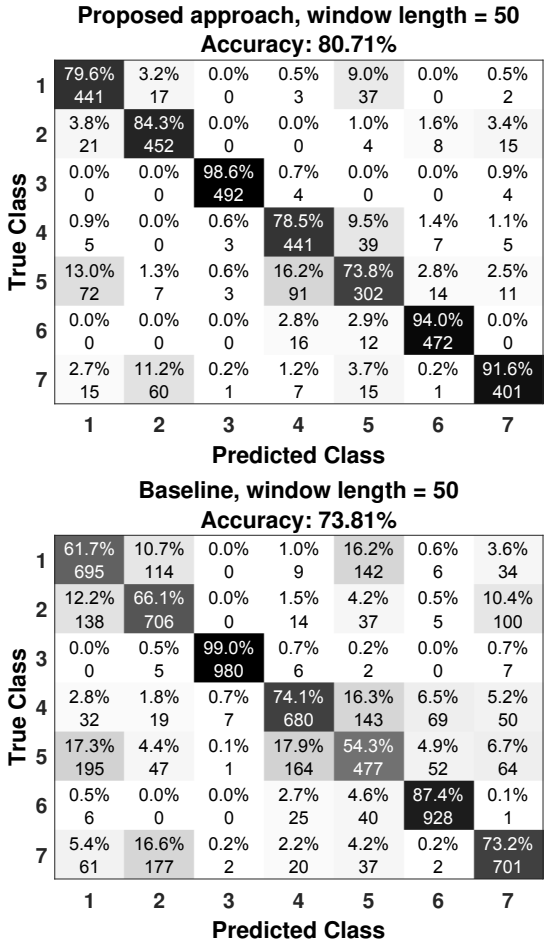


Figure 5. Comparison of confusion matrices with precision rates of proposed approach (top) and baseline (bottom) with **window length 50**. Please note that the confusion between the first two classes – which differ in temporal rates – is reduced using the proposed approach.

proach. In our set of activity classes, note that the first two differ in temporal rates– walking on a treadmill at 1mph and 4mph. The confusion matrices obtained with different window lengths are shown in Figure 4 and 5. As seen in both figures, our approach results in increased ability to distinguish between them.

One of the reasons for the improvement in recognition accuracy is likely due to the improved frequency domain entropy feature. For example, walking activities in the same-class carried out at varying rates may have different distribution of FFT components in the frequency domain. After temporal alignment, the spectra of same-class signals may become more similar with each other, and as a result, the frequency domain features become more concentrated within the same-class, leading to better classification result. We computed frequency domain entropy and compared the values before and after alignment. The results, shown in Figure 6 and Table 3, supported our assumption. In Figure 6, both plots show the variance of the feature within each class. The first box-plot shows the variance before alignment, whereas the second box-plot corresponds to after alignment. The second box-plot shows that the standard deviations of the features are smaller, compared to the first box-plot. We also demonstrate the quantitative comparisons of $(mean \pm std)$ in Table 3. We observe that the standard deviations are reduced for all classes.

For further visualization of impact of alignment, we illustrate the overall performance improvement for individual classes across a range of window lengths in Figure 7. We note that the higher improvements in recognition accuracy are obtained for classes 1, 2, and 7.

Analysis of testing time: We also conducted an analysis of the increased computational load for the proposed

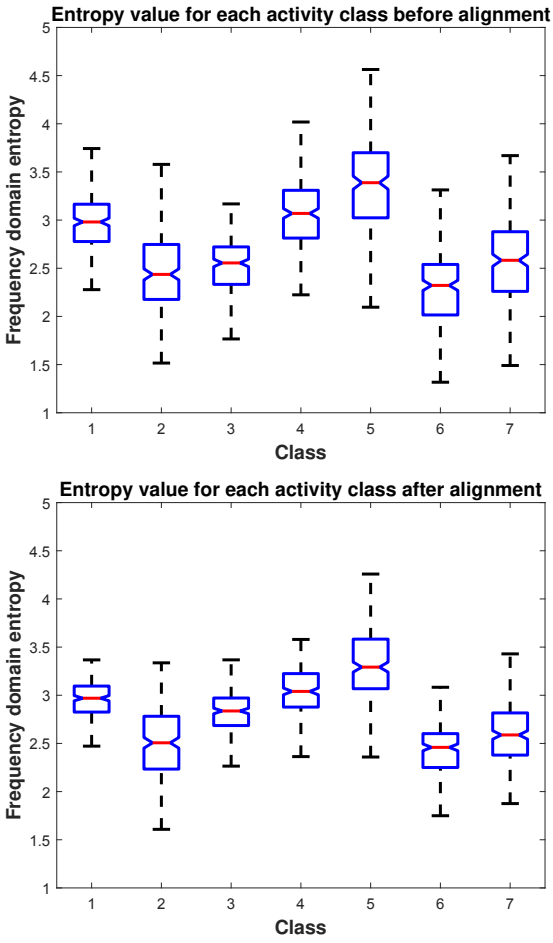


Figure 6. Frequency domain entropy feature values before and after alignment. We observe reduction of variance in this feature, across all classes, while not altering the separation of dominant modes in feature space. Numerical details appear in table 3.

methods. We report average testing time per test-window in Table 2. We note that the implementation was based in MATLAB 2017a, on a Windows desktop, with an Intel Core i7-2600 CPU(3.40 GHz). We observe there is an increase in computation time, as expected. This implies that the proposed method, even with unoptimized code, is at least suited for offline analysis applications. But with optimization of the pipeline, it is conceivable that the approach could work in real-time or near real-time situations. For reducing testing time, the fast DP algorithm that runs in $O(N)$ time can be applied with N nodes [3].

5. Conclusion

In this paper, we present a fusion of classical methods with modern functional analysis method for elastic temporal alignment, for activity recognition from accelerometer data. We find an overall improvement in recognition performance on a dataset of 7 activities with 50 subjects. The

Class No.	Before alignment $(mean \pm std)$	After alignment $(mean \pm std)$
1	$2.96 \pm$ **0.296**	$2.95 \pm$ **0.215**
2	$2.46 \pm$ **0.406**	$2.50 \pm$ **0.370**
3	$2.49 \pm$ **0.338**	$2.82 \pm$ **0.242**
4	$3.07 \pm$ **0.374**	$3.02 \pm$ **0.278**
5	$3.37 \pm$ **0.514**	$3.30 \pm$ **0.409**
6	$2.28 \pm$ **0.423**	$2.42 \pm$ **0.317**
7	$2.58 \pm$ **0.492**	$2.62 \pm$ **0.363**

Table 3. Comparison of mean and standard deviation of frequency-domain entropy feature before and after alignment. The means and standard deviations were calculated for all subjects with **window length 200**.

Figure 7. The performance improvements for all classes with respect to different window lengths – 500, 200, 100, 50.

improvement benefits are due to: (1) superior discrimination between activities that differ in temporal rates (2) increased robustness to window lengths. An extension of this work would be moving toward new representation spaces, afforded by new kinds of sensors. For example, time-series of rotation-matrices, which may be obtained by sensor-fusion from accelerometer, gyros, and magnetometers, may be treated as Riemannian manifold-valued signals. Extensions of functional methods to manifold-valued functions can be adopted for such new kinds of data. Further, optimization and/or modification of the proposed approach that results in close to real-time classification is another avenue of work.

Acknowledgements

The work of HC, QW, PT was supported by NSF grants 1452163 and 1617999. The work of AS was supported by NSF grant 1617397.

References

[1] R. Anirudh, P. Turaga, J. Su, and A. Srivastava. Elastic functional coding of Riemannian trajectories. *IEEE transactions on Pattern Analysis and Machine Intelligence*, 39(5):922–936, 2017. 2

[2] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data. In *International Conference on Pervasive Computing*, pages 1–17. Springer, 2004. 1, 2

[3] J. Bernal, G. Dogan, and C. R. Hagwood. Fast dynamic programming for elastic registration of curves. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 111–118, 2016. 8

[4] D. J. Berndt and J. Clifford. Using dynamic time warping to find patterns in time series. In *Workshop on Knowledge Discovery in Databases*, volume 10, pages 359–370. Seattle, WA, 1994. 1

[5] M. P. Buman, E. B. Hekler, W. L. Haskell, L. Pruitt, T. L. Conway, K. L. Cain, J. F. Sallis, B. E. Saelens, L. D. Frank, and A. C. King. Objective light-intensity physical activity associations with rated health in older adults. *American Journal of Epidemiology*, 172(10):1155–1165, 2010. 1

[6] M. P. Buman, S. L. Mullane, M. J. Toledo, S. A. Rydell, G. A. Gaesser, N. C. Crespo, P. Hannan, L. Feltes, B. Vuong, and M. A. Pereira. An intervention to reduce sitting and increase light-intensity physical activity at work: design and rationale of the stand & move at workgroup randomized trial. *Contemporary clinical trials*, 53:11–19, 2017. 1

[7] P. Casale, O. Pujol, and P. Radeva. Human activity recognition from accelerometer data using a wearable device. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 289–296. Springer, 2011. 2

[8] C.-C. Chang and C.-J. Lin. LIBSVM: A library for support vector machines. 2001. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm. 5

[9] Y. Chen and Y. Xue. A deep learning approach to human activity recognition based on single accelerometer. In *IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 1488–1492, 2015. 2

[10] P. T. Fletcher, C. Lu, S. M. Pizer, and S. Joshi. Principal geodesic analysis for the study of nonlinear statistics of shape. *IEEE Transactions on Medical Imaging*, 23(8):995–1005, 2004. 2

[11] P. Gupta and T. Dallas. Feature selection and activity recognition system using a single triaxial accelerometer. *IEEE Transactions on Biomedical Engineering*, 61(6):1780–1786, 2014. 4

[12] N. Y. Hammerla, S. Halloran, and T. Ploetz. Deep, convolutional, and recurrent models for human activity recognition using wearables. *arXiv preprint arXiv:1604.08880*, 2016. 2

[13] S. Kurtek, E. Klassen, J. C. Gore, Z. Ding, and A. Srivastava. Elastic geodesic paths in shape space of parameterized surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9):1717–1730, 2012. 2

[14] J. Lester, T. Choudhury, and G. Borriello. A practical approach to recognizing physical activities. In *International Conference on Pervasive Computing*, pages 1–16, 2006. 1, 2

[15] A. Mannini, M. Rosenberger, W. Haskell, A. Sabatini, and S. Intille. Activity recognition in youth using single accelerometer placed at wrist or ankle. *Medicine and Science in Sports and Exercise*, 49(4):801, 2017. 2

[16] J. S. Marron, J. O. Ramsay, L. M. Sangalli, A. Srivastava, et al. Functional data analysis of amplitude and phase variation. *Statistical Science*, 30(4):468–484, 2015. 2

[17] C. Pham, T. Plötz, and P. Olivier. A dynamic time warping approach to real-time activity recognition for food preparation. In *International Joint Conference on Ambient Intelligence*, pages 21–30. Springer, 2010. 2

[18] A. Som, R. Anirudh, Q. Wang, and P. Turaga. Riemannian geometric approaches for measuring movement quality. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 43–50, 2016. 2

[19] A. Srivastava, S. H. Joshi, W. Mio, and X. Liu. Statistical shape analysis: Clustering, learning, and testing. *IEEE Transactions on pattern analysis and machine intelligence*, 27(4):590–602, 2005. 2

[20] A. Srivastava, E. Klassen, S. H. Joshi, and I. H. Jermyn. Shape analysis of elastic curves in euclidean spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(7):1415–1428, July 2011. 2, 5

[21] A. Srivastava, P. Turaga, and S. Kurtek. On advances in differential-geometric approaches for 2d and 3d shape analyses and activity recognition. *Image and Vision Computing*, 30(6-7):398–416, 2012. 2

[22] A. Srivastava, W. Wu, S. Kurtek, E. Klassen, and J. Marron. Registration of functional data using fisher-rao metric. *arXiv preprint arXiv:1103.3817*, 2011. 1, 5

[23] A. Veeraraghavan, A. Srivastava, A. K. Roy-Chowdhury, and R. Chellappa. Rate-invariant recognition of humans and their activities. *IEEE Trans. Image Processing*, 18(6):1326–1339, 2009. 4

[24] R. Vemulapalli and R. Chellapa. Rolling rotations for recognizing human actions from 3d skeletal data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4471–4479, 2016. 2

[25] Q. Wang, S. Lohit, M. J. Toledo, M. P. Buman, and P. Turaga. A statistical estimation framework for energy expenditure of physical activities from a wrist-worn accelerometer. In *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the*, pages 2631–2635. IEEE, 2016. 4