34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Universal Hybrid Decision Tree Classifier Design for Human Activity Classification

Chieh Chien, and Gregory J. Pottie, *Fellow, IEEE*

***Abstract*— A system that reliably classifies daily life activities can contribute to more effective and economical treatments for patients with chronic conditions or undergoing rehabilitative therapy. We propose a universal hybrid decision tree classifier for this purpose. The tree classifier can flexibly implement different decision rules at its internal nodes, and can be adapted from a population-based model when supplemented by training data for individuals. The system was tested using seven subjects each monitored by 14 triaxial accelerometers. Each subject performed fourteen different activities typical of daily life. Using leave-one-out cross validation, our decision tree produced average classification accuracies of 89.9%. In contrast, the MATLAB personalized tree classifiers using Gini's diversity index as the split criterion followed by optimally tuning the thresholds for each subject yielded 69.2%.**

I. Introduction

Home monitoring of activities can both improve medical care and reduce costs, through feedback to both individuals and health care providers. Therefore, increased research effort is going into the creation of systems that record human motions with feasible cost, classify activities with good accuracy, and then analyze these activities with respect to different rules [1] [2] [3] [4].

Some systems [1] [2] have classified small numbers of daily activities using naïve Bayes classifiers with accuracy up to 90%. As the number of classes grows, using a single-stage classifier becomes problematic at many levels, not least in the large volume of training data required. Decision tree classifiers [3] [4] [5] [6] [7] better handle complex decision regions by partitioning them into smaller sets with low dimensional hypothesis spaces at each stage. Advantages include reduced training set size, robustness to outliers in training data, extensibility of target classes, and invariance under monotone transformations. Nonetheless, decision tree methodologies that apply a single classifier type at each node can suffer from mismatches between assumed and actual distributions for different sets of classes, resulting in unacceptable accuracy. Another issue is the generalizability of a model. In a clinical trial, due to very high logistical costs, one can acquire extensive ground truth only for a small set of subjects; for the rest, at best only short training is feasible. However, large classifier accuracy gains result when models are adapted to individuals. One approach is to create a decision tree structure that fits the population, and then tune only the decision thresholds using the short training sequences from individuals. This was attempted in [3], but with inadequate accuracy.

In this paper, we report a complete procedure for daily life activity classification, from data collection, feature extraction, tree structure and feature selection, to testing. The resulting classifier is generalizable and has high accuracy. We conclude with possible extensions of the work.

II. Data Collection and Feature Extraction

A. Data Collection

We used the Gulf Coast Data Concept USB Accelerometer X6-2mini with a built-in tri-axial accelerometer [8] to collect the data at the sample rate 160 Hz, resolution 16 bits, and gain ±6g. We put accelerometers on 14 parts of the body, as described in Table I. In the training and testing processes, each sensor collected x, y, and z directions of acceleration, thus producing in total 42 channels of data. Seven people took part in data collection, with 2 hours of measurement for each person. Each was asked to perform the series of common activities listed in Table II. An annotator followed the subject to label the activities using an Android program to log start and stop time of each activity. During the experiment, start and end markers were manually added.

B. Feature Extraction

After collecting the annotated data, we extracted features from the measured acceleration using a moving window of length 4 s and of step size 1 s. The window size of 4 s ensured that we captured more than a complete cycle for every activity, to enable similar features for each class. The step size of 1 s yields activity classification

978-1-4577-1787-1/12/$26.00 ©2012 IEEE

resolution of one prediction per second. In each window, thirty-one features for each accelerometer were calculated. These features include signal mean, variance, energy, max and min value, dominant frequency, etc.

TABLE I. SENSOR PLACEMENTS

Upper limb and head	Lower limb
Forehead	Left and right pockets
Chest	Left and right knees
Right and left elbows	Left and right ankles
Right and left wrists	Left and right toes

TABLE II. COLLECTED ACTIVITIES

Motion	Stationary
Walk slowly	Stand
Walk fast	Sit upright
Run	Sit while slouching
Walk up slope	Sit while hunching
Walk down slope	Lie on back
Walk upstairs	Lie on stomach
Walk downstairs	Lie on side

III. HYBRID TREE FORMATION

The proposed hybrid tree classifier T includes l internal nodes. An internal node of a tree is a node that is not a leaf node. T can be thought as a set of l single-stage classifiers, each with its subset of classes, features and the decision rules used for the node. One can write

$$T = \{C(t), F(t), D(t)\} \quad (1)$$

where C is the subset of classes of node t, indicating how to group classes in that node; F is the feature set used for node t; and D is the decision rule of node t. Forming a tree classifier consists of deciding upon $C(t)$, $F(t)$ and $D(t)$ for each internal node based on prior knowledge and observation of the training data.

We manually determined the tree based on our knowledge of the activities, and found feature sets and decision rules for each internal node. That is, we fix the set of classes C for all nodes and try to find the feature set F^* and decision rules D^* that minimize the overall classification error. In contrast to most designs of decision tree classifiers [9] [10] where only one type of decision rule is used, we used both the naïve Bayes classifier and support vector machine (SVM) at internal nodes. The naïve Bayes classifiers were used to classify motion activities, since after observing various features of the training data given each class, we found Gaussian distributions formed adequate descriptions for such changing features; SVMs were used to classify stationary activities, such as standing, lying, and sitting, since the training data are concentrated, and fewer outliers occur.

IV. UNIVERSAL HYBRID DECISION TREES

After creating a hybrid tree classifier for classifying various activities, we then tried to find a single tree that can classify multiple sets of testing data from many subjects. This is important since with this tree we can specialize the classifier to individuals with minimal additional training, therefore making the model more easily applied to the general public. Since machines do not easily learn how to generalize a model, automatically created trees are often too specific to the trained data. For each tested subject, we tuned the decision thresholds for each internal node, and the final classifier model was formed. Although this seems time-consuming, the generalizability of the classifier actually will produce a huge time saving as we collect more data, since annotation of accurate ground truth will dwarf all other human effort in large-scale studies.

We now present an algorithm for systematic design of the tree. Let TD_j be the training data for subject j, $j=1,\ldots,M$ (M=number of subjects). There are N manually structured trees, each with $l(i)$ internal nodes for i=1,2,…,N. Each tree T_i can be written as

$$T_i = \{F(t), D(t), C(t)\}$$
$$t = 1,2,\cdots,l, i = 1,2,\cdots,N \quad (2)$$

with $l(i)$ internal nodes. In every tree, the class subset for each node $C(t)$ is determined for every internal node. Let $TD_{j,t}$ be part of the training data TD_j whose classes that are involved in node t of tree T. $P_{e_{j,t}}\left(F(t), D(t), TD_{j,t}\right)$ is the probability of error of node t when applying feature set $F(t)$ and decision rule $D(t)$ on training data $TD_{j,t}$. The algorithm can be stated as follows:

Begin

1. Given a set of possible decision trees, randomly pick a tree T with l internal nodes.
2. For $t = 1$ to l

   Find the optimal set $\left(F^*(t), D^*(t)\right)$ that minimizes the weighted probability of error

$$\left(F^*(t), D^*(t)\right) = \arg \min_{(F(t),D(t))} \sum_{j=1}^{M} w_j \cdot P_{e_{j,t}}\left(F(t), D(t), \mathrm{TD}_{j,t}\right)$$

   where w_j is the weighting function for the subject j, indicating the weighting of that type of people to the general public.
3. If $\sum_{j=1}^{M} w_j \cdot P_{e_{j,t}}\left(F^*(t), D^*(t), TD_{j,t}\right) > th_{\mathrm{err}}$

Terminate the for loop, go to step 1 and try the next tree T, where th_{err} is the predefined error threshold

End If

End For

4. Output the tree classifier

$$T^* = \left\{F^*(t), D^*(t), C(t)\right\}, t = 1,2,\cdots,l$$

End Begin

This algorithm provides a means to find a compromise tree that accounts for the differences among people, while maintaining a satisfactory error rate. After creating this universal hybrid decision tree classifier, when there is a test subject with only small amount of training, we can then apply the tree classifier, include the tree structure, its separating features and decision rules, to the test subject. All that is changed is the decision threshold for each internal node. The threshold is determined specifically for each subject, while maintaining the decision tree structure.

V. Results

A. Method

Three different decision tree classifier mechanisms were used in this study: the custom universal hybrid decision tree, automatically generated trees for each subject, and automated trees with tuned thresholds for individuals. The classification results were calculated using leave-one-out cross-validation (LOOCV).

The custom decision tree (Figure 1) consisted of 27 nodes, where 13 were internal nodes with binary separation. We first manually determined the tree structure, and then used all data from all subjects except the testing one to determine the tree structure and features giving the highest accuracies. Afterward, for each subject, we determined decision thresholds for internal nodes of the tree using 40% of the data from the testing subject. Thresholds for each subject have to be determined individually since properties of each set of data are different from other sets. In this tree, we first separate motion activities from stationary activities. In upper part of the tree (motion activities), we used naïve Bayes classifiers on each branch, and assumed equal prior probabilities; in the lower part (stationary activities), we used SVMs with the Gaussian radial basis function kernel. For nodes using naïve Bayes classifiers, we selected two features that gave the highest weighted accuracy in separating classes; for nodes

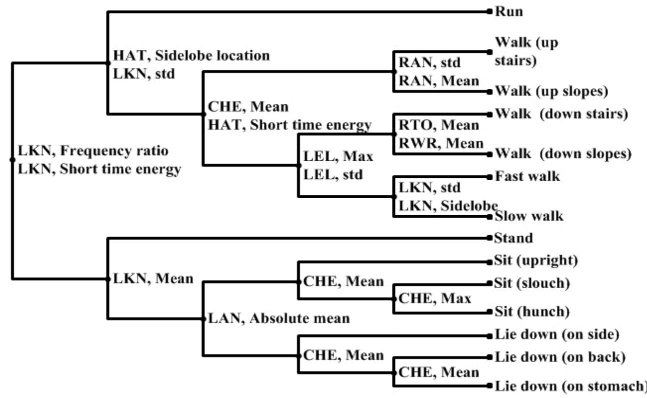

Figure 1. Custom universal hybrid decision tree

with SVM classifiers, we only selected one feature due to computational concerns.

Automatically generated decision trees were created using the MATLAB built-in function "classregtree." This function uses Gini's diversity index [11] as the separation criterion. It produces trees that are specific to the target training data so that each subject has a unique decision tree. The acquired data varied from person to person, even from different parts when we chopped it. On average the tree had 19.9 internal nodes ranging from 16 to 26, and on average 20.9 leaf nodes with a range from 16 to 27. In LOOCV, we found the automated decision trees including decision thresholds by using data of all except one subject, and then tested on the targeted subject.

To compare to the universal tree structure, we kept the same structure of automatically generated trees from the previous section and tuned their threshold values according to testing data. We called MATLAB to generate the automated decision tree for all subjects except the testing one, and then changed the thresholds of internal nodes of the tree using data of the testing subject. In LOOCV, we found the decision trees using all but one subject's data, then determined the thresholds using 40% of the last subject's data, and then tested on the remaining 60%.

B. Result

Table III shows the recall (true positive/(true positive + false positive)) and precision (true positive/(true positive + false negative)) for each class of the custom universal hybrid trees, automatically generated trees and automatically generated trees with tuned thresholds using LOOCV. The F-score, which is the harmonic mean of recall and precision, is also shown. The overall accuracies are 89.9%, 73.0%, and 69.2% for

TABLE III. SIMULATION RESULT

	Recall			Precision			F-measure		
Activity	***Custom tree***	***Auto tree***	***Auto tree with tuned threshold***	***Custom tree***	***Auto tree***	***Auto tree with tuned threshold***	***Custom tree***	***Auto tree***	***Auto tree with tuned threshold***
Walk slow	94.8%	60.9%	53.0%	83.5%	64.1%	72.7%	88.8%	62.4%	61.3%
Walk fast	70.2%	98.5%	32.8%	85.0%	53.3%	76.1%	76.9%	69.2%	45.8%
Walk down	76.8%	99.8%	50.8%	87.4%	44.5%	33.6%	81.8%	61.5%	40.4%
Walk up	72.6%	98.1%	75.0%	67.1%	56.1%	46.9%	69.7%	71.4%	57.7%
Stairs down	86.5%	99.8%	52.3%	92.1%	57.0%	57.9%	89.2%	72.5%	55.0%
Stairs up	67.3%	99.7%	62.2%	68.9%	66.2%	63.8%	68.1%	79.5%	63.0%
Run	98.6%	100.0%	99.5%	97.3%	94.9%	99.7%	98.0%	97.4%	99.6%
Stand	99.6%	100.0%	100.0%	89.2%	99.6%	99.3%	94.1%	99.8%	99.6%
Lie back	97.2%	100.0%	86.7%	87.3%	85.6%	85.5%	92.0%	92.2%	86.1%
Lie side	100.0%	100.0%	89.8%	100.0%	100.0%	100.0%	100.0%	100.0%	94.6%
Lie stomach	86.2%	99.5%	85.3%	97.2%	92.0%	89.2%	91.4%	95.6%	87.2%
Sit slouch	99.1%	100.0%	41.8%	99.5%	65.4%	60.7%	99.3%	79.1%	49.5%
Sit hunch	99.4%	100.0%	70.1%	99.5%	84.0%	73.9%	99.5%	91.3%	72.0%
Sit upright	99.1%	100.0%	61.9%	99.8%	54.5%	41.5%	99.4%	70.5%	49.7%

custom generated trees and automatically generated trees with tuned thresholds respectively.

The static activities are more easily classified than motion activities. This is because for the motion activities there are many variations in activity performance, and some are difficult to quantify, e.g., speed for walking fast vs. slowly, or the slope of walking up and down ramps. For the custom universal hybrid decision tree vs. the automated generated tree, we found that generally the performance of the custom tree was better than the automatic tree, indicating the automatic model was overfit to the trained data, and failed to classify activities that are not included in the training dataset. But it is worth noticing that the F-score of walking up ramp and walking upstairs for the custom tree was lower than that of the automatic model. This was due to the variable slopes of ramps and thus collected data spread over a wide range in the feature space.

Since the automatic model is very subject specific, it is fine-tuned to the training data and thus it is hard to generalize. Even with a short period of training data for the target subject, with full knowledge of other sets of data, we still cannot apply a general model generated from the dataset and simply change decision thresholds for each internal node of the tree.

VI. CONCLUSION

The proposed universal hybrid tree structure provides flexibility at the expense of the use of intuition or domain knowledge in its construction. The effort is rewarded in relative ease of tuning it to new individuals with modest additional training. Similar approaches may thus prove attractive in large clinical studies. A number of future research directions are suggested. It is of interest how to reduce the effort involved in model construction when the number of classes becomes large, as is the minimal ground truth required to personalize models to new subjects. There is a tradeoff against the size and variety represented in the dataset used to construct the base model. With a large enough set, multiple models could be employed to represent different populations. One might in such cases be able to bootstrap from the best model and use classification decisions to then adjust decision thresholds, even without any ground truth.

REFERENCES

[1] L. Atallah, B. Lo, R. King, and G.-Z. Yang, "Sensor placement for activity detection using wearable accelerometers," in *Proc. Wearable*.

[2] Long, X., Yin, B., and Aarts, R.M. 2009. Single accelerometer-based daily physical activity classification. In *31st Annual International Conference of the IEEE EMBS*, 6107-6110.

[3] J. Pärkkä, M. Ermes, P. Korpipää, J. M¨antyjärvi, J. Peltola, and I. Korhonen, "Activity classification using realistic data from wearable sensors," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 119–128, Jan. 2006.

[4] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Proceedings 2nd Int. Conference on. Pervasive Computing*. Springer, 2004, pp. 1-17.

[5] G. Landerweerd, T. Timmers, E. Gelsema, M. Bins and M. Halic, "Binary tree versus single level tree classification of while blood cells," *Pattern Recognition* vol. 16, 571-577 (1983).

[6] X. Li and R. C. Dubes, "Tree classifier design with a Permutation statistic," *Pattern Recognition* vol. 19, 229-235 (1986).

[7] AS.R. Safavian and D. Landgrebe, "A survey of decision tree classifier methodology." *IEEE Trans. Systems Man Cybernet*. 21 (1991), pp. 660–674.

[8] http://www.gcdataconcepts.com/

[9] Breiman, L., J. Friedman, R. Olshen, and C. Stone. *Classification and Regression Trees*. Boca Raton, FL: CRC Press, (1984).

[10] De'Ath, G. & Fabricius, K.E. (2000) Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology*, 81, 3178–3192.

[11] L. Breiman, J.H. Friedman, R.A. Olshen, C.J. Stone, Classificationm and Regression Trees, Chapman & Hall, 1984.