INTERNATIONAL JOURNAL OF RESEARCHES IN BIOSCIENCES, AGRICULTURE AND TECHNOLOGY
© VISHWASHANTI MULTIPURPOSE SOCIETY (Global Peace Multipurpose Society) R. No. MH-659/13(N)
www.vmsindia.org

TRAFFIC DENSITY STATE ESTIMATION USING MULTIPLE CONTIGUOUS FEATURE VECTOR FRAMES

Prashant Borkar[1], M. V. Sarode[2] and L. G. Malik[1]
[1]GHRCE, Department of CSE, Nagpur, India
[2]JCOET, Department of CSE, Yawatmal, India

Abstract:
To estimate vehicular traffic density state as low, medium and heavy, cumulative vehicular acoustic signal is collected from roadside installed Omni-directional microphone followed by acoustic feature extraction using Mel Frequency Cepstral Coefficients (MFCC) for varying combination of frame size and shift size. Classification is performed using ANFC and further performance improved using feature selection (FS) where linguistic hedges are employed for FS. Consideration of multiple contiguous frames will definitely increase the accuracy but with cost of computational time and to reduce this computational time for medium to large scale datasets, SSCG is employed in this research work. In SSCG gradient estimation decreases the training time; however, it does not cause any performance degradation for the classifier.
Keywords: Traffic density, acoustic, neuro-fuzzy, scaled conjugate gradient

Introduction:

Congestion due to traffic has significant impact on both environment and economy. This results in increasing usage of vehicles worldwide and vehicles CO_2 emission is becoming the main reason for global warming (Li and Shimamoto, 2011). Reducing such congestion can significantly improve traffic flow, reduces travel time and the environmental impact. In developed countries, traffic is characterized by lane-driven. Use of intrusive techniques such as video cameras, inductive loop detectors, magnetic sensors, and speed guns are suggested and also proved to be efficient approach for density state estimation but the installation, operational, and maintenance cost of these sensors is very expensive. Therefore researchers have been developing several techniques, which have number of significant advantages and disadvantages relative to each other. Nonintrusive traffic-monitoring technologies based on Radar, Ultrasound, Laser, and Acoustic signals, present various characteristics in terms of robustness to changes in manufacture, installation, environmental conditions; safety regulation compliance, and so forth (Valcarce et al., 2004).

This research work considers a problem of vehicular traffic density state estimation, modality of adaptive neuro-fuzzy classifier is employed to perform the classification task. In this work, Roadside acquired vehicular noises has to undergo feature extraction mechanism using MFCC, it results in huge dimensionality of dataset if multiple contiguous feature vectors frames were considered for training of classifier. To reduce the computation time per iteration network based methods can be used and these methods are trained by gradient based algorithms. Heuristics and numerical are the two broad categories of these algorithms, some algorithms are based on gradient-descent methods and some algorithms are based on linear least-square methods. Scaled Conjugate Gradient (SCG) is one of the most popular second-order gradient supervised procedure (Moller, 1993). As Scaled Conjugate Gradient computes two first-order gradients for the parameters to determine the second-order information, therefore training time is higher than steepest-descent method. To increase the performance Speeding up Scaled Conjugate Gradient (SSCG) is proposed in (Cetis & Barkana, 2010). SSCG shortens the training time per iteration; however convergence rate is not affected. In case of vehicular density estimation using acoustics, the consideration of multiple feature vectors frames will certainly results in better classification performance but at the cost of training time per iteration. In such cases, Modality of Linguistic hedges for feature selection and then employing SSCG on selected feature dataset will be better and alterative solution.

VEHICULAR ACOUSTIC SIGNAL & TRAFFIC DENSITY STATE ESTIMATION:

A vehicular acoustic signal is occurrence and mixture of various signals such as tyre noise, engine noise, exhaust noise, noise due to mechanical effects (e.g., axle rotation, brake, and suspension), air-turbulence noise and the honks. The traffic density state condition at any location is depends upon the mixture weighting of spectral components of various vehicular acoustics.

Jien Kato proposed method for traffic density estimation based on recognition of temporal variations that appear on the power signals in accordance with vehicle passes through reference point. HMM is used for observation of local temporal variations over small periods of time, extracted by wavelet transformation. Experimental results show 166 detection of passage of vehicles out of 176 i.e. 94% but in case of occlusion this detection rate is lower to 78.9%. System only considers the overlapping of two vehicles each in opposite direction. This system requires a camera to mount on pathway which makes system costly; such system may not applicable for traffic jam situation as occlusion rate is high (Kato, 2005).

The detailed design of an acoustic sensing hardware prototype which has been deployed by the side of the road is presented (Sen et al., 2011). This unit samples and processes road noise to compute various metrics like amount of vehicular honks and vehicle speed distribution and sends the metrics to a remote server every alternate minute. Traffic density state as congested and free-flow is estimated.

EMPLOYING SSCG:

MFCC: Acoustic feature extraction is carried out with the help of Mel Frequency Cepstral Coefficients (MFCC). MFCC has proven to be one of the most successful feature representations in speech related recognition tasks. Mel-Frequency Cepstral Coefficients (MFCC), which are the Discrete Cosine Transform (DCT) coefficients of a Mel-filter smoothed logarithmic power spectrum. First ten to fifteen Cepstral coefficients of a signal's spectrum sufficiently capture the smooth spectral envelope information (Nakagawa et a;., 2012).

Feature Selection using Linguistic Hedges: The Adaptive Neuro-Fuzzy Classifier with LHs is based on fuzzy rules. Example, for two inputs{x_1, x_2} and one output y is defined with LHs as IF x_1 is A_1 with p_1 hedge AND x_2 is A_2 with p_2 hedge THEN y is C_1 class, where A_1 and A_2 denote linguistic terms that are defined on X_1 and X_2 feature space; p_1 and p_2 denote linguistic hedges, respectively; C_1 denotes the class label of the output y.

1 Describe single classification rule for every class.
2 Set $p_{ij} = 0.5$, for i = 1, 2,. . .,C and j = 1,2,. . .,F, where C is the number of classes and F is the number of features.
3 K be the selected features
4 Use LH to train neuro-fuzzy classifier. In training, $0 \leq p_{ij} \leq 1$.
5 For i = 1 to C. Find the j^th^ feature that satisfies the maximum p value for the i^th^ class. Take the j^th^ feature into the individual discriminative features set.
6 The (K–C) features which are having biggest hedge value are selected
7 There are K discriminative features. The new training X_{new} and testing data are created by the selected features from the original data.

A neural-fuzzy system is a constitution of neural networks and fuzzy systems. The combination is such that the neural networks algorithms are used to determine parameters of fuzzy system. Layers in ANFC with linguistic hedges are as follows (Cetisli 2010),

EXPERIMENTAL SETUP AND RESULTS:

An omnidirectional microphone was placed on the pedestrian sidewalk at about 1 m height, and it recorded the cumulative signal at 16000 Hz sampling frequency. A small area segment (64, Ring Road, Rana Pratap Square to 505/506/507 Wardha road) of Nagpur city, India is considered. Total 180 acoustic samples were collected for three traffic density states. Each sample acoustic signal is of time span approx 30ms. Different combination primary window size and shift size were considered while performing windowing and segmentation of acoustic signals. The possible combinations were considered are 200_100MFCC, 500_200MFCC, 1000_300MFCC (ex. Dataset 200_100MFCC: Primary window size of 200ms and shift by 100ms, features were extracted using MFCC). In single frame, first 13 cepstral coefficients were extracted. All feature vector frames were considered for study. Dimensionalities of above datasets and selected features were presented in table 1.Table 2 compares the training time of SCG and SSCG algorithm per iteration.

Table 1. Dimensionality of Original datasets, selected features and reduced datasets

Dataset	Approx Number of Frames	Dimension of Dataset (Number of Features by Instances)	Selected Features using linguistic hedges	Reduced Dimension of Dataset
200_100 MFCC	305	3965 by 180	1, 9, 7, 3, 8, 13, 6	2135 by 180
500_200 MFCC	150	1950 by 180	1, 9, 13, 3, 7, 11, 6	1050 by 180
1000_300 MFCC	100	1300 by 180	1, 3, 9, 6, 7, 4, 10	700 by 180

Table 2. Classification accuracy of ANFC_SCG and ANFC_SSCG for varying window size and shift (in %)

Data	Training Algorithm	Training Set Recognition (%)	Testing Set Recognition (%)	Unit Iteration time (in second)	Shorten training time(%)	Cluster size per class
200_100 MFCC	SCG	86.677	86.443	47.173	76.06	3
	SSCG	86.086	86.181	11.292		
500_200 MFCC	SCG	88.445	88.260	18.343	65.27	3
	SSCG	88.121	88.047	6.370		
1000_300 MFCC	SCG	89.702	89.544	12.8793	62.85	3
	SSCG	89.456	89.342	4.7843		

Conclusion:

In this research work, Speeded Scaled Conjugate Gradient (SSCG) is employed to reduce network training time per iteration and adaptive neuro-fuzzy classifier is modeled to classify traffic density states as low, medium and heavy corresponds to speed range 40 km/hr and above, 20 – 40 km/hr and 0 -20 km/hr respectively. Multiple feature vector frames were considered and are passed to classifier. Generally consideration multiple frames will increase the classification performance but the training time required per iteration is significantly high. To overcome this issue we have employed SSCG, which reduces the training time and also it preserves the classification performance.

References:

1. Cetisli, B. (2010): Development of an adaptive neuro-fuzzy classifier using linguistic hedges: Part 1. Expert Systems with Applications, doi:10.1016/j.eswa.2010.02.108.
2. Cetisli B., Barkana A. (2010): Speeding up the scaled conjugate gradient algorithm and its application in neuro-fuzzy classifier training, Soft Computing 14:365–378, DOI 10.1007/s00500-009-0410.
3. Kato, J. (2005): An Attempt to Acquire Traffic Density by Using Road Traffic Sound. In Proceedings of the 2005 International Conference on Active Media Technology. DOI: http://dx.doi.org/10.1109/AMT.2005.1505370, 353-358.
4. Li, C., Shimamoto, S. (2011): A Real Time Traffic Light Control Scheme for Reducing VehiclesCO2 Emissions. In Proceedings of the 8th Annual IEEE Consumer Communications and Networking Conference - Emerging and Innovative Consumer Technologies and Applications.
5. Møller M. (1993): A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw 6(4):525–533. doi:10.1016/S0893-6080(05)80056-5.
6. Nakagawa, S., Longbiao, W., Ohtsuka, S. (2012): Speaker Identification and Verification by Combining MFCC and Phase Information. IEEE Transaction on Audio, Speech, and Language Processing, Vol 20 (4).
7. Sen, R., Siriah, P., Raman, B. (2011): Roadsoundsense: Acoustic sensing based road congestion monitoring in developing regions. In IEEE SECON.
8. Valcarce, R., Mosquera, C., Gonzalez, R. (2004): Estimation of road vehicle speed using two omnidirectional microphones A maximum likelihood approach. EURASIP Journal on Advances in Signal Processing, DOI: http://dx.doi.org/10.1155/S1110865704311133, 2004: 1059-1077.