Information Sciences 294 (2015) 600–611

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Automated intelligent system for sound signalling device quality assurance

Tomasz Maniak [a,*], Chrisina Jayne [b], Rahat Iqbal [b], Faiyaz Doctor [b]

[a] Engineering and Quality Department, Nippon Seiki, UK-NSI Co. Ltd, Redditch, Worcs B98 9HL, UK
[b] Faculty of Engineering and Computing, Coventry University, Priory Street, Coventry CV1 5FB, UK

ARTICLE INFO

Article history:
Received 7 January 2014
Received in revised form 4 September 2014
Accepted 25 September 2014
Available online 14 October 2014

Keywords:
Non-speech sound recognition
Audio signal processing
Artificial neural networks
Auto-encoder
Manufacturing
Automation

ABSTRACT

This paper presents a novel approach to the detection and recognition of faulty audio signalling devices as part of an automated industrial manufacturing quality assurance process. The proposed system outperforms other well-established automated systems based on mel-frequency cepstrum coefficients (MFCC) and multi-layer perceptron (MLP). It uses both unlabelled sound data and labelled historical data acquired from human experts in detecting faulty signalling devices. The unlabelled data is used to train a deep neural network generative model to create multiple levels of hierarchical feature extractors which are used to train an MLP classifier, with the intent to model the human reasoning and judging processes in respect to sound classification. This paper presents the results of real world experiments based on data pertaining to the audio signalling quality assurance process for car instrument cluster manufacturing. These results show that the proposed system is able to successfully classify speakers into two groups: "Good" and "No good" depending on the part quality. The proposed system proves to be capable enough to eliminate the need for a manual inspection within the manufacturing process and is shown to be able to diagnose a fault with a high degree of accuracy. This work can be extended to other areas of automotive inspection where there is a need for a robust solution to sound detection and where an output signal is represented by a complex and changing frequency spectrum even with significant environmental noise.

1. Introduction

The manual inspection of sound produced by an audio device is standard practice in the car instrument cluster manufacturing process. This process generally involves a human operator with expert knowledge performing classification routines by deciding which of the parts are faulty and which are not. This approach to detection of faulty devices although widely accepted lacks many important characteristics of an automated test such as repeatability, accuracy and cost efficiency. It has been a long term challenge for companies and engineers to automate the inspection process of these devices. Until recently all the solutions provided have not been sufficiently advanced enough for use in production of car instrument clusters either because of their deficiencies in performance, accuracy or speed [9,21,32].

* Corresponding author at: UK-NSI, Research, Merse Road, North Moons Moat, Redditch B989HL, UK.
E-mail addresses: maniakt@uni.coventry.ac.uk, tom.maniak@uk-nsu.co.uk (T. Maniak), chrisina.jayne@coventry.ac.uk (C. Jayne), r.iqbal@coventry.ac.uk (R. Iqbal), fayaz.doctor@coventry.ac.uk (F. Doctor).

http://dx.doi.org/10.1016/j.ins.2014.09.042
0020-0255/

There are many types of audio signalling devices available on the market, examples of these include buzzers, speakers, whistles, foghorns, tone sounders, just to name a few. These devices are very often embedded into more generic electronic appliances such as doorbells, alarm clocks, washing machines, computers, cameras and instrument clusters. As with any other part they have to be tested at some point during the manufacturing process in order to maintain product quality and meet design requirements.

One can distinguish between two main types of audio signalling devices used in electronics: buzzers or speakers. The main difference between the two is the complexity of the frequency envelope they can produce. Speakers are designed for a large range of frequencies whereas buzzers very often have an output of only one frequency. The main advantage of buzzers is that they are very cheap to manufacture and can create higher pitched sounds at lower powers.

A signalling device very often produces a complex output signal resulting from non-linear characteristics of the sound itself. This non linearity is partially caused by the way the waves of pressure propagate through compressible media, the physical characteristic of waves (such as frequency, amplitude, harmonic constitution) and the human perception of the sound. The propagation of the sound can further cause the waves to refract, reflect or attenuate creating more complicated patterns. As a result the output signal produced cannot be described by a simple linear function, and therefore cannot be easily assessed by simple voltage measurement techniques.

Therefore the main difficulty in automating the manual inspection process could be in identifying the salient features of the sound signal. These features can be used to classify a faulty device and model the expert decision making process of the human operator. Although the methods for testing speaker quality with the use of multi-layer perceptron and mel-frequency cepstral coefficients capable of modelling the expert knowledge have been investigated and successfully applied in industrial environment [28] there is still need to improve the classification rate and flexibility.

Advances in training artificial neural networks and the development of Deep Learning [14] have opened new possibilities to revisit well-established methods for faulty audio signalling device detection. Using features other than MFCCs has been a focus of research in the sound recognition community [12]. Successful application of deep neural networks (DNN) to sound recognition provides a valid alternative [17,30] for extracting features different than MFCCs.

This paper introduces a novel approach to automated sound signalling device quality analysis through the use of a layer by layer learning algorithm to train a multi-layer generative model with a fixed length sound spectrogram. The method described in this work utilizes a four layer architecture pre-trained with the Boltzmann machine (RBMs) algorithm. The Boltzmann Machine belongs to a class of graphical models known as Markov random fields [22]. The pre-trained generative model is then used to create a deep auto-encoder and the network parameters are fine-tuned using a back-propagation algorithm to reconstruct the sound spectrogram. The network created in this way can be used for the efficient compression/coding of sound. The bottleneck encoder layer is used as an input to an MLP classifier and trained with a supervised learning algorithm thereby minimizing the discrepancies between target and output class labels. A fully trained network is able to capture human expert knowledge and classify the output of each sound signalling device based on the quality of the sound produced or the fault type. The system is then tested on an industrial application which controls the flow of the production line based on the sound classification result.

The method presented in this work has been proven to reduce the overhead and inaccuracies of the current manual approach thus improving product quality, performance of the production line and ultimately reducing manufacturing costs. The shift from manual sound checking to a reliable automated approach would allow a company to utilize some of the limited human resources in other areas of the production process and increase first time yield (FTY) which is defined as a number of goods produced divided by the total number of units being processed. The solution provided in this paper can be used both for buzzers and speakers. The experiments were performed on a real production line and the performance of the trained system was measured with uniform and uneven background noise. The experiments were then extended to measure the performance of the system with different white noise levels for two types of microphone. The automated system was finally compared to both a manual inspection quality check (with the use of inter-rater agreement statistic for ten operators) and an automated sound signalling device quality tool based on MFCC and MLP.

The rest of this paper is organized as follows. Section 2 presents the literature review. Section 3 introduces the methods utilized in the work: deep belief nets, Restricted Boltzmann Machines, auto-encoders and classification techniques for sound detection. Section 4 explains the manual evaluation and classification process with its inefficiencies. Section 5 describes the proposed automated inspection system and outlines the methods, system architecture and stages included in the automated inspection process. Section 6 presents the experiments including the results and performance of the proposed system observed on a real production line. Concluding remarks and future work considerations can be found in section 7.

2. Literature review

Previous research in the application of machine learning and artificial intelligence to the field of environmental sound detection seems to have produced very few publications [34]. In the domain of automated sound detection there is much emphasis placed on speech detection including many feature extraction techniques such as perceptual linear prediction features, linear prediction cepstral coefficients (LPCC) or mel-frequency cepstral coefficients (MFCC). These, although in some cases relevant to non-vocal sound detection, are not necessarily optimal and more sophisticated methods of sound feature extraction can yield better results [1]. The reason for their low performance in environment sound detection can be attrib-