ABSTRACT

In this paper, a new real-time approach for audio recognition of waterbird species in noisy environments, based on a Texas Instruments DSP, i.e. TMS320C6713 is proposed. For noise estimation in noisy water bird’s sound, a tonal region detector (TRD) using a sigmoid function is introduced. This method offers flexibility since the slope and the mean of the sigmoid function can be adapted autonomously for a better trade-off between noise overvaluation and undervaluation. Then, the features Mel Frequency Cepstral Coefficients post processed by Spectral Subtraction (MFCC-SS) were extracted for classification using Support Vector Machine classifier. A development of the Simulink analysis models of classic MFCC and MFCC-SS is described. The audio recognition system is implemented in real time by loading the created models in DSP board, after being converted to target C code using Code Composer Studio. Experimental results demonstrate that the proposed TRD-MFCC-SS feature is highly effective and performs satisfactorily compared to conventional MFCC feature, especially in complex environment.

KEYWORDS

Code Composer Studio, Mel Frequency Cepstral Coefficients, Noise Estimation, Spectral Subtraction, Support Vector Machine, TMS320C6713 DSK, Tonal Region Detection, Waterbird Recognition

INTRODUCTION

Algeria is home to a great diversity of wetlands, which are important sites for the wintering and the breeding of Palearctic migratory waterbirds. Tonga Lake, of El Kala National Park (northeast of Algeria), is a very special place for tens of thousands of different waterbirds species to overwinter or to a temporary halt. With an area of 2700 ha, Tonga lake is a refuge for over thirty-six different waterbird species and a major nesting site for Africa endangered waterbirds, namely: White-headed Duck (*Oxyura leucocephala*) and Ferruginous Duck (*Aythya nyroca*) which are threatened species ranked as “Endangered” and “Near Threatened” in the International union for conservation of nature red list, respectively. Currently, Tonga Lake and Garaet Hadj Tahar (northeast of Algeria) host the highest numbers of this species in Algeria and North Africa (Chettibi, 2013; Aissaoui et al., 2011).
One of primary element of ecological environment is environmental sound, which contains a great amount of rich information. Birds living in complex ecological environments can produce several types of sounds, which also contain a wealth of information closely related to the human survival environments. By examining, analyzing and recognizing bird sounds, pertinent information of birds can be acquired.

Apart from counting of individuals of each species, categorization of bird species and learning their life habits, automatic audio recognition of bird species has great potential significance in evaluating and predicting the surrounding living, as well as countless benefits on the protection development of ecological environment, which positively influence the development of economy and industry.

Recently and due to numerous analogies that exist between human speech and birdsongs (Acevedo et al., 2009), various techniques designed for automatic speech recognition have been adapted to automatic bird sound recognition as in (Lee et al., 2008; Chu & Blumstein, 2011; Jančovič & Kokuer, 2011; Patti, 2013). In a majority of these studies, databases used in the experiments consist on pretreated and enhanced sounds of different bird species (Stowell & Plumbley, 2014; Ptacek et al. 2016). However, in the real case where the birds are in their natural habitat, recordings of their sounds are subject to numerous environmental noises (e.g. sound of wind, rain, thunder, etc…) thus distorting the individual’s recognition operation.

The ultimate objective of this work is to develop a method based on embedded systems to survey the population of endangered waterbirds in natural habitat. The identification of different waterbird species can be done from an automatic recognition of their sounds. These kind of applications necessitate high-speed DSK (DSP Starter Kit). For this reason, in this study TMS320C6713 digital signal processor kit (DSK) has been used. The floating point TMS320C6713 DSP is a signal-processing platform of Texas Instrument. It features sufficient processing power for this project. As, it has a low current consumption to build an embedded system for bird monitoring in natural environment based on this DSP. Indeed, unlike the Raspberry board, that is rather a small computer for general use, with an operating system based on the Linux system, the TMS320C6713 DSP is a microprocessor optimized to perform different digital signal processing applications in real-time as in (Kumar, Singh, & Bhattacharya, 2015; Li, Bai, & Gao, 2015; López-Robles, Rodríguez-Rivas, Peralta-Sánchez, & Carranza-Castillo, 2015).

Considering the presence of noise in real world, many studies thought of taking advantage of the approaches of noise reduction (de Oliveira, 2015; Zhang and Li, 2015; Ventura et al., 2015; Potamitis, 2015). However, these methods have not been designed for embedded system with limited capacity. In this work, a tonal region detection (TRD) approach to identify the foreground bird sound of interest in complex environments is proposed. It is shown that a flexible sigmoid function can be used to replace the sound presence probability (SPP) algorithm in (Gerkmann & Hendriks, 2012). This function offers the possibility to alter the slope and the mean of the SPP independently to achieve a desired trade-off between noise overestimation and underestimation. It is also argued that a soft SPP is insufficient for the noise tracking, and improve it by employing harder decisions based on conditional smoothing.

For feature extraction step, a new method designed for waterbird sounds with very low SNR (signal to noise ratio) called MFCC-SS (for Mel frequency cepstral coefficients with spectral subtraction integration) is introduced. This method is developed to be initially, robust to natural noise (wind, heavy rain, thunder, etc…) and to have a very low computational complexity, making it suitable for implementation on embedded systems. The principle behind this technique is the integration inside the popular algorithm MFCC, a spectral subtraction (Boll, 1989) for bird sound enhancement. In fact, the spectral subtraction method of single channel speech enhancement is the most widely used conventional method for reducing additive noise (Upadhyay & Karmakar, 2015).
Application of Red, Green, and Blue Color Channels in 3D Shape Measurement
www.igi-global.com/chapter/application-red-green-blue-color/64393?camid=4v1a