Comparison of the Vocal Calls of Alexandrine and African Grey Species of Parrots using LPC Based Analysis Approach

Kamaljit Singh Arora, Randhir Singh, Parveen Lehana

Abstract—Speech analysis is one of the interesting analytical approaches in the areas of digital signal processing and it has been explored for various research applications including modeling of speech signals, phonetics research, understanding the speech production mechanism, speech coding and speech recognition processes, etc. Birds are completely dependent on their vocal signals in order to fulfill their survival requirements like nesting, food, protection from any threat or danger and other mutual communicating activities, etc. Birds communicate with their vocal signals because they have a greater range of sounds than humans. Moreover, bird’s calls contain a lot of useful information, but it cannot be easily recognized by human ear as time and frequency resolution of our auditory system is limited. The basic mechanism of sound production in birds is almost similar to that of humans in many contexts. LPC model has been used as an efficient speech model for the analysis of human voice. So, it may be used for the analysis of bird’s vocal calls. This paper examines an investigating method for exploring the application of phonetics research on parrot’s vocal calls. LPC based analysis approach has been applied on the vocal calls of Alexandrine Parakeet (Psittacula Eupatria) and African Grey (Psittacus Erithacus) species of parrots and investigations have been successfully carried out for comparing the number of phonemes in the calls of both the species using line spectral frequencies (LSF) vectors and Euclidean distances. Line spectral frequencies can be used to encode speech spectral information more efficiently than other transmission parameters. A classical method, known as vector quantization has also used for performing efficient speech analysis.

Index Terms—Speech analysis, Vocal calls, Linear predictive coding (LPC), Line spectral frequencies (LSF).

I. INTRODUCTION

Speech signal processing is a special form of digital signal processing in which the speech signals are processed in digital forms so as to fetch some useful information using different speech processing methods. It has been significantly used for various speech applications including speech recognition, speech coding, speech analysis, speech synthesis and speech enhancement, etc [1]. Speech analysis is an analytical process in which speech signals are usually processed for the extraction of time varying parameters or desirable speech information [2]. It has been used for various speech applications including modeling of speech signals, phonetics research, understanding the speech production mechanism, speech coding and speech recognition, etc [3].

Different species of birds can understand and produce sounds within a few days of birth. They have greater range of sounds than humans because their time resolution ability is ten times better than us. Vocal sounds help birds for the need of their survival requirements like nesting, food, protection from any threat or danger and any other mutual communicating activities, etc. Interpretation of vocal calls may provide various information like change of season, danger, surrounding environment, group size of the birds, emotional state of the bird, etc. Birds and their sounds play a significant role in the life of humans also. Bird’s calls give us a stress free realization and always attached us with the nature beauties. People can recognize and enjoy different calls and songs of birds in different places including villages, mountains and cities. Bird’s calls are a source of inspiration for many children, poets, artists, writers, and music composers [4], [5].

The basic vocal sound production mechanism in birds is similar to that of humans in many contexts. As we know that, human’s speech are composed of small sound units called phonemes, the bird’s sounds are also comprised of these small units. Humans use vocal cords for excitation of their vocal tract, but there are no vocal cords present inside birds. Instead, they have a special important organ, known as syrinx and its position can be naturally present in the intersection of the trachea. Syrinx is an important sound organ used for the production of sounds in birds and it also gives information about the internal structure of different birds because different bird’s species have different anatomy or structure of syrinx [6], [7]. The main internal parts of speech for the sound production mechanism in parrots include lungs, bronchi, tracheal syrinx, larynx, trachea, mouth, and beak, are shown in Fig.1. Different interiors of the syrinx can be used by different parrots to produce calls and songs, and sound production is controlled by non homologous part of their brain. Tracheal syrinx is generally found in parrots. The parrot tracheal syrinx is mostly composed of different muscles including superficial syringeal muscle,
Comparison of the Vocal Calls of Alexandrine and African Grey Species of Parrots using LPC Based Analysis Approach

II. VOCAL COMMUNICATION IN PSITTACINES

Parrots are biologically known as Psittacines. Vocal communication in psittacines is composed of different varieties of calls generally short or innate vocalizations and songs are comprised of longer or complex vocalizations which are used to convey some specific kind of information. There are more than 300 species of parrots across the world. Parrots occupy small to extra large range of sizes. For example, Pygmy parrot is only four inches (10 cm) long whereas Hyacinth Macaw can grow even three feet (one m) long. Parrots are also popular for their bright colors. Some parrots are red, blue, green, orange and yellow in color. Moreover, some parrots can be overall white. These species of birds occupy a hearing frequency range from 1 KHz to 4 KHz [13], [14].

The Alexandrine Parakeets (Psittacula eupatria) are larger version of P. krameri manillensis. They are about 20 inches long. These species are generally found in eastern Afghanistan, Sri Lanka, India and china counties, etc. They have an attracting appearance attributable to their length and coloring. They have a black stripe under their cheek and a bright pink collar around their neck [15], [16]. African grey parrots (Psittacus Erithacus) have a length of about 13 inches and are native to Africa and can be found from the Ivory Coast to Angola, Kenya and Tanzania, etc. Some shadows of black and grey usually found in them. The African grey parrot is a beautiful bird and some people recognized them as excellent talkers of all the birds with voices mostly resembling with human’s speech. They found out to be very alert and intelligent [15].

Vocal communication in Parrots involves different types of sounds, known as bird calls which are given for special communicating activities. Mostly, there are different varieties of calls produced by parrots for their vocal communication. But some of the meaningful calls identified in them are mentioned in Table I.

<table>
<thead>
<tr>
<th>S.N.</th>
<th>Parrot call</th>
<th>S.N.</th>
<th>Parrot call</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alarm call</td>
<td>5</td>
<td>Mating call</td>
</tr>
<tr>
<td>2</td>
<td>Pre-flight call</td>
<td>6</td>
<td>Agonistic protest</td>
</tr>
<tr>
<td>3</td>
<td>Begging call</td>
<td>7</td>
<td>Soft contact call</td>
</tr>
<tr>
<td>4</td>
<td>Distress call</td>
<td>8</td>
<td>Loud contact call</td>
</tr>
</tbody>
</table>

Alarm call is one of the interesting calls of parrots and it is usually produced by parrots when they feel any threat or danger near them. It is also produced to protect or save other species of birds from some kind of danger. Pre-flight call is given by parrots when they just started flying in the sky and it is recognized as a loud and harsh flight call. Begging call is one of the needy calls of parrots and it is generally given by hungry parrots when there is a lack of food. Distress call identifies that the parrot is presently of distressed nature. Agonistic protest is...
mostly given because some species of parrots are of angry or disturbed nature due to some personal reasons like jealousy, temperature, etc. Some species of parrots also make duet with each other and produced a call which identify their sexual behavior, known as mating call. Soft contact call coordinates the movement of flock members when they are flying in a group for the search of different foods or a particular type of vegetation. This call is identified as low in amplitude and can be repeatedly given with or without any responses by other flock members. Mostly all species of parrots can produce a special type of call, known as loud contact call and it is generally mutually communicated with some other species of flying birds. It is usually given to make a vocal connection with other species of birds and it is recognized as the loudest of all the calls [18], [19], [20].

Their calls are also composed of small units called phonemes. Many researchers have found out that bird’s sounds resemble human language. In both birds and humans, the individual sounds have less importance as compared to the combination of sounds. Moreover, by comparing human language, bird’s sounds are only similar at the phonetic level and not at the syntactic level. In human language, words are composed of combination of phonemes, but each distinct phoneme has no inherent meaning. Bird’s vocal calls are similar to the phonetic sub-component of human language while it creates differences at syntax and semantics that give expressive power to their communication [21].

III. LPC BASED ANALYSIS APPROACH

Linear predictive coding model of speech was implemented by United States Govt. Department of Defense in 1984 using a federal standard LPC-10. LPC model has been used in many speech recognition systems because it gives an excellent modeling of speech signals. It can reduce calculations in speech recognition process with increased accuracy. LPC can be used for the analysis of speech signals with the use of a source-filter model. LPC can be used to find out vocal tract area functions, fundamental frequency, bandwidths and frequencies of spectral poles and zeros, but it is mostly used to determine small set of parameters of speech which represents vocal tract configuration. The main reason behind the use of LPC is to reduce the sum of squared differences between the original speech signal and estimated speech signal within a finite duration [22], [23]. The block diagram of LPC based analysis approach is shown in Fig.3. A digital filter is used at the initial stage in order to compensate the high frequency portion of the speech signal that was suppressed during the mechanism of speech production. A window size is generally adjusted in order to reduce the dissimilarities in the speech signal at the edges of each frame of speech and hamming window is mostly used for LPC analysis. The autocorrelation function (ACF) can be used to provide a warped autocorrelation so that frequency resolution can be represented by the varying lambda. Levinson Durbin recursion algorithm can be used for calculating auto regression (AR) coefficients and reflection coefficients (RC). The LPC to LSF block converts linear prediction coefficients to line spectral frequencies. After that, spectral parameters are used for vector quantization. Vector quantizer follows the principle of block coding and has the ability to provide lossy data compression. It provides fixed length algorithm and functions like an approximator [24], [25], [26].

IV. RESEARCH METHODOLOGY

Experimental investigations may be divided into three sub-parts: Material recording, Segmentation process and Phonemic analysis of vocal calls. The corresponding flow chart is shown in Fig.4.

A. Material Recording

The Analysis was carried out by recording vocal calls of Alexandrine (Psittacula Eupatria) and African Grey (Psittacus Erithacus) species at different locations using a high quality voice recorder i.e. Sony digital voice recorder of series ICD-UX513F. It provides 4GB UX digital voice recording. It has expandable memory capabilities with high voice quality. Acoustically shielded environment is used for the recording of vocal calls. A sampling frequency of 16 KHz was used for the sampling process. After that, recorded calls were labeled and stored in wav format.

B. Segmentation

After the recording process, speech segmentation is required to outline the visual information for making efficient speech analysis. The segmented calls were then labeled, stored and processed for investigating phonemic analysis. Speech segmentation is the process for identifying the boundaries in a particular speech signal between words, syllables, or
phonemes. It is defined as the mechanism of labeling speech signal areas with symbolic information in some applications [27]. Segmentation is usually performed for the analysis of spectrograms. A spectrogram is a computer generated plot which shows different frequencies of speech signal at each instant of time. The spectrograms can be significantly used to recognize speech signals on the basis of their amplitude, frequency and time duration. It can be used for the visual analysis of speech signals and represented as the square of their absolute value of short time Fourier transform (STFT) [28].

C. Phonemic Analysis of vocal calls

For the phonemic analysis, the calls were analyzed using LPC model by adjusting order ‘L’ of line spectral frequencies using a hamming window with some shifting level. Different classes were recognized using LSF and the centroids of classes separated by lesser Euclidean distances were clubbed together. Experiment was repeated for different class numbers until we obtain a level where we cannot merge the centroids further.

![Fig.4. Experimental Methodology](image)

After that, the classes were merged using LSF vector quantizer. Linear predictive coding parameters can be used in many speech processing applications for identifying the spectral envelope information of speech signals. LSFs provide an alternative parameterization of the analysis filter used for linear predictive coding (LPC) of speech. They can be used to encode speech spectral information more efficiently than other transmission parameters. Moreover, LSF’s have the properties of frame interpolation with smooth spectral changes due to their frequency domain interpretation [39], [30]. Vector quantization technique has been used for LPC analysis and coding applications. Vector quantization is a classical quantization process which allows the modeling of probability density function by the distribution of prototype vector. It is an efficient process for making speech compression and has been successfully used in various applications involving VQ based encoding and recognition [31], [32].

V. RESULTS AND ANALYSIS

Investigations were carried out for the phonemic analysis of Alexandrine and African Grey vocal calls using Euclidean distances and LSF vector quantization (VQ). It was observed that the analysis may carry out by adjusting 21 order line spectral frequencies using a hamming window of 25 ms with 1 ms shifting. Before processing the calls, the distances among the classes were made zero if it was less than the threshold.

The numbers of zero’s Euclidean distances were calculated for classes 2 to 32 for both the species. But, for example, the calculation of the no. of zero’s Euclidean distances of some classes for Alexandrine and African Grey vocal calls are shown in Table II and Table III respectively. It is clear from the Table II that the mean of zero’s Euclidean distances remains one (1) up to class 10 and suddenly increases when moving from class 10 to 11, etc. This indicates that there are around 10 phonemes in the Alexandrine calls On the other hand; the mean of zero’s Euclidean distances remains one (1) up to class 14 in Table III and suddenly increases when moving from class 14 to 15 and so on. This indicates that there are around 14 phonemes in the African Grey calls.

The stem plots for different classes ranging from 2 to 32 were also analyzed for both the species. But for example, stem plots showing Euclidean distances of LSF vectors for class 9, 10 and 11 are shown in Fig.5, Fig.6 and Fig.7 respectively for the analysis of Alexandrine vocal calls. The stem plots showing Euclidean distances of LSF vectors for class 13, 14 and 15 are also shown in Fig.8, Fig.9 and Fig.10 respectively for African Grey vocal calls. So, it is clear from the stem plots that when the number of classes goes beyond 10 and 14, more and more Euclidean distances becomes zero, in vocal calls of Alexandrine and African Grey species of parrots respectively.

The total number of classes represents the total number of phonemes in the calls. So, it can be seen from the ribbon plots that the numbers of classes were around 10 and 14 in both the species respectively. The ribbon plots for classes 10 and 14 for both the species are also shown in Fig.11 and Fig.12 respectively.

<table>
<thead>
<tr>
<th>Table II. Calculation of the number of zero’s Euclidean distances for Alexandrine vocal calls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class No.</strong></td>
</tr>
<tr>
<td>7</td>
</tr>
<tr>
<td>8</td>
</tr>
<tr>
<td>9</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table III. Calculation of the number of zero’s Euclidean distances for African Grey vocal calls</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Class No.</strong></td>
</tr>
<tr>
<td>11</td>
</tr>
<tr>
<td>12</td>
</tr>
<tr>
<td>13</td>
</tr>
<tr>
<td>14</td>
</tr>
<tr>
<td>15</td>
</tr>
<tr>
<td>16</td>
</tr>
</tbody>
</table>
Fig. 5. Stem plot for Euclidean distances of LSF vectors for class 9, when $m = 1$
(in Alexandrine vocal calls)

Fig. 6. Stem plot for Euclidean distances of LSF vectors for class 10, when $m = 1$
(in Alexandrine vocal calls)

Fig. 7. Stem plot for Euclidean distances of LSF vectors for class 11, when $m = 4$
(in Alexandrine vocal calls)

Fig. 8. Stem plot for Euclidean distances of LSF vectors for class 13, when $m = 1$
(in African Grey vocal calls)

Fig. 9. Stem plot for Euclidean distances of LSF vectors for class 14, when $m = 1$
(in African Grey vocal calls)

Fig. 11. Ribbon plot indicating 10 numbers of classes
(in Alexandrine vocal calls)
Comparison of the Vocal Calls of Alexandrine and African Grey Species of Parrots using LPC Based Analysis Approach

Bird’s vocal calls contain a lot of useful information which is beyond human’s hearing sense because birds have excellent time resolution ability. The unprocessed or natural calls do not provide us efficient knowledge regarding their vocal behaviour and communicating activities. Parrots are very intelligent birds and can produce complex vocalizations. They can also mimic human speech. Different vocal calls of Alexandrine Parakeet and African Grey species of parrots were processed using LPC model and LSF vector quantizer. The parameters were adjusted as 21 order line spectral frequencies using a hammering window of 25 ms with 1 ms shifting. So, LPC based analysis has been recognized as an efficient approach for making a phonemic comparison between the two species. From the analysis, the number of phonemes found out in Alexandrine and African Grey species of parrots were calculated around 10 and 14 phonemes respectively. Phonemic analysis can be used for fetching the knowledge regarding vocal behavioral information’s of different bird’s varieties. It can also be used for identifying unknown cross breeding species from their phonemes. It should be noted that the number and types of phonemes may be slightly varying for different species of Psittacines belonging to different habitats. Researches can be extended in future for fetching more and more information from their vocal calls since the calls are comprised of some meaningful information’s, for e.g. change of season, danger, surrounding environment, emotional state of a particular bird, size of the bird group, etc.

VII. ACKNOWLEDGEMENT

First and above all, I would like to bow my head in front of God, the Almighty for providing me this work to proceed successfully. I would like to pay a special thanks to my parents for everything they have done for me because without them everything wouldn’t be possible. And last but not the least; I am very much thankful to IJIEAT Editor and team members for providing us this prestigious platform.

REFERENCES

AUTHOR PROFILE

Er. Kamaljit Singh (M.Tech Engineer) resident of Jammu (J&K) received his B.Tech degree in Electronics and Communication Engineering with an aggregate percentage of 72% (First division) from M.B.S College of Engineering and Technology (MBSCET) in the year 2008. He also have training experience in Rural Electrification Corporation of India Ltd., Bharat Sanchar Nigam Limited (Central Govt.) and passed his M.Tech degree in Electronics and Communication Engineering with an aggregate percentage of 68% (First division) from Sri Sai College of Engineering and Technology (SSCET), Punjab, affiliated to Punjab Technical University, Jalandhar (Punjab) in the year 2012. His research interests include Speech signal processing, Digital signal Processing, Image processing, Analog and Digital Communication, Electronics and control systems, and having more than 10 publications in national/international conferences and journals.

Er. Randhir Singh (Assistant Professor and Head) received his M.Tech. Degree in Electronics and Communication Engineering from Beant College of Engineering and Technology, Gurdaspur, Punjab affiliated to Punjab Technical University, Jalandhar (Punjab). He is presently working as Assistant Professor (H.O.D.), Electronics and Communication Engineering Department, Sri Sai College of Engineering and Technology, Pathankot (Punjab) and pursuing PhD in Punjab Technical University. His research interests include Speech signal processing, Digital signal Processing, Image processing, Analog and Digital Communication, Electronics and control systems, etc. and having more than 10 publications in national/international conferences and journals.

Dr. Parveen Lehana (Associate Professor) received his Master’s degree in Electronics from Kurushetra University in 1992. He worked as lecturer in Guru Nanak Khalsa College, Yamuna nagar, Haryana for next two years. He qualified NET-JRF in Physical science in 1994 and got selected as permanent lecturer in A. B. College, Pathankot, where he worked for one year. He also qualified NET-JRF in Electronic Science and presently working as Associate Professor in Physics and Electronics Department, University of Jammu and did Ph.D. degree from IIT, Bombay in Speaker Transformation. He also invited in many colleges for attending national and international conferences and also guided many students for short term certification programs in embedded systems. He also invited for conducting workshops on MATLAB/simulink in different esteemed institutions/colleges. His research interests include Speech recognition, Speaker transformation, Signal processing, Speech signal processing, Analog and Digital signal processing, Nanowires characterization, Robotics, Image processing, Analog communication, Digital communication, Microwave and Antennas, Electronics and control systems, Instrumentation, Electronics system design, etc. and having more than 100 publications in national/international conferences and journals. He has a lot of experience in guiding M.Tech, M.Phil, Ph.D. students and other researchers also.


©BEIESP