Automatic Classification of African Elephant (*Loxodonta africana*) Follicular and Luteal Rumbles

Patrick J. Clemins[1]* and Michael T. Johnson[1]

[1] *Dept. of Electrical and Computer Engineering, Marquette University, P.O. Box 1881, Milwaukee, WI 53201*

INTRODUCTION

Recent research in African elephant vocalizations has shown that there is evidence for acoustic differences in the rumbles of females based on the phase of their estrous cycle (1). One reason for these differences might be to attract a male for reproductive purposes. Since rumbles have a fundamental frequency near 10Hz, they attenuate slowly and can be heard over a distance of several kilometers. This research exploits differences in the rumbles to create an automatic classification system that can determine whether a female rumble was made during the luteal or follicular phase of the ovulatory cycle. This system could be used as the basis for a non-invasive technique to determine the reproductive status of a female African elephant.

The classification system is based on current state-of-the-art human speech processing systems. Standard features and models are applied with the necessary modifications to account for the physiological, anatomical and language differences between humans and African elephants. The long-term goal of this research is to develop a universal analysis framework and robust feature set for animal vocalizations that can be applied to many species. This research represents an application of this framework.

The vocalizations used for this study were collected from a group of three female captive elephants. The elephants are fitted with radio-transmitting microphone collars and released into one of three naturalistic yards on a daily basis. Although this data collection setup is good for determining the speaker of each vocalization, it suffers from many potential noise sources such as RF interference, passing vehicles, and the flapping of the elephant's ears against the collar.

FEATURE EXTRACTION AND MODEL DETERMINATION

As is common in speech processing systems, features are extracted from the rumbles using a moving Hamming window. The window size used here is 300ms with one-third overlap between each window. In each frame, a modified set of ten Mel-Frequency Cepstral Coefficients (MFCCs) is calculated (2). The first change from standard MFCCs is that the Fast Fourier Transform of the window is zero padded to four times its original size in order to interpolate and smooth the magnitude spectrum in the lower frequencies. This interpolation is done because of the relatively high sample rate (7518Hz) in comparison to the frequencies of interest (10Hz – 150Hz). The other change is that the Mel-Frequency spaced filter banks are compressed to span the range from 10Hz to 150Hz. This change was made to compensate for the much lower range of frequencies rumbles reside in compared to that of human speech (about 100Hz – 4000Hz).

Five-state Hidden Markov Models (HMMs) are used to model the rumbles from each phase of the African elephant ovulatory cycle. The states of the HMMs are represented by a Gaussian distribution of ten cepstral coefficients. One HMM is trained for each class of rumble using the Baum-Welsh Expectation Maximization algorithm (3, 4). During evaluation, the Viterbi algorithm (5) is used to find the likelihood of each test vocalization coming from each trained HMM.

For this experiment, three different HMMs were trained for three different phases of the female African elephant ovulatory cycle. Female African elephants have two hormonal peaks per ovulatory cycle although they only ovulate during the second hormonal peak. In order to determine if and how the vocalizations very between each hormonal phase, three classes were used for classification; luteal, anovulatory follicular and ovulatory follicular. The training and testing vocalizations for the luteal class were chosen from rumbles made ±6 days from the midpoint of the luteal phase and the vocalizations chosen for the follicular classes were chosen from rumbles made ±6 days from each hormonal peak. An additional HMM was trained to model the silence before and after each vocalization.

	Anovul.	Ovul.	Luteal
Anovul.	8	7	4
Ovul.	6	19	6
Luteal	3	3	12

Figure 1. Confusion Matrix for 3-class classification

	Foll.	Luteal
Foll.	40	10
Luteal	5	13

Figure 2. Confusion Matrix for 2-class classification

RESULTS

Experiments were performed on 68 female African elephant rumbles. The rumbles were made in various behavioral contexts. Therefore, some errors could result from the fact that some rumbles in the dataset were not meant to attract a male at all, but instead to maintain contact with a member of the family group or other various reasons. Leave-one-out verification was used to generate confusion matrices for each of the three classes of rumbles. The confusion matrix of the 3-class classifier is in Figure 1. As can be seen from the matrix, the two classes of follicular rumbles are easily confused. However, the confusion between the two follicular classes and the luteal class is small. This would seem to imply that although rumbles made in the two follicular phases are similar, they differ significantly from rumbles made in the luteal phase.

In order determine how similar the two follicular classes are, the two follicular classes were collapsed into a single class. The confusion matrix of the two-class classifier is shown in Figure 2. The classification accuracy between follicular and luteal of this classifier is similar to that of the 3-class classifier (77.9% vs. 76.5%). This supports the assumption that the rumbles made in the two follicular classes are extremely similar.

DISCUSSION

By using an automatic classification system based on current state-of-the-art speech recognition systems, the hormonal status of female African elephant rumbles can be determined with 77.9% accuracy. The rumbles made during both follicular phases are extremely similar, but the follicular rumbles can be easily separated from rumbles made during the luteal phase.

There are many modifications that can be made to the current classification framework in order to improve classification accuracy. Although MFCCs performed well for this experiment, it is generally agreed that the fundamental frequency and harmonic structure of animal vocalizations are important for meaning. Therefore, the addition of the fundamental frequency and harmonic information as features could improve the accuracy of animal vocalization classifiers. Another future modification to the framework is the addition of noise correction algorithms. Noise correction can either be applied to the actual waveform or to the extracted feature vectors. Numerous algorithms developed for both types of noise correction can be found in human speech processing literature.

The underlying goal of this research is to develop a universal analysis framework and robust feature set for animal vocalizations that can be applied to many species. This successful application of the framework to the classification of African elephant rumbles shows the potential of such a framework.

ACKNOWLEDGMENTS

The authors would like to thank Anne Savage, Kirsten Leong, and the rest of the staff at Disney's Animal Kingdom™ for the collection and organization of the acoustic data used in this research.

REFERENCES

1. Leong, K. M., Ortolani, A., Graham, L. H., and Savage, A. (**2003**). "The Use of Low-frequency Vocalizations to Signal Female Reproductive Status in a Group of Captive African Elephants *(Loxodonta africana)*," Hormones and Behavior, in press.
2. Hidden Markov Model Toolkit (HTK) Version 3.1.1 User's Guide. (**2002**). Cambridge University Engineering Department.
3. Baum, L. E., Petrie, T., Soules, G., and Weiss, N. (**1970**). "A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains," The Annals of Mathematical Statistics, **41**, 164-171.
4. Moon, T. K. (**1996**). "The Expectation-Maximization Algorithm," IEEE Signal Processing Magazine, **13**(6), 47-60.
5. Forney, G. D. (**1973**). "The Viterbi Algorithm," Proceedings of the IEEE, **61**, 268-278.