Stress Detection and Classification of Laying Hens by Sound Analysis

Jonguk Lee, Byeongjoon Noh, Suin Jang, Daihee Park*, Yongwha Chung, and Hong-Hee Chang1
Department of Computer and Information Science, College of Science and Technology,
Korea University, Sejong 339-700, Korea

ABSTRACT: Stress adversely affects the wellbeing of commercial chickens, and comes with an economic cost to the industry that cannot be ignored. In this paper, we first develop an inexpensive and non-invasive, automatic online-monitoring prototype that uses sound data to notify producers of a stressful situation in a commercial poultry facility. The proposed system is structured hierarchically with three binary-classifier support vector machines. First, it selects an optimal acoustic feature subset from the sound emitted by the laying hens. The detection and classification module detects the stress from changes in the sound and classifies it into subsidiary sound types, such as physical stress from changes in temperature, and mental stress from fear. Finally, an experimental evaluation was performed using real sound data from an audio-surveillance system. The accuracy in detecting stress approached 96.2%, and the classification model was validated, confirming that the average classification accuracy was 96.7%, and that its recall and precision measures were satisfactory. (Key Words: Laying Hens, Stress Recognition, Sound Analysis, Monitoring System)

INTRODUCTION
Stress not only affects the wellbeing of commercial chickens, it also results in an economic cost to the industry that cannot be ignored. In general, captive chickens are exposed to a variety of stressful conditions in commercial poultry facilities. Some of the routine management practices themselves are stressful to the birds, and these are coupled with environmental pressures (Otu-Nyarko, 2010). Stress refers the way an organism responds to environmental stimuli that it perceives as a real or anticipated threat to its survival or wellbeing (Harvey et al., 1984). Environmental stressors, such as temperature and fear, have deleterious effects on the productive performance of laying hens (Mushtaq et al., 2013). In particular, heat stress depresses egg production (Muiruri and Harrison, 1991), egg weight (Balnave and Muheereza, 1997), and shell quality (Mahmoud et al., 1996). This is generally accompanied by a suppressed feed intake, which is a likely cause or the decline in production. Therefore, understanding environmental conditions is crucial to successful poultry production and welfare.

There have been several attempts to measure stress responses in animals (Gutierrez, 2013). However, conventional methods for measuring stress are not necessarily good indicators of welfare because they detect stress only after it has negatively affected the animals. One of the major challenges in assessing physiological responses to stress is that collecting data from livestock is often stressful in its own right (Freeman, 1976). Therefore there is a push to measure stress non-invasively and early enough to enable the farmer to remove the stressor before it has an adverse effect. The sound produced by animals is a candidate bio-signal that can be easily measured from a distance, and will not thereby cause any additional stress for the animals (Blahová et al., 2007). Furthermore, in recent years, sound analysis has become an increasingly important tool for interpreting the behavior, health condition, and wellbeing of animals (Steen et al., 2012; Chung et al., 2013a, b; Lee et al., 2014).

The field of bioacoustics, in particular, the study of animal vocalizations, has received increasing attention in...
recent years with the advent of new recording and analysis technologies (Otu-Nyarko, 2010). Bioacoustics is the study of the acoustic characteristics and biological significance of sounds emitted by living organisms (Tefera, 2012). Birds are one of the few animal groups known to exhibit vocal learning. They rely on acoustic communication for territoriality, mate choice, offspring recognition, alarm signaling, and individual recognition to make their presence known to one another (Waldvogel, 2000). Such avian expressiveness can be used as a tool to understand the bird’s wellbeing especially under stressful conditions.

In this paper, we develop a non-invasive, inexpensive, and automatic online-monitoring prototype that monitors the avian vocalizations in a commercial poultry facility and notifies the producer of a stressful situation when it occurs in the coop. The proposed system includes a detection and classification model that arranges three binary-classifier support vector machines (SVMs) hierarchically. The model begins by selecting an optimal acoustic feature subset from the sound emitted by the laying hens in the chicken coop. This process occurs offline, and it is unnecessary during the subsequent real-time online process. The detection and classification module detects the stress-related sound and classifies it into subsidiary sound types that are arranged hierarchically. For instance, the detected sound might be caused by physical stress from low or high temperatures, or mental stress resulting from fear. Finally, an experimental evaluation was performed using real sound data from an audio-surveillance system. The accuracy of the stress detection approached 96.2%, and the stress classification was validated, confirming that its recall and precision measures were satisfactory.

MATERIALS AND METHODS

Sample sound collection

The experiment was conducted in a commercial poultry production farm located in Jinju, South Korea. A total of 120 chickens, specifically, 44-week-old Hyline brown layers, were randomly housed in battery type metal wire cages. The dimensions for each cage were 60×60×40 cm. The cages were located in a controlled chamber (4×4×4.2 m) with constant temperatures set at 10°C±2, 21°C±2, and 34°C±2 during the entire period of the study. Each individual birdcage offered ad libitum access to feeding and drinking stations. The experiment consisted of 4 groups with 15 replications, and each replication included 2 birds per cage. Each group was exposed to physical stressors by changing the environmental conditions, that is, with temperature changes of 10°C±2, 21°C±2, and 34°C±2. One group was exposed to mental stressors resulting in fear. This was accomplished by hitting the cage with a stick while the temperature remained at 21°C±2.

The sounds emitted by the brown layers were recorded with a digital camcorder (Sony HDR-XR160, Tokyo, Japan) that was placed inside the chamber facing the cages for at least 30 to 60 minutes. The recorded video files were converted to a MP3 file by using a free video to MP3 converter, available online (v. 5.0.17 build 903 www.dvdvideoosoft.com). The converted MP3 files were then digitalized using Cool Edit (Adobe, San Jose, CA, USA) in a PC with a standard soundcard (Realtek AC97) at 16 bits and a 44.1 kHz sampling rate. The sounds collected were classified using a manual labeling method. This procedure is based on an acoustic analysis combined with a visual spectral analysis to extract specific sounds from the entire recording. A human operator was tasked with listening to the recorded files in their entirety and making and describing every sound. Labeling was done offline to extrapolate only those sounds that could be classified as calls from the operator’s visual observation of the spectrogram and auditory confirmation.

Acoustic features and feature subset selection method

The universal sound features set was initially established from the popular features found in acoustic literatures. The following features were derived from the time domain: Root mean square (RMS), power, energy, absolute extremum, intensity, shimmer, jitter, harmonic-to-noise ratio (HNR), and pitch. In addition, the following features were derived from the frequency domain: Formant F1 to F9, and the power spectral density (PSD) P1 to P39. A brief summary of these sound features is provided in what follows (Boersma, 2002; Slocombe and Zuberbühler, 2006; Guyer, 2009):

Time domain features:

i) RMS: The RMS is an amplitude modulated by a Gaussian random process.

ii) Power: The power is defined as $1/(t_2 - t_1) \int_{t_1}^{t_2} x^2(t) dt$, where $x(t)$ is the amplitude of the sound and $(t_1, t_2)$ is the time range.

iii) Energy: The energy is defined as $\int_{t_1}^{t_2} x^2(t) dt$, where $x(t)$ is the amplitude of the sound.

iv) Absolute extremum: The absolute extremum refers to the absolute value of the maximum amplitude from the sound.

v) Intensity: The sound intensity is the sound power per unit area.

vi) Shimmer: The shimmer is the average absolute difference between the amplitudes of consecutive periods, divided by the average amplitude.

vii) Jitter: The jitter is the average absolute difference between consecutive periods, divided by the average period.

viii) HNR: The HNR measures the ratio of the harmonic
signal power and the noise power in the observation

ix) Pitch: The pitch is the relative concept of frequency.

**Frequency domain features:**

i) Formant F1 to F4: Formants are characterized by the frequency of the peak, the resonance factor, and the relative amplitude level of the sound. The frequency of acoustic resonance was extracted between 0 to 5,000 Hz.

ii) PSD, PSD1 to PSD39: The PSD is the average power for the sound within a certain time and frequency range, expressed as $Pa^2/Hz$. We used 39 PSDs that were extracted every 100 Hz between 100 to 4,000Hz.

Selecting attribute (or feature) subset efficiently for pattern recognition is an important issue (George and Bo, 2008). Attribute selection involves selecting a subset of attributes from a feature set in order to provide a compact, precise, and fast recognizer, with minimal performance degradation, by removing the attributes that are ineffectual, redundant, or least-used (Hall, 1999). Reducing the dimensionality of the data reduces the size of the hypothesis space and allows algorithms to operate faster and more effectively. In this paper, we used correlation-based feature selection (CFS), which has been verified as the best attribute subset selection method (Hall, 1999; Yu et al., 2010). The CFS uses the features' predictive performances and inter-correlations to guide its search for a suitable feature subset. This heuristic takes into account the usefulness of individual features for predicting the class label, along with the level of inter-correlation among them. The CFS first calculates a matrix of feature-class and feature-feature correlations from the training data and then searches the feature subset space using a best first search. The version of CFS used in this paper includes a heuristic that considers locally predictive features and avoids a re-introduction of redundancy.

**Binary classifier support vector machine**

A SVM is presented in the proposed section, warranting a brief review of some of the basic literature on SVMs (Cristianini and Shawe-Taylor, 2000; Lee et al., 2013). In order to explain the principles for SVMs, we shall examine the simplest case, a two-class problem, where the classes are linearly separable. The goal in this case is to separate the two classes using a function that is induced from the available examples. Many possible linear classifiers can separate the data, but only one can maximize the margin. This linear classifier is called the optimal separating hyperplane. Consider the problem of separating the set of training vectors that belong to two separate classes,

$$D = \{(x^1, y^1), \ldots, (x^l, y^l)\}, x \in \mathbb{R}^n, y \in \{-1, 1\},$$  \hspace{1cm} (1)

with a hyperplane,

$$<w, x>+b = 0.$$  \hspace{1cm} (2)

If the set of vectors is separated without error and the distance between the vectors closest to the hyperplane is maximal then, this set is defined as optimally separated by the hyperplane. Separating the hyperplane in canonical form satisfies the following constraints:

$$y^i[<w, x^i>+b] \geq 1, \ i = 1, \ldots, l.$$  \hspace{1cm} (3)

The distance $d(w, b; x)$ of a point $x$ from the hyperplane $(w, b)$ is

$$d(w; b, x) = \frac{|<w, x>+b|}{\|w\|}.$$  \hspace{1cm} (4)

Hence, the hyperplane that optimally separates the data is the one that minimizes the following:

$$\text{minimize} \quad \Phi(w) = \frac{1}{2} \|w\|^2$$  \hspace{1cm} (5)

subject to $d(w^i, x^i+b) \geq 1$ for $i = 1, \ldots, l$.

The solution to the optimization problem from Equation (5) is given by the saddle point of the Lagrange function,

$$\Phi(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^{l} \alpha_i (y^i[<w, x^i>+b]-1),$$  \hspace{1cm} (6)

where $\alpha$ denotes the Lagrange multipliers.

The hard classifier is then given by

$$f(x) = \text{sgn}(<w^*, x>+b).$$  \hspace{1cm} (7)

The approach here described for a linear SVM can be extended for the creation of a nonlinear SVM to classify linearly inseparable data (Cristianini and Shawe-Taylor, 2000).

**System for recognizing stress levels in hens**

The proposed automatic stress recognition system is composed of three modules (Figure 1): the preprocessor, the feature generator, and the stress detector and classifier. During preprocessing, the real sounds of laying hens are obtained from an audio sensor or a CCTV camera. During feature generation, various acoustic sound features (from
both time and frequency domains) are first extracted from the recorded sounds emitted by the chickens. Subsequently, the optimal acoustic feature subsets are selected by means of the CFS algorithm. This process is unnecessary during the on-line process in the proposed real-time system. The third module invokes the stress detector and classifier to detect the stress sounds and classify them hierarchically into subsidiary sound types, such as the sounds associated with physical stress resulting from low and high temperatures and mental stress as a result of fear. In this study, we used a multi-class hierarchical SVM with a one-against-all classification structure. Figure 2 shows the overall architecture of the SVM-based stress recognition system.

RESULTS AND DISCUSSION

In our experiments, we used 407 temperature-induced sound samples (149 at 10°C±2 and 258 at 34°C±2), 114 fear-induced samples, and 136 normal sound samples (from the thermal comfort zone at 21°C±2). Figure 3 shows the respective waveforms and spectrograms for normal and stressed sound samples for Korean laying hens using Praat 5.3.52 (Boersma, 2002). In spectrograms, the amplitude of a frequency is coded by increasingly darker shades of grey. The different amplitudes of certain frequency ranges are the result of resonance- and filter-properties from the vocal tract (Schrader and Hammerschmidt, 1997). We extracted various acoustic features (in both time and frequency domains) using Praat from actual hens’ vocalizations. To select the optimal acoustic feature subset, we used the CFS from Weka 3.6 (http://www.cs.waikato.ac.nz/ml). The acoustic- feature subset obtained is (F1, F3, RMS, Mean Pitch, Max. Pitch, Shimmer, Jitter, PSD38). Notice that the dimension of the selected optimal feature subset is reduced from 54 to 8. The proposed system was realized by using a PC (3.5 GHz Intel core i7, 8 GB memory), and Weka 3.6 was used for the sequential minimal optimization to solve the SVM. Furthermore, we used a ten-fold cross validation in our experiments.

For the performance evaluation of the proposed method, we used three important formulae to measure the detection accuracy: the stress detection rate (SDR), false positive rate (FPR), and false negative rate (FNR). The formulae are given as follows (Han et al., 2012):

\[ SDR = \frac{\sum_{i=1}^{m} T_i}{\sum_{i=1}^{m} I_i} \times 100 \]  
(8)

\[ FPR = \frac{\sum_{i=1}^{m} P_i}{\sum_{i=1}^{m} N_i} \times 100 \]  
(9)

\[ FNR = \frac{\sum_{i=1}^{m} F_i}{\sum_{i=1}^{m} T_i} \times 100 \]  
(10)
In the above equations, \( I \) is an individual stress sound sample, and \( N \) is a normal sound sample. \( T \) represents stress sound samples that are correctly classified as such by the system, \( P \) indicates normal sound samples that are misclassified as stress samples, and \( F \) indicates stress sound samples that are misclassified as normal.

Our experimental results show that the SDR for the proposed system is 96.2\%, with the FPR and FNR averaging 9.6\% and 3.8\%, respectively. The detector in this experiment is identified as SVM 1 in Figure 2. We used a Puk-kernel, and the trade-off constant \( C \) was set at 4.5 in this experiment. A summary of the detection results is provided in Table 1.

Furthermore, we classified the stress vocalizations from the laying hens into three hierarchical subsidiary sound types: physical stress as result of a low temperature, physical stress as a result of a high temperature, and mental stress from fear. To measure the classification accuracy of

the proposed system, precision and recall were used as performance measurements (Han et al., 2012) with their respective formulae as follows:

\[
\text{Precision} = \frac{TP}{TP + FP} \times 100,
\]

\[
\text{Recall} = \frac{TP}{TP + FN} \times 100
\]

Table 1. Performance measurement for stress detection

<table>
<thead>
<tr>
<th>Stress detectors</th>
<th>Optimal subset: (F1, F3, RMS, Mean pitch, Maximum pitch, Shimmer, Jitter, PSD38)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM 1 (C = 4.5)</td>
<td>SDR (%), FPR (%), FNR (%)</td>
</tr>
<tr>
<td></td>
<td>96.2, 9.6, 3.8</td>
</tr>
</tbody>
</table>

RMS, root mean square; PSD, power spectral density; SDR, stress detection rate; FPR, false positive rate; FNR, false negative rate; SVM, support vector machine; C, trade off constant.
In a given class, the number of correctly classified objects is the number of true positives (TP). The number of falsely identified objects is the number of false positives (FP). The number of objects from a class that are falsely labeled as belonging to another class is the number of false negatives (FN). Precision refers to the ratio of TP to the sum of TP and FP. This determines the number of correctly identified objects. Recall is the ratio of TP to the sum of TP and FN. This determines the number of misclassified objects in a class.

Our experimental results show that the average classification accuracy for the proposed system is 96.7%, with precision and recall averaging 96.7% and 97.1%, respectively. The classifiers in this experiment are identified as SVM 2 and 3 in Figure 2. We used a normalized polynomial kernel, and the trade-off constant C was set at 3.8 and 4.5 for SVM 2 and SVM 3, respectively. The classification results are provided in Table 2.

By way of summarizing our experimental results, an optimal acoustic feature subset (F1, F3, RMS, Mean pitch, Maximum pitch, Shimmer, Jitter, PSD38) was selected from the sound emitted by hens in a coop. Incidentally, this represents the first algorithmic attempt to find useful acoustic features in the feature subset space to recognize the stress calls of laying hens, as far as we know. The stress detection accuracy of our proposed system was 96.2%, with the FPR and FNR averaging 9.6% and 3.8%, respectively. In addition, the average classification accuracy of the proposed system was 96.7%, with precision and recall averaging 96.7% and 97.1%, respectively. Even sound data acquired with an inexpensive microphone can detect stress accurately and efficiently without causing any additional stress to the laying hens. Moreover, our method can be used in a commercial poultry production farm, either as a standalone solution or to complement other known methods. To the best of our knowledge, this system has never been investigated before. Moreover, this study might confirm that an analysis of laying hens’ sounds is a creditable method for understanding the current health condition of livestock.

It is well known that the efficiency of poultry production can be adversely affected by high ambient temperature. Some studies have further reported that decreases in the environmental temperature (i.e., stress resulting from the cold) negatively influence some indices of the performance and circulatory systems in chickens (Blahová et al., 2007). In addition, Elrom (2000) noted that high levels of fear adversely affect bird plumage, egg production, egg shell quality, growth, and feed conversion efficiency. Therefore, it is of considerable importance to detect and classify stress into subsidiary sound types such as the physical stress resulting from changes in temperature, and the mental stress resulting from fear. Finally, even were a new stress class to emerge, it could be easily adapted for incremental updating and scaling in our proposed system without reconstructing the entire system.

### CONCLUSION

Early detection of health anomalies is an important issue in the management of group-housed livestock. In particular, failure to detect stress in laying hens in a timely and accurate manner can be a serious and limiting factor for achieving efficient reproductive performance. In this study, we developed a low-cost, non-invasive, and automatic online prototype that monitors the vocalizations in a commercial poultry facility to notify the producer of stressful situations in the coop. Offline, the proposed system preprocesses an optimal acoustic-feature subset (F1, F3, RMS, Mean pitch, Maximum pitch, Shimmer, Jitter, PSD38) from the sound emitted from laying hens. On-line, a recognition module detects the stress sounds and classifies them hierarchically into subsidiary sound types such as physical stress from changes in temperature and mental stress resulting from fear. In our experiments, we found that the stress detection accuracy of the proposed system is 96.2%, and the stress classification measures were satisfactory.

### ACKNOWLEDGMENTS

This research was supported by a Korea University grant.

### REFERENCES


---

**Table 2. Performance measurement for stress classification**

<table>
<thead>
<tr>
<th>Stress classifier</th>
<th>Class</th>
<th>Optimal subset:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>(F1, F3, RMS, Mean pitch, Maximum pitch, Shimmer, Jitter, PSD38)</td>
</tr>
<tr>
<td>SVM 2 (C = 3.8)</td>
<td>10°C±2</td>
<td>99.3</td>
</tr>
<tr>
<td>SVM 3 (C = 4.5)</td>
<td>34°C±2</td>
<td>97.7</td>
</tr>
<tr>
<td></td>
<td>Fear</td>
<td>93.9</td>
</tr>
<tr>
<td></td>
<td>Avg.</td>
<td>96.7</td>
</tr>
</tbody>
</table>

RMS, root mean square; PSD, power spectral density; SVM, support vector machine; C, trade off constant; Avg., average.


Han, J., M. Kamber, and J. Pei. 2012. Data Mining: Concepts and Techniques, 3rd edition, Morgan Kaufmann, Burlington, MA, USA.


