Journal of Intelligent Learning Systems and Applications, **2010, 2, 221-228**
doi:10.4236/jilsa.2010.24025 Published Online November 2010 (http://www.scirp.org/journal/jilsa)

Scientific Research

A Feasible Approach for Automatic Detection and Recognition of the Bengalese Finch Songnotes and Their Sequences

Khan Md. Mahfuzus Salam[1,*], Tetsuro Nishino[1], Kazutoshi Sasahara[2], Miki Takahasi[2], Kazuo Okanoya[2]

[1]Department of Information and Communication Engineering, the University of Electro-Communications, Tokyo, Japan; [2]Laboratory for Bioliguistics, RIKEN Brain Science Institute (BSI), Saitama, Japan. *corresponding author
E-mail: mahfuz@ice.uec.ac.jp

Received March 23rd, 2010; revised September 14th, 2010; accepted September 18th, 2010.

ABSTRACT

The Bengalese finch song has been widely studied for its unique features and similarity to human language. For computational analysis the songs must be represented in songnote sequences. An automated approach for this purpose is highly desired since manual processing makes human annotation cumbersome, and human annotation is very heuristic and easily lacks objectivity. In this paper, we propose a new approach for automatic detection and recognition of the songnote sequences via image processing. The proposed method is based on human recognition process to visually identify the patterns in a sonogram image. The songnotes of the Bengalese finch are dependent on the birds and similar pattern does not exist in two different birds. Considering this constraint, our experiments on real birdsong data of different Bengalese finch show high accuracy rates for automatic detection and recognition of the songnotes. These results indicate that the proposed approach is feasible and generalized for any Bengalese finch songs.

Keywords: *Birdsong Analysis, Bengalese Finch Song, Songnote Detection and Recognition, Pattern Recognition*

1. Introduction

Birdsong has been actively studied via analysis of songnote sequences to understand the language model of birds. The songs of the Bengalese finch (*Lonchura striata var. domestica*) – a popular fowl in Japan, is widely employed for this purpose. The song of the Bengalese finch has a complex structure as compared with those of other songbirds such as zebra finches (*Taeniopygia guttata*) [1]. Thus, Bengalese finch songs have been studied as a model of human language. According to the recent studies, the courtship songs of Bengalese finches have unique features and similarity with a human language [2]. In birdsong research, acoustic song analysis is necessary to find the song elements and their sequence for carrying out an analysis to understand the song syntax [3] and the learning process of the song. The current research is focused on automatic detection and recognition of the songnote and its sequence. Previous studies that employed sound processing had drawbacks as an automated approach. This paper introduces a new generalized approach that employs image processing to overcome the drawbacks.

2. Preliminaries

This section briefly introduces the theoretical foundations of a birdsong, its representations, image basics, and the recognition process by humans as we focused on the recognition process that is manually carried out by humans.

2.1. Birdsong Representation

In birdsong analysis, the song data is recorded in an appropriate environment – special cage equipped with automated recording system and also to avoid noise. From the recorded sound data, we obtain the sonogram image of the song. For further computational analysis, the obtained sonogram image is used as the standard representation of the song [4].

The following of this section briefly explains some general terms that are used in birdsong research.

Songnote: An independent pattern appearing in sonogram which is assigned a symbol is called a *songnote*. It is also referred as a *song element* or a *behavioral element*. From the definition, we can say the text data consisting of symbols (such as *a*, *b*, *c*, and so on) are called songnote sequence. Songnotes are analogous to phonemes in human language.

Chunk: A fixed sequence of song notes is called a *chunk*. In **Figure 1**, for example, the chunks are *ab*, *cde* and *fg*. Chunks are analogous to words in human language.

Song unit: A *song unit* consists of chunks. Song units are analogous to sentences in human language.

Sonogram: A sonogram is an image that shows how the spectral density of a signal varies with time. It is also known as a spectrogram, voiceprint, or voicegram. Sonogram are used to identify phonetic sounds to analyze the animal cries and also in the fields of speech processing, music, sonar/radar, seismology, etc.

There are many variations in the format of the sonogram. Sometimes, the vertical and horizontal axes are switched; sometimes, the amplitude is represented as the height of a 3D surface instead of color or intensity. The frequency and amplitude axes can be either linear or logarithmic, depending on what the graph is being used for. For instance, audio would usually be represented with a logarithmic amplitude axis, and frequency would be linear in order to emphasize harmonic relationships, or logarithmic to emphasize musical, tonal relationships. The most common format is a graph with two geometric dimensions: the horizontal axis represents time, and the vertical axis is frequency; a third dimension indicating the amplitude of a particular frequency at a particular time is represented by the intensity or color of each point in the image. For the birdsong research this common format is used. **Figure 1** shows a sample grayscale sonogram image of a Bengalese finch courtship song.

2.2. Bengalese Finch Song

Recent studies on Bengalese finches show that the songs of male Bengalese finches are neither monotonous nor random; they consist of chunks, each of which is a fixed

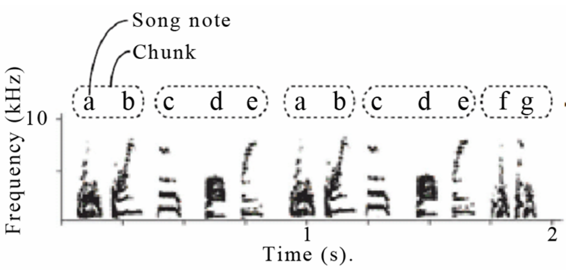


Figure 1. Grayscale sonogram image of a Bengalese finch song.

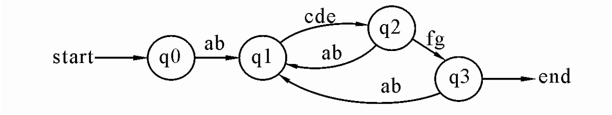


Figure 2. Courtship song syntax represented by an automaton

sequence of a few song notes. The song of each individual can be represented by a finite automaton, which is called song syntax (see **Figure 2**) [2]. The songs of Bengalese finches have double articulation – a sentence consists of words, and each word consists of phonemes, which is also one of the important faculties of human language.

The song syntax is manipulated by the song control nuclei in the brain. The hierarchy of the song control nuclei directly corresponds to the song hierarchy [5]. Because of the structural and functional similarities of vocal leaning between songbirds and humans, the former have been actively studied as a good model of a human language [6]. In particular, the song syntax of Bengalese finches sheds light on the biological foundations of syntax.

2.3. Detection and Recognition

Human vision is one of the most important and perceptive mechanisms. It provides information required for the relatively simple tasks (e.g., object recognition) and for very complex tasks as well. In bird song research, the songnote recognition is carried out by humans by inspecting the patterns visually represented in a sonogram image [4].

2.3.1. Image Feature Extraction

Digital image processing denotes the analysis carried out on the basis of the pixel property of the image irrespective of the image type. A digital image has a finite set of digital values called picture elements or pixels. The image contains a fixed number of rows and columns of pixels. Pixels are the smallest individual elements in an image, holding quantized values that represent the brightness of a given color at any specific point. Typically, the pixels are stored in computer memory as a raster image or raster map, a two-dimensional array of small integers. These values are often transmitted or stored in a compressed form.

Each pixel of a raster image is typically associated with a specific position in some 2D region and has a value of one or more quantities related to that position. Digital images can be classified according to the number and nature of such samples into different categories like *Binary*, *Grayscale*, *Color* and *False-color*. In our research, we use a grayscale sonogram image.

Grayscale Image: A grayscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. In fact a gray color is one in which the red, green, and blue components all have equal intensity in the RGB space, and hence, it is only necessary to specify a single intensity value for each pixel, as opposed to the three intensities needed to specify each pixel in a full color image.

Pixel Values: For a grayscale image, the pixel value is a single number that represents the brightness of the pixel. The intensity of a pixel is expressed within a given range between a minimum and a maximum. Presently, grayscale images are commonly stored with 8 bits per sampled pixel, which allows 256 different intensities (*i.e.*, shades of gray). The binary representations assume that 0 is black and the maximum value 255 is white.

2.3.2. Image Matching and Recognition

Pattern recognition aims to classify data or patterns on the basis of either a priori knowledge or statistical information extracted from the patterns. The patterns to be classified are usually groups of measurements or observations, defining points in an appropriate multidimensional space. This is in contrast to pattern matching, where the pattern is rigidly specified. Pattern recognition is used to test whether things have a desired structure, to find relevant structure, to retrieve the aligning parts, and to substitute the matching part with something else.

In human vision-based recognition of an image, the first thing that will catch the attention is something that is familiar. To be recognized, an object must have some feature that our consciousness can assign. Behind this process, the mental model captures the important characteristics of the object. It is unfortunate that, in many scientific experiments, the task assigned to human vision is not the recognition of familiar objects, but the detection and description of unfamiliar ones, which is far more difficult. According to the McCulloch and Pitts simplified neuron model, the weighted sum of many inputs exceeds a threshold, and then the output is turned on. (see **Figure 3**). Learning consists of adjusting the weights, which can be either positive or negative [7].

The current research applies image processing methodology based on grayscale image features of the sonogram. The motivation of applying such image processing is to find a simple and generalized way for the automation as a human brain does in the recognition process by applying pattern matching.

3. Methodology

The proposed automation process is divided into two steps. First, from the song sonogram image, we detect the song elements on the basis of the local property of the sonogram image. Then, on the basis of the detected elements, we apply image matching to assign a label to the extracted elements, and thus, we obtain the songnote sequence of the song. **Figure 4** shows the process flow diagram of the proposed methodology.

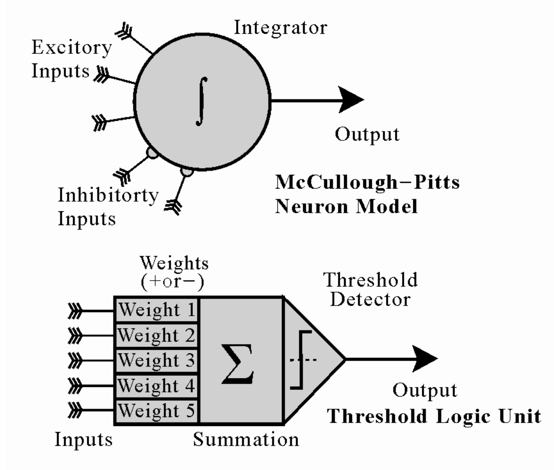


Figure 3. McCulloch and Pitts simplified model of a neuron and its implementation as a threshold logic unit [7].

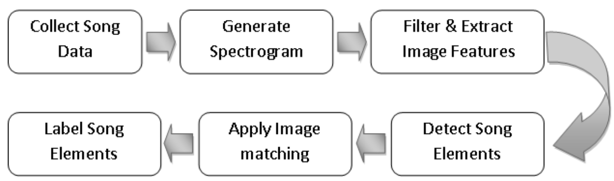


Figure 4. Process flow diagram of the songnote detection and recognition.

3.1. Songnote Detection

From the sonogram image, we first detect the elements. On the basis of the extracted statistical features of the detected elements, we carry out the recognition process. For this reason, the detection process is very important.

3.1.1. Detection Method

The detection process is carried out by analyzing the sonogram image for intensity values; we can obtain a graph for the average pixel intensity value. If the sonogram image has many noises at the beginning, which are ignored in the visual inspection by human, the present system does not ignore them as noises. For this reason, we pre-process the sonogram image. Then, if we take the average intensity value along the vertical line and draw a graph where the Y-axis represents the average intensity value or gray value and the X-axis represents the pixel index x, which is the distance from the (0, 0) pixel along the X-axis, we have a graph as follows:

The above graph (see **Figure 5**) is generated from the sample sonogram image shown in **Figure 6**. It is clearly visible that from the graph we can find some clear gaps between the elements. By defining parameters (see **Figure 6**) such as minimum element width, minimum gap between elements, and the intensity threshold, we can execute our algorithm to find the song elements. If some region does not fit with the three above mentioned parameters, we consider it to be noise. Note that these parameters can vary from bird to bird. The detected song elements and the features of the elements, such as width information, are used for the recognition process.

3.1.2. Detection Algorithm

The song element detection algorithm takes the array of the average intensity values as the input. On the basis of the defined parameter values, the proposed detection algorithm produces an unlabeled list of song elements.

Detection Algorithm

Input: array of intensity values.
Output: a list of elements.
Procedure:
(1) Initialize the parameters.
(2) If the intensity value exceeds threshold and next is not a gap
- set start element flag true;
- set start index to current index;

(3) If start element flag is true and next minimum gap is detected
- set start element flag false;
- set end index to current index;
- add to element list;

(4) Continue step 2 and 3 until end of the intensity array
(5) Return element list.

3.2. Songnote Recognition

For extracting the songnote sequence from the sonogram image, we extract local statistical features and then carry out the statistical pattern matching for recognition.

3.2.1. Recognition Method

As discussed in the previous section, similar patterns are assigned with the same label in the recognition process. Our recognition method is based on the local property of the sonogram image. By executing the note detection algorithm, we obtain element list information. This unlabeled element list provides the start pixel and the end pixel information for every element.

As for the Bengalese finch song, note patterns differ from bird to bird. Therefore, we decided not to use any prior knowledge; rather, we use the statistical information extracted from the patterns. See **Figure 7**. First, we divide every note into N regions, and every region is

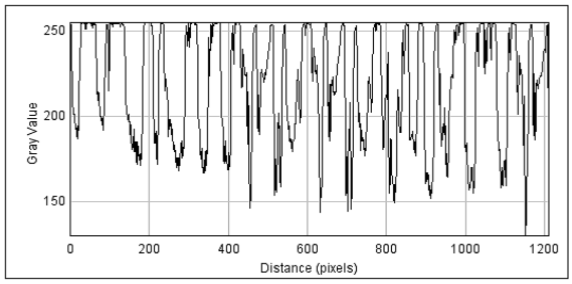


Figure 5. Average intensity value graph derived from the sonogram image.

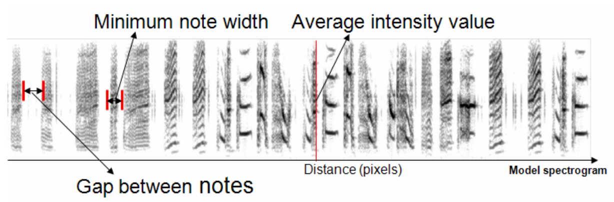
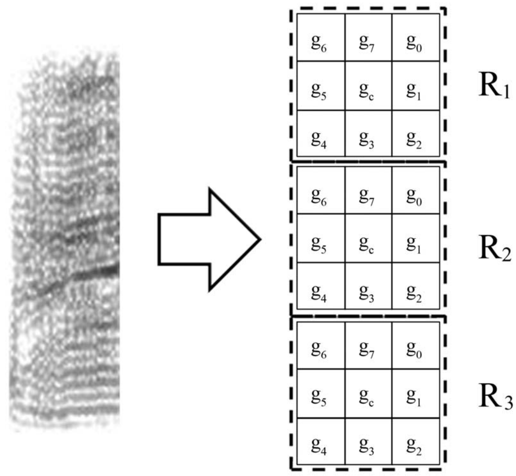


Figure 6. Sample sonogram image and the parameters.



$$R_1 \cup R_2 \cup R_3 = \{R_1 g_0, ..., R_1 g_7, R_1 g_c, R_2 g_0, ..., R_2 g_7, R_2 g_c, R_3 g_0, ..., R_3 g_7, R_3 g_c\}$$

Figure 7. Explains the procedure while N = 3.

divided into nine (3×3) cells. We denote the center cell as g_c and the other cells as g_n in a clockwise direction, where n = 0, 1, …, 7. Thus, we obtain a set of values for every single element. Then, we apply a statistical test called the chi-square test to find the similarity between elements. Note that the value of N should not be greater than 3 because if the set size exceeds thirty, the Chi-square distribution tends toward a normal distribution.

3.2.2. Chi-square Goodness Fit Test

The chi-square test (χ^2) is a statistical hypothesis test whose results are evaluated by reference to the chi-square distribution. Pearson's chi-square test is the best-known of several chi-square tests. Its properties were first investigated by Karl Pearson [8]. Pearson's chi-square is the original and most widely-used chi-square test.

When an analyst attempts to fit a statistical model to observed data to find how well the model actually reflects the data, *i.e.*, how *close* are the observed values to those that would be expected under the fitted model, one statistical test that addresses this issue is the chi-square goodness of fit test. This test is commonly used to test the association of variables in two-way tables, where the assumed model of independence is evaluated against the observed data. In general, the chi-square test statistic is of the following form:

$$\chi^2 = \sum \frac{(Observed - Expected)^2}{Expected}$$

In the equation, χ^2 is of the form:

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

where,

χ^2 = the test statistic that asymptotically approaches a χ^2 distribution.

O_i = an observed frequency;

E_i = an expected frequency, asserted by the null hypothesis;

n = the number of possible outcomes;

The chi-square statistic is calculated by finding the difference between each observed and theoretical frequency for each possible outcome, squaring these values, dividing each by the theoretical frequency, and taking the sum of the results. The chi-square statistic can be used for calculating a p-value by comparing the value of the statistic to a chi-square distribution. The number of degrees of freedom is equal to the number of possible outcomes minus 1. If the computed test statistic is larger than the chi-square table [8] with (n – 1) degrees of freedom, the observed and expected values are not close and the model is a poor fit to the data.

For pattern recognition one famous statistical machine learning approach is Support Vector Machines (SVM). SVM separates the data space into two clusters over a separation boundary defined by a non-linear function [9]. We can apply SVM when supervised learning is possible and also the number of clusters is known. It is difficult to apply this technique in our application where there are several cluster exist which cannot be predefined. During our experiment we also try to employ another image pattern recognition technique presented by Ojala *et al*. [10] but unfortunately, we could not obtain good result. As the songnote patterns are dependent to the birds we have limitations for preparing the training set. For that reason, we find that chi-square test is suitable for our application comparing to other state-of-the-art techniques for pattern recognition.

3.2.3. Recognition Algorithm

The songnote recognition algorithm takes the unlabeled list of song elements. It applies the goodness of fit test to find the similarity between elements and produces the songnote sequence.

Recognition Algorithm

Input: unlabeled list of elements.
Output: labeled list of elements.
Procedure:
(1) For each element in element list divide into N × 9 cells where 0 < N < 4
(2) Calculate the average intensity value for every cell
(3) For each element until there is any unlabeled element
- set one as expected and others as observed;
- if expected is not labeled set it with a new label;
- test the Chi-square statistics;

(4) If the observed element pass the test then set the element with same label
(5) Return updated element list

4. Results

In this section, we present the results of our methodology for analyzing the Bengalese finch song. First, we explain the nature of our real song data, and then discuss the results of the automatic detection and recognition of the songnote.

4.1. Description of Data

For testing our proposed method we use five different song unit or phrase for each of the three matured Bengalese finch song; the names of the finches are *Hikari* 52, *Hikari* 49 and *Kuro* 0362. The song data were recorded at the *Okanoya laboratory* of RIKEN. The spectrogram image of the matured Bengalese finch has similar properties, where the note patterns are clearly visible and almost each songnote is separated by considerable blank space. **Figure 8** shows the partial sonogram images for the three birds.

Table 1 shows the sample sonogram images contains forty six to fifty four notes for *Hikari* 52, fifty one to fifty nine notes for *Hikari* 49 and fifty two to sixty one notes for *Kuro* 0362. From the sample sonogram image **Figure 8**, it is clearly visible that the sonogram image of

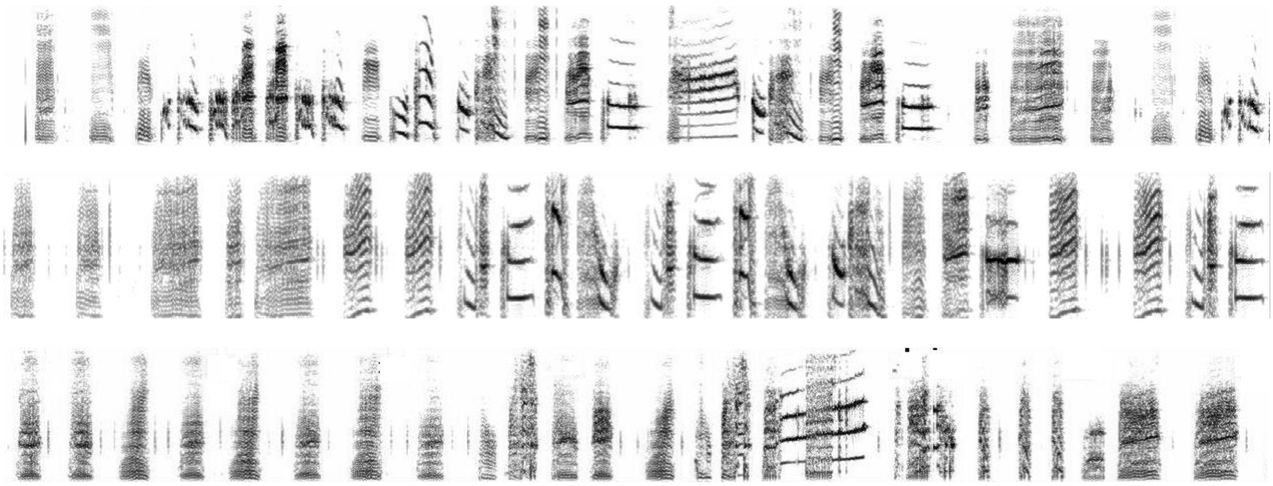


Figure 8. Spectrogram for *Hikari* 49 (top), *Hikari* 52 (middle) and *Kuro* 0362 (bottom).

Table 1. Results of automatic detection of song elements.

Bird name	Number of appeared elements	Average accuracy rate
Hikari 52	46 - 54	98%
Hikari 49	51 - 59	90%
Kuro 0362	52 - 61	95%

Hikari 49 is more complex than that of *Hikari* 52 and *Kuro* 0362, *i.e.*, for *Hikari* 52 and *Kuro* 0362, the song notes are almost clearly separated from one to another, but for *Hikari* 49, the song notes are not clearly separated from one another.

By applying our methodology, we implemented an application in JAVA, which takes the sonogram image as an input and provides extracted song elements and their sequence as the output. *ImageJ API* [11] is used for analyzing the image property.

4.2. Songnote Detection

In Section 3.1, we discussed the song note extraction methodology and explained the algorithm used for extracting the song notes from a sonogram image. We used parameters such as minimum note width, intensity threshold, and minimum gap between notes. We set the parameter values for minimum note width as 10 pixels, intensity threshold as 250, and minimum gap between notes as 5 pixels for every bird. After executing the algorithm mentioned in Section 3.1.2, we obtain the result for the best case as follows:

Now in the case of *Hikari* 52, when we inspect the extracted patterns, we find that there are some noises with the extracted patterns although we have a good accuracy rate. To avoid the noise, if we apply a cutoff level of 30 at the intensity value graph, we obtain 40 extracted elements while the original numbers of elements are 46. Therefore, the accuracy rate decreases, and certain elements lose some necessary information, which is not desirable. **Figure 9** describes the noise situation.

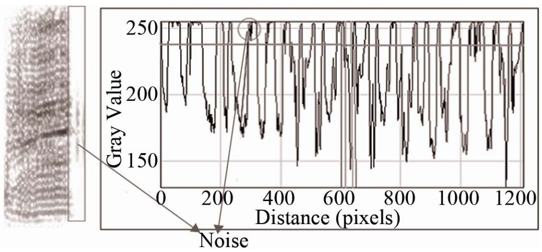


Figure 9. Description of noise and effect of applying cutoff level for *Hikari* 52.

In the case of *Hikari* 49, when we inspect the extracted patterns, we find that some song notes are not extracted correctly. Initially, we have an accuracy rate of 75% with our default parameter value as the gaps between the elements are too short to separate. **Figure 9** describes the errors in the detection process.

Figure 10, except **Figure 10(d)**, shows some incorrect extracted notes for *Hikari* 49. If we carefully inspect **Figure 10**, we can observe that **Figure 10(a)** and **Figure 10(b)** should be extracted as two different elements because the right pattern in **Figure 10(a)** and **Figure 10(b)** appears separately (see **Figure 10(c)**) in the sonogram image, and **Figure 10(c)** should be extracted as three different elements However, **Figure 10(d)** is considered to be extracted as a right pattern although it has the same nature as the patterns shown in **Figure 10(a**, **b**, and **c)** because the two patterns are very close and the left and the right patterns do not appear separately in the song.

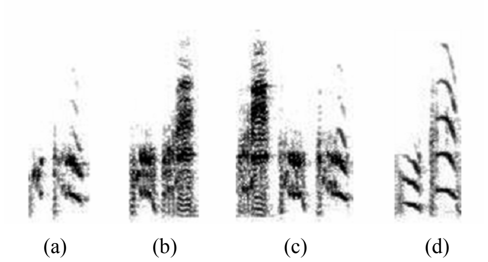


Figure 10. Description of the error in the detection for *Hikari* 49.

Table 2. Results of the song note recognition.

Bird name	Accuracy rate
Hikari 52	86%
Hikari 49	85%
Kuro 0362	78%

We adjust the default parameter value of the minimum gap between the notes to be two pixels and use the cutoff level of nine. Thus, for the best case result we obtain an accuracy rate of 90%.

4.3. Songnote Recognition

In Section 3.2, we discussed the songnote recognition methodology and explained the algorithm. The first step is to divide every extracted element into N parts, and then calculate the average intensity value for every region. Thus, for every element, we have a set of 27 element while N = 3. Then, we apply the Chi-square test considering the note width information. In the proposed method, we compare the elements if the note width is greater than three-fourths or smaller than five-fourths of the observed element. After executing the algorithm mentioned in Section 3.2.3, we obtain the songnote sequences.

We can summarize the result for the recognition as follows (See **Table 2**):

Notice that for *Hikari* 49, the result is based on extracted patterns in the previous step. If we consider the wrong extracted pattern, then the accuracy rate become around 70%.

For further discussion the songnote sequence of one song unit that is produced by our system and the sequence by human annotation for *Hikari* 52 have been shown below where the bold letters show the different outcomes in recognition.

System (*Hikari* 52):

*AAB***AC***DDEFGHEFGHIBJK***L***DEFAAB***AC***D-DEFGHEFGHI***B***JK***L***DEF*

Correct (*Hikari* 52):

*AAB***LB***DDEFGHEFGHICJK***D***DE-FAAB***LB***DDEFGHEFGHI***C***JK***D***DEF*

If we inspect the wrong decisions made by the system for *Hikari* 52, we find that note *B* is labeled as *C* and note *L* is labeled as *D*. This is because the incorrectly labeled note contains a considerable noise (white part), which affects the matching process. In the case of incorrectly labeling note *L* note *A* for *Hikari* 52, by carefully observing each note, we find that the intensity density is the same for both the notes (see **Figure 11**).

From **Figure 11** it is clearly visible that the distribution of intensity density is the same for both the notes. This causes the recognition error and is a limitation of the proposed image matching algorithm. Notice that *Note* 1 and *Note* 4 are recognized as *A*, but originally by human annotation by inspecting the image and hearing to the song, *Note* 4 was labeled *L*. We notice a similar recognition error in the case of *Hikari* 49 and *Kuro* 0362.

5. Conclusions and Discussion

The present study proposes a brand-new approach to automatic recognition of song elements and its sequences other then sound processing, and by applying image processing, we obtain good results for the approach. There are good possibilities to improve the accuracy rate for both the extraction and the recognition methods to some extent. From the obtained results, we find that the element extraction process is very important and has a significant effect on the recognition process. The major advantage of the proposed approach is its simplicity and feasibility. The approach is focused on a generalized (does not depend on the bird) process same as humans do.

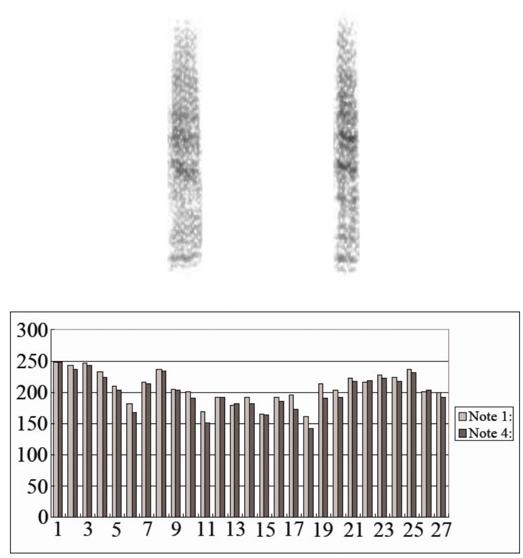


Figure 11. Note 1 (A, top left), note 4 (L, top right) and distribution of intensity density value (bottom) for *Hikari* 52.

Table 3(a). Comparison results for automatic detection.

Bird name	accuracy rate (sound processing)	accuracy rate (proposed method)
Hikari 52	96%	98%
Hikari 49	94%	90%

Table 3(b). Comparison results for automatic recognition.

Bird name	accuracy rate (sound processing)	accuracy rate (proposed method)
Hikari 52	83%	86%
Hikari 49	--*	85%

* Satisfactory clustering was not possible for *Hikari* 49 based on the parameter used.

The accuracy rate of the proposed approach is better than that of other methods such as sound processing which was previously carried out at our laboratory. The following tables show comparison of the accuracy rate between sound processing and our proposed method for sngnote detection (see **Table 3(a)**) and recognition (see **Table 3(b)**). For comparison we use the song data of Hikari 52 and Hikari 49.

For the detection process sound processing method uses *Amplitude* and *Wiener Entropy*. For the recognition process it applies K-means clustering algorithm which uses *Duration*, *Amplitude, Wiener Entropy, Mean Frequency* and *Harmonic Pitch* as parameters. However, the sound processing requires considerable human effort for fixing the parameter values (manual labeling has to be done once if k-means algorithm has been applied) or for training the system (each songnote has to be manually separated to build the database if HMM approach has been applied) for detecting and recognizing the songnotes for every bird. Furthermore, it is not possible to make a corpus for bird phonemes. If we employ sound processing the only thing we can do is to train the system for a specific bird family as similar patterns do not appear in different bird families. This is not practical for an automated system. In contrast, the proposed methodology is almost automated and feasible for songbirds as our approach represents the human inspection method and does not depend on birds. The default parameter values have been used for detecting the songnotes is almost good for any bird but can be changed by couple of click by the user if necessary.

For the element detection process, the accuracy rate is 100% for some birds, and for other birds, the accuracy rate is also satisfactorily high. Thus, our approach saves time and is practical as an automated system. In the recognition process, we obtain a high accuracy rate of more than 80%.

REFERENCES

[1] E. Honda and K. Okanoya, "Acoustical and Syntactical Comparisons between Songs of the White-Backed Munia (Lonchura striata) and Its Domesticated Strain, the Bengalese Finch (Lonchura Striata Var. Domestica)," *Zoological Science*, Vol. 16, 1999, pp. 319-326.

[2] K. Okanoya, "Song Syntax in Bengalese Finches: Proximate and Ultimate Analyses," *Advances in the Study of Behavior*, Vol. 34, 2004, pp. 297-346.

[3] Y. Kakishita, K. Sasahara, T. Nishino, M. Takahasi and K. Okanoya, "Ethological Data Mining: an Automata-Based Approach to Extract Behavioral Units and Rules," *Data Mining and Knowledge Discovery*, Vol. 18, No. 3, 2009, pp. 446-471.

[4] C. K. Catchpole and P. J. B. Slater, "Bird Song: Biological Themes and Variations," 2nd Edition, Cambridge University Press, UK, 2003.

[5] J. Nishikawa and K. Okanoya, *Ornithological Science*, 2006, pp. 95-103.

[6] J. Doupe and P. K. Kuhl, "Birdsong and Human Speech: Common Themes and Mechanisms," *Annual Reviews Neuroscience*, Vol. 22, 1999, pp. 567-631.

[7] J. C. Russ, "The Image Processing Handbook," 5th Edition, CRC Press, 2006.

[8] S. M. Ross, "Introduction to Probability and Statistics for Engineers and Scientists," 3rd Edition, Elsevier Academic Press, USA, 2004.

[9] J. Shawe-Taylor and N. Cristianini, "An Introduction to Support Vector Machines and Other Kernel-Based Learning Methods," 1st Edition, Cambridge University Press, UK, 2000.

[10] T. Ojala, M. Pietikainen and D. Harwood, "A Comparative Study of Texture Measures with Classification Based on Feature Distributions," *Pattern Recognition*, Vol. 29, No. 1, 1996, pp. 51-59.

[11] ImageJ 1.41, National Institutes of Health, USA, 2010. http://rsbweb.nih.gov/ij/

