Identification of European woodpecker species in audio recordings from their drumming rolls

Juliette Florentin [a,*], Thierry Dutoit [b], Olivier Verlinden [a]

[a] Department of Theoretical Mechanics, Dynamics and Vibrations, University of Mons, Mons, Belgium
[b] Department of Circuit Theory and Signal Processing, University of Mons, Mons, Belgium

ARTICLE INFO

ABSTRACT

Drumming sounds are substantial clues when searching audio recordings for the presence of woodpeckers. Woodpeckers use drumming for territory defence and mate attraction to such an extent that some species have no vocalisations for these functions. This implies that drumming bears species markers. This hypothesis stands at the root of our project to develop an autonomous program for the identification of drumming species. To proceed, we assembled a database of 361 recordings from open-access bird sound archives. The recordings were for nine drumming species found on the European continent. Focusing on the signal below 1500 Hz, we reviewed all audio files and extracted 2665 drumming rolls. For recordings where multiple birds were present, the drumming rolls were attributed to individual birds. This allowed keeping track of the time interval between successive rolls. The characteristic traits of drumming are decidedly temporal. Consequently, the spectral features that have been successful in other recent bird identification studies were not applicable in our case. We developed specialized drumming parameters and automated their calculation. We then performed a t-SNE dimensionality reduction to visualise the dataset and to demonstrate that our parameters detached the different classes properly. Eventually, a k-NN algorithm accurately labelled 87.2% of the submitted test samples. The time structure within the drumming rolls (speed, acceleration) provided the critical features. The duration of the rolls followed in importance. The results compare well to existing literature and attest to the feasibility of monitoring European woodpecker species by tracking drumming.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The continuous acoustic monitoring of wildlife and birds in particular has generated terabytes of data (Aide et al., 2013; Jahn et al., 2013; Towsey et al., 2014; Florentin et al., 2015), of which a large part is still awaiting exhaustive processing. The dormant information pertains to the presence or absence of species within certain areas, the evolution of bird communities over seasons and the potential degradation of habitats (Farina et al., 2011). Hence, in recent years, researchers have steered towards the development of robust algorithms that would be capable of identifying all species and most notably all bird species captured on audio recordings (Blumstein et al., 2011). Such algorithms are required to perform well on the two critical functions of 1) detecting bird sounds in the audio stream and 2) identifying the species emitting these sounds. The work we present in the present paper primarily addresses the second function, species identification, in a context where short extracts of bird sounds have already been made available.

It must be noted however that birdsong detection techniques sometimes inherently include the species identification step. This is the case for spectrogram cross-correlation. This well-established technique allows searching long audio recordings for one specific sound. It is well-suited to sounds that produce repeatable spectrogram patterns, such as stereotyped songs (Ulloa et al., 2016). The concept is to have a template of the target spectrogram image sliding over a continuous audio stream until a maximum in cross-correlation is reached. Swiston and Mennill (2009) used it to detect double-knocks from two species of woodpeckers (*Campephilus guatemalensis* and *Campephilus principalis*). There was no expectation that potential species traits in double-knocks would play a role and the same template was used for both species. Mean detection rates of 24% and 8% respectively were achieved, while the number of false positives was 77 times greater than with human observers.

If species identification is to be addressed separately, then its object is to automate the classification of short sound extracts lasting between a few seconds and a few minutes. Recent works in this area engaged in the identification of hundreds of bird species in collections of up to ten thousand audio files (Potamitis, 2014; Stowell and Plumbley, 2014; Lasseck, 2015). For this type of task, the choice of classifier evolved from template comparisons (Kogan and Margoliash, 1998) towards

hidden Markov models (Somervuo et al., 2006; Brandes, 2008), artificial neural networks (Fox et al., 2008; Ranjard and Ross, 2008) and finally single- or multi-label random forests (Potamitis, 2014; Stowell and Plumbley, 2014). The most common acoustic features are the Short-Time Fourier Transforms (STFT), either in full (Stowell and Plumbley, 2014) or compressed in the form of Mel-Frequency Cepstral Coefficients (MFCC) (Kogan and Margoliash, 1998; Somervuo et al., 2006; Fox et al., 2008; Ranjard and Ross, 2008; Lee et al., 2013). The time dimension is then handled by calculating the mean and standard deviation of the acoustic features over the duration of the sound extract (Stowell and Plumbley, 2014) or other statistics such as moments and regressions (Lasseck, 2015). Spectrogram images are an alternative possible basis for acoustic features (Lee et al., 2013; Potamitis, 2014; Lasseck, 2015). Potamitis (2014) and Lasseck (2015) derived their features from cross-correlation scores with a set of template images. Lasseck (2015) showed that these outperform spectral features in the classification of syllables and elements of songs. Multiple authors reported difficulties pertaining to the "variability" of songs (Kogan and Margoliash, 1998; Somervuo et al., 2006; Potamitis, 2014). The following species were offered as difficult cases: canaries (Kogan and Margoliash, 1998), Finnish blackcaps and pied flycatchers (Somervuo et al., 2006) and European tits (Potamitis, 2014). All are passerines, i.e. the order with the most elaborate songs. A typical percentage of correct identifications for passerines is 70%–80% (Somervuo et al., 2006; Brandes, 2008; Fox et al., 2008).

Overall, the success of the above classification studies is contingent on the existence of generic acoustic features that can grasp the species traits in the vocalisations of any species. On the other hand, Kogan and Margoliash (1998) suggested that their classification results would have been improved by the use of biological features. These are the features that birds themselves use to recognize their conspecifics. They are specific to species or to a group of species (Catchpole and Slater, 2008). Lasseck (2015) reselected a subset of features independently for each species to optimize the classification score. Bardeli et al. (2010) advocated redefining the features for each new species to improve the recognition rate. Ulloa et al. (2016) observed that spectrogram cross-correlation was not appropriate for all vocalisations and Lasseck (2015) that it did not render temporal structures and repetition rates in bird songs. All these comments reflect a need to fall back on differentiated features. The consequence is that it is more realistic to mine audio streams for the presence of a predetermined limited set of species than to aim at an exhaustive analysis. This sets the philosophy behind the present work: classification algorithms need to be built from the bottom-up, starting from sub-groups of species. We chose to target European woodpeckers.

2. Materials

2.1. European woodpeckers

Woodpeckers are members of the *Picidae* family within the order of the Piciformes. A list of the eleven species present on the European continent is found in Table 1.

Dendrocopos minor and *Jynx torquilla* are the smallest specimens, 14–16 cm and 16–19 cm in length, respectively. *D. martius* is the largest with a length of 45–50 cm. Hybrids exist between *D. major* and *D. syriacus*, between *Picus canus* and *P. viridis* and between *P. sharpei* and *P. viridis*. Until 2012, *P. sharpei* was a sub-species of *P. viridis* (Gorman, 2014). The reasons that make woodpeckers an interesting target for acoustic monitoring are plenty. They are valued as ecosystem keystones (Gorman, 2014) and indicators of forest health (Mikusinski and Angelstam, 1998). Some species are targeted by regional conservation programs, e.g. *D. medius* in Sweden (Pettersson, 1985) or *P. canus* in Belgium.

Woodpeckers have relatively simple vocalisations (Gorman, 2014). Their peculiarity is their use of drumming for territory marking and mate attraction (Zabka, 1980; Tremain et al., 2008). Some species altogether forego the use of vocal signals for these functions (*D. major, D. syriacus, D. leucotos, Picoides tridactylus*) (Table 1). It follows that drumming sounds have been thought to carry the species and individual information (Zabka, 1980; Dodenhoff et al., 2001). Only *J. torquilla* does not use drumming. For the other species, both sexes drum. Male-female pairs have synchronized drumming duets during the mating season. Drumming contests also occur between neighbours and males of different species. Drumming Rolls (DRs) are easily recognizable in spectrograms. The example in Fig. 1 shows the time parameters of drumming: the time between DRs, the time between strokes and the DR duration. These traits are straightforward features to classify DRs. Woodpeckers also produce isolated strokes (while foraging or digging holes) and demonstrative tapping. The latter consists in constant-speed rolls which are shorter, slower and quieter than drumming rolls (Zabka, 1980). Demonstrative tapping is hypothesized to achieve near-by communication (Zabka, 1980; Tremain et al., 2008) whereas drumming would primarily be for long-distance communication (Zabka, 1980; Stark et al., 1998). Our analysis focuses on drumming only.

Zabka (1980) ran a previous study on woodpecker drumming using 240 DRs from the same nine species as in the present work, i.e. all but *J. torquilla* and *P. sharpei*. *D. major* made up almost half of his collection (Table 8). The time intervals between strokes and the duration of rolls were measured manually on spectrograms and signal envelopes. This author rejected the time between rolls as a viable acoustic feature because of excessive variability. His findings were that *D. major* had the shortest DR, and *D. leucotos* and *D. martius* the longest. The time structure (evolution of the time between strokes) followed either a linear law or a decreasing exponential law and was a critical species trait. Some woodpeckers accelerated (e.g. *D. major*); others decelerated (e.g. *D. minor*). There was a significant difference in DR duration between *D. major* males and females, and between *D. major* male neighbours. Stark et al. (1998) performed another statistical analysis on drumming parameters for 11 woodpecker species occurring in California (3347 DRs). DR duration, number of strokes per DR, average time interval between strokes and cadence (strokes per second) were considered. Cadence was found to be the best indicator for species differentiation. 78% of all samples were correctly reclassified using their

Table 1
Woodpecker species and drumming database composition.

Index	Species	Drumming	Advertising call	Original files	Drumming rolls
1	*Dendrocopos leucotos*	Yes	No	43	248
2	*Dendrocopos major*	Yes	No	115	818
3	*Dryocopus martius*	Yes	Yes	27	84
4	*Dendrocopos medius*	Rare	Yes	3	8
5	*Dendrocopos minor*	Yes	Yes	67	832
6	*Dendrocopos syriacus*	Yes	No	3	8
7	*Jynx torquilla*	No	Yes	0	0
8	*Picus canus*	Yes	Yes	29	104
9	*Picus sharpei*	Rare	Yes	0	0
10	*Picoides tridactylus*	Yes	No	68	547
11	*Picus viridis*	Rare	Yes	6	16
			Total	361	2665