Towards the Automatic Acoustical Avian Monitoring System

Robert Wielgat[a,*], Daniel Król[a], Tomasz Potempa[a], Paweł Kozioł[b], Agnieszka Lisowska-Lis[a]

[a]*State Higher Vocational School in Tarnów, ul. Mickiewicza 8, Tarnów 33-100, Poland*
[b]*Regional Directorate for Environmental Protection in Kraków, division in Tarnów, ul. Solidarności 5-9, Tarnów 33-100, Poland*
*Corresponding author: rwielgat@poczta.onet.pl

Abstract

One of the crucial aspects of the environmental protection is continuous monitoring of environment. Specific aspect is estimation of the bird species population. It is particularly important for bird species being in danger of extinction. Avian monitoring programs are time and money consuming actions which usually base on terrain expeditions. Certain remedy for this can be automatic acoustical avian monitoring system, described in the paper. Main components of the designed system are: digital audio recorder for bird voices acquisition, computer program automatically recognizing bird species by its signals emitted (voices or others) and object-relational database accessed via the Internet. Optional system components can be: digital camera and camcorder, bird attracting device, wireless data transmission module, power supply with solar panel, portable weather station. The system records bird voices and sends the recordings to the database. Recorded bird voices can be also provoked by the attracting device. Application of wireless data transmission module and power supply with solar panel allows long term operation of digital sound recorder in a hard accessible terrain. Recorded bird voices are analysed by the computer program and labelled with the automatically recognized bird species. Recognition accuracy of the program can be optionally enhanced by an expert system. Besides of labelled sound recordings, database can store also many other information like: photos and films accompanying recorded bird voices/ sounds, information about localization of observation/ recordings (GPS position, description of a place of an observation), information about bird features and behaviour, meteorological information, etc. Database on the base of geographical/ geological digital maps can generate actual maps of bird population (presence, number of individuals of each species). Moreover data-base can trigger alerts in case of rapidly decreasing bird population. It is also possible to obtain new knowledge about bird species with data mining methods.
The paper presents collected data on observed bird species (audio recordings, photos and films) as well as results of experiments testing particular components of the automatic acoustical avian monitoring system.
Key words: bird voices recognition, bird song recognition, hidden Markov models, dynamic time warping, HFCC, MFCC

---

1. Introduction

Environmental protection becomes nowadays more and more important issue in government policy of many countries. One of the crucial aspects of this policy is protection of avifauna biodiversity [1, 2]. In order to protect bird species especially endangered with extinction the knowledge about actual number and the map of presence as well as the count of particular bird species should be available. In order to gain these knowledge many terrain expeditions, actions of bird counting etc. are undertaken. These actions involve either skilled expert or volunteers and are usually time consuming and expensive [3]. Some remedy for this situation can be an automatic acoustical bird monitoring system [4-6].

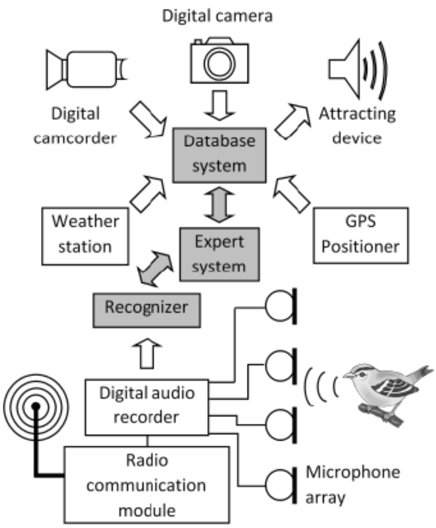

Figure 1. Block diagram of the Automatic Acoustical Avian Monitoring System

The general concept of the system presented in Fig. 1 is as follows. Digital audio recorder located in the vicinity of observation place collects the sounds caused by bird voices via microphone array. Sounds can be also recorded by the digital devices during terrain expedition. Recorded sounds are afterwards analysed and recognized (labelled with bird species) by the program automatically recognizing bird voices. Information about recognized (labelled) birds voices is inserted into the data base.

The most crucial component of the system is the program automatically recognizing bird voices. Among different bird voices recognition methods the most reasonable seem to be dynamic time warping (DTW) [7] or hidden Markov models (HMM) [8]. These methods allow not only bird species recognition but also the type of bird communication signal identification.

The paper has following structure. Chapter 2 presents description of the system. Chapter 3 contains simple summary of the expeditions and observations. There are also presented results of experiments with automatic recognition of bird voices and with preselection. Those experiments were crucial for establishing the structure of a whole system. Moreover Chapter 3 presents discussion of the obtained results as well as a survey of the state of the art of acoustical bird monitoring systems. Chapter 4 contains conclusion and most important tasks to be done in the nearest future.

2. Materials and Methods

This chapter is divided into 3 parts. Part 2.1 describes expeditions and observations carried out in order to acquire material and subsequently to collect patterns in bird voices recognition program and in order to evaluate all the system. Part 2.2 presents electronic devices used in the system. One of the device is digital audio recorder automatically recording bird voices which are recognized by the computer program described in Part 2.3. Recognition process is supported by database and expert system presented in Part 2.4.

2.1. Expeditions and Observations

Application of bird voices recognition needs robust patterns of bird voices to achieve high recognition accuracy. Moreover appropriately described set of recordings is required for evaluation of the system performance. Therefore an effort has been undertaken to collect large set of recordings. At the beginning of the work there was assumed that recordings should be realized in various localizations, year and day period, weather conditions, levels of noise, kind of vocalizations etc.

2.1.1. Point-transects and observation points

Following the above assumption, in the year 2006-2011 there were designed some routs with observation points, so called point-transects [3]. Recordings, films, photos and observations has been carried out at each observation point. Locations for transects were tested and then modified if necessary. Alterations of transects were necessary because: they were too long, contained too many points, some points turned out to be not attractive for recordings, some areas were difficult to reach (especially during wintertime or early spring). The transects were arranged to cover the most varied ecosystems, and ones that were approachable for recordings. Every transect consisted of minimum 8 up to 10 fixed observation points. Distance between every two observation points was set at between 50 m and 400 m. Six transects were selected for frequent observation:

1. Tarnowiec Village – Radlna Village
2. Krzyskie Ponds
3. Stone Town Reserve
4. Polichty (Ciężkowicko-Rożnowski Landscape Park)
5. Rożnów Lake
6. Styr Reserve

Beside of transects with fixed observation points there have been visited less frequently transects with no fixed observation points. At these transects recordings and observations have been carried out when some interesting bird species have been heard or seen. During such an observation or recording, GPS position has been registered as well. The transects without fixed observation points had following locations:

7. Mydlniki
8. Wola Rzędzińska
9. Sanguszko Park in Tarnow
10. Niepołomice Forest
11. Biebrza National Park – Gugny Village
12. Biebrza National Park – Osowiec Fortress
13. Biebrza National Park – Dolistowo Village
14. Biebrza National Park – Dolistowo Stare Village
15. Biebrza National Park – Surroundings of Goniądz
16. Biebrza National Park – Ławki Swamp
17. Biebrza National Park – Honczarowska Dyke
18. Lahema National Park near Viitna Village – Estonia
19. City Park in Tallin – Estonia

Every transect was chosen to cover a specific ecosystem, and more than one ecosystem if possible. The ecosystems of interest were: fields and agricultural area, meadows, forests, swamps, river valleys, ponds and lakes, sea side, parks, town or village area. Field and agriculture area are covered in transects: 1, 2, 4, 5, 8, 11, 13, 15. Meadows are present in transects 1, 2, 3, 4, 5, 6, 8, 11, 13, 15. Forests are represented in transects: 2, 3, 4, 5, 6, 8, 10, 11, 13, 17, 18. Ponds and lakes are in tran-sects: 2, 4, 5, 7, 8, 12, 18. River valleys are in transects: 1, 3, 14, 15. Swamp are in transects: 11, 12, 14, 15, 16, 17, 18. Parks are present in transects: 9, 19. Sea side is present in transect 19.

Besides of transects there have been chosen places being singular observation points. These points were chosen to make recordings in the specific area, or to record interesting bird species outside the transects. Point counts methodology of observation and recording was planned and carried out according to the point-transects methodology described in [3]. Point counts were chosen for recording in the area of parks as well as in more urban areas such as villages and towns. These points from the year 2006 to 2011 were: Zbylitowska Gora, Radlna village, Żabno town, State

Higher Vocational School (in town), Strzeleckich Park in Tarnow, Tarnow city centre, Cathedral in Tarnow. The observation point in Radlna Village is especially interesting because of external camcorder monitoring area in the neighbourhood of feeder (Fig. 2). It gives good opportunity to observe more shy bird species.

Figure 2. Digital stationary internet camcorder mounted in observation point in Radlna Village near the feeder

In the spring of the year 2009 also extra expedition was undertaken to swampy Pawęzów meadow. During this expedition some interesting recordings were done of the species typical for the wetlands and swampy meadow. Two individual recording points were chosen in years 2009-2010 for recordings in industrial and waste disposal areas: Waste Disposal in Krzyż – where solid waste has been deposed, Settlings Basins "Czajki" where the former industrial liquid waste disposal used to be, now under recultivation.

2.1.2. Observation Methodology

Observations and recordings were carried out going through the transect during the expedition. Frequency of the expedition was the highest in spring and the lowest in winter. The motivation was vocal activity of birds and number of species: the highest in spring and the lowest in winter. Exact number of expedition in each season of year is presented in table 1 in chapter „Results and Discussion".

There were also different number of expeditions dependently on the period of day. The greatest number of expeditions was carried out in the morning, noon and evening when the vocal activity of birds was the highest, and the least number of expeditions was carried out in the night. Table 2 in chapter "Results and Discussion" presents exact number of expeditions dependently on the period of day.

In order to discover some correlations between the period of a day and the bird vocalization, some of the expeditions (in spring) were repeated in the same place three times a day: in the morning, at noon and in the evening.

Every walk through the transect with fixed observation points took from about 2 hours to 4 hours. Walk through the transect with no fixed point or visit in the singular observation point was usually shorter: from about 1 hour to 2 hours.

The audio recordings lasted from about 5 min to 10 min (dependently on bird vocal activity), and were done at each observation point on the transect with fixed observation points. It gave approximately up to 2 hours of audio material per every expedition. The recordings were done digitally, as the WAV files, with sampling frequency 48 kHz or 96 kHz (dependently on recording device) and with 16 bit/sample resolution. Recording devices and other electronic equipment are described in details in chapter 2.2. Electronic Equipment.

Beside of audio recordings expedition participants tried to make films and photos of singing birds, what was helpful in later labelling of audio recordings. Moreover an approximated number of individuals of observed bird species was noted down even when they were not audio recorded, video recorded or photographed. Other optional information might be collected concerning the birds, and their behaviour: sex, age, type of sound, specific behaviour accompanying the vocalization, mimicry.

The GPS position has been estimated once for every singular observation point, and points on the transect with fixed observation point. For transects with no fixed observation points the GPS position has been estimated every time when some recordings of singing bird has been done.

For every expedition the weather data have been also determined and recorded. The local temperature, relative humidity and atmospheric pressure have been stored when possible. Beside of this weather from ICM has been also obtained.

For some transects with fixed observation points the general description of transect (concerning geology, soil, flora and fauna) and the flora description of particular observation points have been done.

2.1.3. Processing of Gathered Material

Processing of the audio recordings starts with marking the most representative for expedition bird voices. Adobe Audition 3.0 program has been used for marking the bird voices. This operation can be performed even by a medium skilled person. Marked bird voices have been published on the project web site, and the team of experts labelled the voices with Latin species name. When the experts did not agree with each other than the bird voices were not labelled. Marked voices were examples for labelling the rest of voices collected during expedition.

Bird voices could be compared to available photos or films when necessary. Properly labelled bird voices has been stored in database to become patterns and evaluation recordings for automatic bird voices recognition program.

2.2. Electronic Equipment

In order to record the bird voices properly and collect another information in the automatic acoustical avian monitoring system good electronic equipment is required. The most important is audio recording device capable to record bird voices with relatively high quality. Digital camera and digital camcorder are also important devices because images and films obtained by them can help to recognize bird species which voice is recorded. Not singing birds can be registered by this equipment as well. Another equipment like bird attracting device, GPS, weather station are of less importance, but can increase efficiency and quality of observations. More detailed description of every electronic device is presented in the subsequent chapters.

2.2.1. Audio Recording Device

The most important electronic device in the presented system is audio recording device. Bird voices have been recorded by one of two recording devices: notebook equipped with EDIROL-25 sound card or solid state recorder MARANTZ. The sounds have been saved in the uncompressed WAV format in order not to lose important information embedded in bird voices. Important parameter of audio recording devices is sampling frequency. In case of MARANTZ it is 48 kHz what allows to record the voices of highest spectral components up to 24 kHz (being the half of sampling frequency according to sampling theorem). This is basically sufficient for recording majority bird voices. However there are specific vocalization in some species of birds which contains spectral components laying in ultrasound range above 24 kHz. Therefore some records have been carried out by notebook equipped with EDIROL-25 sound card capable to record sounds with sampling frequency 96 kHz. In case of both recording devices hipercardioidal microphone TONSIL MC 358 has been used.

In spite of quite good parameters of used recording devices there were some drawbacks of the recording methods used so far during the expeditions. The most important were: sampling frequency limited to 48 kHz (in case of MARANTZ), big dimensions and not convenient way of recordings (in case of laptop and EDIROL-25 sound card), possibility to record only 2 channels. Therefore the new digital audio recorder has been constructed.

Digital Audio Recorder (Fig. 3) has been designed as 4-channel recording audio device. There is possible to save 1, 2, and 4-channel WAV files. 4-channel recording can be carried out by microphone array being specifically spaced set of four microphones. Microphone array has been constructed using Panasonic Electret Capsules WM-61A which have linear frequency response in broad band. Using properly designed algorithm of processing the signals from microphone arrays helps to improve quality of digitized acoustic signal coming from chosen direction. Some algorithms e.g. time difference of arrival (TDOA) [9] can even enhance quality of signal coming from chosen point. The ability to enhance signal from chosen point can be very useful while recording bird voices from feeder or nest. Moreover there are also the methods which can locate sound source using microphone array [10]. This can be helpful to some extent while studying behaviour of particular singing bird.

Bit resolution of the registered signal was 16 bit per channel and maximal sampling frequency was 96 kHz. Recorded signals were saved in the uncompressed WAV format onto the SD/SDHC memory card. The largest available SDHC card with 32 GB memory allows non-stop recording for about 11,5 hours with the highest sound quality (4 channels and 96 kHz sampling frequency) and for about 185 hours (nearly 8 days) with the lowest sound quality (1 channel and 24 kHz sampling frequency). Recordings with the lowest sound quality can be very useful for long-term recording low frequency bird voices e.g. owls, doves, crows and others. It is especially important when the terrain is hard accessible and hostile.

Constructed audio recorder convert analogue sound signals acquired by microphones into digital form using SAR A/D converters. SAR A/D converter parameters such as: linearity of frequency characteristics, and SNR proved to be better in comparison with classical sigma-delta A/D converter [11, 12]. Slightly better performance of the SAR converter over the sigma-delta one has been also observed in bird voices recognition experiments [13].

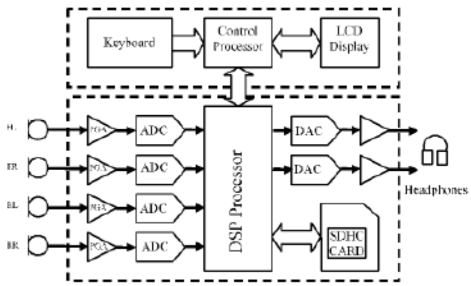

Figure 3. Four-channel Digital Audio Recorder

2.2.2. Bird Attracting Device

Bird attracting device (Fig. 4) can significantly increase efficiency of bird voices recording, especially shy species. However this device was not used in every expedition. Bird attracting device used in expedition was standard portable CD or MP3 player. Yet these device are not ergonomic and the sounds emitted by them has limited bandwidth and rather low volume.

Therefore bird attracting device has been designed and built. This is digital device which allows playback of bird attracting signals stored on the SD/SDHC card. Attracting device has been built using 32-bit microcontroller/DSP with ARM7 core. In order to make the work comfortable, device has been equipped with big LCD display and user-friendly menu allowing quick access to the given bird attracting sound. The device has got built-in amplifier, and loudspeaker for attracting sounds playback. Played signal has 16-bit resolution and 96 kHz sampling rate what makes possible to emit sounds with ultrasound spectral components.

Audio recording device and attracting device is powered by 6 recharged batteries or AC adapter. Batteries make possible approximately 14 hours of continuous work. In order to extend working time of the device solar panel with appropriate electronic charging system can be used.

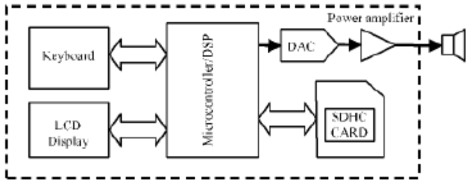

Figure 4. Bird attracting device

2.2.3. Digital Camera and Camcorder

Photos have been made by digital single-lens reflex (DSLR) camera equipped with telephoto lens. Three types of cameras have been used during expeditions: Canon EOS 20D and Canon EOS 5D with Canon EF 100-400mm f/4,5-5,6L IS USM lens and Canon EF 300mm f/2.8 IS USM lens (fig. 5), Nikon D80 with Nikkor AF-S 70-300mm f/4.5-5.6G IF-ED VR lens. Bird pictures support audio recordings and help the team of ornithologists to recognize observed bird species. The birds photography is one of the most difficult task in bird monitoring process. Pictures of birds are taken in terrain at various times of year in variable weather conditions.

In the birds photography DSLR cameras are mostly used nowadays. The DSLR should be fast and it can shot 6 to 10 frames per second for captures dynamic scenes from a birds life. Second important aspect is large memory buffer. The buffer gives possibility to capture series of photos in the internal memory faster than saving them into memory card. The camera should be easy in operating for changing parameters in short time and without removing eyes from camera viewfinder. Inside the viewfinder of DSLR should be displayed the most important parameters like a Shutter Speed, Aperture, Exposure Mode, Focus Mode, ISO Sensitivity. Next most important aspect in dynamic photography is speed and accuracy of an autofocus. Resolution of Picture Sensor should be minimum 10 megapixels. With the Picture sensor connected is second parameter – Crop Factor. This commonly used term denotes the size ratio between a smaller sensor and a full frame 24x36 mm sensor. The crop factor is most often used to determine the "35mm equivalent" field of view for any given lens when used on the camera with sensor smaller than 24x36mm. Range of crop factor in not full-frame DSLRs is 1.3–2x.

Figure 5. Digital camera in use

The most important elements of the equipment in birds photography are lens. Taking pictures of small wild animals such as birds for example, usually require an expensive camera lens. Minimal Focal Length of telephoto lens used in the birds photography is 300 mm. The best telephoto lenses for bird photography have 500 mm and 600 mm. These are expensive and large and heavy camera lenses. The lens should be bright (F/2.8 or F/4.0) and sharp with very fast autofocus. Enlarge focal length of lenses is possible by Teleconverter. The teleconverter is a secondary lens which is mounted between the camera and the photographic lens. Most teleconverters have factor of 1.4 or 2.0.

Films made with digital camcorders provide additional information about appearance and behaviour of the observed bird species. These pieces of information together with digital picture support audio recordings. Films were made with hand-held camcorders type SONY DCR-SR210 with 15-fold optical zoom. Films were saved in MPEG-2 format. In one observation point on the Tarnowiec-Radlna transect there was also possibility to monitor the birds on-line via internet with stationary camcorder of type SONY SNC-RZ 50P with 26- fold optical zoom. Stationary internet camcorder allows observation without presence of human in place of an observation.

2.2.4. Auxiliary Devices

Among auxiliary devices GPS positioner, portable weather station and radio communication module seem to be the most important.

GPS positioner is an additional device used together with digital audio recorder. The positioner allows to estimate GPS position of the recording place. The GPS data can be saved on the hard disc using XML format describing exact recording position. While recording in the fixed observation point GPS positioner is necessary only once during the first walk through the transect.

Birds vocalization strongly depends on weather. The most crucial measured parameters are: temperature, barometric pressure and relative humidity. These parameters were locally measured during some expeditions by the portable weather station. Most often these weather data have been obtained from ICM weather forecast [14]. Good portable weather station should be designed to monitor outdoor temperature, humidity, barometric pressure, wind speed, rainfall and chill factor. The station should also have possibility of connection with personal computer by RS232 interface, or USB interface.

Radio communication module allows wireless control of digital audio recorder. This is an optional module in the system. It may be built into the audio recorder. It provides wireless control of the audio recorder, for example when the recorder is hidden. The person standing in the distance may switch it on and off at any time, with the remote control.

2.3. Automatic Bird Voices Recognition

Probably the most important element in automatic acoustical avian monitoring system is computer program capable of automatically recognizing bird voices. Using this program, bird species or communication signal recognition can be done completely automatically without the need of human presence. There are different recognition methods. In case of bird voices recognition the most reasonable methods seem to be dynamic time warping (DTW) [7, 14] or hidden Markov models (HMM) [8]. These methods allow not only bird species recognition but also the type of bird communication signal identification.

Recognition process can be divided into two phases:

– training phase
– classification phase

During the training phase of the recognition procedure assumed in the system, digital signals recorded for every bird species are detected and labelled with bird species label and signal type label. Moreover for every type of signal the minimal and maximal frequency of the signal should be determined. Some other information of the signal like time and date of recording, GPS position during recording, weather data etc. are also determined. Certain parameters from the signal are calculated too. Process of the labelling, determination of minimal and maximal frequency and giving additional information is carried out manually or semi-automatically. Appropriately designed software can be very helpful in this process.

Afterwards signal is digitally filtered by the filter of the lower and upper frequency being the minimal and maximal frequency of the signal. Afterwards feature extraction procedure on the filtered signal is carried out. Popular feature extraction method are MFCC [16] and HFCC [15]. After feature extraction procedure signal representation in a sequence of feature vector form is obtained. This representation is called a pattern. In case of DTW recognition method, labelled original signal, filtered signal, signal parameters and other data are saved in the database. In case of HMM method addi-tionally for every signal type HMM model is created from the patterns representing signal type.

Classification phase depends on the task to accomplish. The simplest but practically very useful task is species detection. Detection of ural owl (strix uralensis) by voice can be given as an example. Species detection procedure can be carried out as follows:

1. Filtering the whole signal in the frequency range specific for given signal type belonging to detected species.
2. Signal detection which is usually based on energy threshold method.
3. Preselection step which allows to calculate signal parameters like signal length, X coordinate of squared signal mass center, X coordinate of squared FFT mass center, signal-to-noise ratio. These parameters can help decrease significantly the number of patterns or recognized class resulting in lower number of wrong recognitions and shorter recognition time [17].
4. Feature extraction usually based on Mel frequency cepstral coefficient (MFCC) method. The newer method which proved to be better than MFCC in bird voices recognition is Human factor cepstral coefficients (HFCC).
5. Classification which can be DTW or HMM method. For signal rejected at the preselection step DTW in its word spotting form can be used. HMM method in case of bird voices recognition uses whole-word Markov models. For DTW method decision that the recognized signal belongs to detected signal type is taken when the DTW distance between recognized signal and at least one pattern is below predefined threshold. For HMM method decision that the recognized signal belongs to detected signal type is taken when the probability that the signal belongs to the HMM model for signal type is above predefined threshold. When the DTW distance is above the threshold or HMM probability is below the threshold then the signal can be eventually recognized by the word spotting procedure. When this procedure do not recognize the signal then it is marked as unrecognized.
6. Point from 1 to 5 are repeated for every signal type belonging to detected species until any signal type is detected. If none signal type is detected then the species is not recognized.

More complicated recognition task is species recognition by voice. This procedure can be realized like species detection procedure repeated for all the species. In order to limit number of species to be checked out an expert system can be used.

2.4. Database and Expert System

Labelled recognized bird voices with associated data are inserted into the database. Here after some analysis all the data can give useful pieces of information. The most important pieces of information seem to be maps of occurrence of particular bird species, and alerts in case of drop in population of bird species. Moreover expert system using data from database can significantly enhance recognition accuracy. It is also possible to obtain by data mining methods new knowledge embedded in data from database.

2.4.1. Expert System

In the presented acoustical avian monitoring system an expert system provides additional support in preselection process based on metadata and context information. There is used rule-based representation formalism. Knowledge base is built on the basis of ornithologists expertise but can also be discovered with an expert system. The inference mechanism is based on a forward chaining strategy. The strategy searches the knowledge base until it finds a rule where the antecedent is known to be true. After any rule is found the consequent is obtained and therefore the conclusion can be made. Exemplary rules of the expert system have been shown below:

R1: IF observation_date BETWEEN first_day_of_october AND last_day_of_march THEN NOT Golden Oriole (Confidence fac-tor: 99%);

R2: IF height_of_bird < 100 cm OR height_of_bird > 125 cm THEN NOT White_Stork (Confidence factor: 95%);

In acoustical avian monitoring system Prolog logic programming language was chosen.

2.4.2. Database

The acoustical avian monitoring system is implemented with MVC (Model-View-Controller) architectural pattern. Database, represented in MVC architecture in Model tier, describes entities as well as relationships be-tween them. Database model represents two categories of information:

- encyclopedical data;
- experimental data.

Encyclopedical data includes well-known knowledge about bird species, especially taxonomy, appearance and anatomical features, usual habitats and nutrition, typical singing time, approximated migration and breeding dates as well as voice characteristics. Experimental data includes information which are gathered during observations. Since there were proposed precise procedures, defining how data during observations are collected, data-base had to be adopted for storing information coming from different sources and with distinctive structure. Some of the procedures are used for observations with audio recording, some of them are used for presence of species ascertainment and the other ones are used for bird counting. All procedures delivers very useful information which are utilized by recognizer, especially in preselection process but also in data mining. In acoustical avian monitoring system PostgreSQL DBMS (Database Management System) was chosen.

2.4.3. Data Mining Engine

Data mining engine is used to extract hidden, concealed knowledge from entered sound recordings and metadata, context information. Techniques of classification and clusterization in data mining process are used. Since the mining process is ineffective if the sound recordings as well as metadata and context information sets are not a good representation of the sets universum it is necessary to gather as much sound recordings, metadata and context information as possible. Therefore there is proposed alternative path of collecting information. Besides entering alphanumerical (e.g. text, numbers) and multimedia data via recognizer, alphanumerical data can be entered via internet browser and application implemented with Flex and Java technology. Java technology represents second tier (i.e. domain logic) in MVC architectural pattern, whereas Flex technology represents third tier (i.e. user interface view). In data mining, classification with neural networks or naïve Bayes could be used to determine preselection based both on signal parameters as well as metadata and context information. Classification based on metadata and context information are sources of experimental knowledge which, after formalization, can formulate new rules in an expert system.

Results and Discussion

In the period from August 2006 till August 2011, 118 scientific expeditions dedicated to bird species recordings were undertaken. The collected and acoustically analysed material was about 250 hours of recordings. 88 bird species vocalizations were recorded and analysed. This was accompanied with the compulsory and optional information collected.

Table 1. Water contact angle (θ) and roughness (R_a) of top and bottom surface of PTMC-PLA films

Season of the year	Number of expeditions
Spring	94
Summer	45
Autumn	20
Winter	20

Table 2. The number of expedition dependently on the time of the day between August 2006 and August 2011

Time of the day	Number of expeditions
Morning	56
Noon	54
Evening	48
Night	6

Table 3. List of bird species audio recorded (A), photographed (P) and video registered (V) between August 2006 and August 2011

No.	Species	A	P	V	No.	Species	A	P	V
1.	Cygnus olor	+	+	+	61.	Pica pica	+	+	+
2.	Anas platyrhynchos	+	+	+	62.	Corvus monedula	+	+	+
3.	Anas clypeata	-	+	+	63.	Corvus frugilegus	+	+	+
4.	Aythya ferina	-	+	+	64.	Corvus cornix	-	+	+
5.	Aythya fuligula	-	+	+	65.	Corvus corax	+	+	+
6.	Mergus merganser	-	+	+	66.	Cinclus cinclus	-	-	+
7.	Apus apus	+	-	-	67.	Fringilla coelebs	+	+	+
8.	Botaurus Stelaris	+	-	-	68.	Serinus serinus	+	+	+
9.	Egretta alba*	-	+	+	69.	Carduelis chloris	+	+	+
10.	Ardea cinerea	-	+	+	70.	Carduelis carduelis	+	+	+
11.	Ciconia nigra	-	+	+	71.	Carduelis spinus	+	+	+
12.	Ciconia ciconia	-	+	+	72.	Carduelis cannabina	-	+	-
13.	Columba livia	+	+	+	73.	Carpodacus erythrinus	+	+	+
14.	Columba palumbus	+	+	+	74.	Pyrrhula pyrrhula	+	+	+
15.	Streptopelia decaocto	+	+	+	75.	Coccothraetus coccothraetus	-	+	-
16.	Streptopelia turtur	+	+	+	76.	Riparia riparia	+	+	+
17.	Upupa epops	+	-	+	77.	Hirundo rustica	+	+	+
18.	Cuculus canorus	+	-	-	78.	Delichon urbicum	+	+	+
19.	Phasianus colchicus	+	+	+	79.	Motacilla flava	-	+	+
20.	Podiceps cristatus	+	+	+	80.	Motacilla cinerea	-	+	+
21.	Phalacrocorax carbo	-	+	+	81.	Motacilla alba	+	+	+
22.	Circus aeruginosus	-	+	+	82.	Bombycilla garrulus	+	+	+
23.	Accipiter gentilis	-	+	+	83.	Troglodytes troglodytes	+	+	+
24.	Accipiter nisus	-	+	-	84.	Erithacus rubecula	+	+	-
25.	Buteo buteo	+	+	+	85.	Luscinia luscinia	+	+	-
26.	Aquilla chrysaetos	-	+	+	86.	Luscinia megarhynchos	+	-	-
27.	Falco tinnunculus	+	+	+	87.	Luscinia svecica	+	+	+

28.	Falco peregrinus	+	+	+	88.	Phoenicurus ochruros	+	+	+
29.	Aquila pomarina	-	-	+	89.	Saxicola rubetra	-	+	-
30.	Porzana porzana	+	-	-	90.	Saxicola rubicola	+	+	+
31.	Crex crex	+	-	-	91.	Turdus merula	+	+	+
32.	Gallinula chloropus	+	+	+	92.	Turdus pilaris	+	+	+
33.	Fulica atra	+	+	+	93.	Turdus philomelos	+	+	+
34.	Grus grus	+	-	+	94.	Turdus viscivorus	-	+	+
35.	Charadrius dubius	+	+	+	95.	Acrocephalus paludicola	+	-	-
36.	Vanellus vanellus	+	+	+	96.	Acrocephalus schoeno-baenus	+	-	-
37.	Calidris alpina	+	+	+	97.	Acrocephalus palustris	+	-	-
38.	Gallinago gallinago	+	+	+	98.	Acrocephalus scirpaceus	+	-	-
39.	Scolopax rusticola	+	-	-	99.	Acrocephalus arundinaceus	+	+	-
40.	Limosa limosa	+	-	-	100.	Sylvia communis	+	+	+
41.	Numenius arquata	-	+	+	101.	Sylvia borin	+	-	-
42.	Larus ridibundus	+	+	+	102.	Sylvia atricapilla	+	+	+
43.	Sterna hirundo	-	+	+	103.	Phylloscopus collybita	+	+	+
44.	Chlidonias niger	+	+	+	104.	Phylloscopus trochilus	-	+	-
45.	Chlidonias leucop-terus	+	+	+	105.	Phylloscopus sibilatrix	+	+	+
46.	Strix aluco	+	-	-	106.	Regulus regulus	+	+	+
47.	Strix uralensis	+	-	-	107.	Muscicapa striata	-	+	-
48.	Asio otus	+	+	+	108.	Ficedula albicollis	-	+	-
49.	Jynx torqilla	+	+	+	109.	Aegithalos caudatus	+	+	+
50.	Picus viridis	+	+	+	110.	Parus caeruleus	+	+	+
51.	Dryocopus martius	+	+	+	111.	Parus major	+	+	+
52.	Dendrocopos major	+	+	+	112.	Remiz pendulinus	-	-	+
53.	Dendrocopos syriacus	-	+	+	113.	Sitta europaea	+	+	+
54.	Dendrocopos medius	-	+	+	114.	Oriolus oriolus	+	+	-
55.	Emberiza citrinella	+	+	+	115.	Lanius collurio	+	+	+
56.	Emberiza hortulana	+	-	-	116.	Lanius excubitor	+	+	+
57.	Emberiza schoeniclus	+	-	+	117.	Sturnus vulgaris	+	+	+
58.	Emberiza calandra	-	+	-	118.	Passer domesticus	+	+	+
59.	Alauda arvensis	+	+	+	119.	Passer montanus	+	+	+
60.	Garrulus glandarius	+	+	+					

Table 1 presents exact number of expedition in each season of the year in the time between August 2006 and August 2011. Exact number of expedition dependently on the time of the day in the same years presents Table 2.

The vocalizations were recognized and marked using ADOBE AUDITION tools. The compulsory and optional information was stored in database. Such information are collected in the database, to join them with the vocalizations. Simultaneously with the audio recordings the audio-video recordings and pictures were prepared. In the time between August 2006 and August 2011 about 100 hours of audio-video recordings were done, which includes 91 bird species. For the mentioned time 97 bird species were pictured. Table 3 presents list of species audio recorded (A), photographed (P) or video registered (V). The history of expeditions, representative bird audio samples, video recordings as well as pictures are presented on the project web page [18].

3.1. Recognition results

In order to evaluate influence of various factors on recognition accuracy a couple of experiments have been carried out.

3.1.1. Type of A/D Converter

The first recognition experiment concerned the influence of A/D converters. Two types of A/D converters sigma-delta (SD) and SAR have been tested. Comparison of spectrograms for sigma-delta and SAR converters is presented on Fig. 6. As can be noticed the SAR converter indicates lower noise level than the sigma-delta one in the higher frequency range.

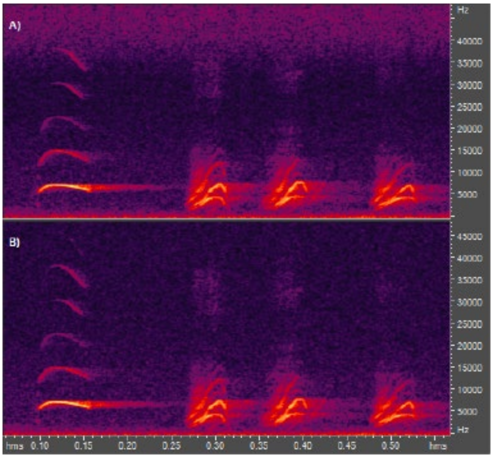

Figure 6. Comparison of spectrograms of bird voices for a) sigma-delta A/D converter, b) SAR A/D converter

Table 4. Recognition accuracies for accomplished experiments. Abbreviations not explained in the text: Filtr. – filtration type, N – none, HP – high pass, Class. – classifier, Ex. No. – experiment number, Avg. accu. – average recognition accuracy, CF – *Corvus frugilegus*, CM – *Corvus monedula*, PMa – *Parus major*, PMo – *Passer montanus, TM – Turdus Merula.*

ADC	Filtr.	Class.	Accuracy [%]					
			Species					Avg. accu.
			CF	CM	PMa	PMo	TM	
SAR	N	HMM	100	95,25	100	15,18	56,11	**77,71**
		DTW	99,87	96,75	100	100	100	**99,32**
	HP	HMM	100	97,25	100	90,63	91,11	**95,80**
		DTW	100	96,75	100	100	100	**99,35**
SD	N	HMM	100	94,25	100	23,61	51,39	**75,06**
		DTW	100	97,25	100	99,38	100	**99,32**
	HP	HMM	100	97,25	100	92,50	89,17	**95,78**
		DTW	100	96,5	100	100	100	**99,30**

Table 5. Error rate for accomplished experiments. Abbreviations not explained in the text: Avg. error – average error rate., the rest of abbreviations like in Table 4

ADC	Filtr.	Class.	Error rate [%]					
			Species					Avg. error
			CF	CM	PMa	PMo	TM	
SAR	N	HMM	0	4,75	0	85,63	43,89	**26,85**
		DTW	0,13	3,25	0	0	0	**0,68**
	HP	HMM	0	2,75	0	4,69	8,89	**3,10**
		DTW	0	3,25	0	0	0	**0,65**
SD	N	HMM	0	5,75	0	77,50	48,61	**26,37**
		DTW	0	0	0	2,75	0	**0,55**
	HP	HMM	0	2,75	0	4,81	10,83	**3,63**
		DTW	0	3,5	0	0	0	**0,7**

As can be seen from Tables 4 and 5 slightly better overall results have been obtained in case of SAR converter over the Sigma-delta (SD) one. Besides of influence of the type of A/D converter on recognition accuracy there were also examined influence of high-pass filtering and classification method (DTW or HMM). Comprehensive description of the research can be found in [13].

3.1.2. Prefiltration

Another type of experiment examined influence of band pass filtration of the bird voice acoustical signal on recognition accuracy [19]. This process was called prefiltration. Results are summarized in table 6. As a classification method HMM was used.

Table 6. Summary of recognition accuracies and error rates for accomplished experiments. Abbreviations: NF – not filtered (reference results), BB6(4) – prefiltration with 6(4)-order filter in broad band, NB4 – prefiltration with 4-order filter in narrow band, GLP – global low-pass prefiltration with 200 Hz cut-off frequency, GBP – global band-pass prefiltration with 200 Hz and 17000 Hz cut-off frequencies

	Recognition accuracy				
	NF	BB6	NB4	GLP	GBP
Total	86,34	91,47	92,12	93,24	93,33
	Error rate				
	NF	BB4	NB4	GLP	GBP
Total	0,21	0,14	0,13	0,12	0,11

Results from Table 6 show that after prefiltration a significant improvement in bird voices recognition accuracy can be achieved. However, high recognition accuracy and low error rate strongly depend on filter order what has been shown on Fig. 7.

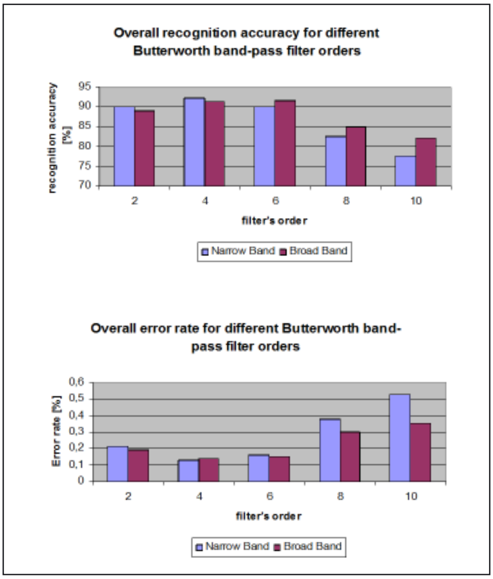

Figure 7. Recognition accuracies and error rates for prefiltration with IIR band pass Butterworth filters of different orders

3.1.3. Preselection

Besides of prefiltration techniques also preselection has been applied in order to enhance bird voices recognition accuracy. Detailed description of preselection methods is beyond scope of this paper and can be found in [17].

Here only the best achieved results of preselection based on the signal length as the preselec-tion parameter are presented in Table 7. Parameters: Pc, Sn, Sp, Ac and Fm result from ROC (Receiver Operating Characteristic) analysis [19]. Definitions of parameters PW, WW, SP, WZW and WPK are given in [17]. Discussion on most important aspects of preselection pro-cess is presented in chapter 3.2.

3.1.4. Choice of feature extraction and microphone type

Five birds species were recorded and analysed in performed experiments:

- House Sparrow – Passer domesticus
- Common Swift – Apus apus (flush)
- Great Tit – Parus major
- Rock-dove – Columba livia
- Chaffinch – Fringilla coelebs

Table 7. Preselection quality coefficients with signal length as the preselection parameter for different bird species in interval method. Pc – precision, Sn – sensitivity, Sp- specificity, Ac – accuracy, F-measure

BS[a]	PM	DM	CF	TM	GG	AS	FC	mean
PW	216	352	242	349	213	259	298	-
WW	687	418	1013	1827	1199	776	1631	-
SP	100	100	100	100	100	100	100	100
WZW	69.7	88	60.5	39.2	41.8	63.2	44.5	53.7
WPK	17.7	45.5	16.8	18.1	17.4	18.1	17.2	21.5
Pc	0.31	0.84	0.24	0.19	0.18	0.33	0.18	0.33
Sn	0.7	0.88	0.61	0.39	0.42	0.63	0.44	0.58
Sp	0.75	0.97	0.68	0.72	0.68	0.79	0.67	0.75
Ac	0.74	0.96	0.67	0.68	0.64	0.77	0.64	0.73
Fm	0.43	0.86	0.34	0.26	0.25	0.44	0.26	0.41

[a]BS – bird species, PM – Parus major, DM - Dendrocopos major, CF - Corvus frugilegus, TM - Turdus merula, GG - Garrulus glandarius, AS - Acrocephalus scirpaceus, FC - Fringilla coelebs

Recordings were done in Tarnow, in the town centre (city park: Parus major, Columba livia, Fringilla coelebes, and high buildings: Apus apus (sounding of flush), Passer domesticus), in August 2006. A 24bit/96kHz sigma delta analog-to-digital converter and two capacitive microphones were used in experiments. The first microphone was of unidirectional hyper-cardioid type, while the second one was a cardioidal microphone. There were 10 examples of bird voice per one species in the training set recorded by two microphones. Because there was limited time of experiments and limited number of particular individuals of bird species, the number of birds voices per one species in testing set is different. The structure of the testing set is shown in Table 8.

Table 8. Structure of the testing set

Bird species	Number of examples
Passer domesticus	17
Apus apus	64
Parus major	25
Columba Livia	63
Fringilla Celebes	95

After the acquisition, MFCC and HFCC feature extraction methods had been used. As a classification method the DTW method has been used. Detailed description of the recognition procedure can be found in [15]. Results of the recognition are presented in table 9. As can be seen the best result has been achieved with HFCC features.

Table 9. Overall accuracy of the bird species recognition by different microphones and signal features

	Hipercardioidal microphone	Cardioidal microphone
MFCC	82.97 %	87.98 %
HFCC	91.91 %	88.65 %

In the next experiment the problem of frequency range has been investigated using only HFCC as signal features (since they were observed to be more efficient than MFCC). In this experiment upper limit of the recognized signal frequency band has been changed from 4 kHz to 48 kHz. Obtained results are shown in Fig. 8. The best recognition result has been achieved with hipercardioidal microphone where the frequency band has been limited to about 24 kHz.

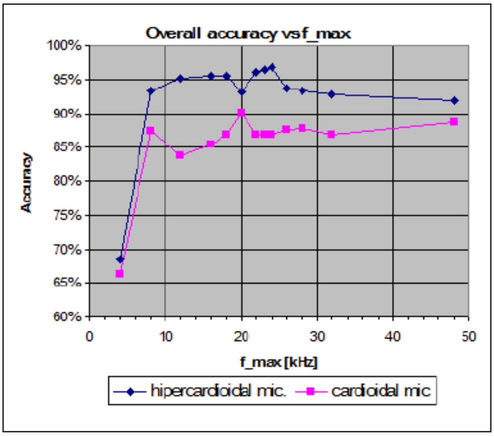

Figure 8. Overall accuracy versus maximal frequency for two types of microphone used in the research

Presented recognition results are relatively good but they were obtained in the closed set experiment without another real world sounds. Therefore, a preliminary open set experiment was carried out only for the hipercardioidal microphone recordings using optimal feature extraction parameters values found in previous experiments. Results are presented in table 10. This experiment reflects to some extent real world application of the investigated methods where system does not take decision about recognized sound if the smallest DTW distance between sound pattern and recognized sound falls above some threshold. It gives some possibility to reject the sounds not being bird voices under consideration.

Table 10. Overall accuracy of the open-set bird spe-cies recognition for recordings made by hipercar-dioidal microphone and optimal parameters values. Th – DTW distance threshold, CR – correctly recognized, R – rejected, FR – Falsely recognized

Bird species	Th	CR	R	FR	Recognition accuracy
Columba livia	5,7	23	40	0	36,51%
Apus apus	2,4	7	57	0	10,94%
Parus major	8,5	14	11	0	56,00%
Passer domesticus	2,2	1	16	0	5,88%
Fringilla coelebs	8,1	48	47	0	50,53%

During recognition in open set a DTW distance thresholds have been experimentally established for every bird species patterns. Threshold values were optimized in order to achieve 0 % false classification. Recognition accuracy in open set experiment was calculated according to the following formula:

$$A = \frac{C}{C+R} \cdot 100\% \quad (1)$$

here:

A – recognition accuracy in [%]

C – number of the correctly recognised sounds

R – number of the rejected sounds.

3.2. Discussion

Majority of expeditions have been carried out in natural environment. Only a few bird species have been recorded, video registered or photographed in somewhat artificial environment like an aviary is.

From Table 1 it is evident that the majority of expeditions took place in spring-summer season. In the autumn-winter season there have been only 40 expeditions. That was because of lower vocal activity and presence of lower number of species in autumn-winter season. During this season there are also usually hard weather conditions like continuous rains, snow, frost, strong wind in Poland. Such a conditions can cause damage of the equipment and are bad for making audio recordings. These conditions cause also lower vocal activity of birds.

Lower vocal activity of birds or its lack in the night is also the reason for lower number of expeditions carried out in the night (Table 2). These expeditions are mainly for recording of species with typical night style of life like owls, nightingales or some crakes.

Looking at Table 3 it can be stated that the representatives of majority of orders observed in Poland are present among the listed species. There are only no representatives of Gaviiformes, Procellariiformes, Phoenicopteriformes, Pterocliformes because birds of these orders are very rarely seen in Poland. The number of bird species with registered voices includes nearly 20 % of the species ever seen in Poland, what seems to be quite a big number. However the number of registered bird voices should contain complete set of bird voices and some set of sounds not being bird voices in order to create good automatic avian monitoring system. From authors' experience every type of sound should be repeated about 100 times in various conditions.

The gathered material is continuously labelled and the patterns for automatic recognizer are being prepared. This is however very difficult and time consuming process, therefore many recordings, films or photos have not been analysed yet.

Recognition experiments have been carried out according to somewhat different methods than that presented in chapter 2. Results of the chronologically first experiment presented in table 9 are relatively good but they were obtained in the closed set experiment without other real world sounds. Moreover the signals have not been filtered. This experiment was carried out only to find the best microphone and feature extraction method. From the results conclusion can be drawn, that the most promising features are HFCC coefficients and hipercardioidal microphone is better than cardioidal one. Hipercardioidal microphone has more narrow directivity pattern. In the future research comparison of hipercardioidal microphone and microphone array is planned.

The second experiment confirmed the positive influence of filtration for recognition accuracy. As one can see in Fig. 8, decreasing the maximal frequency has led to increasing accuracy to the point close to the end of microphone pass-band. After this point recognition accuracy decreases again. Explanation of this phenomena can be following: decreasing maximal frequency limits the energy of noises of spectrum above upper boundary of the microphone frequency response. Because this noises bear nearly no useful information and cause disturbances in the classification process, their limitation causes increasing of the recognition accuracy. On the other hand, decreasing maximal frequency below the upper boundary of the microphone causes cutting off important information laying in the bird voice spectrum thereby decreasing overall recognition accuracy.

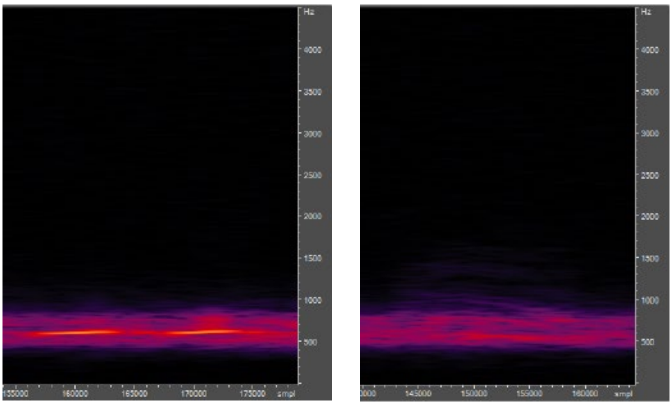

Figure 9. Spectra of signals filtered by 10th order band-pass Butterworth filter. a) Upupa epops voice; b) Corvus frugilegus voice

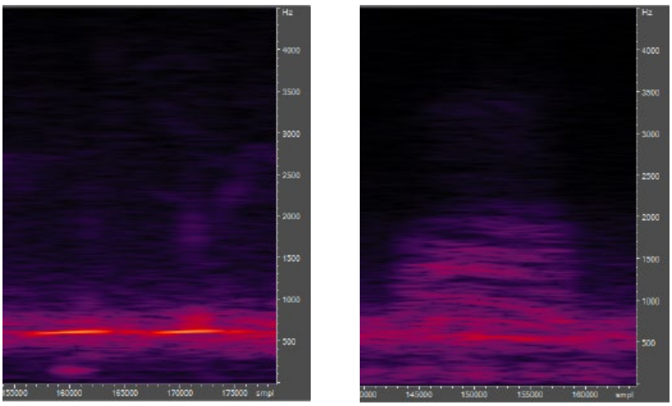

Figure 10. Spectra of signals filtered by 4th order band-pass Butterworth filter a) Upupa epops voice; b) Corvus frugilegus voice

Positive influence of the filtration has been also proved in another research [19]. In these experiments a little bit surprising results have been achieved as for the optimal order of the species specific band pass filters. It turned out that very selective and demanding filtration by 10th order filter gave worse results than less selective and less computationally demanding filtration by 4th order filters. An explanation of this phenomena can be as follows: prefiltration with higher order filters removes many noises and disturbances from outside of the species specific frequency range and this is of course positive effect of prefiltration. However, very selective prefiltration has some serious negative effects which are especially strong while filtering in the narrow bands. Spectra of two voices belonging to different species become similar as shown in Fig. 9. Similarity of the spectra causes similarity of the MFCC coefficients and decreasing recognition accuracies. The same spectra become much less similar when filtering with lower order filters as shown in Fig. 10. Therefore two signals filtered with lower order filters are easier to distinguish and in effect easier to recognize.

Results for open set recognition are presented in Table 10. This results are much worse from results obtained in the closed set experiment (Table 10). However it should be noted that in open set experiment, thresholds were optimised in order to obtain 0 % falsely recognised species. The recognizer took only two types of decision: signal is recognized and signal is rejected. For some bird voices there was big number of the rejected signals. The reason for this was low SNR of the signal. However big number of the rejected signals is not harmful for some applications. If we think about open set recognition procedure as of the procedure indicating presence or absence of the particular species then it is much better to have information about presence of particular bird species with 100 % confidence factor once a day than to have many pieces of information with for instance 80 % confidence factor. The second situation can generate many false alarms which can cause not necessary waste of time and money.

Number of patterns needed for recognition process can be very high. In order to limit this number preselection process is exploited. The first step in preselection process excludes very low probable patterns using information not included in the signal. For example recognized species rather can't be the voice of golden oriole in January and rather can't be the voice of house sparrow at midnight. Similarly the bird voice recorded in Poland in the forest far away from the river rather can't be the voice of kingfisher. Of course the rules excluding some species can be more sophisticated than that presented above.

Beside of the preselection based on the information not included in the signal there can be preselection based on the easy to calculate parameters included in the signal. These are usually parameters related to length of signal, spectral features and signal energy. In table 7 there are presented preselection quality parameters for signal parameter being simply the length of signal. The most important parameter is SP being preselection efficiency. The value of this parameter should be 100 % what means that for every recognized signal the list of preselected patterns includes at least one pattern belonging to the class of recognized signal. The rest of the parameters are also important but not so critical as the SP parameter.

3.3. Acoustical Bird Monitoring – State of the Art

At the end of the article it is worthy to look at existing solutions for acoustical bird moni-toring. Among presently existing bird monitoring systems, the most significant group are systems for bird conservation purposes [22, 23, 25, 32]. However systems constructed for protection of birds against wind turbines are also popular [21, 33]. There are also systems for protection of airports against large bird flocks endangering taking off airplanes [9].

Existing acoustical bird monitoring systems differ as for hardware and software construc-tion. In hardware equipment the most important component is audio recorder. Usually these is one [22, 26] or two [32] microphone device which does not give capability to cope with overlapping signals. More complex solutions include microphone arrays [23] capable to separate voices of the particular individuals and to enhance acoustic signal quality. Maximal range of audio recorder is estimated on 250 m [25].

Processes of data acquisition and bird voices recognition are usually separated [24, 25, 32]. It means that data collection is accomplished by a standalone microprocessor or microcontroller device but acoustic signal analysis and recognition is performed by the stationary computer or notebook of the more powerful computing performance. However, solutions based on autonomous platforms like Raspberry Pi [22] or smartphone [32] have recently appeared.

Another hardware component is power supply. The majority of the systems use only battery supply [22]. Systems designed for long-term work are usually equipped with solar panel recharge device. Small wind turbines are not used because of noise generated by them.

Supplementary devices could be digital camera [5, 21], weather station [5, 33] and radar detector [21, 28, 31, 33] being additional source information supporting acoustical bird monitoring process. Hardware equipment is more and more often connected in sensor networks [28]. An interesting device is a balloon where the electronic equipment can be placed [29].

Software components of the bird monitoring acoustical system are basically: signal detector, feature extractor, classifier and optionally expert system.

Signal detector is usually based on simple energy threshold detection. However more advanced methods are also encountered. Interesting approach is presented in [34] where several detection methods depending on the kind of signal have been proposed. These methods rely on time-frequency spectral analysis. To some extent they allow overlapping signal detection. It seems that the most promising detection strategy is using microphone arrays for signal enhancement [23] together with prefiltration [19] and spectrogram analysis [34] in order to find the region of interest.

Among feature detection methods the MFCC one is the most popular [9, 13]. Good results have been also obtained by using similar HFCC features [15]. There also exist approaches extracting features directly from spectrograms [22]. Some principal component analysis of acoustic features is encountered too [30].

The most often encountered classification method are hidden Markov models HMM [9, 19, 26]. However this method has serious limitation because syntactic of bird songs and calls is not very well recognized. Similar method to HMM is GMM [9, 22] – also often used by researchers. Another methods include ANN [28], SVM [30] and DTW [15].

Summarizing all the consideration in this chapter, it can be stated that general system of bird voices recognition capable to recognize all the voices of birds species from a given geo-graphic region does not exists. Such systems are rather complementary to the traditional point-count methods [27]. The most promising overall strategy of automatic bird voices recognition seems to include: microphone array signal separation, detection of region of interest from spectrogram, MFCC feature extraction and HMM or DTW classification. However in specific cases another strategy can be more efficient. In order to enhance recognition process it is worthy to use non-acoustic information like GPS position, time of the day, season, weather and so on. Nevertheless high efficiency bird voices recognition system still remains an open question.

4. Conclusion

General concept of the automatic acoustical avian monitoring system has been presented in the paper. There were also presented actual state of the work and results of the experiments tending to build such a system.

It can be concluded that it is possible to build automatic acoustical avian monitoring system which can make ornithological research easier. The system can be especially useful in determining presence or absence of the particular bird species in a given area. The system will be able to generate warnings about decrease or increase of bird species population. Moreover the system should be valuable tool for discovering cause of decreasing or increasing population and for discovering various phenomenas during ornithological research.

The most important tasks in the future include:

- integration of the prepared recognition procedures with database system
- evaluation of the recognition procedures integrated with database system on the larger set of recordings including about 90 bird species.
- evaluation of the database system on the gathered material
- terrain tests of the constructed audio recording device.
- continuing terrain expeditions in order to collect complete set of voices of bird species in Poland.

Acknowledgements

This work has been partially cofunded from the Polish Ministry of Science and Higher Education grant no N N519 402934. The project was partially implemented in cooperation with Complex of Foothills Landscape Parks.

References

1. Directive 2009/147/EC of the European Parliament and of the Council of 30 November 2009 on the conservation of wild birds.
2. Council Directive 92/43/EEC on the Conservation of natural habitats and of wild fauna and flora 1992.
3. C. Bibby, N. Burgess, D. Hill, S. Mustoe, Bird Census Techniques. Elsevier 2007.
4. A. Lisowska-Lis, R. Wielgat, D. Król, T. Potempa, Acoustical Bird Monitoring System – Recordigns Methodology and Speceis Chosen for Recordings. 5th International Conference on Interdisciplinarity in Education 6-8 May 2010, Tallinn, Estonia.
5. D. Król, R. Wielgat, T. Potempa, A. Lisowska-Lis, Acoustical Bird Monitoring System – Elec-tronic Equipment. 5th International Conference on Interdisciplinarity in Education 6-8 May 2010, Tallinn, Estonia.
6. T. Potempa, R. Wielgat, A. Lisowska-Lis, D. Król, Acoustical Bird Monitoring System – Data Base Aided Signal Recognition. 5th International Conference on Interdisciplinarity in Education 6-8 May 2010. Tallinn, Estonia.
7. H. Tyagi, R. M. Hegde, H. A. Murthy, A. Prabhakar Automatic identification of bird calls using spectral ensemble average voiceprints. Proceedings of the 13th European Signal Pro-cessing Conference (EUSIPCO '06), Florence, Italy, September 2006.
8. Ch.-H. Chou, Ch.-H. Lee, H.-W Ni, Bird Species Recognition by Comparing the HMMs of the Syllables. ICICIC, pp.143, Second International Conference on Innovative Computing, Information and Control (ICICIC 2007), 2007.
9. C. Kwan, K. Ho, An Automated Acoustic System to Monitor and Classify Birds. Bird Strike Committee Proceedings, 5th Joint Annual Meeting, Toronto, ONT, 2003.
10. J.-M. Valin, F. Michaud, J. Rouat, D. Létourneau, Robust sound source localization using

a microphone array on a mobile robot. in Proceedings International Conference on Intelli-gent Robots and Systems, 2003.
11. D. Król, On superiority of Successive Approximation Register over Sigma Delta AD con--verter in standard audio measurements using Maximum Length Sequences. International Conference on Signals and Electronic Systems, ICSES'08, 14-17 September 2008, Kraków, Poland.
12. D. Król, Choice of analog-to-digital converters for audio measurements using MLS algorithm. 15th European Signal Processing Conference, EUSIPCO 2007, 3-7 September 2007, Poznań, Poland.
13. D. Król, R. Wielgat, T. Potempa, P. Świętojański, Analysis of Ultrasonic Components in Voices of Chosen Bird Species, Forum Acusticum 2011, 26 June-1 July 2011, Aalborg, Denmark.
14. http://www.meteo.pl
15. R. Wielgat, T. P. Zieliński, T. Potempa, A. Lisowska-Lis and D. Król, *Signal Processing Algorithms, Architectures, Arrangements, and Applications*, 2007, 129-134.
16. Ch.-H. Lee, Y.-K. Lee, R.-Z Huang. *Journal of Information Technology and Applications*, 2006, **1**, 17-23.
17. T. Potempa, R. Wielgat, D. Król, P. Kozioł, A. Lisowska-Lis, *STUDIA INFORMATICA. Zeszyty Naukowe Politechniki Śląskiej, seria INFORMATYKA*, 2010, **31 (2A)**, 375-391.
18. http://www.birdsmond.pwsztar.edu.pl
19. R. Wielgat, T. Potempa, P. Świętojański and D. Król, "On using prefiltration in HMM-based bird species recognition," 2012 International Conference on Signals and Electronic Systems (ICSES), Wroclaw, 2012, 1-5.
20. T. Fawcett, *Pattern Recognition Letters*, 2006, **27**, 861-874.
21. G. Mirzaei, M. M. Jamali, J. Ross, P. V. Gorsevski and V. P. Bingman, *IEEE Sensors Journal*, 2015, **15**, 6625-6632.
22. C. wa Maina, D. Muchiri, P. Njoroge, *Biodiversity Data Journal,* 2016, **4**, e9906.
23. D. T. Blumstein, D. J. Mennill, P. Clemins, L. Girod, K. Yao, G. Patricelli, J. L. Deppe, A. H. Krakauer, C. Clark, K. A. Cortopassi, S. F. Hanser, B. McCowan, A. M. Ali, A. N. G. Kirschel, *Journal of Applied Ecology*, 2011, **48**, 758-767.
24. C. E. Sanders and D. J. Mennill, *Ornithological Applications*, 2016, **116**, 371-383.
25. M. A. Pryde, T. C. Greene, *New Zealand Journal of Ecology*, 2016, **40**, 100-107.
26. T. A. Mitchell et al. "Real-Time Bioacoustics Monitoring and Automated Species Iden-tification." Ed. Xiaolei Huang. PeerJ 1 (2013): e103. PMC. Web. 10 Apr. 2017.
27. E. C. Leach, C. J. Burwell, L. A. Ashton, D. N. Jones, R. L. Kitching, (2016) Comparison of point counts and automated acoustic monitoring: detecting birds in a rainforest biodiversity survey. Emu 116, 305-309.
28. A. Boulmaiz, D. Messadeg, N. Doghmane, et al., *International Journal of Speech Technology*, 2016, **19**, 631-645.
29. S. C. Prevost, Estimating Avian Populations with Passive Acoustic Technology and Song Behavior. Master's Thesis, University of Tennessee, 2016.
30. J. Salamon, J. P. Bello, A. Farnsworth, M. Robbins, S. Keen, H. Klinck, et al., *PLoS ONE*, 2016, **11(11)**: e0166866.
31. K.A. Williams, E.M. Adams, J. Fiely, D. Yates, P.B. Chilson, C. Kuster, D.C. Evers. 2013. Migratory Bird and Bat Monitoring in the Thousand Islands Region of New York State: Final Report, March 2013. Report to the U.S. Fish and Wildlife Service Columbus, Ohio Field Office. Report BRI 2013-11, Biodiversity Research Institute, Gorham, Maine.
32. https://www.wildlifeacoustics.com
33. https://www.fws.gov/radar/acoustic
34. S. Duan, Automated species recognition in environmental recordings, PhD thesis, Queensland University of Technology, 2014.