APPROACH TO DESIGN OF DISTRIBUTED MULTI-AGENT SYSTEM FOR PROCESSING SOUND INFORMATION OF THE ENVIRONMENT

Belarusian state University of Informatics and Radioelectronics, Minsk, Republic of Belarus

The processes of measurement, recording and analysis of different sound levels are considered. The amplitude and effect of sound waves vary considerably in continuous space-time measurements. Modeling different types of sounds and their spatiotemporal effects becomes important for assessing the sound situation both in working spaces and in recreation areas. Developing a model that reflects the characteristics of sounds, their sources, and the rules that govern their distribution in different environments would help track sound variations and predict their future changes for spatiotemporal states. Similar works abroad are given, but they are of a private nature. There are many features that you can use to describe audio signals. We consider a wide range of objects to evaluate the effect of each object and select the appropriate set of objects to distinguish between classes. Two estimations of a sound situation are given: on the basis of short-term energy and average speed of change. Three different classification methods are investigated: KNearest neighbors, Gaussian mixture model and Support vector machine.

Multi-agent system (MAS) characteristics are given, the classification, trends in the use of multi-agent intelligent technologies for information processing are presented. Authors propose the use of MAS for sound information (MASSI) monitoring. MFSSI structure includes many agents of sound transformation, analysis of information received from them and decision-making. MASSI can handle noise levels in the urban space and to help in the study of noise pollution in many areas.

Keywords: sound level, sound information, multi-agent system, agents of sound transformation, processing, decision-making.

Introduction

Measuring, registering and analyzing the various levels of sounds and their effect on the surrounding areas is a very complex process. In fact, the amplitude and effect of sound waves vary considerably across the continuous spatio-temporal dimensions. For instance, the noise produced by a taking-off airplane is perceived by its neighborhood with varying amplitudes over time: it starts loud, then decreases gradually while it is flying away. On the other hand, industrial machines typically expose the workers to continuous or repetitive cycles of noise. Hence, modeling the various types of sounds and their spatiotemporal effects becomes crucial. Using a model that embeds the representation of sounds properties on sources, and the rules that govern their propagation across the various surrounding mediums would help in both, tracking the historical variations and predicting the future changes of sounds properties along the spatio-temporal dimensions. In fact, such a model can represent levels of noise in a large urban space and help in studying noise pollution at various layers: inside a given building, in a specific public park or around the whole city. It shall also help in predicting how spatio-temporal changes may affect the levels of noise pollution at any of these layers, for instance when a new building complex or a compound community take place in the city.

The region of Metz in France made a study on the geographical distribution of noise in the region [1], in accordance with the decision of the French ministry of ecology amended in 2006, and which obliges all the French municipalities and regions to perform exhaustive studies on urban noise effects. The study showed, not only the effect of the detected high levels of noise pollution, but also the need for more elaborated models and tools to further understand the collected data and
represent its time and space variations. Another study made in Dalian Municipality of Northeast China [2] showed the potential of the «Land Use Regression Method» in accurately studying noise effects at three special scales, yet the temporal aspects of the noise variations were not completely taken into account.

**Sound information from the environment**

Consider the task of recognizing environment sounds for the understanding of a scene (or context) surrounding an audio sensor. By auditory scenes, refer to a location with different acoustic characteristics such as a coffee shop, park or quiet hallway. Consider, for example, applications in robotic navigation and obstacle detection, assistive robots, surveillance, and other mobile device based services. Many of these systems are dominantly vision-based. To understand unstructured environments, their robustness or utility will be lost if visual information is compromised or totally absent. Audio data could be easily acquired, in spite of challenging external conditions such as poor lighting or visual obstruction, and is relatively cheap to store and compute than visual signals. To enhance the system’s context awareness, we need to incorporate and adequately utilize such audio information [1, 2].

Research in general audio environment recognition has received some interest in the last few years [3], but the activity is much less as compared to that for speech or music. Other applications include those in the domain of wearable’s and context-aware applications [4]. Unstructured environment characterization is still in its infancy. Most research in environmental sounds has centered mostly on recognition of specific events or sounds [5]. To date, only a few systems have been proposed to model raw environment audio without pre-extracting specific events or sounds [6]. Similarly, our focus is not in analyzing and recognition of discrete sound events, but rather on characterizing the general acoustic environment types as a whole.

**Input and encoding information from the environment**

Despite the reduced emphasis in modern times on listening to sounds other than speech or music, humans retain the ability to identify a wide range of sounds, and can still make subtle discrimination among sounds that have particular importance. One major issue in building a recognition system for multimedia data is the choice of proper signal features that are likely to result in effective discrimination between different auditory environments. Environmental sounds are considered unstructured data, where the differences in the characteristics to each of these contexts are caused by random physical environment or activities from humans or nature. Unlike music or speech, there exist neither predictable repetitions nor harmonic sounds. Because of the nature of unstructured data, it is very difficult to form a generalization to quantify them. In order to obtain insights into these data, we performed an analysis to evaluate the characteristics from a signal processing point of view. There are many features that can be used to describe audio signals. The choice of these features is crucial in building a pattern recognition system. Therefore, we examined a wide range of features in order to evaluate the effect of each feature and to select a suitable feature set to discriminate between the classes.

Acoustic features can be grouped into two categories according to the domain in which they are extracted from: frequency-domain (spectral features) and time-domain (temporal features). The temporal information is obtained by reading the amplitudes of the raw samples. Two common measures are the energy and zero-crossing rates [7].

**Short-time energy**

\[
E_n = \frac{1}{N} \sum_{m} [x(m)w(n-m)]^2
\]

Where: \(x(m)\) is the discrete time audio signal, \(n\) is the time index of the short-time energy, and \(w(m)\) is the window of length \(N\) Energy provides a convenient representation of the amplitude variation over time. Zero-crossings occur when successive samples have different signs. It is a simple measure of the frequency content of a signal.

**Short-time average zero-crossing rate (ZCR):**

\[
Z_n = \frac{1}{2} \sum_{m} [\text{sgn}[x(m) - \text{sgn}[x(m-1)]]]w(n-m)
\]

Where

\[
\text{sgn}[x(n)] = \begin{cases} 
1, & x(n) \geq 0, \\
-1, & x(n) < 0.
\end{cases}
\]

Similarly, \(w(m)\) is the window of length \(N\). Since the energy level varies depending on the
distance from the sound source, we use the range of the short-time energy as a measure and feature, instead of the average.

**Classification of sound information from the environment**

Three different classification methods were investigated: KNearest Neighbors (KNN) [8], Gaussian Mixture Models (GMM) [9], and Support Vector Machine (SVM) [10]. For KNN, we used the Euclidean distance as the distance measure and the 1-nearest neighbor queries to obtain the results. As for GMM, we set the number of mixtures for both training and testing to 5. For the SVM classifiers, we used a 2-degree polynomial as its kernel with regularization parameter $C=10^2$ and the epsilon $\varepsilon = 10^{-7}$, which controls the width of the e-insensitive zone, which used to fit the training data, affecting the number of support vectors used. Since SVM is a two-class classifier, we used the one-against-the-rest algorithm [Burges] for our multi-class classification in all of the experiments. We performed leave-one-out cross-validation on the data. The recognition accuracy using leave-one-out cross-validation was found from calculating:

$$\text{Accuracy} = \frac{\text{# of correctly classified}}{\text{Tota# of dataset}}.$$  

There are many features that can be used to describe audio signals.

**Multi-Agent Systems**

A Multi-Agent System (MAS) is an extension of the agent technology where a group of loosely connected autonomous agents act in an environment to achieve a common goal. This is done either by cooperating or competing, sharing or not sharing knowledge with each other. Multi-agent systems have been widely adopted in many application domains because of the beneficial advantages offered. Some of the benefits available by using MAS technology in large systems [8] are:

1. An increase in the speed and efficiency of the operation due to parallel computation and asynchronous operation.
2. A graceful degradation of the system when one or more of the agent fail. It thereby increases the reliability and robustness of the system.
3. Scalability and flexibility- Agents can be added as and when necessary.

4. Reduced cost- This is because individual agents cost much less than a centralized architecture.

5. Reusability-Agents have a modular structure and they can be easily replaced in other systems or be upgraded more easily than a monolithic system.

Though multi-agent systems have features that are more beneficial than single agent systems, they also present some critical challenges. Some of the challenges are highlighted in the following section.

**Environment:** In a multi-agent system, the action of an agent not only modifies its own environment but also that of its neighbors. This necessitates that each agent must predict the action of the other agents in order to decide the optimal action that would be goal directed. This type of concurrent learning could result in non-stable behavior and can possibly cause chaos. The problem is further complicated, if the environment is dynamic. Then each agent needs to differentiate between the effects caused due to other agent actions and variations in environment itself.

**Perception:** In a distributed multi-agent system, the agents are scattered all over the environment. Each agent has a limited sensing capability because of the limited range and coverage of the sensors connected to it. This limits the view available to each of the agents in the environment. Therefore decisions based on the partial observations made by each of the agents could be sub-optimal and achieving a global solution by this means becomes intractable.

**Abstraction:** In agent system, it is assumed that an agent knows its entire action space and mapping of the state space to action space could be done by experience. In MAS, every agent does not experience all of the states. To create a map, it must be able to learn from the experience of other agents with similar capabilities or decision making powers. In the case of cooperating agents with similar goals, this can be done easily by creating communication between the agents. In case of competing agents it is not possible to share the information as each of the agents tries to increase its own chance of winning. It is therefore essential to quantify how much of the local information and the capabilities of other agent must be known to create an improved modeling of the environment.

**Conflict resolution:** Conflicts stem from the lack of global view available to each of the agents. An action selected by an agent to modify a specific
internal state may be bad for another agent. Under these circumstances, information on the constraints, action preferences and goal priorities of agents must be shared between to improve cooperation. A major problem is knowing when to communicate this information and to which of the agents.

**Inference:** A single agent system inference could be easily drawn by mapping the State Space to the Action Space based on trial and error methods. However in MAS, this is difficult as the environment is being modified by multiple agents that may or may not be interacting with each other. Further, the MAS might consist of heterogeneous agents, that is agents having different goals and capabilities. These may be not cooperating and competing with each other. Identifying a suitable inference mechanism in accordance of the capabilities of each agent is crucial in achieving global optimal solution.

It is not necessary to use multi-agent systems for all applications. Some specific application domains which may require interaction with different people or organizations having conflicting or common goals can be able to utilize the advantages presented by MAS in its design.

**Classification of Multi Agent System**

The classification of MAS is a difficult task as it can be done based on several different attributes such as Architecture [12], Learning [13–15], Communication [12], Coordination [16]. Architecture based on the internal structures of the particular individual agents forming the multi-agent system, it may be classified as two types: homogeneous structure, heterogeneous structure.

**Hierarchical Organization** [17]: is one of the earliest organizational design in multiagent systems. Hierarchical architecture has been applied to a large number of distributed problems. In the hierarchical agent architecture, the agents are arranged in a typical tree like structure. The agents at different levels on the tree structure have different levels of autonomy. The data from the lower levels of hierarchy typically flow upwards to agents with a higher hierarchy. The control signal or supervisory signals flow from higher to a lower hierarchy [18]. The flow of control signals is from a higher to lower priority agents.

In a holonic multi-agent system, an agent is a single entity and may be composed of many sub-agents bound together by commitments. The sub-agents are not bound by hard constraints or by pre-defined rule but through commitments. These refer to the relationships agreed to by all of the participating agents inside the holon. Each holon appoints or selects a Head Agent that can communicate with the environment or with other agents located in the environment. The selection of the head agent is usually based on the resource availability, communication capability and the internal architecture of each agent. In a homogeneous multi-agent system, the selection can be random and a rotation policy could be employed similar to that used with distributed wireless sensor networks. In the heterogeneous architecture, the selection is based on the capability of the agent. The holons formed may group further in accordance to benefits foreseen in forming a coherent structure.

In coalition architecture, a group of agents come together for a short time to increase the utility or performance of the individual agents in a group. The coalition ceases to exist when the performance goal is achieved. The agents forming the coalition may have either a flat or a hierarchical architecture. Even when using a flat architecture, it is possible to have a leading agent to act as a representative of the coalition group. The overlap of agents among coalition groups is allowed as this increases the common knowledge within the coalition group. It helps to build a belief model.

**Team MAS architecture** [19] is similar to coalition architecture in design except that the agents in a team work together to increase the overall performance of the group. Rather than each working as individual agents. The interactions of the agents within a team can be quite arbitrary and the goals or the roles assigned to each of the agents can vary with time based on improvements resulting from the team performance.

Reference [20] deals with a team based multi-agent architecture having a partially observable environment. In other words, teams that cannot communicate with each other has been proposed for the Arthur’s bar problem. Each team decides on whether to attend a bar by means of predictions based on the previous behavioral pattern and the crowd level experienced which is the reward or the utility received associated with the specific period of time. Based on the observations made in [17], it can be concluded that a large team size is not beneficial under all conditions. Consequently some compromise must be made between the
amount of information, number of agents in the team and the learning capabilities of the agents. Large teams offer a better visibility of the environment and larger amount of relevant information.

Proposals

The authors propose is the structure of multi-agent system of sound information (MASSI) monitoring of surrounding environment. MASSI structure has many agents such as sound transformation, analysis of the information received from sound agents, decision-making. It implements the functions to ensure the required class of protection of people (working or living) and allows you to implement an environmental safety system [21]. MASSI can handle noise levels in urban space and help with learning noise pollution of various areas: inside the building, in a public park or around the entire area, increasing the protection of the space to the required level.

Conclusion

1. Modeling different types of sounds and their spatiotemporal effects becomes important for assessing the sound situation both in working spaces and in recreation areas. Two estimations of a sound situation are given: on the basis of short-term energy and average speed of change. Three different methods of classification are investigated: KNearest neighbors (KNN), Gaussian mixture model (GMM) and support vector machine (SVM), the latter being preferred.

2. The multi-agent system (MAC) is considered as an extension of agent technology, in which a group of related autonomous agents acts in an environment for information processing. Its characteristics are given, the classification of MAC, trends in the use of multi-agent intelligent technologies for information processing are presented.

3. It is proposed to use a multi-agent system for monitoring sound information (MASSI) in the environment, which includes a variety of agents for the transformation of sound information, its analysis and decision-making on environmental assessment. MASSI can adjust noise levels in the urban environment and help in the study of its pollution from various noise.

REFERENCES


ЛИТЕРАТУРА


СИСТЕМНЫЙ АНАЛИЗ И ПРИКЛАДНАЯ ИНФОРМАТИКА 3, 2019
ВИШНЯКОВ В. А., САЙЯ Б. Х.

ПОДХОД К РАЗРАБОТКЕ РАСПРЕДЕЛЕННОЙ МУЛЬТИАГЕНТНОЙ СИСТЕМЫ ОБРАБОТКИ ЗВУКОВОЙ ИНФОРМАЦИИ ОКРУЖАЮЩЕЙ СРЕДЫ

Рассмотрены процессы измерения, записи и анализа различных уровней звука. Амплитуда и эффект звуковых волн значительно различаются при непрерывных пространственно-временных измерениях. Моделирование различных типов звуков и их пространственно-временных эффектов становится важным для оценки звуковой обстановки как в рабочих помещениях, так и в зонах отдыха. Разработка модели, отражающей характеристики звуков, их источники и правила, регулирующие их распределение в различных средах, поможет отслеживать изменения звуков и прогнозировать их будущие изменения для пространственно-временных состояний. Подобные работы за рубежом имеются, но они носят частный характер. Существует множество функций, которые можно использовать для описания звуковых сигналов. Мы рассматриваем широкий спектр объектов, чтобы оценить эффект каждого объекта и выбрать соответствующий набор объектов, чтобы различать классы. Даны две оценки звуковой ситуации: на основе кратковременной энергии и средней скорости изменения. Исследованы три различных метода классификации: модель замешивания, модель гауссовой смеси и машина опорных векторов.

Даны характеристики мультиагентной системы (МАС), классификация, представлены тенденции использования мультиагентных интеллектуальных технологий для обработки информации. Авторы предлагают использовать МАС для мониторинга звуковой информации (МАСИ). Структура МАСИ включает в себя множество агентов звукового преобразования, анализ полученной от них информации и принятия решений. МАСИ может обрабатывать уровни шума в городском пространстве и помогать в изучении шумового загрязнения во многих районах.


Вишняков Владимир Анатольевич – д. т. н., профессор профессора БГУИР, каф. ИКТ. Область научных интересов: информационное управление и безопасность, электронный бизнес, интеллектуальные системы управления. Член 2-х докторских Советов по защите диссертаций. Автор более 400 научных работ, в том числе 6 монографий (1 на английском), 4-х учебных пособий с грифом Министерства образования, 8-и томного учебного комплекса «Информационный менеджмент», 137 научных статей.

Shaya Bahaa – master of technical science, PhD-student of ICT department of Belarusian State University of Informatics and Radioelectronics.

Сайя Бахаа – магистр технических наук, аспирант кафедры ИКТ БГУИР.