

UNIVERSITATEA TEHNICĂ DIN CLUJ-NAPOCA
FACULTATEA DE ELECTRONICĂ
TELECOMUNICAȚII ȘI TEHNOLOGIA INFORMAȚIEI
Specializarea: Tehnologii și Sisteme de Telecomunicații (în limba engleză)

Indoor Audio Events Detection Using MFCC and kNN

Proiect de diplomă

Absolvent,
Toma TELEMBICI

Decan,
Prof.dr.ing. Gabriel OLTEAN

Președinte comisie,
Prof.dr.ing. Sorin HINTEA

2018

UNIVERSITATEA TEHNICĂ DIN CLUJ-NAPOCA
FACULTATEA DE ELECTRONICĂ
TELECOMUNICAȚII ȘI TEHNOLOGIA INFORMAȚIEI

Departamentul Bazele Electronicii

Titlul proiectului de diplomă:

Indoor Audio Events Detection Using MFCC and kNN

Descrierea temei:

Subiectul principal al acestui proiect de diplomă a fost ilustrarea conștientizării contextului ambiental pentru un robot de servicii, pe baza analizei acustice.

Pentru validarea rezultatelor s-au folosit 21 de clase audio ce corespund la 3 scenarii diferite (30 fișiere audio/clasă). Semnalele audio au fost înregistrate cu robotul de servicii TIAGo. Scenariile considerate au fost: *bucătărie*, *cameră* și *electrocasnice*.

Pentru a descrie semnalele audio s-au folosit coeficienții MFCC, iar pentru clasificare s-a utilizat kNN. Rezultatele obținute sunt ilustrate pentru un număr diferit de caracteristici, diferite metode de filtrare aplicate înainte de clasificare, metrice, proceduri de vot și metode de ponderare diferite.

Locul de realizare:

Sala 503, str. Observatorului, nr. 2

Data emiterii temei: 02.10.2017

Data predării temei: 05.07.2018

Absolvent,

Toma TELEMBICI

Director departament,

Prof.dr.ing. Sorin HINTEA

Conducător,

Conf.dr.ing. Lăcrimioara GRAMA

Absolvent: Toma TELEMBICI

Conducător: Conf.dr.ing. Lăcrimioara GRAMA

SINTEZA PROIECTULUI DE DIPLOMĂ

Indoor Audio Events Detection Using MFCC and kNN

The main subject of this diploma project is to illustrate the problem of context awareness for a service robot, based on acoustic analysis. The database used was captured by the TIAGo service robot; it consists in 21 classes of audio files that correspond to 3 different scenarios (30 audio files/class). The *kitchen* scenario contains 8 classes (chair, tap water, drop water, shower water, porcelain dish, cutlery, plastic bag rush, cardboard drop), the *room* scenario contains 8 classes (page turn, Velcro, zip open, zip close, door knock, door key, door open, door close) and the *appliances* scenario contains 5 classes (washing machine, microwave open, microwave close, microwave alarm, toaster alarm).

To describe the audio signals, we have used the liftering Mel frequency cepstral coefficients as features, while for classification the k-Nearest Neighbor was used. The results obtained are illustrated for different number of features, various filtering methods prior classification, different metrics, voting procedures and weighting methods, respectively. The best results are obtained using 37 features, City and Simple Value Difference metric, Inverse Distance voting, Accuracy Based weighting method, and $k=3$. The correct classification rate is improved from 98.25% to 99.21%, by resampling data prior classification.

Detecția evenimentelor acustice indoor folosind MFCC și kNN

Subiectul acestui proiect de diplomă este ilustrarea conștientizării contextului ambiental pentru un robot de servicii, pe baza analizei acustice. Baza de date a fost captată cu robotul TIAGo și constă în 21 de clase audio ce corespund la 3 scenarii diferite (30 fișiere audio/clasă). Scenariul *bucătărie* conține 8 clase (scaun, apă chiuvetă, apă picături, apă duș, farfurie porțelan, tacâmuri, fâșâit pungă, obiect carton), scenariul *cameră* conține 8 clase (pagii carte, scai, deschidere fermoar, închidere fermoar, ciocănit ușă, cheie ușă, deschidere ușă, închidere ușă) și scenariul *electrocasnice* conține 5 clase (mașină de spălat, microunde deschidere, microunde închidere, microunde alarmă, toaster alarmă). Pentru a descrie semnalele audio am folosit coeficienții MFCC, iar pentru clasificare am utilizat kNN. Rezultatele obținute sunt ilustrate pentru un număr diferit de caracteristici, diferite metode de filtrare aplicate înainte de clasificare, metrici, proceduri de vot și metode de ponderare diferite. Cele mai bune rezultate s-au obținut pentru 37 de coeficienți MFCC, metrica City and Simple Value Difference, vot de tip Inverse Distance, metoda de ponderare Accuracy Based și $k=3$. Rata corectă de clasificare este îmbunătățită de la 98.25% la 99.21%, prin reeșantionarea datelor înainte de clasificare.

Avizul conducătorului

Conducător,

Absolvent,

ACKNOWLEDGMENT

This work was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CNCS/CCCDI–UEFISCDI, project number PNIII-P2-2.1-BG-2016-0378, 54BG/2016, within PNCDI III.

FINANȚARE

Această lucrare de diplomă a fost realizată în cadrul grantului finanțat de Autoritatea Națională Română pentru Cercetare Științifică și Inovare, CNCS/CCCDI–UEFISCDI, prin proiectul cu numărul PNIII-P2-2.1-BG-2016- 0378, 54BG/2016, PNCDI III.

Table of Contents

1. Rezumat în limba română	2
2. Work planning	8
3. State of the art	9
4. Theoretical Fundamentals	12
4.1. Audio Signals	12
4.2. Audio Signals Features	13
4.2.1. Cepstral Features Extraction	13
4.2.2. Mel Frequency Cepstral Coefficients	13
4.3. Audio Signals Classification	20
4.3.1. <i>K</i> nearest neighbours (KNN)	20
5. Implementation	24
5.1. Database	26
5.2. Features Extraction and Classification	28
5.2.1. Weka	28
5.2.2. Cross Validation	30
5.2.3. Mel Scale	31
5.2.4. RSESLIB	32
5.2.5. Used Filters in Weka	34
6. Experimental Results	38
7. Conclusions	51
8. References	53