

UNIVERSITATEA TEHNICĂ DIN CLUJ-NAPOCA
FACULTATEA DE ELECTRONICĂ
TELECOMUNICAȚII ȘI TEHNOLOGIA INFORMAȚIEI
Specializarea: Circuite și Sisteme Integrate

Tehnici de recunoaștere a sunetului ambiental

Lucrare de disertație

Absolvent,
Liana-Laura TUNS

Decan,
Prof.dr.ing. Gabriel OLTEAN

Președinte comisie,
Prof.dr.ing. Sorin HINTEA

2017

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

Departamentul Bazele Electronicii

Titlul lucrării de disertație:

Tehnici de recunoaștere a sunetului ambiental

Descrierea temei:

În această lucrare s-a urmărit problema conștientizării contextului ambiental pe baza unei analize de semnal acustic. S-a dorit o rată de recunoaștere de peste 90%. Pentru validarea sistemului propus s-au utilizat cinci clase diferite de semnale audio.

Pentru recunoașterea sunetului din mediul înconjurător s-au utilizat caracteristicile oferite de coeficienții Mel cepstrali. S-a studiat influența numărului de caracteristici asupra performanței sistemului. Pentru clasificare, s-au testat un număr de 44 de clasificatori de tip: Multilayer Perceptron, k-Nearest Neighbor, Random Forests, Support Vector Machines etc.

Pentru primii patru clasificatori, cu cea mai bună rată de recunoaștere a semnalului audio, s-a realizat și optimizarea parametrilor.

Locul de realizare:

Sala 503, str. Observatorului, nr. 2

Data emiterii temei: 03.10.2016

Data predării temei: 03.07.2017

Absolvent,

Liana-Laura TUNS

Director departament,

Prof.dr.ing. Sorin HINTEA

Conducător,

Conf.dr.ing. Lăcrimioara GRAMA

Absolvent: Liana-Laura TUNS

Conducător: Conf.dr.ing. Lăcrimioara GRAMA

SINTEZA LUCRĂRII DE DISERTAȚIE

Tehnici de recunoaștere a sunetului ambiental

În această lucrare de disertație s-a urmărit problema conștientizării contextului ambiental pe baza unei analize de semnal acustic. Pentru validarea sistemului propus s-au utilizat cinci clase diferite de semnale audio. S-a dorit o rată de recunoaștere de peste 90%.

Prima etapă a constat în extragerea caracteristicilor semnalelor acustice: coeficienții Mel cepstrali (MFCC). Numărul coeficienților a fost variat de la 10 la 28 (cu sau fără energie). Unul dintre obiective a fost studiul influenței numărului de coeficienți asupra ratei de recunoaștere. A doua etapă a fost clasificarea propriu-zisă a semnalelor audio. Pentru aceasta s-au testat 44 de clasificatori de tip: Multilayer Perceptron (MLP), k-Nearest Neighbor (kNN), Random Forests (RF), Support Vector Machines (SVM) etc. S-a obținut o acuratețe de peste 90% pentru 14 clasificatori când nu s-a luat în calcul energia, respectiv pentru 27 de clasificatori când s-a considerat și energia, indiferent de numărul coeficienților MFCC utilizați. Un alt obiectiv a constat în optimizarea parametrilor clasificatorilor cu cele mai bune performanțe, pentru cadrul considerat. În urma optimizării, cele mai bune rezultate s-au obținut pentru k-NN: 99.91% (10 coeficienți MFCC, cu energie), SVM: 99.85%, RF: 99.69% și MLP: 99.47% (28 de coeficienți MFCC, cu energie).

Environmental Sound Recognition Techniques

In this dissertation thesis, the problem of environmental awareness was investigated based on acoustic analysis. Five different classes of audio signals were used to validate the proposed system. A recognition rate greater than 90% was desired.

The first step was to extract the acoustic signals features: Mel frequency cepstral coefficients (MFCC). The number of coefficients ranged from 10 to 28 (with or without energy). One of the objectives was to study the influence of the number of coefficients on the recognition rate. The second step was the proper classification of audio signals. For this, 44 classifiers such as Multilayer Perceptron (MLP), k-Nearest Neighbor (kNN), Random Forests (RF), Support Vector Machines (SVM), etc. were tested. An accuracy greater than 90% was obtained for 14 classifiers when energy was not considered, respectively for 27 classifiers when energy was considered, regardless of the number of MFCC coefficients used. Another objective was to optimize the parameters of the best performing classifiers for the considered framework. Following optimization, the best results were obtained for kNN: 99.91% (10 MFCC coefficients, with energy), SVM: 99.85%, RF: 99.69%, and MLP: 99.47% (28 MFCC coefficients, without energy).

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 disertație 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.

Cuprins

1	Summary	3
2	Planificarea activității	9
3	Stadiul actual	10
4	Fundamentare teoretică	15
4.1	Semnale audio	15
4.1.1	Categoriile de semnale audio.....	15
4.1.2	Parametrii semnalelor audio	15
4.2	Caracteristicile semnalului audio – coeficienții Mel cepstrali.....	16
4.2.1	Preaccentuarea semnalului	16
4.2.2	Separarea pe cadre a semnalului.....	17
4.2.3	Ferestuirea semnalului.....	17
4.2.4	Transformata Fourier	19
4.2.5	Filtrarea pe scară Mel	19
4.2.6	Transformata cosinus discretă (DCT).....	20
4.2.7	Energia într-un cadru	21
4.2.8	Compensarea zgomotului	21
4.3	Rețele Neuronale - Multi Layer Perceptron	21
4.3.1	Neuronul biologic	22
4.3.2	Neuronul artificial	23
4.3.3	Straturile unei rețele neuronale.....	24
4.3.4	Metode de învățare	24
4.3.5	Arhitecturi de rețele neuronale	25
4.3.6	Perceptronul multistrat (Multi Layer Perceptron)	26
4.3.7	Avantaje și dezavantaje	27
4.4	Algoritmul k-Nearest Neighbor (k-NN).....	27
4.4.1	Algoritm	28
4.4.2	Selectarea parametrilor.....	29
4.4.3	Avantaje și dezavantaje	29
4.5	Random Forests.....	29
4.5.1	Algoritm	30
4.5.2	Caracteristici.....	31
4.5.3	Avantaje și dezavantaje	31
4.6	Support vector machines (SVM).....	32
4.6.1	Clasificarea liniară.....	32
4.6.2	Clasificarea neliniară pentru date neseparabile	33

4.6.3	Clasificarea multi-clasă	34
4.6.4	Avantaje și dezavantaje	34
5	Implementarea soluției adoptate	35
5.1	Achiziția semnalului acustic	35
5.2	Descrierea setului de date	35
5.3	Extragerea caracteristicilor - obținerea coeficienților Mel cepstrali.....	36
5.4	Clasificarea sunetelor în Weka	37
5.4.1	Crearea fișierelor de tip ARFF	38
5.4.2	Algoritmi de clasificarea în Weka	39
5.4.3	Multi Layer Perceptron (MLP).....	44
5.4.4	k-Nearest Neighbor (k-NN).....	46
5.4.5	Random Forests	47
5.4.6	Support Vector Machines	47
6	Rezultate experimentale	49
6.1	Obținerea coeficienților MFCC	49
6.2	Rezultatele clasificării în Weka	51
6.2.1	Rezultatele clasificării cu Multilayer Perceptron (MLP).....	52
6.2.2	Rezultatele clasificării cu k-Nearest Neighbors (k-NN).....	55
6.2.3	Rezultatele clasificării cu Random Forests	60
6.2.4	Rezultatele obținute cu Support Vector Machines	62
6.3	Rezultate comparative	63
7	Concluzii	66
8	Bibilografie.....	68
9	Anexe.....	71