IJESRT

INTERNATIONAL JOURNAL OF ENGINEERING SCIENCES & RESEARCH TECHNOLOGY

COMPARATIVE STUDY AND NEW APPROACH MULTI CLASSIFIERS: APPLICATION TO THE RECOGNITION OF ARABIC NUMERALS

Abdelmajid LAMKADAM, Mohamed KARIM
Department of Mathematics and Computer Science,
Faculty of Science, P.O Box 1796, Dhar El Mehraz,
University of Sidi Mohamed Ben Abdellah, Fez, Morocco.

DOI : 10.5281/zenodo.54645

ABSTRACT

Generally, most of the problems of recognition are due to the classification of the feature vectors, or to the mismatch between learning conditions and test, so to improve the robustness of the existing recognition systems, it is necessary to efficient extraction, to reduce these differences between the reference corpus (learning) and test, and finally achieve a reliable and valid comparison.
In fact, the task at first is devoted to study and present the different classifications algorithms (classifiers) and fly on the theoretical basis of algorithms recognition, which allows us to address the principle of operation of systems recognition in their totality.

And in a second step, we compared the different classification algorithms to identify the advantages and disadvantages of such algorithms, and assess the level and performance of the system already made, to directly filter and keep only the positive side of each method.

In the end, an analysis and comparison of the results were made along this work, which has led us to propose a hybrid classifier, and provides us a significant increase in the rate of speech recognition performing a biometric authentication tool powerful enough as possible.

KEYWORDS:Classifiers, Speech recognition, MFCC, DTW, HMM, SVM, GMM, ANN, Hybridization Approach.

INTRODUCTION

The speech recognition is based in most current systems [1] [2] on the probability or connectionist approaches, these systems are generally composed of three main phases, the first is devoted to the preparation and setting the signal word, the second is for the extraction [11] acoustic vectors, while the latter is to the classification of acoustic fingerprints already extracted. So the treatment and resolution of problems of speech recognition return to address and resolve the problems identified in these three levels.

Our goal here was based on the third phase has as goal to explore classification techniques to obtain sufficient performance, and use when comparing these corpus robust methods, effective and comprehensive, which resist to noise, variability of speakers, and compatible with the properties of sound.

Indeed, this work consisted of a comparative study and selective existing classification algorithms as: DTW [3] [4], HMM [6] [7] GMM [13] [14], SVM [8] [9] ANN [15] [16] hybridizations already created [18] [26], and see their strength, their consideration of characteristics of the speech signal, so their resistance to noisy risk degrading the quality recognition, then attach to the functioning of the strategies planned hybridizations.

In late correct and fill some gaps identified in the classification process, we arrive at a classification techniques of mixing [8], [11] as a hybridization [17] meets the maximum of comparison criteria corpus [2] test and reference, this contribution allows to slightly raise the rate of recognition of Arabic numerals [2] [10].

MATERIALS AND METHODS

A. Comparison of existing classifiers:

These classifiers are different from, each other according to aspects of the speech signal taken into consideration, there are those which take the temporal aspect, the other account the linear aspect, other is based on the probabilistic aspect, the following table (***Table 1***) organize its various aspects in relation to the famous classifiers selected.

Table 1. Classification of different classifiers according to the most used criteria

CRITERION / CLASSIFIER	TOMPO-RAL	DYNA-MIC	LINEAR	STATIS-TIC & DETER-MINIST	CONNEX-IONNISTE	DISCRIMI-NATIVE	DEPEN-DENT OF TEXT
LTW	✓		✓				
DTW		✓	✓				✓
HMM	✓		✓	✓			✓
GMM	✓		✓	✓		✓	
SVM	✓		✓	✓		✓	
ANN					✓	✓	

B. Hybridizations already realized: (some solutions)

As solutions to some problems of recognition, numerous articles [25] recently addressed the robust recognition problems in terms of transverse, made use of several hybridizations and combinations are cited in this case:

- DTW-LDA: Introduction of the time factor in the DTW procedure via the LDA approach.
- HMM-DTW [18]: Introduction of dynamic aspect in HMM.
- HMM-GMM [19]: Combination is very popular for its ease of implementation and easy of interpretation.
- HMM-ANN [23]: ANN were used to estimate the emission probabilities of the HMM, done in the case of texts independent system, and when there is too much noise.
- HMM-SVM [20] [22]
- LVQ-GMM [21]
- GMM-SVM [22]
- SVM-ANN [24]
- HMM-SVM-CRA [25]
- SVM-HMM-ANN [26]
- HMM-ANN-TDNN [18]
- HMM-GMM-ANN [18] [27]
- ...

These hybridizations are not exhaustive; they are others, but more specific to determined cases and less important.

C. Problems and limitations of current systems:

It is difficult to compare the results of these hybridizations since the databases, the extractors used, the assessment criteria vary considerably from one article to another, as the characteristics of the speech signals used, methods filtering and segmentation are also specified.

Advanced and recent research [25][26] in speech recognition, have open and sensitive issues concerning the problems that still remains unsolved currently. So the problem of speech recognition can be formulated in the question: How to classify more the representative units of the speech signal?

D. Working environment:

1) Test basis:

The ***Table 2*** brings together all the parameters; and their associated values, to establish the basic references containing Arabic numbers (0 to 9). The record of corpus is done with precaution, and for one Moroccan speaker.

Table 2. Acoustic properties of the corpus recording.

PARAMETER	VALUE
Format	Mono, (.wav)
Sampling	8 Khz
Codage	16 Bits
Frames number	50
Recording time	5 secondes/digit
Windowing	Hamming

Corpus	10 Arabic numerals

2) Factor of rating:

So to compare the different extraction methods tested on prepared samples, and to evaluate the obtained empirical results, and determine [26] the threshold of recognition; it's based on factor RPF (Recognition Performance Factor), whose trust is the inverse of that given by distance, as of this factor is high there will be a good recognition and vice versa.

$$\text{RPF} = (\text{Number of recognized trials} / \text{Number of total trials}) * 100$$

3) Choice of extractor MFCC:

The choice of MFCC [5][19] as extractor for different classifiers to be used, is not arbitrary, it is justified by the fact that MFCC has considerable results in ASR (Automatic Speech Recognition), and this extractor is characterized by:

- Frequency Aspect,
- Spectral Analysis,
- No linear Aspect,
- Human auditory aspect: proper and better representation of the speech signal,
- Noise: resists better and more robust to noise,
- Signal Capture: better support limitations related to the problem of the capture of the signal.

E. Approach and proposed solution:

1) New hybridization approach:

New techniques [18] [19] appeared to increase the robustness of the recognition systems [26] [27]: their common feature is the use of multiple classifiers that are recombined [18] to take a final decision.

The essential points in these approaches relate to the optimization of the scores obtained from each classifier and the choice of efficient recombination strategies.
The solution proposed as a new approach is the following hybridization:

DTW-HMM-ANN

2) Why this hybridization?

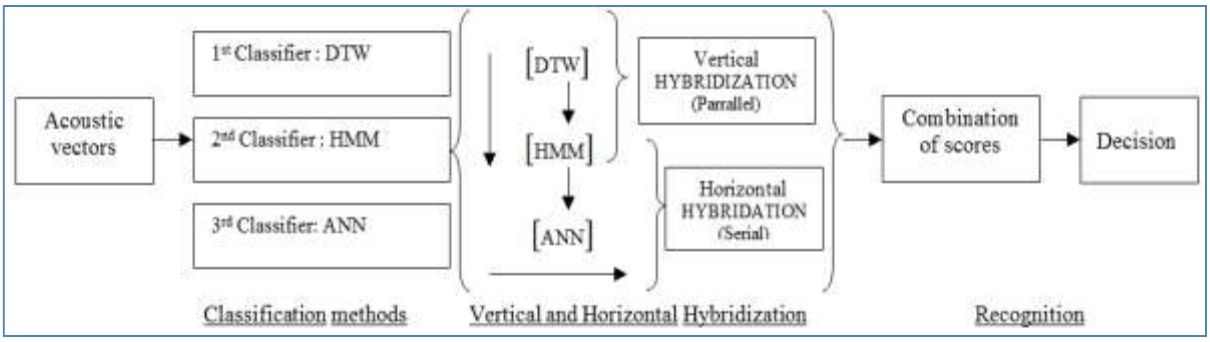
She gathers the advantages of three classifiers:

- DTW: Dynamic aspect,
- HMM: Statistical and Linear aspect,
- ANN: Connexionniste and no linear aspect.

3) Strategy for the chosen hybridization:

During this present manuscript, we became interested in hybridizations multi classifiers, proposing working classifiers in parallel [17] or in serial:

- In parallel: we merged (hybridized) HMM with DTW: DTW is used to calculate the HMM closest distance,
- In serial: we combined HMM with ANN: HMM applied to the output of ANN.



Hybridization model (Multi classifiers Approach).

RESULTS AND DISCUSSION

1) Recognition of Arabic numerals by different methods:

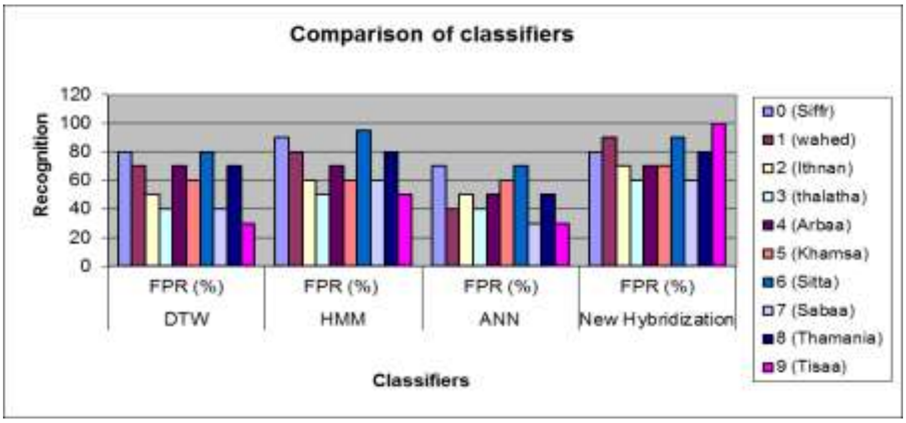
The results of the Arabic recognition digits {0, 1, 2,..., 9}, by the proposed hybridization (new approach), are collected in the following table:

Table 3. Résultats des tests de différentes méthodes de classification.

CLASSIFIER	DTW	HMM	ANN	HYBRIDIZATION
DIGITS	FPR (%)	FPR (%)	FPR (%)	FPR (%)
0 (Siffr)	80	90	70	80
1 (wahed)	70	80	40	90
2 (Ithnan)	50	60	50	70
3 (thalatha)	40	50	40	60
4 (Arbaa)	70	70	50	70
5 (Khamsa)	60	60	60	70
6 (Sitta)	80	95	70	90
7 (Sabaa)	40	60	30	60
8 (Thamania)	70	80	50	80
9 (Tisaa)	30	50	30	100
Average of scores	**59**	**69,5**	**49**	**77**

2) Graphic illustration of test results:

Figure that comes shows the graphic result of ***Table 3***, comparing RPF report of the proposed hybridization, to those old methods, for the Arabic digits recognition.



Comparison reports (RPF) of recognizing numbers {0-9}.

DISCUSSION

- Recognition by HMM is better than ANN or by DTW.
- The rate (average of scores) of recognition by the new hybridization is increased by 7.5 %.

CONCLUSION

In this manuscript, we became interested in the foundations of the theory of classification of acoustic vectors, and we saw what the theoretical foundations of the various used algorithms.

After, we had to compare these various methods of classification according to the applied aspects taken into account, and then flew on different hybridized systems already made over the last decade, by comparing the results of multiple searches and work in several laboratories using well defined contexts. Indeed, these hybridizing work on some methods have presented better and more efficient results.

This has enables us to offer and to chart a way out approach to hybridization of already existing methods, to increase the level of the comparison and to raise the quality of recognition.

With the completed hybridization, results are significant and encouraging, and as future work we can devise other combinations of extractors work and other hybridizations classifiers.

REFERENCES

[1] Laurent BUNIET, “Traitement automatique de la parole en milieu bruité : étude de modèles connexionnistes statiques et dynamiques, ” Soutenue publiquement le lundi 10 février 1997, à l’Université Henri Poincaré - Nancy 1

[2] Lotfi AMIAR, "Un système hybride AG/PMC Pour la reconnaissance de la parole Arabe," Soutenue en 2005 à la Faculté des sciences de l'ingénieur, Université ANNABA, Algérie.
[3] B. KUMAR PADHY & S. KUMAR SAHU, "Analysis of speech recognition techniques," National Institute of Technology Rourkela – 769008, 10 Oct 2009.
[4] C.LEVY, G.LINARES, P.NOCERA, "Comparaison of several acoustic modeling techniques and decoding algorithms for embedded speech recognition systems," Dec 2008 à LIA / CERI Univ. Avignon France.
[5] F.BONNET, B.DEVEZE, M.FOUQUIN, J.JEANY, " Reconnaissance automatique de la parole application : calculatrice vocale, " Nov. 2004 à EPITA 14-16 rue Voltaire 94270 Kremlin-Bicêtre.
[6] Mohamed CHETOUANI, " Codage neuro-prédictif pour l'extraction de caractéristiques de signaux de parole," Soutenue le 14 Décembre 2004, à l'Université Pierre & Marie Curie, France.
[7] C.LEVY, G.LINARES, P.NOCERA, Jean-François BONASTRE, " Reconnaissance de chiffres isolés embarquée dans un téléphone portable," LIA, Avignon, & Stepmind SA, Cannet. Journées d'études de la Parole, JEP, Fez, Maroc, 2004 (ACTN).
[8] Hadi HARB, "Classification du signal sonore en vue d'une indexation par le contenu des documents multimédias ", Lab. LIRIS, Ecole Centrale de Lyon. Décembre 2003.
[9] Lise REGNIER, "Détection de la voix chantée dans un morceau de musique," Master ATIAM 2008, Institut de Recherche et de Coordination Université Paris VI Pierre & Marie Curie- Paris.
[10] H. Satori, M. Harti, and N. Chenfour, "Système de Reconnaissance Automatique de l'arabe basé sur CMU Sphinx,". SCIII 2007, Agadir Maroc. (28-30 Mars 2007).
[11] A. RABAOUI, Z. LACHIRI et N. ETTOUZE, " Classification Robuste des Sons Environnementaux : Amélioration des Paramètres et Adaptation des Modèles, " (TAIMA'07), 21-26 Mai 2007, Hammamet, Tunisie.
[12] Hemant MISRA, "Multi-stream Processing for Noise Robust Speech Recognition," Présentée le 12 Mai 2006 à école polytechnique fédérale de Lausanne.
[13] Ch. CHARBUILLET, " Algorithmes évolutionnistes appliqués à l'extraction de caractéristiques pour la reconnaissance du locuteur ". Soutenue le 5 Décembre 2008 à l'Université Pierre et Marie Curie-Paris 6.
[14] Jamal-Eddine ROUGUI, " Indexation de documents audio : Cas des grands volumes de données, " Soutenue le 16 Juillet 2008, Université de Nantes, France.
[15] Michel WEINFELD, " Réseaux de neurones," H 1990, CNRS, Labo. D'Informatique de l'École polytechnique, Palaiseau.
[16] Drs. TRANI & RAKHA, "MATLAB Toolboxes, Computer Applications in CEE," Spring 2000, Virginia Polytechnic Institute and State University.
[17] Laurent BESACIER, "Un modèle parallèle pour la reconnaissance automatique du locuteur, " Soutenue le 17 Avril 1998 à l'Université d'Avignon et des Pays de Vaucluse.
[18] D.Genoud, F.Bimbot, G.Gravier, G.Chollet, "Combining methods to improve speaker verification decision," CNRS URA820 F-75634 Paris France.
[19] Mohamed FEZARI, Ali Al-DHIOUD, "An Approach For: Improving Voice Command processor Based On Better Features and Classifiers Selection," The 13th International Arab Conference on Information Technology, ACIT 2012.
[20] Bilal TOUNSI, «Inférence d'identité dans le domaine forensique en utilisant un système de reconnaissance automatique du locuteur adapté au dialecte Algérien,". (2008)
[21] Hassan EZAIDI, "Discrimination Parole/Musique et étude de nouveaux paramètres et modèles pour un système d'identification du locuteur dans le contexte de conférences téléphoniques». (2002)
[22] W.M.CAMPBELL, D.E.STURIM, D.A.REYNOLDS, "Support Vector Machines using GMM Supervectors for Speaker Verification," 10 Avr 2006, MIT Lincoln Laboratory 244 Wood Street Lexington, MA 02420.
[23] Ana I. García-Moral, Rubén Solera-Ureña, Carmen Peláez-Moreno, and Fernando Díaz-de-María, " Hybrid Models for Automatic Speech Recognition: a Comparison of Classical ANN and Kernel Based Methods," de AI García-Moral - 2007, EPS-Universidad Carlos III de Madrid.
[24] Yacoub SHERIF, Simske STEVE, Lin XIAOFAN, John BURNS, "Recognition of Emotions in Interactive Voice Response Systems," July 2nd 2003, HP Laboratories Palo Alto, Geneva Switzerland.
[25] Yassine BEN AYED, « Détection de mots clés dans un flux de parole,". Soutenu le 23 Décembre 2003, à LORIA, Ecole nationale supérieur de Télécom.
[26] Christian Raymond et Julien Fayolle, "Reconnaissance robuste d'entités nommées sur de la parole transcrite automatiquement," TALN 2010, Montréal, 19–23 juillet 2010.
[27] ASTRID HAEN, "Robust speech recognition based on multi stream processing," Ecole Polytechnique Fédérale de Lausanne, EPFL 2001.