

Mise en œuvre d'un système d'extraction de règles floues (SERF)

Application à classification phonémique

Mokhtar Bennari¹, Dorra Ben Ayed², Nouredine Ellouze¹

¹Département TIC
Ecole Nationale d'Ingénieur de Tunis
BP-37 Campus universitaire 1002 le belvédère, Tunis
Tunisia
mokhtar.bennari@gmail.com, N.elouze@enit.rnu.tn

²Département Informatique
Institut Supérieur d'Informatique
2 Rue Abou Raihane Bayrouni 2080, Ariana
Tunisia
dorra.mezghani@isi.rnu.tn

ABSTRACT

In this paper, the fuzzy classification method developed by Abe and Lan [1] was implemented. This method extracts fuzzy rules directly from numerical data. This paper describes how generate fuzzy rules and, using the rule inference, how compute the degree of membership. This method is evaluated by using the MFCC coefficient from the TIMIT corpus for phoneme classification.

I. INTRODUCTION

Plusieurs techniques de classification et de décision, se basant sur la logique classique, ont été élaborées [2]. La création des règles floues pouvant servir à une bonne classification constitue une difficulté majeure pour ces techniques. C'est dans cette perspective que cet article s'intéressera au processus d'extraction des règles floues pour la classification phonémique. L'approche proposée est inspirée des travaux des chercheurs M. Shigeo Abe, chercheur au Hitachi Research Laboratory, à Hitachi, et de M. Ming-Shong Lan, membre du staff technique au Science Center, Rockwell International Corporation [1]. Le système ainsi conçu est développé et testé sur l'ensemble

Le présent article est divisé en cinq chapitres. Le premier présente une vue générale du système d'extraction des règles floues. La deuxième partie décrit la phase de préparation des bases utilisées. La troisième partie présente le processus d'extraction des règles floues. Tandis que la quatrième partie développe la phase de classification des échantillons. L'étude expérimentale et la présentation des résultats font l'objet de la cinquième partie.

II. SYSTEME D'EXTRACTION DES REGLES FLOUES (SERF)

Le système d'Extraction des Règles Floues, SERF [3], est constitué, de trois phases :

- la préparation des bases d'apprentissage et de test : L'extraction de connaissance est un but et une nécessité afin d'effectuer l'apprentissage [4] [5]. Dans cette phase, nous avons organisé l'information représentant l'échantillon à expérimenter.
- l'extraction des règles (ERF) : le principe d'extraction est fondé sur le principe de zone ou hyperbox d'inhibition afin de cerner les données de la base d'apprentissage.
- la classification (CSERF) : la classification est basée sur la base des règles ainsi fournie par la phase précédente, ERF.

La figure 1 propose une structure globale du SERF.

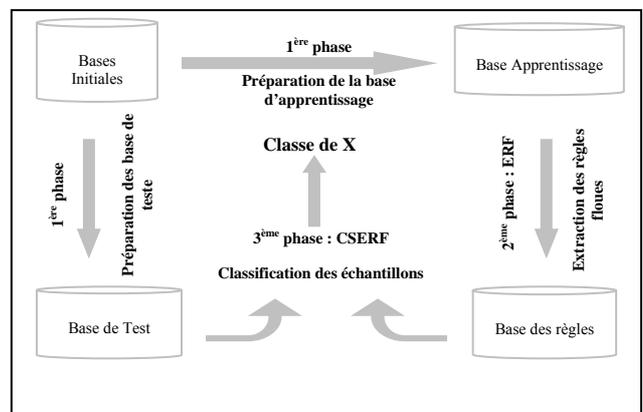


Figure 1 : Structure globale du SERF

III. PREPARATION DES BASES D'APPRENTISSAGE ET DE TEST

La base d'apprentissage est constituée de :

- m colonnes représentant les m propriétés des échantillons.
- Une colonne supplémentaire représentant l'identificateur numérique de la classe à laquelle un échantillon appartient.

Le corpus de son utilisé est celui de TIMIT speech corpus. Nous avons organisé la base d'une manière hiérarchique de sorte qu'une première partie sera réservée comme base d'apprentissage et une deuxième partie comme une base de test. Chacune des bases, d'apprentissage et de test, est divisée en sept groupes phonétiquement homogènes constituant ainsi des sous bases de classes à savoir les *Voyelles*, les *Semi-voyelles*, les *Affriquées*, les *Fricatives*, les *Occlusives*, les *Nasales* et les autres (silences et bruits).

Chacune des sous-bases de ce groupe phonétique est organisée en tables. Chaque table caractérise un ensemble de phonèmes de même type et possède un ensemble de descripteurs caractérisant :

- des informations relatives aux locuteurs ;
- des informations relatives à son identification ;
- des indications sur sa position dans la phrase ;
- le signal audio numérisé ;
- des valeurs caractérisant le phonème.

Chaque phonème est représenté par un ensemble d'attributs constituant ainsi une matrice de paramètres à valeurs réelles. Chaque ligne de cette matrice définit un échantillon de phonème et chaque colonne définit une mesure. De plus, un attribut spécifique est réservé pour caractériser le phonème représentant le *numéro*.

IV. EXTRACTION DES REGLES FLOUES (ERF)

Partons d'une base d'apprentissage à paramètres réels, on applique le processus d'extraction des règles floues fondé sur le principe des hyperbox d'activation et d'inhibition comme le montre l'exemple de la figure 2.

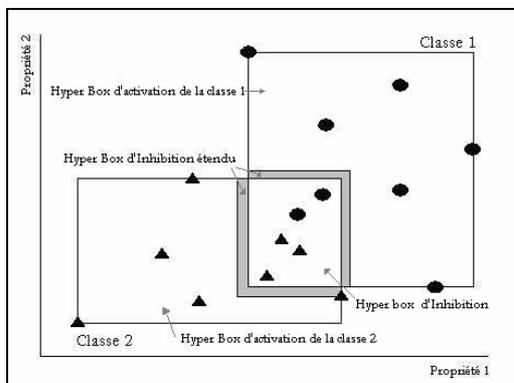


Figure 2 : Hyperbox d'activation et d'inhibition

Supposant que $A_{ii}(1)$ et $A_{jj}(1)$ se chevauchent et soit $J_{ij}(1)$ la région d'inhibition étendue associée à ces deux régions.

On définit la règle floue de niveau 1 associée à la classe C_i comme suit :

Si x est dans $A_{ii}(1)$ Mais x est hors $J_{ij}(1)$ Alors x est de classe C_i .

De la même manière, on définit la règle de classification de niveau 1 à la classe C_j .

Si x est dans $A_{jj}(1)$ Mais x est hors $J_{ij}(1)$ Alors x est de classe C_j .

S'il n'y a qu'un seul hyperbox d'activation ou si les deux hyperbox d'activation ne se chevauchent pas, on définit les règles floues de niveau 2 :

Si x est dans $A_{ij}(2)$ Alors x est de classe C_i .

Si x est dans $A_{ji}(2)$ Alors x est de classe C_j .

Par contre, s'il y a un chevauchement, on construit les deux hyperbox d'inhibition de niveau 2.

Si x est dans $A_{ij}(2)$ Mais x est hors $J_{ij}(2)$ Alors x est de classe C_i .

Si x est dans $A_{ji}(2)$ Mais x est hors $J_{ji}(2)$ Alors x est de classe C_j .

Le processus d'extraction des règles s'arrête :

- s'il n'y a plus de chevauchement entre les hyperbox d'activation ;
- si les deux hyperbox d'activation sont confondus ;
- ou bien si on atteint le seuil de chevauchement fixé a priori.

A la sortie de ce processus, on obtient une base de règles floues. La figure 3 décrit la phase d'extraction de règles floues en suivant le processus ERF.

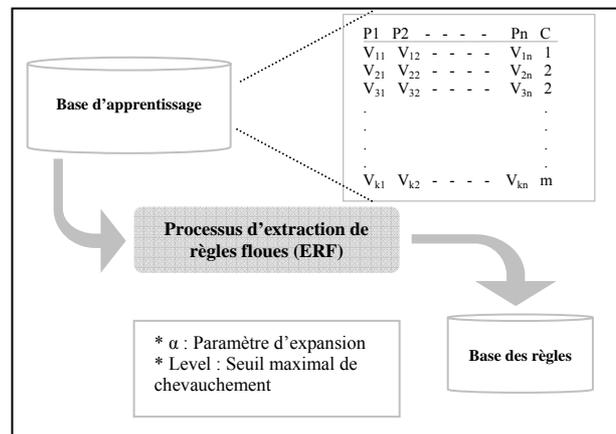


Figure 3 : Vue générale du processus d'extraction des règles floues (ERF)

Chaque règle est une structure composée de :

- **MainClass** : contient l'identifiant de la classe principale
- **AgainstClass** : contient l'identifiant de la classe qui se chevauche avec la classe principale.
- **ActivationBox** : détermine le domaine de l'hyperbox d'activation. Il caractérise l'ensemble A_{ij} défini par les bornes min et max par rapport à une propriété k ; $k \in [1..n]$.

- **InhibitionBox** : détermine le domaine de l'hyperbox d'inhibition. Il caractérise l'ensemble J_{ij} défini par les bornes min et max par rapport à une propriété k .
- **HBox** : détermine le domaine de l'hyperbox d'activation. Il caractérise l'ensemble H_{ij} défini par les bornes min et max par rapport à une propriété k .
- **Level** : contient le niveau de chevauchement.

L'exemple de la figure 4 présente l'organisation d'une règle floue construite par le processus ERF. Dans cet exemple le nombre de propriétés est égal à 12 définissant les 12 coefficients Mel Cepstre. Cette règle décrit un chevauchement au niveau 2 (*Level*) entre la classe 1 (*MainClass*) et la classe 2 (*AgainstClass*). Dans ce cas, tous les hyperbox (*ActivationBox*, *InhibitionBox*, *HBox*) sont remplis. Par contre, dans le cas où il n'y aurait plus de chevauchement entre les classes *MainClass* et *AgainstClass*, seul l'*ActivationBox* serait rempli.

RuleMatrix (l)												
MainClass	1											
AgainstClass	2											
ActivationBox												
Min	0.17	0.33	0.35	0.36	0.00	0.00	0.17	0.30	0.35	0.36	0.00	0.17
Max	0.18	0.98	1.00	0.92	0.79	0.71	0.85	0.98	1.00	0.92	0.79	0.71
InhibitionBox												
Min	0.17	0.33	0.35	0.36	0.01	0.01	0.17	0.17	0.33	0.36	0.01	0.01
Max	0.86	0.98	0.99	0.92	0.78	0.71	0.85	0.99	0.92	0.78	0.71	0.92
HBox												
Min	0.17	0.00	0.00	0.00	0.01	0.01	0.00	0.17	0.00	0.00	0.00	0.01
Max	0.86	0.98	0.99	0.00	0.78	0.71	0.85	0.86	0.98	0.99	0.00	0.78
Level	2											

Figure 4: Structure et exemple d'une règle floue créée par ERF

V. CLASSIFICATION DES ECHANTILLONS (CSERF)

Cette phase consiste à classer un échantillon en déterminant l'identifiant de la classe qu'il pourrait appartenir. La figure 5 donne une vue générale du processus de la phase de classification des échantillons (CSERF).

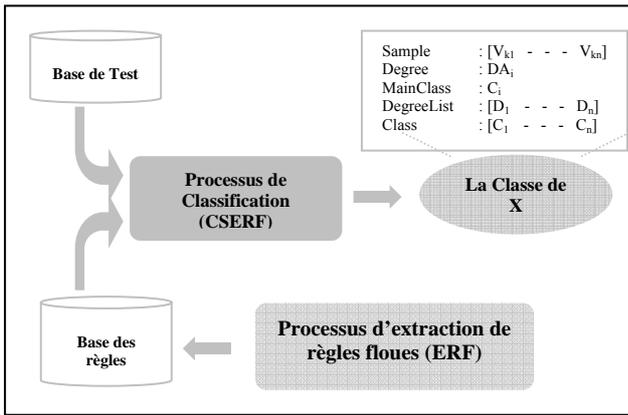


Figure 5: Vue générale du processus de classification (CSERF)

Etant donné un échantillon X inconnu et en se référant à la base des règles fournie par le processus (ERF), le système calcule le degré d'appartenance de l'échantillon d'entrée comme suit :

Tout d'abord, on doit calculer le degré d'appartenance de x à la région A_{ij} de niveau l par rapport à une règle floue, $d_{r_{ij}(l)}(x)$.

Par la suite, on calcule le degré d'appartenance de x à la région A_{ij} , pour tout l'ensemble des règles floues $\{r_{ij}(l)/l=1, \dots\}$. Ce degré est noté $d_{r_{ij}(l)}(x)$, est donné par :

$$d_{r_{ij}}(x) = \max_{l=1, \dots} (d_{r_{ij}(l)}(x))$$

Le degré final est le maximum des degrés d'appartenance par rapport à toutes les règles car l'hyperbox d'activation $A_{ij}(l+1)$, s'il existe, est inclus dans l'hyperbox d'inhibition étendu $J_{ij}(l)$.

Finalement, on détermine le degré d'appartenance de x à la classe i , noté $d_i(x)$, comme suit :

$$d_i(x) = \min_{\substack{j=1, \dots, n \\ i \neq j \\ A_i \cap A_j \neq \emptyset}} (d_{r_{ij}}(x))$$

Si l'hyperbox d'activation de la classe i se chevauche avec ceux des classes j et k , on résout le conflit indépendamment. On commence par le chevauchement entre les classes i et j , on détermine $d_{r_{ij}}(x)$ puis entre les classes i et k et on calcule $d_{r_{ik}}(x)$.

Le degré d'appartenance de x à la classe i est égal au minimum entre les deux degrés.

Et pour terminer, l'échantillon d'entrée x sera attribué à la classe i :

$$d_i(x) = \max_{\substack{j=1, \dots, n \\ i \neq j}} (d_{r_{ij}}(x))$$

CSERF renvoie une structure de données organisée de la manière suivante :

Sample : contient un vecteur de dimension $[1, m]$ où m est le nombre des propriétés.

Degree : est le degré d'appartenance à la classe *MainClass*. Il est le degré maximal par rapport à tous les degrés des autres classes.

MainClass : est l'identifiant de la classe à laquelle l'échantillon *Sample* est classé.

DegreeList : est un vecteur de dimension $[1, n]$, où n est le nombre de classes mises en jeu. Il contient les degrés par rapport à toutes les classes.

Class : est un vecteur de dimension $[1, n]$, où n est le nombre de classes mises en jeu. Il contient les identifiants de toutes les classes organisés dans le même ordre que le DegreeList.

VI. ETUDE EXPERIMENTALE

Nous avons testé le système proposé (SERF) en préparant différentes bases d'apprentissage à paramètres Mel Cepstre (MFCC), à savoir :

- des bases d'apprentissage à locuteur masculin et féminin pour chaque groupe phonémique afin de déterminer l'identifiant d'un phonème en question ;
- des bases d'apprentissage à locuteur masculin et féminin de l'ensemble des phonèmes pour déterminer l'identifiant du groupe de phonème en question.

Le processus ERF a été mis en œuvre :

- en ajustant le paramètre α représentant le rayon de l'hyperbox d'inhibition ;
- en fixant un seuil maximal de chevauchement.

Les figures 6 et 7 présentent les variations du taux de bonne classification par classe phonémique par rapport à α respectivement pour les locuteurs féminins et les locuteurs masculins.

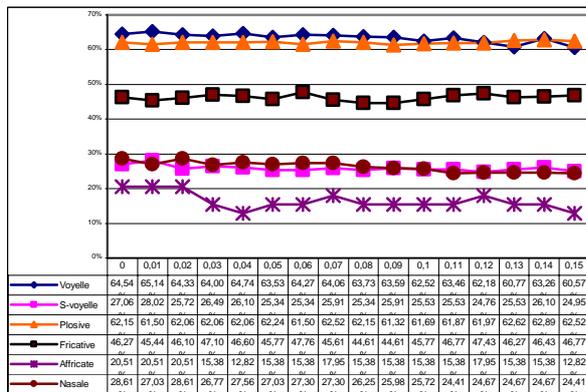


Figure 6: Variation du taux de bonne classification par classe phonémique pour les locuteurs féminins par rapport à α (Base MFCC)

Nous remarquons que les meilleurs taux de bonne classification aperçus sont ceux de la base des phonèmes de type voyelle qui est de l'ordre de 70% et la base des plosives qui est de l'ordre de 60%. Ceci peut s'expliquer par le fait que ces deux bases ont les effectifs les plus élevés par rapport aux autres bases.

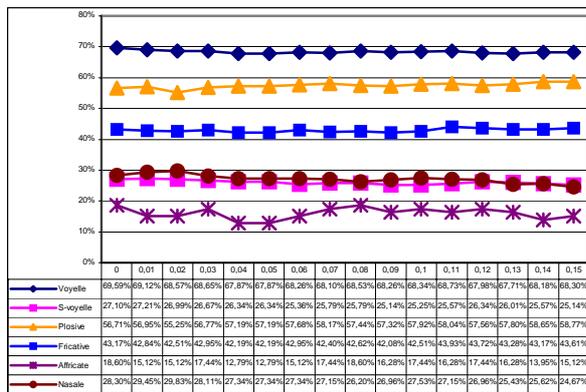


Figure 7 : Variation du taux de bonne classification par classe phonémique pour les locuteurs masculins par rapport à α (Base MFCC)

On remarque aussi que les taux de bonne classification des semi-voyelles, les nasales et les affricates sont faibles vu que les caractéristiques de ces derniers sont très

proches. Le processus ainsi établi ne peut pas attribuer les échantillons aux bonnes classes.

VII. CONCLUSION

Nous nous sommes intéressés dans cet article à la mise en œuvre de la méthode d'extraction de règles floues à partir de données numériques pour aborder la classification phonémique. Ainsi, nous avons présenté les différentes phases constituant le système SERF à savoir la phase préparation des bases d'apprentissage et de test, la d'extraction des règles (ERF) et la phase de classification (CSERF).

Nous avons utilisé le système développé pour la classification des phonèmes en appliquant les coefficients MFCC du corpus TIMIT aussi bien pour des locuteurs masculins que pour des locuteurs féminins.

Les résultats obtenus sont encourageants en terme de bonnes classification et en terme du temps de calcul. Néanmoins, certaines modifications peuvent être apportés à cette méthode afin de d'améliorer les taux de bonne classification (utilisation de la Fuzzy C-means, utilisation des régions ellipsoïdales, introduction de la notion de pondération pour les propriétés ...).

VIII. REFERENCES

- [1] S. Abe, and M.S. Lan, "A method for fuzzy Rules Extraction directly from numerical data and its application to Pattern Classification" IEEE Fuzzy System, Vol. 3, N. 1, pp. 18-28, 1995.
- [2] T.D. Gedeon, H. Kuo, and P.M. Wong, "Rule Extraction Using Fuzzy Clustering for a Sedimentary Rock Data Set", International Journal of Fuzzy Systems, Vol. 4, N. 1, pp. 600-605, March 2002
- [3] M. Bennari, D. Ben Ayed, "Etude et Implémentation d'algorithme d'extraction de règles floues : Application à une base de phonème paramétrées du corpus TIMIT", Mastère en Automatique et traitement du signal, Laboratoire Systèmes et traitement du signal (LSTS), Ecole Nationale d'Ingénieurs de Tunis (ENIT), Juillet 2004.
- [4] W. Fisher, V. Zue, J. Bernstein, D. Pallet, "An acoustic-phonetic database". JASA suppl., A., Vol. 81(S92), 1986.
- [5] H. Prade, C. Testemale, "Fuzzy relational database: Representational issues and reduction using similarity measures", Journal of the American society for information science. 38(2) pp.118-126.
- [6] D. Ben Ayed, N. Ellouze, "Fuzzy C-means Clustering for phoneme classification in continuous speech", first SIAM-EMS Conference, "Applied Mathematics in Our Changing World: Session Speech and Image Recognition", Berlin, September 2-6, 2001.