

République Algérienne Démocratique et Populaire  
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université Mohamed Sadik Benyahia de Jijel  
Faculté des Sciences Exactes et informatique  
Département d'Informatique



*Mémoire de fin d'étude pour l'obtention du  
diplôme Master de Recherche en Informatique.  
Option : Informatique Légale et Multimeèdia*

### Thème

L'utilisation de la logique floue pour  
l'estimation de la qualité objective des  
données audiovisuelles par fusion de  
modalités

**Présenté par :**

AGUIB Mokhtar.

DOUKHANE Sifeddine .

**Encadré par :**

BOUDJERIDA fatima

Promotion : 2019/2020.

## *\* Remerciements \**

Je tiens remercier en premier lieu le Dieu le tout puissant et miséricordieux, qui m'a donné la force et la patience d'accomplir ce Modeste travail.

Je tiens avant tout à exprimer ma reconnaissance à Mme F.BOUDJRIDA enseignante au département de l'informatique à l'Université de Jijel pour avoir accepté de m'encadrer dans cette étude. Je la remercie pour son implication, son soutien et ses encouragements tout au long de ce travail.

Un sincère et honnête merci à nos parents et nos frères et soeurs pour leur soutien indéfectible qui'ils savent nous l'apporter tout au long de nos études et en particulier pendant cette mémoire.

Ce travail est réalisé pour obtenir le diplôme de master , spécialité Informatique légale multimédia au département de l'informatique , université de Jijel.

Nous tenons à remercier vivement les jurys de leur extrême empressement d'évaluer notre travail. Qu'ils ont accepté de venir à évaluer notre travail.

Nous voudrions remercier également tous nos amis de master 2. Nos profonds remerciements vont à nos camarades.

Nos plus vifs remerciements s'adressent aussi à tout le cadre professoral et administratif De l'Université Mohamed Soddik Benyahia de jijel.

Enfin, j'adresse mes plus sincères remerciements tous mes amies, qui m'ont toujours soutenu et encouragé au cours de la réalisation de ce mémoire.

*\* Et Merci \**

*\* Dédicace \**

*\*MOKHTAR \**

Je tiens à dédier ce modeste travail :

A ma très chère **Mère RAZIKA** qui toujours à côté de moi et m'a poussé en avant et à l'esprit de mon père **MOHAMMED \*\*RABI yarhmo\*\*** qui reste toujours dans le coeur .

Mon seul frère **AMMAR** qui me comprend et m'a toujours aidé.

Mes soeurs **NAWALLE** et **HAKIMA** et **MASSOUDA** et **ZAHIRA** et **NADIA**  
et **RAHMA** et leurs enfants.

Mes amis **HICHAM CHOUKI** et **YAACOUB HALLAL** et **BRAHIM TAFER**  
et **WALID DJADAAI** et **MOHAMMED BEN GASOUM** et **YACOUB**  
**MAZGICHE** et **BADIS MAHROUC** et **HOUSAMM CAMEOUNI** et  
**ABDLATIF SLAMNA** .

Que Dieu le tous puissant vos protège et vos réserve un bon avenir.

A mon binôme : **SIFFEDINNE DOUKHANE** Qui je lui considère comme beaucoup plus un Frère que comme mon amie et toute la famille **DOUKHANE** .

*\*E T\**

A la page facebook : **Esay Sciences Algeria** , qui nous a vraiment aidés, et **MS ABDRAHMAN ET OMAR** les deux assistants laboratifs ,c'est avec plaisir que je vous remercie beaucoup. A toute la famille **Doukhane et Aguib** .  
Et tous mes amis dans **université de Jijel**.

## *※SIFFEDINNE ※*

Je tiens à dédier ce modeste travail :

A ma très chère Mère **Zahiya** et à mon cher Père **Abdelrahmane Doukhane**, en témoignage et en gratitude de leurs dévouement, de leurs soutien permanent durant toutes mes années d'études, leurs sacrifices illimités, leurs réconfort moral, eux qui ont consenti tant d'effort pour mon éducation, mon instruction et pour me voir atteindre ce but, pour tout cela et pour ce qui ne peut être dit, mes affectations sans limite.

A ceux qui sont la source de mon inspiration et mon courage, à qui je dois de l'amour et de la reconnaissance.

A mes chers Frères : **Adel, Oussama Zakaria**, Et ma Soeur : **Hadil** Nulle dédicace ne saurait exprimer mon amour et ma profonde affection.

Aux moments inoubliable que nous avons passé ensemble, vous avez toujours été là et où il faut et quand il faut. Les phrases me manquent pour t'exprimer ma grande connaissance et mon admiration profonde

Que Dieu le tous puissant vos protège et vos réserve un bon avenir.

A mon binôme : **Mokhtar Aguib** . Qui je lui considère comme beaucoup plus un Frère que comme mon amie et toute la famille **Aguib**.

A ma très cher amie : **Abdallah Amiour, Hatam Hamlil, Djafer Khen** , Je ne peux trouver les mots justes et sincères pour vous exprimer mon affection et mes pensées, vous êtes pour moi des Frères et des amies sur qui je peux compter. En témoignage de l'amitié qui nous unit et des souvenirs de tous les moments que nous avons passés ensemble, je vous dédie ce travail et je vous souhaite une vie pleine de santé et de bonheur.

Et à ceux qui croiront toujours en moi.

*※Avous tous un grand merci.※*

# Resume

La logique floue propose un environnement mathématique complet, qui permet d'intégrer des données imparfaites dans un processus de décision. Elle offre un cadre pour modéliser la décision humaine tout en s'inspirant de son mode de raisonnement. Les variables du modèle ne peuvent donc pas être représentées par de simples variables booléennes ayant 0 ou 1 comme valeur.

de l'autre côté :

La qualité audiovisuelle (QAV) est essentiellement un problème non linéaire, une solution utilisant la technique de logique floue pourrait lui convenir. Donc nous avons proposé un modèle de logique floue (FL) pour l'évaluation de la qualité d'expérience (QoE). Et il vise à minimiser les effets négatifs des principaux paramètres de QoE (perte de paquets, réduction de bruit,...) en maintenant l'efficacité et la fiabilité des livraisons de données et en améliorant les performances globales du système.

Mots clés : Audiovisuel, Qualité, logique floue, QAV, QV, QA, QoE.

## Abstract

fuzzy logic offers a complete mathematical environment, which allows the integration of imperfect data in a decision process. It provides a framework for modeling human decision while drawing inspiration from their way of reasoning. The variables of the model cannot therefore be represented by simple boolean variables having 0 or 1 as value.

On the other side :

Audiovisual quality (QAV) is essentially a non-linear problem; a solution using the fuzzy logic technique might suit it. Therefore we have proposed a fuzzy logic (FL) model for the evaluation of the quality of experience (QoE). And it aims to minimize the negative effects of the main QoE parameters (packet loss, noise reduction ...) by maintaining the efficiency and reliability of data deliveries and by improving the overall performance of the system.

Keywords : Audiovisual, Quality, fuzzy logic, QAV, QV, QA, QoE.

Table des matières

<b>Table des matières</b>	<b>2</b>
<b>Liste des tableaux</b>	<b>ii</b>
<b>Table des figures</b>	<b>iv</b>
<b>Liste des abréviations</b>	<b>1</b>
<b>Introduction</b>	<b>3</b>
<b>1 Audiovisuel et la qualité perçu</b>	<b>5</b>
<b>Audiovisuel et la qualité perçu</b>	<b>5</b>
1.1 Introduction . . . . .	5
1.2 Définition audiovisuel . . . . .	5
1.3 Les domaines de l'utilisation de l'audiovisuel . . . . .	5
1.3.1 Vidéoconférence . . . . .	5
1.3.2 Face time . . . . .	5
1.3.3 Dans le domaine de l'autorité . . . . .	6
1.4 Les flux de base de L'audiovisuel . . . . .	6
1.4.1 La vidéo . . . . .	6
1.4.2 Codecs . . . . .	6
1.4.3 codecs Vidéo . . . . .	6
1.4.4 Audio . . . . .	7
1.4.5 Codec audio . . . . .	7
1.5 Logiciel pour lire les audiovisuelle . . . . .	8
1.6 La qualité . . . . .	8
1.6.1 qualité de l'expérience (QoE) . . . . .	8
1.7 Evaluation de la qualité . . . . .	9
1.7.1 Evaluation de la qualité subjective . . . . .	9
1.7.2 Evaluation de la qualité objective . . . . .	9
1.8 Conclusion . . . . .	9
<b>2 Evaluation de la qualité audiovisuelle</b>	<b>10</b>

<b>Evaluation de la qualité audiovisuelle</b>	<b>10</b>
2.1 Introduction . . . . .	10
2.2 Les types d'évaluation de la qualité . . . . .	10
2.2.1 Méthodes subjectives . . . . .	11
2.2.2 Métriques objectives . . . . .	15
2.2.3 Performance des modèles d'évaluation objective de la qualité audiovisuelle . . . . .	19
2.3 Conclusion . . . . .	22
<b>3 logique floue</b>	<b>23</b>
<b>logique floue</b>	<b>23</b>
3.1 Introduction . . . . .	23
3.2 Historique de la logique floue . . . . .	24
3.3 Logique floue vs logique classique . . . . .	24
3.4 Théorie de la logique floue . . . . .	25
3.4.1 Les sous-ensembles flous . . . . .	25
3.4.2 Les opérations de base sur les sous-ensembles flous . . . . .	25
3.4.3 Les variables linguistiques . . . . .	27
3.5 Les types de la logique floue . . . . .	28
3.5.1 Type-1 logique floue . . . . .	28
3.6 Type -2- logique floue : . . . . .	31
3.6.1 Fuzzification : . . . . .	31
3.6.2 Mécanisme d'inférence : . . . . .	33
3.6.3 Type-Réduction : . . . . .	33
3.6.4 Défuzzification de type 2 -FL : . . . . .	34
3.7 Conclusion : . . . . .	35
<b>4 Simulation D'une méthode objective pour l'évaluation de la qualité audiovisuelles par la logique floue</b>	<b>36</b>
<b>Simulation D'une méthode objective pour l'évaluation de la qualité audiovisuelles par la logique floue</b>	<b>36</b>
4.1 Introduction . . . . .	36
4.2 Pourquoi la logique floue (FL) . . . . .	36
4.3 Description des modèles proposés . . . . .	36
4.4 La base de données audiovisuelle utilisée . . . . .	38
4.5 Description des paramètres d'entrées et de sortie du système floue pour les trois modèles . . . . .	39
4.6 Analyse et interprétation des résultats . . . . .	45
4.7 conclusion . . . . .	49

<b>Conclusion</b>	<b>50</b>
-------------------	-----------

<b>Bibliographie</b>	<b>51</b>
----------------------	-----------



Liste des tableaux

- 2.1 Echelle d'évaluation de qualité a 9 et 5 niveaux . . . . . 12
- 2.2 Echelle de dégradation à cinq niveaux . . . . . 13
  
- 4.1 Paramètres de compression des médias et dégradation du réseau. [43] . . . . . 38
- 4.2 les résultats de performances des modèles proposées . . . . . 46

## Table des figures

1.1	Diagramme de bloc de codeur H.264. . . . .	6
2.1	Processus d'évaluation objective.[18] . . . . .	10
2.2	Chronogramme de la méthode ACR.[24] . . . . .	12
2.3	Chronogramme de la méthode DCR.[24] . . . . .	13
2.4	Chrono-gramme de la méthode PC.[24] . . . . .	14
2.5	aperçu (a) la méthode de référence complète,(b) la m éthode de référence réduite, (c)la méthode de non-référence.[21] . . . . .	16
2.6	Modèle d'estimation de la qualité multimédia de base . . . . .	18
3.1	Schéma d'ensemble d'un système floue [27] . . . . .	23
3.2	Fonction d'appartenance caractérisant un ensemble classique (a) et d'un en- semble floue (b). . . . .	25
3.3	Fonction d'appartenance . . . . .	25
3.4	Intersection des fonctions d'appartenances . . . . .	26
3.5	Union des fonctions d'appartenances . . . . .	26
3.6	Variable linguistique «par exemple la description de la qualité du service . .	27
3.7	Structure générale d 'un contrôleur flou type-1.[36] . . . . .	28
3.8	Différentes formes de la fonction d'appartenance . . . . .	29
3.9	Exemple d'inférence Max-Min . . . . .	29
3.10	Défuzzification par le centre de gravité . . . . .	30
3.11	fonctions d'appartenance symétriques. . . . .	31
3.12	Structure d'un système de logique floue de type 2. . . . .	31
3.13	forme gaussienne type 2 [41]. . . . .	32
3.14	Footprint of uncertainty et Fonctions d'appartenance supérieures et inférieures.[42]. . . . .	32
3.15	utilise fonctions d'appartenance supérieures et inférieures pour opérateur « ET » .[42] . . . . .	33
3.16	Exemple d'inférence Max-Min.[42] . . . . .	33
3.17	Passage des bornes supérieures aux bornes inférieures et des bornes inférieures aux bornes supérieures..[42] . . . . .	34
4.1	Premier modèle FL pour évaluer la QAV. . . . .	37
4.2	deuxième modèle FL pour évaluer la QAV. . . . .	37

4.3	troisième modèle FL pour évaluer la QAV. . . . .	38
4.4	Évalué la qualité audio par FL . . . . .	39
4.5	Graphe pour intervalle de PLR- audio . . . . .	39
4.6	Fonction d'appartenance de PLR-audio.et Fonction d'appartenance de MOS-audio. . . . .	40
4.7	Fenêtre qui montre un exemple des étapes de la défuzzification pour l'évaluation de QA sur FL. . . . .	40
4.8	Utilisation de la FL pour évaluation de la QV. . . . .	41
4.9	Les meilleurs intervalles pour les inputs : 1=> QP, 2 => FPS, 3 => PLR-vidéo, 4 => NR de vidéo successivement. . . . .	41
4.10	les fonctions d'appartenances A=> FPS, B=>NR, C => QP, D => PLR-vidéo successivement . . . . .	42
4.11	fonction d'appartenance de MOS-vidéo. . . . .	42
4.12	Fenêtre qui montre un exemple des règles d'inférences pour évaluer QV sur FL. . . . .	43
4.13	Fenêtre qui montre un exemple des étapes de la défuzzification pour l'évaluation de QV sur FL. . . . .	43
4.14	fonction d'appartenance de MOS-audio. . . . .	44
4.15	fonction d'appartenance de MOS-vidéo. . . . .	44
4.16	fonction d'appartenance de MOS-audiovisuelle. . . . .	44
4.17	Fenêtre qui montre un exemple des règles d'inférences pour évaluer QAV sur FL. . . . .	45
4.18	graphe définit comment PLR audio et PLR vidéo influence sur la QAV . . . . .	46
4.19	les graphes de MOS-model VS MOS-prédiction pour trois modèles . . . . .	47
4.20	graphe définit comment influencer NR sur la QAV. . . . .	48
4.21	graphe définit comment influencer QP sur la QAV . . . . .	48
4.22	graphe définit comment influencer FPS sur la QAV. . . . .	49

# Liste des abréviations

<b>FL</b>	fuzzy logic
<b>QAV</b>	qualité audiovisuelle
<b>QA</b>	Qualité audio
<b>QV</b>	qualité vidéo
<b>MF</b>	membership function
<b>FIS</b>	système de logique floue
<b>MOS</b>	mean opinion score
<b>INRS</b>	Institut National de la Recherche Scientifique
<b>FPS</b>	frame par seconde
<b>QP</b>	paramètre de quantification
<b>NR</b>	réduction du bruit
<b>PLR</b>	taux de perte de paquets
<b>RMSE</b>	Root Mean Squared Error
<b>RO</b>	outlire ratio
<b>QOE</b>	Quality of Experience
<b>DMOS</b>	Differential Mean Opinion Score
<b>ITU</b>	Union Internationale des Télécommunications
<b>SSCQS</b>	Single-Stimulus Continuous Quality-Scale
<b>ACR</b>	ABSOLUTE CATEGORY RATING
<b>DSCQS</b>	Double Stimulus Continuous Quality-Scale
<b>DCR</b>	Degradation category rating
<b>PC</b>	Pair comparison
<b>FR</b>	Référence complète
<b>RR</b>	Référence réduit
<b>ITS</b>	Institut des sciences des télécommunications
<b>QVGA</b>	Quarter Video Graphics Array

**IPTV** Internet Protocol Television  
**RL** Résolution inférieure  
**HR** Haute résolution  
**DT** Decision Trees  
**GP** Programmation génétique  
**MLP** Multi-Layer perceptron  
**IC** Intervalle de confiance  
**QR** Valeurs aberrantes  
**AV** audiovisuelle  
**SCA** Section cinématographique de l'armée  
**JPEG** Joint Photographic Experts Group  
**MPEG** Moving Picture Experts Group  
**GOP** Group of pictures  
**AVI** Audio Video Interleave  
**MP3** MPEG-1/2 Audio Layer 3  
**AV1** AOMedia Vidéo 1  
**RAW** Real Audio Wrapper  
**WAV** Waveform Audio  
**ASF** Advanced systems format  
**MP4** MPEG Layer-4 Audio  
**AC3** Audio coding3  
**BWF** Broadcast wave format  
**RC** Rate controle  
**CBR** Constant bitrate  
**SBR** Spectral bandwidth replication  
**EQM** Erreur quadratique moyenne  
**PSNR** Peak signal to noise ratio  
**SSCQE** Single stimulus continuous quality evaluation  
**ALAC** Apple lossless audio codec  
**DQ** Dégradation de la qualité  
**SVH** Système visuel humain  
**SSIM** Structural similarity  
**RC** Référence complète  
**DIVIINE** Distorting identification-based image verityand integrity evaluation

# Introduction

Les services multimédias (par exemple, Internet Protocol Télévision, réseaux sociaux, jeux de réalité virtuelle et immersive, conférences multimédias, etc.) ont non seulement gagné en importance à l'ère du sans-fil, mais devraient également croître de façon exponentielle. Ces services et applications peuvent fournir une mauvaise expérience à un utilisateur pour de multiples raisons, par exemple, le réseau (surcharge du réseau) et les dispositifs terminaux appelés éléments de qualité. Ainsi, les communautés de la recherche et de l'industrie élaborent différents programmes pour comprendre comment les utilisateurs perçoivent et subissent les dégradations, et pour leur fournir ensuite une meilleure qualité d'expérience (QoE).

Les approches traditionnelles d'évaluation de la qualité Audiovisuelle QAV peuvent être divisées en deux types : (1) méthodes de qualité vidéo objective et (2) méthodes de qualité vidéo subjective. Toute méthode d'évaluation objective de la qualité audiovisuelle joue un rôle important en offrant des évaluations pour compenser et peut être utilisée comme alternative pour des évaluations subjectives approfondies sous la forme d'un score d'opinion moyen (MOS). Cependant, en raison de la nature intrinsèque de la qualité perceptive étant subjective, il est difficile d'atteindre une métrique de qualité précise en raison de la dépendance à de nombreuses caractéristiques. Tous ces facteurs rendent très difficile la mesure précise et quantitative de la qualité audiovisuelle. En fait, l'évaluation de la qualité des signaux numériques est l'un des problèmes fondamentaux et complexes du traitement multimédia et de ses situations pratiques, telles que l'évaluation et la mise en œuvre de processus, l'optimisation du codage et du décodage, les tests et la surveillance (par exemple en transmission et en fabrication des sites). De plus, comment évaluer la qualité audio et vidéo joue un rôle central dans la conception de la plupart (sinon de la totalité) des services, algorithmes et systèmes multimédias.

La logique floue est une technique populaire qui fournit des mécanismes pour gérer les incertitudes non linéaires qui existent dans les systèmes physiques. En outre, la logique floue peut prendre en charge les descriptions naturelles des entrées et des sorties en termes de langage, ce qui évite la complexité d'identifier les valeurs numériques exactes, à modéliser dans chaque situation. Puisque la qualité audiovisuelle (QAV) est essentiellement un problème non linéaire, une solution utilisant la technique de logique floue pourrait lui

convenir.

Dans ce contexte, nous avons organisé notre mémoire en quatre parties principales, la première partie présente une vue panoramique au l'audiovisuel et sa qualité, ainsi qu'une perspective détailler sur la vidéo et l'audio, ensuite nous avons abordé l'évaluation de la qualité multimédia (audiovisuel), en-suit nous avons détaillé la logique floue et leur types, puis nous avons choisit une méthode pour évaluer la qualité audiovisuel. Plus précisément, ce travail est composé des chapitres suivants :

- **Le premier chapitre** :est consacré à une introduction sur l'audiovisuel tout en passant par ces composants qui sont la vidéo et l'audio, et pour mieux assimiler ce thème nous allons parler de l'évaluation subjective et l'évaluation objective.
- **Le deuxième chapitre** :est dédié à l'évaluation de la qualité audiovisuelle, où on s'est penché sur les différentes méthodes de l'évaluation subjective et les différentes méthodes de l'évaluation objectives, ainsi que les Performance des modèles d'évaluation objective de la qualité audiovisuelle.
- **Le troisième chapitre** : explique la logique floue théoriquement pour se familiariser avec ces astuces et principes de raisonnement. aussi parle un peu de l'histoire de cette technique, et une comparaison avec la logique classique, et ces étapes.
- **Le quatrième chapitre** : est consacré à la description de la méthode choisie pour l'évaluation de la qualité audiovisuelle on prêtera attention aux différentes étapes que nous devons suivre pour obtenir une évaluation, ainsi qu'une comparaison entre les résultats obtenue (évaluation objective) et les résultats qui sont déjà été réalise (évaluation subjective).

Enfin, nous terminons par une conclusion dans laquelle nous dressons un bilan et une synthèse du travail effectué, et nous discutons un ensemble de perspectives qui peuvent être considérées comme des directions de recherche future.

# Audiovisuel et la qualité perçue

## 1.1 Introduction

L'évolution récente des systèmes de communication numérique a conduit à une explosion de services et d'applications multimédias, tels que l'IPTV, le multimédia mobile sur les Smartphones, les réseaux sociaux (par exemple, Face book), la vidéoconférence et les présentations multimédias éducatives, ainsi que l'émergence de plusieurs codec pour audio vidéo et audiovisuelle. Donc ces applications multimédias font désormais partie intégrante (sinon indispensable) de la vie quotidienne et devraient continuer à croître de manière exponentielle.[1]

## 1.2 Définition audiovisuel

Le signal audiovisuel est doublement composé d'un signal vidéo et d'un signal audio, donc audiovisuel sert à désigner tout ce qui est relatif à l'image et/ou au son. Les fichiers audiovisuels s'agit de toutes les formes d'enregistrement du son et/ou des images animées et/ou des images fixes.[2]

## 1.3 Les domaines de l'utilisation de l'audiovisuel

### 1.3.1 Vidéoconférence

En 1968, la vidéoconférence a été introduite pour la première fois et présentée comme une solution commerciale à l'exposition universelle de New York. La technologie introduite s'appelait le Picture phone d'AT&T. Les participants ont pu s'asseoir et communiquer par vidéo avec la personne de l'autre côté pendant 10 minutes à la fois pour faire l'expérience du premier appareil de visiophone conçu pour les masses. Malheureusement, cette machine particulier était ridiculement chère, maladroite et difficile à installer.[3]

### 1.3.2 Face time

Face time est une application de chat vidéo développée par Apple. Apple l'a développé sur un standard ouvert, ce qui signifie que techniquement (sans jeu de mots), Face time pourrait être utilisé sur une gamme de plates-formes, et d'autres fabricants peuvent tirer parti du protocole de Face time. Cependant, en réalité, Fac time reste disponible uniquement pour les utilisateurs de produits Apple.[4]



### 1.3.3 Dans le domaine de l'autorité

La possibilité d'enregistrer des images et du son a naturellement attiré l'attention des autorités militaires. à partir de la fin du XIXe siècle, l'audiovisuel sert à la fois pour les opérations de renseignement militaire et pour la propagande. La Section cinématographique de l'armée (SCA) est créée en 1915.[2]

## 1.4 Les flux de base de L'audiovisuel

### 1.4.1 La vidéo

Une vidéo est une succession d'images à une certaine cadence. L'œil humain a comme caractéristique d'être capable de distinguer environ 20 images par seconde.[4]

### 1.4.2 Codecs

Un codec est un algorithme de compression / décompression d'un signal audiovisuel numérique, La vidéo ou l'audio brut est compressé lors de l'encodage et décompressé (décodé) lors de la lecture. MP3 est un codec audio - une norme de compression que les lecteurs MP3 savent décoder, et les encodeurs MP3 savent encoder.

### 1.4.3 codecs Vidéo

#### 1.4.3.1 Format H264 /AVC

La norme de codage vidéo avancée H.264 / MPEG-4 (H.264 / AVC) est la norme de codage vidéo la plus récente développée conjointement par le Groupe d'experts du codage vidéo UIT-T (VCEG) et le groupe d'experts ISO / IEC Moving Picture (MPEG).[6]

H.264, actuellement l'un des codecs vidéo fréquemment utilisés, est une compression populaire pour la vidéo HD. étant donné que H.264 peut atteindre des vidéos de haute qualité dans des débits binaires relativement bas, il est couramment utilisé dans les caméscopes AVCHD, HDTV, Blu-ray et HD DVD. MP4 (.mp4) est l'un des formats vidéo codés H.264.[7]

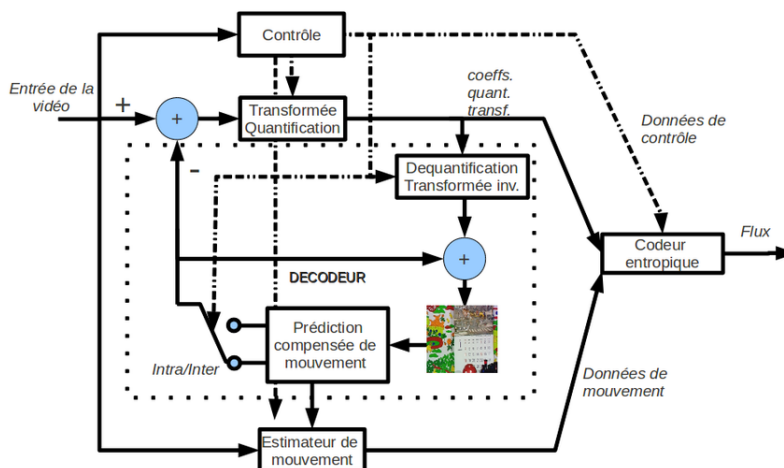


FIGURE 1.1 – Diagramme de bloc de codeur H.264.

### 1.4.3.2 Format MPEG-4

MPEG-4 utilise des techniques similaires à M-JPEG, en ce qui concerne la mise en séquence. Il compare essentiellement deux images compressées, enregistre l'image et enregistre uniquement la différence à partir de chaque image séquentielle supplémentaire, comme le mouvement, ce qui permet d'économiser du temps, de l'espace mémoire et une puissance de traitement.[8] Un taux de compression plus élevé fait partie des avantages de MPEG-4. Il peut synchroniser l'audio et la vidéo, et est idéal pour la visualisation en temps réel. MPEG-4 a été conçu pour prendre en charge les applications à faible bande passante.[7]

### 1.4.3.3 AV1(AOMedia Vidéo 1)

AV1 est actuellement candidat à la normalisation par l'Internet Engineering Task Force (IETF) en tant que vidéo Internet Codec (NetVC). La finalisation du processus de normalisation est prévue fin 2018. Il fournit à la fois un décodeur et un encodeur de référence. Le codeur logiciel AV1 implémente une optimisation de codage non évidente et des algorithmes non normatifs tels que la quantification adaptative ou la structure de sous-gop dynamique. Il peut optimiser le PSNR ou les critères de qualité perceptuelle. Il prend en charge trois modes de fonctionnement RC : VBR, CBR (Rate Control (RC) , Constant Bitrate (CBR), Variable Bitrate (VBR) et Constant (ou de façon équivalente) Qualité.[9]

## 1.4.4 Audio

L'audio est une onde produite par la vibration mécanique d'un support fluide ou solide et propagée grâce à l'élasticité du milieu environnant sous forme d'ondes longitudinales. Par extension physiologique, l'audio désigne la sensation auditive à laquelle cette vibration est susceptible de donner naissance.[10]

## 1.4.5 Codec audio

### 1.4.5.1 Formats audio compressés sans perte

- **RAW** : RAW est un format audio utilisé pour représenter les données de son en modulation d'impulsion codée sans en-tête ni métadonnées.[2]

- **ALAC** (Apple Lossless Audio Codec) C'est un format de codage sans perte de données, créé en 2004 par Apple.[11]

### 1.4.5.2 Formats audio compressés avec perte

- **AC3** : La compression AC3 permet d'utiliser jusqu'à 6 canaux sonores indépendants avec un taux d'échantillonnage de 32, 44,1 ou 48 kHz et avec un taux de transfert allant de 32 à 640 kbit/s. Le Dolby Digital utilise ce principe de codage, c'est pourquoi on le désigne souvent sous ce nom. Format très courant dans les DVD.[2]

- **MP3** : MP3 est l'abréviation de MPEG-1/2 Audio Layer 3, La couche (Layer) III est la couche la plus complexe. Elle est dédiée à des applications nécessitant des débits faibles (128 kbit/s) d'où une adhésion très rapide du monde Internet à ce format de compression. [2]

- **MP3PRO** : Le format mp3PRO, fruit de la collaboration entre Thomson Multimédia et l'Institut Fraunhofer, combine l'algorithme MP3 et un système améliorant la qualité des fichiers comprimés appelé (en)SBR pour Spectral Bandwidth Replication.[2]

### 1.4.5.3 Formats audio non compressés

- **PCM** : La modulation par impulsions et code est une représentation numérique des ondes sonores analogiques. En MIC, l'échantillonnage est utilisé pour convertir les ondes acoustiques sous forme numérique. PCM est instancié par deux capacités principales : la fréquence d'échantillonnage et la profondeur de bits. La fréquence d'échantillonnage permet de mesurer l'amplitude des vagues dans le temps et la profondeur de bits correspond au nombre de bits d'information dans chaque échantillon. PCM est largement utilisé dans la création de CD et de DVD.

- **WAV** : Le format audio Waveform, ou tout simplement WAV, est un format audio brut et non compressé développé par Microsoft et IBM et principalement utilisé sur les plates-formes Windows, WAV a perdu de son attrait en tant que format «de qualité proche de celle d'un CD», mais est toujours très populaire en raison de sa grande disponibilité.

## 1.5 Logiciel pour lire les audiovisuelle

VideoPad Vidéo editor : est un logiciel d'édition vidéo grâce auquel vous pourrez rapidement vous lancer sur Youtube ou Facebook. Il contient toutes les options nécessaires pour effectuer un montage rapide et efficace dont des effets de transition ou encore un effet de stabilisation. VLC media Player : VLC est un Framework et un lecteur multimédia multiplateforme gratuit et open source qui lit la plupart des fichiers multimédias ainsi que des DVD, CD audio, VCD et divers protocoles de streaming

## 1.6 La qualité

La qualité est généralement utilisée dans l'optique d'une ingénierie, car elle est un critère essentiel pour évaluer les systèmes, les services ou les applications au cours des phases de conception et d'exploitation. Fondamentalement, la qualité est le résultat d'un jugement humain basé sur divers critères. La qualité perçue, et plus largement la QoE, devient un élément clé qu'il faut par conséquent savoir mesurer.[12]

### 1.6.1 qualité de l'expérience (QoE)

La QoE est une mesure du jugement personnel de l'utilisateur selon son expérience vécue, sur la qualité globale du service fourni par les opérateurs et fournisseurs de services Internet.

En effet, la notion de l'expérience utilisateur a été introduite pour la première fois par le Dr Donald Norman, évoquant l'importance de la conception d'un service centré utilisateur [13]. Gulliver et Ghinea [14] décomposent la QoE en trois composantes : l'assimilation, le jugement et la satisfaction.

La qualité d'assimilation est une mesure de la clarté du contenu d'un point de vue informatif. Le jugement de qualité reflète la qualité de présentation. La satisfaction indique le degré d'appréciation globale de l'utilisateur.

Pour évaluer la qualité audiovisuel il existe essentiellement deux catégories d'évaluation à savoir les méthodes subjectives qui impliquent des observateurs humains pour évaluer la qualité des contenus multimédias et des méthodes objectives qui calculent la qualité automatiquement à l'aide de modèles mathématiques.[2]

## 1.7 Evaluation de la qualité

### 1.7.1 Evaluation de la qualité subjective

Afin de mesurer de manière fiable la qualité perceptuelle par les systèmes auditifs et/ou visuels humains, les tests subjectifs sont effectués lorsque des groupes d'observateurs humains formés ou naïfs fournissent des cotes de qualité [15]. Cette procédure d'évaluation est connue comme évaluation de la qualité subjective qui vise à quantifier gamme d'opinions que les utilisateurs expriment quand ils voient entendre le contenu numérique et elle est généralement effectuée dans un environnement bien contrôlé 'a l'aide de recommandations normalisées.

Les modèles subjectifs d'évaluation de la qualité des multimédia les plus performants et les plus utilisés sont normalisés par l'Union Internationale des Télécommunications (UIT). Cet organisme est chargé de la normalisation et de la planification des télécommunications dans le monde. Elle établit les normes de ce secteur et diffuse toutes les informations techniques nécessaires pour permettre l'exploitation des services mondiaux de télécommunications. D'autres modèles sont aussi proposés par des laboratoires universitaires ou encore par des sociétés privées, mais ils ne sont pas validés par les normes de l'UIT.[2]

### 1.7.2 Evaluation de la qualité objective

Bien que l'évaluation de la qualité subjective fournisse des indices fiables de la qualité de la perception humaine, il ne peut pas être appliqué dans l'évaluation de la qualité en temps réel en service. Ainsi, les méthodes d'évaluation de la qualité objective ont été mis au point pour remplacer le panneau humain par un modèle de calcul pour prédire les résultats d'un test subjectif. A savoir, l'objectif de l'évaluation objective de la qualité est d'estimer automatiquement les valeurs MOS, qui sont aussi proches que possible de scores de qualité obtenus à partir évaluation de la qualité subjective.[16]

## 1.8 Conclusion

Dans ce chapitre nous avons abordé certaines définitions associées à l'audiovisuel notamment la vidéo et l'audio tout on passe par sa qualité qui contient la qualité subjective et objective et qui seront des points essentiels dans la suite de notre travail. Dans le chapitre suivant nous avons voir les différents métrique d'évaluation de la qualité audiovisuelle (QAV).

# Evaluation de la qualité audiovisuelle

## 2.1 Introduction

Il est possible de mesurer la QoE (Quality of Experience) à l'aide de deux indicateurs : les tests subjectifs ou les métriques objectives. L'évaluation de la qualité par les tests subjectifs consiste à demander à un groupe de personnes d'attribuer une note de qualité au service qu'ils utilisent selon leur degré de satisfaction. D'autre part, le processus d'évaluation est automatisé par des algorithmes dans les mesures dites objectives. Les performances des métriques objectives sont mesurées par rapport aux résultats des tests subjectifs. [17] Dans ce chapitre nous avons présente des notions de base sur l'évaluation de la qualité audiovisuel. Cette qualité contient les tests subjectives et les mériques objective qui discuter au long du chapitre.

## 2.2 Les types d'évaluation de la qualité

Il existe deux types d'évaluation, subjective et objective .[2]

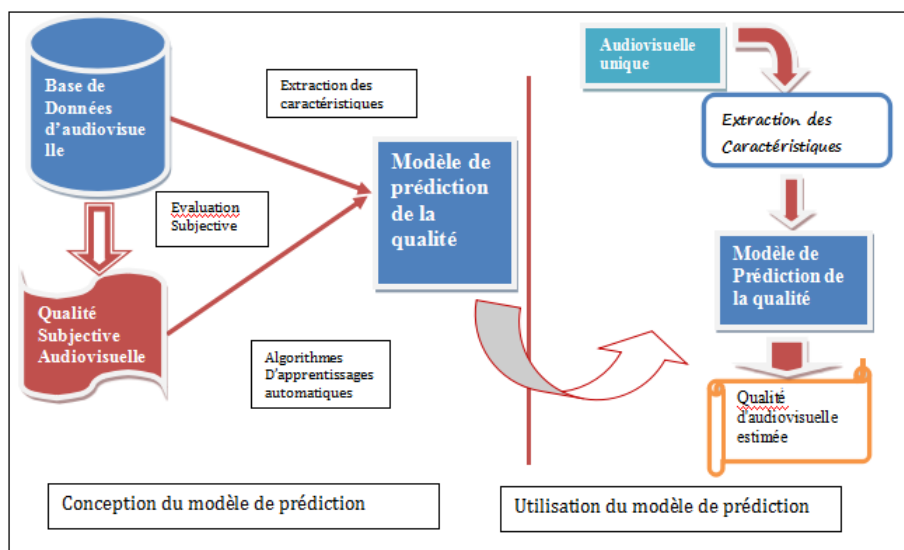


FIGURE 2.1 – Processus d'évaluation objective.[18]

## 2.2.1 Méthodes subjectives

Le moyen le plus naturel d'évaluer la qualité audiovisuelle est de demander à un groupe d'observateurs humains de noter l'audiovisuel suivant un protocole bien défini ; puis d'effectuer la moyenne des notes données par chacun des observateurs : la note moyenne trouvée est appelée à score moyen d'opinion à (MOS : Mean Opinion Score) ou à score moyen d'opinion différentiel à (DMOS : Differential Mean Opinion Score). Cette méthode est appelée méthode d'évaluation subjective du fait que les résultats qu'elle produit sont fortement impactés par la subjectivité du jugement humain.

Plusieurs facteurs implicites influençant l'évaluation de la qualité des audiovisuelle par des observateurs humains, peuvent impacter la note finale donnée à l'audiovisuelle. Les facteurs examinés ci-après permettent de mieux les distinguer

- **L'écran d'affichage** : le choix de l'écran d'affichage, de sa taille, de sa calibration, de sa capacité de reproduction des couleurs, de son contraste. Ainsi est-il important lors d'une évaluation subjective de la qualité d'avoir toutes les informations sur les écrans d'affichage utilisés, afin de tenir compte des artefacts liés aux écrans et à leur calibration.
- **La distance d'observation** : la visibilité de l'audiovisuelle lors de l'évaluation dépend fortement de la distance entre l'observateur et l'écran ; il est recommandé de fixer cette distance entre 4 à 6 fois la hauteur de l'image à projeter, pour permettre une visibilité optimale c'est-à-dire ni très proche, ni très éloigné de l'écran. De plus, une fois fixée, cette distance doit être maintenue tout au long de l'expérience.
- **Les conditions de visualisation** : les salles dans lesquelles sont menées les différentes expériences, l'éclairage, ainsi que les couleurs de fond jouent un rôle important dans l'évaluation de la qualité audiovisuelle. Elles influencent fortement les notes données à audiovisuelle.
- **La durée** : il est préférable de mener des expériences de courte durée pour éviter la fatigue chez l'observateur et partant de fausser les résultats. L'Union Internationale des Télécommunications (ITU) recommande de ne pas faire des séances de plus de 30 minutes, afin d'éviter la fatigue chez l'observateur.
- **Les observateurs** idéalement, le choix des observateurs doit dépendre du domaine d'application visé. Aussi, avant de débiter l'expérience, une phase d'apprentissage doit être faite avec les observateurs pour leur expliquer clairement l'objectif du test ainsi que le protocole utilisé, sans pour autant influencer leur jugement.

Il importe aussi d'avoir un nombre suffisant d'observateurs qui participent à l'expérience. Il reste entendu qu'un test n'est statistiquement valide que si l'on a au moins 16 observateurs pour l'expérience. [69][70]

L'évaluation subjective dépend aussi d'autres facteurs tel que l'humeur, l'âge, la culture, le niveau intellectuel, .... Les méthodes d'évaluation subjectives ainsi que le protocole d'évaluation de ces méthodes subjectives diffèrent suivant la base de données audiovisuelle considérée. On distingue les protocoles à simple et double stimulus, ainsi que des protocoles à stimulus comparatif. [18]

### 2.2.1.1 Protocole d'évaluation

Il y a essentiellement trois grandes familles communes d'évaluation subjective définies Par l'UIT : échelle continue de la qualité sur stimulus double (DSCQS), échelle de dégradation sur stimulus double (DSIS) et évaluation continue de la qualité sur stimulus unique (SSCQE). [2]

1. **Le protocole d'évaluation à simple stimulus** : encore appelé à Single-Stimulus Continuous Quality-Scale à (SSCQS), il permet de juger de la qualité audiovisuelle dégradée par un stimulus unique[18]. La méthodologie à stimulus unique est plus utile dans un environnement de test réaliste, comme les tests conversationnels dans lesquels deux sujets à content et parlent de manière interactive via le système de transmission en cours d'évaluation pour fournir une qualité.[2]

Le processus d'évaluation est le suivant : un audiovisuelle s'affiche, puis l'observateur a un temps de latence pour donner une note à cet audiovisuelle en fonction de l'échelle des notes proposée dans l'expérience, et ensuite on passe à l'audiovisuelle suivante. Des échelles à 5, 6, 10 ou même 100 niveaux peuvent être utilisées.[18] ,En générale, on utiliser la méthode suivant :

- **Méthode ACR : ABSOLUTE CATEGORY RATING**

La méthode ACR ou méthode d'évaluation par catégories absolues consiste a attribuer une note de qualité après chaque séquence AV visualisé e/entendue. La note de jugement attribuée doit rejeter l'opinion du participant quant a la qualité audiovisuelle globale perçue, c'est-a dire la qualité audio et vidéo combinée. Cette évaluation est réalisée sur une échelle caté gorielle de cinq ou neuf points (intervalles) explicitée par cinq items (Excellent-Bon-Satisfaisant-Médiocre -Mauvais). Il est recommande d'utiliser l'échelle en neuf points lorsqu'une plus grande puissance de discrimination est nécessaire, typiquement, lorsque l'on souhaite évaluer des codages a bas débit.[20]

9	Excellent
8	
7	Bon
6	
5	Satisfaisant
4	
3	Médiocre
2	
1	Mauvais

5	Excellent
4	Bon
3	Satisfaisant
2	Médiocre
1	Mauvais

TABLE 2.1 – Echelle d'évaluation de qualité a 9 et 5 niveaux

La norme recommande des séquences d'une durée comprise entre huit et dix secondes, l'intervalle de temps conseille pour le vote est égal ou inférieur à dix secondes. Le chronogramme recommande par la norme UIT-T P.911.[19]

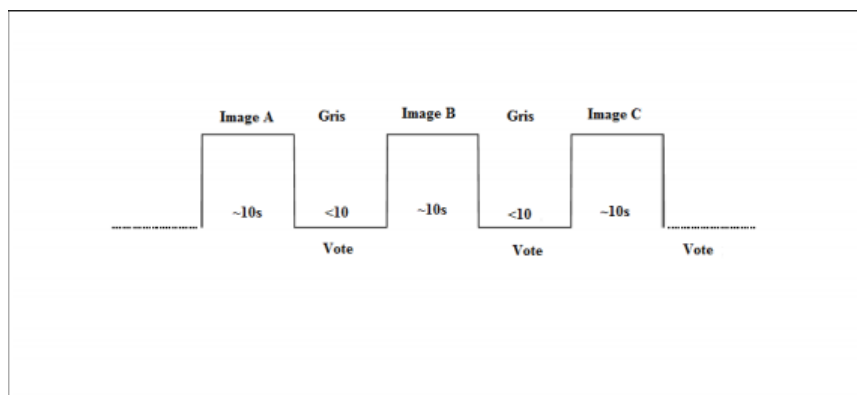


FIGURE 2.2 – Chronogramme de la méthode ACR.[24]

2. **Le protocole d'évaluation à double stimulus** : également appelé à Double Stimulus Continuous Quality-Scale à (DSCQS).[18] Le but principal de la méthode DSCQS est de mesurer la qualité des systèmes par rapport à une référence. Les personnes qui sont montrées paires de séquence audiovisuelle (la séquence de référence et la séquence altérée) dans un ordre aléatoire. Il est largement accepté comme une méthode de test précis avec peu de sensibilité aux effets de contexte, en tant que spectateurs sont présentes deux fois la séquence. Les téléspectateurs sont invités à évaluer la qualité de chaque séquence de la paire après la deuxième projection. Il est également utilisé pour mesurer la qualité du codage audiovisuelle stéréoscopique. [2], la méthode proposée dans le cadre de cette norme est :

**- METHODE DCR : DEGRADATION CATEGORY RATING**

La méthode DCR ou méthode par évaluation de catégories de dégradations propose une présentation des séquences AV de test par paires. Les séquences constituant la paire sont identiques à la différence que la première est toujours présentée sans dégradations (référence) tandis que la seconde est traitée par le système à évaluer (donc susceptible de comporter des dégradations). La séquence traitée est toujours présentée après la référence.[24]

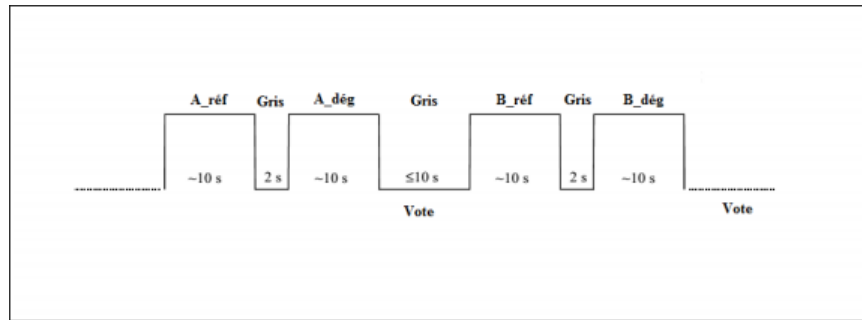


FIGURE 2.3 – Chronogramme de la méthode DCR.[24]

Seule la séquence traitée est évaluée par les participants en comparaison avec la condition de référence. L'échelle d'évaluation correspond ici à une échelle de perceptibilité de la dégradation comme présente par le tableau 2.2.

5	Imperceptible
4	Perceptible
3	Peu dégradée
2	Dégradée
1	Très dégradée

TABLE 2.2 – Echelle de dégradation à cinq niveaux

3. **Le protocole d'évaluation à stimulus comparatif** : les méthodes comparatives permettent d'évaluer la qualité audiovisuelle en fonction d'une ou plusieurs autres audiovisuelles, venant toutes de la même audiovisuelle de référence.[18], comme la méthode suivante :

**METHODE PC : PAIR COMPARISON**

La méthode des comparaisons par paires implique que les séquences d'essai soient présentées en paires. Chaque paire est formée de la même séquence, présentée d'abord au moyen d'un système d'essai puis au moyen d'un autre système. La séquence de référence (sans dégradation) peut être incluse et sera traitée comme un système à l'essai additionnel.



Toutes les combinaisons de paires de séquences A, B, C, etc... Devront être évaluées associées selon toutes les  $n(n-1)$  combinaisons possibles (AB, BA, CA, etc.) et présentées dans les deux ordres possibles (AB, BA, etc.). Le jugement de qualité AV globale est ici exprimé à travers un jugement de préférence pour l'une ou l'autre séquence de la paire qui doit réaliser après la présentation de chaque paire. Cette méthode est notamment préconisée pour la comparaison de systèmes quasi-équivalents et/ou de haute qualité. La durée recommandée pour les séquences de test est d'environ dix secondes, celle du temps de vote doit être inférieure ou égale à dix secondes :[2]

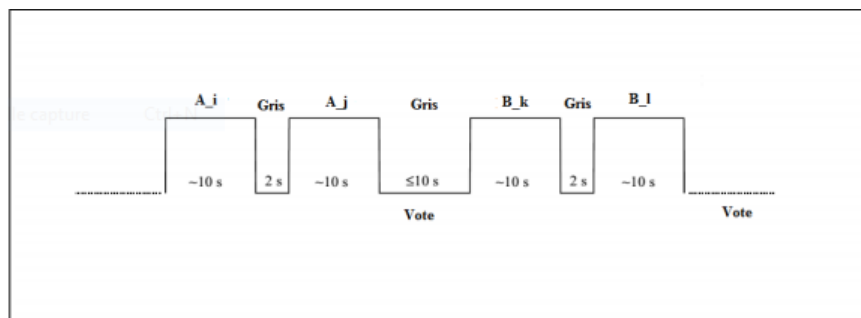


FIGURE 2.4 – Chrono-gramme de la méthode PC.[24]

### 2.2.1.2 Comparaison entre les différentes méthodes de l'évaluation de la qualité audiovisuelle

Lors du choix d'une méthode d'essai, un critère important est la différence fondamentale entre méthodes faisant appel à des références explicites (par exemple DCR) et méthodes ne faisant pas appel à des références explicites (par exemple ACR, PC). Cette deuxième classe de méthodes ne contrôle ni la transparence ni la fidélité.

Il convient d'utiliser la méthode DCR lorsque l'on contrôle la fidélité de transmission par rapport au signal de source. Ce facteur présente souvent de l'importance pour l'évaluation de systèmes de haute qualité. D'autres méthodes peuvent être utilisées pour évaluer les systèmes de haute qualité. Les observations spécifiques de l'échelle DCR (dégradation imperceptible/perceptible) sont précieuses lorsque la détection d'une dégradation par l'observateur est un facteur important.

Lorsqu'il importe de vérifier la fidélité par rapport au signal de source, il convient donc d'utiliser la méthode DCR. La méthode DCR sera également appliquée pour l'évaluation de systèmes de haute qualité, dans le contexte des communications multimédias, cela grâce à la discrimination entre dégradation imperceptible/perceptible sur l'échelle DCR ainsi que grâce à la comparaison avec la qualité de référence.

La méthode ACR est facile et d'application rapide. Sa présentation des stimuli est semblable à celle de l'usage courant des systèmes. La méthode ACR convient donc bien pour des essais de qualification. Le principal mérite de la méthode PC est son haut pouvoir discriminatoire, qui est particulièrement précieux lorsque plusieurs objets d'essai sont de qualité presque égale.

Lorsqu'il faut évaluer un grand nombre d'objets au cours du même essai, la procédure fondée sur la méthode PC tend à être longue. Dans ce cas, un essai ACR ou DCR peut d'abord être effectué avec un nombre limité d'observateurs, suivi d'un essai PC effectué seulement sur les objets qui ont reçu à peu près la même note d'évaluation.[19]

## 2.2.2 Métriques objectives

Bien qu'une évaluation subjective de la qualité indices de qualité de la perception humaine, il ne peut pas être appliqué dans l'évaluation en temps réel de la qualité en service. Ainsi, des méthodes objectives d'évaluation de la qualité ont été développées pour remplacer le panel humain par un modèle informatique pour prédire les résultats d'un test subjectif. à savoir, le but de l'évaluation objective de la qualité est d'estimer automatiquement les valeurs MOS (mean opinion score), qui sont aussi proches que possible des scores de qualité obtenus à partir de l'évaluation subjective de la qualité. Les mesures numériques de la qualité obtenues à partir de la méthode objective (également appelées MOS objectives ou prédites) devraient mieux correspondre à la subjectivité humaine. Il existe différentes mesures pour mesurer la relation entre le MOS subjectif et le MOS prédit. Les deux paramètres statistiques les plus couramment utilisés pour rendre compte des performances des méthodes d'évaluation objective de la qualité sont à l'erreur quadratique moyenne (RMSE) et la corrélation de Pearson un algorithme d'évaluation de la qualité objective ayant une forte corrélation (généralement supérieure à 0,8) est considéré comme efficace.[23]

Deux principaux avantages d'une évaluation objective de la qualité utilisation définissent la signification du MOS pour une application donnée (c.-à-d. que les gens savent ce qu'un MOS de 3 signifie en termes de qualité) et la prédiction reproductible du MOS (c.-à-d.les personnes utilisant l'outil pour les mêmes échantillons de test obtiennent les mêmes résultats). Les techniques objectives de mesure de la qualité peuvent être classées en cinq groupes, en fonction du type de données d'entre utilisées par les paramètres thématiques :

- i. Modèles de couche média :** les modèles de cette catégorie ne exigent des informations sur le système en question. En particulier, ces modèles n'utilisent que des échantillons audio ou vidéo pour estimer la qualité et peuvent être appliqués à des applications telles que l'optimisation et la comparaison de codecs.
- ii. Modèles de couches de paquets paramétriques :** Les solutions pour prédire la qualité dans ce groupe sont légères, car les modèles de couches de paquets paramétriques doivent uniquement traiter les informations d'en-tête de paquet sans traiter avec les médias signaux.
- iii. Modèles de planification paramétrique :** Ces modèles utilisent encodage et paramètres réseaux pour prédire la qualité. Ils demandent donc une connaissance a priori du système Dans la question.
- iv. Modèles de couche binaire :** Ces modèles prédisent la qualité à l'aide des informations codées de la couche binaire et de la couche paquet qui sont utilisées dans les modèles paramétriques de la couche paquet.
- v. Modèles hybrides :** les modèles de cette classe intègrent généralement deux ou plusieurs des modèles mentionnés ci-dessus.  
D'autre part, les techniques objectives d'évaluation de la qualité peuvent également être classées en trois catégories : référence complète (FR), référence réduite (RF) et sans référence (NR) selon la disponibilité de la référence (originale / idéale), des informations partielles sur la référence, ou aucune référence pour évaluer la qualité, respectivement.[21]

### 2.2.2.1 Evaluation objective de la qualité audiovisuel :

Les métriques de qualité objective peuvent être classées en trois catégories principales en fonction de la disponibilité du signal de référence non déformé : référence complète (FR), référence réduite (RR) et sans

référence (NR). Les métriques FR comparent un signal de référence à un signal déformé afin de calculer la différence de qualité entre les deux. Les algorithmes FR sont généralement les plus précis et relativement simples ce qui contribue à leur utilisation généralisée. Cependant, dans de nombreuses applications réelles (par exemple, vidéoconférence, IPTV, etc.), les modèles FR ne peuvent pas être utilisés car le signal de référence n'est tout simplement pas disponible pour la comparaison. Dans de tels cas, les métriques NR sont généralement utilisées. Les méthodes NR sont une mesure absolue des caractéristiques et des caractéristiques d'un signal dégradé et sont souvent axées sur un type de dégradation spécifique (par exemple, flou, bloc) et l'analyse des réglages des paramètres de codage. En raison de l'absence d'un signal de référence, elles peuvent être moins précises que d'autres approches, mais sont plus efficaces à calculer. Les algorithmes RR, au lieu d'une référence complète, utilisent des caractéristiques de qualité extraites de la référence et des signaux déformés. Ces caractéristiques sont ensuite comparées afin de générer un seul score de qualité. Les modèles RR sont généralement adoptés dans les cas où le signal de référence complet ne peut pas être utilisé (par exemple dans une transmission avec une bande passante limitée).

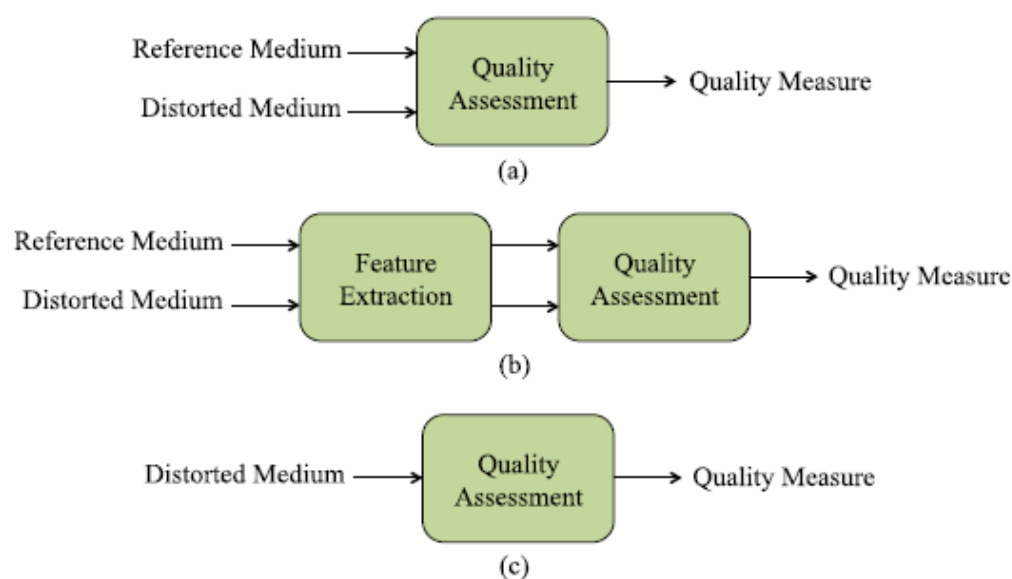


FIGURE 2.5 – aperçu (a) la méthode de référence complète, (b) la méthode de référence réduite, (c) la méthode de non-référence. [21]

### 2.2.2.2 L'approche de prévision de la qualité audiovisuelle :

L'UIT-T a proposé certains modèles normalisés de prévision de la qualité audiovisuelle, par exemple, les Recommandations UIT-T P.1201 [58], UIT-T G.1070 [59] et UIT-TG.1071 [60] :

- Le modèle UIT-T P.1201 : C'est un modèle non intrusif d'information en tête de paquet visant la surveillance de service et le benchmarking du streaming UDP. Le modèle prend en charge les applications à basse résolution comme la télévision mobile et les applications à plus haute résolution comme l'IPTV. Il utilise les informations extraites de l'entête du paquet et les informations fournies hors bande. Il fournit des prédictions distinctes de la qualité audio, vidéo et audiovisuelle sous forme de résultat en terme du MOS à 5 points. Le modèle a été validé pour la compression, la perte de paquets et le buffering des altérations de l'audio et de la

vidéo avec des débits différents. Le modèle ITU-T Rec. P.1201 après des testes . Les valeurs de corrélation RMSE et de Pearson pour la modélisation audiovisuelle ont été évaluées respectivement à 0,470 et 0,852 pour les applications à résolution inférieure et à 0,435 et 0,911 pour les applications à plus haute résolution.

- Le modèle UIT-T G.1070 : ce modèle propose un algorithme a estimé la qualité de l'expérience et la qualité de services. il contient trois fonctions principales d'évaluation de la qualité audio, de la qualité vidéo et de la qualité multimédia globale. La fonction d'estimation de la qualité de la audio prend comme paramètres d'entrée le type de codec vocal, le taux de perte de paquets, le débit binaire et le niveau sonore d'écho de la parole. La fonction vidéo prend comme paramètres d'entrée le format vidéo, la taille d'affichage, le type du codec, le taux de perte de paquets, le débit binaire, l'intervalle d'images et le taux d'images. La fonction multimédia intègre séparément la qualité audio et la qualité vidéo en incluant l'asynchronisme audiovisuel (audiovisual asynchrony) et le délai de bout en bout. Sur des ensembles de données précis, la précision du modèle d'évaluation de la qualité des communications multimédias en terme de corrélation de Pearson est de 0,83 pour QVGA et de 0,91 pour la résolution QQVGA. L'application du modèle est limitée à la planification de la QoE et de la QoS.
- le modèle UIT-T G.1071 :est recommandé pour la planification réseau des services de diffusion audio et vidéo. Cette recommandation concerne les domaines d'application à plus haute résolution (HR) comme l'IPTV et les domaines d'application de résolution inférieure (LR) comme la TV mobile. L'application des modèles est limitée à la planification de la qualité d'expérience (QoE)/qualité de service (QoS). Le modèle prend en entrée les hypothèses de planification de réseau comme la résolution vidéo, les types et profils de codecs audio et vidéo, les débits audio et vidéo et le taux de perte de paquets. Il fournit en sortie des prédictions distinctes de la qualité audio, vidéo et audiovisuelle définies sur l'échelle MOS à 5 points. Notre note que les tests ont montré que les applications à basse résolution utilisant les bases de données d'apprentissage et le test ITU-T P.1201.1. Les valeurs de corrélation RMSE et de Pearson pour la modélisation audiovisuelle ont été évaluées respectivement à 0.5 et 0.83 Pour les applications à haute résolution et 0,51 et 0,87 pour des bases de données d'apprentissage et de validation ITU-T P.1201.2.[2]

### 2.2.2.3 L'approche d'apprentissage automatique :

Le monde de l'apprentissage automatique consiste un nombre incalculable d'algorithmes ainsi que de leurs implémentations dans diverses bibliothèques. Certaines de ces méthodes sont destinées uniquement aux problèmes de classification. Cependant, plusieurs algorithmes sont adaptés à des problèmes de classification et de régression. Ce sont quelques-uns méthodes d'apprentissage automatique

- **Méthodes d'ensemble basées sur l'arbre de décision** : Les arbres de décision (DecisionTrees, DT) sont des structures de données hiérarchiques utilise pour des problèmes de classification et de régression par stratégie de **Diviser-et-conquérir** (divide-and-conquer). Pour évaluer la QAV demribilek a généré deux modèles basés sur les forêts d'arbres décisionnels et deux modèles basés sur les techniques de bootstrap. Ces modèles utilisent comme caractéristiques soit les paramètres indépendants (5 caractéristiques), soit tous les paramètres extraits (34 caractéristiques). Et par la suite il a comparé les résultats obtenus. Pour des besoins de simplification, il nomme le modèle de forêt d'arbres décisionnels qui utilisent tous les paramètres (paramètres indépendants et supplémentaires) par le modèle RF1. Le modèle RF2 va référer au modèle de forêt d'arbres décisionnels qui utilise uniquement les paramètres indépendants. Avec la même logique, il appellera le modèle basé sur les techniques de bootstrap et utilisant tous les paramètres (indépendants et supplémentaires) le modèle BG1. Le modèle basé également sur les techniques de bootstrap et qui utilise uniquement les paramètres indépendants sera nommé par le modèle BG2.[61]

- **Régression symbolique et programmation génétique** : La programmation génétique est une technique de calcul, permet de trouver une solution à un problème sans connaître la forme de la solution. baser par l'évolution d'une population de programmes informatiques ou les populations sont transformées aléatoirement à nouvelles populations génération par génération. Pour évaluer la QAV et découvrir le meilleur modèle demirbilek a généré deux modèles comme Apprentissage profond, un modèle utilise uniquement les variables indépendantes et un modèle utilise tout les variables et après il a comparé les résultats.[61].

- **Apprentissage profond** :

L'apprentissage profond remonte aux années 1940, reflétant l'influence de différents chercheurs et de différentes perspectives. Cette appellation spécifique est très récente, tels que typique d'un modèle d'apprentissage profond est **feed forward Deep Network** ou le perceptron multicouche (Multi-Layer perceptron, MLP) [62]. Dans cette apprentissage demirbilek a généré deux modèles, un modèle utilise uniquement les variables indépendantes et un modèle utilise tout les variables et après il a comparé les résultats, qui découvrira que le modèle utilise uniquement les variables indépendantes ont obtenu de meilleurs résultats que le modèle utilise toutes les variables.

#### 2.2.2.4 L'approche de la fusion des deux modalités :

Les études empiriques montrent que les domaines auditif et visuel ont une influence mutuelle sur la qualité audiovisuelle globale perçue.

Cependant, la majorité des chercheurs ont adopté la théorie de la fusion tardive, dans laquelle les canaux auditifs et visuels sont traités en interne pour produire des valeurs de qualité respectives qui sont intégrées à un stade avancé pour former une seule qualité globale perçue [61].

Les modèles de calcul susmentionnés prédisent automatiquement la qualité perceptuelle à l'aide d'opérations mathématiques. De telles opérations sont souvent faites à l'usage, d'un modèle du système visuel humain (HVS) et du système auditif.

La plupart des mesures de qualité objectives existantes se concentrent uniquement sur une seule modalité, audio ou vidéo, et ne tiennent pas compte de la forte influence mutuelle des deux dans le processus d'évaluation de la qualité. En cas de stimulus multimodales, notre cerveau utilise plusieurs sources d'informations sensorielles dérivés de plusieurs modalités différentes.

La perception multimodale n'est pas une simple combinaison linéaire de perceptions de modalités uniques. La plupart des recherches indiquent qu'à un moment donné du traitement perceptuelle, toutes ces différentes sources d'information s'intègrent pour former un percept cohérent et robuste (fusion perceptuelle). Au cours de ce processus, une modalité peut modifier et compléter la perception dérivée d'une autre modalité. En cas d'évaluation de la qualité multimodale, un tel effet transmodal de l'intégration multi sensorielle peut fortement influencer (positivement ou négativement) la QoE.

La qualité audiovisuelle est donc décrite comme une fusion de deux dimensions (qualités audio et vidéo), comme illustré à la Figure 2.6 :

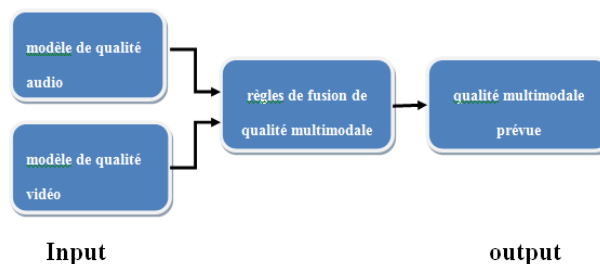


FIGURE 2.6 – Modèle d'estimation de la qualité multimédia de base

Le modèle de fusion le plus couramment utilisé et adopté dans plusieurs études est celui présenté dans l'équation suivante [61][62][63] :

$$QAV = \alpha + \beta QA + \gamma QV + \delta QAQV \quad (2.1)$$

Où QAV, QA, QV et  $\alpha, \beta, \gamma, \delta$  sont respectivement la qualité audiovisuelle, la qualité audio, la qualité vidéo et les poids prédits. Les valeurs rapportées dans la littérature vont de  $\alpha = [-3.34, 4.26]$ ,  $\beta = [-0.19, 0.85]$ ,  $\gamma = [0, 0.89]$ ,  $\delta = [-0.01, 0.26]$ . Peu d'études suggèrent que les canaux audio et vidéo pourraient être intégrés dans une phase précoce de la formation de la perception humaine. Sur cette base, plusieurs chercheurs [61], ont proposé des modèles de Qualité audiovisuelle comme une multiplication de qualité audio et vidéo d'égale importance, comme le montre l'équation suivante :

$$QAV = \alpha + \beta QAQV \quad (2.2)$$

De même, Martineza et al [64], ont proposé trois mesures de la qualité perçue audiovisuelle. Le premier modèle est un modèle linéaire simple donné par l'équation suivante :

$$QAV = \alpha + \beta QA + \gamma QV \quad (2.3)$$

La deuxième mesure est basée sur le modèle pondéré de Minkowski comme suit :

$$QAV = ((\beta QA)^P + (\gamma QV)^P)^{1/P} \quad (2.4)$$

Où les valeurs de la puissance de Minkowski (P) sont toutes comprises entre 1 et 1,2. Sur la base de ces résultats, nous avons varié la valeur de P dans la plage de 0,9 à 1,3 et répété la procédure d'ajustement pour chacune de ces valeurs. Comme certaines études, suggèrent que la modalité visuelle peut être plus dominante que l'audio dans la formation de la qualité audiovisuelle perçue, en particulier pour les vidéos avec des données de mouvement élevées, ainsi les auteurs dans présentent ainsi l'équation suivante :

$$QAV = \alpha + \beta QV + \gamma QAQV \quad (2.5)$$

malgré les modèles dans les équations prétendants atteindre assez précisément la qualité audiovisuelle prévue dans certaines études lorsque des durées de qualité audio et vidéo sont les mêmes, il ne reflète pas les différences de l'influence de seulement audio et vidéo uniquement des stimuli sur la qualité globale.

### 2.2.3 Performance des modèles d'évaluation objective de la qualité audiovisuelle

Un aspect important de la modélisation de la qualité perçue est qu'un modèle objectif ne devrait pas prédire une opinion moyenne subjective de manière plus précise qu'un sujet de test moyen. L'incertitude des votes subjectifs est calculée par l'écart-type et l'intervalle de confiance (IC) correspondant. Ces paramètres statistiques visent à déterminer l'incertitude des sujets par fichier, ou par condition de test [74]. La performance d'un modèle est évaluée via trois métriques statistiques, utilisées pour informer de la précision du modèle, de sa consistance et de sa linéarité/monotonie [74][75] :

- **précision** : saisit la capacité du modèle à prédire les évaluations de qualité subjectives avec de faibles erreurs.
- **consistance** : reflète le degré auquel le modèle maintient l'exactitude des prévisions sur la plage des séquences de test.
- **la monotonie** : correspond au degré auquel les prédictions du modèle conviennent avec l'ampleur relative

des évaluations subjectives de la qualité.

Lorsque les données sont tirées de données de test avec une distribution proche de la normale, ces critères sont obtenus en calculant respectivement l'erreur de prédiction, le rapport de valeurs aberrantes (outlier ratio) et le coefficient de corrélation de Pearson[75]. Lorsqu'il n'est pas possible de vérifier que les données sont tirées d'une distribution proche de la normale, le coefficient de Spearman Rank est utilisé dans la littérature au lieu du coefficient de corrélation de Pearson comme mesure de la monotonie [74][75]. il est recommandé d'utiliser l'erreur de prédiction pour la précision, le rapport de valeurs aberrantes (OR) ou la distribution d'erreur résiduelle pour la cohérence et le coefficient de corrélation de Pearson pour la linéarité[74].

### 2.2.3.1 Exactitude du modèle (précision) :

L'erreur de prédiction (c'est-à-dire l'exactitude) est obtenue à l'aide de l'erreur quadratique moyenne (RMSE) (Root Mean Square Error). La précision d'un modèle est habituellement déterminée par une interprétation statistique de la différence entre les valeurs MOS du test subjectif et sa prédiction sur une échelle généralisée. Un modèle précis a pour but de prédire la qualité avec l'erreur la plus faible en terme de RMSE lors des tests subjectifs [74].

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Predicted_i - Actual_i)^2}{N}} \quad (2.6)$$

Où  $i$  est l'index de la séquence, et  $N$  est le nombre de séquences utilisées pour comparer les scores de qualité estimés aux scores subjectifs tandis que la division à  $(N - 1)$  assure un estimateur sans biais pour rmse avec un intervalle de confiance à 95%.

Le rmse est approximativement caractérisé par un  $\chi^2(n)$ , où  $n$  représente les degrés de liberté et est défini par l'équation :  $n=N - d$ , où  $d = 4$  indique les degrés de liberté de la fonction de cartographie (fonction polynomiale de 3 e ordre). En utilisant la distribution  $\chi^2(n)$ , l'intervalle de confiance de 95 % pour la rmse est donné par l'équation[74] :

$$\frac{(rmse\sqrt{N-1})}{\sqrt{\chi_{0.975}^4(N-d)}} < rmse < \frac{(rmse\sqrt{N-1})}{\sqrt{\chi_{0.025}^4(N-d)}} \quad (2.7)$$

Cette mesure de rmse dépend de l'échelle de notation utilisée lors des tests subjectifs. Par conséquent, pour comparer deux valeurs rmse, les scores de qualité doivent d'abord être convertis à la même échelle. rmse est toujours positif, et des valeurs rmse plus faibles indiquent une plus grande précision. Pour tenir compte du degré d'incertitude des jugements des sujets, les valeurs de qualité dites epsilon-modifiées Root-Mean-Square-Error (rmse\*) entre les valeurs prédites et subjectives peuvent être calculées à la place de rmse. Rmse\* est similaire à rmse, mais avec :

$$Perror(i) = \max(0, |Mos(i) - Mosp(i)| - ci95(i)) \quad (2.8)$$

Où  $ci95$  est l'intervalle de confiance à 95 % de la séquence  $i$ .

Cette métrique pour comparer les performances des modèles sur la base de bases de données de tests subjectifs avec des intervalles de confiance très variables [75] pour l'évaluation de la qualité à référence complète.

### 2.2.3.2 Consistance du modèle (cohérence et) :

La consistance du modèle est obtenue en calculant soit le rapport des valeurs aberrantes (Outlier Ratio OR), soit la distribution des erreurs résiduelles [74][75].

$$OR = \frac{TotalNoOutliers}{N} \quad (2.9)$$

Les valeurs OR sont définies comme les points pour lesquels l'erreur de prévision  $P_{error}$  dépasse l'intervalle de confiance de 95 % de la valeur MOS moyenne, c-à-d.

$$|P_{error}(i)| > \frac{z\sigma(MOS(i))}{\sqrt{N_{subj}}} \quad (2.10)$$

$$\sigma(MOS(i)) = \sqrt{\frac{MOS(i)(1 - MOS(i))}{N}} \quad (2.11)$$

Où  $\sigma(MOS(i))$  représente l'écart-type des scores individuels associés à l'échantillon de médias  $i$ , et  $N_{subj}$  est le nombre d'électeurs par échantillon de médias  $i$ . La limite d'intervalle de confiance de 95 % définie par la variable  $z$  est déterminée en fonction de  $N_{subj}$ . Si  $N_{subj} > 30$ , alors la distribution gaussienne peut être utilisée, et donc  $z=1.96$ . Si  $N_{subj} < 30$ , la distribution t-Student est utilisée et la variable  $z = t$  et sa valeur dépend du  $N_{subj}$ , respectivement le degré de liberté  $df=N_{subj} - 1$  [74][75].

### 2.2.3.3 Modèle monotonie (linéarité) :

Dans la littérature, deux métriques couramment utilisées pour le calcul de la linéarité d'un modèle existent : le coefficient de Spearman et le coefficient de corrélation de Pearson. Le coefficient de corrélation de Pearson est utilisé chaque fois que les données échantillonnées ont une distribution presque normale. Dans d'autres cas, le coefficient de Spearman est utilisé pour qualifier la linéarité entre les scores de qualité subjective prédits et réels [74][75].

$$R = \frac{\sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})}{(\sum_{i=1}^n (X_i - \bar{X})) * (\sum_{i=1}^n (Y_i - \bar{Y}))} \quad (2.12)$$

$X_i$  indique le score subjectif MOS et  $Y_i$  le score objectif (MOSp).  $N$  représente le nombre total d'échantillons pris en compte dans l'analyse.

Le coefficient de corrélation de Spearman est défini comme suit [74][75] :

$$R_s = \frac{\sum_{i=1}^n (RO(i) - \bar{RO})(RO_e(i) - \bar{RO}_e)}{(\sum_{i=1}^n (RO_i - \bar{RO})^2) * (\sum_{i=1}^n (RO_e - \bar{RO}_e)^2)} \quad (2.13)$$

Cette formule est similaire au coefficient de corrélation de Pearson, sauf le fait que l'ordre de classement des



scores (rank order) de qualité subjectifs (RO(i)) et pr édites (ROe(i)) est pris au lieu des scores de qualité eux-mêmes. Cette métrique mesure donc si l'augmentation (diminution resp.) d'une variable est associée à l'augmentation (diminution resp. ) de l'autre variable, indépendamment du surface de l'augmentation (diminution resp. ). Cette mesure est une mesure non paramétrique de la monotonie [74][75] :

$$z = 0.5 \ln\left(\frac{1+R}{1-R}\right) \quad (2.14)$$

$$\sigma_z = \sqrt{\frac{1}{N-3}} \quad (2.15)$$

L'intervalle de confiance de 95 % pour le coefficient de corrélation est déterminé à l'aide de la distribution gaussienne, qui caractérise la variable z et est donnée par l'équation [74][75] :

$$z \pm 1.96 * \sigma_z \quad (2.16)$$

## 2.3 Conclusion

Dans ce chapitre nous avons présenté les méthode d'évaluation de la qualité audiovisuel qui se divise aux deux familles importantes : les méthodes subjectives qui se divisent sur trois protocoles d'évaluation : protocole d'évaluation simple stimulus qui contient la méthode ACR, protocole d'évaluation double stimulus qui contient la méthode DCR, protocole d'évaluation stimulus comparatif qui contient la méthode PC et les méthodes objectives basée trois approches : L'approche de prévision de la qualité audiovisuelle, L'approche d'apprentissage automatique, L'approche de la fusion des deux modalités .

# logique floue

## 3.1 Introduction

La logique floue est apparue comme un outil très puissant pour traiter des problèmes complexes. Récemment, le rôle de l'inférence dans la gestion de l'incertitude dans les applications d'ingénierie gagne en importance. Les ingénieurs et les scientifiques sont généralement confrontés à des problèmes impossibles à résoudre numériquement à l'aide de règles mathématiques traditionnelles. En utilisant la logique floue, on peut caractériser et contrôler un système dont le modèle n'est pas connu ou est mal défini.[25]

La théorie floue a la capacité de saisir l'imprécision des termes linguistiques dans les énoncés du langage naturel. Cela a fourni une plus grande capacité à modéliser le raisonnement de bon sens humain et la prise de décision [26]. La figure 10 définit Schéma d'ensemble d'un système floue.

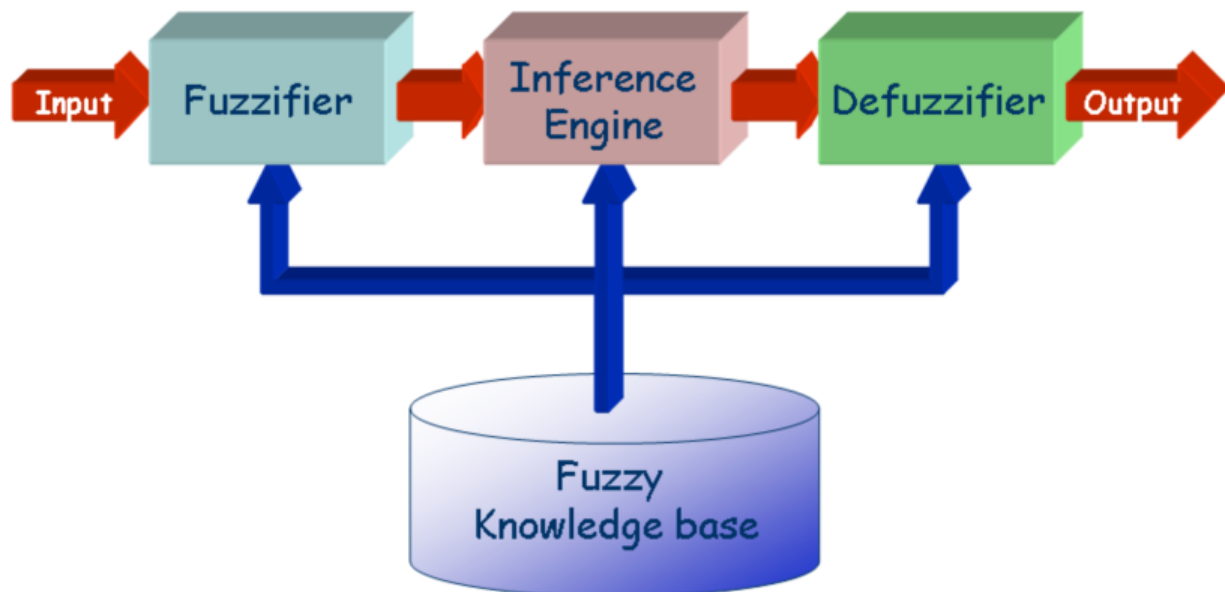


FIGURE 3.1 – Schéma d'ensemble d'un système floue [27]

-**fuzzification** : à ce stade, l'entrée système est transformée de son nombre net en ensembles flous

- **Knowledge base (rules)** : qui est représentée par un ensemble de règles IF-THEN fournies par des experts.
- **Inferences engine** : qui simule le processus de raisonnement humain. Cela se fait en faisant une inférence floue sur les entrées en utilisant les règles.
- **défuzzification** : il transforme l'ensemble flou obtenu par le moteur d'inférence en une valeur nette.[28]

## 3.2 Historique de la logique floue

Depuis longtemps l'homme cherche à maîtriser les incertitudes et les imperfections inhérentes à sa nature. La première réelle manifestation de la volonté de formaliser la prise en compte des connaissances incertaines fut le développement de la théorie des probabilités à partir du XVII<sup>e</sup> siècle. Mais les probabilités ne peuvent maîtriser les incertitudes psychologiques et linguistiques. On a donc assisté aux développements des théories de probabilité subjective (dans les années 50) puis de l'évidence (dans les années 60).[29] Puis la Logique Floue est apparue en 1965 à Berkeley dans le laboratoire de Lotfi Zadeh avec la théorie des sous-ensembles flous puis en 1978 avec la théorie des possibilités. Ces deux théories constituent aujourd'hui ce que l'on appelle Logique Floue .[30]

La Logique Floue permet la formalisation des imprécisions dues à une connaissance globale d'un système très complexe et l'expression du comportement d'un système par des mots.

Elle permet donc la standardisation de la description d'un système et du traitement de données aussi bien numériques qu'exprimées symboliquement par des qualifications linguistiques .[31]

## 3.3 Logique floue vs logique classique

Dans la logique classique, les variables gérées sont Booléennes. C'est à dire qu'elles ne prennent que deux valeurs 0 ou 1 [32]. C'est-à-dire que la réponse de l'instruction est 0 ou 1. De la manière de l'échantillonneur, elle est définie comme une valeur vraie ou fausse. Mais en cas de flou nous pourrions prendre la valeur intermédiaire. En valeur courte entre 0 ou 1.

Pour une meilleure compréhension, prenons l'exemple :

Supposons que quelqu'un court. Prenons la valeur normalisée de la vitesse de course dans une plage de 0 à 1. Supposons que la vitesse soit supérieure à 0,5, alors considérons-la rapidement Et si la vitesse est inférieure à 0,5, elle est considérée comme lente Ici, la valeur nous définissons deux valeurs nettes qui sont lentes et rapides en définissant une valeur de seuil comme 0,5 Ce n'est rien d'autre que la logique classique. Définissez comme rapide ou lent. En cas de logique floue. Nous pouvons prendre une valeur intermédiaire comme lente, moyenne, rapide et très rapide. Ici, nous définissons 4 valeurs au lieu de seulement 2 valeurs. Avec floue, nous pouvons définir plus que ce que nous avons défini ici dans cet exemple :

De 0 à 0,25 -> lent.

Pour 0,25 à 0,5 -> moyen.

Pour 0,5 à 0,75 -> rapide.

Pour 0,75 à 1 -> très rapide.

La valeur intermédiaire ci-dessus relève de la logique floue.

La logique floue est comme la décision de notre cerveau humain. Supposons que quelqu'un vous a donné différents plats à tester et demandez votre avis pour ces plats.

## 3.4 Théorie de la logique floue

### 3.4.1 Les sous-ensembles flous

La logique floue repose sur la théorie des ensembles flous, qui sont une généralisation de la théorie des ensembles classiques. Les ensembles classiques sont également appelés ensembles nets, par opposition à floue, et de même la logique classique est également appelée logique booléenne ou binaire :[33]

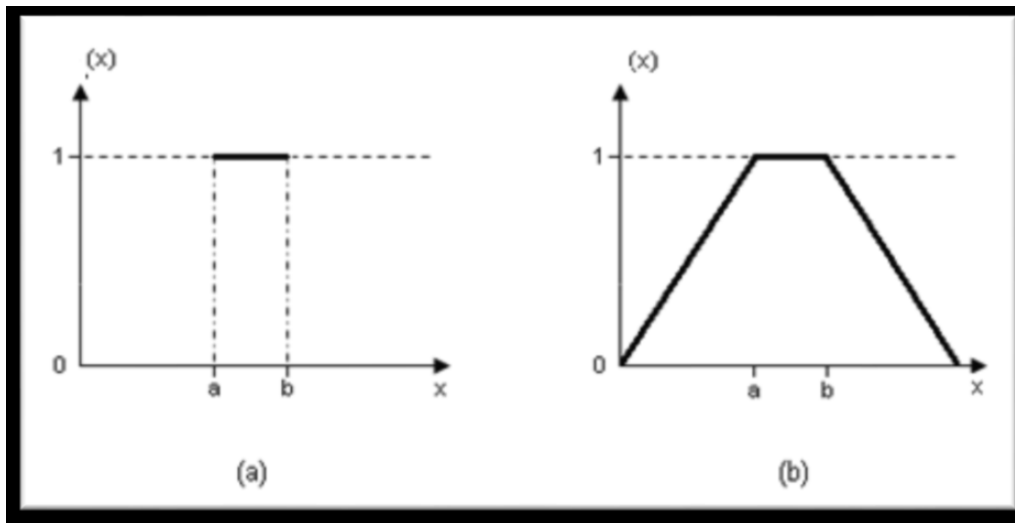


FIGURE 3.2 – Fonction d'appartenance caractérisant un ensemble classique (a) et d'un ensemble floue (b).

### 3.4.2 Les opérations de base sur les sous-ensembles flous

La théorie mathématique sur les sous-ensembles flous définit de nombreuses opérations sur ces sous-ensembles et sur les fonctions d'appartenance qui rendent ces notions utilisables.

Nous ne présentons ici que les opérations de base de cette théorie .[30]

Si A et B sont deux sous-ensembles flous et  $\mu(A)$  et  $\mu(B)$  leur fonction d'appartenance, on définit :

**-Le complémentaire de A, par la fonction d'appartenance :**

$$\mu(\bar{A}) = 1 - \mu(A) \quad (3.1)$$

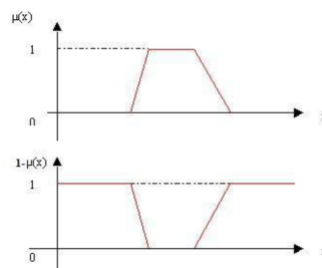


FIGURE 3.3 – Fonction d'appartenance

-Le sous-ensemble A et B, par la fonction d'appartenance :

$$(A \cap B) = \min(\mu(A), \mu(B)) \quad (3.2)$$

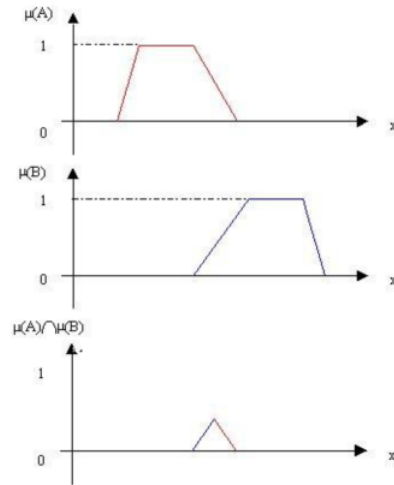


FIGURE 3.4 – Intersection des fonctions d'appartenances

-Le sous-ensemble A ou B,  $A \cup B$ , par la fonction d'appartenance :

$$(A \cup B) = \max(\mu(A), \mu(B)) \quad (3.3)$$

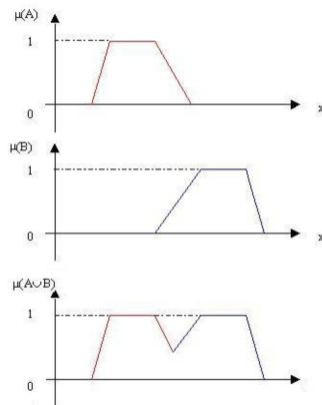


FIGURE 3.5 – Union des fonctions d'appartenances

Ces définitions sont celles qui sont les plus communément utilisées mais parfois, pour certains cas, d'autres sont plus appropriées. Par exemple, l'intersection peut être définie par le produit des fonctions d'appartenance et l'union par la moyenne arithmétique des fonctions d'appartenance. Ces différentes techniques de calcul engendrent une énorme capacité d'adaptation des raisonnements flous [33].

**Opérateur ET :** L'opérateur ET dans la logique floue correspond à l'opération minimum.

$$\mu_E(z) = \mu_A \quad \text{et} \quad B(z) = \beta \quad \min[\mu_A(x), \mu_B(y)] + \left[\frac{(1-\beta)}{2}\right][\mu_A(x) + \mu_B(y)] \quad [34] \quad (3.4)$$

**-Opérateur OU :** Selon la théorie des ensembles, l'opération OU est utilisée souvent pour exprimer l'union entre deux ensembles. Dans la logique floue, cette opération est définie par l'opération maximum.

$$\mu_O(z) = \mu_A \quad \text{et} \quad B(z) = \beta \quad \max[\mu_A(x), \mu_B(y)] + \left[\frac{(1-\beta)}{2}\right][\mu_A(x) + \mu_B(y)] \quad [34] \quad (3.5)$$

Avec le facteur  $\beta \in [0, 1]$ , il est possible de pondérer l'influence des deux termes. Pour  $\beta = 1$ , on aboutit respectivement à l'opérateur minimal ou maximal. Par contre, pour  $\beta = 0$ , on obtient pour les deux opérateurs la moyenne arithmétique correspondant à l'opérateur somme selon 3.3 et 3.4

### 3.4.3 Les variables linguistiques

Une variable linguistique est une variable définie à base de mots ou des phrases au lieu des nombres. En effet, la description d'une certaine situation, d'un phénomène ou d'un procédé contient en général des expressions floues comme "quelque, beaucoup, souvent, chaud, froid, rapide, lent, grand, petit etc.". Ce genre d'expressions forme ce qu'on appelle des variables linguistiques de la logique floue[35]

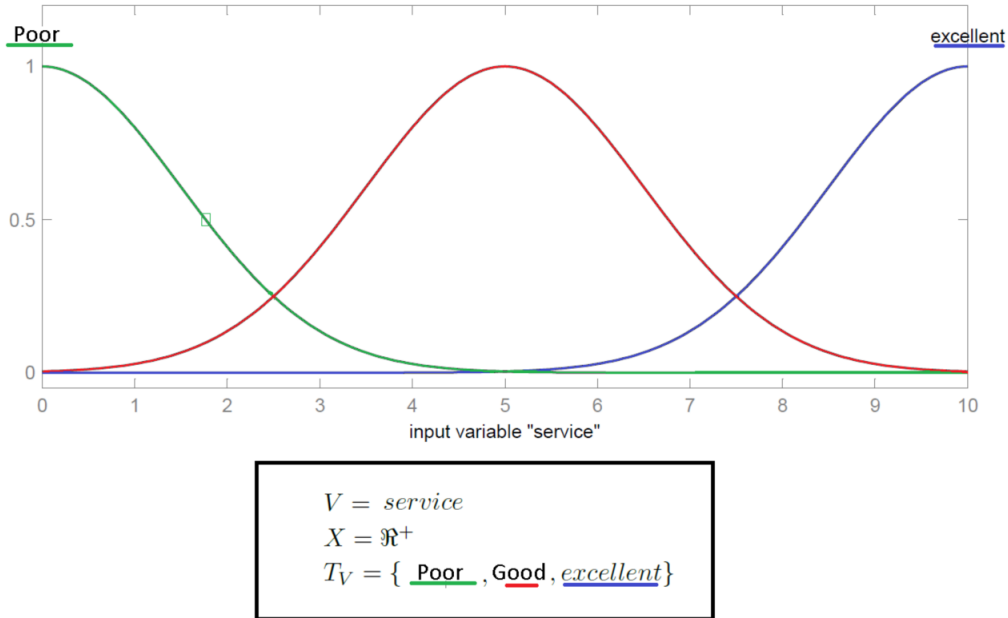


FIGURE 3.6 – Variable linguistique «par exemple la description de la qualité du service

## 3.5 Les types de la logique floue

### 3.5.1 Type-1 logique floue

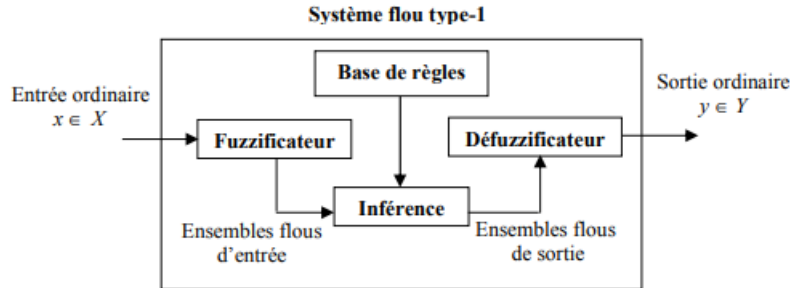


FIGURE 3.7 – Structure générale d'un contrôleur flou type-1.[36]

#### 3.5.1.1 Fuzzification

Elle transforme toute variable numérique d'entrée en variable floue.

**Base de règles :** En logique floue, les règles stipulent la relation entre les ensembles flous d'entrée et les ensembles flous de commande correspondante. Une règle prend habituellement la forme d'instruction IF -THEN comme suit :

**IF**  $x_1$  is  $A_1^1$  **and**... $x_p$  is  $A_p^1$ , **THEN**  $y$  is  $Y^1$   $1 = 1...M$   $x \in X$  ,  $y \in Y$  [41]

Où A et Y sont des valeurs linguistiques définies par des ensembles flous. Par conséquent, le nombre de règles définies dépend directement du nombre de sous ensembles défini pour chaque variable d'entrée et de sortie.[37]

**Définition des fonctions d'appartenance :** Un ensemble flou est défini par sa fonction d'appartenance qui correspond à la notion de fonction caractéristique en logique classique, elle permet de mesurer le degré d'appartenance d'un élément à l'ensemble flou. En toute généralité, une fonction d'appartenance d'un ensemble flou est désignée par  $\mu_A(x)$ . L'argument  $x$  se rapporte à la variable caractérisée, alors que l'indice  $A$  indique l'ensemble concerné [33].

Les fonctions d'appartenance peuvent avoir différentes formes :

a) **Fonction d'appartenance triangulaire ;**

$$\mu(X) = \begin{cases} \frac{X-a}{b-a} & a < X \leq b \\ \frac{c-X}{c-b} & b < X \leq c \\ 0 & \text{ailleurs} \end{cases} \quad (3.6)$$

b) **Fonction d'appartenance trapézoïdale ;**

$$\mu(X) = \begin{cases} \frac{X-a}{b-a} & a < X \leq b \\ 1 & b < X \leq c \\ \frac{X-a}{b-a} & c < X \leq d \\ 0 & \text{ailleurs} \end{cases} \quad (3.7)$$

c) **Fonction d'appartenance gaussienne.**

$$\mu(X) = EXP \left[ - \left( \frac{X - m}{\delta} \right)^2 \right] \quad -\infty < X < +\infty \quad (3.8)$$

La Figure (3.8) représente les formes de ces trois types de fonctions d'appartenance :

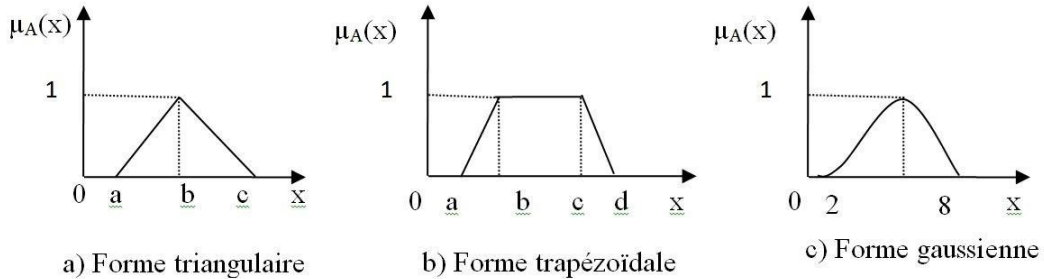


FIGURE 3.8 – Différentes formes de la fonction d'appartenance

**3.5.1.2 Mécanisme d'inférence :**

Consiste à calculer les valeurs de vérité des différentes règles et à associer à chacune d'elle une valeur de sortie qui dépend de sa partie conclusion. L'utilisation des opérateurs et / ou et l'implication dans les règles floues, tels qu'ils sont définis par les fonctions élémentaires (min, max, produit), Il existe plusieurs possibilités pour réaliser les opérateurs qui combinent les valeurs d'entrée et les valeurs de sortie, C'est ce qu'on appelle la méthode d'inférence .les méthodes les plus utilisées sont :

- Méthode d'inférence MAX-MIN.
- Méthode d'inférence MAX-PROD.
- Méthode d'inférence SOMME-PROD. [30]

**- Méthode d'inférence Max-Min :**

Cette méthode réalise l'opérateur "ET" par la fonction "Min", la conclusion "ALORS" de chaque règle par la fonction "Min" et la liaison entre toutes les règles (opérateur "OU») par la fonction Max. La dénomination de cette méthode, dite Max-Min ou "implication de Mamdani", est due à la façon de réaliser les opérateurs ALORS et OU de l'inférence [38] :

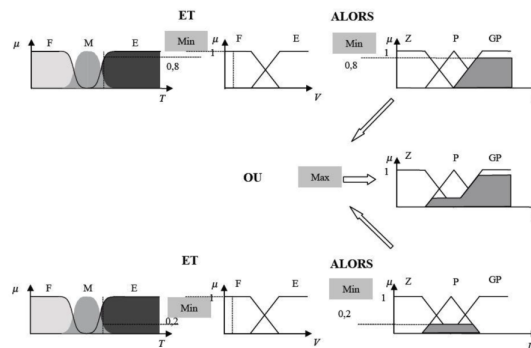


FIGURE 3.9 – Exemple d'inférence Max-Min



### 3.5.1.3 Défuzzification type 1- FL :

Cette étape consiste à réaliser l'opération inverse de la fuzzification, c'est-à-dire, obtenir une valeur physique de la sortie à partir de la surface obtenue. Plusieurs méthodes de défuzzification existent. Comme pour tous les opérateurs flous, le concepteur du système doit choisir parmi plusieurs définitions possibles de défuzzification[33]. Les plus utilisées sont :

- Méthode du maximum
- Méthode de la moyenne des maximums
- Méthode du centre de gravité
- **Méthode du centre de gravité** La méthode du centre de gravité est la plus utilisée. Cette méthode consiste à trouver le centre de gravité de la surface obtenue. L'abscisse du centre de gravité de la sortie peut se déterminer à l'aide de la relation générale[35],[39] :

$$\mu = \frac{\int_{x_0}^{x_1} X\mu(X)dx}{\int_{x_0}^{x_1} \mu(X)dx} \quad (3.9)$$

L'intégrale au dénominateur donne la surface, tandis que l'intégrale au numérateur correspond au moment de la surface.

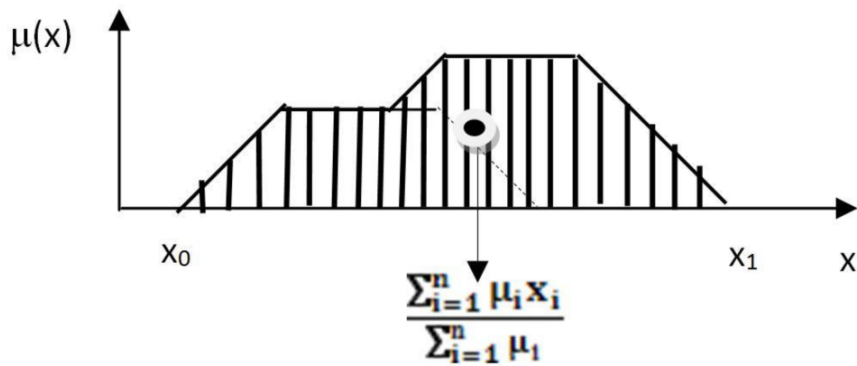


FIGURE 3.10 – Défuzzification par le centre de gravité

- **Méthode de la moyenne des maximums** La méthode est également connue comme le milieu des maxima Dans cette méthode, la valeur défuzzifiée est considérée comme l'élément ayant les valeurs d'appartenance les plus élevées. Lorsqu'il y a plus d'un élément ayant des valeurs d'appartenance maximales, la valeur moyenne des maxima est prise. La sortie est définie par :

$$x^* = \frac{\sum_{i=1}^n \bar{x}_i}{n} \quad [40] \quad (3.10)$$

Dans l'équation 3.9, la sortie défuzzifiée serait :

$$x^* = \frac{a + b}{2} \quad (3.11)$$

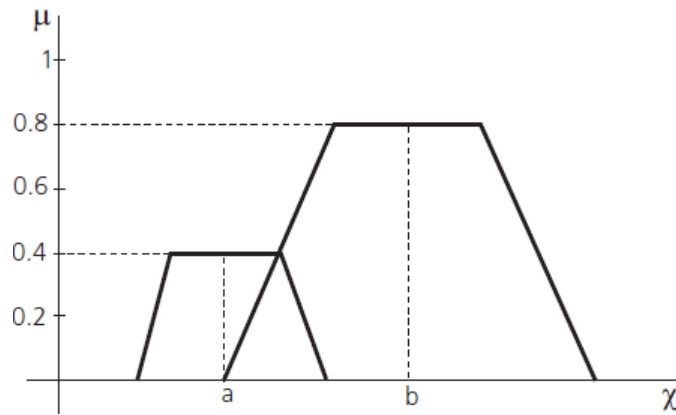


FIGURE 3.11 – fonctions d’appartenance symétriques.  
[40]

- **Méthode du maximum** Cette méthode est également connue sous le nom de méthode de la hauteur. La sortie est définie par[40] :

$$\mu_F(x^*) \geq \mu_F(x), \text{ for all } x \in X \tag{3.12}$$

### 3.6 Type -2- logique floue :

Une logique floue de type-2 inclut au moins un ensemble flou type-2 dans les fonctions d’appartenance définissant ses prémisses ou ses conséquences. L’inférence floue, dans ce type de système, conduit à l’obtention d’ensembles flous de sortie de type-2. La structure d’un système flou type-2 est très similaire à celle d’un système flou type1 La seule différence.

Dans le cas du type-2, le bloc de défuzzification est précédé d’un bloc réduction de type(figure3.12)[36] :

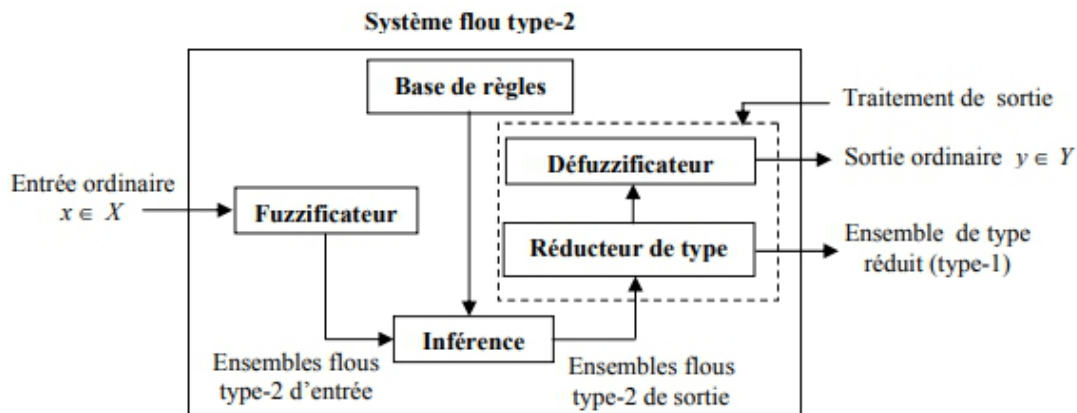


FIGURE 3.12 – Structure d’un système de logique floue de type 2.  
[36]

#### 3.6.1 Fuzzification :

- **Base de règles** La structure des règles dans un FL de type 1 et un FL de type 2 est la même, mais dans ce dernier, les antécédents et les conséquents sont représentés par des ensembles flous de type 2 :

**IF**  $x_1$  is  $\tilde{A}_1^1$  **and**...  $x_p$  is  $\tilde{A}_p^1$  **THEN**  $y$  is  $\tilde{y}^1$  1 = 1... M [41]

- **Les fonctions d'appartenance** Les bases de la logique floue ne changent pas d'ensembles flous de type 1 à type 2 et, en général, ne changeront pour aucun type n (Karnik et Mendel 1998)[41] . si le nombre de type plus élevé il modifie la nature des fonctions d'appartenance, les opérations qui dépendent des fonctions d'appartenance changent, si on appelle MF de type 1 « primaire MF » alors on appelle MF de type 2 « secondaire MF»,Le FL de type 2 peut également être représenté sur la base de triangulaires, gaussiens, trapézoïdaux.

- **Fonction d'appartenance gaussienne**

$$\mu_2(x, \mu_1) = e - \frac{1}{2} \left[ (\mu_1 - m(x)) / \sigma m \right]^2 \quad \mu_1 \in [0, 1] [41] \quad (3.13)$$

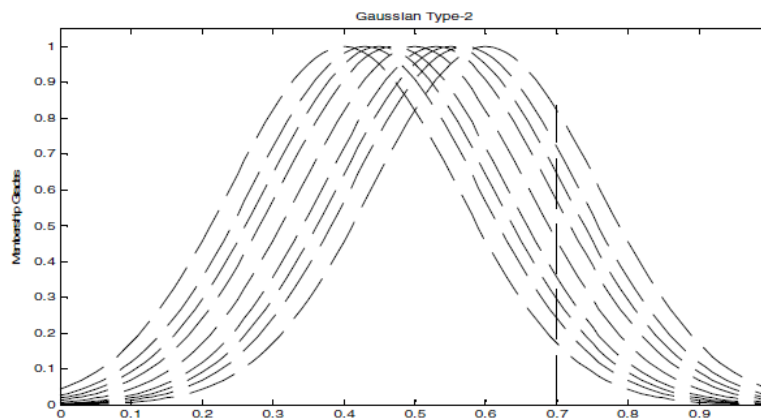


FIGURE 3.13 – forme gaussienne type 2 [41].

- **Footprint of uncertainty** it is the union of all primary membership functions
- **Fonctions d'appartenance supérieures et inférieures** Une «fonction d'appartenance supérieure» et une «fonction d'appartenance inférieure» sont deux fonctions d'appartenance de type 1 qui sont des bornes pour le FOU d'un ensemble flou de type 2  $\tilde{A}$ . La fonction d'appartenance supérieure est associée à la limite supérieure de FOU ( $\tilde{A}$ ). La fonction d'appartenance inférieure est associée à la limite inférieure de FOU ( $\tilde{A}$ ). [41]

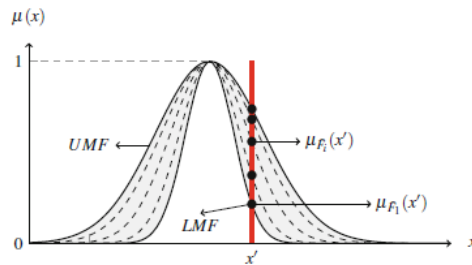


FIGURE 3.14 – Footprint of uncertainty et Fonctions d'appartenance supérieures et inférieures.[42].

### 3.6.2 Mécanisme d'inférence :

Les méthodes inférence pour FL type 2 le même pour le FL type1, la seule différence le type 2 les valeurs d'entrée et les valeurs de sortie sont des ensembles  $\tilde{A}$  donc il utilise les Fonctions d'appartenance supérieures et inférieures pour efficacité de ces méthodes, nous donnons un exemple par la Méthode d'inférence Max-Min permis les méthodes les plus utilisé :

- **Méthode d'inférence Max-Min** Cette méthode réalise l'opérateur "ET" par la fonction "Min", la conclusion "ALORS" de chaque règle par la fonction "Min" et la liaison entre toutes les règles (opérateur "OU») par la fonction Max [38] :

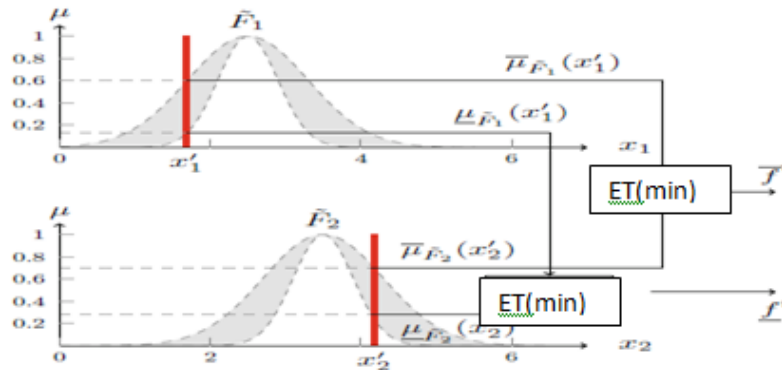


FIGURE 3.15 – utilise fonctions d'appartenance supérieures et inférieures pour opérateur « ET » .[42]

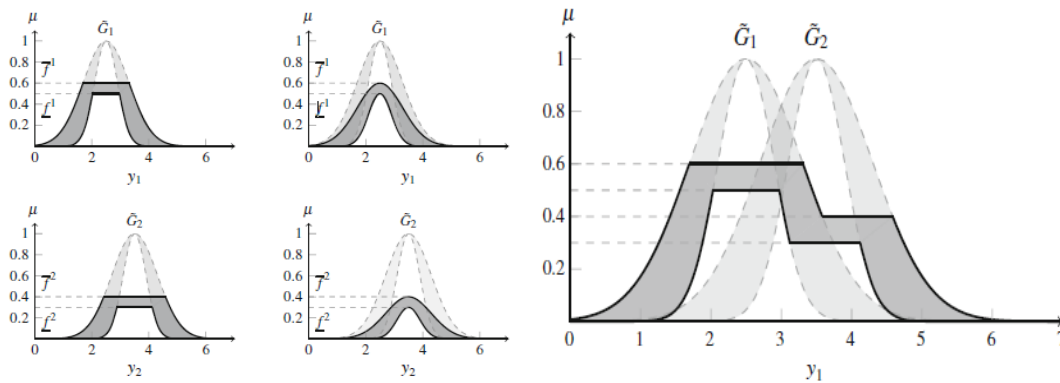


FIGURE 3.16 – Exemple d'inférence Max-Min.[42]

### 3.6.3 Type-Réduction :

Les sortie d'inférence c'est ensemble de type 2 et defuzzification permet d'obtenir une valeur numérique (ordinaire). Cette opération transforme un ensemble flou type-2 à un ensemble flou type-1, les type de réduction sont les méthodes de défuzzification de type1, sauf la sortie des méthodes de type 2 donne deux point final, qu'il defuzzifié pour obtenir un nombre ordinaire.]

- **Centre d'un ensemble Type-Réduction** Dans cette section, nous allons généraliser le concept du centroïde d'un ensemble flou type-1 vers celui du type-2. est généralement la méthode employée,

car il effectue une plus petite quantité d'opérations, la seule procédure qui doit être effectuée après chaque nouvelle entrée dans le système est une moyenne pondérée des centroïdes stockés selon une combinaison des fonctions de MF supérieur et inférieur de chaque règle. De la même manière que la procédure de défuzzification du Centroid, le Centroid TR commence par obtenir  $K$  échantillons d'un FL de Type-2 à travers le FOU d'un FL de Type-2 embarque plusieurs FL de Type-1, pour effectuer le TR, il faut obtenir deux Type-1 FL, et la meilleure façon est les limites supérieures et inférieures du centroïde FL de type 2. la procédure commence en utilisant ses bornes supérieure et inférieure échantillonnées pour trouver les valeurs optimales pour les points de commutation  $[L, R]$  [42] :

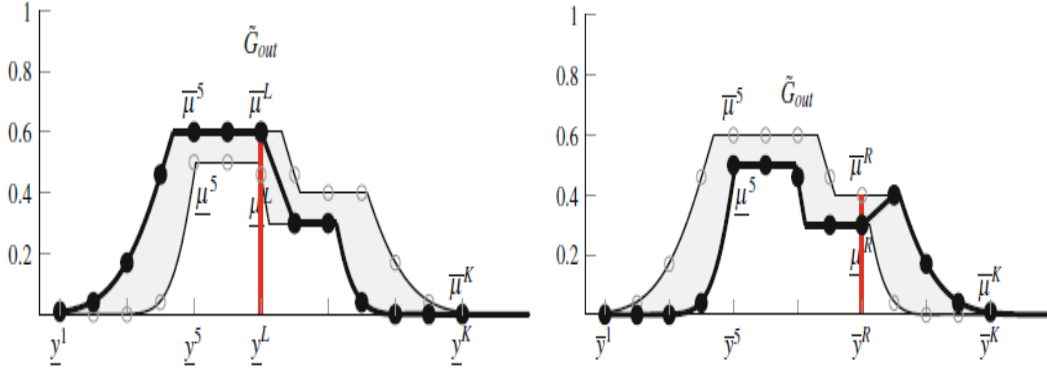


FIGURE 3.17 – Passage des bornes supérieures aux bornes inférieures et des bornes inférieures aux bornes supérieures..[42]

$$y_l(k) = \frac{\sum_{i=1}^k \bar{y}^i \bar{\mu}^i \bar{G}_{out} + \sum_{i=k+1}^k y^i \mu^i \bar{G}_{out}}{\sum_{i=1}^k \bar{\mu}^i \bar{G}_{out} + \sum_{i=k+1}^k \mu^i \bar{G}_{out}}$$

$$y_r(k) = \frac{\sum_{i=1}^k \bar{y}^i \mu^i \bar{G}_{out} + \sum_{i=k+1}^k \bar{y}^i \bar{\mu}^i \bar{G}_{out}}{\sum_{i=1}^k \mu^i \bar{G}_{out} + \sum_{i=k+1}^k \bar{\mu}^i \bar{G}_{out}} \quad (3.14)$$

Où  $k$  est un entier dans l'intervalle  $[1, K - 1]$  et  $K$  représente le nombre de points de discrétisation. Ensuite, les limites d'intervalle optimales peuvent être obtenues par  $y_l$  et  $y_r$ , comme suit :[42]

$$y_l = \min_{k \in [1, M - 1]} y_l(k) \equiv y(L) \equiv \frac{\sum_{i=1}^L \bar{y}^i \bar{\mu}^i \bar{G}_{out} + \sum_{i=L+1}^k y^i \mu^i \bar{G}_{out}}{\sum_{i=1}^L \bar{\mu}^i \bar{G}_{out} + \sum_{i=L+1}^k \mu^i \bar{G}_{out}}$$

$$y_r = \max_{k \in [1, M - 1]} y_r(k) \equiv y(R) \equiv \frac{\sum_{i=1}^R \bar{y}^i \mu^i \bar{G}_{out} + \sum_{i=R+1}^k \bar{y}^i \bar{\mu}^i \bar{G}_{out}}{\sum_{i=1}^R \mu^i \bar{G}_{out} + \sum_{i=R+1}^k \bar{\mu}^i \bar{G}_{out}} \quad (3.15)$$

Où  $L$  et  $R$  sont des points de commutation satisfaisants :

$$y^L \leq y_l < y^{L+1}$$

$$y^R \leq y_r < y^{R+1}$$

### 3.6.4 Défuzzification de type 2 -FL :

Cette étape consiste à réaliser l'opération inverse de la fuzzification, après la transformation un ensemble flou type-2 à un ensemble flou type-1 par TR il donne deux point finale, donc la défuzzification de ces point c'est la moyen  $Y$  comme équation suivant :

$$Y = \frac{Y_l + Y_r}{2} \quad [17] \quad (3.16)$$

### 3.7 Conclusion :

La logique floue est une approche au raisonnement humain. Dans ce chapitre on a vu les étapes et les méthodes de type-1 et type-2 de la logique floue pour la réalisation d'un système basé sur la logique floue. Les fonctions d'appartenances sont le cœur de la logique floue, et ces eux qui mettent la différence entre la logique classique (booléen) et la logique floue. Et aussi entres les types de la logique floue. L'inférence est l'expérience humaine détermine la manière de raisonnement du système. Et l'agrégation c'est le système prend le rôle de l'être humain et commence à raisonner et de défuzzifier en suite pour donner des résultats.

# Simulation D'une méthode objective pour l'évaluation de la qualité audiovisuelles par la logique floue

## 4.1 Introduction

Dans ce chapitre, nous allons présenter le déroulement du travail réalisé sur l'évaluation de la qualité audiovisuelle en utilisant la logique floue, et nous discutons les résultats obtenus. Nous avons proposé trois modèles pour faire l'évaluation objective base sur la logique floue : le premier modèle est un modèle globale qui génère la qualité audiovisuelle globalement sans passer par l'évaluation des deux qualités visuelle et auditive, par contre les deux autre modèles que nous avons créés sont des modèles objectives basés sur les qualités auditive et visuelle en utilisant la logique floue qui nous a donné des bons résultats par rapport le premier modèle les trois modèles sont appliqués sur la même base donnée INRS (Institut National de la Recherche Scientifique) réalisé par Edip Demirbilek [43]. A noter que le système flou proposé a été simulé sous Matlab en utilisant Fuzzy Logic Toolbox.

## 4.2 Pourquoi la logique floue (FL)

On utilise la logique floue pour évaluer la qualité audiovisuelle (QAV) car la logique floue est basée sur le langage naturel, la base de FL est la base de la communication humaine. Cette observation sous-tend de nombreuses autres déclarations sur la FL. Parce que la FL construite sur les structures de description qualitatives utilisées dans le langage courant et la FL est facile à utiliser et aussi la logique floue peut modéliser des fonctions non linéaires de complexité arbitraire. Vous pouvez créer un système flou pour correspondre à n'importe quel ensemble de données d'entrée-sortie. Ce processus est rendu particulièrement facile par des techniques adaptatives disponibles dans le logiciel Fuzzy Logic Toolbox[44]. Nous avons proposé trois modèles de QAV dont le but est d'attendre les meilleures performances de QAV.

## 4.3 Description des modèles proposés

Le premier modèle est un modèle globale qui génère la qualité audiovisuelle globalement sans passer par l'évaluation des deux qualités visuelle et auditive, vous avez utilisé un seul système de logique floue (FIS) selon le schéma de figure 4.1 :

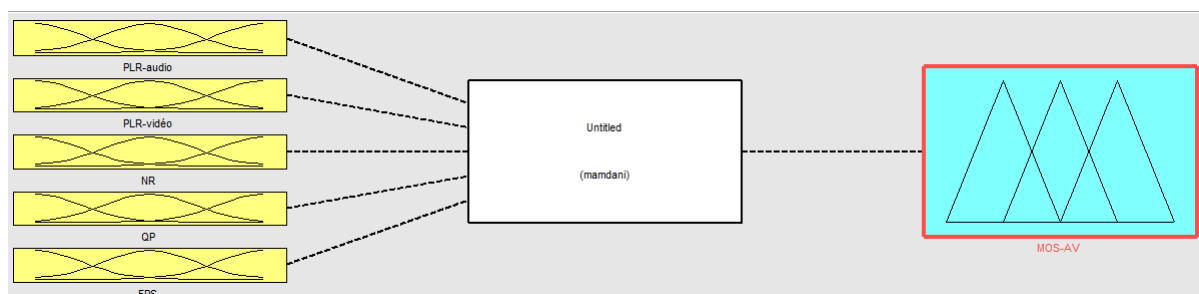


FIGURE 4.1 – Premier modèle FL pour évaluer la QAV.

Pour les deux autres modèles, nous avons utilisé L'audio et vidéo comme des paramètres d'entrés mais on a connu que audiovisuelle concaténé de deux flux : audio et vidéo donc il faut évaluer la qualité audio (QA) et qualité vidéo (QV) par FL premièrement et après nous utilisons les résultats de leur évaluation pour évaluer la qualité audiovisuelle par FL.

Pour le deuxième modèle comme la figure suivante :

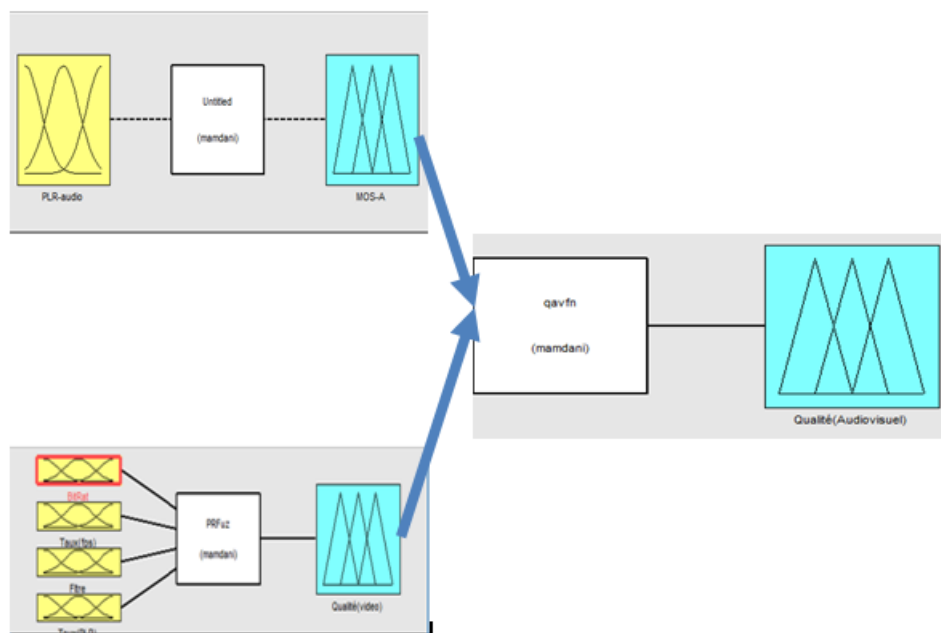


FIGURE 4.2 – deuxième modèle FL pour évaluer la QAV.

Pour troisième modèle, selon le schéma suivant :



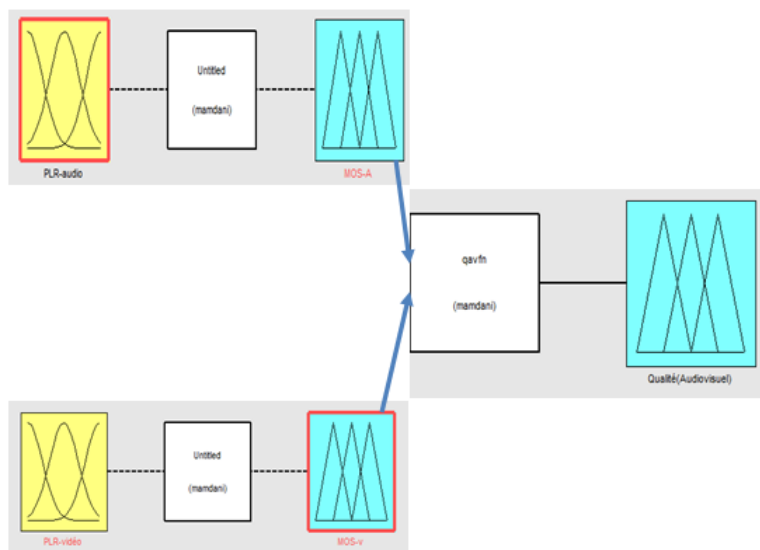


FIGURE 4.3 – troisième modèle FL pour évaluer la QAV.

## 4.4 La base de données audiovisuelle utilisée

La base donnée audiovisuelle de l'INRS[43], est composé de 160 configurations uniques pour le contenu audiovisuel, comprenant divers paramètres de compression du média et de distorsion du réseau, tels que la fréquence d'images, la quantification et les paramètres de réduction du bruit, ainsi que le taux de perte de paquets. Les valeurs de plage de paramètres de compression et de distorsion réseau sont sélectionnées pour correspondre aux cas d'utilisation de communications en temps réel. Le codec vidéo H.264 avec une résolution de 720p et le codec audio AMR-WB sont utilisés pour coder les flux vidéo et audio. 30 observateurs ont évalué la qualité audiovisuelle globale sur l'échelle de qualité ACR (Absolute Catégorie Rating) sur 5 niveaux dans un environnement contrôlé. L'ensemble de données comprend les valeurs MOS, les taux de perte de paquets mesurés au niveau du flux de bits pour les flux audio et vidéo.

Le tableau suivant présent les cinq niveaux des facteurs existe dans la base donnée INRS qui influent sur QA, QV :

- - - -	Video	Audio
FPS	10 , 15 ,20 ,25	Mono ,16 kHz , 24 kbps
QP	23 , 27 , 31 ,35	Mono ,16 kHz , 24 kbps
NR	0 , 999	Mono ,16 kHz , 24 kbps
PLR %	0 , 0.1 , 0.5 , 1.5	0 , 0.1 , 0.5 , 1.5

TABLE 4.1 – Paramètres de compression des médias et dégradation du réseau. [43]

A partir des facteurs existants dans la base INRS, nous avons choisi, frame par seconde (FPS : frame rate), paramètre de quantification (QP : bitrate), réduction du bruit (NR : noise réduction filtre analyse), taux de perte de paquets (PLR : packet loss rate) comme des paramètres entrés parce qu'ils jouent le rôle le plus important dans l'évaluation de la qualité[45].

## 4.5 Description des paramètres d'entrés et de sortie du système flou pour les trois modèles

On observe qu'un PLR supérieur à 0,5% réduit considérablement la qualité [2=3]. Et sélectionné FPS [10; 15; 20; 25], QP [23; 27; 30; 35] et NR [0 et 999], on observe si le FPS supérieur à 10 ou QP supérieur à 23 réduit considérablement la qualité.[46], on utilise Fuzzy Logic Toolbox version : 2.2.13 qui disponible sur le logiciel MATLAB Version 7.12.0.635 (R2011a). Vous pouvez créer ces systèmes à l'aide d'outils graphiques ou de fonctions de ligne de commande, la figure suivant présente comment évalué la qualité audio par FL :

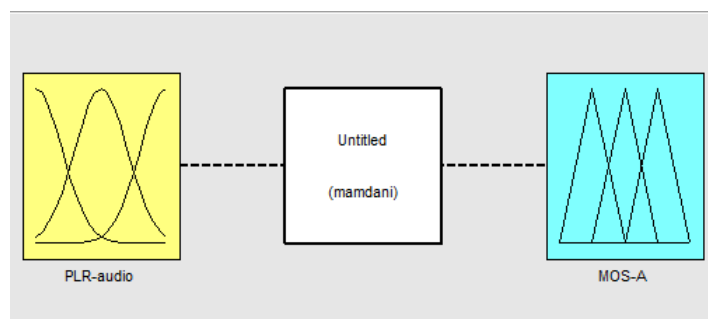


FIGURE 4.4 – Évalué la qualité audio par FL

A partir des paramètres existe de la base INRS. On choisit, PLR-audio, comme de paramètre entré pour notre système d'évaluation de la QA. Et pour la sortie on a pris MOS-audio . Et on précise trois variables linguistiques pour paramètre entré PLR-audio :[faible, moyen, élevé], et cinq pour paramètre sortie [mouvais, médiocre, satisfaisant, bien, excellent].

Avant de coder les fonctions appartenances, on a défini les meilleurs intervalles pour les paramètres entrés a partir de [47], et le graphe illustre dans la figure 4.6 qui est créé à partir des facteurs choisis dans la base donnée INRS :

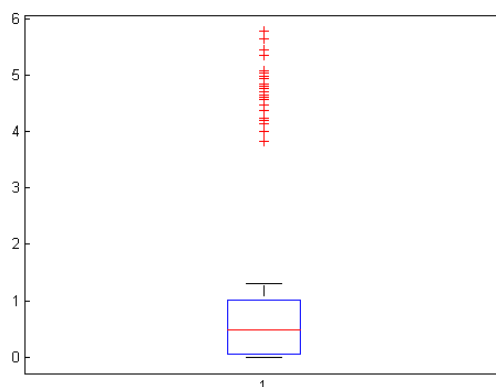


FIGURE 4.5 – Graphe pour intervalle de PLR- audio

On observe que le meilleur intervalle pour input PLR- audio est entre 0 - 1.5 % et que entre 0 - 0.5% c'est les note faible, la ligne rouge dans intervalle c'est la moyenne de l'intervalle égale 0.5%, supérieur a 0.5% les

note élevé, pour les note dans intervalle 4-6 c'est des note Hors intervalle.

Pour crier les Fonctions d'appartenances, on a utilisé la fonction trapézoïdale car il est simple et plus flexible et aussi permit les trois fonctions la plus utilise, La figure 4.6 illustre la fonction d'appartenance de « PLR-audio » et « MOS-audio ».

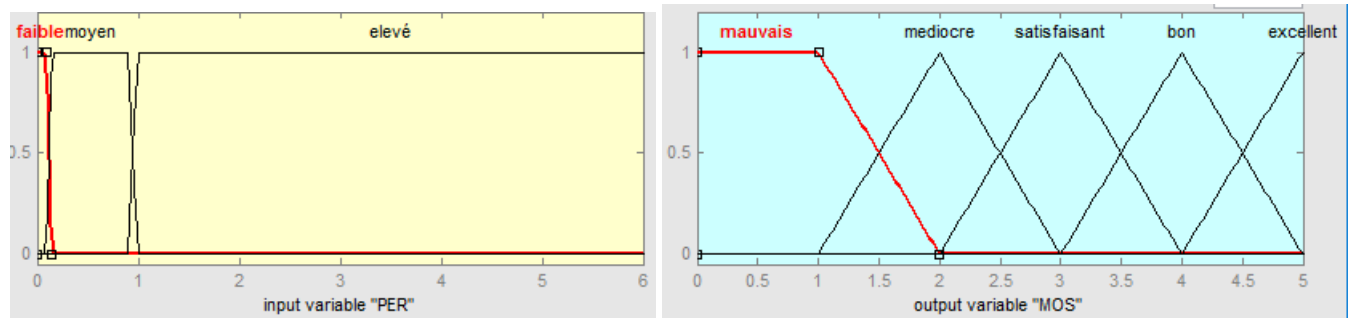


FIGURE 4.6 – Fonction d'appartenance de PLR-audio.et Fonction d'appartenance de MOS-audio.

Pour les règles d'inférence, on a fait entrée les règles floues qui relié les sous-ensembles d'entrées et sortie. Cette étapes est dans le but de déterminer les relations entre l'ensemble d'entré et la sortie en se basant sur ce qu'on a obtenu comme résultats en pratique.

Les opérateurs utilisés dans ce cas sont les opérateurs de type « Mamdani » ou s'appelle aussi « min-max » car il a un système d'inférence comme le comportement humain et peut donc être utilisé pour modéliser des systèmes non linéaires très compliqués[48]. Donc le moteur d'inférence sera comme l'exemple suivant :

1. If(PLR-audio is faible) then (MOS-audio is excellent).
2. If (PLR-audio is Moyne) then (MOS-audio is satisfaisant).
3. If (PLR-audio is élevé) then (MOS-audio is mauvais).

-En fin la Défuzzification, Dans cette étape, le système flou utilise la méthode du centre de gravité car c'est le plus utilisé dans le système FL pour la défuzzification, plus simple et flexible. Cette méthode a été expliquée dans le chapitre précédent. La figure 4.8 montre la fenêtre des règles sous Matlab où on peut choisir nos entrés pour que le système nous prédit des résultats.

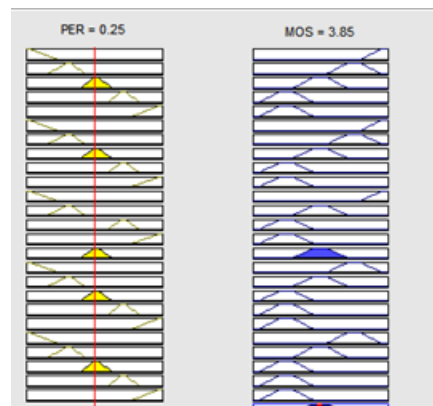


FIGURE 4.7 – Fenêtre qui montre un exemple des étapes de la défuzzification pour l'évaluation de QA sur FL.

Comme l'évaluation de la qualité audio, on a évalué la qualité vidéo la seul différence sur les paramètres entrés, on utilise les facteurs influents sur la qualité vidéo qui sont QP, FPS, NR,PLR-vidéo.

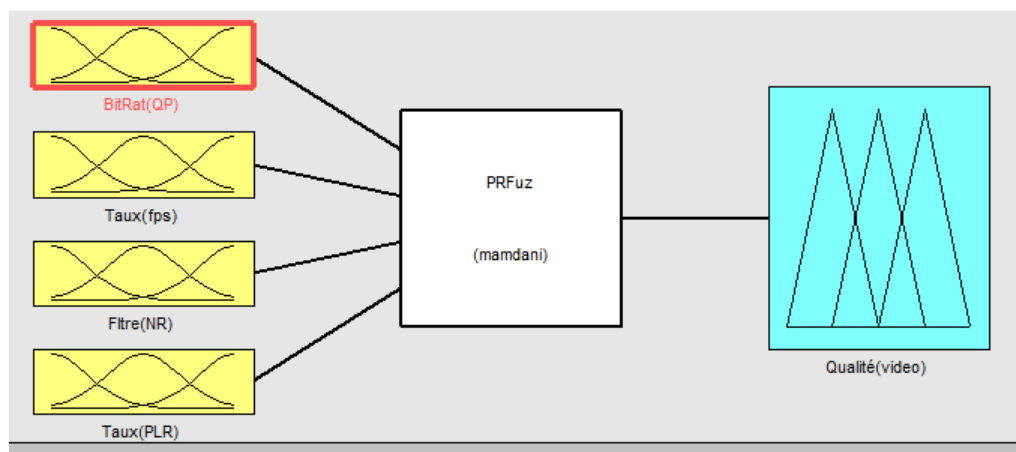


FIGURE 4.8 – Utilisation de la FL pour évaluation de la QV.

On précise trois variables linguistiques pour les paramètres entrés faible, moyen, élevé, et cinq pour paramètre sortie [ **mouvais, médiocre,satisfaisant,bien, excellent** ]. à partir des paramètres utilisés dans la base INRS, on choisit, **QP** , **FPS** , **NR** ,**PLR-vidéo** , comme des entrés pour notre système. Et pour la sortie on a pris MOS-vidéo.

Comme audio on a définit les meilleurs intervalles pour les paramètres entrés, selon les graphes illustrés dans la figure 4.10 et selon[47], on observé que le meilleur intervalle pour QP c'est 23 - 35 kbps, FPS c'est 10 - 25 fps, PLR-vidéo c'est 0 - 1.%, et pour NR c'est 0 - 999. Mais les notes qui ne sont pas dans ces intervalles sont des hors intervalles :

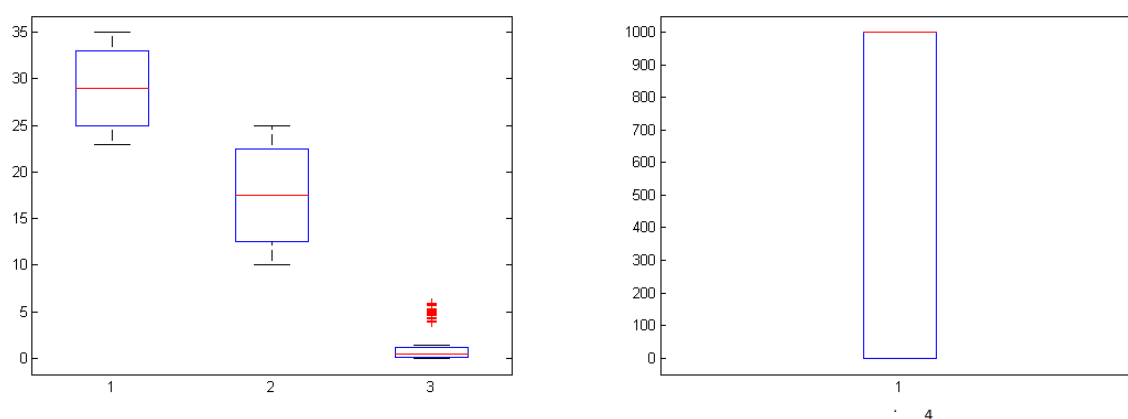


FIGURE 4.9 – Les meilleurs intervalles pour les inputs : 1=> QP, 2 => FPS, 3 => PLR-vidéo, 4 => NR de vidéo successivement.

On utilise la fonction trapézoïdale pour crier les Fonctions d'appartenances les deux figures suivants présentent les fonctions d'appartenances pour les entrées et les sorties :

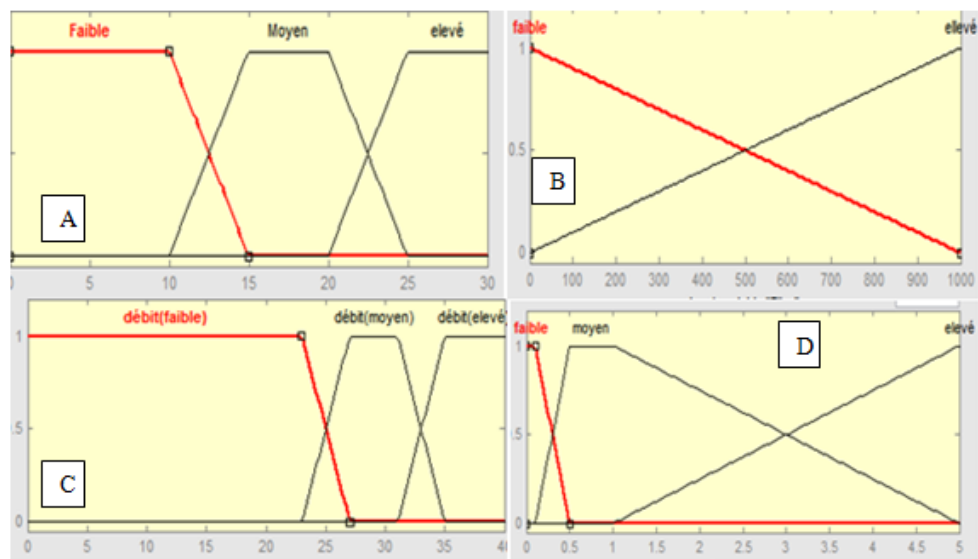


FIGURE 4.10 – les fonctions d'appartenances A=> FPS, B=>NR, C => QP, D => PLR-vidéo successivement

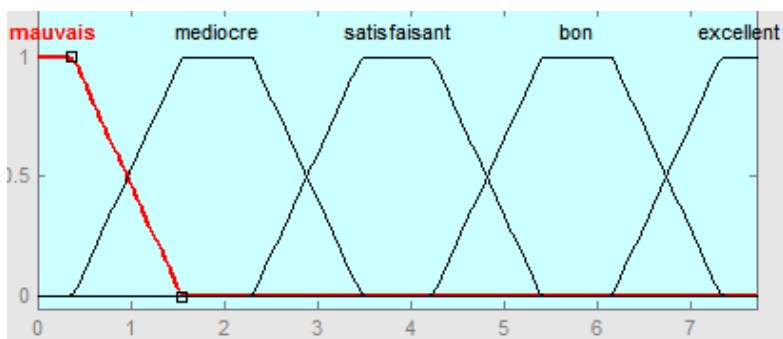


FIGURE 4.11 – fonction d'appartenance de MOS-vidéo.

Pour les règles d'inférence La figure suivante expliqué les règles d'inférence qui relient les entrés avec les sorties.

3. If (BitRat(QP) is not débit(faible)) and (Taux(fps) is Faible) and (Taux(PLR) is élevé) then (Qualité(video) is mauvais) (1)
4. If (BitRat(QP) is débit(faible)) and (Taux(fps) is Moyen) and (Taux(PLR) is faible) then (Qualité(video) is satisfaisant) (1)
5. If (BitRat(QP) is débit(faible)) and (Taux(fps) is Moyen) and (Taux(PLR) is moyen) then (Qualité(video) is médiocre) (1)
6. If (BitRat(QP) is débit(faible)) and (Taux(fps) is Moyen) and (Taux(PLR) is élevé) then (Qualité(video) is mauvais) (1)
7. If (BitRat(QP) is débit(elevé)) and (Taux(fps) is élevé) and (Taux(PLR) is faible) then (Qualité(video) is satisfaisant) (1)
8. If (BitRat(QP) is débit(elevé)) and (Taux(fps) is élevé) and (Taux(PLR) is moyen) then (Qualité(video) is médiocre) (1)
9. If (BitRat(QP) is débit(elevé)) and (Taux(fps) is élevé) and (Taux(PLR) is élevé) then (Qualité(video) is mauvais) (1)
10. If (BitRat(QP) is débit(faible)) and (Taux(fps) is Faible) and (Taux(PLR) is faible) then (Qualité(video) is bon) (1)
11. If (BitRat(QP) is débit(faible)) and (Taux(fps) is Faible) and (Taux(PLR) is moyen) then (Qualité(video) is médiocre) (1)
12. If (BitRat(QP) is débit(faible)) and (Taux(fps) is Faible) and (Taux(PLR) is élevé) then (Qualité(video) is mauvais) (1)
13. If (BitRat(QP) is débit(moyen)) and (Taux(fps) is Moyen) and (Taux(PLR) is faible) then (Qualité(video) is bon) (1)
14. If (BitRat(QP) is débit(moyen)) and (Taux(fps) is Moyen) and (Taux(PLR) is moyen) then (Qualité(video) is satisfaisant) (1)
15. If (BitRat(QP) is débit(moyen)) and (Taux(fps) is Moyen) and (Taux(PLR) is élevé) then (Qualité(video) is médiocre) (1)
16. If (BitRat(QP) is débit(elevé)) and (Taux(fps) is Moyen) and (Taux(PLR) is not élevé) then (Qualité(video) is satisfaisant) (1)
17. If (BitRat(QP) is débit(elevé)) and (Taux(fps) is Moyen) and (Taux(PLR) is élevé) then (Qualité(video) is médiocre) (1)
18. If (Taux(fps) is élevé) and (Taux(PLR) is élevé) then (Qualité(video) is mauvais) (1)

FIGURE 4.12 – Fenêtre qui montre un exemple des règles d’inférences pour évaluer QV sur FL.

Nous avons atteint la dernière étape « défuzzification », on peut choisir nos entrées pour que le système nous prédit des résultats. La figure suivante illustre ça :

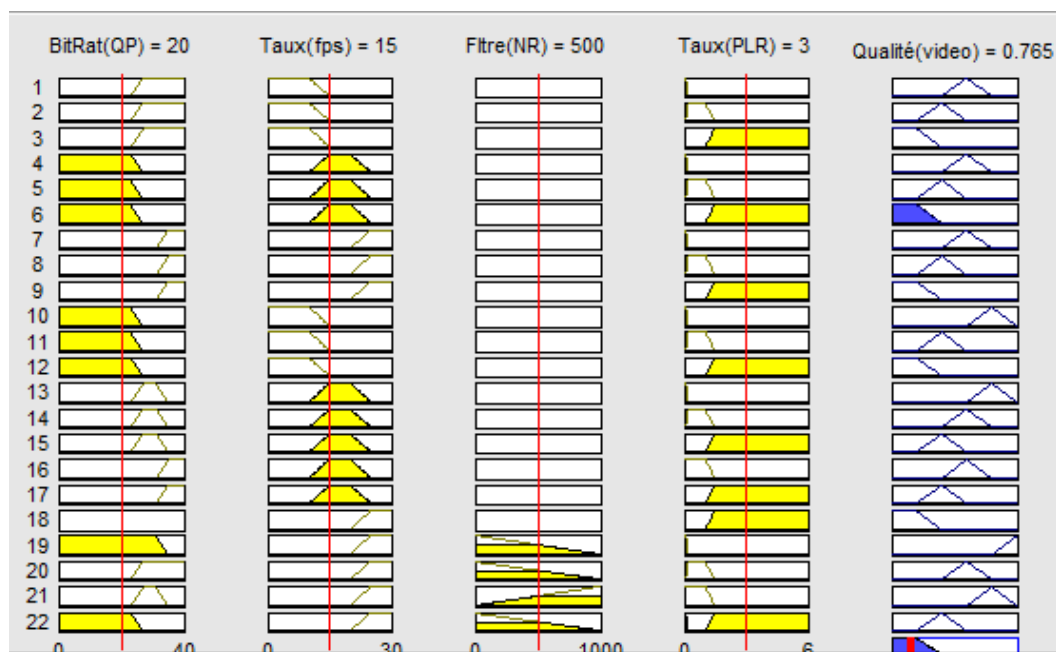


FIGURE 4.13 – Fenêtre qui montre un exemple des étapes de la défuzzification pour l’évaluation de QV sur FL.

Après l’évaluation de la QA et la QV, on utilise les résultats de chaque évaluation de QA et QV des paramètres entrés pour évaluer la QAV comme on a vu dans la figure 4.2 précédente de deuxième modèle. Donc pour les variables linguistiques on précise cinq variables mauvais, médiocre, satisfaisant, bien, excellent et aussi utilise la fonction trapézoïdale pour créer les fonctions d’appartenance comme présente dans les figure suivantes :

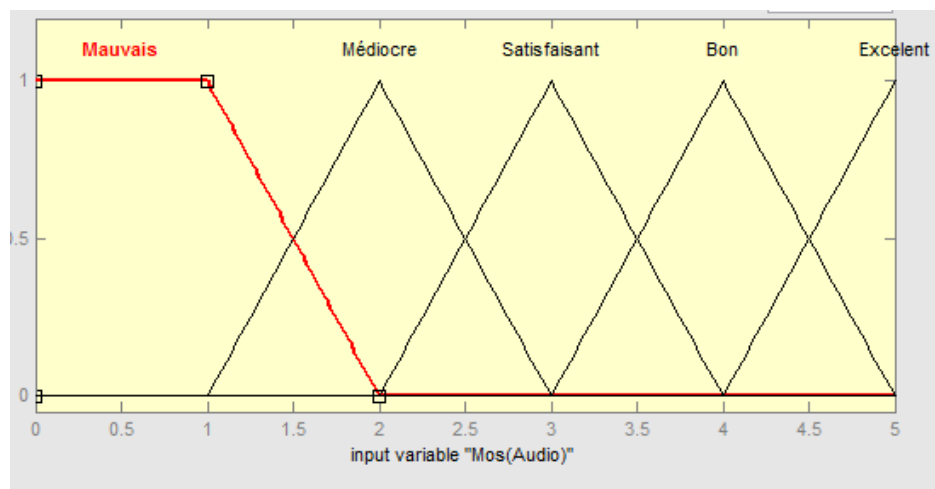


FIGURE 4.14 – fonction d'appartenance de MOS-audio.

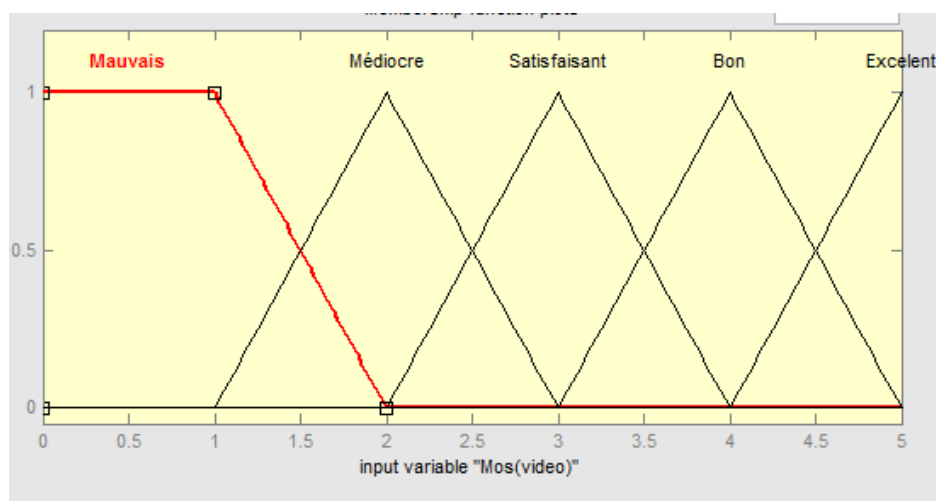


FIGURE 4.15 – fonction d'appartenance de MOS-vidéo.

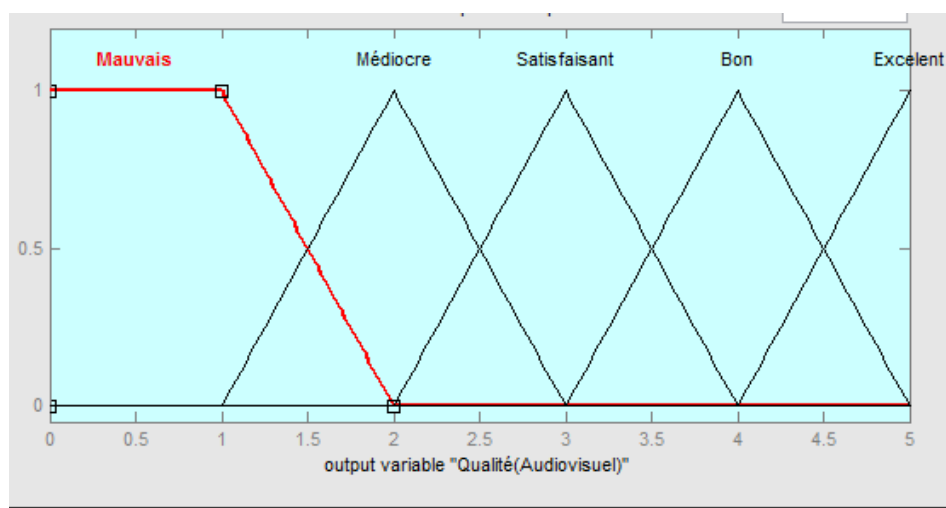


FIGURE 4.16 – fonction d'appartenance de MOS-audiovisuelle.

-Pour les règles d'inférences la figure suivant illustré ça :

```

1. If (Mos(video) is Mauvais) and (Mos(Audio) is Mauvais) then (Qualité(Audiovisuel) is Mauvais) (1)
2. If (Mos(video) is Mauvais) and (Mos(Audio) is Médiocre) then (Qualité(Audiovisuel) is Mauvais) (1)
3. If (Mos(video) is Mauvais) and (Mos(Audio) is Satisfaisant) then (Qualité(Audiovisuel) is Mauvais) (1)
4. If (Mos(video) is Mauvais) and (Mos(Audio) is Bon) then (Qualité(Audiovisuel) is Mauvais) (1)
5. If (Mos(video) is Mauvais) and (Mos(Audio) is Excelent) then (Qualité(Audiovisuel) is Mauvais) (1)
6. If (Mos(video) is Médiocre) and (Mos(Audio) is Mauvais) then (Qualité(Audiovisuel) is Médiocre) (1)
7. If (Mos(video) is Médiocre) and (Mos(Audio) is Médiocre) then (Qualité(Audiovisuel) is Médiocre) (1)
8. If (Mos(video) is Médiocre) and (Mos(Audio) is Satisfaisant) then (Qualité(Audiovisuel) is Médiocre) (1)
9. If (Mos(video) is Médiocre) and (Mos(Audio) is Bon) then (Qualité(Audiovisuel) is Médiocre) (1)
10. If (Mos(video) is Médiocre) and (Mos(Audio) is Excelent) then (Qualité(Audiovisuel) is Médiocre) (1)
11. If (Mos(video) is Satisfaisant) and (Mos(Audio) is Mauvais) then (Qualité(Audiovisuel) is Satisfaisant) (1)
12. If (Mos(video) is Satisfaisant) and (Mos(Audio) is Médiocre) then (Qualité(Audiovisuel) is Satisfaisant) (1)
13. If (Mos(video) is Satisfaisant) and (Mos(Audio) is Satisfaisant) then (Qualité(Audiovisuel) is Satisfaisant) (1)
14. If (Mos(video) is Satisfaisant) and (Mos(Audio) is Bon) then (Qualité(Audiovisuel) is Satisfaisant) (1)
15. If (Mos(video) is Satisfaisant) and (Mos(Audio) is Excelent) then (Qualité(Audiovisuel) is Satisfaisant) (1)
16. If (Mos(video) is Bon) and (Mos(Audio) is Mauvais) then (Qualité(Audiovisuel) is Bon) (1)

```

FIGURE 4.17 – Fenêtre qui montre un exemple des règles d'inférences pour évaluer QAV sur FL.

Pour l'étape de défuzzification elle fonctionne automatiquement et donne directement le résultat de mos audiovisuel.

## 4.6 Analyse et interprétation des résultats

Dans ce qui suit, nous détaillons les résultats obtenues de nos modèles proposés pour l'évaluation objective de la qualité audiovisuelle.

Donc l'objectif de notre travail premièrement est d'arriver à obtenir des résultats similaires aux résultats subjectives de la base donnée INRS en le système flou et comme deuxième objective est de généraliser cette technique pour prédire d'autres résultats qui sont irréalizable au laboratoire ce qui montre la robustesse du système flou proposé.

A noter que la performance d'un modèle d'évaluation objective de QAV est évaluée via trois métriques statistiques, utilisées pour informer de la précision du modèle via l'erreur quadratique moyenne (RMSE), de sa consistance via le rapport de valeurs aberrantes (outlier ratio) et de sa linéarité/monotonie via le coefficient de corrélation de Pearson (voir le chapitre 2).

Pour les trois modèles de QAV développés, nous calculons la corrélation de Pearson entre Mos-prédictive de chaque modèle et Mos-subjective de la base donnée INRS et aussi nous calculons le RMSE et RO (outlire), c'est calculs aide nous pour que nous puissions découvrir notre meilleur modèle et nous comparons aussi nos modèles avec les modèles objective de Edip Demirbilek qui utilise les réseaux de neurone et l'apprentissage [45]. Le tableau 4.2 montre les résultats obtenus.



—	Corrélation- de Pearson	rmse	RO
Modèle 1	0.7971	0.6671	0.4500
Modèle 2	0.8008	0.7012	0.4000
Modèle 3	0.8504	0.6519	0.4375

TABLE 4.2 – les résultats de performances des modèles proposés

A partir du tableau 4., on remarque que les résultats de performance des trois modèles proposés sont meilleurs, et aussi le troisième modèles qui utilise que les PLR-audio et PLR-vidéo comme paramètres est le meilleur puisque il a une plus grande valeur pour Corrélation de Pearson et des plus petites valeurs pour RMSE et sauf RO de la deuxième modèle est optimale.

D'après les figures 4.18, 4.19 ,4.20 et 4.21, 4.22 on observe à partir des résultats de troisième modèle que les facteurs PLR-audio et PLR-vidéo sont les facteurs le plus influent pour l'estimation de la qualité audiovisuelle. Dans la figure 4.20, on observe que le PLR audio et vidéo est très influents sur la QAV et on voit dans les graphes que si le PLR audio et vidéo inférieur à 0.5%, les Mos-prédictive est élevé et le contraire si PLR audio et vidéo supérieur à 0.5%. Pour les facteurs QP, NR, et FPS, on observé que QP, NR, FPS est moins influents sur la QAV, il y a des valeurs aberrantes entre le 0.5% et 1 ne pas influent clairement cela retour a l'influence léger quand les paramètre se rencontrent .

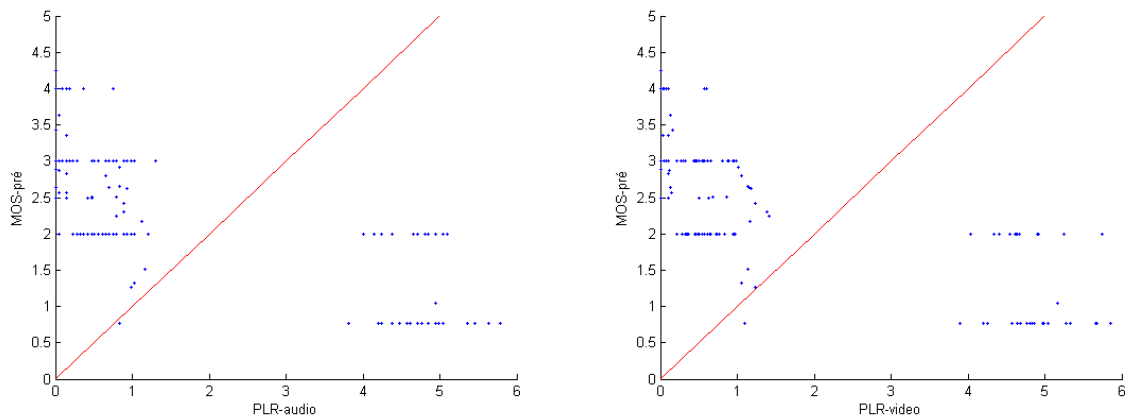


FIGURE 4.18 – graphe définit comment PLR audio et PLR vidéo influence sur la QAV

La différence de performances est beaucoup plus facile à réaliser dans une interprétation graphique. Figures suivant montre le MOS-subjective par rapport au MOS-prévue pour les trois modèles proposés. Les chiffres montrent également que le modèle trois permet une estimation plus précise :

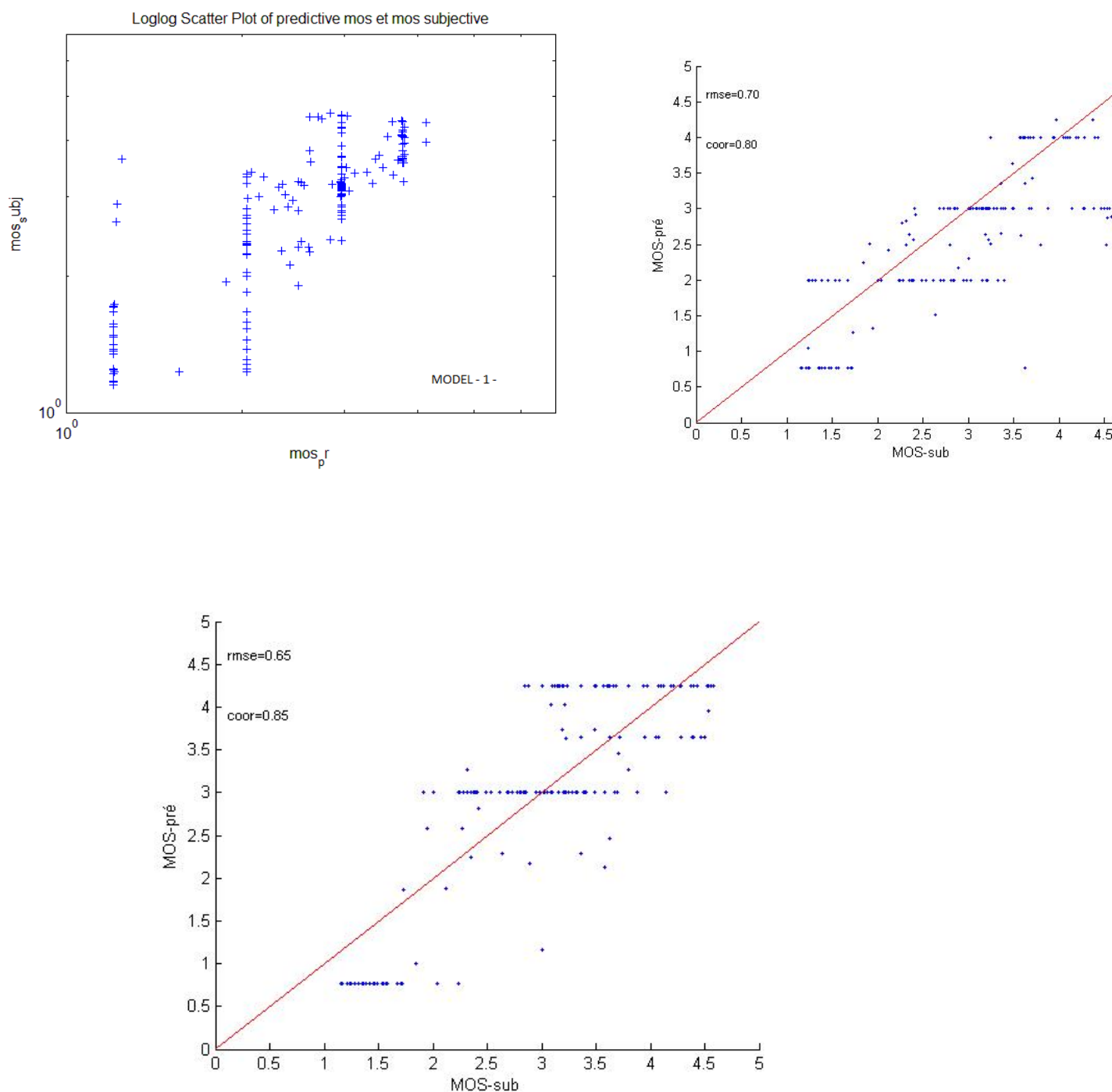


FIGURE 4.19 – les graphes de MOS-model VS MOS-prédiction pour trois modèles

On a calculé la Corrélacion de Personne entre les autres facteurs influents sur la QA, QV, QAV qui est QP, FPS, NR et mos prédictive d'un modèle pour que nous puissions découvrir comment influent ces facteurs sur la QAV, les figures 4.20, 4.21, 4.22 démontrent ça :

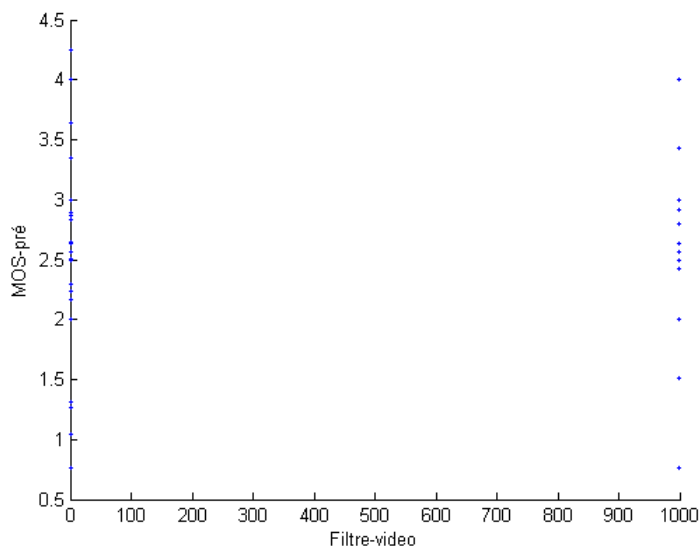


FIGURE 4.20 – graphe définit comment influencer NR sur la QAV.

-D'après la figure 4.20, on remarque que l'influence n'est pas stable c'est-à-dire l'estimation de la qualité indépendant du facteur NR..

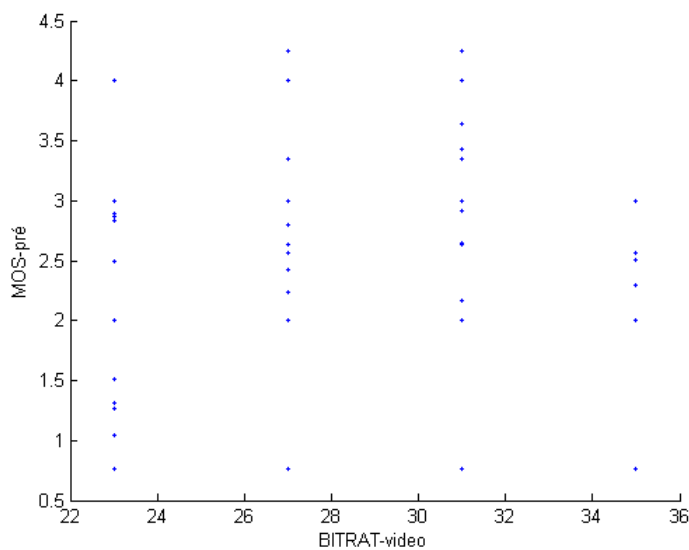


FIGURE 4.21 – graphe définit comment influencer QP sur la QAV

-Ici aussi l'influence n'est pas stable on a remarqué des lignes pour chaque valeur de QP malgré l'intervalle entre 26 à 32 kbps s'affiche des bons résultats mais il reste indécisive, on peut dire que l'estimation de la qualité indépendante du facteur QP.

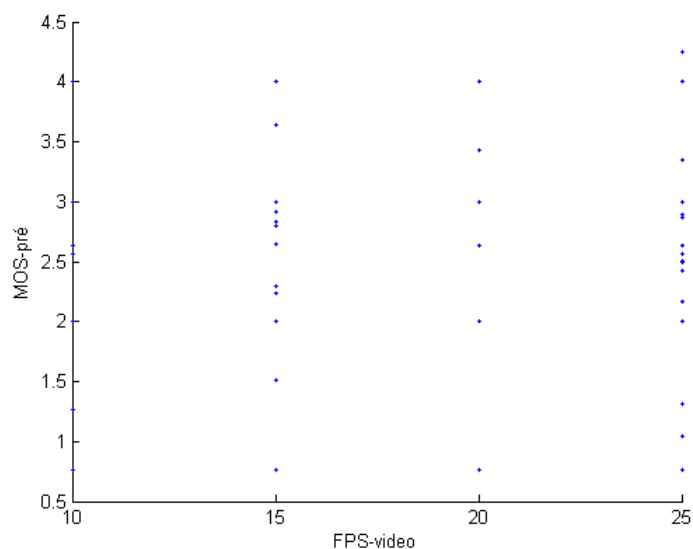


FIGURE 4.22 – graphe définit comment influencer FPS sur la QAV.

-Ici aussi l'influence n'est pas stable on a remarqué des lignes pour chaque valeur de FPS malgré la valeur 25. FPS s'affiche des bons résultats donc on peut dire que l'estimation de la qualité indépendante du facteur FPS .

On observé a partir des figures précédentes que le FPS, QP, NR sont moins influencés sur la QAV par apport à PLR audio et PLR vidéo car on voit un très petites changement dans les résultats de mos prédictive. Donc le PLR c'est le facteur très influencé sur la QAV.

## 4.7 conclusion

Dans ce chapitre montre l'efficacité de la logique floue pour évaluer la QAV, une comparaison avec les résultats des modèles proposés permet de classer les paramètres influent sur la QAV, et aussi de classe notre méthode parmi les meilleures méthodes d'évaluation objectives de la qualité audiovisuelle

# Conclusion

Dans le domaine informatique et plus exactement le domaine multimédia, l'audiovisuel a pris une grande place dans les temps moderne grâce à la fusion de la vidéo ainsi que l'audio.

Une poussée récente de services multimédias sur des réseaux filaires et sans fil a également déclenché une étude d'évaluation de la qualité perceptuelle. En particulier, il existe une forte demande pour des méthodes capables d'estimer et de quantifier la qualité du codage, de la transmission et du décodage (réception), des services, de l'expérience et de la satisfaction perçues par l'utilisateur final. Bien que l'évaluation de la qualité multimédia perceptuelle se soit avéré une tâche difficile, de nombreux efforts de recherche et de développement ont été consacrés à cette tâche et à ses applications, permettant ainsi des progrès significatifs dans ce domaine.

En fin, la logique floue a éprouvé l'efficacité pour d'estimer et de quantifier la qualité audiovisuelle. Aussi nous avons vu que la conception d'un système flou nécessite un nombre important des tests pratique. Avec ces différents donnée collectées des plusieurs tests et leur analyse permet, en particulier de définir les fonctions d'appartenance.

La puissance de la logique floue rendre possible la mise en place du système d'inférence dans les décisions sont plus proche du comportement humain que ne l'est la logique classique. De plus, les règles de la matrice des décisions sont exprimées en langage naturel, ce qui comporte de nombreux avantages, comme par exemple inclure des connaissances d'un expert non informaticien au cœur d'un système décisionnel ou encore modéliser plus aisément certains aspects du langage naturel.

En perspectives de travail :

- Dans le futur travail il est question de recherche sur les deux améliorations suivantes de notre méthode : (type-2-logique floue et système adaptatifs d'inférence neuro-floue (ANFIS)), on les compare avec notre méthode (l'approche de la fusion des deux modalités) pour choisir la meilleur pour l'évaluation de la qualité audiovisuel.

- Pour le moment notre résultat est acceptable parce qu'on les a appliqués sur une seule base de données INRS, et pour assurer la fiabilité de notre méthode, il faut encore s'élargir dans nos recherche. Parmi cet élargissement on va mettre l'application de notre méthode sur plusieurs base de donnée audiovisuel et là où on peut voir les bon et mauvais résultats et choisir la meilleur.

# Bibliographie

- [1] M.TAYACHI and Hamida MAKLOUF. *Techniques audiovisuelles*. Université Virtuelle de Tunis, (2006).
- [2] Stefan Winkler and Christof Faller. Audiovisual quality evaluation of low-bitrate video. In *Human Vision and Electronic Imaging X*, volume 5666, pages 139–148. International Society for Optics and Photonics, 2005.
- [3] KV Rop and NK Bett. Video conferencing and its application in distance learning. June 2012.
- [4] HOWSTUFFWORKS.COM CONTRIBUTORS. What is facetime. "<https://HowStuffWorks.comContributors/WhatisFaceTime>", 2018.
- [5] Jean and Francois Pillou. Introduction à la vidéo numérique. "<https://www.commentcamarche.net/contents/1493-introduction-a-la-video-numerique>", 19 mai 2015.
- [6] Thomas Wiegand, Gary J Sullivan, Gisle Bjontegaard, and Ajay Luthra. Overview of the h. 264/avc video coding standard. *IEEE Transactions on circuits and systems for video technology*, 13(7) :560–576, 2003.
- [7] Mohammed El-Amin Selmi and Mohammed Yala. *Synthèse et étude comparative sur les méthodes de compression vidéo*. PhD thesis, 07-03-2017.
- [8] L Chiariglione. Mpeg-4 project description. In *ISO/IEC JTC1/SC29/WG11N1177 Munich MPEG meeting*, 1996.
- [9] Julien Le Tanou and Médéric Blestel. Analysis of emerging video codecs : Coding tools, compression efficiency. *SMPTE Motion Imaging Journal*, 128(10) :14–24, 2019.
- [10] H.BENZENINE. Traitement du son. "<http://dSPACE.univ-tlemcen.dz>".
- [11] Solène LIMOUSIN. Mieux connaître les formats audio. "<https://www.supinfo.com/articles/single/2680-mieux-connaître-formats-audio>", 16/10/2016.
- [12] Zahid Akhtar and Tiago H Falk. Audio-visual multimedia quality assessment : A comprehensive survey. *IEEE Access*, 5 :21090–21117, 2017.
- [13] Donald A Norman and Stephen W Draper. *User centered system design; new perspectives on human-computer interaction*. L. Erlbaum Associates Inc., 1986.
- [14] Stephen R Gulliver and Gheorghita Ghinea. The perceptual and attentive impact of delay and jitter in multimedia delivery. *IEEE Transactions on Broadcasting*, 53(2) :449–458, 2007.
- [15] document ITU Rec.J.148. *Requirements for an Objective Perceptual Multimedia Quality Model*. ITU Telecommunication Standardization Sector, (2010).
- [16] Søren Bech and Nick Zacharov. *Perceptual audio evaluation : theory, method and application*. Wiley Online Library, 2006.
- [17] Lahouari Benyettou. Mise en œuvre d'un modèle de corrélation qos - qoe.
- [18] Ghislain Takam Tchendjou. *Contrôle des performances et conciliation d'erreurs dans les décodeurs d'image*. PhD thesis, 2018.

- [19] UIT T.Recommandation P.911. Méthodes d'évaluation subjective de la qualité audiovisuelle pour applications multimédias. "Http://www.itu.int/pub/T-REC," , 1998.
- [20] UIT T.Recommandation P.911. Méthodes d'évaluation subjective de la qualité audiovisuelle pour applications multimédias, corrigendum 1. "Http://www.itu.int/pub/T-REC," , 1999.
- [21] Zahid Akhtar and Tiago H Falk. Audio-visual multimedia quality assessment : A comprehensive survey. *IEEE Access*, 5 :21090–21117, 2017.
- [22] Adam Borowiak, Ulrich Reiter, and U Peter Svensson. Quality evaluation of long duration audiovisual content. In *2012 IEEE Consumer Communications and Networking Conference (CCNC)*, pages 337–341. IEEE, 2012.
- [23] Sebastian Möller and Alexander Raake. *Quality of experience : advanced concepts, applications and methods*. Springer, 2014.
- [24] Borhen Eddine Dakkar and Fella Hachouf. *Évaluation de la qualité perceptuelle des signaux multimédias*. PhD thesis, l'université frères Mentouri Constantine, 2017.
- [25] Masaki Togai and Stephen Chiu. A fuzzy logic chip and a fuzzy inference accelerator for real-time approximate reasoning. In *Proc. of 17th. International Symposium of Multiple Valued Logic*, pages 25–29, 1987.
- [26] Robert F Calusdian. Fuzzy reality. *IEEE Potentials*, 14(2) :20–22, 1995.
- [27] Franck Deroncourt. Introduction to fuzzy logic. *Massachusetts Institute of Technology*, 21, 2013.
- [28] Ghislain Takam Tchendjou, Rshdee Alhakim, and Emmanuel Simeu. Fuzzy logic modeling for objective image quality assessment. In *2016 Conference on Design and Architectures for Signal and Image Processing (DASIP)*, pages 98–105. IEEE, 2016.
- [29] Rémi Dubois. Application des nouvelles méthodes d'apprentissage à la détection précoce d'anomalies en électrocardiographie. *These de Doctorat de l'Université Pierre et Marie Curie*, 2004.
- [30] Fatima Zohra OUISSI. Classification des arythmies cardiaques par les arbres de décision flous. 06 décembre 2011.
- [31] Ramzi Deghnouche. *étude de l'influence d'une barrière isolante sur l'amorç age d'un intervalle pointe-plan*. PhD thesis, juin 2007.
- [32] Halim MERABTI. *Etude des systèmes flous à intervalle*. PhD thesis, Université de Constantine 1, 2008.
- [33] Franck Deroncourt. Introduction to fuzzy logic. *Massachusetts Institute of Technology*, 21, 2013.
- [34] Gonzalo Cabodevila. *Commande et identification floue*, 1993.
- [35] FAID Sabrina. *Commande par la logique floue de la machine synchrone a aimants permanents*. PhD thesis, Université Mohamed Boudiaf-M'Sila, 2005.
- [36] Houssam Eddine-Soufiane BOUZID-BENMERIEM et al. *Application de la technique de la logique floue pour la prédiction de l'amorçage des intervalles d'air pointes-plans*. PhD thesis.
- [37] Salim Boukaka. *Réalisation d'une bibliothèque de lois de commande adaptative our MSAP*. PhD thesis, Université du Québec à Trois-Rivières, 2015.
- [38] Lotfi Baghli. *Contribution à la commande de la machine asynchrone, utilisation de la logique floue, des réseaux de neurones et des algorithmes génétiques*. PhD thesis, 1999.
- [39] M. Kerouaz. *Commande vectorielle d'une machine synchrone à aimants permanents alimentée par onduleur de tension*. PhD thesis, 2000.
- [40] Ahmad Ibrahim. *Fuzzy logic for embedded systems applications*. Newnes, 2004.
- [41] Oscar Castillo, Patricia Melin, Janusz Kacprzyk, and Witold Pedrycz. Type-2 fuzzy logic : theory and applications. In *2007 IEEE International Conference on Granular Computing (GRC 2007)*, pages 145–145. IEEE, 2007.

- [42] Rômulo Antão. *Type-2 Fuzzy Logic : Uncertain Systems Modeling and Control*. Springer, 2017.
- [43] Edip Demirbilek and Jean-Charles Grégoire. Inrs audiovisual quality dataset. In *Proceedings of the 24th ACM international conference on Multimedia*, pages 167–171, 2016.
- [44] Inc MathWorks and Wei-cheng Wang. *Fuzzy Logic Toolbox : for Use with MATLAB : User's Guide*. Mathworks, Incorporated, 1998.
- [45] Edip Demirbilek and Jean-Charles Grégoire. Perceived audiovisual quality modelling based on decision trees, genetic programming and neural networks. *arXiv preprint arXiv :1801.05889*, 2017.
- [46] Edip Demirbilek and Jean-Charles Grégoire. Towards reduced reference parametric models for estimating audiovisual quality in multimedia services. In *2016 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2016.
- [47] Winarno Sugeng, Jazi Eko Istiyanto, Khabib Mustofa, and Ahmad Ashari. The impact of qos changes towards network performance. *International Journal of Computer Networks and Communications Security*, 3(2) :48–53, 2015.
- [48] Muhammet OZTURK, Mustafa Utku YILMAZ, Evren Ozgur, and Nilcan Akatas. Using fuzzy logic approach on evaporation modeling. *International Journal of Natural and Engineering Sciences*, 11(3) :04–06, 2017.
- [49] ANSI-Accredited committee T1 contribution. Report on extension of combined audio vidéo quality model. Technical report, Bellcore, 1993.
- [50] ANSI-Accredited committee T1 contribution. Report on extension of combined audio vidéo quality model. Technical report, Bellcore, 1993.
- [51] ANSI-Accredited committee T1 contribution. Report on extension of combined audio vidéo quality model. Technical report, Bellcore, 1995.
- [52] Margaret H Pinson, William Ingram, and Arthur Webster. Audiovisual quality components. *IEEE Signal Processing Magazine*, 28(6) :60–67, 2011.
- [53] Nobuhiko Kitawaki, Yusuke Arayama, and Takeshi Yamada. Multimedia opinion model based on media interaction of audio-visual communications. *MESAQIN2005*, pages 5–10, 2005.
- [54] Marie-Neige Garcia, Robert Schleicher, and Alexander Raake. Impairment-factor-based audiovisual quality model for iptv : influence of video resolution, degradation type, and content type. *EURASIP Journal on Image and Video Processing*, 2011(1) :629284, 2011.
- [55] Takanori Hayashi, Kazuhisa Yamagishi, Toshiko Tominaga, and Akira Takahashi. Multimedia quality integration function for videophone services. In *IEEE GLOBECOM 2007-IEEE Global Telecommunications Conference*, pages 2735–2739. IEEE, 2007.
- [56] ITU-T REC P931. Multimedia communication delay, 1998.
- [57] ITU-T REC J148. Requirements for an objective perceptual multimedia quality model, 2003.
- [58] ITU-T REC P.1201. Parametric non-intrusive assessment of audiovisual media streaming quality, 2012.
- [59] ITU-T REC G.1070. Opinion model for video-telephony applications, 2012.
- [60] ITU-T REC G.1071. Opinion model for network planning of video and audio streaming applications, 2015.
- [61] David S Hands. A basic multimedia quality model. *IEEE Transactions on multimedia*, 6(6) :806–816, 2004.
- [62] Benjamin Belmudez. *Audiovisual quality assessment and prediction for videotelephony*. Springer, 2014.
- [63] Junyong You, Ulrich Reiter, Miska M Hannuksela, Moncef Gabbouj, and Andrew Perkis. Perceptual-based quality assessment for audio–visual services : A survey. *Signal Processing : Image Communication*, 25(7) :482–501, 2010.



- [64] Helard Becerra Martinez and Mylène C Farias. Full-reference audio-visual video quality metric. *Journal of Electronic Imaging*, 23(6) :061108, 2014.
- [65] Yoshua Bengio, Ian J Goodfellow, and Aaron Courville. Deep learning an mit press book in preparation. *Draft chapters available at*, 2015.
- [66] M Sajid Mushtaq, Brice Augustin, and Abdelhamid Mellouk. Empirical study based on machine learning approach to assess the qos/qoe correlation. In *2012 17th European Conference on Networks and Optical Communications*, pages 1–7. IEEE, 2012.
- [67] Haiqing Du, Chang Guo, Yixi Liu, and Yong Liu. Research on relationship between qoe and qos based on bp neural network. In *2009 IEEE International Conference on Network Infrastructure and Digital Content*, pages 312–315. IEEE, 2009.
- [68] Riccardo Poli, William B Langdon, Nicholas F McPhee, and John R Koza. *A field guide to genetic programming*. Lulu. com, 2008.
- [69] Michael Schmidt and Hod Lipson. Symbolic regression of implicit equations. In *Genetic Programming Theory and Practice VII*, pages 73–85. Springer, 2010.
- [70] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2014.
- [71] Leo Breiman. Random forests. *Machine learning*, 45(1) :5–32, 2001.
- [72] Andy Liaw, Matthew Wiener, et al. Classification and regression by randomforest. *R news*, 2(3) :18–22, 2002.
- [73] Pedro Domingos. A few useful things to know about machine learning. *Communications of the ACM*, 55(10) :78–87, 2012.
- [74] ITU-T REC P.1401. Methods, metrics and procedures for statistical evaluation, qualification and comparison of objective quality prediction models, 07/2012.
- [75] Marie-Neige Garcia. *Parametric packet-based audiovisual quality model for IPTV services*. Springer, 2014.