



SORBONNE UNIVERSITÉ
Campus Sciences et Ingénierie, Jussieu

COMPTE RENDU TME 1 ITQOS

KITOKO DAVID ET DIZDAREVIC ADRIAN

ITQOS

E-Model

1 – Équation principale du R-factor et rôle des termes

L'équation générale de l'E-Model est donnée par :

$$R = R_0 - I_s - I_d - I_{e,\text{eff}} + A$$

- R_0 : qualité de base du signal sans perturbations (ex : bruit de fond).
- I_s : dégradations dues aux conditions d'écoute et d'émission indépendamment du délai (ex : volume).
- I_d : dégradations dues au délai (ex : latence).
- $I_{e,\text{eff}}$: dégradations dues aux codecs et à la perte de paquets.
- A : facteur d'avantage (compensation contextuelle, ex : mobilité).

2 – Prise en compte du taux de perte et du délai e2e

Perte de paquets audio

Oui — elle est prise en compte dans le terme $I_{e,\text{eff}}$. La recommandation (Section 7.5, équation (7-29)) donne la formule :

$$I_{e,\text{eff}} = I_e + (95 - I_e) \cdot \frac{P_{pl}}{P_{pl}/\text{BurstR} + B_{pl}}$$

où :

- I_e : impairment intrinsèque du codec à 0% de pertes,
- P_{pl} : probabilité de perte de paquets,
- BurstR : burst ratio (mesure du regroupement des pertes),
- B_{pl} : facteur de robustesse du codec face aux pertes.

Délai réseau end-to-end

Oui — il est pris en compte dans le terme I_d , plus précisément dans le sous-terme I_{dd} , qui modélise l'impairment dû au **délai absolu** T_a (en l'absence d'écho). La section 7.4 de la recommandation ITU-T G.107 fournit les équations correspondantes (équations 7-18 à 7-27).

3 – Valeurs de R pour juger la qualité

Selon l'annexe / tableau d'interprétation (Annexe B / Table B.1 et tables de G.107) :

- Très bonne qualité : $R \geq 90$ (MOS ≈ 4.3 ou plus).
- Acceptable : $R \approx 80$ (MOS ≈ 4.0).
- Entre 70 et 80 : qualité moyenne, quelques insatisfaits.
- En dessous de 60 : qualité mauvaise.

4 – MOS-CQE, GoB, PoW

- **MOS-CQE** : Mean Opinion Score – Conversational Quality Estimate. Conversion du R-factor en une note MOS pour la conversation.
- **GoB (Good-or-Better)** : probabilité qu'un utilisateur juge la qualité « bonne ou meilleure ».
- **PoW (Poor-or-Worse)** : probabilité qu'un utilisateur juge la qualité « mauvaise ou pire ».

Ces estimateurs permettent de traduire le R-factor en mesures perceptibles pour l'utilisateur.

5 – MOS correspondant à $R = 75$

Formule de conversion $R \rightarrow \text{MOS}$:

$$\text{MOS}(R) = 1 + 0.035R + 7 \times 10^{-6}R(R - 60)(100 - R)$$

Pour $R = 75$:

$$\text{MOS}(75) \approx 3.82$$

Donc la qualité est acceptable, mais avec quelques perturbations.

6 – Calcul de R pour différents délais (sans pertes)

Formule simplifiée :

$$R = 93.2 - I_{dd} - I_{e,\text{eff}}$$

Ici $P_{pl} = 0$, $I_e = 0 \Rightarrow I_{e,\text{eff}} = 0$.

- $T = 50 \text{ ms}$: $I_{dd} = 0 \Rightarrow R = 93.2$.
- $T = 150 \text{ ms}$: $I_{dd} \approx 0.0001 \Rightarrow R \approx 93.2$.
- $T = 300 \text{ ms}$: $I_{dd} \approx 0.049 \Rightarrow R \approx 93.15$.
- $T = 500 \text{ ms}$: $I_{dd} \approx 0.462 \Rightarrow R \approx 92.74$.

Interprétation : même avec des délais élevés, R reste très haut dans cette version simplifiée. En pratique, l'interactivité souffre dès que le délai dépasse $\approx 150 \text{ ms}$.

7 – Calcul de R pour différents taux de perte (codec G.729)

Données : $T = 150$ ms (donc $I_{dd} \approx 0$), $BurstR = 1$, $I_e = 11$, $B_{pl} = 19$.

Formule :

$$I_{e,\text{eff}} = I_e + (95 - I_e) \frac{P_{pl}}{P_{pl}/BurstR + B_{pl}}$$

- $P_{pl} = 1\%$: $I_{e,\text{eff}} \approx 15.2 \Rightarrow R \approx 78.0$.
- $P_{pl} = 5\%$: $I_{e,\text{eff}} \approx 28.5 \Rightarrow R \approx 64.7$.
- $P_{pl} = 10\%$: $I_{e,\text{eff}} \approx 40.0 \Rightarrow R \approx 53.2$.

Interprétation : la perte de paquets a un effet très fort sur la qualité. Au-delà de 5% de pertes, la conversation devient difficile voire inacceptable.

Conclusion

Le modèle E-Model (G.107) fournit une approche analytique pour estimer la qualité conversationnelle. Le R-factor intègre plusieurs sources d'impairment (bruit, délais, pertes, codec) et peut être converti en MOS pour une interprétation utilisateur. Les résultats montrent que la **perte de paquets** a un impact beaucoup plus sévère que le délai, ce qui est critique pour la VoIP.

PESQ

8 – tests subjectifs et objectifs

8.1 - Procédure et principe des tests subjectifs et objectifs

— Tests subjectifs :

- Réalisés par un panel d'auditeurs humains.
- Les auditeurs écoutent des séquences audio (originales et dégradées) et attribuent une note de qualité perçue.
- Les notes sont généralement données sur une échelle de 1 (très mauvaise qualité) à 5 (excellente qualité).
- La moyenne des notes est appelée MOS (Mean Opinion Score).
- Méthode de référence, mais coûteuse, longue et dépendante de la subjectivité humaine.

— Tests objectifs :

- Réalisés automatiquement à l'aide d'algorithmes mathématiques.
- Comparent le signal original avec le signal dégradé.
- Exemples d'algorithmes : MNB1, MNB2, PESQ.
- Méthode rapide, peu coûteuse et reproductible.

8.2 - Interprétation des résultats du graphique

- L'axe des abscisses représente le MOS subjectif (référence humaine).
- L'axe des ordonnées représente le MOS objectif (valeurs calculées par les algorithmes).
- Si un algorithme est fiable, les points doivent être proches de la diagonale $x = y$.
- D'après le graphique :
 - Les algorithmes **MNB1** et **MNB2** présentent une dispersion importante et sont souvent éloignés de la diagonale, donc leur corrélation avec les résultats subjectifs est faible.
 - L'algorithme **PESQ** est globalement mieux aligné avec la diagonale, donc il prédit mieux la qualité perçue.

Conclusion : L'algorithme PESQ est le plus conseillé pour les futurs tests de qualité audio.

9 – Notation PESQ

D'après la recommandation **ITU-T P.862** (Perceptual Evaluation of Speech Quality, PESQ), le score PESQ peut varier dans l'intervalle suivant :

$$-0.5 \leq \text{PESQ} \leq 4.5$$

Ainsi, la note maximale pouvant être fournie par l'algorithme **PESQ** est :

$$\boxed{4.5}$$

10 – Notation Fichier Audio

Note attribué par Adrian :

3

Note attribué par David :

3.5

note moyenne obtenue :

$$\boxed{3.25}$$

11 – Notation Fichier Audio Bis

Note attribué par Adrian :

2

Note attribué par David :

3

note moyenne obtenue :

$$\boxed{2.5}$$

12 – Compilation des fichiers

12.1 - Introduction

+8000 et +16000 correspondant aux valeurs du taux d'échantillonnage qui permet de reconstituer le signal original.

12.2 - (+8000 ou +16000)

Il faut choisir +8000 car les fichiers audio ont un échantillonnage de 8000Hz, généralement utilisé lors des appels téléphoniques qui utilisent des bandes étroites. Pour des bandes larges on aurait pris la valeur de +16000.

12.4 - les tests

```
Reading reference file /users/Etu2/28706212/P862_annex_A_2005_CD/conform/u_am1s01.wav...done.
Reading degraded file /users/Etu2/28706212/P862_annex_A_2005_CD/conform/u_am1s01b2c8.wav...done.
Level normalization...
IRS filtering...
Variable delay compensation...
Acoustic model processing...

P.862 Prediction (Raw MOS, MOS-LQ0): = 2.198  1.806
28706212@ppti-14-502-08:/users/nfs/Etu2/28706212/P862_annex_A_2005_CD/source$
./presq +8000 ~/P862_annex_A_2005_CD/conform/u_am1s01.wav ~/P862_annex_A_2005_CD/conform/u_am1s01b2c8.wav
```

FIGURE 1 – Calcul de la valeur du MOS du fichier dégradé u_am1s01b2c8.wav

```
Reading reference file /users/Etu2/28706212/P862_annex_A_2005_CD/conform/u_am1s01.wav...done.
Reading degraded file /users/Etu2/28706212/P862_annex_A_2005_CD/conform/u_am1s01b2c1.wav...done.
Level normalization...
IRS filtering...
Variable delay compensation...
Acoustic model processing...

P.862 Prediction (Raw MOS, MOS-LQ0): = 4.300  4.415
28706212@ppti-14-502-08:/users/nfs/Etu2/28706212/P862_annex_A_2005_CD/source$
./presq +8000 ~/P862_annex_A_2005_CD/conform/u_am1s01.wav ~/P862_annex_A_2005_CD/conform/u_am1s01b2c1.wav
```

FIGURE 2 – Calcul de la valeur du MOS du fichier dégradé u_am1s01b2c1.wav

12.4 - Résultats

Nous observons un résultat assez proche pour le fichier u_am1s01b2c8.wav mais pour le fichier u_am1s01b2c1.wav, nous avons des résultats totalement différents. Cette disparité peut s'expliquer par le fait que malgré des bonnes méthodes de calcul pour l'estimation de la satisfaction des utilisateurs. Il y a d'autres paramètres tel que la santé auditif qui ne rentre pas dans l'évaluation de la qualité de l'écoute.

13 – POLQA

- POLQA est conçu comme la succession de PESQ, pour couvrir des cas que PESQ ne gère pas bien.
- POLQA supporte l'évaluation de la qualité pour les réseaux **narrowband**, **wideband**, **super-wideband** et **fullband**, couvrant des bandes de fréquences plus larges que PESQ.
- POLQA est mieux adapté pour les **codecs modernes**, les **réseaux IP**, les communications **voix haute définition (HD-Voice)** et les scénarios où des effets de **time-warping**, de **décalage**, d'**alignement temporel** ou de **variation d'échantillonnage** doivent être pris en compte.
- Selon P.863.1, le mode **narrowband** de POLQA peut être utilisé pour assurer la compatibilité avec les résultats PESQ (scénarios de bande étroite), mais ne doit pas être utilisé pour **wideband** ou **fullband**.
- Le mode **fullband** de POLQA est recommandé pour les scénarios modernes et peut même être appliqué sur des signaux narrowband pour les évaluer sur une échelle commune.

Conclusion

- Utiliser **POLQA** plutôt que PESQ lorsqu'on veut évaluer la qualité vocale dans des conditions modernes : codecs wideband, super-wideband ou fullband, réseaux IP/VoIP, ou présence de variations temporelles et décalages que PESQ gère mal.
- **PESQ** reste acceptable pour les scénarios traditionnels en bande étroite (narrowband) ou pour assurer la compatibilité avec des mesures historiques.

15 – MOS subjectif

Pour les conditions de tests, pour ce qui est du contenu on a bien dans le même audio deux phrases sans liaison évidente, comme recommander. Nous avons également des phrases ni trop courte ni trop longues, de deux à trois secondes. Pour ce qui est du nombre de locuteurs, nous n'en avons qu'un seul. Du coup nous avons bien respecté les conditions de test préconisées.