**TP –** Analyse subjective-objective de signaux audio

**Etude du bruit de toqué[[1]](#footnote-1) de planches de bord**

Le TP s’effectue sur MATLAB et/ou XLSTAT (disponible uniquement en salle D005 ECN). Les fichiers nécessaires à la réalisation du TP sont situés sur le serveur pédagogique (cours AUDIO).

* *Texte\_TP\_audio2022.docx*
* *TP-sons.xls*
* *les 16 sons au format.wav (dans un zip)*

*XLSTAT est un logiciel de statistique basé sur Excel. Il permet d’obtenir très rapidement toutes les analyses classiques de données, comparé à Matlab qui demande de la programmation, moins immédiate.*

|  |  |
| --- | --- |
| Le « bruit de toqué » est obtenu lorsqu’un client « sonne » la planche de bord d’une voiture en frappant dessus avec ses doigts. Il tire de cette information une évaluation subjective de la qualité perçue. L’objectif du TP est de comprendre ce qui fonde, du point de vue du signal, les différences perceptives entre différentes planches, et d’expliquer les préférences utilisateur. Le but ultime de l’étude est de savoir comment reconcevoir la planche de bord (matériaux, montage) pour augmenter la préférence utilisateur. |  |

16 sons ont été enregistrés dans des conditions identiques sur différents véhicules (enregistrement avec une tête artificielle – frappe avec un marteau au mouvement calibré). Ces sons ont été évalués par un panel d’experts sensoriels[[2]](#footnote-2) (source PSA) selon 5 descripteurs (tableau 1), sur une échelle de 0 à 10 (voir onglet « profil »).

|  |  |
| --- | --- |
| Descripteur | définition |
| d1 : hauteur (grave à aigu) | Hauteur perçue, de grave à aigu |
| d2 : détonant | Niveau de détonation du bruit |
| d3 : attaquant | Evaluation de l’attaque du bruit |
| d4 : éloignement | Distance à laquelle on estime le bruit |
| d5 : intensité | Niveau sonore global du bruit |

Tableau 1 : liste des descripteurs

# Ecoute-visualisation des sons

Sur Matlab, réaliser une fonction permettant de visualiser le signal temporel, l’analyse spectrale, et le spectrogramme d’un son .wav donné.

**Indications et fonctions Matlab utiles**

Pour sélectionner un fichier son et le charger dans le vecteur  s :

[Fn,Pn] = uigetfile('\*.wav', 'Choisir le fichier son à traiter');

nomfichier=[Pn,Fn] ;

[s,Fe] = audioread(nomfichier) ;

Pour charger automatiquement un fichier son à partir de son nom :

[s,Fe] = audioread(‘∖∖chemin acces∖nom.wav’) ;

ou bien

[s,Fe] = audioread(strcat(num2str(i),’.wav’));

ou bien (pour ne pas à avoir a renumeroter les noms de fichier)

filedir=dir('\*.wav')

for i=1,length(filedir)

[s,Fe] = audioread(filedir(i));

….

% calcul des descripteurs

end

*Attention les sons comportent 2 pistes (stéréo) similaires : ne prendre qu’une seule piste pour les analyses.*

% pour ne prendre qu'une seule piste du signal

s=s(:,1);

Pour jouer un son :

sound(s,Fe);% Attention tous les éléments de s doivent être dans l’intervalle [−1, 1] sinon il y a du « clipping » (saturation).

La fonction soundsc(s,Fe)(sc pour ’scale’) réalise une mise à l’échelle de s. Elle divise s par max(abs(s))) et permet de ramener les valeurs du vecteur son s dans l’intervalle [−1, 1].

**Analyse temporelle**

% T : durée du signal en (s)

T = (length(s)-1)/Fe ;

% t : temps discrétisé pour la durée du signal (en n points)

t = (1:n)/n\*T;

% tracé du signal temporel

figure(1);

plot(t,s);

xlabel('Temps'); ylabel('Amplitude');

**Analyse spectrale (analyse de Fourier avec la fonction fft – Fast Fourier Transform)**

analyse spectrale d’un son *s(t)* de durée *T*, et tracé sur la bande de fréquence [0, 2000Hz]

n=length(s) ;

% normalisation en 1/n pour que les coefficients de la fft ne soient pas proportionnels au nombre de point du signal (l’amplitude de la fft d’une composante sinus sera a/2 si l’amplitude initiale est a)

z = fft(s/n) ;

T = (length(s)-1)/Fe ;

fr = 0 :1/T :2000 ;

nf = length(fr) ;

plot(fr,abs(z(1 :nf))) ;

**Analyse temps-frequence (spectrogramme)**

[S,f,t,p]=spectrogram(s,1024,512,1024,Fe,'yaxis')%,(y,window,noverlap,nfft,Fe);

surf(t,f,10\*log10(abs(p)),’EdgeColor’,’none’) ;

colormap(jet) ;

xlabel('Temps'); ylabel('Fréquence');

axis([0 T 0 2000]);

Consulter le document spectre-spectrogramme.pdf (serveur pédagogique) pour plus d’information sur le sujet. En particulier, prenez conscience du compromis à réaliser entre précision temporelle et précision fréquentielle (pour avoir une bonne précision fréquentielle, il faut une trame longue, comportant beaucoup de points – pour avoir une bonne précision temporelle, il faut considérer beaucoup de trames, courtes, donc de nombreuses fenêtres d’analyse).

# Analyse subjective des sons

## Analyse en composantes principales. Représentation des différences entre les sons

Sur XLSTAT[[3]](#footnote-3), utiliser les données de l’onglet « profil ».

Effectuer une ACP normée sur le tableau de données croisant les sons en lignes (observations = les sons) et les valeurs moyennes des évaluations selon les 5 descripteurs en colonnes (variables = les descripteurs). *XLSTAT>analyse de données>analyse en composantes principales>type d’ACP : choisir « pearson (n-1) » (réalise une ACP normée = ACP sur les données centrées réduites)*.

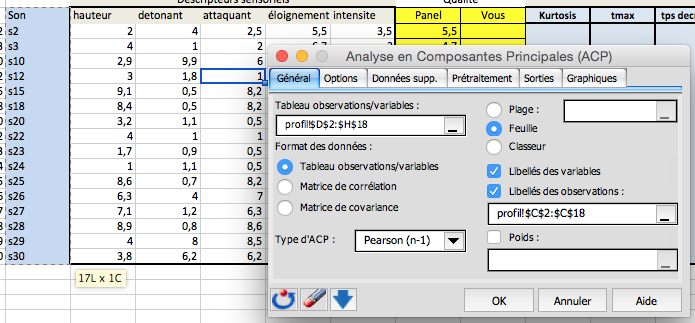
Analyse des composantes principales

* + % d’inertie retenu par les 2 premiers axes
  + analyser les vecteurs propres (contribution des variables à créer les composantes principales)
  + Comment se caractérise un son situé aux extrémités de la CP 1 ?
  + Même question pour la CP 2
  + Interpréter les composantes principales

Etude des individus (les sons)

* + Quels sons contribuent le plus à la CP1 ? à la CP2 ?
  + Y a t-il des groupes de sons qui se dégagent de cette analyse multivariée ?

Aide pour la sélection des zones de données



# Analyse objective des sons

## Calcul de descripteurs du signal.

On propose d’extraire de chaque son les descripteurs suivants.

* + La Valeur LRMS du signal :

* + Le niveau maximum Lmax du signal :

* + Le niveau en dB SPL (Sound pressure level)

* + le Kurtosis K (ou coefficient d’aplatissement) : moment d’ordre 4 de la variable centrée réduite. Il pourrait quantifier « l’impulsivité » du signal :

* + le temps maximum *Tmax* : temps au bout duquel le maximum d’amplitude du signal est atteint (représentation temporelle).
  + Le temps de décroissance *Tdec*: temps ou bout duquel le signal ne varie plus beaucoup. De manière pratique, on cherchera le temps au bout duquel la variance locale du signal (glissante sur 100 échantillons) devient inférieure à un seuil (on prendra 0,05).
  + La puissance par bande de fréquence : on considère k bandes de fréquences dyadiques (de largeur L, 2L, 4L, …2k-1L. Les résultats pour 8 bandes de fréquence sont donnés dans le fichier Excel (vous pouvez re-implémenter cette fonction si vous le souhaitez – voir si dessous). Pour un signal s comportant N points, la puissance sur la bande b est donnée par :

* + Le niveau sonore en dBA. Utiliser pour cela la fonction rundBA.m (qui appelle les fonctions analyzeSignal.m, estimatelevel.m, filterA.m). Transférer le dBA.zip disponible sur le serveur. Pour plus d’information sur le calcul des dBA, consulter l’article SLM.pdf situé dans le zip.

NB : attention pour le calcul des temps : le son ne commence pas au début du fichier, il y a une période de « blanc », variable, dont il faut tenir compte.

1) Implémentez ces fonctions sur Matlab (pour la puissance par bandes, voir ci après) et calculez les descripteurs pour les 16 sons de toqué.

2) Sur le modèle du centre de gravité spectral CGS et des tristimuli TR1, TR2, TR3 définis dans le cours, calculer, à l’aide des puissances par bandes pbi, les descripteurs suivants :

- Centre de gravité de la puissance par bande :

- Tristimulus TR1 :

- Tristimulus TR2 :

- Tristimulus TR3 :

3) Implémentez d’autres descripteurs que vous jugerez utile pour caractériser les différences entre les signaux. Vous pouvez par exemple déterminer

- le temps au bout duquel le niveau sonore en dBA se situe en dessous de xxxdBA.

- la fréquence pour laquelle l’amplitude du spectre atteint un maximum (on suspecte que cette grandeur soit corrélée à la hauteur perçue du son)

**fonctions matlab utiles :**

k = kurtosis(y) ; kurtosis de y

v = var(y) ; variance de y

détection d’un seuil pour le début du son

nseuil=min(find(abs(y)>0.01));

tseuil = (nseuil-1)/Fe ;

Puissance par bande

% yreduit : signal utile (en retirant le début – car le son ne commence pas au début du fichier)

nr = length(yreduit);

Treduit=(nr-1)/Fe

nbandes=8;

% largeur de la première bande

bandwidth = floor(length(yreduit)/2/((2^nbandes)-1));

% fréquence de la première bande F1

F1=bandwidth\*Fe/nr ;

yreduit = yreduit - mean(yreduit);

spec = abs(fft(yreduit)/nr);

spec = spec(1:floor(nr/2)+1); % De 0 a 0.5 fe

spband = [];

band = 1;

for i=1:nbandes

spband (i) = sum(spec(band:band + bandwith -1));

band = band + bandwith;

bandwith = bandwith \*2;

end;

spband(nbandes)=spband(nbandes) + sum(spec(band:end))

Pour exporter une matrice M au format xlsx (Excel)

xlswrite(‘descripteur.xlsx’,M);

# Etude des corrélations

Etudier les corrélations (coefficient de corrélation linéaire de Pearson) entre les descripteurs sensoriels et les descripteurs du signal.

*(sur XLSTAT: fonction XLStat>description de données>matrices de similarité dissimilarité)*

*sur matlab : fonction* corrcoef(A,B);)

Analyser les résultats. Les corrélations sont elles cohérentes, logiques ?

Pour chaque descripteur sensoriel, déterminer un descripteur du signal qui l’explique (corrélation suffisamment importante – 0,7 voir 0,8, en positif ou négatif). Revenir si besoin à l’étape 3 précédente et proposer de nouveaux descripteurs des signaux.

NB : vous devez déterminer un descripteur signal adapté qui explique chacun des descripteurs sensoriels. Pour cela, écoutez les sons extrêmes selon le descripteur sensoriel, utilisez des représentations graphiques, pour déterminer les attributs qui les différencient. Cela doit vous aider à définir la forme que doit prendre le descripteur signal.

# Etude de la qualité des sons

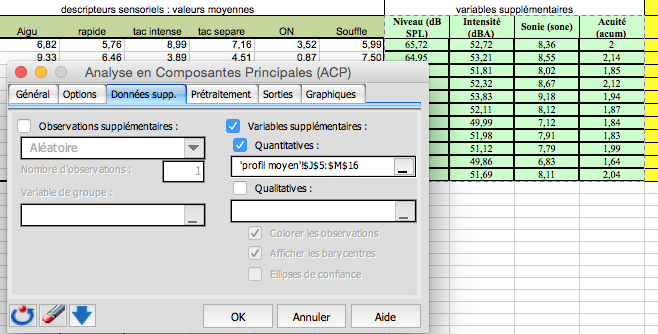
Une évaluation de la qualité perçue des sons a été réalisée par un panel d’utilisateurs.

Les évaluations moyennes de la qualité sont données dans l’onglet « profil », colonne « panel ».

1) Par une écoute attentive de chaque son, effectuez vous même une cotation de la qualité des sons, sur une échelle de 0 (très mauvaise qualité) à 10 (très bonne qualité) (colonne « vous »)

2) Projeter les cotations de qualité en variables supplémentaires sur le plan de l’ACP des descripteurs sensoriels (voir onglet *« Données supplémentaires »* de la fonction ACP d’XLStat). Conclusions ? Comment caractériser la qualité du son pour le panel ? pour vous ?

*Aide pour la projection de variables supplémentaires en ACP avec XLSTAT (onglet Données supp.)*



3) Sur un graphe 2D, représenter la cotation de qualité en fonction de chaque descripteur sensoriel (faire 5 graphes) *(fonction XLStat>Visualisation de données>nuage de points).* Conclusions ?

Calculer la matrice des coefficients de corrélation (linéaire de Pearson) entre les cotations de qualité et les descripteurs sensoriels *(sur XLSTAT: fonction XLStat>description de données>matrices de similarité dissimilarité).* Conclusions ?

A l’aide d’une régression linéaire simple, proposer un modèle explicatif expliquant la qualité par un des descripteurs sensoriels. *(fonction XLStat>modélisation de données>régression linéaire).*

Variable dépendante Y : la qualité

Variable explicative X: le descripteur sensoriel

Equation du modèle :

(1)

Conclusions ? L’ajustement du modèle sur les données est-il correct (R2) ? Comment se caractérise la qualité du son pour le panel d’un point de vue sensoriel ? Pour vous ?

4) Dans cette question, on cherche à modéliser plus finement la qualité en fonction des descripteurs sensoriels. On propose de réaliser un modèle de la qualité donnée par le « panel » et par « vous » non pas sen fonction d’un seul descripteur (question précédente), mais en fonction des composantes principales F1 et F2 de l’ACP des descripteurs sensoriels (voir cours sur les cartographies de préférences).

* 1. Vérifier que les données reportées en colonnes F1 et F2 sont bien les coordonnées factorielles des sons sur les deux premières composantes principales F1 et F2 (cf. question 2 ACP)
  2. Pour chaque cas (panel et « vous »), réaliser une régression linéaire multiple

Q = aF1 + bF2+c (2)

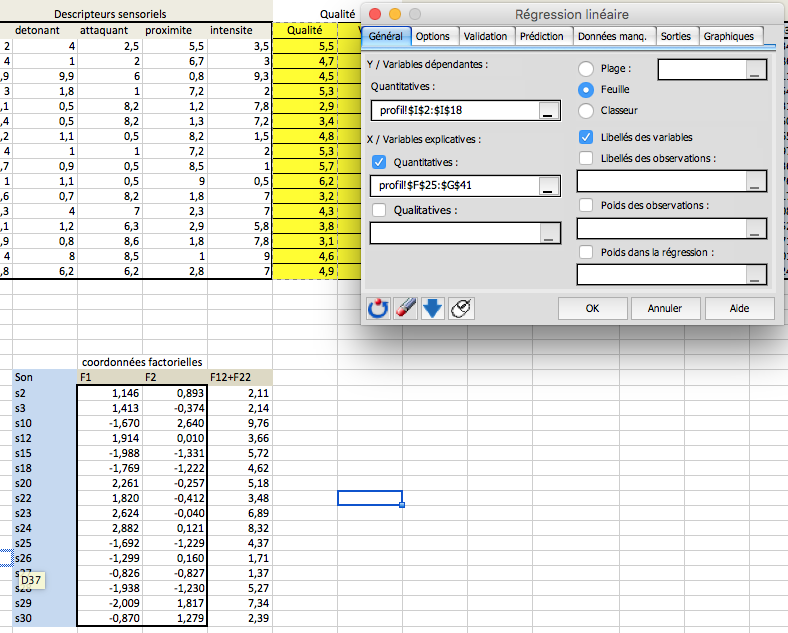
expliquant la qualité Q par les coordonnées factorielles F1 et F2 (modèle vectoriel) (*XLSTAT>modélisation de données>régressions linéaires*).

Remplir le tableau suivant. Conclusions ? Quel pourcentage de la variance est expliqué par la régression (coefficient de détermination R2) ?

Conclusions concernant l’interprétation de la qualité ? (modèle vecteur)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | **a** | **b** | **c** | **R2** | **conclusions** |
| panel |  |  |  |  |  |
| vous |  |  |  |  |  |

*Indication pour la saisie des données d’une régression linéaire multiple avec XLSTAT.*



Si le pourcentage de variance expliquée par le modèle, caractérisée par R2, n’est pas suffisant (inférieur à 0,7), nous allons réaliser un modèle plus complet, un modèle circulaire (degré 2).

* 1. Pour les sujets pour lesquels le R2 est inférieur à 0.7 (70%), ajuster un modèle circulaire (*XLSTAT>modélisation de données>régressions linéaires*) (vérifier au passage que la nouvelle variable explicative introduite, colonne F12+F22, correspond bien à la quantité F12+F22)

Q = aF1 + bF2+d(F12+F22)+c (2)

L’extremum de ce modèle est situé au point I de coordonnées : I (-a/2d ; -b/2d)

* + Si d >0 : il s’agit d’un point de rejet (minimum de la qualité)
  + Si d<0 : il s’agit d’un point idéal (maximum de qualité).

Caractériser les extrema d’un point de vue sensoriel.

Reporter les résultats dans le tableau suivant. Conclusions sur l’optimum de qualité ?

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **a** | **b** | **d** | **c** | **R2** | **conclusions** |
| panel |  |  |  |  |  |  |
| Vous |  |  |  |  |  |  |

A la lumière de l’analyse précédente, proposer des niveaux pour les descripteurs du signal pour optimiser la qualité du « toqué ». Comment se caractérise un son de qualité du point de vue du signal ?

Rédiger un compte rendu par binôme. A envoyer par email en .pdf à J-F PETIOT [jean-francois.petiot@ls2n.fr](mailto:jean-francois.petiot@ls2n.fr) - (bureau LS2N S417) pour ce soir 3/02/2022.

Conseil pour la rédaction

* + Numéroter les figures et tableaux
  + Commenter les figures
  + présenter toutes les données permettant au lecteur de comprendre les conclusions
  + appuyer vos analyses/discussions sur des faits présentés dans des figures ou des tableaux
  + ne pas faire figurer des données avec 10 chiffres significatifs après la virgule
  + ne pas mettre de programmes dans le compte rendu (mettre en annexe si nécessaire)

1. Voir la thèse de doctorat de Marie-Céline BEZAT (entreprise PSA) sur HAL. Perception de bruits d’impact. Application au bruit de fermeture de porte automobile, Université de Provence Aix Marseille I, 2007. [↑](#footnote-ref-1)
2. Cette évaluation a été effectuée par le panel d’experts de PSA, les valeurs sont les données réelles. [↑](#footnote-ref-2)
3. Pour lancer XLStat la 1iere fois, il est parfois nécessaire d’abaisser le niveau de sécurité des macros d’Excel.

   1) Lancer Excel 2010 - 2) Aller dans Options>centre de gestion...>paramètre du centre...>paramètres des macros>activer les macros – 3) lancer Xlstat – compléments>XLSTAT [↑](#footnote-ref-3)